

# «Ψηφιακές Εφαρμογές και Καινοτομία»

Πρόγραμμα Μεταπτυχιακών Σπουδών (ΠΜΣ)



Capstone Project

## **«Μοντέλα διάγνωσης διαταραχών της ισορροπίας με χρήση πολυδιάστατων δεδομένων»**



Φοιτήτρια: Κρητικού Αγγελική

ΑΜ:ΠΜΚ201904

Επιβλέπων καθηγητής: Έξαρχος Θέμης

## Περιεχόμενα

|                                                                                                                                                                                                                                                                                |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| «Μοντέλα διάγνωσης διαταραχών της ισορροπίας με χρήση πολυδιάστατων δεδομένων».....                                                                                                                                                                                            | 1  |
| Περίληψη.....                                                                                                                                                                                                                                                                  | 5  |
| Εισαγωγή.....                                                                                                                                                                                                                                                                  | 5  |
| Ισορροπία.....                                                                                                                                                                                                                                                                 | 6  |
| Αιθουσαίο σύστημα.....                                                                                                                                                                                                                                                         | 6  |
| Αιθουσαίες διαταραχές.....                                                                                                                                                                                                                                                     | 7  |
| Ίλιγγος.....                                                                                                                                                                                                                                                                   | 8  |
| Τι συμβαίνει και έχουμε ίλιγγο.....                                                                                                                                                                                                                                            | 8  |
| Τι προκαλεί τον ίλιγγο.....                                                                                                                                                                                                                                                    | 9  |
| Ποια είναι τα συχνότερα αίτια.....                                                                                                                                                                                                                                             | 9  |
| Καλοήθης παροξυσμικός ίλιγγος θέσεως- Benign Paroxysmal Posistional Vertigo (BPPV).....                                                                                                                                                                                        | 9  |
| Αιθουσαία νευρωνίτιδα.....                                                                                                                                                                                                                                                     | 10 |
| Νόσος του Meniere.....                                                                                                                                                                                                                                                         | 11 |
| Αιθουσαία Ημικρανία.....                                                                                                                                                                                                                                                       | 12 |
| Επίμονη Ζάλη Ορθοστατικής Αντίληψης (Persistent Postural-Perceptual Dizziness – PPPD).....                                                                                                                                                                                     | 12 |
| Εγκεφαλικό.....                                                                                                                                                                                                                                                                | 12 |
| Στρες- Ψυχογενής ίλιγγος.....                                                                                                                                                                                                                                                  | 12 |
| Ναυτία ταξιδιωτών.....                                                                                                                                                                                                                                                         | 12 |
| Συνήθειες που επιδεινώνουν τον ίλιγγο.....                                                                                                                                                                                                                                     | 13 |
| Μεθοδολογίες εξέτασης.....                                                                                                                                                                                                                                                     | 13 |
| Το πρόβλημα : Δυσκολία διάγνωσης.....                                                                                                                                                                                                                                          | 13 |
| Αντιμετώπιση:Κλινικά Συστήματα Υποστήριξης Απόφασης - Clinical Decision Support Systems (CDSS)<br>.....                                                                                                                                                                        | 14 |
| Παρουσίαση ερευνών με συναφή προβλήματα.....                                                                                                                                                                                                                                   | 15 |
| Ανάπτυξη Διαγνωστικού Συστήματος Υποστήριξης Απόφασης για καλοήθη παροξυσμικό ίλιγγο θέσεως χρησιμοποιώντας μοντέλο βαθιάς μάθησης (2019).....                                                                                                                                 | 15 |
| Μηχανογραφημένο σύστημα κλινικών αποφάσεων και εφαρμογή για κινητά με εξειδικευμένη υποστήριξη για τη βελτιστοποίηση της διαχείρισης του ίλιγγου στην πρωτοβάθμια φροντίδα: πρωτόκολλο μελέτης για μια πραγματιστική κατά συστάδα-τυχαιοποιημένη ελεγχόμενη δοκιμή (2020)..... | 16 |
| Δεδομένα διαταραχών ισορροπίας εξόρυξης για την ανάπτυξη διαγνωστικών συστημάτων υποστήριξης αποφάσεων - EMBalance διαγνωστικό σύστημα υποστήριξης αποφάσεων (2016) .....                                                                                                      | 17 |
| Διαφορική μέθοδος διαγνωστικής συλλογιστικής για καλοήθη παροξυσμικό ίλιγγο θέσεως με βάση δυναμικό αβέβαιο γράφημα αιτιότητας (2019).....                                                                                                                                     | 18 |

|                                                                                                                                                                                          |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Σχεδιασμός ενός συστήματος υποστήριξης κλινικών αποφάσεων για τη διαχείριση και τη θεραπεία ασθενών με κύριο παράπονο τον ίλιγγο (2018).....                                             | 19 |
| Ανάπτυξη και επικύρωση ενός αλγορίθμου ταξινόμησης για τη διάγνωση και τη διαφοροποίηση αυθόρμητων επεισοδίων σύνδρομων ίλιγγου: αποτελέσματα από το μητρώο ασθενών DizzyReg (2020)..... | 20 |
| Μηχανική μάθηση.....                                                                                                                                                                     | 20 |
| Προεπεξεργασία δεδομένων.....                                                                                                                                                            | 21 |
| Εξισορρόπηση.....                                                                                                                                                                        | 21 |
| Αλγόριθμοι κατηγοριοποίησης.....                                                                                                                                                         | 22 |
| Κατηγοριοποίηση κατά Κ-πλησιέστερων γειτόνων Κ-NN. ....                                                                                                                                  | 22 |
| Κατηγοριοποίηση με Δέντρο Απόφασης (Decision Tree).....                                                                                                                                  | 23 |
| Κατηγοριοποίηση κατά Naive Bayes. ....                                                                                                                                                   | 23 |
| Κατηγοριοποίηση με Random Forest .....                                                                                                                                                   | 24 |
| Μέτρα ακριβείας.....                                                                                                                                                                     | 24 |
| Ακρίβεια Ταξινόμησης.....                                                                                                                                                                | 25 |
| Ευαισθησία ή αλλιώς Ανάκληση (Sensitivity or Recall).....                                                                                                                                | 25 |
| Ειδικότητα-Specificity .....                                                                                                                                                             | 25 |
| Ακρίβεια-Precision .....                                                                                                                                                                 | 25 |
| Μέτρο F - F-measure.....                                                                                                                                                                 | 26 |
| Αξιολόγηση.....                                                                                                                                                                          | 26 |
| Κ-πλη διασταυρωμένη επικύρωση.....                                                                                                                                                       | 27 |
| Εφαρμογή στο WEKA.....                                                                                                                                                                   | 28 |
| Ερευνητικός χώρος.....                                                                                                                                                                   | 28 |
| Αποτελέσματα αλγορίθμων για την ασθένεια BPPV-1 (Παροξυσμικός ίλιγγος θέσεως τύπου 1) χωρίς επιλογή χαρακτηριστικών.....                                                                 | 29 |
| Αποτελέσματα αλγορίθμων για την ασθένεια BPPV-2 (Παροξυσμικός ίλιγγος θέσεως τύπου 2) χωρίς την επιλογή χαρακτηριστικών.....                                                             | 33 |
| Αποτελέσματα αλγορίθμων για την ασθένεια Meniere χωρίς επιλογή χαρακτηριστικών.....                                                                                                      | 36 |
| Αποτελέσματα αλγορίθμων για την ασθένεια Vestibular Migraine χωρίς την επιλογή χαρακτηριστικών.....                                                                                      | 39 |
| Αποτελέσματα αλγορίθμων για την ασθένεια Vestibular Neuronitis χωρίς την επιλογή χαρακτηριστικών.....                                                                                    | 43 |
| Επιλογή χαρακτηριστικών ασθενειών.....                                                                                                                                                   | 46 |
| Αποτελέσματα αλγορίθμων για την ασθένεια BPPV-1 (Παροξυσμικός ίλιγγος θέσεως τύπου 1) μετά την επιλογή χαρακτηριστικών.....                                                              | 46 |

|                                                                                                                              |    |
|------------------------------------------------------------------------------------------------------------------------------|----|
| Αποτελέσματα αλγορίθμων για την ασθένεια BPPV-2 (Παροξυσμικός ίλιγγος θέσεως τύπου 2) μετά την επιλογή χαρακτηριστικών. .... | 48 |
| Αποτελέσματα αλγορίθμων για την ασθένεια Meniere μετά την επιλογή χαρακτηριστικών. ....                                      | 50 |
| Αποτελέσματα αλγορίθμων για την ασθένεια Vestibular Migraine μετά την επιλογή χαρακτηριστικών. ....                          | 51 |
| Αποτελέσματα αλγορίθμων για την ασθένεια Vestibular Migraine μετά την επιλογή χαρακτηριστικών ....                           | 53 |
| Συμπεράσματα- Μελλοντική Έρευνα. ....                                                                                        | 56 |
| Βιβλιογραφία. ....                                                                                                           | 58 |

## Περίληψη.

*“Υγεία σημαίνει ισορροπία”*

Η ισορροπία των σωμάτων του σύμπαντος ήταν ένα θέμα που απασχολούσε ακόμα και τους αρχαίους Έλληνες φιλόσοφους. Από την αρχαιότητα, Έλληνες φιλόσοφοι όπως ο Αριστοτέλης, ο Γαλιλαίος αλλά μεταγενέστερα και ο Ισαάκ Νεύτων, πρότειναν διάφορες θεωρίες για την κίνηση και την ισορροπία των σωμάτων. Μία από τις πιο σημαντικές λειτουργίες και του ίδιου του ανθρώπινου σώματος, είναι ο έλεγχος της ισορροπίας του και η κινητική του στο χώρο. Η ισορροπία όπως έχει αποδειχθεί είναι το μέσο για μια αρμονική ζωή.

Σκοπός της εργασίας αυτής είναι η κατανόηση της έννοιας της ισορροπίας, η οποία είναι το βασικό χαρακτηριστικό της κινητικής συμπεριφοράς. Κατανοώντας τη σημασία της ισορροπίας και τα συστατικά από τα οποία απαρτίζεται, το επόμενο στάδιο είναι να δούμε πως μπορεί ένας γιατρός να ερμηνεύσει τις διαταραχές ισορροπίας των ασθενών και την προέλευση τους, έτσι ώστε να φτάσει στο στάδιο της διάγνωσης και της σωστής θεραπείας. Η απώλεια ισορροπίας μπορεί να οφείλεται σε μια πλειάδα αιτιών όπως νευρολογικές παθήσεις, προβλήματα μυοσκελετικού συστήματος και ωτολογικές παθήσεις. Εξαιτίας της πολυπλοκότητας των διαταραχών ισορροπίας, η προσέγγιση του θεράποντα ιατρού (εξειδικευμένου ή μη) δεν είναι αρκετή. Για το λόγο αυτό, τέτοιου είδους προβλήματα απαιτούν “ένα χέρι βοήθειας”. Εκεί συμβάλλει η επιστήμη της “Εξόρυξης Δεδομένων” όπου τα τελευταία χρόνια έχει γίνει απαραίτητη στον Ιατρικό κλάδο. Σύμφωνα με έρευνα της PubMed, η εξόρυξη δεδομένων γίνεται όλο και πιο δημοφιλής στην υγειονομική περίθαλψη, αν όχι όλο και πιο απαραίτητη. Η εξόρυξη δεδομένων περιλαμβάνει τη δημιουργία κανόνων συσχέτισης, τη χρήση κριτηρίων υποστήριξης και εμπιστοσύνης για τον εντοπισμό των πιο σημαντικών σχέσεων στα δεδομένα. Έτσι δημιουργήθηκαν τα Κλινικά Συστήματα Υποστήριξης Αποφάσεων (DCSS- Clinical Decision Support Systems). Ένα τέτοιο σύστημα είναι ένας τύπος συστήματος λογισμικού που υποστηρίζει τη λήψη αποφάσεων του γιατρού ή του ειδήμονα γιατρού. Αυτά τα συστήματα ορίζονται συνήθως ως οποιοσδήποτε τύπος συστήματος εφαρμογών που παρουσιάζει αναλυτικά δεδομένα για να βοηθήσει τους γιατρούς ή άλλους επαγγελματίες του ιατρικού τομέα να λάβουν αποφάσεις. Στην εργασία αυτή λοιπόν παρουσιάζεται η σημασία της επιστήμης της Εξόρυξης Δεδομένων στον Ιατρικό Κλάδο και συγκεκριμένα στην εξακρίβωση των γενεσιουργών αιτιών των αιθουσαίων διαταραχών χάρη στα κατάλληλα Κλινικά Συστήματα Υποστήριξης Αποφάσεων.

## Εισαγωγή.

Ο στόχος του παρόντος Capstone Project είναι η ανάλυση συνόλου πολυδιάστατων δεδομένων από ασθενείς με 12 διαφορετικές διαταραχές της ισορροπίας, με τεχνικές μηχανικής μάθησης, με σκοπό την ανάπτυξη και την αξιολόγηση μοντέλων διάγνωσης. Εν συνεχεία θα αναλυθούν οι έννοιες της ισορροπίας αλλά και των διαταραχών της. Ιδιαίτερη σημασία θα δοθεί στα προβλήματα διάγνωσης που αντιμετωπίζουν οι θεράποντες ιατροί (ειδήμονες ή μη) λόγω της πολυπλοκότητας των αιτιών των διαταραχών ισορροπίας. Απώτερος σκοπός είναι η διαχείριση και αξιοποίηση των πολυδιάστατων δεδομένων (προσωπικά στοιχεία, ερωτηματολόγια, αποτελέσματα κλινικών εξετάσεων αλλά και

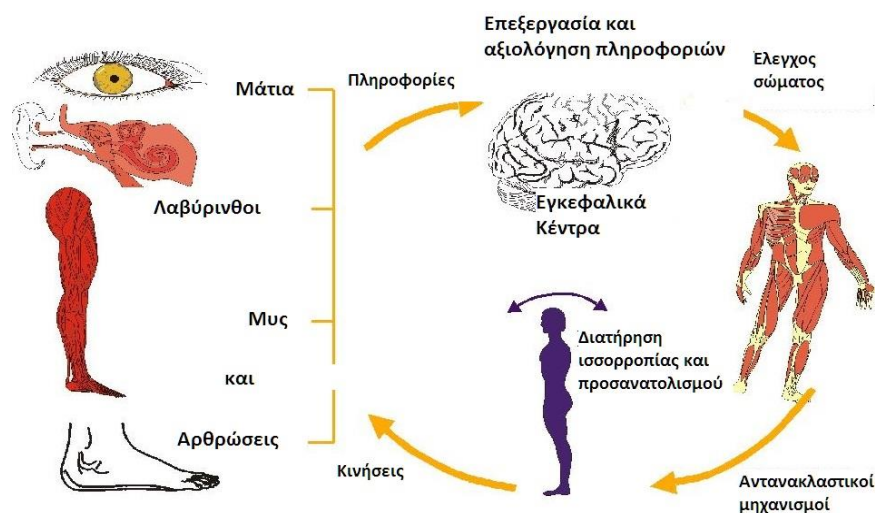
συμπτωμάτων ) των ασθενών μέσω των Κλινικών Συστημάτων Υποστήριξης Αποφάσεων έτσι ώστε να επιτευχθεί η παραγωγή πολύτιμης γνώσης και η άντλησή της για μελλοντική αξιοποίηση στον ιατρικό κλάδο.

*Λέξεις κλειδιά: Αιθουσαίες διαταραχές, Αλγόριθμοι Ταξινόμησης, Weka, CDSS – Κλινικά Συστήματα Υποστήριξης Απόφασης.*

## Ισορροπία.

### Ορισμός:

Ισορροπία είναι η διαδικασία με την οποία το άτομο διατηρεί τη θέση (στατική) ή/και την κίνηση (δυναμική) του σώματος του, σε μια συγκεκριμένη σχέση προς το περιβάλλον. Επηρεάζεται από τη δύναμη της βαρύτητας, τις διαταραχές της βοηθητικής μετακίνησης (π.χ. στροφή, κάμψη, έκταση κορμού, κτλ.), και τις αλληλεπιδράσεις του περιβάλλοντος (π.χ. ολισθήσεις, συγκρούσεις, ωθήσεις, κτλ.).



**Εικόνα 1. Μηχανισμοί που συμβάλλουν στην διατήρηση της ισορροπίας<sup>1</sup>**

## Αιθουσαίο σύστημα.

Το σύστημα ισορροπίας ή αλλιώς το αιθουσαίο σύστημα (vestibular system) είναι πιθανότατα το κυριότερο αισθητηριακό σύστημα του ανθρώπινου σώματος ,παρόλα αυτά είναι και το σύστημα για το οποίο υπάρχει η λιγότερη γνώση από την πλειοψηφία της ανθρωπότητας.

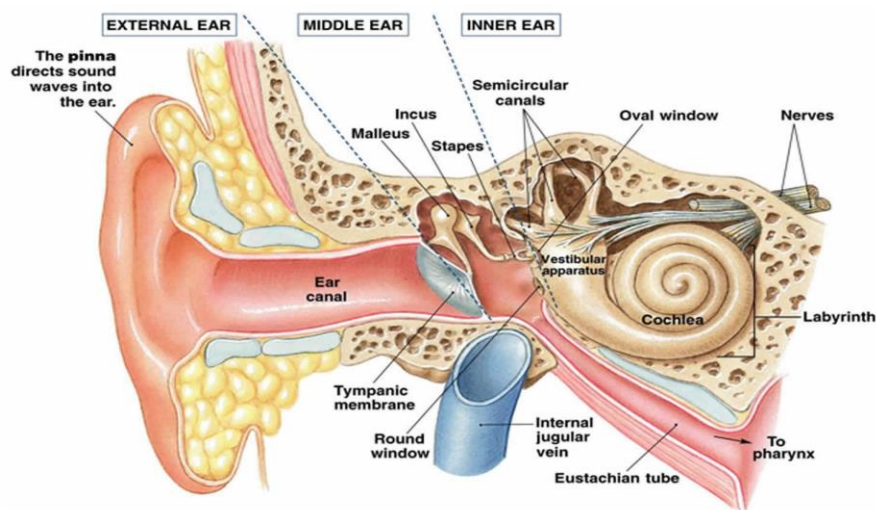
<sup>1</sup> <https://www.vardouniotis.gr/index.php?id=57>

Το αιθουσαίο σύστημα λοιπόν, είναι υπεύθυνο για να μεταφέρει την πληροφορία για τη θέση της κεφαλής σε σχέση με το χώρο, να ανιχνεύει την κίνηση και επομένως να έχει τη δυνατότητα να αναγνωρίζει τις πιθανές αλλαγές της θέσης της κεφαλής.

Αξίζει να σημειωθεί ότι οι τύποι όλων των αισθήσεων (όραση, ακοή, αφή, όσφρηση, γεύση) επεξεργάζονται σύμφωνα με τις πληροφορίες που δέχεται το αιθουσαίο σύστημα, επομένως είναι προφανές ότι όλα τα συστήματα συνεργάζονται με αυτό. Επομένως για την σωστή ένταξη των υπολοίπων αισθητηριακών συστημάτων είναι απαραίτητο ένα υγιές αιθουσαίο σύστημα. Καθώς, είναι η βάση για τον μυϊκό τόνο, την ισορροπία και τον αμφίπλευρο συντονισμό.

Με λίγα λόγια έχει τη δυνατότητα να αντιλαμβάνεται εάν το σώμα κινείται, εάν στέκεται όρθιο, εάν βρίσκεται σε ισορροπία, κλπ. Όταν κάποιος έχει ένα «κανονικό» ή καλά αναπτυγμένο αιθουσαίο σύστημα, όλες αυτές οι πληροφορίες κατηγοριοποιούνται σωστά και δίνεται η κατάλληλη απόκριση.

Σύμφωνα με τα παραπάνω λοιπόν, σε περίπτωση μη σωστής λειτουργίας του αιθουσαίου συστήματος, ο άνθρωπος είτε γίνεται υπερευαίσθητος σε κάθε κίνηση είτε υπο-ανταποκρίνεται. Αυτό έχει ως αποτέλεσμα το άτομο να μη μπορεί να ανταποκριθεί στις καθημερινές του δραστηριότητες σε περίπτωση απώλειας ή και ακόμη διαταραχής της αιθουσαίας λειτουργίας.



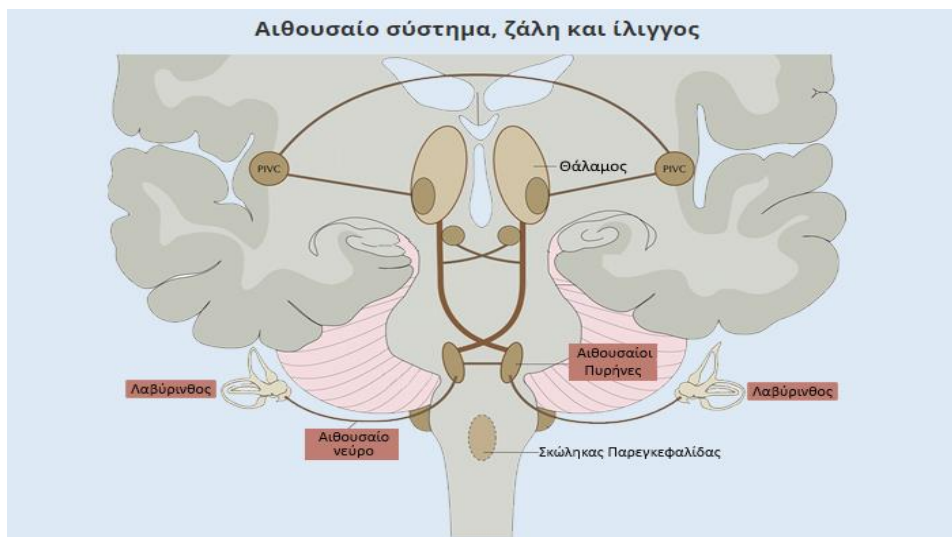
Εικόνα 2. Αιθουσαίο σύστημα<sup>2</sup>

## Αιθουσαίες διαταραχές

Σύμφωνα με έρευνες έχει καταγραφεί ότι μεγάλο ποσοστό του πληθυσμού βιώνει συχνά συμπτώματα όπως ανισορροπία, διαταραχή όρασης, ζάλη και ίλιγγο. Τα συμπτώματα αυτά οφείλονται στις αιθουσαίες διαταραχές και είναι υπεύθυνα για την υποβάθμιση της καθημερινότητας των ανθρώπων αφενός και αφετέρου για την κατώτερη ποιότητα ζωής τους. Άτομα με τέτοιου είδους συμπτώματα έχει παρατηρηθεί ότι ακολουθούν ένα καθιστικό τρόπο ζωής, αποσκοπώντας στην αποφυγή επιδείνωσης

<sup>2</sup> <https://healthjade.net/vestibular-system/>

τους. Ως αποτέλεσμα όμως φανερώνεται η επιδείνωση των συμπτωμάτων τους, αφού η κατάσταση αυτή οδηγεί σε μειωμένη μυϊκή δύναμη, μειωμένη ευκινησία και σε μειωμένη κινητικότητα των αρθρώσεων.



Εικόνα 3. Επισκόπηση αιθουσαίου συστήματος<sup>3</sup>

## Ίλιγγος.

Ένα από τα βασικά συμπτώματα των αιθουσαίων διαταραχών είναι ο ίλιγγος. Ο ίλιγγος είναι η ψευδαίσθηση της κίνησης, καθώς το άτομο βρίσκεται σε αδράνεια. Το άτομο έχει την αίσθηση ότι το ίδιο ή το περιβάλλον του βρίσκεται σε κίνηση.

Ο ίλιγγος συνήθως έρχεται με συνοδεία από διάφορες νευροφυτικές εκδηλώσεις, όπως ναυτία, εμετό, εφίδρωση αλλά και ωχρότητα. Ανάλογα με την αιτία της βλάβης, τα συμπτώματα αυτά γίνονται αρκετά δυσάρεστα, και επιπλέον μπορεί να συνοδεύονται από βαρηκοΐα, και εμβοές.

Αντικειμενικό σημείο του είναι ο νυσταγμός. Ως νυσταγμός, ορίζονται οι ακούσιες εκκρεμοειδείς κινήσεις των οφθαλμικών βολβών, όταν το άτομο τρέφει τους οφθαλμούς προς μια κατεύθυνση.

## Τι συμβαίνει και έχουμε ίλιγγο.

Ο ίλιγγος όπως προαναφέρθηκε είναι σύμπτωμα, όχι ασθένεια. Μπορεί να διαρκέσει από μερικά λεπτά έως και ώρες, μέρες ακόμη και εβδομάδες σε περιπτώσεις που χαρακτηρίζονται ως σοβαρές. Η διάρκεια του λοιπόν εξαρτάται από την αιτία που τον προκαλεί, όσο πιο κακοήθης είναι η αιτία τόσο μεγαλύτερη είναι η διάρκεια του.

Όσον αφορά το κομμάτι της διάγνωσης, η διάρκεια του ίλιγγου βοηθά τον γιατρό στο να εντοπίσει την αιτία του. Για να αντιμετωπιστούν οι ίλιγγοι, πρέπει να εντοπιστεί και να θεραπευτεί η αιτία που τους

<sup>3</sup> <https://neurologos.gr/iliggos-ayxenas-sindromo/>

προκαλεί. Ανεξάρτητα από την αιτία που τους προκαλεί όμως, συχνά υποχωρούν μόνοι τους ή μπορεί να αντιμετωπιστούν επιτυχώς.

### Τι προκαλεί τον ίλιγγο.

Οι ίλιγγοι χωρίζονται σε δύο κύριες κατηγορίες σε κεντρικούς και περιφερικούς. Οι κεντρικοί ίλιγγοι έχουν ως αιτίες ανισορροπίες στο νευρικό σύστημα, ενώ από την άλλη, οι περιφερικοί έχουν ως αιτίες ανισορροπίες στο εσωτερικό του αυτιού. Αξίζει να σημειωθεί ότι οι περιφερικοί ίλιγγοι είναι συχνότεροι από τους κεντρικούς, παρόλο που υπάρχουν πολλές κατηγορίες.

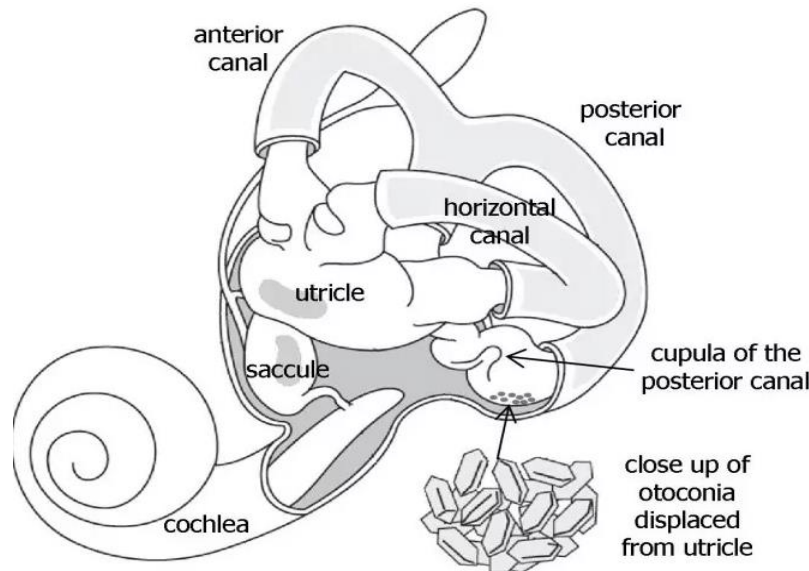
Αναλυτικότερα, ο ίλιγγος στο 80% των περιπτώσεων, οφείλεται σε ενδολαβυρινθικές βλάβες, παθήσεις που αφορούν το αυτί ή βλάβες της αιθουσαίας οδού, διαταραχές οι οποίες πραγματοποιούνται κατά τη μεταφορά των ερεθισμάτων στον εγκέφαλο από το αυτί. Επίσης, μπορεί να οφείλεται σε ψυχογενή αίτια ή σε παθήσεις του κεντρικού νευρικού συστήματος όπως αποστήματα εγκεφάλου, αιμορραγίες, όγκοι, αγγειακά εγκεφαλικά επεισόδια ακόμα και σκλήρυνση κατά πλάκας. Τέλος, ίλιγγος μπορεί να προκληθεί από χαμηλή αρτηριακή πίεση, χαμηλό σάκχαρο, αναιμία, υπογλυκαιμία, λήψη κάποιων φαρμάκων ή ακόμη και stress.

### Ποια είναι τα συχνότερα αίτια.

Η νόσος του Meniere, η Αιθουσαία Νευρωνίτιδα και ο Καλοήθης Παροξυσμικός Ίλιγγος Θέσης είναι οι συνηθέστερες αιτίες παθογένειας του εσωτερικού αυτιού (περιφερικός ίλιγγος) ενώ η Αιθουσαία Ημικρανία είναι περισσότερο μια μορφή κεφαλαλγίας. Στην πλειοψηφία των περιπτώσεων οι ασθένειες αυτές είναι που προκαλούν ίλιγγο και διαταραχή της ισορροπίας. Παρακάτω, θα γίνει μια αναφορά ασθενειών που μπορεί να προκαλέσουν ίλιγγο:

### Καλοήθης παροξυσμικός ίλιγγος θέσεως- Benign Paroxysmal Posistional Vertigo (BPPV).

Αφορά το 20% των περιπτώσεων ίλιγγου και κύριο χαρακτηριστικό είναι τα σύντομα επεισόδια ίλιγγου, όταν αλλάζει απότομα η θέση της κεφαλής. Το BPPV είναι ένα μηχανικό πρόβλημα στο εσωτερικό αυτί. Εμφανίζεται όταν μερικοί από τους κρυστάλλους ανθρακικού ασβεστίου (ωτοκονία) που κανονικά είναι ενσωματωμένοι σε γέλη στη μήτρα αποσπώνται και μεταναστεύουν σε ένα ή περισσότερα από τα 3 ημικυκλικά κανάλια γεμάτα υγρό, στα οποία υποτίθεται ότι δεν είναι. Όταν αρκετά από αυτά τα σωματίδια συσσωρευτούν σε ένα από τα κανάλια, παρεμβαίνουν στην κανονική κίνηση ρευστού που χρησιμοποιούν αυτά τα κανάλια για να ανιχνεύσουν την κίνηση της κεφαλής, προκαλώντας στο εσωτερικό αυτί να στέλνει ψευδή σήματα στον εγκέφαλο.



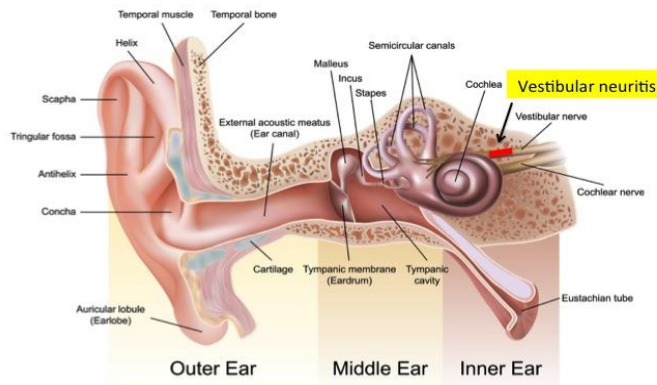
*Εικόνα 4. Ανατομία εσωτερικού αυτιού. Η ωτοκονία (ωτόλιθοι) μεταναστεύει από τη μήτρα, συνήθως εγκαθίσταται στον οπίσθιο ημικυκλικό σωλήνα (φαίνεται) ή πιο σπάνια στους πρόσθιους ή οριζόντιους ημικυκλικούς σωλήνες. Η αποσπασμένη ωτοκονία μετατοπίζεται όταν το κεφάλι κινείται, διεγείροντας τον κύπελλο για να στείλει ψευδή σήματα στον εγκέφαλο που δημιουργούν μια αίσθηση ίλιγγου<sup>4</sup>.*

Αιθουσαία νευρωνίτιδα.

Η αιθουσαία νευρίτιδα, επίσης γνωστή με την ονομασία αιθουσαία νευρωνίτιδα, πιστεύεται ότι προκαλείται από φλεγμονή του αιθουσαίου τμήματος του όγδου κρανιακού νεύρου και παρουσιάζεται με ίλιγγο, ναυτία και ανισορροπία βάδισης. Πιστεύεται ότι σχετίζεται με προηγούμενη ή συνοδευτική ιογενή λοίμωξη. Θεωρείται καλοήθης, αυτοπεριοριζόμενη κατάσταση που συνήθως διαρκεί αρκετές ημέρες, αλλά μπορεί να διαρκέσει εβδομάδες έως μήνες για να επιλυθούν πλήρως όλα τα αιθουσαία συμπτώματα. Δεν σχετίζεται με απώλεια ακοής, ενώ μπορεί να παρατηρηθεί οριζόντιος νυσταγμός κατά την κλινική εξέταση.

<sup>4</sup> <https://vestibular.org/article/diagnosis-treatment/types-of-vestibular-disorders/benign-paroxysmal-positional-vertigo-bppv/>

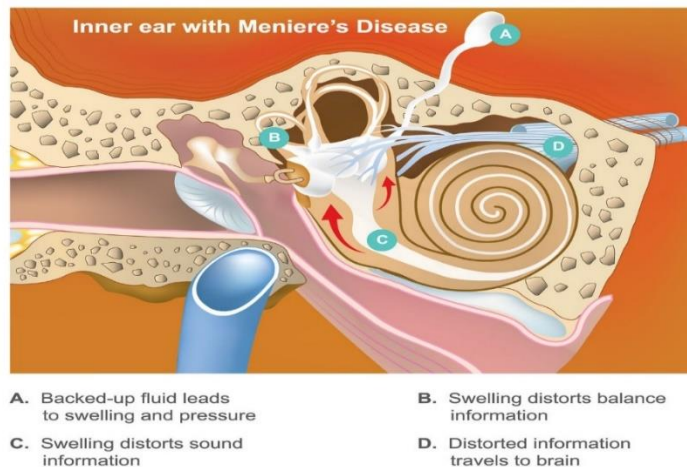
## Vestibular neuritis



**Εικόνα 5. Αιθουσαία Νευρωνίτιδα<sup>5</sup>**

## Νόσος του Meniere.

Η νόσος του Meniere είναι μια διαταραχή που επηρεάζει το έσω αυτί. Το εσωτερικό αυτί είναι υπεύθυνο για την ακοή και την ισορροπία. Η κατάσταση προκαλεί ίλιγγο και οδηγεί επίσης σε προβλήματα ακοής και έναν ήχο κουδουνίσματος στο αυτί. Η διάρκεια του επεισοδίου είναι από 30 λεπτά έως 24 ώρες. Η νόσος του Meniere προσβάλλει συνήθως μόνο ένα αυτί. Το Εθνικό Ινστιτούτο Κωφών και Άλλες Διαταραχές Επικοινωνίας (NIDCD) εκτιμά ότι 615.000 άνθρωποι στις Ηνωμένες Πολιτείες έχουν νόσο του Meniere. Περίπου 45.500 άτομα διαγιγνώσκονται κάθε χρόνο. Είναι πιο πιθανό να συμβεί σε άτομα ηλικιακής ομάδας μεταξύ 40 και 50 ετών. Η αιτία της νόσου του Meniere δεν είναι γνωστή, αλλά οι επιστήμονες πιστεύουν ότι προκαλείται από αλλαγές στο υγρό στους σωλήνες του εσωτερικού αυτιού. Άλλες προτεινόμενες αιτίες περιλαμβάνουν αυτοάνοση ασθένεια, αλλεργίες και γενετική.



**Εικόνα 6. Εσωτερικό αυτί με νόσο του Meniere<sup>6</sup>**

<sup>5</sup> <https://lmhofmeyr.co.za/conditions/dizziness-and-balance-disorders/vestibular-neuritis/>

<sup>6</sup> <https://www.healthyhearing.com/help/tinnitus/Menieres-disease>

## Αιθουσαία Ημικρανία.

Η αιθουσαία ημικρανία είναι ένα πρόβλημα του νευρικού συστήματος που προκαλεί επαναλαμβανόμενη ζάλη (ή ίλιγγο) σε άτομα που έχουν ιστορικό συμπτωμάτων ημικρανίας. Σε αντίθεση με τις παραδοσιακές ημικρανίες, οι αιθουσαίες ημικρανίες δεν προκαλούν πάντα πονοκέφαλο. Παράλληλα μπορεί να υπάρξουν συμπτώματα ζάλης που διαρκεί περισσότερο από λίγα λεπτά, ναυτία και έμετο, προβλήματα ισορροπίας και ευαισθησία στον ήχο. Η αιθουσαία ημικρανία διαγιγνώσκεται βάσει του ιστορικού του ασθενούς, απαντήσεις σε ερωτήσεις σχετικά με το πότε και πού εμφανίζονται τα συμπτώματα και με φυσική εξέταση.

## Επίμονη Ζάλη Ορθοστατικής Αντίληψης (Persistent Postural-Perceptual Dizziness – PPPD).

Η επίμονη ζάλη ορθοστατικής αντίληψης- (PPPD) είναι καθορισμένο διαγνωστικό σύνδρομο που ενοποιεί βασικά χαρακτηριστικά της χρόνιας υποκειμενικής ζάλης, του φοβικού ορθοστατικού ίλιγγου και των σχετικών διαταραχών. Περιγράφει μια κοινή χρόνια δυσλειτουργία του αιθουσαίου συστήματος και του εγκεφάλου που προκαλεί επίμονη ζάλη, μη περιστρεφόμενο ίλιγγο και / ή αστάθεια. Η διαταραχή αποτελεί μακροχρόνια κακή προσαρμογή σε ένα νευρο-ωτολογικό, ιατρικό ή ψυχολογικό συμβάν που προκάλεσε αιθουσαία συμπτώματα. Ενώ οι διαγνωστικές εξετάσεις και η συμβατική απεικόνιση συνήθως παραμένουν αρνητικές, οι ασθενείς με PPPD παρουσιάζονται με χαρακτηριστικό τρόπο που αντιστοιχεί σε θετικά διαγνωστικά κριτήρια. Οι ασθενείς συχνά αναπτύσσουν δευτερογενή λειτουργική διαταραχή βάδισης, άγχος, συμπεριφορά αποφυγής και σοβαρή αναπηρία.

Άλλες αιτίες που μπορεί να προκαλέσουν ίλιγγο είναι:

### Εγκεφαλικό

Μετά από ένα επεισόδιο εγκεφαλικού υπάρχει πιθανότητα εμφάνισης ίλιγγου. Η εμφάνιση του σε αυτή την περίπτωση είναι ξαφνική και μπορεί να διαρκέσει αρκετές ημέρες. Όταν το επεισόδιο αφορά την πρόσθια κάτω παρεγκεφαλιδική αρτηρία ή την έσω ακουστική αρτηρία μονόπλευρη απώλεια ακοής μπορεί να προκληθεί.

### Στρες- Ψυχογενής ίλιγγος.

Ένας πολύ σημαντικός παράγοντας επιδείνωσης ίλιγγου είναι το άγχος. Συνήθης αίτια εμφάνισης του είναι το πρωτοπαθές άγχος (γενικευμένο άγχος, κατάθλιψη, διαταραχή πανικού, αγοραφοβία κ.ά.) και το δευτεροπαθές άγχος ως μόνιμη εκδήλωση μετά από προηγούμενο αιθουσαίο ίλιγγο.

Το 25-50% των ασθενών, που είχαν νοσήσει από αιθουσαία νευρωνίτιδα και δεν ακολούθησαν ενεργητική αποκατάσταση, προσβάλλονται από άγχος και/ή κατάθλιψη και μη ειδικό ίλιγγο, που εκλύεται πιο εύκολα. Οι ασθενείς γενικεύουν το φόβο πτώσης τους (ψυχοκαταπόνηση). Ο ασθενής παρερμηνεύει τα ερεθίσματα της ιδιοδεκτικότητας.

### Ναυτία ταξιδιωτών.

Κατά τη διάρκεια ενός μεγάλου ταξιδιού με πλοίο, αεροπλάνο ή ακόμη και με αυτοκίνητο πολλοί άνθρωποι βιώνουν ίλιγγο. Η διάρκεια του ίλιγγου είναι βραχεία, αλλά συχνά προκαλεί ανησυχία.

## Συνήθειες που επιδεινώνουν τον ίλιγγο.

Ως επί το πλείστον οι αιτίες των ιλίγγων δεν μπορούν να ελεγχθούν από τους νοσούντες, όμως υπάρχουν κάποιες συνήθειες που τον επιδεινώνουν όπως για παράδειγμα η μεγάλη κατανάλωση αλκοόλ, καφεΐνης και αλατιού. Επίσης πρέπει να επισημανθεί ότι ο ίλιγγος μπορεί να προκληθεί ή να επιδεινωθεί από το άγχος, τις αλλαγές στη βαρομετρική πίεση, την αφυδάτωση και τη στέρηση του ύπνου. Επομένως συνήθειες όπως ο διαλογισμός και η γιόγκα, η ξεκούραση, η καλή ενυδάτωση, μπορούν να βοηθήσουν στη διαχείριση του.

## Μεθοδολογίες εξέτασης.

Για τη διάγνωση του ιλίγγου κατά την πρώτη επαφή με τον ασθενή γίνονται μια σειρά από εξετάσεις. Σε πρώτη φάση είναι η εξειδικευμένη νευροτολογική-ακοολογική εξέταση όπου πραγματοποιούνται δοκιμασίες ελέγχου αντανakλαστικών (αιθουσο-νωτιαίων και αιθουσο-οφθαλμικών) και πρόκλησης ιλίγγου θέσεως. Κάνοντας αυτές τις διαδικασίες ξεδιπλώνεται μια εικόνα για το ποιο σύστημα έχει υποστεί βλάβη από αυτά που βοηθούν στη διατήρηση της ισορροπίας του σώματος (οπτικό, αιθουσαίο, παρεγκεφαλιδικό, σωματοαισθητικό).

Για να ληφθούν όλες οι πληροφορίες σχετικά με την ακοή του ασθενούς η ακοολογική εξέταση είναι ενδεδειγμένη. Πιο συγκεκριμένα πραγματοποιείται ακοόγραμμα, τυμπανόγραμμα, ωτοακουστικές εκπομπές και ακουστικά προκλητά δυναμικά εγκεφαλικού στελέχους. Έτσι λαμβάνονται πληροφορίες τόσο για το αυτί, όσο και για την ακουστική οδό μεταφοράς των ερεθισμάτων στον εγκέφαλο. Τέλος σε εξέταση επέρχεται η λειτουργία του λαβυρίνθου με θερμικούς διακλυσμούς, ταλαντευόμενο έδρανο και ηλεκτρονυσταγμογραφία, που καθορίζουν τον πάσχοντα λαβύρινθο. Η δυναμική ισορροπομετρία σε συγκεκριμένες περιπτώσεις βοηθά στο διαχωρισμό των βλαβών των διαφόρων συστημάτων ισορροπίας.

Όλες οι παραπάνω εξετάσεις είναι εξειδικευμένες, αλλά σχετικά απλές και είναι απαραίτητες προκειμένου να αποκλειστούν παθήσεις με παρόμοια κλινική εικόνα, όμως πολλές φορές δεν είναι αρκετές για να καταλήξουμε στην αιτία του ιλίγγου.

## Το πρόβλημα : Δυσκολία διάγνωσης.

Ο ίλιγγος και η ζάλη είναι από τα πιο κοινά συμπτώματα στην ιατρική και συχνά έχουν σημαντικό αντίκτυπο στις δραστηριότητες της καθημερινής ζωής. Στη διάγνωση και στη θεραπεία των ιλίγγων εμπλέκονται νευρολόγοι, ωτορινολαρυγγολόγοι, παθολόγοι, καρδιολόγοι, ορθοπαιδικοί και ψυχίατροι. Κάθε γιατρός ερμηνεύει συνήθως τα συμπτώματα του ασθενούς από τη σκοπιά της ειδικότητάς του. Γι' αυτό για να βοηθηθεί αποτελεσματικά ο ασθενής, απαιτείται η συνεργασία πολλών ειδικοτήτων.

Στην πρώτη επαφή με τον ασθενή όπως προαναφέρθηκε απαιτείται εμπεριστατωμένη κλινική νευρολογική εξέταση που θα συμπεριλαμβάνει και δοκιμασίες πρόκλησης ιλίγγου που τις περισσότερες φορές θα δώσουν στοιχεία για το σύστημα που έχει υποστεί βλάβη: οπτικό, αιθουσαίο, παρεγκεφαλιδικό, σωματοαισθητικό.

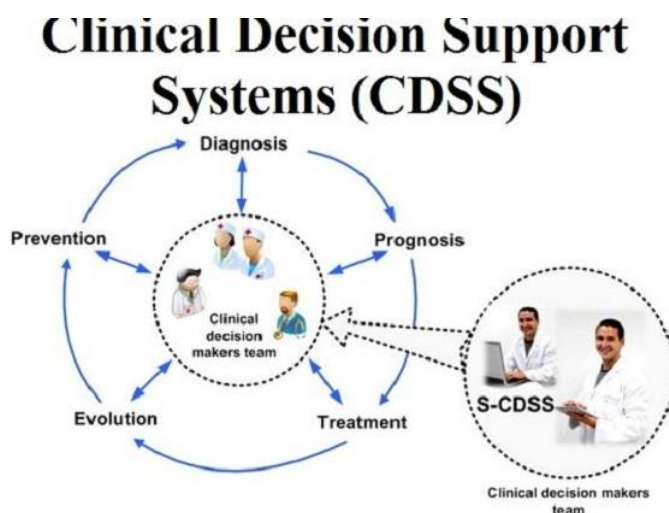
Κλινικές νευροφυσιολογικές εξετάσεις, όπως το ηλεκτρονυσταθογράφημα μπορούν να συμβάλλουν στη διάγνωση βλάβης κάποιων προαναφερθέντων συστημάτων. Η συσκευή αυτή μπορεί να δώσει πληροφορίες για βλάβες λαβυρίνθου και αιθουσαίου νεύρου αλλά έχει την δυνατότητα να δώσει πληροφορίες σχετικά με βλάβες του στελέχους και των ημισφαιρίων στην καθοδήγηση των κινήσεων των οφθαλμών. Πληροφορίες οι οποίες είναι πολύ σημαντικές καθώς έχει αποδειχθεί ότι ο προσανατολισμός του ασθενή στο χώρο επηρεάζεται από ειδικές οφθαλμοκινητικές βλάβες.

Επίσης οι βλάβες των αισθητικών οδών του κεντρικού και περιφερικού νευρικού συστήματος μπορούν να διαγνωστούν μέσω μυογραφήματος. Άλλες εξετάσεις που μπορούν να φανούν χρήσιμες σε πιο περίπλοκες καταστάσεις ιλίγγου είναι η μέθοδος της δυναμικής ποστουρογραφίας ή κινήσιογραφίας, που σε συνδυασμό με την κλινική νευρολογική εξέταση θα βοηθήσει στο διαχωρισμό των βλαβών των διαφόρων συστημάτων ισορροπίας. Από την άλλη το ακουόγραμμα και τα ακουστικά προκλητά δυναμικά δίνουν έμμεσα πληροφορίες για βλάβες στο λαβύρινθο και άμεσα για βλάβες στο 8<sup>ο</sup> εγκεφαλικό (αιθουσαίο και ακουστικό) νεύρο.

Τέλος, για την ανεύρεση στοιχείων σχετικά με διαταραχές στην αιμάτωση του εγκεφάλου χρήσιμη είναι η υπερηχογραφική διερεύνηση (Triplex καρωτίδων, σπονδυλικών αρτηριών, διακρανιακός υπέρηχος για τα ενδοκρανιακά αγγεία, υπέρηχος καρδιάς).

Σύμφωνα με όσα προαναφέρθηκαν ο ίλιγγος μπορεί να χαρακτηριστεί αρκετά περίπλοκος, καθώς τα πιθανά αίτια πρόκλησής του είναι πολλά και για τη διάγνωση θα απαιτηθεί συχνά ο συνδυασμός γνώσης προσωπικών στοιχείων του ασθενή, αποτελέσματα κλινικών εξετάσεων, ιστορικό υγείας καθώς και συμπτωμάτων αλλά και η γνώση και η εμπειρία γιατρών διαφορετικών ειδικοτήτων όπως προαναφέρθηκε.

Αντιμετώπιση: Κλινικά Συστήματα Υποστήριξης Απόφασης - Clinical Decision Support Systems (CDSS)



Εικόνα 7. Επισκόπηση Κλινικού Συστήματος Υποστήριξης Απόφασης<sup>7</sup>

<sup>7</sup> <https://www.kenresearch.com/blog/2019/11/global-clinical-decision-support-systems-market/>

Τα Κλινικά Συστήματα Υποστήριξης Αποφάσεων είναι βασικά εργαλεία που βοηθούν κλινικούς ιατρούς, διοικητικό προσωπικό, ασθενείς, φροντιστές ή άλλα μέλη της ομάδας φροντίδας. Τέτοιου τύπου συστήματα έχουν σχεδιαστεί για να βοηθήσουν στη διαλογή τεράστιων ποσοτήτων ψηφιακών δεδομένων για να προτείνουν τα επόμενα βήματα για θεραπείες, να ειδοποιήσουν τους παρόχους για τις διαθέσιμες πληροφορίες που μπορεί να μην έχουν δει ή να πιάσουν πιθανά προβλήματα, όπως επικίνδυνες αλληλεπιδράσεις φαρμάκων.

Τα CDSS σήμερα χρησιμοποιούνται κυρίως στο σημείο της φροντίδας, για να συνδυάσει ο γιατρός τις γνώσεις του με πληροφορίες ή προτάσεις που παρέχονται από το CDSS. Όλο και περισσότερο, υπάρχουν CDSS που αναπτύσσονται με την ικανότητα να αξιοποιούν δεδομένα και παρατηρήσεις που διαφορετικά δεν είναι εφικτές ή δεν ερμηνεύονται από τον άνθρωπο.

Ιατρικά λάθη και εσφαλμένες διαγνώσεις είναι ζητήματα που έχουν αρχίσει να εξετάζονται σοβαρά. Η μείωση τους θα συνέβαλλε σε άμεση βελτίωση της διάγνωσης αλλά και κατά συνέπειας της θεραπείας των ασθενών.

Σύμφωνα με έκθεση του ακαδημαϊκού περιοδικού BMJ εκτιμάται ότι το ιατρικό σφάλμα είναι η τρίτη κύρια αιτία θανάτου στις Ηνωμένες Πολιτείες, υπεύθυνη για περισσότερους από 400.000 θανάτους ετησίως. Επίσης η Εθνική Ακαδημία Επιστημών, Μηχανικών και Ιατρικής (NASEM) ανέφερε ότι τα διαγνωστικά σφάλματα συμβάλλουν περίπου στο 10% των θανάτων ασθενών στις ΗΠΑ και αντιπροσωπεύουν έως και το 17% των ανεπιθύμητων ενεργειών στο νοσοκομείο.

Τα Κλινικά Συστήματα Υποστήριξης Απόφασης λοιπόν, μπορούν να βελτιώσουν την πιθανότητα διαγνωστικών λαθών χάρη στην δυνατότητα ενσωμάτωσης δεδομένων και πληροφοριών σε κλινικά σχετικές γνώσεις προσφέροντας ιατρική ακρίβεια και οικονομική και εξατομικευμένη φροντίδα ασθενών.

Η πολυπλοκότητα της αναγνώρισης των αιτιών που προκαλούν ίλιγγο οδήγησε στην ανάγκη για εύρεση προτύπων και κανόνων. Κανόνες και πρότυπα που θα συνδυάζουν όλα τα απαραίτητα χαρακτηριστικά του ασθενή –προσωπικές πληροφορίες, ιατρικό ιστορικό, συμπτώματα αλλά και αποτελέσματα εξετάσεων και εμπειρία του εκάστοτε ιατρού. Διάφορες έρευνες έχουν δημοσιευθεί και παρουσιάζουν μοντέλα πρόβλεψης ίλιγγου αλλά και Κλινικά Συστήματα Υποστήριξης Απόφασης που βοηθούν στην διάγνωση της αιτίας πρόκλησής του.

## Παρουσίαση ερευνών με συναφή προβλήματα.

### Ανάπτυξη Διαγνωστικού Συστήματος Υποστήριξης Απόφασης για καλοήγη παροξυσμικό ίλιγγο θέσεως χρησιμοποιώντας μοντέλο βαθιάς μάθησης (2019)

Στην έρευνα αυτή στόχος ήταν η ανάπτυξη ενός Διαγνωστικού Συστήματος Υποστήριξης Απόφασης που θα διαγνώσκει τον καλοήγη παροξυσμικό ίλιγγο θέσεως χρησιμοποιώντας μοντέλο βαθιάς μάθησης. Η ακριβής ερμηνεία του νυσταγμού που προκαλείται από δοκιμές θέσης μπορεί να έχει ως αποτέλεσμα τη διάγνωση καλοήθους παραροξυσμικού ίλιγγου θέσης. Ωστόσο, δυσκολίες στην ερμηνεία της κίνησης των ματιών μπορεί συχνά να προκύψουν στην πρακτική πρωτοβάθμιας φροντίδας ή στο δωμάτιο

έκτακτης ανάγκης. Για την επίλυση του προβλήματος η χρήση της μηχανικής μάθησης θα ήταν χρήσιμη για την ερμηνεία.

Έτσι από την αποθήκη κλινικών δεδομένων, ελήφθησαν 91.778 βίντεο νυσταγμού από 3467 ασθενείς με ζάλη, στα οποία η τρισδιάστατη κίνηση του νυσταγμού σχολιάστηκε από τέσσερις ωτολογικούς εμπειρογνώμονες. Από κάθε βίντεο με ετικέτα, 30 δυνατότητες άλλαξαν σε 255 εικόνες πλέγματος που εισήχθησαν στο επίπεδο εισόδου του νευρικού δικτύου για το σύνολο δεδομένων εκπαίδευσης. Για την επικύρωση του μοντέλου, συλλέχθηκαν σύνολο δεδομένων βίντεο 3566 οριζόντιων, 2068 κατακόρυφων και 720 κυκλοστροφικών κινήσεων από 1005 ασθενείς με BPPV.

Τα αποτελέσματα έδειξαν ότι το μοντέλο είχε ευαισθησία και ειδικότητα  $0,910 \pm 0,036$  και  $0,919 \pm 0,032$  για οριζόντιο νυσταγμό,  $0,879 \pm 0,029$  και  $0,894 \pm 0,025$  για κάθετο νυσταγμό και  $0,783 \pm 0,040$  και  $0,799 \pm 0,038$  για κυκλοστροφικό νυσταγμό, αντίστοιχα. Το προσβεβλημένο κανάλι προβλέφθηκε με ευαισθησία  $0,806 \pm 0,010$  και ειδικότητα  $0,971 \pm 0,003$ .

Συμπερασματικά λοιπόν, η έρευνα αυτή έδειξε ότι δεδομένου ότι το μοντέλο βαθιάς μάθησης είχε υψηλή ευαισθησία και ειδικότητα για την ταξινόμηση του νυσταγμού και τον εντοπισμό του προσβεβλημένου καναλιού σε ασθενείς με BPPV, μπορεί να έχει ευρεία κλινική εφαρμογή.

Πηγή: «*Developing a Diagnostic Decision Support System for Benign Paroxysmal Positional Vertigo Using a Deep-Learning Model*», <https://pubmed.ncbi.nlm.nih.gov/31072056/>

Μηχανογραφημένο σύστημα κλινικών αποφάσεων και εφαρμογή για κινητά με εξειδικευμένη υποστήριξη για τη βελτιστοποίηση της διαχείρισης του ίλιγγου στην πρωτοβάθμια φροντίδα: πρωτόκολλο μελέτης για μια πραγματιστική κατά συστάδα-τυχαιοποιημένη ελεγχόμενη δοκιμή (2020).

Το πρόβλημα ορθής διάγνωσης των αιτιών που προκαλούν ίλιγγο απασχόλησε προσφάτως τους ερευνητές στην Γερμανία. Αν και πολλές αιτίες ίλιγγου και ζάλης μπορούν εύκολα να αναγνωριστούν, οι ασθενείς συχνά λαμβάνουν ακατάλληλη και αναποτελεσματική θεραπεία. Οι λόγοι για αυτό είναι διάφοροι. Επειδή ο ίλιγγος είναι ένα διεπιστημονικό σύμπτωμα και υπάρχει έλλειψη τυποποιημένων διαγνωστικών εργαλείων, είναι εύκολο να χαθεί η επισκόπηση των πιθανών διαφορικών διαγνώσεων. Ωστόσο, υπάρχουν ενδείξεις ότι η διαχείριση ασθενών με ίλιγγο μπορεί να βελτιστοποιηθεί χρησιμοποιώντας τυποποιημένες οδούς φροντίδας με ψηφιακή υποστήριξη.

Στο πλαίσιο της “PoiSe - prevention, online feedback, and interdisciplinary therapy of acute vestibular syndromes by e-health” («Πρόληψη της διαδικτυακής ανατροφοδότησης και της διεπιστημονικής θεραπείας των οξέων αιθουσαίων συνδρόμων από την ηλεκτρονική υγεία») η έρευνα αυτή στοχεύει στην αξιολόγηση της εφαρμογής ενός προγράμματος με πολλά αλληλοσυνδεόμενα στοιχεία.

Τα τρία βασικά στοιχεία είναι ένα μηχανογραφημένο σύστημα κλινικών αποφάσεων, μια εφαρμογή για κινητά και ένα συμβουλευτικό και διεπιστημονικό εκπαιδευτικό πρόγραμμα που αναπτύχθηκε από το Γερμανικό Κέντρο Διαταραχών του Ίδρυματος και της Ισορροπίας (DSGZ). Η μελέτη είναι μια κατά

συστάδα-τυχαιοποιημένη ελεγχόμενη δοκιμή με σχεδιασμό παράλληλων ομάδων, καθώς και μια λεπτομερή αξιολόγηση της διαδικασίας. Οι συστάδες περιλαμβάνουν πρακτικές ιατρών πρωτοβάθμιας περίθαλψης στη Βαυαρία της Γερμανίας.

Στο πλαίσιο της μελέτης θα αξιολογηθεί η αποτελεσματικότητα, η αποδοχή και η αποδοτικότητα της παρέμβασης. Αναμένεται ότι η παρέμβαση θα βελτιώσει την ποιότητα και την αποτελεσματικότητα της διαχείρισης των ζαλισμένων ασθενών. Η υψηλότερη διαγνωστική ακρίβεια, η βελτιστοποιημένη θεραπεία και η παρακολούθηση της εξέλιξης της νόσου αναμένεται να βελτιώσει τα αποτελέσματα που σχετίζονται με τον ασθενή και να μειώσει το κόστος της υγειονομικής περίθαλψης.

---

#### ΤΥΧΑΙΟΠΟΙΗΜΕΝΗ ΕΛΕΓΧΟΜΕΝΗ ΔΟΚΙΜΗ

(ΑΓΓΛΙΚΟΣ ΟΡΟΣ: RANDOMIZED CONTROLLED TRIAL)

ΟΡΙΣΜΟΣ: ΠΕΙΡΑΜΑΤΙΚΗ ΜΕΛΕΤΗ ΓΙΑ ΝΑ ΕΚΤΙΜΗΘΟΥΝ ΤΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΜΙΑΣ ΣΥΓΚΕΚΡΙΜΕΝΗΣ ΜΕΤΑΒΛΗΤΗΣ (Π.Χ. ΦΑΡΜΑΚΟΥ Η ΘΕΡΑΠΕΙΑΣ), ΣΤΗΝ ΟΠΟΙΑ ΤΑ ΑΤΟΜΑ ΕΧΟΥΝ ΤΥΧΑΙΑ ΕΝΤΑΧΘΕΙ ΣΤΗΝ ΠΕΙΡΑΜΑΤΙΚΗ ΟΜΑΔΑ, ΣΤΗΝ ΟΜΑΔΑ PLACEBO, Η ΣΤΗΝ ΟΜΑΔΑ ΕΛΕΓΧΟΥ. Η ΠΕΙΡΑΜΑΤΙΚΗ ΟΜΑΔΑ ΛΑΜΒΑΝΕΙ ΤΟ ΦΑΡΜΑΚΟ Η ΑΚΟΛΟΥΘΕΙ ΤΗ ΔΙΑΔΙΚΑΣΙΑ· ΤΟ ΦΑΡΜΑΚΟ ΤΗΣ ΟΜΑΔΑΣ PLACEBO ΕΧΕΙ ΚΑΛΥΦΘΕΙ ΩΣΤΕ ΝΑ ΠΡΟΣΟΜΟΙΑΖΕΙ ΣΤΟ ΦΑΡΜΑΚΟ ΠΟΥ ΕΡΕΥΝΑΤΑΙ. Η ΟΜΑΔΑ ΕΛΕΓΧΟΥ ΔΕΝ ΛΑΜΒΑΝΕΙ ΤΙΠΟΤΕ. ΕΡΓΑΣΤΗΡΙΑΚΕΣ ΕΞΕΤΑΣΕΙΣ Η ΚΛΙΝΙΚΕΣ ΕΚΤΙΜΗΣΕΙΣ ΓΙΝΟΝΤΑΙ ΣΤΙΣ ΟΜΑΔΕΣ (ΣΥΧΝΑ ΧΡΗΣΙΜΟΠΟΙΩΝΤΑΣ ΤΗ ΔΙΠΛΗ ΤΥΦΛΗ ΤΕΧΝΙΚΗ) ΓΙΑ ΝΑ ΚΑΘΟΡΙΣΘΟΥΝ ΤΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ ΦΑΡΜΑΚΟΥ Η ΤΗΣ ΔΙΑΔΙΚΑΣΙΑΣ. ΣΥΝΤΟΜΟΓΡΑΦΙΑ : RCT

---

Πηγή: «Computerized clinical decision system and mobile application with expert support to optimize management of vertigo in primary care: study protocol for a pragmatic cluster-randomized controlled trial», <https://pubmed.ncbi.nlm.nih.gov/32719973/>

Δεδομένα διαταραχών ισορροπίας εξόρυξης για την ανάπτυξη διαγνωστικών συστημάτων υποστήριξης αποφάσεων - EMBalance διαγνωστικό σύστημα υποστήριξης αποφάσεων (2016)

Σε αυτήν την εργασία αφού έγινε αναφορά σε προγενέστερα διαγνωστικά συστήματα υποστήριξης αποφάσεων (DCSS) παρουσιάστηκε η μεθοδολογία την ανάπτυξη ενός νέου του συστήματος του EMBalance.

Για την ανάπτυξη του παραπάνω συστήματος χρησιμοποιήθηκαν τεχνικές εξόρυξης δεδομένων σε ιατρικά δεδομένα από ασθενείς με διαταραχές ισορροπίας . Η διαφοροποίηση στην έρευνα αυτή βρίσκεται στα χαρακτηριστικά των ασθενών , δεν λαμβάνει υπόψη μόνον τα συμπτώματα του ασθενούς

αλλά συνδυάζει από δημογραφικά χαρακτηριστικά έως κλινική εξέταση, ακουστικές και αιθουσαίες δοκιμές, προκειμένου να παρέχει μια ακριβή διάγνωση.

Το σύστημα στοχεύει στην δημιουργία ενός συστήματος παροχής υποστήριξης αποφάσεων σε γενικούς ιατρούς και ειδικούς στη διάγνωση διαταραχών ισορροπίας, καθώς και στην παροχή συστάσεων για τις κατάλληλες πληροφορίες και δεδομένα που πρέπει να ζητούνται σε κάθε βήμα της διαγνωστικής διαδικασίας.

Λεπτομερή αποτελέσματα παρέχονται για τη διάγνωση 12 διαταραχών ισορροπίας, τόσο για ιατρούς όσο και για ειδικούς. Συνολικά, η αναφερόμενη ακρίβεια κυμαίνεται από 59,3 έως 89,8% για τους ιατρούς και από 74,3 έως 92,1% για τους ειδικούς.

Πηγή: «Mining balance disorders' data for the development of diagnostic decision support systems», <https://pubmed.ncbi.nlm.nih.gov/27619194/>

Διαφορική μέθοδος διαγνωστικής συλλογιστικής για καλοήγη παροξυσμικό ίλιγγο θέσεως-BPPV με βάση δυναμικό αβέβαιο γράφημα αιτιότητας (2019).

Οι υπότυποι του καλοήθους παροξυσμικού ίλιγγου θέσεως διακρίνονται ανάλογα από το ποιος ημικυκλικός σωλήνας συμμετέχει στην πρόκληση του ίλιγγου και από το εάν η αποσπασθείσα ωτοκονία αιωρείται μέσα στην ενδολέμφο του ημικύκλιου σωλήνα, που έχει προσβληθεί (καναλολιθίαση) ή είναι προσκολλημένη στη λήκυθο (ληκυθολιθίαση).

Η ακριβής διαφοροποίηση των υποτύπων του καλοήθους παροξυσμικού ίλιγγου θέσεως (BPPV) μπορεί να βελτιώσει σημαντικά την αποτελεσματικότητα της επανατοποθέτησης ελιγμών στη θεραπεία του και έτσι να μειώσει τις περιττές κλινικές δοκιμές και ακατάλληλα φάρμακα.

Σε αυτήν τη μελέτη, έχουν γίνει προσπάθειες για την ανάπτυξη προσεγγίσεων μοντελοποίησης αιτιότητας και διαγνωστικού συλλογισμού σχετικά με τις αβεβαιότητες που μπορεί να προκύψουν από ιατρικές πληροφορίες.

Κατασκευάστηκε ένα δυναμικό αβέβαιο μοντέλο διαφορικής διάγνωσης βασισμένο σε γραφήματα αιτιότητας για το BPPV, το οποίο περιλαμβάνει 354 μεταβλητές και 885 τόξα αιτιότητας.

Για να γίνει διάκριση μεταξύ των διαφόρων αιτιών των ιλιγγικών διαταραχών με παρόμοια συμπτώματα, η προτεινόμενη μέθοδος περιλαμβάνει μια γραφική και συμπαγή αναπαράσταση, εκτός από λογικούς αλγόριθμους συλλογιστικής για παθολογικούς μηχανισμούς και σχετικές αβέβαιες αιτίες. Νέοι αλγόριθμοι προτάθηκαν επίσης για διαφορική διάγνωση μέσω λογικών και πιθανολογικών συμπερασμάτων, με έμφαση στην επίλυση των προβλημάτων των περίπλοκων και συγχυτικών παραγόντων της νόσου, ελλιπών κλινικών παρατηρήσεων και ανεπαρκών δειγμάτων δεδομένων. Αυτή η μελέτη χρησιμοποιεί περαιτέρω περιπτώσεις ίλιγγου για να ελέγξει την απόδοση της προτεινόμενης μεθόδου στην κλινική πρακτική.

Τα αποτελέσματα δείχνουν υψηλή ακρίβεια, ικανοποιητική ικανότητα διάκρισης για BPPV και ευνοϊκή ευρωστία όσον αφορά ελλείψεις ιατρικές πληροφορίες. Οι υποκείμενοι παθολογικοί μηχανισμοί και η

σημασιολογία της αιτιότητας επαληθεύονται χρησιμοποιώντας συμπαγή γραφική αναπαράσταση και διαδικασία συλλογισμού, που ενισχύουν την ερμηνεία των συμπερασμάτων της διάγνωσης.

Πηγή: « *Differential Diagnostic Reasoning Method for Benign Paroxysmal Positional Vertigo Based on Dynamic Uncertain Causality Graph*», <https://www.hindawi.com/journals/cmmm/2020/1541989/>

### Σχεδιασμός ενός συστήματος υποστήριξης κλινικών αποφάσεων για τη διαχείριση και τη θεραπεία ασθενών με κύριο παράπονο τον ίλιγγο (2018)

Άλλη μια μελέτη που είχε ως αντικείμενο ζήτημα την αιτιολογία του ίλιγγου και τη σωστή ερμηνεία του από τους γιατρούς στα τμήματα έκτακτης ανάγκης (επείγοντα) πραγματοποιήθηκε το 2017 στο Τμήμα Έκτακτης Ανάγκης, νοσοκομείο Rasoul-e Akram, Τεχεράνη, Ιράν. Η χρήση φορητού λογισμικού που εγκαθίσταται στο κινητό μπορεί να βοηθήσει τους κλινικούς ιατρούς να μειώσουν τις πιθανότητες λάθους διάγνωσης στην επείγουσα θεραπεία ασθενών.

Αφού δημιουργήθηκε πρωτόκολλο διαχείρισης ασθενών με ίλιγγους βάση των προτύπων και της βιβλιογραφίας, το πρωτόκολλο αυτό εγκρίθηκε από την ειδική ομάδα. Για τη θεραπεία και τη διαχείριση της νόσου. Σύνολο 190 ασθενών με ίλιγγους διαγνώστηκαν χρησιμοποιώντας δύο μεθόδους για τη θεραπεία και τη διαχείριση της νόσου. Για την αξιολόγηση των δύο ομάδων χρησιμοποιήθηκε σχεδιασμένο σύστημα υποστήριξης κλινικών αποφάσεων για την πρώτη ομάδα – δηλαδή οι ασθενείς είχαν στο κινητό εγκατεστημένο το φορητό λογισμικό και μέθοδο ρουτίνας για τη δεύτερη ομάδα αντίστοιχα. Η διάρκεια της θεραπείας, η διάρκεια συντήρησης και οι διαφορές μεταξύ της πρωτογενούς και της τελικής διάγνωσης συγκρίθηκαν μεταξύ των ομάδων.

Οι διαφορές μεταξύ της πρωτογενούς και της τελικής διάγνωσης ήταν σημαντικές και στις δύο ομάδες ( $P < 0,050$ ). Η μέση διάρκεια θεραπείας ήταν 3,38 στην πρώτη ομάδα χρησιμοποιώντας σύστημα υποστήριξης ενώ 4,96 ώρες στη δεύτερη ομάδα χρησιμοποιώντας μεθόδους ρουτίνας, αντίστοιχα ( $P < 0,001$ ). Η μέση περίοδος νοσηλείας ήταν σημαντικά μικρότερη στην ομάδα του συστήματος υποστήριξης (9,37 ώρες) σε σύγκριση με την ομάδα που ακολούθησε μεθόδους ρουτίνας (11,17 ώρες) ( $P < 0,001$ ). Το επίπεδο ικανοποίησης του ιατρού με το σύστημα υποστήριξης ήταν κατά μέσο όρο (47,9).

Συμπέρασμα: Η χρήση του συστήματος υποστήριξης κλινικών αποφάσεων μπορεί να βοηθήσει τους γιατρούς να βελτιώσουν τη διάγνωση, να μειώσουν την περίοδο νοσηλείας και να διαχειριστούν τους ασθενείς με το κύριο παράπονο του ίλιγγου πολύ καλύτερα.

Πηγή: « *Designing a clinical decision support system for managing and treating patients with the chief complaint of vertigo*», <http://eprints.iums.ac.ir/6924/>

Ανάπτυξη και επικύρωση ενός αλγορίθμου ταξινόμησης για τη διάγνωση και τη διαφοροποίηση αυθόρμητων επεισοδίων σύνδρομων ιλίγγου: αποτελέσματα από το μητρώο ασθενών DizzyReg (2020).

Ακόμα και οι πιο έμπειροι γιατροί είναι δύσκολο να διαχωρίσουν τα αυθόρμητα επεισόδια ιλίγγου, δηλαδή να διαχωρίσουν αν η αιτία του ιλίγγου προέρχεται από αιθουσαία ημικρανία ή αν είναι νόσος του Meniere. Η λήψη αυτών τέτοιου είδους αποφάσεων- που απαρτίζονται από σύνθετες διαγνωστικές πληροφορίες- μπορούν να υποστηριχτούν από αυτοματοποιημένα συστήματα. Έτσι η διάγνωση αιθουσαίας ημικρανίας και νόσος του Meniere θα μπορεί να πραγματοποιηθεί με ακεραιότητα με τη βοήθεια της μηχανικής μάθησης.

Τα δεδομένα αυτής της μελέτης προέρχονται από το υποψήφιο μητρώο ασθενών του Γερμανικού Κέντρου για Διαταραχές του Ίλιγγου και της Ισορροπίας, ένα εξειδικευμένο κέντρο τριτοβάθμιας θεραπείας στο Πανεπιστημιακό Νοσοκομείο του Μονάχου. Η αποστολή ταξινόμησης ήταν η διαφοροποίηση περιπτώσεων αιθουσαίας ημικρανίας και νόσος του Meniere από άλλες οντότητες αιθουσαίων νόσων. Τα βαθιά νευρωνικά δίκτυα (Deep Neural Networks -DNN) και τα ενισχυμένα δέντρα απόφασης (Boosted Decision Trees -BDT) χρησιμοποιήθηκαν για την ταξινόμηση.

Αποτελέσματα: Συμπεριλήφθηκαν συνολικά 1357 ασθενείς (μέση ηλικία 52,9, SD 15,9, 54,7% γυναίκες), 9,9% με νόσο του Meniere και 15,6% με αιθουσαία ημικρανία. Τα μοντέλα DNN απέδωσαν ακρίβεια ταξινόμησης (accuracy)  $98,4 \pm 0,5\%$ , ακρίβεια (precision)  $96,3 \pm 3,9\%$ , και ευαισθησία (recall)  $85,4 \pm 3,9\%$  για αιθουσαία ημικρανία, και ακρίβεια ταξινόμησης  $98,0 \pm 1,0\%$ , ακρίβεια (precision)  $90,4 \pm 6,2\%$  και ευαισθησία (recall)  $89,9 \pm 4,6\%$  για νόσο του Meniere. Το BDT απέδωσε ακρίβεια ταξινόμησης (accuracy)  $84,5 \pm 0,5\%$ , ακρίβεια (precision)  $51,8 \pm 6,1\%$ , ευαισθησία  $16,9 \pm 1,7\%$  για αιθουσαία ημικρανία, και ακρίβεια ταξινόμησης (accuracy)  $93,3 \pm 0,7\%$ , ακρίβεια (precision)  $76,0 \pm 6,7\%$ , ευαισθησία (recall)  $41,7 \pm 2,9\%$  για νόσο του Meniere.

Η σωστή διάγνωση των αυθόρμητων επεισοδίων αιθουσαίων συνδρόμων είναι πρόκληση στην κλινική πρακτική. Οι σύγχρονες μέθοδοι μηχανικής μάθησης θα μπορούσαν να αποτελέσουν τη βάση για την ανάπτυξη διαγνωστικών συστημάτων που βοηθούν τους επαγγελματίες και τους ιατρούς στις καθημερινές αποφάσεις τους για θεραπεία.

Πηγή: « *Development and validation of a classification algorithm to diagnose and differentiate spontaneous episodic vertigo syndromes: results from the DizzyReg patient registry*», <https://pubmed.ncbi.nlm.nih.gov/32661715>

## Μηχανική μάθηση.

Η δημιουργία μοντέλων ή προτύπων από ένα σύνολο δεδομένων ονομάζεται Μηχανική Μάθηση. Η Μηχανική Μάθηση πρόκειται για υποπεριοχή της Τεχνητής Νοημοσύνης. Πρόκειται για μεθόδους αναπαράστασης και αλγορίθμους που βοηθούν στη βελτίωση ενός υπολογιστικού συστήματος ,αφού

μαθαίνει να λύνει μόνο του το πρόβλημα. Εφαρμόζεται στο τομέα ανάλυσης δεδομένων αλλά και στο πλαίσιο συστημάτων υποστήριξης απόφασης. Εφαρμογή των μεθοδολογιών της μηχανικής μάθησης σε μεγάλες βάσεις δεδομένων αποτελεί η Εξόρυξη Δεδομένων (Data Mining) ή αλλιώς Ανακάλυψη Γνώσης σε Βάσεις Δεδομένων (Knowledge Discovery in Databases). Έχουν αναπτυχθεί πολλές τεχνικές μηχανικής μάθησης που χρησιμοποιούνται ανάλογα με τη φύση του προβλήματος και εμπίπτουν σε ένα από τα παρακάτω δυο είδη:

- ❖ μάθηση με επίβλεψη (supervised learning) ή μάθηση με παραδείγματα (learning from examples),
- ❖ μάθηση χωρίς επίβλεψη (unsupervised learning) ή μάθηση από παρατήρηση (learning from observation).

Κύριας σημασίας αποτελεί η Μηχανική Μάθηση στον κλάδο της Ιατρικής. Μοντέλα πρόβλεψης εξέλιξης ασθένειας ή Κλινικά Συστήματα Υποστήριξης Απόφασης είναι εφικτό να κατασκευαστούν από δεδομένα ασθενών-εξετάσεις, βιομετρικά χαρακτηριστικά, συμπτώματα.

### Προεπεξεργασία δεδομένων.

Κατά την διαδικασία εξόρυξης γνώσης πρώτο και κύριο βήμα είναι η προεπεξεργασία των δεδομένων. Τα δεδομένα πολλές φορές είναι απαραίτητο να μετασχηματιστούν σε μορφή καταλληλότερη της αρχικής έτσι ώστε να χρησιμοποιηθούν στην συνέχεια της ανάλυσης μας. Οι μετασχηματισμοί των δεδομένων εξαρτώνται είτε από τη μορφή των δεδομένων είτε από τη φύση των αλγορίθμων εκμάθησης. Οι αλγόριθμοι μπορεί να είναι μαθηματικοί ή λογικοί ή ακόμα και μετασχηματισμοί βασισμένοι στη γνώση του πεδίου. Παρατηρείται συχνά στις εφαρμογές εξόρυξης γνώσης ότι η διαχείριση των χαρακτηριστικών από τους αλγορίθμους που εφαρμόζονται είναι αδύνατη λόγω του μεγάλου όγκου τους. Συνήθως επιλέγεται ένα υποσύνολο δεδομένων προς χρήση για στη διαδικασία εκμάθησης αφού τα περισσότερα χαρακτηριστικά είναι ευκρινώς μη συσχετιζόμενα ή περιττά. Αν και οι μέθοδοι εκμάθησης επιτελούν τη διαδικασία επιλογής των κατάλληλων χαρακτηριστικών και απόρριψης των υπολοίπων, η αποδοτικότητά τους μπορεί να βελτιωθεί κατά πολύ.

### Εξισορρόπηση.

Η ταξινόμηση της προγνωστικής μοντελοποίησης περιλαμβάνει πρόβλεψη μιας ετικέτας τάξης για μια δεδομένη παρατήρηση. Ένα ανισορροπημένο πρόβλημα ταξινόμησης είναι ένα παράδειγμα ενός προβλήματος ταξινόμησης όπου η κατανομή των παραδειγμάτων στις γνωστές τάξεις είναι προκατειλημμένη ή παραμορφωμένη. Η κατανομή μπορεί να ποικίλει από μια μικρή προκατάληψη έως μια σοβαρή ανισορροπία, όπου υπάρχει ένα παράδειγμα στην τάξη των μειονοτήτων για εκατοντάδες, χιλιάδες ή εκατομμύρια παραδείγματα στην τάξη ή τις τάξεις της πλειοψηφίας.

Ο λόγος για τον οποίο εντοπίζουμε την ανισορροπημένη ταξινόμηση ως πρόβλημα είναι επειδή μπορεί να επηρεάσει την απόδοση στους αλγόριθμους Μηχανικής Μάθησης. Ένας τρόπος που η ανισορροπία μπορεί να επηρεάσει τον αλγόριθμο Μηχανικής Μάθησης είναι όταν ο αλγόριθμος αγνοεί εντελώς την κατηγορία μειονοτήτων, και αυτό διότι η τάξη των μειονοτήτων είναι συχνά η τάξη που μας ενδιαφέρει περισσότερο. Για την αντιμετώπιση της ανισορροπημένης ταξινόμησης υπάρχουν δύο μέθοδοι:

- ❖ Υπερ δειγματοληψία - Αναπαραγωγή δειγμάτων από την κατηγορία μειοψηφίας
- ❖ Υποδειγματοληψία- Διαγραφή δειγμάτων από την κατηγορία πλειοψηφίας

Η μη ισορροπημένη ταξινόμηση είναι το πρόβλημα της ταξινόμησης όταν υπάρχει άνιση κατανομή τάξεων στο σύνολο δεδομένων εκπαίδευσης. Η ανισορροπία στην κατανομή τάξης μπορεί να ποικίλλει, αλλά μια σοβαρή ανισορροπία είναι πιο δύσκολο να μοντελοποιηθεί και μπορεί να απαιτεί εξειδικευμένες τεχνικές.

Τόσο η υπερ-δειγματοληψία όσο και η υπο-δειγματοληψία περιλαμβάνουν την εισαγωγή μιας προκατάληψης για την επιλογή περισσότερων δειγμάτων από μια τάξη από μια άλλη, για την αντιστάθμιση μιας ανισορροπίας που είτε υπάρχει ήδη στα δεδομένα είτε είναι πιθανό να αναπτυχθεί εάν ληφθεί ένα καθαρά τυχαίο δείγμα.

### Αλγόριθμοι κατηγοριοποίησης.

Η κατηγοριοποίηση αποτελεί μια από τις βασικότερες τεχνικές της εξόρυξης δεδομένων, η οποία εξετάζοντας τα γνωρίσματα ενός στιγμιότυπου το αντιστοιχεί σε μία προκαθορισμένη κλάση. Η κατηγοριοποίηση συναντάται στη βιβλιογραφία και ως ταξινόμηση. Ένα μοντέλο κατηγοριοποίησης κατασκευάζεται ως εξής :ο αλγόριθμος επεξεργάζεται τα δεδομένα του συνόλου εκπαίδευσης και χτίζει ένα μοντέλο, αφού έχει κατασκευαστεί το μοντέλο ελέγχεται η ικανότητα του να προβλέπει την κλάση άγνωστων παρατηρήσεων. Σε περίπτωση που το μοντέλο αποδίδει ικανοποιητικά τότε συνίσταται η χρήση του μοντέλου για τη διατύπωση προβλέψεων. Οι μέθοδοι κατηγοριοποίησης αξιολογούνται με βάση τη ακρίβεια πρόβλεψης-δηλαδή την ικανότητα του μοντέλου να προβλέπει την κατηγορία μιας νέας περίπτωσης, τη ταχύτητα, την ερμηνευσιμότητα, την επεκτασιμότητα και την ανθεκτικότητα. Το σημαντικότερο μέτρο μέτρησης απόδοσης είναι η ακρίβεια στην πρόβλεψη της κλάσης. Το μέγεθος της απόδοσης του συστήματος στα εξεταζόμενα δεδομένα αντικατοπτρίζει το πόσο ένα μοντέλο που καλύπτει καλά ένα «ικανό» σύνολο δεδομένων καλύπτει καλά και το σύνολο των δεδομένων λειτουργίας.

### Κατηγοριοποίηση κατά Κ-πλησιέστερων γειτόνων K-NN.

Ο Κ-πλησιέστερος γείτονας (K- NN) ταξινομητής είναι ένας από τους απλούστερους αλγόριθμους μηχανικής μάθησης που βασίζονται στην τεχνική εποπτευόμενης μάθησης. Ο αλγόριθμος K-NN αναλαμβάνει την ομοιότητα μεταξύ της νέας υπόθεσης των δεδομένων και των διαθέσιμων περιπτώσεων και θέτει τη νέα υπόθεση στην κατηγορία που είναι πιο παρόμοια με τις διαθέσιμες κατηγορίες.

Ο αλγόριθμος K-NN αποθηκεύει όλα τα διαθέσιμα δεδομένα και ταξινομεί ένα νέο σημείο δεδομένων με βάση την ομοιότητα. Αυτό σημαίνει ότι όταν εμφανίζονται νέα δεδομένα τότε μπορούν εύκολα να ταξινομηθούν σε μια νέα κατηγορία χρησιμοποιώντας τον αλγόριθμο K-NN.Ο αλγόριθμος KNN στη φάση

της εκπαίδευσης αποθηκεύει μόνο το σύνολο δεδομένων και όταν λαμβάνει νέα δεδομένα, τότε ταξινομεί αυτά τα δεδομένα σε μια κατηγορία που μοιάζει πολύ με τα νέα δεδομένα.

Ο αλγόριθμος αυτός μπορεί να χρησιμοποιηθεί τόσο για την παλινδρόμηση όσο και για την ταξινόμηση, αλλά ως επί το πλείστον χρησιμοποιείται για τα προβλήματα ταξινόμησης. Ονομάζεται επίσης αλγόριθμος τεμπέλης εκμάθησης επειδή δεν μαθαίνει αμέσως από το εκπαιδευτικό σύνολο, αλλά αποθηκεύει το σύνολο δεδομένων και κατά τη στιγμή της ταξινόμησης, εκτελεί μια ενέργεια στο σύνολο δεδομένων.

### Κατηγοριοποίηση με Δέντρο Απόφασης (Decision Tree).

Ο ταξινομητής Δέντρο Απόφασης (Decision Tree) μπορεί να χρησιμοποιηθεί τόσο για προβλήματα ταξινόμησης όσο και για προβλήματα παλινδρόμησης, αλλά ως επί το πλείστον προτιμάται για την επίλυση προβλημάτων ταξινόμησης. Είναι ένας ταξινομητής με δομή δέντρου όπου οι εσωτερικοί κόμβοι αντιπροσωπεύουν τα χαρακτηριστικά ενός συνόλου δεδομένων, οι κλάδοι αντιπροσωπεύουν τους κανόνες απόφασης και κάθε κόμβος φύλλων αντιπροσωπεύει το αποτέλεσμα.

Σε ένα Δέντρο Απόφασης, υπάρχουν δύο κόμβοι, ο κόμβος απόφασης και ο κόμβος φύλλων. Οι κόμβοι απόφασης χρησιμοποιούνται για τη λήψη οποιασδήποτε απόφασης και έχουν πολλούς κλάδους, ενώ οι κόμβοι φύλλων είναι το αποτέλεσμα αυτών των αποφάσεων και δεν περιέχουν περαιτέρω κλάδους. Οι αποφάσεις ή το τεστ εκτελούνται βάσει των χαρακτηριστικών του δεδομένου συνόλου δεδομένων. Είναι μια γραφική παράσταση για την εξεύρεση όλων των πιθανών λύσεων σε ένα πρόβλημα / απόφαση βάσει συγκεκριμένων συνθηκών. Τέλος ένα δέντρο αποφάσεων απλώς θέτει μια ερώτηση, και με βάση την απάντηση (Ναι / Όχι), χωρίζει περαιτέρω το δέντρο σε υποδέντρα.

### Κατηγοριοποίηση κατά Naive Bayes.

Οι ταξινομητές Naive Bayes είναι μια συλλογή αλγορίθμων ταξινόμησης που βασίζονται στο Θεώρημα του Bayes.

*Το Θεώρημα του Bayes βρίσκει την πιθανότητα ενός συμβάντος, δεδομένης της πιθανότητας ενός άλλου γεγονότος που έχει ήδη συμβεί. Το θεώρημα του Bayes δηλώνεται μαθηματικά ως η ακόλουθη εξίσωση:*

$$P(A | B) = \{P(B | A) P(A)\} / \{P(B)\}$$

όπου  $A$  και  $B$  είναι συμβάντα .

Βασικά, είναι η πιθανότητα του συμβάντος  $A$ , δεδομένου ότι το συμβάν  $B$  είναι αλήθεια. Το συμβάν  $B$  ονομάζεται επίσης ως απόδειξη (evidence).

Δεν είναι ένας μοναδικός αλγόριθμος, αλλά μια οικογένεια αλγορίθμων όπου όλοι τους μοιράζονται μια κοινή αρχή, ότι κάθε ζεύγος χαρακτηριστικών που ταξινομούνται είναι ανεξάρτητο το ένα από το άλλο. Τα χαρακτηριστικά θεωρείται ότι είναι ανεξάρτητα και κάθε χαρακτηριστικό έχει το ίδιο βάρος ή σημασία. Έτσι κανένα από τα χαρακτηριστικά δεν είναι άσχετο και θεωρείται ότι συμβάλλει εξίσου στο αποτέλεσμα.

## Κατηγοριοποίηση με Random Forest

Ο Random Forest ταξινομητής δημιουργεί ένα σύνολο δέντρων αποφάσεων από τυχαία επιλεγμένο υποσύνολο του εκπαιδευτικού συνόλου. Στη συνέχεια συγκεντρώνει τις ψήφους από διαφορετικά δέντρα αποφάσεων για να αποφασίσει την τελική τάξη του αντικειμένου δοκιμής.

Ας υποθέσουμε ότι το σετ εκπαίδευσης δίνεται ως:  $[X1, X2, X3, X4]$  με αντίστοιχες ετικέτες ως  $[L1, L2, L3, L4]$ , ο Random Forest μπορεί να δημιουργήσει τρία δέντρα αποφάσεων λαμβάνοντας υπόψη το υποσύνολο για παράδειγμα,

$[X1, X2, X3]$

$[X1, X2, X4]$

$[X2, X3, X4]$

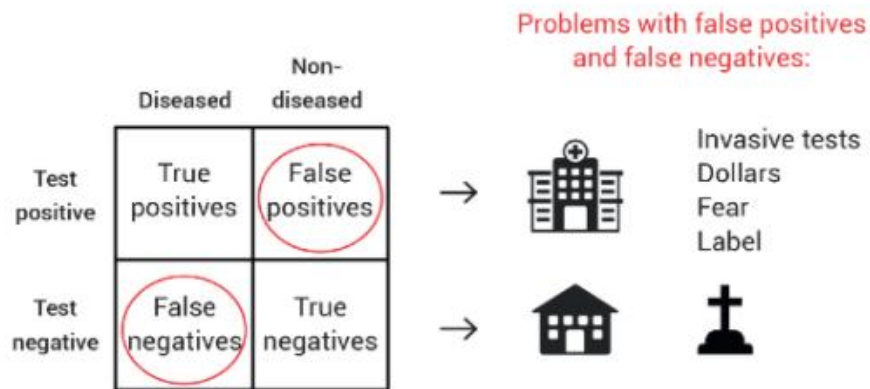
Τελικά, προβλέπει με βάση την πλειοψηφία των ψήφων από καθένα από τα δέντρα απόφασης που έχουν ληφθεί. Αυτό λειτουργεί καλά επειδή ένα δέντρο αποφάσεων μπορεί να είναι επιρρεπές σε θόρυβο, αλλά το σύνολο πολλών δέντρων αποφάσεων μειώνει την επίδραση του θορύβου δίνοντας πιο ακριβή αποτελέσματα. Τα υποσύνολα σε διαφορετικά δέντρα αποφάσεων που δημιουργούνται ενδέχεται να αλληλεπικαλύπτονται.

Εναλλακτικά, ο Random Forest μπορεί να εφαρμόσει την έννοια του βάρους για να εξετάσει τον αντίκτυπο του αποτελέσματος από οποιοδήποτε δέντρο αποφάσεων. Δέντρο με υψηλό ποσοστό σφάλματος δίνεται χαμηλή τιμή βάρους και αντίστροφα. Αυτό θα αύξανε την επίπτωση στην απόφαση των δέντρων με χαμηλό ποσοστό σφάλματος.

## Μέτρα ακριβείας.

Η απόδοση ενός μοντέλου ταξινόμησης μπορεί να περιγραφεί μέσω ενός πίνακα, τον πίνακα Σύγχυσης (Confusion matrix). Στον πίνακα αυτό εμφανίζεται η κατανομή των αποτελεσμάτων του μοντέλου. Οι ποσότητες οι οποίες είναι σημαντικές για τα αποτελέσματα του μοντέλου δίνονται ως εξής :

- ❖ Αληθώς θετικά (TP-True Positives): Οι περιπτώσεις στις οποίες η πρόβλεψη ήταν θετική και η πραγματική έξοδος ήταν επίσης θετική .
- ❖ Αληθώς αρνητικά (TN-True Negative): Οι περιπτώσεις στις οποίες η πρόβλεψη ήταν αρνητική και η πραγματική έξοδος ήταν αρνητική.
- ❖ Ψευδώς θετικά (FP-False Positives): Οι περιπτώσεις στις οποίες η πρόβλεψη ήταν θετική και η πραγματική έξοδος ήταν αρνητική.
- ❖ Ψευδώς αρνητικά (FN-False Negatives): Οι περιπτώσεις στις οποίες η πρόβλεψη ήταν αρνητική και η πραγματική έξοδος ήταν θετική.



*Εικόνα 8. Πίνακας Σύγχυσης-Confusion Matrix*

Με βάση τις παραπάνω ποσότητες δημιουργούνται μέτρα ακριβείας της απόδοσης του μοντέλου, τα οποία παρατείνονται παρακάτω :

### Ακρίβεια Ταξινόμησης

Η ακρίβεια (accuracy) ορίζεται ως ο λόγος του αριθμού των σωστών προβλέψεων προς τον συνολικό αριθμό των δειγμάτων εισαγωγής.

$$\frac{TP+TN}{\text{Total Sample}}$$

### Ευαισθησία ή αλλιώς Ανάκληση (Sensitivity or Recall)

Ευαισθησία ή αλλιώς ανάκληση ή αλλιώς αληθινό σωστό ποσοστό (Sensitivity or Recall or True Positive Rate) αντιστοιχεί στην αναλογία των θετικών σημείων δεδομένων που θεωρούνται σωστά ως θετικά, σε σχέση με όλα τα θετικά σημεία δεδομένων.

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

### Ειδικότητα-Specificity

Ειδικότητα ή αλλιώς αληθινό αρνητικό ποσοστό (Specificity or True Negative Rate) αντιστοιχεί στην αναλογία των αρνητικών σημείων δεδομένων που θεωρούνται σωστά ως αρνητικά, σε σχέση με όλα τα αρνητικά σημεία δεδομένων.

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

### Ακρίβεια-Precision

Ακρίβεια (Precision) αντιστοιχεί στην αναλογία των σωστών θετικών αποτελεσμάτων σε σχέση με τον αριθμό των θετικών αποτελεσμάτων που προβλέπονται από τον ταξινομητή.

$$TP/(TP+FP)$$

Μέτρο F - F-measure

Το μέτρο F (F-measure) είναι ένα μέτρο της ακρίβειας ενός μοντέλου σε ένα σύνολο δεδομένων. Χρησιμοποιείται για την αξιολόγηση δυαδικών συστημάτων ταξινόμησης, τα οποία ταξινομούν τα παραδείγματα σε «θετικά» ή «αρνητικά». Το μέτρο F είναι ένας τρόπος συνδυασμού της ακρίβειας και της ανάκλησης του μοντέλου και ορίζεται ως ο αρμονικός μέσος της ακρίβειας και ανάκλησης του μοντέλου.

$$2 * \text{Precision} * \text{Recall} / (\text{Precision} + \text{Recall})$$

|                                                     |  |
|-----------------------------------------------------|--|
| Ευαισθησία ή ποσοστό αληθώς θετικών (TPR)           |  |
| $TPR = \frac{TP}{P} = \frac{TP}{TP + FN}$           |  |
| Ποσοστό ψευδώς θετικών (FPR)                        |  |
| $FPR = \frac{FP}{N} = \frac{FP}{FP + TN}$           |  |
| Ειδικότητα ή ποσοστό αληθώς αρνητικών (SPC)         |  |
| $SPC = \frac{TN}{N} = \frac{TN}{FP + TN} = 1 - FPR$ |  |
| Ακρίβεια (Accuracy)                                 |  |
| $ACC = \frac{TP + TN}{P + N}$                       |  |

Εικόνα 9. Πίνακας μέτρων ακριβείας

Αξιολόγηση.

Ανεξάρτητα από το πόσο ισχυρό είναι ένα μοντέλο μηχανικής μάθησης, δεν μπορεί ποτέ να έχει αρκετά εκπαιδευτικά δεδομένα. Αφού εκπαιδεύσει το μοντέλο με το σετ προπόνησης, ο χρήστης θα προχωρήσει στην επικύρωση των αποτελεσμάτων και στο συντονισμό των υπερπαραμέτρων με το σετ επικύρωσης έως ότου ο χρήστης φτάσει σε ικανοποιητική μέτρηση απόδοσης. Μόλις ολοκληρωθεί αυτό το στάδιο, ο χρήστης θα προχωρήσει στη δοκιμή δεδομένων. Τυχαίος θόρυβος (δηλ. σημεία δεδομένων που καθιστούν δύσκολη την εμφάνιση ενός μοτίβου), η χαμηλή συχνότητα μιας συγκεκριμένης κατηγορικής μεταβλητής, η χαμηλή συχνότητα της κατηγορίας στόχου (εάν η μεταβλητή στόχος είναι κατηγορηματική) και οι λανθασμένες αριθμητικές τιμές κ.λπ. είναι μόνο μερικοί λόγοι που μπορούν να βλάψουν ένα μοντέλο. Ενώ η διαδικασία επικύρωσης δεν μπορεί να βρει άμεσα τι είναι λάθος, η διαδικασία μπορεί να μας δείξει μερικές φορές ότι υπάρχει πρόβλημα με τη σταθερότητα του μοντέλου.

Η πιο βασική μέθοδος επικύρωσης των δεδομένων (δηλαδή συντονισμός των υπερπαραμέτρων πριν από τη δοκιμή του μοντέλου) είναι η εκτέλεση διαχωρισμού προπόνησης/ επικύρωσης / δοκιμής στα δεδομένα. Μια τυπική αναλογία για αυτό μπορεί να είναι 80/10/10 για να βεβαιωθεί ότι υπάρχουν ακόμα του μοντέλου με το σετ δοκιμών για να προβλέψει και να αξιολογήσει την απόδοση.

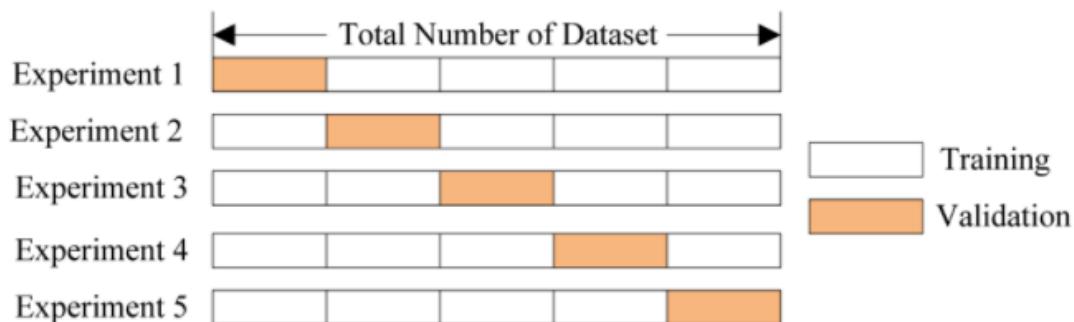


*Εικόνα 10. Εκτέλεση διαχωρισμού προπόνησης/επικύρωσης/δοκιμής δεδομένων*

### Κ-πλη διασταυρωμένη επικύρωση.

Η διασταυρούμενη επικύρωση είναι μια στατιστική μέθοδος που χρησιμοποιείται για την εκτίμηση της ικανότητας των μοντέλων μηχανικής μάθησης. Χρησιμοποιείται συνήθως στην εφαρμοσμένη μηχανική εκμάθηση για να συγκρίνει και να επιλέξει ένα μοντέλο για ένα δεδομένο πρόβλημα προγνωστικής μοντελοποίησης, επειδή είναι εύκολο να γίνει κατανοητό, εύκολο να εφαρμοστεί και οδηγεί σε εκτιμήσεις δεξιοτήτων που γενικά έχουν χαμηλότερη προκατάληψη από άλλες μεθόδους.

Κατά την Κ-πλη διασταυρωμένη επικύρωση ένα δεδομένο σύνολο δεδομένων χωρίζεται σε έναν αριθμό Κ πτυχών, όπου κάθε πτυχή χρησιμοποιείται ως σύνολο δοκιμών σε κάποιο σημείο. Στο σενάριο της 5-δυπλής εγκάρσιας επικύρωσης ( $K = 5$ ), το σύνολο δεδομένων χωρίζεται σε 5 πτυχές. Στην πρώτη επανάληψη, το πρώτο δίπλωμα χρησιμοποιείται για τη δοκιμή του μοντέλου και τα υπόλοιπα χρησιμοποιούνται για την εκπαίδευση του μοντέλου. Στη δεύτερη επανάληψη, η 2η πτυχή χρησιμοποιείται ως σύνολο δοκιμών, ενώ τα υπόλοιπα χρησιμεύουν ως σύνολο προπόνησης. Αυτή η διαδικασία επαναλαμβάνεται έως ότου κάθε πτυχή των 5 πτυχών να χρησιμοποιηθεί ως σύνολο δοκιμών.



*Εικόνα 11. Παράδειγμα 5-πλής διασταυρωμένης επικύρωσης*

## Εφαρμογή στο WEKA.

### Ερευνητικός χώρος.

Ο ίλιγγος είναι ένας συγκεκριμένος τύπος ζάλης που χαρακτηρίζεται από την ξαφνική αίσθηση ότι γυρίζεται ή ότι ο κόσμος γύρω σας περιστρέφεται. Το συναίσθημα μπορεί να έρθει και να φύγει ή μπορεί να διαρκέσει για ώρες ή ημέρες. Μαζί με τη ζάλη, οι άνθρωποι μπορεί επίσης να παρουσιάσουν ναυτία ή έμετο, πονοκεφάλους, διπλή όραση ή κτύπο της καρδιάς. Οι αιτίες που μπορούν να τον προκαλέσουν είναι πολλές . Μια σειρά από εξετάσεις σε συνδυασμό με τα συμπτώματα του ασθενούς είναι η πιο αποτελεσματική μέθοδος διάγνωσης των αιτιών πρόκλησης ίλιγγου. Ωστόσο η πολυπλοκότητα ερμηνείας του συμπτώματος οδηγεί πολλές φορές σε λάθος αποτελέσματα.

Για να μειωθεί ο κίνδυνος λανθασμένων διαγνώσεων ,αρκετά Κλινικά Συστήματα Υποστήριξης Απόφασης βασισμένα στην μηχανική μάθηση έχουν προταθεί τα τελευταία χρόνια. Τέτοιου τύπου συστήματα βοηθούν τους γιατρούς στην απόφασή τους για το τύπο της αιτίας πρόκλησης του ίλιγγου .

Στα πλαίσια αυτής της εργασίας θα γίνει αναφορά σε ένα σύνολο δεδομένων που αφορά στοιχεία από 987 ασθενείς που συλλέχθηκαν από το National Hospital of Neurology, Queen, Square, την 1<sup>η</sup> Οτορινολαρυγγολογική Πανεπιστημιακή Κλινική Αθηνών ,Ελλάδα, Πανεπιστήμιο της Αμβέρσας του Βελγίου και της Πανεπιστημιακής κλινικής του Φράιμπουργκ, Γερμανία.

Πρόκειται για σύνολο δεδομένων ασθενών και μπορεί να χρησιμοποιηθεί για να προβλέψει ποια χαρακτηριστικά συσχετίζονται με κάθε μία από τις 13 αιτίες πρόκλησης του ίλιγγου.

| Diseases                                                     |    |
|--------------------------------------------------------------|----|
| Disease bppv posterior                                       | 1  |
| Disease meniere                                              | 2  |
| Disease vestibular neuronitis                                | 3  |
| Disease bilateral vestibular                                 | 4  |
| Disease vestibular migraine                                  | 5  |
| Disease cpa lesion acousticus neurinoma                      | 6  |
| Disease cerebellar pontine lesion                            | 7  |
| Disease vestibular paroxysmia                                | 8  |
| Disease chronic subjective dizziness pppd                    | 9  |
| Disease mental disorders                                     | 10 |
| Disease unilateral peripheral vestibular dysfunction failure | 11 |
| Disease bppv-1                                               | 12 |
| Disease bppv-2                                               | 13 |

*Εικόνα 12 . Πίνακας μεταβλητών ασθενειών*

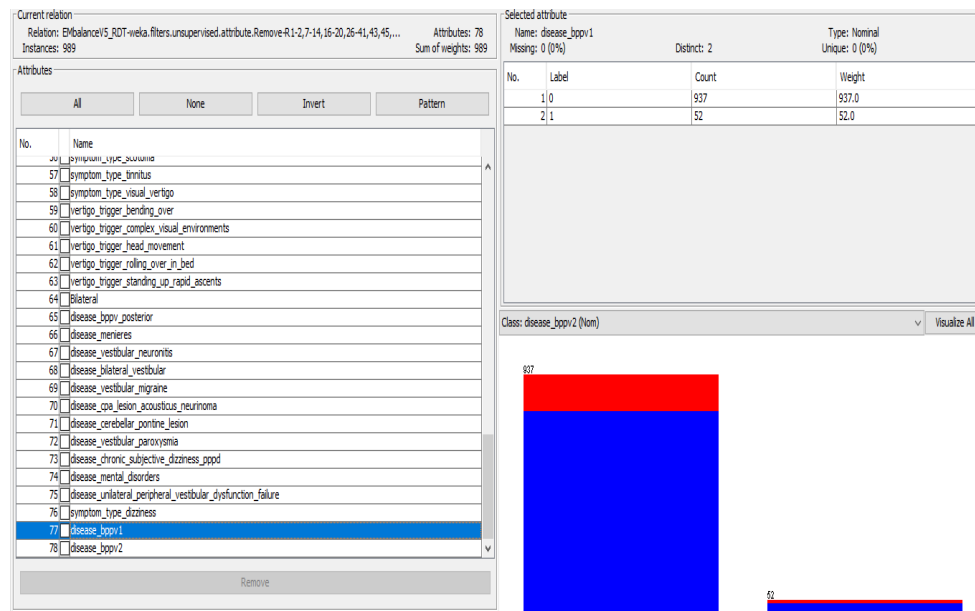
Στόχος είναι να εξετάσουμε την δυνατότητα πρόβλεψης ασθένειας ή μη ασθένειας που προκαλεί ίλιγγο με δεδομένα κάποια γνωρίσματα ασθενών μελετώντας την ακρίβεια του κάθε αλγορίθμου που θα

χρησιμοποιήσουμε. Για κάθε αλγόριθμο θα εξεταστεί από στατιστικής απόψεως η απόδοσή του μοντέλου δοκιμάζοντας τη μείωση της διάστασης, επιλέγοντας δηλαδή συγκεκριμένα χαρακτηριστικά εισόδου κάθε φορά. Για να επιτευχθεί αυτό θα χρησιμοποιηθούν αλγόριθμοι αξιολόγησης χαρακτηριστικών, καθώς πολλές φορές η ύπαρξη μεγάλου αριθμού γνωρισμάτων όπως και η ύπαρξη γνωρισμάτων που δε σχετίζονται με το υπό εξέταση χαρακτηριστικό εξόδου μπορεί να δημιουργήσει προβλήματα στην εκπαίδευση ορισμένων μοντέλων και να μειωθεί η αξιοπιστία του.

Με τη βοήθεια του εργαλείου WEKA θα κατασκευαστούν μοντέλα πρόβλεψης χαρακτηριστικών που συνδέονται με τις παρακάτω ασθένειες που προκαλούν ίλιγγο : disease BPPV-1, disease BPPV-2, disease Meniere ,disease Vestibular Migraine και disease Vestibular Neuronitis.

### 1ο Βήμα: Explorer =>Preprocess => Open File

Και “φορτώνουμε” το αρχείο “embalance\_grouped\_patient\_Task5\_3.csv”

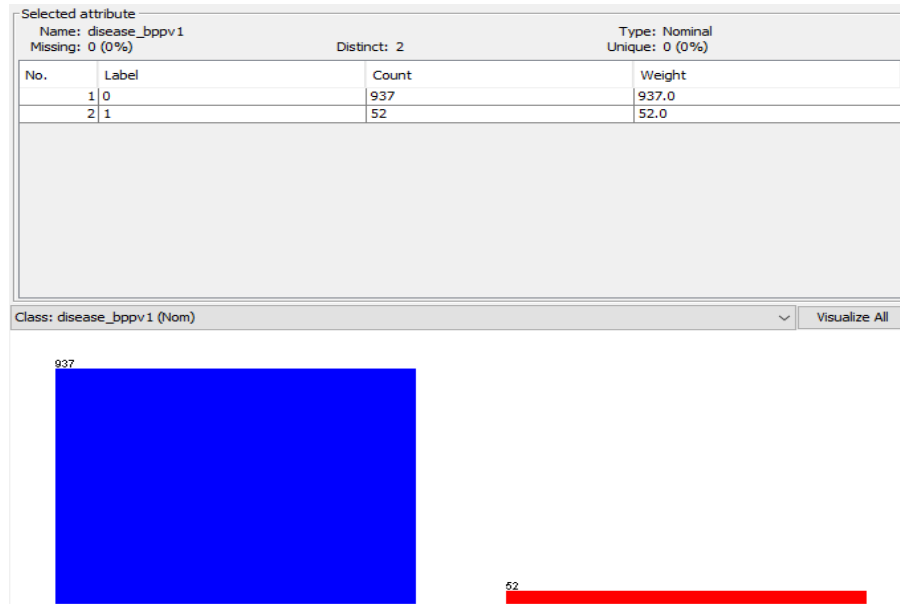


Εικόνα 13. Αρχικό σεν δεδομένων

Αποτελέσματα αλγορίθμων για την ασθένεια BPPV-1 (Παροξυσμικός ίλιγγος θέσεως τύπου 1) χωρίς επιλογή χαρακτηριστικών.

### Μεταβλητή στόχος: BPPV-1.

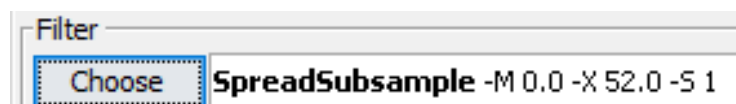
Επιλέγουμε το χαρακτηριστικό disease BPPV-1 ως αυτό το οποίο δείχνει σε ποια κλάση ανήκει κάθε φορά το υπόδειγμα και αφαιρούμε τις υπόλοιπες 12 μεταβλητές που δηλώνουν ασθένεια από το δείγμα. Το χαρακτηριστικό “ disease BPPV-1” είναι μία μεταβλητή δίτιμη με 1 να δηλώνει ότι ο εξεταζόμενος έχει την ασθένεια και 0 ότι ο εξεταζόμενος δεν έχει την ασθένεια BPPV-1.



*Εικόνα 14. Γράφημα εξόδου BPPV-1 πριν την προεπεξεργασία*

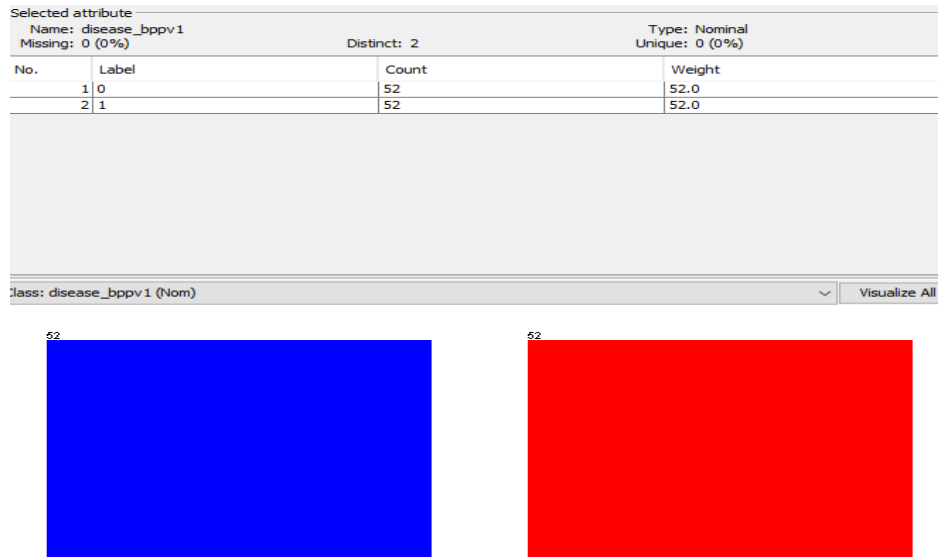
Η έξοδος του σετ δεδομένων- BPPV-1 παίρνει 2 διαφορετικές τιμές (βλ.Εικόνα14). Εφαρμόστηκε σε αυτό το σετ δεδομένων ο K-NN ταξινομητής με "default" παραμέτρους.

Παρατηρήθηκε ότι ο ταξινομητής δεν χωρίζει σωστά τις κλάσεις αφού υπάρχει ανισοκατανομή στα αποτελέσματα των κατηγοριών στα TP και FP . Από την μεριά της Μηχανικής μάθησης απαιτούνται ισορροπημένες κλάσεις για να γίνει εκπαίδευση. Επομένως για να υπάρχει ισοκατανομή στην έξοδο εφαρμόστηκε η τεχνική εξισορρόπησης SpreadSubsample κάνοντας undersampling –υπό δειγματοληψία (διαγραφή δειγμάτων από την κατηγορία πλειοψηφίας).



*Εικόνα 15. Φίλτρο SpreadSubsample*

Επιλέχθηκαν μόνο 52 χαρακτηριστικά από τη μεταβλητή με τιμή 0 και το γράφημα μετά την εξισορρόπηση έχει ως εξής:



*Εικόνα 16. Γράφημα εξόδου BPPV-1 μετά την εξισορρόπηση*

**2ο Βήμα:** Επιλέγουμε στο πάνελ την καρτέλα «Classify» όπου περιλαμβάνονται ποικίλες τεχνικές ταξινόμησης με διαφορετικές δυνατότητες για το χρήστη ή κάθε μία. Εφαρμόστηκαν τεχνικές στις οποίες η κατηγοριοποίηση των δεδομένων υλοποιείται μέσω διαφορετικών αλγορίθμων ταξινόμησης.

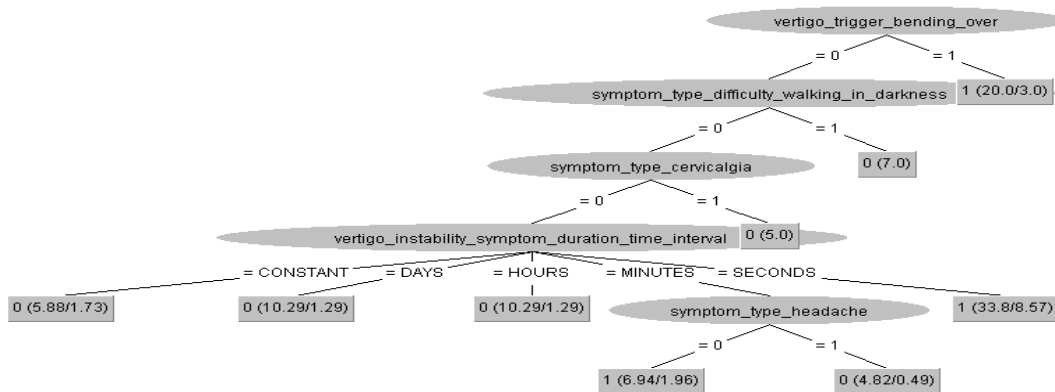
Ακολουθούμε την παρακάτω διαδικασία:

**Choose=>Weka=>Classifiers**

Οι αλγόριθμοι που χρησιμοποιήθηκαν είναι οι εξής:

- ❖ Ταξινόμηση μέσω του αλγορίθμου του πλησιέστερου γείτονα (K-NN)
- ❖ Ταξινόμηση μέσω του αλγορίθμου J-48
- ❖ Ταξινόμηση μέσω του αλγορίθμου Naive Bayes
- ❖ Ταξινόμηση μέσω του αλγορίθμου Random Forest

Για τον έλεγχο απόδοσης των παραπάνω αλγορίθμων χρησιμοποιήθηκε η τεχνική 10 fold cross validation (10-πλη διασταυρωτική) κατά την οποία το σύνολο δεδομένων διασπάται σε 10 ίσα σύνολα δεδομένων. Κατόπιν επιλέγεται ένα από αυτά τυχαία ως σύνολο εκπαίδευσης και το δέντρο εκπαιδεύεται με βάση αυτό και εξετάζει τα υπόλοιπα 9 σύνολα. Αυτό επαναλαμβάνεται 10 φορές ώστε κάθε υποσύνολο να έχει την ευκαιρία να είναι σύνολο εκπαίδευσης. Στην συνέχεια εξάγονται τα στατιστικά στοιχεία για την απόδοση των αλγορίθμων σε επιτυχία πρόβλεψης νέων τιμών δεδομένων αλλά και το Δέντρο Απόφασης του J-48.



**Εικόνα 17. Δέντρο Απόφασης για την ασθένεια BPPV-1**

Σύμφωνα με το Δέντρο Απόφασης η ασθένεια BPPV τύπου 1 εμφανίζεται αν ο ασθενής κάνοντας κάμψεις έχει ίλιγγο . Αλλιώς αν δεν έχει ίλιγγο αλλά έχει δυσκολία να περπατήσει στο σκοτάδι . Αν δεν έχει ίλιγγο κάνοντας κάμψεις ,δεν έχει δυσκολία να περπατήσει και έχει και πόνο στον αυχένα τότε έχει BPPV τύπου 1. Αν πάλι δεν έχει ίλιγγο κάνοντας κάμψεις , δεν έχει δυσκολία να περπατήσει, δεν έχει πόνο στον αυχένα και το χρονικό διάστημα διάρκειας αστάθειας του ίλιγγου είναι είτε συνέχεια ,είτε μέρες ,είτε ώρες ,είτε δευτερόλεπτα είτε λεπτά αλλά με συμπτώματα πονοκεφάλου τότε ο ασθενής έχει BPPV-1 .

| Μέτρα Ακριβείας BPPV-1 |                          |                                           |                                  |         |         |           |             |           |
|------------------------|--------------------------|-------------------------------------------|----------------------------------|---------|---------|-----------|-------------|-----------|
| classifier             |                          | Correctly Classified Instances - Accuracy | Incorrectly Classified Instances | TP Rate | FP Rate | Precision | Specificity | F-Measure |
| 1-nn                   | before feature selection | 74,04%                                    | 25,96%                           | 0,74    | 0,26    | 0,75      | 0,71        | 0,74      |
| 8-nn                   | before feature selection | 75,96%                                    | 24,04%                           | 0,76    | 0,24    | 0,77      | 0,72        | 0,76      |
| J48                    | before feature selection | 71,15%                                    | 28,85%                           | 0,71    | 0,29    | 0,72      | 0,67        | 0,71      |
| Naïve Bayes            | before feature selection | 75,00%                                    | 25,00%                           | 0,75    | 0,75    | 0,75      | 0,76        | 0,75      |
| Random Forest          | before feature selection | 76,92%                                    | 23,08%                           | 0,77    | 0,23    | 0,77      | 0,75        | 0,77      |

**Πίνακας 1. Συγκριτικά αποτελέσματα μεταξύ των αλγορίθμων ταξινόμησης για την ασθένεια BPPV-1**

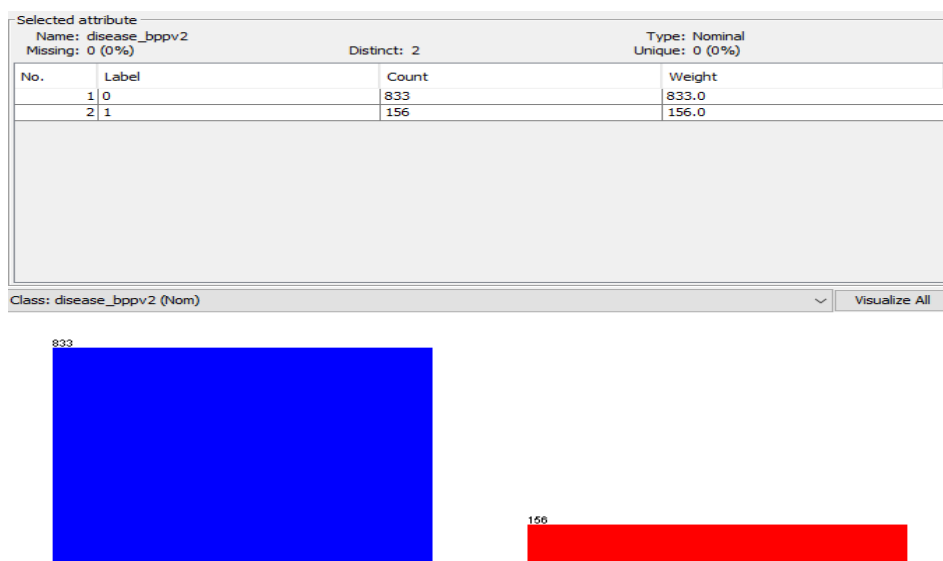
Από τα αποτελέσματα τα οποία προέκυψαν, τα οποία παρουσιάζονται συγκεντρωτικά στον παραπάνω πίνακα, διαπιστώνουμε ότι:

1. Μεταξύ των αλγορίθμων του πλησιέστερου γείτονα καλύτερη απόδοση είχε ο 8<sup>ος</sup> πλησιέστερος γείτονας, του οποίου τα αποτελέσματα ήταν τα πιο υψηλά.
2. Χειρότερη απόδοση διακρίνουμε στον αλγόριθμο J-48 του οποίου τα αποτελέσματα ταξινομήθηκαν λιγότερο από τα αποτελέσματα των υπόλοιπων αλγορίθμων.
3. Ο αλγόριθμος με την καλύτερη απόδοση και μέγιστη προβλεπτική ικανότητα είναι ο αλγόριθμος Random Forest για τον οποίο παρατηρείται το μεγαλύτερο ποσοστό ακρίβειας, μεγαλύτερο ποσοστό σωστά ταξινομημένων υποδειγμάτων- 76,93%, το μεγαλύτερο ποσοστό των θετικών υποδειγμάτων που έχουν αναγνωριστεί σωστά ως θετικά (μεγαλύτερη ευαισθησία) –0,77, υψηλότερες τιμές για όλα τα υπό εξέταση μέτρα ακρίβειας (precision, recall, f-measure).
4. Ο ταξινομητής με το μεγαλύτερο ποσοστό των αρνητικών υποδειγμάτων που έχουν αναγνωριστεί ως αρνητικά (μεγαλύτερη ειδικότητα) είναι ο Naive Bayes με 0,76.

Αποτελέσματα αλγορίθμων για την ασθένεια BPPV-2 (Παροξυσμικός ίλιγγος θέσεως τύπου 2) χωρίς την επιλογή χαρακτηριστικών.

#### Μεταβλητή στόχος: BPPV-2

Ακολουθώντας την ίδια μεθοδολογία επιλέγουμε το χαρακτηριστικό disease BPPV-2 ως αυτό το οποίο δείχνει σε ποια κλάση ανήκει κάθε φορά το υπόδειγμα και αφαιρούμε όπως είναι απαραίτητο τις υπόλοιπες 12 μεταβλητές που δηλώνουν ασθένεια από το δείγμα. Το χαρακτηριστικό “disease BPPV-2” είναι επίσης μία μεταβλητή δίτιμη με 1 να δηλώνει ότι ο εξεταζόμενος έχει την ασθένεια και 0 ότι ο εξεταζόμενος δεν έχει την ασθένεια BPPV-2.



Εικόνα 18. Γράφημα εξόδου BPPV-2 πριν την προεπεξεργασία

Η έξοδος του σετ δεδομένων- BPPV-2 παίρνει 2 διαφορετικές τιμές (βλ.Εικόνα 18).Εφαρμόστηκε σε αυτό το σετ δεδομένων ο K-NN ταξινομητής με default παραμέτρους.

Παρουσιάστηκε ξανά το πρόβλημα ανισοκατανομής στην ταξινόμηση των κλάσεων όπως φαίνεται από τα ποσοστά των κατηγοριών TP και FP. Επομένως εφαρμόστηκε όπως παραπάνω η τεχνική εξισορρόπησης SpreadSubsample κάνοντας undersampling –υπό δειγματοληψία (διαγραφή δειγμάτων από την κατηγορία πλειοψηφίας).

Επιλέχθηκαν μόνο 156 χαρακτηριστικά από τη μεταβλητή με τιμή 0 και το γράφημα μετά την εξισορρόπηση έχει ως εξής:



Εικόνα 19. Γράφημα εξόδου BPPV-2 μετά την εξισορρόπηση

**2ο Βήμα:** Επιλέγουμε στο πάνελ την καρτέλα «Classify» όπου περιλαμβάνονται ποικίλες τεχνικές ταξινόμησης με διαφορετικές δυνατότητες για το χρήστη ή κάθε μία. Θα εφαρμόσουμε τεχνικές στις οποίες η κατηγοριοποίηση των δεδομένων υλοποιείται μέσω διαφορετικών αλγορίθμων ταξινόμησης.

Ακολουθούμε την παρακάτω διαδικασία:

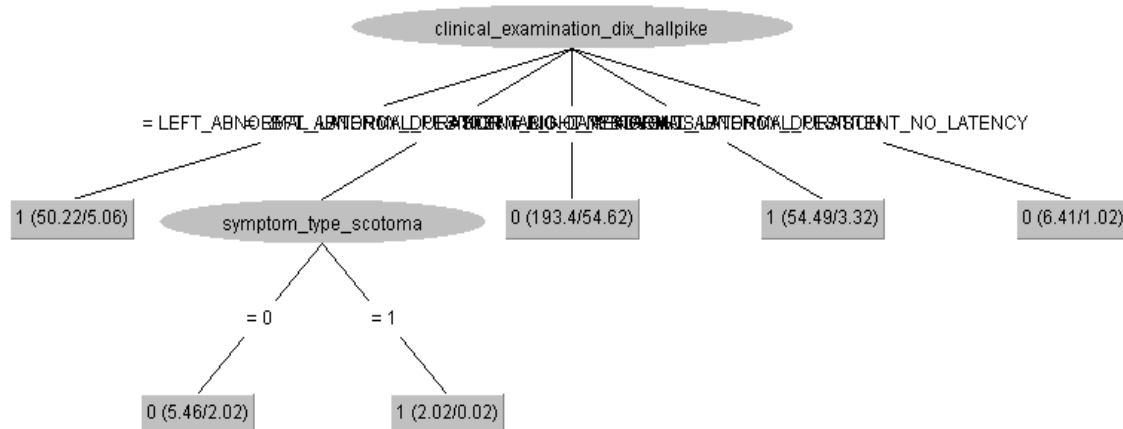
**Choose=>Weka=>Classifiers**

Οι αλγόριθμοι που χρησιμοποιήθηκαν είναι οι εξής:

- ❖ Ταξινόμηση μέσω του αλγορίθμου του πλησιέστερου γείτονα (K-NN)
- ❖ Ταξινόμηση μέσω του αλγορίθμου J-48

- ❖ Ταξινόμηση μέσω του αλγορίθμου Naive Bayes
- ❖ Ταξινόμηση μέσω του αλγορίθμου Random Forest

Για τον έλεγχο απόδοσης των παραπάνω αλγορίθμων χρησιμοποιήθηκε ξανά η τεχνική 10 fold cross validation (10-πλη διασταυρωτική). Στην συνέχεια εξάγονται τα στατιστικά στοιχεία για την απόδοση των αλγορίθμων σε επιτυχία πρόβλεψης νέων τιμών δεδομένων αλλά και το Δέντρο Απόφασης του J-48.



**Εικόνα 20. Δέντρο Απόφασης για την ασθένεια BPPV-2**

Ο καλοήθης παροξυσμικός ίλιγγος θέσεως διαγιγνώσκεται με την προσεκτική παρατήρηση ενός χαρακτηριστικού τύπου νυσταγμού κατά την διαδικασία της δοκιμασίας Dix-Hallpike. Υπάρχει ένας χαρακτηριστικός λανθάνον χρόνος, ένα έως δύο δευτερόλεπτα, πριν την έναρξη του νυσταγμού και του ίλιγγου. Σύμφωνα με το Δέντρο Απόφασης η ασθένεια BPPV τύπου 2 εμφανίζεται αν ο ασθενής κατά την εξέταση dix-hallpike εμφανίζει αριστερά μη φυσιολογική διάρκεια καθυστέρησης ή δεξιά μη φυσιολογική διάρκεια καθυστέρησης ή αριστερά ασυνήθιστη επιμονή καθυστέρησης ή δεξιά ασυνήθιστη επιμονή καθυστέρησης. Αν κατά την εξέταση Dix -Hallpike ο ασθενής δεν εμφανίσει νυσταγμό αλλά έχει συμπτώματα σκοτοδίνης τότε πάσχει από BPPV-2.

| Μέτρα Ακρίβειας BPPV-2 |                          |                                           |                                  |         |         |           |             |           |
|------------------------|--------------------------|-------------------------------------------|----------------------------------|---------|---------|-----------|-------------|-----------|
| classifier             |                          | Correctly Classified Instances - Accuracy | Incorrectly Classified Instances | TP Rate | FP Rate | Precision | Specificity | F-Measure |
| 1-nn                   | before feature selection | 75,64%                                    | 24,36%                           | 0,76    | 0,24    | 0,76      | 0,73        | 0,76      |
| 4-nn                   | before feature selection | 83,97%                                    | 16,03%                           | 0,84    | 0,16    | 0,84      | 0,76        | 0,84      |
| J48                    | before feature selection | 75,64%                                    | 24,36%                           | 0,76    | 0,24    | 0,78      | 0,86        | 0,75      |
| Naïve Bayes            | before feature selection | 83,97%                                    | 16,03%                           | 0,84    | 0,16    | 0,84      | 0,84        | 0,84      |
| Random Forest          | before feature selection | 85,26%                                    | 14,74%                           | 0,85    | 0,15    | 0,87      | 0,95        | 0,85      |

**Πίνακας 2. Συγκριτικά αποτελέσματα μεταξύ των αλγορίθμων ταξινόμησης για την ασθένεια BPPV-2**

Από τα αποτελέσματα τα οποία προέκυψαν, τα οποία παρουσιάζονται συγκεντρωτικά στον παραπάνω πίνακα, διαπιστώνουμε ότι:

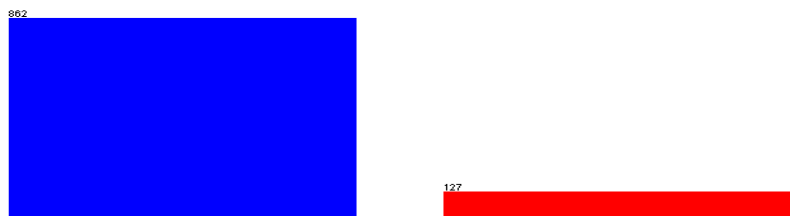
1. Μεταξύ των αλγορίθμων του πλησιέστερου γείτονα καλύτερη απόδοση είχε ο 4<sup>ος</sup> πλησιέστερος γείτονας, του οποίου τα αποτελέσματα ήταν τα πιο υψηλά.
2. Χειρότερη απόδοση διακρίνουμε στον αλγόριθμο J-48 του οποίου τα αποτελέσματα ταξινομήθηκαν λιγότερο από τα αποτελέσματα των υπόλοιπων αλγορίθμων.
3. Ο αλγόριθμος με την καλύτερη απόδοση και μέγιστη προβλεπτική ικανότητα είναι ο αλγόριθμος Random Forest για τον οποίο παρατηρείται το μεγαλύτερο ποσοστό ακρίβειας, μεγαλύτερο ποσοστό σωστά ταξινομημένων υποδειγμάτων- 85,26%, το μεγαλύτερο ποσοστό των θετικών υποδειγμάτων που έχουν αναγνωριστεί σωστά ως θετικά (μεγαλύτερη ευαισθησία) –0,77, το μεγαλύτερο ποσοστό των αρνητικών υποδειγμάτων που έχουν αναγνωριστεί ως αρνητικά (μεγαλύτερη ειδικότητα)-0,95, υψηλότερες τιμές για όλα τα υπό εξέταση μέτρα ακρίβειας (precision, recall, f-measure).

## Αποτελέσματα αλγορίθμων για την ασθένεια Meniere χωρίς επιλογή χαρακτηριστικών.

### Μεταβλητή στόχος: Meniere.

Ακολουθώντας την ίδια μεθοδολογία επιλέγουμε το χαρακτηριστικό disease Meniere ως αυτό το οποίο δείχνει σε ποια κλάση ανήκει κάθε φορά το υπόδειγμα και αφαιρούμε όπως είναι απαραίτητο τις υπόλοιπες 12 μεταβλητές που δηλώνουν ασθένεια από το δείγμα. Το χαρακτηριστικό “disease Menieres” είναι επίσης μία μεταβλητή δίτιμη με 1 να δηλώνει ότι ο εξεταζόμενος έχει την ασθένεια και 0 ότι ο εξεταζόμενος δεν έχει την ασθένεια Meniere

| Selected attribute     |       |                |               |
|------------------------|-------|----------------|---------------|
| Name: disease_menieres |       | Distinct: 2    | Type: Nominal |
| Missing: 0 (0%)        |       | Unique: 0 (0%) |               |
| No.                    | Label | Count          | Weight        |
| 1                      | 0     | 862            | 862.0         |
| 2                      | 1     | 127            | 127.0         |



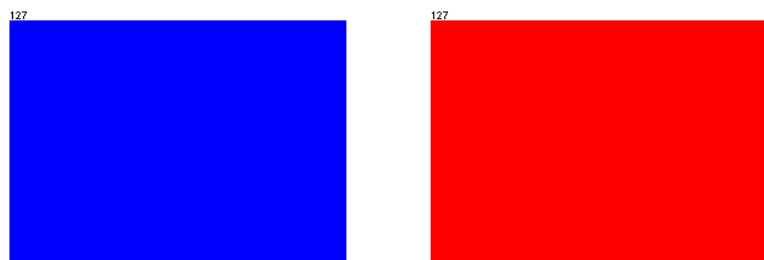
**Εικόνα 21 . Γράφημα εξόδου Meniere πριν την προεπεξεργασία**

Η έξοδος του σετ δεδομένων- Meniere παίρνει 2 διαφορετικές τιμές (βλ.Εικόνα 21). Εφαρμόστηκε σε αυτό το σετ δεδομένων ο K-NN ταξινομητής με default παραμέτρους.

Παρουσιάστηκε ξανά το πρόβλημα ανισοκατανομής στην ταξινόμηση των κλάσεων όπως φαίνεται από τα ποσοστά των κατηγοριών TP και FP. Επομένως εφαρμόστηκε όπως παραπάνω η τεχνική εξισορρόπησης SpreadSubsample κάνοντας undersampling –υπό δειγματοληψία (διαγραφή δειγμάτων από την κατηγορία πλειοψηφίας).

Επιλέχθηκαν μόνο 127 χαρακτηριστικά από τη μεταβλητή με τιμή 0 και το γράφημα μετά την εξισορρόπηση έχει ως εξής:

| Selected attribute     |       |                |               |
|------------------------|-------|----------------|---------------|
| Name: disease_menieres |       | Distinct: 2    | Type: Nominal |
| Missing: 0 (0%)        |       | Unique: 0 (0%) |               |
| No.                    | Label | Count          | Weight        |
| 1                      | 0     | 127            | 127.0         |
| 2                      | 1     | 127            | 127.0         |

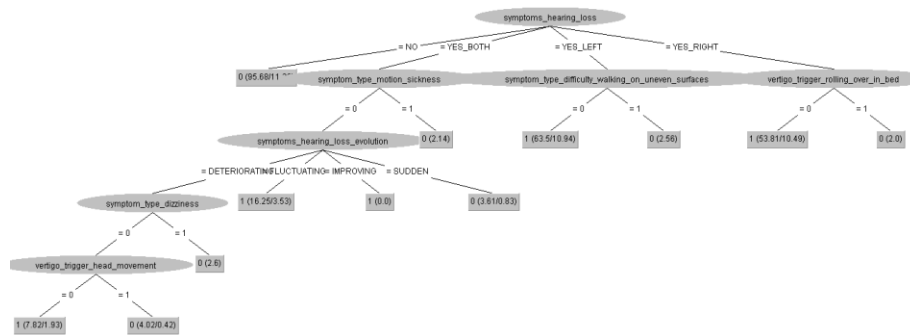


**Εικόνα 22. Γράφημα εξόδου Meniere μετά την εξισορρόπηση**

Ακολουθούμε την παρακάτω διαδικασία:

Οι αλγόριθμοι που χρησιμοποιήθηκαν είναι οι εξής:

- Για τον έλεγχο απόδοσης των παραπάνω αλγορίθμων χρησιμοποιήθηκε ξανά η τεχνική 10 fold cross validation (10-πλη διασταυρωτική) Στην συνέχεια εξάγονται τα στατιστικά στοιχεία για την απόδοση των αλγορίθμων σε επιτυχία πρόβλεψης νέων τιμών δεδομένων αλλά και το Δέντρο Απόφασης του J-48.



Η ασθένεια του Meniere εμφανίζεται εάν ο ασθενής έχει απώλεια ακοής στο δεξί αυτί και παράλληλα παθαίνει ίλιγγο καθώς περιστρέφεται στο κρεβάτι, αλλιώς αν έχει απώλεια ακοής στο αριστερό αυτί και παράλληλα αντιμετωπίζει δυσκολίες να περπατήσει σε ανώμαλες επιφάνειες ή αλλιώς αν αντιμετωπίζει απώλεια ακοής και από τα δύο αυτιά και παράλληλα έχει συμπτώματα ναυτίας, συμπτώματα τύπου ζάλης και ίλιγγο καθώς η κεφαλή βρίσκεται σε κίνηση.

| Μέτρα Ακρίβειας Meniere Disease |                          |                                           |                                  |         |         |           |             |           |
|---------------------------------|--------------------------|-------------------------------------------|----------------------------------|---------|---------|-----------|-------------|-----------|
| classifier                      |                          | Correctly Classified Instances - Accuracy | Incorrectly Classified Instances | TP Rate | FP Rate | Precision | Specificity | F-Measure |
| 1-nn                            | before feature selection | 77,56%                                    | 22,44%                           | 0,78    | 0,22    | 0,80      | 0,71        | 0,77      |
| 8-nn                            | before feature selection | 81,89%                                    | 18,11%                           | 0,82    | 0,18    | 0,84      | 0,75        | 0,82      |
| J48                             | before feature selection | 80,71%                                    | 19,29%                           | 0,81    | 0,19    | 0,81      | 0,77        | 0,81      |
| Naïve Bayes                     | before feature selection | 87,01%                                    | 12,99%                           | 0,87    | 0,13    | 0,87      | 0,87        | 0,87      |
| Random Forest                   | before feature selection | 87,01%                                    | 12,99%                           | 0,87    | 0,13    | 0,87      | 0,89        | 0,87      |

**Πίνακας 3. Συγκριτικά αποτελέσματα μεταξύ των αλγορίθμων ταξινόμησης για την ασθένεια Meniere**

Από τα αποτελέσματα τα οποία προέκυψαν, τα οποία παρουσιάζονται συγκεντρωτικά στον παραπάνω πίνακα, διαπιστώνουμε ότι:

1. Μεταξύ των αλγορίθμων του πλησιέστερου γείτονα καλύτερη απόδοση είχε ο 8<sup>ος</sup> πλησιέστερος γείτονας , του οποίου τα αποτελέσματα ήταν τα πιο υψηλά
2. Χειρότερη απόδοση διακρίνουμε στον αλγόριθμο J-48 του οποίου τα αποτελέσματα ταξινομήθηκαν λιγότερο από τα αποτελέσματα των υπόλοιπων αλγορίθμων
3. Οι αλγόριθμοι με την καλύτερη απόδοση και μέγιστη προβλεπτική ικανότητα είναι ο αλγόριθμος Random Forest και ο αλγόριθμος Naive Bayes για τους οποίους παρατηρείται το μεγαλύτερο ποσοστό ακρίβειας, μεγαλύτερο ποσοστό σωστά ταξινομημένων υποδειγμάτων- 87,01%, το μεγαλύτερο ποσοστό των θετικών υποδειγμάτων που έχουν αναγνωριστεί σωστά ως θετικά (μεγαλύτερη ευαισθησία) –0,87 ,υψηλότερες τιμές για όλα τα υπό εξέταση μέτρα ακρίβειας (precision, recall, f-measure).
4. Μεταξύ των Random Forest και Naive Bayes καλύτερος ταξινομητής είναι ο Random Forest αφού έχει την μεγαλύτερη ειδικότητα -0,89 (ποσοστό των αρνητικών υποδειγμάτων που έχουν αναγνωριστεί ως αρνητικά).

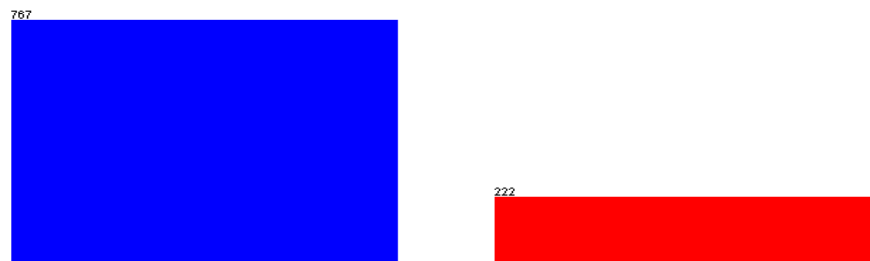
Αποτελέσματα αλγορίθμων για την ασθένεια Vestibular Migraine χωρίς την επιλογή χαρακτηριστικών.

**Μεταβλητή στόχος: Vestibular Migraine.**

Ακολουθώντας την ίδια μεθοδολογία επιλέγουμε το χαρακτηριστικό disease Vestibular Migraine ως αυτό το οποίο δείχνει σε ποια κλάση ανήκει κάθε φορά το υπόδειγμα και αφαιρούμε όπως είναι απαραίτητο τις υπόλοιπες 12 μεταβλητές που δηλώνουν ασθένεια από το δείγμα. Το χαρακτηριστικό “disease Vestibular Migraine” είναι επίσης μία μεταβλητή δίτιμη με 1 να δηλώνει ότι ο εξεταζόμενος έχει την ασθένεια και 0 ότι ο εξεταζόμενος δεν έχει την ασθένεια Vestibular Migraine.

| Selected attribute                |       |                |        |
|-----------------------------------|-------|----------------|--------|
| Name: disease_vestibular_migraine |       | Type: Nominal  |        |
| Missing: 0 (0%)                   |       | Distinct: 2    |        |
|                                   |       | Unique: 0 (0%) |        |
| No.                               | Label | Count          | Weight |
| 1                                 | 0     | 767            | 767.0  |
| 2                                 | 1     | 222            | 222.0  |

Class: disease\_vestibular\_migraine (Nom) Visualize All

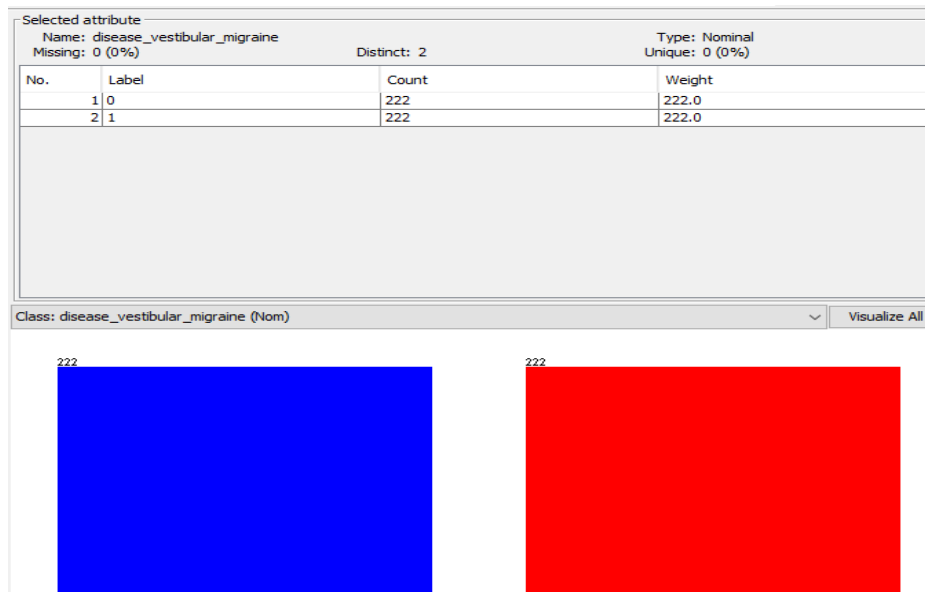


**Εικόνα 24 . Γράφημα εξόδου Vestibular Migraine πριν την προεπεξεργασία**

Η έξοδος του σετ δεδομένων Vestibular Migraine παίρνει 2 διαφορετικές τιμές όπως φαίνεται στην εικόνα . Εφαρμόστηκε σε αυτό το σετ δεδομένων ο K-NN ταξινομητής με “default” παραμέτρους.

Παρουσιάστηκε ξανά το πρόβλημα ανισοκατανομής στην ταξινόμηση των κλάσεων όπως φαίνεται από τα τα ποσοστά των κατηγοριών TP και FP (βλ.εικόνα 24). Επομένως εφαρμόστηκε όπως παραπάνω η τεχνική εξισορρόπησης SpreadSubsample κάνοντας undersampling –υπό δειγματοληψία (διαγραφή δειγμάτων από την κατηγορία πλειοψηφίας).

Επιλέχθηκαν μόνο 222 χαρακτηριστικά από τη μεταβλητή με τιμή 0 και το γράφημα μετά την εξισορρόπηση έχει ως εξής:



**Εικόνα 25. Γράφημα εξόδου Vestibular Migraine μετά την εξισορρόπηση**

**2ο Βήμα:** Επιλέγουμε στο πάνελ την καρτέλα «Classify» όπου περιλαμβάνονται ποικίλες τεχνικές ταξινόμησης με διαφορετικές δυνατότητες για το χρήστη ή κάθε μία. Θα εφαρμόσουμε τεχνικές στις οποίες η κατηγοριοποίηση των δεδομένων υλοποιείται μέσω διαφορετικών αλγορίθμων ταξινόμησης.

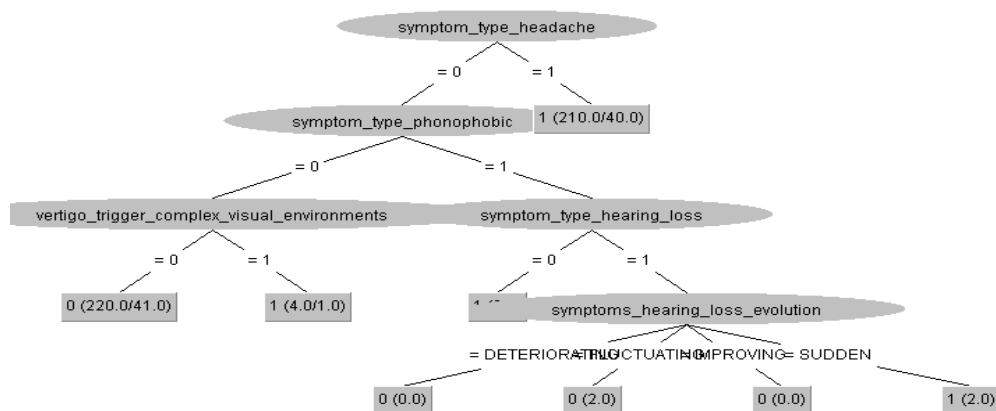
Ακολουθούμε την παρακάτω διαδικασία:

**Choose=>Weka=>Classifiers**

Οι αλγόριθμοι που χρησιμοποιήθηκαν είναι οι εξής:

- ❖ Ταξινόμηση μέσω του αλγορίθμου του πλησιέστερου γείτονα (K-NN)
- ❖ Ταξινόμηση μέσω του αλγορίθμου J-48
- ❖ Ταξινόμηση μέσω του αλγορίθμου Naive Bayes
- ❖ Ταξινόμηση μέσω του αλγορίθμου Random Forest

Για τον έλεγχο απόδοσης των παραπάνω αλγορίθμων χρησιμοποιήθηκε ξανά η τεχνική 10 fold cross validation (10-πλη διασταυρωτική) Στην συνέχεια εξάγονται τα στατιστικά στοιχεία για την απόδοση των αλγορίθμων σε επιτυχία πρόβλεψης νέων τιμών δεδομένων αλλά και το Δέντρο Απόφασης του J-48.



**Εικόνα 2 . Δέντρο Απόφασης για την ασθένεια Vestibular Migraine**

Σύμφωνα με το Δέντρο Απόφασης κάποιος μπορεί να πάσχει από Vestibular Migraine αν έχει συμπτώματα πονοκεφάλου .Διαφορετικά αν δεν έχει συμπτώματα πονοκεφάλου αλλά έχει συμπτώματα τύπου φωνοφοβικού, συμπτώματα τύπου ξαφνικής απώλειας ακοής και συμπτώματα εξέλιξης της απώλειας ακοής τότε μπορεί να πάσχει από την ασθένεια ή εναλλακτικά αν δεν έχει συμπτώματα πονοκεφάλου και δεν έχει συμπτώματα τύπου φωνοφοβικού αλλά έχει ίλιγγο που προκαλείται από σύνθετα οπτικά περιβάλλοντα τότε έχει την ασθένεια.

| Μέτρα Ακρίβειας Vestibular Migraine |                          |                                           |                                  |         |         |           |             |           |
|-------------------------------------|--------------------------|-------------------------------------------|----------------------------------|---------|---------|-----------|-------------|-----------|
| classifier                          |                          | Correctly Classified Instances - Accuracy | Incorrectly Classified Instances | TP Rate | FP Rate | Precision | Specificity | F-Measure |
| 1-nn                                | before feature selection | 74,10%                                    | 25,90%                           | 0,74    | 0,26    | 0,74      | 0,73        | 0,74      |
| 6-nn                                | before feature selection | 81,08                                     | 18,92                            | 0,81    | 0,19    | 0,81      | 0,66        | 0,81      |
| J48                                 | before feature selection | 79,05%                                    | 20,95%                           | 0,79    | 0,21    | 0,79      | 0,80        | 0,79      |
| Naïve Bayes                         | before feature selection | 81,76%                                    | 18,24%                           | 0,82    | 0,18    | 0,82      | 0,79        | 0,82      |
| Random Forest                       | before feature selection | 82,43%                                    | 17,57%                           | 0,82    | 0,18    | 0,83      | 0,84        | 0,82      |

**Πίνακας 4. Συγκριτικά αποτελέσματα μεταξύ των αλγορίθμων ταξινόμησης για την ασθένεια Vestibular Migraine**

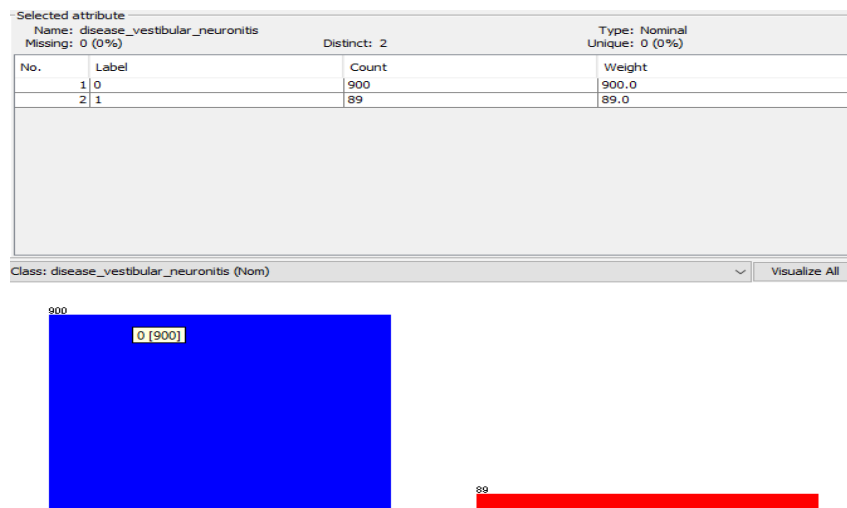
Από τα αποτελέσματα τα οποία προέκυψαν, τα οποία παρουσιάζονται συγκεντρωτικά στον παραπάνω πίνακα, διαπιστώνουμε ότι :

1. Μεταξύ των αλγορίθμων του πλησιέστερου γείτονα καλύτερη απόδοση είχε ο 6<sup>ος</sup> πλησιέστερος γείτονας , του οποίου τα αποτελέσματα ήταν τα πιο υψηλά.
2. Χειρότερη απόδοση διακρίνουμε στον αλγόριθμο J-48 του οποίου τα αποτελέσματα ταξινομήθηκαν λιγότερο από τα αποτελέσματα των υπόλοιπων αλγορίθμων.
3. Ο αλγόριθμος με την καλύτερη απόδοση και μέγιστη προβλεπτική ικανότητα είναι ο αλγόριθμος Random Forest και για τον οποίο παρατηρείται το μεγαλύτερο ποσοστό ακρίβειας, μεγαλύτερο ποσοστό σωστά ταξινομημένων υποδειγμάτων- 82,43%, το μεγαλύτερο ποσοστό των θετικών υποδειγμάτων που έχουν αναγνωριστεί σωστά ως θετικά (μεγαλύτερη ευαισθησία) -0,82 ,το μεγαλύτερο ποσοστό των αρνητικών υποδειγμάτων που έχουν αναγνωριστεί ως αρνητικά (μεγαλύτερη ειδικότητα)-0,84 και υψηλότερες τιμές για όλα τα υπό εξέταση μέτρα ακρίβειας (precision, recall, f-measure).

### Αποτελέσματα αλγορίθμων για την ασθένεια Vestibular Neuronitis χωρίς την επιλογή χαρακτηριστικών.

#### Μεταβλητή στόχος: Vestibular Neuronitis.

Ακολουθώντας την ίδια μεθοδολογία επιλέγουμε το χαρακτηριστικό disease Vestibular Neuronitis ως αυτό το οποίο δείχνει σε ποια κλάση ανήκει κάθε φορά το υπόδειγμα και αφαιρούμε όπως είναι απαραίτητο τις υπόλοιπες 12 μεταβλητές που δηλώνουν ασθένεια από το δείγμα. Το χαρακτηριστικό “disease Vestibular Neuronitis” είναι επίσης μία μεταβλητή δίτιμη με 1 να δηλώνει ότι ο εξεταζόμενος έχει την ασθένεια και 0 ότι ο εξεταζόμενος δεν έχει την ασθένεια Vestibular Neuronitis.

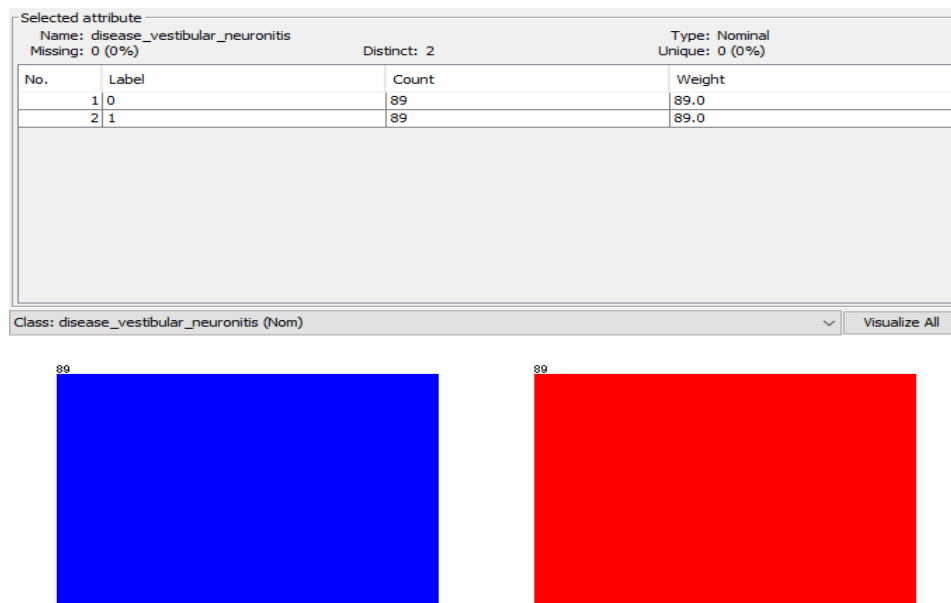


Εικόνα 27. Γράφημα εξόδου Vestibular Migraine πριν την προεπεξεργασία

Η έξοδος του σετ δεδομένων Vestibular Neuronitis παίρνει 2 διαφορετικές τιμές όπως φαίνεται στην εικόνα . Εφαρμόστηκε σε αυτό το σετ δεδομένων ο K-NN ταξινομητής με default παραμέτρους.

Παρουσιάστηκε ξανά το πρόβλημα ανισοκατανομής στην ταξινόμηση των κλάσεων όπως φαίνεται από τα ποσοστά των κατηγοριών TP και FP (βλ.εικόνα 27). Επομένως εφαρμόστηκε όπως παραπάνω η τεχνική εξισορρόπησης SpreadSubsample κάνοντας undersampling –υπό δειγματοληψία (διαγραφή δειγμάτων από την κατηγορία πλειοψηφίας).

Επιλέχθηκαν μόνο 89 χαρακτηριστικά από τη μεταβλητή με τιμή 0 και το γράφημα μετά την εξισορρόπηση έχει ως εξής:



*Εικόνα 28. Γράφημα εξόδου Vestibular Neuronitis μετά την εξισορρόπηση*

**2ο Βήμα:** Επιλέγουμε στο πάνελ την καρτέλα «Classify» όπου περιλαμβάνονται ποικίλες τεχνικές ταξινόμησης με διαφορετικές δυνατότητες για το χρήστη η κάθε μία. Θα εφαρμόσουμε τεχνικές στις οποίες η κατηγοριοποίηση των δεδομένων υλοποιείται μέσω διαφορετικών αλγορίθμων ταξινόμησης.

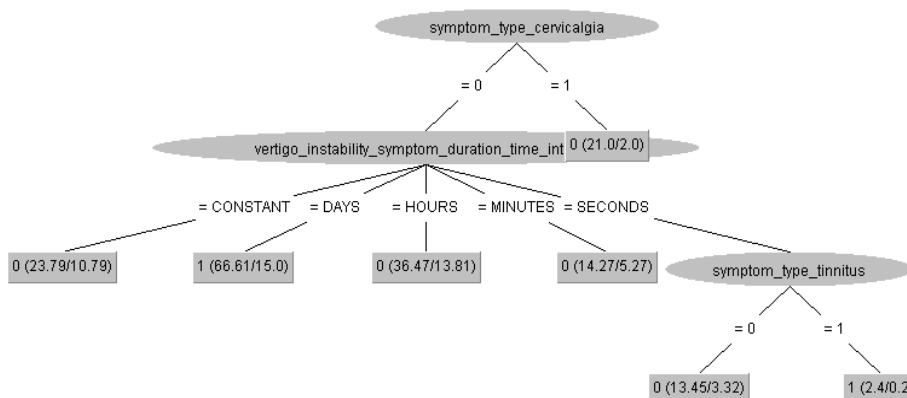
Ακολουθούμε την παρακάτω διαδικασία:

**Choose=>Weka=>Classifier**

Οι αλγόριθμοι που χρησιμοποιήθηκαν είναι οι εξής:

- ❖ Ταξινόμηση μέσω του αλγορίθμου του πλησιέστερου γείτονα (K-NN).
- ❖ Ταξινόμηση μέσω του αλγορίθμου J-48.
- ❖ Ταξινόμηση μέσω του αλγορίθμου Naive Bayes.
- ❖ Ταξινόμηση μέσω του αλγορίθμου Random Forest

Για τον έλεγχο απόδοσης των παραπάνω αλγορίθμων χρησιμοποιήθηκε ξανά η τεχνική 10 fold cross validation (10-πλη διασταυρωτική). Στην συνέχεια εξάγονται τα στατιστικά στοιχεία για την απόδοση των αλγορίθμων σε επιτυχία πρόβλεψης νέων τιμών δεδομένων αλλά και το Δέντρο Απόφασης του J-48.



**Εικόνα 29 . Δέντρο Απόφασης για την ασθένεια Vestibular Neuronitis**

Σύμφωνα με το Δέντρο Απόφασης κάποιος μπορεί να πάσχει από Vestibular Neuronitis αν έχει συμπτώματα πόνου στον αυχένα. Σε διαφορετική περίπτωση πάσχει από την ασθένεια αν δεν έχει πόνο στον αυχένα αλλά έχει συμπτώματα αστάθειας ιλίγγου συνέχεια, ημέρες ώρες ή λεπτά ή αλλιώς αν η διάρκεια της αστάθειας ιλίγγου είναι δευτερόλεπτα αλλά έχει συμπτώματα εμβοών

| Μέτρα Ακρίβειας Vestibular Neuronitis |                          |                                           |                                  |         |         |           |             |           |
|---------------------------------------|--------------------------|-------------------------------------------|----------------------------------|---------|---------|-----------|-------------|-----------|
| classifier                            |                          | Correctly Classified Instances - Accuracy | Incorrectly Classified Instances | TP Rate | FP Rate | Precision | Specificity | F-Measure |
| 1-nn                                  | before feature selection | 57,30%                                    | 42,70%                           | 0,57    | 0,43    | 0,58      | 0,55        | 0,56      |
| 3-nn                                  | before feature selection | 68,54%                                    | 31,46%                           | 0,69    | 0,32    | 0,72      | 0,63        | 0,67      |
| J48                                   | before feature selection | 60,67%                                    | 39,33%                           | 0,61    | 0,39    | 0,62      | 0,58        | 0,60      |
| Naïve Bayes                           | before feature selection | 74,16%                                    | 25,84%                           | 0,74    | 0,26    | 0,75      | 0,71        | 0,74      |
| Random Forest                         | before feature selection | 73,60%                                    | 26,40%                           | 0,74    | 0,26    | 0,74      | 0,70        | 0,73      |

**Πίνακας 5. Συγκριτικά αποτελέσματα μεταξύ των αλγορίθμων ταξινόμησης για την ασθένεια Vestibular Neuronitis**

Από τα αποτελέσματα τα οποία προέκυψαν, τα οποία παρουσιάζονται συγκεντρωτικά στον παραπάνω πίνακα, διαπιστώνουμε ότι :

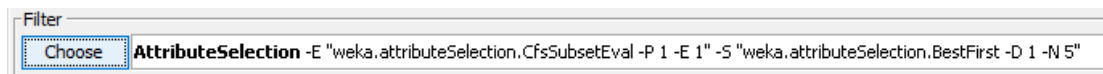
1. Μεταξύ των αλγορίθμων του πλησιέστερου γείτονα καλύτερη απόδοση είχε ο 3<sup>ος</sup> πλησιέστερος γείτονας , του οποίου τα αποτελέσματα ήταν τα πιο υψηλά.
2. Χειρότερη απόδοση διακρίνουμε στον αλγόριθμο J-48 του οποίου τα αποτελέσματα ταξινομήθηκαν λιγότερο από τα αποτελέσματα των υπόλοιπων αλγορίθμων.
3. Ο αλγόριθμος με την καλύτερη απόδοση και μέγιστη προβλεπτική ικανότητα είναι ο αλγόριθμος Naive Bayes και για τον οποίο παρατηρείται το μεγαλύτερο ποσοστό ακρίβειας, μεγαλύτερο ποσοστό σωστά ταξινομημένων υποδειγμάτων- 74,15%, το μεγαλύτερο ποσοστό των θετικών υποδειγμάτων που έχουν αναγνωριστεί σωστά ως θετικά (μεγαλύτερη ευαισθησία) -0,74 ,το μεγαλύτερο ποσοστό των αρνητικών υποδειγμάτων που έχουν αναγνωριστεί ως αρνητικά (μεγαλύτερη ειδικότητα)-0,71 και υψηλότερες τιμές για όλα τα υπό εξέταση μέτρα ακρίβειας (precision, recall, f-measure).

### Επιλογή χαρακτηριστικών ασθενειών.

Έπειτα από την εισαγωγή των δεδομένων για κάθε ασθένεια στους αλγορίθμους ταξινόμησης και έχοντας εικόνα της απόδοσης των μοντέλων που δημιουργήθηκαν ,εξετάστηκε η επιλογή χαρακτηριστικών ανά ασθένεια .Στόχος αυτής της ενέργειας είναι να βρεθούν τα πραγματικά χαρακτηριστικά από τα οποία εξαρτάται η εμφάνιση κάθε ασθένειας που προκαλεί ίλιγγο και η δημιουργία αποδοτικότερων μοντέλων πρόβλεψης πρόκλησης ασθενειών για μελλοντική χρήση τους σε Κλινικά Συστήματα Υποστήριξης Αποφάσεων .

### Αποτελέσματα αλγορίθμων για την ασθένεια BPPV-1 (Παροξυσμικός ίλιγγος θέσεως τύπου 1) μετά την επιλογή χαρακτηριστικών.

Στο λογισμικό Weka για την επιλογή χαρακτηριστικών επιλέγουμε το tab με τίτλο “select attributes” το οποίο εφαρμόζει διάφορους αλγόριθμους επιλογής χαρακτηριστικών, και έχει ως στόχο τη μείωση της διάστασης της βάσης δεδομένων. Επιλέγεται η μέθοδος εκτίμησης χαρακτηριστικών απ’ το πεδίο “attribute evaluator”, καθώς και η μέθοδο αναζήτησης, από το πεδίο “search method” για την επιλογή των χαρακτηριστικών.

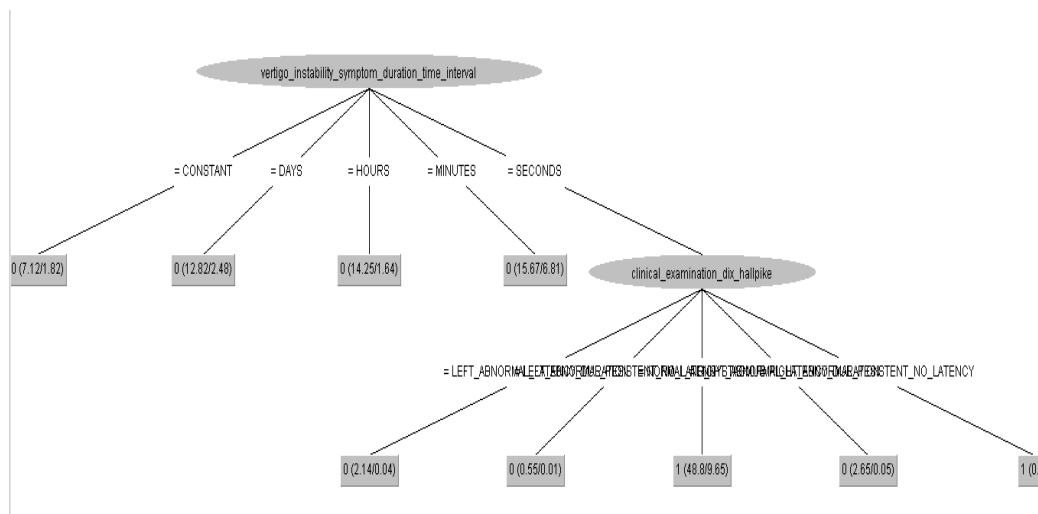


Εικόνα 30. Φίλτρο επιλογής χαρακτηριστικών –AttributeSelection.

Χρησιμοποιήθηκε η μέθοδος CfsSubsetEval”, η οποία αξιολογεί την προβλεπτική ικανότητα κάθε χαρακτηριστικού ξεχωριστά και το βαθμό πλεονασμού μεταξύ αυτών, προτιμώντας τα σύνολα

χαρακτηριστικών που έχουν υψηλή συσχέτιση με την τάξη, αλλά έχουν χαμηλή ενδοσυσχέτιση (intercorrelation).

Μετά την εφαρμογή του φίλτρου AttributeSelection επιλέχθηκαν τα κατάλληλα χαρακτηριστικά για την ασθένεια BPPV-1 (βλ. Πίνακα 11). Μετά την επιλογή χαρακτηριστικών για την ασθένεια BPPV-1 εφαρμόστηκαν τεχνικές κατηγοριοποίησης δεδομένων και η αξιολόγηση της απόδοσης των μοντέλων θα γίνει όπως και στην προηγούμενη φάση με χρήση της 10-πλής διασταυρωτικής. Στην συνέχεια εξάγονται τα στατιστικά στοιχεία για την απόδοση των αλγορίθμων σε επιτυχία πρόβλεψης νέων τιμών δεδομένων αλλά και το Δέντρο Απόφασης του J-48.



**Εικόνα 31. Δέντρο Απόφασης για την ασθένεια BPPV-1 μετά την επιλογή χαρακτηριστικών**

Μετά την επιλογή χαρακτηριστικών και σύμφωνα με το Δέντρο Απόφασης ασθενής πάσχει από BPPV-1 όταν έχει συμπτώματα αστάθειας ιλίγγου συνέχεια ,ημέρες ,ώρες ,λεπτά ή δευτερόλεπτα αλλά η τα αποτελέσματα της κλινικής εξέτασης Dix-hallpike εμφανίζει αριστερά μη φυσιολογική διάρκεια καθυστέρησης ή δεξιά μη φυσιολογική διάρκεια καθυστέρησης ή αριστερά ασυνήθιστη επιμονή καθυστέρησης ή δεξιά ασυνήθιστη επιμονή καθυστέρησης.

| Μέτρα Ακρίβειας BPPV-1 |                          |                                           |                                  |         |         |           |             |           |
|------------------------|--------------------------|-------------------------------------------|----------------------------------|---------|---------|-----------|-------------|-----------|
| classifier             |                          | Correctly Classified Instances - Accuracy | Incorrectly Classified Instances | TP Rate | FP Rate | Precision | Specificity | F-Measure |
| 1-nn                   | before feature selection | 74,04%                                    | 25,96%                           | 0,74    | 0,26    | 0,75      | 0,71        | 0,74      |
|                        | after feature selection  | 75%                                       | 25%                              | 0,75    | 0,25    | 0,76      | 0,71        | 0,75      |
| 8-nn                   | before feature selection | 75,96%                                    | 24,04%                           | 0,76    | 0,24    | 0,77      | 0,72        | 0,76      |
|                        | after feature selection  | 64,42%                                    | 35,58%                           | 0,64    | 0,36    | 0,72      | 0,59        | 0,61      |
| J48                    | before feature selection | 71,15%                                    | 28,85%                           | 0,71    | 0,29    | 0,72      | 0,67        | 0,71      |
|                        | after feature selection  | 74,04%                                    | 25,96%                           | 0,74    | 0,26    | 0,75      | 0,70        | 0,74      |
| Naïve Bayes            | before feature selection | 75,00%                                    | 25,00%                           | 0,75    | 0,75    | 0,75      | 0,76        | 0,75      |
|                        | after feature selection  | 82,69%                                    | 17,31%                           | 0,83    | 0,17    | 0,83      | 0,83        | 0,83      |
| Random Forest          | before feature selection | 76,92%                                    | 23,08%                           | 0,77    | 0,23    | 0,77      | 0,75        | 0,77      |
|                        | after feature selection  | 79,81%                                    | 20,19%                           | 0,80    | 0,20    | 0,80      | 0,79        | 0,80      |

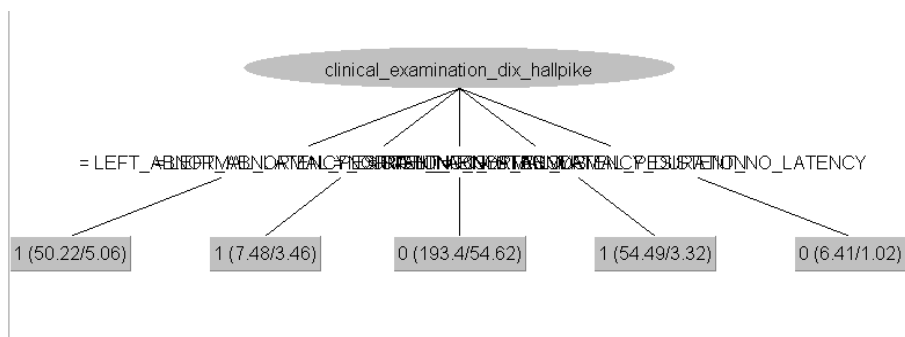
*Πίνακας 6 .Συγκριτικά αποτελέσματα για τα μέτρα ακριβείας της ασθένειας BPPV-1 μετά την επιλογή χαρακτηριστικών*

Μετά την επιλογή χαρακτηριστικών παρατηρείται ότι καλύτερη απόδοση εμφανίζει ο αλγόριθμος Naive Bayes για τον οποίο παρατηρείται το μεγαλύτερο ποσοστό ακρίβειας, 82,69%, το μεγαλύτερο ποσοστό των θετικών υποδειγμάτων που έχουν αναγνωριστεί σωστά ως θετικά (μεγαλύτερη ευαισθησία) –0,83 ,το μεγαλύτερο ποσοστό των αρνητικών υποδειγμάτων που έχουν αναγνωριστεί ως αρνητικά (μεγαλύτερη ειδικότητα)–0,83 και υψηλότερες τιμές για όλα τα υπό εξέταση μέτρα ακρίβειας (precision, recall, f-measure).

### *Αποτελέσματα αλγορίθμων για την ασθένεια BPPV-2 (Παροξυσμικός ίλιγγος θέσεως τύπου 2) μετά την επιλογή χαρακτηριστικών.*

Ενεργώντας όπως παραπάνω για την επιλογή χαρακτηριστικών για την ασθένεια BPPV-2 επιλέγουμε “select attributes” με μέθοδο “CfsSubsetEval”.Μετά την εφαρμογή του φίλτρου “AttributeSelection” επιλέχθηκαν τα κατάλληλα χαρακτηριστικά για την ασθένεια BPPV-2 (βλ. Πίνακα 11 ).Μετά την επιλογή χαρακτηριστικών για την ασθένεια BPPV-2 εφαρμόστηκαν τεχνικές κατηγοριοποίησης δεδομένων και η αξιολόγηση της απόδοσης των μοντέλων θα γίνει όπως και προηγουμένως με χρήση της 10-πλής

διασταυρωτικής. Στην συνέχεια εξάγονται τα στατιστικά στοιχεία για την απόδοση των αλγορίθμων σε επιτυχία πρόβλεψης νέων τιμών δεδομένων αλλά και το Δέντρο Απόφασης του J-48.



Εικόνα 32. Δέντρο Απόφασης για την ασθένεια BPPV-2 μετά την επιλογή χαρακτηριστικών

Μετά την επιλογή χαρακτηριστικών βλέπουμε ότι η πιθανότητα εμφάνισης ασθένειας BPPV-2 εξαρτάται μόνο από τα αποτελέσματα της κλινικής εξέτασης Dix-Hallpike. Ο ασθενής πάσχει από την ασθένεια αν εμφανίζει αριστερά μη φυσιολογική διάρκεια καθυστέρησης ή δεξιά μη φυσιολογική διάρκεια καθυστέρησης ή αριστερά ασυνήθιστη επιμονή καθυστέρησης ή δεξιά ασυνήθιστη επιμονή καθυστέρησης ή νυσταγμό.

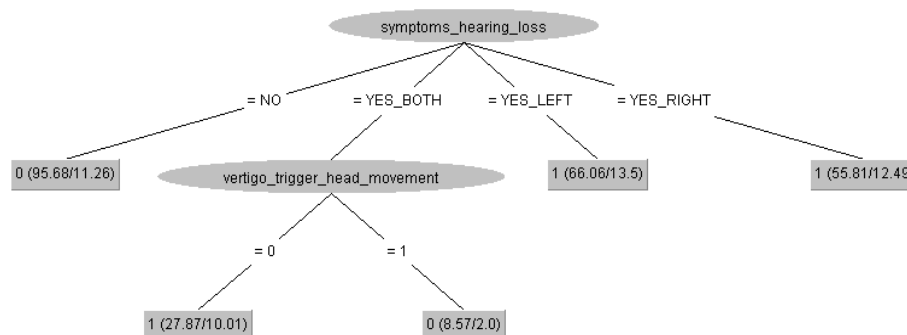
| Μέτρα Ακρίβειας BPPV-2 |                          |                                           |                                  |         |         |           |             |           |
|------------------------|--------------------------|-------------------------------------------|----------------------------------|---------|---------|-----------|-------------|-----------|
| classifier             |                          | Correctly Classified Instances - Accuracy | Incorrectly Classified Instances | TP Rate | FP Rate | Precision | Specificity | F-Measure |
| 1-nn                   | before feature selection | 75,64%                                    | 24,36%                           | 0,76    | 0,24    | 0,76      | 0,73        | 0,76      |
|                        | after feature selection  | 81,41%                                    | 18,59%                           | 0,81    | 0,19    | 0,82      | 0,84        | 0,81      |
| 4-nn                   | before feature selection | 83,97%                                    | 16,03%                           | 0,84    | 0,16    | 0,84      | 0,76        | 0,84      |
|                        | after feature selection  | 81,41%                                    | 18,59%                           | 0,81    | 0,19    | 0,82      | 0,84        | 0,81      |
| J48                    | before feature selection | 75,64%                                    | 24,36%                           | 0,76    | 0,24    | 0,78      | 0,86        | 0,75      |
|                        | after feature selection  | 77,24%                                    | 22,76%                           | 0,77    | 0,23    | 0,80      | 0,89        | 0,77      |
| Naïve Bayes            | before feature selection | 83,97%                                    | 16,03%                           | 0,84    | 0,16    | 0,84      | 0,84        | 0,84      |
|                        | after feature selection  | 81,73%                                    | 18,27%                           | 0,82    | 0,18    | 0,82      | 0,83        | 0,82      |
| Random Forest          | before feature selection | 85,26%                                    | 14,74%                           | 0,85    | 0,15    | 0,87      | 0,95        | 0,85      |
|                        | after feature selection  | 80,45%                                    | 19,55%                           | 0,80    | 0,20    | 0,81      | 0,82        | 0,80      |

### Πίνακας 7. Συγκριτικά αποτελέσματα για τα μέτρα ακριβείας της ασθένειας BPPV-2 μετά την επιλογή χαρακτηριστικών

Μετά την επιλογή χαρακτηριστικών καλύτερη απόδοση εμφανίζει ο αλγόριθμος Naive Bayes με ποσοστό ακριβείας 81,73% . Παρατηρείται όμως ότι ο Random Forest χωρίς επιλογή χαρακτηριστικών έφερε τα καλύτερα αποτελέσματα ,καθώς έχει υψηλότερες τιμές για όλα τα υπό εξέταση μέτρα ακρίβειας (precision, recall, f-measure),το μεγαλύτερο ποσοστό ακρίβειας, 85,26%, το μεγαλύτερο ποσοστό των θετικών υποδειγμάτων που έχουν αναγνωριστεί σωστά ως θετικά (μεγαλύτερη ευαισθησία) –0,85 ,το μεγαλύτερο ποσοστό των αρνητικών υποδειγμάτων που έχουν αναγνωριστεί ως αρνητικά (μεγαλύτερη ειδικότητα)-0,95.

### Αποτελέσματα αλγορίθμων για την ασθένεια Meniere μετά την επιλογή χαρακτηριστικών.

Ενεργώντας όπως παραπάνω για την επιλογή χαρακτηριστικών για την ασθένεια Meniere επιλέγουμε “select attributes” με μέθοδο “CfsSubsetEval”.Μετά την εφαρμογή του φίλτρου “AttributeSelection” επιλέχθηκαν τα κατάλληλα χαρακτηριστικά για την ασθένεια Meniere (βλ.Πίνακα 11).Μετά την επιλογή χαρακτηριστικών για την ασθένεια Meniere εφαρμόστηκαν τεχνικές κατηγοριοποίησης δεδομένων και η αξιολόγηση της απόδοσης των μοντέλων θα γίνει όπως και προηγουμένως με χρήση της 10-πλής διασταυρωτικής.Στην συνέχεια εξάγονται τα στατιστικά στοιχεία για την απόδοση των αλγορίθμων σε επιτυχία πρόβλεψης νέων τιμών δεδομένων αλλά και το Δέντρο Απόφασης του J-48.



Εικόνα 33. Δέντρο Απόφασης για την ασθένεια Meniere μετά την επιλογή χαρακτηριστικών

Μετά την επιλογή χαρακτηριστικών η εμφάνιση ασθένειας Meniere εξαρτάται από την εμφάνιση συμπτωμάτων απώλειας ακοής. Αν ο ασθενής έχει συμπτώματα ακοής δεξιού αυτιού ή αριστερού αυτιού ή και από τα δύο σε συνδυασμό με αίσθηση ιλίγγου καθώς περιστρέφεται η κεφαλή τότε πάσχει από την ασθένεια Meniere.

| Μέτρα Ακριβείας Meniere Disease |                          |                                           |                                  |         |         |           |             |           |
|---------------------------------|--------------------------|-------------------------------------------|----------------------------------|---------|---------|-----------|-------------|-----------|
| classifier                      |                          | Correctly Classified Instances - Accuracy | Incorrectly Classified Instances | TP Rate | FP Rate | Precision | Specificity | F-Measure |
| 1-nn                            | before feature selection | 77,56%                                    | 22,44%                           | 0,78    | 0,22    | 0,80      | 0,71        | 0,77      |
|                                 | after feature selection  | 82,68%                                    | 17,32%                           | 0,83    | 0,17    | 0,84      | 0,77        | 0,83      |
| 8-nn                            | before feature selection | 81,89%                                    | 18,11%                           | 0,82    | 0,18    | 0,84      | 0,75        | 0,82      |
|                                 | after feature selection  | 84,25%                                    | 15,75%                           | 0,84    | 0,16    | 0,86      | 0,83        | 0,84      |
| J48                             | before feature selection | 80,71%                                    | 19,29%                           | 0,81    | 0,19    | 0,81      | 0,77        | 0,81      |
|                                 | after feature selection  | 78,74%                                    | 21,26%                           | 0,79    | 0,21    | 0,80      | 0,75        | 0,79      |
| Naïve Bayes                     | before feature selection | 87,01%                                    | 12,99%                           | 0,87    | 0,13    | 0,87      | 0,87        | 0,87      |
|                                 | after feature selection  | 88,19%                                    | 11,81%                           | 0,88    | 0,12    | 0,88      | 0,87        | 0,88      |
| Random Forest                   | before feature selection | 87,01%                                    | 12,99%                           | 0,87    | 0,13    | 0,87      | 0,89        | 0,87      |
|                                 | after feature selection  | 84,25%                                    | 15,75%                           | 0,84    | 0,16    | 0,84      | 0,85        | 0,84      |

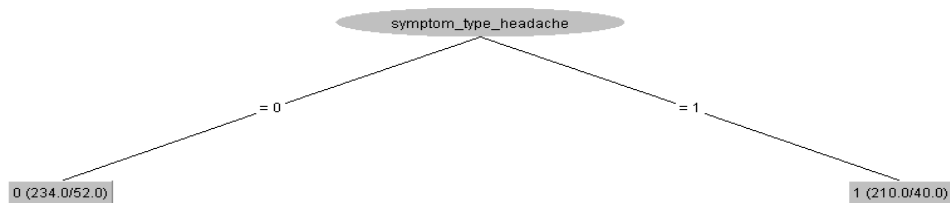
*Πίνακας 8 .Συγκριτικά αποτελέσματα για τα μέτρα ακριβείας της ασθένειας Meniere μετά την επιλογή χαρακτηριστικών*

Μετά την επιλογή χαρακτηριστικών καλύτερη απόδοση εμφανίζει ο αλγόριθμος Naive Bayes με ποσοστό ακριβείας 88,19% ο οποίος επιστρέφει τα καλύτερα αποτελέσματα . Εμφανίζει τις υψηλότερες τιμές για όλα τα υπό εξέταση μέτρα ακρίβειας (precision, recall, f-measure), το μεγαλύτερο ποσοστό των θετικών υποδειγμάτων που έχουν αναγνωρισθεί σωστά ως θετικά (μεγαλύτερη ευαισθησία) –0,88 και το μεγαλύτερο ποσοστό των αρνητικών υποδειγμάτων που έχουν αναγνωρισθεί ως αρνητικά (μεγαλύτερη ειδικότητα)–0,87.

*Αποτελέσματα αλγορίθμων για την ασθένεια Vestibular Migraine μετά την επιλογή χαρακτηριστικών.*

Ενεργώντας όπως παραπάνω για την επιλογή χαρακτηριστικών για την ασθένεια Meniere επιλέγουμε “select attributes” με μέθοδο “CfsSubsetEval”.Μετά την εφαρμογή του φίλτρου “AttributeSelection” τα κατάλληλα χαρακτηριστικά για την ασθένεια Vestibular Migraine (βλ. Πίνακα 12).Μετά την επιλογή χαρακτηριστικών για την ασθένεια Vestibular Migraine εφαρμόστηκαν τεχνικές κατηγοριοποίησης

δεδομένων και η αξιολόγηση της απόδοσης των μοντέλων θα γίνει όπως και προηγουμένως με χρήση της 10-πλής διασταυρωτικής. Στην συνέχεια εξάγονται τα στατιστικά στοιχεία για την απόδοση των αλγορίθμων σε επιτυχία πρόβλεψης νέων τιμών δεδομένων αλλά και το Δέντρο Απόφασης του J-48.



*Εικόνα 34 Δέντρο Απόφασης για την ασθένεια ασθένειας Vestibular Migraine μετά την επιλογή χαρακτηριστικών*

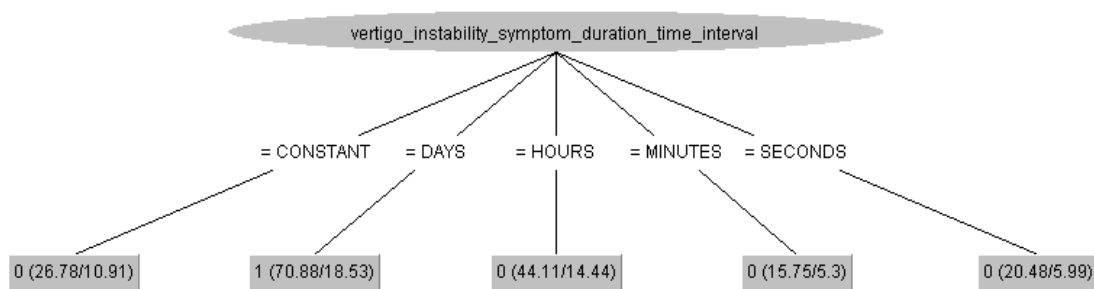
| Μέτρα Ακριβείας Vestibular Migraine |                          |                                           |                                  |         |         |           |             |           |
|-------------------------------------|--------------------------|-------------------------------------------|----------------------------------|---------|---------|-----------|-------------|-----------|
| classifier                          |                          | Correctly Classified Instances - Accuracy | Incorrectly Classified Instances | TP Rate | FP Rate | Precision | Specificity | F-Measure |
| 1-nn                                | before feature selection | 74,10%                                    | 25,90%                           | 0,74    | 0,26    | 0,74      | 0,73        | 0,74      |
|                                     | after feature selection  | 82,88%                                    | 17,12%                           | 0,83    | 0,83    | 0,83      | 0,84        | 0,83      |
| 6-nn                                | before feature selection | 81,08                                     | 18,92                            | 0,81    | 0,19    | 0,81      | 0,66        | 0,81      |
|                                     | after feature selection  | 84,56                                     | 15,54                            | 0,85    | 0,16    | 0,85      | 0,84        | 0,85      |
| J48                                 | before feature selection | 79,05%                                    | 20,95%                           | 0,79    | 0,21    | 0,79      | 0,80        | 0,79      |
|                                     | after feature selection  | 78,60%                                    | 21,40%                           | 0,79    | 0,21    | 0,79      | 0,81        | 0,79      |
| Naïve Bayes                         | before feature selection | 81,76%                                    | 18,24%                           | 0,82    | 0,18    | 0,82      | 0,79        | 0,82      |
|                                     | after feature selection  | 85,36%                                    | 14,64%                           | 0,85    | 0,15    | 0,85      | 0,85        | 0,85      |
| Random Forest                       | before feature selection | 82,43%                                    | 17,57%                           | 0,82    | 0,18    | 0,83      | 0,84        | 0,82      |
|                                     | after feature selection  | 79,96%                                    | 20,05%                           | 0,80    | 0,20    | 0,80      | 0,82        | 0,80      |

*Πίνακας 9. Συγκριτικά αποτελέσματα για τα μέτρα ακριβείας της ασθένειας Vestibular Migraine μετά την επιλογή χαρακτηριστικών*

Μετά την επιλογή χαρακτηριστικών καλύτερη απόδοση εμφανίζει ο αλγόριθμος Naive Bayes με ποσοστό ακριβείας 85,36% ο οποίος επιστρέφει τα καλύτερα αποτελέσματα . Εμφανίζει τις υψηλότερες τιμές για όλα τα υπό εξέταση μέτρα ακρίβειας (precision, recall, f-measure), το μεγαλύτερο ποσοστό των θετικών υποδειγμάτων που έχουν αναγνωριστεί σωστά ως θετικά (μεγαλύτερη ευαισθησία) –0,85 και το μεγαλύτερο ποσοστό των αρνητικών υποδειγμάτων που έχουν αναγνωριστεί ως αρνητικά (μεγαλύτερη ειδικότητα)–0,85.

### Αποτελέσματα αλγορίθμων για την ασθένεια Vestibular Migraine μετά την επιλογή χαρακτηριστικών

Ενεργώντας όπως παραπάνω για την επιλογή χαρακτηριστικών για την ασθένεια Meniere επιλέγουμε “select attributes” με μέθοδο “CfsSubsetEval”.Μετά την εφαρμογή του φίλτρου “AttributeSelection” επιλέχθηκαν τα κατάλληλα χαρακτηριστικά για την ασθένεια Vestibular Migraine (βλ. Πίνακα 12) .Μετά την επιλογή χαρακτηριστικών για την ασθένεια Meniere εφαρμόστηκαν τεχνικές κατηγοριοποίησης δεδομένων και η αξιολόγηση της απόδοσης των μοντέλων θα γίνει όπως και προηγουμένως με χρήση της 10-πλής διασταυρωτικής. Στην συνέχεια εξάγονται τα στατιστικά στοιχεία για την απόδοση των αλγορίθμων σε επιτυχία πρόβλεψης νέων τιμών δεδομένων αλλά και το Δέντρο Απόφασης του J-48.



**Εικόνα 35. Δέντρο Απόφασης για την ασθένεια ασθένειας Vestibular Neuronitis μετά την επιλογή χαρακτηριστικών**

| Μέτρα Ακρίβειας Vestibular Neuronitis |                          |                                           |                                  |         |         |           |             |           |
|---------------------------------------|--------------------------|-------------------------------------------|----------------------------------|---------|---------|-----------|-------------|-----------|
| classifier                            |                          | Correctly Classified Instances - Accuracy | Incorrectly Classified Instances | TP Rate | FP Rate | Precision | Specificity | F-Measure |
| 1-nn                                  | before feature selection | 57,30%                                    | 42,70%                           | 0,57    | 0,43    | 0,58      | 0,55        | 0,56      |
|                                       | after feature selection  | 69,66%                                    | 30,34%                           | 0,70    | 0,30    | 0,71      | 0,66        | 0,69      |
| 3-nn                                  | before feature selection | 68,54%                                    | 31,46%                           | 0,69    | 0,32    | 0,72      | 0,63        | 0,67      |
|                                       | after feature selection  | 70,22%                                    | 29,78%                           | 0,70    | 0,30    | 0,73      | 0,63        | 0,69      |
| J48                                   | before feature selection | 60,67%                                    | 39,33%                           | 0,61    | 0,39    | 0,62      | 0,58        | 0,60      |
|                                       | after feature selection  | 55,62%                                    | 44,38%                           | 0,56    | 0,44    | 0,56      | 0,58        | 0,55      |
| Naïve Bayes                           | before feature selection | 74,16%                                    | 25,84%                           | 0,74    | 0,26    | 0,75      | 0,71        | 0,74      |
|                                       | after feature selection  | 73,03%                                    | 26,97%                           | 0,73    | 0,27    | 0,74      | 0,70        | 0,73      |
| Random Forest                         | before feature selection | 73,60%                                    | 26,40%                           | 0,74    | 0,26    | 0,74      | 0,70        | 0,73      |
|                                       | after feature selection  | 70,22%                                    | 29,78%                           | 0,70    | 0,30    | 0,71      | 0,67        | 0,70      |

Πίνακας 10 .Συγκριτικά αποτελέσματα για τα μέτρα ακριβείας της ασθένειας Vestibular Neuronitis μετά την επιλογή χαρακτηριστικών

Μετά την επιλογή χαρακτηριστικών καλύτερη απόδοση εμφανίζει ο αλγόριθμος Random Forest με ποσοστό ακριβείας 73,60%. Τα καλύτερα αποτελέσματα συνολικά όμως επιστρέφονται από τον αλγόριθμο Naïve Bayes χωρίς την επιλογή χαρακτηριστικών , όπου το ποσοστό ακριβείας κυμαίνεται στο 74,16% και εμφανίζει τις υψηλότερες τιμές για όλα τα υπό εξέταση μέτρα ακρίβειας (precision, recall, f-measure), το μεγαλύτερο ποσοστό των θετικών υποδειγμάτων που έχουν αναγνωριστεί σωστά ως θετικά (μεγαλύτερη ευαισθησία) –0,74 και το μεγαλύτερο ποσοστό των αρνητικών υποδειγμάτων που έχουν αναγνωριστεί ως αρνητικά (μεγαλύτερη ειδικότητα)-0,71.

| Disease                                                        | Features                                           |                                                    |                                             |
|----------------------------------------------------------------|----------------------------------------------------|----------------------------------------------------|---------------------------------------------|
| <b>BPPV-1</b><br><i>Attributes 11</i><br><i>Instances 104</i>  | Clinical examination head thrust                   | Clinical examination dix-hallpike                  | Auditory test PTA 8000 AC right             |
|                                                                | Personaldisease bundle name                        | Vertigo trigger bending over                       | Symptom type phonophobic                    |
|                                                                | Symptom type scotoma                               | Vertigo trigger rolling over in bed                | Vertigo trigger head movement               |
|                                                                | Vertigo instability symptom duration time interval |                                                    |                                             |
| <b>BPPV-2</b><br><i>Attributes 5</i><br><i>Instances 104</i>   | Vertigo trigger complex visual environments        | Vertigo instability symptom duration time interval | Symptom type difficulty walking in darkness |
|                                                                | Clinical examination dix-hallpike                  |                                                    |                                             |
| <b>MENIERE</b><br><i>Attributes 10</i><br><i>Instances 254</i> | Symptom type lightheaded                           | Symptoms hearing loss                              | Imaging imaging result                      |
|                                                                | Vertigo trigger head movement                      | Vertigo symptom type bundle name                   | Symptom type hearing loss                   |
|                                                                | Vertigo instability symptom duration time interval | Vertigo trigger complex visual environments        | Patient sex                                 |

Πίνακας 11 . Πίνακας χαρακτηριστικών ασθενειών

| Disease                                                                     | Features                                           |                                           |                          |
|-----------------------------------------------------------------------------|----------------------------------------------------|-------------------------------------------|--------------------------|
| <b>Vestibular Migraine</b><br><i>Attributes 8</i><br><i>Instances 44</i>    | Personaldisease bundle name                        | Vertigo trigger rolling over in bed       | Symptom type headache    |
|                                                                             | Symptom type difficulty walking on uneven surfaces | Vertigo trigger standing up rapid ascents | Symptom type lightheaded |
|                                                                             | Patient sex                                        |                                           |                          |
| <b>Vestibular Neuronitis</b><br><i>Attributes 7</i><br><i>Instances 178</i> | Vertigo symptom type bundle name                   | Symptom type hearing loss                 | Symptom type cervicalgia |
|                                                                             | Vertigo instability symptom duration time interval | Personaldisease bundle name               | Bilateral                |

Πίνακας 12. Πίνακας χαρακτηριστικών ασθενειών

## Συμπεράσματα- Μελλοντική Έρευνα.

Η διαδικασία της ανακάλυψης γνώσης από βάσεις δεδομένων συνιστά ένα πολύ χρήσιμο και ενδιαφέρον πεδίο. Το πλήθος, η ποιότητα των αλγορίθμων αυτού του είδους, αλλά και των εργαλείων ανακάλυψης γνώσης εξελίσσονται διαρκώς ενώ παράλληλα παραμένουν προσιτοί στους χρήστες. Φυσικά, ο κρίσιμος και ευαίσθητος τομέας της υγείας αποτελεί σημαντικό πεδίο εφαρμογής της εξόρυξης δεδομένων.

Στην συγκεκριμένη εργασία πραγματοποιήθηκε έρευνα στην κατασκευή διαγνωστικών μεθόδων, αναπτύσσοντας διάφορα μοντέλα ταξινόμησης με αλγορίθμους μηχανικής μάθησης, επιτρέποντας την εκτίμηση των αιτιών πρόκλησης ίλιγγου για την ανίχνευση και διάγνωση της σωστής ασθένειας. Οι διαγνωστικές μέθοδοι που παρουσιάζονται στην παρούσα εργασία φάνηκε ότι μπορούν να αποτελέσουν αξιόπιστα διαγνωστικά εργαλεία.

Σε πρώτο στάδιο έγινε αναφορά σχετικά με το αιθουσαίο σύστημα και την έννοια της ανθρώπινης ισορροπίας. Αφού αναλύθηκε η έννοια του ίλιγγου και τα πιθανά αίτια πρόκλησης του, αναφέρθηκε το πρόβλημα της δυσκολίας διάγνωσης των ασθενειών στις οποίες οφείλεται. Σε επόμενο στάδιο, αναφορικά παρατέθηκαν έρευνες που έχουν πραγματοποιηθεί στο παρελθόν και οι οποίες στοχεύουν στην δημιουργία Κλινικών Συστημάτων Υποστήριξης Απόφασης που βοηθούν τους ιατρούς στην διάγνωση των ασθενειών που προκαλούν ίλιγγο. Τέλος, ακολουθήθηκαν τα στάδια της διαδικασίας εξόρυξης γνώσης και χρησιμοποιήθηκε ένα πολύ γνωστό και ελεύθερο λογισμικό εργαλείο ανακάλυψης γνώσης: το Weka για την επεξεργασία των δεδομένων (dataset), την εφαρμογή των αλγορίθμων, την παρουσίαση και την αξιολόγηση των συμπερασμάτων.

Τα κυριότερα συμπεράσματα που προέκυψαν από την έρευνα, μπορούν να συνοψιστούν στα εξής:

Από την παρουσίαση των αποτελεσμάτων στους βασικούς αλγόριθμους, υπερτέρησε σε απόδοση ο Random Forest ο οποίος έφερνε τα καλύτερα αποτελέσματα πρόβλεψης για όλες τις ασθένειες πριν την επιλογή χαρακτηριστικών, ενώ μετά την επιλογή χαρακτηριστικών ως καλύτερος ταξινομητής διακρίθηκε ο Naive Bayes. Τα ποσοστά πρόβλεψης ακριβείας χαρακτηρίζονται ως υψηλά, αφού κυμαίνονται από 55,618% έως 88,189%. Στη συνέχεια, με την παρουσίαση αποτελεσμάτων των συνδυαστικών μεθόδων, διαπιστώθηκε ότι εκμεταλλευόμενοι τα πλεονεκτήματα της μεθόδου επιλογής χαρακτηριστικών, επιτεύχθηκε βελτίωση της απόδοσης πρόβλεψης των βασικών αλγορίθμων, ανιχνεύοντας πολύ πιο αποτελεσματικά τις κατηγορίες ασθενειών από ότι πετυχαίνει ο κάθε ένας αλγόριθμος χωρίς την επιλογή χαρακτηριστικών, εξάγοντας ακριβέστερα και πιο αξιόπιστα διαγνωστικά συμπεράσματα ως επί τον πλείστον.

Κρίνοντας από τα αποτελέσματα, συμπεραίνουμε πως η επιτυχής ανίχνευση και διάγνωση ασθένειας που προκαλεί ίλιγγο με χρήση του προτεινόμενου μοντέλου είναι εφικτή. Τα αποτελέσματα αυτής της εργασίας υποδεικνύουν πόσο χρήσιμες είναι οι τεχνικές των αλγορίθμων μηχανικής μάθησης και τις δυνατότητες που παρέχει το εργαλείο εξόρυξης δεδομένων Weka στην ανάλυση των λειτουργικών παραμέτρων για την εφαρμογή του διαγνωστικού συστήματος ανίχνευσης και διάγνωσης ασθενειών που προκαλούν ίλιγγο.

Η εφαρμογή των αλγορίθμων εξόρυξης γνώσης σε ιατρικά δεδομένα μπορούν να φανούν χρήσιμες για την παροχή ποιοτικότερων υπηρεσιών περίθαλψης στους ασθενείς και πιο αποτελεσματική διαχείριση των διοικούντων των παρόχων των υπηρεσιών υγείας. Τα αποτελέσματα από την επεξεργασία και

μελέτη πραγματικών ιατρικών δεδομένων μπορούν να εφαρμοστούν στην καθημερινή ιατρική πράξη με στόχο την κατηγοριοποίηση των ασθενών , των περιπτώσεων ή και της φαρμακευτικής αγωγής ώστε να προσφέρεται ποιοτικότερη περίθαλψη. Με στόχο την επίτευξη πιθανής διάγνωσης της κατάλληλης ασθένειας που προκαλεί ίλιγγο μπορεί η παραπάνω μεθοδολογία η οποία βασίζεται σε τεχνικές εξόρυξης δεδομένων να χρησιμοποιηθεί για την ανάπτυξη εργαλείου πρότασης και ενός Κλινικού Συστήματος Υποστήριξης Απόφασης για να βοηθήσει τους γιατρούς αφενός για την σωστή εύρεση πληροφοριών από τον εκάστοτε ασθενή, αφετέρου για την υποστήριξη διάγνωσης ισορροπίας.

## Βιβλιογραφία.

"[Ιλιγγος: Αίτια, διάγνωση και θεραπεία](#)". Κυπριακή Εταιρεία Ωτορινολαρυγγολογίας. Ανακτήθηκε στις 17 Νοεμβρίου 2020

"[Mining Balance Disorders' data for the development of Diagnostic Decision Support Systems](#)" City, University of London Institutional Repository. Ανακτήθηκε στις 22 Δεκεμβρίου 2020

"[Clinical practice guideline: Benign paroxysmal positional vertigo](#)" Otolaryngology–Head and Neck Surgery. Ανακτήθηκε στις 20 Νοεμβρίου 2020

"[Benign paroxysmal positional vertigo \(benign paroxysmal nystagmus\)](#)" Cancer Therapy Advisor. Ανακτήθηκε στις 13 Νοεμβρίου 2020

"[ΙΛΙΓΓΟΣ. ΤΑ ΣΥΧΝΟΤΕΡΑ ΑΙΤΙΑ ΤΟΥ ΚΑΙ Η ΑΝΤΙΜΕΤΩΠΙΣΗ ΤΟΥ](#)". ΔΡ ΜΑΡΙΔΑΚΗ ΧΡΥΣΟΥΛΑ. Ανακτήθηκε στις 26 Νοεμβρίου 2020

"[Ακουστικό Νευρίνωμα](#)" Νευροχειρουργική Νευρολογία. Ανακτήθηκε στις 28 Νοεμβρίου 2020

"[Test, training and validation sets](#)" Brains to Bytes. Ανακτήθηκε στις 27 Ιανουαρίου 2021

"[Cross Validation](#)" Kaggle. Ανακτήθηκε στις 23 Ιανουαρίου 2021

"[An overview of clinical decision support systems: benefits, risks, and strategies for success](#)" Digital Medicine. Ανακτήθηκε στις 03 Φεβρουαρίου 2021

"[Τι είναι το Αιθουσαίο και Ιδιοδεκτικό σύστημα; Δραστηριότητες ευόδωσης τους](#)" IPaidia.gr. Ανακτήθηκε στις 11 Ιανουαρίου 2021

"[Menieres Disease](#)" Heath Line. Ανακτήθηκε στις 05 Δεκεμβρίου 2020

"[Persistent postural-perceptual dizziness \(PPPD\): a common, characteristic and treatable cause of chronic dizziness](#)" BMJ Journals. Ανακτήθηκε στις 16 Ιανουαρίου 2021

"[Ένας χρήσιμος οδηγός για τις μελέτες, τις έρευνες και τα επιστημονικά περιοδικά](#)" ellinikahoaxes.gr. Ανακτήθηκε στις 03 Φεβρουαρίου 2021

"[Machine Learning](#)" Medium. Ανακτήθηκε στις 24 Ιανουαρίου 2021

"[Decision Tree Classification Algorithm](#)" Java Point Ανακτήθηκε στις 17 Δεκεμβρίου 2020