



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

Ερμηνεύσιμη τεχνητή νοημοσύνη για την απεικονιστική μελέτη νευροεκφυλιστικών ασθενειών σχετιζόμενων με γήρανση

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Μαρία Ελευθερία Ι. Βλόντζου

Επιβλέπουσα : Κωνσταντίνα Νικήτα
Καθηγήτρια Ε.Μ.Π.

Αθήνα, Οκτώβριος 2022



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

Ερμηνεύσιμη τεχνητή νοημοσύνη για την απεικονιστική μελέτη νευροεκφυλιστικών ασθενειών σχετιζόμενων με γήρανση

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Μαρία Ελευθερία Ι. Βλόντζου

Επιβλέπουσα : Κωνσταντίνα Νικήτα
Καθηγήτρια Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 18^η Οκτωβρίου 2022

.....
Κωνσταντίνα Νικήτα
Καθηγήτρια Ε.Μ.Π.

.....
Γεώργιος Στάμου
Καθηγητής Ε.Μ.Π.

.....
Αθανάσιος Βουλόδημος
Επίκουρος Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2022

.....
Μαρία Ελευθερία Ι. Βλόντζου

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

Copyright © Μαρία Ελευθερία Ι. Βλόντζου, 2022

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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

Περίληψη

Πάνω από 50 εκατομμύρια άνθρωποι παγκοσμίως έχουν διαγνωστεί με κάποιας μορφής άνοια, ενώ ο αριθμός αυτός πρόκειται να ξεπεράσει τα 152 εκατομμύρια έως το 2050. Η Ήπιας Μορφής Νοητική Εξασθένηση (Mild Cognitive Impairment - MCI) και η άνοια ανήκουν στις κυριότερες νευροεκφυλιστικές ασθένειες που οδηγούν σε νοητική δυσλειτουργία. Καθώς οι ασθενείς παραμένουν ασυμπτωματικοί για μεγάλο χρονικό διάστημα πριν τη διάγνωση, η απουσία αποτελεσματικών διαγνωστικών εργαλείων για την έγκαιρη διάγνωση των νευροεκφυλιστικών ασθενειών αποτελεί σημαντικό ζήτημα. Η παρούσα μελέτη στοχεύει στην ανάπτυξη ενός μοντέλου τεχνητής νοημοσύνης για τον διαχωρισμό ασθενών με φυσιολογική εγκεφαλική γήρανση, με Ήπιας Μορφής Νοητική Εξασθένηση και άνοια. Η προτεινόμενη μέθοδος αξιοποιεί ογκομετρικά δεδομένα από T1-σταθμισμένες εικόνες μαγνητικής τομογραφίας του εγκεφάλου και γενετικά δεδομένα και χρησιμοποιεί σχήματα διαίρεσης του προβλήματος πολλαπλών κλάσεων σε δυαδικά σε συνδυασμό με μία μέθοδο συλλογικής μάθησης, κατά την οποία συνδυάζονται οι προβλέψεις πολλαπλών ταξινομητών. Στόχος της τεχνικής αυτής είναι η αύξηση της προβλεπτικής ικανότητας, πρώτον, αντιμετωπίζοντας την αρχική ανισορροπία των δεδομένων και δεύτερο, ελαττώνοντας τη διακύμανση των αποτελεσμάτων πρόβλεψης. Επιπλέον, η μελέτη εστιάζει στη χρήση μεθόδου ερμηνευσιμότητας, ώστε να παρέχει φιλικές προς τον άνθρωπο επεξηγήσεις για τις προβλέψεις και να αυξήσει την αξιοπιστία του μοντέλου, που είναι ιδιαίτερα σημαντική στα ιατρικά προβλήματα. Το τελικό μοντέλο παρείχε ισορροπημένη ακρίβεια ίση με 87.5% και f1-σκορ ίσο με 90.8%. Τα αποτελέσματα των μεθόδων ερμηνευσιμότητας εστίασαν στην επίδραση των χαρακτηριστικών για την πρόβλεψη και βρέθηκε πως η ατροφία ορισμένων περιοχών όπως ο ιππόκαμπος ή η αμυγδαλή, η διεύρυνση των πλευρικών κοιλιών και διαφοροποιήσεις του DNA που σχετίζονται με την παρουσία του γονιδίου της απολιποπρωτεΐνης E συσχετίζονται με την ανάπτυξη των συγκεκριμένων νευροεκφυλιστικών παθήσεων, ενώ η παρουσία διαφοροποιήσεων σχετικών με άλλα γονίδια λειτουργεί προστατευτικά έναντι αυτών των ασθενειών.

Λέξεις Κλειδιά

γήρανση εγκεφάλου, νευροεκφυλιστικές ασθένειες, ήπιας μορφής νοητική εξασθένηση, άνοια, μηχανική μάθηση, ερμηνεύσιμη τεχνητή νοημοσύνη, συλλογική μάθηση, bagging

Abstract

Over 50 million people worldwide have been diagnosed with some type of dementia, while this number will probably exceed 152 millions by 2050. Mild Cognitive Impairment (MCI) and Dementia belong in the most prominent neurodegenerative diseases leading to cognitive dysfunction. Since patients remain asymptomatic for a long period of time before diagnosis, the lack of effective diagnostic tools for the identification of neurodegenerative diseases in early stages remains a serious issue. This study aims to assist in the diagnosis of brain atrophy by creating artificial intelligence algorithms that make a distinction between healthy individuals, patients with Mild Cognitive Impairment and Dementia patients. The proposed method utilises volumetric measurements from T1-weighted brain imaging and genetic data and uses multiclass classification decomposition schemes combined with an ensemble learning technique, which aggregates multiple classifiers' predictions. The purpose of this technique is to increase the prediction accuracy, firstly, by combating the original imbalance of the data and secondly, by decreasing the predictions' variance. Furthermore, this study engages artificial intelligence interpretability methods in order to provide human-friendly explanations for the predictions and increase the model's reliability, which is particularly important in the medical field. The proposed model yielded 87.5% balanced accuracy and 90.8% f1-score. The interpretability results showed how certain features affect the models' predictions, indicating that the atrophy of brain regions, such as the hippocampus or the amygdala, the enlargement of the lateral ventricles and the presence of DNA polymorphisms related to the apolipoprotein E gene lead to the development of the specific neurodegenerative diseases, while the presence of DNA polymorphisms regarding other genes can play a protective role against those diseases.

Keywords

brain ageing, neurodegenerative diseases, mild cognitive impairment, dementia, machine learning, interpretable artificial intelligence, ensemble learning, bagging

Αφιερώνεται στη γιαγιά μου, Ελευθερία
και στη μνήμη του πατέρα μου, Γιάννη

Ευχαριστίες

Θα ήθελα να ευχαριστήσω θερμά την επιβλέπουσα καθηγήτρια κα Κωνσταντίνα Νικήτα για την εμπιστοσύνη που μου έδειξε και την ευκαιρία που μου έδωσε να εκπονήσω τη διπλωματική μου εργασία υπό την επίβλεψή της σε ένα τόσο ενδιαφέρον θέμα, όπως επίσης και τον κ. Νταβατζικό για την πολύτιμη βοήθειά του.

Επίσης, θα ήθελα να ευχαριστήσω ιδιαίτερα τη μεταδιδακτορική ερευνήτρια κα Καλλιόπη Δαλακλείδη και την υποψήφια διδάκτορα κα Ιωάννα Σκαμπαρδώνη για τη συνεχή καθοδήγηση και την ανεκτίμητη βοήθειά τους. Είμαι ευγνώμων για την εξαιρετική συνεργασία που είχαμε και τον πολύτιμο χρόνο που αφιέρωσαν καθόλη τη διάρκεια της παρούσας εργασίας.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου και τους φίλους μου Γιώργο, Φώτη και Ελευθερία για την αμέριστη στήριξή τους καθόλη τη διάρκεια των σπουδών μου και ιδιαίτερα τη Χρύσα για τη συνεχή ενθάρρυνσή της και την κοινή μας φοιτητική πορεία τα τελευταία χρόνια.

Πίνακας Περιεχομένων

Περίληψη	6
Abstract	8
Ευχαριστίες	10
Κατάλογος Εικόνων	14
Κατάλογος Πινάκων.....	16
1. Εισαγωγή.....	17
2. Γήρανση Εγκεφάλου και Σχετιζόμενες Ασθένειες.....	19
2.1 Ανατομικές Περιοχές του Ανθρώπινου Εγκεφάλου.....	19
2.2 Πολυμορφισμοί Ενός Νουκλεοτιδίου	23
2.3 Φυσιολογική Γήρανση εγκεφάλου	23
2.4 Ασθένειες Σχετιζόμενες με τη Γήρανση του Εγκεφάλου.....	24
2.4.1 Ήπιας Μορφής Νοητική Εξασθένηση	24
2.4.2 Άνοια	28
3. Θεωρητικό υπόβαθρο	30
3.1 Μέθοδοι Ταξινόμησης	30
3.1.1 Αλγόριθμοι Μηχανικής Μάθησης	30
3.1.1.1 Τυχαία Δάση	30
3.1.1.2 Λογιστική Παλινδρόμηση.....	35
3.1.1.3 Μηχανές Διανυσμάτων Υποστήριξης.....	36
3.1.1.4 Gradient Boosting	40
3.1.1.5 XGBoost.....	42
3.1.2 Αλγόριθμοι Βαθιάς Μάθησης	44
3.1.2.1 Πολυστρωματικό Perceptron	44
3.1.3 Συλλογική Μάθηση.....	46
3.1.3.1 Bagging	47
3.1.4 Προβλήματα ταξινόμησης πολλαπλών κλάσεων	48
3.1.4.1 Ένα-προς-Ένα	49
3.1.4.2 Ένα-προς-Όλα.....	50
3.2 Μετρικές	51
3.3 Μέθοδοι Ερμηνευσιμότητας.....	54
3.3.1 Σημαντικότητα Gini.....	55
3.3.2 Πρόσθετες Shapley Επεξηγήσεις.....	55
3.3.3 Τοπικές Ερμηνεύσιμες Επεξηγήσεις Ανεξάρτητες από το Μοντέλο.....	58

3.3.4	Διάγραμμα Μερικής Εξάρτησης	59
3.3.5	Αντιπαραθετικά παραδείγματα Permute Attack	60
4.	Μέθοδος	64
4.1	Σύνολο Δεδομένων.....	64
4.2	Προεπεξεργασία Δεδομένων	65
4.3	Ταξινόμηση Πολλαπλών Κλάσεων	67
4.3.1	Ένα-προς-Όλα	67
4.3.2	Ένα-προς-Ένα.....	67
4.4	Συλλογική Μάθηση.....	68
4.5	Προσδιορισμός Βέλτιστων Υπερπαραμέτρων και Σχήμα Αξιολόγησης.....	69
4.6	Ερμηνευσιμότητα	72
4.6.1	Σημαντικότητα Gini.....	72
4.6.2	Πρόσθετες Shapley Επεξηγήσεις.....	73
4.6.3	Τοπικές Ερμηνεύσιμες Επεξηγήσεις Ανεξάρτητες από το Μοντέλο.....	73
4.6.4	Διάγραμμα Μερικής Εξάρτησης	73
4.6.5	Αντιπαραθετικές Επεξηγήσεις Permute Attack.....	74
5.	Αποτελέσματα	75
5.1	Αποτελέσματα Μοντέλων Ταξινόμησης	75
5.1.1	Σύγκριση τεχνικών.....	75
5.1.2	Σύγκριση για διαφορετικά χαρακτηριστικών του συνόλου δεδομένων	77
5.2	Αποτελέσματα Μοντέλων Ερμηνευσιμότητας	77
5.2.1	Gini Σημαντικότητα Χαρακτηριστικών	77
5.2.1.1	Τυχαία Δάση	78
5.2.1.2	Gradient Boosting	79
5.2.1.3	XGBoost	80
5.2.2	Πρόσθετες Shapley Επεξηγήσεις.....	81
5.2.2.1	Σημαντικότητα Χαρακτηριστικών SHAP	81
5.2.2.2	Γραφική Παράσταση Σύνοψης	82
5.2.3	Τοπικές Ερμηνεύσιμες Επεξηγήσεις Ανεξάρτητες από το Μοντέλο.....	84
5.2.4	Διάγραμμα Μερικής Εξάρτησης	86
5.2.5	Αντιπαραθετικές Επεξηγήσεις Permute Attack.....	89
6.	Συμπεράσματα και Συζήτηση.....	93
6.1	Σύνοψη	93
6.2	Συμπεράσματα	94
6.3	Περιορισμοί	95

6.4 Μελλοντικές Επεκτάσεις.....	95
7. Βιβλιογραφία	97
Πηγές Εικόνων	102

Κατάλογος Εικόνων

Εικόνα 1: Τα ανατομικά μέρη του εγκεφάλου (αριστερό ημισφαίριο). Διακρίνεται ο εγκέφαλος (cerebrum), η παρεγκεφαλίδα (cerebellum) και το στέλεχος του εγκεφάλου (brainstem).....	19
Εικόνα 2: Πλευρική όψη των περιοχών του εγκεφαλικού φλοιού (αριστερό ημισφαίριο). Σε κάθε έλικα έχει αποδοθεί ξεχωριστό χρώμα.....	21
Εικόνα 3: Επίπεδη νευροανατομία σε αξονικό προσανατολισμό -12mm κάτω από το επίπεδο του πρόσθιου συνδέσμου (anterior commissure) που συνδέει τα δύο ημισφαίρια.....	22
Εικόνα 4: Εξέλιξη των απεικονιστικών χαρακτηριστικών από τη φυσιολογική κατάσταση στην ήπια νοητική εξασθένηση και την άνοια.....	27
Εικόνα 5: Το βέλτιστο υπερεπίπεδο (Μηχανές Διανυσμάτων Υποστήριξης.....	37
Εικόνα 6: Η δομή ενός νευρωνικού δικτύου πρόσθιας τροφοδότησης.....	45
Εικόνα 7: Σχηματική απεικόνιση της μεθόδου bagging και των σταδίων bootstrapping και aggregating από τα οποία προκύπτει ο τελικός ταξινομητής συλλογικής μάθησης.....	48
Εικόνα 8: Σχηματική αναπαράσταση των τεχνικών διαίρεσης Ένα-προς-Όλα (πάνω) και Ένα-προς-Ένα (κάτω).....	50
Εικόνα 9: Παράδειγμα πίνακα σύγχυσης για ταξινόμηση πολλαπλών κλάσεων.....	51
Εικόνα 10: Ιστόγραμμα των ηλικιών των συμμετεχόντων.....	64
Εικόνα 11: Διαγράμματα Διασποράς των τιμών όγκου του αριστερού ιπποκάμπου συναρτήσει της ηλικίας των συμμετεχόντων: α) πριν την προεπεξεργασία β) μετά τη γραμμική διόρθωση γ) μετά την κανονικοποίηση. Με μπλε συμβολίζονται τα υγιή άτομα (CN), με πράσινο τα άτομα με Ήπια μορφή νοητική εξασθένηση (MCI) και με κόκκινο τα άτομα με άνοια (DEM).....	66
Εικόνα 12: Σχηματική αναπαράσταση του μοντέλου πρόβλεψης και του διαχωρισμού του συνόλου εκπαίδευσης κατά την εφαρμογή της τεχνικής bagging και του σχήματος Ένα-προς-Ένα.....	69
Εικόνα 13: Σχηματική αναπαράσταση του μοντέλου εμφωλευμένης διασταυρούμενης επικύρωσης που εφαρμόστηκε για τον εντοπισμό των βέλτιστων υπερπαραμέτρων κάθε ταξινομητή με τη μέθοδο αναζήτησης πλέγματος και τον εντοπισμό του βέλτιστου μοντέλου.....	70
Εικόνα 14: Κατάταξη Σημαντικότητας Χαρακτηριστικών Gini κατά την ταξινόμηση με το μοντέλο Τυχαίων Δασών για το ζεύγος κλάσεων α) CN και MCI, β) MCI και DEM και γ) CN και DEM.....	78
Εικόνα 15: Κατάταξη Σημαντικότητας Χαρακτηριστικών Gini κατά την ταξινόμηση με το μοντέλο Gradient Boosting για το ζεύγος κλάσεων α) CN και MCI, β) MCI και DEM και γ) CN και DEM.....	79
Εικόνα 16: Κατάταξη Σημαντικότητας Χαρακτηριστικών Gini κατά την ταξινόμηση με το μοντέλο XGBoost για το ζεύγος κλάσεων α) CN και MCI, β) MCI και DEM και γ) CN και DEM.....	80
Εικόνα 17: Κατάταξη Σημαντικότητας Χαρακτηριστικών SHAP κατά την ταξινόμηση με το μοντέλο SVM για το ζεύγος κλάσεων α) CN και MCI, β) MCI και DEM και γ) CN και DEM.....	82
Εικόνα 18: Διαγράμματα Σύνοψης SHAP κατά την ταξινόμηση με το μοντέλο SVM για το ζεύγος κλάσεων α) CN και MCI, β) MCI και DEM και γ) CN και DEM.....	83
Εικόνα 19: Αποτελέσματα του μοντέλου LIME για την ταξινόμηση με ταξινομητή Τυχαίων Δασών για τις κλάσεις CN και MCI. Ορθώς θετική περίπτωση, όπου ένα MCI άτομο ταξινομείται σωστά. Αριστερά: η προβλεπόμενη πιθανότητα της συγκεκριμένης περίπτωσης για τις δύο κλάσεις, Μέση: τα χαρακτηριστικά που ωθούν την πρόβλεψη στην κλάση MCI (πορτοκαλί) και στην κλάση CN (μπλε) αντίστοιχα, Δεξιά: μια κατάταξη σημαντικότητας των χαρακτηριστικών και αντίστοιχες τιμές τους. Τα χαρακτηριστικά πορτοκαλί χρώματος ωθούν την πρόβλεψη στην κλάση MCI και τα χαρακτηριστικά μπλε χρώματος ωθούν την πρόβλεψη στην κλάση CN.....	84

Εικόνα 20: Αποτελέσματα του μοντέλου LIME για την ταξινόμηση με ταξινομητή Μηχανών Διανυσμάτων Υποστήριξης για τις κλάσεις CN και MCI. Ορθώς θετική περίπτωση, όπου ένα MCI άτομο ταξινομείται σωστά. Αριστερά: η προβλεπόμενη πιθανότητα της συγκεκριμένης περίπτωσης για τις δύο κλάσεις, Μέση: τα χαρακτηριστικά που ωθούν την πρόβλεψη στην κλάση MCI (πορτοκαλί) και στην κλάση CN (μπλε) αντίστοιχα, Δεξιά: μια κατάταξη σημαντικότητας των χαρακτηριστικών και αντίστοιχες τιμές τους. Τα χαρακτηριστικά πορτοκαλί χρώματος ωθούν την πρόβλεψη στην κλάση MCI και τα χαρακτηριστικά μπλε χρώματος ωθούν την πρόβλεψη στην κλάση CN.84

Εικόνα 21: Αποτελέσματα του μοντέλου LIME για την ταξινόμηση με ταξινομητή Τυχαίων Δασών για τις κλάσεις MCI και DEM. Εσφαλμένα αρνητική περίπτωση, όπου ένα άτομο με άνοια (DEM) ταξινομείται στην κλάση MCI. Αριστερά: η προβλεπόμενη πιθανότητα της συγκεκριμένης περίπτωσης για τις δύο κλάσεις, Μέση: τα χαρακτηριστικά που ωθούν την πρόβλεψη στην κλάση DEM (πορτοκαλί) και στην κλάση MCI (μπλε) αντίστοιχα, Δεξιά: μια κατάταξη σημαντικότητας των χαρακτηριστικών και αντίστοιχες τιμές τους. Τα χαρακτηριστικά πορτοκαλί χρώματος ωθούν την πρόβλεψη στην κλάση DEM και τα χαρακτηριστικά μπλε χρώματος ωθούν την πρόβλεψη στην κλάση MCI.85

Εικόνα 22: Αποτελέσματα του μοντέλου LIME για την ταξινόμηση με ταξινομητή Μηχανών Διανυσμάτων Υποστήριξης για τις κλάσεις MCI και DEM. Εσφαλμένα αρνητική περίπτωση, όπου ένα άτομο με άνοια (DEM) ταξινομείται στην κλάση MCI. Αριστερά: η προβλεπόμενη πιθανότητα της συγκεκριμένης περίπτωσης για τις δύο κλάσεις, Μέση: τα χαρακτηριστικά που ωθούν την πρόβλεψη στην κλάση DEM (πορτοκαλί) και στην κλάση MCI (μπλε) αντίστοιχα, Δεξιά: μια κατάταξη σημαντικότητας των χαρακτηριστικών και αντίστοιχες τιμές τους. Τα χαρακτηριστικά πορτοκαλί χρώματος ωθούν την πρόβλεψη στην κλάση DEM και τα χαρακτηριστικά μπλε χρώματος ωθούν την πρόβλεψη στην κλάση MCI.85

Εικόνα 23: Αποτελέσματα του μοντέλου LIME για την ταξινόμηση με ταξινομητή Τυχαίων Δασών για τις κλάσεις CN και DEM. Ορθώς αρνητική περίπτωση, όπου ένα υγιές άτομο ταξινομείται σωστά. Αριστερά: η προβλεπόμενη πιθανότητα της συγκεκριμένης περίπτωσης για τις δύο κλάσεις, Μέση: τα χαρακτηριστικά που ωθούν την πρόβλεψη στην κλάση DEM (πορτοκαλί) και στην κλάση CN (μπλε) αντίστοιχα, Δεξιά: μια κατάταξη σημαντικότητας των χαρακτηριστικών και αντίστοιχες τιμές τους. Τα χαρακτηριστικά πορτοκαλί χρώματος ωθούν την πρόβλεψη στην κλάση DEM και τα χαρακτηριστικά μπλε χρώματος ωθούν την πρόβλεψη στην κλάση CN.....86

Εικόνα 24: Αποτελέσματα του μοντέλου LIME για την ταξινόμηση με ταξινομητή Μηχανών Διανυσμάτων Υποστήριξης για τις κλάσεις CN και DEM. Εσφαλμένα θετική περίπτωση, όπου το ίδιο υγιές άτομο με την Εικόνα 23, ταξινομείται λανθασμένα στην κλάση DEM. Αριστερά: η προβλεπόμενη πιθανότητα της συγκεκριμένης περίπτωσης για τις δύο κλάσεις, Μέση: τα χαρακτηριστικά που ωθούν την πρόβλεψη στην κλάση DEM (πορτοκαλί) και στην κλάση CN (μπλε) αντίστοιχα, Δεξιά: μια κατάταξη σημαντικότητας των χαρακτηριστικών και αντίστοιχες τιμές τους. Τα χαρακτηριστικά πορτοκαλί χρώματος ωθούν την πρόβλεψη στην κλάση DEM και τα χαρακτηριστικά μπλε χρώματος ωθούν την πρόβλεψη στην κλάση CN.....86

Εικόνα 25: Διαγράμματα Μερικής Εξάρτησης πέντε σημαντικών χαρακτηριστικών όπως προέκυψαν από την ταξινόμηση των κλάσεων CN και DEM με το μοντέλο Τυχαίων Δασών. Στον άξονα x αναπαρίστανται οι τιμές των χαρακτηριστικών και στον άξονα y η πιθανότητα ενός δείγματος να ανήκει στην κλάση DEM συναρτήσει της εκάστοτε τιμής του χαρακτηριστικού, επομένως τα διαγράμματα αυτά δείχνουν την επίδραση συγκεκριμένων χαρακτηριστικών στην τελική πρόβλεψη.... 87

Εικόνα 26: Διάγραμμα Μερικής Εξάρτησης που απεικονίζει την επίδραση του rs429358 SNP, το οποίο ανήκει στο γονίδιο ApoE [54] για την πιθανότητα ταξινόμησης στην κλάση των ασθενών με άνοια, κατά το δυαδικό υποπρόβλημα του ζεύγους κλάσεων CN και DEM.88

Εικόνα 27: Διάγραμμα Μερικής Εξάρτησης που απεικονίζει την επίδραση του rs983392 SNP, το οποίο ανήκει στο γονίδιο MS4A6A [55] για την πιθανότητα ταξινόμησης στην κλάση των ασθενών με MCI, κατά το δυαδικό υποπρόβλημα του ζεύγους κλάσεων CN και MCI.88

Εικόνα 28: Διάγραμμα Μερικής Εξάρτησης που απεικονίζει την επίδραση του rs6024870 SNP, το οποίο ανήκει στο γονίδιο CASS4 [56] για την πιθανότητα ταξινόμησης στην κλάση των ασθενών με

άνοια, κατά το δυαδικό υποπρόβλημα του ζεύγους κλάσεων MCI και DEM	88
Εικόνα 29: Αντιπαραθετικά παραδείγματα Permute Attack για ανατροπή της πρόβλεψης ενός CN δείγματος και την ταξινόμηση του στην κλάση MCI. Στην πρώτη γραμμή παρουσιάζονται οι αρχικές τιμές των χαρακτηριστικών και οι αντίστοιχες πιθανότητες πρόβλεψης για κάθε κλάση, ενώ στις υπόλοιπες τρεις γραμμές οι τιμές που έχουν επισημανθεί με πορτοκαλί υποδεικνύουν τις αλλαγές στις τιμές των χαρακτηριστικών και τις νέες πιθανότητες για κάθε κλάση.	89
Εικόνα 30: Αντιπαραθετικά παραδείγματα Permute Attack για ανατροπή της πρόβλεψης ενός MCI δείγματος και την ταξινόμηση του στην κλάση DEM. Στην πρώτη γραμμή παρουσιάζονται οι αρχικές τιμές των χαρακτηριστικών και οι αντίστοιχες πιθανότητες πρόβλεψης για κάθε κλάση, ενώ στις υπόλοιπες τρεις γραμμές οι τιμές που έχουν επισημανθεί με πορτοκαλί υποδεικνύουν τις αλλαγές στις τιμές των χαρακτηριστικών και τις νέες πιθανότητες για κάθε κλάση.	89
Εικόνα 31: Αντιπαραθετικά παραδείγματα Permute Attack για ανατροπή της πρόβλεψης ενός DEM δείγματος και την ταξινόμηση του στην κλάση CN. Στην πρώτη γραμμή παρουσιάζονται οι αρχικές τιμές των χαρακτηριστικών και οι αντίστοιχες πιθανότητες πρόβλεψης για κάθε κλάση, ενώ στις υπόλοιπες τρεις γραμμές οι τιμές που έχουν επισημανθεί με πορτοκαλί υποδεικνύουν τις αλλαγές στις τιμές των χαρακτηριστικών και τις νέες πιθανότητες για κάθε κλάση.	90
Εικόνα 32: Ραβδόγραμμα της συχνότητας με την οποία κάθε χαρακτηριστικό επιλέχθηκε να μεταβληθεί στα αντιπαραθετικά παραδείγματα, τα οποία δημιουργήθηκαν για το σύνολο ελέγχου της ταξινόμησης των κλάσεων CN και MCI. Τα χαρακτηριστικά έχουν τοποθετηθεί σε φθίνουσα σειρά συχνότητα. Μία πράσινη ράβδος αντιπροσωπεύει μία συνολικά θετική μεταβολή του εκάστοτε χαρακτηριστικού, ενώ μία κόκκινη, συνολικά αρνητική μεταβολή.....	91
Εικόνα 33: Ραβδόγραμμα της συχνότητας με την οποία κάθε χαρακτηριστικό επιλέχθηκε να μεταβληθεί στα αντιπαραθετικά παραδείγματα, τα οποία δημιουργήθηκαν για το σύνολο ελέγχου της ταξινόμησης των κλάσεων MCI και DEM. Τα χαρακτηριστικά έχουν τοποθετηθεί σε φθίνουσα σειρά συχνότητα. Μία πράσινη ράβδος αντιπροσωπεύει μία συνολικά θετική μεταβολή του εκάστοτε χαρακτηριστικού, ενώ μία κόκκινη, συνολικά αρνητική μεταβολή.....	92

Κατάλογος Πινάκων

Πίνακας 1 - Κατανομή των κλάσεων ανά ηλικιακή ομάδα.....	64
Πίνακας 2: Σύγκριση αποτελεσμάτων των μετρικών του σταθμισμένου F1 σκορ και της ισορροπημένης ακρίβειας, κατά την ταξινόμηση με τα μοντέλα Τυχαίων Δασών, Μηχανών Διανυσμάτων Υποστήριξης, Gradient Boosting, Λογιστικής Παλινδρόμησης, Πολυστρωματικού Perceptron και XGBoost ανάμεσα στις τεχνικές Ένα-προς-Όλα (OVA), Ένα-προς-Ένα (OVO) και Bagging με OVO με διαχωρισμό σε σύνολα εκπαίδευσης και ελέγχου (TTS), καθώς και μετά τον προσδιορισμό των βέλτιστων υπερπαραμέτρων με διασταυρούμενη επικύρωση (CV)	75
Πίνακας 3: Συμμετρικός πίνακας των p-value που προέκυψαν ανά ζεύγος ταξινομητών από το στατιστικό t-test για τις πέντε τιμές της ισορροπημένης ακρίβειας από τη διασταυρούμενη επικύρωση για τους έξι ταξινομητές.	76
Πίνακας 4: Σύγκριση αποτελεσμάτων των μετρικών του σταθμισμένου F1 σκορ και της ισορροπημένης ακρίβειας χρησιμοποιώντας μόνο τα SNP χαρακτηριστικά, μόνο τα VOI χαρακτηριστικά και τον συνδυασμό των δύο ειδών χαρακτηριστικών κατά την ταξινόμηση με το μοντέλο των Τυχαίων Δασών.....	77

1. Εισαγωγή

Πάνω από 50 εκατομμύρια άνθρωποι παγκοσμίως πάσχουν από κάποια μορφή άνοιας, ενώ το νούμερο αυτό αναμένεται να αυξηθεί στα 152 εκατομμύρια έως το 2050 [1]. Η γνωστική δυσλειτουργία οφείλεται σε ποικίλους παράγοντες. Η Ήπιας Μορφής Νοητική Εξασθένηση (Mild Cognitive Impairment - MCI) και η άνοια αποτελούν τις πιο αντιπροσωπευτικές νευροεκφυλιστικές διαταραχές της νοητικής λειτουργίας [2]. Μέχρι στιγμής, δεν υπάρχει αποτελεσματική θεραπεία που να επιβραδύνει ή να σταματάει την εξέλιξη της άνοιας. Επομένως, είναι καίριας σημασίας, η ιατρική επιστήμη να εστιάσει στα πρώιμα στάδια της νόσου, ώστε να υπάρχει έγκαιρη παρέμβαση και καθυστέρηση της εξέλιξής της. Η κλινική διάγνωση της άνοιας βασίζεται στο λεπτομερές ιατρικό ιστορικό που παρέχουν οι ίδιοι οι ασθενείς ή οι οικείοι τους, καθώς και στις νευρολογικές και νευροψυχολογικές εξετάσεις. Άλλες εξετάσεις, όπως αιματολογικές ή απεικονιστικές πραγματοποιούνται προκειμένου να αποκλειστούν άλλα αίτια που μπορεί να προκαλούν την άνοια.

Η γήρανση του εγκεφάλου σχετίζεται με πολύπλοκες αλλαγές στη δομή και τη λειτουργία του, όπως μοτίβα ατροφίας της φαιάς ουσίας σε ορισμένες περιοχές του εγκεφάλου και αγγειακές αλλαγές στη λευκή ουσία. Υπάρχουν ενδείξεις της ασθένειας που μπορούν να αναγνωριστούν κατά τη διάρκεια μίας απεικονιστικής εξέτασης, όπως η μαγνητική τομογραφία (magnetic resonance imaging -MRI) και η τομογραφία εκπομπής ποζιτρονίου (positron emission tomography - PET) και οι οποίες οδηγούν στη διάγνωση της άνοιας. Τα συνηθέστερα μακροσκοπικά χαρακτηριστικά της ατροφίας του εγκεφάλου περιλαμβάνουν μεγέθυνση των αυλακώσεων του μετωπικού και κροταφικού φλοιού και την ατροφία των ελικών. Η ατροφία επηρεάζει επίσης το προσφηνοειδές λόβιο και την έλικα του προσαγωγίου, με αποτέλεσμα τη διεύρυνση των πλευρικών κοιλιών και την ατροφία της αμυγδαλής και του ιπποκάμπου [3].

Η ατροφία του εγκεφάλου είναι ενδεικτική νευροεκφυλιστικών διαταραχών, όπως η άνοια, η οποία χαρακτηρίζεται από δυσλειτουργία της μνήμης και νοητική εξασθένηση που μπορεί να επηρεάσει τη συμπεριφορά του ατόμου, την ικανότητα ομιλίας του και το κινητικό του σύστημα, όμως πριν την εμφάνιση των συμπτωμάτων αυτών, ο ασθενής είναι ασυμπτωματικός για μεγάλο χρονικό διάστημα. Η απουσία διαγνωστικών εργαλείων για την έγκαιρη αναγνώριση των νευροεκφυλιστικών παθήσεων σε πρώιμα στάδια, προκειμένου να χορηγηθεί η κατάλληλη θεραπεία που μπορεί να αναχαιτίσει την εξέλιξη των νόσων, αποτελεί σημαντικό πρόβλημα [3], επομένως, είναι απαραίτητη η ανάπτυξη αποτελεσματικότερων διαγνωστικών εργαλείων.

Αυτή τη στιγμή η τεχνητή νοημοσύνη έχει αποκτήσει εξαιρετικά σημαντικό ρόλο ως εργαλείο για την αντιμετώπιση πραγματικών προβλημάτων σε διάφορους τομείς, συμπεριλαμβανομένου του τομέα της υγείας [4]. Η εφαρμογή της μηχανικής μάθησης στον τομέα της υγείας μπορεί να προσφέρει αποτελεσματικούς τρόπους για την αξιοποίηση των πολύπλοκων ιατρικών δεδομένων, ώστε οι διαγνώσεις να πραγματοποιούνται με ακρίβεια [5]. Τα μοντέλα ταξινόμησης έχουν δύο βασικούς στόχους, την υψηλή επίδοση, δηλαδή να μπορούν με όσο μεγαλύτερη ακρίβεια γίνεται, να προβλέψουν την έξοδο για νέα δεδομένα εισόδου και την ερμηνευσιμότητα, δηλαδή την ικανότητα να παρέχουν επεξηγήσεις για τη συσχέτιση ανάμεσα στα χαρακτηριστικά εισόδου και στην έξοδο. Συνήθως, υπάρχει κάποια αντιστάθμιση ανάμεσα στους δύο αυτούς στόχους, καθώς τα απλούστερα μοντέλα είναι εύκολο να ερμηνευθούν αλλά έχουν χαμηλότερη επίδοση από τα πολυπλοκότερα, τα οποία αντίθετα έχουν υψηλότερη προβλεπτική ικανότητα, αλλά είναι λιγότερο ερμηνεύσιμα. [6]. Στην προσπάθεια, λοιπόν, να αυξηθεί η προβλεπτική ικανότητα των αλγορίθμων μηχανικής μάθησης, έχει αυξηθεί και η πολυπλοκότητα των μοντέλων αυτών, γεγονός που έχει οδηγήσει σε επιπλέον δυσκολία κατά την ερμηνεία των προβλέψεων. Το γεγονός αυτό αποτελεί σημαντικό πρόβλημα, ιδιαίτερα στον τομέα της υγείας, όπου η ανάγκη για αξιόπιστα, εύρωστα, αποτελεσματικά και κατ' επέκταση επεξηγήσιμα ή ερμηνεύσιμα μοντέλα είναι ιδιαίτερα αυξημένη. Είναι σημαντικό, όχι μόνο να παράγουν προβλέψεις σχετικά με το κλινικό αποτέλεσμα για τον ασθενή, αλλά και να λαμβάνουν

υπόψιν τους τα χαρακτηριστικά του ασθενούς με έναν επεξηγήσιμο και μετρήσιμο τρόπο, όπως επίσης και να παρέχουν εφαρμόσιμες συστάσεις για την πρόληψη β].

Η συγκεκριμένη μελέτη έχει ως στόχο την ανάπτυξη ενός μοντέλου τεχνητής νοημοσύνης για τον διαχωρισμό ατόμων με φυσιολογική γήρανση του εγκεφάλου, με Ήπιας Μορφής Νοητική Εξασθένηση (MCI) και ασθενών με άνοια. Η προτεινόμενη μέθοδος χρησιμοποιεί αριθμητικά δεδομένα ογκομετρικών μετρήσεων των ανατομικών περιοχών του εγκεφάλου που έχουν προκύψει από εικόνες μαγνητικής τομογραφίας εγκεφάλου, καθώς και γενετικά δεδομένα που αφορούν σε πολυμορφισμούς ενός νουκλεοτιδίου σχετιζόμενους με τη νόσο Αλτσχάιμερ. Έχει αναπτυχθεί μία μέθοδος συλλογικής μάθησης, στην οποία εκπαιδεύονται πανομοιότυποι ταξινομητές σε δύο ισορροπημένα σύνολα δεδομένων που προέρχονται από το αρχικό μη ισορροπημένο σύνολο και οι προβλέψεις τους συναθροίζονται για να παραχθεί η τελική πρόβλεψη. Στόχος την συγκεκριμένης μεθόδου είναι η αύξηση της προβλεπτικής ικανότητας του μοντέλου, πρώτον αντιμετωπίζοντας την αρχική ανισορροπία των κλάσεων στο σύνολο δεδομένων και δεύτερον, ελαττώνοντας τη διακύμανση των προβλέψεων. Επιπλέον, η μελέτη εστιάζει στη χρήση μεθόδων ερμηνευσιμότητας των μοντέλων ταξινόμησης, ώστε να αποκτηθούν φιλικές προς τον άνθρωπο επεξηγήσεις σχετικά με τις προβλέψεις των μοντέλων και να αυξηθεί η αξιοπιστία τους, η οποία είναι ιδιαίτερα σημαντική στον τομέα της υγείας.

2. Γήρανση Εγκεφάλου και Σχετιζόμενες Ασθένειες

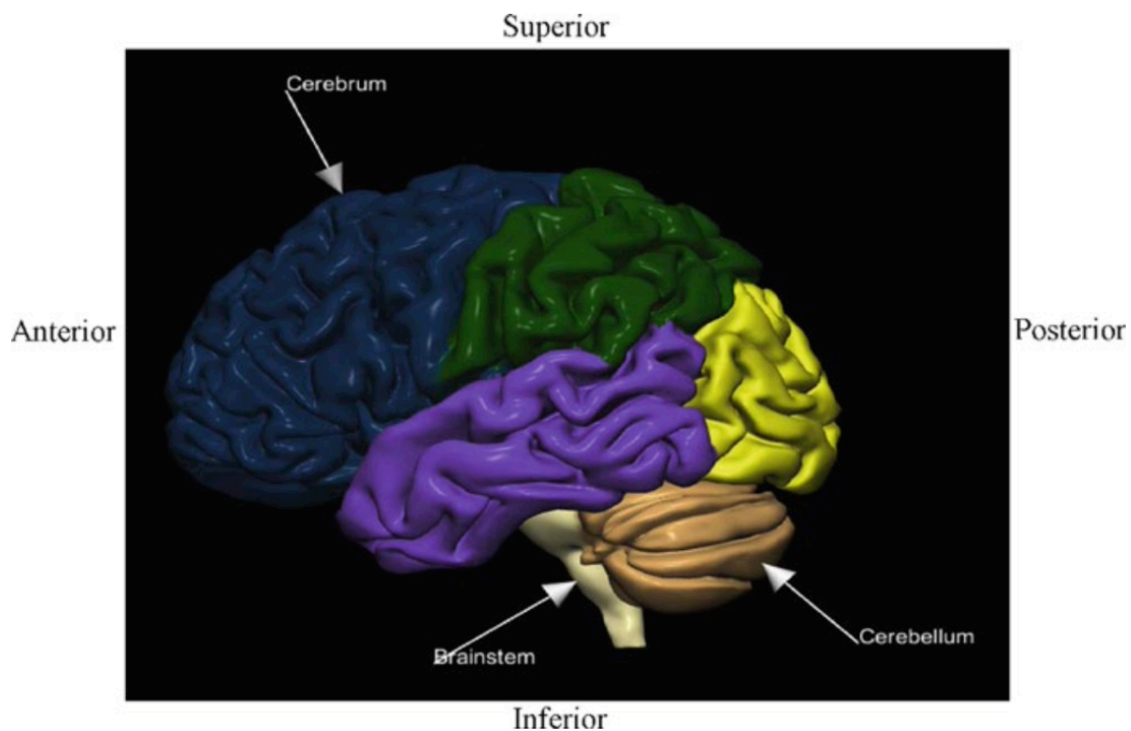
Στο παρόν κεφάλαιο, θα παρουσιαστούν συνοπτικά οι ανατομικές περιοχές του ανθρώπινου εγκεφάλου, καθώς και οι πολυμορφισμοί ενός νουκλεοτιδίου (SNP) που αποτελούν τα χαρακτηριστικά από τα οποία απαρτίζεται το σύνολο δεδομένων της συγκεκριμένης μελέτης. Στη συνέχεια, θα περιγραφούν τα βασικά γνωρίσματα και η παθολογία της φυσιολογικής γήρανσης του εγκεφάλου, καθώς και ασθενειών που σχετίζονται με τη γήρανση, όπως η Ήπιας Μορφής Νοητική Εξασθένηση (Mild Cognitive Impairment - MCI) και η άνοια.

2.1 Ανατομικές Περιοχές του Ανθρώπινου Εγκεφάλου

Το κεντρικό νευρικό σύστημα (ΚΝΣ) του ανθρώπου, ύστερα από 600 εκατομμύρια χρόνια εξέλιξης, αποτελεί το πιο σύνθετο όργανο ζωντανού οργανισμού και έχει μελετηθεί εκτενώς ανά τους αιώνες [7].

- Δομική νευροανατομία του εγκεφάλου:

Το κεντρικό νευρικό σύστημα αποτελείται από τον εγκέφαλο και τον νωτιαίο μυελό. Ο εγκέφαλος περικλείει το σύστημα των κοιλιών και χωρίζεται σε τρία κύρια μέρη: τον εγκέφαλο (cerebrum), την παρεγκεφαλίδα (cerebellum) και το στέλεχος του εγκεφάλου (brainstem).



Εικόνα 1: Τα ανατομικά μέρη του εγκεφάλου (αριστερό ημισφαίριο). Διακρίνεται ο εγκέφαλος (cerebrum), η παρεγκεφαλίδα (cerebellum) και το στέλεχος του εγκεφάλου (brainstem).

Ο εγκέφαλος αποτελείται από: το αριστερό και δεξιό εγκεφαλικό ημισφαίριο (cerebral hemispheres), τον διάμεσο εγκέφαλο (diencephalon) μεταξύ του εγκεφάλου και του εγκεφαλικού στελέχους και εν τω βάθει δομές (deep gray nuclei) [7].

- Εγκέφαλος

- Εγκεφαλικά ημισφαίρια

Τα εγκεφαλικά ημισφαίρια συνιστούν το μεγαλύτερο τμήμα του εγκεφάλου και συνδέονται μεταξύ τους με ίνες λευκής ουσίας. Αποτελούνται από: εξωτερική φαιά ουσία που ονομάζεται εγκεφαλικός φλοιός (cerebral cortex) και εσωτερική λευκή ουσία που περιλαμβάνει τις εν τω βάθει δομές.

Η φαιά ουσία περιλαμβάνει κυρίως σώματα νευρικών κυττάρων, ενώ η λευκή ουσία αποτελείται κυρίως από νευράξονες. Ο εγκεφαλικός φλοιός είναι εξαιρετικά δαιδαλώδης. Οι πτυχές δημιουργούν έλικες (gyri), οι οποίες χωρίζονται από αυλακώσεις που ονομάζονται αύλακες (sulci) ή σχισμές (fissures), δηλαδή βαθύτερες αύλακες. Τα εγκεφαλικά ημισφαίρια χωρίζονται σε πέντε επιμέρους λοβούς: τον μετωπιαίο (frontal lobe), τον κροταφικό (temporal lobe), τον βρεγματικό (parietal lobe), τον ινιακό (occipital lobe) και τον λοβό μεταιχμιακού συστήματος (limbic lobe).

Οι λοβοί οριοθετούνται εν μέρει από τις αύλακες/σχισμές. Η κεντρική αύλακα (central sulcus) χωρίζει τον πρόσθιο μετωπιαίο λοβό (frontal lobe anterior) από τον οπίσθιο βρεγματικό λοβό (parietal lobe posterior). Η πλευρική σχισμή (sylvian or lateral fissure) διαχωρίζει τον κροταφικό λοβό από τον μετωπιαίο και βρεγματικό λοβό. Η βρεγματο-ινιακή σχισμή (parieto-occipital fissure) οριοθετεί τον πρόσθιο βρεγματικό λοβό (parietal lobe anterior) από τον οπίσθιο ινιακό λοβό (occipital lobe posterior). Τέλος, η αύλακα προσαγωγίου (cingulate sulcus) διαχωρίζει τον μετωπιαίο λοβό από τον λοβό μεταιχμιακού συστήματος [7].

- Διάμεσος εγκέφαλος

Ο διάμεσος εγκέφαλος περιλαμβάνει τον θάλαμο (thalamus), την επιθαλάμια χώρα (subthalamus) και τον υποθάλαμο (hypothalamus) [7].

- Παρεγκεφαλίδα

Η παρεγκεφαλίδα αποτελείται από το αριστερό και το δεξιό παρεγκεφαλιδικό ημισφαίριο (cerebellar hemispheres) και το μεσαίο στέλεχος (midline vermis) που τα ενώνει [7].

- Στέλεχος εγκεφάλου

Το εγκεφαλικό στέλεχος διαιρείται στον μέσο εγκέφαλο (midbrain), τη γέφυρα (pons) και τον μυελό (medulla) [7].

- Δομική νευροανατομία των περιοχών του φλοιού:

Ο φλοιός του εγκεφάλου αποτελείται από τρεις επιφάνειες: την πλευρική (lateral), τη μεσαία (medial) και την κατώτερη (inferior), η οποία ονομάζεται επίσης κοιλιακή (ventral). Επιπλέον οι μεταβατικές περιοχές σχηματίζουν τον μετωπιαίο, τον κροταφικό και τον ινιακό πόλο.

- Πλευρική επιφάνεια

Στην πλευρική επιφάνεια του φλοιού εμφανίζονται τέσσερις λοβοί, ο μετωπιαίος, ο κροταφικός, ο βρεγματικός και ο ινιακός.

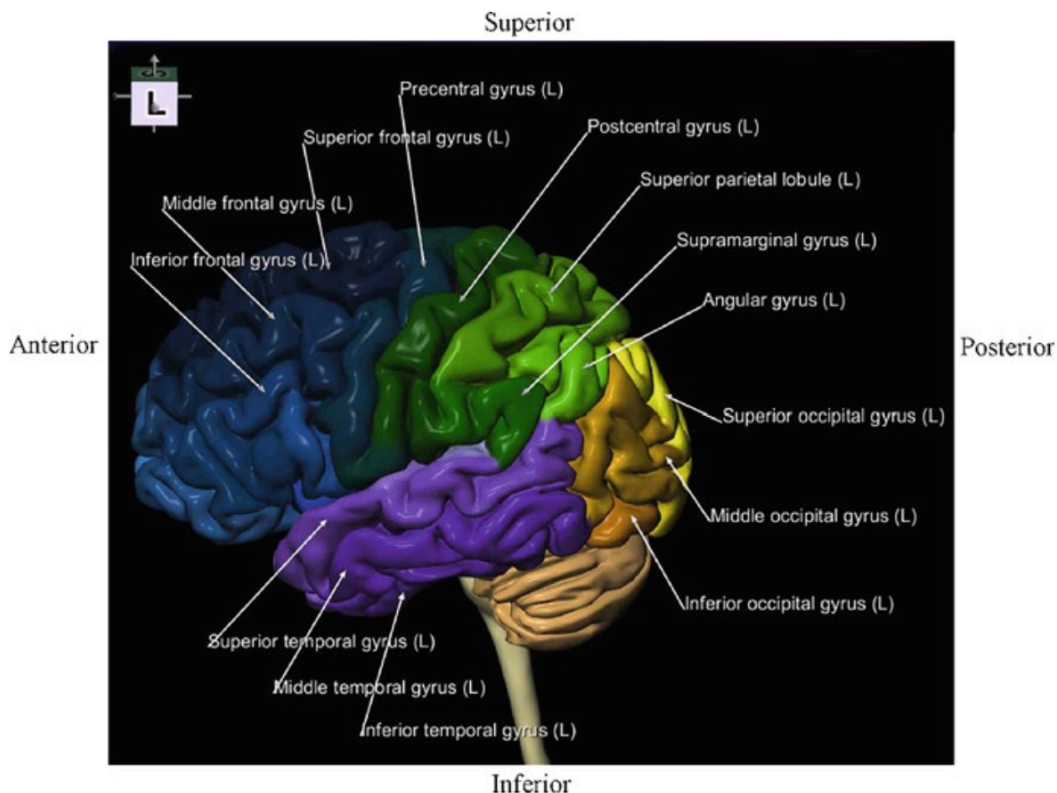
Η πλευρική επιφάνεια του μετωπιαίου λοβού διαχωρίζεται από τρεις αύλακες (την ανώτερη μετωπιαία, την κατώτερη μετωπιαία και την προκεντρική αύλακα) σε τέσσερις έλικες: την ανώτερη μετωπιαία (superior frontal gyrus), τη μέση μετωπιαία (middle frontal gyrus), την κατώτερη μετωπιαία (inferior frontal gyrus) και την προκεντρική έλικα (precentral gyrus).

Η πλευρική επιφάνεια του κροταφικού λοβού χωρίζεται από δύο αύλακες (την ανώτερη κροταφική και την κατώτερη κροταφική αύλακα) σε τρεις έλικες: την ανώτερη κροταφική (superior temporal gyrus), τη μέση κροταφική (middle temporal gyrus) και την κατώτερη κροταφική έλικα (inferior temporal gyrus).

Η πλευρική επιφάνεια του βρεγματικού λοβού διαιρείται από την ενδοβρεγματική αύλακα (intraparietal sulcus) σε τρεις έλικες: τη μετακεντρική (postcentral gyrus), την ανώτερη βρεγματική (superior parietal gyrus) και την κάτω βρεγματική έλικα (inferior parietal gyrus). Τέλος, η πλευρική επιφάνεια του ινιακού λοβού, χωρίζεται από δύο αύλακες (την ανώτερη ινιακή και την κατώτερη ινιακή αύλακα) σε τρεις έλικες: την ανώτερη ινιακή (superior occipital gyrus), τη μέση ινιακή (middle occipital gyrus) και την κατώτερη ινιακή έλικα (inferior occipital gyrus) [7].

- Μεσαία επιφάνεια

Οι λοβοί της μεσαίας επιφάνειας του φλοιού είναι ο μετωπιαίος, ο βρεγματικός, ο ινιακός και ο λοβός του μεταιχμιακού συστήματος. Ο τελευταίος περιλαμβάνει τις έλικες που βρίσκονται στην εσωτερική άκρη του ημισφαιρίου, δηλαδή την υπομεσολόβια έλικα (subcallosal gyrus), την έλικα του προσαγωγίου (cingulate gyrus), τον ισθμό της έλικας του προσαγωγίου (isthmus of cingulate gyrus) και την έλικα του παραίπποκάμπου (parahippocampal gyrus). Η ανώτερη μετωπιαία έλικα (superior frontal gyrus) καταλαμβάνει το μεγαλύτερο μέρος της μεσαίας επιφάνειας του μετωπιαίου λοβού. Ο βρεγματικός λοβός περιλαμβάνει το προσφηνοειδές λόβιο (precuneus), ενώ ο ινιακός λοβός αποτελείται από το σφηνοειδές λόβιο (cuneus) και τη γλωσσική έλικα (lingual gyrus) [7].



Εικόνα 2: Πλευρική όψη των περιοχών του εγκεφαλικού φλοιού (αριστερό ημισφαίριο). Σε κάθε έλικα έχει αποδοθεί ξεχωριστό χρώμα.

- Κατώτερη επιφάνεια

Η κατώτερη επιφάνεια του φλοιού περιλαμβάνει τον μετωπιαίο, τον κροταφικό και τον ινιακό λοβό.

Στην επιφάνεια αυτή ο μετωπιαίος λοβός αποτελείται από την ευθεία έλικα (straight gyrus), τις τροχιακές έλικες (orbital gyri), οι οποίες χωρίζονται στην πρόσθια, μέση, πλευρική και οπίσθια τροχιακή έλικα.

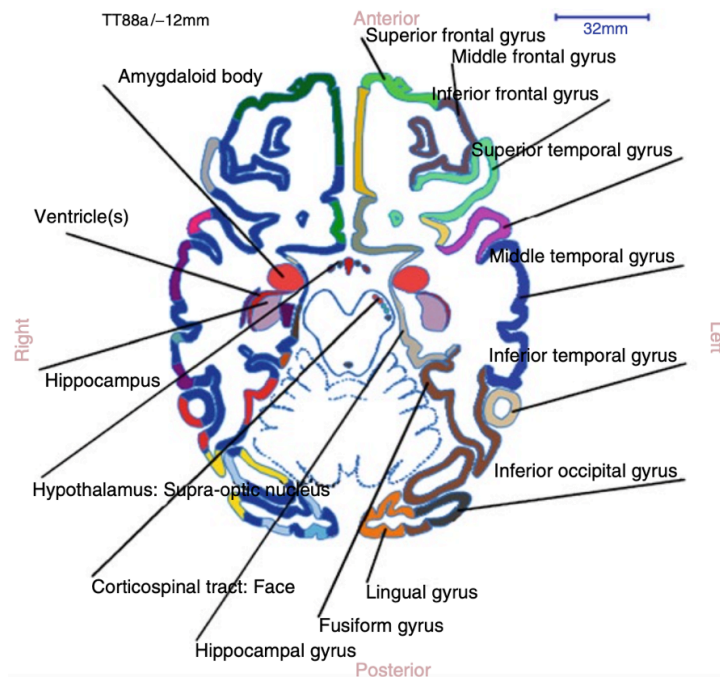
Ο κροταφικός και ινιακός λοβός χωρίζονται από δύο αύλακες (την πλευρική και τη μέση ινιακο-κροταφική αύλακα) σε τρεις έλικες: τη μέση ινιακο-κροταφική έλικα (medial occipitotemporal gyrus), της οποίας το κροταφικό τμήμα αποτελείται από την έλικα του παραίπποκάμπου και το ινιακό τμήμα της γλωσσικής έλικας, την πλευρική ινιακο-κροταφική έλικα (lateral occipitotemporal gyrus) και την κατώτερη κροταφική έλικα (inferior temporal gyrus) [7].

- Δομική νευροανατομία των εν τω βάθει δομών:

Οι εν τω βάθει δομές είναι δομές φαιάς ουσίας με κυριότερες: τα βασικά γάγγλια (basal ganglia), τον θάλαμο (thalamus), τον ιππόκαμπο (hippocampus) και την αμυγδαλή (amygdala). Τα βασικά γάγγλια είναι ο κερκοφόρος πυρήνας (caudate nucleus) και οι φακοειδείς πυρήνες (lentiform nuclei) που αποτελούνται από το κέλυφος (putamen) και την ωχρή σφαίρα (globus pallidus) [7].

- Δομική νευροανατομία του κοιλιακού συστήματος

Το κοιλιακό σύστημα περιλαμβάνει τέσσερις διασυνδεδεμένες εγκεφαλικές κοιλίες (κοιλότητες) που περιλαμβάνουν το εγκεφαλονωτιαίο υγρό. Αυτές είναι η αριστερή και δεξιά πλευρική κοιλία (left and right lateral ventricle), η τρίτη (third ventricle) και η τέταρτη κοιλία (fourth ventricle).



Εικόνα 3: Επίπεδη νευροανατομία σε αξονικό προσανατολισμό -12mm κάτω από το επίπεδο του πρόσθιου συνδέσμου (anterior commissure) που συνδέει τα δύο ημισφαίρια.

Το εγκεφαλονωτιαίο υγρό εκκρίνεται κυρίως στο χοριοειδές πλέγμα (choroid plexus) (ένα δίκτυο αγγείων) και κυκλοφορεί από τις πλευρικές κοιλίες (οι οποίες είναι και οι μεγαλύτερες) μέσω των μεσοκοιλιακών τρημάτων (interventricular foramina) προς την τρίτη κοιλία και στη συνέχεια προς την τέταρτη κοιλία [7].

2.2 Πολυμορφισμοί Ενός Νουκλεοτιδίου

Θεωρείται πως τα γονιδιώματα μεταξύ δύο τυχαία επιλεγμένων ατόμων περιλαμβάνουν περίπου 0.1% διαφοροποιήσεις. Οι διαφοροποιήσεις αυτές καλείται πολυμορφισμοί και προκαλείται λόγω μεταλλάξεων. Η απλούστερη μορφή διαφοροποίησης του DNA ανάμεσα σε άτομα είναι η αντικατάσταση ενός νουκλεοτιδίου με ένα άλλο. Αυτού του είδους η αλλαγή ονομάζεται πολυμορφισμός του ενός νουκλεοτιδίου (single nucleotide polymorphism - SNP) και φαίνεται να είναι συνηθέστερη από άλλα είδη πολυμορφισμών [8].

Εκτιμάται πως τα SNP συναντώνται με συχνότητα περίπου 1 στα 1000 ζεύγη βάσεων σε όλο το γονιδίωμα και περισσότερα από 3 εκατομμύρια SNP έχουν καταγραφεί μέχρι στιγμής. Αυτές οι μικρές διαφοροποιήσεις θεωρείται πως είναι σταθερές και όχι επιβλαβείς για τους οργανισμούς. Το 50% των SNP εμφανίζεται σε μη κωδικοποιούσες περιοχές, το 25% προκαλεί παρανοηματικές (missense) μεταλλάξεις και το υπόλοιπο 25% σιωπηλές (silent) μεταλλάξεις. Τα σιωπηλά αυτά SNP καλούνται επίσης “συνώνυμα”, καθώς δεν αλλάζουν τα κωδικοποιημένα αμινοξέα. Παλιότερα δεν θεωρούνταν πως έχουν επίδραση στη γονιδιακή λειτουργία και τον φαινότυπο, ωστόσο, υπάρχουν πλέον ενδείξεις πως κάτι τέτοιο δεν ισχύει. Τα μη συνώνυμα SNP, τα οποία μεταβάλλουν τα αμινοξέα, μπορούν να επηρεάσουν τη γονιδιακή έκφραση, καθώς και τη σταθερότητα και τη μεταφραστική αποδοτικότητα του αγγελιαφόρου RNA (mRNA). Επομένως, μπορεί να είναι υπεύθυνα για την ευαισθησία του ατόμου στην ανάπτυξη ασθενειών [8].

Ο εντοπισμός γονιδίων που συνδέονται με νευροεκφυλιστικές ασθένειες σχετιζόμενες με τη γήρανση, έχει οδηγήσει σε πολύτιμη γνώση που αφορά στην παθογένεια των νόσων αυτών [9]. Πλήθος ερευνών ανάλυσης συσχέτισης γονιδιώματος (genome-wide association analysis - GWAS) έχουν εστιάσει στον εντοπισμό πολυμορφισμών του ενός νουκλεοτιδίου που καθιστούν τους ανθρώπους επιρρεπείς στην ανάπτυξη νευροεκφυλιστικών ασθενειών, όπως για παράδειγμα η νόσος Αλτσχάιμερ ή σχετίζονται με μακροπρόθεσμες νοητικές μεταβολές, ανάλογα με το πλήθος των άλλων (0, 1 ή 2) που διαθέτουν [10]. Οι μελέτες αυτές αφορούν την ισχύ της στατιστικής συσχέτισης ορισμένων πολυμορφισμών του ενός νουκλεοτιδίου με τη διάγνωση νευροεκφυλιστικών ασθενειών [11]. Παρόλο που τα SNP δεν έχουν απαραίτητα άμεση σχέση με έναν συγκεκριμένο φαινότυπο, αποτελούν συχνά δείκτες που σχετίζονται με τον φαινότυπο [12].

2.3 Φυσιολογική Γήρανση εγκεφάλου

Η γήρανση αποτελεί μία σύνθετη βιολογική διαδικασία, η οποία χαρακτηρίζεται από συσσωρευση μοριακών και κυτταρικών φθορών κατά τη διάρκεια της ζωής ενός ατόμου. Η αδυναμία του σώματος να αποκαταστήσει τις φθορές αυτές οδηγεί στην απώλεια ορισμένων σωματικών λειτουργιών, όπως οι αισθητηριακές, οι κινητικές και οι γνωστικές, γεγονός που συντείνει στην επιδείνωση της ποιότητας ζωής του ατόμου. Η γήρανση συνιστά, επίσης, έναν βασικό παράγοντα κινδύνου για πλήθος ασθενειών, συμπεριλαμβανομένου του καρκίνου, των καρδιαγγειακών νόσων και των νευροεκφυλιστικών ασθενειών [13].

Ο εγκέφαλος είναι ιδιαίτερα ευαίσθητος στην επίδραση της γήρανσης, καθώς αυτή εκδηλώνεται με αλλαγές στη δομή και τις νοητικές λειτουργίες. Οι συχνότερες αλλαγές που σχετίζονται με τη γήρανση είναι η ατροφία του εγκεφάλου (δηλαδή η ελάττωση του όγκου της φαιάς ουσίας και λέπτυνση του φλοιού), η μείωση της ακεραιότητας και του

όγκου της λευκής ουσίας και μη φυσιολογική λειτουργική συνδεσιμότητα [13]. Ο εγκέφαλος των ατόμων με φυσιολογικές νοητικές λειτουργίες παρουσιάζει αλλαγές σχετιζόμενες με τη γήρανση, οι οποίες περιλαμβάνουν μία συνολική ελάττωση στον όγκο και το βάρος του εγκεφάλου και διεύρυνση των κοιλιών. Οι αλλαγές αυτές αποτελούν μερικώς αποτέλεσμα της απώλειας νευρικών κυττάρων. Επιπλέον, παρατηρείται απώλεια συνάψεων και δενδριτών, η οποία δεν είναι συνολική, αλλά εντοπίζεται σε συγκεκριμένες περιοχές [14]. Όταν οι αλλαγές αυτές είναι ισχυρές, μπορούν να θεωρηθούν ως ένδειξη επιταχυνόμενης γήρανσης ή εξέλιξης κάποιας υποκείμενης νόσου [13].

- Ανατομικές αλλαγές και γενετικοί παράγοντες της φυσιολογικής γήρανσης του εγκεφάλου:

Η απώλεια του όγκου συνοδεύεται από αύξηση του όγκου των κοιλιών και άλλων περιοχών του εγκεφαλονωτιαίου υγρού. Οι περιοχές που επηρεάζονται περισσότερο είναι ο ιππόκαμπος και οι μετωπιαίοι λοβοί. Στο διάστημα ηλικιών από 30 έως 90 έτη η απώλεια όγκων είναι 14% στον εγκεφαλικό φλοιό, 35% στον ιππόκαμπο και 26% στη λευκή ουσία, ενώ υπάρχει σχετικά μεγαλύτερη απώλεια λευκής ουσίας μετά την ηλικία των 50 ετών [14].

Έχει επίσης αποδειχθεί πως η φυσιολογική γήρανση του εγκεφάλου σχετίζεται επίσης με απώλεια νευρώνων κατά 10% στον φλοιό του εγκεφάλου. Όσο προχωράει η γήρανση, οι νευρώνες σε ορισμένες περιοχές του εγκεφάλου χάνονται, συμπεριλαμβανομένων εκείνων του ιππόκαμπου, του εγκεφαλικού φλοιού και της αμυγδαλής. Άλλοι νευρώνες παρουσιάζουν συρρίκνωση χωρίς απώλεια, ενώ σε άλλες περιοχές δεν εμφανίζεται συρρίκνωση ή απώλεια νευρώνων [14].

Έχει παρατηρηθεί απώλεια συνάψεων έως και κατά 20% σε ορισμένες περιοχές του φλοιού. Ωστόσο, υπάρχουν στοιχεία που υποδεικνύουν πως η πυκνότητα των συνάψεων στον ιππόκαμπο δεν ελαττώνεται με τη φυσιολογική γήρανση σε περιοχές που λαμβάνουν ερεθίσματα από νευρώνες της στιβάδας II του ενδοκρινικού φλοιού, όπου παρατηρείται απώλεια νευρώνων. Έχει επίσης διαπιστωθεί πως οι συνάψεις στο έλασμα III και V του ανθρώπινου μετωπιαίου φλοιού δεν παρουσιάζουν αλλαγές που σχετίζονται με την φυσιολογική γήρανση [14].

Επιπλέον, υπάρχουν ενδείξεις πως γενετικοί παράγοντες επηρεάζουν τη γήρανση του εγκεφάλου, με χαρακτηριστικότερο παράδειγμα την εμφανή συνεισφορά του γονιδίου ApoE στη σχετιζόμενη με τη γήρανση εξασθένηση των νοητικών λειτουργιών [14].

2.4 Ασθένειες Σχετιζόμενες με τη Γήρανση του Εγκεφάλου

Καθώς το προσδόκιμο ζωής αυξάνεται, ολοένα ενισχύεται η προσπάθεια για βελτίωση της ποιότητας ζωής των ανθρώπων μεγάλης ηλικίας. Μία πολύ σημαντική πλευρά της προσπάθειας αυτής είναι ο πιο έγκαιρος εντοπισμός της πιθανής νοητικής εξασθένησης των ατόμων αυτών, ώστε να εκτελεστούν άμεσα στοχευμένες θεραπευτικές ενέργειες [15]. Πλήθος προσπαθειών έχουν πραγματοποιηθεί προκειμένου να οριστεί η εξασθένηση των νοητικών ικανοτήτων που σχετίζεται με τη γήρανση [16].

2.4.1 Ήπιας Μορφής Νοητική Εξασθένηση

Η έννοια της Ήπιας Μορφής Νοητικής Εξασθένησης (Mild Cognitive Impairment - MCI) έχει εξελιχθεί, ώστε να αντιπροσωπεύει την ενδιάμεση κατάσταση συμπτωματολογίας ανάμεσα στις αλλαγές της γνωστικής λειτουργίας λόγω της φυσιολογικής γήρανσης και στην άνοια [15]. Τα συμπτώματα του MCI αφορούν γνωστική εξασθένηση μεγαλύτερη από την αναμενόμενη σύμφωνα με την ηλικία του ατόμου, η οποία όμως δεν επηρεάζει σημαντικά τις δραστηριότητες της καθημερινότητας. Επομένως, αποτελεί διακριτή νόσο από την άνοια, της οποίας η νοητική εξασθένηση είναι πιο σοβαρή και εκτεταμένη, με

αποτέλεσμα να επηρεάζει σημαντικά την καθημερινή λειτουργικότητα του ατόμου. Ωστόσο, όταν το MCI συνοδεύεται από ενοχλήσεις σχετικά με ελλείμματα της μνήμης (γνωστό και ως αμνησιακό MCI), όπως έχει συστηματικά φανεί, υπάρχει μεγαλύτερη πιθανότητα εξέλιξης σε άνοια, και ειδικότερα άνοια της νόσου Αλτσχάιμερ [16].

Ορισμένοι ασθενείς με MCI φαίνεται να παραμένουν σταθεροί ή να επανέρχονται στη φυσιολογική κατάσταση με την πάροδο του χρόνου, όμως περισσότεροι από τους μισούς παρουσιάζουν άνοια μέσα σε διάστημα πέντε ετών. Η Ήπιας Μορφής Νοητική Εξασθένηση μπορεί, επομένως, να θεωρηθεί ως μία κατάσταση επικινδυνότητας για την ανάπτυξη άνοιας και κατ' επέκταση ο εντοπισμός του MCI μπορεί να οδηγήσει σε έμμεση πρόληψη, μέσω του ελέγχου των σχετικών παραγόντων κινδύνου, όπως η συστολική υπέρταση [16].

- Διάγνωση:

Τα βασικότερα κριτήρια που έχουν οριστεί για την αναγνώριση του MCI είναι 1) το άτομο να μην βρίσκεται ούτε στη φυσιολογική νοητική κατάσταση, αλλά ούτε να πάσχει από άνοια, 2) να υπάρχουν στοιχεία επιδείνωσης των νοητικών ικανοτήτων του, τα οποία να έχουν αντληθεί είτε από αντικειμενικά μετρούμενη εξασθένηση με την πάροδο του χρόνου, είτε από αναφορά της εξασθένησης από τον ίδιο τον ασθενή ή πρόσωπο του περιβάλλοντός του, σε συνδυασμό με αντικειμενική δυσλειτουργία των νοητικών ικανοτήτων και 3) να πραγματοποιούνται σταθερά χωρίς δυσκολία οι καθημερινές δραστηριότητες του ατόμου, και οι σύνθετες λειτουργίες να έχουν επηρεαστεί είτε καθόλου, είτε σε πολύ μικρό βαθμό. Τα κριτήρια αυτά, στοχεύουν στην επέκταση της έννοιας του MCI, ώστε να περιλαμβάνει γνωστικούς τομείς εκτός της μνήμης και να θεωρείται πρόδρομος των διάφορων τύπων της άνοιας [16].

Η διαδικασία της διάγνωσης, είθισται να ξεκινά με τον ασθενή ή κάποιο κοντινό του πρόσωπο να εκφράζει την ανησυχία για τις νοητικές ικανότητες του ασθενούς. Όταν οι ανησυχίες αυτές παρουσιάζονται στον κλινικό ιατρό, πρέπει πρώτα να καθοριστεί εάν η νοητική κατάσταση του ατόμου είναι φυσιολογική, ή υπάρχει υπόνοια άνοιας. Αυτό επιτυγχάνεται λαμβάνοντας το ιστορικό του ασθενούς και πραγματοποιώντας μία εξέταση της νοητικής του κατάστασης, πιθανότατα σε συνδυασμό με κάποια νευροψυχολογική εξέταση. Εάν ο ιατρός προσδιορίσει πως το άτομο δεν είναι ούτε υγιές, ούτε πάσχει από άνοια, όμως έχει βιώσει νοητική εξασθένηση με βάση το ιστορικό με τις λειτουργικές δραστηριότητές του να έχουν παραμείνει ανεπηρέαστες, τότε ο ασθενής διαγιγνώσκεται με MCI [15].

Το MCI δεν αντικατοπτρίζει μία μακροχρόνια μειωμένη νοητική λειτουργία, αλλά την αλλαγή στις νοητικές ικανότητες του ατόμου. Για τον λόγο αυτό είναι πολύ σημαντική η παρουσία δεδομένων για το ιστορικό του ασθενούς σε βάθος χρόνου, ώστε ο κλινικός ιατρός να μπορεί να εστιάσει στο είδος των γνωστικών αλλαγών που έχουν παρατηρηθεί στον ασθενή σε ορισμένο χρονικό διάστημα. Αν, για παράδειγμα, το βασικότερο σύμπτωμα αφορά σε δυσλειτουργία της μνήμης, τότε ο ιατρός θα πρέπει να εστιάσει σε σχετικά περιστατικά, τα οποία να είναι πρόσφατα, δηλαδή να έχουν λάβει χώρα τον τελευταίο μισό έως έναν χρόνο [17].

Ύστερα από τον κλινικό χαρακτηρισμό των συμπτωμάτων του ασθενούς, το επόμενο βήμα περιλαμβάνει τον καθορισμό της αιτιολογίας των συμπτωμάτων. Η διαδικασία αυτή συνήθως βασίζεται το ιστορικό του ασθενούς και περιλαμβάνει κλινικές εξετάσεις για άλλες πιθανές διαταραχές που μπορεί να ευθύνονται για τη νοητική εξασθένηση, καθώς και νευροαπεικονιστικές μελέτες. Στη συνέχεια, ο ιατρός αποφαινεται αν η αιτία του MCI είναι εκφυλιστική, αγγειακή, ψυχιατρική ή δευτερογενής λόγω άλλων ιατρικών διαταραχών [15].

- Τύποι:

Μόλις η διάγνωση του MCI επιβεβαιωθεί, ο επόμενος στόχος αφορά στην αναγνώριση του υποτύπου. Ο κλινικός ιατρός πρέπει πρώτα να προσδιορίσει εάν η μνήμη έχει επηρεαστεί,

καθώς η εξασθένηση της μνήμης αποτελεί ισχυρή ένδειξη προδιάθεσης της νόσου Αλτσχάιμερ. Αυτό μπορεί να προσδιοριστεί πραγματοποιώντας τις κατάλληλες εξετάσεις και εάν διαπιστωθεί πως υπάρχει πράγματι εξασθένηση της μνήμης δεδομένης της ηλικίας του ατόμου, τότε θεωρείται πως ο ασθενής πάσχει από αμνησιακού (amnesic) τύπου MCI. Αντιθέτως, αν διαπιστωθεί πως η μνήμη έχει παραμείνει ανεπηρέαστη, τότε το άτομο πάσχει από εξασθένηση γνωστικών τομέων ανεξάρτητων από τη μνήμη, όπως η γλώσσα, η εκτελεστική λειτουργία ή οι οπτικοχωρικές ικανότητες και άρα από MCI μη αμνησιακού (non-amnesic) τύπου [15].

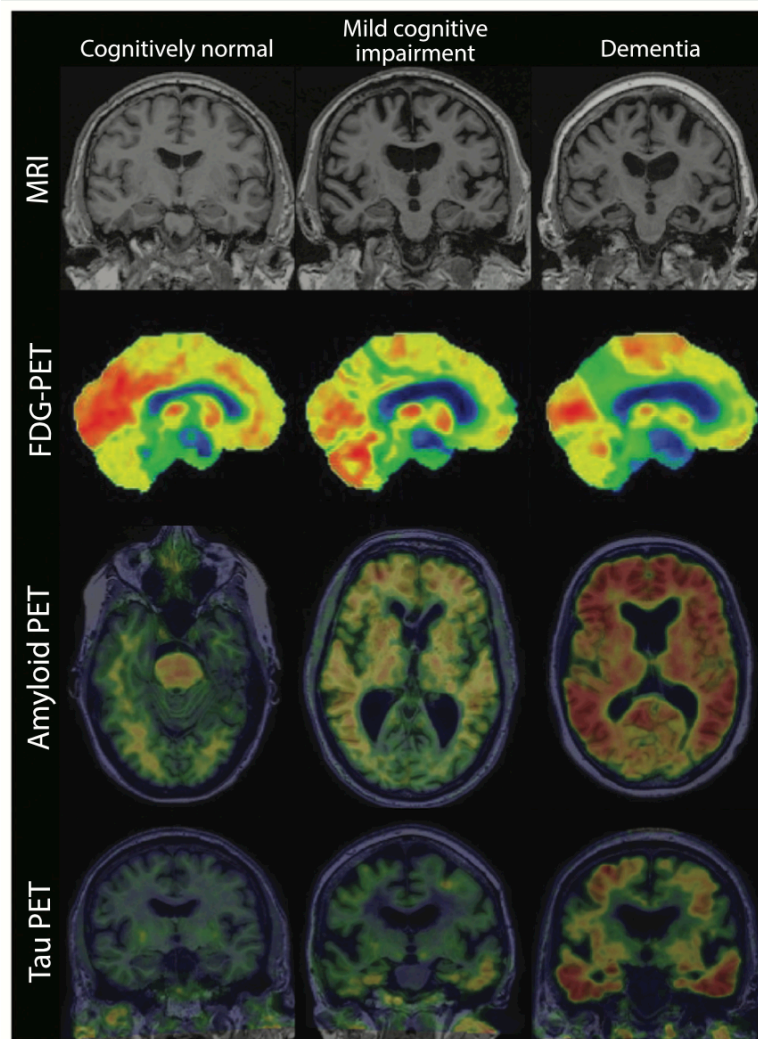
Τέλος, ο ιατρός πρέπει να προσδιορίσει εάν υπάρχουν άλλοι γνωστικοί τομείς που να έχουν επηρεαστεί, γεγονός που μπορεί να διαπιστωθεί πραγματοποιώντας, μεταξύ άλλων, νευροψυχολογικά τεστ. Η διάγνωση αφορά το αμνησιακού τύπου MCI μεμονωμένου τομέα (amnesic MCI single domain), εάν η μνήμη αποτελεί τον μοναδικό τομέα που έχει εξασθενήσει, ενώ αφορά το αμνησιακού τύπου MCI πολλαπλών τομέων (amnesic MCI multiple domain), εάν υπάρχει τουλάχιστον ακόμη ένα γνωστικό πεδίο εκτός της μνήμης που να έχει υποστεί εξασθένηση, όπως η γλώσσα, η εκτελεστική λειτουργία ή οι οπτικοχωρικές ικανότητες. Ομοίως, η διάγνωση του μη αμνησιακού MCI μεμονωμένου τομέα (non-amnesic MCI single domain) αφορά στην εξασθένηση ενός μόνο νοητικού τομέα εκτός της μνήμης και του μη αμνησιακού MCI πολλαπλών τομέων (non-amnesic MCI multiple domain), αφορά την εξασθένηση πολλαπλών νοητικών πεδίων εκτός της μνήμης [15].

- Εξέλιξη σε άνοια:

Το μεμονωμένου τομέα και πολλαπλών τομέων αμνησιακό MCI με εκφυλιστική αιτία, αποτελεί ένδειξη για πιθανή ανάπτυξη της νόσου Αλτσχάιμερ, ενώ οι μη αμνησιακοί υποτύποι του MCI, όπου η εξασθένηση αφορά γνωστικούς τομείς πέραν της μνήμης, οδηγούν σε υψηλότερη πιθανότητα ανάπτυξης μορφών άνοιας μη σχετιζόμενων με τη νόσο Αλτσχάιμερ, όπως η μετωποκροταφική άνοια ή η άνοια με σωμάτια Lewy. Επιπλέον, τα άτομα με αμνησιακού τύπου MCI πολλαπλών τομέων είναι πιθανότερο να αναπτύξουν ταχύτερα άνοια (συνήθως με ρυθμό 10% έως 15% ανά έτος), σε σχέση με τα άτομα που έχουν αμνησιακού τύπου MCI μεμονωμένου τομέα. Ο συνδυασμός του κλινικού τύπου και της πιθανής αιτιολογίας της ασθένειας μπορεί να αποβεί ιδιαίτερα χρήσιμος κατά την πρόβλεψη του τελικού τύπου της άνοιας, στην οποία θα εξελιχθούν οι παραπάνω ασθένειες [15].

Η δυνατότητα αναγνώρισης των ασθενών με MCI, οι οποίοι πρόκειται να αναπτύξουν άνοια ή νόσο Αλτσχάιμερ πιο ραγδαία από άλλους παραμένει μία σημαντική περιοχή ενδιαφέροντος στο πεδίο της έρευνας γύρω από το MCI [15].

Τα τελευταία χρόνια, η νευροαπεικόνιση έχει φανεί ιδιαίτερα χρήσιμη στην πρόβλεψη της εξέλιξης του MCI σε νόσο Αλτσχάιμερ [15]. Ορισμένες μέθοδοι είναι ευαίσθητες στο MCI, όπως η απεικόνιση του εγκεφάλου με MRI, η τομογραφία εκπομπής ποζιτρονίου και το ηλεκτροεγκεφαλογράφημα. Οι νευροαπεικονιστικές και ηλεκτροφυσιολογικές εξετάσεις για την ανάλυση του MCI συχνά ταυτίζονται με αυτές που χρησιμοποιούνται στα αρχικά στάδια της άνοιας [16]. Έχει αποδειχθεί πως η ατροφία του ιπποκάμπου αποτελεί σημαντικό προβλεπτικό παράγοντα της εξέλιξης του αμνησιακού MCI σε νόσο Αλτσχάιμερ, ενώ στοιχεία, όπως ο όγκος ολόκληρου του εγκεφάλου ή οι όγκοι των κοιλιών, συνεισφέρουν επίσης στην πρόβλεψη της εξέλιξης σε νόσο Αλτσχάιμερ, γεγονός που αποδεικνύει τη χρησιμότητα του δομικού MRI, αλλά και του FDG-PET [15]. Παράλληλα, η ατροφία του μέσου κροταφικού λοβού στη μαγνητική τομογραφία και του υπομεταβολισμού της φθοριοδεοξυγλυκόζης στην τομογραφία εκπομπής ποζιτρονίου έχουν καταγραφεί σε ασθενείς με MCI και η παρουσία αυτών των στοιχείων έχει υψηλή προβλεπτική αξία για την εξέλιξη του MCI σε άνοια [16]. Επιπλέον, οι μοριακές απεικονιστικές τεχνικές που επιτρέπουν την οπτικοποίηση της εξέλιξης της παθολογικής διαδικασίας έχουν θεωρηθεί ιδιαίτερα ενδιαφέρουσες [15].



Εικόνα 4: Εξέλιξη των απεικονιστικών χαρακτηριστικών από τη φυσιολογική κατάσταση στην ήπια νοητική εξασθένιση και την άνοια

Αξίζει επίσης να σημειωθεί πως, η κατάσταση φορέα της απολιποπρωτεΐνης E-E4 (ApoE4) αποτελεί έναν ισχυρά εδραιωμένο παράγοντα κινδύνου, καθώς έχει αποδειχθεί πως συνεισφέρει στην πρόβλεψη της εξέλιξης του MCI σε νόσο Αλτσχάιμερ [15]. Οι μεταλλάξεις στα άλληλα της απολιποπρωτεΐνης E εμφανώς αυξάνουν τον κίνδυνο εξέλιξης του αμνησιακού MCI σε νόσο Αλτσχάιμερ, καθώς μεταβάλλουν τη μεταφορά χοληστερόλης και τη συναπτική πλαστικότητα [16]. Έχει, επίσης, αποδειχθεί πως συσχετίζονται με πιο ραγδαία εξέλιξη της ατροφίας του ιπποκάμπου στην απεικόνιση MRI σε γνωστικά φυσιολογικά άτομα [15].

Παρόλο που τα γνωστικά συμπτώματα και η αναγνώρισή τους με τις αντίστοιχες εξετάσεις, έχουν αποτελέσει τα κυριότερα χαρακτηριστικά της Ήπιας Μορφής Νοητικής Εξασθένισης έως τώρα, δίνεται ολοένα αυξανόμενη προσοχή στον τομέα της συμπεριφοράς, ο οποίος περιλαμβάνει το άγχος, την κατάθλιψη, την ευερεθιστότητα και την απάθεια. Η παρουσία συμπεριφορικών και ψυχολογικών ενδείξεων αποτελεί ένδειξη υψηλής πιθανότητας για ανάπτυξη άνοιας, ενώ το MCI συνοδεύεται επίσης και από αλλαγές στην ισορροπία και τον συντονισμό, ενώ ευρήματα επιδημιολογικών ερευνών έχουν δείξει πως δυσκολίες στην εκτέλεση καθημερινών δραστηριοτήτων εντοπίζονται συχνά σε άτομα με MCI 2 χρόνια πριν από τη διάγνωση της άνοιας. [16].

2.4.2 Άνοια

Η άνοια αποτελεί μία επίκτητη απώλεια νοητικών ικανοτήτων σε πολλαπλούς γνωστικούς τομείς, η οποία επηρεάζει έντονα τη λειτουργικότητα του ατόμου. Συνιστά μία εξαιρετικά κοινή ασθένεια που πλήττει πάνω από 50 εκατομμύρια ανθρώπους παγκοσμίως και η οποία συνήθως αποδίδεται στη νόσο Αλτσχάιμερ. Η άνοια χαρακτηρίζεται από χρόνια επίκτητη απώλεια τουλάχιστον μίας από τις γνωστικές ικανότητες που προκαλείται από ασθένεια του εγκεφάλου ή τραυματισμό και πρόκειται για ένα κλινικό σύνδρομο με ποικίλες εκδηλώσεις. Αξίζει να σημειωθεί πως ο διαχωρισμός της νόσου Αλτσχάιμερ από άλλα αίτια που προκαλούν άνοια είναι ευκολότερος στα αρχικά στάδια της ασθένειας, καθώς οι περισσότερες μορφές άνοιας σε προχωρημένο στάδιο είναι όμοιες [18].

- Παράγοντες κινδύνου και νευροπαθολογία:

Η γήρανση αποτελεί έναν από τους σημαντικότερους παράγοντες για κάθε τύπο άνοιας, ενώ επίσης σημαντικούς παράγοντες αποτελούν η υπέρταση και ο διαβήτης, η διατροφή και οι περιορισμένες νοητικές, σωματικές και κοινωνικές δραστηριότητες.

Η παθολογικά “μικτή” άνοια αποτελεί τη συνηθέστερη μορφή άνοιας, η οποία εντοπίζεται στο 46% των ατόμων που είναι κλινικά διεγνωσμένα με νόσο Αλτσχάιμερ και συνοδεύεται από νευροεκφυλισμό της νόσου Αλτσχάιμερ και εγκεφαλοαγγειακές παθήσεις. Μετά τη νόσο Αλτσχάιμερ, οι συνηθέστερες παθήσεις νευροεκφυλιστικής άνοιας είναι η άνοια με σωματίδια Lewy, η οποία αφορά το 17% των περιπτώσεων και η εκφύλιση του μετωποκροταφικού λοβού που αφορά λιγότερο από 5%. Η άνοια με σωματίδια Lewy χαρακτηρίζεται από χρόνια διαταραχή της συμπεριφοράς του ύπνου REM με πρώιμη οπτικοχωρική δυσλειτουργία και παρκινσονισμό, ενώ η μετωποκροταφική άνοια χαρακτηρίζεται από διαταραχές της συμπεριφοράς ή σπανιότερα γλωσσική δυσλειτουργία [18].

- Διάγνωση:

Η αξιολόγηση της άνοιας απαιτεί ένα σύντομο ιατρικό ιστορικό, καθώς και γνωστική και νευρολογική εξέταση. Το ιστορικό παραμένει το σημαντικότερο διαγνωστικό εργαλείο και θα πρέπει να αποκτάται και από τον ίδιο τον ασθενή, αλλά και από ένα κοντινό του πρόσωπο, καθώς σε αντίθεση με ορισμένους ασθενείς που αναγνωρίζουν τα συμπτώματα απώλειας μνήμης, υπάρχουν περιπτώσεις όπου ο ασθενής αδυνατεί να θυμηθεί λεπτομέρειες σχετικά με το ιστορικό του ή παρουσιάζει ανογνωσία, δηλαδή αδυναμία να αντιληφθεί την πάθησή του [18].

Το ιστορικό θα πρέπει να εστιάζει στις ιατρικές παθήσεις που επηρεάζουν τις νοητικές λειτουργίες, συμπεριλαμβανομένων των παραγόντων κινδύνου που σχετίζονται με αγγειακές νόσους (όπως η υπέρταση και ο διαβήτης), των υπαρχουσών εγκεφαλικών παθήσεων (όπως το εγκεφαλικό επεισόδιο, η νόσος Πάρκινσον ή κάποια κranioεγκεφαλική κάκωση), καθώς επίσης και πιθανής φαρμακευτικής αγωγής που μπορεί να επηρεάζει τις νοητικές λειτουργίες (για παράδειγμα αγχολυτικά, όπως οι βενζοδιαζεπίνες, τα αναλγητικά που περιέχουν κωδεΐνη και άλλα) [18].

Από τη γνωστική εξέταση μπορεί να αξιολογηθεί η παρουσία, η σοβαρότητα και η φύση της νοητικής δυσλειτουργίας, ενώ από τη νευρολογική εξέταση μπορούν να εντοπιστούν αντικειμενικά στοιχεία νευρογνωστικών προβλημάτων, όπως η αφασία, η απραξία και η αγνωσία. Η αξιολόγηση περιλαμβάνει επίσης σωματική εξέταση, ώστε να εντοπιστεί πιθανή συστηματική αγγειακή νόσος και συστηματικά σημάδια που μπορεί να σχετίζονται με σπανιότερα αίτια άνοιας. Η εξέταση ρουτίνας περιλαμβάνει συνήθως ορισμένες αιματολογικές εξετάσεις (για παράδειγμα B12 και TSH) και κάποια νευροαπεικονιστική εξέταση, ώστε να αναγνωριστεί η ατροφία του φλοιού και του ιπποκάμπου (όπως είθισται να παρατηρείται στη νόσο Αλτσχάιμερ) ή νευροπαθολογία που αφορά δυνητικά θεραπεύσιμες αιτίες άνοιας [18].

Για ασθενείς των οποίων η διάγνωση δεν είναι συνεπής με τη νόσο Αλτσχάιμερ ή υπάρχει μικρή βεβαιότητα όσον αφορά τη διάγνωση, οι κλινικοί ιατροί μπορούν να εξετάσουν το ενδεχόμενο παραπομπής σε ειδικό και διενέργειας περαιτέρω εξετάσεων [18].

Η λειτουργική νευροαπεικόνιση, όπως η τομογραφία εκπομπής ποζιτρονίου (PET) μπορεί να δείξει αλλαγές που υποδηλώνουν νόσο Αλτσχάιμερ και σχετίζονται συνήθως με ασύμμετρο αμφίπλευρη κροταφο-παρεγκεφαλικό υπομεταβολισμό που ανιχνεύεται με συνήθεις ιχνηθέτες, όπως η φθοριοδεοξυγλυκόζη (FDG) [18].

Για άτομα με μετωποκροταφική άνοια, το FDG-PET συνήθως δείχνει ελαττωμένο, ασύμμετρο υπομεταβολισμό του μετωπιαίου λοβού στους ασθενείς που έχουν τον τύπο με διαταραχή συμπεριφοράς και υπομεταβολισμό του πρόσθιου κροταφικού λοβού σε εκείνους που πάσχουν από μετωποκροταφική άνοια με γλωσσική διαταραχή [18].

Σε ορισμένες περιπτώσεις μπορεί να καταστεί απαραίτητη και η εξέταση εγκεφαλονωτιαίου υγρού για τον εντοπισμό στοιχείων που υποδεικνύουν τη νόσο Αλτσχάιμερ (όπως χαμηλά επίπεδα αμυλοειδούς και υψηλά επίπεδα πρωτεΐνης ταυ), άλλες νευροεκφυλιστικές ασθένειες ή άλλες αιτίες άνοιας. Τέλος, μπορεί να είναι χρήσιμος ο γενετικός έλεγχος, ιδιαίτερα σε ασθενείς νεαρής ηλικίας με κληρονομικότητα από συγγενείς πρώτου βαθμού [18].

3. Θεωρητικό υπόβαθρο

3.1 Μέθοδοι Ταξινόμησης

Τα τελευταία χρόνια στον τομέα της υγείας έχει γίνει διαθέσιμος μεγάλος όγκος δεδομένων, τα οποία μπορεί να είναι είτε δομημένα είτε μη δομημένα. Τα δομημένα ιατρικά δεδομένα αφορούν πληροφορία που εύκολα ταξινομείται σε βάσεις δεδομένων, ενώ τα μη δομημένα, που αποτελούν και την πλειοψηφία, είναι δύσκολο να οργανωθούν με τρόπο τέτοιο, ώστε να μπορούν να αξιοποιηθούν προς όφελος του τομέα της υγείας [19]. Για τον λόγο αυτό, οι αλγόριθμοι μηχανικής μάθησης μπορούν να φανούν εξαιρετικά χρήσιμοι, καθώς δύνανται να εντοπίσουν περίπλοκα πρότυπα ανάμεσα σε μεγάλο όγκο δεδομένων. Ενσωματώνονται σε συστήματα διάγνωσης με τη βοήθεια υπολογιστή (Computer Aided Diagnosis - CAD) και χρησιμοποιούνται συχνά για την ανίχνευση ασθενειών, την πρόβλεψη της επικινδυνότητας ορισμένων παθήσεων και τη λήψη αποφάσεων σχετικά με θεραπευτικά σχήματα των ασθενών [20].

Προκειμένου να αποφανθεί ένας αλγόριθμος μηχανικής μάθησης αν ένας ασθενής πάσχει ή όχι από μία ασθένεια πρέπει να επιλύσει ένα πρόβλημα ταξινόμησης. Η ταξινόμηση των δεδομένων αφορά την απόφαση σχετικά με την κλάση y' στην οποία ανήκει ένα άγνωστο στοιχείο δεδομένων x' με βάση ένα σύνολο δεδομένων $D = ((x_1, y_1), \dots, (x_n, y_n))$ από στοιχεία x_i με γνωστές κλάσεις y_i . Χάριν ευκολίας, θεωρείται πως η ταξινόμηση είναι δυαδική και οι ετικέτες των κλάσεων είναι είτε 0 είτε 1. Τα x_i είθισται να είναι διανύσματα m διαστάσεων των οποίων τα στοιχεία ονομάζονται ανεξάρτητες μεταβλητές ή μεταβλητές εισόδου. Στα περισσότερα προβλήματα ταξινόμησης, δεν υφίσταται συναρτησιακή σχέση $y = f(x)$ ανάμεσα στο y και το x . Σε αυτή την περίπτωση, η σχέση ανάμεσά τους πρέπει να περιγραφεί με γενικότερο τρόπο από μία κατανομή πιθανοτήτων $P(x, y)$, υποθέτοντας ότι το σύνολο δεδομένων D περιλαμβάνει δείγματα ανεξάρτητα από την P και από στατιστικής άποψης, η βέλτιστη απόφαση συμμετοχής ενός δεδομένου σε μία κλάση αφορά την επιλογή της ετικέτας y που μεγιστοποιεί την εκ των υστέρων κατανομή $P(y|x)$ [21].

Υπάρχουν δύο διαφορετικές προσεγγίσεις στην ταξινόμηση δεδομένων. Η πρώτη αφορά την δυαδική ταξινόμηση, δηλαδή τον διαχωρισμό ανάμεσα σε δύο κλάσεις και αναθέτει σε άγνωστα στοιχεία δεδομένων ετικέτες 0 ή 1, ενώ η δεύτερη προσπαθεί να μοντελοποιήσει την πιθανότητα $P(y|x)$ και αποδίδει εκτός από μία ετικέτα και την πιθανότητα συμμετοχής ενός άγνωστου στοιχείου σε μία κλάση [21]. Παρακάτω παρουσιάζεται ένα σύνολο από αλγορίθμους μηχανικής μάθησης για την ταξινόμηση δεδομένων.

3.1.1 Αλγόριθμοι Μηχανικής Μάθησης

3.1.1.1 Τυχαία Δάση

Τα Τυχαία Δάση (Random Forests - RF) αποτελούν έναν αλγόριθμο που στηρίζεται στη μη παραμετρική εκτίμηση παλινδρόμησης (regression). Έστω ένα δείγμα εκπαίδευσης $D_n = ((X_1, Y_1), \dots, (X_n, Y_n))$ που αποτελείται από ανεξάρτητες τυχαίες μεταβλητές, με την ίδια κατανομή όπως το ανεξάρτητο πρωτότυπο ζεύγος (X, Y) . Στόχος είναι η εκτίμηση της συνάρτησης παλινδρόμησης $m(x) = \mathbb{E}[Y|X=x]$, με τη χρήση του συνόλου δεδομένων D_n . Θεωρείται πως η συνάρτηση παλινδρόμησης είναι συνεπής αν το μέσο τετραγωνικό σφάλμα $\mathbb{E}[m_n(X) - m(X)]^2 \rightarrow 0$ όσο $n \rightarrow \infty$ [22].

Τα Τυχαία Δάση απαρτίζονται από ένα σύνολο M τυχαιοποιημένων δέντρων παλινδρόμησης. Για το j -οστό δέντρο, η τιμή πρόβλεψης για το σημείο x είναι $m_n(x; \Theta_j, D_n)$, όπου $\Theta_1, \dots, \Theta_M$ είναι ανεξάρτητες τυχαίες μεταβλητές, με την ίδια κατανομή με μία γενική τυχαία μεταβλητή Θ και ανεξάρτητες από το D_n . Στην πράξη, η μεταβλητή Θ χρησιμοποιείται για την επαναδειγματοληψία (resampling) του συνόλου δεδομένων εκπαίδευσης πριν την ανάπτυξη μεμονωμένων δέντρων και για τη διαλογή διαδοχικών κατευθύνσεων για τη διάσπαση ενός δέντρου [22].

Η εκτίμηση του j -οστού δέντρου έχει τη μαθηματική μορφή:

$$m_n(x; \Theta_j, D_n) = \sum_{i \in D_n^*(\Theta_j)} \frac{\mathbf{1}_{x_j \in A_n(x; \Theta_j, D_n)} Y_i}{N_n(x; \Theta_j, D_n)},$$

όπου $D_n^*(\Theta_j)$ είναι το σύνολο των σημείων των δεδομένων που έχουν επιλεγεί πριν τη δημιουργία του δέντρου, $A_n(x; \Theta_j, D_n)$ είναι το κελί που περιέχει το x και $N_n(x; \Theta_j, D_n)$ είναι το πλήθος των προεπιλεγμένων σημείων που ανήκουν στο κελί $A_n(x; \Theta_j, D_n)$ [22].

Τα δέντρα συνδυάζονται, ώστε να σχηματίσουν την πεπερασμένη εκτίμηση του δάσους:

$$m_{M,n}(x; \Theta_1, \dots, \Theta_M, D_n) = \frac{1}{M} \sum_{j=1}^M m_n(x; \Theta_j, D_n)$$

Κάθε δέντρο απόφασης κατασκευάζεται από ένα τυχαίο διάνυσμα παραμέτρων. Ένα τυχαίο δάσος μπορεί να κατασκευαστεί, για παράδειγμα, με τυχαία δειγματοληψία ενός υποσυνόλου χαρακτηριστικών για κάθε δέντρο απόφασης ή με τυχαία δειγματοληψία ενός υποσυνόλου δεδομένων εκπαίδευσης για κάθε δέντρο απόφασης [23].

Όλοι οι κόμβοι ενός δέντρου συνδέονται μεταξύ τους με ορθογώνια κελιά, ώστε σε κάθε βήμα της κατασκευής ενός μεμονωμένου τυχαίου δέντρου, το σύνολο των κελιών που συνδέονται με τους τερματικούς κόμβους (φύλλα) να σχηματίζει μία διαμέριση (partition) [24]. Η ρίζα του δέντρου είναι το ίδιο το X και σε κάθε βήμα της κατασκευής δέντρων, ένας κόμβος (ή ισοδύναμα το αντίστοιχο κελί) διασπάται σε δύο μέρη.

Πριν την κατασκευή κάθε δέντρου, επιλέγονται τυχαία a_n δείγματα από το αρχικό σύνολο δεδομένων. Μόνο αυτά τα δείγματα λαμβάνονται υπόψιν κατά την κατασκευή των δέντρων. Στη συνέχεια, σε κάθε κελί του δέντρου πραγματοποιείται μία διάσπαση μεγιστοποιώντας το κριτήριο CART (βλ. παρακάτω) για το πλήθος των πιθανών κατευθύνσεων διάσπασης σε κάθε κόμβο κάθε δέντρου που επιλέγονται ομοιόμορφα με τυχαίο τρόπο ανάμεσα στις p αρχικές. Η κατασκευή μεμονωμένων δέντρων τερματίζεται όταν κάθε κελί περιέχει λιγότερα σημεία από το ελάχιστο πλήθος δειγμάτων που έχει προκαθοριστεί για τον τερματισμό των διασπάσεων σε κάθε κόμβο. Για κάθε σημείο $x \in X$, κάθε δέντρο παλινδρόμησης προβλέπει τον μέσο όρο των Y_i που βρίσκονταν ανάμεσα στα a_n δείγματα, για τα οποία τα αντίστοιχα X_i εμπίπτουν στο κελί του x [22].

Όσον αφορά το κριτήριο CART (Δέντρα Ταξινόμησης και Παλινδρόμησης - Classification And Regression Trees), έστω A ένα κελί και $N_n(A)$ το πλήθος των δειγμάτων που εμπίπτουν στο A . Μία τομή στο A έχει τη μορφή (j, z) , όπου ως j συμβολίζεται κάποια διάσταση από το $\{1, \dots, p\}$ και z η θέση της τομής στη j -οστή συντεταγμένη, εντός των ορίων του A . Έστω C_A το σύνολο όλων των πιθανών τομών στο A . Τότε με τον

συμβολισμό $\mathbf{X}_i = X_i^{(1)}, \dots, X_i^{(p)}$ για οποιοδήποτε $(j, z) \in C_A$, το κριτήριο CART λαμβάνει τη μορφή:

$$L_{reg,n}(j, z) = \frac{1}{N_n(A)} \sum_{i=1}^n (Y_i - \bar{Y}_A)^2 \mathbf{1}_{X_i \in A} - \frac{1}{N_n(A)} \sum_{i=1}^n (Y_i - \bar{Y}_{A_L} \mathbf{1}_{X_i^{(j)} < z} - \bar{Y}_{A_R} \mathbf{1}_{X_i^{(j)} \geq z})^2 \mathbf{1}_{X_i \in A} \quad (1)$$

όπου $A_L = \{\mathbf{x} \in A : \mathbf{x}^{(j)} < z\}$, $A_R = \{\mathbf{x} \in A : \mathbf{x}^{(j)} \geq z\}$ και \bar{Y}_A (και $\bar{Y}_{A_L}, \bar{Y}_{A_R}$ αντίστοιχα) είναι ο μέσος όρος των Y_i , ώστε το \mathbf{X}_i να ανήκει στο A (A_L, A_R αντίστοιχα) με τη σύμβαση ότι ο μέσος όρος ισούται με 0 όταν κανένα σημείο \mathbf{X}_i δεν ανήκει στο A . Για κάθε κελί A , η βέλτιστη τομή (j_n^*, z_n^*) επιλέγεται μεγιστοποιώντας το $L_{reg,n}(j, z)$ για το πλήθος των πιθανών κατευθύνσεων διάσπασης σε κάθε κόμβο κάθε δέντρου και για το σύνολο C_A κάθε πιθανής τομής [22].

Με τον τρόπο αυτό, σε κάθε κελί κάθε δέντρου, ο αλγόριθμος επιλέγει ομοιόμορφα στην τύχη τόσες συντεταγμένες στο $\{1, \dots, p\}$, όσες το πλήθος των πιθανών κατευθύνσεων διάσπασης, αξιολογεί το παραπάνω κριτήριο (1) ανάμεσα σε όλες τις πιθανές διασπάσεις για το παραπάνω πλήθος κατευθύνσεων και επιστρέφει τη βέλτιστη. Το παραπάνω ποσοτικό κριτήριο υπολογίζει την κανονικοποιημένη διαφορά ανάμεσα στη διακύμανση του κόμβου πριν και μετά τη διάσπαση.

Στην περίπτωση της δυαδικής ταξινόμησης, η τυχαία απόκριση Y λαμβάνει τιμές στο $\{0, 1\}$ και δεδομένου του \mathbf{X} πρέπει να γίνει πρόβλεψη της τιμής του Y . Στην περίπτωση αυτή ο ταξινομητής λαμβάνεται μέσω ψήφου πλειοψηφίας ανάμεσα στα δέντρα ταξινόμησης, δηλαδή είναι:

$$m_{M,n}(\mathbf{x}; \Theta_j, D_n) = \begin{cases} 1 & \text{αν } \sum_{i \in D_n^*(\Theta_j)} \mathbf{1}_{\mathbf{x}_j \in A, Y_i = 1} > \sum_{i \in D_n^*(\Theta_j)} \mathbf{1}_{\mathbf{x}_j \in A, Y_i = 0}, \mathbf{x} \in A \\ 0 & \text{διαφορετικά} \end{cases}$$

όπου το $D_n^*(\Theta_j)$ περιλαμβάνει τα δείγματα που επιλέχθηκαν στο στάδιο της επαναδειγματοληψίας, δηλαδή σε κάθε φύλλο έχει εφαρμοστεί ψήφος πλειοψηφίας σε όλα τα (\mathbf{X}_i, Y_i) για τα οποία το \mathbf{X}_i ανήκει στην ίδια περιοχή [22].

Όσον αφορά το κριτήριο CART, χρησιμοποιεί τη μέθοδο του συντελεστή Gini, προκειμένου να πραγματοποιήσει τη διάσπαση των κόμβων. Ο συντελεστής Gini αφορά στην πιθανότητα ενός τυχαία επιλεγμένου δείγματος του συνόλου δεδομένων να διασπαστεί. Όσο μικρότερος είναι ο δείκτης Gini, τόσο χαμηλότερη είναι η πιθανότητα το επιλεγμένο δείγμα να διασπαστεί και άρα τόσο υψηλότερη η καθαρότητα του συνόλου και αντιστρόφως. Ο δείκτης Gini, δηλαδή ο δείκτης καθαρότητας ισούται με την πιθανότητα ενός δείγματος να επιλεγεί επί την πιθανότητα εσφαλμένης ταξινόμησης του δείγματος.

$$Gini(p) = \sum_{k=1}^K p_k(1 - p_k) = 1 - \sum_{k=1}^K p_k^2$$

Το p_k υποδηλώνει την πιθανότητα το επιλεγμένο δείγμα να ανήκει στην κατηγορία k από τις K το πλήθος κατηγορίες του συνόλου δεδομένων, επομένως η πιθανότητα το δείγμα να διαχωριστεί είναι $(1 - p_k)$.

Το κριτήριο CART, στην περίπτωση της ταξινόμησης, αποτελεί ένα δυαδικό δέντρο, επομένως όταν ένα χαρακτηριστικό χρησιμοποιείται για τον διαχωρισμό ενός συνόλου

δειγμάτων, υπάρχουν δύο σύνολα, το D_1 που ισούται με μία δεδομένη τιμή χαρακτηριστικού και το σύνολο D_2 που δεν ισούται με τη δεδομένη τιμή.

Για καθεμία από τις παραπάνω διασπάσεις, η καθαρότητα του διαχωρισμού του συνόλου δεδομένων D σε δύο υποσύνολα με βάση το χαρακτηριστικό διαμέρισης ισούται με μία τιμή που υπολογίζεται ως εξής:

$$Gini(D, V) = \frac{|D_1|}{|D|} \cdot Gini(D_1) + \frac{|D_2|}{|D|} \cdot Gini(D_2)$$

Επομένως, για ένα χαρακτηριστικό με πολλαπλές τιμές, είναι απαραίτητο να υπολογιστεί η καθαρότητα $Gini(D, V_i)$ του υποσυνόλου μετά τον διαχωρισμό του δείγματος D με κάθε τιμή να αποτελεί σημείο διαχωρισμού (όπου το V_i αναπαριστά πιθανή τιμή του χαρακτηριστικού v).

Επομένως, ο μικρότερος δείκτης Gini, είναι αυτός που προκύπτει υπολογίζοντας όλα τα πιθανά $Gini(D, V_i)$. Το σημείο διαχωρισμού σε αυτή τη διαμέριση είναι το βέλτιστο σημείο διαχωρισμού του συνόλου δειγμάτων D με βάση το χαρακτηριστικό V [25].

Τα Τυχαία Δάση περιλαμβάνουν πλήθος υπερπαραμέτρων που καθορίζουν τη δομή κάθε μεμονωμένου δέντρου, το μέγεθος του δάσους (δηλαδή το πλήθος των δέντρων που το απαρτίζουν), όπως και την τυχαιότητά τους. Η επίδραση των υπερπαραμέτρων αυτών έχει μελετηθεί σε πλήθος ερευνών. Για τις περισσότερες από τις υπερπαραμέτρους είναι απαραίτητη η εύρεση των βέλτιστων τιμών, καθώς συχνά αυτές εξαρτώνται από το σύνολο δεδομένων. Υπερπαραμέτροι του αλγορίθμου είναι το πλήθος των υποψήφιων χαρακτηριστικών του συνόλου δεδομένων που λαμβάνονται υπόψιν σε κάθε διάσπαση, το σχήμα δειγματοληψίας, το ελάχιστο μέγεθος κόμβου και οι σχετικές με αυτό παράμετροι, το πλήθος των δέντρων και το κριτήριο με βάση το οποίο γίνεται διάσπαση των δέντρων [26].

- Πλήθος των τυχαία επιλεγμένων υποψήφιων χαρακτηριστικών

Μία από τις βασικές υπερπαραμέτρους των Τυχαίων Δασών ορίζεται ως το πλήθος των τυχαία επιλεγμένων υποψήφιων μεταβλητών από τις οποίες καθορίζεται η κάθε διάσπαση κατά την ανάπτυξη ενός δέντρου (`max_features`). Χαμηλότερες τιμές της υπερπαραμέτρου οδηγούν σε δέντρα με μικρότερη συσχέτιση, επομένως προσφέρουν καλύτερη σταθερότητα κατά τη συνένωση των δέντρων. Τα δάση που κατασκευάζονται με χαμηλές τιμές `max_features`, παρουσιάζουν καλύτερη αξιοποίηση των μεταβλητών που έχουν ήπια επίδραση στη μεταβλητή απόκρισης, η οποία θα καλυπτόταν από μεταβλητές με ισχυρή επίδραση, αν αυτές ήταν υποψήφιες για τον καθορισμό της διάσπασης. Ωστόσο, χαμηλές τιμές της παραμέτρου, οδηγούν επίσης σε δέντρα που κατά μέσο όρο έχουν χειρότερη επίδοση, αφού έχουν κατασκευαστεί με βάση μη βέλτητα χαρακτηριστικά που επιλέχθηκαν από ένα μικρό σύνολο τυχαία επιλεγμένων υποψήφιων μεταβλητών, με αποτέλεσμα την πιθανή εκλογή μη σημαντικών χαρακτηριστικών. Επομένως, για την επιλογή του πλήθους των υποψήφιων χαρακτηριστικών πρέπει να ληφθεί υπόψιν το αντιστάθμισμα ανάμεσα στη σταθερότητα και την ακρίβεια των μεμονωμένων δέντρων. Για μεγάλες τιμές της υπερπαραμέτρου, οι μεταβλητές με μικρή επίδραση, ίσως να μην έχουν τη δυνατότητα να συνεισφέρουν στην πρόβλεψη, καθώς ισχυρότερα χαρακτηριστικά προτιμώνται για τον καθορισμό της διάσπασης και επισκιάζουν αυτά με τις μικρότερες επιδράσεις. Μεγάλες τιμές του `max_features` εξασφαλίζουν ότι με μεγάλη πιθανότητα θα υπάρχει τουλάχιστον ένα ισχυρό χαρακτηριστικό στο σύνολο των υποψήφιων μεταβλητών [26].

- Μέγεθος Δείγματος

Το μέγεθος του δείγματος (`min_samples_split`) αποτελεί μία υπερπαράμετρο που καθορίζει, πόσα δείγματα επιλέγονται για την εκπαίδευση κάθε δέντρου. Έχει παρόμοια επίδραση όπως η παράμετρος `max_features`. Ελαττώνοντας το μέγεθος του δείγματος, προκύπτουν δέντρα με μεγαλύτερη ποικιλομορφία, και άρα με μικρότερη συσχέτιση ανάμεσά τους, γεγονός που επιδρά θετικά στην ακρίβεια της πρόβλεψης κατά τη συνένωση των δέντρων. Ωστόσο, η ακρίβεια ενός μεμονωμένου δέντρου ελαττώνεται, καθώς χρησιμοποιούνται λιγότερα δείγματα κατά την εκπαίδευση. Επομένως, και πάλι για την επιλογή του μεγέθους του δείγματος λαμβάνεται υπόψιν το αντιστάθμισμα ανάμεσα στη σταθερότητα και την ακρίβεια των δέντρων [26].

- Μέγεθος Κόμβου

Η υπερπαράμετρος του μεγέθους κόμβου (`min_samples_leaf`) καθορίζει το ελάχιστο πλήθος δειγμάτων σε έναν τερματικό κόμβο (φύλλο). Θέτοντας μία μικρή τιμή, προκύπτουν δέντρα με μεγαλύτερο βάθος, γεγονός που σηματοδοτεί μεγαλύτερο πλήθος διασπάσεων μέχρι τους τερματικούς κόμβους. Πειραματικά έχει προκύψει πως το υπολογιστικό κόστος ελαττώνεται σχεδόν εκθετικά με την αύξηση του μεγέθους κόμβου. Ειδικότερα, σε μεγάλα σύνολα δεδομένων, ο χρόνος εκτέλεσης ελαττώνεται σημαντικά όταν η συγκεκριμένη υπερπαράμετρος λαμβάνει τιμές μεγαλύτερες από την προκαθορισμένη και μάλιστα συχνά χωρίς απώλεια στην επίδοση της πρόβλεψης. Μία υπερπαράμετρος στενά συνδεδεμένη με το μέγεθος κόμβου είναι το μέγιστο βάθος δέντρου (`max_depth`), το οποίο αφορά το μέγιστο πλήθος διασπάσεων μέχρι έναν τερματικό κόμβο [26].

- Πλήθος δέντρων

Η υπερπαράμετρος του πλήθους δέντρων (`n_estimators`) πρέπει να λαμβάνει επαρκώς μεγάλες τιμές, καθώς έχει αποδειχθεί και θεωρητικά πως με την αύξηση του πλήθους των δέντρων αυξάνεται και η επίδοση της πρόβλεψης. Ο ρυθμός σύγκλισης και κατά συνέπεια το πλήθος των δέντρων που απαιτούνται για τη βέλτιστη επίδοση εξαρτώνται από τις ιδιότητες του συνόλου δεδομένων. Απαιτούνται περισσότερα δέντρα για να ληφθούν ξεκάθαρες προβλέψεις για κάθε δείγμα, επομένως υψηλότερος αριθμός δέντρων συνεπάγεται γρηγορότερη σύγκλιση. Το υπολογιστικό κόστος αυξάνεται γραμμικά με την αύξηση του πλήθους των δέντρων [26].

- Κριτήριο διαχωρισμού

Το κριτήριο διαχωρισμού δεν αποτελεί κλασική υπερπαράμετρο, καθώς πρόκειται για μία από τις βασικές ιδιότητες που χαρακτηρίζουν το Τυχαίο Δάσος, ωστόσο ενίοτε μπορεί να θεωρηθεί και ως κατηγορική υπερπαράμετρος. Το συνηθέστερο κριτήριο εφαρμόζεται επιλέγοντας από όλες τις πιθανές διασπάσεις των υποψήφιων χαρακτηριστικών, αυτή που ελαχιστοποιεί την καθαρότητα του δείκτη Gini (στην περίπτωση της ταξινόμησης) ή τη σταθμισμένη διακύμανση (στην περίπτωση της παλινδρόμησης). Η μέθοδος αυτή ευνοεί την επιλογή χαρακτηριστικών με πολλές πιθανές διασπάσεις σε σχέση με αυτά που οδηγούν σε λιγότερες διασπάσεις [26].

Τα Τυχαία Δάση εφαρμόζονται σε ένα μεγάλο εύρος προβλημάτων πρόβλεψης και απαιτούν βελτιστοποίηση λίγων μόνο υπερπαράμετρων. Αποτελούν μία απλή μέθοδο με υψηλή ακρίβεια πρόβλεψης και την ικανότητα να διαχειρίζονται μικρά και πολλών διαστάσεων σύνολα δεδομένων. Επιπλέον, πρόκειται για μία μέθοδο εύκολα παραλληλοποιήσιμη και άρα με προοπτική αντιμετώπισης μεγάλης κλίμακας πραγματικών προβλημάτων [22].

3.1.1.2 Λογιστική Παλινδρόμηση

Η Λογιστική Παλινδρόμηση (Logistic Regression - LR) είναι ένα μοντέλο, που στην περίπτωση της δυαδικής ταξινόμησης, υπολογίζει την πιθανότητα ενός στοιχείου να ανήκει σε μία από τις δύο κλάσεις δεδομένων. Δίνει σε κάθε προβλεπτικό παράγοντα (predictor) έναν συντελεστή που υπολογίζει την ανεξάρτητη συνεισφορά του στη διακύμανση της εξαρτημένης μεταβλητής. Η εξαρτημένη μεταβλητή Y , λαμβάνει τιμή 1 ή 0. Η μορφή του μοντέλου για τις προβλεφθείσες πιθανότητες εκφράζεται ως φυσικός λογάριθμος του λόγου των πιθανοτήτων:

$$\ln\left[\frac{P(Y)}{1 - P(Y)}\right] = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \dots + \beta_k \cdot X_k \Rightarrow$$

$$\frac{P(Y)}{1 - P(Y)} = e^{\beta_0 + \beta_1 \cdot X_1 + \dots + \beta_k \cdot X_k} \Rightarrow$$

$$P(Y) = \frac{e^{\beta_0 + \beta_1 \cdot X_1 + \dots + \beta_k \cdot X_k}}{1 + e^{\beta_0 + \beta_1 \cdot X_1 + \dots + \beta_k \cdot X_k}},$$

όπου το $\ln\left[\frac{P(Y)}{1 - P(Y)}\right]$ αποτελεί τον λογάριθμο των αποτελεσμάτων, Y είναι το δυαδικό

αποτέλεσμα, X_1, X_2, \dots, X_k είναι οι μεταβλητές πρόβλεψης και $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ είναι οι συντελεστές του μοντέλου παλινδρόμησης.

Στην παραπάνω εξίσωση, το μοντέλο λογιστικής παλινδρόμησης συσχετίζει άμεσα την πιθανότητα του Y με τις μεταβλητές πρόβλεψης. Στόχος του μοντέλου είναι ο υπολογισμός των $k + 1$ άγνωστων παραμέτρων β στην παραπάνω εξίσωση. Αυτό επιτυγχάνεται με την εκτίμηση μέγιστης πιθανοφάνειας που περιλαμβάνει την εύρεση του συνόλου των παραμέτρων για τις οποίες η πιθανότητα των υπό παρατήρηση δεδομένων είναι μέγιστη. Οι συντελεστές παλινδρόμησης υποδεικνύουν τον βαθμό συσχέτισης ανάμεσα σε κάθε ανεξάρτητη μεταβλητή και στο αποτέλεσμα. Κάθε συντελεστής αναπαριστά ποσοτικά την αλλαγή που αναμένεται στην μεταβλητή απόκρισης αν η μεταβλητή πρόβλεψης μεταβαλλόταν κατά μία μονάδα [27].

Στόχος της λογιστικής παλινδρόμησης είναι η σωστή πρόβλεψη της κλάσης των μεμονωμένων στοιχείων του συνόλου δεδομένων με τη χρήση του βέλτιστου μοντέλου. Προκειμένου να επιτευχθεί αυτό, δημιουργείται ένα μοντέλο που περιλαμβάνει όλες τις μεταβλητές πρόβλεψης που είναι χρήσιμες στην πρόβλεψη της μεταβλητής απόκρισης. Η λογιστική παλινδρόμηση υπολογίζει την πιθανότητα επιτυχίας έναντι της πιθανότητας αποτυχίας. Τα αποτελέσματα της ανάλυσης δίνονται στη μορφή ενός λόγου πιθανοτήτων [27].

Η δυαδική εξαρτημένη μεταβλητή λαμβάνει τιμές 0 και 1 και η προβλεφθείσα τιμή πρέπει να περιοριστεί στο ίδιο διάστημα. Προκειμένου να οριστεί η σχέση περιορισμού από τα 0 και 1, η λογιστική παλινδρόμηση χρησιμοποιεί τη λογαριθμική καμπύλη για την αναπαράσταση της σχέσης μεταξύ της ανεξάρτητης και της εξαρτημένης μεταβλητής. Σε πολύ χαμηλές τιμές της ανεξάρτητης μεταβλητής, η πιθανότητα προσεγγίζει το 0, χωρίς να γίνεται, όμως, ίση. Ομοίως, αν η ανεξάρτητη μεταβλητή αυξηθεί, οι τιμές που προβλέφθηκαν αυξάνονται στην καμπύλη και προσεγγίζουν το 1, χωρίς να γίνονται ίσες με 1 [27].

Ο λογαριθμικός μετασχηματισμός εξασφαλίζει ότι οι τιμές που υπολογίστηκαν δεν παίρνουν τιμές έξω από τα όρια του διαστήματος (0,1). Αυτό εξασφαλίζεται επαναδιατυπώνοντας την πιθανότητα με βάση τον λόγο της πιθανότητας να συμβεί ένα γεγονός προς την πιθανότητα να μη συμβεί. Προκειμένου να περιοριστούν οι τιμές που προβλέφθηκαν στο εύρος (0,1), ο παραπάνω λόγος μετατρέπεται εκ νέου σε πιθανότητα. Επιπλέον, προκειμένου ο λόγος των πιθανοτήτων να μην λάβει τιμή κάτω από το 0, η λογαριθμική τιμή υπολογίζεται λαμβάνοντας τον λογάριθμο του λόγου πιθανοτήτων. Για τιμές του λόγου μικρότερες του 1, η λογαριθμική τιμή είναι αρνητική, για τιμές

μεγαλύτερες του 1, θετική, ενώ όταν είναι ίσος με 1, ο αντίστοιχος λογάριθμος είναι ίσος με 1 [27].

Σε πολλές περιπτώσεις, το γεγονός εξόδου ταξινομείται εύκολα με βάση το αν έχει συμβεί ή όχι, ωστόσο, σε άλλες περιπτώσεις το αποτέλεσμα μπορεί να αντιμετωπίζεται ως δυαδικό, αλλά στην πραγματικότητα να προέρχεται από την αποκοπή συνεχών δεδομένων, δηλαδή να έχει τεθεί ένα κριτήριο αποκοπής και τα δεδομένα να έχουν κωδικοποιηθεί εκ νέου από συνεχή σε κατηγορικά στο σημείο αποκοπής. Σε αυτές τις περιπτώσεις, η επιλογή της μεταβλητής απόκρισης μπορεί να γίνει πιο περίπλοκη [27].

Μία άλλη παράμετρος που πρέπει να ληφθεί υπόψη στην ανάπτυξη ενός μοντέλου λογιστικής παλινδρόμησης είναι η επιλογή των μεταβλητών που θα αναλυθούν ως πιθανοί παράγοντες πρόβλεψης του αποτελέσματος. Το αποτέλεσμα της λογιστικής παλινδρόμησης εξαρτάται από τις μεταβλητές που έχουν επιλεγεί ως πιθανοί παράγοντες πρόβλεψης. Εάν μία μεταβλητή δεν έχει επιλεγεί για ανάλυση, τότε δεν συμπεριλαμβάνεται στο τελικό μοντέλο και η επιλογή σχετικά με τη συμμετοχή της ή όχι στην πρόβλεψη επηρεάζει το τελικό αποτέλεσμα. Ωστόσο, η ένταξη όσο το δυνατόν περισσότερων μεταβλητών, δεν αποτελεί λύση στο παραπάνω πρόβλημα, καθώς η ένταξη μεταβλητών που δεν σχετίζονται με το ζητούμενο αποτέλεσμα τείνει να επηρεάζει αρνητικά την προβλεπτική ικανότητα του τελικού μοντέλου [27].

3.1.1.3 Μηχανές Διανυσμάτων Υποστήριξης

Οι συμβατικές μέθοδοι εκπαίδευσης λειτουργούν με τέτοιο τρόπο, ώστε κάθε ζεύγος εισόδου εξόδου ενός μοντέλου να ταξινομείται σωστά στην κλάση στην οποία ανήκει. Ωστόσο, στην περίπτωση υπερεκπαίδευσης (overfitting) ενός ταξινομητή, το μοντέλο αρχίζει να απομνημονεύει τα δεδομένα εκπαίδευσης, αντί να μαθαίνει με βάση αυτά, υποβαθμίζοντας έτσι την ικανότητα γενίκευσής του. Αντιθέτως, οι Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines - SVM), στοχεύουν στον διαχωρισμό του συνόλου εκπαίδευσης σε κλάσεις με μία επιφάνεια που μεγιστοποιεί το περιθώριο ανάμεσά τους και κατ'επέκταση στη μεγιστοποίηση της ικανότητας γενίκευσης του μοντέλου [28].

Για τη γραμμικά διαχωρίσιμη περίπτωση, έστω n το πλήθος δείγματα εκπαίδευσης. Κάθε δείγμα αποτελείται από το ζεύγος του διανύσματος εισόδου x_i και της αντίστοιχης ετικέτας y_i . Επομένως, το σύνολο εκπαίδευσης έχει τη μορφή: $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, δηλαδή $X = \{x_i, y_i\}_{i=1}^n$, όπου $x_i \in \mathbb{R}^d$ και $y_i \in \{+1, -1\}$ (για λόγους καλύτερης οπτικοποίησης η είσοδος έχει θεωρηθεί διδιάστατη, δηλαδή $x \in \mathbb{R}^2$). Στην περίπτωση των γραμμικά διαχωρίσιμων δεδομένων, υπάρχουν άπειρα υπερεπίπεδα που μπορούν να επιτελέσουν τον διαχωρισμό. Ωστόσο, η ικανότητα γενίκευσης του μοντέλου εξαρτάται από τη θέση του βέλτιστου υπερεπιπέδου, δηλαδή εκείνου με το μέγιστο περιθώριο διαχωρισμού, το οποίο ορίζεται από την εξίσωση $w^T \cdot x_i + b = 0$. Η βελτιστοποίηση του γεωμετρικού περιθωρίου, επιτυγχάνεται θέτοντας το συναρτησιακό περιθώριο ίσο με 1, δηλαδή ίσο με τον γραμμικό ταξινομητή $y_i = 1$.

Οι εξισώσεις $\langle w \cdot x^+ \rangle + b = 1$ και $\langle w \cdot x^- \rangle + b = -1$ μπορούν να συνδυαστούν στην ανισότητα $y_i(\langle w \cdot x_i \rangle + b) \geq 1, \forall i$.

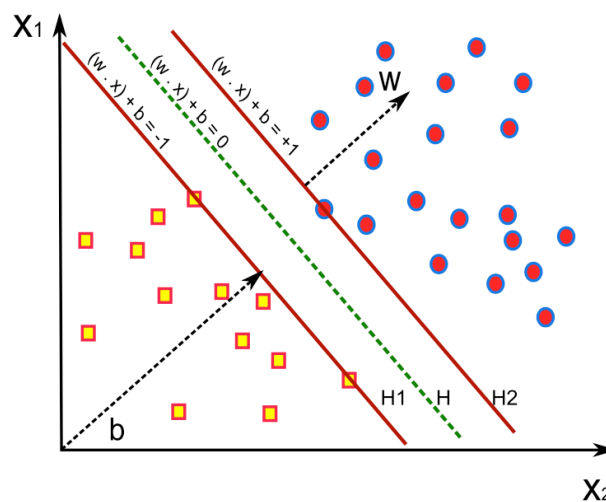
Το γεωμετρικό περιθώριο είναι

$$\gamma_i = \frac{1}{2} \left(\left\langle \frac{w}{\|w\|} \cdot x^+ \right\rangle - \left\langle \frac{w}{\|w\|} \cdot x^- \right\rangle \right) = \frac{1}{2\|w\|} [\langle w \cdot x^+ \rangle - \langle w \cdot x^- \rangle] = \frac{1}{\|w\|},$$

όπου το w συμβολίζει το βέλτιστο υπερεπίπεδο διαχωρισμού και το b την πόλωση. Η απόσταση μεταξύ του υπερεπιπέδου και των δεδομένων εκπαίδευσης που βρίσκονται εγγύτερα στο υπερεπίπεδο, αποτελεί το περιθώριο. Η ικανότητα γενίκευσης του μοντέλου μεγιστοποιείται αν το βέλτιστο υπερεπίπεδο επιλεγεί για τον διαχωρισμό. Η

βελτιστοποίηση του γεωμετρικού περιθωρίου συνεπάγεται την ελαχιστοποίηση της νόρμας του διανύσματος των βαρών. Επομένως, προκύπτει πρόβλημα τετραγωνικού προγραμματισμού για την εύρεση του βέλτιστου υπερεπιπέδου και των δύο παράλληλων υπερεπιπέδων H_1 και H_2 [28].

Η απόσταση μεταξύ των H_1 και H_2 μεγιστοποιείται όταν δεν υπάρχουν δεδομένα ανάμεσα στα δύο υπερεπιπέδα. Όταν η απόσταση αυτή μεγιστοποιείται, κάποια σημεία βρίσκονται πάνω από το H_1 και κάποια πάνω από το H_2 . Τα σημεία αυτά ονομάζονται διανύσματα υποστήριξης, καθώς συμμετέχουν άμεσα στον καθορισμό του βέλτιστου υπερεπιπέδου, ενώ τα υπόλοιπα σημεία μπορούν να διαγραφούν ή να μεταβληθούν χωρίς να διαπεραστούν τα υπερεπιπέδα H_1 και H_2 και χωρίς να μεταβληθεί η ικανότητα γενίκευσης του ταξινομητή. Κάθε υπερεπιπέδο μπορεί να οριστεί από τα w, x και b , όπου w ένα διάνυσμα που είναι κάθετο στο υπερεπιπέδο και το οποίο αποτελεί τη γεωμετρική αναπαράσταση του προβλήματος τετραγωνικού προγραμματισμού. Η Εικόνα 5 απεικονίζει τα υπερεπιπέδα H (βέλτιστο), H_1 και H_2 . Με τον τρόπο αυτό, το αρχικό πρόβλημα βελτιστοποίησης γίνεται:



Εικόνα 5: Το βέλτιστο υπερεπιπέδο (Μηχανές Διανυσμάτων Υποστήριξης)

Για τη γραμμικά διαχωρίσιμη περίπτωση $S = [(x_1, y_1), \dots, (x_l, y_l)]$, αν το (w, b) αποτελεί τη λύση $\min_{w,b} \langle w \cdot w \rangle = \|w\|^2$, με την προϋπόθεση $y_i(\langle w \cdot x_i \rangle + b) \geq 1$, τότε το μέγιστο περιθώριο δίνεται από το $\gamma = \frac{1}{\|w\|}$.

Το πρόβλημα μπορεί να μετατραπεί στο δυϊκό πρόβλημα με τη μέθοδο των πολλαπλασιαστών Lagrange, αφενός διότι διευκολύνει τους υπολογισμούς και αφετέρου, γιατί μετασχηματίζοντας το πρόβλημα με αυτόν τον τρόπο, τα δεδομένα εκπαίδευσης θα εμφανίζονται μόνο με τη μορφή εσωτερικού γινόμενου των διανυσμάτων, γεγονός που θα επιτρέψει τη γενίκευση της διαδικασίας για τη μη γραμμική περίπτωση. Η συνάρτηση κόστους λαμβάνει τη μορφή:

$$L(w, b, a) = \frac{1}{2} \langle w \cdot w \rangle - \sum_{i=1}^l a_i \cdot [y_i(\langle w \cdot x_i \rangle + b - 1)],$$

όπου a_i είναι οι πολλαπλασιαστές Lagrange.

Παραγωγίζοντας ως προς w και b , προκύπτει:

$$\frac{\partial L(w, b, a)}{\partial w} = w - \sum_{i=1}^l a_i \cdot y_i \cdot x_i = 0 \Rightarrow w = \sum_{i=1}^l a_i \cdot y_i \cdot x_i$$

και

$$\frac{\partial L(w, b, a)}{\partial b} = - \sum_{i=1}^l a_i \cdot y_i = 0 \Rightarrow \sum_{i=1}^l a_i \cdot y_i = 0$$

Επομένως, αντικαθιστώντας τις παραπάνω σχέσεις στη συνάρτηση κόστους, προκύπτει:

$$L(w, b, a) = \frac{1}{2} \langle w \cdot w \rangle - \sum_{i=1}^l a_i \cdot [y_i (\langle w \cdot x_i \rangle + b - 1)] \Rightarrow$$

$$L(w, b, a) = \frac{1}{2} \sum_{i,j=1}^l a_i \cdot y_i \cdot a_j \cdot y_j \langle x_i \cdot x_j \rangle - \sum_{i,j=1}^l a_i \cdot y_i \cdot a_j \cdot y_j \langle x_j \cdot x_i \rangle - \sum_{i=1}^l a_i \cdot y_i \cdot b + \sum_{i=1}^l a_i \Rightarrow$$

$$L(w, b, a) = \sum_{i,j=1}^l a_i \cdot y_i \cdot a_j \cdot y_j \langle x_i \cdot x_j \rangle + \sum_{i=1}^l a_i$$

Τα σημεία με $x_i \geq 0$ ονομάζονται διανύσματα υποστήριξης, ορίζουν τα υπερεπίπεδα H_1 και H_2 και βρίσκονται πιο κοντά στο υπερεπίπεδο διαχωρισμού. Σε όλα τα υπόλοιπα δεδομένα εκπαίδευσης $a_i = 0$.

Για τη γραμμικά διαχωρίσιμη περίπτωση $S = [(x_1, y_1), \dots, (x_l, y_l)]$, αν το a_i^* αποτελεί τη λύση στο τετραγωνικό πρόβλημα βελτιστοποίησης

$$\max_{a_i} - \frac{1}{2} \sum_{i,j=1}^l a_i \cdot y_i \cdot a_j \cdot y_j \langle x_i \cdot x_j \rangle + \sum_{i=1}^l a_i, \text{ με την προϋπόθεση } \sum_{i=1}^l a_i \cdot y_i = 0, \text{ τότε}$$

το $\|w\|^2$ ορίζει το ελάχιστο $w^* = \sum_{i=1}^l a_i^* \cdot y_i \cdot x_i$ και το γεωμετρικό περιθώριο

$$\gamma^* = \frac{1}{\|w^*\|} \text{ μεγιστοποιείται.}$$

Με βάση τις συνθήκες Karush-Kuhn-Tucker (KKT), οι οποίες διαδραματίζουν πολύ σημαντικό ρόλο στην εύρεση της βέλτιστης λύσης σε ένα γενικό πρόβλημα βελτιστοποίησης. Από τις συνθήκες αυτές για ένα γραμμικά διαχωρίσιμο σύνολο δεδομένων, αποδεικνύεται ότι $\|w^*\|^2 = \langle w^* \cdot w^* \rangle = \left(\sum_{i \in sv} a_i^* \right)^{-\frac{1}{2}}$.

Επομένως, η μέγιστη απόσταση ενός υπερεπιπέδου είναι $\frac{1}{\|w^*\|} = \left(\sum_{i \in sv} a_i^* \right)^{-\frac{1}{2}}$.

Ενώ τα προβλήματα με γραμμικά διαχωρίσιμη είσοδο είναι πολύ σπάνια, υπάρχουν προβλήματα στα οποία ένα γραμμικό υπερεπίπεδο διαχωρισμού μπορεί να προσφέρει ικανοποιητικά αποτελέσματα ακόμη και όταν υπάρχουν τομές στο σύνολο δεδομένων. Ωστόσο, οι παραπάνω προσεγγίσεις τετραγωνικού προγραμματισμού δεν μπορούν να αξιοποιηθούν στην περίπτωση τομών, καθώς δεν μπορεί να ικανοποιηθεί η συνθήκη $y_i (\langle w \cdot x_i \rangle + b) \geq 1$. Τα δεδομένα που βρίσκονται στην τομή δεν μπορούν να ταξινομηθούν σωστά και για τα εσφαλμένα ταξινομημένα δεδομένα x_i , το αντίστοιχο a_i θα τείνει στο άπειρο. Επομένως, για τον εντοπισμό του μέγιστου περιθωρίου θα χρειαστεί ένα χαλαρό περιθώριο και θα οριστούν μη αρνητικές μεταβλητές χαλάρωσης $\xi_i \geq 0$, ώστε

$y_i(\langle w \cdot x_i \rangle + b) \geq 1 - \xi_i, \forall i$. Με τη χρήση των μεταβλητών χαλάρωσης ξ_i υπάρχει πάντα εφικτή λύση, όμως για τα δεδομένα εκπαίδευσης x_i , αν ισχύει $0 < \xi_i < 1$, τότε τα δεδομένα δεν έχουν το μέγιστο περιθώριο, αλλά ταξινομούνται σωστά [28].

Στην περίπτωση μη γραμμικά διαχωρίσιμων δεδομένων, προκειμένου να μεγιστοποιηθεί ο χώρος ανάμεσα στις κλάσεις, ο αρχικός χώρος δεδομένων μετασχηματίζεται σε έναν χώρο πολλών διαστάσεων (feature space). Η βασική ιδέα της κατασκευής μη γραμμικών μηχανών διανυσμάτων υποστήριξης είναι ο μετασχηματισμός των διανυσμάτων εισόδου $x \in \mathbb{R}^n$ σε διανύσματα $\Phi(x)$ ενός χώρου πολλών διαστάσεων F , όπου το Φ αναπαριστά την απεικόνιση: $\mathbb{R}^n \rightarrow \mathbb{R}^f$ και η επίλυση ενός προβλήματος γραμμικής ταξινόμησης στον χώρο F με μηχανές διανυσμάτων υποστήριξης $x \in \mathbb{R}^n \rightarrow \Phi(x) = [\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)]^T \in \mathbb{R}^f$.

Για ένα σύνολο υποθέσεων $f(x) = \sum_{i=1}^l w_i \cdot \varphi_i(x) + b$, όπου $\varphi : X \rightarrow F$ είναι μία μη

γραμμική συνάρτηση απεικόνισης από τον χώρο εισόδου στον χώρο F . Με δυϊκή αναπαράσταση, το παραπάνω σύνολο υποθέσεων εκφράζεται ως γραμμικός συνδυασμός των δεδομένων εκπαίδευσης, επομένως ο κανόνας απόφασης μπορεί να γραφεί με τη

βοήθεια εσωτερικού γινομένου $f(x) = \sum_{i=1}^l a_i \cdot y_i \cdot \langle \varphi(x_i) \cdot \varphi(x) \rangle + b$.

Προκειμένου το γινόμενο $\langle \varphi(x_i) \cdot \varphi(x) \rangle$ να απεικονιστεί απευθείας ως συνάρτηση των δεδομένων εισόδου στον χώρο F , χρησιμοποιείται η συνάρτηση πυρήνα. Μία συνάρτηση πυρήνα K , αποτελεί συνάρτηση, τέτοια ώστε για κάθε $x, z \in X, K(x, z) = \langle \varphi(x) \cdot \varphi(z) \rangle$. Στην προσέγγιση αυτή είναι βασικό να βρεθεί μία συνάρτηση πυρήνα, που να μπορεί να αξιολογηθεί αποδοτικά. Τότε ο κανόνας απόφασης γίνεται:

$$f(x) = \sum_{i=1}^l a_i \cdot y_i \cdot K(x_i, x_j) + b.$$

Όσον αφορά τις υπερπαραμέτρους των μηχανών διανυσμάτων υποστήριξης, ορισμένες από αυτές παρουσιάζονται παρακάτω:

- Σταθερά ποινης C

Η σταθερά C, αποτελεί μία υπερπαραμέτρο ελέγχει το πλάτος του χαλαρού περιθωρίου, ορίζοντας αντίστοιχη ποινή στις μεταβλητές χαλάρωσης [28]. Η παράμετρος αυτή ελέγχει την αντιστάθμιση ανάμεσα στην πολυπλοκότητα του υπερεπιπέδου και των σφαλμάτων εκπαίδευσης και ελαττώνει την πιθανότητα υπερεκπαίδευσης [29].

- Πυρήνας

Η συνάρτηση πυρήνα $K(x, z) = \langle \varphi(x) \cdot \varphi(z) \rangle$ αποτελεί επίσης παράμετρο που πρέπει να προσδιοριστεί. Μερικές από τις συνηθέστερες συναρτήσεις πυρήνα είναι ο Γραμμικός Πυρήνας $K(x_i, x_j) = (x_i \cdot x_j)$, ο Πολυωνυμικός Πυρήνας $K(x_i, x_j) = (x_i \cdot x_j + 1)^p$, ο

Γκαουσιανός Πυρήνας $K(x_i, x_j) = e^{-\frac{\|x_i - x_j\|^2}{2 \cdot \sigma^2}}$, ο RBF Πυρήνας $K(x_i, x_j) = e^{-\gamma(x_i - x_j)^2}$ και ο

Σιγμοειδής Πυρήνας $K(x_i, x_j) = \tanh(\eta \cdot x_i \cdot x_j + \nu)$. Η επιλογή του κατάλληλου πυρήνα πρέπει να βασίζεται στις ιδιότητες του συνόλου δεδομένων και για τη βέλτιστη επίδοση του μοντέλου, πρέπει να προσδιοριστούν και οι βέλτιστες παράμετροι του πυρήνα που χρησιμοποιείται [28].

Οι Μηχανές Διανυσμάτων Υποστήριξης διακρίνονται για την ικανότητα τους για γενίκευση, όπως και για τη διακριτική τους ικανότητα. Αποτελούν ένα ισχυρό εργαλείο για την

επίλυση δυαδικών προβλημάτων και χάρη στο ισχυρό θεωρητικό τους υπόβαθρο, αλλά και την ικανότητά τους για γενίκευση, συμπεριλαμβάνονται στις πιο δημοφιλείς μεθόδους ταξινόμησης. Η μέθοδος αυτή ελαχιστοποιεί τα σφάλματα της ταξινόμησης των δεδομένων εκπαίδευσης και διαφέρει σημαντικά από άλλες μεθόδους, ειδικότερα όταν το σύνολο δεδομένων εκπαίδευσης έχει μικρό μέγεθος. Ωστόσο, σε περιπτώσεις σημαντικής ανισορροπίας των δεδομένων, η επίδοση του μοντέλου επηρεάζεται σε μεγάλο βαθμό, καθώς είναι δύσκολο να βρεθεί το βέλτιστο υπερεπίπεδο διαχωρισμού [28].

3.1.1.4 Gradient Boosting

Στο μοντέλο Gradient Boosting Machines (GBM), η διαδικασία μάθησης προσαρμόζει διαδοχικά νέα μοντέλα, ώστε να παρέχει μία ακριβέστερη εκτίμηση της μεταβλητής απόκρισης. Η κύρια ιδέα του αλγορίθμου είναι η κατασκευή νέων βασικών μοντέλων μάθησης, τα οποία να συσχετίζονται στον μέγιστο βαθμό με την αρνητική κλίση της συνάρτησης σφάλματος που αφορά ολόκληρο το σύνολο των μοντέλων μάθησης. Η συνάρτηση σφάλματος που εφαρμόζεται μπορεί να είναι αυθαίρετη, ωστόσο αν οριστεί ως η τετραγωνική απώλεια σφάλματος, η διαδικασία μάθησης μπορεί να οδηγηθεί σε διαδοχική προσαρμογή σφάλματος [30].

Όσον αφορά στο πρόβλημα της εκτίμησης συναρτήσεων στην επιβλεπόμενη μάθηση, έστω το σύνολο δεδομένων $(x, y)_{i=1}^N$, όπου $x = (x_1, \dots, x_d)$ οι μεταβλητές εισόδου και y οι αντίστοιχες ετικέτες της μεταβλητής απόκρισης. Ο στόχος είναι η ανακατασκευή της άγνωστης συναρτησιακής εξάρτησης $x \rightarrow^f y$, με την εκτίμηση $\hat{f}(x)$, τέτοια ώστε κάποια δοσμένη συνάρτηση $\Psi(y, f)$ να ελαχιστοποιείται: $\hat{f}(x) = y$, $\hat{f}(x) = \operatorname{argmin}_{f(x)} \Psi(y, f(x))$. Αν το πρόβλημα εκτίμησης γραφεί εκ νέου με όρους προσδοκόμενων τιμών, η ισοδύναμη σχέση θα αφορούσε την ελαχιστοποίηση της προσδοκόμενης συνάρτησης απώλειας έναντι της μεταβλητής απόκρισης $E_y(\Psi[y, f(x)])$ με την προϋπόθεση των παρατηρούμενων δεδομένων x :

$$\hat{f}(x) = \operatorname{argmin}_{f(x)} E_x[E_y(\Psi[y, f(x)]) | x].$$

Η μεταβλητή απόκρισης y μπορεί να προέρχεται από διαφορετικές κατανομές, γεγονός που συνεπάγεται και τον ορισμό διαφορετικών συναρτήσεων απώλειας Ψ . Συγκεκριμένα, αν η μεταβλητή απόκρισης είναι δυαδική και ανήκει για παράδειγμα στο $y \in \{0, 1\}$, τότε χρησιμοποιείται διωνυμική συνάρτηση απώλειας, ενώ αν είναι συνεχής, μπορεί να χρησιμοποιείται, μεταξύ άλλων, η τετραγωνική συνάρτηση σφάλματος [30].

Προκειμένου να καταστεί το πρόβλημα προσιτό, μπορεί ο χώρος αναζήτησης των συναρτήσεων να περιοριστεί σε μία παραμετρική οικογένεια συναρτήσεων $f(x, \theta)$. Αυτό θα μετέτρεπε το πρόβλημα βελτιστοποίησης της συνάρτησης, σε πρόβλημα εκτίμησης της παραμέτρου, με τον ακόλουθο τρόπο: $\hat{f}(x) = f(x, \hat{\theta})$ και $\hat{\theta} = \operatorname{argmin}_{\theta} E_x[E_y(\Psi[y, f(x, \theta)]) | x]$. Προκειμένου να εκτελεστούν υπολογισμοί, απαιτούνται επαναληπτικές αριθμητικές διαδικασίες [30].

Έστω M βήματα επανάληψης, η εκτίμηση των παραμέτρων μπορεί να γραφεί: $\hat{\theta} = \sum_{i=1}^M \hat{\theta}_i$.

Η απλούστερη και συχνότερα χρησιμοποιούμενη διαδικασία εκτίμησης παραμέτρων είναι ο αλγόριθμος απότομης καθόδου, γνωστός και ως κάθοδος βασισμένη στην κλίση (gradient descent). Με δεδομένα N σημεία δεδομένων $(x, y)_{i=1}^N$, απαιτείται η ελάττωση της εμπειρικής συνάρτησης απώλειας $J(\theta)$ για τα παρατηρούμενα δεδομένα:

$$J(\theta) = \sum_{i=1}^N \Psi(y_i, f(x_i, \hat{\theta})).$$

Η κλασική διαδικασία βελτιστοποίησης με τη μέθοδο απότομης καθόδου βασίζεται σε διαδοχικές βελτιώσεις στην κατεύθυνση της κλίσης της συνάρτησης απωλειών $\nabla J(\theta)$. Με τον κάτω δείκτη $\hat{\theta}_t$ συμβολίζεται το t-οστό βήμα της εκτίμησης $\hat{\theta}$, ενώ με τον άνω δείκτη $\hat{\theta}^t$ η συνολική εκτίμηση όλου του συνόλου, δηλαδή το άθροισμα όλων των εκτιμήσεων από το βήμα 1 έως το βήμα t. Η διαδικασία βελτιστοποίησης με την μέθοδο της απότομης καθόδου είναι: Αρχικοποίηση των εκτιμήσεων $\hat{\theta}_0$ των παραμέτρων και για κάθε επανάληψη t, υπολογισμός μίας συνολικής εκτίμησης $\hat{\theta}^t$ από όλες τις προηγούμενες επαναλήψεις:

$$\hat{\theta}^t = \sum_{i=0}^{t-1} \hat{\theta}_i, \text{ αξιολόγηση της κλίσης της συνάρτησης απωλειών } \nabla J(\theta), \text{ δεδομένων των}$$

εκτιμήσεων των παραμέτρων που έχουν προκύψει: $\nabla J(\theta) = \{ \nabla J(\theta_i) \} = \left[\frac{\partial J(\theta)}{\partial J(\theta_i)} \right]_{\theta=\hat{\theta}^t}$.

Το επόμενο βήμα είναι ο υπολογισμός της νέας εκτίμησης παραμέτρων $\hat{\theta}_t$: $\hat{\theta}_t \leftarrow - \nabla J(\theta)$ και η προσθήκη της νέας εκτίμησης $\hat{\theta}_t$ στο σύνολο [30].

Η βασική διαφορά ανάμεσα στις μεθόδους ενίσχυσης (boosting) και τις συμβατικές τεχνικές μηχανικής μάθησης, έγκειται στο γεγονός ότι η βελτιστοποίηση πραγματοποιείται στον χώρο των συναρτήσεων, καθώς παραμετροποιείται η εκτίμηση συνάρτησης \hat{f} στην

$$\text{ακόλουθη μορφή: } \hat{f}(x) = \hat{f}(x)^M = \sum_{i=0}^M \hat{f}_i(x).$$

Στην παραπάνω αναπαράσταση, M είναι το πλήθος των επαναληπτικών βημάτων, \hat{f}_0 είναι η αρχική υπόθεση και $\{\hat{f}_i\}_{i=1}^M$ οι προσαυξήσεις της συνάρτησης, οι οποίες ονομάζονται “boosts”.

Προκειμένου να καταστεί εφικτή στην πράξη η προσέγγιση αυτή, μπορεί να ακολουθηθεί αντίστοιχη στρατηγική παραμετροποίησης της οικογένειας συναρτήσεων. Στο σημείο αυτό εισάγονται οι παραμετροποιημένες βασικές συναρτήσεις μάθησης $h(x, \theta)$, ώστε να υπάρχει διαχωρισμός από τις συνολικές εκτιμήσεις συναρτήσεων $\hat{f}(x)$. Μπορούν να επιλεγούν διάφορες οικογένειες βασικών συναρτήσεων μάθησης, όπως για παράδειγμα δέντρα απόφασης [30].

Στη συνέχεια διατυπώνεται η άπληστη προσέγγιση επαύξησης της συνάρτησης με τις βασικές συναρτήσεις μάθησης. Για τον σκοπό αυτό καθορίζεται το βέλτιστο μέγεθος βήματος σε κάθε επανάληψη. Για την εκτίμηση συνάρτησης στην t-οστή επανάληψη, ο κανόνας βελτιστοποίησης

$$\text{ο ρί ζ ε τ α ι} \quad \omega \varsigma : \quad \hat{f}_t \leftarrow \hat{f}_{t-1} + \rho_t \cdot h(x, \theta_t) \quad \mu \epsilon$$

$$(\rho_t, \theta_t) = \underset{\rho, \theta}{\operatorname{argmin}} \sum_{i=1}^N \Psi(y_i, \hat{f}_{t-1} + \rho \cdot h(x_i, \theta)).$$

Όσον αφορά τον αλγόριθμο Gradient Boosting, τόσο η συνάρτηση απωλειών, όσο και η βασική συνάρτηση μάθησης (base learner) μπορούν να καθοριστούν αυθαίρετα. Με δεδομένη μία συγκεκριμένη συνάρτηση απωλειών $\Psi(y, f)$ και μία βασική συνάρτηση μάθησης $h(x, \theta)$, η λύση για τις εκτιμήσεις παραμέτρων μπορεί να είναι δύσκολο να προσδιοριστεί. Προκειμένου να αντιμετωπιστεί το ζήτημα αυτό, έχει προταθεί η επιλογή μίας νέας συνάρτησης $h(x, \theta_t)$, η οποία να είναι παράλληλη στην αρνητική κλίση $\{\hat{g}_t(x_i)\}_{i=1}^N$ μαζί με τα παρατηρούμενα δεδομένα:

$$g_t(x) = E_y \left[\frac{\partial \Psi(y, f(x))}{\partial f(x)} \mid x \right]_{f(x)=\hat{f}^{t-1}(x)}.$$

Αντί για τη γενική λύση, μπορεί να επιλεγεί μία νέα συνάρτηση επαύξησης, η οποία να έχει τη μεγαλύτερη συσχέτιση με τη $-g_t(x)$, γεγονός που επιτρέπει την αντικατάσταση μίας πιθανά δύσκολης διαδικασίας βελτιστοποίησης με την κλασική μέθοδο ελαχιστοποίησης

$$\text{ελαχίστων τετραγώνων: } (\rho_t, \theta_t) = \underset{\rho, \theta}{\operatorname{argmin}} \sum_{i=1}^N [-g_t(x_i) + \rho \cdot h(x_i, \theta)]^2.$$

Επομένως, τα βήματα του αλγορίθμου gradient boosting είναι: η αρχικοποίηση της αρχικής υπόθεσης \hat{f}_0 με μία σταθερά και για M επαναλήψεις υπολογισμός της αρνητικής κλίσης $g_t(x)$, προσαρμογή στη νέα βασική συνάρτηση μάθησης $h(x, \theta_t)$, εύρεση του βέλτιστου μεγέθους βήματος της καθόδου βασισμένης στην κλίση (gradient descent) και η ενημέρωση της εκτίμησης συνάρτησης $\hat{f}_t \leftarrow \hat{f}_{t-1} + \rho_t \cdot h(x, \theta_t)$.

Για τον σχεδιασμό ενός μοντέλου Gradient Boosting Machine, πρέπει να καθοριστούν οι συναρτησιακές παράμετροι $\Psi(y, f)$ και $h(x, \theta)$, δηλαδή πρέπει να προσδιοριστεί ποιο στοιχείο απαιτεί βελτιστοποίηση και ποια θα είναι η μορφή της συνάρτησης που θα χρησιμοποιηθεί για την κατασκευή της λύσης. Οι παραπάνω επιλογές είναι καθοριστικές για τις ιδιότητες του μοντέλου [30].

Όσον αφορά τις βασικές συναρτήσεις μάθησης (base learners), στο μοντέλο Gradient Boosting Machine είθισται να χρησιμοποιούνται γραμμικές, ομαλές συναρτήσεις ή δέντρα απόφασης. Η χρήση των δέντρων απόφασης είναι συνήθης και συνιστά έναν υπολογιστικά εφικτό τρόπο να αποτυπωθούν οι αλληλεπιδράσεις μεταξύ των μεταβλητών στο μοντέλο [30].

Στα δέντρα απόφασης, όπως περιγράφηκε παραπάνω, πραγματοποιείται διαμέριση του χώρου των μεταβλητών εισόδου σε κελιά με βάση τον κανόνα απόφασης του συστήματος. Κάθε διάσπαση αντιστοιχεί σε μία συνθήκη “αν-τότε” για μία μεταβλητή εισόδου. Η δομή ενός δέντρου απόφασης κωδικοποιεί και μοντελοποιεί τις αλληλεπιδράσεις μεταξύ των μεταβλητών πρόβλεψης. Τα δέντρα αυτά, συχνά παραμετροποιούνται με το πλήθος των διασπάσεων και τις αντίστοιχες υπερπαραμέτρους που περιγράφηκαν παραπάνω [30].

3.1.1.5 XGBoost

Το μοντέλο XGBoost αποτελεί μία εξελιγμένη εφαρμογή του Gradient Boosting και χρησιμοποιεί στα δέντρα απόφασης ως βασικό ταξινομητή. Χρησιμοποιείται συχνά χάρη στην ταχύτητα, την απόδοση και τη δυνατότητα κλιμάκωσής του [31].

Για ένα σύνολο δεδομένων, χρησιμοποιούνται αθροιστικές συναρτήσεις για την πρόβλεψη της εξόδου. Για μία δεδομένη είσοδο, η τελική πρόβλεψη υπολογίζεται αθροίζοντας τα βάρη των φύλλων του δέντρου. Η βασική ιδέα του μοντέλου Gradient Boosting είναι η ελαχιστοποίηση της συνάρτησης απωλειών [31]. Η μεγαλύτερη διαφορά από το Gradient Boosting μοντέλο που χρησιμοποιεί Δέντρα Απόφασης, έγκειται στο γεγονός ότι στο GBM η συνάρτηση απωλειών χρησιμοποιεί μόνο την πρώτη παράγωγο της συνάρτησης απωλειών κατά τον υπολογισμό της αντικειμενικής συνάρτησης. Το XGBoost προσεγγίζει τη συνάρτηση απωλειών με τη χρήση του αναπτύγματος Taylor δεύτερης τάξης:

$$f(x + \Delta x) \cong f(x) + f'(x)\Delta x + \frac{1}{2} f''(x)\Delta x^2.$$

Ένα δέντρο στο μοντέλο XGBoost, διασπάται στα δομικά μέρη q και στα φύλλα w .

$$\varphi(x) = w_{q(x)}, w \in \mathbb{R}^T q : \mathbb{R}^D \rightarrow \{1, \dots, t\},$$

όπου q η συνάρτηση δομής, στην οποία ο δείκτης υποδεικνύει την είσοδο στο φύλλο, w το κλάσμα των φύλλων που αντιστοιχεί σε κάθε δείκτη, T το πλήθος των κόμβων που αποτελούν φύλλα στο δέντρο και D η διάσταση των χαρακτηριστικών [32].

Η πολυπλοκότητα του δέντρου ορίζεται ως:

$$\Omega(\varphi(x)) = \gamma \cdot T + \frac{1}{2} \lambda \sum_{t=1}^T w_t^2, \text{ όπου } \gamma \text{ είναι ο κανονικός συντελεστής L1 και ορίζει την}$$

ελάχιστη μείωση απωλειών προκειμένου να επιτραπεί περαιτέρω διαμέριση, λ ο κανονικός συντελεστής L2 με τα γ και λ να καθορίζουν την πολυπλοκότητα του δέντρου [31], [32].

Προσθέτοντας τον όρο κανονικοποίησης Ω στη συνάρτηση απωλειών, προστίθεται ποινή στην πολυπλοκότητα του μοντέλου. Αν η ελάττωση των απωλειών είναι μικρότερη του γ , τότε το XGBoost σταματά. Ομοίως με το γ , υπάρχει υπερπαράμετρος που ελέγχει το βάθος του δέντρου, και της οποίας αν η τιμή είναι σημαντικά μεγάλη μπορεί να καταστήσει το μοντέλο πιο “συντηρητικό” όσον αφορά τις διασπάσεις [31].

Επομένως, η αντικειμενική συνάρτηση είναι:

$$\begin{aligned} J(\theta) &= \sum_{i=1}^N L(f(x_i; \theta), y_i) + \Omega(\theta) \\ J(\theta) &\cong \sum_{i=1}^N [g_{m,i} \varphi(x_i) + \frac{1}{2} h_{m,i} \varphi(x_i)^2] + \gamma \cdot T + \frac{1}{2} \lambda \sum_{t=1}^T w_t^2 \\ J(\theta) &\cong \sum_{i=1}^T [\sum_{i \in I_t} g_{m,i} w_t + \frac{1}{2} \sum_{i \in I_t} h_{m,i} w_t^2 + \frac{1}{2} \lambda \sum_{t=1}^T w_t^2] + \gamma \cdot T \\ J(\theta) &\cong \sum_{i=1}^T [\sum_{i \in I_t} g_{m,i} w_t + \frac{1}{2} (\sum_{i \in I_t} h_{m,i} + \lambda) w_t^2] + \gamma \cdot T, \end{aligned}$$

όπου η $g_{m,i} = [\frac{\partial^2 L(f(x_i), y_i)}{\partial^2 f(x_i)}]_{f=f_{m-1}}$ αντιπροσωπεύει την πρώτη παράγωγο της συνάρτησης απωλειών και η $h_{m,i} = [\frac{\partial^2 L(f(x_i), y_i)}{\partial^2 f(x_i)}]_{f=f_{m-1}}$ τη δεύτερη παράγωγο.

Με δεδομένα τη δομή του γνωστού δέντρου q , την εξαγωγή της αντικειμενικής

συνάρτησης και την παράγωγο, το αποτέλεσμα είναι ίσο με 0: $w_t^* = - \frac{\sum_{i \in I_{T-t}} g_{m,i}}{\sum_{i \in I_t} h_{m,i} + \lambda}$

Θέτοντας $w_t = w_t^*$, τελικά προκύπτει: $J(\theta) = - \frac{1}{2} \sum_{t=1}^T \frac{(\sum_{i \in I_{T-t}} g_{m,i})}{\sum_{i \in I_t} h_{m,i} + \lambda} + \lambda T$

Από την παραπάνω σχέση, όσο πιο μικρό είναι το w_t , τόσο μικρότερη είναι η αντικειμενική συνάρτηση και όσο υψηλότερη είναι η ακρίβεια της πρόβλεψης, τόσο καλύτερη είναι η ανθεκτικότητα του μοντέλου χάρη στην ύπαρξη τακτικών στοιχείων [32].

Εφόσον η βασική συνάρτηση μάθησης του Gradient Boosting πρόκειται για δέντρο απόφασης, αρκετές από τις υπερπαραμέτρους σχετίζονται με τη δομή των δέντρων.

Όπως περιγράφηκε παραπάνω (βλ. 3.1.1.1), σε αυτές συμπεριλαμβάνονται η υποδειγματοληψία, το μέγιστο πλήθος φύλλων και το μέγιστο βάθος του δέντρου, των οποίων η επιλογή στοχεύει στην ελάττωση της υπερεκπαίδευσης και στη βελτίωση του μοντέλου. Η υποδειγματοληψία με βάση τις στήλες (ή αλλιώς με βάση τα χαρακτηριστικά - `colsample_bytree`) που χρησιμοποιείται στο XGBoost μπορεί να καταστεί πιο αποδοτική στην αποφυγή της υπερεκπαίδευσης από την κλασική υποδειγματοληψία με βάση τις γραμμές που εφαρμόζεται στο απλό μοντέλο Gradient Boosting. Ο ρυθμός μάθησης που ελέγχει την ανάθεση βαρών των νέων δέντρων εντάσσεται επίσης στο μοντέλο και βοηθά στην ελάττωση του ρυθμού με τον οποίο το μοντέλο προσαρμόζεται στα δεδομένα εκπαίδευσης [31].

3.1.2 Αλγόριθμοι Βαθιάς Μάθησης

3.1.2.1 Πολυστρωματικό Perceptron

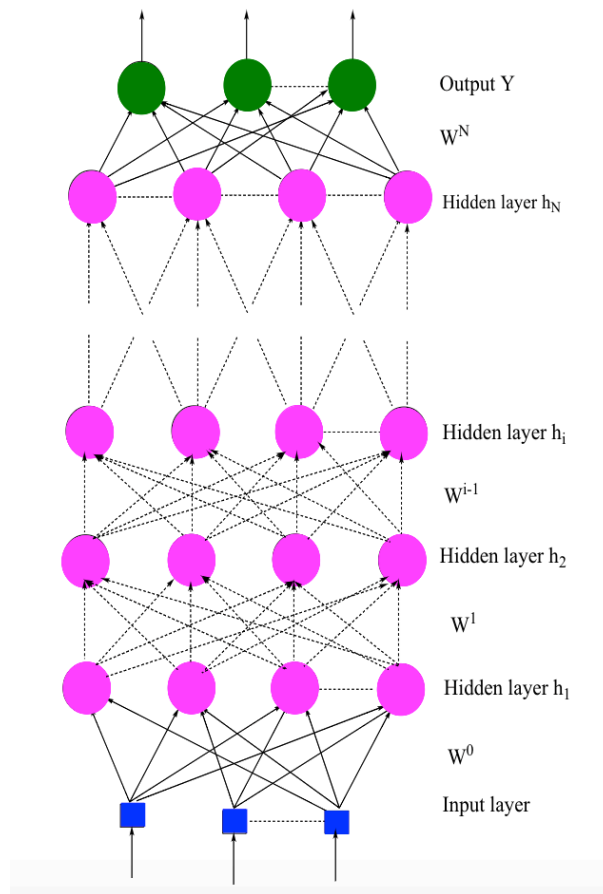
Τα τελευταία χρόνια, τα νευρωνικά δίκτυα έχουν προσελκύσει σημαντική προσοχή, καθώς έχουν αποδειχθεί πολύ σημαντικά σε προβλήματα ταξινόμησης και παλινδρόμησης και συγκεκριμένα σε εφαρμογές όπως η αναγνώριση και βελτιστοποίηση προτύπων. Το Πολυστρωματικό Perceptron (Multilayer Perceptron - MLP) αποτελεί το πιο διαδεδομένο μοντέλο νευρωνικών δικτύων που χρησιμοποιεί τον αλγόριθμο οπισθοδιάδοσης για την εκπαίδευση [33].

Τα δεδομένα ενός προβλήματος ταξινόμησης προτύπων αποτελούνται από ένα πλήθος προτύπων και την σωστή ταξινόμηση τους, ενώ κάθε πρότυπο απαρτίζεται από ένα πλήθος μετρήσεων. Στόχος είναι η κατασκευή ενός ταξινομητή που λαμβάνει τις μετρήσεις ενός προτύπου και παρέχει τη σωστή ταξινόμησή του ως έξοδο. Ένας δημοφιλής τύπος ταξινομητή είναι τα νευρωνικά δίκτυα πρόσθιας τροφοδότησης [33].

Ένα νευρωνικό δίκτυο πρόσθιας τροφοδότησης αποτελείται από ένα επίπεδο εισόδου που απαρτίζεται από νευρώνες, ένα αυθαίρετο πλήθος κρυφών επιπέδων και ένα επίπεδο εξόδου. Τα συγκεκριμένα νευρωνικά δίκτυα για τον σκοπό της ταξινόμησης προτύπων, αποτελούνται από τόσους νευρώνες εισόδου, όσες οι μετρήσεις των προτύπων του συνόλου δεδομένων, δηλαδή για κάθε μέτρηση υπάρχει αντίστοιχα ένας ακριβώς νευρώνας εισόδου. Το επίπεδο εξόδου αποτελείται από τόσους νευρώνες, όσες και οι κλάσεις ταξινόμησης. Δεδομένων των βαρών όλων των συνδέσεων μεταξύ των νευρώνων, προκειμένου να ταξινομηθεί ένα πρότυπο, οι μετρήσεις παρέχονται ως είσοδος στους νευρώνες εισόδου και τα σήματα εξόδου διαδίδονται από επίπεδο σε επίπεδο, ώσπου να προκύψουν τα σήματα εξόδου των νευρώνων εξόδου. Κάθε νευρώνας εξόδου ταυτίζεται με μία από τις πιθανές κλάσεις του προβλήματος και ο νευρώνας εξόδου που παράγει το υψηλότερο σήμα εξόδου ταξινομεί το αντίστοιχο πρότυπο [33].

Το MLP αποτελεί μία παραλλαγή του Perceptron, η οποία έχει ένα ή περισσότερα κρυφά επίπεδα ανάμεσα στα επίπεδα εισόδου και εξόδου. Οι νευρώνες είναι οργανωμένοι σε επίπεδα, των οποίων οι συνδέσεις έχουν την κατεύθυνση από τα χαμηλότερα στα υψηλότερα επίπεδα, ενώ οι νευρώνες που ανήκουν στο ίδιο επίπεδο δεν έχουν μεταξύ τους συνδέσεις. Το πλήθος των νευρώνων στο επίπεδο εισόδου ισούται με το πλήθος των μετρήσεων του προβλήματος προτύπων και το πλήθος των νευρώνων στο επίπεδο εξόδου ισούται με τον αριθμό των κλάσεων [33].

Η εκπαίδευση ενός MLP συνίσταται στην διαδικασία προσαρμογής των βαρών των συνδέσεων, ώστε να προκύπτει η ελάχιστη δυνατή διαφορά ανάμεσα στην έξοδο του δικτύου και στην επιθυμητή έξοδο.



Εικόνα 6: Η δομή ενός νευρωνικού δικτύου πρόσθιας τροφοδότησης.

Έστω ένα επίπεδο εισόδου με n_0 το πλήθος νευρώνες $X = (x_0, x_1, \dots, x_{n_0})$ και μία σιγμοειδής συνάρτηση ενεργοποίησης $f(x) = \frac{1}{1 + e^{-x}}$. Για να προκύψει η έξοδος του δικτύου πρέπει πρώτα να υπολογιστεί η έξοδος σε κάθε νευρώνα κάθε επιπέδου. Έστω ένα σύνολο κρυφών επιπέδων (h_1, h_2, \dots, h_N) , υποθέτοντας ότι n_i είναι το πλήθος των νευρώνων σε κάθε κρυφό επίπεδο h_i , τότε για την έξοδο του πρώτου κρυφού επιπέδου έχουμε:

$$h_i^j = f\left(\sum_{k=1}^{n_{i-1}} w_{k,j}^0 \cdot x_k\right), \quad j = 1, \dots, n_i.$$

Οι έξοδοι h_i^j των νευρώνων στα κρυφά επίπεδα υπολογίζονται ως εξής:

$$h_i^j = f\left(\sum_{k=1}^{n_{i-1}} w_{k,j}^{i-1} \cdot h_{i-1}^k\right), \quad i = 2, \dots, N \text{ και } j = 1, \dots, n_i$$

Όπου $w_{k,j}^i$ είναι το βάρος ανάμεσα στον νευρώνα k στο κρυφό επίπεδο i και στον νευρώνα j στο κρυφό επίπεδο $i+1$ και n_i είναι το πλήθος των νευρώνων στο i -οστό κρυφό επίπεδο. Η έξοδος του i -οστού επιπέδου υπολογίζεται ως εξής: $h_i = \iota(h_i^1, h_i^2, \dots, h_i^{n_i})$.

Η έξοδος του δικτύου υπολογίζεται ως: $y_i = f\left(\sum_{k=1}^{n_N} w_{k,j}^N \cdot h_N^k\right)$ με

$$Y = (y_1, \dots, y_j, \dots, y_{N+1}) = F(W, X),$$

όπου $w_{k,j}^N$ είναι το βάρος ανάμεσα στον νευρώνα k στο N -οστό κρυφό επίπεδο και στον νευρώνα j στο επίπεδο εξόδου, n_N το πλήθος των νευρώνων στο N -οστό κρυφό επίπεδο, Y το διάνυσμα του επιπέδου εξόδου, F η συνάρτηση μεταφοράς και W ο πίνακας των βαρών που ορίζεται ως:

$$W = [W^0, \dots, W^j, \dots, W^N]$$

$$W^i = (w_{j,k}^i) \quad 0 \leq i \leq N \text{ με } w_{j,k}^i \in \mathbb{R}.$$

$$1 \leq j \leq n_{i+1}$$

$$1 \leq k \leq n_i$$

Απλουστεύοντας μπορεί να θεωρηθεί ότι $n = n_i \forall i = 1, \dots, N$ για κάθε κρυφό επίπεδο, όπου X είναι η είσοδος του νευρωνικού δικτύου, f η συνάρτηση ενεργοποίησης, W^i ο πίνακας των βαρών ανάμεσα στο i -οστό και στο $(i+1)$ -οστό κρυφό επίπεδο για $i = 1, \dots, N-1$, W^0 ο πίνακας βαρών ανάμεσα στο επίπεδο εισόδου και στο πρώτο κρυφό επίπεδο και W^N ο πίνακας βαρών ανάμεσα στο N -οστό κρυφό επίπεδο και στο επίπεδο εξόδου [33].

Αξίζει να σημειωθεί πως η επιλογή της αρχιτεκτονικής και άρα των υπερπαραμέτρων του νευρωνικού δικτύου επηρεάζει σε μεγάλο βαθμό τη σύγκλιση του. Η αρχιτεκτονική ενός MLP καθορίζεται από την επιλογή του πλήθους των επιπέδων, του πλήθους των κρυφών νευρώνων σε καθένα από αυτά τα επίπεδα και της επιλογής της αντικειμενικής συνάρτησης. Μικρός αριθμός συνδέσεων μπορεί να καταστήσει το δίκτυο μη ικανό για την επίλυση του προβλήματος εξαιτίας ανεπαρκών προσαρμοσίμων παραμέτρων, ενώ ένα υπερβολικά μεγάλο πλήθος νευρώνων μπορεί να προκαλέσει υπερεκπαίδευση του μοντέλου στα δεδομένα εκπαίδευσης [33].

3.1.3 Συλλογική Μάθηση

Η ανισορροπία των κλάσεων στα προβλήματα ταξινόμησης αποτελεί ένα σύννητες και απαιτητικό πρόβλημα και αφορά τις περιπτώσεις όπου μία από τις κλάσεις των δεδομένων περιλαμβάνει σημαντικά περισσότερα στιγμιότυπα από άλλες. Η κλάση με τα περισσότερα δείγματα ονομάζεται κλάση πλειοψηφίας, ενώ οι μικρότερες κλάσεις μειοψηφίας [34].

Οι μέθοδοι που χρησιμοποιούνται για την αντιμετώπιση τέτοιων προβλημάτων διακρίνονται σε τρεις κατηγορίες: στο επίπεδο του αλγορίθμου, στο επίπεδο των δεδομένων και στη μάθηση που είναι ευαίσθητη στο κόστος. Η προσέγγιση στο επίπεδο του αλγορίθμου αφορά τη δημιουργία ή την τροποποίηση του αλγορίθμου ώστε να λαμβάνει υπόψιν με κατάλληλο τρόπο τις επιμέρους κλάσεις. Σε μεθόδους που δρουν στο επίπεδο των δεδομένων προστίθεται ένα βήμα δειγματοληψίας κατά το στάδιο της προεπεξεργασίας. Η επαναδειγματοληψία εξισορροπεί στην κατανομή των δεδομένων μειώνοντας την επίδραση της ανισορροπίας κατανομής δεδομένων στο μοντέλο. Στην προσέγγιση μάθησης ευαίσθητης στο κόστος, πραγματοποιείται συνδυασμός αλγορίθμων και μεθόδων στο επίπεδο των δεδομένων, ώστε να ενσωματωθεί κόστος στα σφάλματα ταξινόμησης κάθε κλάσης στα σύνολα εκπαίδευσης [34].

Οι μέθοδοι συλλογικής μάθησης (ensemble learning) στοχεύουν στη βελτίωση της προβλεπτικής ικανότητας ενός μοντέλου και λειτουργούν συνδυάζοντας ένα σύνολο από μεμονωμένους ταξινομητές, οι οποίοι αφού εκπαιδευτούν, τα αποτελέσματα της πρόβλεψής τους συνδυάζονται ώστε να προκύψει η τελική πρόβλεψη με κάποιο σχήμα ψήφου [34].

3.1.3.1 Bagging

Από τις πιο διαδεδομένες μεθόδους συλλογικής μάθησης είναι οι bagging και boosting, οι οποίες επιτυγχάνουν βελτίωση της ακρίβειας ταξινόμησης. Η τεχνική bagging είναι ένας συνδυασμός των μεθόδων bootstrap και aggregating και έχει τη δυνατότητα να ελαττώσει τη διακύμανση των αποτελεσμάτων της ταξινόμησης. Οι μέθοδοι στο επίπεδο των δεδομένων διακρίνονται σε τρεις κατηγορίες, τις βασισμένες σε boosting, τις βασισμένες σε bagging και τις υβριδικές, ενώ παρακάτω θα αναλυθούν οι τεχνικές βασισμένες σε bagging [34].

- Μέθοδος Δειγματοληψίας

Η προσέγγιση της δειγματοληψίας αποτελεί μία τεχνική που δεν απαιτεί την τροποποίηση του αλγορίθμου και η οποία χρησιμοποιείται πολύ συχνά για την αντιμετώπιση προβλημάτων με ανισόρροπες κλάσεις. Η τεχνική αυτή μεταβάλλει την κατανομή των δεδομένων ή τροποποιεί το σύνολο δεδομένων καθιστώντας το ισορροπημένο. Η διαδικασία της επαναδειγματοληψίας πραγματοποιείται στο στάδιο της προεπεξεργασίας και διακρίνεται σε δύο είδη, στην υποδειγματοληψία (undersampling), κατά την οποία αφαιρούνται κάποια δεδομένα με τυχαίο τρόπο από την κλάση πλειοψηφίας και στην υπερδειγματοληψία (oversampling), κατά την οποία προστίθενται ή αντιγράφονται τυχαία κάποια στιγμιότυπα των κλάσεων μειοψηφίας, ώστε να εξισορροπηθεί το πλήθος των δεδομένων σε κάθε κλάση. Μειονέκτημα της υποδειγματοληψίας αποτελεί η απώλεια πληροφορίας που προκύπτει αφαιρώντας τα δεδομένα, ενώ το πρόβλημα που προκύπτει από την υπερδειγματοληψία σχετίζεται με την υπερεκπαίδευση, αφού παράγεται μεγάλο πλήθος δεδομένων [34].

- Μέθοδοι συλλογικής μάθησης βασισμένες στο bagging

Το bagging αποτελεί μία σύντηξη των bootstrap και aggregating, με το bootstrap ή αλλιώς δειγματοληψία να εκτελείται στα αρχικά δεδομένα n φορές με αντικατάσταση, ώστε να δημιουργηθούν τα σύνολα εκπαίδευσης. Μετά την εκπαίδευση των μεμονωμένων ταξινομητών, ξεκινά τη διαδικασία της άθροισης (aggregating), είτε τη διαδικασία των μέγιστων ψήφων στην περίπτωση της ταξινόμησης, είτε με τον υπολογισμό του μέσου όρου στην περίπτωση της παλινδρόμησης. Η μέθοδος bagging εφαρμόζεται συχνά στα προβλήματα ταξινόμησης με στόχο τη βελτίωση της ακρίβειας ταξινόμησης, συνδυάζοντας μεμονωμένους ταξινομητές. Παρακάτω παρουσιάζονται τεχνικές bagging σε συνδυασμό με μεθόδους επαναδειγματοληψίας:

- UnderBagging

Η μέθοδος UnderBagging αποτελεί έναν συνδυασμό της υποδειγματοληψίας και του bagging, όπου από το σύνολο δεδομένων εκπαίδευσης κατασκευάζεται σύνολο δεδομένων που περιλαμβάνει όλα τα δεδομένα των κλάσεων μειοψηφίας μαζί με τυχαία επιλεγμένα δεδομένα από την κλάση πλειοψηφίας με ή χωρίς αντικατάσταση, ελαττώνοντας, δηλαδή, τα δεδομένα της κλάσης πλειοψηφίας που συμμετέχουν στα τελικά σύνολα δεδομένων [34].

- OverBagging

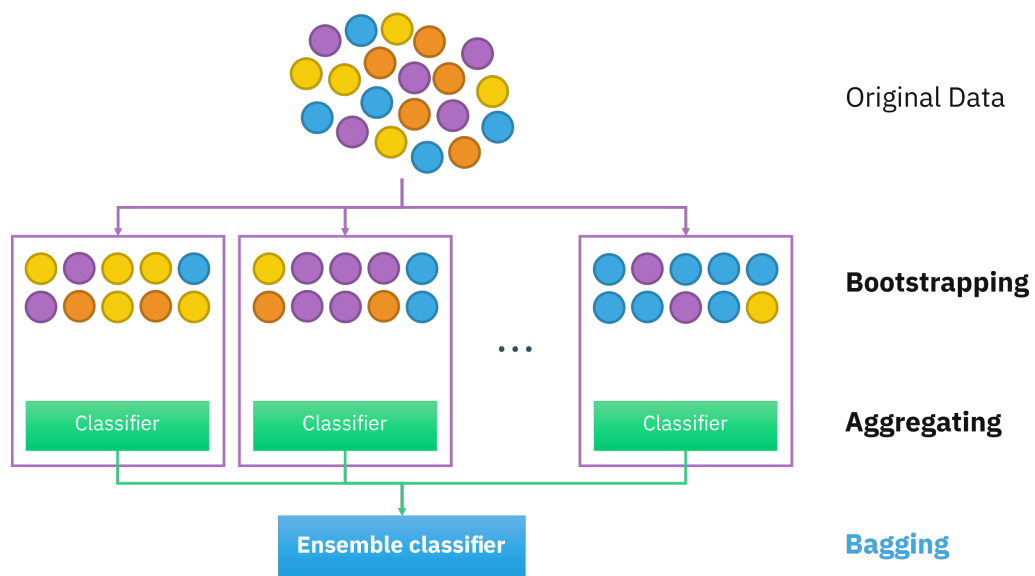
Αντίστοιχα, η μέθοδος OverBagging αποτελεί έναν συνδυασμό της υπερδειγματοληψίας και του bagging, όπου αυξάνονται τα δεδομένα της κλάσης μειοψηφίας και κατασκευάζεται σύνολο δεδομένων που περιλαμβάνει όλα τα δεδομένα της κλάσης πλειοψηφίας μαζί με τα δεδομένα της κλάσης μειοψηφίας που έχουν αυξηθεί, ώστε να υπάρχει ισορροπία [34].

- SMOTEBagging

Η μέθοδος SMOTEBagging αποτελεί έναν συνδυασμό των τεχνικών SMOTE (Synthetic Minority Oversampling Technique) και bagging που αφορά την παραγωγή συνθετικών δεδομένων. Συνιστά μία μέθοδο υπερδειγματοληψίας κατά την οποία επαναδειγματοληφτούν τα δεδομένα της κλάσης μειοψηφίας είτε διπλασιάζοντάς τα είτε δημιουργώντας νέα συνθετικά με τυχαίο τρόπο. Δημιουργείται, επομένως, σύνολο δεδομένων εκπαίδευσης που περιλαμβάνει όλα τα στιγμιότυπα της κλάσης πλειοψηφίας και ισάριθμα δεδομένα από την κλάση μειοψηφίας, τα οποία έχουν αυξηθεί με τον παραπάνω τρόπο [34].

- Bagging Ensembles Variation

Η συγκεκριμένη τεχνική αποτελεί τμήμα του UnderBagging, δηλαδή είναι επίσης συνδυασμός της υποδειγματοληψίας και του bagging. Στόχος είναι να αντιμετωπιστούν τα προβλήματα που σχετίζονται με τη διαδικασία της ταξινόμησης, ειδικά στις περιπτώσεις της ανισορροπίας κλάσεων. Στηρίζεται στην αύξηση των δεδομένων της κλάσης μειοψηφίας χωρίς τη δημιουργία συνθετικών δεδομένων ή την τροποποίηση του συστήματος ταξινόμησης. Για παράδειγμα, αν υπάρχει ανισορροπία στο σύνολο δεδομένων σε ποσοστό 20:80, τότε η κλάση πλειοψηφίας χωρίζεται σε 4 μέρη, και καθένα από αυτά συνδυάζεται με την κλάση μειοψηφίας, ώστε τελικά να έχουν δημιουργηθεί 4 σύνολα δεδομένων με ποσοστά 20:20 του αρχικού συνόλου, δηλαδή τα δεδομένα των επιμέρους κλάσεων να βρίσκονται τελικά σε ποσοστό 50%:50% [34].



Εικόνα 7: Σχηματική απεικόνιση της μεθόδου bagging και των σταδίων bootstrapping και aggregating από τα οποία προκύπτει ο τελικός ταξινομητής συλλογικής μάθησης.

3.1.4 Προβλήματα ταξινόμησης πολλαπλών κλάσεων

Υπάρχει πληθώρα προβλημάτων ταξινόμησης που αφορούν περισσότερες από δύο κλάσεις δεδομένων. Υπάρχουν αρκετά μοντέλα ταξινομητών που είναι σχεδιασμένα για την αντιμετώπιση, τόσο δυαδικών, όσο και προβλημάτων με πολλαπλές κλάσεις, ωστόσο κάποια μοντέλα δεν είναι εύκολο να επεκταθούν στην περίπτωση των πολλαπλών κλάσεων, όπως για παράδειγμα οι μηχανές διανυσμάτων υποστήριξης (SVM). Όσον αφορά στον αριθμό των κλάσεων, ένα δυαδικό πρόβλημα αποτελεί την απλούστερη περίπτωση, ενώ όσο αυξάνεται το πλήθος των κλάσεων του προβλήματος, τόσο αυξάνεται και η δυσκολία του. Αυτός είναι ο λόγος που αναπτύχθηκαν τεχνικές διαίρεσης τέτοιων προβλημάτων σε μικρότερα και συνήθως οι τεχνικές αυτές έχουν στόχο τη μετατροπή του αρχικού προβλήματος σε απλούστερα δυαδικά (binarization techniques) [35].

Οι τεχνικές διαίρεσης σε δυαδικά προβλήματα αποτελούνται από δύο ξεχωριστά βήματα. Το πρώτο αφορά τη στρατηγική διαίρεσης με τις δημοφιλέστερες τεχνικές να είναι η Ένα-προς-Ένα (One-vs-One - OVO) και η Ένα-προς-Όλα (One-vs-All - OVA), ενώ υπάρχουν και άλλες μέθοδοι, όπως ο Κώδικας Εξόδου για Διόρθωση Σφάλματος (Error Correcting Output Code - ECOC) ή ορισμένες ιεραρχικές τεχνικές που δεν είναι τόσο συνηθισμένες. Το δεύτερο βήμα σχετίζεται με την πραγματοποίηση της τελικής πρόβλεψης από τις εξόδους των δυαδικών ταξινομητών, για την οποία είναι ιδιαίτερα σημαντικός ένας κατάλληλος συνδυασμός επιμέρους ταξινομητών [35].

3.1.4.1 Ένα-προς-Ένα

Το σχήμα διαίρεσης Ένα-προς-Ένα (One-vs-One - OVO) αρχικά χωρίζει ένα πρόβλημα m κλάσεων σε $m \cdot \frac{m-1}{2}$ δυαδικά προβλήματα. Για καθένα από αυτά χρησιμοποιείται ένας δυαδικός ταξινομητής που είναι υπεύθυνος για τον διαχωρισμό των δεδομένων ανάμεσα στο συγκεκριμένο ζεύγος κλάσεων. Η εκπαίδευση των ταξινομητών πραγματοποιείται χρησιμοποιώντας ως δεδομένα εκπαίδευσης, μόνο τα στιγμιότυπα του αρχικού συνόλου δεδομένων εκπαίδευσης που ανήκουν στις κλάσεις του συγκεκριμένου ζεύγους, ενώ τα υπόλοιπα στιγμιότυπα αγνοούνται για αυτό το δυαδικό πρόβλημα. Στη φάση της επικύρωσης (validation), σε κάθε δυαδικό ταξινομητή παρέχεται ένα πρότυπο. Η έξοδος ενός ταξινομητή δίνεται από το $r_{i,j} \in [0,1]$, που αποτελεί την αξιοπιστία του ταξινομητή για τον διαχωρισμό των κλάσεων i και j υπέρ της i . Η αξιοπιστία του ταξινομητή για την κλάση j υπολογίζεται από τη σχέση $r_{j,i} = 1 - r_{i,j}$, ενώ η κλάση με τη μεγαλύτερη τιμή αξιοπιστίας θα αποτελέσει την έξοδο του ταξινομητή [35]. Οι έξοδοι αυτές αναπαριστώνται στον πίνακα R :

$$R = \begin{pmatrix} - & r_{1,2} & \cdots & r_{1,m} \\ r_{2,1} & - & \cdots & r_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m,1} & r_{m,2} & \cdots & - \end{pmatrix}$$

Η τελική έξοδος του συστήματος προκύπτει από τον παραπάνω πίνακα με διαφορετικές τεχνικές άθροισης (aggregation), όπως οι δύο απλούστερες που παρουσιάζονται παρακάτω:

- Στρατηγική Ψήφου (Voting Strategy -VOTE)

Η στρατηγική ψήφου αποτελεί την απλούστερη μέθοδο για τον υπολογισμό της τελικής πρόβλεψης από τους δυαδικούς Ένα-προς-Ένα ταξινομητές. Κάθε δυαδικός ταξινομητής αποδίδει μία ψήφο για την κλάση πρόβλεψης και στη συνέχεια οι ψήφοι που λαμβάνει η κάθε κλάση αθροίζονται, με αποτέλεσμα η κλάση με το μέγιστο πλήθος ψήφων να επιλέγεται ως τελική έξοδος για ένα συγκεκριμένο στιγμιότυπο. Η μαθηματική μορφή της μεθόδου είναι [35]:

$$Class = \underset{i=1,\dots,m}{\operatorname{argmax}} \sum_{1 \leq j \neq i \leq m} s_{ij}, \text{ όπου } s_{ij} \text{ ισούται με } 1 \text{ αν } r_{ij} > r_{ji} \text{ και } 0 \text{ διαφορετικά.}$$

- Σταθμισμένη Στρατηγική Ψήφου (Weighted Voting Strategy - WV)

Στη σταθμισμένη στρατηγική ψήφου, κάθε δυαδικός ταξινομητής ψηφίζει και για τις δύο κλάσεις ανάμεσα στις οποίες καλείται να ταξινομήσει τα δεδομένα. Το βάρος για κάθε ψήφο δίνεται από την αξιοπιστία του ταξινομητή κατά την πρόβλεψη κάθε κλάσης. Η κλάση με το μεγαλύτερο άθροισμα αποτελεί την κλάση εξόδου, επομένως ο κανόνας απόφασης είναι ο ακόλουθος [35]:

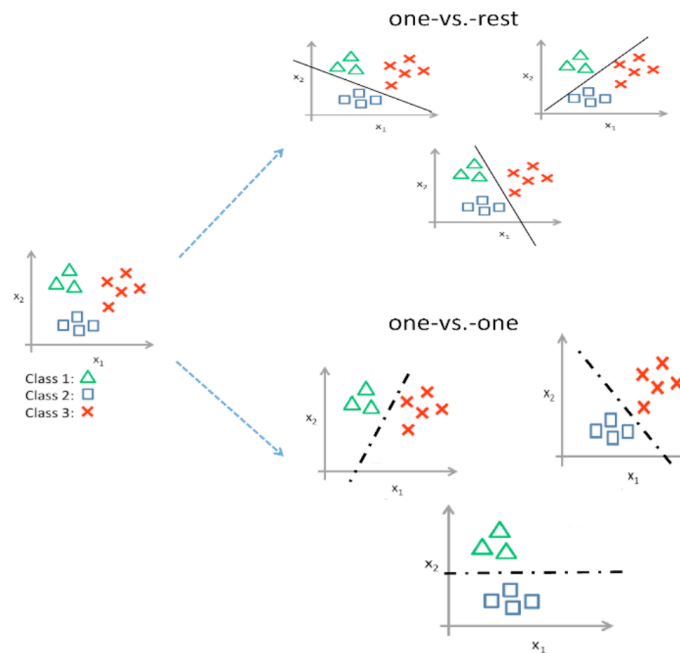
$$Class = \underset{i=1, \dots, m}{\operatorname{argmax}} \sum_{1 \leq j \neq i \leq m} r_{ij}$$

3.1.4.2 Ένα-προς-Όλα

Η τεχνική διαίρεσης Ένα-προς-Όλα (One-vs-All - OVA) χωρίζει ένα πρόβλημα m κλάσεων σε m δυαδικά προβλήματα. Για καθένα από αυτά χρησιμοποιείται ένας δυαδικός ταξινομητής, ο οποίος είναι υπεύθυνος για τον διαχωρισμό μίας κλάσης από όλες τις υπόλοιπες. Η εκπαίδευση του μοντέλου απαιτεί ολόκληρο το σύνολο δεδομένων εκπαίδευσης, με τα δεδομένα της μεμονωμένης κλάσης να λαμβάνουν “θετικές” ετικέτες και αυτά όλων των υπόλοιπων κλάσεων να λαμβάνουν “αρνητικές”, γεγονός που μπορεί να προκαλέσει ανισορροπία στις κλάσεις κάθε δυαδικού προβλήματος. Στο στάδιο του ελέγχου, κάθε στιγμιότυπο δίνεται σε καθέναν από τους δυαδικούς ταξινομητές και στη συνέχεια ο ταξινομητής που δίνει “θετική” έξοδο υποδεικνύει την κλάση εξόδου. Σε πολλές περιπτώσεις, η “θετική” έξοδος δεν είναι μοναδική και για τον λόγο αυτό απαιτείται κάποια τεχνική για την επίλυση πιθανής ισοπαλίας, με τη δημοφιλέστερη τεχνική να χρησιμοποιεί την αξιοπιστία των ταξινομητών για την τελική πρόβλεψη. Αντιθέτως με την Ένα-προς-Ένα τεχνική, η Ένα-προς-Όλα δεν χρησιμοποιεί πίνακα με τιμές αξιοπιστίας, αλλά ένα διάνυσμα $R = (r_1, r_2, \dots, r_i, \dots, r_m)$ της αξιοπιστίας για κάθε κλάση, όπου $r_i \in [0,1]$ η αξιοπιστία για την πρόβλεψη της κλάσης i [35].

- Στρατηγική μέγιστης αξιοπιστίας (Maximum Confidence Strategy - MAX)

Η στρατηγική μέγιστης αξιοπιστίας αποτελεί την απλούστερη και συνηθέστερη μέθοδο τελικής πρόβλεψης για την Ένα-προς-Όλα τεχνική και προσομοιάζει στη σταθμισμένη στρατηγική ψήφου των Ένα-προς-Ένα μοντέλων. Μετά την εκπαίδευση των m ταξινομητών για τον διαχωρισμό κάθε κλάσης από όλες τις υπόλοιπες, κάθε στιγμιότυπο δίνεται σε κάθε ταξινομητή και η κλάση εξόδου αποτιμάται από τον ταξινομητή με τη μέγιστη θετική έξοδο, δηλαδή: $Class = \underset{i=1, \dots, m}{\operatorname{argmax}} r_i$ [35].



Εικόνα 8: Σχηματική αναπαράσταση των τεχνικών διαίρεσης Ένα-προς-Όλα (πάνω) και Ένα-προς-Ένα (κάτω).

3.2 Μετρικές

Οι δείκτες της επίδοσης είναι ιδιαίτερα χρήσιμοι για την αξιολόγηση και τη σύγκριση μοντέλων ταξινόμησης και τεχνικών μηχανικής μάθησης. Οι μετρικές αυτές μπορούν να χρησιμεύσουν σε διαφορετικά στάδια της ανάπτυξης ενός μοντέλου, όπως στην ανάλυση της συμπεριφοράς του ή στη βελτιστοποίηση των υπεραπαραμέτρων του [36].

Σε προβλήματα ταξινόμησης πολλαπλών κλάσεων, υπολογίζεται η πιθανότητα ενός στοιχείου να ανήκει σε καθεμία από τις πιθανές κλάσεις και στη συνέχεια εφαρμόζεται ένας κανόνας ταξινόμησης προκειμένου να ανατεθεί μία κλάση σε κάθε στοιχείο του συνόλου δεδομένων [36].

- Πίνακας Σύγχυσης

Πολλές μετρικές βασίζονται στον πίνακα σύγχυσης (Confusion Matrix), ο οποίος περιλαμβάνει όλη τη σχετική πληροφορία που αφορά την επίδοση της ταξινόμησης ενός αλγορίθμου. Ο πίνακας σύγχυσης καταγράφει ποσοτικά το πραγματικό πλήθος δεδομένων σε κάθε κλάση και το πλήθος των δεδομένων σε κάθε κλάση με βάση την πρόβλεψη του αλγορίθμου [36].

		PREDICTED classification				Total
		Classes	a	b	c	
ACTUAL classification	a	6	0	1	2	9
	b	3	9	1	1	14
	c	1	0	10	2	13
	d	1	2	1	12	16
Total		11	11	13	17	52

Εικόνα 9: Παράδειγμα πίνακα σύγχυσης για ταξινόμηση πολλαπλών κλάσεων.

- Θετική Προβλεπτική Ικανότητα (Positive Predictive Value) & Ανάκληση (Recall)

Η ανάκληση και η θετική προβλεπτική ικανότητα αποτελούν τα δομικά στοιχεία για τις μετρικές ισορροπημένη ακρίβεια (Balanced Accuracy) και F1 Score. Στην περίπτωση της δυαδικής ταξινόμησης η θετική προβλεπτική ικανότητα ορίζεται ως το κλάσμα των Αληθώς Θετικών (True Positive - TP) στοιχείων προς το συνολικό πλήθος των στοιχείων που έχουν θετική πρόβλεψη. Συγκεκριμένα, ως Ορθώς Θετικά (TP) ορίζονται τα στοιχεία που λαμβάνουν “θετική” ετικέτα από το μοντέλο και ανήκουν πράγματι στη θετική κλάση και ως Λανθασμένα Θετικά (False Positive - FP) τα στοιχεία που έχουν λάβει “θετική” ετικέτα ενώ ανήκουν στην αρνητική κλάση.

$$PPV = \frac{TP}{TP + FP}$$

Η θετική προβλεπτική ικανότητα εκφράζει πόσο αξιόπιστο είναι ένα μοντέλο κατά την ταξινόμηση ενός στοιχείου ως “θετικού”.

Η ανάκληση αποτελεί το κλάσμα των TP στοιχείων προς το συνολικό πλήθος των στοιχείων που έχουν ανήκουν στη “θετική” κλάση:

$$Recall = \frac{TP}{TP + FN},$$

όπου ως Λανθασμένα Αρνητικά (False Negative - FN) ορίζονται τα στοιχεία που έχουν λάβει από το μοντέλο “αρνητική” ετικέτα, ενώ στην πραγματικότητα ανήκουν στη “θετική” κλάση. Η ανάκληση εκτιμά την ακρίβεια πρόβλεψης του μοντέλου για τη θετική κλάση, δηλαδή την ικανότητα του μοντέλου να εντοπίζει όλα τα “θετικά” στοιχεία του συνόλου δεδομένων [36].

Όσον αφορά τις μετρικές για τα προβλήματα πολλαπλών κλάσεων:

- Ακρίβεια

Η ακρίβεια (Accuracy) αποτελεί μία από τις δημοφιλέστερες μετρικές στα προβλήματα πολλαπλών κλάσεων, υπολογίζεται απευθείας από τον πίνακα σύγχυσης και ορίζεται ως:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}.$$

Η παραπάνω σχέση λαμβάνει υπόψιν τα Αληθώς Θετικά και Αληθώς Αρνητικά (True Negative - TN) δεδομένα στον αριθμητή, καθώς το συνολικό πλήθος όλων των δεδομένων του πίνακα σύγχυσης στον αριθμητή. Τα TP και TN δεδομένα αποτελούν τα στοιχεία που έχουν ταξινομηθεί σωστά από το μοντέλο και ανήκουν στην κύρια διαγώνιο του πίνακα σύγχυσης, ενώ ο παρονομαστής, από την κύρια διαγώνιο του πίνακα, λαμβάνει επίσης υπόψιν όλα τα στοιχεία που έχουν ταξινομηθεί λάθος από το μοντέλο. Η ακρίβεια εκφράζει την πιθανότητα να είναι σωστή μία πρόβλεψη που πραγματοποιεί το μοντέλο.

Η συγκεκριμένη μετρική αποδίδει το ίδιο βάρος σε κάθε στοιχείο του συνόλου δεδομένων, επομένως καθένα από αυτά συνεισφέρει εξίσου στον υπολογισμό της ακρίβειας. Ωστόσο, όταν το σύνολο δεδομένων περιλαμβάνει κλάσεις με ανισορροπία στο πλήθος των δειγμάτων τους, στον υπολογισμό της ακρίβειας οι πολυπληθέστερες κλάσεις έχουν μεγαλύτερη βαρύτητα. Επομένως, η συγκεκριμένη μετρική θεωρείται κατάλληλη μόνο όταν χρειάζεται να ληφθούν υπόψιν τα μεμονωμένα δείγματα και όχι όταν το πρόβλημα αφορά πολλαπλές κλάσεις και μάλιστα με ανισορροπία στο πλήθος των δειγμάτων τους. Εξαιτίας τις μικρότερης βαρύτητας των κλάσεων με τα λιγότερα δεδομένα, η μετρική της ακρίβειας τείνει να αποκρύπτει τα σφάλματα ταξινόμησης που αφορούν τις μικρές κλάσεις, επομένως δεν είναι εφικτό να εντοπιστούν οι κλάσεις για τις οποίες η επίδοση του αλγορίθμου υστερεί [36].

- Ισορροπημένη Ακρίβεια (Balanced Accuracy)

Η ισορροπημένη ακρίβεια αποτελεί μία μετρική που χρησιμοποιείται τόσο στα δυαδικά προβλήματα ταξινόμησης, όσο και στα προβλήματα πολλαπλών κλάσεων. Υπολογίζεται επίσης από τον πίνακα σύγχυσης και ορίζεται ως:

$$Balanced Accuracy = \frac{\frac{TP}{Total_{row1}} + \frac{TN}{Total_{row2}} + \dots + \frac{TN}{Total_{rowN}}}{N}$$

Στην πραγματικότητα η παραπάνω σχέση εκφράζει τον μέσο όρο της ανάκλησης για όλες τις κλάσεις, δηλαδή την πιθανότητα ένα στοιχείο του συνόλου δεδομένων να ταξινομηθεί σωστά.

Αν το σύνολο δεδομένων είναι σχετικά ισορροπημένο όσον αφορά το πλήθος των στοιχείων σε κάθε κλάση, η ακρίβεια και η ισορροπημένη ακρίβεια τείνουν να συγκλίνουν, ενώ όταν υπάρχει ανισορροπία, οι δύο μετρικές διαφέρουν. Η διαφορά αυτή έγκειται στα βάρη που εφαρμόζει η ανάκληση σε κάθε κλάση. Με τον τρόπο αυτό, σε κάθε κλάση

αποδίδεται ίδια βαρύτητα στον τελικό υπολογισμό της ισορροπημένης ακρίβειας και κάθε κλάση αντιπροσωπεύεται από την ανάκλησή της, ανεξάρτητα από το μέγεθός της, συνεπώς η ισορροπημένη ακρίβεια δεν είναι ευαίσθητη στην ανισόρροπη κατανομή των κλάσεων. Αντιθέτως, η ακρίβεια συνήθως ευνοεί την κλάση πλειοψηφίας, καθώς εξαρτάται κυρίως από την επίδοση του αλγορίθμου στην κλάση με τα περισσότερα δείγματα, ενώ η επίδοση στις μικρότερες κλάσεις δεν επηρεάζει ισχυρά την ακρίβεια λόγω του μικρού τους βάρους [36].

- Σκορ F1

Το σκορ F1 αξιολογεί την επίδοση του μοντέλου ταξινόμησης ξεκινώντας από τον πίνακα σύγχυσης και συνθέτοντας τη θετική προβλεπτική ικανότητα και την ανάκληση στην έννοια του αρμονικού μέσου:

$$F1 \text{ Score} = \left(\frac{2}{ppv^{-1} + recall^{-1}} \right) = 2 \cdot \left(\frac{ppv \cdot recall}{ppv + recall} \right)$$

Η σχέση του σκορ F1 μπορεί να ερμηνευθεί ως ο σταθμισμένος μέσος ανάμεσα στη θετική προβλεπτική ικανότητα και την ανάκληση. Η σχετική συνεισφορά των δύο μετρικών είναι ίση στον υπολογισμό του σκορ F1 και ο αρμονικός μέσος είναι χρήσιμος στην εύρεση της αντιστάθμισης ανάμεσά τους.

Η θετική προβλεπτική ικανότητα και η ανάκληση μπορεί να αναφέρονται και σε δυαδικά προβλήματα ταξινόμησης, όπου λαμβάνονται υπόψιν μόνο τα δείγματα της “θετικής” κλάσης και σε προβλήματα πολλαπλών κλάσεων, όπου χρησιμοποιούνται όλα τα στοιχεία του πίνακα σύγχυσης, δηλαδή όλες οι κλάσεις μία προς μία. Έχει παρατηρηθεί ότι το σκορ F1 δίνει μεγαλύτερη έμφαση στις μικρές κλάσεις και ευνοεί περισσότερο μοντέλα τα οποία έχουν κοντινές τιμές ανάκλησης και θετικής προβλεπτικής ικανότητας.

Στην περίπτωση των πολλαπλών κλάσεων το σκορ F1 πρέπει να συμπεριλαμβάνει όλες τις κλάσεις και προκειμένου να καταστεί αυτό εφικτό είναι απαραίτητη η εισαγωγή ενός αντίστοιχου μέτρου για την ανάκληση και θετική προβλεπτική ικανότητα στον αρμονικό μέσο [36].

- Macro F1 σκορ & Σταθμισμένο macro F1 σκορ

Η μετρική macro F1 υπολογίζεται ως ο αρμονικός μέσος της macro ανάκλησης και macro θετικής προβλεπτικής ικανότητας, οι οποίες ορίζονται ως ο αριθμητικός μέσος της κάθε μετρικής. Η προσέγγιση αυτή λαμβάνει υπόψιν όλες τις κλάσεις ως βασικά στοιχεία του υπολογισμού και σε κάθε κλάση αποδίδεται ίδια βαρύτητα στον μέσο όρο, επομένως δεν υπάρχει διαφοροποίηση ανάμεσα στις κλάσεις με πολλά και με λιγότερα δείγματα. Ωστόσο, αν υπολογιστεί και το σταθμισμένο macro F1 σκορ, που αποτιμάται εξαγοντας τις μετρικές για κάθε κλάση και βρίσκοντας τον σταθμισμένο μέσο τους κατά τη διαίρεση με το πλήθος των στοιχείων σε κάθε κλάση, λαμβάνεται υπόψιν πιθανή ανισορροπία [36].

- Micro F1 σκορ

Η μετρική micro F1 προϋποθέτει τον υπολογισμό της micro ανάκλησης και micro θετικής προβλεπτικής ικανότητας και λαμβάνει υπόψιν συγκεντρωτικά όλα τα στοιχεία TP, FN και FP. Η micro ανάκληση αποδεικνύεται ακριβώς ίδια με τη micro θετική προβλεπτική ικανότητα και κατ' επέκταση το micro F1 σκορ είναι τελικά ίσο με την μετρική της ακρίβειας [36].

3.3 Μέθοδοι Ερμηνευσιμότητας

Τα μοντέλα ταξινόμησης έχουν δύο κύριους στόχους που περιλαμβάνουν την υψηλή επίδοση του συστήματος, δηλαδή την όσο το δυνατόν ακριβέστερη εκτίμηση της εξόδου για νέα δεδομένα εισόδου και την ερμηνευσιμότητα, δηλαδή την ικανότητα να παρέχουν μία κατανοητή σχέση ανάμεσα στα δεδομένα εισόδου και στην έξοδο [6]. Συνήθως υπάρχει αντιστάθμιση ανάμεσα στην επίδοση ενός μοντέλου μηχανικής μάθησης και στην ικανότητα του να παράγει επεξηγήσιμες και ερμηνεύσιμες προβλέψεις, καθώς η υψηλή προβλεπτική ικανότητα συνήθως επιτυγχάνεται με αυξημένη πολυπλοκότητα των μοντέλων. Για παράδειγμα, ένα μοντέλο βαθιάς μάθησης ανήκει στις πιο σύγχρονες μεθόδους μηχανικής μάθησης και επιτρέπει την αυτόματη ανακάλυψη, μάθηση και εξαγωγή ιεραρχικών αναπαραστάσεων των δεδομένων που απαιτούνται για τους σκοπούς της ταξινόμησης. Η διαδικασία της αύξησης της πολυπλοκότητας σε συνδυασμό με τη χρήση πολύ μεγάλων συνόλων δεδομένων που χρησιμοποιούνται για την εκπαίδευση, συντείνουν στην κατασκευή πολύπλοκων συστημάτων που αυξάνουν σε μεγάλο βαθμό την προβλεπτική ικανότητα του μοντέλου και μειώνουν εγγενώς την ικανότητα επεξήγησης των εσωτερικών μηχανισμών που οδηγούν στην τελική πρόβλεψη, καθιστώντας το σκεπτικό της απόφασης δυσνόητο και τις προβλέψεις δύσκολο να ερμηνευθούν [4], [37], [38].

Από τη μία πλευρά υπάρχουν τα μοντέλα μαύρου κουτιού που συμπεριλαμβάνουν τα μοντέλα βαθιάς μάθησης και συλλογικής μάθησης και από την άλλη τα μοντέλα λευκού κουτιού ή γυάλινου κουτιού που εύκολα παράγουν επεξηγήσιμα αποτελέσματα, όπως τα γραμμικά μοντέλα ή εκείνα που βασίζονται σε δέντρα απόφασης, τα οποία ενώ είναι πιο εύκολα επεξηγήσιμα και ερμηνεύσιμα, δεν είναι τόσο ισχυρά, καθώς αποτυγχάνουν να παρέχουν τόσο υψηλή επίδοση όσο τα μοντέλα μαύρου κουτιού [4].

Είναι δυσκολότερο να εμπιστευτεί κανείς, τα συστήματα των οποίων οι αποφάσεις δεν μπορούν να ερμηνευθούν με ευκολία, ειδικότερα σε τομείς όπως η υγεία, όπου είναι εξαιρετικά σημαντική, όχι μόνο η πρόβλεψη του κλινικού αποτελέσματος των ασθενών, αλλά και το να ληφθούν υπόψη τα χαρακτηριστικά ενός ασθενούς με έναν επεξηγήσιμο και μετρήσιμο τρόπο, ενώ ταυτόχρονα να δοθεί καθοδήγηση για την προστασία του από πιθανές ασθένειες. Η ταξινόμηση ενός ασθενούς σε μία ορισμένη κατάσταση υγείας δεν αρκεί, καθώς είναι σημαντικό να παρέχονται επεξηγήσεις που αφορούν χαρακτηριστικά που μπορούν να βελτιωθούν, ώστε να αποφευχθεί μία ανεπιθύμητη κατάσταση υγείας ή να εντοπιστεί έγκαιρα η επικινδυνότητά της [6], [39].

Η ανάγκη για αξιόπιστα, δίκαια και υψηλής επίδοσης συστήματα, προκειμένου να επιλυθούν πραγματικά προβλήματα, έχει οδηγήσει στη δημιουργία του πεδίου της Ερμηνεύσιμης Τεχνητής Νοημοσύνης (eXplainable Artificial Intelligence - XAI), το οποίο εστιάζει στην κατανόηση και ερμηνεία της συμπεριφοράς των συστημάτων τεχνητής νοημοσύνης [40].

Οι όροι ερμηνευσιμότητα (interpretability) και επεξηγησιμότητα (explainability) χρησιμοποιούνται συνήθως χωρίς διάκριση. Ωστόσο, ενώ οι δύο όροι σχετίζονται στενά μεταξύ τους, υπάρχει διάκριση ανάμεσα στις δύο έννοιες. Η ερμηνευσιμότητα αναφέρεται στον βαθμό κατά τον οποίο ένας άνθρωπος μπορεί να κατανοήσει την αιτία μίας απόφασης του μοντέλου και συνδέεται περισσότερο με τη διαίσθηση γύρω από την έξοδό του, δηλαδή όσο πιο ερμηνεύσιμο είναι ένα μοντέλο, τόσο πιο εύκολη είναι η αναγνώριση των σχέσεων αιτίου και αιτιατού ανάμεσα στην είσοδο και την έξοδο. Η έννοια της επεξηγησιμότητας σχετίζεται με την εσωτερική λογική και τους μηχανισμούς του συστήματος μηχανικής μάθησης. Όσο πιο επεξηγήσιμο είναι ένα μοντέλο, τόσο βαθύτερα μπορεί να κατανοήσει ένας άνθρωπος τις εσωτερικές διεργασίες που λαμβάνουν χώρα κατά την εκπαίδευση ή την λήψη αποφάσεων του μοντέλου. Ένα ερμηνεύσιμο μοντέλο δεν αποτελεί οπωσδήποτε ένα μοντέλο, του οποίου τους εσωτερικούς μηχανισμούς μπορεί να κατανοήσει ένας άνθρωπος, γι' αυτό και στα συστήματα μηχανικής μάθησης η ερμηνευσιμότητα δεν συμπεριλαμβάνει αξιωματικά την επεξηγησιμότητα, ούτε το αντίθετο.

Ένας ιδιαίτερα σημαντικός διαχωρισμός των μεθόδων ερμηνευσιμότητας αφορά τα είδη των αλγορίθμων στους οποίους μπορούν να εφαρμοστούν. Αν οι εφαρμογές μίας μεθόδου περιορίζονται σε μία συγκεκριμένη οικογένεια αλγορίθμων, τότε αυτές ονομάζονται συγκεκριμένων-μοντέλων (model-specific), ενώ αντίθετα οι μέθοδοι ερμηνευσιμότητας που εφαρμόζονται σε κάθε είδος αλγορίθμου ονομάζονται ανεξαρτήτως-μοντέλου (model-agnostic). Επιπλέον, μία σημαντική πτυχή του διαχωρισμού των μοντέλων ερμηνευσιμότητας βασίζεται στην κλίμακα της ερμηνείας, δηλαδή αν ένα μοντέλο παρέχει επεξηγήσεις για ένα μόνο στιγμιότυπο του συνόλου δεδομένων, τότε ονομάζεται τοπικό (local), ενώ αν εξηγεί ολόκληρο το σύστημα, τότε είναι καθολικό (global). Τέλος, μια σημαντική διάσταση, που πρέπει να ληφθεί υπόψιν στον διαχωρισμό είναι το είδος των δεδομένων στα οποία εφαρμόζονται οι μέθοδοι, με συχνότερα τα αριθμητικά, τα δεδομένα εικόνων και τα δεδομένα κειμένου [4].

3.3.1 Σημαντικότητα Gini

Το πιο δημοφιλές μέτρο σημαντικότητας χαρακτηριστικών που χρησιμοποιείται από τον αλγόριθμο των Τυχαίων Δασών είναι η σημαντικότητα Gini (Gini Importance). Η σημαντικότητα Gini αποτελεί μία καθολική μέθοδο ερμηνευσιμότητας και προκύπτει άμεσα από τον δείκτη Gini του ταξινομητή των Τυχαίων Δασών, ο οποίος αποτελεί τη συνάρτηση που καθορίζει με βάση ποια χαρακτηριστικά θα εκτελεστεί η διάσπαση του δέντρου στη φάση της εκπαίδευσης. Ο δείκτης Gini υπολογίζει τον βαθμό καθαρότητας των δειγμάτων που έχουν ανατεθεί σε έναν κόμβο με βάση τη διάσπαση στον γονικό του κόμβο [41]. Για παράδειγμα, στην περίπτωση της δυαδικής ταξινόμησης, έστω p_1 η πιθανότητα ενός στιγμιότυπου να ανήκει στη θετική κλάση και p_2 η πιθανότητα να ανήκει στην αρνητική κλάση, τότε ο δείκτης Gini ορίζεται ως:

$$Gini = p_1 \cdot (1 - p_1) + p_2 \cdot (1 - p_2).$$

Ο δείκτης ελαχιστοποιείται όταν κάποια από τις δύο πιθανότητες προσεγγίζει το μηδέν [6]. Όσο υψηλότερη είναι η καθαρότητα του κόμβου, τόσο μικρότερη είναι η τιμή του δείκτη Gini και κάθε φορά που πραγματοποιείται μία διάσπαση κόμβου με βάση ένα συγκεκριμένο χαρακτηριστικό, η τιμή Gini για τους δύο απογόνους είναι μικρότερη από αυτή του κόμβου γονέα. Η τιμή Gini ενός χαρακτηριστικού σε ένα μεμονωμένο δέντρο προκύπτει ως άθροισμα της ελάττωσης του δείκτη Gini από τον γονικό κόμβο στον κόμβο παιδί σε όλους τους κόμβους που χρησιμοποιούν το συγκεκριμένο χαρακτηριστικό για διάσπαση. Η συνολική σημαντικότητα σε ένα δέντρο ορίζεται είτε ως άθροισμα, είτε ως μέσος όρος της τιμής σημαντικότητας όλων των δέντρων στο δάσος [41].

Η εκπαίδευση σε ιατρικά δεδομένα χαρακτηρίζεται συχνά από μεγάλο πλήθος χαρακτηριστικών και μικρό πλήθος στιγμιότυπων στο σύνολο δεδομένων. Ως απλή εκτίμηση της σημαντικότητας των χαρακτηριστικών σε ένα πρόβλημα ταξινόμησης, η σημαντικότητα Gini αποτελεί πολύ δημοφιλή επιλογή. Ωστόσο, υπάρχουν περιπτώσεις όπου η σημαντικότητα Gini μεροληπτεί υπέρ ορισμένων κατηγορικών μεταβλητών, όταν αυτές έχουν μεγαλύτερο πλήθος κατηγοριών από άλλες.

3.3.2 Πρόσθετες Shapley Επεξηγήσεις

Οι πρόσθετες Shapley επεξηγήσεις (SHapley Additive exPlanations - SHAP) αποτελούν μία μέθοδο επεξήγησης μεμονωμένων προβλέψεων βασισμένη στις τιμές Shapley της θεωρίας παιγνίων συμμαχίας. Οι SHAP μέθοδοι παρέχουν καθολικές επεξηγήσεις για την έξοδο του μοντέλου, καθώς προκύπτουν ως συνάθροιση των τιμών Shapley που χρησιμοποιούνται για την ερμηνεία μεμονωμένων προβλέψεων [42].

Στόχος των SHAP είναι να εξηγήσουν την πρόβλεψη ενός στιγμιότυπου x υπολογίζοντας τη συνεισφορά κάθε χαρακτηριστικού στην πρόβλεψη. Η SHAP μέθοδος υπολογίζει τις

τιμές Shapley (της θεωρίας παιγνίων συμμαχίας), όπου οι τιμές των χαρακτηριστικών ενός στιγμιότυπου των δεδομένων δρουν ως παίκτες μίας συμμαχίας και υποδεικνύουν πώς κατανέμεται δίκαια η πρόβλεψη ανάμεσα στα χαρακτηριστικά. Ένας παίκτης μπορεί να αντιστοιχεί σε μία μεμονωμένη τιμή χαρακτηριστικού, όπως στην περίπτωση των αριθμητικών δεδομένων, ή σε ένα σύνολο τιμών, όπως στην περίπτωση των ομάδων από πίξελ στα δεδομένα εικόνων. Το SHAP αποτελεί ένα γραμμικό μοντέλο, καθώς αναπαρίσταται ως αθροιστική μέθοδος της συνεισφοράς των χαρακτηριστικών:

$$g(z') = \varphi_0 + \sum_{j=1}^M \varphi_j \cdot z'_j,$$

όπου το g αποτελεί το μοντέλο επεξήγησης, το $z' \in \{0,1\}^M$ το διάνυσμα συμμαχίας, M είναι το μέγιστο μέγεθος συμμαχίας ή αλλιώς το πλήθος των απλοποιημένων χαρακτηριστικών εισόδου και $\varphi_j \in \mathbb{R}$ η συνεισφορά του χαρακτηριστικού j , δηλαδή η τιμές Shapley [43]. Τα μοντέλα επεξήγησης είθισται να χρησιμοποιούν απλοποιημένες εισόδους x' που ανατίθενται στις αρχικές εισόδους μέσω μίας συνάρτησης αντιστοίχισης $x = h_x(x')$ [42]. Στο διάνυσμα συμμαχίας, ένα στοιχείο με τιμή 1 υποδεικνύει πως η αντίστοιχη τιμή του χαρακτηριστικού είναι “παρούσα”, ενώ ένα στοιχείο με τιμή 0 δηλώνει “απούσα” τιμή χαρακτηριστικού. Για τον υπολογισμό των τιμών Shapley χρησιμοποιείται η υπόθεση πως μόνο κάποιες τιμές χαρακτηριστικών είναι “παρούσες” και κάποιες “απούσες”. Η αναπαράσταση των συμμαχιών με τη μορφή γραμμικού μοντέλου γίνεται για τον υπολογισμό των τιμών φ . Αν για το ζητούμενο στιγμιότυπο x , το διάνυσμα συμμαχίας του αποτελείται μόνο από 1, τότε όλες οι τιμές χαρακτηριστικών είναι “παρούσες” και η παραπάνω σχέση απλοποιείται ως εξής [43]:

$$g(x') = \varphi_0 + \sum_{j=1}^M \varphi_j.$$

Μία σημαντική ιδιότητα του SHAP ως μοντέλου συνάθροισης της συνεισφοράς των χαρακτηριστικών είναι ύπαρξη μίας μοναδικής λύσης που ικανοποιεί τις τρεις επιθυμητές ιδιότητας της τοπικής ακρίβειας (local accuracy), της έλλειψης (missingness) και της συνέπειας (consistency) που θα περιγραφούν παρακάτω [42]:

- Τοπική ακρίβεια

Κατά την προσέγγιση του αρχικού μοντέλου f για μια συγκεκριμένη είσοδο x , η τοπική ακρίβεια απαιτεί το μοντέλο επεξήγησης g να ταιριάζει με την έξοδο του f για την απλοποιημένη είσοδο x'

$$f(\hat{x}) = g(x') = \varphi_0 + \sum_{j=1}^M \varphi_j \cdot x'_j$$

Το μοντέλο επεξήγησης $g(x')$ ταιριάζει με την αρχική έξοδο $f(x)$ όταν $x = h_x(x')$, όπου $\varphi_0 = f(h_x(0))$ αναπαριστά την έξοδο του μοντέλου αγνοώντας όλες τις απλοποιημένες εισόδους.

- Έλλειψη

Στα χαρακτηριστικά που απουσιάζουν στην αρχική είσοδο, αποδίδεται μηδενική σημασία. $x'_j = 0 \Rightarrow \varphi_j = 0$

Η ιδιότητα της έλλειψης περιορίζει τα χαρακτηριστικά για τα οποία ισχύει $x'_j = 0$ να έχουν μηδενική συνεισφορά [42]. Με όρους συμμαχίας, προκειμένου όλες οι τιμές χαρακτηριστικών x'_j ενός στιγμιότυπου να εξηγηθούν, θα πρέπει να έχουν τιμή 1 στο διάνυσμα συμμαχίας. Η παρουσία ενός 0 θα σήμαινε πως η τιμή του χαρακτηριστικού

απουσιάζει για το ζητούμενο στιγμιότυπο. Θεωρητικά, μία απουσιάζουσα τιμή θα μπορούσε να λάβει αυθαίρετα μία Shapley τιμή, χωρίς αυτό να επηρεάσει την ιδιότητα της τοπικής ακρίβειας, καθώς εκτελείται πολλαπλασιασμός με το $x'_j = 0$ [43].

- Συνέπεια

Σύμφωνα με την ιδιότητα της συνέπειας, ακόμα κι αν μεταβληθεί ένα μοντέλο ώστε ένα χαρακτηριστικό να έχει μεγαλύτερη επίδραση στην έξοδο, η συνεισφορά που ανατίθεται στο χαρακτηριστικό, δηλαδή η Shapley τιμή του επίσης αυξάνεται. Έστω $f_x(z') = f(h_x(z'))$ και $z' \setminus j$ υποδηλώνει ότι $z'_j = 0$.

Για οποιαδήποτε δύο μοντέλα f και f' αν $f'_x(z') - f'_x(z' \setminus j) \geq f_x(z') - f_x(z' \setminus j)$ για κάθε είσοδο $z' \in \{0,1\}^M$, τότε $\varphi_j(f', x) \geq \varphi_j(f, x)$ [42].

Ο υπολογισμός των SHAP τιμών πραγματοποιείται ως ενοποιημένο μέτρο της σημαντικότητας χαρακτηριστικών, ορίζοντας $f_x(S) = f(h_x(z')) = E[f(x) | x_S]$, όπου S είναι το σύνολο των μη μηδενικών δεικτών στο z' και $E[f(x) | x_S]$ η αναμενόμενη τιμή της συνάρτησης δεδομένου του υποσυνόλου S των δεδομένων εισόδου. Οι τιμές SHAP συνδυάζουν τις αναμενόμενες αυτές τιμές με τις κλασικές τιμές Shapley, ώστε να αποδώσουν τη συνεισφορά φ_j σε κάθε χαρακτηριστικό. Μόνο ένα πιθανό μοντέλο επεξήγησης g ικανοποιεί και τις τρεις παραπάνω ιδιότητες:

$$\varphi_j = \sum_{S \subseteq N \setminus \{j\}} \frac{|S|!(M - |S| - 1)!}{M!} \cdot [f_x(S \cup \{j\}) - f_x(S)],$$
 όπου N το σύνολο όλων των δεδομένων εισόδου [42].

Ο ακριβής υπολογισμός των τιμών SHAP είναι ιδιαίτερα απαιτητικός, ωστόσο οι τιμές μπορούν να προσεγγιστούν με τη λογική της μεθόδου άθροισης της συνεισφοράς χαρακτηριστικών. Η SHAP αποτελεί μία ενοποιημένη προσέγγιση για την επεξήγηση του αποτελέσματος ενός μοντέλου, η οποία συνδέει τη θεωρία παιγνίων με τα τοπικά μοντέλα επεξήγησης σε συνδυασμό με μερικές μεθόδους αναπαράστασης της μεθόδους άθροισης συνεισφοράς των χαρακτηριστικών [42]. Δύο από αυτές τις μεθόδους αναπαράστασης είναι η γραφική παράσταση της σημαντικότητας χαρακτηριστικών που προκύπτει από τις τιμές Shapley (Feature Importance Plot) και η γραφική παράσταση σύνοψης (Summary Plot).

- Σημαντικότητα χαρακτηριστικών SHAP

Σύμφωνα με τη σημαντικότητα χαρακτηριστικών SHAP, τα χαρακτηριστικά που έχουν υψηλή απόλυτη τιμή Shapley έχουν σημαντική επίδραση στο μοντέλο. Για να αποκτηθεί η καθολική σημαντικότητα του χαρακτηριστικού, υπολογίζεται ο μέσος όρος των απόλυτων τιμών Shapley ανά χαρακτηριστικό για όλα τα δεδομένα, ως εξής:

$$I_j = \frac{1}{n} \cdot \sum_{i=1}^n |\varphi_j^{(i)}|. \text{ Στη}$$

συνέχεια ταξινομούνται τα χαρακτηριστικά σε φθίνουσα σημαντικότητα και κατασκευάζεται η αντίστοιχη γραφική παράσταση. Η γραφική αυτή παράσταση, εκτός από την σημαντικότητα των χαρακτηριστικών, δεν προσφέρει επιπλέον πληροφορία, γι' αυτό και χρησιμοποιείται επιπλέον η γραφική παράσταση σύνοψης (Summary Plot) [43].

- Γραφική Παράσταση Σύνοψης (Summary Plot)

Η γραφική παράσταση σύνοψης συνδυάζει τη σημαντικότητα των χαρακτηριστικών με την επίδραση τους. Κάθε σημείο αποτελεί μία τιμή Shapley για ένα χαρακτηριστικό ενός στιγμιότυπου. Η θέση στον άξονα y καθορίζεται από το χαρακτηριστικό, καθώς όσο πιο

σημαντικό είναι, τόσο υψηλότερα βρίσκεται στον άξονα και η θέση στον άξονα x καθορίζεται από την τιμή Shapley. Το χρώμα του σημείου αναπαριστά την τιμή του χαρακτηριστικού με τα κόκκινα σημεία να υποδεικνύουν υψηλή τιμή και όσο το χρώμα τείνει προς το μπλε το σημείο αντιστοιχεί σε χαμηλότερες τιμές. Τα επικαλυπτόμενα σημεία συσσωρεύονται στην κατεύθυνση του άξονα y , επομένως αποκτάται μία εικόνα της κατανομής των τιμών Shapley ανά χαρακτηριστικό. Στη συγκεκριμένη γραφική παράσταση αναδεικνύονται οι σχέσεις ανάμεσα στην τιμή ενός χαρακτηριστικού και στην επίδρασή της στην τελική πρόβλεψη. Ωστόσο, αξίζει να αναφερθεί πως υπάρχει και γραφική παράσταση που έχει αυτόν ακριβώς τον σκοπό, η γραφική παράσταση εξάρτησης (SHAP Dependence Plot) [43].

3.3.3 Τοπικές Ερμηνεύσιμες Επεξηγήσεις Ανεξάρτητες από το Μοντέλο

Η μέθοδος Τοπικών Ερμηνεύσιμων Επεξηγήσεων Ανεξάρτητων από το Μοντέλο (Local Interpretable Model-agnostic Explanations - LIME) αποτελεί ένα σύγχρονο εργαλείο που χρησιμοποιείται για την επεξήγηση μεμονωμένων προβλέψεων οποιωνδήποτε μοντέλων μαύρου κουτιού δημιουργώντας τοπικά υποκατάστατα μοντέλων. Τα υποκατάστατα αυτά εκπαιδεύονται, ώστε να προσεγγίζουν τις προβλέψεις του αρχικού μοντέλου μαύρου κουτιού και αντί να πραγματοποιείται εκπαίδευση ενός καθολικού υποκατάστατου μοντέλου, η μέθοδος LIME εστιάζει στην εκπαίδευση τοπικών υποκατάστατων για την επεξήγηση μεμονωμένων προβλέψεων [44].

Το LIME ελέγχει τη συμπεριφορά του μοντέλου μηχανικής μάθησης στην έξοδο όταν δίνονται σε αυτό παραλλαγές των δεδομένων. Για κάθε στιγμιότυπο και την αντίστοιχη πρόβλεψή του, παράγονται παραλλαγμένα δεδομένα και πραγματοποιούνται νέες προβλέψεις για τα παραλλαγμένα δείγματα, οι οποίες σταθμίζονται με βάση την εγγύτητά τους στο αρχικό στιγμιότυπο [4]. Ένα ερμηνεύσιμο τοπικό μοντέλο εκπαιδεύεται πάνω στα νέα παραλλαγμένα δεδομένα. Το εκπαιδευμένο αυτό μοντέλο πρέπει να αποτελεί καλή τοπική προσέγγιση των προβλέψεων του αρχικού μοντέλου μηχανικής μάθησης (τοπική αξιοπιστία - local fidelity), χωρίς να απαιτείται να έχει καλή καθολική προσέγγιση [44]. Ερμηνεύοντας το τοπικό μοντέλο, κατ' επέκταση γίνεται ερμηνεύσιμο και το αρχικό μοντέλο μαύρου κουτιού [4].

Τα υποκατάστατα μοντέλων στη μέθοδο LIME εκφράζονται μαθηματικά ως εξής:

$$explanation(x) = \underset{g \in G}{\operatorname{argmin}} L(f, g, \pi_x) + \Omega(g),$$

όπου f η πρόβλεψη του αρχικού μοντέλου για το στιγμιότυπο x και g η επεξήγηση που ελαχιστοποιεί την απώλεια L . Προκειμένου να επιτευχθεί, τόσο η ερμηνευσιμότητα, όσο και η τοπική αξιοπιστία του μοντέλου, η απώλεια, για τον υπολογισμό της οποίας μετράται η εγγύτητα της επεξήγησης ως προς την πρόβλεψη του αρχικού μοντέλου, πρέπει να διατηρείται σε χαμηλές τιμές, ενώ με το $\Omega(g)$ επιβάλλεται ποινή στην πολυπλοκότητα του μοντέλου g και πρέπει επίσης να λαμβάνει χαμηλή τιμή, προκειμένου να είναι αρκετά κατανοητή από τον άνθρωπο η επεξήγηση g . Το G συμβολίζει την οικογένεια όλων των πιθανών επεξηγήσεων και το π_x αποτελεί το μέτρο της εγγύτητας που καθορίζει πόσο μεγάλη θα είναι η γειτονιά γύρω από το προς επεξήγηση στιγμιότυπο x [44].

Το διάγραμμα επεξήγησης που παράγεται από το μοντέλο LIME αντικατοπτρίζει τη συνεισφορά κάθε χαρακτηριστικού στην πρόβλεψη ενός στιγμιότυπου του συνόλου δεδομένων, αναδεικνύοντας τις αλλαγές στα χαρακτηριστικά που θα είχαν την ισχυρότερη επίδραση στην πρόβλεψη. Η έξοδος του μοντέλου δίνει αρχικά την πιθανότητα του στιγμιότυπου να ταξινομηθεί σε καθεμία από τις δύο κλάσεις (στην περίπτωση ενός δυαδικού προβλήματος). Επιπλέον, παράγεται ραβδόγραμμα κατά μήκος του άξονα y με τα πιο επιδραστικά χαρακτηριστικά, ενώ στον άξονα x δίνεται το βάρος καθενός από αυτά.

Το χρώμα κάθε ράβδου υποδεικνύει αν το χαρακτηριστικό αυξάνει ή μειώνει την πιθανότητα του στιγμιότυπου να ταξινομηθεί στην αντίστοιχη κλάση.

3.3.4 Διάγραμμα Μερικής Εξάρτησης

Το διάγραμμα μερικής εξάρτησης (Partial Dependence Plot - PDP) οπτικοποιεί την επίδραση ενός (ή περισσότερων) χαρακτηριστικών στην τελική πρόβλεψη ενός μοντέλου μηχανικής μάθησης αγνοώντας την επίδραση κάθε άλλου χαρακτηριστικού. Μπορεί να αναδείξει αν η σχέση μεταξύ της εξόδου και του χαρακτηριστικού είναι γραμμική, μονοτονική ή πιο πολύπλοκη. Η συνάρτηση του διαγράμματος μερικής εξάρτησης για την περίπτωση της παλινδρόμησης περιγράφεται ως εξής:

$$f_S(\hat{x}_S) = E_{X_C}[f(\hat{x}_S, X_C)] = \int f(\hat{x}_S, X_C) \cdot d\mathbb{P}(X_C).$$

Το x_S συμβολίζει τα χαρακτηριστικά για τα οποία δημιουργείται το διάγραμμα μερικής εξάρτησης και το X_C τα υπόλοιπα χαρακτηριστικά που αξιοποιούνται στο μοντέλο μηχανικής μάθησης \hat{f} , τα οποία λειτουργούν ως τυχαίες μεταβλητές. Η μερική εξάρτηση λειτουργεί περιθωριοποιώντας την έξοδο του μοντέλου μηχανικής μάθησης στην κατανομή των χαρακτηριστικών του συνόλου C , ώστε η συνάρτηση να αναδεικνύει τη σχέση μεταξύ των χαρακτηριστικών του συνόλου S και του αποτελέσματος της πρόβλεψης. Με τον τρόπο αυτό προκύπτει μία συνάρτηση που εξαρτάται μόνο από τα χαρακτηριστικά του συνόλου S [45].

Για την εκτίμηση της μερικής συνάρτησης \hat{f}_S υπολογίζεται ο μέσος όρος στα δεδομένα εκπαίδευσης:

$$\hat{f}_S(x_S) = \frac{1}{n} \sum_{i=1}^n \hat{f}(x_S, x_C^{(i)})$$

Η μερική συνάρτηση υποδεικνύει ποια είναι η μέση οριακή επίδραση στη πρόβλεψη για δεδομένες τιμές των χαρακτηριστικών του συνόλου S . Οι τιμές $x_C^{(i)}$ αποτελούν τις πραγματικές τιμές των χαρακτηριστικών τα οποία δεν αφορούν το διάγραμμα εξάρτησης και το n συμβολίζει το πλήθος των δειγμάτων στο σύνολο δεδομένων. Μία υπόθεση που πραγματοποιεί το διάγραμμα μερικής εξάρτησης PDP είναι πως τα χαρακτηριστικά του συνόλου C δεν συσχετίζονται με αυτά του συνόλου S [45].

Το διάγραμμα μερικής εξάρτησης αποτελεί μία καθολική μέθοδο ανεξάρτητη του μοντέλου, καθώς λαμβάνει υπόψιν όλα τα στιγμιότυπα και παραθέτει μία επεξήγηση που αφορά την καθολική σχέση ενός χαρακτηριστικού με την τελική πρόβλεψη [45].

Η μεμονωμένη υπό όρους προσδοκία (Individual Conditional Expectation - ICE) αποτελεί την αντίστοιχη τοπική μέθοδο, καθώς αντιστοιχεί μία γραμμή σε κάθε στιγμιότυπο και αναπαριστά τη αλλαγή της πρόβλεψης καθώς μεταβάλλεται η τιμή ενός χαρακτηριστικού. Το διάγραμμα ICE οπτικοποιεί την εξάρτηση της πρόβλεψης από κάποιο χαρακτηριστικό για κάθε στιγμιότυπο μεμονωμένα, επομένως το διάγραμμα PDP αποτελεί τον μέσο όρο των γραμμών ενός διαγράμματος ICE. Οι τιμές μίας γραμμής υπολογίζονται διατηρώντας σταθερά όλα τα άλλα χαρακτηριστικά και δημιουργώντας παραλλαγές του στιγμιότυπου, ώστε να πραγματοποιηθούν προβλέψεις από το μοντέλο με αυτά τα νέα παραλλαγμένα στιγμιότυπα.

Ο λόγος που τα διαγράμματα ICE συνυπάρχουν με τα καθολικότερα PDP έγκειται στο γεγονός ότι τα διαγράμματα μερικής εξάρτησης μπορεί να αποκρύψουν την ετερογενή αλληλεπίδραση ανάμεσα στα χαρακτηριστικά των συνόλων S και C . Τα PDP μπορούν μόνο να καταδείξουν την μέση σχέση ανάμεσα σε ένα χαρακτηριστικό και στο αποτέλεσμα της

πρόβλεψης και είναι αξιόπιστα μόνο στην περίπτωση που οι αλληλεπιδράσεις ανάμεσα στα χαρακτηριστικά είναι ασθενείς, ενώ τα διαγράμματα ICE επιτρέπουν να παρατηρηθεί η διαφορετική συμπεριφορά του ζητούμενου χαρακτηριστικού ανάλογα με τις διαφορετικές τιμές που έχουν τα χαρακτηριστικά του συνόλου C για κάθε στιγμιότυπο [46].

3.3.5 Αντιπαραθετικά παραδείγματα Permute Attack

Τα Αντιπαραθετικά παραδείγματα Permute Attack (Permute Attack Counterfactual Examples) αποτελούν μία μέθοδο για την παραγωγή αντιπαραθετικών επεξηγήσεων για αριθμητικά δεδομένα. Τα αντιπαραθετικά παραδείγματα είναι υποθετικά παραδείγματα που οδηγούν το μοντέλο σε διαφορετική πρόβλεψη από την αρχική για ένα δείγμα, παρέχοντας παραλλαγές της εισόδου με διαταραγμένες τις τιμές ορισμένων χαρακτηριστικών. Μία τέτοια τεχνική επιτρέπει τον εντοπισμό ασθενών που βρίσκονται σε υψηλό κίνδυνο για την ανάπτυξη ορισμένων ασθενειών. Δημιουργώντας μία παραλλαγή του στιγμιότυπου ενός υγιούς ατόμου, το άτομο μπορεί να μεταπηδήσει στην κλάση των ασθενών και εάν το παραλλαγμένο στιγμιότυπο είναι κοντά στο αρχικό, αυτό σημαίνει πως το άτομο πιθανώς βρίσκεται σε υψηλό κίνδυνο για την ανάπτυξη της συγκεκριμένης νόσου.

Το αντιπαραθετικό παράδειγμα πρέπει να είναι όσο το δυνατόν εγγύτερα στο αρχικό δείγμα όσον αφορά στις τιμές των χαρακτηριστικών, γεγονός που καθιστά απαραίτητο κάποιο μέτρο απόστασης ανάμεσα στα δύο δείγματα. Το παράδειγμα θα πρέπει, επιπλέον, να έχει αλλαγές στις τιμές όσο το δυνατόν λιγότερων χαρακτηριστικών και αυτό μπορεί να επιτευχθεί επιλέγοντας το κατάλληλο μέτρο απόστασης, όπως η απόσταση Μανχάτταν, ενώ η τελευταία προϋπόθεση που πρέπει να πληροί ένα αντιπαραθετικό παράδειγμα είναι οι τιμές των χαρακτηριστικών να λαμβάνουν αληθοφανείς τιμές. Ο αλγόριθμος Permute Attack βασίζεται στη βελτιστοποίηση γενετικού αλγορίθμου και, συνεπώς, δεν εξαρτάται από υποθέσεις διαφοροποιησιμότητας για το εκάστοτε μοντέλο [47].

Για τη μέθοδο παραγωγής των διαταραχών στις τιμές των χαρακτηριστικών, έστω ο χώρος εισόδου $X \subset \mathbb{R}^d$ και $Y = \{0, 1, \dots, K - 1\}$ το σύνολο των ετικετών. Υποθέτοντας ότι υπάρχει μία κατανομή P στο X από την οποία λαμβάνονται δείγματα των εισόδων, και η ετικέτα της κάθε κλάσης καθορίζεται από τη συνάρτηση αληθείας $f^* : X \rightarrow Y$. Για μία είσοδο x επιλεγμένη από το σύνολο X , η αντίστοιχη ετικέτα κλάσης προκύπτει ως $Y = f^*(x)$. Έστω ένας ντετερμινιστικός ταξινομητής $f : X \rightarrow Y$ (για παράδειγμα μία συναρτησιακή σχέση που προκύπτει ως αποτέλεσμα μάθησης από τη διαδικασία εκπαίδευσης). Η επίδοση του f μπορεί να εκτιμηθεί από την απώλεια μηδέν-ένα: $L(f, x) = I(f(x) \neq f^*(x))$, όπου το $I(x)$ αποτελεί τη συνάρτηση δείκτη στο X [47].

Για την ταξινόμηση με κατανομή δεδομένων P και συνάρτηση αληθείας f^* , μία γεννήτρια αντιπαραθετικών διαταραχών για έναν ταξινομητή f είναι μία μετρήσιμη αντιστοιχία $g : X \rightarrow X$, η οποία διατηρεί τις ετικέτες των δειγμάτων, δηλαδή ισχύει $f^*(x) = f^*(g(x))$. Η παραδοχή αυτή, η οποία καλείται παραδοχή εγγύτητας, αντικατοπτρίζεται στην πράξη στην παραδοχή ότι οι αντιπαραθετικές διαταραχές τιμών παραμένουν στην ε-γειτονιά του αρχικού δεδομένου x (ως προς κάποια μετρική), υποθέτοντας ότι η ετικέτα που προκύπτει για το $g(x)$ είναι ίδια με αυτή του x [47].

Ο αλγόριθμος Permute Attack βασίζεται στη βελτιστοποίηση γενετικού αλγορίθμου χωρίς κλίση. Οι γενετικοί αλγόριθμοι είναι εμπνευσμένοι από τη φυσική επιλογή και τη θεωρία της εξέλιξης. Η βελτιστοποίηση ξεκινάει από μία δεξαμενή δειγμάτων (πληθυσμός P), τα οποία εξελίσσονται επαναλαμβανόμενα, ώστε να αυξηθεί συνάρτηση καταλληλότητας (για παράδειγμα μία συνάρτηση βελτιστοποίησης κόστους). Σε κάθε επανάληψη, η εκάστοτε γενιά (δηλαδή ο πληθυσμός της συγκεκριμένης επανάληψης) περνά από το στάδιο της

επιλογής, της δισταύρωσης και της μετάλλαξης. Στο βήμα της επιλογής ο νέος πληθυσμός επιλέγεται με τυχαίο τρόπο με βάση την καταλληλότητα (fitness) που υπολογίζεται από τη συνάρτηση καταλληλότητας, γεγονός που σημαίνει πως δείγματα με χαμηλότερο κόστος έχουν υψηλότερη πιθανότητα να επιλεγούν. Η επόμενη γενιά παράγεται από τον συνδυασμό της δισταύρωσης και της μετάλλαξης. Το στάδιο της δισταύρωσης συνδυάζει τα χαρακτηριστικά δύο γονίων με τυχαία σειρά, ενώ κατά τη μετάλλαξη, εφαρμόζονται κάποιες μεταβολές σε ορισμένα χαρακτηριστικά, τα οποία επιλέγονται με τυχαίο τρόπο [47].

Η βασικότερη διαφορά ανάμεσα στον αλγόριθμο Permute Attack και σε άλλους αλγόριθμους αντιπαραθετικών παραδειγμάτων είναι ο τρόπος με τον οποίο τα δείγματα μεταβάλλονται. Συνήθως τα δείγματα μεταβάλλονται με μικρό θόρυβο μέσα σε μία l_p σφαίρα, στην οποία μπορεί να δημιουργηθούν ορισμένες μη πιθανές τιμές, ειδικότερα στην περίπτωση των κατηγορικών χαρακτηριστικών. Αντίθετα, το Permute Attack αλλάζει τις τιμές των δειγμάτων, μεταβάλλοντας τυχαία επιλεγμένα χαρακτηριστικά, των οποίων οι νέες τιμές επιλέγονται από πιθανές τιμές των χαρακτηριστικών στα δεδομένα εκπαίδευσης, ώστε να είναι έγκυρες και ρεαλιστικές και η κατανομή πιθανότητας των χαρακτηριστικών να παραμένει ίδια στις νέες γενιές [47].

Έστω $\Pi = \{\pi_1, \pi_2, \dots, \pi_n\}$ ένα σύνολο διανυσμάτων μήκους n , με το καθένα να περιλαμβάνει μία διαφορετική μετάθεση του συνόλου $\{1, 2, \dots, n\}$ και το $X = [X_1, X_2, \dots, X_m] \subset \mathbb{R}^{d \times m}$ να αποτελεί την είσοδο μήκους d (όπου κάθε γραμμή συνιστά ένα στιγμιότυπο των δεδομένων) και m στηλών (όπου οι στήλες αντιπροσωπεύουν τα χαρακτηριστικά). Ορίζουμε $X_i[\pi_l]$, με $[\pi_l] \in \Pi$, ως μία μετάθεση του X_i , γεγονός που σημαίνει πως η τιμή του έχει επιλεγεί τυχαία από τις πιθανές τιμές στη στήλη i . Το f αναπαριστά το μοντέλο το οποίο έχει προσαρμοστεί σε ζεύγη (X, Y) , όπου $Y \in y$, $y \subset \mathbb{R}^K$ είναι η ετικέτα στόχος με K πιθανές κλάσεις και $f(x) \subset \mathbb{R}^K$ το μοντέλο πρόβλεψης για την είσοδο $x = \{x_1, \dots, x_m\} \subset \mathbb{R}^m$. Επιπλέον, ορίζεται ως $x_{perm}^{(i)} = \{x_1, \dots, X_i[\pi_l, 0], \dots, x_m\}$ ένα δείγμα με την i -οστή στήλη να μεταβάλλεται κατά π_l και να επιλέγεται η τιμή της πρώτης γραμμής [47].

Στόχος του αλγορίθμου Permute Attack είναι η εύρεση του x_{perm} , το οποίο έχει το μικρότερο πλήθος αλλαγμένων στηλών και άρα χαρακτηριστικών $\delta_{0,max}$ και την ελάχιστη μεταβολή μέσα σε μία σφαίρα l_2 των $\delta_{2,max}$. Το παραλλαγμένο δείγμα που ικανοποιεί τις παραπάνω συνθήκες αποτελεί ένα αντιπαραθετικό παράδειγμα x_{cnt} .

$argmax_{x \in C} f(x_{cnt}) = t$, τέτοιο ώστε $\|x_{orig} - x_{cnt}\| \leq \delta_{0,max}$ και $\|x_{orig} - x_{cnt}\|_2 \leq \delta_{2,max}$

Οι κύριες συναρτήσεις του αλγορίθμου Permute Attack είναι:

- Η συνάρτηση καταλληλότητας

Η συνάρτηση καταλληλότητας αξιολογεί την ποιότητα κάθε μέλους του πληθυσμού. Η σύααρτηση καταλληλότητας του Permute Attack ορίζεται ως:

$Fitness(x) = \|f(x)_t - f(x_{orig})_t\|_2 - \rho_0 \cdot \|x_{orig} - x\|_0 - \rho_1 \cdot \|x_{orig} - x\|_2$, όπου $f(x)_t$

είναι το αποτέλεσμα για την κλάση στόχο t . Οι είσοδοι με τις υψηλότερες τιμές καταλληλότητας είναι τα δείγματα που επιλέγονται στο βήμα της επιλογής [47].

- Η επιλογή

Τα στοιχεία με υψηλότερη τιμή της συνάρτησης καταλληλότητας είναι πιθανότερο να ανήκουν στην επόμενη γενιά, ενώ τα μέλη με χαμηλότερη καταλληλότητα είναι πιθανότερο να αντικατασταθούν. Η πιθανότητα επιλογής για κάθε στοιχείο του πληθυσμού υπολογίζεται από τη softmax συνάρτηση των τιμών καταλληλότητας, ώστε να λάβουν τη μορφή κατανομής πιθανότητας. Επιπλέον, το στοιχείο με την υψηλότερη καταλληλότητα στην εκάστοτε γενιά, είναι εγγυημένο πως θα αποτελέσει μέρος της επόμενης γενιάς [47].

- Η διασταύρωση

Τα χαρακτηριστικά των επιλεγμένων γονέων ανταλλάσσονται με τυχαίο τρόπο [47].

- Η μετάλλαξη

Κάποια από τα χαρακτηριστικά των μελών της γενιάς επιλέγονται με βάση τη σημαντικότητά τους για την αλλαγή του αποτελέσματος και η τιμή τους αντικαθίσταται τυχαία με μία από τις πιθανές τιμές του χαρακτηριστικού από τα δεδομένα εκπαίδευσης. Η πιθανότητα επιλογής υπολογίζεται ως: $softmax(\{e_1, e_2, \dots, e_m\})$, όπου e_i είναι οι αλλαγές στο $f(x)_i$ όταν το i -οστό χαρακτηριστικό έχει μεταβληθεί στην προηγούμενη γενιά. Επομένως, τα χαρακτηριστικά με μικρότερη επίδραση στην αλλαγή του αποτελέσματος, είναι λιγότερο πιθανό να επιλεγούν για μετάλλαξη [47].

- Η ενημέρωση των παραμέτρων

Αν οι συνθήκες που εφαρμόζονται στο l_0 και στο l_2 (που αναπαρίστανται ως $\rho = \{\rho_0, \rho_1\}$) είναι αυστηρές και ο αλγόριθμος αδυνατεί να βρει κάποιο αντιπαραθετικό παράδειγμα, η τιμή του ρ θα εξασθενίσει κατά έναν προκαθορισμένο παράγοντα. Αυτό χαλαρώνει τις συνθήκες και επιτρέπει στον αλγόριθμο να βρει παραδείγματα με μεγαλύτερο πλήθος αλλαγμένων χαρακτηριστικών, με μεγαλύτερες αποκλίσεις από τις αρχικές τιμές [47].

- Στατιστικές ανώτερης τάξης

Η ανεξάρτητη μεταβολή χαρακτηριστικών αγνοεί τις στατιστικές σχέσεις ανώτερης τάξης ανάμεσα στα χαρακτηριστικά. Αυτό αποτελεί μία συνηθισμένη παραδοχή για αυτό το είδος αλγορίθμων, γεγονός που καθιστά τα παραγόμενα δείγματα μη ρεαλιστικά. Επομένως, τα παραδείγματα αυτά μπορούν να εντοπιστούν με τη χρήση ενός αλγορίθμου ανίχνευσης στοιχείων εκτός της κατανομής.

Προκειμένου να βελτιωθεί η ποιότητα των αντιπαραθετικών παραδειγμάτων, η δειγματοληψία πρέπει να πραγματοποιείται από μία ρεαλιστική, κοινή κατανομή δεδομένων. Στον αφελή αλγόριθμο Permute Attack τα χαρακτηριστικά θεωρούνται ανεξάρτητα και κάθε χαρακτηριστικό μεταβάλλεται με ανεξάρτητο τρόπο. Προκειμένου να επιλυθεί αυτό, η μεταβολή περιορίζεται σε κοντινότερες τιμές του χαρακτηριστικού.

Για να επιτευχθεί ο στόχος αυτός, όλα τα συνεχή χαρακτηριστικά διακριτοποιούνται σε μικρότερα “δοχεία” (bins), ώστε να υπάρχει ίσος αριθμός από δείγματα σε κάθε “δοχείο” χαρακτηριστικών. Χρησιμοποιώντας τα διακριτοποιημένα δεδομένα, προκειμένου να μεταβληθεί η τιμή του i -οστού χαρακτηριστικού, οι πιθανές τιμές για τη μεταβολή επιλέγονται ανάμεσα στα δείγματα των οποίων όλα τα διακριτοποιημένα χαρακτηριστικά, εκτός του i -οστού, είναι ίδια, δηλαδή $x_j = x'_j, \forall \{j : j \in \{1, \dots, m\} \wedge j \neq i\}$. Συνεπώς, η δειγματοληψία γίνεται από την υπό συνθήκη κατανομή πιθανοτήτων $p(x_i | x_{-i})$, όπου το x_{-i} συμβολίζει όλα τα χαρακτηριστικά εκτός του i -οστού.

Η υπό συνθήκη κατανομή βελτιώνει τη δειγματοληψία για ένα χαρακτηριστικό δεδομένων όλων των άλλων χαρακτηριστικών. Ωστόσο, συχνά στις επαναλήψεις μετάλλαξης και

διασταύρωσης μεταβάλλονται περισσότερο από ένα χαρακτηριστικά. Προκειμένου να πραγματοποιηθεί δειγματοληψία πολλαπλών χαρακτηριστικών μπορεί να ακολουθηθεί η ίδια διαδικασία. Για παράδειγμα, αν το i -οστό και το j -οστό χαρακτηριστικό μεταβάλλονται από κοινού, οι τιμές τους θα δειγματοληπτηθούν από την κατανομή $p(x_{i,j} | x_{-i,j})$ διατηρώντας σταθερά όλα τα χαρακτηριστικά εκτός από τα i και j και επιλέγοντας τις τιμές των $x_{i,j}$ από τα δείγματα που έχουν κοινές παρόμοιες διακριτοποιημένες τιμές σε όλα τα χαρακτηριστικά εκτός των i και j .

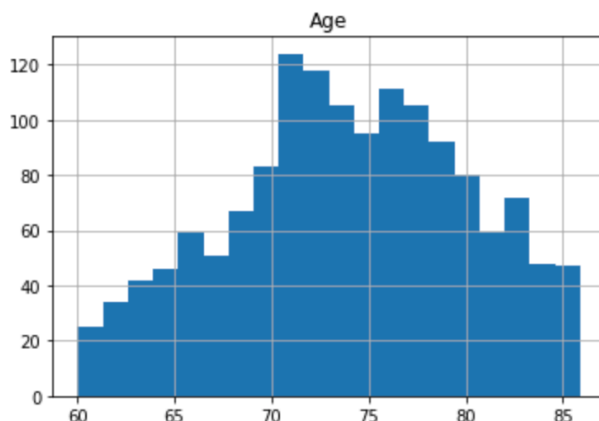
Ωστόσο, για να αυξηθεί η ακρίβεια της από κοινού δειγματοληψίας χαρακτηριστικών, χρησιμοποιείται η δειγματοληψία Gibbs. Για να εκτιμηθεί ένα δείγμα της από κοινού κατανομής, η δειγματοληψία Gibbs επιλέγει επαναληπτικά δείγματα από τις υπό συνθήκη κατανομές. Η διαδικασία αυτή αποδεικνύεται πως συγκλίνει στην από κοινού κατανομή. Στην πράξη, η σύγκλιση αυτή επιτυγχάνεται ύστερα από μερικές επαναλήψεις.

Η στρατηγική δειγματοληψίας μπορεί να χρησιμοποιηθεί απευθείας στο στάδιο της μετάλλαξης, όμως στη βήμα της διασταύρωσης, οι τιμές των τυχαία επιλεγμένων χαρακτηριστικών στα δείγματα γονείς ανταλλάσσονται με τυχαίο τρόπο. Προκειμένου να επιβεβαιωθεί πως αυτή η προσέγγιση είναι κοντά στην πραγματική κατανομή των δεδομένων, ακολουθείται η ίδια στρατηγική δειγματοληψίας, όμως αντί να επιλέγονται τυχαίες τιμές από τις πιθανές επιλογές, επιλέγεται η εγγύτερη τιμή σε αυτήν που πρόκειται να αλλαχθεί [47].

4. Μέθοδος

4.1 Σύνολο Δεδομένων

Το σύνολο δεδομένων που χρησιμοποιήθηκε για τη συγκεκριμένη έρευνα προέρχεται από το Alzheimer's Disease Neuroimaging Initiative (ADNI). Αποτελείται από 1463 άτομα, εκ των οποίων 449 είναι γνωσιακά υγιή άτομα (controls - CN), 740 έχουν διαγνωστεί με ήπιας μορφής νοητική εξασθένηση (Mild Cognitive Impairment - MCI) και 274 πάσχουν από άνοια (Dementia - DEM). Οι ηλικίες των συμμετεχόντων ανήκουν στο εύρος 60 με 86 έτη και παρακάτω δίνεται το αντίστοιχο ιστόγραμμα της κατανομής των ηλικιών τους (Εικόνα 10):



Εικόνα 10: Ιστόγραμμα των ηλικιών των συμμετεχόντων.

Το σύνολο δεδομένων αποτελείται από δημογραφικά, κλινικά και γενετικά δεδομένα. Τα δημογραφικά δεδομένα αφορούν την ηλικία και το φύλο των ασθενών, τα κλινικά έχουν ληφθεί από T1 εικόνες Μαγνητικής Τομογραφίας του εγκεφάλου, ενώ τα γενετικά αποτελούνται από πολυμορφισμούς ενός νουκλεοτιδίου (Single Nucleotide Polymorphisms - SNPs) που αντιστοιχούν σε γονίδια σχετιζόμενα με τη νόσο Αλτσχάιμερ. Πιο συγκεκριμένα, τα κλινικά δεδομένα απαρτίζονται από τον συνολικό όγκο του εγκεφάλου και από 145 ογκομετρικές περιοχές ενδιαφέροντος (Volumes Of Interest - VOIs) που αντιστοιχούν στον όγκο που καταλαμβάνουν συγκεκριμένες ανατομικές περιοχές, όπως για παράδειγμα ο ιππόκαμπος ή η αμυγδαλή, ενώ τα γενετικά από 54 νουκλεοτιδικούς πολυμορφισμούς που περιγράφουν πόσα αλληλόμορφα διαθέτει ο κάθε ασθενής από το κάθε SNP.

Ηλικία	Γνωσιακά υγιή Άτομα (% άντρες)	MCI (% άντρες)	Άνοια (% άντρες)
60-69	96 (39.58%)	236 (51.69%)	55 (60.00%)
70-79	287 (49.82%)	366 (65.85%)	151(53.64%)
80-86	66 (66.67%)	138 (58.69%)	68 (67.65%)
Σύνολο	449 (50.11%)	740 (60.00%)	274 (58.39%)

Πίνακας 1 - Κατανομή των κλάσεων ανά ηλικιακή ομάδα

Όλα τα κλινικά δεδομένα είναι αριθμητικά (tabular), δηλαδή λαμβάνουν συνεχείς αριθμητικές τιμές, ενώ από τα δημογραφικά δεδομένα, το φύλο των ατόμων αποτελεί κατηγορική μεταβλητή που κωδικοποιήθηκε ώστε η τιμή 1 να αντιστοιχεί στην κατηγορία “γυναίκα” και η τιμή 0 στην κατηγορία “άντρας”. Επιπλέον, όσον αφορά στα γενετικά δεδομένα, σε αυτά αποδόθηκαν διακριτές αριθμητικές τιμές στο διάστημα [0,1] που υποδεικνύουν αν το άτομο έχει 0, 1 ή 2 αλληλόμορφα του συγκεκριμένου SNP.

Ο Πίνακας 1 παρουσιάζει το πλήθος των συμμετεχόντων που ανήκουν σε κάθε κλάση ανά ηλικιακή ομάδα και το αντίστοιχο ποσοστό των αντρών στις τρεις κλάσεις. Όπως φαίνεται, υπάρχει αρκετά σημαντική ανισορροπία ανάμεσα στο πλήθος των ατόμων των τριών κλάσεων, με τα MCI άτομα να αποτελούν την κλάση πλειοψηφίας και τα CN και DEM να συνιστούν τις κλάσεις μειοψηφίας. Τα υγιή άτομα (CN) αντιστοιχούν στο 30.69% των συνολικών συμμετεχόντων, τα MCI στο 50.58% και τα άτομα με άνοια (DEM) στο 18.73%. Η ανισορροπία αυτή μπορεί σε μεγάλο βαθμό να επηρεάσει την επίδοση ενός μοντέλου ταξινόμησης, γι’ αυτό και πρέπει να αντιμετωπιστεί με κατάλληλη μέθοδο.

4.2 Προεπεξεργασία Δεδομένων

Προκειμένου να εξαχθούν τα αριθμητικά δεδομένα του όγκου των περιοχών ενδιαφέροντος του εγκεφάλου, είχε προηγηθεί διαδικασία προεπεξεργασίας κατά την οποία κάθε εικόνα μαγνητικής τομογραφίας εγκεφάλου υπέστη διόρθωση της ανομοιογένειας των εντάσεων της, ενώ στη συνέχεια, σε κάθε εικόνα εφαρμόστηκε αλγόριθμος αφαίρεσης του κρανίου [48], [49].

Το επόμενο στάδιο της προεπεξεργασίας που εφαρμόστηκε στα αριθμητικά δεδομένα των VOI είναι η γραμμική διόρθωση συμμεταβλητών που είχε ως στόχο την εξάλειψη της επίδρασης της ηλικίας, του φύλου και του συνολικού όγκου του εγκεφάλου στα δεδομένα. Αναλυτικότερα, επειδή η ανατομία του εγκεφάλου και ο όγκος της λευκής και της φαιάς ουσίας κάθε περιοχής διαφέρει ανάλογα με την ηλικία ή το φύλο ενός ασθενούς [50], [51], ήταν απαραίτητη η προσαρμογή των δεδομένων, ώστε να διατηρηθούν μόνο οι σχετιζόμενες με τις υπό μελέτη νόσους νευροανατομικές διαφοροποιήσεις του όγκου των περιοχών ενδιαφέροντος.

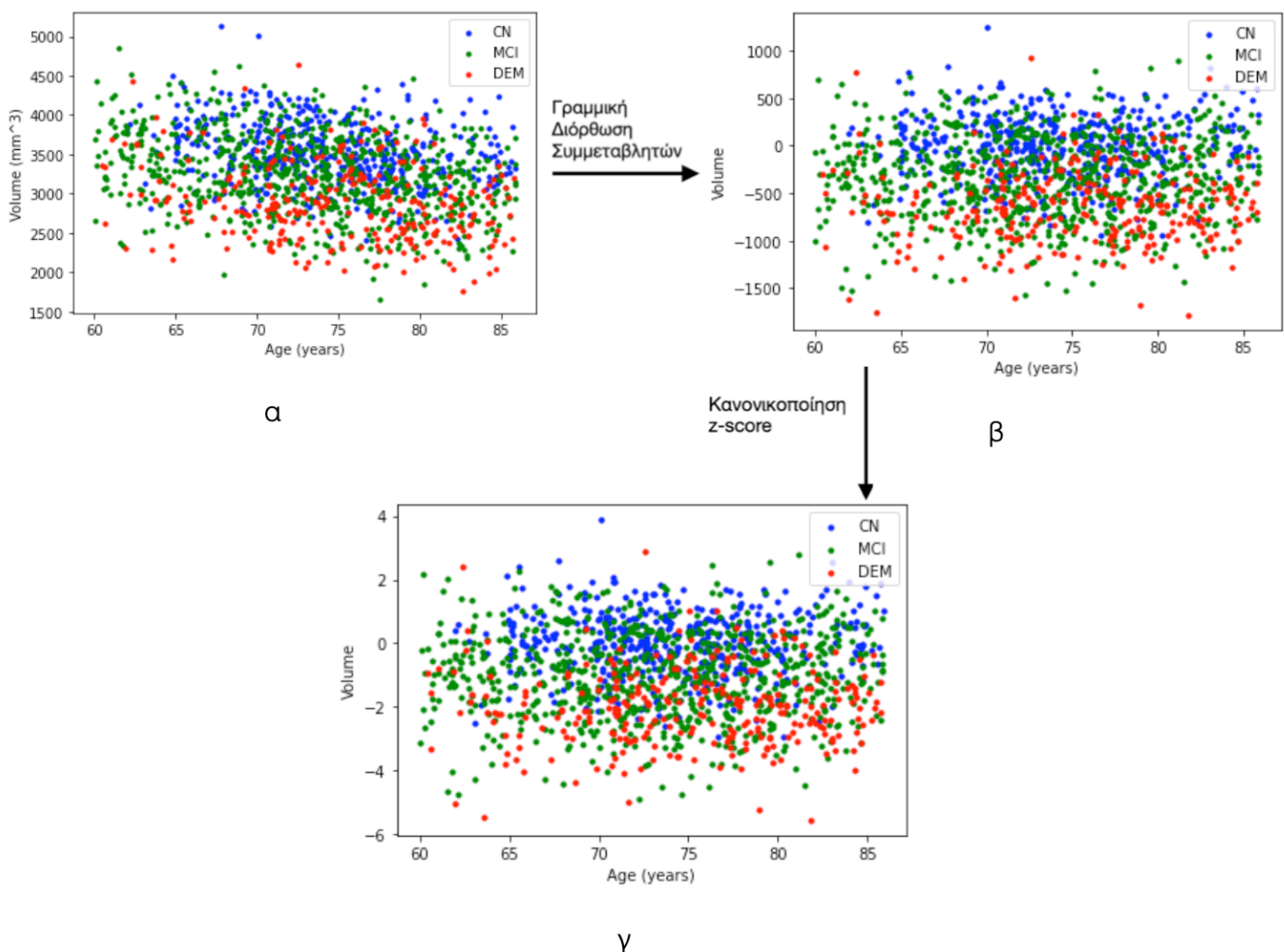
Για τον σκοπό αυτό ένα μοντέλο Γραμμικής Παλινδρόμησης (Linear Regression) εκπαιδεύτηκε με τις συμμεταβλητές της ηλικίας, του φύλου και του συνολικού όγκου του εγκεφάλου των 449 υγιών ατόμων ως προβλεπτικούς παράγοντες και τα VOIs του εγκεφάλου ως έξοδο. Πραγματοποιήθηκε προσαρμογή του μοντέλου στους προβλεπτικούς παράγοντες και στη συνέχεια το μοντέλο της Γραμμικής Παλινδρόμησης εφαρμόστηκε σε ολόκληρο το σύνολο δεδομένων, ώστε να προκύψει η πρόβλεψη για τον όγκο των περιοχών ενδιαφέροντος κάθε ατόμου. Στη συνέχεια, η τιμή πρόβλεψης για κάθε άτομο αφαιρέθηκε από την αρχική τιμή για καθένα από τα VOIs και προέκυψε η ζητούμενη τιμή υπολοίπου στην οποία είχε εξαλειφθεί η επίδραση της ηλικίας και του φύλου.

Το τελευταίο βήμα της προεπεξεργασίας αφορά στην κανονικοποίηση των δεδομένων. Προκειμένου να διατηρηθεί η κατανομή των αρχικών δεδομένων, δηλαδή η διαφοροποίηση που προκύπτει στον όγκο των περιοχών ενδιαφέροντος και στην αντίστοιχη ατροφία ανάμεσα στις τρεις κλάσεις, βρέθηκε ο μέσος όρος και η τυπική απόκλιση στην τιμή κάθε VOI χαρακτηριστικού για τα υγιή άτομα (κλάση CN) και υπολογίστηκε η κανονικοποίηση z-score για όλα τα άτομα, με τον ακόλουθο τρόπο:

$$z_score_i = \frac{x_i - \mu_{CN}}{\sigma_{CN}},$$

όπου x_i η τιμή του χαρακτηριστικού για την οποία υπολογίζεται το z_score, μ_{CN} ο μέσος όρος της αντίστοιχης τιμής χαρακτηριστικού για την κλάση των υγιών ατόμων CN και σ_{CN} η τυπική απόκλιση του χαρακτηριστικού για τη CN κλάση.

Στην Εικόνα 11 διακρίνονται τα διαγράμματα διασποράς των τιμών του όγκου του αριστερού ιπποκάμπου σε συνάρτηση με την ηλικία των συμμετεχόντων. Οι τιμές που απεικονίζονται με μπλε προέρχονται από την κλάση των υγιών ατόμων (CN), οι τιμές με πράσινο αντιστοιχούν στα άτομα από την MCI κλάση και με κόκκινο στα δεδομένα των ατόμων με άνοια (κλάση DEM). Το διάγραμμα (α) προέκυψε από τις αρχικές τιμές των δεδομένων πριν την προεπεξεργασία και φαίνεται πως όσο αυξάνεται η ηλικία των ατόμων, τόσο χαμηλότερες τιμές όγκου λαμβάνει η περιοχή ενδιαφέροντος, δηλαδή υπάρχει μία φθίνουσα σχέση ανάμεσα στον όγκο του ιπποκάμπου και στη ηλικία των συμμετεχόντων. Το διάγραμμα (β) προέκυψε από τις τιμές του αριστερού ιπποκάμπου όπως αυτές διαμορφώθηκαν μετά τη γραμμική διόρθωση συμμεταβλητών που εφαρμόστηκε και είναι εμφανές πως η πρότερη φθίνουσα σχέση που είχαν οι τιμές του χαρακτηριστικού με την ηλικία έχει εξαλειφθεί, αλλά παραμένει γραμμική. Τέλος, στο διάγραμμα (γ) αναπαρίστανται οι τιμές, όπως έχουν διαμορφωθεί μετά την κανονικοποίηση και φαίνεται πως σε σχέση με το διάγραμμα (β) δεν έχει διαταραχθεί η κατανομή των δεδομένων. Από τα διαγράμματα είναι επίσης εμφανές ότι τα υγιή άτομα (κλάση CN) καταλαμβάνουν το ανώτερο μέρος του διαγράμματος, έχουν δηλαδή υψηλότερες συγκριτικά τιμές όγκου του ιπποκάμπου, τα άτομα από την κλάση MCI απαντώνται σε όλη την επιφάνεια του διαγράμματος και οι ασθενείς με άνοια (κλάση DEM) τοποθετούνται στην περιοχή του διαγράμματος που βρίσκεται χαμηλότερα, δηλαδή έχουν τις μικρότερες συγκριτικά τιμές του χαρακτηριστικού.



Εικόνα 11: Διαγράμματα Διασποράς των τιμών όγκου του αριστερού ιπποκάμπου συναρτήσει της ηλικίας των συμμετεχόντων: α) πριν την προεπεξεργασία β) μετά τη γραμμική διόρθωση γ) μετά την κανονικοποίηση. Με μπλε συμβολίζονται τα υγιή άτομα (CN), με πράσινο τα άτομα με Ήπιας μορφής νοητική εξασθένιση (MCI) και με κόκκινο τα άτομα με άνοια (DEM).

4.3 Ταξινόμηση Πολλαπλών Κλάσεων

Προκειμένου να αναζητηθεί ο βέλτιστος ταξινομητής για τον διαχωρισμό των τριών κλάσεων δεδομένων, έπρεπε το πρόβλημα να διαμορφωθεί με τέτοιο τρόπο, ώστε να μπορούν να εφαρμοστούν στα δεδομένα δυαδικοί ταξινομητές. Για τον λόγο αυτό, χρησιμοποιήθηκαν δύο μέθοδοι διαίρεσης προβλημάτων πολλαπλών κλάσεων σε απλούστερα δυαδικά προβλήματα.

4.3.1 Ένα-προς-Όλα

Όπως προαναφέρθηκε (βλ. 3.1.4.2) το σχήμα διαίρεσης Ένα-προς-Όλα χωρίζει ένα πρόβλημα m κλάσεων σε m δυαδικά προβλήματα. Δημιουργούνται επομένως, m σύνολα δεδομένων, στα οποία μία από τις αρχικές κλάσεις αντιπαρατίθεται σε όλες τις άλλες που συναποτελούν την αντίπαλη κλάση δεδομένων. Στη προκειμένη περίπτωση, δημιουργήθηκαν τα ακόλουθα 3 σύνολα δεδομένων: CN - (MCI & DEM), MCI - (CN & DEM) και DEM - (CN & MCI), στα οποία εκπαιδεύτηκαν τρεις πανομοιότυποι ταξινομητές, ένας για το κάθε δυαδικό υποπρόβλημα. Επομένως, καθεμία από τις κλάσεις αντιπαρατέθηκε στο σύνολο των άλλων δύο και αποδόθηκε σε αυτήν η ετικέτα 1 (“θετική” κλάση), ενώ στις υπόλοιπες η ετικέτα 0 (“αρνητική κλάση”). Κάθε στιγμιότυπο του αρχικού συνόλου εμφανίζεται από μία φορά και στα τρία δυαδικά υποπροβλήματα. Για την απόκτηση της τελικής πρόβλεψης, χρησιμοποιήθηκε η τεχνική της στρατηγικής μέγιστης αξιοπιστίας (βλ. 3.1.4.2), όπου η κλάση εξόδου αποτιμάται από τον ταξινομητή με τη μέγιστη θετική έξοδο, δηλαδή με τη μεγαλύτερη συγκριτικά πιθανότητα που αποδόθηκε στο στιγμιότυπο να ανήκει σε μία από τις “θετικές” κλάσεις.

Έγιναν δοκιμές με διαφορετικά μοντέλα ταξινομητών, με τρία κυριότερα τα Τυχαία Δάση (Random Forests - RF), τις Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines -SVM) και το μοντέλο Gradient Boosting. Ωστόσο, η προβλεπτική ικανότητα των ταξινομητών με το σχήμα Ένα-προς-Όλα δεν ήταν ικανοποιητική.

4.3.2 Ένα-προς-Ένα

Το επόμενο σχήμα διαίρεσης που εφαρμόστηκε είναι το Ένα-προς-Ένα (βλ. 3.1.4.1), το οποίο χωρίζει ένα πρόβλημα m κλάσεων σε $m \cdot \frac{m-1}{2}$ δυαδικά προβλήματα, λαμβάνοντας υπόψιν όλα τα πιθανά ζεύγη κλάσεων. Στην προκειμένη περίπτωση, δημιουργήθηκαν 3 ξεχωριστά σύνολα δεδομένων, τα οποία περιλαμβάνουν τα ζεύγη κλάσεων CN - MCI, MCI - DEM και CN - DEM, επομένως κάθε στιγμιότυπο εμφανίζεται σε δύο από τα τρία σύνολα. Για κάθε ζεύγος κλάσεων, ένας δυαδικός ταξινομητής, πανομοιότυπος με τους άλλους δύο, εκπαιδεύτηκε πάνω στο αντίστοιχο σύνολο δεδομένων και υπολόγισε την πιθανότητα ένα δείγμα να ανήκει σε καθεμία από τις δύο κλάσεις. Στη συνέχεια, για την απόκτηση της τελικής πρόβλεψης, επιλέχθηκε η σταθμισμένη στρατηγική ψήφου που περιγράφηκε παραπάνω (βλ. 3.1.4.1), όπου από τις δύο φορές που ταξινομήθηκε το κάθε δείγμα, υπολογίστηκε το άθροισμα των πιθανοτήτων του να ανήκει σε καθεμία από τις τρεις κλάσεις και του αποδόθηκε η ετικέτα της κλάσης με το μεγαλύτερο άθροισμα πιθανοτήτων.

Ο λόγος που επιλέχθηκε η σταθμισμένη στρατηγική ψήφου έναντι της απλής στρατηγικής ψήφου, η οποία υπολογίζει πόσες ψήφους έχει λάβει η κάθε κλάση για την ταξινόμηση ενός στιγμιότυπου και του αποδίδει την ετικέτα αυτής με το μεγαλύτερο πλήθος ψήφων, είναι διότι το κάθε δείγμα ανήκει σε άρτιο πλήθος δυαδικών υποπροβλημάτων, με αποτέλεσμα να δημιουργούνται πολλές ισοψηφίες κατά τον υπολογισμό της τελικής πρόβλεψης.

Οι ταξινομητές που χρησιμοποιήθηκαν για την εφαρμογή της μεθόδου Ένα-προς-Ένα ήταν, μεταξύ άλλων, τα Τυχαία Δάση (Random Forests - RF), οι Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines - SVM) και το μοντέλο Gradient Boosting και παρείχαν βελτιωμένα αποτελέσματα σε σύγκριση με την τεχνική Ένα-προς-Όλα. Ωστόσο, προκειμένου να βελτιωθεί περαιτέρω η ακρίβεια της ταξινόμησης, ήταν απαραίτητη και η αντιμετώπιση του προβλήματος της ανισορροπίας κλάσεων του συνόλου δεδομένων.

4.4 Συλλογική Μάθηση

Όπως φάνηκε από τον Πίνακα 1, η κλάση MCI είναι παραπάνω από διπλάσια σε πλήθος από την κλάση DEM και 65% μεγαλύτερη από την κλάση CN, γεγονός που υποδεικνύει πως υπάρχει μεγάλη ανισορροπία κλάσεων στο σύνολο δεδομένων με κλάση πλειοψηφίας την MCI. Το γεγονός αυτό καθιστά απαραίτητη την εφαρμογή μίας τεχνικής που να αντιμετωπίζει αποτελεσματικά το πρόβλημα της ανισορροπίας και να βελτιώνει την προβλεπτική ικανότητα του μοντέλου.

Για τον λόγο αυτό εφαρμόστηκε η μέθοδος συλλογικής μάθησης bagging (βλ. 3.1.3) που επιδρά στο επίπεδο των δεδομένων και αποτελείται από ένα στάδιο δειγματοληψίας (bootstrap) και ένα στάδιο συνδυασμού των επιμέρους προβλέψεων για τα διαφορετικά σύνολα δεδομένων που δημιουργήθηκαν κατά τη δειγματοληψία (aggregating).

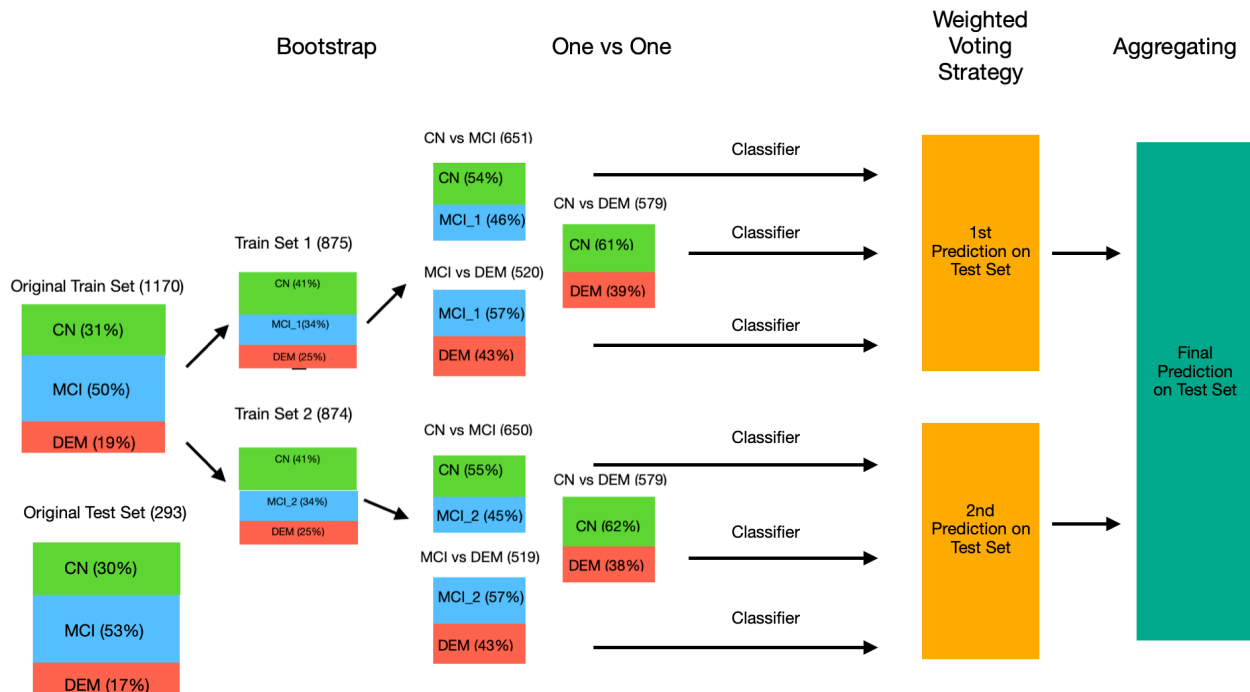
Από το αρχικό σύνολο δεδομένων τα στιγμιότυπα χωρίστηκαν, με τυχαίο τρόπο, στο σύνολο εκπαίδευσης (train set) και στο σύνολο δειγμάτων ελέγχου (test set). Ως προς το βήμα της δειγματοληψίας, επιλέχθηκε η τεχνική Bagging Ensembles Variation (βλ. 3.1.3.1), κατά την οποία η κλάση πλειοψηφίας, εν προκειμένω η κλάση MCI, χωρίζεται σε επιμέρους τμήματα και δημιουργούνται νέα σύνολα εκπαίδευσης που αποτελούνται από ένα από τα τμήματα της κλάσης πλειοψηφίας και όλα τα δείγματα των κλάσεων μειοψηφίας (CN και DEM). Στην προκειμένη περίπτωση, η κλάση MCI χωρίστηκε σε δύο ίσα τμήματα και από αυτά, σε συνδυασμό με όλα τα δείγματα από τις άλλες δύο κλάσεις, προέκυψαν δύο νέα σύνολα εκπαίδευσης. Στην Εικόνα 12, διακρίνεται η νέα σύσταση των συνόλων εκπαίδευσης μετά το βήμα της δειγματοληψίας (bootstrap). Όπως φαίνεται, η μεγάλη ανισορροπία του αρχικού συνόλου ελαττώθηκε σημαντικά, όμως μέχρι εκείνο το βήμα, εξακολουθεί να υπάρχει ανισορροπία με κλάση πλειοψηφίας αυτή τη φορά, την κλάση των υγιών ατόμων (CN).

Στη συνέχεια, καθένα από τα επιμέρους σύνολα εκπαίδευσης, χωρίστηκε σε τρία νέα σύνολα εκπαίδευσης με τη τεχνική Ένα-προς-Ένα που περιγράφηκε παραπάνω, ώστε μετά το βήμα αυτό, τα δυαδικά σύνολα εκπαίδευσης να είναι, τελικά, σχεδόν ισορροπημένα, όπως φαίνεται στο βήμα “One vs One” της Εικόνας 12. Για καθένα από τα δύο σύνολα εκπαίδευσης, επιστρατεύτηκαν τρεις πανομοιότυποι ταξινομητές, των οποίων οι προβλέψεις από τα επιμέρους δυαδικά υποπροβλήματα, συμψηφίστηκαν σε μία με τη σταθμισμένη στρατηγική ψήφου (weighted voting strategy). Από τις δύο προβλέψεις που προέκυψαν για το σύνολο ελέγχου, με παρόμοιο τρόπο, αυτές συνδυάστηκαν στο βήμα aggregating της bagging τεχνικής, ώστε να προκύψει μία τελική πρόβλεψη.

Η μέθοδος συλλογικής μάθησης που περιγράφηκε, αύξησε σημαντικά την ακρίβεια πρόβλεψης, καθώς αντιμετώπισε αποτελεσματικά το πρόβλημα της ανισορροπίας και ελάττωσε τη διακύμανση των αποτελεσμάτων της ταξινόμησης, αφού τα αποτελέσματα διασταυρώθηκαν δύο φορές από τις επιμέρους προβλέψεις, προτού προκύψει η τελική πρόβλεψη και άρα μειώθηκαν τα σφάλματα της ταξινόμησης.

Οι ταξινομητές που χρησιμοποιήθηκαν για την εφαρμογή της παραπάνω μεθόδου ήταν τα Τυχαία Δάση (Random Forests - RF), οι Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines - SVM), η Λογιστική Παλινδρόμηση (Logistic Regression - LR), το μοντέλο Gradient Boosting (GB), το μοντέλο XGBoost (XGB) και ένα Πολυστρωματικό Perceptron (Multilayer Perceptron - MLP). Ωστόσο, για την περαιτέρω βελτίωση του μοντέλου, κρίθηκε

απαραίτητη η εύρεση των βέλτιστων υπερπαραμέτρων των επιμέρους ταξινομητών, ώστε να μπορεί να εκτιμηθεί ποιοι από αυτούς είναι οι καταλληλότεροι για την επίλυση του συγκεκριμένου προβλήματος και ποια είναι η παραμετροποίηση που πρέπει να εφαρμοστεί σε αυτούς.



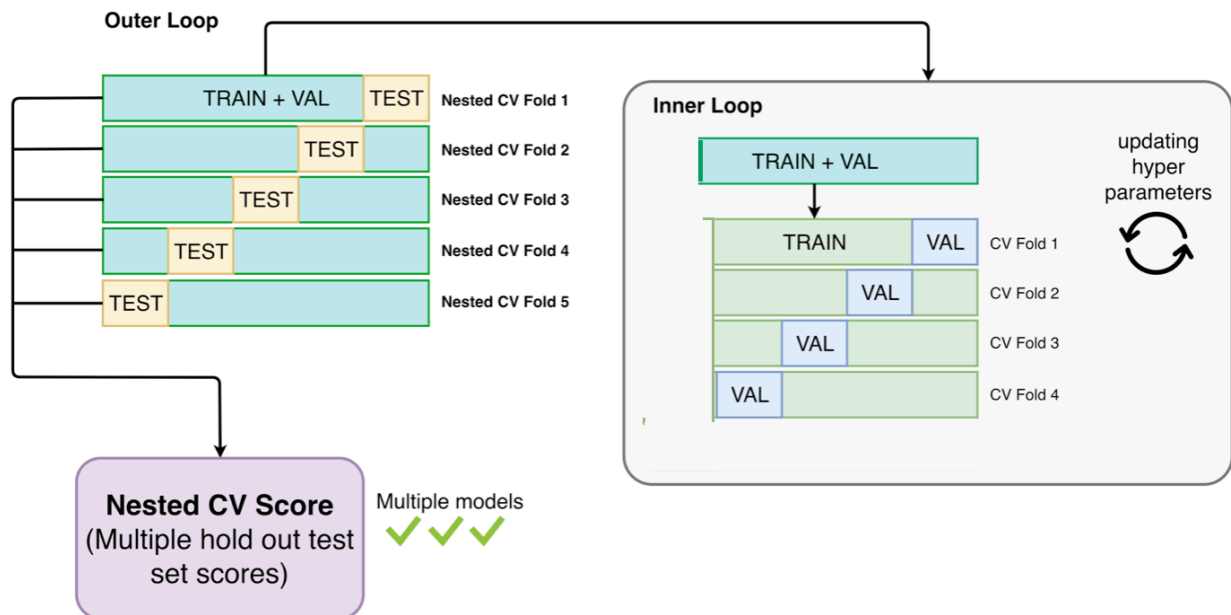
Εικόνα 12: Σχηματική αναπαράσταση του μοντέλου πρόβλεψης και του διαχωρισμού του συνόλου εκπαίδευσης κατά την εφαρμογή της τεχνικής bagging και του σχήματος Ένα-προς-Ένα.

4.5 Προσδιορισμός Βέλτιστων Υπερπαραμέτρων και Σχήμα Αξιολόγησης

Προκειμένου να προσδιοριστούν οι βέλτιστες υπερπαραμέτροι και στη συνέχεια να προκύψει ο βέλτιστος ταξινομητής, επιστρατεύθηκε η μέθοδος της εμφωλευμένης διασταυρούμενης επικύρωσης (nested cross validation). Συγκεκριμένα, το σύνολο δεδομένων χωρίστηκε σε 5 ίσα μέρη για να εφαρμοστεί διασταυρούμενη επικύρωση 5 πτυχών (5-fold cross validation). Το σύνολο εκπαίδευσης καθενός από αυτά τα folds, χωρίστηκε εκ νέου σε 4 τμήματα, προκειμένου να εφαρμοστεί σε αυτό διασταυρούμενη επικύρωση 4 πτυχών και να προσδιοριστούν οι βέλτιστες υπερπαραμέτροι για καθένα από τα 5 αρχικά folds. Η διαδικασία αυτή περιγράφεται σχηματικά στην Εικόνα 13.

Αρχικά, το σύνολο δεδομένων χωρίστηκε 5 φορές σε σύνολο εκπαίδευσης και σύνολο ελέγχου σε αναλογία 80:20 με διαφορετικό τμήμα των δεδομένων να αποτελεί κάθε φορά το σύνολο ελέγχου. Για καθένα από τα 5 εξωτερικά folds, τα δεδομένα εκπαίδευσης χωρίστηκαν εκ νέου 4 φορές σε σύνολο εκπαίδευσης και σύνολο επικύρωσης με αναλογία 75:25 και με διαφορετικό τμήμα να είναι κάθε φορά το σύνολο επικύρωσης. Πραγματοποιήθηκε διασταυρούμενη επικύρωση 4 πτυχών με εφαρμογή της τεχνικής αναζήτησης πλέγματος (grid search) για την εύρεση των βέλτιστων υπερπαραμέτρων. Από τις 4 επαναλήψεις, υπολογίστηκε ο μέσος όρος των τιμών του σκορ Συνάρτ και της ισορροπημένης ακρίβειας για την πρόβλεψη στο εκάστοτε σύνολο επικύρωσης, ώστε να επιλεγούν οι τιμές των υπερπαραμέτρων που μεγιστοποιούν τη μέση τιμή των μετρικών

επίδοσης. Για καθέναν από τους έξι ταξινομητές που χρησιμοποιήθηκαν, ορίστηκε ένα σύνολο από διακριτές τιμές για συγκεκριμένες παραμέτρους του κάθε ταξινομητή (χώρος υπερπαραμέτρων) και πραγματοποιήθηκε μία πλήρης αναζήτηση, με εξαντλητική δοκιμή όλων των συνδυασμών για τις τιμές των παραμέτρων που ορίστηκαν στον χώρο των υπερπαραμέτρων.



Εικόνα 13: Σχηματική αναπαράσταση του μοντέλου εμφωλευμένης διασταυρούμενης επικύρωσης που εφαρμόστηκε για τον εντοπισμό των βέλτιστων υπερπαραμέτρων κάθε ταξινομητή με τη μέθοδο αναζήτησης πλέγματος και τον εντοπισμό του βέλτιστου μοντέλου.

Το βέλτιστο μοντέλο για καθένα από τα 5 folds ορίστηκε ως εκείνο που έδωσε τη μέγιστη μέση τιμή ισορροπημένης ακρίβειας. Στη συνέχεια, πραγματοποιήθηκε η διασταυρούμενη επικύρωση 5 πτυχών, εφαρμόζοντας κάθε φορά στο εκάστοτε σύνολο ελέγχου το βέλτιστο μοντέλο του αντίστοιχου fold. Ύστερα από τη σύγκριση των τιμών της ισορροπημένης ακρίβειας από τα 5 folds, προέκυψε το τελικό βέλτιστο μοντέλο. Η μετρική της ισορροπημένης ακρίβειας επιλέχθηκε για την αξιολόγηση, καθώς λαμβάνει υπόψιν και την ανισορροπία των κλάσεων (βλ. 3.2). Η παραπάνω διαδικασία ακολουθήθηκε για καθέναν από τους έξι ταξινομητές.

Τέλος, για τις πέντε τιμές ισορροπημένης ακρίβειας και σκορ f1 που προέκυψαν από τη διασταυρούμενη επικύρωση 5 πτυχών για τους 6 ταξινομητές, υπολογίστηκε ο μέσος όρος και η τυπική απόκλιση, ώστε να τεθούν προς σύγκριση. Επιπλέον, πραγματοποιήθηκε για καθένα από τα 15 ζεύγη ταξινομητών ένα t-τεστ ανεξάρτητων δειγμάτων για τις 5 τιμές καθεμίας από τις δύο μετρικές, ώστε να διαπιστωθεί εάν υπάρχει στατιστικά σημαντική διαφορά ανάμεσα στη μέση τιμή της κάθε μετρικής για τους επιμέρους ταξινομητές. Το t-test ανεξάρτητων δειγμάτων χρησιμοποιείται για τη σύγκριση δύο συνόλων, των οποίων οι μέσες τιμές είναι ανεξάρτητες μεταξύ τους [52], [53].

Όσον αφορά τις υπερπαραμέτρους των επιμέρους ταξινομητών, αυτές προέκυψαν ως εξής:

- Τυχαία Δάση

Για τον ταξινομητή Τυχαίων Δασών (RF), οι υπερπαράμετροι που επιλέχθηκαν για την παραμετροποίηση ήταν το πλήθος δέντρων ($n_estimators$), το μέγιστο βάθος δέντρου (max_depth), το μέγεθος κόμβου ($min_samples_leaf$) και το μέγεθος δείγματος ($min_samples_split$) (βλ. 3.1.1.1). Για το πλήθος δέντρων δοκιμάστηκαν τέσσερις διαφορετικές τιμές με εύρος από 100 έως 800 δέντρα, καθώς όπως προαναφέρθηκε ένα μεγάλο πλήθος δέντρων οδηγεί σε γρηγορότερη σύγκλιση του ταξινομητή [26] και βρέθηκε πως η βέλτιστη τιμή ήταν τα 600 δέντρα. Το βέλτιστο βάθος δέντρου αναζητήθηκε ανάμεσα από τέσσερις διακριτές τιμές ανάμεσα στο 5 και το 100 και υπολογίστηκε η τιμή 100 ως βέλτιστη, γεγονός που επιβεβαιώνει πως όσο βαθύτερο είναι το δέντρο, τόσα περισσότερες διασπάσεις έχει και άρα τόσο περισσότερη πληροφορία συλλέγεται από τον ταξινομητή για τα δεδομένα. Για την παράμετρο του μεγέθους κόμβου ελέγχθηκαν τρεις τιμές από το 1 έως το 4, με βέλτιστη την τιμή 4. Τέλος, για την υπερπαράμετρο του μεγέθους δείγματος δοκιμάστηκαν διακριτές τιμές από το 2 έως το 10 και υπολογίστηκε ως βέλτιστη η τιμή 5. Άλλωστε, όπως προαναφέρθηκε, από ένα σχετικά μικρό μέγεθος δείγματος προκύπτουν δέντρα με μεγαλύτερη ποικιλομορφία, γεγονός που επιδρά θετικά στην ακρίβεια της πρόβλεψης [26].

- Λογιστική Παλινδρόμηση

Όσον αφορά τη Λογιστική Παλινδρόμηση (LR), οι υπερπαράμετροι, των οποίων οι βέλτιστες τιμές αναζητήθηκαν ήταν ο όρος εφαρμογής ποινής ($penalty$), η σταθερά C που καθορίζει την ισχύ της ποινής, ο αλγόριθμος επίλυσης ($solver$) και το μέγιστο πλήθος επαναλήψεων για τη σύγκλιση του αλγορίθμου επίλυσης (max_iter). Η ποινή στο πλέγμα αναζήτησης έλαβε την τιμή 'l2' που αντιπροσωπεύει τον όρο ποινής L_2 και 'none' με την οποία δεν εφαρμόζεται καμία ποινή και υπολογίστηκε πως η τιμή 'l2', και άρα η εφαρμογή ποινής, ήταν η βέλτιστη. Οι τιμές της σταθεράς C αναζητήθηκαν ανάμεσα σε πέντε τιμές στο εύρος 0.0001 έως 10000, με βέλτιστη το 0.01. Ο αλγόριθμος επίλυσης αναζητήθηκε ανάμεσα στους 'lbfgs', 'newton-cg' και 'sag' με καλύτερο τον 'lbfgs' και τέλος το βέλτιστο πλήθος των μέγιστων επαναλήψεων υπολογίστηκε ίσο με 100, ανάμεσα στις τιμές 100, 1000 και 2500.

- Μηχανές Διανυσμάτων Υποστήριξης

Οι υπερπαράμετροι των Μηχανών Διανυσμάτων Υποστήριξης (SVM) που ελέγχθηκαν ήταν η σταθερά ποινής C που ελέγχει την αντιστάθμιση ανάμεσα στην πολυπλοκότητα του υπερεπιπέδου και των σφαλμάτων εκπαίδευσης [29], η συνάρτηση πυρήνα ($kernel$) και ο συντελεστής $gamma$ για ορισμένους τύπους πυρήνων, ο οποίος σχετίζεται με το αντίστροφο της ακτίνας στην οποία τα διανύσματα υποστήριξης μπορούν να επιδρούν. Η σταθερά ποινής C έλαβε τέσσερις τιμές από το 0.1 έως το 100 κατά την αναζήτησή πλέγματος με την τιμή 1 να προκύπτει ως βέλτιστη. Ως πυρήνες δοκιμάστηκαν ο γραμμικός ('linear') και ο πυρήνας 'rbf' (βλ. 3.1.1.3) και εκτιμήθηκε πως ο 'rbf' υπερτερεί έναντι του γραμμικού στη συγκεκριμένη περίπτωση. Τέλος, όσον αφορά τη σταθερά $gamma$, οι πέντε τιμές που ελέγχθηκαν ανήκαν στο εύρος 0.0001 έως 1 με βέλτιστη την 0.001, γεγονός που συνεπάγεται πως επιλέχθηκε μία ευρεία ακτίνα για την καμπύλη του ορίου απόφασης.

- Πολυστρωματικό Perceptron

Οι υπερπαράμετροι που αφορούν στο Πολυστρωματικό Perceptron (MLP) σχετίζονται με την αρχιτεκτονική του νευρωνικού δικτύου, δηλαδή με το πλήθος των κρυφών επιπέδων και των νευρώνων από τους οποίους αυτά απαρτίζονται ($hidden_layer_sizes$), με το είδος της συνάρτησης ενεργοποίησης ($activation$), τη συνάρτηση επίλυσης ($solver$), τη σταθερά ισχύος του όρου ποινής L_2 ($alpha$) και με τον ρυθμό εκμάθησης ($learning_rate$). Δοκιμάστηκε η αρχιτεκτονική τριών κρυφών επιπέδων με 50 νευρώνες έκαστο, τριών κρυφών επιπέδων με 50, 100 και 50 νευρώνες αντίστοιχα και η αρχιτεκτονική ενός κρυφού επιπέδου με 100 νευρώνες και επιλέχθηκε η πρώτη ως βέλτιστη. Όσον αφορά στη

συνάρτηση ενεργοποίησης, επιλέχθηκε η 'relu' μεταξύ των 'relu' και 'tanh', ενώ ως συνάρτηση επίλυσης για τη βελτιστοποίηση των βαρών ο 'sgd' μεταξύ των 'sgd' και 'adam'. Μεταξύ των τιμών 0.0001 και 0.05 για τον συντελεστή alpha, η τιμή 0.05 βρέθηκε να είναι η βέλτιστη και ο ρυθμός εκμάθησης για την ενημέρωση των βαρών, ο οποίος συντέλεσε την υψηλότερη επίδοση του μοντέλου ήταν ο σταθερός ('constant') έναντι του προσαρμοστικού ('adaptive').

- Gradient Boosting

Για τον ταξινομητή Gradient Boosting οι υπερπαράμετροι των οποίων οι βέλτιστες τιμές προσδιορίστηκαν, παρόμοια με τα Τυχαία Δάση, αφορούν το πλήθος των δέντρων ($n_estimators$), το μέγιστο βάθος δέντρου (max_depth) και τον ρυθμό εκμάθησης ($learning_rate$). Το πλήθος των ταξινομητών αναζητήθηκε ανάμεσα σε τρεις τιμές στο εύρος 50 έως 500, με βέλτιστο πλήθος το 500, το μέγιστο βάθος ανάμεσα σε τέσσερις τιμές από το 3 έως το 9 με βέλτιστο το 3 και η τιμή του βαθμού εκμάθησης που απέφερε τη μέγιστη επίδοση του μοντέλου ήταν το 0.01 από το εύρος 0.01 έως 1.

- XGBoost

Οι υπερπαράμετροι του XGBoost στις οποίες εφαρμόστηκε η αναζήτηση πλέγματος ήταν ο ρυθμός εκμάθησης ($learning_rate$), το μέγιστο βάθος δέντρου (max_depth), το ελάχιστο βάρος παιδιού (min_child_weight) που χρησιμοποιείται για τον περιορισμό της υπερεκπαίδευσης, ο συντελεστής $gamma$, ο οποίος σχετίζεται με την ελάχιστη ελάττωση απωλειών, ώστε να πραγματοποιηθεί μία διάσπαση και η υποδειγματοληψία με βάση τις στήλες ($colsample_bytree$) (βλ. 3.1.1.5). Ο βέλτιστος ρυθμός εκμάθησης είχε τιμή 0.05 από τρεις τιμές στο εύρος 0.05 έως 0.25, το μέγιστο βάθος δέντρου που επιλέχθηκε ήταν 3 στο εύρος 3 έως 100, το ελάχιστο βάρος παιδιού υπολογίστηκε ίσο με 1 στο εύρος 1 έως 5, ο συντελεστής $gamma$ έλαβε βέλτιστη τιμή 0.3 έναντι του 0.1 και η υποδειγματοληψία με βάση τις στήλες υπολογίστηκε πως αποδίδει βέλτιστα στο 0.5 των δεδομένων έναντι του 0.3.

4.6 Ερμηνευσιμότητα

Μετά την εύρεση των ιδανικότερων ταξινομητών και των βέλτιστων τιμών των υπερπαραμέτρων τους για την επίλυση του συγκεκριμένου προβλήματος ταξινόμησης, εφαρμόστηκε ένα σύνολο από μεθόδους ερμηνευσιμότητας, ώστε να εξηγηθούν οι αποφάσεις των μοντέλων και να ληφθούν υπόψιν τα χαρακτηριστικά των ασθενών με έναν ερμηνεύσιμο και μετρήσιμο τρόπο.

4.6.1 Σημαντικότητα Gini

Αρχικά αναζητήθηκαν τα πιο σημαντικά χαρακτηριστικά στα οποία στηρίχθηκαν οι προβλέψεις των ταξινομητών. Για τον ταξινομητή των Τυχαίων Δασών και τους ταξινομητές που χρησιμοποιούν δέντρα απόφασης ως βασική συνάρτηση μάθησης, δηλαδή το μοντέλο Gradient Boosting και το XGBoost, χρησιμοποιήθηκε ο δείκτης σημαντικότητας Gini (βλ. 3.3.1). Για την απόκτηση της κατάταξης των σημαντικότερων χαρακτηριστικών, το πρόβλημα χωρίστηκε κατά την Ένα-προς-Ένα τεχνική σε τρία ξεχωριστά υποπροβλήματα, όπου οι κλάσεις εμφανίζονται ανά δύο. Επομένως, καθένας από τους τρεις ταξινομητές εκπαιδεύτηκε με τις βέλτιστες τιμές των υπερπαραμέτρων του στα τρία δυαδικά σύνολα δεδομένων με ζεύγη κλάσεων τα CN - MCI, MCI - DEM και CN - DEM και προέκυψε από κάθε υποπρόβλημα μία κατάταξη των χαρακτηριστικών σε φθίνουσα σειρά της τιμής του δείκτη Gini.

4.6.2 Πρόσθετες Shapley Επεξηγήσεις

Επιπλέον, αναζητήθηκαν τα σημαντικότερα χαρακτηριστικά για τον ταξινομητή SVM και για τον λόγο αυτό επιστρατεύθηκε η μέθοδος των πρόσθετων Shapley επεξηγήσεων (SHAP) (βλ. 3.3.2). Στα τρία δυαδικά υποπροβλήματα της μεθόδου Ένα-προς-Ένα εκπαιδεύτηκαν ξεχωριστά τρεις ταξινομητές με τις βέλτιστες τιμές υπερπαραμέτρων που υπολογίστηκαν κατά τη διασταυρούμενη επικύρωση και υπολογίστηκαν οι τιμές SHAP που σχετίζονται με τη συνεισφορά του κάθε χαρακτηριστικού στην πρόβλεψη των επιμέρους δειγμάτων. Προκειμένου να αποκτηθεί η συνολική σημαντικότητα του κάθε χαρακτηριστικού υπολογίζεται ο μέσος όρος των απόλυτων τιμών Shapley ανά χαρακτηριστικό για όλα στιγμιότυπα. Επομένως, για κάθε ζεύγος κλάσεων προέκυψε η αντίστοιχη γραφική παράσταση με τα χαρακτηριστικά να κατατάσσονται σε φθίνουσα σειρά σημαντικότητας.

Ωστόσο, οι πρόσθετες επεξηγήσεις Shapley, καθώς δρουν σε τοπικό επίπεδο, εκτός από τη συνολική σημαντικότητα χαρακτηριστικών, μπορούν να παρέχουν επιπλέον πληροφορία για το μοντέλο και την επίδραση των τιμών των χαρακτηριστικών κάθε επιμέρους στιγμιότυπου. Για τον λόγο αυτό, για κάθε ζεύγος κλάσεων με ταξινομητή το μοντέλο SVM, εκτός από την κατάταξη σημαντικότητας, εξήχθησαν και γραφικές παραστάσεις σύνοψης (summary plots) (βλ. 3.3.2), από τις οποίες προκύπτει επιπλέον πληροφορία σχετικά με το πώς επιδρούν οι τιμές του κάθε χαρακτηριστικού στην πρόβλεψη.

4.6.3 Τοπικές Ερμηνεύσιμες Επεξηγήσεις Ανεξάρτητες από το Μοντέλο

Σε συνέχεια της αναζήτησης επεξηγήσεων σε τοπικό επίπεδο, χρησιμοποιήθηκε το μοντέλο LIME, από το οποίο προέκυψαν επεξηγήσεις για μεμονωμένα στιγμιότυπα. Το μοντέλο εφαρμόστηκε στα τρία ξεχωριστά δυαδικά υποπροβλήματα με προβλέψεις από δύο ταξινομητές, τα Τυχαία Δάση και τις Μηχανές Διανυσμάτων Υποστήριξης και αναζητήθηκαν και στιγμιότυπα που ταξινομήθηκαν ορθά και άλλα που ταξινομήθηκαν εσφαλμένα. Συγκεκριμένα, επιλέχθηκαν για μελέτη δείγματα ορθώς θετικά (TP), ορθώς αρνητικά (TN), εσφαλμένα θετικά (FP) και εσφαλμένα αρνητικά (FN) για τα οποία αποκτήθηκαν τα διαγράμματα επεξήγησης LIME (βλ. 3.3.3) που ανέδειξαν την επίδραση των τιμών των σημαντικότερων -σύμφωνα με το μοντέλο LIME- χαρακτηριστικών στην αύξηση ή την ελάττωση της πιθανότητας του εκάστοτε δείγματος να ανήκει σε κάποια από τις κλάσεις.

4.6.4 Διάγραμμα Μερικής Εξάρτησης

Μετά από την ανάδειξη των σημαντικότερων χαρακτηριστικών από τις παραπάνω μεθόδους, καθώς και της επίδρασης αυτών στην τελική πρόβλεψη για μεμονωμένα δείγματα, θεωρήθηκε σημαντικό να μελετηθεί η επίδραση των τιμών των σημαντικών χαρακτηριστικών σε καθολικό επίπεδο. Για τον σκοπό αυτό, χρησιμοποιήθηκαν διαγράμματα μερικής εξάρτησης (PDP) (βλ. 3.3.4), τα οποία υπολογίζουν την επίδραση ενός χαρακτηριστικού, αγνοώντας όλα τα υπόλοιπα. Για καθένα από τα τρία δυαδικά υποπροβλήματα εκπαιδεύτηκε ένας ταξινομητής Τυχαίων Δασών και επιλέχθηκαν ορισμένα χαρακτηριστικά που αναδείχθηκαν από τις προηγούμενες μεθόδους ερμηνευσιμότητας. Πιο συγκεκριμένα, εκτός από τα σημαντικότερα χαρακτηριστικά που αφορούσαν κυρίως ογκομετρικές περιοχές ενδιαφέροντος (VOIs), μελετήθηκε μέσω των διαγραμμάτων εξάρτησης και η επίδραση ορισμένων χαρακτηριστικών που συνιστούν πολυμορφισμούς ενός νουκλεοτιδίου (SNP). Με τον τρόπο αυτό, ερευνήθηκε περαιτέρω η σχέση των συγκεκριμένων χαρακτηριστικών με το αποτέλεσμα της ταξινόμησης των

δειγμάτων και η προστατευτική ή επιβαρυντική τους επίδραση στην πιθανότητα των ασθενών να διαγνωστούν με κάποια από τις σχετιζόμενες με τη γήρανση του εγκεφάλου νόσους.

4.6.5 Αντιπαραθετικές Επεξηγήσεις Permute Attack

Τέλος, επιστρατεύτηκε η μέθοδος αντιπαραθετικών (counterfactual) επεξηγήσεων Permute Attack (βλ. 3.3.5), μέσω της οποίας μελετήθηκαν και αντιστράφηκαν οι προβλέψεις για συγκεκριμένα παραδείγματα ασθενών μέσω αλλαγών στις τιμές ορισμένων χαρακτηριστικών τους. Σε καθένα από τα τρία δυαδικά σύνολα δεδομένων εκπαιδεύτηκε ένας SVM ταξινομητής με τις βέλτιστες υπερπαραμέτρους και εξήχθησαν αντιπαραθετικά παραδείγματα για συγκεκριμένα στιγμιότυπα. Επιπλέον, δημιουργήθηκαν, για κάθε υποπρόβλημα, διαγράμματα που περιγράφουν τη συχνότητα με την οποία μεταβλήθηκε το κάθε χαρακτηριστικό, ώστε να αντιστραφούν τα αποτελέσματα των προβλέψεων. Σκοπός της εφαρμογής της συγκεκριμένης μεθόδου ήταν η αναγνώριση ασθενών που βρίσκονται σε υψηλό κίνδυνο ανάπτυξης ήπιας μορφής νοητικής εξασθένησης ή άνοιας και ο εντοπισμός των χαρακτηριστικών που καθιστούν τον κίνδυνο αυτό υψηλό. Επιπλέον, μελετώντας συνολικά τα χαρακτηριστικά, τα οποία μεταβλήθηκαν περισσότερες φορές για την αντιστροφή των προβλέψεων, προέκυψε ένα επιπλέον μέτρο σημαντικότητας των επιμέρους χαρακτηριστικών και της επίδρασης των τιμών τους στις προβλέψεις του μοντέλου.

5. Αποτελέσματα

Παρακάτω παρουσιάζονται οι τιμές των μετρικών για τις προβλέψεις των διαφορετικών τεχνικών ταξινόμησης που εφαρμόστηκαν, ενώ στη συνέχεια αναλύονται τα αποτελέσματα των μεθόδων ερμηνευσιμότητας που χρησιμοποιήθηκαν για την απόκτηση επεξηγήσεων σχετικά με τα επιμέρους μοντέλα πρόβλεψης και τα χαρακτηριστικά του συνόλου δεδομένων.

5.1 Αποτελέσματα Μοντέλων Ταξινόμησης

5.1.1 Σύγκριση τεχνικών

Όπως περιγράφηκε στο προηγούμενο κεφάλαιο (βλ. 4.3), αρχικά εφαρμόστηκαν δύο διαφορετικές τεχνικές διαίρεσης ενός προβλήματος πολλαπλών κλάσεων σε δυαδικά, η τεχνική Ένα-προς-Όλα (OVA) και η τεχνική Ένα-προς-Ένα (OVO). Προκειμένου να βελτιωθούν περαιτέρω τα αποτελέσματα της ταξινόμησης σε σχέση με την απλή εφαρμογή της τεχνικής Ένα-προς-Ένα, επιστρατεύτηκε η μέθοδος συλλογικής μάθησης Bagging (βλ. 4.4). Τα αποτελέσματα των τεχνικών αυτών προέκυψαν από την ταξινόμηση στο σύνολο ελέγχου κατά τον διαχωρισμό του συνόλου δεδομένων σε σύνολο εκπαίδευσης και σύνολο ελέγχου (Train Test Split - TTS). Στη συνέχεια για να βελτιωθεί η προβλεπτική ικανότητα του μοντέλου, προσδιορίστηκαν με την τεχνική της εμφωλευμένης διασταυρούμενης επικύρωσης (Cross Validation - CV) με πλέγμα αναζήτησης (βλ. 4.5) οι βέλτιστες τιμές των υπερπαραμέτρων των μοντέλων που εφαρμόστηκαν για την ταξινόμηση των δεδομένων. Με τον τρόπο αυτό, προέκυψε η μέση τιμή των μετρικών f1 σκορ και ισορροπημένης ακρίβειας από τις επαναλήψεις της διασταυρούμενης επικύρωσης 5 πτυχών, η μέγιστη τιμή των μετρικών για το βέλτιστο μοντέλο, καθώς και η τυπική απόκλιση των διαφορετικών τιμών των μετρικών ανάμεσα στις 5 επαναλήψεις της διασταυρούμενης επικύρωσης.

	Ένα προς Όλα TTS	Ένα προς Ένα TTS	Bagging TTS	Bagging CV (Μέση)	Bagging CV (Μέγιστη)	Bagging CV (STD)
	f1 σκορ Ισ. Ακρ.	f1 σκορ Ισ. Ακρ.	f1 σκορ Ισ. Ακρ.	f1 σκορ Ισ. Ακρ.	f1 σκορ Ισ. Ακρ.	f1 σκορ Ισ. Ακρ.
RF	54.4% 47.9%	75.1% 72.1%	88.8% 84.3%	90.6% 86.9%	92.9% 88.4%	1.4% 1.4%
LR	53.6% 51.1%	77.1% 76.9%	77.3% 78.7%	88.3% 85.3%	90.9% 88.3%	1.6% 1.8%
MLP	50.5% 47.2%	76.5% 75.0%	81.2% 82.9%	85.2% 83.3%	86.3% 84.5%	1.2% 1.0%
SVM	53.8% 48.5%	78.0% 76.1%	89.9% 86.1%	90.8% 87.5%	92.6% 89.6%	2.3% 1.9%
GB	55.9% 50.8%	79.2% 76.7%	86.9% 84.6%	88.2% 85.6%	93.1% 90.2%	2.6% 2.8%
XGB	53.9% 48.3%	80.7% 78.7%	86.7% 84.2%	89.6% 86.9%	93.8% 90.8%	2.4% 2.3%

Πίνακας 2: Σύγκριση αποτελεσμάτων των μετρικών του σταθμισμένου F1 σκορ και της ισορροπημένης ακρίβειας, κατά την ταξινόμηση με τα μοντέλα Τυχαίων Δασών, Μηχανών Διανυσμάτων Υποστήριξης, Gradient Boosting, Λογιστικής Παλινδρόμησης, Πολυστρωματικού Perceptron και XGBoost ανάμεσα στις τεχνικές Ένα-προς-Όλα (OVA), Ένα-προς-Ένα (OVO) και Bagging με OVO με διαχωρισμό σε σύνολα εκπαίδευσης και ελέγχου (TTS), καθώς και μετά τον προσδιορισμό των βέλτιστων υπερπαραμέτρων με διασταυρούμενη επικύρωση (CV)

Στον Πίνακα 2 παρουσιάζονται συγκριτικά τα αποτελέσματα της μετρικής του σταθμισμένου f1 σκορ και της ισορροπημένης ακρίβειας, για την ταξινόμηση με τις παραπάνω τεχνικές χρησιμοποιώντας ως ταξινομητές τα μοντέλα των Τυχαίων Δασών, της Λογιστικής Παλινδρόμησης, του Πολυστρωματικού Perceptron, των Μηχανών Διανυσμάτων Υποστήριξης, του Gradient Boosting και του XGBoost.

Όπως παρατηρεί κανείς, συγκριτικά με την τεχνική Ένα προς Όλα, στο σχήμα διαίρεσης Ένα-προς-Ένα τα αποτελέσματα των δύο μετρικών είναι πολύ υψηλότερα και για τους έξι ταξινομητές, ενώ στη συνέχεια με την εφαρμογή της μεθόδου συλλογικής μάθησης bagging υπάρχει σημαντική βελτίωση των αποτελεσμάτων. Ύστερα από τον προσδιορισμό των βέλτιστων υπερπαραμέτρων για τους ταξινομητές και την εφαρμογή της εμφωλευμένης διασταυρούμενης επικύρωσης (βλ. 4.5) στο μοντέλο της τεχνικής bagging με τη μέθοδο Ένα-προς-Ένα, προέκυψε η μέση τιμή των μετρικών για τις 5 επαναλήψεις της διασταυρούμενης επικύρωσης, καθώς και για το καλύτερο μοντέλο ανάμεσα στις επαναλήψεις αυτές. Παρατηρεί κανείς πως υπάρχει περαιτέρω βελτίωση των μετρικών και στη μέση τιμή αλλά και στη μέγιστη.

Για τον εντοπισμό του βέλτιστου μοντέλου η μέθοδος της διασταυρούμενης επικύρωσης εφαρμόστηκε σε έξι διαφορετικούς ταξινομητές, για τους οποίους ελήφθη μέση και η βέλτιστη τιμή, καθώς και η τυπική απόκλιση για τις μετρικές του σταθμισμένου f1 σκορ και της ισορροπημένης ακρίβειας. Από τον παραπάνω πίνακα, είναι εμφανές πως το μοντέλο με την υψηλότερη μέση τιμή τις δύο μετρικές είναι οι Μηχανές Διανυσμάτων Υποστήριξης, ενώ ακολουθούν τα Τυχαία Δάση και το μοντέλο XGBoost. Παρόλο που τα μοντέλα XGBoost και Gradient Boosting έχουν υψηλότερη τιμή μέγιστης ισορροπημένης ακρίβειας από τις Μηχανές Διανυσμάτων Υποστήριξης, η τελική αξιολόγηση εστιάζει στη μέση τιμή που προέκυψε για τις μετρικές από το σχήμα διασταυρούμενης επικύρωσης, χάρη στο οποίο έχει ελαττωθεί η προκατάληψη (bias). Άλλωστε, οι Μηχανές Διανυσμάτων Υποστήριξης παρουσιάζουν και χαμηλότερη τιμή στην τυπική απόκλιση των μετρικών σε σχέση με τους δύο αυτούς ταξινομητές.

	Τυχαία Δάση	Λογιστική Παλινδρόμηση	Πολυστρωματικό Perceptron	Μηχανές Διανυσμάτων Υποστήριξης	Gradient Boosting	XGBoost
Τυχαία Δάση	1.0000	0.1745	0.0026	0.6609	0.3894	0.9935
Λογιστική Παλινδρόμηση	0.1745	1.0000	0.0885	0.1312	0.8797	0.2896
Πολυστρωματικό Perceptron	0.0026	0.0885	1.0000	0.0050	0.1687	0.0201
Μηχανές Διανυσμάτων Υποστήριξης	0.6609	0.1312	0.0050	1.0000	0.2833	0.7252
Gradient Boosting	0.3894	0.8797	0.1687	0.2833	1.0000	0.4622
XGBoost	0.9935	0.2896	0.0201	0.7252	0.4622	1.0000

Πίνακας 3: Συμμετρικός πίνακας των p-value που προέκυψαν ανά ζεύγος ταξινομητών από το στατιστικό t-test για τις πέντε τιμές της ισορροπημένης ακρίβειας από τη διασταυρούμενη επικύρωση για τους έξι ταξινομητές.

Για την τελική αξιολόγηση των μοντέλων, εφαρμόστηκε στη μέση τιμή της ισορροπημένης ακρίβειας, για καθένα από τα 15 ζεύγη ταξινομητών στατιστικό t-τεστ, ώστε να προσδιοριστεί εάν υπάρχει στατιστικά σημαντική διαφορά ανάμεσα στις μέσες τιμές των μετρικών των έξι ταξινομητών. Τα αποτελέσματα της p-value του t-τεστ, για τη μετρική της ισορροπημένης ακρίβειας, φαίνονται στον Πίνακα 3, όπου και έχουν επισημανθεί με πράσινο χρώμα τα ζεύγη ταξινομητών για τα οποία η p-value είναι μικρότερη του 0.05, γεγονός που υποδεικνύει στατιστικά σημαντική διαφορά για τις μέσες τιμές του εκάστοτε ζεύγους.

Όπως προκύπτει από τον Πίνακα 3, ο μοναδικός ταξινομητής που παρουσιάζει στατιστικά σημαντική διαφορά για την ισορροπημένη ακρίβεια, είναι το Πολυστρωματικό Perceptron, το οποίο είχε συνολικά και τη χαμηλότερη επίδοση, όπως φαίνεται και στον Πίνακα 3. Οι υπόλοιποι ταξινομητές, δεν παρουσιάζουν μεταξύ τους στατιστικά σημαντική διαφορά ως προς τη μέση τιμή της ισορροπημένης ακρίβειας, ωστόσο μπορεί να θεωρηθεί, όπως προαναφέρθηκε, πως το βέλτιστο μοντέλο είναι αυτό που χρησιμοποιεί τον ταξινομητή των Μηχανών Διανυσμάτων Υποστήριξης, ακολουθούμενο από τα Τυχαία Δάση και το XGBoost.

5.1.2 Σύγκριση για διαφορετικά χαρακτηριστικών του συνόλου δεδομένων

Σημειώνεται πως πραγματοποιήθηκαν δοκιμές για την ταξινόμηση του συνόλου δεδομένων εφαρμόζοντας τη μέθοδο Bagging με την Ένα-προς-Ένα τεχνική, χρησιμοποιώντας τον ταξινομητή των Τυχαίων Δασών, για την εκπαίδευση με τα διαφορετικά είδη χαρακτηριστικών του συνόλου δεδομένων. Αρχικά, απομονώθηκαν τα SNP χαρακτηριστικά για την εκπαίδευση του μοντέλου και πραγματοποιήθηκε πρόβλεψη με βάση αυτά, στη συνέχεια απομονώθηκαν τα χαρακτηριστικά των ογκομετρικών περιοχών ενδιαφέροντος (VOIs) και τέλος αξιοποιήθηκαν συνδυαστικά και τα δύο είδη χαρακτηριστικών για την εκπαίδευση και την εκτέλεση της πρόβλεψης. Τα αποτελέσματα των δοκιμών αυτών παρουσιάζονται στον Πίνακα 4, όπου είναι εμφανές ότι η ταξινόμηση με τα χαρακτηριστικά SNP αποφέρει πολύ χαμηλά αποτελέσματα πρόβλεψης, ενώ η ταξινόμηση μόνο με τα VOI χαρακτηριστικά στο σύνολο εκπαίδευσης παρέχει ικανοποιητικά αποτελέσματα και πολύ κοντινά σε αυτά που προκύπτουν με τον συνδυασμό και των δύο ειδών χαρακτηριστικών. Επομένως, συμπεραίνει κανείς πως τα SNP χαρακτηριστικά συνεισφέρουν σε μικρό βαθμό στην πρόβλεψη.

	Σταθμισμένο F1- Σκορ	Ισορροπημένη Ακρίβεια
Μόνο SNP	63.7%	59.2%
Μόνο VOI	88.6%	84.9%
Συνδυασμός χαρακτηριστικών	89.0%	85.2%

Πίνακας 4: Σύγκριση αποτελεσμάτων των μετρικών του σταθμισμένου F1 σκορ και της ισορροπημένης ακρίβειας χρησιμοποιώντας μόνο τα SNP χαρακτηριστικά, μόνο τα VOI χαρακτηριστικά και τον συνδυασμό των δύο ειδών χαρακτηριστικών κατά την ταξινόμηση με το μοντέλο των Τυχαίων Δασών.

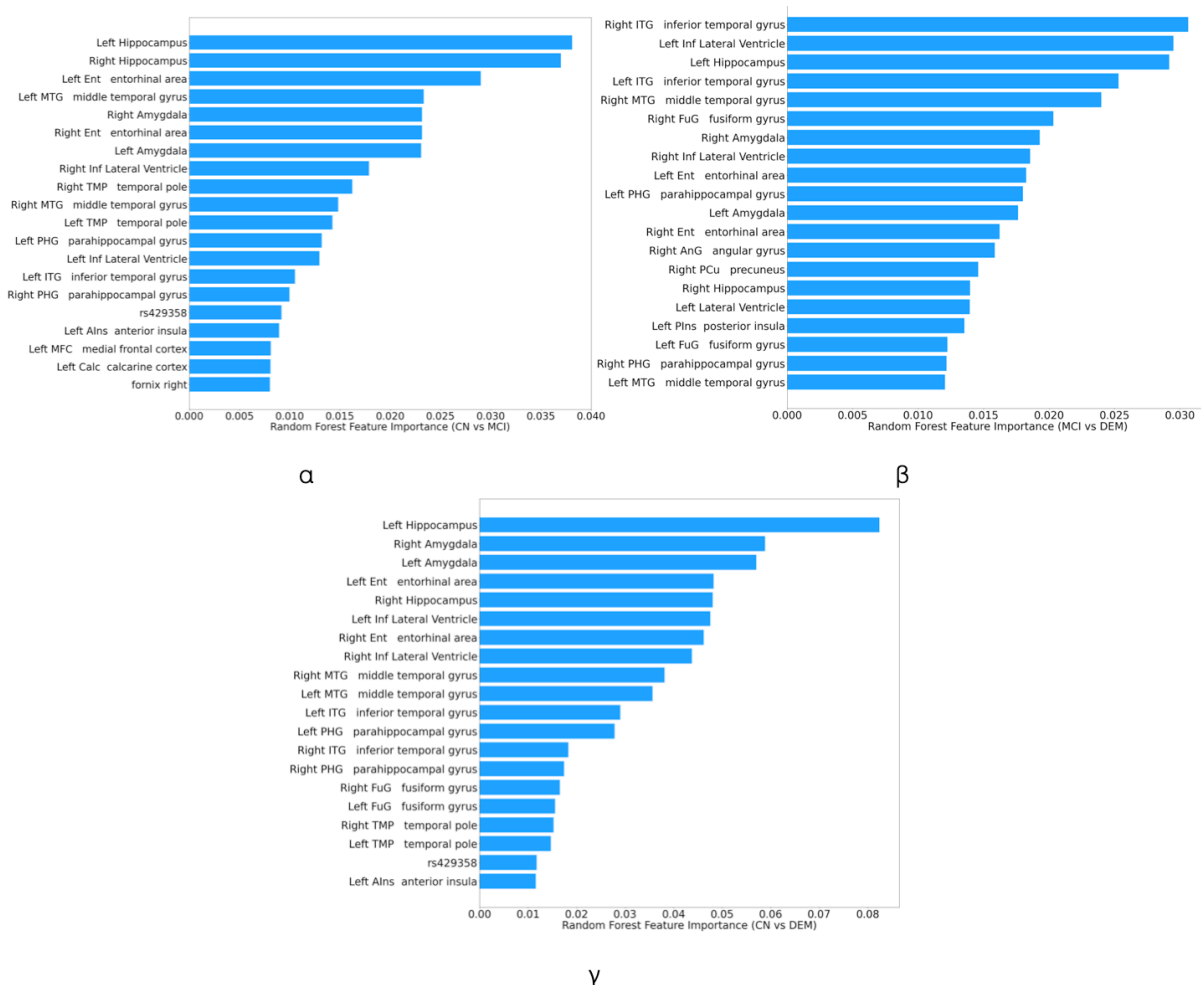
5.2 Αποτελέσματα Μοντέλων Ερμηνευσιμότητας

5.2.1 Gini Σημαντικότητα Χαρακτηριστικών

Παρακάτω παρατίθενται τα αποτελέσματα της σημαντικότητας των χαρακτηριστικών όπως προέκυψαν από την ταξινόμηση των μοντέλων Τυχαίων Δασών, Gradient Boosting και XGBoost. Χάρη στη δενδρική δομή των τριών αυτών μοντέλων, η κατάταξη σημαντικότητας αποκτήθηκε με βάση τον δείκτη καθαρότητας Gini (βλ. 3.3.1).

5.2.1.1 Τυχαία Δάση

Όπως φαίνεται από το διάγραμμα (α) της Εικόνας 14, όπου παρουσιάζονται τα 20 σημαντικότερα χαρακτηριστικά, κατά τη δυαδική ταξινόμηση ανάμεσα στις κλάσεις CN και MCI, τα σημαντικότερα χαρακτηριστικά αφορούν στις ογκομετρικές περιοχές ενδιαφέροντος (VOIs) του αριστερού και δεξιού ιπποκάμπου (left and right hippocampus), της αριστερής και δεξιάς ενδορινικής περιοχής (left and right entorhinal area), της αριστερής μεσαίας κροταφικής έλικας (left middle temporal gyrus) και της δεξιάς και αριστερής αμυγδαλής (right and left amygdala). Αξίζει, επίσης, να σημειωθεί πως μία μεγάλη πλειοψηφία των SNP χαρακτηριστικών παρέμειναν στις τελευταίες θέσεις της κατάταξης σημαντικότητας, γεγονός που υποδεικνύει πως διαδραματίζουν μικρό ρόλο στον διαχωρισμό ανάμεσα στις κλάσεις CN και MCI. Ωστόσο, ο πολυμορφισμός ενός νουκλεοτιδίου (SNP) rs429358, ο οποίος ανήκει στο γονίδιο της απολιποπρωτεΐνης E (ApoE) [54], κατατάχθηκε ανάμεσα στις 16 πρώτες θέσεις της κατάταξης σημαντικότητας χαρακτηριστικών.



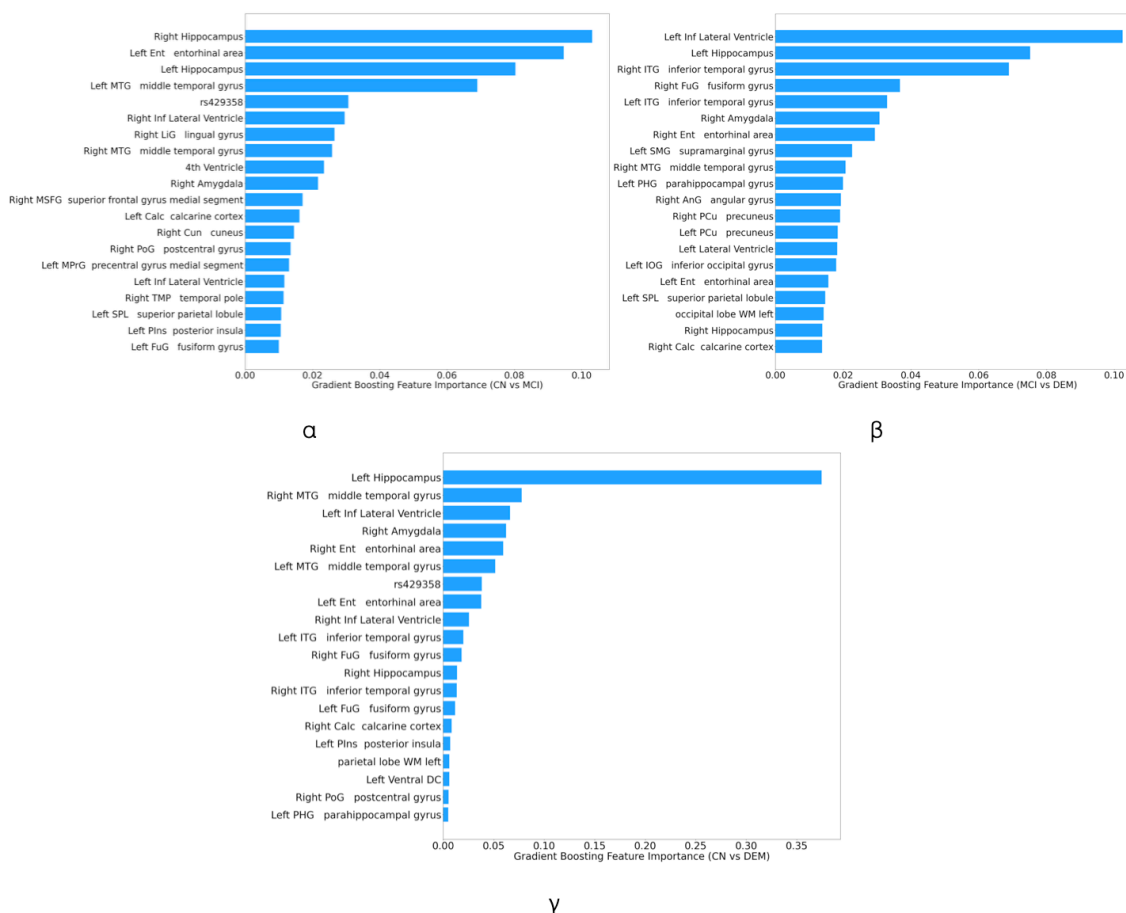
Εικόνα 14: Κατάταξη Σημαντικότητας Χαρακτηριστικών Gini κατά την ταξινόμηση με το μοντέλο Τυχαίων Δασών για το ζεύγος κλάσεων α) CN και MCI, β) MCI και DEM και γ) CN και DEM

Όσον αφορά τον διαχωρισμό ανάμεσα στις κλάσεις MCI και DEM, όπως φαίνεται και στο διάγραμμα (β) της Εικόνας 14, τα επικρατέστερα χαρακτηριστικά αποτελούνται από τη δεξιά κατώτερη κροταφική έλικα (right inferior temporal gyrus), την αριστερή και δεξιά κάτω πλευρική κοιλία (left and right inferior lateral ventricle), το αριστερό τμήμα του ιπποκάμπου, τη δεξιά μέση κροταφική έλικα με την αμυγδαλή και τον δεξιό ιππόκαμπο να λαμβάνουν λίγο χαμηλότερες θέσεις στην κατάταξη σε σύγκριση με το ζεύγος CN - MCI. Τα SNP χαρακτηριστικά βρίσκονται στις τελευταίες θέσεις της κατάταξης και δεν απαντώνται καθόλου στο διάγραμμα των πρώτων 20 θέσεων.

Τέλος, κατά τον διαχωρισμό μεταξύ των CN και DEM ατόμων, όπως φαίνεται στο διάγραμμα (γ) της Εικόνας 14, οι ογκομετρικές περιοχές ενδιαφέροντος (VOIs) που κατατάχθηκαν υψηλότερα ήταν ο αριστερός και δεξιός ιππόκαμπος, η δεξιά και αριστερή αμυγδαλή, η αριστερή και δεξιά ενδορινική περιοχή και η αριστερή και δεξιά κάτω πλευρική κοιλία, ενώ από τα SNP χαρακτηριστικά μόνο ο rs429358 πολυμορφισμός του γονιδίου ApoE εμφανίζεται στις πρώτες 20 θέσεις της κατάταξης.

5.2.1.2 Gradient Boosting

Τα 20 σημαντικότερα χαρακτηριστικά που προέκυψαν κατά την ταξινόμηση ανάμεσα στις κλάσεις CN και MCI από το μοντέλο Gradient Boosting, με τις το βέλτιστο μοντέλο του ταξινομητή αυτού, φαίνονται στο διάγραμμα (α) της Εικόνας 15. Τα κυριότερα χαρακτηριστικά περιλαμβάνουν τον δεξιό και αριστερό ιππόκαμπο, την αριστερή ενδορινική περιοχή, την αριστερή μεσαία κροταφική



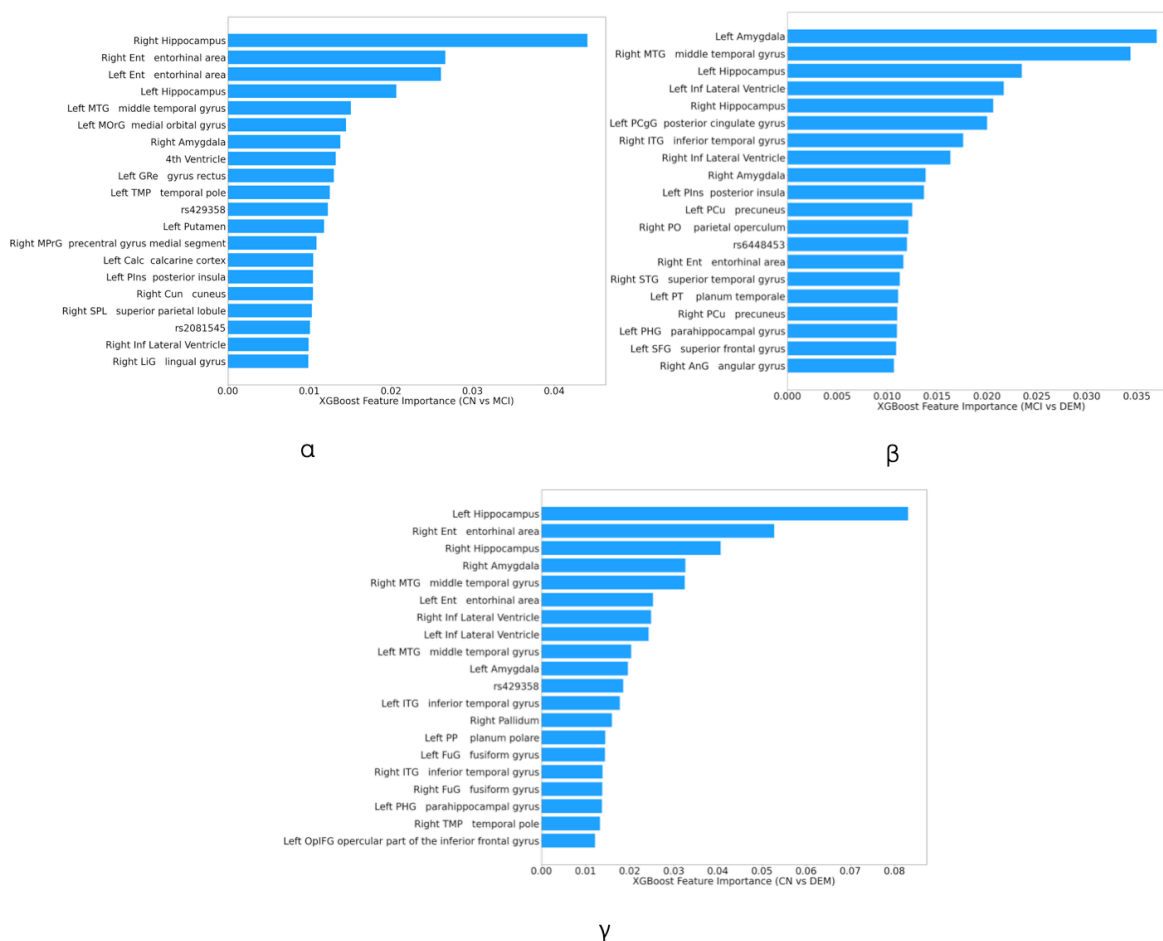
Εικόνα 15: Κατάταξη Σημαντικότητας Χαρακτηριστικών Gini κατά την ταξινόμηση με το μοντέλο Gradient Boosting για το ζεύγος κλάσεων α) CN και MCI, β) MCI και DEM και γ) CN και DEM

έλικα και το rs429358 SNP που ανήκει στο γονίδιο ApoE, και συνιστά το μοναδικό SNP που απαντάται στις πρώτες θέσεις της κατάταξης.

Παρόμοια συμπεράσματα μπορούν να εξαχθούν και για τα άλλα δύο ζεύγη κλάσεων (MCI - DEM και CN - DEM). Από το διάγραμμα (β) της Εικόνας 15, μπορεί κανείς να παρατηρήσει για τον διαχωρισμό των κλάσεων MCI και DEM, πως τα 5 κυριότερα χαρακτηριστικά αποτελούνται από την αριστερή κάτω πλευρική κοιλία, τον αριστερό ιππόκαμπο, την αριστερή και δεξιά κατώτερη κροταφική έλικα και τη δεξιά ατρακτοειδή έλικα (right fusiform gyrus). Επιπλέον, σε αντίθεση με τα άλλα δύο ζεύγη, στα πρώτα είκοσι σημαντικότερα χαρακτηριστικά δεν εμφανίζεται το SNP rs429358.

Αντίστοιχα, για το ζεύγος CN και DEM, από το διάγραμμα (γ) της Εικόνας 15, εντοπίζει κανείς ως σημαντικότερες τις περιοχές του αριστερού ιπποκάμπου, της δεξιάς και αριστερής μεσαίας κροταφικής έλικας, της αριστερής κάτω πλευρικής κοιλίας, της δεξιάς αμυγδαλής, της δεξιάς ενδορινικής περιοχής και σε αρκετά υψηλή θέση το SNP rs429358 που σχετίζεται με το γονίδιο ApoE. Αξίζει να σημειωθεί, πως οι τιμές σημαντικότητας που αποδίδονται στα χαρακτηριστικά στην περίπτωση του συγκεκριμένου μοντέλου και για τα τρία ζεύγη κλάσεων, είναι σχεδόν μία τάξη μεγέθους μεγαλύτερες συγκριτικά με τα άλλα δύο μοντέλα (Τυχαία Δάση και XGBoost) που βασίζονται στη σημαντικότητα Gini, γεγονός που ευθυγραμμίζεται με την παρατήρηση πως αρκετά από τα χαρακτηριστικά λαμβάνουν μηδενική τιμή σημαντικότητας.

5.2.1.3 XGBoost



Εικόνα 16: Κατάταξη Σημαντικότητας Χαρακτηριστικών Gini κατά την ταξινόμηση με το μοντέλο XGBoost για το ζεύγος κλάσεων α) CN και MCI, β) MCI και DEM και γ) CN και DEM

Όπως προκύπτει από το διάγραμμα (α) της Εικόνας 16 για το ζεύγος κλάσεων CN και MCI με μοντέλο ταξινόμησης το XGBoost, τα βασικότερα χαρακτηριστικά στα οποία στηρίχτηκε η ταξινόμηση ήταν ο δεξιός και αριστερός ιππόκαμπος, η αριστερή και δεξιά ενδορινική περιοχή, η αριστερή μέση κροταφική έλικα και η αριστερή μέση τροχιακή έλικα (left medial orbital gyrus). Δύο πολυμορφισμοί ενός νουκλεοτιδίου (SNPs) συμπεριλαμβάνονται στις πρώτες θέσεις της κατάταξης σημαντικότητας, όπως, το rs429358 που σχετίζεται με το γονίδιο ApoE [54], το rs2081545, το οποίο ανήκει στο γονίδιο MS4A6A [9].

Όσον αφορά τον διαχωρισμό ανάμεσα στις κλάσεις MCI και DEM, σύμφωνα με το διάγραμμα (β) της Εικόνας 16, τα σημαντικότερα χαρακτηριστικά περιλαμβάνουν την αριστερή αμυγδαλή, τη δεξιά μέση κροταφική έλικα, τον ιππόκαμπο και την κάτω πλευρική κοιλία. Από την κατάταξη απουσιάζει το rs429358 του γονιδίου ApoE, αλλά έχει συμπεριληφθεί το SNP rs6448453, το οποίο ανήκει στο γονίδιο CLNK [9].

Επιπλέον, για την ταξινόμηση ανάμεσα στις κλάσεις CN και DEM, όπως φαίνεται και από το διάγραμμα (γ) της Εικόνας 16, τα σημαντικότερα χαρακτηριστικά συμπίπτουν με την κατάταξη των δύο παραπάνω μοντέλων ταξινόμησης, καθώς αποτελούνται από τον αριστερό και δεξιό ιππόκαμπο, τη δεξιά ενδορινική περιοχή, τη δεξιά αμυγδαλή, και τη δεξιά μέση κροταφική έλικα, ανάμεσα σε άλλες σημαντικές ογκομετρικές περιοχές ενδιαφέροντος. Αξίζει να σημειωθεί, πως για το συγκεκριμένο δυαδικό υποπρόβλημα, ο μόνος πολυμορφισμός ενός νουκλεοτιδίου που εμφανίζονται στις 20 υψηλότερες θέσεις της κατάταξης είναι το rs429358 του γονιδίου ApoE.

5.2.2 Πρόσθετες Shapley Επεξηγήσεις

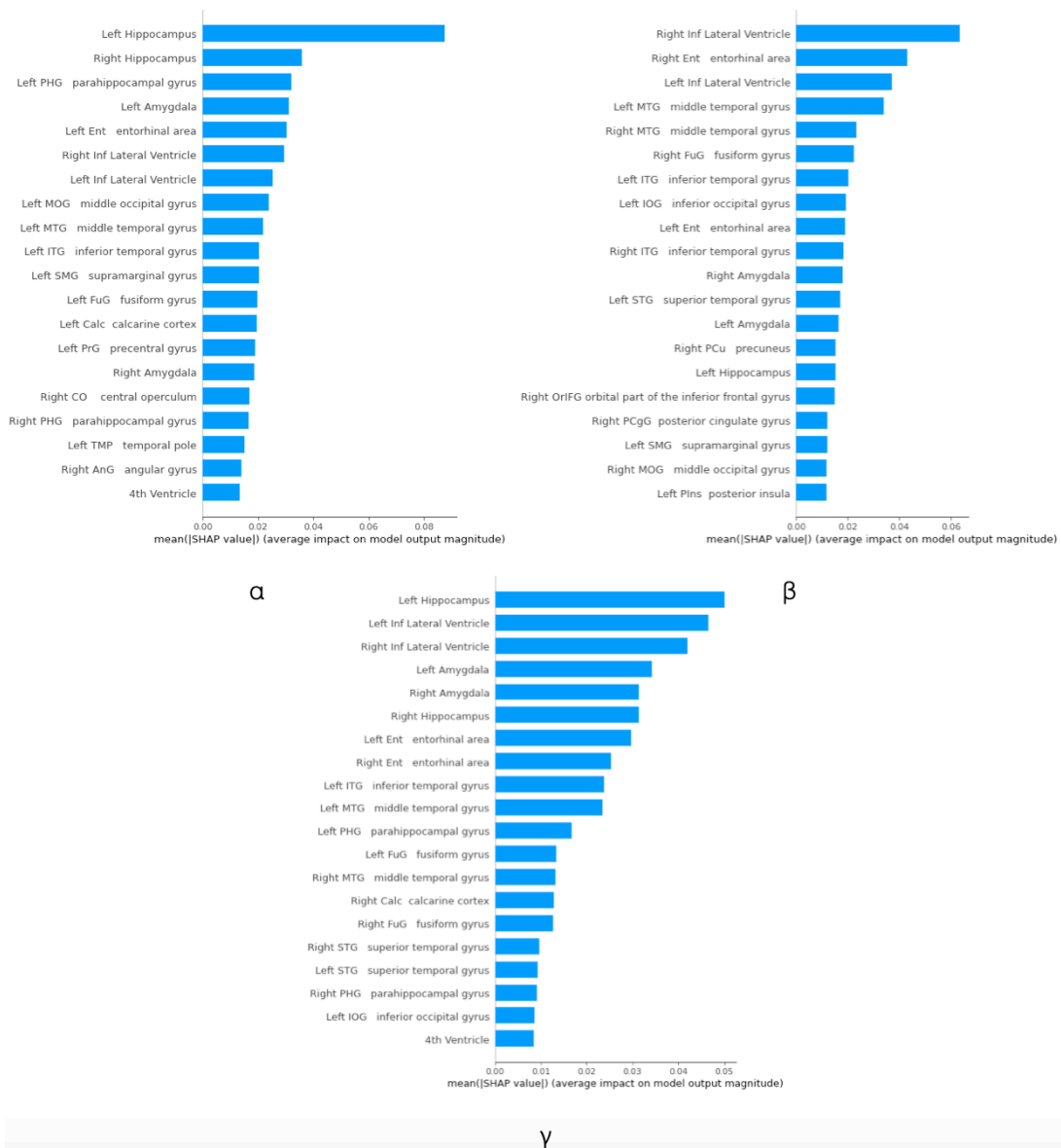
5.2.2.1 Σημαντικότητα Χαρακτηριστικών SHAP

Όσον αφορά στην κατάταξη σημαντικότητας χαρακτηριστικών για τον ταξινομητή των Μηχανών Διανυσμάτων Υποστήριξης και για κάθε ζεύγος κλάσεων, αυτή προέκυψε από τη μέθοδο Πρόσθετων Shapley Επεξηγήσεων (SHAP) (βλ. 3.3.2) και συγκεκριμένα από τον υπολογισμό του μέσου όρου των απόλυτων SHAP τιμών, όπως περιγράφηκε στο προηγούμενο κεφάλαιο (βλ. 4.6.2).

Όπως προκύπτει από το διάγραμμα (α) της Εικόνας 17, κατά τον διαχωρισμό των κλάσεων CN και MCI, παρόμοια με τα προηγούμενα μοντέλα, τα κυριότερα χαρακτηριστικά περιλαμβάνουν τον αριστερό και δεξιό ιππόκαμπο, την αριστερή παραϊπποκαμπική έλικα (left parahippocampal gyrus), την αριστερή αμυγδαλή, την αριστερή ενδορινική περιοχή, τη δεξιά και αριστερή κάτω πλευρική κοιλία και άλλες ογκομετρικές περιοχές ενδιαφέροντος, όπως η αριστερή μέση ινιακή έλικα (left middle occipital gyrus) και η αριστερή μέση κροταφική έλικα. Όλα τα χαρακτηριστικά στις πρώτες είκοσι θέσεις της κατάταξης, αποτελούνται αποκλειστικά από ογκομετρικές περιοχές ενδιαφέροντος και όχι από πολυμορφισμούς ενός νουκλεοτιδίου.

Κατά τον διαχωρισμό των MCI και DEM κλάσεων, τα επικρατέστερα χαρακτηριστικά, σύμφωνα με το διάγραμμα (β) της Εικόνας 17, είναι η δεξιά και αριστερή κάτω πλευρική κοιλία, η δεξιά ενδορινική περιοχή, η αριστερή και δεξιά μέση κροταφική έλικα, ενώ η δεξιά και αριστερή αμυγδαλή και ο αριστερός και δεξιός ιππόκαμπος βρίσκονται σε σχετικά χαμηλότερες θέσεις συγκριτικά με τα υπόλοιπα γραφήματα.

Αντίστοιχα, κατά τον διαχωρισμό των CN και DEM κλάσεων, σύμφωνα με το διάγραμμα (γ) της Εικόνας 17, τα σημαντικότερα χαρακτηριστικά αφορούν αποκλειστικά ογκομετρικές περιοχές ενδιαφέροντος, οι οποίες είναι, μεταξύ άλλων, ο αριστερός και δεξιός ιππόκαμπος, η αριστερή και δεξιά κάτω πλευρική κοιλία, η αριστερή και δεξιά αμυγδαλή και η αριστερή και δεξιά ενδορινική περιοχή.

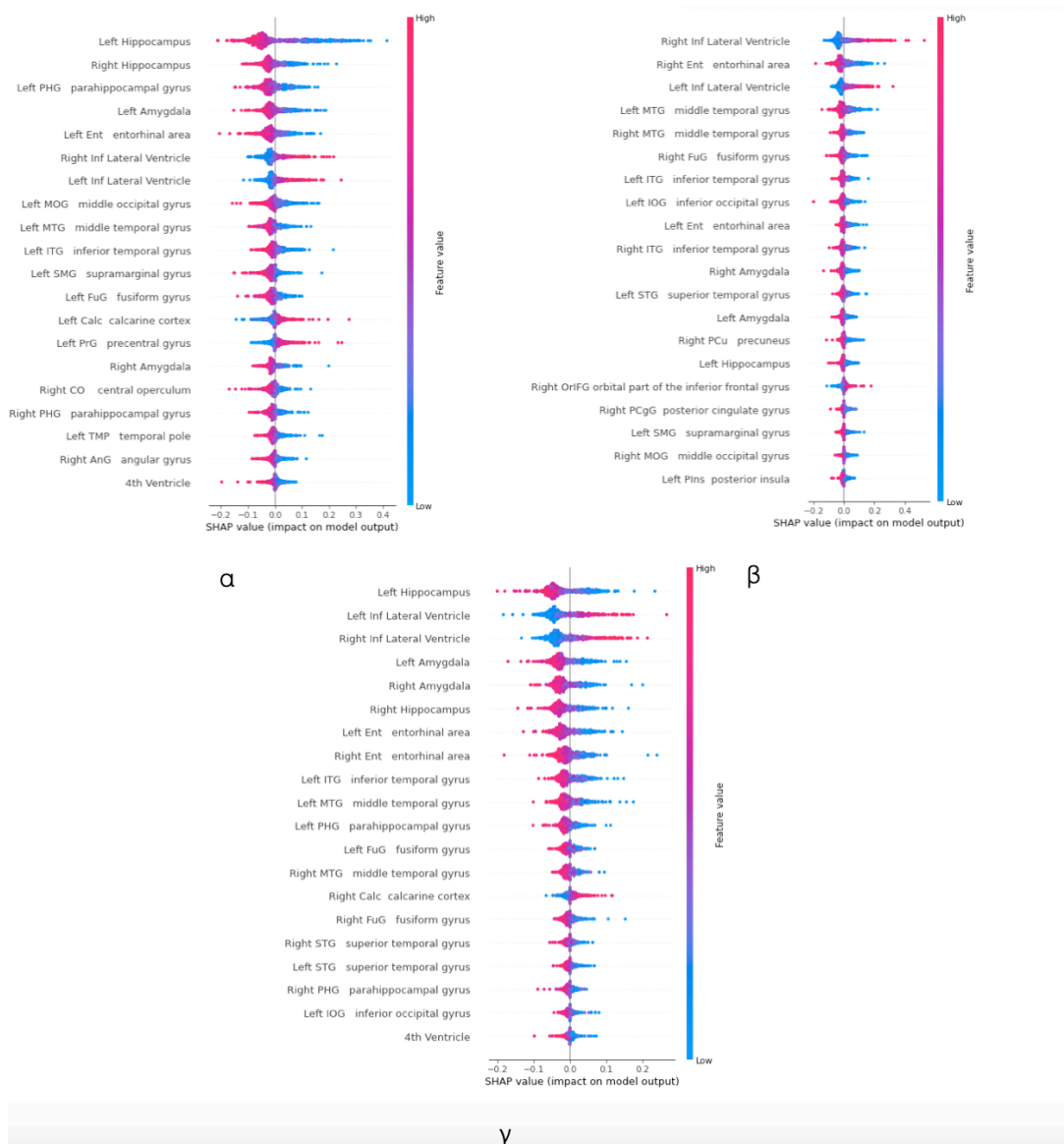


Εικόνα 17: Κατάταξη Σημαντικότητας Χαρακτηριστικών SHAP κατά την ταξινόμηση με το μοντέλο SVM για το ζεύγος κλάσεων α) CN και MCI, β) MCI και DEM και γ) CN και DEM.

5.2.2.2 Γραφική Παράσταση Σύνοψης

Στη συνέχεια, όπως περιγράφηκε στο προηγούμενο κεφάλαιο (βλ. 4.6.2), εξήχθησαν οι γραφικές παραστάσεις σύνοψης για καθένα από τα τρία δυαδικά υποπροβλήματα, όπου μπορεί κανείς να παρατηρήσει την επίδραση των τιμών κάθε χαρακτηριστικού για τα επιμέρους στιγμιότυπα. Στη γραφική παράσταση σύνοψης, τα χαρακτηριστικά είναι διατεταγμένα σε φθίνουσα σειρά σημαντικότητας στον άξονα y και κάθε σημείο αντιπροσωπεύει την SHAP τιμή του χαρακτηριστικού για ένα μεμονωμένο στιγμιότυπο. Όσο πιο κόκκινο είναι το χρώμα με το οποίο αναπαρίστανται ένα σημείο, τόσο υψηλότερη είναι τιμή του αντίστοιχου χαρακτηριστικού, ενώ όσο τείνει προς το μπλε χρώμα, τόσο χαμηλότερη είναι η τιμή. Όταν ένα σημείο βρίσκεται στο διάγραμμα στα αριστερά του άξονα x, τότε αντιπροσωπεύει ένα στιγμιότυπο με αρνητική SHAP τιμή για το συγκεκριμένο χαρακτηριστικό, γεγονός που υποδεικνύει πως η τιμή του χαρακτηριστικού ωθεί το δείγμα να ταξινομηθεί στην “αρνητική” κλάση και το αντίθετο. Στο ζεύγος CN και MCI η “θετική” κλάση είναι η MCI, στο ζεύγος CN και DEM “θετική” είναι η DEM, ενώ στο MCI και DEM ως “θετική” θεωρείται και πάλι η DEM κλάση.

Όπως είναι εμφανές, στα τρία διαγράμματα της Εικόνας 18, η κατάταξη της σημαντικότητας των χαρακτηριστικών ταυτίζεται με αυτή των διαγραμμάτων της Εικόνας 17 για τα τρία ζεύγη κλάσεων. Συνολικά από τα τρία διαγράμματα (α), (β) και (γ) της Εικόνας 18, μπορεί κανείς να παρατηρήσει ότι οι υψηλές τιμές του ιπποκάμπου, της αμυγδαλής, της ενδορινικής περιοχής, της παραϊπποκαμπικής έλικας, της μέσης κροταφικής έλικας και άλλων σημαντικών ογκομετρικών περιοχών ενδιαφέροντος ωθούν την πρόβλεψη προς την “αρνητική”, δηλαδή την “υγιέστερη” κλάση (είτε CN είτε MCI αναλόγως το ζεύγος). Αντιθέτως, οι χαμηλότερες τιμές (μπλε) των ίδιων χαρακτηριστικών, είναι ενδεικτικές της ατροφίας του εγκεφάλου και συνεισφέρουν στην ταξινόμηση του αντίστοιχου ασθενούς στη “θετική” κλάση (είτε MCI, είτε DEM ανάλογα με το δυαδικό υποπρόβλημα). Επιπλέον, όπως φαίνεται από την Εικόνα 18, οι χαμηλές τιμές της αριστερής και δεξιάς κάτω πλευρικής κοιλίας υποδεικνύουν ένα υγιές άτομο, ενώ οι υψηλότερες τιμές σημαίνουν πως το αντίστοιχο άτομο είναι πιθανότερο να ανήκει στη “θετική” κλάση (MCI ή DEM αναλόγως το ζεύγος).



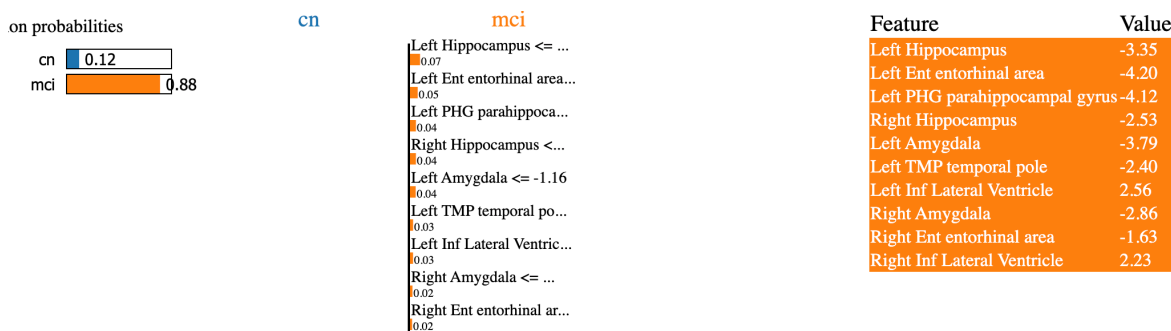
Εικόνα 18: Διαγράμματα Σύνοψης SHAP κατά την ταξινόμηση με το μοντέλο SVM για το ζεύγος κλάσεων α) CN και MCI, β) MCI και DEM και γ) CN και DEM.

5.2.3 Τοπικές Ερμηνεύσιμες Επεξηγήσεις Ανεξάρτητες από το

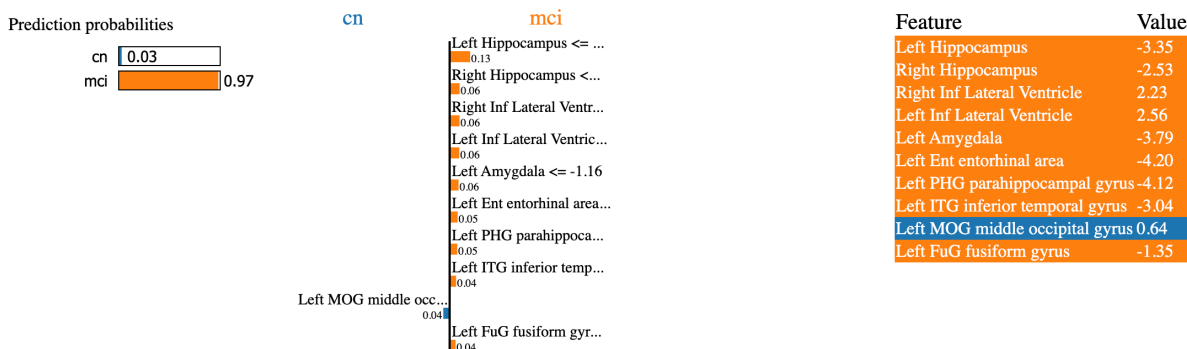
Μοντέλο

Όπως περιγράφηκε στο προηγούμενο κεφάλαιο (βλ. 4.6.3) το μοντέλο LIME εφαρμόστηκε ξεχωριστά για κάθε ζεύγος κλάσεων για δύο μοντέλα ταξινομητών.

Κατά τον διαχωρισμό ανάμεσα στις κλάσεις CN και MCI, ελήφθησαν τοπικές επεξηγήσεις για τέσσερις περιπτώσεις ασθενών, για ένα ορθώς θετικό (TP) στιγμιότυπο, ένα ορθώς αρνητικό (TN), ένα ψευδώς θετικό (FP) και ένα ψευδώς αρνητικό (FN), όπου τα CN άτομα θεωρείται ότι αποτελούν την “αρνητική” κλάση και τα MCI τη “θετική”. Οι Εικόνες 19 και 20 παρουσιάζουν μία περίπτωση του υποπροβλήματος του ζεύγους CN και MCI, όπου ο ίδιος ασθενής ταξινομήθηκε και από τους δύο ταξινομητές (Τυχαία Δάση και Μηχανές Διανυσμάτων Υποστήριξης στην σωστή κλάση των ατόμων που πάσχουν από MCI. Από τις δύο εικόνες φαίνεται πως χαρακτηριστικά που ευνοούν την ταξινόμηση ενός στιγμιότυπου στην MCI κλάση έναντι της CN, είναι οι χαμηλές τιμές όγκου του αριστερού και δεξιού ιπποκάμπου, της αριστερής ενδορινικής περιοχής και της αριστερής παραϊπποκαμπικής έλικας, ανάμεσα σε άλλες περιοχές ενδιαφέροντος, ενώ το χαρακτηριστικό που αυξάνει την πιθανότητα ενός δείγματος να ανήκει στην CN κλάση, είναι η υψηλή τιμή της αριστερής μέσης ινιακής έλικας.

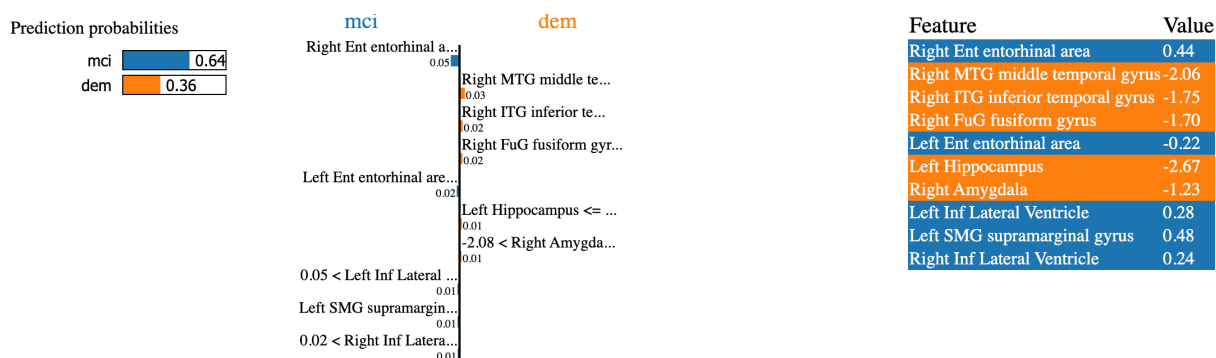


Εικόνα 19: Αποτελέσματα του μοντέλου LIME για την ταξινόμηση με ταξινομητή Τυχαίων Δασών για τις κλάσεις CN και MCI. Ορθώς θετική περίπτωση, όπου ένα MCI άτομο ταξινομείται σωστά. Αριστερά: η προβλεπόμενη πιθανότητα της συγκεκριμένης περίπτωσης για τις δύο κλάσεις, Μέση: τα χαρακτηριστικά που ωθούν την πρόβλεψη στην κλάση MCI (πορτοκαλί) και στην κλάση CN (μπλε) αντίστοιχα, Δεξιά: μια κατάταξη σημαντικότητας των χαρακτηριστικών και αντίστοιχες τιμές τους. Τα χαρακτηριστικά πορτοκαλί χρώματος ωθούν την πρόβλεψη στην κλάση MCI και τα χαρακτηριστικά μπλε χρώματος ωθούν την πρόβλεψη στην κλάση CN.

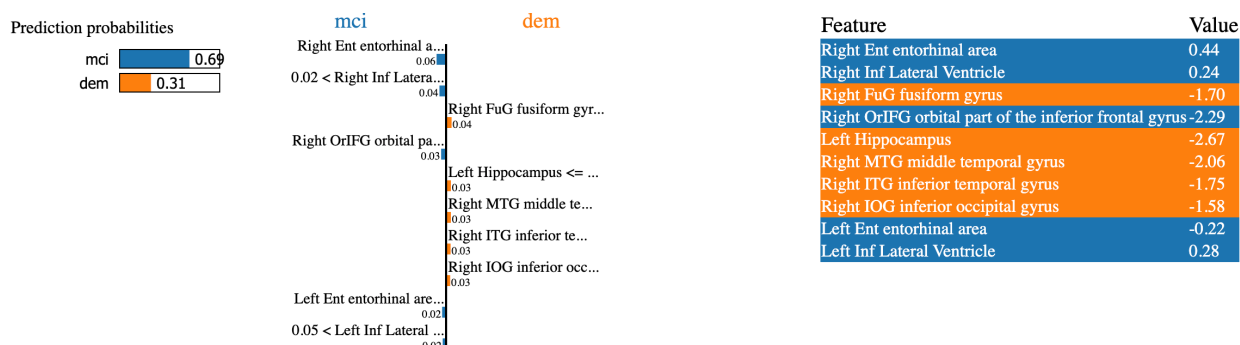


Εικόνα 20: Αποτελέσματα του μοντέλου LIME για την ταξινόμηση με ταξινομητή Μηχανών Διανυσμάτων Υποστήριξης για τις κλάσεις CN και MCI. Ορθώς θετική περίπτωση, όπου ένα MCI άτομο ταξινομείται σωστά. Αριστερά: η προβλεπόμενη πιθανότητα της συγκεκριμένης περίπτωσης για τις δύο κλάσεις, Μέση: τα χαρακτηριστικά που ωθούν την πρόβλεψη στην κλάση MCI (πορτοκαλί) και στην κλάση CN (μπλε) αντίστοιχα, Δεξιά: μια κατάταξη σημαντικότητας των χαρακτηριστικών και αντίστοιχες τιμές τους. Τα χαρακτηριστικά πορτοκαλί χρώματος ωθούν την πρόβλεψη στην κλάση MCI και τα χαρακτηριστικά μπλε χρώματος ωθούν την πρόβλεψη στην κλάση CN.

Κατά την ταξινόμηση των κλάσεων MCI και DEM, οι επεξηγήσεις που ελήφθησαν από το μοντέλο LIME, έδειξαν πως οι ογκομετρικές περιοχές που αυξάνουν την πιθανότητα ενός δείγματος να ανήκει στην κλάση DEM έναντι της MCI, ήταν μεταξύ άλλων η δεξιά μέση κροταφική έλικα, η δεξιά κατώτερη κροταφική έλικα, η δεξιά ατρακτοειδής έλικα και ο αριστερός ιππόκαμπος, ενώ αντίθετα οι τιμές της αριστερής και δεξιάς ενδορινικής περιοχής και της αριστερής και δεξιάς πλευρικής κοιλίας, ωθούν το συγκεκριμένο δείγμα προς την κλάση MCI έναντι της DEM. Οι επεξηγήσεις αυτές ευθυγραμμίζονται με τα αποτελέσματα της σημαντικότητας χαρακτηριστικών για το ίδιο ζεύγος κλάσεων. Στις Εικόνες 21 και 22 παρουσιάζονται οι επεξηγήσεις που προέκυψαν από το μοντέλο LIME για τον ίδιο ασθενή με τους δύο διαφορετικούς ταξινομητές, οι οποίοι πραγματοποίησαν εσφαλμένες προβλέψεις ταξινομώντας τον ασθενή στην κλάση MCI, αντί της DEM στην οποία ανήκει.



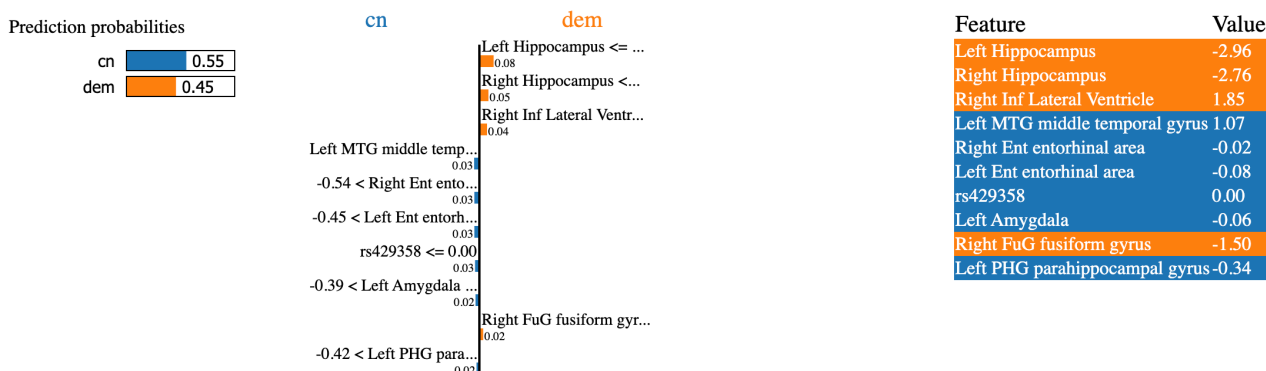
Εικόνα 21: Αποτελέσματα του μοντέλου LIME για την ταξινόμηση με ταξινομητή Τυχαίων Δασών για τις κλάσεις MCI και DEM. Εσφαλμένα αρνητική περίπτωση, όπου ένα άτομο με άνοια (DEM) ταξινομείται στην κλάση MCI. Αριστερά: η προβλεπόμενη πιθανότητα της συγκεκριμένης περίπτωσης για τις δύο κλάσεις, Μέση: τα χαρακτηριστικά που ωθούν την πρόβλεψη στην κλάση DEM (πορτοκαλί) και στην κλάση MCI (μπλε) αντίστοιχα, Δεξιά: μια κατάταξη σημαντικότητας των χαρακτηριστικών και αντίστοιχες τιμές τους. Τα χαρακτηριστικά πορτοκαλί χρώματος ωθούν την πρόβλεψη στην κλάση DEM και τα χαρακτηριστικά μπλε χρώματος ωθούν την πρόβλεψη στην κλάση MCI.



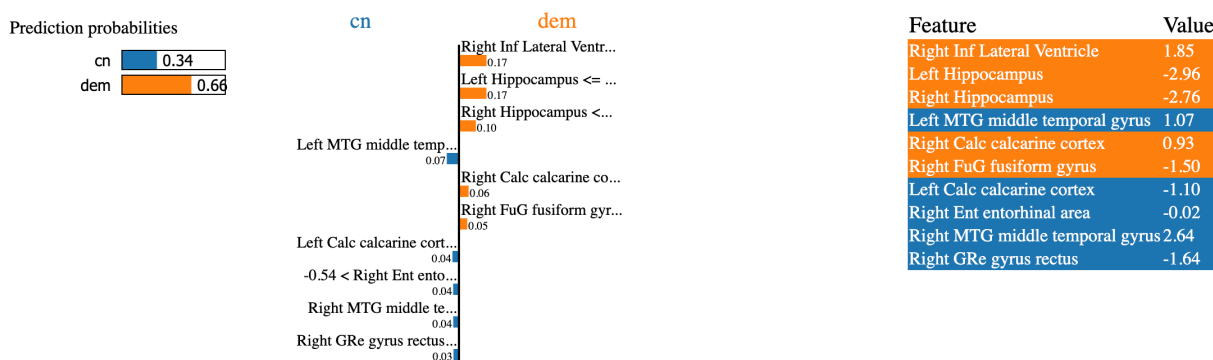
Εικόνα 22: Αποτελέσματα του μοντέλου LIME για την ταξινόμηση με ταξινομητή Μηχανών Διανυσμάτων Υποστήριξης για τις κλάσεις MCI και DEM. Εσφαλμένα αρνητική περίπτωση, όπου ένα άτομο με άνοια (DEM) ταξινομείται στην κλάση MCI. Αριστερά: η προβλεπόμενη πιθανότητα της συγκεκριμένης περίπτωσης για τις δύο κλάσεις, Μέση: τα χαρακτηριστικά που ωθούν την πρόβλεψη στην κλάση DEM (πορτοκαλί) και στην κλάση MCI (μπλε) αντίστοιχα, Δεξιά: μια κατάταξη σημαντικότητας των χαρακτηριστικών και αντίστοιχες τιμές τους. Τα χαρακτηριστικά πορτοκαλί χρώματος ωθούν την πρόβλεψη στην κλάση DEM και τα χαρακτηριστικά μπλε χρώματος ωθούν την πρόβλεψη στην κλάση MCI.

Αντίστοιχα, κατά την ταξινόμηση CN και DEM δειγμάτων, σύμφωνα με το μοντέλο LIME, η πιθανότητα ενός δείγματος να ταξινομηθεί ως ασθενής με άνοια, αυξάνεται με βάση τις τιμές όγκου του αριστερού και δεξιού ιπποκάμπου και της δεξιάς κάτω πλευρικής κοιλίας, ενώ η πιθανότητα να ταξινομηθεί ως υγιής αυξάνεται από τις τιμές της δεξιάς και αριστερής ενδορινικής περιοχής, της αριστερής και δεξιάς μέσης κροταφικής έλικας και της απουσίας άλλων (0) του SNP rs429358 που σχετίζεται με το γονίδιο ApoE. Στις

Εικόνας 23 και 24 παρουσιάζεται ένα δείγμα υγιούς ατόμου που ταξινομήθηκε σωστά στην κλάση CN από τον ταξινομητή Τυχαίων Δασών, αλλά εσφαλμένα από τις Μηχανές Διανυσμάτων Υποστήριξης, στην κλάση των ασθενών με άνοια.



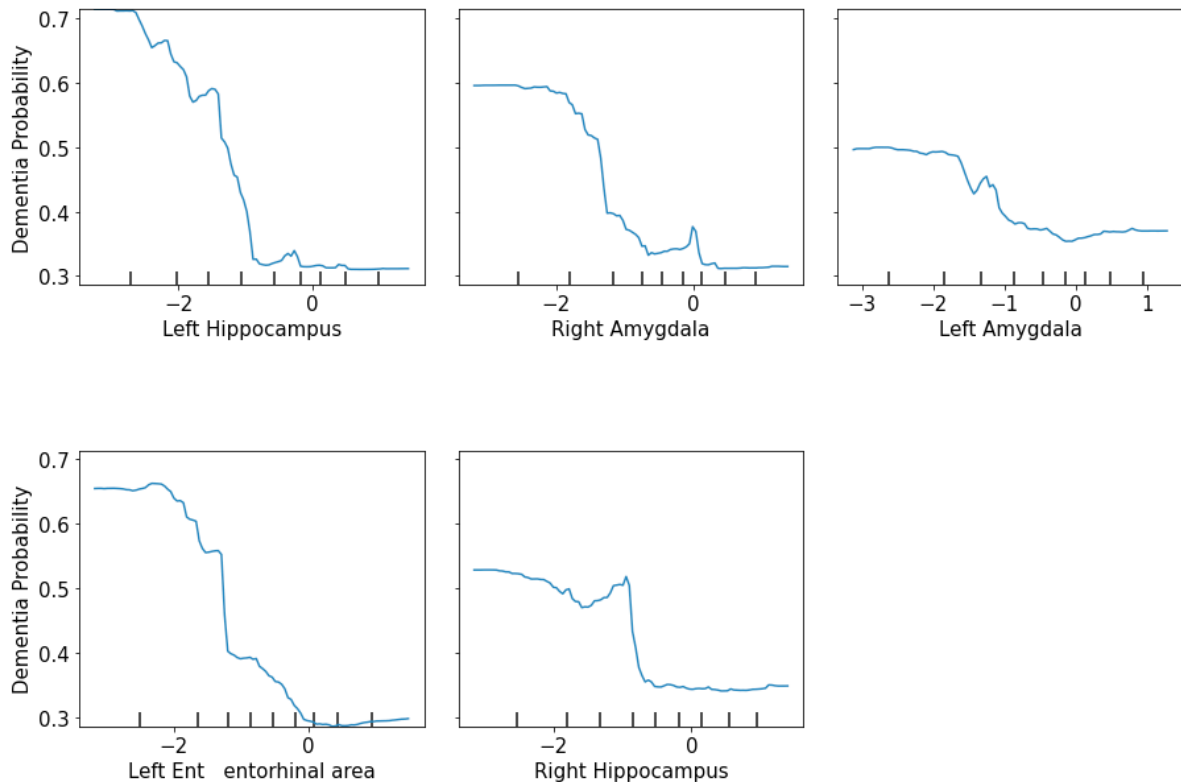
Εικόνα 23: Αποτελέσματα του μοντέλου LIME για την ταξινόμηση με ταξινομητή Τυχαίων Δασών για τις κλάσεις CN και DEM. Ορθώς αρνητική περίπτωση, όπου ένα υγιές άτομο ταξινομείται σωστά. Αριστερά: η προβλεπόμενη πιθανότητα της συγκεκριμένης περίπτωσης για τις δύο κλάσεις, Μέση: τα χαρακτηριστικά που ωθούν την πρόβλεψη στην κλάση DEM (πορτοκαλί) και στην κλάση CN (μπλε) αντίστοιχα, Δεξιά: μια κατάταξη σημαντικότητας των χαρακτηριστικών και αντίστοιχες τιμές τους. Τα χαρακτηριστικά πορτοκαλί χρώματος ωθούν την πρόβλεψη στην κλάση DEM και τα χαρακτηριστικά μπλε χρώματος ωθούν την πρόβλεψη στην κλάση CN.



Εικόνα 24: Αποτελέσματα του μοντέλου LIME για την ταξινόμηση με ταξινομητή Μηχανών Διανυσμάτων Υποστήριξης για τις κλάσεις CN και DEM. Εσφαλμένα θετική περίπτωση, όπου το ίδιο υγιές άτομο με την Εικόνα 23, ταξινομείται λανθασμένα στην κλάση DEM. Αριστερά: η προβλεπόμενη πιθανότητα της συγκεκριμένης περίπτωσης για τις δύο κλάσεις, Μέση: τα χαρακτηριστικά που ωθούν την πρόβλεψη στην κλάση DEM (πορτοκαλί) και στην κλάση CN (μπλε) αντίστοιχα, Δεξιά: μια κατάταξη σημαντικότητας των χαρακτηριστικών και αντίστοιχες τιμές τους. Τα χαρακτηριστικά πορτοκαλί χρώματος ωθούν την πρόβλεψη στην κλάση DEM και τα χαρακτηριστικά μπλε χρώματος ωθούν την πρόβλεψη στην κλάση CN.

5.2.4 Διάγραμμα Μερικής Εξάρτησης

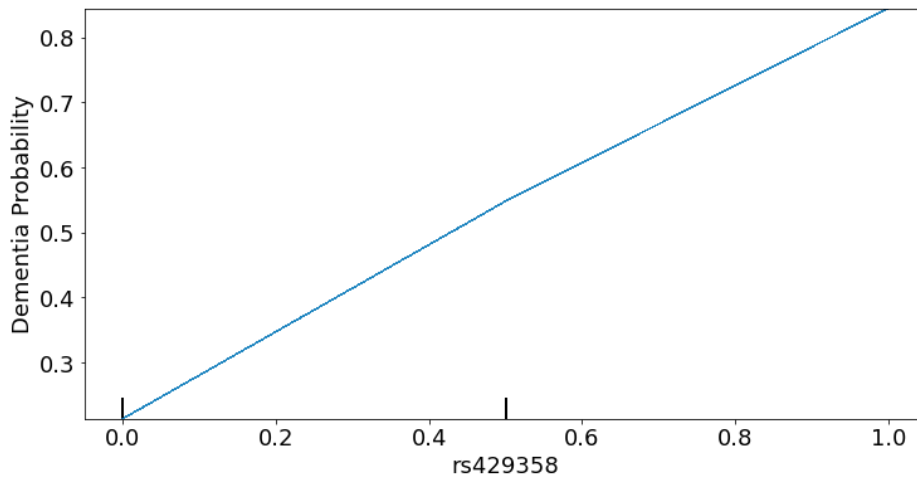
Το μοντέλο των Διαγραμμάτων Μερικής Εξάρτησης εφαρμόστηκε για κάθε δυαδικό υποπρόβλημα χρησιμοποιώντας τον ταξινομητή Τυχαίων Δασών, όπως περιγράφηκε στο κεφάλαιο 4.6.4.



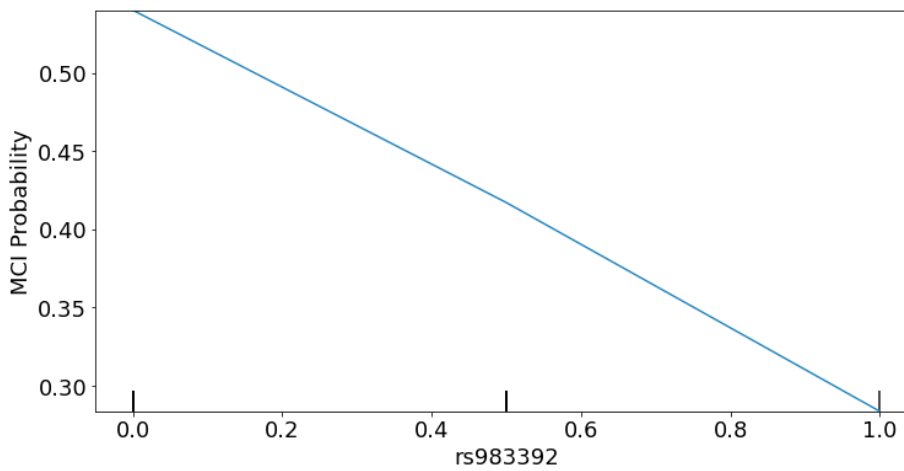
Εικόνα 25: Διαγράμματα Μερικής Εξάρτησης πέντε σημαντικών χαρακτηριστικών όπως προέκυψαν από την ταξινόμηση των κλάσεων CN και DEM με το μοντέλο Τυχαίων Δασών. Στον άξονα x αναπαρίστανται οι τιμές των χαρακτηριστικών και στον άξονα y η πιθανότητα ενός δείγματος να ανήκει στην κλάση DEM συναρτήσει της εκάστοτε τιμής του χαρακτηριστικού, επομένως τα διαγράμματα αυτά δείχνουν την επίδραση συγκεκριμένων χαρακτηριστικών στην τελική πρόβλεψη.

Αρχικά, ελήφθησαν διαγράμματα μερικής εξάρτησης για κάποια από τα σημαντικότερα χαρακτηριστικά, όπως τα διαγράμματα της Εικόνας 25, όπου παρουσιάζονται τα διαγράμματα της επίδρασης ορισμένων χαρακτηριστικών κατά την ταξινόμηση ανάμεσα στις κλάσεις CN και DEM. Επαληθεύτηκε για ακόμη μία φορά, πως όσο μικρότερος είναι ο όγκος περιοχών όπως ο ιππόκαμπος, η αμυγδαλή ή η ενδορινική περιοχή, τόσο υψηλότερη είναι η πιθανότητα του δείγματος να ταξινομηθεί στην κλάση των ασθενών με άνοια.

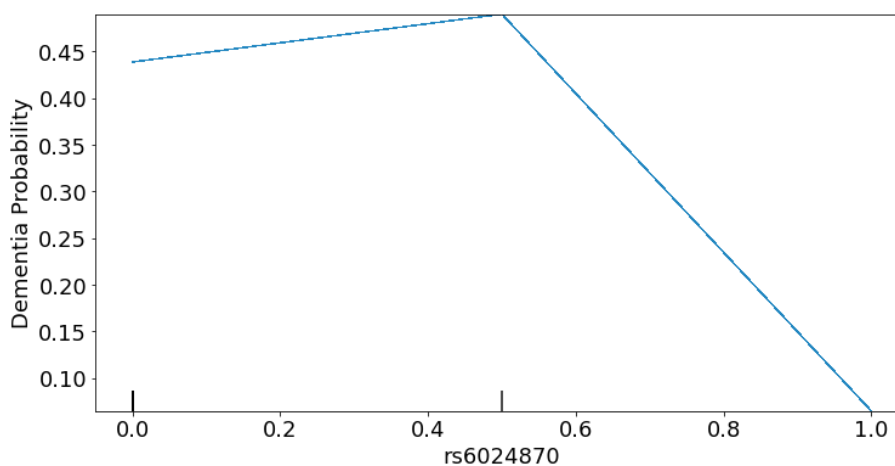
Στη συνέχεια, τα διαγράμματα μερικής εξάρτησης χρησιμοποιήθηκαν για τη μελέτη της επίδρασης ορισμένων SNP στην πρόβλεψη του μοντέλου. Ορισμένα SNP φάνηκαν να ενισχύουν την πιθανότητα ενός ατόμου να έχει διαγνωστεί με ήπιας μορφής νοητική εξασθένηση (MCI) ή άνοια, ενώ αντίθετα, η παρουσία άλλων SNP φάνηκε να ελαττώνει την πιθανότητα αυτή. Πιο συγκεκριμένα, η παρουσία ενός ή δύο άλληλων (τιμή 0.5 και 1 αντίστοιχα στον άξονα x) του rs429358 που σχετίζεται με το γονίδιο APOE αυξάνει γραμμικά την πιθανότητα ενός ασθενούς να ταξινομηθεί στην κλάση των ατόμων με άνοια (Εικόνα 26). Αντίθετα, η παρουσία (με ένα ή δύο άλληλα) του rs983392 που ανήκει στο γονίδιο MS4A6A [55] και του rs6024870 που σχετίζεται με το γονίδιο CASS4 [56] φάνηκε να ελαττώνει την πιθανότητα του ατόμου να ανήκει στην κλάση MCI ή DEM στα τρία δυαδικά υποπροβλήματα (Εικόνες 26-28).



Εικόνα 26: Διάγραμμα Μερικής Εξάρτησης που απεικονίζει την επίδραση του rs429358 SNP, το οποίο ανήκει στο γονίδιο ApoE [54] για την πιθανότητα ταξινόμησης στην κλάση των ασθενών με άνοια, κατά το δυαδικό υποπρόβλημα του ζεύγους κλάσεων CN και DEM.



Εικόνα 27: Διάγραμμα Μερικής Εξάρτησης που απεικονίζει την επίδραση του rs983392 SNP, το οποίο ανήκει στο γονίδιο MS4A6A [55] για την πιθανότητα ταξινόμησης στην κλάση των ασθενών με MCI, κατά το δυαδικό υποπρόβλημα του ζεύγους κλάσεων CN και MCI.



Εικόνα 28: Διάγραμμα Μερικής Εξάρτησης που απεικονίζει την επίδραση του rs6024870 SNP, το οποίο ανήκει στο γονίδιο CASS4 [56] για την πιθανότητα ταξινόμησης στην κλάση των ασθενών με άνοια, κατά το δυαδικό υποπρόβλημα του ζεύγους κλάσεων MCI και DEM

5.2.5 Αντιπαραθετικές Επεξηγήσεις Permute Attack

Η μέθοδος αντιπαραθετικών επεξηγήσεων Permute Attack εφαρμόστηκε ξεχωριστά στα τρία δυαδικά προβλήματα χρησιμοποιώντας τον ταξινομητή SVM, όπως περιγράφηκε στο κεφάλαιο 4.6.5. Αρχικά, εξήχθησαν αντιπαραθετικά παραδείγματα για ορισμένα στιγμιότυπα, στα οποία φαίνονται οι αλλαγές στις τιμές των χαρακτηριστικών που είναι ικανές να ανατρέψουν το αποτέλεσμα της πρόβλεψης (Εικόνες 29 έως 31).

Στην Εικόνα 29 φαίνεται πως ένα υγιές άτομο, με πιθανότητα 24.6% να ταξινομηθεί στην κλάση MCI, θα μπορούσε να θεωρηθεί MCI αν η τιμή ορισμένων από τα ακόλουθα χαρακτηριστικά μεταβαλλόταν, ώστε να αυξήσει τον κίνδυνο για ανάπτυξη MCI: δεξιά αμυγδαλή, αριστερός ιππόκαμπος, δεξιά και αριστερή κάτω πλευρική κοιλία, δεξιά ενδορινική περιοχή, αριστερή κατώτερη ινιακή έλικα, αριστερή μέση κροταφική έλικα και άλλα.

	Right Amygdala	Left Hippocampus	Right Inf Lateral Ventricle	Left Inf Lateral Ventricle	Right Ent entorhinal area	Left IOG inferior occipital gyrus	Left MTG middle temporal gyrus	Right OrIFG orbital part of the inferior frontal gyrus	Right PCGg posterior cingulate gyrus	CN Probability	MCI Probability
Original Value	-0.061461	0.080467	0.008348	1.214025	0.127498	0.721089	-0.611474	-0.066917	1.477293	0.754064	0.245936
1st Counterfactual Value	-2.552078	0.080467	0.008348	1.214025	-3.401019	-1.694175	-2.316634	2.575087	1.477293	0.490281	0.509719
2nd Counterfactual Value	-1.995737	-2.586924	3.278754	1.214025	0.127498	-1.337064	-0.611474	-0.066917	-2.299880	0.500000	0.500000
3rd Counterfactual Value	-0.061461	0.080467	3.624476	6.010917	-3.401019	0.721089	-0.611474	-0.066917	1.477293	0.480510	0.519490

Εικόνα 29: Αντιπαραθετικά παραδείγματα Permute Attack για ανατροπή της πρόβλεψης ενός CN δείγματος και την ταξινόμηση του στην κλάση MCI. Στην πρώτη γραμμή παρουσιάζονται οι αρχικές τιμές των χαρακτηριστικών και οι αντίστοιχες πιθανότητες πρόβλεψης για κάθε κλάση, ενώ στις υπόλοιπες τρεις γραμμές οι τιμές που έχουν επισημανθεί με πορτοκαλί υποδεικνύουν τις αλλαγές στις τιμές των χαρακτηριστικών και τις νέες πιθανότητες για κάθε κλάση.

Η Εικόνα 30, παρουσιάζει ένα αντιπαραθετικό παράδειγμα, ενός ατόμου της MCI κλάσης με σχετικά χαμηλή πιθανότητα να ανήκει στην κλάση των ασθενών με άνοια. Όπως παρατηρεί κανείς, μία ισχυρή μεταβολή στις τιμές των χαρακτηριστικών είναι απαραίτητη για την ανατροπή της πρόβλεψης, και την ταξινόμηση του ατόμου στην κλάση DEM. Καθένα από τα τρία αντιπαραθετικά παραδείγματα μεταβάλλει τέσσερα ή πέντε από τα ακόλουθα χαρακτηριστικά: αριστερή και δεξιά κάτω πλευρική κοιλία, δεξιά κατώτερη κροταφική έλικα, αριστερή μέση κροταφική έλικα και αριστερή ανώτερη κροταφική έλικα (left superior temporal gyrus).

	Right Inf Lateral Ventricle	Left Inf Lateral Ventricle	Right ITG inferior temporal gyrus	Left MTG middle temporal gyrus	Left STG superior temporal gyrus	rs7810606	MCI Probability	Dementia Probability
Original Value	-0.846350	-0.944551	0.090685	-0.032223	1.852486	0.500000	0.769456	0.230544
1st Counterfactual Value	0.256800	5.636389	0.090685	-4.239847	-0.904284	0.500000	0.490463	0.509537
2nd Counterfactual Value	1.632706	4.947804	-3.705698	-1.738529	-2.355812	0.500000	0.441131	0.558869
3rd Counterfactual Value	3.885828	4.342009	0.090685	-0.032223	-1.684376	1.000000	0.488352	0.511648

Εικόνα 30: Αντιπαραθετικά παραδείγματα Permute Attack για ανατροπή της πρόβλεψης ενός MCI δείγματος και την ταξινόμηση του στην κλάση DEM. Στην πρώτη γραμμή παρουσιάζονται οι αρχικές τιμές των χαρακτηριστικών και οι αντίστοιχες πιθανότητες πρόβλεψης για κάθε κλάση, ενώ στις υπόλοιπες τρεις γραμμές οι τιμές που έχουν επισημανθεί με πορτοκαλί υποδεικνύουν τις αλλαγές στις τιμές των χαρακτηριστικών και τις νέες πιθανότητες για κάθε κλάση.

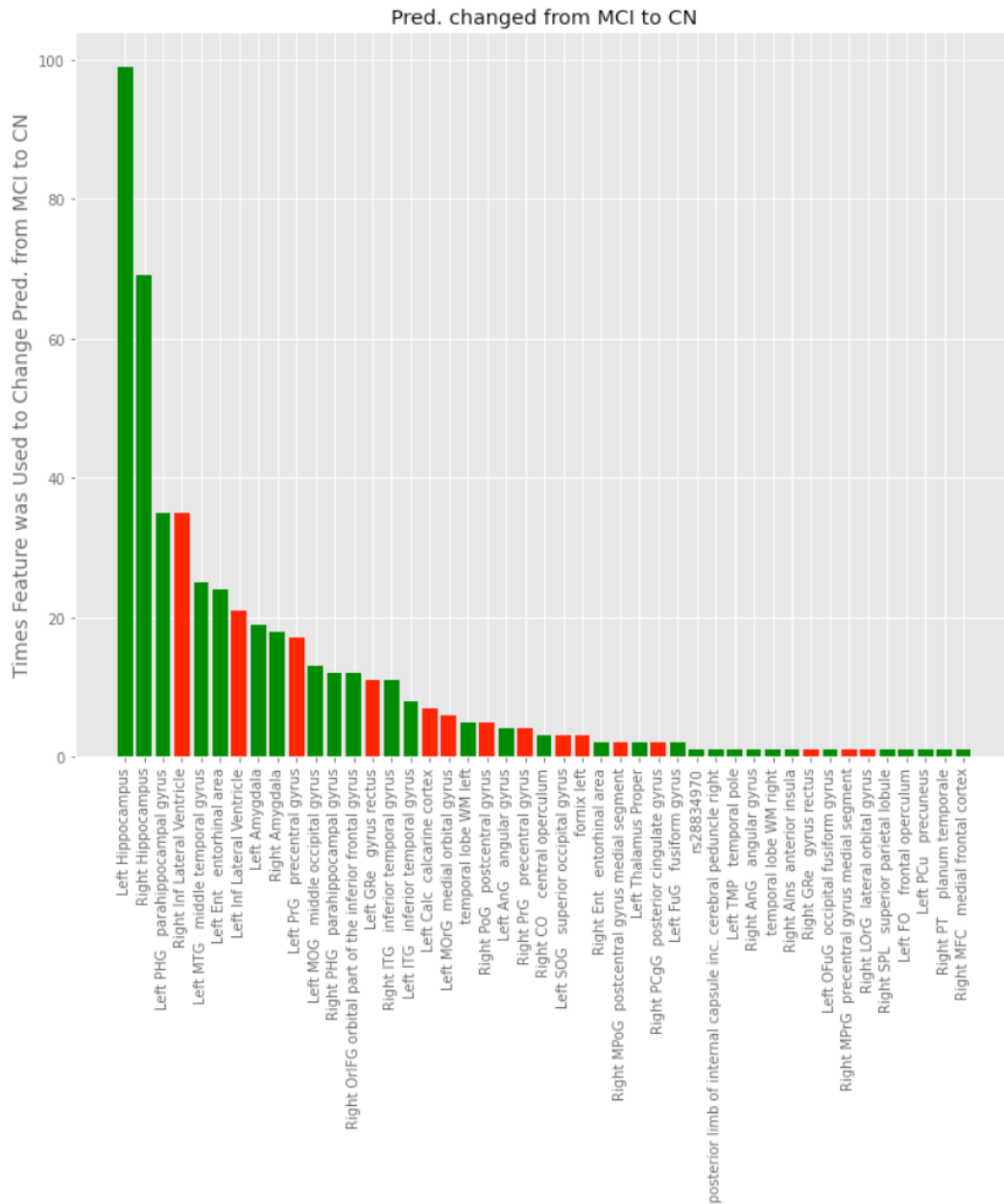
Στην Εικόνα 31, διακρίνεται ένα άτομο με άνοια, το οποίο προκειμένου να ταξινομηθεί στην κλάση των υγιών ατόμων, η μόνη αλλαγή που απαιτείται είναι η αύξηση στην τιμή του αριστερού ιπποκάμπου, γεγονός που σημαίνει πως το συγκεκριμένο χαρακτηριστικό διαδραματίζει εξαιρετικά σημαντικό ρόλο στην αναγνώριση ενός ασθενούς με άνοια και τη διάκρισή του από τα υγιή άτομα.

	Left Hippocampus	CN Probability	Dementia Probability
Original Value	-1.606740	0.353712	0.646288
1st Counterfactual Value	0.391685	0.510204	0.489796
2nd Counterfactual Value	1.629365	0.604601	0.395399
3rd Counterfactual Value	1.702757	0.609918	0.390082

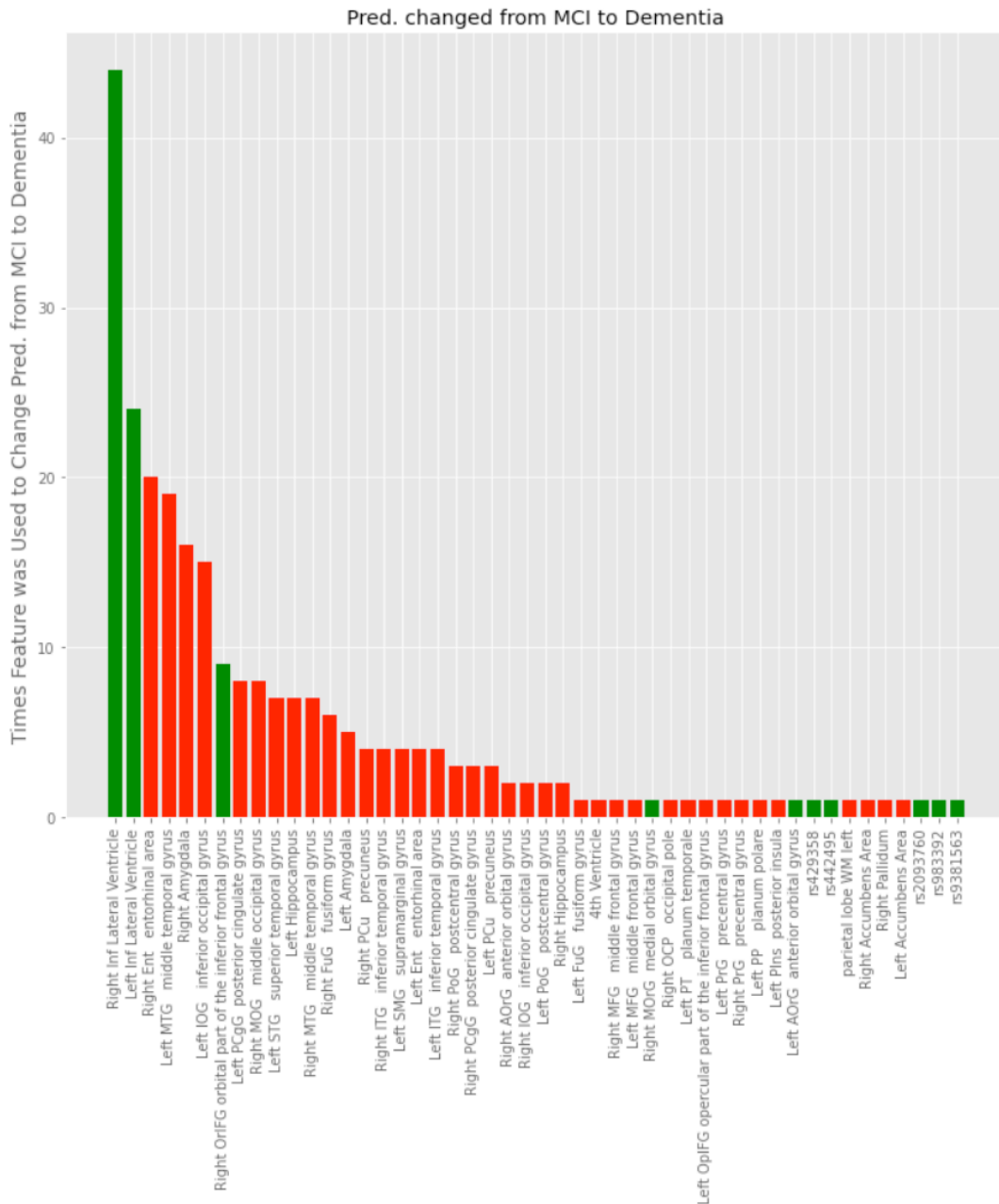
Εικόνα 31: Αντιπαραθετικά παραδείγματα Permute Attack για ανατροπή της πρόβλεψης ενός DEM δείγματος και την ταξινόμηση του στην κλάση CN. Στην πρώτη γραμμή παρουσιάζονται οι αρχικές τιμές των χαρακτηριστικών και οι αντίστοιχες πιθανότητες πρόβλεψης για κάθε κλάση, ενώ στις υπόλοιπες τρεις γραμμές οι τιμές που έχουν επισημανθεί με πορτοκαλί υποδεικνύουν τις αλλαγές στις τιμές των χαρακτηριστικών και τις νέες πιθανότητες για κάθε κλάση.

Στη συνέχεια, για κάθε ζεύγος κλάσεων, εξήχθη από ένα αντιπαραθετικό παράδειγμα για κάθε στιγμιότυπο του συνόλου ελέγχου (test set) και υπολογίστηκε το πλήθος των φορών που κάθε χαρακτηριστικό μεταβλήθηκε ώστε να συνεισφέρει στην ανατροπή της πρόβλεψης. Οι υπολογισμοί αυτοί παρουσιάζονται στα ραβδογράμματα των Εικόνων 32 και 33, όπου τα χαρακτηριστικά έχουν ταξινομηθεί σε φθίνουσα σειρά της συχνότητας εμφάνισής τους στα αντιπαραθετικά παραδείγματα. Μία πράσινη ράβδος αντιπροσωπεύει μία συνολικά θετική μεταβολή του εκάστοτε χαρακτηριστικού, ενώ μία κόκκινη, συνολικά αρνητική μεταβολή.

Καθίσταται σαφές, πως τα επικρατέστερα χαρακτηριστικά είναι αυτά τα οποία έχουν λάβει από τα προηγούμενα μοντέλα ερμηνευσιμότητας υψηλές θέσεις στην κατάταξης σημαντικότητας χαρακτηριστικών. Η αύξηση στον όγκο του ιπποκάμπου, της αριστερής παράϊπποκαμπικής έλικας, της ενδορινικής περιοχής ή της αμυγδαλής των ασθενών και η ελάττωση του όγκου των κάτω πλευρικών κοιλιών φαίνεται να συνεισφέρουν στην ανατροπή της πρόβλεψης MCI και να οδηγούν στην ταξινόμηση ενός ατόμου στην κλάση CN στην Εικόνα 32 για το συγκεκριμένο ζεύγος κλάσεων. Αντίστοιχα, η ελάττωση του όγκου του ιπποκάμπου, της ενδορινικής περιοχής ή της αριστερής μέσης κροταφικής έλικας, μεταξύ άλλων, και η αύξηση του όγκου των κάτω πλευρικών κοιλιών θα οδηγούσαν έναν MCI ασθενή να ταξινομηθεί στην κλάση των ατόμων με άνοια, σύμφωνα με την Εικόνα 33 για το συγκεκριμένο δυαδικό υποπρόβλημα. Μπορεί, να προκύψει το συμπέρασμα, ότι όσο μεγαλύτερη είναι η συχνότητα εμφάνισης ενός χαρακτηριστικού στα αντιπαραθετικά παραδείγματα, τόσο υψηλότερη είναι η σημαντικότητα του για τον διαχωρισμό των αντίστοιχων κλάσεων.



Εικόνα 32: Ραβδόγραμμα της συχνότητας με την οποία κάθε χαρακτηριστικό επιλέχθηκε να μεταβληθεί στα αντιπαραθετικά παραδείγματα, τα οποία δημιουργήθηκαν για το σύνολο ελέγχου της ταξινόμησης των κλάσεων CN και MCI. Τα χαρακτηριστικά έχουν τοποθετηθεί σε φθίνουσα σειρά συχνότητα. Μία πράσινη ράβδος αντιπροσωπεύει μία συνολικά θετική μεταβολή του εκάστοτε χαρακτηριστικού, ενώ μία κόκκινη, συνολικά αρνητική μεταβολή.



Εικόνα 33: Ραβδόγραμμα της συχνότητας με την οποία κάθε χαρακτηριστικό επιλέχθηκε να μεταβληθεί στα αντιπαραθετικά παραδείγματα, τα οποία δημιουργήθηκαν για το σύνολο ελέγχου της ταξινόμησης των κλάσεων MCI και DEM. Τα χαρακτηριστικά έχουν τοποθετηθεί σε φθίνουσα σειρά συχνότητα. Μία πράσινη ράβδος αντιπροσωπεύει μία συνολικά θετική μεταβολή του εκάστοτε χαρακτηριστικού, ενώ μία κόκκινη, συνολικά αρνητική μεταβολή.

6. Συμπεράσματα και Συζήτηση

6.1 Σύνοψη

Η συγκεκριμένη μελέτη αφορά στην ταξινόμηση υγιών ατόμων και ασθενών με Ήπια Νοητική Εξασθένιση (Mild Cognitive Impairment - MCI) ή άνοια με τη χρήση μεθόδων τεχνητής νοημοσύνης. Χρησιμοποιήθηκαν και συγκρίθηκαν έξι διαφορετικοί ταξινομητές (Τυχαία Δάση, Λογιστική Παλινδρόμηση, Μηχανές Διανυσμάτων Υποστήριξης, Πολυστρωματικό Perceptron, Gradient Boosting και XGBoost) κατά τη δοκιμή των μεθόδων που αναπτύχθηκαν. Αρχικά δοκιμάστηκαν οι τεχνικές διαίρεσης προβλημάτων με πολλαπλές κλάσεις Ένα-προς-Όλα και Ένα-προς-Ένα, με καταλληλότερη την Ένα-προς-Ένα, η οποία παρείχε σημαντικά υψηλότερη επίδοση. Στη συνέχεια, αντιμετωπίστηκε επιτυχώς το υπάρχον πρόβλημα ανισορροπίας των κλάσεων και βελτιώθηκαν περαιτέρω τα αποτελέσματα της ταξινόμησης, εφαρμόζοντας τη μέθοδο συλλογικής μάθησης Bagging με το σχήμα Ένα-προς-Ένα. Για κάθε ταξινομητή που χρησιμοποιήθηκε στη μέθοδο αυτή, εντοπίστηκαν οι βέλτιστες υπερπαραμέτροι, εφαρμόστηκε το σχήμα αξιολόγησης εμφωλευμένης διασταυρούμενης επικύρωσης και προέκυψε το βέλτιστο μοντέλο με βάση τη μέση τιμή ισορροπημένης ακρίβειας για τις 5 πτυχές της διασταυρούμενης επικύρωσης, όπως περιγράφηκε παραπάνω. Τη βέλτιστη προβλεπτική ικανότητα παρουσίασε το μοντέλο με τη χρήση του ταξινομητή των Μηχανών Διανυσμάτων Υποστήριξης, ενώ αντίστοιχα υψηλά αποτελέσματα παρείχαν τα Τυχαία Δάση και ο ταξινομητής XGBoost. Τέλος, ελέγχθηκε εάν υπάρχει στατιστικά σημαντική διαφορά ανάμεσα στη μέση τιμή της ισορροπημένης ακρίβειας για τους έξι ταξινομητές, εφαρμόζοντας ένα στατιστικό t-τεστ για κάθε ζεύγος ταξινομητών. Φάνηκε πως με εξαίρεση το Πολυστρωματικό Perceptron, το οποίο είχε συνολικά τη χαμηλότερη επίδοση, δεν υπήρχε στατιστικά σημαντική διαφορά στην προβλεπτική ικανότητα ανάμεσα στους υπόλοιπους ταξινομητές.

Σημαντικό κομμάτι της μελέτης αποτέλεσαν οι μέθοδοι ερμηνευσιμότητας που εφαρμόστηκαν για την κατανόηση της επίδρασης των χαρακτηριστικών στην πρόβλεψη που πραγματοποίησαν τα μοντέλα. Για τους ταξινομητές Τυχαία Δάση, Gradient Boosting και XGBoost αποκτήθηκε η κατάταξη σημαντικότητας χαρακτηριστικών που βασίζεται στο κριτήριο Gini, ενώ για τον ταξινομητή Μηχανές Διανυσμάτων Υποστήριξης προέκυψε η αντίστοιχη κατάταξη με την εφαρμογή της μεθόδου Πρόσθετων Shapley Επεξηγήσεων (Shapley Additive ExPlanations - SHAP). Τα αποτελέσματα της κάθε κατάταξης συγκλίνουν, με κάποιες διαφοροποιήσεις, ως προς τα σημαντικότερα χαρακτηριστικά, τα οποία φάνηκε να είναι, μεταξύ άλλων, ο ιππόκαμπος, η αμυγδαλή, η ενδορινική περιοχή, οι κάτω πλευρικές κοιλίες, η μέση κροταφική έλικα, καθώς και πολυμορφισμοί του ενός νουκλεοτιδίου, όπως το rs429358 που ανήκει στο γονίδιο ApoE [54], το rs6448453 που σχετίζεται με το γονίδιο CLNK [9] και το rs2081545, το οποίο ανήκει στο γονίδιο MS4A6A [9].

Από τη μέθοδο SHAP εξήχθησαν, επιπλέον, γραφικές παραστάσεις σύνοψης για τον ταξινομητή Μηχανών Διανυσμάτων Υποστήριξης, στις οποίες εκτός από τη σημαντικότητα των χαρακτηριστικών, φαίνεται και η επίδραση της τιμής του κάθε χαρακτηριστικού στην πρόβλεψη του μοντέλου. Παρατηρήθηκε πως η ατροφία ορισμένων ανατομικών περιοχών του εγκεφάλου, όπως για παράδειγμα, ο ιππόκαμπος, η αμυγδαλή και η ενδορινική περιοχή και η διεύρυνση περιοχών όπως οι κάτω πλευρικές κοιλίες ωθούν το μοντέλο προς την πρόβλεψη πως το άτομο πάσχει από MCI ή άνοια και το αντίθετο.

Στη συνέχεια, χρησιμοποιήθηκε το μοντέλο LIME, ώστε εξαχθούν, σε τοπικό επίπεδο, συμπεράσματα για τον τρόπο με τον οποίο τα επιμέρους χαρακτηριστικά επηρεάζουν την πρόβλεψη των μοντέλων. Για τα Τυχαία Δάση και τις Μηχανές Διανυσμάτων Υποστήριξης, αποκτήθηκαν τοπικές επεξηγήσεις για μεμονωμένα στιγμιότυπα, από τις οποίες προέκυψαν τα σημαντικότερα χαρακτηριστικά για το εκάστοτε δείγμα και φάνηκε ποια από αυτά αυξάνουν την πιθανότητα του στιγμιότυπου να ταξινομηθεί στην κλάση των υγιών ατόμων, των ασθενών με MCI ή άνοια αντίστοιχα. Επιβεβαιώθηκαν οι παρατηρήσεις της

μεθόδου SHAP σχετικά με τις χαμηλές ή υψηλές τιμές των ογκομετρικών περιοχών ενδιαφέροντος που προαναφέρθηκαν και παρατηρήθηκε επιπλέον πως η απουσία του rs429358 του γονιδίου ApoE ωθεί την πρόβλεψη του δείγματος στην κλάση των υγιών ατόμων, ενώ η παρουσία του με ένα ή δύο άλλα στις κλάσεις του MCI και της άνοιας.

Προκειμένου να μελετηθεί περαιτέρω η επίδραση των συγκεκριμένων χαρακτηριστικών, επιστρατεύθηκε η μέθοδος των Διαγραμμάτων Μερικής Εξάρτησης, από την οποία προέκυψαν οι γραφικές παραστάσεις που δείχνουν τη συσχέτιση της τιμής του κάθε χαρακτηριστικού μεμονωμένα στην πιθανότητα να ταξινομηθεί ένα δείγμα στην εκάστοτε κλάση και επιβεβαιώνουν τις παραπάνω παρατηρήσεις. Η μέθοδος αυτή επιβεβαίωσε τις παρατηρήσεις των προηγούμενων μεθόδων και παρείχε επιπλέον στοιχεία για τα SNP χαρακτηριστικά, καθώς ανέδειξε πολυμορφισμούς ενός νουκλεοτιδίου, των οποίων η παρουσία ελαττώνει την πιθανότητα ένα δείγμα να ταξινομηθεί στις κλάσεις των ασθενών, όπως το rs983392 που ανήκει στο γονίδιο MS4A6A [55] και το rs6024870 που σχετίζεται με το CASS4.

Τέλος με τη μέθοδο αντιπαραθετικών παραδειγμάτων Permute Attack αποκτήθηκαν δείγματα με παραλλαγμένες τιμές χαρακτηριστικών, ικανές για να αλλάξουν την πρόβλεψη για το εκάστοτε στιγμιότυπο. Δημιουργήθηκαν διαγράμματα, τα οποία απεικονίζουν πόσες φορές συνολικά για όλα τα δείγματα μεταβλήθηκε το κάθε χαρακτηριστικό, καθώς και αν η μεταβολή αυτή ήταν θετική ή αρνητική. Επομένως, αποδείχθηκε πως τα χαρακτηριστικά αυτά που έχουν αναδειχθεί ως τα σημαντικότερα από τα προηγούμενα μοντέλα, είναι πράγματι καθοριστικά για την ταξινόμηση των ατόμων, καθώς μία μεταβολή στις τιμές τους είναι ικανή να ανατρέψει την πρόβλεψη. Με τη μέθοδο αυτή, εντοπίστηκαν επίσης άτομα, τα οποία βρίσκονται σε υψηλό κίνδυνο, είτε εμφάνισης MCI ή άνοιας, είτε εξέλιξης του MCI από το οποίο πάσχουν σε άνοια, καθώς με αλλαγές στις τιμές ορισμένων ανατομικών περιοχών του εγκεφάλου τους, μπορεί το μοντέλο να τα κατατάξει στην κλάση των ασθενών.

6.2 Συμπεράσματα

Από την παραπάνω μελέτη, μπορούν να εξαχθούν σημαντικά συμπεράσματα σχετικά με την ατροφία του εγκεφάλου και την ανάπτυξη νευροεκφυλιστικών ασθενειών σχετιζόμενων με τη γήρανση. Όλες οι παραπάνω παρατηρήσεις που αφορούν στις μεταβολές του όγκου ορισμένων ανατομικών περιοχών του εγκεφάλου, καθώς και στην παρουσία ή απουσία συγκεκριμένων πολυμορφισμών του ενός νουκλεοτιδίου, συμπίπτουν με τη σχετική ιατρική γνώση, γεγονός που επικυρώνει την αξιοπιστία των προτεινόμενων μοντέλων.

Η ατροφία ορισμένων ανατομικών περιοχών του εγκεφάλου, όπως ο ιππόκαμπος, η αμυγδαλή και οι ενδορινικές περιοχές συνδέονται ισχυρά με την νευροπαθολογία νευροεκφυλιστικών ασθενειών, όπως η Ήπια Νοητική Εξασθένηση και η άνοια. Επιπλέον, η παρατήρηση πως η διεύρυνση των κάτω πλευρικών κοιλιών αυξάνει την πιθανότητα του ατόμου να ανήκει στην κλάση των ασθενών, συμβαδίζει με το γεγονός ότι η συσσωρευτική απώλεια των νευρώνων και των συνάψεων οδηγεί σε διεύρυνση των κοιλιών, η οποία είναι ισχυρά συνδεδεμένη με την εξασθένηση των νοητικών ικανοτήτων και αποτελεί δείκτη της εξέλιξης νευροεκφυλιστικών νόσων [57].

Παρόλο που τα περισσότερα χαρακτηριστικά των πολυμορφισμών ενός νουκλεοτιδίου δεν κατέκτησαν υψηλές θέσεις στην κατάταξη σημαντικότητας χαρακτηριστικών των διάφορων μοντέλων, εξήχθησαν ενδιαφέροντα συμπεράσματα από τα διαγράμματα που αφορούν στην επίδραση της απουσίας ή παρουσίας ορισμένων από αυτά στην πιθανότητα των δειγμάτων να ταξινομηθούν σε κάποια από τις κλάσεις των ασθενών με MCI ή άνοια. Το SNP που αναγνωρίστηκε από όλα σχεδόν τα μοντέλα ερμηνευσιμότητας πως αυξάνει την πιθανότητα του ατόμου να ταξινομηθεί στην κλάση του MCI ή της άνοιας ήταν το rs429358 του γονιδίου ApoE, το οποίο έχει αποδειχθεί πως σχετίζεται με αυξημένο κίνδυνο ανάπτυξης άνοιας [58]. Επιπλέον, το SNP rs6448453 που σχετίζεται με το γονίδιο CLNK [59] και το οποίο αναδείχθηκε από τη σημαντικότητα χαρακτηριστικών του ταξινομητή

XGBoost, έχει επίσης συσχετιστεί με τον κίνδυνο ανάπτυξης νευροεκφυλιστικών ασθενειών [9]. Ωστόσο, η παρουσία ενός ή δύο άλλων των πολυμορφισμών ενός νουκλεοτιδίου που σχετίζονται με τα γονίδια CASS4 και MS4A6A, η οποία φάνηκε από το Διάγραμμα Μερικής Εξάρτησης να οδηγεί σε μειωμένη πιθανότητα για ταξινόμηση στις κλάσεις των δύο ασθενειών, έχει αποδειχθεί πως σχετίζεται με χαμηλότερο κίνδυνο ανάπτυξης νευροεκφυλιστικών νόσων και διαδραματίζει σημαντικό ρόλο ενάντια στην ατροφία του εγκεφάλου [56], [60].

6.3 Περιορισμοί

Ένας από τους περιορισμούς της συγκεκριμένης μελέτης, αφορά στον εντοπισμό των βέλτιστων υπερπαραμέτρων για τους επιμέρους ταξινομητές. Επειδή η προτεινόμενη μέθοδος συλλογικής μάθησης με σχήμα Ένα-προς-Ένα χρησιμοποιεί έξι πανομοιότυπους ταξινομητές, η εκπαίδευση του μοντέλου ήταν σχετικά πιο χρονοβόρα από ένα απλό μοντέλο ταξινόμησης. Επομένως, η αναζήτησή πλέγματος που πραγματοποιήθηκε κατά την εμφωλευμένη διασταυρούμενη επικύρωση είχε μεγάλο χρονικό κόστος, γεγονός που περιόρισε το πλήθος των συνδυασμών υπερπαραμέτρων που θα μπορούσαν να είχαν δοκιμαστεί για την περαιτέρω βελτίωση της επίδοσης του μοντέλου με καθέναν από τους επιμέρους ταξινομητές. Ωστόσο, η προβλεπτική ικανότητα του μοντέλου με τους βέλτιστους συνδυασμούς που εντοπίστηκαν για κάθε ταξινομητή ήταν υψηλή.

Ένας ακόμη περιορισμός, αφορά στη χρήση των μεθόδων ερμηνευσιμότητας στο μοντέλο συλλογικής μάθησης. Επειδή το σύνολο εκπαίδευσης χωρίζεται σε δύο μέρη κατά τη μέθοδο bagging, οι τεχνικές ερμηνευσιμότητας εφαρμόζονται κάθε φορά έχοντας εκπαιδεύσει τον εκάστοτε ταξινομητή σε ένα μόνο από τα δύο σύνολα εκπαίδευσης. Επιπλέον, επειδή η μελέτη αφορά σε πρόβλημα πολλαπλών κλάσεων, στο τελικό μοντέλο συλλογικής μάθησης όπου χρησιμοποιήθηκε η τεχνική διαίρεσης Ένα-προς-Ένα, οι τρεις κλάσεις μελετήθηκαν ανά δύο. Συνεπώς, τα συμπεράσματα που εξήχθησαν από τις μεθόδους ερμηνευσιμότητας αφορούσαν σε κάθε ζεύγος κλάσεων ξεχωριστά και όχι στη σημαντικότητα και επίδραση των χαρακτηριστικών συνολικά στην πρόβλεψη για τις τρεις κλάσεις. Παρά τους περιορισμούς αυτούς, τα αποτελέσματα των μεθόδων ερμηνευσιμότητας, όπως προαναφέρθηκε, συμπίπτουν με την υπάρχουσα ιατρική γνώση, γεγονός που τα καθιστά αξιόπιστα.

6.4 Μελλοντικές Επεκτάσεις

Μελλοντικές επεκτάσεις της συγκεκριμένης έρευνας θα μπορούσαν να αφορούν τον εντοπισμό της ετερογένειας των υπό μελέτη ασθενειών, με την αναγνώριση των υποτύπων του MCI (αμνησιακό, μη αμνησιακό, μεμονωμένου τομέα, πολλαπλών τομέων) και της άνοιας (της Νόσου Αλτσχάιμερ, με σωματία Lewy, εκφύλιση του μετωποκροταφικού λοβού), καθώς και άλλων διαφοροποιήσεων που υφίστανται για τις συγκεκριμένες νόσους. Ενδιαφέρουσα θα ήταν και η προσέγγιση των συγκεκριμένων ασθενειών με τη χρήση μη εποπτευόμενης μάθησης, για τον εντοπισμό νέων κατηγοριοποιήσεων της φυσιολογικής γήρανσης, του MCI και της άνοιας. Η μελέτη της ετερογένειας των ασθενειών αυτών θα βοηθούσε στην περαιτέρω κατανόηση της αιτιολογίας τους και στην ανάπτυξη εξατομικευμένων και συνεπώς αποτελεσματικότερων θεραπευτικών σχημάτων για τους ασθενείς.

Σημαντική θα ήταν, επίσης, η προσπάθεια περαιτέρω βελτίωσης της προβλεπτικής ικανότητας των προτεινόμενων μοντέλων. Ορισμένα βήματα προς αυτή την κατεύθυνση, μπορεί να περιλαμβάνουν, μεταξύ άλλων, τη χρήση περισσότερων δειγμάτων στο σύνολο δεδομένων για την εκπαίδευση των μοντέλων, την επιλογή χαρακτηριστικών [61] και την εφαρμογή τεχνικών μείωσης της διαστατικότητας [62], [63], [64] όπως οι αυτοκωδικοποιητές (autoencoders), καθώς το σύνολο δεδομένων αποτελείται από πολύ μεγάλο πλήθος χαρακτηριστικών (199). Επιπλέον, η ενσωμάτωση στο σύνολο δεδομένων,

χαρακτηριστικών άλλου είδους, όπως δεδομένα που αφορούν το ιατρικό ιστορικό των ατόμων που μπορεί να σχετίζεται με την αιτιολογία ανάπτυξης νευροεκφυλιστικών ασθενειών που σχετίζονται με τη γήρανση.

Ως προς την ερμηνευσιμότητα των μοντέλων, θα μπορούσε να μελετηθεί η ευστάθεια και αξιοπιστία (robustness) των τεχνικών ερμηνευσιμότητας που εφαρμόστηκαν. Κάτι τέτοιο θα ενίσχυε την αξιοπιστία ολόκληρου του μοντέλου ταξινόμησης, η οποία είναι εξαιρετικά σημαντική ιδιαίτερα σε εφαρμογές που αφορούν στην ιατρική διάγνωση.

7. Βιβλιογραφία

- [1] Bloch, L., & Friedrich, C. M. (2021). Data analysis with Shapley values for automatic subject selection in Alzheimer's disease data sets using interpretable machine learning. *Alzheimer's Research and Therapy*, 13(1). <https://doi.org/10.1186/s13195-021-00879-4>
- [2] Yim, D., Yeo, T. Y., & Park, M. H. (2020). Mild cognitive impairment, dementia, and cognitive dysfunction screening using machine learning. *Journal of International Medical Research*, 48(7). <https://doi.org/10.1177/0300060520936881>
- [3] Deture, M. A., & Dickson, D. W. (n.d.). The neuropathological diagnosis of Alzheimer's disease. <https://doi.org/10.1186/s13024-019-0333-5>
- [4] Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2021). Explainable ai: A review of machine learning interpretability methods. In *Entropy* (Vol. 23, Issue 1, pp. 1–45). MDPI AG. <https://doi.org/10.3390/e23010018>
- [5] Stamate, D., Alghambdi, W., Ogg, J., Hoile, R., & Murtagh, F. (2019). A Machine Learning Framework for Predicting Dementia and Mild Cognitive Impairment. *Proceedings - 17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018*, 671–678. <https://doi.org/10.1109/ICMLA.2018.00107>
- [6] Saarela, M., & Jauhiainen, S. (2021). Comparison of feature importance measures as explanations for classification models. *SN Applied Sciences*, 3(2). <https://doi.org/10.1007/s42452-021-04148-9>
- [7] 'Nowinski, W. L. '. (2011). *Biomechanics of the Brain* (K. Miller, Ed.). Springer New York. <https://doi.org/10.1007/978-1-4419-9997-9>
- [8] 'Shastry, B. S. '. (2009). SNPs: Impact on Gene Function and Phenotype. In A. A. Komar (Ed.), *Single Nucleotide Polymorphisms Methods and Protocols* (Vol. 578, pp. 3–22). Humana Press. <https://doi.org/10.1007/978-1-60327-411-1>
- [9] Jansen, I. E., Savage, J. E., Watanabe, K., Bryois, J., Williams, D. M., Steinberg, S., Sealock, J., Karlsson, I. K., Hägg, S., Athanasiu, L., Voyle, N., Proitsi, P., Witoelar, A., Stringer, S., Aarsland, D., Almdahl, I. S., Andersen, F., Bergh, S., Bettella, F., ... Posthuma, D. (2019). Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nature Genetics*, 51(3), 404–413. <https://doi.org/10.1038/s41588-018-0311-9>
- [10] Lee, E., Giovanello, K. S., Saykin, A. J., Xie, F., Kong, D., Wang, Y., Yang, L., Ibrahim, J. G., Doraiswamy, P. M., & Zhu, H. (2017). Single-nucleotide polymorphisms are associated with cognitive decline at Alzheimer's disease conversion within mild cognitive impairment patients. *Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring*, 8, 86–95. <https://doi.org/10.1016/j.dadm.2017.04.004>
- [11] Li, H., Wetten, S., Li, L., St Jean, P. L., Upmanyu, R., Surh, L., Hosford, D., Barnes, M. R., David Briley, J., Borrie, M., Coletta, N., Delisle, R., Dhalla, D., Ehm, M. G., Feldman, H. H., Fornazzari, L., Gauthier, S., Goodgame, N., Guzman, D., ... Roses, A. D. (2008). Candidate Single-Nucleotide Polymorphisms From a Genomewide Association Study of Alzheimer Disease. In *Arch Neurol* (Vol. 65, Issue 1). <http://www.GSK.com>
- [12] Coelho, L., Goertzel, B., Pennachin, C., & Heward, C. (n.d.). Classifier Ensemble Based Analysis of a Genome-Wide SNP Dataset Concerning Late-Onset Alzheimer Disease. <http://code.google.com/p/openbiomind/>

- [13] Wrigglesworth, J., Ward, P., Harding, I. H., Nilaweera, D., Wu, Z., Woods, R. L., & Ryan, J. (2021). Factors associated with brain ageing - a systematic review. *BMC Neurology*, 21(1). <https://doi.org/10.1186/s12883-021-02331-4>
- [14] Anderton, B. H. (2002). Ageing of the brain. In *Mechanisms of Ageing and Development* (Vol. 123). www.elsevier.com/locate/mechagedev
- [15] Petersen, R. C., & Negash, S. (2008). Mild cognitive impairment: An overview. In *CNS Spectrums* (Vol. 13, Issue 1, pp. 45–53). MBL Communications. <https://doi.org/10.1017/S1092852900016151>
- [16] Gauthier, S., Reisberg, B., Zaudig, M., Petersen, R. C., Ritchie, K., Broich, K., Belleville, S., Brodaty, H., Bennett, D., Chertkow, H., Cummings, J. L., de Leon, M., Feldman, H., Ganguli, M., Hampel, H., Scheltens, P., Tierney, M. C., Whitehouse, P., Winblad, B., & International Psychogeriatric Association Expert Conference on mild cognitive impairment (2006). Mild cognitive impairment. *Lancet (London, England)*, 367(9518), 1262–1270. [https://doi.org/10.1016/S0140-6736\(06\)68542-5](https://doi.org/10.1016/S0140-6736(06)68542-5)
- [17] Petersen, R. C. (2016). Mild Cognitive Impairment. *CONTINUUM: Lifelong Learning in Neurology*, 22(2, Dementia), 404–418. <https://doi.org/10.1212/CON.0000000000000313>
- [18] Arvanitakis, Z., Shah, R. C., & Bennett, D. A. (2019). Diagnosis and Management of Dementia: Review. *JAMA - Journal of the American Medical Association*, 322(16), 1589–1599. <https://doi.org/10.1001/jama.2019.4782>
- [19] [Bhardwaj, R., Nambiar, A. R., & Dutta, D. (2017). A Study of Machine Learning in Healthcare. *Proceedings - International Computer Software and Applications Conference*, 2, 236–241. <https://doi.org/10.1109/COMPSAC.2017.164>
- [20] Shailaja, K., Seetharamulu, B., & Jabbar, M. A. (2018). Machine Learning in Healthcare: A Review. *Proceedings of the 2nd International Conference on Electronics, Communication and Aerospace Technology, ICECA 2018*, 910–914. <https://doi.org/10.1109/ICECA.2018.8474918>
- [21] Dreiseitl, S., & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: A methodology review. *Journal of Biomedical Informatics*, 35(5–6), 352–359. [https://doi.org/10.1016/S1532-0464\(03\)00034-0](https://doi.org/10.1016/S1532-0464(03)00034-0)
- [22] Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test*, 25(2), 197–227. <https://doi.org/10.1007/s11749-016-0481-7>
- [23] Bernard, S., Heutte, L., & Adam, S. (2009). Influence of hyperparameters on random forest accuracy. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5519 LNCS, 171–180. https://doi.org/10.1007/978-3-642-02326-2_18/COVER
- [24] Biau, G., & Fr, G. B. (2012). Analysis of a Random Forests Model. In *Journal of Machine Learning Research* (Vol. 13).
- [25] Dai, B., Chen, R. C., Zhu, S. Z., & Zhang, W. W. (2019). Using random forest algorithm for breast cancer diagnosis. *Proceedings - 2018 International Symposium on Computer, Consumer and Control, IS3C 2018*, 449–452. <https://doi.org/10.1109/IS3C.2018.00119>
- [26] Probst, P., Wright, M. N., & Boulesteix, A. L. (2019). Hyperparameters and tuning strategies for random forest. In *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* (Vol. 9, Issue 3). Wiley-Blackwell. <https://doi.org/10.1002/widm.1301>

- [27] Boateng, E. Y., & Abaye, D. A. (2019). A Review of the Logistic Regression Model with Emphasis on Medical Research. *Journal of Data Analysis and Information Processing*, 07(04), 190–207. <https://doi.org/10.4236/jdaip.2019.74012>
- [28] Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*. <https://doi.org/10.1016/j.neucom.2019.10.118>
- [29] Pisner, D. A., & Schnyer, D. M. (2019). Support vector machine. In *Machine Learning: Methods and Applications to Brain Disorders* (pp. 101–121). Elsevier. <https://doi.org/10.1016/B978-0-12-815739-8.00006-7>
- [30] Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in Neuroinformatics*, 7(DEC). <https://doi.org/10.3389/fnbot.2013.00021>
- [31] Wang, Y., Ni, S., & Sherry Ni, X. (2019). A XGBoost risk model via feature selection and Bayesian hyper-parameter optimization. *International Journal of Database Management Systems (IJDBMS)*, 11(1). <https://doi.org/10.5121/ijdbms.2019.11101>
- [32] Zhang, R., Li, B., & Jiao, B. (2019). Application of XGboost Algorithm in Bearing Fault Diagnosis. *IOP Conference Series: Materials Science and Engineering*, 490(7). <https://doi.org/10.1088/1757-899X/490/7/072062>
- [33] Ramchoun, H., Amine, M., Idrissi, J., Ghanou, Y., & Ettaouil, M. (2016). Multilayer Perceptron: Architecture Optimization and Training. *International Journal of Interactive Multimedia and Artificial Intelligence*, 4(1), 26. <https://doi.org/10.9781/ijimai.2016.415>
- [34] Hakim, L., Sartono, B., & Saefuddin, A. (2017). Bagging Based Ensemble Classification Method on Imbalance Datasets. *IJCSN-International Journal of Computer Science and Network*, 6(6). www.IJCSN.orgImpactFactor:1.5670
- [35] Galar, M., Fernández, A., Barrenechea, E., Bustince, H., & Herrera, F. (n.d.). *Aggregation schemes for binarization techniques. Methods' description.*
- [36] Grandini, M., Bagli, E., & Visani, G. (2020). Metrics for Multi-Class Classification: an Overview. <http://arxiv.org/abs/2008.05756>
- [37] Rizi, F. Y., Au, J., Yli-Ollila, H., Golemati, S., Makūnaitė, M., Orkisz, M., Navab, N., MacDonald, M., Laitinen, T. M., Behnam, H., Gao, Z., Gastounioti, A., Jurkonis, R., Vray, D., Laitinen, T., Sérusclat, A., Nikita, K. S., & Zahnd, G. (2020). Carotid Wall Longitudinal Motion in Ultrasound Imaging: An Expert Consensus Review. *Ultrasound in Medicine & Biology*, 46(10), 2605–2624. <https://doi.org/10.1016/j.ultrasmedbio.2020.06.006>
- [38] Golemati, S., Patelaki, E., Gastounioti, A., Andreadis, I., Liapis, C. D., & Nikita, K. S. (2020). Motion synchronisation patterns of the carotid atheromatous plaque from B-mode ultrasound. *Scientific Reports*, 10(1), 11221. <https://doi.org/10.1038/s41598-020-65340-2>
- [39] Panayides, A. S., Amini, A., Filipovic, N. D., Sharma, A., Tsaftaris, S. A., Young, A., Foran, D., Do, N., Golemati, S., Kurc, T., Huang, K., Nikita, K. S., Veasey, B. P., Zervakis, M., Saltz, J. H., & Pattichis, C. S. (2020). AI in Medical Imaging Informatics: Current Challenges and Future Directions. *IEEE Journal of Biomedical and Health Informatics*, 24(7), 1837–1857. <https://doi.org/10.1109/JBHI.2020.2991043>
- [40] Athanasiou, M., Thanopoulou, A. C., Sfrintzeri, K., Nikita, K. S., & Zarkogianni, K. (n.d.). An explainable XGBoost-based approach towards assessing the risk of cardiovascular disease in patients with Type 2 Diabetes Mellitus.

- [41] Yanjun Qi. (2012). *Ensemble Machine Learning* (Cha Zhang & Yunqian Ma, Eds.). Springer New York, NY.
- [42] Vega García, M., & Aznarte, J. L. (2020). Shapley additive explanations for NO2 forecasting. *Ecological Informatics*, 56. <https://doi.org/10.1016/j.ecoinf.2019.101039>
- [43] Molnar, C. (2022). SHAP (SHapley Additive exPlanations). In *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable* (2nd ed.).
- [44] Molnar, C. (2022). *Local Surrogate (LIME)*. In *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable* (2nd ed.).
- [45] Molnar, C. (2022). *Partial Dependence Plot (PDP)*. In *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable* (2nd ed.).
- [46] Molnar, C. (2022). *Individual Conditional Expectation (ICE)*. In *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable* (2nd ed.).
- [47] Hashemi, M., & Fathi, A. (2020). PermuteAttack: Counterfactual Explanation of Machine Learning Credit Scorecards. <http://arxiv.org/abs/2008.10138>
- [48] Yang, Z., Nasrallah, I. M., Shou, H., Wen, J., Doshi, J., Habes, M., Erus, G., Abdulkadir, A., Resnick, S. M., Wolk, D., & Davatzikos, C. (n.d.). Disentangling brain heterogeneity via semi-supervised deep-learning and MRI: dimensional representations of Alzheimer's Disease.
- [49] Bounias, D., Singh, A., Bakas, S., Pati, S., Rathore, S., Akbari, H., Bilello, M., Greenberger, B. A., Lombardo, J., Chitalia, R. D., Jahani, N., Gastounioti, A., Hershman, M., Roshkovan, L., Katz, S. I., Yousefi, B., Lou, C., Simpson, A. L., Do, R. K. G., ... Davatzikos, C. (2021). Interactive Machine Learning-Based Multi-Label Segmentation of Solid Tumors and Organs. *Applied Sciences*, 11(16), 7488. <https://doi.org/10.3390/app11167488>
- [50] Shiino, A., Chen, Y. W., Tanigaki, K., Yamada, A., Vigers, P., Watanabe, T., Tooyama, I., & Akiguchi, I. (2017). Sex-related difference in human white matter volumes studied: Inspection of the corpus callosum and other white matter by VBM. *Scientific Reports*, 7. <https://doi.org/10.1038/srep39818>
- [51] Kennedy, K. M., Erickson, K. I., Rodrigue, K. M., Voss, M. W., Colcombe, S. J., Kramer, A. F., Acker, J. D., & Raz, N. (2009). Age-related differences in regional brain volumes: A comparison of optimized voxel-based morphometry to manual volumetry. *Neurobiology of Aging*, 30(10), 1657–1676. <https://doi.org/10.1016/j.neurobiolaging.2007.12.020>
- [52] Gerald, B. (2018). A Brief Review of Independent, Dependent and One Sample t-test. *International Journal of Applied Mathematics and Theoretical Physics*, 4(2), 50. <https://doi.org/10.11648/j.ijamtp.20180402.13>
- [53] Asvestas, P., Golemati, S., Matsopoulos, G. K., Nikita, K. S., & Nicolaidis, A. N. (2002). Fractal dimension estimation of carotid atherosclerotic plaques from B-mode ultrasound: a pilot study. *Ultrasound in medicine & biology*, 28(9), 1129–1136. [https://doi.org/10.1016/s0301-5629\(02\)00550-1](https://doi.org/10.1016/s0301-5629(02)00550-1)
- [54] Sarahi Alavez-Rubio, J., Martínez-Rodríguez, N., Escobedo-De-La-Peña, J., Garrido-Acosta, O., & Juárez-Cedillo, T. (2020). *Relationship Between Genetic Variants of ACAT1 and APOE with the Susceptibility to Dementia (SADEM Study)*. <https://doi.org/10.1007/s12035-020-02162-3/Published>

- [55] Lambert, J. C., Ibrahim-Verbaas, C. A., Harold, D., Naj, A. C., Sims, R., Bellenguez, C., Jun, G., DeStefano, A. L., Bis, J. C., Beecham, G. W., Grenier-Boley, B., Russo, G., Thornton-Wells, T. A., Jones, N., Smith, A. v., Chouraki, V., Thomas, C., Ikram, M. A., Zelenika, D., ... Seshadri, S. (2013). Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nature Genetics*, 45(12), 1452–1458. <https://doi.org/10.1038/ng.2802>
- [56] Giri, M., Zhang, M., & Lü, Y. (2016). Clinical Interventions in Aging Dovepress Genes associated with Alzheimer's disease: an overview and current status. *Clinical Interventions in Aging*, 11–665. <https://doi.org/10.2147/CIA.S105769>
- [57] Apostolova, L. G., Green, A. E., Babakchian, S., Hwang, K. S., Chou, Y. Y., Toga, A. W., & Thompson, P. M. (2012). Hippocampal atrophy and ventricular enlargement in normal aging, mild cognitive impairment (MCI), and Alzheimer disease. *Alzheimer Disease and Associated Disorders*, 26(1), 17–27. <https://doi.org/10.1097/WAD.0b013e3182163b62>
- [58] Packard, C. J., Westendorp, R. G. J., Stott, D. J., Caslake, M. J., Murray, H. M., Shepherd, J., Blauw, G. J., Murphy, M. B., Bollen, E. L. E. M., Buckley, B. M., Cobbe, S. M., Ford, I., Gaw, A., Hyland, M., Jukema, J. W., Kamper, A. M., Macfarlane, P. W., Jolles, J., Perry, I. J., ... Twomey, C. (2007). Association between apolipoprotein E4 and cognitive decline in elderly adults. *Journal of the American Geriatrics Society*, 55(11), 1777–1785. <https://doi.org/10.1111/j.1532-5415.2007.01415.x>
- [59] Jansen, I. E., Savage, J. E., Watanabe, K., Bryois, J., Williams, D. M., Steinberg, S., Sealock, J., Karlsson, I. K., Hägg, S., Athanasiu, L., Voyle, N., Proitsi, P., Witoelar, A., Stringer, S., Aarsland, D., Almdahl, I. S., Andersen, F., Bergh, S., Bettella, F., ... Posthuma, D. (2019). Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nature Genetics*, 51(3), 404–413. <https://doi.org/10.1038/s41588-018-0311-9>
- [60] Andrews, S. J., Fulton-Howard, B., & Goate, A. (2019). Protective Variants in Alzheimer's Disease. *Current Genetic Medicine Reports*, 7(1), 1–12. <https://doi.org/10.1007/s40142-019-0156-2>
- [61] Golemati, S., Tegos, T. J., Sassano, A., Nikita, K. S., & Nicolaidis, A. N. (2004). Echogenicity of B-mode sonographic images of the carotid artery: work in progress. *Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine*, 23(5), 659–669. <https://doi.org/10.7863/jum.2004.23.5.659>
- [62] Stoitsis, J., Valavanis, I., Mougiakakou, S. G., Golemati, S., Nikita, A., & Nikita, K. S. (2006). Computer aided diagnosis based on medical image processing and artificial intelligence methods. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 569(2), 591–595. <https://doi.org/10.1016/j.nima.2006.08.134>
- [63] Mougiakakou, S. Gr., Golemati, S., Gousias, I., Nicolaidis, A. N., & Nikita, K. S. (2007). Computer-aided diagnosis of carotid atherosclerosis based on ultrasound image statistics, Laws' texture and neural networks. *Ultrasound in Medicine & Biology*, 33(1), 26–36. <https://doi.org/10.1016/j.ultrasmedbio.2006.07.032>
- [64] Tzalavra, A. G., Andreadis, I., Dalakleidi, K., Constantinidis, F., Zacharaki, E., & Nikita, K. (2022). Dynamic contrast enhanced-magnetic resonance imaging radiomics combined with a hybrid adaptive <scp>neuro-fuzzy</scp> inference system-particle swarm optimization approach for breast tumour classification. *Expert Systems*, 39(4). <https://doi.org/10.1111/exsy.12895>

Πηγές Εικόνων

Εικόνα 1: 'Nowinski, W. L. '. (2011). *Biomechanics of the Brain* (K. Miller, Ed.). Springer New York. <https://doi.org/10.1007/978-1-4419-9997-9>

Εικόνα 2: 'Nowinski, W. L. '. (2011). *Biomechanics of the Brain* (K. Miller, Ed.). Springer New York. <https://doi.org/10.1007/978-1-4419-9997-9>

Εικόνα 3: 'Nowinski, W. L. '. (2011). *Biomechanics of the Brain* (K. Miller, Ed.). Springer New York. <https://doi.org/10.1007/978-1-4419-9997-9>

Εικόνα 4: Petersen, R. C. (2016). Mild Cognitive Impairment. *CONTINUUM: Lifelong Learning in Neurology*, 22(2, Dementia), 404–418. <https://doi.org/10.1212/CON.0000000000000313>

Εικόνα 5: Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*. <https://doi.org/10.1016/j.neucom.2019.10.118>

Εικόνα 6: Ramchoun, H., Amine, M., Idrissi, J., Ghanou, Y., & Ettaouil, M. (2016). Multilayer Perceptron: Architecture Optimization and Training. *International Journal of Interactive Multimedia and Artificial Intelligence*, 4(1), 26. <https://doi.org/10.9781/ijimai.2016.415>

Εικόνα 7: Bootstrap aggregating - Wikipedia. (n.d.). Retrieved September 2, 2022, from https://en.wikipedia.org/wiki/Bootstrap_aggregating

Εικόνα 8: Armi, L., & Fekri-Ershad, S. (2019). TEXTURE IMAGE ANALYSIS AND TEXTURE CLASSIFICATION METHODS-A REVIEW. In *International Online Journal of Image Processing and Pattern Recognition* (Vol. 2, Issue 1).

Εικόνα 9: Grandini, M., Bagli, E., & Visani, G. (2020). Metrics for Multi-Class Classification: an Overview. <http://arxiv.org/abs/2008.05756>