



Πανεπιστήμιο
Κύπρου

**EXTRACTING EXPLAINABLE ASSESSMENTS OF ALZHEIMER'S DISEASE
VIA MACHINE LEARNING ON BRAIN MRI DATA**

ΓΙΩΡΓΟΣ ΚΟΥΜΗΣ

Η Διατριβή αυτή Υποβλήθηκε προς Μερική Εκπλήρωση των Απαιτήσεων
για την Απόκτηση Τίτλου Σπουδών Master σε Προηγμένες Τεχνολογίες
Πληροφορικής στο Πανεπιστήμιο Κύπρου

Συστήνεται προς Αποδοχή από το Τμήμα Πληροφορικής

Λευκωσία, Ιούνιος 2024

ΣΕΛΙΔΑ ΕΓΚΡΙΣΗΣ

Διατριβή Master

EXTRACTING EXPLAINABLE ASSESSMENTS OF ALZHEIMER'S DISEASE VIA MACHINE LEARNING ON BRAIN MRI DATA

Παρουσιάστηκε από

Γιώργο Κουμή

Εξεταστική Επιτροπή:

Ερευνητικός Σύμβουλος: Καθηγητής Κωνσταντίνος Παττίχης

Μέλος Επιτροπής: Καθηγήτρια Ελπίδα Κεραυνού-Παπαηλιού

Μέλος Επιτροπής: Καθηγητής Χρίστος Χριστοδούλου

Πανεπιστήμιο Κύπρου

Ιούνιος, 2024

ABSTRACT

This master's thesis focuses on the development of machine learning models and argumentation frameworks for the early diagnosis of Alzheimer's disease (AD) using brain MRI data. AD is the most common type of dementia, often detected at advanced stages when symptoms are apparent, and the neurodegenerative process has significantly progressed. Early diagnosis is crucial for improving the quality of life for AD patients.

Therefore, the main objectives of this study are to identify the most significant features for early AD diagnosis, develop machine learning models to predict the disease, and create argumentation models based on the machine learning models to provide explanations for these predictions.

The study begins with an introduction to the concepts of dementia and AD, followed by a literature review on machine learning applications in AD diagnosis, highlighting the approaches and results of previous studies. Next, the data analysis process is described, focusing on identifying key features that contribute to the diagnosis of AD, along with the preparation and preprocessing of data for model training.

In addition, the thesis presents the methodology for developing the machine learning models, including the selection and training of models using appropriate techniques and parameters to predict AD. The models' performance is evaluated based on various metrics to assess their accuracy and reliability. Notably, when cognitive test score data is used, the accuracy of these models reaches 98%.

Furthermore, the thesis introduces the ArgEML argumentation tool, designed to provide clear and understandable explanations for the models' predictions. Explainability in artificial intelligence, particularly in the medical field, is critical for ensuring transparency and inspiring trust among medical experts and patients. The models developed through the ArgEML tool contribute to this goal by providing arguments that facilitate the interpretation of results. These explanations are achieved using rules derived from tree-based machine learning models.

Finally, a comparison is made between argumentation models and traditional machine learning models used for classification. Argumentation models demonstrate comparable performance in terms of prediction accuracy and transparency. An important distinction is that while some predictions from argumentation models may lack a definitive answer, the provided explanations for this uncertainty can further enhance the models' accuracy and reliability.

ΠΕΡΙΛΗΨΗ

Η παρούσα διατριβή επικεντρώνεται στην ανάπτυξη μοντέλων μηχανικής μάθησης και μοντέλων επιχειρηματολογίας για την έγκαιρη διάγνωση της νόσου Αλτσχάιμερ (NA), χρησιμοποιώντας δεδομένα MRI του εγκεφάλου. Η NA είναι μια μορφή άνοιας που συχνά εντοπίζεται σε προχωρημένα στάδια, όταν τα συμπτώματα είναι εμφανή και η νευροεκφυλιστική διαδικασία έχει ήδη προχωρήσει σημαντικά. Η ανάγκη για έγκαιρη διάγνωση είναι ζωτικής σημασίας για τη βελτίωση της ποιότητας ζωής των ασθενών με NA.

Επομένως, οι κύριοι στόχοι της παρούσας μελέτης περιλαμβάνουν τον εντοπισμό των σημαντικότερων χαρακτηριστικών του εγκεφάλου για την έγκαιρη διάγνωση της NA, την ανάπτυξη μοντέλων μηχανικής μάθησης για την πρόβλεψη της νόσου, καθώς και την ανάπτυξη μοντέλων επιχειρηματολογίας τα οποία είναι βασισμένα σε μοντέλα μηχανικής μάθησης και παρέχουν εξηγήσεις για τις προβλέψεις που δίνουν.

Η μελέτη ξεκινά κάνοντας μια εισαγωγή στις έννοιες της άνοιας και της NA. Γίνεται μια βιβλιογραφική ανασκόπηση στον τομέα της μηχανικής μάθησης και της NA, καταγράφοντας τις προσεγγίσεις και τα αποτελέσματα προηγούμενων μελετών. Στη συνέχεια, γίνεται ανάλυση των δεδομένων με σκοπό την εύρεση των σημαντικότερων χαρακτηριστικών που συμβάλλουν στη διάγνωση της νόσου και περιγράφεται η προετοιμασία και επεξεργασία των δεδομένων για την εκπαίδευση των μοντέλων.

Ακολούθως, περιγράφεται η μεθοδολογία ανάπτυξης των μοντέλων μηχανικής μάθησης που περιλαμβάνει την επιλογή και την εκπαίδευση μοντέλων χρησιμοποιώντας κατάλληλες τεχνικές, για την πρόβλεψη της NA. Η απόδοση των μοντέλων αξιολογείται βάσει διαφόρων μετρικών για να ελεγχθούν η ακρίβεια και η αξιοπιστία τους. Αξίζει να σημειωθεί ότι στις περιπτώσεις που χρησιμοποιείται και το γνωστικό τεστ ως δεδομένο, η ακρίβεια των μοντέλων αυτών φτάνει το 98%.

Επίσης, παρουσιάζεται το εργαλείο επιχειρηματολογίας ArgEML που έχει ως στόχο να παρέχει σαφείς και κατανοητές εξηγήσεις για τις προβλέψεις των μοντέλων. Η ανάγκη για παροχή εξηγήσεων στον τομέα της τεχνητής νοημοσύνης, και ιδιαίτερα στην ιατρική, είναι κρίσιμη, καθώς τα συστήματα αυτά χρειάζονται να παρέχουν διαφάνεια και να εμπνέουν εμπιστοσύνη στους ιατρούς και τους ασθενείς. Τα μοντέλα που αναπτύχθηκαν μέσω του εργαλείου επιχειρηματολογίας ArgEML σε αυτή τη μελέτη συμβάλλουν σε αυτή την κατεύθυνση, παρέχοντας εξηγήσεις (επιχειρήματα) που διευκολύνουν την ερμηνεία των αποτελεσμάτων. Αυτό το επιτυγχάνουν χρησιμοποιώντας κανόνες που εξάγονται από τα μοντέλα μηχανικής μάθησης τύπου δέντρων αποφάσεων.

Τέλος, γίνεται σύγκριση των μοντέλων επιχειρηματολογίας με τα παραδοσιακά μοντέλα μηχανικής μάθησης που χρησιμοποιούνται για ταξινόμηση. Παρατηρείται ότι τα μοντέλα επιχειρηματολογίας μπορούν να δώσουν εξίσου καλά αποτελέσματα, επιδεικνύοντας την αποτελεσματικότητά τους στην παροχή ακριβών προβλέψεων και διαφανών εξηγήσεων. Μια κύρια διαφορά τους, είναι ότι ένα ποσοστό των προβλέψεων των μοντέλων επιχειρηματολογίας δεν παρέχουν ξεκάθαρη απάντηση, δίνοντας επίσης εξηγήσεις γι' αυτό, κάτι το οποίο αυξάνει ακόμα περισσότερο την ακρίβεια του μοντέλου.

ΕΥΧΑΡΙΣΤΙΕΣ

Με την παρούσα μεταπτυχιακή διατριβή ολοκληρώνεται το μεταπτυχιακό πρόγραμμα μου στο Τμήμα Πληροφορικής του Πανεπιστημίου Κύπρου. Η ολοκλήρωση της εργασίας αυτής ήταν αποτέλεσμα της προσπάθειας αλλά και της συνεργασίας μου με διάφορα άτομα που η συμβολή τους ήταν καταλυτική.

Αρχικά, θα ήθελα να εκφράσω την εκτίμηση μου στον επιβλέποντα καθηγητή μου κ. Κωνσταντίνο Παττίχη για την εμπιστοσύνη που μου έδειξε και για την δυνατότητα που μου έδωσε να ασχοληθώ με ένα τόσο ενδιαφέρον θέμα. Ο επαγγελματισμός και οι γνώσεις του ήταν μια συνεχής πηγή στήριξης και έμπνευσης για μένα. Με την επιστημονική του κατάρτιση ήταν πάντα πρόθυμος να με βοηθήσει και με ενθάρρυνε σε όλα τα στάδια της μελέτης μου.

Επίσης, θα ήθελα να ευχαριστήσω ιδιαίτερα τον καθηγητή κ. Αντώνη Κάκα, για τη καθοδήγηση και τη βοήθεια που μου παρείχε σε σημαντικά ζητήματα της εργασίας μου. Τα σχόλια και οι παρατηρήσεις του ήταν σημαντικές για την ολοκλήρωση της διατριβής μου. Ακόμη, ευχαριστώ τον Δρ. Στέφανο Λεάνδρου για τη βοήθεια που μου παρείχε. Μέσα από συζητήσεις μας και την έρευνα του, μου επέτρεψε να εξερευνήσω βαθύτερα το συγκεκριμένο θέμα.

Τέλος, θα ήθελα να ευχαριστήσω θερμά την οικογένεια μου και τους φίλους μου που βρίσκονταν πάντα στο πλάι μου κατά τη διάρκεια του μεταπτυχιακού μου προγράμματος. Στάθηκαν δίπλα μου στις δύσκολες στιγμές, δείχνοντας υπομονή και κατανόηση.

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1: Εισαγωγή.....	1
1.1 Γενική Εισαγωγή.....	1
1.1.1 Εισαγωγή στη Άνοια	1
1.1.2 Νόσος Αλτσχάιμερ.....	1
1.1.3 Διάγνωση και θεραπεία της νόσου	2
1.2 Στόχοι παρούσας μελέτης.....	3
ΚΕΦΑΛΑΙΟ 2: Σχετική εργασία.....	5
ΚΕΦΑΛΑΙΟ 3: Ανάλυση δεδομένων	9
3.1 Βάσεις δεδομένων και μέθοδοι εξαγωγής δεδομένων.....	9
3.1.2 Βάσεις δεδομένων που χρησιμοποιήθηκαν	9
3.1.3 Περιγραφή εικόνων MRI που χρησιμοποιήθηκαν για την εξαγωγή δεδομένων ...	10
3.1.4 Περιγραφή τεχνικών που χρησιμοποιήθηκαν για τη εξαγωγή χαρακτηριστικών..	10
3.1.5 Εργαλείο FreeSurfer που χρησιμοποιήθηκε για τμηματοποίηση ROI	12
3.2 Ανάλυση και Περιγραφή δεδομένων που χρησιμοποιήθηκαν	13
3.2.1 Περιγραφή δεδομένων και Γνωστικό τεστ MMSE	13
3.2.2 Ανάλυση και σύγκριση των δεδομένων που χρησιμοποιήθηκαν	17
3.2.3 Επιλογή χαρακτηριστικών (Feature selection).....	20
3.2.4 Διαχωρισμός των δεδομένων για εκπαίδευση των μοντέλων μηχανικής μάθησης	23
ΚΕΦΑΛΑΙΟ 4: Μοντέλα μηχανικής μάθησης, μοντέλα επιχειρηματολογίας και μετρικές αξιολόγησής τους.....	24
4.1 Περιγραφή Μοντέλων μηχανικής μάθησης	24
4.2 Μοντέλα μηχανικής μάθησης με επιχειρηματολογία.....	26
4.2.1 Περιγραφή επιχειρηματολογίας	26
4.2.2 Περιγραφή εργαλείου Gorgias και εργαλείου επιχειρηματολογίας ArgEML.....	27
4.3 Περιγραφή μετρικών αξιολόγησής μοντέλων	29
ΚΕΦΑΛΑΙΟ 5: Ανάπτυξη και Αποτελέσματα Μοντέλων μηχανικής μάθησης χρησιμοποιώντας δεδομένα από βάση ADNI και βάση OASIS	32
5.1. Μεθοδολογία ανάπτυξης μοντέλων μηχανικής μάθησης.....	32
5.2 Αποτελέσματα μοντέλων μηχανικής μάθησης για διαχωρισμό NC και NA περιπτώσεων.....	33
5.3 Σύγκριση αποτελεσμάτων μοντέλων μηχανικής μάθησης.....	36

ΚΕΦΑΛΑΙΟ 6: Υλοποίηση και αποτελέσματα μοντέλων επιχειρηματολογίας χρησιμοποιώντας το εργαλείο Gorgias ArgEML	38
6.1 Μεθοδολογία ανάπτυξης μοντέλων επιχειρηματολογίας χρησιμοποιώντας το εργαλείο Gorgias ArgEML	38
6.1.1 Διαδικασία ανάπτυξης μοντέλου μηχανικής μάθησης με επιχειρηματολογία παρέχοντας εξήγηση	38
6.1.2 Μετατροπή δεδομένων από συνεχή σε κατηγορικά δεδομένα.....	39
6.1.3 Μέθοδος αξιολόγησης και επιλογή των κανόνων που εξάγονται από το εργαλείο TE2Rules για αρχικοποίηση του εργαλείου ArgEML.....	42
6.2 Παρουσίαση και Αποτελέσματα Μοντέλων επιχειρηματολογίας χρησιμοποιώντας δεδομένα ιππόκαμπου από βάση δεδομένων OASIS	44
6.2.1 Ανάπτυξη μοντέλων επιχειρηματολογίας με σκοπό τη μείωση των διλημάτων .53	
6.3 Παρουσίαση και Αποτελέσματα Μοντέλων επιχειρηματολογίας χρησιμοποιώντας δεδομένα ιππόκαμπου από βάση δεδομένων ADNI.....	55
6.3.1 Σύγκριση μοντέλων επιχειρηματολογίας με βάση των αριθμό κατηγοριών που επιλέχθηκαν για τη κατηγοροποίηση τους.....	61
6.3.2 Σύγκριση μοντέλων επιχειρηματολογίας με βάση το μοντέλο μηχανικής μάθησης που χρησιμοποιήθηκε αρχικά για τη αρχικοποίηση των κανόνων	63
6.4 Παρουσίαση και Αποτελέσματα Μοντέλων επιχειρηματολογίας χρησιμοποιώντας δεδομένα ιππόκαμπου και γνωστικού τεστ από βάση δεδομένων OASIS.....	64
6.5 Παρουσίαση και Αποτελέσματα Μοντέλων επιχειρηματολογίας χρησιμοποιώντας δεδομένα ιππόκαμπου και γνωστικού τεστ από βάση δεδομένων ADNI	70
ΚΕΦΑΛΑΙΟ 7: Συμπεράσματα	75
Βιβλιογραφία.....	79

ΚΑΤΑΛΟΓΟΣ ΔΙΑΓΡΑΜΜΑΤΩΝ

1. Διάγραμμα 3.1: Σημαντικότητα χαρακτηριστικών υφής και όγκου του ιππόκαμπου στο διαχωρισμό περιπτώσεων NC και AD από βάση δεδομένων ADNI
2. Διάγραμμα 3.2: Σημαντικότητα χαρακτηριστικών υφής του ιππόκαμπου και γνωστικού τεστ στο διαχωρισμό περιπτώσεων NC και AD από βάση δεδομένων ADNI
3. Διάγραμμα 3.3: Σημαντικότητα χαρακτηριστικών υφής και όγκου του ενδοκρινικού φλοιού στο διαχωρισμό περιπτώσεων NC και AD από βάση δεδομένων ADNI
4. Διάγραμμα 3.4: Σημαντικότητα χαρακτηριστικών υφής, όγκου του ενδοκρινικού φλοιού και γνωστικού τεστ στο διαχωρισμό περιπτώσεων NC και AD από βάση δεδομένων ADNI
5. Διάγραμμα 6.1: Αρχική κατανομή δεδομένων του όγκου του ιππόκαμπου από βάση ADNI
6. Διάγραμμα 6.2: Κατανομή δεδομένων του όγκου του ιππόκαμπου από βάση δεδομένων ADNI, μετά την κατηγοριοποίηση των δεδομένων σε 5 κατηγορίες
7. Διάγραμμα 6.3: Αρχική κατανομή δεδομένων γνωστικού τεστ πριν να χωριστεί σε κατηγορίες από βάση από βάση δεδομένων OASIS
8. Διάγραμμα 6.4: Αρχική κατανομή δεδομένων γνωστικού τεστ πριν να χωριστεί σε κατηγορίες από βάση δεδομένων ADNI

ΚΕΦΑΛΑΙΟ 1: Εισαγωγή

1.1 Γενική Εισαγωγή

1.1.1 Εισαγωγή στη Άνοια

Η άνοια είναι ένας γενικός όρος που σχετίζεται με τη απώλεια μνήμης, γλώσσας, ικανότητας επίλυσης προβλημάτων και άλλων ικανοτήτων σκέψης γνωστές και ως γνωστικές ικανότητες. Όλα αυτά τα συμπτώματα προκαλούνται από μη φυσιολογικές εγκεφαλικές αλλαγές και είναι αρκετά σοβαρά ώστε να παρεμβαίνουν στην καθημερινή ζωή του ασθενή. Επηρεάζουν επίσης τη συμπεριφορά, τα συναισθήματα και τις σχέσεις. Η συγκεκριμένη ασθένεια είναι συνήθως προοδευτική και συχνά δεν είναι εμφανές στα πρώιμα στάδιά της. Τα σημάδια της γνωστικής εξασθένησης ξεκινούν αργά και σταδιακά χειροτερεύουν με την πάροδο του χρόνου, οδηγώντας σε άνοια.

Η άνοια προκαλείται από μια ποικιλία ασθενειών που παρεμβαίνουν στα εγκεφαλικά κύτταρα εμποδίζοντας τη μεταξύ τους επικοινωνία. Όταν τα εγκεφαλικά κύτταρα δεν μπορούν να επικοινωνήσουν φυσιολογικά τότε οι γνωστικές ικανότητες του ασθενή εξασθενούν και η σκέψη, η συμπεριφορά και τα συναισθήματα του μπορεί να επηρεαστούν. Υπάρχουν διάφοροι τύποι άνοιας οι οποίοι κατηγοριοποιούνται με βάση το τύπο βλάβης που υπάρχει στα εγκεφαλικά κύτταρα και τις περιοχές του εγκεφάλου στη οποία εμφανίζεται [1].

1.1.2 Νόσος Αλτσχάιμερ

Ο πιο κοινός τύπος άγνοιας, είναι η νόσος του Αλτσχάιμερ που αντιπροσωπεύει το 80% των περιπτώσεων άνοιας. Στη νόσο του Αλτσχάιμερ, αυτά που δυσκολεύουν τα εγκεφαλικά κύτταρα να παραμείνουν υγιή και να επικοινωνήσουν μεταξύ τους είναι τα υψηλά επίπεδα ορισμένων πρωτεϊνών εντός και εκτός των εγκεφαλικών κυττάρων. Τα εγκεφαλικά κύτταρα

στην περιοχή του ιππόκαμπου είναι συχνά από τα πρώτα που καταστρέφονται. Ο ιππόκαμπος είναι περιοχή του εγκεφάλου και αποτελεί το κέντρο της μάθησης και της μνήμης του. Έτσι, λόγω του ότι τις περισσότερες φορές η απώλεια μνήμης είναι συχνά ένα από τα πρώτα συμπτώματα του Αλτσχάιμερ [2], ο ιππόκαμπος είναι ένα από τα μέρη που μελετάται εκτενέστερα για τη διάγνωση της νόσου. Εκτός, από τη περιοχή του ιππόκαμπου, ένα άλλο μέρος εντός του κροταφικού λοβού (MTL) ο ενδορινικός φλοιός, αναγνωρίζεται ως μια περιοχή που επηρεάζεται επίσης σοβαρά από την νόσο Αλτσχάιμερ και αναφέρεται ότι είναι ο πιο σοβαρά κατεστραμμένος φλοιός στην συγκεκριμένη ασθένεια [3]. Η ατροφία του ενδορινικού φλοιού προβλέπεται να συμβεί πριν από τη βλάβη του ιππόκαμπου, είναι ένα από τα πρώτα σημάδια εκδήλωσης της νόσου και επηρεάζει επίσης τη σοβαρότητα των γνωστικών συμπτωμάτων [4]

1.1.3 Διάγνωση και θεραπεία της νόσου

Δεν υπάρχει κανένα τεστ που να μπορεί να διαπιστώσει με ακρίβεια εάν κάποιος έχει άνοια. Οι γιατροί διαγιγνώσκουν το Αλτσχάιμερ και άλλους τύπους άνοιας με βάση προσεκτικό ιατρικό ιστορικό, φυσική εξέταση, εργαστηριακές εξετάσεις και την καθημερινή λειτουργία. Επίσης, η διάγνωση της νόσου βασίζεται σε κλινικές και νευροψυχολογικές εξετάσεις [5] που αξιολογούν τη μνήμη και τις γλωσσικές ικανότητες. Είναι πιο δύσκολο να προσδιοριστεί ο ακριβής τύπος άνοιας επειδή τα συμπτώματα και οι εγκεφαλικές αλλαγές διαφορετικών ανοιών μπορεί να αλληλεπικαλύπτονται. Οι γιατροί μπορούν να προσδιορίσουν ότι ένα άτομο έχει άνοια με υψηλό επίπεδο βεβαιότητας και η ασθένεια μπορεί να επιβεβαιωθεί μόνο με μεταθανάτιο υλικό [6]. Ωστόσο, με τη χρήση ειδικών τεχνικών απεικόνισης, είναι πλέον εφικτό να ανιχνευθεί η νόσος Αλτσχάιμερ αρκετά χρόνια πριν εμφανιστούν κλινικά συμπτώματα και ίσως η τροποποίηση της θεραπείας θα μπορούσε να είναι πιο ωφέλιμη. Η πρόοδος στην ιατρική απεικόνιση, τη γενετική, τα βιολογικά δείγματα και τα κλινικά

δεδομένα δίνουν αυτή την ευκαιρία στην ερευνητική κοινότητα [7]. Σχετικά με τη θεραπεία της νόσου, δυστυχώς, ακόμη και στη σημερινή εποχή, δεν έχει βρεθεί θεραπεία που να μπορεί να σταματήσει τον προοδευτικό νευροεκφυλισμό της νόσου Αλτσχάιμερ, αλλά μόνο για να επιβραδύνει την επιδείνωση των συμπτωμάτων της [7].

1.2 Στόχοι παρούσας μελέτης

Η διάγνωση της ασθένειας-άνοιας βασίζεται κυρίως σε γνωστικά τεστ που έχουν όμως πολλούς περιορισμούς. Επίσης, οι εργαστηριακές εξετάσεις και η ποιοτική απεικόνιση δεν θα παράσχουν έγκαιρη διάγνωση επειδή ο ακτινολόγος θα αντιληφθεί την ατροφία του εγκεφάλου σε ένα όψιμο στάδιο της νόσου. Έστω και αν υπάρχουν πολλά διαθέσιμα δεδομένα υψηλών διαστάσεων και απεικονίσεις, όλα αυτά είναι δύσκολο να αναλυθούν με τις υφιστάμενες ιατρικές μεθόδους. Ως αποτέλεσμα, η υψηλή ζήτηση υπολογιστικής ανάλυσης, έχει εξελίξει τη χρήση μεθόδων υπολογιστικής μηχανικής μάθησης (ML) για την ολοκληρωτική ανάλυση αυτών των δεδομένων.

Σήμερα, οι μέθοδοι μηχανικής μάθησης χρησιμοποιούνται για να αξιοποιήσουν τα δεδομένα υψηλών διαστάσεων, την ενσωμάτωση δεδομένων από διαφορετικές πηγές, τη μοντελοποίηση της αιτιολογικής και κλινικής ετερογένειας και στοχεύουν την ανακάλυψη νέων βιοδεικτών στην αξιολόγηση της νόσου Αλτσχάιμερ [6].

Ο 1ος στόχος της μελέτης αυτής είναι η ανακάλυψη των σημαντικότερων χαρακτηριστικών όπως ηλικία, φύλο, γνωστικό τεστ, χαρακτηριστικά ενδορινικού φλοιού και του ιππόκαμπου μπορεί να μας βοηθήσουν στο εντοπίσουμε έγκαιρα τη ύπαρξη Αλτσχάιμερ στο άτομο, ώστε γνωρίζοντας να προβεί έγκαιρα σε θεραπεία και να προλάβει τη ανάπτυξη της πάθησης.

Ο 2ος στόχος αυτής της διατριβής είναι να γίνει χρήση μοντέλων μηχανικής μάθησης ώστε να αξιολογηθεί η διαγνωστική και η προγνωστική ικανότητα των χαρακτηριστικών του ιππόκαμπου του εγκεφάλου για τη έγκαιρη διάγνωση της νόσου Αλτσχάιμερ.

Ωστόσο, στα μοντέλα που χρησιμοποιούνται με μηχανική μάθηση και δεν παρέχεται κάποια εξήγηση και επιχειρήματα που θα στηρίζουν τη απόφαση που πάρθηκε.

Έτσι, ένας 3ος στόχος της μελέτης, είναι να χρησιμοποιηθούν μοντέλα τα οποία θα μπορούν να δώσουν και εξήγηση μαζί με τα αποτελέσματα, προσφέροντας έτσι διαφάνεια (transparency). Αυτό είναι απαραίτητο για ηθικούς λόγους να υπάρχει και είναι σημαντικό για του γιατρούς, αφού μέσω της φυσικής αλληλεπίδρασης τους με το σύστημα μπορούν να γνωρίζουν για ποιους λόγους αποφασίστηκε η συγκεκριμένη απάντηση και να καταλάβουν πως σκέφτεται το σύστημα. Έτσι μπορούν να κατανοήσουν αναθεωρήσουν , συμβάλουν στη απόφαση αυξάνοντας τη ακρίβεια στη διάγνωση.

ΚΕΦΑΛΑΙΟ 2: Σχετική εργασία

Οι έρευνες που διεξάγονται σχετικά με τις ασθένειες της άνοιας και ιδιαίτερα η νόσος του Αλτσχάιμερ αντιπροσωπεύουν σημαντικό τομέα της ιατρικής επιστημονικής έρευνας, καθώς οι ασθένειες αυτές αποτελούν μια αυξανόμενη παγκόσμια πρόκληση στο τομέα της υγείας. Οι επιστήμονες διεξάγουν εκτεταμένες έρευνες με στόχο την κατανόηση των βιολογικών, γενετικών και περιβαλλοντικών παραγόντων που σχετίζονται με την εμφάνιση και εξέλιξη της νόσου. Τα τελευταία χρόνια, έχουν σημειωθεί σημαντικές εξελίξεις στην ανίχνευση πρόωρων συμπτωμάτων της νόσου, καθώς και στην ανάπτυξη νέων τεχνικών διάγνωσης.

Ο Λεάνδρου Σ. [7] είχε μελετήσει τη χρησιμότητα της ποσοτικής απεικόνισης μαγνητικής τομογραφίας (MRI) και συγκεκριμένα των χαρακτηριστικών της υφής στην αξιολόγηση της Ήπιας Γνωσιακής Εξασθένησης (MCI) και της νόσου του Αλτσχάιμερ. Στη συγκεκριμένη έρευνα, αρχικά εξετάστηκε εάν ένα ισχυρότερο μαγνητικό πεδίο (MRI) θα ωφελούσε την ποσοτική ανάλυση απεικόνισης που προέρχεται από χαρακτηριστικά υφής του ενδορινικού φλοιού [8]. Μετέπειτα, έγινε έρευνα στη διάγνωση και την πρόβλεψη των περιπτώσεων MCI και AD με βάση χαρακτηριστικά υφής του ενδορινικού φλοιού, κάτι το οποίο δεν είχε γίνει μέχρι στιγμής σε άλλες έρευνες [9]. Βάση της έρευνας, έδειξε ότι ο συνδυασμός χαρακτηριστικών της υφής του ενδορινικού φλοιού παρείχε καλύτερα αποτελέσματα στην ταξινόμηση των περιπτώσεων NC vs AD, MCI vs MCIc, MCI vs AD σε σύγκριση με τον όγκο και το πάχος του ιππόκαμπου ή του ενδορινικού φλοιού. Επίσης, έδειξε ότι με αυτά τα χαρακτηριστικά υφής μπορεί να γίνει καλύτερη πρόβλεψη της μετατροπής από κατάσταση MCI σε AD.

Μια άλλη έρευνα που έγινε η οποία σχετίζεται άμεσα με τη παρούσα διατριβή, είναι αυτή της Κ. Γ. Αχιλλέως [10]. Σκοπός της μελέτης αυτής ήταν να δείξει τη χρησιμότητα της

εξαγωγής κανόνων στην αξιολόγηση της AD με βάση τα χαρακτηριστικά της μαγνητικής τομογραφίας και να παρουσιάσει τα θετικά αποτελέσματα της χρήσης του συμβολικού συλλογισμού που βασίζεται στην επιχειρηματολογία για τη σύνθεση και την ερμηνεία των αποτελεσμάτων μηχανικής μάθησης. Συγκεκριμένα, αυτή η μελέτη εστιάζει στην εξαγωγή μοντέλων τα οποία θα διαχωρίζουν τις περιπτώσεις υγιούς ατόμων (NC) και ατόμων με νόσο Αλτσχάιμερ (AD) δίνοντας μαζί με το αποτέλεσμα και εξήγηση. Αρχικά, η έρευνα κτίζει μοντέλα μηχανικής μάθησης, χρησιμοποιώντας αλγόριθμους δέντρων αποφάσεων (DT) και τυχαίων δασών (RF) μέσω της πλατφόρμας ανάλυσης KNIME. Αφού δημιουργήσει αυτά μοντέλα, στη συνέχεια από αυτά εξάγει τους κανόνες σε ένα πλαίσιο συλλογιστικής που βασίζεται σε επιχειρηματολογία. Στη συνέχεια, οι κανόνες ερμηνεύονται μέσω του εργαλείου Gorgias ώστε τα αποτελέσματα να είναι κατανοητά από τους χρήστες. Το μοντέλο επιχειρηματολογίας που εφαρμόστηκε πέτυχε μέση ακρίβεια (accuracy) 91%, παρουσιάζοντας βελτιωμένα αποτελέσματα σε σύγκριση με τα αντίστοιχα μοντέλα DT και RF μηχανικής μάθησης.

Στο πίνακα 2.1, φαίνονται επιπλέον μελέτες μαγνητικής τομογραφίας όπου χρησιμοποιήθηκαν τεχνικές μηχανικής μάθησης (ML) και ραδιομικά χαρακτηριστικά για την αξιολόγηση της νόσου [6]. Στις πλείστες έρευνες είχαν χρησιμοποιηθεί μοντέλα μηχανικής μάθησης όπως Logistic regression, Decision tree και δεδομένα από MRI. Παρατηρούμε ότι στις έρευνες στις οποίες χρησιμοποιήθηκε και το γνωστικό τεστ ως δεδομένο τα αποτελέσματα ήταν πολύ καλύτερα.

Συνοψίζοντας, πολλές είναι οι έρευνες που έχουν γίνει σχετικά με την ασθένεια άνοιας. Οι ερευνητικές αυτές προσπάθειες σχετικά με τον AD αναδεικνύουν τη σημασία της πολυπλοκότητας της νόσου και την ανάγκη για ολοκληρωμένες προσεγγίσεις. Οι προσφορές

των προαναφερθέντων ερευνητών ανοίγουν νέες προοπτικές για την κατανόηση, πρόληψη και θεραπεία της νόσου, παρέχοντας σημαντική πρόοδο στον τομέα της νευροεπιστήμης και της ψυχιατρικής.

References	Subjects	Description	Split	Methodology	Results
Bogdanovic et al., 2022	Total 9,592 subjects (NC, EMCI, LMCI, SMC. AD)	Structural MRI, PET, gene expression and cognitive measures.	Training (70%) and testing (30%)	Correlation analysis	f1-score:0.84
Shu et al., 2021	MCI: 357 AD:154	Structural MRI, CSF, APOE ε4, cognitive measures.	Training (70%) and Testing (30%)	Logistic regression	AUC of 0.814, sensitivity of 0.726, and specificity of 0.798.
Achilleos et al., 2020	NC: 144, AD: 69	Haralick features from hippocampus.	10-fold cross validation	Decision tree and random forests	Accuracy: 0.770
Khan and Zubair, 2020	343 sessions–150 subjects (NC: 72, AD:78)	Structural MRI, cognitive measures, demographics.	Random selection allocation for train, validate and test	Random Forest	Accuracy: 0.868 precision: 0.941 recall: 0.8 AUC: 0.872
Battineni et al., 2020	373 sessions–150 subjects (NC:72,	Structural MRI, cognitive measures, demographics.	10-fold cross validation	Hybrid modeling	Accuracy: 0.980 precision: 0.981 recall: 0.980

	AD:64, MCIc:14)				ROC: 0.991
Kim et al., 2019	NC:146, AD: 143	Structural MRI.	10-fold cross validation	Principal component and linear discriminant analysis	Accuracy: 75.8%
Spasov et al., 2018	NC: 184, MCI: 409 AD: 192	Structural MRI, CSF, APOE ε4, cognitive measures.	10-fold cross validation	Convolution neural network	AUC of 0.925, accuracy: 86%, sensitivity: 87.5% and specificity: 85%

Πίνακας 2.1: Προηγούμενες έρευνες στις οποίες χρησιμοποιήθηκαν τεχνικές μηχανικής μάθησης (ML) και ραδιομικά χαρακτηριστικά για την αξιολόγηση της ΝΑ

ΚΕΦΑΛΑΙΟ 3: Ανάλυση δεδομένων

3.1 Βάσεις δεδομένων και μεθόδοι εξαγωγής δεδομένων

3.1.2 Βάσεις δεδομένων που χρησιμοποιήθηκαν

Τα δεδομένα που χρησιμοποιούνται με σκοπό μελέτη της νόσου Αλτσχάιμερ συλλέγονται από βάσεις δεδομένων που είναι διαθέσιμες διαδικτυακά. Μερικές από τις κυριότερες βάσεις δεδομένων προέρχονται από τη Πρωτοβουλία Νευροαπεικόνισης για τη Νόσο του Alzheimer (Alzheimer's Disease Neuroimaging Initiative - ADNI), από τη Australian Imaging Biomarker and Lifestyle flagship και το έργο Open Access Series of Imaging Studies (OASIS) κ.α.

Στη μελέτη αυτή, μια βάση δεδομένων που χρησιμοποιείται είναι η Alzheimer's Disease Neuroimaging Initiative (ADNI). Η πρωτοβουλία ADNI ξεκίνησε το 2004 ως συνεργασία δημόσιου και ιδιωτικού τομέα, με επικεφαλής τον κύριο ερευνητή Michael W. Weiner, MD. Ο πρωταρχικός στόχος της ADNI ήταν να εξεταστεί κατά πόσον η σειριακή απεικόνιση μαγνητικού συντονισμού (MRI), η τομογραφία εκπομπής ποζιτρονίων (PET), άλλοι βιολογικοί δείκτες και η κλινική και νευροψυχολογική αξιολόγηση μπορούν να συνδυαστούν για τη μέτρηση της εξέλιξης της ήπιας νοητικής αναπηρίας (Mild Cognitive Impairment, MCI) και της πρώιμης νόσου Αλτσχάιμερ. Άλλος ένας στόχος της είναι να υποστηρίξει τις προόδους στην παρέμβαση, την πρόληψη και τη θεραπεία της AD μέσω της εφαρμογής νέων διαγνωστικών μεθόδων στα πιο πρώιμα δυνατά στάδια όταν η παρέμβαση μπορεί να είναι πιο αποτελεσματική. Από την έναρξή της μέχρι και σήμερα έχει συνεισφέρει σημαντικά στην έρευνα AD, επιτρέποντας την ανταλλαγή δεδομένων μεταξύ ερευνητών σε όλο τον κόσμο [11].

Μια άλλη βάση δεδομένων που εξετάστηκε στη μελέτη αυτή από το έργο Open Access Series of Imaging Studies (OASIS). Το Open Access Series of Imaging Studies (OASIS) είναι ένα έργο

που στοχεύει να καταστήσει διαθέσιμα στην επιστημονική κοινότητα σύνολα δεδομένων νευροαπεικόνισης του εγκεφάλου. Συγκεντρώνοντας και διανέμοντας ελεύθερα σύνολα δεδομένων νευροαπεικόνισης, στόχος της βάσης δεδομένων αυτής είναι να διευκολύνει μελλοντικές ανακαλύψεις στη βασική και κλινική νευροεπιστήμη παρόμοιες με άλλες πρωτοβουλίες όπως η βάση δεδομένων ADNI. Η προετοιμασία και η απελευθέρωση του συνόλου δεδομένων που πάρθηκε από αυτή τη βάση δεδομένων, ακολουθεί τα πρότυπα που καθιερώθηκαν με την αρχική έκδοση, όπως προσεκτικός ποιοτικός έλεγχος, λεπτομερής τεκμηρίωση, παραδείγματα εικόνων μεταγενέστερης επεξεργασίας, πλήρης ανωνυμοποίηση, πολλαπλή πρόσβαση [12].

3.1.3 Περιγραφή εικόνων MRI που χρησιμοποιήθηκαν για την εξαγωγή δεδομένων

Οι εικόνες μαγνητικής τομογραφίας εγκεφάλου που λήφθηκαν ήταν μέρος του πλήρους συνόλου δεδομένων ADNI-I. Όλα τα άτομα είχαν ένα τυποποιημένο πρωτόκολλο [11] σε μονάδες μαγνητικής τομογραφίας 1,5-T από τη Siemens Medical Solutions και τη General Electric Healthcare. Τα πρωτόκολλα μαγνητικής τομογραφίας περιλάμβαναν υψηλής ανάλυσης (συνήθως $1,25 \times 1,25 \times 1,25 \text{ mm}^3$ voxels) ογκομετρική τρισδιάστατο οβελιαία μαγνητισμό με στάθμιση T1 που προετοίμασε σαρώσεις ταχείας διαβάθμισης ήχου (MPRAGE). Οι τεχνικές απόκτησης δεδομένων MRI τυποποιήθηκαν σε διαφορετικές τοποθεσίες σύμφωνα με το πρωτόκολλο ADNI.

3.1.4 Περιγραφή τεχνικών που χρησιμοποιήθηκαν για τη εξαγωγή χαρακτηριστικών

Μέσω της εξαγωγής χαρακτηριστικών (feature extraction) είναι δυνατή η ανάκτηση σημαντικών δεδομένων που μπορούν να βοηθήσουν στον εντοπισμό μιας παθολογίας. Οι μεθοδολογίες εξαγωγής χαρακτηριστικών αναλύουν αντικείμενα ή εικόνες με σκοπό να εξαγάγουν τα πιο σημαντικά χαρακτηριστικά που είναι αντιπροσωπευτικά των διαφόρων κατηγοριών αυτών των αντικειμένων.

Οι επιλεγμένες μέθοδοι εξαγωγής χαρακτηριστικών που χρησιμοποιούνται επί του παρόντος στην αξιολόγηση της ΝΑ μπορούν να χωριστούν σε τρεις κατηγορίες, ανάλογα με τον τύπο των χαρακτηριστικών που εξάγονται από τη μαγνητική τομογραφία: (i) Βάσει Voxel εάν τα χαρακτηριστικά προέρχονται από GM, WM ή CSF, (ii) Βάσει Vertex εάν τα χαρακτηριστικά προέρχονται από την επιφάνεια του φλοιού, όπως μετρήσεις πάχους και (iii) με βάση τη περιοχή ενδιαφέροντος ROI, εάν τα χαρακτηριστικά προέρχονται από ROI [13]. Όσο και ο ιππόκαμπος τόσο και ο ενδορινικός φλοιός είναι δομές που χρησιμοποιούνται σήμερα από πολλές μελέτες για την εκτίμηση της AD.

Στη συγκεκριμένη μελέτη, για την εξαγωγή των χαρακτηριστικών επιλέχθηκε η μέθοδος της ποσοτικής απεικόνισης με βάση τη περιοχή ενδιαφέροντος (ROI). Η περιοχή ενδιαφέροντος περιλαμβάνει κάποια ποιοτικά χαρακτηριστικά όπως το μέγεθος, το σχήμα, η εικασία, η σπηλαίωση ή η ενίσχυση της αντίθεσης, με τα οποία οι ακτινολόγοι μπορούν να περιγράψουν και να ερμηνεύσουν διαγνωστικές ιατρικές εικόνες. Στη συνέχεια, από αυτά τα ROI εξάγονται ποσοτικά χαρακτηριστικά για την ανάπτυξη διαγνωστικών ή προγνωστικών μοντέλων. Η αναγκαιότητα της ποσοτικής απεικόνισης στην αξιολόγηση της ΝΑ προέρχεται από το γεγονός ότι το ανθρώπινο μάτι δεν μπορεί να αντιληφθεί τις ανατομικές αλλαγές μέσω της ποιοτικής απεικόνισης στα αρχικά στάδια της νόσου. Παράλληλα, στην ιατρική απεικόνιση, υπάρχει η ραδιομική ανάλυση η οποία υποκινείται από την ιδέα ότι οι βιοϊατρικές εικόνες περιέχουν κρυφές πληροφορίες που αντικατοπτρίζουν την υποκείμενη παθοφυσιολογία και ότι αυτές οι σχέσεις μπορούν να αποκαλυφθούν μέσω ποσοτικής ανάλυσης εικόνων. Με άλλα λόγια, η ραδιομική ανάλυση αναφέρεται στο εκχύλισμα ενός μεγάλου αριθμού ποσοτικών χαρακτηριστικών που θα χρησιμοποιηθούν για τη βελτίωση της διάγνωσης, της πρόγνωσης και της υποστήριξης αποφάσεων. Έτσι, μέσω αυτής της τεχνικής μπορούν να εξαχθούν μοναδικές πληροφορίες που ενδέχεται να περιλαμβάνουν

νευροεκφυλιστικές αλλαγές σε μικροσκοπικό επίπεδο ακόμα και πριν συμβεί η ατροφία του εγκεφάλου [10]. Παρέχονται δηλαδή στο κλινικό ιατρό πολύτιμες πληροφορίες και χαρακτηριστικά (μοτίβα) τα οποία μπορούν να αναγνωριστούν και να εξαχθούν από αλγορίθμους υπολογιστών αλλά όχι μέσω του ανθρώπινου ματιού. Τέλος, μέσω της ποσοτικής απεικόνισης μπορούν να εξαχθούν δεδομένα υψηλών διαστάσεων, όπως ιστογράμματα, χαρακτηριστικά υφής, κυματομορφές, μετασχηματισμοί Laplacian ή συναρτήσεις Minkowski [6].

3.1.5 Εργαλείο FreeSurfer που χρησιμοποιήθηκε για τμηματοποίηση ROI

Η τμηματοποίηση της περιοχής ενδιαφέροντος (ROI) που προαναφέρθηκε, έγινε μέσω του εργαλείου Freesurfer v6.0 (<https://surfer.nmr.mgh.harvard.edu/>). Το Freesurfer εξακολουθεί να είναι μέχρι και σήμερα, το πιο συχνά χρησιμοποιούμενο λογισμικό στην ποσοτική νευροαπεικόνιση και αντιπροσωπεύει μια αυτοματοποιημένη τεχνική που παρέχει τμηματοποίηση αναθέτοντας μια ετικέτα νευροανατομίας σε κάθε voxel στον όγκο της MRI εικόνας. Για την παραγωγή τμηματοποίησης, χρησιμοποιείται ένας χάρτης πιθανοτήτων μαζί με τον αλγόριθμο συμπερασμάτων Bayesian από ένα σύνολο εκπαίδευσης δεδομένων με ετικέτα. Αυτός ο πιθανοτικός αλγόριθμος μπορεί να χρησιμοποιηθεί για να ορίσει πληροφορίες καμπυλότητας του εγκεφαλικού φλοιού σε νευροανατομικές περιοχές που βασίζονται σε γυροσκόπιο, οι οποίες αντιπροσωπεύουν τυπικές μετρήσεις του πάχους και της επιφάνειας του φλοιού [14]. Οι πληροφορίες έντασης και συνέχειας από τις διαδικασίες τμηματοποίησης και παραμόρφωσης χρησιμοποιούνται για την παραγωγή αναπαραστάσεων των ορίων του φλοιού. Το πάχος του φλοιού υπολογίζεται ως η πλησιέστερη απόσταση από το όριο GM/WM έως το όριο GM/CSF σε κάθε κορυφή της πτυσσόμενης επιφάνειας. Μέχρι στιγμής, το FreeSurfer είναι

το πιο συχνά χρησιμοποιούμενο λογισμικό για την ανάλυση του πάχους του φλοιού στην AD [10].

Τα χαρακτηριστικά της υφής υπολογίστηκαν χρησιμοποιώντας την πλατφόρμα KNIME Analytics [15]. Υπολογίστηκαν τα ακόλουθα χαρακτηριστικά υφής [16]: Γωνιακή Δεύτερη Στιγμή (ASM), Αντίθεση, Συσχέτιση, Διακύμανση, Μέσος όρος Αθροίσματος, Διακύμανση αθροίσματος και Εντροπία. Ο μέσος όρος τους υπολογίστηκε σε τέσσερις κατευθύνσεις (0°, 45°, 90°, 135°) με την απόσταση μεταξύ γειτονικών εικονοστοιχείων να έχει οριστεί στο 1.

3.2 Ανάλυση και Περιγραφή δεδομένων που χρησιμοποιήθηκαν

3.2.1 Περιγραφή δεδομένων και Γνωστικό τεστ MMSE

Όπως αναφέρθηκε στη προηγούμενη ενότητα, προκειμένου να εξαχθούν οι πιθανοί παράγοντες κινδύνου και οι αλληλεπιδράσεις τους που συμβάλλουν στην πιθανότητα εμφάνισης Νόσου Αλτσχάιμερ, απαιτείται μια μεγάλη βάση δεδομένων. Στην περίπτωση του συνόλου δεδομένων μας, τα δεδομένα συλλέχθηκαν από τη βάση ADNI και από τη βάση OASIS. Τα πλείστα δεδομένα μας ήταν αριθμητικά (Numerical) και κάποια ήταν τύπου Boolean. Τα χαρακτηριστικά τύπου Boolean χρησιμοποιούν τον αριθμό μηδέν για να αναπαραστήσουν μια αρνητική κατάσταση και τον αριθμό ένα για να αναπαραστήσουν μια κατάσταση θετική. Για παράδειγμα, εάν ένας ασθενής προβλεφθεί ως υγιείς (NC) το μοντέλο που έδινε τη τιμή 0 ως αποτέλεσμα, διαφορετικά θα έδινε τιμή 1 που σημαίνει ότι ο ασθενής που προβλέπεται να αναπτύξει νόσο Αλτσχάιμερ (θετική περίπτωση).

Τα δεδομένα που χρησιμοποιήθηκαν σε αυτή τη μελέτη, αποτελούνται από τα χαρακτηριστικά του ιππόκαμπου και του ενδοκρινικού φλοιού που αποτελούν μέρη του εγκεφάλου. Συγκεκριμένα, σε αυτή τη μελέτη, χρησιμοποιήθηκαν ο όγκος του ιππόκαμπου,

ο όγκος του ενδορινικού φλοιού τα οποία περιγράφονται στο κεφάλαιο 1. Ακόμη, χρησιμοποιήθηκαν οκτώ χαρακτηριστικά υφής (texture features) για την ποσοτική αξιολόγηση των χαρακτηριστικών υφής του ενδορινικού φλοιού και του ιππόκαμπου. Τα χαρακτηριστικά αυτά αναγράφονται στο πίνακα 3.1 και περιγράφονται λεπτομερώς από τους J.-H. Cai [17]. Επίσης, ένα άλλο σημαντικό δεδομένο που χρησιμοποιήθηκε είναι η τιμή της εξέτασης Mini Mental State (MMSE) η οποία περιγράφεται αναλυτικά στη συνέχεια.

Feature Description	Type
Ηλικία	Numerical
Φύλο	Boolean (M/F)
Όγκος	Numerical
ASM	Numerical
Contrast	Numerical
Corelation	Numerical
Variance	Numerical
Sum Average	Numerical
Entropy	Numerical
Cluster Shade	Numerical
MMSE	Numerical

Πίνακας 3.1: Χαρακτηριστικά που επιλέχθηκαν για τη μελέτη. Τα Volume, ASM, Contrast, Corelation baseline, Variance, Sum Average, Sum Variance, Entropy, Cluster Shade αφορούν τα texture features και στη περίπτωση μελέτης του ιππόκαμπου και στη περίπτωση μελέτης του ενδορινικού φλοιού

Μέχρι σήμερα, η διάγνωση της ΝΑ βασίζεται στα κλινικά χαρακτηριστικά της νόσου που αφορούν παρουσία διαταραχής μνήμης και έκπτωσης, τα οποία παρεμβαίνουν στην κοινωνική λειτουργία ή τις δραστηριότητες της καθημερινής ζωής. Η εξέταση Mini Mental State (MMSE) και η αξιολόγηση κλινικής άνοιας (CDR) είναι δύο από τις πιο συχνά χρησιμοποιούμενες εξετάσεις στην κλινική πράξη για την αξιολόγηση της νόσου [18].

Στα δεδομένα τη μελέτης αυτής, χρησιμοποιήθηκε η εξέταση MMSE. Το MMSE αποτελείται από μια σειρά κλινικών και ψυχομετρικών αξιολογήσεων μέσω νευροψυχολογικών τεστ, τα οποία αξιολογούν τις γλωσσικές και μνημονικές ικανότητες και την ικανότητα επίλυσης προβλημάτων. Η μέγιστη βαθμολογία MMSE είναι 30 βαθμοί. Μια βαθμολογία μικρότερη από 12 υποδηλώνει σοβαρή άνοια, 13 έως 20 συνιστά μέτρια άνοια, 20 έως 24 υποδηλώνει ήπια άνοια και 24 έως 30 αντιπροσωπεύει κανονική περίπτωση (NC) (βλ. Πίνακα 3.2) [18].

Δυστυχώς, η γνωστική αξιολόγηση δεν είναι αντικειμενική, καθώς μπορεί να επηρεαστεί από εξωτερικούς παράγοντες, όπως από την ψυχολογική κατάσταση του ασθενούς. Αυτό οδηγεί σε μεγάλη διακύμανση στον προσδιορισμό της άνοιας ειδικά των υποκειμένων MCI περιπτώσεων. Έτσι, η ερευνητική κοινότητα έχει οδηγηθεί στη αναζήτηση διαγνωστικών δεικτών απεικόνισης. Τα αναθεωρημένα κριτήρια για τη διάγνωση της νόσου προτάθηκαν το 2007 από την ομάδα εργασίας του Εθνικού Ινστιτούτου Νευρολογικών Διαταραχών και Εγκεφαλικού-Αλτσχάιμερ και Σχετικών Διαταραχών [19]. Σύμφωνα με αυτά τα κριτήρια, η κλινική αξιολόγηση πρέπει να περιλαμβάνει τουλάχιστον ένα υποστηρικτικό χαρακτηριστικό από τα ακόλουθα: (i) ατροφία μεσολαβητικού κροταφικού λοβού (MTL) όπως φαίνεται στη δομική μαγνητική τομογραφία, (ii) κροταφοβρεγματικός υπομεταβολισμός όπως φαίνεται στο PET, λόγω του ότι στο νευρωνικό θάνατο ο εγκέφαλος δεν απορροφά το 18F-FDG

ραδιοϊχνηλάτη, (iii) Θετικότητα στην απεικόνιση αμυλοειδούς όπως φαίνεται στο PET και iv)

Μη φυσιολογικοί νευρωνικοί δείκτες CSF.

Βαθμολογία	Στάδιο άνοιας	Communication skills / Impairment
24-30	Mild Cognitive Impairment (MCI)	Προβλήματα συγκέντρωσης/μειωμένο εύρος προσοχής. Αρχίζει να έχει δυσκολία εύρεσης λέξεων. Δεν θα αναγνωρίζονταν βλάβες σε μια προσωπική συνέντευξη Συνήθως γνωρίζουν προβλήματα και μπορεί να προσπαθήσουν να κρύψουν ή να αποζημιώσουν.
20-24	Mild	Μειωμένες οπτικές/χωρικές ικανότητες Ακατάλληλες κοινωνικές ενδείξεις (π.χ. να στέκεστε πολύ κοντά στο άτομο κατά τη διάρκεια της συνομιλίας). Δυσκολία εύρεσης λέξεων Χάνετε το συρμό της σκέψης στη συνομιλία, Επαναλαμβάνεται Συνήθως γνωρίζει προβλήματα και μπορεί να προσπαθήσει να κρύψει ή να αποζημιώσει.
13-20	Moderate	Δυσκολία παρακολούθησης συνομιλίας. Απώλεια λεξιλογίου, ειδικά ουσιαστικών. Περισσότερη δυσκολία εύρεσης λέξεων Αντικατάσταση λέξης ή δημιουργία νέων λέξεων Δυσκολία παρακολούθησης ιστορίας ή ταινίας. Κακή ανάκληση Δυσκολία ακολουθώντας τις οδηγίες. Τάση να μιλάμε για τίποτα ή να περιπλανιόμαστε.
0-12	Severe	Τάση για περιπλανήσεις ή επανάληψη λέξεων. Αυξανόμενη απώλεια λεξιλογίου Δύσκολο να ακολουθήσεις οτιδήποτε άλλο εκτός από απλή συνομιλία/ οδηγίες. Δεν είναι δυνατή η παρακολούθηση μιας ιστορίας ή μιας ταινίας Σημαντικές αλλαγές προσωπικότητας/ συμπεριφοράς. Τερματική άνοια Αδυναμία ομιλίας Δυσκολία κατανόησης όταν μιλούν Κυρίως μη λεκτική επικοινωνία

Πίνακας 3.2: Ερμηνευση του Γνωστικού τεστ (MMSE Score)

3.2.2 Ανάλυση και σύγκριση των δεδομένων που χρησιμοποιηθήκαν

Στο πίνακα 3.3 παρουσιάζονται όλα τα δεδομένα που συλλέχθηκαν από τις βάσεις ADNI και OASIS. Επίσης, για σκοπούς σύγκρισης παρουσιάζονται και τα δεδομένα από τη βάση ADNI που είχαν χρησιμοποιηθεί σε άλλη μελέτη [7]

Διάγνωση	Δεδομένα ADNI	Δεδομένα OASIS	Δεδομένα προηγ. μελέτης (ADNI)
Normal Control	194 (96 males and 98 females)	527 (226 males and 301 females)	153 (73 males and 80 females)
MCI	200 (127 males and 73 females)	44 (23 males and 21 females)	141 (95 males and 46 females)
MCIc	84 (49 males and 35 females)	---	77 (43 males and 34 females)
AD	130 (60 males and 70 females)	113 (54 males and 59 females)	84 (40 males and 44 females)
TOTAL	608 (332 males and 276 females)	684 (303 males and 381 females)	455 (251 males and 204 females)

Πίνακας 3.3: Περιγραφή των δεδομένων που χρησιμοποιήθηκαν στη μελέτη από τη βάση ADNI και από τη βάση της OASIS. Τα δεδομένα προηγούμενης μελέτης αφορούν τη μελέτη [7]

Σε κάποιες από τις πιο πάνω περιπτώσεις του πίνακα 3.3, κάποια δεδομένα ήταν ελλιπή και περιείχαν ερωτηματικό αντί κάποια τιμή, οπότε ανάλογα με τα δεδομένα που χρησιμοποιήθηκαν σε κάθε μοντέλο μας, οι ελάχιστες περιπτώσεις που είχαν ερωτηματικό στα δεδομένα μας αγνοούνταν. Στον Πίνακα 3.4 και 3.5 παρουσιάζονται τα βασικά δημογραφικά στοιχεία με βάση τα δεδομένα από τη βάση δεδομένων της ADNI και από τη βάση δεδομένων OASIS αντίστοιχα. Οι πίνακες συμπεριλαμβάνουν χαρακτηριστικά φύλου, ηλικίας, MMSE. Όλα τα βασικά χαρακτηριστικά (εκτός από την ηλικία) έχουν σημαντικές διαφορές όπως φαίνεται. Όπως αναμενόταν, οι ασθενείς με NA είχαν μικρότερους όγκους

από τα άτομα με MCI και τα άτομα με MCI είχαν μικρότερους όγκους από τα υγιή άτομα (NC).

Μέσος όρος (±Τυπική Απόκλιση)	NC	MCI	MCI_C	AD
Φύλο(Α/Θ)	96/98	127/73	49/35	60/70
Ηλικία	76.17 (5.20)	74.75 (7.18)	74.88 (7.30)	76.02 (7.34)
Γνωστικό τεστ (MMSE)	29.11 (1.0)	27.20 (1.82)	26.51 (1.84)	23.14 (2.16)
Όγκος Ιππόκαμπου (mm ³)	3539.24 (414.0)	3243.68 (461.83)	2941.59 (461.82)	2892.56 (474.24)
Όγκος Ενδοκρινικού φλοιού (mm ³)	1931.17 (284.0)	1719.38 (384.85)	1545.20 (338.96)	1418.09 (348.62)

Πίνακας 3.4: Μέσος όρος και τυπική απόκλιση δεδομένων από τη βάση ADNI για κάθε κατηγορία (Υγιή άτομα NC, Άτομα με MCI, Άτομα με MCIc, άτομα με AD)

Μέσος όρος (±Τυπική Απόκλιση)	NC	MCI	AD
Φύλο(Α/Θ)	226 /301	23 /21	54 /59
Ηλικία	66.74 (8.90)	71.71 (6.16)	74.22 (8.04)
Γνωστικό τεστ (MMSE)	29.05 (1.17)	27.42 (2.24)	24.06 (4.81)
Όγκος Ιππόκαμπου (mm ³)	3755.87 (509.65)	3471.08 (442.33)	2904.98 (546.31)
Όγκος Ενδοκρινικού φλοιού (mm ³)	1677.45 (320.85)	1605.52 (342.44)	1359.97 (353.12)

Πίνακας 3.5: Μέσος όρος και τυπική απόκλιση δεδομένων από τη βάση OASIS για κάθε κατηγορία (Υγιή άτομα NC, Άτομα με MCI, Άτομα με MCIc, άτομα με AD)

Στο πίνακα 3.6 και 3.7, παρουσιάζονται οι μέσες διαφορές στα χαρακτηριστικά υφής (texture features) για τα δεδομένα που αφορούν τις περιπτώσεις NC και AD για το μέρος του ιππόκαμπου και του ενδορινικού φλοιού αντίστοιχα.

Χαρακτηριστικό	Δεδομένα ADNI	Δεδομένα OASIS	Δεδομένα προηγ. μελέτης (ADNI Data)
ASM	-0.0135	-0.0135	-0.13
Contrast	-1.97	-1.97	-2.00
Corelation	0.00024	0.00024	0.0004
Variance	-2.40	-2.40	-2.40
Sum Average	1.65	1.65	1.64
Sum Variance	-6.30	-6.30	-6.3
Entropy	0.10	0.103	0.103
Cluster Shade	1009.72	1009.72	1009.00
Volume	646.68	646.68	646

Πίνακας 3.6: Μέσες διαφορές μεγεθών της αρχικής σάρωσης-εξέτασης του ιππόκαμπου για τις περιπτώσεις NC και AD (Mean differences at baseline scans for hippocampus NC Vs AD). Τα δεδομένα προηγούμενης μελέτης αφορούν τη μελέτη [7]

Χαρακτηριστικό	ADNI Data	Δεδομένα OASIS	Δεδομένα προηγ. μελέτης (ADNI Data)
ASM	-0.016	-0.016	-0.17
Contrast	-24.33	-24.33	-24.3
Corelation	0.052	0.052	0.051
Variance	0.20	0.20	0.201
Sum Average	1.44	1.44	1.44
Sum Variance	27.70	27.70	27.70
Entropy	0.137	0.137	0.137
Cluster Shade	471.34	471.34	471
Volume	513.10	513.10	513

Πίνακας 3.7: Μέσες διαφορές μεγεθών της αρχικής σάρωσης-εξέτασης του ενδορινικού φλοιού για τις περιπτώσεις NC και AD (Mean differences at baseline scans for entorhinal cortex NC Vs AD). Τα δεδομένα προηγούμενης μελέτης αφορούν τη μελέτη [7]

3.2.3 Επιλογή χαρακτηριστικών (Feature selection)

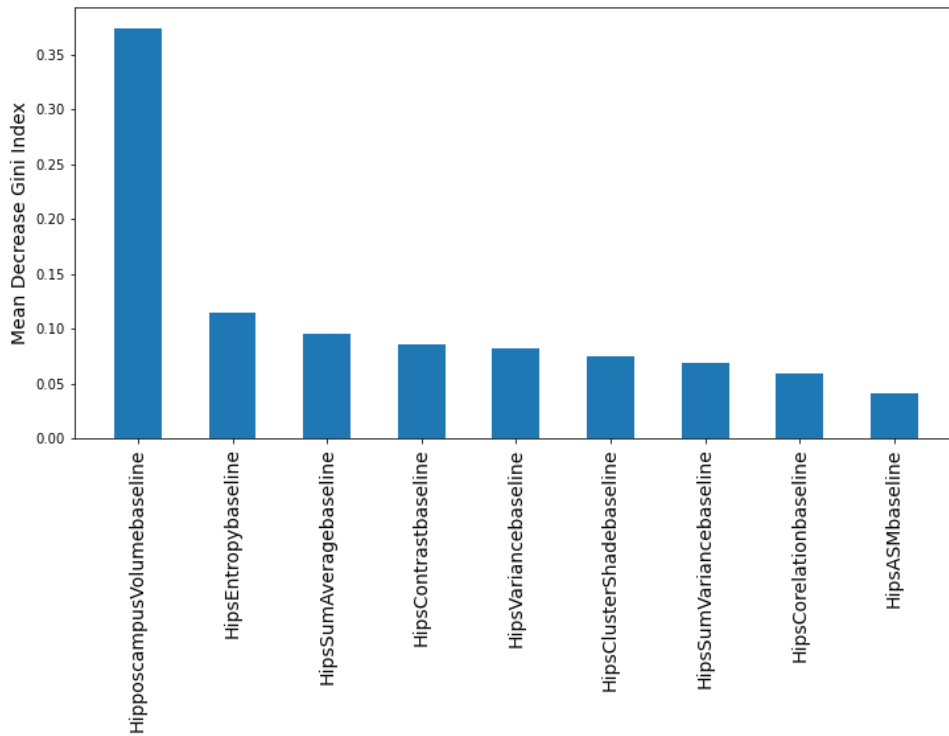
Για τον εντοπισμό και επιλογή των σημαντικότερων χαρακτηριστικών, τα οποία έχουν τη μεγαλύτερη επιρροή στο διαχωρισμό των κανονικών περιπτώσεων από τις περιπτώσεις Αλτσχάιμερ χρησιμοποιήθηκε ο δείκτης Decrease Gini index. Η μέση μείωση του συντελεστή Gini αφορά τα μοντέλα δέντρων αποφάσεων τα οποία περιγράφονται στο κεφάλαιο 4. Χρησιμοποιείται για να μετρήσει τον βαθμό που ένα συγκεκριμένο χαρακτηριστικό (ή μια συνάρτηση) μειώνει τον ακανόνιστη κατανομή των κλάσεων στα φύλλα του δέντρου. Όσο υψηλότερη είναι η τιμή της μέσης μείωσης της βαθμολογίας του συντελεστή Gini, τόσο μεγαλύτερη είναι η σημασία του χαρακτηριστικού στο μοντέλο που δημιουργείται για τη λύση του προβλήματος ταξινόμησης [20].

Στη περίπτωση χαρακτηριστικών υφής του ιππόκαμπου, παρατηρούμαι στο διάγραμμα 3.1 ότι το χαρακτηριστικό με τη μεγαλύτερη σημασία είναι η εντροπία ενώ αυτό με τη λιγότερη σημασία είναι το ASM (Angular Second Moment). Λαμβάνοντας υπόψη και τον όγκο του ιππόκαμπου παρατηρούμε ότι ο όγκος έχει με διαφορά μεγαλύτερη επιρροή στο μοντέλο σχέση με τα υπόλοιπα χαρακτηριστικά, όπως ήταν αναμενόμενο.

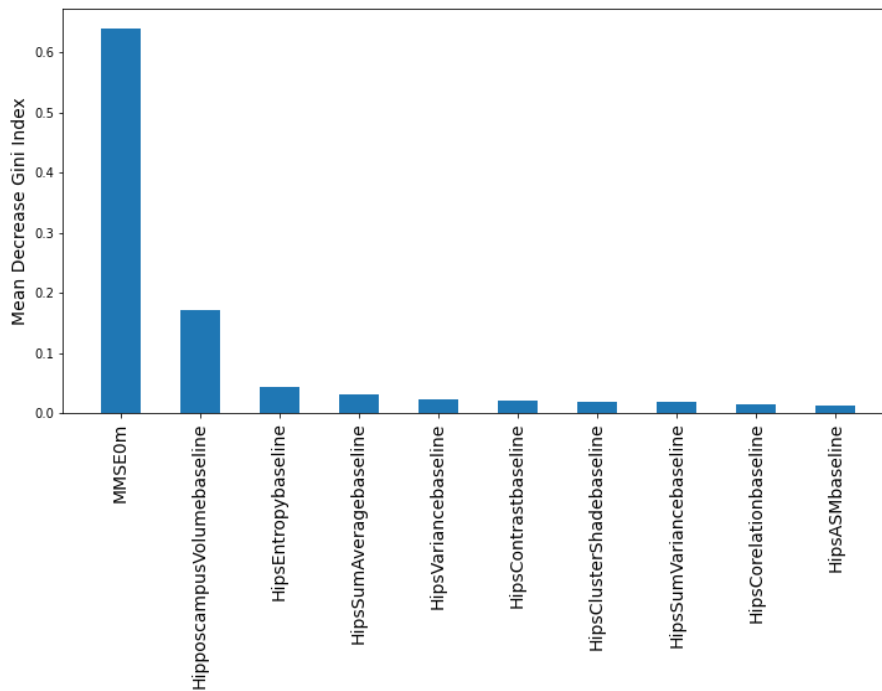
Στη περίπτωση χαρακτηριστικών υφής του ενδοκρινικού φλοιού, παρατηρούμαι στο διάγραμμα 3.3 ότι το χαρακτηριστικό με τη μεγαλύτερη σημασία είναι το πάχος (thickness) ενώ αυτό με τη λιγότερη σημασία είναι το ASM (Angular Second Moment). Λαμβάνοντας υπόψη και τον όγκο του ενδοκρινικού φλοιού παρατηρούμε και σε αυτή τη περίπτωση ότι ο όγκος έχει με διαφορά μεγαλύτερη επιρροή στο μοντέλο σχέση με τα υπόλοιπα χαρακτηριστικά, όπως ήταν αναμενόμενο.

Και στις 2 περιπτώσεις (χαρακτηριστικών ιππόκαμπου και χαρακτηριστικών ενδοκρινικού φλοιού) εάν ληφθεί υπόψη το γνωστικό τεστ MSSE score όπως φαίνεται στα διαγράμματα

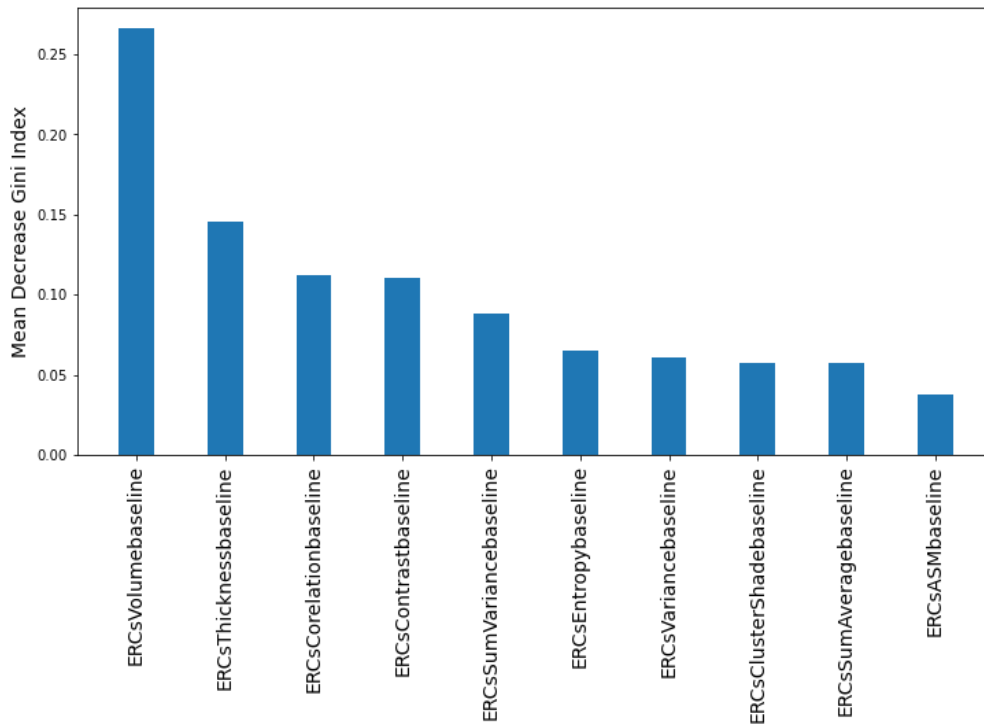
3.2 και 3.4, παρατηρούμε ότι αυτό έχει την μεγαλύτερη συμβολή με διαφορά από τα υπόλοιπα χαρακτηριστικά, με ποσοστό πάνω από 50%.



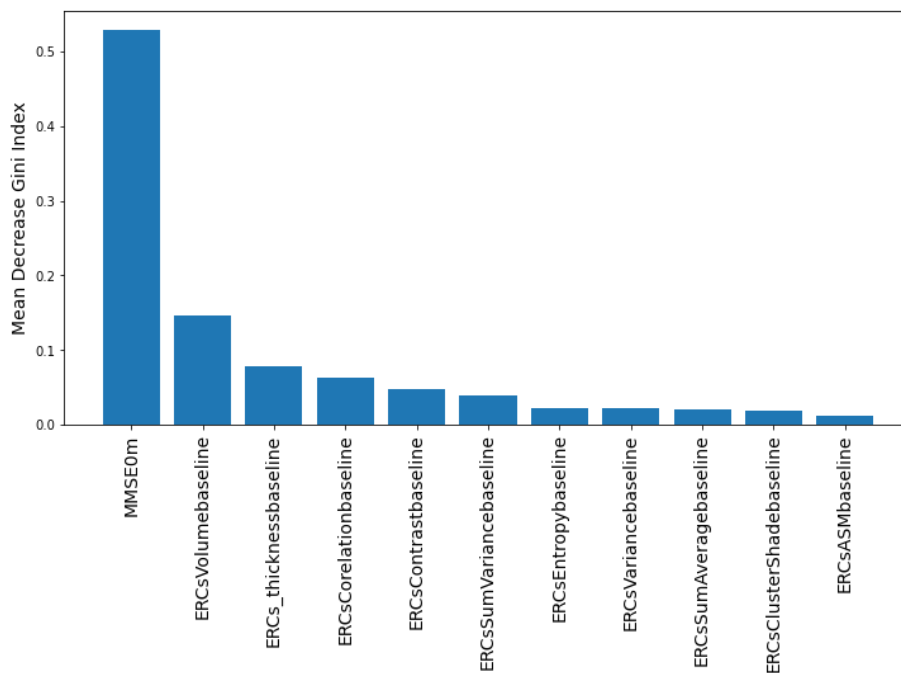
Διάγραμμα 3.1: Σημαντικότητα χαρακτηριστικών υφής και όγκου του ιππόκαμπου στο διαχωρισμό περιπτώσεων NC και AD από βάση δεδομένων ADNI



Διάγραμμα 3.2: Σημαντικότητα χαρακτηριστικών υφής του ιππόκαμπου και γνωστικού τεστ στο διαχωρισμό περιπτώσεων NC και AD από βάση δεδομένων ADNI



Διάγραμμα 3.3: Σημαντικότητα χαρακτηριστικών υφής και όγκου του ενδοκρινικού φλοιού στο διαχωρισμό περιπτώσεων NC και AD από βάση δεδομένων ADNI



Διάγραμμα 3.4: Σημαντικότητα χαρακτηριστικών υφής, όγκου του ενδοκρινικού φλοιού και γνωστικού τεστ στο διαχωρισμό περιπτώσεων NC και AD από βάση δεδομένων ADNI

3.2.4 Διαχωρισμός των δεδομένων για εκπαίδευση των μοντέλων μηχανικής μάθησης

Για τα μοντέλα μηχανικής μάθησης που χρησιμοποιήθηκαν με σκοπό τον διαχωρισμό των περιπτώσεων υγιή ατόμων από ατόμων με Νόσο Αλτσχάιμερ, τα δεδομένα χρησιμοποιήθηκαν ως εξής:

Αρχικά, χρησιμοποιήθηκαν μόνο τα δεδομένα ADNI, τα οποία χωρίστηκαν τυχαία σε σετ εκπαίδευσης ~80% των συμμετεχόντων και σε σετ αξιολόγησης (test set) ~20% των συμμετεχόντων. Στο σετ αξιολόγησης (test set), υπήρχε ίδιος αριθμός NC και NA όπως φαίνεται στο πίνακα 3.8.

Στη συνέχεια, δοκιμάστηκε μια άλλη περίπτωση, στην οποία χρησιμοποιήθηκαν σε συνδυασμό και τα δεδομένα από ADNI και από OASIS . Με τον ίδιο τρόπο (ποσοστό) διαχωρίστηκαν και αυτά τα δεδομένα φαίνεται στο πίνακα 3.9.

	Training δεδομένα	Test δεδομένα
NC	161	33
NA	97	33
Σύνολο	258 (80%)	66 (20%)

Πίνακας 3.8: Διαχωρισμός των δεδομένων βάσης ADNI σε train set για εκπαίδευση του μοντέλου και σε test set για αξιολόγηση του μοντέλου για το διαχωρισμό NC και NA περιπτώσεων

	Training δεδομένα	Test δεδομένα
NC	636	85
NA	158	85
Σύνολο	796 (80%)	170 (20%)

Πίνακας 3.9: Διαχωρισμός των δεδομένων βάσης ADNI και OASIS (συνδυαστικά) σε train set για εκπαίδευση του μοντέλου και σε test set για αξιολόγηση του μοντέλου για το διαχωρισμό NC και NA περιπτώσεων

ΚΕΦΑΛΑΙΟ 4: Μοντέλα μηχανικής μάθησης, μοντέλα επιχειρηματολογίας και μετρικές αξιολόγησης τους

4.1 Περιγραφή Μοντέλων μηχανικής μάθησης

Τα μοντέλα μηχανικής μάθησης που χρησιμοποιήθηκαν είναι τα Decision Trees, Random Forest και XGBoost (eXtreme Gradient Boosting) τα οποία βασίζονται σε δέντρα αποφάσεων και χρησιμοποιούνται σε προβλήματα ταξινόμησης. Ο σκοπός που χρησιμοποιήθηκαν αυτά τα μοντέλα ήταν για μπορούμε να κάνουμε εξαγωγή τους κανόνες από κανόνες από αυτά και αναπτύξουμε τα μοντέλα επιχειρηματολογίας. Επίσης, χρησιμοποιήθηκε το μοντέλο SVM (Support Vector Machine) το οποίο συμπεριλήφθηκε για σκοπούς σύγκρισης με τα υπόλοιπα μοντέλα.

Τα Δέντρα Αποφάσεων (DT) είναι μια μη παραμετρική εποπτευόμενη μέθοδος μάθησης που χρησιμοποιείται για ταξινόμηση και παρεμβολή. Ο στόχος είναι να δημιουργηθεί ένα μοντέλο που προβλέπει την τιμή μιας μεταβλητής στόχου μαθαίνοντας απλούς κανόνες απόφασης που συνάγονται από τα χαρακτηριστικά δεδομένων. Τα δέντρα αποφάσεων μαθαίνουν από δεδομένα και προσπαθούν για παράδειγμα να προσεγγίσουν μια καμπύλη ημιτόνου με ένα σύνολο κανόνων απόφασης if-then-else. Όσο πιο βαθιά είναι το δέντρο, τόσο πιο περίπλοκοι είναι οι κανόνες απόφασης και τόσο πιο κατάλληλο είναι το μοντέλο. Γενικά, είναι απλά στην κατανόηση και στην ερμηνεία [21].

Το μοντέλο Random Forest (τυχαίο δάσος) αποτελεί είδος μοντέλο συνόλου δέντρων απόφασης (ensemble tree) και είναι μια μέθοδος εκμάθησης συνόλου για ταξινόμηση, παρεμβολή και άλλες εργασίες που λειτουργεί κατασκευάζοντας ένα πλήθος δέντρων αποφάσεων κατά το χρόνο εκπαίδευσης. Ένα τυχαίο υποσύνολο των δεδομένων και ένα τυχαίο υποσύνολο χαρακτηριστικών χρησιμοποιούνται για την εκπαίδευση κάθε δέντρου

απόφασης στο Random Forest. Στη συνέχεια δημιουργείται η τελική πρόβλεψη συνδυάζοντας τις προβλέψεις όλων των δέντρων. Αυτό βοηθά στη μείωση της υπερπροσαρμογής (overfitting) και στη βελτίωση της ακρίβειας του μοντέλου. Τα Random Forest είναι γνωστά για την αξιοπιστία και την ικανότητά τους να χειρίζονται δεδομένα υψηλών διαστάσεων και χρησιμοποιούνται συχνά για εφαρμογές ταξινόμησης και παρεμβολή [22].

Το τρίτο μοντέλο μηχανικής μάθησης XGBoost (eXtreme Gradient Boosting) που πρόκειται να διερευνηθεί είναι μια δημοφιλής τεχνική μηχανικής μάθησης για εποπτευόμενες εργασίες εκμάθησης, ιδιαίτερα για ζητήματα ταξινόμησης και παρεμβολής. Για τη δημιουργία ενός ισχυρότερου μοντέλου, αυτό το μοντέλο ενώνει επαναληπτικά τα αδύναμα δέντρα αποφάσεων χρησιμοποιώντας ένα πλαίσιο ενίσχυσης κλίσης. Προκειμένου να μειωθεί η υπερπροσαρμογή και να ενισχυθεί η γενίκευση (generalization), το XGBoost ενσωματώνει επιπλέον τεχνικές κανονικοποίησης (regularization). Είναι πλέον μια τυπική προσέγγιση για πολλούς διαγωνισμούς επιστήμης δεδομένων και χρησιμοποιείται εκτενώς στις επιχειρήσεις για διάφορους σκοπούς, συμπεριλαμβανομένων συστημάτων συστάσεων, ανίχνευσης απάτης και πρόβλεψης εκτροπής πελατών [23].

Τα μοντέλα Support Vector Machines (SVM) είναι ένα σύνολο εποπτευόμενων μεθόδων εκμάθησης που χρησιμοποιούνται για ταξινόμηση, παρεμβολή και ανίχνευση ακραίων τιμών. Είναι αποτελεσματικά σε χώρο υψηλών διαστάσεων, όταν δηλαδή χρειάζεται να λάβουμε υπόψη πολλά χαρακτηριστικά (features) δεδομένων. Η δημοτικότητα των SVM πιθανότατα οφείλεται στην ευελιξία τους στη θεωρητική ανάλυση, στην ευελιξία τους να εφαρμόζονται σε μια μεγάλη ποικιλία εργασιών, συμπεριλαμβανομένων των προβλημάτων δομημένης πρόβλεψης [21].

4.2 Μοντέλα μηχανικής μάθησης με επιχειρηματολογία

4.2.1 Περιγραφή επιχειρηματολογίας

Η επιχειρηματολογία μπορεί να θεωρηθεί ως μια διαδικασία παραγωγής εξηγήσεων. Είναι μια μορφή προγραμματισμού που έχει δημιουργηθεί ως επέκταση του Λογικού Προγραμματισμού. Αναφέρεται στον προγραμματισμό της κοινής λογικής, όπου τα συμπεράσματα εξάγονται χρησιμοποιώντας λογικό συλλογισμό μέσω της χρήσης επιχειρημάτων, δηλαδή κανόνων που αντιπροσωπεύουν ισχυρισμούς που υποστηρίζουν ή αντιτίθενται. Δεδομένου ενός θέματος, η διαδικασία της επιχειρηματολογίας μπορεί να θεωρηθεί ως ο εντοπισμός σχετικών πληροφοριών και η δημιουργία μιας εξήγησης για το θέμα χρησιμοποιώντας επιχειρήματα. Ως εκ τούτου, θα μπορούσαμε να ορίσουμε την επιχειρηματολογία ως «το πλαίσιο πάνω στο οποίο δρουν τα επιχειρήματα. Βασίζεται στην ικανότητα του ανθρώπου να δημιουργεί επιχειρήματα για να υποστηρίξει απόψεις και θέσεις, να εκφράζει σκέψεις και να αλληλοεπιδρά με άλλα μέλη της κοινωνίας». Τα επιχειρήματα που παράγονται μέσω της μάθησης μπορούν να αξιολογηθούν σε σχέση με τα υπάρχοντα επιχειρήματα των ειδικών-επαγγελματιών (experts) για την καθοδήγηση της μαθησιακής διαδικασίας [24].

Ένα βασικό πλεονεκτήματα της επιχειρηματολογίας είναι φυσική αλληλεπίδραση με τον χρήστη ώστε να γνωρίζει για ποιους λόγους αποφασίστηκε το συγκεκριμένη απάντηση και να καταλάβει πως σκέφτεται το σύστημα. Επίσης, αναπαριστά τις γνώσεις που παίρνει το σύστημα μας καθιστώντας το πιο ευέλικτο και οδηγεί στην οργάνωση των πληροφοριών που μαθαίνονται σε συλλογισμούς ώστε να μπορούν να αξιοποιηθούν καλύτερα. Τέλος, η επιχειρηματολογία και η εξήγηση που προσφέρει είναι απαραίτητη να υπάρχει στα συστήματα τεχνητής νοημοσύνης (T.N.) λόγω ηθικών λόγων.

Τα επιχειρήματα εμφανίστηκαν πρόσφατα στον τομέα της Ιατρικής Πληροφορικής, κυρίως για να υποστηρίξουν την απόφαση των συστημάτων τεχνητής νοημοσύνης. Για παράδειγμα, το Medica είναι ένα σύστημα επιχειρηματολογίας που αναπτύχθηκε στο πλαίσιο Gorgias, για να αποφασίσει εάν ένα άτομο μπορεί να έχει πρόσβαση σε ευαίσθητους ιατρικούς φακέλους. Άλλα παραδείγματα περιλαμβάνουν μια προσέγγιση Defeasible Reasoning και Argumentation Theory για τη λήψη αποφάσεων σε ένα σενάριο υγειονομικής περίθαλψης για τη διάγνωση του καρκίνου του μαστού. Οι επιχειρηματολογικές αποφάσεις είναι από τη φύση τους εξηγήσιμες στον άνθρωπο – αποτελούμενες από περιγραφές σε συνδυασμό με το ίχνος λογικής του συστήματος, δηλαδή κλινικά στοιχεία και ιατρικά αρχεία που είναι αποδεκτά ως επιχειρήματα. Έτσι, επιχειρηματολογία αποτελεί ένα από τους κύριους μηχανισμούς όταν πρόκειται για την εξήγηση της απόφασης ενός συστήματος τεχνητής νοημοσύνης [25].

4.2.2 Περιγραφή εργαλείου Gorgias και εργαλείου επιχειρηματολογίας ArgEML

Το εργαλείο Gorgias είναι ένα πλαίσιο επιχειρηματολογίας που συνδυάζει με επιτυχία τις αρχές της συλλογιστικής προτίμησης (reasoning) και της απαγωγής (abduction), διατηρώντας παράλληλα τα οφέλη και των δύο. Προσφέρει τη δυνατότητα δημιουργίας συλλογισμών, χρησιμοποιώντας προσαρμοστικές πολιτικές προτιμήσεων, ακόμη και με απουσία πλήρους πληροφοριών. Μπορεί να μοντελοποιήσει "αρχές υψηλότερης τάξης", λαμβάνοντας υπόψη τις προσωπικές προτιμήσεις ή ρυθμίσεις ενός χρήστη στη διαδικασία λήψης αποφάσεων [26].

Το εργαλείο επιχειρηματολογίας ArgEML είναι βασισμένο στο εργαλείο Gorgias και υποκινείται από πολλά από τα προαναφερθέντα στοιχεία επιχειρηματολογίας, στοχεύοντας στη σύνθεση διαφορετικών στοιχείων από τη σύνδεση μεταξύ επιχειρηματολογίας και

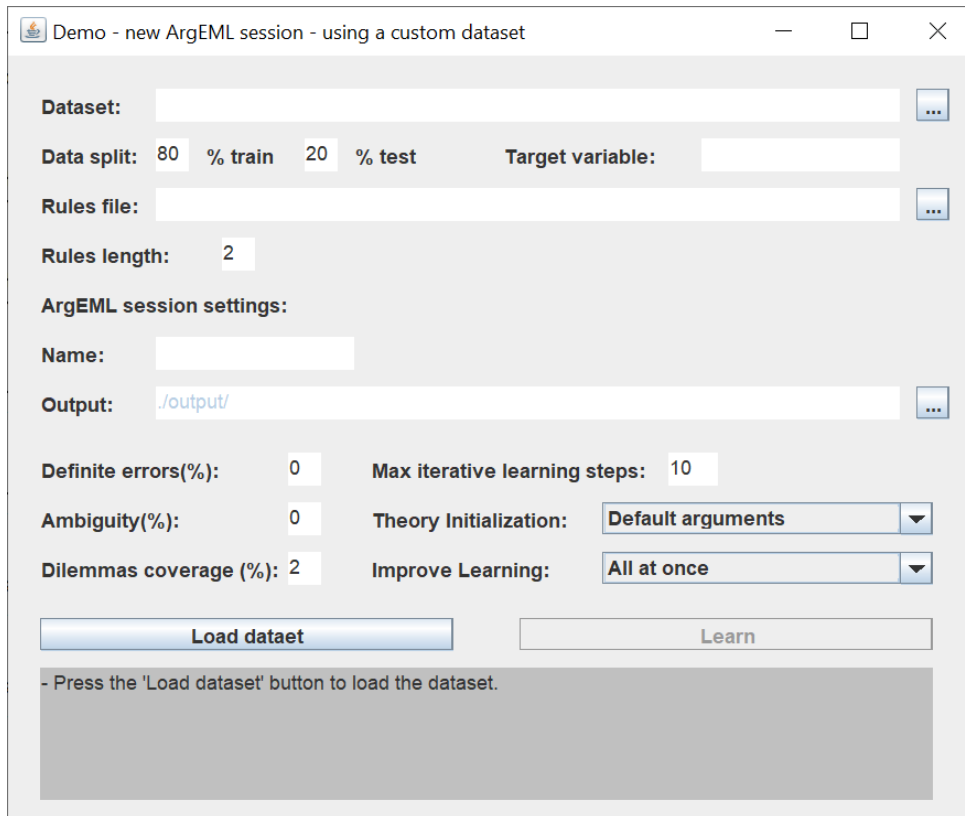
μάθησης. Στοχεύει στην παροχή μιας ολοκληρωμένης προσέγγισης όπου αξιοποιούνται αρκετοί από τους δεσμούς επιχειρηματολογίας και μάθησης, όπως ο τρόπος με τον οποίο η επιχειρηματολογία ως απορριπτική μορφή συλλογισμού παρακινεί νέες μετρήσεις αξιολόγησης για τη μαθησιακή διαδικασία και πώς η επιχειρηματολογία μπορεί να παρέχει ένα πλαίσιο για ante-hoc ή post-hoc επεξηγησιμότητα για τις μηχανές μάθησης [25].

Για να ξεκινήσει το εργαλείο να αναπτύσσει τη θεωρία του, χρειάζεται αρχικά να του δώσουμε όλα τα δεδομένα. Επίσης, χρειάζεται να ρυθμιστούν κατάλληλα οι παράμετροι (όρια) όπως φαίνεται στην εικόνα 4.1. Επίσης, στο εργαλείο μπορούμε να δώσουμε και κανόνες που παρήγαγε το μοντέλο μηχανικής μάθησης που υποστηρίζουν τις 2 κλάσεις. Έτσι το εργαλείο, με βάση τα προαναφερθέντα, ξεκινά να αρχικοποιεί μια θεωρία και μετά κάθε φορά βασισμένο στη θεωρία που αναπτύχθηκε, δημιουργεί μια νέα θεωρία προσπαθώντας να βελτιώσει την ακρίβεια και να μειώσει τον αριθμό των διλημάτων με βάση τα όρια που του δώσαμε [25].

Το εργαλείο τρέχει μέχρι να αναπτύξει το μοντέλο με τη θεωρία που θα έχει τα επιθυμητά αποτελέσματα με βάση τα όρια σύγκλισης που του θέσαμε στη αρχή. Αν δεν καταφέρει να συγκλίνει στα όρια που του θέσαμε, τότε ο αλγόριθμος θα σταματήσει να αναπτύσσει τη θεωρία μετά από ένα αριθμό βημάτων ('maximum iterative steps') που του θέσαμε. Τα επιτρεπτά όρια που του θέτουμε είναι το 'Definite errors' που αφορά το μέγιστο ποσοστό των λανθασμένων προβλέψεων και το 'Ambiguity' το οποίο αφορά το μέγιστο ποσοστό των διλημάτων.

Η παράμετρος 'Dilemmas coverage' αφορά το κάθε τύπο διλήματος ξεχωριστά. Για παράδειγμα, έστω ότι σε 10% των δεδομένων, δίνεται ως απάντηση το ίδιο δίλημμα με την ίδια εξήγηση. Τότε αν εμείς θέσαμε όριο 5% τη παράμετρο 'Dilemmas coverage', τότε ο

αλγόριθμος θα συνεχίζει να αναπτύσσει τη θεωρία μέχρι το συγκεκριμένο δίλημμα, να δίνεται ως απάντηση σε λιγότερο από 5% των προβλέψεων των δεδομένων μας.



Εικόνα 4.1: Αρχικοποίηση παραμέτρων εργαλείου επιχειρηματολογίας ArgEML

4.3 Περιγραφή μετρικών αξιολόγησης μοντέλων

Οι μετρικές αξιολόγησης που χρησιμοποιήθηκαν για τα όλα τα μοντέλα είναι το ROC-AUC, Accuracy, Sensitivity (Recall) και το Specificity, τα οποία περιγράφονται αναλυτικά στη συνέχεια [21]. Επίσης, για τα μοντέλα επιχειρηματολογίας λήφθηκε υπόψη και μια άλλη μετρική ακρίβειας (LA) που εξηγείται στη συνέχεια.

ROC (Receiver operating characteristic) είναι μια μέτρηση που δείχνει την απόδοση ενός μοντέλου δυαδικής ταξινόμησης μετρώντας την περιοχή κάτω από την καμπύλη που σχηματίζεται από τη γραφική παράσταση του πραγματικού θετικού ρυθμού (sensitivity)

έναντι του ψευδώς θετικού ρυθμού, σε διαφορετικές τιμές ορίων. Μια υψηλότερη βαθμολογία AUC-ROC (πιο κοντά στο 1) υποδηλώνει καλύτερη διάκριση μεταξύ θετικών και αρνητικών κατηγοριών.

$$ROC = \frac{TPR}{FPR} = \frac{\frac{True\ Positives(TP)}{True\ Positives(TP) + False\ Negatives\ (FN)}}{\frac{False\ Positive(FP)}{False\ Positive(FP) + True\ Negative(TN)}}$$

Το Accuracy στην ταξινόμηση πολλαπλών ετικετών, δείχνει ποσό σωστά προβλέπει το μοντέλο μας. Αντιστοιχεί στη πιθανότητα το μοντέλο μας να προβλέψει σωστά τη κατάσταση ενός ατόμου. Όσο πιο υψηλή τιμή τόσο πιο πολλές είναι οι σωστές προβλέψεις του μοντέλου μας.

$$Accuracy = \frac{True\ Positives(TP) + True\ Negative(TN)}{no\ of\ positive\ cases + no\ of\ negative\ cases}$$

Sensitivity (Recall) είναι η ικανότητα του ταξινομητή να βρίσκει προβλέπει τα θετικά δείγματα, δηλαδή αντιστοιχεί στη πιθανότητα το μοντέλο να προβλέπει θετικά ένα άτομο, υπό τον όρο ότι το άτομο είναι πραγματικά θετικό.

$$Sensitivity = \frac{True\ Positives(TP)}{no\ of\ Positive\ cases(P)}$$

Specificity είναι η ικανότητα του ταξινομητή να βρίσκει προβλέπει τα αρνητικά δείγματα, δηλαδή αντιστοιχεί στη πιθανότητα το μοντέλο να προβλέπει αρνητικά ένα άτομο, υπό τον όρο ότι το άτομο είναι πραγματικά αρνητικό.

$$Specificity = \frac{True\ Negative(TN)}{no\ of\ Negative\ cases(N)}$$

Στη συνέχεια, παρουσιάζεται μια μετρική ακρίβειας που χρησιμοποιείται στα μοντέλα επιχειρηματολογίας και η οποία εκτός από τις σωστές προβλέψεις, λαμβάνει υπόψη και τα διλήμματα που δίνονται από το μοντέλο [25].

$$LA = \frac{\#Definite\ Correct\ Predictions + \#dilemmas * w_a}{Total\ number\ of\ predictions}$$

Όπου το w_a είναι ίσο με 0.5 για δυαδική ταξινόμηση (binary classification)

ΚΕΦΑΛΑΙΟ 5: Ανάπτυξη και Αποτελέσματα Μοντέλων μηχανικής μάθησης χρησιμοποιώντας δεδομένα από βάση ADNI και βάση OASIS

Σε αυτό το κεφάλαιο, αναπτύχθηκαν τα μοντέλα μηχανικής μάθησης και τα μοντέλα επιχειρηματολογίας για διάφορα χαρακτηριστικά του ανθρώπινου εγκεφάλου. Παρουσιάζονται τα αποτελέσματα και γίνεται σύγκριση τους.

5.1. Μεθοδολογία ανάπτυξης μοντέλων μηχανικής μάθησης

Όπως προαναφέρθηκε τα δεδομένα χωρίστηκαν τυχαία σε σετ εκπαίδευσης (training set) 80% και σε σετ αξιολόγησης (test set) 20%. Ο διαχωρισμός των δεδομένων σε train και σε test σετ επαναλήφθηκε 5 φορές για να παρέχει 5 μοναδικές ομάδες εκπαίδευσης και αξιολόγησης του μοντέλου μας. Αυτή η τεχνική στη μηχανική μάθηση είναι γνωστή ως cross validation. Ο κύριος λόγος για τη χρήση της cross validation τεχνικής ήταν να αξιολογήσουμε την ικανότητα γενίκευσης του μοντέλου μας και να λάβουμε μια πιο ακριβή εκτίμηση για το πόσο καλά θα αποδώσει το μοντέλο σε νέα δεδομένα, επειδή το μοντέλο δοκιμάζεται σε πολλά διαφορετικά υποσύνολα δεδομένων.

Σε κάθε επανάληψη (fold) εκπαίδευσης και αξιολόγηση του μοντέλου γίνονταν parametrized tuning για τη εύρεση των κατάλληλων παραμέτρων (estimators και max_depth) του μοντέλου μηχανικής μάθησης.

Τα δεδομένα από βάση OASIS ήταν ανισόρροπα (imbalanced dataset) αφού το 83% των δεδομένων ήταν αρνητικές περιπτώσεις (περιπτώσεις Normal Control). Για να διορθωθεί αυτό το πρόβλημα, χρησιμοποιήθηκε η τεχνική η τεχνική oversampling/undersampling.

Τα μοντέλα μηχανικής μάθησης που χρησιμοποιήθηκαν ήταν το Decision Trees, Random Forest, XGBoost και SVM.

5.2 Αποτελέσματα μοντέλων μηχανικής μάθησης για διαχωρισμό NC και NA περιπτώσεων

Στους πίνακες παρουσιάζονται τα αποτελέσματα των μοντέλων που χρησιμοποιήθηκαν. Ο κάθε πίνακας αποτελεσμάτων αφορά διαφορετικό υποσύνολο επιλεγμένων χαρακτηριστικών που χρησιμοποιήθηκαν. Επίσης, στο κάθε πίνακα παρουσιάζονται τα αποτελέσματα που προκύψαν από τα δεδομένα βάσης ADNI και αποτελέσματα που προκύψαν από τη σύνθεση των δεδομένων βάσης ADNI και OASIS. Για σκοπούς σύγκρισης σε κάθε πίνακα παρουσιάζονται και τα αποτελέσματα από παρελθοντικές μελέτες που είχαν γίνει, τα οποία αφορούν αποτελέσματα μοντέλων που χρησιμοποιούσαν δεδομένα από τη βάση ADNI.

Επιλεγμένα Χαρακτηριστικά: Ηλικία, Φύλο, Hippocampus Volume				
Μοντέλο	ADNI δεδομένα			
	AUC	Accuracy	Sensitivity	Specificity
Decision Trees	0.77	0.74	0.62	0.87
Random Forest	0.78	0.78	0.69	0.88
XGBoost	0.77	0.77	0.70	0.84
SVM	0.81	0.79	0.69	0.89
Αποτελέσματα μελέτης [7] Logistic Regression	0.869	-	-	-
Αποτελέσματα Άρθρου Juottonen [27]	-	0.86	0.80	0.91
ADNI-OASIS δεδομένα				
	AUC	Accuracy	Sensitivity	Specificity
Random Forest	0.75	0.75	0.83	0.67

Πίνακας 5.1: Αποτελέσματα των μοντέλων παρούσας μελέτης για το διαχωρισμό NC και NA περιπτώσεων. Παρουσιάζονται τα αντίστοιχα αποτελέσματα των μελετών [7],[27] για σκοπούς σύγκρισης. Επιλεγμένα χαρακτηριστικά ήταν: Ηλικία, Φύλο, Όγκος ιππόκαμπου

Επιλεγμένα Χαρακτηριστικά: Ηλικία, Φύλο, HipsScampusVolumebaseline, HipsASMBaseline, HipsContrastbaseline, HipsCorelationbaseline, HipsVariancebaseline, HipsSumAveragebaseline, HipsSumVariancebaseline, HipsEntropybaseline, HipsClusterShadebaseline				
Μοντέλο	ADNI δεδομένα			
	AUC	Accuracy	Sensitivity	Specificity
Decision Trees	0.76	0.74	0.63	0.85
Random Forest	0.79	0.79	0.70	0.89
XGBoost	0.77	0.77	0.71	0.82
SVM	0.78	0.77	0.75	0.80
Άρθρο [10] Decision Trees	-	0.77	0.66	0.88
ADNI-OASIS δεδομένα				
Random Forest	0.74	0.74	0.87	0.60

Πίνακας 5.2: Αποτελέσματα των μοντέλων παρούσας μελέτης για το διαχωρισμό NC και NA περιπτώσεων. Παρουσιάζονται τα αντίστοιχα αποτελέσματα της μελέτης [10] για σκοπούς σύγκρισης. Επιλεγμένα χαρακτηριστικά ήταν: Ηλικία, Φύλο, Όγκος ιππόκαμπου και τα χαρακτηριστικά υφής ιππόκαμπου

Επιλεγμένα Χαρακτηριστικά: Ηλικία, Φύλο, ERC Volume				
Μοντέλο	ADNI δεδομένα			
	AUC	Accuracy	Sensitivity	Specificity
Decision Trees	0.81	0.79	0.69	0.88
Random Forest	0.80	0.80	0.73	0.87
XGBoost	0.74	0.74	0.64	0.84
SVM	0.82	0.80	0.70	0.91
Αποτελέσματα μελέτης [7] Logistic Regression	0.888	-	-	-
Αποτελέσματα Άρθρου Juottonen [27]	-	0.87	0.80	0.94
ADNI-OASIS δεδομένα				
Random Forest	0.68	0.68	0.76	0.61

Πίνακας 5.3: Αποτελέσματα των μοντέλων παρούσας μελέτης για το διαχωρισμό NC και NA περιπτώσεων. Παρουσιάζονται τα αντίστοιχα αποτελέσματα των μελετών [7],[27] για σκοπούς σύγκρισης. Επιλεγμένα χαρακτηριστικά ήταν: Ηλικία, Φύλο, Όγκος ενδοκρινικού φλοιού

Επιλεγμένα Χαρακτηριστικά: Ηλικία, Φύλο, ERCsASMbaseline, ERCsCorelationbaseline, ERCsVariancebaseline, ERCsSumAveragebaseline, ERCsClusterShadebaseline				
Μοντέλο	ADNI δεδομένα			
	AUC	Accuracy	Sensitivity	Specificity
Decision Trees	0.75	0.74	0.68	0.80
Random Forest	0.75	0.75	0.67	0.83
XGBoost	0.74	0.74	0.66	0.82
SVM	0.78	0.77	0.71	0.82
Αποτελέσματα μελέτης [7] Logistic Regression	0.872	-	-	-
ADNI-OASIS δεδομένα				
Random Forest	0.70	0.70	0.71	0.68

Πίνακας 5.4: Αποτελέσματα των μοντέλων παρούσας μελέτης για το διαχωρισμό NC και NA περιπτώσεων. Παρουσιάζονται τα αντίστοιχα αποτελέσματα της μελέτης [7] για σκοπούς σύγκρισης. Επιλεγμένα χαρακτηριστικά ήταν: Ηλικία, Φύλο και τα χαρακτηριστικά υψής ενδοκρινικού φλοιού

Επιλεγμένα Χαρακτηριστικά: Ηλικία, Φύλο, ERCsASMbaseline, ERCsCorelationbaseline, ERCsVariancebaseline, ERCsSumAveragebaseline, ERCsClusterShadebaseline, ERCsVolumebaseline				
Μοντέλο	ADNI δεδομένα			
	AUC	Accuracy	Sensitivity	Specificity
Decision Trees	0.78	0.76	0.67	0.86
Random Forest	0.81	0.81	0.72	0.90
XGBoost	0.79	0.79	0.75	0.84
SVM	0.86	0.85	0.82	0.89
Αποτελέσματα μελέτης [7] Logistic Regression	0.914	-	-	-
ADNI-OASIS δεδομένα				
Random Forest	0.68	0.68	0.84	0.52

Πίνακας 5.5: Αποτελέσματα των μοντέλων παρούσας μελέτης για το διαχωρισμό NC και NA περιπτώσεων. Παρουσιάζονται τα αντίστοιχα αποτελέσματα της μελέτης [7] για σκοπούς σύγκρισης. Επιλεγμένα χαρακτηριστικά ήταν: Ηλικία, Φύλο, Όγκος και χαρακτηριστικά υψής ενδοκρινικού φλοιού

Επιλεγμένα Χαρακτηριστικά: Ηλικία, Φύλο, Hippocampus Volume, ERC Volume, MMSE				
Μοντέλο	ADNI δεδομένα			
	AUC	Accuracy	Sensitivity	Specificity
Decision Trees	0.98	0.98	0.98	0.97
Random Forest	0.98	0.98	0.98	0.97
XGBoost	0.97	0.97	0.97	0.98
SVM	0.98	0.98	0.97	0.98
ADNI-OASIS δεδομένα				
Random Forest	0.89	0.89	0.91	0.87

Πίνακας 5.6: Αποτελέσματα των μοντέλων παρούσας μελέτης για το διαχωρισμό NC και NA περιπτώσεων. Επιλεγμένα χαρακτηριστικά ήταν: Ηλικία, Φύλο, Όγκος ενδοκρινικού φλοιού και γνωστικό σκορ (MSSE)

Επιλεγμένα Χαρακτηριστικά: Ηλικία, Φύλο, Hippocampus Volume, ERC Volume				
Μοντέλο	ADNI δεδομένα			
	AUC	Accuracy	Sensitivity	Specificity
Decision Trees	0.80	0.78	0.65	0.90
Random Forest	0.85	0.85	0.79	0.92
XGBoost	0.82	0.82	0.77	0.87
SVM	0.86	0.85	0.80	0.90
ADNI-OASIS δεδομένα				
Random Forest	0.75	0.75	0.83	0.67

Πίνακας 5.7: Αποτελέσματα των μοντέλων παρούσας μελέτης για το διαχωρισμό NC και NA περιπτώσεων. Επιλεγμένα χαρακτηριστικά ήταν: Ηλικία, Φύλο, Όγκος ιππόκαμπου, Όγκος ενδοκρινικού φλοιού

5.3 Σύγκριση αποτελεσμάτων μοντέλων μηχανικής μάθησης

Από τα πιο πάνω αποτελέσματα επιβεβαιώνεται ότι τα χαρακτηριστικά που μας δίνουν το καλύτερο μοντέλο για τη πρόβλεψη της νόσου είναι ο όγκος του ιππόκαμπου, ο όγκος του

ενδορινικού φλοιού και το γνωστικό τεστ (MMSE Score). Το γνωστικό τεστ είναι αυτό που μπορεί να ανεβάσει τη απόδοση του μοντέλου μας κατά 12%, δίνοντας ένα πολύ αποδοτικό μοντέλο ακρίβειας 98%.

Όσον αφορά τα μοντέλα που χρησιμοποιήθηκαν, παρατηρείται ότι τα μοντέλα SVM και Random Forest είναι αυτά που δίνουν τα καλύτερα αποτελέσματα στις πλείστες περιπτώσεις.

Σχετικά με τα δεδομένα που χρησιμοποιήθηκαν, παρατηρείται ότι τα δεδομένα από τη βάση ADNI δίνουν καλύτερα αποτελέσματα από ότι αν χρησιμοποιηθούν σε συνδυασμό με τα δεδομένα από τη βάση OASIS. Αυτό οφείλεται κυρίως στο ότι τα δεδομένα από τη βάση OASIS δεν ήταν τόσο ποιοτικά όπως θα εξηγηθεί και στη συνέχεια. Ωστόσο, τα μοντέλα που χρησιμοποίησαν και αυτά τα δεδομένα ενδέχεται να είναι πιο γενικευμένα.

Συγκρίνοντας τα αποτελέσματα των μοντέλων της μελέτης αυτής σε σχέση με τα αποτελέσματα που έδωσαν άλλες μελέτες, παρουσιάζεται μια μικρή διαφορά στην απόδοση των μοντέλων. Συγκεκριμένα, στη μελέτη αυτή τα αποτελέσματα δείχνουν μικρότερη απόδοση των μοντέλων από ότι σε άλλες μελέτες. Αυτό, οφείλεται στο γεγονός, ότι χρησιμοποιήθηκαν διαφορετικά δεδομένα στις άλλες μελέτες καθώς και το γεγονός ότι χρησιμοποιήθηκαν διαφορετικές τεχνικές ή μοντέλα ανάπτυξης των μοντέλων μηχανικής μάθησης. Στη μελέτη αυτή τα μοντέλα που χρησιμοποιήθηκαν ήταν κυρίως δέντρα αποφάσεων αφού ο σκοπός ήταν να επεκταθούν αυτά τα μοντέλα για να δοθεί εξήγηση.

ΚΕΦΑΛΑΙΟ 6: Υλοποίηση και αποτελεσμάτα μοντέλων επιχειρηματολογίας χρησιμοποιώντας το εργαλείο Gorgias ArgEML

Σε αυτό το κεφάλαιο εξηγείται η διαδικασία που ακολουθήθηκε για ανάπτυξη και υλοποίηση των μοντέλων επιχειρηματολογίας τα οποία εκτός από τη πρόβλεψη που παρέχουν στη κάθε περίπτωση, παρέχουν επίσης και εξήγηση.

6.1 Μεθοδολογία ανάπτυξης μοντέλων επιχειρηματολογίας χρησιμοποιώντας το εργαλείο Gorgias ArgEML

6.1.1 Διαδικασία ανάπτυξης μοντέλου μηχανικής μάθησης με επιχειρηματολογία παρέχοντας εξήγηση

Για τη ανάπτυξη του μοντέλου με εξήγηση αρχικά χρησιμοποιήθηκε ένα μοντέλο μηχανικής μάθησης (π.χ. Random Forest, XGB boost). Για τη επιλογή των τεχνικών παραμέτρων (estimators, max depth) του μοντέλου μηχανικής μάθησης, η μια προσέγγιση ήταν να γίνει μέσω παραμετροποίησης parametrized tuning. Η δεύτερη προσέγγιση, ήταν να χρησιμοποιηθούν μικρές παράμετροι, ώστε από το TE2Rules να παράγονται πιο απλοποιημένοι κανόνες (ο λόγος θα εξηγηθεί στη συνέχεια). Είναι σημαντικό να υπάρχει μια ισορροπία μεταξύ των δυο παραγόντων που αναφέρθηκαν, ώστε οι κανόνες που θα παραχθούν να είναι σχετικά απλοί αλλά ταυτόχρονα να υπάρχει και ικανοποιητική ακρίβεια ώστε να αντιπροσωπεύουν σωστά το σετ των δεδομένων μας.

Στη συνέχεια, από αυτό το μοντέλο μηχανικής μάθησης που αναπτύχθηκε, παράχθηκαν κανόνες μέσω του εργαλείου TE2Rules. Το εργαλείο TE2Rules παράγει κανόνες οι οποίοι αντιπροσωπεύουν μόνο τις θετικές περιπτώσεις. Για το λόγο αυτό, η διαδικασία που προαναφέρθηκε έγινε και 2^η φορά ώστε να παραχθούν κανόνες και για τις περιπτώσεις που

το αποτέλεσμα είναι αρνητικό. Αυτό επιτεύχθηκε αντιστρέφοντας τις τιμές της στήλης διάγνωσης από 0 σε 1 και αντίστροφα.

Αφού έγινε η εξαγωγή των κανόνων και για τις θετικές περιπτώσεις (περιπτώσεις θετικές σε Alzheimer) και για τις αρνητικές περιπτώσεις (περιπτώσεις αρνητικές σε Alzheimer), έγινε η αξιολόγηση των κανόνων. Με βάση την αξιολόγηση των κανόνων, επιλέχθηκε ένα υποσύνολο των κανόνων που παρήχθησαν από το εργαλείο TE2Rules. Στη συνέχεια, έγινε η εισαγωγή αυτού του υποσυνόλου των κανόνων στο εργαλείο ArgEML μαζί με τα αντίστοιχα δεδομένα.

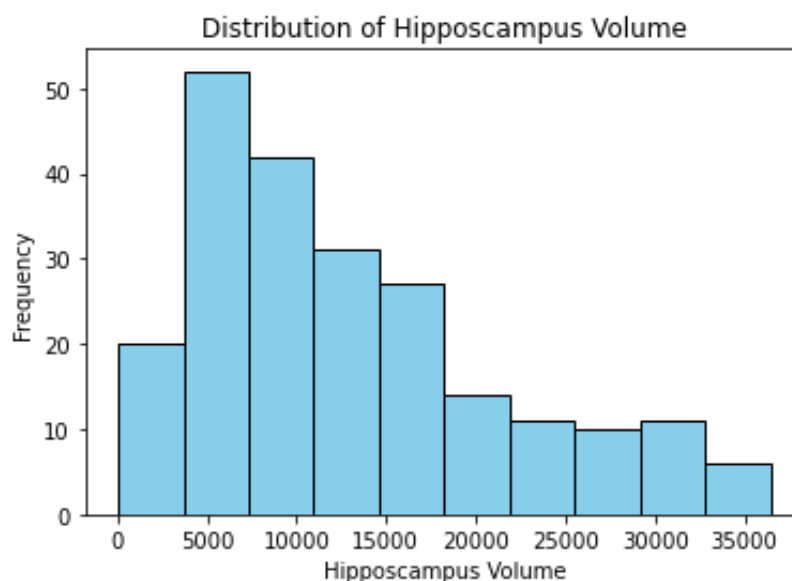
Τέλος, αφού επιλέχθηκαν οι κατάλληλες παράμετροι υπολογισμού στο εργαλείο ArgEML Gorgias, το πρόγραμμα εκτελέστηκε για να αναπτύξει το μοντέλο επιχειρηματολογίας με επεξήγηση και να δώσει τα αποτελέσματα.

6.1.2 Μετατροπή δεδομένων από συνεχή σε κατηγορικά δεδομένα

Το εργαλείο ArgEML χρησιμοποιεί κατηγορικά δεδομένα για τη λειτουργία του. Για το λόγο αυτό, τα δεδομένα που χρησιμοποιήθηκαν σε αυτή την έρευνα, τα οποία είχαν αρχικά συνεχή τιμές μετατράπηκαν σε κατηγορικά δεδομένα.

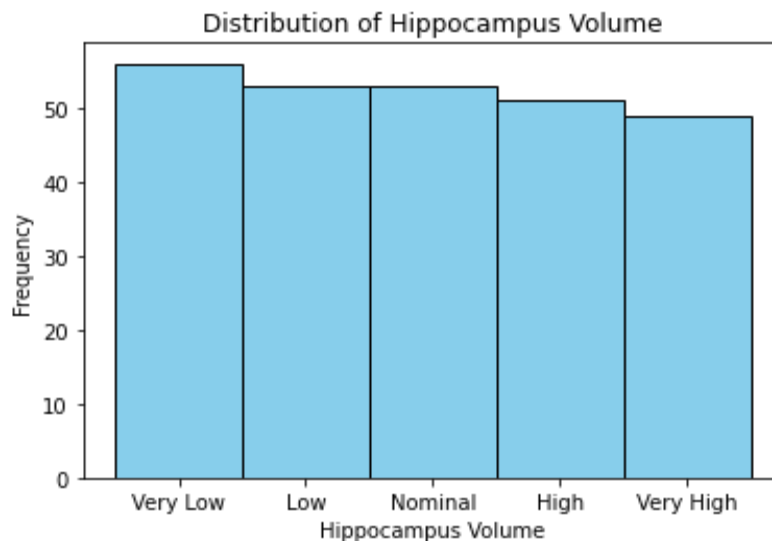
Είναι σημαντικός ο διαχωρισμός των δεδομένων μας να γίνει κατάλληλα σε κατηγορίες. Εάν ο διαχωρισμός των δεδομένων γίνει σε μικρότερο αριθμό κατηγοριών από ότι πρέπει ενδέχεται να χαθεί μεγάλο μέρος της πληροφορίας και να έχουμε υπεραπλουστεύσεις. Αντίθετα, αν ο διαχωρισμός των δεδομένων γίνει σε πιο πολλές κατηγορίες από ότι χρειάζεται, αυτό μπορεί να οδηγήσει σε υπερπροσαρμογή (overfitting) στα δεδομένα μας, οδηγώντας και πάλι τα μοντέλα μηχανικής μάθησης και επιχειρηματολογίας που αναπτύσσονται σε λάθος συμπεράσματα.

Λήφθηκαν υπόψη 2 κριτήρια για τη επιλογή του αριθμού των κατηγοριών στις οποίες θα κατηγοριοποιηθούν τα δεδομένα. Ο πρώτος παράγοντας είναι το ιατρικό πεδίο (medical domain) και δηλαδή το τι αντιπροσωπεύουν τα δεδομένα που χρησιμοποιούνται. Σε αυτό λαμβάνεται υπόψη και η εξήγηση που θέλουμε να δώσουμε ώστε να είναι κατανοητή και ερμηνεύσιμη στο συγκεκριμένο πρόβλημα ταξινόμησης υγείων ασθενών και ασθενών με Alzheimer. Σύμφωνα με αυτά, ο ιδανικός τρόπος διαχωρισμού των δεδομένων είναι να χωριστούν σε 5 κατηγορίες [10]. Ένας άλλος παράγοντας που επηρεάζει την κατηγοριοποίηση των δεδομένων, είναι η κατανομή (distribution) των τιμών του κάθε χαρακτηριστικού. Μέσα από τη κατανομή αυτή, μπορούμε να κατανοήσουμε την φύση των δεδομένων και τις διακυμάνσεις τους. Βασικές τεχνικές που μπορούν να χρησιμοποιηθούν ώστε να επιλεγεί η σωστός αριθμός κατηγοριών είναι οι τεχνικές Sturges formula και Freedman-Diaconis [28]. Αυτές οι τεχνικές είναι καλές ειδικά για δεδομένα που δεν έχουν κανονική κατανομή όπως παρατηρούμε και στα δεδομένα του όγκου του ιππόκαμπου στο διάγραμμα 6.1.



Διάγραμμα 6.1: Αρχική κατανομή δεδομένων του όγκου του ιππόκαμπου από βάση ADNI

Με βάση τα πιο πάνω κριτήρια, τα δεδομένα του όγκου του ιππόκαμπου στις πλείστες περιπτώσεις χωρίστηκαν σε 5 κατηγορίες, όπως φαίνεται στο διάγραμμα 6.2.



Διάγραμμα 6.2: Κατανομή δεδομένων του όγκου του ιππόκαμπου από βάση δεδομένων ADNI, μετά την κατηγοριοποίηση των δεδομένων σε 5 κατηγορίες

Οι 5 κατηγορίες του όγκου ιππόκαμπου (Hippocampus volume) χωρίστηκαν όπως φαίνεται πιο κάτω:

1. Κατηγορία 1 - Very Low: 1811.0 έως 2767.2
2. Κατηγορία 2 -Low: 2767.2 έως 3147.4
3. Κατηγορία 3 - Nominal: 3147.4 έως 3463.60
4. Κατηγορία 4 - High: 3463.60 έως 3789.2
5. Κατηγορία 5 - Very High: 3789.2 έως 4605.0

Ακολουθήθηκε η αντίστοιχη μεθοδολογία και για τα υπόλοιπα χαρακτηριστικά.

Στη μελέτη αυτή λαμβάνοντας υπόψη και τα 2 κριτήρια κατηγοριοποίησης που προαναφέρθηκαν, δοκιμάστηκαν διαφορετικές περιπτώσεις κατηγοριοποίησης των

δεδομένων και εξαγωγής των αντίστοιχων αποτελεσμάτων. Σε κάποιες περιπτώσεις έγινε διαχωρισμός των δεδομένων σε 3 κατηγορίες, σε κάποιες περιπτώσεις έγινε διαχωρισμός των δεδομένων σε 5 κατηγορίες και σε κάποια άλλη περίπτωση δοκιμάστηκε ο διαχωρισμός των δεδομένων σε 10 κατηγορίες. Περισσότερη έμφαση δόθηκε στη περίπτωση που τα δεδομένα χωρίστηκαν σε 5 κατηγορίες καθώς ήταν η πιο αντιπροσωπευτική σύμφωνα και με το ιατρικό πεδίο που αφορά την ασθένεια Alzheimer.

Για σκοπούς επαλήθευσης, έγινε σύγκριση των αποτελεσμάτων των μοντέλων μηχανικής μάθησης στις περιπτώσεις πριν και μετά το διαχωρισμό των δεδομένων από συνεχή σε κατηγορικά. Παρατηρήθηκε ότι τα αποτελέσματα και οι μετρικές ακρίβειας ήταν αρκετά κοντά και στις 2 περιπτώσεις και αυτό επιβεβαιώνει ότι ο διαχωρισμός των δεδομένων σε κατηγορίες έγινε σωστά. Στη περίπτωση που χρησιμοποιήθηκαν κατηγορικά δεδομένα, η ακρίβεια του μοντέλου μειώθηκε κατά 4%. Αυτό ήταν αναμενόμενο βέβαια, αφού μετατρέποντας τα δεδομένα από συνεχή τιμές σε κατηγορίες, ένα μέρος της πληροφορίας χάνεται, με αποτέλεσμα η ακρίβεια του μοντέλου μηχανικής μάθησης να μειώνεται.

6.1.3 Μέθοδος αξιολόγησης και επιλογή των κανόνων που εξάγονται από το εργαλείο TE2Rules για αρχικοποίηση του εργαλείου ArgEML

Για την παραγωγή των κανόνων από το εκπαιδευμένο μοντέλο μηχανικής μάθησης, χρησιμοποιήθηκε το εργαλείο TE2Rules στη Python. Το TE2Rules (Tree Ensemble to Rules) είναι μια βιβλιοθήκη στη Python που εξηγεί τα μοντέλα Tree Ensemble, εξάγοντας μια λίστα κανόνων από το εκπαιδευμένο μοντέλο. Στο εργαλείο TE2Rules επιλέχθηκαν κατάλληλοι παράμετροι 'min precision' και 'number of stages' ώστε να παραχθούν σχετικά λίγοι κανόνες που να μην είναι πολύ περίπλοκοι. Αυτό επιδιώχθηκε ώστε να υπάρχει μέγιστη αποδοτικότητα του μοντέλου επιχειρηματολογίας που θα αναπτυχθεί στο εργαλείο ArgEML. Επίσης για το σκοπό αυτό, χρειάζεται να γίνει κατάλληλη επιλογή και αξιολόγηση των

κανόνων που παράγονται από το TE2rules, πριν δοθούν στο εργαλείο ArgEML. Έτσι, το μοντέλο επιχειρηματολογίας που θα αναπτυχθεί και θα βασιστεί σε αυτούς κανόνες, ενδέχεται να είναι πιο αποδοτικό και ερμηνεύσιμο, δίνοντας απλές και κατανοητές εξηγήσεις. Παράλληλα, έτσι θα αποφευχθεί το μεγάλο ποσοστό διλημάτων στα οποία το συγκεκριμένο μοντέλο δεν δίνει ξεκάθαρη απάντηση.

Τα κριτήρια αξιολόγησης και επιλογής του κάθε κανόνα που εξάχθηκε από το μοντέλο μηχανικής μάθησης, μέσω του εργαλείου TE2Rules, ήταν:

- Ο κανόνας να έχει καλή ακρίβεια.
- Ο κανόνας να μην είναι πολύ σύνθετος, και δηλαδή να περιλαμβάνει όσο το δυνατότερο λιγότερες συνθήκες (κατά προτίμηση να περιλαμβάνει το πολύ 5 συνθήκες).
- Ο κανόνας να ικανοποιεί όσο το δυνατότερο περισσότερες περιπτώσεις (records) οι οποίες δεν ικανοποιούνται από κάποιο άλλο κανόνα
- Παράλληλα, ο κανόνας να μην επικαλύπτεται από άλλο κανόνα. (να μην υπάρχει redundancy)
- Ο κανόνας σε συνδυασμό με τους υπόλοιπους κανόνες που επιλέχθηκαν να καλύπτει τουλάχιστον το 90% των δεδομένων

6.2 Παρουσίαση και Αποτελέσματα Μοντέλων επιχειρηματολογίας χρησιμοποιώντας δεδομένα ιππόκαμπου από βάση δεδομένων OASIS

Σε αυτή τη ενότητα τα δεδομένα που χρησιμοποιήθηκαν ήταν:

- Φύλο
- Ηλικία
- Texture Features ιππόκαμπου (HipsASMbaseline, HipsContrastbaseline, HipsCorrelationbaseline, HipsVariancebaseline, HipsSumAveragebaseline, HipsSumVariancebaseline, HipsEntropybaseline, HipsClusterShadebaseline)
- Όγκος ιππόκαμπου

Τα δεδομένα χωρίστηκαν σε 5 κατηγορίες [0,1,2,3,4] που αντιστοιχούν στις κατηγορίες [very Low, Low, Nominal, High, Very High]. Εφαρμόστηκε ίση κατανομή (equal distribution) των κατηγοριών.

Τα train και τα test δεδομένα χωρίστηκαν σε 80% και 20% αντίστοιχα. Αυτό έγινε και στη περίπτωση ανάπτυξης των μοντέλων μηχανικής μάθησης και στη περίπτωση ανάπτυξης του μοντέλου επιχειρηματολογίας στο εργαλείο ArgEML.

	Training δεδομένα	Test δεδομένα
NC	415	104
NA	88	22
Σύνολο	503 (80%)	126 (20%)

Πίνακας 6.1: Χωρισμός δεδομένων σε training και σε test δεδομένα

Όσες περιπτώσεις δεδομένων είχαν έλλειψη από τιμές (είχαν null values) τότε διαγραφτήκαν επειδή αυτό επηρέαζε τη κατηγοριοποίηση των δεδομένων και την εξαγωγή κανόνων. Αυτές οι περιπτώσεις αφορούσαν το ~15% των δεδομένων μας. Μια άλλη προσέγγιση θα ήταν να δοθούν μέσες τιμές του συνόλου στα δεδομένα που ήταν κενά.

Οι κανόνες στους πίνακες 6.2 και 6.3 παράχθηκαν από το μοντέλο Random Forest, χρησιμοποιώντας το εργαλείο TE2 Rules:

	Rule for negative cases	Correctness	Coverage (%)
1	hipposcampusvolumebaseline > 0.5	0.94	81.31
2	hipposcampusvolumebaseline < 2	0.64	39.37

Πίνακας 6.2: Κανόνες που υποστηρίζουν τις αρνητικές περιπτώσεις

	Rule for positive cases	Correctness	Coverage (%)
1	ageatscreening > 2 & hipposcampusvolumebaseline < 1 & hipssumvariancebaseline < 2 & hipsvariancebaseline < 2	0.80	7.95
2	ageatscreening > 3 & hipposcampusvolumebaseline < 1 & hipscontrastbaseline < 3	0.83	5.77
3	hipposcampusvolumebaseline < 1 & hipsclustersshadebaseline > 0 & hipscorrelationbaseline < 1	0.86	4.18
4	hipposcampusvolumebaseline < 1 & hipscontrastbaseline < 3 & hipscorrelationbaseline < 4 & hipsentropybaseline < 2 & hipssumvariancebaseline < 3	5.36	0.78
5	gender < 1 & hipposcampusvolumebaseline < 1 & hipscorrelationbaseline < 3 & hipsentropybaseline < 2	0.83	3.58
6	ageatscreening <= 3.5 & ageatscreening > 1.5 & hipposcampusvolumebaseline <= 1.5 & hipscontrastbaseline <= 3.5 & hipscorrelationbaseline <= 2.5 & hipsentropybaseline <= 0.5 & hipssumaveragebaseline <= 0.5 & hipsvariancebaseline <= 2.5	1	< 1
7	ageatscreening <= 3.5 & hipposcampusvolumebaseline <= 1.5 & hipsasmbaseline <= 1.5 & hipsclustersshadebaseline <= 1.5 & hipssumaveragebaseline > 2.5 & hipssumvariancebaseline <= 3.5	1	< 1
8	ageatscreening > 0.5 & hipposcampusvolumebaseline = 1 & hipsclustersshadebaseline > 2.5 & hipscontrastbaseline > 1.5 & hipsentropybaseline > 3.5 & hipssumaveragebaseline <= 3.5	1	< 1
9	ageatscreening > 3.5 & gender > 0.5 & hipposcampusvolumebaseline <= 1.5 & hipsclustersshadebaseline > 1.5 & hipssumaveragebaseline <= 2.5 & hipssumvariancebaseline > 2.5	1	< 1
10	ageatscreening > 2.5 & gender <= 0.5 & hipposcampusvolumebaseline > 1.5 & hipsasmbaseline > 3.5 & hipsclustersshadebaseline > 1.5 & hipssumaveragebaseline > 1.5	1	< 1
11	hipposcampusvolumebaseline <= 1.5 & hipposcampusvolumebaseline > 0.5 & hipsclustersshadebaseline > 0.5 & hipsentropybaseline <= 1.5 & hipssumaveragebaseline > 0.5 & hipssumvariancebaseline <= 0.5	1	< 1

12	ageatscreening <= 3.5 & hipposcampusvolumebaseline > 0.5 & hipsasmbaseline > 2.5 & hipsclustersshadebaseline > 2.5 & hipscorelationbaseline > 1.5 & hipssumaveragebaseline > 2.5 & hipssumvariancebaseline <= 3.5 & hipsvariancebaseline > 2.5	1	< 1
13	ageatscreening > 1.5 & hipposcampusvolumebaseline > 0.5 & hipsasmbaseline <= 2.5 & hipsasmbaseline > 1.5 & hipsclustersshadebaseline <= 1.5 & hipscontrastbaseline <= 2.5 & hipscorelationbaseline > 0.5 & hipssumaveragebaseline <= 0.5	1	< 1
14	ageatscreening > 2.5 & hipposcampusvolumebaseline > 2.5 & hipsclustersshadebaseline > 3.5 & hipssumaveragebaseline <= 3.5	1	< 1
15	hipsasmbaseline > 1.5 & hipscontrastbaseline > 0.5 & hipscorelationbaseline > 3.5 & hipsentropybaseline > 3.5 & hipssumaveragebaseline > 3.5	1	< 1
16	ageatscreening <= 1.5 & hipposcampusvolumebaseline > 0.5 & hipsasmbaseline > 2.5 & hipsclustersshadebaseline <= 1.5 & hipscorelationbaseline <= 1.5 & hipsentropybaseline > 0.5	1	< 1
17	ageatscreening <= 3.5 & ageatscreening > 1.5 & hipposcampusvolumebaseline <= 1.5 & hipscontrastbaseline <= 3.5 & hipscorelationbaseline <= 2.5 & hipsentropybaseline <= 0.5 & hipssumaveragebaseline <= 0.5 & hipsvariancebaseline <= 2.5	1	< 1
18	ageatscreening <= 3.5 & hipposcampusvolumebaseline <= 1.5 & hipsasmbaseline <= 1.5 & hipsclustersshadebaseline <= 1.5 & hipssumaveragebaseline > 2.5 & hipssumvariancebaseline <= 3.5	1	< 1
19	ageatscreening > 0.5 & hipposcampusvolumebaseline <= 1.5 & hipposcampusvolumebaseline > 0.5 & hipsclustersshadebaseline > 2.5 & hipscontrastbaseline > 1.5 & hipsentropybaseline > 3.5 & hipssumaveragebaseline <= 3.5	1	< 1
20	ageatscreening > 3.5 & gender > 0.5 & hipposcampusvolumebaseline <= 1.5 & hipsclustersshadebaseline > 1.5 & hipssumaveragebaseline <= 2.5 & hipssumvariancebaseline > 2.5	1	< 1

Πίνακας 6.3: Κανόνες που υποστηρίζουν τις θετικές περιπτώσεις.

Για του κανόνες που θα δοθούν ως είσοδος στο εργαλείο ArgEML για τη ανάπτυξη του μοντέλου επιχειρηματολογίας, χρειάστηκε να γίνει επιλογή από τους προαναφερθέντες κανόνες. Δοκιμάστηκαν 3 διαφορετικά μοντέλα, όπου στο κάθε μοντέλο επιλέχθηκε διαφορετικό σεν κανόνων.

➤ 1ο μοντέλο - Τελικοί κανόνες που επιλέχθησαν

	Rule for negative cases	Correctness	Coverage %
1	hipposcampusvolumebaseline > 0.5	0.94	81.31
2	hipposcampusvolumebaseline < 2	0.64	39.37

Πίνακας 6.4: Κανόνες που επιλέχθηκαν που υποστηρίζουν τις αρνητικές περιπτώσεις

	Rule for positive cases	Correctness	Coverage %
3	ageatscreening > 2 & hipposcampusvolumebaseline < 1 & hipssumvariancebaseline < 2 & hipsvariancebaseline < 2	0.80	7.95
4	ageatscreening > 3 & hipposcampusvolumebaseline < 1 & hipscontrastbaseline < 3	0.83	5.77
5	hipposcampusvolumebaseline < 1 & hipsclustersshadebaseline > 0 & hipscorelationbaseline < 1	0.86	4.18
6	hipposcampusvolumebaseline < 1 & hipscontrastbaseline < 3 & hipscorelationbaseline < 4 & hipsentropybaseline < 2 & hipssumvariancebaseline < 3	5.36	0.78
7	gender < 1 & hipposcampusvolumebaseline < 1 & hipscorelationbaseline < 3 & hipsentropybaseline < 2	0.83	3.58

Πίνακας 6.5: Κανόνες που επιλέχθηκαν που υποστηρίζουν τις θετικές περιπτώσεις

- Όλοι κανόνες καλύπτουν το 100% των δεδομένων
- Συγκρούσεις μεταξύ κανόνων (Conflicts) σε όλα τα δεδομένα (train + test set)
 - κανόνας Hipposcampus Volume > 0.5:

Δεν συγκρούεται με κάποιο άλλο κανόνα που υποστηρίζει τις θετικές περιπτώσεις
 - κανόνας hipposcampus volume < 2:
 - Συγκρούεται σε 31 περιπτώσεις με τον κανόνα 3
 - Συγκρούεται σε 43 περιπτώσεις με τον κανόνα 4
 - Συγκρούεται σε 11 περιπτώσεις με τον κανόνα 5
 - Συγκρούεται σε 40 περιπτώσεις με τον κανόνα 6
 - Συγκρούεται σε 18 περιπτώσεις με τον κανόνα 7

Πιο κάτω βλέπουμε τις περιπτώσεις που καλύπτει ο κάθε κανόνας και παράλληλα αυτές οι περιπτώσεις καλύπτονται και από κάποιο άλλο κανόνα

	Κανόνας 1	Κανόνας 2
Κανόνας 1	-	383
Κανόνας 2	128	-

Πίνακας 6.6: Επαναληψιμότητα (Redundancy) μεταξύ κανόνων αρνητικών περιπτώσεων

	Κανόνας 3	Κανόνας 4	Κανόνας 5	Κανόνας 6	Κανόνας 7
Κανόνας 3	0	15	5	15	6
Κανόνας 4	15	0	0	23	4
Κανόνας 5	5	0	0	0	3
Κανόνας 6	15	23	0	0	6
Κανόνας 7	6	4	3	6	0

Πίνακας 6.7: Επαναληψιμότητα (Redundancy) μεταξύ κανόνων θετικών περιπτώσεων

➤ 2ο μοντέλο - Τελικοί κανόνες που επιλέχθησαν

	Rule for negative cases	Correctness	Coverage %
1	hippocampusvolumebaseline > 0.5	0.94	81.31

Πίνακας 6.8: Κανόνες που επιλέχθηκαν που υποστηρίζουν τις αρνητικές περιπτώσεις

	Rule for positive cases	Correctness	Coverage %
2	ageatscreening > 2 & hippocampusvolumebaseline < 1 & hipssumvariancebaseline < 2 & hipsvariancebaseline < 2	0.80	7.95
3	ageatscreening > 3 & hippocampusvolumebaseline < 1 & hipscontrastbaseline < 3	0.83	5.77
4	hippocampusvolumebaseline < 1 & hipsclustershadebaseline > 0 & hipscorelationbaseline < 1	0.86	4.18
5	hippocampusvolumebaseline < 1 & hipscontrastbaseline < 3 & hipscorelationbaseline < 4 & hipsentropybaseline < 2 & hipssumvariancebaseline < 3	5.36	0.78
6	gender < 1 & hippocampusvolumebaseline < 1 & hipscorelationbaseline < 3 & hipsentropybaseline < 2	0.83	3.58

Πίνακας 6.9: Κανόνες που επιλέχθηκαν που υποστηρίζουν τις θετικές περιπτώσεις

- Όλοι κανόνες καλύπτουν το 94% των δεδομένων
- Δεν υπάρχουν συγκρούσεις μεταξύ των κανόνων

	Κανόνας 3	Κανόνας 4	Κανόνας 5	Κανόνας 6	Κανόνας 7
Κανόνας 3	0	15	5	15	6
Κανόνας 4	15	0	0	23	4
Κανόνας 5	5	0	0	0	3
Κανόνας 6	15	23	0	0	6
Κανόνας 7	6	4	3	6	0

Πίνακας 6.10: Επαναληψιμότητα (Redundancy) μεταξύ κανόνων θετικών περιπτώσεων

➤ 3^ο μοντέλο - δεν δόθηκαν αρχικά κανόνες

Στο 3ο μοντέλο ArgEML, δεν δόθηκαν αρχικοί κανόνες από το TE2Rules στο εργαλείο ArgEML. Έτσι στο εργαλείο ArgEML, δόθηκε μόνο το σετ των δεδομένων και το εργαλείο από μόνο του εξαρχής άρχισε να αναπτύσσει τη θεωρία επιχειρηματολογίας.

Αποτελέσματα των 3 μοντέλων

Τα αποτελέσματα και η ακρίβεια των 3 μοντέλων που δοκιμάστηκαν φαίνονται στους πιο κάτω πίνακες.

	Train data =503 Test data = 126	Accuracy (LA)	Definite Accuracy	Correct Predictions	False predictions	Dilemmas
ArgEML -1 ^ο Μοντέλο	Train data	0.80	0.93	329	25	30%
	Test data	0.74	0.83	77	16	26%
ArgEML -2 ^ο Μοντέλο	Train data	0.80	0.92	329	27	29%
	Test data	0.73	0.83	72	15	31%
ArgEML -3 ^ο Μοντέλο (δεν δοθηκαν αρχικοί κανόνες)	Train data	0.86	0.90	411	45	9.3%
	Test data	0.85	0.89	101	13	9.5%

Πίνακας 6.11: Αποτελέσματα μοντέλων επιχειρηματολογίας από βάση δεδομένων OASIS χρησιμοποιώντας δεδομένα όγκου και χαρακτηριστικά υφής ιππόκαμπου

	AUC	Definite Accuracy (LA)	Sensitivity	Specificity	Dilemmas
ArgEML -1 ^ο Μοντέλο	0.55	0.83 (0.74)	0.20	0.90	26%
ArgEML -2 ^ο Μοντέλο	0.71	0.83 (0.73)	0.54	0.88	31%
ArgEML -3 ^ο Μοντέλο (δεν δοθηκαν αρχικοί κανόνες)	0.62	0.89 (0.85)	0.27	0.98	9.5%

Πίνακας 6.12: Αποτελέσματα μετρικών αξιολόγησης μοντέλων επιχειρηματολογίας από βάση δεδομένων OASIS χρησιμοποιώντας δεδομένα όγκου και χαρακτηριστικά υφής ιππόκαμπου

Παρατηρούμε ότι το μοντέλο 3 που αναπτύχθηκε εξολοκλήρου από το εργαλείο ArgEML έχει καλύτερη ακρίβεια (accuracy) από το 1^ο και 2^ο μοντέλο. Όμως, το Sensitivity (recall) του συγκεκριμένου μοντέλου είναι ίσο με 0.27 και είναι πάρα πολύ χαμηλό. Αυτή η μετρική είναι

πολύ σημαντικό να είναι υψηλή στο τομέα της ιατρικής καθώς δείχνει το πόσο καλά το μοντέλο ανιχνεύει τις θετικές περιπτώσεις.

Φαίνεται ότι λόγω του ότι στα δεδομένα μας, το 83% των περιπτώσεων (519 από τις 629 συνολικά περιπτώσεις) είναι αρνητικές περιπτώσεις, στο μοντέλο επιχειρηματολογίας μας δημιουργείται ο κανόνας 'negative is generally true'. Λόγω αυτού και λόγω του ότι το εργαλείο ArgEML εστιάζει στο accuracy, το sensitivity είναι μικρό. Σε αυτό πιθανόν να έπαιξε ρόλο και ο διαχωρισμός των δεδομένων σε 5 κατηγορίες.

Συμπερασματικά, το 2ο μοντέλο φαίνεται να έχει συνολικά τη καλύτερη απόδοση από ό,τι τα άλλα μοντέλα. Και ο λόγος είναι ότι κρατάει μια ισορροπία στη πρόβλεψη θετικών και αρνητικών περιπτώσεων έχοντας σχετικά καλή ακρίβεια και παραμέτρους Sensitivity και Specificity μεγαλύτερες του 0.50.

Κάτι που σημαίνει ότι η επιλογή των κανόνων που έγινε στο 2ο μοντέλο ήταν καλύτερη. Αυτό δείχνει ότι η μη χρήση του κανόνα 2 από το TE2rules (που υποστηρίζει τις αρνητικές περιπτώσεις και έχει χαμηλή ακρίβεια) στο ArgEML μοντέλο, επηρέασε θετικά το μοντέλο καθώς η θεωρεία που αναπτύχθηκε ισορρόπησε τα επιχειρήματα που αναπτύχθηκαν για τις θετικές και τις αρνητικές περιπτώσεις. Για αυτό, και η μετρική Sensitivity έφτασε τη τιμή 0.54.

Μέσω του εργαλείου ArgEML και των αποτελεσμάτων, μπορούμε να αναλύσουμε τη θεωρεία που αναπτύχθηκε και να δούμε με ποιους κανόνες έγιναν οι προβλέψεις. Πολλές προβλέψεις έγιναν με τον ίδιο τρόπο και δηλαδή μέσω του ίδιου κανόνα από το εργαλείο.

Κάθε ομάδα χαρακτηρίζεται από έναν μοναδικό τύπο ή μοτίβο επεξήγησης. Τα τελικά αποτελέσματα (προβλέψεις) του test set της καλύτερης περίπτωσης (2ο μοντέλο) φαίνονται

στο πίνακα 6.13. Στο πίνακα αυτό παρουσιάζονται οι δημοφιλέστερες επεξηγήσεις που δόθηκαν τις περισσότερες φορές.

Αυτά μπορούν να δομηθούν σε διαφορετικά υποτμήματα. Στη συγκεκριμένη μαθημένη θεωρία για παράδειγμα είχαμε τις πιο κάτω περιπτώσεις εξηγήσεων, τα οποία παρουσιάζονται στο πίνακα 6.13:

- Σε κάποιες περιπτώσεις έδωσε ως απάντηση θετική (παράδειγμα φαίνεται στη Κατηγορία 7 στο πίνακα 6.13).
- Σε περισσότερες περιπτώσεις το εργαλείο έδωσε απάντηση δίλημμα για διαφορετικούς λόγους κάθε φορά. Παράδειγμα φαίνεται στη Κατηγορία 6 στο πίνακα 6.13.
- Στις πλείστες περιπτώσεις το εργαλείο έδωσε ως απάντηση αρνητική για διαφορετικούς λόγους κάθε φορά (βλ. Κατηγορία 1 έως 5 στο πίνακα 6.13).

Expl Group		Prediction	Coverage	Accuracy	Explanation
1	r2d1;p7; r1	Negative	7.4%	0.97	The model derived a definite prediction, negative, because, gender is equal to male.
2	r6;p34;r1	Negative	4.2%	0.95	The model derived a definite prediction, negative, because, hippocampus volume is equal to low.
3	r4d1;p21; r1	Negative	3.6%	0.78	The model derived a definite prediction, negative, because, hips entropy is equal to nominal AND hippocampus volume is equal to low.
4	r2d1;p13; r1	Negative	6.8%	0.97	The model derived a definite prediction, negative, because, gender is equal to male.
5	r1	Negative	28%	0.99	The model derived a definite prediction, negative, because, negative is generally true. (In this case no rule supporting positive cases applies)
6	r6;p30;r1 p31;r4;p1 2; r2	dilemma	1.3%	-	<u>For this case the model is in dilemma.</u> <u>Various predictions are equally supported:</u> <ul style="list-style-type: none"> • positive, because, while the condition(s) [hippocampus volume is equal to low] suggest the prediction negative, the condition(s) [hippocampus volume is equal to low] that support the prediction positive are stronger. • negative, because, hippocampus volume is equal to low.
7	r8;p44;r2	Positive	0.4%	0.50	The model derived a definite prediction, positive, because, hipssumaveragebaseline is equal to verylow AND gender is equal to male.

Πίνακας 6.13: Υποομάδες Δεδομένων που προσδιορίζονται από συγκεκριμένη εξήγηση από το εργαλείο ArgEML

6.2.1 Ανάπτυξη μοντέλων επιχειρηματολογίας με σκοπό τη μείωση των διλημάτων

Σε συγκεκριμένες περιπτώσεις, η θεωρία επιχειρηματολογίας που ανέπτυξε το μοντέλο ArgEML είχε ως αποτέλεσμα πολλά διλήματα. Για να αποφευχθούν τα διλήματα, χρειάστηκε στις αρχικές παραμέτρους (ρυθμίσεις) που δίνονται στο εργαλείο ArgEML, να αυξηθεί κατάλληλα η παράμετρος 'definite error', με σκοπό η θεωρία που αναπτύσσεται να έχει ένα περιθώριο σφάλματος να αναπτυχθεί ώστε να μην δίνει πολλά διλήματα.

Όπως φαίνεται στις περιπτώσεις που παρουσιάζονται στο πίνακα 6.14, χρειάστηκε να υπάρχει μια ισορροπία μεταξύ σφάλματος και ακρίβειας ώστε το μοντέλο να είναι αποδοτικό. Δοκιμάστηκαν σύνολο 4 διαφορετικές περιπτώσεις μοντέλων επιχειρηματολογίας που αναπτύχθηκαν στο εργαλείο ArgEML όπου στη κάθε περίπτωση δόθηκε διαφορετικό όριο σφαλμάτων στο εργαλείο ArgEML. Οι κανόνες που δόθηκαν και στις 4 περιπτώσεις ήταν αυτοί που αναφέρθηκαν στο 'μοντέλο ArgEML 2' στη προηγούμενη παράγραφο. Καθώς αυξάνεται το όριο ανοχής σφαλμάτων που δίνεται στο εργαλείο, το ποσοστό των διλημάτων στις προβλέψεις του εργαλείου μειώνεται. Παράλληλα όμως, αλλάζει η θεωρία ανάπτυξης του μοντέλου και επηρεάζονται όλες οι μετρικές αξιολόγησης του μοντέλου. Πιο κάτω παρατηρείται ότι η 3η περίπτωση είναι αυτή που έχει τα καλύτερα αποτελέσματα, όμως το ποσοστό των διλημάτων σε αυτή τη περίπτωση, εξακολουθεί να είναι μεγάλο. Η 2η περίπτωση έχει ικανοποιητικό ποσοστό διλημάτων (30%) όμως η μετρική Sensitivity είναι αρκετά μικρή (0.54).

Model		Accuracy (LA)	Definite Accuracy	Correct Predictions	Errors (False predictions)	Dilemmas	AUC	Specificity	Sensitivity
ArgEML 1 st Περίπτωση	Train data	0.78	0.84	351	68	17%	0.63	0.89	0.39
	Test data	0.77	0.83	86	18	17%			
ArgEML - 2 ^η Περίπτωση	Train data	0.80	0.92	329	27	29%	0.71	0.88	0.54
	Test data	0.73	0.83	72	15	31%			
ArgEML – 3 ^η Περίπτωση	Train data	0.75	0.90	283	31	38%	0.88	0.88	0.89
	Test data	0.72	0.88	67	9	40%			
ArgEML 4 th Περίπτωση	Train data	0.71	0.90	237	26	48%	0.72	0.86	0.58
	Test data	0.67	0.81	57	13	44%			

Πίνακας 6.14: Συνοπτικά αποτελέσματα χρησιμοποιώντας δεδομένα όγκου και χαρακτηριστικά υφής υπόκαμπου από βάση OASIS

1^η περίπτωση: Έγινε εισαγωγή κανόνων από TE2Rules και στις ρυθμίσεις παραμέτρων επιλέχθηκε definite error = 10%

2^η περίπτωση: Έγινε εισαγωγή κανόνων από TE2Rules και στις ρυθμίσεις παραμέτρων επιλέχθηκε definite error = 7%

3^η περίπτωση: : Έγινε εισαγωγή κανόνων από TE2Rules και στις ρυθμίσεις παραμέτρων επιλέχθηκε definite error = 4%

4^η περίπτωση: : Έγινε εισαγωγή κανόνων από TE2Rules και στις ρυθμίσεις παραμέτρων επιλέχθηκε definite error = 0%

Στο πίνακα 6.15 γίνεται σύγκριση αποτελεσμάτων μοντέλου επιχειρηματολογίας που αναπτύχθηκε με εξήγηση σε σχέση με μοντέλα μηχανικής μάθησης Random Forest. Αξίζει να σημειωθεί ότι στο μοντέλο Random Forest έχει μεγαλύτερη ακρίβεια (κατά 4%) στη περίπτωση που χρησιμοποιούνται δεδομένα με συνεχή τιμές. Όπως ήταν αναμενόμενο η απόδοση του μοντέλου μειώνεται όταν χρησιμοποιούνται τα κατηγορικά δεδομένα.

Επίσης, το μοντέλο επιχειρηματολογίας που αναπτύχθηκε έχει καλύτερη ακρίβεια από τα μοντέλα Random Forest, ειδικά στη περίπτωση στη οποία δεν λαμβάνονται υπόψη τα διλήμματα. Τέλος, όπως φαίνεται στο πίνακα 6.15, σε σύγκριση με αντίστοιχο μοντέλο επιχειρηματολογίας που είχε γίνει στην έρευνα [10] όπου είχαν χρησιμοποιηθεί παρόμοια

χαρακτηριστικά ιπποκάμπου, αυτό το μοντέλο επιχειρηματολογίας που αναπτύχθηκε φαίνεται να έχει παρόμοια αποτελέσματα σύμφωνα με τις μετρικές αξιολόγησης Accuracy, Sensitivity και Specificity.

Μοντέλο	OASIS dataset				Dilemmas (%)
	AUC	Definite Accuracy (LA)	Sensitivity (Recall)	Specificity	
Μοντέλο Μηχανικής Μάθησης Random Forest - 10 CV	0.71	0.87	0.464	0.96	0%
Μοντέλο Μηχανικής Μάθησης Random Forest 2- 10 CV -	0.71	0.75	0.72	0.78	0%
ArgEML - 2 ^ο Μοντέλο (με όριο ανοχής σφάλματος 7%)	0.71	0.83 (0.73)	0.54	0.88	31%
ArgEML - 2^ο Μοντέλο (με όριο ανοχής σφάλματος 4%)	0.88	0.88 (0.72)	0.89	0.88	40%
Αποτελέσματα άρθρου [10]	-	0.91	0.87	0.95	-

Πίνακας 6.15: Σύγκριση μοντέλων μηχανικής μάθησης και μοντέλων επιχειρηματολογίας χρησιμοποιώντας δεδομένα όγκου και χαρακτηριστικά υφής ιπποκάμπου από βάση OASIS. Στο Μοντέλο Μηχανικής Μάθησης Random Forest 2 ρυθμίστηκαν κατάλληλα οι παράμετροι ώστε να δοθεί ιδιαίτερη έμφαση στη μετρική Sensitivity. Η μελέτη [10] είχε χρησιμοποιήσει παρόμοια χαρακτηριστικά και αφορούσε επίσης μοντέλο επιχειρηματολογίας

6.3 Παρουσίαση και Αποτελέσματα Μοντέλων επιχειρηματολογίας χρησιμοποιώντας δεδομένα ιπποκάμπου από βάση δεδομένων ADNI

Σε αυτή τη ενότητα τα δεδομένα που χρησιμοποιήθηκαν ήταν:

- Φύλο
- Ηλικία
- Texture Features ιπποκάμπου (HipsASMBaseline, HipsContrastbaseline, HipsCorelationbaseline, HipsVariancebaseline, HipsSumAveragebaseline, HipsSumVariancebaseline, HipsEntropybaseline, HipsClusterShadebaseline)
- Όγκος ιπποκάμπου

Τα δεδομένα χωρίστηκαν σε 3 κατηγορίες [0,1,2] που αντιστοιχούν στις τιμές [Low, Nominal, High].

Για παράδειγμα, οι κατηγορίες στις οποίες χωρίστηκε ο όγκος του ιππόκαμπου φαίνονται στο πίνακα 6.16.

	Όγκος Ιππόκαμπου	Αριθμός Περιπτώσεων
Κατηγορία 0	1811.00 έως 3037.33	103
Κατηγορία 1	3037.33 έως 3587.67	102
Κατηγορία 2	3587.67 έως 4605.00	103

Πίνακας 6.16: Κατηγορίες στις οποίες χωρίστηκε ο όγκος ιππόκαμπου

Τα train και τα test δεδομένα χωρίστηκαν σε 80% (208 σύνολο δεδομένα) και 20% (54 σύνολο δεδομένα) αντίστοιχα. Αυτό έγινε και στη περίπτωση ανάπτυξης των μοντέλων μηχανικής μάθησης και στη περίπτωση ανάπτυξης του μοντέλου επιχειρηματολογίας στο εργαλείο ArgEML.

ADNI database	Training δεδομένα	Test δεδομένα
NC	120	31
NA	88	23
Σύνολο	208 (80%)	54 (20%)

Πίνακας 6.17: Διαχωρισμός των δεδομένων από το εργαλείο ArgEML

Όσες περιπτώσεις δεδομένων είχαν έλλειψη από τιμές (είχαν null values) τότε διαγραφτήκαν επειδή αυτό επηρέαζε τη κατηγοριοποίηση των δεδομένων και την εξαγωγή κανόνων. Αυτές οι περιπτώσεις αφορούσαν το ~10% των δεδομένων μας. Μια άλλη προσέγγιση θα ήταν να δοθούν μέσες τιμές στα δεδομένα που ήταν κενά.

Οι κανόνες στους πίνακες 6.18 και 6.19 παράχθηκαν από μοντέλο Random Forest, χρησιμοποιώντας το εργαλείο TE2 Rules:

	Rule	Correctness	Coverage (%)
1	hipposcampusvolumebaseline >= 1 & hipsasmbaseline < 2	0.79	54.50
2	hipposcampusvolumebaseline >= 1 & hipsvariancebaseline >= 1	0.80	42.58
3	hipsclustersshadebaseline >= 1 & hipsentropybaseline >= 1 & hipssumvariancebaseline >= 1	0.77	29.19
4	ageatscreening >= 1 & hipscontrastbaseline >= 1 & hipsentropybaseline >= 1	0.81	21.05
5	hipscorelationbaseline >= 1 & hipsentropybaseline > 1	0.78	17.70
6	ageatscreening > 1.5 & hipsclustersshadebaseline > 0.5 & hipssumvariancebaseline > 1.5	0.88	3.82
7	hipscontrastbaseline <= 0.5 & hipsentropybaseline <= 0.5 & hipsvariancebaseline > 1.5	1	0.48
8	ageatscreening <= 0.5 & hipsclustersshadebaseline > 0.5 & hipssumaveragebaseline > 0.5 & hipssumvariancebaseline > 1.5	0.78	4.30

Πίνακας 6.18: Κανόνες που υποστηρίζουν τις αρνητικές περιπτώσεις

	Rule	Correctness	Coverage (%)
1	hipposcampusvolumebaseline < 1 & hipscorelationbaseline < 2	0.84	21.5
2	hipposcampusvolumebaseline < 2 & hipsentropybaseline < 1 & hipsvariancebaseline < 2	0.83	14.40
3	hipposcampusvolumebaseline < 1 & hipsclustersshadebaseline < 1	0.86	10.53
4	ageatscreening < 1 & hipposcampusvolumebaseline < 2 & hipscontrastbaseline < 2 & hipscontrastbaseline > 0 & hipsentropybaseline > 0 & hipssumvariancebaseline < 2	0.89	4.31
5	hipscorelationbaseline > 0 & hipsentropybaseline < 2 & hipssumvariancebaseline < 1 & hipsvariancebaseline < 1	0.86	3.35
6	ageatscreening < 1 & hipposcampusvolumebaseline < 2 & hipsentropybaseline < 1	0.91	5.74
7	gender <= 0.5 & hipposcampusvolumebaseline <= 1.5 & hipposcampusvolumebaseline > 0.5 & hipsclustersshadebaseline <= 0.5 & hipssumaveragebaseline > 0.5 & hipssumvariancebaseline > 1.5	1	1.44
8	hipscontrastbaseline <= 0.5 & hipsentropybaseline > 0.5 & hipssumaveragebaseline <= 1.5 & hipssumaveragebaseline > 0.5 & hipssumvariancebaseline > 0.5 & hipsvariancebaseline <= 0.5	1	0.96
9	gender > 0.5 & hipsasmbaseline > 1.5 & hipscontrastbaseline <= 1.5	0.875	3.83
10	ageatscreening <= 0.5 & hipsclustersshadebaseline <= 0.5 & hipscontrastbaseline <= 1.5	0.89	4.30
11	hipposcampusvolumebaseline > 1.5 & hipscontrastbaseline > 1.5 & hipscorelationbaseline > 1.0 & hipsentropybaseline <= 0.5	1	0.50

Πίνακας 6.19: Κανόνες που υποστηρίζουν τις θετικές περιπτώσεις

Οι τελικοί κανόνες που επιλέχθηκαν για να εισαχθούν στο εργαλείο ArgEML, είναι οι εξής:

	Rule	Correctness	Coverage (%)
1	hipposcampusvolumebaseline >= 1 & hipsasmbaseline < 2	0.79	54.50
2	hipposcampusvolumebaseline >= 1 & hipsvariancebaseline >= 1	0.80	42.58
3	hipsclustersshadebaseline >= 1 & hipsentropybaseline >= 1 & hipssumvariancebaseline >= 1	0.77	29.19
4	ageatscreening >= 1 & hipscontrastbaseline >= 1 & hipsentropybaseline >= 1	0.81	21.05
5	hipscorelationbaseline >= 1 & hipsentropybaseline > 1	0.78	17.70

Πίνακας 6.20: Κανόνες που επιλέχθηκαν που υποστηρίζουν τις αρνητικές περιπτώσεις

	Rule	Correctness	Coverage (%)
1	hipposcampusvolumebaseline < 1 & hipscorelationbaseline < 2	0.84	21.5
2	hipposcampusvolumebaseline < 2 & hipsentropybaseline < 1 & hipsvariancebaseline < 2	0.83	14.40
3	hipposcampusvolumebaseline < 1 & hipsclustersshadebaseline < 1	0.86	10.53
4	ageatscreening < 1 & hipposcampusvolumebaseline < 2 & hipscontrastbaseline < 2 & hipscontrastbaseline > 0 & hipsentropybaseline > 0 & hipssumvariancebaseline < 2	0.89	4.31
5	hipscorelationbaseline > 0 & hipsentropybaseline < 2 & hipssumvariancebaseline < 1 & hipsvariancebaseline < 1	0.86	3.35
6	ageatscreening < 1 & hipposcampusvolumebaseline < 2 & hipsentropybaseline < 1	0.91	5.74

Πίνακας 6.21: Κανόνες που επιλέχθηκαν που υποστηρίζουν τις θετικές περιπτώσεις

Όλοι κανόνες ικανοποιούν το 99% των περιπτώσεων. Πιο κάτω βλέπουμε τις περιπτώσεις που καλύπτει ο κάθε κανόνας και παράλληλα αυτές οι περιπτώσεις καλύπτονται και από κάποιο άλλο κανόνα.

	Κανόνας 1	Κανόνας 2	Κανόνας 3	Κανόνας 4	Κανόνας 5
Κανόνας 1	0	77	59	44	36
Κανόνας 2	77	0	56	36	13
Κανόνας 3	59	56	0	35	23
Κανόνας 4	44	36	35	0	12
Κανόνας 5	36	13	23	12	0

Πίνακας 6.22: Επικάλυψη (Redundancy) μεταξύ κανόνων αρνητικών περιπτώσεων

Αν και φαίνεται να υπάρχει μεγάλη επικάλυψη μεταξύ κανόνων αρνητικών περιπτώσεων στους κανόνες που επιλέχθηκαν, παρόλαυτα όλοι κανόνες που επιλέχθηκαν είναι σημαντικοί

αφού καλύπτουν δεδομένα τα οποία δεν καλύπτονται από άλλους κανόνες και ταυτόχρονα βοηθούν στο να υποστηρίξουν τις αρνητικές περιπτώσεις, αυξάνοντας τη ακρίβεια.

	Κανόνας 6	Κανόνας 7	Κανόνας 8	Κανόνας 9	Κανόνας 10	Κανόνας 11
Κανόνας 6	0	17	15	4	1	8
Κανόνας 7	17	0	9	0	1	9
Κανόνας 8	15	9	0	1	0	9
Κανόνας 9	4	0	1	0	1	0
Κανόνας 10	1	1	0	1	0	0
Κανόνας 11	8	9	9	0	0	0

Πίνακας 6.23: Επικάλυψη (Redundancy) μεταξύ κανόνων θετικών περιπτώσεων

	Κανόνας 1	Κανόνας 2	Κανόνας 3	Κανόνας 4	Κανόνας 5
Κανόνας 6	0	0	12	11	7
Κανόνας 7	2	9	0	0	0
Κανόνας 8	0	0	0	3	1
Κανόνας 9	5	2	3	0	1
Κανόνας 10	9	0	0	1	0
Κανόνας 11	1	2	0	0	0

Πίνακας 6.24: Συγκρούσεις μεταξύ κανόνων (Conflicts) σε όλα τα δεδομένα (train + test set)

Αποτελέσματα μοντέλου που αναπτύχθηκε

	Accuracy (LA)	Definite Accuracy	Correct Predictions	Errors (False predictions)	Dilemmas
Train data	0.71	0.81	55%	12%	33%
Test data	0.70	0.82	51%	10%	47%

Πίνακας 6.25: Αποτελέσματα μοντέλου επιχειρηματολογίας χρησιμοποιώντας δεδομένα υποκάμπτου από βάση δεδομένων ADNI

Μέσω του εργαλείου ArgEML Gorgias και των αποτελεσμάτων, μπορούμε να αναλύσουμε τη θεωρία που αναπτύχθηκε και να δούμε με ποιους κανόνες έγιναν οι προβλέψεις. Πολλές προβλέψεις έγιναν με τον ίδιο τρόπο και δηλαδή μέσω του ίδιου κανόνα από το εργαλείο. Οι κυριότερες προβλέψεις φαίνονται στο πίνακα 6.26. Στη συγκεκριμένη περίπτωση, βλέπουμε ότι η θεωρία επιχειρηματολογίας που αναπτύχθηκε δεν είναι τόσο βοηθητική και

ξεκάθαρη. Αυτό ίσως να οφείλεται και στο ότι τα δεδομένα κατηγοριοποιήθηκαν σε 3 κατηγορίες και αυτό ίσως να δυσκόλεψε το εργαλείο να αναπτύξει κατάλληλα τη θεωρία. Σε κάποιες από τις περιπτώσεις που τα δεδομένα χωρίζονταν σε περισσότερες κατηγορίες δεν παρατηρήθηκε κάτι αντίστοιχο, ωστόσο σε εκείνες τις περιπτώσεις η ακρίβεια δεν ήταν τόσο μεγάλη, όπως θα παρουσιαστεί στη συνέχεια.

explGroup	Rule	prediction	coverage	accuracy	explanation
1	r2	positive	35%	79%	The model derived a definite prediction, positive, because, positive is generally true.
2	r1 r2	dilemma	33%	-	For this case the model is in dilemma. Various predictions are equally supported: <ul style="list-style-type: none"> • positive, because, positive is generally true. • negative, because negative is generally true.
3	r1	negative	32%	86%	The model derived a definite prediction, negative, because negative is generally true.

Πίνακας 6.26: Υποομάδες Δεδομένων που προσδιορίζονται από συγκεκριμένη εξήγηση από το εργαλείο ArgEML

Μοντέλο	ADNI δεδομένα				Dilemmas
	AUC	Definite Accuracy (LA)	Sensitivity (Recall)	Specificity	
Μοντέλο Μηχανικής Μάθησης Random Forest	0.79	0.79	0.70	0.89	0%
Μοντέλο ArgEML	0.82	0.82 (0.70)	0.80	0.84	47%

Πίνακας 6.27: Σύγκριση καλύτερου μοντέλου επιχειρηματολογίας που αναπτύχθηκε χρησιμοποιώντας δεδομένα ιππόκαμπου από βάση δεδομένων ADNI σε σχέση με μοντέλο μηχανικής μάθησης

Στη συγκεκριμένη περίπτωση, αν δεν ληφθούν υπόψη τα διλήμματα, τότε παρατηρούμε στο πίνακα 6.27 ότι το μοντέλο επιχειρηματολογίας που αναπτύχθηκε είναι πιο ακριβές από ότι το μοντέλο μηχανικής μάθησης Random Forest που είχε αναπτυχθεί αρχικώς. Βέβαια, όπως

είχε αναφερθεί προηγουμένως το μοντέλο επιχειρηματολογίας στο 47% των περιπτώσεων δεν μπορεί να δώσει ξεκάθαρη απάντηση (θετική ή αρνητική) δίνοντας δίλημμα.

6.3.1 Σύγκριση μοντέλων επιχειρηματολογίας με βάση των αριθμό κατηγοριών που επιλέχθηκαν για τη κατηγοροποίηση τους

Στο πίνακα 6.28, παρουσιάζονται τα αποτελέσματα που έδωσε το μοντέλο επιχειρηματολογίας στις περιπτώσεις που τα δεδομένα χωρίστηκαν σε 3, 5 και 10 κατηγορίες. Για τη ανάπτυξη του κάθε μοντέλου επιχειρηματολογίας, έγινε η εξαγωγή των κανόνων από το μοντέλο Random Forest μέσω TE2Rules και αφού επιλέχθηκε ένα υποσύνολο των κανόνων, αυτό δόθηκε ως εισαγωγή στο εργαλείο ArgEML για να αναπτυχθεί το μοντέλο επιχειρηματολογίας.

Όπως φαίνεται, κατηγοριοποιώντας τα δεδομένα μας σε 3 κατηγορίες, το μοντέλο επιχειρηματολογίας που αναπτύσσεται αποδίδει καλύτερα. Ειδικά, αν εξαιρεθεί το μεγάλο ποσοστό διλημάτων (47%), το μοντέλο αυτό φαίνεται να είναι πιο ισορροπημένο έχοντας υψηλά ποσοστά Accuracy, AUC, Sensitivity, Specificity κάτι το οποίο δεν έχουν τα υπόλοιπα μοντέλα.

Model		Accuracy (LA)	Definite Accuracy	Correct Predictions	Errors (False predictions)	Dilemmas	AUC	Specificity	Sensitivity
1st ArgEML model - 3 categories	Train data	0.71	0.81	55%	12%	33%	-	-	-
	Test data	0.70	0.82	52%	11%	47%	0.82	0.84	0.80
2 st ArgEML model- 5 categories Negative rules	Train data	0.77	0.82	80%	17%	10%	-	-	-
	Test data	0.80	0.85	81%	15%	4%	0.73	1.0	0.46
3 st ArgEML Model - 10 categories	Train data	0.81	1	63%	0%	37%	-	-	-
	Test data	0.69	0.83	46%	1%	44%	0.81	0.89	0.73

Πίνακας 6.28: Σύγκριση μοντέλων επιχειρηματολογίας ανάλογα με το αριθμό κατηγοριών που χωρίστηκαν αρχικά τα δεδομένα χρησιμοποιώντας δεδομένα υπόκαμπου από βάση δεδομένων ADNI

Στο πίνακα 6.29, παρουσιάζονται τα αποτελέσματα που έδωσε το μοντέλο επιχειρηματολογίας στις περιπτώσεις που τα δεδομένα χωρίστηκαν σε 3, 5 και 10 κατηγορίες. Και στις 3 εκείνες περιπτώσεις, δεν δόθηκαν κανόνες από το εργαλείο TE2rules και δηλαδή το εργαλείο ArgEML ανέπτυξε τη θεωρία από μόνο του με βάση τα δεδομένα που του δόθηκαν.

Model		Accuracy (LA)	Definite Accuracy	Correct Predictions	Errors (False predictions)	Dilemmas	AUC	Specificity	Sensitivity
ArgEML - 3 categories	Train data	0.71	0.82	111	25	35%	-	-	-
	Test data	0.72	0.82	31	7	30%	0.82	0.88	0.76
ArgEML- 5 categories	Train data	0.79	0.94	128	8	35%	-	-	-
	Test data	0.57	0.60	15	10	48%	0.73	0.81	0.53
ArgEML- 10 categories	Train data	0.78	0.91	135	13	31%	-	-	-
	Test data	0.6	0.67	22	11	39%	0.73	1	0.47

Πίνακας 6.29: Σύγκριση μοντέλων επιχειρηματολογίας ανάλογα με το αριθμό κατηγοριών που χωρίστηκαν αρχικά τα δεδομένα, χρησιμοποιώντας δεδομένα ιππόκαμπου από βάση δεδομένων ADNI. Τα μοντέλα αυτά αναπτύχθηκαν εξολοκλήρου από το εργαλείο ArgEML χωρίς να γίνει εισαγωγή κανόνων από το εργαλείο TE2rules

6.3.2 Σύγκριση μοντέλων επιχειρηματολογίας με βάση το μοντέλο μηχανικής μάθησης που χρησιμοποιήθηκε αρχικά για τη αρχικοποίηση των κανόνων

Στη συνέχεια γίνεται σύγκριση των μοντέλων XGBoost και Random Forest που συνδυάζονται μαζί με το εργαλείο ArgEML για να παραχθεί εξήγηση. Και στις 2 περιπτώσεις, έγινε χρήση δεδομένων από τη βάση ADNI τα οποία κατηγοριοποιήθηκαν σε 3 κατηγορίες και έγινε κατάλληλη επιλογή ορίου ανοχής σφάλματος με σκοπό να υπάρχουν όσο το δυνατό καλύτερα αποτελέσματα.

Model		LA	Definite Accuracy	Correct Predictions	Errors (False predictions)	Dilemmas	AUC	Specificity	Sensitivity
		ArgEML μοντέλο Μέσω Random Forest	Train data	0.71	0.81	114	26	33%	-
	Test data	0.70	0.82	28	6	47%	0.82	0.84	0.80
ArgEML Μέσω XGBoost	Train data	0.70	0.81	109	25	36%	-	-	-
	Test data	0.74	0.82	33	7	26%	0.82	0.86	0.77

Πίνακας 6.30: Σύγκριση των μοντέλων επιχειρηματολογίας στα οποία οι κανόνες που δόθηκαν για τη αρχικοποίησή τους, ήταν διαφορετικοί λόγω διαφορετικού μοντέλου μηχανικής μάθησης που χρησιμοποιήθηκε αρχικά.

Στο πίνακα 6.30, στη 1^η περίπτωση έγινε χρήση του μοντέλου Random Forest, εξαγωγή κανόνων μέσω TE2Rules και εισαγωγή υποσυνόλου κανόνων στο εργαλείο ArgEML για ανάπτυξη του μοντέλου επιχειρηματολογίας και στη 2^η περίπτωση έγινε χρήση του μοντέλου XGBoost, εξαγωγή κανόνων μέσω TE2Rules και εισαγωγή υποσυνόλου κανόνων στο εργαλείο ArgEML για ανάπτυξη του μοντέλου επιχειρηματολογίας.

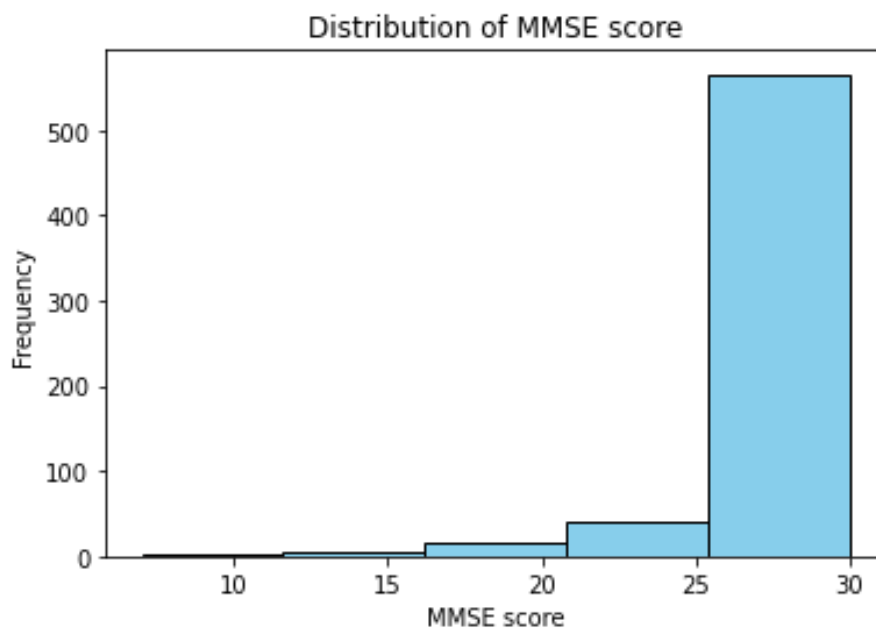
Όπως φαίνεται και στο πίνακα 6.30, τα δύο μοντέλα επιχειρηματολογίας που αναπτύχθηκαν έχουν παρόμοια αποτελέσματα. Φαίνεται ότι τόσο μοντέλο XGBoost όσο και το μοντέλο Random Forest που χρησιμοποιήθηκαν αρχικά για τη εξαγωγή κανόνων είχαν παρόμοια συμπεριφορά και αποδοτικότητα.

6.4 Παρουσίαση και Αποτελέσματα Μοντέλων επιχειρηματολογίας χρησιμοποιώντας δεδομένα ιππόκαμπου και γνωστικού τεστ από βάση δεδομένων OASIS

Τα δεδομένα που χρησιμοποιήθηκαν ήταν:

- Φύλο
- Ηλικία
- Texture Features ιππόκαμπου (HipsASMbaseline, HipsContrastbaseline, HipsCorelationbaseline, HipsVariancebaseline, HipsSumAveragebaseline, HipsSumVariance baseline, HipsEntropybaseline, HipsClusterShadebaseline)
- Όγκος ιππόκαμπου
- Γνωστικό τεστ MMSE Score

Τα δεδομένα χωρίστηκαν σε 5 κατηγορίες [0,1,2,3,4] που αντιστοιχούν στις κατηγορίες [very Low, Low, Nominal, High, Very High]. Εφαρμόστηκε ίση κατανομή (equal distribution) των κατηγοριών σε όλα τα δεδομένα, εκτός από το γνωστικό τεστ το οποίο χωρίστηκε όπως φαίνεται στο πίνακα 6.31.



Διάγραμμα 6.3: Αρχική κατανομή δεδομένων γνωστικού τεστ πριν να χωριστεί σε κατηγορίες

	MMSE Score	Αριθμός Περιπτώσεων
Κατηγορία 0	0-25	50
Κατηγορία 1	25-27	81
Κατηγορία 2	28	94
Κατηγορία 3	29	156
Κατηγορία 4	30	248

Πίνακας 6.31: Οι κατηγορίες που χωρίστηκε το γνωστικό τεστ (MMSE)

Τα train και τα test δεδομένα χωρίστηκαν σε 80% και 20% αντίστοιχα από το εργαλείο ArgEML.

	Training δεδομένα	Test δεδομένα
NC	415	104
NA	88	22
Σύνολο	503 (80%)	126 (20%)

Πίνακας 6.32: Διαχωρισμός των δεδομένων σε training και test δεδομένα

Οι πιο κάτω κανόνες παράχθηκαν από μοντέλο Random Forest, χρησιμοποιώντας το εργαλείο TE2 Rules:

	Rule	Correctness	Coverage
1	MMSE_score > 0.5	91%	93%
2	hippocampusvolumebaseline > 0.5	94%	81%

Πίνακας 6.33: Κανόνες που υποστηρίζουν τις αρνητικές περιπτώσεις

	Rule	Correctness	Coverage
1	MMSE_score <= 0.5 & hippocampusvolumebaseline <= 1.5	97%	6.8%
2	MMSE_score <= 1.5 & hippocampusvolumebaseline <= 0.5 & hipscontrastbaseline > 0.5 & hipsentropybaseline > 0.5	95%	4%
3	MMSE_score <= 1.5 & ageatscreening > 2.5 & hippocampusvolumebaseline <= 2.5 & hipsclustershadebaseline <= 3.5 & hipsentropybaseline <= 1.5	96%	5%

4	ageatscreening > 2.5 & hipposcampusvolumebaseline <= 0.5 & hipscontrastbaseline <= 3.5 & hipscorelationbaseline <= 1.5	100%	2.4%
5	ageatscreening > 2.5 & gender <= 0.5 & hipposcampusvolumebaseline <= 0.5 & hipsentropybaseline <= 1.5 & hipsentropybaseline > 0.5	100%	1.4%
6	MMSE_score <= 1.5 & ageatscreening <= 3.5 & ageatscreening > 1.0 & hipsclustersshadebaseline > 3.5 & hipscorelationbaseline > 1.5	100%	<1%
7	MMSE_score <= 1.5 & ageatscreening > 3.5 & hipsclustersshadebaseline > 1.5 & hipsentropybaseline <= 2.5 & hipssumaveragebaseline <= 2.5 & hipssumvariancebaseline > 1.5	100%	<1%
8	MMSE_score > 1.5 & ageatscreening <= 1.5 & hipposcampusvolumebaseline <= 1.5 & hipsclustersshadebaseline <= 2.5 & hipscontrastbaseline > 3.5 & hipssumvariancebaseline <= 2.5	100%	<1%
9	MMSE_score > 3.5 & gender <= 0.5 & hipposcampusvolumebaseline <= 0.5 & hipscontrastbaseline <= 2.5	100%	<1%
10	MMSE_score <= 2.5 & ageatscreening <= 3.5 & ageatscreening > 2.5 & hipposcampusvolumebaseline <= 0.5 & hipsentropybaseline <= 1.5	100%	1.8%
11	hipposcampusvolumebaseline <= 1.5 & hipposcampusvolumebaseline > 0.5 & hipsasmbaseline <= 1.5 & hipsclustersshadebaseline > 1.5 & hipscontrastbaseline > 1.5 & hipsentropybaseline > 3.5 & hipssumaveragebaseline <= 3.5 & hipsvariancebaseline > 0.5	100%	<1%

Πίνακας 6.34: Κανόνες που υποστηρίζουν τις θετικές περιπτώσεις. Επιλέχθηκαν οι 5 πρώτοι κανόνες για τη ανάπτυξη του 1^{ου} μοντέλου της υποενοότητας αυτής, όπου ο καθένας καλύπτει μεγάλο ποσοστό των δεδομένων και είχε σχετικά καλή ακρίβεια.

Για του κανόνες που θα δοθούν ως είσοδος στο εργαλείο ArgEML για τη ανάπτυξη του μοντέλου επιχειρηματολογίας, χρειάστηκε να γίνει επιλογή από τους παραπάνω κανόνες.

➤ 1ο μοντέλο - Τελικοί κανόνες που επιλέχθησαν

	Rule for negative cases	Correctness	Coverage (of total dataset)
1	MMSE_score > 0.5	93%	90%
2	hipposcampusvolumebaseline > 0.5	81%	94%

Πίνακας 6.35: Κανόνες που επιλέχθηκαν που υποστηρίζουν τις αρνητικές περιπτώσεις

	Rule for positive cases	Correctness	Coverage (of total dataset)
3	MMSE_score < 0.5 & hipposcampusvolumebaseline < 1.5	97%	6.8%
4	MMSE_score < 1.5 & hipposcampusvolumebaseline < 0.5 & hipscontrastbaseline > 0.5 & hipsentropybaseline > 0.5	95%	4%
5	MMSE_score < 1.5 & ageatscreening > 2.5 & hipposcampusvolumebaseline < 2.5 & hipsclustersshadebaseline < 3.5 & hipsentropybaseline < 1.5	96%	5%

6	ageatscreening > 2.5 & hipposcampusvolumebaseline < 0.5 & hipsontrastbaseline < 3.5 & hipscorelationbaseline < 1.5	100%	2.4%
7	ageatscreening > 2.5 & gender < 0.5 & hipposcampusvolumebaseline <= 0.5 & hipsentropybaseline < 1.5 & hipsentropybaseline > 0.5	100%	1.4%
8	MMSE_score < 2.5 & ageatscreening < 3.5 & ageatscreening > 2.5 & hipposcampusvolumebaseline < 0.5 & hipsentropybaseline < 1.5	100%	1.8%

Πίνακας 6.36: Κανόνες που επιλέχθηκαν που υποστηρίζουν τις θετικές περιπτώσεις

- Οι κανόνες που υποστηρίζουν τις αρνητικές περιπτώσεις καλύπτουν το 95% των αρνητικών περιπτώσεων των δεδομένων
- Οι κανόνες που υποστηρίζουν τις θετικών περιπτώσεις καλύπτουν το 74% των θετικών περιπτώσεων των δεδομένων
- Όλοι κανόνες καλύπτουν το 100% των δεδομένων

Πιο κάτω βλέπουμε τις περιπτώσεις που καλύπτει ο κάθε κανόνας και παράλληλα αυτές οι περιπτώσεις καλύπτονται και από κάποιο άλλο κανόνα

	Κανόνας 1	Κανόνας 2
Κανόνας 1	-	489
Κανόνας 2	489	-

Πίνακας 6.37: Επαναληψιμότητα (Redundancy) μεταξύ κανόνων αρνητικών περιπτώσεων

Δηλαδή ο κανόνας 1 και κανόνας 2 καλύπτουν 489 κοινές περιπτώσεις

	Κανόνας 3	Κανόνας 4	Κανόνας 5	Κανόνας 6	Κανόνας 7	Κανόνας 8
Κανόνας 3	0	13	21	10	4	21
Κανόνας 4	13	0	6	7	3	8
Κανόνας 5	21	6	0	9	3	29
Κανόνας 6	10	7	9	0	3	10
Κανόνας 7	4	3	3	3	0	4
Κανόνας 8	21	8	29	10	4	0

Πίνακας 6.38: Επαναληψιμότητα (Redundancy) μεταξύ κανόνων θετικών περιπτώσεων

➤ 2ο μοντέλο: Δεν δόθηκαν κανόνες για αρχικοποίηση θεωρείας στο εργαλείο ArgEML

Το 2^ο μοντέλο επιχειρηματολογίας που αναπτύχθηκε, αξίζει να σημειωθεί ότι είχε αναπτυχθεί εξολοκλήρου από το εργαλείο ArgEML, χωρίς να δοθούν κανόνες που εξάχθηκαν από το εργαλείο TE2Rules.

Αποτελέσματα μοντέλων επιχειρηματολογίας

	Train data=503 Test data=126	Accuracy (LA)	Definite Accuracy	Correct Predictions	False prediction	Dilemmas
ArgEML Model 1 (Rules given from TE2Rules)	Train data	0.82	0.95	68%	3%	29%
	Test data	0.81	0.92	68%	6%	26%
ArgEML Model 2 (No initial rules given from TE2Rules)	Train data	0.88	0.92	83%	7%	10%
	Test data	0.88	0.94	82%	5%	13%

Πίνακας 6.39: Αποτελέσματα μοντέλων επιχειρηματολογίας χρησιμοποιώντας δεδομένα ιππόκαμπου και γνωστικού τεστ από βάση δεδομένων OASIS

Μοντέλο	ADNI δεδομένα				Dilemmas
	AUC	Definite Accuracy (LA)	Sensitivity (Recall)	Specificity	
Μοντέλο Μηχανικής Μάθησης Random Forest	0.75	0.84	0.55	0.95	0%
ArgEML Model 1 (rules given from TE2Rules)	0.85	0.92 (0.81)	0.75	0.95	26%
ArgEML Model 2 (No initial rules given from TE2Rules)	0.65	0.94 (0.88)	0.30	1	13%

Πίνακας 6.40: Αποτελέσματα μοντέλων επιχειρηματολογίας χρησιμοποιώντας δεδομένα ιππόκαμπου και γνωστικού τεστ από βάση δεδομένων OASIS

Παρατηρείται ότι το μοντέλο επιχειρηματολογίας ArgEML 1 είναι αυτό που αποδίδει καλύτερα από ό,τι τα υπόλοιπα μοντέλα. Ωστόσο, το ποσοστό διλημάτων σε αυτό το μοντέλο είναι κοντά στο 26%.

Στο πίνακα 6.41 παρουσιάζονται οι δημοφιλέστερες υποομάδες δεδομένων που προβλέφθηκαν με τη ίδια εξήγηση από το εργαλείο. Παρατηρείται ότι αν και η εξήγηση στις απόλυτες προβλέψεις (θετικό ή αρνητικό) είναι κατανοητή, στις περιπτώσεις που η απάντηση είναι δίλημμα, η εξήγηση δεν είναι και τόσο ξεκάθαρη.

explGroup	prediction	coverage	accuracy	explanation
r3;p18;r2	positive	0.014	0.29	The model derived a definite prediction, positive, because, hips entropy is equal to very low.
p17;r1 r3;p18;r2	dilemma	0.036	-	For this case the model is in dilemma. Various predictions are equally supported: <ul style="list-style-type: none"> • positive, because, hipsentropy baseline is equal to verylow. • negative, because, while the condition(s) [hipsentropy baseline is equal to verylow] suggest the prediction positive, the condition(s) [negative is generally true] that support the prediction negative are stronger.
r1	negative	0.64	0.98	The model derived a definite prediction, negative, because, negative is generally true.
p13;r1 p14;r1d1	dilemma	0.014	-	For this case the model is in dilemma. Various predictions are equally supported: <ul style="list-style-type: none"> • positive, because, while the condition(s) [negative is generally true] suggest the prediction negative, the condition(s) [mmse_score is equal to low] that support the prediction positive are stronger. • negative, because, while the condition(s) [mmse_score is equal to low] suggest the prediction positive, the condition(s) [negative is generally true] that support the prediction negative are stronger.

Πίνακας 6.41: Υποομάδες Δεδομένων που προσδιορίζονται από συγκεκριμένη εξήγηση από το εργαλείο ArgEML

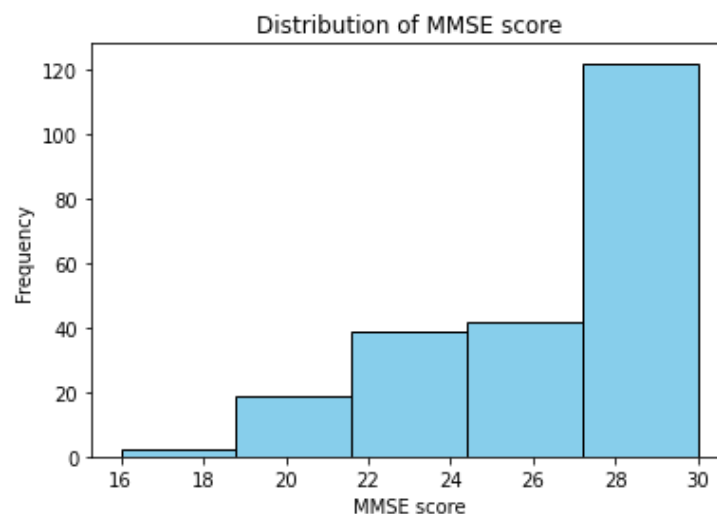
6.5 Παρουσίαση και Αποτελέσματα Μοντέλων επιχειρηματολογίας χρησιμοποιώντας δεδομένα ιππόκαμπου και γνωστικού τεστ από βάση δεδομένων ADNI

Τα δεδομένα που χρησιμοποιήθηκαν ήταν:

- Φύλο
- Ηλικία
- Texture Features ιππόκαμπου (HipsASMbaseline, HipsContrastbaseline, HipsCorelationbaseline, HipsVariancebaseline, HipsSumAveragebaseline, HipsSumVariancebaseline, HipsEntropybaseline, HipsClusterShadebaseline)
- Όγκος ιππόκαμπου
- Γνωστικό τεστ MMSE Score

Όλα τα δεδομένα (εκτός από το γνωστικό τεστ) κατηγοριοποιήθηκαν σε 5 κατηγορίες [0,1,2,3,4] που αντιστοιχούν στις τιμές [Very Low, Low, Nominal, High, Very High].

Τα δεδομένα που αφορούσαν το γνωστικό τεστ (MMSE score) κατηγοριοποιήθηκαν σε 3 κατηγορίες [Low, Nominal, High] λόγω της διακύμανσης τους που είχαν, όπως φαίνεται στο διάγραμμα 6.4.



Διάγραμμα 6.4: Αρχική κατανομή δεδομένων γνωστικού τεστ πριν να χωριστεί σε κατηγορίες από βάση δεδομένων ADNI

	MMSE Score	Αριθμός Περιπτώσεων
Κατηγορία 0	16.0 έως 26.0	133
Κατηγορία 1	26.0 έως 29.0	108
Κατηγορία 2	29.0 έως 30.0	82

Πίνακας 6.42: Κατηγορίες στις οποίες χωρίστηκε το γνωστικό τεστ

Τα train και τα test δεδομένα χωρίστηκαν σε 80% και 20% αντίστοιχα.

	Training δεδομένα	Test δεδομένα
NC	120	31
NA	88	23
Σύνολο	208 (80%)	54 (20%)

Πίνακας 6.43: Χωρισμός δεδομένων σε training και σε test δεδομένα

Οι ακόλουθοι κανόνες παράχθηκαν από μοντέλο Random Forest, χρησιμοποιώντας το

εργαλείο TE2 Rules:

	Rule	Correctness	Coverage of negative cases	Coverage of total dataset
1	MMSE_score > 0	98.4%	97%	60%

Πίνακας 6.44: Κανόνες που υποστηρίζουν τις αρνητικές περιπτώσεις

	Rule	Correct	Coverage of positive cases	Coverage of total dataset
2	MMSE_score < 1	100%	99%	41%
3	MMSE_score > 0 & ageatscreening > 3 & hippocampusvolumebaseline < 3 & hipsasmbaseline < 4 & hipsclustershadebaseline < 2	100%	1%	1%

Πίνακας 6.45: Κανόνες που υποστηρίζουν τις θετικές περιπτώσεις

- Όλοι κανόνες ικανοποιούν το 100% των περιπτώσεων
- Ο κανόνας 1 συγκρούεται με το κανόνα 3 σε δύο μόνο περιπτώσεις

Δοκιμάστηκαν 3 διαφορετικά μοντέλα επιχειρηματολογίας σε αυτή τη υποενοότητα:

- 1^ο Μοντέλο: Επιλέχθηκαν οι 2 πρώτοι κανόνες (χωρίς το κανόνα 3) για αρχικοποίηση ArgEML
- 2^ο Μοντέλο: Επιλέχθηκαν και οι 3 κανόνες για αρχικοποίηση του ArgEML
- 3^ο Μοντέλο: Δεν δόθηκαν κανόνες για αρχικοποίηση

Model	Train data=208 Test data=54	Accuracy (LA)	Definite Accuracy	Correct Predictions	False predictions	Dilemmas
ArgEML Model 1	Test data	0.98	1	200	0	4%
	Test data	0.93	0.94	51	3	1%
ArgEML Model 2	Train data	0.99	0.99	204	1	1.4%
	Test data	0.96	0.96	52	2	0%
ArgEML Model 3	Train data	0.98	0.99	202	1	2%
	Test data	0.97	0.98	52	1	2%

Πίνακας 6.46: Σύγκριση μοντέλων επιχειρηματολογίας χρησιμοποιώντας δεδομένα ιππόκαμπου και γνωστικού τεστ από βάση δεδομένων ADNI

Μοντέλο	ADNI δεδομένα				Dilemmas
	AUC	Definite Accuracy (LA)	Sensitivity	Specificity	
Μοντέλο Μηχανικής Μάθησης Random Forest	0.95	0.94 (0.94)	1	0.89	0%
ArgEML Model 1	0.94	0.94 (0.93)	0.90	0.97	1%
ArgEML Model 2	0.96	0.96 (0.96)	0.96	0.97	0%
ArgEML Model 3	0.98	0.98 (0.97)	1	0.97	2%

Πίνακας 6.47: Σύγκριση μοντέλων επιχειρηματολογίας και μηχανικής μάθησης χρησιμοποιώντας δεδομένα ιππόκαμπου και γνωστικού τεστ από βάση δεδομένων ADNI

Παρατηρούμε ότι το μοντέλο 3 επιχειρηματολογίας στο οποίο δεν δόθηκαν κανόνες για αρχικοποίηση του μοντέλου ArgEML έχει τα καλύτερα αποτελέσματα. Ωστόσο, οι τελικές εξηγήσεις που δίνονται από το συγκεκριμένο μοντέλο 3 δεν ήταν πολύ ξεκάθαρες. Το μοντέλο 2 επιχειρηματολογίας φαίνεται να αποδίδει και αυτό καλά δίνοντας και αρκετά

καλές επεξηγήσεις όπως φαίνονται στο πίνακα 6.48. Μάλιστα, αποδίδει καλύτερα από το μοντέλο μηχανικής μάθησης Random Forest αν ληφθούν υπόψη όλες οι μετρικές αξιολόγησης έχοντας 0% ποσοστό διλημμάτων στο σετ αξιολόγησης (test set).

Μέσω του εργαλείου ArgEML και των αποτελεσμάτων, μπορούμε να αναλύσουμε τη θεωρία που αναπτύχθηκε και να δούμε με ποιους κανόνες έγιναν οι προβλέψεις. Πολλές προβλέψεις έγιναν με τον ίδιο τρόπο και δηλαδή μέσω του ίδιου κανόνα από το εργαλείο. Στο πίνακα 6.48 παρουσιάζονται οι δημοφιλέστερες υποομάδες δεδομένων που προβλέφθηκαν με τη ίδια εξήγηση από το εργαλείο για το μοντέλο 2 που περιεγράφηκε πιο πάνω.

explGroup		prediction	coverage	accuracy	explanation
1	r3	negative	0.567	1	The model derived a definite prediction, negative, because, mmse_score is equal to one nominal, high.
2	r1d1 r1	dilemma	0.0048	-	For this case the model is in dilemma. Various predictions are equally supported: <ul style="list-style-type: none"> • positive, because, mmse_score is equal to low. • negative, because, hippocampusvolumebaseline is equal to veryhigh AND mmse_score is equal to low.
3	r3 r2	dilemma	0.0096	-	For this case the model is in dilemma. Various predictions are equally supported: <ul style="list-style-type: none"> • positive, because, ageatscreening is equal to veryhigh AND hippocampusvolumebaseline is equal to one of verylow,low,nominal AND hipsasmbaseline is equal to one of verylow,low,nominal,high AND hipsclustershadebaseline is equal to one of verylow,low. <ul style="list-style-type: none"> • negative, because, mmse_score is equal to one of nominal,high.
4	r1d1	negative	0.0048	1	The model derived a definite prediction, negative, because, hippocampusvolumebaseline is equal to veryhigh AND mmse_score is equal to verylow.
5	r1	positive	0.41	0.99	The model derived a definite prediction, positive, because, mmse_score is equal to verylow.

Πίνακας 6.48: Υποομάδες Δεδομένων που προσδιορίζονται από συγκεκριμένη εξήγηση από το εργαλείο ArgEML

ΚΕΦΑΛΑΙΟ 7: Συμπεράσματα

Πιο κάτω, παρουσιάζονται συνοπτικά κάποια συμπεράσματα και αναφέρονται χρήσιμα θέματα που μπορούν να αποτελέσουν βάση για περαιτέρω μελέτη στο μέλλον.

- Σε αυτή τη παράγραφο, γίνεται μια σύνοψη της σύγκρισης των δεδομένων που χρησιμοποιήθηκαν στη μελέτη αυτή, τα οποία είχαν εξαχθεί από τις βάσεις δεδομένων OASIS και ADNI. Στο σετ δεδομένων από τη βάση OASIS τα δεδομένα παρατηρήθηκε ότι ήταν ανισόρροπα (imbalance) καθώς υπήρχε μειονότητα των περιπτώσεων Alzheimer σε σχέση με τις κανονικές περιπτώσεις. Αυτό, παρ' όλων των τεχνικών που χρησιμοποιήθηκαν (oversampling/undersampling) δυσκόλεψε αρκετά τη ανάπτυξη των μοντέλων κυρίως στο να επιτύχουν υψηλή ακρίβεια. Τα δεδομένα που προήλθαν από τη βάση ADNI ήταν πιο ισορροπημένα. Στα δεδομένα από τη βάση OASIS, το MSSE score δεν ήταν ομοιόμορφα κατανομημένο όσο ήταν στα δεδομένα ADNI. Αυτό επηρέασε επίσης σημαντικά τα αποτελέσματα. Και για αυτό τα μοντέλα που αναπτύχθηκαν μέσω των δεδομένων βάσης ADNI χρησιμοποιώντας το γνωστικό τεστ έδωσαν πολύ καλύτερα αποτελέσματα σε σύγκριση με τα δεδομένα της βάσης OASIS.
- Όπως φαίνεται και στα αποτελέσματα, τα σημαντικότερα χαρακτηριστικά που μας βοηθούν να διαγνώσουμε τη νόσο Alzheimer είναι όγκος του ιππόκαμπου, ο όγκος του ενδορινικού φλοιού και το γνωστικό τεστ. Ειδικά το γνωστικό τεστ έχει πολύ μεγάλη επιρροή στο διαχωρισμό κανονικών περιπτώσεων από περιπτώσεις AD. Αντίθετα τα χαρακτηριστικά υφής (texture features) δεν έχουν τόσο μεγάλη επιρροή στο να βελτιώσουν τα μοντέλα. Ωστόσο, στη περίπτωση ασθενούς με AD τα χαρακτηριστικά αυτά έχουν πιο πρόωρη μεταβολή από ότι ο όγκος, έτσι μέσω αυτών μπορούμε πιο γρήγορα να διαγνώσουμε τη ασθένεια.

- Τα μοντέλα επιχειρηματολογίας που αναπτύσσονται στο εργαλείο ArgEML βελτιώνουν την ακρίβεια των αποτελεσμάτων σε σχέση με τα μοντέλα μηχανικής μάθησης. Αν και όπως αναφέρθηκε ένα ποσοστό των προβλέψεων τους είναι διλήμματα, χωρίς να δίνεται δηλαδή ξεκάθαρη απάντηση. Επίσης, σε κάθε πρόβλεψη ή δίλημμα που δίνει το μοντέλο επιχειρηματολογίας εξηγεί το λόγο που δόθηκε η συγκεκριμένη απάντηση, παρουσιάζοντας λογικά επιχειρήματα. Αυτό αποτελεί βασικό πλεονέκτημα, ειδικά στο τομέα της ιατρικής, όπου ο ειδικός (medical expert) χρειάζεται να ξέρει τους λόγους που δόθηκε η συγκεκριμένη απάντηση ώστε να είναι σίγουρος για τη διάγνωση. Έτσι ο ειδικός γνωρίζοντας τα επιχειρήματα που υποστηρίζουν τη συγκεκριμένη πρόβλεψη του εργαλείου μπορεί να τα κρίνει και αναλόγως να συμφωνήσει ή να διαφωνήσει με τη απάντηση. Η συνεργασία αυτή, μεταξύ εργαλείου επιχειρηματολογίας και ειδικού ενδέχεται να μειώσει τις λανθασμένες διαγνώσεις σε σύγκριση με τα μοντέλα μηχανικής μάθησης που λειτουργούν ως μαύρα κουτιά. Ένα άλλο παράδειγμα, είναι ότι ακόμα και αν το μοντέλο μηχανικής μάθησης έχει ακρίβεια 95% , ο γιατρός δεν μπορεί πάλι να το εμπιστευτεί χωρίς να κοιτάξει τους λόγους που αποφασίστηκε η συγκεκριμένη απάντηση, καθώς υπάρχει πιθανότητα (έστω και 5%) η απάντηση να είναι λανθασμένη και ο ασθενής να μην διαγνωσθεί έγκαιρα αν είναι θετικός.
- Γενικά, είναι καλό να υπάρχει μια καλή ισορροπία (trade off) μεταξύ του ποσοστού των λανθασμένων προβλέψεων και του ποσοστού των διλημάτων που δίνουν τα μοντέλα επιχειρηματολογίας. Αξίζει να σημειωθεί, ότι είναι δυνατό μέσω των ρυθμίσεων που προσφέρει το εργαλείο ArgEML, αυτά τα ποσοστά να προσαρμοστούν και να ισορροπηθούν αναλόγως με τις ανάγκες και το πεδίο στο οποίο εφαρμόζεται το μοντέλο επιχειρηματολογίας.

- Μια πρόκληση που υπήρξε στη μελέτη αυτή ήταν ότι λόγω της μειονότητας των θετικών περιπτώσεων, τα μοντέλα που αναπτύσσονταν είχαν χαμηλό Sensitivity. Αυτό στα μοντέλα μηχανικής μάθησης διορθώθηκε εφαρμόζοντας τη τεχνική παραμετροποίησης Parametrize tuning, στην οποία οι παράμετροι (π.χ. max depth, estimators) επιλέχθηκαν κατάλληλα ώστε το μοντέλο που αναπτύσσεται να εστιάζει σε υψηλή τιμή (score) sensitivity. Μια άλλη μέθοδος που χρησιμοποιήθηκε σε αυτά τα μοντέλα μηχανικής μάθησης, ήταν να αλλάξω το όριο αποφάσεων που χρησιμοποιείται για την ταξινόμηση των προβλέψεων. Συγκεκριμένα, για να θεωρήσω μια πρόβλεψη ως θετική, όρισα ένα χαμηλότερο κατώφλι. Για παράδειγμα, αν η πιθανότητα μιας πρόβλεψης να είναι θετική ήταν μεγαλύτερη από αυτό το όριο (π.χ. 0.30), τότε αυτή η πρόβλεψη έδινε θετικό αποτέλεσμα (ασθένεια AD). Αυτό βοήθησε το μοντέλο να ανιχνεύει περισσότερες θετικές περιπτώσεις, αυξάνοντας έτσι τη μετρική sensitivity. Στα μοντέλα επιχειρηματολογίας οι προηγούμενες τεχνικές δεν μπορούσαν να υλοποιηθούν, καθώς το εργαλείο ArgEML λειτουργεί διαφορετικά. Η μετρική Sensitivity σε κάποιες περιπτώσεις ήταν χαμηλότερη από 50%. Για να αυξηθεί η μετρική Sensitivity σε αυτές τις περιπτώσεις, χρειάστηκε να επιλεγθούν και να εισαχθούν στο εργαλείο ArgEML, κατάλληλοι και περισσότεροι κανόνες που υποστήριζαν τις θετικές περιπτώσεις. Επίσης, για τον ίδιο λόγο χρειάστηκε να μην χρησιμοποιηθούν κάποιοι κανόνες που είχαν χαμηλή ακρίβεια και υποστήριζαν τις αρνητικές περιπτώσεις.
- Τέλος, όπως αναφέρθηκε σε αυτή τη μελέτη έγινε ανάπτυξη μοντέλων μηχανικής μάθησης και μοντέλων επιχειρηματολογίας βασισμένα στα μοντέλα μηχανικής μάθησης, με σκοπό τη πρόβλεψη της ασθένειας Alzheimer. Θα έχει ενδιαφέρον μελλοντικά να χρησιμοποιηθούν μοντέλα μηχανικής μάθησης με επιχειρηματολογία για το

διαχωρισμό των περιπτώσεων του MCI και περιπτώσεων NA, κάτι το οποίο δεν έχει μελετηθεί σε πολλές μελέτες.

Βιβλιογραφία

- [1] Emmady PD, Schoo C, Tadi P. Major Neurocognitive Disorder (Dementia) [Updated 2022 Nov 19]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 Jan, <https://www.ncbi.nlm.nih.gov/books/NBK557444/>
- [2] Alzheimer Association, <https://www.alz.org/alzheimers-dementia>
- [3] G. W. Van Hoesen, B. T. Hyman, and A. R. Damasio, "Entorhinal cortex pathology in Alzheimer's disease," *Hippocampus*, vol. 1, no. 1, pp. 1–8, Jan. 1991, doi: 10.1002/hipo.450010102
- [4] X. Li, J. Jiao, S. Shimizu, I. Jibiki, K.-I. Watanabe, and T. Kubota, Correlations between atrophy of the entorhinal cortex and cognitive function in patients with Alzheimer's disease and mild cognitive impairment, *Psychiatry Clin. Neurosci.*, vol. 66, no. 7, pp. 587–593, Dec. 2012, doi: 10.1111/pcn.12002.
- [5] M. F. Folstein, S. E. Folstein, and P. R. McHugh, "Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician, *J Psychiatr Res*, vol. 12, no. 3, pp. 189–198, Nov. 1975.
- [6] Leandrou S, Lamnisos D, Bougias H, Stogiannos N, Georgiadou E, Achilleos KG, Pattichis CS and Alzheimer's Disease Neuroimaging Initiative (2023) A cross-sectional study of explainable machine learning in Alzheimer's disease: diagnostic classification using MR radiomic features. *Front. Aging Neurosci.* 15:1149871. doi: 10.3389/fnagi.2023.1149871
- [7] Leandrou S, Quantitative Texture analysis in MR Imaging in the assessment of Alzheimer's disease, PHD Dissertation, 2021
- [8] Leandrou, S., Lamnisos, D., Kyriacou, P. A., Constanti, S. & Pattichis, C. S. (2020). Comparison of 1.5 T and 3 T MRI hippocampus texture features in the assessment of Alzheimer's disease. *Biomedical Signal Processing and Control*, 62, 102098. doi: 10.1016/j.bspc.2020.102098
- [9] Leandrou S, Lamnisos D, Mamais I, Kyriacou PA and Pattichis CS, Assessment of Alzheimer's Disease Based on Texture Analysis of the Entorhinal Cortex., 2020, *Front. Aging Neurosci.* 12:176. doi: 10.3389/fnagi.2020.00176
- [10] K. G. Achilleos, S. Leandrou, N. Prentzas, P. A. Kyriacou, A. C. Kakas and C. S. Pattichis, "Extracting Explainable Assessments of Alzheimer's disease via Machine Learning on brain MRI imaging data," 2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE), Cincinnati, OH, USA, 2020, pp. 1036-1041, doi: 10.1109/BIBE50027.2020.00175
- [11] Alzheimer's Disease Neuroimaging Initiative (ADNI), <https://adni.loni.usc.edu>, 2024
- [12] Washington University in St. Louis, <https://sites.wustl.edu/oasisbrains/>, 2024

- [13] R. Cuingnet et al., Automatic classification of patients with Alzheimer’s disease from structural MRI: A comparison of ten methods using the ADNI database, *NeuroImage*, vol. 56, no. 2, pp. 766–781, May 2011, doi: 10.1016/j.neuroimage.2010.06.013.
- [14] FreeSurfer tool, <https://surfer.nmr.mgh.harvard.edu/>
- [15] Berthold M. R., Cebron N., Dill F., Gabriel T. R., Kötter T., Meinl T., et al. (2008). KNIME: The Konstanz information miner, in *Data analysis, machine learning and applications, studies in classification, data analysis, and knowledge organization*, eds Preisach C., Burkhardt H., Schmidt-Thieme L., Decker R. (Berlin: Springer;), 319–326.
- [16] Haralick R. M., Shanmugam K., Dinstein I. (1973). Textural Features for Image Classification. *IEEE Trans. Syst. Man Cybernet. SMC* 3 610–621. 10.1109/TSMC.1973.4309314
- [17] J.-H. Cai et al., “Magnetic Resonance Texture Analysis in Alzheimer’s disease,” *Academic Radiology*, Feb. 2020, doi: 10.1016/j.acra.2020.01.006.
- [18] M. F. Folstein, S. E. Folstein, and P. R. McHugh, Mini-mental state. A practical method for grading the cognitive state of patients for the clinician, *J Psychiatr Res*, vol. 12, no. 3, pp. 189–198, Nov. 1975.
- [19] B. Dubois et al., Research criteria for the diagnosis of Alzheimer’s disease: revising the NINCDS-ADRDA criteria, *Lancet Neurol*, vol. 6, no. 8, pp. 734–746, Aug. 2007, doi: 10.1016/S1474-4422(07)70178-3.
- [20] Tangirala, Suryakanthi. “Evaluating the Impact of GINI Index and Information Gain on Classification using Decision Tree Classifier Algorithm.” *International Journal of Advanced Computer Science and Applications* 11 (2020): n. pag.
- [21] Scikit-learn, Machine Learning in Python <https://scikit-learn.org/>
- [22] S. J. Rigatti, ‘Random Forest’, *Journal of Insurance Medicine*, vol. 47, no. 1, pp. 31–39, Jan. 2017, doi: 10.17849/in-sm-47-01-31-39.1.
- [23] W. Dong, Y. Huang, B. Lehane, and G. Ma, ‘XGBoost algorithm-based prediction of concrete electrical resistivity for structural health monitoring’, *Automation in Construction*, vol. 114, p. 103155, Jun. 2020, doi: 10.1016/j.autcon.2020.103155.
- [24] Carrera, Á., Iglesias, C.A. A systematic review of argumentation techniques for multi-agent systems research. *Artif Intell Rev* 44, 509–535 (2015). <https://doi.org/10.1007/s10462-015-9435-9>
- [25] Prentzas N, *Advancing AI Explainability with an Argumentation-Based Machine Learning Approach and Its Application in the Medical Domain*, PhD Dissertation, 2024,
- [26] A. C. Kakas, P. Moraitis, and N. I. Spanoudakis, "GORGIAS: Applying argumentation," *Argument & Computation*, t. 10, αρ. 1, pg. 55-81, 2018. Available doi: 10.3233/AAC-181006

[27] K. Juottonen, M. P. Laakso, K. Partanen, and H. Soininen, "Comparative MR analysis of the entorhinal cortex and hippocampus in diagnosing Alzheimer disease," *AJNR Am J Neuroradiol*, vol. 20, no. 1, pp. 139–144, Jan. 1999.

[28] Valentina Zelaya Mendizábal, Marc Boullé, Fabrice Rossi, Fast and fully-automated histograms for large-scale data sets, Volume 180,2023, <https://doi.org/10.1016/j.csda.2022.107668>