

Ελληνικό Μεσογειακό Πανεπιστήμιο

Πολυτεχνική Σχολή

Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Ανάπτυξη ενός Ασφαλούς Συστήματος Υποστήριξης Κλινικών Αποφάσεων με χρήση τεχνολογιών Μηχανικής Μάθησης και Blockchain

Διπλωματική Εργασία

Μπιτζαράκης Γεώργιος (TH20074)

Υπεύθυνος Επιβλέπων: Δρ. Κωνσταντίνος Καραμπίδης



Εκτός αν ορίζεται διαφορετικά, το περιεχόμενο της παρούσας διατριβής διατίθεται υπό την άδεια Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0).
Οι απόψεις και τα συμπεράσματα που περιέχονται στο παρόν έγγραφο εκφράζουν αποκλειστικά τον συγγραφέα και δεν πρέπει να εκλαμβάνονται ως εκπροσώπηση των επίσημων θέσεων του Τμήματος Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών της Σχολής Μηχανικών του Ελληνικού Μεσογειακού Πανεπιστημίου.

Ευχαριστίες

Το να σπουδάζει κανείς παράλληλα με την εργασία του απαιτεί θυσίες, τόλμη, αφοσίωση, υπομονή και επιμονή στον στόχο. Όμως πέρα από όσα είναι κανείς διατεθειμένος να κάνει για φτάσει εκεί που θέλει, η ηθική και πρακτική στήριξη των ανθρώπων του δεν πρέπει να περνάει απαρατήρητη διότι έχει καθοριστικό ρόλο στην επιτυχή έκβαση μιας τέτοιας προσπάθειας, ενώ η βαρύτητα που της δίνει κάθε φορά εμπλουτίζει το ταξίδι της μάθησης με εμπειρίες που τον συνοδεύουν για πάντα.

Καταρχάς, θα ήθελα να ευχαριστήσω το εκπαιδευτικό προσωπικό του Ελληνικού Μεσογειακού Πανεπιστημίου για τη συνεχή υποστήριξή τους σε όλη τη διάρκεια των σπουδών μου. Ιδιαίτερες ευχαριστίες οφείλω στον επιβλέποντα καθηγητή μου, Δρ. Κωνσταντίνο Καραμπίδη, για την καθοδήγηση, την ενθάρρυνση και την εμπιστοσύνη που έδειξε καθ' όλη της πορείας αυτής της εργασία, αλλά κυρίως για την υπομονή του στις αναπόφευκτες περιόδους αδράνειας που προέκυψαν κατά την προετοιμασία της διπλωματικής διατριβής. Η ευελιξία και κατανόηση του απέναντι στις ανάγκες της συγκεκριμένης έρευνας αποτέλεσαν πολύτιμα εφόδια και μαθήματα ταυτόχρονα για εμένα.

Πάνω απ' όλα όμως, θέλω να ευχαριστήσω αυτούς που ήταν από την αρχή εκεί, την οικογένεια μου που καθημερινά αποτελεί στήριγμα από τα πρώτα μου βήματα σ 'αυτή τη διαδρομή και ανθρώπους που εμφανίστηκαν μετά, την κοπέλα μου, Μαριάννα, για την αμέριστη υπομονή, τη πίστη και αγάπη της σε στιγμές που οι απαιτήσεις της καθημερινότητας δεν άφηναν πολλά περιθώρια για τίποτε άλλο. Κλείνοντας, ιδιαίτερη ευχαρίστηση οφείλω να απευθύνω στους ανθρώπους που βρίσκονται καθημερινά στο κύκλο μου και έχω την τιμή να αποτελώ κι εγώ κομμάτι της δικής τους ζωής· χωρίς τη δική τους παρουσία, το ταξίδι αυτό θα ήταν φτωχότερο.

Περίληψη

Η παρούσα διπλωματική διατριβή επικεντρώνεται στη σχεδίαση και την υλοποίηση ενός καινοτόμου Συστήματος Υποστήριξης Κλινικών Αποφάσεων που συνδυάζει τρεις τεχνολογίες αιχμής, τη Τεχνητή Νοημοσύνη μέσω του κλάδου της Μηχανικής Μάθησης, τη νεοεμφανιζόμενη τεχνολογία κατανεμημένων συστημάτων Blockchain και σύγχρονα εργαλεία δημιουργίας διαδικτυακών εφαρμογών. Για την πειραματική παρουσίαση της πραγματοποιήθηκε εστίαση στο κλινικό σύνδρομο της Καρδιακής Ανεπάρκειας, η οποία ακόμη και σήμερα αποτελεί μια από τις κρισιμότερες αιτίες νοσηρότητας και θνησιμότητας παγκοσμίως, καθιστώντας την πρόωρη πρόγνωση του κινδύνου σημαντική για τη κλινική καθημερινότητα. Παρά την ακρίβεια των σύγχρονων αλγορίθμων Μηχανικής Μάθησης, η υιοθέτηση τους από την κλινική κοινότητα παρεμποδίζεται εξαιτίας του προβλήματος του «μαύρου κουτιού», προκαλώντας την ανησυχία των κλινικών σχετικά με την ερμηνευσιμότητα, εγκυρότητα και ακεραιότητα των ιατρικών δεδομένων.

Για την αντιμετώπιση των προαναφερόμενων προκλήσεων, στην παρακάτω εργασία προτείνονται λύσεις με την μορφή αρχιτεκτονικής τριών επιπέδων. Για την κάλυψη της εγκυρότητας, αξιοποιήθηκε η διεθνώς αναγνωρισμένη βάση ιατρικών δεδομένων MIMIC-IV για την εκπαίδευση τεσσάρων μοντέλων, με τη βελτιστοποίηση των υπερπαραμέτρων να πραγματοποιείται με τη βιβλιοθήκη Optuna. Η ερμηνευσιμότητα επιτυγχάνεται με τη ενσωμάτωση της βιβλιοθήκης εξηγήσιμης τεχνητής νοημοσύνης SHAP που παρέχει στους ιατρούς διαφανείς επεξηγήσεις στις διαγνώσεις σε τοπικό και καθολικό επίπεδο. Παράλληλα για την διασφάλιση της ακεραιότητας και ιδιωτικότητας, ενσωματώνεται η τεχνολογία Blockchain μέσω έξυπνων συμβολαίων στη γλώσσα προγραμματισμού Solidity, εφαρμόζοντας έλεγχο πρόσβασης βάσει ρόλων και ιδιαίτερη έμφαση δίνεται στη συμμόρφωση με τον Γενικό Κανονισμό Προστασίας Δεδομένων, που χρησιμοποιήθηκαν δομές δεδομένων όπως τα Merkle δέντρα για την εξισορρόπηση μεταξύ της προστασίας και τα ταχύτητας των συναλλαγών.

Τα αποτελέσματα της υπάρχουσας προτεινόμενης πλατφόρμας καταδεικνύουν την ταυτόχρονη επίτευξη υψηλής προγνωστικής ακρίβειας, της κλινικής ερμηνευσιμότητας και την αμετάβλητη καταγραφή δεδομένων σε ένα ενιαίο σύστημα λήψης αποφάσεων, δίνοντας επιπλέον τη δυνατότητα επέκτασης του σε περισσότερες ιατρικές εφαρμογές.

Λέξεις Κλειδιά: Καρδιακή Ανεπάρκεια, Μηχανική Μάθηση, Εξηγήσιμη Τεχνητή Νοημοσύνη, SHAP, Blockchain, Έξυπνα Συμβόλαια, Σύστημα Υποστήριξης Κλινικών Αποφάσεων, MIMIC-IV, Γενικός Κανονισμός Προστασίας Δεδομένων, Merklee Δέντρο.

Abstract

This master's thesis focuses on the design and implementation of an innovative Clinical Decision Support System that integrates three cutting-edge technologies: Artificial Intelligence through Machine Learning, the emerging field of Blockchain distributed systems, and modern web application development tools. The experimental focus of this research is on Heart Failure, a clinical syndrome that remains a leading cause of morbidity and mortality worldwide, making early risk prediction vital for daily clinical practice. Despite the high accuracy of modern Machine Learning algorithms, their clinical adoption is often hindered by the "black box" problem, raising concerns among healthcare professionals regarding the explainability, validity, and integrity of medical data.

To address these challenges, this thesis proposes a solution in the form of a three-layer architecture. To ensure validity, the internationally recognized MIMIC-IV medical database was utilized to train four candidate models, with hyperparameter optimization performed using the Optuna library. Explainability is achieved through the integration of the SHAP framework, providing clinicians with transparent, local, and global explanations for each diagnosis. Concurrently, to ensure data integrity and privacy, Blockchain technology is integrated via Smart Contracts authored in the Solidity programming language. The system implements Role-Based Access Control and places significant emphasis on compliance with the General Data Protection Regulation, utilizing data structures such as Merkle Trees to achieve a balance between data protection and transaction efficiency.

The results of the proposed platform demonstrate the simultaneous achievement of high predictive accuracy, clinical explainability, and immutable data logging within a unified decision-making environment. Furthermore, this architecture offers significant scalability and potential for expansion to a wider range of medical applications.

Keywords: Heart Failure, Machine Learning, Explainable AI, SHAP, Blockchain, Smart Contracts, Clinical Decision Support System, MIMIC-IV, GDPR, Merkle Trees.

Περιεχόμενα

Κεφάλαιο 1: Εισαγωγή....20
1.1 Πλαίσιο και Κίνητρο....20
1.1.1Ψηφιακός Μετασχηματισμός της Υγείας....20
1.1.2 Τεχνητή Νοημοσύνη: Όρισμος και Εφαρμογές....20
1.1.3 Τεχνητή Νοημοσύνη στην Ιατρική....23
1.1.4 Τεχνολογία Blockchain και ακεραιότητα ιατρικών δεδομένων....24
1.2 Ορισμός του Προβλήματος....26
1.2.1 Στόχοι Εργασίας....27
1.3 Συνεισφορά της Εργασίας....28
1.4 Δομή της Εργασίας....28
Κεφάλαιο 2 : Ορισμοί και Βιβλιογραφική Ανασκόπηση....29
2.1 Καρδιακή Ανεπάρκεια: Κλινικό Πλαίσιο....29
2.1.2 Φαρμακευτική Θεραπεία για ασθενείς με καρδιακή ανεπάρκεια μειωμένου κλάσμα εξώθησης (HFrEF)....33
2.1.3 Θεραπεία για ασθενείς με καρδιακή ανεπάρκεια ελαφρώς μειωμένου κλάσμα εξώθησης (HFmrEF)....35
2.1.4 Θεραπεία για ασθενείς με καρδιακή ανεπάρκεια διατηρημένου κλάσματος εξώθησης (HFpEF)....36
2.1.5 Επιδημιολογία της Καρδιακής Ανεπάρκειας....37
2.1.6 Αιτιολογία και Παράγοντες Κινδύνου....38
2.1.7 Βιοδείκτες: Νατριουρητικά Πεπτίδια και Τροπονίνες....40
2.2 Μηχανική Μάθηση και Ερμηνεύσιμη Τεχνητή Νοημοσύνη στην Κλινική Πράξη....40
2.2.1 Η Μηχανική Μάθηση στην Υγειονομική Περίθαλψη....41
2.2.2 Μηχανική Μάθηση στη πρόβλεψη της Καρδιακής Ανεπάρκειας....45
2.2.3 Το πρόβλημα του “Μαύρου Κουτιού“....46
2.2.4 Μέθοδοι Εξηγήσιμης Τεχνητής Νοημοσύνης....47
2.2.5 Η εφαρμογή της Εξηγήσιμης Τεχνητής Νοημοσύνης στην πρόβλεψη της Καρδιακής Ανεπάρκειας....49
2.2.6 Η Εμπιστοσύνη των κλινικών στα συστήματα Εξηγήσιμης Τεχνητής Νοημοσύνης...49
2.3 Η Τεχνολογία Blockchain στην Υγεία....53

2.3.1 Γιατί να χρησιμοποιήσει κάποιος την τεχνολογία Blockchain στην Υγεία; 53
2.3.2 Blockchain και Ηλεκτρονικά Ιατρικά Αρχεία 55
2.3.3 Η συμβολή των έξυπνων συμβολαίων στο χώρο της υγείας 55
2.3.4 Blockchain και GDPR: Η σύγκρουση της αμεταβλητότητας και του δικαιώματος διαγραφής.....58
2.3.5 Περιορισμοί της τεχνολογίας Blockchain και η σύνδεση της με τα υπάρχουσα CDSS 59
2.4 Συστήματα Υποστήριξης Κλινικών Αποφάσεων (CDSS).....61
2.4.1 Ορισμός και Ταξινόμηση 61
2.4.2 Κλινική Αποτελεσματικότητα των CDSS.....62
2.4.3 Εμπόδια υιοθέτησης των CDSS στην κλινική καθημερινότητα 63
2.4.4 Συστήματα CDSS βασισμένα σε Τεχνητή Νοημοσύνη 64
2.4.5 Πρότυπα διαλειτουργικότητας σε Ηλεκτρονικούς Φακέλους Υγείας.....65
2.5 Υπάρχουσα Κατάσταση 65
2.5.1 Συναφείς Εργασίες.....65
2.5.2 Ερευνητικά Κενά.....66
2.5.3 Συγκριτικός Πίνακας.....66
Κεφάλαιο 3: Θεωρητικό Υπόβαθρο.....67
3.1 Αλγόριθμοι Μηχανικής Μάθησης.....67
3.1.1 Λογιστική Παλινδρόμηση (Logistic Regression).....67
3.1.2 Αλγόριθμος Τυχαίου Δάσους (Random Forest) 70
3.1.3 Αλγόριθμος XGBoost 72
3.1.4 Αλγόριθμος LightGBM.....73
3.1.5 Βαθμονόμηση Πιθανότητων 73
3.2 Μετρικές Αξιολόγησης.....74
3.2.1 Μετρικές Διακριτικής Ικανότητας.....75
3.2.2 Μετρικές Βαθμονόμησης.....76
3.2.3 Μετρική Καθαρού Οφέλους.....77
3.2.4 Κλινική Χρησιμότητα.....77
3.3 SHAP.....77
3.3.1 Θεωρία Shapley Values.....78

3.3.2 Υπολογιστική Στρατηγική 78
3.3.3 Οπτικοποίηση και Σύνθεση Ερμηνειών 79
3.4 Τεχνολογία Blockchain 79
3.4.1 Βασικές Αρχές Λειτουργίας 80
3.4.2 Στρατηγική Αποθήκευσης και GDPR Compliance 80
3.5 Έξυπνα Συμβόλαια 81
3.5.1 Γλώσσα Προγραμματισμού Solidity 82
3.5.2 Έλεγχος Προσβασιμότητας 83
3.5.3 Συμβάντα και Βελτιστοποίηση Κόστους 83
3.6 Αρχιτεκτονικές Διαδικτυακών Εφαρμογών 84
3.6.1 SPA και REST API 84
Κεφάλαιο 4: Μεθοδολογία & Αρχιτεκτονική 84
4.1 Επισκόπηση Συστήματος 85
4.1.1 Αρχιτεκτονική και ροή εργασίας των τριών επιπέδων 86
4.1.2 Πρωτόκολλα και ροές επικοινωνίας 86
4.2 Σύνολο Δεδομένων (MIMIC-IV) 89
4.2.1 Πηγή Δεδομένων Εκπαίδευσης και Ηθική Ζητήματα 89
4.2.2 Αξιοποιημένοι πίνακες 90
4.2.3 Μεταβλητή Στόχου (label) 92
4.3 Αρχιτεκτονική Συστήματος 95
4.3.1 Διάγραμμα Συνιστωσών 95
4.3.2 Διάγραμματα Ροής Δεδομένων 98
4.3.3 Διαγράμματα Ακολουθίας 104
Κεφάλαιο 5: Πειραματικό Μέρος 113
5.1 Προεπεξεργασία Δεδομένων 113
5.1.1 Δημιουργία Γενικού Πίνακα Νοσηλειών 113
5.1.2 Εξαγωγή Εργαστηριακών Τιμών σε πλαίσιο 24 Ωρών 115
5.1.3 Δημιουργία Σύνθετων Ιατρικών Δεικτών 117
5.1.4 Χρονικά Ασφαλής Κατασκευή Ιστορικού Συννοσηροτήτων 120
5.1.5 Τελική Διαμόρφωση του Συνόλου Εκπαίδευσης 122
5.2 Εκπαίδευση Μοντέλων Μηχανικής Μάθησης 124

5.2.1 Αρχιτεκτονική της Ροής Εκπαίδευσης 124
5.2.2 Διαμόρφωση των Τεσσάρων Υποψήφιων Μοντέλων 127
5.2.3 Μηχανισμός Πρόωρου Τερματισμού των Ενισχυμένων Μοντέλων 130
5.2.4 Βελτιστοποίηση Υπερπαραμέτρων μέσω Optuna 132
5.2.5 Επιλογή Μοντέλου Πρωταθλητή 137
5.2.6 Βαθμονόμηση Πιθανοτήτων του Μοντέλου Πρωταθλητή 138
5.3 Ερμηνευσιμότητα μέσω των Διαγραμμάτων SHAP 140
5.3.1 Αρχιτεκτονική Ενσωμάτωσης της μεθόδου SHAP στο Pipeline 140
5.3.2 Τεχνικά Ζητήματα Υλοποίησης 143
5.3.3 Στατικά Γραφήματα SHAP 145
5.4 Backend API 152
5.4.1 Αρχιτεκτονική του API 152
5.4.2 Τερματικά σημεία διαχείρισης 153
5.4.3 Τερματικά Σημεία Πρόβλεψης και Καταγραφής της 153
5.5 Έξυπνο Συμβόλαιο 153
5.5.1 Διεπαφή Backend με το Έξυπνο Συμβόλαιο 154
5.5.2 Επισκόπηση του Έξυπνου Συμβολαίου MedicalOracle 155
5.5.3 Έλεγχος Πρόσβασης βάσει Ρόλων 158
5.5.4 Λογική Συγκατάθεση κατά GDPR 158
5.5.5 Τμηματοποίηση μέσω Merkle Trees 160
5.5.6 Έκτακτη Πρόσβαση και Καταγραφή Ενεργειών 161
5.5.7 Ανάπτυξη Έξυπνου Συμβολαίου 161
5.6 Διεπαφή Χρήστη της Κλινικής Πλατφόρμας 168
5.6.1 Αρχιτεκτονική της Διεπαφής Χρήστη 168
5.6.2 Προσβασιμότητα και Αυθεντικοποίηση 169
5.6.3 Μεταφόρτωση CSV και Καταγραφή Ασθενή 172
5.6.4 Πίνακας Ανάλυσης Στατιστικών 174
5.6.5 Αποτελέσματα και Λίστα ασθενών 176
5.6.6 Καρτέλα Blockchain 179
5.6.7 Παραγωγή PDF ανά ασθενή 181
5.6.8 Πάνελ Διαχείρισης Χρηστών 183

Κεφάλαιο 6: Αποτελέσματα και Αξιολόγηση 184
6.1 Υπολογιστικό Περιβάλλον και Σύνολο Δεδομένων 185
6.2 Αξιολόγηση των Τεσσάρων Υποψήφιων Μοντέλων 187
6.3 Επίδραση του Optuna 188
6.4 Επίδοση του Μοντέλου Πρωταθλητή 192
6.4.1 Μετρικές αποτελεσμάτων του συνόλου ελέγχου 192
6.4.2 Επαλήθευση Μετρικών μέσω της μεθόδου Bootstrap CI 194
6.4.3 Διαγνωστικά Διαγράμματα Απόδοσης στο Σύνολο Ελέγχου 196
6.5 Αξιολόγηση και Ερμηνεία των τελικών Διαγραμμάτων SHAP 200
6.5.1 Διάγραμμα μέσης σημαντικότητας χαρακτηριστικών 200
6.5.2 Διάγραμμα διασποράς ανά Χαρακτηριστικό (Beeswarm) 205
6.5.3 Διάγραμμα Καταρράκτη ανά Ασθενή 208
6.6 Αξιολόγηση του Blockchain Επιπέδου 212
Κεφάλαιο 7: Συζήτηση 213
7.1 Συνολική Αποτίμηση των Αποτελεσμάτων 214
7.2 Ερμηνεία στο Πλαίσιο της Βιβλιογραφίας 214
7.3 Η αξία της Εξηγήσιμης Τεχνητής Νοημοσύνης στη Λήψη Κλινικών Αποφάσεων 215
7.4 Συνεισφορά Blockchain Επιπέδου 215
7.5 Σχεδιαστική Συμμόρφωση υπό τα Σύγχρονα Κανονιστικά Πλαίσια 216
7.6 Περιορισμοί της Εργασίας 217
7.7 Μελλοντικές Κατευθύνσεις και Τελικά Συμπεράσματα 217

Κατάλογος Εικόνων

Εικόνα 1 : Απεικόνιση της αρχιτεκτονικής των σταδιών που ακολουθεί η διαδικασία της μηχανικής μάθησης 21
Εικόνα 2 : Ένα συνελικτικό νευρωνικό δίκτυο που εκπαιδεύεται στην ερμηνεία εικόνων. Η έξοδος του χρησιμοποιείται από ένα επαναληπτικό νευρωνικό δίκτου που έχει εκπαιδευτεί στην δημιουργία λεζάντων κειμένου. Η ακολουθία στο κάτω μέρος παρουσιάζει την εστιάση του δικτύου, λέξη προς λέξη. 22
Εικόνα 3 : Παράδειγμα αρχιτεκτονικής του Blockchain με συνεχής κόμβους. 24
Εικόνα 4 : Δομή του Μπλοκ 25
Εικόνα 5 : Αναπαράσταση ενός Μοντέλου Συγκεντρωμένων Παραμέτρων (Lumped Parameter Model) του καρδιακού συστήματος, με σκοπό την προσομείωση της αιμοδυναμικής 30
Εικόνα 6 : Διαγνωστικός αλγόριθμος της καρδιακής ανεπάρκειας 33
Εικόνα 7 : Θεραπευτικός αλγόριθμος διαχείρισης των ασθενών που πάσχουν από Καρδιακή ανεπάρκεια με μειωμένο κλάσμα εξώθησης (HFrEF). 34
Εικόνα 8 : Αναλυτικό Διάγραμμα θεραπείας για ασθενείς με καρδιακή ανεπάρκεια ελαφρώς μειωμένου κλάσματος εξώθησης. 36
Εικόνα 9 : Ταξινόμηση της Καρδιακής Ανεπάρκειας σύμφωνα με το Κολλέγιο Καρδιολογίας της Αμερικής/ Αμερικανική Καρδιολογική Εταιρεία (AHA/ACC) και τη Καρδιολογική Εταιρεία της Νέας Υόρκης (NYHA). 39
Εικόνα 10 : Σχηματική αναπαράσταση της λειτουργίας ενός δικτύου Long Short-Term Memory (LSTM) με στόχο την ανάλυση κλινικών δεδομένων. 42
Εικόνα 11 : Σχηματική αναπαράσταση του μοντέλου SVM για την πρόβλεωη της θνησιμότητας σε ασθενείς που πάσχουν από Καρδιακή Ανεπάρκεια 44
Εικόνα 12 : Συνολική εξηγήσιμη ανάλυση επιβίωσης σε ασθενείς με καρδιακή ανεπάρκεια, με βάση τις τιμές των χαρακτηριστικών που επιλέχθηκαν με τη χρήση του SurvSHAP 49
Εικόνα 13 : Συγκριτικό διάγραμμα των μετρικών αξιολόγησης ανά ομάδα. (α) Κλίμακα Εμπιστοσύνης για XAI, (β) Κλίμακα Ικανοποίησης από την Εξήγηση, (γ) Κλίμακα Ευχρηστίας Συστήματος. 51
Εικόνα 14 : Παράδειγμα συστήματος Ευχρηστίας που συνδυάζει τα αποτελέσματα της Τεχνητής Νοημοσύνης, με Διαγράμματα Καταρράκτη SHAP και τη κλινική εξήγηση σε φυσική γλώσσα για τη πρόβλεψη ανάγκης και μετάγγισης αιμοσφαρίνων σε θωρακοχειρουργική επέμβαση. 52
Εικόνα 15 : Ταξινόμηση των ερευνητικών δημοσιεύσεων σε χρονική εξέλιξη (2016-2019) ανά τομέα εφαρμογής της τεχνολογίας Blockchain στην υγεία 54
Εικόνα 16 : Διάγραμμα ροής λειτουργίας της πλατφόρμας ADvoCATE 57
Εικόνα 17 : a) Ο αριθμός άρθρων που χρησιμοποιούν τεχνικές Τεχνητής Νοημοσύνης στα δημοσιευμένα άρθρα. b) Η κατανομή των τεχνικών Τεχνητής Νοημοσύνης που χρησιμοποιούνται στα άρθρα με θέμα «Ενσωμάτωση τεχνητής νοημοσύνης και αλυσίδας μπλοκ στον τομέα της υγειονομικής περίθαλψης» (2017–2023). AI: Τεχνητή Νοημοσύνη, ML:

Μηχανική μάθηση, DL: Βαθιά μάθηση, FL: Ομοσπονδιακή μάθηση, RL: Ενισχυτική μάθηση, XAI: Εξηγήσιμη τεχνητή νοημοσύνη 60
Εικόνα 18 : Αναπαράσταση των Συστημάτων Υποστήριξης Κλινικών Αποφάσεων με γνώμονα τη γνώση. 62
Εικόνα 19 : Απεικόνιση της αρχιτεκτονικής του Τυχαίου Δάσους.....71
Εικόνα 20 : Αρχιτεκτονική τριών επιπέδων του συστήματος υποστήριξης κλινικής απόφασης (CDSS) για τη πρόβλεψη της καρδιακής ανεπάρκειας. Παρουσιάζονται διακριτά τα τρία επίπεδα της αρχιτεκτονικής (Presentation, Business Logic, Data) και οι υπηρεσίες ενδιάμεσου λογισμικού. 88
Εικόνα 21 : Επισκόπηση ανάπτυξης της ιατρικής βάσης δεδομένων MIMIC. Τα δεδομένα συλλέγονται από το αποθετήριο δεδομένων του BIDMC, το σύστημα πληροφοριών της ΜΕΘ (Metavision) και εξωτερικές πηγές (external) και μετατρέπονται σε μια κοινή ενοποιημένη μορφή. Κατά την αποιδιοποιήση (deidentification) αφαιρούνται όλα τα προσωπικά δεδομένα που μπορούν να ταυτοποιήσουν έναν ασθενή. Η τελική οργάνωση σε πίνακες του MIMIC χωρίζεται σε δύο μεγάλα υποσύνολα, το HOSPITAL και το ICU, ενώ το υποσύνολο NOTE περιέχει σημειώσεις σχετικά με τους ασθενείς. 90
Εικόνα 22 : Διάγραμμα Συνιστωσών του CDSS. Αλληλεπίδραση μεταξύ των επιπέδων διεπαφής, του συστήματος πρόβλεψης και του αποκεντρωμένου δικτύου Blockchain 97
Εικόνα 23 : Διάγραμμα ροής δεδομένων αλληλεπίδρασης του χρήστη με το ρόλο του «γιατρού» και του υποσυστήματος τεχνητής νοημοσύνης του CDSS 99
Εικόνα 24 : Διάγραμμα ροής δεδομένων αλληλεπίδρασης του χρήστη με το ρόλο του «ερευνητή» και του CDSS. 101
Εικόνα 25 : Διάγραμμα ροής δεδομένων αλληλεπίδρασης του χρήστη με το ρόλο του «διαχειριστή» και του CDSS 102
Εικόνα 26 : Διάγραμμα ροής δεδομένων αλληλεπίδρασης του χρήστη με το ρόλο του «διαχειριστή» και του CDSS. 103
Εικόνα 27 : Διάγραμμα ροής δεδομένων αλληλεπίδρασης του χρήστη με το ρόλο του «διαχειριστή» και του CDSS. 104
Εικόνα 28 : Διάγραμμα ακολουθίας της ροής ταυτοποίησης χρήστη και ανάκτησης ρόλων πρόσβασης (RBAC) 106
Εικόνα 29 : Διάγραμμα ακολουθίας της ροής πρόβλεψης κινδύνου καρδιακής ανεπάρκειας και ερμηνευσιμότητας SHAP, από τη μεταφόρτωση του αρχείου .CSV έως την οπτική παρουσίαση των διαγραμμάτων 109
Εικόνα 30 : Διάγραμμα ακολουθίας ανάρτησης πρόβλεψης στο έξυπνο συμβόλαιο MedicalOracle, καλύπτοντας τις φάσεις σύνδεσης πορτοφολιού και δημοσίευσης προβλέψεων. 113
Εικόνα 31 : Αρχιτεκτονική της διαδικασίας εκπαίδευσης του μοντέλου «Πρωταθλητή». Η διαδικασία καλύπτει την εισαγωγή δεδομένων στο MIMIC-IV έως την παραγωγή του βαθμονομημένου μοντέλου πρωταθλητή και των αντίστοιχων αποτελεσμάτων. 127

Εικόνα 32 : Διάγραμμα παραγωγής στατικών γραφημάτων SHAP από τη συνάρτηση _make_shap_plots() 143
Εικόνα 33 : Παράδειγμα απεικόνισης Beeswarm γραφήματος SHAP με γνώμονα τους βιοδείκτες συνόλου βαθμονόμησης που εκπαιδεύτηκε το μοντέλο 146
Εικόνα 34 :Παράδειγμα απεικόνισης ραβδογράμματος σπουδαίοτητας SHAP με γνώμονα τους βιοδείκτες συνόλου βαθμονόμησης που εκπαιδεύτηκε το μοντέλο 149
Εικόνα 35 : Παράδειγμα απεικόνισης διαγράμματος καταρράκτη SHAP με γνώμονα τους βιοδείκτες συνόλου βαθμονόμησης που εκπαιδεύτηκε το μοντέλο 151
Εικόνα 36 : Στιγμιότυπο οθόνης του τοπικού περιβάλλοντος Ganache πριν την ανάπτυξη του συμβολαίου.....162
Εικόνα 37 : Στιγμιότυπο του αρχείου .env στο περιβάλλον ανάπτυξης κώδικα Visual Studio .163
Εικόνα 38 Έξοδος του τερματικού PowerShell κατά την εκτέλεση των αρχείων deploy_contract.py και setup_consent.py 164
Εικόνα 39 Έξοδος του τερματικού PowerShell κατά την εκτέλεση του sub_storage.py.....164
Εικόνα 40 : Κατάσταση του Ganache GUI μετά την ολοκλήρωση των 5 αναρτήσεων. 165
Εικόνα 41 : Καρτέλα Συναλλαγών του Ganache.....166
Εικόνα 42 : Έξοδος του τερματικού PowerShell κατά τον συγχρονισμό του backend με το frontend μέσω του REST API.....166
Εικόνα 43 : Καρτέλα Blockchain της προτεινόμενης κλινικής πλατφόρμας. Στην καρτέλα απεικονίζεται και η λίστα των «Individual Patient Actions» που αφορά τους 5 ασθενείς που καταχωρήθηκαν στην αλυσίδα και τις ενέργειες που μπορούν να πραγματοποιηθούν πάνω στα δεδομένα τους. 167
Εικόνα 44 : Κεντρική σελίδας του συστήματος υποστήριξης κλινικών αποφάσεων με το όνομα «Diagnosis» 169
Εικόνα 45 : Σελίδα σύνδεσης χρήστη στην κλινική πλατφόρμα υποστήριξης αποφάσεων με όνομα «Diagnosis».....170
Εικόνα 46 : Σελίδα εγγραφής χρήστη στην κλινική πλατφόρμα υποστήριξης αποφάσεων με όνομα «Diagnosis» και επιλογή ρόλου εξουσιοδότησης. 171
Εικόνα 47 : Σελίδα ανάκτησης κωδικού χρήστη στην κλινική πλατφόρμα υποστήριξης αποφάσεων με όνομα «Diagnosis».....171
Εικόνα 48 : Μήνυμα απόρριψης πρόσβασης στη κλινική πλατφόρμα 172
Εικόνα 49 : Καρτέλα μεταφόρτωσης και καταγραφής δεδομένων του ασθενή. 174
Εικόνα 50 : Καρτέλα Dashboard απεικόνισης των αποτελεσμάτων ενός πληθυσμού ασθενών με συγκεντρωτικά γραφήματα.....176
Εικόνα 51 : Λίστα ασθενών στην καρτέλα των αποτελεσμάτων (Results). Η λίστα δίνει την δυνατότητα της ταξινόμησης και του φιλτραρίσματων των ασθενών. 177
Εικόνα 52 : Εξατομικευμένο SHAP Διάγραμμα Καταρράκτη για ασθενή. Στο συγκεκριμένο απεικονίζεται η υψηλή πιθανότητα κινδύνου και το ιατρικό ιστορικό του. 179
Εικόνα 53 : Καρτέλα Blockchain της εφαρμογής. Μετά την σύνδεση με το πορτοφόλι Metamask ο κλινικός γιατρός έχει τη δυνατότητα να επιβλέπει τις εγγραφές νοσηλείας των ασθενών. 180

Εικόνα 54 : Αναδειόμενο παράθυρο ενημέρωσης της κατάστασης των ασθενών. Στους ασθενείς που εμφανίζουν υψηλό ποσοστό κινδύνου καρδιακής ανεπάρκειας αναδύεται δεύτερο παράθυρο ειδοποίησης.....181
Εικόνα 55 Παράδειγμα ασθενούς από την λίστα κινδύνου των ασθενών στην καρτέλα Results 182
Εικόνα 56 : Παράδειγμα παραγόμενης κλινικής αναφοράς PDF του ασθενή DEMO1023. Η αναφορά περιλαμβάνει τα δεδομένα του ασθενούς, τη πρόβλεψη του μοντέλου και την ερμηνεία της μέσω των κυρίαρχω παραγόντων SHAP.....183
Εικόνα 57 : Πάνελ Διαχειριστή. Ο Διαχειριστής (Admin) διαχειρίζεται την πρόσβαση και εξουσιοδότηση ρόλων από υπάρχοντες χρήστες 184
Εικόνα 58 : Καμπύλες ROC των τεσσάρων υποψήφιων μοντέλων επί του συνόλου βαθμονόμησης στη 1η δοκιμή. 188
Εικόνα 59 : Οι Κατανομές Bootstrap των πέντε μετρικών απόδοσης του μοντέλου «Πρωταθλητή» επί του συνόλου ελέγχου. 195
Εικόνα 60 : Καμπύλη ROC του μοντέλου «Πρωταθλητή» Τυχαίου Δάσους στο σύνολο ελέγχου. Το AUROC=0854 επιβεβαιώνει τη διακριτική ικανότητα του μοντέλου. 197
Εικόνα 61 : Η καμπύλη βαθμονόμησης πιθανοτήτων στο σύνολο ελέγχου υπό δεκαπλή κβαντοποίηση. Η σχεδόν τέλεια ευθγράμμιση με τη διαγώνιο επιβεβαιώνει οπτικά τη χαμηλή τιμή Brier=0.094 198
Εικόνα 62 : Ανάλυση Καμπυλών Απόφασης (DCA) του μοντέλου έναντι των στρατηγικών αναφοράς. Το μοντέλο παρουσιάζεται να διατηρεί υψηλό κέρδος δικτύου σε ολόκληρη τη κλινική περιοχή κατωφλίων (0.01-0.50) 199
Εικόνα 63 : Διάγραμμα Μέσης Σημαντικότητας χαρακτηριστικών (SHAP Importance Bar) του μοντέλου Λογιστικής Παλινδρόμησης επί του συνόλου βαθμονόμησης.....201
Εικόνα 64 : Διάγραμμα Μέσης Σημαντικότητας χαρακτηριστικών (SHAP Importance Bar) του μοντέλου «Πρωταθλητή» Τυχαίου Δάσους επί του συνόλου βαθμονόμησης.....202
Εικόνα 65 : Διάγραμμα Διασποράς ανά χαρακτηριστικό (Beeswarm) του μοντέλου Λογιστικής Παλινδρόμησης επί του συνόλου βαθμονόμησης.....206
Εικόνα 66 : Διάγραμμα Διασποράς ανά χαρακτηριστικό (Beeswarm) του μοντέλου «Πρωταθλητή» Τυχαίου Δάσους επί του συνόλου βαθμονόμησης. 207
Εικόνα 67 : Διάγραμμα Καταρράκτη του ασθενούς με υψηλό κίνδυνο ΚΑ από το μοντέλο Λογιστική Παλινδρόμηση. Η βασική τιμή είναι E[f(X)] = -0.849 και η τελική πρόβλεψη f(x) = +12.919 εκφρασμένες στο χώρο των λογιστικών πιθανοτήτων 210
Εικόνα 68 : Διάγραμμα Καταρράκτη του ασθενούς με υψηλό κίνδυνο ΚΑ του μοντέλου «πρωταθλητή» Τυχαίου Δάσους. Η βασική τιμή είναι E[f(X)] = 0.308 και η τελική πρόβλεψη είναι f(x) =0.951 εκφρασμένες στο χώρο των πιθανοτήτων. 211

Κατάλογος Στιγμιότυπων Κώδικα

Κώδικας 1 Διαμόρφωση Pipeline της Λογιστικής Παλινδρόμησης....127
Κώδικας 2 Διαμόρφωση των υπεραπαραμέτρων του RandomForestClassifier....128
Κώδικας 3 Διαμόρφωση των υπερπαραμέτρων του XGBClassifier....129
Κώδικας 4 Διαμόρφωση των υπερπαραμέτρων του LGBMClassifier....130
Κώδικας 5 Διαχωρισμός συνόλου εκπαίδευσης για πρόωρο τερματισμό....131
Κώδικας 6 Πρόωρος τερματισμός εκπαίδευσης XGBoost....131
Κώδικας 7 Πρόωρος τερματισμός εκπαίδευσης LightGBM....132
Κώδικας 8 Αρχικοποίηση TPESampler και MedianPruner του Optuna....133
Κώδικας 9 Χώρος αναζήτησης των υπερπαραμέτρων του XGBoost μέσω της βιβλιοθήκης Optuna....134
Κώδικας 10 Χώρος αναζήτησης των υπερπαραμέτρων LightGBM μέσω της βιβλιοθήκης Optuna....135
Κώδικας 11 Βρόχος σταυροεπαλήθευσης και εκπαίδευσης μοντέλου ανά fold....136
Κώδικας 12 Συνάρτηση υπολογισμού της μετρικής σκορ του Optuna....136
Κώδικας 13 Αξιολόγηση των υποψήφιων μοντέλων στο σύνολο βαθμονόμησης....137
Κώδικας 14 Σχεδίαση καμπύλων ROC των υποψήφιων μοντέλων....138
Κώδικας 15 Βαθμονόμηση των πιθανοτήτων μέσω της CalibratedClassifierCV....139
Κώδικας 16 Σειριοποίηση του αρχείου που περιλαμβάνει το συμπιεσμένο εκπαιδευμένο μοντέλο model_bundle.pkl....140
Κώδικας 17 Δημιουργία SHAP Explainer ανά τύπο μοντέλου....144
Κώδικας 18 Επιλογή θετικής κλάσης σε τιμές SHAP....145
Κώδικας 19 Παραγωγή γραφήματος Beeswarm SHAP....147
Κώδικας 20 Παραγωγή ραβδογράμματος σημαντικότητας SHAP....148
Κώδικας 21 Παραγωγή γραφήματος καταρράκτη SHAP υψηλού κινδύνου....151
Κώδικας 22 Κατασκευή payload μεταφοράς πρόβλεψης στο Blockchain....154
Κώδικας 23 Κατασκευή payload μεταφοράς πρόβλεψης στο Blockchain....156
Κώδικας 24 Ορισμοί mappings αποθήκευσης (patientRecords, patientConsents, emergencyLogs)....156
Κώδικας 25 Ορισμός ρόλων ελέγχου πρόσβασης (RBAC)....158
Κώδικας 26 Συνάρτηση καταγραφής συγκατάθεσης ασθενούς setPatientConsent....159
Κώδικας 27 Συνάρτηση τροποποίησης ελέγχου συγκατάθεσης setPatientConsent....160
Κώδικας 28 Συνάρτηση πρόσβασης έκτακτης ανάγκης emergencyAccess....161
Κώδικας 29 Συνάρτηση επικύρωσης του ασθενή verifyPatientInBatch....161

ΕΛΛΗΝΙΚΑ ΑΚΡΩΝΥΜΙΑ

Συντομογραφία	Επεξήγηση
ΚΑ	Καρδιακή Ανεπάρκεια
ΤΝ	Τεχνητή Νοημοσύνη
ΜΜ	Μηχανική Μάθηση
ΜΕΘ	Μονάδα Εντατικής Θεραπείας
ΔΚ	Δεξιά Κοιλία
ΧΑΠ	Χρόνια Αποφρακτική Πνευμονοπάθεια
ΧΝΝ	Χρόνια Νεφρική Νόσος
ΗΠΑ	Ηνωμένες Πολιτείες Αμερικής

ΙΑΤΡΙΚΟΙ / ΚΛΙΝΙΚΟΙ ΟΡΟΙ

Συντομογραφία	Επεξήγηση	Ελληνική Ορολογία
HF	Heart Failure	Καρδιακή Ανεπάρκεια
HFrEF	Heart Failure with reduced Ejection Fraction	ΚΑ μειωμένου κλάσματος εξώθησης
HFmrEF	Heart Failure with mildly reduced Ejection Fraction	ΚΑ ελαφρώς μειωμένου κλάσματος εξώθησης
HFpEF	Heart Failure with preserved Ejection Fraction	ΚΑ διατηρημένου κλάσματος εξώθησης
LVEF	Left Ventricular Ejection Fraction	Κλάσμα Εξώθησης Αριστερής Κοιλίας
EF	Ejection Fraction	Κλάσμα Εξώθησης
BNP	B-type Natriuretic Peptide	Νατριουρητικό Πεπτίδιο τύπου Β
NT-proBNP	N-terminal pro B-type Natriuretic Peptide	
BUN	Blood Urea Nitrogen	Άζωτο Ουρίας Αίματος
RDW	Red cell Distribution Width	Εύρος Κατανομής Ερυθρών

eGFR	estimated Glomerular Filtration Rate	Εκτιμώμενος Ρυθμός Σπειραματικής Διήθησης
ESC	European Society of Cardiology	Ευρωπαϊκή Καρδιολογική Εταιρεία
AHA	American Heart Association	Αμερικανική Καρδιολογική Εταιρεία
ACC	American College of Cardiology ()	Αμερικάνικο Κολέγιο Καρδιολογίας
NYHA	New York Heart Association	Καρδιολογική Εταιρεία της Νέας Υόρκης
MAGGIC	Meta-Analysis Global Group in Chronic Heart Failure	Παγκόσμια Ομάδα Μετα-Ανάλυσης για τη Χρόνια Καρδιακή Ανεπάρκεια
ICD	International Classification of Diseases	Διεθνής Ταξινόμηση Νόσων

ΑΚΡΩΝΥΜΙΑ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ / ΤΕΧΝΗΤΗΣ ΝΟΗΜΟΣΥΝΗΣ

Συντομογραφία	**Επεξήγηση**	Ελληνική Ορολογία
AI	Artificial Intelligence	Τεχνητή Νοημοσύνη
ML	Machine Learning	Μηχανική Μάθηση
DL	Deep Learning	Βαθιά Μάθηση
XAI	Explainable Artificial Intelligence	Εξηγήσιμη Τεχνητή Νοημοσύνη
FL	Federated Learning	Ομοσπονδιακή Μάθηση
RL	Reinforcement Learning	Ενισχυτική Μάθηση
SHAP	SHapley Additive exPlanations	

LR	Logistic Regression	Λογιστική Παλινδρόμηση
RF	Random Forest	Τυχαίο Δάσος
XGBoost	eXtreme Gradient Boosting	Ακραία Ενίσχυση Κλίσης
LightGBM	Light Gradient Boosting Machine	Μηχανή Ενίσχυσης Ελαφριάς Κλίσης
SVM	Support Vector Machine	Μηχανή Διανυσμάτων Υποστήριξης
LSTM	Long Short-Term Memory	
SMOTE	Synthetic Minority Over-sampling Technique	

ΜΕΤΡΙΚΕΣ ΑΞΙΟΛΟΓΗΣΗΣ

Συντομογραφία	**Επεξήγηση**	**Ελληνική Ορολογία**
AUROC	Area Under the Receiver Operating Characteristic curve	Περιοχή κάτω από την καμπύλη χαρακτηριστικών λειτουργίας του δέκτη
AUPRC	Area Under the Precision-Recall Curve	Περιοχή κάτω από την καμπύλη ακρίβειας-ανάκλησης
ROC	Receiver Operating Characteristic	Χαρακτηριστική καμπύλη λειτουργίας δέκτη
AUC	Area Under the Curve	Περιοχή κάτω από την καμπύλη
DCA	Decision Curve Analysis	Ανάλυση Καμπυλών Απόφασης
CI	Confidence Interval	Διάστημα Εμπιστοσύνης

ΑΚΡΩΝΥΜΙΑ BLOCKCHAIN

Συντομογραφία	**Επεξήγηση**	**Ελληνική Ορολογία**

EVM	Ethereum Virtual Machine	Εικονική Μηχανή Ethereum
ABI	Application Binary Interface	Δυαδική Διεπαφή Εφαρμογής
RBAC	Role-Based Access Control	Έλεγχος Πρόσβασης Βάσει Ρόλων
RPC	Remote Procedure Call	Απομακρυσμένη Κλήση Διαδικασίας
SHA	Secure Hash Algorithm	Αλγόριθμος Ασφαλούς Κατακερματισμού
HMAC	Hash-based Message Authentication Code	Κώδικας Αυθεντικοποίησης Μηνύματος Βασισμένος σε Κατακερματισμό.

ΑΚΡΩΝΥΜΙΑ ΛΟΓΙΣΜΙΚΟΥ

Συντομογραφία	**Επεξήγηση**	**Ελληνική Ορολογία**
CDSS	Clinical Decision Support System	Σύστημα Υποστήριξης Κλινικών Αποφάσεων
EHR	Electronic Health Records	Ηλεκτρονικά Αρχεία Υγείας
PHR	Personal Health Records	Προσωπικά Αρχεία Υγείας
API	Application Programming Interface	Διεπαφή Προγραμματισμού Εφαρμογών
REST	Representational State Transfer	Μεταφορά Αναπαράστασης Κατάστασης
SPA	Single Page Application	Μονοσέλιδη Εφαρμογή Σελίδας
GUI	Graphical User Interface	Γραφική Διεπαφή Χρήστη
JSON	JavaScript Object Notation	Σημειογραφία Αντικειμένων JavaScript
CSV	Comma-Separated Values	Τιμές διαχωρισμένες με κόμμα
PDF	Portable Document Format	Μορφή Φορητού Εγγράφου
HTTP	HyperText Transfer Protocol	Πρωτόκολλο Μεταφοράς Υπερκειμένου

Κεφάλαιο 1: Εισαγωγή

Στο παρόν εισαγωγικό κεφάλαιο παρουσιάζονται το πλαίσιο και κίνητρο υλοποίησης της εργασίας, ο ορισμός του ερευνητικού προβλήματος, η διατύπωση των στόχων, η ανάδειξη της συνεισφοράς και η σύντομη περιγραφή της δομής της διπλωματικής εργασίας.

1.1 Πλαίσιο και Κίνητρο

1.1.1Ψηφιακός Μετασχηματισμός της Υγείας

Η σύγχρονη υγειονομική περίθαλψη ακολουθεί την ραγδαία εξέλιξη της τεχνολογίας και βρίσκεται σε φάση ψηφιακού μετασχηματισμού. Η μετάβαση αυτή αποσκοπεί στην ταχύτερη εξυπηρέτηση και διαχείριση των υπηρεσιών υγείας, καθώς και την ικανοποίηση των ασθενών. Στοιχεία αυτής της μετάβασης αποτελούν η ευρεία υιοθέτηση των Ηλεκτρονικών Ιατρικών Αρχείων (Electronic Health Records, EHR), η πρόοδος της τηλεϊατρικής και τα συστήματα παρακολούθησης ασθενών σε πραγματικό χρόνο, που οδηγούν στη παραγωγή ενός εκθετικά αυξανόμενου όγκου κλινικών δεδομένων. Στις Ηνωμένες Πολιτείες, ο νόμος HITECH (Health Information Technology for Economic and Clinical Heath) που ψηφίστηκε το 2009, αποτέλεσε προσπάθεια για την αύξηση αποδοχής των Ηλεκτρονικών Ιατρικών Αρχείων (EHRs) ως κύρια πηγή αποθήκευσης δεδομένων για το ιατρικό προσωπικό του ασθενή. Πράγματι, ο νόμος οδήγησε στην χρήση EHRs από περίπου 9% σε πάνω από 96% για τα νοσοκομεία σε λιγότερο από μία δεκαετία [1]. Ταυτόχρονα, ο παγκόσμιος όγκος ιατρικών δεδομένων διπλασιαζόταν κάθε 73 μέρες ήδη από τα μέσα της δεκαετίας του 2010, τροφοδοτούμενος κυρίως στη σύγκλιση τεχνολογιών όπως η ψηφιακή απεικόνιση, η γονιδιωματική και τα φορητά αισθητήρια [2].

Η ανάγκη αξιοποίησης των ιατρικών δεδομένων αναγνωρίζεται σε διεθνές επίπεδο. Ο Παγκόσμιος Οργανισμός Υγείας (ΠΟΥ), στην Παγκόσμια Στρατηγική Ψηφιακής Υγείας 2020-2025, καταγράφει ότι 129 χώρες διαθέτουν εθνικές στρατηγικές ψηφιακής υγείας [3]. Παράλληλα, η Ευρωπαϊκή Ένωση θέσπισε τον Κανονισμό για τον Ευρωπαϊκό Χώρο Δεδομένων Υγείας (EHDS, Regulation 2025/327), το οποίο αποτέλεσε πλαίσιο για την δευτερογενή χρήση δεδομένων υγείας σε έρευνα και καινοτομία [4]. Όλα τα παραπάνω καθιστούν επιτακτική την αξιοποίηση τεχνολογιών ικανών να επεξεργαστούν μεγάλα σύνολα κλινικών δεδομένων, με κυριότερη εξ αυτών την Τεχνητή Νοημοσύνη (ΤΝ).

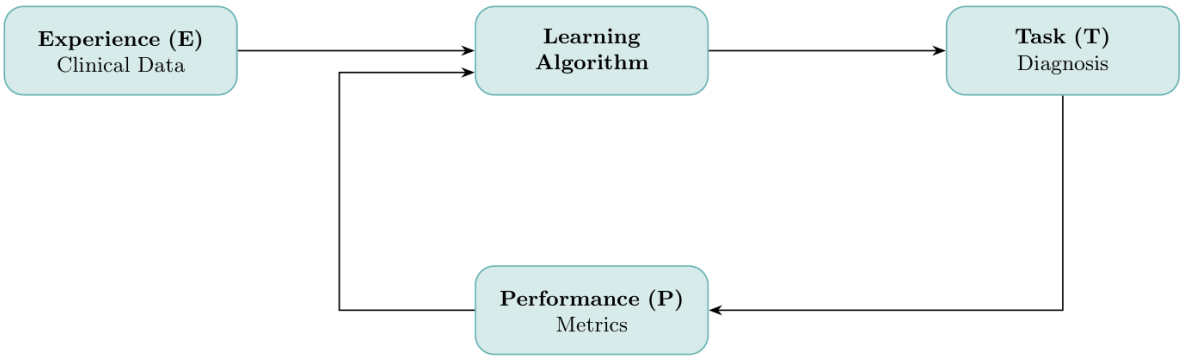
1.1.2 Τεχνητή Νοημοσύνη: Ορισμός και Εφαρμογές

Η Τεχνητή Νοημοσύνη (Artificial Intelligence, συντομογραφικά A.I.) αποτελεί κλάδο της επιστήμης υπολογιστών που αποσκοπεί στην σχεδίαση και υλοποίηση ευφυών συστημάτων να εκτελούν λειτουργίες βασισμένες σε νοητικές ικανότητες. Η πρώτη εμφάνιση της ΤΝ ως

ερευνητικό θέμα ανάγεται στο συνέδριο του Darthmouth το 1956 [5]. Εκεί, ο John McCarthy και η ερευνητική ομάδα του εισήγαγαν την θεμελιώδη παραδοχή πως κάθε πτυχή της ανθρώπινης μάθησης μπορεί να περιγράφει με τόση ακρίβεια, ώστε να είναι εφικτή η υπολογιστική προσομοίωση του από μια μηχανή. Στην σύγχρονη επιστημονική βιβλιογραφία, οι Russel και Norvig διακρίνουν τα συστήματα ΤΝ σε εκείνα που σκέφτονται ή δρουν είτε με γνώμονα την ανθρώπινη συμπεριφορά, είτε με γνώμονά την ορθολογικότητα για την επίτευξη του βέλτιστου δυνατού αποτελέσματος [6].

Η ΤΝ αποτελείται από ένα πλήθος υποπεδίων, με τα κυριότερα να είναι τα εξής:

1) **Μηχανική Μάθηση (Machine Learning):** Η μηχανική μάθηση (ML) αφορά την σχεδίαση και μελέτη αλγορίθμων και στατιστικών μοντέλων που επιτρέπουν σε υπολογιστές να εκτελούν διεργασίες βασιζόμενοι σε πρότυπα (patterns) και συμπερασμούς (inferences), αντί για ρητές οδηγίες προγραμματισμού [7]. Σύμφωνα με τον κλασικό ορισμό, η διαδικασία περιγράφεται ως η βελτιστοποίηση μιας συνάρτησης απόδοσης **P** για μια συγκεκριμένη κατηγορία εργασιών **T**, η οποία επιτυγχάνεται βάσει της εμπειρίας **E** που αντλείται από τα δεδομένα εισόδου [7].



Εικόνα 1 : Απεικόνιση της αρχιτεκτονικής των σταδίων που ακολουθεί η διαδικασία της μηχανικής μάθησης [7].

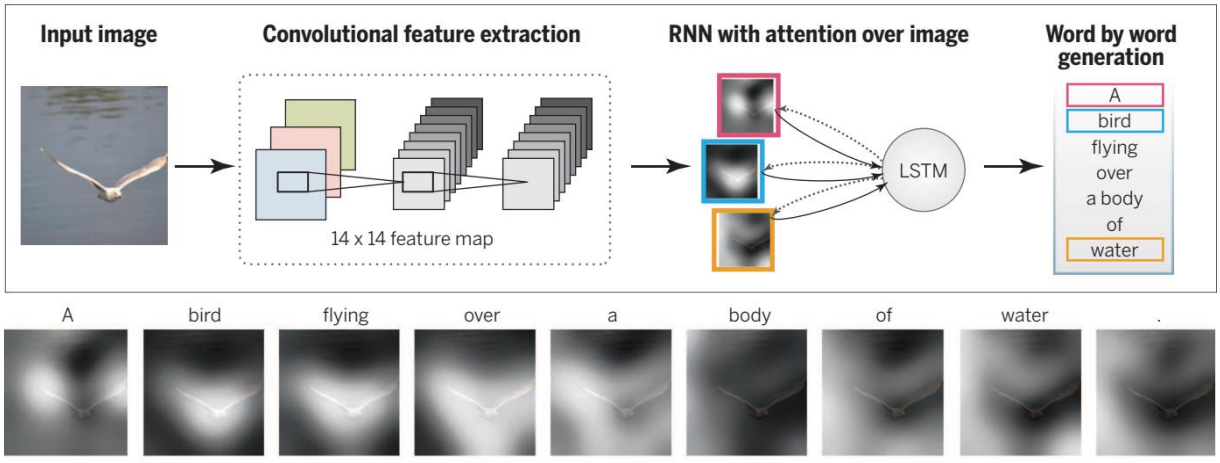
Στο τομέα της βιοιατρικής, η Μηχανική Μάθηση αποτελεί τον κεντρικό πυλώνα από την απλή ψηφιοποίηση στην ευφυή ανάλυση κλινικών παραμέτρων, συμβάλλοντας στην πρόβλεψη σύνθετων παθολογικών καταστάσεων.

2) **Βαθιά Μάθηση (Deep Learning):** Η Βαθιά Μάθηση (DL) αποτελεί υποκατηγορία της Μηχανικής Μάθησης που βασίζεται στην αρχιτεκτονική των Τεχνητών Νευρωνικών Δικτύων (Artificial Neural Networks - ANNs) με πολλαπλά στρώματα (layers) επεξεργασίας και ανάλυσης [8]. Η ειδοποιός διαφορά μεταξύ της DL και των κλασσικών μεθόδων ML έγκειται στην ικανότητα αυτόματης εξαγωγής χαρακτηριστικών (feature extraction). Ενώ για παράδειγμα στην Μηχανική Μάθηση απαιτείται συχνά η χειροκίνητη σχεδίαση των

χαρακτηριστικών από ειδικούς (feature engineering), τα νευρωνικά δίκτυα εκπαιδεύονται στην ιεραρχική αναπαράσταση των δεδομένων, κατά την οποία τα πρώτα επίπεδα εντοπίζουν απλά πρότυπα (π.χ. ακμές σε μια ιατρική εικόνα) και τα βαθύτερα επίπεδα συνθέτουν αυτά τα πρότυπα για την αναγνώριση πολυπλοκότερων εννοιών ή παθολογιών [8].

Το μαθηματικό υπόβαθρο τους στηρίζεται στη χρήση μη γραμμικών μετασχηματισμών σε υψηλής διάστασης δεδομένα, κάτι που θέτει την DL ιδιαίτερα αποτελεσματική στην επεξεργασία μη δομημένων δεδομένων, όπως οι ιατρικές εικόνες, τα σήματα ηλεκτροκαρδιογραφήματος. Μάλιστα, στον ιατρικό τομέα η χρήση της DL επιδεικνύει όλο και αποδοτικότερη πρόβλεψη κλινικών εκβάσεων και διάγνωση παθήσεων, εξαιτίας της διαχείρισης που παρέχει στην εγγενή πολυπλοκότητα και θόρυβο των βιολογικών δεδομένων.

Παρά την υπεροχή της, η αποτελεσματικότητα των μοντέλων DL κρίνεται από το μέγεθος των δεδομένων και την πρόσβαση σε υπολογιστικούς πόρους [8][9]. Η απουσία επαρκούς όγκου δεδομένων καθιστά τα μοντέλα επιρρεπή σε φαινόμενα όπως η υπερπροσαρμογή (overfitting), ελαττώνοντας την αξιοπιστία και την ακρίβεια τους σε εφαρμογές που έχουμε περιορισμένα δεδομένα. Τα μοντέλα βαθιάς μάθησης παρουσιάζονται ως συστήματα υψηλής υπολογιστικής πολυπλοκότητας (computational complexity), για τον λόγο αυτόν η εκπαίδευση τους απαιτεί εξειδικευμένο υλικό για την επεξεργασία των εκατομμύριων πολλές φορές παραμέτρων που τα απαρτίζουν.



Εικόνα 2 : Ένα συνελικτικό νευρωνικό δίκτυο που εκπαιδεύεται στην ερμηνεία εικόνων. Η έξοδος του χρησιμοποιείται από ένα επαναληπτικό νευρωνικό δίκτυο που έχει εκπαιδευτεί στην δημιουργία λεζάντων κειμένου. Η ακολουθία στο κάτω μέρος παρουσιάζει την εστίαση του δικτύου, λέξη προς λέξη [8].

3) **Επεξεργασία Φυσικής Γλώσσας (Natural Language Processing):** Η Επεξεργασία Φυσικής Γλώσσας αποτελεί διεπιστημονικό κλάδο της Τεχνητής Νοημοσύνης και της Υπολογιστικής Γλωσσολογίας που έχει ως κύριο στόχο την μοντελοποίηση ανάμεσα στην ανθρώπινη γλώσσα και τα υπολογιστικά συστήματα [10]. Η βασική λειτουργία της NLP έγκειται στο

μετασχηματισμό της αδόμητης πληροφορίας (unstructured data) σε δομημένη. Με αυτόν τον τρόπο, δίνεται η δυνατότητα στους υπολογιστές να εκτελούν πολύπλοκες εργασίες όπως η συντακτική ανάλυση, η σημασιολογική κατανόηση και η αυτόματη παραγωγή κειμένου [10]. Στην σύγχρονη έρευνα, ο ρόλος της έχει αναβαθμιστεί από τα παραδοσιακά συστήματα των αυστηρών γραμματικών κανόνων, στα στατιστικά μοντέλα πιθανοτήτων που σε συνδυασμό με την Βαθιά Μάθηση, επιτρέπουν την κατανόηση περιεχομένου (context) και ερμηνεύουν την απόχρωση του λόγου [10]. Στην σημερινή εποχή, η NLP εμφανίζει προστατικό ρόλο στον τομέα της υγείας, συγκεκριμένα στους Ηλεκτρονικούς Φακέλους Υγείας, καθώς το μεγαλύτερο μέρος της ιατρικής πληροφορίας παραμένει «εγκλωβισμένο» σε μορφή ελευθέρου κειμένου (π.χ. ιατρικές γνωματεύσεις, ιστορικά ασθενών). Οι ικανότητες των αλγορίθμων NLP να εξάγουν αυτοματοποιημένα κλινικά ευρήματα, σε υψηλή ακρίβεια και να τα συσχετίζουν με διαγνωστικά πρότυπα, καθιστά την ύπαρξη τους κρίσιμη. Εκτός από την επεξεργασία και ερμηνεία ευαίσθητων δεδομένων, αναγκαία παρουσιάζεται η εφαρμογή πρωτόκολλών που συνάδουν ακεραιότητα και ασφάλεια, όπως η Αλυσίδα Συστοιχιών (Blockchain).

Οι εφαρμογές την Τεχνητής Νοημοσύνης εκτείνονται σε πλήθος τομέων: τα χρηματοοικονομικά (ανίχνευση απάτης, αλγοριθμικές συναλλαγές), στις μεταφορές (αυτόνομα οχήματα), στην ενέργεια (ευφυή δίκτυα) και στη βιομηχανία (Industry 4.0). Σύμφωνα με την ετήσια Έκθεση ΤΝ 2025 του Stanford HAI, το 78% των οργανισμών χρησιμοποίησε ΤΝ το 2025, ενώ οι ιδιωτικές επενδύσεις στις ΗΠΑ ξεπέρασαν τα 110 δισεκατομμύρια δολάρια [11].

1.1.3 Τεχνητή Νοημοσύνη στην Ιατρική

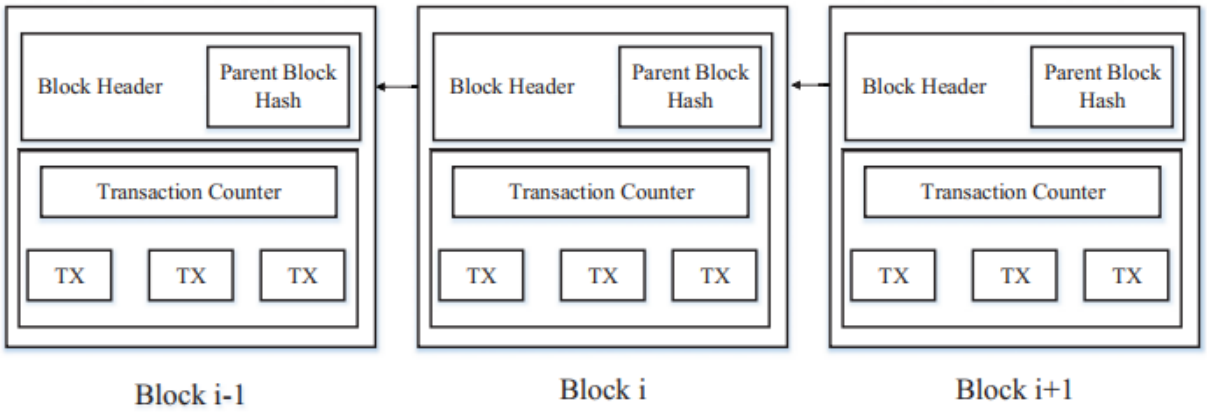
Η εφαρμογή της Τεχνητής Νοημοσύνης (ΤΝ) στον τομέα της υγειονομικής περίθαλψης έχει σημειώσει σημαντική πρόοδο. Ο Topol στην έρευνα του παρουσίασε ότι η ΤΝ έχει πραγματοποιήσει επιδόσεις ισοδύναμες ή ανώτερες από αυτές των ειδικών σε κλάδους όπως η ακτινολογία, η δερματολογία και η οφθαλμολογία, τονίζοντας πως λειτουργεί ενισχυτικά στη λήψη κλινικών αποφάσεων [12]. Παρά την τεκμηριωμένη ωριμότητα των αλγορίθμων, Οι Rajpukar και η ερευνητική ομάδα του επιβεβαίωσαν πως υπάρχουν προκλήσεις στη γενικευσιμότητα (generability) των μοντέλων σε διαφορετικούς πληθυσμούς και την ερμηνευσιμότητα (interpretability) των αποφάσεων τους [13].

Στο πεδίο της καρδιολογίας, η ΤΝ αναδεικνύεται σε καταλυτικό παράγοντα για την ανάλυση πολυδιάστατων δεδομένων, με εφαρμογές που εκτείνονται από την αυτοματοποιημένη ερμηνεία ηλεκτροκαρδιογραφημάτων και υπερηχοκαρδιογραφημάτων έως την αξιοποίηση δεδομένων από γονιδιωματικές αναλύσεις και φορητούς αισθητήρες (wearables) [14]. Η αναγκαιότητα αυτών των προηγμένων εργαλείων καθίσταται επιτακτική υπό το πρίσμα της παγκόσμιας επιδημιολογικής κρίσης της καρδιακής ανεπάρκειας. Με τον παγκόσμιο επιπολασμό να υπερβαίνει τα 64 εκατομμύρια ασθενείς και την πενταετή θνησιμότητα να αγγίζει το 50%, η HF συνιστά μία από

τις κρισιμότερες προκλήσεις για τα συστήματα δημόσιας υγείας παγκοσμίως [14][15]. Η ανησυχητική δυναμική της νόσου υπογραμμίζεται από το γεγονός ότι ο απόλυτος αριθμός των περιπτώσεων σημείωσε αύξηση της τάξης του 91,9% κατά την περίοδο 1990-2017. Τα δεδομένα αυτά τεκμηριώνουν την ανάγκη για τη μετάβαση από τα παραδοσιακά στατιστικά μοντέλα σε δυναμικά συστήματα πρόβλεψης κινδύνου βασισμένα στη Μηχανική Μάθηση, τα οποία δύνανται να εντοπίσουν σύνθετα παθοφυσιολογικά πρότυπα και να επιτρέψουν την έγκαιρη κλινική παρέμβαση [16].

1.1.4 Τεχνολογία Blockchain και ακεραιότητα ιατρικών δεδομένων

Η τεχνολογία Blockchain είναι είδος των κατανεμημένων καθολικών συστημάτων (Distributed Ledger Technology -- DLT), η οποία επιτρέπει την ασφαλή καταγραφή συναλλαγών σε ένα δίκτυο ομότιμων κόμβων (peer-to-peer), αφαιρώντας την ανάγκη για την κεντρικού ελέγχου [17]. Για να την κατανόηση της λειτουργίας του δικτύου εξετάζουμε αναλυτικά την εσωτερική αρχιτεκτονική του στην παρακάτω **Εικόνα 3**.



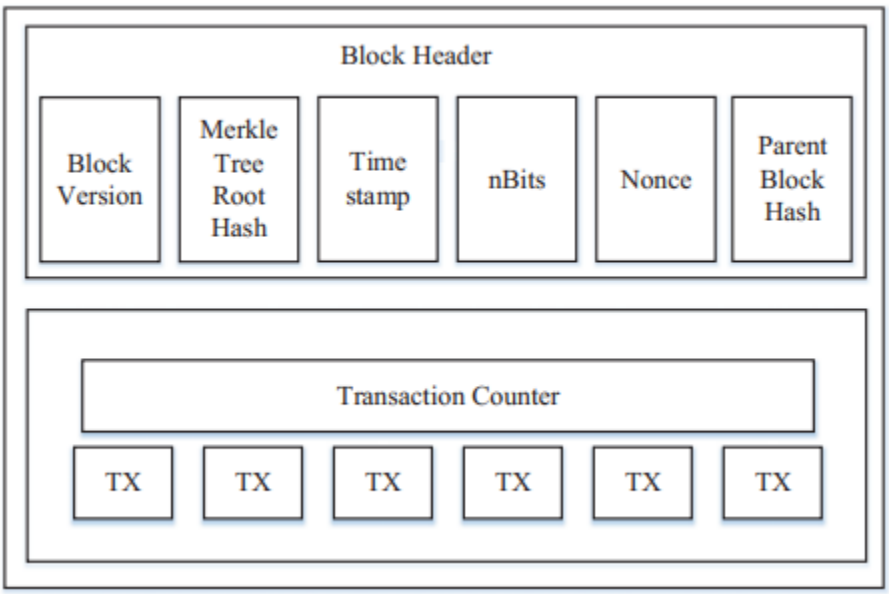
Εικόνα 3 : Παράδειγμα αρχιτεκτονικής του Blockchain με συνεχής κόμβους [17].

Όπως απεικονίζεται στην **Εικόνα 3**, το Blockchain αποτελείται μια συνεχή σειρά μπλοκ (Block i-1, i+1), όπου το καθένα περιλαμβάνει μια πληθώρα από πληροφορίες. Κάθε μπλοκ χωρίζεται σε δύο βασικά μέρη:

1. **Κεφαλίδα Μπλοκ:** Περιλαμβάνει κρίσιμα μεταδεδομένα, με κυριότερο τον Γονικός Κρυπτογραφικός Κατακερματισμός (Parent Block Hash). Το συγκεκριμένο έχει τον ρόλο του ψηφιακού συνδέσμου που συνδέει το τρέχον μπλοκ με το προηγούμενό του, δημιουργώντας την αλυσίδα. Η συνάρτηση κατακερματισμού που χρησιμοποιείται είναι η SHA-256. Πέραν του Parent Block Hash παρουσιάζονται και τα εξής πεδία:
    a. Έκδοση του Μπλοκ (Block Version): Ο αριθμός έκδοσης πρωτοκόλλου

b. Κατερματισμός Merkle Tree (Merkle Tree Root Hash): Ένα ενιαίο hash που αντιπροσωπεύει όλες τις συναλλαγές (TX) του μπλοκ.
c. Χρονική Σφραγίδα (Timestamp): Η χρονική σφραγίδα δημιουργίας του μπλοκ
d. nBits: το απαιτούμενο όριο για την επικύρωση hash του μπλοκ.
e. Nonce: Τυχαίος αριθμός που αλλάζει μέχρι το hash να πληροί το nBits

2. **Σώμα Μπλοκ:** Περιλαμβάνει έναν μετρητή συναλλαγών και τις συναλλαγές (TX), στην παρούσα εργασία αντιπροσωπεύουν τις ιατρικές εγγραφές ή τα δεδομένα ασθενών.



Εικόνα 4 : Δομή του Μπλοκ [17]

Η ασφάλεια του συστήματος στηρίζεται από την μέθοδο κατακερματισμού που εφαρμόζεται, δηλαδή αν τουλάχιστον ένα ψηφίο πληροφορίας αλλάξει στο Μπλοκ i-1, το hash θα αλλάξει ολοκληρωτικά. Εξαιτίας της σύνδεσης που παρουσιάζει το μπλοκ i-1 με το i, η σύνδεση θα σπάσει αμέσως, δείχνοντας την απάτη άμεσα ανιχνεύσιμη από όλους τους κόμβους του δικτύου [17].

Η μαθηματική υπόσταση του κατακερματισμού εμφανίζεται και στην εξίσωση (1):

$$H_n = SHA256(SHA256(Block_Header_n)) \qquad (1)$$

Όπου,

$$Block_Header_n = \{Version||H_{n-1}||MerkleRoot||Timestamp||Bits||Nonce\}$$

Η απουσία κεντρικού κόμβου στο Blockchain δίκτυο καθιστά την καθολικότητα στους κατανεμημένους κόμβους αδύνατη. Για τον παραπάνω λόγο, απαιτούνται ορισμένα πρωτόκολλα,

για να διασφαλιστεί η συνέπεια μεταξύ των κόμβων. Αυτά τα πρωτόκολλα ονομάζονται **Μηχανισμοί Συναίνεσης** [18]. Οι ίδιοι είναι αυστηρά μαθηματικά πρωτόκολλα που εξασφαλίζουν πως οι κόμβοι συμφωνούν με την κατάσταση του καθολικού δικτύου. Στην ερευνητική βιβλιογραφία εμφανίζονται διάφοροι μηχανισμοί με τους κυριότερους να είναι οι παρακάτω:

- **Απόδειξη Εργασίας (Proof of Work ή PoW):** Είναι ο πρώτος και πιο διαδεδομένος μηχανισμός (χρησιμοποιείται στο Bitcoin). Οι κόμβοι (miners) ανταγωνίζονται για την επίλυση ενός εξαιρετικά δύσκολου μαθηματικού γρίφου. Ο πρώτος που θα βρει τη λύση αποκτά το δικαίωμα να προσθέσει το επόμενο μπλοκ στην αλυσίδα. Αν και προσφέρει το υψηλότερο επίπεδο ασφάλειας, η διαδικασία αυτή απαιτεί τεράστια υπολογιστική ισχύ και κατανάλωση ενέργειας [18].
- **Απόδειξη Επικύρωσης (Proof of Validation ή PoV):** Αποτελεί την πιο ενεργειακά αποδοτική εναλλακτική. Αντί για υπολογιστική ισχύ, οι επικυρωτές (validators) επιλέγονται με βάση το ποσοστό των ψηφιακών νομισμάτων που «δεσμεύουν» στο δίκτυο. Όσο μεγαλύτερο είναι το μερίδιο συμμετοχής (stake), τόσο αυξάνονται οι πιθανότητες επιλογής για την επικύρωση του επόμενου μπλοκ, μειώνοντας δραστικά το περιβαλλοντικό αποτύπωμα του συστήματος [18].
- **Απόδειξη Αυθεντικότητας (Proof of Authenticity ή PoA):**
  Εδώ η συναίνεση βασίζεται στην ταυτότητα και τη φήμη των επικυρωτών. Μόνο εγκεκριμένοι και αξιόπιστοι κόμβοι (π.χ. πιστοποιημένα ιατρικά ιδρύματα) έχουν το δικαίωμα να επικυρώνουν συναλλαγές. Είναι ιδανικός μηχανισμός για την υγεία, καθώς συνδυάζει την ταχύτητα με την αυστηρή διακυβέρνηση των δεδομένων [18].

Γιατί η επιλογή του μηχανισμού είναι κρίσιμη για την εργασία;

Στην καρδιολογία και τη διαχείριση κρίσιμων ιατρικών δεδομένων, η επιλογή του μηχανισμού συναίνεσης αποτελεί μια ισορροπία μεταξύ τριών παραγόντων: **ασφάλειας**, **ταχύτητας** και **κλιμακωσιμότητας**. Για παράδειγμα, ενώ το PoW είναι αδιάβλητο, η καθυστέρηση στην επικύρωση μπορεί να είναι απαγορευτική για ένα σύστημα υποστήριξης κλινικών αποφάσεων που απαιτεί δεδομένα σε πραγματικό χρόνο [19]. Αντιθέτως, μηχανισμοί όπως το PBFT ή το PoA προσφέρουν την απαραίτητη ταχύτητα και ιδιωτικότητα που απαιτούν οι Ηλεκτρονικοί Φάκελοι Υγείας, διασφαλίζοντας ότι η πληροφορία παραμένει αμετάβλητη και άμεσα διαθέσιμη για τους αλγορίθμους Τεχνητής Νοημοσύνης [19].

1.2 Ορισμός του Προβλήματος

Η καρδιακή ανεπάρκεια παραμένει μία από τις κύριες αιτίες θανάτου παγκοσμίως, παρά τις σημαντικές εξελίξεις στη θεραπεία της — μια πραγματικότητα που αναδεικνύει την ανάγκη για πιο αποτελεσματικές προσεγγίσεις πρόβλεψης κινδύνου. Οι παραδοσιακοί κλινικοί δείκτες, όπως ο MAGGIC, στηρίζονται σε απλές στατιστικές μεθόδους. Η αρχική ανάπτυξη του MAGGIC, από δεδομένα 39.372 ασθενών, παρουσίασε δείκτη C-statistic (μέτρο διακριτικής ικανότητας ενός μοντέλου) περίπου 0,73, ενώ εξωτερικές επικυρώσεις σε 51.043

ασθενείς επιβεβαίωσαν C-index 0,741. Αντιθέτως, μετα-αναλύσεις δείχνουν ότι μοντέλα μηχανικής μάθησης επιτυγχάνουν μέσο c-index 0,777, αναδεικνύοντας σαφή υπεροχή [18].
Ωστόσο, η ενσωμάτωση μοντέλων ML στην κλινική πράξη παραμένει περιορισμένη, κυρίως λόγω ανησυχιών για διαφάνεια, ασφάλεια δεδομένων και συμμόρφωση με κανονισμούς. Πρώτον, τα περισσότερα μοντέλα λειτουργούν ως «μαύρα κουτιά» (black boxes), παρέχοντας προβλέψεις χωρίς αιτιολόγηση — γεγονός που υπονομεύει την εμπιστοσύνη των κλινικών [22]. Δεύτερον, τα συστήματα αποθήκευσης ιατρικών δεδομένων βασίζονται σε κεντρικές βάσεις δεδομένων ευάλωτες σε κυβερνοεπιθέσεις: σύμφωνα με την έκθεση IBM 2024, ο κλάδος υγείας εμφάνισε το υψηλότερο μέσο κόστος παραβίασης (9,77 εκατομμύρια δολάρια) για 14 συνεχόμενα έτη [23]. Τρίτον, η συμμόρφωση με τον Γενικό Κανονισμό Προστασίας Δεδομένων (GDPR) απαιτεί ψευδωνυμοποίηση, διαφάνεια και ελέγξιμη διαχείριση συναίνεσης — απαιτήσεις που τα υπάρχοντα συστήματα σπανίως πληρούν [24].
Η παρούσα εργασία αποσκοπεί στην αντιμετώπιση αυτών των προβλημάτων μέσω ενός ολοκληρωμένου Συστήματος Υποστήριξης Κλινικών Αποφάσεων (CDSS) που συνδυάζει μηχανική μάθηση, blockchain και ερμηνεύσιμη ΤΝ. Αρχικά, θα αναλυθούν οι περιορισμοί των υπαρχόντων μοντέλων, στη συνέχεια θα παρουσιαστεί η προτεινόμενη λύση και τέλος θα συζητηθούν οι προκλήσεις εφαρμογής.

1.2.1 Στόχοι Εργασίας

Το πείραμα της παρούσας διπλωματικής εργασίας διακρίνεται σε τέσσερις βασικούς πυλώνες που αλληλοεπιδρούν, εκ των οποίων ο καθένας αποτελεί ένα ξεχωριστό τεχνολογικό υποσύστημα.

Στόχος 1 – Ανάπτυξη Pipeline Μηχανικής Μάθησης: Η εκπαίδευση και αξιολόγηση πολλαπλών μοντέλων (Logistic Regression, Random Forest, XGBoost, LightGBM) σε πραγματικά κλινικά δεδομένα από τη βάση MIMIC-IV, με τριπλό διαχωρισμό δεδομένων (εκπαίδευση, βαθμονόμηση, έλεγχος) σε επίπεδο ασθενούς, βελτιστοποίηση υπερπαραμέτρων.

Στόχος 2 – On-Chain Αποθήκευσης προβλέψεων: Ο σχεδιασμός έξυπνων συμβολαίων σε Solidity με μηχανισμό ελέγχου πρόσβασης βάση ρόλων (RBAC), λογική συναίνεσης ασθενή συμβατή με GDPR, μαζική δημοσίευση μέσω Merkle Trees και ελεγκτικό ίχνος ενεργειών.

Στόχος 3 – Ενσωμάτωση Ερμηνεύσιμης Τεχνητής Νοημοσύνης (XAI): Η εφαρμογή της μεθόδου SHAP σε δύο επίπεδα: πληθυσμιακό (global feature importance) και ατομικό (per-patient waterfall), ώστε κάθε πρόβλεψη να συνοδεύεται από τεκμηριωμένη αιτιολόγηση κατανοητή από κλινικούς ιατρούς.

Στόχος 4 – Διαδικτυακή Εφαρμογή (CDSS): Τριεπίπεδη αρχιτεκτονική (React frontend, FastAPI backend, blockchain/database layer) με αυθεντικοποίηση, ρόλους χρηστών, οπτικοποίηση SHAP αποτελεσμάτων και εξαγωγή αναφορών σε PDF.

1.3 Συνεισφορά της Εργασίας

Η παρούσα διπλωματική εργασία έρχεται να συνεισφέρει στην δημιουργία και ανάπτυξη ενός ολοκληρωμένου, ασφαλούς και ερμηνεύσιμου Συστήματος Υποστήριξης Κλινικών Αποφάσεων (CDSS). Αρχικά η παρούσα εργασία ξεχωρίζει για την ανάπτυξη και βαθμονόμηση μηχανικής μάθησης σε ευρείας κλίμακας δεδομένα, όπου σε αντίθεση με προσεγγίσεις της βιβλιογραφίας που περιορίζονται σε μικρά, ελλιπή ή τεχνητά δείγματα, η παρούσα διπλωματική έρευνα χρησιμοποιεί μια από τις δημοφιλέστερες διεθνώς βάσεις ιατρικών δεδομένων, τη MIMIC-IV, στην τελευταία έκδοση της. Σε αυτό το πλαίσιο επιλέχθηκαν 4 μοντέλα ως περισσότερο εμφανιζόμενα στη βιβλιογραφία και πραγματοποιήθηκε η σύγκριση τους, εφαρμόζοντας αυστηρές μεθόδους βαθμονόμησης, όπως η Ισοτονική και Σιγμοειδής Παλινδρόμηση, για την επίτευξη της ακρίβειας και της ευστοχίας των προβλέψεων. Ιδιαίτερα σημαντική προσθήκη είναι η ενσωμάτωση εξηγήσιμης τεχνητής νοημοσύνης για να αντιμετωπιστεί το πρόβλημα του «Μαύρου Κουτιού» των μοντέλων μηχανικής μάθηση, δηλαδή να επιτευχθεί καθολική και τοπική ερμηνεία των διαγνώσεων. Παράλληλα η εργασία δεν περιορίζεται στην απλή υλοποίηση ενός πειραματικού προτύπου Συστήματος Υποστήριξης Κλινικών Αποφάσεων αλλά ενσωματώνει στην αρχιτεκτονική του την τεχνολογία Blockchain. Η τελευταία έρχεται να διασφαλίσει την ακεραιότητα και τη μη αποκήρυξη των κλινικών προβλέψεων μέσω ενός έξυπνου συμβολαίου, γραμμένο σε γλώσσα Solidity, που περιλαμβάνει μηχανισμούς προστασίας των προσωπικών δεδομένων GDPR, έλεγχο πρόσβασής βάσει ρόλων (RBAC) και επιτρέπει τη φόρτωση δεδομένων μέσω αρχείων .csv, από τα οποία οι ευάλωτες πληροφορίες ενός ασθενούς, όπως τα προσωπικά στοιχεία του και η πιθανότητα κινδύνου αποθηκεύονται πάνω στην αλυσίδα (on-chain), ενώ οι υπόλοιπες πληροφορίες αποθηκεύονται εκτός αλυσίδας (off-chain).

1.4 Δομή της Εργασίας

Για την ομαλή και συστημική παρουσίαση της διπλωματικής εργασίας η οργάνωση της διαρθρώνεται σε επτά εξειδικευμένα κεφάλαια, τα οποία περιγράφονται ακολούθως. Αρχικά, το 1ο Κεφάλαιο αποτελεί την εισαγωγή και θέτει το γενικότερο πλαίσιο και κίνητρο πίσω από το ψηφιακό μετασχηματισμό της υγείας. Στο 2ο Κεφάλαιο πραγματοποιείται μια εκτενής αναδρομή στην υπάρχουσα επιστημονική βιβλιογραφία, αναλύονται σε βαθύ κλινικό υπόβαθρό ιατρικοί ορισμοί όπως η επιδημιολογία και η φαρμακευτική διαχείριση της Καρδιακής Ανεπάρκειας. Στο 3ο Κεφάλαιο παρουσιάζεται το θεωρητικό υπόβαθρο των εργαλείων που αξιοποιήθηκαν. Ειδικότερα περιλαμβάνονται η μαθηματική ανάλυση των αλγορίθμων ταξινόμησης, οι μετρικές αξιολόγησης της διακριτικής ικανότητας και βαθμονόμησης, η θεωρία παιγνίων πίσω από τις τιμές Shapley (SHAP), οι κρυπτογραφικές αρχές της τεχνολογίας Blockchain, όπως ο κατακερματισμός

SHA-256, HASH PEPPER SHA-256 και η Απόδειξη συμμετοχής (Proof-of-Stake - PoS). Στο 4ο Κεφάλαιο γίνεται μια εκτενής ανάλυση της αρχιτεκτονικής του προτεινόμενου συστήματος υποστήριξης κλινικών αποφάσεων, καθώς χρησιμοποιούνται Διαγράμματα Συνιστωσών, Διαγράμματα Ροής Δεδομένων και Διαγράμματα Ακολουθίας, μοντελοποιόντας τις δυναμικές αλληλεπιδράσεις σε ροές πληροφορίας μεταξύ της διεπαφής χρήστη, του backend και του έξυπνου συμβολαίου. Το 5ο Κεφάλαιο εστιάζει στο πειραματικό μέρος και την ανάπτυξη του λογισμικού, περιγράφοντας τη διαδικασία προ επεξεργασίας και μηχανικής χαρακτηριστικών της βάσης MIMIC-IV, η ροή εκπαίδευσης του μοντέλου και η πρόωρη παύση της, οι τεχνικές εξηγήσιμης τεχνητής νοημοσύνης, η προσβασιμότητα με βάση τον εξουσιοδοτημένο ρόλο και η ανάπτυξη της διεπαφής μεταξύ του React και ξεχωριστού API. Το 6ο Κεφάλαιο παρουσιάζει τα πειραματικά αποτελέσματα της εργασίας με την αναλυτική αξιολόγηση των επιδόσεων των μοντέλων βάση των μετρικών και των καμπυλών AUROC/AUPRC και της Bootstrap CI, οι καμπύλες απόφασης ενώ ακόμη εκτιμάται και η αποδοτικότητα του Blockchain ως προς την κατανάλωση καυσίμων (gas) και η βελτιστοποίηση κόστους μέσω Merkle δέντρων.

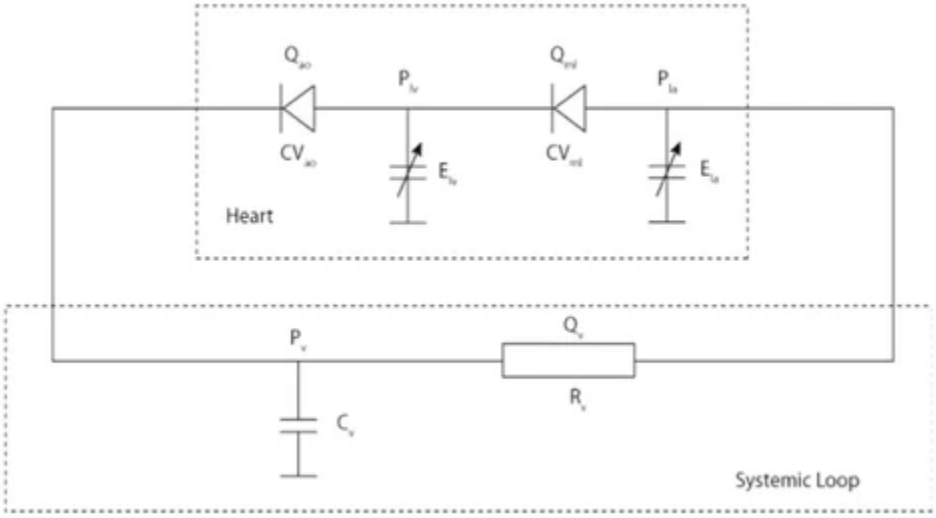
Κεφάλαιο 2 : Ορισμοί και Βιβλιογραφική Ανασκόπηση

2.1 Καρδιακή Ανεπάρκεια: Κλινικό Πλαίσιο

Η Καρδιακή Ανεπάρκεια (ΚΑ, Heart Failure) δεν αποτελεί μεμονωμένη παθολογική νόσο αλλά κλινικό σύνδρομο που οφείλεται στη δομική/λειτουργική ανωμαλία της καρδιάς και έχει ως αποτέλεσμα την αυξημένη ενδοκάρδια πίεση ή ανεπαρκή καρδιακή παροχή σε ηρεμία ή κατά τη διάρκεια της άσκησης. Σύμφωνα με τις κατευθυντήριες οδηγίες που δόθηκαν από την Ευρωπαϊκή Καρδιολογική Εταιρεία (European Society of Cardiology, ESC), το 2021, η ΚΑ ορίζεται ως κλινικό σύνδρομο που αποτελείται από κύρια συμπτώματα (δύσπνοια, οίδημα των αστραγάλων, κόπωση) τα οποία πολλές φορές συνοδεύονται από κλινικά σημεία (αυξημένη σφαγιτιδική φλεβική πίεση, πνευμονικοί ρόγχοι, περιφερειακό οίδημα) [26]. Η πιο συνηθισμένη πρόκληση καρδιακής ανεπάρκειας οφείλεται στη διαστολική ή συστολική δυσλειτουργία του μυοκαρδίου. Παράλληλα, άλλες ανωμαλίες όπως η παθολογία των βαλβίδων, του περικαρδίου και του ενδοκαρδίου, του καρδιακού ρυθμού και αγωγιμότητας συμβάλουν στην εμφάνιση ΚΑ σε ένα άτομο [26].

Στο παρελθόν έχει πραγματοποιηθεί σημαντική προσπάθεια από ερευνητές με σκοπό την συστημική προσέγγιση της καρδιακής λειτουργίας μέσω μοντέλων συγκεντρωμένων παραμέτρων (lumpted parameter models). Τα μοντέλα αυτά επιτρέπουν την προσομοίωση της αιμοδυναμικής συμπεριφοράς μέσω ηλεκτρικών ανάλογων [32]. Στην Εικόνα 5 παρουσιάζεται η ηλεκτρική αναπαράσταση του κυκλοφοριακού συστήματος της αριστερής καρδιάς, στο οποίο οι ανατομικές δομές αναπαρίστανται ως ηλεκτρικά στοιχεία. Ο αριστερός κόλπος (Ela,Elv) μοντελοποιούνται ως μεταβλητές ελαστικότητας (time-varying elastances) [32]. Στο κύκλωμα οι μεταβλητές αυτές

χαρακτηρίζονται ως πυκνωτές με μεταβαλλόμενη χωρητικότητα, παρομοιάζοντας τις διαδικασίες της συστολής και διαστολής, με τις διαδικασίες φόρτισης και εκ φόρτισης. Η αορτική (CVao) και μιτροειδής (CVmi) βαλβίδα αναπαρίστανται από διόδους, που επιτρέπουν τη μονοκατευθυντική διέλευση του "ρεύματος" (αίματος) και εμποδίζουν την παλινδρόμηση [32]. Η μοντελοποίηση της αιματικής κυκλοφορίας γίνεται ως ένα δίκτυο RC, στο οποίο η αντίσταση (Rv) εκφράζει το ιξώδες και τις τριβές των αγγείων, ενώ ο πυκνωτής (Cv) εκφράζει την ενδοτικότητα (compliance), δηλαδή η ικανότητα των αγγείων να αποθηκεύουν τον όγκο υπό πίεση. Τέλος η πίεση (P) συνάδει με την ηλεκτρική τάση και η παροχή (Q) συνάδει στην ένταση του ηλεκτρικού ρεύματος. Η μαθηματική μοντελοποίηση της καρδιακής λειτουργίας καθιστά δυνατή την παραγωγικότερη έρευνα, γεφυρώνοντας το χάσμα μεταξύ μηχανικής και ιατρικής [32].



Εικόνα 5 : Αναπαράσταση ενός Μοντέλου Συγκεντρωμένων Παραμέτρων (Lumped Parameter Model) του καρδιακού συστήματος, με σκοπό την προσομοίωση της αιμοδυναμικής [32].

Η απλούστερη ορολογία που εφαρμόζεται για την περιγραφή σοβαρότητας της καρδιακής ανεπάρκειας είναι η λειτουργική ταξινόμηση της Καρδιολογικής Εταιρείας της Νέας Υόρκης (NYHA), η οποία με βάση την ύπαρξη συμπτωμάτων και τη φυσική κατάσταση του ασθενούς διακρίνεται σε τέσσερις κλάσεις [26]. Στην **Κλάση Ι** αντιστοιχούν σε ασθενείς που δεν παρουσιάζουν κανένα περιορισμό στη σωματική δραστηριότητα. Στην **Κλάση ΙΙ** οι ασθενείς υπόκεινται σε ήπιο περιορισμό στην σωματική δραστηριότητα, κυρίως όταν ξεπερνούν την κατάσταση ηρεμίας [26]. Στην **Κλάση ΙΙΙ** οι ασθενείς παρουσιάζουν περιορισμό της σωματικής άσκησης, ενώ σε κατάσταση ηρεμίας εμφανίζουν άνεση, αλλά ακόμα και η λιγότερη από τη συνηθισμένη δραστηριότητα τους προκαλεί δύσπνοια, κόπωση ή έντονο αίσθημα παλμών. Τέλος, η **Κλάση ΙV** παρουσιάζεται ως η πιο κρίσιμη κατάσταση ενός ασθενούς που κινδυνεύει από ΚΑ, καθώς τα καρδιαγγειακά συμπτώματα εμφανίζονται ακόμα και σε κατάσταση ηρεμίας, με την οποιαδήποτε δυσφορία να αυξάνεται σε περίπτωση ελάχιστης σωματικής άσκησης [26].

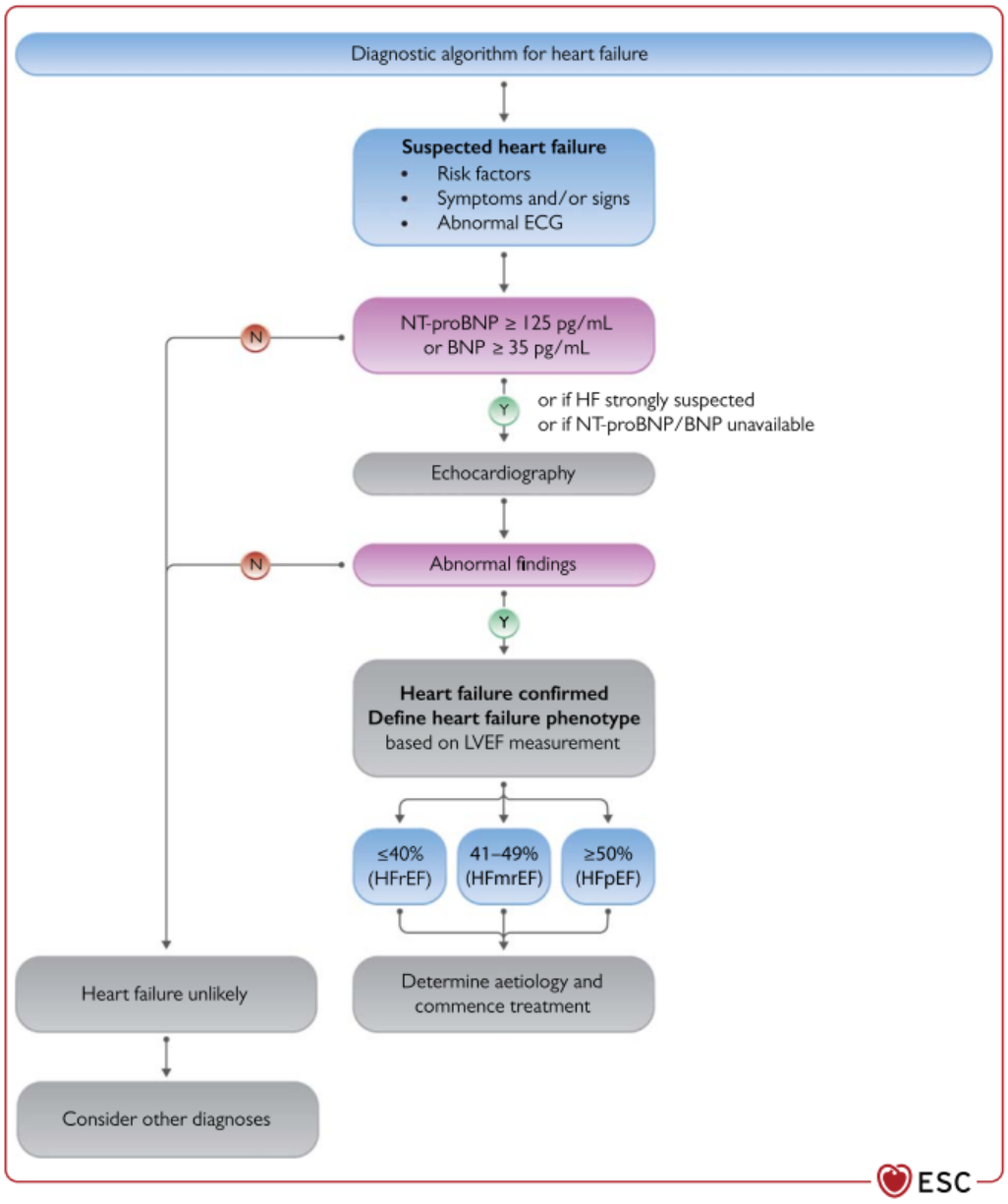
Παράλληλα η ΚΑ διακρίνεται σε φαινοτύπους με γνώμονα τη μέτρηση του κλάσματος εξώθησης της αριστερής κοιλίας (LVEF). Ως κλάσμα εξώθησης (Eject Fraction, EF) ορίζουμε το πηλίκο του ποσοστού αίματος που εξωθείται από μια καρδιακή κοιλία κατά την συστολή (Όγκος Παλμού, ΟΠ – Stroke Volume) προς τη συνολική ποσότητα αίματος που υπήρχε στην κοιλιά στο τέλος της διαστολής (Τελ διαστολικός Όγκος, LVESV) [26][28]. Η επιλογή της αριστερής κοιλίας ως κεντρικός δείκτης αναφοράς για την ΚΑ οφείλεται στην ιδιότητα που κατέχει ως κύριος κινητήρας της συστηματικής κυκλοφορίας του οξυγονωμένου αίματος. Οι τρείς διακριτοί φαινότυποι χωρίζονται σε τρείς ζώνες:

1. Αν το κλάσμα εξώθησης της αριστερής κοιλίας (LVEF) ορίζεται ως $\leq 40\%$, τότε οι ασθενείς εμφανίζονται με μειωμένη συστολική λειτουργία στην αριστερή κοιλία και η περίπτωση χαρακτηρίζεται ως HFrEF [26][28].
2. Αν το LVEF βρίσκεται μεταξύ των τιμών 41% και 49%, τα άτομα έχουνε ελαφρώς μειωμένη συστολική λειτουργία της αριστερής κοιλίας, δηλαδή HFmrEF. Οι τυχαιοποιημένες ελεγχόμενες δοκιμές (RCT) με LVEF στο παραπάνω διάστημα μας δείχνουν ότι οι θεραπείες που εφαρμόζονται στους ασθενείς είναι παρόμοιες με αυτές των HFpEF και HFrEF περιπτώσεων [26][28].
3. Αν το LVEF ορίζεται $\geq 50\%$, οι ασθενείς εμφανίζουν συμπτώματα και σημεία καρδιακής ανεπάρκειας ή/και αυξημένα νατριουρητικά πεπτίδια (NPs), τότε ανήκουν στην κατηγορία HFpEF (καρδιακή ανεπάρκεια με διατηρημένο κλάσμα εξώθησης) [26][27].

Πίνακας 2.1: Πίνακας Ταξινόμησης περιπτώσεων καρδιακής ανεπάρκειας με γνώμονα το κλάσμα εξώθησης [26][27]

Τύπος HF	**HFrEF**	**HFmrEF**	**HFpEF**
ΚΡΙΤΗΡΙΑ	1 Συμπτώματα ± Σημάδια	Συμπτώματα ± Σημάδια	Συμπτώματα ± Σημάδια
	2 LVEF ≤ 40%	41% ≤ LVEF ≤ 49%	LVEF≥ 50%
	3		Ενδείξεις καρδιακών δοκιμών ή/και λειτουργικών ανωμαλιών που συνάδουν με τη παρουσία διαστολικής δυσλειτουργίας της αριστερής κοιλίας, ύπαρξη αυξημένων πιέσεων πλήρωσης και αυξημένων νατριουρητικών πεπτιδίων

Η διάκριση αυτή έχει θεμελιώδη κλινική σημασία, καθώς τα θεραπευτικά πρωτόκολλα διαφέρουν ουσιαστικά μεταξύ των κατηγοριών. Πιο συγκεκριμένα στον διαγνωστικό αλγόριθμο της Ευρωπαϊκής Καρδιολογικής Εταιρείας στην **Εικόνα 6**, η κλινική υποψία κινδύνου επικυρώνεται μέσω της μέτρησης βιοδεικτών που τηρούν τα όρια αναφοράς των NT-proBNP≥ 125 pg/mL ή το BNP ≥ 25 pg/mL. Η διαγνωστική διαδρομή βαίνει προς την ολοκλήρωση της με τον υπερηχοκαρδιογραφικό έλεγχο που επιτρέπει τον ακριβή προσδιορισμό του φαινοτύπου με γνώμονα το κλάσμα εξώθησης και διακρίνεται στις περιπτώσεις HFrEF (≤ 40%), HFmrEF (41 − 49%) και HFpEF (≥ 50%). Οι τρείς περιπτώσεις καθορίζουν την θεραπευτική παρέμβαση στην οποία θα προβεί το ιατρικό προσωπικό.



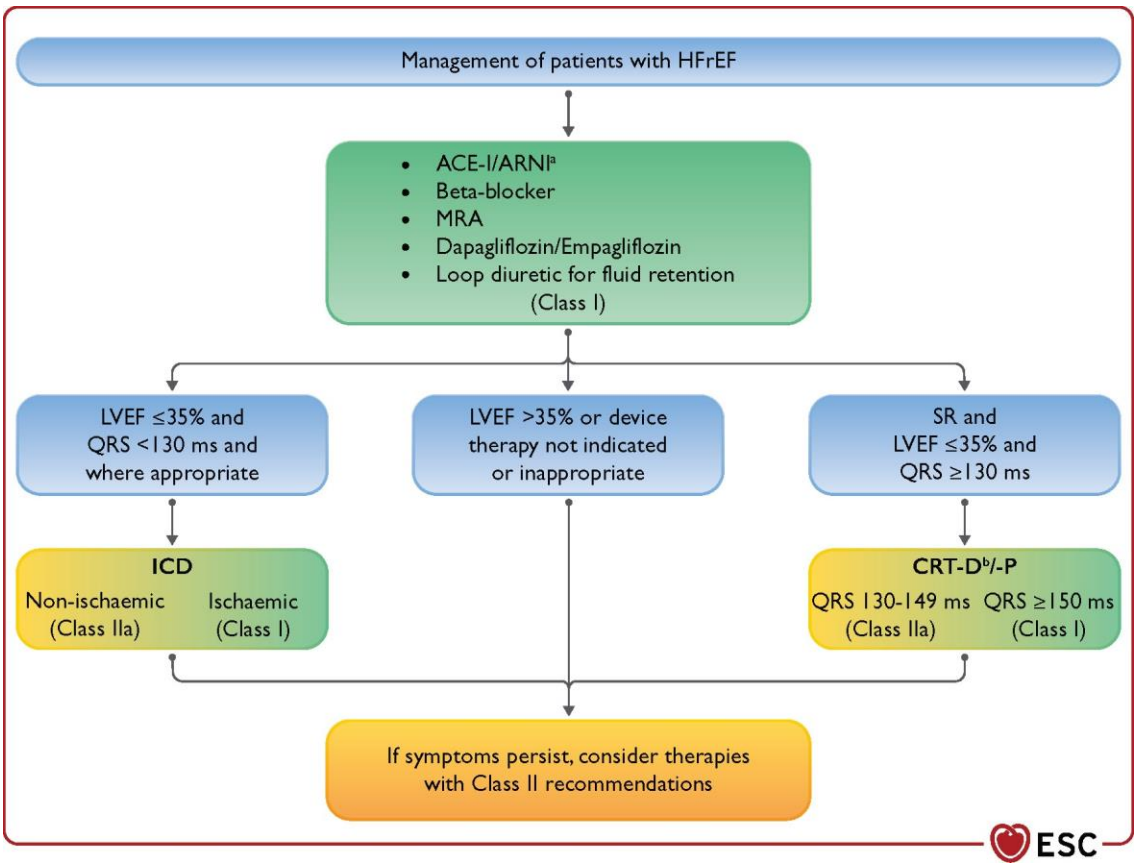
Εικόνα 6 : Διαγνωστικός αλγόριθμος της καρδιακής ανεπάρκειας [26].

2.1.2 Φαρμακευτική Θεραπεία για ασθενείς με καρδιακή ανεπάρκεια μειωμένου κλάσμα εξώθησης (HFrEF)

Η φαρμακευτική αγωγή χαρακτηρίζεται ως ακρογωνιαίος λίθος της θεραπείας για ασθενείς με HFrEF και πρέπει να πραγματοποιείται πριν από οποιαδήποτε άλλη θεραπεία με υποβολή σε συσκευές ή μη φαρμακολογίες παρεμβάσεις [15][29]. Η θεραπεία αυτή αποσκοπεί στη μείωση της θνησιμότητας, στην πρόληψη επαναλαμβανόμενων νοσηλειών λόγω επιδείνωσης της καρδιακής ανεπάρκειας και στη βελτίωση της κλινικής κατάστασης, της λειτουργικής ικανότητας και της

ποιότητας ζωής. Σε αυτήν εφαρμόζεται η συ χορήγηση 5 θεμελιωδών κατηγοριών φαρμάκων (Κατηγορία Σύστασης I):[16]

1. Αναστολείς των υποδοχέων αγγειοτενσίνης και νεπριλυσίνης (ARNI) ή αναστολείς μετατρεπτικού ενζύμου αγγειοτενσίνης (ACE-I). Εάν καμία από τις παραπάνω χορηγήσεις δεν είναι ανεκτή από τον ασθενή, συνιστώνται οι αναστολείς υποδοχέων αγγειοτενσίνης II (ARBs)
2. Β-αποκλειστές (Β)
3. Ανταγωνιστές υποδοχέων αλατοκορτικοειδών (MRAs)
4. Αναστολείς SGLT2 (δαπαγλιφλοζίνη ή εμπαγλιφλοζίνη), που μειώνουν τον κίνδυνο θανάτου και νοσηλειών.
5. Τα διουρητικά αγκύλης χορηγούνται αποκλειστικά για την ανακούφιση από τα συμπτώματα της συμφόρησης.

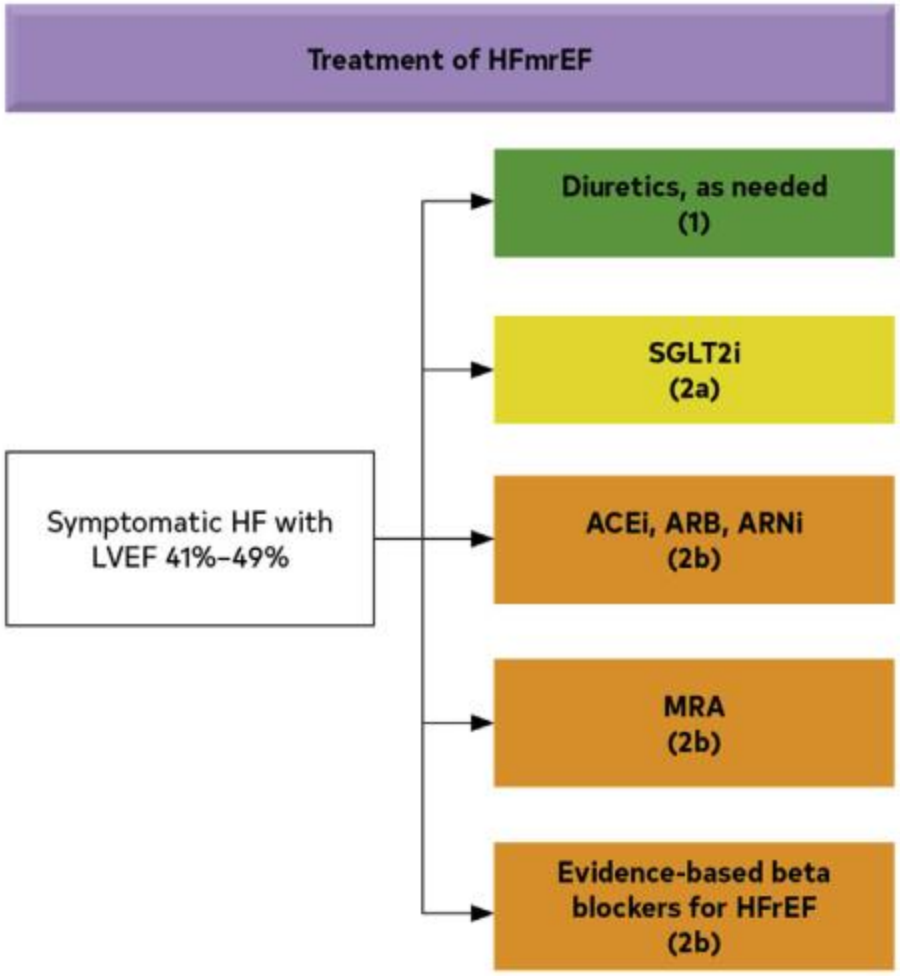


Εικόνα 7 : Θεραπευτικός αλγόριθμος διαχείρισης των ασθενών που πάσχουν από Καρδιακή ανεπάρκεια με μειωμένο κλάσμα εξώθησης (HFrEF)[26][27].

Μάλιστα η ερευνητική βιβλιογραφία μας αναφέρει πως η καρδιακή ανεπάρκεια μπορεί να προκληθεί από τη δυσλειτουργία της δεξιάς κοιλίας (ΔΚ) εξαιτίας της υπερφόρτωσης με πίεση ή όγκο μεταβάλλοντας την μηχανική λειτουργία της. Εκτός από αυτήν, υπάρχουν και άλλες αιτίες δυσλειτουργίας της ΔΚ που οδηγούν σε ΚΑ, όπως το έμφραγμα του μυοκαρδίου, αρρυθμογενής μυοκαρδιοπάθεια δεξιάς κοιλίας (ARVC) ή βαλβιδοπάθεια [15][29].

2.1.3 Θεραπεία για ασθενείς με καρδιακή ανεπάρκεια ελαφρώς μειωμένου κλάσμα εξώθησης (HFmrEF)

Η θεραπεία αφορά ασθενείς με συμπτώματα και LVEF μεταξύ των 41% και 49%. Ιδιαίτερη αναφορά γίνεται στην αναθεώρηση του όρου από «μεσαίου εύρους» (mid-range) σε «ήπια μειωμένο» (mildly reduced), διότι αναδρομικές αναλύσεις παρουσίασαν πως οι θεραπείες είναι παρόμοιες με εκείνες της HFrEF [25][31]. Η κατηγορία εμφανίζει δυναμικό χαρακτήρα, διότι περιλαμβάνει ασθενείς που το κλάσμα εξώθησης αριστερής κοιλίας μειώθηκε ελάχιστα λιγότερο από 50%, όσο και αυτούς που βελτιώθηκαν από την μειωμένη συστολή (LVEF≥ 40%). Σύμφωνα με την επικαιροποίηση του 2023 της ESC και του 2022 της AHA/ACC, οι αναστολείς SGLT2 (δαπαγλιφλοζίνη ή εμπαγλιφλοζίνη) λαμβάνονται με ισχυρή σύσταση για την μείωση του κινδύνου θανάτου και νοσηλειών. Από την άλλη πλευρά, τα κύρια φάρμακα που χρησιμοποιούνται και στην HFrEF (ACE-I,ARBs,Beta-blockers,ARNI,MRAs) διατηρούν ασθενότερη σύσταση. Δεν διατίθενται αρκετά δεδομένα για το πόρισμα που να συνιστούν χρήση CRT ή ICD συσκευών [26][31].



Εικόνα 8 : Αναλυτικό Διάγραμμα θεραπείας για ασθενείς με καρδιακή ανεπάρκεια ελαφρώς μειωμένου κλάσματος εξώθησης.

2.1.4 Θεραπεία για ασθενείς με καρδιακή ανεπάρκεια διατηρημένου κλάσματος εξώθησης (HFpEF)

Η θεραπευτική προσέγγιση της συγκεκριμένης κατηγορίας αποτελεί την μεγαλύτερη θεραπευτική πρόκληση, καθώς τα κλασικά φάρμακα, όπως αναστολείς ACE και ARBs, αποτυγχάνουν να καταδείξουν μείωση της θνησιμότητας ή της νοσηρότητας. Παρ' όλα αυτά, οι νεότερες κατευθυντήριες οδηγίες από την αναθεώρηση του ESC το 2023 εισάγουν νέες μεθόδους θεραπείας [15]. Οι αναστολείς SGLT2, συγκεκριμένα η δαπαγλιφλοζίνη και η εμπαγλιφλοζίνη αποτελούν την κύρια κατηγορία φαρμάκων για ασθενείς με HFpEF, αποδεικνύοντας με έρευνες την σημαντική μείωση στον κίνδυνο νοσηλειών και θάνατο λόγω ΚΑ. Παράλληλα, η χορήγηση διουρητικών παραμένει ο βασικότερος πυλώνας για την ανακούφιση από την κατακράτηση υγρών [21]. Σημαντικός εμφανίζεται ο έλεγχος συχνότητας για την διαχείριση της κολπικής μαρμαρυγής,

καθώς και η πρόληψη της παχυσαρκίας και του σακχαρώδους διαβήτη, ώστε να βελτιωθεί η κλινική κατάσταση του ασθενούς. Τέλος, βάσει των οδηγιών (AHA 2022), σε συγκεκριμένες περιπτώσεις ασθενών με HFpEF, μπορεί να χορηγηθούν Ανταγωνιστικοί Υποδοχείς Αλατοκορτικοειδών (MRAs), Αναστολείς ARNi ή ARBs , προκειμένου να ελαττωθεί ο αριθμός των νοσηλειών [31].

2.1.5 Επιδημιολογία της Καρδιακής Ανεπάρκειας

Σύμφωνα με την σύγχρονη ανασκόπηση του Savarese και των συνεργατών του, ο συνολικός αριθμός των ασθενών ΚΑ υπερβαίνει τα 64 εκατομμύρια παγκοσμίως, σε σύγκριση με το 1990 που καταγράφηκαν 33.5 εκατομμύρια περιστατικά, προκαλώντας αύξηση της τάξης του 91.9%. Παρ' όλα αυτά, το ηλικιακά σταθμισμένο ποσοστό επιπολασμού (age-standardized prevalence rate) ελαττώθηκε κατά 7.2% στο ίδιο περίπου χρονικό διάστημα, φτάνοντας τα 831 περιστατικά ανά 100.000 άτομα το 2017. Αποδεικνύοντας με τον τρόπο αυτόν ότι η ραγδαία αύξηση του συνολικού πληθυσμού ασθενών οφείλεται σε φαινόμενα γήρανσης του πληθυσμού και της πληθυσμιακής αύξησης [26].

Στις αναπτυγμένες χώρες, ο επιπολασμός κυμαίνεται μεταξύ 1-3% του ενήλικου πληθυσμού, ενώ η επίπτωση αναφέρεται σε περίπου 5 ανά 1000 ανθρωποέτη. Η νόσος εξαρτάται σε έντονο βαθμό από την ηλικία, καθώς ο επιπολασμός ανέρχεται από 1% στις ηλικίες άνω των 55 ετών σε πάνω από 10% στους ασθενείς άνω των 70. Στην Ευρώπη και τις ΗΠΑ, η επίπτωση κυμαίνεται από 1 έως 9 νέα περιστατικά ανά 1000 ανθρωποέτη. Ειδικότερα, στον ευρωπαϊκό χώρο εκτιμάται σε 3/1000 ανθρωποέτη για το ηλικιακό σύνολο ή 5/1000 στον ενήλικο πληθυσμό. Η επίπτωση παρουσιάζει μια σταθερότητα στις νέες διαγνώσεις ανά έτος, η οποία αποδίδεται στην καλύτερη πρόληψη των καρδιαγγειακών νοσημάτων [26][27].

Στην ερευνητική βιβλιογραφία, ως δείκτης σύγκρισης της επιδημιολογίας μεταξύ των χωρών (μελέτη Global Burden of Disease) χρησιμοποιείται ο Κοινωνικοδημογραφικός Δείκτης (SDI). Ο δείκτης κυμαίνεται στο εύρος από 0 (για τις αναπτυσσόμενες) έως το 1 (περισσότερο ανεπτυγμένες χώρες) και στηρίζεται σε τρεις παραμέτρους: Τη μέση εκπαίδευση που διαθέτουν οι πολίτες άνω των 15 ετών ενός πληθυσμού, την γονιμότητα του γυναικείου πληθυσμού κάτω των 25 ετών και το κατά κεφαλήν εισόδημα. Η ανάλυση των δεδομένων πάνω σε επιδημιολογικές έρευνες αναδεικνύει για την περίοδο 1990-2017 κρίσιμες αποκλίσεις στην παγκόσμια βαρύτητα της ΚΑ. Οι αναπτυγμένες χώρες που εμφανίζουν υψηλό SDI, ο σταθμισμένος ηλικιακά επιπολασμός σημείωσε σημαντική μείωση κατά 20.3%. Εξίσου σημαντική μείωση εμφάνισαν και τα ηλικιακά σταθμισμένα Έτη Ζωής με Αναπηρία (Years Lived with Disability - YLDs) των νοσούντων από ΚΑ, κατά 15.8%. Η βελτίωση των προηγούμενων ποσοστών οφείλεται στην εξέλιξη της τεχνολογίας στις θεραπείες και τις αποτελεσματικές μεθόδους πρόληψης. Αντιθέτως, καταγράφηκε αύξηση των YLDs κατά 10.7% σε αναπτυσσόμενες περιοχές με χαμηλό SDI, κάτι που οφείλεται στην έξαρση παραγόντων κινδύνου, όπως ο διαβήτης, η αρτηριακή υπέρταση, το κάπνισμα και η παχυσαρκία σε συνδυασμό με ητν περιορισμένη διαθεσιμότητα ιατρικών μέσων,

ακόμη και θεμελιωδών καρδιαγγειακών φαρμάκων (όπως η ασπιρίνη, beta-blockers, αναστολείς ACE και στεάτινες) [26][27].

Παρατηρείται μια εμφανής επιδημιολογική μεταστροφή, ο επιπολασμός της ΚΑ με Διατηρημένο Κλάσμα Εξώθησης (HFpEF) αυξάνεται διαρκώς, ενώ η ΚΑ με μειωμένο κλάσμα (HFrEF) σταθεροποιείται ή μειώνεται. Μάλιστα, σε κάποιες έρευνες παρατηρείται πως η HFpEF φτάνει να αποτελεί το 54% των περιπτώσεων [28].

Οι μακροοικονομικές ανισότητες απεικονίζουν την δυσμενή κλινική πρόγνωση και υψηλή θνησιμότητα. Αναφορές από μελέτες όπως η INTER-CHF και το REPORT-HF καταγράφουν πως η προσαρμοσμένη ανά ηλικία θνησιμότητα παρουσιάζεται αισθητά υψηλότερη κατά 26% στις χώρες χαμηλού και μεσαίου εισοδήματος (όπως η Αφρική και η Ινδία) σε σχέση με τις αναπτυγμένες χώρες που ανέρχεται στο 20% και 17%. Εκτός από το οικονομικό κεφάλαιο, άλλοι παράγοντες που επηρεάζουν την έκβαση της ΚΑ είναι η κοινωνική απομόνωση και ο αναλφαβητισμός, που διαδραματίζουν καθοριστικό ρόλο στην έκβαση της νόσου. Επομένως, καθίσταται επιτακτική η ανάγκη για επενδύσεις σε ενέργειες ενίσχυσης της δημόσιας υγείας, με βάση το γεωγραφικό βάρος της ΚΑ, που θα συμβάλλουν στην αντιμετώπιση της [15][35].

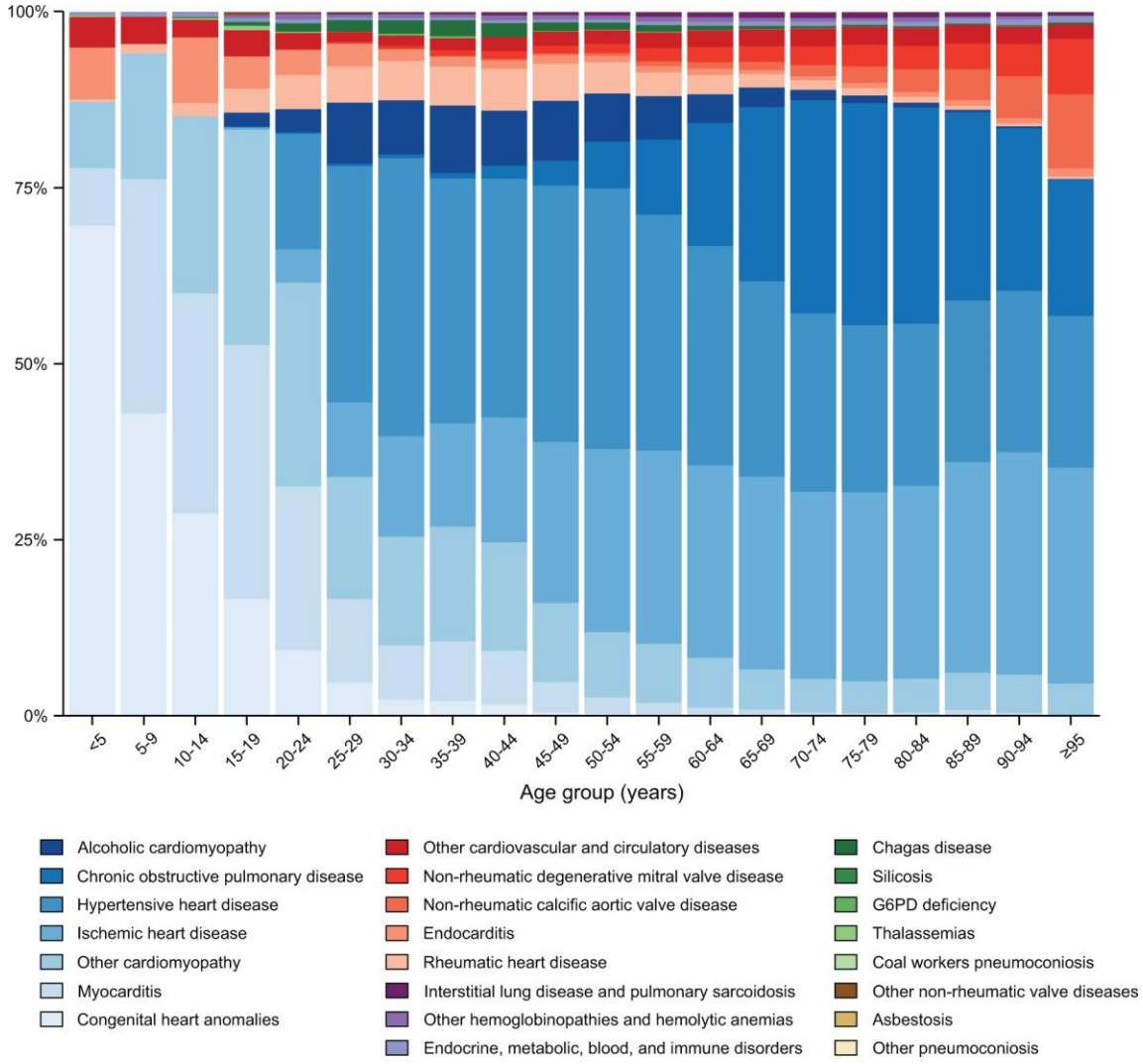
2.1.6 Αιτιολογία και Παράγοντες Κινδύνου

Η Ευρωπαϊκή Καρδιολογική Εταιρεία, κατατάσσει την αιτιολογία της ΚΑ σε τέσσερις κύριες κατηγορίες: την ισχαιμική καρδιοπάθεια, την αρτηριακή υπέρταση, τις βαλβιδοπάθειες και τις μυοκαρδιοπάθειες. Κυρίαρχη απειλή πρόκλησης της νόσου αποτελεί η ισχαιμική καρδιοπάθεια, στην οποία ανήκουν το έμφραγμα του μυοκαρδίου και η στεφανιαία νόσος, ευθυνόμενη για το 26.5% των περιπτώσεων παγκοσμίως. Με μικρή ποσοστιαία διαφορά ακολουθεί η υπερτασική καρδιοπάθεια, κατά 26.2%, με κύρια αιτία της ΚΑ στις αναπτυσσόμενες χώρες. Έπειτα ακολουθούν οι μυοκαρδιοπάθειες κατά 6.5% περίπου, συμπεριλαμβανομένης των διαιτητικών, των υπερτροφικών και διηθητικών νόσων [15][16]. Ενώ κρίσιμη αιτία στους γηραιότερους πληθυσμούς εμφανίζονται οι βαλβιδοπάθειες σε ποσοστό 6.8% περίπου, με σημαντικές παθήσεις να αποτελούν η μη ρευματική εκφυλιστική νόσος της μιτροειδούς βαλβίδας (NDMVD), η μη ρευματική απασβεστωμένη νόσος της αορτικής βαλβίδας (NCAVD) και η ρευματική καρδιοπάθεια (Rheumatic heart disease) [15][16].

Παράλληλα η δραματική επιδείνωση της ΚΑ προκαλείται από τους παράγοντες κινδύνου. Ο σακχαρώδης διαβήτης, η παχυσαρκία και η χρόνια νεφρική νόσος αποτελούν καθοριστικούς παράγοντες κινδύνου. Ενδεικτικά, τα ιστορικά δεδομένα από την έρευνα του Wang και των συνεργατών του παρουσιάζουν αύξηση στον επιπολασμό του διαβήτη από 25% σε 38%, της παχυσαρκίας από 33% σε 47% και της νεφρικής νόσου από 35% σε 46%. Οι 3 αυτοί παράγοντες χαρακτηρίζονται από ένα κοινό σύνθετο παθοφυσιολογικό υπόβαθρο, όπως η ενδοθηλιακή δυσλειτουργία και η συστηματική φλεγμονή, που δυσχεραίνουν την πρόγνωση της ΚΑ. Επίσης η Χρόνια Αποφρακτική Πνευμονοπάθεια (ΧΑΠ) παρουσιάζει υψηλά ποσοστά συν νοσηρότητας με

την ΚΑ με έρευνες να δείχνουν ότι το 20.5% των ηλικιωμένων ασθενών που πάσχουν από ΧΑΠ ενδέχεται να πάσχουν και από μη αναγνωρισμένη ΚΑ [26][30].

Τέλος, η δημογραφική γήρανση αποτελεί τον κύριο παράγοντα αύξησης του επιπολασμού της ΚΑ σε παγκόσμια κλίμακα. Ο κίνδυνος ανάπτυξης ΚΑ παρουσιάζει εκθετική αύξηση με την ηλικία. Σε άτομα με ηλικία άνω των 70 ετών, ο επιπολασμός υπερβαίνει το 10%. Στη μελέτη NHANES, παρατηρείται κλιμάκωση των ποσοστών σε μεγαλύτερες ηλικιακές ομάδες, όπου για άτομα των 80 ετών και άνω, ο επιπολασμός να ανέρχεται σε 10.6% στους άνδρες και 13.5% στις γυναίκες [26][28].



Εικόνα 9 : Ταξινόμηση της Καρδιακής Ανεπάρκειας σύμφωνα με το Κολλέγιο Καρδιολογίας της Αμερικής/ Αμερικανική Καρδιολογική Εταιρεία (AHA/ACC) και τη Καρδιολογική Εταιρεία της Νέας Υόρκης (NYHA).

2.1.7 Βιοδείκτες: Νατριουρητικά Πεπτίδια και Τροπονίνες

Τα νατριουρητικά πεπτίδια, ιδιαίτερα τα Β-τύπου νατριουρητικό πεπτίδιο (BNP) και το Ν-τελικό προ πεπτίδιο αυτού (NT-proBNP), αποτελούν τον «χρυσό κανόνα» στην πρόωρη διάγνωση της καρδιακής ανεπάρκειας. Ο παθοφυσιολογικός μηχανισμός έκκρισης και σύνθεσης τους πυροδοτείτε από την αυξημένη μηχανική τάση και διάταση του μυοκαρδιακού τοιχώματος, λειτουργώντας ως δείκτης του αιμοδυναμικού στρες. Σε μη οξεία περιστατικά, η εμφάνιση τιμών κάτω των καθορισμένων ορίων (< 35 pg/mL για το BNP και <125 pg/mL για το NT-proBNP) παρουσιάζουν εξαιρετικά αρνητική προγνωστική αξία, κυμαινόμενη μεταξύ των 94% και 98%, καθιστώντας την εμφάνιση νόσου απίθανη [26].

Ωστόσο, η ερμηνευσιμότητα των επιπέδων των νατριουρητικών πεπτιδίων αντιμετωπίζει προκλήσεις λόγω της ύπαρξης «γκρίζων ζωνών» (gray zones), που οι τιμές επηρεάζονται από εξωγενείς και ενδογενείς παράγοντες. Η προχωρημένη ηλικία, η κολπική μαρμαρυγή και η χρόνια νεφρική ανεπάρκεια, οδηγούν σε ψευδώς υψηλές τιμές και αποτελούν παράγοντες μείωσης της διαγνωστικής ακρίβειας. Σε αντίθετα αποτελέσματα οδηγεί η παχυσαρκία, όπου η ενισχυμένη κάθαρση των πεπτιδίων και των λιπώδη ιστών αποδίδει δυσανάλογες συγκεντρώσεις νατριουρητικών πεπτιδίων. Αυτή η πολυπλοκότητα των βιοδεικτών καθιστά κρίσιμη την ενσωμάτωση τους σε σύνθετα υπολογιστικά μοντέλα που υπολογίζουν ταυτόχρονα πολλαπλές κλινικές παραμέτρους [26].

Παράλληλα με τα νατριουρητικά πεπτίδια, κομβικό ρόλο στην αξιολόγηση κινδύνου (risk stratification) της νόσου αποτελούν οι καρδιακές τροπονίνες υψηλής ευαισθησίας (hs-cTnT/l). Τα αυξημένα επίπεδα τροπονίνης, έστω και σε υποκλινικό επίπεδο, υποδηλώνουν εκτεταμένη μυοκαρδιακή βλάβη ή νέκρωση και σχετίζονται ανεξάρτητα με δυσμενή πρόγνωση και αυξημένα ποσοστά θνητότητας. Επομένως, ο συνδυασμός των νατριουρητικών πεπτιδίων (NPs) και οι τροπονίνες συμβάλλουν καταλυτικά στην καλύτερη αξιολόγηση της κλινικής κατάστασης του ασθενούς, παρέχοντας βασικά δεδομένα για την εκπαίδευση μοντέλων πρόβλεψης που σκοπός τους είναι η ανίχνευση της νόσου [26][28].

2.2 Μηχανική Μάθηση και Ερμηνεύσιμη Τεχνητή Νοημοσύνη στην Κλινική Πράξη

Στην παρούσα ενότητα γίνεται ανασκόπηση στη χρήση της Τεχνητής Νοημοσύνης, ειδικότερα των υποπεδίων της Μηχανικής Μάθησης και της Ερμηνεύσιμης Τεχνητής Νοημοσύνης, στον ιατρικό τομέα. Συγκεκριμένα εξετάζεται η χρήση των τεχνολογιών αυτών της ΤΝ στην διάγνωση, πρόγνωση και πρόβλεψη της καρδιακής ανεπάρκειας, αναδεικνύοντας την θετική επίδραση τους στην κλινική πράξη. Παράλληλα, διερευνάται η προσφορά την Ερμηνεύσιμης Τεχνητής Νοημοσύνης, ενός κλάδου που έρχεται να γεφυρώσει το χάσμα μεταξύ της υπολογιστικής απόδοσης των μοντέλων της μηχανικής μάθησης και της ανθρώπινης εμπιστοσύνης του ιατρικού προσωπικού.

2.2.1 Η Μηχανική Μάθηση στην Υγειονομική Περίθαλψη

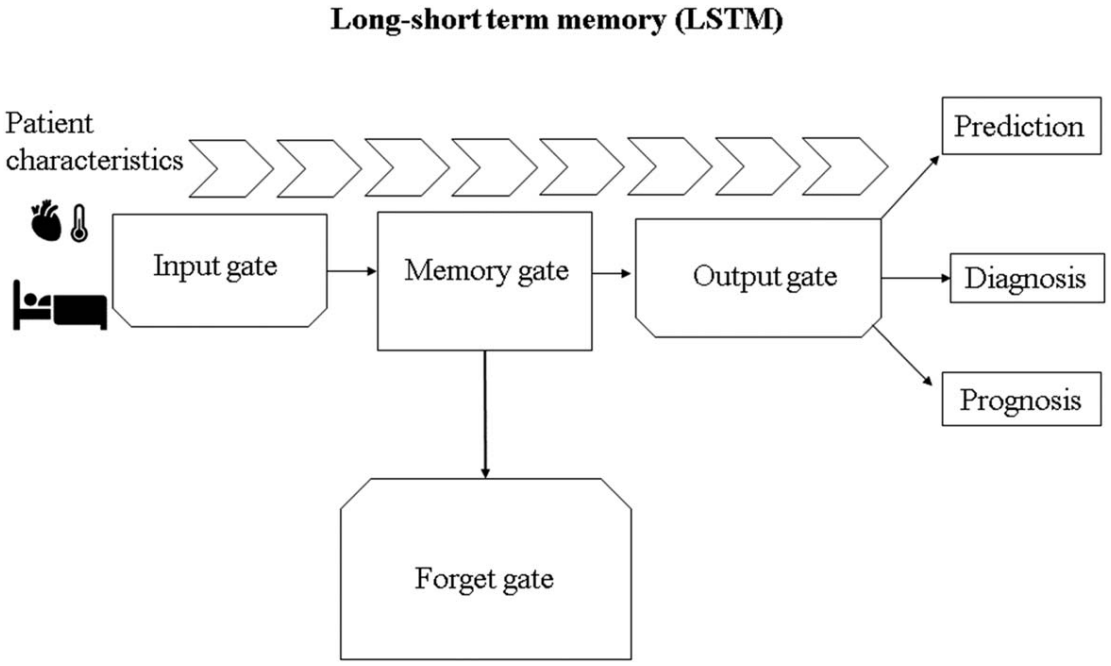
Η κύρια εφαρμογή της Μηχανικής Μάθησης στον ιατρικό τομέα είναι η ικανότητα της εξαγωγής γνώσης μέσω πολύπλοκων αλγορίθμων που εντοπίζουν κρυμμένα μοτίβα (pattern recognition) από μεγάλα ιατρικά δεδομένα, βελτιώνοντας σημαντικά την ακρίβεια της διάγνωσης και της πρόγνωσης. Με γνώμονα την έρευνα του Topol, διανύουμε την εποχή της "ιατρικής υψηλών επιδόσεων" (high-performance medicine), κατά την οποία η σύγκλιση της ανθρώπινης και τεχνητής νοημοσύνης αναδιαμορφώνει εποικοδομητικά την υγειονομική περίθαλψη, ως προς την παραγωγικότητα και την βελτίωση της ακρίβειας των αποτελεσμάτων [12][13].

Ειδικότερα στην ανάλυση ιατρικών εικόνων, η "βαθιά μάθηση" (deep learning), που είναι υποπεδίο της μηχανικής μάθησης, σημειώνει εντυπωσιακή ικανότητα στην αναγνώριση και ταξινόμηση στοιχείων. Οι βασικότεροι τομείς ιατρικής που βλέπουμε χρήση της βαθιάς μάθησης είναι η Ακτινολογία & Παθολογοανατομία, η Δερματολογία, η Οφθαλμολογία, η Καρδιολογία. Στον κλάδο της Ακτινολογίας, χρησιμοποιούνται αλγόριθμοι βασισμένοι σε αξονικές τομογραφίες με σκοπό την διάγνωση εγκεφαλικών προβλημάτων, την ερμηνεία μαστογραφιών και τον προ συμπτωματικό έλεγχο για καρκίνο στον πνεύμονα. Στην Οφθαλμολογία, η ανάλυση φωτογραφιών Βυθού (Fundus) αποτελεί χαρακτηριστικό παράδειγμα της αυτόνομης ΤΝ στην ανίχνευση διαβητικής αμφιβληστροειδοπάθειας, ενώ έχουν αναπτυχθεί μοντέλα για τη διάγνωση καταρράκτη. Επιπλέον, στη Δερματολογία έρευνες επισημαίνουν την ακρίβεια των συνελικτικών νευρωνικών δικτύων για τη ταξινόμηση υπότυπων δερματικού καρκίνου και τον εντοπισμό μελανώματος. Τέλος, στη Καρδιολογία, οι δυνατότητες της ΤΝ ποικίλλουν, σε μερικές με το να γίνεται ιδιαίτερη αναφορά όπως η ανίχνευση αρρυθμιών από ηλεκτροκαρδιογραφήματα, ο αυτόματος προσδιορισμός του κλάσματος εξώθησης μέσω υπερηχοκαρδιογραφίας και η πρόγνωση κινδύνου για ασθενείς από την καρδιακή ανεπάρκεια [12][13].

Πέραν της αξιοσημείωτης ερμηνείας των απεικονιστικών δεδομένων, ο Rajpurkar et al. στην έρευνα του τονίζει ότι η πρόκληση μετατοπίζεται στα πολυδιάστατα δεδομένα, όπως τα Ηλεκτρονικά Αρχεία Υγείας (Electronic Health Records) [13]. Τα Ηλεκτρονικά Αρχεία Υγείας ή αλλιώς Ηλεκτρονικά Ιατρικά Αρχεία (Electronic Medical Records), είναι η ψηφιακή έκδοση καταγραφής του ιατρικού ιστορικού ενός ασθενούς. Δεν είναι απλά μια στατική καταγραφή, αλλά μια δυναμική συλλογή πληροφοριών που μπορεί να περιλαμβάνει ετερογενή δεδομένα, όπως η φαρμακευτική αγωγή, τα εργαστηριακά αποτελέσματα, οι εμβολιασμοί και συνταγογραφήσεις, τα δημογραφικά στοιχεία και άλλα τα οποία συνοδεύουν το άτομο σε όλη την διάρκεια της ζωής του. Αναφέραμε προηγουμένως, ότι τα μοντέλα μηχανικής μάθησης μπορούν και αντλούν γνώση από πολύπλοκά δεδομένα έναντι των παραδοσιακών μοντέλων στατιστικής [13]. Πράγματι, το μεγαλύτερο μέρος κρίσιμων πληροφορίων στα Ηλεκτρονικά Αρχεία Υγείας βρίσκεται σε μη δομημένη μορφή, όπως ένα ελεύθερο κείμενο, γεγονός που οδηγεί στην συνήθη εμφάνιση ελλιπών δεδομένων (missing data). Αυτό ορίζεται ως ένα κοινό πρόβλημα των κοινών στατιστικών μοντέλων και των πολυτροπικών μοντέλων μηχανικής μάθησης. Παρόλα αυτά, η πρόοδος στην

επεξεργασία φυσικής γλώσσας έχει επιτρέψει την αποτελεσματική εξαγωγή πληροφοριών και τη πρόβλεψη κρίσιμων κλινικών συμβατών που υπό άλλες συνθήκες ενδέχεται να περνούσαν απαρατήρητες. Τέτοιες περιπτώσεις είναι ο κίνδυνος μιας απρογραμμάτιστης επανεισαγωγής ενός ασθενούς, η ελάχιστή ή υπερβολική διάρκεια νοσηλείας του ή η ενδονοσκομειακή θνησιμότητα [13].

Μια από τις κρισιμότερες παθήσεις που εμφανίζει σημαντική χρησιμότητα η μηχανική μάθηση είναι η Καρδιακή Ανεπάρκεια (ΚΑ). Η εξέλιξη της ΚΑ επηρεάζεται από πολλούς παράγοντες με αποτέλεσμα η σωστή αξιοποίηση των ηλεκτρονικών φακέλων και της μηχανικής μάθησης να παρέχει τεράστιες δυνατότητες στην θεραπεία των ασθενών. Σε πρόσφατη ανασκόπηση των Saqib et al. (2024) [36], αποδεικνύεται η αποτελεσματικότητα των Βαθιών Νευρωνικών Δικτύων (Deep Neural Networks - DNNs), πιο συγκεκριμένα μιας ειδικής αρχιτεκτονικής τους των δικτύων Μακράς Βραχυπρόθεσμης Μνήμης (Long Short-Term Memory - LSTM), στην ανάλυση ακολουθιακών δεδομένων (sequential data) [36].

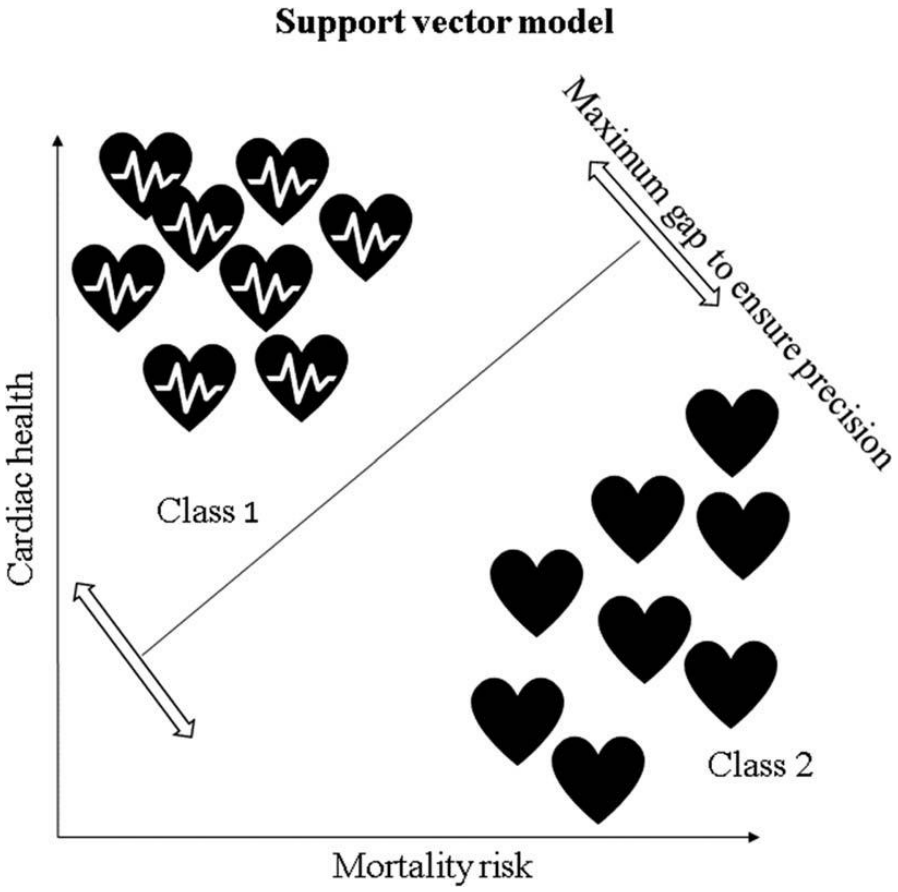


Εικόνα 10 : Σχηματική αναπαράσταση της λειτουργίας ενός δικτύου Long Short-Term Memory (LSTM) με στόχο την ανάλυση κλινικών δεδομένων[36].

Το διάγραμμα αναπαριστά την δομή λειτουργίας του νευρωνικού δικτύου LSTM. Αρχικά λαμβάνει τα δεδομένα του ασθενή μέσω της πύλης εισόδου (input gate). Στην συνέχεια στην πύλη μνήμης τροφοδοτείται από τα επιλεγμένα δεδομένα της πύλης εισόδου. Οι ερευνητές παρομοιάζουν ρητά την λειτουργία της πύλης μνήμης ως ένας "ιμάντας μεταφοράς" (conveyor belt) που μεταφέρει τις πληροφορίες παράλληλα με τον χρόνο. Ταυτόχρονα με την μεταφορά των

δεδομένων μέσω της πύλης μνήμης πραγματοποιείται η διαγραφή των δεδομένων που δεν θεωρούνται πλέον απαραίτητα από τον αλγόριθμο μέσω της πύλης λήθης (forget gate) [36]. Για παράδειγμα, αν μια παλιά μέτρηση δεν έχει πια καμία σημασία για την κατάσταση του ασθενή, το σύστημα δεν συνεχίζει να τη "θυμάται". Τέλος, η πύλη εξόδου (output gate) λαμβάνει το τελικό αποτέλεσμα από τον ιμάντα μνήμης . Στη περίπτωση της καρδιακής ανεπάρκειας, το αποτέλεσμα που καταλήγει στην πύλη εξόδου θα μπορούσε να είναι μια πρόβλεψη, μια διάγνωση ή την πρόγνωση για εξέλιξη της νόσου του ασθενούς [36]. Άλλο ενδεικτικό παράδειγμα είναι η έρευνα των Wang et al. που χρησιμοποιεί τον αλγόριθμο LSTM σε ένα μεγάλο εθνικό σύνολο δεδομένων από εμπορικές ασφαλιστικές εταιρείες των ΗΠΑ για να εντοπίσει ασθενείς που πάσχουν από ΚΑ με μειωμένο κλάσμα εξώθησης (HFrEF), αναλύοντας τις διαχρονικές μεταβολές των ζωτικών σημείων τους [36].

Επιπλέον, ιδιαίτερη παρουσιάζεται η εφαρμογή των Μηχανών Διανυσμάτων Υποστήριξης (Support Vector Machines - SVM) στην ακριβή ταξινόμηση του κινδύνου των ασθενών. Στο συγκεκριμένο πλαίσιο, οι Gao et al. κατασκεύασαν ένα προγνωστικό μοντέλο SVM που προέβλεπε με υψηλή ακρίβεια την θνησιμότητα, σε χρονικό ορίζοντα διετίας, σε ασθενείς με ΚΑ διατηρημένου κλάσματος εξώθησης (HFpEF), εξετάζοντας συνδυαστικά τους κυκλοφορόντες βιοδείκτες. Αξιοσημείωτο εύρημα παρουσιάζεται η ανάδειξη του μοντέλο για την προγνωστική αξία της ενδογλίνης (endoglin). Η εμφάνιση υψηλών επιπέδων ενδογλίνης αντιστοιχήθηκε άμεσα με διάφορες καρδιολογικές παθολογίες, όπως η φλεγμονή, η ενδοθηλιακή δυσλειτουργία, η ίνωση μυοκαρδίου και η αγγειακή αναδιαμόρφωση, όντας ισχυρό δείκτη για τη σοβαρότητα της καρδιακής βλάβης και το μακροπρόθεσμο κίνδυνο [36][37].



Εικόνα 11 : Σχηματική αναπαράσταση του μοντέλου SVM για την πρόβλεψή της θνησιμότητας σε ασθενείς που πάσχουν από Καρδιακή Ανεπάρκεια [36].

Σε μια άλλη έρευνα των Mpanya et al., στο ενδεχόμενο πρόβλεψης της θνησιμότητας, το SVM μοντέλο τους διαχώρισε τους ασθενείς με βάση συγκεκριμένους κλινικούς παράγοντες. Οι δείκτες με την μεγαλύτερη βαρύτητα με βάση το μοντέλο ήταν η φαρμακευτική αγωγή (π.χ. χρήση φουροσεμίδης, β-αναστολέων και σπιρονολακτόνης), τα ευρήματα από τη φυσική εξέταση (όπως το πρώιμο διαστολικό φύσημα), οι συν νοσηρότητες (όπως η στεφανιαία νόσος) [38].

Η μελέτη των Li et al. [66], εφάρμοσε τον αλγόριθμο ενισχυμένης μάθησης XGBoost, ο οποίος προσπέρασε σημαντικά τα παραδοσιακά μοντέλα λογιστικής παλινδρόμησης (LR) στη πρόβλεψη της θνησιμότητας ασθενών με ΚΑ εντός των Μονάδων Εντατικής Θεραπείας (ΜΕΘ), ξεχωρίζοντας το επίπεδο αζώτου ουρίας αίματος (BUN) ως τον βαρύτερο προγνωστικό βιολογικό παράγοντα [66].

Ωστόσο η εφαρμογή των μοντέλων αυτών, λόγω προκλήσεων, έρχεται αντιμέτωπη με περιορισμούς. Ένα από τους βασικότερους περιορισμούς εντοπίζεται στη γενικευσιμότητα (generalization), δηλαδή στο κατά πόσο ένα μοντέλο που προέρχεται από διαφορετικά νοσοκομεία ή ετερογενείς πληθυσμούς μπορεί να ανταποκριθεί με την ίδια ακρίβεια. Επίσης, το πρόβλημα της εγγενής αδιαφάνεια των μοντέλων, γνωστό και ως πρόβλημα του "μαύρου κουτιού" (black box), προκαλεί επιφυλακτικότητα και δυσπιστία στους γιατρούς, θέτοντας επιτακτική την ανάγκη για

ενσωμάτωση των εξηγήσιμων συστημάτων (explainable AI) ώστε να τεκμηριώνουν με σαφήνεια τις αποφάσεις τους [39].

2.2.2 Μηχανική Μάθηση στη πρόβλεψη της Καρδιακής Ανεπάρκειας

Η Μηχανική Μάθηση παρουσιάζεται ως ένα ισχυρό και υποσχόμενο εργαλείο για την πρόβλεψη, διάγνωση και πρόγνωση της ΚΑ. Συγκεκριμένα, τα μοντέλα ML αξιοποιούνται στη ταξινόμηση των τριών βασικών φαινοτύπων της ΚΑ, την έγκαιρη ανίχνευση νόσου από ηλεκτρονικά αρχεία υγείας και τη πρόβλεψη κρίσιμων παραγόντων, όπως η θνητότητα (mortality) και ο κίνδυνος επανεισαγωγής σε κλινικές μονάδες όπως το νοσοκομείο (readmission) [36]. Η δυνατότητα διαχείρισης των μη γραμμικών σχέσεων μεταξύ εκατοντάδων μεταβλητών θέτει τα μοντέλα σημαντικά αποτελεσματικά στη διαστρωμάτωση πολυδιάστατων δεδομένων των ασθενών [36].

Η βελτιστοποίηση των μοντέλων μηχανικής μάθησης αποτελεί καθημερινή πρόκληση για την επιστημονική κοινότητα, καθώς η πολυπλοκότητα των αλγορίθμων της δεν μεταφράζεται με την αποδοτικότητα, χαρακτηριστικό που συγκρούεται με τα παραδοσιακά μοντέλα. Ο Sun και οι συνεργάτες τους σε μια πρόσφατη μετάναλυση με θέμα τη σύγκριση των μοντέλων ΜΜ και των παραδοσιακών μοντέλων πρόβλεψης καταρρίπτει τον μύθο πως τα μοντέλα ΜΜ είναι παντού καλύτερα [21]. Το μεγάλο πλεονέκτημα των μοντέλων ΜΜ είναι η ικανότητα ενσωμάτωσης πολυδιάστατων και μη γραμμικών δεδομένων στην εκπαίδευση τους, βρίσκοντας περίπλοκα μοτίβα, τα οποία δεν είναι εφικτό να εντοπιστούν από τις παραδοσιακές μεθόδους. Αυτό φάνηκε και στην πρόβλεψη θανάτου της έρευνας στην οποία τα μοντέλα ΜΜ πήραν βαθμό C-index 0.777, υψηλότερο από το 0.733 της στατιστικής [21]. Ο θάνατος χαρακτηρίζεται ως ένα απόλυτο και εξαιρετικά πολύπλοκο γεγονός, όπου ο συνδυασμός δεκάδων εξετάσεων συμβάλει καθοριστικά στην πρόβλεψης της έκβασής του [21].

Η μελέτη ορόσημο που δημοσιεύθηκε στο JAMA από τους Desai et al. [40] επιβεβαιώνει τα παραπάνω ευρήματα [40]. Η έρευνα έγινε πάνω σε πραγματικά δεδομένα 9502 ηλικιωμένων ασθενών με Καρδιακή Ανεπάρκεια. Στην μελέτη αυτή, η ερευνητική ομάδα συνέκρινε δύο προηγμένα μοντέλα, το Gradient Boosting Machine (GBM) και τη Λογιστική Παλινδρόμηση (LR). Τα αρχικά αποτελέσματα έδειξα πως το GBM παρείχε σταθερά υψηλότερη ικανότητα διάκρισης στην πρόβλεψη νοσηλείας για ΚΑ με C-index 0.745 έναντι 0.707 της LR και στη πρόβλεψη θνησιμότητας με C-index 0.727 έναντι 0.724 της LR [40]. Οι ερευνητές απέδειξαν πως η μέθοδος της ΜΜ υπερείχε οριακά στην κλινική πράξη, καθώς τα περισσότερα δεδομένα ήταν ρηχά διοικητικά/ασφαλιστικά. Για παράδειγμα, η πρόβλεψη θνησιμότητας με το GBM ήταν C-index 0.727 και με το LR ήταν 0.724. Η δυναμικότητα των αλγορίθμων ξεκλειδώθηκε όταν το μοντέλο GBM εκπαιδεύτηκε σε κλινικά δεδομένα από Ηλεκτρονικούς Φακέλους Υγείας (EMR), που περιλάμβαναν συνεχείς χρονικά μεταβλητές, όπως τα επίπεδα αζώτου ουρίας αίματος (BUN), καλίου και το κλάσμα εξώθησης. Η μη γραμμική φύση του μοντέλου επέτρεψε τη πρόβλεψη της νοσηλείας φτάνοντας ακόμη σε βαθμό C-index 0.778 [40].

Εντούτοις, τα μοντέλα ΜΜ κρίθηκαν να υστερούν στην πρόβλεψη επανεισαγωγής στο νοσοκομείο, στην οποία τα παραδοσιακά μοντέλα στέφθηκαν νικητές με βαθμό C-index 0.678 έναντι 0.660 [21]. Ο λόγος είναι ότι η επανεισαγωγή στο νοσοκομείο επηρεάζεται από ευρύ φάσμα παραγόντων πέραν των βιολογικών, όπως η οικονομική κατάσταση ενός ασθενή, η προσβασιμότητα σε κλινική μονάδα [21].

Μέσα από την ερευνητική βιβλιογραφία μας εντοπίστηκα τρία μεγάλα μεθοδολογικά προβλήματα της μηχανικής μάθησης. Πρώτον, η ερευνητική κοινότητα πετάνε μέσα στον αλγόριθμο μηχανικής μάθησης μεταβλητές χωρίς να τις επεξεργαστούν, πολλές φορές χωρίς κλινική λογική. Η αλόγιστη χρήση μεταβλητών μειώνει την πρακτικότητα του μοντέλου, διότι σε επείγοντα περιστατικά ο γιατρός δεν έχει την χρονική πολυτέλεια να εισάγει 100 διαφορετικές μεταβλητές σε ένα σύστημα [21]. Δεύτερον, τα μοντέλα ΜΜ παρουσιάζονται πιο επιρρεπή σε σφάλματα, όπως η υπερπροσαρμογή και η ανισορροπία κλάσεων αν δεν ελεγχθούν κατάλληλα πριν την εφαρμογή τους. Στις περισσότερες έρευνες για να αποφευχθεί το προαναφερόμενο ελάττωμα χρησιμοποιούνται εργαλεία ελέγχου σαν το PROBAST [21]. Τρίτον, η προ γνωστικότητα των μοντέλων ΜΜ δεν είναι ομοιόμορφη στην πλειοψηφία των υποομάδων των ασθενών, στοιχείο που συχνά παραβλέπετέ [21]. Ο περιορισμός αυτός επιβεβαιώνεται στην κλινική μελέτη των Desai et al.,[40] όπου παρατηρήσαμε σημαντική μείωση στη διακριτική ικανότητα των μοντέλων (C-index και ROC καμπύλης) στους γηραιότερους ασθενείς, μεγαλύτερη ή ίση με 75 ετών, συγκριτικά με νεότερους. Αυτό αιτιολογείται με βάση την ανασκόπηση των Mpanya et al. [38] στην εκπαίδευση των αλγορίθμων από δεδομένα που οι ασθενείς έχουν ηλικία 72 έτη, εγείροντας ερωτήματα για τη γενικευσιμότητα και ασφάλεια χρήσης των μοντέλων σε πρόβλεψης για νεότερης ηλικίας άτομα.

2.2.3 Το πρόβλημα του "Μαύρου Κουτιού"

Παρά την αλματώδη πρόοδο της Τεχνητής Νοημοσύνης (AI) στην ιατρική, η υιοθέτηση της στην κλινική πράξη καθυστερεί. Ο κύριος λόγος καθυστέρησης στην μετάβαση αυτή είναι η έλλειψη ερμηνευσιμότητας των μοντέλων και εμπιστοσύνης από την πλευρά του ιατρικού προσωπικού, το λεγόμενο πρόβλημα του "μαύρου κουτιού" (black box) [42]. Ως Μαύρο Κουτί στην επιστήμη των δεδομένων ορίζουμε ένα προγνωστικό μοντέλο του οποίου η μαθηματική φόρμουλα είναι είτε υπερβολικά περίπλοκη για να την κατανοήσει ο ανθρώπινος νους, όπως η αναδρομικότητα των βαθιών νευρωνικών δικτύων, είτε αποτελεί κλειστή εμπορική ιδιοκτησία, στην οποία οι χρήστες δεν έχουν πρόσβαση [22]. Στην πράξη, ο χρηστής (γιατρός στην περίπτωση μας) εισάγει τα δεδομένα ενός ασθενούς ως είσοδο σε ένα υπολογιστικό σύστημα και λαμβάνει ως έξοδο μιας απόφαση (για παράδειγμα ένα σκορ κινδύνου) χωρίς να γνωρίζει τους υπολογισμούς που πραγματοποιήθηκαν στο ενδιάμεσο. Η αδιαφάνεια αυτή εγκυμονεί σοβαρούς κινδύνους, από την εισαγωγή απλών τυπογραφικών έως και την μεροληψία στα μοντέλα, που μπορεί να οδηγήσουν σε λανθασμένα αποτελέσματα, συνεπώς λανθασμένες κλινικές αποφάσεις [22].

Η Cynthia Rudin, σε μια δημοσίευση ορόσημο , [22] υποστήριξε ότι πρέπει να σταματήσει η χρήση μοντέλων "Μαύρου Κουτιού" όταν λαμβάνονται κρίσιμες αποφάσεις που επηρεάζουν άμεσα ή έμμεσα ανθρώπινες ζωές [22]. Άντυτού, η επιστημονική και εμπορική κοινότητα πρέπει να επιδιώκει την ανάπτυξη εξ αρχής "ερμηνεύσιμων" (Interpretable) μοντέλων, ώστε η δομή τους να είναι κατανοητή γενικότερα, υπακούοντας στη λογική του εκάστοτε επιστημονικού πεδίου. Η Rudin παρουσιάζει ως επιχειρηματικό παράδειγμα την υπόθεση του Glenn Rodriguez, στην οποία το λογισμικό COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) που χρησιμοποιούν τα δικαστήρια των ΗΠΑ για την πρόβλεψη κινδύνου υποτροπής ενός προ της αποφυλάκισης κατηγορούμενου, λανθασμένα κατέταξε τον Rodriguez στη κατηγορία υψηλού κινδύνου λόγω τυπογραφικού λάθους και του αρνήθηκαν την αναστολή (parole) οι δικαστές [22].

Ως απάντηση, η Rudin και η επιστημονική ομάδα της ανέπτυξαν ένα σύστημα τεχνητής νοημοσύνης, το CORELS (Cerifiably Optimal RulE ListS), που είναι ερμηνεύσιμο εκ φύσεως του, καθώς η λειτουργία βασίζεται σε μια απλή λίστα λογικών κανόνων (if-then) που χρησιμοποιεί μόνο δύο ή τρεις μεταβλητές έναντι των 130 παραγόντων εισαγωγής του COMPAS. Μάλιστα, το CORELS σημείωσε παρόμοια επίπεδα ακρίβειας με το COMPAS και με άλλα πολύπλοκα μοντέλα μηχανικής μάθησης, με κυμαινόμενο μέσο όρο ακριβείας στο 0.66-0.67, αποδεικνύοντας πως δεν η ερμηνευσιμότητα και η ακρίβεια μπορούν να γίνουν συμβατές έννοιες [22].

Οι Bienefed et al. αναδεικνύουν την σύγκρουση μεταξύ των προγραμματιστών και των κλινικών γιατρών σχετικά με την εμπιστοσύνη στα μοντέλα μηχανικής μάθησης [43]. Οι πρώτοι προσπαθούν να ερμηνεύσουν τα μοντέλα, στηρίζοντας τις αποφάσεις τους σε μετρικές αξιολογήσεις . Από την άλλη πλευρά, οι γιατροί αναζητούν τη Κλινική Αληθοφάνεια (Clinical Plausibility), δεδομένου ότι η μη δυνατότητα κατανόησης του αλγόριθμού τους αποτρέπει από την πλήρη αποδοχή του στην κλινική καθημερινότητα [43]. Επίσης για τους προγραμματιστές η απόλυτη αλήθεια σχετικά με τη κατάσταση ενός ασθενή προκύπτει από τα δεδομένα του συστήματος ενώ οι ιατροί τονίζουν πως η απόλυτη αλήθεια προκύπτει από τον ίδιο τον ασθενή.

Συμπερασματικά, το πρόβλημα του "μαύρου κουτιού" αποτελεί πολυδιάστατο πρόβλημα που επισκιάζει την ευρεία υιοθέτηση των συστημάτων τεχνητής νοημοσύνης στον ιατρικό τομέα και απαιτεί την ανάγκη γεφύρωσης του χάσματος με την ανάπτυξη εξηγήσιμής τεχνητής νοημοσύνης (Explainable Artificial Intelligence), η οποία αναλύεται στο επόμενο κεφάλαιο [42].

2.2.4 Μέθοδοι Εξηγήσιμης Τεχνητής Νοημοσύνης

Ως Εξηγήσιμη Τεχνητή Νοημοσύνη (Explainable Artificial Intelligence ή XAI) ορίζεται το πεδίο που αναπτύσσει μεθόδους και τεχνικές για την κατανόηση, την ερμηνεία και διαφάνεια των αποτελεσμάτων που παράγουν τα μοντέλα Τεχνητής Νοημοσύνης [44]. Ο ρόλος της καθίσταται ιδιαίτερα κρίσιμος για την αποφυγή λαθών που μπορεί να οδηγήσουν σε μοιραίες συνέπειες της υγείας του ασθενή, την μείωση λαθών και προκαταλήψεων, καθώς και για την ενίσχυση εμπιστοσύνης μεταξύ των μεθόδων ΤΝ και του ιατρικού προσωπικού [44].

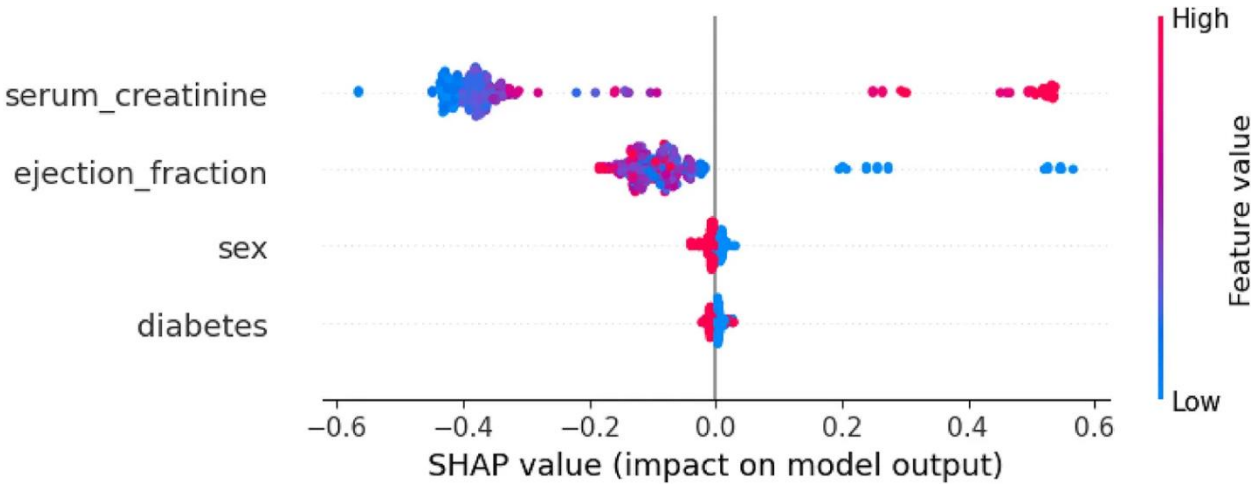
Οι Noor et al. καταγράφουν πως οι μέθοδοι της XAI γενικότερα ταξινομούνται σε πέντε κατηγορίες, του σταδίου, της εφαρμοσιμότητας, του εύρους, της μορφής και του τύπου [45]. Πρώτον, η κατηγορία του σταδίου αναφέρεται στο διαχωρισμό των μεθόδων σε αυτές που ερμηνεύονται αφού έχουν βγάλει μια απόφαση (post-hoc) και σε αυτές που το μοντέλο έχει σχεδιαστεί εγγενές ερμηνεύσιμο (ante-hoc). Ένα καλό παράδειγμα εγγενούς ερμηνεύσιμου μοντέλου είναι το TabNet [45]. Το 2022 χρησιμοποίησε τη μέθοδο TabNet για να προβλέψει τον κίνδυνο ασθενών από διαβήτη. Το μοντέλο έδινε ταυτόχρονα την πρόβλεψη αλλά και την απόδειξη της. Δεύτερον η εφαρμοσιμότητα περιγράφει αν η μέθοδος είναι γενικευμένης εφαρμογής (model-agnostic) ή αν εφαρμόζεται σε συγκεκριμένο μοντέλο (model-specific) [45]. Τρίτον, το εύρος κατηγοριοποιεί τις μέθοδούς σε ολικές (global) που εξηγούν τη συμπεριφορά του μοντέλου στο σύνολο των ασθενών ή εξηγούν τοπικά (local) μεμονωμένα περιστατικά. Τέταρτον, η μορφή που αφορά τον τρόπο παρουσίασης της εξήγησης. Αν δίνεται με ξεκάθαρες λογικές συνθήκες (rule-based) ή μέσω οπτικοποιήσεων. Πέμπτων και τελευταίο, ο τύπος της XAI μεθόδου αφορά το αν πρόκειται για λευκό κουτί (white-box), γκρίζο κουτί (grey-box), ή μαύρο κουτί (black-box) μοντέλο. Το white-box προσφέρει υψηλή επεξηγησιμότητα και το black-box χαμηλή, ενώ το grey-box να βρίσκεται ενδιάμεσα [45].

Σύμφωνα με τις ανασκοπήσεις των Chaddad et al. και Noor et al.,[44][45] οι τέσσερείς κύριες post-hoc μέθοδοι στην υγειονομική περίθαλψη είναι η SHAP, η LIME, η Grad-CAM και οι Μηχανισμοί Προσοχής:

1) Η SHAP (SHapley Additive exPlanations) βασίζεται στη θεωρία των παιγνίων και αναπτύχθηκε και αποδίδει δίκαια σε κάθε χαρακτηριστικό μια τιμή σημασίας [46]. Η τιμή φανερώνει την συμβολή του χαρακτηριστικού στη τελική πρόβλεψη του αποτελέσματος [46]. Θεωρείται η περισσότερο συνεπής μέθοδος, καθώς με μαθηματική θεμελίωση αποδεικνύει τη συνεισφορά της κάθε μεταβλητής στο τελικό αποτέλεσμα [46].
2) Η LIME (Local Interpretable Model-agnostic explanations) παρέχει τοπική εξήγηση για οποιονδήποτε ταξινομητή ή παλινδρομητή κοντά σε μια συγκεκριμένη περίπτωση χρησιμοποιώντας ερμηνεύσιμα μοντέλα, δημιουργώντας ένα τοπικό προσεγγιστικό μοντέλο που εξηγεί γιατί το αρχικό μοντέλο έκανε μια συγκεκριμένη πρόβλεψη για ένα μεμονωμένο δείγμα δεδομένων [44].
3) Η Grad-CAM αποτελεί μέθοδο που παράγει οπτικούς χάρτες θερμότητας (heatmaps) ανάλογα με τις κλίσεις (gradients) των νευρωνικών δικτύων, αναδεικνύοντας τις κρισιμότερες περιοχές που τα δεδομένα συνέβαλαν στην απόφαση του μοντέλου [44].
4) Οι Μηχανισμοί Προσοχής που αποσκοπούν σε κατηγορία ενσωματωμένων τεχνικών στην αρχιτεκτονική του μοντέλου που εξυπηρετούν στην εστίαση του μοντέλου σε δεδομένα με μεγαλύτερη βαρύτητα συνεισφοράς στην πρόβλεψη [44].

2.2.5 Η εφαρμογή της Εξηγήσιμης Τεχνητής Νοημοσύνης στην πρόβλεψη της Καρδιακής Ανεπάρκειας

Η αξιοποίηση των μεθόδων Εξηγήσιμης Τεχνητής Νοημοσύνης στα Συστήματα Υποστήριξης Κλινικών Αποφάσεων (Clinical Decision Support Systems ή CDSS) έχει αναδειχθεί ως κρίσιμος παράγοντας στη διαμόρφωση του μέλλοντος της ιατρικής [49]. Στην μελέτη των Moreno-Sanchez, η μέθοδος SHAP φανερώνεται ως ιδανική μέθοδο για την ανάλυση των βιοδεικτών, όπως το BNP, η κρεατινίνη, η ηλικία δίνοντας αναλυτική περιγραφή της κατάστασης του ασθενούς σε ατομικό επίπεδο. Παράλληλα, στην ερευνά του Moreno-Sanchez χρησιμοποιείται η βιβλιοθήκη SurvSHAP επιτρέπει τη δημιουργία γραφημάτων που απεικονίζουν τη συνολική εξηγησιμότητα των διάφορων χαρακτηριστικών που εξετάζονται [35].



Εικόνα 12 : Συνολική εξηγήσιμη ανάλυση επιβίωσης σε ασθενείς με καρδιακή ανεπάρκεια, με βάση τις τιμές των χαρακτηριστικών που επιλέχθηκαν με τη χρήση του SurvSHAP [35].

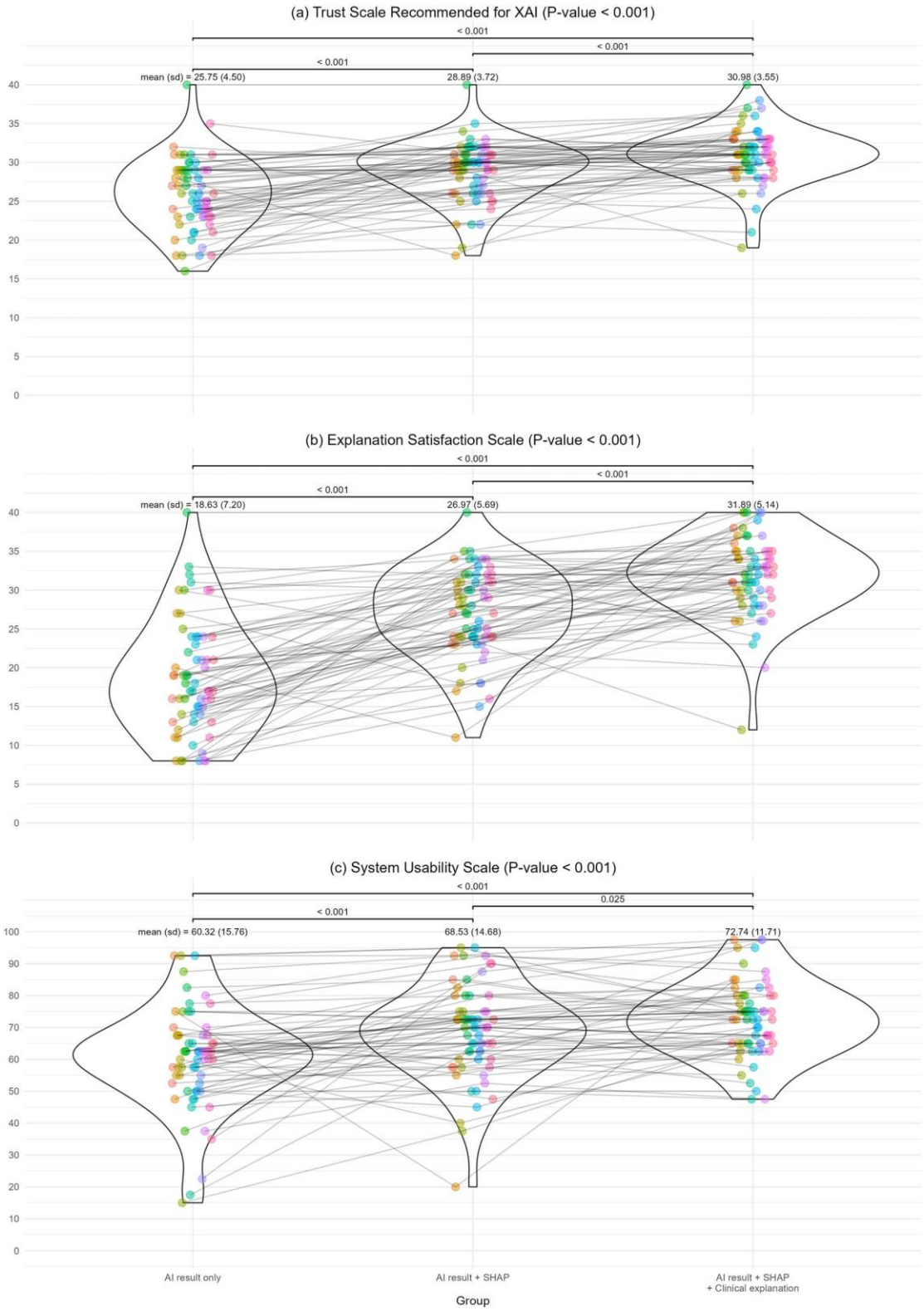
2.2.6 Η Εμπιστοσύνη των κλινικών στα συστήματα Εξηγήσιμης Τεχνητής Νοημοσύνης

Η ανάγκη για θεμελίωση εμπιστοσύνης του ιατρικού προσωπικού στα συστήματα XAI είναι μείζον θέμα για την ταχύτερη εξέλιξη της ιατρικής. Για τον λόγο αυτό, η ερευνητική κοινότητα παρουσιάζει αυξημένη έρευνα στο πεδίο αυτό. Η μελέτη των Hur et al. [47], δημοσιευμένη στο επιστημονικό περιοδικό Nature Digital Medicine, αφορούσε την ανάλυση και σύγκριση τριών κύριων μεθόδων παρουσίασης εξηγήσεων σε 63 γιατρών (N=63) που είχαν συνταγογραφήσει προϊόντα αίματος και εμπιστεύθηκαν τις αποφάσεις τους σε μια από τις τρεις μεθόδους εξήγησης CDSS (Clinical Decision Support System). Η μέθοδος SHAP πρωταγωνίστησε ως εργαλείο εξαιτίας της δυνατότητας διάκρισης των κλινικών χαρακτηριστικών, όπως τα επίπεδα NT-proBNP, κρεατίνινης και ηλικίας, σε μεμονωμένα. Οι τρείς μέθοδοι είναι οι ακόλουθες:

Η πρώτη μορφή είναι οι “Μόνο Αποτέλεσμα“ (Result Only ή RO) η οποία είναι η παρουσίαση του τελικού αποτελέσματος χωρίς αιτιολόγηση. Η δεύτερη είναι η “Αποτέλεσμα + SHAP“ (Result

+ SHAP ή RS) η οποία συνδυάζει το τελικό αποτέλεσμα με τεχνική οπτικοποίηση των παραγόντων που επηρεάζουν στη πρόβλεψη. Η τρίτη είναι η "Αποτέλεσμα + SHAP + Κλινική Εξήγηση" (Result + SHAP + Clinical Explanation ή RSC), η οποία είναι ο συνδυασμός της RS με την επιπλέον κλινική εξήγηση παρέχοντας πλήρη εξήγηση στο ιατρικό προσωπικό [47].

Ως μετρική αξιολόγησης χρησιμοποιήθηκε το Βάρος Συμβουλής (Weight of Advice ή WOA), το οποίο μα δείχνει την αλλαγή της απόφασης του γιατρού μετά τη πρόταση της μεθόδου XAI. Με βάση τα γραφήματα η μέθοδος RSC παρουσίασε την μεγαλύτερη εμπιστοσύνη των κλινικών ιατρών, βελτιώνοντας σημαντικά την αξιοπιστία με την επεξηγηματικότητα που παρέχει [47].



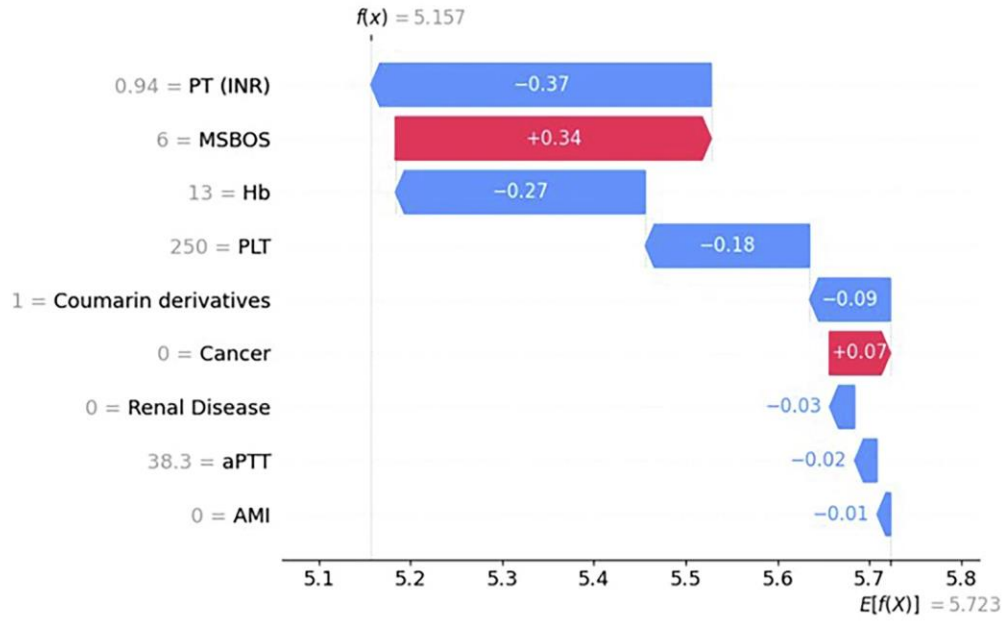
Εικόνα 13 : Συγκριτικό διάγραμμα των μετρικών αξιολόγησης ανά ομάδα. (α) Κλίμακα Εμπιστοσύνης για XAI, (β) Κλίμακα Ικανοποίησης από την Εξήγηση, (γ) Κλίμακα Ευχρηστίας Συστήματος[47].

AI Result

- **MSBOS: Open and other replacement of mitral valve [3524]** 6 units.
- **pMSBOS-TS:** 5 units.

Details

Explainable AI



Clinical explanation

This AI model predict red blood cell demand for thoracic surgery patients by considering input data. Based on the patient data inputted, the AI model predicted that **5** red blood cells would be needed (the actual value was rounded from 5.16).

- SHAP is an explainable AI, and the input that had the greatest impact were PT(INR), MSBOS, and Hb.

1. PT(INR) (Input: 0.94, Contribution relative to average: -0.37) [Reference: Min 0.9/ Max 1.1]
PT(INR) is one of the standardized methods used to measure blood clotting time. INR represents this measurement as an international standardized ratio. It is primarily used for diagnosing blood clotting disorders and when there is a deficiency or dysfunction in clotting factors, the clotting time is prolonged. In this case, the PT(INR) result was 0.94 before surgery, indicating a decrease in blood demand by 0.37 compared to the average.

2. MSBOS (Input: 6, Contribution relative to average: 0.34)
MSBOS (Maximum Surgical Blood Ordering Schedule) is used to determine the maximum amount of blood that should be ordered for a patient undergoing a specific surgery. In this case, the MSBOS for the surgery is 6, indicating an increase in blood demand by 0.34.

3. Hb (Input: 13g/dL, Contribution relative to average: -0.27) [Reference: Male 13.6-17.4 / Female 11.2-14.8]
Hemoglobin is transporting oxygen to the tissues. In this case, the preoperative Hb result was 13g/dL, indicating a decrease in blood demang by 0.27 compared to the average.

Εικόνα 14 :Παράδειγμα συστήματος Ευχρηστίας που συνδυάζει τα αποτελέσματα της Τεχνητής Νοημοσύνης, με Διαγράμματα Καταρράκτη SHAP και τη κλινική εξήγηση σε φυσική γλώσσα για τη πρόβλεψη ανάγκης και μετάγγισης αιμοσφαρίνων σε θωρακοχειρουργική επέμβαση.

Επίσης, στη συστηματική ανασκόπηση των Rosenbacke et al., επισημαίνεται η XAI αυξάνει ή μειώνει την εμπιστοσύνη των ιατρών, ανάλογα με την ποιότητα της εξήγησης που παρέχει [48]. Παρατηρήθηκε πως το ιατρικό σύνολο αντιμετωπίζει την κατάσταση με υπερβολή. Ειδικότερα , η υπεραπλούστευση που παρέχουν μερικές μέθοδοι XAI μπορεί να οδηγήσουν στην τυφλή εμπιστοσύνη (blind trust) που οι γιατροί τείνουν να αποδέχονται τις προτάσεις της Τεχνητής Νοημοσύνης χωρίς την επαρκή επαλήθευση των αποτελεσμάτων. Από την άλλη πλευρά, όταν οι εξηγήσεις γίνονται περίπλοκες, οι κλινικοί συνηθίζουν να δείχνουν δυσπιστία και να απορρίπτουν την οποιαδήποτε σύσταση παρήγαγε το μοντέλο. Η λύση φάνηκε στην η επίτευξη ενός κατάλληλου επιπέδου εμπιστοσύνης (optimal trust) που θα δημιουργεί το μοντέλο XAI [48].

Συνεπώς, η εφαρμογή των μεθόδων Εξηγήσιμης Τεχνητής Νοημοσύνης έρχεται για να γεφυρώσει το χάσμα μεταξύ των κρυφών πολύπλοκων υπολογιστικών μοντέλων και της ανάγκης των κλινικών για διαφανή και αξιόπιστα αποτελέσματα. Έτσι ξεκλειδώνεται η δυνατότητα πλήρης ενσωμάτωσης των μοντέλων ΜΜ στην κλινική πράξη και η άμεση ανάπτυξη της παραγωγικότητας στον χώρο της υγειονομικής περίθαλψης.

2.3 Η Τεχνολογία Blockchain στην Υγεία

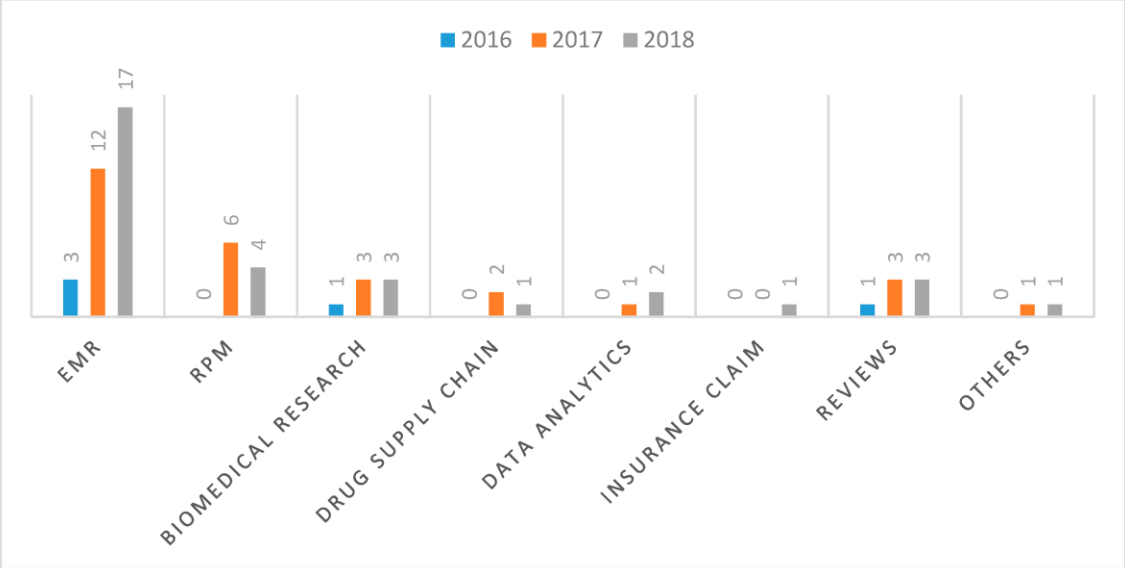
2.3.1 Γιατί να χρησιμοποιήσει κάποιος την τεχνολογία Blockchain στην Υγεία;

Η τεχνολογία Blockchain ορίζεται ως ένα αμετάβλητο και διαμοιρασμένο καθολικό μητρώο δεδομένων που καταγράφει συναλλαγές με κρυπτογραφικό τρόπο, χωρίς την ανάγκη κεντρικής αρχής ελέγχου. Η τεχνολογία αυτή αποτελεί είδος των κατανεμημένων συστημάτων και ξεχωρίζει ιδιαίτερα στον χώρο της υγειονομικής περίθαλψης λόγω των εγγενών χαρακτηριστικών της, όπως η ακεραιότητα και αμεταβλητότητα, η αποκεντρωτικότητα, η διαφάνεια και η ιχνηλασιμότητα.

Αρχικά, κάθε καταγραφή δεδομένων, αποτελεί μια συναλλαγή που κρυπτογραφείται και ενσωματώνεται σε ένα μπλοκ δεδομένων, το οποίο συνδέεται το προηγούμενο μπλοκ, δημιουργώντας μια αλυσίδα κατακερματισμού (hash chain). Με αυτόν τον τρόπο εξασφαλίζεται η ακεραιότητα και αμεταβλητότητα του δικτύου, καθώς η τροποποίηση ή διαγραφή πληροφοριών είναι αδύνατη χωρίς την έγκριση του. Επίσης, η αποθήκευση των δεδομένων γίνεται σε ένα ομότιμο δίκτυο (peer-to-peer) κόμβων, γεγονός που αποτρέπει την πρόκληση προβλημάτων του μοναδικού σημείου αποτυχίας (single point of failure). Ακόμη, η προέλευση των δεδομένων (provenance) είναι δυνατή, διασφαλίζοντας πως κάθε πρόσβαση ή μετατροπή είναι εντοπίσιμη και επαληθεύσιμη.

Σύμφωνα με ανασκόπηση πεδίου (scoping review) των Hasselgren et al., οι εφαρμογές του blockchain ως εργαλείο στο χώρο της υγείας παρουσιάζει εκθετική αύξηση [45]. Συγκεκριμένα, η διαχείριση στα Ηλεκτρονικά Αρχεία Υγείας (EHR) και τα Προσωπικά Αρχεία Υγείας (PHR)

παρουσιάζει το μεγαλύτερο ποσοστό εφαρμογής της τεχνολογίας Blockchain κατά 48%. Ο λόγος είναι η συμβατότητα που παρέχει μεταξύ διαφορετικών συστημάτων υγείας, δίνοντας σε ασθενείς να μπορούν να ελέγχουν τα δεδομένα τους μέσω των έξυπνων συμβολαίων (smart contracts) που ορίζουν τη προσβασιμότητα. Συγχρόνως, προστατεύεται η ακεραιότητα των ερευνητικών δεδομένων, προστατεύοντας τη λειτουργία των κλινικών δοκιμών. Επίσης, σε μεγάλες φαρμακευτικές εταιρίες προσφέρεται η δυνατότητα ιχνηλάτισης των προϊόντων τους, καταπολεμώντας την παραποίηση τους και προστατεύοντας την εφοδιαστική αλυσίδα της εταιρείας.



Εικόνα 15 : Ταξινόμηση των ερευνητικών δημοσιεύσεων σε χρονική εξέλιξη (2016-2019) ανά τομέα εφαρμογής της τεχνολογίας Blockchain στην υγεία [19].

Το διάγραμμα εμφανίζει 8 κύριες κατηγορίες εφαρμογής του Blockchain. Με βάση την **Εικόνα 15** παρατηρείται πως τα Ηλεκτρονικά Ιατρικά Αρχεία κυριαρχούν στην έρευνα φτάνοντας μέχρι τις 17 μελέτες το 2018. Στην 2η και 3η θέση του διαγράμματος ακολουθούν απομακρυσμένη παρακολούθηση ασθενών και διαχείριση της εφοδιαστικής αλυσίδας φαρμάκων. Ολοκληρώνοντας, οι Agbo et al. συμπεραίνουν πως οι περισσότερες από τις χρήσεις του Blockchain παραμένουν σε πειραματικό επίπεδο, υποδεικνύοντας την ανάγκη υλοποίησης περισσότερων προτύπων που είναι αποτελεσματικά σε πραγματικά κλινικά περιβάλλοντα [19].

Πίνακας 2.2: Blockchain πλατφόρμες που έχουν χρησιμοποιηθεί στην ανάπτυξη ιατρικών εφαρμογών [19].

Λειτουργικό Πλαίσιο	**Παραδείγματα Εφαρμογών**
Ethereum	MedRec, Ancile, DPS, FHIRChaine, MedShare

Hyperledger Fabric	HealthChain, MedicalChain
Proprietary	Guardtime, MedBlock, BlockHIE

2.3.2 Blockchain και Ηλεκτρονικά Ιατρικά Αρχεία

Προηγουμένως αναφέρθηκε η έντονη παρουσία που εμφανίζει η τεχνολογία Blockchain στα Ηλεκτρονικά Αρχεία Υγείας (EHR) και τα Προσωπικά Αρχεία Υγείας (PHR). Στη συστηματική ανασκόπηση των Durneva et al. καταδεικνύεται η αύξηση του ενδιαφέροντος για την χρήση του Blockchain στην αντιμετώπιση προβλημάτων στην Πληροφορική της Υγείας (HIT), με το 51% να επικεντρώνεται στη διαχείριση των ιατρικών πληροφοριών, ενώ το 53% σε ζητήματα ασφαλείας και ιδιωτικότητας [46]. Παρομοίως οι Ettaloui et al. στηρίζουν την άποψη πως το Blockchain επιλύει αποτελεσματικά ζητήματα ασφαλείας, ιδιωτικότητα και διαθεσιμότητα που αντιμετωπίζουν τα παραδοσιακά συστήματα EHR [47].

Σημαντικό παρουσιάζεται το δίλημμα αφορά την αποθήκευση των δεδομένων on-chain και off-chain. Για παράδειγμα, η αλυσίδα του Bitcoin διαθέτει περιορισμένο μέγεθος μπλοκ (1 MB) και χρειάζεται 10 λεπτά η εξόρυξη νέου μπλοκ, καθιστώντας τη πλήρη αποθήκευση ιατρικών πληροφοριών μη πρακτική. Αντ' αυτού, για να αντιμετωπιστεί το ζήτημα αυτό οι Zheng et al. προτείνουν μια υβριδική προσέγγιση [17]. Η καταγραφή μόνο των δεικτών, των αποτυπωμάτων και σημαντικών μεταδεδομένων να γίνεται On-chain, ενώ η αποθήκευση μεγάλων αρχείων δεδομένων να γίνεται σε εξωτερικές βάσεις, για παράδειγμα στο σύννεφο ή IPFS, ενώ το Blockchain να λειτουργεί μόνο για την επαλήθευση ακεραιότητας των δεδομένων. Η προσέγγιση αυτή επιτρέπει τον υψηλότερο όγκο συναλλαγών και την μείωση του συνολικού μεγέθους του Blockchain. Οι συστηματικές ανασκοπήσεις των Durneva et al. και του Ettaloui et al., αναφέρουν τις αλυσίδες Ethereum και Hyperledger Fabric ως τις πιο ιδανικές και δημοφιλής πλατφόρμες καταγραφής Ηλεκτρονικών Ιατρικών Αρχείων [46][47]. Στην έρευνα τους, το Hyperledger Fabric εμφανίζεται περισσότερο κατάλληλη λόγω της βελτιστοποιήσεις σε απόδοση και έλεγχο ταυτότητας που προσφέρει σε σύγκριση με το Ethereum που χρησιμοποιεί χρονοβόρους μηχανισμούς συναίνεσης όπως το Proof-of-Work.

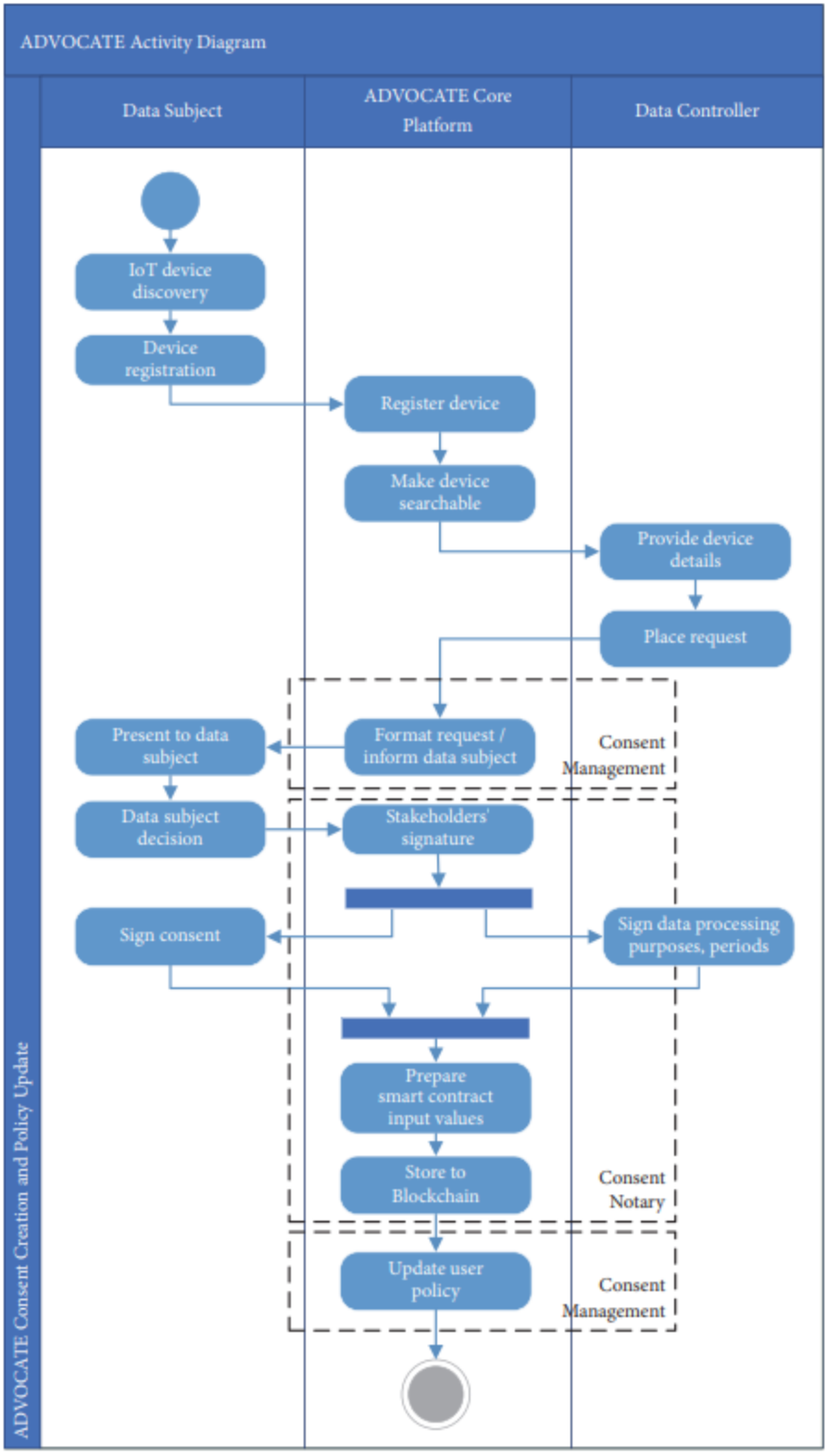
Το κυριότερο πρόβλημα, με βάση την υπάρχουσα βιβλιογραφία, αναγνωρίζεται πως είναι η κλιμακωσιμότητα (scalability). Το Blockchain δεν έχει την δυνατότητα επεξεργασίας εκατομμύριων συναλλαγών σε πραγματικό χρόνο. Η λύση στην πρόκληση αυτή παρουσιάζεται στην πρόταση δημιουργίας μηχανισμών βελτιστοποίησης αποθήκευσης που καταργούν παλαιές εγγραφές και στον ανασχεδιασμό της αρχιτεκτονικής με την χρήση πρωτοκόλλων όπως το Bitcoin-NG που βελτιστοποιούν σημαντικά την διαδικασία.

2.3.3 Η συμβολή των έξυπνων συμβολαίων στο χώρο της υγείας

Ως γνωστόν, τα έξυπνα συμβόλαια αποτελούν αυτοεκτελούμενα ψηφιακά πρωτόκολλα αποθήκευσης δεδομένων στο Blockchain δικτύο και ενεργοποιούνται αυτομάτως όταν

πληρούνται οι όροι του συμβολαίου. Στον υγειονομικό τομέα λειτουργούν ως μηχανισμοί αυτοματοποίησης της επιχειρηματικής λογικής (business logic) της συναίνεσης, διασφαλίζοντας την αξιοπιστία και διαφάνεια των υπηρεσιών χωρίς τη μεσολάβηση τρίτων [45][53].

Ένα χαρακτηριστικό παράδειγμα αυτόματης διαχείρισης συναινέσεων παρουσιάζεται το πλαίσιο ADvoCATE, που χρησιμοποιεί το blockchain για την προστασία της ακεραιότητας και την έκδοση (versioning), εξυπηρετώντας τους ασθενείς στην διατήρηση του ελέγχου των δεδομένων τους [53]. Το πλαίσιο επιτρέπει στους ασθενείς να κατέχουν τον έλεγχο του δικαιώματος πρόσβασης στα δεδομένα τους, εξασφαλίζοντας ότι η πρόσβαση από τους φορείς υγείας γίνεται με την συγκατάθεση των ασθενών [45]. Σε άλλη έρευνα το δικαίωμα πρόσβασης γίνεται μέσω των έξυπνων συμβολαίων που χρησιμοποιούν το μοντέλο RBAC, όπου το δικαίωμα πρόσβασης ανατίθεται με βάση τον λειτουργικό ρόλο του χρήστη. Παραδείγματος χάριν, σε μια νοσοκομειακή μονάδα υπάρχουν διάφοροι λειτουργικοί ρόλοι όπως ο ιατρός, ο ερευνητής, ο νοσηλευτής και άλλοι [45][53].



Εικόνα 16 : Διάγραμμα ροής λειτουργίας της πλατφόρμας ADvoCATE [53].

Ακόμη, τα έξυπνα συμβόλαια έχουν τη δυνατότητα εκπομπής συμβάντων (events). Συστήματα όπως το AuditChain χρησιμοποιούν τα συμβόλαια για να δημιουργεί λεπτομερείς αναφορές που περιλαμβάνουν χρονοσήμανση, τη ταυτότητα του χρήστη και το αποτέλεσμα στη πρόσβαση, δίνοντας το πλήρες ελεγκτικό ίχνος (audit trail).

2.3.4 Blockchain και GDPR: Η σύγκρουση της αμεταβλητότητας και του δικαιώματος διαγραφής

Η ιδιότητα της αμεταβλητότητας (immutability) που παρουσιάζει το blockchain προκαλεί νομική σύγκρουση με τις απαιτήσεις του Γενικού Κανονισμού Προστασίας Δεδομένων (GDPR) [19]. Η κύρια ένταση εντοπίζεται στο Άρθρο 17 που κατοχυρώνει το "δικαίωμα στη λήθη" (right to be forgotten). Πιο συγκεκριμένα, η αλυσίδα του Blockchain δεν δέχεται τροποποιήσεις ή διαγραφές καταγεγραμμένων εγγραφών. Έτσι διασφαλίζεται η ακεραιότητα των ιατρικών αρχείων και γίνεται αντιπαραγωγική στην ενδεχόμενη ολική διαγραφή του ιατρικού ιστορικού ενός ασθενή.

Το νομικό και τεχνολογικό χάσμα έρχεται να γεφυρώσει η έρευνα των Agbo et al. με την πρόταση των τροποιήσιμων blockchain (redactable blockchains) [19]. Η βασική λύση που προτείνουν είναι η ανάπτυξη δικτύων που προσαρμόζονται και υποστηρίζουν την ελεγχόμενη επεξεργασία. Αυτό επιτυγχάνεται με την ανάλογη αρχιτεκτονική του έξυπνου συμβολαίου, το οποίο θα επιτρέπει την τροποποίηση των δεδομένων και παράλληλα την διατήρηση κάθε λεπτομέρειας σε ίχνη ελέγχου (audit trails) [19].

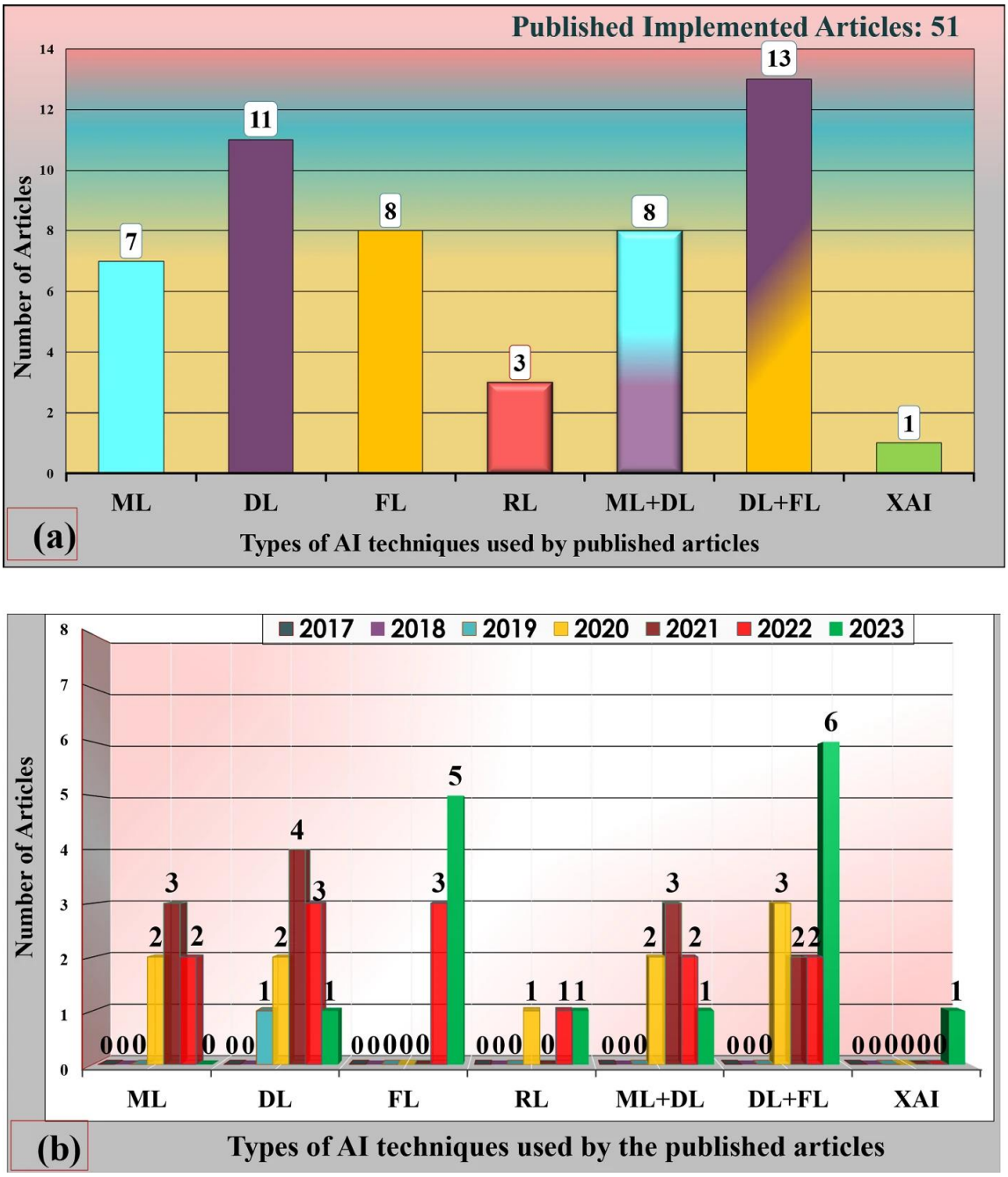
Η έρευνα των Ettaloui et al. τονίζει την σημασία υιοθέτησης ενός υβριδικού μοντέλου που συνδυάζει on-chain και off-chain αποθήκευση ελαχιστοποιώντας το κόστος καυσίμων (gas fees) και κλιμακώνοντας την επεκτασιμότητα, με συμμόρφωση στις απαιτήσεις του GDPR [47]. Στην ανασκόπηση με τίτλο "*A Systematic Literature Review for Blockchain-Based Healthcare Implementations*" των Shaikh et al. προβάλλεται το European Blockchain Sandbox που είναι ένα υποσχόμενο και ελεγχόμενο περιβάλλον δοκιμής καινοτόμων λύσεων της τεχνολογίας Blockchain σε νεοφυείς επιχειρήσεις (startups), μεγάλες πολυεθνικές ή ρυθμιστικές αρχές [48]. Η ιδιαιτερότητα αυτού του περιβάλλοντος είναι η δημιουργία συστημάτων που συμμορφώνονται πλήρως με το GDPR και το eIDAS (Ευρωπαϊκός κανονισμός ηλεκτρονικής ταυτοποίησης).

Η υλοποίηση έξυπνων συμβολαίων που βασίζονται την προσάρτηση των προηγούμενων προτάσεων της ερευνητικής βιβλιογραφίας κρίνεται αναγκαία, καθώς θα συμβάλλει στην επίλυση ενός καίριου για την εξέλιξη της τεχνολογίας Blockchain τόσο στον ιατρικό τομέα όσο και στη βιομηχανία γενικότερα.

2.3.5 Περιορισμοί της τεχνολογίας Blockchain και η σύνδεση της με τα υπάρχουσα CDSS

Πέραν του αυξανόμενου ερευνητικού ενδιαφέροντος, η εφαρμογή της τεχνολογίας Blockchain στον τομέα της υγειονομικής περίθαλψης παραμένει κατά κύριο λόγο σε επίπεδο Proof-of-Concept (PoC). Η πρόσφατη ανασκόπηση των Shaikh et al. ακολουθεί ως κατευθυντήρια γραμμή το αυστηρό μεθοδολογικό πλαίσιο PRISMA-2020 και αναφέρει πως τα πρότυπα αρχικού σταδίου της υλοποίησης Blockchain είναι ιδανικά για την διαμόρφωση νέων ιδεών και δοκιμή υποθέσεων, αλλά δεν προσφέρουν ολοκληρωμένη εικόνα της τεχνολογίας σε εφαρμογή στον πραγματικό κόσμο [48].

Ιδιαίτερα σπανίως παρουσιάζεται ο συνδυασμός της τεχνολογίας Blockchain και της τεχνητής νοημοσύνης στα υγειονομικά συστήματα. Στην αφηγηματική ανασκόπηση των Bathula et al. παρατηρείται ότι ανάμεσα σε 402 επιστημονικών άρθρων, τα 79 έχουν ως θέμα την τεχνητή νοημοσύνη και το Blockchain στην υγεία, από τα οποία τα 51 αφορούν υλοποιήσιμα έργα [49]. Ωστόσο, παρατηρώντας την **Εικόνα 17** διαπιστώνεται ότι 1 από τις 51 υλοποιήσιμες μελέτες ενσωματώνουν την Εξηγήσιμη Τεχνητή Νοημοσύνη (XAI). Αυτό αναδεικνύει πως ενώ κάθε τεχνολογία (MM, Blockchain, XAI) έχουν εξεταστεί μεμονωμένα, ο συνδυασμός στους σε ένα ενιαίο κλινικό σύστημα παραμένει σπάνιος [49].



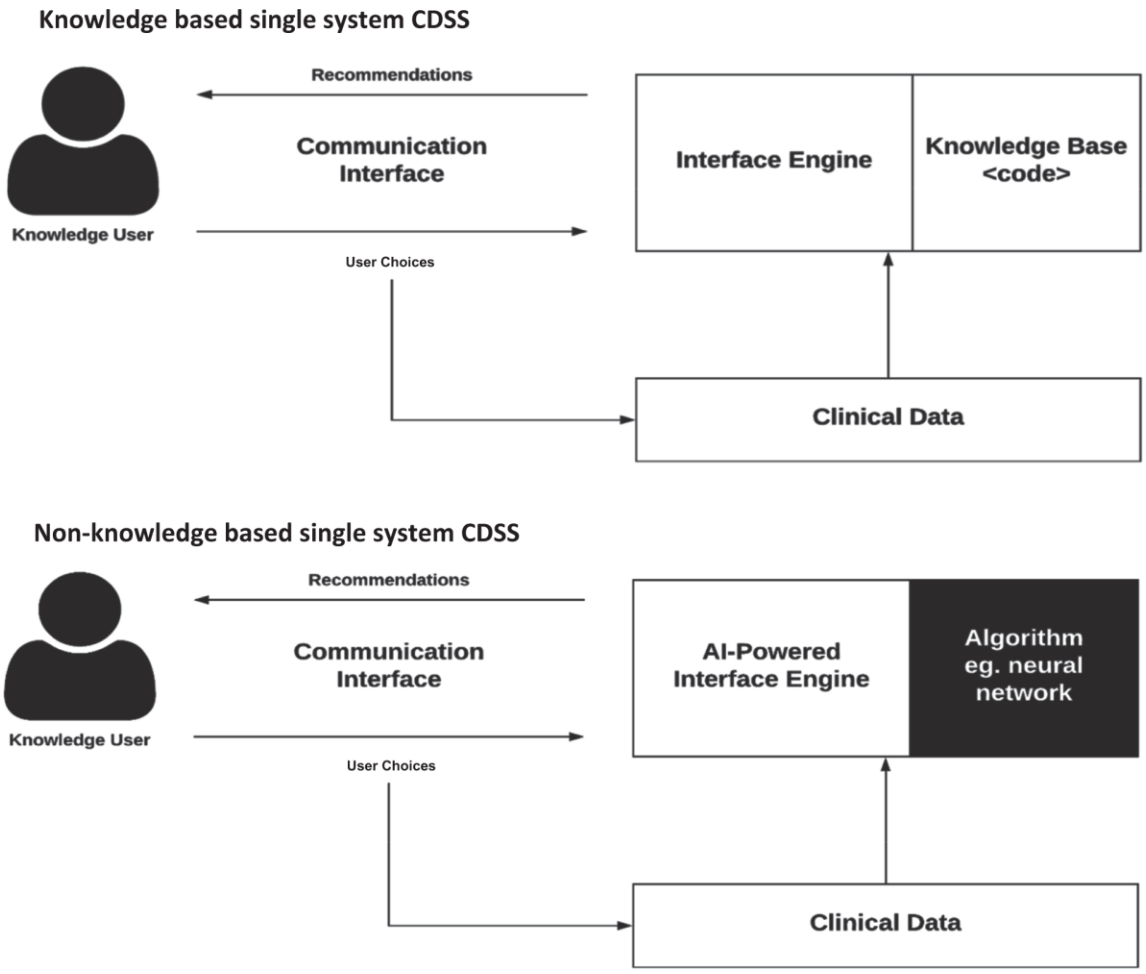
Εικόνα 17 : a) Ο αριθμός άρθρων που χρησιμοποιούν τεχνικές Τεχνητής Νοημοσύνης στα δημοσιευμένα άρθρα. b) Η κατανομή των τεχνικών Τεχνητής Νοημοσύνης που χρησιμοποιούνται στα άρθρα με θέμα «Ενσωμάτωση τεχνητής νοημοσύνης και αλυσίδας μπλοκ στον τομέα της υγειονομικής περίθαλψης» (2017–2023). AI: Τεχνητή Νοημοσύνη, ML: Μηχανική μάθηση, DL: Βαθιά μάθηση, FL: Ομοσπονδιακή μάθηση, RL: Ενισχυτική μάθηση, XAI: Εξηγήσιμη τεχνητή νοημοσύνη [56].

2.4 Συστήματα Υποστήριξης Κλινικών Αποφάσεων (CDSS)

2.4.1 Ορισμός και Ταξινόμηση

Ένα Σύστημα Υποστήριξης Κλινικών Αποφάσεων (CDSS) έχει ως στόχο τη βελτίωση των παροχών υγειονομικής περίθαλψης, βελτιώνοντας την ιατρική γνωμάτευση παρέχοντας εξατομικευμένη κλινική γνώση. Σύμφωνα με τον διεξοδικό απολογισμό και έρευνα των Sutton et al., που έχει δημοσιευθεί στο Digital Medicine, τα Συστήματα Υποστήριξης Κλινικών Αποφάσεων ταξινομούνται σε δύο θεμελιώδεις κατηγορίες: τα συστήματα βασισμένα στη γνώση (Knowledge based) και τα μη βασισμένα σε γνώση (non-knowledge-based) [50].

Τα συστήματα βασισμένα στη γνώση λειτουργούν υπό την αρχή της λογικής εξαγωγής συμπερασμάτων, αξιοποιώντας συγκεκριμένες λογικούς κανόνες (εντολές IF-THEN). Τα συστήματα αυτού του είδους ανακτούν τα απαραίτητα δεδομένα, τα περνάνε από αξιολόγηση των κανόνων και συνεχίζουν παράγοντας μια ενέργεια. Η δημιουργία των κανόνων εξαρτάται από τη βιβλιογραφία ή το κλινικό ιστορικό του ασθενή. Αφετέρου, στα Συστήματα Υποστήριξης Κλινικών Αποφάσεων που δεν στηρίζονται στη γνώση, την θέση των λογικών κανόνων παίρνουν οι αλγόριθμοι μηχανικής μάθησης, της τεχνητής νοημοσύνης και τα στατιστικά πρότυπα αναγνώρισης. Αυτό καθιστά τα CDSS που δεν βασίζονται στη γνώση γεμάτα προκλήσεις, όπως η κατανόηση της λογικής πίσω από τα μοντέλα Τεχνητής Νοημοσύνης, ειδικά στην εξαγωγή των αποτελεσμάτων, το γνωστό πρόβλημα των “μαύρων κουτιών” και στη διαθεσιμότητα των δεδομένων. Παρά τις προκλήσεις, τα συστήματα που δεν είναι βασισμένα σε γνώσεις αποτελούν το μέλλον των CDSS και είναι η βάση της παρούσας διπλωματικής εργασίας [50].



Εικόνα 18 : Αναπαράσταση των Συστημάτων Υποστήριξης Κλινικών Αποφάσεων με γνώμονα τη γνώση. [50]

2.4.2 Κλινική Αποτελεσματικότητα των CDSS

Η αξία των Συστημάτων Υποστήριξης Κλινικών Αποφάσεων (CDSS) τεκμηριώνεται μέσα από πολλές συστηματικές ανασκοπήσεις και μετα-αναλύσεις, με τελικό συμπέρασμα τους την επείγουσα ενσωμάτωση τους στην κλινική καθημερινότητα. Χαρακτηριστική είναι η δυναμική αύξηση των ερευνών στο τομέα των κλινικών συστημάτων αποφάσεων, καθώς μεταξύ του 2011 και 2021 οι δημοσιεύσεις διπλασιάστηκαν.

Η λεπτομερής ανάλυση των Chen et al. που δημοσιεύτηκε στο επιστημονικό περιοδικό “Journal of the American Medical Informatics Association (JAMIA)“, επιβεβαιώνει τη συμβολή των CDSS στην βελτίωση της ποιότητας των υπηρεσιών υγείας [51]. Στην συγκεκριμένη έρευνα αξιολογούνται 76 επιμέρους μελέτες από τις οποίες 32 είναι τυχαιοποιημένες ελεγχόμενες δοκιμές και 44 είναι σε θεωρητικό επίπεδο, που επικεντρώνονται σε CDSS τα οποία έχουν ως πηγή δεδομένων τους τους Ηλεκτρονικούς Φακέλους Υγείας (EHR). Το 70% των μελετών εμφάνιζε την εφαρμογή των CDSS σε δομές πρωτοβάθμιας περίθαλψης, με έμφαση στη διαχείριση χρόνιων

νοσημάτων με καρδιακούς παράγοντες κινδύνου (25%), τον σακχαρώδη διαβήτη (18%) και τη χρόνια νεφρική νόσο (13%) [51].

Αναφορικά με την λειτουργικότητα τους, η ερευνητική "πλάστιγγα" έγειρε υπέρ της χρήσης των CDSS έναντι της τυπικής φροντίδας (ομάδων ελέγχου). Ο λόγος που συνέβη αυτό είναι η παροχή άμεσων διαδικασιών φροντίδας (process-of-care), ειδικά στη φαρμακολογική διαχείριση που το σύστημα παρέχει σημαντικές κλινικές κατευθυντήριες οδηγίες στους χρήστες του, αυξάνοντας την ορθότητα της συνταγογράφησης και τα ποσοστά προληπτικού ελέγχου. Απεναντίας, στα μακροπρόθεσμα τελικά αποτελέσματα, όπως η θνησιμότητα, τα CDSS εμφανίζουν ασαφή επίδραση. Αυτό οφείλεται στο ότι το 66% των μελετών αφορούσε αποτελέσματα με χρονικό ορίζοντα έως 12 μήνες, ενώ τα χρόνια νοσήματα επηρεάζουν τους ασθενείς σε βάθος χρόνου [51].

Επομένως, ενώ η χρήση των CDSS σε τακτικά περιστατικά ή βραχείες χρονικές περιόδους παρουσιάζεται λειτουργική και αποδοτική, κρίνεται απαραίτητη η χρηματοδότηση μελετών υψηλής ποιότητας, που θα παρακολουθούν σε μεγάλο χρονικό ορίζοντα τους ασθενείς με ουσιαστικούς δείκτες αξιολόγησης (meaning evaluation measures).

2.4.3 Εμπόδια υιοθέτησης των CDSS στην κλινική καθημερινότητα

Πάρα τα τεκμηριωμένα πλεονεκτήματα της υιοθέτησης των CDSS, η εφαρμογή τους στη κλινική πράξη παρουσιάζει κρίσιμα εμπόδια. Σύμφωνα με την ανασκόπηση των Meunier et al., που είναι δημοσιευμένη στο περιοδικό Annals of Family Medicine, συμπεριέλαβε συνολικά 48 μελέτες που αξιολογούσαν 45 CDSS και διαπιστώθηκαν 186 διακριτοί ανασταλτικοί και διευκολυντικοί παράγοντες [61]. Λόγω της πληθώρας τους και για την αποτελεσματική κατηγοριοποίηση αυτών των στοιχείων, οι ερευνητές επωφελήθηκαν από το πλαίσιο HOT-fit, που διαχωρίζει τους αλληλεξαρτώμενους παράγοντες που επηρεάζουν την ενσωμάτωση των πληροφοριακών συστημάτων υγείας. Οι τέσσερις παράγοντες που στηρίζεται το πλαίσιο είναι ο Ανθρώπινος (Human), Οργανωτικός (Organization), ο Τεχνολογικός (Technology), τα Καθαρά Οφέλη (Net Benefit) [61].

Πρώτα απ' όλα, ο ανθρώπινος παράγοντας περιγράφει τη χρήση του συστήματος από επαγγελματίες (system use) και την ικανοποίηση των χρηστών (user satisfaction). Επίσης, ο οργανωτικός παράγοντας (organization) σχετίζεται με το ευρύτερο εργασιακό περιβάλλον και τη δομή της κλινικής. Επιπλέον, ο τεχνολογικός παράγοντας αφορά την συνολική ποιότητα του συστήματος (system quality), δηλαδή την ποιότητα των δεδομένων που παρέχει (information quality) και τη ποιότητα των υπηρεσιών υποστήριξης (service quality). Τελειώνοντας, τα καθαρά οφέλη αναφέρονται στη διαφορά των θετικών και αρνητικών επιπτώσεων των συστημάτων CDSS στον πρωτογενή υγειονομικών τομέα (PCP) [61].

Το τελικό συμπέρασμα της έρευνας ανέφεραν πως τα περισσότερα εμπόδια εμφανίζονται ως ανθρώπινα, οργανωτικά για παράδειγμα διαταραχή της συνήθους ροής εργασίας και τεχνολογικά όπως η φιλικότητα προς τον χρήση των CDSS. Το μεγαλύτερο από τα εμπόδια είναι ο αυξημένος

φόρτος εργασίας. Παράλληλα ο τεχνολογικός παράγοντας είχε ουδέτερο αντίκτυπο και η διάσταση των καθαρών οφελών θετικό.

Πίνακας 2.3: Σύγκριση μέσης επίδρασης των κατηγοριών HOT-fit στα συστήματα υποστήριξης λήψης κλινικών αποφάσεων (CDSS) από τους παθολόγους [61].

	Άνθρωπος	**Οργανισμός**	**Τεχνολογία**	**Καθαρά Οφέλη**
Μέση επίδραση των κατηγοριών HOT-fit στα συστήματα υποστήριξης κλινικών αποφάσεων (CDSS)	Ελαφρώς αρνητικη	Ελαφρώς αρνητική	Ουδέτερη	Θετική
Μέση διαφορά μεταξύ εμποδίων και παραγόντων διευκόλυνσης (95% CI)	-1.5 (-2.2 έως -0.8)	-1.9 (-2.6 έως -1.1)	-0.5 (-1.5 έως 0.5)	+3.1 (2.2 έως 3.9)

Φλέγον ζήτημα αποτελεί η κόπωση από ειδοποιήσεις (alert fatigue). Οι Sutton et al. επισημαίνουν πως έως και το 95% των ειδοποιήσεων στα CDSS αποτελούν συνήθως ασήμαντες (incosequential), γεγονός που οδηγεί το ιατρικό προσωπικό στην δυσπιστία ή αδιαφορία τους, υπονομεύοντας τη λειτουργία του συστήματος. Οι συγκεκριμένοι παράγοντες καθιστούν σαφείς την ανάγκη για σχεδίαση πιο διαφανών και φιλικών κλινικών διεπαφών. Μια εξαιρετική λύση σ ’αυτό παρουσιάζονται οι οπτικοποιήσεις που παρέχουν τα μοντέλα εξηγήσιμης τεχνητής νοημοσύνης SHAP [50].

2.4.4 Συστήματα CDSS βασισμένα σε Τεχνητή Νοημοσύνη

Η μετάβαση των Συστημάτων Υποστήριξης Κλινικών Αποφάσεων από τα παραδοσιακά συστήματα των λογικών κανόνων στα συστήματα με μη βασισμένη γνώση κατατάσσεται ως μια από τις πιο αναπτυσσόμενες ερευνητικές τάσεις στο τομέα της βιοπληροφορικής. Η σκοπευτική ανασκόπηση των Susanto et al., δημοσιευμένη στο περιοδικό ιατρικής JAMIA (Journal of the American Medical Informatics Association), από τις 1255 έρευνες που δημοσιεύθηκαν από τον Ιανουάριο του 2016 έως τον Απρίλιο του 2021, οι 32 παρουσίαζαν αξιολογημένα κλινικά συστήματα CDSS βασισμένα στην Τεχνητή Νοημοσύνη. Η ανασκόπηση κατηγοριοποιούσε την αναγνώριση και ερμηνεία εικόνας (n=12.38%), την πρόβλεψη κινδύνου (n=9.28%), τη σήψη (n=5%) [62].

Η ανασκόπηση μας έδειξε ότι καμία από τα 32 συστήματα δεν διέθετε την λειτουργία αυτόνομης λήψης αποφάσεων, όλα τα συστήματα λειτουργούσαν ως εργαλεία υποστήριξης, με το γιατρό να διατηρεί τον πλήρη έλεγχο. Επίσης, η ανασκόπηση υπογράμμισε την περιορισμένη τεκμηρίωση σε σχέση με την ακρίβεια στις προβλέψεις που παρείχαν οι αλγόριθμοι μηχανικής μάθησης [62]. Ακόμη, η δυσκολία ένταξης των μοντέλων ΜΜ σε υπάρχουσες ροές κλινικής εργασίας (workflow) αποτελεί ζήτημα στην ορθή λειτουργία τους, καθώς δημιουργούνται κατακερματισμένες ροές εργασίας που ενώ αυξάνουν των όγκο σε γνώσεις, μειώνουν τη χρονική διάρκεια επαφής με τον ασθενή [62].

Τα παραπάνω συμπεράσματα υποδεικνύουν τη κρισιμότητα υλοποίησης ενός κλινικού συστήματος που περιλαμβάνει στην αρχιτεκτονική του μεθόδους εξηγήσιμης τεχνητής νοημοσύνης για διαφάνεια, την τεχνολογία blockchain για ακεραιότητα και σύγχρονα πλαίσια διεπαφής χρήστη για την ενημέρωση σε πραγματικό χρόνο.

2.4.5 Πρότυπα διαλειτουργικότητας σε Ηλεκτρονικούς Φακέλους Υγείας

Η δια λειτουργικότητα ανάμεσα στα υπάρχοντα συστήματα πληροφορικής στην υγεία και τα αναδυόμενα CDSS είναι σημαντική για την αποτελεσματική λειτουργία και των δύο. Σύμφωνα με τους Yoo et al., η βέλτιστη δια λειτουργικότητα πραγματοποιείται με τη χρήση κοινών μοντέλων δεδομένων, όπως το OMOP CDM, που χαρτογραφούν τα δεδομένα των Ηλεκτρονικών Φακέλων Υγείας σε ένα μοντέλο [54]. Μια άλλη λύση που πρότειναν οι συγκεκριμένοι ερευνητές είναι η κωδικοποίηση δεδομένων και η θεσμοθέτηση αναγνωρισμένων ορολογιών. Μερικά παραδείγματα είναι το SNOMED CT που χρησιμοποιείται για κλινικά ευρήματα, το LOINC για εργαστηριακές εξετάσεις, το RxNorm για τη ταξινόμηση φαρμάκων και η πιο γνωστή η ICD-10 που αφορά την ταξινόμηση των νόσων. Επίσης απαραίτητη είναι η χρήση του προτύπου HL7 FHIR που λειτουργεί ως κοινή ψηφιακή γλώσσα και επιτρέπει σε διαφορετικά ιατρικά συστήματα να επικοινωνούν μεταξύ τους [54].

Επομένως, η υιοθέτηση καθιερωμένων προτύπων διαλειτουργικότητας και τυποποιημένων ορολογιών είναι θεμελιώδης για να λειτουργούν οι CDSS ανεξάρτητα, χωρίς χρονικές καθυστερήσεις και με μειωμένο κόστος.

2.5 Υπάρχουσα Κατάσταση

Η ανασκόπηση του Κεφαλαίου 2 ανέδειξε την έντονη ερευνητική δραστηριότητα που υπάρχει γύρω από τις τεχνολογίες της Μηχανικής Μάθησης, της Ερμηνεύσιμης Τεχνητής Νοημοσύνης, του Blockchain και των Συστημάτων Υποστήριξης Κλινικών Αποφάσεων στον ιατρικό τομέα. Ωστόσο, ο ολοκληρωμένος συνδυασμός τους σε ένα ενιαίο πλαίσιο παραμένει ελάχιστα διερευνημένος. Στην παρούσα ενότητα πραγματοποιείται η σύγκριση της παρούσας διπλωματικής εργασίας με άλλες συναφείς εργασίες και εντοπίζει τα κύρια ερευνητικά κενά.

2.5.1 Συναφείς Εργασίες

Στον κλάδο της μηχανικής μάθησης και ερμηνευσιμότητας για πρόβλεψη της καρδιακής ανεπάρκειας στην έρευνα των Moreno-Sanchez et al εφαρμόστηκαν SHAP μοντέλα, συγκεκριμένα το Extra Trees, για τη πρόβλεψη επιβίωσης 299 ασθενών από καρδιακή ανεπάρκεια [27]. Η έρευνα παρείχε τοπική και καθολική ερμηνεία μέσω Beeswarm και Waterfall διαγραμμάτων σε μεταβλητές όπως ο χρόνος παρακολούθησης, η κρεατινίνη ορού και το κλάσμα

εξώθησης, χωρίς να αξιοποιείται κάποιο αποθηκευτικό μέσο για ακεραιότητα των δεδομένων [49]. Η ανασκόπηση των Bathula et al. εξέτασε τη σύγκλιση της τεχνητής νοημοσύνης και του Blockchain από τις 51 υλοποιήσιμες μελέτες που εντοπίστηκαν, μόλις 1 ενσωμάτωνε ερμηνεύσιμη τεχνητή νοημοσύνη, δείχνοντας το πρόβλημα που υπάρχει στη κλινική αδιαφάνεια [55]. Η μελέτη των Otoum et al, πρότεινε ένα πλαίσιο Federated Learning που συνδυάζει το Blockchain με τη ασφαλή εκπαίδευση μοντέλων μηχανικής μάθησης σε κατανεμημένα κλινικά δεδομένα. Ενώ το σύστημα που προτάθηκε παρέχει ιδιωτικότητα δεδομένων μέσω αποκεντρωμένης εκπαίδευσης, δεν παρέχει ερμηνεύσιμες εξηγήσεις. Τέλος, οι Muneer et al προγραμμάτισαν ένα chatbot βασισμένο σε μηχανική μάθηση για την πρόβλεψη καρδιαγγειακού κίνδυνού, με γραφική διεπαφή χρήστη [56].

2.5.2 Ερευνητικά Κενά

Η σύνθεση της ερευνητικής βιβλιογραφίας προβάλλει σημαντικά ερευνητικά κενά. Καταρχήν, η πλειοψηφία των μοντέλων μηχανικής μάθησης δεν ενσωματώνει μεθόδους ερμηνευσιμότητας. Η μετα-ανάλυση των Abbas et al. κατέδειξε πως η πλειοψηφία των ερευνών εξηγήσιμης τεχνητής νοημοσύνης σε κλινικά συστήματα δεν εμφανίζει πιστότητα (fidelity), εμπιστοσύνη (trust) ούτε χρησιμότητα (usability) [42]. Επίσης, το κλινικό προσωπικό έδειξε πως προτιμάει τις εξηγήσεις των κλινικών καταστάσεων σε φυσική γλώσσα από τα αποτελέσματα τεχνητής νοημοσύνης για λόγους εμπιστευτικότητας και διαφάνειας. Επιπλέον, οι περισσότερες υλοποιήσεις στην υγεία παραμένουν σε proof-of-concept και τα κλινικά συστήματα υποστήριξης αποφάσεων δεν εξασφαλίζουν αμεταβλητότητα στα αποτελέσματα ούτε ελεγκτικό ίχνος.

2.5.3 Συγκριτικός Πίνακας

Ο Πίνακας 2.4 συνοψίζει τα χαρακτηριστικά των εργασιών σε ένα κεντρικό παραλληλισμό για να γίνουν περισσότερο κατανοητά τα ερευνητικά κενά που αναφέραμε προηγουμένως. Πρώτά απ' όλα, ο συνδυασμός της μηχανικής μάθησης, της εξηγήσιμης τεχνητής νοημοσύνης και του Blockchain είναι μοναδικός. Οι Moreno-Sanchez δεν έχουν blockchain ενώ οι Bathula et al. (2024) και Otoum et al. (2023) έχουν μηχανική μάθηση και blockchain αλλά χωρίς εξηγησιμότητα. Επίσης, καμία δημοσίευση δεν διαθέτει κλινική διεπαφή, εκτός του Muneer et al., παραμένοντας σε επίπεδο proof-of-concept. Καμία απολύτως δημοσίευση δεν παρέχει μηχανισμό ασφαλείας προσωπικών δεδομένων με την συναίνεση του ασθενή, κάτι που η προτεινόμενη πλατφόρμα το παρουσιάζει. Ακόμη, καμία επιστημονική δημοσίευση δεν χρησιμοποιεί μεγάλο κλινικό σύνολο δεδομένων, αμφισβητώντας με αυτόν τον τρόπο τη γενικευσιμότητα των αποτελεσμάτων τους. Επιπλέον, η αρχιτεκτονική της αποκεντρωμένης εφαρμογής της πλατφόρμας, δηλαδή η δυνατότητα υπογραφής συναλλαγών μέσω MetaMask εξασφαλίζει τη καταγραφή ως μη αποκήρυξή και αμετάβλητη, κάτι που στις έρευνες των Otoum και Bathula δεν συμβαίνει με την αποθήκευση των κλειδιών να γίνεται στην πλευρά του διακομιστή. Συνεπώς, η προτεινόμενη

πλατφόρμα της υπάρχουσας διπλωματικής εργασίας προσδίδει μια καινοτομία και εξέλιξη, καθώς έχει ξεφύγει από τα ερευνητικά πλαίσια και εμφανίζεται πρότυπη και λειτουργική.

Πίνακας 2.4: Συγκριτική ανάλυση συναφών εργασιών. ✓ = υποστηρίζεται, ✗ = δεν υποστηρίζεται, — = δεν εφαρμόζεται.

Εργασία	**ML HF Prediction**	**XAI (SHAP)**	**Blockchain**	**Κλινικό UI**	**GDPR Consent**	**Μεγάλο Dataset**
Moreno-Sánchez (2023)	✓	✓	✗	✗	✗	✗ (n=299)
[49]Bathula et al. (2024)	✓	✗	✓	✗	✗	—
[55]Otoum et al. (2023)	✓	✗	✓	✗	✗	—
[56]Muneer et al. (2024)	✓	✗	✗	✓	✗	✗
[40]Desai et al. (2020)	✓	✗	✗	✗	✗	✓ (n=9.502)
Παρούσα εργασία	✓	✓	✓	✓	✓	✓ (MIMIC-IV)

Κεφάλαιο 3: Θεωρητικό Υπόβαθρο

3.1 Αλγόριθμοι Μηχανικής Μάθησης

3.1.1 Λογιστική Παλινδρόμηση (Logistic Regression)

Η Λογιστική Παλινδρόμηση (Logistic Regression-LR) αποτελεί έναν αλγόριθμο εποπτευόμενης μηχανικής μάθησης που εφαρμόζεται σε προβλήματα ταξινόμησης. Ο συγκεκριμένος αλγόριθμος μοντελοποιεί τις λογαριθμικές πιθανότητες ενός συμβάντος ως έναν γραμμικό συνδυασμό μίας ή περισσότερων ανεξάρτητων μεταβλητών. Η πιο συνηθισμένη εκδοχή της παρουσιάζει η δυαδική ταξινόμηση η οποία ως έξοδο αποκτά δύο πιθανές κατηγορίες. Αυτό επιτυγχάνεται μέσω της λογιστικής (sigmoid) συνάρτησης, η οποία μετατρέπει οποιαδήποτε πραγματική τιμή στο διάστημα (0,1).

Παρουσιάζονται τρείς τύποι Λογιστικής Παλινδρόμησης. Η Διωνυμική,η Πολυωνυμική και η Τακτική. Η Διωνυμική Λογιστική Παλινδρόμηση (Binomial Logistic Regression) η εξαρτημένη μεταβλητή της συνάρτησης ανήκει σε δύο μόνο πιθανές κατηγορίες. Η ίδια αποτελεί την πιο συνηθισμένη μορφή της Λογιστικής Παλινδρόμησης. Η Πολυωνυμική Λογιστική Παλινδρόμηση (Multinomial Logistic Regression) είναι η εκδοχή λογιστικής παλινδρόμησης που η εξαρτημένη μεταβλητή εμφανίζει 3 ή περισσότερες πιθανές κατηγορίες που δεν είναι κατηγοριοποιημένες σε κάποια σειρά. Επεκτείνει τη δυαδική λογική της για την διαχείριση πολλαπλών κλάσεων. Η Τακτική (Ordinal Logistic Regression) εφαρμόζεται όταν 3 ή περισσότερες κατηγορίες κατανέμονται με φυσική σειρά.

Μαθηματική Εξήγηση

Έστω Χ πίνακας χαρακτηριστικών που ορίζεται από την εξίσωση (2) ως είσοδος:

$$X = \begin{bmatrix} X_{11} & \cdots & X_{1m} \\ \vdots & \ddots & \vdots \\ X_{n1} & \cdots & X_{nm} \end{bmatrix} \qquad (2)$$

Και Υ η εξαρτημένη μεταβλητή της συνάρτησης για την οποία ισχύει

$$Y = \begin{cases} 0 \; \alpha\nu \; Τάξη \; 1 \\ 1 \; \alpha\nu \; Τάξη \; 2 \end{cases}$$

Έπειτα, εφαρμόζουμε τη πολυγραμμική συνάρτηση της εξίσωση (3) στις μεταβλητές εισόδου Χ.

$$z = (\sum_{i=1}^{n} w_i x_i) + b \qquad (3)$$

Όπου ,

- xi = ιοστή παρατήρηση του Χ
- Wi = [w1,w2,…,wm] είναι τα βάρη ή συντελεστής
- b = bias όρος γνωστός ως παρεμβολή

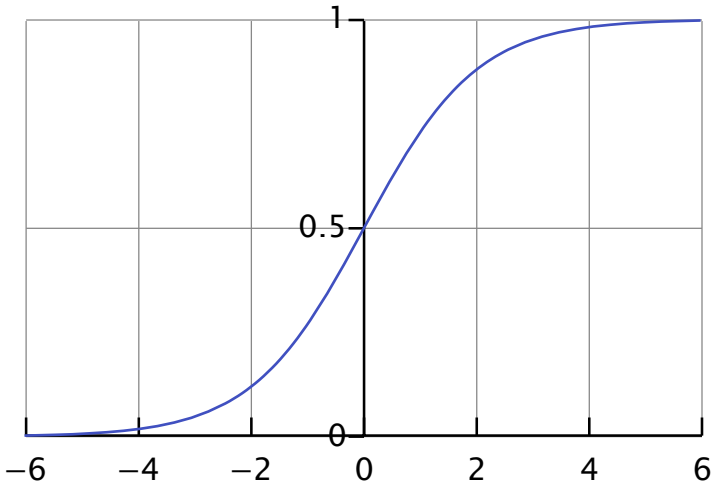
Επομένως ο τύπος αποκτά την γενική μορφή της εξίσωσης (3): $z = w \cdot X + b$

Μέχρι τώρα το z αποτελεί συνεχή τιμή με βάση την Γραμμική Παλινδρόμηση. Από εδώ και έπειτα εφαρμόζεται η λογιστική παλινδρόμηση μέσω της σιγμοειδής συνάρτησης με σκοπό τη μετατροπή της πιθανότητας ανάμεσα σε 0 και 1 ώστε να προβλέψουμε την κλάση που ανήκει.

Σιγμοειδής Συνάρτηση

Οι σιγμοειδείς συναρτήσεις χαρακτηρίζονται από τη μορφή ενός τελικού σίγμα και αποτελούν συνεχείς και παραγωγίσιμες προσεγγίσεις της βηματικής συνάρτησης. Μάλιστα στο όριο που η κλίση γίνεται πολύ μεγάλη, η σιγμοειδής ταυτίζεται με τη βηματική. Το συγκεκριμένο είδος μαθηματικών συναρτήσεων χρησιμοποιείται σε πληθώρα εφαρμογών μηχανικής μάθησης και τεχνητών νευρωνικών δικτύων. Παρατηρούνται στη βιβλιογραφία ότι χωρίζεται σε δύο βασικούς τύπους, τη Λογιστική και την Υπερβολική. Η πρώτη παρουσιάζει Στην παρούσα διπλωματική εργασία χρησιμοποιούμε τη Λογιστική, η οποία έχει την μορφή της εξίσωσης (4):

$$\sigma(z) = \frac{1}{1+e^{-z}} \qquad (4)$$



Στηριζόμενη στη γραφική παράσταση της, η συνάρτηση εμφάνιζε τις εξής καταστάσεις εξόδου:

- σ(z) τείνει προς τη τιμή 1 όταν $z \to \infty$
- σ(z) τείνει προς τη τιμή 0 όταν $z \to -\infty$
- σ(z) είναι πάντα μεταξύ των τιμών 0 και 1

Ενώ η πιθανότητα της μπορεί να μετρηθεί ως:

$$P(y = 1) = \sigma(z)$$

$$P(y = 0) = 1 - \sigma(z)$$

Έστω τώρα το αποτέλεσμα της πιθανότητας να είναι 1 δοσμένη της είσοδο x. Ισχύει η παρακάτω σχέση

$$p(x) = p(y = 1|x)$$

Η λογιστική συνάρτηση μοντελοποιεί τις πιθανότητες εμφάνισης του εξαρτώμενου γεγονότος, που είναι ο λόγος της πιθανότητας εμφάνισης του γεγονότος προς την πιθανότητα μη εμφάνισής του.

$$\frac{p(x)}{1 - p(x)} = e^z$$

Όπου λογαριθμίζοντας φυσικά τον λόγο πιθανοτήτων καταλήγουμε στη συνάρτηση logit (log odds) γνωστή και ως αντίστροφος της $g = \sigma^{-1}$ της τυπικής λογιστικής συνάρτησης. Συνεχίζοντας τις πράξεις καταλήγουμε στα εξής αποτελέσματα της εξίσωσης (5):

$$\log\left(\frac{p(x)}{1 - p(x)}\right) = z \leftrightarrow \log\left(\frac{p(x)}{1 - p(x)}\right) = w \cdot X + b \leftrightarrow \frac{p(x)}{1 - p(x)} = e^{w \cdot X + b} \leftrightarrow$$

$$p(x) = e^{w \cdot X + b}\big(1 - p(x)\big) \leftrightarrow p(x) = \frac{e^{w \cdot X + b}}{1 + e^{w \cdot X + b}}$$

Συνδυάζοντας το αποτέλεσμα με την υποθετική σχέση που κάναμε προηγουμένως, η τελική μορφή της συνάρτησης μας είναι

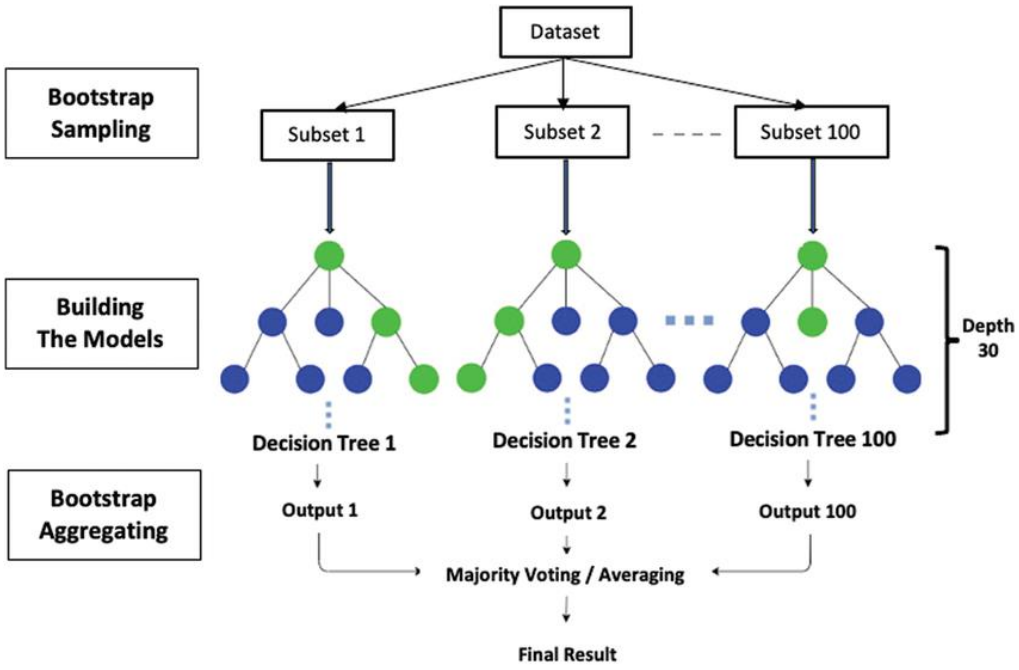
$$P(y = 1 \mid x) = \sigma\left(w^T x + b\right) = \frac{1}{1+e^{-(w^T x+b)}} \qquad (5)$$

3.1.2 Αλγόριθμος Τυχαίου Δάσους (Random Forest)

Ο αλγόριθμος Τυχαίου Δάσους αποτελεί μέθοδο ενσωματωμένης μάθησης (ensemble learning) η οποία χρησιμοποιείται ευρέως για ταξινόμηση, παλινδρόμηση και άλλες εργασίες που χρησιμοποιεί πολλά δέντρα αποφάσεων για να επιτύχει καλύτερες προβλέψεις. Η βασική ιδέα του αλγορίθμου είναι να συνδυάζει πολλά δέντρα απόφασης (decision trees), ώστε η συλλογική τους απόφαση να οδηγεί σε περισσότερο ορθές και αξιόπιστες προβλέψεις από ένα μεμονωμένο μοντέλο.

Η ορθή λειτουργία του εισάχθηκε σωστά στην έρευνα του Leo Breiman, μια από τις πιο αναγνωρισμένες ερευνητικές δημοσιεύσεις παγκοσμίως. Η εργασία ουσιαστικά αναφέρει πως ο αλγόριθμος δημιουργεί ένα δάσος από μη συσχετισμένα δέντρα μέσω της διαδικασίας CART (Classification and Regression Tree) συνδυάζοντας την μέθοδο bagging (bootstrap aggregating) και τυχαιοποίησης βέλτιστων κόμβων. Δηλαδή κάθε δέντρο χρησιμοποιεί διαφορετικό σύνολο δεδομένων και των αποτελεσμάτων συνδυασμένα για ταξινόμηση ή βρίσκοντας τον μέσο όρο για παλινδρόμηση που καθιστά η τεχνική ενσωματωμένης μάθησης. Για λειτουργίες ταξινόμησης, η έξοδος του τυχαίου δάσους είναι η κλάση που επιλέγεται από την πλειοψηφία του συνόλου των δέντρων, ενώ για λειτουργίες παλινδρόμησης, η έξοδος ειναι ο μέσος όρος των προβλέψεων της πλειοψηφίας των δέντρων.

Με την προσέγγιση αυτή μειώνεται το φαινόμενο της υπερπροσαρμογής (overfitting), το οποίο συνήθως εμφανίζει σε μεμονωμένα δέντρα αποφάσεων. Παράλληλα , ο αλγόριθμος παρουσιάζει αυξημένη αντοχή σε θόρυβο και ακραίες τιμές, ενώ προσμετρά στους υπολογισμούς του την σημαντικότητα χαρακτηριστικών (feature importance), που βασίζονται σε μείωση της ακαθαρσίας Gini κατά των διαχωρισμό των κόμβων.



Εικόνα 19 : Απεικόνιση της αρχιτεκτονικής του Τυχαίου Δάσους

Μαθηματική Διατύπωση

Έστω ένα σύνολο εκπαίδευσης :

$$D = \{(xi, yi)\}_{i=1}^{N}$$

με $x_i \in R^d$ διάνυσμα χαρακτηριστικών (features) της i-οστής παρατήρησης και y_i η αντίστοιχη κλάση (ή τιμή στόχου σε περίπτωση παλινδρόμησης)

Το D χωρίζεται σε πολλά τυχαία υποσύνολα με επανατοποθέτηση. Έστω Dt το bootstrap σύνολο δεδομένων (training dataset) που εκπαιδεύεται το t-οστού δέντρου Επομένως ισχύει,

$$D_t \subseteq D$$

Σε κάθε δέντρο απόφασης $h_t(x)$ εκπαιδεύεται πάνω στο bootstrap σύνολο D_t. Κατά τη διάρκεια της δημιουργίας του δέντρου, κάθε κόμβος τοποθετείται τυχαία ένα υποσύνολο χαρακτηριστικών m από το συνολικό πλήθος χαρακτηριστκών m από το γενικό πλήθος χαρακτηριστικών d, όπου $m < d$ και υπολογίζεται με βέλτιστο τρόπο ο διαχωρισμός μεταξύ αυτών των χαρακτηστικών. Η τελική πρόβλεψη του μοντέλου προκύπτει από τον συνδυασμό των επιμέρους προβλέψεων των δέντρων.

Σε περίπτωση ταξινόμησης, η πρόβλεψη κλάσης υπολογίζεται μέσω πλειοψηφικής ψήφου (majority voting) του συνόλου των δέντρων. Μαθηματικά εκφράζεται ως προς την Εξίσωση (6),

$$\hat{y} = mode\{h_1(x), h_2(x), \dots, h_T(x)\} \qquad (6)$$

Όπου:

- $h_T(x)$ η τιμή πρόβλεψης του t-οστού δέντρου
- T είναι ο συνολικό αριθμός δέντρων στο δάσος

Η συνάρτηση mode βασίζεται στη σχέση της εξίσωσης (7):

$$\hat{y} = \text{argmax} \sum_{t=1}^{T} I(h_t(x) = y) \qquad (7)$$

Με τον δείκτη I να παίρνει την τιμή 1 όταν η πρόβλεψη του δέντρου είναι ίση με την κλάση y και 0 διαφορετικά.

Σε περίπτωση παλινδρόμησης από κάθε δέντρο παράγεται μια αριθμητική πρόβλεψη. Ο μέσος όρος των προβλέψεων αποτελεί την τελική πρόβλεψη του αλγορίθμου στην εξίσωση (8):

$$\hat{y} = \frac{1}{T} \sum_{t=1}^{T} h_t(x) \qquad (8)$$

Η διαδικασία αυτή οδηγεί στην εξομάλυνση των αποκλίσεων μέσω του μέσου όρου ενώ τα σφάλματα που προκύπτουν από τα μεμονωμένα δέντρα μειώνονται.

3.1.3 Αλγόριθμος XGBoost

Το XGBoost (eXtreme Gradient Boosting) αποτελεί μια βιβλιοθήκη ανοιχτού κώδικα, η οποία παρέχει ένα πλαίσιο κανονικοποίησης για την ενισχυμένη διαβάθμιση. Αναγνωρίζεται για την αποδοτικότητα, τη ταχύτητα και τη μέγιστη λειτουργικότητα του και ανήκει στις μεθόδους boosting, οι οποίες εκπαιδεύουν δέντρα διαδοχικά, με κάθε νέο δέντρο να επικεντρώνεται στην επίλυση σφαλμάτων του προηγούμενου, βασιζόμενο στα κατάλοιπα (residuals). Το τελικό μοντέλο της συνάρτησης εκφράζεται ως προς την εξίσωση (9):

$$F(x) = \sum_t \eta_t \cdot f_t(x) \quad (9)$$

Όπου:

- f_t: το αποφασιστικό δέντρο
- n_t: ο βαθμός μάθησης

Η καινοτομία του XGBoost βασίζεται στη χρήση προσεγγίσεων $2^{ης}$ τάξης Taylor για την ελαχιστοποίηση του βαθμού της συνάρτησης, επιτυγχάνοντας την βέλτιστη σύγκλιση. Επίσης, περιλαμβάνει όρο κανονικοποίησης που ελέγχει τη πολυπλοκότητα κάθε δέντρου (αριθμό φύλλων και μέγεθος βαρών), ελαχιστοποιώντας το overfitting.

Η παρούσα υλοποίηση του εστιάζει στη χρήση του early stopping. Η λειτουργία αυτή τερματίζει την εκπαίδευση του μοντέλου όταν η απόδοση του συνόλου βαθμονόμησης δεν βελτιώνεται μετά

από ένα προκαθορισμένο αριθμό γύρων. Η υπερπαράμετρος scale_pos_weight ρυθμίζεται αυτόματα βάση της αναλογίας αρνητικών/ θετικών δειγμάτων, επιλύοντας την ανισορροπία των κλάσεων που χαρακτηρίζουν τα κλινικά δεδομένα.

3.1.4 Αλγόριθμος LightGBM

Το LightGBM ανήκει στο είδος Δέντρων Αποφάσεων με Ενίσχυση Διαβάθμισης (GBDT) και αποτελεί βελτιστοποιημένη εκδοχή τους. Παρουσιάζεται ιδιαίτερα δημοφιλής αλγόριθμος μηχανικής μάθησης, κυρίως για την αποτελεσματικότητα του. Εισάγονται δύο τεχνικές επιτάχυνσης: Gradient-based One-Side Sampling (GOSS) και Exclusive Feature Bundling (EFB). Η πρώτη τεχνική εστιάζει στη διατήρηση των δειγμάτων σε υψηλές κλίσεις, λόγω της υψηλής συνεισφοράς στη μάθηση του μοντέλου και απορρίπτει τυχαία τις μικρές κλίσεις, μειώνοντας το υπολογιστικό φόρτο κάνοντας το μοντέλο ταχύτερο στην εύρεση αποτελεσμάτων. Από την άλλη, η δεύτερη τεχνική Exclusive Feature Bundling (EFB) ελαττώνει τη διαστατικότητα του προβλήματος δεσμεύοντας τα αμοιβαία αποκλειόμενα χαρακτηριστικά, δηλαδή χαρακτηριστικά που σπάνια παίρνουν μη μηδενική τιμή την ίδια χρονική στιγμή για να μειώσει τον αριθμό των χαρακτηριστικών. Με αυτόν τον τρόπο δεν διακυβεύεται η ακρίβεια του σημείου διαχωρισμού.

Η κυριότερη διαφορά από το XGBoost μοντέλο κρίνεται στην αρχιτεκτονική και είναι η στρατηγική ανάπτυξη των δέντρων. Στον LightGBM τα δέντρα αναπτύσσονται κατά φύλλο (leaf-wise) αντί ανά επίπεδο (level-wise) που γίνεται στον XGBoost, επιλέγοντας το φύλλο με τη μεγαλύτερη μείωση απώλειας. Αυτές οι ιδιότητες καθιστούν το LightGBM αποδοτικό και ιδανική λύση για την χρήση μεγάλων δεδομένων συνόλων, όπως το MIMIC-IV με εκατοντάδες χιλιάδες νοσηλείες.

3.1.5 Βαθμονόμηση Πιθανοτήτων

Το θέμα είναι το εξής: η βαθμονόμηση της πιθανότητας είναι απολύτως κρίσιμη όταν πρόκειται για την εφαρμογή μοντέλων μηχανικής μάθησης σε κλινικά περιβάλλοντα. Πρόκειται για το να διασφαλιστεί ότι οι προβλεπόμενες πιθανότητες αντιστοιχούν στην πραγματική συχνότητα με την οποία συμβαίνουν τα γεγονότα. Για παράδειγμα, εάν ένα καλά βαθμονομημένο μοντέλο δίνει ως αποτέλεσμα $P(y=1|x) = 0.70$, τότε από τους ασθενείς που χαρακτηρίζει με αυτόν τον κίνδυνο, περίπου το 70% θα πρέπει να καταλήξει να αναπτύξει καρδιακή ανεπάρκεια. Αυτό έχει εξίσου μεγάλη σημασία με την ικανότητα του μοντέλου να διακρίνει (όπως το AUROC), επειδή όταν ένας κλινικός γιατρός βλέπει μια αριθμητική βαθμολογία κινδύνου, βασίζει τις αποφάσεις του σε αυτήν — επομένως, η ακρίβεια σε αυτές τις πιθανότητες είναι καθοριστική.

Τώρα, τα μοντέλα gradient boosting τείνουν να δημιουργούν προβλήματα εδώ — συχνά παράγουν πιθανότητες που είναι πολύ κοντά στο 0 ή στο 1, γεγονός που χαλάει τη βαθμονόμηση. Για να διορθωθεί αυτό, χρησιμοποιείται συνήθως η ισοτονική παλινδρόμηση μέσω του CalibratedClassifierCV του scikit-learn. Αυτό που κάνει είναι να προσαρμόζει μια μονοτονικά

αυξανόμενη, μη παραμετρική συνάρτηση που αντιστοιχίζει τις ακατέργαστες βαθμολογίες στο διάστημα [0,1]. Και το ωραίο είναι ότι δεν γίνονται υποθέσεις σχετικά με το σχήμα αυτής της συνάρτησης — είναι αρκετά ευέλικτη. Για να ελέγξουμε πόσο καλά λειτουργεί αυτή η βαθμονόμηση, βασιζόμαστε σε μετρήσεις όπως το Brier Score, το οποίο ουσιαστικά μετρά τη μέση τετραγωνική διαφορά μεταξύ των προβλεπόμενων πιθανοτήτων και των πραγματικών αποτελεσμάτων, καθώς και σε καμπύλες βαθμονόμησης που συγκρίνουν τις προβλεπόμενες συχνότητες με τις παρατηρούμενες σε δείγματα ομοιόμορφου μεγέθους.

Στην παρούσα ρύθμιση, αυτό το βήμα βαθμονόμησης εφαρμόζεται μόνο στο επιλεγμένο μοντέλο «Champion», χρησιμοποιώντας ένα σύνολο δεδομένων βαθμονόμησης που παραμένει αμετάβλητο καθ' όλη τη διάρκεια της εκπαίδευσης. Πρόκειται για ένα λεπτό κομμάτι του παζλ, αλλά αναμφισβήτητα ζωτικής σημασίας, διότι χωρίς καλά βαθμονομημένες πιθανότητες, η λήψη κλινικών αποφάσεων γίνεται στα τυφλά.

3.2 Μετρικές Αξιολόγησης

Η αξιολόγηση ενός προγνωστικού μοντέλου μηχανικής μάθησης σε κλινικό περιβάλλον απαιτεί τη χρήση μετρικών που βασίζεται σε τρεις διακριτές ιδιότητες, την διακριτική ικανότητα, την βαθμονόμηση πιθανοτήτων και τη κλινική χρησιμότητα. Κάθε ιδιότητα αντιπροσωπεύει διαφορετική λειτουργία και δεν υποκαθιστά τις υπόλοιπες. Στην παρούσα μελέτη, χρησιμοποιούμε ένα σύνολο δεδομένων μεγάλης χωρητικότητας. Σε περίπτωση λοιπόν που η πλειοψηφία των νοσηλειών εμφανίζουν χαμηλή πιθανότητα πρόκλησης καρδιακής ανεπάρκειας, η χρήση της ακρίβειας ως μετρική δεν είναι ορθή. Αυτό στηρίζεται στο γεγονός ότι το μοντέλο αγνοεί πλήρως τους ασθενείς υψηλού κινδύνου, λόγω του ότι πετυχαίνει υψηλότερη ακρίβεια με την πλειοψηφία των υγειών νοσηλευμένων. Η μαθηματική απόδειξη της παραπάνω τεκμηρίωσης αποδεικνύεται ανατρέχοντας στην Εξίσωση (10).

$$\text{Ακρίβεια (Accuracy)} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \qquad (10)$$

Στην περίπτωση μας, ο TN (True Negative) αποτελεί τεράστιο αριθμό, οπότε το μοντέλο επιβραβεύεται βρίσκοντας ορθά τους ασθενείς που ανήκουν στο TN, αγνοώντας τους ασθενής του TP συνόλου που βρίσκονται σε πραγματικό κίνδυνο.

Για τον λόγο αυτόν, οι μετρικές AUROC,AUPRC, Brier Score και DCA αξιολογούν εγκυρότερα την αποδοτικότητα του μοντέλου ανεξαρτήτως την αναλογία κλάσεων, εστιάζοντας στο μοναδικό πράγμα που έχει πραγματική σημασία στον κλινικό τομέα, την ορθή αναγνώριση ασθενών με υψηλό κίνδυνο.

3.2.1 Μετρικές Διακριτικής Ικανότητας

Η διακριτική ικανότητα είναι η ιδιότητα του μοντέλου να αποδίδει υψηλό ρίσκο κινδύνου σε ασθενείς που πραγματικά κινδυνεύουν από πρόκληση καρδιακής ανεπάρκειας, έναντι αυτών που δεν κινδυνεύουν πραγματικά. Η ίδια αξιολογείται μέσω δύο εξειδικευμένων μετρικών που έχουν συμπληρωματικό χαρακτήρα, τις AUROC και AUPRC που αξιολογούν την πιθανότητα μέσω της ίδιας διάστασης, της καμπύλης ROC, αλλά από διαφορετική οπτική γωνία.

Η καμπύλη ROC (Receiver Operating Characteristic) αποτελεί γραφική παράσταση που απεικονίζει τη σχέση μεταξύ του ποσοστού αληθών θετικών (True Positive Rate - TPR) (ευαισθησία) και του ποσοστού ψευδών θετικών (False Positive Rate - FPR) σε διαφορετικα όρια ταξινόμησης. Κατά την μείωση των ορίων ταξινόμησης, το μοντέλο γίνεται πιο ευαίσθητο, αυξάνοντας την ανίχνευση θετικών περιπτώσεων, αλλά και των ψευδών θετικών.

Η AUROC εκφράζει την ικανότητα διάκρισης που διαθέτει ένας ταξινομητής μεταξύ θετικών και αρνητικών κατηγοριών, αξιολογώντας την σχέση των ποσοστών αληθών θετικών (TP) και ψευδών θετικών (FP) σε διαφορετικά κατώφλια. Συνοπτικό μέτρο διακριτικής ικανότητας παρουσιάζεται το εμβαδόν κάτω από την καμπύλη ROC. Έστω ότι έχουμε τα εξής σημεία που ανήκουν στην ROC καμπύλη

$(\mathrm{FPR1}, \mathrm{TPR1}), (\mathrm{FPR2}, \mathrm{TPR2}), \dots, (\mathrm{FPRn}, \mathrm{TPRn})$ με n: συνολικό αριθμό σημείων

Εφαρμόζεται η μέθοδος του τραπεζίου (trapezoidal rule) για τον υπολογισμό του εμβαδού και προκύπτει η μαθηματική Εξίσωση (11).

$$AUC = \sum_{i=1}^{n-1} (FPR_{i+1} - FPR_i) \cdot \frac{TPR_{i+1} + TPR_i}{2} \qquad (11)$$

Πριν τον υπολογισμό όμως του εμβαδού ακολουθεί η ταξινόμηση των προβλεπόμενων πιθανοτήτων, μεταβάλλοντας την τιμή του κατωφλίου (threshold) από 0 έως 1.Για κάθε αλλαγή του ορίου υπολογίζονται οι τιμές των εξισώσεων (12) και (13):

$$TPR = \frac{TP}{TP+FN} \quad , \; FPR = \frac{FP}{FP+TN} \quad (12),(13)$$

Ωστόσο, παρουσιάζεται ένας σημαντικός περιορισμός στα ανισόρροπα σύνολα δεδομένων, ο μεγάλος αριθμός αρνητικών δειγμάτων καθιστά τον FPR χαμηλό, η AUROC παραμένει υψηλή ακόμα και στο ενδεχόμενο κακής απόδοσης της θετικής κλάσης. Τον συγκεκριμένο περιορισμό έρχεται να αναιρέσει η AUPRC, η οποία παρουσιάζεται αυστηρότερη στην κατάταξη των ασθενών, καθώς η Precision, που είναι ο αριθμός των πραγματικά θετικών προς το σύνολο των προβλεφθέντων θετικών δεν επηρεάζεται από το πλήθος των αρνητικών δειγμάτων. Επικεντρώνεται αποκλειστικά στην απόδοση της θετικής κλάσης. Υψηλή AUPRC υποδηλώνει πως το μοντέλο εντοπίζει αξιόπιστα τους ασθενείς υψηλού κινδύνου. Για τους παραπάνω λόγους

και εξαιτίας της ανισορροπίας κλάσεων του MIMIC-IV, η AUPRC επιλέγεται ως κύρια μετρική σύγκρισης τον μοντέλων.

3.2.2 Μετρικές Βαθμονόμησης

Σε ένα ιατρικό περιβάλλον, για την επαλειμμένη αξιοπιστία των προβλέψεων δεν αρκεί μόνο η χρήση της διακριτικής ικανότητας. Ένα μοντέλο ταξινόμησης μπορεί να κατατάσσει σωστά τους ασθενείς στις ανάλογες κατηγορίες, χωρίς να παρέχει σωστά τις αριθμητικές πιθανότητες. Στην περίπτωση μας, η ακριβής πρόβλεψη κινδύνου καρδιακής ανεπάρκειας είναι κρίσιμη για την απόφαση που θα λάβει ο ιατρός, η βαθμονόμηση των πιθανοτήτων κρίνεται απαραίτητη.

Ο Brier Score παρουσιάζεται κατάλληλος κανόνας βαθμολόγησης που μετρά την ακρίβεια των πιθανοτήτων προβλέψεων. Για μονοδιάστατες προβλέψεις , είναι ισοδύναμος με το μέσο τετραγωνικό σφάλμα (Mean Square Error) που εφαρμόζεται στις προβλεπόμενες πιθανότητες. Ορίζεται από την εξίσωση (15)

$$BS = \frac{1}{N}\sum_{t=1}^{N}(f_t - O_t)^2 \qquad (15)$$

Όπου p_i είναι η προβλεπόμενη πιθανότητα, O_t το πραγματικό αποτέλεσμα της t-οστής έκβασης και N ο αριθμός των περιπτώσεων έκβασης. Η παραπάνω φόρμουλα παρουσιάζεται ιδανική στο πείραμα της διπλωματικής εργασίας καθώς εμφανίζεται ιδανική για γεγονότα δυαδικής ταξινόμησης, όπως η πιθανότητα «εκδήλωσης» ή «μη εκδήλωσης» καρδιακής ανεπάρκειας. Ο δείκτης εφαρμόζεται πριν και μετά την διαδικασία της isotonic regression που περιγράφηκε στο 3.1.5 – χρησιμοποιείται για να αποδειχθεί η βελτίωση της βαθμονόμησης στις προβλεπόμενες πιθανότητες.

Η καμπύλη βαθμονόμησης (calibration curve) παρέχει γραφική αξιολόγηση της εγκυρότητας του μοντέλου. Το μοντέλο παράγει προβλεπόμενες πιθανότητες με βάση τον τύπο της εξίσωσης (16)

$$\widehat{p_i} = P(Y = 1|x_i) \qquad (16)$$

για κάθε δείγμα, όπου i-οστή παρατήρηση. Στην συνέχεια οι προβλεπόμενες πιθανότητα ες θα ομαδοποιηθούν σε διαστήματα (bins), ίσου πλάτους στο διάστημα [0,1]. Για την γραφική απεικόνιση του κάθε bin ως σημειο (pin,fbin) στο επίπεδο υπολογίζονται δύο βασικά μεγέθη των εξισώσεων (17) και (18),

$$\overline{p_{bin}} = \frac{1}{n_{bin}}\sum_{i=1}^{n_{bin}}\widehat{p_i} \quad , \; f_{bin} = \frac{\sum y_i}{n_{bin}} \qquad (17), (18)$$

Πραγματοποιούνται συγκρίσεις μεταξύ των bins και της διαγώνιος y=x που η ίδια εκφράζει ένα τέλειο βαθμονομημένο μοντέλο. Οι αποκλίσεις κάτω από τη διαγώνιο υποδηλώνει υπερεκτίμηση των αποτελεσμάτων του μοντέλου. Το αντίστροφο υποδηλώνουν οι αποκλίσεις πάνω από τη διαγώνιο.

3.2.3 Μετρική Καθαρού Οφέλους

Η Ανάλυση Καμπύλης Απόφασης, σε αντίθεση με τα προηγούμενα διαγράμματα, ποσοτικοποιεί τη κλινική χρησιμότητα, δηλαδή το καθαρό όφελος που προκύπτει από τη χρήση του μοντέλου ως υποστηρικτικού εργαλείου απόφασης, συγκρίνοντας δύο στρατηγικές αναφοράς: την «Θεραπεία Όλων», που αφορά τη θεραπεία κάθε ασθενή χωρίς διάκριση και τη «Καμία Θεραπεία», που αφορά την παράλειψη παρέμβασης. Σύμφωνα με τη μεθόδο του Vickers, το καθαρό όφελος υπολογίζεται από την εξίσωση (19):

$$NB(p_t) = \frac{TP}{N} - \frac{FP}{N} \cdot \frac{p_t}{1-p_t} \qquad (19)$$

Όπου p_t αποτελεί το κατώφλι πιθανότητας πάνω στο οποίο θεωρείται απαραίτητη η παρέμβαση και ο όρος $\frac{p_t}{1-p_t}$ αντικατοπτρίζει τη σχετική στάθμιση ζημίας μεταξύ μη-εντοπισμένης πάθησης (FN) και άσκοπης παρέμβασης (FP).

3.2.4 Κλινική Χρησιμότητα

Ωστόσο, ακόμη και ένα μοντέλο με υψηλή διακριτική ικανότητα και καλή βαθμονόμηση ενδέχεται να μην είναι κλινικά χρήσιμο. Ένα τέτοιο μοντέλο μπορεί να αποδίδει χειρότερα από την απλή, συχνά αφελή, στρατηγική της παρέμβασης σε όλους τους ασθενείς. Σε αυτό το σημείο, η Ανάλυση Καμπύλης Απόφασης (Decision Curve Analysis, DCA), που προτάθηκε από τους Vickers και Elkin (2006), συμβάλλει στη μέτρηση μιας τρίτης διάστασης: κατά πόσο το μοντέλο προσφέρει πραγματικό κλινικό όφελος σε σύγκριση με αυτές τις ακραίες εναλλακτικές λύσεις. Η DCA συγκρίνει τρεις στρατηγικές λήψης αποφάσεων. Η πρώτη, η «θεραπεία όλων», υποθέτει ότι ο κλινικός γιατρός παρεμβαίνει σε κάθε περίπτωση. Αυτή η στρατηγική δεν απαιτεί τη χρήση μοντέλου, αλλά οδηγεί σε πολλές περιττές παρεμβάσεις. Η δεύτερη, η «μη θεραπεία κανενός», προϋποθέτει πλήρη αποφυγή παρεμβάσεων, αποτρέποντας όλες τις περιττές ενέργειες, αλλά χωρίς να εντοπίζει ασθενείς υψηλού κινδύνου. Η τρίτη στρατηγική αφορά το ίδιο το μοντέλο, το οποίο παρεμβαίνει επιλεκτικά μόνο σε ασθενείς με προβλεπόμενη πιθανότητα πάνω από ένα προκαθορισμένο όριο pt. Εάν η καμπύλη του μοντέλου υπερτερεί των δύο ακραίων στρατηγικών στο κλινικά σχετικό εύρος τιμών pt, τότε το μοντέλο θεωρείται κλινικά χρήσιμο.

3.3 SHAP

Η διασφάλιση της ερμηνείας των μοντέλων καθίσταται κύρια προϋπόθεση για την ενσωμάτωση τους σε κλινικά περιβάλλοντα λήψης αποφάσεων. Τα μοντέλα που λειτουργούν ως «μαύρα κουτιά» (black boxes), παράγουν προβλέψεις χωρίς επαρκή αιτιολόγηση, μη παρέχοντας παράλληλα την απαραίτητη διαφάνεια που απαιτείται από το κλινικό προσωπικό για την κατανόηση των υποκείμενων βιοδεικτών που οδηγούν σε εκτίμηση κινδύνου. Το πλαίσιο SHAP εισήχθη από τους Lundberg και Lee παρουσιάζεται ως μια ενοποιημένη προσέγγιση για την

ερμηνεία των προβλέψεων, συνδυάζοντας την αυστηρότητα της θεωρίας παιγνίων με την απόδοση των χαρακτηριστικών, καθιστώντας το ένα από τα κορυφαία εργαλεία Εξηγήσιμης Τεχνητής Νοημοσύνης (XAI) που υπάρχουν στην βιβλιογραφία.

3.3.1 Θεωρία Shapley Values

Η θεωρητική προσέγγιση του SHAP βασίζεται στη Θεωρία Συνεργατικών Παιγνίων, η οποία εισήχθη από τον Lloyd Shapley. Στο πλαίσιο της μηχανικής μάθησης, τα χαρακτηριστικά εισόδου (input features) αντιμετωπίζονται ως «παίκτες» σε παίγνιό, όπου η συνολική «αμοιβή» αντιστοιχεί στην απόκλιση μεταξύ της πρόβλεψης του μοντέλου και της μέσης τιμής του. Η τιμή Shapley (φ_i) για το i-οστό χαρακτηριστικό ορίζεται μαθηματικά ως σταθμισμένη μέση οριακή συνεισφορά που παρέχει σε όλους τους δυνατούς συνδυασμούς χαρακτηριστικών της εξίσωσης (20):

$$\varphi_i(f,x) = \sum_{S\subseteq F\{i\}} \frac{|S|!(p-|S|-1)!}{p!}[f(S\cup\{i\}) - f(S)] \qquad (20)$$

Όπου:

- F είναι το σύνολο των χαρακτηριστικών p = |F|
- S είναι το υποσύνολο χαρακτηριστικών που δεν περιλαμβάνουν το χαρακτηριστικό i.
- $f(S)$ είναι η προβλεπόμενη τιμή από την πρόβλεψη των δεδομένων των χαρακτηριστικών στο σύνολο S.

Η μέθοδος SHAP ικανοποιεί εύστοχα τέσσερις κρίσιμες ιδιότητες: της τοπικής ακρίβειας (local accuracy/efficiency), της έλλειψης ευαισθησίας (missingness/dummy), της συνέπειας (consistency/monotonicity) και της συμμετρίας (symmetry). Η ιδιότητα της ακρίβειας εξασφαλίζει πως το άθροισμα των τιμών Shapley του συνόλου των χαρακτηριστικών ισούται με τη διαφορά μεταξύ της τρέχουσας πρόβλεψης και της μέσης προβλεπόμενης τιμής.

3.3.2 Υπολογιστική Στρατηγική

Πέραν της θεωρητικής υπεροχής τους, η εκθετική πολυπλοκότητα O(2^p) των τιμών Shapley καθιστά τον ακριβή υπολογισμό τους ανέφικτο για μεγάλο αριθμό μεταβλητών. Η ανάπτυξη εξειδικευμένων αλγορίθμων (explainers) που εκμεταλλεύονται τη δομή συγκεκριμένων αρχιτεκτονικών μοντέλων.

Το **TreeExplainer** αποτελεί ιδανική μέθοδο για την ερμηνεία μοντέλων που βασίζονται σε δέντρα αποφάσεων (π.χ. XGBoost, Random Forest, LightGBM). Ο αλγόριθμος χρησιμοποιεί την ιεραρχική δομή των δέντρων για τον υπολογισμό των τιμών Shapley σε πολυωνιμικό χρόνο O($T \cdot L \cdot p$). Μέσω της προσέγγισης αυτή διατηρείται η μαθηματική ακρίβεια των τιμών χωρίς την

ανάγκη δειγματοληψίας, αποφεύγοντας την εισαγωγή στατιστικού σφάλματος και την επιβράδυνση του συστήματος.

Το **LinearExplainer** εφαρμόζεται σε γραμμικά μοντέλα, όπως η Λογιστική Παλινδρόμηση. Οι τιμές SHAP υπολογίζονται αναλυτικά λαμβάνοντας υπόψη τη συν διακύμανση μεταξύ των μεταβλητών, γεγονός που επιτρέπει την ακριβή αποτύπωση της συνεισφοράς κάθε βιοδείκτη στο τελικό αποτέλεσμα.

3.3.3 Οπτικοποίηση και Σύνθεση Ερμηνειών

Η πρακτική αξία του SHAP έγκειται στην ικανότητα μετάφρασης πολύπλοκων μαθηματικών μεγεθών σε εποπτικά διαγράμματα, διευκολύνοντας πολλαπλές πτυχές της κλινικής αξιολόγησης. Η ερμηνεία τους χωρίζεται σε δύο πτυχές.

Καθολική Ερμηνεία. Τα διαγράμματα Beeswarm επιτρέπουν τη συνολική επισκόπηση του μοντέλου, κατατάσσοντας τα χαρακτηριστικά σε ολόκληρο σύνολο δεδομένων βάσει της σπουδαιότητας αποκαλύπτοντας την κατεύθυνση της επίδρασης τους (θετική ή αρνητική συσχέτιση με το στόχο). Παράλληλα, τα διαγράμματα Dependence Plots αναδεικνύουν σχέση μεταξύ των μεταβλητών.

Τοπική Ερμηνεία. Τα διαγράμματα επικεντρώνεται στην εξήγηση μεμονωμένων προβλέψεων για συγκεκριμένους ασθενείς. Τέτοιου είδους διαγράμματα παρουσιάζονται τα Waterfall Plots και Force Plots που απεικονίζουν πώς κάθε χαρακτηριστικό «ωθεί» (push) την πρόβλεψη μακριά από τη μέση τιμή, λειτουργώντας ως οπτική απόδειξη των παραγόντων που διαμόρφωσαν τη διάγνωση για ένα συγκεκριμένο ασθενή.

Τα παραπάνω εργαλεία ενσωματώνονται στη μεθοδολογία της παρούσας διπλωματικής διατριβής, παρέχοντας στο ιατρικό προσωπικό τους κρισιμότερους παράγοντες κινδύνου ανά περίπτωση, ενισχύοντας την διαφάνεια και τη συμμόρφωση στις κανονιστικές απαιτήσεις (π.χ. GDPR) περί αυτοματοποιημένης λήψης αποφάσεων και της εξηγήσιμης τεχνητής νοημοσύνης.

3.4 Τεχνολογία Blockchain

Η ένταξη της τεχνολογίας της Αλυσίδας Συστοιχιών (Blockchain) σε συστήματα υποστήριξης κλινικών αποφάσεων επιλύει το κρίσιμο πρόβλημα του Σημείου Μοναδικού Σφάλματος (Single Point of Failure) που χαρακτηρίζεται ως κύριο μειονέκτημα των συμβατικών κεντρικών αρχιτεκτονικών. Παραδείγματος χάριν, σε ένα σύστημα μιας ιατρικής κλινικής με κεντρικό κατανεμητή, ο διαχειριστής της βάσης δεδομένων διατηρεί τον πλήρη έλεγχο εγγραφών, γεγονός που καθιστά δεδομένα όπως το ιστορικό των ασθενών ή τα αποτελέσματα των προβλέψεων, ευάλωτα σε περίπτωση κυβερνοεπίθεσης, αναδρομικής τροποποίησης ή διαγραφής. Η Αλυσίδα Συστοιχιών αντιμετωπίζει αυτό το ελάττωμα δημιουργώντας ένα απαραβίαστο ίχνος ελέγχου

(audit trail), όπου η ακεραιότητα των δεδομένων εγγυάται μαθηματικά και όχι μέσω εμπιστοσύνης σε μια κεντρική αρχή.

3.4.1 Βασικές Αρχές Λειτουργίας

Η Αλυσίδα Συστοιχιών λειτουργεί ως αποκεντρωμένη κατανεμημένη βάση δεδομένων, στην οποία τα δεδομένα καταγράφονται κρυπτογραφημένα σε πολλούς υπολογιστές, καθιστώντας την ανθεκτική σε παραβιάσεις. Ουσιαστικά επιτρέπει την αποθήκευση των συναλλαγών και τον εντοπισμό οποιοδήποτε είδους συναλλαγής που εγκρίνεται από τον μηχανισμό συναίνεσης (consensus mechanism), προσδίδοντας συμφωνίας κατά μήκος του δικτύου. Κάθε συναλλαγή ομαδοποιείται σε μπλοκ (block), που είναι συνδεδεμένα μεταξύ τους. Το κάθε μπλοκ απαρτίζεται από τα εξής χαρακτηριστικά: έναν δείκτη (index), μια χρονοσήμανση (timestamp), τα δεδομένα συναλλαγής και το hash του προηγούμενου block. Το hash σύνδεσης χαρακτηρίζεται ως «ταυτότητα» του μπλοκ και είναι ο κύριος παράγοντας σύνδεσης μεταξύ των blocks. Βασίζεται στον αλγόριθμο κατακερματισμού **SHA-256**, ο οποίος παρουσιάζει κύριο χαρακτηριστικό του το φαινόμενο της χιονοστιβάδας (avalanche effect), κατά το οποίο η παραμικρή τροποποίηση έστω και ενός bit στα δεδομένα του block μεταβάλλει πλήρως το παραγόμενο hash, ακυρώνοντας την μαθηματική συνοχή του με το επόμενο block της αλυσίδας. Αυτή η αναδρομική εξάρτηση (backward dependency) αποτελεί τον πυρήνα της αμεταβλητότητας (immutability) του συστήματος, εξασφαλίζοντας την ακεραιότητα, την προστασία των δεδομένων σε κρίσιμες υποδομές.

Η συναίνεση στο δίκτυο πραγματοποιείται μέσω αποκεντρωμένων πρωτοκόλλων, με την παρούσα εργασία να βασίζεται στο **Proof of Stake** του Ethereum. Το Proof of Stake ήρθε να λύσει το πρόβλημα του υψηλού υπολογιστικού κόστους που παρουσιάζει ο μηχανισμός Proof-of-Work (PoW) κατά την ιστορική μετάβαση «The Merge» στις 15 Σεπτεμβρίου 2022. Χρησιμοποιούνται επαληθευτές (validators) έναντι των miners που χρησιμοποιεί το proof-of-work blockchain. Οι επαληθευτές λαμβάνουν ανταμοιβή για την εργασία τους μια ορισμένη ποσότητα token της αλυσίδας μπλοκ, με αποτέλεσμα οι πιθανοί επιτιθέμενοι να πρέπει να αποκτήσουν ένα μεγάλο μέρος της αλυσίδας συστοιχιών για να πραγματοποιήσουν μια επίθεση. Αυτό καθιστά την επίθεση χρονοβόρα και με τεράστιο ενεργειακό κόστος, μη συμφέροντας τις περισσότερες φορές τον επιτεθειμένο να μπει στην διαδικασία να την πραγματοποιήσει. Η επιλογή του PoS προσφέρει σημαντική ενεργειακή απόδοση άνω του 99.5% και αυξημένη ασφάλεια μέσω των οικονομικών κινήτρων (staking), θέτοντας το σύστημα βιώσιμο για μεγάλης κλίμακας κλινικά συστήματα.

3.4.2 Στρατηγική Αποθήκευσης και GDPR Compliance

Το κρισιμότερο σημείο της υπάρχουσας αρχιτεκτονικής είναι ο διαχωρισμός των δεδομένα με βάση τον τρόπο αποθήκευσης τους, σε On-chain και Off-chain. Η αποθήκευση πλήρων κλινικών δεδομένων απευθείας στην αλυσίδα συστοιχιών, θα τον Γενικό Κανονισμό Προστασίας

Δεδομένων (GDPR) της Ευρωπαϊκής Ένωσης, διότι η αμεταβλητότητα της υπάρχουσας τεχνολογίας συγκρούεται με μια σειρά δικαιωμάτων, όπως αυτό της Λήθης (Right to be Forgotten – Άρθρο 17), της Διόρθωσης (Right to Rectification – Άρθρο 16), της Ενημέρωσης (Right to be Informed – Άρθρο 13-14) και άλλα πολλά.

Για τον λόγο αυτό, ακολουθείται η στρατηγική **Privacy by Design,** στην οποία τα ευαίσθητα δεδομένα παραμένουν off-chain, ενώ στο blockchain καταχωρείται ένα μόνο κρυπτογραφικό αποτύπωμα (fingerprint):

{patient_hash, risk_score, data_integrity_hash, model_version}

Αυτή η προσέγγιση διασφαλίζει τα εξής:

1. Η ακεραιότητα της πρόβλεψης είναι επαληθεύσιμη από οποιονδήποτε μέσω του hash.
2. Η εμπιστευτικότητα προστατεύεται, καθώς δεν αποκαλύπτονται ιατρικά δεδομένα on-chain.
3. Η συμμόρφωση με τον GDPR διατηρείται, καθώς τα μη επεξεργασμένα δεδομένα μπορούν να διαγραφούν από την off-chain βάση χωρίς να επηρεαστεί η δομή της αλυσίδας.

Σε αντίθεση με μία απλή κρυπτογραφημένη βάση δεδομένων, που προστατεύει μόνο την πρόσβαση, το blockchain διασφαλίζει ότι καμία καταχωρημένη πρόβλεψη δεν μπορεί να αναιρεθεί ή να αλλοιωθεί από κακόβουλο διαχειριστή, μετατοπίζοντας την εμπιστοσύνη από τον ανθρώπινο παράγοντα στους αλγορίθμους.

3.5 Έξυπνα Συμβόλαια

Τα Έξυπνα Συμβόλαια είναι αυτοεκτελούμενα προγράμματα (self-executing) τα οποία εκτελούνται για την αυτοματοποίηση μιας συμφωνίας, ώστε όλοι οι συμμετέχοντας να μπορούν να είναι σίγουροι για το αποτέλεσμα, χωρίς διαμεσολάβηση τρίτων προσώπων ή ύπαρξη απώλειας χρόνου. Αυτοματοποιούν την ροή μιας λειτουργίας, ενεργοποιώντας την επόμενη ενέργεια εφόσον πληρούνται προκαθορισμένες συνθήκες. Ο όρος «Έξυπνο Συμβόλαιο» αναφέρθηκε για πρώτη φορά από τον επιστήμονα πληροφορικής Nick Szabo το 1996, αναφερμένος σε συμβόλαια τα οποία θα εκτελούνταν με φυσική ιδιοκτησία (όπως υλικό ή λογισμικό) αντί με το νόμο. Ο Szabo περιέγραψε τη λειτουργία των αυτόματων πωλητών ως ένα παράδειγμα αυτής της έννοιας. Στην πραγματικότητα, τα Έξυπνα συμβόλαια αποτελούν πρωτόκολλα υπολογιστών που έχουν σχεδιαστεί για την διευκόλυνση, επικύρωση και επιβολή μιας ψηφιακής διαπραγμάτευσης ή την εκτέλεση ενός συμβολαίου, κληρονομώντας τρία βασικά χαρακτηριστικά της τεχνολογίας blockchain. Πρώτον, λειτουργούν χωρίς παραβιάσεις (tamper-proof), καθώς δεν γίνονται μετατροπές στον κώδικα του συμβολαίου μετά την ανάπτυξή του. Δεύτερον, τα δεδομένα που καταχωρούνται σε ένα συμβόλαιου του blockchain αποθηκεύονται μόνιμα σε αυτό. Τρίτον, προσδίδουν διαφάνεια με τον κώδικα να είναι ανοιχτός και ορατός, δίνοντας την δυνατότητα επαλήθευσης (validate) στοιχείων και την αέναη λειτουργία του συμβολαίου όσο υπάρχουν

κόμβοι στο δίκτυο. Η εμφάνιση τους είναι συνηθέστερη στον χώρο τον κρυπτονομισμάτων, των αποκεντρωμένων οικονομιών (DeFi) και στις μη ανταλλάξιμες token (non-fungible token) εφαρμογές.

Στο πλαίσιο ενός κλινικού συστήματος λήψης αποφάσεων, τα έξυπνα συμβόλαια έχουν τον ρόλο της επιχειρηματικής λογικής (business logic), εξασφαλίζοντας πως οι προβλέψεις του μοντέλου καταγράφονται αμετάβλητες και είναι προσβάσιμες μόνο από εξουσιοδοτημένους χρήστες.

3.5.1 Γλώσσα Προγραμματισμού Solidity

Το έξυπνο συμβόλαιο που αναπτύχθηκε στο πλαίσιο της υπάρχουσας διπλωματικής πραγματοποιήθηκε στη γλώσσα Solidity. Η Solidity είναι μια στατιστικά αποτυπωμένη γλώσσα προγραμματισμού, υψηλού επιπέδου που χρησιμοποιείται για την σχεδίαση έξυπνων συμβολαίων που εκτελούνται πάνω σε εικονική μηχανή του Ethereum (EVM) ή άλλες συμβατές εικονικές μηχανές, όπως το Hyperledger Fabric που είναι αποκλειστικά για επιχειρήσεις. Μάλιστα, οργανισμοί παγκόσμιας εμβέλειας. όπως η SWIFT (εταιρεία τηλεπικοινωνιών), που έχουν πειραματιστεί δημιουργώντας το δικό τους proof of concept με χρήση της Solidity για τη διασύνδεση τραπεζικών συστημάτων.

Η Solidity παρουσιάζει σημαντικές ομοιότητες με την ECMAScript, κανοντάς την προσιτή σε προγραμματιστές διαδικτυακών εφαρμογών. Ιδιαίτερη εμφανίζεται η χρήση της γλώσσας, λόγω της αποκεντροποίησης και ενίσχυσης της ιδιωτικότητας, σε μια νέα αναδυόμενη εκδοχή του Παγκόσμιου Ιστού, το Web3. Από την άλλη, σημαντική είναι η διαφορά από τις συνηθισμένες γλώσσες προγραμματισμού διαδικτυακών εφαρμογών, εξαιτίας της υποστήριξης μεταβλητών τύπων επιστροφής και αυστηρής διαχείρισης μνήμης. Στην παρούσα υλοποίηση αξιοποιήθηκαν σημαντικά χαρακτηριστικά της γλώσσας, όπως:

- **Σύνθετες Δομές Δεδομένων:** Οι ιεραρχικές απεικονίσεις (mappings) και δομές (structs), που υποστηρίζει η γλώσσα επιτρέπουν την αποθήκευση κρυπτογραφικών αποτυπωμάτων (hashes) των ασθενών με τρόπο αποδοτικό ως προς το κόστος Gas.
- **Κληρονομικότητα:** Η ιδιότητα της γραμμικοποίησης C3 (C3 linearization) παρέχει τη δυνατότητα εφαρμογής της πολλαπλής κληρονομικότητας, επιτρέποντας την ενσωμάτωση ελεγμένων και έτοιμων προτύπων ασφαλείας από τη βιβλιοθήκη OpenZeppelin (π.χ. AccessControl, Ownable).
- **Διεπαφή Δυαδικής Εφαρμογής:** Με την εκτέλεση του έξυπνου συμβολαίου Solidity παράγεται το ABI (Application Binary Interface), που λειτουργεί ως «μεταφραστής» ανάμεσα στο έξυπνο συμβόλαιο και τις αποκεντρωμένες εφαρμογές που επικοινωνούν μαζί του. Μέσω αυτού του μηχανισμού διεπαφής, οι συναρτήσεις του συμβολαίου καλούνται με ασφάλεια τύπων.
- **NatSpec:** Το NatSpec αποτελεί σύστημα σχολίων που μπορεί να γράψει ο προγραμματιστής περιγραφές σε φυσική γλώσσα εντός του περιβάλλον συγγραφής κώδικα

για την διευκόλυνση της κατανόησης των λειτουργιών του συμβολαίου από εξωτερικούς χρήστες. Η πλειοψηφία των κλινικών χρηστών εμφανίζει απουσία γνώσης σχετικά με τον τομέα πληροφορικής. Αυτό το χάσμα γεφυρώνεται με την χρήση του αναφερόμενου συστήματος.

Η επιλογή της Solidity έναντι άλλων γλωσσών του EVM, όπως η Serpent ή Mutan, βασίστηκε στην ολοένα και ευρύτερη υποστήριξη της κοινότητας και της δυνατότητας εφαρμογής προηγμένων μηχανισμών ασφαλείας, που είναι απαραίτητοι για τη διαχείριση κλινικών δεδομένων.

3.5.2 Έλεγχος Προσβασιμότητας

Η διαχείριση των ιατρικών δεδομένων απαιτεί την εφαρμογή αυστηρών μηχανισμών αυθεντικοποίησης. Για τον σκοπό αυτό, χρησιμοποιείται το πρότυπο Role-Based Access Control μέσω της βιβλιοθήκης OpenZeppelin AccessControl. Η αρχιτεκτονική αυτή διαχωρίζει τα δικαιώματα σε διαφορετικά επίπεδα:

- **ADMIN_ROLE:** Ο Διαχειριστής (Admin) διαχειρίζεται το συμβόλαιο και αποδίδει αρμοδιότητες μέσω ρόλων.
- **DOCTOR:** Ο γιατρός είναι χρήστης που συνδέεται μέσω του frontend στην εφαρμογή. Ο συγκεκριμένος ρόλος έχει το δικαίωμα της επαλήθευσης, της καταγραφής συμφωνίας ή διαφωνίας με τα αποτελέσματα των προβλέψεων, την πρόσβαση στο ιστορικό των ασθενών και την δυνατότητα έκτακτης πρόσβασης με υποχρεωτική αιτιολόγηση.
- **RESEARCHER:** Ο ερευνητής έχει αποκλειστικά πρόσβαση στα ανώνυμα δεδομένα, εφόσον έχει παρθεί η έγκριση από τον ασθενή. Δεν καταγράφει on-chain.
- **ORACLE_ROLE:** Είναι η διεύθυνση Ethereum που ανήκει στο FastAPI backend. Είναι πρόγραμμα, όχι φυσική οντότητα που διαθέτει αποκλειστικό δικαίωμα την καταχώρηση ρίσκων κινδύνου στην αλυσίδα πάνω (on-chain) μέσω συναρτήσεων.

Τα πρότυπα της OpenZeppelin αντιμετωπίζει την ευπάθεια του συστήματος έναντι επιθέσεων σε τρωτά σημεία (vulnerabilities), με την χρήση εκτενών ελέγχων (security audits).

3.5.3 Συμβάντα και Βελτιστοποίηση Κόστους

Τα συμβάντα (Events) αποτελούν κύριο μηχανισμό επικοινωνίας του έξυπνου συμβολαίου με το εξωτερικό περιβάλλον. Στο ενδεχόμενο καταχώρησης μιας νέας πρόβλεψης, το συμβόλαιο εκπέμπει ένα συμβάν (event), το οποίο με την σειρά του καταγράφεται στα logs της αλυσίδας, ώστε η εφαρμογή να ενημερώνει τον χρήστη, περί της κατάστασης του συστήματος, μέσω της διεπαφής σε πραγματικό χρόνο χωρίς την ανάγκη συνεχούς αναζήτησης (polling). Η βελτιστοποίηση του κόστους (Gas) αποσκοπεί στην βιωσιμότητα του συστήματος. Η χρήση ορθά οργανωμένων δομών δεδομένων, παραδείγματος χάριν mappings έναντι arrays στην αποθήκευση των hashes, ελαττώνει σημαντικά το υπολογιστικό κόστος εκτέλεσης των συναλλαγών. Επίσης,

με την μετάβαση του Ethereum στο PoS, η οικονομική και ενεργειακή επιβάρυνση μειώθηκε σημαντικά, καθιστώντας την on-chain επαλήθευση κλινικών δεδομένων προσιτή.

3.6 Αρχιτεκτονικές Διαδικτυακών Εφαρμογών

Ένα σύγχρονο Σύστημα Υποστήριξης Κλινικών Αποφάσεων, όπως αυτό που προτείνεται στην υπάρχουσα διπλωματική, οφείλει να υποστηρίζει την επεκτασιμότητα, την υψηλή διαθεσιμότητα και απόλυτη ακεραιότητα των δεδομένων. Την κάλυψη των προηγούμενων ζητημάτων έρχεται να δώσει η υβριδική αρχιτεκτονική που υιοθετεί το παρόν σύστημα, διαχωρίζοντας την αποθήκευση των ευαίσθητων δεδομένων (off-chain) από την επαλήθευση της εγκυρότητας τους (on-chain). Με αυτόν τον τρόπο, αντιμετωπίζονται περιορισμοί των παραδοσιακών κεντρικοποιημένων βάσεων δεδομένων.

3.6.1 SPA και REST API

Η υλοποίηση της διεπαφής χρήστη πραγματοποιείται με την χρήση της βιβλιοθήκης React και είναι τύπου Single Page Application (SPA). Η κύρια αιτία της συγκεκριμένης προσέγγισης είναι η διαχείριση μεγάλου όγκου δεδομένων σε πραγματικό χρόνο, ειδικότερα σε περιβάλλοντα που εστιάζουν οι τεχνολογίες Διαδικτύου των Αντικειμένων (Internet of Things) και Μεγάλων Δεδομένων (Big Data) στην υγεία. Στα περιβάλλοντα αυτά, η ταχύτητα απόκρισης επηρεάζει άμεσα την λήψη αποφάσεων. Από την πληθώρα λειτουργιών που διαθέτει η React, τα δυναμικά γραφήματα (όπως το SHAP Waterfall Plot), ενισχύουν την εμπιστοσύνη των κλινικών στο σύστημα.

Η σύνδεση μεταξύ των μοντέλων μηχανικής μάθησης και των εξωτερικών διεπαφών βασίζεται στο FastAPI. Το FastAPI επιλέχθηκε εξαιτίας της δυνατότητας που προσφέρει για ταυτόχρονη επεξεργασία αιτημάτων πρόβλεψης και παραγωγής κρυπτογραφικών αποτυπωμάτων χωρίς καθυστερήσεις.

Κεφάλαιο 4: Μεθοδολογία & Αρχιτεκτονική

Στο παρόν κεφάλαιο αναλύεται εκτενώς η αρχιτεκτονική δομή του προτεινόμενου συστήματος και η μεθοδολογία του. Η τελευταία παρουσιάζεται με κλιμακωτή προσέγγιση. Αρχικά γίνεται η στατική μοντελοποίηση του συστήματος με την χρήση διαγράμματος των συνιστωσών (component diagram) και ακολουθεί η δυναμική μοντελοποίηση με την χρήση διαγραμμάτων ροής δεδομένων. Τα τρία διαγράμματα ροής δεδομένων που χρησιμοποιούνται αποσκοπούν στη σαφή κατανόηση της δρομολόγησης πληροφορίας μεταξύ των χρηστών, των υποσυστημάτων τεχνητής νοημοσύνης και των βάσεων δεδομένων. Σε αυτό το πλαίσιο επίσης υλοποιείται η εκτός

λειτουργία του δικτύου, η ανάπτυξη του μοντέλου, καταδεικνύοντας το πλήρη κύκλο μέχρι τη παραγωγή του τελικού εκτελέσιμου προτύπου. Επιπροσθέτως, τη δυναμική μοντελοποίηση του συστήματος περιγράφουν και τρία διαγράμματα αλληλουχίας (sequence diagrams). Σκοπός τους είναι η διασαφήνιση των ροών λειτουργίας του λογισμικού, αρχικά με τη ασφαλή ταυτοποίηση και διαχείριση των ρόλων πρόσβασης (RBAC) και καταλήγοντας στο μηχανισμό επαλήθευσης και αποθήκευσης των κρίσιμων δεδομένων στο ασφαλές δίκτυο Blockchain.

4.1 Επισκόπηση Συστήματος

Το προτεινόμενο σύστημα αποτελεί μια σύγχρονη και ολοκληρωμένη πλατφόρμα υποστήριξης κλινικών αποφάσεων (Clinical Decision Support System - CDSS) με κύριο ρόλο την πρόβλεψη κινδύνου καρδιακής ανεπάρκειας. Η ιδιοτυπία του σε σύγκριση με τα συνηθισμένα κλινικά συστήματα λήψης αποφάσεων είναι ο συνδυασμός τεχνικών μηχανικής μάθησης και της τεχνολογίας του Blockchain. Ο συνδυασμός αυτός αποσκοπεί στη διεύρυνση των συνόρων της παρεχόμενης ιατρικής περίθαλψης ενώ παράλληλα εξασφαλίζεται η ασφάλεια και δυνατότητα προσδιορισμού προέλευσης των δεδομένων του ασθενούς. Η αρχιτεκτονική του συστήματος δομείται βασιζόμενη το καθιερωμένο μοντέλο τριών επιπέδων (three-tier-architecture), οριοθετώντας με σαφήνεια το σύστημα στα επίπεδα παρουσίασης (presentation layer), του επιπέδου επιχειρησιακής λογικής (business logic layer) και του επιπέδου δεδομένων (data layer).

Όπως γίνεται αντιληπτό από την **Εικόνα 20**, το σύστημα οριοθετείτε από ένα μεγάλο πλαίσιο με την ετικέτα CDSS – Heart Failure Prediction System. Τα στοιχεία που βρίσκονται εντός του πλαισίου αποτελούν τα δομικά στοιχεία της υλοποιημένης εφαρμογής. Παράλληλα, οι οντότητες του κλινικού χρήστη (clinical user) και του συνόλου δεδομένων MIMIC-IV τοποθετούνται εξωτερικά του πλαισίου. Ο λόγος διάκρισης των δύο οντοτήτων από το πλαίσιο εφαρμογής είναι η ασύγχρονη και σύγχρονη λειτουργία τους ως εξωτερικοί παράγοντες. Η οριοθέτηση αυτή ικανοποιεί το βασικό κανόνα της μηχανικής λογισμικού, χωρίζοντας διακριτά την επικοινωνία (integration points) μεταξύ του συστήματος και του εξωτερικού οικοσυστήματος.

Το γενικό πλαίσιο του συστήματος περιλαμβάνει ένα υποπλαίσιο που είναι το σύνολο υπηρεσιών ενδιάμεσου λογισμικού (Middleware Services). Οι υπηρεσίες αυτές καλύπτουν λειτουργίες και για τα τρία επίπεδα της αρχιτεκτονικής , όπως η ταυτοποίηση και εξουσιοδότηση (authentication), η καταγραφή συμβάντων και αιτημάτων (logging), η επικύρωση των πληροφοριών που δίνονται στην είσοδο (input validation) και ο έλεγχος πρόσβασης διαφορετικών πηγών (Cross Origin Resource Sharing ή CORS). Κατά τη διάρκεια υλοποίησης τους, ο μηχανισμός του middleware του FastAPI υλοποιεί διαχωρισμό των ενδιαφερόντων, εφαρμόζοντας τις εν λόγω υπηρεσίες εγκάρσιως σε όλα τα αιτήματα χωρίς παρεμβολή στο κώδικα των επιμέρους endpoints.

4.1.1 Αρχιτεκτονική και ροή εργασίας των τριών επιπέδων

Η πρώτη ενέργεια της ροής εργασίας του συστήματος είναι η διεπαφή (user interface ή UI) μεταξύ του χρήστη και του επιπέδου παρουσίασης. Το επίπεδο παρουσίασης έχει υλοποιηθεί με την γνωστή βιβλιοθήκη της Javascript, την React και τη γλώσσα Typescript. Ο τελικός χρήστης, μέσα από το αποκρίσιμο και εύχρηστο περιβάλλον αλληλεπίδρασης, μπορεί να φορτώσει τα δεδομένα των ασθενών στην πλατφόρμα, να πραγματοποιηθεί η πρόβλεψη και στη συνέχεια να επιθεωρήσει τα αποτελέσματα μαζί με τα στοιχεία καταχώρισής τους στο Blockchain. Συνεχίζοντας, το επίπεδο επιχειρησιακής λογικής πραγματοποιείται από ένα εξυπηρετητή REST API, που έχει αναπτυχθεί με το FastAPI, γραμμένο στη γλώσσα προγραμματισμού Python. Ο εξυπηρετητής διαθέτει τον ρόλο του ενορχηστρωτή στο σύστημα, γιατί αναλαμβάνει την παραλαβή και επικύρωση των δεδομένων, την προ επεξεργασία τους, την εκτέλεση του εκπαιδευμένου μοντέλου μηχανικής μάθησης για πρόβλεψη κινδύνου καρδιακής ανεπάρκειας και τέλος τη συνεχή διατήρηση της επικοινωνίας με το Blockchain για τη καταχώρηση ή ανάληψη των αποτελεσμάτων. Το τρίτο και τελευταίο επίπεδο, το επίπεδο δεδομένο, παρουσιάζει διττή φύση. Από τη μία πλευρά, η σχεσιακή βάση δεδομένων της εφαρμογής χρησιμοποιείται για την αποθήκευση των λειτουργικών δεδομένων της εφαρμογής. Η βάση δεδομένων έρχεται σε επαφή με την άλλη πλευρά του επιπέδου, το ιδιωτικό Blockchain δικτύο , το οποίο στην υπάρχουσα εργασία έχει υλοποιηθεί σε τοπικό περιβάλλον Ganache. Το Blockchain μέσω ενός έξυπνου συμβολαίου γραμμένο σε γλώσσα προγραμματισμού Solidity, έχει τη πολλαπλή λειτουργικότητα να αποθηκεύει, να κατοχυρώνει και να επαληθεύει εισερχόμενα έγγραφα σε λογική σειρά. Με αυτό τον τρόπο λειτουργεί ως μια ελεγκτική διαδρομή καταγραφής (audit trail). Ο διαχωρισμός των δύο αποθηκευτικών μέσων πραγματοποιήθηκε με σκοπό τη βέλτιστη διαχείριση των ιατρικών δεδομένων, χωρίς τη ύπαρξη χρονικών καθυστερήσεων και τη διατήρηση αξιοπιστίας και ιχνηλασιμότητας των αποτελεσμάτων πρόβλεψης.

4.1.2 Πρωτόκολλα και ροές επικοινωνίας

Ροή (1): Αλληλεπίδραση χρήστη με τη διαδικτυακή εφαρμογή

Ο χρήστης αλληλοεπιδρά με τη διαδικτυακή εφαρμογή του CDSS μέσω γραφικής διεπαφής χρήστη (graphical user interface ή GUI). Η είσοδος πραγματοποιείται από τον ανάλογο ρόλο (επομένως και τα δικαιώματα) του χρήστη αποκλειστικά μέσω προγράμματος περιήγησης (web browser), εξοικονομώντας χωρητικότητα στον αποθηκευτικό χώρο του υπολογιστή αλλά και διευκολύνοντας την ευρεία διάθεση της εφαρμογής. Από το frontend πλαίσιο της πλατφόρμας, ο χρήστης μπορεί εισάγει τα δεδομένα του ασθενή για την πρόβλεψη κινδύνου και επιθεωρεί τα αποτελέσματα που επιστρέφει.

Ροή (2): Επικοινωνία διαδικτυακής εφαρμογής με τον εξυπηρετητή

Με την ολοκλήρωση της εισαγωγής των δεδομένων, η διαδικτυακή εφαρμογή (frontend) αποστέλλει ένα αίτημα HTTP προς τον εξυπηρετητή REST. Ειδικότερα, χρησιμοποιείται η έκδοση

HTTPS (HTTP Secure), που αποτελεί την έκδοση του HTTP μηνύματος υπό κρυπτογράφηση, εξασφαλίζοντας την διακριτικότητα των μεταδιδόμενων δεδομένων μέσω του πρωτοκόλλου TLS (Transport Layer Security). Ο λόγος της επιλογής αυτής είναι πως το HTTPS αποτρέπει την υποκλοπή και αλλοίωση των δεδομένων κατά τη μεταφορά τους από τρίτους, το λεγόμενο man-in-the-middle attacks.

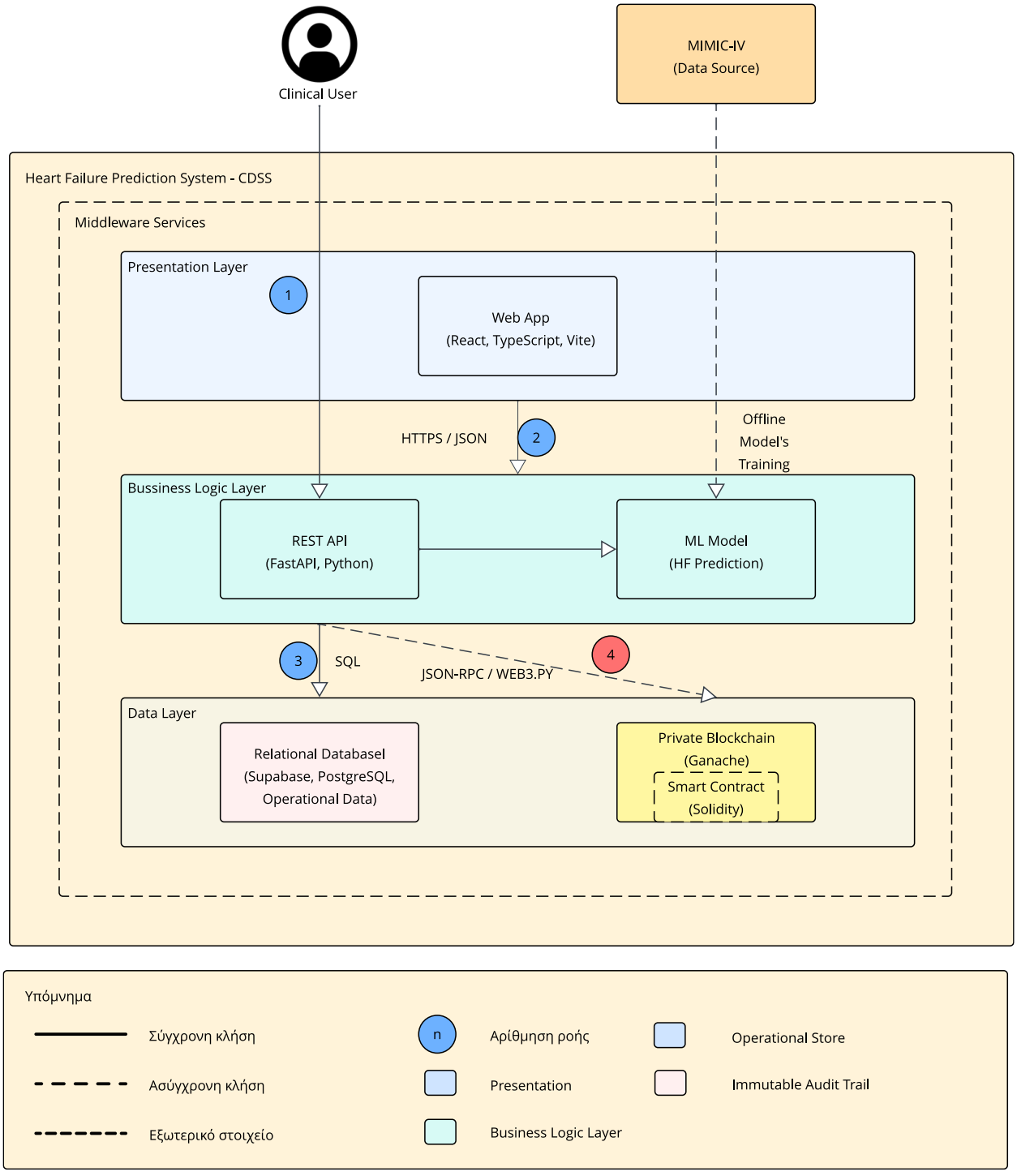
Επιπλέον η σειριοποίηση των αποκρίσεων και των αιτημάτων γίνεται σε μορφή JSON (Javascript Object Notation), η οποία είναι εύκολα αναγνωρίσιμη από τον άνθρωπο, ελαφρύ και διαδεδομένο στις σύγχρονες διεπαφές. Το JSON εμφανίζεται να επιβαρύνει λιγότερο σε εύρος από το XML και της εγγενούς συμβατότητας του με τη γλώσσα Javascript που χρησιμοποιήθηκε στην πλειοψηφία του πηγαίου κώδικα της εφαρμογής.

Ροή (3): Αποθήκευση δεδομένων στη σχεσιακή βάση

Ο εξυπηρετητής παραλαμβάνει το εγκεκριμένο αίτημα και προχωράει στη μόνιμη αποθήκευση των ανάλογων δεδομένων. Η αποθήκευση πραγματοποιείται στη σχεσιακή βάση Supabase, η οποία είναι γραμμένη στη γλώσσα προγραμματισμού PostgreSQL, μέσω σύγχρονων κλήσεων SQL. Ο όρος «σύγχρονη κλήση» αναφέρεται στο γεγονός που η ροή εκτέλεσης μπλοκάρεται μέχρι να επιβεβαιωθεί η αποθήκευση του προηγούμενου βήματος εκτέλεσης. Η επιλογή σύγχρονης επικοινωνίας ήταν αναγκαία διότι σε ένα περιβάλλον κρίσιμης υποδομής η διαθεσιμότητα των δεδομένων είναι σημαντική. Η σχεσιακή βάση υποστηρίζει συναλλαγές που έχουν πρότυπο το ACID (Atomicity, Consistency, Isolation, Durability), οι οποίες εγγυώνται πως η συνέπεια των δεδομένων θα διατηρηθεί ακάθεκτη, καθώς απαιτείται οι λειτουργίες εγγραφής να είναι ολοκληρωμένες στο σύνολο τους.

Ροή (4): Καταχώριση τελικού αποτελέσματος στο Blockchain

Η ροή 4 συμβαίνει σε παράλληλο χρόνο με τη ροή 3 φανερώνοντας όμως διαφορές που αξίζουν ιδιαίτερη ανάλυση. Την αποθήκευση στη σχεσιακή βάση δεδομένων ακολουθεί η καταχώριση της πρόβλεψης στο Blockchain. Η επικοινωνία με τον κόμβο Ethereum του Ganache πραγματοποιείται μέσω του πρωτοκόλλου JSON-RPC 2.0 πάνω από το HTTP. Μέσω του JSON-RPC 2.0 ο κάθε κόμβος Ethereum θέτει ένα σύνολο από RPC μεθόδους (eth_sendTransaction, eth_call, eth_getTransactionReceipt, κτλ.), που επιτρέπουν σε εξωτερικές εφαρμογές να αλληλοεπιδρούν με την αλυσίδα του Blockchain. Ο ρόλος της Web3.py, με το υψηλού επιπέδου API της ενθυλακώνει τις κλήσεις JSON-RPC, αυτοματοποιεί την ABI κωδικοποίηση, προσφέροντας βολικές συναρτήσεις για συναλλαγές. Ακόμη, μόλις η συναλλαγή αποσταλεί στο κόμβο, τοποθετείται σε ουρά αναμονής (mempool) και αναμένει να ενσωματωθεί σε νέο μπλοκ. Με λίγα λόγια εφαρμόζεται το μοντέλο fire-and-forget.



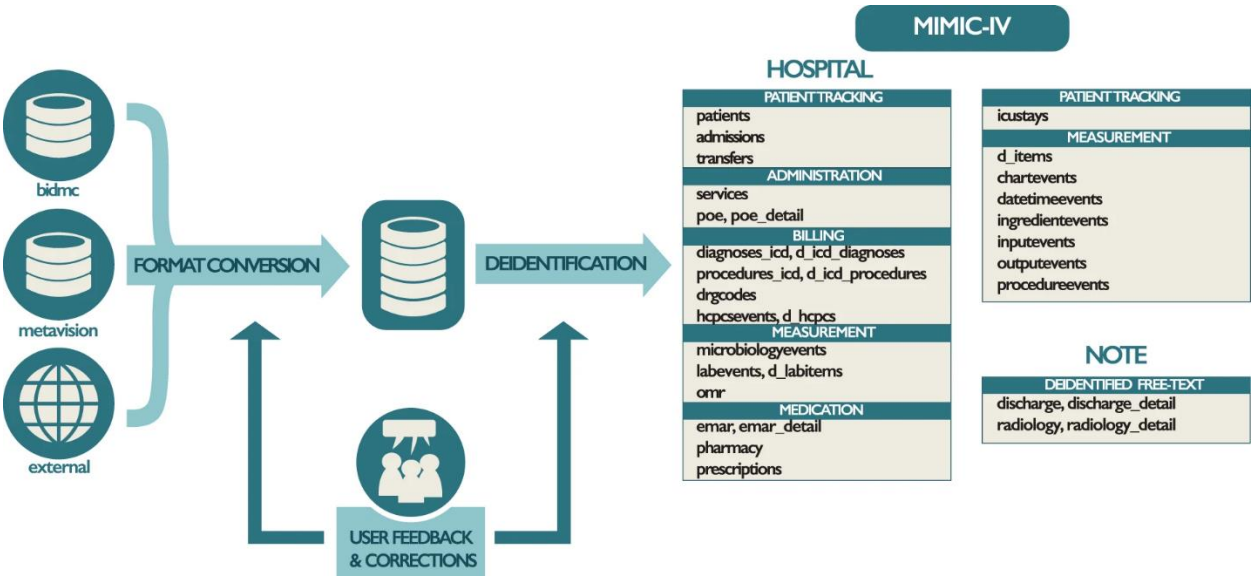
Εικόνα 20 : Αρχιτεκτονική τριών επιπέδων του συστήματος υποστήριξης κλινικής απόφασης (CDSS) για τη πρόβλεψη της καρδιακής ανεπάρκειας. Παρουσιάζονται διακριτά τα τρία επίπεδα της αρχιτεκτονικής (Presentation, Business Logic, Data) και οι υπηρεσίες ενδιάμεσου λογισμικού.

4.2 Σύνολο Δεδομένων (MIMIC-IV)

Η αποτελεσματικότητα των μοντέλων μηχανικής μάθησης στο ιατρικό περιβάλλον βασίζεται στην ποιότητα της εκπαίδευσης που έχουν λάβει. Πρακτικά, αυτό σημαίνει πως το μοντέλο πρέπει να έχει τροφοδοτηθεί με μεγάλο όγκο δεδομένων που θα είναι καθαρό και απαλλαγμένο από μεροληψίες, με υψηλή σχετικότητα και σωστή δημογραφική εκπροσώπηση. Για τον λόγο αυτό, στη παρούσα διπλωματική εργασία επιλέχθηκε μια από τις πιο αναγνωρισμένες και χρησιμοποιούμενες βάσεις κλινικών δεδομένων, η MIMIC-IV (Medical Information Mart for Intensive Care, 4^{η} έκδοση). Στην συνέχεια του κεφαλαίου αναλύονται η προέλευση και τα ηθικά ζητήματα που προκύπτουν, η δομή των πινάκων που αξιοποιήθηκε και τα βήματα ορισμού της μεταβλητής στόχου label για την βέλτιστη εκπαίδευση του μοντέλου.

4.2.1 Πηγή Δεδομένων Εκπαίδευσης και Ηθική Ζητήματα

Η βάση δεδομένων MIMIC-IV αποτελεί το αποτέλεσμα της ερευνητικής συνεργασίας μεταξύ του ιατρικού κέντρου Beth Israel Deaconess (BIDMC) της Βοστόνης και του εργαστηρίου υπολογιστικής φυσιολογίας του Ινστιτούτου Τεχνολογίας της Μασαχουσέτης (MIT). Το MIMIC ξεκίνησε ως μια βάση δεδομένων κυματομορφών που περιλαμβάνει τα δημογραφικά στοιχεία για πάνω από 90 ασθενείς. Στην εξέλιξη της το MIMIC-II και MIMIC-III ακολούθησαν με σημαντικά αυξημένο μέγεθος δειγμάτων και εύρος πληροφοριών, εντάσσοντας και πίνακες SQL για την ευκολότερη επεξεργασία και στατιστική ανάλυση των δεδομένων. Στη συγκεκριμένη διπλωματική εργασία χρησιμοποιήθηκε η έκδοση MIMIC-IV v3.1, η οποία είναι από τις τελευταίες ενημερωμένες έκδοσης της σειράς και περιέχει αποιδιοποιημένα κλινικά δεδομένα για πάνω από 364.000 ασθενείς που νοσηλεύτηκαν στο BIDMC την χρονική περίοδο 2008-2022. Η συλλογή των δεδομένων πραγματοποιήθηκε από γενικά τμήματα νοσηλείας του νοσοκομείου και μονάδες εντατικής θεραπείας (ΜΕΘ), καλύπτοντας ένα ευρύ πλαίσιο ιατρικών περιστατικών.



Εικόνα 21 : Επισκόπηση ανάπτυξης της ιατρικής βάσης δεδομένων MIMIC. Τα δεδομένα συλλέγονται από το αποθετήριο δεδομένων του BIDMC, το σύστημα πληροφοριών της ΜΕΘ (Metavision) και εξωτερικές πηγές (external) και μετατρέπονται σε μια κοινή ενοποιημένη μορφή. Κατά την αποιδιοποίηση (deidentification) αφαιρούνται όλα τα προσωπικά δεδομένα που μπορούν να ταυτοποιήσουν έναν ασθενή. Η τελική οργάνωση σε πίνακες του MIMIC χωρίζεται σε δύο μεγάλα υποσύνολα, το HOSPITAL και το ICU, ενώ το υποσύνολο NOTE περιέχει σημειώσεις σχετικά με τους ασθενείς.

Η πρόσβαση στο σύνολο δεδομένων MIMIC-IV διέπετέ από αυστηρούς κανόνες δεοντολογίας. Ειδικότερα, η πρόσβαση παραχωρείται μέσω της πλατφόρμας PhysioNet, ενός έμπιστου αποθετηρίου ιατρικών δεδομένων που φιλοξενεί το MIT. Ύστερα από τη δημιουργία λογαριασμού στη πλατφόρμα, ο χρήστης επιβάλλεται να έχει λάβει επιτυχώς την ολοκλήρωση του εκπαιδευτικού προγράμματος CITI προστασίας ανθρωπίνων υποκειμένων στην έρευνα. Με την συμφωνία χρήσης δεδομένων (Data Use Agreement ή DUA), δεσμεύει στην αποκλειστική χρήση δεδομένων για ερευνητικούς σκοπούς. Συγχρόνως, τα δεδομένα του MIMIC έχουν υποστεί αποταυτοποίηση, σύμφωνα με τα πρότυπα του Health Insurance Portability and Accountability Act (HIPAA) των Ηνωμένων Πολιτειών. Συγκεκριμένα, έχει γίνει η αφαίρεση 18 συγκεκριμένων αναγνωριστικών στοιχείων, όπως τα ονόματα, οι διευθύνσεις και τα τηλέφωνα, για την διασφάλιση της εμπιστευτικότητας.

4.2.2 Αξιοποιημένοι πίνακες

Προηγουμένως αναφέρθηκε πως τα δύο βασικά υποσύνολα της βάσης είναι το MIMIC-IV Hosp, περιέχει δεδομένα από τη χρονική διάρκεια νοσηλείας ενός ασθενή στο νοσοκομείο και το MIMIC-IV ICU, που εστιάζει σε πιο λεπτομερή δεδομένα κατά την νοσηλεία ενός ασθενή στις μονάδες εντατικής θεραπείας. Για τις ανάγκες της εργασίας αξιοποιήθηκαν πέντε βασικοί πίνακες στην ενότητα hosp του MIMIC, διότι διαθέτει ευρύτητα πληθυσμού και ενέχει τους πίνακες με τα διαγνωστικά κωδικοποιημένα δεδομένα για τον ορισμό της μεταβλητής στόχου. Οι πίνακες που αξιοποιήθηκαν είναι οι εξής:

1) **Πίνακας Ασθενών (patients):** ο πίνακας εμπεριέχει τα βασικά δημογραφικά στοιχεία κάθε ασθενούς, όπως το φύλο, η ημερομηνία γέννησης (anchor_age) ανά μοναδική αναγνωριστικό (subject_id). Η μεγαλύτερη συσχέτιση με το κίνδυνο εμφάνισης της ΚΑ οφείλεται στα θεμελιώδη χαρακτηριστικά της ηλικίας και του φύλου.

2) **Πίνακας Εισαγωγών (admissions):** ο πίνακας περιέχει μια εγγραφή για κάθε νοσηλεία ενός ασθενή. Προφανώς, ένας ασθενής μπορεί να παρουσιάζει περισσότερες από μία εγγραφές αν έχει νοσηλευτεί πολλαπλές φορές. Τα βασικότερα πεδία του πίνακα είναι το μοναδικό αναγνωριστικό νοσηλείας (hadm_id), οι χρόνοι εισαγωγής και εξιτηρίου (admittime και dischtime), ο δείκτης επιβίωσης του ασθενή κατά τη διάρκεια νοσηλείας του (hospital_expire_flag).

3) **Πίνακας Διαγνώσεων (diagnoses_icd):** ο πίνακας είναι κρίσιμος για τον ορισμό της μεταβλητής στόχου, καθώς περιλαμβάνει όλους τους διαγνωστικούς κωδικούς που έχουν αποδοθεί σε κάθε νοσηλεία. Η κάθε εγγραφή συνδέει μια νοσηλεία (hadm_id) με έναν διαγνωστικό κωδικό (icd_code) και δείχνει την έκδοση κωδικοποίησης μέσω του icd_version, χρησιμοποιώντας κωδικοποίηση ICD-9 ή ICD-10, αναλόγως το πρότυπο ICD (International Classification of Diseases) που εφαρμόστηκε κατά τη νοσηλεία του. Η μετάβαση από το ICD-9 στο ICD-10 πραγματοποιήθηκε το 2015, επομένως ο πίνακας περιέχει κωδικοποίησης και για τα δύο.

4) **Πίνακας Εργαστηριακών Εκδηλώσεων (labevents):** Ο πίνακας αποτελεί το ιστορικό των αποτελεσμάτων των εργαστηριακών εξετάσεων κάθε ασθενή που διεκπεραιώθηκαν κατά τη νοσηλεία του. Αποτελεί τον μεγαλύτερο σε μέγεθος πίνακα της βάσης, καθώς περιέχει εκατομμύρια εγγραφές. Ο συγκεκριμένος πίνακας αποτελεί τη πηγή εξαγωγής εργαστηριακών χαρακτηριστικών (lab features) που χρησιμοποιήθηκαν ως είσοδοι στο μοντέλο πρόβλεψης, όπως η κρεατινίνη, το νάτριο, το κάλιο, η αιμοσφαιρίνη και οποιασδήποτε άλλος βιοδείκτης που σχετίζεται με τη καρδιακή λειτουργία.

5) **Πίνακας d_labitems:** ο πίνακας λειτουργεί ως λεξικό των labevents, παρέχοντας τη περιγραφή και το κάθε κωδικό εργαστηριακής εξέτασης (itemid) και τη κατηγορία της (category). Ακολουθεί το πρότυπο LOINC (Logical Observation Identifiers Names and Codes), που έχει διεθνή ισχύ για τη κωδικοποίηση των εργαστηριακών παρατηρήσεων.

Η ένωση των πινάκων πραγματοποιείται μέσω της σχεσιακής σύνδεσης join στα χαρακτηριστικά subject_id και hadm_id, ενώ η όλη επεξεργασία, φόρτωση, σύνδεση και μετασχηματισμός των δεδομένων βασίστηκε στη βιβλιοθήκη pandas της γλώσσας Python.

4.2.3 Μεταβλητή Στόχου (label)

Ο ορισμός της μεταβλητής στόχου επικεντρώνεται στη διάκριση κάθε νοσηλείας, σε θετικής ή αρνητικής εμφάνισης κινδύνου καρδιακής ανεπάρκειας. Η καταχώριση των ασθενών ως πάσχοντες ΚΑ εκτελέστηκε σε επίπεδο νοσηλείας και όχι το επίπεδο ασθενούς, ακολουθώντας τα διεθνή πρότυπα διάγνωσης ICD που υπάρχουν στο πίνακα diagnoses_icd. Εξαιτίας της αλλαγής του προτύπου από το ICD-9 στο ICD-10, για να εξασφαλιστεί η συστηματική συνοχή της μεταβλητής στόχου χρονικά καλύφτηκε χρησιμοποιώντας συγχρόνως και τις δύο κωδικοποιήσεις για την αποφυγή απώλειας δεδομένων. Κατά τον ορισμό της μεταβλητής στόχου συμπεριλήφθηκαν τόσο οι βασικοί κωδικοί καρδιακής ανεπάρκειας όσο και η αιτιολογία υπερτασικής μορφολογίας της νόσου, καλύπτοντας το πλήρες φάσμα των κλινικών περιπτώσεων. Αυτό επιτεύχθηκε με την ένταξη των κωδικών I11.0, I13.0, I13.2, 402.x, 404.x. Αξίζει να σημειωθεί ότι η αντιστοίχιση των κωδικών καλύπτει πλήρως τους επίσημους χάρτες μετάβασης General Equivalence Mappings (GEMs) που εκδίδονται από το Center for Medicare & Medicaid Services (CMS) των Ηνωμένων Πολιτειών, τηρώντας τη σημασιολογική συνέπεια μεταξύ των δύο εκδόσεων.

Πίνακας 4.1 : Αναλυτικός πίνακας των κωδικών ICD-10 που χρησιμοποιήθηκαν στη μεταβλητή στόχου

Κωδικός ICD	Κλινική Περιγραφή
I50.1	Αριστερή κοιλιακή ανεπάρκεια
I50.20	Συστολική καρδιακή ανεπάρκεια (HFrEF) μη καθορισμένης φάσης
I50.21	Οξεία συστολική καρδιακή ανεπάρκεια (HFrEF)
I50.22	Χρόνια συστολική καρδιακή ανεπάρκεια (HFrEF)
I50.23	Οξεία χρόνια συστολική καρδιακή ανεπάρκεια (HFrEF)
I50.30	Διαστολική καρδιακή ανεπάρκεια (HFpEF) μη καθορισμένης φάσης
I50.31	Οξεία διαστολική καρδιακή ανεπάρκεια (HFpEF)
I50.32	Χρόνια διαστολική καρδιακή ανεπάρκεια (HFpEF)
I50.33	Οξεία χρόνια διαστολική καρδιακή ανεπάρκεια (HFpEF)
I50.40	Συνδυασμός συστολική και διαστολικής καρδιακής

	ανεπάρκειας χωρίς καθορισμένη φάση
I50.41	Οξεία συνδυασμένη συστολική και διαστολική καρδιακή ανεπάρκεια
I50.42	Χρόνια συνδυασμένη συστολική και διαστολική καρδιακή ανεπάρκεια
I50.43	Χρόνια οξεία συνδυασμένη συστολική και καρδιακή ανεπάρκεια
I50.8	Καθορισμένες μορφές καρδιακής ανεπάρκειας
I50.9	Μη καθορισμένη καρδιακή ανεπάρκεια
I11.0	Υπερτασική καρδιοπάθεια με καρδιακή ανεπάρκεια
I13.0	Υπερτασική καρδιακή και χρόνια νεφρική νόσος με καρδιακή ανεπάρκεια (ΧΝΝ 1-4 σταδίου ή καθορισμένη)
I13.2	Υπερτασική καρδιακή και χρόνια νεφρική νόσος με καρδιακή ανεπάρκεια (ΧΝΝ 5 σταδίου ή τελικού σταδίου)

Πίνακας 4.2 : Αναλυτικός πίνακας των κωδικών ICD-9 που χρησιμοποιήθηκαν στη μεταβλητή στόχου

Κωδικός ICD	Κλινική Περιγραφή
428.0	Συμφορητική καρδιακή ανεπάρκεια μη καθορισμένη
428.1	Αριστερή καρδιακή ανεπάρκεια
428.20	Συστολική καρδιακή ανεπάρκεια μη καθορισμένης φάσης
428.21	Οξεία συστολική καρδιακή ανεπάρκεια
428.22	Χρόνια συστολική καρδιακή ανεπαρκεια
428.23	Οξεία και χρόνια συστολική καρδιακή ανεπάρκεια
428.30	Διαστολική καρδιακή ανεπάρκεια μη καθορισμένης φάσης
428.31	Οξεία διαστολική καρδιακή ανεπάρκεια
428.32	Χρόνια διαστολική καρδιακή ανεπάρκεια
428.33	Οξεία και χρόνια διαστολική καρδιακή ανεπάρκεια

428.40	Συνδυασμός συστολικής και διαστολικής καρδιακής ανεπάρκειας μη καθορισμένης φάσης
428.41	Συνδυασμός οξείας συστολικής και διαστολικής καρδιακής ανεπάρκειας
428.42	Συνδυασμός χρόνιας συστολικής και διαστολικής καρδιακής ανεπάρκειας
428.43	Οξεία και χρόνια συστολική και διαστολική ανεπάρκεια
428.9	Καρδιακή ανεπάρκεια μη καθορισμένης φάσης
402.0	Κακοήθης υπερτασική καρδιοπάθεια με καρδιακή ανεπάρκεια
402.1	Καλοήθης υπερτασική καρδιοπάθεια με καρδιακή ανεπάρκεια
402.9	Μη καθορισμένη υπερτασική καρδιακή και νεφρική νόσος με καρδιακή ανεπάρκεια
404.0	Κακοήθης υπερτασική καρδιακή και νεφρική νόσος με καρδιακή ανεπάρκεια
404.1	Καλοήθης υπερτασική καρδιακή και νεφρική νόσος με καρδιακή ανεπάρκεια
404.9	Μη καθορισμένη υπερτασική καρδιακή και νεφρική νόσος με καρδιακή ανεπάρκεια

Η αντιστοίχιση νοσηλειών της μεταβλητής στόχου υλοποιήθηκε με τη μέθοδο των φαινοτύπων. Αρχικά, διαπιστώνεται η έκδοση του κωδικού μέσω του πεδίου icd_version για κάθε διάγνωση από τον πίνακα diagnoses_icd. Εφόσον η διάγνωση πληρή ένα από τα κριτήρια που έχουμε ορίσει, σημαίνει ότι είναι θετική στην εμφάνιση κινδύνου πρόκλησης καρδιακής ανεπάρκειας.

Αξίζει να σημειωθεί ότι αυτή η μέθοδος ορισμού της μεταβλητής στόχου εισάγει ορισμένους εγγενείς περιορισμούς. Πρώτον, βασίζεται αποκλειστικά στην κλινική κωδικοποίηση, η οποία ενδέχεται να εμπεριέχει σφάλματα ή παραλείψεις, ιδιαίτερα σε περιπτώσεις ήπιας ή πρώιμης νόσου που δεν έχουν ακόμη επίσημα διαγνωστεί. Δεύτερον, επειδή οι κωδικοί ICD στο αμερικανικό σύστημα υγείας προέρχονται κυρίως από διαδικασίες χρέωσης ασφαλιστικών εταιρειών, ενδέχεται να παρουσιάζουν συστηματικές μεροληψίες (systematic biases) υπέρ ή κατά ορισμένων διαγνώσεων, ανάλογα με τις πολιτικές αποζημίωσης. Τρίτον, δεν λαμβάνεται υπόψη η χρονική σειρά εμφάνισης της νόσου σε σχέση με τις εργαστηριακές μετρήσεις που χρησιμοποιούνται ως είσοδοι του μοντέλου, γεγονός που μπορεί δυνητικά να εισαγάγει στοιχεία διαρροής δεδομένων αν δεν αντιμετωπιστεί κατάλληλα κατά τη φάση της προ επεξεργασίας. Τα ζητήματα αυτά αντιμετωπίστηκαν με τη χρήση κατάλληλων χρονικών παραθύρων κατά την κατασκευή των χαρακτηριστικών εισόδου, όπως αναλύεται στο επόμενο κεφάλαιο.

4.3 Αρχιτεκτονική Συστήματος

4.3.1 Διάγραμμα Συνιστωσών

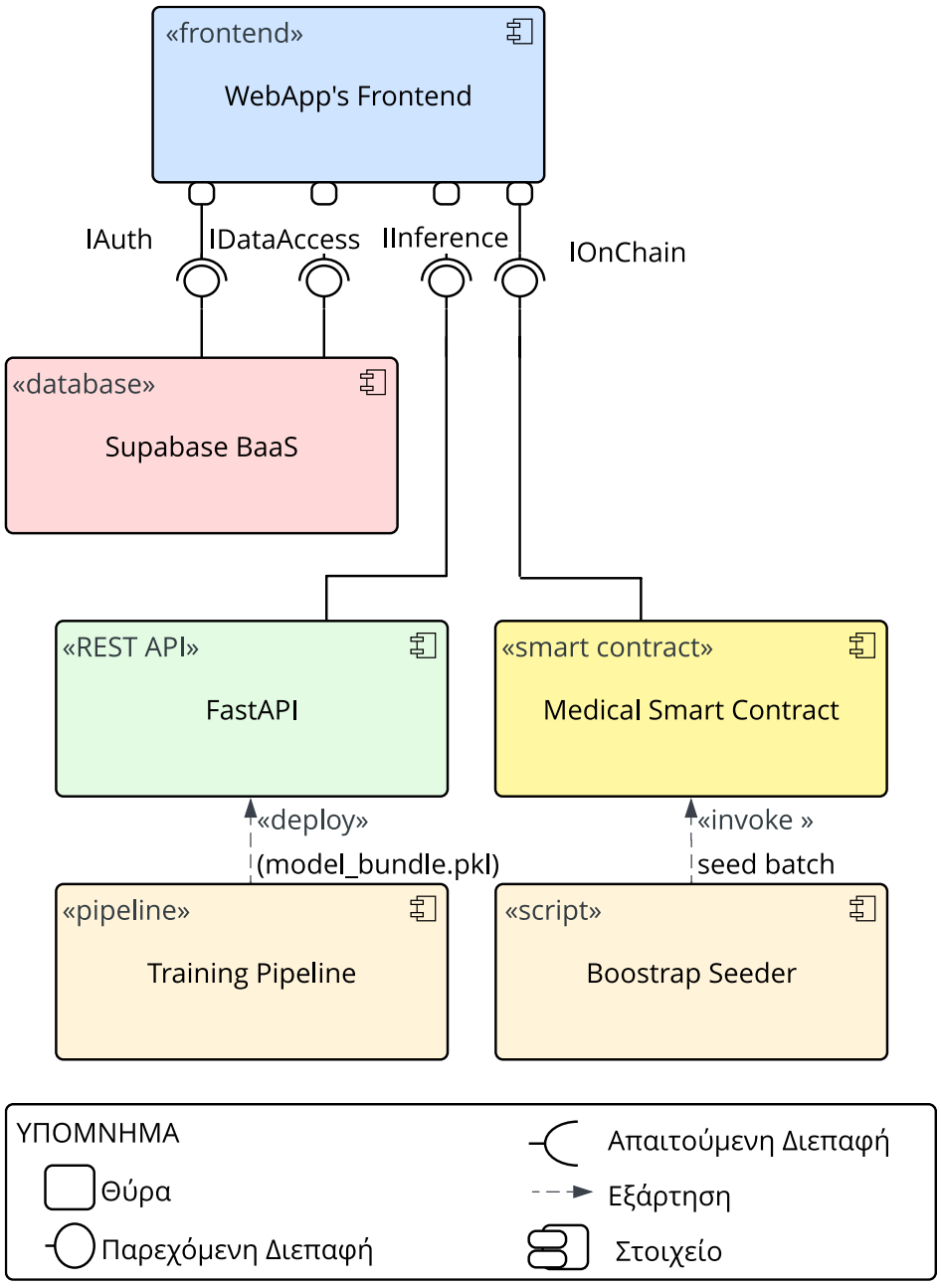
Η επιλογή του διαγράμματος συνιστωσών έναντι του διαγράμματος κλάσεων πραγματοποιήθηκε με σκοπό να στραφεί η προσοχή του αναγνώστη στη φυσική οργάνωση του κώδικα και την αλληλεπίδραση των εξαρτημάτων (components) σύνθεσης της εφαρμογής μεταξύ των διεπαφών. Στην Εικόνα 22 παρουσιάζεται το διάγραμμα συνιστωσών του λειτουργικού προτύπου Diagnosis. Η αρχιτεκτονική του προτύπου οργανώνεται σε τρία επίπεδα, διατεταγμένα κατακόρυφα προς τη βάση. Αρχικά, το ανώτερο επίπεδο αποτελεί τον μοναδικό πελάτη (client) του συστήματος, το συστατικό «WebApp's Frontend» με στερεότυπο «frontend». Το αμέσος κατώτερο επίπεδο χωρίζεται σε τρείς ανεξάρτητες συνιστώσες του backend κατά την εκτέλεση που εξυπηρετούν του πελάτη, οι συνιστώσες «Supabase BaaS», το «FastAPI» και το «Medical Smart Contract», με τα αντίστοιχα στερεότυπα «database», «REST API», «smart contract». Στο τελευταίο επίπεδο τοποθετούνται δύο εξαρτήματα που λειτουργούν εκτός σύνδεσης με την εφαρμογή: το «Training Pipeline» με στερεότυπο «pipeline» και το «Bootstrap Seeder» με στερεότυπο «script». Η επιλογή της κατακόρυφης διάταξης δεν είναι τυχαία καθώς απεικονίζει τη ροή λειτουργίας της εφαρμογής.

Την κορυφή του διαγράμματος διακατέχει ο μοναδικός αιτών για υπηρεσίες του συστήματος, τη συνιστώσα «WebApp's Frontend», το οποίο εκτελείται πάνω στο πρόγραμμα περιήγησης με τη μορφή μονοσέλιδης εφαρμογής (Single Page Application). Ο κώδικας της εφαρμογής έχει γραφτεί με χρήση της βιβλιοθήκης React, την γλώσσα Typescript ενώ η διανομή της εφαρμογής γίνεται μέσω του εργαλείου Vite. Ιδιαίτερα σημαντική παρατήρηση κρίνεται ο διαχωρισμός των απαιτούμενων διεπαφών που αναπαρίστανται ως ημικύκλια συνδεδεμένα μέσω θυρών στο κάτω μέρος της συνιστώσας. Ο λόγος που επιλέχθηκε η συγκεκριμένη αρχιτεκτονική, που εφαρμόζει την Αρχή Διαχωρισμού Διεπαφών (Interface Segregation Principle), είναι η αμεταβλητότητα και σταθερότητα κατάστασης της πλατφόρμας σε ενδεχόμενο τροποποίησης κάποιας διεπαφής. Η πρώτη απαιτούμενη διεπαφή, η IAuth, αντιπροσωπεύει το σύνολο λειτουργιών αυθεντικοποίησης και διαχείρισης συνεδρίας του χρήστη. Σε περιβάλλον κώδικα η διεπαφή χρησιμοποιεί το React hook useAuth, που εκθέτει τις βασικές μεθόδους σύνδεσης και εγγραφής signUp, signIn, signOut και τη μεταβλητή user, που λειτουργεί ως ψηφιακή ταυτότητα του χρήστη. Για παράδειγμα, ο γιατρός συνδέεται με το ρόλο «doctor» και η μεταβλητή ενημερώνει κάθε συνιστώσα για τον χρήστη και τις επιτρεπόμενες δυνατότητες που έχει στο σύστημα. Η δεύτερη απαιτούμενη διεπαφή, η IDataAccess εκφράζει το σύνολο λειτουργιών CRUD των πινάκων της βάσης δεδομένων Supabase του συστήματος (πίνακες όπως patients, profiles, user_roles) και υλοποιείται μέσω του hook usePatients. Ως πρωταρχικός ρόλος της ορίζεται η αφαίρεση της πολυπλοκότητας διαχείρισης των δεδομένων από τις οπτικές συνιστώσες της εφαρμογής και προσδίδει επεκτασιμότητα στο σύστημα, γιατί σε περίπτωση αλλαγής της Supabase βάσης με Firebase ή MongoDB, το σύστημα παραμένει ως έχει. Η τρίτη απαιτούμενη διεπαφή, η IInference αποκτά τον ρόλο του σημείου διασύνδεσης του frontend και του υποσυστήματος τεχνητής νοημοσύνης (main.py). Ουσιαστικά ενθυλακώνει τη κλήση των δύο endpoints του FastAPI (/api/score,

/api/score-json) και χρησιμοποιεί μόνο το CsvUploader μέσω fetch API του προγράμματος περιήγησης. Η τέταρτη και τελευταία διεπαφή, η IOnChain, αντιπροσωπεύει το σύνολο των λειτουργιών (setPatientConsent, publishRiskScore, publishBatch, getHistory, verifyPrediction) που γίνονται on-chain του έξυπνου συμβολαίου και χρησιμοποιείται μέσω της βιβλιοθήκης ethers.js και του ηλεκτρονικού πορτοφολιού κρυπτονομισμάτων Metamask.

Στο μεσαίο επίπεδο του διαγράμματος διακρίνονται τρεις ανεξάρτητες συνιστώσες που ειδικεύονται σε διαφορετικό τομέα και αποτελούν το backend κομμάτι του συστήματος. Αρχικά, η συνιστώσα «Supabase BaaS» εμφανίζει στη κορυφή της δύο διακριτές παρεχόμενες διεπαφές, τις IAuth και IDataAccess που χρησιμοποιούνται στη ταυτοποίηση και πρόσβαση το πελάτη στη βάση δεδομένων. Με την παροχή δύο διεπαφών αντικατοπτρίζει τη φύση της συνιστώσας ως Backend-as-a-Service, η οποία συνδυάζει τις υπηρεσίες της αυθεντικοποίησης (supabase auth) με τη βάση δεδομένων γραμμένη σε PostgreSQL και των αυτόματων παραγόμενων RESTful endpoints (PostgREST). Δεύτερον, η συνιστώσα «FastAPI» εκφράζει τον εξυπηρετητή Python και υλοποιείται με το αρχείο api.py. Στην παρούσα εργασία, η «FastAPI» δεν αναλαμβάνει καμία άλλη υπηρεσία πέραν την ενσωμάτωση του μοντέλου μηχανικής μάθησης, στηριζόμενη στην αρχή της απομόνωσης των ευθυνών (Separation of Concerns), διευκολύνοντας σημαντικά τη συντήρηση, τη κλιμακωσιμότητα και τη δοκιμή του συστήματος. Τρίτον, η συνιστώσα «Medical Smart Contract» παρέχει την παρεχόμενη διεπαφή IOnChain και αναπαριστά το έξυπνο συμβόλαιο που έχει αναπτυχθεί στο ιδιωτικό blockchain Ethereum του τοπικού παρόχου Ganache. Το έξυπνο συμβόλαιο λειτουργεί με δικό του κύκλο ζωής, δικά του δικαιώματα πρόσβασης μέσω της βιβλιοθήκης OpenZeppelin και δική του κατάσταση που διατηρείται αμετάβλητη στο blockchain.

Κάτω από τις συνιστώσες του μεσαίου επιπέδου παρουσιάζονται δύο τύπου εκτός λειτουργίας συνιστώσες που είναι απαραίτητες στην προετοιμασία των τεχνουργημάτων (artifacts) των υπηρεσιών χρόνου εκτέλεσης (runtime services). Το Training Pipeline υλοποιείται μέσω του script main.py, το οποίο με τη σειρά του έχει αναλάβει την ανάγνωση των δεδομένων από το MIMIC-IV, τον καθαρισμό, την προ επεξεργασία, την εκπαίδευση, την βαθμονόμηση πολλαπλών μοντέλων και την επιλογή του βέλτιστου μοντέλου βάσει μετρικών αξιολόγησης. Από την άλλη πλευρά, ο Bootstrap Seeder, υλοποιείται μέσω του sub_storage.py script. Η κύρια λειτουργία του είναι η παραγωγή η μαζική εγγραφή των αποτελεσμάτων που έχουν παραχθεί από το training pipeline, το αρχείο blockchain_payload.json, στο blockchain. Ο διαχωρισμός του seeding, που είναι λειτουργία χαμηλής συχνότητας, ως ξεχωριστή συνιστώσα λειτουργίας δείχνει την απομόνωση από τυχόν αιτήματα των χρηστών που αποτελούν λειτουργίες υψηλής συχνότητας.

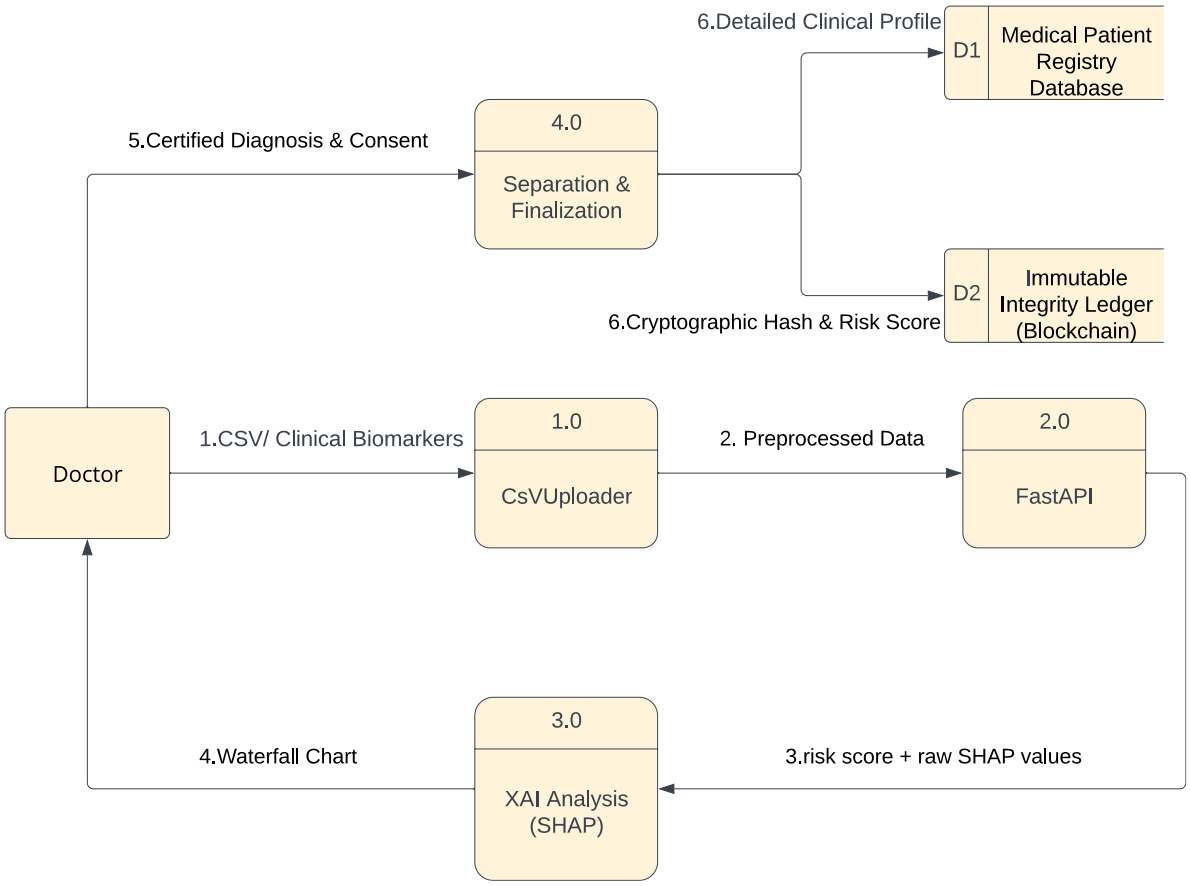


Εικόνα 22 : Διάγραμμα Συνιστωσών του CDSS. Αλληλεπίδραση μεταξύ των επιπέδων διεπαφής, του συστήματος πρόβλεψης και του αποκεντρωμένου δικτύου Blockchain

4.3.2 Διαγράμματα Ροής Δεδομένων

4.3.2.1 Διάγραμμα Ροής Δεδομένων με ρόλο χρήστη τον «Γιατρό»

Στην **Εικόνα 23** αποτυπώνεται το διάγραμμα ροής δεδομένων επιπέδου 1 (DFD Level 1). Το πρώτο διάγραμμα ροής δεδομένων αναλύει τη δυναμική αλληλεπίδραση του κλινικού ιατρού με το σύστημα και εστιάζει στη ενορχήστρωση της πληροφορίας από την είσοδο της έως τη τελική αποθήκευση στη βάση δεδομένων. Η κυκλική ροή λειτουργίας στο διάγραμμα φανερώνει πως η τεχνητή νοημοσύνη λειτουργεί επικουρικά προς τον κλινικό ιατρό (human-in-the-loop), εξασφαλίζοντας τη διαφάνεια και επαληθευσιμότητα σε κάθε στάδιο του. Το πρώτο βήμα, μετά την αυθεντικοποίηση του χρήστη ως ιατρός, είναι η μεταφόρτωση αρχείο CSV ή η συμπλήρωση των κλινικών βιοδεικτών σε μορφή ηλεκτρονικής φόρμας. Μετά τη διεπαφή μεταφόρτωσης CsvUploader τα δεδομένα υφίστανται καθαρισμό από ελλιπές και ακραίες τιμές, εξαλείφοντας τον θόρυβο και διορθώνοντας τυχόν σφάλματα κωδικοποίησης και κανονικοποίηση ώστε όλες οι τιμές να βρίσκονται στην ίδια κλίμακα και να αποφευχθεί η ανισορροπία των κλάσεων. Ύστερα τα δεδομένα προωθούνται στον εξυπηρετητή FastAPI, ο οποίος λειτουργεί ως διαμεσολαβητής που φορτώνει το εκπαιδευμένο μοντέλο (model_bundle.pkl), για να χρησιμοποιηθεί στη πρόβλεψη. Από εκεί παράγεται το σκορ κινδύνου και οι ακατέργαστές τιμές SHAP και η συνιστώσα 3.0 αναλαμβάνει τη μορφοποίηση τους σε εποπτικά διαγράμματα Waterfall. Η οπτικοποίηση επιτρέπει στο κλινικό προσωπικό να καταλάβει τη βαρύτητα κάθε βιοδείκτη. Ο γιατρός, λοιπόν, εκδίδει τη πιστοποιημένη διάγνωση των διαγραμμάτων και με την συγκατάθεση του ασθενούς, για τη θωράκιση της διαδικασίας σύμφωνα με τον GDPR, συνεχίζει στην αποθήκευση του πλήρους κλινικού προφίλ σε μια σχεσιακή βάση δεδομένων (D1 Medical Patient Registry) και το κρυπτογραφικό αποτύπωμα (hash) τους με το σκορ κινδύνου καταχωρούνται στο Blockchain ledger (D2 Immutable Integrity Ledger). Ο διαχωρισμός αυτός αποθήκευσης των δεδομένων διασφαλίζει τη αμεταβλητότητα και ιχνηλασιμότητα των ευαίσθητων προσωπικών δεδομένων πάνω στην αλυσίδα (on-chain), ενώ παράλληλα διασφαλίζεται η άμεση διαθεσιμότητα των δεδομένων για παρακολούθηση.

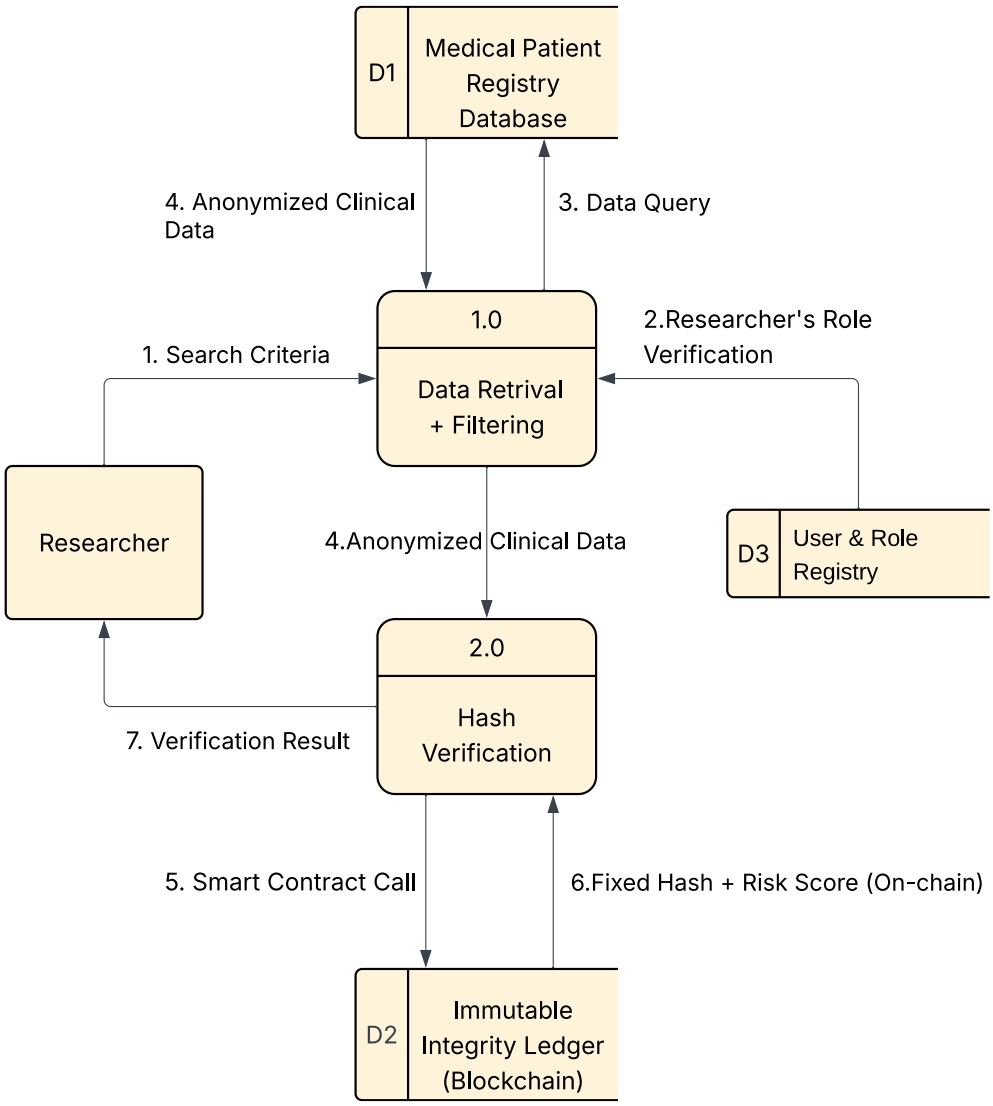


Εικόνα 23 : Διάγραμμα ροής δεδομένων αλληλεπίδρασης του χρήστη με το ρόλο του «γιατρού» και του υποσυστήματος τεχνητής νοημοσύνης του CDSS

4.3.2.2 Διάγραμμα Ροής Δεδομένων με ρόλο χρήστη τον «Ερευνητή»

Στην **Εικόνα 24** παρουσιάζεται το διάγραμμα ροής δεδομένων με ρόλο χρήστη ως ερευνητής και είναι επιπέδου 2 (DFD Level 2). Το διάγραμμα αντικατοπτρίζει τη κύρια διαφορά που έχει ο ρόλος του ερευνητή από τον ρόλο του γιατρού. Ο ερευνητής εξουσιοδοτείται μόνο με τα δικαιώματα της ανάγνωσης και επαλήθευσης των δεδομένων. Στην ροή εργασίας αρχικά ο ερευνητής προκειμένου να απομονώσει υποσύνολα δεδομένων από τη βάση MIMIC-IV, εισάγει τα απαραίτητα κριτήρια αναζήτησης, που είναι ερωτήματα προς τη βάση δεδομένων. Τα κριτήρια αναζήτησης μπορεί να είναι δημογραφικά χαρακτηριστικά, διάφορες κλινικές διαγνώσεις με συγκεκριμένους κωδικούς όπως η ICD-9 ή ICD-10, εργαστηριακές τιμές όπως η κρεατινίνη, το νάτριο, το κάλιο και η αιμοσφαιρίνη, η διάρκεια νοσηλείας και τα τελικά αποτελέσματα των προβλέψεων. Ταυτόχρονα επικοινωνεί με το μητρώο χρηστών, το οποίο υλοποιείται μέσω του Supabase Auth και περιλαμβάνει πίνακες με τα μοναδικά αναγνωριστικά χρηστών με συγκεκριμένους ρόλους. Η επαλήθευση του ρόλου πραγματοποιείται σε τρία επίπεδα, το επίπεδο ταυτοποίησης, το επίπεδο αιτήματος και το επίπεδο βάσης δεδομένων. Στο επίπεδο ταυτοποίησης ελέγχονται τα διαπιστευτήριά του χρήστη και στην συγκεκριμένη περίπτωση το σύστημα επιστρέφει το

κρυπτογραφημένο διαβατήριο-token με τον ρόλο του ερευνητή. Στο επίπεδο αιτήματος όπου ο ερευνητής πραγματοποιεί αναζήτηση στη πλατφόρμα, το frontend αλληλοεπιδρά με το FastAPI, υπό την προϋπόθεση πως ο χρήστης έχει πράγματι το ρόλο που απαιτείται για την ανάκτηση ανωνυμοποιημένων δεδομένων. Τέλος στο επίπεδο βάσης δεδομένων πραγματοποιείται διασταύρωση με το μητρώο χρηστών και ανατίθενται τα ανάλογα δικαιώματα. Εφόσον έχει δοθεί η έγκριση αναζήτησης η συνιστώσα 1.0 εκτελεί τα ανάλογα ερωτήματα δεδομένων (data query) στη βάση δεδομένων D1 που περιλαμβάνει το ιστορικό του ασθενή. Από εκείνη τα αποτελέσματα επιστρέφονται σε ανώνυμη μορφή, ώστε καμία ευαίσθητη πληροφορία δεν παρέχεται για ταυτοποίηση του ασθενή πέραν του επιπέδου δεδομένων. Συνεχίζοντας τα ανώνυμα δεδομένα περνάνε στη συνιστώσα 2.0, κατά την οποία γίνεται κλήση στο έξυπνο συμβόλαιο της αλυσίδας συστοιχειών (D2 Immutable Integrity Ledger). Η τελευταία επιστρέφει το αμετάβλητο αποτύπωμα (fixed hash) και το σκορ κινδύνου (risk score) που είχαν καταγραφεί από την αρχική πρόβλεψη. Ως τελευταία ενέργεια στη ροή δεδομένων πραγματοποιείται η σύγκριση των κλινικών δεδομένων με τα ανακτώμενα δεδομένα από το Blockchain και το τελικό αποτέλεσμα επαλήθευσης παραδίδεται στον ερευνητή επιβεβαιώνοντας την ανθεκτικότητα των δεδομένων.

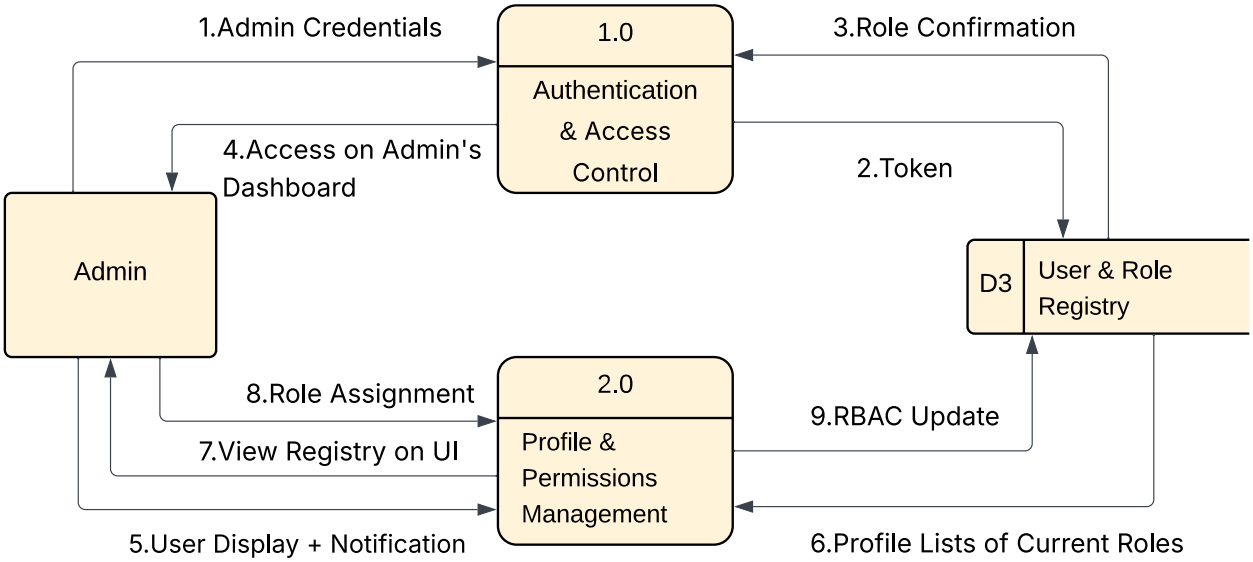


Εικόνα 24 : Διάγραμμα ροής δεδομένων αλληλεπίδρασης του χρήστη με το ρόλο του «ερευνητή» και του CDSS.

4.3.2.3 Διάγραμμα Ροής Δεδομένων με ρόλο χρήστη τον «Διαχειριστή»

Στην **Εικόνα 25** παρουσιάζεται το διάγραμμα ροής δεδομένων όταν ο χρήστης έχει τον ρόλο του διαχειριστή. Ο κύριος ρόλος αυτού του ρόλου είναι η διαχείριση των χρηστών και η απόδοση των δικαιωμάτων μέσω του μοντέλου RBAC. Αρχικά, γίνεται η εισαγωγή των διαπιστευτηρίων διαχειριστή στη συνιστώσα 1.0. Η συνιστώσα αντιπροσωπεύει την επαλήθευση του διακριτικού ασφαλείας (token) μέσω του μητρώου χρηστών D3. Με την επιτυχημένη ταυτοποίηση του χρήστη ως διαχειριστής του δίνεται η πρόσβαση στον πίνακα ελέγχου. Ο πίνακας ελέγχου αποτελεί το κεντρικό περιβάλλον διεπαφής (UI) που εκτελούνται όλες οι διοικητικές ενέργειες. Στην συνιστώσα 2.0 το σύστημα ανακτά τις λίστες προφίλ των τρεχόντων ρόλων και τις προβάλλει στο πίνακα ελέγχου. Έτσι ο διαχειριστής μπορεί και εποπτεύει το ιατρικό και ερευνητικό προσωπικό που έχουν εγγραφεί στη πλατφόρμα. Μια επιπλέον ενέργεια που κάνει ο διαχειριστής είναι η

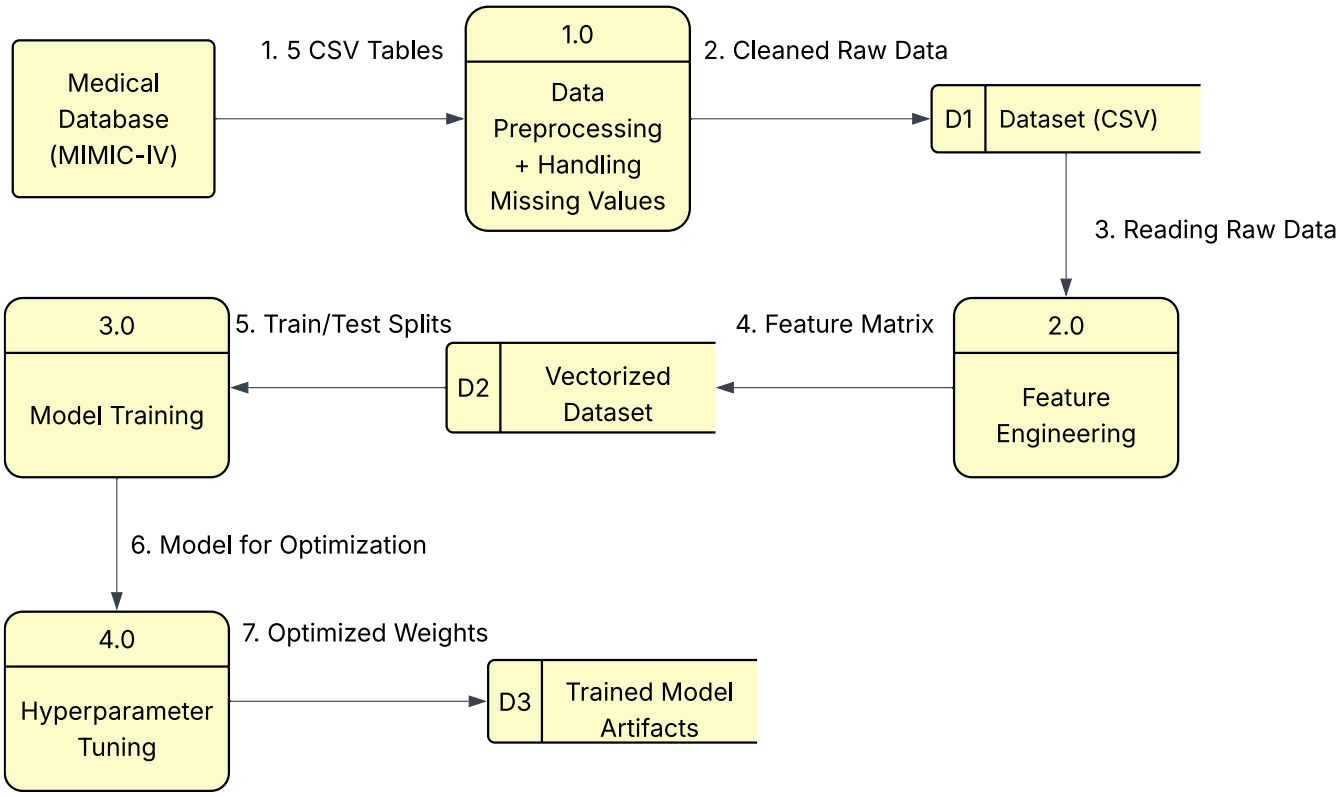
ενημέρωση των δικαιωμάτων εξουσιοδότησης των χρηστών. Αυτό αντιπροσωπεύει η συνιστώσα 2.0.



Εικόνα 25 : Διάγραμμα ροής δεδομένων αλληλεπίδρασης του χρήστη με το ρόλο του «διαχειριστή» και του CDSS

4.3.2.4 Διάγραμμα Ροής Δεδομένων Εκπαίδευσης και Προετοιμασίας Μοντέλου

Το διάγραμμα στην **Εικόνα 26** αφορά την ροή δεδομένων εκπαίδευσης και δημιουργίας του μοντέλου μηχανικής μάθησης που θα χρησιμοποιηθεί στο υποσύστημα τεχνητής νοημοσύνης και κατ' επέκταση από το ιατρικό προσωπικό. Η διαδικασία ξεκινά με την ανάκτηση δεδομένων στη πρωτογενή μορφή CSV πινάκων, από τη βάση δεδομένων D1, τα οποία περιλαμβάνουν χιλιάδες νοσηλείες με τα εργαστηριακά αποτελέσματα των εξετάσεων τους. Ακολουθεί ο καθαρισμός και η συμπλήρωση των ελλιπών τιμών στη συνιστώσα 1.0 και παράγονται τα «καθαρά πρωτογενή δεδομένα» (Cleaned Raw Data). Στην επόμενη συνιστώσα πραγματοποιείται η ανάγνωση αυτών των δεδομένων και η μηχανική χαρακτηριστικών (Feature engineering), στην οποία εφαρμόζονται η στατιστική μοντελοποίηση και η προ επεξεργασία των δεδομένων, όπως ο μετασχηματισμός των δεδομένων με μέση τιμή 0 και τυπική απόκλιση 1 (Standard Scaling). Τα αποτελέσματα του νέου πίνακα χαρακτηριστικών αποθηκεύονται σε ένα διανυσματικό σύνολο δεδομένων που θα αξιοποιηθεί από το μοντέλο για την εκπαίδευση του. Η συνιστώσα 3.0 αποτελεί την καρδιά του pipeline του υποσυστήματος τεχνητής νοημοσύνης. Σε αυτό πραγματοποιούνται η ρύθμιση των υπερπαραμέτρων, η πειραματική εκπαίδευση και σύγκριση των αλγορίθμων, η τελική αξιολόγηση απόδοσης του μοντέλου και η επιλογή του ακριβέστερου μοντέλου πρόβλεψης. Η τελευταία διεργασία της ροής δεδομένων είναι η μετατροπή του εκπαιδευμένου μοντέλου σε στατικό αρχείο με κατάληξη .pkl. Το συγκεκριμένο αρχείο φορτώνεται στο κανάλι επικοινωνίας του εξυπηρετητή FastAPI και παρέχει ζωντανές προβλέψεις καρδιακής ανεπάρκειας στον χρήστη.

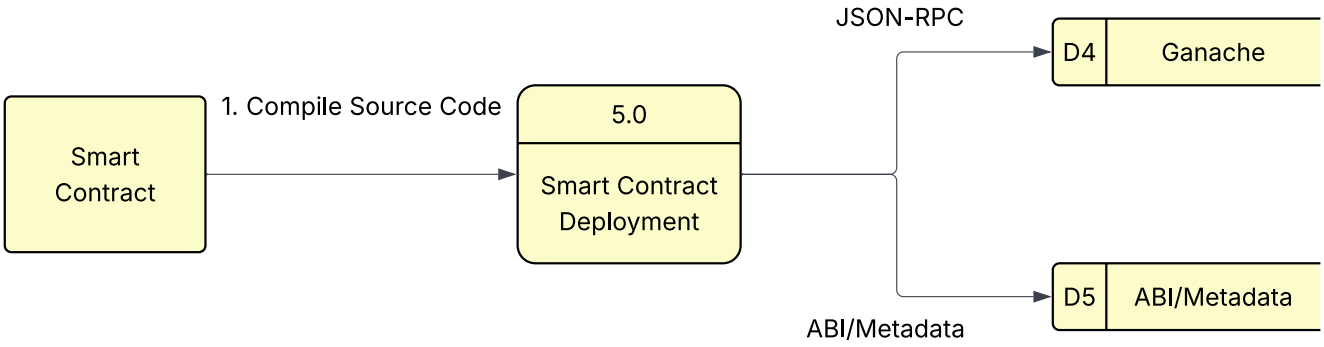


Εικόνα 26 : Διάγραμμα ροής δεδομένων αλληλεπίδρασης του χρήστη με το ρόλο του «διαχειριστή» και του CDSS.

4.3.2.5 Διάγραμμα Ροής Δεδομένων στο Έξυπνο Συμβόλαιο

Το διάγραμμα στην **Εικόνα 27** απεικονίζει την ανάπτυξη του έξυπνου συμβολαίου στο ιδιωτικό Blockchain του CDSS. Η οντότητα smart contract αποτελεί το έξυπνο συμβόλαιο που ο κώδικας του υλοποιεί την συγκατάθεση ασθενών, τη δημοσίευση σκορ κινδύνου, την επαλήθευση προβλέψεων από ιατρούς ή ερευνητές και το μηχανισμό ελέγχου πρόσβασης βάσει ρόλων (RBAC) μέσω του OpenZeppelin AccessControl. Στη συνιστώσα «Smart Contract Deployment», αναπαρίστανται ο κεντρικός κόμβος επεξεργασίας. Κατά τη φάση αυτή, ο Solidity compiler μεταγλωττίζει τον πηγαίο κώδικα σε bytecode που αναπτύσσεται σε τοπικό κόμβο Ethereum στο δίκτυο. Ο τελικός κόμβος παράγει δύο ξεχωριστές εξόδους που κατευθύνονται σε αντίστοιχους χώρους αποθήκευσης. Η πρώτη έξοδος κατευθύνεται προς τον αποθηκευτικό χώρο του Ganache. Η ροή δεδομένων στο Ganache πραγματοποιείται μέσω του πρωτοκόλλου JSON-RPC. Μετά την ανάπτυξη του, το έξυπνο συμβόλαιο γίνεται μόνιμα διαθέσιμο στο δίκτυο και δέχεται κλήσεις τόσο από το frontend μέσω ethers.js όσο και scripts που λειτουργούν εκτός λειτουργίας στο σύστημα. Η δεύτερη έξοδος κατευθύνεται προς το χώρο αποθήκευσης D5. Στο χώρο αυτό αποθηκεύονται τα μεταδεδομένα που έχουν παραχθεί κατά τη μεταγλώττιση του συμβολαίου. Τα σημαντικότερα μεταδεδομένα που αποθηκεύονται στο χώρο αποθήκευσης D5 είναι το Application Binary Interface (ABI) του συμβολαίου, το οποίο αντιστοιχεί στη περιγραφή JSON όλων των δημόσιων συναρτήσεων, event και δομών δεδομένων του έξυπνου συμβολαίου και η διεύθυνση

ανάπτυξης. Το ABI αποτελεί ουσιαστικά το κλειδί επικοινωνίας του frontend με το blockchain, επιτρέποντας στη βιβλιοθήκη ethers.js να κωδικοποιεί ορθά τις κλήσεις των on-chain συναρτήσεων. Ακόμη, η διεύθυνση ανάπτυξης χρησιμοποιείται από το frontend και τα scripts που λειτουργούν εκτός λειτουργίας με το σύστημα.



Εικόνα 27 : Διάγραμμα ροής δεδομένων αλληλεπίδρασης του χρήστη με το ρόλο του «διαχειριστή» και του CDSS.

4.3.3 Διαγράμματα Ακολουθίας

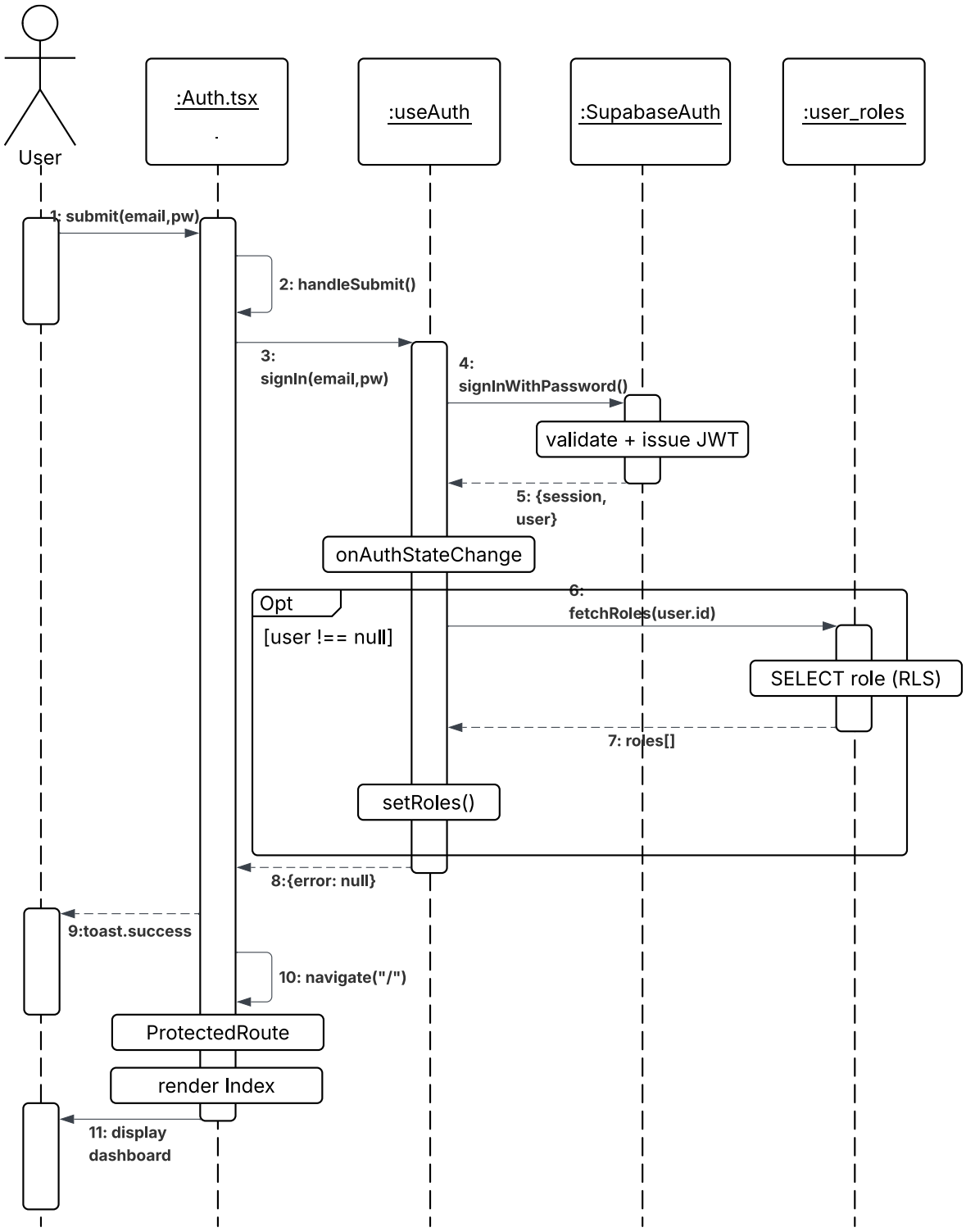
4.3.3.1 Διάγραμμα Ακολουθίας Ταυτοποίησης, Σύνδεσης και Ανάθεσης Ρόλων

Στο πρώτο διάγραμμα ακολουθίας, στην **Εικόνα 28**, αποτυπώνεται η ροή λειτουργίας του συστήματος από την ταυτοποίηση και είσοδο του χρήστη στην πλατφόρμα έως την ανάθεση ρόλου εξουσιοδότησης του. Το διάγραμμα αυτό αποτελεί το πρώτο σημείο επαφής του χρήστη με την κλινική πλατφόρμα και εστιάζει στο μηχανισμό Role-Based Access Control (RBAC) που είναι κομβικό σημείο για τις μεταγενέστερες αλληλεπιδράσεις μεταξύ της βάσης δεδομένων και του έξυπνου συμβολαίου.

Στο διάγραμμα απεικονίζονται 5 γραμμές ζωής (lifelines) και η καθεμία αντιστοιχεί σε μια ξεχωριστή αρμοδιότητα. Οι γραμμές ζωής είναι οι εξής: ο «Χρήστης» (User) είναι η εξωτερική οντότητα του συστήματος και αναπαριστά τον χρήστη ανεξαρτήτου ρόλου. Το «Auth.tsx» αποτελεί συνιστώσα της βιβλιοθήκης React με κύρια λειτουργία την γραφική διεπαφή εισόδου, παρέχοντας τη φόρμα καταχώρισης των διαπιστευτηρίων, όπως η ηλεκτρονική διεύθυνση email ή ο κωδικός πρόσβασης. Ακολουθεί το «useAuth» που αποτελεί hook της βιβλιοθήκης React και παρέχει τις κεντρικές συναρτήσεις σύνδεσης, αποσύνδεσης και ανάθεσης ρόλου (signIn, signUp, signOut, hasRole). Στη συνέχει το «SupabaseAuth» αποτελεί την υπηρεσία ταυτοποίησης στη βάση δεδομένων της Supabase. Τελευταία γραμμή ζωής είναι ο πίνακας «user_roles» που αποθηκεύονται τα μοναδικά αναγνωριστικά των ανατιθέμενων ρόλων.

Η εκκίνηση της ροής πραγματοποιείται με την συμπλήρωση του κωδικού πρόσβασης και της ηλεκτρονικής διεύθυνσης του χρήστη στη φόρμα σύνδεσης και την υποβολή της. Η συνιστώσα «Auth.tsx» δεσμεύει τα διαπιστευτήρια και ενεργοποιεί τη διαδικασία ταυτοποίησης. Η

συγκεκριμένη αρχιτεκτονική προσέγγιση πραγματοποιήθηκε ώστε η πλατφόρμα να τηρεί τις αρχές της διαστρωματωμένης ασφάλειας και του διαχωρισμού των ανησυχιών (separation of concerns). Η διασφάλιση αφαίρεσης της επιχειρησιακής λογικής από τα συστατικά στοιχεία της διεπαφής δίδεται από το hook «useAuth». Το «useAuth» έχει τον ρόλο του ενδιάμεσου διαμεσολαβητή για την Supabase και πραγματοποιεί την επαλήθευση των στοιχείων έναντι κρυπτογραφημένων δεδομένων και έκδοση διακριτικού ασφαλείας JSON Web Token. Με την επιστροφή των {session, user} στη γραμμή ζωής της useAuth, ο μηχανισμός «onAuthStateChange» δίνει την έγκριση αυθεντικοποίησης του χρήστη σε όλη τη διάρκεια της εφαρμογής. Με λίγα λόγια λειτουργεί ως πάσο εισόδου. Το επόμενο βήμα είναι η προϋπόθεση τήρησης της συνθήκης [user !== null] που στελεχώνεται στο πλαίσιο opt. Εφόσον ο χρήστης ικανοποιεί την συνθήκη ακολουθεί η κλήση του ρόλου εξουσιοδότησης. Αυτό γίνεται μέσω την κλήσης της συνάρτησης fetchRoles() προς τον πίνακα user_roles. Η ανάκτηση των εξουσιοδοτημένων δικαιωμάτων γίνεται βασιζόμενη σε πολιτικές Row-Level Security (RLS), δηλαδή ο χρήστης με βάση τον ρόλο που του ανατίθεται έχει και πρόσβαση στις ανάλογες λειτουργίες. Από εκεί και έπειτα, η τοπική αποθήκευση των ρόλων επιτρέπει τη δυνατότητα άμεσης λήψης αποφάσεων πρόσβασης από το βοηθητικό ελεγκτή. Η διαδικασία ολοκληρώνεται στο στοιχείο ProtectedRoute, του οποίου η κύρια λειτουργία είναι η διασφάλιση εγκυρότητας της συνεδρίας και του JWT. Εν κατακλείδι, η επιλεγμένη αρχιτεκτονική ροής επιτυγχάνει την ταυτόχρονη συνέπεια της off-chain επιχειρησιακής λογικής και της on-chain λογικής εξουσιοδότησης, ενισχύοντας την αμεταβλητότητα και την εμπιστοσύνη διαχείρισης των κλινικών δεδομένων από τους ασθενείς.



Εικόνα 28 : Διάγραμμα ακολουθίας της ροής ταυτοποίησης χρήστη και ανάκτησης ρόλων πρόσβασης (RBAC)

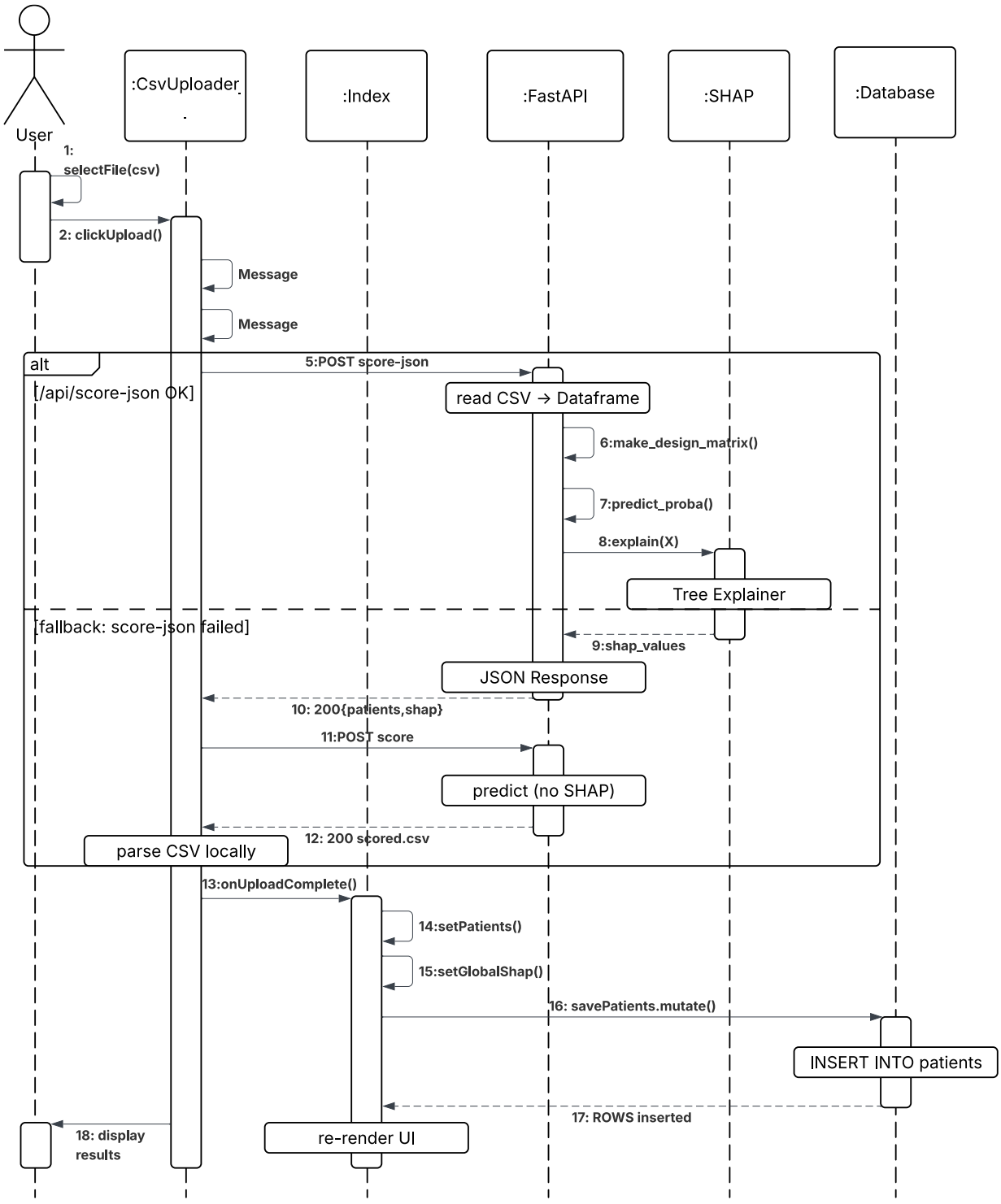
4.3.3.2 Διάγραμμα Ακολουθίας Πρόβλεψης Κινδύνου με Εξηγησιμότητα SHAP

Το διάγραμμα ακολουθίας στην **Εικόνα 29** απεικονίζει τον πυρήνα του συστήματος που είναι η μεταφόρτωση συνόλου δεδομένων ασθενών σε μορφή CSV από τον εξουσιοδοτημένο χρήστη και ως αποτέλεσμα λαμβάνει τις ατομικές προβλέψεις κινδύνου καρδιακής ανεπάρκειας μαζί με τις αντίστοιχες εξηγήσεις SHAP. Στη ροή αναπαρίστανται και τα τρία δομικά επίπεδα αρχιτεκτονικής της εφαρμογής, της παρουσίασης, της επιχειρησιακής λογικής και των δεδομένων.

Το διάγραμμα εμφανίζει έξι γραμμές ζωής, με την καθεμιά να έχει έναν ξεχωριστό ρόλο στην λειτουργία του. Οι γραμμές ζωής είναι οι εξής: η εξωτερική οντότητα του «Χρήστη» που αντιπροσωπεύει τον εξουσιοδοτημένο ρόλο που έχει ο χρήστης. Η συνιστώσα «CsvUploader» στην οποία υλοποιείται η γραφική διεπαφή μεταφόρτωσης αρχείου CSV. Η συνιστώσα «Index» αποτελεί τον κεντρικό διαχειριστή της πλατφόρμας, καθώς συγκεντρώνει τη συνολική κατάσταση των ασθενών και των SHAP τιμών και την αυτόματη ενημέρωση της κατάστασης της σελίδας. Ακολουθεί ο «FastAPI» που λειτουργεί ως κεντρικός διαχειριστής του επιχειρησιακού επιπέδου λογικής. Ο ίδιος λαμβάνει το μεταφορτωμένο CSV, εκτελεί την προ επεξεργασία και πρόβλεψη του φορτωμένου μοντέλου XGBoost και επιστρέφει μια δομημένη απόκριση JSON. Η συνιστώσα της βιβλιοθήκης «SHAP» που έχει σκοπό την ερμηνεία των ακατέργαστων τιμών SHAP και η συνιστώσα «Database» που εκφράζει τον εκτός αλυσίδας (off-chain) αποθηκευτικό χώρο.

Η διαδικασία εκκινεί με την επιλογή ενός αρχείου CSV από τον χρήστη και την μεταφόρτωση του. Η συνιστώσα CsvUploader πραγματοποιεί τοπικό έλεγχο και μέσω της βιβλιοθήκης react-hot-toast απορρίπτει οποιοδήποτε αρχείο δεν είναι της μορφής CSV. Στην συνέχεια πραγματοποιούνται δύο μηνύματα κατάστασης, ένα για την ενημέρωση της επιλογής αρχείου και ένα για την έναρξη της επεξεργασίας. Σκοπός ύπαρξης των μηνυμάτων είναι η συνδεσιμότητα και βελτίωση της εμπειρίας του χρήστη. Το πλαίσιο alt αποτελεί το κορμό λειτουργίας του διαγράμματος και ενσωματώνει τη λογική της σταδιακής υποβάθμισης (graceful degradation). Με βάση τη συγκεκριμένη λογική αν αποτύχει η υλοποίηση μιας ενέργειας τότε η ροή συνεχίζει με μια απλούστερη εναλλακτική. Έστω ότι επιτυγχάνεται η αποστολή και καταγραφή της πρόβλεψης κινδύνου στο σύστημα [/api/score-json]. Σε αυτήν την υλοποίηση το CsvUploader το αρχείο .json ενσωματώνεται σε ένα πολυμερές σώμα αιτήματος και αποστέλλεται στο endpoint με αίτημα POST. Μόλις το αίτημα φτάσει στο endpoint, ο εξυπηρετητής FastAPI προχωράει στην ανάγνωση του CSV και μετατροπή του σε δομή Pandas Dataframe. Για την σωστή μετατροπή του εφαρμόζει τη συνάρτηση make_design_matrix() για τη διαμόρφωση του πίνακα σχεδιασμού και εκτελεί την predict_proba() επί του καλιμπραρισμένου ταξινομητή (calibrator). Αμέσως μετά ο FastAPI καλεί τη μέθοδο explain(X) προς τη βιβλιοθήκη SHAP. Χρησιμοποιείται η μέθοδος Tree Explainer που είναι ένας ειδικός αλγόριθμος της βιβλιοθήκης SHAP, ο οποίος εξαιτίας της πολυωνυμικής πολυπλοκότητας προσφέρει υψηλή μαθηματική ακρίβεια στον υπολογισμό των SHAP τιμών σε πραγματικό χρόνο παράλληλα με το σύστημα. Με την επιστροφή των νέων επεξεργασμένων SHAP τιμών ο εξυπηρετητής δημιουργεί ένα τελικό αρχείο JSON, με την αρχική απόκριση να επιστρέφει στη συνιστώσα «CsvUploader». Στην περίπτωση που το σύστημα εμφανίσει [fallback:score-json failed], δηλαδή αποτύχει η αποστολή του, ο CsvUploader επαναλαμβάνει το

αίτημα προς το endpoint /api/score. Η απάντηση που θα λάβει θα είναι το CSV χωρίς τις τιμές SHAP, καθώς δεν δόθηκε η δυνατότητα ευθυγράμμισης τους. Ο CsvUploader θα πραγματοποιήσει τη τοπική ανάλυση του επιστρεφόμενου CSV, εξασφαλίζοντας ότι η ανάλυση τους θα γίνει και στο ενδεχόμενο αποτυχίας εξηγησιμότητας των τιμών SHAP. Συνεχίζοντας τη ροή του, τα επεξεργασμένα δεδομένα μεταβιβάζονται στη συνιστώσα «Index», όπου από εκεί και έπειτα υλοποιούνται δύο διαδοχικές ενημερώσεις, η setPatients() που ουσιαστικά αποθηκεύει το πίνακα ασθενών στη τοπική μνήμη και η setGlobalShap() που διατηρεί και τις τιμές SHAP για την οπτικοποίηση τους μέσω του GlobalShapChart. Το Index χρησιμοποιεί τη συνάρτηση μετάλλαξης mutate() για τη δημιουργία ενός διανύσματος εγγραφών που απεικονίζει έναν ασθενή με τα βιοδεικτικά χαρακτηριστικά του και η συνεδρία με τη βάση δεδομένων εξασφαλίζει την ιχνηλασιμότητα της προέλευσης τους. Συνεχίζοντας πραγματοποιείται η εισαγωγή δεδομένων στο πίνακα PostgreSQL μέσω της INSERT. Μόλις ολοκληρωθεί η αποθήκευση πραγματοποιείται ενημέρωση της διεπαφής και προβολή των τελικών αποτελεσμάτων.



Εικόνα 29 : Διάγραμμα ακολουθίας της ροής πρόβλεψης κινδύνου καρδιακής ανεπάρκειας και ερμηνευσιμότητας SHAP, από τη μεταφόρτωση του αρχείου .CSV έως την οπτική παρουσίαση των διαγραμμάτων

4.3.3.3 Διάγραμμα Ακολουθίας Ανάρτησης Πρόβλεψης στο Έξυπνο Συμβόλαιο

Στο τρίτο διάγραμμα ακολουθίας, στην **Εικόνα 30**, αποτυπώνεται το διάστημα ροής λειτουργίας του συστήματος από την παραλαβή ενός αποτελέσματος από το μοντέλο πρόβλεψης έως την ανάρτηση του στο έξυπνο συμβόλαιο «MedicalOracle». Το διάγραμμα διακρίνεται σε δύο διακριτές φάσεις, τη φάση εγκατάστασης του ηλεκτρονικού πορτοφολιού, που εκτελείται μία φορά ανά συνεδρία του χρήστη και τη φάση ανάρτησης, που επαναλαμβάνεται για κάθε νοσηλεία με υποχρεωτική την υπογραφή του αρμόδιου κλινικού ιατρού και στις περισσότερες περιπτώσεις την νόμιμη συγκατάθεση του πάσχοντος ασθενή.

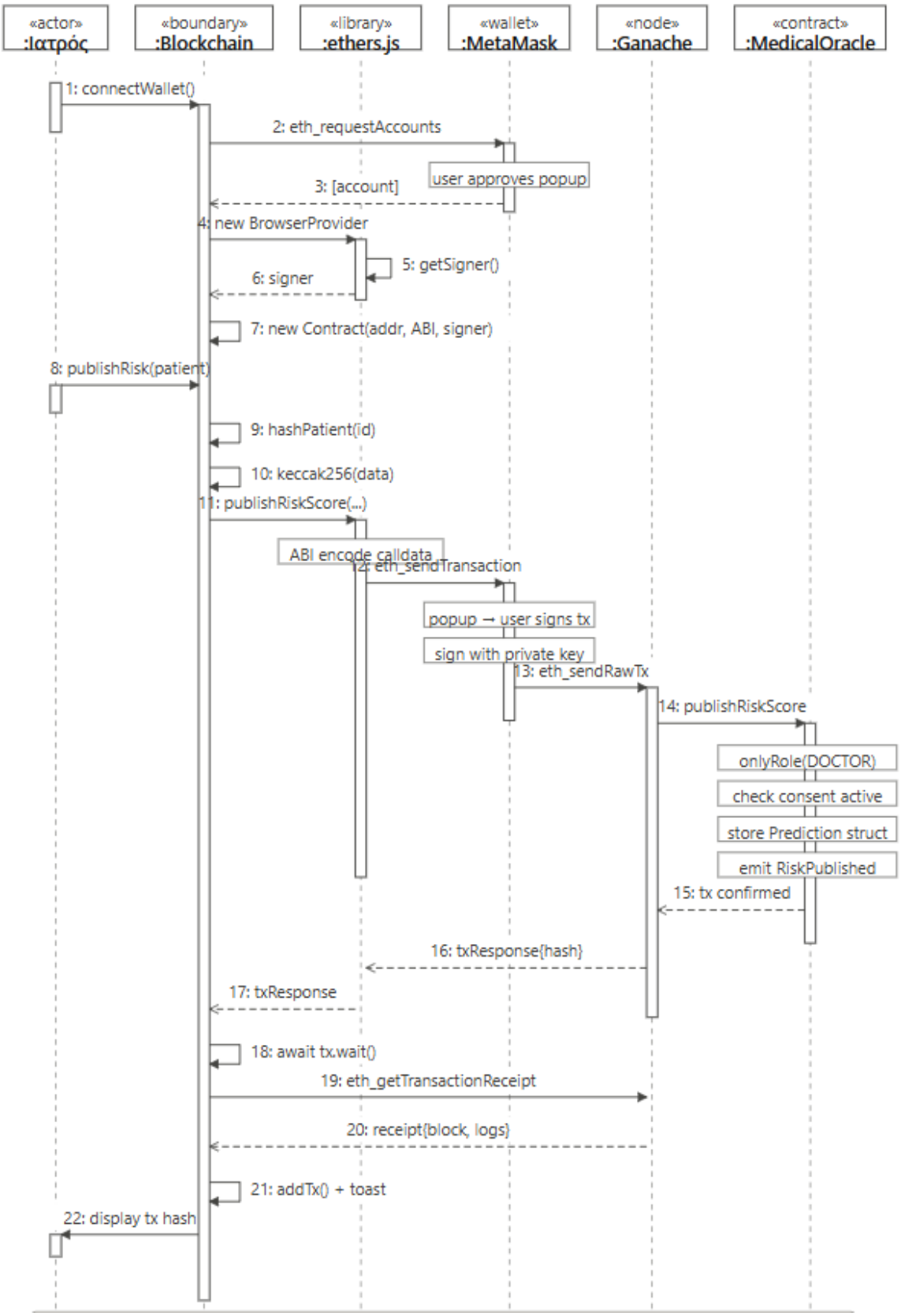
Η ροή του διαγράμματος χωρίζεται σε έξι γραμμές ζωής, με την καθεμιά να υποστηρίζει ξεχωριστή αρμοδιότητα στην αρχιτεκτονική του Blockchain επιπέδου. Η γραμμή ζωής εκκίνησης της ροής του διαγράμματος ξεκινάει από την εξωτερική οντότητα του «Ιατρού» που αναπαριστά τον εξουσιοδοτημένο χρήστη με ρόλο DOCTOR που έχει το δικαίωμα ανάρτησης των προβλέψεων. Η επόμενη συνιστώσα ονομάζεται «Blockchain» και αποτελεί την γέφυρα μεταξύ του χρήστη ιατρού και της αλυσίδας. Η βιβλιοθήκη «ethers.js» προσφέρει τις απαραίτητες συναρτήσεις και κλάσεις για τη δημιουργία και διαχείριση των αντικειμένων BrowserProvider, Signer και Contract. Παράλληλα αναλαμβάνει τις κωδικοποιημένες κλήσεις του ABI που παράγεται από τον solc (Solidity Compiler) του περιβάλλοντος Remix IDE. Η συνιστώσα «MetaMask» λειτουργεί ως διαχειριστής των ιδιωτικών κλειδιών του χρήστη και αποστέλλει αναδυόμενα παράθυρα έγκρισης των συναλλαγών δια χειροκίνητης υπογραφής. Το «Ganache» αποτελεί τοπικό περιβάλλον προσομοίωσης της Ethereum αλυσίδας. Μέσα σ ’αυτόν φιλοξενείται σε κόμβο το έξυπνο συμβόλαιο που τρέχει το RPC Endpoint και εκτελεί το EVM Bytecode που έχει παραχθεί από το Remix IDE επίσης για την επιβεβαίωση των συναλλαγών σε προσομοιωμένα μπλοκ. Τέλος το έξυπνο συμβόλαιο παρουσιάζεται στη τελευταία γραμμή ζωής «MedicalOracle» και αποτελεί το έξυπνο συμβόλαιο που αναπτύσσεται στη αλυσίδα και υλοποιεί τη λογική επιβολής των ρόλων, του ελέγχου συγκατάθεσης και τη μόνιμη αποθήκευση των προβλέψεων.

Η πρώτη φάση εκκινεί με τη κλήση connectWallet() από το κλινικό ιατρό προς τη συνιστώσα της Blockchain, σηματοδοτώντας τη πρόθεση για σύνδεση με το τοπικό κόμβο της αλυσίδας. Η Blockchain διαβιβάζει το αίτημα με την μορφή eth_requestAccounts στο αρχείο ethers.js και εκείνο με τη κατάλληλη συνάρτηση λειτουργίας το προωθεί στο πορτοφόλι MetaMask. Αφού το Metamask αποστείλει το αίτημα προς έγκριση για την αποδοχή καταγραφής νοσηλείας με το ανάλογο κόστος στην αλυσίδα και ο χρήστης ιατρός το αποδεχτεί, τότε επιστρέφεται από το Metamask και μετά την ethers.js η εκάστοτε διεύθυνση του λογαριασμού χρήστη. Συνεχίζοντας, η συνιστώσα Blockchain κατασκευάζει το αντικείμενο BrowserProvider, το οποίο με τη σειρά του καλεί τη μέθοδο getSigner() του ethers.js για τη ανάκτηση της υπογραφής που αντιπροσωπεύει το αντικείμενο «Signer». Ο Signer παρέχει την ικανότητα υπογραφής των συναλλαγών μέσω ιδιωτικού κλειδιού του χρήστη, όχι το ίδιο το κλειδί. Εφόσον συγκεντρωθούν και τα τρία αντικείμενα, διεύθυνση συμβολαίου, ABI και Signer, η Blockchain δημιουργεί ένα νέο συμβόλαιο (Contract), που εκθέτει τις δημόσιες και εξωτερικές συναρτήσεις της Solidity σε κανονικές μεθόδους Javascript. Με την ολοκλήρωση του βήματος 7, έχει γίνει η εγκατάσταση του

πορτοφολιού στο σύστημα και τα τρία βασικά αντικείμενα της βιβλιοθήκης ethers.js, Provider,Signer και Contract, με τις διακριτές λειτουργίες τους παραμένουν σταθερά για το υπόλοιπο της συνεδρίας.

Η δεύτερη φάση, ο ιατρός δημοσιεύει τη πρόβλεψη κινδύνου με τη συνάρτηση publishRisk() και στη συνιστώσα Blockchain παράγονται δύο ξεχωριστές κρυπτογραφήσεις. Η πρώτη αφορά τη συνάρτηση hashPatient(id) που παράγει ψευδωνυμοποιημένο το hash του ασθενή, όπως αναφέρθηκε στη ενότητα 5.5.1 και τη keccak256(data) για τη δημιουργία του data_integrity_hash, που είναι το hash όλης της πληροφορίας σχετικά με τον ασθενή και τα αποτελέσματα του. Μόλις παραλάβει τα κρυπτογραφημένα μηνύματα, η ethers.js τα κωδικοποιεί σε ένα ABI-encoded calldata και προωθεί το eth_sendTransaction στο πορτοφόλι MetaMask, που εμφανίζεται αναδυόμενο παράθυρο στο οποίο ο γιατρός υπογράφει χειροκίνητα τη συναλλαγή με το ιδιωτικό κλειδί του. Όταν υπογράφεται η συναλλαγή προωθείται στο Ganache μέσω του eth_sendRawTransaction και εκτελείται από το έξυπνο συμβόλαιο. Στο έξυπνο συμβόλαιο εκτελούνται τέσσερις κρίσιμες ενέργειες σε τέσσερα κρίσιμα στάδια. Πρώτον, η συνάρτηση modifier onlyRole(DOCTOR) επιβεβαιώνει πως ο υπογράφων είναι ο εγκεκριμένος ρόλος και η συνάρτηση hasStorageConsent να επικυρώνει τη συγκατάθεση των αποτελεσμάτων του ασθενή. Τα αποτελέσματα της πρόβλεψης της νοσηλείας και τα στοιχεία ταυτοποίησης του ασθενή αποθηκεύονται στο νέο αντικείμενο Prediction, που περιλαμβάνει τον πίνακα patientRecords και η εκπομπή των αποτελεσμάτων γίνεται με το συμβάν RiskPublished. Μόλις ολοκληρωθούν αυτά, το Ganache επιστρέφει το μήνυμα επιβεβαίωσης συναλλαγής, στο βήμα 15 tx confirmed, μαζί με το αντίστοιχο κατακερματισμό συναλλαγής.

Στα τελευταία βήματα, το Blockchain εκτελεί την συνάρτηση αναμονής tx.wait() που ουσιαστικά δηλώνει την αναμονή της επικύρωσης της συναλλαγής εντός block. Το αποτέλεσμα της ενέργειας αυτής είναι η απόδειξη eth_getTransactionReceipt προς το Ganache, που επιστρέφει το αποδεικτικό με τον αριθμό μπλοκ και συνδέσεις μέσω της εσωτερικής συνάρτησης addTx() και αποστέλλει ειδοποίηση με αναδυόμενο παράθυρο για την επιβεβαίωση της κρυπτογραφημένης συναλλαγής. Έτσι βλέποντας την επιβεβαίωση ο γιατρός μπορεί να προχωρήσει σε επόμενη ανάρτηση ή να επαληθεύσει μια εγγραφή νοσηλείας στην καρτέλα συναλλαγών του Ganache. Με τη συγκεκριμένη αρχιτεκτονική ροής επιτυγχάνεται η ταυτόχρονη μη διαρροή ιδιωτικών κλειδιών, καθώς η υπογραφή της συναλλαγής απομονώνεται στο ασφαλές ηλεκτρονικό πορτοφόλι Metamask, η διπλή επικύρωση μέσω των συναρτήσεων on-chain onlyRole και hasStorageConsent δημιουργώντας ένα διπλό στρώμα ασφαλείας στη διατήρηση της ομαλής λειτουργίας του συστήματος. Επίσης η αμετάκλητη ιχνηλασιμότητα κάθε δημοσιευμένης πρόβλεψης μέσω τον κρυπτογραφημένων συναλλαγών και των εκπεμπόμενων συμβάντων (events) δημιουργεί μια σύζευξη μεταξύ off-chain επιχειρησιακής λογικής και on-chain αξιόπιστης καταγραφής.


«actor»
:Ιατρός
«boundary»
:Blockchain
«library»
:ethers.js
«wallet»
:MetaMask
«node»
:Ganache
«contract»
:MedicalOracle
1: connectWallet()
2: eth_requestAccounts
user approves popup
3: [account]
4: new BrowserProvider
5: getSigner()
6: signer
7: new Contract(addr, ABI, signer)
8: publishRisk(patient)
9: hashPatient(id)
10: keccak256(data)
11: publishRiskScore(...)
ABI encode calldata
12: eth_sendTransaction
popup → user signs tx
sign with private key
13: eth_sendRawTx
14: publishRiskScore
onlyRole(DOCTOR)
check consent active
store Prediction struct
emit RiskPublished
15: tx confirmed
16: txResponse{hash}
17: txResponse
18: await tx.wait()
19: eth_getTransactionReceipt
20: receipt{block, logs}
21: addTx() + toast
22: display tx hash

Εικόνα 30 : Διάγραμμα ακολουθίας ανάρτησης πρόβλεψης στο έξυπνο συμβόλαιο MedicalOracle, καλύπτοντας τις φάσεις σύνδεσης πορτοφολιού και δημοσίευσης προβλέψεων.

Κεφάλαιο 5: Πειραματικό Μέρος

5.1 Προ επεξεργασία Δεδομένων

Η προ επεξεργασία των δεδομένων αποτελεί το πρώτο και καθοριστικό στάδιο για την αξιοπιστία της προτεινόμενης κλινικής πλατφόρμας. Όπως αναφέρθηκε προηγουμένως, το προτεινόμενο σύστημα υποστήριξης κλινικών αποφάσεων χρησιμοποιεί το σύνολο δεδομένων MIMIC-IV. Το ίδιο συγκαταλέγεται ανάμεσα στις πιο αναγνωρισμένες βάσεις κλινικών δεδομένων διεθνώς, καθώς αντανακλά τη λειτουργική πραγματικότητα ενός νοσοκομειακού περιβάλλοντος και περιλαμβάνει εκατομμύρια ετερογενείς εγγραφές, αποτελέσματα εργαστηριακών εξετάσεων, πληθώρα νοσηλειών ανά ασθενή, ελλιπείς τιμές και ανισορροπία κλάσεων προς τη μεταβλητή στόχου. Η φύση του συνόλου δεδομένων καθορίζει σημαντικό τον μετασχηματισμό του σε μορφή κατάλληλή για εποπτευόμενη μάθηση. Για την ορθότητα του μετασχηματισμού απαιτείται να πραγματοποιηθούν προσεκτικά σχεδιασμένες ενέργειες. Στις ενότητες που ακολουθούν αξιοποιείται η ενοποίηση των πινάκων του MIMIC-IV σε έναν νέο βασικό πίνακα νοσηλειών, η εξαγωγή εργαστηριακών τιμών από βιοδείκτες σε καθορισμένο χρονικό ορίζοντα, η δημιουργία νέων σύνθετων κλινικών δεικτών, η κατασκευή ιστορικού συν νοσηροτήτων και η διαχείριση των ελλιπών τιμών, η εξισορρόπηση των κλάσεων και η τμηματοποίηση σε επίπεδο ασθενούς.

5.1.1 Δημιουργία Γενικού Πίνακα Νοσηλειών

Η κατασκευή ενός γενικού πίνακα νοσηλειών αποτελεί το πρώτο εκτελεστικό βήμα της προ επεξεργασίας των δεδομένων και υλοποιείται εξ ολοκλήρου από τη συνάρτηση build_dataset_mimic της ενότητας main.py. Η λειτουργία της συνάρτησης χωρίζεται σε τέσσερις διαδοχικές ενέργειες: τη φόρτωση των πηγαίων πινάκων, τη σύζευξη με αναγνωριστικό κλειδί ασθενούς, τον υπολογισμό της ηλικίας εισαγωγής και τα φίλτρα εγκυρότητας.

Για την κατασκευή του πίνακα νοσηλειών ενοποιούνται 3 βασικοί πίνακες της βάσης MIMIC-IV v3.1, οι «admissions», «patients», «diagnoses_icd». Για την εξοικονόμηση μνήμης κατά την επεξεργασία μεγάλων δεδομένων επιλέχτηκαν οι στήλες που αναφέρονται παρακάτω:

- Admissions { subject_id , hadm_id , admittime }
- Patients { gender , anchor_age , anchor_year }
- Diagnoses_Icd { subject_id , hadm_id , icd_code , icd_version }

Οι πίνακες admissions και patients συνενώνονται και με αριστερή σύζευξη με κλειδί το subject_id, διασφαλίζοντας με αυτόν τον τρόπο πως όλες οι νοσηλείες καταγράφονται σε αυτό το στάδιο. Για την προστασία των προσωπικών δεδομένων των ασθενών , η ηλικία τους υπολογίζεται διανυσματικά στο MIMIC-IV με βάση τον μαθηματικό τύπο της εξίσωσης (21):

$$age = anchor_year + (year(admittime) - anchor_year) \qquad (21)$$

Στην παραπάνω ισότητα, οι απόλυτες τιμές admittime και anchor_year έχουν μετατοπιστεί χρονικά και γραμμικά κατά ένα σταθερό αριθμό ετών για κάθε ασθενή. Η διαφορά τους ισούται με τη πραγματική διαφορά ετών μεταξύ του έτους αναφοράς (anchor_year) και του έτους εισαγωγής (admittime). Το πεδίο «gender» στο πίνακα Patients θα μετονομαστεί σε «sex» και θα γίνει δυαδική κατηγορική κωδικοποίηση η οποία αποδίδει την τιμή 1 για κάθε εγγραφή με αρχικό χαρακτήρα "M" και τη τιμή 0 για κάθε εγγραφή με αρχικό χαρακτήρα "F". Ως "M" ορίζεται το ανδρικό φύλο (Male). Ενώ "F" το θηλυκό φύλο (Female). Επίσης, για την εγκυρότητα των αποτελεσμάτων του μοντέλου εφαρμόστηκαν δύο διαδοχικά φίλτρα στον ενοποιημένο πίνακα. Το πρώτο περιορίζει τον ενήλικο πληθυσμό (ηλικία ≥ 18 $έτη$), αποκλείοντας τις παιδιατρικές νοσηλείες. Έτσι και αλλιώς με βάση τη παθοφυσιολογία αυτές διαφοροποιούνται ουσιωδώς στη κλινική διαχείριση τους. Το δεύτερο φίλτρο απορρίπτει εγγραφές που δεν έχουν τιμές στα πεδία subject_id, hadm_id, admittime, age, sex. Τα δύο προηγούμενα φίλτρα ποιότητας εφαρμόζονται πάντοτε από τη συνάρτηση που φτιάχνει το mimic και αποτελούν τα ελάχιστα κριτήρια για την βέλτιστη και εγκυρότερη αξιοποίηση του συνόλου. Επιπρόσθετα ο κώδικας υποστηρίζει τη παράμετρο max_hadm που επιτρέπει τη τυχαία δειγματοληψία ενός υποσυνόλου νοσηλειών μετά την εφαρμογή των φίλτρων, με στόχο την επιτάχυνση των ενδιάμεσων δοκιμών κατά τη φάση ανάπτυξης. Στα συμπεράσματα της τελικής φάσης εκπαίδευσης, η παράμετρος έχει απενεργοποιηθεί, ώστε η ροή εκπαίδευσης των μοντέλων μηχανικής μάθησης να βασιστεί στο πλήρες σύνολο των 546.028 νοσηλειών του MIMIC-IV v3.1 για τη μέγιστη στατιστική ισχύ των μετρήσεων. Εάν η παράμετρος ήταν ενεργοποιημένη ο γενικός πληθυσμός του συνόλου των νοσηλειών θα μειωνόταν από 546.028 σε 462.890 νοσηλείες, με τις 61.630 εγγραφές να απορρίπτονται λόγω ηλικιακών κριτηρίων. Το τελευταίο βήμα προ επεξεργασίας των δεδομένων είναι η ενοποίηση του γενικού πίνακα νοσηλειών μέσω αριστερής σύζευξης με το κλειδί hadm_id ($Dataset \bowtie_{hadm_id} Heart_Failure_Labels$) και όσες νοσηλείες δεν εμφανίζουν κανένα διαγνωστικό κωδικό λαμβάνουν τιμή ετικέτας False, δηλαδή 0.

Συνοψίζοντας, ο τελικός πίνακας νοσηλειών περιέχει 546.028 εγγραφές με 231.770 μοναδικούς ασθενείς. Τα τελικά αποτελέσματα φανερώνουν πως οι 66.270 νοσηλείες (12.14%) φέρουν θετική ετικέτα της ΚΑ, επομένως είναι πάσχοντες και οι 479.760 αρνητικές περιπτώσεις (87.87%). Λόγω της ανισορροπίας αυτής εισάγεται η στρατηγική υπερδειγματοληψίας SMOTE που αφορά αποκλειστικά το σύνολο εκπαίδευσης. Ο **Πίνακας 5.1** σηματοδοτεί τη μη-ανεξαρτησία των παρατηρήσεων σε επίπεδο νοσηλείας. Αυτό συμβαίνει λόγω της φύσης του MIMIC-IV, στο οποίο η χρήση συμβατικού τυχαίου διαχωρισμού θα οδηγούσε σε διαρροή δεδομένων. Για την αποφυγή αυτού του φαινομένου εφαρμόστηκε η στρατηγική GroupShuffleSplit με βάση το subject_id, ώστε να επιβεβαιωθεί ότι όλες οι καταγραφές νοσηλείας ενός ασθενούς ανήκουν σε ένα από τα δύο σύνολα εκπαίδευσης ή αξιολόγησης και επιτρέπεται η χρήση του μοντέλου σε πραγματικά άγνωστους πληθυσμούς. Το μοντέλο γίνεται περισσότερο αξιόπιστο και ακριβή για την χρήση του σε πραγματικές συνθήκες, κάτι που κρίνεται σημαντικό για την εφαρμογή του σε νοσοκομειακή μονάδα.

Πίνακας 5.1: Ταξινόμηση πληθυσμού ασθενών με γνώμονα τον αριθμό νοσηλειών ανά ασθενή

Αριθμός νοσηλειών ανά ασθενή	Πλήθος Ασθενών	Ποσοστό
1	111.371	56.67%
2	38.560	19.62%
$\boldsymbol{N \geq 3}$	46.596	23.71%
$\boldsymbol{Σύνολο > 1}$	85.156	43.33%

5.1.2 Εξαγωγή Εργαστηριακών Τιμών σε πλαίσιο 24 Ωρών

Η εξαγωγή των εργαστηριακών τιμών παρουσιάζεται ως η πιο απαιτητική φάση της προ επεξεργασίας δεδομένων, καθώς ο πίνακας labevents της MIMIC-IV περιλαμβάνει εκατοντάδες εκατομμύρια εγγραφές μετρήσεων. Λόγω της κλίμακας μεγέθους των εγγραφών απαιτείται η στρατηγική τμηματοποίηση και φιλτράρισμα που υλοποιείται η συνάρτηση extract_lab_values στο backend/main.py . Η συνάρτηση υλοποιεί τέσσερις διαδοχικές ενέργειες, κατά τις οποίες η κάθε μια επηρεάζει την επόμενη: η χαρτογράφηση των βιοδεικτών προς τους κωδικούς της MIMIC-IV, η τμηματική σάρωση του πίνακα labevents, η εφαρμογή χρονικού παραθύρου εικοσιτεσσάρων ωρών από την εισαγωγή και η αναδιάρθρωση των δεδομένων για την κατάλληλη μοντελοποίηση.

Η εκπαίδευση του προγνωστικού μοντέλου περιλαμβάνει 29 βιοδείκτες. Οι βιοδείκτες αυτοί ομαδοποιούνται σε πέντε κατηγορίες με γνώμονα την βιολογική λειτουργία τους.

Πίνακας 5.2: Ταξινόμηση βιοδεικτών που χρησιμοποιήθηκαν με βάση τη κατηγορία τους

Κατηγορία Βιοδεικτών	**Βιοδείκτες**
Αιματολογικοί Δείκτες	Αιμοσφαιρίνη, Αιματοκρίτης, RDW, Λευκά Αιμοσφαίρια, Αιμοπετάλια, MCV, MCH, MCHC
Νεφρική και Ηπατική λειτουργία	Κρεατινίνη, BUN, Αλβουμίνη, Ολική Χολερυθρίνη, ALT, AST
Δείκτες πήξης	INR, PT, PTT
Αέρια αίματος	pH, pCO2, Po2, ΓΑΛΑΚΤΙΚΆ
Ηλεκτρολύτες (με οξεοβασική ισορροπία)	Νάτριο, Κάλιο, Χλώριο, Ασβέστιο, Διττανθρακικά, Anion gap, Γλυκόζη

Για την χαρτογράφηση του πίνακα labevents, δεδομένου πως ο πίνακας δεν περιέχει ονοματικά τα αποτελέσματα των εξετάσεων από τις νοσηλείες των ασθενών παρά μόνον το αναγνωριστικό της itemid χρειάζεται να γίνει η μεσολάβηση του πίνακα d_labitems. Η συνάρτηση κατασκευάζει ένα αντίστροφο λεξικό με όλες τις πιθανές ονομασίες των βιοδεικτών και το χαρτογραφεί στο αντίστοιχο όνομα του βιοδείκτη. Έπειτα αποθηκεύονται οι εγγραφές του d_labitems στο σύνολο target_itemids. Ένα άλλο εμπόδιο είναι το μέγεθος του labevents πίνακα είναι το υπερβολικό μέγεθος του (17,9 GB). Ο πίνακας καθίσταται δύσχρηστος και η φόρτωση του στη μνήμη

συμβατικών υπολογιστικών μονάδων παρουσιάζεται πολλές φορές αδύνατη. Εξαιτίας αυτού πραγματοποιείται η τμηματοποίηση του πίνακα σε 500.000 εγγραφές μέσω της chunksize παραμέτρου της pd.read_csv() συνάρτησης. Εφόσον γίνουν τα παραπάνω, το κάθε τμήμα υφίσταται δύο διαδοχικά φιλτραρίσματα. Στο πρώτο διατηρούνται οι βιοδείκτες ενδιαφέροντος ενώ το δεύτερο περιορίζει τις εγγραφές των εργαστηριακών εξετάσεων σε ένα σύνολο νοσηλειών με το χαρακτηριστικό keep_hadm. Η συνεισφορά των φίλτρων στην μείωση του όγκου δεδομένων είναι τεράστια καθώς οι εκατοντάδες εκατομμύρια εγγραφές μειώνονται σε δεκάδες εκατομμύρια εγγραφές. Παράλληλα οι μοναδικές στήλες που θα φορτωθούν στη μνήμη είναι οι (hadm_id, itemid, valuenym, charttime) και εκείνες που δεν χρησιμοποιούνται ορίζονται ως flag ή σχόλιο για την αποφυγή του κόστους μεταφοράς των πεδίων.

Μια από τις κρισιμότερες ενέργειες που πραγματοποιήθηκαν στη ροή της προ επεξεργασίας των δεδομένων είναι η επιλογή του χρονικού παραθύρου των 24 ωρών. Ο λόγος που πραγματοποιήθηκε η συγκεκριμένη επιλογή είναι η εξασφάλιση κλινικής σκοπιμότητας του μοντέλου, καθώς οι περισσότερες κλινικές αποφάσεις, ειδικά στα άμεσα περιστατικά, λαμβάνονται εντός του πρώτου 24ώρου. Παραδείγματος χάριν, σε περιπτώσεις εμφράγματος η θεραπεία ξεκινάει εντός των πρώτων ωρών, η λεγόμενη «Golden Hour» περίπτωση. Επίσης, τα διεθνώς αναγνωρισμένα προγνωστικά συστήματα, όπως το SAPS II που υπολογίζει τις χειρότερες τιμές των ζωτικών σημείων και το APACHE II που υπολογίζει τον κίνδυνο θνησιμότητας, εστιάζουν στις πρώτες 24 ώρες. Ακόμη, στη περίπτωση που δινόταν πρόσβαση στο μοντέλο σε μεταγενέστερες του 24ώρου τιμές το μοντέλο ενδέχεται να δει αποτελέσματα εξετάσεων που έχουν ήδη διαγνωστεί, οπότε οι μετρικές αξιολογήσεις του μοντέλου θα είναι υπεραισιόδοξες και πλασματικές. Με άλλα λόγια, η προσέγγιση αυτή εξαλείφει τον κίνδυνο του φαινομένου διαρροής της πληροφορίας από το μέλλον (look-ahead bias). Για τον ακριβή υπολογισμό των ωρών που έχει εισαχθεί ενάς ασθενής χρησιμοποιούμε τη νέα στήλη hours_from_admit που υπολογίζεται από τη μαθηματική εξίσωση (22):

$$hours_from_admit = \frac{charttime - admittime}{3600\ sec} \qquad (22)$$

με βάση τον οποίο οι εγγραφές που διατηρούνται είναι αυτές που ικανοποιούν τη συνθήκη $0 \leq hours_{from_{admit}} \leq 24$.

Σ ’ένα παράθυρο 24 ωρών ενδέχεται να υπάρχουν πολλαπλές μετρήσεις ενός βιοδείκτη. Έτσι επιλέχθηκε της πρώτης χρονολογικά διαθέσιμης τιμής. Πρακτικά, ο πίνακας κατηγοριοποιείται στα πεδία hadm_id, lab_standardized και charttime, ενώ εφαρμόζεται η συνάρτηση groupby().first() για να κρατηθεί μια μόνο εγγραφή ανά νοσηλεία. Με την επιλογή αυτή αποτυπώνεται η κλινική βαρύτητα του μοντέλου σε έγκαιρες προβλέψεις. Μετά απ’ όλα αυτά η τελική ενέργεια που πραγματοποιείται είναι η αναδιάταξη των δεδομένων σε ευρεία μορφή, δηλαδή στο pivot_table οι 29 βιοδείκτες παύουν να είναι λέξεις σε μια στήλη και γίνονται τίτλοι στο νέο σύνολο δεδομένων, το lab_wide. Το lab_wide με τη σειρά του αντικατοπτρίζει τη πλήρη κλινική κατάσταση του ασθενή στο πρώτο 24ωρο. Οι ελλιπείς τιμές γεμίζουν τα κενά τους με τιμή

NaN. Αποτέλεσμα της τροποποίησης αυτής είναι η δημιουργία ενός νέου πίνακα χαρακτηριστικών που είναι το τελικό προϊόν προ επεξεργασίας καθώς κάθε γραμμή είναι μία νοσηλεία, κάθε στήλη είναι ένα χαρακτηριστικό και η τελευταία ετικέτα αποτελεί δείκτη ένδειξης της καρδιακής ανεπάρκειας.

5.1.3 Δημιουργία Σύνθετων Ιατρικών Δεικτών

Η παθοφυσιολογία μας αποδεικνύει ότι τα κρίσιμα συμβάντα και οι παθήσεις είναι αποτελέσματα πολύπλοκων αλληλεπιδράσεων μεταξύ διαφορετικών βιολογικών συστημάτων του οργανισμού. Επομένως η αποτελεσματικότητα ενός προγνωστικού μοντέλου με κλινική χρησιμότητα δεν κρίνεται αποκλειστικά από τον όγκο των αρχικών πληροφοριών αλλά την ικανότητα του να προσαρμόζεται σε σύνθετες κλινικές περιπτώσεις. Για την επίτευξη της μέγιστης αποτελεσματικότητας του μοντέλου, προσθέτουμε δώδεκα νέες στήλες με τους **σύνθετους δείκτες** στο πίνακα δεδομένων που προορίζονται για εκπαίδευση. Οι **σύνθετοι δείκτες** εφαρμόζονται στη κλινική καθημερινότητα και προσφέρουν στο μοντέλο μηχανικής μάθησης συμπυκνωμένες κλινικές γνώσεις, από επικυρωμένη έρευνα της βιβλιογραφίας, επιτρέποντας τη δυνατότητα το μοντέλο να «βλέπει» την παθολογία του ασθενούς με κλινική εμπειρία και όχι μόνο να «προβλέπει». Επίσης, η χρήση τους συμβάλλει στην εξοικείωση του ιατρικού προσωπικού με το μοντέλο μηχανικής μάθησης.

Την κατασκευή του νέου πίνακα που περιλαμβάνει τους σύνθετους δείκτες υλοποιεί η συνάρτηση compute_composite_indices(). Ουσιαστικά η συνάρτηση λαμβάνει τον αρχικό πίνακα δεδομένων και τον επιστρέφει εμπλουτισμένο με τις νέες δώδεκα στήλες των σύνθετων δεικτών. Ο πίνακας περιλαμβάνει μερικές “δικλείδες” ασφαλείας, όπως ο υπολογισμός κάθε βιοδείκτη απαιτεί την παρουσία όλων των απαραίτητων βιοδεικτών και η αποφυγή σφαλμάτων διαίρεσης μέσω της μετατροπής μηδενικών παρονομαστών με τιμές NaN.

Οι σύνθετοι βιοδείκτες έχουν χωριστεί σε λόγους δύο βιοδεικτών που είναι απλοί βιοδείκτες που περνούν κλασματική μορφή και οι δείκτες κλινικής βαρύτητας. Οι λόγοι δύο βιοδεικτών είναι οκτώ και έχουν αξιοποιηθεί σε διάφορες κλινικές περιπτώσεις:

1) **Λόγος RDW προς Αλβουμίνη ($RDW/Albumin$)**
   Συμβολίζεται με το ακρωνύμιο RAR (Red Cell Distribution Width to Albunim Ratio) και ορίζεται μαθηματικά ως ο λόγος του εύρους κατανομής ερυθρών αιμοσφαιρίων προς τη συγκέντρωση αλβουμίνης. Στη βιβλιογραφία μας, ο RAR αναδεικνύεται ως ισχυρός προγνωστικός δείκτης θνητότητας σε καρδιοπαθείς.

2) **Λόγος αιμοσφαιρίνης προς RDW ($Hemoglobin/RDW$)**
   Συμβολίζεται με το ακρωνύμιο HRR (Hemoglobin to Red Cell Distribution Width Ratio) και ορίζεται μαθηματικά ως ο λόγος της ποσότητας των αιμοσφαιρίων προς τη ποιοτική ομοιογένεια των ερυθροκυττάρων. Οι χαμηλές τιμές του δείκτη υποδηλώνουν αναιμία και

ανισοκυττάρωση. Στη βιβλιογραφία μας, ο RDW χρησιμοποιείται ως προγνωστικός δείκτης της στεφανιαίας νόσου και της χρόνιας καρδιακής ανεπάρκειας.

3) **Λόγος BUN προς κρεατιτίνης ($BUN/Creatinine$)**
Συμβολίζεται με το ακρωνύμιο BCR (Blood Urea Nitrogen to Creatinine Ratio) και ορίζεται μαθηματικά ως ο λόγος της συγκέντρωσης του αζώτου ουρίας αίματος προς τη συγκέντρωση της κρεατινίνης. Είναι ο κλασικότερος δείκτης για το Καρδιονεφρικό Σύνδρομο, με τον αυξημένο λόγο BCR να υποδεικνύει την μειωμένη αιμάτωση των νεφρών λόγω χαμηλής καρδιακής παροχής.

4) **Λόγος BUN προς αλβουμίνη ($BUN/Albumin$)**
Συμβολίζεται με το ακρωνύμιο BAR (Blood Urea Nitrogen to Albumin Ratio) και ορίζεται μαθηματικά ως ο λόγος της συγκέντρωσης του αζώτου ουρίας αίματος προς τη συγκέντρωση της αλβουμίνης. Συνδυάζει τη νεφρική λειτουργία καταβολικού στρες με τη θρεπτική κατάσταση και τη φλεγμονή. Οι έρευνες στη βιβλιογραφία μας αναδεικνύουν την κρίσιμη συμβολή του ως προγνωστικός δείκτης σοβαρότητας σε ασθενείς με σήψη, πνευμονία και οξεία νεφρική βλάβη, αποτελώντας ισχυρό εργαλείο πρώιμης στρωμάτωσης κινδύνου.

5) **Λόγος αλβουμίνης προς κρεατινίνης ($Albumin/Creatinine$)**
Συμβολίζεται με το ακρωνύμιο ACR (Albumin to Creatinine Ratio) και ορίζεται μαθηματικά ως ο λόγος της συγκέντρωσης της αλβουμίνης προς τη συγκέντρωση της κρεατινίνης. Χρησιμοποιείται στο πλαίσιο των πρώτων εικοσιτεσσάρων ωρών νοσηλείας και ο δείκτης αυτός εκφράζει την αναλογική σχέση μεταξύ της συνθετικής λειτουργίας του ήπατος και των νεφρικών απεκκρίσεων, προσφέροντας μια πλήρη εικόνα της οργανικής επάρκειας. Χαμηλές τιμές του ACR προδιαθέτουν τη προχωρημένη καρδιακή ανεπάρκεια και το ηπατονεφρικό σύνδρομο.

6) **Λόγος κρεατινίνης προς ολική χολερυθρίνη**
Συμβολίζεται με το ακρωνύμιο CBR και μαθηματικά ορίζεται ως ο λόγος της συγκέντρωσης της αλβουμίνης προς τη συγκέντρωση κρεατινίνης προς τη χολερυθρίνη. Εφαρμόζεται κυρίως στη προσέγγιση των ασθενών με συνυπάρχουσα νεφρική και ηπατική δυσλειτουργία. Οι καταστάσεις αυτές κάνουν την εμφάνιση τους σε προχωρημένο στάδιο της καρδιακής ανεπάρκειας, λόγω της φλεβικής συμφόρησης που προκαλείται μεταξύ των δύο οργάνων.

7) **Λόγος ALT προς AST**
Συμβολίζεται με το ακρωνύμιο AAR (AST-to-ALT Ratio) ή αλλιώς αναφέρεται και ως δείκτης De Ritis και ορίζεται ως ο λόγος των τρανσαμινών αλανινοτρανσαμινάσης προς την ασπαρτική τρανσαμινάση. Τα δύο αυτά αποτελούν ηπατικά ένζυμα που

απελευθερώνονται κατά την ηπατοκυτταρική βλάβη. Σε περιπτώσεις που η τιμή είναι κάτω της μονάδας ($ALT < AST$) τότε υποδηλώνεται ισχαιμική ηπατίτιδά, δηλαδή η καρδιά δεν στέλνει αρκετό οξυγόνο στο ήπαρ, οπότε μπορεί να προκληθεί καρδιακή ανεπάρκεια.

8) **Λόγος γαλακτικών προς αλβουμίνη**
Συμβολίζεται με το ακρωνύμιο LAR (Lactate to Albunim Ratio) και ορίζεται μαθηματικά ως το πηλίκο της συγκέντρωσης των γαλακτικών οξέων προς τη συγκέντρωση της αλβουμίνης. Ο λόγος αυτός αποτελεί συνδυασμό δύο δεικτών, έναν ιστικής υποαιμάτωσης και έναν δείκτη συνθετικής ηπατικής λειτουργίας με χρόνια φλεγμονή. Αναδεικνύεται ως προγνωστικός δείκτης θνητότητας σε ασθενείς με σηπτικό σοκ και κυκλοφορική ανεπάρκεια.

Οι τέσσερις δείκτες κλινικής βαρύτητας έχουν δημοσιευθεί ως αυτόνομα προγνωστικά εργαλεία σε κλινικές μελέτες και απαρτίζονται από δυσκολότερους μαθηματικούς τύπους:

1. Ο **δείκτης ABIC** (Age-Bilirubin-INR-Creatinine) σχεδιάστηκε αρχικά για την πρόγνωση αλκοολικής ηπατίτιδας και στο MIMIC-IV χρησιμεύει δείκτης αναπαράστασης για την πολυοργανικής επιβάρυνσης του ασθενούς, καθώς τη ηλικία, με τη λειτουργία του ήπατος (Bilirubin), της πήξης (INR) και των νεφρών (Creatinine). Υπολογίζεται από την μαθηματική εξίσωση (23):

$$ABIC = 0.1 \cdot Age + 0.66 \cdot Bilirubin + 1.71 \cdot INR + 0.3 \cdot Creatinine \quad (23)$$

Στην παρούσα υλοποίηση οι απουσιάζουσες τιμές των τριών παραγόντων της συνάρτησης συμπληρώνονται με φυσιολογικές τιμές αναφοράς (1,0) μέσω της συνάρτησης fillna().

2. Ο **δείκτης ACAG** (Albumin- Corrected-Anion-Gap) μετράει την οξεοβασική ισορροπία. Επειδή η αλβουμίνη είναι αρνητικά φορτισμένη πρωτεΐνη διορθώνει το Anion Gap αποκαλύπτοντας τη πραγματική μεταβολική οξέωση του ασθενούς. Υπολογίζεται από τη μαθηματική εξίσωση (24):

$$ACAG = AG + 2.5 \times (4.5 - Albumin) \qquad (24)$$

3. Ο **δείκτης ALBI** (Albumin-Bilirubin) εξυπηρετεί στην εκτίμηση της συνθετικής λειτουργίας του ήπατος από την εξίσωση (25):

$$ALBI = 0.66 \cdot log_{10}(Bilirubin) - 0.085 \cdot Albumin \qquad (25)$$

4. Ο **δείκτης eGFR** (estimated Glomerular Filtration Rate) ονομάζεται εκτιμώμενος ρυθμός σπειραματικής διήθησης αποτελεί έναν δείκτη με ιδιαίτερα σύνθετη φύση για τρείς

κύριους λόγους. Κατέχει σημαντικό ρόλο στη κλινική σταδιοποίηση της Χρόνιας Νεφρικής Νόσου με βασή τα κριτήρια KDIGO, εμφανίζει τη συνθετότερη δομή από τους προηγούμενους και αποτελεί το κεντρικό διαμεσολαβητή του καρδιονεφρικού συνδρόμου που χαρακτηρίζει σημαντικό ποσοστό των ασθενών με καρδιακή ανεπάρκεια. Παράλληλα, χρησιμοποιεί το ρυθμό σπειραματικής διήθησης για τη νεφρική λειτουργία. Υπολογίζεται από την μαθηματική εξίσωση (26):

$$eGFR = 175 \cdot Creatinine^{-1.154} \cdot Age^{-0.203} \cdot k_{sex} \qquad (26)$$

Με τον συντελεστή φύλου k_{sex} να ισούται με 0.742 για τις γυναίκες και 1.0 για τους άνδρες. Η διαφορά στο συντελεστή μεταξύ των δύο φύλων οφείλεται στη λιγότερη μυική μάζα και στη μικρότερη παραγωγή κρεατινίνης που εμφανίζουν οι γυναίκες από τους άνδρες.

5.1.4 Χρονικά Ασφαλής Κατασκευή Ιστορικού Συννοσηροτήτων

Ένας άλλος σημαντικός παράγοντας που δημιουργείται με την εγγραφή ενός ασθενή σε πολλαπλές νοσηλείες είναι το ιατρικό ιστορικό του. Στο προτεινόμενο σύστημα, ο ασθενής για κάθε νοσηλεία του μπορεί να καθοριστεί αν έπασχε από έξι κλινικές παθήσεις: σακχαρώδη διαβήτη, την αρτηριακή υπέρταση, τη χρόνια νεφρική νόσο, χρόνια αποφρακτική πνευμονοπάθεια, τη στεφανιαία νόσο και τη κολπική μαρμαρυγή. Παρόλα αυτά, η καταγραφή ιστορικών χαρακτηριστικών για την εκπαίδευση του μοντέλου μας ενέχει έναν κίνδυνο προς την εγκυρότητα των αποτελεσμάτων του, ο οποίος είναι η διαρροή δεδομένων από το μέλλον προς το παρελθόν (look-ahead bias). Στην παρούσα ενότητα χρησιμοποιήθηκαν οι τελεστές cummax(), shift(), fillna() για την αποφυγή της διαρροής δεδομένων.

Αρχικά πραγματοποιήθηκε η αντιστοίχιση των ICD για κάθε νοσηρότητα που αναφέρθηκε στο λεξικό COMORB_SETS του backend/main.py. Το MIMIC-IV περιέχει ιατρικά δεδομένα από το 2008 έως το 2019 και την 1η Οκτωβρίου 2015 έγινε η επίσημη μετάβαση από την κωδικοποίηση ICD-9-CM στην ICD-10-CM, με τον νόμο HIPAA. Για την αποφυγή παράλειψης συννοσηροτήτων από το σύνολο δεδομένων, κατά την υλοποίηση του pipeline της μηχανικής μάθησης δημιουργήθηκε μια διπλή λίστα που εξασφαλίζει τη διαχρονική συνοχή των δεδομένων μας. Η μία λίστα αντιπροσωπεύει την κωδικοποίηση ICD-9 και η άλλη την ICD-10. Παραδείγματος χάριν, ο σακχαρώδης διαβήτης αντιστοιχεί στο “250” στο λεξικό COMORB_SETS και στους κωδικούς ICD-10 E10-E14, η αρτηριακή υπέρταση αντιστοιχεί στους ICD-9 κωδικούς 401-405 και στους ICD-10 κωδικούς I10-I15. Οι κωδικοποιήσεις που υλοποιήθηκαν στο κώδικα εξυπηρετούν σημαντικά στην ευστάθεια και την κλινική ερμηνευσιμότητα του μοντέλου. Πρώτα απ’ όλα, παρουσιάζεται μείωση της δια στασιμότητας διότι το MIMIC-IV έχει πάνω από 70.000 μοναδικούς κωδικούς ICD-10, οπότε σε περίπτωση που το μοντέλο προσπαθήσει να εκπαιδευτεί σε κάθε περίπτωση ξεχωριστά θα χρειαζόταν

εκατομμύρια ασθενείς για να έχει στατιστική σημασία ο κάθε κωδικός του. Με την ομαδοποιημένη κωδικοποίηση που εφαρμόστηκε οι κωδικοί συμπεριλαμβάνονται σε 10-20 ουσιαστικές κλινικές μεταβλητές. Παράλληλα με τον τρόπο αυτό αντιμετωπίζεται και η αραιότητα των δεδομένων, γιατί όλες οι σπάνιες περιπτώσεις που ενδέχεται να εμφανίζονται σε αυτούς τους χιλιάδες κωδικούς τοποθετούνται κάτω από μια κοινή ομπρέλα.

Αρχικά στη κάθε κατηγορία συννοσηροτήτων δημιουργείται μια δυαδική στήλη (_this_hadm), η οποία παίρνει τη τιμή 1 αν ο ασθενής λάβει τον διαγνωστικό κωδικό της αντίστοιχης κατηγορίας σε μια συγκεκριμένη νοσηλεία ή 0 αν δεν τον λάβει. Οι υπολογισμοί γίνονται με χρήση των διανυσματικών τεχνικών Boolean που συνδυάζουν τους κωδικούς ICD (icd_version) με τον έλεγχο του προθέματος (startswith) και συναθροίζονται μέσω της συνάρτησης ομαδοποίησης groupby([‘subject_id’,’hadm_id’]).any(). Πρέπει να διευκρινιστεί ότι η στήλη αναπαριστά τη διάγνωση τρέχουσας νοσηλείας όχι κάποιο προγνωστικό χαρακτηριστικό. Για τη μετατροπή των διαγνώσεων σε αξιοποιήσιμα προγνωστικά χαρακτηριστικά πραγματοποιείται μέσω ενός αυστηρού μηχανισμού χρονικής ευθυγράμμισης.

Αναλυτικότερα, η cummax() μέθοδος χρησιμοποιείται στον υπολογισμό της συσσωρευτικής μέγιστης τιμής της δυαδικής στήλης του _this_hadm. Δεδομένου ότι οι τιμές της στήλης βρίσκονται μεταξύ του εύρους [0,1], η συμπεριφορά της συνάρτησης είναι ισοδύναμη με τη λογική σύζευξη OR. Δηλαδή σε περίπτωση εμφάνισης της τιμής 1, που σημαίνει οτι ο ασθενής πάσχει από κάποια συν νοσηρότητα, τότε όλες οι μεταγενέστερες τιμές της στήλης γίνονται 1. Με τον τρόπο αυτό διατηρείται η διάγνωση μιας χρόνιας νόσου στο ιστορικό του ασθενή. Συνεχίζοντας, στο αποτέλεσμα του cummax() μετατοπίζεται κατά μία θέση προς τα κάτω μέσω της συνάρτησης shift(1). Ο λόγος που συμβαίνει αυτό είναι για να αποφευχθεί το μοντέλο να κοιτάξει το ιστορικό των προηγούμενων νοσηλειών και να μην βασιστεί στην τρέχουσα νοσηλεία για να προβλέψει το ρίσκο κινδύνου από ΚΑ σε αυτήν. Επιπλέον, η μετατόπιση των νοσηλειών αφήνει ην πρώτη νοσηλεία του ασθενή είναι κενή, με τιμή NaN. Η συνάρτηση fillna(False) αντικαθιστά στις τιμές NaN με την τιμή 0 ώστε να τηρηθεί η έλλειψη διάγνωσης κάποιας χρόνιας νόσου κατά τη πρώτη επαφή του με το σύστημα.

Η εφαρμογή του προηγούμενη μηχανισμού δεν είναι απλώς θεωρητική. Η σημασία του φαίνεται από την shift(1) και τη ποσοτική σύγκριση μεταξύ της σωστής, με χρήση της shift() και της εσφαλμένης, χωρίς τη χρήση της shift(). Ο παρακάτω **Πίνακας 4.3** παρουσιάζει το ποσοστό καταγραφής εσφαλμένων διαγνώσεων ως προ υπάρχουσες από τωρινές.

Πίνακας 4.3: Ταξινόμηση συννοσηροτήτων με βάση τις νοσηλείες με μεροληψία πρόβλεψης του μέλλοντος

Συννοσηρότητα	**Νοσηλείες με μεροληψία πρόβλεψης του μέλλοντος**	**Ποσοστό**
Αρτηριακή Υπέρταση (prior_htn)	55.440	11.98%
Στεφανιαία Νόσος (prior_cad)	30.094	6.50%

Σακχαρώδης Διαβήτης (prior_dm)	29.890	6.46%
Κολπική Μαρμαρυγή (pror_afib)	23.323	5.04%
Χρόνια Νεφρική Νόσος (prior_ckd)	19.796	4.28%
Χρόνια Αποφρατική Πνευμονοπάθεια (prior_copd)	13.915	3.01%

Επομένως, από τις 546.028 νοσηλείες, οι 125.310 (ποσοστό του 22.95% του συνόλου) θα εμφάνιζαν νοσηλεία με μεροληψία πρόβλεψης του μέλλοντος σε τουλάχιστον μία συν νοσηρότητα εάν παραλειπόταν ο τελεστής shift().

Πέραν των έξι δυαδικών στηλών που δημιουργήθηκαν για τις συν νοσηρότητες, δημιουργήθηκαν δύο πρόσθετα χαρακτηριστικά που σκοπεύουν στην πλήρη κάλυψη του ιατρικού ιστορικού του ασθενούς. Το πρώτο είναι το prior_comborb_count που είναι το άθροισμα των συνολικών συννοσηροτήτων που βαρύνουν τον ασθενή, δημιουργόντας μια πολυνοσηρότητα που συμβάλει στην πρόγνωση του μοντέλου. Το δεύτερο είναι το num_prior_adm, το οποίο υπολογίζεται με τη συνάρτηση cumcount() και εκφράζει τον αριθμό προηγούμενων νοσηλειών του ασθενούς πριν την τρέχουσα. Η ιδιότητα της καθιερωμένης αρχικής τιμής της cumcount() με τη τιμή μηδέν εξυπηρετεί στην υλοποίηση της προ επεξεργασίας των δεδομένων χωρίς περαιτέρω χειρισμούς αλλά και στη αναγνώριση συσχέτισης μεταξύ της συχνότητας προηγούμενων νοσηλειών και κινδύνου επανεισαγωγής. Ένα τελικό βήμα που έχει δημιουργηθεί είναι η προαιρετική ενσωμάτωση του ιστορικού συννοσηροτήτων στην εκπαίδευση του μοντέλου. Το include_history λειτουργεί ως Boolean flag, σε περίπτωση που είναι False, η εκπαίδευση του μοντέλου βασίζεται σε βιοδείκτες ενώ στη περίπτωση True βασίζεται και στο επιπλέον ιστορικό συννοσηροτήτων.

5.1.5 Τελική Διαμόρφωση του Συνόλου Εκπαίδευσης

Τα βήματα των προηγούμενων ενοτήτων οδήγησαν σε ένα οργανωμένο σύνολο δεδομένων των 546.028 νοσηλειών με πίνακα χαρακτηριστικών πλήρως κατασκευασμένο. Πριν γίνει η τροφοδοσία του σε αλγόριθμο μηχανικής μάθησης πρέπει να γίνουν τρείς αλληλένδετες ενέργειες, η διαχείριση των ελλιπών τιμών, η τμηματοποίηση του συνόλου σε μικρότερα υποσύνολα εκπαίδευσης και η εξισορρόπηση των κλάσεων σε θετική και αρνητική κλάση. Στην ενότητα αυτή αναλύονται οι τρεις τεχνικές επεξεργασίας, ή αλλιώς μετασχηματιστές, που χρησιμοποιήθηκαν για τη αντιμετώπιση των παραπάνω, με ιδιαίτερη έμφαση να γίνεται στη σειρά υλοποίηση και αλληλεπίδρασης τους.

Ως σχεδιαστική αρχή της διπλωματικής εφαρμόστηκε πρώτα η τμηματοποίηση και ακολούθως οι μετασχηματισμοί με βάση την κατανομή. Αρχικά πραγματοποιήθηκε η απομόνωση του συνόλου

ελέγχου από το σύνολο εκπαίδευσης. Η διαμέριση τους έγινε μέσω της συνάρτησης group_split_indices του backend/main.py σε δύο διαδοχικές τμηματοποιήσεις με χρήση του μετασχηματισμού GroupShuffleSplit. Στη πρώτη τμηματοποίηση το σύνολο χωρίζεται σε δύο τμήματα, το ένα σύνολο είναι η εκπαίδευση και το καλιμπράρισμα, ενώ το άλλο το σύνολο ελέγχου, σε αναλογία (80/20)% αντίστοιχα. Η δεύτερη τμηματοποίηση γίνεται με το διαχωρισμό του συνόλου εκπαίδευσης και καλιμπραρίσματος σε δύο ξεχωριστά σύνολα, με αναλογία (75/25)%. Επομένως ο τελικός διαχωρισμός αποτελείται από τρία ξεχωριστά σύνολα με ποσοστά 60%, 20% και 20% επί του αρχικού συνόλου. Η επιλογή χαρακτηριστικών κατά την ομαδοποίηση στήριζεται στο subject_id για κλειδί της τμηματοποίησης διότι το 43.3% των ασθενών, όπως αναφέρθηκε στην 5.1.1, διαθέτει πολλαπλές νοσηλείες και για να ικανοποιηθεί η στατιστική ανεξαρτησία μεταξύ των υποσυνόλων. Το κάθε subject_id είναι το μοναδικό χαρακτηριστικό που ανήκει σε ένα και μόνο σύνολο.

Πίνακας 4.4: Ταξινόμηση των τριών υποσυνόλων με την GroupShuffleSplit

Σύνολο	**Νοσηλείες**	**Μοναδικοί Ασθενείς**	**Ποσοστό Καρδιακής Ανεπάρκειας**
Εκπαίδευση	326.741	138.700	12.09%
Βαθμονόμηση	110.003	46.700	12.13%
Έλεγχος	109.284	46.400	12.26%
Σύνολο	546.028	231.770	12.14%

Επαληθεύοντας την ορθότητα του μετασχηματισμού παρατηρήθηκε μηδενική επικάλυψη των ασθενών μεταξύ των τριών υποσυνόλων και τα ποσοστά καρδιακής ανεπάρκειας παρέμειναν σε σταθερό εύρος 12.09-12.26%, παρόλο που ο μετασχηματισμός GroupShuffleSplit δεν εφαρμόζει ρητή στρώματωση προς το χαρακτηριστικό κλειδί subject_id. Η σταθερότητα αυτή οφείλεται στον μεγάλο όγκο δεδομένων που περιλαμβάνουν τα τρία υποσύνολα. Σε μικρότερα υποσύνολα δεδομένων η χρήση StratifiedGroupKFold θα ήταν αναγκαία.

Το επόμενο βήμα της τελικής διαμόρφωσης του συνόλου που προορίζεται για την εκπαίδευση του μοντέλου πραγματοποιείται το «γέμισμα» των κενών που αφήνουν οι ελλιπείς τιμές. Η ύπαρξη ελλιπών τιμών οφείλεται σε απουσία κλινικής ένδειξης όχι μηδενικής μέτρησης. Για παράδειγμα, ο θεράπων ιατρός παραγγέλνει αποκλειστικά τις εξετάσεις που κρίνει αναγκαίες, όχι δεν του είναι όλες οι μετρήσεις χρήσιμες. Για την αντιμετώπιση των κενών χρησιμοποιείται η κλάση SimpleImputer της βιβλιοθήκης scikit-learn με χρήση της διαμέσου τιμών κάθε στήλης και η αντικατάσταση των ελλιπών τιμών της με αυτήν. Η επιλογή της διαμέσου αντί για τον αριθμητικό μέσο όρο στηρίζεται στις λοξές κατανομές των εργαστηριακών δεδομένων. Αξιοσημείωτη είναι η αποκλειστική εφαρμογή της fit_trasnform() μεθόδου στο σύνολο εκπαίδευσης και η εφαρμογή της transform() στα υπόλοιπα σύνολα ώστε να αποφευχθεί η διαρροή δεδομένων.

Το τελικό βήμα είναι η εξασφάλιση της ποιότητας των αποτελεσμάτων του προγνωστικού μοντέλου. Όπως αναδείχθηκε στην ενότητα 5.1.1, το ποσοστό θετικών περιπτώσεων καρδιακής

ανεπάρκειας είναι της κλίμακας του 12.14% επί του συνόλου του. Αυτό αντιστοιχεί σε αναλογία αρνητικών προς θετικών σε 7.24 προς 1, δηλαδή 287.067 αρνητικών έναντι 39.674 θετικών. Η εκπαίδευση ενός μοντέλου σε ένα τέτοιο ανισόρροπο σύνολο παράγει μοντέλα που εστιάζουν στη πλειοψηφία, η οποία εμφανίζει απουσία καρδιακής ανεπάρκειας, κάτι το οποίο προκαλεί τη φαινομενική μεγιστοποίηση της ακρίβειας. Για την αντιμετώπιση της εφαρμόζεται η τεχνική SMOTE που υλοποιείται στη βιβλιοθήκη imbalanced-learn και ενσωματώνεται στο σύνολο εκπαίδευσης αποκλειστικά της ροής προ επεξεργασίας δεδομένων. Η παραγωγή σύνθετων δειγμάτων, αντί της απλής δειγματοληψίας, εμπλουτίζει το σύνολο εκπαίδευσης και διατηρεί την γεωμετρία των θετικών δειγμάτων, θέτοντας την ιδανική για κλινικές εφαρμογές. Στο τελικό σύνολο εκπαίδευσης, η SMOTE παράγει 247.393 συνθετικά δείγματα της θετικής κλάσης, εξισορροπώντας πλήρως τις δύο κλάσεις σε 287.067 δείγματα έκαστη και διογκώνοντας το συνολικό μέγεθος του συνόλου εκπαίδευσης από 326.741 σε 574.134 δείγματα.

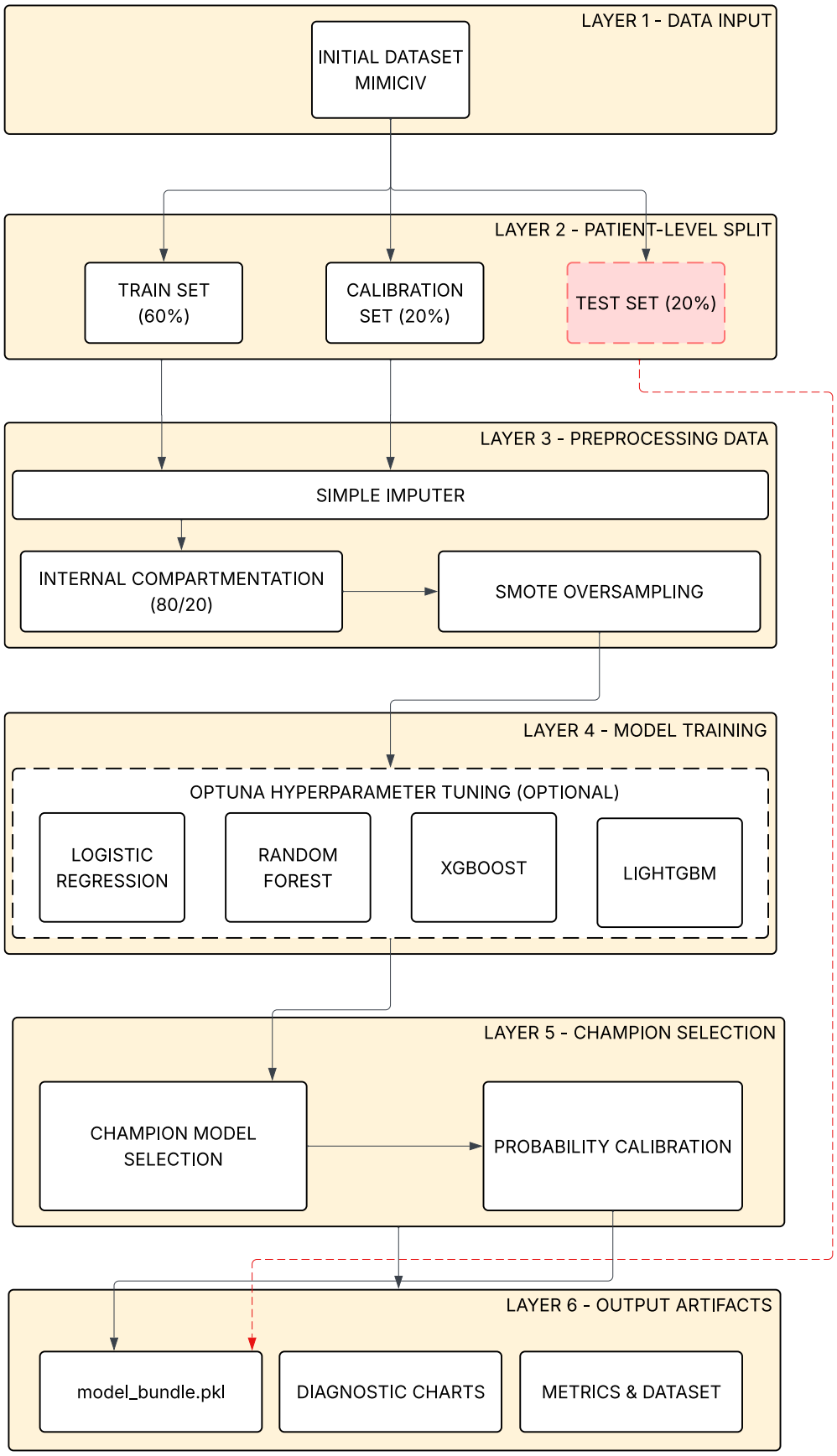
5.2 Εκπαίδευση Μοντέλων Μηχανικής Μάθησης

5.2.1 Αρχιτεκτονική της Ροής Εκπαίδευσης

Η ενότητα αυτή αποσκοπεί στην αποτύπωση της συνολικής αρχιτεκτονικής του pipeline της πλατφόρμας για την καλύτερη κατανόηση του αναγνώστη περί της ροής δεδομένων. Η εκπαίδευση των μοντέλων υλοποιείται στη συνάρτηση train_and_select του backend/main.py. Η συνάρτηση train_and_select εκτελεί μια αλληλουχία βημάτων, ξεκινώντας με τη διαμόρφωση των δεδομένων, την εκπαίδευση πολλαπλών υποψηφίων αλγορίθμων και την επιλογή βέλτιστου μοντέλου με την βαθμονόμηση των τελικών πιθανοτήτων, καλύπτοντας το πλήρες εύρος του πυρήνα της μηχανικής μάθησης της παρούσας διπλωματικής εργασίας.

Εν πρώτοις στο 1ο επίπεδο πραγματοποιείται διαμοιρασμός του αρχικού συνόλου δεδομένων σε επίπεδο ασθενούς, μέσω της συνάρτησης group_split_indices για την παραγωγή τριών διακριτών υποσυνόλων, το υποσύνολο δεδομένων προς εκπαίδευσης (60%), το υποσύνολο καλιμπραρίσματος (20%), το υποσύνολο ελέγχου (20%). Συνεχίζοντας στο 2ο επίπεδο γίνεται η διαχείριση των ελλιπών τιμών μέσω του εργαλείου SimpleImputer της βιβλιοθήκης scikit-learn της Python εφαρμόζοντας ως κύρια στρατηγική την εύρεση της διαμέσου της κάθε στήλης και των τριών συνόλων μέσω του μετασχηματισμού transform. Έπειτα στο 3ο επίπεδο ακολουθεί ο εσωτερικός διαμοιρασμός για πρόωρη παύση (early stopping) του συνόλου εκπαίδευσης. Από το υποσύνολο εκπαίδευσης αποσπάται ένα 20% του για την επικύρωση του, ικανοποιώντας τις αναγκαίες του πρόωρου τερματισμού των ενισχυμένων μοντέλων που αναφέρονται στην ενότητα 5.2.3. Η περαιτέρω ανάλυση προτίμησης του 20% του υποσυνόλου εκπαίδευσης έναντι του υποσυνόλου καλιμπραρίσματος για πρόωρη παύση της εκπαίδευσης τεκμηριώνεται στην προαναφερθείσα ενότητα. Εφόσον τελειώσει η διαδικασία πρόωρης παύσης στο εκπαιδευτικό υποσύνολο εφαρμόζεται η τεχνική υπερδειγματοληψίας SMOTE για την ταυτόχρονη διατήρηση της φυσικής κατανομής και βελτίωσης της ακρίβειας του προγνωστικού μοντέλου. Τα δεδομένα μετά εισέρχονται στο 4ο επίπεδο, η πρώτη ενέργεια είναι η βελτιστοποίηση των υπερπαραμέτρων

μέσω του εργαλείου Optuna, η οποία αναλύεται στην ενότητα 5.2.4 και η χρήση της είναι προαιρετική. Ύστέρα από την παραμετροποίηση ακολουθεί, η εκπαίδευση των τεσσάρων υποψήφιων μοντέλων και η επιλογή ενός μοντέλου πρωταθλητή. Η επιλογή βασίζεται στην αξιολόγηση των μετρικών και αποτελεσμάτων εκπαίδευσης από το υποσύνολο του καλιμπραρίσματος, σύμφωνα με το οποίο πρωταγωνιστής επιλέγεται το μοντέλο με την υψηλότερη τιμή AUROC. Το τελευταίο βήμα του 5^{ου} επιπέδου είναι η βαθμονόμηση των πιθανοτήτων του πρωταγωνιστικού μοντέλου με σκοπό την ταύτιση της καμπύλης των προβλέψεων όσο το δυνατόν γίνεται στη πραγματική συχνότητα εμφάνισης της νόσου. Στο 6^{ο} επίπεδο και τελευταίο το πλήρες εκπαιδευμένο και βαθμονομημένο μοντέλο εφαρμόζεται το σύνολο ελέγχου και παράγονται οι τελικές μετρικές που θα παρουσιαστούν στο Κεφάλαιο 6.


LAYER 1 - DATA INPUT
INITIAL DATASET MIMICIV
LAYER 2 - PATIENT-LEVEL SPLIT
TRAIN SET (60%)
CALIBRATION SET (20%)
TEST SET (20%)
LAYER 3 - PREPROCESSING DATA
SIMPLE IMPUTER
INTERNAL COMPARTMENTATION (80/20)
SMOTE OVERSAMPLING
LAYER 4 - MODEL TRAINING
OPTUNA HYPERPARAMETER TUNING (OPTIONAL)
LOGISTIC REGRESSION
RANDOM FOREST
XGBOOST
LIGHTGBM
LAYER 5 - CHAMPION SELECTION
CHAMPION MODEL SELECTION
PROBABILITY CALIBRATION
LAYER 6 - OUTPUT ARTIFACTS
model_bundle.pkl
DIAGNOSTIC CHARTS
METRICS & DATASET

Εικόνα 31 : Αρχιτεκτονική της διαδικασίας εκπαίδευσης του μοντέλου «Πρωταθλητή». Η διαδικασία καλύπτει την εισαγωγή δεδομένων στο MIMIC-IV έως την παραγωγή του βαθμονομημένου μοντέλου πρωταθλητή και των αντίστοιχων αποτελεσμάτων.

5.2.2 Διαμόρφωση των Τεσσάρων Υποψήφιων Μοντέλων

Ένας αλγόριθμος μηχανικής μάθησης που χρησιμοποιείται για την δημιουργία ενός προγνωστικού μοντέλου οφείλει η ακρίβεια των αποτελεσμάτων του να είναι τεκμηριωμένη. Για τον λόγο αυτόν εκπαιδεύτηκαν παράλληλα τέσσερα μοντέλα που εκπροσωπούν διακριτές οικογένειες της εποπτευόμενης μάθησης. Η επιλογή πολλαπλής υποψηφιότητας μοντέλων έναντι της μεμονωμένης συνέβη για την πειραματική κάλυψη της τελικής επιλογής του μοντέλου πρωταγωνιστή ως αποδεδειγμένου καλύτερου μοντέλου πρόβλεψης της καρδιακής ανεπάρκειας. Οι θεωρητικές αρχές κάθε αλγορίθμου έχουν αναπτυχθεί στο Κεφάλαιο 3 και η παρούσα ενότητα εστιάζει στην επιλογή των παραμέτρων υλοποίησης τους.

Ως πρώτο υποψήφιο μοντέλο παρουσιάζεται η Λογιστική Παλινδρόμηση και υλοποιείται ως αλυσίδα δύο στάδιο μέσω της βιβλιοθήκης scikit-learn.

```
1. Pipeline([
2.     ("scaler", StandardScaler()),
3.     ("lr", LogisticRegression(solver="liblinear",
random_state=cfg.seed)),
4. ])
```

Κώδικας 1 Διαμόρφωση Pipeline της Λογιστικής Παλινδρόμησης

Πριν την εκπαίδευση του αλγορίθμου λογιστικής παλινδρόμησης πραγματοποιείται η διαδικασία της κλιμάκωσης μέσω της συνάρτησης StandardScaler. Ο αλγόριθμος βελτιστοποιεί μια συνάρτηση κόστους που χρησιμοποιεί ευκλείδεια απόσταση στον χώρο των βαρών, δηλαδή σε χαρακτηριστικά που παρουσιάζουν μεγάλη διαφορά στην τάξη μεγέθους, όπως η ηλικία στα δεκάδες έτη έναντι της αλβουμίνης σε δεκαδικές τιμές g/dL, η StandardScaler αποδίδει σε όλα τα χαρακτηριστικά μέση τιμή 0 και τυπική απόκλιση 1 επιτρέποντας στον αλγόριθμο να συγκλίνει γρήγορα και να αποδώσει δίκαια βάρη σε κάθε κλινική παράμετρο. Με τον τρόπο αυτό καλύπτουμε την ισότιμη επεξεργασία των δεδομένων σε όλο το φάσμα του συνόλου εκπαίδευσης, αποφεύγοντας φαινόμενα διαστρέβλωσης των βαρών και αυξημένου υπολογιστικού κόστους. Η μετατροπή του συνόλου εκπαίδευσης σε πρόβλημα δυαδικής ταξινόμησης χρήζει αποδοτικότερη την επιλογή του λύτη liblinear αντί για τον lbfgs που συνηθίζεται η επιλογή του. Ο liblinear εκτελεί συντονιστική κατάβαση, δηλαδή διορθώνει μία παράμετρο τη φορά αντί να προσπαθήσει να αλλάξει όλες τις παραμέτρους του μοντέλου ταυτόχρονα όπως η lbfgs. Έτσι το μοντέλο αντιλαμβάνεται ποιο βάρος πρέπει να μηδενιστεί ή να μειωθεί χωρίς να χρειαστούν πολυπλοκότεροι υπολογισμοί δεύτερης παραγώγου και η αποθήκευση τεράστιων πινάκων. Επιπλέον η ενσωμάτωση τους στη Pipeline εξυπηρετεί στην καθαρότερη ερμηνευσιμότητα από την μέθοδο SHAP ώστε οι τιμές της να υπολογίζονται στο σωστό χώρο χαρακτηριστικών.

Ως δεύτερο υποψήφιο μοντέλο παρουσιάζεται ο αλγόριθμος Τυχαίου Δάσους μέσω της κλάσης RandomForestClassifier. Η κλάση παρουσιάζει τις ακόλουθες υπερπαραμέτρους σύμφωνα με τον

κώδικα:

```
RandomForestClassifier(
    n_estimators=400,
    max_depth=12,
    random_state=cfg.seed,
    n_jobs=-1,
    class_weight="balanced",
)
```

Κώδικας 2 Διαμόρφωση των υπεραπαραμέτρων του RandomForestClassifier

Η επιλογή των παραπάνω παραμέτρων έγινε ύστερα από επανελλειμένες δοκιμές. Αρχικά η επιλογή των 400 δέντρων αποτελεί την «χρυσή» τομή στην ευρωστία και το κόστος, καθώς η αύξηση των δέντρων πέραν αυτού του ορίου παράγει οριακή βελτίωση της μετρικής AUROC με πρόσθετη επιβάρυνση χρόνο εκπαίδευσης. Η επιλογή του βάθους στο 12 αναγκάζει τα κλινικά μοτίβα να ισχύουν σε όλες τις περιπτώσεις, διατηρώντας στην εκπαίδευση όλες τις απαραίτητες λεπτομέρειες και δημιουργώντας μια δικλείδα ασφαλείας στην υπερπροσμαρμογή του μοντέλου. Επίσης, η παράμετρος n_jobs= -1, αποτελεί εντολή εφαρμογής όλων των διαθέσιμων πυρήνων του επεξεργαστή στην παράλληλη εκπαίδευση των δέντρων, εξοικονομώντας χρόνο. Ακόμη, ο ορισμός της παραμέτρου class_weight ως ισορροπημένη (“balanced”) αποτελεί τη κρισιμότερη επιλογή διότι αντισταθμίζει την ανισορροπία μεταξύ των κλάσεων και λειτουργεί ενισχυτικά στην τεχνική SMOTE. Ουσιαστικά εξασφαλίζει την ανάθεση ανάλογου βάρους στις πραγματικές περιπτώσεις σε περίπτωση που η εξισορρόπηση των συνθετικών δεδομένων του SMOTE δεν γίνει τέλεια.

Ως τρίτο υποψήφιο μοντέλο παρουσιάζει ο αλγόριθμος Ακραίας Ενίσχυσης Κλίσης, ή αλλιώς XGBoost και εφαρμόζεται στη κλάση XGBClassifier. Οι υπερπαράμετροι που τέθηκαν στην εκπαίδευση του είναι οι εξής με βάση τον παρακάτω κώδικα:

```
XGBClassifier(
    n_estimators=800,
    learning_rate=0.05,
    max_depth=4,
    subsample=0.9,
    colsample_bytree=0.9,
    objective="binary:logistic",
    eval_metric="auc",
    tree_method="hist",
    scale_pos_weight=(1 - pos_ratio) / pos_ratio,
```

```
11.     random_state=cfg.seed,
12. )
```

Κώδικας 3 Διαμόρφωση των υπερπαραμέτρων του XGBClassifier

Η ειδοποιός διαφορά μεταξύ του αλγορίθμου Τυχαίου Δάσους και της Ακραίας Ενίσχυσης Κλίσης είναι στην εκπαίδευση των δέντρων. Στο Τυχαίο Δάσος, τα δέντρα εκπαιδεύονται τυχαία το ένα με το άλλο, ενώ στον αλγόριθμο Ακραίας Ενίσχυσης Κλίσης το κάθε δέντρο δεν είναι αυτόνομο αλλά όταν εκπαιδεύεται διορθώνει τα σφάλματα των προηγούμενων. Η διόρθωση γίνεται σταδιακά και απαιτείται μεγαλύτερος αριθμός δειγμάτων για την επαρκή σύγκλιση του μοντέλου και την ορθότητα του. Με την μείωση του ρυθμού μάθησης στο μισό επιτυγχάνεται η ορθότερη πρόβλεψη και αποφεύγεται ο θόρυβος. Ακόμη, η ρηχή σε βάθος εκπαίδευση των δέντρων του XGBoost (max_depth=4) βασίζεται στη δομική λειτουργία του αλγορίθμου ώστε το τελικό μοντέλο να έχει γίνει «σοφότερο» και να μην έχει επιβαρύνει λειτουργικά το σύστημα. Στον αλγόριθμο Τυχαίου Δάσους, τα δέντρα λειτουργούν ως αυτόνομη εκτιμητές και πρέπει να έχουν βάθος οι γνώσεις του κάθε δέντρου, για αυτό και τα λιγότερα δέντρα. Ο XGBoost όμως δουλεύει σειριακά και τα δέντρα λειτουργούν σε μικρότερο βάθος με περισσότερες διορθώσεις. Η μέθοδος χρησιμοποιεί ιστογραφική προσέγγιση των συνεχών χαρακτηριστικών με βάση τη παράμετρο tree_method=”hist” και επιτυγχάνεται μηδενική απώλεια ακρίβειας, ιδιαίτερα σημαντική στις κλινικές εφαρμογές. Για παράδειγμα, σε αρκετά ιατρικά δεδομένα παρατηρούνται πολύ μικρές διαφορές στην τιμές των εργαστηριακών εξετάσεων οι οποίες μπορεί να προκαλέσουν θόρυβο. Με την ιστογραφική προσέγγιση ομαδοποιούνται παρόμοιες τιμές ώστε το μοντέλο να μην κάνει διαχωρισμό σε ασήμαντες λεπτομέρειες και να εστιάζει στα δεδομένα που έχουν μεγαλύτερη βαρύτητα για την πρόβλεψη της καρδιακής ανεπάρκειας. Τέλος ο τελεστής scale_pos_weight λειτουργεί αντισταθμιστικά στη ανισορροπία των κλάσεων, δρώντας στο εσωτερικό της συνάρτησης κόστους του αλγορίθμου.

Ως τέταρτο υποψήφιο μοντέλο παρουσιάζεται ο αλγόριθμος LightGBM που υλοποιείται στη κλάση του LGBMClassifier. Οι παράμετροι ορίζονται στο κώδικα:

```
1. LGBMClassifier.
 2. LGBMClassifier(
 3.     n_estimators=800,
 4.     learning_rate=0.05,
 5.     max_depth=-1,
 6.     subsample=0.9,
 7.     colsample_bytree=0.9,
 8.     objective="binary",
 9.     scale_pos_weight=(1 - pos_ratio) / pos_ratio,
10.     random_state=cfg.seed,
```

```
11.     verbose=-1,
12. )
```

Κώδικας 4 Διαμόρφωση των υπερπαραμέτρων του LGBMClassifier

Αρχικά η τιμή της παραμέτρου max_depth=-1 σηματοδοτεί απουσία στο ανώτατο όριο βάθους. Αυτό άλλωστε αποτελεί και σχεδιαστική αρχή της LightGBM καθώς η κύρια διαφορά με την XGBoost είναι πως εκείνη αναπτύσσει τα δέντρα κατά φύλλο αντί επιπέδου. Με την στρατηγική αυτήν παράγονται δέντρα βαθύτερα τα οποία εκπαιδεύονται σε μεγαλύτερη πληροφορία. Η πολυπλοκότητα του αλγορίθμου ρυθμίζεται κυρίως από τις υπερπαραμέτρους num_leaves και min_child_samples που προσαρμόζονται αναλόγως στο Optuna. Επιπλέον, η παράμετρος verbose=-1 περιορίζεται τις πολυάριθμες κλήσης κατά τη βελτιστοποίηση των παραμέτρων και τη δημιουργία δέντρων. Ιδιαίτερη παρατήρηση χρήζει η επιλογή κοινών αρχικών τιμών στις βασικές υπερπαραμέτρους των δύο ενισχυμένων μοντέλων n_estimators=800, learning_rate=0.05, subsample=0.9, colsample_bytree=0.9. Ο κύριος λόγος είναι επειδή οι δύο αλγόριθμοι ενίσχυσης ανήκουν στην ίδια οικογένεια και η εκκίνηση της εκπαίδευσης από το ίδιο σημείο επιτρέπει στον αναγνώστη να δει τις διαφορές στην αποδοτικότητα που εμφανίζουν. Στο τέλος και οι δύο αλγόριθμοι βελτιστοποιούνται μέσω του εργαλείου Optuna και επιστρέφεται η καλύτερη δυνατή εκδοχή τους.

5.2.3 Μηχανισμός Πρόωρου Τερματισμού των Ενισχυμένων Μοντέλων

Στους αλγορίθμους ενισχυμένης εκπαίδευσης , την Ακραία Ενίσχυση Κλίσης και τη Μηχανή Ελαφριάς Ενίσχυσης Κλίσης, η ρύθμιση των υπερπαραμέτρων για την εκπαίδευση των δέντρων αποτελεί την κρισιμότερη επιλογή. Το μοντέλο κινδυνεύει να υποπροσαρμοστεί σε περίπτωση που η ποσότητα των δέντρων είναι ελάχιστη ή να υπερπροσαρμοστεί αν η ποσότητα των δέντρων είναι υπερβολική. Στην πρώτη περίπτωση το μοντέλο δεν προλαβαίνει να μάθει τα μοτίβα των δεδομένων, ενώ στη δεύτερη τα δέντρα μετά από ένα σημείο παύουν να παράγουν χρήσιμη γνώση και αρχίζουν την απομνημόνευση θορύβου. Για την αποφυγή αυτών των καταστάσεων εφαρμόστηκε ο πρόωρος τερματισμός που αποτελεί καθιερωμένη λύση στο δίλημμα, δηλαδή αντί να καθορίζεται ντετερμινιστικά ο αριθμός των δέντρων εκ των προτέρων η εκπαιδευτική διαδικασία παρακολουθείται από ένα ξεχωριστό σύνολο επικύρωσης και τερματίζεται αυτόματα μόλις παρατηρηθεί παύση στη βελτίωση της απόδοσης.

Γιατί η επιλογή ενός νέου εσωτερικού συνόλου επικύρωσης;

Ο κύριος λόγος που εφαρμόστηκε προσθέτως διαχωρισμός στο σύνολο εκπαίδευσης παρά την ήδη ύπαρξη συνόλου καλιμπραρίσματος δεν είναι άλλος από την αποφυγή ύπουλης υπερπροσαρμογής του μοντέλου κατά την εκπαίδευση του. Η επιλογή του συνόλου καλιμπραρίσματος θα παραβίαζε την αρχή της απομόνωσης, επηρεάζοντας τον αριθμό δέντρων του κάθε ενισχυμένου μοντέλου διότι το ίδιο σύνολο θα χρησιμοποιούνταν μετά στην επιλογή του μοντέλου πρωταγωνιστή αλλά και στην βαθμονόμηση των τελικών πιθανοτήτων. Επίσης ο διαχωρισμός πραγματοποιήθηκε σε

ίδια αναλογία θετικών και αρνητικών δειγμάτων, αλλά σε επίπεδο νοσηλείας όχι επίπεδο ασθενούς που είχε γίνει μέσω του GroupShuffleSplit, όπως είχε γίνει στα προηγούμενα δύο υποσύνολα εκπαίδευσης και ελέγχου ώστε το αποτέλεσμα της μετρικής AUC να υπολογίζεται μέσω του αντιπροσωπευτικού δείγματος και να αντικατοπτρίζεται η πραγματική κατανομή των κλάσεων. Ο πρόωρος τερματισμός γίνεται πριν την εφαρμογή της υπερδειγματοληψίας SMOTE και μέσω της κλάσης train_test_split της scikit-learn βιβλιοθήκης με στρωμάτωση.

```
1. X_tr_inner, X_es, y_tr_inner, y_es = train_test_split(
2.     X_train, y_train,
3.     test_size=0.20,
4.     random_state=cfg.seed,
5.     stratify=y_train,
6. )
```

Κώδικας 5 Διαχωρισμός συνόλου εκπαίδευσης για πρόωρο τερματισμό

Η αναλογία διαχωρισμού του συνόλου εκπαίδευσης σε 80/20 δεν είναι τυχαία. Ένα μεγαλύτερο σύνολο πρόωρης παύσης, της τάξης των 30% για παράδειγμα, θα παρείχε στατιστικά σταθερότερη εκτίμηση της μετρικής AUC ανά γύρο αλλά θα μείωνε τον κίνδυνο πρόωρου τερματισμού λόγω του θορύβου από τον υπολογισμό. Το σύνολο της πρόωρης παύσης με την αναλογία 80/20 είναι περίπου 55.700 δείγματα. Για την υλοποίηση του XGBoost σε πρόωρη παύση πραγματοποιείται μέσω της συνάρτησης _xgb_fit_with_early_stopping. Σημαντική είναι η διαχείριση της παραμέτρου early_stopping_rounds ως όρισμα της μεθόδου fit(), ενώ σε νεότερες έχει μετακινηθεί στον κατασκευαστή XGBClassifier. Ο λόγος αυτής της ρύθμισης είναι η δυνατότητα ευκολότερων παρεμβάσεων στο μέλλον.

```
 1. fit_params = inspect.signature(XGBClassifier.fit).parameters
 2. early_stopping_in_fit = "early_stopping_rounds" in
fit_params
 3.
 4. if early_stopping_in_fit:
 5.     model.fit(X_tr, y_tr, eval_set=[(X_va, y_va)],
 6.               verbose=False, early_stopping_rounds=50)
 7. else:
 8.     model.set_params(early_stopping_rounds=50)
 9.     model.fit(X_tr, y_tr, eval_set=[(X_va, y_va)],
verbose=False)
10.
```

Κώδικας 6 Πρόωρος τερματισμός εκπαίδευσης XGBoost

Για την υλοποίηση στο LightGBM ο πρόωρος τερματισμός υλοποιείται μέσω του μηχανισμού ανακλήσεων της βιβλιοθήκης.

```
1. model.fit(
2.     X_fit, y_fit,
3.     eval_set=[(X_es, y_es)],
4.     eval_metric="auc",
5.     callbacks=[lgb.early_stopping(50, verbose=False)],
6. )
```

Κώδικας 7 Πρόωρος τερματισμός εκπαίδευσης LightGBM

Η τιμή αυτή αντιπροσωπεύει το λεγόμενο παράθυρο υπομονής δηλαδή πόσοι διαδοχικοί γύροι χωρίς βελτίωση στη μετρική AUC του συνόλου πρόωρης παύσης απαιτούνται προτού η εκπαίδευση τερματιστεί. Η τιμή 50 παρατηρείται ως καθιερωμένη επιλογή στη βιβλιογραφία ταξινόμησης με ενισχυμένη κλίση και προσφέρει ισορροπία στην ευρωστία και την υπολογιστική αποδοτικότητα.

5.2.4 Βελτιστοποίηση Υπερπαραμέτρων μέσω Optuna

Το Optuna αποτελεί μια ανοιχτού κώδικα βιβλιοθήκη γραμμένη στη γλώσσα Python που χρησιμοποιείται για την βελτιστοποίηση υπερπαραμέτρων και κατασκευάστηκε από την Preferred Networks το 2018, μια ιαπωνική εταιρεία που δραστηριοποιείται στην εφαρμογή της βαθιάς μάθησης σε διάφορα πεδία. Στην παρούσα διπλωματική εργασία η χρήση της βιβλιοθήκης περιορίζεται στη στην συστηματική αναζήτηση καλύτερων παραμέτρων στα μοντέλα των αλγορίθμων XGBoost και LightGBM, εξαιτίας της μεγαλύτερης ευαισθησίας που εμφανίζουν σε σχέση με την Λογιστική Παλινδρόμηση και το Τυχαίο Δάσος.

Η συγκεκριμένη βιβλιοθήκη αποτέλεσε εξαιρετική επιλογή έναντι των πολυεφαρμοσμένων προσεγγίσεων της Εξαντλητικής Αναζήτησης Πλέγματος (GridSearchCV) και της Τυχαίας Αναζήτησης (RandomizedSearchCV). Στην πρώτη μέθοδο ο ντετερμινιστικός χαρακτήρας της καθίσταται πρόβλημα σε αλγορίθμους με μεγάλο αριθμό υπερπαραμέτρων, όπως ο XGBoost λόγω του εκθετικού υπολογιστικού κόστους. Από την άλλη, η Τυχαία Αναζήτηση παρουσιάζεται αποδοτικότερη από την Εξαντλητική Αναζήτηση Πλέγματος αλλά αγνοεί την ιστορία των προηγούμενων δοκιμών. Το Optuna εστιάζει λειτουργεί με την εφαρμογή πολλαπλών δοκιμών (trials) και εστιάζει στα αποτελέσματα τους, κατευθύνοντας τις επόμενες σε υποσχόμενες περιοχές του χώρου για να ελαχιστοποιηθεί το υπολογιστικό κόστος και να επιτευχθούν καλύτερα αποτελέσματα παράλληλα.

Στην συγκεκριμένη υλοποίηση της διπλωματικής εργασίας το Optuna χρησιμοποιεί δύο αλγορίθμους Μπευζιανής Βελτιστοποίησης (Bayesian Optimazation), τον εκτιμητή Tree-Structured Parzen (TPE) και τον MedianPruner. Ο εκτιμητής TPE αποτελεί τον λειτουργικό πυρήνα του Optuna και κατηγοριοποιεί την συνάρτηση στόχου σε κλάσμα δύο διακριτών

συναρτήσεων πυκνότητας πιθανότητας. Στον αριθμητή παρουσιάζεται η κατανομή παραμέτρων με εξαιρετικά αποτελέσματα, για παράδειγμα υψηλή τιμή στη μετρική AUROC, ενώ στο παρανομαστή παρουσιάζεται η κατανομή παραμέτρων με φτωχότερη απόδοση. Ο σκοπός αυτού του αλγορίθμου βελτιστοποίησης είναι η μεγιστοποίηση του λόγου των δύο συναρτήσεων, θέτοντας τον ιδανικό για κλινικά δεδομένα που χαρακτηρίζονται από μικτή φύση στις παραμέτρους τους, δηλαδή την ύπαρξη συνεχών, διακριτών και λογαριθμικών χαρακτηριστικών ταυτόχρονα. Παράλληλα, ο MedianPruner, αποτελεί διακόπτη πρόωρης παύσης των δοκιμών σε περίπτωση που η ενδιάμεση απόδοση ελαχιστοποιηθεί σε σχέση με την διάμεσο των προηγούμενων αποδόσεων. Ο συνδυασμός τους παρουσιάζεται παρακάτω. Η επιλογή των πρώτων 5 δοκιμών να μείνουν ανέπαφες και το 10% των συνολικών δοκιμών να πραγματοποιηθούν χωρίς κίνδυνο πρόωρης παύσης, ώστε να έχει ο αλγόριθμος να έχει μια αξιόπιστη βάσης σύγκρισης μεταξύ «καλών» και «κακών» δοκιμών.

```
1. sampler = optuna.samplers.TPESampler(seed=opt_cfg.seed)
2. pruner  = optuna.pruners.MedianPruner(
3.     n_startup_trials=max(5, opt_cfg.trials // 10),
4.     n_warmup_steps=1,
5. )
6. study = optuna.create_study(direction="maximize",
sampler=sampler, pruner=pruner, ...)
```

Κώδικας 8 Αρχικοποίηση TPESampler και MedianPruner του Optuna

Για τον χώρο αναζήτησης του XGBoost χρησιμοποιείται η συνάρτηση tune_xgb_with_optuna η οποία ορίζει 11 υπερπαραμέτρους προς βελτιστοποίηση. Οι υπερπαράμετροι έχουν επιλέγει με ιδιαίτερη προσοχή για την ύπαρξη βέλτιστων τιμών σε όλους. Αρχικά παρατηρείται η εφαρμογή λογαριθμικής κλίμακας (log=True) στις παραμέτρους learning_rate, min_child_weight, gamma, reg_alpha και reg_lambda. Η επιλογή αυτή βασίζεται στην φύση των παραμέτρων, καθώς οι μικρές αλλαγές σε μικρές τιμές τους μπορεί να έχει μεγάλο αντίκτυπο οδηγώντας σε υπερσυχνή δειγματοληψία μεγάλων τιμών και υπό συχνή δειγματοληψία μικρών τιμών αντίστοιχα. Οι παράμετροι subsample και colsample δειγματοληπτούνται σε γραμμική κλίμακα στο εύρος [0.5, 1.0] αφού είναι ποσοστά και έχουν αναλογική επίδραση. Οι παράμετροι reg_alpha και reg_lambda αντιστοιχούν στις παραμέτρους κανονικοποίησης L1 και L2 αντίστοιχα, αποτελούν όρια για την αποφυγή της απομνημόνευσης κάθε λεπτομέρειας από το μοντέλο. Τέλος, η παράμετρος gamma ορίζει ένα ελάχιστο κέρδος στην νέα διακλάδωση στο δέντρο εκπαίδευσης του αλγόριθμο. Αν η gamma εμφανίζει μεγάλη τιμή, το δέντρο παρουσιάζεται ρηχό και συντηρητικό, αντιθέτως με την χαμηλή τιμή του που το χαρακτηρίζει ως βαθύ και υπεργενικευμένο.

```
 1. params: Dict[str, object]  = {
 2.     "n_estimators":      trial.suggest_int("n_estimators",
300, 3000),
 3.     "learning_rate":
trial.suggest_float("learning_rate", 0.005, 0.3, log=True),
 4.     "max_depth":         trial.suggest_int("max_depth", 2,
10),
 5.     "min_child_weight":
trial.suggest_float("min_child_weight", 0.1, 30.0, log=True),
 6.     "subsample":         trial.suggest_float("subsample",
0.5, 1.0),
 7.     "colsample_bytree":
trial.suggest_float("colsample_bytree", 0.5, 1.0),
 8.     "gamma":             trial.suggest_float("gamma", 1e-8,
10.0, log=True),
 9.     "reg_alpha":         trial.suggest_float("reg_alpha",
1e-4, 10.0, log=True),
10.     "reg_lambda":        trial.suggest_float("reg_lambda",
1e-3, 100.0, log=True),
11.     "max_delta_step":    trial.suggest_int("max_delta_step",
0, 10),
12.     ...
13. }
```

Κώδικας 9 Χώρος αναζήτησης των υπερπαραμέτρων του XGBoost μέσω της βιβλιοθήκης Optuna

Για τον χώρο αναζήτησης του LightGBM χρησιμοποιείται η συνάρτηση tune_lgbm_with_optuna και ορίζει επίσης 11 υπερπαραμέτρους. Η μια από τις δύο κύριες διαφοροποιήσεις από τις υπερπαραμέτρους του XGBoost είναι η ανάπτυξη του δέντρου ανά φύλλο αντί για ανά επίπεδο, θέτοντας την παράμετρο num_leaves ως την κύρια παράμετρο για τον έλεγχο της πολυπλοκότητας του μοντέλου. Η άλλη είναι η παράμετρος min_child_samples που ουσιαστικά ορίζει τον ελάχιστο αριθμό δειγμάτων που απαιτούνται σε έναν τερματικό κόμβο, ενώ στον XGBoost ελέγχεται ο ελάχιστος αριθμός βαρών του.

```
 1. params: Dict[str, object] = {
 2.             "n_estimators":
trial.suggest_int("n_estimators", 300, 5000),
 3.             "learning_rate":
trial.suggest_float("learning_rate", 0.005, 0.2, log=True),
 4.             "num_leaves": trial.suggest_int("num_leaves",
15, 255),
```

```
            "max_depth": trial.suggest_int("max_depth", -1,
12),
            "min_child_samples":
trial.suggest_int("min_child_samples", 5, 150),
            "subsample": trial.suggest_float("subsample",
0.5, 1.0),
            "subsample_freq":
trial.suggest_int("subsample_freq", 1, 10),
            "colsample_bytree":
trial.suggest_float("colsample_bytree", 0.5, 1.0),
            "reg_alpha": trial.suggest_float("reg_alpha",
1e-4, 10.0, log=True),
            "reg_lambda": trial.suggest_float("reg_lambda",
1e-3, 100.0, log=True),
            "min_split_gain":
trial.suggest_float("min_split_gain", 1e-8, 1.0, log=True),
            "objective": "binary",
            "random_state": opt_cfg.seed,
            "n_jobs": -1,
            "scale_pos_weight": scale_pos_weight,
            "verbose": -1,
        }
```

Κώδικας 10 Χώρος αναζήτησης των υπερπαραμέτρων LightGBM μέσω της βιβλιοθήκης Optuna

Σημαντική παρατήρηση αξίζει να δοθεί στην σχεδιαστική επιλογή της υλοποίησης group-based cross-validation σε κάθε δοκιμή του Optuna συμβάλλει στην αποφυγή διαρροής δεδομένων στο σύνολο καλιμπραρίσματος ή στο σύνολο ελέγχου, στην διατήρηση των ασθενών με πολλαπλές νοσηλείες σε μια ομάδα με την χρήση του subject_id.

```
for fold_i, (tr_idx, va_idx) in enumerate(splitter.split(X,
y, groups=groups), start=1):
            imp = SimpleImputer(strategy="median")
            X_tr = imp.fit_transform(X.iloc[tr_idx])
            y_tr = y[tr_idx]
            X_va = imp.transform(X.iloc[va_idx])
            y_va = y[va_idx]

            if cfg.use_smote and SMOTE_AVAILABLE:
                sm = SMOTE(random_state=cfg.seed)
                X_tr, y_tr = sm.fit_resample(X_tr, y_tr)

```

```
12.             model = XGBClassifier(**params)
13.             _xgb_fit_with_early_stopping(model, X_tr, y_tr,
X_va, y_va)
14.
15.             proba = model.predict_proba(X_va)[:, 1]
16.             score = _score_for_optuna(opt_cfg.metric, y_va,
proba)
17.             fold_scores.append(score)
```

Κώδικας 11 Βρόχος σταυροεπαλήθευσης και εκπαίδευσης μοντέλου ανά fold

Μετά το πέρας της χρήσης του Optuna και για την προβολή των τελικών αποτελεσμάτων του έχει κατασκευαστεί η συνάρτηση _score_for_optuna. Η συνάρτηση αυτή εντοπίζει τις προβλέψεις του μοντέλου και τις μετατρέπει σε τιμή των μετρικών AUPRC, AUROC, BRIER σκορ. Για την μετρική AUPRC επιστρέφει τη μέση βαθμολογία ακρίβειας (average_precision_score), για την μετρική AUROC επιστρέφει την περιοχή κάτω από την καμπύλη ROC και για τη μετρική BRIER επιστρέφει την τιμή Brier.

```
 1. def _score_for_optuna(metric: str, y_true: np.ndarray,
proba: np.ndarray) -> float:
 2.     """Return a scalar score for Optuna optimization.
 3.
 4.     For AUROC/AUPRC: higher is better (maximize).
 5.     For Brier: lower is better; we return negative Brier so
we can maximize consistently.
 6.     """
 7.     metric = metric.upper().strip()
 8.     if metric == "AUROC":
 9.         return float(roc_auc_score(y_true, proba)) if
len(np.unique(y_true)) > 1 else float("nan")
10.     if metric == "AUPRC":
11.         return float(average_precision_score(y_true, proba))
if len(np.unique(y_true)) > 1 else float("nan")
12.     if metric == "BRIER":
13.         return -float(brier_score_loss(y_true, proba))
14.     raise ValueError(f"Unknown metric: {metric}")
```

Κώδικας 12 Συνάρτηση υπολογισμού της μετρικής σκορ του Optuna

5.2.5 Επιλογή Μοντέλου Πρωταθλητή

Μετά την φάση εκπαίδευσης του επιπέδου 4, όπως μας δείχνει και την **Εικόνα 31,** υπάρχουν 4 πλήρως εκπαιδευμένα μοντέλα, πιθανών και με βελτιστοποιημένες παραμέτρους μέσω Optuna, η Λογιστική Παλινδρόμηση, το Τυχαίο Δάσος, το XGBoost και το LightGBM. Από αυτά, ένα μόνο επιλέγεται ως πρωταγωνιστής και προωθείται για την αξιολόγηση, τη βαθμονόμηση και την εξαγωγή του για παραγωγική χρήση. Στην ενότητα αυτή αναλύεται εκτενώς η διαδικασία επιλογής που βασίζεται σε μερικές κρίσιμες μεθοδολογικές επιλογές.

Η θεμελιώδης αρχή που διέπει την διαδικασίας επιλογής του μοντέλου είναι πως κανένας από τα μοντέλα δεν αξιολογούνται για τη βελτίωση τους από το υποσύνολο ελέγχου, αλλά αξιολογείται από το υποσύνολο καλιμπραρίσματος. Το υποσύνολο ελέγχου δεν πρέπει να συμβάλει, ούτε έμμεσα ούτε άμεσα στη βελτιστοποίηση του μοντέλου, διότι στη περίπτωση που το συνέβαλε, η μετρική του πρωταγωνιστή θα ήταν πλασματική και συστηματικά μεροληπτική αυξημένη. Το φαινόμενο αυτό αναφέρεται στη βιβλιογραφία ως «test set peeking» και αναφέρεται ως μια από τις δημοφιλέστερες αιτίες υπεραισιόδοξων αποτελεσμάτων σε πειραματικές μετρήσεις ερευνών που έχουν θέμα την μηχανική μάθηση. Για τον λόγω αυτό, όπως αναφέρθηκε στην ενότητα 5.1, το υποσύνολο εκπαίδευσης διασπάται και ένα κομμάτι του 20% γίνεται υποσύνολο καλιμπραρίσματος για την αξιολόγηση και βαθμονόμηση των πιθανοτήτων του μοντέλου πρωταθλητή.

Η διαδικασία επιλογής του μοντέλου πρωταθλητή ξεκινάει από την συνάρτηση train_and_select, στην οποία μετά από την εκπαίδευση των υποψηφίων μοντέλων εκτελείται η πρόβλεψη επί του συνόλου καλιμπραρίσματος και έπειτα από υπολογισμούς παράγονται πέντε μετρικές μέσω της compute_metrics.

```
1. proba_calib = model.predict_proba(X_calib)[:, 1]
2.         m = compute_metrics(y_calib, proba_calib)
3.         m["model"] = name
4.         metrics_rows.append(m)
5.         trained[name] = model
```

Κώδικας 13 Αξιολόγηση των υποψήφιων μοντέλων στο σύνολο βαθμονόμησης

Οι μετρικές που παράγονται είναι η AUROC, AUPRC, Brier Score, Ακριβεια και F1 score. Με την σειρά, η AUROC αποτυπώνει την διακριτική ικανότητα του μοντέλου, η AUPRC εστιάζει στη θετική κλάση, το Brier Score εκτιμά την ειλικρίνεια των πιθανοτήτων του μοντέλου, η Ακρίβεια (Accuracy) και το F1 Score αντιπροσωπεύουν μετρικές της απόδοσης του μοντέλου σε σταθερό κατώφλι και δείχνουν πως λειτουργεί το μοντέλο στο πραγματικό κόσμο. Από τις πέντε μετρικές ως κριτήριο επιλογής χρησιμοποιείται η AUROC. Η επιλογή της δικαιολογείται από τον στόχο του μοντέλου να διαχωρίζει αποτελεσματικότερα τις κλάσεις, δηλαδή να έχει μεγαλύτερη διακριτική ικανότητα. Πρωτίστως είχε χρησιμοποιηθεί η μετρική AUPRC διότι στόχος ήταν η μεγιστοποίηση της ικανότητας εντοπισμού της θετικής κλάσης σε ανισόρροπα δεδομένα για τη

βελτιστοποίηση του μοντέλου από το Optuna. Τα αποτελέσματα των μετρικών αποθηκεύονται στο αρχείο model_comparison_metrics.csv, κατά φθίνουσα τάξη AUROC, ώστε η επανεξέταση τους από κάποιον ερευνητή να είναι απρόσκοπτη. Συμπληρωματικά με την αριθμητική σύγκριση γίνεται και η οπτική σύγκριση μέσω των συναρτήσεων save_model_comparison_roc, η οποί παράγει γράφημα με επικαλυπτόμενες ROC καμπύλες (model_comparison_roc.png).

```
 1. for name, model in trained_models.items():
 2.             try:
 3.                 proba = model.predict_proba(X)[:, 1]
 4.                 fpr, tpr, _ = roc_curve(y, proba)
 5.                 auc = roc_auc_score(y, proba) if
len(np.unique(y)) > 1  else float("nan")
 6.                 plt.plot(fpr, tpr, label=f"{name}
(AUC={auc:.3f})")
 7.                 any_plotted = True
 8.             except Exception as e:
 9.                 print(f"WARNING: ROC plot skipped for
{name}: {e}")
```

Κώδικας 14 Σχεδίαση καμπύλων ROC των υποψήφιων μοντέλων

Το γράφημα αυτό χρησιμεύει ως εργαλείο τεκμηρίωσης των σχεδιαστικών επιλογών της ροής εκπαίδευσης των μοντέλων μηχανικής μάθησης και εξηγείται περαιτέρω στο Κεφάλαιο 6. Η συνάρτηση train_and_select αποθηκεύει το μοντέλο που επιλέγει ως πρωταθλητής στο αρχείο model_bundle.pkl και κρατάει παράλληλα στη μνήμη της τα τέσσερα εκπαιδευμένα μοντέλα σε μορφή λεξικού σε περίπτωση που επιθυμεί η ανάκληση τους για τη σύγκριση τους.

5.2.6 Βαθμονόμηση Πιθανοτήτων του Μοντέλου Πρωταθλητή

Η επιλογή του πρωταγωνιστικού μοντέλου αποτελεί το πρώτο βήμα προς την παραγωγή ενός εκπαιδευμένου ταξινομητή που εκτιμά πιθανότητες θετικής κλάσης. Παρά την εκτιμόμενη διακριτική ικανότητα του, με βάση τη μετρική AUROC που διατυπώθηκε στην ενότητα 5.2.5, είναι κρίσιμο οι πιθανότητες που εξάγει το μοντέλο να ανταποκρίνονται σε πραγματικές πιθανότητες εμφάνισης καρδιακής ανεπάρκειας. Στην ενότητα αυτή παρουσιάζεται το τελικό βήμα της ροής μηχανικής μάθησης, όπου το μοντέλο πρωταθλητής υφίσταται βαθμονόμηση επί των πιθανοτήτων του για να επιβεβαιωθεί η άμεση κλινική ερμηνευσιμότητα του.

Ένας από τους κυριότερος λόγους που η βαθμονόμηση πιθανοτήτων του μοντέλου πρωταθλητή κρίνεται απαραίτητη είναι ότι οι αλγόριθμοι ενισχυμένης κλίσης και ο αλγόριθμος Τυχαίου

Δάσους σε μικρότερο βαθμό παράγουν υπερβολικές αυτοεκτιμήσεις στις πιθανότητες τους. Αυτό προκαλείται από την προσπάθεια των μοντέλων να ελαχιστοποιήσουν τη συνάρτηση σφάλματος της Λογαριθμικής Απώλειας (Log-Loss). Στην περίπτωση της παρούσας εργασίας ομώς που προτείνεται ένα διαγνωστικό εργαλείο πρέπει η πρόβλεψη κινδύνου του σε κλινικό περιβάλλον να είναι ακριβής και όχι κατά προσέγγιση. Επιπλέον ιδιαίτερη επιλογή παρουσιάζεται η βαθμονόμηση μέσω της κλάσης CalibratedClassifierCV της scikit-learn. Η κλάση CalibratedClassifierCV λειτουργεί ως εκτιμητής των πιθανοτήτων του εκπαιδευμένου μοντέλου και τροποποιεί τις επεμβαίνει όπου χρειαστεί για να αντικατοπτρίζουν οι πιθανότητες στη συχνότητα εμφάνισης ενός γεγονότος του πραγματικού κόσμου. Η κλάση επανεκπαιδεύει τον βασικό ταξινομητή με τη μέθοδο της αν επικύρωσης (cross-validation) επί του συνόλου του, κάτι που προκαλεί καταστροφικά αποτελέσματα στη περίπτωση μας καθώς θα παραγόταν ένα τελείως διαφορετικό μοντέλο με νέες παραμέτρους, αποκλειστικά εκπαιδευμένο από το υποσύνολο καλιμπραρίσματος κάνοντας το μοντέλο πιο «χαζό». Επίσης, θα ήταν άσκοπος οι υπολογιστικοί πόροι που ξοδεύτηκαν για την ήδη εκπαίδευση και βαθμονόμηση του μοντέλου πρωταθλητή. Για να αποφευχθούν οι δύο προηγούμενες καταστάσεις εφαρμόστηκε η κλάση FrozenEstimator η οποία διατηρεί όλες τις μαθημένες παραμέτρους και λειτουργεί ως αμετάβλητος, επιτρέποντας μόνο στη συνάρτηση μετασχηματισμού των πιθανοτήτων της CalibratedClassifierCV να εκπαιδεύεται επί του υποσυνόλου καλιμπραρίσματος. Εξίσου σημαντική είναι η προσθήκη συμβατότητας με παλαιότερες εκδόσεις της βιβλιοθήκης scikit-learn μέσω της εξαίρεσης, όπως παρατηρείται στον παρακάτω κώδικα.

```
1. try:
2.         calibrator = 
CalibratedClassifierCV(estimator=FrozenEstimator(base_model), 
method=cfg.calibration_method)
3.     except Exception:
4.         # Older scikit-learn fallback
5.         calibrator = CalibratedClassifierCV(base_model, 
method=cfg.calibration_method, cv="prefit")
6.     calibrator.fit(X_calib, y_calib)
```

Κώδικας 15 Βαθμονόμηση των πιθανοτήτων μέσω της CalibratedClassifierCV

Στην παρούσα διπλωματική επιλέχθηκε η Σιγμοειδής μέθοδος ως παραμετρική μέθοδο βαθμονόμησης για την τοποθετήσει τις μη βαθμονομημένες εξόδους του βασικού ταξινομητή σε έγκυρε κατανομές πιθανοτήτων. Αν και ο κώδικας της εργασίας επιτρέπει την εναλλαγή σε ισοτονική μέθοδο, οι 110.003 νοσηλείες του υποσυνόλου καλιμπραρίσματος προκαλούν την εμφάνιση του κινδύνου υπερπροσαρμογής. Ακόμη η Σιγμοειδής μέθοδος παράγει ομαλές κατανομές, χωρίς απότομα άλματα ή ασυνέχειες που παρατηρούνται σε μη παραμετρικές μεθόδους όπως η Ισοτονική Παλινδρόμηση. Γενικότερα στα διαγνωστικά εργαλεία οι απότομες

ασυνέχειες στις πιθανότητες δεν είναι θεμιτές. Επιπλέον, η Σιγμοειδής μέθοδος επιτυγχάνει χαμηλότερο Brier σκορ από την Ισοτονική.

Με την ολοκλήρωση του pipeline, ο μετασχηματιστής προ επεξεργασίας, ο βασικός εκτιμητής, το μοντέλο δευτερογενούς βαθμονόμησης και τα μεταδεδομένα του σχήματος σειριοποιούνται στο model_bundle.pkl. Το πακέτο αυτό αποκτά πολλαπλή χρησιμότητα καθώς μπορεί να φορτώνεται σε μελλοντικά αιτήματα πρόβλεψης από το FastAPI και εξυπηρετεί στην υλοποίηση των SHAP διαγραμμάτων, στην οποία χρειαζόμαστε άμεση πρόσβαση στο μοντέλο εσωτερικά του.

```
1. bundle = {
2.         "imputer": imputer,
3.         "base_model": base_model,
4.         "calibrator": calibrator,
5.         "feature_names": feature_cols,
6.         "best_model_name": best_model_name,
7.         "model_version": cfg.model_version,
8.         "seed": cfg.seed,
9.     }
```

Κώδικας 16 Σειριοποίηση του αρχείου που περιλαμβάνει το συμπιεσμένο εκπαιδευμένο μοντέλο model_bundle.pkl

5.3 Ερμηνευσιμότητα μέσω των Διαγραμμάτων SHAP

5.3.1 Αρχιτεκτονική Ενσωμάτωσης της μεθόδου SHAP στο Pipeline

Η ενσωμάτωση της μεθόδου ερμηνευσιμότητας SHAP αποτελεί μια από τις πιο ενδιαφέροντες προσθήκες στο σύστημα υποστήριξης κλινικών αποφάσεων που προτείνεται. Η ερμηνευσιμότητα που προσφέρει η μέθοδος παράγει διπλό έργο λειτουργίας στο σύστημα μας. Η πρώτη λειτουργία είναι η παραγωγή στατιστικών γραφημάτων τεκμηρίωσης των αποτελεσμάτων και η δεύτερη αφορά την παραγωγή δυναμικών γραφημάτων για κάθε αίτημα πρόβλεψης ασθενούς με βάση τα δεδομένα εισαγωγής του. Η υβριδική λειτουργία του συστήματος το προορίζει ως ιδανικό για εμπορική και ερευνητική χρήση.

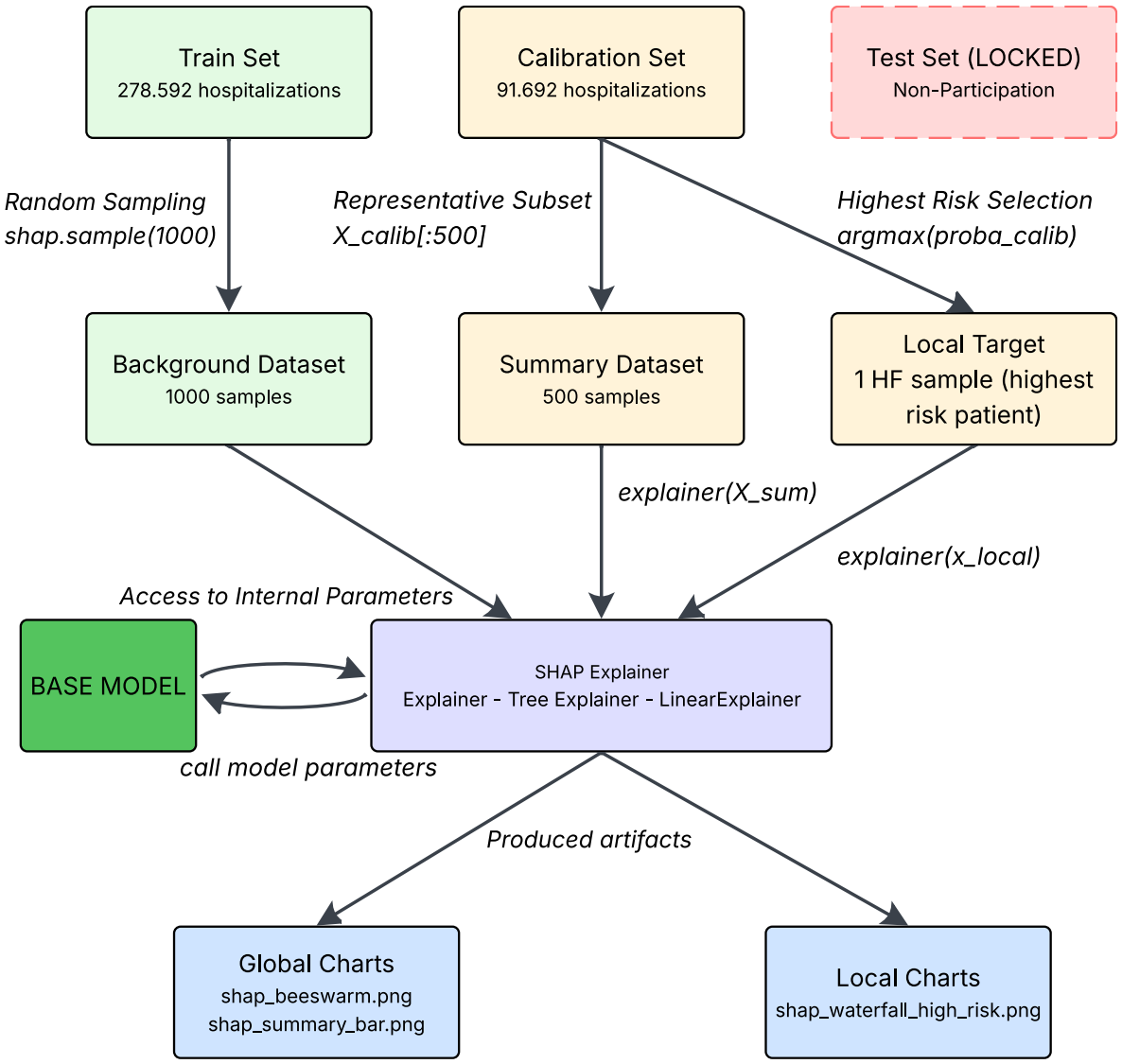
Μετά την ολοκλήρωση της βαθμονόμησης που έγινε στην ενότητα 5.2.6 πραγματοποιείται η δημιουργία SHAP διαγραμμάτων στη βοηθητική συνάρτηση _make_shap_plots, την οποία ενθυλακώνει η train_and_select. Η τοποθέτηση της συνάρτησης _make_shap_plots στο συγκεκριμένο σημείο της γονικής train_and_select δεν είναι τυχαία καθώς στόχος της είναι η εξαγωγή των ερμηνείων SHAP από το σύνολο βαθμονόμησης (σύνολο καλιμπραρίσματος) και όχι το εκπαιδευτικό σύνολο. Σε περίπτωση που οι τιμές SHAP παράγονταν από το σύνολο εκπαίδευσης, η μέθοδος SHAP θα ερμηνεύει τα ήδη απομνημονευμένα δεδομένα του μοντέλου και θα υπέκυπτε στο ενδεχόμενο υπερπροσαρμογής, ενώ με την επιλογή του συνόλου

βαθμονόμησης, λόγω της τυχαιότητας του, η μέθοδος ερμηνεύει τις αποφάσεις στο πραγματικό κόσμο, βασιζόμενο σε νέους ασθενείς κάθε φορά. Πιο απλά, οι τιμές SHAP θα αποτυπώνουν την πραγματική γενικευτική ικανότητα του μοντέλου. Παράλληλα, η κλήση του Explainer πραγματοποιείται όχι στο σύνολο της βαθμονόμησης αλλά στο Base μοντέλο μας. Ο λόγος είναι πως ο TreeExplainer εκτελεί αναδρομική διάσχιση όλων των πιθανών διαδρομών των δέντρων ταυτόχρονα. Εάν εξεταζόταν μόνο στο CalibratedClassifierCV, ο TreeExplainer δεν θα είχε πρόσβαση στα δέντρα του μοντέλου παρά μόνο στην σιγμοειδή συνάρτηση βαθμονόμησης που το εκφράζει μαθηματικά Ο TreeExplainer λειτουργεί με ακριβείς υπολογισμούς στους κόμβους διαχωρισμού, τη δομή των κλαδιών, τα βάρη των φύλλων των δέντρων και η φύση του «μαύρου κουτιού» που έχει το δέντρο θα οδηγούσε τον αλγόριθμο σε αδιέξοδο. Αντί λοιπόν να εφαρμοστεί ο KernelExplainer που λειτουργεί κατά προσέγγιση και απαιτεί χιλιάδες τυχαίες προσομοιώσεις για την ακριβή ανάγνωση των χαρακτηριστικών των δέντρων, ο TreeExplainer χρησιμοποιεί την Fast Exact Tree Attribution μέθοδο, η οποία πραγματοποιεί αναδρομική διάσχιση και υπολογίζει την ακριβή συνεισφορά κάθε χαρακτηριστικού στη τελική πρόβλεψη. Περαιτέρω ανάλυση γίνεται στην ενότητα 5.3.2.

Το διάγραμμα ροής υλοποίησης της μεθόδου SHAP παρουσιάζει πως δουλεύει ο αλγόριθμος στη πλήρη εκδοχή του. Αρχικά παρατηρείται ότι βασίζεται σε τρία ξεχωριστά σύνολα δεδομένων που το καθένα έχει διαφορετική λειτουργία. Το αναφορικό σύνολο (Background Dataset) έχει σύνολο 1000 δείγματα που δειγματοληπτούνται τυχαία στο σύνολο εκπαίδευσης και χρησιμοποιείται από τους Explainers για τον υπολογισμό της αναμενόμενης πρόβλεψης όταν ένα χαρακτηριστικό λείπει. Δεύτερον, το σύνολο εξήγησης (Summary Dataset) έχει 500 δείγματα τα οποία δειγματοληπτούνται από το σύνολο βαθμονόμησης και χρησιμοποιείται στον υπολογισμό των τιμών SHAP για τα γραφήματα beeswarm και importance bar. Η επιλογή του σύνολο βαθμονόμησης έναντι του συνόλου εκπαίδευσης οφείλεται στην γενικευσιμότητα που πρέπει να έχουν τα μοντέλα και να ανταποκρίνονται σε δεδομένα που δεν έχουν ξαναδεί. Τρίτον και τελευταίο ο στόχος τοπικής εξήγησης (Test Set) περιλαμβάνει τον ασθενή με το μεγαλύτερο ποσοστό υψηλού κινδύνου καρδιακής ανεπάρκειας στο σύνολο βαθμονόμησης. Στο waterfall διάγραμμα η χρήση του ασθενή με το μεγαλύτερο ποσοστό κινδύνου δείχνει με σαφήνεια τους παράγοντες που ώθησαν το μοντέλο μακριά από τη μέση τιμή. Η επιλογή των 1000 δειγμάτων για αναφορικό σύνολο και 500 δειγμάτων ως σύνολο εξήγησης δεν είναι τυχαία. Με την επιλογή των 1000 δειγμάτων στο αναφορικό σύνολο καλύπτεται επαρκώς η διαδικασία, χωρίς αυξημένα υπολογιστικά κόστη και τα 500 δείγματα είναι οριακά πάνω από το όριο πριν κορεστή το διάγραμμα beeswarm με θόρυβο αντί για ευκρίνεια.

Μόλις ολοκληρωθεί η εκτέλεση του SHAP Explainer η συνάρτηση _make_shap_plots παράγει τρία αρχεία εικόνων. Το shap_beeswarm.png και shap_summary_bar.png αρχείο αποτελούν εικόνες που φέρουν καθολικά γραφήματα, ενώ το shap_waterfall_high_risk.png αποτελεί εικόνα που φέρει τοπικό γράφημα. Τα καθολικά γραφήματα εξηγούν την συνολική στρατηγική του μοντέλου στη λήψη αποφάσεων και τα τοπικά αναλύουν μεμονωμένες αποφάσεις. Καταρχάς, το διάγραμμα SHAP Beeswarm απεικονίζει την συμβολή των χαρακτηριστικών σε μια πρόβλεψη.

Στο διάγραμμα αυτό η κατανομή λαμβάνει οριζόντια, το κόκκινο χρώμα συμβολίζει την υψηλή τιμή και το μπλε χρώμα συμβολίζει τη χαμηλή τιμή. Όταν οι τιμές τους βρίσκονται στην θετική πλευρά του άξονα αυξάνεται ο κίνδυνος και στην αρνητική πλευρά του άξονα μειώνεται ο κίνδυνος. Η πυκνότητα του γραφήματος δείχνει τη ποσότητα των ασθενών. Επιπλέον, το διάγραμμα SHAP Summary Bar από το αρχείο shap_summary_bar.png αποτελεί ραβδόγραμμα που υπολογίζει την μέση απόλυτη τιμή επίδρασης της κάθε μεταβλητής στο μοντέλο και την κατατάσσει τα χαρακτηριστικά του συνόλου εξήγησης κατά φθίνουσα σειρά. Αποτελεί το πρώτο γράφημα δημιουργεί ο αλγόριθμος και ταξινομεί ορθά τις κλινικές μεταβλητές, όπως η ηλικία, η αιμοσφαιρίνη, η κρεατινινή και άλλοι πολλοί, αφαιρώντας τυχόν θόρυβο που παρεμβάλλεται στην διαδικασία της εκπαίδευσης. Ακόμη το διάγραμμα SHAP Waterfall αποτελεί το κύριο ιατρικό εργαλείο που περιγράφει την διάγνωση του ασθενή με υψηλό κίνδυνο εμφάνισης καρδιακής ανεπάρκειας που εντόπισε το μοντέλο και τον αναπαριστά εικονικά μέσω της διεπαφής της πλατφόρμας. Στο κάτω μέρος του διαγράμματος το E[f(X)] παρουσιάζεται η αναμενόμενη τιμή του μοντέλου, η οποία είναι η τιμή στην οποία το μοντέλο δεν έχει κανένα δεδομένο για τον ασθενή. Είναι με άλλα λόγια η τιμή εκκίνησης που ξεκινάνε οι αλληλεπιδράσεις των βιολογικών χαρακτηριστικών εμφάνισης καρδιακής ανεπάρκειας του ασθενή. Κοινά χαρακτηριστικά αποτελεί η ιδιότητα των χρωμάτων, της αναμενόμενης τιμής και των βιολογικών χαρακτηριστικών, αποτελώντας και τα τρία μαζί διαγράμματα, τα οποία προσφέρουν διαφάνεια και ευκρινής οπτικοποίηση των διαγνώσεων γεφυρόντας εν μέρη το χάσμα μεταξύ κλινικών ιατρών και των πληροφοριακών συστημάτων τεχνητής νοημοσύνης.



Εικόνα 32 : Διάγραμμα παραγωγής στατικών γραφημάτων SHAP από τη συνάρτηση _make_shap_plots()

5.3.2 Τεχνικά Ζητήματα Υλοποίησης

Η μέθοδος SHAP παρουσιάζει μερικά τεχνικά ζητήματα που οφείλονται στο ετερογενές σύνολο μοντέλων. Στην ενότητα αυτή αναφέρεται τις προκλήσεις που προκύπτουν αλλά και τους μηχανισμούς ασφαλείας που δημιουργήθηκαν για την αντιμετώπιση τους.

Για την αποδοτικότερη χρήση του αλγορίθμου η βιβλιοθήκη της μεθόδου SHAP προσφέρει διαφορετικούς Explainers με διαφορετικά πεδία εφαρμογής, ώστε να καλύπτεται πλήρως το χάσμα της υπολογιστικής πολυπλοκότητας. Η πλειοψηφία των αλγορίθμων που δοκιμάζονται βασίζονται σε δέντρα απόφασης, όπως ο XGBoost, ο LightGBM, ο Τυχαίου Δάσους, με πολυωνυμική πολυπλοκότητα και υπάρχει η Λογιστική Παλινδρόμηση που είναι γραμμικό μοντέλο. Επομένως,

έχει δημιουργηθεί ένας αυτόματος ταξινομητής Explainer που προσπαθεί να επιλέξει αυτόματα τον κατάλληλο είδος explainer που ταιριάζει με βάση τον αλγόριθμο εκπαίδευσης του μοντέλου.

```
 1. try:
 2.             explainer = shap.Explainer(model_for, X_bg_for,
feature_names=feature_names)
 3.         except Exception:
 4.             try:
 5.                 explainer = shap.TreeExplainer(model_for)
 6.             except Exception:
 7.                 explainer = shap.LinearExplainer(model_for,
X_bg_for)
 8.
 9.         shap_values = explainer(X_sum_for)
```

Κώδικας 17 Δημιουργία SHAP Explainer ανά τύπο μοντέλου

Η λειτουργία αυτή εξυπηρετεί σε δύο σκοπούς. Αν αποτύχει η αυτόματη ανίχνευση, όπως συμβαίνει σε ορισμένες εκδόσεις SHAP που έχουν δυσκολίες με νεότερες εκδόσεις XGBoost ή LightGBM, ο TreeExplainer ορίζεται ως σταθερή και σωστή επιλογή. Παράλληλα, το σύστημα είναι ταχύτερο καθώς ο Explainer επιλέγει συνήθως αυτόματα τον βέλτιστο Explainer διασφαλίζοντας την ολοκλήρωση της ροής με τουλάχιστον explainer.

Σε δυαδικά προβλήματα ταξινόμησης, ιδιαίτερα στον TreeExplainer, παρατηρείται πως ορισμένοι Explainers θα επιστρέψουν θετικές και αρνητικές τιμές SHAP, ενώ μερικοί επιστρέφουν μόνο θετικές τιμές. Η διαφορά αυτή πρέπει να αντισταθμιστεί ώστε να διασφαλιστεί πως τα τελικά γραφήματα αναφέρονται στη θετική κλάση πάντοτε. Για τον λόγο αυτόν δημιουργήθηκε η συνάρτηση _shap_pick_positive_class, η οποία ελέγχει τις διαστάσεις του πίνακα τιμών SHAP και στο ενδεχόμενο ύπαρξης τρίτης διάστασης διατηρείται μόνο η τομή που αντιστοιχεί στη θετική κλάση. Με αυτή τη ρύθμιση η ερμηνεία των γραφημάτων παραμένει ομοιόμορφη και κλινικά ορθή.

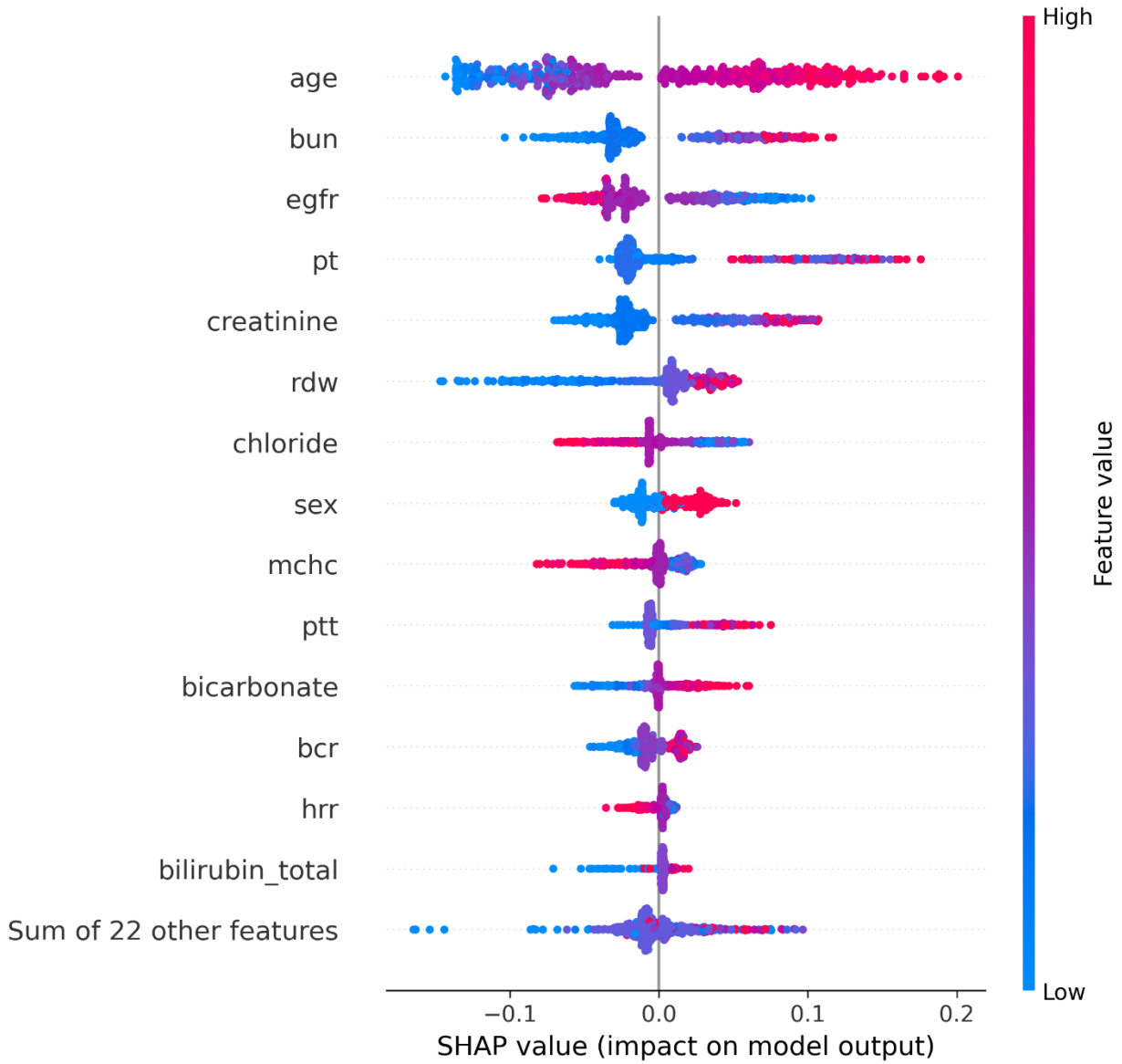
```
1. def _shap_pick_positive_class(exp):
2.     try:
3.         vals = getattr(exp, "values", None)
4.         if vals is not None and hasattr(vals, "ndim") and
vals.ndim == 3 and vals.shape[-1] >= 2:
5.             return exp[..., 1]
6.     except Exception:
```

```
            pass
    return exp
```

Κώδικας 18 Επιλογή θετικής κλάσης σε τιμές SHAP

5.3.3 Στατικά Γραφήματα SHAP

Η συνάρτηση _make_shap_plots παρήγαγε τρία γραφήματα εξηγήσιμης τεχνητής νοημοσύνης. Η ενότητα αυτή επιδιώκει την αναλυτική αιτιολόγηση των σχεδιαστικών επιλογών, των παραμέτρων και των κρίσιμων λεπτομερειών που επηρέασαν τα τελικά γραφήματα SHAP. Το πρώτο από τα γραφήματα που παράγεται μέσω της συνάρτησης shap.summary_plot έχει τύπο εμφάνισης dot και επίσημη υλοποίηση του beeswarm της βιβλιοθήκης SHAP. Πρώτον , παράμετρος max_display=20 περιορίζει την απεικόνιση των χαρακτηριστικών στα 20 κατά μέση απόλυτη τιμή SHAP. Δεύτερον το dpi=160 εξυπηρετεί στην οπτική ευκρίνεια και τη βέλτιστη χωρητικότητα για την αποθήκευση του διαγράμματος. Η συνάρτηση tight_layout() προσαρμόζει τις παραμέτρους των γραφημάτων ώστε οι ετικέτες των αξόνων, οι τίτλοι και οποιαδήποτε άλλα στοιχεία του γραφήματος να μην επικαλύπτονται, ενώ η παράμετρος bbox_inches='tight' περικόπτει αυτόματα το περιττό λευκό περιθώριο. Το συγκεκριμένο παράγει συμπυκνωμένες και σημαντικές κλινικές πληροφορίες με το κάθε σημείο να αντιπροσωπεύει έναν ασθενή και την οριζόντια θέση του να κωδικοποιεί η τιμή SHAP, με το ανάλογο χρώμα για το ποσοστό της τιμής.



Εικόνα 33 : Παράδειγμα απεικόνισης Beeswarm γραφήματος SHAP με γνώμονα τους βιοδείκτες συνόλου βαθμονόμησης που εκπαιδεύτηκε το μοντέλο

```
 1. try:
 2.             plt.figure(figsize=(10, 8))
 3.             shap.plots.beeswarm(shap_values, show=False,
max_display=15)
 4.             plt.tight_layout()
 5.             plt.savefig(outdir / "shap_beeswarm.png",
dpi=300, bbox_inches="tight")
```

```
            plt.close()
        except Exception:
            # Compatibility fallback
            plt.figure(figsize=(10, 8))
            shap.summary_plot(
                getattr(shap_values, "values", shap_values),
                X_sum_for,
                feature_names=feature_names,
                show=False,
                max_display=15,
            )
            plt.tight_layout()
            plt.savefig(outdir / "shap_beeswarm.png",
dpi=300, bbox_inches="tight")
            plt.close()
```

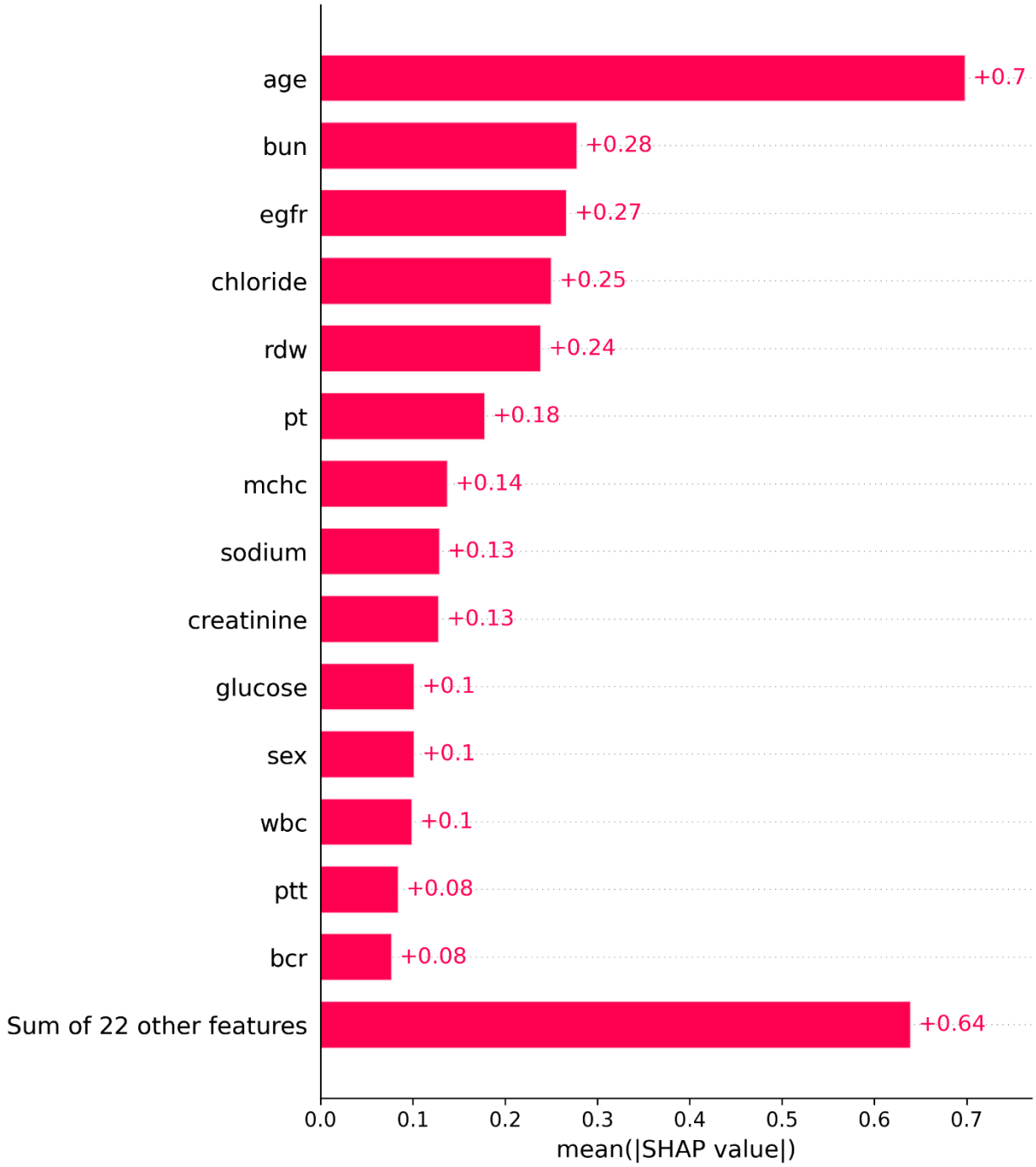
Κώδικας 19 Παραγωγή γραφήματος Beeswarm SHAP

Το διάγραμμα σπουδαιότητας (importance bar) αποτελεί το δεύτερο γράφημα που παράγει η συνάρτηση _make_shap_plots, έχει τύπο εμφάνισης bar και απεικονίζει τη μέση απόλυτη τιμή SHAP ανά χαρακτηριστικό. Η μορφή του διαγράμματος είναι το ραβδόγραμμα και χαρακτηρίζει από τη καθαρή ιεραρχική κατάταξη των χαρακτηριστικών, συν τη πληροφορία κατεύθυνσης τους. Είναι το πιο οικείο γράφημα για άτομα που έρχονται πρώτη φορά σε επαφή με τα SHAP διαγράμματα και στη παρούσα διπλωματική έχει καθολικό χαρακτήρα, δηλαδή ιεραρχεί τα χαρακτηριστικά που εκπαιδεύτηκε το μοντέλο επί του βαθμονομημένου σύνολού. Ιδιαιτερότητα των δύο καθολικών γραφημάτων που αναφέρθηκαν προηγουμένως είναι πως αλληλοσυμπληρώνονται, με το ραβδόγραμμα ως πρώτη εικόνα κατάταξης χαρακτηριστικών και το beeswarm για βαθύτερη κατανόηση της επιρροής των κατευθυνόμενων χαρακτηριστικών του γραφήματος.

```
try:
            plt.figure(figsize=(10, 8))
            shap.plots.bar(shap_values, show=False,
max_display=15)
            plt.tight_layout()
            plt.savefig(outdir / "shap_importance_bar.png",
dpi=300, bbox_inches="tight")
            plt.close()
        except Exception:
```

```
            plt.figure(figsize=(10, 8))
            shap.summary_plot(
                getattr(shap_values, "values", shap_values),
                X_sum_for,
                feature_names=feature_names,
                show=False,
                plot_type="bar",
                max_display=15,
            )
            plt.tight_layout()
            plt.savefig(outdir / "shap_importance_bar.png",
dpi=300, bbox_inches="tight")
            plt.close()
```

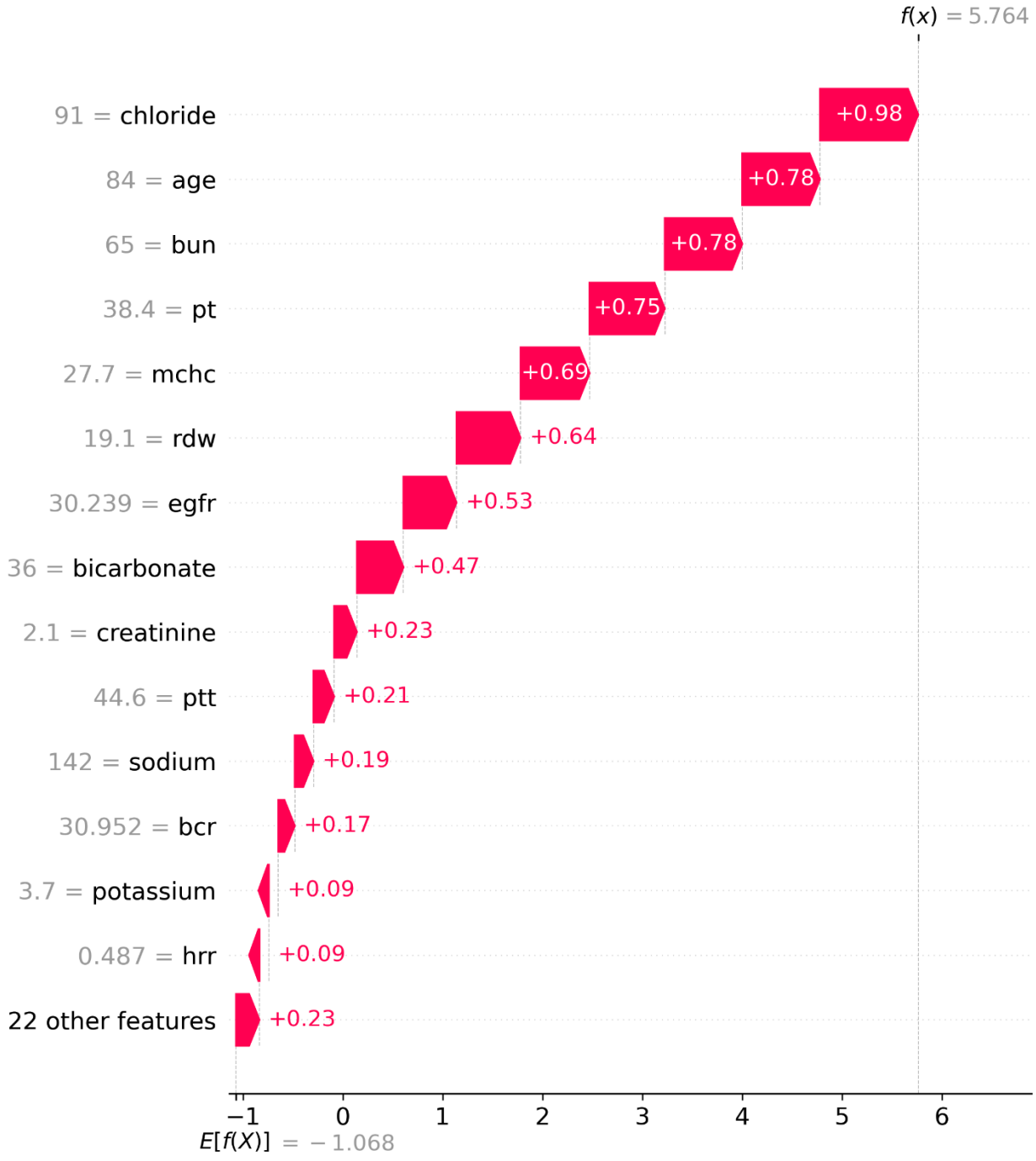
Κώδικας 20 Παραγωγή ραβδογράμματος σημαντικότητας SHAP



Εικόνα 34 :Παράδειγμα απεικόνισης ραβδογράμματος σπουδαιότητας SHAP με γνώμονα τους βιοδείκτες συνόλου βαθμονόμησης που εκπαιδεύτηκε το μοντέλο

Το τελευταίο παραγόμενο διάγραμμα της _make_shap_plots ξεφεύγει από το καθολικό χαρακτήρα των διαγραμμάτων και μετατοπίζεται στο τοπικό, συνθέτοντας την απεικόνιση μίας μεμονωμένης πρόβλεψης. Ο ασθενής που επιλέγεται για την παρουσίαση είναι η μεγαλύτερη προβλεπόμενη πιθανότητα για εμφάνιση καρδιακής ανεπάρκειας. Ειδικότερα, η τεχνική λεπτομέρεια που αξίζει να αναφερθεί είναι η κλήση του πρώτου δείγματος shap_local[0], καθώς αποτελεί

ένα «πακέτο» πληροφοριών που περιέχει τις τιμές SHAP (φ_i) για κάθε ένα από τα χαρακτηριστικά του, την αναμενόμενη τιμή $E[f(X)]$ ως σημείο εκείνης του διαγράμματος Waterfall και τις ονομαστικές τιμές που είχε ο ασθενής κατά την εισαγωγή του. Μαζί με αυτό, σημαντική τεχνική λεπτομέρεια είναι και ο περιορισμός των χαρακτηρηστικών στις 15 από τις 20 γραμμές διότι οι 15 γραμμές χαρακτηριστικών δίνουν επαρκή οπτικό χώρο για την απεικόνιση και των ενδείξεων συμβολής. Η κλινική αξία του διαγράμματος Καταρράκτη (Waterfall) παραπέμπει στην απεικόνιση της πορείας του ασθενή υψηλού κινδύνου, από τη μέση τιμή πρόβλεψης του μοντέλου στη τελική προβλεπόμενη πιθανότητα, αναγνωρίζοντας την συνεισφορά στην μείωση ή ενίσχυση εκτίμησης κινδύνου του μοντέλου.



Εικόνα 35 : Παράδειγμα απεικόνισης διαγράμματος καταρράκτη SHAP με γνώμονα τους βιοδείκτες συνόλου βαθμονόμησης που εκπαιδεύτηκε το μοντέλο

```
 1. try:
 2.             if len(X_test) > 0:
 3.                 hi = int(np.argmax(proba_test))
 4.                 X_hi = X_test_for.iloc[[hi]]
 5.                 shap_hi = explainer(X_hi)
 6.                 shap_hi = _shap_pick_positive_class(shap_hi)
 7.
 8.                 # shap_hi[0] is the Explanation for the
single sample
 9.                 plt.figure(figsize=(10, 6))
10.                 shap.plots.waterfall(shap_hi[0], show=False,
max_display=15)
11.                 plt.tight_layout()
12.                 plt.savefig(outdir /
"shap_waterfall_high_risk_patient.png", dpi=300,
bbox_inches="tight")
13.                 plt.close()
14.         except Exception as e:
15.             print(f"WARNING: SHAP waterfall failed: {e}")
16.
17.     except Exception as e:
18.         print(f"WARNING: SHAP failed: {e}")
```

Κώδικας 21 Παραγωγή γραφήματος καταρράκτη SHAP υψηλού κινδύνου

Πίνακας 5.3: Χαρακτηριστικά εισόδου γραφημάτων SHAP του προγνωστικού μοντέλου

Όνομα	Επιστημονική Ονομασία	Μονάδα Μέτρησης	Κατηγορία
Age	Ηλικία Ασθενούς	Έτη	Δημογραφικό
Mch	Μέση Πυκνότητα Αιμοσφαιρίνης Ερυθρών	Pg/cell	Αιματολογικός Δείκτης
Mcv	Μέσος Όγκος Ερυθρών	fL	Αιματολογικός Δείκτης
Hemoglobin	Αιμοσφαιρίνη	g/dL	Αιματολογικός Δείκτης
Hrr	Λόγος Αιμοσφαιρίνης προς RDW	Αδιάσταστο	Σύνθετος Δείκτης

Chloride	Χλώριο ορού	mEq/L	Ηλεκτρολύτης
Anion_gap	Χάσμα Ανιόντων	mEq/L	Οξεοβασική Ισορροπία
Egfr	Εκτιμώμενος Ρυθμός Σπειραματικής Διήθησης	mL/min/1.73m^2	Σύνθετος δείκτης νεφρικής λειτουργίας
Acag	Διορθωμένο Χάσμα Ανιόντων για Αλβουμίνη	mEq/L	Σύνθετος Δείκτης
Sodium	Νάτριο ορού	mEq/L	Ηλεκτρολύτης
Sex	Φύλο	Δυαδικό	Δημογραφικό
Bun	Άζωτο Ουρίας Αίματος	mg/dL	Νεφρικός Δείκτης
Mchc	Μέση Πυκνότητα Αιμοσφαιρίνης Ερυθρών κατά Όγκο	g/dL	Αιματολογικός Δείκτης
Hematocrit	Αιματοκρίτης	%	Αιματολογικός Δείκτης
Glucose	Γλυκόζη Ορού	mg/dL	Μεταβολικός Δείκτης
Bicarbonate	Διαττανθρακικά	mEq/L	Οξεοβασική Ισορροπία
platelet	Αιμοπετάλια	$\times 10^9$/L	Αιματολογικός Δείκτης

Συνεπώς, το κάθε διάγραμμα έχει τον σκοπό του. Το ραβδόγραμμα του διαγράμματος σπουδαιότητας εξυπηρετεί στην εποπτική κατανόηση λειτουργίας του μοντέλου, το Beeswarm συμβάλλει στη κατεύθυνση επιρροής των χαρακτηριστικών του και το διάγραμμα καταρράκτη απομονώνει τη διάγνωση του κάθε ασθενή ξεχωριστά.

5.4 Backend API

Το FastAPI Backend που χρησιμοποίει η παρούσα κλινική πλατφόρμα υλοποιείται στο αρχείο api.py και αποτελεί τον δίαυλο επικοινωνίας μεταξύ του frontend της εφαρμογής και του εκπαιδευμένου μοντέλου. Το api.py μέσω των RESTful endpoints, το αρχείο επιτρέπει την ασύγχρονη επικοινωνία μέσω HTTP, διευκολύνοντας την άντληση των προβλέψεων κινδύνου και των ερμηνειών SHAP σε πραγματικό χρόνο.

5.4.1 Αρχιτεκτονική του API

Ο εξυπηρετητής έχει την φιλοσοφία της μονής φόρτωσης του αρείο και της εξυπηρέτησης πολλαπλών φορών. Το αρχείο model_bundle.pkl που παράγεται από τη συνάρτηση train_and_select, φορτώνεται μία φορά κατά την εκκίνηση του εξυπηρετητή μέσω συνάρτησης _bootstrap και αποσειριοποιείται με την αποθήκευση των μεταβλητών και να γίνεται καθολική για να επαναχρησιμοποιούνται σε κάθε επερχόμενο αίτημα. Αν το model_bundle.pkl λείπει ή

διαφθαρεί ο εξυπηρετητής εκκινεί κανονικά αλλά σε υποβαθμισμένη (degraded) κατάσταση επιστρέφοντας ένα μήνυμα σφάλματος αντί για ολική κατάρρευση. Επίσης η ευρεία πολιτική του επιτρέπει την επικοινωνία με οποιαδήποτε οικοδεσπότη μέσω του μηχανισμού CORSMiddleware.

5.4.2 Τερματικά σημεία διαχείρισης

Στο κώδικα υπάρχουν δύο GET σημεία που παρέχουν πληροφορίες κατάστασης χωρίς να εκτελούν συμπερασμό. Το /api/health επιστρέφει ένα πεδίο τριών αντικειμένων (status, model_loaded, error) και αξιοποιείται από το frontend για να ελέγχεται η διαθεσιμότητα του εξυπηρετητή. Από την άλλη, το /api/features που με την κλήση του επιστρέφει τα ονόματα χαρακτηριστικών και την ονομασία του προγνωστικού μοντέλου, δίνοντας τη δυνατότητα στο frontend να κατασκευάζει δυναμικά τις φόρμες εισαγωγής των δεδομένων. Ο μηχανισμός αυτός υπηρετεί στη διατήρηση της δομής του frontend σε περίπτωση που το μοντέλο εκπαιδευτεί για παραπάνω κλινικές εφαρμογές.

5.4.3 Τερματικά Σημεία Πρόβλεψης και Καταγραφής της

Το /api/score αποτελεί το τερματικό σημείο που αποδέχεται ένα μεταφορτωμένο αρχείο CSV με τους βιοδείκες του ασθενή και επιστρέφει ένα επεξεργασμένο CSV εμπλουτισμένο σε πίνακα δύο στηλών, το risk_score που περιλαμβάνει τη πιθανότητα στο εύρος τιμών [0,1] και το risk_percent με την ίδια τιμή πιθανότητας στρογγυλοποιημένη προς δύο δεκαδικά. Μόλις βγει ο πίνακας με τα σκορ κινδύνου φορτώνεται στη ροή επεξεργασίας του. Πρώτα από όλα ξεκινάει με την ανάγνωση του αρχείο CSV και μέσω της _make_design_matrix() μετατρέπονται τα ακατέργαστα δεδομένα στην αυστηρή μορφή που απαιτεί το μοντέλο δημιουργώντας έναν νέο πίνακα. Έπειτα ακολουθεί η συμπλήρωση των ελλειπών τιμών με τη κλάση Imputer και παράγονται οι βαθμονομημένες πιθανότητες μέσω του calibrator. Η μορφή της απόκρισης επιστρέφεται ως Streaming Response

5.5 Έξυπνο Συμβόλαιο

Το έξυπνο συμβόλαιο MedicalOracle.sol αποτελεί τον βασικό μηχανισμό αμετάβλητης καταγραφής, ασφάλειας και ελέγχου πρόσβασης στα δεδομένα της προτεινόμενης κλινικής πλατφόρμας υποστήριξης ιατρικών αποφάσεων. Το συμβόλαιο έχει αναπτυχθεί στη γλώσσα προγραμματισμού Solidity, για εκδόσεις ^0.8.0 και χρησιμοποιεί το έξυπνο συμβόλαιο AccessControl.sol για τις βασικές λειτουργίες προσβασιμότητας. Στο παρόν κεφάλαιο, αναλύεται ο σχεδιασμός και η υλοποίηση του, με πρώτο σημείο αναφοράς η γέφυρα που δημιουργεί μεταξύ του pipeline μηχανικής μάθησης και του έξυπνου συμβολαίου και καταλήγει στις υπόλοιπες λειτουργίες του.

5.5.1 Διεπαφή Backend με το Έξυπνο Συμβόλαιο

Η μεταφορά των αποτελεσμάτων από το pipeline της μηχανικής μάθησης στο Blockchain πραγματοποιείται σε δύο διαδοχικά στάδια. Στο πρώτο στάδιο, η συνάρτηση generate_blockchain_payload() λαμβάνει τη πιθανότητα επικινδυνότητας (riskScore) που προέβλεψε το μοντέλο, τη ταυτότητα του ασθενή (patientId), τις τιμές των βιοδεικτών που καταγράφηκαν στην φόρμα εισαγωγής και την έκδοση του αλγορίθμου μηχανικής μάθησης. Μετά από τις απαραίτητες διεργασίες η συνάρτηση κατασκευάζει ένα payload που περιέχει 8 πεδία χαρακτηριστικών, το blockchain_payload.json.

```
payloads.append(
            {
                "patient_hash": patient_hash,
                "admission_hash": admission_hash,
                "risk_score": round(p, 6),
                "risk_score_int": int(round(p * SCALE)),
                "data_integrity_hash": data_hash,
                "model_version": cfg.model_version,
                "feature_schema_hash":
hashlib.sha256("|".join(feature_cols).encode("utf-
8")).hexdigest(),
                "inference_time_utc": now_iso,
            }
        )
```

Κώδικας 22 Κατασκευή payload μεταφοράς πρόβλεψης στο Blockchain

Στο παραπάνω στιγμιότυπο κώδικα, η ταυτότητα του ασθενή έχει κρυπτογραφηθεί HMAC-SHA256 με κρυπτογραφικό «pepper» το οποίο φορτώνεται από την καθολική μεταβλητή HASH_PEPPER στο main.py. Το admission_hash αποτελεί μοναδικό κρυπτογραφικό αποτύπωμα για συγκεκριμένη νοσηλεία, διαχωρίζοντας τις εισαγωγές που μπορεί να έχει πραγματοποιήσει ένας ασθενής. Το risk_score αντιπροσωπεύει την δεκαδική πιθανότητα που έβγαλε το μοντέλο και επειδή η Solidity δεν έχει την ικανότητα να την διαβάσει το πεδίο risk_score_int αναλαμβάνει τον πολλαπλασιασμό του σκορ επί 10000 ώστε να μετατραπεί σε ακέραιο αριθμό και να αποθηκευτεί στην αλυσίδα. Ιδιαίτερο ρόλο εμφανίζει το πεδίο data_integrity_hash καθώς αποτελεί την εγγύηση ακεραιότητας αν κάποιος αλλάξει έστω και μια τιμή βιοδείκτη στο μέλλον, το hash δεν επαληθεύεται, διασφαλίζοντας πως τα αποτελέσματα δεν έχουν διαστρεβλωθεί κατά την εξαγωγή τους και τα δεδομένα που βασίστηκε η πρόβλεψη είναι τεκμηριωμένα. Το model_version είναι το πεδίο που αποκαλύπτει την έκδοση του προβλεπτικού μοντέλου ενώ το feature_schema_hash αποτελεί μια έξυπνη προσθήκη διότι είναι το κρυπτογραφικό αποτύπωμα της σειράς εκπαίδευσης των βιοδεικτών στο μοντέλο. Έτσι σε περίπτωση προσθήκης ενός επιπλέον βιοδείκτη στο μοντέλο, το κρυπτογραφικό αποτύπωμα θα αλλάξει, προειδοποιώντας πως το μοτίβο των βιοδεικτών δεν είναι το ίδιο. Ως τελευταίο πεδίο παρουσιάζεται το inference_time_utc που είναι

η ακριβής χρονική στιγμή της πρόβλεψης σε μορφή ISO. Αυτό μετά από μετατροπή που θα υποστεί σε unix timestamp θα αποθηκευτεί στο έξυπνο συμβόλαιο.

Η μεταφορά του payload στο έξυπνο συμβόλαιο γίνεται με το script sub_storage.py. Το script αυτό διαβάζει το blockchain_payload.json και εκτελεί μια συναλλαγή καταγραφής πάνω στην αλυσίδα (on-chain) μέσω της βιβλιοθήκης web3.py, μέσω κλήσης της μεθόδου publishRiskScore. Η μέθοδος publishRiskScore επιστρέφει ένα λεξικό τεσσάρων ορισμάτων {patient_hash, risk_score_int, data_integrity_hash, model_version}. Για αποδοτικότερη ανάρτηση μεγάλων όγκων προβλέψεων εφαρμόζεται η τμηματοποίηση των δεδομένων υπό τη μορφή δομής Merkle Tree, όπως αναφέρεται στην ενότητα 5.5.5. Η δομή Merkle Tree επιτρέπει την αποτελεσματικότερη και ασφαλή επαλήθευση ακεραιότητας των δεδομένων μεγάλη κλίμακας κατανεμημένα συστήματα.

5.5.2 Επισκόπηση του Έξυπνου Συμβολαίου MedicalOracle

Το MedicalOracle είναι αρχείο τύπου .sol και αποτελεί την «καρδιά» του Blockchain. Αρχικά χαρακτηρίζεται από την αμετάβλητη καταγραφή των προβλέψεων κινδύνου, τον έλεγχο προσβασιμότητας βάση ρόλων, τη συγκατάθεση ασθενούς συμβατή με τον Γενικό Κανονισμό για την Προστασία Δεδομένων (GDPR) και την καταγραφή ενεργειών έκτακτης ανάγκης. Η αρχιτεκτονική του συμβολαίου οργανώνεται σε τέσσερις θεμελιώδεις δομές δεδομένων.

```
1. //---STATE VARIABLE---
 2.     bytes32 public constant DOCTOR = keccak256("DOCTOR");
 3.     bytes32 public constant ADMIN = keccak256("ADMIN");
 4.     bytes32 public constant ORACLE_ROLE =
keccak256("ORACLE_ROLE");
 5.     bytes32 public constant RESEARCHER =
keccak256("RESEARCHER");
 6.
 7.
 8.     struct Prediction{
 9.         uint256 timestamp;
10.         uint256 riskScore;
11.         bytes32 dataHash;
12.         bytes32 modelVersion;
13.         bool isVerified;
14.         address verifiedBy;
15.         bool doctorAgrees;
16.     }
17.
18.     struct PatientConsent{
19.         bool dataStorageAllowed;
20.         bool researchUseAllowed;
```

```
21.         uint256 consentTimestamp;
22.         uint256 retentionDays; //0=indefinite
23.         bool isActive;
24.     }
25.
26.     struct EmergencyAccessLog {
27.         address doctor;
28.         uint256 timestamp;
29.         string reason;
30.     }
31.
32.     struct Batch {
33.         uint256 batchId;
34.         uint256 timestamp;
35.         bytes32 merkleRoot;
36.         string ipfsCid;
37.         uint256 recordCount;
38.         string modelVersion;
39.     }
```

Κώδικας 23 Κατασκευή payload μεταφοράς πρόβλεψης στο Blockchain

Η δομή της αποθήκευσης οργανώνεται σε τρία mappings. Το πρώτο ήταν το patientRecords, το οποίο για κάθε κρυπτογραφημένο ασθενή (patient_hash) διατηρεί έναν πίνακα από αντικείμενα Prediction. Η χρήση πίνακα και όχι μια απλή καταγραφή τιμής είναι πως ένας ασθενείς μπορεί να εμφανίζεται σε πολλαπλές εισαγωγές στη διάρκεια του χρόνου, όπως αναφέρθηκε 43.3% των ασθενών >1 νοσηλειών. Με την δομή του πίνακα διατηρεί ολόκληρο το ιατρικό ιστορικό των προβλέψεων, κάτι που χρησιμεύει στη παρακολούθηση τάσης κινδύνου, στη σύγκριση μεταξύ μοντέλων διαφορετικών εκδόσεων και νομική ιχνηλασιμότητα. Ιδιαίτερη παρατήρηση αξίζει να δοθεί στο τύπο του mapping που είναι ιδιωτικό (private).

```
1.mapping(bytes32 => Prediction[]) private patientRecords;

2.mapping(bytes32 => PatientConsent) public patientConsents;

3.mapping(bytes32 => EmergencyAccessLog[]) public emergencyLogs;
```

Κώδικας 24 Ορισμοί mappings αποθήκευσης (patientRecords, patientConsents, emergencyLogs)

Η δεύτερη δομή αποθήκευσης είναι το mapping patientConsents, που είναι ο πίνακας των καταγεγραμμένων συγκαταθέσεων ανά ασθενή. Ο συγκεκριμένος πίνακας αποθηκεύει για κάθε ασθενή μία μοναδική εγγραφή της τρέχουσας κατάστασής του. Σε σχέση με το patientsRecords, η δομή αποθήκευσης είναι δημόσια (public), καθώς η συγκατάθεση πρέπει να είναι διαφανής και

επαληθεύσιμη. Η κάθε εγγραφή περιέχει διάφορα υποπεδία, με διαφορετικές λειτουργίες το καθένα, σύμφωνα με τον παρακάτω πίνακα

Πίνακας 5.4: Πίνακας πεδίων λειτουργίας της patientsRecords

Πεδίο	Λειτουργία
dataStorageAllowed	Επιτρέπει την αποθήκευση των προβλέψεων on-chain
researchUseAllowed	Επιτρέπει την δημοσίευση των δεδομένων για ερευνητικούς σκοπούς
consentTimestamp	Χρονική στιγμή της συγκατάθεσης
retentionDays	Διάρκεια διατήρησης νοσηλείας
isActive	Ενεργή συγκατάθεση

Η τρίτη δομή αποθήκευσης είναι το mapping EmergencyAccessLog. Αφορά τον πίνακα καταγραφής πρόσβασης έκτακτης ανάγκης. Κάθε φορά που ένας γιατρός, μόνο ο γιατρός έχει το δικαίωμα κλήσης έκτακτης ανάγκης όπως θα αναλυθεί στην ενότητα 5.5.3, οι κανόνες συγκατάθεσης παρακάμπτονται και μια νέα εγγραφή προσαρτάται. Το είδος του πίνακα είναι public για λόγους διαφάνειας

Σημαντική σχεδιαστική αρχή είναι ότι οι ασθενείς δεν είναι χρήστες του Blockchain, με την επιλογή του κλειδιού ως τύπου bytes32 αντί για κανονικής διεύθυνσης (address). Ακόμη το συμβόλαιο διαθέτει έξι τύπους events που επιτρέπουν την αλληλεπίδραση των παρατηρητών με το frontend να παρακολουθούν ασύγχρονα την κατάσταση της αλυσίδας χωρίς να κάνουν κάποιου είδους παρέμβαση. Τα events και οι λειτουργίες τους περιγράφονται στον πίνακα παρακάτω.

Πίνακας 5.5: Πίνακας λειτουργιών των events του MedicalOracle

Είδος Event	Λειτουργία
RiskPublished	Ενεργοποιείται κάθε φορά που το backend αναρτά μια νέα πρόβλεψη κινδύνου για έναν ασθενή.
PredictionVerified	Ενεργοποιείται όταν ένας γιατρός επικυρώνει ή απορρίπτει μια υπάρχουσα πρόβλεψη.
ConsentUpdated	Ενεργοποιείται κάθε φορά που η συγκατάθεση ενός ασθενούς τροποποιείται ή ανακαλείται
EmergencyAccessGranted	Ενεργοποιείται κάθε φορά που ένας γιατρός παρακάμπτει τους κανόνες συγκατάθεσης.
ResearchAccessGranted	Ενεργοποιείται όταν ένας ερευνητής ανακτά τα δεδομένα ενός ασθενούς για έρευνα
BatchPublished	Ενεργοποιείται για την ενημέρωση εξωτερικών χρηστών για την ύπαρξη νέο batch merkle root

5.5.3 Έλεγχος Πρόσβασης βάσει Ρόλων

Το έξυπνο συμβόλαιο κληρονομεί την κλάση AccessControl της OpenZeppelin και ορίζει τέσσερις διακριτούς ρόλους

```
1. bytes32 public constant DOCTOR = keccak256("DOCTOR");
2.     bytes32 public constant ADMIN = keccak256("ADMIN");
3.     bytes32 public constant ORACLE_ROLE =
keccak256("ORACLE_ROLE");
4.     bytes32 public constant RESEARCHER =
keccak256("RESEARCHER");
```

Κώδικας 25 Ορισμός ρόλων ελέγχου πρόσβασης (RBAC)

Ο κάθε ρόλος εξουσιοδοτείται με διαφορετικά δικαιώματα πάνω στη κλινική πλατφόρμα με γνώμονα τη λειτουργική ευθύνη του. Αρχικά, ο ORACLE_ROLE ανατίθεται σε λειτουργία του backend που διαχειρίζεται τις προβλέψεις των μοντέλων μηχανικής μάθησης και δικαιούται να καλέσει το ποσοστό κινδύνου καρδιακής ανεπάρκειας και της δημοσίευσης του batch. Ο Γιατρός (DOCTOR) δικαιούται να πράξει τις περισσότερες ενέργειες, που είναι κυρίως κλινικές. Ο Διαχειριστής (ADMIN) έχει τον ρόλο του επόπτη της πλατφόρμας, με τη διαχείριση των συγκαταθέσεων των ασθενών και της ανάκλησης τους, ενώ ο Ερευνητής (RESEARCHER) δικαιούται πρόσβαση σε ιατρικά δεδομένα, εφόσον έχει υπάρξει η συγκατάθεση του ασθενή.

5.5.4 Λογική Συγκατάθεση κατά GDPR

Η συμμόρφωση με το Γενικό Κανονισμό Προστασίας Προσωπικών Δεδομένων (GDPR) αντιμετωπίζεται μέσω της συνάρτηση τροποποίησης hasStorageConsent, η οποία εφαρμόζεται κάθε αποθήκευση δεδομένων, ασχέτως του ρόλου εξουσιοδότησης και επιβάλλει τρείς ελέγχους τήρησης του κανονισμού. Στον πρώτο έλεγχο εξασφαλίζεται το δικαίωμα ανάκλησης της συγκατάθεσης, του άρθρου 7 του GDPR, ανά πάσα στιγμή από τον ρόλο του Διαχειριστή (ADMIN) μπλοκάροντας παράλληλα και όλες τις μεταγενέστερες προσπάθειες ανάρτησης. Ο δεύτερος μηχανισμός εξυπηρετεί στη διάκριση μεταξύ διαφορετικών τύπων συγκατάθεσης, οι οποίες ορίζονται μέσω των σημαιών Boolean dataStorageAllowed και researchUseAllowed στη δομή PatientConsent, που αναφέρθηκε προηγουμένως στην ενότητα 5.5.2. Μάλιστα η χρήση δύο ανεξάρτητων Boolean εξυπηρετούν στον διαχωρισμό των λειτουργιών αποθήκευσης των προβλέψεων και τη χρήση τους για ερευνητικούς σκοπούς. Η εφαρμογή τους ξεκινάει στη συνάρτηση setPatientConsent, η οποία έχει ως κύρια λειτουργία τον ορισμό συγκατάθεσης του ασθενή και δίνει την δυνατότητα σε αυτόν τον ασθενή να επιλέξει τις παρακάτω διακριτές επιλογές χρήσης των δεδομένων του. Έτσι τηρείται η αρχή περιορισμού του σκοπού του άρθρου 5 του GDPR. Τρίτον στη συνάρτηση τροποποίησης hasStorageConsent επιβάλλει την αρχή περιορισμού της χρονικής περίοδο διατήρησης, του άρθρου 5 του GDPR. Σύμφωνα με αυτήν, ο

τύπος συγκατάθεσης του ασθενούς απορρίπτεται αυτόματα μετά την περίοδο διατήρησης που όρισε ο ίδιος, χωρίς δυνατότητα παράκαμψης. Η μοναδική δυνατότητα μόνιμης διατήρησης αορίστου χρόνου μιας συγκατάθεσης γίνεται με τον ορισμό της μεταβλητής retentionDays=0, δηλαδή την ρητή αποδοχή αορίστου χρόνου από τον ασθενή. Οι τρείς έλεγχοι μετατρέπουν τον Γενικό Κανονισμό Προστασίας Προσωπικών Δεδομένων από νόμο σε εκτελέσιμη συνθήκη διασφαλίζοντας συλλογικά την προστασία των προσωπικών δεδομένων που καταγράφονται πάνω ή εκτός αλυσίδας.

```
 1. function setPatientConsent(
 2.     bytes32 _patientHash,
 3.     bool _storageAllowed,
 4.     bool _researchAllowed,
 5.     uint256 _retentionDays
 6. ) public {
 7.     require(hasRole(ADMIN, msg.sender) || hasRole(DOCTOR,
msg.sender),
 8.             "Consent isn't manageable");
 9.     patientConsents[_patientHash] = PatientConsent({
10.         dataStorageAllowed: _storageAllowed,
11.         researchUseAllowed: _researchAllowed,
12.         consentTimestamp: block.timestamp,
13.         retentionDays: _retentionDays,
14.         isActive: true
15.     });
16.     emit ConsentUpdated(_patientHash, _storageAllowed,
_researchAllowed, _retentionDays);
17. }
```

Κώδικας 26 Συνάρτηση καταγραφής συγκατάθεσης ασθενούς setPatientConsent

Πίνακας 5.6: Πίνακας λειτουργιών του μηχανισμού συγκατάθεσης κατά GDPR. Παρουσιάζονται οι τέσσερις πιθανοί συνδυασμοί των σημαιών dataStorageAllowed και researchUseAllowed

Flag: dataStorageAllowed	Flag: researchUseAllowed	Λειτουργία
True	True	Ταυτόχρονη αποθήκευση προβλέψεων και έρευνα
True	False	Δυνατότητα αποθήκευσης χωρίς την ερευνητική χρήση
False	True	Ανώνυμη ερευνητική χρήση χωρίς αποθήκευση
False	False	Πλήρης άρνηση καταγραφής

```
 1. modifier hasStorageConsent(bytes32 _patientHash) {
 2.     PatientConsent memory c = patientConsents[_patientHash];
```

```
 3.      require(c.isActive, "Consent is NOT active");
 4.      require(c.dataStorageAllowed, "Storage NOT allowed by
patient");
 5.      if (c.retentionDays > 0) {
 6.          require(block.timestamp <= c.consentTimestamp +
(c.retentionDays * 1 days),
 7.                  "Consent has EXPIRED");
 8.      }
 9.      _;
10. }
```

Κώδικας 27 Συνάρτηση τροποποίησης ελέγχου συγκατάθεσης setPatientConsent

5.5.5 Τμηματοποίηση μέσω Merkle Trees

Η ανάρτηση μίας συναλλαγής στο Blockchain ανά πρόβλεψη διακρίνεται απαγορευτικά ακριβή στη περίπτωση μιας νοσοκομειακής μονάδας που υποδέχεται χιλιάδες ασθενείς. Για την εξοικονόμηση κόστους ανάρτησης (gas fees) και την ευκολότερη εφαρμογή του έξυπνου συμβολαίου σε πραγματικές κλινικές πλατφόρμες, υιοθετείται η λογική της ομαδοποιημένης δημοσίευσης μέσω Merkle δέντρων. Δηλαδή, το έξυπνο συμβόλαιο MedicalOracle διαθέτει τη συνάρτηση publishBatch η οποία υπολογίζει την merkle ρίζα πολλαπλών εγγραφών και τη δημοσιεύει μόνο αυτή μέσω της publishBatch. Η εξατομικευμένη επαλήθευση μιας συγκεκριμένης εγγραφής γίνεται μέσω της συνάρτησης verifyPatientInBatch, στην οποία ένας αιτούντας αποστέλλει ένα merkle αποδεικτικό και επιβεβαιώνει αν το ζεύγος χαρακτηριστικών (patient_hash, risk_score) ανήκει στο δηλωθέν batch.

```
 1. function emergencyAccess(bytes32 _patientHash, string
calldata _reason) public onlyRole(DOCTOR) returns(Prediction[]
memory){
 2.         require(bytes(_reason).length > 5,"Reason
Required");
 3.

4.         emergencyLogs[_patientHash].push(EmergencyAccessLog({
 5.             doctor: msg.sender,
 6.             timestamp: block.timestamp,
 7.             reason: _reason
 8.         }));
 9.         emit EmergencyAccessGranted(_patientHash,
msg.sender, _reason);
```

```
10.         return patientRecords[_patientHash];
11.     }
```

Κώδικας 28 Συνάρτηση πρόσβασης έκτακτης ανάγκης emergencyAccess

```
 1. function verifyPatientInBatch(
 2.         uint256 _batchId,
 3.         bytes32[] calldata _proof,
 4.         bytes32 _patientHash,
 5.         uint256 _riskScore
 6.     ) public view returns (bool){
 7.         require(_batchId < batches.length, "Batch doesn't
exist");
 8.         bytes32 root = batches[_batchId].merkleRoot;
 9.         bytes32 leaf =
keccak256(abi.encodePacked(_patientHash, _riskScore));
10.         return verifyCalldata(_proof, root, leaf);
11.     }
```

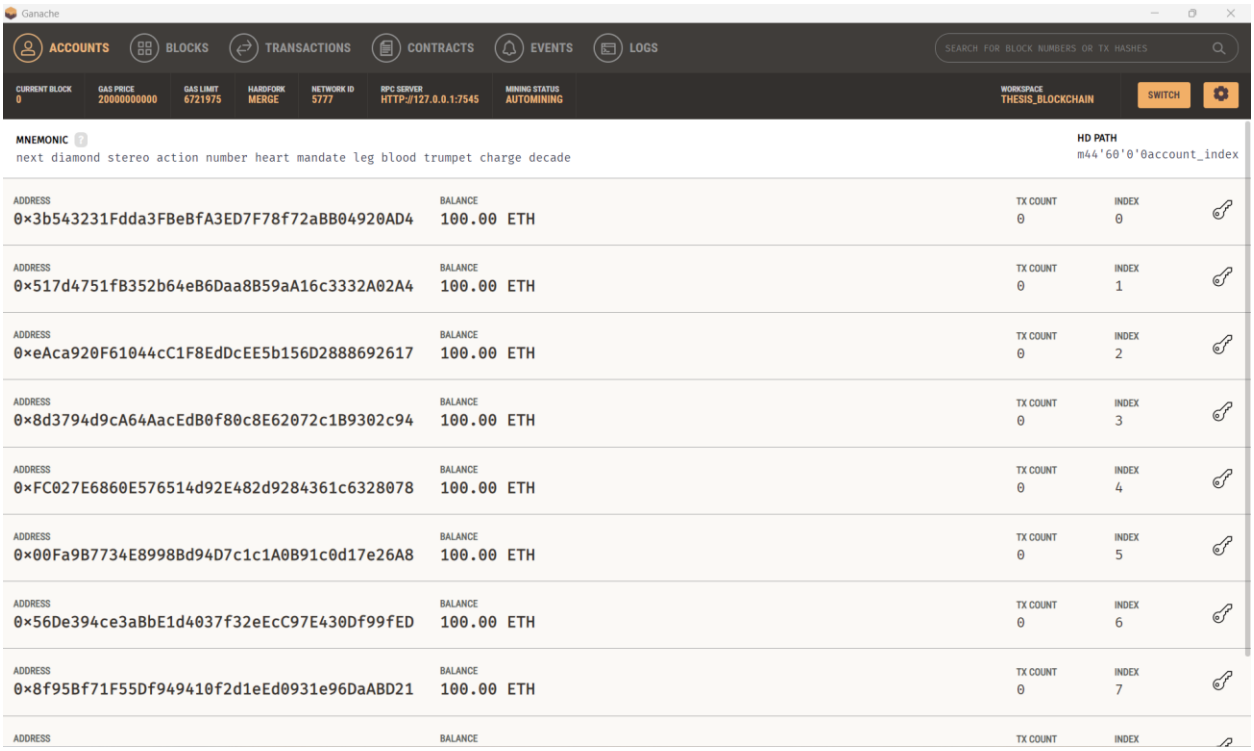
Κώδικας 29 Συνάρτηση επικύρωσης του ασθενή verifyPatientInBatch

5.5.6 Έκτακτη Πρόσβαση και Καταγραφή Ενεργειών

Σε μονάδες εντατικής θεραπείες η άμεση πρόσβαση ή καταγραφή των ιατρικών δεδομένων ενός ασθενή πολλές φορές παρουσιάζεται απαραίτητη. Για την αποφυγή, λοιπόν, χρονικών καθυστερήσεων από τη διαδικασία συγκατάθεσης, η συνάρτηση emergencyAccess() επιτρέπει την ανάκτηση του ιατρικού ιστορικού χωρίς την ενεργή συγκατάθεση του ασθενή. Επίσης καταγράφει την κάθε απόπειρα τέτοια ενέργειας καταγράφεται στο mapping emergencyLogs και εκπέμπει το event EmergencyAccessGranted. Με αυτόν τον τρόπο δίνεται η δυνατότητα στον γιατρό να ασκήσει την κρίση του σε επείγουσα κατάσταση και ταυτόχρονα να γνωρίζει πως η ενέργεια θα καταγραφεί μόνιμα στο σύστημα, παραμένοντας ορατή στους αρμόδιους.

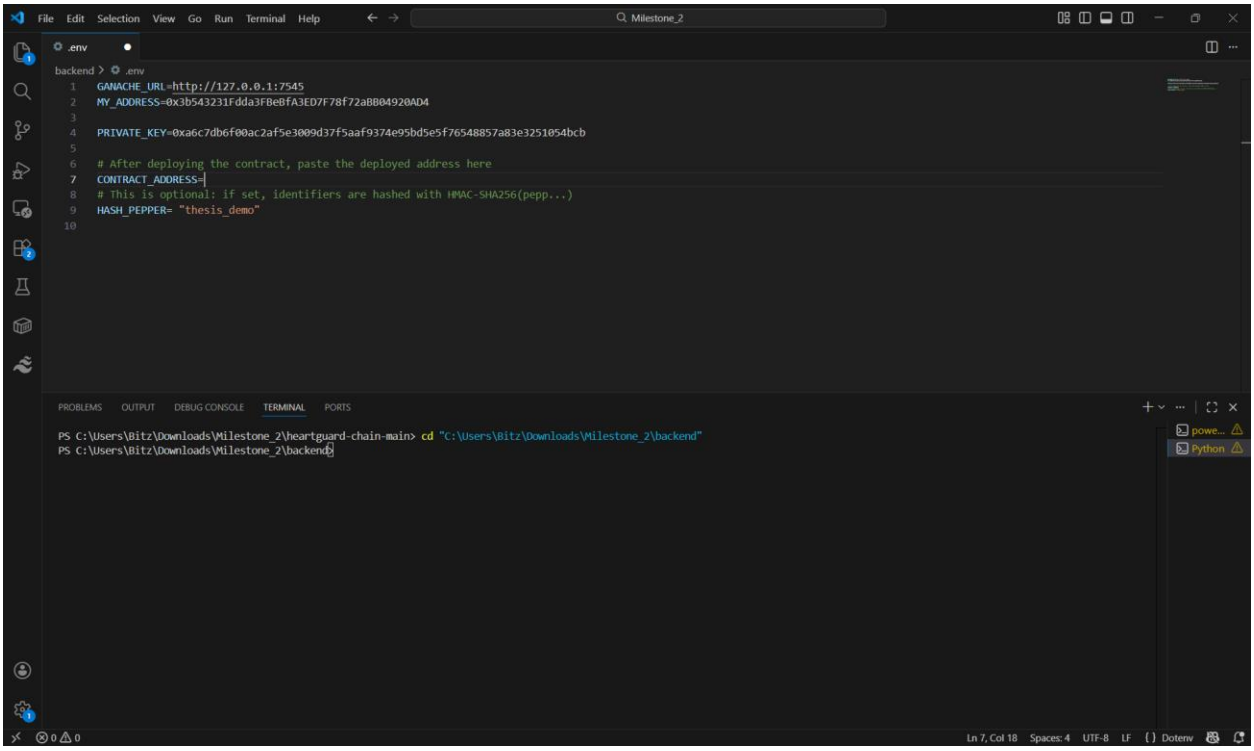
5.5.7 Ανάπτυξη Έξυπνου Συμβολαίου

Το έξυπνο συμβόλαιο MedicalOracle αναπτύσσεται σε ένα τοπικό περιβάλλον προσομοίωσης του δικτύου Ethereum, το Ganache. Το Ganache είναι εργαλείο ανάπτυξης Blockchain για πειραματικούς σκοπούς, με τις συναλλαγές να γίνονται δωρεάν σε αυτό και άμεσα χωρίς τις καθυστερήσεις των πραγματικών ανθρακωρύχων (miners). Η επιλογή του ήταν ιδανική για την επείδη της end-to-end λειτουργίας της προτεινόμενης κλινικής πλατφόρμας. Ως πρώτο βήμα δημιουργήθηκε ένας τοπικός χώρος εργασίας (workspace) με το όνομα thesis_blockchain, με διεύθυνση RPC την HTTP://127.0.0.1:7545 και chain Id το 5777. Το τοπικό δίκτυο παρέχει 10 λογαριασμούς με υπόλοιπο 100 ETH ώστε να υπάρχει η δυνατότητα πραγματοποίησης πληθώρας συναλλαγών.



Εικόνα 36 : Στιγμιότυπο οθόνης του τοπικού περιβάλλοντος Ganache πριν την ανάπτυξη του συμβολαίου

Πριν αναπτυχθεί το συμβόλαιο γίνεται η αντιγραφή των στοιχείων του πρώτου λογαριασμού του Ganache στο αρχείο .env της εργασίας. Πιο συγκεκριμένα, αποθηκεύονται η διεύθυνση Ganache, η διεύθυνση του λογαριασμού διαχειριστή και το ιδιωτικό κλειδί που θα χρησιμοποιηθεί για την υπογραφή των συναλλαγών ανάπτυξης και αρχικοποίησης. Για την πρόσθετη ασφάλεια της κρυπτογράφησης ορίζεται η μεταβλητή «HASH_PEPPER» που χρησιμοποιεί την ψευδωνυμοποίηση των ταυτοτήτων των ασθενών μέσω της κρυπτογράφησης HMAC-SHA256, όπως αναφέρθηκε στην ενότητα 5.5.1. Το πεδίο της διεύθυνσης του συμβολαίου CONTRACT_ADDRESS παραμένει κενό ώστε να συμπληρωθεί στη συνέχεια με την ανάπτυξη του συμβολαίου.



Εικόνα 37 : Στιγμιότυπο του αρχείου .env στο περιβάλλον ανάπτυξης κώδικα Visual Studio

Η ανάπτυξη του έξυπνου συμβολαίου ξεκινά με την εκτέλεση του αρχείου deploy_contract.py. Το script φορτώνει το ABI από το MO_abi.json και το προ-μεταγλωττισμένο bytecode του συμβολαίου MedicalOracle και χρησιμοποιεί τη βιβλιοθήκη Web3.py για να συνδεθεί στο τοπικό δίκτυο Ganache , ώστε να αποσταλεί η συναλλαγή ανάπτυξης υπογεγραμμένη με το ιδιωτικό κλειδί του διαχειριστή. Μόλις ολοκληρωθεί επιτυχώς η εκτέλεση το συμβόλαιο αποκτά μοναδική διεύθυνση πάνω στην αλυσίδα, on-chain. Συνεχίζοντας το script setup_consent.py διαβάζει το αρχείο blockchain_payload.json και κάνει τις εγγραφές στην αλυσίδα με βάση τις αρχικές δηλώσεις συγκατάθεσης των 5 πρώτων μοναδικών ασθενών, σύμφωνα με τη λογική του GDPR που αναπτύχθηκε στην ενότητα 5.5.4. Η κάθε εγγραφή κοστίζει περίπου 95.700 μονάδες καυσίμου και παράγονται συνολικά 478.503 μονάδες καυσίμου.

```
PS C:\Users\Bitz\Downloads\Milestone_2> &
C:/Users/Bitz/AppData/Local/Programs/Python/Python311/python.exe
c:/Users/Bitz/Downloads/Milestone_2/backend/deploy_contract.py
Deploying contract...
Deployed. Contract address:
0x3f08B8F883eC0b4E0E5e36D78A38321c43ff3D3F
Add this to your .env as CONTRACT_ADDRESS=...
PS C:\Users\Bitz\Downloads\Milestone_2>
```

```
PS C:\Users\Bitz\Downloads\Milestone_2> &
C:/Users/Bitz/AppData/Local/Programs/Python/Python311/python.exe
c:/Users/Bitz/Downloads/Milestone_2/backend/setup_consent.py
Loaded payload records: 462890 from
C:\Users\Bitz\Downloads\Milestone_2\backend\hf_outputs\biomarker
_only\blockchain_payload.json
Unique patients to process: 196527
----------------------------------------
[1/196527] OK | Gas: 95703 | Patient: 3a09ddc8...
[2/196527] OK | Gas: 95703 | Patient: e8f14b92...
[3/196527] OK | Gas: 95691 | Patient: 4022e5aa...
[4/196527] OK | Gas: 95703 | Patient: ee5b6e91...
[5/196527] OK | Gas: 95703 | Patient: 51fd6bb8...
SETUP FINISHED.
Success: 5
Failed: 0
Total Gas Used: 478503
PS C:\Users\Bitz\Downloads\Milestone_2>
```

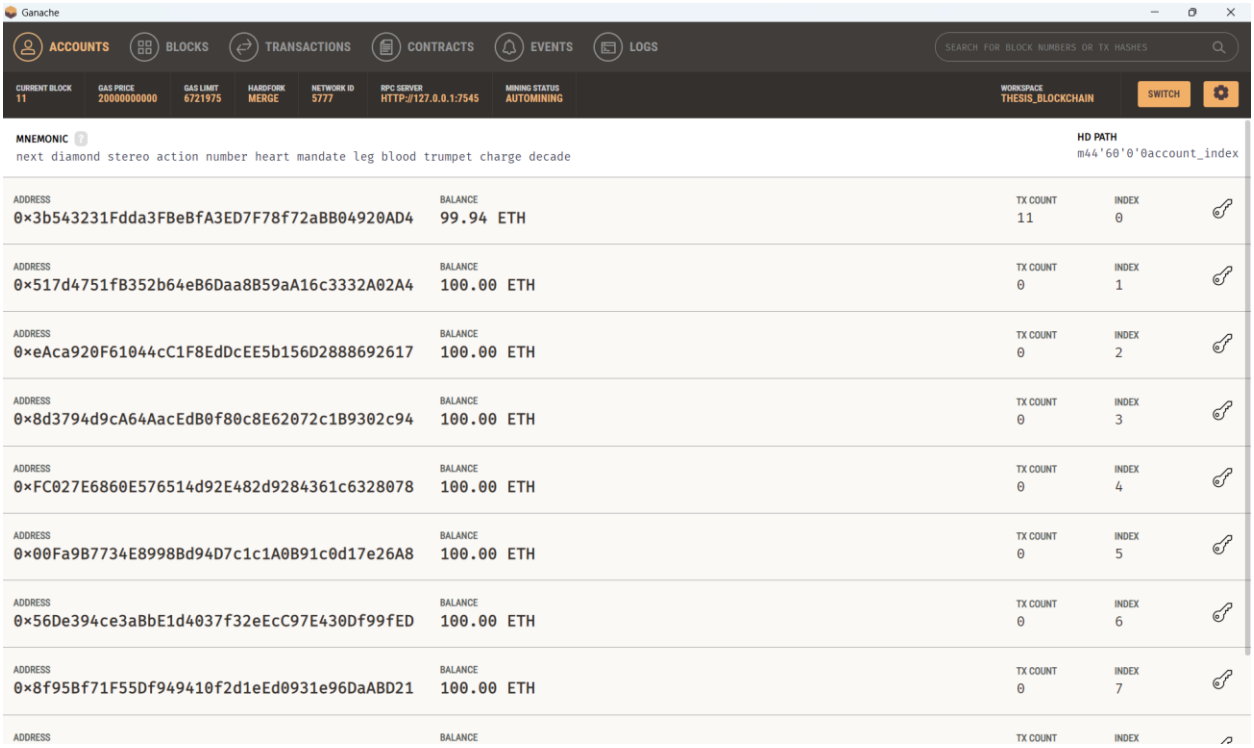
Εικόνα 38 Έξοδος του τερματικού PowerShell κατά την εκτέλεση των αρχείων deploy_contract.py και setup_consent.py

```
PS C:\Users\Bitz\Downloads\Milestone_2> &
C:/Users/Bitz/AppData/Local/Programs/Python/Python311/python.exe
c:/Users/Bitz/Downloads/Milestone_2/backend/sub_storage.py
Loaded payload records: 462890 from
C:\Users\Bitz\Downloads\Milestone_2\backend\hf_outputs\biomarker
_only\blockchain_payload.json
Uploading 5 records to contract
0x3f08B8F883eC0b4E0E5e36D78A38321c43ff3D3F
[1/5] OK block=7 gas=149897 tx=107d7d3be0...
[2/5] OK block=8 gas=132809 tx=e32bccf22f...
[3/5] OK block=9 gas=132809 tx=a756db1b95...
[4/5] OK block=10 gas=132809 tx=7f0a0cdf7c...
[5/5] OK block=11 gas=149885 tx=36d0b9dbbc...
============================================================
Upload finished. success=5 failed=0 total_gas=698209
PS C:\Users\Bitz\Downloads\Milestone_2>
```

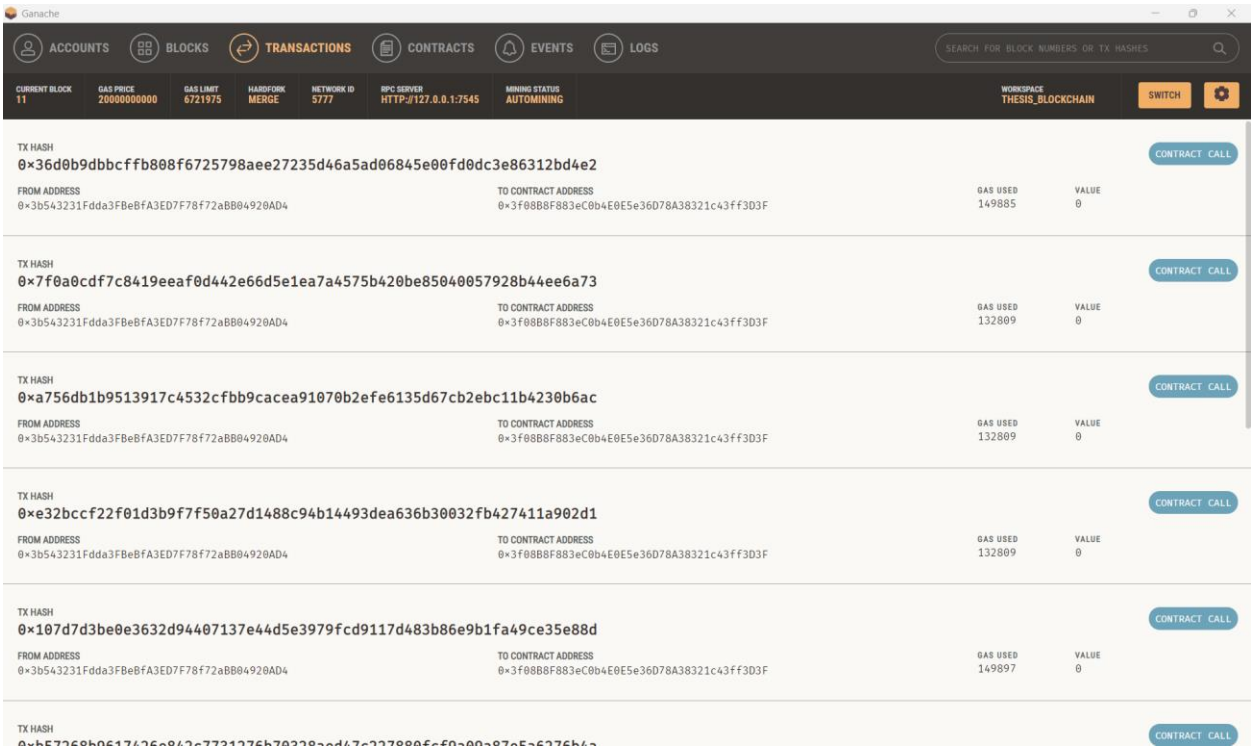
Εικόνα 39 Έξοδος του τερματικού PowerShell κατά την εκτέλεση του sub_storage.py

Μόλις γίνει η απαραίτητη καταχώριση των συγκαταθέσεων των ασθενών, εκτελείται το script sub_storage.py που αναλαμβάνει τη μεταφόρτωση των πέντε αντίστοιχων προβλέψεων κινδύνου

ΚΑ στο έξυπνο συμβόλαιο. Κάθε συναλλαγή περιλαμβάνει τέσσερα ορίσματα, τη κρυπτογραφημένη ταυτότητα του ασθενή, το σκορ κινδύνου, το κρυπτογράφημα των συνολικών δεδομένων και την έκδοση του μοντέλου. Ιδιαίτερη παρατήρηση χρήζει η αύξηση του κόστους της πρώτης και τελευταίας συναλλαγής σε σχέση με τις ενδιάμεσες κατά 17088 καύσιμα, η οποία οφείλεται στο επιπλέον κόστος αρχικοποίησης νέας θέσης μνήμης για το πρώτο ασθενή. Παραπάνω ανάλυση της διαδικασίας πραγματοποιείται στο Κεφάλαιο 6. Στην **Εικόνα 40** φαίνεται η ολοκληρωμένη κατάσταση των πέντε αναρτήσεων. Στη καρτέλα των λογαριασμών, ο πρώτος λογαριασμός εμφανίζει μειωμένο υπόλοιπο κατά 99.94 ETH και αυξημένο αριθμό συναλλαγών (κατά 11 συναλλαγές γιατί είχαν γίνει και προηγούμενες συναλλαγές κανονικά έπρεπε να είναι 5). Παράλληλα, στη καρτέλα των συναλλαγών του Ganache, εμφανίζεται η αναλυτική λίστα των συναλλαγών που είναι καταγεγραμμένες πάνω στην αλυσίδα. Κάθε συναλλαγή φέρει το μοναδικό κρυπτογράφημα συναλλαγής, τη διεύθυνση αποστολέα και τη διεύθυνση προορισμού. Στην **Εικόνα 41** διακρίνονται τα πέντε διαδοχικά κρυπτογραφικά αποτυπώματα συναλλαγών που είχαν καταχωρηθεί από τη sub_storage.py, παρέχοντας την οπτική επαλήθευση της αντιστοιχία μεταξύ του script και της αλυσίδας.



Εικόνα 40 : Κατάσταση του Ganache GUI μετά την ολοκλήρωση των 5 αναρτήσεων.

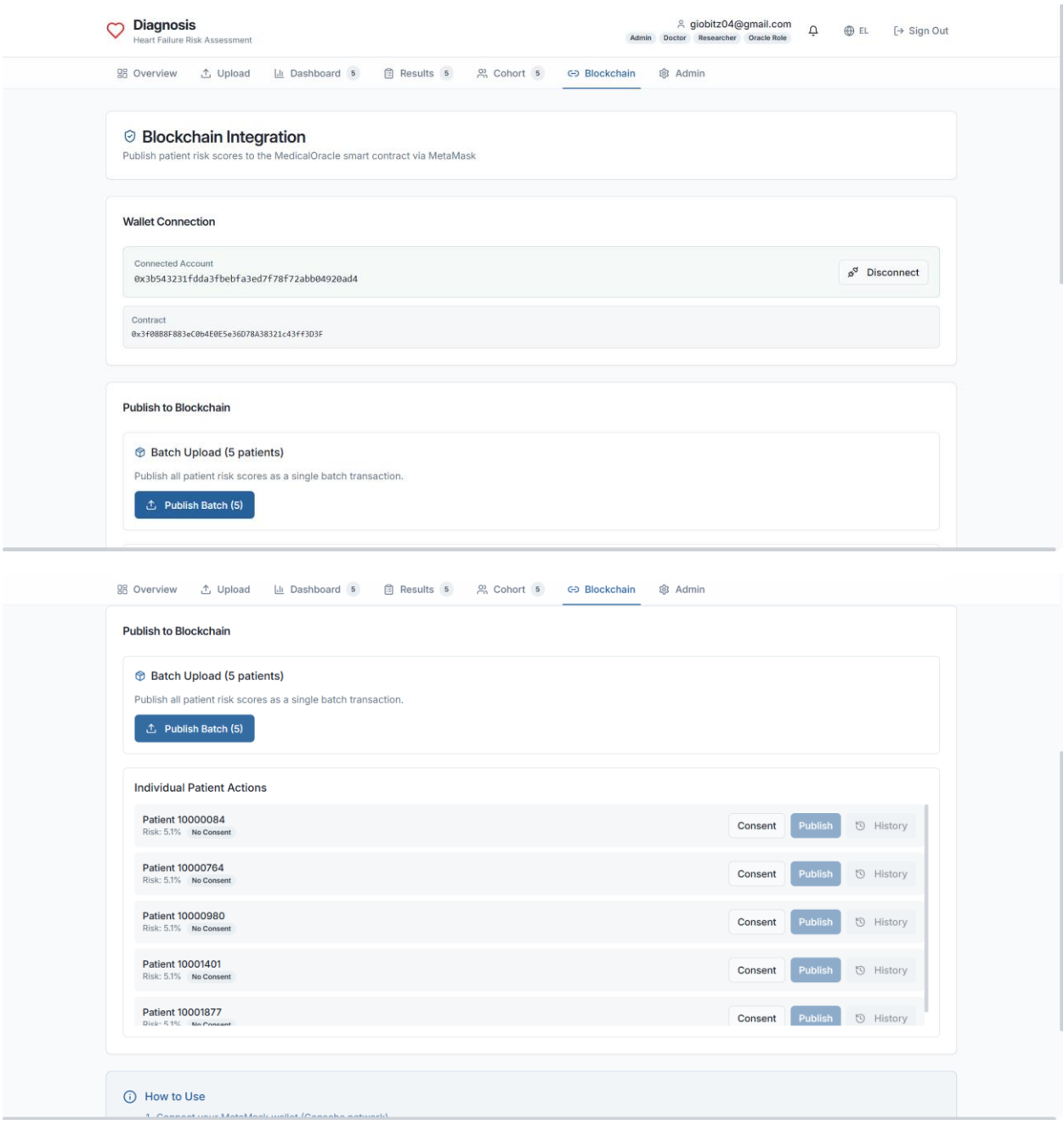


Εικόνα 41 : Καρτέλα Συναλλαγών του Ganache

```
PS C:\Users\Bitz\Downloads\Milestone_2\backend> uvicorn api:app
--host 0.0.0.0 --port 8000 --reload
INFO:     Will watch for changes in these directories:
['C:\\Users\\Bitz\\Downloads\\Milestone_2\\backend']
INFO:     Uvicorn running on http://0.0.0.0:8000 (Press CTRL+C
to quit)
INFO:     Started reloader process [13236] using StatReload
INFO:     Started server process [28560]
INFO:     Waiting for application startup.
INFO:     Application startup complete.
```

Εικόνα 42 : Έξοδος του τερματικού PowerShell κατά τον συγχρονισμό του backend με το frontend μέσω του REST API

Για τον συγχρονισμό των εγγραφών που βρίσκονται πάνω στην αλυσίδα με την υπόλοιπη κλινική πλατφόρμα ενεργοποιείται ο εξυπηρετητής REST API της κλινικής πλατφόρμας μέσω της εντολής uvicorn api:app. Στην **Εικόνα 42** η εντολή αυτή ενεργοποιεί τον Fast API εξυπηρετητή στη θύρα 8000, που αναλαμβάνει την επικοινωνία μεταξύ της διεπαφής χρήστη και του backend συνδυασμού μηχανικής μάθησης και Blockchain. Η μεταβλητή σημαίας Boolean –reload εξασφαλίζει την αυτόματη επαναφόρτιση του κώδικα σε κάθε φάση τροποποίησης.



Εικόνα 43 : Καρτέλα Blockchain της προτεινόμενης κλινικής πλατφόρμας. Στην καρτέλα απεικονίζεται και η λίστα των «Individual Patient Actions» που αφορά τους 5 ασθενείς που καταχωρήθηκαν στην αλυσίδα και τις ενέργειες που μπορούν να πραγματοποιηθούν πάνω στα δεδομένα τους.

Στο τελευταίο στάδιο ανάπτυξης πραγματοποιείται η αλληλεπίδραση μεταξύ του κλινικού ιατρού και της καρτέλας Blockchain, η οποία αποτελεί το τελικό κρίκο της end-to-end λειτουργίας. Μέσω

της βιβλιοθήκης ethers.js και του ηλεκτρονικού πορτοφολιού MetaMask, η διεπαφή συνδέεται άμεσα με το αναπτυγμένο συμβόλαιο του Ganache, παρουσιάζοντας τα στοιχεία της διεύθυνσης λογαριασμού του χρήστη και του έξυπνου συμβολαίου. Για την πρόσβαση σε παραπάνω λειτουργίες, ο γιατρός θα πρέπει να πάρει την έγκριση από τον ασθενή και να προχωρήσει στη συγκατάθεση των ιατρικών αποτελεσμάτων του μέσω του κουμπιού «Consent». Ύστερα έχει την δυνατότητα δημοσίευσης των αποτελεσμάτων για ερευνητικούς σκοπούς μέσω του κουμπιού «Publish» και το αναλυτικό ιατρικό «History». Ακόμη, γίνεται η δημοσίευση ολόκληρης της λίστας αποτελεσμάτων των ασθενών με το κουμπί «Publish Batch».

5.6 Διεπαφή Χρήστη της Κλινικής Πλατφόρμας

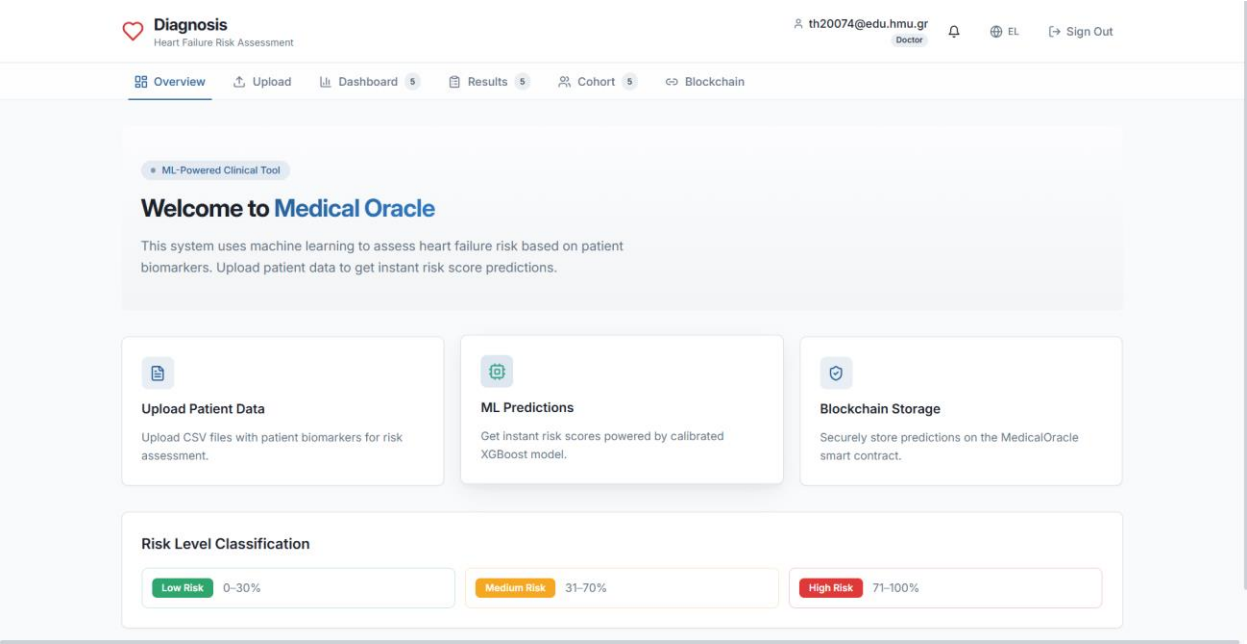
Η διεπαφή χρήστη του κλινικού συστήματος υποστήριξης αποφάσεων υλοποιείται ως μονοσέλιδη εφαρμογή με τη βιβλιοθήκη React έκδοση 18, με το μεγαλύτερο μέρος της εφαρμογής να έχει γραφτεί στη γλώσσα προγραμματισμού Typescript χρησιμοποιώντας το Vite ως εργαλείο ανάπτυξης. Ο συνδυασμός αυτών των τριών τεχνολογιών επιτρέπει την αποφυγή συντακτικών λαθών του κώδικα, την ευρεία ποικιλία έτοιμων και δοκιμασμένων εργαλείων για το περιβάλλον εργασίας σε σύγχρονο υβριδικό σύστημα, εκμηδενίζοντας χρόνους αναμονής και βελτιστοποιώντας την πλατφόρμα γρηγορότερα. Οι υποενότητες που ακολουθούν παρέχουν τη γραπτή και οπτική προβολή της λειτουργίας του συστήματος.

5.6.1 Αρχιτεκτονική της Διεπαφής Χρήστη

Η αρχιτεκτονική της διεπαφής χρήστη οργανώνεται σε τρία στρώματα που αλληλοεπιδρούν μεταξύ τους. Αρχικά ο AuthProvider υλοποιείται στο React Context API και δημιουργεί μια καθολική μεταβλητή κατάστασης σε μορφή λεξικού που περιέχει τα εξής: {user, role, session, isAuthenticated}. Ουσιαστικά ελέγχει ποιος είναι συνδεδεμένος την δεδομένη χρονική στιγμή και ποιο ρόλο έχει. Ο καθολικός χαρακτήρας αυτής της συνιστώσας επιτρέπει την μεταφορά της κατάστασης ενός χρήστη χωρίς να χρειάζεται να καταγραφή χειροκίνητα κάθε φορά. Στο κομμάτι ροής του εκτελέσιμου κώδικα της εφαρμογής παρουσιάζεται υπεύθυνο το ErrorBoundary. Το ίδιο είναι συνιστώσα κλάσης, διότι υλοποιείται στον κύκλο ζωής του React μόνο και λειτουργεί με μηχανισμό try-catch, ώστε αν κάποιο «παιδί» συνιστώσα πετάξει κατά το rendering κάποιο σφάλμα, το ErrorBoundary υποχωρεί την διεπαφή χρήστη, αποφεύγοντας το «ρήξιμο» της εφαρμογής. Έπειτα το BrowserRoute είναι ο κύριος κατανεμητής του περιβάλλοντος χρήστη με βάση το ιστορικό που λαμβάνει από το API. Δηλαδή, διαχειρίζεται την πλοήγηση μεταξύ των σελίδων. Επίσης το ProtectedRoute που χρησιμοποιεί η συνιστώσα BrowserRoute παρέχει προστασία του χρήστη κατά την πλοήγηση του στην εφαρμογή.

Η κεντρική σελίδα της εφαρμογής αποτελείται από συνιστώσες διεπαφής και με βάση αυτές χωρίζεται στις καρτέλες CSV Upload, Patient List, Dashboard Analytics, Blockchain και Admin Panel μόνο για τους διαχειριστές. Η επιλογή των καρτελών κάνει ευκολότερη τη χρήση της

εφαρμογής από το ιατρικό προσωπικό και η σειρά τοποθέτησης τους αντιστοιχεί στην πλήρη ανάλυση μιας συνεδρίας, με τον γιατρό να φορτώνει τα δεδομένα στην καρτέλα Upload, να τα εξετάζει στο Results και καταχωρεί τα αποτελέσματα που εγκρίνει ο ίδιος στο Blockchain. Η ηλεκτρονική διεύθυνση και ο ρόλος που έχει γίνει η εγγραφή του κλινικού προσώπου τοποθετούνται δεξιά πάνω, ενώ δίνεται η δυνατότητα αλλαγής της γλώσσας από ελληνικά σε αγγλικά μέσω του εικονιδίου EL.

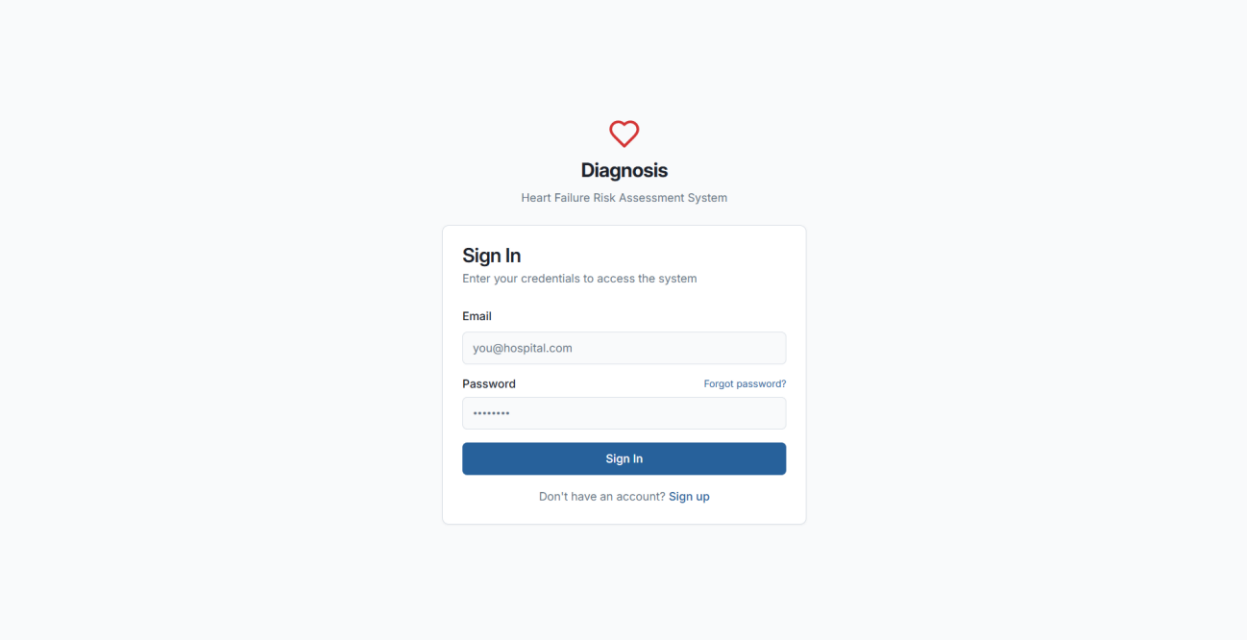


Εικόνα 44 : Κεντρική σελίδας του συστήματος υποστήριξης κλινικών αποφάσεων με το όνομα «Diagnosis»

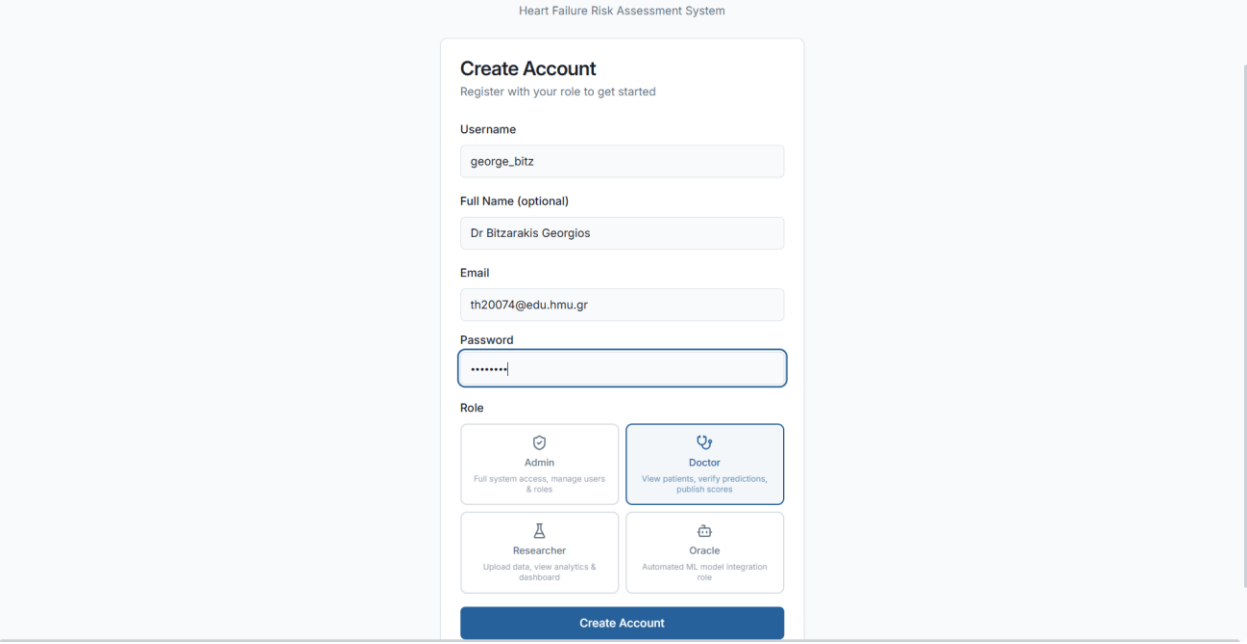
5.6.2 Προσβασιμότητα και Αυθεντικοποίηση

Η κλινική πλατφόρμα για τον έλεγχο πρόσβασης και την ταυτοποίηση του εκάστοτε χρήστη χρησιμοποιεί το Supabase. Το Supabase αποτελεί υπηρεσία διαχείρισης του backend (Backend-as-a-Service) που βασίζεται στην σχεσιακή βάση δεδομένων PostgreSQL και δίνει τη δυνατότητα στον προγραμματιστή να χρησιμοποιήσει έτοιμες υποδομές API και SDK, εστιάζοντας το ενδιαφέρον του κυρίως στον προγραμματισμό της διεπαφής χρήστη. Κατά την επιτυχή σύνδεση ο Auth Server στέλνει ένα JWT token. Η αυθεντικοποιήση χρήστη γίνεται μέσω του hook useAuth, το οποίο διαχειρίζεται δύο ασύγχρονες λειτουργίες. Η μία λειτουργία είναι ότι κάθε φορά που ο χρήστης πραγματοποιεί ενέργειες στην διεπαφή χρήστη, παραδείγματος χάρην κάνει αποσύνδεση από την πλατφόρμα, αλλάζει ή ανανεώνει την σελίδα, το Supabase ενημερώνει αυτόματα με event την κατάσταση της εφαρμογής. Μετά την αυθεντικοποίηση πραγματοποιείται η δεύτερη λειτουργία που είναι η ανάκτηση ρόλου του χρήστη. Ξεχωριστή λεπτομέρεια αξίζει να δοθεί πως ο ρόλος δεν αποθηκεύεται μέσα στο JWT token αλλά αποθηκεύεται μέσα στη βάση δεδομένων,

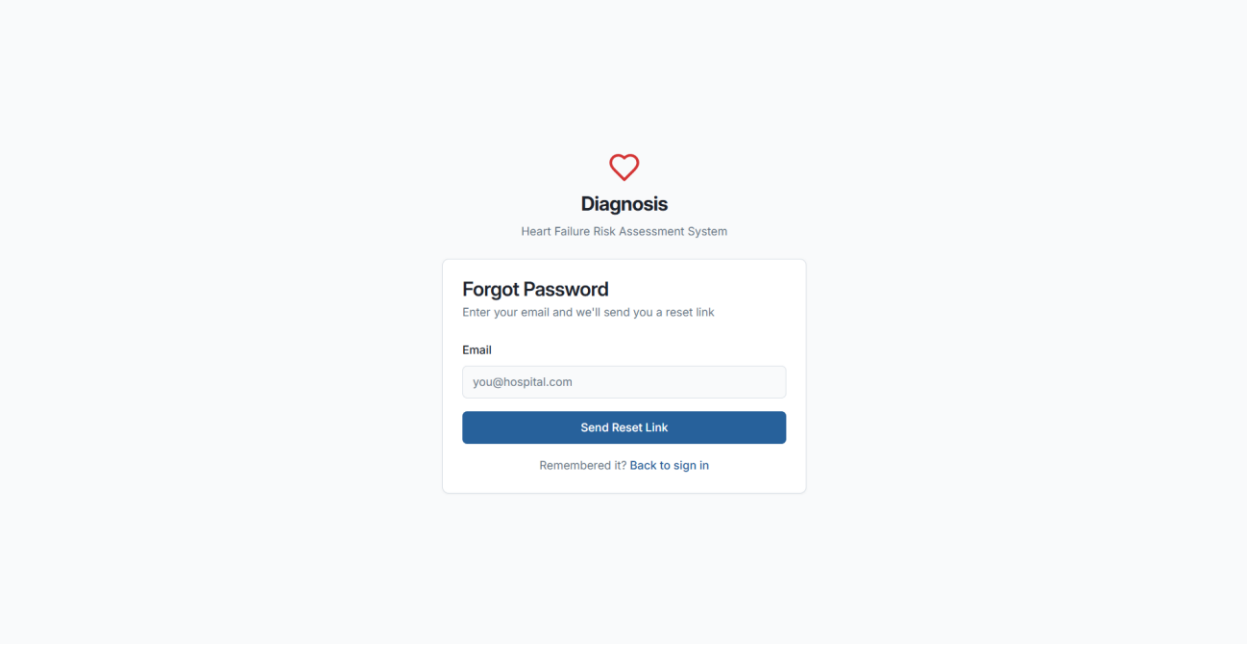
με αποτέλεσμα σε περίπτωση που αλλάξει ρόλο ο χρήστης, η αλλαγή ισχύει αμέσως χωρίς να χρειαστεί ανανέωση του token. Ύστερα τα αποτελέσματα των δύο λειτουργιών αποθηκεύονται και αποδίδονται οι αλλαγές τους σε κάθε re-render. Επιπλέον, το ProtectedRoute υλοποιεί αυθεντικοποίηση δύο παραγόντων για την πλήρη δρομολόγηση και ασφάλεια του χρήστη. Οι παραπάνω ενέργειες ικανοποιούν τις αρχές του Διαχωρισμού των Αρμοδιοτήτων (Separation of Concers), τη Μοναδική Πηγή Αλήθειας (Single Source of Truth) και Προεπιλογές Ασφαλείας (Fail-Safe Defaults).Στην **Εικόνα 45** απεικονίζεται η φόρμα σύνδεσης του χρήστη στην εφαρμογή, στην οποία δίνεται η δυνατότητα εγγραφής του σε περίπτωση που είναι πρώτη φόρα και ανάκτησης του κωδικού πρόσβασης. Στην επιλογή της νέας εγγραφής χρήστη μαζί με τα στοιχεία σύνδεσης γίνεται και η επιλογή ρόλου του, ο οποίος επιβεβαιώνεται με ηλεκτρονική διεύθυνση επαλήθευσης στο κλινικό πρόσωπο που κάνει εγγραφή στη πλατφόρμα.



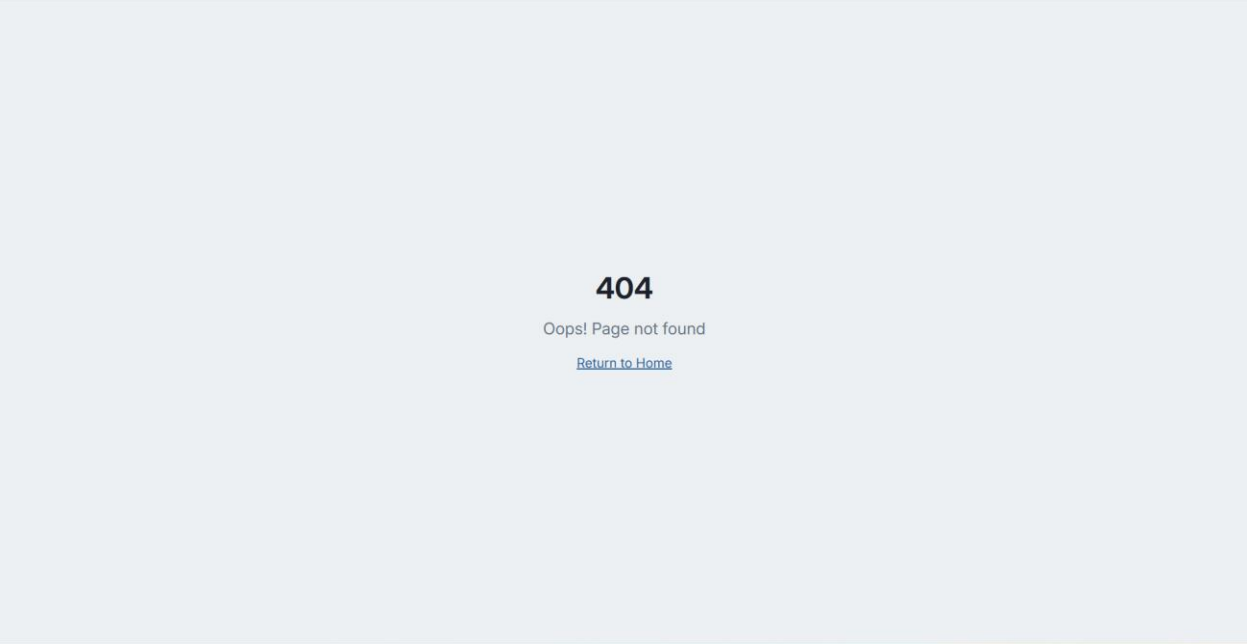
Εικόνα 45 : Σελίδα σύνδεσης χρήστη στην κλινική πλατφόρμα υποστήριξης αποφάσεων με όνομα «Diagnosis»



Εικόνα 46 : Σελίδα εγγραφής χρήστη στην κλινική πλατφόρμα υποστήριξης αποφάσεων με όνομα «Diagnosis» και επιλογή ρόλου εξουσιοδότησης.



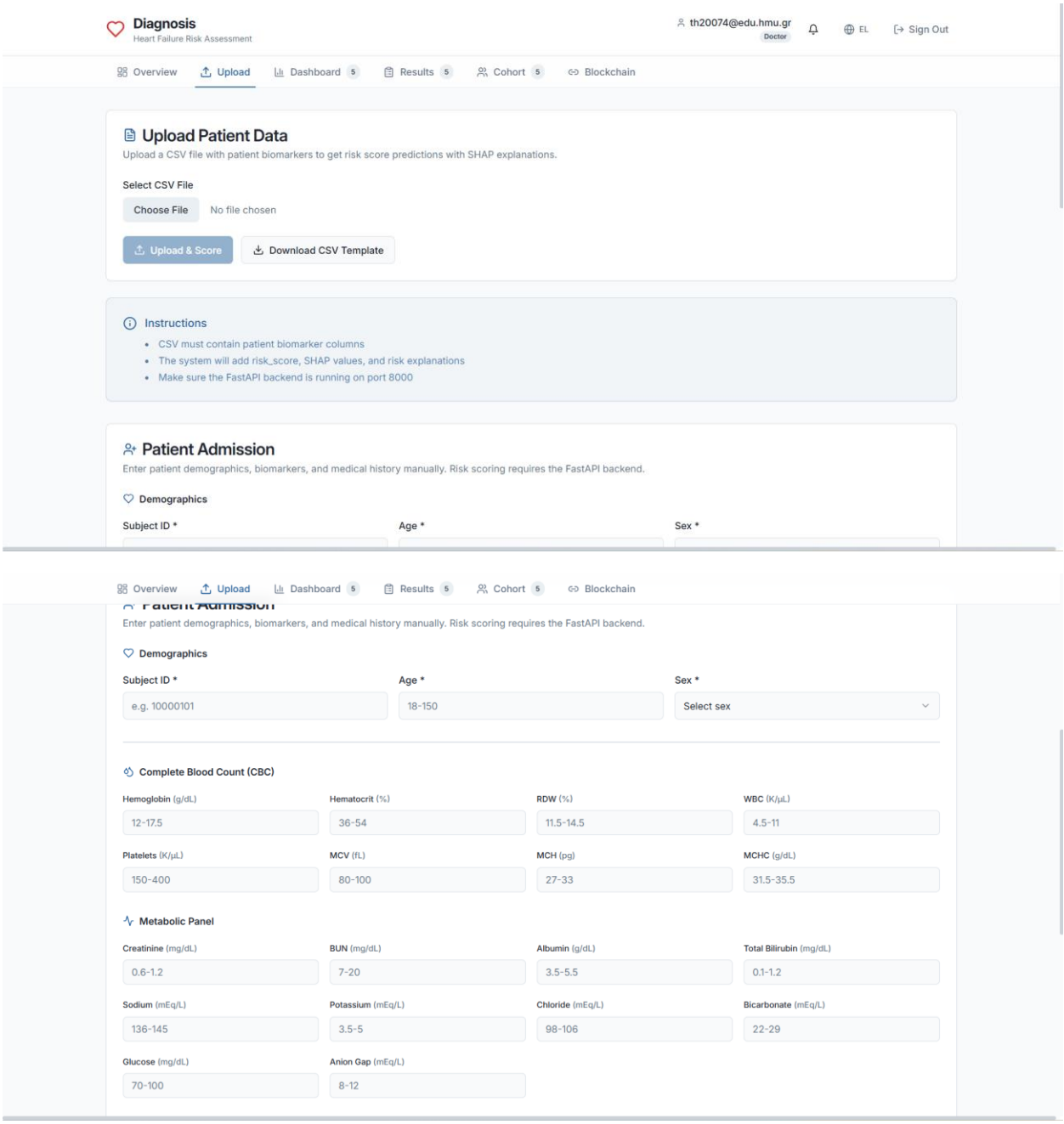
Εικόνα 47 : Σελίδα ανάκτησης κωδικού χρήστη στην κλινική πλατφόρμα υποστήριξης αποφάσεων με όνομα «Diagnosis»



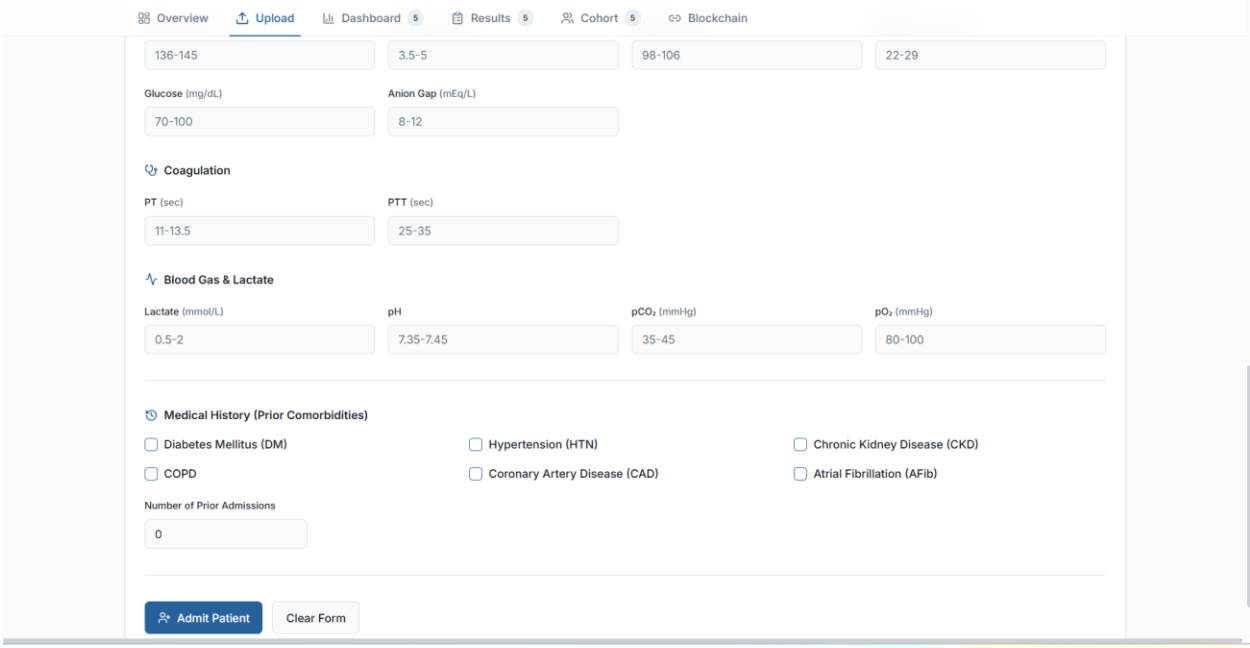
Εικόνα 48 : Μήνυμα απόρριψης πρόσβασης στη κλινική πλατφόρμα

5.6.3 Μεταφόρτωση CSV και Καταγραφή Ασθενή

Στην **Εικόνα 49** ο κλινικός γιατρός μπορεί και μεταφορτώνει αρχείο CSV με δεδομένα ασθενών μέσω drag-and-drop ή την επιλογή του από τοπικό φάκελο μέσω του «Choose File».Το αρχείο μεταφόρτωσης του παραδείγματος της εικόνας περιλαμβάνει 60 ασθενείς.Το σύστημα της πλατφόρμας στέλνει POST αίτημα στο backend και κατά την επεξεργασία του ο χρήστης ενημερώνεται με αναδυόμενο παράθυρο στο δεξί πάνω τμήμα της ιστοσελίδας. Το backend αποκρίνεται με JSON αρχείο που περιέχει την προβλεπόμενη πιθανότητα εμφάνισης καρδιακής ανεπάρκειας, τις τοπικές τιμές SHAP, το μοντέλο πρόγνωσης και τα αποτελέσματα παρουσιάζονται αναλυτικότερα στις καρτέλες Dashboard και Results. Πέραν της μεταφόρτωσης αρχείου, η κλινική πλατφόρμα δίνει την δυνατότητα χειροκίνητης καταγραφής του χρήστη.


Diagnosis
Heart Failure Risk Assessment
th20074@edu.hmu.gr
Doctor
EL
Sign Out
Overview
Upload
Dashboard 5
Results 5
Cohort 5
Blockchain
Upload Patient Data
Upload a CSV file with patient biomarkers to get risk score predictions with SHAP explanations.
Select CSV File
Choose File
No file chosen
Upload & Score
Download CSV Template
Instructions
CSV must contain patient biomarker columns
The system will add risk_score, SHAP values, and risk explanations
Make sure the FastAPI backend is running on port 8000
Patient Admission
Enter patient demographics, biomarkers, and medical history manually. Risk scoring requires the FastAPI backend.
Demographics
Subject ID *
Age *
Sex *

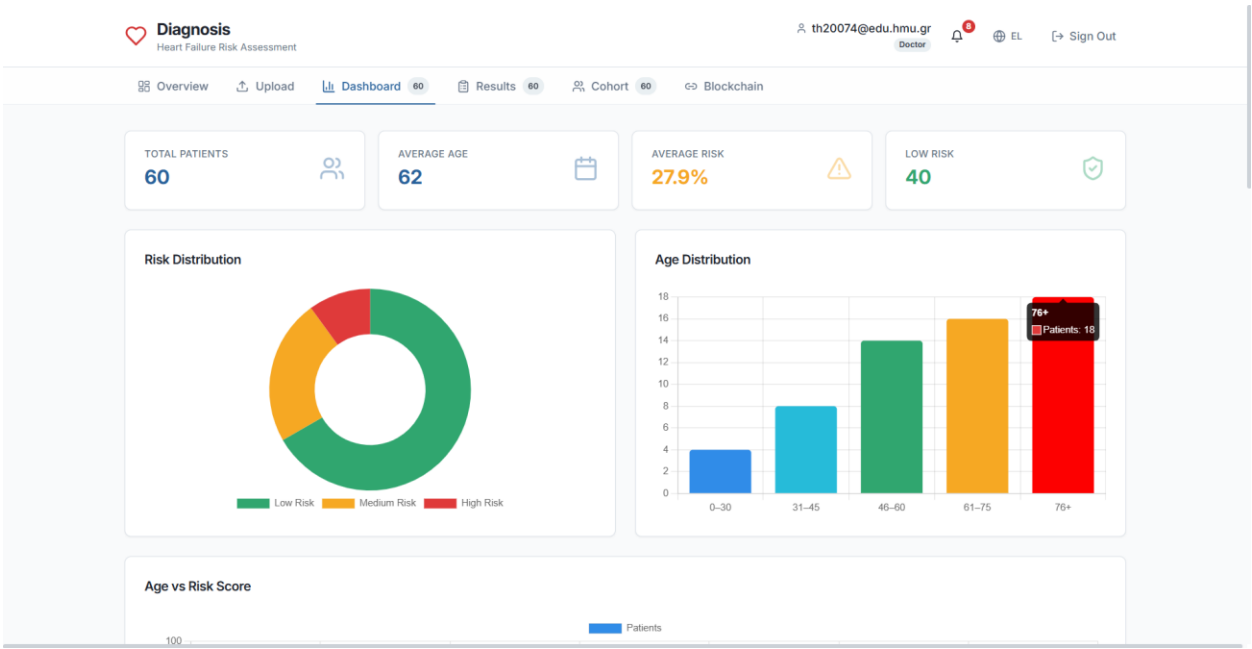
Overview
Upload
Dashboard 5
Results 5
Cohort 5
Blockchain
Enter patient demographics, biomarkers, and medical history manually. Risk scoring requires the FastAPI backend.
Demographics
Subject ID *
e.g. 10000101
Age *
18-150
Sex *
Select sex
Complete Blood Count (CBC)
Hemoglobin (g/dL)
12-17.5
Hematocrit (%)
36-54
RDW (%)
11.5-14.5
WBC (K/µL)
4.5-11
Platelets (K/µL)
150-400
MCV (fL)
80-100
MCH (pg)
27-33
MCHC (g/dL)
31.5-35.5
Metabolic Panel
Creatinine (mg/dL)
0.6-1.2
BUN (mg/dL)
7-20
Albumin (g/dL)
3.5-5.5
Total Bilirubin (mg/dL)
0.1-1.2
Sodium (mEq/L)
136-145
Potassium (mEq/L)
3.5-5
Chloride (mEq/L)
98-106
Bicarbonate (mEq/L)
22-29
Glucose (mg/dL)
70-100
Anion Gap (mEq/L)
8-12

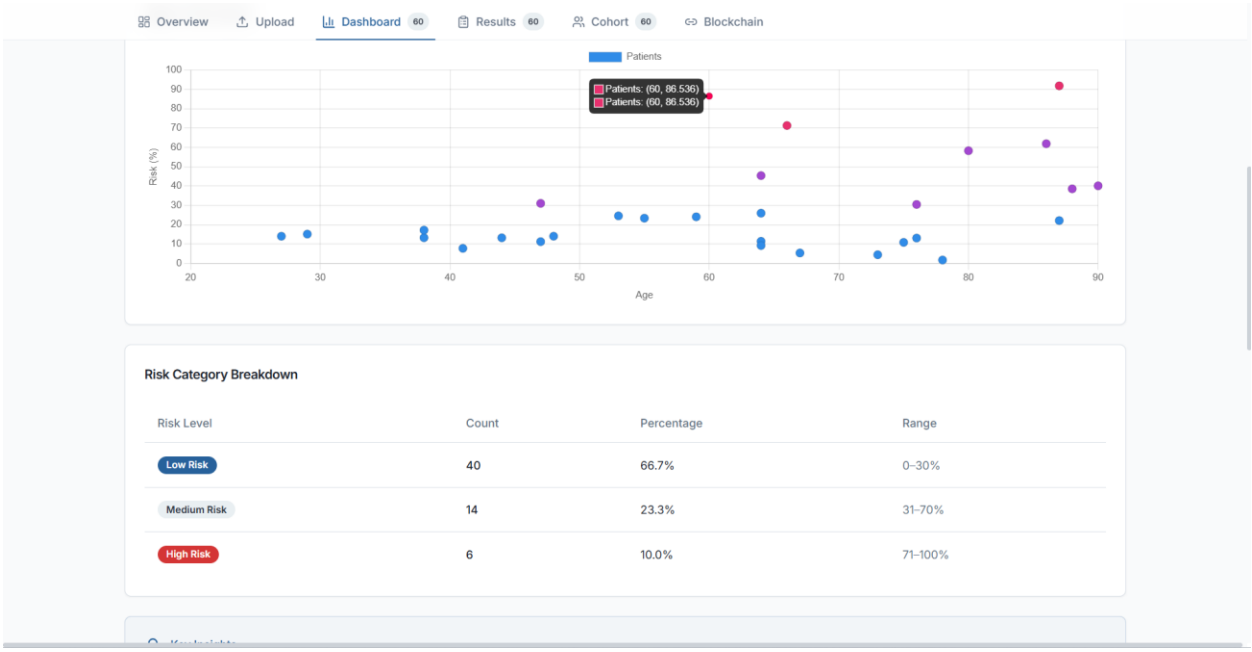


Εικόνα 49 : Καρτέλα μεταφόρτωσης και καταγραφής δεδομένων του ασθενή.

5.6.4 Πίνακας Ανάλυσης Στατιστικών

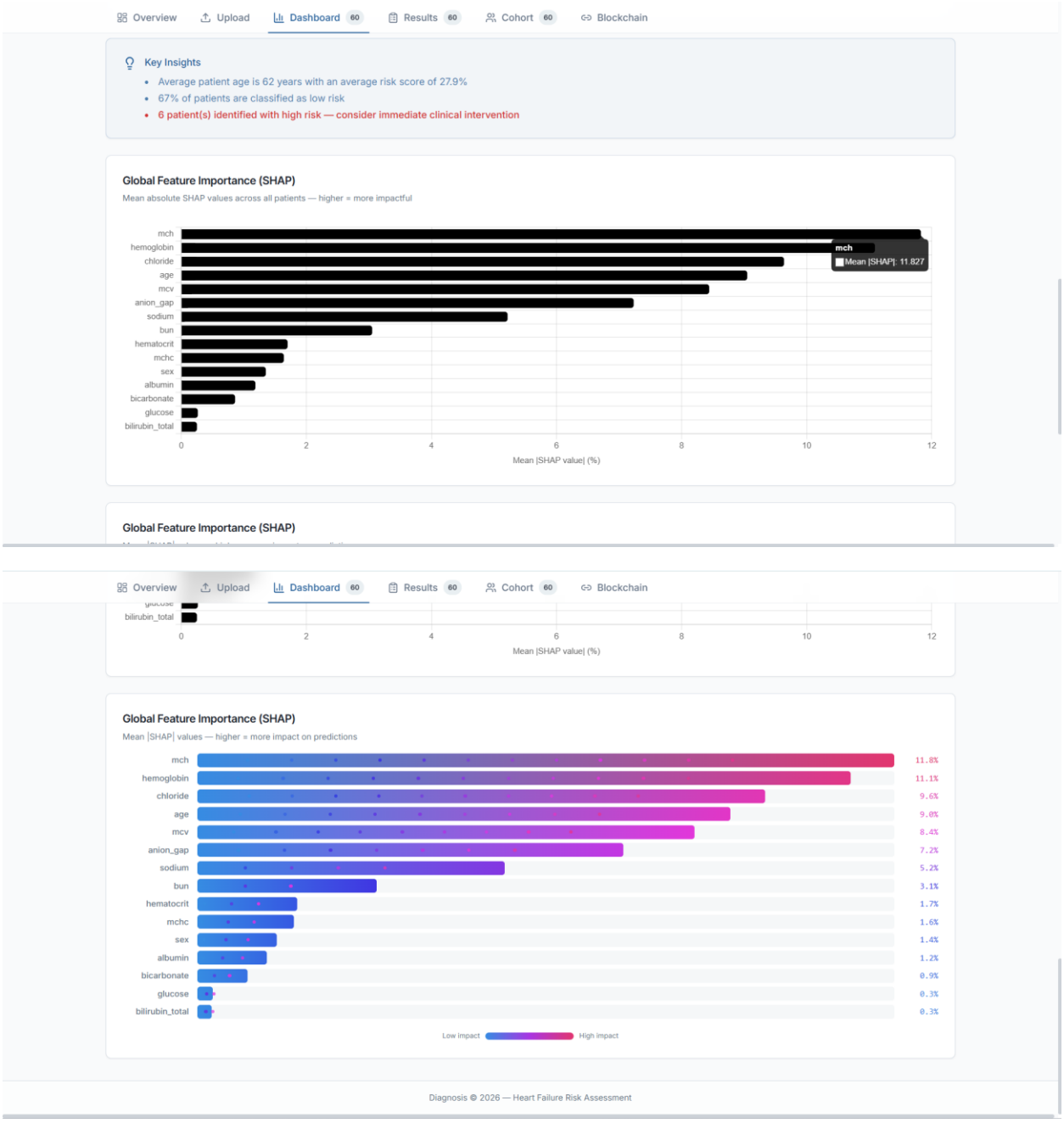
Η καρτέλα dashboard παρέχει την εποπτική ανάλυση των αποτελεσμάτων πρόβλεψης του ολικού πληθυσμού των ασθενών που έχει καταχώρηση ο κλινικός χρήστης, μέσω δια δραστικών γραφημάτων της βιβλιοθήκης Chart.js της React. Αναλυτικότερα τα τέσσερα κύρια πλαίσια στο πάνω μέρος της καρτέλας δείχνουν τον συνολικό αριθμό των ασθενών, την μέση ηλικία τους, το μέσο ρίσκο καρδιακής ανεπάρκειας που εμφανίζουν και τελευταίο παρουσιάζει το ποσοστό πιθανότητας καρδιακής ανεπάρκειας που εμφανίζει η πλειοψηφία των ασθενών. Παρακάτω δίνονται 5 χρήσιμα γραφήματα απεικόνισης των αποτελεσμάτων και ταξινόμησης των ασθενών με βάση τα χαρακτηριστικά τους. Στην κατανομή ρίσκου (Risk Distribution), ένα διάγραμμα δακτυλίου εμφανίζει την κατανομή των ασθενών σε τρείς ζώνες κινδύνου, τη χαμηλή, τη μέτρια και την υψηλή. Στο ιστόγραμμα ηλικίας δείχνει τη μέση πιθανότητα καρδιακής ανεπάρκειας ανά ηλικιακή ομάδα. Το διάγραμμα διασποράς ταξινομεί τους ασθενείς ως σημεία στο διάγραμμα και συγκρίνει τα χαρακτηριστικά της πιθανότητας καρδιακής ανεπάρκειας και ηλικίας. Στο τέλος απεικονίζονται δύο SHAP διαγράμματα που κατατάσσουν τους 15 πιο κρίσιμους βιοδείκτες που επηρεάζουν στην πιθανότητα πρόβλεψης της καρδιακής ανεπάρκειας. Η ύπαρξη δύο ξεχωριστών διαγραμμάτων SHAP οφείλεται στην παροχή περισσότερης οπτικοποίησης των τιμών SHAP, καθώς το πρώτο διάγραμμα απεικονίζει τη μέση τιμή SHAP και τον αντίκτυπο που έχει κάθε τιμή με χρωματική απεικόνιση.





Risk Category Breakdown

Risk Level	Count	Percentage	Range
Low Risk	40	66.7%	0–30%
Medium Risk	14	23.3%	31–70%
High Risk	6	10.0%	71–100%

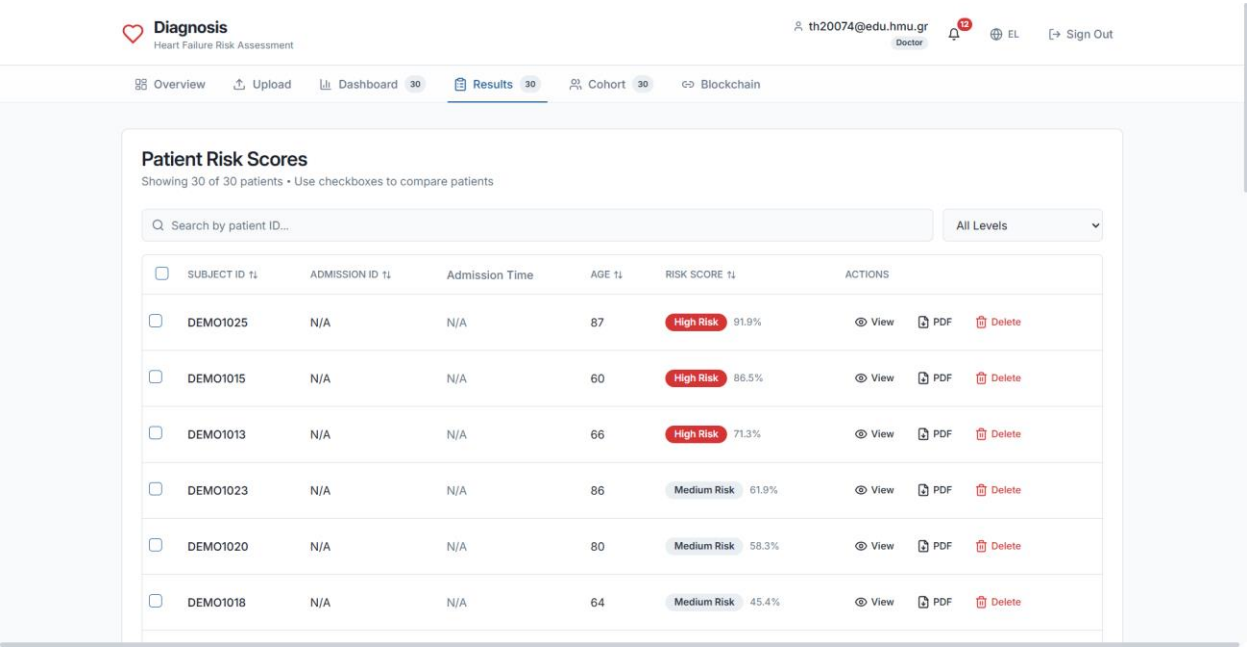


Εικόνα 50 : Καρτέλα Dashboard απεικόνισης των αποτελεσμάτων ενός πληθυσμού ασθενών με συγκεντρωτικά γραφήματα

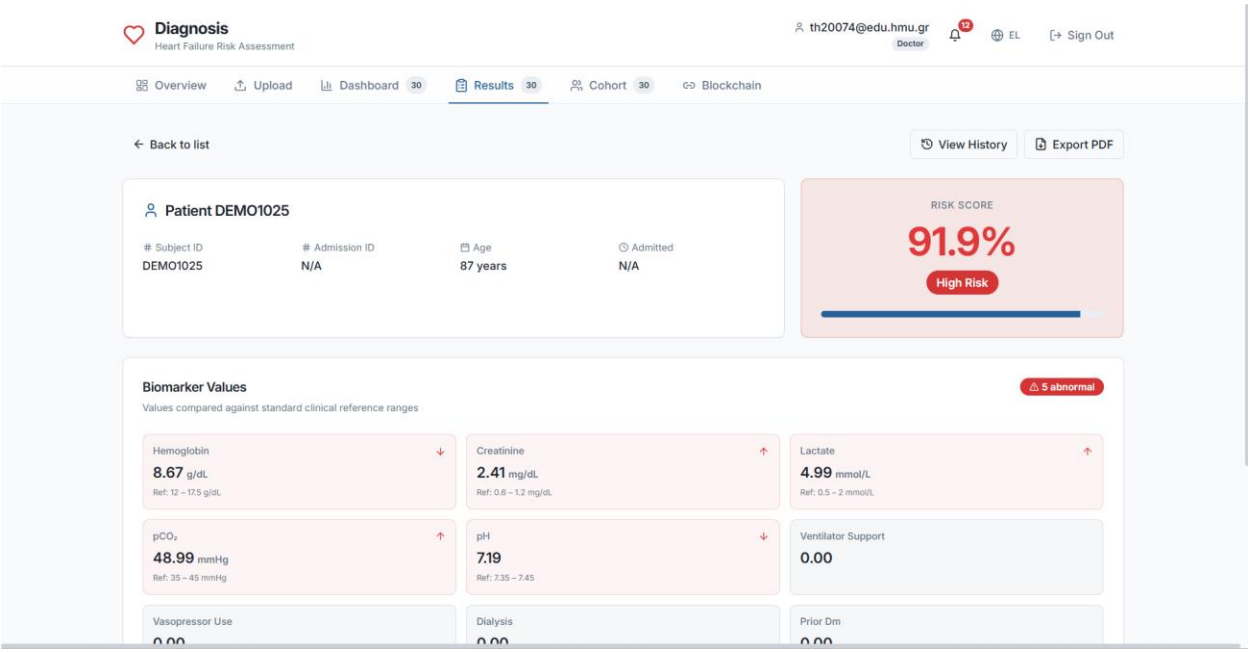
5.6.5 Αποτελέσματα και Λίστα ασθενών

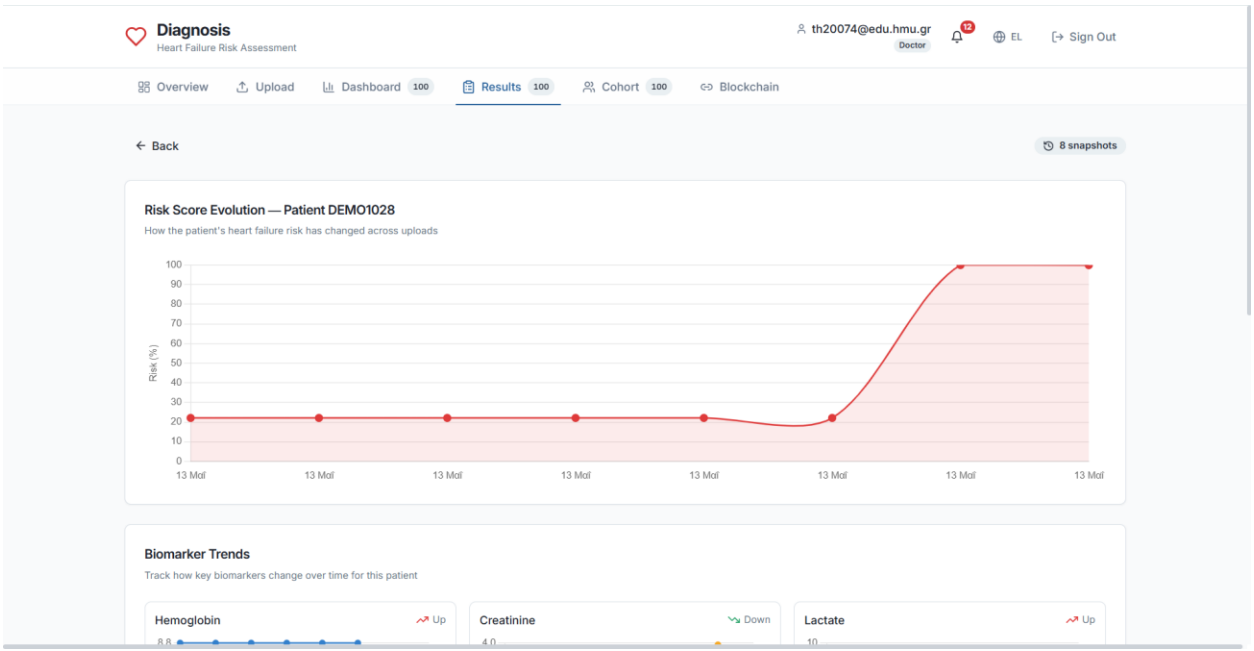
Στη καρτέλα Results απεικονίζονται τα αποτελέσματα των καταχωρήσεων που πραγματοποίησε ο ιατρός. Για την απεικόνιση των αποτελεσμάτων οφείλεται η συνιστώσα PatientList που εμφανίζει τον πίνακα ασθενών με χρωματική κωδικοποίηση κινδύνου. Με την επιλογή συγκεκριμένου

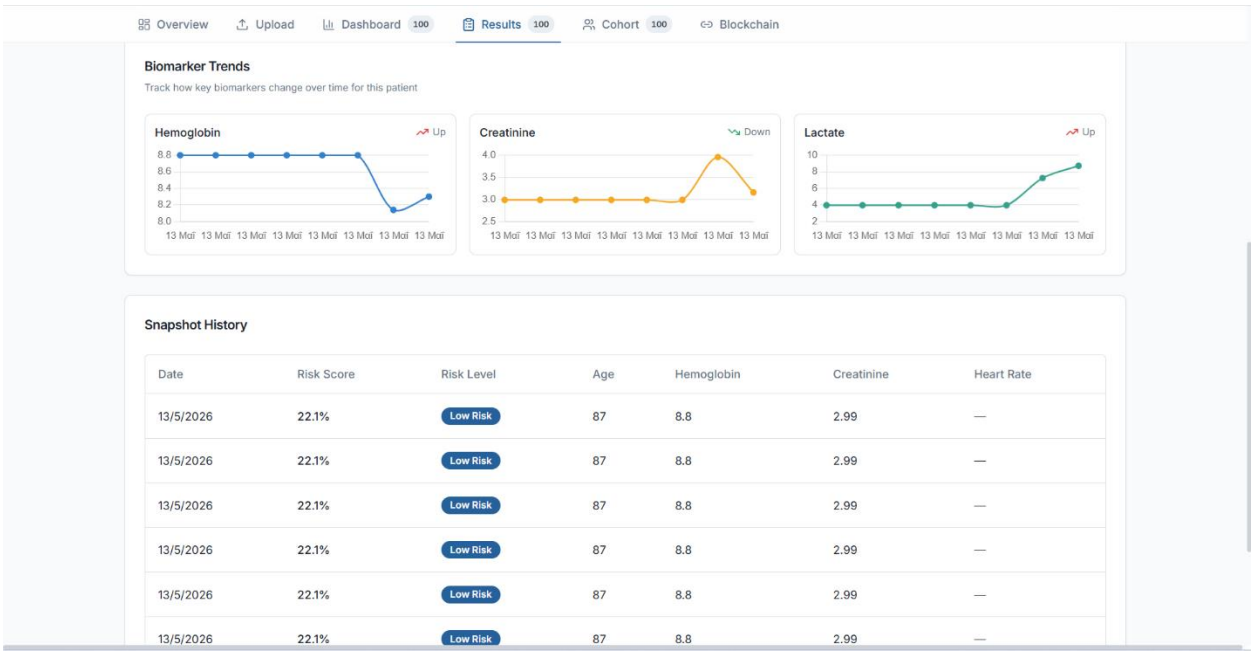
ασθενή, ενεργοποιείται η συνιστώσα PatientDetail που παρουσιάζει το πλήρες προφίλ του και κατασκευάζει ένα ειδικό διάγραμμα καταρράκτη SHAP με βάση τα δεδομένα του ασθενή.

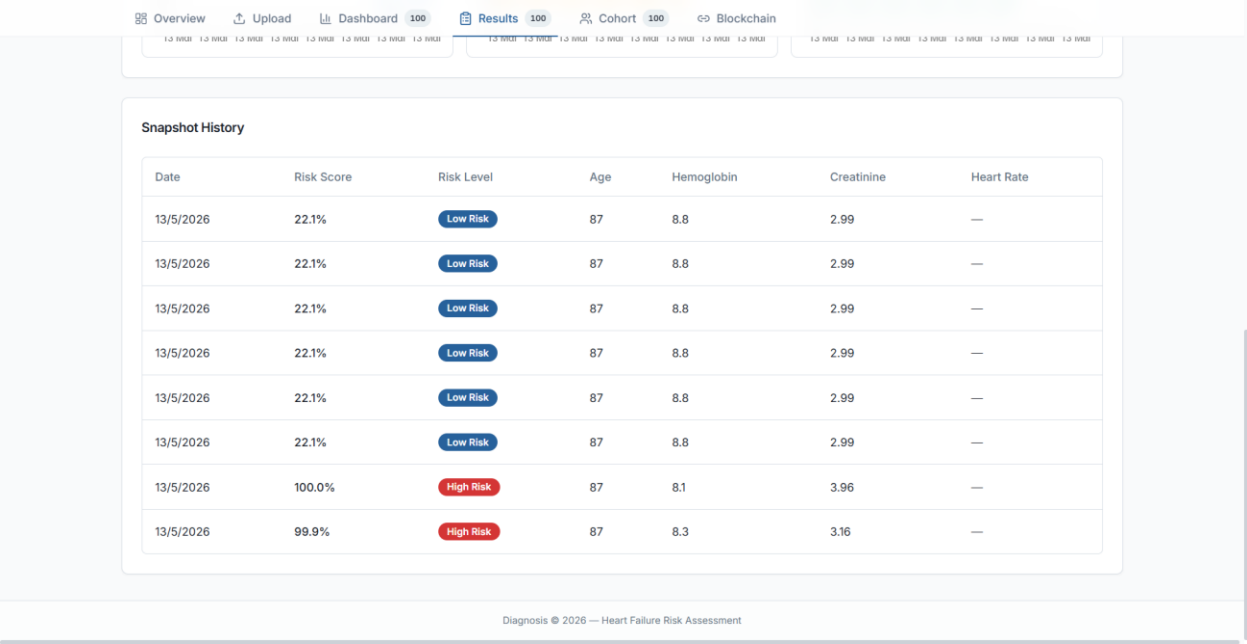


Εικόνα 51 : Λίστα ασθενών στην καρτέλα των αποτελεσμάτων (Results). Η λίστα δίνει την δυνατότητα της ταξινόμησης και του φιλτραρισμάτων των ασθενών.




Diagnosis
Heart Failure Risk Assessment
th20074@edu.hmu.gr
Doctor
12
EL
Sign Out
Overview
Upload
Dashboard 100
Results 100
Cohort 100
Blockchain
Back
8 snapshots
Risk Score Evolution — Patient DEMO1028
How the patient's heart failure risk has changed across uploads
Risk (%)
100
90
80
70
60
50
40
30
20
10
0
13 Maï
Biomarker Trends
Track how key biomarkers change over time for this patient
Hemoglobin
Up
Creatinine
Down
Lactate
Up

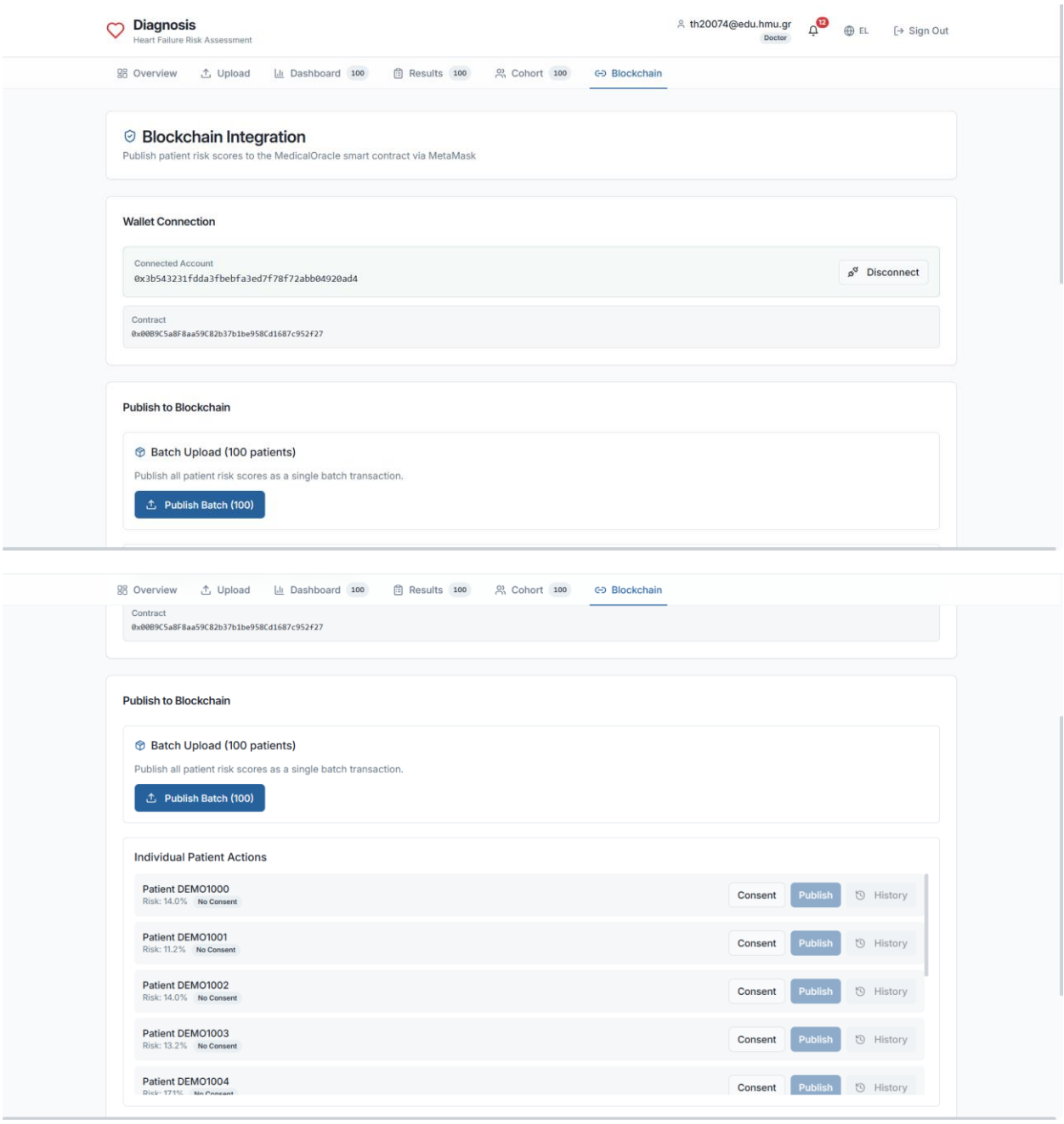

Overview
Upload
Dashboard 100
Results 100
Cohort 100
Blockchain
Biomarker Trends
Track how key biomarkers change over time for this patient
Hemoglobin
Up
8.8
8.6
8.4
8.2
8.0
Creatinine
Down
4.0
3.5
3.0
2.5
Lactate
Up
10
8
6
4
2
13 Maï
Snapshot History
Date
Risk Score
Risk Level
Age
Hemoglobin
Creatinine
Heart Rate
13/5/2026
22.1%
Low Risk
87
8.8
2.99
—



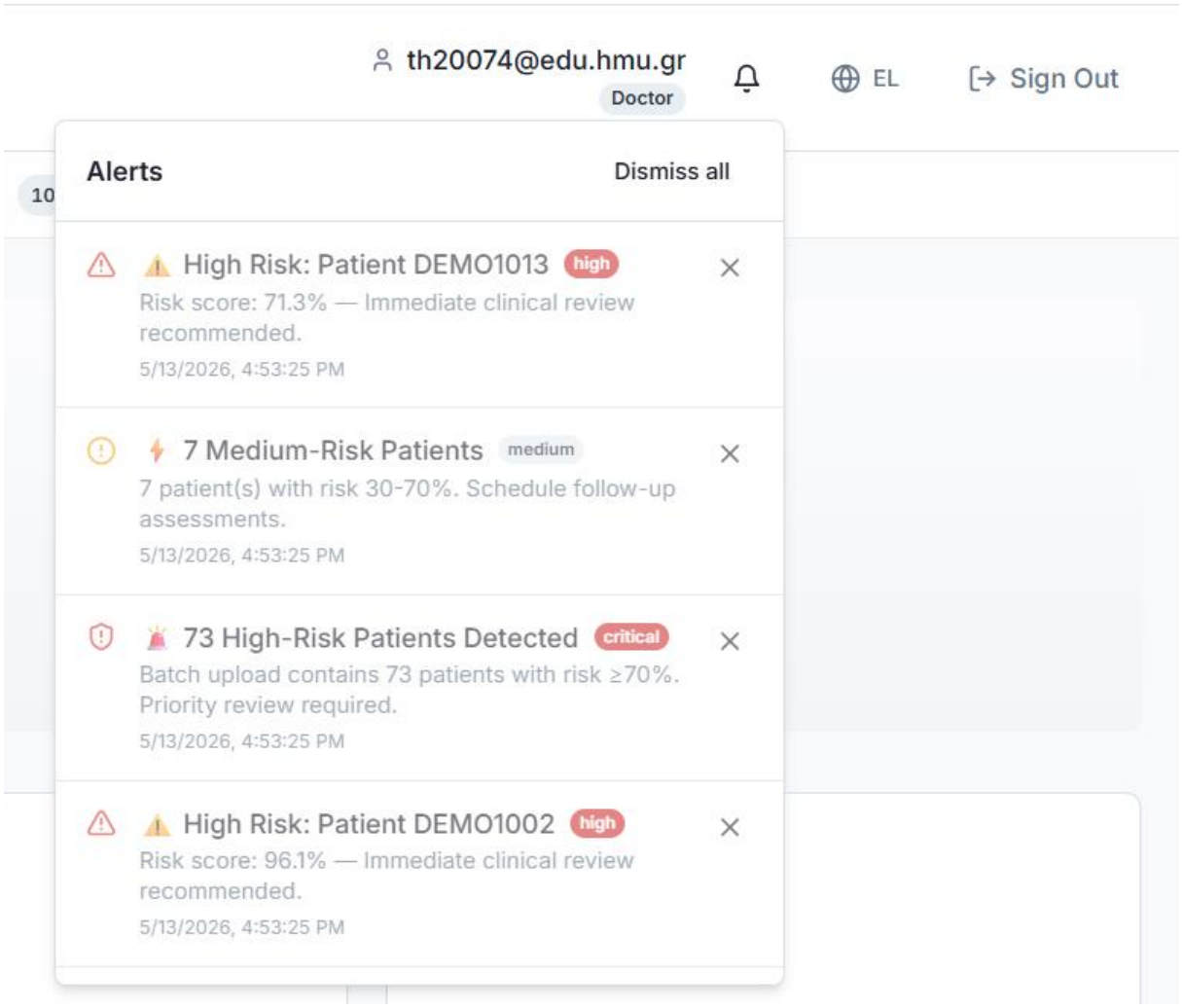
Εικόνα 52 : Εξατομικευμένο SHAP Διάγραμμα Καταρράκτη για ασθενή. Στο συγκεκριμένο απεικονίζεται η υψηλή πιθανότητα κινδύνου και το ιατρικό ιστορικό του.

5.6.6 Καρτέλα Blockchain

Στην καρτέλα Blockchain πραγματοποιείται η σύνδεση της αλυσίδας με την διεπαφή της κλινικής πλατφόρμας. Η καρτέλα ζητάει από το χρήστη την σύνδεση του Metamask πορτοφολιού πατώντας το κουμπί «Connect» και με την συμπλήρωση των στοιχείων πρόσβασης του πορτοφολιού, η συνιστώσα BrowserProvider παρέχει ζωντανή σύνδεση με το έξυπνο συμβόλαιο που τρέχει στο Ganache. Πιο συγκεκριμένα, η συνιστώσα διαβάζει την διεύθυνση του ανατιθέμενου συμβολαίου και το ABI από στατικό αρχείο που υπάρχει στη παρούσα διπλωματική εργασία και μετατρέπει τις δημόσιες και εξωτερικές συναρτήσεις της Solidity σε εκτελέσιμες μεθόδους Javascript. Ο διαχωρισμός των κλήσεων ανάγνωσης και των συναλλαγών γραφής ειναι σαφής. Οι κλήσεις αναγνώσεις καλούνται χωρίς να κοστίζουν gas. Σε αυτές ο γιατρός ανοίγει την καρτέλα Blockchain και μέσω της βιβλιοθήκης ethers.js ο γιατρός ανοίγει την καρτέλα Blockchain και διαβάζει τις εγγραφές που είναι γραμμένες on-chain ή αυτές που έχουν επαληθευτεί. Εάν κάποια καταγεγραμμένη νοσηλεία δεν έχει καταγραφή ο γιατρός αναλαμβάνει την ανάρτηση της στην αλυσίδα χειροκίνητα με ψηφιακή υπογραφεί στο αναρτώμενο μήνυμα του Metamask. Παράλληλα με αυτά δίνεται η δυνατότητα στο κλινικό γιατρό να κοινοποιήσει τα αποτελέσματα για την αξιοποίηση τους από το ερευνητικό προσωπικό με την επιλογή «Publish», εφόσον ο ασθενής έχει συμφωνήσει με τη συγκατάθεση δημοσίευσης των αποτελεσμάτων πρόβλεψης. Για το ταχύτερο ανέβασμα των αποτελεσμάτων γίνεται τα αποτελέσματα να ανέβουν στο δίκτυο σε παρτίδες, μέσω του μηχανισμού MerkleTree του έξυπνου συμβολαίου.



Εικόνα 53 : Καρτέλα Blockchain της εφαρμογής. Μετά την σύνδεση με το πορτοφόλι Metamask ο κλινικός γιατρός έχει τη δυνατότητα να επιβλέπει τις εγγραφές νοσηλείας των ασθενών.

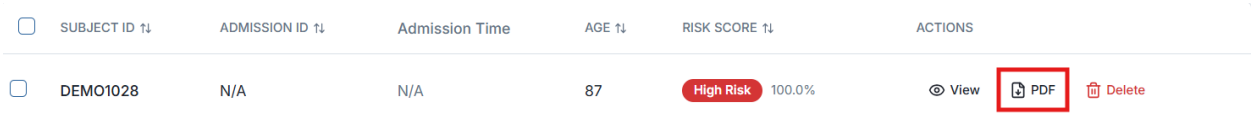


Εικόνα 54 : Αναδυόμενο παράθυρο ενημέρωσης της κατάστασης των ασθενών. Στους ασθενείς που εμφανίζουν υψηλό ποσοστό κινδύνου καρδιακής ανεπάρκειας αναδύεται δεύτερο παράθυρο ειδοποίησης.

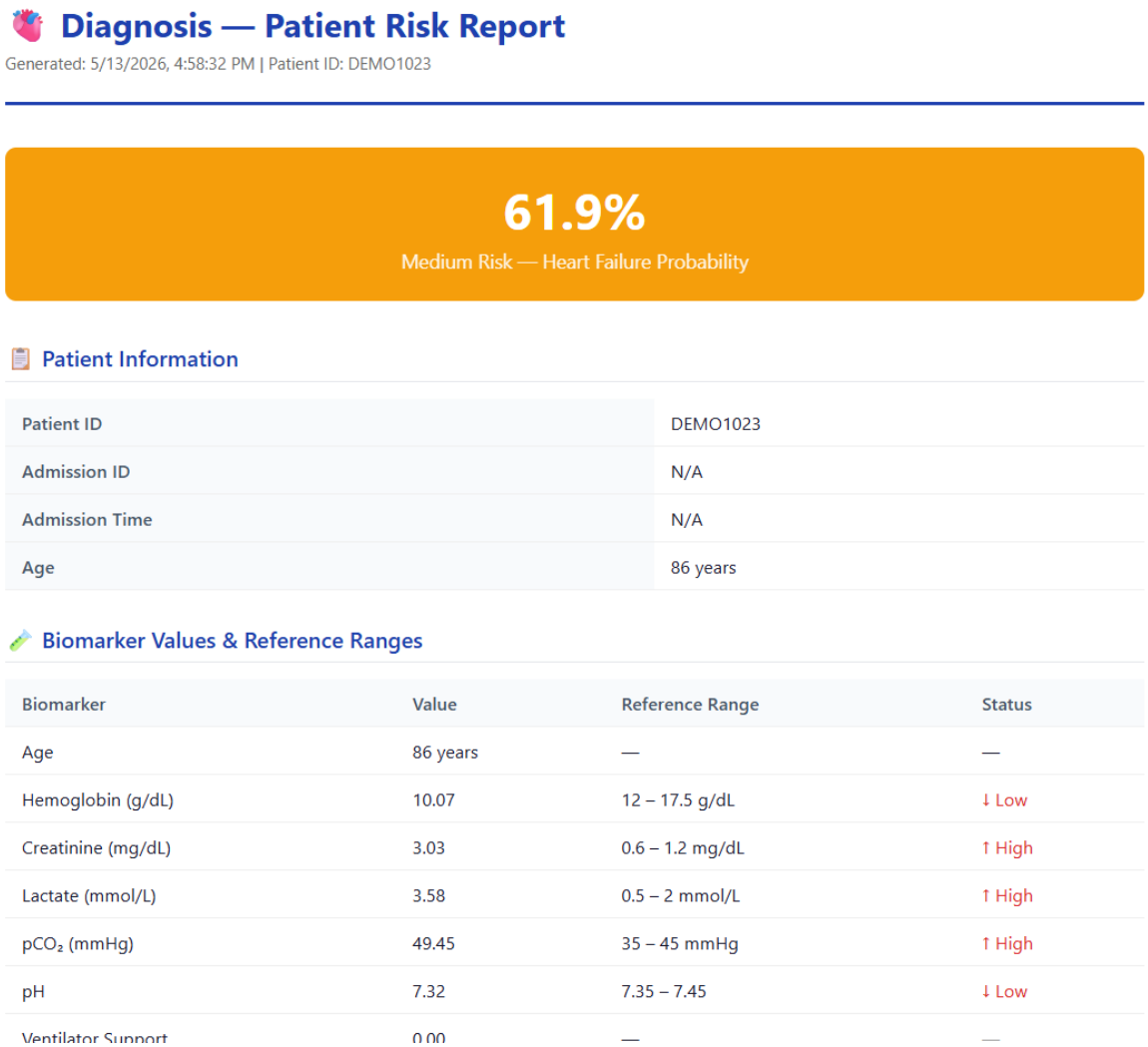
5.6.7 Παραγωγή PDF ανά ασθενή

Όταν ο κλινικός επιθυμεί να αποθηκεύσει την ανάλυση ενός ασθενούς σε φάκελο του συστήματος, η εφαρμογή παράγει ένα PDF αρχείο, με την επιλογή του ανάλογου κουμπιού, μέσω της βιβλιοθήκης jsPDF. Στο παραγόμενο PDF περιλαμβάνονται τα δημογραφικά χαρακτηριστικά και οι βιοδείκτες του ασθενούς, η τιμή ρίσκου κινδύνου ΚΑ, η χρονοσφραγίδα καταγραφής της εξέτασης, η ταυτότητα του αρμόδιου ιατρού και οι πέντε σημαντικότεροι παράγοντες, είτε δημογραφικοί είτε βιολογικοί που επηρέασαν στη πρόβλεψη. Η επιλογή του τύπου PDF εξυπηρετεί στην κλινική τεκμηρίωση των αποτελεσμάτων με την δυνατότητα εκτύπωσης σε φυσική μορφή και αποθήκευση των αποτελεσμάτων στο φάκελο του ασθενούς ή την κοινοποιήσει τους σε άλλους θεραπεύοντες μέσω ηλεκτρονικού ταχυδρομείου. Η παραγωγή του αρχείου γίνεται

εξ ολοκλήρου στη πλευρά του χρήστη (client-side) και όχι στου εξυπηρετητή, εξασφαλίζοντας την αποδοτικότερη διαχείριση των ιατρικών αποτελεσμάτων.



Εικόνα 55 Παράδειγμα ασθενούς από την λίστα κινδύνου των ασθενών στην καρτέλα Results



pCO_2 (mmHg)	49.45	35 – 45 mmHg	↑ High
pH	7.32	7.35 – 7.45	↓ Low
Ventilator Support	0.00	—	—
Vasopressor Use	0.00	—	—
Dialysis	0.00	—	—
Prior Dm	0.00	—	—
Prior Htn	0.00	—	—
Prior Ckd	0.00	—	—
Prior Copd	0.00	—	—
Prior Cad	0.00	—	—
Prior Afib	0.00	—	—
Num Prior Adm	0.00	—	—

Feature Importance (Model Weights)

Feature	Coefficient	Impact
MCH (pg)	-1.2862	
MCV (fL)	+1.0887	
Age (years)	+0.9310	
Chloride (mEq/L)	-0.8162	
Hemoglobin (g/dL)	+0.8057	

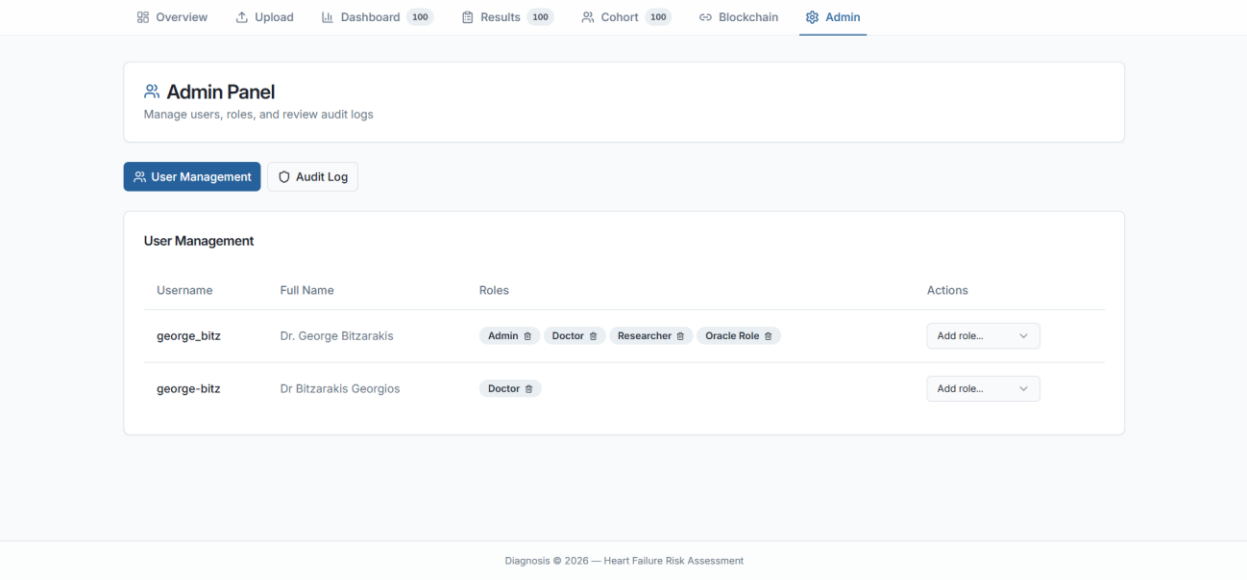
⚠ **Clinical Disclaimer:** This report is generated by a machine learning model (XGBoost) and should be reviewed by a qualified clinician before any clinical decisions are made. SHAP values indicate model reasoning, not causal relationships.

Εικόνα 56 : Παράδειγμα παραγόμενης κλινικής αναφοράς PDF του ασθενή DEMO1023. Η αναφορά περιλαμβάνει τα δεδομένα του ασθενούς, τη πρόβλεψη του μοντέλου και την ερμηνεία της μέσω των κυρίαρχω παραγόντων SHAP

5.6.8 Πάνελ Διαχείρισης Χρηστών

Στην σελίδα http://localhost:8080/admin πρόσβαση έχει μόνο ο εξουσιοδοτημένος διαχειριστής της κλινικής πλατφόρμας. Ο διαχειριστής διαθέτει στην διεπαφή χρήστη μια επιπρόσθετη καρτέλα την «Admin» που από εκεί διαχειρίζεται την ανάθεση ή αφαίρεση ρόλων στους υπάρχοντες λογαριασμούς. Η ταξινόμηση των ρόλων γίνεται μέσω του πίνακα user_roles της Supabase που αναφέρθηκε προηγουμένως. Εκτός από τη διαχείριση ρόλων, το Πάνελ Διαχειριστή προσφέρει στοιχεία εποπτείας της λειτουργίας της πλατφόρμας, όπως τα στατιστικά συνδέσεων, το ιστορικό μεταφορτώσεων CSV αρχείο ανά χρήστη και κατάλογο πρόσφατων ενεργειών του συστήματος σε

πραγματικό χρόνο. Οι συγκεκριμένες λειτουργίες προσφέρουν στον διαχειριστή την δυναμική εποπτεία του συστήματος για την διασφάλιση της ομαλής λειτουργίας του.



Εικόνα 57 : Πάνελ Διαχειριστή. Ο Διαχειριστής (Admin) διαχειρίζεται την πρόσβαση και εξουσιοδότηση ρόλων από υπάρχοντες χρήστες

Κεφάλαιο 6: Αποτελέσματα και Αξιολόγηση

Το παρόν κεφάλαιο αποσκοπεί στην αξιολόγηση της προτεινόμενης πλατφόρμας υποστήριξης κλινικών αποφάσεων και εξετάζει τη συνολική εμπειρία του χρήστη. Αντίθετα με το Κεφάλαιο 5 που αποτελούσε την πειραματική υλοποίηση της, το Κεφάλαιο 6 παρουσιάζει τα μετρήσιμα αποτελέσματα που προκύπτουν από την εκτέλεση του ολοκληρωμένου συστήματος και στην απόδοση των μοντέλων που εκπαιδεύτηκαν στα πραγματικά κλινικά δεδομένα του MIMIC-IV. Το κεφάλαιο οργανώνεται σε επτά υποενότητες για ευρεία αξιολόγηση ευχρηστίας της πλατφόρμας, καλύπτοντας από την πειραματική διαμόρφωση και σύγκριση των υποψήφιων μοντέλων μηχανικής μάθησης, μέχρι την κλινική ερμηνευσιμότητα των προβλέψεων και λειτουργία του Blockchain επιπέδου.

Η βασική αρχή που διέπει την ορθή αξιολόγηση είναι η αυστηρή απομόνωση του συνόλου ελέγχου. Όλες οι εκπαιδεύσεις, συγκρίσεις των μοντέλων μηχανικής μάθησης και βελτιστοποιήσεις παραμέτρων που πραγματοποιήθηκαν στο pipeline χρησιμοποιούν αποκλειστικά τα σύνολα βαθμονόμησης και ελέγχου, όπως αναφέρεται στην ενότητα 5.1.5. Ο λόγος είναι η αμερόληπτη παραγωγή κάθε μετρικής, κάθε γραφήματος και τιμής που παρουσιάζεται στη συνέχεια.

6.1 Υπολογιστικό Περιβάλλον και Σύνολο Δεδομένων

Στην παρούσα ενότητα τεκμηριώνονται τα αποτελέσματα σε τρεις διαστάσεις της διαμόρφωσης, το υπολογιστικό περιβάλλον, τα δεδομένα εισόδου και τις μετρικές αξιολόγησης. Για την μέγιστη αποτελεσματικότητα του πειράματος χρησιμοποιήθηκε η πλατφόρμα cloud computing Paperspace της DigitalOcean. Το υπολογιστικό περιβάλλον στο οποίο διεξήχθησαν οι πειραματικές εκτελέσεις αναφέρεται στους παρακάτω πίνακες.

Πίνακας 6.1: Πίνακας Υλικής υποδομής που χρησιμοποιήθηκε για την εκπαίδευση των υποψήφιων μοντέλων μηχανικής μάθησης.

Υλικό	**Προδιαγραφή**
GPU	NVIDIA A4000 16GB VRAM
CPU	Intel Xeon Gold 5315Y (3.20 GHz Base, 3.60 GHz Turbo)
RAM	45 GB
Αποθηκευτικός Χώρος	SSD

Πίνακας 6.2: Πίνακας Βιβλιοθηκών και Λογισμικού που χρησιμοποιήθηκε για την εκπαίδευση των υποψήφιων μοντέλων μηχανικής μάθησης.

Βιβλιοθήκη	**Έκδοση**
Python	3.11
Scikit-learn	1.8.0
XGBoost	2.1.0
LightGBM	4.6.0
Optuna	4.8.0
SHAP	0.51.0
Pandas	3.0.3
Numpy	2.4.4
Matplotlib	3.7.3
Imbalanced-Learn	0.14.1

Πίνακας 6.3: Πίνακας Blockchain Λογισμικού που χρησιμοποιήθηκε για την εκπαίδευση των υποψήφιων μοντέλων μηχανικής μάθησης.

Χαρακτηριστικά	**Τιμή**
Περιβάλλον υλοποίησης	Ganache GUI v2.7
RPC Endpoint	http://127.0.0.1:7545
Solidity Compiler	Solc 0.8.20
Smart Contract	MedicalOracle

Πίνακας 6.4: Πίνακας Παραμέτρων που χρησιμοποιήθηκαν κατά την εκπαίδευση των μοντέλων μηχανικής μάθησης.

Παράμετρος	**Τιμή**
Random Seed	42

Τμηματοποίηση	GroupShuffleSplit
Υπερδειγματοληψία	SMOTE
Σπόρος πρόωρης διακοπής	42
Μοντέλα Μηχανικής Μάθησης	Logistic Regression (LR), Random Forest (RF), XGBoost, LightGBM

Ο τελικός πίνακας αξιολόγησης του συνόλου δεδομένων προέκυψε από την εκτέλεση του pipeline της μηχανικής μάθησης στην ενότητα 5.1 περιέχει 546.028 νοσηλείες, στις οποίες οι 231.770 είναι μοναδικοί ασθενείς, με συχνότητα θετικής ετικέτας καρδιακής ανεπάρκειας 12.14%, δηλαδή 66.270 νοσηλείες. Το σύνολο δεδομένων δια μερίστηκε σε τρία σύνολα μέσω της GroupShuffleSplit χωρίς να επικαλύπτεται το ένα σύνολο με το άλλο. Τα μεγέθη παρουσιάζονται στον **πίνακα 6.5** ενώ η αναλογία θετικών περιπτώσεων παραμένει εντός του εύρους 12.09%-12.26% και αποτελεί δείκτη της επαρκούς μεγέθους δείγματος που εξασφαλίζει αντιπροσωπευτική κατανομή ακόμη και με απλή τυχαία διαμέριση.

Πίνακας 6.5: Τελικός πίνακας παραμέτρων που χρησιμοποιήθηκαν κατά την εκπαίδευση των μοντέλων μηχανικής μάθησης

Σύνολο	**Νοσηλείες**	**Ασθενείς**	**Ποσοστό ΚΑ**	**Χρήση**
Εκπαίδευσης	326.741	138.700	12.09%	Προσαρμογή Παραμέτρων μοντέλων
Βαθμονόμησης	110.003	46.700	12.13%	Επιλογή Μοντέλου Πρωταγωνιστή & Βαθμονόμηση Πιθανοτήτων
Ελέγχου	109.284	46.400	12.26%	Τελική Αξιολόγηση
Σύνολικο Σύνολο Δεδομένων	546.028	231.770	12.14%	-

Η προγνωστική ικανότητα του μοντέλου αξιολογείται από πέντε συμπληρωματικές μετρικές, που αναπτύσσονται θεωρητικά στην ενότητα 3.2 και καθεμία αποτυπώνει διαφορετική διάσταση της απόδοσης:

- AUROC (Area Under the Receiver Operating Characteristic Curve):
- AUPRC (Area Under the Precision-Recall Curve):
- Brier Score
- F1 Score
- Accuracy

6.2 Αξιολόγηση των Τεσσάρων Υποψήφιων Μοντέλων

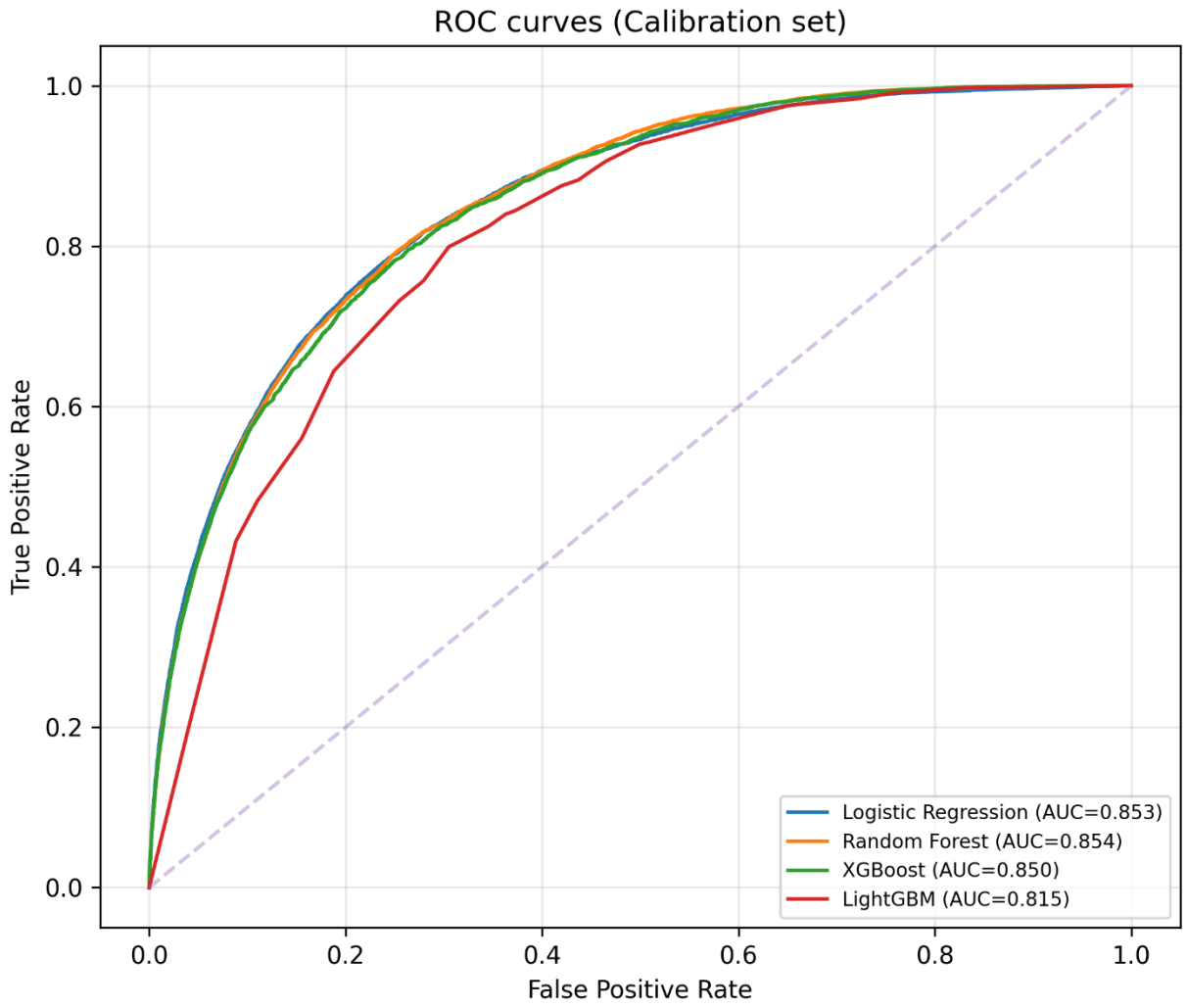
Στην ενότητα αυτή αξιολογούνται τα τέσσερα υποψήφια μοντέλα επί του συνόλου βαθμονόμησης, για την τελική επιλογή του μοντέλου πρωταθλητή. Όπως αναφέρθηκε και στην ενότητα 5.2.5, η σύγκριση των αποτελεσμάτων εκπαίδευσης έγινε αυστηρά στο σύνολο βαθμονόμησης για την αποφυγή της διαρροής δεδομένων και παραγωγής ψευδών μελλοντικών αποτελεσμάτων. Η σύγκριση των μοντέλων πραγματοποιήθηκε υπό κανονικοποιημένες συνθήκες, σταθερά ορισμένες και χωρίς την εφαρμογή υπερδειγματοληψίας. Με βάση τον **Πίνακα 6.6**, η συγκριτική απόδοση των τεσσάρων μοντέλων στο σύνολο βαθμονόμησης αναδεικνύει ως μοντέλο πρωταθλητή το Τυχαίο Δάσος με τιμή AUROC 0.8542, ενώ σε ελάχιστη απόκλιση βρίσκεται η Λογιστική Παλινδρόμηση με AUROC 0.8529. Η διαφορά των μοντέλων είναι πραγματικά οριακή, μόλις 0.0013, καθιστώντας τη μεταξύ τους σχέση κλίμακας ισοτιμίας. Η αναλυτική παρουσίαση της συμπεριφοράς των μοντέλων μηχανικής μάθησης αποτυπώνεται οπτικά στην **Εικόνα 58** που απεικονίζονται οι καμπύλες ROC των υποψήφιων μοντέλων. Αξίζει να σχολιαστεί η σχεδόν πλήρης ταύτιση που εμφανίζουν οι καμπύλες των αλγορίθμων Λογιστικής Παλινδρόμησης και του Τυχαίου Δάσους, ενώ η καμπύλη του LightGBM εμφανώς υστερεί.

Πίνακας 6.6: Πίνακας απόδοσης των τεσσάρων μοντέλων επί του συνόλου βαθμονόμησης

Μοντέλο	AUROC	AUPRC	Brier Score	Ακρίβεια	F1 Score
Τυχαίο Δάσος	0.8542	0.5360	0.1241	0.8333	0.5296
Λογιστική Παλινδρόμηση	0.8529	0.5396	0.1564	0.7740	0.5002
XGBoost	0.8505	0.5302	0.2890	0.5469	0.3821
LightGBM	0.8151	0.3820	0.2504	0.4456	0.3421

Παρά την καλή απόδοση των μετρικών AUROC και AUPRC, το Brier σκορ του εμφανίζεται πολύ υψηλό, σχεδόν διπλάσιο με αυτό της Λογιστικής Παλινδρόμησης. Αυτό οφείλεται στην υπερβολική σιγουριά των μοντέλων ενισχυμένης κλίσης. Για τον λόγο αυτό η βαθμονόμηση κρίνεται απαραίτητη και όχι προαιρετική, όπως περιγράφεται στην ενότητα 5.2.6. Στον **Πίνακα 6.6** παρατηρείται και μια φανερή απόκλιση του LightGBM στην μετρική AUROC από τα υπόλοιπα μοντέλα, κάτι που δικαιολογεί την μη επιλογή του ως μοντέλο πρωταθλητή. Συνολικά, με τις υπάρχουσες συνθήκες και τα αποτελέσματα των μετρικών, το Τυχαίο Δάσος παίρνει την θέση του πρωταθλητή με οριακή υπεροχή έναντι της Λογιστικής Παλινδρόμησης, διαφορά των 0.0013 μονάδων της μετρικής AUROC. Η μικρή του διαφορά δείχνει ότι τα δύο μοντέλα έχουν ισοδύναμη διακριτική ικανότητα, με την ειδοποιός διαφορά τους να είναι η σύλληψη μη-γραμμικών αλληλεπιδράσεων από το Τυχαίο Δάσος μεταξύ των σύνθετων βιοδεικτών που υπολογίστηκαν στην ενότητα 5.1.3, ενώ ταυτόχρονα διατηρεί πολύ χαμηλό Brier σκορ, 0.1241 έναντι 0.1564 της Λογιστικής Παλινδρόμησης. Η χαμηλή τιμή Brier σκορ δείχνει τη καλύτερη βαθμονόμηση των πιθανοτήτων υπό τον αλγόριθμο Τυχαίου Δάσους. Το γεγονός πως ένα γραμμικό μοντέλο όπως η Λογιστική Παλινδρόμηση φτάνει σε τόσο κοντινή απόδοση

επιβεβαιώνει πως οι σύνθετοι βιοδείκτες έχουν ενσωματώσει σημαντικό μέρος της μη-γραμμικότητας του προβλήματος στα δεδομένα εισόδου.



Εικόνα 58 : Καμπύλες ROC των τεσσάρων υποψήφιων μοντέλων επί του συνόλου βαθμονόμησης στη 1η δοκιμή.

6.3 Επίδραση του Optuna

Στην ενότητα αυτή αναλύονται τα αποτελέσματα της συμβολής του μηχανισμού Μπευζιανής βελτιστοποίησης υπερπαραμέτρων Optuna στην συνολική αρχιτεκτονική του συστήματος και τη ροή της μηχανικής μάθησης. Όπως αναπτύχθηκε και στην ενότητα 5.2.4, η βιβλιοθήκη Optuna στηρίζεται στον Δενδροειδή Εκτιμητή Parzen (TPE) που επιτρέπει την εξαντλητική και πλήρως καθοδηγούμενη αναζήτηση σε πολυδιάστατούς χώρους.

Η εφαρμογή του υπάρχοντος μηχανισμού αξιοποιείται αποκλειστικά από τα δύο πιο σύνθετα μοντέλα που χρησιμοποιούνται, το XGBoost και το LightGBM. Τα μοντέλα από τη φύση τους

διαθέτουν δεκάδες υπερπαραμέτρους που μπορούν να υπόκεινται σε τροποποίηση, όπως ο αριθμός δέντρων, ο ρυθμός μάθησης, το μέγιστο βάθος, μερικοί παράμετροι κανονικοποίησης και η στρατηγική δειγματοληψίας. Η ποικιλία υπερπαραμέτρων των δύο αυτών τεχνικών προσδίδει την ανάγκη σταδιακού πειραματισμού απέναντι στη Λογιστική Παλινδρόμηση που έχει ελάχιστες ρυθμίσεις και το Τυχαίο Δάσος που έχει προεπιλεγμένες τιμές της βιβλιοθήκης scikit-learn.

Για την αποφυγή υπερπροσαρμογής ή υπό προσαρμογής που μπορεί να υποστούν τα μοντέλα πραγματοποιείται συστηματική διερεύνηση της επίδρασης διαφορετικών παραμετροποιήσεων του Optuna στη τελική απόδοση των μοντέλων. Συγκεκριμένα πραγματοποιήθηκαν 7 ανεξάρτητες πειραματικές δοκιμές. Καθεμία διαφοροποιήθηκε σε 4 διαστάσεις: τη μετρική βελτιστοποίησης (AUROC, AUPRC, Brier σκορ), το σύνολο των δοκιμών, τη μέθοδο διασταυρωμένης επικύρωσης και τη προαιρετική ενεργοποίηση ή απενεργοποίηση της υπερδειγματοληψίας SMOTE κατά τη φάση τροποποίησης. Στον παρακάτω πίνακα απεικονίζονται οι πειραματικές δοκιμές που συνέβησαν στο Optuna.

Πίνακας 6.7: Πίνακας των πειραματικών δοκιμών του μηχανισμού Optuna, με τις μετρικές βελτιστοποίησης, τον αριθμό δοκιμών, το σχήμα cross-validation, το SMOTE και τα μοντέλα.

Δοκιμή	Μετρική Βελτιστοποίησης	Δοκιμές	CV	SMOTE	Μοντέλα υπό Optuna
T1	AUPRC	50	Απλό (default)	Ναι	XGB, LGBM
T2	AUPRC	100	Απλό	Ναι	XGB, LGBM
T3	AUPRC	50	5-fold	Ναι	XGB, LGBM
T4	AUROC	50	Απλό	Ναι	XGB, LGBM
T5	Brier	50	Απλό	Ναι	XGB, LGBM
T6	AUPRC	50	Απλό	Οχι	XGB, LGBM
T7	AUPRC	50	Απλό	Ναι	LR,RF,XGB,LGBM

Η δοκιμή T1 αποτελεί τη βασική αναφορά με 50 δοκιμές, AUPRC ως μετρική βελτιστοποίησης, απλό σχήμα διασταυρωμένης επικύρωσης με ενεργοποιημένη την υπερδειγματοληψία SMOTE. Η πρώτη δοκιμή εφαρμόζεται σε δύο μόνο μοντέλα. Στην T2 δοκιμή εφαρμόζεται διπλασιασμός του αριθμού δοκιμών για αξιολόγηση σύγκλισης του αλγορίθμου Δενδροειδούς Εκτίμησης Parzen. Η δοκιμή T3 κρατάει τον αριθμό των δοκιμών στις 50 αλλά εφαρμόζει αυστηρότερη αξιολόγηση μέσω διασταυρούμενης επικύρωσης 5 πτυχών για την αξιολόγηση της ακρίβειας των μοντέλων στα δεδομένα μας. Η T4 περιλαμβάνει τα ίδια ορίσματα, με την μόνη διαφορά ότι η μετρική βελτιστοποίησης είναι η AUROC με σκοπό τη διάκριση αντί ισορροπίας πρόβλεψης και ανάκλησης. Η δοκιμή T5 χρησιμοποιεί την εναλλακτική μετρική βελτιστοποίησης Brier σκορ για τη βελτιστοποίηση της ποιότητας των πιθανοτήτων και ενώ η T6 απενεργοποιεί την υπερδειγματοληψία και βασίζεται στη μετρική βελτιστοποίησης AUPRC με σκοπό τον έλεγχο συμβολής του SMOTE στη βελτίωση των αποτελεσμάτων. Τέλος, η τελευταία δοκιμή T9 εφαρμόζει τα ίδια ορίσματα με την T1 δοκιμή με τη διαφορά ότι εκτελείται πλήρως και στα τέσσερα μοντέλα για την επιλογή τελικού μοντέλου πρωταθλητή.

Πίνακας 6.8: Πίνακας των βέλτιστων υπερπαραμέτρων του αλγορίθμου XGBoost για κάθε πειραματική δοκιμή,

Παράμετρος	T1	T2	T3	T4	T5	T6
Αριθμός Δέντρων	2.274	2.657	2.053	823	1.116	2.999
Ρυθμός Μάθησης	0,0185	0,0132	0,0056	0,0136	0,1896	0,0122
Μέγιστο Βάθος	7	10	6	7	10	8
Ελάχιστο Βάρος Φύλλου	0,341	0,128	0,838	4,761	0,275	0,113
Ποσοστό Δειγμάτων	0,7180	0,6202	0,5594	0,5735	0,5959	0,5368
Ποσοστό Χαρακτηριστικών	0,5044	0,5014	0,5869	0,5052	0,5298	0,7639
Ελάχιστη Μείωση Απώλειας	0,090	1,772	1,106	0	0	2,354
Κανονικοποίηση L1	7,911	4,934	1,632	0,001	0,001	0
Κανονικοποίηση L2	6,588	2,747	14,490	2,489	0,107	77,733
Μέγιστο Βήμα Δέλτα	10	6	8	5	4	0

Η ανάγνωση του πίνακα των βέλτιστων υπερπαραμέτρων XGBoost ανά πειραματική δοκιμή του Optuna αποκαλύπτει τρείς σημαντικές παρατηρήσεις. Πρώτον ο αριθμός των δέντρων κυμαίνεται από 823 έως 2999 σημαντική διαφορά τάξη μεγέθους που αντικατοπτρίζει την αντιστάθμιση μεταξύ μάθησης και κανονικοποίησης. Όταν η βελτιστοποιούμενη μετρική απαιτεί επιθετική προσαρμογή στα δεδομένα της δοκιμής T6, το Optuna παράγει πολλά δέντρα ενώ η μετρική σύγκλισης σταματάει νωρίς την παραγωγή των δέντρων στα 800 περίπου. Επίσης, ο ρυθμός μάθησης ξεχωρίζει στη δοκιμή T5 που εφαρμόζεται η πιο επιθετική στρατηγική σύγκλισης (0.1896) ενώ στις υπόλοιπες δοκιμές ο ρυθμός μάθησης παραμένει χαμηλός στο εύρος [0,005-0,019]. Τρίτον και σημαντικότερο, η L2 Κανονικοποίηση παρουσιάζει ακραίες μεταβολές, από 0,11 έως 77,73, σε σχέση με τη Κανονικοποίηση L1, 0 έως 7,911, υποδηλώνοντας πως το Optuna εκμεταλλεύτηκε πραγματικά την ελευθερία του χώρου αναζήτησης για να βρει διαφορετικές λύσεις ανάλογα τις συνθήκες.

Πίνακας 6.9: Πίνακας των βέλτιστων υπερπαραμέτρων του αλγορίθμου LightGBM για κάθε πειραματική δοκιμή.

Παράμετρος	T1	T2	T3	T4	T5	T6
Αριθμός Δέντρων	1.872	3.029	2.730	3.807	4.829	1.910
Ρυθμός Μάθησης	0,1118	0,1460	0,1377	0,1349	0,0615	0,0095
Μέγιστο Βάθος	11	12	9	7	12	9
Αριθμός Φύλλων	231	215	206	211	172	238
Ελάχιστα Δείγματα Φύλλου	26	127	91	6	6	105
Ποσοστό Δειγμάτων	0,8553	0,7047	0,8861	0,8897	0,5021	0,5283
Ποσοστό Χαρακτηριστικών	0,6466	0,7133	0,8577	0,7953	0,8826	0,5
Κανονικοποίηση L1	0,258	0,093	0,364	1,598	0,834	0,001
Κανονικοποίηση L2	86,18	61,676	46,591	0,052	18,592	0,229

Σε αντίθεση με το προφίλ συμπεριφοράς του αλγορίθμου XGBoost, η εικόνα του LightGBM προσδίδει μια διαφορετική οπτική. Αρχικά, ο αριθμός των δέντρων είναι σταθερά υψηλότερος από τον αλγόριθμο XGBoost με εύρος [1.872-4.829] έναντι [823-2.999] του τελευταίου. Η διαφορά αυτή στη ποσότητα των δέντρων οφείλεται στο γεγονός πως ο LightGBM στηρίζεται στην στρατηγική ανάπτυξης ανά φύλλο (leaf-wise) που δημιουργεί μεγαλύτερο πλήθος ασθενέστερων δέντρων σε αντίθεση με τη στρατηγική ανάπτυξης ανά επίπεδο του XGBoost που είναι οριζόντιας ανάπτυξης. Επίσης, ο αριθμός φύλλων παραμένει σχετικά σταθερός στο εύρος [172 - 238], φανερώνοντας πως ανεξάρτητα από τις άλλες υπερπαραμέτρους το Optuna συγκλίνει σε παρόμοιο επίπεδο πολυπλοκότητας. Ακόμη, η διαφορετική στρατηγική του αλγορίθμου προσδίδει μια έντονη μεταβλητότητα στον αριθμό ελάχιστων δειγμάτων φύλλου με εύρος [6 - 127], δείχνοντας πως ο μηχανισμός κανονικοποίησης μέσω ελάχιστου μεγέθους φύλλων αξιοποιείται διαφορετικά σε κάθε δοκιμή για να αντιμετωπιστούν τα ανισόρροπα δεδομένα

Στις προσπάθειες που πραγματοποιήθηκαν υπήρξε σημαντική επένδυση υπολογιστικών πόρων, παρόλα αυτά τα τελικά αποτελέσματα έδειξαν, όπως παρουσιάζεται στην ενότητα 6.4, το Τυχαίο Δάσος ως μοντέλο «Πρωταθλητή» με τις προεπιλεγμένες παραμέτρους που είχε από τη βιβλιοθήκη scikit-learn. Η κατάσταση αυτή παρουσιάζεται απροσδόκητα παράδοξη στην αρχή, όμως φανερώνει δύο πραγματικότητες της μηχανικής μάθησης στα ιατρικά δεδομένα μεγάλης κλίμακας. Η πρώτη εικόνα αφορά την ανθεκτικότητα που δείχνει το Τυχαίο Δάσος λόγω της συλλογικής φύσης του να ξαναλύνει ένα πρόβλημα φτιάχνοντας εκατοντάδες δέντρα παρόμοια, με το κάθε δέντρο να εκπαιδεύεται σε ένα ελαφρώς διαφορετικό τυχαίο κομμάτι των δεδομένων. Είναι σαν να πραγματοποιεί μια κανονικοποίηση παραμέτρων από μόνο του κάθε φορά που συχνά καθιστά την επιθετική ρύθμιση περιττή. Η δεύτερη εικόνα αφορά το μικρό περιθώριο βελτίωσης που ελαττώθηκε εξαιτίας της εκτενής φάσης προ επεξεργασίας χαρακτηριστικών που έγινε στην ενότητα 5.1.3 και η κατασκευή μη-γραμμικής πληροφορίας μέσω της κατασκευής σύνθετων κλινικών δεικτών (eGFR, RAR, BCR, ALBI, HRR, ACAG) αφήνοντας μικρότερο περιθώριο βελτίωσης μέσω αλγοριθμικής πολυπλοκότητας.

Η συνολική πειραματική διαδικασία δεν ακυρώνεται από τα τελικά αποτελέσματα της. Χωρίς τη βελτιστοποίηση επτά ανεξάρτητων δοκιμών, δε θα μπορούσε να υποθεί με αποδείξεις ότι το Τυχαίο Δάσος αποτελεί τη βέλτιστη επιλογή. Θα επικρατούσε η συνηθισμένη άποψη ότι τα ενισχυμένα μοντέλα είναι ανώτερα των κλασσικών συλλογικών και γραμμικών μοντέλων. Με τις παρούσες εμπειρικές δοκιμές αποδείχθηκε στη τελευταία δοκιμή που εκπαιδεύτηκαν όλα τα μοντέλα, ο νικητής Τυχαίο Δάσος επικράτησε υπό συγκεκριμένες συνθήκες και ενισχύει την επιστημονική εγκυρότητα της ενότητας 6.4.

6.4 Επίδοση του Μοντέλου Πρωταθλητή

6.4.1 Μετρικές αποτελεσμάτων του συνόλου ελέγχου

Με την ολοκλήρωση της φάσης επιλογής, η μέθοδος Optuna ανέδειξε πως η υπερπαραμετροποίηση των παραμέτρων του μοντέλου κρίθηκε απαραίτητη για την αληθοφάνεια των πιθανοτήτων κίνδυνου σε σχέση με τις αρχικές σταθερές παραμέτρους που εμφάνιζε η 1η δοκιμή. Το μοντέλο Τυχαίου Δάσους ανακηρύσσεται μοντέλο «Πρωταθλητής» και με τη βελτιστοποίηση και βαθμονόμηση του, το σύνολο ελέγχου που ήταν κλειδωμένο ενεργοποιείται και συμβάλλει για πρώτη και μοναδική φορά στη παρούσα ενότητα. Με την ενεργοποίηση του παράγονται αμερόληπτες εκτιμήσεις της γενικευσιμότητας χρήσης του κλινικού συστήματος σε πραγματικές συνθήκες. Άλλωστε αυτός ήταν και ο λόγος που στην ενότητα 5.1.5 αναφέρθηκε πως το σύνολο ελέγχου των 109.284 νοσηλειών παρέμεινε πλήρως απομονωμένο από όλες τις προηγούμενες αποφάσεις σχεδιασμού, ώστε οι μετρικές που παρουσιάζονται παρακάτω να αντικατοπτρίζουν τη συμπεριφορά του συστήματος σε νέους ασθενείς, χωρίς διαρροή πληροφορίας από οποιοδήποτε ενδιάμεσο στάδιο. Στο **Πίνακα 6.8** παρουσιάζονται οι πέντε μετρικές του βαθμονομημένου μοντέλου πρωταθλητή.

Πίνακας 6.10: Πίνακας σύγκρισης των μετρικών του μοντέλου Πρωταθλητή σε σύνολο βαθμονόμησης και ελέγχου.

Μετρική	Σύνολο Βαθμονόμησης (Calibration Set)	Σύνολο Ελέγχου (Test Set)	Διαφορά
AUROC	0.8542	0.8541	-0.0001
AUPRC	0.5360	0.5378	+0.0018
Brier	0.1241	0.0941	-0.0300
Ακρίβεια	0.8333	0.8701	+0.0368
F1	0.5296	0.4441	-0.0855

Εν πρώτοις αξιοσημείωτη είναι η πολύ μικρή διαφορά μεταξύ των δύο συνόλων προς τη μετρική AUROC μόλις σε 0.0001 μονάδες. Η σχεδόν ταύτιση των τιμών μας, 0.8542 έναντι 0.8541, επιβεβαιώνει την ορθή επιστημονική εφαρμογή της μεθοδολογίας τμηματοποίησης GroupShuffleSplit που πραγματοποιήθηκε σε επίπεδο ασθενούς στην ενότητα 5.1.5 και την εξαιρετική γενικευσιμότητα που δείχνει το μοντέλο όταν δέχεται στην είσοδο του νέους ασθενείς. Συνεχίζοντας στον πίνακα, η 2η μετρική AUPRC παρουσιάζει ελαφρά βελτίωση, σχεδόν μηδαμινή στο σύνολο ελέγχου, από 0.5378 έναντι 0.5360. Γνωρίζοντας ότι η θετική κλάση κινδύνου ΚΑ στο γενικό πληθυσμιακό σύνολο είναι 12.14%, η τιμή AUPRC του σύνολό ελέγχου που είναι 0.5378 ισοδυναμεί με βελτίωση τάξης μεγέθους λίγο πιο πάνω από 4.4 φορές σε σχέση με τη τυχαία πρόβλεψη, ενισχύοντας την βαρύτητα της πρόγνωσης που εξάγει το μοντέλο σε ασθενείς που τείνουν ή πάσχουν από ΚΑ. Η παρατήρηση αυτή παρουσιάζεται ως ύψιστης σημασίας για την αναγνώριση της κλινικής σημασίας της μειοψηφικής κλάσης. Πραγματοποιώντας μια αναδρομική ματιά στην ενότητα 3.2.2 όπου εξηγείται μαθηματικά η μετρική Brier

Ένα από τα παράδοξα που συνέβη στην φάση εκπαίδευσης του μοντέλου Τυχαίου Δάσους κατά τη βαθμονόμηση του από το σύνολο ελέγχου είναι η μείωση της τιμής Brier σκορ, από 0.1241 έναντι 0.0941. Ανατρέχοντας στην ενότητα 3.2.2 όπου εξηγείται μαθηματικά η μετρική Brier σκορ

γίνεται αντιληπτό πως η μετρική είναι ουσιαστικά το Μέσο Τετραγωνικό Σφάλμα (MSE) των προβλεπόμενων πιθανοτήτων. Η βελτίωση αυτή προκαλείται από την αλλαγή της κατανομής των δεδομένων μεταξύ των δύο συνόλων και όχι επειδή το μοντέλο έγινε «πιο έξυπνο». Αρχικά παρουσιάζεται αλλαγή στη συχνότητα των ασθενών με ΚΑ στο 12.13% του συνόλου βαθμονόμησης και στη συχνότητα των ασθενών με ΚΑ στο σύνολο ελέγχου που είναι μεγαλύτερη στο 12.26%.. Επιπλέον, επί τη βαθμονόμηση του μοντέλου Τυχαίου Δάσους από το σύνολο ελέγχου, οι προβλέψεις που παρήχθησαν στην έξοδο του μοντέλου είχανε μειωμένη διασπορά μεταξύ τους, δηλαδή ήταν περισσότερο συγκεντρωμένες σε συγκεκριμένα σημεία τιμών. Με βάση το μαθηματικό τύπο της Brier σκορ, όσο οι προβλέψεις εμφανίζουν μικρή διασπορά και είναι πιο κοντά στο γενικό μέσο όρο τιμών του συνόλου, το τετραγωνικό σφάλμα μειώνεται για τη συντριπτική πλειοψηφία των ασθενών που αποτελούν το 87% περίπου του συνόλου της αρνητικής κλάσης. Με βάση τη βιβλιογραφία, το μοντέλο θεωρείται τέλεια βαθμονομημένο όταν το Brier σκορ ισούται με την αναμενόμενη διακύμανση της κλάσης. Παίρνοντας τον τύπο, λοιπόν, από την ενότητα 3.2.3 της διακύμανσης παρουσιάζεται ότι το τελικό Brier σκορ του 0.0941 που πέτυχε το μοντέλο είναι μικρότερο ακόμα και από το θεωρητικό όριο των 0.107, αποδεικνύοντας την πραγματική και βαθιά προγνωστική ικανότητα του μοντέλου πάνω σε κάθε ασθενή ξεχωριστά, μειώνοντας την αβεβαιότητα κάτω από το επίπεδο του εγγενούς σφάλματος.

Το δεύτερο παράδοξο που συνέβη κατά την εκπαίδευση του μοντέλου πρωταθλητή είναι η φαινομενική ψαλίδας μεταξύ των μετρικών της Ακρίβειας και του F1 σκορ. Αναλυτικότερα, στην εκπαίδευση επί του συνόλου βαθμονόμησης το μοντέλο παρουσίασε ακρίβεια της κλίμακας των 83.33% και F1 σκορ 0.5296, δηλαδή των 52.96%, ενώ στην εκπαίδευση επί του συνόλου ελέγχου παρατηρήθηκε αύξηση κατά 3.68% με την ακρίβεια να φτάνει το 87.01% και το F1 σκορ να ελαττώνεται ακόμη περισσότερο στο 0.4441, εμφανίζοντας πτώση 8.55%.

Ένας αναγνώστης που διαβάζει πρώτη φορά την αυτούσιά διπλωματική εργασία θα αναρωτιόταν πως γίνεται το μοντέλο να γίνεται συνολικά ακριβέστερο αλλά η ικανότητα του να προβλέπει σωστά χειροτέρεψε; Προφανώς η απάντηση στο ερώτημα του θα βρισκόταν στον μαθηματικό τύπο της ακρίβειας στην ενότητα 3.2. Στα προηγούμενα κεφάλαια αναφέρθηκε η ύπαρξη έντονης ανισορροπίας στα δεδομένα, με τους ασθενείς που είναι στη θετική κλάση καρδιακής ανεπάρκειας του συνόλου ελέγχου να αντιστοιχούν στο 12.26% του και το υπόλοιπο 87.7% ανήκει στους ασθενείς που ανήκουν στην αρνητική κλάση καρδιακής ανεπάρκειας. Εάν το μοντέλο μας κοιτούσε κατά τη πλειοψηφία των υγειών νοσηλειών, θα γινόταν πιο συντηρητικό και θα προέβλεπε με μεγαλύτερη ευκολία ότι κάποιος είναι υγιής, εκτοξεύοντας την ακρίβεια του. Αυτό όμως είναι λάθος. Για την εξισορρόπηση της πλασματικής ακρίβειας συμβάλλει το F1 σκορ. Το ίδιο δεν ενδιαφέρεται για την πρόβλεψη της αρνητικής κλάσης ασθενών αλλά εστιάζει πλήρως στη θετική κλάση, με άλλα λόγια σε αυτούς που πάσχουν από καρδιακή ανεπάρκεια. Στο μαθηματικό ορισμό του αποκαλύπτονται οι δύο έννοιες στις οποίες στηρίζεται η επιστημονική θεμελίωση του, η ακρίβεια και η ανάκληση. Με την ακρίβεια του F1 σκορ να επηρεάζεται από τη διαδικασία της βαθμονόμησης που το όριο απόφασης που έχει τεθεί στο 0.5, δηλαδή το μοντέλο επιλέγει ως ασθενείς κινδύνου εμφάνισης καρδιακής ανεπάρκειας αυτούς με πιθανότητα πάνω

από 50%. Έτσι το μοντέλο έχασε μερικούς πραγματικούς ασθενείς, αφού ελαττώθηκε η ανάκληση του, με την ελάττωση να συνεχίζει στο F1 σκορ φτάνοντας στη τιμή των 0.4441. Επομένως η συμπεριφορά των δύο μετρικών παρουσιάζεται λογικός συστημική και προβλέψιμη, κάτι το οποίο αποδεικνύεται και αναλύεται περισσότερο στην επόμενη ενότητα που χρησιμοποιείται η μέθοδος Bootstrap.

6.4.2 Επαλήθευση Μετρικών μέσω της μεθόδου Bootstrap CI

Για να εξασφαλισθεί η στατιστική ευστάθεια και γενικευσιμότητα του μοντέλου πρόβλεψης που βγήκε πρωταθλητής στο παρών σύστημα υποστήριξης κλινικών αποφάσεων, η αξιολόγηση του μοντέλου δεν πρέπει να περιοριστεί σε απλές σημειακές τιμές εκτίμησης των μετρικών απόδοσης του συνόλου ελέγχου. Η μη-παραμετρική τεχνική Bootstrap με επανατοποθέτηση αποτελεί μια ιδανική λύση, καθώς η μη-παραμετρική φύση της σε συνδυασμό με τη πολυπλοκότητα και την εγγενούς ανισορροπίας των ιατρικών δεδομένων της βάσης MIMIC-IV κρίνει απαραίτητο τον προσδιορισμό αυστηρών Διαστημάτων Εμπιστοσύνης. Η τεχνική υλοποιείται σε επίπεδο ασθενούς με N=2000 επαναλήψεις.

Αναλυτικότερα, από το αρχικό σύνολο ελέγχου, που αποτελεί το 20% (109.284 νοσηλείες) του συνολικού πληθυσμού δεδομένων της μελέτης, παράγονται 2000 ανεξάρτητα υποδείγματα ίσου μεγέθους με τυχαία δειγματοληψία και επανατοποθετήθηκαν εκ νέου πέντε μετρικές απόδοσης στο καθένα. Οι νέες τιμές που προκύπτουν για κάθε μετρική ταξινομούνται σε κατανομές bootstrap στο όριο του 95% με τη μέθοδο των εκατοστημορίων, δηλαδή απομονώνοντας το κεντρικό 95% της κατανομής, στο 2.5ο κάτω εκατοστημόριο έως το 97.5ο άνω εκατοστημόριο.

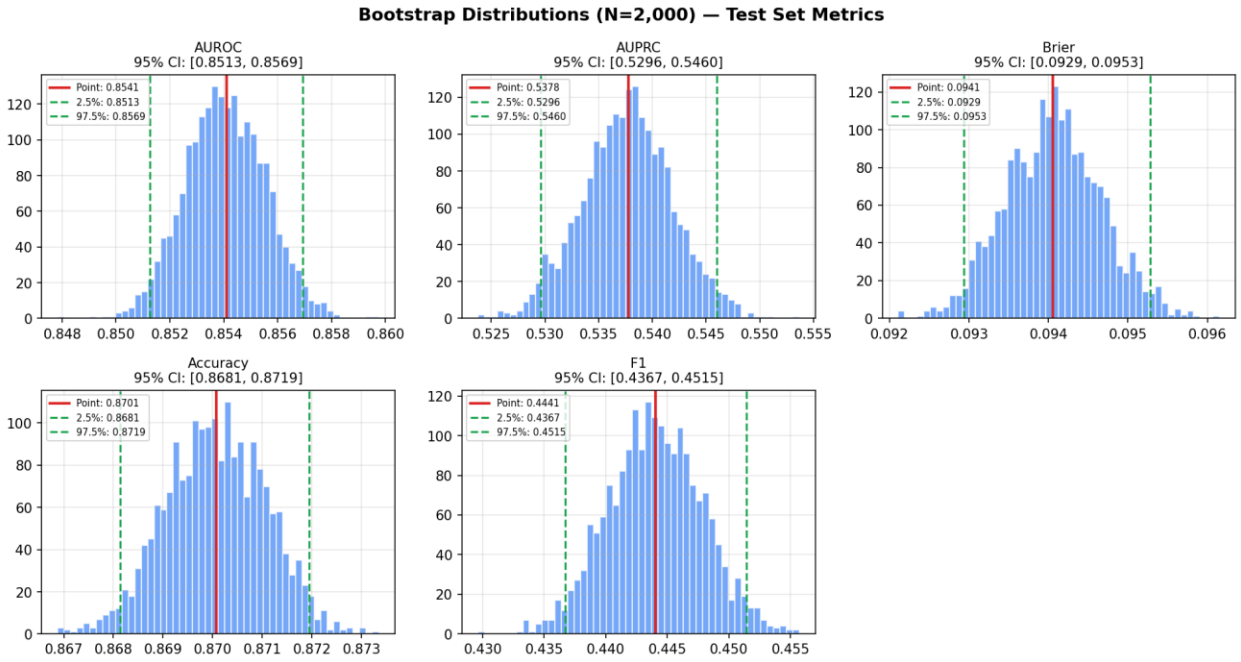
Πίνακας 6.11: Πίνακας Ανάλυσης των μετρικών του μοντέλου «Πρωταθλητή» επί του συνόλου ελέγχου με Bootstrap σε όριο 95%

Μετρική Αξιολόγησης	Σημειακή Εκτίμηση	Μέσος Όρος Bootstrap	Μέση Τιμή Bootstrap	Τυπική Απόκλιση Bootstrap (SD)	Κάτω Όριο (CI 2.5%)	Άνω Όριο (CI 97.5%)	Εύρος Διασποράς (CI Width)
AUROC	0.8541	0.8541	0.854	0.0015	0.8513	0.8569	0.0057
AUPRC	0.5378	0.5377	0.5378	0.0042	0.5296	0.546	0.0164
Brier	0.0941	0.0941	0.0941	0.0006	0.0929	0.0953	0.0023
Ακρίβεια	0.8701	0.9701	0.8701	0.001	0.8681	0.8719	0.0038
F1	0.4441	0.444	0.444	0.0038	0.4367	0.4515	0.0147

Στην ανάλυση των αποτελεσμάτων μέσω της τεχνικής Bootstrap στον **Πίνακα 6.11** αναδεικνύεται η υψηλή σταθερότητα και αξιοπιστία του κλινικού συστήματος, με τη στενότητα που εμφανίζουν τα διαστήματα εμπιστοσύνης με εύρος διασποράς μικρότερο του 0.02 (CI_Width<0.02) στις μετρικές AUROC, Brier και Ακρίβεια αποδεικνύεται πως η σταθερή απόδοση του μοντέλου δεν οφείλεται σε ευνοϊκό δείγμα ελέγχου. Ειδικότερα, η μετρική AUROC ανέρχεται σε 0.8541 με το διάστημα εμπιστοσύνης να κυμαίνεται σε εξαιρετικά στενό εύρος [95% CI: 0.8513, 0.8569] δείχνοντας την υψηλή διακριτική ικανότητα του μοντέλου να ξεχωρίζει τους ασθενείς υψηλού κινδύνου για ΚΑ, αφού και η χειρότερη πιθανή εκτίμηση υπερβαίνει το όριο των 0.85, όριο που

στη βιβλιογραφία της μοντελοποίησης προγνωστικών μοντέλων καρδιακής ανεπάρκειας θεωρείται κλινικά αξιοποιήσιμο. Ακόμη η ικανότητα του μοντέλου δεν επηρεάζεται από τυχόν στατιστικές διακυμάνσεις. Λαμβάνοντας, λοιπόν, υπόψη τη φύση του συνόλου, η κλάση στόχος παρουσιάζει μια ανισορροπία, την οποία η μετρική AUPRC που ισούται με 0.5378 στο εύρος [95% CI: 0.5296, 0.5460] επιβεβαιώνει πως το μοντέλο μπορεί να διατηρεί υψηλή ακρίβεια ακόμα και σε χαμηλά επίπεδα ανάκλησης, χωρίς αστάθεια. Ομοίως, η βαθμονόμηση του μοντέλου επικυρώνεται μέσω του Brier σκορ και καταγράφει υπερβολικά χαμηλή τιμή ίση με 0.0941 [95% CI: 0.0929, 0.0953]. Η καταγραφή αυτή επιβεβαιώνει την εξαιρετική βαθμονόμηση του τελικού μοντέλου, με την πιθανότητα κινδύνου που επιστρέφει ο αλγόριθμος να είναι ευθυγραμμισμένη με την πραγματική συχνότητα εμφάνισης της νόσου στο γενικό πληθυσμό. Τέλος, η υψηλή τιμή Ακρίβειας που εμφανίζεται, ίση με 0.8701 και το χαμηλό F1 σκορ ίσο με 0.4441, δείχνει ότι παρά την ανισορροπία των κλάσεων, με τις περισσότερες νοσηλείες να είναι υγιείς παρά ασθενείς που κινδυνεύουν ή πάσχουν από ΚΑ, οι μετρήσεις Bootstrap δεν είναι καθόλου τυχαίες. Μάλιστα οι συγκεκριμένες μετρήσεις σε τόσο στενό εύρος [95% CI:0.8681, 0.8719] και [95% CI:0.4367, 0.4515] δείχνουν πως οι πιθανότητες κινδύνου που εξάγει το σύστημα υποστήριξης κλινικών αποφάσεων «Diagnosis», ανταποκρίνονται σε πραγματικούς κλινικούς κινδύνους.

Οι κατανομές Bootstrap των μετρικών απεικονίζονται στην Εικόνα 59 όπου στο κάθε υπό διάγραμμα η κόκκινη κάθετη γραμμή σημειώνει τη σημειακή τιμή στο πλήρες σύνολο, οι μπλε στήλες δείχνουν την συχνότητα ιστογράμματος και οι κάθετες πράσινες διακεκομμένες γραμμές αποτελούν τα όρια των εκατοστημορίων 2.5% και 97.5% αντίστοιχα.

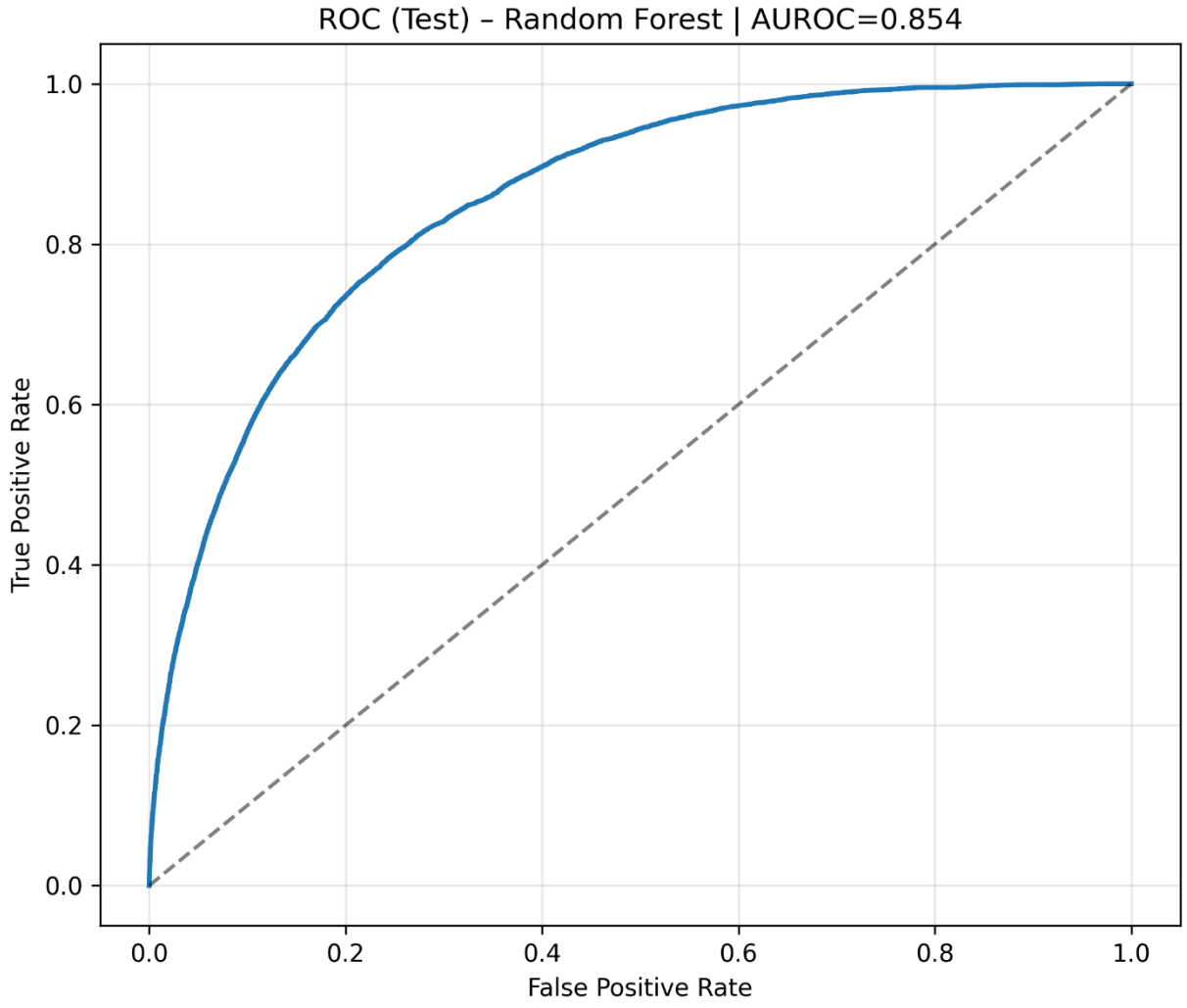


Εικόνα 59 : Οι Κατανομές Bootstrap των πέντε μετρικών απόδοσης του μοντέλου «Πρωταθλητή» επί του συνόλου ελέγχου.

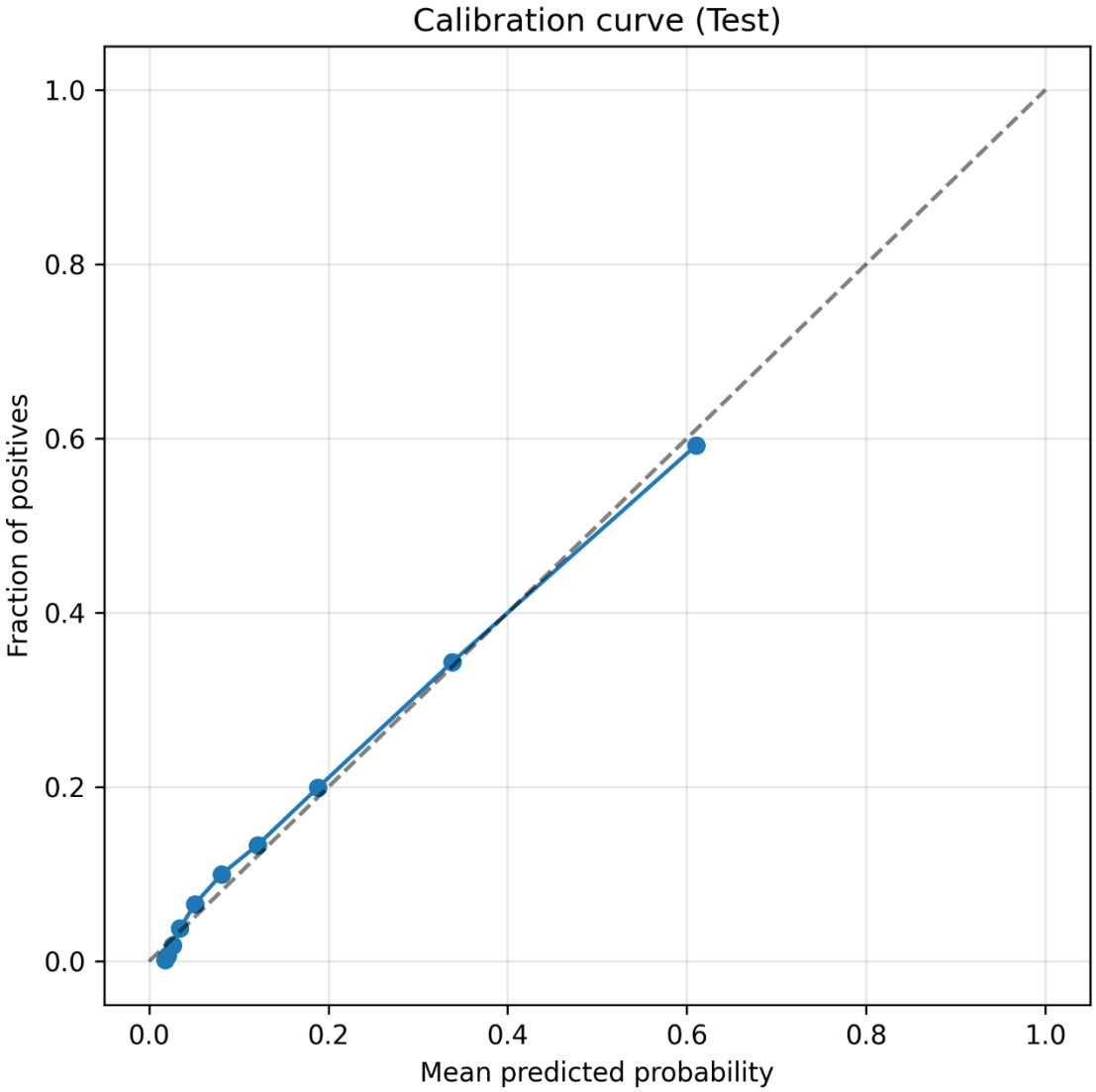
6.4.3 Διαγνωστικά Διαγράμματα Απόδοσης στο Σύνολο Ελέγχου

Το τελευταίο βήμα για την πλήρη αξιολόγηση του μοντέλου πρωταθλητή ολοκληρώνεται με τη προσθήκη τριών διαγραμμάτων διάγνωσης που αλληλοσυμπληρώνουν το ένα το άλλο. Το πρώτο διάγραμμα αφορά την καμπύλη ROC η οποία αξιολογεί τη διακριτική ικανότητα του μοντέλου, το δεύτερο διάγραμμα απεικονίζει τη καμπύλη βαθμονόμησης που αξιολογεί την αξιοπιστία των παραγόμενων πιθανοτήτων και η τρίτη αποτελεί την ανάλυση καμπύλης απόφασης (DCA) που αποτιμά τη κλινική χρησιμότητα του μοντέλου σε ρεαλιστικά σενάρια ανάγκης λήψης αποφάσεων.

Η κυρτότητα της καμπύλης πάνω από τη διαγώνιο της τυχαίας πρόβλεψης προς την πάνω αριστερή γωνία του διαγράμματος αποτελεί τη σχηματική απεικόνιση της υψηλής διακριτικής ικανότητα του μοντέλου. Στη κρίσιμη περιοχή χαμηλών ψευδώς θετικών (FPR<0.2), το μοντέλο επιτυγχάνει ποσοστό αληθών θετικών άνω του 0.70, δηλαδή εντοπίζει το 70% των ασθενών με ΚΑ αποδεχόμενο το 20% ψευδώς θετικών. Η αναλογία αυτή τοποθετεί το μοντέλο ευνοϊκό για κλινική χρήση διότι στα ψευδώς θετικά συνεπάγονται περιττές διαγνώσεις και διάφορες ψυχομετρικές μετρήσεις.

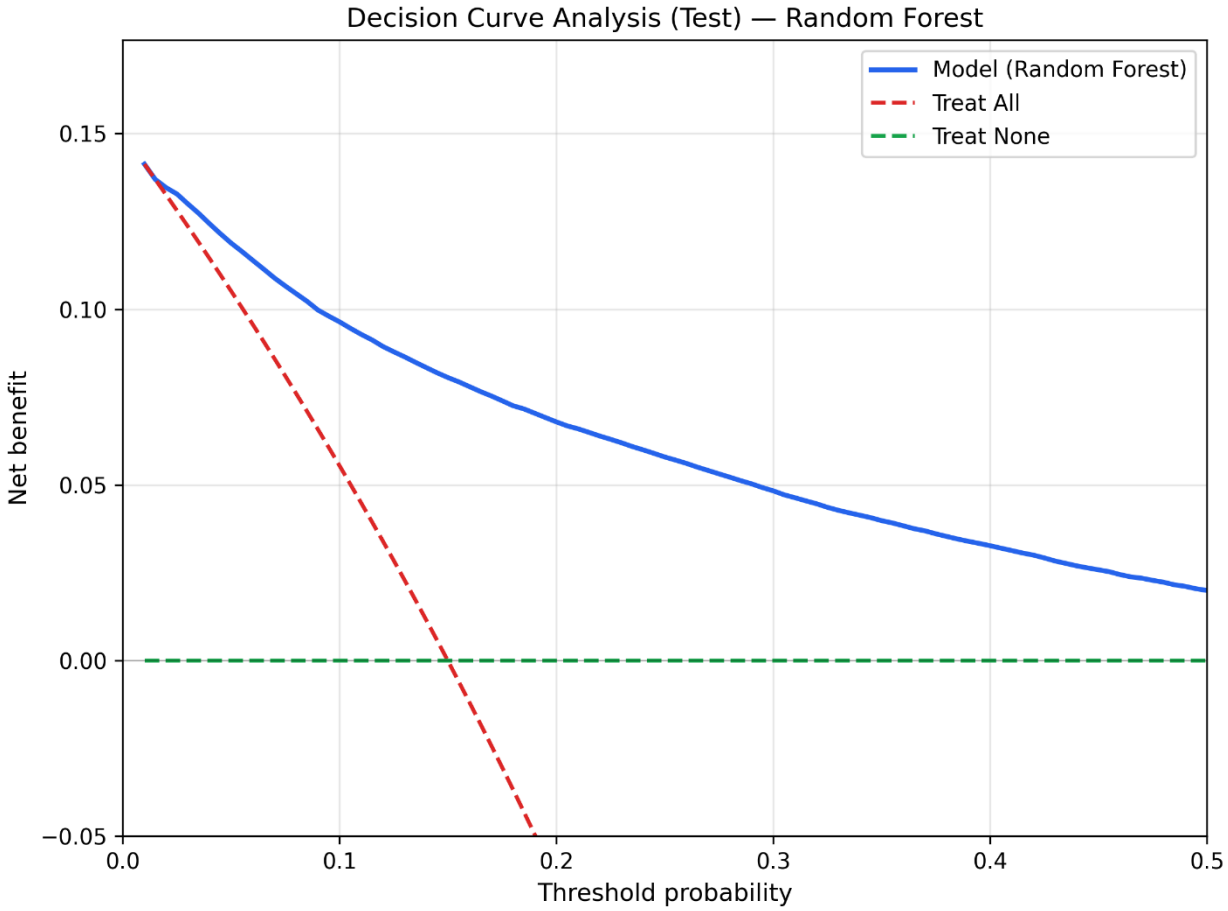


Εικόνα 60 : Καμπύλη ROC του μοντέλου «Πρωταθλητή» Τυχαίου Δάσους στο σύνολο ελέγχου. Το AUROC=0854 επιβεβαιώνει τη διακριτική ικανότητα του μοντέλου.



Εικόνα 61 : Η καμπύλη βαθμονόμησης πιθανοτήτων στο σύνολο ελέγχου υπό δεκαπλή κβαντοποίηση. Η σχεδόν τέλεια ευθυγράμμιση με τη διαγώνιο επιβεβαιώνει οπτικά τη χαμηλή τιμή Brier=0.094

Ένα βαθμονομημένο μοντέλο θεωρείται τέλειο όταν τα σημεία της καμπύλης βρίσκονται επί της διαγώνιου y=x, δηλαδή έστω η πιθανότητα κινδύνου που προβλέπει το μοντέλο να είναι ακριβώς ταυτόσημη με το πραγματικό ποσοστό θετικών περιπτώσεων. Στην αρχή η πλειονότητα των ασθενών παρουσιάζεται στο εύρος 0.02-0.12 με μηδενική πρακτικά απόκλιση των σημείων από τη διαγώνιο, ακόμη και στο υψηλό σημείο μέσης πρόβλεψης που είναι περίπου 0.60, η εμπειρική συχνότητα εμφάνισης ΚΑ παραμένει εντός 3 ποσοστιαίων μονάδων της προβλεπόμενης, κάτι που αιτιολογεί τη τιμή Brier σκορ που ισούται με 0.0941 όπως αναφέρθηκε στον Πίνακα [X] της ενότητας 6.4.1. Η συνολική καμπύλη βαθμονόμησης αποδεικνύει πως οι παραγόμενες πιθανότητες είναι εντελώς αξιοποιήσιμες σε πραγματικές εκτιμήσεις κινδύνου.



Εικόνα 62 : Ανάλυση Καμπυλών Απόφασης (DCA) του μοντέλου έναντι των στρατηγικών αναφοράς. Το μοντέλο παρουσιάζεται να διατηρεί υψηλό κέρδος δικτύου σε ολόκληρη τη κλινική περιοχή κατωφλίων (0.01-0.50)

Η γενική ανάλυση της καμπύλης εκτείνεται στο εύρος κατωφλίων [0.01 – 0.50] που καλύπτει όλες τις ρεαλιστικές κλινικές περιοχές απόφασης για ένα συμβάν με επιπολασμό περίπου 12% και στηρίζεται στον τύπο της ενότητας 3.2.3. Αρχικά, η καμπύλη του μπλε μοντέλου βρίσκεται σταθερά πάνω από αμφότερες τις στρατηγικές αναφοράς σε ολόκληρη τη κλινικά σχετική περιοχή, αποδεικνύοντας πως η χρήση του παρέχει καθαρό όφελος έναντι της εμπειρικής λήψης απόφασης. Επίσης παρατηρείται μια αρνητική κλήση της καμπύλης μετά το κατώφλι 0.12, κοντά στον επιπολασμό καρδιακής ανεπάρκειας της στρατηγικής «Treat All» των 12.26%, παρουσιάζοντας πως οποιαδήποτε αδιάκριτη παρέμβαση σε όλους τους ασθενείς πάνω από αυτό το όριο θα προκαλούσε περισσότερη ζημία πρός όφελος. Το μοντέλο διατηρεί θετικό Κέρδος Δικτύου (Net Benefit) σε όλο το εύρος, με σταδιακή μείωση συναρτήσει αυξανόμενης αυστηρότητας. Τρίτον και τελευταίο, η ποσοτική υπεροχή στο κατώφλι 0.10 του Κέρδος Δικτύου σε 0.096 από 0.068 που είχε στο κατώφλι 0.20. Σε άλλα σενάρια, υπογραμμίζεται πως η χρήση του προτεινόμενου συστήματος προσφέρει όφελος ζωών έναντι των απλούστερων στρατηγικών αναφοράς.

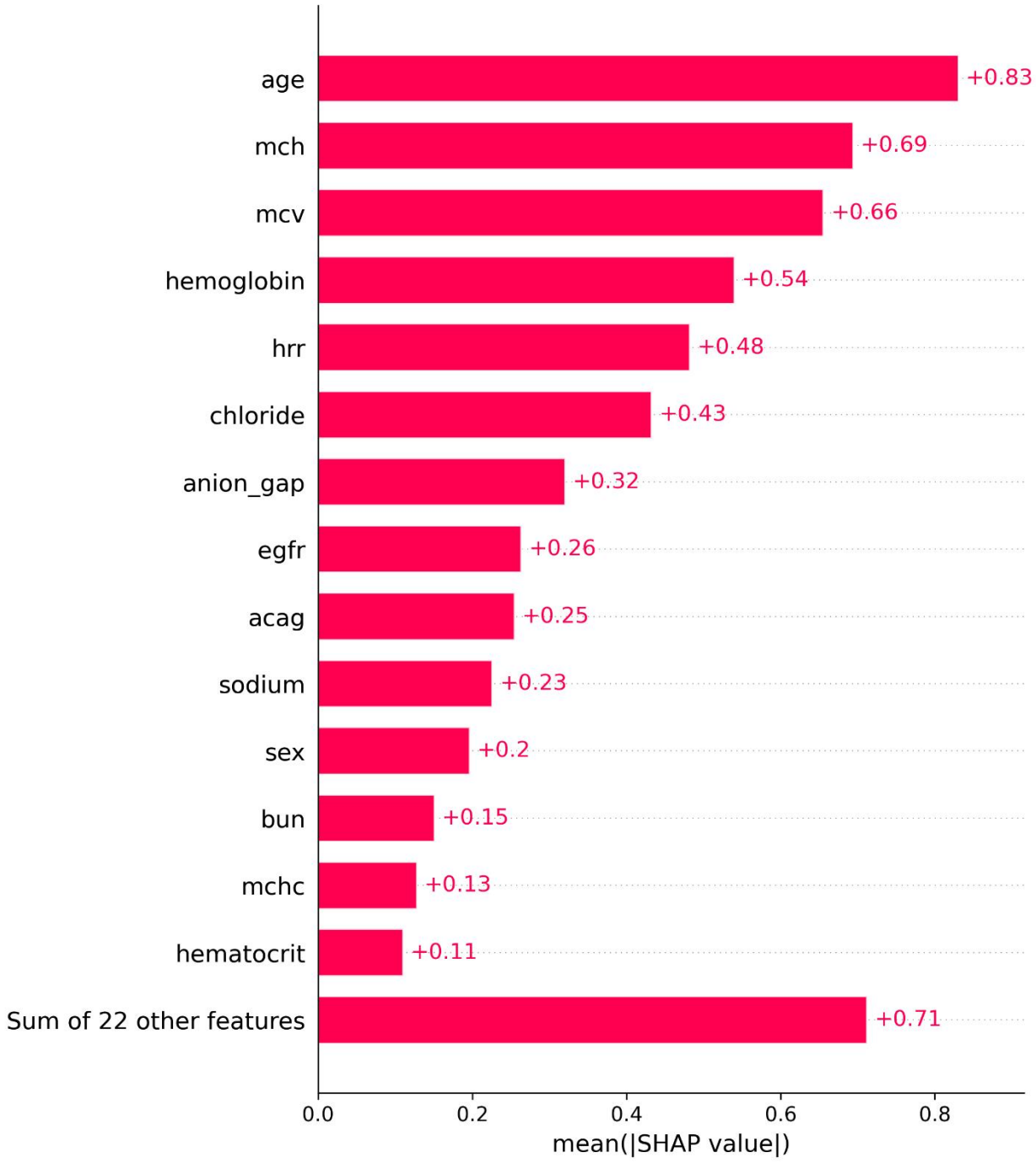
Συνολικά, η σύνθεση των τριών διαγνωστικών διαγραμμάτων δημιουργεί μία ολοκληρωμένη εικόνα ποιότητας του μοντέλου πρωταθλητή, καλύπτοντας όλες τις διαστάσεις που απαιτούνται για μια αξιόπιστη κλινική εφαρμογή.

6.5 Αξιολόγηση και Ερμηνεία των τελικών Διαγραμμάτων SHAP

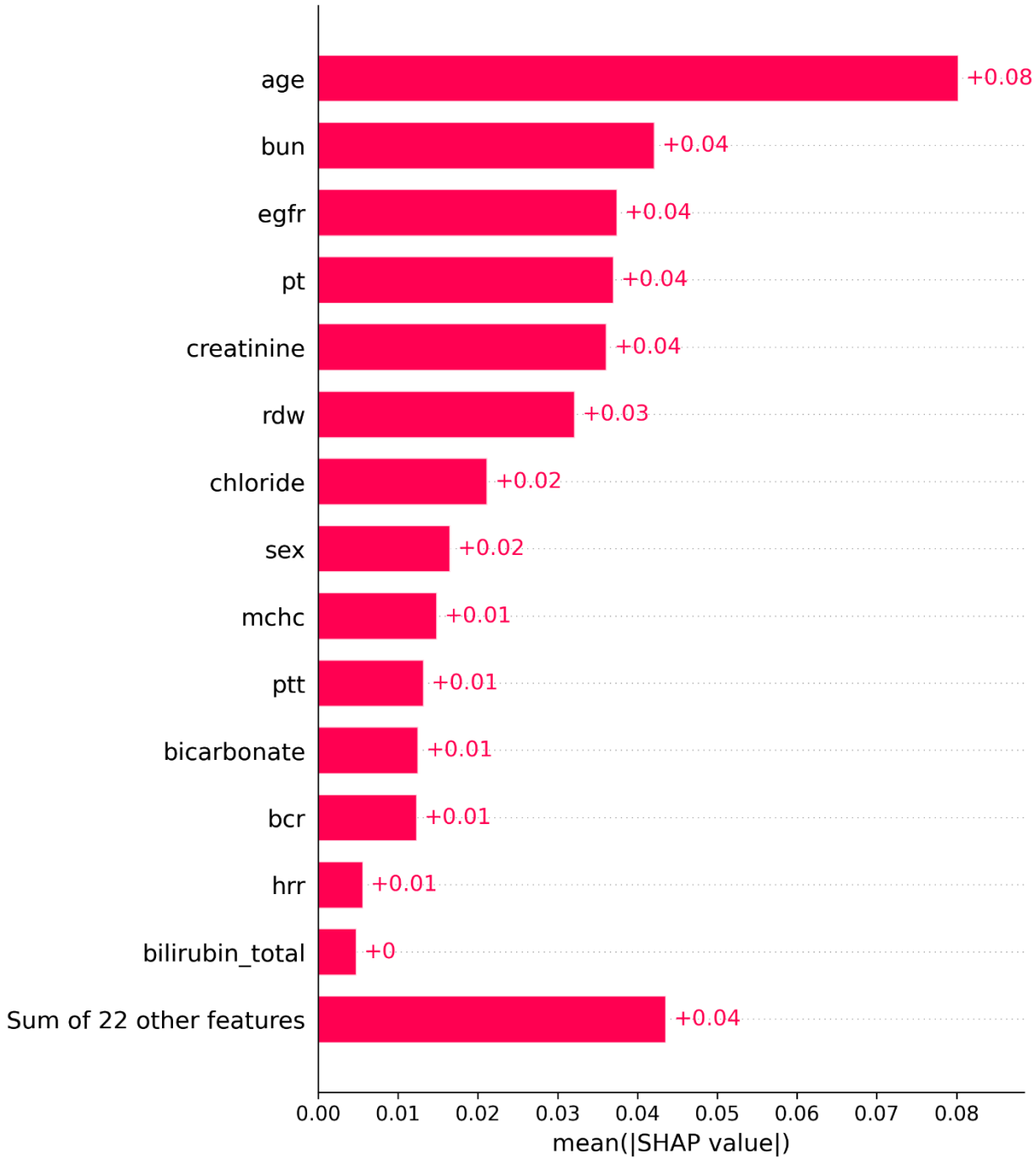
Η φύση του «μαύρου κουτιού» που χαρακτηρίζει τα μοντέλο μηχανικής μάθησης αποτρέπει την υιοθέτηση του από την ιατρική κοινότητα, παρά τις ακριβείς προβλέψεις που παρέχει. Η ερμηνευσιμότητα των διαγνωστικών αποτελεσμάτων αποτελεί απαραίτητη συνθήκη για την εμπιστοσύνη και εγκυρότητα τους, τοποθετώντας την Εξηγήσιμη Τεχνητή Νοημοσύνη στο επίκεντρο. Στην παρούσα διπλωματική εργασία η αξιολόγηση της ερμηνευσιμότητας οργανώνεται σε 6 συμπληρωματικά διαγράμματα SHAP, 2 από κάθε είδος, που παρήχθησαν από τη ροή λειτουργίας της μηχανικής μάθησης και καθένα αποτυπώνει διαφορετική όψη της λογικής του μοντέλου. Οι 3 τύποι διαγράμματος είναι το διάγραμμα μέσης σημαντικότητας χαρακτηριστικών, το διάγραμμα διασποράς ανά χαρακτηριστικό και το διάγραμμα ανάλυσης κινδύνου ανά ασθενή. Σε κάθε υποενότητα πραγματοποιείται η σύγκριση των αποτελεσμάτων από την πρώτη κανονικοποιημένη δοκιμή, με μοντέλο πρωταθλητή τη Λογιστική Παλινδρόμηση και τη τελευταία δοκιμή με υπερπαραμετροποίηση μέσω Optuna που είχε επιλογή μοντέλου πρωταθλητή τον αλγόριθμο Τυχαίου Δάσους.

6.5.1 Διάγραμμα μέσης σημαντικότητας χαρακτηριστικών

Το διάγραμμα μέσης σημαντικότητας αντιπροσωπεύει τη καθολική σημασία του κάθε χαρακτηριστικού στη πιθανότητα πρόβλεψης και κατατάσσει τα χαρακτηριστικά εισόδου με βάση τη μέση απόλυτη τιμή SHAP. Η συμπυκνωμένη εικόνα που δημιουργούν τα οριζοντίως τοποθετημένα χαρακτηριστικά οπτικοποιεί τον κλινικό ρόλο τους στη παθοφοφυσιολογία της ΚΑ του πάσχοντος ασθενή. Παρακάτω, στην **Εικόνα 63** παρουσιάζονται τα αποτελέσματα των δύο δοκιμών, πρώτης και τελευταίας ενώ στο πιο κάτω **Πίνακα 6.10** γίνεται η σύγκριση τους.



Εικόνα 63 : Διάγραμμα Μέσης Σημαντικότητας χαρακτηριστικών (SHAP Importance Bar) του μοντέλου Λογιστικής Παλινδρόμησης επί του συνόλου βαθμονόμησης.



Εικόνα 64 : Διάγραμμα Μέσης Σημαντικότητας χαρακτηριστικών (SHAP Importance Bar) του μοντέλου «Πρωταθλητή» Τυχαίου Δάσους επί του συνόλου βαθμονόμησης.

Πίνακας 6.12: Πίνακας ταξινόμησης κατά φθίνουσα σειρά μέσης τιμής |SHAP| των δεκατεσσάρων κορυφαίων χαρακτηριστικών εισόδου του μοντέλου Λογιστικής Παλινδρόμησης

Όνομα Χαρακτηριστικού	Επιστημονική Ονομασία	Μέση |SHAP|	Κατηγορία
Age	Ηλικία Ασθενούς	+0.83	Δημογραφικό
Mch	Μέση Πυκνότητα Αιμοσφαιρίνης Ερυθρών	+0.69	Αιματολογικός Δείκτης
Mcv	Μέσος Όγκος Ερυθρών	+0.66	Αιματολογικός Δείκτης
Hemoglobin	Αιμοσφαιρίνη	+0.54	Αιματολογικός Δείκτης
Hrr	Λόγος Αιμοσφαιρίνης προς RDW	+0.48	Σύνθετος Δείκτης
Chloride	Χλώριο Ορού	+0.43	Ηλεκτρολύτης
Anion_gap	Χάσμα Ανιόντων	+0.32	Οξεοβασική Ισορροπία
Egfr	Εκτιμώμενος Ρυθμός Σπειραματικής Διήθησης	+0.26	Σύνθετος Δείκτης
Acag	Διορθωμένο Χάσμα Ανιόντων για Αλβουμίνη	+0.25	Σύνθετος Δείκτης
Sodium	Νάτριο Ορού	+0.23	Ηλεκτρολύτης
Sex	Φύλο	+0.20	Δημογραφικό
Bun	Άζωτο Ουρίας Αίματος	+0.15	Νεφρικός Δείκτης
Mchc	Μέση Συγκέντρωση Αιμοσφαιρίνης κατά Όγκο	+0.13	Αιματολογικός Δείκτης
Hematocrit	Αιματοκρίτης	+0.11	Αιματολογικός Δείκτης

Πίνακας 6.13: Πίνακας ταξινόμησης κατά φθίνουσα σειρά μέσης τιμής |SHAP| των δεκατεσσάρων κορυφαίων χαρακτηριστικών εισόδου του μοντέλου «Πρωταθλητή» Τυχαίου Δάσους.

Όνομα Χαρακτηριστικού	Επιστημονική Ονομασία	Μέση τιμή |SHAP|	Κατηγορία
Age	Ηλικία Ασθενούς	+0.08	Δημογραφικό
Bun	Άζωτο Ουρίας Αίματος	+0.04	Νεφρικός Δείκτης
Egfr	Εκτιμώμενος Ρυθμός Σπειραματικής Διήθησης	+0.04	Σύνθετος Δείκτης
Pt	Χρόνος Προθρομβίνης	+0.04	Δείκτης Πήξης
Creatinine	Κρεατινίνη Ορού	+0.04	Νεφρικός Δείκτης

Rdw	Εύρος Κατανομής Ερυθρών Αιμοσφαιρίων	+0.04	Αιματολογικός Δείκτης
Chloride	Χλώριο Ορού	+0.03	Ηλεκτρολύτης
Sex	Φύλο	+0.02	Δημογραφικό
Mchc	Μέση Συγκέντρωση Αιμοσφαιρίνης κατά Όγκο	+0.02	Αιματολογικός Δείκτης
Ptt	Μερικός Χρόνος Θρομβοπλαστίνης	+0.01	Δείκτης Πήξης
Bicarbonate	Διττανθρακικά	+0.01	Οξεοβασική Ισορροπία
Bcr	Λόγος Αζώτου Ουρίας / Κρεατινίνης	+0.01	Σύνθετος Δείκτης
Hrr	Λόγος Αιμοσφαιρίνης προς RDW	+0.01	Σύνθετος Δείκτης
Bilirubim_total	Ολική Χολερυθρίνη	+0.01	Ηπατικός Δείκτης

Ξεκινώντας και συγκρίνοντας τη πρώτη εμφανή διαφορά στην απόλυτη κλίμακα τιμών |SHAP| παρατηρούμε ότι οι τιμές διαφέρουν κατά μία τάξη μεγέθους ανά δοκιμή. Στο μοντέλο της Λογιστικής Παλινδρόμησης, οι τιμές κυμαίνονται στο εύρος [+0.11 - +0.83] και στο μοντέλο Τυχαίου Δάσους στο [+0.01 - +0.08]. Η διαφορά αυτή δεν επηρεάζει τη προγνωστική ισχύ των μοντέλων αλλά οφείλεται σε διαφορετική κλίμακα εξόδου, με τη Λογιστική Παλινδρόμηση να παράγει τιμές SHAP στο χώρο του λογάριθμου των σχετικών πιθανοτήτων (log-odds), με τις μεταβολές μερικών μονάδων να αντιστοιχούν σε σημαντικές αλλαγές πιθανότητας. Από την άλλη πλευρά, το Τυχαίο Δάσος παράγει τιμές SHAP στον δυαδικό χώρο πιθανοτήτων [0,1], με τις μεταβολές να είναι σημαντικά μικρότερες.

Η ηλικία διατηρείται κυρίαρχο χαρακτηριστικό και στις δύο δοκιμές, καθώς στη δοκιμή της Λογιστικής Παλινδρόμησης είναι 1^{η} θέση με τιμή ίσή 0.83, πολύ μεγαλύτερη του δεύτερου χαρακτηριστικού που είναι 0.63 και στη δοκιμή του Τυχαίου Δάσους εξακολουθεί να κατέχει τη 1^{η} θέση με τιμή 0.08, διπλάσια από το δεύτερο χαρακτηριστικό που είναι το BUN με τιμή 0.04. Η σύμπτωση αυτή τεκμηριώνει πως οι δύο τεχνικά διαφορετικοί αλγόριθμοι καταλήγουν στο ίδιο επιστημονικά αποδεκτό συμπέρασμα.

Επίσης παρατηρείται πως ο Εκτιμώμενος Ρυθμός Σπειραματικής Διήθησης (eGFR), εμφανίζεται και στα δύο μοντέλα, αλλά σε διαφορετικές θέσεις στο καθένα. Πρώτα απ' όλα στη πρώτη δοκιμή ο σύνθετος δείκτης κατατάσσεται στην 8^{η} θέση και στη τελευταία δοκιμή του μοντέλου Τυχαίου Δάσους κατατάσσεται στη 3^{η} θέση, υπογραμμίζοντας τη βαρύτητα που έχει. Ένα βιολογικό χαρακτηριστικό που αλληλοεπιδρά με τον eGFR είναι η Κρεατινίνη (Creatinine), κατά την οποία στη πρώτη δοκιμή απουσιάζει εντελώς από τα 14 κορυφαία χαρακτηριστικά και στη τελευταία δοκιμή την βλέπουμε να καταλαμβάνει τη 5^{η} θέση. Η συμπεριφορά αυτή των αλγορίθμων δεν είναι καθόλου τυχαία και είναι κλινικά συμβατή, αντικατοπτρίζοντας τη διαφορετική ικανότητα

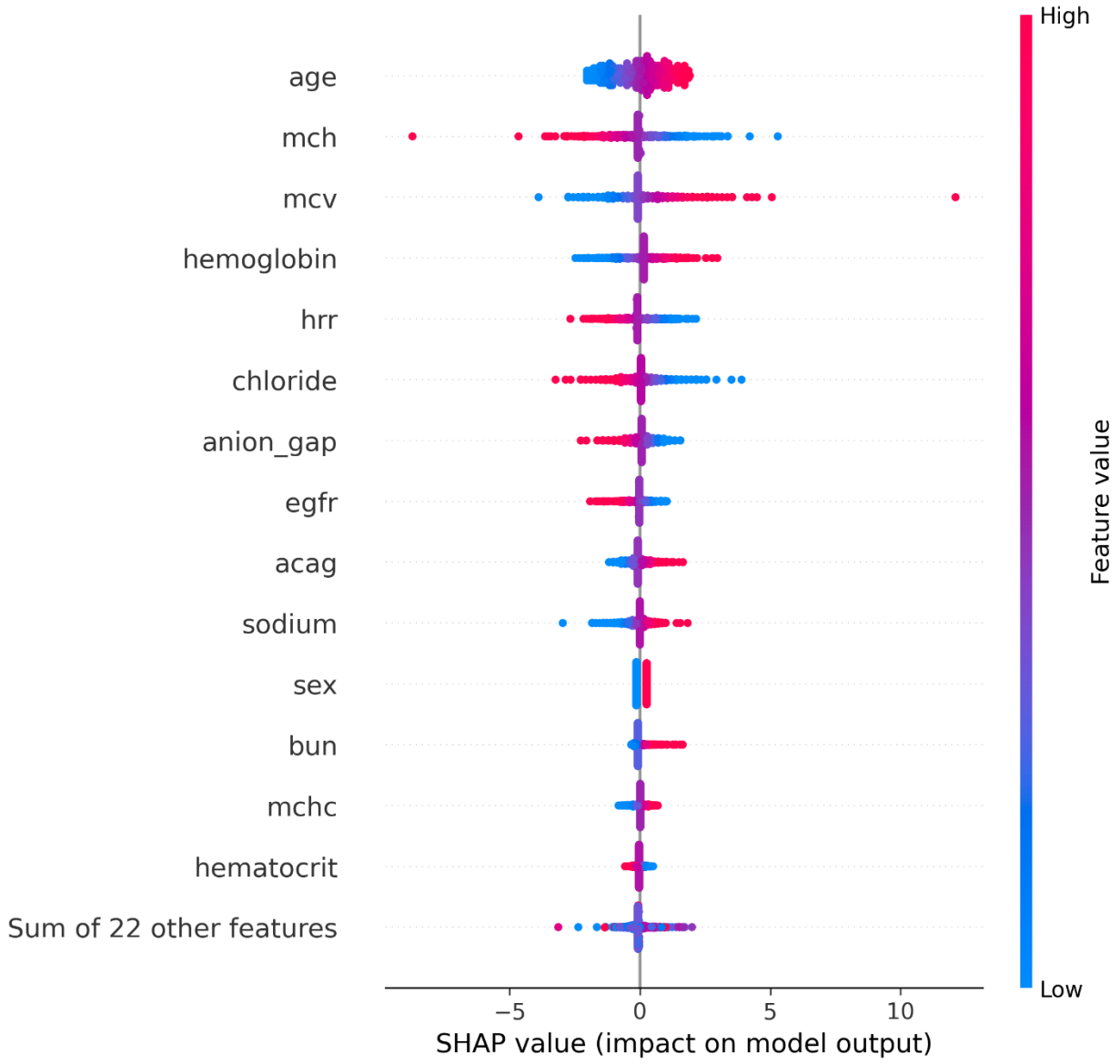
αξιοποίησης συσχετισμένων μεταβλητών που έχει ο καθένας ξεχωριστά. Χαρακτηριστικό παράδειγμα παρουσιάζεται το μη-γραμμικό σύνολο δέντρων του Τυχαίου Δάσους που μπορεί να αξιοποιήσει και τα δύο χαρακτηριστικά παράλληλα, εκμεταλλευόμενο τυχόν μη-γραμμικές αλληλεπιδράσεις, δίνοντας τη δυνατότητα παρατήρησης της διαφοράς στη σχέση Κρεατινίνης-eGFR σε ηλικιωμένους ασθενείς με χαμηλή μυϊκή μάζα. Αυτό είναι κάτι που η Λογιστική Παλινδρόμηση αδυνατεί να πραγματοποιήσεις λόγω της γραμμικής φύσης της, παραβλέπει πολλές φορές τις μεταβλητές που οι τιμές τους δεν είναι «πολύ» συσχετισμένες.

Επιπλέον διαφορετικές παθοφυσιολογικές οδοί των δύο προγνωστικών μοντέλων παρατηρούνται στην πάθηση της αναιμίας. Η λογιστική παλινδρόμηση δίνει εξέχουσα θέση στους ερυθροκυτταρικούς δείκτες, τη Μέση Πυκνότητα Αιμοσφαιρίνης Ερυθρών (Mhc), το Μέσος Όγκος Ερυθρών (Mcv), την Αιμοσφαιρίνη (Hemoglobin) και το Λόγο Αιμοσφαιρίνης προς RDW (Hrr) όπου η γραμμικότητα του μοντέλου σχετίζει τον κίνδυνο ένας ασθενής να πάθει καρδιακή ανεπάρκεια στην αναιμία. Στη τελευταία δοκιμή, έχουμε τελείως διαφορετική κατάταξη των χαρακτηριστικών και των τιμών βαρύτητας SHAP που εμφανίζουν, με τους ερυθρυκυτταρικούς δείκτες που αναφέρθηκαν προηγουμένως να απουσιάζουν και τη θέση τους να παίρνουν οι δείκτες πήξης (Pt,Ptt), ο ηπατικός δείκτης (Bilirubin_total) και ο σύνθετος δείκτης BCR (BUN/Creatinine ratio). Η ποικιλομορφία των νέων δεικτών που προσδέθηκαν συλλαμβάνει επιπρόσθετους μηχανισμούς που συνδέονται με τη συνολική οργανική επιβάρυνση του οργανισμού, όπως η πηκτικότητα αίματος, η δυσλειτουργία του ήπατος και η νεφρική λειτουργία. Οι μηχανισμοί αυτοί προσδίδουν μια ολιστικότερη εικόνα στους προχωρημένους ασθενείς και διασφαλίζουν την πλήρη κάλυψη της ορθής διάγνωσης.

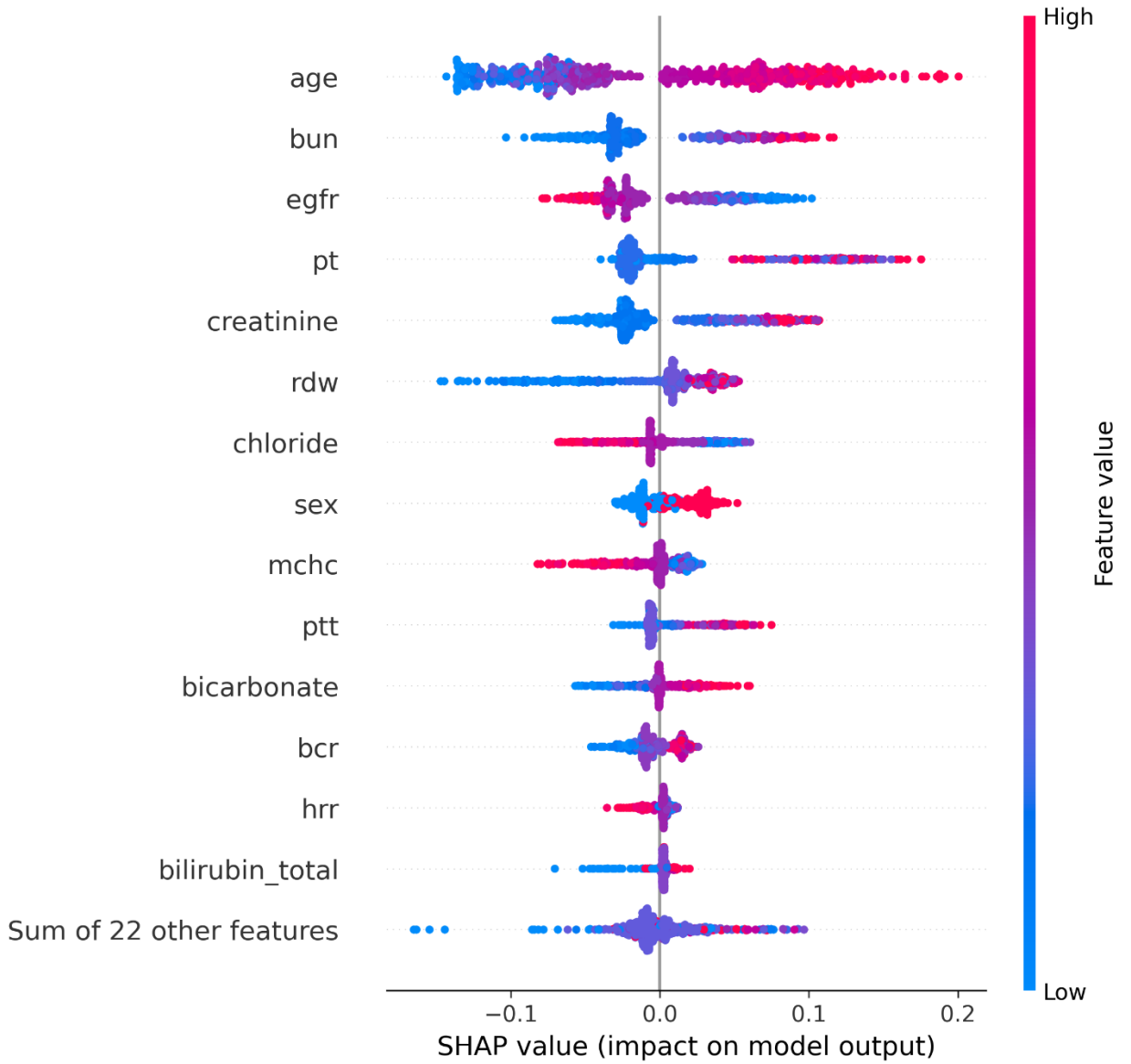
Τα υπόλοιπα επτά χαρακτηριστικά των δύο δοκιμών, τουτέστιν οι δύο ηλεκτρολυτικοί και οξεοβασικοί δείκτες του Χλωρίου (Chloride) και το διαττανθρακικό νάτριο (Bicarbonate), το οποίο περιέχεται στο anion_gap της πρώτης δοκιμής, η ηλικία, ο eGFR, το φύλο, το άζωτο ουρίας αίματος (BUN), η μέση συγκέντρωση αιμοσφαιρίνης κατά όγκο (MCHC) και ο λόγος αιμοσφαιρίνης προς RDW (HRR) δείχνουν την ευρύτερη επικράτηση και ανθεκτικότητα τους σε παραλλαγές των προβλεπτικών αλγορίθμων.

6.5.2 Διάγραμμα Διασποράς ανά Χαρακτηριστικό (Beeswarm)

Το διάγραμμα διασποράς ανά χαρακτηριστικό έρχεται να προσφέρει βαθύτερη ανάλυση από το απλό διάγραμμα μέσης σημασίας των χαρακτηριστικών, οπτικοποιώντας και την κατεύθυνση της επίδρασης κάθε χαρακτηριστικού ανάλογα με τη τιμή του. Κάθε σημείο στο διάγραμμα αναπαριστά μια ξεχωριστή νοσηλεία, η κάθετη θέση που βρίσκεται υποδηλώνει το χαρακτηριστικό ενώ η οριζόντια θέση την επίδραση που έχει στη πρόβλεψη, για την ακρίβεια η θετική πρόβλεψη τείνει προς τα δεξιά και η αρνητική προς τα αριστερά. Ένας ακόμη δείκτης βαρύτητας του χαρακτηριστικού είναι το χρώμα, με το κόκκινο χρώμα να αντιπροσωπεύει τις υψηλότερες τιμές και το μπλε τις χαμηλότερες.



Εικόνα 65 : Διάγραμμα Διασποράς ανά χαρακτηριστικό (Beeswarm) του μοντέλου Λογιστικής Παλινδρόμησης επί του συνόλου βαθμονόμησης.



Εικόνα 66 : Διάγραμμα Διασποράς ανά χαρακτηριστικό (Beeswarm) του μοντέλου «Πρωταθλητή» Τυχαίου Δάσους επί του συνόλου βαθμονόμησης.

Κατά την ανάγνωση των διαγραμμάτων ο αναγνώστης θα αντιληφθεί ότι οι κατευθύνσεις επιρροής συμφωνούν με τη κλινική παθοφυσιολογία. Πιο συγκεκριμένα, στο χαρακτηριστικό της ηλικίας που εμφανίζεται στη πρώτη γραμμή του διαγράμματος τα κόκκινα σημεία συγκεντρώνονται στο θετικό ημιεπίπεδο του SHAP άξονα, υποδεικνύοντας πως οι γηραιότεροι ασθενείς κινδυνεύουν περισσότερο να πασχίσουν ΚΑ, όπως αναμένεται και στη πραγματικότητα. Πέραν αυτού, αντίστροφη συμπεριφορά εμφανίζουν τα χαρακτηριστικά της αιμοσφαιρίνης και του αιματοκρίτη. Στο πρώτο διάγραμμα της 1ης δοκιμής παρατηρείται πως τα μπλε σημεία είναι αυτά που ωθούν τις

προβλέψεις προς τα δεξιά, επιβεβαιώνοντας πως η χαμηλή αιμοσφαιρίνη, οδηγεί σε αναιμία και αυξάνει τη προβλεπόμενη πιθανότητα καρδιακής ανεπάρκειας.

Αξιοσημείωτη είναι η μεγάλη οριζόντια έκταση των χαρακτηριστικών της Μέσης Πυκνότητας Αιμοσφαιρίνης Ερυθρών (MCH) και του Μέσος Όγκος Ερυθρών (MCV) και υποδηλώνει την διαφοροποιημένη επίδραση των χαρακτηριστικών ανά ασθενή. Δηλαδή λόγω των μη γραμμικών αλληλεπιδράσεων μεταξύ των χαρακτηριστικών, το μοντέλο Τυχαίου Δάσους μπορεί και μαθαίνει συνδυασμούς όπως ο υψηλός μέσος όγκος ερυθρών εμφανίζεται σε ασθενή νεαρής ηλικίας χωρίς αναιμία και η μέση πυκνότητα αιμοσφαιρίνης ερυθρών είναι υψηλή σε ηλικιωμένα άτομα. Αυτή ακριβώς η δυνατότητα λήψης συμπερασμάτων και τιμών εκτιμήσεις από διαφορετικές πιθανολογικές καταστάσεις ξεχωρίζει το μοντέλο Τυχαίου Δάσους έναντι της Λογιστικής Παλινδρόμησης ως μοντέλο πρωταθλητή.

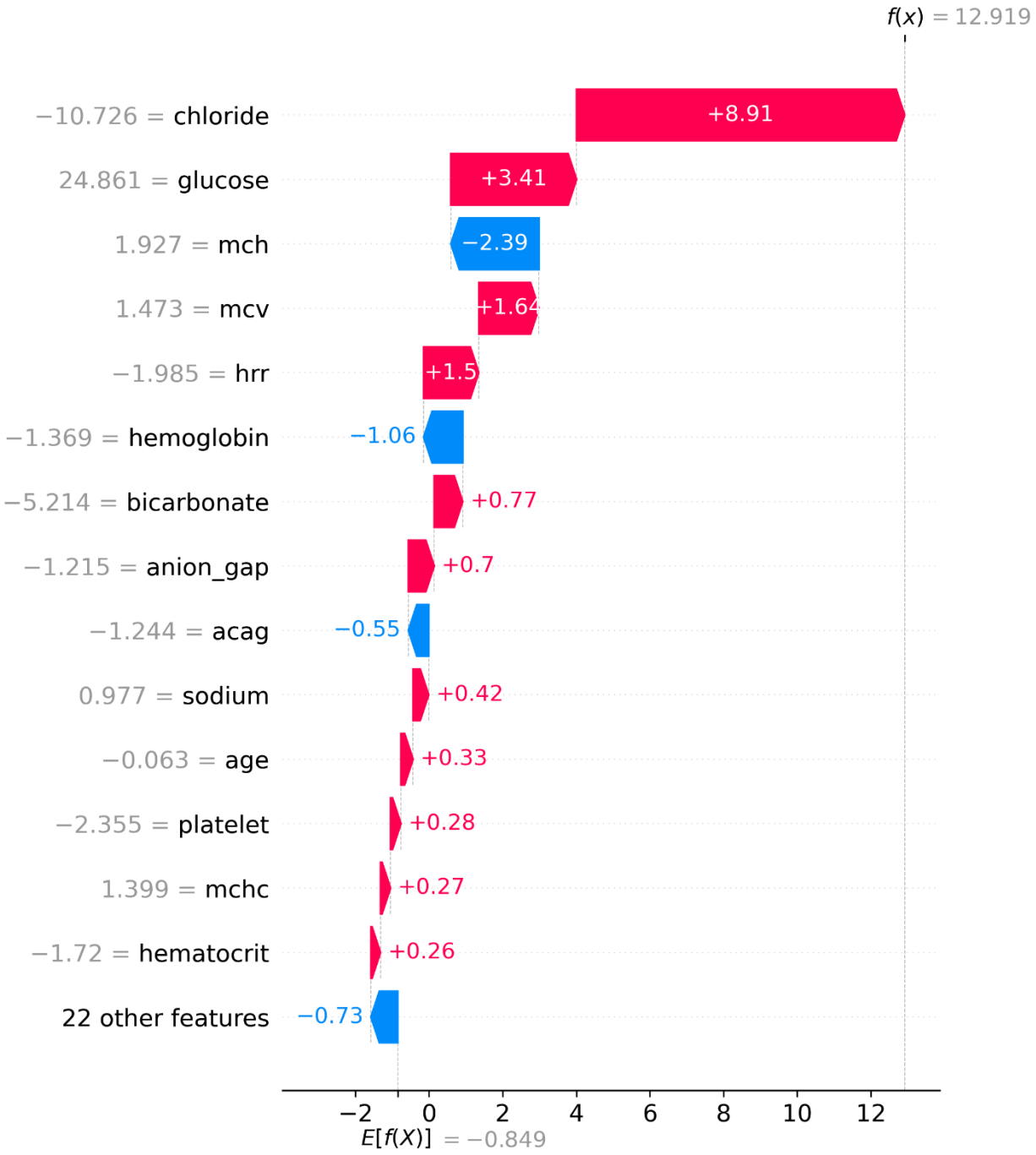
6.5.3 Διάγραμμα Καταρράκτη ανά Ασθενή

Το τρίτο και εξατομικευμένο διάγραμμα ανάλυσης SHAP ασθενούς υψηλού κινδύνου εμφάνισης καρδιακής ανεπάρκειας, το διάγραμμα Καταρράκτη (waterfall), παρουσιάζει μια διαφορετική οπτική γωνία από τις δύο προηγούμενες υποενότητες. Στην παρούσα ανάλυση το κάθε μοντέλο εστιάζει στο ασθενή με την υψηλότερη προβλεπόμενη πιθανότητα καρδιακής ανεπάρκειας στο σύνολο ελέγχου και αναλύει τη συμβολή κάθε χαρακτηριστικού στη τελική πρόβλεψη. Στο κάθε διάγραμμα οριζόντια παρατηρείται η τιμή του μοντέλου E[f(X)] προς τη κάθετη τελική τιμή πρόβλεψης f(x), με κόκκινες ράβδους για τις θετικές συμβολές και μπλε τις αρνητικές.

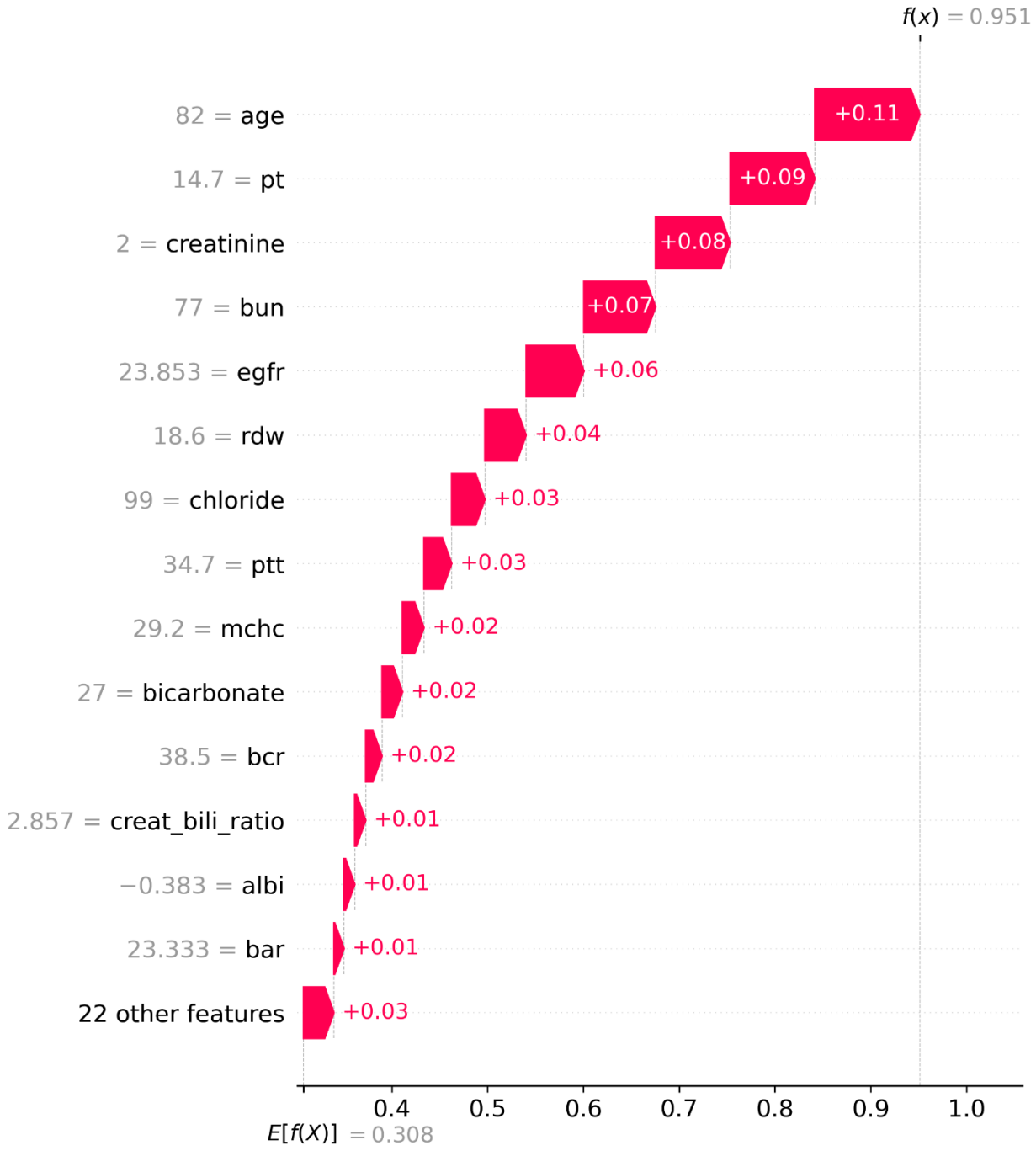
Στο πρώτο διάγραμμα της 1^{ης} δοκιμής παρουσιάζεται η μέση πιθανότητα εμφάνισης καρδιακής ανεπάρκειας στο χώρο των Λογιστικών Πιθανοτήτων με βασική τιμή E[f(X)] = -0.849 να αντιστοιχεί περίπου στο 30% του πληθυσμού. Αντίθετα το διάγραμμα Τυχαίου Δάσους παρουσιάζει συμβολές στο χώρο των πιθανοτήτων, όπου η βασική τιμή E[f(X)] = 0.308 και η τελική τιμή f(x) = 0.951, πιθανότητες που είναι άμεσα ερμηνεύσιμες ως 30.8% και 95.1%. Επιπρόσθετα οι τιμές χαρακτηριστικών που συνοδεύουν τις ράβδους στο διάγραμμα Καταρράκτη Λογιστικής Παλινδρόμησης είναι κανονικοποιημένες, ενώ στο διάγραμμα Καταρράκτη Τυχαίου Δάσους εμφανίζονται οι πραγματικές κλινικές τιμές του ασθενή, διευκολύνοντας τη κλινική ερμηνεία. Η σημαντικότερη διαφορά που κάνει ξεχωριστά αυτά τα διαγράμματα είναι πως τα δύο μοντέλα επέλεξαν ως «πιο υψηλού κινδύνου» δύο εντελώς διαφορετικούς ασθενείς του συνόλου ελέγχου. Ο ασθενής που εντοπίστηκε από τη Λογιστική Παλινδρόμηση παρουσιάζει τη κλινική εικόνα του πάσχοντος από ηλεκτρολυτική διαταραχή λόγω εντατικής ούρησης. Αυτό φαίνεται από τα ακραία υποχλωραιμία συνοδευόμενα από υπεργλυκαιμία, με τιμή z ίση με 24.86 και συμβολή +3.41 και μερική μεταβολική οξέωση από τα χαμηλά διττανρθακικά, με τιμή z ίση με -5.21 και συμβολή +0.77. Στο κάτω διάγραμμα ακριβώς, ο ασθενής που εντοπίστηκε από το Τυχαίο Δάσος παρουσιάζει την εικόνα ενός ηλικιωμένου ασθενούς με πολυοργανική επιβάρυνση. Ο ασθενής είναι 82 ετών, με κορυφαία συμβολή της στο διάγραμμα ίση με +0.11, αυξημένη κρεατινίνη 2.0 mg/dL με συνοδό αύξηση BUN στα 77 mg/Dl και παράλληλη πτώση του δείκτη eGFR στα 23.85

mL/min/1.73m^2. Τα συγκεκριμένα χαρακτηριστικά προσδίδουν την εικόνα χρόνιας νεφρικής νόσου σταδίου 4.

Η συνολική σύνθεση των δύο εξατομικευμένων αναλύσεων παρουσιάζει πως δεν υπάρχει ένας μοναδικός τύπος ασθενούς υψηλού κινδύνου αλλά η καρδιακή ανεπάρκεια εκδηλώνεται μέσω διαφορετικών παθοφυσιολογικών περιπτώσεων. Ταυτόχρονα φανερώνεται ότι η ικανότητα του Τυχαίου Δάσους να εστιάζει σε πολυοργανικά προφίλ με ομοιόμορφες συμβολές αποτελεί σημαντικό πλεονέκτημα στη κλινική πράξη, καθώς αυτός είναι ο πιο διαδεδομένες τύπος ένταξης καρδιακής ανεπάρκειας στις ΜΕΘ.



Εικόνα 67 : Διάγραμμα Καταρράκτη του ασθενούς με υψηλό κίνδυνο ΚΑ από το μοντέλο Λογιστική Παλινδρόμηση. Η βασική τιμή είναι E[f(X)] = -0.849 και η τελική πρόβλεψη f(x) = +12.919 εκφρασμένες στο χώρο των λογιστικών πιθανοτήτων



Εικόνα 68 : Διάγραμμα Καταρράκτη του ασθενούς με υψηλό κίνδυνο ΚΑ του μοντέλου «πρωταθλητή» Τυχαίου Δάσους. Η βασική τιμή είναι E[f(X)] = 0.308 και η τελική πρόβλεψη είναι f(x) =0.951 εκφρασμένες στο χώρο των πιθανοτήτων.

Πίνακας 6.14: Συγκριτικός Πίνακας των διαγραμμάτων Καταρράκτη της αρχικής και τελευταίας δοκιμής

Στοιχείο	Τιμές 1ης Δοκιμής	Τιμές Τελευταίας Δοκιμής
E[f(X)]	-0.849	0.308
F(x)	+12.919	0.951
Κλίμακα	Χώρος λογιστικών πιθανοτήτων	Χώρος πιθανοτήτων
Πιθανότητα ΚΑ	≅ 99.9%	95.1%
1ο χαρακτηριστικό συμβολής	Χλωριούχο (Chloride) +8.91	Ηλικία (Age) +0.11
2ο χαρακτηριστικό συμβολής	Γλυκόζη (Glucose) +3.41	Δείκτης Πήξης (pt) +0.09
3ο χαρακτηριστικό συμβολής	Μέση Πυκνότητα Αιμοσφαιρίνης (Mch) -2.39	Κρεατινίνη (Creatinine) +0.08

6.6 Αξιολόγηση του Blockchain Επιπέδου

Σκοπός του Blockchain επιπέδου είναι η επιτυχής λειτουργικότητα end-to-end του συστήματος, η συγκριτική αξιολόγηση του κόστους συναλλαγών στην παρούσα υλοποίηση και η εμπειρική επικύρωση των μηχανισμών ελέγχου, που είχαν σχεδιαστεί στην ενότητα 5.5. Οι μετρικές εδώ αφορούν την υπολογιστική απόδοση του έξυπνου συμβολαίου του Blockchain με το υπόλοιπο σύστημα. Η ροή λειτουργίας του Blockchain αναλύεται στην αλληλουχία σταδίων στη ενότητα 5.5.7. Κατά την διάρκεια των συναλλαγών μεταξύ του έξυπνου συμβολαίου και του Blockchain δικτύου, όλες οι συναλλαγές εκτελέστηκαν επιτυχώς, με ορατά αποτελέσματα στο Ganache GUI και στη διεπαφή της κλινικής πλατφόρμας. Συνολικά αναρτήθηκαν πέντε προβλέψεις καρδιακής ανεπάρκειας στο έξυπνο συμβόλαιο MedicalOracle, μέσω της συνάρτησης publishRiskScore, παράγοντας το blockchain_payload.json με τα αποτελέσματα των μοντέλων, μετρικών και γραφημάτων.

Στον **Πίνακα 6.12** παρουσιάζεται η ανάλυση του κόστους σε gas για τις διαδοχικές συναλλαγές που εκτελέστηκαν στο πειραματικό μέρος της διπλωματικής εργασίας.

Πίνακας 6.15: Πίνακας εκτίμησης του χρησιμοποιημένου κόστος έναντι της συναλλαγής

Συναλλαγή	Χρησιμοποιημένο Gas	Παρατηρήσεις
1η	149.885	Πρώτη εγγραφή. Οι θέσεις στο κόμβο αποθήκευσης της αλυσίδας είναι κενές
2η (ίδιος ασθενής)	132.809	Επόμενες εγγραφές στον ίδιο πίνακα
3η	132.809	-
4η	132.809	-
5η και τελευταία	149.897	Νέος ασθενής
Μέσο τυπικό κόστος συναλλαγών	133.000	Μεταξύ διαδοχικών εγγραφών
Μέσο κόστος αρχικής εγγραφής	150.000	Πρώτη εγγραφή ασθενούς

Η διαφορά κόστους ανάμεσα στην πρώτη και τις επόμενες εγγραφές, κατά 17.000 gas περίπου, δείχνει ότι η πρώτη αποθήκευση πρόβλεψης του πίνακα ενός ασθενή είναι ακριβότερη από όλες τις επόμενες. Αναμενόμενο θα έλεγε κανείς, καθώς η εικονική μηχανή Ethereum χρεώνει σημαντικά υψηλότερο κόστος για εγγραφή σε αρχικές κενές θέσεις σε σύγκριση με τις χρησιμοποιημένες. Συνήθως το κόστος καυσίμων στις αρχικές θέσεις ενός πίνακα όταν είναι κενές είναι περίπου στα 22.100, ενώ σε ήδη χρησιμοποιημένες θέσεις κατεβαίνει στα 5.000. Στην ενότητα 5.5.5, αποδείχτηκε πως η ανάρτηση μεμονωμένων συναλλαγών για κάθε πρόβλεψη γίνεται απαγορευτική, με την πλήρη παρτίδα των 109,284 του συνόλου ελέγχου, η εκτίμηση κόστους ανέρχεται σύμφωνα με την εξίσωση (27):

$$Συνολικό\ Κόστος_{Ατομικής\ Καταχώρησης} = 109{,}284 \times 133.000 \cong 14.500.000.000\ gas \quad (27)$$

Η εναλλακτική προσέγγιση καταχωρεί μόνο τη Merkle ρίζα και τα ελάχιστα μεταδεδομένα της παρτίδας, με σταθερό κόστος ανεξάρτητο του πλήθους εγγραφών. Με την δομή της παρτίδας που κατασκευάστηκε στο έξυπνο συμβόλαιο, η εκτίμηση κόστους κυμαίνεται περί τα 80.000-100.000 gas ανεξαρτήτως αν η παρτίδα περιέχει 10 ή 10.000 εγγραφές. Συγκρίνοντας τη διαφορά με γνώμονα την υπολογιστική πολυπλοκότητα, η εξοικονόμηση είναι της τάξης από O(N) σε O(1), δηλαδή διαφορά τεσσάρων τάξεων μεγέθους. Η κλιμακωσιμότητα αυτή αναδεικνύεται σημαντική τεχνική για την εμπορευματοποίηση του έξυπνου συμβολαίου, εξοικονομώντας αρκετά χρήματα παράλληλα εκτός από υπολογιστική ισχύ. Παρακάτω αναλύεται περισσότερο η αξιοποίηση του έξυπνου συμβολαίου σε δικτύου μεγάλης κλίμακας, όπως το Sepolia testnet και σε L2 επιπέδου δικτύου το Polygon. Η μεταφορά της υλοποίησης σε δημόσιο δίκτυο αποτελεί φυσική επέκταση του πειράματος της παρούσας εργασίας. Στην ενότητα 6.7 γίνεται περισσότερο ανάλυση των πλεονεκτημάτων μιας τέτοιας επέκτασης.

Επιπλέον, ο μηχανισμός επικύρωσης συγκατάθεσης του GDPR που έχει αναπτυχθεί στην ενότητα 5.5.4 λειτουργεί ως απαραίτητο βήμα ενεργοποίησης πριν την καταχώρηση την χειροκίνητη αποθήκευση των καταγραφών στην αλυσίδα, μέσω του sub_storage.py. Η λειτουργία αυτή αποδεικνύει την πραγματική εφαρμογή του μηχανισμού και όχι την απλή θεωρητική χρήση του, διότι ακόμα και ένας προγραμματιστής με πλήρη πρόσβαση στο δίκτυο και τον πηγαίο κώδικα δεν μπορεί να καταχωρήσει δεδομένα ασθενούς χωρίς ισχύουσα συγκατάθεση.

Κεφάλαιο 7: Συζήτηση

Η προτεινόμενη κλινική πλατφόρμα αφορούσε την ανάπτυξη ενός ολοκληρωμένου κλινικού συστήματος λήψης αποφάσεων, με έμφαση στη πρόβλεψη της καρδιακής ανεπάρκειας και την αποθήκευση των ιατρικών πληροφοριών μέσω της τεχνολογίας Blockchain. Στον παρόν κεφάλαιο συγκεντρώνονται τα κρισιμότερα ευρήματα του Κεφαλαίου 6 και πραγματοποιείται η ερμηνεία τους υπό το φως της επιστημονικής βιβλιογραφίας, ενώ παράλληλα συζητείται η τεχνική σημασία της πλατφόρμας.

7.1 Συνολική Αποτίμηση των Αποτελεσμάτων

Η σύγκριση τεσσάρων υποψήφιων μοντέλων έδειξε ότι ο αλγόριθμος Τυχαίου Δάσους είναι το μοντέλο νικητής. Ο αλγόριθμος έδωσε τιμή Περιοχή κάτω από την καμπύλη λήψης-αποδοχής ίση με 0.8541 στο σύνολο ελέγχου των 109.284 νοσηλειών. Η τιμή ήρθε με στενό 95% διάστημα εμπιστοσύνης: 0.8513 έως 0.8569, μετά από 2000 επαναλήψεις Bootstrap. Το μοντέλο έδωσε πιθανότητες κινδύνου με σκορ Brier ίσο με 0.0941. Η καμπύλη βαθμονόμησης του μοντέλου σχεδόν ταιριάζει με τη διαγώνιο, όπως φαίνεται στο Σχήμα 6.4. Η καμπύλη δείχνει ότι το μοντέλο ανταποκρίνεται στην εμφάνιση της νόσου, όχι μόνο σε στατιστικές εκτιμήσεις. Βλέπω ότι το μοντέλο έχει σχεδόν καθαρό κλινικό όφελος. Το μοντέλο κρατά το όφελος σε όλο το εύρος του κατωφλίου απόφασης 0.01-0.50 στην Ανάλυση Καμπυλών Απόφασης. Το μοντέλο ξεπερνά τις στρατηγικές «Θεραπεία Όλων» και «Καμία Θεραπεία». Στο Blockchain το σύστημα λειτουργεί από την αρχή μέχρι το τέλος. Το σύστημα ανεβάζει πέντε προβλέψεις στο τοπικό δίκτυο Ganache. Το σύστημα πληρώνει μέσο κόστος 133.000 gas για κάθε ανάρτηση. Το σύστημα βελτιώνει τη χρήση του δέντρου Merkle. Η βελτίωση περνά από O(N) σε O(1). Τα αποτελέσματα του Κεφαλαίου 6 δείχνουν. Μου φαίνεται ότι τρεις τεχνολογίες μπορούν να συναντηθούν. Η Εξηγήσιμη Τεχνητή Νοημοσύνη, το Blockchain, η Μηχανική Μάθηση. Όλες τρεις μπλέκονται σε μια ενιαία αρχιτεκτονική. Η αρχιτεκτονική αυτή μπορεί να δουλέψει στην επιχείρηση με επιχειρησιακή βιωσιμότητα.

7.2 Ερμηνεία στο Πλαίσιο της Βιβλιογραφίας

Κατά την άποψή μου, η σύγκριση του αποτελέσματος της διπλωματικής εργασίας με τη βιβλιογραφία δείχνει ξεχωριστές θέσεις έρευνας που το προτεινόμενο σύστημα μπορεί να καλύψει. Πρώτα, οι Moreno-Sanchez και οι συνεργάτες [27] έδειξαν παρόμοιο συνδυασμό μηχανικής μάθησης και SHAP με μικρό σύνολο ασθενών, 299 άτομα. Έπειτα, οι Desai και οι συνεργάτες [40] χρησιμοποίησαν σύνολο 9 502 ασθενών, αλλά δεν πρόσθεσαν μηχανισμό εξηγήσιμης τεχνητής νοημοσύνης. Η παρούσα εργασία χρησιμοποίησε 192 527 μοναδικούς ασθενείς από τη βάση MIMIC-IV. Το μέγεθος αυτό ξεπερνά κατά πολλές τάξεις το μέγεθος των προαναφερθέντων μελετών. Το Rudin [22] σημειώνει ότι σε σύνολα δεδομένων με καλά χαρακτηριστικά τα απλά γραμμικά μοντέλα συχνά αποδίδουν καλύτερα από τα πολύπλοκα μοντέλα. Στην εργασία μας η διαφορά αυτή προέρχεται από το Τυχαίο Δάσος. Το Τυχαίο Δάσος πιάνει τις μη γραμμικές αλληλεπιδράσεις των σύνθετων βιοδεικτών στην υποενότητα 5.1.3. Το Τυχαίο Δάσος δίνει πιο ακριβή εκτίμηση πιθανοτήτων, με το Brier 0.1241 και το αποτέλεσμα της Λογιστικής Παλινδρόμησης 0.1564. Επιπλέον το AUROC φτάνει στο 0.854, τιμή που ξεπερνά το καθιερωμένο MAGGIC σκορ ρίσκου. Το σύνολο ασθενών 51.043 έδωσε το C-index 0.741 [20]. Το C-index 0.741 δείχνει ότι η ενσωμάτωση νέων αλγορίθμων μηχανικής μάθησης στο σκορ καρδιακής ανεπάρκειας φέρνει μετρήσιμη βελτίωση.

7.3 Η αξία της Εξηγήσιμης Τεχνητής Νοημοσύνης στη Λήψη Κλινικών Αποφάσεων

Πέρα από τη μικρή διαφορά του Τυχαίου Δάσους, ένα πιο χρήσιμο πλεονέκτημα εμφανίστηκε στην σύγκριση των SHAP διαγραμμάτων στην ενότητα 6.5. Τα SHAP του Τυχαίου Δάσους είναι κατευθείαν στο διάστημα [0, 1]. Η αρχική τιμή είναι E[f(X)]=0.308. Η τελική πρόβλεψη είναι f(x)=0.951 στο Σχήμα 6.11. Αυτές οι τιμές δείχνουν 30.8% και 95.1% πιθανότητα καρδιακής ανεπάρκειας. Είναι σαφές. Η πιθανότητα καρδιακής ανεπάρκειας είναι άμεσα κατανοητή. Ο κλινικός χρήστης δεν χρειάζεται να κάνει μαθηματική αλλαγή. Το SHAP του Τυχαίου Δάσους δίνει τις τιμές κατευθείαν, έτσι ο κλινικός χρήστης βλέπει την πιθανότητα χωρίς πρόσθετο υπολογισμό. Αντίθετα, τιμές SHAP της Λογιστικής Παλινδρόμησης βρίσκονται στο χώρο των λογαριθμικών πιθανοτήτων (log-odds). Τιμές SHAP φτάνουν στο f(x)=+12.919. Να καταλάβει τιμές SHAP χρειάζεται ένα επιπλέον βήμα που δεν ξέρει το μέσο κλινικό προσωπικό. Από την εμπειρία μου, η αμεσότητα ταιριάζει με τα ευρήματα των Hur et al. [47] και Bienefeld et al. [43]. Τα ευρήματα δείχνουν ότι οι κλινικοί προτιμούν εξηγήσεις που είναι άμεσα κατανοητές, όχι εξηγήσεις που είναι πολύ αυστηρές αλλά δύσκολες. Από την προσωπική μου άποψη, η ηλικία είναι ο πιο σημαντικός κίνδυνος. Πέρα από τη κλίμακα, η ηλικία εμφανίζεται ως κορυφαίος παράγοντας κινδύνου και στα δύο μοντέλα. Η ηλικία ταιριάζει με τη γνωστή κλινική γνώση. Η ηλικία ταιριάζει με τη βαρύτητα που δίνει το MAGGIC Risk Score [20]. Η ηλικία ενισχύει την κλινική εγκυρότητα του προτεινόμενου συστήματος. Το Τυχαίο Δάσος φέρνει νέους δείκτες συ νοσηροτήτων. Το Τυχαίο Δάσος δείχνει το BUN, τη κρεατινίνη και το eGFR. Το BUN, η κρεατινίνη και το eGFR δείχνουν την αλληλεπίδραση καρδιάς και νεφρού (cardiorenal syndrome), που έχει τεκμηριωθεί στη βιβλιογραφία. [30] δείχνει ότι το μοντέλο δεν εντοπίζει στατιστικές συσχετίσεις, αλλά εντοπίζει κλινικά σημαντικά μοτίβα. Για αυτό, η αξία της εξηγήσιμης τεχνητής νοημοσύνης στην προτεινόμενη πλατφόρμα δεν περιορίζεται στη συμμόρφωση με τις σύγχρονες απαιτήσεις διαφάνειας. Η αξία της εξηγήσιμης τεχνητής νοημοσύνης στην προτεινόμενη πλατφόρμα λειτουργεί ως εργαλείο που συνδέει την απόσταση μεταξύ του μαύρου κουτιού των αλγορίθμων και της κλινικής σκέψης του ιατρικού προσωπικού. Το βλέπω χρήσιμο στην καθημερινή δουλειά.

7.4 Συνεισφορά Blockchain Επιπέδου

Το να βάλω το επίπεδο Blockchain κάνει τη δουλειά αυτή να διαφέρει από τις άλλες υλοποιήσεις Συστήματος Υποστήριξης Κλινικών Αποφάσεων. Στο άρθρο των Bathula et al. [56] διαβάζω ότι από 51 μελέτες που συνδυάζουν τεχνητή νοημοσύνη και Blockchain στην υγεία, μόνο μία προσθέτει εξηγήσιμη τεχνητή νοημοσύνη. Αυτό κάνει το τριπλό συνδυασμό μηχανικής μάθησης, SHAP και έξυπνου συμβολαίου στη δική μας υλοποίηση κάτι μοναδικό. Πέρα από το ότι τα δεδομένα παραμένουν αμετάβλητα στο δίκτυο, το πιο σημαντικό που φέρνει το επίπεδο Blockchain είναι η αλλαγή των απαιτήσεων του. Καθώς γράφω, νιώθω τη σημασία του βήματος αυτού. Καθώς δουλεύω με το GDPR, βλέπω ότι το GDPR μετατρέπει το νομικό κείμενο σε εκτελέσιμη συνθήκη με τη λογική συγκατάθεση της ενότητας 5.5.4. Η εμπειρική επικύρωση της

ενότητας 6.6 δείχνει ότι ακόμα και ένας χρήστης με όλη πρόσβαση στον πηγαίο κώδικα και στα στοιχεία του δικτύου δεν μπορεί να βάλει δεδομένα ασθενούς χωρίς έγκυρη συγκατάθεση. Το GDPR ενσωματώνει την αρχή της προστασίας από σχεδιασμού (privacy-by-design) του Άρθρου 25 στο επίπεδο της εικονικής μηχανής Ethereum [24]. Επιπλέον, η βελτιστοποίηση της κλιμάκωσης με τη δομή Merkle Tree μειώνει το κόστος υπολογισμού από O(N) σε O(1). Η διαφορά είναι τέσσερις τάξεις. Η διαφορά κάνει το σύστημα πιο εφικτό. Η διαφορά λύνει τα προβλήματα κόστους που η βιβλιογραφία δείχνει ως κύριο εμπόδιο στην υιοθέτηση του Blockchain στην υγεία [57], [58]. Η συνδυασμένη λειτουργία βάζει το έξυπνο συμβόλαιο MedicalOracle σε καλή θέση για μελλοντική μετάπτωση σε δίκτυα παραγωγής. Η συνδυασμένη λειτουργία ακολουθεί την κατεύθυνση που προτείνεται στην επόμενη ενότητα.

7.5 Σχεδιαστική Συμμόρφωση υπό τα Σύγχρονα Κανονιστικά Πλαίσια

Η σωστή αξιολόγηση κάθε ερευνητικής εργασίας απαιτεί ειλικρινή αποδοχή των περιορισμών της μεθοδολογίας. Για μένα, η ανάγκη να παραδεχθούμε τους περιορισμούς είναι ξεκάθαρη. Εγώ, όταν κοιτάζω το σύνολο εκπαίδευσης, βλέπω ότι το σύνολο εκπαίδευσης είναι πολύ μεγάλο. Παρά το μεγάλο μέγεθος, τα δεδομένα της βάσης MIMIC-IV προέρχονται από έναν τριτοβάθμιο νοσηλευτικό φορέα των Ηνωμένων Πολιτειών. Τα δεδομένα δείχνουν μια γεωγραφική και μια δημογραφική σύνθεση. Στο σύνολο ελέγχου, ο επιπολασμός καρδιακής ανεπάρκειας είναι 12.26%. Η τιμή είναι σαφώς υψηλότερη από τον επιπολασμό στον γενικό πληθυσμό, που σύμφωνα με τις σύγχρονες επιδημιολογικές μελέτες κυμαίνεται σε 1-2% [15]. Μου φαίνεται ότι η χρήση του μοντέλου σε άλλα ιατρικά μέρη, ειδικά σε εξωτερικά ιατρεία πρωτοβάθμιας φροντίδας, μπορεί να χρειαστεί επαναρύθμιση των πιθανοτήτων. Το pipeline υποστηρίζει πλήρως τη διαδικασία της επαναρύθμισης. Δεύτερον, τα στενά διαστήματα εμπιστοσύνης Bootstrap αποδεικνύουν τη σταθερότητα του μοντέλου στο συγκεκριμένο σύνολο ελέγχου. Η σταθερότητα του μοντέλου στο συγκεκριμένο σύνολο ελέγχου δεν αντικαθιστά την εξωτερική επικύρωση σε ανεξάρτητο νοσοκομειακό σύνολο. Η εξωτερική επικύρωση σε ανεξάρτητο νοσοκομειακό σύνολο είναι το επόμενο βήμα της έρευνας. Τρίτον, πραγματοποιήσαμε την ανάλυση των διαγραμμάτων Καταρράκτη της ενότητας 6.5.3 σε διαφορετικούς αντιπροσωπευτικούς ασθενείς ανά μοντέλο. Η επιλογή έδειξε καθαρά την παθοφυσιολογική ποικιλομορφία των περιπτώσεων υψηλού κινδύνου. Αλλά αφήνει ανοιχτή τη δυνατότητα μελλοντικής διπλής ανάλυσης στον ίδιο ασθενή. Έτσι μπορούμε να συγκρίνουμε άμεσα τις ερμηνευτικές αποδόσεις των δύο μοντέλων. Από τη δική μου εμπειρία, η δυνατότητα είναι πολύ χρήσιμη. Τέλος, αναπτύξαμε το έξυπνο συμβόλαιο σε τοπικό περιβάλλον προσομοίωσης το Ganache. Η επιλογή εξυπηρέτησε την από αρχής μέχρι το τέλος επικύρωση χωρίς κόστος δημόσιου δικτύου. Ωστόσο δεν αξιολογήσαμε την απόδοση σε πραγματική συμφόρηση δικτύου. Από την άποψή μου, οι περιορισμοί δεν ακυρώνουν τα ευρήματα της εργασίας. Οι περιορισμοί ορίζουν σαφώς το πεδίο εγκυρότητας των ευρημάτων. Οι περιορισμοί ανοίγουν τις προοπτικές για μελλοντική επέκταση.

7.6 Περιορισμοί της Εργασίας

Η ορθή ακαδημαϊκή αποτίμηση κάθε ερευνητικής εργασίας απαιτεί την ειλικρινή αναγνώριση των μεθοδολογικών της περιορισμών. Πρώτον, παρά το ιδιαίτερα μεγάλο μέγεθος του συνόλου εκπαίδευσης, τα δεδομένα της βάσης MIMIC-IV προέρχονται από έναν τριτοβάθμιο νοσηλευτικό φορέα των Ηνωμένων Πολιτειών και αντικατοπτρίζουν συγκεκριμένη γεωγραφική και δημογραφική σύνθεση, με τον επιπολασμό καρδιακής ανεπάρκειας στο σύνολο ελέγχου να ανέρχεται σε 12.26%, τιμή σαφώς υψηλότερη του αντίστοιχου επιπολασμού στον γενικό πληθυσμό που σύμφωνα με τις σύγχρονες επιδημιολογικές μελέτες κυμαίνεται σε 1-2% [15]. Η παρατήρηση αυτή υποδηλώνει ότι η εφαρμογή του μοντέλου σε διαφορετικά κλινικά περιβάλλοντα, και ιδιαίτερα σε εξωτερικά ιατρεία πρωτοβάθμιας φροντίδας, ενδέχεται να απαιτεί επανα-βαθμονόμηση των πιθανοτήτων, διαδικασία η οποία υποστηρίζεται πλήρως από την αρχιτεκτονική του pipeline. Δεύτερον, η στενότητα των διαστημάτων εμπιστοσύνης Bootstrap τεκμηριώνει την εσωτερική σταθερότητα του μοντέλου επί του συγκεκριμένου συνόλου ελέγχου, χωρίς ωστόσο να υποκαθιστά την εξωτερική επικύρωση σε ανεξάρτητο νοσοκομειακό σύνολο, βήμα που αποτελεί φυσική επέκταση της παρούσας έρευνας. Τρίτον, η ανάλυση των διαγραμμάτων Καταρράκτη της ενότητας 6.5.3 πραγματοποιήθηκε σε διαφορετικούς αντιπροσωπευτικούς ασθενείς ανά μοντέλο, επιλογή που ανέδειξε με ξεκάθαρο τρόπο την παθοφυσιολογική ποικιλομορφία των περιπτώσεων υψηλού κινδύνου αλλά αφήνει ανοιχτή τη δυνατότητα μελλοντικής διπλής ανάλυσης στον ίδιο ασθενή για άμεση σύγκριση των ερμηνευτικών αποδόσεων των δύο μοντέλων. Τέλος, η ανάπτυξη του έξυπνου συμβολαίου πραγματοποιήθηκε σε τοπικό περιβάλλον προσομοίωσης Ganache, επιλογή που εξυπηρέτησε με ιδανικό τρόπο την end-to-end επικύρωση χωρίς το κόστος δημόσιου δικτύου, χωρίς ωστόσο να αξιολογεί την απόδοση σε συνθήκες πραγματικής συμφόρησης δικτύου. Οι περιορισμοί αυτοί δεν αναιρούν τα ευρήματα της εργασίας αλλά οριοθετούν με σαφήνεια το πεδίο εγκυρότητας τους και διαμορφώνουν τις προοπτικές μελλοντικής επέκτασης.

7.7 Μελλοντικές Κατευθύνσεις και Τελικά Συμπεράσματα

Οι περιορισμοί που βρέθηκαν στην προηγούμενη ενότητα δημιουργούν φυσικές επεκτάσεις για τη δουλειά αυτή σε τέσσερα επίπεδα. Στο επίπεδο της μηχανικής μάθησης, η εξωτερική δοκιμή του μοντέλου σε ανεξάρτητα κλινικά σύνολα δεδομένων, όπως το σύνολο δεδομένων eICU ή το τοπικό σύνολο δεδομένων ελληνικών νοσοκομείων, θα δείξει τη δυνατότητα του μοντέλου να δουλεύει παντού. Το μοντέλο θα αποδείξει ότι μπορεί να γενικεύσει. Το μοντέλο θα κάνει τη χρήση στην κλινική πιο σίγουρη. Εγώ πιστεύω ότι η προσθήκη τεχνικών ομοσπονδιακής μάθησης (Federated Learning) όπως προτείνουν οι Otoum και άλλοι [55] θα επιτρέψει τη μάθηση μοντέλων από πολλά νοσοκομεία χωρίς να μεταφερθούν ευαίσθητα δεδομένα. Στο Blockchain, η μετάβαση του έξυπνου συμβολαίου από το Ganache στο Sepolia testnet ή στο Polygon δίνει μια πιο αληθινή εκτίμηση του κόστους στην παραγωγή. Το Ganache είναι τοπικό. Το Sepolia testnet είναι δημόσιο. Το Polygon είναι Layer-2 δίκτυο. Το Polygon δείχνει κόστος συναλλαγής που είναι τρεις φορές

μικρότερο από το Ethereum. Στη διεπαφή χρήστη, η τακτική αξιολόγηση ευχρηστίας με κλινικούς ιατρούς και το System Usability Scale δείχνει πώς η πλατφόρμα γίνεται αποδεκτή από την κλινική πλευρά. Για μένα, είναι σημαντικό να δω το έξυπνο συμβόλαιο να δουλεύει σε πραγματικά δίκτυα. Το αποτέλεσμα είναι μια πλατφόρμα που είναι αποδεκτή από κλινικούς. Τέλος, στην κανονιστική συμμόρφωση, η προετοιμασία της πλατφόρμας θα ταξινομήσει την πλατφόρμα ως ιατροτεχνολογικό λογισμικό σύμφωνα με το Κανονισμό (ΕΕ) 2017/745 για τα Ιατροτεχνολογικά Προϊόντα (MDR). Πιστεύω ότι η προετοιμασία θα ολοκληρώσει τη μετάβαση. Η μετάβαση θα φέρει το σύστημα στην κλινική χρήση. Στο τέλος, η διπλωματική εργασία δείχνει με πραγματικά δεδομένα ότι ο συνδυασμός καλής ικανότητας με το AUROC ίσο με 0.854, η σωστή βαθμονόμηση με το Brier σκορ 0.094, οι κλινικά ερμηνεύσιμες εξηγήσεις με το SHAP, η σταθερή καταγραφή στο Blockchain με συμμόρφωση με το GDPR και η πλήρης τριεπίπεδη διαδικτυακή υλοποίηση, είναι τεχνικά εφικτή και ακαδημαϊκά ώριμη. Καθώς έγραφα τη διπλωματική εργασία, είχα την αίσθηση ότι η πλατφόρμα δεν αντικαθιστά τη κλινική κρίση του ιατρού. Η πλατφόρμα προσφέρει τεκμηριωμένη υποστήριξη αποφάσεων. Η πλατφόρμα διασφαλίζει την ακεραιότητα των καταγραφών. Η πλατφόρμα προστατεύει τα δεδομένα. Προστατεύουμε τα ευαίσθητα προσωπικά δεδομένα των ασθενών. Καλύπτουμε το κενό στην έρευνα που βρέθηκε στην ενότητα 2.5.2. Το έργο τοποθετεί τη μηχανική μάθηση, την εξηγήσιμη τεχνητή νοημοσύνη και την τεχνολογία Blockchain στην κορυφή της σύγκλισης στον τομέα της υγείας. Καθώς δουλεύω, βλέπω τη μηχανική μάθηση, την εξηγήσιμη τεχνητή νοημοσύνη και τη τεχνολογία Blockchain να αλλάζουν την υγεία..

Βιβλιογραφία

[1] J. Adler-Milstein and A. K. Jha, "HITECH Act drove large gains in hospital electronic health record adoption," Health Affairs, vol. 36, no. 8, pp. 1416–1422, Aug. 2017. doi: 10.1377/hlthaff.2017.0395.

[2] C. S. Kruse, P. Goswamy, Y. Raval, and S. Marber, "Challenges and opportunities of big data in health care: A systematic review," JMIR Med. Inform., vol. 4, no. 4, e38, Nov. 2016. doi: 10.2196/medinform.5359.

[3] World Health Organization, "Global Strategy on Digital Health 2020–2025," Geneva, 2021. ISBN: 978-92-4-002092-4.

[4] European Parliament and Council of the EU, "Regulation (EU) 2025/327 on the European Health Data Space," Official Journal of the EU, Mar. 5, 2025.

[5] J. McCarthy, M. L. Minsky, N. Rochester, and C. E. Shannon, "A proposal for the Dartmouth Summer Research Project on AI," 1955, repr. AI Magazine, vol. 27, no. 4, pp. 12–14, 2006. doi: 10.1609/aimag.v27i4.1904.

[6] S. J. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 4th ed. Hoboken, NJ: Pearson, 2020.

[7] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," Science, vol. 349, no. 6245, pp. 255–260, Jul. 2015. doi: 10.1126/science.aaa8415.

[8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, pp. 436–444, May 2015. doi: 10.1038/nature14539.

[9] E. Hossain et al., "Natural language processing in EHR in relation to healthcare decision-making: A systematic review," Comput. Biol. Med., vol. 155, p. 106649, Mar. 2023. doi: 10.1016/j.compbiomed.2023.106649.

[10] C. Zhang and Y. Lu, "Study on artificial intelligence: The state of the art and future prospects," J. Industrial Inform. Integration, vol. 23, p. 100224, Sep. 2021. doi: 10.1016/j.jii.2021.100224.

[11] Stanford HAI, "AI Index Report 2025," Stanford Univ., Apr. 2025. Available: https://hai.stanford.edu/ai-index/2025-ai-index-report.

[12] E. J. Topol, "High-performance medicine: The convergence of human and artificial intelligence," Nature Med., vol. 25, no. 1, pp. 44–56, Jan. 2019. doi: 10.1038/s41591-018-0300-7.

[13] P. Rajpurkar, E. Chen, O. Banerjee, and E. J. Topol, "AI in health and medicine," Nature Med., vol. 28, no. 1, pp. 31–38, Jan. 2022. doi: 10.1038/s41591-021-01614-0.

[14] P. Elias et al., "AI for cardiovascular care—Part 1: Advances," J. Amer. Coll. Cardiol., vol. 83, no. 24, pp. 2472–2486, Jun. 2024. doi: 10.1016/j.jacc.2024.03.400.

[15] G. Savarese et al., "Global burden of heart failure: A comprehensive and updated review of epidemiology," Cardiovasc. Res., vol. 118, no. 17, pp. 3272–3287, Dec. 2022. doi: 10.1093/cvr/cvac013.

[16] N. L. Bragazzi et al., "Burden of heart failure and underlying causes in 195 countries 1990 to 2017," Eur. J. Prev. Cardiol., vol. 28, no. 15, pp. 1682–1690, 2021. doi: 10.1093/eurjpc/zwaa147.

[17] Z. Zheng et al., "An overview of blockchain technology: Architecture, consensus, and future trends," in Proc. IEEE Int. Congress Big Data, 2017, pp. 557–564. doi: 10.1109/BigDataCongress.2017.85.

[18] V. Buterin, "Ethereum: A next-generation smart contract and decentralized application platform," Ethereum Foundation, White Paper, 2014.

[19] C. C. Agbo, Q. H. Mahmoud, and J. M. Eklund, "Blockchain technology in healthcare: A systematic review," Healthcare, vol. 7, no. 2, art. 56, Apr. 2019. doi: 10.3390/healthcare7020056.

[20] S. J. Pocock et al., "Predicting survival in heart failure: A risk score based on 39 372 patients from 30 studies," Eur. Heart J., vol. 34, no. 19, pp. 1404–1413, May 2013. doi: 10.1093/eurheartj/ehs337.

[21] Z. Sun et al., "Comparing ML models and statistical models for predicting heart failure events: A systematic review and meta-analysis," Frontiers Cardiovasc. Med., vol. 9, 812276, Mar. 2022. doi: 10.3389/fcvm.2022.812276.

[22] C. Rudin, "Stop explaining black box ML models for high stakes decisions and use interpretable models instead," Nature Machine Intell., vol. 1, no. 5, pp. 206–215, May 2019. doi: 10.1038/s42256-019-0048-x.

[23] IBM Security and Ponemon Institute, "Cost of a Data Breach Report 2024," IBM Corp., Jul. 2024.

[24] European Parliament and Council of the EU, "Regulation (EU) 2016/679 (GDPR)," Official Journal of the EU, L 119, pp. 1–88, May 4, 2016.

[25] A. E. W. Johnson et al., "MIMIC-IV, a freely accessible EHR dataset," Scientific Data, vol. 10, no. 1, p. 1, Jan. 2023. doi: 10.1038/s41597-022-01899-x.

[26] T. A. McDonagh et al., "2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure," Eur. Heart J., vol. 42, no. 36, pp. 3599–3726, Sep. 2021. doi: 10.1093/eurheartj/ehab368.

[27] T. A. McDonagh et al., "2023 Focused Update of the 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure," Eur. Heart J., vol. 44, no. 37, pp. 3627–3639, Oct. 2023. doi: 10.1093/eurheartj/ehad195.

[28] P. A. Heidenreich et al., "2022 AHA/ACC/HFSA Guideline for the Management of Heart Failure," Circulation, vol. 145, no. 18, pp. e895–e1032, May 2022. doi: 10.1161/CIR.0000000000001063.

[29] B. Shahim, C. J. Kapelios, G. Savarese, and L. H. Lund, "Global public health burden of heart failure: An updated review," Cardiac Failure Rev., vol. 9, e11, Jul. 2023. doi: 10.15420/cfr.2023.05.

[30] R. Aimo et al., "Biomarkers for the diagnosis and management of heart failure," Heart Failure Rev., vol. 27, pp. 1–16, 2022. doi: 10.1007/s10741-021-10105-w

[31] S. J. Pocock et al., "Predicting survival in heart failure: A risk score based on 39 372 patients from 30 studies," Eur. Heart J., vol. 34, no. 19, pp. 1404–1413, May 2013. doi: 10.1093/eurheartj/ehs337.

[32] Warriner et al., 2017 “Computer model for the cardiovascular system: development of an e-learning tool for teaching of medical students”. doi: https://doi.org/10.1186/s12909-017-1058-1

[33] Gianluigi Pironti “State-of-the-art methodologies used in preclinical studies to assess left ventricular diastolic and systolic function in mice, pitfalls and troobleshooting” doi:10.3389/fcvm.2023.1228789

[34] Castiglione “Biomarkers for the diagnosis and management of heart failure”. doi: https://doi.org/10.1007/s10741-021-10105-w

[35] Moreno-Sanchez et al. “Improvement of a prediction model for heart failure survival though explainable artificial intelligence” doi: 10.3389/fcvm.2023/1219586

[36] Saqib et al. “Machine Learning in Heart Failure Diagnosis, Prediction and Prognosis: Review” doi: 10.1097/ms9.0000000000002138

[37] Gao et al. “SGLT2 Inhibitors, Functional Capacity, and Quality of Life in Patients With Heart Failure: A Systematic Review and Meta-Analysis” doi: 10.1001/jamanetworkopen.2024.5135

[38] Mpanya et al. “Predicting mortality and hospitalization in heart failure using machine learning: A systematic literature review” doi: 10.1016/j.ijcha.2021.100773

[39] Jiang et al. “Artifical Intelligence in Healthcare: Past, Present and Future”, doi: 10.1136/svn-2017-000101

[40] R. J. Desai et al., "Comparison of ML methods with traditional models to predict HF outcomes," JAMA Netw. Open, vol. 3, no. 1, e1918962, 2020. doi: 10.1001/jamanetworkopen.2019.18962.

[41] Jelle P. Man “Digital Solutions to Optimize Guideline-Directed Medical Therapy Prescriptions in Heart Failure Patients: Current Applications and Future Directions” doi: 10.1007/s11897-024-00649-x

[42]S. Kundu, "AI in medicine must be explainable," Nature Med., vol. 27, p. 1328, 2021. doi: 10.1038/s41591-021-01461-z.

[43] N. Bienefeld et al., "Solving the explainable AI conundrum by bridging clinicians' needs and developers' goals," npj Digital Med., vol. 6, 94, 2023. doi: 10.1038/s41746-023-00837-4.

[44] A. Chaddad et al., "A review of Explainable Artificial Intelligence in healthcare," Comput. Electr. Eng., vol. 118, 109370, 2024. doi: 10.1016/j.compeleceng.2024.109370

[45] F. Noor et al., "Unveiling Explainable AI in Healthcare: Current Trends, Challenges, and Future Directions," WIREs Data Mining and Knowledge Discovery, vol. 15, e70018, 2025. doi: 10.1002/widm.70018

[46] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in NeurIPS 2017, pp. 4766–4777.

[47] S. Hur et al., "Comparison of SHAP and clinician-friendly explanations reveals effects on clinical decision behaviour," npj Digital Med., 2025. doi: 10.1038/s41746-025-01958-8.

[48] R. Rosenbacke et al., "How XAI can increase or decrease clinicians' trust in AI applications in health care: Systematic review," JMIR AI, vol. 3, e53207, Oct. 2024. doi: 10.2196/53207

[49] Abbas et al., “Explainable AI in Clinical Decision Support Systems: A Meta-Analysis of Methods, Applications and Usability Challenges” doi: 10.3390/healthcare13172154

[50] Zheng et al. “An Overview of Blockchain Technlogy: Architecture, Consensus and Future Trends” doi: 10.1109/BigDataCongress.2017.85

[51] P. Durneva, K. Cousins, and M. Chen, "The current state of research of blockchain technology in patient care: Systematic review," J. Med. Internet Res., vol. 22, no. 7, e18619, Jul. 2020. doi: 10.2196/18619.

[52] T. Ettaloui, A. Lachhab, and B. Ouchetto, "Blockchain-Based Electronic Health Record: Systematic Literature Review," Human Behavior and Emerging Technologies, vol. 2024, art. 4734288, 2024. doi: 10.1155/hbe2/4734288.

[53] S. Rantos, G. Drosatos, K. Demertzis, C. Ilioudis, A. Papanikolaou, and A. Kritsas, "A systematic review of blockchain for consent management," Healthcare, vol. 7, no. 2, 2019. doi: 10.3390/healthcare7020013. (PMC systematic review on blockchain consent management)

[54] Various, "Smart Contracts, Blockchain, and Health Policies: Past, Present, and Future," Information, vol. 16, no. 10, art. 853, Oct. 2025. doi: 10.3390/info16100853.

[55] M. Shaikh et al., "A systematic literature review for blockchain-based healthcare implementations," Healthcare, vol. 13, no. 9, art. 1087, May 2025. doi: 10.3390/healthcare13091087.

[56] A. Bathula et al., "Blockchain, AI, and healthcare: The tripod of future—A narrative review," Artif. Intell. Rev., vol. 57, no. 9, art. 238, 2024. doi: 10.1007/s10462-024-10873-5.

[57] T. McGhin, K.-K. R. Choo, C. Z. Liu, and D. He, "Blockchain in healthcare applications: Research challenges and opportunities," J. Netw. Comput. Appl., vol. 135, pp. 62–75, Jun. 2019. doi: 10.1016/j.jnca.2019.02.027

[58] A. Hasselgren, K. Kralevska, D. Gligoroski, S. A. Pedersen, and A. Faxvaag, "Blockchain in healthcare and health sciences—A scoping review," Int. J. Med. Inform., vol. 134, art. 104040, Feb. 2020. doi: 10.1016/j.ijmedinf.2019.104040

[59] R. T. Sutton, D. Pincock, D. C. Baumgart, D. C. Sadowski, R. N. Fedorak, and K. I. Kroeker, "An overview of clinical decision support systems: benefits, risks, and strategies for success," npj Digital Med., vol. 3, art. 17, Feb. 2020. doi: 10.1038/s41746-020-0221-y

[60] W. Chen et al., "Design, effectiveness, and economic outcomes of contemporary chronic disease clinical decision support systems: a systematic review and meta-analysis," J. Amer. Med. Inform. Assoc., vol. 29, no. 10, pp. 1757–1772, Oct. 2022. doi: 10.1093/jamia/ocac110

[61] P.-Y. Meunier, C. Raynaud, E. Guimaraes, F. Gueyffier, and L. Letrilliart, "Barriers and facilitators to the use of clinical decision support systems in primary care: A mixed-methods systematic review," Ann. Fam. Med., vol. 21, no. 1, pp. 57–69, Jan. 2023. doi: 10.1370/afm.2908

[62] A. P. Susanto, D. Lyell, B. Widyantoro, S. Berkovsky, and F. Magrabi, "Effects of machine learning-based clinical decision support systems on decision-making, care delivery, and patient outcomes: A scoping review," J. Amer. Med. Inform. Assoc., vol. 30, no. 12, pp. 2050–2063, Dec. 2023. doi: 10.1093/jamia/ocad180

[63] H. R. Strasberg, B. Rhodes, G. Del Fiol, R. A. Jenders, P. J. Haug, and K. Kawamoto, "Contemporary clinical decision support standards using Health Level Seven International Fast Healthcare Interoperability Resources," J. Amer. Med. Inform. Assoc., vol. 28, no. 8, pp. 1796–1806, Aug. 2021. doi: 10.1093/jamia/ocab070

[64] P. Taber, C. Radloff, G. Del Fiol, C. Staes, and K. Kawamoto, "New standards for clinical decision support: A survey of the state of implementation," Yearbook Med. Inform., vol. 30, no. 1, pp. 159–171, Aug. 2021. doi: 10.1055/s-0041-1726502.

[65] A. Muneer et al., "A chatbot based on ML for risk prediction of CVD using a validated dataset," Diagnostics, vol. 14, no. 4, art. 445, 2024.

[66] Li et al. “XGBoost for Predictiong In-Hospital Mortality in Patients with Acute Heart Failure”, https://doi.org/10.1002/ehf2.13887

[67] https://en.wikipedia.org/wiki/Optuna