



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ
ΥΛΙΚΩΝ

«Μοντελοποίηση Παιδικής
Παχυσαρκίας με χρήση
Ερμηνεύσιμης Τεχνητής
Νοημοσύνης»

Διπλωματική εργασία

Του φοιτητή **Βόγκα Δημητρίου**

Επιβλέπων: **καθ. Κα Κωνσταντίνα Νικήτα**

Συμμετοχή στην επίβλεψη: **Δρ. Κωνσταντία Ζαρκογιάννη**

Αθήνα, Ιούλιος 2023



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ
ΥΛΙΚΩΝ

«Μοντελοποίηση Παιδικής
Παχυσαρκίας με χρήση
Ερμηνεύσιμης Τεχνητής
Νοημοσύνης»

Διπλωματική εργασία

Του φοιτητή **Βόγκα Δημητρίου**

Επιβλέπων: **καθ. Κα Κωνσταντίνα Νικήτα**

Συμμετοχή στην επίβλεψη: **Δρ. Κωνσταντία Ζαρκογιάννη**

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την

(Υπογραφή)

(Υπογραφή)

(Υπογραφή)

.....

.....

.....

Κωνσταντίνα Νικήτα

Γεώργιος Στάμου

Αθανάσιος Βουλόδημος

Καθηγήτρια Ε.Μ.Π.

Καθηγητής Ε.Μ.Π.

Επίκουρος Καθηγητής Ε.Μ.Π.

Αθήνα, 13^η Ιουλίου 2023

.....

Δημήτριος Β. Βόγκας

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

Copyright © Δημήτριος Βόγκας, 2023

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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

Περίληψη

Η παιδική παχυσαρκία αποτελεί προδιαθεσικό παράγοντα για την εμφάνιση χρόνιων νοσημάτων (π.χ. σακχαρώδης διαβήτης, καρδιαγγειοπάθεια κ.ά.) μετά την ενηλικίωση, ενώ προκαλεί επιπτώσεις και στην ψυχική υγεία. Συνεκτιμώντας και τη μεγάλη εξάπλωσή της σε παγκόσμιο επίπεδο, καθίσταται αναγκαία η εύρεση αποτελεσματικών παρεμβάσεων για τη διαχείρισή της. Η μεγάλη πρόκληση που έχουν να αντιμετωπίσουν τα πλάνα παρέμβασης είναι η συμμόρφωση στις ιατρικές συστάσεις. Για αυτόν τον λόγο, προτείνεται η χρήση τεχνολογιών κινητής υγείας και τεχνητής νοημοσύνης. Στο πλαίσιο της παρούσας διπλωματικής, διερευνήθηκαν τεχνικές ερμηνεύσιμης τεχνητής νοημοσύνης και ενισχυτικής μάθησης τόσο για την διαστρωμάτωση των παιδιών ως προς το βαθμό συμμόρφωσης στο πλάνο θεραπείας, όσο και για την παραγωγή εξατομικευμένου περιεχομένου για την μετέπειτα προβολή του στους τελικούς χρήστες υπολογιστικών οικοσυστημάτων διαχείρισης της παιδικής παχυσαρκίας. Πιο συγκεκριμένα, στο πρώτο μέρος της εργασίας αναπτύσσονται μοντέλα επιβλεπόμενης μηχανικής μάθησης βασισμένης σε δέντρα αποφάσεων για την πρόβλεψη του βαθμού συμμόρφωσης χρησιμοποιώντας δεδομένα πρώτης επίσκεψης, κατά την έναρξη του πλάνου παρέμβασης. Στο δεύτερο σκέλος, χρησιμοποιείται ενισχυτική μάθηση για την εκμάθηση της βέλτιστης πολιτικής αποστολής προτρεπτικών και εκπαιδευτικών μηνυμάτων με στόχο τη μεγιστοποίηση της συμμόρφωσης και τη βελτιστοποίηση της έκβασης της θεραπείας. Σε όλη την έκταση της εργασίας, δίνεται έμφαση στην ερμηνευσιμότητα των εφαρμοζόμενων τεχνικών μηχανικής μάθησης.

Λέξεις Κλειδιά

Παιδική Παχυσαρκία, Πλάνο Παρέμβασης, Συμμόρφωση, Μηχανική Μάθηση, Επιβλεπόμενη Μάθηση Βασισμένη σε Δέντρα, Ενισχυτική Μάθηση, Σύστημα Συστάσεων

Abstract

Childhood obesity constitutes a risk factor for the occurrence of chronic diseases (e.g. diabetes mellitus, cardiovascular diseases, etc.) in adulthood, while having negative consequences on mental health. Considering its global prevalence, it is imperative, also provoking, to develop effective interventions for managing childhood obesity. A great challenge that needs to be faced constitutes the adherence to the medical recommendations. Therefore, the use of mobile health technologies and artificial intelligence is recommended. In the context of the present diploma thesis, explainable artificial intelligence and reinforcement learning techniques have been investigated towards stratifying obese children according to their degree of adherence to the treatment plan, as well as producing personalized content to be presented to the end users of software ecosystems focusing on childhood obesity management. More specifically, tree-based supervised machine learning models for predicting the level of adherence to the treatment are developed using data at baseline. Also, reinforcement learning has been explored for learning the optimal policy of sending rewarding and educational messages to maximize adherence and optimizing the treatment outcome. Throughout the thesis, the interpretability of the applied machine learning techniques is in depth investigated.

Keywords

Childhood Obesity, Intervention Plan, Adherence, Machine Learning, Tree-based Supervised Learning, Reinforcement Learning, Recommendation System

Ευχαριστίες

Θα ήθελα να ευχαριστήσω την επιβλέπουσα καθηγήτρια κα. Κωνσταντίνα Νικήτα η οποία μου έδειξε εμπιστοσύνη και μου έδωσε την ευκαιρία να εκπονήσω τη διπλωματική μου εργασία στο Εργαστήριο Βιοϊατρικών Προσομοιώσεων και Απεικονιστικής Τεχνολογίας, πάνω σε ένα τόσο ενδιαφέρον θέμα.

Επίσης, θα ήθελα να ευχαριστήσω ιδιαίτερα τη διδάκτορα κα. Κωνσταντία Ζαρκογιάννη που από την πρώτη μέρα με βοήθησε και με καθοδήγησε σε όλα τα στάδια εκπόνησης της διπλωματικής. Ο φετινός υπήρξε ο παραγωγικότερος χρόνος για μένα στο Πολυτεχνείο όσον αφορά τις γνώσεις που αποκόμισα και το πόσο εξελίχθηκα ως μηχανικός και δεν θα τα είχα καταφέρει χωρίς την υποστήριξή της και την πίστη της στις δυνατότητές μου.

Τέλος, θα ήθελα να ευχαριστήσω τη Χρυσομάλλη και τους γονείς μου που ήταν δίπλα μου σε όλη αυτή την προσπάθεια.

0. ΠΕΡΙΕΧΟΜΕΝΑ

Πίνακας Περιεχομένων Πινάκων	12
1. Πρόλογος.....	14
2. Παιδική Παχυσαρκία.....	16
2.1. Αίτια, Επιπτώσεις και Στατιστικά Στοιχεία	16
2.2. Αντιμετώπιση / Διαχείριση	18
3. Επιστημονικό Υπόβαθρο	21
3.1. Τεχνολογίες Επιβλεπόμενης Μάθησης.....	21
3.2. Ερμηνεύσιμη Τεχνητή Νοημοσύνη και Επίπεδο Βεβαιότητας	24
3.3. Ενισχυτική Μάθηση	28
4. Ανάπτυξη Μοντέλου Εκτίμησης της Συμμόρφωσης στο Πλάνο Παρέμβασης	31
4.1. Μεθοδολογία	31
4.1.1. Δεδομένα.....	31
4.1.2. Συμμόρφωση στο Πλάνο Παρέμβασης.....	32
4.1.3. Περιγραφή της Διαδικασίας.....	34
4.1.4. Διαδικασία Εκπαίδευσης και Αξιολόγησης.....	35
4.1.5. Επίπεδο Βεβαιότητας	36
4.1.6. Ερμηνευσιμότητα	38
4.2. Αποτελέσματα και Συζήτηση.....	40
4.2.1. Πρόβλεψη Συμμόρφωσης με XGBoost	40
4.2.2. Πρόβλεψη Συμμόρφωσης με Optimal Decision Trees.....	64
4.2.3. Ανίχνευση Προκαταλήψεων στα Δεδομένα	87
5. Ανάπτυξη Συστήματος Παραγωγής Εξατομικευμένου Περιεχομένου	91
5.1. Μεθοδολογία	91
5.1.1. Δεδομένα.....	91
5.1.2. Επιλογή του αλγορίθμου.....	92
5.1.3. Διατύπωση του προβλήματος.....	94
5.1.4. Περιγραφή της διαδικασίας εκπαίδευσης.....	97
5.1.5. Ιεραρχική ενισχυτική μάθηση για μείωση του χώρου ενεργειών	99
5.1.6. Απόσταση στον χώρο ενεργειών.....	101
5.2. Αποτελέσματα	101
5.2.1. Αρχικός χώρος ενεργειών.....	102
5.2.2. Ιεραρχική προσέγγιση	104
6. Συμπεράσματα – Μελλοντικές Επεκτάσεις	107
7. Βιβλιογραφία.....	110

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ ΠΙΝΑΚΩΝ

Πίνακας 4.1: Αντιστοίχιση χρήσης εφαρμογής κινητού με επίπεδο συμμόρφωσης.	33
Πίνακας 4.2: Αντιστοίχιση χρήσης ιχνηλάτη δραστηριότητας (Fitbit) με επίπεδο συμμόρφωσης.	33
Πίνακας 4.3: Περιγραφή των χαρακτηριστικών του συνόλου δεδομένων με τα οποία εκπαιδεύτηκε το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με XGBoost.	41
Πίνακας 4.4: Μετρικές αξιολόγησης του μοντέλου πρόβλεψης της συνολικής συμμόρφωσης με XGBoost.	42
Πίνακας 4.5: Επεξήγηση των swarm plots για το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με XGBoost.	45
Πίνακας 4.6: Περιγραφή των χαρακτηριστικών του συνόλου δεδομένων με τα οποία εκπαιδεύτηκε το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με XGBoost. ...	49
Πίνακας 4.7: Μετρικές αξιολόγησης του μοντέλου πρόβλεψης της συμμόρφωσης με την εφαρμογή με XGBoost.	51
Πίνακας 4.8: Επεξήγηση των swarm plots για το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με XGBoost.	53
Πίνακας 4.9: Περιγραφή των χαρακτηριστικών του συνόλου δεδομένων με τα οποία εκπαιδεύτηκε το μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με XGBoost.	59
Πίνακας 4.10: Μετρικές αξιολόγησης του μοντέλου πρόβλεψης της συμμόρφωσης με το Fitbit με XGBoost.	60
Πίνακας 4.11: Επεξήγηση των swarm plots για το μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με XGBoost.	61
Πίνακας 4.12: Περιγραφή των χαρακτηριστικών του συνόλου δεδομένων με τα οποία εκπαιδεύτηκε το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με Optimal Decision Trees.	66
Πίνακας 4.13: Μετρικές αξιολόγησης του μοντέλου πρόβλεψης της συνολικής συμμόρφωσης με Optimal Decision Trees.	68
Πίνακας 4.14: Επεξήγηση των swarm plots για το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με Optimal Decision Trees.	71
Πίνακας 4.15: Περιγραφή των χαρακτηριστικών του συνόλου δεδομένων με τα οποία εκπαιδεύτηκε το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με Optimal Decision Trees.	76
Πίνακας 4.16: Μετρικές αξιολόγησης του μοντέλου πρόβλεψης της συμμόρφωσης με την εφαρμογή με Optimal Decision Trees.	77
Πίνακας 4.17: Επεξήγηση των swarm plots για το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με Optimal Decision Trees.	79
Πίνακας 4.18: Περιγραφή των χαρακτηριστικών του συνόλου δεδομένων με τα οποία εκπαιδεύτηκε το μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με Optimal Decision Trees.	83
Πίνακας 4.19: Μετρικές αξιολόγησης του μοντέλου πρόβλεψης της συμμόρφωσης με το Fitbit με Optimal Decision Trees.	83
Πίνακας 4.20: Επεξήγηση των swarm plots για το μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με Optimal Decision Trees.	85

Πίνακας 5.1: Περιγραφή των χαρακτηριστικών του συνόλου δεδομένων τα οποία συνιστούν την κατάσταση σε κάθε χρονική στιγμή για τα μοντέλα συστάσεων αποστολής μηνυμάτων.	95
Πίνακας 5.2: Αξιολόγηση των πολιτικών σε επίπεδο χρηστών για τα μοντέλα του συστήματος συστάσεων όταν ακολουθήθηκε η προσέγγιση με τον αρχικό χώρο ενεργειών. Οι τιμές συσχέτισης με έντονη γραμματοσειρά ήταν οι μέγιστες για τη συγκεκριμένη βαρύτητα συμμόρφωσης μεταξύ των διαφορετικών προσεγγίσεων.	102
Πίνακας 5.3: Συγκεντρωτικός πίνακας για την ομαδοποίηση των δειγμάτων σε κάδους για την προσέγγιση με αναπαράσταση των δεδομένων στον αρχικό χώρο ενεργειών και βαρύτητα της συμμόρφωσης στη συνάρτηση ανταμοιβής 70%. π.κ.: πλήθος κάδων, α.κ.: αριθμός κάδων, μ.α.: μέση ανταμοιβή, π.δ.: πλήθος δειγμάτων.	103
Πίνακας 5.4: Συγκεντρωτικός πίνακας για την ομαδοποίηση των δειγμάτων σε κάδους για την προσέγγιση με αναπαράσταση των δεδομένων στον αρχικό χώρο ενεργειών και βαρύτητα της συμμόρφωσης στη συνάρτηση ανταμοιβής 80%. π.κ.: πλήθος κάδων, α.κ.: αριθμός κάδων, μ.α.: μέση ανταμοιβή, π.δ.: πλήθος δειγμάτων.	103
Πίνακας 5.5: Συγκεντρωτικός πίνακας για την ομαδοποίηση των δειγμάτων σε κάδους για την προσέγγιση με αναπαράσταση των δεδομένων στον αρχικό χώρο ενεργειών και βαρύτητα της συμμόρφωσης στη συνάρτηση ανταμοιβής 100%. π.κ.: πλήθος κάδων, α.κ.: αριθμός κάδων, μ.α.: μέση ανταμοιβή, π.δ.: πλήθος δειγμάτων.	104
Πίνακας 5.6: Αξιολόγηση των πολιτικών σε επίπεδο χρηστών για τα μοντέλα του συστήματος συστάσεων όταν ακολουθήθηκε η ιεραρχική προσέγγιση. Οι τιμές συσχέτισης με έντονη γραμματοσειρά ήταν οι μέγιστες για τη συγκεκριμένη βαρύτητα συμμόρφωσης μεταξύ των διαφορετικών προσεγγίσεων.	104
Πίνακας 5.7: Συγκεντρωτικός πίνακας για την ομαδοποίηση των δειγμάτων σε κάδους για την ιεραρχική προσέγγιση και βαρύτητα της συμμόρφωσης στη συνάρτηση ανταμοιβής 70%. π.κ.: πλήθος κάδων, α.κ.: αριθμός κάδων, μ.α.: μέση ανταμοιβή, π.δ.: πλήθος δειγμάτων.....	105
Πίνακας 5.8: Συγκεντρωτικός πίνακας για την ομαδοποίηση των δειγμάτων σε κάδους για την ιεραρχική προσέγγιση και βαρύτητα της συμμόρφωσης στη συνάρτηση ανταμοιβής 80%. π.κ.: πλήθος κάδων, α.κ.: αριθμός κάδων, μ.α.: μέση ανταμοιβή, π.δ.: πλήθος δειγμάτων.....	105
Πίνακας 5.9: Συγκεντρωτικός πίνακας για την ομαδοποίηση των δειγμάτων σε κάδους για την ιεραρχική προσέγγιση και βαρύτητα της συμμόρφωσης στη συνάρτηση ανταμοιβής 100%. π.κ.: πλήθος κάδων, α.κ.: αριθμός κάδων, μ.α.: μέση ανταμοιβή, π.δ.: πλήθος δειγμάτων.....	106

1. ΠΡΟΛΟΓΟΣ

Η παιδική παχυσαρκία αποτελεί μία από τις πλέον σοβαρές προκλήσεις για τη δημόσια υγεία στον 21ο αιώνα.¹ Υπολογίζεται ότι το 2017, 124 εκατομμύρια παιδιά και έφηβοι στον κόσμο κατατάσσονταν ως παχύσαρκα. Τα τελευταία σαράντα χρόνια, το ποσοστό των παχύσαρκων παιδιών και εφήβων παγκοσμίως έχει οκταπλασιαστεί [1].

Παιδιά με παχυσαρκία έχουν περισσότερες πιθανότητες να γίνουν παχύσαρκοι ενήλικες σε σύγκριση με τα παιδιά φυσιολογικού βάρους [2]. Τα παχύσαρκα παιδιά και οι έφηβοι αντιμετωπίζουν αυξημένο κίνδυνο για ασθένειες σχετιζόμενες με την παχυσαρκία μόνο αν παραμείνουν παχύσαρκα και ως ενήλικες [3]. Επομένως, η έγκαιρη αντιμετώπιση της παχυσαρκίας κατά την παιδική ή εφηβική ηλικία μπορεί να προσδώσει πολλαπλά οφέλη στους ασθενείς και να μειώσει την πιθανότητα επιπλοκών.

Για την αντιμετώπιση της παιδικής παχυσαρκίας προτείνονται προγράμματα ρύθμισης βάρους για παιδιά που ενσωματώνουν τεχνολογικά χαρακτηριστικά και απευθύνονται τόσο στο παιδί, όσο και στους γονείς του. Τα υπολογιστικά συστήματα υποστήριξης αποφάσεων και οι τεχνολογίες κινητής υγείας μπορούν να παρέχουν πολλές δυνατότητες για να ενισχύσουν πλάνα παρέμβασης [4]–[9]. Οι μέθοδοι της τεχνητής νοημοσύνης και της μηχανικής μάθησης προσφέρονται για την μοντελοποίηση του πολυπαραγοντικού και σύνθετου προβλήματος της παιδικής παχυσαρκίας, λόγω της ικανότητάς τους να ανακαλύπτουν χρήσιμες πληροφορίες μέσα από εκτενή και περίπλοκα σύνολα δεδομένων. Τεχνολογίες όπως οι εφαρμογές κινητών συσκευών και οι ιχνηλάτες σωματικής δραστηριότητας μπορούν να συμβάλουν στην αποτελεσματική παρακολούθηση και καταγραφή της τήρησης των πλάνων παρέμβασης, να καταστήσουν ευκολότερη και πιο λειτουργική την επικοινωνία των ασθενών με τις κλινικές ομάδες και να μειώσουν τον φόρτο εργασίας των τελευταίων.

Η παρούσα εργασία κινείται σε δύο βασικούς άξονες. Ο πρώτος άξονας αφορά την ανάπτυξη μοντέλων μηχανικής μάθησης βασισμένα σε δέντρα αποφάσεων για την εκτίμηση της συμμόρφωσης χρησιμοποιώντας δημογραφικά δεδομένα, απαντήσεις σε ερωτηματολόγια, καθώς και κλινικούς και βιοχημικούς δείκτες. Ο δεύτερος εστιάζει στη χρήση τεχνικών ενισχυτικής μάθησης (Reinforcement Learning – RL) με στόχο την επιλογή των παροτρυντικών μηνυμάτων για τη βελτιστοποίηση της συμμόρφωσης και της έκβασης της θεραπείας για κάθε χρήστη.

Τα δεδομένα που αξιοποιούνται στην παρούσα εργασία συλλέχθηκαν από τις προπilotικές και πιλοτικές δοκιμές του πλάνου παρέμβασης ENDORSE [10]–[13], ενός προγράμματος διάρκειας 12 εβδομάδων το οποίο περιλάμβανε τη χρήση μιας εφαρμογής κινητού, ενός ιχνηλάτη σωματικής δραστηριότητας και ενός παιχνιδιού σοβαρού σκοπού για την υποστήριξη της τήρησης του πλάνου διατροφής και άσκησης. Τα δεδομένα προέρχονται από ένα σχετικά μικρό δείγμα παιδιών για τα οποία, ωστόσο, είναι διαθέσιμο μεγάλο πλήθος πληροφοριών. Πιθανώς, τα αποτελέσματα αυτής της προσπάθειας να μπορέσουν να αναδείξουν παράγοντες της συμμόρφωσης σε προγράμματα ρύθμισης βάρους, ενισχύοντας την έρευνα σε αυτόν τον τομέα. Ταυτόχρονα, ένα μοντέλο πρόβλεψης της συμμετοχής σε ένα πλάνο παρέμβασης δύναται να είναι χρήσιμο, ούτως ώστε να βοηθηθούν πιο αποτελεσματικά τα παιδιά με περισσότερες πιθανότητες χαμηλής συμμόρφωσης είτε με

¹ <https://www.who.int/news-room/questions-and-answers/item/noncommunicable-diseases-childhood-overweight-and-obesity>

μεγαλύτερη προσοχή από την κλινική ομάδα είτε ακολουθώντας ένα διαφορετικό πλάνο παρέμβασης. Επιπρόσθετα, τα αποτελέσματα της διαδικασίας εύρεσης της βέλτιστης πολιτικής αποστολής μηνυμάτων, μπορούν να αξιοποιηθούν από παρεμφερείς εφαρμογές με στόχο την αύξηση της συμμόρφωσης και της αποτελεσματικότητας της θεραπείας, αλλά κυρίως ως μεθοδολογία για την εκπαίδευση συστημάτων συστάσεων (Recommendation Systems) αποστολής μηνυμάτων για αντίστοιχες εφαρμογές.

Η εκτίμηση της συμμόρφωσης στην παρούσα μελέτη βασίστηκε στον βαθμό χρησιμοποίησης της εφαρμογής και του ιχνηλάτη σωματικής δραστηριότητας.

Στο Κεφάλαιο 2 πραγματοποιείται μία σύντομη βιβλιογραφική ανασκόπηση πάνω στην παιδική παχυσαρκία, τα πλάνα παρέμβασης για την αντιμετώπισή της και τις εφαρμογές της μηχανικής μάθησης στο συγκεκριμένο πεδίο. Αναφέρονται τα βασικότερα αίτια και οι επιπτώσεις της παιδικής παχυσαρκίας στην υγεία και γενικά στη ζωή ενός παιδιού. Επίσης, παρέχεται μία περιγραφή των πλάνων παρέμβασης βασισμένων σε τεχνολογίες κινητής υγείας που έχουν δημοσιευτεί και αναλύονται ορισμένες σημαντικές εφαρμογές μηχανικής μάθησης για την αντιμετώπιση της παιδικής παχυσαρκίας.

Το τεχνολογικό υπόβαθρο που χρησιμοποιήθηκε στην παρούσα εργασία παρουσιάζεται στο Κεφάλαιο 3. Συγκεκριμένα, γίνεται μια εισαγωγή στις τεχνολογίες της επιβλεπόμενης μάθησης που εφαρμόστηκαν στην παρούσα εργασία, αναλύονται οι έννοιες της ερμηνευσιμότητας μοντέλων και ποσοτικοποίησης της βεβαιότητας και επεξηγούνται τα βασικά της ενισχυτικής μάθησης και του αλγορίθμου Q-learning ο οποίος αποτελεί τον πυρήνα του συστήματος παραγωγής εξατομικευμένου περιεχομένου.

Η μεθοδολογία της διαδικασίας ανάπτυξης μοντέλων μηχανικής μάθησης για την πρόβλεψη της συμμόρφωσης στις ιατρικές συστάσεις, καθώς και τα αποτελέσματά της παρουσιάζονται στο Κεφάλαιο 4. Μετά την περιγραφή των δεδομένων, αναλύονται η διαδικασία εκπαίδευσης και αξιολόγησης των μοντέλων πρόβλεψης της συμμόρφωσης, καθώς και της μεθοδολογίας ερμηνείας και ποσοτικοποίησης της βεβαιότητας των προβλέψεων. Ακόμη, παρέχονται τα αποτελέσματα της διαδικασίας εκπαίδευσης μοντέλων, καθώς και εκτενείς εξηγήσεις για τη μεθοδολογία και σχολιασμοί των ευρημάτων.

Η διαδικασία ανάπτυξης συστήματος παραγωγής εξατομικευμένου περιεχομένου με χρήση ενισχυτικής μάθησης παρουσιάζεται στο Κεφάλαιο 5. Αναλύονται διεξοδικά τα δεδομένα, η μεθοδολογία που ακολουθήθηκε για τη διαδικασία εξαγωγής και αξιολόγησης των πολιτικών αποστολής μηνυμάτων, καθώς και το σκεπτικό πίσω από τις σχετικές αποφάσεις. Παρατίθενται και σχολιάζονται τα αποτελέσματα της διαδικασίας αξιολόγησης.

Στο Κεφάλαιο 6, τονίζονται τα συμπεράσματα που προέκυψαν στο πλαίσιο της παρούσας διπλωματικής εργασίας. Επίσης, αναφέρονται ορισμένοι περιορισμοί και προτείνονται μελλοντικές επεκτάσεις για τη βελτίωση του έργου.

Τέλος, στο Κεφάλαιο 7, παρατίθεται η σχετική βιβλιογραφία.

2. ΠΑΙΔΙΚΗ ΠΑΧΥΣΑΡΚΙΑ

Η υπερβαρότητα και η παχυσαρκία ορίζονται ως «αφύσικη ή υπερβολική συσσώρευση λίπους που παρουσιάζει κίνδυνο για την υγεία». Μέτρο της παχυσαρκίας στην παιδική ηλικία είναι το BMI z-score το οποίο δείχνει την εκατοστιαία θέση (ΕΘ) του δείκτη μάζας σώματος (body mass index – BMI) ενός παιδιού σε σχέση με τους δείκτες μάζας σώματος του πληθυσμού των παιδιών με ίδιο φύλο και ηλικία. Συγκεκριμένα, για παιδιά άνω των 2 ετών, αν έχουν BMI άνω της 85^{ης} ΕΘ και κάτω της 95^{ης} τότε κατατάσσονται ως υπέρβαρα, με BMI άνω της 95^{ης} ΕΘ κατατάσσονται ως παχύσαρκα, ενώ με BMI άνω της 99^{ης} κατατάσσονται ως σοβαρά παχύσαρκα [14].

Στις επόμενες υποενότητες αναφέρονται τα σημαντικότερα αίτια και οι επιπτώσεις της παιδικής παχυσαρκίας, κάποια σχετικά στατιστικά, καθώς και οι τρόποι με τους οποίους οι τεχνολογίες κινητής υγείας και τεχνητής νοημοσύνης μπορούν να συμβάλουν στην αντιμετώπισή της.

2.1. ΑΙΤΙΑ, ΕΠΙΠΤΩΣΕΙΣ ΚΑΙ ΣΤΑΤΙΣΤΙΚΑ ΣΤΟΙΧΕΙΑ

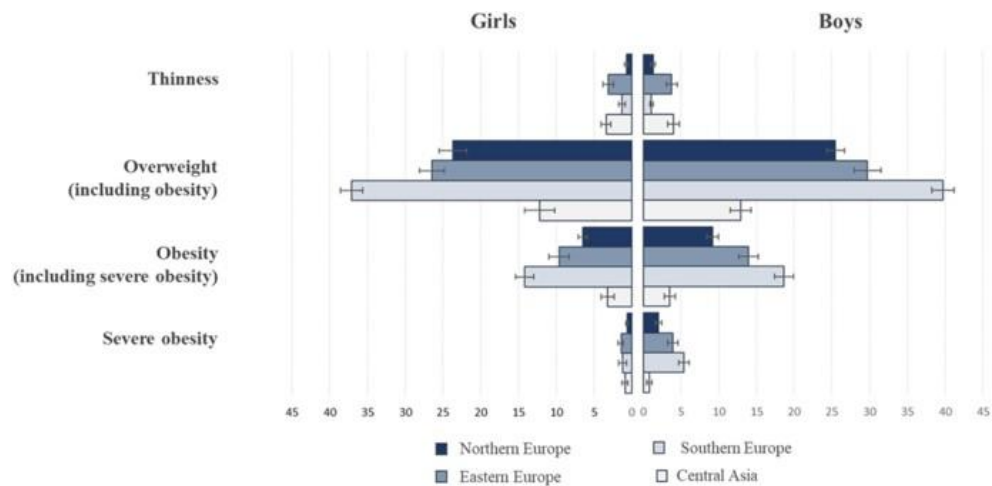
Η παχυσαρκία μπορεί να ευνοήσει την ανάπτυξη πολλαπλών επιπλοκών οι οποίες έχουν αρνητικές επιπτώσεις σε διάφορα συστήματα οργάνων και να οδηγήσει σε μεγαλύτερη θνησιμότητα από την πρόωμη ενήλικη ζωή εξαιτίας καρδιαγγειακών επεισοδίων. Η παιδική παχυσαρκία έχει συνδεθεί με προβλήματα υγείας, όπως η αντίσταση σε ινσουλίνη (insuline resistance), το μεταβολικό σύνδρομο, η αρτηριακή υπέρταση, ο σακχαρώδης διαβήτης τύπου 2, το σύνδρομο πολυκυστικών ωοθηκών, το σύνδρομο αποφρακτικής άπνοιας ύπνου, καθώς και ψυχολογικά προβλήματα κατά την παιδική ηλικία και την εφηβεία [14].

Αν και δευτερεύουσες σε σχέση με τις αντίστοιχες στην υγεία του ασθενή, οι οικονομικές συνέπειες της παιδικής παχυσαρκίας δεν πρέπει να αμελούνται. Εκτιμάται ότι το σύνολο του κόστους εξαιτίας της παιδικής ή εφηβικής παχυσαρκίας ανέρχεται σε περίπου 150,000€ κατά τη διάρκεια της ζωής ενός ασθενή, με το μεγαλύτερο μέρος αυτού του ποσού να οφείλεται σε απώλειες παραγωγικότητας εξαιτίας της παχυσαρκίας και ένα μικρό μέρος στα κόστη θεραπείας ασθενειών ως απόρροια της παχυσαρκίας [15].

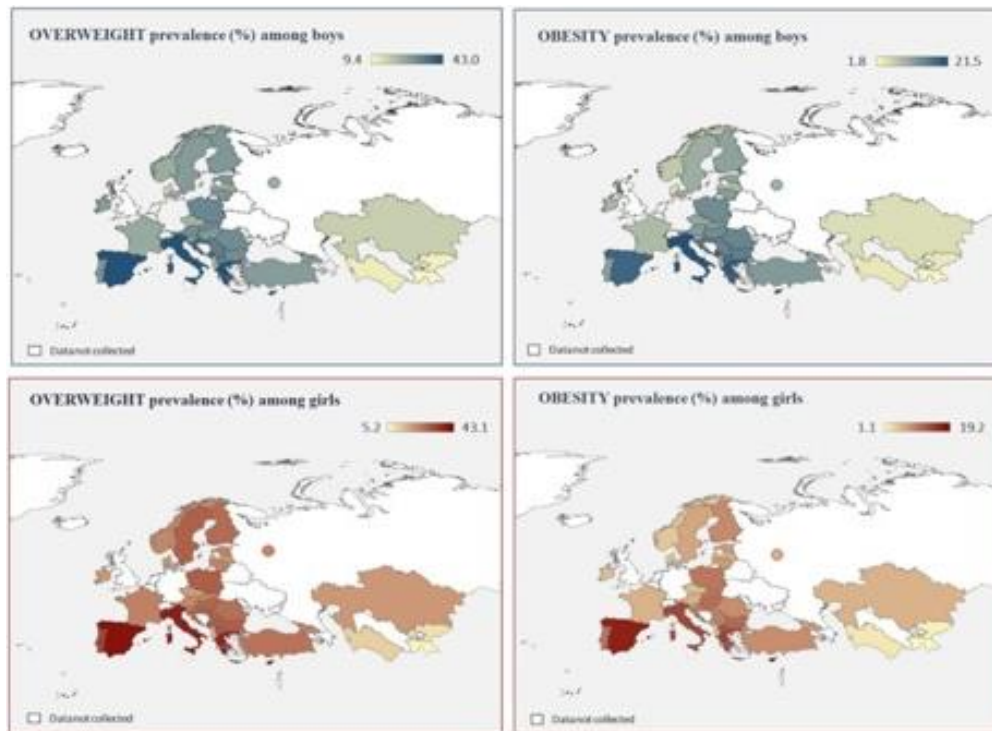
Η παιδική παχυσαρκία είναι μία σύνθετη ασθένεια η οποία οφείλεται σε διάφορα είδη παραγόντων, όπως γενετικοί, περιγεννητικοί, επίκτητοι, διατροφικοί και ψυχοκοινωνικοί. Η ανισορροπία μεταξύ της προσλαμβανόμενης και της καταναλισκόμενης ενέργειας από γενετικά επιρρεπή άτομα αποτελεί μία από τις κύριες αιτίες της παχυσαρκίας [14]. Συγκεκριμένα, η αναντιστοιχία μεταξύ προσλαμβανόμενης ενέργειας και θερμιδικών αναγκών υπέρ της πρώτης οδηγεί σταδιακά σε αύξηση του βάρους [16]. Σε αυτό συμβάλλουν σε μεγάλο βαθμό η κατανάλωση βιομηχανοποιημένων τροφίμων με υψηλή περιεκτικότητα σε πολυακόρεστα λίπη, νάτριο και υψηλούς γλυκαιμικούς δείκτες, σε συνδυασμό με τη μείωση της σωματικής δραστηριότητας και αύξηση των καθιστικών συνηθειών των παιδιών, που επιφέρει ο σύγχρονος τρόπος ζωής. Ωστόσο, σημαντικό ρόλο παίζουν παράγοντες όπως το κάπνισμα της μητρός κατά την εγκυμοσύνη, το αν οι γονείς του παιδιού είναι παχύσαρκοι, η ποιότητα και ποσότητα του ύπνου του παιδιού, ενώ υπολογίζεται ότι μόνο οι γενετικοί παράγοντες μπορούν να είναι υπεύθυνοι μέχρι και για το 80% του βάρους ενός παιδιού [14].

Στη σύγχρονη εποχή είναι αισθητή η εξάπλωση της παιδικής παχυσαρκίας. Συγκεκριμένα, στη νότια Ευρώπη τα ποσοστά της υπερβαρότητας κυμαίνονται στο 39.7% και 37.1% και της

παχυσαρκίας στο 18.6% και 14.2%, για αγόρια και κορίτσια, αντίστοιχα [17]. Στην Ελλάδα, ενώ τα ποσοστά της παιδικής παχυσαρκίας παραμένουν από τα υψηλότερα στην Ευρώπη, τα τελευταία χρόνια παρατηρείται μείωση της εξάπλωσης του φαινομένου. Η πτωτική τάση της παιδικής παχυσαρκίας αποδίδεται σε αρκετές προληπτικές παρεμβάσεις που έχουν εφαρμοστεί σε παιδιά ηλικίας δημοτικού. Για να διατηρηθεί αυτή η τάση, προτείνεται η υιοθέτηση στρατηγικών οι οποίες περιλαμβάνουν τη συμμετοχή της οικογένειας, της κοινότητας και του σχολικού προσωπικού και αξιοποιούν νέες τεχνολογίες, όπως τα Big Data και οι έξυπνες συσκευές [18].



Σχήμα 2.1: Διάγραμμα σύγκρισης των ποσοστών εξάπλωσης της λιποσαρκίας, της υπερβαρότητας, της παχυσαρκίας και της σοβαρής παχυσαρκίας μεταξύ περιοχών της Ευρώπης και της Κεντρικής Ασίας. Τα μεγαλύτερα ποσοστά υπερβαρότητας, παχυσαρκίας και σοβαρής παχυσαρκίας για αγόρια και κορίτσια σημειώνονται στις χώρες της Νότιας Ευρώπης. Πηγή: [17]



Σχήμα 2.2: Αναπαράσταση σε χάρτη της εξάπλωσης της υπερβαρότητας και της παχυσαρκίας σε αγόρια και κορίτσια για ορισμένες χώρες της Ευρώπης και της Κεντρικής Ασίας. Τα υψηλότερα ποσοστά υπερβαρότητας και παχυσαρκίας για αγόρια και κορίτσια παρατηρούνται στις χώρες της Νότιας Ευρώπης συμπεριλαμβανομένης και της Ελλάδας. Πηγή: [17]

2.2. ΑΝΤΙΜΕΤΩΠΙΣΗ / ΔΙΑΧΕΙΡΙΣΗ

Η μεγάλη εξάπλωση της παιδικής παχυσαρκίας τονίζει την ανάγκη για προσιτές και αποτελεσματικές παρεμβάσεις ρύθμισης σωματικού βάρους. Διεπιστημονικές παρεμβάσεις οι οποίες εστιάζουν στην οικογένεια, υγιεινές συνήθειες και αλλαγές συμπεριφοράς μπορούν να συμβάλουν στον περιορισμό της παιδικής παχυσαρκίας. Τα παιδιά και οι οικογένειες που παρευρίσκονται σε περισσότερες συνεδρίες και λαμβάνουν τη θεραπεία για μεγαλύτερες χρονικές περιόδους επιτυγχάνουν τη σημαντικότερη μείωση σωματικού βάρους και βελτίωση της υγείας τους [19].

Δυστυχώς, για πολλούς λόγους, μεταξύ άλλων δυσκολίες στις μετακινήσεις και μη-εκπλήρωση αναγκών ή προσδοκιών, πολλές οικογένειες αποχωρούν πρόωρα από τις παρεμβάσεις. Αυτό έχει ως αποτέλεσμα τον περιορισμό των οφελών υγείας που τα παιδιά θα αποκόμιζαν αν παρέμεναν στη θεραπεία. Οι επαγγελματίες υγείας και οι ερευνητές που δουλεύουν με παιδιά, αναγνωρίζουν ότι η κόπωση (attrition) αποτελεί μεγάλη πρόκληση για την προσπάθεια αντιμετώπισης της παιδικής παχυσαρκίας. Η πρόωρη αποχώρηση ενός ασθενή από μια θεραπεία έχει ως αποτέλεσμα τη μη-επωφελή χρήση των πόρων υγειονομικής περίθαλψης, τη μείωση της παραγωγικότητας των επαγγελματιών υγείας και την αποθάρρυνση των οικογενειών να αξιοποιήσουν μελλοντικά υπηρεσίες ρύθμισης σωματικού βάρους [19].

Η αποτελεσματικότητα των παρεμβάσεων για την αντιμετώπιση της παιδικής παχυσαρκίας βελτιώνεται σημαντικά με την αύξηση της έντασης και της διάρκειας του προγράμματος [20]. Ωστόσο, η αυξημένη ένταση και διάρκεια απαιτούν πολύ χρόνο από τις κλινικές ομάδες, πιθανώς μεγάλη ταλαιπωρία για τις οικογένειες των παιδιών και αυξημένο κόστος για όλα τα εμπλεκόμενα μέρη. Οι εφαρμογές της κινητής υγείας (mobile Health) έρχονται να βοηθήσουν στην προσπάθεια να αντιμετωπιστούν τα παραπάνω. Πολλές μελέτες περιλαμβάνουν τη χρήση κάποιου είδους activity tracker για την καταγραφή της σωματικής δραστηριότητας των παιδιών [21]–[26], καθώς και ειδικών εφαρμογών για smartphone [22], [24], [25], [27]–[32] οι οποίες μπορεί να χρησιμοποιούνται τόσο για την καταγραφή δεδομένων άσκησης, διατροφής ή μετρήσεων από τα παιδιά και τους γονείς όσο και για την επικοινωνία με την υπεύθυνη κλινική ομάδα.

Παρά την αφθονία προγραμμάτων ρύθμισης βάρους για παιδιά και εφήβους στη βιβλιογραφία, ως επί το πλείστον, χρησιμοποιούνται συμβατικές στατιστικές μέθοδοι για την μελέτη της βιωσιμότητας [22], [29]–[31], [33]–[36] και της έκβασης των θεραπειών [22]–[25], [27]–[29], [31], [33], [35], [37]. Συγκεκριμένα, χρησιμοποιούνται τεχνικές παλινδρόμησης, όπως γραμμική παλινδρόμηση (linear regression) [28], [29], [34], γραμμικά μοντέλα μικτών επιδράσεων (linear mixed-effects models) [23]–[25], λογιστική παλινδρόμηση (logistic regression) [24] και γραμμικό μοντέλο γενικευμένης εξίσωσης εκτίμησης (generalized estimating equation linear model) [31], καθώς και τεχνικές ανάλυσης διακύμανσης (analysis of variance) [22], [24], [27]–[31], [33], [35]–[37]. Πιο προηγμένες τεχνικές, όπως η μηχανική μάθηση, θα μπορούσαν να συμβάλουν στην ανάπτυξη ισχυρότερων μοντέλων για την εξαγωγή συμπερασμάτων από τα δεδομένα των παρεμβάσεων.

Τα τελευταία χρόνια, μελετάται η αξιοποίηση μοντέλων μηχανικής μάθησης για την πρόληψη και την αντιμετώπιση της παιδικής παχυσαρκίας. Ο λόγος είναι ότι τα τελευταία δύνανται να αποτελέσουν χρήσιμα εργαλεία στο εν λόγω πεδίο χάρη στην εξαιρετική προβλεπτική τους ισχύ και την ικανότητά τους να μοντελοποιούν σύνθετες, μη-γραμμικές αλληλεξαρτήσεις

μεταξύ μεταβλητών και να χειρίζονται δεδομένα μεγάλων διαστάσεων με μικρό αριθμό δειγμάτων [38].

Ο κύριος όγκος των ερευνών στο πεδίο της παιδικής παχυσαρκίας που χρησιμοποιούν μηχανική μάθηση εστιάζει στην πρόβλεψη της παχυσαρκίας σε μετέπειτα ηλικία [39]–[43] ή τη συσχέτισή της με κάποιες μεταβλητές [44]–[50]. Τα μοντέλα που έχουν αναπτυχθεί έχουν ως έξοδο είτε μία εκτίμηση για το BMI του παιδιού [39], [41]–[43] είτε την κατάταξή του σε κατάλληλη κλάση BMI [39], [40], [43]–[50]. Τρεις έρευνες χρησιμοποίησαν δεδομένα από ηλεκτρονικούς φακέλους υγείας (Electronic Health Records), αξιοποιώντας την πληθώρα διαθέσιμων δειγμάτων για να εκπαιδεύσουν τα μοντέλα τους [39], [40], [43]. Το μοντέλο των Rios-Julian κ.ά. μπορεί με χρήση εύκολα μετρήσιμων ανθρωπομετρικών και δημογραφικών δεδομένων να εκτιμήσει την παχυσαρκία σε παιδιά χωρίς την ανάγκη μέτρησης σωματικού λίπους [48]. Επίσης, το μοντέλο των Park κ.ά. μπορεί με χρήση δεδομένων απεικόνισης λειτουργικού μαγνητικού συντονισμού (fMRI) να προβλέψει σε ικανοποιητικό βαθμό την εξέλιξη του BMI εφήβων, ένα με δύο χρόνια μετά την λήψη του fMRI [41]. Στόχο αυτών των ερευνών αποτελούσε η μελέτη των παραγόντων που οδηγούν στην παχυσαρκία σε παιδικές και εφηβικές ηλικίες. Ωστόσο, παρατηρείται μεγάλη ποικιλομορφία όσον αφορά τα χαρακτηριστικά (features) πάνω στα οποία έχουν βασιστεί τα διαθέσιμα στη βιβλιογραφία μοντέλα μηχανικής μάθησης. Αυτό υπογραμμίζει το γεγονός ότι η παιδική παχυσαρκία εξαρτάται από πληθώρα ετερογενών παραγόντων, καθώς και ότι τα κατάλληλα δεδομένα για την εκπαίδευση ενός συστήματος μηχανικής μάθησης διαφέρουν ανάλογα με το μοντέλο και το ζητούμενο.

Πρόσφατα, έχουν αναπτυχθεί μοντέλα μηχανικής μάθησης για την ανάλυση των δεδομένων και την εξαγωγή συμπερασμάτων σχετικών με παρεμβάσεις που έχουν ως στόχο τη ρύθμιση του βάρους σε παιδιά και εφήβους. Εντούτοις, η χρήση μηχανικής μάθησης για τη βελτίωση παρεμβάσεων για την παιδική παχυσαρκία δεν ακόμα τόσο διαδεδομένη [38].

Το 2018, οι Hasan κ.ά. εφάρμοσαν τόσο πιθανοτικές μεθόδους όσο και τεχνικές μηχανικής μάθησης που βασίζονται σε επαναλαμβανόμενα νευρωνικά δίκτυα (recurrent neural networks), προκειμένου να προβλέψουν το αποτέλεσμα κινητοποιητικών συνεντεύξεων (motivational interviews) σε παχύσαρκα παιδιά, δηλαδή αν η συνέντευξη έληξε με μία θετική ή αρνητική απάντηση εκ μέρους του παιδιού. Χρησιμοποιήθηκαν μοντέλα μακράς βραχυχρόνιας μνήμης (long short-term memory) τα οποία είχαν καλύτερη απόδοση σε σχέση με τα πιθανοτικά που δοκιμάστηκαν, επιτυγχάνοντας μάλιστα θετική διαγνωστική αξία (positive predictive value) της τάξης του 87.3% [51].

Τον ίδιο χρόνο, οι Öksüz κ.ά. εκμεταλλεύτηκαν διάφορες τεχνικές μηχανικής μάθησης για να αναπτύξουν μοντέλα τα οποία θα προβλέπουν αν ένα πρόγραμμα παρέμβασης που απαιτούσε συμμόρφωση του παιδιού με πλάνο άσκησης και διατροφής για διάστημα έξι μηνών θα ήταν επιτυχημένο, δηλαδή αν το BMI του παιδιού θα μειωνόταν κατά 0.4 της μονάδας ή περισσότερο. Τα δεδομένα εισόδου ήταν ανθρωπομετρικά, καθώς και δεδομένα καρδιακών παλμών τα οποία αποκτήθηκαν από τεστ φυσικής κατάστασης το οποίο διενεργήθηκε για τους σκοπούς της έρευνας. Δοκιμάστηκαν αρκετά μοντέλα k κοντινότερων γειτόνων (k nearest neighbors), δέντρα αποφάσεων (decision trees), μοντέλο ενίσχυσης κλίσης (gradient boosting), μηχανές διανύσματος υποστήριξης (support vector machines). Το γραμμικό SVM πέτυχε ακρίβεια 85% στη δυαδική ταξινόμηση σε επιτυχημένη/μη-επιτυχημένη θεραπεία, παρέχοντας καλύτερες προβλέψεις από ειδικούς παιδίατρος στους οποίους δόθηκαν τα ίδια δεδομένα και ζητήθηκε να προβλέψουν την έκβαση των θεραπειών. Επίσης, το μοντέλο αξιολογήθηκε θετικά από ειδικούς ως προς τη χρησιμότητά σε πιθανή

μελλοντική εφαρμογή του στην κλινική πράξη. Σημαντικός περιορισμός του μοντέλου ήταν το μικρό μέγεθος του δείγματος το οποίο αποτελούνταν από 20 εφήβους ηλικίας 11-16 ετών [52].

Στην έρευνα των Rajarakse κ.ά. που δημοσιεύτηκε το 2021, αναλύεται ένα σύστημα για τη διαχείριση της παιδικής παχυσαρκίας το οποίο θα μπορούσε να χρησιμοποιηθεί από τους γονείς και τα παιδιά, χωρίς τη συμμετοχή κάποιου ειδικού. Συγκεκριμένα, αναπτύχθηκε μία εφαρμογή για smartphone η οποία λαμβάνει δεδομένα από έναν ιχνηλάτη δραστηριοτήτων. Η εφαρμογή έχει τις εξής δυνατότητες: να αναγνωρίζει και να καταγράφει τις δραστηριότητες στις οποίες συμμετέχει το παιδί, να παρέχει εξατομικευμένες προτάσεις για δραστηριότητες, να προβλέπει τις θερμιδικές ανάγκες των παιδιών, καθώς και να αποφαινεται αν το παιδί βρίσκεται σε κίνδυνο για παθήσεις σχετιζόμενες με την παχυσαρκία όταν μεγαλώσει. Για να υλοποιηθούν οι τέσσερις παραπάνω λειτουργίες αναπτύχθηκαν μοντέλα μηχανικής μάθησης βασισμένα σε δεδομένα που προέκυψαν από χρήση της εφαρμογής από 11 παιδιά ηλικίας 5-10 ετών. Επιγραμματικά, εκπαιδεύτηκαν μοντέλα συνελκτικών νευρωνικών δικτύων (convolutional neural networks)-LSTM, τυχαίου δάσους (random forest), SVM, καθώς και ενισχυτικής μάθησης και όλα επιτύχαναν ακρίβεια άνω του 90% [26].

Το 2022, οι Fayyaz κ.ά. χρησιμοποίησαν μηχανική μάθηση για να προβλέψουν την κόπωση (attrition) των παιδιών από ένα πρόγραμμα παρέμβασης. Αναλυτικότερα, χρησιμοποιώντας ένα μεγάλο σύνολο δεδομένων από EHR εκπαιδύσαν ένα πολυεπίπεδο νευρωνικό δίκτυο το οποίο μπορεί να προβλέψει δύο ζητούμενα για κάθε παιδί για διάφορα χρονικά διαστήματα: αν θα αποχωρήσει από το πλάνο παρέμβασης πριν από ένα δεδομένο χρονικό διάστημα και αν θα μειωθεί η εκατοστιαία θέση BMI του. Το μοντέλο είχε καλύτερη απόδοση σε σχέση με συμβατικές στατιστικές μεθόδους, επιτυγχάνοντας μάλιστα area under the curve της τάξης του 84% για τα μακρύτερα χρονικά διαστήματα. Τα σημαντικότερα χαρακτηριστικά και για τις δύο προβλέψεις ήταν η ηλικία των παιδιών, με τα μικρότερα παιδιά να παρουσιάζουν μεγαλύτερη συμμετοχή και μείωση του BMI, ο χρόνος μεταξύ των επισκέψεων στο νοσοκομείο για το πλάνο παρέμβασης, με τα παιδιά τα οποία πηγαίνουν πιο αραιά να έχουν περισσότερες πιθανότητες για αποχώρηση και μη-βελτίωση του BMI, και η αρχική πορεία του BMI του παιδιού στην αρχή του πλάνου παρέμβασης, με τις θεραπείες σε παιδιά τα οποία βλέπουν από νωρίς αποτέλεσμα να έχουν περισσότερες πιθανότητες τόσο να ολοκληρωθούν, όσο και να είναι επιτυχημένες [53].

Οι χρήση μηχανικής μάθησης για την παιδική παχυσαρκία είναι πολλά υποσχόμενη. Όμως, επικρατεί μία δυσπιστία από μη-ειδικούς προς τις εφαρμογές τεχνητής νοημοσύνης, ιδίως όταν αφορούν τον χώρο της υγείας, στον οποίο οι αποφάσεις του μοντέλου μπορούν να επηρεάσουν την υγεία του ασθενή. Μία πρόκληση για τους σχεδιαστές συστημάτων μηχανικής μάθησης είναι η αποδοχή τους από την κοινότητα των ασθενών και των επαγγελματιών υγείας [54], [55].

Εξ όσων γνωρίζω, πλην του δημοσιεύματος των Rajarakse κ.ά. [26], η ενισχυτική μάθηση δεν έχει επιστρατευτεί ακόμα σε κάποια άλλη εφαρμογή για την καταπολέμηση της παιδικής παχυσαρκίας. Ωστόσο, έχει αξιοποιηθεί στην προσπάθεια για την αντιμετώπιση της παχυσαρκίας στους ενήλικες [56]–[58]. Οι εφαρμογές αυτές υποδεικνύουν ότι οι δυνατότητες που προσφέρει η ενισχυτική μάθηση στη βελτιστοποίηση ακολουθιακών διαδικασιών με συγκεκριμένο στόχο θα μπορούσαν να συνεισφέρουν αποτελεσματικά στην προσπάθεια περιορισμού της παιδικής παχυσαρκίας.

3. ΕΠΙΣΤΗΜΟΝΙΚΟ ΥΠΟΒΑΘΡΟ

Αυτό το κεφάλαιο αποτελεί μία συνοπτική επισκόπηση των τεχνολογιών που χρησιμοποιήθηκαν για την εκπόνηση της παρούσας εργασίας.

Η μηχανική μάθηση (Machine Learning) συνιστά ένα πεδίο της επιστήμης υπολογιστών το οποίο αναζητά λύσεις για σύνθετα προβλήματα των οποίων ο προγραμματισμός είναι δύσκολος με συμβατικές προγραμματιστικές μεθόδους [59]. Διαισθητικά, η μηχανική μάθηση αποτελεί μία κατηγορία τεχνητής νοημοσύνης η οποία επιτρέπει στους υπολογιστές να σκέφτονται και να μαθαίνουν από μόνοι τους. Ένα πρόγραμμα υπολογιστών λέγεται ότι μαθαίνει από κάποια εμπειρία σχετική με μία εργασία, αν η απόδοσή του στην εργασία βελτιώνεται με την απόκτηση της εμπειρίας αυτής [60]. Για παράδειγμα, στην παρούσα εργασία, τα μοντέλα μηχανικής μάθησης για εκτίμηση της συμμόρφωσης μαθαίνουν από τα δεδομένα για τα παιδιά που συμμετείχαν στο ENDORSE [10]–[13] και τον βαθμό συμμόρφωσης που είχε κάθε παιδί, για να βελτιωθεί η απόδοσή τους στο να προβλέπουν τον τελικό βαθμό συμμόρφωσης ενός παιδιού που συμμετέχει στη μελέτη, όντων γνωστών των αρχικών δεδομένων.

3.1. ΤΕΧΝΟΛΟΓΙΕΣ ΕΠΙΒΛΕΠΟΜΕΝΗΣ ΜΑΘΗΣΗΣ

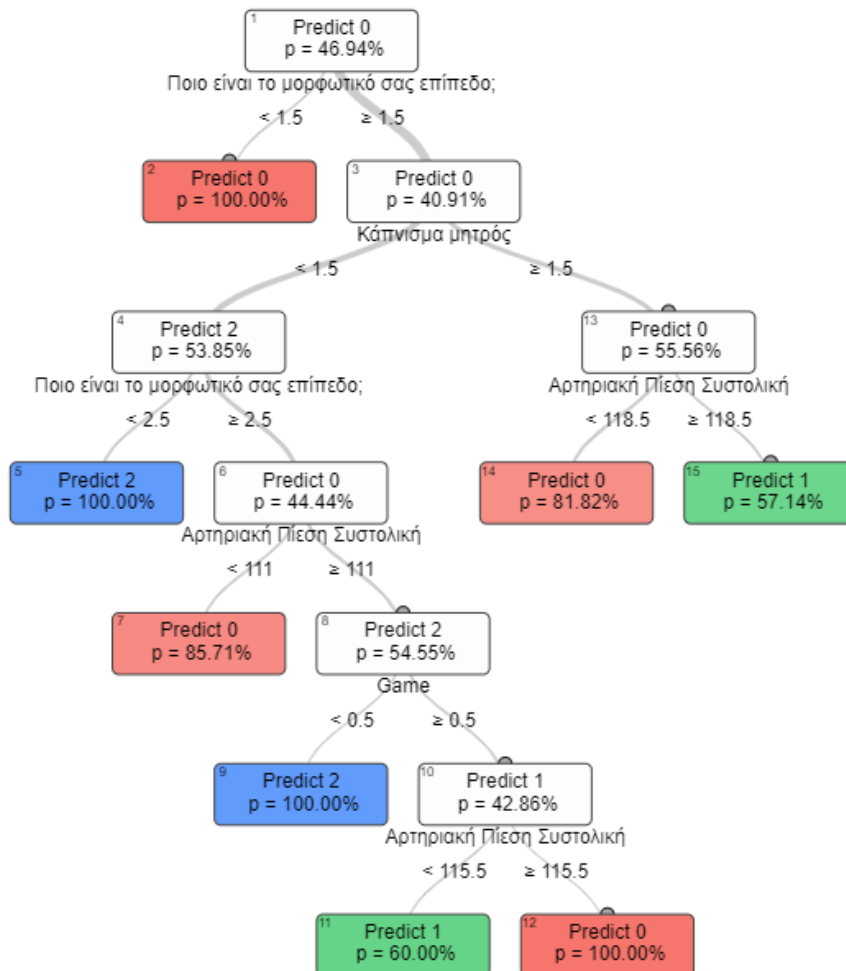
Για την πρόβλεψη της συμμόρφωσης στο πλάνο χρησιμοποιήθηκαν δύο ταξινομητές: ο XGBClassifier του πακέτου xgboost και το Optimal Classification Tree του πακέτου interpretableai.

Βάση των ταξινομητών του XGBoost και του InterpretableAI αποτελούν τα δέντρα αποφάσεων (decision trees). Τα δέντρα αποφάσεων προτιμώνται πολλές φορές στον τομέα της υγείας σε σχέση με άλλα πιο σύνθετα μοντέλα, κυρίως διότι είναι εύληπτα από μη-ειδικούς και εύκολα ερμηνεύσιμα [61].

Ο XGBoost είναι ένα κλιμακώσιμο (scalable) σύστημα μηχανικής μάθησης το οποίο υλοποιεί έναν αλγόριθμο ενίσχυσης κλίσης δέντρων αποφάσεων (gradient boosting decision trees). Ο XGBoost εφαρμόζει μία μέθοδο από κοινού μάθησης (ensemble) κατά την οποία πολλά αδύναμα δέντρα αποφάσεων συνδυάζονται για να προκύψει ένα ισχυρότερο μοντέλο πρόβλεψης. Για την επιλογή των δέντρων που θα συμμετέχουν στην πρόβλεψη, χρησιμοποιείται αλγόριθμος κατάβασης κλίσης (gradient descent), δηλαδή το αρχικό δέντρο συνδυάζεται με το δέντρο το οποίο θα οδηγήσει στη μεγαλύτερη βελτίωση των αποτελεσμάτων. Σε κάθε βήμα, στο μοντέλο προστίθεται με άπληστο τρόπο από τα υποψήφια δέντρα, αυτό το οποίο θα βελτιστοποιήσει τη συνολική απόδοση του μοντέλου, μέχρι να καταλήξει σε τοπικό μέγιστο [62].

Παραδοσιακά, κατά την κατασκευή των δέντρων αποφάσεων εξετάζεται η προσθήκη ενός κόμβου απόφασης ανά βήμα, συνήθως με άπληστο τρόπο, δηλαδή επιλέγεται να προστεθεί ο κόμβος ο οποίος αυξάνει περισσότερο την απόδοση του υπό κατασκευή δέντρου άμεσα, ακόμα και αν η προσθήκη ενός διαφορετικού κόμβου μπορεί να οδηγήσει σε καλύτερα αποτελέσματα σε επόμενα βήματα. Η διαδικασία εξέτασης της προσθήκης ενός κόμβου ανά βήμα μπορεί πολλές φορές να οδηγήσει σε υποβέλτιστα δέντρα. Το Optimal Classification Tree χρησιμοποιεί τεχνικές μικτής ακέραιης βελτιστοποίησης (mixed integer optimization) η οποία λαμβάνει υπόψιν και τους υπόλοιπους κόμβους κατά την κατασκευή ή το κλάδεμα (pruning) ενός κόμβου διαχωρισμού (split node). Με αυτόν τον τρόπο, προκύπτουν δέντρα

αποφάσεων τα οποία έχουν κατασκευαστεί με βάση μια πιο ολοκληρωμένη εικόνα του συνόλου δεδομένων εκπαίδευσης και έχουν καλύτερες επιδόσεις σε σχέση με συμβατικά δέντρα αποφάσεων. [61]. Πέραν του ταξινομητή, το πακέτο `interpretableai` περιλαμβάνει και δυνατότητες διαδραστικής οπτικοποίησης των παραγόμενων δέντρων (Σχήμα 3.1).



Σχήμα 3.1: Παράδειγμα ταξινομητή δέντρου αποφάσεων που προέκυψε μετά από εκπαίδευση του `OptimalTreeClassifier` του πακέτου `InterpretableAI`. Το συγκεκριμένο δέντρο προβλέπει τη συμμόρφωση με την εφαρμογή. Οι κλάσεις συμμόρφωσης Χαμηλή, Μεσαία, Υψηλή έχουν κωδικοποιηθεί σε 0, 1, 2 αντίστοιχα, ενώ έχει πραγματοποιηθεί αριθμητική κωδικοποίηση και για τις κατηγορικές μεταβλητές.

Κατά την ανάπτυξη μοντέλων μηχανικής μάθησης, ο προγραμματιστής μπορεί να επηρεάσει την έξοδο των μοντέλων με δύο τρόπους: με την επιλογή των χαρακτηριστικών (feature selection) και με τον συντονισμό υπερπαραμέτρων (hyperparameter tuning).

Όταν οι διαστάσεις των διαθέσιμων δεδομένων είναι πολλές, ένα σημαντικό μέρος από τα δεδομένα πιθανότατα δεν επηρεάζει κάπως το μετρούμενο αποτέλεσμα. Ειδικότερα, όταν ο αριθμός των χαρακτηριστικών, p , είναι μεγαλύτερος του αριθμού των δειγμάτων, n , για κάθε δείγμα είναι γνωστές περισσότερες πληροφορίες από όσες μπορούν να δώσουν μια αξιόπιστη περιγραφή του προβλήματος. Τα πλεονάζοντα δεδομένα ενδέχεται να μειώσουν την απόδοση του μοντέλου και να οδηγήσουν σε λανθασμένες προβλέψεις και συμπεράσματα σχετικά με τους παράγοντες που επηρεάζουν το ζητούμενο, ως αποτέλεσμα υπερπροσαρμογής (overfitting). Ένα μοντέλο υπερπροσαρμόζεται στα δεδομένα όταν

αποδίδει πολύ καλά για δείγματα από το σύνολο δεδομένων με το οποίο εκπαιδεύτηκε, αλλά όχι ικανοποιητικά για δείγματα εκτός αυτού [63]. Επομένως, είναι σημαντικό, στα μοντέλα, από τα διαθέσιμα χαρακτηριστικά (features) να εισαχθούν αυτά τα οποία σχετίζονται όντως με το πρόβλημα και οδηγούν σε υψηλότερες αποδόσεις των μοντέλων.

Οι οικογένειες μεθοδολογιών επιλογής χαρακτηριστικών στη βιβλιογραφία είναι τρεις [64]:

- Μέθοδοι φιλτραρίσματος (filter methods), με τις οποίες επιλέγονται τα καλύτερα χαρακτηριστικά σύμφωνα με στατιστικές μετρικές, όπως οι συσχετίσεις Pearson και Spearman.
- Μέθοδοι καλύμματος (wrapper methods), οι οποίες χρησιμοποιούν τις επιδόσεις του μοντέλου για να επιλέξουν τα καλύτερα χαρακτηριστικά, όπως η διαδοχική επιλογή χαρακτηριστικών (sequential feature selection) και ο αναδρομικός αποκλεισμός χαρακτηριστικών (recursive feature elimination).
- Ενσωματωμένες μέθοδοι (embedded methods), οι οποίες προκύπτουν μέσα από το μοντέλο, όπως τα χαρακτηριστικά με τις μεγαλύτερες «σημαντικότητες» (feature importances).

Η μέθοδος sequential feature selection ανήκει στην οικογένεια των μεθόδων καλύμματος και χρησιμοποιείται για την επιλογή από ένα σύνολο p χαρακτηριστικών. Συγκεκριμένα, στην πρώτη επανάληψη, δοκιμάζεται εξαντλητικά η εκπαίδευση p μοντέλων καθένα από τα οποία έχει ως είσοδο το τμήμα των δεδομένων του συνόλου εκπαίδευσης που αντιστοιχούν σε ένα χαρακτηριστικό. Ακολουθεί η αξιολόγηση των μοντέλων και το χαρακτηριστικό που οδήγησε στο μοντέλο με τις καλύτερες επιδόσεις χρησιμοποιείται και στα επόμενα βήματα. Στη συνέχεια, εκπαιδεύονται $p - 1$ μοντέλα καθένα από τα οποία με είσοδο δύο χαρακτηριστικά, αυτό που προκρίθηκε κατά την προηγούμενη επανάληψη και ένα ακόμα, και πραγματοποιείται ξανά αξιολόγηση όλων των μοντέλων με στόχο να βρεθεί το ζεύγος χαρακτηριστικών που οδηγεί στις καλύτερες επιδόσεις. Γενικά, στην i -οστή επανάληψη εκπαιδεύονται $p - i + 1$ μοντέλα καθένα από τα οποία έχει ως είσοδο την ένωση του συνόλου των $i - 1$ χαρακτηριστικών που προέκυψε από την προηγούμενη επανάληψη με ένα από τα εναπομείναντα χαρακτηριστικά και κατόπιν τα μοντέλα αξιολογούνται, ώστε να προκύψει ένα σύνολο i χαρακτηριστικών. Η αναζήτηση σταματάει όταν συμπληρωθεί ένας καθορισμένος αριθμός επαναλήψεων ή όταν η προσθήκη κανενός από τα αχρησιμοποίητα χαρακτηριστικά στο σύνολο επιλεγμένων χαρακτηριστικών δεν οδηγεί στην εκπαίδευση ενός μοντέλου με βελτιωμένες επιδόσεις.

Επιπρόσθετα των εισαγόμενων δεδομένων, τα μοντέλα μηχανικής μάθησης εξαρτώνται και από κάποιες υπερπαραμέτρους οι οποίες ορίζονται από τον χρήστη. Για παράδειγμα, σε έναν ταξινομητή ο οποίος χρησιμοποιεί δέντρα αποφάσεων, ο χρήστης μπορεί να ορίζει μέσω μίας υπερπαραμέτρου το μέγιστο επιτρεπόμενο βάθος των δέντρων. Οι διαφορετικές τιμές των υπερπαραμέτρων μπορεί να οδηγήσουν και σε διαφορετικές προβλέψεις από το μοντέλο μηχανικής μάθησης, οπότε είναι σημαντικό να βρεθούν οι τιμές εκείνες οι οποίες βελτιστοποιούν την απόδοση του μοντέλου. Παραδείγματα υπερπαραμέτρων στον XGBoost είναι το μέγιστο βάθος των δέντρων που χρησιμοποιεί, ο αριθμός των εκτιμητών (no of estimators) που θα ληφθούν υπόψη στην τελική πρόβλεψη και ο βαθμός εκμάθησης (learning rate). Οι υπερπαραμέτροι του Optimal Tree Classifier περιλαμβάνουν μεταξύ άλλων το μέγιστο βάθος του δέντρου, το κριτήριο επιλογής των κόμβων που θα προστεθούν και τον ελάχιστο αριθμό δειγμάτων σε ένα φύλλο.

Οι βέλτιστες τιμές των υπερπαραμέτρων μπορεί να βρεθούν με χρήση της μεθόδου της αναζήτησης πλέγματος (grid search). Το grid search αποτελεί έναν συστηματικό τρόπο εξερεύνησης της επίδρασης των διαφορετικών τιμών υπερπαραμέτρων σε ένα μοντέλο, με δοκιμή όλων των πιθανών συνδυασμών που προκύπτουν από έναν πίνακα πιθανών τιμών για κάθε υπερπαραμέτρο οριζόμενο από τον χρήστη.

3.2. ΕΡΜΗΝΕΥΣΙΜΗ ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΚΑΙ ΕΠΙΠΕΔΟ ΒΕΒΑΙΟΤΗΤΑΣ

Τα ισχυρότερα και πλέον περίπλοκα μοντέλα μηχανικής μάθησης, όπως η βαθιά μάθηση και η συλλογική (ensemble) μάθηση παρουσιάζουν καλύτερες επιδόσεις σε διάφορα προβλήματα πρόβλεψης, ταξινόμησης ή παλινδρόμησης σε σύγκριση με απλούστερα μοντέλα μηχανικής μάθησης, όπως τα μοντέλα γραμμικής παλινδρόμησης και τα δέντρα αποφάσεων. Ωστόσο, ο τρόπος με τον οποίο αποφασίζουν και ποιες αποφάσεις λαμβάνουν κατά τη λειτουργία τους τα σύνθετα μοντέλα είναι αδιαφανής [65]. Πολλές φορές κατά την επιλογή ενός μοντέλου είναι απαραίτητο να πραγματοποιηθεί ένας συμβιβασμός μεταξύ των επιδόσεων και της δυνατότητας ερμηνείας του [54].

Η ερμηνεύσιμη τεχνητή νοημοσύνη (Explainable Artificial Intelligence – XAI) περιλαμβάνει τις μεθόδους που αποσκοπούν να κάνουν τα συστήματα τεχνητής νοημοσύνης πιο κατανοητά στους ανθρώπους. Οι μέθοδοι ερμηνεύσιμης τεχνητής νοημοσύνης διέπονται από τις αρχές της ερμηνευσιμότητας (interpretability), της διαφάνειας (transparency) και της επεξηγησιμότητας (explainability). Η ερμηνευσιμότητα σχετίζεται με το κατά πόσο ένα μοντέλο μπορεί να γίνει κατανοητό. Η διαφάνεια αναφέρεται στην ικανότητα παροχής στα ενδιαφερόμενα μέρη πληροφορίες σχετικές με τη λειτουργία του μοντέλου ή του αλγορίθμου. Η επεξηγησιμότητα είναι ιδιότητα ενός μοντέλου που δίνει πληροφορίες για τα αίτια πίσω από τη διαδικασία λήψης αποφάσεων του συστήματος και παρέχει εργαλεία που επιτρέπουν την κατανόηση και ενίοτε την αμφισβήτηση της λογικής πίσω από τα εξαγόμενα αποτελέσματά του [54].

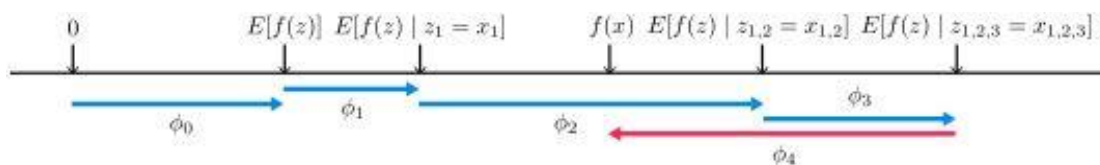
Οι τεχνικές ερμηνεύσιμης τεχνητής νοημοσύνης συμβάλλουν στην κατανόηση της λειτουργίας των μοντέλων και των χαρακτηριστικών των δεδομένων εισόδου τα οποία οδήγησαν στην πρόβλεψη. Επομένως, παρέχεται η δυνατότητα απομόνωσης των σημαντικότερων χαρακτηριστικών στα δεδομένα και χρήση τους για τη δημιουργία ισχυρότερων μοντέλων. Ακόμα, οι τεχνικές ερμηνεύσιμης τεχνητής νοημοσύνης ενισχύουν την αιτιότητα (causality), δηλαδή την ικανότητα να βρεθούν συσχετίσεις μεταξύ διαφόρων χαρακτηριστικών του συστήματος, καθώς και τη δυνατότητα μεταφοράς (transferability) των γνώσεων που αποκτώνται από τη διαδικασία στην προσπάθεια επίλυσης άλλων προβλημάτων [66].

Κατά την ανάπτυξη εφαρμογών στον χώρο της υγείας [4]–[9], είναι βαρύνουσας σημασίας ο τρόπος λειτουργίας των εφαρμογών να είναι εύκολα κατανοητός ακόμα και από άτομα χωρίς τεχνική κατάρτιση. Είναι ιδιαίτερα βοηθητικό μία εφαρμογή να μην είναι τύπου μαύρου κουτιού (black box), όπου δηλαδή ο χρήστης μπορεί να δει μόνο την είσοδο και την έξοδο χωρίς να έχει κάποια πληροφορία για την ενδιάμεση διαδικασία. Η ακρίβεια ενός μοντέλου δεν είναι από μόνη της αρκετή για να ληφθεί μία τεκμηριωμένη απόφαση στον χώρο της υγείας, αλλά είναι σημαντικό να είναι γνωστό και το σύνολο των χαρακτηριστικών που οδήγησαν στην πρόβλεψη, ώστε να δίνεται η δυνατότητα στους επαγγελματίες του χώρου να συγκρίνουν την απόφαση του μοντέλου με τη δική τους κλινική εκτίμηση [54], [55].

Το SHAP είναι μία τεχνική ερμηνεύσιμης τεχνητής νοημοσύνης, η οποία βασίζεται μεταξύ άλλων στην έννοια των τιμών Sharpley (Sharpley values) από τη συνεργατική θεωρία παιγνίων. Έστω μία ομάδα από παίκτες η οποία συνεργάζεται για την ολοκλήρωση ενός παιχνιδιού και λαμβάνει μια ανταμοιβή. Οι τιμές Sharpley μπορούν να εγγυηθούν ότι η κατανομή της ανταμοιβής σε κάθε παίκτη είναι δίκαιη υπολογίζοντας την οριακή συνεισφορά (marginal contribution) του καθενός, δηλαδή το πόσο συνέβαλε ο καθένας στην ολοκλήρωση του παιχνιδιού [66]. Ομοίως, η μέθοδος SHAP αναθέτει σε κάθε χαρακτηριστικό μία τιμή σημαντικότητας για μια συγκεκριμένη πρόβλεψη, ανάλογα με το πόσο συνεισέφερε η τιμή του χαρακτηριστικού στην πρόβλεψη [67].

Στο δημοσίευμα του SHAP, εισάγεται μία νέα οπτική, σύμφωνα με την οποία οποιαδήποτε επεξήγηση μια πρόβλεψης ενός μοντέλου αντιμετωπίζεται η ίδια ως μοντέλο. Αυτή η προσέγγιση επιτρέπει την ενοποίηση προγενέστερων μεθόδων για την ερμηνεία μοντέλων σε μία κλάση μεθόδων αθροιστικής κατανομής χαρακτηριστικών (additive feature attribution methods), η οποία έχει μοναδική λύση και πολλές επιθυμητές ιδιότητες [67].

Αναλυτικότερα, υπολογίζονται οι συντελεστές ϕ_i οι οποίοι δείχνουν την επίδραση κάθε χαρακτηριστικού και προκύπτουν από μια συνάρτηση δεσμευμένου μέσου όρου της εξόδου του μοντέλου, $f(\mathbf{z})$, όπου \mathbf{z} το διάνυσμα των χαρακτηριστικών του μοντέλου. Οι συντελεστές SHAP εξηγούν πώς πρέπει να μεταβληθεί η μέση τιμή $\mathbb{E}[f(\mathbf{z})]$, η οποία είναι η αναμενόμενη τιμή της πρόβλεψης του μοντέλου αν δεν γνωρίζει την τιμή κανενός χαρακτηριστικού προς την τιμή της πρόβλεψης $f(\mathbf{x})$. Οι τιμή ϕ_1 για το πρώτο χαρακτηριστικό ισούται με τη διαφορά της αναμενόμενης τιμής της $f(\mathbf{z})$ δεδομένης της τιμής του χαρακτηριστικού από την αναμενόμενη τιμή με άγνωστες τιμές όλων των χαρακτηριστικών. Αντίστοιχα, η τιμή ϕ_i ισούται με τη διαφορά της αναμενόμενης τιμής της $f(\mathbf{z})$ δεδομένων των τιμών των χαρακτηριστικών $1, 2, \dots, i-1$ από την αναμενόμενη τιμή δεδομένων των τιμών των χαρακτηριστικών $1, 2, \dots, i$. Καθώς η τιμή των ϕ_i ενδέχεται να διαφέρει ανάλογα με τη σειρά με την οποία τα χαρακτηριστικά προστίθενται στον υπολογισμό της αναμενόμενης τιμής για μη-γραμμικά μοντέλα ή δεδομένα στα οποία τα χαρακτηριστικά δεν είναι ανεξάρτητα, οι SHAP τιμές προκύπτουν από τον μέσο όρο των ϕ_i για όλες τις δυνατές αναδιατάξεις [67].



Σχήμα 3.2: Διάγραμμα επεξήγησης του τρόπου υπολογισμού των SHAP τιμών. Πηγή: [67]

Το δημοφιλές πακέτο shap της rython υλοποιεί την ομώνυμη μέθοδο παρέχοντας μία σειρά από κλάσεις για τον αποδοτικό υπολογισμό μίας εκτίμησης των shapley τιμών για κάθε χαρακτηριστικό για κάθε πρόβλεψη. Στην περίπτωση του ταξινομητή πολλαπλών κλάσεων, υπολογίζει τις shapley τιμές για την απόφαση να ανήκει το δείγμα σε κάθε κλάση ξεχωριστά. Δηλαδή, υπολογίζει για κάθε κλάση για ένα χαρακτηριστικό ενός δείγματος μία εκτίμηση της επίδρασης του χαρακτηριστικού στις πιθανότητες να αποφανθεί το μοντέλο ότι το χαρακτηριστικό ανήκει στην εκάστοτε κλάση.

Η μέθοδος SHAP είναι ανεξάρτητη μοντέλου (model-agnostic), δηλαδή μπορεί να υπολογίσει τις SHAP τιμές για ένα μοντέλο προβλέψεων ανεξαρτήτως του είδους του μοντέλου. Ωστόσο, το γεγονός ότι για την εκτίμηση των SHAP τιμών χρειάζεται να υπολογιστεί ο μέσος όρος των ϕ_i για όλες τις δυνατές αναδιατάξεις την καθιστά μη-αποδοτική. Για αυτόν τον λόγο, το

πακέτο `shap` βελτιστοποιεί την αποδοτικότητα για συγκεκριμένες κατηγορίες μοντέλων παρέχοντας εξειδικευμένες κλάσεις οι οποίες εκμεταλλεύονται τη δομή των μοντέλων για να ελαχιστοποιήσουν τον υπολογιστικό χρόνο. Μία τέτοια κλάση είναι ο `treeExplainer`, ο οποίος υλοποιεί την μέθοδο `Tree SHAP`, η οποία υπολογίζει αποδοτικά τις SHAP τιμές για μοντέλα που βασίζονται σε δέντρα αποφάσεων [68].

Η μέθοδος `TreeSHAP` ενδείκνυται για την ερμηνεία μοντέλων που περιλαμβάνουν πολλά δέντρα αποφάσεων και υποστηρίζει κάποια από τα πιο διαδεδομένα μοντέλα, όπως το `XGBoost`. Μάλιστα, έχει χρησιμοποιηθεί σε πληθώρα εφαρμογών στον χώρο της υγείας και δη για πολλά συστήματα υποστήριξης κλινικών αποφάσεων [66].

Είναι σημαντικό μία τεχνική ερμηνεύσιμης τεχνητής νοημοσύνης που προορίζεται για τον χώρο της υγείας να παρέχει δυνατότητες οπτικοποίησης της διαδικασίας εξαγωγής αποτελεσμάτων οι οποίες να ενισχύουν τη διαδικασία λήψης κλινικών αποφάσεων [54]. Το πακέτο `shap` προσφέρει έναν εύκολο τρόπο γραφικής παράστασης των εκτιμήσεων των SHAP τιμών, ώστε να παράγονται γραφήματα μέσω των οποίων μπορεί να οπτικοποιηθεί η επίδραση κάθε χαρακτηριστικού στις τελικές αποφάσεις των μοντέλων. Συγκεκριμένα, ο `treeExplainer` για ταξινομητές πολλαπλών κλάσεων παρέχει δυνατότητα αναπαράστασης στο ίδιο γράφημα του μέτρου της επίδρασης κάθε χαρακτηριστικού στην απόφαση ταξινόμησης σε κάθε κλάση και, για κάθε κλάση, ξεχωριστά τη δυνατότητα παραγωγής `swarm plots` τα οποία απεικονίζουν τόσο την ποσοτική όσο και την ποιοτική επίδραση της τιμής του κάθε χαρακτηριστικού στο δείγμα στις τελικές αποφάσεις του μοντέλου για την αντίστοιχη κλάση.

Εκτός των τεχνικών ενίσχυσης της επεξηγησιμότητας και ερμηνευσιμότητας των μοντέλων, η εμπιστοσύνη στις προβλέψεις τους μπορεί να αυξηθεί με την εισαγωγή ενός τρόπου μετάφρασης της βεβαιότητας των προβλέψεων, δηλαδή της πιθανότητας να είναι ακριβείς. Στον χώρο της υγείας, όπου οι αποφάσεις μπορεί να καθορίσουν την εξέλιξη μίας θεραπείας, η σημασία της ποσοτικοποίησης της βεβαιότητας εντείνεται. Αν ένα μοντέλο μπορεί να υπολογίσει τη βεβαιότητά του σε κάθε πρόβλεψη που κάνει, τότε δίνεται η ευκαιρία στους ασθενείς και τους επαγγελματίες στον χώρο της υγείας να λάβουν ασφαλείς, αποτελεσματικές αποφάσεις βάσει των δεδομένων, αποκομίζοντας τα οφέλη των εξόδων του μοντέλου, είτε αγνοώντας, είτε επανεξετάζοντας τις προβλέψεις με μεγάλη αβεβαιότητα [69].

Στο άρθρο των Bukowski et al. [70] αναπτύσσεται μία μεθοδολογία για την εξαγωγή του επιπέδου βεβαιότητας (`level of confidence`) για κάθε πρόβλεψη ενός ταξινομητή τριών κλάσεων. Συγκεκριμένα, χρησιμοποιούνται οι ενσωματωμένες μέθοδοι των ταξινομητών για την εξαγωγή πιθανοτήτων να ανήκει ένα δείγμα σε κάθε κλάση και μετασχηματίζονται σε έναν αριθμό από το 0 (πλήρης αβεβαιότητα) μέχρι το 1 (απόλυτη σιγουριά), ο οποίος υποδεικνύει τη βεβαιότητα (`confidence`) για κάθε πρόβλεψη. Στο συγκεκριμένο δημοσίευμα, εκπαιδεύεται ένα συνελκτικό νευρωνικό δίκτυο και οι πιθανότητες να ανήκει κάθε δείγμα σε κάθε κλάση υπολογίζονται μέσω της συνάρτησης `softmax`. Ωστόσο, η μεθοδολογία μπορεί να γενικευτεί για οποιονδήποτε τύπο μοντέλου ταξινομητή πολλαπλών κλάσεων από τον οποίο μπορούν να εξαχθούν οι πιθανότητες ταξινόμησης σε κάθε κλάση, καθώς και για οποιονδήποτε αριθμό κλάσεων. Η μοναδική προϋπόθεση είναι να είναι διαθέσιμο για κάθε πρόβλεψη ένα διάνυσμα του οποίου οι τιμές να είναι ίσες με τις πιθανότητες ταξινόμησης του δείγματος σε κάθε κλάση.

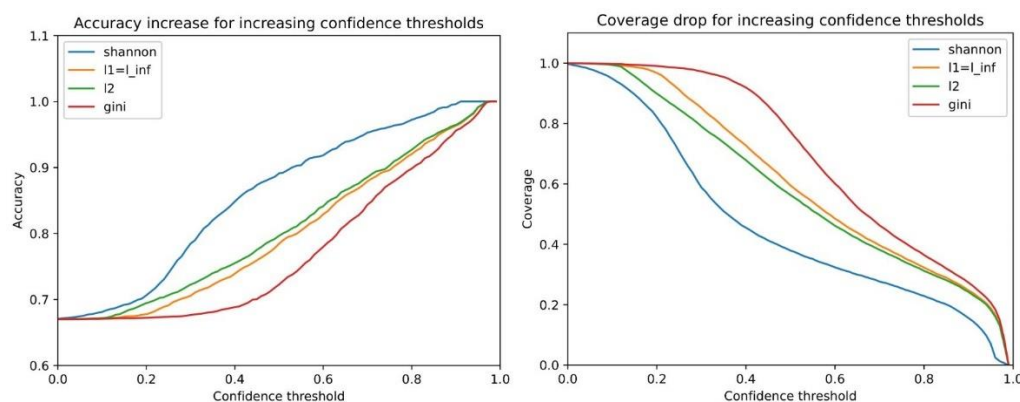
Συγκεκριμένα, εξάγεται το επίπεδο σιγουριάς κάθε πρόβλεψης σύμφωνα με τέσσερις διαφορετικές συναρτήσεις μετασχηματισμού του διανύσματος των πιθανοτήτων του

ταξινομητή σε ένα βαθμωτό επίπεδο βεβαιότητας από 0 έως 1. Όταν το ενδεχόμενο ταξινόμησης σε κάθε κλάση είναι ισοπίθανο, τότε η πρόβλεψη είναι αβέβαιη και οι συναρτήσεις μετασχηματίζουν το διάνυσμα πιθανοτήτων στο μηδενικό επίπεδο βεβαιότητας. Αντιθέτως, όταν η πιθανότητα ταξινόμησης σε μία κλάση είναι 100% και άρα 0% για τις υπόλοιπες, τότε το επίπεδο βεβαιότητας είναι 1. Αναλόγως με το πόσο κοντά είναι μεταξύ τους οι τιμές του διανύσματος πιθανοτήτων στην περίπτωση οι συναρτήσεις το μετασχηματίζουν σε ένα αντίστοιχο επίπεδο βεβαιότητας [70].

Χρησιμοποιήθηκαν οι εξής συναρτήσεις βεβαιότητας (βλ. Υποενότητα 4.1.5):

- **Βεβαιότητα βασισμένη στην εντροπία κατά Shannon.** Η εντροπία αποτελεί μέτρο της συστημικής ή πληροφοριακής αταξίας. Στην περίπτωση του διανύσματος πιθανοτήτων, όσο μεγαλύτερη είναι η ανισότητα μεταξύ των επιμέρους τιμών του, τόσο μικρότερη είναι η τιμή της εντροπίας του [71]. Όσο μικρότερη είναι η εντροπία του διανύσματος, τόσο μεγαλύτερο είναι το επίπεδο βεβαιότητας.
- **Βεβαιότητα βασισμένη στον συντελεστή Gini.** Ο συντελεστής Gini είχε αρχικά εισαχθεί ως μέτρο της ανισότητας εισοδήματος, ωστόσο έχει χρησιμοποιηθεί σε πληθώρα πεδίων στα οποία απαιτείται να ποσοτικοποιηθεί η ανισότητα μεταξύ ενός συνόλου αριθμών [72]. Μεγαλύτερη ανισότητα μεταξύ των τιμών του διανύσματος πιθανοτήτων οδηγεί σε μεγαλύτερες τιμές του συντελεστή Gini και μεγαλύτερο επίπεδο βεβαιότητας.
- **Βεβαιότητα βασισμένη στις νόρμες επιπέδου 1 και 2.** Η πιο απλή συνάρτηση, η οποία υπολογίζει το επίπεδο βεβαιότητας με βάση την απόσταση του διανύσματος πιθανοτήτων από το μοναδιαίο διάνυσμα. Η απόσταση υπολογίζεται ως το άθροισμα των στοιχείων του διανύσματος διαφοράς για το επίπεδο 1 ή την τετραγωνική ρίζα του αθροίσματος του τετραγώνου των στοιχείων του διανύσματος διαφοράς για το επίπεδο 2.

Οι συγγραφείς του [70] αξιολογούν τις συναρτήσεις βεβαιότητας και βρίσκουν ότι ανάλογα με την εφαρμογή μπορεί να διαφέρει το ποια είναι η πλέον κατάλληλη από αυτές. Ο τρόπος αξιολόγησης περιλαμβάνει τον υπολογισμό των μετρικών απόδοσης του μοντέλου για το υποσύνολο των προβλέψεων οι οποίες ξεπερνούν κάποιο κατώφλι βεβαιότητας, καθώς και σύγκριση των μετρικών απόδοσης για διαφορετικές συναρτήσεις βεβαιότητας και τιμές κατωφλίου.



Σχήμα 3.3: Διαγράμματα της ακρίβειας και της κάλυψης του συνόλου δειγμάτων συναρτήσει του κατωφλίου βεβαιότητας για το μοντέλο ταξινόμησης σε τρεις κλάσεις του [70]. Τα επίπεδα βεβαιότητας καθορίζονται σύμφωνα με τις 4 συναρτήσεις βεβαιότητας (Shannon, Gini, νόρμα επιπέδου 1, νόρμα επιπέδου 2). Για όλες τις συναρτήσεις, όσο αυξάνεται η τιμή του κατωφλίου μειώνεται η κάλυψη δείγματος και αυξάνεται η ακρίβεια. Πηγή: [70]

3.3. ΕΝΙΣΧΥΤΙΚΗ ΜΑΘΗΣΗ

Σε αυτήν την υποενότητα, καλύπτονται κάποια προαπαιτούμενα σχετικά με τα βασικά της ενισχυτικής μάθησης.

Η ενισχυτική μάθηση (Reinforcement Learning) αποτελεί υποκατηγορία της μηχανικής μάθησης και ένα εργαλείο με συγκεκριμένο στόχο στο οποίο ένας παράγοντας (agent) ή λήπτης αποφάσεων μαθαίνει μία πολιτική για να βελτιστοποιήσει μία μακροπρόθεσμη αριθμητική ανταμοιβή (reward) μέσω της αλληλεπίδρασής του με το περιβάλλον. Σε κάθε βήμα, ο παράγοντας ενισχυτικής μάθησης λαμβάνει μια αξιολόγηση σχετικά με την απόδοση της ενέργειας (action) που επέλεξε, επιτρέποντάς του να βελτιώσει τις επιδόσεις του στην επιλογή των επόμενων ενεργειών [73].

Η ενισχυτική μάθηση αποσκοπεί στην αναγνώριση μίας σειράς από ενέργειες ώστε να αυξηθεί η πιθανότητα να επιτευχθεί ένας προκαθορισμένος στόχος. Είναι μία τεχνική για την ανάπτυξη ισχυρών λύσεων για πληθώρα πεδίων στον τομέα της υγείας, στα οποία η διάγνωση ή η θεραπεία χαρακτηρίζονται από ακολουθιακές διαδικασίες λήψης αποφάσεων. Ένας παράγοντας (agent) ενισχυτικής μάθησης αλληλεπιδρά με ένα περιβάλλον με στόχο τη μεγιστοποίηση της συνολικής ανταμοιβής ως απόρροια των ενεργειών του [74].

Γενικά, τα προβλήματα ενισχυτικής μάθησης μοντελοποιούνται και επιλύονται μέσω διαδικασιών Markov (Markov Decision Process – MDP) και αποτελούνται από τα εξής 4 συστατικά:

- Την κατάσταση (state): η οποία αντιπροσωπεύει το περιβάλλον σε κάθε χρονική στιγμή.
- Την ενέργεια (action): την οποία λαμβάνει ο παράγοντας σε κάθε χρονική στιγμή και επηρεάζει την επόμενη κατάσταση.
- Την πιθανότητα μετάβασης (transition probability): η οποία παρέχει πληροφορίες σχετικά με την μετάβαση σε μεταγενέστερες καταστάσεις και αντικατοπτρίζει το περιβάλλον με το οποίο αλληλεπιδρά ο παράγοντας.
- Τη συνάρτηση ανταμοιβής (reward function): η οποία παρέχει δεδομένα ανάδρασης (feedback) σχετικά με το πόσο επιθυμητή είναι η τρέχουσα κατάσταση. [74], [75]

Με άλλα λόγια, τα προβλήματα ενισχυτικής μάθησης αποτελούνται από πολλές χρονικές στιγμές (time steps). Σε κάθε χρονική στιγμή, ο παράγοντας βρίσκεται σε μία συγκεκριμένη κατάσταση και καλείται να επιλέξει κάποια ενέργεια. Η ενέργεια που θα επιλέξει σε συνδυασμό με την κατάσταση στην οποία βρίσκεται θα καθορίσει την κατάσταση στην οποία θα βρεθεί την επόμενη χρονική στιγμή, σύμφωνα με τις πιθανότητες μετάβασης, καθώς και την ανταμοιβή που θα λάβει την τρέχουσα χρονική στιγμή. Στόχο της διαδικασίας της ενισχυτικής μάθησης αποτελεί η εύρεση της πολιτικής (policy), δηλαδή μίας συνάρτησης επιλογής ενεργειών δεδομένης της τρέχουσα κατάσταση, η οποία μεγιστοποιεί μακροπρόθεσμα τη συνολική ανταμοιβή.

Σχεδόν όλοι οι αλγόριθμοι ενισχυτικής μάθησης περιλαμβάνουν κάποιας μορφής συνάρτηση αξίας (value function) οι οποία εκτιμά το πόσο καλό είναι για τον παράγοντα να βρίσκεται σε μια δεδομένη κατάσταση ή να επιλέξει μια δεδομένη ενέργεια σε μια δεδομένη κατάσταση. Η αξία μιας κατάστασης ή ενός ζεύγους κατάστασης-ενέργειας καθορίζεται από την εκτιμώμενη μελλοντική ανταμοιβή από εκείνη την κατάσταση και έπειτα [73], [76].

Δεδομένης μιας πολιτικής π , μιας κατάστασης s_t και μίας ενέργειας a_t , ορίζεται η συνάρτηση αξίας του ζεύγους κατάστασης-ενέργειας:

$$q_{\pi}(s_t, a_t) = \sum_{t'=t}^T \mathbb{E}_{\pi}[r(s_{t'}, a_{t'}) | s_t, a_t]$$

Εξίσωση 3.1: Ορισμός της συνάρτησης αξίας ενέργειας (action-value function), η οποία αποτελεί μια εκτίμηση της αξίας ενός ζεύγους κατάστασης-ενέργειας. Το $r(s_{t'}, a_{t'})$ είναι η ανταμοιβή που λαμβάνεται αν στην κατάσταση $s_{t'}$ ληφθεί η ενέργεια $a_{t'}$.

Η συνάρτηση αξίας $q_{\pi}(s_t, a_t)$ ισούται με την αναμενόμενη συνολική ανταμοιβή μέχρι κάποιο χρονικό ορίζοντα T , αν ακολουθούνται οι ενέργειες της πολιτικής π και ξεκινώντας από την κατάσταση s_t , λαμβάνοντας την ενέργεια a_t [73].

$$v_{\pi}(s_t) = \mathbb{E}_{a_t \sim \pi(a_t, s_t)}[q_{\pi}(s_t, a_t)]$$

Εξίσωση 3.2: Ορισμός της συνάρτησης αξίας (value function), η οποία αποτελεί μια εκτίμηση της αξίας μιας κατάστασης.

Αντίστοιχα, η συνάρτηση αξίας $v_{\pi}(s_t)$ ισούται με την αναμενόμενη τιμή της $q_{\pi}(s_t, a_t)$, αν a_t η ενέργεια που υπαγορεύει η πολιτική π [73].

Αν είναι γνωστές οι ακριβείς τιμές κάποιας από τις συναρτήσεις αξίας, η εύρεση της βέλτιστης πολιτικής είναι εύκολη: αρκεί σε κάθε κατάσταση να επιλεγεί η ενέργεια η οποία μεγιστοποιεί τη συνάρτηση αξίας [73].

Στην πράξη, οι τιμές των v_{π} και q_{π} εκτιμώνται από τα δεδομένα. Η πραγματική ανταμοιβή που λαμβάνεται σε μία κατάσταση ή μετά την επιλογή μίας ενέργειας σε μία κατάσταση χρησιμοποιείται για την ενημέρωση των τιμών των συναρτήσεων αξίας. Ο τρόπος με τον οποίο αξιοποιούνται τα δεδομένα για την εκτίμηση των συναρτήσεων αξίας διαφέρει ανάλογα με τον χρησιμοποιούμενο αλγόριθμο ενισχυτικής μάθησης [76].

Το πιο διαδεδομένο είδος ενισχυτικής μάθησης είναι η διαδικτυακή ή σε σύνδεση (online) ενισχυτική μάθηση, στην οποία ο παράγοντας μπορεί να αλληλεπιδρά με το περιβάλλον. Σε κάθε χρονική στιγμή, είναι γνωστή η τρέχουσα κατάσταση και ο παράγοντας επιλέγει μία ενέργεια, η οποία πραγματοποιείται στο πραγματικό περιβάλλον ή σε κάποιο μοντέλο του, από όπου επιστρέφεται η ανταμοιβή για τη λήψη της συγκεκριμένης ενέργειας στη συγκεκριμένη κατάσταση. Η τιμή της επιστρεφόμενης ανταμοιβής χρησιμοποιείται για την αξιολόγηση της λήψης της εν λόγω ενέργειας στην τρέχουσα κατάσταση και η πληροφορία αυτή αξιοποιείται σε μεταγενέστερες χρονικές στιγμές.

Από την άλλη, σε πολλές εφαρμογές στον χώρο της υγείας δεν καθίσταται δυνατή η αλληλεπίδραση με το πραγματικό περιβάλλον, εξαιτίας του γεγονότος ότι οι ενέργειες μπορούν να επηρεάσουν την υγεία των ασθενών, ενώ η πολυπλοκότητα των εφαρμογών μπορεί να καταστήσει αδύνατη την ανάπτυξη ακριβών μοντέλων. Επομένως, για εφαρμογές στον κλάδο της υγείας προτιμάται η ενισχυτική μάθηση εκτός σύνδεσης ή σε δεσμίδες (offline/batch reinforcement learning) [75], η οποία επιτρέπει την εκμάθηση πολιτικών από σταθερά σύνολα δεδομένων.

Οι τεχνικές ενισχυτικής μάθησης μπορούν να χωριστούν σε δύο ακόμα κατηγορίες: τις τεχνικές εντός πολιτικής (on-policy) και εκτός πολιτικής (off-policy). Στις on-policy τεχνικές, εφαρμόζεται μια πολιτική και χρησιμοποιείται το feedback από το περιβάλλον για τη βελτίωση της εφαρμοζόμενης πολιτικής. Αντίθετα, στην off-policy ενισχυτική μάθηση,

χρησιμοποιούνται οι αποκρίσεις του περιβάλλοντος σε μία εφαρμοζόμενη πολιτική για την ανάπτυξη μίας πολιτικής η οποία μπορεί να είναι διαφορετική από αυτήν που εφαρμόστηκε [73], [76]. Προφανώς, στην offline ενισχυτική μάθηση, κατά την οποία υπάρχουν δεδομένα για συγκεκριμένες μόνο πολιτικές, μπορούν να εφαρμοστούν μόνο off-policy μέθοδοι.

Το Q-learning αποτελεί μία off-policy μέθοδο χρονικής διαφοράς (temporal difference – TD). Σε αντίθεση με τις μεθόδους Monte Carlo, οι οποίες περιμένουν μέχρι το τέλος ενός επεισοδίου για την ενημέρωση των συναρτήσεων αξίας, οι μέθοδοι χρονικής διαφοράς πραγματοποιούν κάποια χρήσιμη ενημέρωση σε κάθε χρονικό βήμα. Ένα χαρακτηριστικό των μεθόδων χρονικής διαφοράς είναι ότι πραγματοποιούν χτίσιμο εκ των ενόντων (bootstrapping), το οποίο σημαίνει ότι χρησιμοποιούν ήδη υπάρχουσες εκτιμήσεις για να υπολογίσουν νέες εκτιμήσεις. Όπως υποδεικνύεται από το όνομά του, το Q-learning χρησιμοποιεί τη συνάρτηση αξίας ζεύγους κατάστασης-ενέργειας, την οποία κάθε χρονική στιγμή ενημερώνει σύμφωνα με τον τύπο:

$$q(s, a) \leftarrow q(s, a) + \alpha \left[r(s, a) + \gamma \max_{a' \in A} q(s', a') - q(s, a) \right]$$

Εξίσωση 3.3: Συνάρτηση που περιγράφει την ενημέρωση (update) της συνάρτησης αξίας ενέργειας σε κάθε χρονική στιγμή κατά την εφαρμογή του Q-learning.

Όπου, α ο ρυθμός μάθησης (learning rate) και γ ο παράγοντας έκπτωσης (discount factor), υπερπαραμέτροι καθοριζόμενες από τον προγραμματιστή. Ο ρυθμός μάθησης αποτελεί μέτρο του πόσο γρήγορα μαθαίνει ο παράγοντας (agent), μεγαλύτερες τιμές αντιστοιχούν σε μεγαλύτερες ενημερώσεις των εκτιμήσεων ανά επανάληψη. Μικρότερες τιμές του α καθυστερούν τη σύγκλιση, αλλά μπορούν να οδηγήσουν σε ακριβέστερα μοντέλα. Από την άλλη, ο παράγοντας έκπτωσης αποτελεί μέτρο της δοθείσας προτεραιότητας στην άμεση ανταμοιβή έναντι της μακροπρόθεσμης. Όσο πιο κοντά στο 1 είναι το γ τόσο περισσότερο δίνεται έμφαση στη συνολική μελλοντική ανταμοιβή, ενώ όσο το γ προσεγγίζει το 0 τόσο περισσότερο «μυωπικό» γίνεται το μοντέλο, δηλαδή επιλέγει την άμεση ανταμοιβή έναντι μιας πιθανώς υψηλότερης μελλοντικής [76].

Κατά την εφαρμογή του Q-learning πρέπει να είναι γνωστά τέσσερα πράγματα σε κάθε χρονική στιγμή: η τρέχουσα κατάσταση s , η ληφθείσα ενέργεια a από την τρέχουσα κατάσταση, η επόμενη κατάσταση s' και η ανταμοιβή $r(s, a)$ που ακολούθησε την επιλογή της ενέργειας a όταν ο παράγοντας βρισκόταν στην κατάσταση s . Η προηγούμενη τιμή της συνάρτησης αξίας $q(s, a)$ ενημερώνεται με την προσθήκη του όρου $\alpha \left[r(s, a) + \gamma \max_{a' \in A} q(s', a') - q(s, a) \right]$, ο οποίος αποτελεί το γινόμενο του ρυθμού μάθησης α με το άθροισμα της πραγματικής ανταμοιβής $r(s, a)$ συν την εκτίμηση της συνάρτησης αξίας $q(s', a')$ δεδομένου ότι θα ληφθεί η καλύτερη ενέργεια a' με βάση την τρέχουσα μορφή του πίνακα $q(\cdot, \cdot)$ (αυτός όρος πολλαπλασιάζεται με τον παράγοντα έκπτωσης γ) μείον την τρέχουσα τιμή της συνάρτησης αξίας. Με άλλα λόγια, σε κάθε χρονικό βήμα, για τα σχετικά s και a , πραγματοποιείται μία μετατόπιση της τιμής της εκτίμησης $q(s, a)$ προς την τιμή $r(s, a) + \max_{a' \in A} q(s', a')$, με το μέτρο της μετατόπισης να καθορίζεται από τις υπερπαραμέτρους α και γ . Αξιοποιείται σε κάθε βήμα η πραγματική τιμή της ανταμοιβής $r(s, a)$, αλλά και το γεγονός ότι ακολούθησε η κατάσταση s' , ούτως ώστε η συνάρτηση αξίας $q(s, a)$ να προσεγγίσει την πραγματική αξία για το ζεύγος (s, a) [76].

4. ΑΝΑΠΤΥΞΗ ΜΟΝΤΕΛΟΥ ΕΚΤΙΜΗΣΗΣ ΤΗΣ ΣΥΜΜΟΡΦΩΣΗΣ ΣΤΟ ΠΛΑΝΟ ΠΑΡΕΜΒΑΣΗΣ

Το πρώτο σκέλος της εργασίας αφορά την πρόβλεψη της συμμόρφωσης στο πλάνο παρέμβασης μέσα από μοντέλα μηχανικής μάθησης. Συγκεκριμένα, χρησιμοποιούνται τα δεδομένα που ήταν διαθέσιμα για κάθε χρήστη κατά την έναρξη της μελέτης (baseline data), τα οποία είναι δημογραφικά, οι απαντήσεις στα ερωτηματολόγια και τα αποτελέσματα των κλινικών και εργαστηριακών εξετάσεων ως είσοδοι σε ταξινομητές για την πρόβλεψη της συμμόρφωσης στις ιατρικές συστάσεις. Η συμμόρφωση εκτιμάται από τη συχνότητα χρήσης της εφαρμογής και του ιχνηλάτη σωματικής δραστηριότητας. Χρησιμοποιήθηκαν δύο είδη ταξινομητών επιβλεπόμενης μάθησης: ο XGBoost του πακέτου xgboost, ο οποίος υλοποιήθηκε σε γλώσσα python και ο Optimal Tree Classifier του πακέτου interpretableai, ο οποίος υλοποιήθηκε σε R. Για καθένα από τα δύο είδη ταξινομητών, αναπτύχθηκαν τρία μοντέλα πρόβλεψης: ένα για τη συνολική συμμόρφωση, ένα για τη συμμόρφωση με την εφαρμογή και ένα για τη συμμόρφωση με τον ιχνηλάτη δραστηριότητας.

Δεδομένου του όγκου διαθέσιμων πληροφοριών για κάθε χρήστη, πραγματοποιήθηκε διαδικασία επιλογής χαρακτηριστικών πριν την εκπαίδευση κάθε μοντέλου. Προέκυψε ότι διαφορετικά σύνολα χαρακτηριστικών επιτυγχάνουν βέλτιστες αποδόσεις για διαφορετικούς τύπους ταξινομητών και για διαφορετικό ζητούμενο πρόβλεψης. Το ακριβές σύνολο χαρακτηριστικών που χρησιμοποιήθηκε, αλλά και το πόσο συνεισέφερε κάθε χαρακτηριστικό σε κάθε πρόβλεψη, οπτικοποιήθηκε μέσω του Tree Explainer του πακέτου shap της python. Ακολούθως, η επίδραση κάθε χαρακτηριστικού στις αποφάσεις των μοντέλων επεξηγήθηκε αναλυτικά.

Για κάθε πρόβλεψη των ταξινομητών υπολογίστηκε το επίπεδο βεβαιότητας του ταξινομητή για τις προβλέψεις σύμφωνα με τέσσερις συναρτήσεις. Η σχέση μεταξύ επιπέδου βεβαιότητας και επιτυχίας των προβλέψεων οπτικοποιήθηκε επίσης.

4.1. ΜΕΘΟΔΟΛΟΓΙΑ

Στην τρέχουσα υποενότητα παρουσιάζεται η μεθοδολογία που ακολουθήθηκε για την εκπαίδευση των μοντέλων πρόβλεψης της συμμόρφωσης.

4.1.1. Δεδομένα

Το σύνολο των δεδομένων προέρχεται από τις προπilotικές και πιλοτικές δοκιμές του προγράμματος παρέμβασης ENDORSE [10]–[13]. Το πρόγραμμα ρύθμισης σωματικού βάρους ENDORSE βασίζεται στη χρήση μίας ολοκληρωμένης ηλεκτρονικής πλατφόρμας η οποία αποτελείται από τα εξής:

- i. Ιχνηλάτες σωματικής δραστηριότητας, συγκεκριμένα Fitbit ACE 2 for kids, οι οποίοι δίνονται στα παιδιά που συμμετέχουν.
- ii. Ηλεκτρονικό παιχνίδι σοβαρού σκοπού, εκπαιδευτικού χαρακτήρα, σχεδιασμένο για τις ανάγκες της μελέτης.
- iii. Ειδικά διαμορφωμένη εφαρμογή κινητού για τις μητέρες, η οποία χρησιμοποιείται μεταξύ άλλων για την απάντηση ερωτηματολογίων, για την παρακολούθηση των στόχων της μελέτης και του βάρους του παιδιού και για την επικοινωνία με την κλινική ομάδα.

- iv. Ειδικά διαμορφωμένη εφαρμογή κινητού για την κλινική ομάδα, με αντίστοιχες λειτουργίες όπως η εφαρμογή για τις μητέρες.

Κατά την έναρξη του προγράμματος παρέμβασης, για κάθε συμμετέχοντα έλαβε χώρα πλήρης κλινική εξέταση από παιδοενδοκρινολόγο, καθώς και αιμοληψία. Επίσης, απαντήθηκαν από τις μητέρες ερωτηματολόγια, με στόχο την κοινωνικοδημογραφική και ψυχολογική εκτίμηση των μητέρων και των παιδιών. Συγκεκριμένα:

- i. Ερωτηματολόγιο κοινωνικοδημογραφικών δεδομένων [11].
- ii. Ερωτηματολόγιο Διατροφικών Συνηθειών του παιδιού και άλλων Συμπεριφορών Υγείας [11].
- iii. Ερωτηματολόγιο γονεϊκού τύπου (Parenting Styles and Dimensions Questionnaire – PSDQ), για τις μητέρες.
- iv. Ερωτηματολόγιο γονεϊκών πρακτικών σίτισης (Feeding Practices Questionnaire) [77], [78].
- v. Η γονεϊκή εκδοχή του ερωτηματολογίου Strengths and Difficulties Questionnaire (SDQ), που αφορά τα δυνατά σημεία και τις δυσκολίες των παιδιών [79], [80].
- vi. Το Patient Health Questionnaire 9 (PHQ – 9), για την ανίχνευση της καταθλιπτικής διάθεσης στις μητέρες [81].
- vii. Το Eating Attitudes Test (EAT26), ως κλίμακα των στάσεων διατροφής των μητέρων [82].

Τα αποτελέσματα των κλινικών και αιματολογικών εξετάσεων, όπως και οι απαντήσεις των ερωτηματολογίων ήταν διαθέσιμα και χρησιμοποιήθηκαν ως δεδομένα εισόδου για τα μοντέλα μηχανικής μάθησης.

Αναλυτικότερα, τα δεδομένα προέρχονται από τις προπilotικές και τις πιλοτικές δοκιμές του προγράμματος ENDORSE. Στις προπilotικές δοκιμές συμμετείχαν 20 παιδιά ηλικίας 6-14 ετών στα οποία παρασχέθηκαν ο ιχνηλάτης σωματικής δραστηριότητας και το ηλεκτρονικό παιχνίδι σοβαρού σκοπού, ενώ οι μητέρες τους είχαν πρόσβαση στην ειδικά διαμορφωμένη εφαρμογή του ENDORSE.

Στις πιλοτικές δοκιμές συμμετείχαν 30 παιδιά ηλικίας 6-14 ετών τα οποία χωρίστηκαν σε δύο ομάδες, την ομάδα ελέγχου και την ομάδα παρέμβασης. Στην ομάδα παρέμβασης δόθηκε πρόσβαση στον ιχνηλάτη σωματικής δραστηριότητας, το παιχνίδι και την εφαρμογή. Η ομάδα ελέγχου δεν είχε πρόσβαση στο παιχνίδι, ενώ δεν λάμβανε αυτοματοποιημένα προτρεπτικά μηνύματα μέσω της εφαρμογής.

4.1.2. Συμμόρφωση στο Πλάνο Παρέμβασης

Η εκτίμηση της συμμόρφωσης βασίστηκε στη χρήση της εφαρμογής από τις μητέρες για την καθημερινή αυτοπαρακολούθηση και την εβδομαδιαία καταγραφή βάρους, όπως και στη χρήση του ιχνηλάτη σωματικής δραστηριότητας από τα παιδιά για την καταγραφή των ημερησίων βημάτων και ωρών ύπνου. Να σημειωθεί ότι ο βαθμός χρησιμοποίησης του παιχνιδιού ήταν πολύ μικρός για τη συντριπτική πλειοψηφία των παιδιών, οπότε δεν συνυπολογίστηκε στην παρούσα εργασία. Τα επίπεδα συμμόρφωσης ορίστηκαν ως εξής.

Η συμμόρφωση στην εφαρμογή (mobile app adherence score) ορίστηκε ως το μέγιστο του συνόλου των ημερών κατά τις οποίες εισήχθησαν καταγραφές αυτοπαρακολούθησης (self-monitoring) και του συνόλου των εβδομάδων στις οποίες εισήχθησαν καταγραφές βάρους (weight monitoring):

$$\text{adherence_score_mobile_app} = \max(7 \cdot \text{weight_monitoring}, \text{self_monitoring})$$

Εξίσωση 4.1: Υπολογισμός του σκορ συμμόρφωσης με την εφαρμογή.

Εφαρμόστηκαν κατώφλια για την ταξινόμηση των παιδιών σε επίπεδα συμμόρφωσης με την εφαρμογή ανάλογα με το adherence_score_mobile_app, όπως φαίνεται στον Πίνακα 4.1:

Πίνακας 4.1: Αντιστοίχιση χρήσης εφαρμογής κινητού με επίπεδο συμμόρφωσης.

Βαθμολογία συμμόρφωσης με την εφαρμογή	Επίπεδο συμμόρφωσης με την εφαρμογή
adherence_score < 25 ημέρες	Χαμηλό
25 ημέρες < adherence_score < 50 ημέρες	Μεσαίο
adherence_score > 50 ημέρες	Υψηλό

Η συμμόρφωση με τον ιχνηλάτη δραστηριότητας Fitbit (fitbit adherence score) ορίστηκε ως το γινόμενο του μέσου όρου χρήσης ανά ημέρα σε ώρες (daily average usage hours) επί του συνόλου των ημερών που χρησιμοποιήθηκε (total days of usage) δια 10 ώρες:

$$\text{adherence_score_fitbit} = \frac{\text{daily_average_usage_hours} \cdot \text{total_days_of_usage}}{10 \text{ hours}}$$

Εξίσωση 4.2: Υπολογισμός του σκορ συμμόρφωσης με το Fitbit.

Αντίστοιχα κατώφλια εφαρμόστηκαν και για τη χρήση του ιχνηλάτη δραστηριότητας. Επειδή στους χρήστες της προπilotικής μελέτης δόθηκε το Fitbit για μεγαλύτερο χρονικό διάστημα σε σχέση με αυτούς της πιλοτικής (120 έναντι 84 ημερών), εφαρμόστηκαν μεγαλύτερα κατώφλια για τα αντίστοιχα επίπεδα συμμόρφωσης (40 και 80 ημέρες έναντι 25 και 50 ημερών αντίστοιχα):

Πίνακας 4.2: Αντιστοίχιση χρήσης ιχνηλάτη δραστηριότητας (Fitbit) με επίπεδο συμμόρφωσης.

Βαθμολογία συμμόρφωσης με το Fitbit για τους χρήστες της προπilotικής μελέτης	Βαθμολογία συμμόρφωσης με το Fitbit για τους χρήστες της πιλοτικής μελέτης	Επίπεδο συμμόρφωσης με το Fitbit
adherence_score < 40 ημέρες	adherence_score < 25 ημέρες	Χαμηλό
40 ημέρες < adherence_score < 80 ημέρες	25 ημέρες < adherence_score < 50 ημέρες	Μεσαίο
adherence_score > 80 ημέρες	adherence_score > 50 ημέρες	Υψηλό

Τα επίπεδα συμμόρφωσης με την εφαρμογή και το Fitbit κωδικοποιήθηκαν με τους αριθμούς 1, 2, 3 για τα επίπεδα χαμηλό, μεσαίο και υψηλό αντίστοιχα. Ο βαθμός της συνολικής συμμόρφωσης (overall adherence) ορίστηκε ως το άθροισμα των βαθμών συμμόρφωσης των επιμέρους στοιχείων σύμφωνα με την παραπάνω κωδικοποίηση:

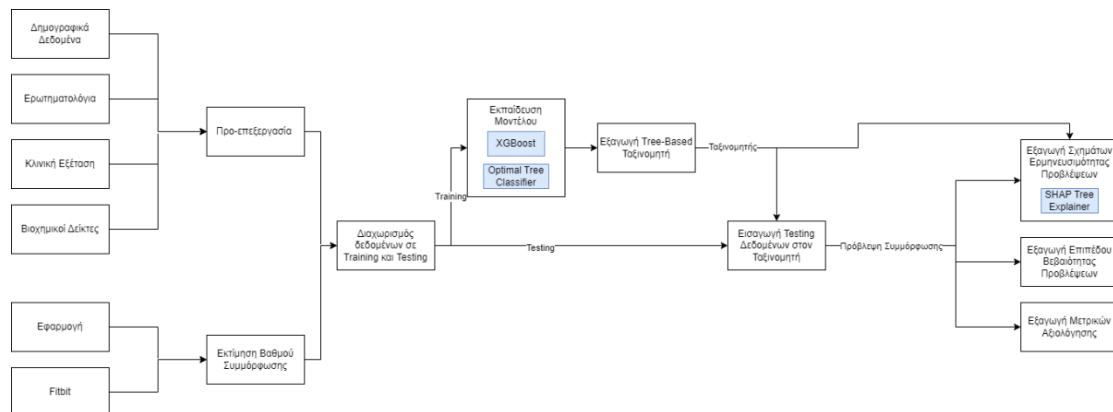
$$\text{overall_adherence} = \text{adherence_level_mobile_app} + \text{adherence_level_fitbit}$$

Εξίσωση 4.3: Υπολογισμός του συνολικού βαθμού συμμόρφωσης ως άθροισμα των επιπέδων συμμόρφωσης με την εφαρμογή (adherence level mobile app) και με τον Fitbit (adherence level fitbit).

Το συνολικός βαθμός συμμόρφωσης χωρίστηκε και αυτός σε τρία επίπεδα. Στο χαμηλό επίπεδο κατατάχθηκαν οι χρήστες με συνολικό βαθμό συμμόρφωσης 2, στο μεσαίο επίπεδο οι χρήστες με βαθμό συμμόρφωση 3 ή 4 και υψηλό επίπεδο οι χρήστες με βαθμό συμμόρφωσης 5 ή 6.

4.1.3. Περιγραφή της Διαδικασίας

Διερευνήθηκαν δύο μοντέλα μηχανικής μάθησης για την πρόβλεψη της συμμόρφωσης με το ENDORSE: το XGBoost και το InterpretableAI. Τα μοντέλα λαμβάνουν ως είσοδο τα δημογραφικά δεδομένα και τα δεδομένα από τα ερωτηματολόγια, την κλινική και την εργαστηριακή εξέταση, τα οποία είναι διαθέσιμα κατά την έναρξη του πλάνου παρέμβασης, για κάθε παιδί και έχουν ως έξοδο τον προβλεπόμενο βαθμό συμμόρφωσης του παιδιού αυτού.



Σχήμα 4.1: Εκπαίδευση, αξιολόγηση και ερμηνεία των μοντέλων επιβλεπόμενης μάθησης για την πρόβλεψη του βαθμού συμμόρφωσης στο πλάνο παρέμβασης.

Τα δεδομένα που συλλέχθηκαν στο πλαίσιο της μελέτης του ENDORSE διέπονται από μεγάλη ποικιλομορφία. Οι τύποι των δεδομένων διαφέρουν (αριθμοί, κείμενο, κατηγορικές μεταβλητές), ενώ, όπως σε πολλές εφαρμογές στον χώρο της υγείας, υπήρχαν ελλείψεις. Ως αποτέλεσμα, χρειάστηκε τα δεδομένα να υποστούν ένα στάδιο προ-επεξεργασίας, ούτως ώστε να είναι συμβατά με τα μοντέλα.

Ένα μοντέλο μηχανικής μάθησης, για να μπορεί να κάνει προβλέψεις, χρειάζεται πρώτα να εκπαιδευτεί. Τα δεδομένα χωρίζονται σε δεδομένα εκπαίδευσης (training data) και δεδομένα ελέγχου (testing data). Στο μοντέλο εισάγονται τα δεδομένα εκπαίδευσης με στόχο το μοντέλο να μάθει να αναγνωρίζει κάποια συσχέτιση μεταξύ των τιμών των δεδομένων εισόδου και της ζητούμενης εξόδου. Για παράδειγμα, στην παρούσα εργασία, εισάγονται τα δεδομένα από τα ερωτηματολόγια και τις εξετάσεις μαζί με τον βαθμό συμμόρφωσης προκειμένου το μοντέλο να αναγνωρίσει μοτίβα στα δεδομένα εισόδου τα οποία οδηγούν στους διαφορετικούς βαθμούς συμμόρφωσης. Τα δεδομένα ελέγχου χρησιμοποιούνται για την αξιολόγηση της απόδοσης του μοντέλου. Πχ. στο ήδη εκπαιδευμένο μοντέλο της εργασίας εισάγονται τα δεδομένα εισόδου από το σύνολο των δεδομένων ελέγχου και συγκρίνονται οι προβλέψεις για τους χρήστες των δεδομένων ελέγχου με τους πραγματικούς βαθμούς συμμόρφωσης.

Τα δεδομένα αντιστοιχούν σε μικρό αριθμό χρηστών ($n = 50$). Ως εκ τούτου, για να μην λείπουν σημαντικές πληροφορίες στη διαδικασία εκπαίδευσης του μοντέλου, εφαρμόστηκε η μέθοδος της διασταυρωμένης επικύρωσης (cross validation) για τον έλεγχο της απόδοσης του μοντέλου. Συγκεκριμένα, ακολουθήθηκε η διαδικασία leave-one-out cross validation, δηλαδή εκπαιδεύτηκαν 50 μοντέλα, με κάθε δείγμα να αποτελεί το σύνολο δεδομένων ελέγχου (μονοσύνολο) ακριβώς ενός μοντέλου το οποίο είχε ως σύνολο δεδομένων εκπαίδευσης τα υπόλοιπα 49 δείγματα. Με αυτόν τον τρόπο, δοκιμάστηκαν οι προβλέψεις σε όλα τα δείγματα, με αποτέλεσμα την αυξημένη εγκυρότητα των δεικτών απόδοσης σε σχέση με το να λαμβάνονταν από ένα μέρος των διαθέσιμων δειγμάτων.

Κάθε φορά που εκπαιδεύονται τα μοντέλα των ταξινομητών πολλαπλών κλάσεων (multi-class classifier) εξάγονταν κάποιες μετρικές για την αξιολόγηση της απόδοσής τους. Αναλυτικότερα, υπολογίστηκαν το area under the receiver operation characteristic curve (roc-auc score) για κάθε κλάση ξεχωριστά και ο μέσος όρος τους, το οποίο αποτελεί μέτρο της διακριτικής ικανότητας ενός ταξινομητή ανεξαρτήτων των κατωφλίων που ορίζονται για την ταξινόμηση των δειγμάτων, η ακρίβεια (accuracy), η οποία ορίζεται ως το ποσοστό των επιτυχών προβλέψεων, καθώς και τα specificity, precision, recall και f1-score για κάθε κλάση.

Συνολικά, αναπτύχθηκαν 6 μοντέλα. Ένα για κάθε έναν από τους ταξινομητές XGBoost και Optimal Tree Classifier (από το πακέτο InterpretableAI) για κάθε ένα από τα τρία εξαγόμενα αποτελέσματα: συνολική συμμόρφωση, συμμόρφωση με την εφαρμογή και συμμόρφωση με το Fitbit. Σε κάθε περίπτωση, επιλέχθηκαν τα χαρακτηριστικά και οι τιμές των υπερπαραμέτρων που βελτιστοποιούσαν την απόδοση του μοντέλου σύμφωνα με τις εξαγόμενες μετρικές.

Για κάθε πρόβλεψη παραγόμενη από τα μοντέλα, υπολογίστηκε το επίπεδο βεβαιότητας (level of confidence), ένας αριθμός ο οποίος ποσοτικοποιεί τη σιγουριά του μοντέλου ότι έχει κάνει τη σωστή πρόβλεψη.

Τέλος, δόθηκε έμφαση στην ερμηνευσιμότητα των μοντέλων. Συγκεκριμένα, χρησιμοποιήθηκε ο Tree Explainer από το δημοφιλές πακέτο της rython shar το οποίο χρησιμοποιεί μια εκτίμηση των shapley τιμών για να ποσοτικοποιήσει την επίδραση κάθε χαρακτηριστικού σε κάθε πρόβλεψη του μοντέλου. Μέσω του shar παράχθηκαν διαγράμματα τα οποία χρησιμοποιήθηκαν για την ερμηνεία των μοντέλων και την ανάλυση του προβλήματος της συμμόρφωσης.

4.1.4. Διαδικασία Εκπαίδευσης και Αξιολόγησης

Σε αυτήν την ενότητα περιγράφεται αναλυτικά η διαδικασία εκπαίδευσης και επικύρωσης των μοντέλων.

Τα διαθέσιμα δεδομένα δεν ήταν εξαρχής σε μορφή συμβατή με τα πακέτα υλοποίησης των μοντέλων μηχανικής μάθησης. Προκειμένου να εκπαιδευτούν τα μοντέλα χρειαζόταν τα δεδομένα εισόδου να είναι σε αριθμητική μορφή. Οι διαδικασίες προ-επεξεργασίας των δεδομένων περιλάμβαναν μεταξύ άλλων:

- Κωδικοποίηση δεδομένων που είχαν μορφή κειμένου.
- One hot encoding των κατηγορικών δεδομένων.
- Αφαίρεση από το σύνολο δεδομένων των επί μέρους ερωτήσεων των ερωτηματολογίων και διατήρηση μόνο των προκυπτουσών βαθμολογιών.
- Έλεγχο της ορθότητας των δεδομένων.

Να σημειωθεί ότι οι ταξινομητές XGBoost και Optimal Tree Classifier μπορούν να διαχειρίζονται από μόνοι τους τα απόντα δεδομένα (missing data), επομένως δεν απαιτήθηκε κάποια διαδικασία συμπερασμού (imputation).

Για να προκύψουν τα σύνολα χαρακτηριστικών για τα μοντέλα δοκιμάστηκε πληθώρα και συνδυασμός μεθόδων και από τις τρεις οικογένειες: φιλτραρίσματος, καλύμματος και ενσωματωμένες.

Για το XGBoost, χρησιμοποιήθηκε συνδυασμός wrapper και embedded μεθόδων για την επιλογή των χαρακτηριστικών που θα εισάγονταν στους ταξινομητές για τη συνολική

συμμόρφωση και τη συμμόρφωση με την εφαρμογή. Το σύνολο των χαρακτηριστικών για την εκπαίδευση του ταξινομητή για τη συμμόρφωση με το Fitbit προέκυψε εύκολα με Sequential Feature Selection.

Για το Optimal Tree Classifier, για κάθε ζητούμενο (είδος συμμόρφωσης) δοκιμάστηκε το σύνολο χαρακτηριστικών που προέκυψε για το αντίστοιχο μοντέλο XGBoost. Για κανένα ζητούμενο δεν προέκυψαν αξιόπιστα μοντέλα. Στη συνέχεια, δοκιμάστηκε Sequential Feature Selection για κάθε ζητούμενο και τα μοντέλα που εκπαιδεύτηκαν με τα προκύπτοντα σύνολα χαρακτηριστικών είχαν ικανοποιητικές επιδόσεις.

Η αξιολόγηση των μοντέλων που προέκυψαν κατά τη διαδικασία του Sequential Feature Selection έγινε κατά κύριο λόγο με βάση την ακρίβεια και κατά δεύτερο με βάση το μέσο area under the curve.

Για κάθε μοντέλο που εκπαιδεύονταν, υπολογίστηκαν κάποιες μετρικές οι οποίες χρησιμοποιήθηκαν για την αξιολόγηση της απόδοσης του μοντέλου. Συγκεκριμένα:

- Accuracy (ακρίβεια): το ποσοστό των επιτυχών προβλέψεων.
- Sensitivity ή recall (ευαισθησία): η πιθανότητα το μοντέλο να προβλέψει ότι ένα δείγμα ανήκει σε μία κλάση δεδομένου ότι το δείγμα ανήκει στην κλάση αυτή.
- Specificity (ειδικότητα): η πιθανότητα το μοντέλο να προβλέψει ότι ένα δείγμα δεν ανήκει σε μία κλάση δεδομένου ότι το δείγμα δεν ανήκει στην κλάση αυτή.
- Positive predictive value (PPV) ή precision (θετική διαγνωστική αξία): η πιθανότητα ένα δείγμα να ανήκει σε μία κλάση δεδομένου ότι το μοντέλο προέβλεψε ότι ανήκει στην κλάση αυτή.
- Area under the curve: ορίζεται ως το εμβαδόν κάτω από την καμπύλη που προκύπτει αν στον έναν άξονα είναι το sensitivity και στον άλλον το specificity για διαφορετικές τιμές κατωφλίου απόφασης για κάθε κλάση.
- f1-score: ο αρμονικός μέσος όρος των precision και recall.

Όλες οι μετρικές, εκτός από το accuracy, υπολογίστηκαν ξεχωριστά για κάθε κλάση (επίπεδο συμμόρφωσης). Για κάθε μετρική υπολογίστηκαν ο μέσος όρος της μεταξύ των τριών κλάσεων, αλλά και ο σταθμισμένος μέσος όρος της με βάση τις συχνότητες των τριών κλάσεων σε κάθε σύνολο δεδομένων.

4.1.5. Επίπεδο Βεβαιότητας

Στην παρούσα εργασία, κρίθηκε σκόπιμο με κάθε πρόβλεψη να εξάγεται και ένας αριθμός ο οποίος υποδεικνύει τη βεβαιότητα για κάθε συμπέρασμα του μοντέλου. Δεδομένου ότι το ζητούμενο αποτελούσε πρόβλημα ταξινόμησης πολλαπλών κλάσεων, ακολουθήθηκε η μεθοδολογία των Bukowski et al. [70].

Αναλυτικότερα, τα μοντέλα που χρησιμοποιήθηκαν, όταν προβλέπουν τη συμμόρφωση ενός δείγματος, υπολογίζουν τις πιθανότητες αυτό το δείγμα να ανήκει σε κάθε κλάση. Ως εκ τούτου, προκύπτει ένα διάνυσμα, το οποίο αποτελείται από τρεις μεταβλητές με άθροισμα 1. Για παράδειγμα, αν ο ταξινομητής αποφανθεί ότι ένα δείγμα έχει 50% πιθανότητα να ανήκει στην κλάση χαμηλής συμμόρφωσης, 30% στην κλάση μεσαίας συμμόρφωσης και 20% στην κλάση υψηλής συμμόρφωσης, τότε λαμβάνεται το διάνυσμα [0.5 0.3 0.2].

Το ζητούμενο είναι κάθε πρόβλεψη να συνοδεύεται από ένα βαθμωτό ποσοστό βεβαιότητας, πχ. το επιθυμητό στο παραπάνω παράδειγμα είναι το μοντέλο να ταξινομήσει το δείγμα στην κλάση χαμηλής συμμόρφωσης και να παράσχει ένα ποσοστό για το πόσο έμπιστη μπορεί να

θεωρηθεί η πρόβλεψή του. Δεν θα μπορούσε να ληφθεί μόνο η πιθανότητα που δίνει ο ταξινομητής να ανήκει το δείγμα στην πιο πιθανή κλάση, διότι τότε θα παραλείπονταν πληροφορίες σχετικές με τη βεβαιότητα της πρόβλεψης. Δηλαδή, το διάνυσμα πιθανοτήτων $[0.5 \ 0.3 \ 0.2]$ δεν μπορεί να θεωρηθεί το ίδιο με το διάνυσμα $[0.5 \ 0.45 \ 0.05]$.

Επομένως, στο [70] χρησιμοποιούνται τέσσερις συναρτήσεις που λαμβάνουν ως είσοδο το διάνυσμα πιθανοτήτων και το μετατρέπουν σε ένα εύχρηστο και εύληπτο βαθμωτό επίπεδο βεβαιότητας. Όλες οι συναρτήσεις λαμβάνουν ως είσοδο ένα διάνυσμα μήκους 3 με άθροισμα στοιχείων 1 και το μετατρέπουν σε έναν αριθμό μεταξύ 0 και 1. Το επίπεδο βεβαιότητας είναι 0, όταν οι πιθανότητες ταξινόμησης είναι ίσες για όλες τις κλάσεις, δηλαδή στο διάνυσμα $[\frac{1}{3} \ \frac{1}{3} \ \frac{1}{3}]$ και 1, όταν οι πιθανότητες ταξινόμησης σε μία κλάση είναι 100% και 0% για τις άλλες, δηλαδή στα διανύσματα $[1 \ 0 \ 0]$, $[0 \ 1 \ 0]$ και $[0 \ 0 \ 1]$.

Συγκεκριμένα, χρησιμοποιήθηκαν για διάνυσμα εισόδου $\mathbf{v} = [p_0 \ p_1 \ p_2]$:

- Η βεβαιότητα βασισμένη στην εντροπία κατά Shannon (Shannon-Based Confidence) [71]:

$$Conf_{Shannon}(\mathbf{v}) = 1.0 + \sum_{i=0}^2 p_i \cdot \log_3 p_i$$

Εξίσωση 4.4: Συνάρτηση βεβαιότητας βασισμένης στην εντροπία κατά Shannon.

- Η βεβαιότητα βασισμένη στον συντελεστή Gini (Gini-Based Confidence) [72]:

$$Conf_{Gini}(\mathbf{v}) = \frac{Gini(\mathbf{v})}{Gini([0 \ 0 \ 1])}$$

Εξίσωση 4.5: Συνάρτηση βεβαιότητας βασισμένης στον συντελεστή Gini.

όπου

$$Gini(\mathbf{v}) = \frac{\sum_{i=0}^2 2(i-1)p_i'}{3 \cdot \sum_{i=0}^2 p_i'}$$

Εξίσωση 4.6: Συνάρτηση υπολογισμού του συντελεστή Gini από ένα διάνυσμα \mathbf{v} . Οι τόνοι στα p_i' υποδηλώνουν τα στοιχεία του \mathbf{v} ταξινομημένα σε αύξουσα σειρά, δηλαδή $p_0' = \min(p_0, p_1, p_2)$, $p_2' = \max(p_0, p_1, p_2)$ και $p_1 \leq p_2$.

- Η βεβαιότητα βασισμένη στη νόρμα (Norm-Based Confidence).
Επιπέδου 1:

$$Conf_{l1-norm}(\mathbf{v}) = \max_{0 \leq i \leq 2} \left(1 - \frac{\|\mathbf{v} - [1_i \ 0 \ 0]\|}{\|[1 \ 0 \ 0] - [\frac{1}{3} \ \frac{1}{3} \ \frac{1}{3}]\|} \right)$$

Εξίσωση 4.7: Συνάρτηση βεβαιότητας βασισμένης στη νόρμα επιπέδου 1.

Επιπέδου 2:

$$Conf_{l2-norm}(\mathbf{v}) = \max_{0 \leq i \leq 2} \left(1 - \frac{\|\mathbf{v} - [1_i \ 0 \ 0]\|_2}{\|[1 \ 0 \ 0]_2 - [\frac{1}{3} \ \frac{1}{3} \ \frac{1}{3}]\|_2} \right)$$

Εξίσωση 4.8: Συνάρτηση βεβαιότητας βασισμένης στη νόρμα επιπέδου 2.

Όπου ο συμβολισμός $[1_i \ 0 \ 0]$ για $0 \leq i \leq 2$ σημαίνει ότι επιλέγεται το διάνυσμα από τα $[1 \ 0 \ 0]$, $[0 \ 1 \ 0]$ και $[0 \ 0 \ 1]$, που μεγιστοποιεί την αριθμητική παράσταση.

Εφαρμόζοντας τις συναρτήσεις προκύπτουν τα επίπεδα βεβαιότητας των ταξινομητών για κάθε πρόβλεψη που κάνουν.

Πραγματοποιήθηκε ανάλογη ανάλυση με το [70] για τον προσδιορισμό του καλύτερου δείκτη βεβαιότητας. Συγκεκριμένα, δοκιμάστηκαν διαφορετικά επίπεδα βεβαιότητας ως κατώφλια για τον καθορισμό υποσυνόλων των δειγμάτων, τα οποία αποτελούνται από δείγματα για τα οποία το μοντέλο μπορεί να κάνει πρόβλεψη με βεβαιότητα πάνω από το κατώφλι. Για κάθε μοντέλο και ζητούμενο παράχθηκαν διαγράμματα τα οποία απεικονίζουν την ακρίβεια, δηλαδή το ποσοστό των σωστών προβλέψεων από αυτές με βεβαιότητα πάνω από το εκάστοτε κατώφλι, και την κάλυψη δείγματος (sample coverage), δηλαδή το πλήθος των προβλέψεων των οποίων η βεβαιότητα ξεπερνά το κατώφλι, συναρτήσει του κατωφλίου που επιλέγεται, όπως ακριβώς στο [70]. Επιπλέον, παράχθηκαν διαγράμματα της ακρίβειας συναρτήσει της κάλυψης.

Δηλαδή για μία τιμή κατωφλίου t ορίζονται:

Η κάλυψη δείγματος:

$$\text{sample_coverage}(t) = N_{\text{προβλέψεων με επίπεδο βεβαιότητας} \geq t}$$

Εξίσωση 4.9: Ορισμός της κάλυψης δείγματος συναρτήσει ενός κατωφλίου βεβαιότητας t . Η κάλυψη δείγματος είναι το πλήθος των προβλέψεων των οποίων το επίπεδο βεβαιότητας ξεπερνά ένα κατώφλι. Η κάλυψη δείγματος για ένα κατώφλι t μπορεί να διαφέρει για διαφορετικές συναρτήσεις ποσοτικοποίησης βεβαιότητας.

Η ακρίβεια συναρτήσει του t :

$$\text{accuracy}(t) = \frac{N_{\text{σωστών προβλέψεων με επίπεδο βεβαιότητας} \geq t}}{\text{sample_coverage}(t)}$$

Εξίσωση 4.10: Ορισμός της ακρίβειας συναρτήσεις ενός κατωφλίου βεβαιότητας t . Η ακρίβεια είναι ο λόγος των σωστών προβλέψεων μεταξύ εκείνων των οποίων το επίπεδο βεβαιότητας ξεπερνά το t . Η ακρίβεια για ένα κατώφλι t μπορεί να διαφέρει για διαφορετικές συναρτήσεις ποσοτικοποίησης βεβαιότητας.

Προφανώς, για όλους τους δείκτες βεβαιότητας όσο αυξάνεται το κατώφλι η κάλυψη μειώνεται από τα n δείγματα (πλήρης κάλυψη) για κατώφλι βεβαιότητας 0 έως τα 0 δείγματα για κατώφλι βεβαιότητας 1, με εξαίρεση αν υπάρχουν δείγματα για τα οποία οι ταξινομητές προβλέπουν με 100% βεβαιότητα. Αντίστοιχα, το αναμενόμενο θα ήταν όσο αυξάνεται το κατώφλι η ακρίβεια να αυξάνεται, ξεκινώντας από το επίπεδο ακρίβειας του εκάστοτε ταξινομητή έως και 100%. Το τελευταίο δεν ισχύει στην περίπτωση της παρούσας εργασίας, πλην των ταξινομητών XGBoost για τη Συνολική Συμμόρφωση και τη Συμμόρφωση με την Εφαρμογή.

4.1.6. Ερμηνευσιμότητα

Δόθηκε έμφαση στην ερμηνευσιμότητα των μοντέλων, καθώς πραγματοποιούν προβλέψεις για τον χώρο της υγείας και επομένως είναι σημαντική η δυνατότητα κατανόησής τους από μη-ειδικούς.

Συγκεκριμένα, χρησιμοποιήθηκε ο Tree Explainer του πακέτου shap της rython, ο οποίος εκτιμά αποδοτικά τις shapley τιμές για ταξινομητές βασισμένους σε δέντρα αποφάσεων. Οι shapley τιμές υπολογίζονται για κάθε χαρακτηριστικό, δείγμα, στην προκειμένη περίπτωση,

στην οποία πραγματοποιείται ταξινόμηση πολλαπλών κλάσεων, για κάθε κλάση και για κάθε ταξινομητή ξεχωριστά. Μία shapley τιμή αποτελεί μέτρο της επίδρασης της τιμής του χαρακτηριστικού στην απόφαση να ταξινομηθεί το δείγμα στην κλάση από τον ταξινομητή.

Όντων γνωστών των εκτιμήσεων των shapley τιμών είναι δυνατό να εξαχθούν και να οπτικοποιηθούν σημαντικές πληροφορίες σχετικές με τη διαδικασία ταξινόμησης, αλλά και την εσωτερική δομή των μοντέλων.

Να σημειωθεί ότι παρόλο που ο Optimal Tree Classifier παράγει δέντρα αποφάσεων, τα οποία είναι απλά και εύκολα ερμηνεύσιμα μοντέλα, αποφασίστηκε η χρήση του shap για την ερμηνεία των προβλέψεών του. Εξαιτίας της διαδικασίας leave-one-out διασταυρωμένης επικύρωσης, η οποία έχει ως αποτέλεσμα την παραγωγή 50 διαφορετικών δέντρων απόφασης για το συγκεκριμένο δείγμα, αξιοποιήθηκε το πακέτο shap ώστε να καταστεί δυνατή η συνοπτική οπτικοποίηση του συνόλου των παραγόμενων δέντρων και να παρασχεθούν πληροφορίες για τις προβλέψεις που προέκυψαν από αυτά.

Για όλους τους ταξινομητές παρήχθησαν ραβδογράμματα στα οποία φαίνεται ποσοτικά η επίδραση κάθε χαρακτηριστικού στην απόφαση ένα δείγμα να ανήκει ή όχι σε κάθε κλάση συμμόρφωσης. Οι χρωματιστές μπάρες σε αυτά δείχνουν την επίδραση κάθε χαρακτηριστικού για κάθε κλάση, με κάθε χρώμα να αντιστοιχεί σε μία κλάση. Συγκεκριμένα, μεγαλύτερο μήκος μπάρας ενός χρώματος υποδεικνύει ότι το συγκεκριμένο χαρακτηριστικό επηρέασε περισσότερο τη διαδικασία λήψης απόφασης για το αν θα ανήκουν τα δείγματα στην αντίστοιχη κλάση. Τα χαρακτηριστικά παρουσιάζονται σε φθίνουσα σειρά συνολικής απόλυτης επίδρασης στις αποφάσεις του ταξινομητή για όλες τις κλάσεις. Να σημειωθεί ότι στα διαγράμματα αυτού του είδους παρουσιάζεται η μέση απόλυτη SHAP τιμή για κάθε χαρακτηριστικό και για κάθε κλάση, οπότε δεν φαίνεται αν το χαρακτηριστικό επηρέασε θετικά ή αρνητικά την απόφαση του μοντέλου.

Για να αντιμετωπιστεί αυτό, για κάθε ταξινομητή παρήχθησαν τρία τα swarm plots, ένα για κάθε επίπεδο συμμόρφωσης. Στα διαγράμματα αυτά φαίνεται η επίδραση (θετική ή αρνητική) που έχει η τιμή κάθε χαρακτηριστικού ενός συγκεκριμένου δείγματος στις πιθανότητες να ταξινομηθεί το δείγμα στην εκάστοτε κλάση. Κάθε κουκίδα αντιστοιχεί στην τιμή του χαρακτηριστικού για ένα δείγμα και το πόσο αυτό το χαρακτηριστικό επηρέασε την απόφαση του ταξινομητή για την εκάστοτε κλάση. Το χρώμα της κουκίδας υποδεικνύει την τιμή του χαρακτηριστικού: η μέγιστη τιμή στο σύνολο τιμών αποτυπώνεται με κόκκινο χρώμα, η ελάχιστη με μπλε και οι ενδιάμεσες με αποχρώσεις ανάμεσα στο κόκκινο και το μπλε, όπως φαίνεται και στην μπάρα στα δεξιά των διαγραμμάτων. Οι απύουσες τιμές παρουσιάζονται με γκριζο. Όσο περισσότερο επηρεάζει ένα χαρακτηριστικό θετικά την απόφαση το δείγμα να ανήκει σε μία κλάση, τόσο πιο δεξιά βρίσκεται η αντίστοιχη κουκίδα και όσο περισσότερο επηρεάζει ένα χαρακτηριστικό αρνητικά την απόφαση τόσο πιο αριστερά βρίσκεται η αντίστοιχη κουκίδα. Αν η τιμή του χαρακτηριστικού δεν επηρεάζει ούτε θετικά ούτε αρνητικά την απόφαση το δείγμα να ανήκει στην αντίστοιχη κλάση, τότε η SHAP τιμή είναι 0 και η κουκίδα βρίσκεται στην αντίστοιχη θέση.

Η χρήση swarm plots διευκολύνει την κατανόηση της διαδικασίας ταξινόμησης. Αν σε ένα διάγραμμα για την ταξινόμηση σε μία συγκεκριμένη κλάση, για ένα χαρακτηριστικό οι περισσότερες κουκίδες βρίσκονται μακριά από το μηδέν τότε το χαρακτηριστικό αυτό επηρέασε πολύ την απόφαση ταξινόμησης στην αντίστοιχη κλάση και το αντίστροφο. Αν οι περισσότερες κουκίδες που βρίσκονται δεξιά του μηδενός έχουν αποχρώσεις κοντά στο κόκκινο και οι περισσότερες κουκίδες που βρίσκονται αριστερά του μηδενός έχουν

αποχρώσεις κοντά στο μπλε, τότε υψηλότερες τιμές του χαρακτηριστικού επηρέασαν θετικά την απόφαση ταξινόμησης στην εν λόγω κλάση και χαμηλότερες τιμές του χαρακτηριστικού την επηρέασαν αρνητικά.

Για την περαιτέρω κατανόηση των μοντέλων, για κάθε ταξινομητή παρέχονται και δύο τύποι αναλυτικών πινάκων. Στον πρώτο τύπο, περιγράφονται τα χαρακτηριστικά τα οποία προέκυψαν από τη διαδικασία επιλογής χαρακτηριστικών και με τα οποία εκπαιδεύτηκαν τα μοντέλα, καθώς και οι τιμές που λαμβάνουν. Με χρήση αυτού του πίνακα μπορούν να αποκωδικοποιηθούν τα χρώματα στα swarm plots. Για παράδειγμα, η κατηγορική μεταβλητή «Ποιο είναι το μορφωτικό σας επίπεδο;» λαμβάνει επτά διακριτές τιμές από 0-6, κάθε μία από τις οποίες αντιστοιχεί σε ένα διαφορετικό μορφωτικό επίπεδο. Στο πεδίο για το συγκεκριμένο χαρακτηριστικό στο swarm plot, είναι γνωστό ότι οι κουκίδες με μπλε αποχρώσεις αντιστοιχούν στα δείγματα με χαμηλότερα μορφωτικά επίπεδα, οι κουκίδες με κόκκινες αποχρώσεις σε δείγματα με υψηλότερα μορφωτικά επίπεδα και οι γκρι αποχρώσεις σε δείγματα για τα οποία απουσίαζε η τιμή για το συγκεκριμένο πεδίο. Μία μπλε κουκίδα αρκετά δεξιά του μηδενός στο swarm plot για την κλάση χαμηλής συμμόρφωσης δείχνει ότι το χαμηλότερο μορφωτικό επίπεδο επηρέασε την απόφαση του ταξινομητή υπέρ της ταξινόμησης στην κλάση χαμηλής συμμόρφωσης.

Επιπρόσθετα, παρέχονται πίνακες ερμηνείας των swarm plots με ρητές εξηγήσεις σχετικά με την ποσοτική και ποιοτική επιρροή κάθε χαρακτηριστικού στην απόφαση να ταξινομηθεί κάθε ταξινομητής τα δείγματα σε κάθε κλάση.

4.2. ΑΠΟΤΕΛΕΣΜΑΤΑ ΚΑΙ ΣΥΖΗΤΗΣΗ

Σε αυτήν την ενότητα, παρουσιάζονται τα αποτελέσματα των τελικών μοντέλων για την πρόβλεψη της συμμόρφωσης.

4.2.1. Πρόβλεψη Συμμόρφωσης με XGBoost

Στην τρέχουσα υποενότητα, παρουσιάζονται τα μοντέλα που προέκυψαν από την εκπαίδευση του ταξινομητή XGBoost. Αναπτύχθηκαν σε γλώσσα rython τρία μοντέλα για τα τρία ζητούμενα: συνολική συμμόρφωση και επιμέρους συμμορφώσεις με την εφαρμογή και το Fitbit.

4.2.1.1. Συνολική Συμμόρφωση

Στον Πίνακα 4.3 παρουσιάζονται τα 12 χαρακτηριστικά με τα οποία εκπαιδεύτηκε το μοντέλο του XGBoost για την πρόβλεψη της συνολικής συμμόρφωσης. Για όλους τους ταξινομητές και ζητούμενα, τα χαρακτηριστικά προέκυψαν μέσω της διαδικασίας επιλογής χαρακτηριστικών.

Στον Πίνακα 4.4 παρουσιάζονται οι μετρικές που προέκυψαν μετά από leave-one-out cross validation για το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης. Το μοντέλο επιτυγχάνει ακρίβεια 78% και μέσο area under the curve 85%.

Όπως φαίνεται στο Σχήμα 4.2, το χαρακτηριστικό που επηρέασε περισσότερο την πρόβλεψη για τη συνολική συμμόρφωση είναι η ηλικία του πατρός, ενώ ακολουθούν το μορφωτικό επίπεδο της μητέρας και ο βαθμός ανταπόκρισης σε ερεθίσματα τροφής. Η ηλικία του πατρός και το μορφωτικό επίπεδο επηρέασαν σε μεγάλο βαθμό τις αποφάσεις για την ταξινόμηση των δειγμάτων στις κλάσεις χαμηλής και μεσαίας συμμόρφωσης, ενώ για την απόφαση ταξινόμησης στην κλάση υψηλής συμμόρφωσης έπαιξαν μεγαλύτερο ρόλο η νόσηση της μητέρας από κάποια χρόνια ασθένεια, τα επίπεδα HDL του παιδιού και το μήκος του σώματός του όταν γεννήθηκε.

Στον Πίνακα 4.5, περιγράφονται οι επιδράσεις των χαρακτηριστικών στο εξαγόμενο αποτέλεσμα, όπως προκύπτουν από τα swarm plots.

Στα σχήματα 4.6 και 4.7 παρουσιάζονται η ακρίβεια και η κάλυψη δείγματος συναρτήσει του κατωφλίου και η ακρίβεια συναρτήσει της κάλυψης δείγματος για διαφορετικά κατώφλια βεβαιότητας για τις τέσσερις μετρικές ποσοτικοποίησης της βεβαιότητας. Παρατηρείται ότι εξαιτίας του μικρού πλήθους δειγμάτων η σχέση μεταξύ ακρίβειας και κάλυψης δεν είναι μονότονη. Ωστόσο, για τον συγκεκριμένο ταξινομητή, η γενική τάση της ακρίβειας είναι φθίνουσα όσο αυξάνεται η κάλυψη δείγματος.

Πίνακας 4.3: Περιγραφή των χαρακτηριστικών του συνόλου δεδομένων με τα οποία εκπαιδεύτηκε το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με XGBoost.

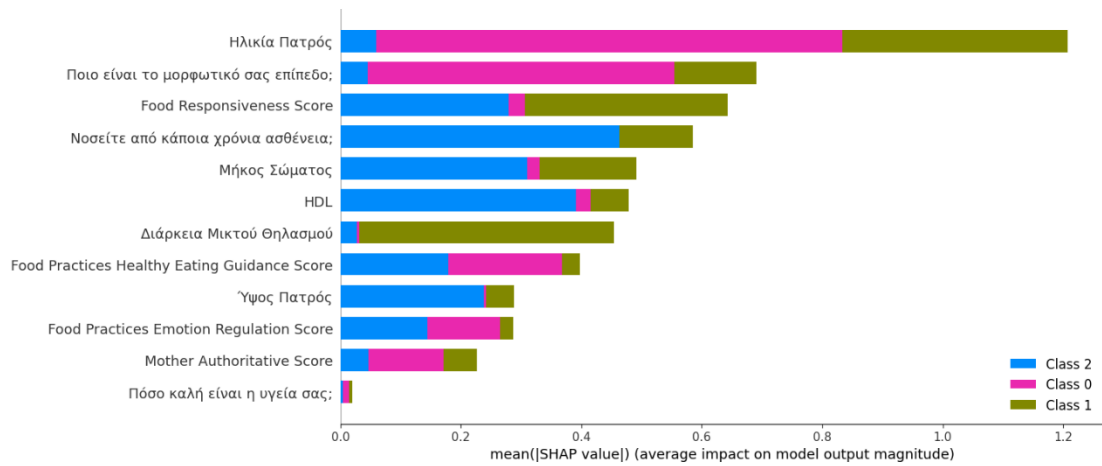
Χαρακτηριστικό	Περιγραφή	Τιμές
Ηλικία Πατρός	Ηλικία του πατρός του παιδιού κατά την έναρξη του προγράμματος παρέμβασης	Ηλικία σε έτη
Ποιο είναι το μορφωτικό σας επίπεδο;	Απάντηση της μητέρας στην ερώτηση «Ποιο είναι το μορφωτικό σας επίπεδο;»	Υψηλότερο μορφωτικό επίπεδο που έχει ολοκληρωθεί από τη μητέρα, κατά σειρά (Δημοτικό, Γυμνάσιο, Γενικό Λύκειο, ΙΕΚ, ΤΕΙ, ΑΕΙ, Κάτοχος Μεταπτυχιακού Διπλώματος), κωδικοποίηση σε αριθμούς 0-6 αντίστοιχα
Food Responsiveness Score	Βαθμολογία απόκρισης σε ερεθίσματα τροφής, όπως προκύπτει από το Ερωτηματολόγιο Διατροφικών Συνηθειών και άλλων Συμπεριφορών Υγείας	Συνεχείς τιμές στο διάστημα 1-5
Νοσείτε από κάποια χρόνια ασθένεια;	Απάντηση της μητέρας στην ερώτηση «Νοσείτε από κάποια χρόνια ασθένεια (πχ. Υπέρταση, Δυσλιπιδαιμία, Διαβήτης);»	Απάντηση ναι (0) ή όχι (1)
Μήκος Σώματος	Μήκος σώματος του παιδιού κατά τη γέννηση	Μήκος σώματος σε εκατοστά
HDL	Επίπεδο υψηλής πυκνότητας σε λιποπρωτεΐνες (High-Density Lipoprotein) χοληστερόλης. Υψηλότερο επίπεδο συνδέεται με τη μείωση των πιθανοτήτων για καρδιακές παθήσεις και εγκεφαλικό.	Χοληστερόλη HDL σε mg/dL
Διάρκεια Μικτού Θηλασμού	Διάρκεια μικτού θηλασμού σε μήνες. Δεν προσμετράται η διάρκεια αποκλειστικού θηλασμού. Οι κενές τιμές υποδεικνύουν ότι τα παιδιά δεν θήλασαν ούτε αποκλειστικά.	Αριθμός μηνών μικτού θηλασμού, το 0 υποδεικνύει ότι υπήρξε διάστημα κατά το οποίο το παιδί θήλαζε αποκλειστικά, η κενή τιμή ότι δεν θήλαζε καθόλου

Πίνακας 4.3 (συνέχεια): Περιγραφή των χαρακτηριστικών του συνόλου δεδομένων με τα οποία εκπαιδεύτηκε το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με ΧGBoost.

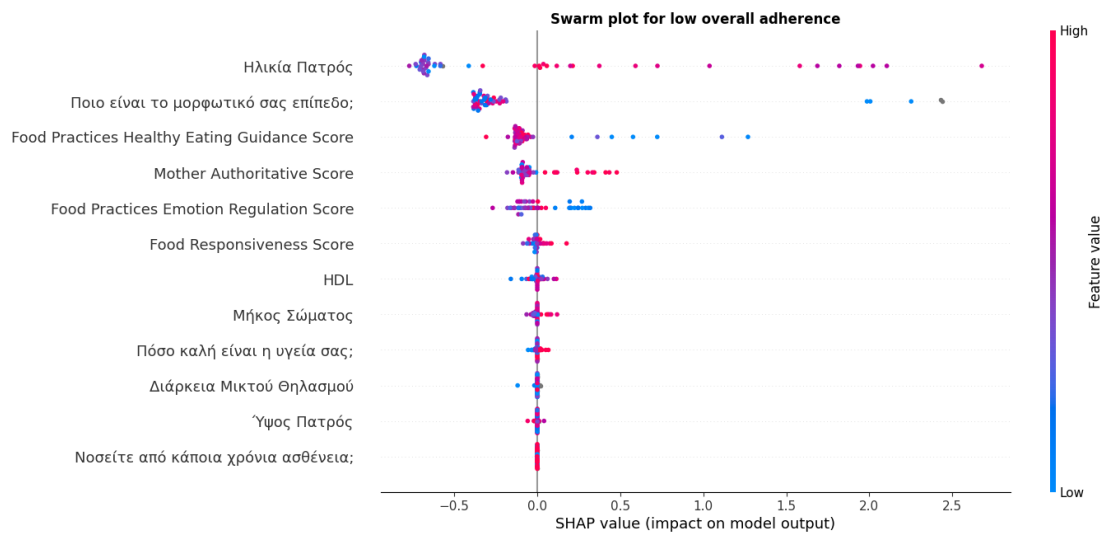
Food Practices Healthy Eating Guidance Score	Βαθμολογία καθοδήγησης υγιεινής διατροφής, όπως προκύπτει από το Ερωτηματολόγιο Γονεϊκών Πρακτικών Σίτισης	Συνεχείς τιμές στο διάστημα 1-5
Ύψος Πατρός	Ύψος του πατρός του παιδιού	Ύψος σε εκατοστά
Food Practices Emotion Regulation Score	Βαθμολογία ρύθμισης συναισθημάτων μέσω του φαγητού, όπως προκύπτει από το Ερωτηματολόγιο Γονεϊκών Πρακτικών Σίτισης	Συνεχείς τιμές στο διάστημα 1-5
Mother Authoritative Score	Βαθμολογία υποστηρικτικού τύπου μητέρας, όπως προκύπτει από το ερωτηματολόγιο γονεϊκού τύπου	Συνεχείς τιμές στο διάστημα 1-5
Πόσο καλή είναι η υγεία σας;	Απάντηση της μητέρας στην ερώτηση «Πόσο καλή είναι η υγεία σας;»	Πιθανές απαντήσεις κατά σειρά (Κακή, Μέτρια, Καλή, Πολύ καλή, Εξαιρετική), κωδικοποίηση σε αριθμούς 0-4 αντίστοιχα

Πίνακας 4.4: Μετρικές αξιολόγησης του μοντέλου πρόβλεψης της συνολικής συμμόρφωσης με ΧGBoost.

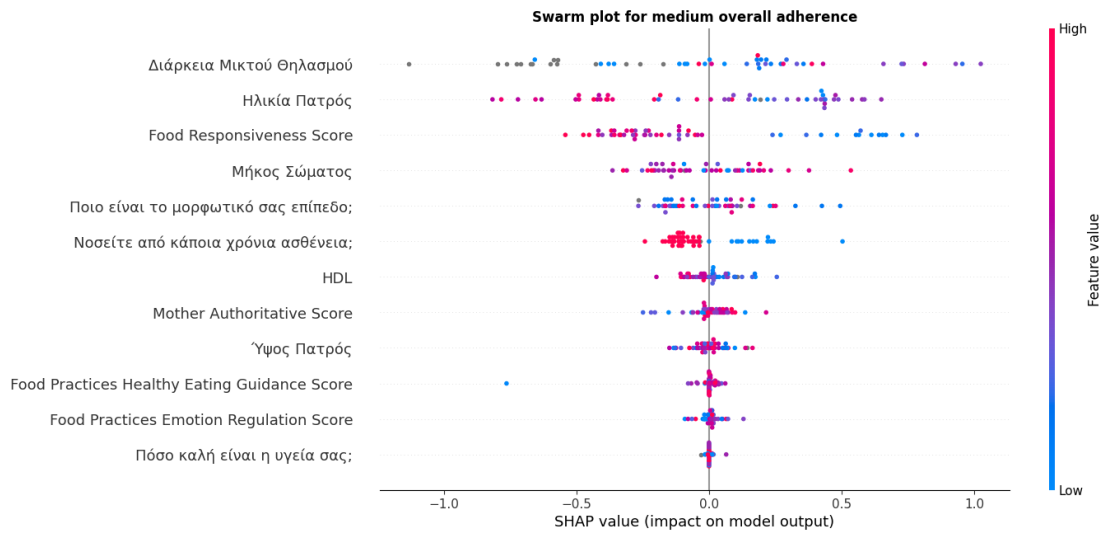
Μετρική / Επίπεδο Συμμόρφωσης	Χαμηλό	Μεσαίο	Υψηλό	Μέσος όρος	Σταθμισμένος Μέσος Όρος
Αριθμός δειγμάτων	14	20	16	—	—
ROC area under the curve	95%	83%	76%	85%	84%
Precision	80%	84%	69%	78%	78%
Recall	86%	80%	69%	78%	78%
f1-score	83%	82%	69%	78%	78%
Specificity	81%	98%	89%	89%	90%
Accuracy	78%			—	—



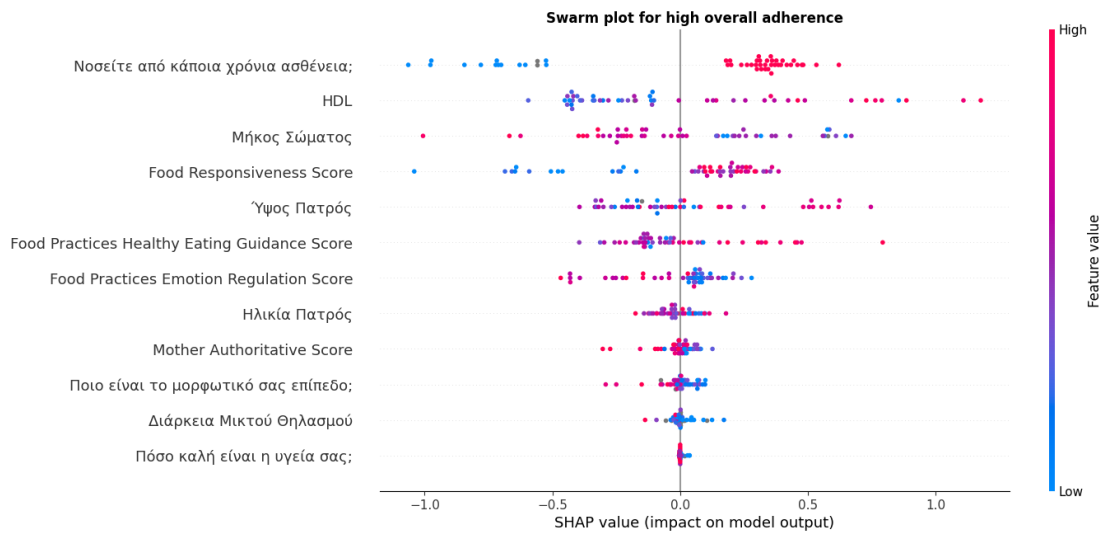
Σχήμα 4.2: Διάγραμμα ποσοτικοποίησης της μέσης επίδρασης κάθε χαρακτηριστικού στο μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με XGBoost.



Σχήμα 4.3: Swarm plot απεικόνισης της επίδρασης κάθε χαρακτηριστικού στην απόφαση ταξινόμησης κάθε δείγματος στην κλάση χαμηλής συμμόρφωσης από το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με XGBoost.



Σχήμα 4.4: Swarm plot απεικόνισης της επίδρασης κάθε χαρακτηριστικού στην απόφαση ταξινόμησης κάθε δείγματος στην κλάση μεσαίας συμμόρφωσης από το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με XGBoost.



Σχήμα 4.5: Swarm plot απεικόνισης της επίδρασης κάθε χαρακτηριστικού στην απόφαση ταξινόμησης κάθε δείγματος στην κλάση υψηλής συμμόρφωσης από το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με XGBoost

Πίνακας 4.5: Επεξήγηση των *swarm plots* για το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με XGBoost.

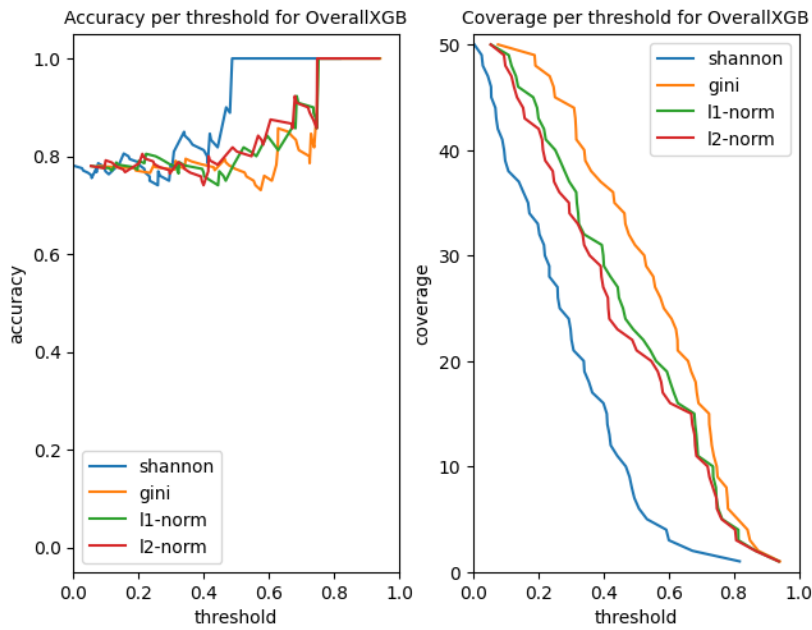
Χαρακτηριστικό / Επίδραση στο Επίπεδο Συμμόρφωσης	Χαμηλό	Μεσαίο	Υψηλό
Ηλικία Πατρός	Επηρεάζει πολύ, με τα παιδιά των οποίων οι πατέρες έχουν μεγαλύτερη ηλικία να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει αρκετά, με τα παιδιά των οποίων οι πατέρες έχουν μεσαία ή μικρή ηλικία να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και αυτά των μεγαλύτερων σε ηλικίες πατεράδων να έχουν λιγότερες	Επηρεάζει λίγο, με τα παιδιά των οποίων οι πατέρες έχουν μικρότερη ηλικία να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Ποιο είναι το μορφωτικό σας επίπεδο;	Επηρεάζει πολύ με τα παιδιά των οποίων οι μητέρες έχουν το χαμηλότερο μορφωτικό επίπεδο ή δεν απάντησαν σε αυτή την ερώτηση να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Φαίνεται να επηρεάζει αρκετά, χωρίς να υπάρχει ξεκάθαρη μονοτονική συσχέτιση	Επηρεάζει πολύ με τα παιδιά των οποίων οι μητέρες έχουν χαμηλότερο μορφωτικό επίπεδο να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Food Responsiveness Score	Επηρεάζει λίγο, με τα υψηλότερα σκορ να αυξάνουν την πιθανότητα για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει αρκετά, με τα χαμηλότερα σκορ να αυξάνουν τις πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει πολύ, με τα υψηλότερα σκορ να αυξάνουν τις πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Νοσείτε από κάποια χρόνια ασθένεια;	Δεν φαίνεται να επηρεάζει	Επηρεάζει αρκετά, με τα παιδιά των οποίων οι μητέρες νοσούν από κάποια χρόνια ασθένεια να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει πολύ, με τα παιδιά των οποίων οι μητέρες δε νοσούν από κάποια χρόνια ασθένεια να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο

Πίνακας 4.5 (συνέχεια): Επεξήγηση των swarm plots για το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με ΧGBoost.

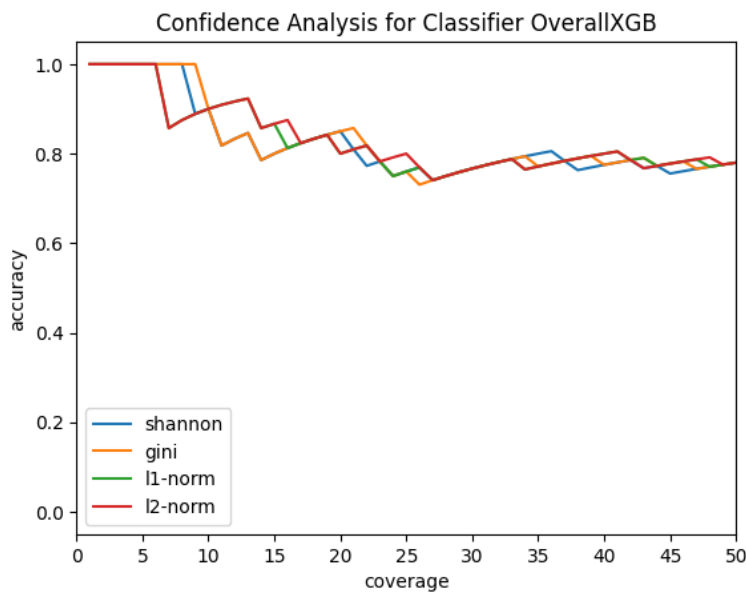
Μήκος Σώματος	Φαίνεται να επηρεάζει λίγο, με τα παιδιά που γεννήθηκαν με μεγαλύτερο μήκος να έχουν μεγαλύτερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Φαίνεται να επηρεάζει αρκετά, χωρίς να υπάρχει ξεκάθαρη μονοτονική συσχέτιση	Επηρεάζει πολύ, με τα παιδιά που γεννήθηκαν με μικρότερο μήκος να έχουν μεγαλύτερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
HDL	Φαίνεται να επηρεάζει λίγο, με τις υψηλότερες τιμές να αυξάνουν τις πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει αρκετά, με τις χαμηλότερες τιμές να αυξάνουν τις πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει πολύ, με τις υψηλότερες τιμές να αυξάνουν τις πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Διάρκεια Μικτού Θηλασμού	Δεν φαίνεται να επηρεάζει	Επηρεάζει πολύ. Τα παιδιά τα οποία θήλασαν για περισσότερους μήνες φαίνεται να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση, ενώ τα παιδιά που δεν θήλασαν καθόλου έχουν λιγότερες	Επηρεάζει ελάχιστα, με τα παιδιά τα οποία θήλασαν λιγότερο να φαίνεται να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Food Practices Healthy Eating Guidance Score	Επηρεάζει πολύ, με τις χαμηλότερες βαθμολογίες να αυξάνουν τις πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Δεν φαίνεται να επηρεάζει	Επηρεάζει αρκετά, με τις υψηλότερες βαθμολογίες να αυξάνουν τις πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο

Πίνακας 4.5 (συνέχεια): Επεξήγηση των swarm plots για το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με XGBoost.

Ύψος Πατρός	Δεν φαίνεται να επηρεάζει	Φαίνεται να επηρεάζει λίγο, χωρίς να υπάρχει ξεκάθαρη μονοτονική συσχέτιση	Φαίνεται να επηρεάζει αρκετά, με τα παιδιά των ψηλότερων πατεράδων να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση σε σχέση με τα υπόλοιπα
Food Practices Emotion Regulation Score	Επηρεάζει αρκετά, με τα χαμηλότερα σκορ να αυξάνουν τις πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Δεν φαίνεται να επηρεάζει	Επηρεάζει αρκετά, με τις χαμηλότερες βαθμολογίες να αυξάνουν τις πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Mother Authoritative Score	Επηρεάζει αρκετά, με τα παιδιά των οποίων οι μητέρες σημείωσαν τα υψηλότερα σκορ να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει αρκετά, με τις υψηλότερες τιμές να φαίνεται να αυξάνουν τις πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει λίγο, με τα παιδιά των οποίων οι μητέρες σημείωσαν τα χαμηλότερα σκορ να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Πόσο καλή είναι η υγεία σας;	Φαίνεται να επηρεάζει ελάχιστα, με τα παιδιά των οποίων οι μητέρες έχουν καλύτερη υγεία να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Δεν φαίνεται να επηρεάζει	Δεν φαίνεται να επηρεάζει ιδιαίτερα



Σχήμα 4.6: Διαγράμματα της ακρίβειας και της κάλυψης του συνόλου δειγμάτων συναρτήσει του κατωφλίου βεβαιότητας για το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με XGBoost. Τα επίπεδα βεβαιότητας καθορίζονται σύμφωνα με τις 4 συναρτήσεις βεβαιότητας (Shannon, Gini, νόρμα επιπέδου 1, νόρμα επιπέδου 2).



Σχήμα 4.7: Διάγραμμα της ακρίβειας του μοντέλου πρόβλεψης της συνολικής συμμόρφωσης με XGBoost για ένα υποσύνολο δειγμάτων συναρτήσει της κάλυψης ολόκληρου το συνόλου. Τα υποσύνολα υπολογίζονται για φθίνοντα κατώφλια βεβαιότητας σύμφωνα με τις 4 συναρτήσεις βεβαιότητας (Shannon, Gini, νόρμα επιπέδου 1, νόρμα επιπέδου 2).

4.2.1.2. Συμμόρφωση με την Εφαρμογή

Στον Πίνακα 4.6 παρουσιάζονται τα χαρακτηριστικά τα οποία προέκυψαν από τη διαδικασία επιλογής χαρακτηριστικών και με τα οποία εκπαιδεύτηκε ο ταξινομητής XGBoost για την πρόβλεψη της συμμόρφωσης με την εφαρμογή.

Ακολουθεί ο Πίνακας 4.7 στον οποίον παρουσιάζονται οι μετρικές απόδοσης του ταξινομητή. Οι επιδόσεις είναι αρκετά χαμηλές για την κλάση μεσαίας συμμόρφωσης εξαιτίας του μικρού αριθμού δειγμάτων στην εν λόγω κλάση. Ωστόσο, ο σταθμισμένος με βάση πλήθος των δειγμάτων μέσος όρος όλων των μετρικών είναι πάνω από 70%.

Το Σχήμα 4.8 είναι το ραβδόγραμμα συνολικής επίδρασης κάθε χαρακτηριστικού στο μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή. Το πιο σημαντικό χαρακτηριστικό στο μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με XGBoost ήταν το επίπεδο TG στο αίμα των παιδιών. Ακολουθούν η βαθμολογία αυταρχικού τύπου μητέρας και το ύψος του παιδιού κατά την πρώτη επίσκεψη. Ωστόσο, για την απόφαση ταξινόμησης στην κλάση χαμηλής συμμόρφωσης η νόσηση της μητέρας από κάποια χρόνια ασθένεια, με τα παιδιά των οποίων οι μητέρες νοσούν να έχουν περισσότερες πιθανότητες να ταξινομηθούν στην κλάση χαμηλής συμμόρφωσης και το αντίστροφο. Για την κλάση μεσαίας συμμόρφωσης η αναλογία περιμέτρου μέσης ανά ύψος ήταν το πλέον σημαντικό χαρακτηριστικό. Το μοντέλο ταξινομεί ευκολότερα τα παιδιά με μεγαλύτερη αναλογία περιμέτρου μέσης ανά ύψος στις κλάσεις μεσαίας και υψηλής συμμόρφωσης σε αντίθεση με την κλάση χαμηλής συμμόρφωσης. Αυτό ενδέχεται να υποδεικνύει ότι οι μητέρες των παιδιών που έχουν πιο έντονο αυτό το χαρακτηριστικό να είναι πιο ευαισθητοποιημένες να τηρήσουν τις ιατρικές συστάσεις που αφορούν την εφαρμογή.

Στον Πίνακα 4.8 περιγράφονται οι επιδράσεις των χαρακτηριστικών στο αποτέλεσμα, όπως προκύπτουν από τα swarm plots.

Ομοίως με την περίπτωση της συνολικής συμμόρφωσης, δεν παρατηρείται μονότονη σχέση μεταξύ της κάλυψης δείγματος και της ακρίβειας (Σχήμα 4.13), ωστόσο η γενική τάση της ακρίβειας συναρτήσει της κάλυψης είναι φθίνουσα.

Πίνακας 4.6: Περιγραφή των χαρακτηριστικών του συνόλου δεδομένων με τα οποία εκπαιδεύτηκε το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με XGBoost.

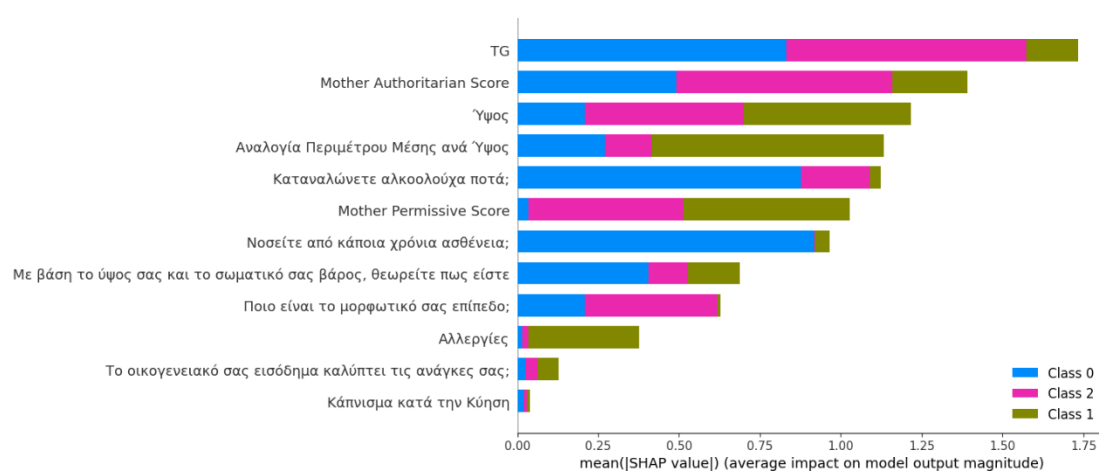
Χαρακτηριστικό	Περιγραφή	Τιμές
TG	Επίπεδο τριγλυκεριδίων. Υψηλότερα επίπεδα TG συνδέονται με παθήσεις, όπως ο διαβήτης τύπου 2, το μεταβολικό σύνδρομο, ο υποθυρεοειδισμός	Επίπεδο TG σε mg/dl
Mother Authoritarian Score	Βαθμολογία αυταρχικού τύπου μητέρας, όπως προκύπτει από το ερωτηματολόγιο γονεϊκού τύπου	Συνεχείς τιμές στο διάστημα 1-5
Ύψος	Ύψος παιδιού κατά την πρώτη επίσκεψη	Ύψος σε εκατοστά
Αναλογία Περιμέτρου Μέσης ανά Ύψος	Λόγος της περιμέτρου μέσης του παιδιού σε εκατοστά προς το ύψος του σε εκατοστά κατά την πρώτη επίσκεψη	Προκύπτων λόγος

Πίνακας 4.6 (συνέχεια): Περιγραφή των χαρακτηριστικών του συνόλου δεδομένων με τα οποία εκπαιδεύτηκε το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με XGBoost.

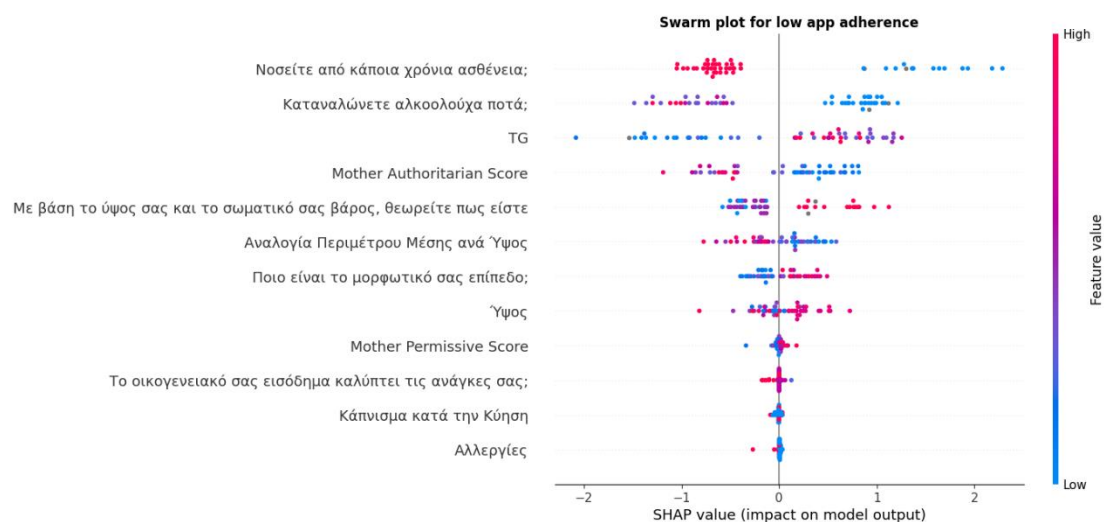
Καταναλώνετε αλκοολούχα ποτά;	Απάντηση της μητέρας στην ερώτηση «Καταναλώνετε αλκοολούχα ποτά;»	Συχνότητα κατανάλωσης αλκοολούχων ποτών από τη μητέρα, κατά σειρά (Καθόλου, 1-3 φορές/μήνα, 1 φορά/εβδομάδα, 2-4 φορές/εβδομάδα), κωδικοποίηση σε 0-3 αντίστοιχα
Mother Permissive Score	Βαθμολογία επιτρεπτικού τύπου μητέρας, όπως προκύπτει από το ερωτηματολόγιο γονεϊκού τύπου	Συνεχείς τιμές στο διάστημα 1-5
Νοσείτε από κάποια χρόνια ασθένεια;	Απάντηση της μητέρας στην ερώτηση «Νοσείτε από κάποια χρόνια ασθένεια (πχ. Υπέρταση, Δυσλιπιδαιμία, Διαβήτη);»	Απάντηση ναι (0) ή όχι (1)
Με βάση το ύψος σας και το σωματικό σας βάρος, θεωρείτε πως είστε	Απάντηση της μητέρας στην ερώτηση «Με βάση το ύψος σας και το σωματικό σας βάρος, θεωρείτε πως είστε»	Αντίληψη της μητέρας σχετικά με το σωματικό της βάρος, κατά σειρά (Φυσιολογικός/η, Υπέρβαρος/η, Παχύσαρκος/η), κωδικοποίηση σε 1-3 αντίστοιχα
Ποιο είναι το μορφωτικό σας επίπεδο;	Απάντηση της μητέρας στην ερώτηση «Ποιο είναι το μορφωτικό σας επίπεδο;»	Υψηλότερο μορφωτικό επίπεδο που έχει ολοκληρωθεί από τη μητέρα, κατά σειρά (Δημοτικό, Γυμνάσιο, Γενικό Λύκειο, ΙΕΚ, ΤΕΙ, ΑΕΙ, Κάτοχος Μεταπτυχιακού Διπλώματος), κωδικοποίηση σε αριθμούς 0-6 αντίστοιχα
Αλλεργίες	Αν το παιδί έχει αλλεργίες	Ναι (1) ή όχι (0)
Το οικογενειακό σας εισόδημα καλύπτει τις ανάγκες σας;	Απάντηση της μητέρας στην ερώτηση «Το οικογενειακό σας εισόδημα καλύπτει τις ανάγκες σας;»	Αντίληψη της μητέρας σχετικά με την ικανοποίηση των αναγκών από το οικογενειακό εισόδημα, κατά σειρά (Καθόλου, Λίγο, Μέτρια, Πολύ, Πάρα πολύ), κωδικοποίηση σε 0-4
Κάπνισμα κατά την Κύηση	Κάπνισμα της μητέρας κατά την κύηση του παιδιού	Ναι (1) ή όχι (0)

Πίνακας 4.7: Μετρικές αξιολόγησης του μοντέλου πρόβλεψης της συμμόρφωσης με την εφαρμογή με XGBoost.

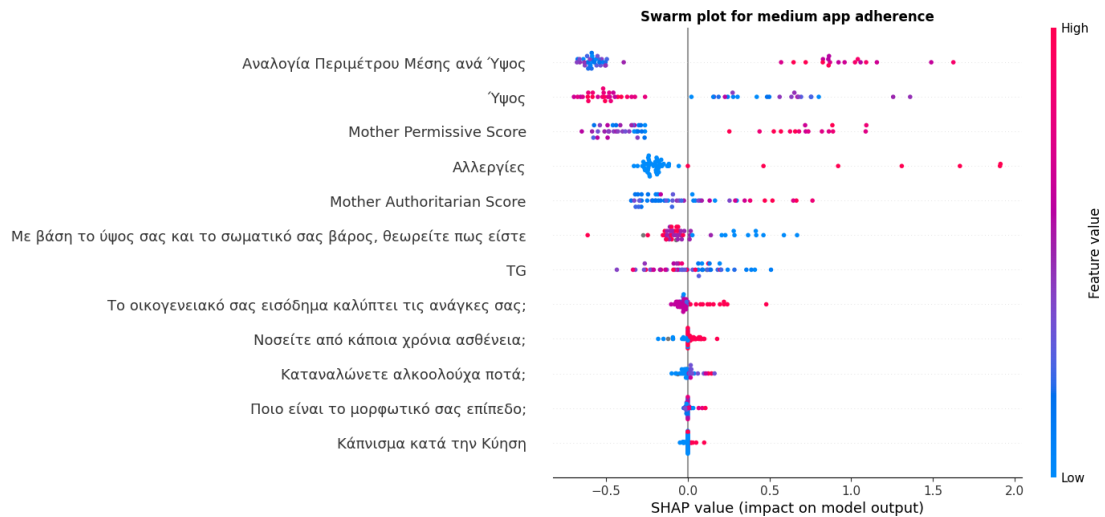
Μετρική / Επίπεδο Συμμόρφωσης	Χαμηλό	Μεσαίο	Υψηλό	Μέσος όρος	Σταθμισμένος Μέσος Όρος
Αριθμός δειγμάτων	24	9	17	—	—
ROC area under the curve	83%	79%	69%	77%	72%
Precision	81%	62%	62%	68%	71%
Recall	88%	56%	59%	68%	72%
f1-score	84%	59%	61%	68%	72%
Specificity	81%	93%	82%	85%	83%
Accuracy	72%			—	—



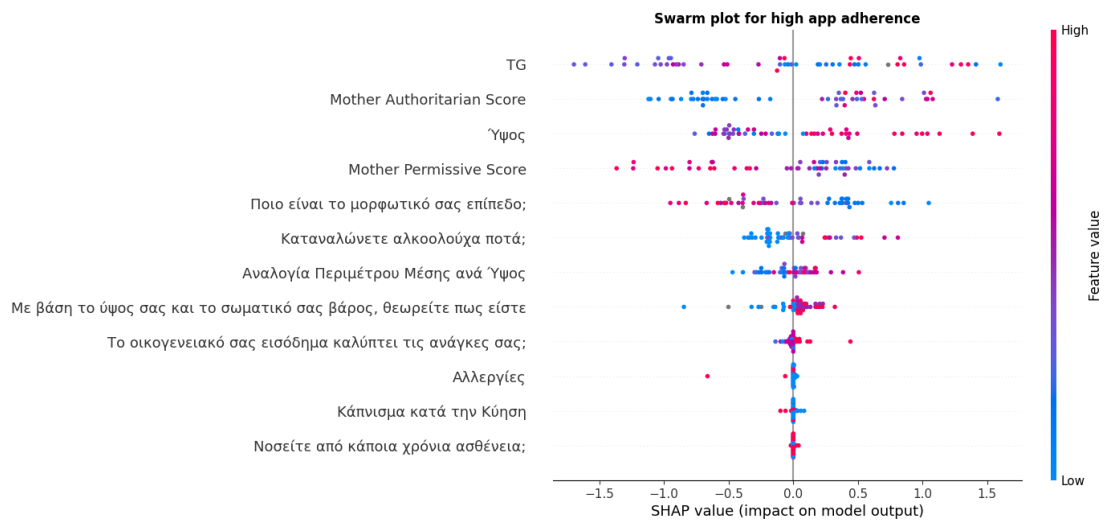
Σχήμα 4.8: Διάγραμμα ποσοτικοποίησης της μέσης επίδρασης κάθε χαρακτηριστικού στο μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με XGBoost.



Σχήμα 4.9: Swarm plot απεικόνισης της επίδρασης κάθε χαρακτηριστικού στην απόφαση ταξινόμησης κάθε δείγματος στην κλάση χαμηλής συμμόρφωσης από το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με XGBoost.



Σχήμα 4.10: Swarm plot απεικόνισης της επίδρασης κάθε χαρακτηριστικού στην απόφαση ταξινόμησης κάθε δείγματος στην κλάση μεσαίας συμμόρφωσης από το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με XGBoost.



Σχήμα 4.11: Swarm plot απεικόνισης της επίδρασης κάθε χαρακτηριστικού στην απόφαση ταξινόμησης κάθε δείγματος στην κλάση υψηλής συμμόρφωσης από το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με XGBoost.

Πίνακας 4.8: Επεξήγηση των *swarm plots* για το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με ΧGBoost.

Χαρακτηριστικό	Χαμηλό	Μεσαίο	Υψηλό
TG	Επηρεάζει πολύ, με τις υψηλότερες τιμές να αυξάνουν τις πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει αρκετά, με τις χαμηλότερες τιμές να αυξάνουν τις πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει πολύ, με τις υψηλότερες και χαμηλότερες τιμές να αυξάνουν τις πιθανότητες για υψηλή συμμόρφωση και τις ενδιάμεσες να τις μειώνουν
Mother Authoritarian Score	Επηρεάζει πολύ, με τα παιδιά των οποίων οι μητέρες σημείωσαν τα χαμηλότερα σκορ να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει αρκετά, με τα παιδιά των οποίων οι μητέρες σημείωσαν τα υψηλότερα σκορ να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει πολύ, με τα παιδιά των οποίων οι μητέρες σημείωσαν υψηλότερα σκορ να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Ύψος	Φαίνεται να επηρεάζει αρκετά, με τις πιθανότητες για χαμηλή συμμόρφωση να αυξάνονται όσο αυξάνεται το ύψος	Επηρεάζει πολύ, με τα κοντύτερα παιδιά να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει πολύ, με τα ψηλότερα παιδιά να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Αναλογία Περιμέτρου Μέσης ανά Ύψος	Επηρεάζει αρκετά, με τα παιδιά που έχουν μικρότερη περίμετρο μέσης σε σχέση με το ύψος τους να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει πολύ, με τα παιδιά που έχουν μεγαλύτερη περίμετρο μέσης σε σχέση με το ύψος τους να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει αρκετά, με τα παιδιά που έχουν μεγαλύτερη περίμετρο μέσης σε σχέση με το ύψος τους να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο

Πίνακας 4.8 (συνέχεια): Επεξήγηση των swarm plots για το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με XGBoost.

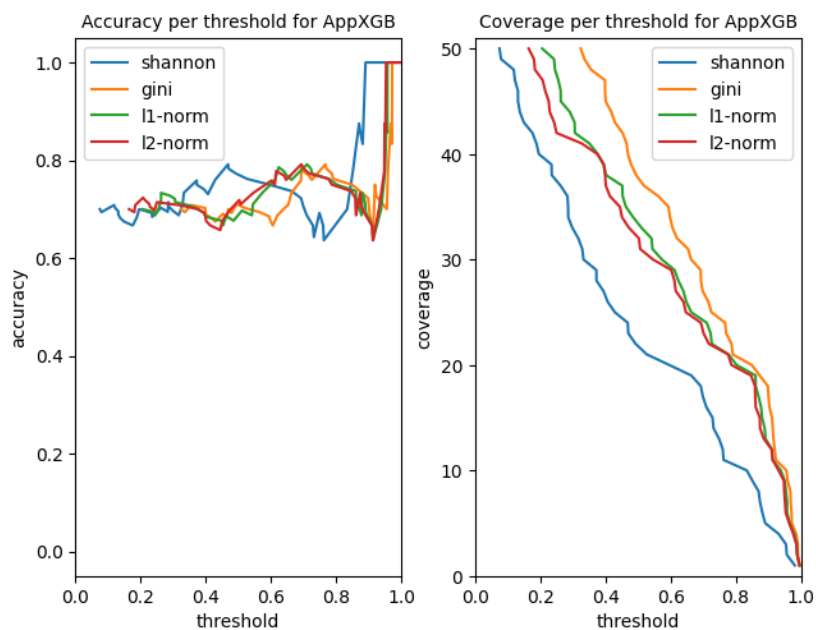
Καταναλώνετε αλκοολούχα ποτά;	Επηρεάζει πολύ, με τα παιδιά των οποίων οι μητέρες δεν καταναλώνουν αλκοολούχα ποτά να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει ελάχιστα, με τα παιδιά των οποίων οι μητέρες καταναλώνουν αλκοολούχα ποτά να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει αρκετά, με τα παιδιά των οποίων οι μητέρες καταναλώνουν αλκοολούχα ποτά συχνότερα να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Mother Permissive Score	Επηρεάζει ελάχιστα, με τα παιδιά των οποίων οι μητέρες σημείωσαν τα υψηλότερα σκορ να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει πολύ, με τα παιδιά των οποίων οι μητέρες σημείωσαν τα υψηλότερα σκορ να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει πολύ, με τα παιδιά των οποίων οι μητέρες σημείωσαν χαμηλότερα σκορ να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Νοσείτε από κάποια χρόνια ασθένεια;	Επηρεάζει πολύ, με τα παιδιά των οποίων οι μητέρες νοσούν από κάποια χρόνια ασθένεια να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει λίγο, με τα παιδιά των οποίων οι μητέρες δε νοσούν από κάποια χρόνια ασθένεια να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Δεν φαίνεται να επηρεάζει

Πίνακας 4.8 (συνέχεια): Επεξήγηση των *swarm plots* για το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με XGBoost.

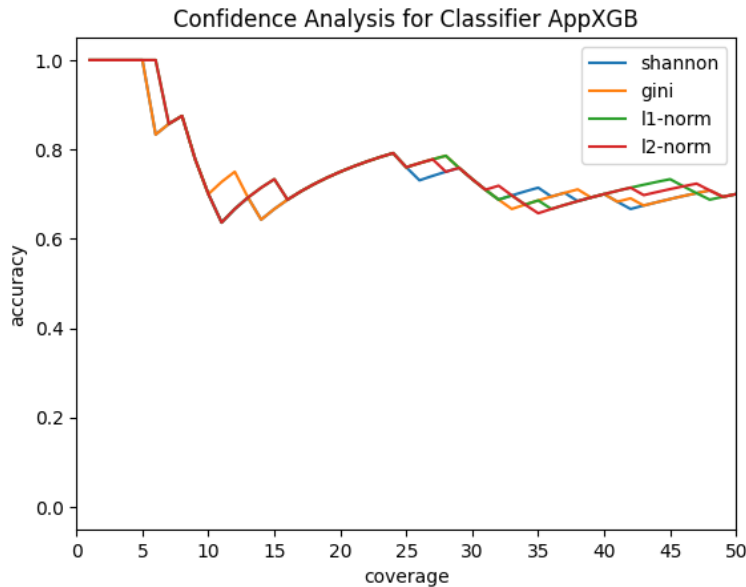
Με βάση το ύψος σας και το σωματικό σας βάρος, θεωρείτε πως είστε	Επηρεάζει πολύ, με τα παιδιά των οποίων οι μητέρες δήλωσαν ότι θεωρούν ότι είναι παχύσαρκες ή δεν απάντησαν σε αυτήν την ερώτηση να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει πολύ, με τα παιδιά των οποίων οι μητέρες δήλωσαν ότι θεωρούν ότι είναι φυσιολογικές να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει αρκετά, με τα παιδιά των οποίων οι μητέρες δήλωσαν ότι θεωρούν ότι είναι υπέρβαρες ή παχύσαρκες να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Ποιο είναι το μορφωτικό σας επίπεδο;	Επηρεάζει αρκετά, με τα παιδιά των οποίων οι μητέρες έχουν υψηλότερο μορφωτικό επίπεδο να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει ελάχιστα, με τα παιδιά των οποίων οι μητέρες έχουν το υψηλότερο μορφωτικό επίπεδο να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση	Επηρεάζει πολύ, με τα παιδιά των οποίων οι μητέρες έχουν χαμηλότερο μορφωτικό επίπεδο να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Αλλεργίες	Δεν φαίνεται να επηρεάζει ιδιαίτερα	Επηρεάζει πολύ, με τα παιδιά που έχουν αλλεργίες να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Φαίνεται να επηρεάζει ελάχιστα, με τα παιδιά που έχουν αλλεργίες να έχουν λιγότερες πιθανότητες για υψηλή συμμόρφωση
Το οικογενειακό σας εισόδημα καλύπτει τις ανάγκες σας;	Επηρεάζει ελάχιστα, με τις χαμηλότερες τιμές να αυξάνουν τις πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει λίγο, με τις υψηλότερες τιμές να αυξάνουν τις πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει λίγο, με τις υψηλότερες τιμές να αυξάνουν τις πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο

Πίνακας 4.8 (συνέχεια): Επεξήγηση των swarm plots για το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με XGBoost.

Κάπνισμα κατά την Κύηση	Δεν φαίνεται να επηρεάζει	Φαίνεται να επηρεάζει ελάχιστα, με τα παιδιά των οποίων οι μητέρες κάπνιζαν κατά την κύηση να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Φαίνεται να επηρεάζει ελάχιστα, με τα παιδιά των οποίων οι μητέρες δεν κάπνιζαν κατά την κύηση να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
-------------------------	---------------------------	--	---



Σχήμα 4.12: Διαγράμματα της ακρίβειας και της κάλυψης του συνόλου δειγμάτων συναρτήσει του κατωφλίου βεβαιότητας για το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με XGBoost. Τα επίπεδα βεβαιότητας καθορίζονται σύμφωνα με τις 4 συναρτήσεις βεβαιότητας (Shannon, Gini, νόρμα επιπέδου 1, νόρμα επιπέδου 2).



Σχήμα 4.13: Διάγραμμα της ακρίβειας του μοντέλου πρόβλεψης της συμμόρφωσης με την εφαρμογή με XGBoost για ένα υποσύνολο δειγμάτων συναρτήσει της κάλυψης ολόκληρου το συνόλου. Τα υποσύνολα υπολογίζονται για φθίνοντα κατώφλια βεβαιότητας σύμφωνα με τις 4 συναρτήσεις βεβαιότητας (Shannon, Gini, νόρμα επιπέδου 1, νόρμα επιπέδου 2).

4.2.1.3. Συμμόρφωση με το Fitbit

Τα 9 χαρακτηριστικά που προέκυψαν με sequential feature selection για την εκπαίδευση του μοντέλου XGBoost για την πρόβλεψη της συμμόρφωσης με το Fitbit παρουσιάζονται στον Πίνακα 4.9.

Ακολουθεί ο Πίνακας 4.10 στον οποίον παρουσιάζονται οι μετρικές απόδοσης του ταξινομητή. Το μοντέλο επιτυγχάνει ακρίβεια 82%. Ο πολύ μικρός αριθμός δειγμάτων από τη μεσαία κλάση ευθύνεται για το γεγονός ότι ο ταξινομητής δεν μπορεί να προβλέψει σωστά κανένα από τα τρία δείγματα της. Οι επιδόσεις του για τις άλλες δύο κλάσεις είναι υψηλές, ιδίως για την κλάση υψηλής συμμόρφωσης, με αποτέλεσμα υψηλούς σταθμισμένους μέσους όρους για όλες τις μετρικές.

Το χαρακτηριστικό με τη μεγαλύτερη επίδραση σε όλες τις κλάσεις είναι η ηλικία του πατρός με την πλειοψηφία των παιδιών νεότερων πατέρων να έχει περισσότερες πιθανότητες να ταξινομηθεί σε κλάσεις υψηλότερης συμμόρφωσης, χωρίς όμως αυτό να ισχύει για το σύνολο του δείγματος, δηλαδή σε ορισμένες περιπτώσεις η νεότερη ηλικία πατρός επηρέασε θετικά την ταξινόμηση στην κλάση χαμηλής συμμόρφωσης. Το σημαντικότερο χαρακτηριστικό για την κλάση χαμηλής συμμόρφωσης είναι η παραμονή του παιδιού σε μονάδα εντατικής νοσηλείας ή αυξημένης φροντίδας μετά τη γέννησή του, με τα παιδιά που παρέμειναν να έχουν λιγότερες πιθανότητες να ταξινομηθούν στην κλάση χαμηλής συμμόρφωσης και το αντίστροφο. Αυτό ενδέχεται να υποδεικνύει ότι τα παιδιά που παρέμειναν σε μονάδες εντατικής νοσηλείας ή αυξημένης φροντίδας είναι ίσως πιο εξοικειωμένα με την τήρηση ιατρικών συστάσεων. Από την άλλη, το σημαντικότερο χαρακτηριστικό για την κλάση υψηλής συμμόρφωσης είναι ο αριθμός των ωρών που κάθε παιδί περνάει μπροστά σε οθόνες τις καθημερινές, με τα παιδιά που περνούν περισσότερες από 4 ώρες να έχουν λιγότερες πιθανότητες να ταξινομηθούν στην κλάση υψηλής συμμόρφωσης έναντι των υπολοίπων. Αυτό μπορεί να αποτελεί ένδειξη του γεγονότος ότι τα παιδιά που περνούν πολλές ώρες μπροστά σε οθόνες να αφιερώνουν λιγότερες ώρες σε άλλα πράγματα, οπότε για αυτό να παρέλειψαν να χρησιμοποιήσουν το Fitbit, ή ότι γενικά οι

περισσότερες ώρες μπροστά σε οθόνες να σχετίζονται με τη μη τήρηση οδηγιών. Στον Πίνακα 4.11 περιγράφονται οι επιδράσεις των χαρακτηριστικών στο αποτέλεσμα, όπως προκύπτουν από τα swarm plots.

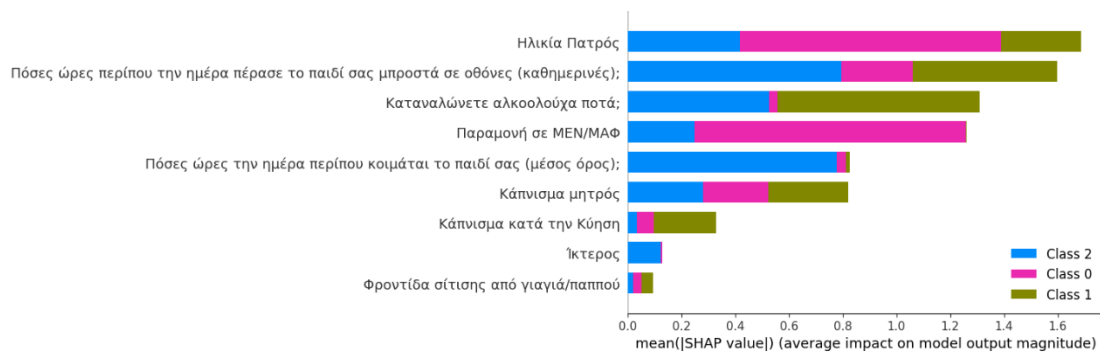
Στα σχήματα 4.18 και 4.19 όχι μόνο δεν μειώνεται μονότονα η ακρίβεια συναρτήσεως της κάλυψης δειγμάτων, αλλά και γίνεται 0 όταν η μοναδική πρόβλεψη άνω του κατωφλίου είναι αυτή με το μεγαλύτερο επίπεδο βεβαιότητας, η οποία είναι λανθασμένη. Παρατηρείται μία ελαφρώς αύξουσα τάση για όλες τις συναρτήσεις ποσοτικοποίησης της βεβαιότητας καθώς το κατώφλι αυξάνεται μέχρι το επίπεδο που αντιστοιχεί στην κάλυψη 20 δειγμάτων, οπότε και παρατηρείται έντονη μείωση της ακρίβειας.

Πίνακας 4.9: Περιγραφή των χαρακτηριστικών του συνόλου δεδομένων με τα οποία εκπαιδεύτηκε το μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με XGBoost

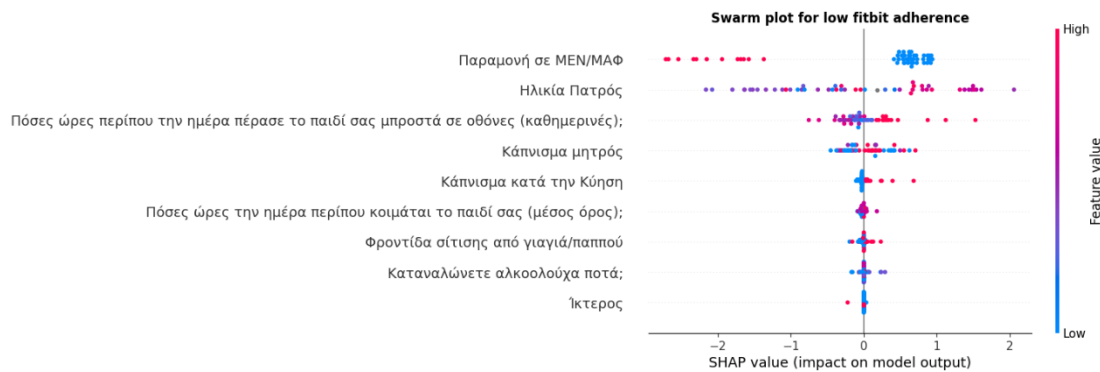
Χαρακτηριστικό	Περιγραφή	Τιμές
Ηλικία Πατρός	Ηλικία του πατρός του παιδιού κατά την έναρξη του προγράμματος παρέμβασης	Ηλικία σε έτη
Πόσες ώρες περίπου την ημέρα πέρασε το παιδί σας μπροστά σε οθόνες (καθημερινές);	Απάντηση της μητέρας στην ερώτηση «Πόσες ώρες περίπου την ημέρα (καθημερινές) πέρασε το παιδί σας μπροστά σε οθόνες; (π.χ. τηλεόραση, τάμπλετ, κινητό)»	Χρονική διάρκεια που πέρασε το παιδί μπροστά από οθόνες τις καθημερινές, κατά σειρά (< 1 ώρα / μέρα, 1-2 ώρες / μέρα, 2-3 ώρες / μέρα, 3-4 ώρες / μέρα, > 4 ώρες / μέρα), κωδικοποίηση σε 0-4 αντίστοιχα
Καταναλώνετε αλκοολούχα ποτά;	Απάντηση της μητέρας στην ερώτηση «Καταναλώνετε αλκοολούχα ποτά;»	Συχνότητα κατανάλωσης αλκοολούχων ποτών από τη μητέρα, κατά σειρά (Καθόλου, 1-3 φορές/μήνα, 1 φορά/εβδομάδα, 2-4 φορές/εβδομάδα), κωδικοποίηση σε 0-3 αντίστοιχα
Παραμονή σε ΜΕΝ/ΜΑΦ	Παραμονή του παιδιού σε Μονάδα Εντατικής Νοσηλείας ή Μονάδα Αυξημένης Φροντίδας κατά τη γέννησή του	Ναι (1) ή όχι (0)
Πόσες ώρες την ημέρα περίπου κοιμάται το παιδί σας (μέσος όρος);	Απάντηση της μητέρας στην ερώτηση «Πόσες ώρες την ημέρα περίπου κοιμάται το παιδί σας (μέσος όρος);»	Χρονική διάρκεια ύπνου του παιδιού ανά μέρα, κατά σειρά (9-11 ώρες, 7-8 ώρες, < 7 ώρες), κωδικοποίηση σε 0-2 αντίστοιχα
Κάπνισμα μητρός	Απάντηση της μητέρας στην ερώτηση «Είστε καπνιστής;»	Καπνίστρια μητέρα, κατά σειρά (Όχι, Πρώην, Ναι), κωδικοποίηση σε 0-2
Κάπνισμα κατά την Κύηση	Κάπνισμα της μητέρας κατά την κύηση του παιδιού	Ναι (1) ή όχι (0)
Ίκτερος	Εμφάνιση νεογνικού ίκτερου (neonatal jaundice) κατά τη γέννηση	Ναι (1) ή όχι (0)
Φροντίδα σίτισης από γιαγιά/παππού	Συμμετοχή της γιαγιάς ή του παππού στη φροντίδα σίτισης του παιδιού	Ναι (1) ή όχι (0)

Πίνακας 4.10: Μετρικές αξιολόγησης του μοντέλου πρόβλεψης της συμμόρφωσης με το Fitbit με XGBoost.

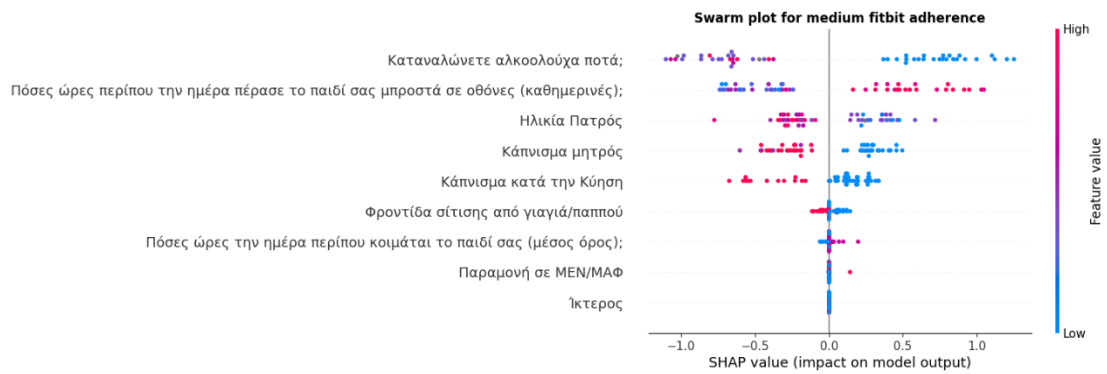
Μετρική / Επίπεδο Συμμόρφωσης	Χαμηλό	Μεσαίο	Υψηλό	Μέσος όρος	Σταθμισμένος Μέσος Όρος
Αριθμός δειγμάτων	24	3	23	-	-
ROC area under the curve	84,93%	65,96%	89,85%	80,25%	86,06%
Precision	81%	0%	87%	56%	79%
Recall	88%	0%	87%	58%	82%
f1-score	84%	0%	87%	57%	80%
Specificity	81%	98%	89%	89%	86%
Accuracy	82%			—	—



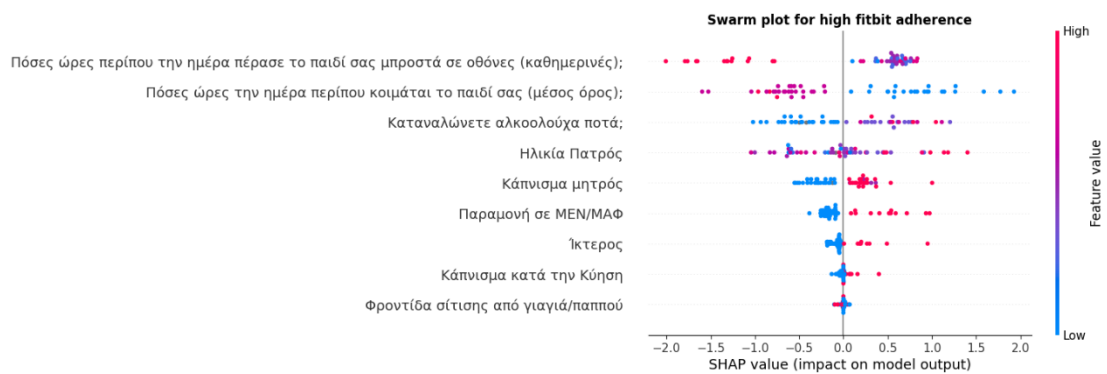
Σχήμα 4.14: Διάγραμμα ποσοτικοποίησης της μέσης επίδρασης κάθε χαρακτηριστικού στο μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με XGBoost.



Σχήμα 4.15: Swarm plot απεικόνισης της επίδρασης κάθε χαρακτηριστικού στην απόφαση ταξινόμησης κάθε δείγματος στην κλάση χαμηλής συμμόρφωσης από το μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με XGBoost.



Σχήμα 4.16: Swarm plot απεικόνισης της επίδρασης κάθε χαρακτηριστικού στην απόφαση ταξινόμησης κάθε δείγματος στην κλάση μεσαίας συμμόρφωσης από το μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με XGBoost.



Σχήμα 4.17: Swarm plot απεικόνισης της επίδρασης κάθε χαρακτηριστικού στην απόφαση ταξινόμησης κάθε δείγματος στην κλάση υψηλής συμμόρφωσης από το μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με XGBoost.

Πίνακας 4.11: Επεξήγηση των swarm plots για το μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με XGBoost.

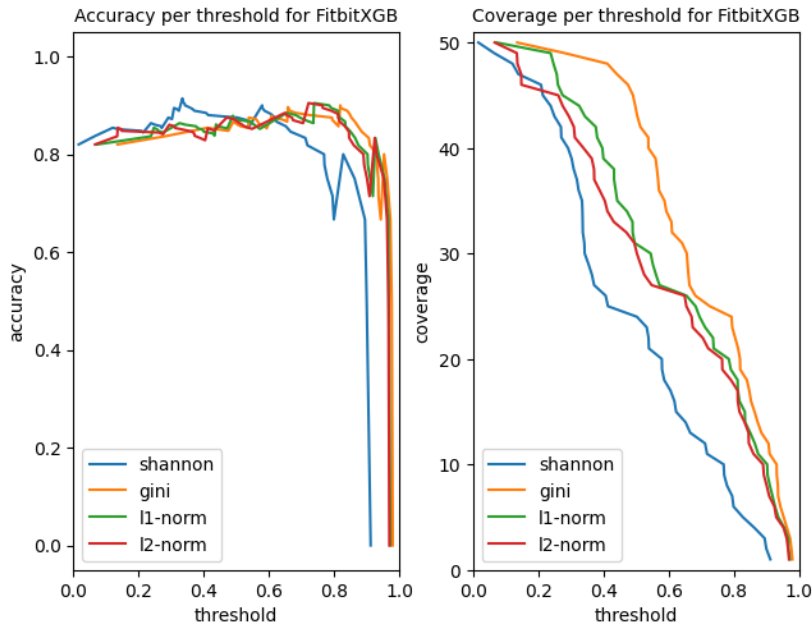
Χαρακτηριστικό	Χαμηλό	Μεσαίο	Υψηλό
Ηλικία Πατρός	Επηρεάζει πολύ, με τα παιδιά των οποίων οι πατέρες έχουν μεγαλύτερη ηλικία να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει πολύ, με τα παιδιά των οποίων οι πατέρες έχουν μικρότερη ηλικία να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει πολύ, χωρίς ξεκάθαρη μονοτονική συσχέτιση

Πίνακας 4.11 (συνέχεια): Επεξήγηση των swarm plots για το μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με XGBoost.

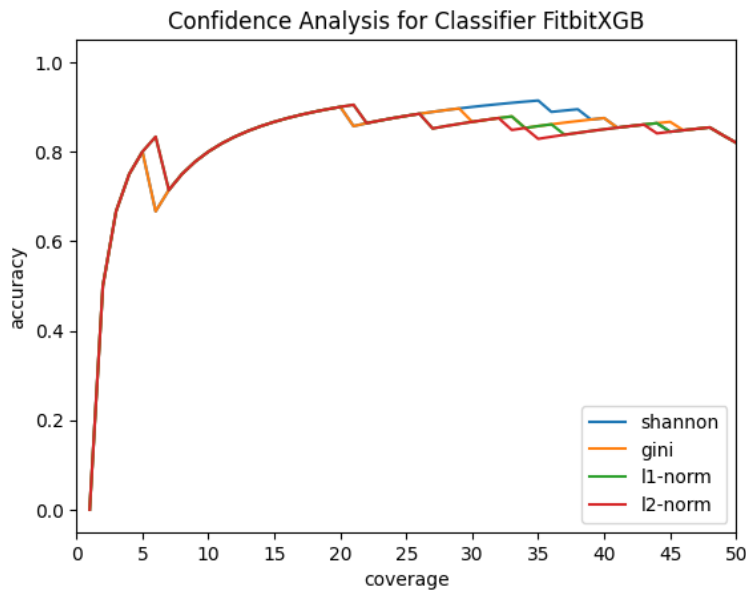
Πόσες ώρες περίπου την ημέρα πέρασε το παιδί σας μπροστά σε οθόνες (καθημερινές);	Επηρεάζει πολύ, με τα παιδιά τα οποία πέρασαν περισσότερες από 4 ώρες μπροστά σε οθόνες να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει αρκετά, με τα παιδιά τα οποία πέρασαν περισσότερες από 3 ώρες μπροστά σε οθόνες να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει πολύ, με τα παιδιά τα οποία πέρασαν περισσότερες από 4 ώρες μπροστά σε οθόνες να έχουν λιγότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Καταναλώνετε αλκοολούχα ποτά;	Φαίνεται να επηρεάζει λίγο, χωρίς ξεκάθαρη μονοτονική συσχέτιση	Επηρεάζει πολύ, με τα παιδιά των οποίων οι μητέρες δεν καταναλώνουν αλκοολούχα ποτά να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει πολύ, με τα παιδιά των οποίων οι μητέρες καταναλώνουν αλκοολούχα ποτά να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Παραμονή σε ΜΕΝ/ΜΑΦ	Επηρεάζει πολύ, με τα παιδιά τα οποία δεν παρέμειναν σε ΜΕΝ ή ΜΑΦ κατά τη γέννησή τους να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Δεν φαίνεται να επηρεάζει	Επηρεάζει αρκετά, με τα παιδιά τα οποία παρέμειναν σε ΜΕΝ ή ΜΑΦ κατά τη γέννησή τους να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Πόσες ώρες την ημέρα περίπου κοιμάται το παιδί σας (μέσος όρος);	Φαίνεται να επηρεάζει λίγο, χωρίς ξεκάθαρη μονοτονική συσχέτιση	Φαίνεται να επηρεάζει λίγο, με τα παιδιά που κοιμούνται 7-8 ώρες να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και τα παιδιά που κοιμούνται λιγότερες από 7 ώρες λιγότερες πιθανότητες για μεσαία συμμόρφωση	Επηρεάζει πολύ, με τα παιδιά που κοιμούνται 9-11 ώρες να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο

Πίνακας 4.11 (συνέχεια): Επεξήγηση των swarm plots για το μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με ΧGBoost.

Κάπνισμα μητρός	Φαίνεται να επηρεάζει αρκετά, χωρίς ξεκάθαρη μονοτονική συσχέτιση	Επηρεάζει αρκετά, με τα παιδιά των οποίων οι μητέρες δεν κάπνιζαν ποτέ να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει αρκετά, με τα παιδιά των οποίων οι μητέρες κάπνιζαν ή έχουν καπνίσει να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Κάπνισμα κατά την Κύηση	Επηρεάζει αρκετά, με τα παιδιά των οποίων οι μητέρες κάπνιζαν κατά την κύηση να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει αρκετά, με τα παιδιά των οποίων οι μητέρες δεν κάπνιζαν κατά την κύηση να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Φαίνεται να επηρεάζει λίγο, με τα παιδιά των οποίων οι μητέρες κάπνιζαν κατά την κύηση να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Ίκτερος	Δεν φαίνεται να επηρεάζει	Δεν φαίνεται να επηρεάζει	Επηρεάζει αρκετά, με τα παιδιά τα οποία παρουσίασαν νεογνικό ίκτερο να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Φροντίδα σίτισης από γιαγιά/παππού	Φαίνεται να επηρεάζει λίγο, με τα παιδιά των οποίων τη φροντίδα σίτισης συμμετέχουν η γιαγιά ή ο παππούς να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Φαίνεται να επηρεάζει λίγο, με τα παιδιά των οποίων τη φροντίδα σίτισης συμμετέχουν η γιαγιά ή ο παππούς να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Φαίνεται να επηρεάζει ελάχιστα, με τα παιδιά των οποίων τη φροντίδα σίτισης δεν συμμετέχουν η γιαγιά ή ο παππούς να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο



Σχήμα 4.18: Διαγράμματα της ακρίβειας και της κάλυψης του συνόλου δειγμάτων συναρτήσεως του κατωφλίου βεβαιότητας για το μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με XGBoost. Τα επίπεδα βεβαιότητας καθορίζονται σύμφωνα με τις 4 συναρτήσεις βεβαιότητας (Shannon, Gini, νόρμα επιπέδου 1, νόρμα επιπέδου 2).



Σχήμα 4.19: Διάγραμμα της ακρίβειας του μοντέλου πρόβλεψης της συμμόρφωσης με το Fitbit με XGBoost για ένα υποσύνολο δειγμάτων συναρτήσεως της κάλυψης ολόκληρου το συνόλου. Τα υποσύνολα υπολογίζονται για φθίνοντα κατώφλια βεβαιότητας σύμφωνα με τις 4 συναρτήσεις βεβαιότητας (Shannon, Gini, νόρμα επιπέδου 1, νόρμα επιπέδου 2).

4.2.2. Πρόβλεψη Συμμόρφωσης με Optimal Decision Trees

Σε αυτήν την ενότητα, παρουσιάζονται τα μοντέλα που προέκυψαν από την εκπαίδευση του ταξινομητή Optimal Tree Classifier του πακέτου interpretableai. Αναπτύχθηκαν σε γλώσσα R μοντέλα για τα τρία ζητούμενα.

4.2.2.1. Συνολική Συμμόρφωση

Σε αυτήν την περίπτωση, η διαδικασία sequential feature selection έδειξε ότι οι καλύτερες επιδόσεις επιτυγχάνονται με 16 χαρακτηριστικά, τα οποία παρουσιάζονται στον Πίνακα 4.12.

Το μοντέλο επιτυγχάνει ακρίβεια 80%. Οι μετρικές είναι υψηλές για όλες τις κλάσεις και ειδικά για αυτή της χαμηλής συμμόρφωσης, για την οποία το f1-score ανέρχεται στο 85%.

Στο Σχήμα 4.20 φαίνεται ότι τρία χαρακτηριστικά κυριάρχησαν έναντι των υπολοίπων 13 στις προβλέψεις του ταξινομητή: η ηλικία του πατρός, το μορφωτικό επίπεδο της μητέρας και το επίπεδο χοληστερόλης του παιδιού. Η σχέση καθενός από τα τρία χαρακτηριστικά και των πιθανοτήτων ταξινόμησης σε κάθε κλάση χαρακτηρίζεται από έλλειψη μονοτονίας. Αυτό σημαίνει ότι η ίδια τιμή ενός χαρακτηριστικού μπορεί να επηρεάσει θετικά την ταξινόμηση ενός δείγματος σε μία κλάση και αρνητικά την ταξινόμηση ενός άλλου δείγματος στην ίδια κλάση. Τα αίτια για αυτή τη συμπεριφορά είναι δύο: το πρώτο έχει να κάνει με τη δομή των δέντρων αποφάσεων και το γεγονός ότι μπορεί δύο κόμβοι να βασίζονται στην απόφασή τους στο ίδιο χαρακτηριστικό. Μπορεί η ίδια τιμή του χαρακτηριστικού στον ένα κόμβο να οδηγεί σε μία απόφαση υπέρ της ταξινόμησης σε μία κλάση και στον άλλον να οδηγεί σε μία απόφαση ενάντια στην ταξινόμηση στην ίδια κλάση. Ανάλογα με τις τιμές και των υπόλοιπων χαρακτηριστικών, δύο δείγματα με την ίδια τιμή σε ένα χαρακτηριστικό μπορεί να οδηγηθούν το καθένα σε διαφορετικό από τους προαναφερόμενους κόμβους. Το δεύτερο αίτιο έχει να κάνει με το γεγονός ότι εκτελείται leave-one-out διασταυρωμένη επικύρωση και άρα εκπαιδεύονται 50 διαφορετικά δέντρα αποφάσεων, τα οποία μπορεί να έχουν διαφορετική συμπεριφορά μεταξύ τους. Στον Πίνακα 4.14 περιγράφονται οι επιδράσεις των χαρακτηριστικών στη συνολική συμμόρφωση, όπως προκύπτουν από τα swarm plots.

Τα σχήματα 4.24 και 4.25 διαφέρουν αρκετά από τα αντίστοιχα για τα μοντέλα που προέκυψαν από το XGBoost. Αυτό οφείλεται στο γεγονός ότι ο Optimal Tree Classifier είναι πιο απλό μοντέλο σε σύγκριση με το XGBoost και άρα υπάρχει μικρότερη ποικιλία στα διανύσματα πιθανοτήτων που είναι δυνατό να εξαχθούν. Συγκεκριμένα, τα διανύσματα πιθανοτήτων προκύπτουν από το φύλλο του δέντρου στο οποίο κατέληξε το δείγμα προς πρόβλεψη. Οι πιθανότητες να ταξινομηθεί το δείγμα σε μία κλάση αντιστοιχούν στο ποσοστό των δειγμάτων του συνόλου δεδομένων εκπαίδευσης που ανήκουν στην κλάση αυτή και θα κατέληγαν στο συγκεκριμένο φύλλο (οι πιθανότητες αναφέρονται στο στοιχείο του διανύσματος πιθανοτήτων που χρησιμοποιείται για τη διαδικασία ποσοτικοποίησης της βεβαιότητας και όχι στην απόφαση πρόβλεψης κατά την οποία το στοιχείο ταξινομείται ντετερμινιστικά στην κλάση που αντιστοιχεί στην πλειοψηφία των δειγμάτων του φύλλου). Επειδή υπάρχουν πολλά φύλλα από τα οποία προκύπτει μοναδιαίο διάνυσμα πιθανοτήτων για πολλά δείγματα, συγκεκριμένα 36, το επίπεδο βεβαιότητας για τα αυτά ήταν 100%.

Πίνακας 4.12: Περιγραφή των χαρακτηριστικών του συνόλου δεδομένων με τα οποία εκπαιδεύτηκε το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με *Optimal Decision Trees*.

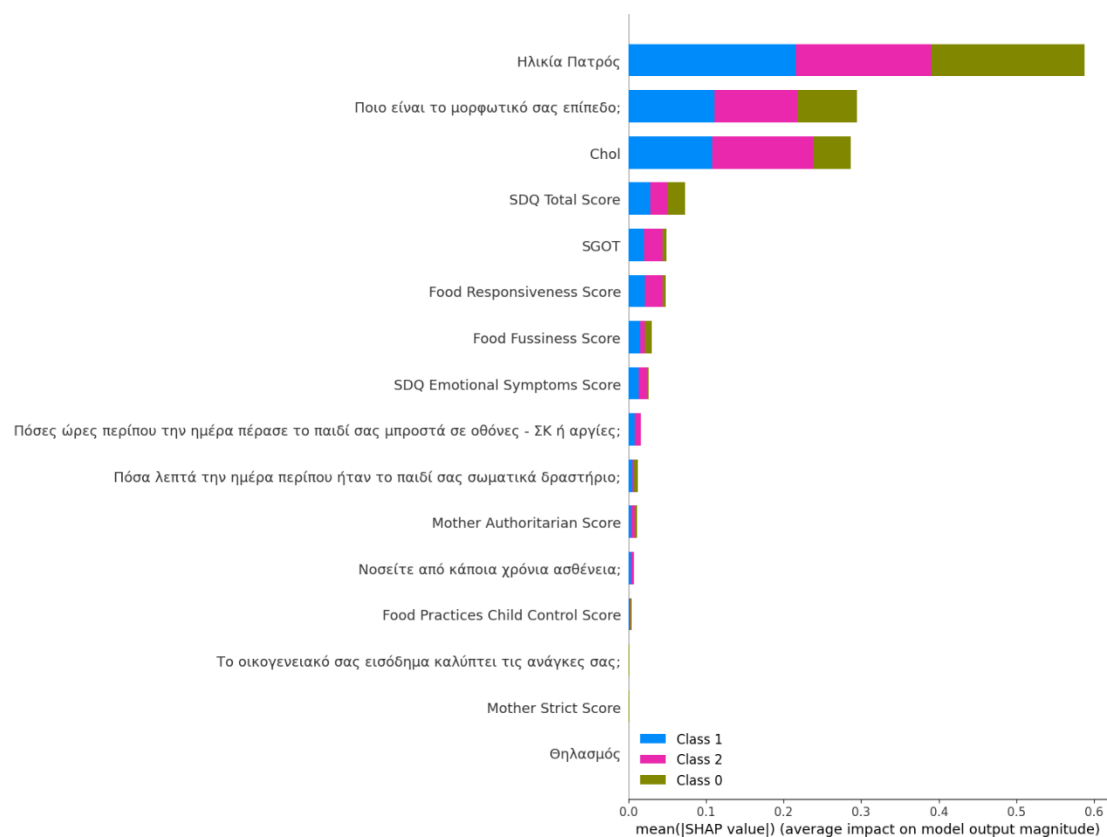
Χαρακτηριστικό	Περιγραφή	Τιμές
Ηλικία Πατρός	Ηλικία του πατρός του παιδιού κατά την έναρξη του προγράμματος παρέμβασης	Ηλικία σε έτη
Ποιο είναι το μορφωτικό σας επίπεδο;	Απάντηση της μητέρας στην ερώτηση «Ποιο είναι το μορφωτικό σας επίπεδο;»	Υψηλότερο μορφωτικό επίπεδο που έχει ολοκληρωθεί από τη μητέρα, κατά σειρά (Δημοτικό, Γυμνάσιο, Γενικό Λύκειο, ΙΕΚ, ΤΕΙ, ΑΕΙ, Κάτοχος Μεταπτυχιακού Διπλώματος), κωδικοποίηση σε αριθμούς 0-6 αντίστοιχα
Chol	Επίπεδο χοληστερόλης στο αίμα. Υψηλή συγκέντρωση χοληστερόλης μπορεί να προκαλέσει αρτηριοσκλήρυνση.	Επίπεδο χοληστερόλης σε mg/dL
SDQ Total Score	Συνολική βαθμολογία ερωτηματολογίου δυνατοτήτων και δυσκολιών (strengths and difficulties questionnaire). Μεγαλύτερες τιμές υποδεικνύουν περισσότερες δυσκολίες.	Ακέραιες τιμές στο διάστημα 0-40
SGOT	Επίπεδο οξαλικής τρανσαμινάσης (SGOT) στο αίμα. Υψηλότερα επίπεδα σχετίζονται με ηπατικά νοσήματα	Επίπεδο SGOT σε U/L
Food Responsiveness Score	Βαθμολογία απόκρισης σε ερεθίσματα τροφής, όπως προκύπτει από το Ερωτηματολόγιο Διατροφικών Συνηθειών και άλλων Συμπεριφορών Υγείας	Συνεχείς τιμές στο διάστημα 1-5
Food Fussiness Score	Βαθμολογία ιδιότροπης συμπεριφοράς με το φαγητό, όπως προκύπτει από το Ερωτηματολόγιο Διατροφικών Συνηθειών και άλλων Συμπεριφορών Υγείας	Συνεχείς τιμές στο διάστημα 1-5
SDQ Emotional Symptoms Score	Βαθμολογία συναισθηματικών συμπτωμάτων, όπως προκύπτει από το ερωτηματολόγιο δυνατοτήτων και δυσκολιών.	Ακέραιες τιμές στο διάστημα 0-10
Πόσες ώρες περίπου την ημέρα πέρασε το παιδί σας μπροστά σε οθόνες - ΣΚ ή αργίες;	Απάντηση της μητέρας στην ερώτηση «Πόσες ώρες περίπου την ημέρα (ΣΚ ή αργίες) πέρασε το παιδί σας μπροστά σε οθόνες; (π.χ. τηλεόραση, τάμπλετ, κινητό)»	Χρονική διάρκεια που πέρασε το παιδί μπροστά από οθόνες το Σαββατοκύριακο ή τις αργίες, κατά σειρά (< 1 ώρα / μέρα, 1-2 ώρες / μέρα, 2-3 ώρες / μέρα, 3-4 ώρες / μέρα, > 4 ώρες / μέρα) κωδικοποίηση σε 0-4 αντίστοιχα

Πίνακας 4.12 (συνέχεια): Περιγραφή των χαρακτηριστικών του συνόλου δεδομένων με τα οποία εκπαιδεύτηκε το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με *Optimal Decision Trees*.

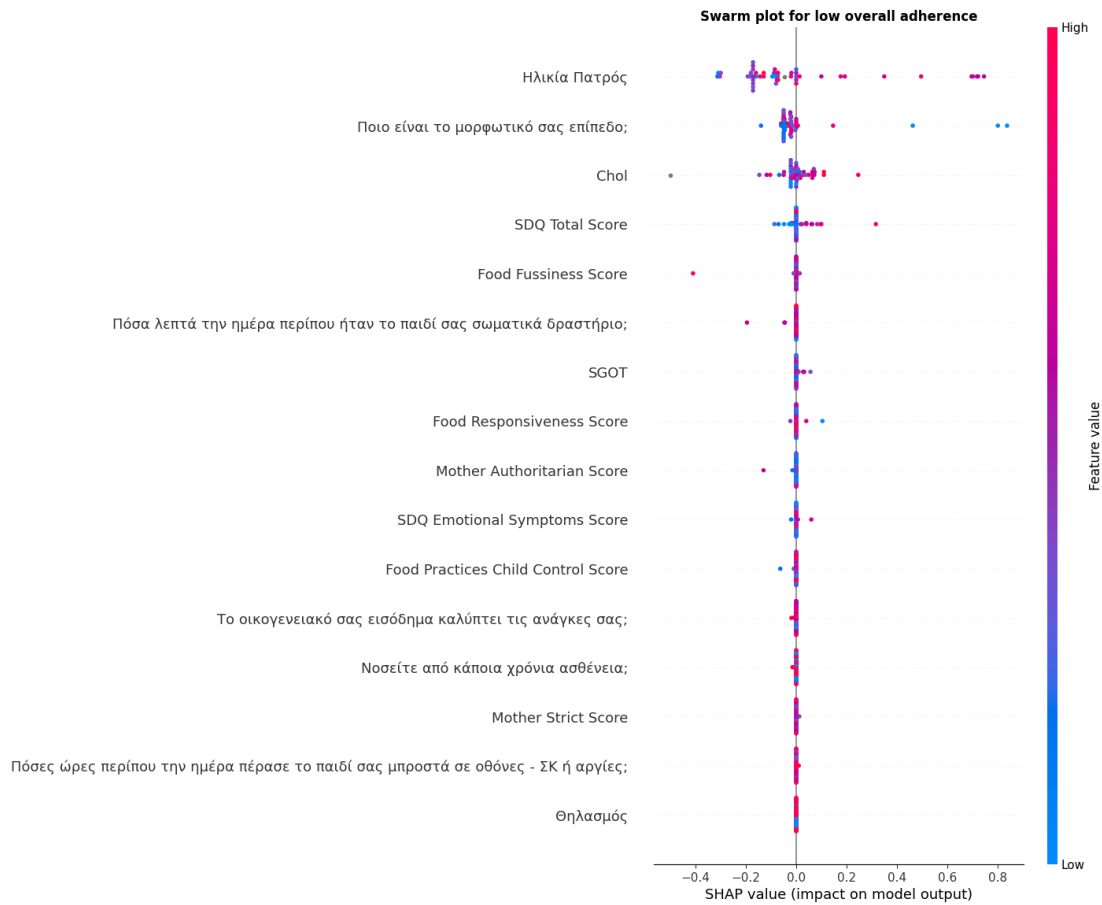
Πόσα λεπτά την ημέρα ήταν το παιδί σας σωματικά δραστήριο;	Απάντηση της μητέρας στην ερώτηση «Πόσα λεπτά την ημέρα περίπου ήταν το παιδί σας σωματικά δραστήριο (περπάτημα, ποδήλατο, πατίνι, παιχνίδια με φίλους που απαιτούν σωματική δραστηριότητα, κηπουρική, μαγειρική, δουλειές σπιτιού);»	Συνολική χρονική διάρκεια σωματικής δραστηριότητας του παιδιού, κατά σειρά (Καθόλου, < 30 λεπτά, 30 - 60 λεπτά, 60-120 λεπτά, > 120 λεπτά) κωδικοποίηση σε 0-4 αντίστοιχα.
Mother Authoritarian Score	Βαθμολογία αυταρχικού τύπου μητέρας, όπως προκύπτει από το ερωτηματολόγιο γονεϊκού τύπου	Συνεχείς τιμές στο διάστημα 1-5
Νοσείτε από κάποια χρόνια ασθένεια;	Απάντηση της μητέρας στην ερώτηση «Νοσείτε από κάποια χρόνια ασθένεια (πχ. Υπέρταση, Δυσλιπιδαιμία, Διαβήτη);»	Απάντηση ναι (0) ή όχι (1)
Food Practices Child Control Score	Βαθμολογία ελέγχου του παιδιού, όπως προκύπτει από το Ερωτηματολόγιο Γονεϊκών Πρακτικών Σίτισης	Συνεχείς τιμές στο διάστημα 1-5
Το οικογενειακό σας εισόδημα καλύπτει τις ανάγκες σας;	Απάντηση της μητέρας στην ερώτηση «Το οικογενειακό σας εισόδημα καλύπτει τις ανάγκες σας;»	Αντίληψη της μητέρας σχετικά με την ικανοποίηση των αναγκών από το οικογενειακό εισόδημα, κατά σειρά (Καθόλου, Λίγο, Μέτρια, Πολύ, Πάρα πολύ), κωδικοποίηση σε 0-4
Mother Strict Score	Βαθμολογία αυστηρού τύπου μητέρας, όπως προκύπτει από το ερωτηματολόγιο γονεϊκού τύπου	Συνεχείς τιμές στο διάστημα 1-5
Θηλασμός	Θηλασμός του παιδιού μετά τη γέννηση	Απάντηση ναι (1) ή όχι (0)

Πίνακας 4.13: Μετρικές αξιολόγησης του μοντέλου πρόβλεψης της συνολικής συμμόρφωσης με *Optimal Decision Trees*.

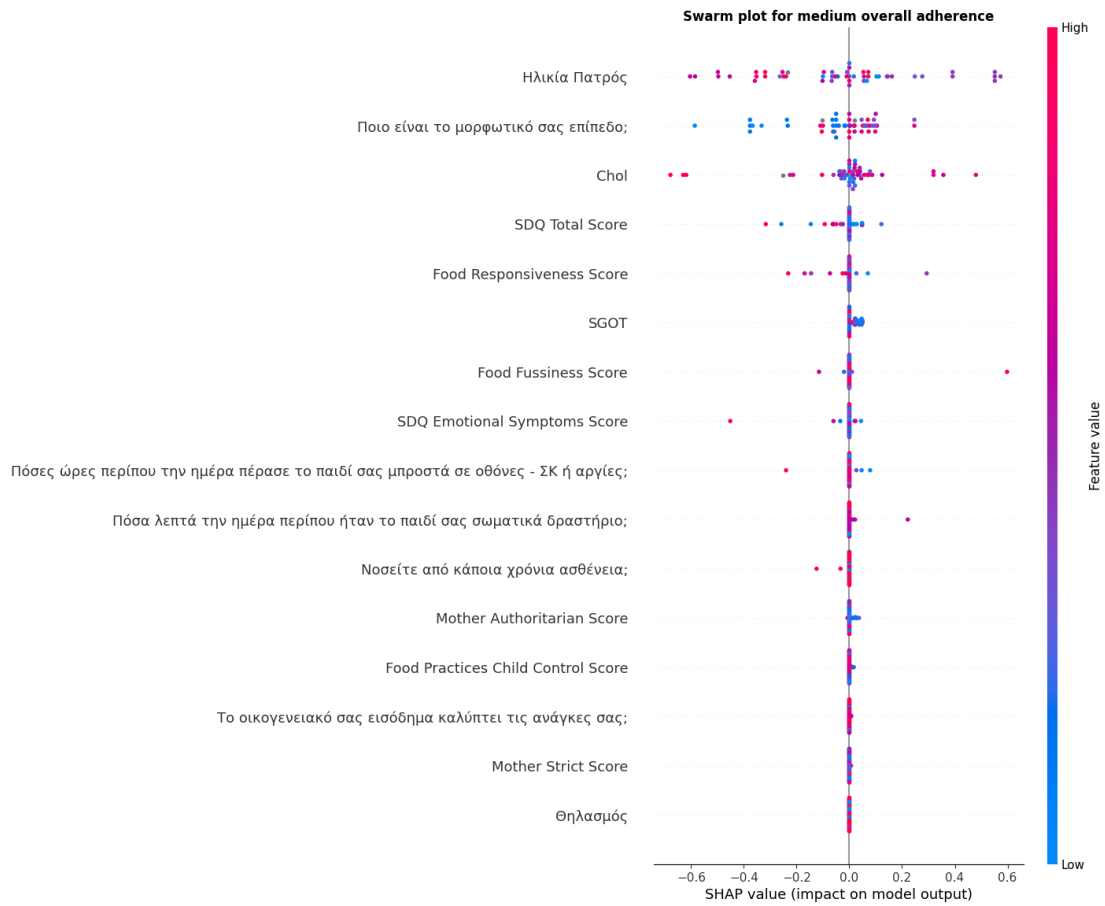
Μετρική / Επίπεδο Συμμόρφωσης	Χαμηλό	Μεσαίο	Υψηλό	Μέσος όρος	Σταθμισμένος Μέσος Όρος
Αριθμός δειγμάτων	14	20	16	—	—
ROC area under the curve	83%	82%	82%	83%	82%
Precision	92%	80%	72%	81%	81%
Recall	79%	80%	81%	80%	80%
f1-score	85%	80%	76%	80%	80%
Specificity	97%	87%	85%	90%	89%
Accuracy	80%			—	—



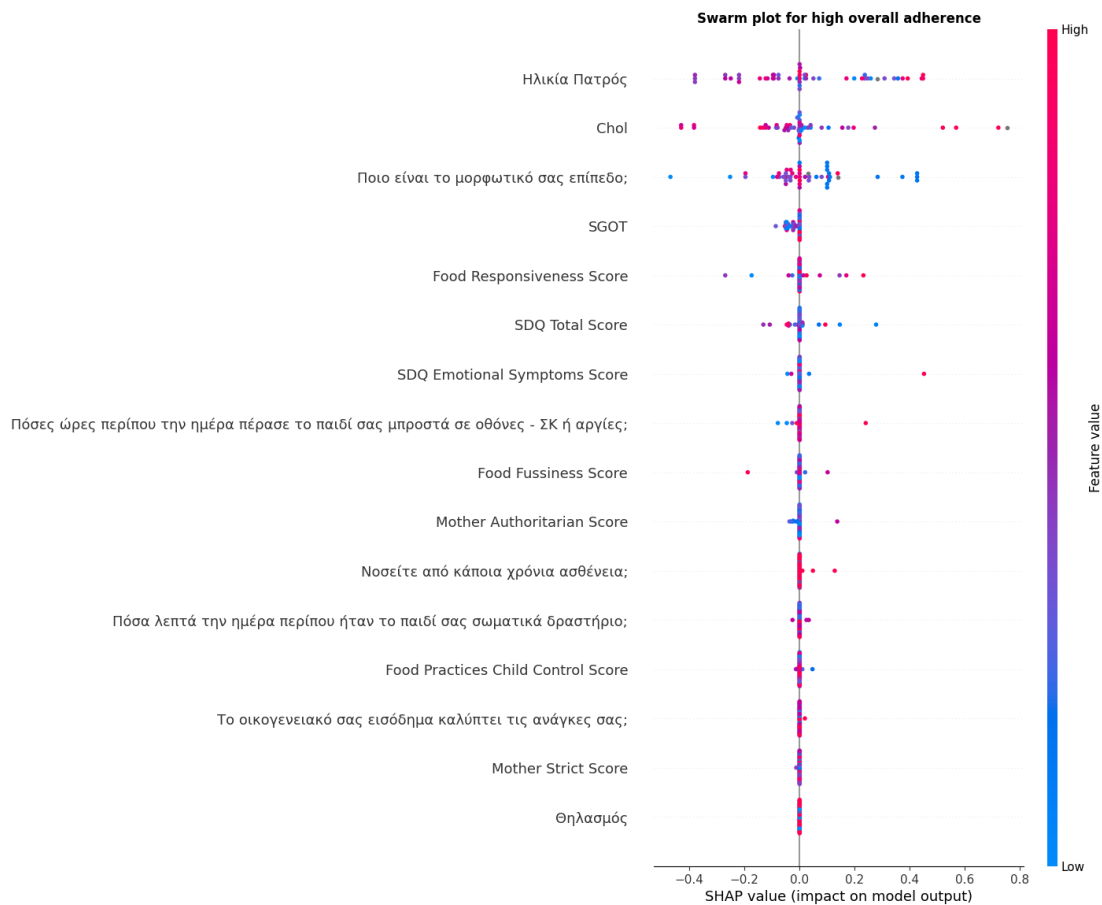
Σχήμα 4.20: Διάγραμμα ποσοτικοποίησης της μέσης επίδρασης κάθε χαρακτηριστικού στο μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με *Optimal Decision Trees*.



Σχήμα 4.21: *Swarm plot* απεικόνισης της επίδρασης κάθε χαρακτηριστικού στην απόφαση ταξινόμησης κάθε δείγματος στην κλάση χαμηλής συμμόρφωσης από το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με *Optimal Decision Trees*.



Σχήμα 4.22: Swarm plot απεικόνισης της επίδρασης κάθε χαρακτηριστικού στην απόφαση ταξινόμησης κάθε δείγματος στην κλάση μεσαίας συμμόρφωσης από το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με *Optimal Decision Trees*.



Σχήμα 4.23: Swarm plot απεικόνισης της επίδρασης κάθε χαρακτηριστικού στην απόφαση ταξινόμησης κάθε δείγματος στην κλάση υψηλής συμμόρφωσης από το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με *Optimal Decision Trees*.

Πίνακας 4.14: Επεξήγηση των swarm plots για το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με *Optimal Decision Trees*.

Χαρακτηριστικό	Χαμηλή Συμμόρφωση	Μεσαία Συμμόρφωση	Υψηλή Συμμόρφωση
Ηλικία Πατρός	Επηρεάζει αρκετά, χωρίς ξεκάθαρη μονότονη σχέση	Επηρεάζει αρκετά, χωρίς ξεκάθαρη μονοτονική σχέση	Επηρεάζει αρκετά, με τα παιδιά των οποίων οι πατέρες έχουν μεγαλύτερες ή μικρότερες ηλικίες να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και τα παιδιά των οποίων οι πατέρες έχουν ενδιάμεσες ηλικίες να έχουν λιγότερες

Πίνακας 4.14 (συνέχεια): Επεξήγηση των *swarm plots* για το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με *Optimal Decision Trees*.

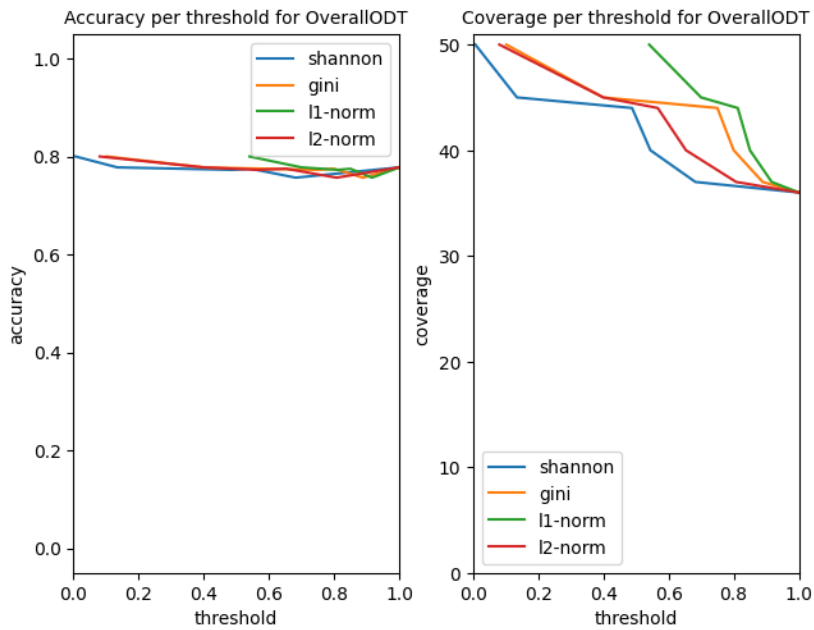
<p>Ποιο είναι το μορφωτικό σας επίπεδο;</p>	<p>Επηρεάζει αρκετά, με τα παιδιά των οποίων οι μητέρες έχουν το χαμηλότερο μορφωτικό επίπεδο να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο</p>	<p>Επηρεάζει αρκετά, με τα παιδιά των οποίων οι μητέρες έχουν υψηλότερο μορφωτικό επίπεδο να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο</p>	<p>Επηρεάζει αρκετά, με τα παιδιά των μητέρων των οποίων το υψηλότερο μορφωτικό επίπεδο είναι το γενικό λύκειο να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και τα παιδιά των μητέρων των οποίων το υψηλότερο μορφωτικό επίπεδο είναι το γυμνάσιο να έχουν λιγότερες πιθανότητες για υψηλή συμμόρφωση</p>
<p>Chol</p>	<p>Επηρεάζει αρκετά, με τα παιδιά που έχουν τις υψηλότερες τιμές να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση</p>	<p>Επηρεάζει αρκετά, χωρίς ξεκάθαρη μονοτονική σχέση</p>	<p>Επηρεάζει αρκετά, χωρίς ξεκάθαρη μονοτονική σχέση</p>
<p>SDQ Total Score</p>	<p>Επηρεάζει λίγο, με τα παιδιά τα οποία έχουν μεγαλύτερο σκορ να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο</p>	<p>Επηρεάζει λίγο, με τα παιδιά τα οποία έχουν μικρότερο σκορ να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο</p>	<p>Επηρεάζει λίγο, με τα παιδιά τα οποία έχουν τα χαμηλότερα σκορ να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο</p>
<p>SGOT</p>	<p>Φαίνεται να επηρεάζει ελάχιστα</p>	<p>Φαίνεται να επηρεάζει ελάχιστα, με τα παιδιά τα οποία έχουν χαμηλότερες τιμές να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση</p>	<p>Επηρεάζει λίγο, με τα παιδιά τα οποία έχουν χαμηλότερες τιμές να έχουν λιγότερες πιθανότητες για υψηλή συμμόρφωση</p>

Πίνακας 4.14 (συνέχεια): Επεξήγηση των swarm plots για το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με Optimal Decision Trees.

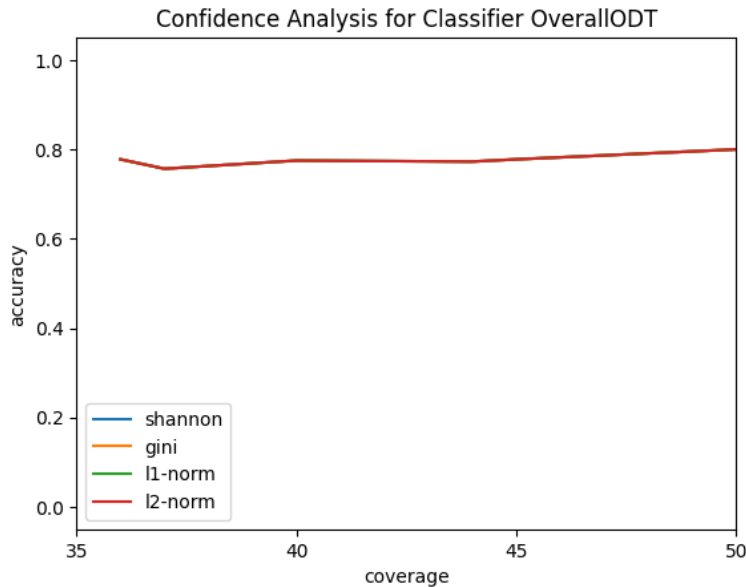
Food Responsiveness Score	Φαίνεται να επηρεάζει ελάχιστα	Επηρεάζει λίγο, με τα παιδιά τα οποία έχουν χαμηλότερο σκορ να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Φαίνεται να επηρεάζει λίγο, με τα παιδιά τα οποία έχουν υψηλότερα σκορ να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Food Fussiness Score	Φαίνεται να επηρεάζει αρνητικά μόνο για το παιδί με την υψηλότερη τιμή	Φαίνεται να επηρεάζει θετικά μόνο για το παιδί με την υψηλότερη τιμή	Φαίνεται να επηρεάζει αρνητικά μόνο για το παιδί με την υψηλότερη τιμή
SDQ Emotional Symptoms Score	Φαίνεται να επηρεάζει ελάχιστα	Φαίνεται να επηρεάζει αρνητικά μόνο για το παιδί με την υψηλότερη τιμή	Φαίνεται να επηρεάζει θετικά μόνο για το παιδί με την υψηλότερη τιμή
Πόσες ώρες περίπου την ημέρα πέρασε το παιδί σας μπροστά σε οθόνες - ΣΚ ή αργίες;	Δεν φαίνεται να επηρεάζει	Φαίνεται να επηρεάζει ελάχιστα, με τα παιδιά τα οποία πέρασαν λιγότερες ώρες να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Φαίνεται να επηρεάζει ελάχιστα, με τα παιδιά τα οποία πέρασαν περισσότερες ώρες να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Πόσα λεπτά την ημέρα ήταν το παιδί σας σωματικά δραστήριο;	Φαίνεται να επηρεάζει αρνητικά μόνο για δύο παιδιά με μεσαίες τιμές	Φαίνεται να επηρεάζει ελάχιστα	Φαίνεται να επηρεάζει ελάχιστα
Mother Authoritarian Score	Φαίνεται να επηρεάζει ελάχιστα	Φαίνεται να επηρεάζει ελάχιστα	Φαίνεται να επηρεάζει ελάχιστα
Νοσεείτε από κάποια χρόνια ασθένεια;	Δεν φαίνεται να επηρεάζει	Φαίνεται να επηρεάζει ελάχιστα	Φαίνεται να επηρεάζει θετικά μόνο για παιδιά των οποίων οι μητέρες νοσοούν
Food Practices Child Control Score	Φαίνεται να επηρεάζει ελάχιστα	Δεν φαίνεται να επηρεάζει	Φαίνεται να επηρεάζει ελάχιστα

Πίνακας 4.14 (συνέχεια): Επεξήγηση των swarm plots για το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με Optimal Decision Trees.

Το οικογενειακό σας εισόδημα καλύπτει τις ανάγκες σας;	Δεν φαίνεται να επηρεάζει	Δεν φαίνεται να επηρεάζει	Δεν φαίνεται να επηρεάζει
Mother Strict Score	Δεν φαίνεται να επηρεάζει	Δεν φαίνεται να επηρεάζει	Δεν φαίνεται να επηρεάζει
Θηλασμός	Δεν φαίνεται να επηρεάζει	Δεν φαίνεται να επηρεάζει	Δεν φαίνεται να επηρεάζει



Σχήμα 4.24: Διαγράμματα της ακρίβειας και της κάλυψης του συνόλου δειγμάτων συναρτήσει του κατωφλίου βεβαιότητας για το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με Optimal Decision Trees. Τα επίπεδα βεβαιότητας καθορίζονται σύμφωνα με τις 4 συναρτήσεις βεβαιότητας (Shannon, Gini, νόρμα επιπέδου 1, νόρμα επιπέδου 2).



Σχήμα 4.25: Διάγραμμα της ακρίβειας του μοντέλου πρόβλεψης της συνολικής συμμόρφωσης με *Optimal Decision Trees* για ένα υποσύνολο δειγμάτων συναρτήσει της κάλυψης ολόκληρου το συνόλου. Τα υποσύνολα υπολογίζονται για φθίνοντα κατώφλια βεβαιότητας σύμφωνα με τις 4 συναρτήσεις βεβαιότητας (Shannon, Gini, νόρμα επιπέδου 1, νόρμα επιπέδου 2).

4.2.2.2. Συμμόρφωση με την Εφαρμογή

Ακολουθούν η περιγραφή των χαρακτηριστικών, το διάγραμμα συνολικής επίδρασης των χαρακτηριστικών, τα swarm plots και η επεξήγησή τους για τον ταξινομητή πρόβλεψης της συμμόρφωσης με την εφαρμογή, ο οποίος εκπαιδεύτηκε με 6 χαρακτηριστικά τα οποία προέκυψαν μέσω sequential feature selection.

Το μοντέλο επιτυγχάνει ακρίβεια 80% και, όπως και στο μοντέλο XGBoost για την πρόβλεψη της συμμόρφωσης με την εφαρμογή, οι μετρικές για το μεσαίο επίπεδο συμμόρφωσης είναι χαμηλότερες εξαιτίας του μικρού αριθμού δειγμάτων της αντίστοιχης κλάσης.

Όπως φαίνεται στα σχήματα 4.26, 4.27, 4.28 και 4.29, τα τρία σημαντικότερα χαρακτηριστικά για τις αποφάσεις του μοντέλου ήταν το μορφωτικό επίπεδο της μητέρας και το αν καπνίζει, καθώς και η συστολική αρτηριακή πίεση που μετρήθηκε στα παιδιά. Τα παιδιά των οποίων οι μητέρες είναι νυν καπνίστριες ταξινομούνται από το μοντέλο με μεγαλύτερη πιθανότητα στην κλάση χαμηλής συμμόρφωσης και το αντίστροφο. Αυτό ενδεχομένως να δείχνει ότι οι μητέρες που καπνίζουν σημειώνουν μικρότερη συμμόρφωση στις ιατρικές οδηγίες.

Ομοίως με την περίπτωση του αντίστοιχου μοντέλου για την πρόβλεψη της συνολικής συμμόρφωσης με τον *Optimal Tree Classifier*, τα δείγματα με επίπεδο βεβαιότητας μικρότερο του 100% είναι μόλις 19, ένας αριθμός πολύ μικρός για να έχουν οι γραφικές παραστάσεις των σχημάτων 4.30 και 4.31 την αναμενόμενη συμπεριφορά.

Πίνακας 4.15: Περιγραφή των χαρακτηριστικών του συνόλου δεδομένων με τα οποία εκπαιδεύτηκε το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με *Optimal Decision Trees*.

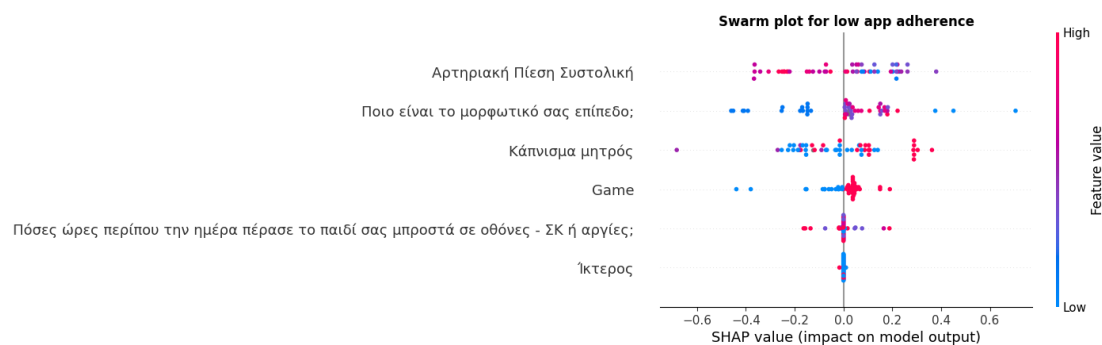
Χαρακτηριστικό	Περιγραφή	Τιμές
Ποιο είναι το μορφωτικό σας επίπεδο;	Απάντηση της μητέρας στην ερώτηση «Ποιο είναι το μορφωτικό σας επίπεδο;»	Υψηλότερο μορφωτικό επίπεδο που έχει ολοκληρωθεί από τη μητέρα, κατά σειρά (Δημοτικό, Γυμνάσιο, Γενικό Λύκειο, ΙΕΚ, ΤΕΙ, ΑΕΙ, Κάτοχος Μεταπτυχιακού Διπλώματος), κωδικοποίηση σε αριθμούς 0-6 αντίστοιχα
Κάπνισμα μητρός	Απάντηση της μητέρας στην ερώτηση «Είστε καπνιστής;»	Καπνίστρια μητέρα, κατά σειρά (Όχι, Πρώην, Ναι) κωδικοποίηση σε 0-2
Αρτηριακή Πίεση Συστολική	Μέτρηση διαστολικής αρτηριακής πίεσης. Η μακροχρόνια αύξηση της πίεσης προκαλεί βλάβες στο καρδιαγγειακό σύστημα. [83]	Επίπεδο χοληστερόλης σε mmHg
Game	Πρόσβαση του παιδιού στο παιχνίδι. Τα παιδιά της προπilotικής μελέτης και της ομάδας παρέμβασης της πιλοτικής είχαν πρόσβαση, ενώ τα παιδιά της ομάδας ελέγχου της πιλοτικής μελέτης δεν είχαν.	Είχαν πρόσβαση (1) ή δεν είχαν πρόσβαση (0)
Πόσες ώρες περίπου την ημέρα πέρασε το παιδί σας μπροστά σε οθόνες - ΣΚ ή αργίες;	Απάντηση της μητέρας στην ερώτηση «Πόσες ώρες περίπου την ημέρα (ΣΚ ή αργίες) πέρασε το παιδί σας μπροστά σε οθόνες; (π.χ. τηλεόραση, τάμπλετ, κινητό)»	Χρονική διάρκεια που πέρασε το παιδί μπροστά από οθόνες το Σαββατοκύριακο ή τις αργίες, κατά σειρά (< 1 ώρα / μέρα, 1-2 ώρες / μέρα, 2-3 ώρες / μέρα, 3-4 ώρες / μέρα, > 4 ώρες / μέρα) κωδικοποίηση σε 0-4 αντίστοιχα
Ίκτερος	Εμφάνιση νεογνικού ίκτερου (neonatal jaundice) κατά τη γέννηση	Ναι (1) ή όχι (0)

Πίνακας 4.16: Μετρικές αξιολόγησης του μοντέλου πρόβλεψης της συμμόρφωσης με την εφαρμογή με *Optimal Decision Trees*.

Μετρική / Επίπεδο Συμμόρφωσης	Χαμηλό	Μεσαίο	Υψηλό	Μέσος όρος	Σταθμισμένος Μέσος Όρος
Αριθμός δειγμάτων	24	9	17	—	—
ROC area under the curve	88%	73%	78%	80%	82%
Precision	85%	62%	81%	76%	79%
Recall	92%	56%	76%	75%	80%
f1-score	88%	59%	79%	75%	80%
Specificity	85%	93%	91%	89%	88%
Accuracy	80%			—	—



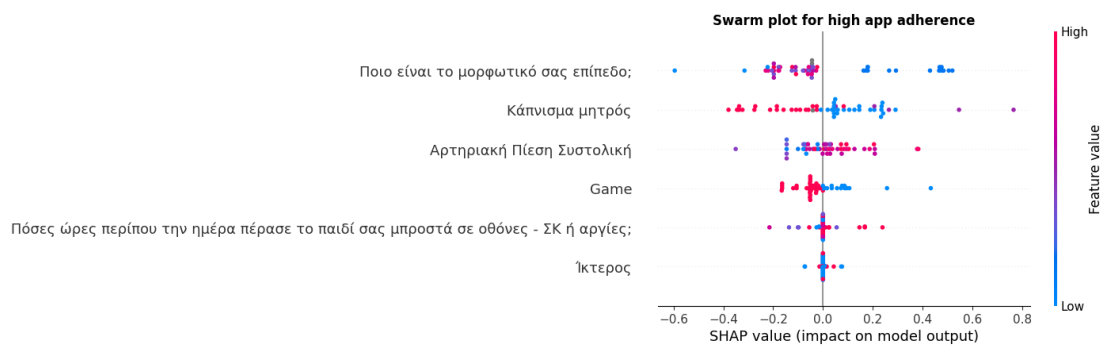
Σχήμα 4.26: Διάγραμμα ποσοτικοποίησης της μέσης επίδρασης κάθε χαρακτηριστικού στο μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με *Optimal Decision Trees*.



Σχήμα 4.27: *Swarm plot* απεικόνισης της επίδρασης κάθε χαρακτηριστικού στην απόφαση ταξινόμησης κάθε δείγματος στην κλάση χαμηλής συμμόρφωσης από το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με *Optimal Decision Trees*.



Σχήμα 4.28: Swarm plot απεικόνισης της επίδρασης κάθε χαρακτηριστικού στην απόφαση ταξινόμησης κάθε δείγματος στην κλάση μεσαίας συμμόρφωσης από το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με *Optimal Decision Trees*.



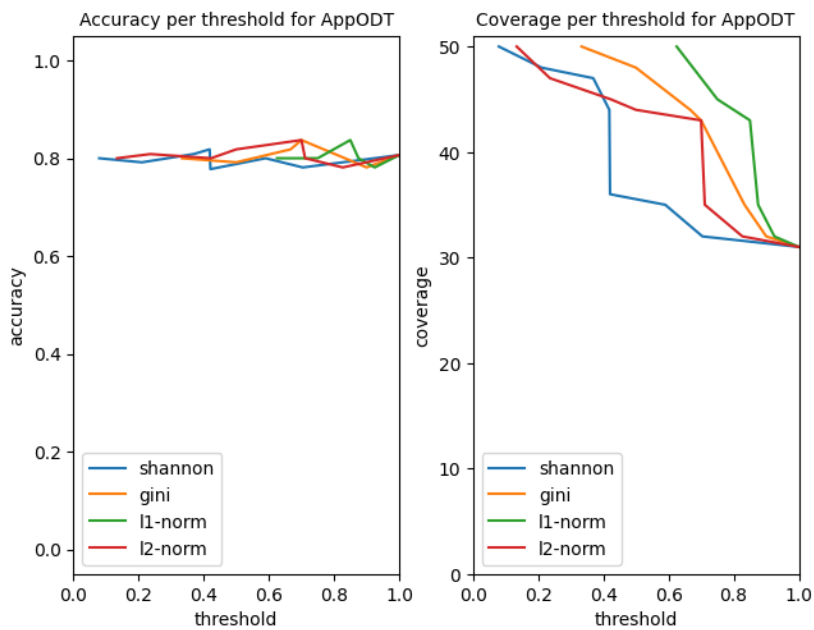
Σχήμα 4.29: Swarm plot απεικόνισης της επίδρασης κάθε χαρακτηριστικού στην απόφαση ταξινόμησης κάθε δείγματος στην κλάση υψηλής συμμόρφωσης από το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με *Optimal Decision Trees*.

Πίνακας 4.17: Επεξήγηση των *swarm plots* για το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με *Optimal Decision Trees*.

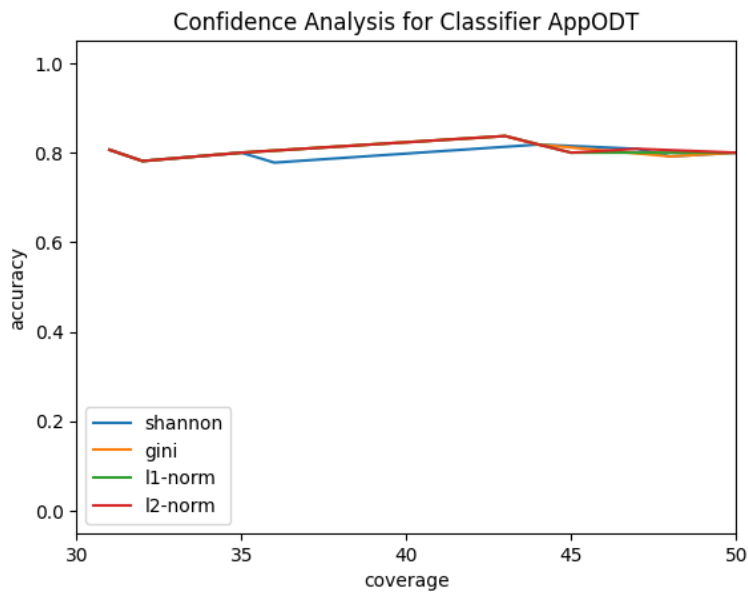
Χαρακτηριστικό	Χαμηλή Συμμόρφωση	Μεσαία Συμμόρφωση	Υψηλή Συμμόρφωση
Ποιο είναι το μορφωτικό σας επίπεδο;	Επηρεάζει αρκετά, με τα παιδιά των οποίων οι μητέρες έχουν δηλώσει ως υψηλότερο μορφωτικό επίπεδο το Γενικό Λύκειο να έχουν μικρότερες πιθανότητες για χαμηλή συμμόρφωση και τα παιδιά των οποίων οι μητέρες έχουν δηλώσει υψηλότερο ή χαμηλότερο μορφωτικό επίπεδο από το Γενικό Λύκειο να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση	Επηρεάζει αρκετά, με τα παιδιά των οποίων οι μητέρες έχουν υψηλότερο μορφωτικό επίπεδο να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει πολύ, με τα παιδιά των οποίων οι μητέρες έχουν δηλώσει ως υψηλότερο μορφωτικό επίπεδο το Γενικό Λύκειο να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και τα παιδιά των οποίων οι μητέρες έχουν δηλώσει υψηλότερο ή χαμηλότερο μορφωτικό επίπεδο από το Γενικό Λύκειο να έχουν λιγότερες πιθανότητες για υψηλή συμμόρφωση
Κάπνισμα μητρός	Επηρεάζει αρκετά, με τα παιδιά των οποίων οι μητέρες είναι νυν καπνίστριες να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει αρκετά, με τα παιδιά των οποίων οι μητέρες είναι νυν καπνίστριες να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει πολύ, με τα παιδιά των οποίων οι μητέρες δεν είναι νυν καπνίστριες να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο

Πίνακας 4.17 (συνέχεια): Επεξήγηση των *swarm plots* για το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με *Optimal Decision Trees*.

Αρτηριακή Πίεση Συστολική	Επηρεάζει αρκετά, με τα παιδιά τα οποία έχουν χαμηλότερες τιμές να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει αρκετά, με τα παιδιά τα οποία έχουν υψηλότερες τιμές να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει αρκετά, με τα παιδιά τα οποία έχουν τις υψηλότερες τιμές να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Game	Επηρεάζει αρκετά, με τα παιδιά τα οποία είχαν πρόσβαση στο παιχνίδι να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει λίγο, με τα παιδιά που είχαν πρόσβαση στο παιχνίδι να φαίνεται να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει αρκετά, με τα παιδιά τα οποία δεν είχαν πρόσβαση στο παιχνίδι να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Πόσες ώρες περίπου την ημέρα πέρασε το παιδί σας μπροστά σε οθόνες - ΣΚ ή αργίες;	Επηρεάζει λίγο, χωρίς ξεκάθαρη μονοτονική σχέση	Φαίνεται να επηρεάζει λίγο, με τα παιδιά που πέρασαν τις περισσότερες ώρες να έχουν λιγότερες πιθανότητες για μεσαία συμμόρφωση	Φαίνεται να επηρεάζει λίγο, χωρίς ξεκάθαρη μονοτονική σχέση
Ίκτερος	Φαίνεται να επηρεάζει ελάχιστα	Φαίνεται να επηρεάζει λίγο, χωρίς ξεκάθαρη μονοτονική σχέση	Φαίνεται να επηρεάζει λίγο, χωρίς ξεκάθαρη μονοτονική σχέση



Σχήμα 4.30: Διαγράμματα της ακρίβειας και της κάλυψης του συνόλου δειγμάτων συναρτήσει του κατωφλίου βεβαιότητας για το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή με *Optimal Decision Trees*. Τα επίπεδα βεβαιότητας καθορίζονται σύμφωνα με τις 4 συναρτήσεις βεβαιότητας (Shannon, Gini, νόρμα επιπέδου 1, νόρμα επιπέδου 2).



Σχήμα 4.31: Διάγραμμα της ακρίβειας του μοντέλου πρόβλεψης της συμμόρφωσης με την εφαρμογή με *Optimal Decision Trees* για ένα υποσύνολο δειγμάτων συναρτήσει της κάλυψης ολόκληρου το συνόλου. Τα υποσύνολα υπολογίζονται για φθίνοντα κατώφλια βεβαιότητας σύμφωνα με τις 4 συναρτήσεις βεβαιότητας (Shannon, Gini, νόρμα επιπέδου 1, νόρμα επιπέδου 2).

4.2.2.3. Συμμόρφωση με το Fitbit

Ακολουθούν η περιγραφή των χαρακτηριστικών, το διάγραμμα συνολικής επίδρασης των χαρακτηριστικών, τα swarm plots και η επεξήγησή τους για τον ταξινομητή πρόβλεψης της συμμόρφωσης με το Fitbit. Ομοίως με την περίπτωση πρόβλεψης της επιμέρους

συμμόρφωσης με την εφαρμογή, το sequential feature selection οδήγησε σε μοντέλο 6 χαρακτηριστικών.

Το μοντέλο επιτυγχάνει ακρίβεια 82%. Το συγκεκριμένο μάλιστα, δεν ταξινομήσε κανένα δείγμα στην κλάση μεσαίας συμμόρφωσης, όπως φαίνεται και από το γεγονός ότι για αυτήν την κλάση σημειώνονται μηδενικά precision και recall και το specificity score είναι 100%. (Πίνακας 4.19)

Οι προβλέψεις του μοντέλου επηρεάστηκαν αποκλειστικά από 4 χαρακτηριστικά: την διάρκεια αποκλειστικού θηλασμού, το στάδιο τριχοφυΐας εφηβαίου κατά Tanner, το κάπνισμα κατά την κύηση και την απάντηση της μητέρας στην ερώτηση «Το οικογενειακό σας εισόδημα καλύπτει τις ανάγκες σας;». Τα υπόλοιπα δύο χαρακτηριστικά, Φάρμακα / Βιταμίνες / Συμπληρώματα και Θηλασμός, δεν επηρέασαν καθόλου τις προβλέψεις (Σχήμα 4.32). Ωστόσο, η παρουσία αυτών των δύο χαρακτηριστικών οδήγησε σε μοντέλο βελτιωμένων επιδόσεων. Αυτό οφείλεται στη μη-ντετερμινιστική φύση του αλγορίθμου κατασκευής των δέντρων με αποτέλεσμα τη δημιουργία καλύτερων δέντρων αποφάσεων με χρήση μόνο των 4 χαρακτηριστικών παρουσία και των δύο επιπλέον στο σύνολο δεδομένων σε σχέση με τα δέντρα που δημιουργήθηκαν χωρίς αυτά.

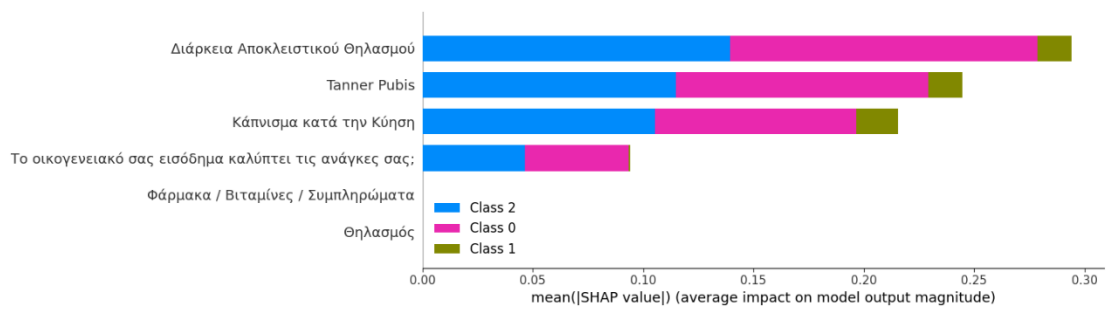
Φαίνεται ότι τα παιδιά που θήλασαν για περισσότερο καιρό αποκλειστικά έχουν περισσότερες πιθανότητες να ταξινομηθούν σε κλάση υψηλότερης συμμόρφωσης και το αντίστροφο.

Πίνακας 4.18: Περιγραφή των χαρακτηριστικών του συνόλου δεδομένων με τα οποία εκπαιδεύτηκε το μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με Optimal Decision Trees.

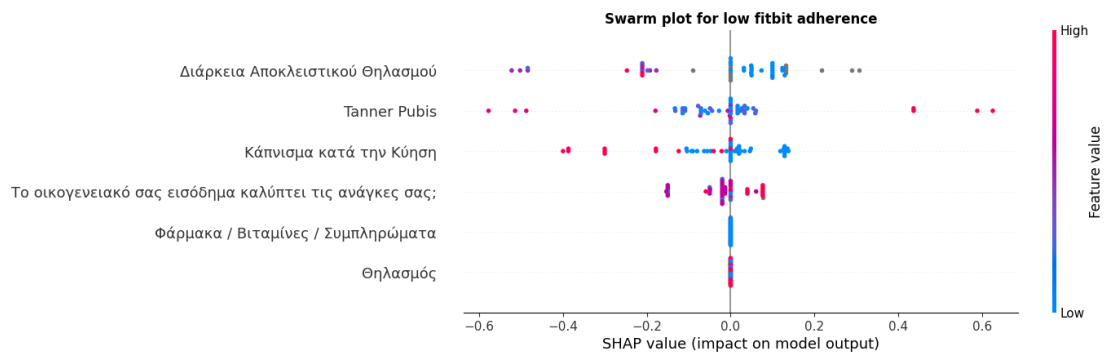
Χαρακτηριστικό	Περιγραφή	Τιμές
Διάρκεια Αποκλειστικού Θηλασμού	Διάρκεια αποκλειστικού θηλασμού σε μήνες. Οι κενές τιμές υποδεικνύουν ότι τα παιδιά δεν θήλασαν ούτε μεικτά.	Αριθμός μηνών αποκλειστικού θηλασμού, το 0 υποδεικνύει ότι υπήρξε διάστημα κατά το οποίο το παιδί θήλαζε μεικτά, η κενή τιμή ότι δεν θήλασε καθόλου
Tanner Pubis	Στάδιο τριχοφυΐας εφηβαίου κατά Tanner. Αποτελεί μέτρο της ανάπτυξης των δευτερευόντων σεξουαλικών χαρακτηριστικών κατά την εφηβεία [84]	Διακριτός αριθμός από 1 έως 5 για το αντίστοιχο στάδιο Tanner
Κάπνισμα κατά την Κύηση	Κάπνισμα της μητέρας κατά την κύηση του παιδιού	Ναι (1) ή όχι (0)
Το οικογενειακό σας εισόδημα καλύπτει τις ανάγκες σας;	Απάντηση της μητέρας στην ερώτηση «Το οικογενειακό σας εισόδημα καλύπτει τις ανάγκες σας;»	Αντίληψη της μητέρας σχετικά με την ικανοποίηση των αναγκών από το οικογενειακό εισόδημα, κατά σειρά (Καθόλου, Λίγο, Μέτρια, Πολύ, Πάρα πολύ) κωδικοποίηση σε 0-4
Φάρμακα / Βιταμίνες / Συμπληρώματα	Λήψη φαρμάκων, βιταμινών ή συμπληρωμάτων από το παιδί	Ναι (1) ή όχι (0)
Θηλασμός	Θηλασμός του παιδιού μετά τη γέννηση	Απάντηση ναι (1) ή όχι (0)

Πίνακας 4.19: Μετρικές αξιολόγησης του μοντέλου πρόβλεψης της συμμόρφωσης με το Fitbit με Optimal Decision Trees.

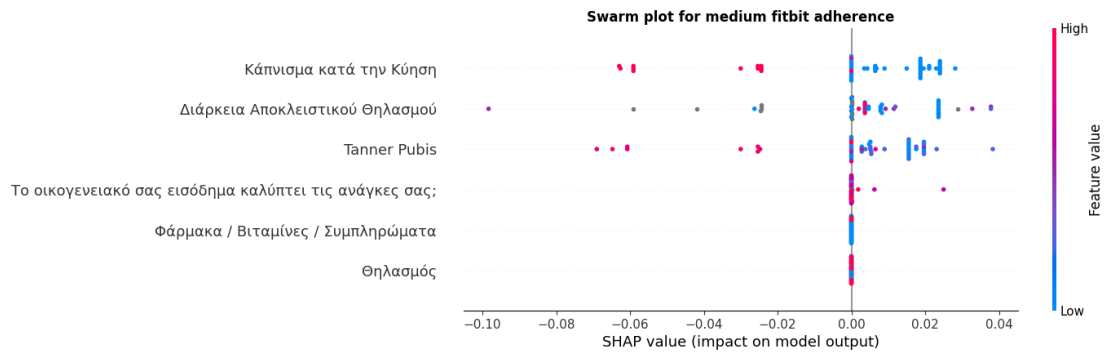
Μετρική / Επίπεδο Συμμόρφωσης	Χαμηλό	Μεσαίο	Υψηλό	Μέσος όρος	Σταθμισμένος Μέσος Όρος
Αριθμός δειγμάτων	24	3	23	—	—
ROC area under the curve	74%	24%	78%	59%	73%
Precision	79%	0%	86%	55%	77%
Recall	92%	0%	83%	58%	82%
f1-score	85%	0%	84%	56%	79%
Specificity	77%	100%	89%	89%	84%
Accuracy	82%			—	—



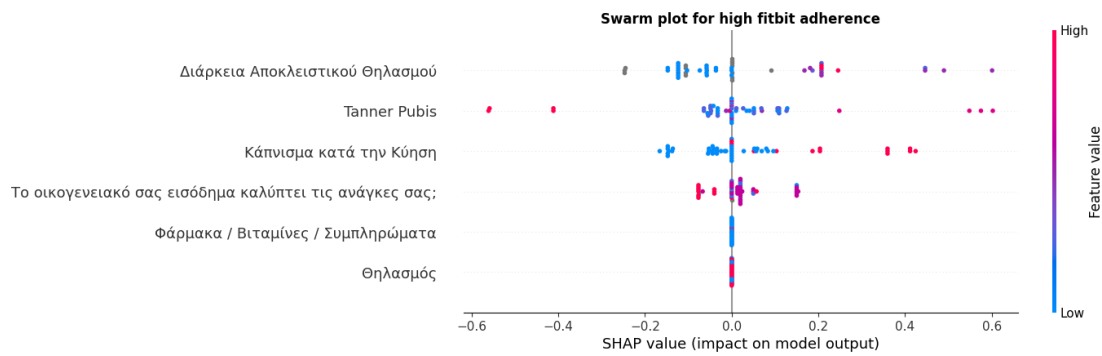
Σχήμα 4.32: Διάγραμμα ποσοτικοποίησης της μέσης επίδρασης κάθε χαρακτηριστικού στο μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με *Optimal Decision Trees*.



Σχήμα 4.33: *Swarm plot* απεικόνισης της επίδρασης κάθε χαρακτηριστικού στην απόφαση ταξινόμησης κάθε δείγματος στην κλάση χαμηλής συμμόρφωσης από το μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με *Optimal Decision Trees*.



Σχήμα 4.34: *Swarm plot* απεικόνισης της επίδρασης κάθε χαρακτηριστικού στην απόφαση ταξινόμησης κάθε δείγματος στην κλάση μεσαίας συμμόρφωσης από το μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με *Optimal Decision Trees*.



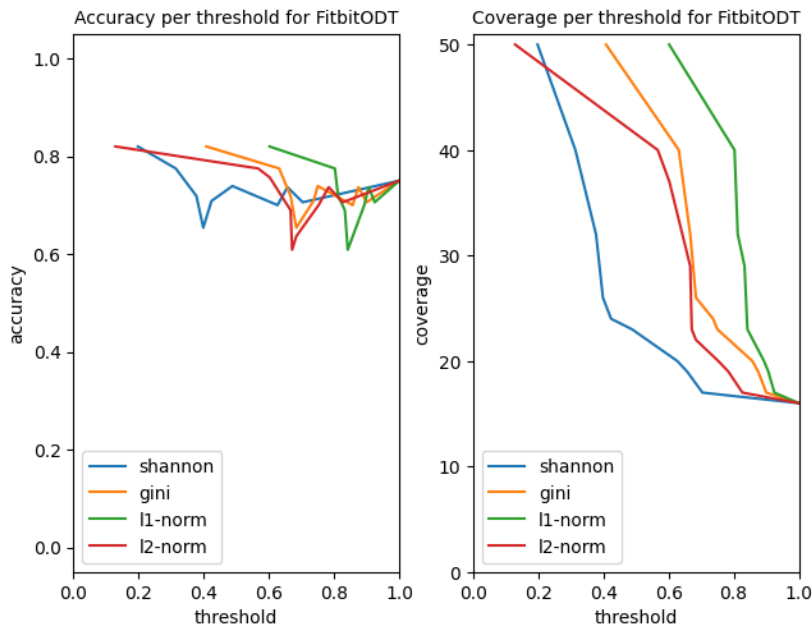
Σχήμα 4.35: Swarm plot απεικόνισης της επίδρασης κάθε χαρακτηριστικού στην απόφαση ταξινόμησης κάθε δείγματος στην κλάση υψηλής συμμόρφωσης από το μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με Optimal Decision Trees.

Πίνακας 4.20: Επεξήγηση των swarm plots για το μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με Optimal Decision Trees.

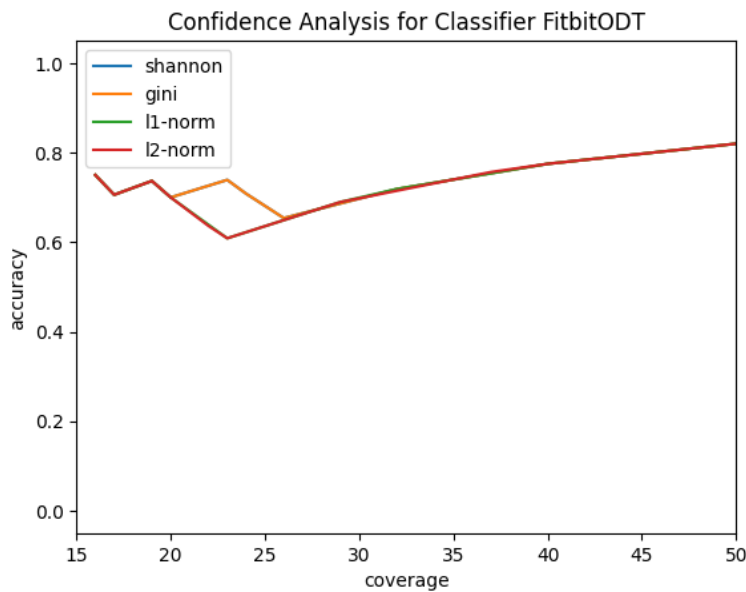
Χαρακτηριστικό	Χαμηλή Συμμόρφωση	Μεσαία Συμμόρφωση	Υψηλή Συμμόρφωση
Διάρκεια Αποκλειστικού Θηλασμού	Επηρεάζει αρκετά, με τα παιδιά τα οποία θήλασαν αποκλειστικά για περισσότερους μήνες να έχουν λιγότερες πιθανότητες για χαμηλή συμμόρφωση και τα παιδιά που θήλασαν για λιγότερους ή δεν θήλασαν να έχουν περισσότερες	Φαίνεται να επηρεάζει λίγο, χωρίς ξεκάθαρη μονότονη σχέση	Επηρεάζει αρκετά, με τα παιδιά τα οποία θήλασαν αποκλειστικά για περισσότερους μήνες να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Tanner Pubis	Φαίνεται να επηρεάζει αρκετά, με τα παιδιά τα οποία ήταν στο στάδιο 5 να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και τα παιδιά τα οποία ήταν στο στάδιο 4 να έχουν λιγότερες πιθανότητες για χαμηλή συμμόρφωση	Επηρεάζει λίγο, με τα παιδιά τα οποία ήταν σε χαμηλότερα στάδια να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Φαίνεται να επηρεάζει αρκετά, με τα παιδιά τα οποία ήταν στο στάδιο 4 να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και τα παιδιά τα οποία ήταν στο στάδιο 5 να έχουν λιγότερες πιθανότητες για υψηλή συμμόρφωση

Πίνακας 4.20 (συνέχεια): Επεξήγηση των swarm plots για το μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με Optimal Decision Trees.

Κάπνισμα κατά την Κύηση	Επηρεάζει αρκετά, με τα παιδιά των οποίων οι μητέρες δεν κάπνιζαν κατά την κύηση να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και το αντίστροφο	Επηρεάζει λίγο, με τα παιδιά των οποίων οι μητέρες δεν κάπνιζαν κατά την κύηση να έχουν περισσότερες πιθανότητες για μεσαία συμμόρφωση και το αντίστροφο	Επηρεάζει αρκετά, με τα παιδιά των οποίων οι μητέρες κάπνιζαν κατά την κύηση να έχουν περισσότερες πιθανότητες για υψηλή συμμόρφωση και το αντίστροφο
Το οικογενειακό σας εισόδημα καλύπτει τις ανάγκες σας;	Φαίνεται να επηρεάζει λίγο, με τα παιδιά των οποίων οι μητέρες απάντησαν 'Πολύ' ή 'Πάρα πολύ' να έχουν περισσότερες πιθανότητες για χαμηλή συμμόρφωση και τα υπόλοιπα λιγότερες	Φαίνεται να επηρεάζει ελάχιστα	Φαίνεται να επηρεάζει λίγο, με τα παιδιά των οποίων οι μητέρες απάντησαν 'Πολύ' ή 'Πάρα πολύ' να έχουν λιγότερες πιθανότητες για υψηλή συμμόρφωση και τα υπόλοιπα λιγότερες
Φάρμακα / Βιταμίνες / Συμπληρώματα	Δεν φαίνεται να επηρεάζει	Δεν φαίνεται να επηρεάζει	Δεν φαίνεται να επηρεάζει
Θηλασμός	Δεν φαίνεται να επηρεάζει	Δεν φαίνεται να επηρεάζει	Δεν φαίνεται να επηρεάζει



Σχήμα 4.36: Διαγράμματα της ακρίβειας και της κάλυψης του συνόλου δειγμάτων συναρτήσει του κατωφλίου βεβαιότητας για το μοντέλο πρόβλεψης της συμμόρφωσης με το Fitbit με *Optimal Decision Trees*. Τα επίπεδα βεβαιότητας καθορίζονται σύμφωνα με τις 4 συναρτήσεις βεβαιότητας (*Shannon*, *Gini*, νόρμα επιπέδου 1, νόρμα επιπέδου 2).



Σχήμα 4.37: Διάγραμμα της ακρίβειας του μοντέλου πρόβλεψης της συμμόρφωσης με το Fitbit με *Optimal Decision Trees* για ένα υποσύνολο δειγμάτων συναρτήσει της κάλυψης ολόκληρου το συνόλου. Τα υποσύνολα υπολογίζονται για φθίνοντα κατώφλια βεβαιότητας σύμφωνα με τις 4 συναρτήσεις βεβαιότητας (*Shannon*, *Gini*, νόρμα επιπέδου 1, νόρμα επιπέδου 2).

4.2.3. Ανίχνευση Προκαταλήψεων στα Δεδομένα

Εξαιτίας του μικρού αριθμού δειγμάτων, είναι πολύ πιθανό τα δεδομένα να μεροληπτούν σε κάποιες από τις μετρούμενες διαστάσεις. Τα μοντέλα μηχανικής μάθησης δύνανται να αποτελέσουν ένα χρήσιμο εργαλείο για την ανίχνευση προκαταλήψεων (*biases*) στο σύνολο

δεδομένων. Επομένως, στην παρούσα ενότητα θα παρουσιαστούν πιθανές προκαταλήψεις στις διαστάσεις του διαθέσιμου συνόλου δεδομένων.

4.2.3.1. Ταξινομητής XGBoost για Πρόβλεψη Συνολικής Συμμόρφωσης

Αρχικά, τα υψηλότερα σκορ υποστηρικτικού τύπου μητέρας φαίνεται να επηρεάζουν θετικά τις πιθανότητες ένα παιδί να έχει χαμηλή συνολική συμμόρφωση. Αυτό επιβεβαιώνεται και από το σύνολο δεδομένων. Από τα 9 παιδιά των οποίων οι μητέρες έχουν σκορ υποστηρικτικού τύπου 4.77 και άνω, τα 5 (55.6%) παρουσίασαν χαμηλή συμμόρφωση, ενώ από τα υπόλοιπα 41 μόνο τα 9 (22.0%) παρουσίασαν χαμηλή συμμόρφωση. Αντίστοιχη μεροληψία φαίνεται να υπάρχει και με τα 11 παιδιά τα οποία γεννήθηκαν 53 εκατοστά και άνω, με το 45.5% αυτών να παρουσιάζει χαμηλή συμμόρφωση, ενώ από τα υπόλοιπα το 23.1%.

Όσον αφορά το μήκος σώματος των παιδιών, παρόμοιο bias φαίνεται να υπάρχει και για την υψηλή συμμόρφωση με το 47.6% των 21 παιδιών που γεννήθηκαν 50 εκατοστά ή λιγότερο να παρουσιάζουν υψηλή συμμόρφωση, με το αντίστοιχο ποσοστό να είναι 21.4% για τα 28 παιδιά που γεννήθηκαν μεγαλύτερα από 50 εκατοστά. Σχετικά με το ύψος των πατέρων το μοντέλο φαίνεται να είναι πιο πιθανό να ταξινομήσει ένα δείγμα στην κλάση της υψηλής συμμόρφωσης αν το ύψος του πατρός είναι μεταξύ 183 και 186 εκατοστά. Πράγματι, στο διάστημα αυτό το ποσοστό των παιδιών που παρουσίασαν υψηλή συμμόρφωση είναι 58.3%, ενώ για τα υπόλοιπα ύψη είναι 21.6%.

4.2.3.2. Ταξινομητής XGBoost για Πρόβλεψη Συμμόρφωσης με την Εφαρμογή

Το μοντέλο πρόβλεψης της συμμόρφωσης με την εφαρμογή είναι πιο πιθανό να ταξινομήσει ένα παιδί σε υψηλότερη κλάση συμμόρφωσης αν η μητέρα του καταναλώνει συχνότερα αλκοόλ. Πράγματι, από τις 25 μητέρες που δήλωσαν ότι δεν καταναλώνουν καθόλου αλκοόλ τα παιδιά των 15 (60.0%) είχαν χαμηλή συμμόρφωση, ενώ από τις 23 που δήλωσαν ότι καταναλώνουν αλκοόλ τα παιδιά των 7 (30.4%) είχαν χαμηλή συμμόρφωση. Και τα δύο δείγματα για τα οποία η απάντηση της μητέρας στη σχετική ερώτηση απουσίαζε ανήκουν στην κλάση της χαμηλής συμμόρφωσης.

Πιθανώς να υπάρχει αντίστοιχη μεροληψία και στην ερώτηση «Ποιο είναι το μορφωτικό σας επίπεδο;». Τα παιδιά 13 από τις 29 (44.8%) μητέρες που δήλωσαν ότι το υψηλότερο μορφωτικό τους επίπεδο είναι Γυμνάσιο, Γενικό Λύκειο ή ΙΕΚ παρουσίασαν υψηλή συμμόρφωση, εν αντιθέσει με τα παιδιά 4 από τις 19 (21.1%) μητέρες που δήλωσαν ότι το υψηλότερο μορφωτικό τους επίπεδο είναι ΤΕΙ, ΑΕΙ ή Μεταπτυχιακό Δίπλωμα.

4.2.3.3. Ταξινομητής XGBoost για Πρόβλεψη Συμμόρφωσης με το Fitbit

Όπως και για τον ταξινομητή πρόβλεψης της συμμόρφωσης με την εφαρμογή, έτσι και για τον ταξινομητή πρόβλεψης της συμμόρφωσης με το Fitbit φαίνεται να υπάρχει μεροληψία όσον αφορά την ερώτηση «Καταναλώνετε αλκοολούχα ποτά;». Οι μητέρες και των τριών παιδιών που παρουσίασαν μεσαίο επίπεδο συμμόρφωσης με το fitbit δήλωσαν ότι δεν καταναλώνουν αλκοόλ, ενώ από τις 27 μητέρες συνολικά που είτε δήλωσαν ότι δεν καταναλώνουν αλκοόλ είτε λείπει η απάντησή τους, τα παιδιά των 10 (37.0%) είχαν υψηλό επίπεδο συμμόρφωσης. Από τις 23 που δήλωσαν ότι καταναλώνουν αλκοόλ, τα παιδιά των 13 (56.5%) παρουσίασαν υψηλό επίπεδο συμμόρφωσης.

Τα δεδομένα που αφορούν τον *ίκτηρο* φαίνεται να είναι *biased*, καθώς από τα 9 παιδιά που είχαν εμφανίσει νεογνικό *ίκτηρο*, τα 7 (77.8%) παρουσίασαν υψηλή συμμόρφωση, ενώ από τα υπόλοιπα 41 μόνο τα 16 (39.0%) παρουσίασαν υψηλή συμμόρφωση. Αντίστοιχα, 9 από τα 12 (75.0%) παιδιά που *παρέμειναν σε Μονάδα Εντατικής Νοσηλείας ή Μονάδα Αυξημένης Φροντίδας* μετά τη γέννησή τους παρουσίασαν υψηλή συμμόρφωση, με το αντίστοιχο ποσοστό για τα υπόλοιπα να ανέρχεται μόλις στο 36.8%.

4.2.3.4. Ταξινομητής *Optimal Tree Classifier* για Πρόβλεψη Συνολικής Συμμόρφωσης

Το μοντέλο πρόβλεψης της συνολικής συμμόρφωσης με *Optimal Decision Trees* είναι πιο πιθανό να ταξινομήσει ένα δείγμα στη μεσαία κλάση συμμόρφωσης αν οι τιμές του δείκτη *SGOT* είναι χαμηλότερες για εκείνο το παιδί. Στο σύνολο δεδομένων, από τα 16 δείγματα τα οποία είχαν δείκτη *SGOT* μικρότερο από 20 μονάδες ανά λίτρο (U/L) τα 11 (62.5%) σημείωσαν μεσαία συμμόρφωση, έναντι 9 δειγμάτων (26.5%) από τα 34 με δείκτη *SGOT* μεγαλύτερο των 20 U/L.

Μέσω των πιθανοτήτων ταξινόμησης του μοντέλου πρόβλεψης της συνολικής συμμόρφωσης ανιχνεύεται το γεγονός ότι τα παιδιά και των τριών μητέρων οι οποίες δήλωσαν ότι το υψηλότερο *μορφωτικό επίπεδο* που έχουν ολοκληρώσει είναι το Γυμνάσιο σημείωσαν χαμηλή συνολική συμμόρφωση, ενώ από τα παιδιά των 15 μητέρων οι οποίες δήλωσαν το Γενικό Λύκειο ως το *μορφωτικό* του επιπέδο, τα 7 (46.7%) σημείωσαν υψηλή συμμόρφωση, έναντι 25.7% για τα υπόλοιπα 35 παιδιά.

4.2.3.5. Ταξινομητής *Optimal Tree Classifier* για Πρόβλεψη Συμμόρφωσης με την Εφαρμογή

Αντίστοιχα με τον *Optimal Tree Classifier* για τη συνολική συμμόρφωση, από τις 15 μητέρες οι οποίες δήλωσαν ως *μορφωτικό επίπεδο* το Γενικό Λύκειο οι 9 (60%) παρουσίασαν υψηλή συμμόρφωση με την εφαρμογή, έναντι 8 (22.9%) από τις υπόλοιπες 35. Από την άλλη, μόλις 4 (26.7%) από τις 15 μητέρες που δήλωσαν το Γενικό Λύκειο ως το *μορφωτικό* τους επίπεδο σημείωσαν χαμηλή συμμόρφωση με την εφαρμογή, έναντι 20 (57.1%) από τις υπόλοιπες 35. Τα παραπάνω έρχονται να συμπληρώσουν την προκατάληψη που έδειξε ο ταξινομητής *XGBoost* στη σχέση μεταξύ του *μορφωτικού επιπέδου* και της συμμόρφωσης με την εφαρμογή.

Από τα 15 παιδιά των οποίων η *συστολική αρτηριακή πίεση* μετρήθηκε μεγαλύτερη από 120 *mmHg* τα 4 (26.7%) και τα 7 (46.7%) σημείωσαν χαμηλή και υψηλή συμμόρφωση με την εφαρμογή, αντίστοιχα. Για τα υπόλοιπα 35 παιδιά, ο αντίστοιχος αριθμός δειγμάτων ήταν 20 (57.1%) για χαμηλή και 10 (28.6%) για υψηλή συμμόρφωση.

4.2.3.6. Ταξινομητής *Optimal Tree Classifier* για Πρόβλεψη Συμμόρφωσης με το *Fitbit*

Από τα αποτελέσματα του μοντέλου πρόβλεψης της συμμόρφωσης με το *Fitbit*, φαίνεται να υπάρχει μία προκατάληψη στα δεδομένα που αφορούν το *κάπνισμα κατά την κύηση*. Τα παιδιά 7 (53.8%) από τις 13 μητέρες οι οποίες κάπνιζαν κατά την κύηση σημείωσαν υψηλή συμμόρφωση, έναντι 16 (43.2%) παιδιών από τις 37 μητέρες οι οποίες δεν κάπνιζαν.

4 (80%) από τα 5 παιδιά που κατά την έναρξη του πλάνου παρέμβασης βρίσκονταν στο στάδιο 5 τριχοφυΐας εφηβαίου κατά *Tanner* ανήκουν στην κλάση χαμηλής συμμόρφωσης με το *Fitbit* και 5 (83.3%) από τα 6 παιδιά που βρίσκονταν στο στάδιο 4 ανήκουν στην κλάση υψηλής συμμόρφωσης. Το ποσοστό χαμηλής συμμόρφωσης για τα 45 παιδιά που δεν είχαν

φτάσει ακόμα το στάδιο 5 ανέρχεται σε 44.4% και το ποσοστό υψηλής συμμόρφωσης για τα 44 παιδιά που δεν βρίσκονταν στο στάδιο 4 είναι 40.1%. Κανένα από τα τρία παιδιά με μέτρια συμμόρφωση με το Fitbit δεν βρισκόταν ακόμα στα στάδια 4 ή 5.

Τέλος, φαίνεται να υπάρχει μία μικρή πώλωση όσον αφορά τις απαντήσεις στην ερώτηση «*Το οικογενειακό σας εισόδημα καλύπτει τις ανάγκες σας;*». Συγκεκριμένα, από τα 12 δείγματα για τα οποία οι απαντήσεις ήταν «Πολύ» ή «Πάρα Πολύ», τα 5 (41.7%) ανήκουν στην κλάση υψηλής συμμόρφωσης με το Fitbit, ενώ από τα 36 δείγματα με απαντήσεις που δείχνουν μικρότερο βαθμό κάλυψης των αναγκών, τα 18 (50%) σημείωσαν υψηλή συμμόρφωση (έλειπε η απάντηση στην ερώτηση για δύο δείγματα).

5. ΑΝΑΠΤΥΞΗ ΣΥΣΤΗΜΑΤΟΣ ΠΑΡΑΓΩΓΗΣ ΕΞΑΤΟΜΙΚΕΥΜΕΝΟΥ ΠΕΡΙΕΧΟΜΕΝΟΥ

Αναπτύχθηκε σύστημα συστάσεων προτρεπτικών μηνυμάτων με στόχο την αύξηση της συμμόρφωσης με τη θεραπεία και της αποτελεσματικότητάς της. Εκπαιδεύτηκαν παράγοντες ενισχυτικής μάθησης χρησιμοποιώντας τα δεδομένα των μηνυμάτων που έλαβαν τα παιδιά και των καταγραφών της εφαρμογής και του Fitbit. Μέσα από την υλοποίηση ενός συστήματος βασισμένου στον αλγόριθμο Q-learning για διαφορετικές μοντελοποιήσεις του προβλήματος εξάχθηκαν και αξιολογήθηκαν πολιτικές με στόχο τη βελτιστοποίηση της συμμόρφωσης και της έκβασης της θεραπείας.

5.1. ΜΕΘΟΔΟΛΟΓΙΑ

5.1.1. Δεδομένα

Καθημερινά, μέσω της πλατφόρμας ENDORSE αποστέλλονταν, τόσο στις μητέρες μέσω της εφαρμογής, όσο και στα παιδιά μέσω του παιχνιδιού, μηνύματα, τα οποία είχαν εκπαιδευτικό και προτρεπτικό χαρακτήρα. Για παράδειγμα, οι μητέρες λάμβαναν μηνύματα, όπως «Το παιδί συστήνεται να αφιερώνει τουλάχιστον μια ώρα τη μέρα για σωματική δραστηριότητα μέτριας ή υψηλής έντασης!» ή «Γάλα, γιαούρτι ή τυρί καθημερινά», ενώ τα παιδιά λάμβαναν μηνύματα, όπως «Σωματική δραστηριότητα και μια ισορροπημένη διατροφή αποτελούν τον καλύτερο τρόπο για να ρυθμίσω το βάρος μου και να μειώσω το αποθηκευμένο λίπος!» ή «Είναι καλό να μην παραλείπεις κύρια γεύματα!». Τα υποψήφια προς αποστολή μηνύματα ήταν καθορισμένα εκ των προτέρων και χωρισμένα σε 11 κατηγορίες, με κάθε κατηγορία να περιέχει συναφή μηνύματα, απευθυνόμενα στις μητέρες και στα παιδιά.

Να σημειωθεί ότι οι χρήστες της ομάδας ελέγχου της πιλοτικής μελέτης δεν λάμβαναν μηνύματα μέσω του παραπάνω συστήματος. Επομένως, το παρόν κομμάτι της εργασίας αφορά τους συνολικά 35 χρήστες της προπιλοτικής και της ομάδας παρέμβασης της πιλοτικής μελέτης.

Στους 20 χρήστες της πιλοτικής μελέτης αποστέλλονταν κάθε μέρα 6 συνολικά μηνύματα, 3 στη μητέρα και 3 στο παιδί, αν και υπήρξαν μέρες που μπορεί σε κάποιους χρήστες να στάλθηκαν περισσότερα (το πολύ 8) ή λιγότερα (τουλάχιστον 1) συνολικά μηνύματα. Οι 15 χρήστες της ομάδας παρέμβασης της πιλοτικής μελέτης έλαβαν κάθε μέρα 4 ακριβώς μηνύματα, 2 η μητέρα και 2 το παιδί.

Η επιλογή των μηνυμάτων προς αποστολή γινόταν μέσω ενός γενετικού αλγορίθμου (genetic algorithm) στον οποίο κάθε γονίδιο ήταν μια κατηγορία μηνυμάτων, ενώ η αξιολόγηση και η ενημέρωση των κατηγοριών προς αποστολή πραγματοποιούνταν κάθε εβδομάδα, με βάση τη συμμόρφωση με την εφαρμογή και το fitbit, καθώς και την επίτευξη των στόχων βάρους, βημάτων, ύπνου και του εβδομαδιαίου στόχου σωματικής άσκησης ή γεύματος [10], [11].

Ταυτόχρονα, για κάθε μέρα υπήρχαν διαθέσιμα για κάθε χρήστη δεδομένα αυτοπαρακολούθησης από την εφαρμογή, όπως και οι ημερήσιες καταγραφές βημάτων και ύπνου από το Fitbit. Κάθε εβδομάδα ανατίθονταν στα παιδιά ένας με δύο εβδομαδιαίοι στόχοι, οι οποίοι αφορούσαν είτε τη σωματική άσκηση είτε την κατανάλωση υγιεινών τροφών σε ένα συγκεκριμένο γεύμα (πρωινό, δεκατιανό, μεσημεριανό, απογευματινό,

βραδινό). Οι χρήστες της προπλοτικής μελέτης μπορούσαν να επιλέγουν έως και δύο εβδομαδιαίους στόχους μέσω της εφαρμογής, ενώ σε όλους τους χρήστες της πιλοτικής μελέτης ανατιθόταν ο ίδιος στόχος, προκαθορισμένος για κάθε εβδομάδα. Οι μητέρες έπρεπε να καταγράφουν στην εφαρμογή πληροφορίες σχετικές με τον ανατιθέμενο εβδομαδιαίο στόχο: π.χ. αν ο στόχος ήταν το πρωινό, οι μητέρες έπρεπε να καταγράφουν τις τροφές που έφαγαν τα παιδιά τους για πρωινό κάθε μέρα της εβδομάδας και αν οι τροφές αυτές ήταν από τις διαθέσιμες επιλογές του διατροφικού πλάνου [11]. Επομένως, μεταξύ άλλων λεπτομερειών οι οποίες δεν αξιοποιήθηκαν στην παρούσα εργασία, από την εφαρμογή ήταν διαθέσιμο για κάθε μέρα το αν η μητέρα προχώρησε στην ημερήσια καταγραφή του στόχου, το αν επιτεύχθη ο στόχος, ο οποίος είχε τεθεί από την κλινική ομάδα, και η καταγραφή του βάρους του παιδιού.

Από το Fitbit, ήταν διαθέσιμα για κάθε μέρα ο αριθμός των βημάτων που καταγράφηκαν, ο αριθμός των λεπτών ύπνου και οι συνολικές ώρες χρήσης του Fitbit.

Συνολικά στάλθηκαν 11,236 μηνύματα στους 35 χρήστες. Ο αριθμός των διαθέσιμων σειρών (rows) δεδομένων ήταν 2,321, με κάθε σειρά να αντιστοιχεί σε μία μέρα για κάθε ξεχωριστό χρήστη. Τα δεδομένα που περιέχει κάθε σειρά είναι ο αριθμός μηνυμάτων από κάθε κατηγορία που έλαβε ο χρήστης, τα δεδομένα αυτοπαρακολούθησης από την εφαρμογή και οι μετρήσεις του Fitbit για τη συγκεκριμένη ημέρα.

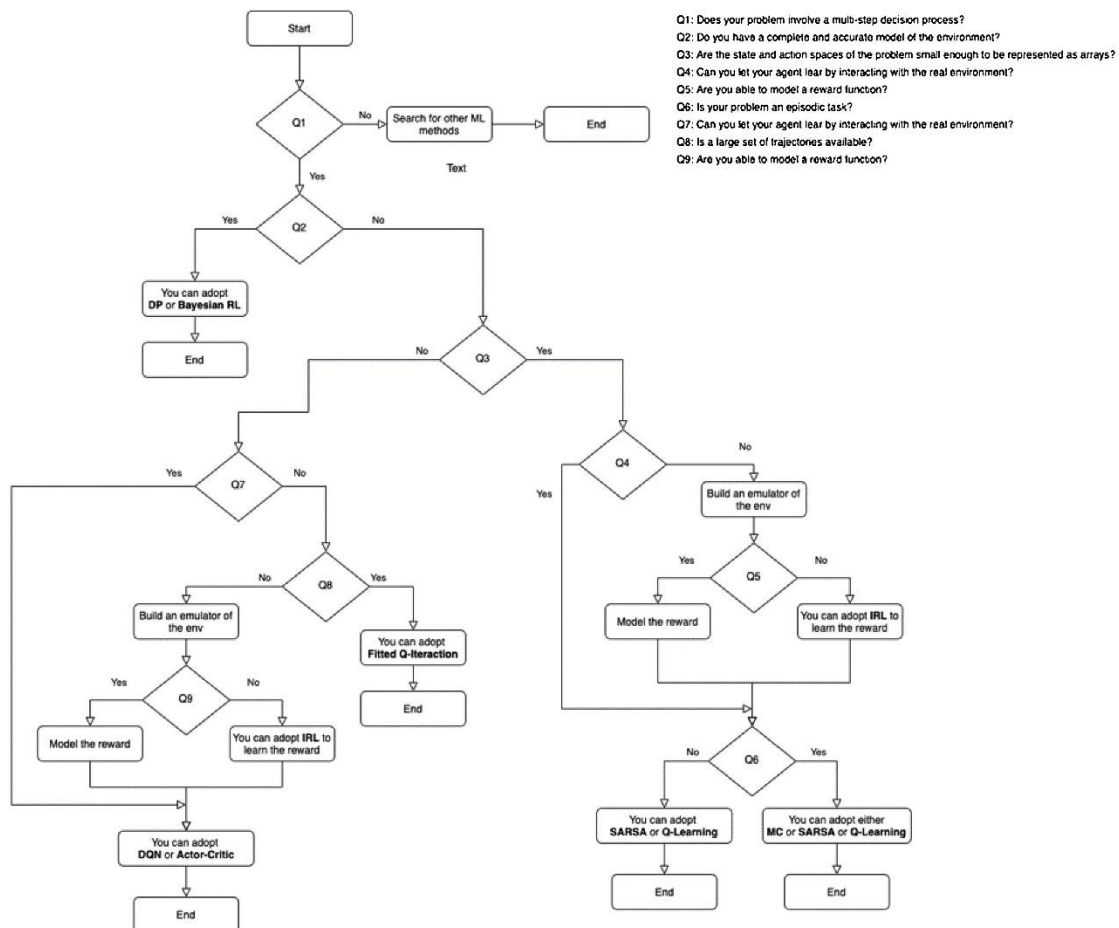
Παρόλο που η διάρκεια του πλάνου παρέμβασης ήταν 12 εβδομάδες για όλους τους χρήστες, έλειπαν δεδομένα για αρκετές ημέρες. Βασικότερη αιτία υπήρξε το γεγονός ότι το σύστημα ENDORSE ήταν εκτός λειτουργίας για κάποιες εβδομάδες κατά τη διάρκεια της προπλοτικής μελέτης, με αποτέλεσμα να μη λειτουργεί η εφαρμογή. Ως εκ τούτου, στο συγκεκριμένο χρονικό διάστημα, δεν κατέστη δυνατή η αποστολή μηνυμάτων και η καταγραφή δεδομένων αυτοπαρακολούθησης. Επομένως, για κάποιους χρήστες υπάρχουν λιγότερες σειρές δεδομένων. Ο μικρότερος αριθμός σειρών που αντιστοιχούν σε έναν χρήστη είναι 53 και ο μεγαλύτερος 84.

Σημειώνεται πως, εκτός από τα παραγόμενα από τον γενετικό αλγόριθμο μηνύματα, οι χρήστες λάμβαναν και επιπλέον μηνύματα, τα οποία περιλάμβαναν συμβουλές σχετικές με τον εβδομαδιαίο στόχο, παρείχαν εβδομαδιαίες αναφορές στις μητέρες σχετικά με τις καταγραφές του Fitbit και τη χρήση του παιχνιδιού από τα παιδιά και εξυπηρέτούσαν ενημερωτικούς και διαχειριστικούς σκοπούς σχετικά με τη διαθεσιμότητα της πλατφόρμας. Καθώς τα μηνύματα των παραπάνω ειδών αποστέλλονταν σε συγκεκριμένες χρονικές στιγμές, κοινές για όλους τους χρήστες, δεν μπορούν να αξιοποιηθούν για την εκπαίδευση μοντέλων μηχανικής μάθησης, άρα και για τους σκοπούς της παρούσας εργασίας.

5.1.2. Επιλογή του αλγορίθμου

Στην παρούσα εργασία, ήταν διαθέσιμο ένα σταθερό σύνολο δεδομένων, επομένως μπορούσε να εφαρμοστεί μόνο offline ενισχυτική μάθηση.

Προκειμένου να επιλεγεί η κατάλληλη τεχνική ενισχυτικής μάθησης ακολουθήθηκε η μεθοδολογία από το δημοσίευμα των Coronato κ.ά., η οποία απεικονίζεται στο παρακάτω διάγραμμα [85].



Σχήμα 5.1: Αναπαράσταση ως δέντρο αποφάσεων της μεθοδολογίας για την επιλογή του κατάλληλου αλγορίθμου ενισχυτικής μάθησης για εφαρμογές στον χώρο της υγείας. Πηγή: [85]

Συγκεκριμένα, το πρόβλημα της παρούσας εργασίας περιλαμβάνει μία διαδικασία λήψης αποφάσεων πολλαπλών βημάτων (Q1) και δεν είναι διαθέσιμο κάποιο πλήρες και ακριβές μοντέλο του περιβάλλοντος (Q2).

Όσον αφορά την (Q3), οι χώροι καταστάσεων και ενεργειών (state/action spaces) δεν ήταν εξαρχής αρκετά μικροί για να απεικονισθούν ως πίνακες (arrays), ωστόσο ήταν δυνατό να μειωθούν χωρίς σημαντική απώλεια πληροφορίας, όπως και έγινε για μια σειρά από λόγους:

- Με κατάλληλη μείωση των χώρων καταστάσεων και ενεργειών διατηρούνται τα πλέον σημαντικά κομμάτια πληροφορίας και μειώνεται η πιθανότητα υπερπροσαρμογής (overfitting) του μοντέλου, δηλαδή να έχει υψηλή απόδοση στα δεδομένα εκπαίδευσης, αλλά χαμηλή σε οποιαδήποτε νέα δεδομένα, για το περιορισμένο διαθέσιμο σύνολο δεδομένων. [63]
- Μειώνεται η πολυπλοκότητα του μοντέλου επιτρέποντας καλύτερη εκπαίδευση για περιορισμένους υπολογιστικούς πόρους.
- Ακόμα, καθίσταται δυνατή η εφαρμογή περισσότερων και απλούστερων αλγορίθμων. Η χρήση απλούστερων αλγορίθμων αυξάνει σημαντικά την ερμηνευσιμότητα του μοντέλου, μια ιδιότητα ιδιαίτερα σημαντική για την αποδοχή μιας εφαρμογής στον χώρο της υγείας.

Όπως προαναφέρθηκε, δεν ήταν δυνατή η αλληλεπίδραση του παράγοντα με το πραγματικό περιβάλλον (Q4). Το συστηνόμενο βήμα ανάπτυξης προσομοιωτή για το περιβάλλον

παρακάμφθηκε, καθώς στη συγκεκριμένη εφαρμογή είναι καταλληλότερη η offline εκπαίδευση του μοντέλου. Η μοντελοποίηση μιας συνάρτησης ανταμοιβής ήταν δυνατή (Q5), αλλά και να μην ήταν εύκολη, δεν θα ήταν δυνατό να χρησιμοποιηθεί αντίστροφη ενισχυτική μάθηση (inverse reinforcement learning), καθώς στα δεδομένα ακολουθείται η πολιτική που καθορίζεται από τον γενετικό αλγόριθμο και όχι από κάποιον ειδικό. Τέλος, το πρόβλημα είναι επεισοδιακό (episodic) (Q6), δηλαδή μπορεί να χωριστεί σε επεισόδια με πεπερασμένο χρονικό ορίζοντα, εν προκειμένω το τέλος της παρέμβασης για κάθε χρήστη, σε αντίθεση με τις συνεχείς (continuous) διαδικασίες, οι οποίες αποσκοπούν στη μεγιστοποίηση της συνολικής ανταμοιβής για αόριστο χρόνο [76].

Επομένως, σύμφωνα πάντα με την προτεινόμενη μεθοδολογία, προτείνεται η υιοθέτηση ενός αλγορίθμου από τρεις υποψήφιες κατηγορίες: Monte Carlo, SARSA και Q-learning.

Οι αλγόριθμοι ενισχυτικής μάθησης Monte Carlo χρησιμοποιούνται αποκλειστικά για επεισοδιακές διαδικασίες. Οι συναρτήσεις αξίας αρχικοποιούνται και ενημερώνονται μόνο μετά την ολοκλήρωση ενός επεισοδίου, όταν και είναι γνωστή η πραγματική συνολική ανταμοιβή του επεισοδίου. Στο τέλος, η συνάρτηση αξίας για μία κατάσταση (ή ένα ζεύγος κατάστασης-ενέργειας) λαμβάνει τη μέση τιμή της συνολικής ανταμοιβής μεταξύ πολλών επεισοδίων στα οποία έχει εμφανιστεί η κατάσταση (ή το ζεύγος κατάστασης-ενέργειας). Η χρήση αλγορίθμου Monte Carlo απορρίφθηκε, λόγω του μικρού αριθμού διαθέσιμων επεισοδίων στο σύνολο δεδομένων, 35, ένα για κάθε χρήστη. Ο αλγόριθμος SARSA μαθαίνει on-policy, οπότε είναι εξ' ορισμού ακατάλληλος για offline ενισχυτική μάθηση. Κατέστη δυνατή, ωστόσο, η χρήση Q-learning για το σύνολο δεδομένων. [76], [85]

5.1.3. Διατύπωση του προβλήματος

Όταν χρησιμοποιείται προγραμματισμός για την επίλυση ενός προβλήματος, είναι σημαντική η αφαίρεση (abstraction) των σημαντικότερων πληροφοριών και η κωδικοποίησή τους σε μορφή αναγνωρίσιμη από τον υπολογιστή. Ειδικά στην ενισχυτική μάθηση, η επιλογή των χαρακτηριστικών τα οποία θα περιγράψουν τους χώρους καταστάσεων και ενεργειών (state/action spaces) και τη συνάρτηση ανταμοιβής (reward function) ενδέχεται να καθορίσει και τις τεχνικές οι οποίες μπορούν να εφαρμοστούν.

Η κατάσταση κάθε χρονική στιγμή, εν προκειμένω μια συγκεκριμένη μέρα για κάποιον συγκεκριμένο χρήστη, αποτελείται από 7 μεταβλητές, οι οποίες παρουσιάζονται στον Πίνακα 5.1:

Πίνακας 5.1: Περιγραφή των χαρακτηριστικών του συνόλου δεδομένων τα οποία συνιστούν την κατάσταση σε κάθε χρονική στιγμή για τα μοντέλα συστάσεων αποστολής μηνυμάτων.

Όνομα	Χαρακτηριστικό	Τιμές	Προέλευση	Ζητούμενο
daily monitoring	Ημερήσια καταγραφή αυτοπαρακολούθησης	0 αν δεν εισήχθη καταγραφή, 1 αν εισήχθη	Εφαρμογή	Συμμόρφωση
weight monitoring	Εβδομαδιαία καταγραφή βάρους	0 αν δεν εισήχθη καταγραφή, 1 αν εισήχθη	Εφαρμογή	Συμμόρφωση
fitbit hours	Ώρες χρήσης του Fitbit	0 για 0-1 ώρες, 1 για 1-10 ώρες, 2 για 10 ώρες και άνω	Fitbit	Συμμόρφωση
daily target	Επίτευξη ημερήσιου στόχου αυτοπαρακολούθησης	0 αν δεν εισήχθη καταγραφή ή εισήχθη και δεν επιτεύχθη ο στόχος, 1 αν εισήχθη καταγραφή και επιτεύχθη ο στόχος	Εφαρμογή	Έκβαση
weight difference	Διαφορά βάρους	0 αν δεν εισήχθη καταγραφή ή εισήχθη και το βάρος αυξήθηκε, 1 αν εισήχθη καταγραφή και το βάρος μειώθηκε ή έμεινε σταθερό	Εφαρμογή	Έκβαση
step count	Αριθμός βημάτων	0 για 0-10,000 βήματα, 1 για 10,000 βήματα και άνω	Fitbit	Έκβαση
sleep hours	Ώρες ύπνου	0 για 0-7 ώρες, 1 για 7-9 ώρες, 2 για 9 ώρες και άνω	Fitbit	Έκβαση

Όπως φαίνεται, το σύνολο χαρακτηριστικών αποτελείται από 5 δυαδικές και 2 τριαδικές μεταβλητές, με αποτέλεσμα έναν χώρο 288 διαφορετικών καταστάσεων, εκ των οποίων δεν θα είχαν νόημα καταστάσεις στις οποίες:

- Η ημερήσια καταγραφή αυτοπαρακολούθησης είναι 0 και η επίτευξη του ημερήσιου στόχου είναι 1.
- Η εβδομαδιαία καταγραφή βάρους είναι 0 και η διαφορά βάρους είναι 1.
- Οι ώρες χρήσεις του Fitbit είναι 0 και ο αριθμός βημάτων είναι 1 ή οι ώρες ύπνου είναι 1 ή 2.

Τηρουμένων των παραπάνω περιορισμών, το σύνολο πιθανών καταστάσεων είναι 117 εκ των οποίων οι 50 εμφανίζονται στο σύνολο δεδομένων.

Στην παρούσα εργασία, η συνάρτηση ανταμοιβής για μία χρονική στιγμή εξαρτάται αποκλειστικά από την κατάσταση την τρέχουσα χρονική στιγμή. Επομένως, η επιλογή των χαρακτηριστικών, όπως και η διακριτοποίησή τους έγιναν με στόχο τη δημιουργία της κατάλληλης συνάρτησης ανταμοιβής.

Τρεις μεταβλητές της κατάστασης σχετίζονται με τη συμμόρφωση με το πλάνο παρέμβασης ENDORSE: η ημερήσια καταγραφή αυτοπαρακολούθησης, η εβδομαδιαία καταγραφή βάρους και οι ώρες χρήσης του Fitbit. Τα κατώφλια για την κατηγοριοποίηση του χρόνου χρήσης του Fitbit ορίστηκαν στη 1 και στις 10 ώρες, ούτως ώστε να υπάρχει μια ξεχωριστή κλάση για τις ημέρες στις οποίες τα παιδιά έκαναν μηδενική ή υπερβολικά βραχεία χρήση του Fitbit, αλλά και για τις ημέρες στις οποίες τα παιδιά φόρεσαν το Fitbit για πάνω από 10 ώρες (πιθανότατα κατά τη διάρκεια του ύπνου συν κάποιες ώρες ακόμα).

Από την άλλη, οι υπόλοιπες τέσσερις μεταβλητές της κατάστασης αποτελούν ενδείξεις τις έκβασης της θεραπείας, δηλαδή ότι τα παιδιά αναπτύσσουν υγιεινές συνήθειες και σημειώνουν μείωση του BMI τους. Με το σκεπτικό ότι για παιδιά τα οποία ακόμα ψηλώνουν, το να μείνει σταθερό το βάρος τους μπορεί να σημαίνει μείωση του BMI, οι ημέρες στις οποίες δεν παρατηρήθηκε μεταβολή στην καταγραφή του βάρους ταξινομήθηκαν στην ίδια κλάση με αυτές στις οποίες μειώθηκε. Το κατώφλι για τον αριθμό των βημάτων ορίστηκε στα 10,000 βήματα, τα οποία αποτελούν τον ημερήσιο στόχο του ENDORSE [11], αλλά και την ελάχιστη σύσταση για παιδιά των συγκεκριμένων ηλικιών στη βιβλιογραφία [86], [87]. Τα κατώφλια για τις ώρες ύπνου ορίστηκαν στις 7 και 9 ώρες, όπως και οι πιθανές απαντήσεις της αντίστοιχης ερώτησης στο σχετικό ερωτηματολόγιο της μελέτης (βλ. Πίνακας 4.9). Παρόλο που η σύσταση ήταν 9-11 ώρες ύπνου [11], οι περιπτώσεις περισσότερων από 11 ώρες ύπνου ταξινομήθηκαν και αυτές στην κλάση του συστηνόμενου ύπνου, καθώς υπάρχουν στοιχεία ότι ο ύπνος για περισσότερες από 11 ώρες συνδέεται με μειωμένα ποσοστά παχυσαρκίας σε παιδιά [88].

Η ενέργεια που λήφθηκε σε κάθε χρονική στιγμή είναι ένα διάνυσμα μήκους 11 το οποίο δείχνει το πλήθος των μηνυμάτων κάθε κατηγορίας που στάλθηκε. Πχ. αν στάλθηκαν 2 μηνύματα από την κατηγορία 10, 1 μήνυμα από τις κατηγορίες 7 και 8 και κανένα μήνυμα από τις υπόλοιπες κατηγορίες, τότε η ενέργεια κωδικοποιείται με το διάνυσμα: [0, 0, 0, 0, 0, 0, 1, 1, 0, 2, 0].

Αν ληφθούν υπόψιν, για έναν χρήστη σε μια συγκεκριμένη μέρα ο μέγιστος αριθμός μηνυμάτων που στάλθηκε από κάθε κατηγορία (2 για τις κατηγορίες 1-3, 1 για τις κατηγορίες 4-7 και 6 για τις κατηγορίες 8-11), ο μέγιστος αριθμός συνολικών μηνυμάτων (8 μηνύματα) και ο μέγιστος αριθμός διαφορετικών κατηγοριών (5 κατηγορίες), προκύπτει ένας χώρος 17,333 πιθανών ενεργειών εκ των οποίων οι 274 εμφανίζονται στο σύνολο δεδομένων.

Δεδομένου ότι οι χρήστες λάμβαναν τα μηνύματα συνήθως πρωινές ώρες και πριν προλάβουν να προχωρήσουν στις καταγραφές, θεωρήθηκε η ενέργεια που ακολουθεί την κατάσταση μιας μέρας να αποτελείται από τα μηνύματα που στάλθηκαν την επόμενη. Με άλλα λόγια, μία σειρά δεδομένων περιέχει μία ενέργεια και την κατάσταση που την ακολούθησε.

Όσον αφορά τη συνάρτηση ανταμοιβής (reward function), αποφασίστηκε να εξαρτάται αποκλειστικά από το διάνυσμα κατάστασης – το να εξαρτάται και από τη ληφθείσα ενέργεια, δηλαδή τα μηνύματα που στάλθηκαν, δεν θα είχε και πολύ νόημα για ένα σύστημα που

ανταμείβει την αυξημένη συμμόρφωση και τη συμπεριφορά σύμφωνα με τις προτεινόμενες οδηγίες.

Όπως αναφέρθηκε, οι μεταβλητές κατάστασης χωρίστηκαν σε κατηγορίες ανάλογα με το αν αφορούσαν τη συμμόρφωση ή την έκβαση της θεραπείας. Το ποσοστό που κάθε κατηγορία συνεισφέρει στην τελική συνάρτηση ανταμοιβής αποφασίστηκε να είναι μεταβλητό και να καθορίζεται από μία υπερπαράμετρο. Επειδή ένα από τα ζητούμενα της μελέτης του ENDORSE ήταν η διερεύνηση της συμμόρφωσης [11], αλλά και επειδή η σωστή επιλογή μηνυμάτων μπορεί να επηρεάσει την ενασχόληση του χρήστη με την πλατφόρμα, αποφασίστηκε να δοθεί μεγαλύτερη βαρύτητα στις μεταβλητές που περιγράφουν τη συμμόρφωση.

Μια περαιτέρω κατηγοριοποίηση των μεταβλητών είναι το αν τα δεδομένα προέρχονται από την εφαρμογή ή από το Fitbit. Αποφασίστηκε η εφαρμογή και το Fitbit να έχουν ίση συνεισφορά τόσο στην ανταμοιβή για τη συμμόρφωση όσο και σε αυτήν για την έκβαση.

Τέλος, για λόγους απλότητας και ερμηνευσιμότητας, η ημερήσια ανταμοιβή αποφασίστηκε να κυμαίνεται μεταξύ 0 και 1, με τις υψηλότερες ανταμοιβές να είναι προτιμότερες.

Ακολουθεί η συνάρτηση υπολογισμού της ανταμοιβής για έναν χρήστη για μία συγκεκριμένη κατάσταση s :

$$r(s) = adh \cdot (0.25 \cdot \text{daily_monitoring} + 0.25 \cdot \text{weight_monitoring} + 0.25 \cdot \text{fitbit_hours}) \\ + (1 - adh) \cdot (0.25 \cdot \text{daily_target} + 0.25 \cdot \text{weight_difference} + 0.25 \\ \cdot \text{step_count} + 0.125 \cdot \text{sleep_hours})$$

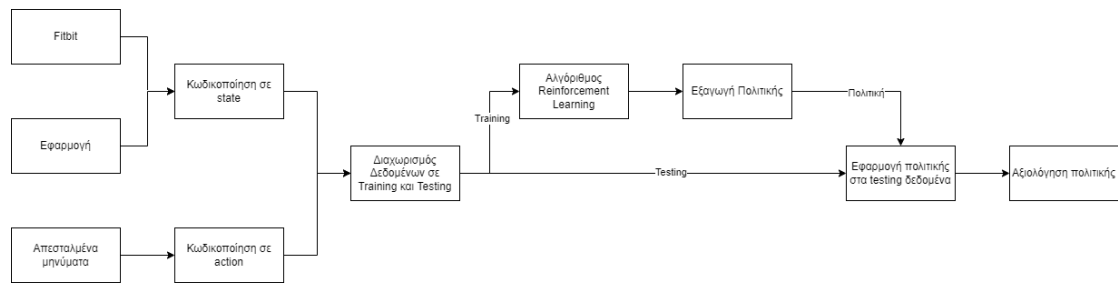
Εξίσωση 5.1: Συνάρτηση υπολογισμού της συνάρτησης ανταμοιβής. Η τιμή της ανταμοιβής εξαρτάται από τις τιμές των μεταβλητών κατάστασης και της υπερπαραμέτρου βαρύτητας της συμμόρφωσης adh .

Όπου adh η υπερπαράμετρος που καθορίζει το ποσοστό συνεισφοράς των μεταβλητών που αφορούν τη συμμόρφωση στη συνολική ανταμοιβή. Εκπαιδεύτηκαν μοντέλα για $adh = 0.7, 0.8, 1$, δηλαδή βαρύτητα της συμμόρφωσης 70%, 80% και 100%, αντίστοιχα. Οι υπόλοιπες μεταβλητές επεξηγούνται στον Πίνακα 5.1. Να σημειωθεί ότι οι συντελεστές των τριαδικών μεταβλητών $fitbit_hours$ και $sleep_hours$ είναι διαιρεμένοι δια 2 σε σχέση με αυτό που θα ήταν αν ήταν δυαδικές, καθώς μπορούν να πάρουν και την τιμή 2 σε αντίθεση με τις υπόλοιπες μεταβλητές. Δηλαδή, η $fitbit_hours$ μπορεί να συνεισφέρει

ει έως και 50% στην ανταμοιβή σχετιζόμενη με τη συμμόρφωση, διότι είναι η μοναδική μεταβλητή που δείχνει τη συμμόρφωση με το Fitbit. Από την άλλη η μεταβλητή $sleep_hours$ μπορεί να συνεισφέρει έως και 25% στην ανταμοιβή σχετιζόμενη με την έκβαση της θεραπείας, όσο και η $step_count$, διότι αυτές οι δύο μαζί συνιστούν τις μετρήσεις του Fitbit που δίνουν στοιχεία για την έκβαση.

5.1.4. Περιγραφή της διαδικασίας εκπαίδευσης

Στο παρακάτω διάγραμμα, περιγράφεται συνοπτικά η διαδικασία εκπαίδευσης των παραγόντων και αξιολόγησης των πολιτικών που έμαθαν.



Σχήμα 5.2: Αρχιτεκτονική της διαδικασίας εκπαίδευσης των μοντέλων συστάσεων αποστολής μηνυμάτων.

Τα δεδομένα που καθορίζουν την κατάσταση προέρχονται από την εφαρμογή και το Fitbit. Αυτά υπέστησαν ένα στάδιο προ-επεξεργασίας, ούτως ώστε να κωδικοποιηθούν και να έρθουν σε μορφή κατάλληλη για να περιγράψουν το πρόβλημα ενισχυτικής μάθησης. Τα δεδομένα που αφορούσαν μέρες στις οποίες δεν στάλθηκαν μηνύματα αφαιρέθηκαν από το σύνολο δεδομένων. Η ενέργεια σε κάθε χρονική στιγμή καθορίζεται από τα μηνύματα που στάλθηκαν. Ομοίως, έπρεπε τα μηνύματα να υποστούν επεξεργασία και να μετατραπούν σε διανύσματα που περιγράφουν ενέργειες. Τα παραπάνω, συνιστούν το σύνολο δεδομένων το οποίο περιέχει 2,321 σειρές.

Στη συνέχεια, πραγματοποιείται διαχωρισμός των δεδομένων σε δεδομένα εκπαίδευσης και δεδομένα ελέγχου. Ο διαχωρισμός ακολουθεί τη μέθοδο της διασταυρωμένης επικύρωσης και συγκεκριμένα *leave-one-out cross validation*, αλλά σε επίπεδο χρηστών. Δηλαδή, πραγματοποιούνται 35 επαναλήψεις της διαδικασίας, σε κάθε μία από τις οποίες οι γραμμές δεδομένων που αφορούν έναν χρήστη αποτελούν τα δεδομένα ελέγχου και οι υπόλοιπες που αφορούν τους άλλους 34 χρήστες τα δεδομένα εκπαίδευσης.

Στα δεδομένα εκπαίδευσης, εφαρμόζεται ο αλγόριθμος ενισχυτικής μάθησης, συγκεκριμένα Q-learning, από τον οποίο προκύπτει ο πίνακας με τις τιμές της συνάρτησης αξίας $q(s, a)$, για κάθε κατάσταση s και ενέργεια a . Ο πίνακας των τιμών της q καθορίζει την πολιτική που μαθαίνει ο παράγοντας από τα δεδομένα.

Η μαθημένη πολιτική εφαρμόζεται στα δεδομένα ελέγχου ως εξής. Σε κάθε κατάσταση s η οποία συναντάται στα δεδομένα, αντιστοιχίζεται η ενέργεια a η οποία μεγιστοποιεί τη συνάρτηση $q(s, a)$ με σταθερό s . Αυτό σημαίνει ότι αν ο παράγοντας βρισκόταν στην κατάσταση s , τότε θα επέλεγε την ενέργεια που έχει μάθει ότι θα μεγιστοποιήσει την μακροπρόθεσμη ανταμοιβή που θα λάβει.

Εξαιτίας του γεγονότος ότι δεν ήταν δυνατή η εφαρμογή των εξαγόμενων πολιτικών στο πραγματικό σύστημα, η αξιολόγηση του μοντέλου πραγματοποιήθηκε μόνο με χρήση των διαθέσιμων δεδομένων ελέγχου. Συγκεκριμένα, ακολουθήθηκε μεθοδολογία αντίστοιχη με το δημοσίευμα των Nemati et al [89], στο οποίο για την offline αξιολόγηση της μαθημένης πολιτικής μελετάται η σχέση της απόστασης των πραγματικών ενεργειών από τις προτεινόμενες (αυτές που συμφωνούν με τη μαθημένη πολιτική) με την ανταμοιβή που ακολούθησε τις πραγματικές ενέργειες. Αν όσο αυξάνεται η απόσταση της πραγματικής ενέργειας από την προτεινόμενη, μειώνεται η ανταμοιβή, τότε σημαίνει ότι η μαθημένη πολιτική είναι καλή, διότι οι ενέργειες που συμφωνούν με την πολιτική δίνουν μεγαλύτερη ανταμοιβή από αυτές που διαφωνούν.

Στην παρούσα εργασία, πραγματοποιείται αξιολόγηση της πολιτικής σε επίπεδο σειρών δεδομένων και σε επίπεδο χρηστών. Οι σειρές δεδομένων χωρίζονται σε κάδους (buckets) ανάλογα με την απόσταση της πραγματικής ενέργειας από την προτεινόμενη. Για κάθε

μοντέλο οι σειρές δεδομένων χωρίζονται σε 2 έως 6 κάδους αυξανόμενης απόστασης πραγματικής ενέργειας από την προτεινόμενη. Στη συνέχεια, υπολογίζεται η μέση ανταμοιβή για κάθε κάδο και αν υπάρχει φθίνουσα σχέση μεταξύ της απόστασης που αντιστοιχεί σε κάθε κάδο και της μέσης ανταμοιβής για τον κάδο, τότε σημαίνει ότι όσο περισσότερο συμφωνούν οι ενέργειες με την πολιτική του μοντέλου, τόσο μεγαλύτερη είναι η ανταμοιβή που απέδωσαν. Σε επίπεδο χρηστών, υπολογίζονται η μέση απόσταση μεταξύ πραγματικής και προτεινόμενης ενέργειας και η μέση ανταμοιβή για κάθε χρήστη (όχι η συνολική γιατί δεν αντιστοιχεί ο ίδιος αριθμός γραμμών σε κάθε χρήστη) και υπολογίζονται οι συσχετίσεις Pearson και Spearman μεταξύ των δύο αυτών ποσοτήτων. Πάλι, το ζητούμενο είναι μια αρνητική συσχέτιση, όσο το δυνατών πιο κοντά στο -1.

5.1.5. Ιεραρχική ενισχυτική μάθηση για μείωση του χώρου ενεργειών

Το Q-learning αποτελεί μία μέθοδο ενισχυτικής μάθησης για προβλήματα διακριτού χώρου καταστάσεων και ενεργειών. Αυτό σημαίνει ότι κάθε κατάσταση και ενέργεια αντιμετωπίζονται από τον αλγόριθμο ως τελείως διαφορετικές η μία από την άλλη. Για παράδειγμα, οι $[1, 1, 2, 0, 0, 0, 0, 0, 0, 0]$, $[1, 0, 2, 1, 0, 0, 0, 0, 0, 0]$ και $[0, 0, 0, 0, 0, 0, 0, 0, 2, 2]$ αντιμετωπίζονται όλες ως τελείως διαφορετικές ενέργειες, παρόλο που οι δύο πρώτες περιλαμβάνουν την αποστολή μηνυμάτων από κοινές κατηγορίες.

Η προφανής προσέγγιση είναι η εκπαίδευση ενός μοντέλου το οποίο αντιμετωπίζει κάθε κατάσταση και ενέργεια ως ξεχωριστή. Σε αυτήν την προσέγγιση, αφαιρέθηκαν από το σύνολο δεδομένων 168 σειρές, οι οποίες αφορούσαν μέρες στις οποίες στάλθηκαν λιγότερα από 4 μηνύματα. Οι ενέργειες με λιγότερα από 4 μηνύματα έλαβαν χώρα ορισμένες μέρες για κάποιους χρήστες στην προπilotική μελέτη, χωρίς να υπάρχουν πολλά δείγματα από αυτές. Ο μικρός αριθμός δειγμάτων ενδέχεται να μειώσει την αξιοπιστία των q -τιμών για τις εν λόγω ενέργειες, ενώ και το πρόγραμμα ENDORSE είναι σχεδιασμένο να στέλνει τουλάχιστον 4 μηνύματα ανά ημέρα, επομένως θεωρήθηκε καλύτερο να μην συμπεριληφθούν αυτές οι ενέργειες στις πολιτικές που προτείνονται από τα μοντέλα. Εν τέλει, το μοντέλο εκπαιδεύεται για τις 50 καταστάσεις και τις 215 ενέργειες που υπάρχουν στο ανανεωμένο σύνολο δεδομένων.

Ωστόσο, πέραν της παραπάνω προσέγγισης, ακολουθήθηκε και μία ακόμα η οποία αξιοποιεί την ιδέα της ιεραρχικής ενισχυτικής μάθησης. Εξαιτίας της διακύμανσης στον αριθμό απεσταλμένων μηνυμάτων κάθε μέρα στο σύνολο δεδομένων, προέκυψε το ζήτημα ότι ο αλγόριθμος στη μορφή που ήταν δεν θα μπορούσε να χρησιμοποιήσει δεδομένα ημερών στις οποίες στάλθηκε ένας συγκεκριμένος αριθμός μηνυμάτων για να καθορίσει την ενέργεια για κάποια μέρα στην οποία στάλθηκε διαφορετικός αριθμός μηνυμάτων. Για παράδειγμα, τα δεδομένα ενός χρήστη της προπilotικής μελέτης ο οποίος λάμβανε 6 μηνύματα σε μία μέρα δεν θα μπορούσαν να αξιοποιηθούν για να μάθει ο παράγοντας να διαχειρίζεται τις περιπτώσεις της πιλοτικής μελέτης στην οποία στέλνονταν 4 μηνύματα ανά ημέρα σε κάθε χρήστη. Προκειμένου να μειωθεί περαιτέρω ο χώρος ενεργειών, αλλά και να αυξηθεί η συχνότητα εμφάνισης κάθε ενέργειας στο σύνολο δεδομένων, αποφασίστηκε κάθε μεταβλητή των διανυσμάτων ενέργειας να είναι δυαδική και να αφορά την αποστολή ή όχι μηνύματα της αντίστοιχης κατηγορίας. Για παράδειγμα, οι ενέργειες $[2, 2, 0, 0, 0, 0, 0, 0, 0, 0]$ και $[1, 1, 0, 0, 0, 0, 0, 0, 0, 0]$, θα αναπαρίστανται και οι δύο με το διάνυσμα $[1, 1, 0, 0, 0, 0, 0, 0, 0, 0]$.

Με τον παραπάνω τρόπο, το σύνολο δεδομένων περιλαμβάνει 144 ενέργειες από έναν χώρο 1,023 πιθανών ενεργειών, ο οποίος προκύπτει με τον περιορισμό ότι σε μία ενέργεια δεν μπορούν να σταλούν μηνύματα από περισσότερες από 5 διαφορετικές κατηγορίες.

Με τη μείωση του χώρου ενεργειών επιτυγχάνεται η αύξηση της πυκνότητας δειγμάτων για κάθε ενέργεια, αλλά και η αξιοποίηση των γνώσεων από την επίδραση των ενεργειών οποιουδήποτε συνολικού αριθμού μηνυμάτων για τη μάθηση πολιτικής αποστολής συγκεκριμένου ημερησίου συνολικού αριθμού μηνυμάτων. Πχ. τα δείγματα που περιλαμβάνουν ενέργειες στις οποίες στάλθηκαν 3 ή 6 συνολικά μηνύματα σε μία μέρα μπορούν να αξιοποιηθούν για την εκπαίδευση ενός συστήματος που στέλνει πάντα 4 μηνύματα ανά ημέρα σε κάθε χρήστη, όπως συνέβαινε και στην πιλοτική μελέτη. Με αυτό το σκεπτικό, οι σειρές δεδομένων που αφορούν ημέρες στις οποίες στάλθηκαν λιγότερα από 4 μηνύματα σε έναν χρήστη συμπεριλήφθηκαν κανονικά στο σύνολο δεδομένων στην ιεραρχική προσέγγιση.

Η ιεραρχική ενισχυτική μάθηση (hierarchical reinforcement learning) είναι μία προσέγγιση σε προβλήματα ενισχυτικής μάθησης η οποία εμπεριέχει την εκπαίδευση και τον σχεδιασμό πολιτικών σε πολλαπλά αφαιρετικά επίπεδα. Μία βασική ιδέα είναι ο διαχωρισμός των διαδικασιών του προβλήματος σε μικρότερες υποδιαδικασίες, για κάθε μία από τις οποίες ο παράγοντας μπορεί να εκπαιδευτεί ευκολότερα σε σχέση με το συνολικό πρόβλημα. Η εκπαίδευση στις υποδιαδικασίες που πραγματοποιείται σε ένα επίπεδο ιεραρχικής ενισχυτικής μάθησης αξιοποιείται από το αμέσως υψηλότερο το οποίο συνδυάζει τις συσσωρευμένες γνώσεις και δεξιότητες για τον καθορισμό μιας υψηλότερου επιπέδου πολιτικής [90].

Στην παρούσα εργασία, υλοποιήθηκαν δύο επίπεδα ιεραρχίας. Στο χαμηλότερο επίπεδο της ιεραρχίας, η κάθε ενέργεια σπάει σε 11 ενέργειες, η κάθε μία από τις οποίες αφορά την αποστολή ή όχι μηνύματος από μια κατηγορία. 11 παράγοντες εκπαιδεύονται, δηλαδή, σε ολόκληρο το σύνολο δεδομένων, με τον καθένα να μπορεί να πραγματοποιήσει δύο πιθανές ενέργειες, την αποστολή ή όχι μηνύματος από την εκάστοτε κατηγορία. Με αυτόν τον τρόπο, μελετάται η επίδραση κάθε κατηγορίας μηνυμάτων ξεχωριστά και εξάγονται 11 πίνακες q -τιμών. Για μία συγκεκριμένη κατηγορία μηνυμάτων και μία κατάσταση s , η τιμή $q(s, 0)$ αποτελεί την προσδοκώμενη συνολική μελλοντική ανταμοιβή μετά την κατάσταση s , αν μετά την s δεν αποσταλεί μήνυμα από την εν λόγω κατηγορία και η τιμή $q(s, 1)$, την προσδοκώμενη συνολική μελλοντική ανταμοιβή αν αποσταλεί.

Στο υψηλότερο ιεραρχικό επίπεδο, επιλέγεται το διάνυσμα της προτεινόμενης ενέργειας με βάση τις q -τιμές των 11 παραγόντων. Εάν μία από τις χαμηλότερου επιπέδου πολιτικές υπαγορεύει στην τρέχουσα κατάσταση την αποστολή μηνύματος της αντίστοιχης κατηγορίας τότε η κατάλληλη θέση στο διάνυσμα της προτεινόμενης ενέργειας λαμβάνει την τιμή 1, αλλιώς λαμβάνει την τιμή 0. Ενέργειες με τις οποίες θα αποστέλλονταν μηνύματα πολλών κατηγοριών θα ήταν καλό να αποφευχθούν, διότι δεν θα βοηθούσαν στην υπενθύμιση και την εστίαση της προσοχής των χρηστών σε επιλεγμένες εκφάνσεις της θεραπείας κάθε φορά, που είναι και ο στόχος ενός έξυπνου συστήματος εξατομικευμένων συστάσεων. Ως εκ τούτου, αποφασίστηκε περιορισμός στην αποστολή μηνυμάτων από μέχρι και 4 κατηγορίες σε έναν χρήστη ανά ημέρα (το όριο δεν τέθηκε στο 5, όπως και ο μέγιστος αριθμός κατηγοριών στο σύνολο δεδομένων, για να έχει εφαρμογή και σε συστήματα στα οποία αποστέλλονταν 4 μηνύματα ανά ημέρα, όπως στην πιλοτική μελέτη). Στην περίπτωση στην οποία το ιεραρχικό σχήμα υπαγόρευε την αποστολή μηνυμάτων από περισσότερες από 4 κατηγορίες σε μία κατάσταση s , επιλέγονταν οι 4 με τις υψηλότερες τιμές $q(s, 1)$.

Κάποιοι λόγοι για την υιοθέτηση της ιεραρχικής προσέγγισης είναι οι εξής:

- **Ξεχωριστή επίδραση κάθε κατηγορίας μηνυμάτων:** Λαμβάνεται υπόψιν από τον αλγόριθμο η επίδραση στη συμμόρφωση και στην έκβαση της θεραπείας κάθε κατηγορίας μηνυμάτων ξεχωριστά.
- **Αυξημένη ευελιξία ενεργειών:** Επιτρέπεται η πρόταση ενεργειών όταν ο παράγοντας βρίσκεται σε μία κατάσταση, ακόμα και αν το ζεύγος ενέργειας-κατάστασης δεν συναντάται στο σύνολο δεδομένων, αλλά και ενεργειών οι οποίες δεν υπάρχουν καθόλου στο σύνολο δεδομένων. Ως αποτέλεσμα, αυξάνεται η κάλυψη του χώρου ενεργειών.
- **Αυξημένη πυκνότητα δειγμάτων:** Κάθε παράγοντας του χαμηλότερου επιπέδου εκπαιδεύεται σε ένα σύνολο 2,321 σειρών, σε κάθε μία από τις οποίες λαμβάνεται μία από δύο ενέργειες. Σε ένα τέτοιο σύνολο δεδομένων είναι πιο πιθανό ένα ζεύγος κατάστασης-ενέργειας να συναντάται περισσότερες φορές σε σύγκριση με ένα σύνολο δεδομένων με περισσότερες ενέργειες. Όσο αυξάνεται η συχνότητα των ζευγών κατάστασης-ενέργειας στο σύνολο δεδομένων ενδέχεται να αυξηθεί και η αξιοπιστία των q -τιμών.

5.1.6. Απόσταση στον χώρο ενεργειών

Ο ορισμός μίας έννοιας της απόστασης μεταξύ των ενεργειών είναι απαραίτητος για την αξιολόγηση των μαθημένων πολιτικών σύμφωνα με τη μεθοδολογία που αναφέρθηκε.

Στον αρχικό χώρο καταστάσεων, οι ενέργειες έχουν τη μορφή διανυσμάτων μήκους 11. Η απόσταση μεταξύ δύο ενεργειών, δηλαδή πόσο κοντινές ή συναφείς είναι δύο ενέργειες, αποφασίστηκε να είναι η νόρμα επιπέδου 2 του διανύσματος της διαφοράς των διανυσμάτων που τις ορίζουν. Δηλαδή, η απόσταση μεταξύ των διανυσμάτων ενέργειας $x = [x_1, x_2, \dots, x_{11}]$ και $y = [y_1, y_2, \dots, y_{11}]$ είναι $\|x - y\|_2 = \sqrt{\sum_{i=1}^{11} |x_i - y_i|^2}$, η τετραγωνική ρίζα του αθροίσματος των τετραγώνων των αντίστοιχων συνιστωσών.

Στην ιεραρχική προσέγγιση, οι συνιστώσες των διανυσμάτων είναι δυαδικές μεταβλητές. Σε αυτήν την περίπτωση, θα είχε νόημα για τον ορισμό της απόστασης μεταξύ δύο ενεργειών η χρήση της νόρμας επιπέδου 1 του διανύσματος της διαφοράς των διανυσμάτων που τις ορίζουν. Η απόσταση μεταξύ των διανυσμάτων ενέργειας x και y ορίζεται ως $\|x - y\|_1 = \sum_{i=1}^{11} |x_i - y_i|$, δηλαδή το πλήθος των συνιστωσών στις οποίες τα x και y έχουν διαφορετική τιμή, εν προκειμένω το πλήθος των κατηγοριών μηνυμάτων τα οποία θα σταλούν σύμφωνα με τη μία ενέργεια και όχι με την άλλη.

5.2. ΑΠΟΤΕΛΕΣΜΑΤΑ

Σε αυτήν την ενότητα, αναλύονται τα αποτελέσματα της εκπαίδευσης των παραγόντων ενισχυτικής μάθησης για το σύστημα συστάσεων αποστολής μηνυμάτων.

Όπως αναφέρθηκε, ακολουθήθηκαν δύο προσεγγίσεις: στην πρώτη εφαρμόστηκε Q-learning στους αρχικούς χώρους καταστάσεων και ενεργειών, ενώ στη δεύτερη οι μεταβλητές του χώρου ενεργειών μετατράπηκαν σε δυαδικές και εκπαιδεύτηκαν 11 παράγοντες με ιεραρχικό τρόπο, ένας για κάθε μεταβλητή. Για κάθε προσέγγιση, μαθεύτηκαν τρεις πολιτικές, για τρεις διαφορετικές τιμές της υπερπαραμέτρου που καθορίζει την βαρύτητα της συμμόρφωσης στη συνάρτηση ανταμοιβής: 70%, 80% και 100%.

5.2.1. Αρχικός χώρος ενεργειών

Σε αυτήν την υποενότητα, αξιολογούνται οι πολιτικές που μαθεύτηκαν με την απλή προσέγγιση στην οποία κάθε ενέργεια αντιμετωπίζεται ως ξεχωριστή.

Στον Πίνακα 5.2, φαίνονται οι συσχετίσεις Pearson και Spearman μεταξύ της μέσης απόστασης πραγματικής από την προτεινόμενη ενέργεια (νόρμα επιπέδου 2) και της μέσης ανταμοιβής που έλαβε κάθε χρήστης για τις διαφορετικές βαρύτητες της συμμόρφωσης:

Πίνακας 5.2: Αξιολόγηση των πολιτικών σε επίπεδο χρηστών για τα μοντέλα του συστήματος συστάσεων όταν ακολουθήθηκε η προσέγγιση με τον αρχικό χώρο ενεργειών. Οι τιμές συσχέτισης με έντονη γραμματοσειρά ήταν οι μέγιστες για τη συγκεκριμένη βαρύτητα συμμόρφωσης μεταξύ των διαφορετικών προσεγγίσεων.

Βαρύτητα συμμόρφωσης	Pearson		Spearman	
	Συσχέτιση	p-τιμή	Συσχέτιση	p-τιμή
70%	-0,824	$1,2 \cdot 10^{-9}$	-0,829	$7,9 \cdot 10^{-10}$
80%	-0,841	$2,5 \cdot 10^{-10}$	-0,841	$2,6 \cdot 10^{-10}$
100%	-0,828	$8,4 \cdot 10^{-10}$	-0,855	$6,3 \cdot 10^{-11}$

Φαίνεται ότι για όλες τις τιμές βαρύτητας της συμμόρφωσης που δοκιμάστηκαν, η μέση απόσταση μεταξύ της πραγματικής από την προτεινόμενη ενέργεια και της μέσης ανταμοιβής για έναν χρήστη είναι σε μεγάλο βαθμό αρνητικά συσχετισμένες. Οι p-τιμές δείχνουν ότι η πιθανότητα να υπάρχει τόσο μεγάλη συσχέτιση σε περίπτωση ισχύος της μηδενικής υπόθεσης (null hypothesis - ότι οι δύο μεταβλητές είναι ανεξάρτητες) είναι απειροελάχιστη.

Η μεγάλη αρνητική συσχέτιση σημαίνει ότι όσο αυξάνεται η μέση ανταμοιβή του χρήστη, τόσο μειώνεται η απόσταση των πραγματικών ενεργειών από τις προτεινόμενες από τις μαθημένες πολιτικές. Δηλαδή, οι κατηγορίες των πραγματικών μηνυμάτων που έλαβαν οι χρήστες με τις υψηλότερες μέσες ανταμοιβές συμφωνούσαν σε μεγαλύτερο βαθμό με τις προτεινόμενες κατηγορίες από όσο συμφωνούσαν οι ενέργειες των χρηστών με τις χαμηλότερες μέσες ανταμοιβές.

Στη συνέχεια, παρουσιάζεται η κατηγοριοποίηση των σειρών του συνόλου δεδομένων σε κάδους με βάση την απόσταση της πραγματικής ενέργειας από την προτεινόμενη και η μέση ανταμοιβή για κάθε κάδο. Η ταξινόμηση των δεδομένων πραγματοποιήθηκε 5 φορές, σε 2, 3, 4, 5 και 6 κάδους, με βάση την απόσταση της πραγματικής ενέργειας από την προτεινόμενη. Να σημειωθεί ότι από αυτήν την προσέγγιση έχουν εξαιρεθεί τα 168 δείγματα στα οποία έχουν σταλεί λιγότερα από 4 μηνύματα ανά ημέρα, ενώ από την ταξινόμηση απουσιάζουν και τα 35 δείγματα που αντιστοιχούν στην πρώτη μέρα θεραπείας κάθε χρήστη, αφού δεν υπάρχει προηγούμενη κατάσταση για αυτά. Ο συνολικός αριθμός δειγμάτων που ταξινομήθηκαν είναι 2.118. Η μέση ανταμοιβή ήταν 0.165, 0.175 και 0.195 για βαρύτητες συμμόρφωσης 70%, 80% και 100% αντίστοιχα.

Πίνακας 5.3: Συγκεντρωτικός πίνακας για την ομαδοποίηση των δειγμάτων σε κάδους για την προσέγγιση με αναπαράσταση των δεδομένων στον αρχικό χώρο ενεργειών και βαρύτητα της συμμόρφωσης στη συνάρτηση ανταμοιβής 70%. π.κ.: πλήθος κάδων, α.κ.: αριθμός κάδων, μ.α.: μέση ανταμοιβή, π.δ.: πλήθος δειγμάτων.

π.κ.	2		3		4		5		6	
	μ.α.	π.δ.	μ.α.	π.δ.	μ.α.	π.δ.	μ.α.	π.δ.	μ.α.	π.δ.
0	0,224	1170	0,267	302	0,221	187	0,291	91	0,208	50
1	0,093	948	0,171	1503	0,225	983	0,227	662	0,279	252
2			0,041	313	0,099	854	0,180	696	0,209	868
3					0,038	94	0,074	628	0,118	635
4							0,032	41	0,045	288
5									0,000	25

Πίνακας 5.4: Συγκεντρωτικός πίνακας για την ομαδοποίηση των δειγμάτων σε κάδους για την προσέγγιση με αναπαράσταση των δεδομένων στον αρχικό χώρο ενεργειών και βαρύτητα της συμμόρφωσης στη συνάρτηση ανταμοιβής 80%. π.κ.: πλήθος κάδων, α.κ.: αριθμός κάδων, μ.α.: μέση ανταμοιβή, π.δ.: πλήθος δειγμάτων.

π.κ.	2		3		4		5		6	
	μ.α.	π.δ.	μ.α.	π.δ.	μ.α.	π.δ.	μ.α.	π.δ.	μ.α.	π.δ.
0	0,236	1161	0,281	292	0,229	182	0,302	88	0,206	48
1	0,102	957	0,182	1513	0,237	979	0,238	656	0,296	244
2			0,046	313	0,109	865	0,193	703	0,221	869
3					0,039	92	0,082	632	0,129	644
4							0,029	39	0,049	288
5									0,000	25

Πίνακας 5.5: Συγκεντρωτικός πίνακας για την ομαδοποίηση των δειγμάτων σε κάδους για την προσέγγιση με αναπαράσταση των δεδομένων στον αρχικό χώρο ενεργειών και βαρύτητα της συμμόρφωσης στη συνάρτηση ανταμοιβής 100%. π.κ.: πλήθος κάδων, α.κ.: αριθμός κάδων, μ.α.: μέση ανταμοιβή, π.δ.: πλήθος δειγμάτων.

π.κ.	2		3		4		5		6	
	μ.α.	π.δ.	μ.α.	π.δ.	μ.α.	π.δ.	μ.α.	π.δ.	μ.α.	π.δ.
0	0,262	1166	0,306	279	0,259	185	0,326	79	0,260	50
1	0,113	952	0,204	1529	0,263	981	0,260	646	0,317	229
2			0,050	310	0,120	861	0,222	726	0,248	887
3					0,047	91	0,091	629	0,144	642
4							0,033	38	0,053	284
5									0,019	26

Όπως φαίνεται από τους πίνακες 5.3, 5.4 και 5.5, η μέση ανταμοιβή για τα δείγματα ενός κάδου μειώνεται όσο αυξάνεται ο αριθμός του κάδου, άρα και η απόσταση της πραγματικής ενέργειας από την προτεινόμενη στα δείγματα που περιέχει. Αυτό σημαίνει ότι ενέργειες πιο κοντινές στις προτεινόμενες από τις πολιτικές προηγήθηκαν των καταστάσεων με μεγαλύτερες ανταμοιβές. Εξαίρεση στη φθίνουσα συμπεριφορά της μέσης ανταμοιβής συναρτήσει του αριθμού του κάδου, αποτελούν ορισμένες περιπτώσεις με μικρό αριθμό δειγμάτων, πχ. ο κάδος 0 για την περίπτωση του διαχωρισμού σε 6 κάδους, αλλά η γενική τάση της ανταμοιβής όσο αυξάνεται η απόσταση είναι πτωτική.

5.2.2. Ιεραρχική προσέγγιση

Σε αυτήν την ενότητα, αξιολογούνται οι πολιτικές που μαθεύτηκαν όταν ακολουθήθηκε η ιεραρχική προσέγγιση. Σε αυτήν την προσέγγιση, ο χώρος ενεργειών είχε μειωθεί, ώστε να λαμβάνεται υπόψιν μόνο αν στάλθηκε μήνυμα από κάθε κατηγορία και όχι το πλήθος τους, και εκπαιδεύτηκαν 11 παράγοντες, ένας για κάθε κατηγορία, με τις πολιτικές που έμαθαν να συνδυάζονται για να προκύψει η τελική προτεινόμενη πολιτική.

Στον Πίνακα 5.6 φαίνονται οι συσχετίσεις Pearson και Spearman μεταξύ της μέσης απόστασης πραγματικής από την προτεινόμενη ενέργεια (νόρμα επιπέδου 1) και της μέσης ανταμοιβής που έλαβε κάθε χρήστης για τις διαφορετικές βαρύτητες της συμμόρφωσης:

Πίνακας 5.6: Αξιολόγηση των πολιτικών σε επίπεδο χρηστών για τα μοντέλα του συστήματος συστάσεων όταν ακολουθήθηκε η ιεραρχική προσέγγιση. Οι τιμές συσχέτισης με έντονη γραμματοσειρά ήταν οι μέγιστες για τη συγκεκριμένη βαρύτητα συμμόρφωσης μεταξύ των διαφορετικών προσεγγίσεων.

Βαρύτητα συμμόρφωσης	Pearson		Spearman	
	Συσχέτιση	p-τιμή	Συσχέτιση	p-τιμή
70%	-0,860	$3.8 \cdot 10^{-11}$	-0,864	$2,4 \cdot 10^{-11}$
80%	-0,863	$2.6 \cdot 10^{-11}$	-0,862	$2,9 \cdot 10^{-11}$
100%	-0,771	$6.1 \cdot 10^{-8}$	-0,819	$1,8 \cdot 10^{-9}$

Όπως και στην περίπτωση της εκπαίδευσης στον αρχικό χώρο ενεργειών, παρατηρείται ισχυρή αρνητική συσχέτιση μεταξύ της μέσης απόστασης πραγματικής από την προτεινόμενη ενέργεια και τις μέσης ανταμοιβής για κάθε χρήστη, καθώς και αμελητέα πιθανότητα να σημειωθούν αυτές οι τιμές συσχέτισης υπό τη μηδενική υπόθεση. Σε σύγκριση με την προηγούμενη προσέγγιση, ο βαθμός της αρνητικής συσχέτισης είναι μεγαλύτερος όταν η συμμόρφωση συνεισφέρει στη συνάρτηση ανταμοιβής 70% και 80% και μικρότερος όταν η συνάρτηση ανταμοιβής εξαρτάται αποκλειστικά από τις μεταβλητές της συμμόρφωσης.

Ακολουθεί η ταξινόμηση των δειγμάτων σε κάδους, ανάλογα με την απόσταση της πραγματικής ενέργειας από την προτεινόμενη. Ομοίως με την προηγούμενη προσέγγιση, δοκιμάστηκαν πλήθη κάδων 2, 3, 4, 5 και 6. Η ταξινόμηση πραγματοποιήθηκε σε ολόκληρο το διαθέσιμο σύνολο δειγμάτων, πλην των 35 δειγμάτων που αντιστοιχούσαν στην πρώτη μέρα θεραπείας κάθε χρήστη. Άρα, ταξινομήθηκαν 2,286 δείγματα. Η μέση ανταμοιβή ήταν 0.182, 0.193 και 0.215 για βαρύτητες συμμόρφωσης 70%, 80% και 100% αντίστοιχα.

Πίνακας 5.7: Συγκεντρωτικός πίνακας για την ομαδοποίηση των δειγμάτων σε κάδους για την ιεραρχική προσέγγιση και βαρύτητα της συμμόρφωσης στη συνάρτηση ανταμοιβής 70%. π.κ.: πλήθος κάδων, α.κ.: αριθμός κάδων, μ.α.: μέση ανταμοιβή, π.δ.: πλήθος δειγμάτων.

π.κ.	2		3		4		5		6	
	μ.α.	π.δ.	μ.α.	π.δ.	μ.α.	π.δ.	μ.α.	π.δ.	μ.α.	π.δ.
0	0,298	927	0,411	221	0,411	221	0,435	72	0,435	72
1	0,102	1359	0,235	1071	0,262	706	0,356	383	0,399	149
2			0,073	994	0,113	1071	0,229	472	0,262	706
3					0,061	288	0,113	1071	0,181	365
4							0,061	288	0,078	706
5									0,061	288

Πίνακας 5.8: Συγκεντρωτικός πίνακας για την ομαδοποίηση των δειγμάτων σε κάδους για την ιεραρχική προσέγγιση και βαρύτητα της συμμόρφωσης στη συνάρτηση ανταμοιβής 80%. π.κ.: πλήθος κάδων, α.κ.: αριθμός κάδων, μ.α.: μέση ανταμοιβή, π.δ.: πλήθος δειγμάτων.

π.κ.	2		3		4		5		6	
	μ.α.	π.δ.	μ.α.	π.δ.	μ.α.	π.δ.	μ.α.	π.δ.	μ.α.	π.δ.
0	0,315	908	0,436	218	0,436	218	0,445	71	0,445	71
1	0,112	1378	0,248	1065	0,277	690	0,380	376	0,431	147
2			0,080	1003	0,124	1090	0,243	461	0,277	690
3					0,064	288	0,124	1090	0,195	375
4							0,064	288	0,087	715
5									0,064	288

Πίνακας 5.9: Συγκεντρωτικός πίνακας για την ομαδοποίηση των δειγμάτων σε κάδους για την ιεραρχική προσέγγιση και βαρύτητα της συμμόρφωσης στη συνάρτηση ανταμοιβής 100%. π.κ.: πλήθος κάδων, α.κ.: αριθμός κάδων, μ.α.: μέση ανταμοιβή, π.δ.: πλήθος δειγμάτων.

π.κ.	2		3		4		5		6	
α.κ.	μ.α.	π.δ.	μ.α.	π.δ.	μ.α.	α.κ.	μ.α.	π.δ.	μ.α.	π.δ.
0	0,341	931	0,458	236	0,458	236	0,442	77	0,442	77
1	0,128	1355	0,275	1073	0,301	695	0,402	417	0,465	159
2			0,090	977	0,146	1070	0,264	437	0,301	695
3					0,062	285	0,146	1070	0,227	378
4							0,062	285	0,102	692
5									0,062	285

Ομοίως με την αρχική προσέγγιση, και στην ιεραρχική, η μέση ανταμοιβή ενός κάδου μειώνεται όσο μεγαλώνει η απόσταση στην οποία αντιστοιχεί. Μάλιστα, εξαίρεση σε αυτή τη συμπεριφορά συναντάται μόνο στην περίπτωση των 6 κάδων για βαρύτητα συμμόρφωσης 100% μεταξύ των κάδων 0 και 1. Από τα αποτελέσματα, φαίνεται ότι οι πραγματικές ενέργειες που είναι περισσότερο σύμφωνες με τις προτεινόμενες από τις πολιτικές ενέργειες ακολουθήθηκαν από καταστάσεις με μεγαλύτερη ανταμοιβή.

6. ΣΥΜΠΕΡΑΣΜΑΤΑ – ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ

Διερευνήθηκαν τεχνικές μηχανικής μάθησης για την πρόβλεψη της συμμόρφωσης σε ένα πλάνο παρέμβασης για την παιδική παχυσαρκία, το οποίο βασίζεται σε τεχνολογίες κινητής υγείας, καθώς και για την παραγωγή εξατομικευμένων συστάσεων για τους χρήστες του. Τα δεδομένα προήλθαν από την πλατφόρμα ENDORSE [10]–[13].

Βρέθηκε ότι η συμμόρφωση με το πλάνο παρέμβασης μπορεί να προβλεφθεί μέσω μεθόδων επιβλεπόμενης (supervised) μηχανικής μάθησης βασισμένων σε δέντρα αποφάσεων (tree-based). Ακολουθήθηκαν μέθοδοι για την εξαγωγή των κατάλληλων χαρακτηριστικών από ένα μεγάλο σύνολο με στόχο την καλύτερη προσαρμογή των μοντέλων. Φάνηκε ότι τα χαρακτηριστικά που αποτελούν ενδείξεις της συμμόρφωσης ενδέχεται να διαφέρουν για διαφορετικές εκφάνσεις της συμμόρφωσης, αλλά και ότι τα διαφορετικά μοντέλα δεν αναδεικνύουν τα ίδια χαρακτηριστικά για το ίδιο ζητούμενο.

Η διαδικασία της πρόβλεψης της συμμόρφωσης θα μπορούσε να επεκταθεί μελλοντικά με ορισμένους τρόπους.

Αρχικά, προκειμένου να αξιολογηθεί ορθότερα η προβλεπτική ικανότητα των μοντέλων, θα ήταν χρήσιμος ο έλεγχός τους με ένα εξωτερικό σύνολο δεδομένων, το οποίο θα μπορούσε να προέρχεται, πχ., από μελλοντικές θεραπείες με βάση το πλάνο παρέμβασης ENDORSE [63]. Τα δεδομένα επικύρωσης δεν συμμετέχουν στη διαδικασία εκπαίδευσης του μοντέλου, αλλά λαμβάνονται υπόψιν στη διαδικασία επιλογής των χαρακτηριστικών.

Ακόμα, γνώσεις ειδικών στα θέματα της παιδικής παχυσαρκίας και της συμπεριφοράς θα μπορούσαν να αξιοποιηθούν για την εκ των προτέρων μείωση των χαρακτηριστικών του συνόλου δεδομένων και ως εκ τούτου μετρίαση της επίδρασης της κατάρας της διαστατικότητας. Με τη διατήρηση στο σύνολο δεδομένων των χαρακτηριστικών τα οποία είναι γνωστό ότι σχετίζονται με τη συμμόρφωση θα μπορούσε να μειωθεί ο χρόνος εκπαίδευσης των μοντέλων, καθιστώντας εφικτές περισσότερες μεθόδους επιλογής χαρακτηριστικών και συντονισμού υπερπαραμέτρων, αλλά και ο περιορισμός της υπερπροσαρμογής, διότι η επιλογή των χαρακτηριστικών θα βασίζεται σε a priori γνώσεις και όχι στα ίδια τα δεδομένα.

Όπως στα περισσότερα προβλήματα μηχανικής μάθησης, η υπερπροσαρμογή των μοντέλων πρόβλεψης της συμμόρφωσης θα μπορούσε να περιοριστεί αν ήταν διαθέσιμα περισσότερα δείγματα [63]. Η υπερπροσαρμογή συμβαίνει επειδή ο αριθμός των δεδομένων δεν επαρκεί για να αποκτήσει το μοντέλο μία πλήρη εικόνα του συνολικού προβλήματος και μαθαίνει μόνο με βάση τα διαθέσιμα δείγματα. Επομένως, ενδεχόμενες μελλοντικές εφαρμογές του πλάνου παρέμβασης ENDORSE είναι δυνατό να παράσχουν δεδομένα πολύτιμα για βελτιωμένη μοντελοποίηση του προβλήματος της συμμόρφωσης.

Επίσης, χρησιμοποιήθηκε ενισχυτική μάθηση για την κατά δύναμη αντιμετώπιση του προβλήματος της συμμόρφωσης, μέσω του συστήματος συστάσεων μηνυμάτων. Δείχθηκε ότι η χρήση ενός εκ των απλούστερων αλγορίθμων offline ενισχυτικής μάθησης, του Q-learning, δύναται να χρησιμοποιηθεί για τη μάθηση πολιτικών αποστολής μηνυμάτων, τα οποία ενδέχεται να αυξήσουν τη συμμόρφωση, αλλά και να βελτιώσουν την έκβαση της θεραπείας.

Όπως για την περίπτωση των μοντέλων πρόβλεψης της συμμόρφωσης, το σύστημα συστάσεων αποστολής μηνυμάτων θα μπορούσε να βελτιωθεί σημαντικά αν εφαρμοστεί ενισχυτική μάθηση σε περισσότερα δεδομένα. Μάλιστα, στο συγκεκριμένο πεδίο εφαρμογής, είναι εφικτή και η online εφαρμογή του αλγορίθμου, στην οποία η προτεινόμενη πολιτική εφαρμόζεται στο πραγματικό περιβάλλον και αξιοποιούνται τα δεδομένα ανάδρασης για την άμεση βελτίωσή της.

Άλλες πιθανές βελτιώσεις του συστήματος περιλαμβάνουν τη χρήση κάποιου πολυπλοκότερου αλγορίθμου για την εκπαίδευση του παράγοντα. Το διπλό Q-learning (double Q-learning) αποτελεί μία πολύ απλή βελτίωση η οποία αντιμετωπίζει την μεροληψία που υπάρχει στο Q-learning λόγω της χρήση του ίδιου του πίνακα q -τιμών για την ενημέρωση των q -τιμών. Στο διπλό Q-learning, χρησιμοποιούνται εναλλάξ δύο πίνακες q -τιμών: ένας για την q -τιμή που ενημερώνεται και ένας για τον όρο $\max_{a \in A} q(s', a)$ και το αντίστροφο (βλ. Εξίσωση 3.3). Αποδεικνύεται ότι το διπλό Q-learning μπορεί να επιτύχει ταχύτερη σύγκλιση και καλύτερα αποτελέσματα σε σχέση με το απλό Q-learning [76].

Η κυριότερη οικογένεια μεθόδων με την οποία αντιμετωπίζεται η κατάρα της διαστατικότητας που επιφέρουν οι μεγάλοι χώροι καταστάσεων και ενεργειών είναι η χρήση τεχνικών προσέγγισης συνάρτησης (function approximation) [76]. Μία πολύ δημοφιλής μέθοδος προσέγγισης συνάρτησης είναι το Deep Q-Network, στο οποίο χρησιμοποιείται ένα τεχνητό νευρωνικό δίκτυο (artificial neural network) με είσοδο τις μεταβλητές κατάστασης και έξοδο μία q -τιμή για κάθε δυνατή ενέργεια [73]. Η συνάρτηση που προσεγγίζεται από το νευρωνικό δίκτυο είναι για κάθε ενέργεια a , η q -τιμή της a συναρτήσεως των μεταβλητών της κατάστασης.

Επιπρόσθετα, θα μπορούσε να παρακαμφθεί το στάδιο της διακριτοποίησης των δεδομένων και να αξιοποιηθούν μέθοδοι ενισχυτικής μάθησης σε συνεχείς χώρους κατάστασης και ενέργειας, στους οποίους το πραγματικό περιβάλλον αναπαρίσταται με μεγαλύτερη λεπτομέρεια [73], [76].

Οι πολυπλοκότεροι αλγόριθμοι ενισχυτικής μάθησης μπορούν να λάβουν υπόψιν περισσότερη πληροφορία, να μάθουν από ακριβέστερες αναπαραστάσεις του προβλήματος και ενδεχομένως να δώσουν καλύτερα αποτελέσματα. Από την άλλη, όσο πολυπλοκότερος είναι ένας αλγόριθμος, τόσο δυσχεραίνει η διαδικασία μάθησης της πολιτικής, αλλά και υποβαθμίζεται η ερμηνευσιμότητα του συστήματος. Στη συγκεκριμένη εργασία, στην οποία διερευνάται το αν η χρήση ενισχυτικής μάθησης για τη βελτίωση τη συμμόρφωσης με ένα πλάνο παρέμβασης είναι εφικτή, επιλέχθηκε ένας εκ των απλούστερων και πλέον ερμηνεύσιμων αλγορίθμων.

Παρουσία περισσότερων δεδομένων και με χρήση πολυπλοκότερων αλγορίθμων, ίσως θα ήταν δυνατή η συμπερίληψη στον χώρο κατάστασης μεταβλητών που παρέχουν επιπλέον πληροφορίες. Τέτοιες πληροφορίες μπορεί να αφορούν την ημέρα της εβδομάδας στην οποία στάλθηκαν μηνύματα, το χρονικό σημείο μέσα στη διάρκεια του πλάνου παρέμβασης, τον εβδομαδιαίο στόχο και δεδομένα που είναι διαθέσιμα από την έναρξη, όπως δημογραφικά ή κλινικά. Επίσης, θα μπορούσαν να αξιοποιηθούν οι προβλέψεις του πρώτου σκέλους για την μάθηση ξεχωριστών πολιτικών με βάση τον προβλεπόμενο βαθμό συμμόρφωσης.

Ο ορισμός της συνάρτησης ανταμοιβής, μολονότι διερευνήθηκαν διαφορετικές βαρύτητες μεταβλητών, πραγματοποιήθηκε με μάλλον απλοϊκό τρόπο. Δεδομένου ότι η επιλογή της

κατάλληλης συνάρτησης ανταμοιβής κατά τη διατύπωση ενός προβλήματος ενισχυτικής μάθησης μπορεί να διαδραματίσει σημαίνοντα ρόλο στην πολιτική που μαθαίνεται, θα μπορούσε να δοθεί μεγαλύτερη έμφαση στον ορισμό της κατάλληλης συνάρτησης. Ένας τρόπος θα μπορούσε να περιλαμβάνει τη χρήση γνώσεων ειδικών για τον καθορισμό της βαρύτητας κάθε μεταβλητής στην ανταμοιβή κάθε κατάστασης. Διαφορετικά, θα μπορούσαν να χρησιμοποιηθούν εξεζητημένες μέθοδοι για τον ορισμό της συνάρτησης ανταμοιβής, με κόστος την αύξηση της πολυπλοκότητας της διαδικασίας και τη μείωση της ερμηνευσιμότητας [91], [92].

7. ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] H. J. Lim, H. Xue, and Y. Wang, "Global Trends in Obesity," in *Handbook of Eating and Drinking*, Cham: Springer International Publishing, 2020, pp. 1217–1235. doi: 10.1007/978-3-030-14504-0_157.
- [2] M. Simmonds, A. Llewellyn, C. G. Owen, and N. Woolacott, "Predicting adult obesity from childhood obesity: a systematic review and meta-analysis," *Obesity Reviews*, vol. 17, no. 2, pp. 95–107, Feb. 2016, doi: 10.1111/obr.12334.
- [3] A. Umer, G. A. Kelley, L. E. Cottrell, P. Giacobbi, K. E. Innes, and C. L. Lilly, "Childhood obesity and adult cardiovascular disease risk factors: a systematic review with meta-analysis," *BMC Public Health*, vol. 17, no. 1, p. 683, Dec. 2017, doi: 10.1186/s12889-017-4691-z.
- [4] M. Athanasiou, G. Fragkozidis, K. Zarkogianni, and K. S. Nikita, "Long Short-term Memory-Based Prediction of the Spread of Influenza-Like Illness Leveraging Surveillance, Weather, and Twitter Data: Model Development and Validation.," *J Med Internet Res*, vol. 25, p. e42519, Feb. 2023, doi: 10.2196/42519.
- [5] G. Fico *et al.*, "What do healthcare professionals need to turn risk models for type 2 diabetes into usable computerized clinical decision support systems? Lessons learned from the MOSAIC project.," *BMC Med Inform Decis Mak*, vol. 19, no. 1, p. 163, Aug. 2019, doi: 10.1186/s12911-019-0887-8.
- [6] K. Zarkogianni, M. Athanasiou, and A. C. Thanopoulou, "Comparison of Machine Learning Approaches Toward Assessing the Risk of Developing Cardiovascular Disease as a Long-Term Diabetes Complication.," *IEEE J Biomed Health Inform*, vol. 22, no. 5, pp. 1637–1647, Sep. 2018, doi: 10.1109/JBHI.2017.2765639.
- [7] K. Dalakleidi, K. Zarkogianni, A. Thanopoulou, and K. Nikita, "Comparative assessment of statistical and machine learning techniques towards estimating the risk of developing type 2 diabetes and cardiovascular complications," *Expert Syst*, vol. 34, no. 6, p. e12214, Dec. 2017, doi: 10.1111/exsy.12214.
- [8] K. Dalakleidi, K. Zarkogianni, A. Thanopoulou, and K. Nikita, "Comparative assessment of statistical and machine learning techniques towards estimating the risk of developing type 2 diabetes and cardiovascular complications," *Expert Syst*, vol. 34, no. 6, p. e12214, Dec. 2017, doi: 10.1111/exsy.12214.
- [9] K. Zarkogianni *et al.*, "A Review of Emerging Technologies for the Management of Diabetes Mellitus.," *IEEE Trans Biomed Eng*, vol. 62, no. 12, pp. 2735–49, Dec. 2015, doi: 10.1109/TBME.2015.2470521.
- [10] P. Pervanidou *et al.*, "The Impact of the ENDORSE Digital Weight Management Program on the Metabolic Profile of Children and Adolescents with Overweight and Obesity and on Food Parenting Practices," *Nutrients*, vol. 15, no. 7, p. 1777, Apr. 2023, doi: 10.3390/nu15071777.

- [11] K. Zarkogianni *et al.*, “The ENDORSE Feasibility Study: Exploring the Use of M-Health, Artificial Intelligence and Serious Games for the Management of Childhood Obesity,” *Nutrients*, vol. 15, no. 6, p. 1451, Mar. 2023, doi: 10.3390/nu15061451.
- [12] IA. Vasilakis *et al.*, “The ENDORSE feasibility pilot trial – Adherence to study protocol,” *ATTD*, 2023.
- [13] E. Kalafatis *et al.*, “Artificial Intelligence Based Procedural Content Generation in Serious Games for Health: The Case of Childhood Obesity,” in *Wireless Mobile Communication and Healthcare*, A. Cunha, N. M. Garcia, J. Marx Gómez, and S. Pereira, Eds., Cham: Springer Nature Switzerland, 2023, pp. 207–219.
- [14] W. J. Morales Camacho, J. M. Molina Díaz, S. Plata Ortiz, J. E. Plata Ortiz, M. A. Morales Camacho, and B. P. Calderón, “Childhood obesity: Aetiology, comorbidities, and treatment,” *Diabetes Metab Res Rev*, vol. 35, no. 8, Nov. 2019, doi: 10.1002/dmrr.3203.
- [15] D. Hamilton, A. Dee, and I. J. Perry, “The lifetime costs of overweight and obesity in childhood and adolescence: a systematic review,” *Obesity Reviews*, vol. 19, no. 4, pp. 452–463, Apr. 2018, doi: 10.1111/obr.12649.
- [16] A. Basolo, S. Magno, F. Santini, and G. Ceccarini, “Ketogenic Diet and Weight Loss: Is There an Effect on Energy Expenditure?,” *Nutrients*, vol. 14, no. 9, p. 1814, Apr. 2022, doi: 10.3390/nu14091814.
- [17] A. Spinelli *et al.*, “Thinness, overweight, and obesity in 6- to 9-year-old children from 36 countries: The World Health Organization European Childhood Obesity Surveillance Initiative—COSI 2015–2017,” *Obesity Reviews*, vol. 22, no. S6, Nov. 2021, doi: 10.1111/obr.13214.
- [18] M. Hassapidou, “Prevalence of childhood obesity in Greece: Results from WHO Childhood Obesity Surveillance Initiative 2010–2020,” *Public Health and Toxicology*, vol. 2, no. Supplement 1, May 2022, doi: 10.18332/pht/149699.
- [19] G. D. C. Ball *et al.*, “Strategies to reduce attrition in managing paediatric obesity: A systematic review,” *Pediatr Obes*, vol. 16, no. 4, Apr. 2021, doi: 10.1111/ijpo.12733.
- [20] D. E. Wilfley *et al.*, “Dose, Content, and Mediators of Family-Based Treatment for Childhood Obesity,” *JAMA Pediatr*, vol. 171, no. 12, p. 1151, Dec. 2017, doi: 10.1001/jamapediatrics.2017.2960.
- [21] A. Kuhlemeier *et al.*, “Individual differences in the effects of the ACTION-PAC intervention: an application of personalized medicine in the prevention and treatment of obesity,” *J Behav Med*, vol. 45, no. 2, pp. 211–226, Apr. 2022, doi: 10.1007/s10865-021-00274-2.
- [22] L. Johansson, E. Hagman, and P. Danielsson, “A novel interactive mobile health support system for pediatric obesity treatment: a randomized controlled feasibility trial,” *BMC Pediatr*, vol. 20, no. 1, p. 447, Dec. 2020, doi: 10.1186/s12887-020-02338-9.

- [23] S. G. Trost, D. Sundal, G. D. Foster, M. R. Lent, and D. Vojta, "Effects of a Pediatric Weight Management Program With and Without Active Video Games," *JAMA Pediatr*, vol. 168, no. 5, p. 407, May 2014, doi: 10.1001/jamapediatrics.2013.3436.
- [24] J.-L. Chen, C. M. Guedes, and A. E. Lung, "Smartphone-based Healthy Weight Management Intervention for Chinese American Adolescents: Short-term Efficacy and Factors Associated With Decreased Weight," *Journal of Adolescent Health*, vol. 64, no. 4, pp. 443–449, Apr. 2019, doi: 10.1016/j.jadohealth.2018.08.022.
- [25] D. R. Lubans *et al.*, "Assessing the sustained impact of a school-based obesity prevention program for adolescent boys: the ATLAS cluster randomized controlled trial," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 13, no. 1, p. 92, Dec. 2016, doi: 10.1186/s12966-016-0420-8.
- [26] R. M. M. P. K. Rajapakse, J. M. A. I. Mudalige, L. A. D. Y. S. Perera, R. N. A. M. S. C. B. Warakagoda, and S. Siriwardana, "Mobile Based Solution to Weight Loss Planning for Children (with Obesity) in Sri Lanka," in *2021 3rd International Conference on Advancements in Computing (ICAC)*, IEEE, Dec. 2021, pp. 222–227. doi: 10.1109/ICAC54203.2021.9671115.
- [27] A. Mohammed Nawi and F. I. Che Jamaludin, "Effect of Internet-based Intervention on Obesity among Adolescents in Kuala Lumpur: A School-based Cluster Randomised Trial.," *Malays J Med Sci*, vol. 22, no. 4, pp. 47–56, 2015.
- [28] G. L. Tripicchio *et al.*, "Technology Components as Adjuncts to Family-Based Pediatric Obesity Treatment in Low-Income Minority Youth," *Childhood Obesity*, vol. 13, no. 6, pp. 433–442, Dec. 2017, doi: 10.1089/chi.2017.0021.
- [29] V. Cueto, C. J. Wang, and L. M. Sanders, "Impact of a Mobile App-Based Health Coaching and Behavior Change Program on Participant Engagement and Weight Status of Overweight and Obese Children: Retrospective Cohort Study," *JMIR Mhealth Uhealth*, vol. 7, no. 11, p. e14458, Nov. 2019, doi: 10.2196/14458.
- [30] S. Browne *et al.*, "Mobile Health Apps in Pediatric Obesity Treatment: Process Outcomes From a Feasibility Study of a Multicomponent Intervention.," *JMIR Mhealth Uhealth*, vol. 8, no. 7, p. e16925, Jul. 2020, doi: 10.2196/16925.
- [31] A. P. Vidmar *et al.*, "An addiction model-based mobile health weight loss intervention in adolescents with obesity," *Pediatr Obes*, vol. 14, no. 2, Feb. 2019, doi: 10.1111/ijpo.12464.
- [32] N. Gómez-Del-Río, C. S. González-González, P. A. Toledo-Delgado, V. Muñoz-Cruz, and F. García-Peñalvo, "Health Promotion for Childhood Obesity: An Approach Based on Self-Tracking of Data.," *Sensors (Basel)*, vol. 20, no. 13, Jul. 2020, doi: 10.3390/s20133778.
- [33] A. E. Staiano, A. M. Marker, R. A. Beyl, D. S. Hsia, P. T. Katzmarzyk, and R. L. Newton, "A randomized controlled trial of dance exergaming for exercise

training in overweight and obese adolescent girls," *Pediatr Obes*, vol. 12, no. 2, pp. 120–128, Apr. 2017, doi: 10.1111/ijpo.12117.

- [34] L. Saez *et al.*, "Reach and Acceptability of a Mobile Reminder Strategy and Facebook Group Intervention for Weight Management in Less Advantaged Adolescents: Insights From the PRALIMAP-INÈS Trial.," *JMIR Mhealth Uhealth*, vol. 6, no. 5, p. e110, May 2018, doi: 10.2196/mhealth.7657.
- [35] S. Armstrong, A. Mendelsohn, G. Bennett, E. M. Taveras, A. Kimberg, and A. R. Kemper, "Texting Motivational Interviewing: A Randomized Controlled Trial of Motivational Interviewing Text Messages Designed to Augment Childhood Obesity Treatment," *Childhood Obesity*, vol. 14, no. 1, pp. 4–10, Jan. 2018, doi: 10.1089/chi.2017.0089.
- [36] F. Duman, M. H. Kokaçya, E. Doğru, N. Katayıfci, Ö. Canbay, and F. Aman, "The Role of Active Video-Accompanied Exercises in Improvement of the Obese State in Children: A Prospective Study from Turkey," *J Clin Res Pediatr Endocrinol*, vol. 8, no. 3, pp. 334–340, Sep. 2016, doi: 10.4274/jcrpe.2284.
- [37] M. Miguet *et al.*, "Cognitive restriction accentuates the increased energy intake response to a 10-month multidisciplinary weight loss program in adolescents with obesity," *Appetite*, vol. 134, pp. 125–134, Mar. 2019, doi: 10.1016/j.appet.2018.12.015.
- [38] G. Colmenarejo, "Machine Learning Models to Predict Childhood and Adolescent Obesity: A Review," *Nutrients*, vol. 12, no. 8, p. 2466, Aug. 2020, doi: 10.3390/nu12082466.
- [39] R. Hammond *et al.*, "Predicting childhood obesity using electronic health records and publicly available data," *PLoS One*, vol. 14, no. 4, p. e0215571, Apr. 2019, doi: 10.1371/journal.pone.0215571.
- [40] X. Pang, C. B. Forrest, F. Le-Scherban, and A. J. Masino, "Understanding Early Childhood Obesity via Interpretation of Machine Learning Model Predictions," in *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, IEEE, Dec. 2019, pp. 1438–1443. doi: 10.1109/ICMLA.2019.00235.
- [41] B. Park, C.-S. Chung, M. J. Lee, and H. Park, "Accurate neuroimaging biomarkers to predict body mass index in adolescents: a longitudinal study," *Brain Imaging Behav*, vol. 14, no. 5, pp. 1682–1695, Oct. 2020, doi: 10.1007/s11682-019-00101-y.
- [42] B. Singh and H. Tawfik, "Machine Learning Approach for the Early Prediction of the Risk of Overweight and Obesity in Young People," 2020, pp. 523–535. doi: 10.1007/978-3-030-50423-6_39.
- [43] M. Gupta, T.-L. T. Phan, T. Bunnell, and R. Beheshti, "Obesity Prediction with EHR Data: A deep learning approach with interpretable elements," Dec. 2019, doi: 10.1145/3506719.
- [44] C. Kim, F. J. Costello, K. C. Lee, Y. Li, and C. Li, "Predicting Factors Affecting Adolescent Obesity Using General Bayesian Network and What-If Analysis," *Int*

J Environ Res Public Health, vol. 16, no. 23, p. 4684, Nov. 2019, doi: 10.3390/ijerph16234684.

- [45] I. Lee, K.-S. Bang, H. Moon, and J. Kim, "Risk Factors for Obesity Among Children Aged 24 to 80 months in Korea: A Decision Tree Analysis," *J Pediatr Nurs*, vol. 46, pp. e15–e23, May 2019, doi: 10.1016/j.pedn.2019.02.004.
- [46] Z. Zheng and K. Ruggiero, "Using machine learning to predict obesity in high school students," in *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, IEEE, Nov. 2017, pp. 2132–2138. doi: 10.1109/BIBM.2017.8217988.
- [47] P. Wiechmann, K. Lora, P. Branscum, and J. Fu, "Identifying Discriminative Attributes to Gain Insights Regarding Child Obesity in Hispanic Preschoolers Using Machine Learning Techniques," in *2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI)*, IEEE, Nov. 2017, pp. 11–15. doi: 10.1109/ICTAI.2017.00014.
- [48] N. Rios-Julian, A. Alarcon-Paredes, G. A. Alonso, D. Hernandez-Rosales, and I. P. Guzman-Guzman, "Feasibility of a screening tool for obesity diagnosis in Mexican children from a vulnerable community of Me'Phaa ethnicity in the State of Guerrero, Mexico," in *2017 Global Medical Engineering Physics Exchanges/Pan American Health Care Exchanges (GMEPE/PAHCE)*, IEEE, Mar. 2017, pp. 1–6. doi: 10.1109/GMEPE-PAHCE.2017.7972105.
- [49] F. S. Abdullah *et al.*, "Data Mining Techniques for Classification of Childhood Obesity Among Year 6 School Children," 2017, pp. 465–474. doi: 10.1007/978-3-319-51281-5_47.
- [50] M. Hasan, A. Kotov, A. I. Carcone, M. Dong, and S. Naar, "Predicting the Outcome of Patient-Provider Communication Sequences using Recurrent Neural Networks and Probabilistic Models.," *AMIA Jt Summits Transl Sci Proc*, vol. 2017, pp. 64–73, 2018.
- [51] M. Hasan, A. Kotov, A. I. Carcone, M. Dong, and S. Naar, "Predicting the Outcome of Patient-Provider Communication Sequences using Recurrent Neural Networks and Probabilistic Models.," *AMIA Jt Summits Transl Sci Proc*, vol. 2017, pp. 64–73, 2018.
- [52] N. Öksüz, I. Shcherbatyi, T. Kowatsch, and W. Mass, "A Data-analytical System to Predict Therapy Success for Obese Children," in *ICIS*, 2018.
- [53] H. Fayyaz, T.-L. T. Phan, H. T. Bunnell, and R. Beheshti, "Who will Leave a Pediatric Weight Management Program and When? -- A machine learning approach for predicting attrition patterns," Feb. 2022.
- [54] A. M. Antoniadis *et al.*, "Current Challenges and Future Opportunities for XAI in Machine Learning-Based Clinical Decision Support Systems: A Systematic Review," *Applied Sciences*, vol. 11, no. 11, p. 5088, May 2021, doi: 10.3390/app11115088.
- [55] S. Tonekaboni, S. Joshi, M. D. McCradden, and A. Goldenberg, "What Clinicians Want: Contextualizing Explainable Machine Learning for Clinical End

Use,” in *Proceedings of the 4th Machine Learning for Healthcare Conference*, F. Doshi-Velez, J. Fackler, K. Jung, D. Kale, R. Ranganath, B. Wallace, and J. Wiens, Eds., in *Proceedings of Machine Learning Research*, vol. 106. PMLR, Jun. 2019, pp. 359–380. [Online]. Available: <https://proceedings.mlr.press/v106/tonkaboni19a.html>

- [56] E. Yom-Tov, G. Feraru, M. Kozdoba, S. Mannor, M. Tennenholtz, and I. Hochberg, “Encouraging Physical Activity in Patients With Diabetes: Intervention Using a Reinforcement Learning System,” *J Med Internet Res*, vol. 19, no. 10, p. e338, Oct. 2017, doi: 10.2196/jmir.7994.
- [57] M. Zhou *et al.*, “Evaluating Machine Learning–Based Automated Personalized Daily Step Goals Delivered Through a Mobile Phone App: Randomized Controlled Trial,” *JMIR Mhealth Uhealth*, vol. 6, no. 1, p. e28, Jan. 2018, doi: 10.2196/mhealth.9117.
- [58] F. Gasparetti, L. M. Aiello, and D. Quercia, “Personalized weight loss strategies by mining activity tracker data,” *User Model User-adapt Interact*, vol. 30, no. 3, pp. 447–476, Jul. 2020, doi: 10.1007/s11257-019-09242-7.
- [59] G. Rebal, A. Ravi, and S. Churiwala, “Machine Learning Definition and Basics,” in *An Introduction to Machine Learning*, Cham: Springer International Publishing, 2019, pp. 1–17. doi: 10.1007/978-3-030-15729-6_1.
- [60] J. Alzubi, A. Nayyar, and A. Kumar, “Machine Learning from Theory to Algorithms: An Overview,” *J Phys Conf Ser*, vol. 1142, p. 012012, Nov. 2018, doi: 10.1088/1742-6596/1142/1/012012.
- [61] D. Bertsimas and J. Dunn, “Optimal classification trees,” *Mach Learn*, vol. 106, no. 7, pp. 1039–1082, Jul. 2017, doi: 10.1007/s10994-017-5633-9.
- [62] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” Mar. 2016, doi: 10.1145/2939672.2939785.
- [63] J. Grus, *Data Science from Scratch: First Principles with Python*, 1st ed. O’Reilly Media, Inc., 2015.
- [64] B. Venkatesh and J. Anuradha, “A Review of Feature Selection and Its Methods,” *Cybernetics and Information Technologies*, vol. 19, no. 1, pp. 3–26, Mar. 2019, doi: 10.2478/cait-2019-0001.
- [65] F. Y. Okay, M. Yildirim, and S. Ozdemir, “Interpretable Machine Learning: A Case Study of Healthcare,” in *2021 International Symposium on Networks, Computers and Communications (ISNCC)*, IEEE, Oct. 2021, pp. 1–6. doi: 10.1109/ISNCC52172.2021.9615727.
- [66] H. W. Loh, C. P. Ooi, S. Seoni, P. D. Barua, F. Molinari, and U. R. Acharya, “Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011–2022),” *Comput Methods Programs Biomed*, vol. 226, p. 107161, 2022, doi: <https://doi.org/10.1016/j.cmpb.2022.107161>.
- [67] S. Lundberg and S.-I. Lee, “A Unified Approach to Interpreting Model Predictions,” May 2017.

- [68] S. M. Lundberg, G. G. Erion, and S.-I. Lee, "Consistent Individualized Feature Attribution for Tree Ensembles," Feb. 2018.
- [69] T. J. Loftus *et al.*, "Uncertainty-aware deep learning in healthcare: A scoping review," *PLOS Digital Health*, vol. 1, no. 8, p. e0000085, Aug. 2022, doi: 10.1371/journal.pdig.0000085.
- [70] M. Bukowski, J. Kurek, I. Antoniuk, and A. Jegorowa, "Decision Confidence Assessment in Multi-Class Classification," *Sensors*, vol. 21, no. 11, p. 3834, Jun. 2021, doi: 10.3390/s21113834.
- [71] S. Mishra and B. M. Ayyub, "Shannon Entropy for Quantifying Uncertainty and Risk in Economic Disparity," *Risk Analysis*, vol. 39, no. 10, pp. 2160–2181, Oct. 2019, doi: 10.1111/risa.13313.
- [72] R. B. Bendel, S. S. Higgins, J. E. Teberg, and D. A. Pyke, "Comparison of skewness coefficient, coefficient of variation, and Gini coefficient as inequality measures within populations," *Oecologia*, vol. 78, no. 3, pp. 394–400, 1989, doi: 10.1007/BF00379115.
- [73] S. Liu, K. Y. Ngiam, and M. Feng, "Deep Reinforcement Learning for Clinical Decision Support: A Brief Survey," Jul. 2019.
- [74] S. Datta *et al.*, "Reinforcement learning in surgery," *Surgery*, vol. 170, no. 1, pp. 329–332, Jul. 2021, doi: 10.1016/j.surg.2020.11.040.
- [75] S. Liu, K. C. See, K. Y. Ngiam, L. A. Celi, X. Sun, and M. Feng, "Reinforcement Learning for Clinical Decision Support in Critical Care: Comprehensive Review," *J Med Internet Res*, vol. 22, no. 7, p. e18477, Jul. 2020, doi: 10.2196/18477.
- [76] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction, 2nd ed.* in Adaptive computation and machine learning. Cambridge, MA, US: The MIT Press, 2018.
- [77] D. Musher-Eizenman and S. Holub, "Comprehensive Feeding Practices Questionnaire: Validation of a New Measure of Parental Feeding Practices," *J Pediatr Psychol*, vol. 32, no. 8, pp. 960–972, Apr. 2007, doi: 10.1093/jpepsy/jsm037.
- [78] M. Michou, D. B. Panagiotakos, E. Mamalaki, M. Yannakoulia, and V. Costarelli, "Development and validation of the Greek version of the comprehensive parental feeding questionnaire," *Med J Nutrition Metab*, vol. 12, no. 2, pp. 211–221, Jun. 2019, doi: 10.3233/MNM-180275.
- [79] R. Goodman, "The Strengths and Difficulties Questionnaire: A Research Note," *Journal of Child Psychology and Psychiatry*, vol. 38, no. 5, pp. 581–586, Jul. 1997, doi: 10.1111/j.1469-7610.1997.tb01545.x.
- [80] I. Bibou-Nakou, G. Kiosseoglou, and A. Stogiannidou, "Strengths and difficulties of school-aged children in the family and school context.," *Psychology: The Journal of the Hellenic Psychological Society*, vol. 8, pp. 506–525, 2001.

- [81] K. Kroenke, R. L. Spitzer, and J. B. W. Williams, "The PHQ-9," *J Gen Intern Med*, vol. 16, no. 9, pp. 606–613, Sep. 2001, doi: 10.1046/j.1525-1497.2001.016009606.x.
- [82] D. M. Garner, M. P. Olmsted, Y. Bohr, and P. E. Garfinkel, "The Eating Attitudes Test: psychometric features and clinical correlates," *Psychol Med*, vol. 12, no. 4, pp. 871–878, Nov. 1982, doi: 10.1017/S0033291700049163.
- [83] G. Stergiou, C. J. Stefanidis, S. Stabouli, and S. Douma, "Practical Recommendations for the Diagnosis, Investigation and Management of Hypertension in Children and Adolescents: Hellenic Society of Hypertension Consensus Document," *Hellenic J Cardiol*, vol. 54, pp. 199–211, 2013, [Online]. Available: <https://www.researchgate.net/publication/236920065>
- [84] M. Emmanuel and B. R. Bokor, *Tanner Stages*. 2023.
- [85] A. Coronato, M. Naeem, G. De Pietro, and G. Paragliola, "Reinforcement learning for intelligent healthcare applications: A survey," *Artif Intell Med*, vol. 109, p. 101964, Sep. 2020, doi: 10.1016/j.artmed.2020.101964.
- [86] C. Tudor-Locke *et al.*, "How many steps/day are enough? for children and adolescents," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 8, no. 1, p. 78, 2011, doi: 10.1186/1479-5868-8-78.
- [87] M. P. da Silva, F. E. Fontana, E. Callahan, O. Mazzardo, and W. De Campos, "Step-Count Guidelines for Children and Adolescents: A Systematic Review," *J Phys Act Health*, vol. 12, no. 8, pp. 1184–1191, Aug. 2015, doi: 10.1123/jpah.2014-0202.
- [88] R. von Kries, A. Toschke, H. Wurmser, T. Sauerwald, and B. Koletzko, "Reduced risk for overweight and obesity in 5- and 6-y-old children by duration of sleep—a cross-sectional study," *Int J Obes*, vol. 26, no. 5, pp. 710–716, May 2002, doi: 10.1038/sj.ijo.0801980.
- [89] S. Nemati, M. M. Ghassemi, and G. D. Clifford, "Optimal medication dosing from suboptimal clinical examples: A deep reinforcement learning approach," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, Aug. 2016, pp. 2978–2981. doi: 10.1109/EMBC.2016.7591355.
- [90] C. Diuk and M. Littman, "Hierarchical Reinforcement Learning," in *Encyclopedia of Artificial Intelligence*, IGI Global, 2009, pp. 825–830. doi: 10.4018/978-1-59904-849-9.ch122.
- [91] J. Eschmann, "Reward Function Design in Reinforcement Learning," 2021, pp. 25–33. doi: 10.1007/978-3-030-41188-6_3.
- [92] A. Gupta, A. Pacchiano, Y. Zhai, S. Kakade, and S. Levine, "Unpacking Reward Shaping: Understanding the Benefits of Reward Engineering on Sample Complexity," in *Advances in Neural Information Processing Systems*, S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, Eds., Curran Associates, Inc., 2022, pp. 15281–15295. [Online]. Available:

https://proceedings.neurips.cc/paper_files/paper/2022/file/6255f22349da5f2126dfc0b007075450-Paper-Conference.pdf