



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

**ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ**

**Ανάπτυξη Μοντέλου Ταξινόμησης Ημερήσιου
Γλυκαιμικού Προφίλ Ατόμων που Πάσχουν από
Σακχαρώδη Διαβήτη Τύπου 1**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

ΚΩΝΣΤΑΝΤΙΝΑΣ Κ. ΔΙΔΥΜΙΩΤΟΥ-ΚΑΟΥΚΑΚΗ

Επιβλέπουσα : Κωνσταντίνα Νικήτα
Καθηγήτρια Ε.Μ.Π.

Αθήνα, Φεβρουάριος 2021



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

Ανάπτυξη Μοντέλου Ταξινόμησης Ημερήσιου Γλυκαιμικού Προφίλ Ατόμων που Πάσχουν από Σακχαρώδη Διαβήτη Τύπου 1

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

ΚΩΝΣΤΑΝΤΙΝΑΣ Κ. ΔΙΔΥΜΙΩΤΟΥ-ΚΑΟΥΚΑΚΗ

Επιβλέπουσα : Κωνσταντίνα Νικήτα
Καθηγήτρια Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 26η Φεβρουαρίου 2021

(Υπογραφή)

.....
Κωνσταντίνα Νικήτα
Καθηγήτριας Ε.Μ.Π.

(Υπογραφή)

.....
Διονύσιος - Δημήτρης
Κουτσούρης
Καθηγητής Ε.Μ.Π

(Υπογραφή)

.....
Γεώργιος Στάμου
Καθηγητής Ε.Μ.Π

Αθήνα, Φεβρουάριος 2021



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

(Υπογραφή)

.....

Διδυμιώτου - Καουκάκη Κωνσταντίνα

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

Copyright © Διδυμιώτου - Καουκάκη Κωνσταντίνα, 2021

Με την 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 το συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις τους Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Αντικείμενο της παρούσας διπλωματικής εργασίας αποτελεί η ανάπτυξη ενός μοντέλου ταξινόμησης οκτάωρων γλυκαιμικών προφίλ που λαμβάνονται καθημερινά από Διατάξεις Συνεχούς Μέτρησης Γλυκόζης (ΔΣΜΓ) ατόμων που πάσχουν από Σακχαρώδη Διαβήτη Τύπου 1 (ΣΔΤ1) με στόχο την ανάπτυξη ενός συστήματος που λειτουργεί ως εργαλείο υποστήριξης της διαχείρισης του μεταβολισμού γλυκόζης. Η ανάπτυξη του μοντέλου βασίστηκε στην ανάλυση πραγματικών δεδομένων γλυκόζης ως Δεδομένα Σύνθεσης και στην εφαρμογή τεχνικών Μηχανικής Μάθησης. Για κάθε οκτάωρη περίοδο υπολογίστηκε ένα διάνυσμα (σύνθεση) που φανερώνει πως κατανεμήθηκε ο χρόνος που πέρασε ο ασθενής σε έξι διαστήματα γλυκόζης σχετικά με υπογλυκαιμία, ευγλυκαιμία και υπεργλυκαιμία. Ο διαχωρισμός των διανυσμάτων σε συστάδες, που προέκυψε από την εφαρμογή αλγορίθμου συσταδοποίησης, χρησιμοποιήθηκε για την ανάπτυξη μοντέλων ταξινόμησης οκτάωρων γλυκαιμικών προφίλ. Τα μοντέλα ταξινόμησης αξιολογήθηκαν με τη μέθοδο διασταυρούμενης επικύρωσης LOO-CV. Από τους ταξινομητές που ελέγχθηκαν καλύτερη επίδοση παρουσίασε το μοντέλο SVM με ακρίβεια 93% κατά μέσο όρο. Παράλληλα, διερευνήθηκε πιθανοτικό μοντέλο για την εκτίμηση των μεταβάσεων μεταξύ διαδοχικών οκτάωρων περιόδων μέσα στην μέρα. Για την απόδοση των μοντέλων συσταδοποίησης χρησιμοποιήθηκαν κυρίως κλινικά κριτήρια με σκοπό να διαπιστωθεί η ομοιογένεια των κλάσεων, ενώ για την αξιολόγηση της απόδοσης των ταξινομητών χρησιμοποιήθηκαν μαθηματικά κριτήρια. Τα αποτελέσματα έδειξαν ότι η ανάλυση Δεδομένων Σύνθεσης είναι αρκετά ικανοποιητική για την κατηγοριοποίηση γλυκαιμικών προφίλ.

Λέξεις Κλειδιά

Σακχαρώδης Διαβήτης Τύπου 1, Διατάξεις Συνεχούς μέτρησης Γλυκόζης, Ανάλυση Δεδομένων Σύνθεσης, Τεχνικές Μηχανικής Μάθησης, Συσταδοποίηση, Ταξινόμηση

Abstract

The aim of the current diploma thesis is the development of a classification model of 8-hour glyceic profiles obtained from continuous glucose monitoring systems (CGMS) worn by people suffering from Type 1 Diabetes Mellitus (T1DM). The applied methodology is based on compositional data analysis (CoDa) and machine learning techniques. Compositional data analysis is applied to three 8-hour periods. Glucose values for each period were distributed in six different ranges, relative to hypoglycemia, normoglycemia and hyperglycemia, which determined the composition vector. A clustering algorithm was applied to the centered log-ratio coordinates of the compositions to group 8-hour periods into categories with similar characteristics. A set of classifiers were applied, using the valid 8-hour compositions and the respective categories, in order to test their capacity to automatically categorize never before seen 8-hour periods. Leave one out cross validation (LOO-CV) method was performed for the evaluation. From all the models examined the best performance was demonstrated by the SVM model with an average accuracy of 93%. A probabilistic model of transitions between the categories of the preceding 8-hour period to the categories of the following 8-hour period was obtained. Results showed that the compositional data analysis of the time spent in different glucose ranges is suitable for the categorization of glucose profiles into groups and it could be a useful tool for the management of glucose metabolism.

Keywords

Type 1 Diabetes Mellitus, Continuous Glucose Monitoring Systems, Compositional Data Analysis, Machine Learning Techniques, Clustering, Classification

Ευχαριστίες

Θα ήθελα καταρχήν να ευχαριστήσω θερμά την επιβλέπουσα καθηγήτρια μου κα. Κωνσταντίνα Νικήτα για την ευκαιρία που μου έδωσε να εκπονήσω και να ανακαλύψω σε βάθος ένα θέμα που με ενδιέφερε πολύ. Επίσης, ευχαριστώ ιδιαίτερα την διδάκτορα κα. Κωνσταντία Ζαρκογιάννη και τον υποψήφιο διδάκτορα Κωνσταντίνο Μήτση για την εξαιρετική συνεργασία που είχαμε και την διαρκή καθοδήγησή και υποστήριξη τους καθ' όλη την διάρκεια της εκπόνησης της εργασίας.

Δεν θα μπορούσα να μην ευχαριστήσω τους φίλους μου που έκαναν τα φοιτητικά μου χρόνια λίγο πιο ανέμελα, εφοδιάζοντας τα με εμπειρίες, στιγμές και αναμνήσεις. Ακόμη, θέλω να εκφράσω την ευγνωμοσύνη μου στον Κωνσταντίνο Σπανό που με ενθάρρυνε πάντα να κάνω όσα ονειρεύομαι. Τέλος, το μεγαλύτερο και πιο εγκάρδιο ευχαριστώ το οφείλω στην οικογένεια μου για την υπομονή και την αμέριστη συμπαράσταση τους όλα αυτά τα χρόνια, αλλά κυρίως γιατί δεν έπαψαν ποτέ να πιστεύουν σε μένα.

Η παρούσα διπλωματική εργασία εκπονήθηκε στο πλαίσιο του ερευνητικού προγράμματος ENDORSE, που χρηματοδοτείται από το ΕΣΠΑ 2014-2020 (Συμφωνία Επιχορήγησης: Τ1ΕΔΚ-03695).

*...στον μπαμπά μου,
που μου λείπει πολύ*

Περιεχόμενα

Περίληψη.....	6
Abstract	7
Ευχαριστίες	9
Περιεχόμενα	13
Υπότιτλοι Σχημάτων	16
Υπέριτλοι Πινάκων.....	19
Πρόλογος	21
1ο Κεφάλαιο	23
Μεταβολισμός Γλυκόζης.....	23
Εισαγωγή.....	23
1.1 Ομοιόσταση της Γλυκόζης	23
1.1.1 Ορμόνες: Ινσουλίνη και Γλυκαγόνη.....	24
1.2 Σακχαρώδης Διαβήτης	28
1.2.1 Ταξινόμηση του Σακχαρώδους Διαβήτη	29
1.2.2 Διαγνωστικά Κριτήρια Σακχαρώδους Διαβήτη.....	31
1.2.3 Επιπλοκές του Σακχαρώδους Διαβήτη.....	32
1.2.4 Επιδημιολογικά Στοιχεία.....	34
2ο Κεφάλαιο	39
Γλυκαιμικός Έλεγχος.....	39
Εισαγωγή.....	39
2.1 Διατάξεις Μέτρησης Γλυκόζης	40
2.1.1 Συμβατικές Διατάξεις.....	40
2.1.2 Διατάξεις Συνεχούς Μέτρησης Γλυκόζης	41
2.2 Μέθοδοι Εξωγενούς Χορήγησης Ινσουλίνης.....	47
2.2.1 Ημερήσιες Πολλαπλές Ενέσιμες Δόσεις Ινσουλίνης	49
2.2.2 Αντλίες Συνεχούς Έγχυσης Ινσουλίνης.....	51
2.3 Μηχανική Μάθηση και Γλυκαιμικός Έλεγχος.....	53
3ο Κεφάλαιο	55

Θεωρητικό Υπόβαθρο	55
Εισαγωγή.....	55
3.1 Ανάλυση Δεδομένων Σύνθεσης	55
3.1.1 Βασικές Έννοιες.....	55
3.1.2 Γεωμετρία Aitchison	57
3.1.3 Γραμμικοί Μετασχηματισμοί.....	60
3.2 Μηχανική Μάθηση.....	64
3.2.1 Μάθηση με Επίβλεψη.....	64
3.2.2 Μάθηση χωρίς Επίβλεψη	65
3.2.3 Συσταδοποίηση.....	66
3.2.4 Ταξινόμηση	70
4ο Κεφάλαιο	75
Μεθοδολογία	75
Εισαγωγή.....	75
4.1 Συλλογή και Προεπεξεργασία Δεδομένων	76
4.1.1 Συλλογή Δεδομένων.....	76
4.1.2 Προεπεξεργασία Δεδομένων	77
4.2 Ανάλυση των Χρονοσειρών Γλυκόζης ως Δεδομένα Σύνθεσης.....	77
4.2.1 Καθορισμός διανυσμάτων σύνθεσης.....	77
4.2.2 Διαχείριση μηδενικών τιμών	78
4.2.3 Μετασχηματισμοί για δημιουργία διανυσμάτων εισόδου.....	79
4.3 Μοντέλα Συσταδοποίησης	81
4.4 Μοντέλα Ταξινόμησης	84
4.5 Πιθανοτικό Μοντέλο Μεταβάσεων.....	85
5ο Κεφάλαιο	87
Αποτελέσματα	87
5.1 Σύνολα Δεδομένων.....	87
5.2 Αποτελέσματα Συσταδοποίησης	88
5.3 Αποτελέσματα Ταξινομητών.....	119

5.4	Πιθανότητες Μετάβασης.....	120
6ο	Κεφάλαιο	123
	Συζήτηση.....	123
6.1	Κατηγοριοποίηση οκτάωρων γλυκαιμικών προφίλ.....	123
6.1.1	Διαγράμματα Biplots.....	123
6.1.2	Διαγράμματα Barplots και Καμπύλες Γλυκόζης.....	125
6.1.3	Κλινική Ανάλυση των κλάσεων.....	127
6.2	Αξιολόγηση Ταξινομητών.....	130
6.3	Αξιολόγηση Μεταβάσεων.....	132
7ο	Κεφάλαιο	135
	Συμπεράσματα - Μελλοντική Έρευνα.....	135
7.1	Συμπεράσματα.....	135
7.2	Μελλοντική Έρευνα.....	136
	Βιβλιογραφία.....	138

Υπότιτλοι Σχημάτων

Σχήματα	σελ.
Σχήμα 1.1: Μηχανισμός διατήρησης σταθερής συγκέντρωσης της γλυκόζης στο αίμα	25
Σχήμα 1.2: Νήσος του Langerhans. Με φθορίζοντα αντισώματα φωτίζονται οι διαφορετικοί τύποι κυττάρων. Τα α-κύτταρα που εκκρίνουν γλυκαγόνη φαίνονται στην περιφέρεια (κόκκινα) και τα β-κύτταρα που εκκρίνουν ινσουλίνη αποτελούν την πλειονότητα των κυττάρων και παρουσιάζονται σε κεντρική θέση (μπλε)	25
Σχήμα 1.3: Μεταβολικές δράσεις της ινσουλίνης στον μυϊκό, λιπώδη και ηπατικό ιστό	27
Σχήμα 1.4: Εξάπλωση του ΣΔ για το έτος 2019 ανά γεωγραφική περιοχή	35
Σχήμα 1.5: Εξάπλωση του μη διαγνωσμένου ΣΔ για το έτος 2019 ανά γεωγραφική περιοχή	35
Σχήμα 1.6: Εξάπλωση του ΣΔ για τα έτη 2019, 2030, 2045 ανά ηλικιακή ομάδα (έτη)	36
Σχήμα 1.7: Αριθμός ατόμων με ΣΔ (ηλικίας 20-79 ετών) που ζουν σε αστικές και αγροτικές περιοχές το 2019, 2030 και 2045	36
Σχήμα 1.8: Ποσοστά νέων περιπτώσεων ΣΔΤ1 (ανά 100.000 πληθυσμό ετησίως) σε παιδιά ηλικίας 0-14 ετών για το έτος 2019	37
Σχήμα 1.9: Συνολικές υγειονομικές δαπάνες που σχετίζονται με τον ΣΔ για ενήλικες ηλικίας 20-70 ετών που πάσχουν από ΣΔ	38
Σχήμα 1.10: Συνολικές δαπάνες υγείας που σχετίζονται με τον ΣΔ για ενήλικες ηλικίας 20-70 ετών που πάσχουν από ΣΔ (ΔΟΔ περιοχές)	38
Σχήμα 2.1: (α) Μετρητής γλυκόζης finger sticks, (β) Διαδικασία μέτρησης με διάτρηση δακτύλου	40
Σχήμα 2.2: Αντιπροσωπευτικά δεδομένα παρακολούθησης της γλυκόζης στο αίμα (BG mg/dL) που λαμβάνονται με ΔΣΜΓ (συνεχής μπλε γραμμή) και με συμβατικές διατάξεις (πράσινες κουκκίδες). Οι διακεκομμένοι κύκλοι υποδηλώνουν υπεργλυκαιμικά ή υπογλυκαιμικά επεισόδια, τα οποία χρησιμοποιώντας συμβατικές διατάξεις μέτρησης γλυκόζης δεν είναι ανιχνεύσιμα	42
Σχήμα 2.3: Εξέλιξη της ακρίβειας των σύγχρονων συστημάτων συνεχούς μέτρησης της γλυκόζης με την πάροδο των χρόνων. Από αριστερά: Medtronic Enlite, Abbott Freestyle Navigator, Dexcom G4 Platinum, Abbott Freestyle Libre, Dexcom G4 Platinum με λογισμικό 505, Senseonics Eversense, Dexcom G5, Dexcom G6	43
Σχήμα 2.4: Προηγμένες διατάξεις συνεχούς καταγραφής της συγκέντρωσης της γλυκόζης. (α) Abbott-Freestyle Libre, (β) Medtronic-Guardian Sensor, (γ) Dexcom-	45

Σχήμα 2.5: ΔΣΜΓ που αποτελείται από τον αισθητήρα μέτρησης της γλυκόζης και την αντλία ινσουλίνης. Ο αισθητήρας (Glucose Sensor), υπό την μορφή βελόνας, είναι τοποθετημένος κάτω από το δέρμα στον υποδόριο ιστό της κοιλιάς. Ανιχνεύει την συγκέντρωση γλυκόζης στο μεσοκυττάριο υγρό (Tissue Fluid). Λόγω της μέτρησης στο διάμεσο υγρό υπάρχει καθυστέρηση στο αποτέλεσμα σε σχέση με το τριχοειδικό αίμα περίπου 10-20 λεπτά	45
Σχήμα 2.6: Ενδογενής έκκριση ινσουλίνης (basal και bolus)	48
Σχήμα 2.7: Γράφημα που απεικονίζει τις διαφορές στους χρόνους δράσης μεταξύ των σκευασμάτων ινσουλίνης	49
Σχήμα 2.8: Σύριγγες χορήγησης ινσουλίνης	50
Σχήμα 2.9: Στυλό Ινσουλίνης	50
Σχήμα 2.10: Συσκευή χορήγησης ινσουλίνης με τη μορφή πίδακα υψηλής πίεσης	51
Σχήμα 2.11: Προηγμένες αντλίες ινσουλίνης με σετ έγχυσης. (α) Medtronic 670G, (β) Tandem t:slim x2	52
Σχήμα 2.12: Προηγμένες αντλίες ινσουλίνης τύπου patch. (α) V-GO, (β) Omnipod	52
Σχήμα 3.1: Απεικόνιση S^3 σε τριμερές διάγραμμα (Ternary)	56
Σχήμα 3.2: Απεικόνιση ενός τριμερούς simplex S^3 . Υπάρχουν 3 τμήματα, x_1, x_2, x_3 που αντιπροσωπεύουν τμήματα κάποιου συνόλου. Τα A,B,C,D,E είναι πέντε διαφορετικές συνθέσεις εντός του simplex. Τα A,B,C είναι ισοδύναμα μεταξύ τους όπως και τα D,E.	57
Σχήμα 3.3: Παράδειγμα του πίνακα προσήμου για $D = 6$, που χρησιμοποιείται για να κωδικοποιήσει μία διαδοχική δυαδική διχοτόμηση της σύνθεσης $x = [x_1, x_2, \dots, x_D]$ για την δημιουργία ορθοκανονικής βάσης. Το κάτω τμήμα του πίνακα δείχνει τον πίνακα Φ	64
Σχήμα 3.4: Δενδρόγραμμα ιεραρχικής ομαδοποίησης	68
Σχήμα 3.5: Αποτέλεσμα αλγορίθμου k-Μέσων και διαχωρισμός συνόλου δεδομένων σε 3 ομάδες. Οι γραμμές που χωρίζουν τις ομάδες αποτελούν τα όρια αποφάσεων	69
Σχήμα 3.6: Παράδειγμα δέντρου απόφασης για το αν παραμένει ή φεύγει ένας συνδρομητής δεδομένου την διάρκεια και το είδος του συμβολαίου του	71
Σχήμα 3.7: Παράδειγμα ταξινομητή k-NN. Το δείγμα x' ταξινομείται ως θετικό αν $k = 1$ και ως αρνητικό αν $k = 5$	72
Σχήμα 4.1: Σχεδιάγραμμα της προτεινόμενης μεθοδολογίας	76
Σχήμα 5.1: Clr biplots των διανυσμάτων που αντιστοιχούν στην 8-h περίοδο 00:00-07:59 για κάθε ασθενή. Οι κουκκίδες αναπαριστούν τα διανύσματα εισόδου στο δισδιάστατο χώρο. Ο διαχωρισμός τους σε ομάδες έχει προκύψει ύστερα από την εφαρμογή αλγορίθμου συσταδοποίησης	90
Σχήμα 5.2: Bar plots των μέσων clr συνιστωσών, για κάθε cluster και για κάθε ασθενή, για την πρώτη 8-h περίοδο της ημέρας (00:00-07:59). Σύγκριση του κέντρου κάθε ομάδας με το συνολικό κέντρο (μηδέν). Θετικές τιμές υποδηλώνουν υψηλές	94

τιμές στο συγκεκριμένο εύρος γλυκόζης, ενώ αρνητικές τιμές υποδηλώνουν χαμηλές τιμές	
Σχήμα 5.3: 8-h χρονοσειρές γλυκόζης (διάστημα 00:00-07:59) για κάθε ασθενή που δείχνουν τα γλυκαιμικά προφίλ κάθε κλάσης A, B, C	98
Σχήμα 5.4: Clr biplots των διανυσμάτων που αντιστοιχούν στην 8-h περίοδο 08:00-15:59 για κάθε ασθενή. Οι κουκκίδες αναπαριστούν τα διανύσματα εισόδου στο διδιάστατο χώρο. Ο διαχωρισμός τους σε ομάδες έχει προκύψει ύστερα από την εφαρμογή αλγόριθμου συσταδοποίησης	100
Σχήμα 5.5: Bar plots των μέσων clr συνιστωσών, για κάθε cluster και για κάθε ασθενή, για την δεύτερη 8-h περίοδο της ημέρας (08:00-15:59). Σύγκριση του κέντρου κάθε ομάδας με το συνολικό κέντρο (μηδέν). Θετικές τιμές υποδηλώνουν υψηλές τιμές στο συγκεκριμένο εύρος γλυκόζης, ενώ αρνητικές τιμές υποδηλώνουν χαμηλές τιμές	104
Σχήμα 5.6: 8-h χρονοσειρές γλυκόζης (διάστημα 08:00-15:59) για κάθε ασθενή που δείχνουν τα γλυκαιμικά προφίλ κάθε κλάσης A, B, C	108
Σχήμα 5.7: Clr biplots των διανυσμάτων που αντιστοιχούν στην 8-h περίοδο 16:00-23:59 για κάθε ασθενή. Οι κουκκίδες αναπαριστούν τα διανύσματα εισόδου στο διδιάστατο χώρο. Ο διαχωρισμός τους σε ομάδες έχει προκύψει ύστερα από την εφαρμογή αλγόριθμου συσταδοποίησης	110
Σχήμα 5.8: Bar plots των μέσων clr συνιστωσών, για κάθε cluster και για κάθε ασθενή, για την τρίτη 8-h περίοδο της ημέρας (16:00-23:59). Σύγκριση του κέντρου κάθε ομάδας με το συνολικό κέντρο (μηδέν). Θετικές τιμές υποδηλώνουν υψηλές τιμές στο συγκεκριμένο εύρος γλυκόζης, ενώ αρνητικές τιμές υποδηλώνουν χαμηλές τιμές	114
Σχήμα 5.9: 8-h χρονοσειρές γλυκόζης (διάστημα 16:00-23:59) για κάθε ασθενή που δείχνουν τα γλυκαιμικά προφίλ κάθε κλάσης A, B, C	118

Υπέρτιτλοι Πινάκων

Πίνακες	σελ.
Πίνακας 1.1: Διαγνωστικά Κριτήρια για μέτρηση FPG	32
Πίνακας 1.2: Διαγνωστικά Κριτήρια για μέτρηση OGTT	32
Πίνακας 2.1: Συγκεντρωτικός πίνακας χαρακτηριστικών των τριών διατάξεων μέτρησης γλυκόζης από τις εταιρείες Dexcom, Abbot και Medtronic	45
Πίνακας 2.2: Χαρακτηριστικά σκευασμάτων ινσουλίνης	49
Πίνακας 2.3: Σύνοψη των κύριων χαρακτηριστικών των μοντέλων συνεχούς έγχυσης υποδόριας ινσουλίνης για τους δύο τύπους αντλίας	53
Πίνακας 4.1: Μέθοδος διαδοχικής διχοτόμησης για δημιουργία ilr διανυσμάτων	80
Πίνακας 5.1: Πλήθος ημερών χωρίς κενά σε μετρήσεις, όπως αυτές πρόεκυψαν μετά την προεπεξεργασία του αρχικού συνόλου δεδομένων	88
Πίνακας 5.2: Πλήθος δειγμάτων ανά cluster (περίοδος: 00:00-07:59)	90
Πίνακας 5.3: Σύγκριση του κέντρου κάθε clr συνιστώσας με το κέντρο του συνόλου για κάθε cluster (διάστημα: 00:00-07:59)	91
Πίνακας 5.4: Συγκεντρωτικός πίνακας κλινικών δεδομένων για κάθε ομάδα (διάστημα: 00:00-07:59)	95
Πίνακας 5.5: Πλήθος δειγμάτων ανά cluster (περίοδος: 08:00-15:59)	100
Πίνακας 5.6: Σύγκριση του κέντρου κάθε clr συνιστώσας με το κέντρο του συνόλου για κάθε cluster (διάστημα: 08:00-15:59)	101
Πίνακας 5.7: Συγκεντρωτικός πίνακας κλινικών δεδομένων ανά ομάδα (διάστημα: 08:00-15:59)	105
Πίνακας 5.8: Πλήθος δειγμάτων ανά cluster (περίοδος: 16:00-23:59)	110
Πίνακας 5.9: Σύγκριση του κέντρου κάθε clr συνιστώσας με το κέντρο του συνόλου για κάθε cluster (διάστημα: 16:00-23:59)	111
Πίνακας 5.10: Συγκεντρωτικός πίνακας κλινικών δεδομένων ανά cluster (διάστημα: 16:00-23:59)	115
Πίνακας 5.11: Πιστότητα των ταξινομητών που δοκιμάστηκαν για την κατηγοριοποίηση γλυκαιμικού προφίλ περιόδου 00:00-7:59	119
Πίνακας 5.12: Πιστότητα των ταξινομητών που δοκιμάστηκαν για την κατηγοριοποίηση γλυκαιμικού προφίλ περιόδου 08:00-15:59	119
Πίνακας 5.13: Πιστότητα των ταξινομητών που δοκιμάστηκαν για την κατηγοριοποίηση γλυκαιμικού προφίλ περιόδου 16:00-23:59	120

Πίνακας 5.14: Πιθανότητες μετάβασης από την κλάση της μίας 8-h περιόδου στην κλάση της ακόλουθης. Για την μετάβαση μεταξύ των 8-h διαστημάτων 16:00-23:59 και 00:00-07:59 χρησιμοποιήθηκαν δεδομένα μόνο από συνεχόμενες ημέρες 120

Πρόλογος

Ο Σακχαρώδης Διαβήτης (ΣΔ) είναι μια χρόνια πάθηση άρρηκτα συνδεδεμένη με την διαταραχή του μεταβολισμού της γλυκόζης. Χαρακτηρίζεται από αδυναμία του οργανισμού να διατηρεί τα επίπεδα της γλυκόζης στο αίμα εντός φυσιολογικών ορίων. Η αύξηση της συγκέντρωσης του σακχάρου στο αίμα και γενικότερα η αδυναμία ρύθμισης των επιπέδων γλυκόζης οφείλεται είτε στην ελαττωμένη ή και καθόλου έκκριση ινσουλίνης από τα κύτταρα του παγκρέατος (ινσουλοεξαρτώμενος διαβήτης, ΣΔ Τύπου 1 ή ΣΔΤ1) είτε λόγω της μειωμένης ευαισθησίας των κυττάρων στην ινσουλίνη (ΣΔ Τύπου 2 ή ΣΔΤ2). Οι επιπτώσεις της απορύθμισης του μεταβολισμού της γλυκόζης οδηγούν με την πάροδο του χρόνου σε σοβαρές επιπλοκές της υγείας. Πολύ σημαντικό ρόλο για την αντιμετώπιση του ΣΔ, καθώς και την πρόληψη από μακροχρόνιες παθήσεις που προκαλεί, παίζει ο τακτικός έλεγχος των επιπέδων γλυκόζης στο αίμα και η εξωγενής χορήγηση ινσουλίνης, κοινώς γλυκαιμικός έλεγχος. Για το σκοπό αυτό έχουν αναπτυχθεί διατάξεις συνεχούς μέτρησης της γλυκόζης και αντλίες συνεχούς έγχυσης ινσουλίνης ώστε να προλαμβάνονται υπογλυκαιμικά και υπεργλυκαιμικά επεισόδια. Οι σύγχρονες αυτές διατάξεις στοχεύουν στην αύξηση της απόδοσης του ελέγχου και την απώτερη αναβάθμιση του τρόπου ζωής του ατόμου. Ωστόσο, σημαντικό εμπόδιο για την αντιμετώπιση του ΣΔ είναι η διαφορετική συμπεριφορά του μεταβολισμού γλυκόζης που υφίσταται μεταξύ ασθενών. Έτσι, ζωτικής σημασίας αποτελεί η προσαρμογή της θεραπευτικής διαδικασίας στις μεταβολικές ιδιαιτερότητες και τις εξατομικευμένες ανάγκες του κάθε ατόμου που πάσχει από ΣΔ.

Αντικείμενο της παρούσας διπλωματικής εργασίας αποτελεί η ανάπτυξη μοντέλου κατηγοριοποίησης και ταξινόμησης ημερήσιου γλυκαιμικού προφίλ ατόμων που πάσχουν από ΣΔΤ1 με απώτερο στόχο την εξαγωγή χρήσιμων πληροφοριών που μπορούν να αξιοποιηθούν από το ιατρικό προσωπικό για την παραγωγή εξατομικευμένων συστάσεων. Το μοντέλο βασίζεται στην ανάλυση των δεδομένων γλυκόζης ως Δεδομένων Σύνθεσης και σε τεχνικές Μηχανικής Μάθησης, όπως αλγόριθμοι συσταδοποίησης και ταξινόμησης.

Η εργασία αυτή διαρθρώνεται στα εξής επτά κεφάλαια:

- Στο **Κεφάλαιο 1** περιγράφεται συνοπτικά η φυσιολογία του μηχανισμού ρύθμισης της γλυκόζης στον ανθρώπινο οργανισμό. Επιπλέον, παρουσιάζονται οι τύποι του ΣΔ, οι επιπλοκές που προκαλεί καθώς και κάποια επιδημιολογικά στοιχεία.
- Στο **Κεφάλαιο 2** γίνεται εκτενής αναφορά στον γλυκαιμικό έλεγχο και στις μεθόδους για την ρύθμιση των επιπέδων γλυκόζης στο αίμα. Παρουσιάζονται διατάξεις μέτρησης της γλυκόζης αλλά και τρόποι εξωγενούς χορήγησης ινσουλίνης. Τέλος, γίνεται περιγραφή μεθόδων Μηχανικής Μάθησης για αξιοποίηση δεδομένων συνεχούς μέτρησης γλυκόζης.

- Στο **Κεφάλαιο 3** παρουσιάζεται το θεωρητικό και μαθηματικό υπόβαθρο που σχετίζεται με τις μεθόδους που χρησιμοποιήθηκαν. Συγκεκριμένα, παρουσιάζεται θεωρία για Δεδομένα Σύνθεσης (Compositional Data) και τεχνικές Μηχανικής Μάθησης (Machine Learning).
- Στο **Κεφάλαιο 4** περιγράφεται η προτεινόμενη μεθοδολογία. Πιο συγκεκριμένα, περιγράφεται η διαδικασία συλλογής και προεπεξεργασίας των δεδομένων και η αρχιτεκτονική του μοντέλου που υλοποιήθηκε.
- Στο **Κεφάλαιο 5** πραγματοποιείται μια εκτενής παρουσίαση των αποτελεσμάτων που προέκυψαν.
- Στο **Κεφάλαιο 6** πραγματοποιείται σχολιασμός των αποτελεσμάτων που παρουσιάστηκαν στο προηγούμενο κεφάλαιο.
- Στο **Κεφάλαιο 7** παρουσιάζονται τα συμπεράσματα που προκύπτουν καθώς και περιορισμοί της υπάρχουσας μελέτης, ενώ τίθενται οι βάσεις για μελλοντική επέκταση της έρευνας.

1ο Κεφάλαιο

Μεταβολισμός Γλυκόζης

Εισαγωγή

Ο Σακχαρώδης διαβήτης (ΣΔ) είναι μια χρόνια μεταβολική ασθένεια που σχετίζεται άμεσα με την λειτουργία του παγκρέατος και αποτελεί παγκοσμίως μία από τις σοβαρότερες αιτίες νοσηρότητας και θνησιμότητας [1]. Χαρακτηρίζεται από αυξημένα επίπεδα της γλυκόζης στο αίμα και διαταραχή των υδατανθράκων, των λιπών και των πρωτεϊνών. Συγκεκριμένα, η πάθηση σχετίζεται με την ισορροπία ορμονών (γλυκαγόνη - ινσουλίνη) που παράγονται από την ενδοκρινή μοίρα του παγκρέατος και συντελούν στην παραγωγή και αποθήκευση ενέργειας από τις τροφές (διάσπαση γλυκόζης) και την ρύθμιση του σακχάρου του αίματος [2]. Η δυσλειτουργία του μηχανισμού ομοιόστασης της γλυκόζης επιφέρει σοβαρές επιπτώσεις στο άτομο. Επομένως, είναι σημαντική η ρύθμιση και η διατήρηση των επιπέδων του σακχάρου στο αίμα εντός φυσιολογικών τιμών. Στο κεφάλαιο αυτό γίνεται εκτενής περιγραφή της ομοιόστασης της γλυκόζης σε υγιή άτομα και η παθοφυσιολογία του ΣΔ. Επίσης, περιγράφονται οι τύποι του ΣΔ, οι επιπτώσεις στην υγεία του ατόμου καθώς και άλλα επιδημιολογικά στοιχεία που αφορούν την νόσο.

1.1 Ομοιόσταση της Γλυκόζης

Όλα τα έμβια όντα χρειάζονται ενέργεια και πρώτες ύλες για να επιτελέσουν τις φυσιολογικές λειτουργίες τους. Οι ζωικοί οργανισμοί όπως και ο άνθρωπος εξασφαλίζουν την ενέργεια αυτή μέσω της τροφής τους. Τα βασικά συστατικά των τροφίμων είναι οι πρωτεΐνες, οι υδατάνθρακες, τα λίπη, τα άλατα και τα ιχνοστοιχεία, οι βιταμίνες και οι φυτικές ίνες. Τα συστατικά αυτά μεταβολίζονται, δηλαδή μετατρέπονται με βιοχημικές διαδικασίες σε ενώσεις που αποσκοπούν στην παραγωγή και απελευθέρωση ενέργειας. Η γλυκόζη είναι ένας απλός μονοσακχαρίτης ο οποίος απορροφάται άμεσα στην κυκλοφορία του αίματος κατά την διάρκεια της πέψης. Τα κύτταρα χρησιμοποιούν τον υδατάνθρακα αυτόν ως μέσο ενέργειας και ως μέσο μεταβολισμού [3]. Ειδικά τα εγκεφαλικά κύτταρα, το κεντρικό νευρικό

σύστημα και τα ερυθρά αιμοσφαίρια βασίζονται αποκλειστικά στην γλυκόζη για την κάλυψη των ενεργειακών τους αναγκών υπό φυσιολογικές συνθήκες [4].

Η συγκέντρωση της γλυκόζης στο πλάσμα διατηρείται σταθερή λόγω της ισορροπίας μεταξύ του ρυθμού εισαγωγής και ρυθμού απομάκρυνσης της από το κυκλοφορικό (**Σχήμα 1.1**). Η κυκλοφορία της γλυκόζης στον οργανισμό προέρχεται από τρεις πηγές: από το έντερο κατά τη λήψη γευμάτων, από την διαδικασία της γλυκογονόλυσης και από την διαδικασία της γλυκονεογένεσης, που λαμβάνουν χώρα στο ήπαρ. Η γρήγορη εμφάνιση της γλυκόζης στο κυκλοφορικό οφείλεται στον ρυθμό κένωσης του στομάχου μετά την λήψη γευμάτων. Όπως προαναφέρθηκε, άλλες πηγές γλυκόζης προέρχονται από την γλυκογονόλυση, κατά την οποία διασπάται το γλυκογόνο και απελευθερώνεται γλυκόζη και την γλυκονεογένεση, όπου η γλυκόζη σχηματίζεται και αποθηκεύεται στο ήπαρ υπό την μορφή γλυκογόνου κυρίως από γλυκερόλη, γαλακτικό οξύ και αμινοξέα για χρήση σε περίοδο νηστείας [5].

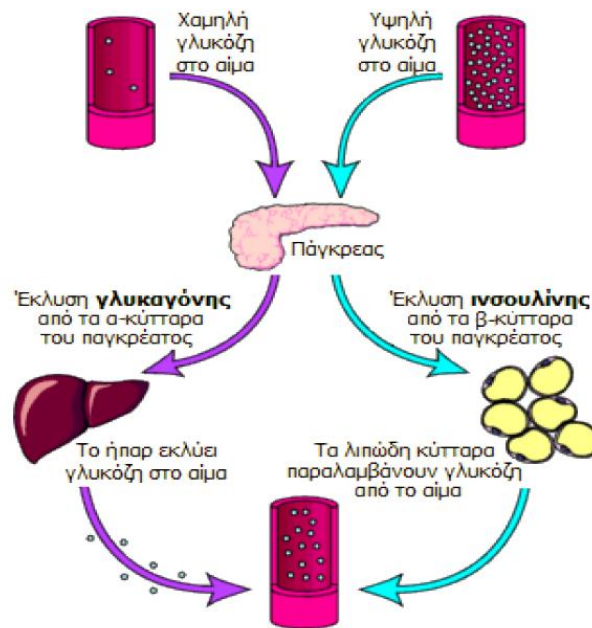
1.1.1 Ορμόνες: Ινσουλίνη και Γλυκαγόνη

Τα αποθέματα γλυκόζης αποθηκεύονται ως πολυμερές γλυκογόνο (γλυκογένεση) το οποίο υπάρχει σε υψηλότερες συγκεντρώσεις στο ήπαρ και στους μύες. Η ρύθμιση του γλυκογόνου και επομένως της γλυκόζης ελέγχεται μέσω των πεπτιδικών ορμονών ινσουλίνης και γλυκαγόνης που παράγονται στο πάγκρεας. Πρόκειται για ενδοκρινείς ορμόνες δηλαδή ορμόνες που εκκρίνονται απ' ευθείας στην κυκλοφορία του αίματος για να δράσουν. Οι δύο ορμόνες έχουν ακριβώς αντίστροφο μηχανισμό δράσης, γεγονός στο οποίο οφείλεται και η ρυθμιστική τους ικανότητα. Υπάρχει, λοιπόν, μια ισορροπία μεταξύ των δύο ορμονών ανάλογα την μεταβολική διεργασία που έχουν να επιτελέσουν τα κύτταρα (νηστεία ή ενεργειακά υψηλή κατάσταση). Η ισορροπία αυτή σε φυσιολογικές συνθήκες είναι υπεύθυνη για την εύρυθμη λειτουργία του μεταβολισμού της γλυκόζης [5][6]. Στο **Σχήμα 1.1** φαίνεται η σχηματική αναπαράσταση του μηχανισμού διατήρησης της σταθερής συγκέντρωσης των επιπέδων γλυκόζης στο αίμα. Σε περίπτωση αύξησης των επιπέδων του σακχάρου στο αίμα εκκρίνεται ινσουλίνη από το πάγκρεας με σκοπό την απομάκρυνση της γλυκόζης από το αίμα και την απορρόφηση της από τους ιστούς. Αντίθετα, σε περίπτωση χαμηλών επιπέδων γλυκόζης στο αίμα το πάγκρεας εκκρίνει γλυκαγόνη με σκοπό την παραγωγή γλυκόζης στο ήπαρ.

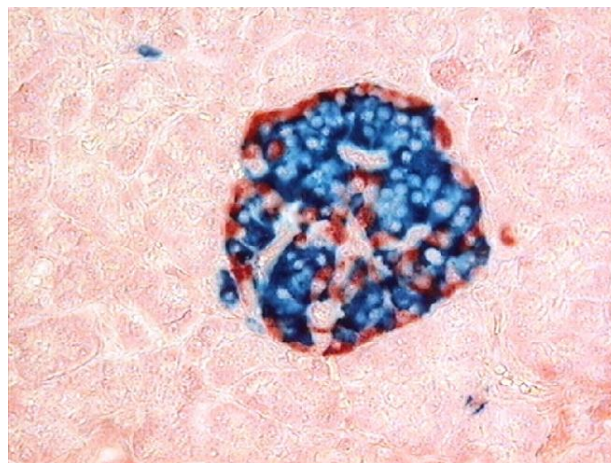
Πάγκρεας

Το πάγκρεας είναι ένας αδένας του πεπτικού συστήματος με ενδοκρινή και εξωκρινή μοίρα ο οποίος παράγει πλήθος σημαντικών ορμονών με κυριότερες την ινσουλίνη και την γλυκαγόνη. Παίξει ρόλο κλειδί στο πεπτικό σύστημα αφού το παγκρεατικό υγρό που εκκρίνει (εξωκρινής αδένας) παρέχει ένζυμα που συμμετέχουν στην διάσπαση υδατανθράκων, λιπιδίων και πρωτεϊνών και συνεπώς στην απορρόφηση θρεπτικών συστατικών από το λεπτό έντερο. Η ενδοκρινής μοίρα του παγκρέατος αποτελείται από ομάδες κυττάρων που ονομάζονται νησίδια του

Langerhans. Τα νησίδια αποτελούνται από κύτταρα καθένα το οποίο παράγει διαφορετικό εκκριτικό προϊόν. Τα β-κύτταρα που εκκρίνουν την ινσουλίνη βρίσκονται σε μεγαλύτερο αριθμό στην κεντρική περιοχή ενώ τα α-κύτταρα που εκκρίνουν την γλυκαγόνη βρίσκονται κυρίως στην περιφέρεια των νησιδίων (**Σχήμα 1.2**). Τα εκκριτικά προϊόντα των νησιδίων περνούν πρώτα από το ήπαρ πριν εισέλθουν στην αιματική κυκλοφορία [7].



Σχήμα 1.1: Μηχανισμός διατήρησης σταθερής συγκέντρωσης της γλυκόζης στο αίμα [9]



Σχήμα 1.2: Νήσος του Langerhans. Με φθορίζοντα αντισώματα φωτίζονται οι διαφορετικοί τύποι κυττάρων. Τα α-κύτταρα που εκκρίνουν γλυκαγόνη φαίνονται στην περιφέρεια (κόκκινα) και τα β-κύτταρα που εκκρίνουν ινσουλίνη αποτελούν την πλειονότητα των κυττάρων και φαίνονται κεντρική θέση (μπλε) [7]

Ινσουλίνη

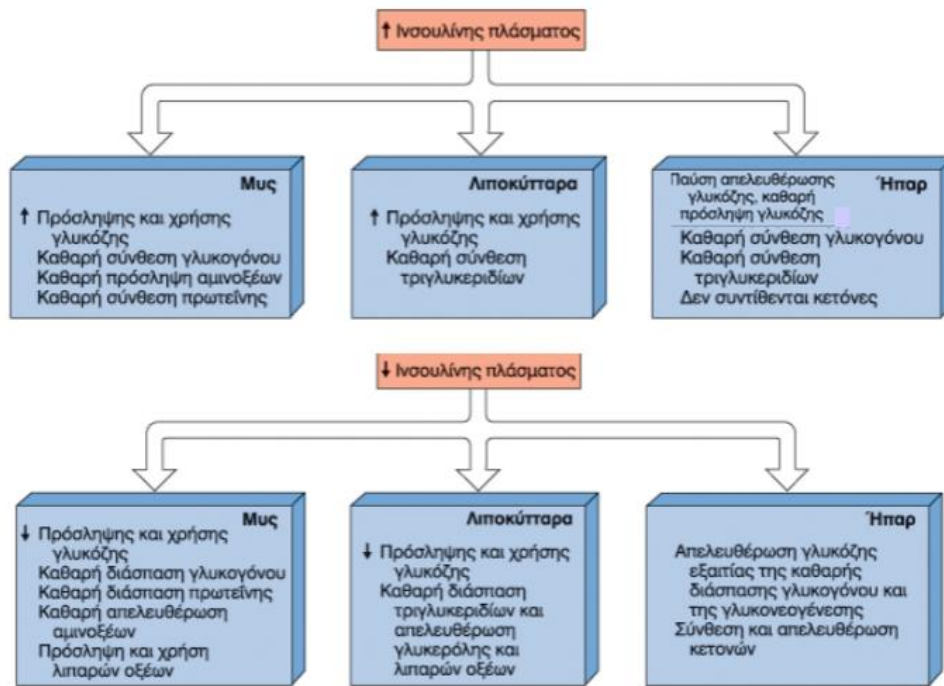
Τα β-κύτταρα του παγκρέατος, που αποτελούν το 60-80% των κυττάρων των νησιδίων του Langerhans, διεγείρονται και εκκρίνουν την ινσουλίνη. Η ινσουλίνη είναι ορμόνη που αποτελείται από 51 αμινοξέα και είναι υπεύθυνη κατά κύριο λόγο για την χρησιμοποίηση της γλυκόζης του αίματος από τα κύτταρα και τους ιστούς, αλλά και για άλλες δράσεις στο ήπαρ και στους μύες. Η δράση της είναι αναβολική, αφού αποσκοπεί στην σύνθεση πρωτεϊνών και λίπους αλλά και στην γενικότερη αποθήκευση ενέργειας [9]. Τα μεταγευματικά αυξημένα επίπεδα γλυκόζης στο αίμα είναι ίσως το ισχυρότερο ερέθισμα για την έκκριση της ινσουλίνης, ωστόσο διεγέρτες μπορεί να ναι συγκεντρώσεις αμινοξέων στο πλάσμα ή ορμόνες που εκκρίνονται από το έντερο κατά την λήψη του γεύματος, όπως οι ορμόνες GLP-1 και GIP [5]. Η έκκριση της ινσουλίνης γίνεται με απώτερο σκοπό την απομάκρυνση της γλυκόζης από το αίμα. Περιλαμβάνει μια ταχεία πρώτη φάση απελευθέρωσης πρόδρομης ινσουλίνης με διάρκεια 10 λεπτών, ακολουθούμενη από παρατεταμένη σύνθεση ινσουλίνης για όσο τα επίπεδα της γλυκόζης παραμένουν ανεβασμένα [2]. Οι μεταβολικές δράσεις της ινσουλίνης στους ιστούς εξηγούνται παρακάτω και αφορούν την ρύθμιση των επιπέδων γλυκόζης, λιπαρών οξέων και αμινοξέων σε περίπτωση που εκκρίνεται φυσιολογικά ινσουλίνη από το πάγκρεας (*Σχήμα 1.3*).

Οι μύες είναι ινσουλοευαίσθητοι ιστοί. Χωρίς την δράση της ινσουλίνης δεν θα λάμβαναν την γλυκόζη που τους είναι απαραίτητη για τις μεταβολικές τους ανάγκες. Στους μύες η ινσουλίνη συντελεί στην απορρόφηση της γλυκόζης, μέσω των υποδοχέων της στην επιφάνεια τους, με σκοπό την οξείδωση της (γλυκόλυση) για απελευθέρωση ενέργειας ή την σύνθεση γλυκογόνου για αποθήκευση ενέργειας [8][9].

Το ήπαρ είναι ο ιστός στόχος και η κύρια περιοχή αποδόμησης της ινσουλίνης. Εκεί ενεργοποιεί την πρόσληψη γλυκόζης και την αποθήκευση της υπό την μορφή γλυκογόνου (αποθήκη ενέργειας σε περίπτωση νηστείας), δηλαδή ενισχύει την γλυκογένεση. Ταυτόχρονα αναστέλλει την έκκριση γλυκαγόνης από τα α-κύτταρα του παγκρέατος εμποδίζοντας τις διαδικασίες της γλυκογονόλυσης και της γλυκονεογένεσης. Κατ' αυτόν τον τρόπο αναστέλλεται η ενδογενής παραγωγή γλυκόζης [8][10]. Ακόμη, αναστέλλει την σύνθεση κετονών δηλαδή την οξείδωση ελευθέρων λιπαρών οξέων.

Η δράση της ινσουλίνης στον λιπώδη ιστό είναι εξίσου σημαντική. Η ινσουλίνη προωθεί την λιπογένεση στα λιποκύτταρα, στα μυϊκά και στα ηπατικά κύτταρα, δηλαδή την σύνθεση των λιπαρών οξέων και την σύνθεση τριγλυκεριδίων. Παράλληλα, αναστέλλει την λιπόλυση, δηλαδή την διάσπαση των τριγλυκεριδίων και την απελευθέρωση λιπαρών οξέων και γλυκερόλης [4][8].

Επιπρόσθετα, η ινσουλίνη συμβάλει και στον μεταβολισμό των πρωτεϊνών, αφού ενισχύοντας την μεταφορά αμινοξέων συντελεί στην πρωτεϊνοσύνθεση στους ιστούς, ενώ αναστέλλει την διάσπαση τους.



Σχήμα 1.3: Μεταβολικές δράσεις της ινσουλίνης στον μυϊκό, λιπώδη και ηπατικό ιστό [8]

Γλυκαγόνη

Η γλυκαγόνη είναι πεπτιδική ορμόνη 29 αμινοξέων που παράγεται από τα α-κύτταρα του παγκρέατος και η δράση της στο μηχανισμό ρύθμισης της γλυκόζης είναι ακριβώς αντίθετη από εκείνη της ινσουλίνης. Η πτώση των επιπέδων γλυκόζης στο αίμα (υπογλυκαιμία) είναι ένα από τα πιο ισχυρά ερεθίσματα έκκρισης γλυκαγόνης. Η υπογλυκαιμία καταστέλλει την έκκριση ινσουλίνης από β-κύτταρα και διεγείρει την έκκριση γλυκαγόνης από τα α-κύτταρα των παγκρεατικών νησιδίων, ομαλοποιώντας τα επίπεδα γλυκόζης στο αίμα. Επίσης, η γλυκαγόνη παίζει σημαντικό ρόλο στην αύξηση των επιπέδων γλυκόζης σε περίοδο νηστείας, έντονης μυϊκής άσκησης ή μετά από γεύματα πλούσια σε πρωτεΐνες και αμινοξέα. Πρόκειται για καταβολική ορμόνη, που αποσκοπεί στην κάλυψη ενεργειακών αναγκών παρέχοντας στην κυκλοφορία γλυκόζη και λιπαρά οξέα.

Η γλυκαγόνη δρα αποκλειστικά στο ήπαρ. Πρακτικά, είναι εκείνη που ενεργοποιεί την ενδογενή (ηπατική) παραγωγή γλυκόζης. Διεγείρει τόσο την γλυκογονόλυση για την διάσπαση του αποθηκευμένου σε ήπαρ (κυρίως) και σε μύες γλυκογόνου σε γλυκόζη, όσο και τη γλυκονογένεση αυξάνοντας την παραγωγή γλυκόζης σε λίγα λεπτά. Υπό ορισμένες συνθήκες, η γλυκαγόνη μπορεί να διεγείρει την παραγωγή κετονικών σωμάτων στο ήπαρ, τα οποία κατά τη διάρκεια της νηστείας ή της παρατεταμένης υπογλυκαιμίας μπορεί να υποκαταστήσουν την γλυκόζη για την κάλυψη ενεργειακών αναγκών του εγκεφάλου [11][12]. Επιπλέον, προωθεί την λιπόλυση, που έχει ως αποτέλεσμα την απελευθέρωση λιπαρών οξέων και γλυκερόλης στην κυκλοφορία και την παραγωγή ενέργειας. Η οποιαδήποτε αύξηση της γλυκόζης στο αίμα αναστέλλει την δράση της γλυκαγόνης. Σε τέτοιες περιπτώσεις

η ανεπαρκής αναστολή της έκκρισης γλυκαγόνης μπορεί να οδηγήσει σε υπεργλυκαιμία.

Άλλες ορμόνες

Υπάρχουν και άλλες ορμόνες που συμβάλουν στον μηχανισμό ομοιόστασης της γλυκόζης, όπως η αμυλίνη και κάποιες γαστρεντερικές ορμόνες όπως τα ινκρετινικά πεπτίδια GIP, GLP-1. Η αμυλίνη εκκρίνεται και αυτή από τα β-κύτταρα του παγκρέατος και συμπληρώνει την δράση της ινσουλίνης μέσω δύο μηχανισμών. Καταστέλλει την έκκριση μεταγευματικής γλυκαγόνης μειώνοντας την ηπατική γλυκόζη. Με αυτόν τον τρόπο αποτρέπει την υπεργλυκαιμία και επιβραδύνει το ρυθμό κένωσης του γαστρικού συστήματος. Οι ορμόνες GLP-1, GIP εκκρίνονται από τα L-κύτταρα και τα K-κύτταρα του λεπτού εντέρου μετά από γεύματα. Οι συγκεντρώσεις τους σε περιόδους νηστείας είναι χαμηλές. Το πεπτίδιο GIP διεγείρει την έκκριση ινσουλίνης σε συνθήκες υπεργλυκαιμίας και συνδράμει στον μεταβολισμό των λιπιδίων, αλλά δεν επιδρά στην διέγερση για την αναστολή της έκκρισης γλυκαγόνης ή την εκκένωση του στομάχου. Από την άλλη, το πεπτίδιο GLP-1 διεγείρει την έκκριση ινσουλίνης και αναστέλλει την μεταγευματική δράση της γλυκαγόνης όταν η συγκέντρωση γλυκόζης στο πλάσμα είναι υψηλή. Επιπρόσθετα, το GLP-1 καθυστερεί την γαστρική κένωση ελέγχοντας έτσι τα μεταγευματικά επίπεδα γλυκόζης [2][5].

Από τα παραπάνω συνάγεται το συμπέρασμα ότι ο μηχανισμός της ομοιόστασης της γλυκόζης είναι πολύπλοκος και απαιτεί την συντονισμένη δράση πολλών ορμονών με βασικότερων αυτών της ινσουλίνης και της γλυκαγόνης. Σε άτομα που πάσχουν από ΣΔ ο ρυθμιστικός μηχανισμός της γλυκόζης υπολειτουργεί και συνεπώς επιφέρει επιπτώσεις στην υγεία του ατόμου. Για παράδειγμα, σε άτομα που πάσχουν από ΣΔ υπάρχει σοβαρή έλλειψη ινσουλίνης. Τότε η γλυκόζη που κυκλοφορεί στο αίμα αδυνατεί να εισέλθει στα κύτταρα και ο οργανισμός αντιλαμβάνεται την κατάσταση αυτή ως νηστεία και αντιδρά με έκκριση ορμονών, όπως η γλυκαγόνη, οι οποίες αυξάνουν κι άλλο τη συγκέντρωση του σακχάρου στο αίμα, μέσω της παραγωγής γλυκόζης από το ήπαρ. Επιπλέον, οι ορμόνες αυτές αφού προάγουν και τη λιπόλυση, παράγουν λιπαρά οξέα τα οποία μέσω της κυκλοφορίας του αίματος κατευθύνονται προς το ήπαρ. Εκεί, και υπό συνθήκες απόλυτης έλλειψης ινσουλίνης και άμεσης ανάγκης για παραγωγή ενέργειας, μετατρέπονται σε κετόνες, οι οποίες αποτελούν μία εναλλακτική πηγή ενέργειας για τα κύτταρα και το μοναδικό υπόστρωμα, εκτός της γλυκόζης, που μπορεί να χρησιμοποιηθεί από τον εγκέφαλο. Ωστόσο, η αυξημένη συγκέντρωσή τους στο αίμα οδηγεί σε κετοξέωση που έχει σοβαρές επιπλοκές για την υγεία του ατόμου.

1.2 Σακχαρώδης Διαβήτης

Όπως έχει αναφερθεί και παραπάνω ο ΣΔ είναι μια παθολογική κατάσταση που προκαλείται λόγω της διαταραχής του μηχανισμού ρύθμισης των επιπέδων της γλυκόζης στο αίμα. Οι μεταβολικές διαταραχές στους υδατάνθρακες, τα λιπίδια και

τις πρωτεΐνες που παρατηρούνται στον ΣΔ οφείλονται στη σημασία της ινσουλίνης ως αναβολικής ορμόνης. Συγκεκριμένα, η ανεπαρκής έκκριση ινσουλίνης από τα β-κύτταρα του παγκρέατος ή/και μειωμένη ευαισθησία των κυττάρων στην ινσουλίνη ευθύνονται για αυτές τις διαταραχές [13][14].

1.2.1 Ταξινόμηση του Σακχαρώδους Διαβήτη

Η ταξινόμηση του ΣΔ εξαρτάται από τις συνθήκες που επικρατούν κατά τη στιγμή της διάγνωσης, ενώ πολλές φορές δεν είναι δυνατή η κατάταξη της περίπτωσης σε μια μόνο κατηγορία. Αυτό όμως που είναι σημαντικό για τον ασθενή και τον ειδικό ιατρό δεν είναι η ταξινόμηση του ΣΔ αλλά η κατανόηση της παθογένειας της υπεργλυκαιμίας. Ο βαθμός της υπεργλυκαιμίας και που αυτή οφείλεται αντανακλά τη βαρύτητα της εκάστοτε υποκείμενης παθολογικής διεργασίας και συνεπώς της αποτελεσματικότητας της θεραπείας που θα επιλεγεί για τον ασθενή. Οι πιο διαδεδομένοι τύποι του ΣΔ παρουσιάζονται παρακάτω.

I. Σακχαρώδης Διαβήτης Τύπου 1 - ΣΔΤ1 (ινσουλοεξαρτώμενος ή νεανικός διαβήτης)

α) Αυτοάνοσος ΣΔ (τύπος 1Α)

Ο τύπος αυτός αφορά μόνο το 5-10% των ασθενών με ΣΔ και κατά κανόνα εμφανίζεται συνήθως σε παιδική και εφηβική ηλικία, χωρίς ωστόσο να αποκλείεται η εμφάνιση του σε οποιαδήποτε ηλικία. Προκαλείται από αυτοάνοση καταστροφή των β-κυττάρων του παγκρέατος που έχουν ως αποτέλεσμα την εμφάνιση αυξημένων επιπέδων γλυκόζης λόγω της ολικής έλλειψης ή ελάχιστης έκκρισης ινσουλίνης στον οργανισμό. Δείκτες της αυτοάνοσης φύσης της νόσου αποτελούν τα αντιησιδιακά αυτοαντισώματα που παρατηρούνται στους ασθενείς της νόσου και αντιτίθενται στην σύνθεση της ινσουλίνης. Η αυτοάνοση καταστροφή αποδίδεται κυρίως σε πολλαπλούς γενετικούς προδιαθεσικούς παράγοντες. Σε μερικούς ασθενείς, κυρίως σε παιδιά και εφήβους, πρώτη κλινική εκδήλωση της ασθένειας είναι η κετοξέωση. Σε άλλους ασθενείς εκδηλώνεται μέτρια υπεργλυκαιμία νηστείας που μπορεί να εξελιχθεί, στο πλαίσιο συνυπάρχουσας λοίμωξης ή άλλου στρεσογόνου παράγοντα, σε σοβαρή υπεργλυκαιμία ή σε κετοξέωση [13][14].

β) Ιδιοπαθής ΣΔ (τύπος 1Β)

Σύμφωνα με τον Αμερικανικό Διαβητικό Οργανισμό (American Diabetes Association, ADA) υπάρχει μια σπάνια μορφή ΣΔΤ1 άγνωστης προέλευσης, λιγότερο σοβαρή νόσος από τον αυτοάνοσο ΣΔΤ1 χωρίς αυτοάνοσο υπόβαθρο. Οι περισσότεροι ασθενείς με αυτόν τον τύπο είναι αφρικάνικης ή ασιατικής καταγωγής και πάσχουν από διάφορους βαθμούς ινσουλοανεπάρκειας και εμφάνιση επεισοδίων κετοξέωσης.

Θεραπεία

Τα άτομα με ΣΔΤ1 απαιτούν εντατική θεραπεία, που χαρακτηρίζεται από εξωγενή χορήγηση ινσουλίνης μέσω πολλαπλών ημερήσιων ενέσεων με

ινσουλίνη ταχείας δράσης μεταγευματικά ή με συνεχή υποδόρια έγχυση ινσουλίνης μέσω αντλιών. Ο σκοπός του γλυκαιμικού ελέγχου μέσω της χορήγησης ινσουλίνης είναι η διατήρηση των επιπέδων γλυκόζης όσο το δυνατόν πιο κοντά στα φυσιολογικά, αποφεύγοντας ταυτόχρονα την υπογλυκαιμία. Η επίτευξη του στόχου αυτού απαιτεί την εξατομίκευση της θεραπείας από ασθενή σε ασθενή [15].

II. Σακχαρώδης Διαβήτης Τύπου 2 - ΣΔΤ2 (μη ινσουλινοεξαρτώμενος διαβήτης)

Αποτελεί τον πιο διαδεδομένο τύπο ΣΔ αφού το 90-95% των διαβητικών ατόμων πάσχουν από αυτόν τον τύπο. Ο ΣΔ τύπου 2 (ΣΔΤ2) χαρακτηρίζεται από αντίσταση στη δράση της ινσουλίνης και συνοδεύεται από σχετική αλλά όχι πλήρη ανεπάρκεια έκκρισης ινσουλίνης. Παλαιότερα, ονομαζόταν ως μη ινσουλινεξαρτώμενος, διότι τα άτομα με ΣΔΤ2 δεν χρειάζονται θεραπεία ινσουλίνης για να επιβιώσουν. Αν και τα ακριβή αίτια της εμφάνισης ΣΔΤ2 παραμένουν αδιευκρίνιστα, είναι δεδομένο ότι δεν παρατηρείται αυτοάνοση καταστροφή των β-κυττάρων των νησιδίων Langerhans. Πρακτικά, ο οργανισμός, λόγω της μειωμένης ευαισθησίας των κυττάρων στην ινσουλίνη, δεν μπορεί να αξιοποιήσει την φυσιολογικά παραγόμενη ινσουλίνη για την απομάκρυνση της γλυκόζης από το αίμα, με αποτέλεσμα να αναπτύσσεται υπεργλυκαιμία, η οποία με την σειρά της δεν προκαλεί την έκκριση περισσότερης ινσουλίνης.

Η ηλικία, η παχυσαρκία, η έλλειψη σωματικής άσκησης και καλής διατροφής αυξάνουν τον κίνδυνο εμφάνισης ΣΔΤ2. Αυτή η μορφή διαβήτη συχνά δεν διαγιγνώσκεται για αρκετά χρόνια επειδή η υπεργλυκαιμία αναπτύσσεται σταδιακά και δεν είναι αρχικά αρκετά σοβαρή για να τον ασθενή για να παρατηρήσει και άλλα συμπτώματα [14].

Θεραπεία

Η αντίσταση στην ινσουλίνη και η υπεργλυκαιμία που αυτή προκαλεί μπορεί να αντιμετωπιστεί είτε με ελάττωση του σωματικού βάρους και καλή διατροφή είτε με φαρμακευτική αγωγή [14]. Ωστόσο, αν και ο ΣΔΤ2 χαρακτηρίζεται ως μη ινσουλινεξαρτώμενος έχει παρατηρηθεί ότι η ινσουλινθεραπεία στην περίπτωση του ΣΔΤ2 μπορεί να έχει ευεργετικά αποτελέσματα στην επιβίωση των β-κυττάρων και στην ευαισθησία στην ινσουλίνη για τον μακροχρόνιο διαβήτη τύπου 2 [16][17].

III. Σακχαρώδης Διαβήτης κύησης

Ο ΣΔ κύησης εμφανίζεται για πρώτη φορά κατά την έναρξη ή κατά την διάρκεια της εγκυμοσύνης και συνήθως υποχωρεί μετά τον τοκετό. Όμοια με τον ΣΔΤ2 που περιγράφηκε παραπάνω, ο ΣΔ κύησης παρουσιάζει ελαττωμένη έκκριση ινσουλίνης και μειωμένη ευαισθησία των κυττάρων στην ινσουλίνη. Μπορεί να προκαλέσει επιπλοκές σε μητέρα και έμβryo-νεογνό. Ακόμη, οι μητέρες με διαβήτη κύησης και τα μωρά που γεννιούνται από τέτοιες μητέρες

έχουν αυξημένο κίνδυνο εμφάνισης διαβήτη αργότερα στη ζωή τους. Οι παράγοντες κινδύνου για εμφάνιση διαβήτη κύησης περιλαμβάνουν παχυσαρκία, οικογενειακό ιστορικό ΣΔ, ηλικία μητέρας, σύνδρομο πολυκυστικών ωοθηκών κ.α. [13][14].

Θεραπεία

Οι θεραπείες για ΣΔ κύησης στοχεύουν την διατήρηση των επιπέδων γλυκόζης στο αίμα κοντά με εκείνα των υγίων εγκύων γυναικών. Η θεραπεία περιλαμβάνει πάντα ειδικά προγράμματα διατροφής και σωματική δραστηριότητα. Επίσης, περιλαμβάνει καθημερινές εξετάσεις για την παρακολούθηση του σακχάρου στο αίμα όπως και εξωγενής χορήγηση ινσουλίνης με ενέσιμες δόσεις [18].

1.2.2 Διαγνωστικά Κριτήρια Σακχαρώδους Διαβήτη

Ο ΣΔ διαγιγνώσκεται κυρίως μέσω μετρήσεων του σακχάρου του αίματος. Ο κλινικός έλεγχος με εξετάσεις σε συνδυασμό με παρουσίαση συνήθων συμπτωμάτων του ΣΔ όπως πολουρία, πολυδιψία, απώλεια βάρους, θόλωση όρασης μπορούν να αποφανθούν για νόσηση από ΣΔ και να προβούν σε έγκαιρη διάγνωση [14]. Όταν απουσιάζουν τα κλασικά συμπτώματα της υπεργλυκαιμίας, οι εξετάσεις μέτρησης πρέπει να επαναλαμβάνονται σε διαφορετική μέρα και να επανεξετάζεται αν τηρούνται τα κριτήρια για διάγνωση ΣΔ.

- **Μέτρηση γλυκοζυλιωμένης αιμοσφαιρίνης (HbA1c)**

Η γλυκοζυλιωμένη αιμοσφαιρίνη (HbA1c ή A1C) είναι μια εργαστηριακή εξέταση η οποία μετράται από απλή αιμοληψία από φλεβικό αίμα ή με εξέταση τριχοειδικού αίματος. Η εξέταση αυτή καθορίζει το μέσο όρο των επιπέδων γλυκόζης στο αίμα τους τελευταίους 2-3 μήνες και δεν επηρεάζεται από βραχυπρόθεσμες διακυμάνσεις των επιπέδων σακχάρου. Έτσι, δεν είναι δυνατόν να αξιοποιηθεί για την παρακολούθηση των οξέως μεταβαλλόμενων επιπέδων σακχάρου, π.χ για τη ρύθμιση της δόσης της ινσουλίνης ή για πρόβλεψη υπογλυκαιμικών επεισοδίων και άρα ενδείκνυται μόνο για τη διάγνωση του ΣΔΤ2. Οι φυσιολογικές τιμές της HbA1c πρέπει να κυμαίνονται σε εύρος 4-6%. Συγκεκριμένα, τιμές μεγαλύτερες του 6.5% υποδηλώνουν ύπαρξη ΣΔ, ενώ τιμές μεταξύ 5.7-6.4% υποδηλώνουν ότι ο ασθενής διατρέχει κίνδυνο εμφάνισης ΣΔ.

- **Μέτρηση της γλυκόζης πλάσματος νηστείας (Fasting Plasma Glucose, FPG)**

Πρόκειται για αιματολογική εξέταση κατά την οποία ο ασθενής πρέπει να είναι νηστικός για 8 ώρες πριν την μέτρηση. Τιμές FPG μεγαλύτερες των 125 mg/dl υποδηλώνουν ΣΔ, ενώ τιμές στο εύρος των 100-125 mg/dl μαρτυρούν την διαταραγμένη γλυκόζη νηστείας (Impaired Fasting Glucose, IFG) και την εμφάνιση προδιαβήτη και άρα αυξημένη πιθανότητα εμφάνισης ΣΔ στο μέλλον.

Πίνακας 1.1: Διαγνωστικά Κριτήρια για μέτρηση FPG

Μέτρηση γλυκόζης πλάσματος νηστείας (mg/dL)	Διάγνωση
< 99	Υγιής
100-125	Προδιαβητικός (IFG)
>126	Διαβητικός

- **Δοκιμασία ανοχής γλυκόζης (Oral Glucose Tolerance Test, OGTT)**

Σύμφωνα με τον Παγκόσμιο Οργανισμό υγείας (ΠΟΥ), η δοκιμασία ανοχής γλυκόζης είναι μια διαδικασία κατά την οποία χορηγείται στον ασθενή ένα υγρό 75 gr υδατανθράκων (άνδρην γλυκόζη διαλυμένη σε νερό) και δύο ώρες μετά την κατανάλωση μετράται η μεταγευματική γλυκόζη πλάσματος. Αν η τιμή είναι μεγαλύτερη ή ίση των 200 mg/dl τότε ο ασθενής διαγιγνώσκεται με ΣΔ. Επίσης, αν η τιμή της γλυκόζης πλάσματος προκύψει εντός του διαστήματος 140-199 mg/dl τότε τα άτομα θεωρούνται ότι έχουν διαταραγμένη ανοχή στην γλυκόζη (Impaired Glucose Tolerance, IGT) που σηματοδοτεί προδιαβήτη.

Πίνακας 1.2: Διαγνωστικά Κριτήρια για μέτρηση OGTT

Μέτρηση μεταγευματικής γλυκόζης πλάσματος μετά από 2 ώρες (mg/dL)	Διάγνωση
< 139	Υγιής
140-199	Προδιαβητικός (IGT)
>200	Διαβητικός

1.2.3 Επιπλοκές του Σακχαρώδους Διαβήτη

Ο ΣΔ έχει σημαντικές βραχυπρόθεσμες και μακροπρόθεσμες συνέπειες στην υγεία του ατόμου. Η καθυστερημένη ή λανθασμένη διαχείριση της νόσου μπορεί να προκαλέσει σοβαρές επιπλοκές στο ανθρώπινο σώμα. Τα σταθερά επίπεδα γλυκόζης στο αίμα μπορούν να οδηγήσουν σε σοβαρές επιπλοκές στην καρδιά, τα αγγεία, τα μάτια, τα νεφρά, τα νεύρα κ.α [19].

Βραχυπρόθεσμες επιπλοκές θεωρούνται οι υπογλυκαιμίες και οι υπεργλυκαιμίες. Η λανθασμένη αντιμετώπιση της νόσου μπορεί προκαλέσει την μείωση των επιπέδων γλυκόζης κάτω από το φυσιολογικό, δηλαδή να προκαλέσει υπογλυκαιμία. Η **υπογλυκαιμία** αν δεν αντιμετωπιστεί άμεσα μπορεί να προκαλέσει ζάλη και απώλεια αισθήσεων αφού η γλυκόζη είναι απαραίτητη για την λειτουργία των εγκεφαλικών διεργασιών. Αντίθετα, η **υπεργλυκαιμία**, που αποτελεί και το κύριο χαρακτηριστικό της νόσου του ΣΔ, μπορεί να οδηγήσει επείγουσες επιπλοκές που χρήζουν άμεσης αντιμετώπισης [14]. Τέτοιες επιπλοκές παρουσιάζονται παρακάτω.

- **Διαβητική κετοξέωση**

Η κετοξέωση είναι μια οξεία, σοβαρή και απειλητική επιπλοκή του ΣΔ. Προκαλείται από την απότομη αύξηση των επιπέδων γλυκόζης στο αίμα, την οποία αδυνατεί να διαχειριστεί ο οργανισμός, λόγω της πλήρους ή σχεδόν πλήρους απουσίας ινσουλίνης (συναντάται κυρίως στον ΣΔΤ1). Χαρακτηρίζεται από κετοναιμία και κετονουρία, δηλαδή παρουσία κετονών στο αίμα και στα ούρα και μεταβολική οξέωση δηλαδή πτώση του pH του αίματος (όξινο αίμα). Η τυπική κλινική εικόνα περιλαμβάνει πολυουρία, πολυδιψία, ναυτία, εμετούς, κοιλιακό άλγος. Χωρίς άμεση ιατρική παρέμβαση, οι ασθενείς μπορούν να οδηγούν σε κώμα ή ακόμα και στον θάνατο.

- **Διαβητικό μη κετωτικό υπερωσμωτικό κώμα**

Αποτελεί επείγουσα επιπλοκή του ΣΔΤ2, χαρακτηρίζεται από βαριά υπεργλυκαιμία χωρίς κέτωση, με υπερόσμωση και βαριά αφυδάτωση. Συνήθως, συμβαίνει σε ηλικιωμένους ασθενείς όπου υπάρχει κάποιο εκλυτικό γεγονός όπως λοίμωξη ή εγκεφαλικό επεισόδιο. Ο ασθενής παρουσιάζεται ληθαργικός και η κατάσταση του μπορεί να εξελιχθεί σε κωματώδη.

Η διακύμανση των επιπέδων γλυκόζης σε μη φυσιολογικό εύρος τιμών είναι ικανή να προκαλέσει με την πάροδο του χρόνου σοβαρές επιπλοκές στην υγεία του ασθενή. Μερικές από τις οποίες παρουσιάζονται στην συνέχεια [20].

- **Διαβητική αμφιβληστροειδοπάθεια**

Πρόκειται για φθορά του αμφιβληστροειδούς χιτώνα στο μάτι που προκαλεί μειωμένη όραση ή και τύφλωση σε περίπτωση μη αντιμετώπισης. Τα σταθερά υψηλά επίπεδα γλυκόζης στο αίμα σε συνδυασμό με υψηλή αρτηριακή πίεση και χοληστερόλη είναι οι κύριες αιτίες της αμφιβληστροειδοπάθειας.

- **Διαβητική νεφροπάθεια**

Προκαλείται από βλάβη στα μικρά αιμοφόρα αγγεία στα νεφρά που έχουν ως συνέπεια να διαταραχθεί η ορθή λειτουργία τους. Η διατήρηση των επιπέδων γλυκόζης σε φυσιολογικά επίπεδα μπορεί να μειώσει σημαντικά τον κίνδυνο νεφρικής νόσου.

- **Διαβητική νευροπάθεια**

Ο διαβήτης μπορεί να προκαλέσει βλάβη τόσο στο περιφερικό όσο και στο αυτόνομο νευρικό σύστημα. Αυτό μπορεί να οδηγήσει σε προβλήματα με την πέψη, την στυτική δυσλειτουργία και άλλες λειτουργίες. Μεταξύ των πιο συχνά προσβεβλημένων περιοχών είναι τα άκρα, ιδίως τα πόδια. Σε αυτές τις περιοχές η νευρική βλάβη ονομάζεται **περιφερική νευροπάθεια** και προκαλεί πόνο, μυρμήγκιασμα και απώλεια αισθήματος. Η απώλεια αισθήματος είναι ιδιαίτερα σημαντική διότι μπορεί να αφήσει τους τραυματισμούς

απαρατήρητους οδηγώντας σε σοβαρές λοιμώξεις και πιθανούς ακρωτηριασμούς (διαβητικό πόδι).

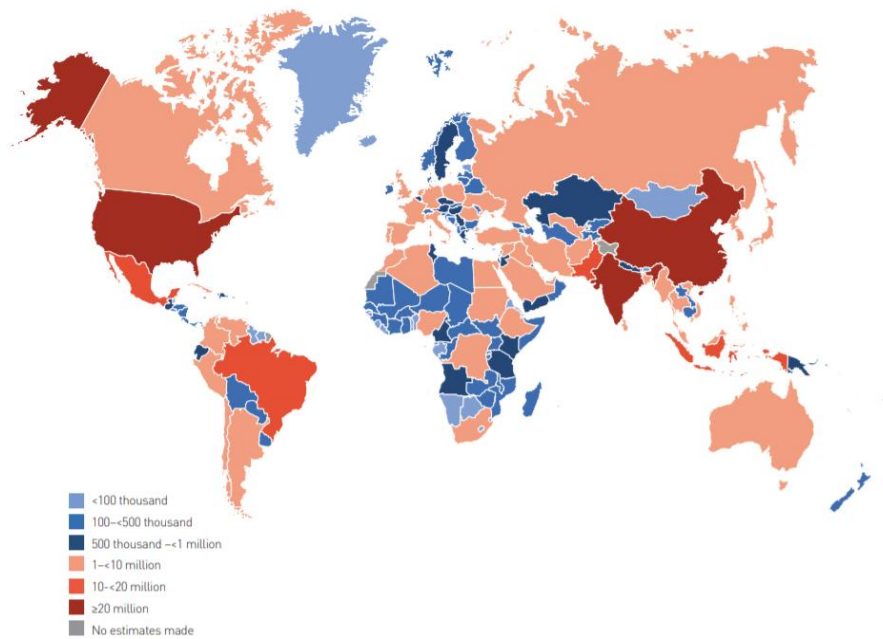
- **Μακροαγγειακές παθήσεις**

Οι μακροαγγειακές παθήσεις σχετίζονται με βλάβες στα αγγεία της καρδιάς, του εγκεφάλου και των περιφερικών αρτηριών. Μπορεί να προκαλέσει θανατηφόρες επιπλοκές όπως **στεφανιαία νόσο** (που οδηγεί σε καρδιακό προσβολή), **εγκεφαλικό επεισόδιο** και **διαλείπουσα χωλότητα** (πόνος στην γάμπα). Η καρδιαγγειακή νόσος είναι η πιο κοινή αιτία θανάτου σε άτομα με ΣΔ [21][22].

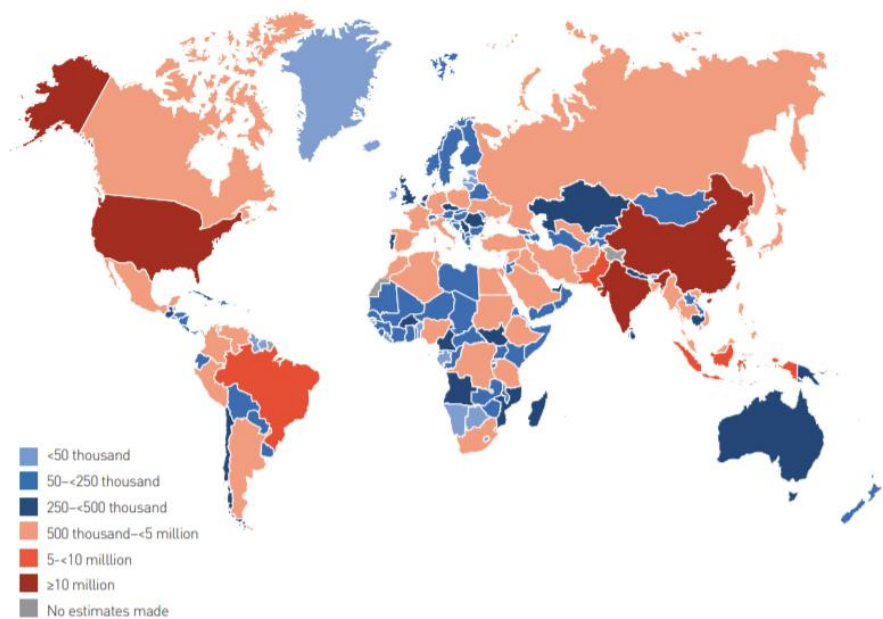
1.2.4 Επιδημιολογικά Στοιχεία

Ο ΣΔ είναι μια νόσος με αλματώδη αύξηση σε παγκόσμιο επίπεδο. Τα αίτια είναι σύνθετα, ωστόσο η έξαρση αποδίδεται σε γενετικούς και προδιαθεσικούς παράγοντες, αλλά κυρίως σε περιβαλλοντικούς όπως η κακή διατροφή, η παχυσαρκία και η έλλειψη σωματικής άσκησης. Ο ΣΔΤ1 είναι πιο σπάνιος ενώ ο ΣΔΤ2 αντιπροσωπεύει το 90% των περιπτώσεων. Σύμφωνα με πρόσφατα στοιχεία της Διεθνούς Ομοσπονδίας Διαβήτη (International Diabetes Federation, IDF) για το έτος 2019 [23], ο αριθμός των ατόμων που πάσχουν από ΣΔ ανέρχεται στα 463 εκατομμύρια ενώ εκτιμάται ότι η "έκρηξη" αυτή θα συνεχιστεί και τα επόμενα χρόνια.

Συγκεκριμένα, εκτιμάται ότι 463 εκατομμύρια ενήλικες ηλικίες 20-79 ετών παγκοσμίως πάσχουν από διαβήτη (9,3% όλων των ενηλίκων σε αυτήν την ηλικιακή ομάδα) εκ των οποίων το 79,4% ζει σε χώρες με χαμηλό και μεσαίο εισόδημα, όπου συχνά στερούνται τα μέσα για αποτελεσματική αντιμετώπιση. Ο ΣΔ αφορά ωστόσο και τις αναπτυσσόμενες χώρες όπου η εμφάνιση της νόσου σχετίζεται με την διαβίωση στα αστικά κέντρα και τον σύγχρονο τρόπο ζωής. Με βάση τις εκτιμήσεις του 2019, προβλέπεται ότι 578,4 εκατομμύρια ενήλικες ηλικίας 20-70 ετών θα πάσχουν από ΣΔ έως το 2030, ενώ μέχρι το 2045 ο αριθμός αυτός θα αγγίζει τα 700,2 εκατομμύρια. Στο χάρτη του **Σχήματος 1.4** παρουσιάζεται η εξάπλωση του ΣΔ κάθε χώρας για το έτος 2019 σύμφωνα με το *Diabetes Atlas* της IDF. Οι πέντε χώρες με το μεγαλύτερο αριθμό παθόντων ήταν η Κίνα (116,4 εκατ.), η Ινδία (77,0 εκατ.), οι Ηνωμένες Πολιτείες (31,0 εκατ.), το Πακιστάν (19,4 εκατ.) και η Βραζιλία (16,8 εκατ.). Το 1970, στην Ελλάδα, το ποσοστό των ατόμων με ΣΔ κυμαίνονταν γύρω στο 2,8%. Σήμερα, ο επιπολασμός (συχνότητα στον γενικό πληθυσμό) της χώρας μας (σε άτομα 20-79 ετών) φτάνει το 8%. Ειδικό επιστήμονες εκτιμούν ότι το ποσοστό αυτό είναι μεγαλύτερο, μπορεί και διπλάσιο, δεδομένου ότι πολλοί άνθρωποι αγνοούν ότι πάσχουν από ΣΔ. Ο χάρτης του **Σχήματος 1.5** απεικονίζει την εξάπλωση του μη διαγνωσμένου ΣΔ ανά χώρα για το έτος 2019.



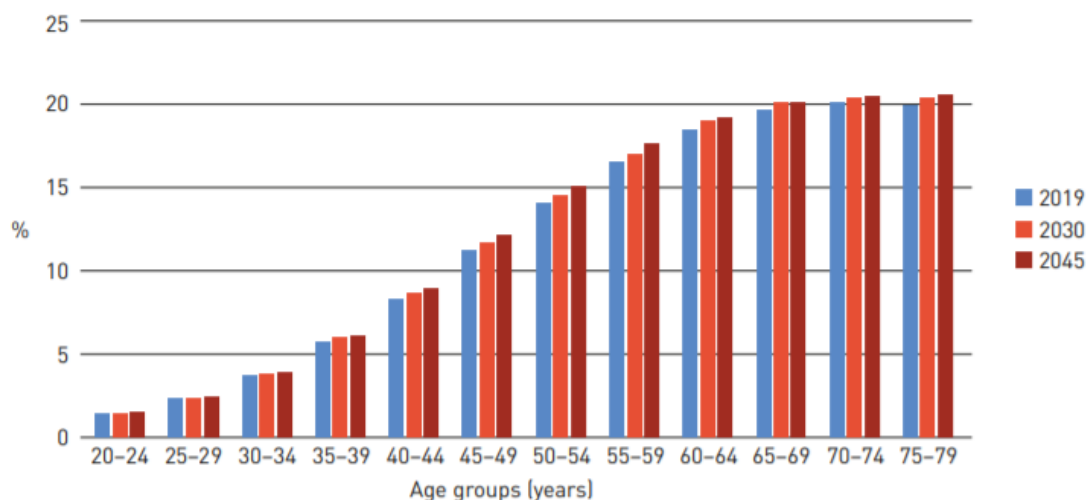
Σχήμα 1.4: Εξάπλωση του ΣΔ για το έτος 2019 ανά γεωγραφική περιοχή [23]



Σχήμα 1.5: Εξάπλωση του μη διαγνωσμένου ΣΔ για το έτος 2019 ανά γεωγραφική περιοχή [23]

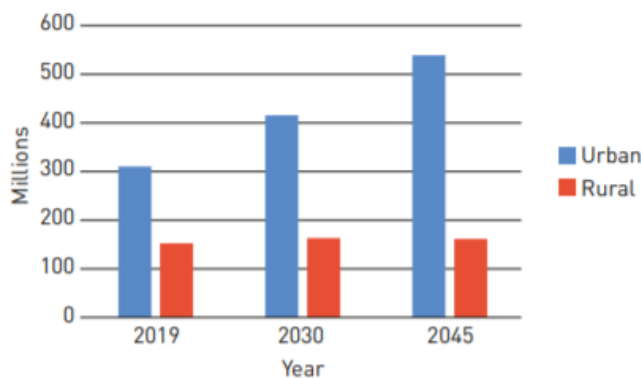
Όσον αφορά την ηλικία εμφάνισης, η εμφάνιση ΣΔ είναι χαμηλότερη σε ενήλικες ηλικίας 20-22 ετών (1,4% το 2019), ενώ το μεγαλύτερο ποσοστό περιπτώσεων ΣΔ (19,9%) παρουσιάζεται σε πληθυσμό ατόμων με ηλικιακό εύρος 75-79 ετών. Η τάση για τα επόμενα έτη είναι αυξητική (**Σχήμα 1.6**). Σχετικά με την εξάπλωση του ΣΔ με βάση το φύλο, οι αριθμοί διαβητικών ανδρών και γυναικών δεν διαφέρουν σημαντικά. Το ποσοστό του ΣΔ σε γυναίκες είναι λίγο χαμηλότερο από ότι στους

άνδρες (9,0% έναντι 9,6%). Το έτος 2019 υπήρχαν περίπου 17,2 εκατομμύρια περισσότεροι άνδρες από γυναίκες με ΣΔ.



Σχήμα 1.6: Εξάπλωση του ΣΔ για τα έτη 2019, 2030, 2045 ανά ηλικιακή ομάδα (έτη) [23]

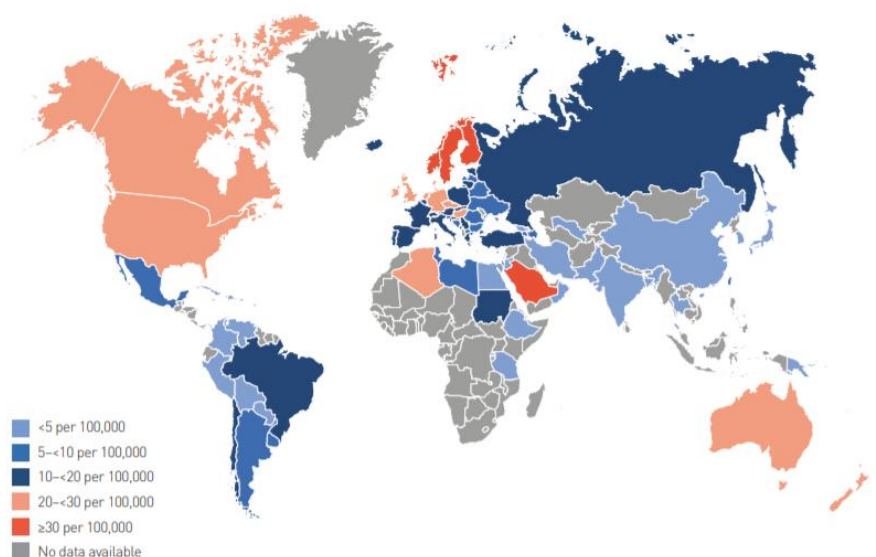
Όπως ήδη αναφέρθηκε, η εξάπλωση του ΣΔ οφείλεται και στην διαβίωση στα αστικά κέντρα, τον τρόπο ζωής και τις διατροφικές συνήθειες. Σύμφωνα με τα επίσημα στοιχεία το 2019 ο επιπολασμός στις αστικές περιοχές είναι 10,8% και στις αγροτικές περιοχές είναι 7,2%. Το 2030 και το 2045 αναμένεται να αυξηθούν τα ποσοστά σε 11,9% και 12,5% αντίστοιχα ως συνέπεια της αστικοποίησης (**Σχήμα 1.7**).



Σχήμα 1.7: Αριθμός ατόμων με ΣΔ (ηλικίας 20-79 ετών) που ζουν σε αστικές και αγροτικές περιοχές το 2019, 2030 και 2045 [23]

Ο αριθμός των παιδιών και των εφήβων με ΣΔ αυξάνεται ετησίως με άνοδο της τάξης του 3%. Ο ΣΔΤ1 είναι η πιο συχνή μορφή παιδικού διαβήτη. Ωστόσο, με την αύξηση των επιπέδων παχυσαρκίας και σωματικής αδράνειας σε νεαρές ηλικίες ο ΣΔΤ2 αυξάνεται με γοργούς ρυθμούς και μπορεί να αποτελέσει παγκόσμιο ζήτημα της δημόσιας υγείας οδηγώντας σε σοβαρά προβλήματα. Σε κάποιες χώρες όπως η

Ιαπωνία, ο ΣΔΤ2 είναι πιο συχνός σε αντίθεση με Ευρωπαϊκές χώρες όπου εμφανίζονται περισσότερα περιστατικά ΣΔΤ1. Εκτιμάται από την IDF ότι συνολικά 1.110.100 παιδιά και έφηβοι νεότεροι των 20 χρόνων παγκοσμίως πάσχουν από ΣΔΤ1. Στον χάρτη του **Σχήματος 1.8** παρουσιάζονται τα ποσοστά εμφάνισης ΣΔΤ1 σε παιδιά και εφήβους κάτω των 15 ετών. Σε χώρες με περιορισμένη πρόσβαση σε ινσουλίνη και ανεπαρκής παροχή υπηρεσιών υγείας, οι πάσχοντες αντιμετωπίζουν σοβαρές επιπλοκές και κατά συνέπεια πρόωρη θνησιμότητα.

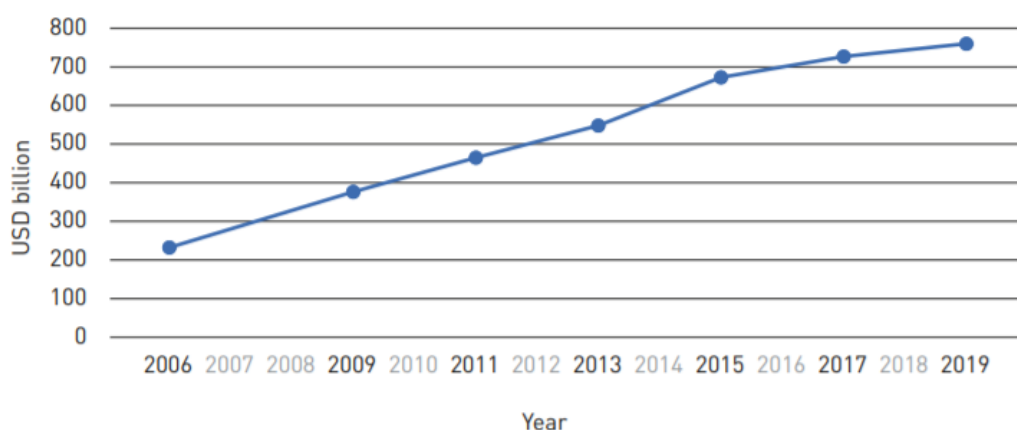


Σχήμα 1.8: Ποσοστά νέων περιπτώσεων ΣΔΤ1 (ανά 100.000 πληθυσμό ετησίως) σε παιδιά ηλικίας 0-14 ετών για το έτος 2019 [23]

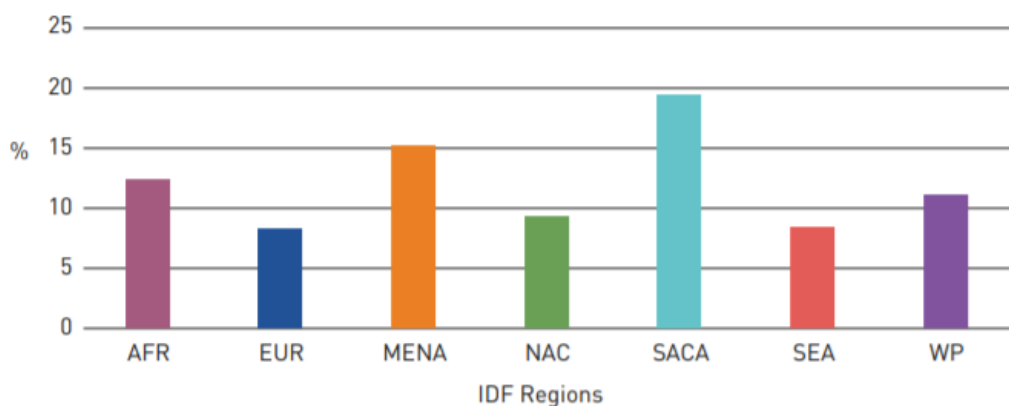
Ο ΣΔ και οι επιπλοκές του αποτελούν κύριες αιτίες πρόωρου θανάτου παγκοσμίως. Τα επίσημα στοιχεία του *Diabetes Atlas 2019* δηλώνουν ότι 4,2 εκατομμύρια ενήλικες ηλικίας 20-79 ετών, πέθαναν από ΣΔ. Αυτό ισοδυναμεί με ένα θάνατο κάθε οκτώ δευτερόλεπτα. Υπολογίζεται ότι ο διαβήτης σχετίζεται με το 11,3% των παγκόσμιων θανάτων αυτών των ηλικιών. Ακόμη, έχει παρατηρηθεί ότι σε παγκόσμιο επίπεδο, υπάρχουν περισσότεροι θάνατοι που σχετίζονται με ΣΔ στις γυναίκες παρά στους άνδρες.

Πέρα από τον αντίκτυπο που έχει ο ΣΔ στην ποιότητα ζωής των ασθενών λόγω των επιπλοκών που προκαλεί, ο διαβήτης επιβάλλει επίσης σημαντικές οικονομικές επιπτώσεις στις χώρες και στο σύστημα υγείας τους. Το 10% της παγκόσμιας δαπάνης για την υγεία αφορά τον διαβήτη. Το κόστος αντιμετώπισης των περιπτώσεων με ΣΔ διαφέρει σημαντικά από χώρα σε χώρα ανάλογα το επίπεδο ανάπτυξης της. Η IDF εκτιμά ότι οι υγειονομικές δαπάνες σχετικές με τον ΣΔ φτάνουν τα 760 δισεκατομμύρια δολάρια. Προβλέπεται ότι οι δαπάνες θα συνεχίσουν να αυξάνονται, με αποτέλεσμα το 2030 να ανέρχονται στα 825 δισεκατομμύρια δολάρια και τα 845 δισεκατομμύρια δολάρια έως το 2045. Το **Σχήμα 1.9** απεικονίζει την ανοδική τάση των συνολικών δαπανών που σχετίζονται με την αντιμετώπιση του ΣΔ από το 2006 μέχρι και το 2019 σύμφωνα με επίσημα στοιχεία της IDF. Στο **Σχήμα**

1.10 φαίνονται τα ποσοστά δαπανών υγείας που σχετίζονται με ΣΔ ως προς το σύνολο των υγειονομικών δαπανών για διάφορες IDF περιοχές, για το έτος 2019. Παρατηρείται ότι η Νότια και Κεντρική Αμερική είχε το 2019 το υψηλότερο ποσοστό δαπανών υγείας για τον ΣΔ. Ακολουθούν οι χώρες της Μέσης Ανατολής και της Βόρειας Αφρικής, ενώ στην τελευταία θέση βρέθηκε η Ευρώπη.



Σχήμα 1.9: Συνολικές υγειονομικές δαπάνες που σχετίζονται με τον ΣΔ για ενήλικες ηλικίας 20-70 ετών που πάσχουν από ΣΔ [23]



Σχήμα 1.10: Συνολικές δαπάνες υγείας που σχετίζονται με τον ΣΔ για ενήλικες ηλικίας 20-70 ετών που πάσχουν από ΣΔ (IDF περιοχές) [23]

2ο Κεφάλαιο

Γλυκαιμικός Έλεγχος

Εισαγωγή

Για την αντιμετώπιση της νόσου του Σακχαρώδους Διαβήτη (ΣΔ) και την πρόληψη των σοβαρών επιπλοκών που προκαλεί, κρίνεται ζωτικής σημασίας ο τακτικός γλυκαιμικός έλεγχος. Ο γλυκαιμικός έλεγχος περιλαμβάνει την μέτρηση της συγκέντρωσης της γλυκόζης στο αίμα ανά τακτά χρονικά διαστήματα κατά την διάρκεια της ημέρας, αλλά και την ρύθμιση της στα φυσιολογικά επίπεδα (70-110 mg/dl), όταν αυτό είναι απαραίτητο. Ωστόσο, η διατήρηση της συγκέντρωσης της γλυκόζης εντός φυσιολογικών ορίων αποτελεί πρόκληση για επιτυχή γλυκαιμικό έλεγχο, διότι επηρεάζεται από πολλές παραμέτρους είτε εξωτερικές, όπως η εξωγενής χορήγηση ινσουλίνης, η φαρμακευτική αγωγή, η διατροφή, οι στρεσογόνες καταστάσεις, η σωματική άσκηση είτε ενδογενείς, όπως η λειτουργία του ήπατος, του εγκεφάλου, των νεφρών, των μυών και του εντέρου. Εξαιτίας της πολυπλοκότητας του μηχανισμού ρύθμισης της γλυκόζης και του μη προβλέψιμου χαρακτήρα των μεταβολών της κατά την διάρκεια της ημέρας, τόσο στο ίδιο άτομο όσο και μεταξύ των ασθενών, ο έλεγχος αυτών των παραγόντων καθίσταται σύνθετος και χρήζει συνεχή παρακολούθηση [24][25].

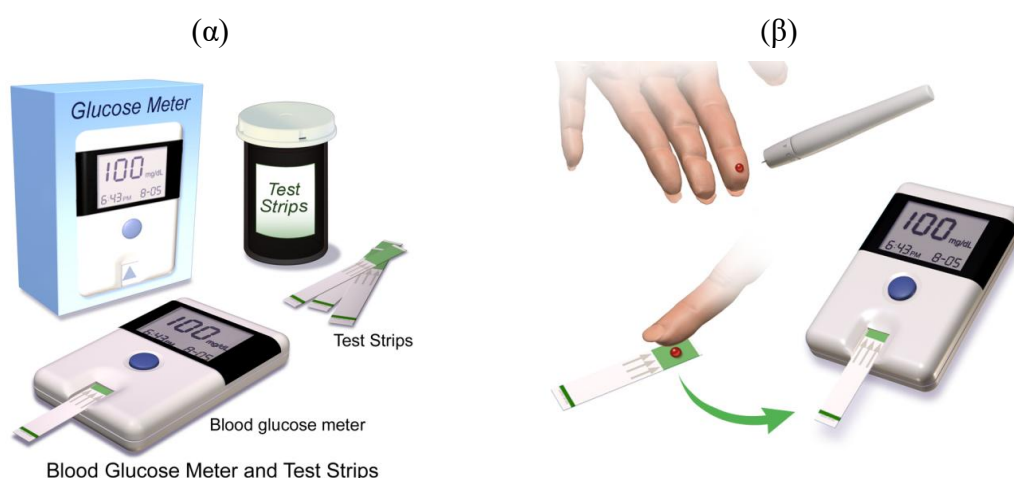
Με βάση αυτά, συνάγεται το συμπέρασμα ότι ο γλυκαιμικός έλεγχος για ασθενείς με ΣΔ, πόσο μάλλον με ΣΔΤ1 (ινσουλινοεξαρτώμενος), απαιτεί διαρκή ενασχόληση τόσο του ίδιου του ατόμου όσο και των κλινικών που του παρέχουν την θεραπευτική αγωγή (π.χ σχήμα ινσουλίνης). Ο ασθενής πρέπει να καταγράφει τακτικά τα επίπεδα γλυκόζης στο αίμα, τις καθημερινές του συνήθειες (γεύματα, ποσότητα υδατανθράκων, σωματική άσκηση), αλλά και να λαμβάνει εξωγενείς δόσεις ινσουλίνης όταν αυτό κρίνεται απαραίτητο, προκειμένου τα επίπεδα σακχάρου στο αίμα να διατηρούνται όσο το δυνατόν πιο κοντά στα φυσιολογικά και να αποφεύγονται υπογλυκαιμικά ή υπεργλυκαιμικά επεισόδια. Οι πρόσφατες τεχνολογικές εξελίξεις επιτρέπουν την συνεχή καταγραφή των επιπέδων γλυκόζης καθώς και την εξωγενή χορήγηση ινσουλίνης χρησιμοποιώντας προηγμένες διατάξεις. Στο Κεφάλαιο 2 αναλύονται οι διατάξεις μέτρησης της γλυκόζης και οι μέθοδοι χορήγησης ινσουλίνης, και πως τα δεδομένα από αυτές τις διατάξεις μπορούν να αξιοποιηθούν χρησιμοποιώντας Μηχανική Μάθηση.

2.1 Διατάξεις Μέτρησης Γλυκόζης

Η γνώση των επιπέδων γλυκόζης στο αίμα ενός διαβητικού ατόμου με ΣΔΤ1 είναι αναγκαία για τον σχεδιασμό της θεραπευτικής αγωγής από τους κλινικούς. Έτσι, έχουν αναπτυχθεί συσκευές-μετρητές της γλυκόζης που παρέχουν στον ασθενή την δυνατότητα να μετράει με αξιοπιστία τα επίπεδα γλυκόζης στο αίμα. Οι εμπορικά διαθέσιμες διατάξεις μέτρησης της γλυκόζης διαχωρίζονται στις συμβατικές διατάξεις μέτρησης της γλυκόζης σε τριχοειδικό αίμα και στις προηγμένες Διατάξεις Συνεχούς Μέτρησης Γλυκόζης (ΔΣΜΓ ή Continuous Glucose Monitoring Systems, CGMS) στον υποδόριο ιστό [24].

2.1.1 Συμβατικές Διατάξεις

Από τις πρώτες, χρονολογικά, διατάξεις μέτρησης της γλυκόζης που χρησιμοποιήθηκαν για γλυκαιμικό έλεγχο είναι οι συμβατικές διατάξεις ή αλλιώς "finger sticks". Ανήκουν στις Διατάξεις Αυτοπαρακολούθησης της γλυκόζης αίματος από τον ίδιο τον ασθενή (ΔΑΓ ή Self-Monitoring of Blood Glucose, SMBG). Οι διατάξεις αυτές είναι επεμβατικές διατάξεις οι οποίες λαμβάνουν μικρή ποσότητα δείγματος από το τριχοειδικό αίμα του ακροδακτύλου του ατόμου ύστερα από διάτρηση με λεπτή βελόνα. Το δείγμα αίματος τοποθετείται σε κάθε μέτρηση σε μια ενζυματική δοκιμαστική ταινία η οποία περιέχει χημικές ουσίες (glucose oxidase ή dehydrogenase ή hexokinase) που αντιδρούν με την γλυκόζη του δείγματος [26]. Έπειτα, η ταινία τοποθετείται σε υποδοχή ψηφιακής συσκευής στην οποία καταγράφεται η μέτρηση της συγκέντρωσης της γλυκόζης, η οποία έχει προκύψει είτε χρησιμοποιώντας κάποιο ηλεκτρικό χαρακτηριστικό όπως το ρεύμα που διαρρέει το δείγμα, είτε μέσω της ανάκλασης του φωτός από το δείγμα αίματος. Τέλος, οι πιο σύγχρονες συσκευές διαθέτουν ενσωματωμένο λογισμικό που επιτρέπει την ενσύρματη ή ασύρματη σύνδεση με υπολογιστή και την μεταφορά των δεδομένων σε αυτόν [27].



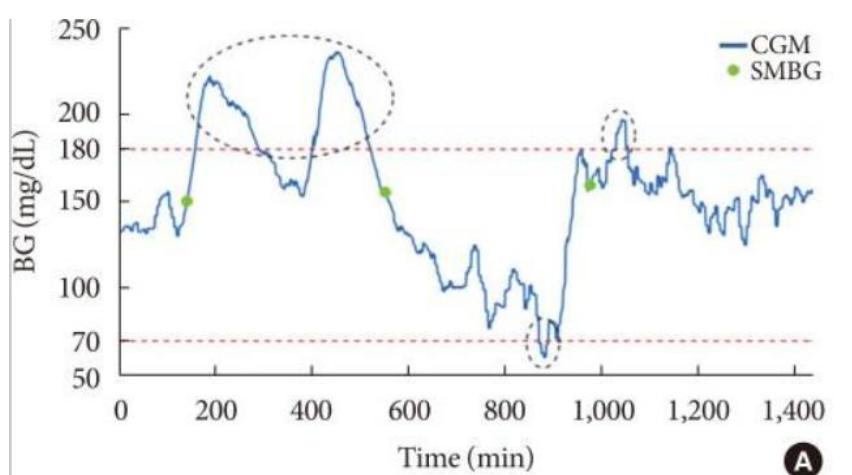
Σχήμα 2.1: (α) Μετρητής γλυκόζης finger sticks, (β) Διαδικασία μέτρησης με διάτρηση δακτύλου [26]

Παραδοσιακά, για τον έλεγχο της γλυκόζης στο αίμα χρησιμοποιούνται οι άκρες των δακτύλων για την διάτρηση. Αυτό οφείλεται στο γεγονός ότι έχουν πολλά τριχοειδή αγγεία και συνήθως παρέχουν αρκετά μεγάλη ποσότητα δείγματος τριχοειδικού αίματος. Ωστόσο, οι άκρες των δακτύλων έχουν επίσης πολλές νευρικές απολήξεις και επομένως είναι αρκετά ευαίσθητες περιοχές. Είναι εύλογο ότι για έναν ασθενή που πάσχει από ΣΔΤ1 που χρειάζεται να καταγράφει τα επίπεδα γλυκόζης του πολλές φορές κατά την διάρκεια της ημέρας, αυτή είναι μια επίπονη διαδικασία. Για το λόγο αυτό, έχουν αναπτυχθεί εναλλακτικές δοκιμές ιστών, όπου η λήψη αίματος γίνεται από ιστούς λιγότερο ευαίσθητους, όπως ο ανώτερος βραχίονας, το αντιβράχιο και ο μηρός. Παρ' όλ' αυτά, οι μεταβολές στην συγκέντρωση της γλυκόζης εμφανίζονται ταχύτερα στο τριχοειδικό αίμα των ακροδαχτύλων από ότι σε άλλα σημεία. Επομένως, σε περιπτώσεις που τα επίπεδα γλυκόζης αλλάζουν γρήγορα, όπως μετά από γεύματα ή ύστερα από χορήγηση ινσουλίνης, ενδέχεται οι μετρητές εναλλακτικού τόπου να είναι λιγότερο ακριβείς. Γι' αυτό το λόγο, οι ειδικοί συμβουλεύουν τους ασθενείς να χρησιμοποιούν τα ακροδάχτυλα για μετρήσεις της συγκέντρωσης της γλυκόζης επειδή παρέχουν γρήγορα και αξιόπιστα αποτελέσματα [28].

2.1.2 Διατάξεις Συνεχούς Μέτρησης Γλυκόζης

Οι συμβατικές διατάξεις μέτρησης της συγκέντρωσης της γλυκόζης παρέχουν μόνο ένα στιγμιότυπο των τρεχόντων επιπέδων της στο αίμα την χρονική στιγμή της μέτρησης. Οι τάσεις στις διακυμάνσεις της γλυκόζης λόγω των γευμάτων, της αγωγής ή/και της σωματικής άσκησης, μπορεί να μην ανιχνεύονται ή να μην αντικατοπτρίζονται με ακρίβεια. Έτσι, αναπτύχθηκαν προηγμένες Διατάξεις Συνεχούς Μέτρησης Γλυκόζης (ΔΣΜΓ), οι οποίες χρησιμοποιούνται κατά κόρο για την διαχείριση του γλυκαιμικού ελέγχου σε ασθενείς με ΣΔΤ1 [29]. Αυτή η νέα τεχνολογία επέτρεψε στους ασθενείς να παρακολουθούν τα επίπεδα σακχάρου στο αίμα καθ' όλη την διάρκεια της ημέρας. Ο αισθητήρας μέτρησης της γλυκόζης έχει την δυνατότητα να λαμβάνει μετρήσεις των επιπέδων γλυκόζης συνεχώς, με αποτέλεσμα να παρέχεται μια ολοκληρωμένη εικόνα της συμπεριφοράς των μεταβολών της γλυκόζης ακόμη και κατά την διάρκεια της νύχτας. Οι περισσότερες σύγχρονες ΔΣΜΓ σήμερα είναι σε θέση να παρέχουν στους ασθενείς δεδομένα σε πραγματικό χρόνο και έξυπνα χαρακτηριστικά, όπως για παράδειγμα σήματα ειδοποίησης υπογλυκαιμίας ή υπεργλυκαιμίας όταν η γλυκόζη βρίσκεται εκτός των φυσιολογικών ορίων, βελτιώνοντας την αυτοδιαχείριση των ασθενών. Από τα παραπάνω, αντιλαμβάνεται κανείς ότι τα δεδομένα που καταγράφονται χρησιμοποιούνται τόσο από τον ίδιο τον ασθενή, όσο και από το ιατρικό προσωπικό βραχυπρόθεσμα, για να αποτρέπονται επιπλοκές που προκαλούνται λόγω των μεταβολών, και μακροπρόθεσμα, ώστε να εκτιμούνται μοτίβα διακυμάνσεων της γλυκόζης κάθε μοναδικού ασθενή με σκοπό να παρέχονται εξατομικευμένες προτάσεις θεραπευτικής αγωγής [29][30][31].

Συνήθως, οι ΔΣΜΓ αποτελούνται από έναν αισθητήρα μέτρησης γλυκόζης, έναν μη εμφυτεύσιμο πομπό και έναν ηλεκτρονικό δέκτη τοποθετημένο πάνω στον ασθενή. Ο αισθητήρας είναι τοποθετημένος ακριβώς κάτω από το δέρμα και κάθε λίγα λεπτά μετρά τα επίπεδα γλυκόζης στον υποδόριο ιστό και στέλνει μέσω του πομπού τις καταγραφές στην οθόνη του δέκτη [27]. Οι περισσότεροι αισθητήρες πραγματοποιούν 288 μετρήσεις το 24ωρο δηλαδή λαμβάνουν μια μέτρηση κάθε 5 λεπτά από το διάμεσο υγρό. Το πλάσμα και το μεσοκυττάριο υγρό παρουσιάζουν διαφορετικά χαρακτηριστικά και πρέπει να θεωρούνται διαφορετικά διαμερίσματα γλυκόζης. Οι τιμές της γλυκόζης στο διάμεσο υγρό προσδιορίζονται από τον ρυθμό διάχυσης της γλυκόζης από το πλάσμα και τον ρυθμό πρόσληψης από τα κύτταρα στον υποδόριο ιστό. Έτσι, λόγω της μέτρησης στο ενδιάμεσο υγρό υπάρχει καθυστέρηση στο αποτέλεσμα σε σχέση με το τριχοειδικό αίμα περίπου 10-20 λεπτά. Μελέτες έχουν δείξει ότι ταχείς μεταβολές της γλυκόζης αίματος αυξάνουν τον χρόνο καθυστέρησης εμφάνισης τους στο μεσοκυττάριο υγρό, με αποτέλεσμα να αποτυπώνονται φυσιολογικές τιμές στην οθόνη τη στιγμή που ο ασθενής ήδη βιώνει τα πρώτα συμπτώματα υπεργλυκαιμίας ή υπογλυκαιμίας. Επίσης, η μείωση της συγκέντρωσης της γλυκόζης τείνει να εμφανίζεται ταχύτερα στο ενδιάμεσο υγρό απ' ότι στο αίμα και άρα η ένδειξη της ΔΣΜΓ να προειδοποιεί για επερχόμενη υπογλυκαιμία. Εξαιτίας της ύπαρξης αυτής της χρονικής υστέρησης, απαιτείται βαθμονόμηση (καλιμπράρισμα) της διάταξης με τιμές τριχοειδικού αίματος, που λαμβάνονται με τις συμβατικές διατάξεις μέτρησης της γλυκόζης 2-4 φορές την ημέρα [32]. Ακόμη, η χρήση συμβατικών διατάξεων για την μέτρηση τριχοειδικού αίματος συνίσταται όταν ο ασθενής πρόκειται να προβεί σε διορθωτικές κινήσεις αύξησης ή μείωσης σακχάρου βασιζόμενος στα αποτελέσματα του αισθητήρα.

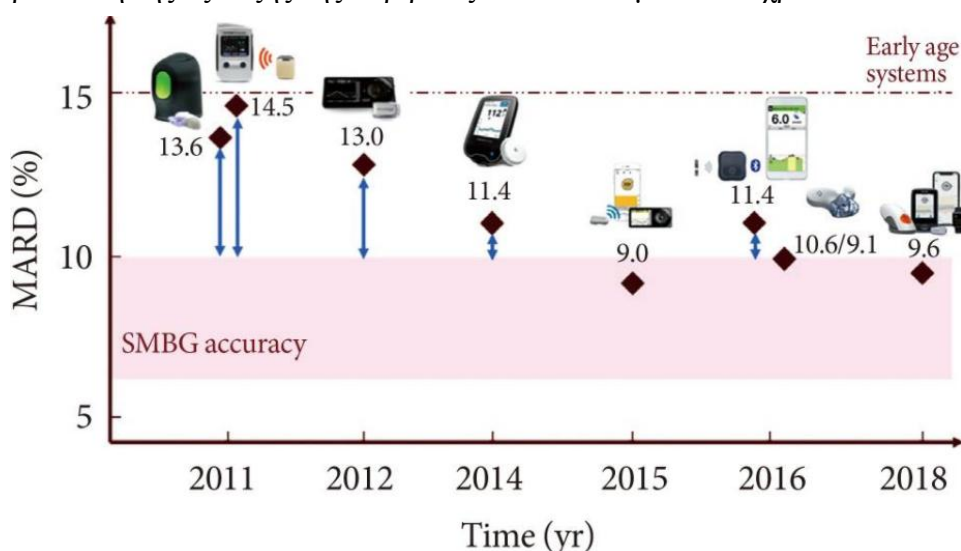


Σχήμα 2.2: Αντιπροσωπευτικά δεδομένα παρακολούθησης της γλυκόζης στο αίμα (BG mg/dL) που λαμβάνονται με ΔΣΜΓ (συνεχής μπλε γραμμή) και με συμβατικές ΔΑΓ (πράσινες κουκκίδες). Οι διακεκομμένοι κύκλοι υποδηλώνουν υπεργλυκαιμικά ή υπογλυκαιμικά επεισόδια, τα οποία χρησιμοποιώντας συμβατικές διατάξεις μέτρησης γλυκόζης δεν είναι ανιχνεύσιμα [31]

Αισθητήρες Συνεχούς Μέτρησης Γλυκόζης

Οι βιοαισθητήρες που χρησιμοποιούνται για την μέτρηση συγκεντρώσεων διαφόρων μορίων αποτελούνται κυρίως από τρία μέρη: τα στοιχεία που χρησιμοποιούνται για την αναγνώριση των μορίων-στόχων παρουσία διαφόρων χημικών ουσιών, έναν μορφοτροπέα που μετατρέπει το συμβάν βιοαναγνώρισης σε μετρήσιμο σήμα και ένα σύστημα επεξεργασίας σήματος που μετατρέπει το σήμα σε αναγνώσιμη μορφή. Γενικά, τα στοιχεία μοριακής αναγνώρισης των βιοαισθητήρων περιλαμβάνουν υποδοχείς, ένζυμα, αντισώματα, νουκλεϊκά οξέα κ.α. Όσον αφορά τους μετατροπείς, διακρίνονται σε ηλεκτροχημικούς, οπτικούς, θερμομετρικούς και μαγνητικούς. Η πλειοψηφία των τρεχόντων βιοαισθητήρων που χρησιμοποιούνται από ΔΣΜΓ είναι ηλεκτροχημικού τύπου, λόγω της καλύτερης ευαισθησίας, της εύκολης συντήρησης και του χαμηλού κόστους. Οι ηλεκτροχημικοί αισθητήρες υποδιαιρούνται σε ποτενσιομετρικούς, αμπερομετρικούς ή αγωγιμομετρικούς. Οι περισσότεροι εμπορικά διαθέσιμοι βιοαισθητήρες γλυκόζης είναι ενζυμικοί αμπερομετρικοί, οι οποίοι έχουν μελετηθεί ευρέως τις τελευταίες δεκαετίες [33].

Για την αξιολόγηση της ακρίβειας και της απόδοσης των αισθητήρων αυτών χρησιμοποιούνται στατιστικά και κλινικά κριτήρια. Η πιο κοινή μέτρηση που χρησιμοποιείται στην βιβλιογραφία για την αξιολόγηση της ακρίβειας της ΔΣΜΓ είναι η μέτρηση της μέσης απόλυτης σχετικής διαφοράς MARD (Mean Absolute Relative Difference). Πρακτικά είναι ο μέσος όρος του απόλυτου σφάλματος μεταξύ όλων των τιμών που προκύπτουν από την ΔΣΜΓ και των αντίστοιχων τιμών αναφοράς. Ένα μικρό ποσοστό MARD υποδεικνύει ότι οι μετρήσεις της ΔΣΜΓ είναι κοντά στην τιμή γλυκόζης αναφοράς, ενώ ένα μεγαλύτερο ποσοστό MARD δείχνει μεγαλύτερες αποκλίσεις [34]. Στο **Σχήμα 2.3** παρουσιάζεται μια γραφική αναπαράσταση της εξέλιξης της ακρίβειας των ΔΣΜΓ μέσα στα χρόνια.



Σχήμα 2.3: Εξέλιξη της ακρίβειας των σύγχρονων συστημάτων συνεχούς μέτρησης της γλυκόζης με την πάροδο των χρόνων. Από αριστερά: Medtronic Enlite, Abbott Freestyle Navigator, Dexcom G4 Platinum, Abbott Freestyle Libre, Dexcom G4 Platinum με λογισμικό 505, Senseonics Eversense, Dexcom G5, Dexcom G6

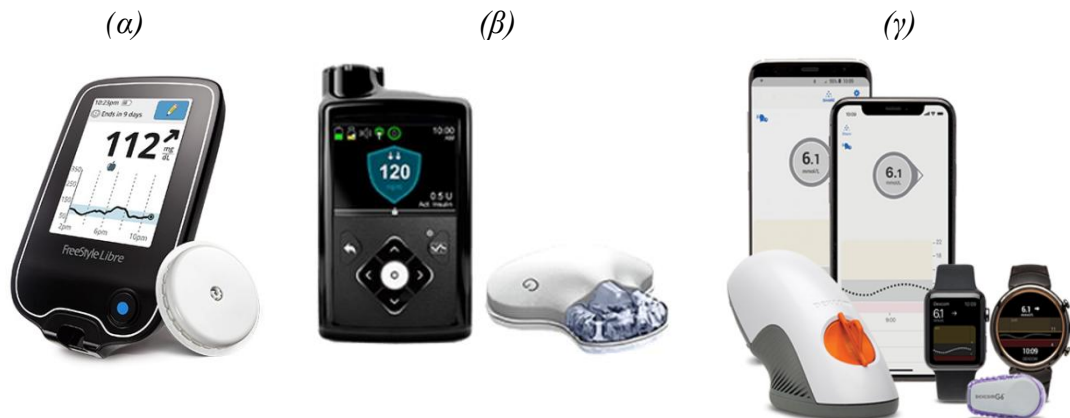
Στη συνέχεια γίνεται μια σύντομη περιγραφή των αισθητήρων που χρησιμοποιούνται από τις διατάξεις συνεχούς καταγραφής των επιπέδων της γλυκόζης. Αυτοί διαχωρίζονται σε επεμβατικούς και μη επεμβατικούς αισθητήρες ανάλογα με τον τρόπο παρέμβασης τους στο ανθρώπινο σώμα.

Επεμβατικοί Αισθητήρες

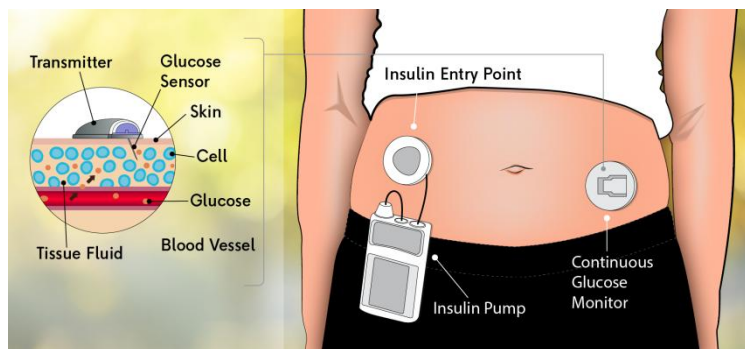
- **Υποδόριοι αισθητήρες που βασίζονται στο ένζυμο οξειδάση της γλυκόζης**

Οι αισθητήρες αυτοί τοποθετούνται κάτω από το δέρμα, στον υποδόριο ιστό και μετρούν την συγκέντρωση της γλυκόζης στο μεσοκυττάριο υγρό. Συγκεκριμένα, χρησιμοποιούνται ηλεκτροχημικοί αισθητήρες οι οποίοι βασίζονται στην χημική αντίδραση της γλυκόζης με κάποιο ένζυμο. Αποτελούνται από ένα ηλεκτρόδιο που χρησιμοποιεί το ένζυμο οξειδάση της γλυκόζης (GO_x). Το ένζυμο GO_x καταλύει την οξείδωση της γλυκόζης με αποτέλεσμα να δημιουργούνται αποδέκτες ή δότες ηλεκτρονίων. Αυτή η κίνηση ηλεκτρονίων μπορεί να αξιοποιηθεί από το ηλεκτρόδιο, το οποίο θα ανιχνεύσει την τάση ή το ρεύμα που δημιουργείται. Όπως ήδη αναφέρθηκε, οι πιο ευρέως διαδεδομένοι βιοαισθητήρες γλυκόζης είναι ενζυμικοί αμπερομετρικοί και άρα ανιχνεύουν ρεύμα κατά την αντίδραση με το ένζυμο. Η συγκέντρωση της γλυκόζης είναι ανάλογη με την ισχύ του παραγόμενου ηλεκτρικού σήματος και αυτό γιατί το υπεροξειδίο του υδρογόνου που παράγεται κατά την χημική αντίδραση είναι ανάλογο της συγκέντρωσης της γλυκόζης [35]. Τα βασικότερα μειονεκτήματα των αισθητήρων, όπως έχουν αναφερθεί και ανωτέρω, είναι η συχνή αλλαγή του αισθητήρα κάθε λίγες μέρες για αποφυγή μόλυνσης και η βαθμονόμηση του αισθητήρα 2-4 φορές την ημέρα χρησιμοποιώντας συμβατικές διατάξεις μέτρησης. Παρά τους περιορισμούς οι συσκευές αυτές μετρούν αρκετά καλά τις συγκεντρώσεις της γλυκόζης και έχει αποδειχθεί ότι βελτιώνουν τον γλυκαιμικό έλεγχο του ασθενούς.

Στο **Σχήμα 2.4** παρουσιάζονται ενδεικτικά οι πιο σύγχρονες συσκευές συνεχούς παρακολούθησης της συγκέντρωσης της γλυκόζης από τις γνωστές εταιρείες Abbott, Medtronic και Dexcom. Ο πομπός (**Σχήμα 2.5**) στέλνει ασύρματα τις πληροφορίες μέτρησης στις θόνες παρακολούθησης ή σε άλλες κινητές συσκευές. Σε περίπτωση αδυναμίας ασύρματης σύνδεσης, η συσκευή παρακολούθησης σαρώνει τον πομπό προκειμένου να καταγραφεί η μέτρηση (Freestyle Libre). Τα δεδομένα εμφανίζονται στις θόνες σε πραγματικό χρόνο και παρέχουν χαρακτηριστικά όπως βέλη ανόδου/καθόδου του σακχάρου στο αίμα και ειδοποιήσεις για υπογλυκαιμία και υπεργλυκαιμία. Στον **Πίνακα 2.1** φαίνονται κάποια από τα χαρακτηριστικά των αισθητήρων αυτών που αφορούν την ακρίβεια τους, κάθε πότε πρέπει να αλλάζονται, κάθε πότε χρειάζονται βαθμονόμηση αλλά και έξυπνα χαρακτηριστικά που παρέχουν.



Σχήμα 2.4: Προηγμένες διατάξεις συνεχούς καταγραφής της συγκέντρωσης της γλυκόζης. (α) Abbott-Freestyle Libre, (β) Medtronic-Guardian Sensor, (γ) Dexcom-G6



Σχήμα 2.5: ΔΣΜΓ που αποτελείται από τον αισθητήρα μέτρησης της γλυκόζης και την αντλία ινσουλίνης. Ο αισθητήρας (Glucose Sensor), υπό την μορφή βελόνας, είναι τοποθετημένος κάτω από το δέρμα στον υποδόριο ιστό της κοιλιάς. Ανιχνεύει την συγκέντρωση γλυκόζης στο μεσοκυττάριο υγρό (Tissue Fluid). Λόγω της μέτρησης στο διάμεσο υγρό υπάρχει καθυστέρηση στο αποτέλεσμα σε σχέση με το τριχοειδικό αίμα περίπου 10-20 λεπτά

Πίνακας 2.1: Συγκεντρωτικός πίνακας χαρακτηριστικών των τριών διατάξεων μέτρησης γλυκόζης από τις εταιρείες Dexcom, Abbot και Medtronic αντίστοιχα [31]

Κατασκευαστής	Μοντέλο	Ακρίβεια (MARD) %	Βαθμονομήσεις	Χρόνος Ζωής Αισθητήρα	Έξυπνα Χαρακτηριστικά
Dexcom	G6	10	-	10	Βέλη της τάσης του σακχάρου, ειδοποιήσεις ρυθμού αλλαγής και υπο/υπεργλυκαιμικών επεισοδίων, απομακρυσμένη παρακολούθηση, ασύρματη επικοινωνία μέχρι 5 συσκευές

Abbott	Freestyle Libre	11.4	-	14	Βέλη της τάσης του σακχάρου, σάρωση για καταγραφή της μέτρησης
Medtronic	Guardian Sensor 3	10.6 (κοιλιά) 9.1 (ώμος)	Κάθε 12 ώρες	7 ημέρες	Βέλη της τάσης του σακχάρου, ειδοποιήσεις ρυθμού αλλαγής και υπο/υπεργλυκαιμικών επεισοδίων, σύνδεση με τις αντλίες ινσουλίνης της Medtronic

- **Εμφυτεύσιμοι αισθητήρες**

Οι υποδόριοι αισθητήρες έχουν κάποια σημαντικά μειονεκτήματα εκ των οποίων είναι η συχνή αντικατάσταση τους και η συχνή βαθμονόμηση τους. Για τον λόγο αυτό, χρήζει αναγκαία η ανάπτυξη και προώθηση νέων τεχνολογιών για την βελτίωση της φροντίδας των ασθενών. Η ανάπτυξη εμφυτεύσιμων αισθητήρων στοχεύει στην αντικατάσταση των υποδόριων αισθητήρων. Η εμφύτευση του αισθητήρα θα επιτρέπει την συνεχή παρακολούθηση των επιπέδων γλυκόζης, καθώς και την επικοινωνία με εμφυτευμένη αντλία ινσουλίνης.

Το 2018 εγκρίθηκε από τον Αμερικανικό Οργανισμό Τροφίμων και Φαρμάκων (U.S Food & Drug Administration, FDA) το σύστημα Eversense CGM που έχει αναπτυχθεί από την εταιρεία Senseonics. Αυτό είναι το πρώτο εγκεκριμένο, από τον FDA, σύστημα συνεχούς μέτρησης γλυκόζης που περιλαμβάνει έναν πλήρως εμφυτεύσιμο αισθητήρα για την ανίχνευση της συγκέντρωσης γλυκόζης. Ο αισθητήρας, ο οποίος εμφυτεύεται ακριβώς κάτω από το δέρμα με διάρκεια ζωής 90 ημέρες, επικαλύπτεται με μια φθορίζουσα χημική ουσία η οποία όταν εκτίθεται σε σάκχαρο παράγει μια μικρή ποσότητα φωτός που μετράται από τον αισθητήρα. Κάθε πέντε λεπτά, οι μετρήσεις αποστέλλονται σε μια κινητή συσκευή, ώστε να ειδοποιούνται οι χρήστες εάν τα επίπεδα γλυκόζης είναι πολύ χαμηλά ή πολύ υψηλά [36].

Μη επεμβατικοί Αισθητήρες

- **Οπτικοί Αισθητήρες**

Οι οπτικοί αισθητήρες χρησιμοποιούν διάφορες ιδιότητες του φωτός για την αλληλεπίδραση με μόρια γλυκόζης. Πρόκειται για μη επεμβατικές διατάξεις οι οποίες χρησιμοποιούν δέσμη φωτός, συγκεκριμένου μήκους κύματος, που προσπίπτει σε κάποιο σημείο του σώματος. Το φως πρέπει να είναι με συγκεκριμένο μήκος κύματος ώστε να διεισδύσει στο αίμα των στρωμάτων του δέρματος και να μην απορροφηθεί από το νερό ή από άλλα

συστατικά του δέρματος. Ιδιότητες όπως η ανάκλαση, η σκέδαση και η απορρόφηση τους φωτός από το δέρμα φανερώνουν τον τρόπο που αλληλεπιδρά το φως με τον ιστό. Η απορρόφηση του φωτός από το δέρμα εξαρτάται από την δομή και τα συστατικά του, συνεπώς και από την συγκέντρωση της γλυκόζης. Επομένως, με βάση τις μεταβολές που παρατηρούνται στην ένταση του ανακλώμενου σήματος, του μήκους κύματος της δέσμης αλλά και της πόλωσης του διερχόμενου σήματος, προσδιορίζεται η συγκέντρωση της γλυκόζης [4][35]. Μερικές από τις φασματοσκοπικές μέθοδοι που χρησιμοποιούνται είναι η Φασματοσκοπία Υπερύθρου, Φασματοσκοπία στο εγγύς υπέρυθρο, Φασματοσκοπία Ραμαν, Φωτοακουστική Φασματοσκοπία, Οπτική Τομογραφία Συνοχής.

- **Δερματικοί Αισθητήρες**

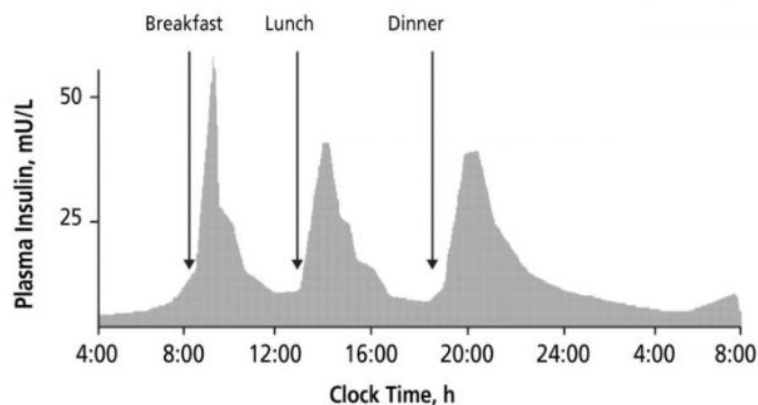
Οι δερματικοί αισθητήρες καταγράφουν τα επίπεδα της γλυκόζης στο αίμα ή στο διάμεσο υγρό χωρίς να παρεμβαίνουν στην επιφάνεια του δέρματος. Οι δυο βασικότερες μεθοδολογίες που χρησιμοποιούνται είναι η φασματοσκοπία της αντίστασης των ιστών και η Αντίστροφη Ιοντοφόρηση. Στην πρώτη τεχνική μικρής έντασης εναλλασσόμενο ρεύμα περνά σε έναν ιστό και επιφέρει αλλαγές στην αντίσταση του, οι οποίες σχετίζονται με τα μόρια γλυκόζης. Η τεχνική της αντίστροφης ιοντοφόρησης εφαρμόζει δύο ηλεκτρόδια στο δέρμα απ' όπου διαπερνάει χαμηλής έντασης ηλεκτρικό ρεύμα με αποτέλεσμα να δημιουργείται ροή ιόντων μέσω του δέρματος η οποία παρασέρνει και ηλεκτρικά ουδέτερα μόρια, όπως η γλυκόζη. Αυτή η μετακίνηση έχει ως αποτέλεσμα την μεταφορά της γλυκόζης έξω από το δέρμα όπου την μετρά ένας ηλεκτροχημικός αισθητήρας που είναι εκεί προσκολλημένος [35].

2.2 Μέθοδοι Εξωγενούς Χορήγησης Ινσουλίνης

Η ινσουλίνη είναι απαραίτητη για τον μεταβολισμό των υδατανθράκων, των πρωτεϊνών και των λιπιδίων. Η ανάγκη για ινσουλίνη διαφέρει ανάλογα με την δυνατότητα έκκρισης ινσουλίνης από το πάγκρεας, την ευαισθησία των κυττάρων στην ινσουλίνη αλλά και άλλων διαφόρων παραγόντων όπως η διατροφή, η σωματική άσκηση και το άγχος. Τα άτομα που πάσχουν από ΣΔΤ1 παράγουν ελάχιστη ή/και καθόλου ποσότητα ινσουλίνης, επομένως η επιβίωση τους εξαρτάται αποκλειστικά από την εξωγενή χορήγηση της ορμόνης. Αντίθετα, τα άτομα με ΣΔΤ2 δεν εξαρτώνται από εξωγενή έγχυση ινσουλίνης, ωστόσο με την πάροδο του χρόνου παρουσιάζουν μειωμένη παραγωγή της ορμόνης και συνεπώς απαιτούν συμπληρωματική δόση για τον έλεγχο της γλυκόζης στο αίμα, ειδικά σε περιόδους άγχους ή ασθένειας. Η ινσουλινοθεραπεία, λοιπόν, είναι πολύ σημαντική για την αντιμετώπιση του ΣΔ. Πρέπει να τονιστεί ότι σε όλες τις περιπτώσεις χορήγησης ινσουλίνης, ο τύπος και η δοσολογία της ινσουλίνης πρέπει να εξατομικεύονται ανάλογα με τις ανάγκες του κάθε ασθενούς, ώστε να προκύπτει το κατάλληλο σχήμα θεραπείας.

Τύποι Ινσουλίνης

Φυσιολογικά, το πάγκρεας παράγει και εκκρίνει στην κυκλοφορία μια μικρή και σταθερή ποσότητα ινσουλίνης, που ονομάζεται βασική ινσουλίνη (**basal**), ώστε να διατηρούνται σταθερά τα επίπεδα της γλυκόζης μεταξύ των γευμάτων κατά την διάρκεια της ημέρας. Μετά τα γεύματα, λόγω αύξησης του σακχάρου, εκκρίνεται στιγμιαία αιχμή ινσουλίνης, δηλαδή δόση εφόδου (**prandial ή bolus**) επαρκούς ύψους και διάρκειας. Στο **Σχήμα 2.6** φαίνεται το πως ακριβώς εκκρίνεται η ινσουλίνη μέσα σε ένα 24ώρο από το πάγκρεας. Η εξωγενής χορήγηση basal και bolus ινσουλίνης μιμείται την φυσιολογική λειτουργία των β-κύτταρων του παγκρέατος.



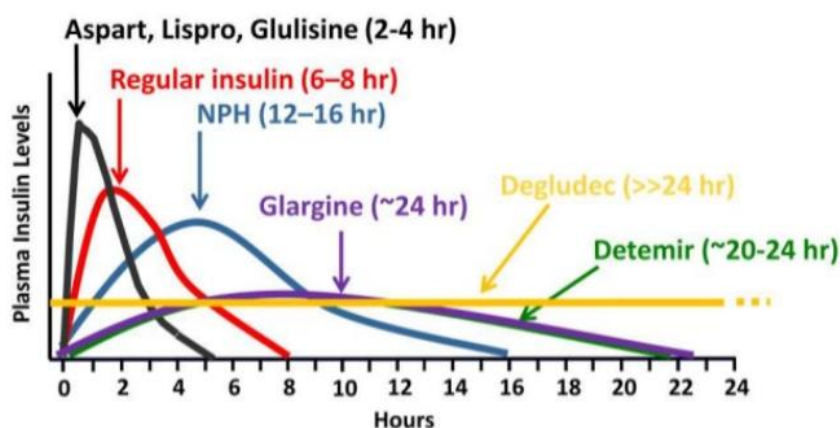
Σχήμα 2.6: Ενδογενής έκκριση ινσουλίνης (*basal και bolus*)

Οι διάφοροι τύποι ινσουλίνης διαχωρίζονται με βάση το πόσο σύντομα αρχίζει και πόσο διαρκεί η δράση τους, αλλά και πότε η επίδραση της ινσουλίνης γίνεται μέγιστη [37].

- **Ταχείας Δράσης Ινσουλίνη:** Η ινσουλίνη ταχείας δράσης, αρχίζει να δρα περίπου 15 λεπτά μετά την έγχυση, κορυφώνεται μέσα σε 1 ή 2 ώρες και διαρκεί μεταξύ 2 με 4 ώρες. Οι δόσεις εφόδου (bolus) χρησιμοποιούν ινσουλίνη ταχείας δράσης και χορηγούνται αμέσως πριν των κυρίων γευμάτων.
- **Βραχείας Δράσης Ινσουλίνη:** Αυτή η ινσουλίνη φτάνει στην κυκλοφορία του αίματος συνήθως εντός 30 λεπτών, κορυφώνεται σε 2-3 ώρες και η δράση της είναι αποτελεσματική για περίπου 3 έως 6 ώρες. Λαμβάνεται μισή ώρα πριν από κάποιο γεύμα.
- **Ενδιάμεσης Δράσης Ινσουλίνη:** Χορηγείται ως basal ινσουλίνη δύο ή περισσότερες φορές μέσα στην μέρα και δρα για 12-18 ώρες. Αρχίζει να δρα στο διάστημα 2-4 ώρες μετά την χορήγηση.
- **Μακράς Δράσης Ινσουλίνη:** Λαμβάνεται μία ή δύο φορές μέσα στην μέρα ως basal δόση και δρα για περίπου 24 ώρες. Η ινσουλίνη μακράς δράσης φτάνει στην κυκλοφορία αρκετές ώρες μετά την χορήγηση.

Πίνακας 2.2: Χαρακτηριστικά σκευασμάτων ινσουλίνης

Σκεύασμα	Τύπος	Έναρξη Δράσης	Κορύφωση Δράσης	Διάρκεια Δράσης
Lispro (Humalog) Aspart (Novolog) Glulisine (Apidra)	Ταχείας Δράσης	10-15 λεπτά	60-90 λεπτά	4-5 ώρες
Regular	Βραχείας Δράσης	30-60 λεπτά	2-4 ώρες	5-8 ώρες
NPH (Humulin)	Ενδιάμεσης Δράσης	2-4 ώρες	4-12 ώρες	10-18 ώρες
Glargine (Lantus) Detemir (Levemir)	Μακράς Δράσης	1 ώρα	Χωρίς Κορύφωση	18-24 ώρες



Σχήμα 2.7: Γράφημα που απεικονίζει τις διαφορές στους χρόνους δράσης μεταξύ των σκευασμάτων ινσουλίνης

Η συνηθέστερη οδός έγχυσης ινσουλίνης είναι η υποδόρια οδός, η οποία εμφανίζει πολλά πλεονεκτήματα σχετικά με την ευκολία, την άνεση και την αξιοπιστία. Ωστόσο, ένα σημαντικό ζήτημα της μεθόδου είναι η χρονική καθυστέρηση για την μεταφορά της ινσουλίνης από τον υποδόριο χώρο στο αίμα, καθώς και ο χρόνος που απαιτείται για την κορύφωση της δράσης της. Σπανιότερα και σε έκτακτες περιπτώσεις, χρησιμοποιείται η ενδοφλέβια οδός για ταχεία μεταφορά της ινσουλίνης στη κυκλοφορία του αίματος. Οι πιο διαδεδομένες μέθοδοι εξωγενούς χορήγησης ινσουλίνης είναι οι ημερήσιες πολλαπλές ενέσιμες δόσεις ινσουλίνης και η συνεχής έγχυση ινσουλίνης μέσω αντλίας, οι οποίες παρουσιάζονται στην συνέχεια.

2.2.1 Ημερήσιες Πολλαπλές Ενέσιμες Δόσεις Ινσουλίνης

Η χρήση υποδόριων ενέσεων ινσουλίνης έχει σχεδιαστεί για να μιμείται την φυσιολογική έκκριση ινσουλίνης του παγκρέατος. Το θεραπευτικό σχήμα που ακολουθείται περιλαμβάνει μία ή δύο ενέσεις με βασική δόση ινσουλίνης μακράς

δράσης κατά την διάρκεια της ημέρας (**basal** δόση ινσουλίνης) και ενέσεις γευματικής ινσουλίνης ταχείας ή βραχείας δράσης πριν από τα γεύματα (**bolus** δόση ινσουλίνης). Ο τύπος, η δοσολογία και το πόσο συχνά πρέπει να χορηγείται ινσουλίνη το προγραμματίζει ο θεράπων ιατρός με βάση προσωπικά δεδομένα του ασθενή. Ωστόσο, ο ασθενής πρέπει να έχει εκπαιδευτεί και να ξέρει να ρυθμίζει μόνος του τις bolus δόσεις, ανάλογα με την ποσότητα υδατανθράκων που πρόκειται να καταναλώσει ή αν πρόκειται να γυμναστεί. Άλλωστε, πρέπει να παρακολουθεί συχνά τα επίπεδα γλυκόζης του, ώστε να πετύχει βέλτιστο γλυκαιμικό έλεγχο.

Πέρα από τις κλασικές σύριγγες για χορήγηση δόσεων ινσουλίνης, κυκλοφορούν στο εμπόριο και άλλες πιο σύγχρονες συσκευές, οι οποίες στοχεύουν σε πιο εύκολη και λιγότερη επίπονη, για τον ασθενή, διαδικασία χορήγησης ινσουλίνης. Τέτοιες συσκευές είναι τα στυλό ινσουλίνης και πίδακες υψηλής πίεσης.

Κλασικές Σύριγγες

Η συμβατική χορήγηση ινσουλίνης περιλαμβάνει υποδόρια ένεση με σύριγγες, μιας χρήσης, συγκεκριμένων μονάδων ινσουλίνης (Units, 1U=36μg ινσουλίνης). Η απαραίτητη ποσότητα ινσουλίνης συγκεντρώνεται σε ένα φιαλίδιο με το οποίο συνδέεται μέσω εμβόλου η βελόνα. Διάφορα μήκη βελόνων είναι διαθέσιμα, ωστόσο σε περίπτωση αλλαγής πρέπει να παρακολουθείται η γλυκόζη αίματος ώστε να εκτιμηθεί η μεταβλητότητα απορρόφησης της ινσουλίνης [38].

Στυλό Ινσουλίνης (Insulin Pens)

Πρόκειται για μια ενιαία μονάδα που μοιάζει με στυλό και αποτελείται από το φυσίγγιο ινσουλίνης και ένα όργανο ένδειξης για την ρύθμιση της χορηγούμενης δοσολογίας. Ο ασθενής τοποθετεί μια μικρή βελόνα στην συσκευή, ρυθμίζει την δόση της ινσουλίνης που πρέπει να χορηγηθεί και στη συνέχεια πιέζει το έμβολο για την έγχυση της δόσης στον υποδόριο χώρο. Οι λεπτές βελόνες καθιστούν την διαδικασία χορήγησης σχετικά ανώδυνη. Τα στυλό ινσουλίνης διαχωρίζονται σε δύο κατηγορίες τα μίας χρήσης και τα επαναχρησιμοποιήσιμα [39]. Τα μίας χρήσης περιέχουν ένα προγεμισμένο φυσίγγιο ινσουλίνης, τα οποία αφότου χρησιμοποιηθούν πρέπει να αντικαθίστανται από νέα. Αντίθετα, τα επαναχρησιμοποιήσιμα στυλό περιέχουν ένα φυσίγγιο ινσουλίνης το οποίο μπορεί να αντικατασταθεί μετά την χρήση.



Σχήμα 2.8: Σύριγγες χορήγησης ινσουλίνης



Σχήμα 2.9: Στυλό Ινσουλίνης

Πίδακες Υψηλής Πίεσης (Jet Injectors)

Οι συσκευές αυτές αντί της χρήσης βελόνας χρησιμοποιούν πίδακα υψηλής πίεσης για την χορήγηση της ινσουλίνης στον υποδόριο χώρο. Ο μηχανισμός ασκεί υψηλή πίεση και ψεκάζει στο δέρμα την κατάλληλη δόση ινσουλίνης. Η τεχνολογία αυτή δεν είναι ιδιαίτερη διαδεδομένη λόγω του υψηλού κόστους αλλά και ορισμένων μειονεκτημάτων όπως η πρόκληση πόνου και τοπικού ερεθισμού.



Σχήμα 2.10: Συσκευή χορήγησης ινσουλίνης με τη μορφή πίδακα υψηλής πίεσης

2.2.2 Αντλίες Συνεχούς Έγχυσης Ινσουλίνης

Η θεραπεία με αντλία ινσουλίνης, γνωστή ως θεραπεία Συνεχούς Υποδόριας Έγχυσης Ινσουλίνης (Continuous Subcutaneous Insulin Infusion, CSII) άλλαξε δραματικά την φιλοσοφία αντιμετώπισης του ΣΔ. Πρόκειται για μια εξελισσόμενη μορφή παροχής ινσουλίνης η οποία έχει αποδειχθεί αποτελεσματική για την διατήρηση της ευγλυκαιμίας και την βελτίωση του τρόπου ζωής των ασθενών.

Οι αντλίες ινσουλίνης είναι μικρές προγραμματιζόμενες σε υπολογιστή συσκευές. Περιλαμβάνουν το δοχείο της ινσουλίνης και μια συσκευή μέσω της οποίας ο ασθενής ρυθμίζει το βασικό ρυθμό έγχυσης και τις προγευματικές δόσεις με βάση τα γεύματα του. Ο μηχανισμός που χρησιμοποιεί μιμείται την φυσιολογική έκκριση ινσουλίνης από το πάγκρεας. Η αντλία λειτουργεί παρέχοντας υποδόρια έγχυση ινσουλίνης ταχείας δράσης η οποία καλύπτει τις ανάγκες του ατόμου σε ινσουλίνη, ανεξάρτητα από την λήψη γευμάτων, κατά την διάρκεια της ημέρας (**basal**) αλλά και πριν τα γεύματα με βάση την περιεκτικότητά τους σε υδατάνθρακες (**γευματικό bolus**). Ακόμη, εκκρίνει και δόσεις ινσουλίνης που επαναφέρουν μια αυξημένη τιμή του σακχάρου στα φυσιολογικά (**διορθωτικό bolus**). Οι αντλίες, συνήθως, χρησιμοποιούνται σε συνδυασμό με ΔΣΜΓ και οι δόσεις προγραμματίζονται εξατομικευμένα με βάση τα δεδομένα από τον αισθητήρα. Ο ρυθμός έγχυσης ινσουλίνης ρυθμίζεται με βάση την ευαισθησία του ασθενή στην ορμόνη καθώς και προσαρμόζεται στο πρόγραμμα του (ψυχολογία, διατροφή και σωματική άσκηση). Ο ασθενής έχει την δυνατότητα να παρακάμψει την προτεινόμενη δόση ινσουλίνης ή να διακόψει την έγχυση όποτε κρίνεται απαραίτητο [40].

Οι αντλίες ινσουλίνης διατίθενται σε διάφορες μορφές οι οποίες περιγράφονται στην συνέχεια:

Αντλίες με σετ έγχυσης

Από τις πιο διαδεδομένες αντλίες ινσουλίνης που χρησιμοποιούνται για την αντιμετώπιση του ΣΔ είναι οι αντλίες που χρησιμοποιούν σετ έγχυσης. Ένα εύκαμπτο λεπτό σωληνάκι συνδέει το δοχείο της ινσουλίνης με την κάνουλα, που έχει τοποθετηθεί κάτω από το δέρμα, μεταφέροντας την ινσουλίνη στον υποδόριο χώρο. Η ίδια η αντλία, μέσω της οποίας προγραμματίζεται η δόση ινσουλίνης, μπορεί να φορεθεί πάνω στην ένδυση του ασθενή [40]. Συχνά, η αντλία συνδέεται ασύρματα και επικοινωνεί με ΔΣΜΓ με σκοπό να αποφεύγονται υπογλυκαιμικά ή υπεργλυκαιμικά επεισόδια.



Σχήμα 2.11: Προηγμένες αντλίες ινσουλίνης με σετ έγχυσης. (α) Medtronic 670G, (β) Tandem t:slim x2

Αντλίες τύπου patch

Οι αντλίες αυτές χρησιμοποιούν ένα έμπλαστρο (patch) και μια ασύρματη συσκευή ελέγχου και ρύθμισης των bolus δόσεων. Το patch προσκολλάται στην επιφάνεια του δέρματος και συνήθως μετά από λίγες μέρες αφαιρείται και απορρίπτεται. Οι αντλίες αυτές δεν χρησιμοποιούν σωλήνωση για την προώθηση της ινσουλίνης, αντίθετα η βελόνα ινσουλίνης είναι μέρος της αντλίας και εισάγεται στο δέρμα όταν συνδέεται το patch. Η απουσία του σετ έγχυσης έχει σημαντικά πλεονεκτήματα, καθώς μειώνει τον κίνδυνο αποσύνδεσης της αντλίας και βελτιώνει τον τρόπο ζωής του ατόμου [40].



Σχήμα 2.12: Προηγμένες αντλίες ινσουλίνης τύπου patch. (α) V-GO, (β) Omnipod

Στον συγκεντρωτικό **Πίνακα 2.3** παρουσιάζονται κάποια από τα χαρακτηριστικά των αντλιών συνεχούς έγχυσης ινσουλίνης που κυκλοφορούν στο εμπόριο.

Πίνακας 2.3: Σύνοψη των κύριων χαρακτηριστικών των μοντέλων συνεχούς έγχυσης υποδόριας ινσουλίνης για τους δύο τύπους αντλίας [40]

Τύπος Αντλίας	Μοντέλο	Τύπος Ινσουλίνης	Σύνδεση με ΔΣΜΓ	Έξυπνα Χαρακτηριστικά
Αντλία με σετ έγχυσης	Medtronic 670G	ταχείας δράσης ινσουλίνη 300 U των U-100	Medtronic Guardian Sensor 3	Αυτόματη λειτουργία για αναστολή έγχυσης σε περίπτωση υπογλυκαιμίας
	Tandem t: slim x2	ταχείας δράσης ινσουλίνη 300 U των U-100	Dexcom G5 & G6	Αναστολή παράδοσης βασικής ινσουλίνης σε προκαθορισμένα όρια και αυτόματη επανεκκίνηση
Αντλία τύπου patch	V-GO	ταχείας δράσης ινσουλίνη Humalog ή Novolog 20, 30 ή 40U των U-100	-	Απελευθερώνει σταθερό ρυθμό basal ινσουλίνη αλλά και bolus ινσουλίνη με το πάτημα ενός κουμπιού
	Omnipod	ταχείας δράσης ινσουλίνη 200U των U-100	ανεξάρτητη χρήση από Dexcom ΔΣΜΓ	Αδιάβροχο έως και 25 πόδια και διάρκεια 60 λεπτών. Υπενθυμίσεις και ειδοποιήσεις για δοσολογία, παροχή βιβλιοθήκης τροφίμων και αριθμομηχανή για υπολογισμό δόσης bolus

2.3 Μηχανική Μάθηση και Γλυκαιμικός Έλεγχος

Ένα συμπέρασμα που απορρέει από τα παραπάνω είναι ότι αποτελεί αδήριτη ανάγκη η ανάπτυξη συστημάτων, τα οποία θα παρέχουν στον πάσχοντα και στους θεράποντες ιατρούς μια ολοκληρωμένη εικόνα σχετικά με τον τρόπο που μεταβάλλονται τα επίπεδα γλυκόζης, ώστε να καθοριστεί το κατάλληλο σχήμα ινσουλίνης. Αυτό σε συνδυασμό με συγκεκριμένα γεύματα ή σωματική άσκηση εξασφαλίζει τη διατήρηση της γλυκόζης εντός επιτρεπτών τιμών. Ωστόσο, η μεταβλητότητα μεταξύ των ασθενών όσο και η μεταβλητότητα στη συμπεριφορά του ίδιου του ασθενούς απαιτεί την ανάπτυξη εξατομικευμένων συστημάτων υποστήριξης που θα έχουν ως στόχο την βελτίωση της αυτοδιαχείρισης του ΣΔ από τους ίδιους τους ασθενείς [41].

Η ευρεία εφαρμογή συστημάτων Τεχνητής Νοημοσύνης για τη συνεχή καταγραφή της γλυκόζης και για υποδόρια έγχυση ινσουλίνης και γλυκαγόνης βελτιώνει σημαντικά την καθημερινότητα των διαβητικών, εξασφαλίζοντας τους μια φυσιολογική ζωή. Μια καινοτόμα ανακάλυψη αποτελεί το «Τεχνητό Πάγκρεας». Αναπτύχθηκε με σκοπό να προσομοιωθεί πλήρως η λειτουργία του παγκρέατος και

άρα να επιτευχθεί η αυτόματη (χωρίς παρέμβαση του ασθενή) χορήγηση ινσουλίνης με τον βέλτιστο ρυθμό και με απόλυτα ελεγχόμενο τρόπο, όπως ακριβώς την εκκρίνει το πάγκρεας. Πρακτικά το «Τεχνητό Πάγκρεας» αποτελεί ένα σύστημα κλειστού βρόγχου μεταξύ των μετρητικών διατάξεων και των αντλιών έγχυσης ινσουλίνης. Συγκεκριμένα, οι πληροφορίες από τους αισθητήρες μέτρησης της γλυκόζης αξιοποιούνται από υπολογιστικό σύστημα και με την βοήθεια αλγορίθμου ελέγχου προβλέπεται η βέλτιστη ποσότητα ινσουλίνης που πρέπει να απελευθερώσει η αντλία [24].

Γενικότερα, η Μηχανική Μάθηση έχει ανοίξει τον δρόμο για την κατανόηση και την διαχείριση του μεγάλου όγκου δεδομένων που προέρχονται από αισθητήρες γλυκόζης και αντλίες ινσουλίνης. Πρακτικά, με την βοήθεια τεχνικών Μηχανικής Μάθησης, τα δεδομένα αυτά αναλύονται, ώστε να ομαδοποιούνται κατάλληλα και να ανακαλύπτονται συσχετίσεις μεταξύ τους, με σκοπό την παραγωγή χρήσιμων και λογικών συμπερασμάτων. Μερικές από τις εφαρμογές της, που παίζουν σημαντικό ρόλο στην καθημερινή αυτοδιαχείριση του ΣΔ, παρουσιάζονται παρακάτω.

Γλυκαιμικός Έλεγχος

Αρκετές μελέτες χρησιμοποιούν διάφορες προσεγγίσεις Μηχανικής Μάθησης για την αξιοποίηση των δεδομένων από τις ΔΣΜΓ, που στοχεύουν στην αυτοματοποίηση του ρυθμού έγχυσης ινσουλίνης και την σύσταση της κατάλληλης δόσης bolus ινσουλίνης [42][43][44]. Για παράδειγμα, οι εφαρμογές αναγνώρισης τροφίμων, οι οποίες χρησιμοποιούν μεθόδους βαθιάς Μηχανικής Μάθησης και ανάλυσης εικόνας, βοηθούν τα άτομα να λαμβάνουν την κατάλληλη δόση υδατανθράκων και την κατάλληλη δόση ινσουλίνης.

Πρόβλεψη γλυκαιμικών επεισοδίων

Πολλές εφαρμογές που είναι ήδη σε εμπορική χρήση χρησιμοποιούν τεχνικές Μηχανικής Μάθησης, ώστε να προβλέψουν επικείμενη υπογλυκαιμία ή υπεργλυκαιμία, με βάση τα δεδομένα γλυκόζης που συλλέγονται από τους αισθητήρες των ΔΣΜΓ, και να ειδοποιούν τον χρήστη [42]. Για παράδειγμα, η πρόβλεψη νυχτερινής υπογλυκαιμίας είναι πολύ σημαντική, διότι θα βοηθήσει τον χρήστη να κάνει τις κατάλληλες διορθωτικές κινήσεις (δόση υδατανθράκων, δόση ινσουλίνης), ώστε να αποφύγει τον κίνδυνο να πέσει σε κώμα.

Πρόβλεψη και διάγνωση επιπλοκών του ΣΔ

Συστήματα που αξιοποιούν βασικά κλινικά και βιοχημικά δεδομένα του κάθε ασθενούς, μπορούν να προβλέψουν τον κίνδυνο νεφροπάθειας, νευροπάθειας, αμφιβληστροειδοπάθειας ή καρδιαγγειακού συμβάντος, γεγονός που είναι πολύ σημαντικό καθώς τα παραπάνω αποτελούν πολύ σοβαρές επιπλοκές του ΣΔ και χρήζουν σοβαρή αντιμετώπιση [42].

3ο Κεφάλαιο

Θεωρητικό Υπόβαθρο

Εισαγωγή

Στην παρούσα εργασία, τα δεδομένα γλυκόζης των ασθενών που προκύπτουν από συστήματα συνεχούς παρακολούθησης (ΔΣΜΓ) θα αναλυθούν ως Δεδομένα Σύνθεσης (Compositional Data, CoDa) τα οποία, στη συνέχεια, θα αξιοποιηθούν ως δεδομένα εισόδου σε αλγορίθμους Μηχανικής Μάθησης. Στο Κεφάλαιο αυτό παρουσιάζεται το θεωρητικό υπόβαθρο των μεθόδων που χρησιμοποιήθηκαν για την ανάπτυξη μοντέλου κατηγοριοποίησης γλυκαιμικού προφίλ ατόμων με ΣΔΤ1.

3.1 Ανάλυση Δεδομένων Σύνθεσης

Ο John Aitchison ήταν ο πρώτος μαθηματικός που εισήγαγε την έννοια των Δεδομένων Σύνθεσης στο βιβλίο του με τίτλο *"The Statistical Analysis of Compositional Data"* το 1986. Πρόκειται για έναν κλάδο των μαθηματικών που ασχολείται με την ανάλυση δεδομένων σύνθεσης, δηλαδή διανυσμάτων αυστηρά θετικών στοιχείων που δείχνουν την συμβολή τμημάτων σε ένα γενικότερο σύνολο. Τα δεδομένα αυτά όταν αθροίζονται δίνουν πάντα ένα σταθερό αριθμό (π.χ μονάδα σε περίπτωση ποσοστών) [45][46]. Η μέθοδος βρίσκει εφαρμογή σε πολλές επιστήμες όπως η γεωλογία (γεωχημικά στοιχεία), η οικονομία (κατανομή δαπανών), η γενετική (συχνότητα γονότυπου) και η βιοϊατρική στην οποία επικεντρώνεται η παρούσα διπλωματική [46].

3.1.1 Βασικές Έννοιες

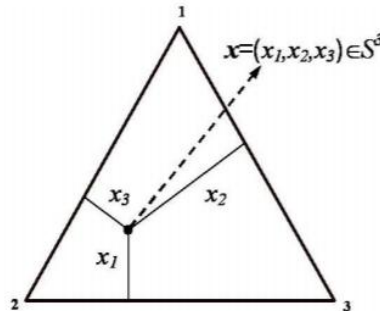
Ορισμοί

Ένα διάνυσμα $x = [x_1, x_2, \dots, x_D]$ με D τμήματα ορίζεται ως σύνθεση, όταν όλες οι συνιστώσες x_i είναι αυστηρά θετικοί αριθμοί και μεταφέρουν σχετική πληροφορία για το συγκεκριμένο τμήμα i . Οι συνιστώσες x_i όταν αθροίζονται δίνουν πάντα έναν σταθερό αριθμό.

Ο δειγματοχώρος των δεδομένων σύνθεσης είναι simplex διάστασης D και ορίζεται ως εξής [47]:

$$S^D = \{x = [x_1, x_2, \dots, x_D] \mid x_i > 0, i = 1, 2, \dots, D, \sum_{i=1}^D x_i = k\} \quad (3.1)$$

Το simplex, στην γεωμετρία, αποτελεί γενικευμένη έννοια τριγώνου ή τετραέδρου σε αυθαίρετες διαστάσεις. Αν $D = 3$ τότε το simplex είναι ένα ισόπλευρο τρίγωνο (Σχήμα 3.1).



Σχήμα 3.1: Απεικόνιση S^3 σε τριμερές διάγραμμα (Ternary) [48]

Κλειστότητα

Η πληροφορία που μεταφέρει ένα διάνυσμα σύνθεσης προκύπτει από τις αναλογίες των στοιχείων της και επομένως η ίδια πληροφορία μπορεί να εξαχθεί πολλαπλασιάζοντας αυτό το διάνυσμα με οποιαδήποτε θετική σταθερά. Κάθε διάνυσμα σύνθεσης έχει ένα πρότυπο δειγματοχώρο στο οποίο μπορεί να αναχθεί. Η κλειστότητα είναι η προβολή του διανύσματος με θετικές συνιστώσες από το \mathbb{R}_+^D στο S^D και ορίζεται ως εξής:

$$C[x_1, x_2, \dots, x_D] = \left[\frac{x_1}{\sum_{i=1}^D x_i}, \frac{x_2}{\sum_{i=1}^D x_i}, \dots, \frac{x_D}{\sum_{i=1}^D x_i} \right] \quad (3.2)$$

Δύο διαφορετικές συνθέσεις $x, y \in \mathbb{R}_+^D$ με D θετικές συνιστώσες ($x_i, y_i \geq 0, i = 1, 2, \dots, D$) είναι ισοδύναμες, αν υπάρχει θετική σταθερά $\lambda \in \mathbb{R}_+^D$ τέτοια ώστε $x = \lambda \cdot y$ και ισοδύναμα $C(x) = C(y)$ [47]. Για παράδειγμα, έστω δύο διανύσματα συνθέσεων $x = [1, 4, 8]$ με $\sum_{i=1}^3 x_i = 13$ και $y = [1.2, 4.8, 9.6]$ με $\sum_{i=1}^3 y_i = 15.6$. Παρατηρείται ότι υπάρχει αναλογία μεταξύ των δύο συνθέσεων, δηλαδή ισχύει ότι $y = \lambda \cdot x = \frac{6}{5} \cdot x$. Ακόμη, η κλειστότητα των δύο αυτών compositions είναι ίδια:

$$C[1.2, 4.8, 9.6] = \left[\frac{1.2}{15.6}, \frac{4.8}{15.6}, \frac{9.6}{15.6} \right] = \left[\frac{1}{13}, \frac{4}{13}, \frac{8}{13} \right] = C[1, 4, 8] \quad (3.3)$$

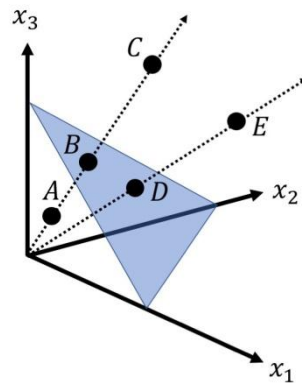
Συνεπώς, με βάση τον ορισμό που προηγήθηκε οι δύο συνθέσεις δεδομένων είναι ισοδύναμες συνθετικά.

Ανεξαρτησία Κλίμακας

Σύμφωνα με τον Aitchison μια συνάρτηση f είναι ανεξάρτητη της κλίμακας εάν για οποιαδήποτε θετική πραγματική τιμή $\lambda \in \mathbb{R}_+$ και για οποιαδήποτε σύνθεση $x \in S^D$ ικανοποιείται η σχέση:

$$f(\lambda x) = f(x), \lambda > 0 \quad (3.4)$$

Με άλλα λόγια η συνάρτηση πρέπει να αποδίδει το ίδιο αποτέλεσμα για όλες τις ισοδύναμες συνθέσεις, ανεξαρτήτως κλίμακας [47].



Σχήμα 3.2: Απεικόνιση ενός τριμερούς simplex S^3 . Υπάρχουν 3 τμήματα, x_1, x_2, x_3 που αντιπροσωπεύουν τμήματα κάποιου συνόλου. Τα A, B, C, D, E είναι πέντε διαφορετικές συνθέσεις εντός του simplex. Τα A, B, C είναι ισοδύναμα μεταξύ τους όπως και τα D, E . [45]

Γεωμετρικός Μέσος

Ο γεωμετρικός μέσος ενός διανύσματος σύνθεσης $x \in S^D$ υπολογίζεται από τον τύπο:

$$g(x) = \left(\prod_{i=1}^D x_i \right)^{1/D} = \exp \left(\frac{1}{D} \sum_{i=1}^D \ln x_i \right) \quad (3.5)$$

3.1.2 Γεωμετρία Aitchison

Η Ευκλείδεια γεωμετρία δεν είναι κατάλληλη για την ανάλυση και κατανόηση των δεδομένων σύνθεσης, διότι ο δειγματοχώρος τους δεν είναι ο Ευκλείδειος χώρος και επομένως δεν ισχύουν οι ιδιότητες και οι πράξεις του (πρόσθεση, αφαίρεση κ.α). Για να γίνει καλύτερα αντιληπτό αυτό, παραθέτεται ένα παράδειγμα. Έστω τέσσερις διαφορετικές συνθέσεις $x, y, z, t \in S^3$:

$$\begin{aligned} x &= [5, 48, 80], y = [10, 43, 85] \\ z &= [50, 25, 8], t = [55, 20, 13] \end{aligned}$$

Η Ευκλείδεια απόσταση μεταξύ των x, y και των z, t είναι ακριβώς ίδια αφού οι συνιστώσες μεταξύ των δύο διανυσμάτων διαφέρουν κατά 5 κάθε φορά. Ωστόσο, φαίνεται ότι στην πρώτη περίπτωση η αναλογία για το πρώτο στοιχείο είναι 2, δηλαδή διπλασιάζεται (αφού $y_1/x_1 = 2$) ενώ στην δεύτερη περίπτωση η αντίστοιχη αύξηση είναι της τάξης του 10% (αφού $t_1/z_1 = 1,1$) [47]. Έτσι, δημιουργήθηκε η

γεωμετρία Aitchison που αποτελείται από δικές τις ιδιότητες και πράξεις συνόλων κατάλληλες για την ανάλυση δεδομένων σύνθεσης. Ακολουθούν κάποιες βασικές πράξεις μεταξύ συνθέσεων στο χώρο simplex S [47].

Πρόσθεση

$$x \oplus y = C[x_1y_1, x_2y_2, \dots, x_Dy_D], \quad \forall x, y \in S^D \quad (3.6)$$

Πολλαπλασιασμός

$$a \odot x = C[x_1^a, x_2^a, \dots, x_D^a], \quad \forall x \in S^D, a \in \mathbb{R} \quad (3.7)$$

Για να χαρακτηριστεί ένας χώρος ως διανυσματικός θα πρέπει οι πράξεις της πρόσθεσης (perturbation) και του βαθμωτού πολλαπλασιασμού (power transformation) να ικανοποιούν κάποια αξιώματα. Στην περίπτωση των δεδομένων σύνθεσης ο χώρος S αποτελεί διανυσματικό χώρο αφού ικανοποιούνται τα παρακάτω αξιώματα.

Αξιώματα Πρόσθεσης

- **Αντιμεταθετική Ιδιότητα**

$$x \oplus y = y \oplus x \quad (3.8)$$

- **Προσεταιριστική Ιδιότητα**

$$(x \oplus y) \oplus z = x \oplus (y \oplus z) \quad (3.9)$$

- **Ουδέτερο Στοιχείο**

$$n = C[1, 1, \dots, 1] = \left[\frac{1}{D}, \frac{1}{D}, \dots, \frac{1}{D} \right] \quad (3.10)$$

Το n αποτελεί το βαρύκεντρο του χώρου S^D και είναι μοναδικό.

- **Αντίθετο Στοιχείο**

Υπάρχει αντίθετο στοιχείο x^{-1} που αν προστεθεί στο x δίνει το ουδέτερο στοιχείο n , δηλαδή $x \oplus x^{-1} = n$

$$x^{-1} = C[x_1^{-1}, x_2^{-1}, \dots, x_D^{-1}] \quad (3.11)$$

Έτσι, προκύπτει και η αντίστοιχη με την αφαίρεση στον Ευκλείδειο χώρο πράξη:

$$x \ominus y = x \oplus y^{-1} = C \left[\frac{x_1}{y_1}, \frac{x_2}{y_2}, \dots, \frac{x_D}{y_D} \right], \quad \forall x, y \in S^D \quad (3.12)$$

Αξιώματα Πολλαπλασιασμού

- Προσεταιριστική Ιδιότητα

$$a \odot (\beta \odot x) = (\alpha \cdot \beta) \odot x \quad (3.13)$$

- Επιμεριστικές Ιδιότητες

$$(\alpha + \beta) \odot x = (\alpha \odot x) \oplus (\beta \odot x) \quad (3.14)$$

$$a \odot (x \oplus y) = (a \odot x) \oplus (a \odot y) \quad (3.15)$$

- Ουδέτερο Στοιχείο

$$1 \odot x = x \quad (3.16)$$

Με βάση τα παραπάνω, το Aitchison Simplex μπορεί να θεωρηθεί ότι σχηματίζει ένα διαστατικό χώρο Hilbert. Ένας χώρος Hilbert είναι ένας αφηρημένος διανυσματικός χώρος εφοδιασμένος με εσωτερικό γινόμενο, το οποίο τον καθιστά πλήρως μετρήσιμο. Έτσι, για τον διανυσματικό χώρο S^D ορίζονται τα εξής:

Εσωτερικό Γινόμενο

$$\langle x, y \rangle_a = \frac{1}{2D} \sum_{i=1}^D \sum_{j=1}^D \ln \frac{x_i}{x_j} \ln \frac{y_i}{y_j}, \quad \forall x, y \in S^D \quad (3.17)$$

Νόρμα

$$\|x\|_a = \sqrt{\frac{1}{2D} \sum_{i=1}^D \sum_{j=1}^D \left(\ln \frac{x_i}{x_j} \right)^2}, \quad \forall x \in S^D \quad (3.18)$$

Απόσταση

$$d_a(x, y) = \sqrt{\frac{1}{2D} \sum_{i=1}^D \sum_{j=1}^D \left(\ln \frac{x_i}{x_j} - \ln \frac{y_i}{y_j} \right)^2}, \quad \forall x, y \in S^D \quad (3.19)$$

Για την απόσταση ισχύουν οι παρακάτω ιδιότητες:

$$d_a(p \oplus x, p \oplus y) = d_a(x, y), \quad \forall x, y \in S^D \text{ και } p \in \mathbb{R} \quad (3.20)$$

$$d_a(\beta \odot x, \beta \odot y) = |\beta| d_a(x, y), \quad \forall x, y \in S^D \text{ και } \beta \in \mathbb{R} \quad (3.21)$$

3.1.3 Γραμμικοί Μετασχηματισμοί

Η στατιστική ανάλυση στο simplex μπορεί να μεταφερθεί στον Ευκλείδειο χώρο μέσω μετασχηματισμών των διανυσμάτων σύνθεσης σε αναλογίες λογαρίθμων στις οποίες εμπεριέχονται οι βασικές πληροφορίες τους και μπορούν να εφαρμοστούν οι κλασικές στατιστικές μέθοδοι. Παρακάτω παρουσιάζονται τρεις βασικοί γραμμικοί μετασχηματισμοί από το Aitchison simplex στον Ευκλείδειο χώρο.

Για την κατανόηση των μετασχηματισμών παραθέτονται, επίσης, κάποιοι βασικοί ορισμοί της άλγεβρας [49].

Ορισμός (Ισομορφισμός):

Ένας **ισομορφισμός** $f: V \rightarrow W$ είναι ένας ομομορφισμός ένα-προς-ένα και επί του W . Δηλαδή, η συνάρτηση f είναι 1-1 και άρα δύο διαφορετικά στοιχεία v_1, v_2 του V θα πρέπει να έχουν διαφορετικά ταίρια $w_1 = f(v_1)$ και $w_2 = f(v_2)$ στο W . Επίσης, κάθε στοιχείο w του W πρέπει να είναι το ταίρι κάποιου στοιχείου v του V , άρα να ισχύει $f(v) = w$.

Ορισμός (Ισομετρία):

Αν V είναι ένα σύνολο στο οποίο έχει οριστεί μια έννοια απόστασης για τα στοιχεία του, ένας μετασχηματισμός f του V λέγεται **ισομετρία** αν η f διατηρεί τις αποστάσεις, δηλαδή αν ισχύει η σχέση $d(x, y) = d(f(x), f(y))$.

Μετασχηματισμός alr

Έστω ένα διάνυσμα $x = [x_1, x_2, \dots, x_D], x \in S^D$, ο προσθετικός λογαριθμικός μετασχηματισμός alr είναι ένας ισομορφισμός $alr: S^D \rightarrow R^{D-1}$ που ορίζεται ως εξής:

$$y = alr(x) = \left[\log\left(\frac{x_1}{x_D}\right), \log\left(\frac{x_2}{x_D}\right), \dots, \log\left(\frac{x_{D-1}}{x_D}\right) \right] \quad (3.22)$$

Ο αντίστροφος μετασχηματισμός $alr^{-1}: R^{D-1} \rightarrow S^D$:

$$x = alr^{-1}(y) = C[\exp(y_1), \exp(y_2), \dots, \exp(y_{D-1}), 1] \quad (3.23)$$

όπου C η πράξη της κλειστότητας (σχέση 3.2).

Το κύριο μειονέκτημα του μετασχηματισμού αυτού είναι ότι ο alr μετασχηματισμός δεν αποτελεί ισομετρία, που σημαίνει ότι οι αποστάσεις των μετασχηματισμένων

τιμών δεν θα είναι ισοδύναμες με τις αποστάσεις των διανυσμάτων στο S^D . Δηλαδή, $d(\text{alr}(x), \text{alr}(y)) \neq d_\alpha(x, y)$ όπου $x, y \in S^D$.

Μετασχηματισμός clr

Έστω ένα διάνυσμα $x = [x_1, x_2, \dots, x_D]$, $x \in S^D$, ο κεντρικός λογαριθμικός μετασχηματισμός clr είναι ένας ισομορφισμός $\text{clr}: S^D \rightarrow U^D$, $U^D \subset R^D$ που ορίζεται ως εξής:

$$z = \text{clr}(x) = \left[\log\left(\frac{x_1}{g(x)}\right), \log\left(\frac{x_2}{g(x)}\right), \dots, \log\left(\frac{x_D}{g(x)}\right) \right] \quad (3.24)$$

όπου $g(x)$ γεωμετρικός μέσος της σύνθεσης. Ο χώρος U ορίζεται ως ένα υπερπλάνο του R^D και ορίζεται ως $U^D = \{[u_1, \dots, u_D], u_1 + u_2 + \dots + u_D = 0\}$. Αυτό σημαίνει ότι τα μετασχηματισμένα δεδομένα z_i θα αθροίζονται στο 0 [50].

Ο αντίστροφος μετασχηματισμός $\text{clr}^{-1}: U^D \rightarrow S^D$ έχει την μορφή:

$$x = \text{clr}^{-1}(z) = C[\exp(z_1), \exp(z_2), \dots, \exp(z_D)] \quad (3.25)$$

όπου C η πράξη της κλειστότητας (σχέση 3.2).

Σε αντίθεση με τον alr μετασχηματισμό, ο clr είναι ισομετρικός μετασχηματισμός του S^D με την μετρική Aitchison, σε ένα υποσύνολο U^D του πραγματικού χώρου με την συνηθισμένη Ευκλείδεια μετρική. Επομένως τα μετασχηματισμένα δεδομένα μπορούν να χρησιμοποιηθούν για την εφαρμογή στατιστικών μεθόδων. Γενικά, για τον clr μετασχηματισμό ισχύουν οι παρακάτω ιδιότητες:

$$\langle x_1, x_2 \rangle_a = \langle \text{clr}(x_1), \text{clr}(x_2) \rangle \quad (3.26\alpha)$$

$$\|x_1\|_a = \|\text{clr}(x_1)\| \quad (3.26\beta)$$

$$d_a(x_1, x_2) = d(\text{clr}(x_1), \text{clr}(x_2)) \quad (3.26\gamma)$$

Μετασχηματισμός ilr

Κανένας από τους παραπάνω μετασχηματισμούς δεν μπορεί να συνδεθεί άμεσα με ένα ορθογώνιο σύστημα συντεταγμένων στο Simplex [47]. Για το λόγο αυτό, αναπτύχθηκε ένας νέος μετασχηματισμός, ονόματι ilr (isometric log-ratio), ο οποίος αποτελεί ισομορφισμό και ισομετρία μεταξύ των χώρων S^D και \mathbb{R}^{D-1} , δηλαδή ισχύει $\text{ilr}: S^D \rightarrow \mathbb{R}^{D-1}$.

Έστω, μια ορθοκανονική βάση $\{e_1, e_2, \dots, e_D\}$ στο Simplex S^D και ο $(D-1, D)$ πίνακας Φ , του οποίου οι γραμμές είναι οι συντελεστές clr των διανυσμάτων e_i (δηλαδή $\Phi = \text{clr}(e_i)$).

Μία ορθοκανονική βάση ικανοποιεί την σχέση:

$$\langle e_i, e_j \rangle_a = \delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} \quad (3.27)$$

Η σχέση 3.27 με βάση την σχέση 3.26α γράφεται ως εξής:

$$\langle e_i, e_j \rangle_a = \langle clr(e_i), clr(e_j) \rangle = \delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} \quad (3.28)$$

Δεδομένου ότι έχει επιλεγεί μια ορθοκανονική βάση $\{e_1, e_2, \dots, e_D\}$, μια σύνθεση $x \in S^D$ μπορεί να εκφραστεί ως:

$$x = \bigoplus_{i=1}^{D-1} x_i^* \odot e_i, \quad x_i^* = \langle x, e_i \rangle_a \quad (3.29)$$

όπου $x^* = [x_1^*, x_2^*, \dots, x_{D-1}^*]$ είναι οι ορθοκανονικές (και καρτεσιανές) συντεταγμένες του x σε σχέση με την επιλεγμένη βάση και αποτελούν τους συντελεστές του ilr μετασχηματισμού.

Έτσι, λοιπόν, ο μετασχηματισμός $ilr: S^D \rightarrow \mathbb{R}^{D-1}$ ορίζεται ως εξής:

$$x^* = ilr(x) = [\langle x, e_1 \rangle_a, \langle x, e_2 \rangle_a, \dots, \langle x, e_{D-1} \rangle_a] \quad (3.30)$$

Από την σχέση 3.28 προκύπτει η ισοδυναμία:

$$\begin{aligned} ilr(x) &= [\langle clr(x), clr(e_1) \rangle, \langle clr(x), clr(e_2) \rangle, \dots, \langle clr(x), clr(e_{D-1}) \rangle] \\ &\Leftrightarrow x^* = ilr(x) = clr(x)\Phi^T \end{aligned} \quad (3.31)$$

όπου Φ ο πίνακας του οποίου η κάθε γραμμή είναι το διάνυσμα $clr(e_i)$, για $i = 1, 2, \dots, D-1$, όπως αναφέρθηκε και παραπάνω.

Ο αντίστροφος μετασχηματισμός $ilr: \mathbb{R}^{D-1} \rightarrow S^D$ δίνεται ως εξής:

$$\begin{aligned} ilr(x) = clr(x)\Phi^T &\Leftrightarrow ilr(x)\Phi = clr(x) \Leftrightarrow x = clr^{-1}(ilr(x)\Phi) \\ x &= C[\exp(ilr(x)\Phi)] \end{aligned} \quad (3.32)$$

Για τον μετασχηματισμό ilr ισχύουν οι παρακάτω ιδιότητες:

$$\langle x_1, x_2 \rangle_a = \langle ilr(x_1), ilr(x_2) \rangle \quad (3.33\alpha)$$

$$\|x_1\|_a = \|ilr(x_1)\| \quad (3.33\beta)$$

$$d_a(x_1, x_2) = d(ilr(x_1), ilr(x_2)) \quad (3.33\gamma)$$

Η διαφορά μεταξύ των ιδιοτήτων για το clr (σχέσεις 3.26) και των ιδιοτήτων για το ilr (σχέσεις 3.33) είναι ότι οι πρώτες αναφέρονται σε διανύσματα στο \mathbb{R}^D , ενώ οι δεύτερες σε διανύσματα στο χώρο \mathbb{R}^{D-1} .

Με βάση όλα τα παραπάνω, για να υπολογιστούν οι ilr συντελεστές ενός διανύσματος σύνθεσης πρέπει πρώτα να καθοριστεί μια ορθοκανονική βάση. Μπορούν να χρησιμοποιηθούν μέθοδοι ορθοκανονικοποίησης από τα μετασχηματισμένα clr δεδομένα, όπως η μέθοδος Gram-Schmidt και η μέθοδος ανάλυσης πίνακα σε ιδιάζουσες τιμές (Singular Value Decomposition, SVD) [45]. Ενδιαφέρον προσέγγιση αποτελεί η δημιουργία συστήματος συντεταγμένων ilr μέσω διαδοχικής δυαδικής διχοτόμησης (Sequential Binary Partition, SBP) ενός διανύσματος σύνθεσης [47]. Η διαδικασία που ακολουθείται περιγράφεται στη συνέχεια. Ένα παράδειγμα του διαδοχικού διαχωρισμού φαίνεται στο **Σχήμα 3.3**.

Οι καρτεσιανές συντεταγμένες μιας σύνθεσης σε τέτοια βάση ονομάζονται ισορροπίες και τα διανύσματα της βάσης ισορροπημένα στοιχεία. Η διαδοχική δυαδική διχοτόμηση αποτελεί μια ιεραρχία των τμημάτων μιας σύνθεσης. Αρχικά, τα τμήματα της σύνθεσης διαχωρίζονται σε δύο ομάδες (ομάδα με +1 και -1 αντίστοιχα στον πίνακα προσήμου). Στα επόμενα βήματα της διαδικασίας κάθε ομάδα διαχωρίζεται σε δύο υποομάδες και η διαδικασία συνεχίζεται μέχρι να μην υφίσταται άλλος διαχωρισμός. Σε κάθε βήμα, δηλαδή σε κάθε επίπεδο της ιεραρχίας, καθορίζεται η ισορροπία μεταξύ των δύο υποομάδων με βάση τον παρακάτω τύπο:

$$b = \sqrt{\frac{rs}{r+s}} \ln \frac{(x_{i1}, x_{i2}, \dots, x_{ir})^{\frac{1}{r}}}{(x_{j1}, x_{j2}, \dots, x_{js})^{\frac{1}{s}}} = \ln \frac{(x_{i1}, x_{i2}, \dots, x_{ir})^{a_+}}{(x_{j1}, x_{j2}, \dots, x_{js})^{a_-}} \quad (3.34)$$

όπου r και s το πλήθος των τμημάτων που αποτελείται κάθε υποομάδα σε κάθε επίπεδο. Τα βάρη που θα παίρνουν τα τμήματα των δύο υποομάδων κάθε φορά θέτονται a_+ και a_- και ισούνται με:

$$a_+ = +\frac{1}{r} \sqrt{\frac{rs}{r+s}}, \quad a_- = -\frac{1}{s} \sqrt{\frac{rs}{r+s}} \quad \text{και} \quad a_0 = 0 \quad (3.35)$$

Έτσι, για κάθε i -οστό επίπεδο της ιεραρχίας προκύπτει η συνιστώσα:

$$b_i = \sum_{j=1}^D a_{ij} \ln x_j \quad (3.36)$$

όπου a_{ij} ισούται με:

- a_+ , αν στο i -οστό επίπεδο της διχοτόμησης ο πίνακας προσήμου έχει +1 στο j -οστό x_j τμήμα της σύνθεσης
- a_- , αν στο i -οστό επίπεδο της διχοτόμησης ο πίνακας προσήμου έχει -1 στο j -οστό x_j τμήμα της σύνθεσης
- a_0 , αν στο i -οστό επίπεδο της διχοτόμησης ο πίνακας προσήμου έχει τιμή 0 στο j -οστό x_j τμήμα της σύνθεσης

Ο πίνακας a_{ij} πρακτικά αποτελεί τον πίνακα Φ της σχέσης 3.31. Οι συντελεστές ισορροπίας b_i αποτελούν τις ilr συνιστώσες.

order	x_1	x_2	x_3	x_4	x_5	x_6	Γ	s
1	+1	+1	-1	-1	+1	+1	4	2
2	+1	-1	0	0	-1	-1	1	3
3	0	+1	0	0	-1	-1	1	2
4	0	0	0	0	+1	-1	1	1
5	0	0	-1	+1	0	0	1	1

1	$+\frac{1}{\sqrt{12}}$	$+\frac{1}{\sqrt{12}}$	$-\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$	$+\frac{1}{\sqrt{12}}$	$+\frac{1}{\sqrt{12}}$		
2	$+\frac{\sqrt{3}}{2}$	$-\frac{1}{\sqrt{12}}$	0	0	$-\frac{1}{\sqrt{12}}$	$-\frac{1}{\sqrt{12}}$		
3	0	$+\frac{\sqrt{2}}{\sqrt{3}}$	0	0	$-\frac{1}{\sqrt{6}}$	$-\frac{1}{\sqrt{6}}$		
4	0	0	0	0	$+\frac{1}{\sqrt{2}}$	$-\frac{1}{\sqrt{2}}$		
5	0	0	$+\frac{1}{\sqrt{2}}$	0	0	$-\frac{1}{\sqrt{2}}$		

Σχήμα 3.3: Παράδειγμα του πίνακα προσήμου για $D = 6$, που χρησιμοποιείται για να κωδικοποιήσει μία διαδοχική δυαδική διχοτόμηση της σύνθεσης $x = [x_1, x_2, \dots, x_D]$ για την δημιουργία ορθοκανονικής βάσης. Το κάτω τμήμα του πίνακα δείχνει τον πίνακα Φ [47]

3.2 Μηχανική Μάθηση

Η Μηχανική Μάθηση αποτελεί βασικό εργαλείο της Επιστήμης Δεδομένων και μπορεί να θεωρηθεί ως η δημιουργία μοντέλων ή προτύπων από ένα σύνολο δεδομένων, από ένα υπολογιστικό σύστημα. Αυτό προσομοιώνει την ανθρώπινη συμπεριφορά όπου ο άνθρωπος προσπαθεί να κατανοήσει το περιβάλλον του παρατηρώντας το και έχει την δυνατότητα να συσχετίζει τις εμπειρίες του και τις παραστάσεις του, δημιουργώντας κατά κάποιον τρόπο μοντέλα και μοτίβα.

Γενικότερα, η Μηχανική Μάθηση αποτελεί μια περιοχή της Τεχνητής Νοημοσύνης που μελετά αλγορίθμους υπολογιστών που μπορούν να βελτιώνονται αυτόματα μέσω της εμπειρίας. Σύμφωνα με τον Mitchell (1997) "Ένα πρόγραμμα υπολογιστή θεωρείται ότι μαθαίνει από την εμπειρία E σε σχέση με μια κατηγορία εργασιών T και μια μετρική απόδοσης P , αν η απόδοση του σε εργασίες της T , όπως μετριοούνται από την P , βελτιώνονται με την εμπειρία E " [51].

Ανάλογα με την φύση του προβλήματος και τον σκοπό της μελέτης που πραγματοποιείται, χρησιμοποιούνται διάφορες τεχνικές Μηχανικής Μάθησης που έχουν αναπτυχθεί. Υπάρχουν δύο βασικά είδη Μηχανικής Μάθησης στα οποία χρησιμοποιούνται διαφορετικές τεχνικές και προσεγγίζεται διαφορετικά το κάθε πρόβλημα. Στη συνέχεια ακολουθεί η ανάλυση τους καθώς και αλγορίθμων που χρησιμοποιούν.

3.2.1 Μάθηση με Επίβλεψη

Στη μάθηση με επίβλεψη το σύστημα καλείται να "μάθει" μια έννοια ή μια συνάρτηση από ένα σύνολο δεδομένων. Ονομάζεται έτσι επειδή θεωρείται ότι υπάρχει κάποιος "επιβλέπων" ο οποίος παρέχει την σωστή τιμή εξόδου της συνάρτησης για τα δεδομένα που εξετάζονται [51]. Τα δεδομένα για τα οποία έχει καθοριστεί η τιμή της μεταβλητής εξόδου χρησιμοποιούνται για την εκπαίδευση του

μοντέλου με σκοπό το σύστημα να μάθει έναν γενικό κανόνα που να χαρτογραφεί τις εισόδους στις εξόδους. Το σύστημα εκπαιδεύεται, όσον το δυνατόν καλύτερα, με σκοπό να μπορεί να γενικεύει και να προβλέπει την έξοδο για άγνωστα δεδομένα εισόδου.

Σε αυτό το είδος Μηχανικής Μάθησης εντάσσονται δύο είδη προβλημάτων τα προβλήματα **ταξινόμησης** και **παρεμβολής**. Τα προβλήματα της ταξινόμησης αφορούν στη δημιουργία μοντέλων πρόβλεψης διακριτών κλάσεων, όπως για παράδειγμα αν ένας όγκος, με κάποια χαρακτηριστικά, ταξινομείται ως καλοήθης ή ως κακοήθης. Αντίθετα, τα προβλήματα παρεμβολής αφορούν στη δημιουργία μοντέλων πρόβλεψης αριθμητικών τιμών, όπως για παράδειγμα την πρόβλεψη του βαθμού ενός μαθητή σε ένα μάθημα. Από τις κυριότερες τεχνικές Μηχανικής Μάθησης με επίβλεψη είναι τα **Δένδρα Απόφασης**, η **Μάθηση κατά Bayes**, η **Γραμμική Παρεμβολή** και οι **Μηχανές Διανυσμάτων Υποστήριξης** [51].

3.2.2 Μάθηση χωρίς Επίβλεψη

Σε αντίθεση με την μάθηση με επίβλεψη, στη μάθηση χωρίς επίβλεψη δεν δίνονται ετικέτες εξόδου στον αλγόριθμο εκμάθησης. Επομένως, το σύστημα καλείται να ανακαλύψει μόνο του συσχετίσεις ή ομάδες σε ένα σύνολο δεδομένων, βασιζόμενο μόνο στις ιδιότητες τους. Αυτό γίνεται με σκοπό να προκύψουν μοτίβα-πρότυπα των δεδομένων, χωρίς, ωστόσο, να είναι γνωστό αν υπάρχουν, πόσα και ποια είναι [51].

Παραδείγματα προτύπων πληροφόρησης είναι οι **κανόνες συσχέτισης** και οι **συστάδες**. Οι κανόνες συσχέτισης επιτρέπουν την δημιουργία συσχετίσεων μεταξύ των αντικειμένων μέσα σε μεγάλες βάσεις δεδομένων. Για παράδειγμα, ένα πρόβλημα θα ήταν να βρεθούν ποια αντικείμενα, από μια βάση δεδομένων σούπερ μάρκετ, αγοράζονται μαζί (δηλαδή στο ίδιο καλάθι αγορών) ή μια ομάδα ταινιών που βαθμολογούνται παρόμοια από τους θεατές. Οι συστάδες είναι ομάδες με κοινά χαρακτηριστικά, οι οποίες προκύπτουν από την διαδικασία της συσταδοποίησης, δηλαδή μια διαδικασία που ασχολείται με την εύρεση δομής ή μοτίβου σε μια συλλογή με μη κατηγοριοποιημένα δεδομένα. Ένα παράδειγμα προβλήματος είναι η συσταδοποίηση ασθενών με βάση το ιστορικό τους [51][52]. Μερικοί από τους πιο συνηθισμένους αλγορίθμους που χρησιμοποιούνται στη μηχανική μάθηση χωρίς επίβλεψη περιλαμβάνουν **Συσταδοποίηση**, **Ανίχνευση ανωμαλιών** και **Νευρωνικά Δίκτυα**. Κάθε προσέγγιση χρησιμοποιεί διάφορες μεθόδους και τεχνικές.

Στη παρούσα διπλωματική εργασία χρησιμοποιήθηκαν, κυρίως, τεχνικές Μηχανικής Μάθησης χωρίς επίβλεψη, δεδομένου του ότι τα δεδομένα εισόδου ήταν μη χαρακτηρισμένα παραδείγματα. Έτσι, η μελέτη επικεντρώθηκε στην χρήση αλγορίθμου συσταδοποίησης για την εύρεση διακριτών ομάδων με κλινικά σημαντικά χαρακτηριστικά. Στη συνέχεια, η κατηγοριοποίηση που προέκυψε χρησιμοποιήθηκε για την δημιουργία μοντέλου ταξινόμησης που θα προβλέπει σε ποια από τις κατηγορίες ανήκει μια άγνωστη είσοδος. Επομένως, κρίνεται αναγκαίο να προηγηθεί ανάλυση μεθόδων και αλγορίθμων συσταδοποίησης και ταξινόμησης.

3.2.3 Συσταδοποίηση

Η ανάλυση συστάδων ή αλλιώς συσταδοποίηση των στοιχείων, είναι η διαδικασία κατά την οποία ένα σύνολο (συνήθως πολυδιάστατων) δεδομένων διαχωρίζεται σε ομάδες, έτσι ώστε σημεία που ανήκουν στην ίδια ομάδα να μοιάζουν όσο περισσότερο γίνεται ως προς κάποια χαρακτηριστικά και σημεία από διαφορετικές ομάδες να διαφέρουν ως προς τα ίδια χαρακτηριστικά [51]. Πρακτικά, οι ομάδες είναι πυκνές περιοχές με σημεία στο χώρο. Η ταξινόμηση κάθε σημείου σε μια τέτοια ομάδα είναι μοναδική και γίνεται με τρόπο τέτοιο ώστε οι αποστάσεις των στοιχείων μέσα στην ίδια ομάδα να ελαχιστοποιούνται, ενώ οι αποστάσεις μεταξύ των διαφορετικών ομάδων να μεγιστοποιούνται.

Η διαδικασία διαμερισμού ενός συνόλου δεδομένων πραγματοποιείται με τη χρήση πολυάριθμων αλγορίθμων με διαφορετικές ιδιότητες ως προς τον τρόπο λειτουργίας τους και τον τρόπο αξιολόγησης. Ανάλογα με την εφαρμογή επιλέγεται ο καταλληλότερος αλγόριθμος. Ωστόσο, πέραν της προσεκτικής επιλογής αλγορίθμου, η ρύθμιση ορισμένων παραμέτρων, όπως η μετρική απόστασης ή ο αριθμός των συστάδων, είναι εξίσου σημαντική διαδικασία. Επομένως, απαιτείται διαδικασία βελτιστοποίησης των παραμέτρων, με την εφαρμογή δοκιμασίας και αποτυχίας, με σκοπό να καθοριστούν οι κατάλληλοι και πιο αποδοτικοί παράμετροι.

Σε περιπτώσεις που η διαστατικότητα είναι πρόβλημα, πραγματοποιείται μια παραγοντική ανάλυση για να περιορίσει τις διαστάσεις των δεδομένων και κατ' επέκταση των μεταβλητών, μεταφέροντάς πρακτικά τα δεδομένα σε χώρο μικρότερης διάστασης. Αυτή η διαδικασία διευκολύνει ποιοτικά την εκτέλεση των αλγορίθμων και είναι αρκετά χρήσιμη για την απεικόνιση των δεδομένων στον δισδιάστατο χώρο και άρα την καλύτερη διαισθητική ερμηνεία τους. Οι τεχνικές που χρησιμοποιούνται για μείωση των διαστάσεων είναι είτε η μέθοδος ανάλυσης πίνακα σε ιδιάζουσες τιμές (Singular Value Decomposition, SVD), είτε η μέθοδος ανάλυσης κύριων συνιστωσών (Principal Components Analysis, PCA) [53].

Η μέθοδος PCA αποτελεί μια γραμμική μέθοδο συμπίεσης δεδομένων, η οποία συνίσταται από τον επαναπροσδιορισμό των συντεταγμένων ενός συνόλου δεδομένων σε ένα άλλο σύστημα συντεταγμένων το οποίο θα είναι καταλληλότερο για την επικείμενη ανάλυση τους. Η μέθοδος αυτή υπολογίζει μια νέα βάση του χώρου η οποία αποτελεί το νέο σύνολο από μεταβλητές που είναι γραμμικά ασυσχέτιστες και ονομάζονται κύριες συνιστώσες. Ο αριθμός των κύριων συνιστωσών επιλέγεται τέτοιος ώστε να μην χάνεται σημαντική πληροφορία και το ποσοστό διακύμανσης να είναι όσο τον δυνατόν μεγαλύτερο. Η μέθοδος SVD είναι μια παρόμοια διαδικασία που αποσυνθέτει τα δεδομένα σε ανεξάρτητες συνιστώσες. Μάλιστα μπορεί να χρησιμοποιηθεί για να βρεθούν τα κυρίαρχα συστατικά της μεθόδου PCA, περικόποντας τα λιγότερο σημαντικά διανύσματα της βάσης στην αρχική μήτρα SVD [54].

Μετά την εφαρμογή των αλγορίθμων που πραγματοποιούν τον διαχωρισμό σε συστάδες ακολουθεί μετά-ανάλυση με σκοπό να περιγραφεί αν υπάρχει στατιστικά σημαντική διαφορά μεταξύ των συστάδων. Η ανάλυση αυτή κρίνεται απαραίτητη,

διότι οι μέθοδοι Μηχανικής Μάθησης χωρίς επίβλεψη δεν διαθέτουν κριτήρια μέτρησης της σωστής προσαρμογής των μοντέλων και η αξιολόγηση τους βασίζεται στην ερμηνεία των χαρακτηριστικών που διαφοροποιούν τις ομάδες [53].

Οι δύο γενικές κατηγορίες αλγορίθμων συσταδοποίησης είναι οι εξής:

- Οι **ιεραρχικοί αλγόριθμοι**, που προσπαθούν με ιεραρχικό τρόπο να ανακαλύψουν την δομή και τον αριθμό των ομάδων.
- Οι αλγόριθμοι που βασίζονται σε **διαχωρισμούς**, που προσπαθούν να βρουν τον βέλτιστο διαμερισμό ενός συνόλου δεδομένων σε ένα συγκεκριμένο αριθμό ομάδων.

Αλγόριθμοι Ιεραρχικής Συσταδοποίησης

Η ιεραρχική ομαδοποίηση βασίζεται στην ιδέα ότι κάποια στοιχεία σχετίζονται περισσότερο με κάποια γειτονικά τους παρά με άλλα σε μεγαλύτερη απόσταση από αυτά. Αυτοί οι αλγόριθμοι συνδυάζουν ομάδες σε μεγαλύτερες ομάδες ή διαιρούν μεγάλες ομάδες σε μικρότερες. Έτσι, το αποτέλεσμα είναι μια ιεραρχία από διαφορετικές ομαδοποιήσεις των δεδομένων, στο ένα άκρο της οποίας βρίσκεται μόνο μια ομάδα που αποτελείται από όλα τα δεδομένα και στο άλλο άκρο τόσες ομάδες όσα και τα δεδομένα του συνόλου. Οι μέθοδοι που χρησιμοποιούνται διαφέρουν ανάλογα με τον τρόπο υπολογισμού των αποστάσεων μεταξύ των στοιχείων. Ακόμη, ο χρήστης πρέπει πέρα από το κριτήριο απόστασης, να αποφασίσει σχετικά με το κριτήριο συνένωσης των ομάδων μεταξύ τους. Οι αλγόριθμοι αυτής της κατηγορίας διαχωρίζονται στους αλγόριθμους **συγχώνευσης** όπου τα αρχικά στοιχεία συγχωνεύονται σε ολοένα μεγαλύτερες ομάδες, και στους αλγορίθμους **διαίρεσης**, όπου από ένα ενιαίο σύνολο τα δεδομένα διαιρούνται σε μικρότερες ομάδες.

Οι πιο διαδεδομένοι αλγόριθμοι ιεραρχικής συσταδοποίησης είναι οι αλγόριθμοι συγχώνευσης. Ο αλγόριθμος ξεκινά με κάθε σημείο να αποτελεί μόνο του μια ομάδα, το οποίο κατόπιν ενώνεται σταδιακά με τα πιο κοντινά του σημεία. Η διαδικασία αυτή συνεχίζεται με την λογική ελαχιστοποίησης ενός κριτηρίου διασύνδεσης, έως ότου προκύψει μια μεγάλη ομάδα. Για τον υπολογισμό της απόστασης μεταξύ δύο σημείων, τυπικά μέτρα απόστασης είναι η απόσταση Μανχάταν και η Ευκλείδεια απόσταση οι οποίες ορίζονται ως εξής:

Απόσταση Μανχάταν

$$d(x, y) = \sum_{i=1}^m |x_i - y_i| \quad (3.37)$$

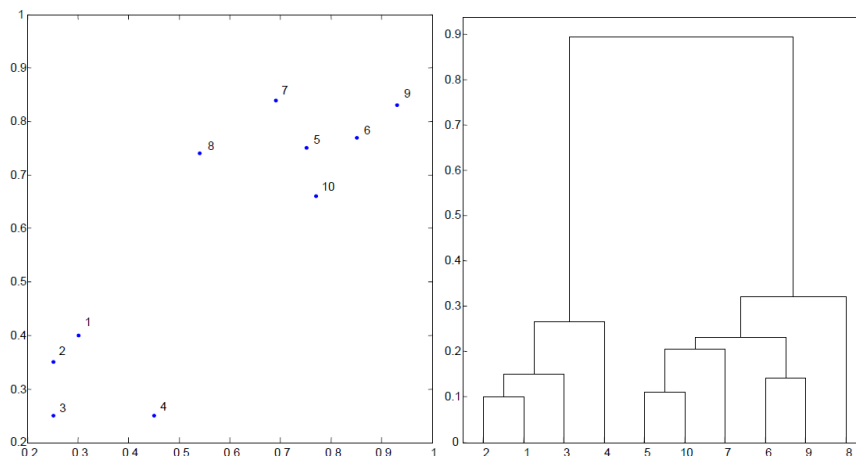
Ευκλείδεια Απόσταση

$$d(x, y) = \sqrt{\sum_{i=1}^m (x_i - y_i)^2} \quad (3.38)$$

Για τον υπολογισμό απόστασης μεταξύ των ομάδων θα μπορούσε να υπολογισθεί είτε η μέση απόσταση μεταξύ των ζευγών των σημείων που ανήκουν στην κάθε

ομάδα (μέση διασύνδεση ή average linkage), είτε η απόσταση μεταξύ των δύο πιο κοντινών (απλή διασύνδεση ή single linkage) ή πιο απομακρυσμένων σημείων (πλήρη διασύνδεση ή complete linkage) των ομάδων [51][53][55].

Για την απεικόνιση των ιεραρχικής δομής που προκύπτει από την εφαρμογή των αλγορίθμων χρησιμοποιούνται γραφήματα δενδρικής μορφής, τα λεγόμενα δενδρογράμματα. Το μήκος των οριζόντιων ακμών του είναι ανάλογο με την απόσταση μεταξύ των κέντρων των συστάδων. Ένα παράδειγμα δενδρογράμματος παρουσιάζεται στο **Σχήμα 3.4**. Το δενδρογράμμα αυτό προκύπτει μετά από ιεραρχική ομαδοποίηση στα στοιχεία του χώρου που φαίνονται αριστερά.



Σχήμα 3.4: Δενδρογράμμα ιεραρχικής ομαδοποίησης [51]

Αλγόριθμοι βασισμένοι σε διαχωρισμούς

Από τους πιο πολυεφαρμοσμένους αλγορίθμους συσταδοποίησης είναι ο αλγόριθμος k-Μέσων. Ο αλγόριθμος αυτός στοχεύει στον διαχωρισμό των δεδομένων σε k συστάδες. Τα δεδομένα που εμφανίζουν μεγάλη ομοιότητα μεταξύ τους ως προς κάποια χαρακτηριστικά ομαδοποιούνται στην ίδια συστάδα. Η απόσταση μεταξύ των δεδομένων αποτελεί το μέτρο ομοιότητας μεταξύ των σημείων. Όσο πιο κοντά είναι τα σημεία δεδομένων, τόσο παρόμοια είναι και άρα είναι πιθανότερο να ταξινομηθούν στην ίδια συστάδα. Η πιο συνηθισμένη μετρική απόστασης που χρησιμοποιείται είναι η τετραγωνική Ευκλείδεια απόσταση:

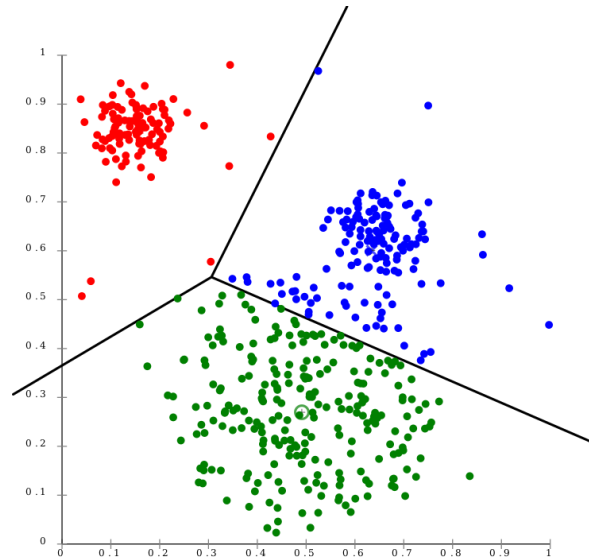
$$d(x, y) = \sum_{i=1}^m (x_i - y_i)^2 = \|\mathbf{x} - \mathbf{y}\|_2^2 \quad (3.39)$$

όπου m είναι η διάσταση των χαρακτηριστικών για κάθε σημείο x, y.

Ο αλγόριθμος ξεκινά διαλέγοντας k τυχαία σημεία-διανύσματα που αποτελούν τα κέντρα των ομάδων. Έπειτα, αναθέτει κάθε σημείο από τα δεδομένα στην ομάδα με το πλησιέστερο κέντρο, έτσι ώστε οι τετραγωνικές αποστάσεις από την ομάδα με αυτό το κέντρο να ελαχιστοποιούνται. Ο παράγοντας αδράνειας που πρέπει να ελαχιστοποιείται σε κάθε βήμα ορίζεται ως εξής:

$$SSE = \sum_{i=1}^n \sum_{j=1}^k w_{i,j} \|x_i - \mu_j\|_2^2 \quad (3.40)$$

όπου μ_j είναι το κέντρο της ομάδας j και $w_{i,j}$ είναι 1 αν το σημείο x_i ανήκει στην ομάδα j και 0 αλλιώς. Στη συνέχεια, υπολογίζει για κάθε ομάδα το μέσο όρο των σημείων της και ορίζει αυτό το μέσο διάνυσμα ως το νέο κέντρο της ομάδας. Τα δύο τελευταία βήματα, η ανάθεση και η ενημέρωση των κέντρων επαναλαμβάνονται έως ότου έρθει σε σύγκλιση ο αλγόριθμος [43]. Στο **Σχήμα 3.5** φαίνεται ένα παράδειγμα διαχωρισμού δεδομένων με τον αλγόριθμο k-Μέσων με $k = 3$.



Σχήμα 3.5: Αποτέλεσμα αλγορίθμου k-Μέσων και διαχωρισμός συνόλου δεδομένων σε 3 ομάδες. Οι γραμμές που χωρίζουν τις ομάδες αποτελούν τα όρια αποφάσεων [55]

Σε αυτό το σημείο να τονιστεί ότι η υπερπαράμετρος k του αλγορίθμου καθορίζεται πριν από την εκτέλεση του αλγορίθμου και η επιλογή της είναι ιδιαίτερα σημαντική για τα αποτελέσματα της συσταδοποίησης. Για την λήψη της απόφασης χρησιμοποιείται συνήθως ένα εμπειρικό κριτήριο που αποκαλείται μέθοδος του "αγκώνα". Η κύρια ιδέα αυτής της μεθόδου είναι η εκτέλεση του k-means αλγορίθμου για το σύνολο δεδομένων για ένα εύρος τιμών του k . Στη συνέχεια, σχεδιάζεται η γραφική παράσταση της αδράνειας ή της διασποράς ως προς το k και ελέγχεται σε ποιο σημείο σχηματίζει "αγκώνα" η γραφική. Η συνολική αδράνεια υπολογίζεται αθροίζοντας τις διασπορές για όλες τις συστάδες C .

$$I = \sum_{C=1}^k I_C$$

όπου I_C η διασπορά των παραδειγμάτων της συστάδας C . Όσο πιο μικρή η μετρική της αδράνειας, τόσο μικρότερη η διασπορά των συστάδων, άρα πιο συμπαγείς ομάδες.

Ωστόσο, η elbow μέθοδος είναι μια ευρετική μέθοδος και, ως εκ τούτου, υπάρχει η πιθανότητα να μην λειτουργεί ικανοποιητικά. Ακόμη, μερικές φορές, σχηματίζονται

περισσότερες γωνίες ή δεν εμφανίζονται καθόλου. Σε αυτές τις περιπτώσεις συνήθως επιλέγεται ο αριθμός k που δίνει καλύτερα αποτελέσματα στο πλαίσιο του προβλήματος, ύστερα από δοκιμές [56].

Ένα από βασικά μειονεκτήματα του αλγορίθμου k -μέσων είναι ότι δεν διαθέτει κάποιον εγγενή τρόπο για να μετράει την αβεβαιότητα της ανάθεσης των δειγμάτων στην εκάστοτε συστάδα. Με άλλα λόγια, ορισμένες φορές μπορεί να υπάρχει επικάλυψη μεταξύ των συστάδων, με αποτέλεσμα η χωροθέτηση των σημείων στο όριο τους να μην γίνεται με απόλυτη εμπιστοσύνη. Εξάλλου, ο αλγόριθμος k -μέσων αποτελεί σκληρό αλγόριθμο συσταδοποίησης. Αντίθετα, οι μαλακοί αλγόριθμοι, δίνουν μια πιθανότητα να ανήκει το κάθε σημείο σε μια κατηγορία. Τα Γκαουσιανά Μοντέλα Μείξης (Gaussian Mixture Models, GMM) είναι μια τέτοια κατηγορία αλγορίθμων που χρησιμοποιούν μείξεις πολυδιάστατων γκαουσιανών κατανομών που μπορούν ως κατώφλια για την κατηγοριοποίηση.

3.2.4 Ταξινόμηση

Το πρόβλημα της ταξινόμησης, όπως ήδη αναφέρθηκε, εντάσσεται στην διαδικασία μάθησης με επίβλεψη. Στόχος της μάθησης είναι να προβλέψει σε ποια κλάση ανήκει ένα άγνωστο αντικείμενο, βάσει των τιμών των χαρακτηριστικών που το προσδιορίζουν, δηλαδή το διάνυσμα εισόδου του. Για την εκμάθηση μιας συνάρτησης στόχου ικανής να κάνει αυτή την αντιστοίχιση, προαπαιτείται ένα σύνολο δεδομένων αντιπροσωπευτικών για κάθε κατηγορία, τα οποία έχουν ταξινομηθεί από τον "επιβλέποντα". Αυτά τα δεδομένα αποτελούν τα δεδομένα εκπαίδευσης τα οποία δίνονται σε έναν αλγόριθμο μηχανικής μάθησης, ο οποίος παρατηρεί και μαθαίνει από τα χαρακτηριστικά τους και παράγει ένα εκτιμώμενο μοντέλο ταξινόμησης, που ονομάζεται ταξινομητής.

Η αξιολόγηση των ταξινομητών γίνεται σε δεδομένα που δεν έχουν συμμετέχει στην εκπαίδευση, έτσι ώστε να υφίσταται η αξιολόγηση της δυνατότητας γενίκευσης τους. Επομένως, απαιτείται ο διαχωρισμός σε δεδομένα εκπαίδευσης, τα οποία χρησιμοποιούνται μόνο κατά την διαδικασία της εκπαίδευσης, και δεδομένα ελέγχου, τα οποία χρησιμοποιούνται μόνο για την τελική εκτίμηση της απόδοσης του ταξινομητή.

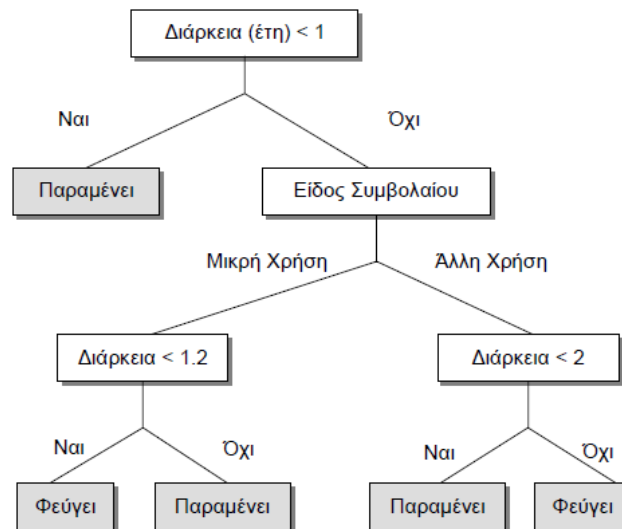
Η ρύθμιση των υπερπαραμέτρων του μοντέλου που θα χρησιμοποιηθεί είναι πολύ σημαντική διαδικασία για την συνολική απόδοση του μοντέλου στην πρόβλεψη. Συνήθως, η επιλογή των παραμέτρων γίνεται με την μέθοδο της Διασταυρούμενης Επικύρωσης (Cross Validation, CV) η οποία χρησιμοποιεί αποτελεσματικά όλα τα δεδομένα εκπαίδευσης. Με αυτήν την τεχνική χωρίζεται το σύνολο εκπαίδευσης σε έναν αριθμό πτυχών, όπου κάθε φορά θα αφήνεται ένα σετ για έλεγχο και τα υπόλοιπα για εκπαίδευση του μοντέλου. Η διαδικασία επαναλαμβάνεται μέχρι ότου να ελεγχθεί η απόδοση για κάθε σετ ελέγχου και ύστερα υπολογίζεται η μέση τιμή της μετρικής σφάλματος. Αυτή η διαδικασία πραγματοποιείται για κάθε τιμή των υπερπαραμέτρων και εντέλει επιλέγονται οι τιμές που πετυχαίνουν την καλύτερη απόδοση. Για αυτές τις τιμές το μοντέλο εκπαιδεύεται ξανά με βάση το σύνολο εκπαίδευσης και στη συνέχεια υπολογίζεται η απόδοση του στο σύνολο ελέγχου.

Η επιλογή του καταλληλότερου μοντέλου εκμάθησης εξαρτάται άμεσα από το εκάστοτε πρόβλημα. Παρακάτω παρουσιάζονται κάποιους από τους συνηθέστερους αλγορίθμους που χρησιμοποιούνται κατά κόρον σε προβλήματα ταξινόμησης.

Δένδρα Απόφασης

Ένα Δέντρο Απόφασης, λαμβάνοντας υπόψη τις αξίες των χαρακτηριστικών εισόδου, παράγει μια ακολουθία κανόνων if-then που χρησιμοποιούνται για την ταξινόμηση των δεδομένων σε κλάσεις. Κάθε κόμβος του δέντρου καθορίζει μια συνθήκη ελέγχου if της τιμής ενός χαρακτηριστικού. Τα κλαδιά του κόμβου αντιστοιχούν σε μια διαφορετική διακριτή τιμή της συνθήκης που υπηρετεί. Αν πρόκειται για κατηγορικά χαρακτηριστικά ο διαχωρισμός γίνεται με βάση τις τιμές του χαρακτηριστικού, ενώ αν πρόκειται για αριθμητικά χαρακτηριστικά επιλέγεται ένα όριο που θα ορίσει την συνθήκη που θα γίνει ο διαχωρισμός.

Ο αλγόριθμος ταξινομεί μια περίπτωση εισόδου ξεκινώντας από την ρίζα του δέντρου και διασχίζοντας τα κλαδιά προς τον επόμενο κόμβο. Σε κάθε κόμβο ελέγχεται η τιμή της περίπτωσης για το χαρακτηριστικό του κόμβου και προχωρά στο αντίστοιχο κλαδί. Η διαδικασία επαναλαμβάνεται, ώστε να καταλήξει σε κάποιο φύλλο (τερματικό κόμβο) που θα περιέχει την τιμή της κατηγορίας όπου και τερματίζεται. Στο **Σχήμα 3.6** παρουσιάζεται ένα παράδειγμα δέντρου ταξινόμησης.



Σχήμα 3.6: Παράδειγμα δέντρου απόφασης για το αν παραμένει ή φεύγει ένας συνδρομητής δεδομένου την διάρκεια και το είδος του συμβολαίου του [51]

Μια μέθοδος ταξινόμησης που βασίζεται στην ιδέα των δέντρων αποφάσεων είναι η τεχνική του Τυχαίου Δάσους (Random Forest). Με την τεχνική αυτή δημιουργούνται πολλά δέντρα αποφάσεων, όπου σε κάθε βήμα της δημιουργίας τους λαμβάνεται υπόψη ένα υποσύνολο των χαρακτηριστικών του συνόλου δεδομένων. Η ποικιλία είναι αυτό που κάνει τον αλγόριθμο αυτόν πιο αποτελεσματικό από ένα μοναδικό δέντρο απόφασης. Στη συνέχεια, κάθε δέντρο απόφασης στο "δάσος" ψηφίζει για την ταξινόμηση μιας νέας περίπτωσης σε μία από τις προκαθορισμένες

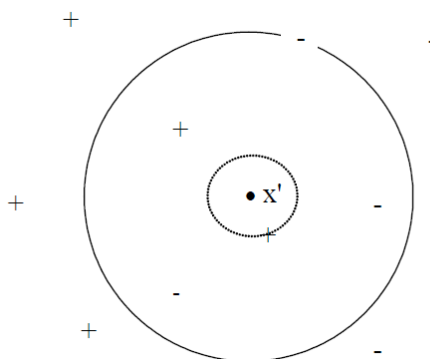
κλάσεις. Το ποσοστό των ψήφων σε κάθε τάξη με όλο το σύνολο είναι το προβλεπτικό διάνυσμα πιθανοτήτων. Η ανάθεση γίνεται σύμφωνα με ένα πλειοψηφικό σύστημα ψηφοφορίας, όπου το νέο δείγμα ανατίθεται στην κλάση με τη μεγαλύτερη συχνότητα [51][57].

Αλγόριθμοι με μάθηση κατά περίπτωση

Στη μάθηση κατά περίπτωση (instance-based learning) δεν υφίσταται η φάση της εκπαίδευσης. Πρακτικά, όταν ένα τέτοιο μοντέλο κληθεί να ταξινομήσει ένα νέο δείγμα στην φάση ελέγχου, εξετάζει την εκείνη την στιγμή την σχέση του με όλα τα δεδομένα εκπαίδευσης.

Χαρακτηριστικός ταξινομητής βασισμένος σε παραδείγματα είναι ο μη παραμετρικός ταξινομητής των k -Κοντινότερων Γειτόνων (k -Nearest Neighbors, k NN). Σύμφωνα με αυτόν τον αλγόριθμο για ένα νέο δείγμα προς ταξινόμηση, πρώτα υπολογίζονται οι k πλησιέστεροι γείτονες του, στον n -διάστατο χώρο των χαρακτηριστικών εισόδου, με βάση κάποια συνάρτηση απόστασης. Συνηθέστερη μετρική απόσταση είναι η Ευκλείδεια απόσταση (σχέση 3.37). Η κλάση του δείγματος θα είναι η κλάση της πλειοψηφίας των k κοντινότερων γειτόνων του, είτε απλά αθροίζοντας τις περιπτώσεις της κάθε κλάσης είτε ζυγίζοντας την απόσταση του κάθε γείτονα.

Η υπερπαραμέτρος k του αλγορίθμου που ορίζει και την γειτονιά του κάθε νέου δείγματος πρέπει να επιλεγεί με προσοχή, ώστε να υπάρχει μια ισορροπία μεταξύ της διακύμανσης (variance) και της απόκλισης (bias) από τα πραγματικά δεδομένα. Για μικρότερες τιμές του k , ο ταξινομητής έχει υψηλότερη διακύμανση αφού τείνει να αγνοεί τη γενικότερη κατανομή και να αποφασίζει αυστηρά μόνο από τα πλησιέστερα δείγματα. Αντίθετα, οι μεγαλύτερες τιμές του k πετυχαίνουν υψηλότερη απόκλιση, αφού ταξινομούν λανθασμένα κάποια αποκλίνοντα δείγματα, με σεβασμό στη συνολική κατανομή [51][57]. Στο **Σχήμα 3.7** απεικονίζεται ένα παράδειγμα που γίνεται αντιληπτή η σημαντικότητα επιλογής της τιμής του k .



Σχήμα 3.7: Παράδειγμα ταξινομητή k -NN. Το δείγμα x' ταξινομείται ως θετικό αν $k = 1$ και ως αρνητικό αν $k = 5$ [51]

Μηχανές διανυσμάτων υποστήριξης

Οι Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines, SVM) αποτελούν μη παραμετρικούς διαχωριστικούς ταξινομητές. Συνήθως, οι SVM ταξινομητές είναι συνήθως δυαδικοί, ωστόσο μπορούν να χρησιμοποιηθούν και σε προβλήματα περισσότερων κλάσεων εφαρμόζοντας τις κατάλληλες τεχνικές. Στην πράξη ο αλγόριθμος που βασίζονται προσπαθεί να βρει μια ευθεία ή μια καμπύλη (σε δύο διαστάσεις) ή μια υπερεπιφάνεια (σε περισσότερες διαστάσεις) που να διαχωρίζει τις κλάσεις μεταξύ τους.

Για γραμμικά διαχωρίσιμα σύνολα δεδομένων, η διαχωριστική γραμμή επιλέγεται έτσι, ώστε να μεγιστοποιείται το περιθώριο μεταξύ των κοντινότερων παραδειγμάτων από τις δύο κλάσεις. Όσα σημεία αγγίζουν το περιθώριο αποκαλούνται διανύσματα υποστήριξης και παίζουν καθοριστικό ρόλο για την εύρεση της βέλτιστης ευθείας. Αντίθετα, τα σημεία που είναι μακριά από το περιθώριο δεν επηρεάζουν την εκπαίδευση. Επομένως, το ίδιο μοντέλο μπορεί να δημιουργηθεί και χωρίς την συμμετοχή τους στην εκπαίδευση.

Ένα από τα σημαντικά πλεονεκτήματα της μεθόδου είναι ότι μπορούν να χρησιμοποιηθούν σε προβλήματα που τα δεδομένα των κλάσεων δεν είναι γραμμικά διαχωρίσιμα. Σε αυτές τις περιπτώσεις χρησιμοποιούν συναρτήσεις πυρήνων με σκοπό να μετασχηματίσουν τον αρχικό χώρο σε έναν άλλο διανυσματικό χώρο, πολύ μεγαλύτερων διαστάσεων. Προβάλλοντας τα δεδομένα σε υψηλότερες διαστάσεις καθίστανται γραμμικά διαχωρίσιμα και άρα μπορεί να βρεθεί ένας γραμμικός διαχωριστής που να ελαχιστοποιεί το σφάλμα ταξινόμησης [51].

4ο Κεφάλαιο

Μεθοδολογία

Εισαγωγή

Ο γλυκαιμικός έλεγχος, όπως αναφέρθηκε στο Κεφάλαιο 2, συμβάλει καθοριστικά στην αντιμετώπιση του διαβήτη και η διαχείριση του είναι ζωτικής σημασίας, ειδικά για τον ΣΔΤ1. Ωστόσο, αν και οι τεχνολογίες για την διαχείριση του έχουν εξελιχθεί συνδυάζοντας συνεχή παρακολούθηση γλυκόζης και συνεχή έγχυση υποδόριας ινσουλίνης, ο βέλτιστος γλυκαιμικός έλεγχος συχνά δεν είναι επιτεύξιμος. Σημαντικά εμπόδια αποτελούν η μεγάλη μεταβλητότητα στην συμπεριφορά του ίδιου του ασθενή, η οποία είναι απόρροια του εξαιρετικά πολύπλοκου μεταβολικού μηχανισμού της γλυκόζης και της ινσουλίνης, καθώς και η αδυναμία πρόβλεψης των καθημερινών συνηθειών του ασθενή από τους ιατρούς που ρυθμίζουν τις δόσεις ινσουλίνης.

Στην παρούσα διπλωματική εργασία, προτείνεται μια μεθοδολογία ανάπτυξης μοντέλου ταξινόμησης οκτάωρων γλυκαιμικών προφίλ που λαμβάνονται καθημερινά από δεδομένα ΔΣΜΓ ατόμου με ΣΔΤ1. Τα δεδομένα γλυκόζης αναλύονται ως δεδομένα σύνθεσης στα οποία εφαρμόζεται αλγόριθμος συσταδοποίησης για να διαχωριστούν σε ομάδες με παρόμοια χαρακτηριστικά. Αυτός ο διαχωρισμός των οκτάωρων περιόδων σε συστάδες χρησιμοποιείται, στη συνέχεια, για την ανάπτυξη μοντέλου ταξινόμησης, ικανό να προβλέπει την κατηγορία που ανήκει ένα άγνωστο οκτάωρο διάστημα μιας ημέρας. Τέλος, η μεθοδολογία περιλαμβάνει ένα πιθανοτικό μοντέλο για τον υπολογισμό των πιθανοτήτων μεταβάσεων μεταξύ διαδοχικών οκτάωρων περιόδων. Συγκεκριμένα, γνωρίζοντας το γλυκαιμικό προφίλ μιας περιόδου, το σύστημα θα μπορεί να ενημερώσει τον ασθενή ή τον θεράποντα ιατρό για το τι προφίλ είναι πιθανότερο να έχει η οκτάωρη περίοδος που ακολουθεί. Κύριος στόχος της μελέτης είναι η ανάπτυξη ενός συστήματος που θα λειτουργεί ως εργαλείο υποστήριξης για την διαχείριση ενός διαστήματος σχετικά με την ινσουλίνη, την διατροφή και τον ύπνο. Η προτεινόμενη μεθοδολογία, που συνοψίζεται στο **Σχήμα 4.1**, εφαρμόστηκε ξεχωριστά για κάθε ασθενή του συνόλου δεδομένων. Τα βήματα και οι μέθοδοι που χρησιμοποιήθηκαν αναλύονται σε αυτό το κεφάλαιο.



Σχήμα 4.1: Σχεδιάγραμμα της προτεινόμενης μεθοδολογίας

4.1 Συλλογή και Προεπεξεργασία Δεδομένων

4.1.1 Συλλογή Δεδομένων

Η μελέτη στηρίχθηκε στην αξιοποίηση των δεδομένων γλυκόζης αισθητήρων από συστήματα συνεχούς μέτρησης της γλυκόζης. Ο αισθητήρας γλυκόζης έχει ρυθμιστεί έτσι, ώστε να καταγράφει δειγματοληπτικά (ανά 5 λεπτά κατά την διάρκεια της μέρας) μετρήσεις γλυκόζης (σε mg/dL) από το υποδόριο υγρό, με αποτέλεσμα να δημιουργούνται καμπύλες γλυκόζης ή αλλιώς χρονοσειρές γλυκόζης. Η κάθε ημέρα πρέπει να περιλαμβάνει 288 μετρήσεις, ώστε να μην υπάρχουν κενά σε αυτές τις καμπύλες. Ωστόσο, τα πραγματικά δεδομένα, τόσο από τις αντλίες ινσουλίνης όσο και τα συστήματα μέτρησης της γλυκόζης, παρουσιάζουν σημαντικές ελλείψεις σε δεδομένα.

Η ανάλυση των δεδομένων γλυκόζης ως δεδομένα σύνθεσης, στο πλαίσιο αυτής εργασίας, απαιτούσε την δημιουργία ενός νέου συνόλου δεδομένων ημερών δίχως κενά σε μετρήσεις ή με κενά που να μπορούν να καλυφθούν, χωρίς όμως να αλλοιώνεται το γλυκαιμικό προφίλ.

4.1.2 Προεπεξεργασία Δεδομένων

Από το αρχικό σύνολο δεδομένων κάθε ασθενούς απορρίφθηκαν οι ημέρες που είχαν κενά μεγαλύτερα των 30 λεπτών (δηλαδή κενά 6 συνεχόμενων δειγμάτων). Στη συνέχεια, ακολουθήθηκε μια μέθοδος αντικατάστασης των τιμών που λείπουν για κάθε ημέρα στο φιλτραρισμένο σύνολο δεδομένων, με σκοπό να μην χαθεί καθόλου πληροφορία.

Αρχικά, τα κενά των 5 λεπτών συμπληρώθηκαν με γραμμική παρεμβολή, με βάση την προηγούμενη και την επόμενη τιμή της γλυκόζης του κενού. Έπειτα, εφαρμόστηκε η μεθοδολογία που προτείνεται από την Sara Zulj και την ομάδα της [58] για την διαχείριση τιμών που λείπουν από δεδομένα γλυκόζης. Σύμφωνα με αυτήν, για την εκτίμηση ενός κενού 10 έως 30 λεπτών (δηλαδή 2 έως 6 συνεχόμενων δειγμάτων) χρησιμοποιήθηκε τμηματική πολυωνυμική-κυβική παρεμβολή (piecewise cubic spline interpolation). Για την προσαρμογή των δεδομένων που λείπουν χρησιμοποιήθηκαν δεδομένα μίας ώρας, δηλαδή 12 δείγματα πριν από το διάστημα που λείπει και δεδομένα 12 δείγματα μετά.

4.2 Ανάλυση των Χρονοσειρών Γλυκόζης ως Δεδομένα Σύνθεσης

Για την ανάπτυξη μοντέλου ταξινόμησης ημερήσιου γλυκαιμικού προφίλ ατόμων με ΣΔΤ1 χρησιμοποιήθηκε μια μεθοδολογία κατηγοριοποίησης των δεδομένων που στηρίζεται στην ανάλυση των δεδομένων γλυκόζης ως Δεδομένα Σύνθεσης, η οποία έχει προταθεί από την Biagi και την ομάδα της [59][60][61]. Σύμφωνα με την θεωρία που παρουσιάστηκε στο Κεφάλαιο 3, ένα διάνυσμα $x = [x_1, x_2, \dots, x_D]$ με D τμήματα ορίζεται ως σύνθεση, όταν όλες οι συνιστώσες x_i είναι αυστηρά θετικοί αριθμοί και το άθροισμα τους δίνει έναν σταθερό αριθμό C . Οι συνιστώσες x_i μεταφέρουν σχετική πληροφορία για την συμβολή του τμήματος i στο γενικό σύνολο C . Στην περίπτωση των χρονοσειρών γλυκόζης, η ανάλυση του χρόνου που δαπανάται σε καθορισμένα εύρη γλυκόζης, κατά την διάρκεια της ημέρας, μπορεί να αναπαρασταθεί ως διάνυσμα σύνθεσης.

4.2.1 Καθορισμός διανυσμάτων σύνθεσης

Το σύνολο δεδομένων για κάθε ασθενή αποτελείται από 288 δείγματα που αναπαριστούν εικοσιτετράωρες χρονοσειρές γλυκόζης. Η κάθε ημέρα διαχωρίζεται σε τρεις οκτάωρες περιόδους με 96 τιμές που ξεκινούν στις 00:00, 08:00 και στις 16:00. Για κάθε οκτάωρη περίοδο της ημέρας υπολογίστηκε ένα διάνυσμα (μια σύνθεση) που φανερώνει πως κατανεμήθηκε ο χρόνος σε έξι διαστήματα γλυκόζης σχετικά με υπογλυκαιμία, ευγλυκαιμία και υπεργλυκαιμία.

- **Υπογλυκαιμία**

Level 2 (σοβαρή): $BG < 60 \text{ mg/dL}$

Level 1 (ήπια): $60 \text{ mg/dL} \leq BG < 70 \text{ mg/dL}$

- **Ευγλυκαιμία**
 $70 \text{ mg/dL} \leq \text{BG} < 140 \text{ mg/dL}$
 $140 \text{ mg/dL} \leq \text{BG} < 180 \text{ mg/dL}$
- **Υπεργλυκαιμία**
Level 1 (ήπια): $180 \text{ mg/dL} \leq \text{BG} < 250 \text{ mg/dL}$
Level 2 (σοβαρή) : $\text{BG} \geq 250 \text{ mg/dL}$

Με βάση τα προαναφερθέντα για καθεμία από τις περιόδους ανάλυσης υπολογίστηκε το διάνυσμα \mathbf{x} με συνιστώσες που υποδηλώνουν τον χρόνο σε κάθε εύρος γλυκόζης. Ο χρόνος που δαπανάται σε κάθε εύρος γλυκόζης μπορεί να μετρηθεί ως ποσοστό, ως λεπτά ή ως ώρες.

$$\mathbf{x}_{8h} = [X_{60}, X_{60.70}, X_{70.140}, X_{140.180}, X_{180.250}, X_{250}]$$

$$\sum_{i=1}^6 X_i = 8 \text{ ώρες ή } 480 \text{ λεπτά ή } 1$$

4.2.2 Διαχείριση μηδενικών τιμών

Επειδή η στατιστική ανάλυση των Δεδομένων Σύνθεσης βασίζεται σε αναλογίες λογαρίθμων μεταξύ των τμημάτων (συνιστωσών) του διανύσματος σύνθεσης, απαιτείται να μην υπάρχουν μηδενικές τιμές. Ωστόσο, σε δεδομένα γλυκόζης είναι πιθανό ο ασθενής να μην έχει βρεθεί σε κάποια διαστήματα κατά την διάρκεια μιας καθορισμένης περιόδου και επομένως η αντίστοιχη συνιστώσα να είναι μηδενική.

Στην παρούσα ανάλυση παρατηρήθηκαν τέτοιες περιπτώσεις μηδενικών τιμών σε κάποια διαστήματα γλυκόζης (συνήθως σε υπογλυκαιμία και σοβαρή υπεργλυκαιμία). Παρ' όλ' αυτά, δεν υπήρχε κάποιο εύρος γλυκόζης που να ήταν πάντα μηδενικό για οποιοδήποτε ασθενή και άρα οι μηδενικές τιμές έπρεπε να αντικατασταθούν. Έτσι, εφαρμόστηκε η τεχνική πολλαπλασιαστικής αντικατάστασης χρησιμοποιώντας την συνάρτηση *multiplicative_replacement* του πακέτου *Composition Statistics* της *skbio* στην *Python*.

Έστω $\mathbf{x} \in S^D$, $\sum x_i = C$ με Z μηδενικά. Η τεχνική θα αντικαταστήσει το διάνυσμα \mathbf{x} με ένα νέο διάνυσμα $\mathbf{r} \in S^D$ χωρίς μηδενικές τιμές σύμφωνα με την σχέση:

$$r_j = \begin{cases} \delta_j, & \text{αν } x_j = 0 \\ \left(1 - \frac{\sum_{k|x_k=0} \delta_k}{C}\right) x_j, & \text{αν } x_j > 0 \end{cases}$$

όπου δ_j είναι μία πολύ μικρή τιμή που αποδόθηκε στη μηδενική συνιστώσα x_j ενώ οι υπόλοιπες συνιστώσες μετασχηματίζονται ώστε να διατηρείται το άθροισμα τους σταθερό ίσο με C [62].

Στην μεταβλητή δ δόθηκε μια πολύ μικρή τιμή ($\delta = 0.0014 = 0,6/480$) η οποία συμβολίζει ότι ο ασθενής από τα 480 λεπτά μιας οκτάωρης περιόδου βρέθηκε για λίγα δευτερόλεπτα σε ένα συγκεκριμένο εύρος γλυκόζης.

4.2.3 Μετασχηματισμοί για δημιουργία διανυσμάτων εισόδου

Για να είναι εφικτή η στατιστική ανάλυση των δεδομένων, κρίθηκε απαραίτητη η μεταφορά τους από τον χώρο Simplex στον Ευκλείδειο χώρο, όπως αυτή περιγράφηκε στο Κεφάλαιο 3. Έτσι, τα διανύσματα σύνθεσης μετά την αντικατάσταση των μηδενικών τιμών μετασχηματίστηκαν σε διανύσματα αναλογιών λογαρίθμων (log-ratio coordinates). Οι γραμμικοί μετασχηματισμοί που χρησιμοποιήθηκαν ήταν οι ισομετρικοί clr και ilr . Η Ευκλείδεια απόσταση των μετασχηματισμένων διανυσμάτων ισούται με την Aitchison απόσταση των αντίστοιχων διανυσμάτων σύνθεσης χωρίς να χάνεται πληροφορία. Επομένως, μπορούν να χρησιμοποιηθούν ως είσοδος σε μοντέλα συσταδοποίησης και σε μοντέλα ταξινόμησης.

Για την δημιουργία των clr διανυσμάτων ακολουθήθηκε η εξής διαδικασία για κάθε ασθενή [47]:

Έστω, X ένας ($n \times 6$) πίνακας που περιέχει n διανύσματα οκτάωρων περιόδων \mathbf{x}_{8h} .

Βήμα 1:

Υπολογισμός του γεωμετρικού μέσου του συνόλου δεδομένων ως ένα μέτρο κεντρικής τάσης.

Το κέντρο του συνόλου δεδομένων ορίζεται ως εξής:

$$g = C[g_1, g_2, g_3, g_4, g_5, g_6]$$

όπου C η πράξη της κλειστότητας (σχέση 3.2) και $g_i = \left(\prod_{j=1}^n x_{ij}\right)^{\frac{1}{n}}$, $i = 1, 2, \dots, 6$
ο γεωμετρικός μέσος κάθε στήλης του πίνακα X .

Βήμα 2:

Μετατόπιση-κεντράρισμα κάθε σύνθεσης γύρω από τον γεωμετρικό μέσο του συνόλου. Για κάθε διάνυσμα-γραμμή του πίνακα X εκτελείται η πράξη της πρόσθεσης με το κέντρο g^{-1} σύμφωνα με την σχέση 3.6:

$$y_i = \mathbf{x}_i \oplus g^{-1} = C \left[\frac{x_{i1}}{g_1}, \frac{x_{i2}}{g_2}, \dots, \frac{x_{i6}}{g_6} \right]$$

$$i = 1, 2, \dots, n$$

Βήμα 3:

Μετασχηματισμός κάθε διανύσματος σε clr συνιστώσες. Σε κάθε γραμμή-διάνυσμα του πίνακα Y εφαρμόζεται clr μετασχηματισμός (σχέση 3.24)

$$clr(y_i) = \left[\log\left(\frac{y_{i1}}{g(y_i)}\right), \log\left(\frac{y_{i2}}{g(y_i)}\right), \dots, \log\left(\frac{y_{i6}}{g(y_i)}\right) \right]$$

$$i = 1, 2, \dots, n$$

Με την παραπάνω διαδικασία δημιουργήθηκε ένας πίνακας Z όπου η κάθε του γραμμή αποτελούσε το μετασχηματισμένο clr -διάνυσμα κάθε οκταώρου του συνόλου δεδομένων. Δηλαδή είχε την μορφή:

$$Z = [clr(y_i)]$$

$$i = 1, 2, \dots, n$$

Τα δύο πρώτα βήματα της διαδικασίας υλοποιήθηκαν χρησιμοποιώντας την συνάρτηση *centralize* και το τρίτο βήμα χρησιμοποιώντας την συνάρτηση *clr* του πακέτου *Composition Statistics* της *skbio* στην *Python*.

Για την δημιουργία των *ilr* διανυσμάτων ακολουθήθηκε η τεχνική της διαδοχικής διχοτόμησης *SBP* που περιγράφηκε στο [60]. Στο πλαίσιο της παρούσας εργασίας η διχοτόμηση καθορίστηκε σύμφωνα με την κλινική ερμηνεία του χρόνου σε διαφορετικά εύρη γλυκόζης, με σκοπό να αναπαραστήσει διαφορετικά περιστατικά υπογλυκαιμικών ή υπεργλυκαιμικών γεγονότων.

Πίνακας 4.1: Μέθοδος διαδοχικής διχοτόμησης για δημιουργία *ilr* διανυσμάτων

j	X.60	X.60.70	X.70.140	X.140.180	X.180.250	X.250	r	s
1	+1	+1	-1	-1	-1	-1	2	4
2	+1	-1	0	0	0	0	1	1
3	0	0	-1	-1	+1	+1	2	2
4	0	0	0	0	-1	+1	1	1
5	0	0	+1	-1	0	0	1	1

Υπολογίζονται οι συνιστώσες $ilr = [ilr_1, ilr_2, ilr_3, ilr_4, ilr_5]$ για κάθε δείγμα του πίνακα X , σύμφωνα με τον τύπο της σχέσης 3.34 του Κεφαλαίου 3. Οι γεωμετρικοί μέσοι σε αριθμητή και παρονομαστή είναι οι κεντρικές τιμές των συνιστωσών για κάθε γκρουπ αυτών των τμημάτων. Ο λόγος μεταξύ τους δείχνει το βάρος της κάθε υποομάδας, ο λογάριθμος παρέχει την κατάλληλη κλίμακα και ο τετραγωνικός συντελεστής είναι μια σταθερά κανονικοποίησης που επιτρέπει την σύγκριση μεταξύ τους [47].

$$ilr_1 = \sqrt{\frac{4}{3}} \ln \frac{(X_{60} \cdot X_{60.70})^{1/2}}{(X_{70.140} \cdot X_{140.180} \cdot X_{180.250} \cdot X_{250})^{1/4}}$$

Η πρώτη συνιστώσα μπορεί να ερμηνευτεί ως η σχέση μεταξύ του χρόνου που περνιέται σε υπογλυκαιμικά εύρη και του χρόνου που περνιέται σε κανονικά και υπεργλυκαιμικά διαστήματα.

$$ilr_2 = \sqrt{\frac{1}{2}} \ln \frac{X_{60}}{X_{60.70}}$$

Η δεύτερη συνιστώσα υποδηλώνει την ισορροπία μεταξύ του χρόνου σε ήπια και σοβαρή υπογλυκαιμία.

$$ilr_3 = \ln \frac{(X_{180.250} \cdot X_{250})^{1/2}}{(X_{70.140} \cdot X_{140.180})^{1/2}}$$

Η τρίτη συνιστώσα δείχνει την ισορροπία μεταξύ του χρόνου σε υπεργλυκαιμία και σε ευγλυκαιμία.

$$ilr_4 = \sqrt{\frac{1}{2}} \ln \frac{X_{250}}{X_{180.250}}$$

Η τέταρτη συνιστώσα υποδηλώνει την ισορροπία μεταξύ του χρόνου σε ήπια και σοβαρή υπεργλυκαιμία.

$$ilr_5 = \sqrt{\frac{1}{2}} \ln \frac{X_{70.140}}{X_{140.180}}$$

Η πέμπτη και τελευταία συνιστώσα είναι ίση με την αναλογία του χρόνου σε εύρος 70-140 mg/dL και σε 140-180mg/dL.

4.3 Μοντέλα Συσταδοποίησης

Μετά τον μετασχηματισμό των διανυσμάτων σύνθεσης σε log-ratio συνιστώσες, ακολούθησε η διαδικασία της συσταδοποίησης. Ο αλγόριθμος που χρησιμοποιήθηκε για την κατηγοριοποίηση των διανυσμάτων ήταν ο k-Μέσων, που περιγράφηκε στην ενότητα 3.2.3. Για την εφαρμογή του αλγορίθμου χρησιμοποιήθηκε η υλοποίηση του *KMeans* από το *scikit-learn* της Python.

Ο αλγόριθμος k-Μέσων στοχεύει στον διαχωρισμό των δεδομένων εισόδου σε k συστάδες με βάση την τετραγωνική Ευκλείδεια απόσταση τους, ώστε κοντινά σημεία να κατηγοριοποιηθούν μαζί. Ο τύπος της Ευκλείδειας απόστασης μεταξύ δύο διανυσμάτων x, y είναι ο εξής:

$$d(x, y) = \sum_{i=1}^m (x_i - y_i)^2 = \|x - y\|_2^2$$

Για κάθε ασθενή, δημιουργήθηκαν τρία διαφορετικά μοντέλα συσταδοποίησης μη επιβλεπόμενης μάθησης ένα για κάθε οκτάωρο διάστημα της ημέρας 00:00-07:59, 08:00-16:00 και 16:00-23:59. Η επιλογή τριών διαφορετικών μοντέλων, ανάλογα με την περίοδο που ελέγχεται, έγινε με σκοπό να γίνεται κάθε φορά ομαδοποίηση σε

όμοια διαστήματα, με ίδιους κερκαδικούς ρυθμούς. Εξάλλου, οι τρεις περίοδοι διαφέρουν ως προς την διαχείριση της ινσουλίνης, των γευμάτων και του ύπνου.

Αλγόριθμος 1: k-Μέσων

Είσοδος: Σύνολο δεδομένων $N = \{x_1, x_2, \dots, x_n\}$, με n διανύσματα εισόδου

Αριθμός Συστάδων k

Έξοδος: Συστάδες $C = \{C_1, C_2, \dots, C_k\}$

1. ανάθεση τυχαίων κέντρων

για $i = 1, 2, \dots, k$ κάνε:

θεώρησε μ_i ένα τυχαίο στοιχείο από το N ως κέντρο μιας ομάδας

2. ομαδοποίηση

για ένα προκαθορισμένο αριθμό επαναλήψεων ή όσο υπάρχουν αλλαγές στο διαχωρισμό των σημείων στις ομάδες C_i κάνε:

2α. ανάθεση σημείων στις ομάδες

για $i = 1, 2, \dots, k$ κάνε:

$$C_i = \{x \in N \mid d(\mu_i, x) \leq d(\mu_j, x) \text{ για όλα τα } j = 1, \dots, k \text{ με } j \neq i\}$$

2β. αναθεώρηση κέντρων των ομάδων

για $i = 1, 2, \dots, k$ κάνε:

$$\mu_i = \text{μέσο διάνυσμα των σημείων που ανήκουν στην ομάδα } C_i$$

Επιπροσθέτως, για κάθε παραγόμενο γκρουπ υπολογίστηκαν τα ακόλουθα κλινικά χαρακτηριστικά τα οποία χρησιμοποιήθηκαν ως κριτήρια ομοιογένειας των συστάδων που προκύπτουν. Αυτά τα χαρακτηριστικά υποδηλώνουν το προφίλ γλυκόζης των δειγμάτων που εμπεριέχει η κάθε ομάδα και δίνουν πληροφορίες σχετικές με τα υπογλυκαιμικά ή τα υπεργλυκαιμικά επεισόδια.

Τιμές γλυκόζης

- μέση τιμή γλυκόζης Avg BG (mg/dL)
- τυπική απόκλιση γλυκόζης SD (mg/dL)

Δείκτες γλυκόζης

- μέσος χαμηλός δείκτης γλυκόζης LBG (Low Blood Glucose Index)
- μέσος υψηλός δείκτης γλυκόζης HBG (High Blood Glucose Index)

Οι δείκτες αυτοί χρησιμοποιούνται για την εκτίμηση κινδύνων εμφάνισης υπογλυκαιμίας και υπεργλυκαιμίας. Για κάθε μέτρηση γλυκόζης υπολογίστηκε η μετασχηματισμένη γλυκόζη και οι δείκτες σύμφωνα με τους τύπους:

$$f(\text{BG}) = 1,509 \cdot [\ln(\text{BG})^{1,084} - 5,381]$$

$$r(\text{BG}) = 10 \cdot f(\text{BG})^2$$

$$rl(\text{BG}) = \begin{cases} r(\text{BG}), & \text{αν } f(\text{BG}) < 0 \\ 0, & \text{αλλού} \end{cases}$$

$$rh(\text{BG}) = \begin{cases} r(\text{BG}), & \text{αν } f(\text{BG}) > 0 \\ 0, & \text{αλλού} \end{cases}$$

Γλυκαιμικά επεισόδια

- μέση τιμή υπογλυκαιμικών επεισοδίων Hypo Events (event/8h period)
- μέση διάρκεια υπογλυκαιμικών επεισοδίων Hypo Duration (min/event)
- μέση τιμή υπεργλυκαιμικών επεισόδια Hyper Events (event/8-h period)
- μέση διάρκεια υπεργλυκαιμικών επεισοδίων Hyper Duration (min/event)

Χρόνος σε φυσιολογικά και μη εύρη γλυκόζης

- ποσοστό χρόνου σε ευγλυκαιμία %Time in Range (70-180 mg/dL)
- ποσοστό χρόνου σε υπογλυκαιμία %Time below 70 mg/dL
- ποσοστό χρόνου σε υπεργλυκαιμία %Time above 180 mg/dL

Η επιλογή του πλήθους των συστάδων k για κάθε μοντέλο συσταδοποίησης βασίστηκε σε συνδυασμό των μεθόδων εύρεσης βέλτιστης τιμής της παραμέτρου αλλά και της κλινικής ποιότητας των αποτελεσμάτων της. Η εφαρμογή των αλγορίθμων για την αξιολόγηση των συστάδων βασίζεται σε απλά μαθηματικά, επομένως δεν συνυπολογίζεται το κλινικό κομμάτι. Το πρόβλημα που μελετήθηκε απαιτούσε την κλινική αξία των αποτελεσμάτων. Γι' αυτόν τον λόγο, εκτός από την ανάλυση silhouette, ακολούθησε και ανάλυση των προαναφερθέντων κλινικών χαρακτηριστικών τα οποία χρησιμοποιήθηκαν και ως κριτήρια για το πόσο ομοιογενείς είναι οι κλάσεις που παράγονται. Κατά την εφαρμογή της μεθόδου "αγκώνα" παρατηρήθηκε ότι για $k=3$ και 4 σχηματιζόταν γωνίες στο γράφημα. Συνεπώς, λήφθηκε η απόφαση να δοκιμαστεί ο αλγόριθμος για αυτές τις δύο τιμές της παραμέτρου k και να ελεγχθεί η κλινική ποιότητα των αποτελεσμάτων. Δεδομένου ότι οι μαθηματικοί αλγόριθμοι δεν είχαν σημαντική διαφορά για αυτές τις δύο τιμές και ότι με τρεις κλάσεις τα αποτελέσματα ήταν καλύτερης κλινικής ποιότητας, επιλέχθηκαν τρεις κλάσεις για κάθε ασθενή και για κάθεμια περίοδο ανάλυσης.

4.4 Μοντέλα Ταξινόμησης

Με την συσταδοποίηση που προηγήθηκε δημιουργήθηκαν, για κάθε ασθενή του συνόλου δεδομένων, οι ετικέτες εξόδου για κάθε διάνυμα εισόδου. Στη συνέχεια, ελέγχθηκε αν αυτά τα ζεύγη εισόδου - εξόδου μπορούν να χρησιμοποιηθούν για την ανάπτυξη μοντέλου ταξινόμησης ενός οκτάωρου διανύσματος σε μία από τις κλάσεις που προέκυψαν από την συσταδοποίηση. Όπως και προηγουμένως, η διερεύνηση έγινε ξεχωριστά για τις τρεις διαφορετικές οκτάωρες περιόδους ανάλυσης.

Για έλεγχο της απόδοσης του εκάστοτε ταξινομητή χρησιμοποιήθηκε η μέθοδος διασταυρούμενης επικύρωσης με πτυχές όσα και τα δείγματα εκπαίδευσης, γνωστή και ως μέθοδος Leave One Out Cross Validation (LOO-CV). Σύμφωνα με την μέθοδο, επιλέγεται κάθε φορά για δοκιμή ένα από τα δείγματα και τα υπόλοιπα χρησιμοποιούνται για την εκπαίδευση του ταξινομητή. Η διαδικασία επαναλαμβάνεται εξαιρώντας κάθε φορά ένα διαφορετικό δείγμα και εκπαιδεύοντας τον ταξινομητή με τα υπόλοιπα. Στο τέλος, υπολογίζεται ο μέσος όρος της απόδοσης του ταξινομητή σε κάθε βήμα. Οι ετικέτες εξόδου των δειγμάτων εκπαίδευσης σε κάθε βήμα λαμβάνονται εκ νέου από την εφαρμογή αλγορίθμου συσταδοποίησης χωρίς την συμμετοχή του δείγματος επικύρωσης. Αντίθετα, η ετικέτα της κλάσης του δείγματος ελέγχου, η οποία θα χρησιμοποιηθεί για την αξιολόγηση της ταξινόμησης, λαμβάνεται από την αρχική συσταδοποίηση με την συμμετοχή όλων των δειγμάτων του συνόλου δεδομένων. Η μέθοδος αυτή επιλέχθηκε γιατί χρησιμοποιεί όλα τα διαθέσιμα δεδομένα για εκπαίδευση και δοκιμή, ενώ διασφαλίζει την ανεξαρτησία μεταξύ αυτών.

Με σκοπό να βρεθεί ο πιο αποδοτικός ταξινομητής, δοκιμάστηκαν διάφοροι αλγόριθμοι ταξινόμησης, μερικοί από τους οποίους είχαν περιγραφεί στην ενότητα 3.2.4. Αλγόριθμοι όπως ο αλγόριθμος κοντινότερων γειτόνων *k*NN, ο αλγόριθμος τυχαίου δάσους Random Forest, ο XGBoost, οι Μηχανές Διανυσμάτων Υποστήριξης SVM και ο γκαουσιανός ταξινομητής Naive Bayes χρησιμοποιήθηκαν με την μέθοδο LOO-CV για να ελεγχθεί η απόδοση τους στην ταξινόμηση ενός αγνώστου οκτάωρου. Κάποιες από τις βασικές παραμέτρους αυτών παρουσιάζονται στη συνέχεια.

k-NN (k Nearest Neighbors)

Χρησιμοποιήθηκε η συνάρτηση *KNeighborsClassifier()* της βιβλιοθήκης *scikit-learn*. Η παράμετρος *k* έπειτα από δοκιμές επιλέχθηκε ίση 3 (*n_neighbors=3*), δηλαδή ένα νέο δείγμα θα ταξινομείται με βάση την πλειοψηφία των κλάσεων των τριών κοντινότερων γειτόνων του με βάση την ευκλείδεια απόσταση (*metric='Euclidean'*).

Random Forest

Χρησιμοποιήθηκε η συνάρτηση *RandomForestClassifier()* της βιβλιοθήκης *scikit-learn*. Χρησιμοποιήθηκαν 100 δέντρα απόφασης (*n_estimators=100*) και ο αριθμός των χαρακτηριστικών για τον διαχωρισμό στα φύλλα επιλέχθηκε ίσος με 3 (*max_features=3*).

XGBoost

Χρησιμοποιήθηκε η συνάρτηση *XGBClassifier()* με τις default παραμέτρους της από το API *xgboost* της Python.

SVM

Χρησιμοποιήθηκε η συνάρτηση *SVC* της βιβλιοθήκης *scikit-learn*. Η συνάρτηση πυρήνα που χρησιμοποιήθηκε ήταν γραμμική (*kernel='linear'*). Ακόμη χρησιμοποιήθηκε η τακτική 'ένα προς ένα' (*decision_function_shape='ovo'*), διότι ο αλγόριθμος SVM χρησιμοποιήθηκε ως ταξινομητής περισσότερων κλάσεων και όχι ως δυαδικός ταξινομητής.

4.5 Πιθανοτικό Μοντέλο Μεταβάσεων

Για την δημιουργία του πιθανοτικού μοντέλου μεταβάσεων μεταξύ των περιόδων, πραγματοποιήθηκε αναδρομική ανάλυση των δεδομένων που προέκυψαν από την διαδικασία της συσταδοποίησης. Το μοντέλο που κατασκευάστηκε υπολόγιζε τις πιθανότητες μετάβασης από τις κλάσεις μιας συγκεκριμένης περιόδου στις κλάσεις της επακόλουθης της, με σκοπό να βρεθεί το πιθανότερο γλυκαιμικό προφίλ που θα έχει ένα οκτάωρο, δεδομένου το προφίλ που είχε το αμέσως προηγούμενο οκτάωρο.

Για τις μεταβάσεις υπολογίσθηκαν οι δεσμευμένες πιθανότητες για κάθε συνδυασμό των ομάδων των διαδοχικών οκτάωρων περιόδων ως ακολούθως:

$$P(G_{n+1}|G_n) = \frac{P(G_n G_{n+1})}{P(G_n)} = \frac{\text{count}(G_n G_{n+1})}{\text{count}(G_n)}$$

όπου G_n και G_{n+1} οι ομάδες των περιόδων n και $n + 1$.

5ο Κεφάλαιο

Αποτελέσματα

Εισαγωγή

Στο παρόν Κεφάλαιο πραγματοποιείται παρουσίαση των αποτελεσμάτων της μεθοδολογίας που αναπτύχθηκε στο Κεφάλαιο 4. Συγκεκριμένα, παρατίθενται πληθώρα γραφημάτων και πινάκων για την κλινική ερμηνεία και χαρακτηρισμό των συστάδων που προκύπτουν μετά την εφαρμογή του αλγορίθμου συσταδοποίησης για τις τρεις διαφορετικές οκτάωρες περιόδους ανάλυσης 00:00-07:59, 08:00:15:59 και 16:00-23:59. Ακόμη, καταγράφεται η απόδοση των ταξινομητών που χρησιμοποιήθηκαν, αλλά και οι πιθανότητες μετάβασης από το ένα προφίλ στο επακόλουθο του.

5.1 Σύνολα Δεδομένων

Πραγματικά δεδομένα 11 παιδιών, ηλικίας 7 έως 14 χρονών, που πάσχουν από ΣΔΤ1 και χρησιμοποιούν συστήματα συνεχούς έγχυσης υποδόριας ινσουλίνης σε συνδυασμό με συστήματα συνεχούς μέτρησης της γλυκόζης, αναλύθηκαν σε αυτήν την εργασία. Συγκεκριμένα, όλοι οι ασθενείς χρησιμοποιούσαν την αντλία ινσουλίνης MiniMed 640G σε συνδυασμό με τον αισθητήρα γλυκόζης Enlite και τον πομπό Guardian 2 Link της εταιρείας Medtronic. Από τα ημερολόγια αυτών των ασθενών συλλέχθηκαν δεδομένα γλυκόζης από τον αισθητήρα του συστήματος που παρείχε πληροφορία για τα επίπεδα γλυκόζης κάθε πέντε λεπτά, δημιουργώντας χρονοσειρές γλυκόζης.

Στη συνέχεια, σύμφωνα με την μεθοδολογία προεπεξεργασίας που αναπτύχθηκε στην ενότητα 4.1.2 πραγματοποιήθηκε αντικατάσταση των τιμών που έλειπαν, με σκοπό το σύνολο δεδομένων για κάθε ασθενή να περιέχει ημέρες χωρίς κενά διαστήματα σε δεδομένα γλυκόζης. Στον **Πίνακα 5.1** φαίνεται το πλήθος ημερών για κάθε ασθενή, όπως προέκυψαν μετά την προεπεξεργασία. Η κάθε ημέρα διαχωρίστηκε σε τρεις οκτάωρες περιόδους, οι οποίες αποτέλεσαν και τα σύνολα εισόδου στα μοντέλα που χρησιμοποιήθηκαν στην μελέτη.

Πίνακας 5.1: Πλήθος ημερών χωρίς κενά σε μετρήσεις, όπως αυτές προέκυψαν μετά την προεπεξεργασία του αρχικού συνόλου δεδομένων

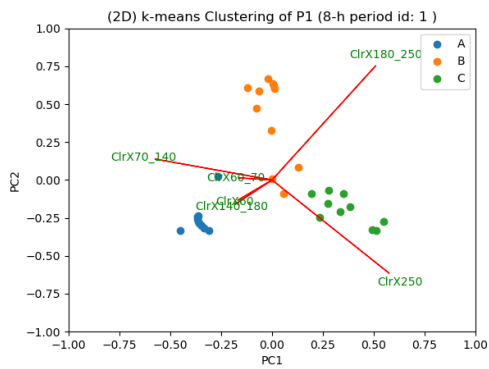
Ασθενής	Πλήθος ημερών
P1	31
P2	29
P3	28
P4	25
P5	21
P6	38
P7	24
P8	24
P9	20
P10	25
P11	62

Η μεθοδολογία που προτάθηκε πραγματοποιήθηκε ξεχωριστά για τον κάθε ασθενή αφού η μεταβλητότητα μεταξύ των ασθενών, όσον αφορά τον μεταβολισμό της γλυκόζης, απαιτεί την ανάπτυξη εξατομικευμένων συστημάτων υποστήριξης. Επιπλέον, όπως αναφέρθηκε παραπάνω, δημιουργήθηκαν τρία διαφορετικά μοντέλα συσταδοποίησης και ταξινόμησης ένα για καθέ οκτάωρο της ημέρας (00:00-07:59, 08:00-15:59, 16:00-23:59). Παρακάτω ακολουθεί παρουσίαση των αποτελεσμάτων όπως αυτά προέκυψαν για κάθε μία από τις τρεις αναλύσεις.

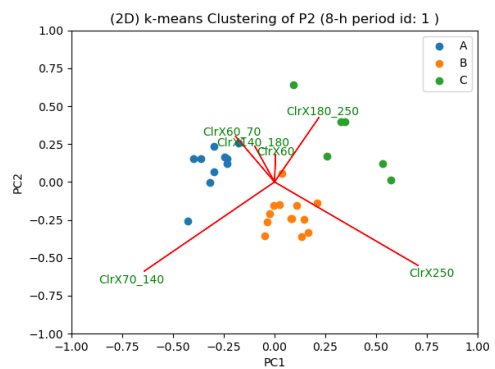
5.2 Αποτελέσματα Συσταδοποίησης

Τα αποτελέσματα που προέκυψαν για κάθε ασθενή ξεχωριστά, έπειτα από συσταδοποίηση με τον αλγόριθμο k-μέσων, παρουσιάζονται ακολούθως. Συγκεκριμένα, τα *Σχήματα 5.1, 5.4 και 5.7* απεικονίζουν *clr biplots* τα οποία αποτελούν τα διαγράμματα διασποράς των διανυσμάτων εισόδου κάθε οκτάωρης περιόδου που μελετάται, καθώς και τον διαχωρισμό τους σε ομάδες. Οι *Πίνακες 5.3, 5.6, 5.9* και τα αντίστοιχα *Σχήματα 5.2, 5.5 και 5.8* υποδηλώνουν ποια *clr* συνιστώσα κυριαρχεί σε κάθε παραγόμενη ομάδα, χαρακτηρίζοντας την ως υπογλυκαιμική, ευγλυκαιμική ή υπεργλυκαιμική κατηγορία. Επιπροσθέτως, οι *Πίνακες 5.4, 5.7 και 5.10* αποτελούν τους συγκεντρωτικούς πίνακες των μέσων τιμών των κλινικών χαρακτηριστικών για κάθε συστάδα, ενώ τα *Σχήματα 5.3, 5.6 και 5.9* απεικονίζουν τις καμπύλες γλυκόζης που ανήκουν σε κάθε κλάση.

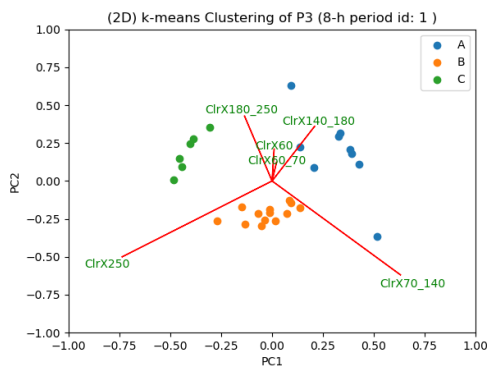
(a) Ασθενής P1



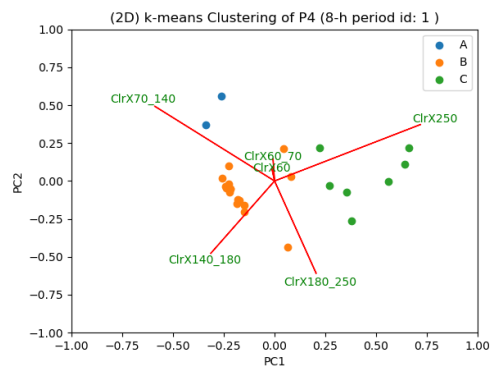
(b) Ασθενής P2



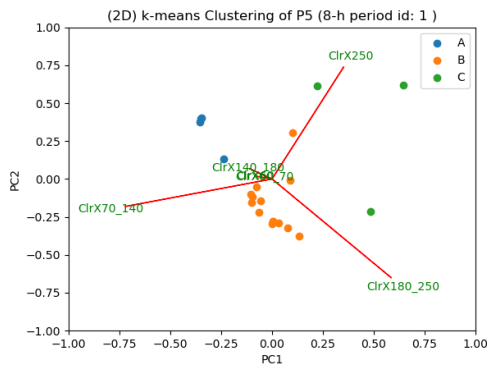
(c) Ασθενής P3



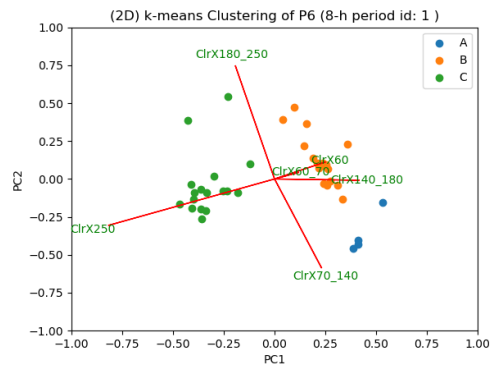
(d) Ασθενής P4



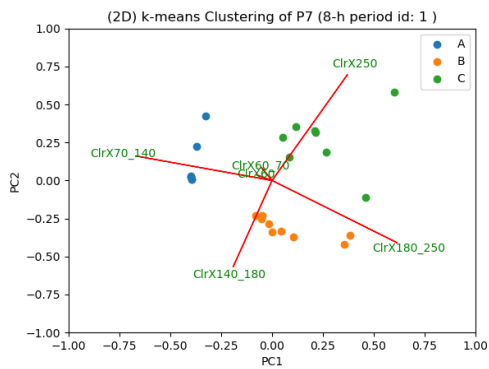
(e) Ασθενής P5



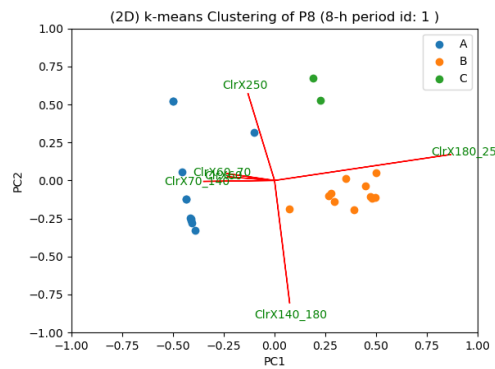
(f) Ασθενής P6

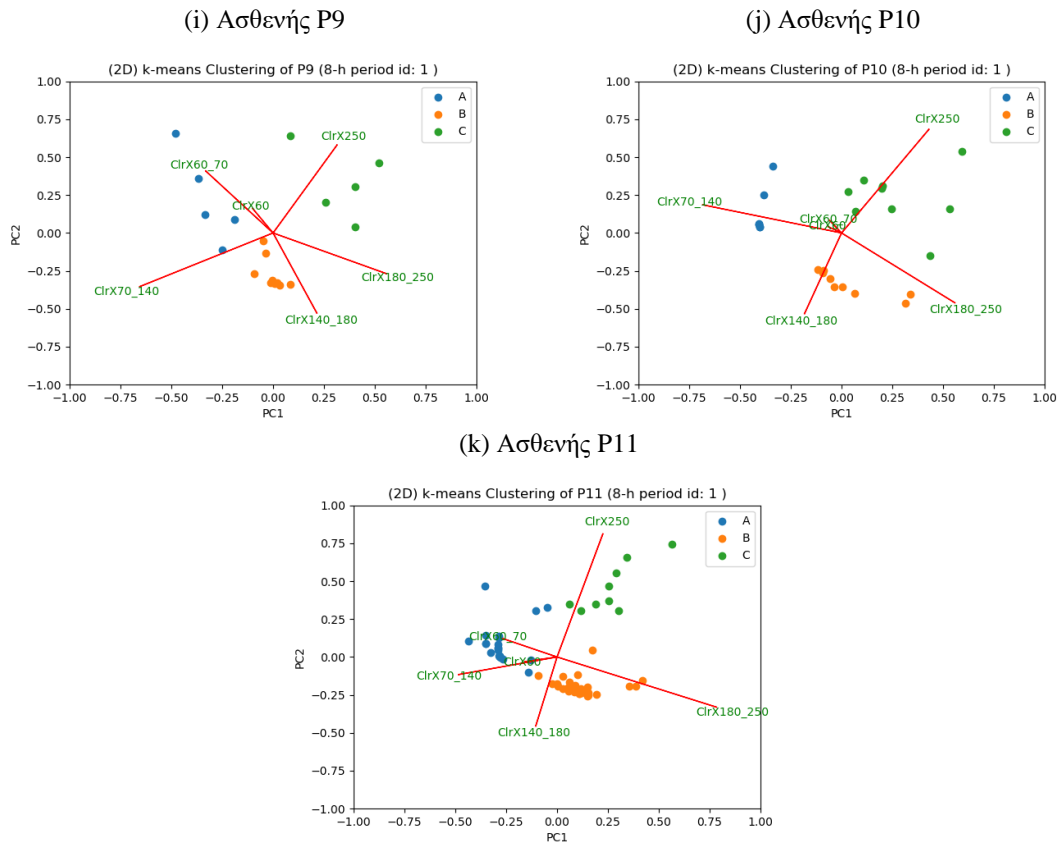


(g) Ασθενής P7



(h) Ασθενής P8





Σχήμα 5.1: Clr biplots των διανυσμάτων που αντιστοιχούν στην 8-h περίοδο 00:00-07:59 για κάθε ασθενή. Οι κουκκίδες αναπαριστούν τα διανύσματα εισόδου στο διδιάστατο χώρο. Ο διαχωρισμός τους σε ομάδες έχει προκύψει ύστερα από την εφαρμογή αλγόριθμου συσταδοποίησης

Στους Πίνακες 5.2, 5.5 και 5.8 παρουσιάζεται ο διαχωρισμός των οκταώρων κάθε περιόδου ανάλυσης στις παραγόμενες συστάδες.

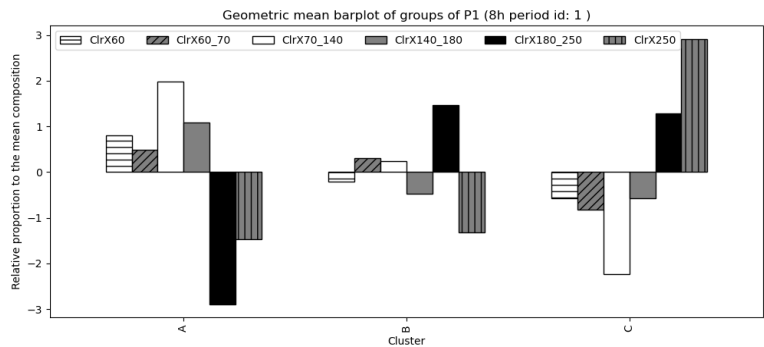
Πίνακας 5.2: Πλήθος δειγμάτων ανά cluster (περίοδος: 00:00-07:59)

Ασθενής	Πλήθος συνολικών ημερών	Πλήθος 8h διανυσμάτων Cluster A	Πλήθος 8h διανυσμάτων Cluster B	Πλήθος 8h διανυσμάτων Cluster C
P1	31	10	11	10
P2	29	10	13	6
P3	28	9	13	6
P4	25	2	16	7
P5	21	14	4	3
P6	38	4	17	17
P7	24	7	9	8
P8	24	11	11	2
P9	20	5	10	5
P10	25	7	9	9
P11	62	22	31	9

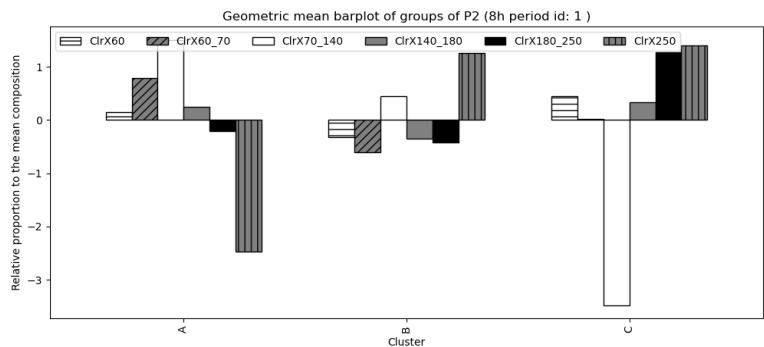
Πίνακας 5.3: Σύγκριση του κέντρου κάθε cluster συνιστώσας με το κέντρο του συνόλου για κάθε cluster (διάστημα: 00:00-07:59)

Ασθενής	Κλάση	X.60	X.60.70	X.70.140	X.140.180	X.180.250	X.250
P1	A	0,805284	0,49133	1,977667	1,08691	-2,89098	-1,47021
	B	-0,20913	0,298677	0,233908	-0,46904	1,46334	-1,31775
	C	-0,57524	-0,81987	-2,23497	-0,57097	1,281304	2,919737
P2	A	0,148529	0,781251	1,50711	0,248565	-0,2081	-2,47735
	B	-0,32355	-0,60785	0,444207	-0,34618	-0,42388	1,257252
	C	0,45348	0,014914	-3,4743	0,335777	1,26525	1,404877
P3	A	0,401721	0,238073	1,094012	1,104197	-0,03851	-2,79949
	B	-0,44888	-0,20372	0,729764	-0,48128	-0,49471	0,898823
	C	0,369997	0,084281	-3,22217	-0,61353	1,129636	2,251785
P4	A	0,85351	0,745562	3,271323	-0,20903	-3,81753	-0,84384
	B	-0,07498	-0,01426	0,771925	0,654885	0,144547	-1,48212
	C	-0,07248	-0,18043	-2,69906	-1,43716	0,760331	3,628805
P5	A	0,630322	0,313716	1,755707	0,485779	-3,07592	-0,1096
	B	-0,14391	-0,12886	0,292461	-0,04892	0,577238	-0,54802
	C	-0,16886	0,183043	-3,70576	-0,41941	1,407453	2,703539
P6	A	1,487655	1,151328	1,65166	1,174411	-3,82276	-1,64229
	B	0,20368	-0,03402	0,058922	0,897042	0,697385	-1,82301
	C	-0,55372	-0,23688	-0,44755	-1,17337	0,202088	2,209428
P7	A	0,646254	0,255078	2,510893	0,577928	-3,15016	-0,83999
	B	-0,1416	-0,13127	-0,689	0,810386	1,779334	-1,62785
	C	-0,40617	-0,07551	-1,42191	-1,41737	0,754642	2,566318
P8	A	0,599153	0,633285	0,740841	-0,14487	-1,98766	0,159255
	B	-0,41206	-0,44094	-0,55919	0,40126	1,799864	-0,78894
	C	-1,02902	-1,0579	-0,9991	-1,41014	1,032897	3,463274
P9	A	0,724877	1,895545	1,702957	-1,07942	-2,69551	-0,54845
	B	-0,33915	-0,68175	0,731588	0,614823	0,546819	-0,87233
	C	-0,04657	-0,53204	-3,16613	-0,15022	1,601872	2,293101
P10	A	0,655598	0,280069	2,6968	0,588721	-3,21762	-1,00357
	B	-0,13226	-0,10628	-0,50309	0,821179	1,711882	-1,79143
	C	-0,37765	-0,11155	-1,59442	-1,27907	0,790707	2,571985
P11	A	0,633047	0,961808	1,162888	0,159363	-2,68707	-0,23003
	B	-0,25043	-0,32875	-0,42397	0,286862	1,551086	-0,8348
	C	-0,68486	-1,21871	-1,38227	-1,37763	1,225775	3,437708

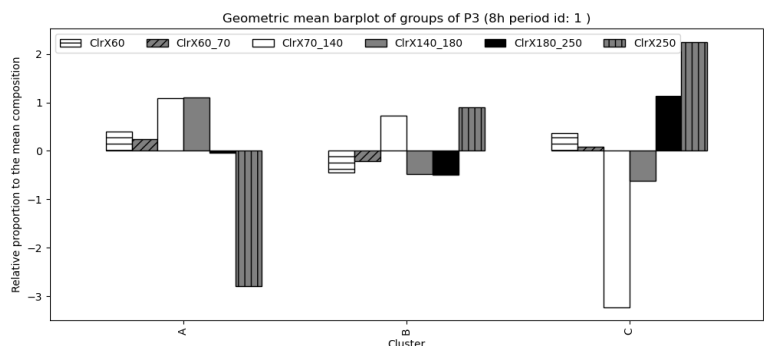
(a) Ασθενής P1



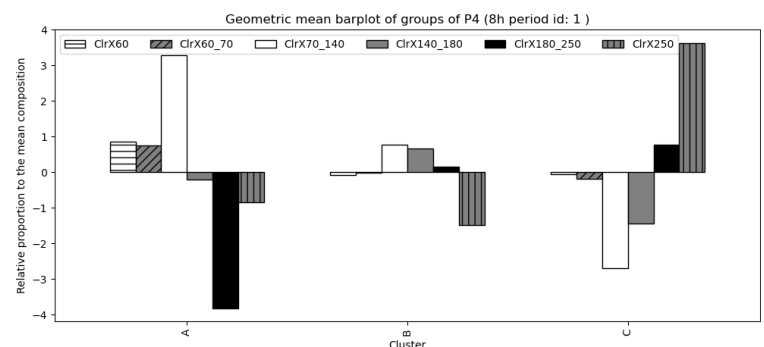
(b) Ασθενής P2



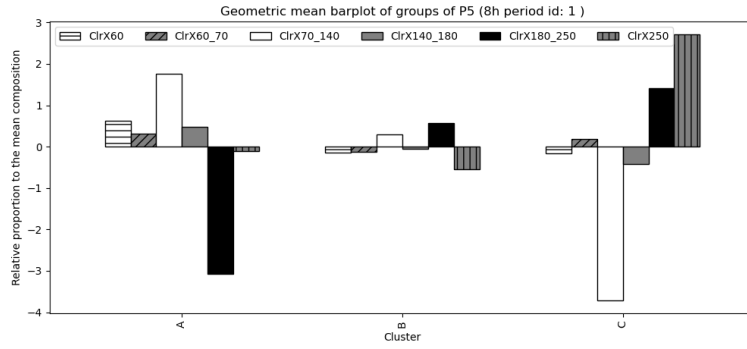
(c) Ασθενής P3



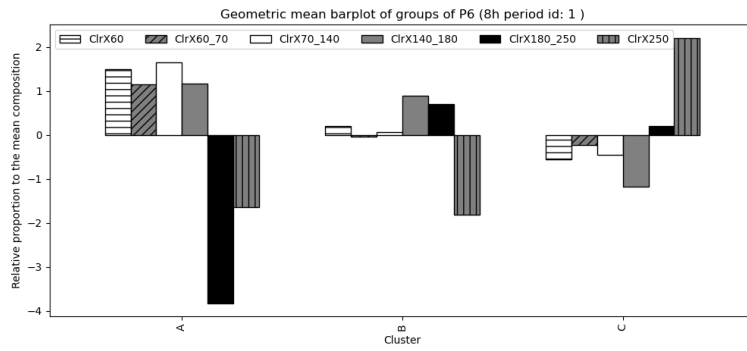
(d) Ασθενής P4



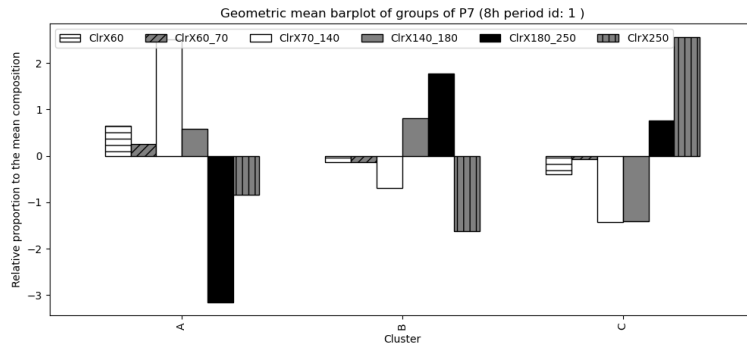
(e) Ασθενής P5



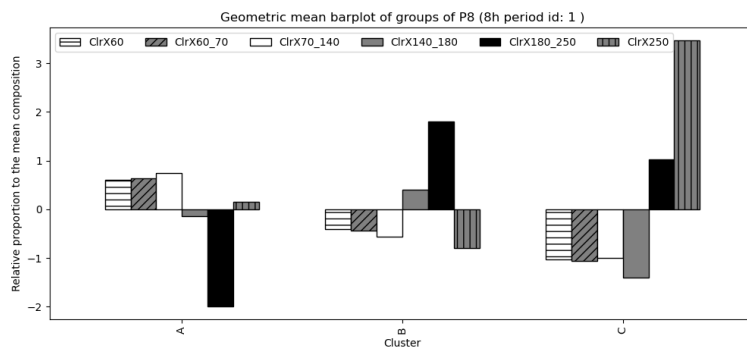
(f) Ασθενής P6



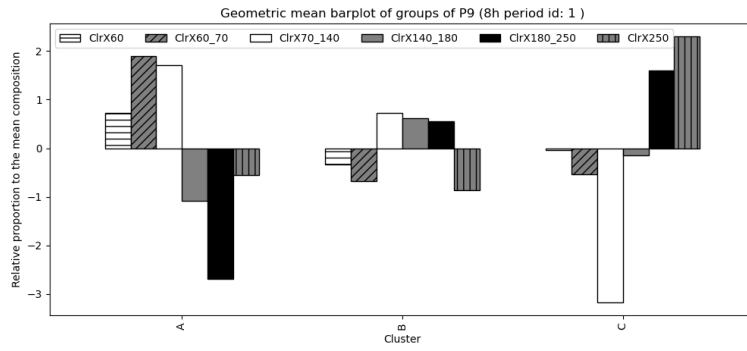
(g) Ασθενής P7



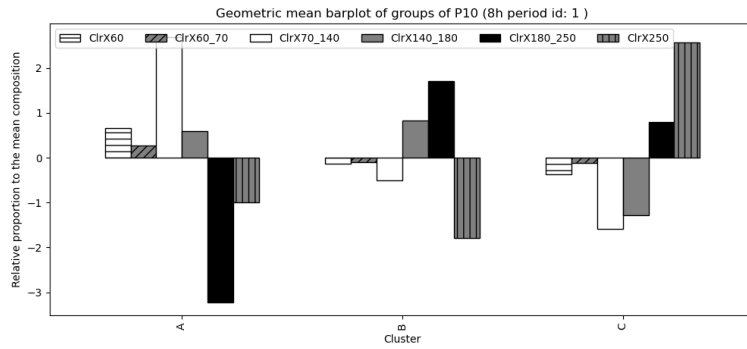
(h) Ασθενής P8



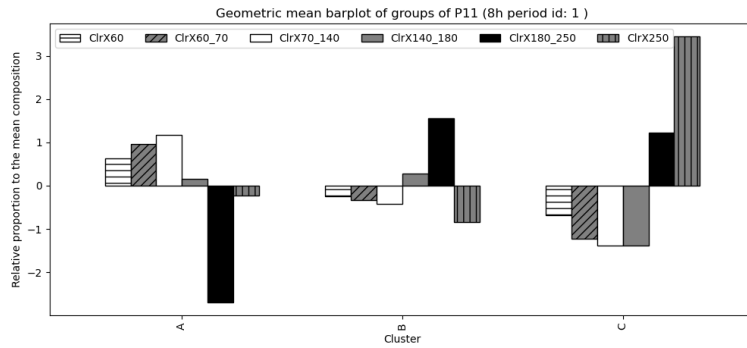
(i) Ασθενής P9



(j) Ασθενής P10



(k) Ασθενής P11



Σχήμα 5.2: Bar plots των μέσων clr συνιστωσών για κάθε cluster και για κάθε ασθενή για την πρώτη 8-h περίοδο της ημέρας (00:00-07:59). Σύγκριση του κέντρου κάθε ομάδας με το συνολικό κέντρο (μηδέν). Θετικές τιμές υποδηλώνουν υψηλές τιμές στο συγκεκριμένο εύρος γλυκόζης, ενώ αρνητικές τιμές υποδηλώνουν χαμηλές τιμές.

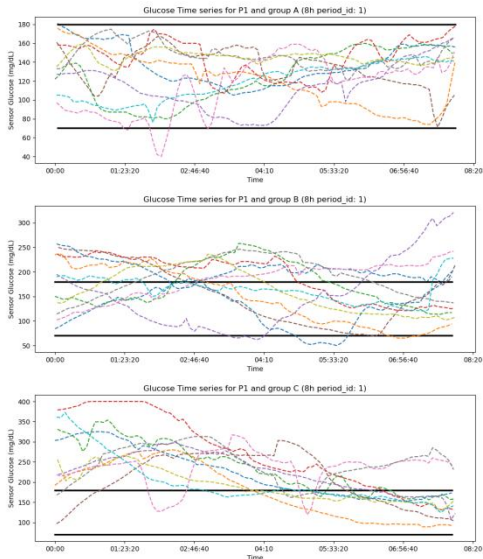
Πίνακας 5.4: Συγκεντρωτικός πίνακας κλινικών δεδομένων για κάθε ομάδα (διάστημα: 00:00-07:59)

Ασθενής	Κλάση	Avg BG (mg/dL)	SD (mg/dL)	LBGI	HBGI	#Hypo (events/period)	T Hypo (min/event)	#Hyper (events/period)	T Hyper (min/event)	% Time in Range	% Time Below 70	% Time Above 180
P1	A (10)	129,95	20,92	0,545	1,632	0,1	2	0	0	98,75	0,88	0,37
	B (11)	167,50	46,24	0,615	7,665	0,091	4,545	0,273	9,091	53,06	2,78	44,16
	C (10)	221,91	57,79	0,027	17,472	0	0	1,4	140,667	32,23	0,28	67,49
P2	A (10)	150,40	37,28	0,464	5,069	0	0	0	0	65,88	2,11	32,01
	B (13)	179,99	59,93	0,305	10,362	0,077	1,538	1,231	62,500	53,55	0,89	45,56
	C (6)	208,72	35,24	0	13,981	0	0	1	60,417	30,21	0,28	69,51
P3	A (9)	160,61	32,45	0,310	6,032	0,111	1,667	0	0	63,64	1,17	35,19
	B (13)	197,11	63,36	0,395	13,771	0,077	1,538	1,231	120,385	41,17	1,76	57,07
	C (6)	243,83	36,58	0	21,419	0	0	1,167	169,583	8,81	0,28	90,91
P4	A (2)	123,17	13,96	0,086	0,661	0	0	0	0	99,44	0,28	0,28
	B (16)	156,75	34,17	0,192	5,224	0,063	0,938	0,188	2,969	71,67	0,77	27,56
	C (7)	245,19	49,88	0,004	22,171	0	0	1,143	187,143	16,89	0,28	82,83
P5	A (4)	118,58	24,63	0,625	1,027	0	0	0	0	99,22	0,28	0,50
	B (14)	152,29	35,43	0,320	4,873	0	0	0	0	72,56	1,31	26,13
	C (3)	211,12	41,80	0,051	14,829	0	0	1	78,333	24,68	0,58	74,74
P6	A (4)	126,68	24,25	2,23	1,833	0,250	32,500	0	0	89,40	10,09	0,51
	B (17)	165,53	31,99	0,22	6,585	0,059	1,471	0	0	64,33	0,63	35,04
	C (17)	201,35	61,03	0,25	14,283	0	0	0,941	118,529	35,11	0,98	63,91

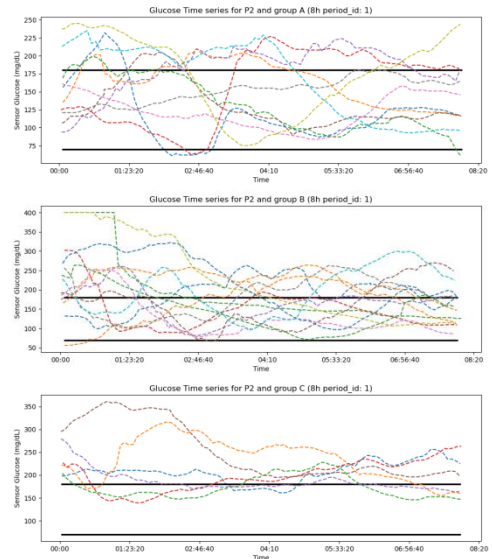
Πίνακας 5.4 (Συνέχεια): Συγκεντρωτικός πίνακας κλινικών δεδομένων για κάθε ομάδα (διάστημα: 00:00-07:59)

Ασθενής	Κλάση	Avg BG (mg/dL)	SD (mg/dL)	LBGI	HBGI	#Hypo (events/period)	T Hypo (min/event)	#Hyper (events/period)	T Hyper (min/event)	% Time in Range	% Time Below 70	% Time Above 180
P7	A (7)	118,84	21,27	0,549	0,897	0	0	0	0	99,44	0,28	0,28
	B (9)	171,22	34,39	0,229	7,713	0	0	0	0	52,36	0,84	46,80
	C (8)	202,12	58,50	0,722	14,988	0,125	3,750	1,125	108,438	33,57	2,68	63,75
P8	A (11)	118,95	21,44	0,498	0,869	0,091	1,364	0	0	98,07	0,82	1,11
	B (11)	144,92	37,87	0,340	4,175	0	0	0	0	75,72	0,28	24,56
	C (2)	158,21	68,70	0,602	8,065	0	0	1	90	64,40	0,28	35,32
P9	A (5)	113,49	24,22	1,314	0,929	0	0	0	0	93,76	4,32	1,92
	B (10)	140,85	37,02	0,495	3,640	0	0	0	0	80,93	0,37	18,70
	C (5)	199,43	42,79	0,290	12,647	0,200	4	1	47,5	29,97	1,68	68,35
P10	A (7)	118,84	21,27	0,550	0,897	0	0	0	0	99,44	0,28	0,28
	B (9)	171,22	34,39	0,228	7,713	0	0	0	0	52,36	0,84	46,80
	C (9)	205,91	56,70	0,641	15,517	0,111	3,330	1,111	121,389	31,01	2,42	66,57
P11	A (22)	119,84	26,67	1,157	1,276	0,182	3,750	0,045	0,682	94,68	3,94	1,38
	B (31)	156,65	37,13	0,360	5,629	0	0	0	0	66,91	0,97	32,12
	C (9)	192,85	49,62	0,156	12,603	0	0	1,333	71,111	40,20	0,28	59,52

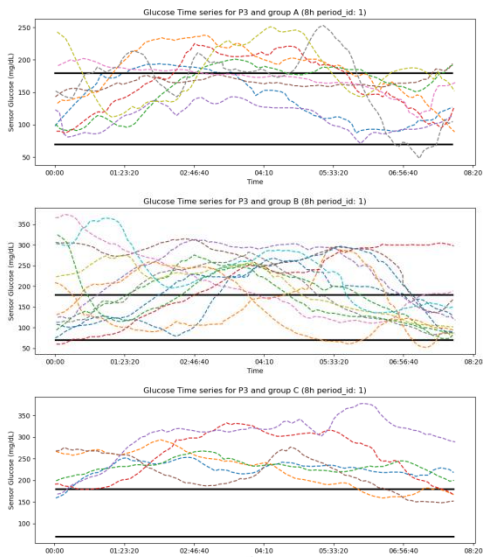
(a) Ασθενής P1



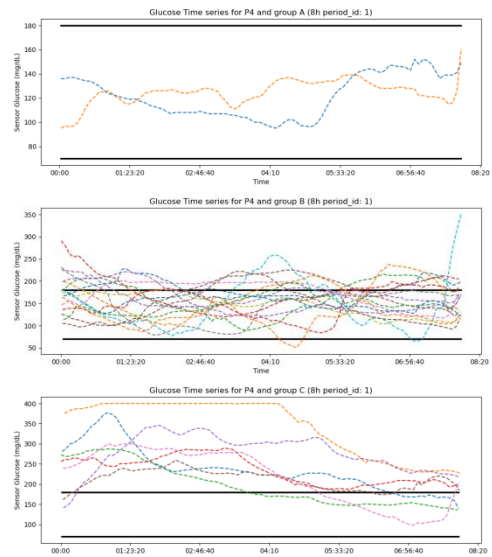
(b) Ασθενής P2



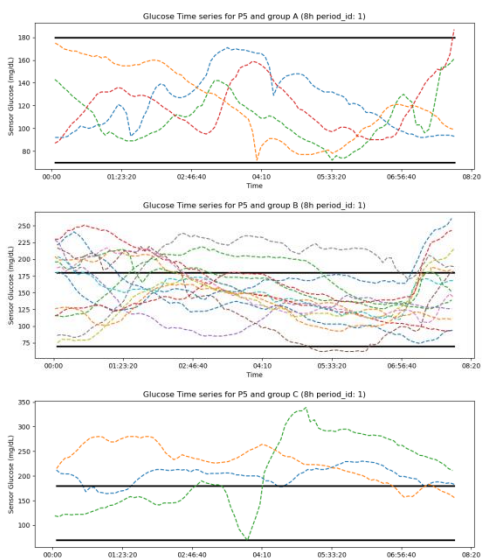
(c) Ασθενής P3



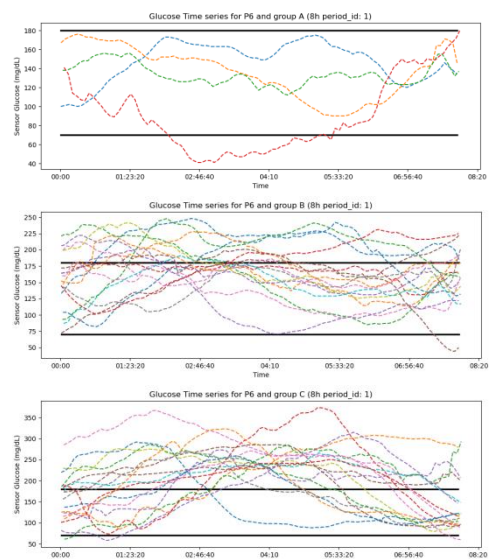
(d) Ασθενής P4

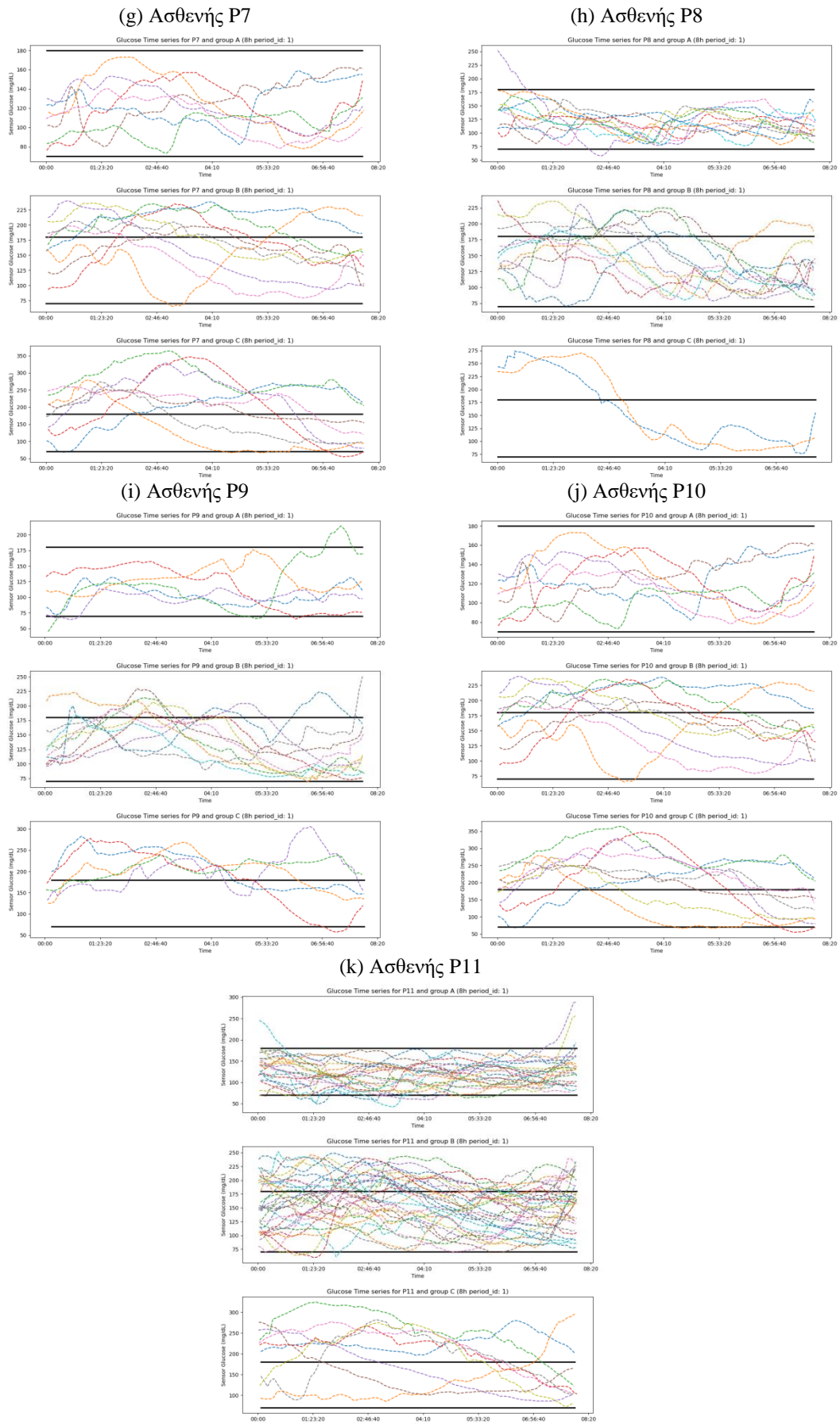


(e) Ασθενής P5



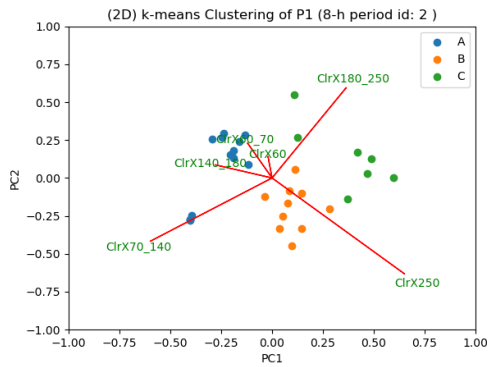
(f) Ασθενής P6



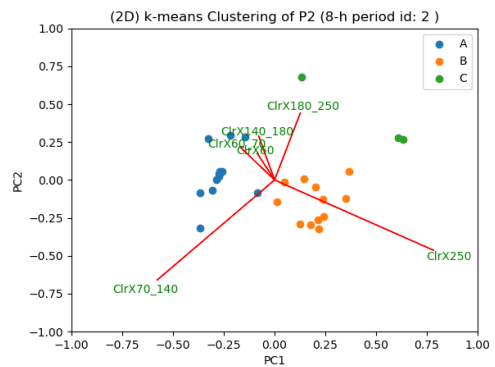


Σχήμα 5.3: 8-h χρονοσειρές γλυκόζης (διάστημα 00:00-07:59) για κάθε ασθενή που δείχνουν τα γλυκαιμικά προφίλ κάθε κλάσης A, B, C

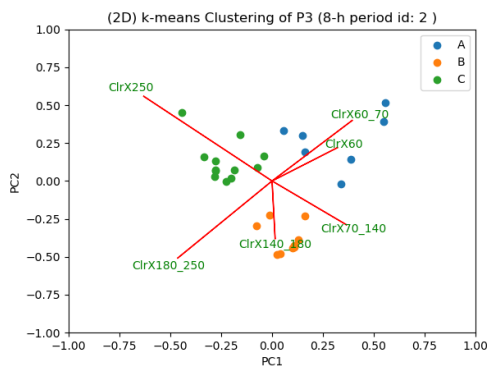
(a) Ασθενής P1



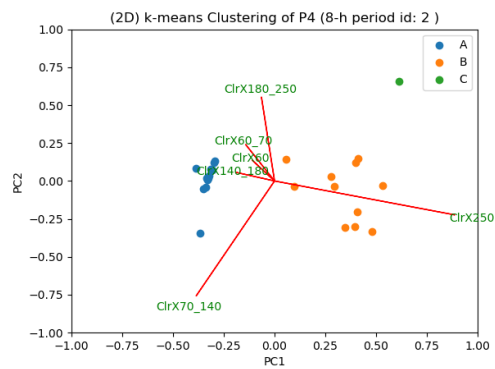
(b) Ασθενής P2



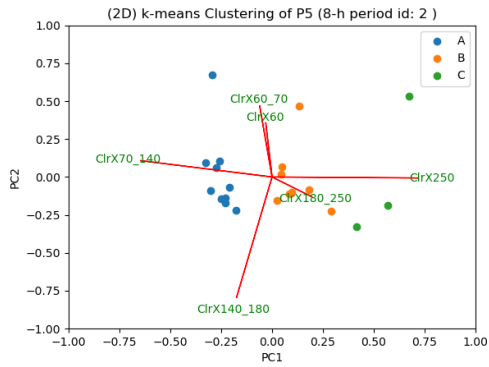
(c) Ασθενής P3



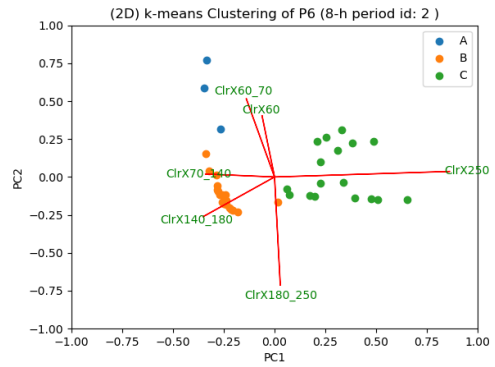
(d) Ασθενής P4



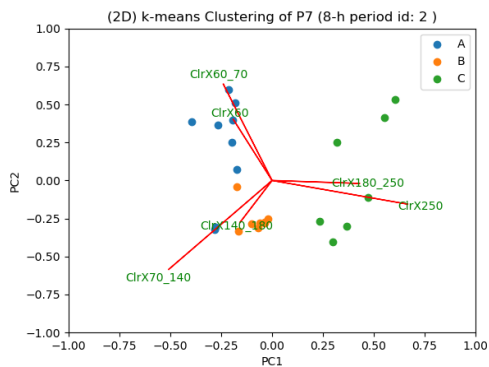
(e) Ασθενής P5



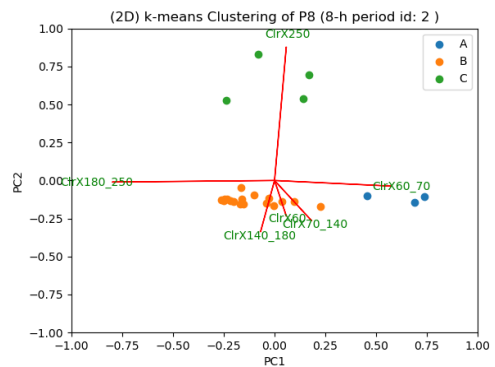
(f) Ασθενής P6

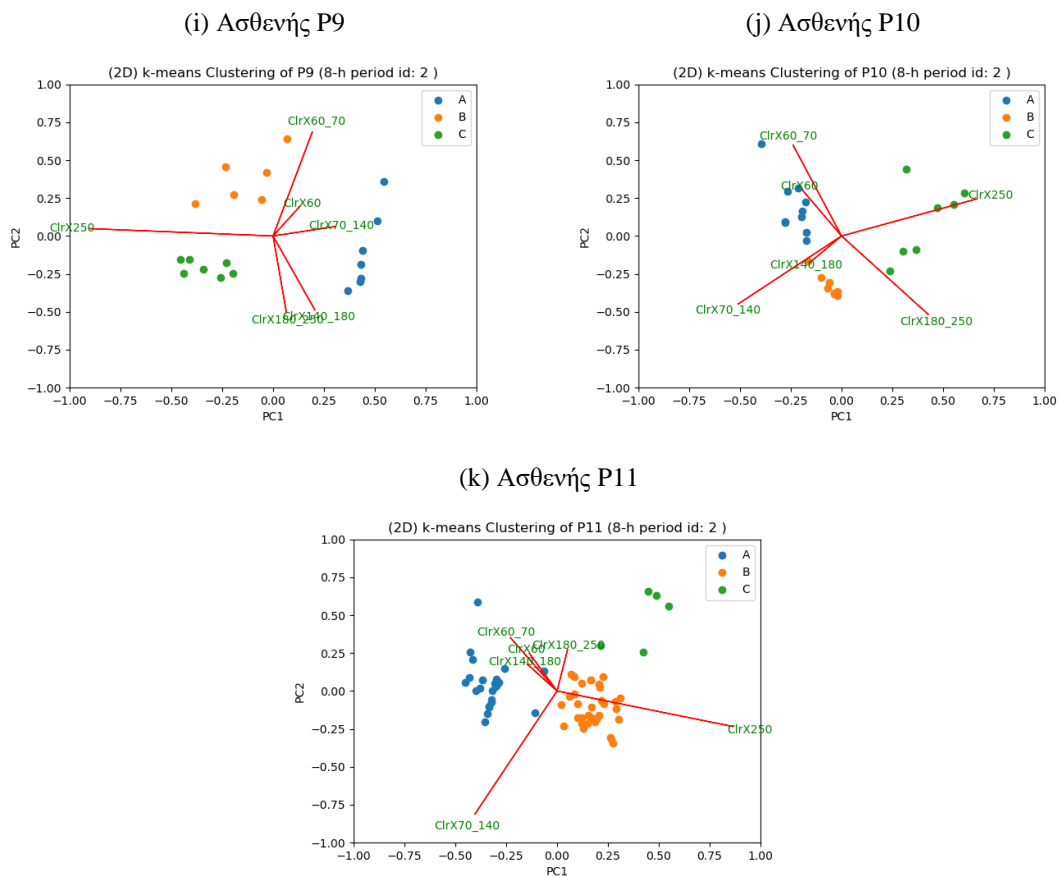


(g) Ασθενής P7



(h) Ασθενής P8





Σχήμα 5.4: Clr biplots των διανυσμάτων που αντιστοιχούν στην 8-h περίοδο 08:00-15:59 για κάθε ασθενή. Οι κουκκίδες αναπαριστούν τα διανύσματα εισόδου στο διδιάστατο χώρο. Ο διαχωρισμός τους σε ομάδες έχει προκύψει ύστερα από την εφαρμογή αλγόριθμου συσταδοποίησης

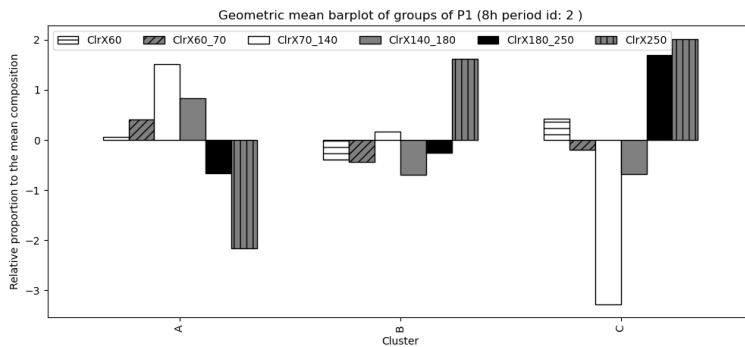
Πίνακας 5.5: Πλήθος δειγμάτων ανά cluster (περίοδος: 08:00-15:59)

Ασθενής	Πλήθος συνολικών ημερών	Πλήθος 8h διανυσμάτων Cluster A	Πλήθος 8h διανυσμάτων Cluster B	Πλήθος 8h διανυσμάτων Cluster C
P1	31	14	10	7
P2	29	14	12	3
P3	28	7	9	12
P4	25	13	11	1
P5	21	10	8	3
P6	38	3	8	17
P7	24	9	8	7
P8	24	3	17	4
P9	20	7	6	7
P10	25	10	8	7
P11	62	24	33	5

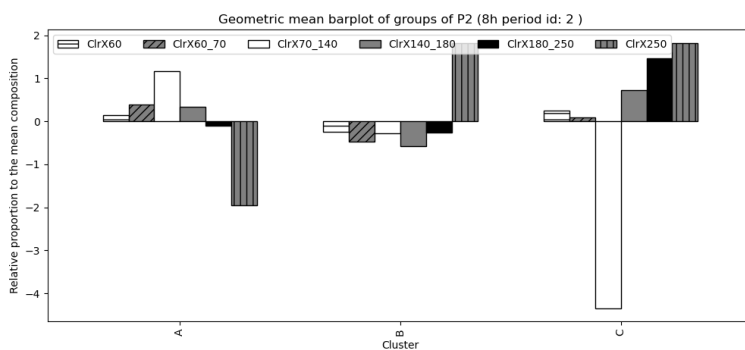
Πίνακας 5.6: Σύγκριση του κέντρου κάθε cluster συνιστώσας με το κέντρο του συνόλου για κάθε cluster (διάστημα: 08:00-15:59)

Ασθενής	Κλάση	X.60	X.60.70	X.70.140	X.140.180	X.180.250	X.250
P1	A	0,067073	0,408297	1,520206	0,834186	-0,66474	-2,16502
	B	-0,39078	-0,43716	0,164793	-0,69209	-0,26111	1,616353
	C	0,424118	-0,19208	-3,27583	-0,67967	1,702507	2,020961
P2	A	0,150597	0,388076	1,172403	0,337661	-0,09703	-1,9517
	B	-0,23775	-0,47575	-0,2832	-0,57334	-0,25466	1,824714
	C	0,24823	0,091971	-4,3384	0,717628	1,471483	1,809088
P3	A	1,695065	1,928265	0,764737	-0,60732	-2,36172	-1,41902
	B	-0,25687	-0,49069	0,770679	0,895348	0,84264	-1,76111
	C	-0,79614	-0,7568	-1,02411	-0,31724	0,745692	2,148594
P4	A	0,254524	0,220575	0,691916	0,385014	0,169039	-1,72107
	B	-0,34902	-0,29179	-0,43433	-0,42892	-0,31636	1,82042
	C	0,530369	0,342256	-4,21723	-0,28712	1,282457	2,349263
P5	A	0,22021	0,287942	1,308743	0,454022	-0,35216	-1,91876
	B	-0,37573	-0,39985	-0,1183	-0,32718	-0,01631	1,23737
	C	0,267908	0,106459	-4,047	-0,64091	1,217339	3,096212
P6	A	2,191697	1,832382	1,432625	-0,04878	-4,01577	-1,39215
	B	-0,24203	-0,11744	0,364799	0,733513	0,811299	-1,55014
	C	-0,13051	-0,19901	-0,63907	-0,76805	-0,15036	1,887004
P7	A	0,961864	1,223471	0,762723	0,056864	-1,50586	-1,49906
	B	-0,38063	-0,64681	0,993059	0,57639	0,552805	-1,09481
	C	-0,80168	-0,83382	-2,11557	-0,73184	1,304329	3,178582
P8	A	0,608601	1,737405	1,054039	0,130799	-3,47579	-0,05505
	B	0,067325	-0,30106	0,026506	0,191763	0,611794	-0,59633
	C	-0,74258	-0,02355	-0,90318	-0,91309	0,00672	2,575693
P9	A	0,095648	0,231038	0,826332	0,918023	0,518065	-2,58911
	B	0,469963	1,20586	-0,2414	-1,31692	-0,96384	0,846338
	C	-0,49847	-1,26463	-0,61942	0,210762	0,308086	1,863674
P10	A	0,823606	1,228464	0,754552	0,099442	-1,47199	-1,43408
	B	-0,3423	-0,83832	0,940772	0,511784	0,755973	-1,02791
	C	-0,78539	-0,79687	-2,1531	-0,72696	1,238868	3,22344
P11	A	0,483429	0,58468	0,94612	0,482945	0,012985	-2,51016
	B	-0,38808	-0,38861	-0,14415	-0,39759	-0,11806	1,436477
	C	0,240836	-0,24165	-3,58999	0,305947	0,716837	2,568019

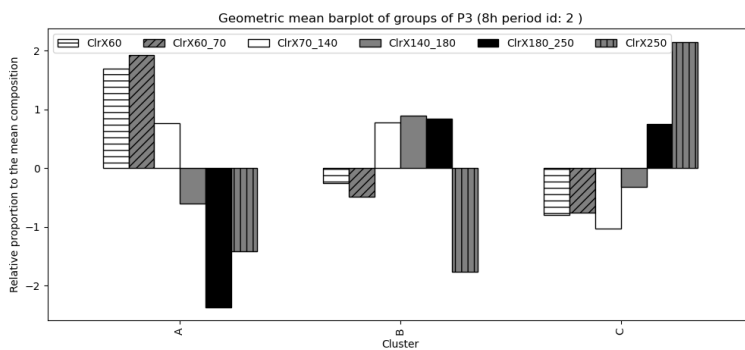
(a) Ασθενής P1



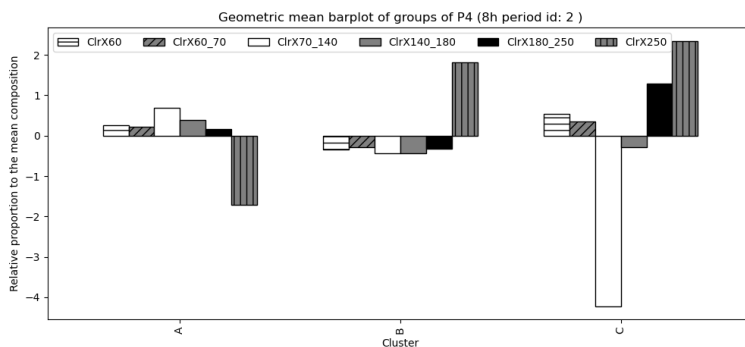
(b) Ασθενής P2



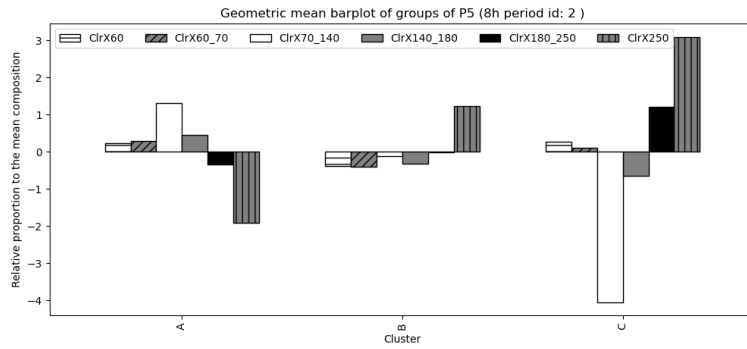
(c) Ασθενής P3



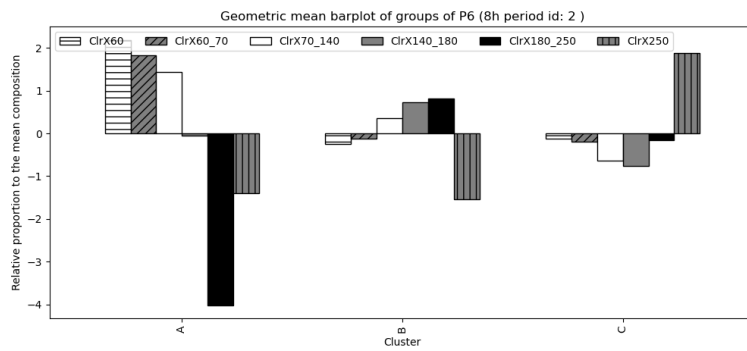
(d) Ασθενής P4



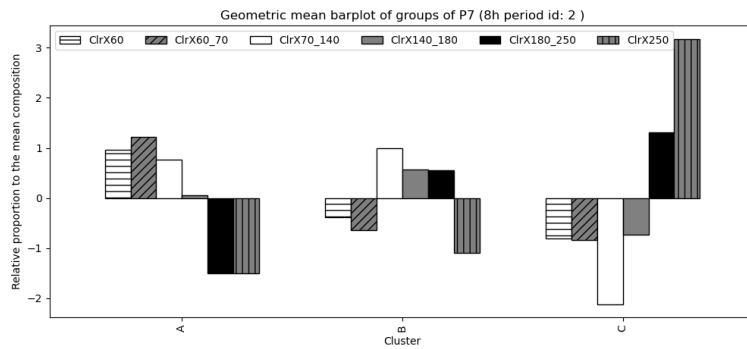
(e) Ασθενής P5



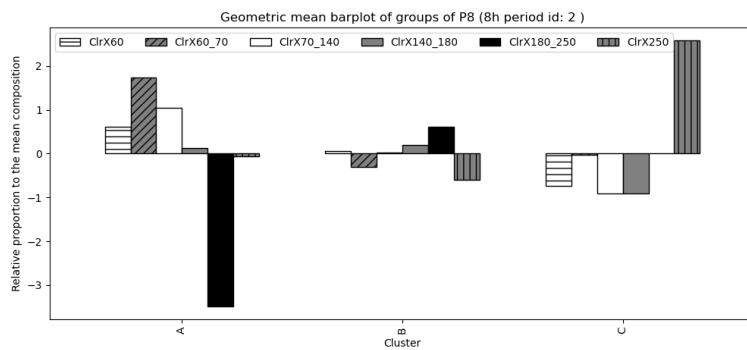
(f) Ασθενής P6



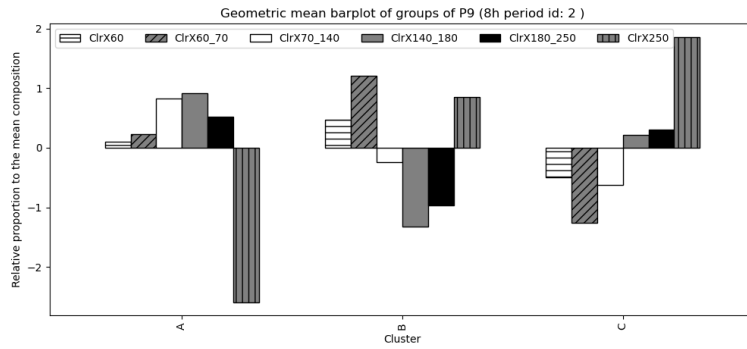
(g) Ασθενής P7



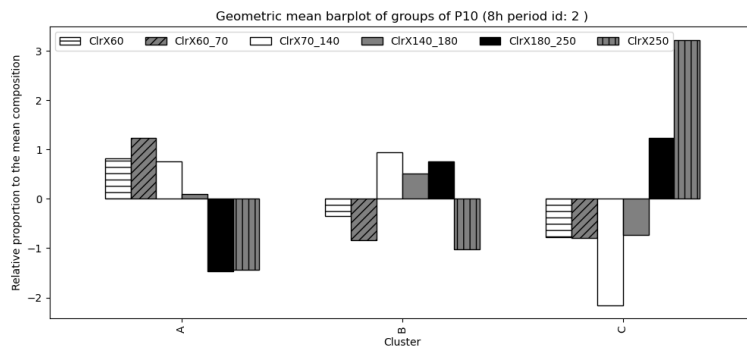
(h) Ασθενής P8



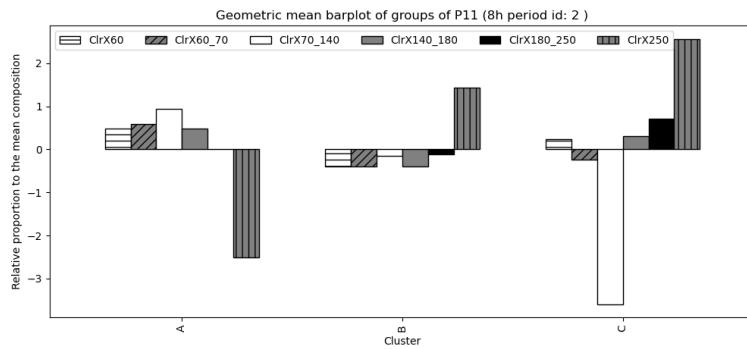
(i) Ασθενής P9



(j) Ασθενής P10



(k) Ασθενής P11



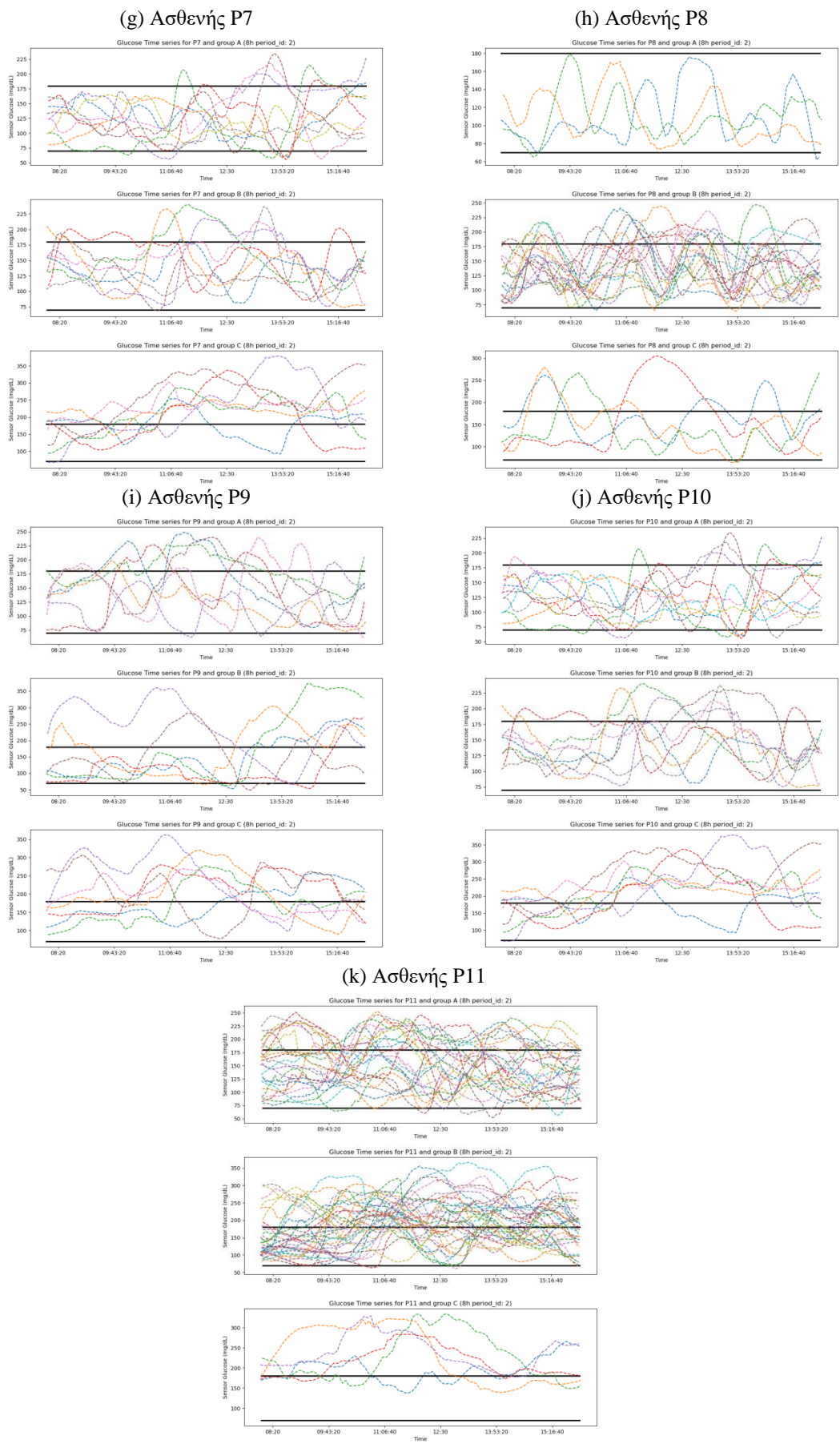
Σχήμα 5.5: Bar plots των μέσων clr συνιστωσών για κάθε cluster και για κάθε ασθενή για την δεύτερη 8-h περίοδο της ημέρας (08:00-15:59). Σύγκριση του κέντρου κάθε ομάδας με το συνολικό κέντρο (μηδέν). Θετικές τιμές υποδηλώνουν υψηλές τιμές στο συγκεκριμένο εύρος γλυκόζης, ενώ αρνητικές τιμές υποδηλώνουν χαμηλές τιμές.

Πίνακας 5.7: Συγκεντρωτικός πίνακας κλινικών δεδομένων ανά ομάδα (διάστημα: 08:00-15:59)

Ασθενής	Κλάση	Avg BG (mg/dL)	SD (mg/dL)	LBGI	HBGI	#Hypo (events/ period)	T Hypo (min/ event)	#Hyper (events/ period)	T Hyper (min/ event)	% Time in Range	% Time Below 70	% Time Above 180
P1	A (14)	144,58	33,97	0,454	3,864	0,071	1,071	0	0	80,02	1,72	18,26
	B (10)	185,42	66,47	0,556	11,692	0,100	4	1,100	96,750	51,14	2,59	46,27
	C (7)	225,74	41,78	0,253	17,831	0,143	4,286	1,143	87,143	16,72	1,43	81,85
P2	A (14)	143,05	36,38	0,379	3,902	0	0	0	0	78,42	0,91	20,67
	B (12)	182,12	63,56	0,427	10,853	0,25	3,542	1,5	68,75	53,26	1,62	45,12
	C (3)	213,97	51,68	0	15,256	0	0	1	75,833	29,51	0,28	70,21
P3	A (7)	124,28	40,32	1,966	2,784	0,857	20,357	0,286	7,857	78,56	9,10	12,34
	B (9)	154,97	38,98	0,447	5,390	0	0	0	0	69,18	0,84	29,98
	C (12)	199,60	53,29	0,196	13,379	0	0	1,417	76,25	35,69	0,85	63,46
P4	A (13)	167,94	39,87	0,173	7,177	0	0	0	0	58,10	0,35	41,56
	B (11)	189,85	54,65	0,146	11,329	0	0	1	46,136	43,82	0,46	55,72
	C (1)	210,44	26,14	0	13,987	0	0	1	30	11,55	0,28	88,17
P5	A (10)	142,97	34,32	0,484	3,831	0,1	2	0	0	79,91	1,07	19,02
	B (8)	180,07	49,84	0,266	9,643	0,125	2,5	0,625	40	48,84	1,29	49,88
	C (3)	211,89	35,22	0	14,536	0	0	1	88,333	29,58	0,28	70,14
P6	A (3)	105,17	24,22	2,207	0,525	0,667	11,667	0	0	92,36	7,36	0,28
	B (18)	151,94	37,31	0,318	4,956	0	0	0	0	70,96	0,78	28,26
	C (17)	170,93	60,12	0,747	9,345	0,294	5,882	1,118	57,696	56,45	2,82	40,73

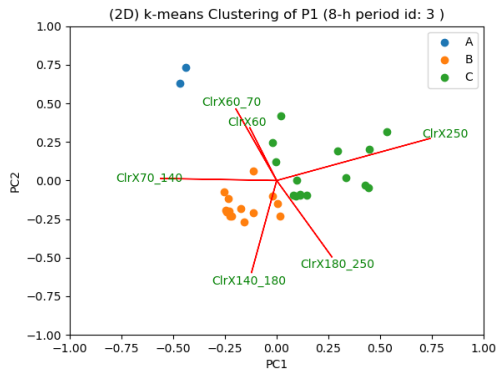
Πίνακας 5.7 (Συνέχεια): Συγκεντρωτικός πίνακας κλινικών δεδομένων ανά ομάδα (διάστημα: 08:00-15:59)

Ασθενής	Κλάση	Avg BG (mg/dL)	SD (mg/dL)	LBGI	HBGI	#Hypo (events/period)	T Hypo (min/event)	#Hyper (events/period)	T Hyper (min/event)	% Time in Range	% Time Below 70	% Time Above 180
P7	A (9)	124,43	34,41	1,246	2,041	0,222	4,444	0	0	87,37	4,93	7,70
	B (8)	141,43	33,83	0,343	3,421	0	0	0	0	82,61	0,39	17,00
	C (7)	214,87	53,07	0,078	16,134	0	0	1,571	65,119	27,50	0,85	71,65
P8	A (3)	109,14	26,62	1,269	0,749	0	0	0	0	97,80	1,92	0,28
	B (17)	139,47	35,79	0,402	3,361	0	0	0	0	82,21	0,56	17,23
	C (4)	153,51	54,20	0,470	6,101	0	0	1	40	70,43	1,77	27,80
P9	A (7)	148,02	40,09	0,731	4,886	0	0	0	0	71,46	1,29	27,27
	B (6)	164,76	72,87	1,520	9,910	0,5	8,333	1,167	71,667	55,35	5,45	39,20
	C (7)	198,61	56,49	0,063	12,987	0	0	1,429	90,952	42,44	0,28	57,28
P10	A (10)	124,84	33,75	1,172	2,004	0,2	4	0	0	87,98	4,56	7,46
	B (8)	145,26	36,09	0,285	3,987	0	0	0	0	78,70	0,28	21,02
	C (7)	214,87	53,07	0,078	16,134	0	0	1,571	65,119	27,50	0,85	71,65
P11	A (24)	152,78	41,15	0,487	5,371	0,042	0,833	0	0	68,21	1,60	30,19
	B (33)	188,20	58,09	0,238	11,526	0	0	1,152	72,35	45,97	0,66	53,37
	C (5)	221,29	47,39	0	16,717	0	0	1,2	119	25,42	0,28	74,30

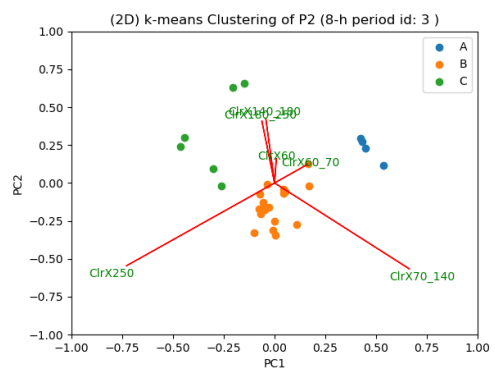


Σχήμα 5.6: 8-h χρονοσειρές γλυκόζης (διάστημα 08:00-15:59) για κάθε ασθενή που δείχνουν τα γλυκαιμικά προφίλ κάθε κλάσης A, B, C

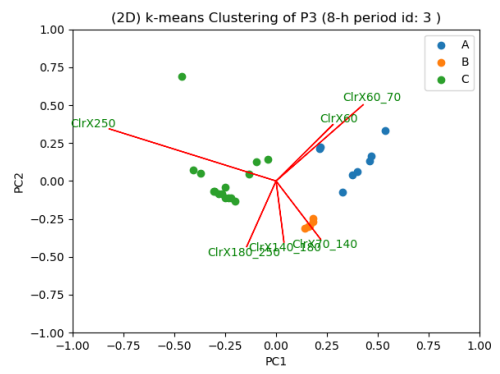
(a) Ασθενής P1



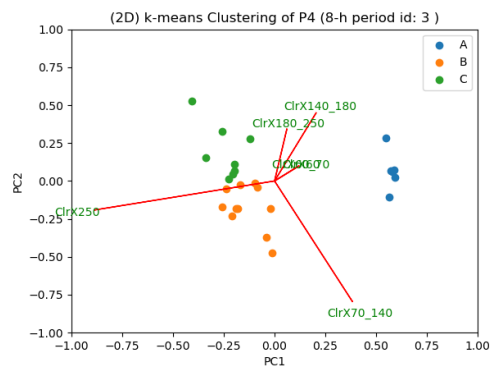
(b) Ασθενής P2



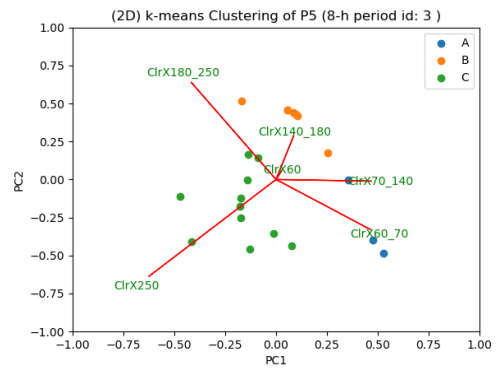
(c) Ασθενής P3



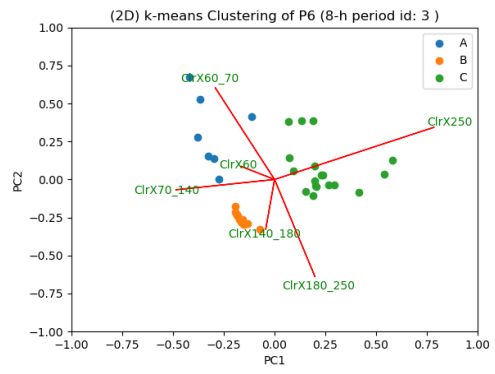
(d) Ασθενής P4



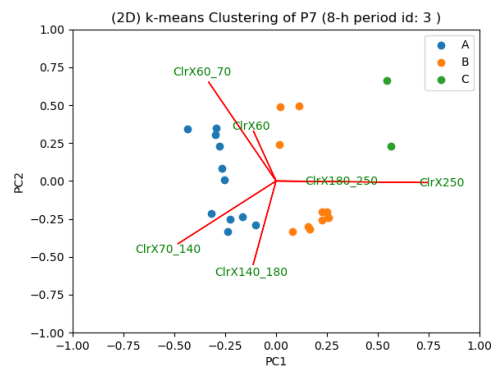
(e) Ασθενής P5



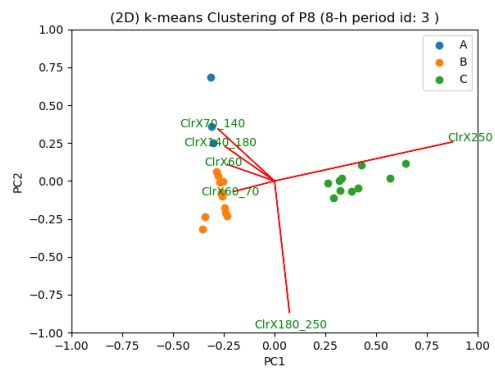
(f) Ασθενής P6

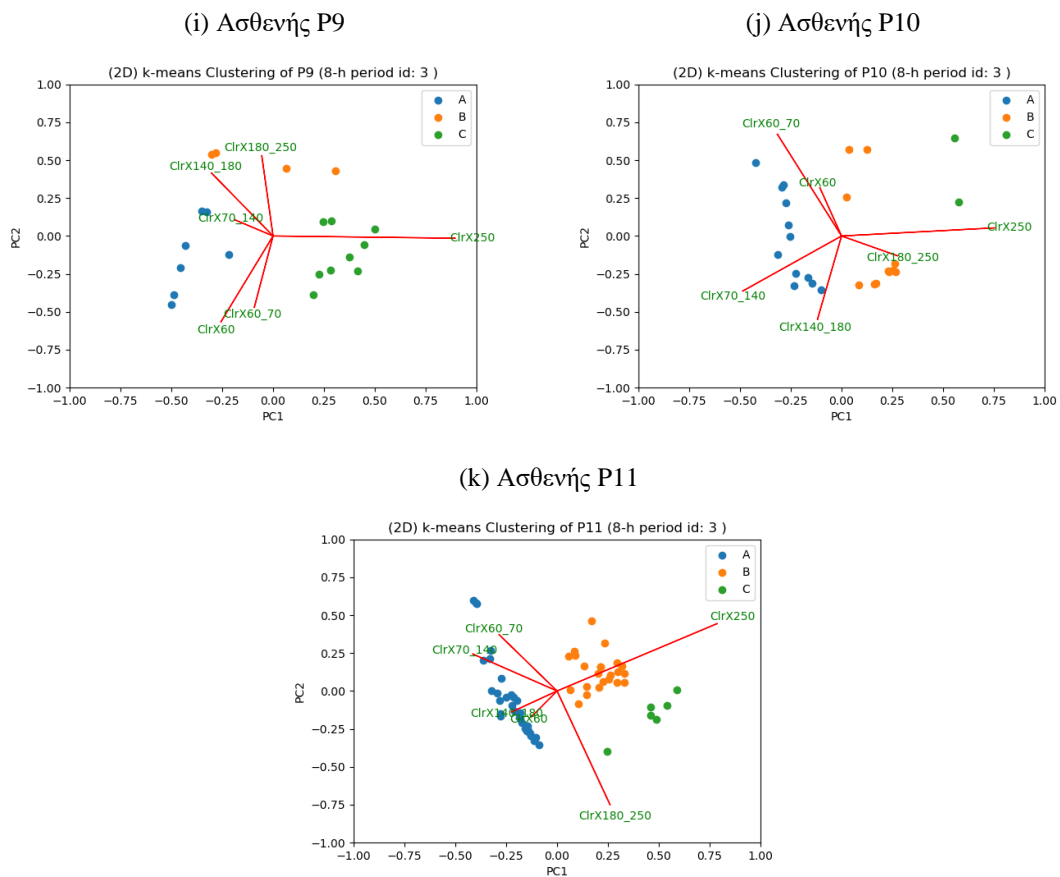


(g) Ασθενής P7



(h) Ασθενής P8





Σχήμα 5.7: Clr biplots των διανυσμάτων που αντιστοιχούν στην 8-h περίοδο 16:00-23:59 για κάθε ασθενή. Οι κουκκίδες αναπαριστούν τα διανύσματα εισόδου στο διδιάστατο χώρο. Ο διαχωρισμός τους σε ομάδες έχει προκύψει ύστερα από την εφαρμογή αλγόριθμου συσταδοποίησης

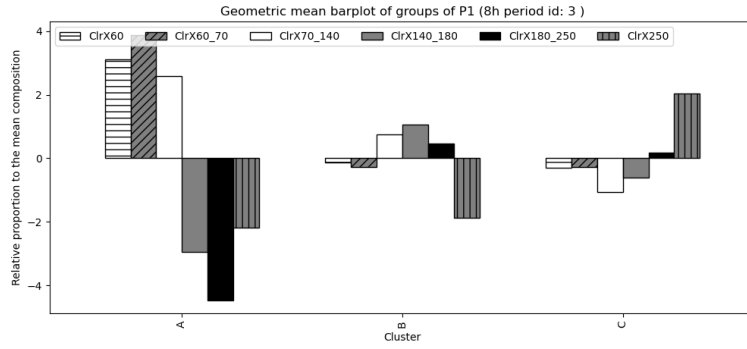
Πίνακας 5.8: Πλήθος δειγμάτων ανά cluster (περίοδος: 16:00-23:59)

Ασθενής	Πλήθος συνολικών ημερών	Πλήθος 8h διανυσμάτων Cluster A	Πλήθος 8h διανυσμάτων Cluster B	Πλήθος 8h διανυσμάτων Cluster C
P1	31	2	14	15
P2	29	4	19	6
P3	28	8	5	15
P4	25	6	11	8
P5	21	3	7	11
P6	38	7	13	18
P7	24	11	11	2
P8	24	3	11	10
P9	20	7	4	9
P10	25	12	11	2
P11	62	33	23	6

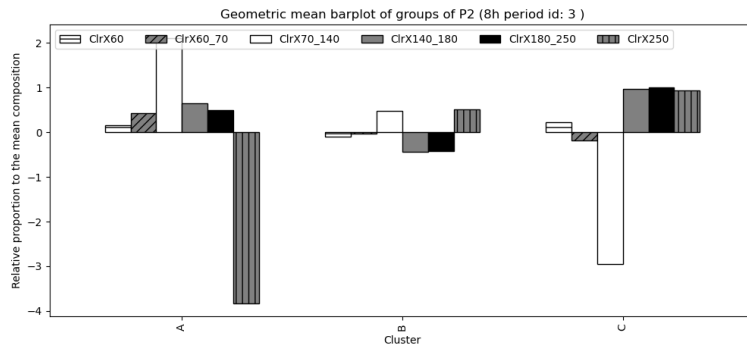
Πίνακας 5.9: Σύγκριση του κέντρου κάθε cluster συνιστώσας με το κέντρο του συνόλου για κάθε cluster (διάστημα: 16:00-23:59)

Ασθενής	Κλάση	X.60	X.60.70	X.70.140	X.140.180	X.180.250	X.250
P1	A	3,110079	3,883541	2,573247	-2,9315	-4,45617	-2,1792
	B	-0,12865	-0,27217	0,759167	1,059097	0,4552	-1,87264
	C	-0,2946	-0,26378	-1,05166	-0,59762	0,169302	2,038358
P2	A	0,147884	0,419832	2,115509	0,640263	0,499739	-3,82323
	B	-0,09908	-0,02926	0,486035	-0,44175	-0,42415	0,508203
	C	0,215156	-0,18722	-2,94945	0,972025	1,009975	0,939509
P3	A	1,445967	2,202177	0,259387	-0,27169	-1,04608	-2,58977
	B	-0,09089	-0,9133	1,478324	0,962782	0,748828	-2,18574
	C	-0,74089	-0,87006	-0,63111	-0,17603	0,308298	2,10979
P4	A	0,466568	0,466568	1,174408	0,821529	0,298664	-3,22774
	B	-0,16007	-0,16007	0,176295	-0,53404	-0,1818	0,859676
	C	-0,12984	-0,12984	-1,12321	0,118164	0,02597	1,238749
P5	A	0,154461	2,956439	1,64133	0,549171	-3,28394	-2,01746
	B	0,212011	-0,28723	0,180593	0,643442	1,211094	-1,95991
	C	-0,17704	-0,62352	-0,56256	-0,55924	0,124923	1,797435
P6	A	1,201965	2,327514	0,676693	-0,58819	-1,59553	-2,02245
	B	0,054306	-0,51293	0,868934	0,722647	0,682532	-1,81549
	C	-0,50665	-0,53469	-0,89072	-0,29317	0,127543	2,097696
P7	A	0,302374	0,858487	1,142522	0,397705	-0,47609	-2,225
	B	-0,34955	-0,78485	-0,37661	-0,10821	0,074645	1,544575
	C	0,259474	-0,40503	-4,21251	-1,59221	2,207961	3,742317
P8	A	0,699945	0,345774	1,013959	1,058441	-2,14185	-0,97627
	B	0,307696	0,235592	0,350361	0,216388	0,440807	-1,55084
	C	-0,54845	-0,36288	-0,68958	-0,55556	0,157666	1,998811
P9	A	0,742249	0,747178	0,693226	0,472108	-0,39957	-2,25519
	B	-0,70831	-1,89277	0,159682	1,327243	1,328761	-0,2146
	C	-0,2625	0,260094	-0,61015	-0,95708	-0,27978	1,849417
P10	A	0,269492	0,750468	1,100379	0,386785	-0,37753	-2,1296
	B	-0,34246	-0,75117	-0,4256	-0,12872	0,020286	1,627666
	C	0,266567	-0,37135	-4,2615	-1,61273	2,153602	3,825408
P11	A	0,400325	0,580415	0,802948	0,565466	-0,31819	-2,03097
	B	-0,68267	-0,77245	-0,06394	-0,58053	-0,00657	2,106151
	C	0,415111	-0,23121	-4,17112	-0,88471	1,77519	3,096742

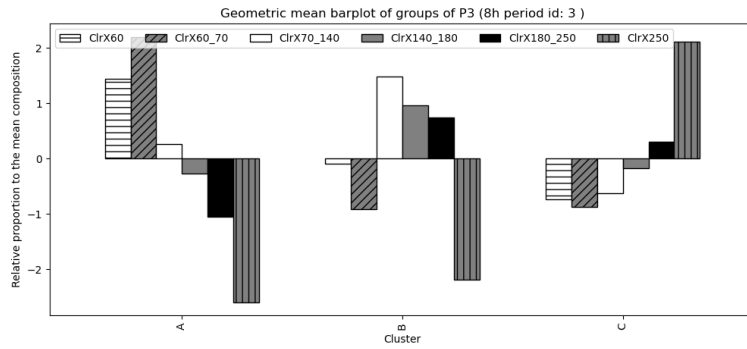
(a) Ασθενής P1



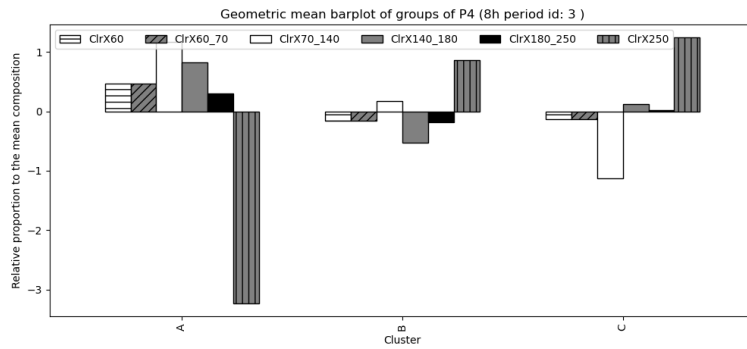
(b) Ασθενής P2



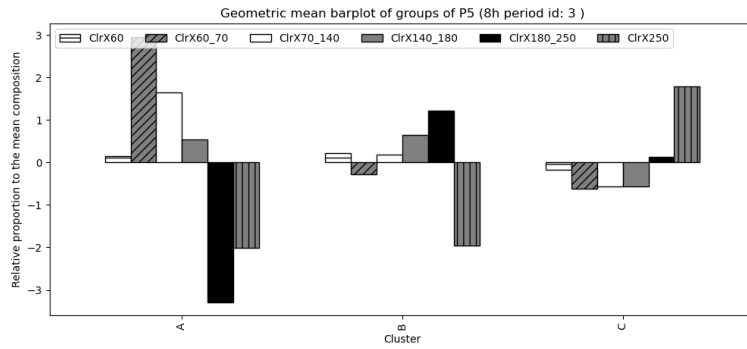
(c) Ασθενής P3



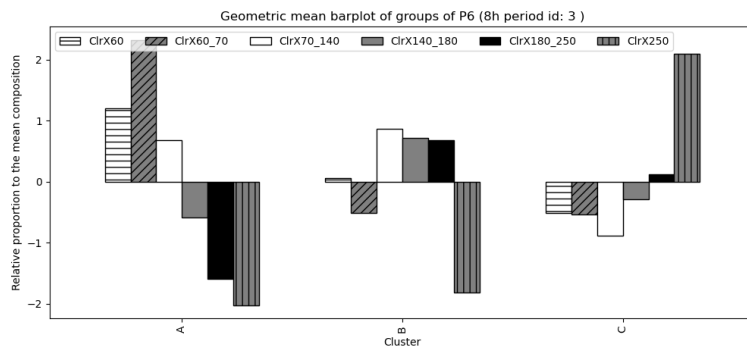
(d) Ασθενής P4



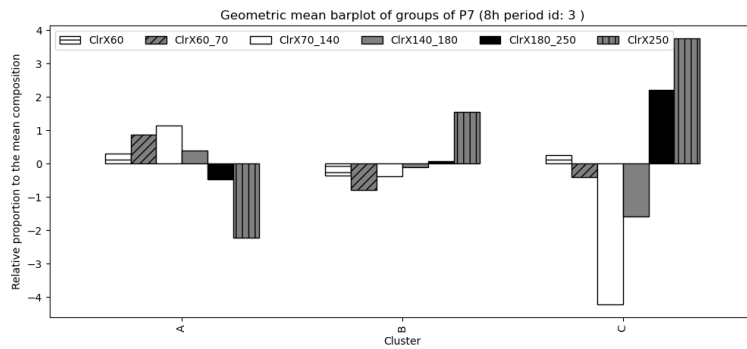
(e) Ασθενής P5



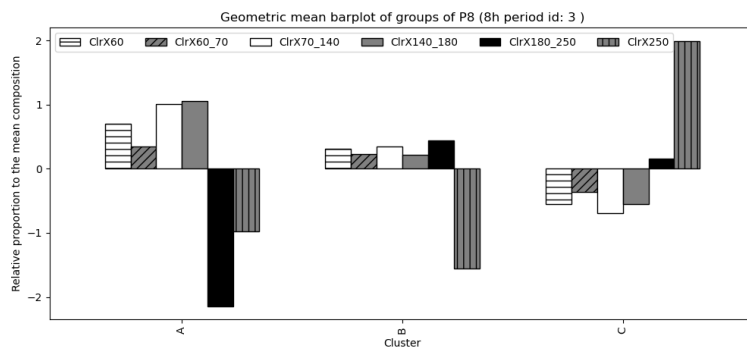
(f) Ασθενής P6



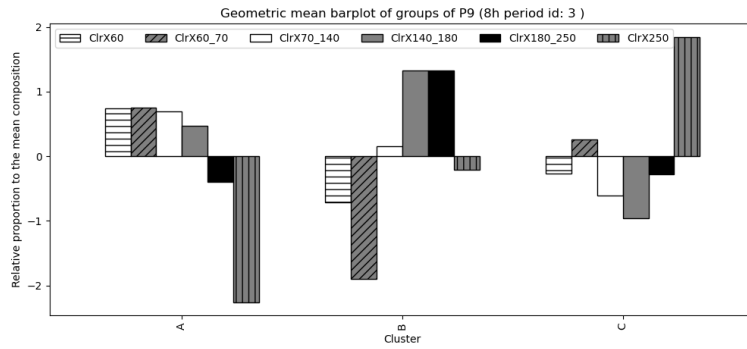
(g) Ασθενής P7



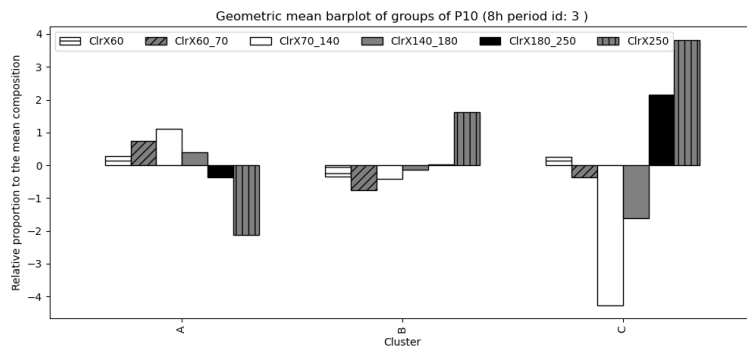
(h) Ασθενής P8



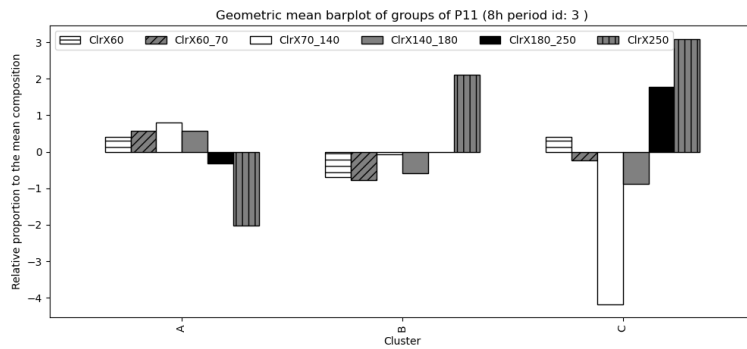
(i) Ασθενής P9



(j) Ασθενής P10



(k) Ασθενής P11



Σχήμα 5.8: Bar plots των μέσων clr συνιστωσών για κάθε cluster και για κάθε ασθενή για την δεύτερη 8-h περίοδο της ημέρας (16:00-23:59). Σύγκριση του κέντρου κάθε ομάδας με το συνολικό κέντρο (μηδέν). Θετικές τιμές υποδηλώνουν υψηλές τιμές στο συγκεκριμένο εύρος γλυκόζης, ενώ αρνητικές τιμές υποδηλώνουν χαμηλές τιμές.

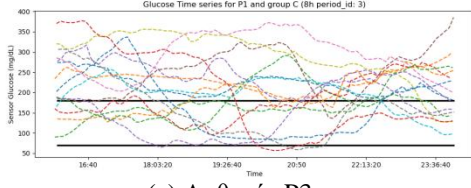
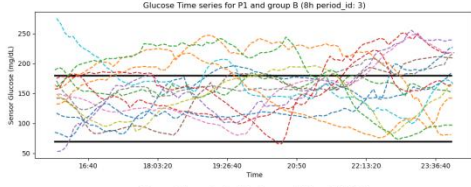
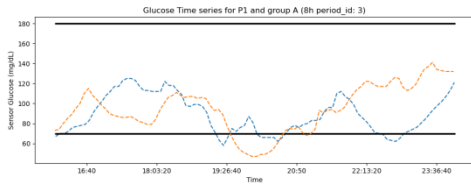
Πίνακας 5.10: Συγκεντρωτικός πίνακας κλινικών δεδομένων ανά cluster (διάστημα: 16:00-23:59)

Ασθενής	Κλάση	Avg BG (mg/dL)	SD (mg/dL)	LBGI	HBGI	#Hypo (events/period)	T Hypo (min/event)	#Hyper (events/period)	T Hyper (min/event)	% Time in Range	% Time Below 70	% Time Above 180
P1	A (2)	91,73	21,36	3,915	0,070	0,5	27,5	0	0	84,15	15,57	0,28
	B (14)	159,22	37,68	0,269	5,950	0,071	1,071	0,071	1,071	67,99	0,62	31,39
	C (15)	207,87	54,26	0,393	15,437	0,067	1,667	1,267	101,5	33,46	1,76	64,78
P2	A (4)	143,63	39,75	0,470	4,190	0	0	0	0	74,46	0,76	24,78
	B (19)	204,20	77,78	0,499	15,529	0,105	2,105	1,421	107,105	42,79	1,90	55,31
	C (6)	226,18	47,80	0	17,865	0	0	1,333	106,111	26,90	0,28	72,82
P3	A (8)	122,85	49,16	3,429	3,135	1	22,188	0	0	67,72	17,87	14,41
	B (5)	141,94	37,87	0,511	3,739	0	0	0	0	76,76	0,28	22,96
	C (15)	202,62	57,99	0,290	14,467	0	0	1,8	91,556	38,40	1,29	60,31
P4	A (6)	157,13	39,82	0,162	5,595	0	0	0	0	66,56	0,28	33,16
	B (11)	188,08	58,04	0,164	11,497	0	0	1,273	71,136	42,87	0,28	56,85
	C (8)	211,18	56,46	0,036	15,072	0	0	1,125	105,313	31,81	0,28	67,91
P5	A (3)	118,59	30,28	1,412	1,238	0	0	0	0	90,63	7,06	2,31
	B (7)	168,18	40,10	0,257	7,309	0	0	0	0	55,01	0,85	44,14
	C (11)	192,12	59,63	0,423	12,746	0,091	3,182	0,909	85,455	44,61	2,22	53,17
P6	A (7)	127,71	41,56	2,004	2,995	0,429	12,857	0,143	2,143	73,77	8,38	17,85
	B (13)	153,50	38,68	0,340	5,204	0	0	0	0	69,42	0,28	30,30
	C (18)	183,70	57,05	0,406	10,970	0	0	1,222	67,314	48,80	0,99	50,21

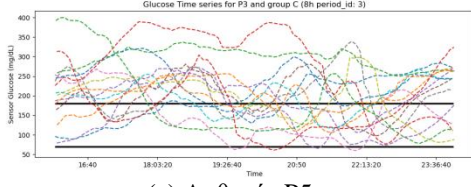
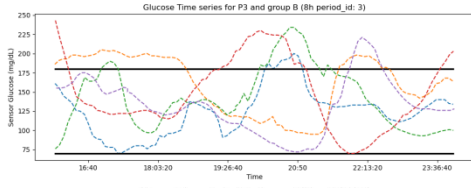
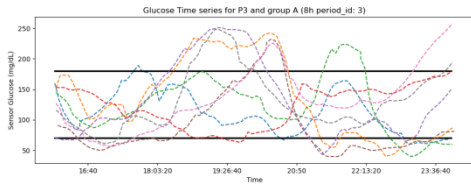
Πίνακας 5.10 (Συνέχεια): Συγκεντρωτικός πίνακας κλινικών δεδομένων ανά cluster (διάστημα: 16:00-23:59)

Ασθενής	Κλάση	Avg BG (mg/dL)	SD (mg/dL)	LBGI	HBGI	#Hypo (events/period)	T Hypo (min/event)	#Hyper (events/period)	T Hyper (min/event)	% Time in Range	% Time Below 70	% Time Above 180
P7	A (9)	130,53	34,52	1,043	2,796	0,182	3,636	0	0	80,58	3,67	15,75
	B (8)	179,57	60,78	0,560	10,468	0,273	4,091	1,091	74,091	51,79	2,48	45,73
	C (7)	246,97	38,45	0	22,041	0	0	3	58,333	4,88	0,28	94,84
P8	A (3)	130,93	20,89	0,257	1,741	0	0	0	0	98,50	0,28	1,22
	B (11)	141,79	36,01	0,307	3,609	0	0	0	0	80,08	0,71	19,20
	C (10)	169,51	55,73	0,176	8,267	0	0	1,1	49,5	62,86	0,67	36,47
P9	A (7)	128,05	38,64	1,549	2,673	0,429	10,714	0	0	82,53	6,02	11,45
	B (4)	162,56	40,42	0,168	6,281	0	0	0,25	12,5	70,59	0,28	29,13
	C (9)	173,95	82,45	1,647	11,320	0,444	10,556	1,222	92,778	52,51	7,12	40,37
P10	A (12)	131,89	36,14	1,012	3,031	0,167	3,333	0	0	79,23	3,39	17,38
	B (11)	179,57	60,78	0,560	10,468	0,273	4,091	1,091	74,091	51,79	2,48	45,73
	C (2)	246,97	38,45	0	22,041	0	0	3	58,333	4,88	0,28	94,84
P11	A (33)	148,04	38,23	0,729	4,695	0,121	2,879	0	0	72,20	2,74	25,06
	B (23)	187,09	59,72	0,255	11,459	0	0	1,087	87,174	46,76	0,70	52,54
	C (6)	224,15	33,26	0	17,057	0	0	1,5	65,417	11,57	0,28	88,15

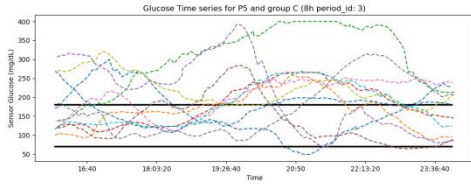
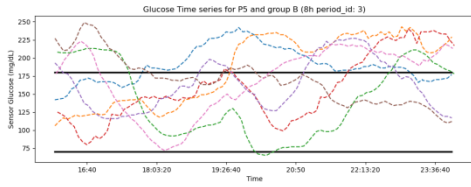
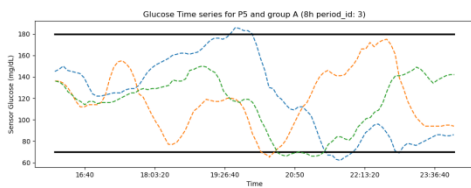
(a) Ασθενής P1



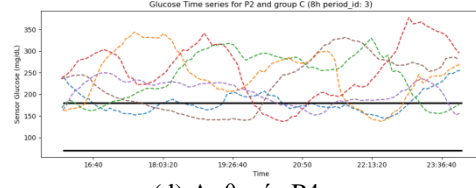
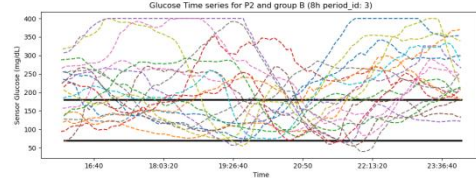
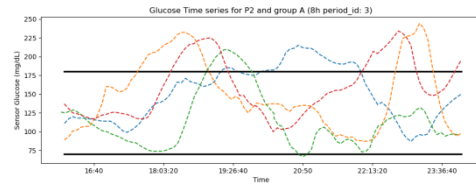
(c) Ασθενής P3



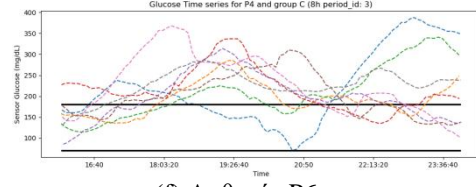
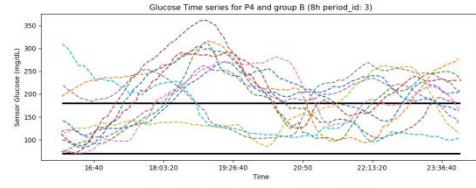
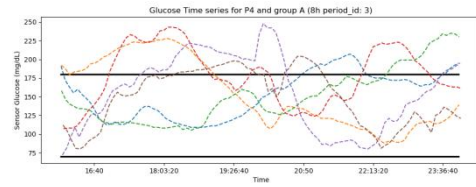
(e) Ασθενής P5



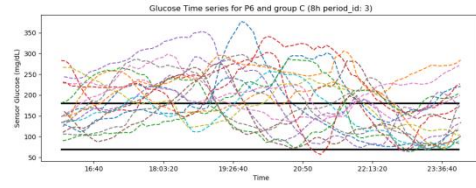
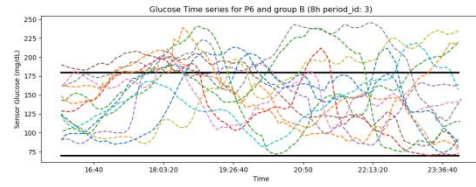
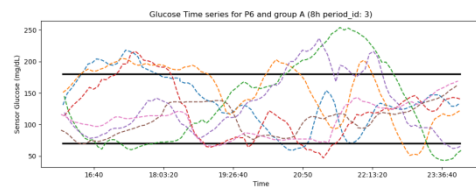
(b) Ασθενής P2

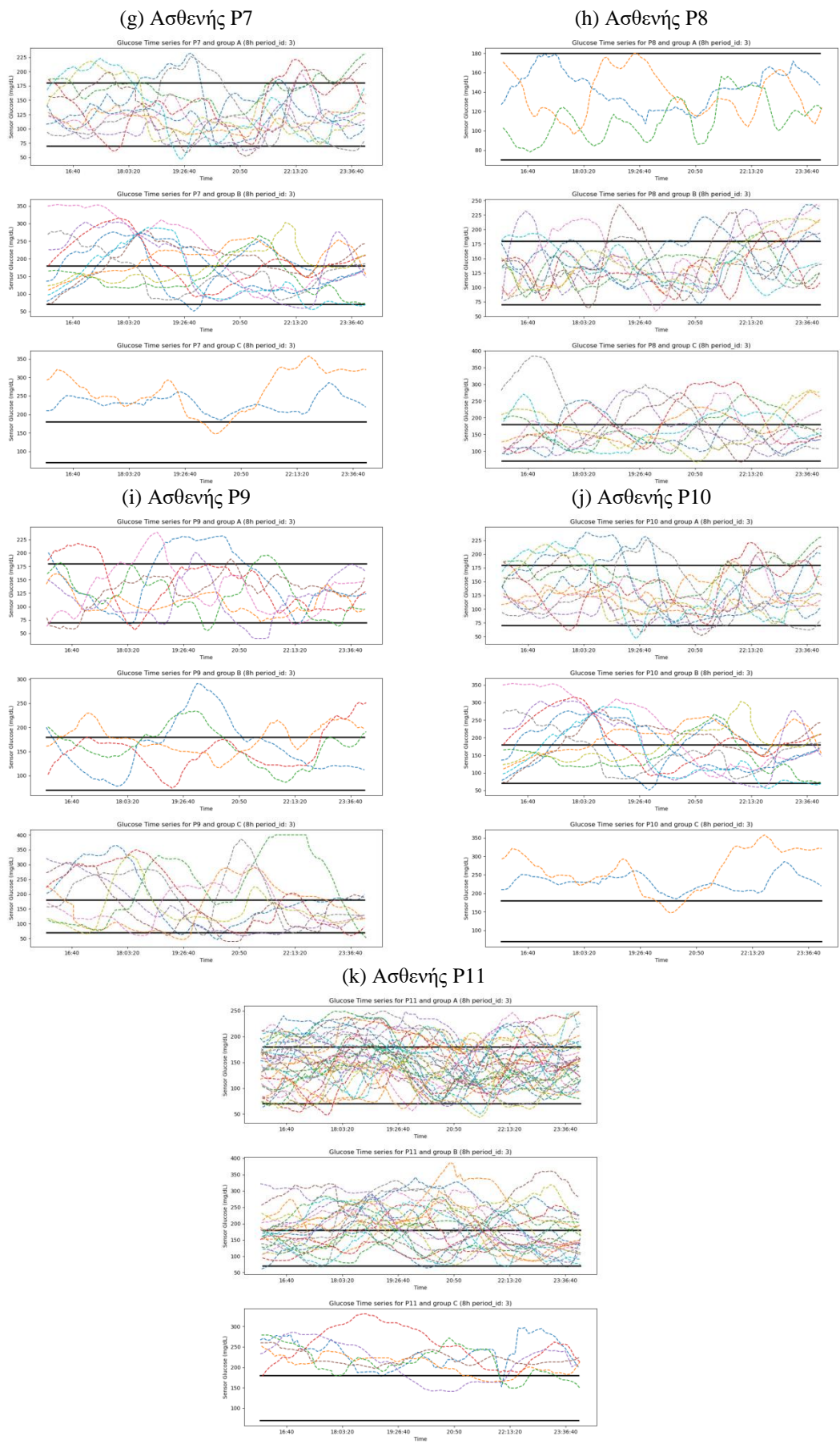


(d) Ασθενής P4



(f) Ασθενής P6





Σχήμα 5.9: 8-*h* χρονοσειρές γλυκόζης (διάστημα 16:00-23:59) για κάθε ασθενή που δείχνουν τα γλυκαιμικά προφίλ κάθε κλάσης A, B, C

5.3 Αποτελέσματα Ταξινόμητων

Στους Πίνακες 5.11, 5.12 και 5.13 παρουσιάζονται οι επιδόσεις των ταξινομητών που ελέγχθηκαν με την μέθοδο LOO-CV για κάθε ασθενή και για καθεμία από τις τρεις περιόδους ανάλυσης ξεχωριστά. Η πιστότητα (accuracy) χρησιμοποιήθηκε ως μετρική απόδοσης των μοντέλων.

Πίνακας 5.11: Πιστότητα των ταξινομητών που δοκιμάστηκαν για την κατηγοριοποίηση γλυκαιμικού προφίλ περιόδου 00:00-7:59

Ασθενής	kNN	Random Forest	XGBoost	SVM	Linear Discriminant	Gaussian Naive Bayes
P1	93,54%	93,54%	87,09%	96,77%	93,54%	90,32%
P2	100%	96,55%	93,10%	96,55%	89,65%	93,10%
P3	92,85%	100%	92,85%	96,42%	92,85%	92,85%
P4	92%	88%	84%	92%	88%	80%
P5	90,47%	85,71%	80,95%	90,47%	85,71%	71,42%
P6	84,21%	84,21%	81,57%	78,94%	84,42%	81,57%
P7	95,83%	100%	100%	100%	95,83%	91,66%
P8	100%	95,83%	91,66%	91,66%	95,83%	91,66%
P9	85%	85%	75%	90%	90%	75%
P10	96%	100%	100%	100%	96%	92%
P11	96,77%	98,38%	96,77%	96,77%	98,38%	96,77%
Μέσος Όρος	92,96%	92,34%	89,40%	93,34%	92,19%	88,33%

Πίνακας 5.12: Πιστότητα των ταξινομητών που δοκιμάστηκαν για την κατηγοριοποίηση γλυκαιμικού προφίλ περιόδου 08:00-15:59

Ασθενής	kNN	Random Forest	XGBoost	SVM	Linear Discriminant	Gaussian Naive Bayes
P1	74,19%	67,74%	70,96%	67,74%	67,74%	64,51%
P2	89,65%	93,10%	86,20%	86,20%	96,55%	93,10%
P3	92,85%	96,42%	82,14%	92,85%	89,28%	96,42%
P4	100%	100%	96%	100%	92%	92%
P5	85,71%	85,71%	90,47%	95,23%	95,23%	80,95%
P6	89,47%	94,73%	94,73%	97,36%	97,36%	92,10%
P7	91,66%	83,33%	75%	87,50%	91,66%	83,33%
P8	95,83%	100%	100%	100%	100%	91,66%
P9	90%	90%	85%	90%	90%	90%
P10	100%	92%	84%	100%	92%	96%
P11	93,54%	96,77%	98,38%	93,54%	96,77%	93,54%
Μέσος Όρος	92,85%	92,23%	89,33%	93,12%	92,07%	88,32%

Πίνακας 5.13: Πιστότητα των ταξινομητών που δοκιμάστηκαν για την κατηγοριοποίηση γλυκαιμικού προφίλ περιόδου 16:00-23:59

Ασθενής	kNN	Random Forest	XGBoost	SVM	Linear Discriminant	Gaussian Naive Bayes
P1	70,96%	80,64%	80,64%	77,41%	67,74%	67,74%
P2	96,55%	89,65%	89,65%	96,55%	100%	89,65%
P3	100%	100%	96,42%	100%	96,42%	96,42%
P4	88%	88%	88%	92%	96%	84%
P5	95,23%	90,47%	85,71%	85,71%	85,71%	80,95%
P6	100%	97,36%	94,73%	100%	97,36%	100%
P7	100%	91,66%	91,66%	95,83%	95,83%	91,66%
P8	91,66%	95,83%	91,66%	100%	95,83%	95,83%
P9	90%	85%	85%	95%	100%	85%
P10	100%	92%	92%	96%	96%	92%
P11	100%	100%	100%	100%	98,38%	100%
Μέσος Όρος	93,36%	92,78%	89,78%	94%	93,11%	88,98%

5.4 Πιθανότητες Μετάβασης

Ο Πίνακας 5.14 παρουσιάζει τις πιθανότητες μετάβασης μεταξύ των κλάσεων διαδοχικών οκτάωρων περιόδων.

Πίνακας 5.14: Πιθανότητες μετάβασης από την κλάση της μίας 8-h περιόδου στην κλάση της ακόλουθης. Για την μετάβαση μεταξύ των 8-h διαστημάτων 16:00-23:59 και 00:00-07:59 χρησιμοποιήθηκαν δεδομένα μόνο από συνεχόμενες ημέρες

n 8-h period	n+1 8-h period	00:00-07:59 (%)			08:00-15:59 (%)			16:00-23:59 (%)		
		A	B	C	A	B	C	A	B	C
P1	A	100	0	0	50	20	30	21,43	57,14	21,43
	B	12,5	37,5	50	54,55	36,36	9,09	0	40	60
	C	36,36	36,36	27,27	30	40	30	0	14,29	85,71
P2	A	100	0	0	30	70	0	21,43	78,57	0
	B	25	62,5	12,5	69,23	15,38	15,38	0	83,33	16,67
	C	0	0	100	33,33	50	16,67	0	66,67	33,33
P3	A	0	75	25	33,34	44,44	22,22	42,86	0	57,14
	B	0	100	0	23,08	30,77	46,15	33,34	22,22	44,44
	C	40	40	20	16,67	16,67	66,66	50	0	50
P4	A	0	100	0	50	50	0	23,08	53,85	23,07
	B	0	85,71	14,29	68,75	31,25	0	0	54,55	45,45
	C	16,67	33,33	50	14,29	71,43	14,29	0	0	100

Πίνακας 5.14 (Συνέχεια): Πιθανότητες μετάβασης από την κλάση της μίας 8-h περιόδου στην κλάση της ακόλουθης. Για την μετάβαση μεταξύ των 8-h διαστημάτων 16:00-23:59 και 00:00-07:59 χρησιμοποιήθηκαν δεδομένα μόνο από συνεχόμενες ημέρες

n 8-h period	n+1 8-h period	00:00-07:59 (%)			08:00-15:59 (%)			16:00-23:59 (%)		
		A	B	C	A	B	C	A	B	C
P5	A	100	0	0	75	25	0	10	30	60
	B	20	80	0	35,71	42,86	21,43	0	50	50
	C	25	50	25	66,67	33,33	0	0	33,33	66,67
P6	A	20	60	20	0	25	75	66,67	33,33	0
	B	33,33	33,33	33,34	11,76	58,82	29,41	22,22	44,44	33,34
	C	0	41,67	58,33	5,88	41,18	52,94	23,53	17,65	58,82
P7	A	40	20	40	28,57	71,43	0	77,78	22,22	0
	B	0	71,43	28,57	44,44	22,22	33,34	62,5	37,5	0
	C	-	-	-	37,5	12,5	50	28,57	71,43	0
P8	A	0	100	0	9,09	81,82	9,09	0	33,33	66,67
	B	28,57	57,14	14,29	18,18	63,64	18,18	5,88	47,06	47,06
	C	71,43	28,57	0	0	50	50	0	0	100
P9	A	0	50	50	20	60	20	42,86	0	57,14
	B	0	50	50	20	30	50	33,33	0	66,67
	C	42,86	57,14	0	80	0	20	14,29	28,57	57,14
P10	A	36,36	18,18	45,46	42,86	57,14	0	80	20	0
	B	0	71,43	28,57	44,44	22,22	33,34	62,5	37,5	0
	C	-	-	-	33,34	22,22	44,44	28,57	57,14	14,29
P11	A	41,67	50	8,33	45,45	54,55	0	66,67	29,17	4,16
	B	36,84	47,37	15,79	38,71	45,16	16,13	36,36	51,52	12,12
	C	0	60	40	22,22	77,78	0	0	40	60

6ο Κεφάλαιο

Συζήτηση

Εισαγωγή

Στο Κεφάλαιο 6 πραγματοποιείται μια ανάλυση των αποτελεσμάτων που παρουσιάστηκαν στο προηγούμενο κεφάλαιο με σκοπό να εξαχθούν χρήσιμες πληροφορίες που θα υποδηλώνουν την κλινική αξία της παρούσας μελέτης.

6.1 Κατηγοριοποίηση οκτώωρων γλυκαιμικών προφίλ

Με σκοπό να χαρακτηριστούν κλινικά οι τρεις κλάσεις που προέκυψαν από την εφαρμογή του αλγορίθμου συσταδοποίησης χρησιμοποιήθηκαν τα διαγράμματα και οι πίνακες της ενότητας 5.2. Παρακάτω ακολουθεί λεπτομερής ανάλυση των συστάδων για κάθε περίοδο, ώστε να ερμηνευθούν κλινικά τα αποτελέσματα της κατηγοριοποίησης. Να σημειωθεί ότι τα αποτελέσματα δεν ερμηνεύονται κατά απόλυτο τρόπο για τον καθένα ασθενή ή για την κάθε περίοδο, αλλά κατά έναν σχετικό τρόπο.

6.1.1 Διαγράμματα Biplots

Τα διαγράμματα διασποράς, που παρατέθηκαν πρώτα στα αποτελέσματα, χρησιμοποιήθηκαν με σκοπό την οπτικοποίηση των δειγμάτων στον διδιάστατο χώρο. Συγκεκριμένα, χρησιμοποιήθηκαν διαγράμματα Biplots ως εργαλεία ερευνητικής ανάλυσης για την ερμηνεία των αποτελεσμάτων. Τα Biplots απεικονίζουν ταυτόχρονα πληροφορίες για την διασπορά των παρατηρήσεων (δειγμάτων) και των μεταβλητών (χαρακτηριστικών) ενός πολυδιάστατου συνόλου δεδομένων. Κατασκευάζονται χρησιμοποιώντας την μέθοδο μείωσης της διαστατικότητας PCA, προκειμένου να επιτευχθεί μια προσέγγιση δεύτερης τάξης του αρχικού πολυδιάστατου χώρου. Οι κάθετοι άξονες αποτελούν τις κύριες συνιστώσες PC1 και PC2 στους οποίους προβάλλονται τα σημεία, ενώ οι γραμμικοί άξονες-ακτίνες με κορυφή O το μηδέν αποτελούν τους άξονες των clr μεταβλητών των διανυσμάτων εισόδου.

Τα μήκη και οι κατευθύνσεις των ακτινών είναι σημαντικά χαρακτηριστικά γιατί υποδεικνύουν πληροφορίες για την διασπορά των σημείων και την συσχέτιση μεταξύ των μεταβλητών. Όσο μεγαλύτερο είναι το μήκος της ακτίνας, τόσο μεγαλύτερη είναι η διασπορά των τιμών για την μεταβλητή που αντιπροσωπεύει. Ακόμη, η γραμμή που ενώνει τις κορυφές δύο αξόνων ονομάζεται σύνδεσμος. Όσο μικρότερο είναι το μήκος του συνδέσμου τόσο λιγότερο διαφέρουν οι τιμές για τις αντίστοιχες μεταβλητές. Ακόμη, το μέγεθος των γωνιών μεταξύ των διανυσμάτων των μεταβλητών καθορίζει τη συσχέτιση μεταξύ τους. Βέλη που δείχνουν προς την ίδια κατεύθυνση, δηλαδή που έχουν μικρή γωνία φανερώνουν θετική συσχέτιση μεταξύ τους. Μια γωνία 90 μοιρών δεν αντιπροσωπεύει συσχέτιση και αντίθετα γωνία 180 μοιρών αντιπροσωπεύει αρνητική συσχέτιση μεταξύ δύο μεταβλητών [47][57][63].

Τα διαγράμματα των *Σχημάτων 5.1, 5.4 και 5.7* αποτελούν τα Biplots για κάθε ασθενή και για καθεμία από την οκτάωρη περίοδο ανάλυσης αντίστοιχα. Η μέθοδος PCA με δύο κύριες συνιστώσες, όπου εξηγούσε γύρω στο 70% της διακύμανσης (πληροφορίας) του συνόλου δεδομένων, χρησιμοποιήθηκε για μείωση της διαστατικότητας. Σημεία πιο κοντά στις καθορισμένες ακτίνες χαρακτηρίζονται από σχετικά υψηλές τιμές στις συνιστώσες που αντιστοιχούν στα αντίστοιχα εύρη γλυκόζης, γεγονός που μπορεί να βοηθήσει στον κλινικό χαρακτηρισμό των κλάσεων.

Biplots που αντιστοιχούν στην ανάλυση των οκτάωρων της περιόδου 00:00-07:59 (Σχήμα 5.1)

Τα μήκη των ακτινών ClrX70_140, ClrX140_180, ClrX180_250 και ClrX250 είναι μεγαλύτερα συγκριτικά με των ClrX60 και ClrX60_70, γεγονός που υποδηλώνει ότι το κάθε διάγραμμα αποτελεί καλή αναπαράσταση για αυτά τα χαρακτηριστικά. Σημεία της κλάσης A φαίνεται να είναι περισσότερο κοντά στους άξονες ClrX70_140, ClrX140_180 που σχετίζονται με ευγλυκαιμία, αλλά και στις μεταβλητές σχετικές με υπογλυκαιμία ClrX60, ClrX60_70. Ωστόσο, παρατηρείται ότι υπάρχουν δείγματα που είναι αρκετά κοντά στην ακτίνα της συνιστώσας ClrX250, όπως φαίνεται στο *Σχήμα 5.1h* για τον ασθενή P8, υποδηλώνοντας ότι ο ασθενής έχει υψηλές τιμές και σε αυτήν την συνιστώσα. Ακόμη, για τους περισσότερους ασθενείς παρατηρείται ότι οι άξονες ClrX180_250 και ClrX250 δείχνουν προς την αντίθετη κατεύθυνση από αυτήν που δείχνουν οι ClrX70_140 και Clr140_180, δηλαδή αναπτύσσουν μεγάλη γωνία και άρα φανερώνουν αρνητική συσχέτιση. Το γεγονός αυτό υποδεικνύει ότι τα σημεία κοντά στις μεταβλητές ClrX180_250 και ClrX250, δηλαδή αυτά της κλάσης C, ανήκουν σε οκτάωρα στα οποία ο ασθενής έχει παρουσιάσει υπεργλυκαιμίες, ενώ ταυτόχρονα έχει περάσει λιγότερο χρόνο σε φυσιολογικά εύρη γλυκόζης. Όσον αφορά τα δείγματα της κλάσης B, για τους ασθενείς P6, P7, P8, P10 και P11 βρίσκονται μεταξύ των ακτινών ClrX140_180 και ClrX180_250, με αυτά του ασθενή P6 να είναι αρκετά κοντά στις ακτίνες σχετικές με υπογλυκαιμία. Για τους ασθενείς P1, P5 και P2, P3 βρίσκονται μεταξύ των αξόνων ClrX70_140, ClrX180_250 και ClrX70_140, ClrX250 αντίστοιχα, ενώ για τους

ασθενείς P4 και P9 βρίσκονται μεταξύ των αξόνων ClrX70_140 και ClrX140_180 που είναι συνυφασμένοι με ευγλυκαιμία.

Biplots που αντιστοιχούν στην ανάλυση των οκτώωρων της περιόδου 08:00-15:59 (Σχήμα 5.4)

Τα μήκη των ακτινών ClrX60, ClrX60_70 φαίνεται να είναι μεγαλύτερα συγκριτικά με αυτά της προηγούμενης περιόδου ανάλυσης. Τα σημεία της κλάσης A βρίσκονται κοντά στους άξονες που σχετίζονται με υπογλυκαιμία και ευγλυκαιμία. Συγκεκριμένα, για τους ασθενείς P1, P2, P4, P11 οι άξονες ClrX60, ClrX60_70, ClrX140_180 έχουν πολύ μικρή γωνία μεταξύ τους και επομένως εμφανίζουν θετική συσχέτιση. Τα διαγράμματα των ασθενών P3, P6, P7, P8 και P10 υποδηλώνουν ότι τα δείγματα της κλάσης A χαρακτηρίζονται από σχετικά υψηλές τιμές σε υπογλυκαιμικά εύρη, εφόσον είναι πλησιέστερα στις αντίστοιχες ακτίνες. Όσον αφορά την κλάση B, φαίνεται ότι χαρακτηρίζεται από σημεία με υψηλές τιμές σε ευγλυκαιμία και υπεργλυκαιμία (P1, P2, P3, P4, P8, P10, P11). Εξαιρέση αποτελούν οι κλάσεις B των ασθενών P5, P9, όπου παρατηρούνται σημεία κοντά και σε υπογλυκαιμικές συνιστώσες εκτός από υπεργλυκαιμικές, και των P6, P7 όπου τα δείγματα βρίσκονται πλησιέστερα στις ακτίνες ClrX70_140, ClrX140_180. Τέλος, τα σημεία που ανήκουν στην κλάση C, για τους περισσότερους ασθενείς, παρατηρούνται μεταξύ των ClrX180_250, ClrX250 και άρα αναμένεται να έχουν υψηλές τιμές σε υπεργλυκαιμικά εύρη.

Biplots που αντιστοιχούν στην ανάλυση των οκτώωρων της περιόδου 16:00-23:59 (Σχήμα 5.7)

Παρόμοια αποτελέσματα εξάγονται και για τα διαγράμματα που αφορούν τις περιόδους που ξεκινούν στις 16:00 και τελειώνουν στις 23:59. Η κλάση A αποτελείται από σημεία που βρίσκονται κατά κόρο πλησίον στις ακτίνες ClrX60, ClrX60_70 και ClrX70_140, επομένως χαρακτηρίζονται από υψηλότερες τιμές σε συνιστώσες σχετικές με φυσιολογικά ή χαμηλά επίπεδα γλυκόζης. Για τους ασθενείς P1, P3, P5, P6 και P9 τα σημεία της κλάσης B βρίσκονται πλησιέστερα στους άξονες ClrX70_140, ClrX140_180 και ClrX180_250. Ακόμη, για τους ασθενείς P2, P4 παρατηρούνται δείγματα της κλάσης B κοντά στην ακτίνα Clr250, φανερώνοντας ότι τα άτομα περνούν χρόνο σε επίπεδα γλυκόζης μεγαλύτερα των 250 mg/dL. Αντίθετα, για τους ασθενείς P7, P10 και P11 κάποια από τα σημεία βρίσκονται πλησίον των ακτινών που σχετίζονται με υπογλυκαιμίες. Τέλος, για τα σημεία της κλάσης C ισχύουν όσα αναφέρθηκαν παραπάνω, δηλαδή χαρακτηρίζονται από υψηλές τιμές σε υπεργλυκαιμία, αφού βρίσκονται κοντά στους άξονες ClrX180_250 και ClrX250.

6.1.2 Διαγράμματα Barplots και Καμπύλες Γλυκόζης

Τα διαγράμματα για κάθε ασθενή των Σχημάτων 5.2, 5.5 και 5.8 αποτελούν τα barplots για τις τιμές που παρουσιάζονται στους αντίστοιχους Πίνακες 5.3, 5.6 και 5.9 αντίστοιχα. Πρακτικά, οι στήλες των πινάκων αυτών αποτελούν τις συνιστώσες του μέσου διανύσματος (κέντρου) της κάθε παραγόμενης κλάσης. Με άλλα λόγια αποτελούν clr συνιστώσες, δηλαδή αναλογίες λογαρίθμων μεταξύ του αντίστοιχου

εύρους γλυκόζης (αριθμητής) και του κέντρου του συνόλου (γεωμετρικός μέσος, παρανομαστής). Οι θετικές τιμές αντικατοπτρίζουν ότι η μέση τιμή για το αντίστοιχο εύρος γλυκόζης υπερβαίνει το γεωμετρικό μέσο του συνόλου, ενώ αρνητικές τιμές υποδηλώνουν το αντίθετο. Τα αποτελέσματα που προκύπτουν συνάδουν με τα αποτελέσματα που εξήχθησαν από τα biplots. Σε συνδυασμό των δύο θα προκύψει ακολούθως ο κλινικός χαρακτηρισμός των ομάδων A, B, C για κάθε 8-h περίοδο.

Από τα διαγράμματα του *Σχήματος 5.2* και τον *Πίνακα 5.3* προκύπτουν τα ακόλουθα συμπεράσματα που αφορούν τον διαχωρισμό των οκταώρων της περιόδου 00:00-07:59. Τα οκτάωρα τύπου A, σχεδόν για όλους τους ασθενείς, παρουσιάζονται με θετικές τιμές στις συνιστώσες που σχετίζονται με υπογλυκαιμία (<60 mg/dL και 60-70mg/dL) και ευγλυκαιμία (70-140 mg/dL και 140-180 mg/dL) και με αρνητικές στις συνιστώσες σχετικές με υπεργλυκαιμία (180-250 mg/dL και >250 mg/dL). Τα οκτάωρα τύπου B, κατά κύριο λόγο, παρουσιάζονται με θετικές τιμές σε συνιστώσες σχετικές με ευγλυκαιμία (140-180 mg/dL) και ήπιας υπεργλυκαιμίας (180-250 mg/dL). Τέλος, τα οκτάωρα τύπου C για τους ασθενείς P3, P6, P7, P8, P9, P10 χαρακτηρίζονται από θετικές και μάλιστα υψηλές τιμές στις δύο συνιστώσες που σχετίζονται με υπεργλυκαιμία (180-250 mg/dL και >250 mg/dL) και αρνητικές σε όλα τα υπόλοιπα. Παρόμοια συμπεριφορά παρατηρείται και για τους ασθενείς P2, P3 με την διαφορά ότι υπάρχουν μικρές μεν, αλλά θετικές τιμές στις συνιστώσες υπογλυκαιμίας (<60 mg/dL και 60-70 mg/dL). Για να γίνει αντιληπτή η διαφορά μεταξύ των δύο ακραίων κλάσεων, A και C, παρατίθενται το εξής παράδειγμα για τον ασθενή P3. Οι τιμές πάνω από 250 mg/dL κατά την διάρκεια των οκταώρων της κλάσης C είναι περίπου 9,48 ($\exp(2,25)$) φορές τον μέσο χρόνο σε αυτό το εύρος γλυκόζης. Η τιμή αυτή μειώνεται σε 8% ($\exp(-2,48)$) για τα δεδομένα της κλάσης A για τον ίδιο ασθενή.

Από τα διαγράμματα του *Σχήματος 5.5* και τον *Πίνακα 5.6* εξάγονται τα ακόλουθα συμπεράσματα που αφορούν τον διαχωρισμό των οκταώρων του διαστήματος 08:00-15:59. Τα οκτάωρα τύπου A, για τους περισσότερους ασθενείς, παρουσιάζονται με θετικές τιμές στις συνιστώσες που σχετίζονται με υπογλυκαιμία (<60 mg/dL και 60-70mg/dL) και ευγλυκαιμία (70-140 mg/dL και 140-180 mg/dL) και με αρνητικές τιμές στις συνιστώσες σχετικές με υπεργλυκαιμία (180-250 mg/dL και >250 mg/dL). Συγκεκριμένα, ένα μέσο οκτάωρο για τους ασθενείς P3, P6, P7 και P10 έχει υψηλότερες τιμές στα υπογλυκαιμικά εύρη από ότι στα φυσιολογικά. Παρατηρείται ότι τα οκτάωρα τύπου B έχουν θετικές τιμές σε συνιστώσες σχετικές με ευγλυκαιμία (70-180 mg/dL ή 140-180 mg/dL) και σοβαρής υπεργλυκαιμίας (250 mg/dL). Ωστόσο, για τον ασθενή P9 παρατηρήθηκαν θετικές τιμές στις συνιστώσες σχετικές με υπογλυκαιμία και αρνητικές τιμές σε συνιστώσες ευγλυκαιμίας. Τα οκτάωρα τύπου C χαρακτηρίζονται από θετικές τιμές στις δύο υπεργλυκαιμικές συνιστώσες (180-250 mg/dL και >250 mg/dL). Παρόλ' αυτά αξίζει να σημειωθεί ότι για τους ασθενείς P1, P2, P4, P5 και P11 οι συνιστώσες που σχετίζονται με σοβαρή υπογλυκαιμία (<60 mg/dL) είναι θετικές μα μικρότερες.

Στη συνέχεια ακολουθούν τα συμπεράσματα που εξάγονται από τα διαγράμματα του *Σχήματος 5.9* και τον *Πίνακα 5.6* για τα οκτάωρα που αφορούν την περίοδο

16:00-23:59. Τα αποτελέσματα είναι παρόμοια με τα προαναφερθέντα. Τα οκτάωρα τύπου A, για τους περισσότερους ασθενείς, παρουσιάζονται με θετικές τιμές στις συνιστώσες που σχετίζονται με υπογλυκαιμία (<60 mg/dL και 60-70mg/dL) και ευγλυκαιμία (70-140 mg/dL ή/και 140-180 mg/dL) και με αρνητικές τιμές στις συνιστώσες σχετικές με υπεργλυκαιμία (180-250 mg/dL και >250 mg/dL). Μάλιστα, για τους ασθενείς P1, P3, P5, P6 και P9 παρατηρούνται υψηλότερες τιμές σε συνιστώσες σχετικές με υπογλυκαιμία, κυρίως στο εύρος 60-70 mg/dL. Το προφίλ των οκτάωρων τύπου B φαίνεται να διαφέρει από ασθενή σε ασθενή. Για τους ασθενείς P1, P2, P3, P4 το προφίλ γλυκόζης τύπου B χαρακτηρίζεται από θετικές τιμές σε συνιστώσες σχετικές με ευγλυκαιμία (70-180 mg/dL ή/και 140-180 mg/dL) και υπεργλυκαιμία (180-250 mg/dL ή >250mg/dL). Παρόμοια συμπεριφορά έχει και το προφίλ γλυκόζης τύπου B των ασθενών P5, P6, P8 με την διαφορά ότι έχουν θετική τιμή, αν και μικρή, στη συνιστώσα σοβαρής υπογλυκαιμίας (<60 mg/dL). Οι υπόλοιποι ασθενείς έχουν θετικές τιμές στα υπεργλυκαιμικά εύρη. Τα οκτάωρα τύπου C χαρακτηρίζονται από θετικές τιμές στις δύο υπεργλυκαιμικές συνιστώσες (180-250 mg/dL και >250 mg/dL). Ωστόσο, για τους ασθενείς P2, P7, P9, P10 και P11 οι συνιστώσες που σχετίζονται με υπογλυκαιμία, αν και με μικρότερες τιμές, είναι επίσης θετικές.

Όλα τα παραπάνω μπορούν να αποτυπωθούν παρατηρώντας, επιπρόσθετα, τις καμπύλες γλυκόζης για κάθε κλάση που παρουσιάζονται στα **Σχήματα 5.3, 5.6 και 5.9**. Στα διαγράμματα των κλάσεων απεικονίζονται οι χρονοσειρές γλυκόζης των οκτάωρων που ανήκουν σε κάθε κατηγορία, όπως αυτές προέκυψαν από τον αλγόριθμο συσταδοποίησης. Οι δύο οριζόντιες συνεχείς γραμμές δηλώνουν τα όρια των φυσιολογικών τιμών γλυκόζης (70-180 mg/dL). Σε ένα γενικότερο πλαίσιο, οι καμπύλες γλυκόζης που ανήκουν στην κλάση A δείχνουν ότι οι περισσότερες τιμές της καμπύλης γλυκόζης βρίσκονται κυρίως εντός των φυσιολογικών ορίων, καθ' όλη την διάρκεια των οκτάωρων, γεγονός που εξηγείται από τις υψηλές τιμές στις συνιστώσες X70_140 και X140_180. Με την ίδια λογική, παρατηρείται ότι το μεγαλύτερο μέρος των καμπυλών που ανήκουν στην κλάση C βρίσκεται πάνω από το ανώτατο φυσιολογικό όριο (>180 mg/dL), γεγονός που επίσης εξηγείται από τις υψηλές τιμές στις συνιστώσες X180_250 και X250. Στην περίπτωση των διαγραμμάτων που αφορούν την κλάση B, παρατηρείται ότι υπάρχουν τμήματα των καμπυλών που βρίσκονται εντός φυσιολογικών ορίων και τμήματα που βρίσκονται πάνω από το ανώτατο όριο των 180 mg/dL. Αυτό πηγάζει από τις υψηλές τιμές στις συνιστώσες X70_140, X140_180 και X180_250 και X250.

6.1.3 Κλινική Ανάλυση των κλάσεων

Στους **Πίνακες 5.4, 5.7 και 5.10** παρουσιάζονται τα κλινικά δεδομένα, για όλους τους ασθενείς του δείγματος, που λαμβάνονται από τις καμπύλες γλυκόζης κάθε κλάσης. Οι αριθμητικοί δείκτες αποτελούν ένα επιπλέον μέτρο αξιολόγησης των αποτελεσμάτων καθώς από αυτές τις τιμές θα φανερωθεί η κλινική ομοιογένεια των ομάδων.

Στον συγκεντρωτικό **Πίνακα 5.4** φαίνονται οι μέσες τιμές των κλινικών χαρακτηριστικών για τα οκτάωρα του διαστήματος 00:00-07:59 που ανήκουν σε καθεμία από τις παραγόμενες ομάδες γλυκόζης A, B ή C. Η χαμηλότερη μεταβλητότητα γλυκόζης (SD (mg/dL)) παρατηρείται για τα προφίλ γλυκόζης της κλάσης υπο/ευγλυκαιμίας A για όλους τους ασθενείς, εκτός από τον P2. Αντίθετα, η μεγαλύτερη τυπική απόκλιση γλυκόζης, για όλους τους ασθενείς εκτός από των P2, P3, παρουσιάζεται στα δείγματα της κλάσης υπεργλυκαιμίας C. Για όλους τους ασθενείς, τα σημεία που ανήκουν στην κλάση A, παρουσιάζονται με την μικρότερη μέση τιμή γλυκόζης (εντός των ορίων 70-180 mg/dL) και τη μικρότερη τιμή του δείκτη HbG1, που σχετίζεται με κίνδυνο για υπεργλυκαιμικά επεισόδια. Αυτό μπορεί να γίνει επίσης αντιληπτό από τα υψηλά ποσοστά του χρόνου σε φυσιολογικά επίπεδα γλυκόζης (%Time in range) αλλά και την μη ύπαρξη υπεργλυκαιμικών επεισοδίων. Για τους ασθενείς P1, P3, P6, P8 και P10 παρατηρείται ότι η κλάση A περιέχει δείγματα με υπογλυκαιμικά επεισόδια, τα οποία ωστόσο είναι μικρής διάρκειας με εξαίρεση εκείνα του ασθενή P6 (μέση διάρκεια 34 min/event) όπου παρατηρείται υψηλότερος κίνδυνος για υπογλυκαιμία (LBGI > 1,1).

Στον αντίποδα βρίσκονται τα οκτάωρα της κλάσης C. Για όλους τους ασθενείς, τα σημεία που ανήκουν στην κλάση C, παρουσιάζονται συγκριτικά με τις υπόλοιπες κλάσεις με την μεγαλύτερη μέση τιμή γλυκόζης (>180 mg/dL) και τη μεγαλύτερη τιμή του δείκτη που σχετίζεται με τον κίνδυνο υπεργλυκαιμικών επεισοδίων HbG1 (HbG1 > 9). Εξάλλου, για όλους τους ασθενείς, το ποσοστό του χρόνου σε υπεργλυκαιμία (%Time above 180) για κάθε οκτάωρο της κλάσης C είναι υψηλότερο από ότι σε φυσιολογικά και υπογλυκαιμικά επίπεδα καθώς, ενώ για όλα τα δείγματα παρατηρούνται υπεργλυκαιμικά επεισόδια μεγάλης διάρκειας.

Όσον αφορά την κλινική συμπεριφορά των κλάσεων τύπου B, παρατηρείται ότι, σχεδόν για όλους τους ασθενείς, τα ποσοστά του χρόνου διαμοιράζονται σε ευγλυκαιμία και σε υπεργλυκαιμία με αυτά της ευγλυκαιμίας να υπερτερούν εκτός από του P3. Οι τιμές HbG1 των κλάσεων B είναι μέτριες ($4,5 < \text{HbG1} < 9$) με εξαίρεση των ασθενών P2 και P3 που είναι υψηλές. Ακόμη, για τους ασθενείς P1, P2, P3 και P4 που παρατηρούνται υπεργλυκαιμικά επεισόδια στα δείγματα της κλάσης B, διαπιστώνεται ότι είναι μικρότερης διάρκειας από τα αντίστοιχα των δειγμάτων της κλάσης C.

Με την ίδια λογική ερμηνεύονται οι αριθμητικοί δείκτες των **Πινάκων 5.7** και **5.10** που αφορούν τις περιόδους 08:00-15:59 και 16:00-23:59 αντίστοιχα. Σε γενικές γραμμές, βγαίνουν τα εξής συμπεράσματα για την κάθε κατηγορία:

- **Κλάση A:**

Θα μπορούσε να χαρακτηριστεί ως κατηγορία ευγλυκαιμίας με στοιχεία υπογλυκαιμιών. Συγκριτικά με τις άλλες κατηγορίες παρουσιάζει

- την χαμηλότερη μέση τιμή γλυκόζης (Avg BG (mg/dL)), τιμές σε εύρος 70-180 mg/dL
- την χαμηλότερη τυπική απόκλιση (SD mg/dL)

- ενίστε την υψηλότερη μέση τιμή του δείκτη σχετικού με υπογλυκαιμία (LBGI), συνήθως τιμές ελάχιστου ή χαμηλού κινδύνου ($LBGI \leq 1,1$ ή $1,1 < LBGI \leq 2,5$)
 - την χαμηλότερη μέση τιμή του δείκτη σχετικού με υπεργλυκαιμία (HBGI), συνήθως τιμές χαμηλού κινδύνου ($HBGI \leq 4,5$)
 - ημέρες με υπογλυκαιμικά επεισόδια (#Hypo) μικρής διάρκειας και κανένα υπεργλυκαιμικό επεισόδιο (#Hyper)
 - ημέρες με υψηλά ποσοστά του χρόνου σε φυσιολογικές τιμές γλυκόζης (% Time in Range)
- **Κλάση B:**
Θα μπορούσε να χαρακτηριστεί ως κατηγορία ευγλυκαιμίας με στοιχεία υπεργλυκαιμίας. Συγκριτικά με τις άλλες κατηγορίες παρουσιάζει
 - την δεύτερη μέση τιμή γλυκόζης (Avg BG (mg/dL)), τιμές σε εύρος 140-250 mg/dL
 - ενίστε την μεγαλύτερη ή την δεύτερη μεγαλύτερη τυπική απόκλιση (SD mg/dL)
 - ενίστε την χαμηλότερη μέση τιμή του δείκτη σχετικού με υπογλυκαιμία (LBGI), συνήθως τιμές ελάχιστου κινδύνου ($LBGI \leq 1,1$)
 - την δεύτερη μεγαλύτερη μέση τιμή του δείκτη σχετικού με υπεργλυκαιμία (HBGI), συνήθως τιμές μέτριου ή υψηλού κινδύνου ($4,5 \leq HBGI \leq 9$ ή $HBGI \geq 9$)
 - συνήθως ημέρες χωρίς υπογλυκαιμικά επεισόδια (#Hypo) ή πολύ μικρής διάρκειας και υπεργλυκαιμικά επεισόδια (#Hyper) με σχετικά μεγάλη διάρκεια
 - ημέρες με υψηλά ποσοστά του χρόνου σε φυσιολογικές τιμές γλυκόζης (% Time in Range) και σε υπεργλυκαιμία (% Time above 180)
 - **Κλάση C:**
Θα μπορούσε να χαρακτηριστεί ως κατηγορία υπεργλυκαιμίας Συγκριτικά με τις άλλες κατηγορίες παρουσιάζει
 - την μεγαλύτερη μέση τιμή γλυκόζης (Avg BG (mg/dL)), τιμές σε εύρος μεγαλύτερο των 180 mg/dL
 - ενίστε την μεγαλύτερη ή την δεύτερη μεγαλύτερη τυπική απόκλιση (SD mg/dL)
 - την χαμηλότερη μέση τιμή του δείκτη σχετικού με υπογλυκαιμία (LBGI), συνήθως τιμές ελάχιστου κινδύνου ($LBGI \leq 1,1$) αλλά και μηδενικές τιμές
 - την μεγαλύτερη μέση τιμή του δείκτη σχετικού με υπεργλυκαιμία (HBGI), τιμές υψηλού κινδύνου ($HBGI \geq 9$)
 - συνήθως ημέρες χωρίς υπογλυκαιμικά επεισόδια (#Hypo) ή πολύ μικρής διάρκειας και υπεργλυκαιμικά επεισόδια (#Hyper) με διάρκεια μεγαλύτερης αυτών της κλάσης B
 - ημέρες με υψηλά ποσοστά του χρόνου σε υπεργλυκαιμία (% Time above 180)

Ωστόσο, η ανάλυση πραγματοποιείται ατομικά για κάθε ασθενή και επομένως είναι λογικό να υπάρχουν διαφοροποιήσεις των παραπάνω αξιολογήσεων από ασθενή σε ασθενή, δεδομένου ότι ο γλυκαιμικός έλεγχος διαφέρει επίσης. Για παράδειγμα παρατηρήθηκε ότι οι τιμές του ασθενή P8 για κάθε κλάση είναι αρκετά μικρότερες από τις αντίστοιχες των άλλων ασθενών και για τις τρεις περιόδους. Πιο συγκεκριμένα, η κλάση C, που γενικά μπορεί να χαρακτηριστεί ως κλάση υπεργλυκαιμίας, χαρακτηρίζεται από μέση τιμή γλυκόζης εντός των φυσιολογικών επιπέδων γλυκόζης (158,21 mg/dL, 153,51 mg/dL, 169,51 mg/dL). Ακόμη, στις άλλες δύο κλάσεις δεν παρατηρούνται ούτε υπογλυκαιμικά ούτε υπεργλυκαιμικά επεισόδια. Όλα τα παραπάνω οδηγούν στο συμπέρασμα ότι ο ασθενής αυτός πετυχαίνει καλύτερο και πιο ισορροπημένο γλυκαιμικό έλεγχο, συγκριτικά με τους υπόλοιπους.

Από τα αποτελέσματα συμπεραίνεται ότι ο διαχωρισμός των οκταώρων κάθε περιόδου ανάλυσης σε τρεις συστάδες δημιουργεί ομάδες που αντιπροσωπεύουν διαφορετικές περιπτώσεις διαχείρισης του μεταβολισμού γλυκόζης που σχετίζονται είτε με υπογλυκαιμίες και υπεργλυκαιμίες, είτε με καλό γλυκαιμικό έλεγχο.

6.2 Αξιολόγηση Ταξινομητών

Για την αξιολόγηση των μοντέλων ταξινόμησης χρησιμοποιήθηκε η μετρική της πιστότητας (accuracy). Την βάση για την εξαγωγή των μετρικών απόδοσης των ταξινομητών αποτελεί ο πίνακας σύγχυσης. Ο πίνακας σύγχυσης C είναι τέτοιος ώστε το στοιχείο $C_{i,j}$ να είναι ίσο με τα δείγματα που ενώ ανήκουν στην κατηγορία i ταξινομούνται στην κατηγορία j . Τα στοιχεία της διαγωνίου είναι αληθινά θετικά δείγματα (true positive ή TP) της κάθε κλάσης. Για κάθε κλάση i τα στοιχεία της γραμμής i , εκτός της διαγωνίου, είναι λανθασμένα αρνητικά δείγματα (false negative ή FN) της κλάσης και τα στοιχεία της στήλης i εκτός της διαγωνίου είναι λανθασμένα θετικά δείγματα (false positive ή FP) της κλάσης.

Πραγματική Κλάση	Προβλεπόμενη Κλάση		
	A	B	C
A	$n_{1,1}$	$n_{1,2}$	$n_{1,3}$
B	$n_{2,1}$	$n_{2,2}$	$n_{2,3}$
C	$n_{3,1}$	$n_{3,2}$	$n_{3,3}$

Η πιστότητα δίνεται από τον παρακάτω τύπο:

$$\text{accuracy} = \frac{n_{1,1} + n_{2,2} + n_{3,3}}{n_{1,1} + n_{1,2} + n_{1,3} + n_{2,1} + n_{2,2} + n_{2,3} + n_{3,1} + n_{3,2} + n_{3,3}} \quad (6.1)$$

όπου ο αριθμητής είναι το πλήθος των σωστών προβλέψεων και ο παρανομαστής το πλήθος όλων δειγμάτων προς ταξινόμηση. Στην παρούσα εργασία που χρησιμοποιήθηκε η τεχνική LOO-CV τα δείγματα προς ταξινόμηση που ελέγχθηκαν αποτελούν όλο το σύνολο δεδομένων.

Από τον παραπάνω πίνακα μπορούν να εξαχθούν ακόμη οι μετρικές της ακρίβειας (precision) και της ανάκλησης (recall) για κάθε κλάση με σκοπό να φανεί η επίδοση των μοντέλων για την κάθε κλάση ξεχωριστά. Οι τύποι των μετρικών αυτών δίνονται από τις εξής σχέσεις:

$$\text{precision} = \frac{TP}{TP + FP} \quad (6.2)$$

$$\text{recall} = \frac{TP}{TP + FN} \quad (6.3)$$

Για παράδειγμα για την κλάση A το precision και το recall θα προκύψουν ως εξής:

$$\text{precision} = \frac{n_{1,1}}{n_{1,1} + n_{2,1} + n_{3,1}}, \quad \text{recall} = \frac{n_{1,1}}{n_{1,1} + n_{1,2} + n_{1,3}}$$

Ωστόσο, μεταξύ της ακρίβειας και της ανάκλησης υπάρχει γενικά trade-off και αρκετά συχνά χρησιμοποιείται το F1 score, το οποίο αποτελεί τον αρμονικό μέσο των δυο τους.

Στους Πίνακες 5.11, 5.12 και 5.13 παρουσιάζονται οι επιδόσεις των αλγορίθμων που χρησιμοποιήθηκαν για την ταξινόμηση οκταώρων των περιόδων 00:00-07:59, 08:00-15:59 και 16:00-23:00. Γενικά, παρατηρούνται υψηλές επιδόσεις, τάξεως 70% και άνω, για όλους τους αλγόριθμους που χρησιμοποιήθηκαν. Εξάιρεση αποτελούν τα αποτελέσματα του ασθενή P1 για τις περιόδους που ξεκινούν στις 08:00 και στις 16:00, όπου εμφανίζουν χαμηλότερη επίδοση (μέχρι και 67%). Πιο συγκεκριμένα, ο ταξινομητής SVM φαίνεται να δουλεύει καλύτερα, κατά γενική ομολογία για όλους τους ασθενείς και για κάθε περίοδο ανάλυσης, πετυχαίνοντας ποσοστά πιστότητας μεγαλύτερα από 90%. Όπως προαναφέρθηκε παραπάνω αρκετά χαμηλότερο ποσοστό (67,74% και 77,41%) παρουσιάζεται για τον ασθενή P1 (Πίνακας 5.12 και 5.13 αντίστοιχα). Για τον συγκεκριμένο ασθενή παρουσιάζονται οι μετρικές απόδοσης για κάθε κλάση που αφορούν την ταξινόμηση της περιόδου 08:00-15:59, με σκοπό να διαπιστωθεί σε ποια κλάση αποτυγχάνει η ταξινόμηση και να ερμηνευθεί το χαμηλό ποσοστό της τάξης του 60%

	precision	recall	F1-score	Δείγματα
A	93%	93%	93%	14
B	60%	30%	40%	10
C	42%	71%	53%	7
		accuracy	68%	31

Τα χαμηλότερα ποσοστά σημειώνονται στο recall της κλάσης B και στο precision της κλάσης C. Το χαμηλό ποσοστό του recall υποδηλώνει ότι ο ταξινομητής αδυνατεί να ταξινομήσει σωστά δείγματα της κλάσης B. Από την άλλη, το μοντέλο ταξινομεί σωστά δείγματα της κλάσης C αλλά αδυνατεί να ταξινομήσει σωστά τα δείγματα των κλάσεων A, B διότι επιστρέφει σταθερά C, γεγονός που το υποδηλώνει η μεγάλη τιμή της ανάκλασης. Από τα αποτελέσματα του παραπάνω πίνακα φαίνεται ότι ο ταξινομητής δεν λειτουργεί ικανοποιητικά για αυτόν τον ασθενή για την πρόβλεψη των κλάσεων B, C, δηλαδή κλάσεων που σχετίζονται με υπεργλυκαιμία.

Με βάση τα παραπάνω εξάγεται το γενικό συμπέρασμα ότι οι συστάδες που δημιουργούνται μπορούν να χρησιμοποιηθούν για την ταξινόμηση αγνώστου οκτάωρου γλυκαιμικού προφίλ. Τα αποτελέσματα των ταξινομητών της παρούσας μελέτης με τα αντίστοιχα της μελέτης που παρουσιάζεται στην βιβλιογραφία [61] δεν διαφέρουν σημαντικά. Συγκεκριμένα, η ακρίβεια που παρουσιάζεται στην παρούσα εργασία είναι λίγο χαμηλότερη, με την διαφορά ότι για την εξαγωγή των αποτελεσμάτων έχουν χρησιμοποιηθεί δεδομένα παιδιών και όχι ενηλίκων των οποίων ο γλυκαιμικός έλεγχος είναι καλύτερα ρυθμισμένος.

6.3 Αξιολόγηση Μεταβάσεων

Στην παρούσα μελέτη αναλύθηκαν οι μεταβάσεις μεταξύ των προφίλ των διαδοχικών οκτάωρων μέσα στην ημέρα. Σε κάθε μια από τις χρονικές στιγμές της ημέρας 00:00, 08:00 και 16:00 οι μεταβάσεις από το προφίλ του προηγούμενου οκτάωρου στο επακόλουθο του μετρήθηκαν και παρουσιάστηκαν με την μορφή πιθανοτήτων μετάβασης. Ο **Πίνακας 5.14** παρουσιάζει τις τιμές των πιθανοτήτων μετάβασης από έναν τύπο οκτάωρου (A, B ή C) σε οποιοδήποτε άλλο τύπο του επόμενου οκτάωρου.

Για παράδειγμα, έστω ότι για τον ασθενή P2 αναλύεται το γλυκαιμικό προφίλ της περιόδου 00:00-07:59 με την μεθοδολογία που προαναφέρθηκε και με την χρήση ενός ταξινομητή κατηγοριοποιείται ως οκτάωρο τύπου B. Από τις τιμές της δεύτερης στήλης του πίνακα, παρατηρείται ότι η πιθανότητα μετάβασης από οκτάωρο τύπου B στο επόμενο οκτάωρο 08:00-15:59 τύπου A είναι 69.23%. Όπως διαπιστώθηκε από την κλινική ανάλυση του **Πίνακα 5.7**, που αφορά τα προφίλ γλυκόζης της πρωινής περιόδου (08:00-15:59), η κλάση A αποτελεί μια κλάση που χαρακτηρίζεται από υψηλά ποσοστά του χρόνου σε ευγλυκαιμία με την ύπαρξη, ωστόσο, υπογλυκαιμικών επεισοδίων. Έτσι, ο ασθενής ξέροντας το πιο πιθανό προφίλ γλυκόζης της περιόδου που ακολουθεί θα δράσει ρυθμίζοντας κατάλληλα την δόση ινσουλίνης αλλά και τα γεύματα του. Αντίστοιχα, αν το προφίλ γλυκόζης της ίδιας περιόδου ταξινομηθεί ως C, τότε το πιο πιθανό προφίλ (50%) της επόμενης περιόδου είναι τύπου B, γεγονός που υποδηλώνει ότι θα περάσει αρκετό διάστημα του χρόνου σε υπεργλυκαιμικά εύρη. Με την ίδια λογική ερμηνεύονται όλες οι μεταβάσεις μεταξύ διαδοχικών περιόδων.

Όσον αφορά την πρώτη στήλη, που σχετίζεται με τις μεταβάσεις στα προφίλ της περιόδου 00:00-07:59, ο υπολογισμός των πιθανοτήτων απαιτούσε την ύπαρξη

διαδοχικών ημερών. Ωστόσο, στο τελικό σύνολο δεδομένων για κάθε ασθενή δεν υπήρχαν πολλές συναπτές ημέρες ώστε να μην χάνεται πληροφορία για την μετάβαση από την περίοδο 16:00-23:59 στην 00:00-07:59 της επόμενης μέρας. Γι αυτό τον λόγο χρησιμοποιήθηκαν δεδομένα από τις διαθέσιμες συνεχόμενες μεταβάσεις. Έτσι, εξηγείται το γεγονός ότι για τον ασθενή P7 και P10 δεν υπάρχουν προφίλ γλυκόζης τύπου C. Η ανάλυση για την συγκεκριμένη μετάβαση έχει ιδιαίτερη βαρύτητα επειδή αφορά το νυχτερινό προφίλ γλυκόζης, όπου ο χρήστης παρεμβαίνει ελάχιστα στον γλυκαιμικό έλεγχο. Εντούτοις, συμπεραίνεται ότι η ανάλυση για την συγκεκριμένη μετάβαση, αν και συμπεριλήφθηκε για λόγους πληρότητας, δεν μπορεί να ληφθεί υπόψη διότι χάνεται σημαντική πληροφορία εξαιτίας του ελάχιστου όγκου διαθέσιμων δεδομένων.

Να σημειωθεί ακόμα ότι σε αρκετές περιπτώσεις οι πιθανότητες μεταβάσεων από κάποια κατηγορία είναι ομοιόμορφα κατανεμημένες μεταξύ των ομάδων A, B και C, με αποτέλεσμα να μην είναι ιδιαίτερα βοηθητικές για τον ασθενή ή τον κλινικό κατά την διαδικασία της λήψης αποφάσεων. Ωστόσο, είναι σημαντικό να τονισθεί ότι στην παρούσα εργασία δεν υπάρχει ένα ικανοποιητικό δείγμα δεδομένων για εξαγωγή χρήσιμης πληροφορίας. Παρόμοια αποτελέσματα παρατηρούνται και στην βιβλιογραφία [61], στην οποία αν και χρησιμοποιείται μεγαλύτερο δείγμα δεδομένων, συμπεριλαμβανομένων δεδομένων συναπτών ημερών, οι πιθανότητες μετάβασης δεν διαφέρουν σημαντικά μεταξύ των κλάσεων και άρα δεν εξάγεται αξιόπιστη πληροφορία ικανή για λήψη απόφασης.

7ο Κεφάλαιο

Συμπεράσματα - Μελλοντική Έρευνα

Ο Σακχαρώδης Διαβήτης (ΣΔ) είναι μια μεταβολική ασθένεια που σχετίζεται με τη δυσλειτουργία του μεταβολισμού της γλυκόζης και συγκεκριμένα με την ελλιπή έκκριση ινσουλίνης από το πάγκρεας. Έχει χρόνια πορεία και προκαλεί πληθώρα βραχυπρόθεσμων και μακροπρόθεσμων επιπλοκών, σοβαρών για την υγεία του ατόμου. Πρωταρχικό ρόλο στην αντιμετώπιση του ΣΔ, πόσο μάλλον για τον ΣΔ τύπου 1 (ινσουλοεξαρτώμενος), παίζει η εξωγενής χορήγηση ινσουλίνης και ο τακτικός γλυκαιμικός έλεγχος. Τα τελευταία χρόνια έχουν υιοθετηθεί καινοτόμες τεχνολογίες που επιτρέπουν την συνεχή παρακολούθηση των επιπέδων γλυκόζης και την εξωγενή χορήγηση ινσουλίνης, βελτιώνοντας κατά πολύ την ποιότητα ζωής του ασθενή. Ακόμη, η τεχνολογία έχει στραφεί προς την κατεύθυνση της αξιοποίησης των δεδομένων γλυκόζης και ινσουλίνης, με σκοπό την δημιουργία συστημάτων που θα χρησιμοποιούνται ως υποστηρικτικά εργαλεία για την λήψη αποφάσεων των θεραπόντων ιατρών αλλά και των ίδιων των ασθενών.

Στο πλαίσιο της παρούσας διπλωματικής εργασίας αναλύθηκαν δεδομένα από αισθητήρες γλυκόζης παιδιών με ΣΔΤ1, με στόχο την ανάπτυξη μοντέλου κατηγοριοποίησης οκτάωρων γλυκαιμικών προφίλ. Η μελέτη αποσκοπεί στην δημιουργία ενός συστήματος που θα λειτουργεί ως συμπληρωματικό εργαλείο υποστήριξης της ικανότητας αυτοδιαχείρισης ατόμου που πάσχει από ΣΔΤ1 και λήψης αποφάσεων σχετικά με την δόσεις ινσουλίνης, την διατροφή και την σωματική άσκηση. Στο κεφάλαιο αυτό παρουσιάζονται τα συμπεράσματα της μελέτης, οι περιορισμοί της, ενώ θέτονται οι μελλοντικοί στόχοι για επέκταση της έρευνας.

7.1 Συμπεράσματα

Στην εργασία αυτή παρουσιάστηκε μια μεθοδολογία που βασίστηκε στην ανάλυση δεδομένων γλυκόζης ως Δεδομένα Σύνθεσης για την κατηγοριοποίηση γλυκαιμικών προφίλ και την αξιοποίησή τους για την ανάπτυξη μοντέλου ταξινόμησης. Η ανάλυση πραγματοποιήθηκε ξεχωριστά για οκτάωρα που αφορούν τις περιόδους που ξεκινούν στις 00:00, 08:00 και 16:00, ώστε να αναλύονται κάθε φορά όμοια προφίλ ως προς τον κερκαδικό ρυθμό. Στη βιβλιογραφία [61] έχει διερευνηθεί η χρήση εξάωρων διαστημάτων, ωστόσο στην παρούσα μελέτη

επιλέχθηκαν οκτώωρα διαστήματα με σκοπό την ενίσχυση της πληροφορίας γλυκαιμικού προφίλ. Τα αποτελέσματα που προέκυψαν από την εφαρμογή αλγορίθμου συσταδοποίησης χρησιμοποιήθηκαν με σκοπό την δοκιμή μοντέλων ταξινόμησης με την μέθοδο LOO-CV. Από τα μοντέλα που δοκιμάστηκαν, το μοντέλο που χρησιμοποιούσε Μηχανές Διανυσμάτων Υποστήριξης έδειξε ότι είναι κατάλληλο για την ταξινόμηση οκτώωρων γλυκαιμικών προφίλ, καθώς για την πλειοψηφία των ασθενών πέτυχε ποσοστά επιτυχίας μεγαλύτερα του 90%. Στη συνέχεια, πραγματοποιήθηκε ανάπτυξη ενός απλού πιθανοτικού μοντέλου που υπολόγιζε τις πιθανότητες μετάβασης των κλάσεων γλυκόζης μεταξύ διαδοχικών οκτώωρων, ώστε να ενημερώνει τον ασθενή ποια είναι η πιθανότερη κατηγορία του οκτώωρου που ακολουθεί γνωρίζοντας την κατηγορία του προηγούμενου οκτώωρου. Τα αποτελέσματα που εξήχθησαν συνάδουν με εκείνα που παρουσιάζονται στην βιβλιογραφία [61] όπου εφαρμόζεται παρόμοια μεθοδολογία για την κατηγοριοποίηση προφίλ γλυκόζης ενήλικων με ΣΔΤ1.

Τα αποτελέσματα της προτεινόμενης μεθοδολογίας, λαμβάνοντας υπόψη το σύνολο δεδομένων, έδειξαν ότι η ανάλυση που ακολουθείται θα μπορούσε να λειτουργήσει ικανοποιητικά για κατηγοριοποίηση οποιουδήποτε οκτώωρου προφίλ γλυκόζης. Η πρόβλεψη της κατηγορίας του προφίλ γλυκόζης, που μπορεί να βρεθεί ο ασθενής σε μεταγενέστερη περίοδο μέσα στην μέρα, θα μπορούσε να βοηθήσει τους ασθενείς να λάβουν εκ των προτέρων διορθωτικά μέτρα σε ανεπιθύμητες καταστάσεις (υπο/υπερ-γλυκαιμικά επεισόδια) και να διαχειριστούν κατάλληλα την διατροφή τους και την χορήγηση ινσουλίνης.

Αν και τα αποτελέσματα είναι αρκετά ενθαρρυντικά, υπάρχουν αρκετοί περιορισμοί της μελέτης. Αρχικά, το σύνολο δεδομένων δεν αποτελεί αντιπροσωπευτικό δείγμα. Συγκεκριμένα, αποτελούνταν μόνο από έντεκα ασθενείς όπου για κάθε ασθενή, λόγω των πολλών ημερών με ελλιπή δεδομένα, αξιοποιήθηκε ένα περιορισμένο σύνολο δεδομένων γλυκόζης. Η επίδραση των περιορισμένων δεδομένων υποδηλώνεται και από την μεγάλη ανομοιογένεια των κλάσεων ως προς το πλήθος των δειγμάτων που περιέχουν. Επίσης, εξαιτίας της έλλειψης διαδοχικών ημερών αλλά και του μικρού αριθμού δεδομένων από διαφορετικές κλάσεις, δεν μπορεί να εξαχθεί ιδιαίτερα χρήσιμη πληροφορία από τους πίνακες με τις πιθανότητες μετάβασης μεταξύ των διαδοχικών οκτώωρων. Τέλος, στην προτεινόμενη μεθοδολογία δεν συμπεριλήφθηκε η ανάλυση δεδομένων ινσουλίνης διότι υπήρχαν αρκετές ελλείψεις σε μετρήσεις από την αντλία του συστήματος των ασθενών, με αποτέλεσμα να μην λαμβάνεται επαρκής πληροφορία για την εξαγωγή χρήσιμων αποτελεσμάτων.

7.2 Μελλοντική Έρευνα

Η μεθοδολογία που αναπτύχθηκε στην παρούσα διπλωματική εργασία επιδέχεται σημαντικές βελτιώσεις και επεκτάσεις. Για αρχή, είναι αρκετά σημαντικό να διευρυνθεί το σύνολο δεδομένων, δηλαδή να συλλεχθούν δεδομένα περισσότερων ασθενών με μεγαλύτερο όγκο δεδομένων γλυκόζης για τον καθένα. Με τον τρόπο

αυτόν, θα υπάρχει περισσότερη ποικιλία γλυκαιμικών προφίλ, που θα χαρακτηρίζονται από σημαντική μεταβλητότητα όσον αφορά τα κλινικά χαρακτηριστικά τους, ώστε να γίνει πιο εύκολη η διαδικασία αξιολόγησης των μοντέλων και η δυνατότητα γενίκευσης τους. Εξάλλου, για την εύρεση των πιθανοτήτων μετάβασης από κατηγορία σε κατηγορία μεταξύ διαδοχικών περιόδων, απαιτείται η ύπαρξη συναπτών ημερών.

Εξίσου απαραίτητη κρίνεται η ενσωμάτωση δεδομένων ινσουλίνης (basal και bolus δόσεις, CHO κ.α) στην ανάλυση. Δεδομένα ινσουλίνης μπορούν να χρησιμοποιηθούν για την εξαγωγή χαρακτηριστικών που θα μπορούσαν να ενισχύσουν την ικανότητα συσταδοποίησης και ταξινόμησης του συστήματος που παρουσιάστηκε στο πλαίσιο της παρούσας διπλωματικής. Η ενίσχυση της ικανότητας παραγωγής συστάδων με μεγαλύτερη ομοιομορφία ως προς τα κλινικά τους χαρακτηριστικά έχει τη δυνατότητα να βελτιώσει την απόδοση και του πιθανοτικού μοντέλου πρόβλεψης μετάβασης από οκτάωρο σε οκτάωρο.

Η μέθοδος που αναπτύχθηκε απαιτεί την ύπαρξη 288 μετρήσεων συνεχούς καταγραφής γλυκόζης μέσα στη μέρα. Στην παρούσα διπλωματική εργασία αξιοποιήθηκαν μέρες οι οποίες δεν περιείχαν διαστήματα κενά από μετρήσεις γλυκόζης άνω της μισής ώρας. Μελλοντική έρευνα θα μπορούσε να αξιολογήσει την ευαισθησία των μεθόδων που παρουσιάστηκαν σε μεγαλύτερα κενά στο ημερήσιο γλυκαιμικό προφίλ. Με χρήση κατάλληλου αλγορίθμου θα μπορούσαν να δημιουργηθούν συνθετικά κενά μέσα στην ημέρα και να συγκριθούν τα αποτελέσματα συσταδοποίησης και ταξινόμησης. Μελλοντική εφαρμογή της μεθόδου με δεδομένα συνεχούς μέτρησης γλυκόζης από αισθητήρες με συχνότητα δειγματοληψίας τα 15 λεπτά αντί τα 5 θα μπορούσε επίσης να βελτιώσει την ποιότητα των αποτελεσμάτων.

Ενδιαφέρουσα προσέγγιση θα αποτελούσε η συλλογή και η αξιοποίηση δεδομένων από αισθητήρες όπως βήματα, θερμίδες και καρδιακός παλμός κατά την διάρκεια του ύπνου ή γενικότερα πληροφορίες σχετικές με σαββατοκύριακα, αργίες, ημέρες με σωματική δραστηριότητα. Δεδομένου του ότι ο διαφορετικός τρόπος ζωής και οι συνήθειες επηρεάζουν την ποιότητα του γλυκαιμικού ελέγχου, είναι πιθανόν η εισαγωγή δεδομένων τέτοιας φύσεως να δημιουργήσει διαφορετικές κατηγορίες που να απέχουν πολύ από ασθενή σε ασθενή.

Συνοψίζοντας, η προτεινόμενη μεθοδολογία, με τις βελτιώσεις που απαιτούνται, θα μπορούσε να χρησιμοποιηθεί ως υποστηρικτικό εργαλείο για τους κλινικούς, βοηθώντας τους να προσαρμόσουν εξατομικευμένα το προφίλ δοσολογίας ινσουλίνης ενός ασθενούς, σύμφωνα με την τελική κατηγοριοποίηση ενός οκταώρου προφίλ γλυκόζης.

Βιβλιογραφία

- [1] World Health Organization, “Diabetes”, 2021. [Online]. Available: https://www.who.int/health-topics/diabetes#tab=tab_1
- [2] Κωνσταντία Ζαρκογιάννη, “Ευφυή Συστήματα Υποστήριξης Εξατομικευμένων Ιατρικών Αποφάσεων για τη Διαχείριση του Σακχαρώδους Διαβήτη”, Διδακτορική Διατριβή, Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, ΕΜΠ, 2011.
- [3] Wikipedia -The Free Encyclopedia, “Glucose” [Online]. Available: <https://en.wikipedia.org/wiki/Glucose>
- [4] Αθανασίου Μαρία, “Ανάπτυξη Ελεγκτή Γλυκόζης για Άτομα με Σακχαρώδη Διαβήτη Τύπου 1: Χρήση Αναλογικού Διαφορικού Ελέγχου με Εξασθένηση Μνήμης”, Διπλωματική Εργασία, Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, ΕΜΠ, 2015.
- [5] S. L. Aronoff, K. Berkowitz, B. Shreiner, and L. Want, “Glucose Metabolism and Regulation: Beyond Insulin and Glucagon,” *Diabetes Spectr.*, vol. 17, no. 3, pp. 183–190, Jul. 2004.
- [6] P. J. Hantzidiamantis and S. L. Lappin, “*Physiology, Glucose*”. StatPearls Publishing, 2019.
- [7] Wikipedia - The Free Encyclopedia, “Pancreas”, [Online]. Available: <https://en.wikipedia.org/wiki/Pancreas>
- [8] “Ενδοκρινής Μοίρα του Παγκρέατος. 21/5/18 Ε. Παρασκευά, Εργ. Φυσιολογίας, Τμήμα Ιατρικής Π.Θ.” [Online]. Available: <http://docplayer.gr/83319699-Endokrinis-moira-toy-pagkreatos-21-5-18-e-paraskeya-erg-fysiologias-tuiua-iatrikis-p-th.html>
- [9] Θ. Βαλαβανίδης, Κ. Ευσταθίου, “Η Χημική Ένωση του Μήνα Ινσουλίνη”, Ιούνιος 2009 .[Online]. Available: http://www.chem.uoa.gr/chemicals/chem_insulin.htm
- [10] M. C. Petersen and G. I. Shulman, “Mechanisms of insulin action and insulin resistance,” *Physiological Reviews*, vol. 98, no. 4. American Physiological Society, pp. 2133–2223, 01-Oct-2018.
- [11] L. Briant, A. Salehi, E. Vergari, Q. Zhang, and P. Rorsman, “Glucagon secretion from pancreatic α -cells,” *Upsala Journal of Medical Sciences*, vol. 121, no. 2. Taylor and Francis Ltd, pp. 113–119, 02-Apr-2016.
- [12] A. D. Cherrington, “Glucagon Physiology,” *Can. J. Diabetes*, vol. 34, no. 3, pp. 187–188, Jan. 2010.

- [13] A. T. Kharroubi, "Diabetes mellitus: The epidemic of the century," *World J. Diabetes*, vol. 6, no. 6, p. 850, 2015.
- [14] American Diabetes Association, "Diagnosis and classification of diabetes mellitus," *Diabetes Care*, vol. 37, no. SUPPL.1, pp. S81–S90, Jan. 2014.
- [15] J. S. Skyler et al., "Differentiation of diabetes by pathophysiology, natural history, and prognosis," *Diabetes*, vol. 66, no. 2. American Diabetes Association Inc., pp. 241–255, 01-Feb-2017.
- [16] Giuseppe Fico, Liss Hernandez, Jorge Cancela, Arianna Dagliati, Lucia Sacchi, Antonio Martinez-Millana, Jorge Posada, Lidia Manero, Jose Verdú, Andrea Facchinetti, Manuel Ottaviano, Konstantia Zarkogianni, Konstantina Nikita, Leif Groop, Rafael Gabriel-Sanchez, Luca Chiovato, Vicente Traver, Juan Francisco Merino-Torres, Claudio Cobelli, Riccardo Bellazzi, Maria Teresa Arredondo, "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 medical informatics and decision making*, vol. 19, pp. 1-16, 2019
- [17] M. Massi-Benedetti and M. Orsini-Federici, "Treatment of type 2 diabetes with combined therapy: what are the pros and cons?," *Diabetes care*, vol. 31 Suppl 2, no. Supplement 2. American Diabetes Association, pp. S131–S135, 01-Feb-2008.
- [18] "How to Treat Gestational Diabetes | ADA." [Online]. Available: <https://www.diabetes.org/diabetes/gestational-diabetes/how-to-treat-gestational-diabetes>
- [19] J. Verdú, F. Sambo, B. Di Camillo, C. Cobelli, A. Facchinetti, G. Fico, R. Bellazzi, L. Sacchi, A. Dagliati, D. Segnani, V. Tibollo, M. Ottaviano, R. Gabriel, L. Groop, J. Postma, A. Martinez, L. Hakaste, T. Tuomi, and K. Zarkogianni, "Predictive, preventive and personalized medicine in diabetes onset and complication (MOSAICproject)", *The EPMA Journal*, vol. 7, Suppl. 1, pp 42-43, 2016
- [20] "International Diabetes Federation - Complications." [Online]. Available: <https://www.idf.org/aboutdiabetes/complications.html>
- [21] K. Zarkogianni, M. Athanasiou, A. C. Thanopoulou, and K. S. Nikita, "Comparison of machine learning approaches towards assessing the risk of developing Cardiovascular disease as a long-term diabetes complication", to appear in *IEEE Journal of Biomedical and Health Informatics*, 2017.
- [22] 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 Systems*, 2017.
- [23] International Diabetes Federation, *IDF Diabetes Atlas*, 9th edition, 2019. [Online]. Available: <https://www.diabetesatlas.org/en/>
- [24] Κωνσταντίνα Νικήτα, "Προσομοίωση Φυσιολογικών Συστημάτων", κεφ. 9, Τζιόλας, 2010

- [25] K. Zarkogianni, E. Litsa, K. Mitsis, P. Wu, C.D. Kaddi, C. Cheng, M. D. Wang, K. S. Nikita, "A Review of Emerging Technologies for the Management of Diabetes Mellitus", *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 12, pp.2735-2749, 2015.
- [26] Wikipedia - The Free Encyclopedia, "Glucose Meter" [Online]. Available: https://en.wikipedia.org/wiki/Glucose_meter
- [27] Wikipedia -The Free Encyclopedia, "Blood Glucose Monitoring" [Online]. Available: https://en.wikipedia.org/wiki/Blood_glucose_monitoring
- [28] "Alternate Site Testing (AST) - What is Alternate Site Testing, Accuracy & How To AST." [Online]. Available: <https://www.diabetes.co.uk/blood-glucose/alternate-site-testing.html>
- [29] V. D. Funtanilla, T. Caliendo, and O. Hilas, "Continuous glucose monitoring: A review of available systems," *P T*, vol. 44, no. 9, pp. 550–553, Sep. 2019.
- [30] D. Olczuk and R. Priefer, "A history of continuous glucose monitors (CGMs) in self-monitoring of diabetes mellitus," *Diabetes and Metabolic Syndrome: Clinical Research and Reviews*, vol. 12, no. 2. Elsevier Ltd, pp. 181–187, 01-Apr-2018.
- [31] G. Cappon, M. Vettoretti, G. Sparacino, and A. Facchinetti, "Continuous glucose monitoring sensors for diabetes management: A review of technologies and applications," *Diabetes Metab. J.*, vol. 43, no. 4, pp. 383–397, Aug. 2019.
- [32] E. Cengiz and W. V. Tamborlane, "A tale of two compartments: Interstitial versus blood glucose monitoring," *Diabetes Technol. Ther.*, vol. 11, no. SUPPL.1, p. S-11, 2009.
- [33] E. H. Yoo and S. Y. Lee, "Glucose biosensors: An overview of use in clinical practice," *Sensors*, vol. 10, no. 5. Multidisciplinary Digital Publishing Institute (MDPI), pp. 4558–4576, May-2010.
- [34] T. S. Bailey, "Clinical implications of accuracy measurements of continuous glucose sensors," *Diabetes Technology and Therapeutics*, vol. 19, no. S2. Mary Ann Liebert Inc., pp. S51–S54, 01-May-2017.
- [35] N. S. Oliver, C. Toumazou, A. E. G. Cass, and D. G. Johnston, "Glucose sensors: A review of current and emerging technology", *Diabetic Medicine*, vol. 26, no. 3. pp. 197–210, Mar-2009.
- [36] [Online]. Available: <https://www.fda.gov/news-events/press-announcements/fda-approves-first-continuous-glucose-monitoring-system-fully-implantable-glucose-sensor-and>
- [37] [Online]. Available: <https://www.diabetes.org/healthy-living/medication-treatments/insulin-other-injectables/insulin-basics>
- [38] "Insulin administration," *Diabetes Care*, vol. 26, no. SUPPL. 1. American Diabetes Association Inc., pp. s121–s124, 01-Jan-2003.
- [39] "Insulin pens: Types, benefits, and how to use them." [Online]. Available: <https://www.medicalnewstoday.com/articles/316607#how-to-use>. [Accessed: 31-Aug-2020].

- [40] N. D. Sora, F. Shashpal, E. A. Bond, and A. J. Jenkins, “Insulin Pumps: Review of Technological Advancement in Diabetes Management,” *American Journal of the Medical Sciences*, vol. 358, no. 5. Elsevier B.V., pp. 326–331, 01-Nov-2019.
- [41] K. Zarkogianni, K. Mitsis, E. Litsa, MT Arredondo, G. Fico, A. Fioravanti, K. S. Nikita, “Comparative assessment of glucose prediction models for Patients with Type 1 Diabetes Mellitus applying sensors for glucose and physical activity monitoring”, *_Medical & Biological Engineering & Computing_*, vol. 53, no. 12, pp. 1333-1343, 2015
- [42] R. Singla, A. Singla, Y. Gupta, and S. Kalra, “Artificial intelligence/machine learning in diabetes care,” *Indian J. Endocrinol. Metab.*, vol. 23, no. 4, pp. 495–497, Jul. 2019.
- [43] K. Zarkogianni, A. Vazeou, S.G. Mougiakakou, A. Prountzou, K.S. Nikita, "An insulin infusion advisory system based on autotuning nonlinear model-predictive control", *_IEEE Transactions on Biomedical Engineering_*, vol. 58, no. 9, pp. 2467-77, 2011.
- [44] S.G. Mougiakakou, C. Bartsocas, E. Bozas, N. Chaniotakis, D. Iliopoulou, I. Kouris, S. Pavlopoulos, A. Prountzou, M. Skevofylakas, A. Tsoukalis, K. Varotsis, A. Vazeou, K. Zarkogianni and K. S. Nikita, "SMARTDIAB: A Communication and Information Technology Approach for the Intelligent Monitoring, Management and Follow-up of Type 1 Diabetes Patients", *_IEEE Transactions on Information Technology in Biomedicine_*, *_Special Issue: New and Emerging Trends in Bioinformatics and Bioengineering_*, vo. 14, no. 3, pp. 622 – 633, 2010.
- [45] Wikipedia - The Free Encyclopedia, “Compositional Data” [Online]. Available: https://en.wikipedia.org/wiki/Compositional_data
- [46] [Online]. Available: <http://www.compositionaldata.com/>
- [47] V. Pawlowsky-Glahn, J. J. Egozcue, and R. Tolosana-Delgado, “Lecture Notes on Compositional Data Analysis,” 2007.
- [48] G. Mateu-Figueras, “Statistical analysis of compositional data” 2014.
- [49] John B. Fraleigh, Μετάφραση: Απόστολος Γιαννόπουλος, “Εισαγωγή στην Άλγεβρα”, σελ. 147-150
- [50] J. J. Egozcue, V. Pawlowsky-Glahn, G. Mateu-Figueras, and C. Barceló-Vidal, “Isometric Logratio Transformations for Compositional Data Analysis,” *Math. Geol.*, vol. 35, no. 3, pp. 279–300, Apr. 2003.
- [51] Ι. Βλαχάβας, Π. Κεφαλάς, Ν. Βασιλειάδης, Φ. Κόκκορας, Η. Σακελλαρίου. “Τεχνητή Νοημοσύνη - Γ' Έκδοση”, σελ 336-365.
- [52] “Unsupervised Machine Learning: What is, Algorithms, Example.” [Online]. Available: <https://www.guru99.com/unsupervised-machine-learning.html>
- [53] Δημήτριος Πετρίδης, “Ανάλυση Πολυμεταβλητών Τεχνικών, Εφαρμογές Περιπτώσεων” Κεφάλαιο 5: Ανάλυση Συστάδων, σελ.158-165.
- [54] “Machine Learning — Singular Value Decomposition (SVD) & Principal Component Analysis (PCA) | by Jonathan Hui | Medium.” [Online]. Available: <https://jonathan-hui.medium.com/machine-learning-singular-value-decomposition-svd-principal-component-analysis-pca-1d45e885e491>

- [55] Wikipedia - The Free Encyclopedia, “Cluster Analysis” [Online]. Available: https://en.wikipedia.org/wiki/Cluster_analysis
- [56] “Beginner’s Guide To K-Means Clustering - Analytics India Magazine.” [Online]. Available: <https://analyticsindiamag.com/beginners-guide-to-k-means-clustering/>
- [57] [Online]. Available: <https://medium.com/@jorgesleoneel/classification-methods-in-machine-learning-58ce63173db8>
- [58] S. Zulj, P. Carvalho, R. Ribeiro, and R. Magjarevic, “Handling Missing Data in CGM Records,” in *IFMBE Proceedings*, 2020, vol. 74, pp. 420–427.
- [59] L. Biagi, A. Bertachi, J. A. Martín-Fernández, and J. Vehí, “Compositional data analysis of type 1 diabetes data,” *CEUR Workshop Proc.*, vol. 2148, pp. 8–12, 2018.
- [60] L. Biagi *et al.*, “Individual categorisation of glucose profiles using compositional data analysis,” *Stat. Methods Med. Res.*, vol. 28, no. 12, pp. 3550–3567, Dec. 2019.
- [61] L. Biagi, A. Bertachi, J. A. Martín-Fernández, and J. Vehí, “Compositional data analysis of glucose profiles of type 1 diabetes patients,” in *IFAC-PapersOnLine*, 2019, vol. 52, no. 1, pp. 1006–1011.
- [62] J. A. Martín-Fernández, C. Barceló-Vidal, and V. Pawlowsky-Glahn, “Dealing with Zeros and Missing Values in Compositional Data Sets Using Nonparametric Imputation,” *Math. Geol.*, vol. 35, no. 3, pp. 253–278, Apr. 2003.
- [63] P. M. Kroonenberg, “Appendix B: Biplots and Their Interpretation,” in *Applied Multiway Data Analysis*, Hoboken, NJ, USA: John Wiley & Sons, Inc., 2008, pp. 491–500.