

Εργαστήριο Ανάλυσης Δεδομένων από Αισθητήρες

Εισαγωγή στη Βιοϊατρική Μηχανική

Πανάγου Ζαχαρίας 03119114

1) Σε αυτό το εργαστήριο ανάλυσης δεδομένων καλούμαστε να διαβάσουμε τις μετρήσεις σακχάρου ατόμων με σακχαρώδη διαβήτη τύπου 1 όπως αυτές μετρήθηκαν από Αισθητήρα μέτρησης περιεκτικότητας του σακχάρου στο αίμα καθώς και τις μετρήσεις από την εφαρμογή υπολογιστικών πειραμάτων στον προσομοιωτή UVa T1DM. Διαβάζουμε τα δεδομένα με Python και τα τοποθετούμε σε Pandas DataFrame. Έπειτα υπολογίζουμε την μέση απόλυτη σχετική διαφορά και τον συντελεστή συσχέτισης μεταξύ των τιμών του αισθητήρα και του προσομοιωτή σε τρεις διαφορετικές περιοχές τιμών, στην περιοχή υπογλυκαιμίας (<70 mg/dl), ευγλυκαιμίας και υπεργλυκαιμίας (>180 mg/dl).

For Blood Glucose values between 60.609 mg/dl and 69.992 mg/dl :

Mean absolute relative distance is : 6.978%

Correlation coefficient between the Glucose and the sensor readings is : 0.6665762433927276

For Blood Glucose values between 70.035 mg/dl and 179.962 mg/dl :

Mean absolute relative distance is : 9.48%

Correlation coefficient between the Glucose and the sensor readings is : 0.9066791694973899

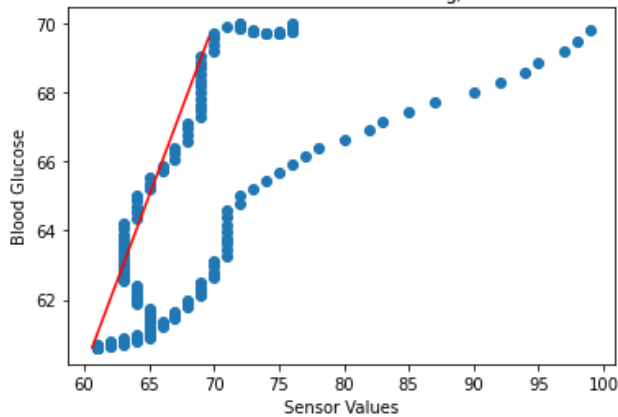
For Blood Glucose values between 180.009 mg/dl and 387.232 mg/dl :

Mean absolute relative distance is : 6.53%

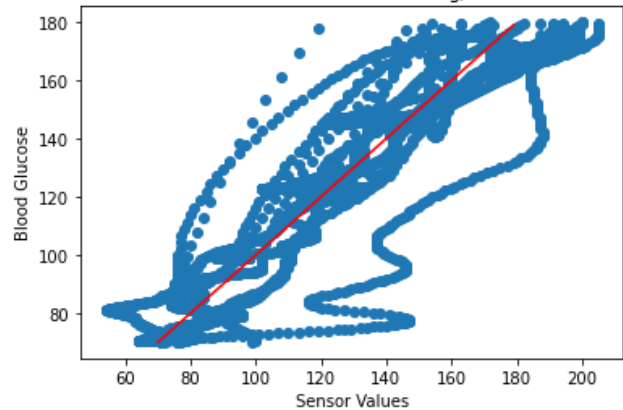
Correlation coefficient between the Glucose and the sensor readings is : 0.8993815576281389

Εδώ βλέπουμε ότι στην περιοχή ευγλυκαιμίας και υπεργλυκαιμίας υπάρχει μια πολύ καλή συσχέτιση μεταξύ της τιμής του αισθητήρα και του προσομοιωτή (περίπου 90%). Σε αυτές τις περιοχές ένα μεγάλος μέρος των μετρήσεων του αισθητήρα προσεγγίζει τις τιμές του UVa T1DM. Αντιθέτως, στη περιοχή υπογλυκαιμίας η συσχέτιση είναι αρκετά ασθενής (66%) πράγμα το οποίο δείχνει ότι ο αισθητήρας δεν ανιχνεύει σε ικανοποιητικό βαθμό τα επίπεδα σακχάρου. Όσων αφορά τη μέση απόλυτη σχετική διαφορά εδώ θα πρέπει να ληφθούν υπ' όψη και το εύρος τιμών της κάθε περιοχής καθώς και το μέγεθος της τιμής (απόκλιση 10mg/dl σε επίπεδα υπεργλυκαιμίας επιφέρει μικρότερη απόλυτη σχετική διαφορά σε σχέση με την ίδια απόκλιση σε επίπεδα υπογλυκαιμίας). Τέλος μπορούμε να απεικονίσουμε τα δεδομένα για να κατανοήσουμε καλύτερα την ακρίβεια των μετρήσεων του αισθητήρα.

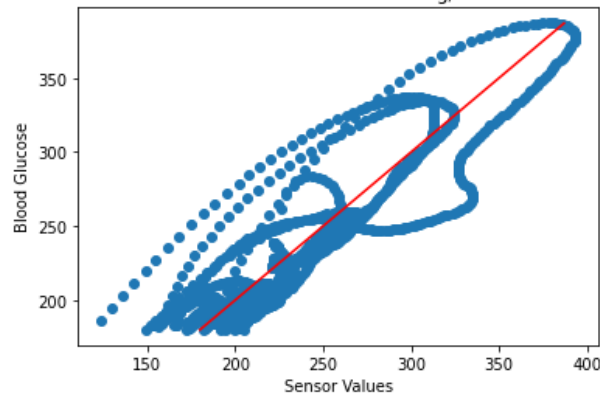
For Blood Glucose values between 60.609 mg/dl and 69.992 mg/dl :



For Blood Glucose values between 70.035 mg/dl and 179.962 mg/dl :



For Blood Glucose values between 180.009 mg/dl and 387.232 mg/dl :

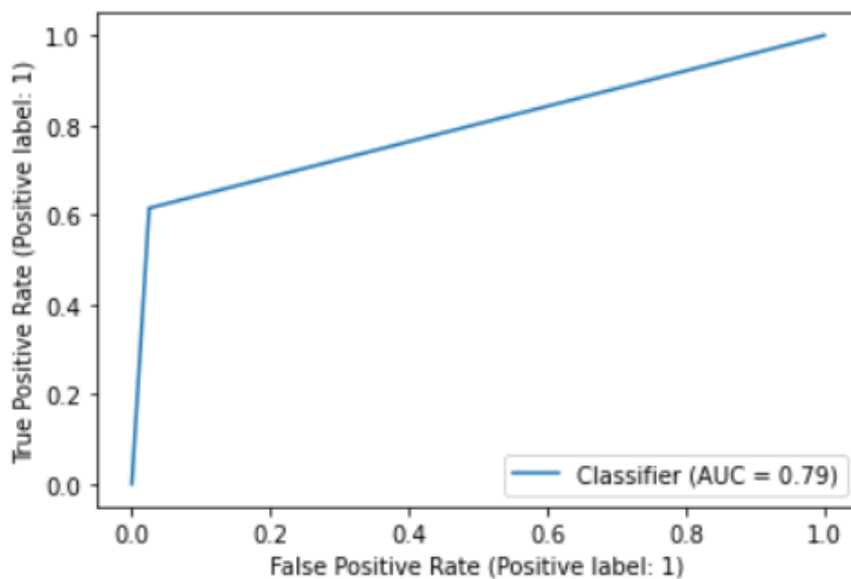


Με κόκκινο εμφανίζεται η ευθεία συσχέτισης για $r=1$.

2) Για να μπορέσουμε να κάνουμε ROC ανάλυση θα χρειαστούμε να αλλάξουμε τα δεδομένα μας. Για αυτό προσθέτουμε τέσσερις νέες στήλες στο DataFrame μας, μια που ονομάζεται **Hypoglycemia** και περιέχει ως τιμές είτε το μηδέν (0), όταν η τιμή του σακχάρου (όπως υπολογίστηκε από τον προσομοιωτή) είναι μεγαλύτερη από 70 mg/dl, είτε το ένα (1) όταν η τιμή του σακχάρου είναι κάτω από 70 mg/dl, και μια που ονομάζεται **Predicted Hypoglycemia** η οποία περιέχει πάλι μηδέν και ένα όπως η προηγούμενη σύμφωνα όμως τώρα με τις τιμές του αισθητήρα. Τέλος οι άλλες δύο ονομάζονται **Hyperglycemia** και **Predicted Hyperglycemia** και συμπληρώνονται αναλόγως, με τις τιμές ένα (1) να βρίσκονται τώρα σε επίπεδα σακχάρου >180 mg/dl.

Έπειτα εφαρμόζουμε ROC ανάλυση για υπογλυκαιμικά επεισόδια (<70 mg/dl) και εμφανίσουμε τους χαρακτηριστικούς δείκτες του Confusion Matrix.

True positive rate for each threshold : [0. 0.61490683 1.]]



Confusion Matrix :

```
[[4758 122]
 [ 62  99]]
```

Sensitivity : 0.6149068322981367

Specificity : 0.975

Positive Predictive Values : 0.4479638009049774

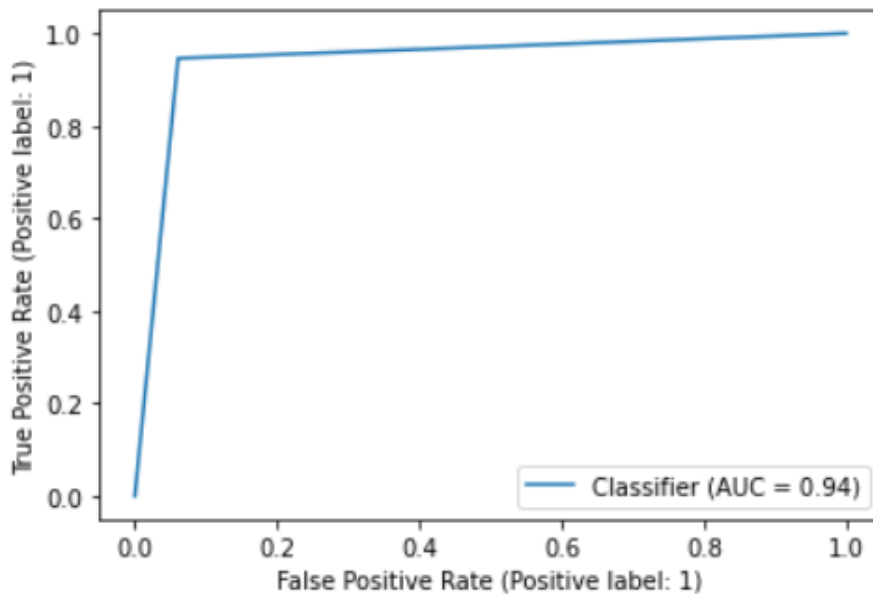
Negative Predictive Values : 0.987136929460581

Accuracy : 0.9634993056933148

Αρχικά βλέπουμε ότι έχει πάρα πολύ καλή ακρίβεια με 96.3%. Παρ' όλα αυτά το sensitivity με τιμή 61.5% μας δείχνει ότι ο αισθητήρας δυσκολεύεται να ανιχνεύσει τα Positives δηλαδή τις περιπτώσεις όπου υπάρχει υπογλυκαιμία. Αντιθέτως το specificity έχει αρκετά υψηλή τιμή με 97.5%, δηλαδή ο αισθητήρας ανίχνευσε σωστά το 97.5% των περιπτώσεων που δεν υπήρχε υπογλυκαιμικό επεισόδιο. Ενώ από όλες τις ανιχνεύσεις που έκανε ότι υπάρχει υπογλυκαιμικό επεισόδιο μόνο το 44.8% ήταν σωστές (PPV= 0.44). Εδώ να σημειωθεί ότι η μεγάλη ακρίβεια είναι παραπλανητική και μπορεί εύκολα να μας οδηγήσει σε λάθος συμπεράσματα. Αυτό οφείλεται στο γεγονός ότι τα υπογλυκαιμικά επεισόδια είναι σπάνια και καταλαμβάνουν ένα πολύ μικρό ποσοστό των δεδομένων μας, μόλις 161 από τις 5041 μετρήσεις. Αν είχαμε για παράδειγμα έναν αισθητήρα που μέτραγε πάντα τιμή σακχάρου π.χ. 90 mg/dl τότε η στήλη **Predicted Hypoglycemia** των δεδομένων μας θα είχε παντού μηδενικά και η τιμή accuracy θα έβγαине 96.8% αλλά δεν θα είχε ανιχνευθεί κανένα υπογλυκαιμικό επεισόδιο. Για αυτό χρειαζόμαστε και τους υπόλοιπους δείκτες αξιολόγησης.

3) Ομοίως για τις στήλες **Hyperglycemia** και **Predicted Hyperglycemia**, δηλαδή για τα υπεργλυκαιμικά επεισόδια.

True positive rate for each threshold : [0. 0.94621639 1.]



Confusion Matrix :

```
[[3233 209]
```

```
[ 86 1513]]
```

Sensitivity : 0.9462163852407754

Specificity : 0.9392794886693783

Positive Predictive Values : 0.8786295005807201

Negative Predictive Values : 0.9740885808978608

Accuracy : 0.9414798651061297

Εδώ τα αποτελέσματα είναι πολύ καλύτερα κάτι το οποίο φαίνεται και από την μορφή της ROC curve. Παρόλο που η ακρίβεια είναι τώρα στο 94.1% ο αισθητήρας ανίχνευσε σωστά το 94.6% των υπεργλυκαιμικών επεισοδίων. Από όλες τις ανιχνεύσεις που έκανε ότι υπάρχει περιεκτικότητα σακχάρου στο αίμα μεγαλύτερη από 180 mg/dl το 87.9% ήταν σωστές (PPV) ενώ για τιμές μικρότερες του 180 mg/dl το 97.4% των ανιχνεύσεων ήταν σωστές (NPV).

4) Σύμφωνα με τα δεδομένα και την παραπάνω ανάλυση μπορούμε να βγάλουμε συμπεράσματα σχετικά με την ακρίβεια του αισθητήρα. Αρχικά είναι εμφανές ότι ανιχνεύει με πολύ μεγαλύτερη ακρίβεια τις υπεργλυκαιμίες από τις υπογλυκαιμίες. Αυτό οφείλεται στο γεγονός ότι η συνεχής καταγραφή γλυκόζης έχει μεγαλύτερη ακρίβεια σε τιμές που αντιστοιχούν σε υπεργλυκαιμία αλλά και σε ευγλυκαιμία σε σχέση με την ακρίβεια καταγραφής της υπογλυκαιμίας. Συνεπώς είναι φυσικό αυτό να επιφέρει και χαμηλότερη ανίχνευση της υπογλυκαιμίας όπως και διαπιστώθηκε.