



Εθνικό Μετσόβιο Πολυτεχνείο

*Σχολή Ηλεκτρολόγων Μηχανικών Μηχανικών
Υπολογιστών*

Εισαγωγή στη Βιοϊατρική Μηχανική

Συγγραφέας:
Σκουρτσίδης Γεώργιος

ΑΜ:
03114307

Γραπτή / Προγραμματιστική Εργασία

Περιγραφή Άσκησης

Σκοπός

Σκοπός της Άσκησης είναι η εξοικείωση του φοιτητή με την εφαρμογή κριτηρίων αξιολόγησης της διακριτικής ικανότητας αισθητήρων καταγραφής φυσιολογικών παραμέτρων.

Περιγραφή

Στο πλαίσιο αυτής της άσκησης, σας δίνονται δεδομένα χρονοσειρών γλυκόζης ατόμων με σακχαρώδη διαβήτη τύπου 1, τα οποία προέκυψαν από την εφαρμογή υπολογιστικών πειραμάτων ανοικτού βρόχου στον προσομοιωτή UVa T1DM. Ο προσομοιωτής UVa T1DM είναι ένα υπολογιστικό περιβάλλον που διαθέτει εικονικούς ασθενείς (παιδιά, έφηβους, ενήλικες) με σακχαρώδη διαβήτη τύπου 1 και έχει λάβει έγκριση από τον Αμερικανικό Οργανισμό Φαρμάκων (FDA) για να χρησιμοποιείται ως υποκατάστατο προκλινικών δοκιμών για την αξιολόγηση ελεγκτών γλυκόζης.

Δεδομένα

Τα δεδομένα που σας δίνονται αντιστοιχούν σε περίπου 3.5 ημέρες και αποτελούνται από τρεις στήλες: (i) στη στήλη A παρουσιάζεται ο χρόνος σε λεπτά, (ii) στη στήλη B, οι τιμές γλυκόζης αίματος σε mg/dl, και (iii) στη στήλη C, οι καταγραφές γλυκόζης από αισθητήρα συνεχούς καταγραφής γλυκόζης σε mg/dl. Συνήθως οι αισθητήρες συνεχούς καταγραφής γλυκόζης παρουσιάζουν ανακρίβειες και αποκλίσεις από την γλυκόζη αναφοράς (γλυκόζη αίματος). Σε αυτήν την άσκηση, καλείστε να αξιολογήσετε την ακρίβεια των αισθητήρων συνεχούς καταγραφής γλυκόζης, εφαρμόζοντας κατάλληλα κριτήρια αξιολόγησης, όπως το Root Mean Square Error και η ROC ανάλυση.

Ερωτήσεις

Προεργασία

Ο κώδικας της παρούσας άσκησης γράφτηκε σε Python Notebook (.ipynb). Για την ευκολότερη χρήση των δεδομένων πραγματοποιήθηκε μετατροπή του αρχείου "Data.xlsx" σε μορφή csv προκειμένου να μετατροπεί στη συνέχεια σε Pandas DataFrame χωρίς προβλήματα. Το παραδοτέο notebook έχει αποθηκευμένα τα αποτελέσματα όλων των κελιών καθώς και καθοδηγητικά σχόλια. Ο ίδιος κώδικας παραδίδεται και σε ".py" αρχείο.

Ερώτηση 1

Εκφώνηση

Να εφαρμόσετε την τετραγωνική ρίζα του μέσου τετραγωνικού σφάλματος (Root Mean Square Error (RMSE)) για να αξιολογήσετε κατά πόσο οι καταγραφές γλυκόζης από τον αισθητήρα προσεγγίζουν τις τιμές γλυκόζης αίματος. Η τετραγωνική ρίζα του μέσου τετραγωνικού σφάλματος (RMSE) δίνεται από την σχέση:

$$\sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{G_{b_i} - G_{s_i}}{N} \right)^2}$$

όπου G_{b_i} και G_{s_i} συμβολίζουν την τιμή γλυκόζης αίματος και την καταγραφή γλυκόζης από τον αισθητήρα την χρονική στιγμή i , αντίστοιχα, ενώ με N συμβολίζεται ο αριθμός των μετρήσεων/καταγραφών γλυκόζης

Απάντηση

Εφαρμόσαμε τον αναφερόμενο τύπο. Τα αποτελέσματα είναι :
 $RMSE = 20.329$

Το RMSE υποδεικνύει την προσαρμογή του μοντέλου στα δεδομένα, δηλαδή πόσο κοντά βρίσκονται τα παρατηρούμενα σημεία στις προβλεπόμενες τιμές του μοντέλου. Το RMSE μπορεί να ερμηνευθεί ως η τυπική απόκλιση της ανεξήγητης διακύμανσης και έχει την χρήσιμη ιδιότητα να βρίσκεται στις ίδιες μονάδες με τα δεδομένα μας. Οι χαμηλότερες τιμές του RMSE υποδεικνύουν καλύτερη προσαρμογή. Το RMSE είναι ένα μέτρο για το πόσο ακριβώς το μοντέλο προβλέπει τα πραγματικά δεδομένα.

Ερώτηση 2

Εκφώνηση

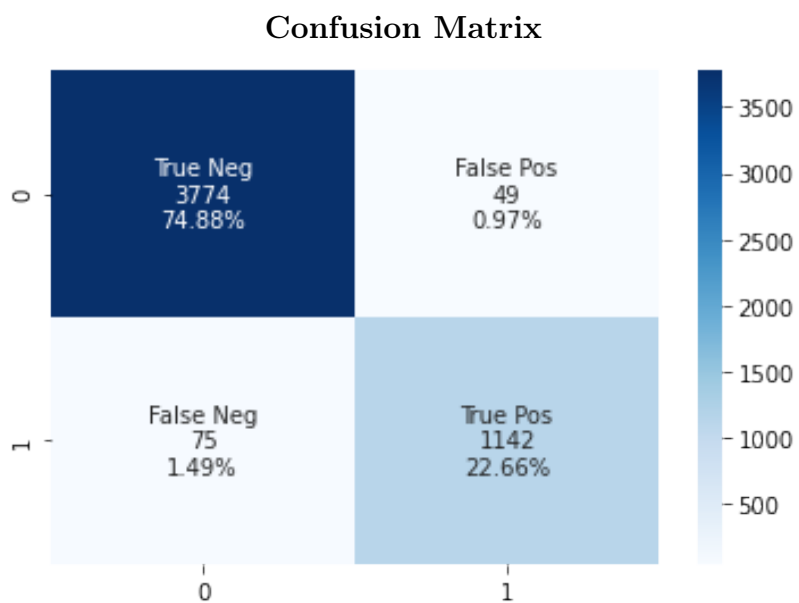
Θεωρώντας ως υπογλυκαιμικά επεισόδια τις τιμές γλυκόζης που είναι μικρότερες από 70 mg/dl, να εφαρμόσετε τη ROC ανάλυση για να αξιολογήσετε την ικανότητα του αισθητήρα συνεχούς καταγραφής γλυκόζης να ανιχνεύει τις υπογλυκαιμίες. Εφαρμόζοντας την τιμή 70 mg/dl ως κατώφλι γλυκόζης, οι χρονοσειρές γλυκόζης αντιμετωπίζονται ως δυαδικές και με αυτό τον τρόπο καθίσταται εφικτή η εφαρμογή της ROC ανάλυσης. Να υπολογίσετε τους δείκτες: Ευαισθησία, Ειδικότητα, Θετική Διαγνωστική Αξία, Αρνητική Διαγνωστική Αξία, Ακρίβεια.

Απάντηση

Σε αυτό το ερώτημα μας ζητείται να ορίσουμε ένα κατώφλι και να αξιολογήσουμε την αποδοτικότητα του ταξινομητή, σχετικά με το αν κατηγοριοποιεί σωστά τα υπογλυκαιμικά επεισόδια.

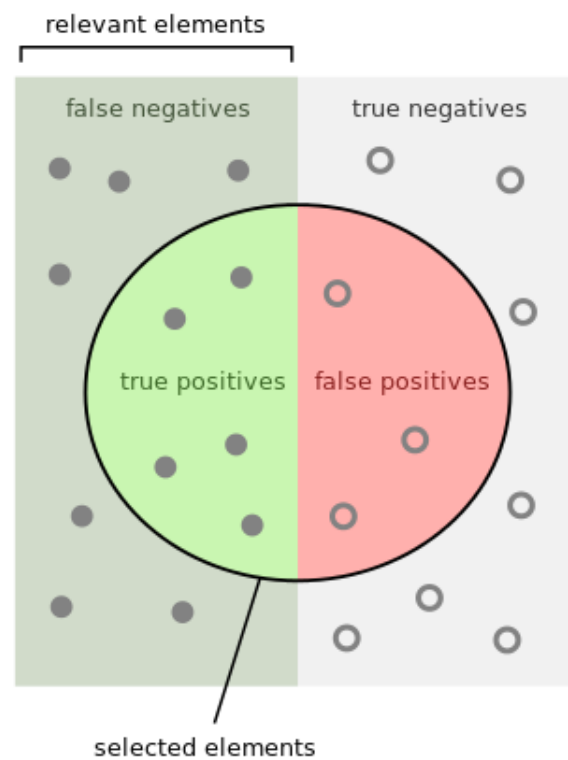
Ορίζουμε πως οποιαδήποτε τιμή κάτω από $70 \frac{mg}{dl}$ θα λαμβάνει την τιμή 1 (υπογλυκαιμικό επεισόδιο), ενώ το αντίθετο λαμβάνει την τιμή 0.

Στη συνέχεια υπολογίζουμε τον confusion matrix, τους δείκτες αξιολόγησης καθώς και σχεδιάζουμε ROC και Precision-Recall γραφήματα. Τα παραπάνω θα μας βοηθήσουν στην καλύτερη ερμηνεία και αξιολόγηση της αποδοτικότητας του ταξινομητή.



Παρατηρούμε πως στον πίνακα σύγχυσης έχουμε αρκετές πληροφορίες σχετικά με την αποτελεσματικότητα του ταξινομητή. Στη συνέχεια υπολογίζουμε τους δείκτες, περιγράφοντας εν συντομία τη σημασία τους.

Γραφική εξήγηση κάποιων δεικτών



How many selected items are relevant?

$$\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$$

How many relevant items are selected?

$$\text{Recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$

Ευαισθησία (True Positive Rate)

Αριθμός δειγμάτων κατηγοριοποιημένος ως Positive / πραγματικός αριθμός δειγμάτων που είναι Positive Ποσοστό των actual θετικών που αναγνωρίστηκαν σωστά (πόσα σχετικά επιλέγονται). Περιγράφει πόσο καλό είναι το μοντέλο στην πρόβλεψη της θετικής κλάσης όταν το πραγματικό αποτέλεσμα είναι θετικό.

$$TPR = \frac{TP}{P} = \frac{TP}{(TP+FN)}$$

Ειδικότητα (True Negative Rate)

Αριθμός δειγμάτων κατηγοριοποιημένος ως Negative / πραγματικός αριθμός δειγμάτων που είναι Negative Αλλιώς λέγεται Ειδικότητα - Specificity.

$$SPC = \frac{TN}{N} = \frac{TN}{(TN+FP)}$$

Θετική Διαγνωστική Αξία

(Positive Predictive Value). Ποσοστό των θετικών που είναι σωστά (πόσα επιλεγμένα είναι σχετικά) Λέγεται επίσης Precision.

$$PPV = \frac{TP}{(TP+FP)}$$

Αρνητική Διαγνωστική Αξία

Λέγεται και Negative Predictive Value

$$NPV = \frac{TN}{(TN+FN)}$$

Ακρίβεια (Accuracy)

Αριθμός σωστών προβλέψεων / συνολικός αριθμός δειγμάτων

$$ACC = \frac{TP+TN}{(TN+TN+FP+FN)}$$

Οι μετρικές για τις υπογλυκαιμίες:

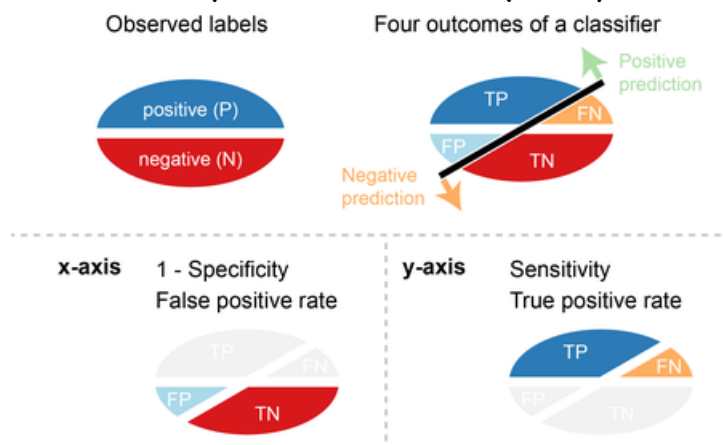
Ακρίβεια	0.975
Ειδικότητα	0.987
Ευαισθησία	0.938
Θετική Διαγνωστική Αξία	0.959
Αρνητική Διαγνωστική Αξία	0.981

Στη συνέχεια σχεδιάζουμε τα γραφήματα ROC και Precision-Recall, αφού πρώτα παρουσιάσουμε τη σημασία τους.

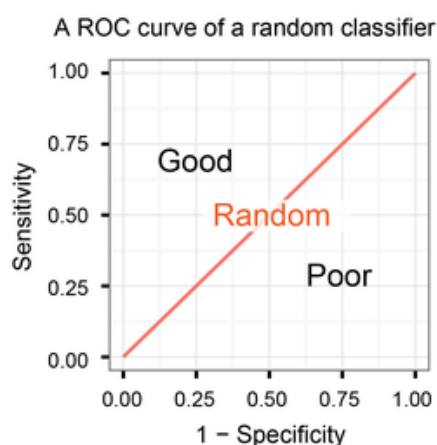
ROC

Το ROC είναι μια γραφική παράσταση του ψευδώς θετικού ρυθμού (ειδικότητα) (άξονας x) έναντι του πραγματικού θετικού ρυθμού (ευαισθησία) (άξονας y) για έναν αριθμό διαφορετικών τιμών κατωφλίου μεταξύ 0 και 1.

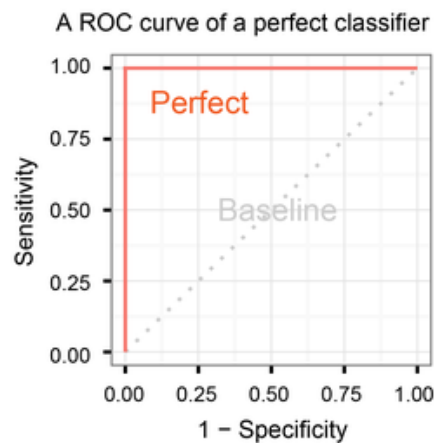
Μετρικές για ROC καμπύλη



Ένας τυχαίος ταξινομητής δείχνει πάντα μια ευθεία γραμμή από την αρχή (0,0) έως την επάνω δεξιά γωνία (1,1). Οι καμπύλες ROC στην περιοχή με την επάνω αριστερή γωνία (0, 1) υποδεικνύουν καλά επίπεδα απόδοσης, ενώ οι καμπύλες ROC στην άλλη περιοχή με την κάτω δεξιά γωνία (1,0) υποδεικνύουν χαμηλά επίπεδα απόδοσης.



Όσο πιο αριστερά βρίσκεται ένα σημείο στο γράφημα ROC τόσο μειώνεται το ποσοστό των False Positives. Ιδανικά θέλουμε ένα σημείο που να βρίσκεται κοντά ή πάνω στον yy άξονα ($x = 0$) και κοντά στην τιμή 1 στον yy ($y = 1$).



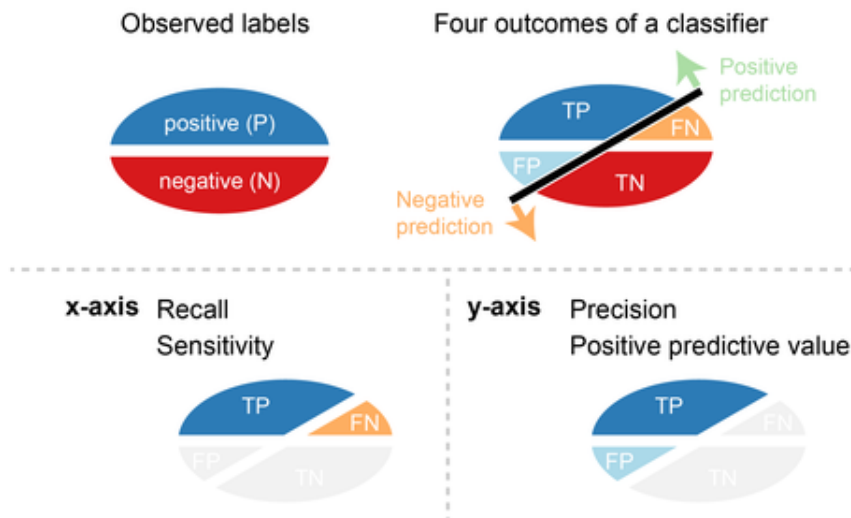
Οι μικρότερες τιμές στον άξονα- x της γραφικής παράστασης δείχνουν χαμηλότερα ψευδώς θετικά και υψηλότερα αληθινά αρνητικά. Μεγαλύτερες τιμές στον άξονα y της γραφικής παράστασης δείχνουν υψηλότερα αληθινά θετικά και χαμηλότερα ψευδώς αρνητικά.

Το γράφημα αυτό μας δείχνει την καλύτερη τιμή για το threshold, ανάλογα με το πόσα False Positive είμαστε διατεθειμένοι να αποδεχτούμε (ανάλογα με το εκάστοτε πρόβλημα).

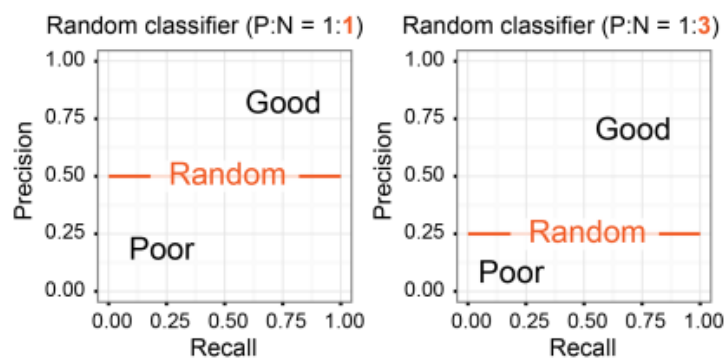
Precision - Recall

Η γραφική παράσταση ακρίβειας-ανάκλησης χρησιμοποιεί ανάκληση στον άξονα x και την ακρίβεια στον άξονα y. Η ανάκληση είναι πανομοιότυπη με την ευαισθησία και η ακρίβεια είναι ίδια με θετική προγνωστική τιμή.

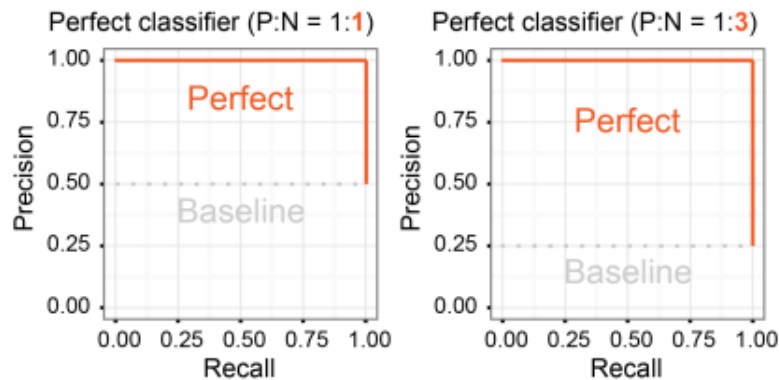
Μετρικές για Precision-Recall



Αντίστοιχα παρουσιάζονται ο τυχαίος ταξινομητής και ο τέλειος ταξινομητής σε precision-recall γράφημα.



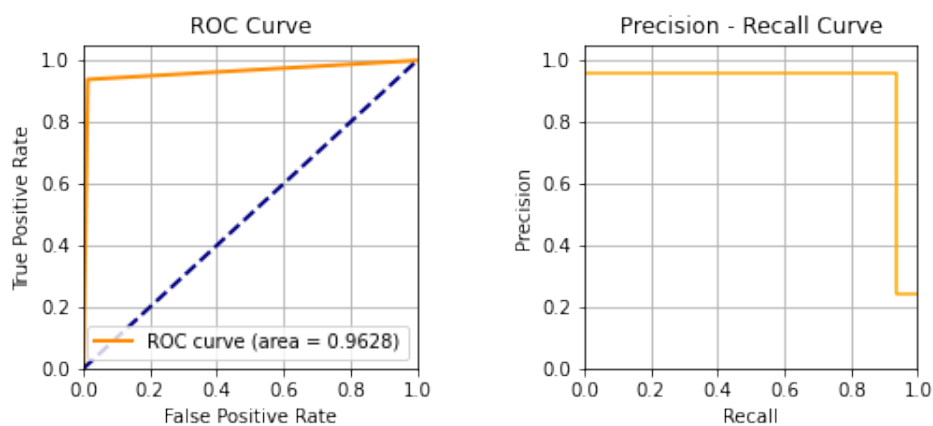
Ένας τυχαίος ταξινομητής δείχνει μια ευθεία γραμμή ως $P / (P + N)$. Για παράδειγμα, η γραμμή είναι $y = 0.5$ όταν ο λόγος θετικών και αρνητικών είναι 1:1, ενώ 0.25 όταν ο λόγος είναι 1:3.



Ένας τέλειος ταξινομητής δείχνει έναν συνδυασμό δύο ευθειών γραμμών. Το τελικό σημείο εξαρτάται από την αναλογία θετικών και αρνητικών. Για παράδειγμα, το τελικό σημείο είναι (1.0, 0.5) όταν ο λόγος θετικών και αρνητικών είναι 1:1, ενώ (1.0, 0.25) όταν ο λόγος είναι 1:3.

Ένα σημείο στο ROC γράφημα έχει 1-1 αντιστοίχιση στο γράφημα Precision-Recall

Παρουσίαση γραφημάτων



Το threshold αυτό μας έδωσε ένα σημείο στο ROC πολύ κοντά στην τιμή (1,0) και στο Precision-Recall ακριβώς πάνω στο (1,1), πράγμα που σημαίνει πως έκανε πολύ καλή δουλειά στην κατηγοριοποίηση των δειγμάτων. Επίσης

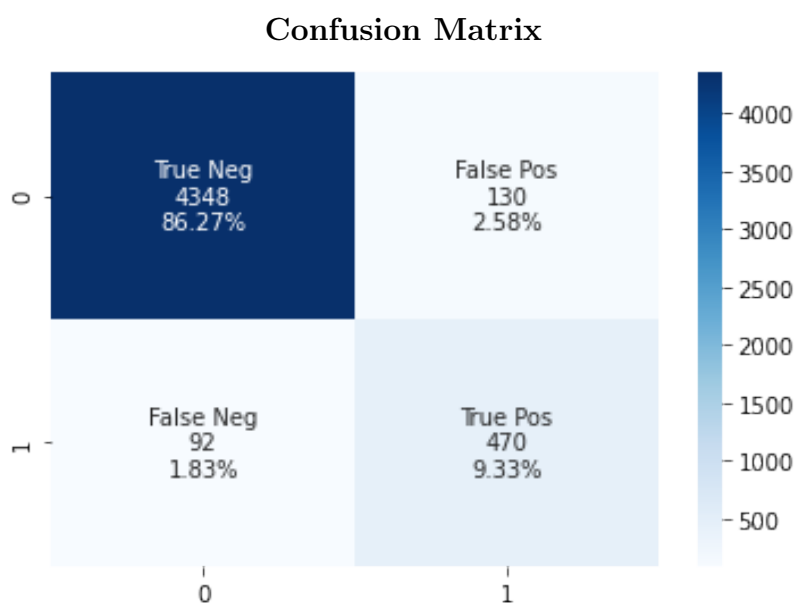
απο το precision-recall γράφημα βλέπουμε πως ο λόγος θετικών και αρνητικών είναι περίπου 1:3.

Ερώτηση 3

Εκφώνηση

Ομοίως, να αξιολογήσετε την ικανότητα του αισθητήρα συνεχούς καταγραφής γλυκόζης να ανιχνεύει τις υπεργλυκαιμίες που αντιστοιχούν σε τιμές γλυκόζης μεγαλύτερες από 180 mg/dl. Να υπολογίσετε τους δείκτες: Ευαισθησία, Ειδικότητα, Θετική Διαγνωστική Αξία, Αρνητική Διαγνωστική Αξία, Ακρίβεια

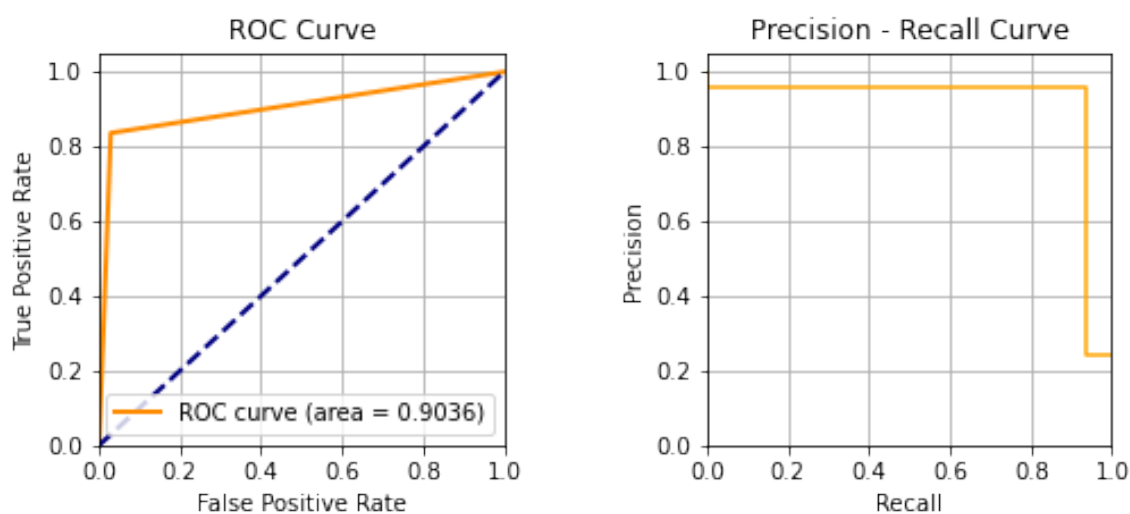
Απάντηση



Δείκτες αξιολόγησης:

Ακρίβεια	0.956
Ειδικότητα	0.971
Ευαισθησία	0.836
Θετική Διαγνωστική Αξία	0.783
Αρνητική Διαγνωστική Αξία	0.979

Γραφικές παραστάσεις



Το threshold αυτό μας έδωσε και πάλι καλά αποτελέσματα. Θετικό είναι το μικρό ποσοστό False Positive Rate καθώς είναι η σημαντικότερη μετρική για αυτό το πρόβλημα. Βλέπουμε πως το True Positive Rate είναι λίγο πάνω από 0.8 αλλά πρωταρχικός στόχος είναι να βρισκόμαστε κοντά στο μηδέν στο x' άξονα.

Ερώτηση 4

Εκφώνηση

Με βάση τα αποτελέσματα των ερωτημάτων 1, 2, και 3, να αναφέρετε τα συμπεράσματά σας σχετικά την ικανότητα του αισθητήρα συνεχούς καταγραφής γλυκόζης να ανιχνεύει τις υπογλυκαιμίες και τις υπεργλυκαιμίες.

Απάντηση

Παρατηρώντας το ROC και Precision-Recall γράφημα βλέπουμε πως τα αποτελέσματα είναι σχεδόν τέλεια, με την περίπτωση των υπογλυκαιμιών να προσεγγίζει τον ιδανικό ταξινομητή. Το ίδιο βλέπουμε και παρατηρώντας τις μετρικές, καθώς πετυχαίνουν υψηλές τιμές.

Η βελτίωση που ίσως θα μπορούσαμε να κάνουμε θα ήταν να μειώσουμε τα False Negatives, καθώς είναι ιδιαίτερα επικίνδυνα σε ένα ιατρικό ζήτημα, όπως το αναφερόμενο. Θα προτιμήσουμε αύξηση του Specificity σε βάρος όλων των άλλων μετρικών. Είναι προτιμότερο το σύστημα να κάνει πιο συχνά "false alarm" παρά να λείπει (έστω και σπάνια) σε έναν άνθρωπο που έχει πρόβλημα πως όλα είναι καλά. Το 1ο είναι ίσως ενοχλητικό, το 2ο είναι επικίνδυνο.