



Εισαγωγή στη Βιοϊατρική Μηχανική

Αξιολόγηση μετρήσεων επιπέδων γλυκόζης στο αίμα

2022-2023

Ηλιόπουλος Ανδρέας

Η εργασία υλοποιήθηκε σε python.

Στο αρχείο .zip υπάρχουν και τα αρχεία με τα δεδομένα καθώς είναι απαραίτητα για να τρέξει ο κώδικας.

Ερώτηση 1:

Η **MARD (Μέση απόλυτη σχετική διαφορά)** μετρά τη μέση σχετική διαφορά μεταξύ των μετρήσεων γλυκόζης του αισθητήρα και των τιμών αναφοράς γλυκόζης αίματος, εκφρασμένη ως ποσοστό. Μια χαμηλότερη τιμή MARD υποδεικνύει ότι οι μετρήσεις γλυκόζης του αισθητήρα είναι πιο ακριβείς και πιο κοντά στις τιμές αναφοράς γλυκόζης αίματος.

Ο **CC (συντελεστής συσχέτισης)** μετρά τον βαθμό γραμμικής σχέσης μεταξύ των μετρήσεων γλυκόζης του αισθητήρα και των τιμών αναφοράς γλυκόζης αίματος. Η τιμή CC κυμαίνεται από -1 έως 1, όπου η τιμή 1 υποδηλώνει τέλεια θετική γραμμική συσχέτιση μεταξύ των δύο μεταβλητών, η τιμή -1 υποδηλώνει τέλεια αρνητική συσχέτιση και η τιμή 0 υποδηλώνει μηδενική συσχέτιση. Μια υψηλότερη τιμή CC υποδηλώνει ισχυρότερη γραμμική σχέση μεταξύ των δύο μεταβλητών, γεγονός που σημαίνει ότι οι μετρήσεις γλυκόζης του αισθητήρα είναι ακριβέστερες στην παρακολούθηση των μεταβολών των επιπέδων γλυκόζης στο αίμα.

Για την επίλυση της άσκησης χωρίσαμε τα δεδομένα σε 3 περιπτώσεις: υπογλυκαιμία, ευγλυκαιμία και υπεργλυκαιμία. Έπειτα υπολογίστηκαν οι τιμές των 2 δεικτών MARD και CC, ώστε να αξιολογήσουμε σε κάθε περίπτωση ξεχωριστά το πόσο καλές είναι οι μετρήσεις του αισθητήρα. Ενδεικτικά ο κώδικας:

```
1 # Define glucose categories
2 hypoglycemia = df[BG < 70]
3 euglycemia = df[(BG >= 70) & (BG <= 180)]
4 hyperglycemia = df[BG > 180]
5
6 # Calculate MARD for each category
7 mard_hypo = np.mean(np.abs((Sensor[hypoglycemia.index] - BG[hypoglycemia.index]) / BG[hypoglycemia.index])) * 100
8 mard_eugly = np.mean(np.abs((Sensor[euglycemia.index] - BG[euglycemia.index]) / BG[euglycemia.index])) * 100
9 mard_hyper = np.mean(np.abs((Sensor[hyperglycemia.index] - BG[hyperglycemia.index]) / BG[hyperglycemia.index])) * 100
10
11 # Calculate CC for each category
12 cc_hypo = np.corrcoef(Sensor[hypoglycemia.index], BG[hypoglycemia.index])[0, 1]
13 cc_eugly = np.corrcoef(Sensor[euglycemia.index], BG[euglycemia.index])[0, 1]
14 cc_hyper = np.corrcoef(Sensor[hyperglycemia.index], BG[hyperglycemia.index])[0, 1]
15
16 print('MARD for hypoglycemia: {:.2f}%'.format(mard_hypo))
17 print('MARD for euglycemia: {:.2f}%'.format(mard_eugly))
18 print('MARD for hyperglycemia: {:.2f}%'.format(mard_hyper))
19 print('CC for hypoglycemia: {:.2f}'.format(cc_hypo))
20 print('CC for euglycemia: {:.2f}'.format(cc_eugly))
21 print('CC for hyperglycemia: {:.2f}'.format(cc_hyper))
```

Οι μετρήσεις που πήραμε:

	Υπογλυκαιμία	Ευγλυκαιμία	Υπεργλυκαιμία
MARD(%)	11.13	8.62	9.25
CC	0.29	0.88	0.54

Στα συγκεκριμένα αποτελέσματα λοιπόν, οι τιμές MARD για την υπογλυκαιμία, την ευγλυκαιμία και την υπεργλυκαιμία είναι 11,13%, 8,62% και 9,25%, αντίστοιχα. Επομένως, οι τιμές MARD < 15% και θεωρούνται αρκετά ακριβείς οι μετρήσεις.

Απ' την άλλη πλευρά, μια τιμή CC μεγαλύτερη από 0,7 θεωρείται καλή, ενώ τιμές μεταξύ 0,4 και 0,7 υποδηλώνουν μέτρια συσχέτιση. Η τιμή CC για υπογλυκαιμία θεωρείται κακή και δηλώνει κακή συσχέτιση, για υπεργλυκαιμία μέτρια και δηλώνει μέτρια συσχέτιση ενώ για ευγλυκαιμία δηλώνει αρκετά καλή συσχέτιση μετρήσεων-αληθινών τιμών.

Ερώτηση 2:

Αρχικά, τα TP, TN, FP και FN είναι μετρικές που χρησιμοποιούνται σε τεχνικές δυαδικής ταξινόμησης για την αξιολόγηση της απόδοσης ενός προγνωστικού μοντέλου. Αντιπροσωπεύουν τον αριθμό των αληθώς θετικών, αληθώς αρνητικών, ψευδώς θετικών και ψευδώς αρνητικών προβλέψεων, αντίστοιχα.

Λίγα λόγια για τα ζητούμενα metrics:

- Η **ευαισθησία(sensitivity)** μας δείχνει το ποσοστό των Positive που όντως εντοπίστηκαν από το σύστημα. Δίνεται απ' τον τύπο: $sensitivity = \frac{tp}{(tp + fn)}$.
- Η **ειδικότητα(specificity)** μας δείχνει το ποσοστό των Negative που όντως εντοπίστηκαν από το σύστημα. Δίνεται απ' τον τύπο: $specificity = \frac{tn}{(tn + fp)}$.
- Η **θετική διαγνωστική αξία(ppv)** μετρά το ποσοστό των ατόμων που διαγιγνώσκονται σωστά ως πάσχοντες από τη νόσο επί του συνόλου των ατόμων που διαγιγνώσκονται ως ασθενείς. Δίνεται απ' τον τύπο: $ppv = \frac{tp}{(tp + fp)}$.
- Η **αρνητική διαγνωστική(npv)** αξία μετρά το ποσοστό των ατόμων που διαγιγνώσκονται σωστά ως μη πάσχοντες από τη νόσο από το σύνολο των ατόμων που διαγιγνώσκονται ως μη πάσχοντες από τη νόσο. Δίνεται απ' τον τύπο:

$$npv = \frac{tn}{(tn + fn)}$$

- Η **ακρίβεια(precision)** που μας υποδεικνύει την γενικότερη απόδοση, την συνολική διακριτική ικανότητα του συστήματός. Δίνεται απ' τον τύπο:

$$precision = \frac{tp}{(tp + fp)}$$

Αρχικά, πρέπει να υπολογίσουμε τον αριθμό των True Positive, True Negative, False Positive και False Negative. Έπειτα, εφαρμόζουμε τους αντίστοιχους τύπους για κάθε μετρική.

```

1 # Count true positive, true negative, false positive, and false negative rates
2 tp = len(df[(BG < 70) & (Sensor < 70)])
3 tn = len(df[(BG >= 70) & (Sensor >= 70)])
4 fp = len(df[(BG >= 70) & (Sensor < 70)])
5 fn = len(df[(BG < 70) & (Sensor >= 70)])
6
7 # Calculate sensitivity, specificity, positive predictive value, negative predictive value, and precision
8 sensitivity = tp / (tp + fn)
9 specificity = tn / (tn + fp)
10 ppv = tp / (tp + fp)
11 npv = tn / (tn + fn)
12 precision = tp / (tp + fp)

```

Τα αποτελέσματα είναι:

```

True Positive (TP): 335
True Negative (TN): 4524
False Positive (FP): 81
False Negative (FN): 100

ROC Analysis:
Sensitivity: 77.01%
Specificity: 98.24%
Positive Predictive Value: 80.53%
Negative Predictive Value: 97.84%
Precision: 80.53%

```

Στο σύνολό τους, τα αποτελέσματα αυτά υποδηλώνουν ότι ο αισθητήρας έχει υψηλή ακρίβεια στον εντοπισμό τόσο θετικών όσο και αρνητικών περιπτώσεων, ωστόσο υπάρχει περιθώριο βελτίωσης.

Ερώτηση 3:

Ομοίως για την περίπτωση της υπεργλυκαιμίας υπολογίστηκαν οι ίδιες μετρικές:

```

True Positive (TP): 524
True Negative (TN): 4275
False Positive (FP): 96
False Negative (FN): 145

ROC Analysis for Hyperglycemia:
Sensitivity: 78.33%
Specificity: 97.80%
Positive Predictive Value: 84.52%
Negative Predictive Value: 96.72%
Precision: 84.52%

```

Σε γενικές γραμμές τα αποτελέσματα αυτά υποδηλώνουν ότι ο αισθητήρας έχει υψηλή ακρίβεια και στις περιπτώσεις υπεργλυκαιμίας.

Ερώτηση 4:

- Με βάση τα αποτελέσματα της ανάλυσης **ROC**, ο αισθητήρας συνεχούς καταγραφής γλυκόζης παρουσιάζει ικανοποιητική ικανότητα ανίχνευσης τόσο της υπογλυκαιμίας όσο και της υπεργλυκαιμίας.
- Η ευαισθησία, η ειδικότητα, η θετική διαγνωστική αξία και η αρνητική διαγνωστική αξία και για τις δύο καταστάσεις είναι πάνω από 77%, υποδεικνύοντας ότι ο αισθητήρας μπορεί να εντοπίσει σωστά τις περισσότερες περιπτώσεις μη φυσιολογικών επιπέδων γλυκόζης.
- Οι τιμές **MARD** και για τις τρεις καταστάσεις είναι κάτω από 12%, υποδεικνύοντας ότι οι ενδείξεις του αισθητήρα είναι αρκετά κοντά στις τιμές αναφοράς. Ο συντελεστής συσχέτισης(**CC**) για την ευγλυκαιμία είναι ο υψηλότερος μεταξύ των τριών συνθηκών, γεγονός που υποδηλώνει ότι η απόδοση του αισθητήρα είναι πιο συνεπής στο φυσιολογικό εύρος γλυκόζης, ωστόσο παρουσιάζει μέτριες προς κακές τιμές για τις περιπτώσεις υπεργλυκαιμίας και υπογλυκαιμίας, δηλώνοντας ότι υπάρχει ακόμα περιθώριο βελτίωσης.
- Συνολικά, ο αισθητήρας συνεχούς γλυκόζης φαίνεται να είναι ένα αρκετά ακριβές εργαλείο για την ανίχνευση της υπογλυκαιμίας και της υπεργλυκαιμίας.