

1η Εργασία - Αναγνώριση Προτύπων

Διδάσκων: Επικ. Καθ. Παναγιώτης Πετραντωνάκης (ppetrant@ece.auth.gr)

Βοηθός διδασκαλίας: Υπ. Διδ. Στέφανος Παπαδόπουλος (stefpapad@iti.gr)

21 Οκτωβρίου 2022

Σε αυτή την εργασία θα εξοικειωθείτε με την Θεωρία Αποφάσεων κατά Bayes. Συγκεκριμένα, σας ζητείται να υλοποιήσετε έναν ταξινομητή κατά Bayes και να τον εφαρμόσετε σε γκαουσιανές κατανομές με συγκεκριμένες παραμέτρους.

Μέρος Α (0.5)

Εργάζεστε σε ένα διαγνωστικό κέντρο ως Μηχανικός Ανάλυσης Δεδομένων. Δουλειά σας είναι να αναπτύσετε διαγνωστικά μοντέλα με βάσει τα βιοιατρικά δεδομένα που σας δίνονται από το Μικροβιολογικό Τμήμα του Κέντρου και να παρέχετε, με βάση αυτά τα μοντέλα, υποστηρικτικό ρόλο στις εξειδικευμένες αποφάσεις των γιατρών του Κέντρου. Καλείστε να αναπτύξετε ένα σύστημα απόφασης κατά Bayes και να το παρουσιάσετε στους αρμόδιους γιατρούς. Τα δεδομένα σας αποτελούν δυο μικροβιολογικούς δείκτες που εξάγονται κατά την εξέταση δείγματος αίματος και είναι γνωστό ότι σχετίζονται με συγκεκριμένη μορφή καρκίνου. Από πολύχρονη ανάλυση δεδομένων έχετε βρει ότι τα δεδομένα σας ακολουθούν γκαουσιανές συναρτήσεις πυκνότητας πιθανότητας $p(\mathbf{x}|\omega_1)$ και $p(\mathbf{x}|\omega_2)$ όπου \mathbf{x} είναι το διάνυσμα των δύο δεικτών. Οι μέσες τιμές για τις κλάσεις ω_1 (υγιής) και ω_2 (πιθανή ύπαρξη καρκίνου) είναι

$$\boldsymbol{\mu}_1 = (0.4, 0.8) \quad \text{και} \quad \boldsymbol{\mu}_2 = (1.5, 2.7),$$

αντίστοιχα. Οι κατανομές έχουν κοινό πίνακα συσχετιστικότητας

$$\Sigma = \begin{pmatrix} 1.5 & 0 \\ 0 & 0.8 \end{pmatrix}.$$

Οι a priori πιθανότητες των κλάσεων είναι $P(\omega_1) = 0.95$ και $P(\omega_2) = 0.05$.

1. Παρουσιάστε στους γιατρούς τις τιμές των κατανομών $p(\mathbf{x}|\omega_1)$ και $p(\mathbf{x}|\omega_2)$ για διακριτές τιμές \mathbf{x} .
2. Υπολογίστε την ολική κατανομή πιθανότητας και παρουσιάστε τη.
3. Υπολογίστε τις a posteriori πιθανότητες $P(\omega_1|\mathbf{x})$ και $P(\omega_2|\mathbf{x})$ και συμπεριλάβετε τις στην παρουσίασή σας.
4. Υπολογίστε το σφάλμα κατά Bayes. Κάνετε ένα σχόλιο για το σφάλμα που παρατηρείτε.

Μέρος Β (0.5)

Σε αυτό το μέρος, θεωρείστε $\Sigma_1 = \Sigma$ και $\Sigma_2 = \Sigma/4$.

1. Επαναλάβετε τα βήματα 1-4 από το μέρος Α.
2. Τι παρατηρείτε για το σφάλμα κατά Bayes σε αυτή τη περίπτωση; Πώς εξηγείτε τη διαφορά με το αντίστοιχο σφάλμα του μέρους Α;

Οδηγίες

1. Η Υλοποίηση της εργασίας θα γίνει σε Python. Επιλέξτε ένα notebook (π.χ., Jupyter, Collab) και γράψτε τον κώδικα όσο και τα σχόλιά σας.
2. Για την παράδοση θα ανεβάσετε ΕΝΑ αρχείο με όνομα: surname1_AEM_surname2_AEM_Assinment1.ipynb (σε περίπτωση που κάνετε την εργασία μόνοι, βάλτε μόνο το επώνυμο και το AEM σας. Αν είστε ομάδα δύο ατόμων, ΜΟΝΟ ένας κατεθέτει την εργασία) με όλες τις απαντήσεις. Στο αρχείο αυτό θα αναγράφονται τα στοιχεία σας (ονοματεπώνυμο, AEM) σε ένα textbox στην αρχή.

3. Κάθε ένα από τα ερωτήματα θα απαντηθεί (κώδικας) σε ξεχωριστό κελί. Και ο κώδικας σε κάθε κελί θα συνοδεύεται από σύντομα σχόλια.
4. Τελική ημερομηνία υποβολής: Παρασκευή 18 Νοεμβρίου, 2022, 23:59.

ΚΑΛΗ ΕΠΙΤΥΧΙΑ!