

2η Εργασία - Αναγνώριση Προτύπων

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

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

18 Νοεμβρίου 2022

Σε αυτή την εργασία θα ασχοληθείτε με την εκτίμηση παραμέτρων με δυο βασικές μεθοδολογίες: την Εκτίμηση Μέγιστης Πιθανοφάνειας και την Εκτίμηση κατά Bayes.

Μέρος Α (0.5)

Εργάζεστε σε μια εταιρεία που παράγει βιντεοπαιχίδια και συγκεκριμένα σε ένα τμήμα που πρόσφατα δημιουργήθηκε στην εταιρεία και ασχολείται με την αναγνώριση του επιπέδου του στρές στους χρήστες με βάση τα μοτίβα συχνότητας και δύναμης πίεσης των πλήκτρων της κονσόλας. Ένας συνάδελφός σας από το ίδιο τμήμα αναλύοντας αυτά τα μοτίβα εξήγαγε έναν δείκτη-αριθμό x και ισχυρίζεται ότι αυτός ο δείκτης μπορεί να χρησιμοποιηθεί σε ένα σύστημα ταξινόμησης για να διαπιστωθεί κάθε φορά αν ο χρήστης αισθάνεται στρες ή όχι. Επίσης από μελέτες που διεξήγαγε ο συνάδελφός σας, παρατήρησε ότι η κατανομή πυκνότητας πιθανότητας που ακολουθεί αυτός ο δείκτης και για τις δυο κλάσεις (χωρίς στρες = κλάση ω_1 , με στρες = κλάση ω_2) είναι:

$$p(x|\theta) = \frac{1}{\pi} \frac{1}{1 + (x - \theta)^2}$$

με τη παράμετρο θ να είναι άγνωστη. Για να διαπιστώσετε αν όντως ο δείκτης που εξήγαγε ο συνάδελφός σας είναι αξιόπιστος δείκτης για το επίπεδο του στρες των χρηστών, ζητήσατε από 12 άλλους συναδέλφους σας στο ίδιο τμήμα να παίξουν ένα συγκεκριμένο παιχνίδι στην κονσόλα που παράγει η εταιρεία και υπολογίσατε τον εν λόγω δείκτη για κάθε χρήστη. Στη συνέχεια ρωτήσατε τους συναδέλφους σας να σας πουν αν ένιωσαν στρες κατά τη διάρκεια του παιχνιδιού ή όχι. Από τους 12 οι 7 δήλωσαν ότι δεν ένιωσαν στρες ενώ οι 5 δήλωσαν ότι το παιχνίδι τους δημιούργησε έντονο στρες. Καλήστε λοιπόν εσείς να υλοποιήσετε ένα ταξινομητή μέγιστης πιθανοφάνειας. Συγκεκριμένα:

1. Εκτιμήστε τις παραμέτρους $\hat{\theta}_1$ και $\hat{\theta}_2$ με την μέθοδο της μέγιστης πιθανοφάνειας και για τις δύο κλάσεις αν για την κλάση ω_1 οι δείκτες είναι $D_1 = [2.8, -0.4, -0.8, 2.3, -0.3, 3.6, 4.1]$ ενώ για την κλάση ω_2 είναι $D_2 = [-4.5, -3.4, -3.1, -3.0, -2.3]$. Απεικονίστε τις $\log p(D_1|\theta)$ και $\log p(D_2|\theta)$ σε συνάρτηση με τη θ .
2. Χρησιμοποιήστε τη συνάρτηση διάκρισης

$$g(x) = \log P(x|\hat{\theta}_1) - \log P(x|\hat{\theta}_2) + \log P(\omega_1) - \log P(\omega_2)$$

και ταξινομήστε τα δύο σύνολα τιμών. Τι παρατηρείτε για το πρόσημο της $g(x)$ σε σχέση με τα δεδομένα σας (απεικονίστε την); Περιγράψτε τον κανόνα απόφασης. Τι παρατηρείτε σε σχέση με την ταξινόμηση των δεδομένων σας με βάση αυτόν τον κανόνα; Βοήθεια: Μπορείτε να δημιουργήσετε μια κλαση Classifier με a) μια συνάρτηση *fit* που θα λαμβάνει ως όρισμα ένα σύνολο D και ένα διάνυσμα από υποψήφιες τιμές θ και θα υπολογίζει τις τιμές μέγιστης πιθανοφάνειας για το θ , β) μια συνάρτηση *predict* που θα λαμβάνει ως όρισμα ένα σύνολο D και τις *a priori* πιθανότητες των κλάσεων και θα επιστρέφει τις τιμές της συνάρτησης g

Μέρος Β (0.5)

Σε αυτό το μέρος καλείστε να υλοποιήσετε ένα νέο ταξινομητή εκτιμώντας την άγνωστη παράμετρο θ με την μέθοδο εκτίμησης κατά Bayes.

Μετά από επίπονο πειραματισμό, διαπιστώσατε ότι οι τιμές τις παραμέτρου θ μπορούν να μοντελοποιηθούν με την συνάρτηση πυκνότητας πιθανότητας (prior)

$$p(\theta) = \frac{1}{10\pi} \frac{1}{1 + (\theta/10)^2}.$$

Έχοντας αυτό το μοντέλο και με βάση τη θεωρία είσατε τώρα σε θέση να υπολογίσετε την *a posteriori* πιθανότητα $p(\theta|D)$ και την πυκνότητα πιθανότητας $p(x|D_j)$, $j = 1, 2$.

1. Απεικονίστε τις εκ των υστέρων πυκνότητες πιθανότητας $p(\theta|\mathcal{D}_1)$ και $p(\theta|\mathcal{D}_2)$. Τι παρατηρείτε σε σχέση με την (prior) $p(\theta)$. *Βοήθεια: Για τον υπολογισμό των ολοκληρωμάτων μπορείτε να χρησιμοποιήσετε τον κανόνα του τραπεζίου*
2. Υλοποιήστε μια συνάρτηση predict που θα υπολογίζει τις τιμές μιας συνάρτησης διάκρισης

$$h(x) = \log P(x|\mathcal{D}_1) - \log P(x|\mathcal{D}_2) + \log P(\omega_1) - \log P(\omega_2).$$

Τι παρατηρείτε τώρα για τις τιμές τις h σε σχέση με τα σύνολα δεδομένων σας (απεικονίστε την); Πως αξιολογείτε την μέθοδο εκτίμησης παραμέτρων κατά Bayes σε σχέση με τη μέθοδο της μέγιστης πιθανοφάνειας για το συγκεκριμένο παράδειγμα; Που πιστεύετε ότι οφείλεται η διαφορά των δύο προσεγγίσεων για το συγκεκριμένο παράδειγμα; *Βοήθεια: Μπορείτε να υιοθετήσετε παρόμοια υλοποίηση με το μέρος A*

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_Assinment2.ipynb (σε περίπτωση που κάνετε την εργασία μόνοι: surname_AEM_Assinment2.ipynb. 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. ΠΡΟΣΟΧΗ: Οι ομάδες πρέπει να είναι οι ίδιες με την εργασία 1!
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 16 Δεκεμβρίου, 2022, 23:59.

ΚΑΛΗ ΕΠΙΤΥΧΙΑ!