Η παράδοση των ασκήσεων θα γίνει στην αρχή του μαθήματος της Τετάρτης 31ης Μαΐου. Επισημαίνεται ότι οι εργασίες είναι **ατομικές** και ότι συνεισφέρουν **8%** του τελικού βαθμού.

---

**Άσκηση 1:** Βρείτε τους εκτιμητές μέγιστης πιθανοφάνειας (maximum likelihood estimators) των παραμέτρων των παρακάτω κατανομών, δοθέντων $N$ στατιστικά ανεξάρτητων δείγμάτων $x_n\,,\ n = 1\,,2\,,...,N$:

(a) Κανονικής κατανομής, δηλ. της:

$$p\,(\,x\,;\,\mu\,,\sigma\,) \;=\; \frac{1}{\sqrt{2\,\pi}\;\sigma}\;\exp\left(-\,\frac{(\,x-\mu\,)^2}{2\,\sigma^2}\right) \quad .$$

(b) Εκθετικής κατανομής, δηλ. της:

$$p\,(\,x\,;\,\theta\,) \;=\; \begin{cases} \dfrac{1}{\theta}\,\exp\left(-\,\dfrac{x}{\theta}\right) \;, & x \,\geq\, 0 \\ 0\;, & \text{αλλού} \end{cases} \quad .$$

Και στις δύο περιπτώσεις, διαπιστώστε αν οι εκτιμητές είναι αμερόληπτοι (unbiased), και, αν όχι, μετατρέψετέ τους σε αμερόληπτους.

---

**Άσκηση 2:** Έστω δύο κλάσεις $\omega_1$ και $\omega_2$ με τα ακόλουθα 2-D διανύσματα χαρακτηριστικών εκπαίδευσης (τέσσερα για κάθε κλάση, γραμμένα σε μορφή $\mathbf{x} = [\,x_1\,,x_2\,]^{\mathrm{T}}$):

$$\omega_1:\quad [\,1\,,1\,]^{\mathrm{T}}\,,\quad [\,2\,,2\,]^{\mathrm{T}}\,,\quad [\,1\,,3\,]^{\mathrm{T}}\,,\quad [\,0\,,2\,]^{\mathrm{T}}\;,$$

$$\omega_2:\quad [\,-1\,,-1\,]^{\mathrm{T}}\,,\quad [\,-2\,,-2\,]^{\mathrm{T}}\,,\quad [\,-1\,,-3\,]^{\mathrm{T}}\,,\quad [\,0\,,-2\,]^{\mathrm{T}}\;.$$

Έστω επίσης και ένα διάνυσμα δοκιμής $\mathbf{x}_{\text{test}} = [\,1/2\,,1/2\,]^{\mathrm{T}}$.

(a) Σχεδιάστε τα παραπάνω διανύσματα στο 2-D χώρο με άξονες τα $x_1\,,\ x_2$. Είναι οι κλάσεις γραμμικά διαχωρίσιμες (με βάση τα διανύσματα εκπαίδευσης); Αν ναι, δώστε μία συνάρτηση διάκρισης $g(\mathbf{x}) = w_1\,x_1 + w_2\,x_2 + w_0$ με διάνυσμα συντελεστών $\mathbf{w} = [\,w_1\,,w_2\,,w_0\,]^{\mathrm{T}}$, που διαχωρίζει τις δύο κλάσεις με $g(\mathbf{x}) > 0$ για $\mathbf{x} \in \omega_1$ και $g(\mathbf{x}) < 0$ για $\mathbf{x} \in \omega_2$ (για τα διανύσματα εκπαίδευσης) και που ταξινομεί το διάνυσμα δοκιμής στην κλάση $\omega_2$.

(b) Ταξινομήστε το διάνυσμα δοκιμής με βάση τον κανόνα του πλησιέστερου γείτονα (1-NN), χρησιμοποιώντας οποιοδήποτε μέτρο απόστασης μεταξύ διανυσμάτων στον δυδιάστατο χώρο επιθυμείτε.

(c) Βρείτε τη γραμμική συνάρτηση διάκρισης των δύο κλάσεων με βάση τον αλγόριθμο perceptron (σε μορφή batch), αρχικοποιώντας τον αλγόριθμο με τιμές $\mathbf{w}(0) = [\,0\,,0\,,\,0\,]^{\mathrm{T}}$ και χρησιμοποιώντας βήμα $\rho = 1/8$. Σχεδιάστε την και ταξινομήστε το $\mathbf{x}_{\text{test}}$ με αυτήν. Σημείωση: Ο συμβολισμός της $g(\mathbf{x})$ και τα πρόσημα επιλογής των κλάσεων είναι όπως στο υπο-ερώτημα (a).

(d) Βρείτε τη γραμμική συνάρτηση διάκρισης των δύο κλάσεων με βάση την μέθοδο των μηχανών διανυσματικής στήριξης (support vector machines), εκφράζοντας το διάνυσμα $[\,w_1\,,w_2\,]^{\mathrm{T}}$ ως γραμμικό συνδυασμό των διανυσμάτων στήριξης (support vectors) και στη συνέχεια βρίσκοντας και το $w_0$. Ποιο είναι το περιθώριο (margin) μεταξύ των κλάσεων που επιτυγχάνει ο αλγόριθμος; Σχεδιάστε την συνάρτηση διάκρισης και ταξινομήστε το $\mathbf{x}_{\text{test}}$ με αυτήν.

(e) Βρείτε τη γραμμική συνάρτηση διάκρισης των δύο κλάσεων με βάση την ελαχιστοποίηση του αθροίσματος τετραγωνικών σφαλμάτων (minimum least squares error). Πόσο είναι το ελάχιστο σφάλμα εκπαίδευσης που πετυχαίνει ο αλγόριθμος; Σχεδιάστε την συνάρτηση διάκρισης και ταξινομήστε το $\mathbf{x}_{\text{test}}$ με αυτήν.

(f) Βρείτε την ευθεία διάκρισης μεταξύ των δύο κλάσεων με βάση την Bayesian προσέγγιση ταξινόμησης, υποθέτοντας κανονικές κατανομές (2-D Gaussian) για τις υπό συνθήκη συναρτήσεις πυκνότητας πιθανότητας $p\,(\,\mathbf{x}\,|\,\omega_1\,)$ και $p\,(\,\mathbf{x}\,|\,\omega_2\,)$, εκτιμώντας τις παραμέτρους τους με αμερόληπτους εκτιμητές βασισμένους στη μέθοδο της μέγιστης πιθανοφάνειας. Στη συνέχεια ταξινομήστε το διάνυσμα δοκιμής, $\mathbf{x}_{\text{test}}$.

(g) Μειώστε την διάσταση των δεδομένων από δύο σε μία χρησιμοποιώντας ανάλυση κύριων συνιστωσών (Principal Component Analysis – PCA). Βρείτε τις προβολές των διανυσμάτων εκπαίδευσης στον μονοδιάστατο χώρο που προκύπτει, και διαπιστώστε αν οι κλάσεις είναι γραμμικά διαχωρίσιμες σε αυτόν τον χώρο ή όχι. Αν ναι, ταξινομήστε το παράδειγμα δοκιμής στην κατάλληλη κλάση, με βάση τον κανόνα του πλησιέστερου γείτονα (1-NN) στον μονοδιάστατο χώρο που προκύπτει.

(h) Επαναλάβετε τα βήματα του υπο-ερωτήματος (g), αλλά χρησιμοποιώντας αυτή τη φορά ανάλυση γραμμικής διάκρισης (Linear Discriminant Analysis – LDA).

---

**Άσκηση 3:** Έστω δύο κλάσεις $\omega_1$ και $\omega_2$ με τα ακόλουθα 2-D διανύσματα χαρακτηριστικών εκπαίδευσης (τέσσερα για κάθε κλάση, γραμμένα σε μορφή $\mathbf{x} = [\,x_1\,,x_2\,]^{\mathrm{T}}$):

$$\omega_1: \quad [\,1\,,1\,]^{\mathrm{T}}, \quad [\,1\,,-1\,]^{\mathrm{T}}, \quad [\,-3\,,1\,]^{\mathrm{T}}, \quad [\,-3\,,-1\,]^{\mathrm{T}}\,,$$

$$\omega_2: \quad [\,3\,,1\,]^{\mathrm{T}}, \quad [\,3\,,-1\,]^{\mathrm{T}}, \quad [\,-1\,,1\,]^{\mathrm{T}}, \quad [\,-1\,,-1\,]^{\mathrm{T}}\,.$$

(a) Σχεδιάστε τα παραπάνω διανύσματα στο 2-D χώρο με άξονες τα $x_1$, $x_2$. Είναι οι κλάσεις γραμμικά διαχωρίσιμες (με βάση τα διανύσματα εκπαίδευσης);

(b) Κατασκευάστε ένα perceptron δύο επιπέδων που να ταξινομεί τα διανύσματα εκπαίδευσης σωστά στις δύο κλάσεις.

(c) Λύστε το παραπάνω πρόβλημα με δίκτυο ακτινωτής συνάρτησης βάσης (RBF).

(d) Σχεδιάστε ένα δέντρο αποφάσεων (decision tree) για το παραπάνω πρόβλημα.

---

**Άσκηση 4:** Τα παρακάτω είναι ανεξάρτητα ερωτήματα:

(a) Έστω τα έξι παρακάτω διανύσματα εκπαίδευσης τα οποία παρουσιάζονται στο βασικό ακολουθιακό αλγοριθμικό σχήμα ομαδοποίησης (BSAS) με την ακόλουθη σειρά:

$\mathbf{x}_1 = [\,1\,,1\,]^{\mathrm{T}}$, $\mathbf{x}_2 = [\,2\,,1\,]^{\mathrm{T}}$, $\mathbf{x}_3 = [\,1\,,2\,]^{\mathrm{T}}$, $\mathbf{x}_4 = [\,-1\,,-1\,]^{\mathrm{T}}$, $\mathbf{x}_5 = [\,-2\,,-1\,]^{\mathrm{T}}$, $\mathbf{x}_6 = [\,-1\,,-2\,]^{\mathrm{T}}$.

Ομαδοποιήστε τα διανύσματα αυτά με βάση τον αλγόριθμο BSAS, χρησιμοποιώντας την απόσταση $L_1$ (Manhattan distance), κατώφλι $\theta = 1.25$, και θεωρώντας ότι οι αντιπρόσωποι των ομάδων που προκύπτουν είναι η μέση τιμή των μελών της ομάδας.

(b) Έστω διανύσματα εκπαίδευσης:

$$\mathbf{x}_1 = [\,1\,,1\,]^{\mathrm{T}}, \quad \mathbf{x}_2 = [\,2\,,2\,]^{\mathrm{T}}, \quad \mathbf{x}_3 = [\,-1\,,-1\,]^{\mathrm{T}}, \quad \mathbf{x}_4 = [\,-2\,,-1\,]^{\mathrm{T}}.$$

Ομαδοποιήστε τα ιεραρχικά, χρησιμοποιώντας τον συσσωρευτικό αλγόριθμο ομαδοποίησης (agglomerative clustering) με βάση την απόσταση $L_1$ (Manhattan distance). Χρησιμοποιείστε δύο διαφοροποιήσεις του αλγορίθμου, όσον αφορά την ενημέρωση των αποστάσεων του πίνακα ανομοιότητας, τον πρώτο με βάση τον αλγόριθμο απλού δεσμού (single link), και τον δεύτερο με βάση τον αλγόριθμο πλήρους δεσμού (complete link). Και στις δύο περιπτώσεις σχεδιάστε τα δεντρογράμματα ανομοιότητας (dendrograms) που προκύπτουν.

(c) Υπολογίστε με διαδικασία δυναμικής χρονικής στρέβλωσης (dynamic time warping) το κόστος σύγκρισης μεταξύ της ακολουθίας τεσσάρων 2-D χαρακτηριστικών (πρότυπο αναφοράς)

$$X = \begin{bmatrix} 1.0 & -1.0 & -1.0 & 0.0 \\ 1.0 & 0.0 & -1.0 & -1.0 \end{bmatrix}$$

και της ακολουθίας πέντε 2-D χαρακτηριστικών (υπό εξέταση πρότυπο)

$$Y = \begin{bmatrix} 2.0 & 0.0 & -1.0 & -1.0 & 1.0 \\ 1.0 & -1.0 & 0.0 & -2.0 & -3.0 \end{bmatrix}$$

υπό τοπικούς περιορισμούς μετάβασης Sakoe-Chiba (σχήμα 7.11.α) του βιβλίου, κόστος σύγκρισης διανυσμάτων την απόσταση $L_1$ (Manhattan distance), και περιορισμούς άκρων που απαιτούν το βέλτιστο μονοπάτι να διέρχεται από τον αρχικό και τον τελικό κόμβο που αποτελούνται από τα ζεύγη των πρώτων και τελευταίων διανυσμάτων των προτύπων αναφοράς και δοκιμής, αντίστοιχα.