

Ακαδημαϊκό Έτος 2017 - 2018

ΑΝΑΓΝΩΡΙΣΗ ΠΡΟΤΥΠΩΝ

ΑΣΚΗΣΗ 1

ΠΡΟΟΔΟΣ 2017

- Να εξηγήσετε πως λειτουργεί ο αλγό. Expectation Maximisation.
- Πως λειτουργεί αίσια και αιτιατά;
- Να εξηγήσετε τη δομή του HMM2.

ΑΣΚΗΣΗ 2

2 ισοκύβαντες Gaussian $\mu_1 = (1,0)$, $\mu_2 = (3,0)$, $\sigma_1^2 = \sigma_2^2 = 2$
α-priori

- Επιφάνεια διαχωρισμού
- Σφάλμα ωραρισης του $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp(-\frac{x^2}{2}) dx$
- Σφάλμα σε προβολή στον x - y
- — — — —

ΑΣΚΗΣΗ 3

Πίνακας Μετάβασης

S: sun
R: rain
F: fog

	S	R	F
S	0.7	0.1	0.2
R	0.1	0.7	0.2
F	0.2	0.3	0.5

O: ο Γωσδέλης έφερε ομπρέλα
 $P(O|R) = 0.8$
 $P(O|S) = 0.2$
 $P(O|F) = 0.3$

α) Τιθανότητα έχθες ήλιο, αύριο να βρέχει.

- Τιθανότητα έχθες να είχε ήλιο και σε 2 μέρες να βρέχει
- — — — — να βρέχει αν έχθες είχε ήλιο και σήμερα ο Γω. έφερε
- — — — — Την 3η ημέρα να βρέχει αν την 1η έχει ήλιο, την 2η ο Γωσδέλης δεν έφερε ομπρέλα και την 3η έφερε.

Θέματα Αναγνώρισης Προτύπων

Πολλαπλής (36/100)

1. Ποιος αλγο ταξινόμησης είναι supervised \rightarrow LDA
2. ML γενικά στην στατιστική χ διανυσμάτων $\rightarrow p(\chi|\theta)$, θ σταθερό ντετερμινιστικό
3. Πολυπλοκότητα αλγο forward-backward $\rightarrow O(N^2T)$
4. Τι κοινό έχουν PCA, LDA \rightarrow Ο υποχώρος προκύπτει από τα ιδιοδιανύσματα κάποιου πίνακα
5. Τι δεν κάνει ο viterbi ? \rightarrow υπολογισμό ακριβής τιμής πιθανότητας $P(O,Q|\lambda)$
5. Αν εφαρμόσω ICA οι συνιστώσες τι δεν κάνουνε \rightarrow η προσέγγιση τους με gaussian ΔΕΝ αυξάνει. (μειώνεται η εντροπία)
6. MLP 2ο και 3ο Layers \rightarrow AND και OR αντίστοιχα
7. SVM εξίσωση διαχωριστικής επιφάνειας με πυρήνα
8. στο LDA ο υποχώρος p παίρνει κάποια συγκεκριμένη τιμή (ή αυτό ρώταγε ή για την διασταση των νέων δεδομένων y) αν D η διάσταση του τωρινού και C ο αριθμός των κλάσεων \rightarrow κανένα από τα παραπάνω (παραπάνω: $C-1$, C , $D/2$)
9. Σύγκλιση Perceptron \rightarrow κλάσεις γραμμικά διαχωρίσιμες
10. GMM και K-means.

Άσκηση 1 (9/100)

1. Να δώσετε μια probabilistic εκδοχή του αλγορίθμου CART
2. Για το πρόβλημα XOR, να σχολιάσετε την απόδοση του NN3, CART, SVM, Eukleidian Class
3. Πότε χρησιμοποιούμε ML (Max.Lik) και πότε EM αλγόριθμο

Άσκηση 2 (18/100)

Δίνονται 4 states HMM και μια ακολουθία 10 παρατηρήσεων.

1. Να υπολογιστούν οι συντελεστές δ_i
2. Να υπολογιστεί η ακολουθία των states που παράγουν αυτές τις παρατηρήσεις
3. Να υπολογιστεί η πιθανότητα $P(Q^*, O|\lambda)$

Άσκηση 3 (23/100)

Δίνονται δύο κλάσεις ω_1 και ω_2 . Η ω_1 θεωρούμε ότι ακολουθεί την κατανομή

$$P(x|\omega_1) = w_{11} * N(\mu_{11}, 1) + w_{12} * N(\mu_{12}, 1)$$

και η ω_2 την εξής :

$$P(x|\omega_2) = N(\mu, 1)$$

Δίνεται επίσης και train set $D1 = \{0, 1, 2, -3, 6.9, 7.2, 8.2, 9.2\}$ και $D2 = \{2.1, 3, 4, 5\}$

1. Να υπολογιστούν με χρήση ML οι παράμετροι μ , μ_{11} , μ_{12} , $P(\omega_1)$ και $P(\omega_2)$
2. Να βρεθούν τα διαστήματα απόφασης για κάθε κατανομή
3. Σχεδίαση CART
3. Να βρεθούν τα διαστήματα απόφασης με NNR3
4. Να βρεθούν με K-means με $\mu_1 = 1$ και $\mu_2 = 5$
5. Να γίνει η σύγκριση CART Error, Bayessian Error και NN3 error
6. Εστω ότι $D = D1$ ένωση $D2$. Να εκτελέσεις EM με αρχικοποιημένες τιμές που έχουν υπολογιστεί από τον K-means

Ασκηση 4 (14/100)

Εστω LDA.

Δίνεται ο τύπος S_w και S_b (δες διαφ 58 σετ compact repr)

Να αποδειχθεί ότι :

$$S_b = P(\omega_1)P(\omega_2)(\mu_1 - \mu_2)(\mu_1 - \mu_2)^T$$

και να λυθεί το γενικευμένο πρόβλημα ιδιοτιμών :

$$\lambda (S_w^{-1} S_b)$$

Ακαδημαϊκό Έτος 2018 - 2019

Εξέταση 2018-2019: Μέρος Α (Πολλαπλής)

Θέμα 1:

Πολλαπλής (40/100)

1. Ποιός αλγόριθμος ταξινόμησης είναι supervised → LDA
2. ML γενικά στην στατιστική χ διανυσμάτων → $p(\chi | \theta)$, θ σταθερό ντετερμινιστικό
3. Πολυπλοκότητα αλγορίθμου forward-backword → $O(N^2T)$
4. Τι κοινό έχουν PCA, LDA → Ο υποχώρος προκύπτει από τα ιδιοδιανύσματα κάποιου πίνακα
5. Τι δεν κάνει ο viterbi ? → υπολογισμό ακριβής τιμής πιθανότητας $P(O, Q|\lambda)$
5. Αν εφαρμόσω ICA οι συνιστώσες τι δεν κάνουνε → η προσέγγιση τους με gaussian ΔΕΝ αυξάνει. (μειώνεται η εντροπία)
6. MLP 2ο και 3ο Layers → AND και OR αντίστοιχα
7. SVM εξίσωση διαχωριστικής επιφάνειας με πυρήνα
8. στο LDA ο υποχώρος p παίρνει κάποια συγκεκριμένη τιμή (ή αυτό ρώταγε ή για την διασταση των νέων δεδομένων y) αν D η διάσταση του τωρινού και C ο αριθμός των κλάσεων → κανένα από τα παραπάνω (παραπάνω: $C-1$, C , $D/2$)
9. Σύγκλιση Perceptron → κλάσεις γραμμικά διαχωρίσιμες
10. GMM και K-means: Τι δεν έχουν κοινό?
 - α) ασυσχέτιστα δεδομένα
 - β) χρησιμοποιούν μέσες τιμές
 - γ)
 - δ)
11. Logistic regression → γραμμικός ταξινομητής , (ανήκει στην κατηγορία supervised learning) δυαδικές τιμές εξόδου ως αποτέλεσμα της σιγμοειδούς
12. Ποια η μορφή της τελικής συνάρτησης του SVM , ενός μη γραμμικά διαχωρισιμου χωρου με διανυσματα επαυξησης διαστάσεων της μορφης $\chi \rightarrow \phi(\chi)$?
 - α) $f(x) = w^T \phi(x)$
 - β) $f(x) = \sum_k w_k^T \phi(x)$
 - γ) $f(x) = w^T \phi(x) + b$
 - δ) $f(x) = \sum_k \lambda_k y_k < \phi(x_k), \phi(x) > + b$
13. Ποια η εκτίμηση της παραμέτρου a της παρακάτω κατανομής με βάση τον ML? → $N/(\sum_i^N x_i)$
 $p(x) = ae^{-ax}$, $x \geq 0$

Θέμα 2ο (36 μονάδες)

Δίνονται τα ακόλουθα δείγματα εκπαίδευσης $D_1 = \{-4, -3, -2, -7, 3, 4.2, 5.6, 3.2\}$ και $D_2 = \{-2.1, -1, 0, 2.1\}$ από τις κατηγορίες ω_1 και ω_2 αντίστοιχα.

(a) Χρησιμοποιήστε τον αλγόριθμο μεγιστοποίησης της πιθανότητας (maximum likelihood) για να υπολογίσετε τις κατανομές $p(x|\omega_1)$ και $p(x|\omega_2)$ και τις a priori πιθανότητες $p(\omega_1)$, $p(\omega_2)$. Υποθέστε ότι η κατανομή ω_1 είναι της μορφής $p(x|\omega_1) = w_{11}\mathcal{N}(\mu_{11}, 1) + w_{12}\mathcal{N}(\mu_{12}, 1)$ και η κατανομή ω_2 είναι της μορφής $\mathcal{N}(\mu_2, 1)$ όπου $\mathcal{N}(\mu, \sigma^2)$ είναι η Γκαουσιανή κατανομή

$$\mathcal{N}(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$

(b) Υπολογίστε τα διαστήματα απόφασης για τις κατηγορίες ω_1 και ω_2 σύμφωνα με τον κανόνα απόφασης του Bayes και τις κατανομές που υπολογίσατε στο (a).

(c) Χρησιμοποιώντας τον αλγόριθμο CART υπολογίστε τον δέντρο απόφασης για τις κατηγορίες ω_1 και ω_2 . Χρησιμοποιείτε ως impurity το γινόμενο $p(\omega_1)p(\omega_2)$ σε κάθε κόμβο.

(d) Υπολογίστε τα διαστήματα απόφασης για τις κατηγορίες ω_1 και ω_2 σύμφωνα με τον κανόνα απόφασης των 3 κοντινότερων γειτόνων NNR-3.

(e) Συγκρίνει (σε λύση) ο αλγόριθμος συναρτήσεων διαχωρισμού perceptron για τις κατηγορίες ω_1 και ω_2 ;

(f) Δεδομένου ότι οι πραγματικές κατανομές των δύο κατηγοριών είναι

$$p(x|\omega_1) = 0.5\mathcal{N}(-4, 1) + 0.5\mathcal{N}(4, 1) \text{ και } p(\omega_1) = \frac{2}{3}$$

$$p(x|\omega_2) = \mathcal{N}(0, 1) \text{ και } p(\omega_2) = \frac{1}{3}$$

ποιός από τους αλγόριθμους Maximum Likelihood (a), CART (c), NNR-3 (d) ελαχιστοποιεί το λάθος ταξινόμησης για το συγκεκριμένο παράδειγμα; Συγκρίνετε το λάθος ταξινόμησης Bayes με το λάθος ταξινόμησης για κάθε έναν από τους αλγορίθμους (συγκρίνοντας τα σημεία απόφασης).

(g) Έστω ένα σύνολο δειγμάτων D_t , το οποίο προκύπτει από την ένωση των συνόλων δειγμάτων D_1 και D_2 . Θεωρείστε ένα πρόβλημα ταξινόμησης δύο κατηγοριών. Εφαρμόστε τον αλγόριθμο k-means χρησιμοποιώντας ως αρχικά κέντρα τα σημεία -2 και 3.

(h) Για το πιο πάνω πρόβλημα, εφαρμόστε τον αλγόριθμο Expectation Maximization (EM) αντί του k-means. Για να προβείτε σε αρχικοποίηση των παραμέτρων του μοντέλου σας εφαρμόστε τον k-means αλγόριθμο. Διευκρίνιση: το προς εκτίμηση μοντέλο υποθέτει Gaussian κατανομές.

Θέμα 3ο (24 μονάδες)

Σε αυτό το πρόβλημα θα χρησιμοποιήσετε ένα HMM για την αποκωδικοποίηση μιας απλής ακολουθίας DNA. Είναι γνωστό ότι το DNA αναπαρίσταται ως μια σειρά από βάσεις $\{A, C, G, T\}$. Ας υποθέσουμε ότι υπάρχει μια κρυφή μεταβλητή S η οποία ελέγχει τη δημιουργία μιας ακολουθίας DNA. Η S μπορεί να πάρει 2 πιθανές καταστάσεις, $\{S_1, S_2\}$. Θεωρήστε επίσης ένα HMM M με πιθανότητες μετάβασης:

$$P(S_1|S_1) = 0.8, P(S_2|S_1) = 0.2, P(S_1|S_2) = 0.2, P(S_2|S_2) = 0.8,$$

πιθανότητες παρατηρήσεων:

$$P(A|S_1) = 0.4, P(C|S_1) = 0.1, P(G|S_1) = 0.4, P(T|S_1) = 0.1$$

$$P(A|S_2) = 0.1, P(C|S_2) = 0.4, P(G|S_2) = 0.1, P(T|S_2) = 0.4$$

και αρχικές πιθανότητες καταστάσεων:

$$P(S_1) = 0.5, P(S_2) = 0.5.$$

Έστω ότι παρατηρείτε την ακολουθία $x = CGTCAG$. Υπολογίστε:

- Την $P(x|M)$ με χρήση του αλγορίθμου forward.
- Τις ύστερες (a posteriori) πιθανότητες $P(\pi_i = S_1|x, M)$ για $i = 1, \dots, 6$.
- Το πιθανότερο μονοπάτι κρυφών καταστάσεων με χρήση του αλγορίθμου Viterbi.

Θέμα 4ο (Μπόνους 10%)

Θεωρήστε το σύνολο $D = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ \star \end{pmatrix} \right\}$ το οποίο έγινε δειγματοληψία από το 1ο τεταρτημόριο του \mathbb{R}^2 . Υποθέστε ότι οι δύο διαστάσεις είναι ανεξάρτητες. Θεωρήστε πως το \star αναπαριστά μια άγνωστη τιμή.

Επίσης από την a-priori μελέτη μας γνωρίζουμε ότι η $x_1 \sim \exp(\theta_1)$ και η $x_2 \sim U(0, \theta_2)$, δηλαδή:

$$\begin{cases} p(x_1) = \begin{cases} \frac{1}{\theta_1} e^{-x_1/\theta_1}, & x_1 \geq 0 \\ 0, & \text{αλλού} \end{cases} \\ p(x_2) = \begin{cases} \frac{1}{\theta_2}, & x_2 \in (0, \theta_2) \\ 0, & \text{αλλού} \end{cases} \end{cases}$$

Στόχος μας είναι να προσδιορίσουμε τις σταθερές με την χρήση του αλγορίθμου Expectation Maximization (EM).

- Έχοντας ως αρχικό μοντέλο το $\theta^0 := \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$ υπολογίστε αναλυτικά την συνάρτηση $Q(\theta, \theta_0)$

- E step -

- Υπολογίστε την τιμή θ η οποία μεγιστοποιεί το $Q(\theta, \theta_0)$ - M step -

Υπόδειξη: $\int e^{-x/t} \ln\left(\frac{e^{-x/m}}{mn}\right) dx = e^{-x/t} \left(\frac{t^2}{m} - t \ln\left(\frac{e^{-x/m}}{mn}\right) \right) + c_0$

Ακαδημαϊκό Έτος 2019 – 2020

Αναγνώριση Προτύπων – Θέματα Χειμερινής 2019-2020

3/2/2020

ΘΕΜΑ 1

Σωστή απάντηση = +3, Λάθος απάντηση = -1, Χωρίς απάντηση = 0

1. Στην Bayesian θεωρία αποφάσεων, αν έχουμε να αποφασίσουμε μεταξύ δύο ανισοπίθανων κατηγοριών ω_1 και ω_2 με δεδομένα \mathbf{x} και το κριτήριο είναι η ελάχιστη πιθανότητα λάθους, ποια είναι η συνθήκη του βέλτιστου κανόνα για να αποφασίσουμε υπέρ της κατηγορίας ω_1 ;
Α. $P(\omega_1) > P(\omega_2)$
Β. $P(\mathbf{x}|\omega_1) > P(\mathbf{x}|\omega_2)$
Γ. $\log P(\mathbf{x}|\omega_1) > \log P(\mathbf{x}|\omega_2)$
Δ. $P(\omega_1|\mathbf{x}) > P(\omega_2|\mathbf{x})$
2. Στη στατιστική θεωρία εκτίμησης παραμέτρων θ που καθορίζουν την pdf των δεδομένων \mathbf{x} , τι ισχύει από τα κατωτέρω για τη μέθοδο maximum Likelihood (ML);
Α. Χρειάζεται και η πρότερη πιθανότητα $P(\theta)$
Β. Οι παράμετροι θ είναι τυχαίες μεταβλητές και βρίσκονται από τη μεγιστοποίηση της $p(\mathbf{x}|\theta)p(\theta)$.
Γ. Οι παράμετροι θ είναι ντετερμινιστικές και βρίσκονται από τη θέση μεγίστου της $p(\mathbf{x}|\theta)$.
Δ. Κανένα από τα ανωτέρω.
3. Ποια από τις ακόλουθες μεθόδους διαχωρισμού δεδομένων σε clusters χρειάζεται επίβλεψη (είναι supervised);
Α. Gaussian Mixture Model (GMM)
Β. Linear Discriminant Analysis (LDA)
Γ. K-means με αρχικοποίηση των κέντρων από τα K πρώτα πρότυπα.
Δ. K-means με τυχαία αρχικοποίηση.
4. Τι από τα ακόλουθα ισχύει για την εφαρμογή του αλγόριθμου EM σε ένα μείγμα Γκαουσιανών GMM (Gaussian Mixture Model);
Α. Η επιλογή των αρχικών τιμών των παραμέτρων των Γκαουσιανών επηρεάζει τις τελικές εκτιμήσεις.
Β. Υπάρχει εγγύηση ότι ο αλγόριθμος θα συγκλίνει.
Γ. Υπάρχει εγγύηση ότι ο αλγόριθμος θα συγκλίνει σε ένα καθολικό μέγιστο.
Δ. Η εκτίμηση των παραμέτρων που προκύπτουν στο τέλος είναι η εκτίμηση μεγίστης πιθανοφάνειας.
5. Σε Hidden Markov Models (HMMs), ο forward backward αλγόριθμος για την εύρεση της πιθανότητας $P(\mathbf{O}|\lambda)$ τι πολυπλοκότητα έχει; (N = αριθμός καταστάσεων, T = μήκος ακολουθίας παρατηρήσεων)
Α. $O(NT)$
Β. $O(N^T)$
Γ. $O(N^2T)$

Δ. $O(TN \log N)$

6. Δίνονται οι ακόλουθοι ισχυρισμοί σχετικά με την Ανάλυση Κύριων Συνιστωσών (Principal Component Analysis - PCA):

(I): Είναι μέθοδος χωρίς επίβλεψη.

(II): Αναζητά αυτές τις κατευθύνσεις έτσι ώστε τα δεδομένα να έχουν τη μεγαλύτερη διασπορά.

(III): Ο μέγιστος αριθμός των κύριων συνιστωσών \leq του αρχικού αριθμού χαρακτηριστικών.

(IV): Όλες οι κύριες συνιστώσες είναι ορθογώνιες μεταξύ τους.

Ποιο από τα παρακάτω ισχύει;

A. I και II.

B. I και III.

Γ. Μόνο III.

Δ. I, II, III και IV.

7. Σε αναπαράσταση δεδομένων με τη μέθοδο Independent Component Analysis (ICA) τι από τα ακόλουθα δεν ισχύει ποτέ;

A. Αποσυσχετίζονται τα δεδομένα.

B. Αυξάνεται η προσέγγιση της κατανομής τους με Gaussian κατανομή.

Γ. Μειώνεται η προσέγγιση της κατανομής τους με Gaussian κατανομή.

Δ. Αυξάνονται οι επιμέρους εντροπίες τους.

8. Δίνονται οι ακόλουθοι ισχυρισμοί σχετικά με τις μεθόδους PCA και LDA:

(I): Και οι δύο μέθοδοι είναι τεχνικές γραμμικού μετασχηματισμού.

(II): Η LDA είναι μέθοδος με επίβλεψη ενώ η PCA χωρίς επίβλεψη.

(III): Η PCA μεγιστοποιεί τη διασπορά των δεδομένων, ενώ η LDA μεγιστοποιεί τον διαχωρισμό μεταξύ των διαφόρων κατηγοριών.

Ποιο από τα παρακάτω ισχύει;

A. I και II.

B. II και III.

Γ. Μόνο III.

Δ. I και II και III.

9. Ποιο από τα παρακάτω ισχύουν για τη μέθοδο Linear Discriminant Analysis (LDA);

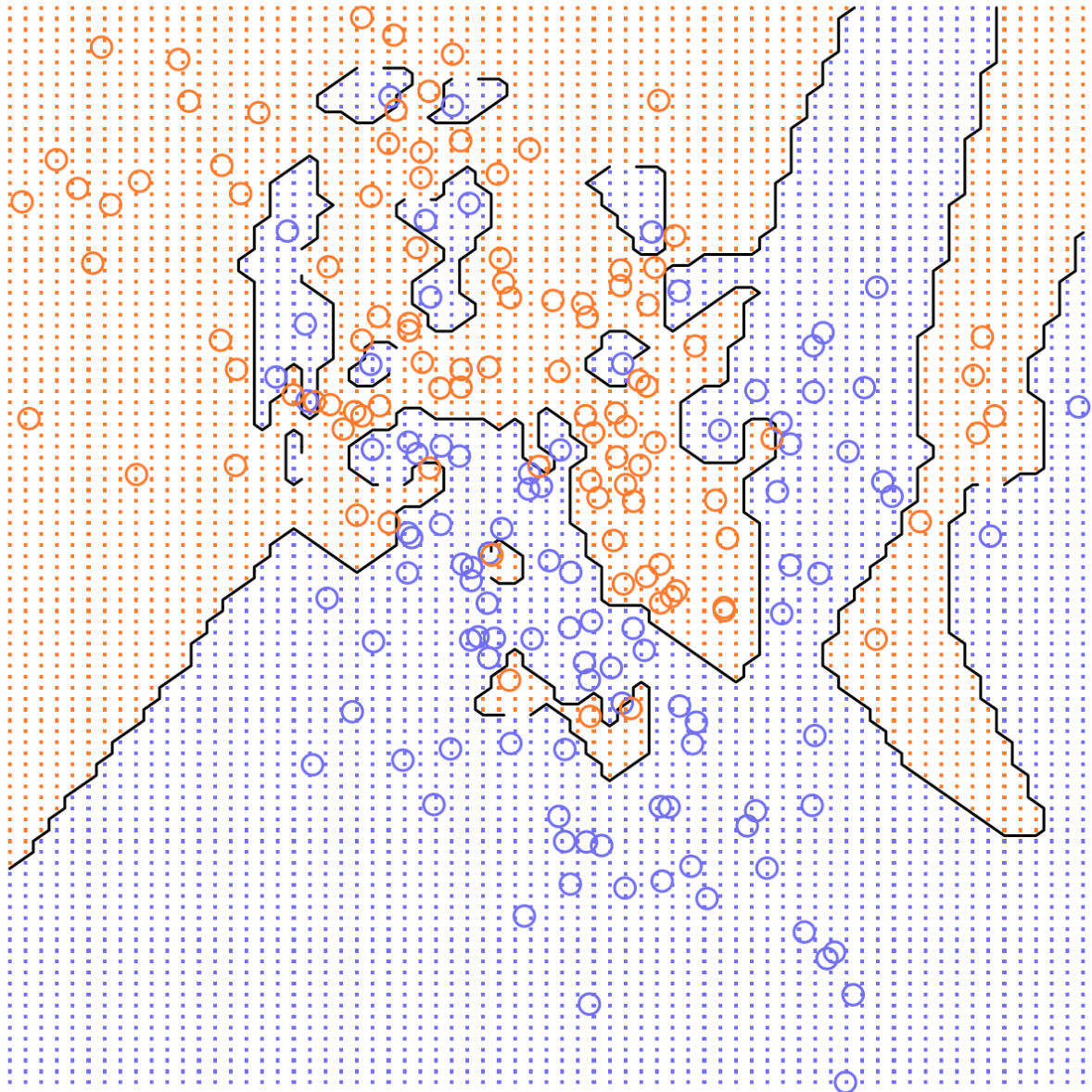
A. Μεγιστοποιεί την απόσταση μεταξύ διαφορετικών κατηγοριών και ελαχιστοποιεί την απόσταση εντός κατηγοριών.

B. Ελαχιστοποιεί τόσο την απόσταση μεταξύ διαφορετικών κατηγοριών όσο και την απόσταση εντός κατηγοριών.

Γ. Ελαχιστοποιεί την απόσταση μεταξύ διαφορετικών κατηγοριών και μεγιστοποιεί την απόσταση εντός κατηγοριών.

Δ. Μεγιστοποιεί τόσο την απόσταση μεταξύ διαφορετικών κατηγοριών όσο και την απόσταση εντός κατηγοριών.

10. Έστω ένα multi-layer perceptron με 3 επίπεδα εκτός της εισόδου. Στο 1ο κρυφό επίπεδο όλοι οι κόμβοι αποτελούνται από απλά perceptrons (με hard thresholding). Ποιος από τους ακόλουθους συνδυασμούς άλλων επιπέδων ισοδυναμεί με μια μηχανή ταξινόμησης προτύπων που μπορεί να εκμάθει περιοχές δεδομένων των οποίων η γεωμετρία στον χώρο των χαρακτηριστικών αποτελείται από αυθαίρετες πολυγωνικές περιοχές;
- A. Στο 2ο κρυφό επίπεδο οι κόμβοι αποτελούνται από απλά perceptrons τύπου Boolean OR ενώ στο επίπεδο εξόδου οι κόμβοι αποτελούνται από απλά perceptrons τύπου Boolean AND.
- B. Στο 2ο κρυφό επίπεδο οι κόμβοι αποτελούνται από απλά perceptrons τύπου Boolean AND ενώ στο επίπεδο εξόδου οι κόμβοι αποτελούνται από απλά perceptrons τύπου Boolean OR.
- Γ. Στο 2ο κρυφό επίπεδο οι κόμβοι αποτελούνται από απλά perceptrons τύπου Boolean AND ενώ στο επίπεδο εξόδου οι κόμβοι αποτελούνται από απλά perceptrons τύπου Boolean XOR.
- Δ. Στο 2ο κρυφό επίπεδο οι κόμβοι αποτελούνται από απλά perceptrons τύπου Boolean XOR ενώ στο επίπεδο εξόδου οι κόμβοι αποτελούνται από απλά perceptrons τύπου Boolean AND.
11. Έστω ότι έχουμε τα εξής δεδομένα $\{(x^{(1)}, y^{(1)}), \dots, (x^{(n)}, y^{(n)})\}$ και τη linear regression λύση ... δεδομένα $y = w_1 x + b_1$. Θεωρήστε, επίσης, και το εξής σύνολο δεδομένων $\{(x^{(1)} + \alpha, y^{(1)} + \beta), \dots, (x^{(n)} + \alpha, y^{(n)} + \beta)\}$ όπου $\alpha > 0, \beta > 0$ και $w_1 \alpha \neq \beta$. Η linear regression λύση για αυτό το σύνολο δεδομένων είναι $y = w_2 x + b_2$. Επιλέξτε ποιος από τους παρακάτω ισχυρισμούς ισχύει για τα w_1, w_2, b_1, b_2 για οποιαδήποτε επιλογή των τιμών α, β εντός των δοθέντων περιορισμών.
- A. $w_1 = w_2, b_1 = b_2$.
- B. $w_1 \neq w_2, b_1 = b_2$.
- Γ. $w_1 = w_2, b_1 \neq b_2$.
- Δ. $w_1 \neq w_2, b_1 \neq b_2$.
12. Ποια είναι η μορφή της τελικής συνάρτησης του SVM, ενός μη γραμμικά διαχωρίσιμου χώρου με διανύσματα επαύξησης διαστάσεων της μορφής $\mathbf{x} \rightarrow \phi(\mathbf{x})$;
- A. $f(\mathbf{x}) = \mathbf{w}^\top \phi(\mathbf{x})$.
- B. $f(\mathbf{x}) = \sum_k \mathbf{w}_k^\top \phi(\mathbf{x})$
- Γ. $f(\mathbf{x}) = \mathbf{w}^\top \phi(\mathbf{x}) + b$
- Δ. $f(\mathbf{x}) = \sum_k \lambda_k y_k \langle \phi(\mathbf{x}_k), \phi(\mathbf{x}) \rangle + b$
13. Υποθέστε έναν χώρο εισόδου τριών διαστάσεων $\mathbf{x} = (x_1, x_2, x_3)^\top$ και έναν μετασχηματισμό των χαρακτηριστικών που ορίζεται ως:
- $$\phi(\mathbf{x}) = (x_1^2, x_2^2, x_3^2, \sqrt{2}x_1x_2, \sqrt{2}x_1x_3, \sqrt{2}x_2x_3)$$
- Ποια είναι η αντίστοιχη kernel function $K(\mathbf{x}, \mathbf{z})$;
- A. $K(\mathbf{x}, \mathbf{z}) = (x_1x_2)^2 + (x_1x_3)^2 + (x_2x_3)^2$
- B. $K(\mathbf{x}, \mathbf{z}) = (\mathbf{x}^\top \mathbf{z})^3$
- Γ. $K(\mathbf{x}, \mathbf{z}) = (\mathbf{x}^\top \mathbf{z})^2$
- Δ. $K(\mathbf{x}, \mathbf{z}) = \mathbf{x}^\top \mathbf{z}$
14. Ποιος από τους ακόλουθους ταξινομητές μπορεί να δημιουργήσει τα επίπεδα απόφασης του σχήματος;



- A. Γραμμικό SVM.
- B. Κοντινότερος Γείτονας 1 (1-NN).
- Γ. Logistic Regression.
- Δ. Κανένας από τους παραπάνω.

ΘΕΜΑ 2

Φανταστείτε ότι εργάζεστε σε ένα εργοστάσιο σοκολάτας. Η δουλειά σας είναι να παρατηρείτε τις σοκολάτες που βγαίνουν από τη γραμμή παραγωγής. Θεωρήστε ότι οι σοκολάτες βγαίνουν μία τη φορά.

Το εργοστάσιο παράγει δύο είδη σοκολάτας, Αμυγδάλου και Απλή. Για τα υποερωτήματα (α), (β) και (γ) υποθέστε ότι γνωρίζετε με 100% βεβαιότητα τι περιέχει η κάθε σοκολάτα απλά παρατηρώντας τη.

Στο δωμάτιο ελέγχου υπάρχει ένας μοχλός που ελέγχει αν η σοκολάτα που παράγεται περιέχει αμύγδαλο ή όχι. Όταν ξεκινά η παραγωγή στην αρχή της μέρας, ο μοχλός βρίσκεται με 50% στην κατάσταση ON και με 50% στην κατάσταση OFF.

Δυστυχώς, σαν μέρος μιας φάρσας, κάποιος άφησε μια μαϊμού στο δωμάτιο ελέγχου και κλείδωσε την πόρτα. Η μαϊμού δεν μπορεί να αλλάξει θέση στον μοχλό όσο παράγεται μια σοκολάτα. Ωστόσο, ανάμεσα στην παραγωγή δύο κομματιών, υπάρχει 30% πιθανότητα η μαϊμού να αλλάξει τη θέση του διακόπτη.

(α) Σχεδιάστε το Μαρκοβιανό μοντέλο που αναπαριστά το σενάριο που περιγράψαμε. Δώστε τον αριθμό των καταστάσεων, τις πρότερες πιθανότητες για κάθε κατάσταση και τον πίνακα μεταβάσεων μεταξύ καταστάσεων.

(β) Τώρα, υποθέστε ότι υπάρχει και ένας δεύτερος μοχλός, ο οποίος ελέγχει αν η παραγόμενη σοκολάτα θα περιέχει φράουλα, έτσι ώστε να υπάρχουν συνολικά 4 τύποι σοκολάτας:

- Απλή
- Αμυγδάλου
- Φράουλα
- Αμυγδάλου + Φράουλα

Και οι δύο μοχλοί έχουν 50% πιθανότητα να είναι ανοιχτοί στην αρχή της ημέρας. Η μαϊμού βρίσκεται ακόμα στο δωμάτιο ελέγχου και αλλάζει τη θέση του κάθε μοχλού με 30% πιθανότητα. Σχεδιάστε το Μαρκοβιανό μοντέλο που περιγράφει το σύστημα, γράψτε τις πρότερες πιθανότητες κάθε κατάστασης και βρείτε τον πίνακα μεταβάσεων μεταξύ καταστάσεων.

(γ) Ποια είναι η πιθανότητα η μηχανή να παράγει τις σοκολάτες με την εξής σειρά (ακολουθία);

Απλή => Αμυγδάλου => Αμυγδάλου => Αμυγδάλου + Φράουλα => Απλή

(δ) Με βάση το μοντέλο του ερωτήματος (β), παράγονται σοκολάτες σε χρώμα Λευκό ή Μαύρο. Σε αυτό το βήμα, όμως, υποθέστε ότι δεν μπορείτε να γνωρίζετε τι ακριβώς περιέχει καθε σοκολάτα. Το μόνο που μπορείτε να καταλάβετε όταν την παρατηρείτε είναι εάν πρόκειται για Λευκή ή Μαύρη σοκολάτα. Χρησιμοποιήστε τον παρακάτω πίνακα που δίνει τις πιθανότητες $P(\text{color}|\text{filling})$. Υποθέστε, επιπλέον, ότι το χρώμα κάθε σοκολάτας είναι ανεξάρτητο από το χρώμα της επόμενης, δεδομένης της γέμισης.


	P(Λευκή)	P(Μαύρη)
Απλή	0.1	0.9
Αμυγδάλου	0.3	0.7
Φράουλα	0.8	0.2
Αμυγδάλου + Φράουλα	0.9	0.1

Αν οι δύο πρώτες σοκολάτες που παράγονται είναι Μαύρη και Λευκή αντίστοιχα, χρησιμοποιήστε τον αλγόριθμο forward/backward για να υπολογίσετε την κατανομή πιθανότητας πάνω στις καταστάσεις στις χρονικές στιγμές 1 και 2.

ΘΕΜΑ 3

Δίνονται τα ακόλουθα δείγματα εκπαίδευσης από τις κατηγορίες ω_1 και ω_2 :

$D_1 = \{0, 1.8, 2, -2.6, 7, 8, 9, 11\}$ και $D_2 = \{2.9, 4.7, 5.5, 6.9\}$

 Χρησιμοποιήστε τον αλγόριθμο μεγιστοποίησης της πιθανότητας (Maximum Likelihood) για να υπολογίσετε τις κατανομές $p(x|\omega_1)$ και $p(x|\omega_2)$ και τις a priori πιθανότητες $p(\omega_1)$ και $p(\omega_2)$. Υποθέστε ότι η κατανομή ω_1 είναι της μορφής $p(x|\omega_1) = w_{11}\mathcal{N}(x|\mu_{11}, 1) + w_{12}\mathcal{N}(x|\mu_{12}, 1)$ και η κατανομή ω_2 είναι της μορφής $\mathcal{N}(x|\mu_2, 1)$ όπου $\mathcal{N}(x|\mu, \sigma^2)$ είναι η Γκαουσιανή κατανομή.

$$\mathcal{N}(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

(β) Υπολογίστε τα διαστήματα απόφασης για τις κατηγορίες ω_1 και ω_2 , σύμφωνα με τον κανόνα απόφασης του Bayes και τις κατανομές που υπολογίσατε στο (α).

(γ) Χρησιμοποιώντας τον αλγόριθμο CART, υπολογίστε το δέντρο απόφασης για τις κατηγορίες ω_1 και ω_2 . Χρησιμοποιήστε ως impurity το γινόμενο $p(\omega_1)p(\omega_2)$ σε κάθε κόμβο.

(δ) Υπολογίστε τα διαστήματα απόφασης για τις κατηγορίες ω_1 και ω_2 , σύμφωνα με τον κανόνα απόφασης των 3 κοντινότερων γειτόνων NNR-3.

(ε) Συγκλίνει (σε λύση) ο αλγόριθμος συναρτήσεων διαχωρισμού perceptron για τις κατηγορίες ω_1 και ω_2 ;

(στ) Δεδομένου ότι οι πραγματικές κατανομές των δύο κατηγοριών είναι:

$$p(x|\omega_1) = 0.5\mathcal{N}(x|3, 1) + 0.5\mathcal{N}(x|9, 1), \quad p(\omega_1) = \frac{2}{3}$$

$$p(x|\omega_2) = \mathcal{N}(x|5, 1), \quad p(\omega_2) = \frac{1}{3}$$

Ποιος από τους αλγόριθμους Maximum Likelihood του ερωτήματος (α), CART του ερωτήματος (γ), και NNR-3 του ερωτήματος (δ) ελαχιστοποιεί το λάθος ταξινόμησης για το συγκεκριμένο παράδειγμα; Συγκρίνετε το λάθος ταξινόμησης Bayes με το λάθος ταξινόμησης για κάθε έναν από τους αλγορίθμους (συγκρίνοντας τα σημεία απόφασης).

(ζ) Έστω ένα σύνολο δειγμάτων D_t , το οποίο προκύπτει από την ένωση των συνόλων δειγμάτων D_1 και D_2 . Πιο συγκεκριμένα:

$$D_t = \{0, 1.8, 2, -2.6, 7, 8, 9, 11, 2.9, 4.7, 5.5, 6.9\}$$

Θεωρήστε ένα πρόβλημα ταξινόμησης δύο κατηγοριών. Εφαρμόστε τον αλγόριθμο k-means χρησιμοποιώντας ως αρχικά κέντρα τα σημεία 1.8 και 5.5.

(η) Για το πιο πάνω πρόβλημα (ζ), εφαρμόστε τον αλγόριθμο Expectation Minimization (EM) αντί του k-means. Για να προβείτε σε αρχικοποίηση των παραμέτρων του μοντέλου σας, εφαρμόστε τον k-means αλγόριθμο. Διευκρίνιση: το προς εκτίμηση μοντέλο υποθέτει Γκαουσιανές κατανομές.

Ακαδημαϊκό Έτος 2020 – 2021

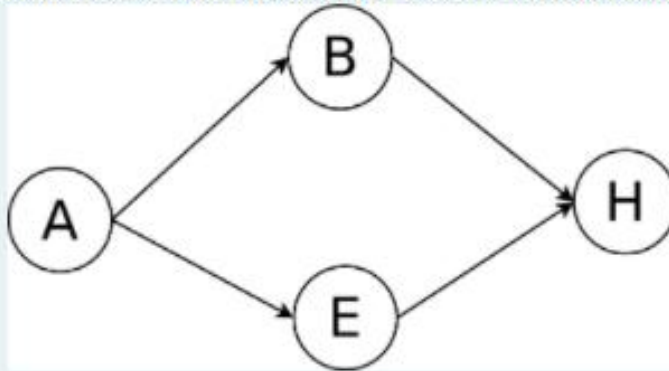
Question 1

Not yet answered

Marked out of 4.00

Flag question

Ενα Μπευσιανό δίκτυο (Bayesian net) έχει την παρακάτω τοπολογία. Ποιά από τις παρακάτω σχέσεις είναι λανθασμένη σύμφωνα με την τοπολογία αυτή



- ☐ a. $P(H|BEA) = P(H|BE)$
- ☐ b. $P(ABEH) = P(H|BE) P(B|A) P(E|A) P(A)$
- ☐ c. $P(AH|BE) = P(A|BE)P(H|BE)$
- ☐ d. $P(HB|E) = P(H|BE) P(B|E)$
- ☐ e. $P(EH|B) = P(E|B) P(H|B)$

Question **2**

Not yet
answered

Marked out of
4.00

🚩 Flag question

Δίδονται τα ακόλουθα δείγματα εκπαίδευσης για δυο κατηγορίες $D_1 = ((0, 0), (0, 1), (0, 2), (0, 3), (0, 4))$ και $D_2 = ((1, 0), (2, 0), (3, 0), (0, 0), (0, 0))$. Ποια από τις παρακάτω ερωτήσεις του ταξινομητή δέντρων απόφασης CART έχει το μικρότερο Gini impurity, δηλ. είναι η ερώτηση που διαχωρίζει καλύτερα τα δείγματα εκπαίδευσης; (Σημείωση x_1, x_2 είναι το 1ο και 2ο χαρακτηριστικό (feature) στα δείγματα εκπαίδευσης)

- ☐ a. $x_1 + x_2 < 2$
- ☐ b. $x_1 > 0$
- ☐ c. $x_2 < 1$
- ☐ d. $x_2 < 0$
- ☐ e. $x_1 > 1$

Question **3**

Not yet
answered

Marked out of
4.00

🚩 Flag question

Πόσο είναι το λάθος του ταξινομητή Bayes για το πρόβλημα ταξινόμησης μεταξύ των κατηγοριών ω_1 και ω_2 που ακολουθούν τις παρακάτω κατανομές $p(x_1, x_2|\omega_1) = 0.2, x_1 \in [0, 5), x_2 \in [0, 1)$ και $p(x_1, x_2|\omega_2) = 0.5, x_1 \in [4, 6), x_2 \in [0, 1)$, με a-priori πιθανότητες ίσες με 0.5 και για τις δυο κατανομές;

- ☐ a. 35%
- ☐ b. 0%
- ☐ c. 25%
- ☐ d. 70%
- ☐ e. 50%

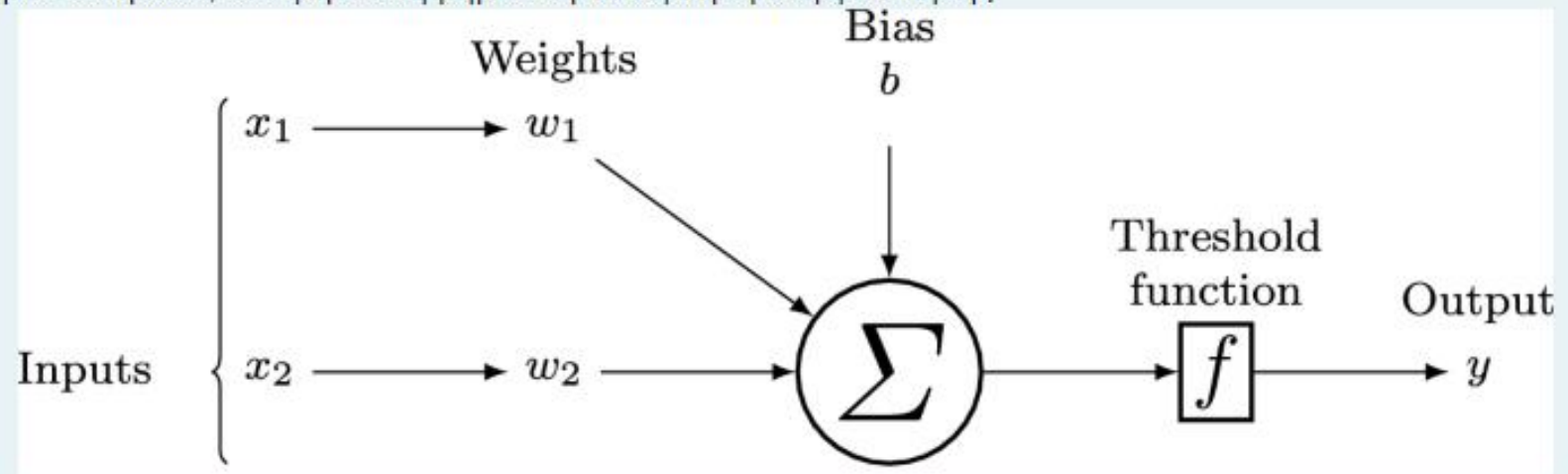
Question 4

Not yet
answered

Marked out of
4.00

Flag question

Ποιες από τις παρακάτω επιλογές των βαρών w_1 , w_2 και bias b είναι οι κατάλληλες, ώστε να σχηματίζεται μια πύλη XOR; Θεωρήστε τη βηματική συνάρτηση ενεργοποίησης.



- ☐ a. Κάποιες άλλες παράμετροι/bias, πέρα από τις επιλογές που δίνονται.
- ☐ b. $w_1 = -5.4, w_2 = -2.7, b = 6$
- ☐ c. Κανένας συνδυασμός βαρών/bias δεν μπορεί να λειτουργήσει.
- ☐ d. $w_1 = 3.2, w_2 = -5.4, b = -0.5$
- ☐ e. $w_1 = 5.4, w_2 = 2.8, b = -1.5$

Time left 0:51:45

Question **5**

Not yet
answered

Marked out of
4.00

🚩 Flag question

Δίδονται δυο κατηγορίες με δείγματα εκπαίδευσης $D_1 = (1, 1, 1, 3)$ και $D_2 = (4, 4, 4, 5)$ αντίστοιχα. Ποιό είναι το σημείο απόφασης για τον ταξινομητή κοντινότερου γείτονα με μέγεθος γειτονιάς 3 (NNR-3);

- ☐ a. 3
- ☐ b. 4
- ☐ c. 5
- ☐ d. 3.5
- ☐ e. 1

Question **6**

Not yet
answered

Marked out of
4.00

🚩 Flag question

Να βάλετε στη σωστή σειρά τα παρακάτω βήματα που απαιτούνται για την εκτέλεση του αλγόριθμου Gradient Descent για την εκπαίδευση ενός νευρωνικού δικτύου:

1. Forward pass των τιμών ενός δείγματος εισόδου και υπολογισμός της εκτιμώμενης εξόδου.
2. Αρχικοποίηση με τυχαία βάρη.
3. Υπολογισμός του σφάλματος ανάμεσα στην πραγματική τιμή και την εκτιμώμενη τιμή.
4. Επανάληψη μέχρις ότου βρεθούν τα βέλτιστα βάρη για το δίκτυο.
5. Για όλους τους νευρώνες που συνεισφέρουν στο σφάλμα, ενημέρωσε τις τιμές των βαρών, ώστε το σφάλμα να ελαττώνεται.

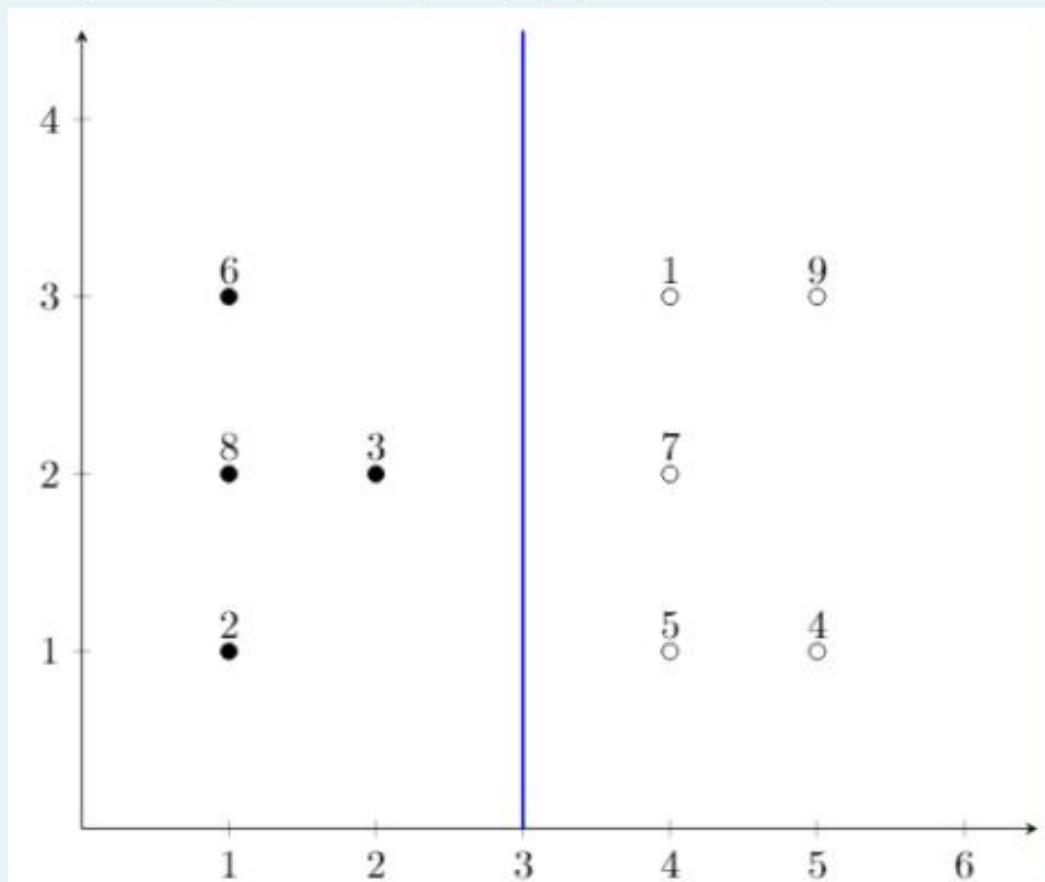
- ☐ a. 2, 1, 3, 5, 4
- ☐ b. 2, 3, 5, 1, 4
- ☐ c. Καμία από τις προτεινόμενες ακολουθίες.
- ☐ d. 1, 2, 3, 4, 5
- ☐ e. 1, 3, 2, 4, 5

Question 7

Not yet
answeredMarked out of
4.00

Flag question

Θεωρούμε ένα πρόβλημα διαχωρισμού διδιάστατων δεδομένων σε δύο κατηγορίες με SVM (γραμμικού πυρήνα) όπως αυτό της εικόνας. Τα δεδομένα είναι 9 σημεία αριθμημένα από το 1 έως το 9.



Ποιο από τα παρακάτω σημεία θα επηρεάσει την ευθεία διαχωρισμού (μπλε γραμμή) αν αφαιρεθεί από το σύνολο δεδομένων;

- ☐ a. Το σημείο 8.
- ☐ b. Καμία από τις υπόλοιπες επιλογές δεν είναι σωστή.
- ☐ c. Οποιοδήποτε σημείο και αν αφαιρεθεί.
- ☐ d. Το σημείο 7.
- ☐ e. Το σημείο 1.

Question **8**

Not yet
answered

Marked out of
4.00

🚩 Flag question

Έστω ότι οι 5 ιδιοτιμές που προκύπτουν από την ανάλυση PCA είναι 0.5, 1.15, 1.3, 1.7 και 2. Το ποσοστό της διασποράς των τιμών που «ερμηνεύεται» από τις 3 κύριες συνιστώσες βρίσκεται στο εύρος:

- ☐ a. 65%-70%
- ☐ b. Δεν αρκεί η πληροφορία της εκφώνησης για να απαντήσουμε.
- ☐ c. 73%-75%
- ☐ d. 75%-78%
- ☐ e. 70%-73%

Time left 0:50:39

Question 9

Not yet
answered

Marked out of
4.00

Flag question

Δίδεται ένα κρυφό μοντέλο Markov (HMM) με δύο καταστάσεις (Q,R) και δύο παρατηρήσεις (A,B). Όλες οι πιθανότητες μετάβασης μεταξύ καταστάσεων είναι ίσες με 0.5 και οι πιθανότητες παρατήρησης είναι $p(A|Q) = 0.3$, $p(B|R) = 0.3$. Ποιά είναι η πιο πιθανή ακολουθία καταστάσεων δεδομένων των παρατηρήσεων (B,A,B);

- ☐ a. (R,Q,R)
- ☐ b. τίποτε από τα υπόλοιπα
- ☐ c. (Q,Q,Q)
- ☐ d. (Q,R,Q)
- ☐ e. (R,R,R)

Question 10

Not yet
answeredMarked out of
4.00

Flag question

Θεωρήστε ότι έχουμε τα εξής τρία σημεία $x_1 = (1, 1)$, $x_2 = (-1, 3)$, $x_3 = (2, 4)$ με τις τιμές που τους αντιστοιχούν $z_1 = 5$, $z_2 = 11$, $z_3 = 18$. Επιθυμούμε να κάνουμε linear regression βρίσκοντας το βέλτιστο επίπεδο (σε 3 διαστάσεις) της μορφής $z = w_1 x + w_2 y$, το οποίο να «προσεγγίζει» τα σημεία αυτά με τον βέλτιστο τρόπο, δηλαδή επιθυμούμε να βρούμε το $w = [w_1 \ w_2]^T$ που ελαχιστοποιεί το σφάλμα $E = \sum_{i=1}^3 (w^T x_i - z_i)^2$.

- ☐ a. Κανένα από τα υπόλοιπα.
- ☐ b. Το βέλτιστο w το οποίο προκύπτει δεν μηδενίζει το σφάλμα E , κάτι που θα συνέβαινε εάν είχαμε θεωρήσει γενικευμένα επίπεδα της μορφής $z = w_1 x + w_2 y + w_3$.
- ☐ c. Το βέλτιστο w το οποίο προκύπτει δεν μηδενίζει το σφάλμα E . Αυτό συμβαίνει επειδή χρειαζόμαστε τουλάχιστον 4 σημεία.
- ☐ d. Δεν μπορεί πάντα να βρεθεί κάποιο w για το οποίο να μηδενίζεται το σφάλμα E , αλλά αυτό συμβαίνει στη συγκεκριμένη περίπτωση αυτών των σημείων.
- ☐ e. Μπορεί πάντα να βρεθεί κάποιο w για το οποίο μηδενίζεται το σφάλμα E , αφού δίνονται 3 σημεία και τα 3 σημεία ορίζουν μοναδικά ένα επίπεδο.

Question 11

Not yet
answeredMarked out of
4.00

Flag question

Τρέξτε τον αλγόριθμο k-means πάνω στα ακόλουθα δεδομένα εκπαίδευσης: $D = (1, 2, 2, 3, 5, 6, 6, 7)$. Εφόσον το πλήθος των κατηγοριών είναι 2 και η αρχικοποίηση για τις δύο κατηγορίες είναι 0 και 8 σε πόσα iterations θα συγκλίνει ο αλγόριθμος k-means.

- ☐ a. 10
- ☐ b. 5
- ☐ c. δεν συγκλίνει ποτέ
- ☐ d. 1
- ☐ e. 2

Question 12

Not yet
answeredMarked out of
4.00

Flag question

Θεωρήστε ένα πρόβλημα ταξινόμησης δύο κλάσεων στον τρισδιάστατο χώρο, στον οποίο οι 2 κλάσεις είναι γραμμικά διαχωρίσιμες. Το βέλτιστο διαχωριστικό υπερεπίπεδο, δηλαδή αυτό με το μεγαλύτερο margin δίνεται από τη σχέση $x + y - 2z - 1 = 0$. Ποιο/α από τα παρακάτω σημεία αποτελούν support vectors;

- ☐ a. Το $(1, -1, 1)$
- ☐ b. Το $(2, -1, -1)$
- ☐ c. Το $(2, -1, 1/2)$
- ☐ d. Κανένα από τα υπόλοιπα σημεία
- ☐ e. Το $(1, -1, -1/2)$

Information

Flag question

Το μέρος αυτό αποτελείται από 3 ερωτήσεις για τις οποίες θα στείλετε την αναλυτική επίλυσή τους χειρόγραφα, την οποία θα κληθείτε να υποβάλετε με **ένα ενιαίο αρχείο pdf** στο μέρος Γ που θα περιλαμβάνει τις λύσεις σας για όλα τα θέματα του Μέρους Β της εξέτασης.

Προσοχή: Το μέγεθος του αρχείου **δεν μπορεί να υπερβαίνει τα 16 MB**.

Ακόμα και αν καταγράψετε τις αναλυτικές απαντήσεις σας online στο αντίστοιχο πλαίσιο κειμένου, και πάλι θα πρέπει να υποβάλετε ένα ενιαίο αρχείο pdf.

Πλήρης βαθμολόγηση κατοχυρώνεται με ορθές και αναλυτικές απαντήσεις και στα 3 θέματα του Μέρους Β.

Η κατανομή των βαθμολογιών έχει ως εξής:

- **Θέμα 1 -> 25%**
- **Θέμα 2 -> 7%**
- **Θέμα 3 -> 20%**

Question 1

Not yet answered

Marked out of 25.00

Flag question



Θεωρείστε σε μια ακολουθία φωνημάτων την μοντελοποίηση της εναλλαγής άφωνων (U=unvoiced), έμφωνων (V=voiced) ήχων και σιωπής (S=silence) με ένα HMM μοντέλο (παραμέτρων λ) 4 καταστάσεων με τις εξής πιθανότητες

	<i>State1</i>	<i>State2</i>	<i>State3</i>	<i>State4</i>
$P(V)$	0.2	0.3	0.2	0.4
$P(U)$	0.5	0.3	0.4	0.3
$P(S)$	0.3	0.4	0.4	0.3

Υποθέτουμε τις ακόλουθες πιθανότητες μετάβασης καταστάσεων:

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{24} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \alpha_{34} \\ \alpha_{41} & \alpha_{42} & \alpha_{43} & \alpha_{44} \end{pmatrix} = \begin{pmatrix} 0.3 & 0.2 & 0.3 & 0.2 \\ 0.1 & 0.5 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.4 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.4 \end{pmatrix}$$

και ίσες πιθανότητες αρχικής κατάστασης:

$$\pi_i = 0.25, \quad i = 1, 2, 3, 4.$$

Παρατηρούμε την ακολουθία

$$\mathbf{O} = (O_1, O_2, \dots, O_5) = (V, U, V, U, S)$$

(a) Να υπολογισθούν οι πιθανότητες

$$\delta_t(i) = \max_{q_1, q_2, \dots, q_{t-1}} P[q_1 q_2 \dots q_{t-1}, q_t = i, O_1 O_2 \dots O_t | \lambda], \quad i = 1, \dots, 4, \quad t = 1, \dots, 5.$$

(b) Να βρεθεί η πιο πιθανή ακολουθία καταστάσεων $\mathbf{Q}^* = (q_1, q_2, \dots, q_5)$.

Για το (a) και (b) χρησιμοποιείτε τον αλγόριθμο Viterbi.

(c) Να υπολογισθεί η πιθανότητα $P^* = P(\mathbf{O}, \mathbf{Q}^* | \lambda)$.

Εξηγήστε αναλυτικά την εργασία σας.

Question 2

Not yet answered

Marked out of 7.00

Flag question

Time left 0:34:56

Θεωρήστε την παρακάτω εικόνα εισόδου διαστάσεων 7×7 .

7	3	1	5	2	0	1
6	0	1	2	0	3	4
4	2	0	4	1	4	3
5	7	3	0	1	4	6
4	1	2	2	4	3	1
0	5	2	3	2	5	7
1	6	3	4	2	0	4

Σε ένα νευρωνικό δίκτυο, ποια θα είναι η έξοδος που θα προκύψει έπειτα από την εφαρμογή ενός σταδίου max pooling διαστάσεων 3×3 με stride ίσο με 2; Συμπληρώστε τα σωστά αποτελέσματα στον πίνακα που ακολουθεί.

<input type="text"/>	<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>	<input type="text"/>

Question 3

Not yet answered

Marked out of 20.00

Flag question

Δίνονται δύο (a-priori) ισοπίθανες δισδιάστατες κανονικές Gaussian κατανομές με διασπορά $\sigma_1^2 = \sigma_2^2 = 2$, και μέση τιμή $\mu_1 = (2, 1)$ και $\mu_2 = (4, 1)$, αντίστοιχα.

1. Ποιά είναι η επιφάνεια απόφασης (Bayes decision boundary);
2. Ποιό είναι το σφάλμα ταξινόμησης Bayes (Bayes error) ως συνάρτηση του $\frac{1}{\sqrt{2\pi}} \int_c^\infty \exp(-\frac{x^2}{2}) dx$;
3. Πόσο είναι το σφάλμα ταξινόμησης Bayes, μετά από προβολή των δύο κατανομών στον άξονα x (δηλαδή ως προς τη μία διάσταση);
4. Πόσο είναι το σφάλμα ταξινόμησης Bayes, μετά από προβολή των δύο κατανομών στον άξονα y (δηλαδή ως προς την άλλη διάσταση);