

- 1) Cosa si intende per iperparametri? Fornire esempi pratici di iperparametri. Come si ottimizzano?
- 2) Con quali tecniche si può estendere SVM da 2 a più classi?
- 3) Cosa si intende per Clustering esclusivo e Clustering soft (o Fuzzy). Quest'ultimo che vantaggi può avere?
- 4) Che cosa codifica la funzione Q nell'ambito dell'approccio Q-learning? Quali sono gli input che ne determinano il valore?

5) Date due distribuzioni multinormali identificate dai seguenti parametri:

$$\mu_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\Sigma_0 = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix}$$

$$P(w_0) = 0,5$$

$$\mu_1 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

$$\Sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix}$$

$$P(w_1) = 0,5$$

Nell'ipotesi dell'impiego di un classificatore di Bayes multinormale, calcolare per il punto $x = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$:

- le densità di probabilità condizionali;
- l'indice della classe restituita in output.

Si ricorda che:

- la densità di probabilità, nel caso della distribuzione multinormale è:

$$p(x) = \frac{1}{(2\pi)^{\frac{d}{2}} \cdot |\Sigma|^{\frac{1}{2}}} \cdot e^{-\frac{1}{2} \cdot (x-\mu)^t \cdot \Sigma^{-1} \cdot (x-\mu)}$$

- l'inversa di una matrice diagonale si ottiene invertendo i singoli elementi;
- il determinante di una matrice diagonale si ottiene moltiplicando gli elementi della diagonale.

6) Un classificatore *Nearest Neighbor* (NN), con un *training set* (TS) composto da $n = 1000$ pattern di dimensionalità $d = 5$, utilizza come metrica la *distanza euclidea*. Calcolare il numero di somme, sottrazioni e moltiplicazioni necessarie (trascurando la radice quadrata) per effettuare la classificazione di un pattern x supponendo che non vengano utilizzate strutture spaziali specifiche per indicizzare il TS.

7) Dati un volume di input e un livello di convoluzione in una CNN, aventi le seguenti caratteristiche:

- *Volume Input*: $3 \times 256 \times 256$
- 64 filtri di dimensione $3 \times 7 \times 7$
- *Stride*: 3
- *Padding*: 0

Si calcoli motivando la risposta:

- La dimensione del volume di output: $F \times W \times H$;
- il numero totale di connessioni e di pesi del livello (senza considerare i bias).

