

1) Quali sono le condizioni necessarie affinché le tecniche di deep learning siano più efficaci di altri approcci di machine learning?

2) La formula di distanza di un pattern dall'iperpiano risultante dal training di un SVM dipende da tutti i pattern del training set o solo da una parte di questi? Motivare la risposta.

3) Descrivere a grandi linee l'algoritmo di Clustering K-means.

4) Qual'è la funzione obiettivo in formato matriciale della multiple linear regression?

5) Data una rete neurale MLP a 3 livelli senza bias composta da:

- 5 neuroni per l'input layer
- 10 neuroni per l'hidden layer
- 4 neuroni di output

Quante somme e moltiplicazioni sono necessarie per il passo forward di un generico pattern trascurando le operazioni effettuate dalla funzione di attivazione? Motivare la risposta riportando anche il numero di operazioni per livello.

6) Date tre distribuzioni multinormali identificate dai seguenti parametri:

$$\boldsymbol{\mu}_1 = \begin{bmatrix} 9,40 \\ 2,18 \\ 6,79 \\ 3,52 \end{bmatrix} \quad \boldsymbol{\mu}_2 = \begin{bmatrix} 5,22 \\ 8,15 \\ 6,82 \\ 2,94 \end{bmatrix} \quad \boldsymbol{\mu}_3 = \begin{bmatrix} 5,66 \\ 0,80 \\ 7,51 \\ 8,10 \end{bmatrix}$$

$\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_3 = I$  (matrice identità) e  $P(w_1) = P(w_2) = P(w_3)$ .

Indicare la classe assegnata ai seguenti pattern da un classificatore di Bayes multinormale (motivandone la risposta):

$$\mathbf{p}_1 = \begin{bmatrix} 7,15 \\ 3,87 \\ 3,56 \\ 2,50 \end{bmatrix}, \mathbf{p}_2 = \begin{bmatrix} 3,21 \\ 5,62 \\ 6,40 \\ 1,29 \end{bmatrix}, \mathbf{p}_3 = \begin{bmatrix} 7,10 \\ 2,05 \\ 0,69 \\ 1,27 \end{bmatrix}$$

7) Supponendo di utilizzare *K-fold Cross-Validation* con  $K = 5$  per suddividere 8500 pattern in *training* e *validation set*, quanti diversi addestramenti (*run*) vengono effettuati? Ad ogni *run* quanti pattern vengono utilizzati per il training e quanti per la validazione?