

- 1) Cosa è possibile apprendere mediante tecniche di reinforcement learning? Fare un esempio.
 - 2) Nell'ambito di classificazione con SVM cosa si intende per pattern linearmente separabili e non linearmente separabili? Fare esempio grafico dei due casi.
 - 3) Qual è l'idea di base dell'algoritmo di clustering EM con Gaussian mixture?
 - 4) Cosa si intende per iperparametri? Fornire esempi pratici di iperparametri. Come si ottimizzano?
-

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

- 16 neuroni per l'input layer
- 32 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 due distribuzioni multinormali identificate dai seguenti parametri:

$$\boldsymbol{\mu}_1 = \begin{bmatrix} 3,8 \\ -2,9 \\ -0,7 \\ 2,8 \end{bmatrix} \quad \boldsymbol{\mu}_2 = \begin{bmatrix} -2,4 \\ 0,5 \\ 2,5 \\ 3,8 \end{bmatrix}$$

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

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

$$\mathbf{p}_1 = \begin{bmatrix} -4,1 \\ 2,1 \\ 3,9 \\ 2,4 \end{bmatrix}, \mathbf{p}_2 = \begin{bmatrix} 3,4 \\ -2,6 \\ 3,1 \\ 0,3 \end{bmatrix}, \mathbf{p}_3 = \begin{bmatrix} 2,0 \\ 0,7 \\ 1,0 \\ 0,6 \end{bmatrix}$$

7) Dato un livello di convoluzione in una CNN con un volume di Input pari a $128 \times 224 \times 16$ (nel formato $Width \times Height \times Depth$), e filtri di dimensioni $5 \times 5 \times 16$. Si calcolino le dimensioni ($Width \times Height$) di ogni *feature map* nel volume di Output considerando un $Padding = 1$ e $Stride = 3$.