

1) Cosa si intende per Clustering esclusivo e Clustering soft (o Fuzzy). Quest'ultimo che vantaggi può avere?

Pag. 6 delle dispense "Clustering"

2) Cosa è possibile apprendere mediante tecniche di reinforcement learning? Fare un esempio.

Pag. 21-26 delle dispense "Deep Learning"

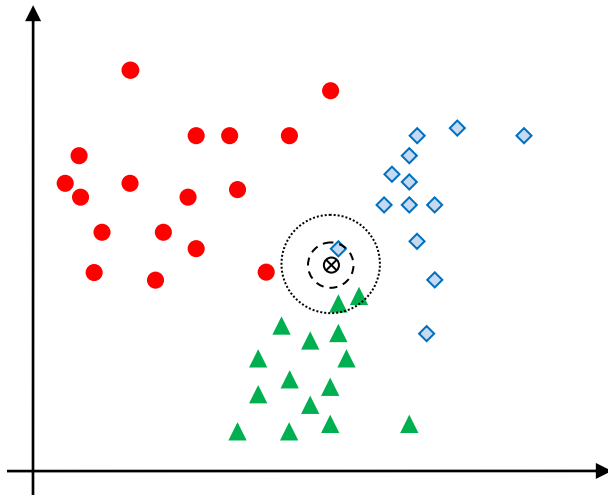
3) Definire cosa si intende per apprendimento supervisionato e non supervisionato.

Pag. 10 delle dispense “Fondamenti”

4) Nelle SVM non lineari cosa si intende per kernel? Quali sono i kernel più utilizzati?

Pag. 11-12 delle dispense “Classificazione (2)”

5) Date le distribuzioni riportate nel grafico sottostante, indicare come viene classificato il pattern  $\otimes$  da due classificatori  $k$ -NN (con  $k$  uguale a 1 e 3) supponendo di utilizzare come metrica la distanza euclidea. Motivare la risposta.



**Svolgimento**

La regola  $k$ -NN determina i  $k$  elementi più vicini al pattern da classificare (in base alla metrica utilizzata). Ognuno dei  $k$  elementi vota per la classe a cui appartiene e il pattern da classificare viene assegnato alla classe che ha ottenuto il maggior numero di voti. Pertanto, al pattern  $\otimes$  verrà assegnata la classe “rombo” nel caso di  $k=1$  (intorno - - -) e la classe “triangolo” nel caso di  $k=3$  (intorno · · ·).

6) Dato un training set composto dai seguenti pattern:

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 3 \\ 1 \\ 6 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 12 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{x}_4 = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}, \quad \mathbf{x}_5 = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$$

a cui sono associate le seguenti osservazioni (variabile dipendente):

$$y_1 = 10, \quad y_2 = 12, \quad y_3 = 15, \quad y_4 = 14, \quad y_5 = 16.$$

formulare il problema di *multiple linear regression* definendo la matrice  $\mathbf{X}$  e il vettore  $\mathbf{y}$ .

**Svolgimento**

La matrice  $\mathbf{X}$  di dimensione  $n \times (d + 1)$  si ottiene disponendo su ciascuna riga gli  $n$  pattern del training set e inserendo a destra una colonna con tutti valori 1:

$$\mathbf{X} = \begin{bmatrix} 1 & 4 & 2 & 1 \\ 3 & 1 & 6 & 1 \\ 12 & 1 & 1 & 1 \\ 1 & 2 & 4 & 1 \\ 0 & 1 & 3 & 1 \end{bmatrix}$$

Il vettore  $\mathbf{y}$  di dimensione  $n$  si ottiene disponendo in colonna i valori  $y_i$ :

$$\mathbf{y} = \begin{bmatrix} 10 \\ 12 \\ 15 \\ 14 \\ 16 \end{bmatrix}$$

Anche se non richiesto dall'esercizio, definita la matrice  $\mathbf{X}$  e il vettore  $\mathbf{y}$  le relazioni tra variabili indipendenti e dipendente possono essere scritte in forma matriciale come:

$$\mathbf{y} = \mathbf{X} \cdot \boldsymbol{\beta}$$

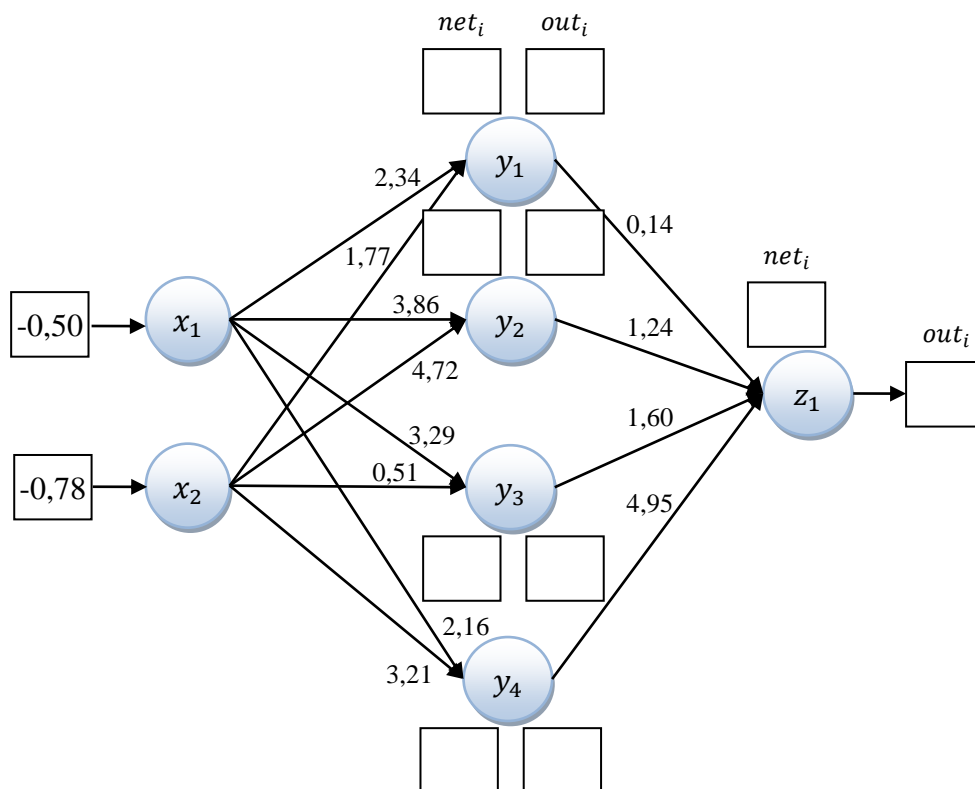
Dove  $\boldsymbol{\beta}$  rappresenta il vettore di dimensione  $d + 1$  dei parametri da determinare:

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix}$$

In formato matriciale la funzione obiettivo da minimizzare (*Least Square*) può essere scritta come:

$$\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2$$

7) Data la seguente rete neurale, calcolare  $net_i$  e  $out_i$  di ogni neurone al seguito del passo forward propagation del pattern di input, utilizzando la *standard logistic function* ( $\frac{1}{1+e^{-net}}$ ) come funzione di attivazione.



## Svolgimento

$$net(y_1) = 2,34 \cdot (-0,50) + 1,77 \cdot (-0,78) = -2,55$$

$$net(y_2) = 3,86 \cdot (-0,50) + 4,72 \cdot (-0,78) = -5,61$$

$$net(y_3) = 3,29 \cdot (-0,50) + 0,51 \cdot (-0,78) = -2,04$$

$$net(y_4) = 2,16 \cdot (-0,50) + 3,21 \cdot (-0,78) = -3,58$$

$$out(y_1) = 0,07$$

$$out(y_2) = 0,00$$

$$out(y_3) = 0,12$$

$$out(y_4) = 0,03$$

$$net(z_1) = 0,14 \cdot 0,07 + 1,24 \cdot 0,00 + 1,60 \cdot 0,12 + 4,95 \cdot 0,03 = 0,35$$

$$out(z_1) = 0,59$$