Machine Learning	**Matricola: ____________________**
10-Set-2019 (90 minuti)	**Cognome: __________ Nome: ____________**

1) Nell'ambito dei multi-classificatori quali sono le più comuni tecniche di fusione a livello di decisione e di confidenza?

Dispense "Classificazione (2)"

2) Nell'ambito di CNN, che cosa si intende con transfer learning? Quali sono le tecniche di transfer learning utilizzabili?

Dispense "Deep Learning (1)"

3) Dare la definizione di Training, Validation e Test set e discutere una possibile suddivisione dei dati nei tre insiemi.

Dispense “Fondamenti”

4) Nella regressione cosa si intende per variabile indipendente e variabile dipendente?

Dispense “Regressione”

Machine Learning **Matricola: ____________________**

10-Set-2019 (90 minuti) **Cognome: __________ Nome: ____________**

5) Dato un insieme di pattern tri-dimensionali composto da 4 elementi:

$$\left\{\begin{bmatrix}-0,5\\-9,1\\-0,9\end{bmatrix},\begin{bmatrix}8,1\\-6,6\\-6,6\end{bmatrix},\begin{bmatrix}6,7\\9,2\\-3,2\end{bmatrix},\begin{bmatrix}-3,4\\-2,1\\-8,9\end{bmatrix}\right\}$$

Calcolare il vettore medio ($\boldsymbol{\mu}$) e la matrice di covarianza ($\boldsymbol{\Sigma} = [\sigma_{ij}]$).

Si ricorda che ogni elemento della matrice di covarianza può essere calcolato come
$\sigma_{ij} = \sigma_{ji} = \frac{1}{n}\sum_{k=1}^{n}(\mathrm{x}_{ki} - \mu_i) \cdot (\mathrm{x}_{kj} - \mu_j)$
dove x_{km} è l'm-esimo elemento del k-esimo pattern, e n il numero di pattern.

Svolgimento

$$\boldsymbol{\mu} = \begin{bmatrix}2,73\\-2,15\\-4,90\end{bmatrix}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix}23,15 & 10,83 & 2,31\\10,83 & 49,23 & -0,28\\2,31 & -0,28 & 9,45\end{bmatrix}$$

6) Data una rete neurale MLP a 3 livelli con bias composta da:

- 3 neuroni per l'input layer
- 12 neuroni per l'hidden layer
- 2 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.

Svolgimento

Per ogni neurone del livello corrente si deve calcolare la seguente formula:

$$net_i = \sum_{j=1..d} w_{ji} \cdot in_j + w_{0i}$$

che comprende una moltiplicazione e una somma per ogni neurone del livello precedente più la somma finale del bias. Pertanto:

Neuroni livello corrente

Una moltiplicazione per ogni neurone livello precedente

Una somma per ogni neurone livello precedente

Somma Bias

Numero operazioni livello hidden: $12 \cdot (3 + 3 + 1) = 84$

Numero operazioni livello di output: $2 \cdot (12 + 12 + 1) = 50$

NOTA: Il calcolo a lato è eseguito trascurando il fatto che un'implementazione ottimizzata della sommatoria potrebbe evitare una somma per il primo elemento della sommatoria (somma con 0).

Totale: **134 (74 somme e 60 moltiplicazioni)**

7) Dato un insieme di pattern tri-dimensionali composto da 5 elementi:

$$\mathbf{p_1} = \begin{bmatrix} 2{,}5 \\ 1{,}1 \\ 4{,}7 \end{bmatrix}, \mathbf{p_2} = \begin{bmatrix} 7{,}7 \\ 6{,}6 \\ 4{,}3 \end{bmatrix}, \mathbf{p_3} = \begin{bmatrix} 3{,}5 \\ 9{,}4 \\ 1{,}0 \end{bmatrix}, \mathbf{p_4} = \begin{bmatrix} 6{,}4 \\ 0{,}3 \\ 2{,}5 \end{bmatrix}, \mathbf{p_5} = \begin{bmatrix} 3{,}2 \\ 9{,}9 \\ 6{,}8 \end{bmatrix}$$

Effettuare la prima iterazione dell'algoritmo K-means supponendo di dover raggruppare i pattern in 3 cluster rappresentati dai seguenti centroidi:

$$\mathbf{c_1} = \begin{bmatrix} 6{,}5 \\ 2{,}8 \\ 6{,}2 \end{bmatrix}, \mathbf{c_2} = \begin{bmatrix} 7{,}0 \\ 7{,}0 \\ 9{,}5 \end{bmatrix}, \mathbf{c_3} = \begin{bmatrix} 0{,}9 \\ 1{,}6 \\ 3{,}8 \end{bmatrix}$$

Riportare il cluster di appartenenza di ogni pattern e le coordinate dei nuovi centroidi calcolate in seguito all'iterazione svolta.

<u>Svolgimento</u>

Un'iterazione del K-means consiste i) nell'assegnare ogni pattern al cluster per cui è minima la distanza dal corrispondente centroide, e ii) nell'aggiornare i centroidi come media dei pattern assegnati al cluster corrispondente.
Utilizzando la distanza euclidea (si può utilizzare la distanza euclidea al quadrato evitando di calcolare la radice quadrata) i pattern vengono attribuiti ai cluster come segue:

$$Cluster_1 = \{\mathbf{p_2}, \mathbf{p_4}\} \quad Cluster_2 = \{\mathbf{p_5}\} \qquad Cluster_3 = \{\mathbf{p_1}, \mathbf{p_3}\}$$

Di conseguenza, i nuovi centroidi saranno:

$$\mathbf{c_1}' = \begin{bmatrix} 7{,}1 \\ 3{,}5 \\ 3{,}4 \end{bmatrix}, \mathbf{c_2}' = \begin{bmatrix} 3{,}2 \\ 9{,}9 \\ 6{,}8 \end{bmatrix}, \mathbf{c_3}' = \begin{bmatrix} 3{,}0 \\ 5{,}3 \\ 2{,}9 \end{bmatrix}$$