

- 1) Come classifica i pattern un classificatore K-NN? Discutere la complessità computazionale.
- 2) Nel classificatore SVM cosa sono i support vectors?
- 3) Cosa si intende per convergenza di un algoritmo di apprendimento iterativo? Accuratezza e loss come si comportano durante le iterazioni in caso di convergenza? Disegnare un semplice grafico.
- 4) Nella regressione cosa si intende per variabile indipendente e variabile dipendente?

5) In una rete CNN, data un'immagine di Input di dimensione  $7 \times 7 \times 3$  (nel formato *Width*  $\times$  *Height*  $\times$  *Depth*) e un livello di convoluzione composto da 1 filtro di dimensioni  $3 \times 3 \times 3$  con *padding* = 0 e *stride* = 2, si calcoli il valore dell'elemento del volume di output indicato con il ?.

**Input**

**Depth 0**

197	103	42	252	27	78	205
114	57	2	195	7	1	130
97	71	179	60	187	22	21
86	84	187	229	208	167	237
25	177	236	250	25	9	87
217	175	190	175	23	10	69
67	127	246	142	4	125	87

**Depth 1**

124	164	158	18	229	152	110
19	111	22	75	167	224	88
136	21	201	237	248	43	136
151	245	140	163	12	207	19
212	197	87	203	42	149	157
12	78	232	52	113	232	198
64	167	99	112	42	236	186

**Depth 2**

105	45	160	12	81	207	228
174	18	111	216	200	91	62
170	191	128	124	74	187	123
199	224	184	134	66	193	87
77	41	50	226	226	88	106
151	182	191	216	198	184	93
60	120	91	168	141	136	150

**Filtro**

**Depth 0**

0.89	0.87	0
0	0.30	0.52
0	0	0

**Depth 1**

0	0.24	0.90
0	0.07	0
0.64	0	0

**Depth 2**

0	0	0
0	0.71	0
0.40	0	0

**Output**

-	-	?
-	-	-
-	-	-

6) Date due distribuzioni multinormali identificate dai seguenti parametri:

$$\boldsymbol{\mu}_0 = \begin{bmatrix} 3.21 \\ -0.57 \end{bmatrix}$$

$$\boldsymbol{\mu}_1 = \begin{bmatrix} 1.27 \\ -2.14 \end{bmatrix}$$

$$\boldsymbol{\Sigma}_0 = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/3 \end{bmatrix}$$

$$\boldsymbol{\Sigma}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1/4 \end{bmatrix}$$

$$P(w_0) = 0.7$$

$$P(w_1) = 0.3$$

Nell'ipotesi dell'impiego di un classificatore di Bayes multinormale, calcolare per il punto  $\boldsymbol{x} = \begin{bmatrix} 3.80 \\ -2.40 \end{bmatrix}$ :

- 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:

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

- 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.

7) Data la seguente porzione di rete neurale (per un problema di regressione), durante il passo forward un pattern ha prodotto le attivazioni  $net_1$  e  $net_2$  indicate in figura. Motivando le risposte si richiede di:

1. Completare il passo forward calcolando  $out_1$ ,  $out_2$ ,  $out_3$  (i neuroni 1 e 2 hanno funzione di attivazione RELU, mentre il neurone 3 non ha funzione di attivazione).
2. Calcolare la loss MSE considerando che il valore atteso per il pattern è  $t = 2$ .
3. Calcolare i nuovi valori che i pesi  $w_1$  e  $w_2$  assumeranno a seguito del passo backward con un learning rate ( $\eta$ ) di 0.1. Si ricorda che il  $\delta_3$  (errore) sul neurone 3 in caso di loss MSE senza funzione di attivazione è uguale alla differenza tra valore atteso e attivazione del neurone ( $t - out_3$ ) e pertanto il gradiente rispetto al peso  $w_i$  corrisponde a  $-(\delta_3 \cdot out_i)$ .

