

1) Per l'addestramento di una rete neurale che cosa si intende con vettore di output desiderato? Come può essere definito? Come si può calcolare l'errore da retro-propagare a partire dal vettore desiderato e dal valore calcolato dalla rete per un pattern?

Dispense "Reti Neurali"

2) Fare esempi pratici di ragionamento induttivo e deduttivo.

Dispense "Introduzione"

3) Nel classificatore di Bayes cosa si intende per densità di probabilità condizionale e probabilità a priori.

Dispense “Classificazione (1)”

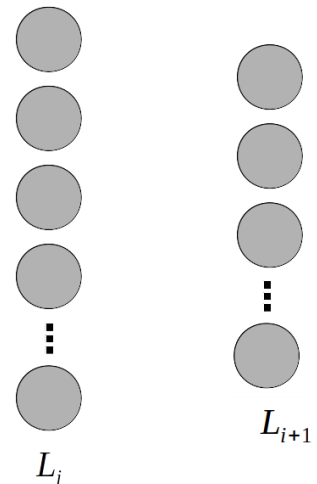
4) Come si misurano le prestazioni di un classificatore?

Dispense “Fondamenti”

5) Dati due livelli di una rete neurale L_i e L_{i+1} costituiti rispettivamente da 12 e 9 neuroni, indicare:

1. Il numero di connessioni
2. Il numero di pesi distinti

sia nel caso i due livelli costituiscano una porzione di una rete MLP, sia nel caso essi appartengano a una CNN dove ogni neurone del livello $i + 1$ è connesso a 5 neuroni del livello i (*receptive field* = 5). Motivare infine la risposta.



Svolgimento

Nel caso in cui i livelli siano “completamente connessi” come accade nelle reti MLP il numero di connessioni è pari a $|L_i| \cdot |L_{i+1}|$ quindi $12 \cdot 9 = 108$ connessioni totali.

Nel caso in cui i livelli siano di una CNN con “receptive field” uguale a 5 allora basta moltiplicare 5 per il numero di neuroni del livello $i + 1$ quindi $5 \cdot 9 = 45$ connessioni totali.

Per quanto riguarda i pesi, nel caso della rete MLP il numero degli stessi è pari al numero di connessioni (quindi 108) mentre per la CNN è pari alla dimensione del *receptive field* (5 - essendo i pesi condivisi).

6) Un multiclassificatore, composto da 3 classificatori combinati a livello di confidenza, viene utilizzato per riconoscere pattern appartenenti a 4 classi (A, B, C, D). Nella tabella seguente sono riportate le confidenze restituite dai singoli classificatori (C_i) dati in input 2 diversi pattern (p_j). Completare la tabella riportando, per ogni metodo di fusione (Somma, Prodotto, Massimo e Minimo), le confidenze ottenute e la classe di output restituita dal multiclassificatore.

	C_1				C_2				C_3			
	A	B	C	D	A	B	C	D	A	B	C	D
p_1	0,39	0,12	0,16	0,33	0,11	0,69	0,18	0,02	0	0,02	0,98	0
p_2	0,18	0,19	0,62	0,01	0,07	0,74	0,11	0,08	0,05	0,67	0,01	0,27

Svolgimento

	Somma					Prodotto					Massimo					Minimo				
	A	B	C	D	Out	A	B	C	D	Out	A	B	C	D	Out	A	B	C	D	Out
p_1	0,50	0,83	1,32	0,35	C	0	0	0,03	0	C	0,39	0,69	0,98	0,33	C	0	0,02	0,16	0	C
p_2	0,30	1,60	0,74	0,36	B	0	0,09	0	0	B	0,18	0,74	0,62	0,27	B	0,05	0,19	0,01	0,01	B

7) Dato un training set composto dai seguenti pattern:

$$\mathbf{x}_1 = \begin{bmatrix} 4,3 \\ -2,3 \\ -2,9 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} -5,0 \\ -3,6 \\ 1,2 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 2,8 \\ -7,4 \\ 4,2 \end{bmatrix}, \quad \mathbf{x}_4 = \begin{bmatrix} -5,9 \\ -8,6 \\ -7,5 \end{bmatrix}, \quad \mathbf{x}_5 = \begin{bmatrix} 8,8 \\ -5,6 \\ 3,6 \end{bmatrix}$$

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

$$y_1 = -3,0, \quad y_2 = 0,4, \quad y_3 = 6,3, \quad y_4 = 3,6, \quad y_5 = -4,1.$$

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:

$$\mathbf{X} = \begin{bmatrix} 4,3 & -2,3 & -2,9 & 1 \\ -5,0 & -3,6 & 1,2 & 1 \\ 2,8 & -7,4 & 4,2 & 1 \\ -5,9 & -8,6 & -7,5 & 1 \\ 8,8 & -5,6 & 3,6 & 1 \end{bmatrix}$$

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

$$\mathbf{y} = \begin{bmatrix} -3,0 \\ 0,4 \\ 6,3 \\ 3,6 \\ -4,1 \end{bmatrix}$$

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:

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