

1) Che cosa denota la matrice Σ nella definizione della distribuzione multinormale?

Dispense “Classificazione 1”

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

Dispense “Classificazione 2”

3) Qual è la differenza sostanziale dell’approccio “On-line” rispetto a “SGD con mini-batch” per il training di reti neurali?

Dispense “Reti neurali”

4) Descrivere le principali criticità e limitazioni dell’algoritmo di Clustering K-means.

Dispense “Clustering”

5) Dato un training set di 12000 pattern, supponendo di addestrare una rete neurale con SGD (mini-batch size=250) per 30 epoche, riportare, motivandone la risposta:

1. il numero di volte in cui ciascun pattern viene visto dalla rete;
2. il numero di volte in cui ciascun peso della rete viene aggiornato.

Svolgimento

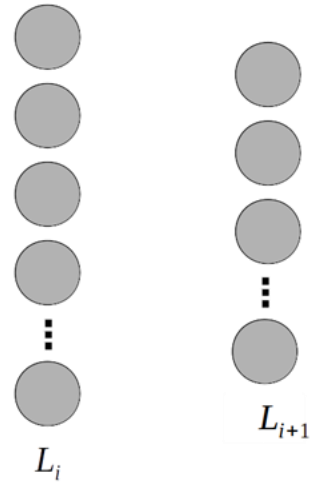
Ad ogni epoca tutti i pattern del training set vengono presentati alla rete esattamente una volta. Pertanto se la rete viene addestrata per 30 epoche, ciascun pattern le verrà mostrato 30 volte.

Nell’approccio SGD, l’aggiornamento dei pesi avviene alla fine di ogni iterazione che corrisponde alla presentazione di un intero mini-batch alla rete. Avendo 12000 pattern raggruppati in mini-batch di dimensione 250, avremo un totale di $\frac{12000}{250} = 48$ iterazioni per epoca. Ne consegue che il numero di volte in cui ogni peso della rete viene aggiornato corrisponde a $30 \cdot 48 = 1440$.

6) Dati due livelli di una rete neurale L_i e L_{i+1} costituiti rispettivamente da 64 e 32 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 11 neuroni del livello i (*receptive field* = 11). 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 $64 \cdot 32 = 2048$ connessioni totali.

Nel caso in cui i livelli siano di una CNN con “*receptive field*” uguale a 11 allora basta moltiplicare 11 per il numero di neuroni del livello $i + 1$ quindi $11 \cdot 32 = 352$ connessioni totali.

Per quanto riguarda i pesi, nel caso della rete MLP il numero degli stessi è pari al numero di connessioni (quindi 2048) mentre per la CNN è pari alla dimensione del *receptive field* (11 - essendo i pesi condivisi).

7) Supponendo di utilizzare *K-fold Cross-Validation* con $K = 12$ per suddividere 24000 pattern in *training* e *validation set*, quanti diversi addestramenti (*run*) vengono effettuati? Ad ogni *run* quanti pattern vengono utilizzati per il training e quanti per la validazione?

Svolgimento

Con $K = 12$ i 24000 pattern sono suddivisi in 12 partizioni da 2000 pattern l’una. Verranno eseguiti 12 addestramenti (*run*) utilizzando ogni volta una partizione diversa come *validation set* e le 11 restanti come *training set*. Pertanto ad ogni *run* verranno utilizzati 22000 pattern per il training e 2000 pattern per la validazione.