

- 1) Che cosa denota la matrice  $\Sigma$  nella definizione della distribuzione multinormale?
  - 2) Nelle SVM non lineari cosa si intende per kernel? Quali sono i kernel più utilizzati?
  - 3) Qual è la differenza sostanziale dell'approccio "On-line" rispetto a "SGD con mini-batch" per il training di reti neurali?
  - 4) Descrivere le principali criticità e limitazioni dell'algoritmo di Clustering K-means.
- 

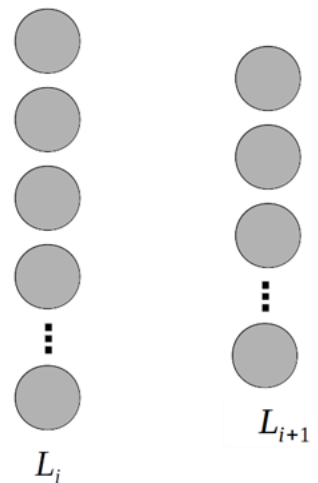
5) Dato un training set di 12000 pattern, supponendo di addestrare una rete neurale con SGD (mini-batch size=250) per 30 epoche, riportare, motivandone la risposta:

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.

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.



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?