

- 1) Cosa si intende per future reward nell'ambito del reinforcement learning? Fare un esempio.
  - 2) Nell'ambito dei multi-classificatori come si può ottenere indipendenza tra i singoli classificatori utilizzati?
  - 3) Che cosa si intende con tokenizzazione negli LLM? Perché si preferisce usare token invece di singoli caratteri o intere parole?
  - 4) Indicare le principali "stagioni" nello sviluppo dell'intelligenza artificiale e machine learning.
- 

5) Data una rete neurale MLP e un training set di 1000000 pattern, si decide di eseguire il training con SGD e mini-batch di 250 pattern. Si eseguono 180 epoche di addestramento. Calcolare il numero di volte in cui viene calcolato il (vettore) gradiente ed aggiornati i pesi durante l'apprendimento, motivando il calcolo.

6) Supponendo di utilizzare *K-fold Cross-Validation* con  $K = 8$  per suddividere 32000 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?

7) Date due distribuzioni multinormali identificate dai seguenti parametri:

$$\mu_0 = \begin{bmatrix} 5.40 \\ -2.48 \end{bmatrix}$$

$$\Sigma_0 = \begin{bmatrix} 1/10 & 0 \\ 0 & 3/5 \end{bmatrix}$$

$$P(w_0) = 0.4$$

$$\mu_1 = \begin{bmatrix} 3.01 \\ -0.93 \end{bmatrix}$$

$$\Sigma_1 = \begin{bmatrix} 1/5 & 0 \\ 0 & 1/10 \end{bmatrix}$$

$$P(w_1) = 0.6$$

Nell'ipotesi dell'impiego di un classificatore di Bayes multinormale, calcolare per il punto  $\mathbf{x} = \begin{bmatrix} 4.30 \\ -1.80 \end{bmatrix}$ :

- 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:

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

- 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.