

- 1) Perché le recenti reti neurali deep sono più efficaci delle MLP a tre livelli?
 - 2) La formula di distanza di un pattern dall'iperpiano risultante dal training di un SVM dipende da tutti i pattern del training set o solo da una parte di questi? Motivare la risposta.
 - 3) Nel classificatore di Bayes cosa si intende per densità di probabilità condizionale e probabilità a priori.
 - 4) Dare la definizione di Training, Validation e Test set e discutere una possibile suddivisione dei dati nei tre insiemi.
-

5) Data una cella di rete ricorrente (RNN) con 32 neuroni e 64 pesi addestrabili, disegnare lo schema grafico del suo *unfolding in time* su 4 stati. Quanti neuroni e pesi addestrabili ha la rete *unfolded*? Giustificare la risposta.

6) Data una rete neurale MLP e un training set di 100000 pattern, si decide di eseguire il training con SGD e mini-batch di 1000 pattern. Si eseguono 150 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.

7) Un problema di multiple linear regression viene risolto ai minimi quadrati ottenendo su un training set i seguenti coefficienti $\beta = [1,2 \quad -0,7 \quad -1,9 \quad 3,5]$ (dove l'ultimo elemento è il termine noto). Dato un test set costituito dai tre pattern \mathbf{x}_1 , \mathbf{x}_2 e \mathbf{x}_3 (di cui y_1 , y_2 e y_3 sono i valori veri della variabile dipendente):

$$\mathbf{x}_1 = [-2,6 \quad 7,9 \quad 5,1], y_1 = 5,8$$

$$\mathbf{x}_2 = [3,8 \quad -0,6 \quad -9,1], y_2 = 1,3$$

$$\mathbf{x}_3 = [0 \quad 10,8 \quad -7,4], y_3 = -3,4$$

determinare l'RMSE sul test set, riportando i principali passaggi intermedi del calcolo.