

1) Quali sono le limitazioni di Q learning per risolvere problemi complessi? Le tecniche di deep learning possono essere di aiuto in questo caso? Come?

Dispense “Deep Learning (2)”

2) Nell’ambito dei multi-classificatori che cosa è e come funziona il Borda count?

Dispense “Classificazione (2)”

3) Fare esempi pratici di ragionamento induttivo e deduttivo.

Dispense “Introduzione”

4) Che cosa si intende per clustering? Fare esempi di applicazioni.

Dispense “Clustering”

5) Dato un training set di 10000 pattern, supponendo di addestrare una rete neurale con SGD (mini-batch size=200) per 15 epoche, riportare, motivandone la risposta:

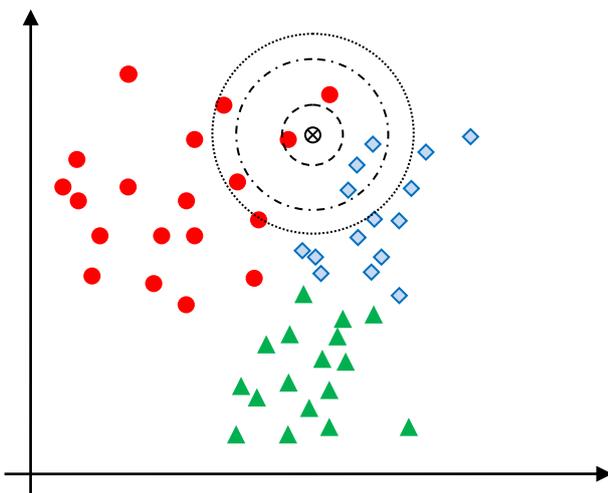
1. il numero di volte in cui ciascun pattern viene visto dalla rete;
2. il numero di volte in cui ciascun peso della rete viene aggiornato.

Svolgimento

Ad ogni epoca tutti i pattern del training set vengono presentati alla rete esattamente una volta. Pertanto se la rete viene addestrata per 15 epoche, ciascun pattern le verrà mostrato 15 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 10000 pattern raggruppati in mini-batch di dimensione 200, avremo un totale di $\frac{10000}{200} = 50$ iterazioni per epoca. Ne consegue che il numero di volte in cui ogni peso della rete viene aggiornato corrisponde a $15 \cdot 50 = 750$.

6) Date le distribuzioni riportate nel grafico sottostante, indicare come viene classificato il pattern \otimes da 3 classificatori k -NN (con k uguale a 1, 5 e 7) supponendo di utilizzare come metrica la distanza euclidea. Motivare la risposta.



Svolgimento

La regola k -NN determina i k elementi più vicini al pattern da classificare (in base alla metrica utilizzata). Ognuno dei k elementi vota per la classe a cui appartiene e il pattern da classificare viene assegnato alla classe che ha ottenuto il maggior numero di voti.

Pertanto, al pattern \otimes verrà assegnata la classe “cerchio” nel caso di $k=1$ (intorno - - -), la classe “quadrato” nel caso di $k=5$ (intorno - · -) e la classe “cerchio” nel caso di $k=7$ (intorno · · ·).

7) Per il training di un classificatore binario SVM si procede con una *grid search* combinata a *k-fold cross-validation* (con $k = 7$). Nell’ottica di voler valutare le seguenti combinazioni di kernel/iperparametri:

1. Lineare
 - $C = \{1, 0.1, 0.01, 0.001\}$
2. RBF
 - $C = \{1, 0.01\}$
 - $\gamma = \{0.5, 0.1, 0.05\}$
3. Polinomiale
 - $C = \{1, 0.1\}$
 - $degree = \{2, 3, 5\}$
 - $\gamma = \{0.2, 0.1\}$
 - $coef0 = \{0\}$

Si determini il numero complessivo di addestramenti da effettuare motivandone la risposta.

Svolgimento

Il numero di combinazioni di iperparametri per il kernel lineare è pari a 4, per il kernel RBF è uguale a $2 \cdot 3 = 6$ mentre per il kernel polinomiale è uguale a $2 \cdot 3 \cdot 2 \cdot 1 = 12$ per un totale di $4 + 6 + 12 = 22$ combinazioni. Per ognuna di queste combinazioni, la *grid search* esegue la *k-fold cross-validation* per un totale di $22 \cdot 7 = 154$ addestramenti.