

Bit Decoding SGD Logistic Regression

Un sistema di intercettazione osserva $K = 48$ features bidimensionali, che sono rappresentate dalla matrice **Xpred** presente nel file **homeworkclass.mat**. Ogni feature deve essere utilizzata per rivelare un bit incognito, vale a dire, una label appartenente al set $\{0, 1\}$.

La sequenza dei K bit corrisponde a 6 caratteri ASCII. La parola composta da tali caratteri è tratta da una celebre frase legata ad un film cult degli anni 80. Per completare l'esercizio, è necessario risalire alla frase originale (per quest'ultimo compito non servono metodi studiati al corso, ma un po' di intuito ed elementi di "cultura generale"...).

Per classificare i bit a partire dalle features si richiede di implementare un algoritmo di classificazione binaria basato sulla regressione logistica. In particolare, si richiede di implementare K classificatori distinti, uno per ciascun bit da classificare.

Per progettare i K classificatori, viene messo a disposizione un training set costituito dalle coppie feature-label

$$\{(X_i, Y_i)\}_{i=1}^n, \quad \text{dove } X_i \in \mathbb{R}^2, \quad Y_i \in \{0, 1\}$$

Le features e le label sono memorizzate rispettivamente nella matrice **Xtrain** e nel vettore **Ytrain** presenti nel file **homeworkclass.mat**.

Il training set contiene K tranches di dati consecutive, di lunghezze incognite, che corrispondono ai K problemi di classificazione. La prima tranche di dati è distribuita in accordo al modello che corrisponde al primo problema di classificazione, cioè a **Xpred(1,:)**; la seconda in accordo al modello che corrisponde al secondo problema di classificazione, cioè a **Xpred(2,:)**; e così via.

Si addestrì un classificatore logistico sul training set. A tal fine, si applichi l'algoritmo del gradiente stocastico con step-size costante. Considerata la sequenza dei vettori $\beta(i)$ prodotti dall'algoritmo del gradiente stocastico, si scelgano K valori di questa sequenza che possano verosimilmente corrispondere ai K problemi di classificazione considerati. Per ottimizzare questa scelta, si valuti il comportamento dell'algoritmo per diverse scelte dello step-size.