CptS 315 - HW 3

Zach Fechko (011711215)

04 November, 2022

Contents

Question 1

Answer the following with a yes or no along with proper justification.

Is the decision boundary of voted perceptron linear? Why or why not?

Is the decision boundary of averaged perceptron linear? Why or why not?

Answer

Question 2

Consider the following setting. You are provided with n training examples, $(x_1, y_1, h_1), (x_2, y_2, h_2), \ldots, (x_n, y_n, h_n)$ where x_i is the input example, y_i is the class label ($+1$ or -1), and $h_i > 0$ is the importance weight of the example. The teacher gave you some additional information by specifying the importance of each training example. How will you modify the perceptron algorithm to be able to leverage this extra information?

Answer

Question 3

Consider the following setting. You are provided with n training examples, $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ where x_i is the input example, y_i is the class label ($+1$ or -1). However the training data is highly imbalanced (say 90 of the examples are negative and 10 are positive) and we care about the accuracy of positive examples. How will you modify the perceptron algorithm to solve this learning problem? Explain your approach.

Answer

Question 4

You were just hired by MetaMind. MetaMind is expanding rapidly, and you decide to use your machine learning skills to assist them in their attempts to hire the best. To do so, you have the following available to you for each candidate i in the pool of candidates $\mathcal{I}$.

Their GPA

Whether they took a Data Mining course and got an A

Whether they took an algorithms course and got an A

Whether they have a job offer from Google

Whether they have a job offer from Facebook

The number of typos in their resume

You decide to repsent each candidate $i \in \mathcal{I}$ by a corresponding 6 dimensional feature vector $f(x^{(i)})$. You believe that if you just knew the right weight vector $w \in \mathbb{R}^6$ you could reliably predict the quality of a candidate i by computing $w \cdot f(x^{(i)})$. To determine w your boss lets you sample pairs of candidates from the pool. For a pair of candidates (k, l) you can have them face off in a "Data Mining Fight". The result is score($k \succ l$), which tells you that candidate k is at least score($k \succ l$) better than candidate l. Note that the score will be negative when l is a better candidate than k. Assume you collected scores for a set of pairs of candidates $\mathcal{P}$.

Describe how you would use a perceptron based algorithm to learn the weight vector w. Make sure to describe the basic intuition; how the weight updates will be done; and psuedocode for the entire algorithm.

Answer

Question 5

Suppose we have n_+ positive training examples and n_- negative training examples. Let C_+ be the center of the positive examples and C_- be the center of the negative examples. i.e. $C_+ = \frac{1}{n_+} \sum_{i:y_i=+1} x_i$ and $C_- = \frac{1}{n_-} \sum_{i:y_i=-1} x_i$. Consider a simple classifier called CLOSE that classifies a test example x by assigning it to the class whose center is closest

Show that the decision boundary of the CLOSE classifier is a linear hyperplane of the form $\text{sign}(w \cdot x + b)$. Compute the values of w and b in terms of C_+ and C_-.

Recall that the weight vector can be written as a linear combination of all the training examples $w = \sum_{i=1}^{n_+ + n_-} \alpha_i \cdot y_i \cdot x_i$. Compute the dual weights (α's). How many of the training examples are support vectors?

Answer

Question 6

Read the following paper and write a brief summary of the main points in at most two pages
A few useful things to know about machine learning

Answer