CIS 419/519: Quiz 4

October 4, 2019

1. We have examples in the form of 10 boolean variables, $< x_1, x_2, ..., x_{10} >$, and know the true function $f(X)$ is in the class of monotone conjunctions. Say we have a "teacher" who knows the true function and must teach the true function through a set of examples; the true function is $y(X) = x_1 \wedge x_2 \wedge x_3 \wedge x_6$. What is the minimum number of examples that is required to learn this function?

   (a) 10

   (b) 11

   (c) 5

   (d) 3

   Solution: (c)

2. Let C be the finite concept class of all monotone conjunctions with up to 3 boolean variables. We are trying to learn f where $f \in C$. Each example in our training set is of the form $(< X >, y)$ where $X =< x_1, x_2, x_3 >, x_i \in \{1, 0\}$ and $y \in \{1, 0\}$. Say we want to use the halving algorithm to reduce the size of consistent concepts in C; our first data point is:

   $$(< 0, 1, 0 >, 1)$$

   Will we make a mistake on this first example? Why?

   (a) Yes we will, because more concepts in C predict 1 than 0

   (b) Yes we will, because more concepts in C predict 0 than 1

   (c) No we won't, because more concepts in C predict 1 than 0

   (d) No we won't, because more concepts in C predict 0 than 1

   Solution: (b)

3. Suppose we have a weight vector $w \in R^2$ with input vectors $x_i \in R^2$ and $y_i \in \{-1, 1\}$, let us initialize our 2-dimensional weight vector to be

$w = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Also, suppose we only have 2 examples in our dataset: ($x_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, y_1 = 1$), ($x_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, y_2 = -1$). After training a model based on the Perceptron algorithm on the above dataset over 1 epoch, which option represents the correct final state of the weight vector if the linear threshold function is $\hat{y} = sgn\{w^T \cdot x \geq 0\}$?

(a) $w^T = [-1, -1]$
(b) $w^T = [0, -1]$
(c) $w^T = [0, 0]$
(d) $w^T = [0, 0]$

Solution: (b)

4. In a mistake-driven algorithm, if we make a mistake on example x_i with label y_i, we can be sure that when the weights are updated we will never make a mistake on this same example if we see it again.

(a) True
(b) False

Solution: (b)

5. Suppose we're using the Averaged Perceptron algorithm. The training data consists of m examples. Assume that after training for 1 epoch, we've made k mistakes on the training data. We have accumulated the following weight vectors: $\{v_1, \ldots v_{k+1}\}$ and their respective consistency counts are $\{c_1, \ldots c_{k+1}\}$. Which statement is true?

(a) $k > m$
(b) $c_1 > c_2 > ... > c_{k+1}$ (decreasing from 1 to k+1)
(c) $c_1 < c_2 < ... < c_{k+1}$ (increasing from 1 to k+1)
(d) None of the above

Solution: (d)