CS M146 Quiz Week 3 Solutions: Linear Models, Perceptron

February 2, 2022

1. [Point: **4**] This question contains the following True/False subquestions about the Perceptron and its learning algorithm presented in class.

   (1) Can a Perceptron model represent boolean function XNOR? The table shows the results of XNOR computation on two independent binary variables.

   A. True

   B. False

   [Answer: **False. For the same reason the Perceptron cannot represent XOR, XNOR (the opposite of XOR) cannot be represented. Eg. Both points $(1,0)$ and $(0,1)$ appear on either side of the line formed by $(0,0)$ and $(1,1)$. This makes it impossible to keep both $(1,0)$ and $(0,1)$ on the same side of a separating line while not making mistakes on $(0,0)$ and $(1,1)$.**]

   (2) The algorithm always converges in at most R^2/γ^2 steps on *any* dataset.

   A. True

   B. False

   [Answer: **False. It does not converge for non-separable data.**]

   (3) An updated vector w_{new} always has lower training error compared to the previous weight vector w.

   A. True

   B. False

   [Answer: **False. While the algorithm may converge, there are no guarantees on the training error at each step. Updating for one mistake may temporarily lead to worse training error.**]

   (4) The algorithm may enter an infinite loop even on linearly separable data.

   A. True

   B. False

   [Answer: **False, due to the mistake bound theorem.**]

2. [Point: **2**] The following diagram represents data and a hyperplane.

   Blue crosses (x) and red circles (o) correspond to labels $y = +1$ and $y = -1$, respectively. The current hyperplane is defined by its normal vector $w = (-1, 1)$ (green arrow). Suppose we are in process of learning a perceptron on the dataset, and we use the algorithm described in the lecture.

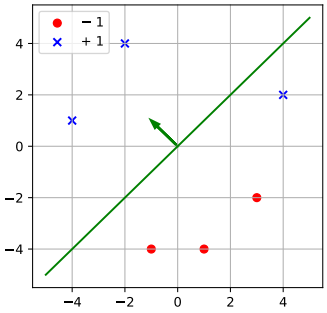
   What will the updated vector w be after the current iteration?

   A. $(2, 0)$

   B. $(2, -1)$

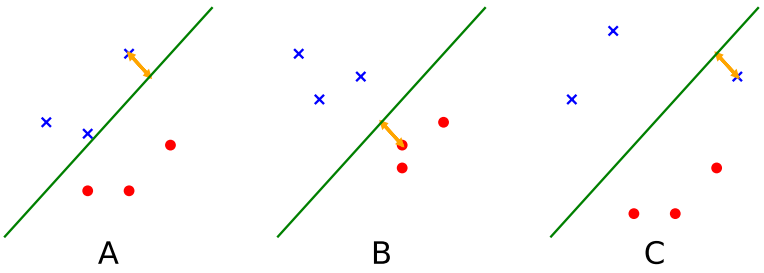
   C. $(3, 3)$

   D. $(-2, -2)$



[Answer: **C. The only incorrectly classified point is** $(4, 2)$**, and** $w_{new} = w + y_n * x_n = (-1, 1) + (1)(4, 2) = (3, 3)$**.**]

3. [Point: **2**] Which of the following figures correctly labels the margin (the orange arrow) *with respect to the hyperplane*?



   A. A
   B. B
   C. C
   D. None of the above

   [Answer: **B. In A, there is a blue cross closer to the line. In C, the line does not separate the points**]

4. [Point: **2**] Given the following three points in $\mathbb{R}^3$:

$$X = (1, 2, 3),$$
$$Y = (4, 5, 6),$$
$$Z = (4, 2, 0),$$

   which one of the following vectors is the normal vector of the plane formed by points X, Y, Z?

   A. $(1, -2, 1)$
   B. $(-1, 2, 1)$
   C. $(1, 1, -1)$
   D. $(1, 2, 1)$

[Answer: **A. The cross product between two conjoined vectors on a plane can inform the normal direction. With some scaling to simplify the vectors:** $\overrightarrow{XY} =< 2, 1, 0 >$ **and** $\overrightarrow{XZ} =< 1, 0, -1 >$**. Then,** $\overrightarrow{XY} \times \overrightarrow{XZ} = -1i + 2j - 1k$**. Then** $< 1, -2, 1 >$ **is one normal vector describing the plane.**]

5. [Point: **2**] Choose the statement that is true about the Voting Perceptron:
   A. The number of parameters grows linearly with the number of points in the dataset.
   B. The number of parameters grows linearly with the number of mistakes made on the dataset.
   C. The voting perceptron ensures it will never make mistakes on linearly non-separable data.

   [Answer: **B. The voting perceptron saves one set of parameters every time it updates ie. when it makes a mistake on a point in the given dataset.**]

6. [Point: **2**] Which of the following statements are true about the Sigmoid function $\sigma(x)$
   A. The domain is $(0, 1)$ and the range is $(0, 1)$
   B. The domain is $(-\infty, +\infty)$ and the range is $(0, 1)$
   C. The domain is $(0, 1)$ and the range is $(-\infty, +\infty)$
   D. When $x = 0$, the sigmoid should be $\sigma(x) = 0$

   [Answer: **B. The sigmoid function has no limitations on the domain, while remaining between** $0 < \sigma(x) < 1$]