Quiz 5

This is a preview of the published version of the quiz

Started: Oct 18 at 5:45pm

Quiz Instructions

Question 1 — 1 pts

Let $(x_1, y_1), \ldots, (x_t, y_t)$ be a sequence of labeled examples where $x_i \in \mathbb{R}^n$, $||x_i|| \leq 10$, and $y_i \in \{1, -1\}$ for all i. Suppose there exists some $u \in \mathbb{R}^n, ||u|| = 1$ such that

$y_i(u^T x_i) \geq 5$ for all i.

Then Perceptron makes at most 100 mistakes on this example sequence.

- ◯ True
- ◯ False

Question 2 — 1 pts

You are tasked with learning a new function over 5 Boolean variables. The function's output is either 0 or 1 and it is given to you that this function belongs to the at least *m-of-n* class of functions. Your friend suggests that they have a good learning algorithm that can learn linear threshold units and suggests that you use it. Is this a good choice?

- ◯ No, since only neural networks can express the type of functions you care about
- ◯ Yes, since all Boolean functions can be represented as LTUs.
- ◯ Yes, since m-of-n functions can be represented as LTUs.
- ◯ No, since the class of LTUs over 5 variables may not express all the functions you care about.

Question 3

1 pts

Let $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ be a dataset where each x_i is a feature vector and $y_i \in \{-1, 1\}$ be the corresponding binary label.

If D is linearly separable, which of the following conditions must be true? Select all that apply.

- ☐ There exist a weight vector w, bias θ, and non-negative constant c such that for all i,
  $y_i(w^T x_i + \theta) \geq c$
- ☐ There exist a weight vector w, bias θ, such that for all i,
  $y_i(w^T x_i + \theta) \geq 0$
- ☐ There exist a weight vector w, bias θ, such that for all i,
  $y_i(w^T x_i + \theta) \leq 0$
- ☐ There exist exactly one weight vector w and bias θ such that for all i,
  $y_i(w^T x_i + \theta) \geq 0$

Question 4

1 pts

Regarding the learning algorithms that we have learned so far, which statement(s) of the following are true? Select all that apply.

- [ ] The averaged perceptron algorithm can be implemented by keeping track of a weighted vector w_{sum} that is a weighted sum of earlier weight vectors.
- [ ] The final weight vector of the averaged perceptron algorithm weighs each of the earlier weight vectors by a weight that is inversely proportional to the "quality" of the weight vector.
- [ ] The perceptron algorithm updates its weight vector by adding an example (times some constant) to the current weight vector.
- [ ] When the examples presented to the Perceptron algorithm are Boolean vectors, all coordinates of the weight vector w are being updated (changed to a different value) when the algorithm makes a mistake.
- [ ] When the examples presented to the Winnow algorithm are Boolean vectors, all coordinates of the weight vector w are being updated (changed to a different value) when the algorithm makes a mistake.

Question 5

1 pts

Consider the following data points:

$x_1 = [1, -1, 3]$

$x_2 = [4, 0, 0]$

$x_3 = [0, 1, -2]$

$x_4 = [2, 2, 0]$

Assume we have a weight vector and bias

$w = [1, -1, 0]$

$\theta = 2$

The distance between a point x and the hyperplane defined by w and θ is

$$\frac{|w^T x + \theta|}{||w||}$$

which example $x_i, i \in [1, 4]$ above is the **farthest** to the hyperplane?

- ○ x_1

- ◯ x_2
- ◯ x_3
- ◯ x_4

Not saved Submit Quiz