

## CS 4780/5780 Homework 2

Due: Tuesday 09/25/18 11:55pm on Gradescope

The following problems concern the perceptron. Using the knowledge you learned in class about when the perceptron may be applied, we will analyze how to algorithmically train the perceptron, and determine when the perceptron converges.

### Problem 1

Five updates as follows:  $[(0, 0), (1, 3), (2, -1), (3, 2), (4, -2), (5, 1)]$

### Problem 2

- (a) Yes. We can compare training error  $\sum_{(x,y) \in D_{TR}} \ell_{0/1}(h(x), y)$  to test error  $\sum_{(x,y) \in D_{TE}} \ell_{0/1}(h(x), y)$ .
- (b) Yes. We do not need to evaluate training error. We know that this "fully trained" perceptron has zero training error because perceptrons finish training only once training error is zero.

### Problem 3

The limit does not exist. The dataset is no longer linearly separable, and thus the perceptron will not converge.

- (a) All points happen to be equidistant from the hyperplane, with a distance of  $\sqrt{2}$  and  $\sqrt{2}/2$  respectively. This is not  $\gamma$  because the points have not been rescaled to lie within the unit circle.
- (b) Rescale by dividing each point by the largest norm  $\|(-1, 1)\| = \sqrt{2}$ . Now the minimum distance from the hyperplane is  $\gamma = 1/2$ , and the upper bound is  $1/\gamma^2 = 4$ .
- (c) The limit does not exist. The dataset is no longer linearly separable, and thus the perceptron will not converge.

### Problem 4

Because the final vector  $\vec{w}$  is equal to  $\sum_{(\vec{x}_i, y_i) \in M} y_i \vec{x}_i$ , where  $M$  is a multiset of all misclassified points in the training process, we see that  $\vec{w} = (0, 0, 0, 0, 4) * 2 + (0, 0, 6, 5, 0) - (3, 0, 0, 0, 0) - (0, 9, 3, 6, 0) - (0, 1, 0, 2, 5) = (-3, -10, 3, -3, 3)$