

## CS/ECE/ME532 Assignment 1

1) A binary classifier decides whether data is from one of two classes, labeled 1 and -1. In this problem the data is described by two features  $x_1$  and  $x_2$  and the classification decision is made as follows. The class with label 1 is decided if  $x_1a_1 + x_2a_2 > b$  while the class with label -1 is decided when  $x_1a_1 + x_2a_2 < b$ . Here  $a_1$ ,  $a_2$ , and  $b$  are given real numbers.

- a) The *decision boundary* is the set of  $\{x_1, x_2\}$  that satisfy  $x_1a_1 + x_2a_2 = b$ . Thus, we may assign the label to the data using the sign of  $y = x_1a_1 + x_2a_2 - b$  since label 1 is decided if  $y > 0$  and label -1 is decided if  $y < 0$ . That is, the label may be obtained as  $\text{sign}\{y\}$ . Express  $y$  as an inner product of a vector  $\mathbf{x}$  containing the features and  $\mathbf{w}$  containing weights, that is, write  $y = \mathbf{x}^T \mathbf{w}$ .
- b) Let  $x_2$  be the vertical axis and  $x_1$  be the horizontal axis. Show that the decision boundary  $y = 0$  is a straight line. Find the slope and intercept with the vertical axis as a function of  $a_1, a_2, b$ .

- c) You classify  $n$  data samples using  $\text{sign}\{\mathbf{y}\}$  where  $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \mathbf{X}\mathbf{w}$ . Suppose  $n = 4$  and the features for the 4 data samples are 1 : (0, 0.3), 2 : (0.2, 0.1), 3 : (0.4, 0.6), 4 : (0.9, 0.8). Write out the matrix  $\mathbf{X}$ .

- d) Suppose  $a_1 = 1$ ,  $a_2 = 2$ , and  $b = 1$ . Sketch the decision boundary in the  $x_1$ - $x_2$  plane assuming  $x_2$  is the vertical axis and  $x_1$  is the horizontal axis. Graph the four data samples from the previous part and classify them.
- e) Download and run the linear classifier script. This script classifies 5000 examples of (randomly generated) data consisting of two features using the linear classifier. Save the figure and include it in your submission. Describe the decision boundary you observe using a sentence.
- f) Change the classifier weights to  $\mathbf{w} = [1.6 \quad 2 \quad -1.6]^T$ . Rerun the scrip. Include the figure in your pdf file. Briefly describe how the change in the weights changed the decision boundary.