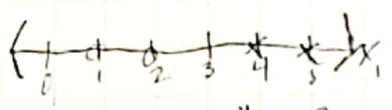
Aidan Daniel
R11679461

1. condition 1: no offset



$\underline{w} = \langle 3\ 0 \rangle$

$x^* = 3 \Rightarrow 3 \cdot 3 - 9 > 0$

condition 2: offset does not cause it to be negative

$\underline{w} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$

$2(2) + 1(1) + 1 = 6 > 0$

offset does not cause this to be negative

$\underline{w} = \begin{bmatrix} 2 \\ 1 \\ -9 \end{bmatrix}$

$2(2) + 1(1) - 9 = -4 < 0$

offset makes it negative

2. One disadvantage of Batch perceptron over online perceptron is having to go through all of the samples before changing your weight vector, large number of samples then slows this down and since it adds up all of the misclassified feature vectors, you require a smaller learning rate.

3

$$J(\underline{w}) = \sum_{\text{misclassified}} \underline{w}^2 \underline{x}^2$$

$$\nabla_{\underline{w}} J(\underline{w}) = \sum_{\text{misclassified}} 2\underline{w}\underline{x}^2$$

$$\underline{w}(k+1) = \underline{w}(k) + \rho \sum_{\text{misclassified}} 2\underline{w}\underline{x}^2$$

4 a. $d_1(\underline{x}^*) = 2(1) + (-1) - (2) + 3 = 2$

$d_2(\underline{x}^*) = -(1) - (-1) + 2(2) + 1 = 5$

$d_3(\underline{x}^*) = +1 + 2 = 1$

$d_4(\underline{x}) = -(-1) + 2(2) + 2 = 7$

$d_4 > d_2 > d_1 > d_3$

$x^* \in C_4$

b. $d_3 > d_4 \Rightarrow d_3 - d_4 > 0$

$x_2 + 2 - (-x_2 + 2x_3 + 2)$

$2x_2 - 2x_3 > 0$

3D training data

4 samples used in training