Homework 5

Name: Yuan Li
ID: N19728558 netID yl6606
Email: foxerlee1@gmail.com

Part I: Written Exercises

1. (a)

Answer:

Because the decision boundary is 0.5, we can get the number of TP, FP, FN, TN should be:

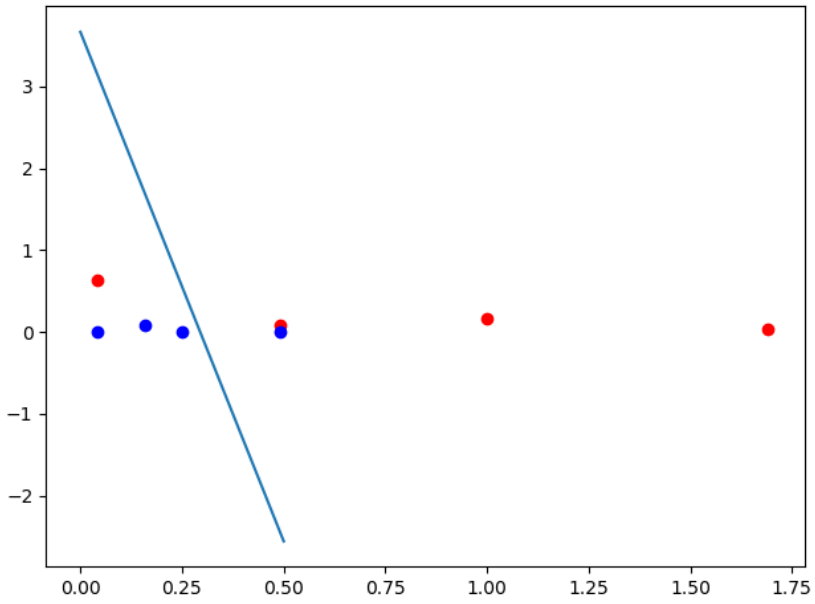
$$TP = 3$$
$$FP = 1$$
$$FN = 1$$
$$TN = 3$$

So the confusion matrix should be:

	pre=1	pre=0
true=1	3	1
true=0	1	3

1. (b)

Answer:



1. (c)

Answer:

$$FPR = \frac{FP}{TN + FP} = \frac{1}{1+3} = 0.25$$

1. (d)

Answer:

$$TPR = \frac{TP}{TP + FN} = \frac{3}{1+3} = 0.75$$

1. (e)

Answer:

$$accuracy = \frac{TP + TN}{TP + FP + FN + TN} = \frac{3+3}{1+3+1+3} = 0.75$$

1. (f)

Answer:

$$recall = \frac{TP}{TP + FP} = \frac{3}{1+3} = 0.75$$

1. (g)

Answer:

$$precision = \frac{TP}{TP + FN} = \frac{3}{1+3} = 0.75$$

1. (h)

Answer:

$$\begin{aligned} l(w) &= -(\sum_{i=1}^{N} y^{(i)} ln(h(x)) + (1 - y^{(i)}) ln(1 - h(x)))) \\ &= -(0 * ln0.389 + 1 * ln(1 - 0.389)) - (0 * ln0.042 + 1 * ln(1 - 0.042)) - (0 * ln0.613 + 1 * ln(1 - 0.613)) \\ &- (0 * ln0.167 + 1 * ln(1 - 0.167)) - (1 * ln0.572 + 0 * ln(1 - 0.572)) - (1 * ln0.526 + 0 * ln(1 - 0.526)) \\ &- (1 * ln0.393 + 0 * ln(1 - 0.393)) - (1 * ln0.638 + 0 * ln(1 - 0.638)) \\ &= 4.252 \end{aligned}$$

2. (a)

Answer:

The learning rate is too large. So we move too far, which means overshoot.

2. (b)

Answer:

The learning rate is too small. So we move slowly, and each step is too small.

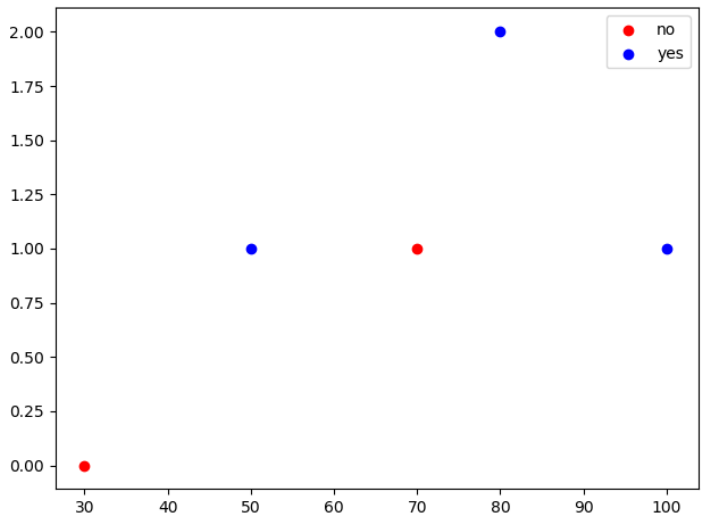
2. (c)

Answer:

This model is not suitable for this data.

3. (a)

Answer:



3. (b)

Answer:

Based on the method to calculate the w for linear regression, we can get:

$$w = (X^T X)^{-1} X^T z$$
$$= \begin{bmatrix} -0.00807018 \\ 0.86666667 \end{bmatrix}$$

In order to fit the form of classifier mentioned in assignment, I set z as:

$$z = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$$

The predict labels are $[0, 1, 1, 1, 1]$, with only one error.

3. (c)

Answer:

i = 0. Because only $z^{(0)} < 0$, $-z^{(0)} > 0$, then $e^{-z^{(0)}}$ is the biggest. After we take the denominator, $\frac{1}{1+e^{-z^{(0)}}}$ should be the smallest.

So sample 0 is the least likely.

3. (d)

Answer:

No they don't change. Because the judgment in b and c is based on the values are positive or negative. If the original w multiplies a positive scalar, it will not change the values' positive or negative relationship. As a result, the result will not change.