CSE514, Fall 2021, HW 2 Name: Anh Le Student ID: 488493
Note: This homework is worth a total of 15 points

**Q1:** Given these results from a model where class T is the positive class, calculate the following performance measures at a threshold of 0.5 (i.e. data points with a score > 0.5 is predicted T) **(0.5pts each, 5pts total)**

| index | class | score | index | class | score |
|---|---|---|---|---|---|
| 1 | T | 0.95 | 11 | T | 0.45 |
| 2 | T | 0.85 | 12 | N | 0.40 |
| 3 | N | 0.80 | 13 | T | 0.38 |
| 4 | T | 0.67 | 14 | N | 0.35 |
| 5 | T | 0.65 | 15 | N | 0.33 |
| 6 | T | 0.60 | 16 | N | 0.30 |
| 7 | N | 0.58 | 17 | T | 0.28 |
| 8 | N | 0.54 | 18 | N | 0.27 |
| 9 | T | 0.52 | 19 | T | 0.26 |
| 10 | N | 0.51 | 20 | N | 0.18 |

False Positive Rate:

$$\frac{FP}{FP + TN} = \frac{4}{4 + 6} = 0.4$$

True Positive Rate:

$$\frac{TP}{TP + FN} = \frac{6}{6 + 4} = 0.6$$

Type I error:

$$= FP = 4$$

Type II error:

$$= FN = 4$$

Precision:

$$\frac{TP}{TP + FP} = \frac{6}{6 + 4} = 0.6$$

Recall:

$$\frac{TP}{TP + FN} = \frac{6}{6 + 4} = 0.6$$

Sensitivity:

$$\frac{TP}{TP + FN} = \frac{6}{6 + 4} = 0.6$$

Specificity:

$$\frac{TN}{TN + FP} = \frac{6}{6 + 4} = 0.6$$

Accuracy:

$$\frac{TP + TN}{TP + FP + FN + TN} = \frac{6 + 6}{6 + 4 + 4 + 6} = 0.6$$

F-score:

$$\frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} = \frac{2}{\frac{1}{0.6} + \frac{1}{0.6}} = 0.6$$

**Q2:** Given this dataset:

And a new data point:
Height = 200
Weight = 200

| Class | Height | Weight |
|---|---|---|
| 1 | 105 | 114 |
| 1 | 92 | 169 |
| 1 | 87 | 140 |
| 2 | 111 | 109 |
| 2 | 79 | 44 |
| 2 | 92 | 55 |
| 3 | 265 | 331 |
| 3 | 330 | 284 |
| 3 | 185 | 309 |

**Q2a (3pts):**
Use the nearest neighbor classifier with the Euclidean distance function to label the new data point. Show your work by giving me the distances you calculated between the new data point and the training data.

$d_1 = \sqrt{(200-105)^2 + (200-114)^2} = 128.14$

$d_2 = \sqrt{(200-92)^2 + (200-169)^2} = 112.36$

$d_3 = \sqrt{(200-87)^2 + (200-140)^2} = 127.94$

$d_4 = \sqrt{(200-111)^2 + (200-109)^2} = 127.29$

$d_5 = \sqrt{(200-79)^2 + (200-44)^2} = 197.43$

$d_6 = \sqrt{(200-92)^2 + (200-55)^2} = 180.8$

$d_7 = \sqrt{(200-265)^2 + (200-331)^2} = 146.24$

$d_8 = \sqrt{(200-330)^2 + (200-284)^2} = 154.78$

$d_9 = \sqrt{(200-185)^2 + (200-309)^2} = (110.03)$

⇒ New Data Point: Class 3

**Q2b (3pts):**
Use 3-nearest neighbors with the Manhattan distance function to label the new data point. Show your work by giving me the distances you calculated between the new data point and the training data.

$d_1 = |200-105| + |200-114| = 181$

$d_2 = |200-92| + |200-169| = (139)$ class 1

$d_3 = |200-87| + |200-140| = (173)$ Class 1

$d_4 = |200-111| + |200-109| = 180$

$d_5 = |200-79| + |200-44| = 277$

$d_6 = |200-92| + |200-55| = 253$

$d_7 = |200-265| + |200-331| = 196$

$d_8 = |200-330| + |200-284| = 214$

$d_9 = |200-185| + |200-309| = (124)$ class 3

∴ New Data Point: Class 1

**Q3: (4pt)** Instead of classification, you're fitting a regression model to predict the weight values from height:

$$f(x) = mx + b$$

Using MSE as your loss function:

$$L(m, b) = \frac{1}{n}\sum_{i=1}^{n}(y_i - (mx_i + b))^2$$

You've randomly started your parameter values at

m = 1

b = 100

Use gradient descent to update the two parameter values by one step, using a learning rate of $\alpha$ = 0.1

| Class | Height | Weight |
|---|---|---|
| 1 | 105 | 114 |
| 1 | 92 | 169 |
| 1 | 87 | 140 |
| 2 | 111 | 109 |
| 2 | 79 | 44 |
| 2 | 92 | 55 |
| 3 | 265 | 331 |
| 3 | 330 | 284 |
| 3 | 185 | 309 |

1, Partial derivative of the loss function:

$$D_m = \frac{1}{n}\sum_{i=0}^{n} 2(y_i - (mx_i + b))(-x_i) = \frac{-2}{n}\sum_{i=0}^{n} x_i(y_i - (mx_i + b))$$

$$= \frac{-2}{9}\left[105(114 - (105 + 100)) + 92(169 - (92 + 100)) + \dots\right] = 22911.33$$

$$D_b = \frac{-2}{n}\sum_{i=0}^{n}(y_i - (mx_i + b)) = \frac{-2}{9}\left[(114 - (105 + 100)) + \dots\right] = 153.56$$

2, Update m and b with $\alpha = 0.1$

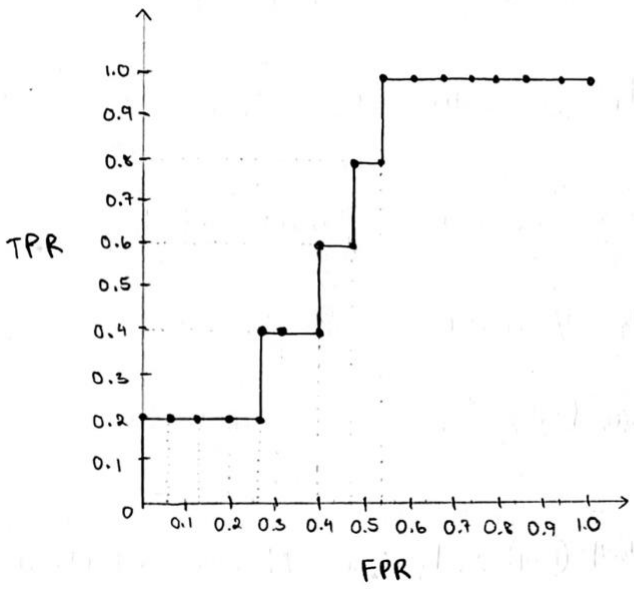
$$m = m - \alpha D_m = 1 - 0.1(22911.33) = \boxed{-2290.133}$$

$$b = b - \alpha D_b = 100 - 0.1(153.56) = \boxed{84.645}$$

**Extra credit:**
Using this dataset, draw the ROC and PR curves **(1 bonus pt each)**

| index | class | score | index | class | score |
|---|---|---|---|---|---|
| 1 | T | 0.95 | 11 | T | 0.45 |
| 2 | N | 0.85 | 12 | N | 0.40 |
| 3 | N | 0.80 | 13 | T | 0.38 |
| 4 | N | 0.67 | 14 | N | 0.35 |
| 5 | N | 0.65 | 15 | N | 0.33 |
| 6 | T | 0.60 | 16 | N | 0.30 |
| 7 | N | 0.58 | 17 | N | 0.28 |
| 8 | N | 0.54 | 18 | N | 0.27 |
| 9 | T | 0.52 | 19 | N | 0.26 |
| 10 | N | 0.51 | 20 | N | 0.18 |

ROC curve:




PR curve: