CSE514, Spring 2022, HW 2 Name: JIANJUN WEI Student ID: 1701040

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, 3pts 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

TP 6	FP 4	
FN 4	TN 6	

True Negative Rate:

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

True Positive Rate:

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

Type I error:

$FP = 4$

Accuracy:

$$\frac{TP+TN}{TP+FP+FN+TN} = \frac{12}{20} = 0.6$$

Precision:

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

False discovery rate:

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

Q2: Given this dataset:

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

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

Q2a (3pts):
Use the nearest neighbor classifier with the Euclidean distance function and k=3 to label the new data point as Class 1, 2, or 3.

Show your work by giving me the distances you calculated between the new data point and the training data.

$dis1\_1 = \sqrt{95^2+86^2} = \sqrt{16421}$

$dis1\_2 = \sqrt{108^2+31^2} = \sqrt{12625}$

$dis1\_3 = \sqrt{113^2+60^2} = \sqrt{16369}$

$dis2\_1 = \sqrt{89^2+91^2} = \sqrt{16202}$

$dis2\_2 = \sqrt{121^2+156^2} = \sqrt{38797}$

$dis2\_3 = \sqrt{108^2+145^2} = \sqrt{32689}$

$dis3\_1 = \sqrt{65^2+131^2} = \sqrt{21386}$

$dis3\_2 = \sqrt{130^2+84^2} = \sqrt{23956}$

$dis3\_3 = \sqrt{15^2+109^2} = \sqrt{12106}$

$k=3$

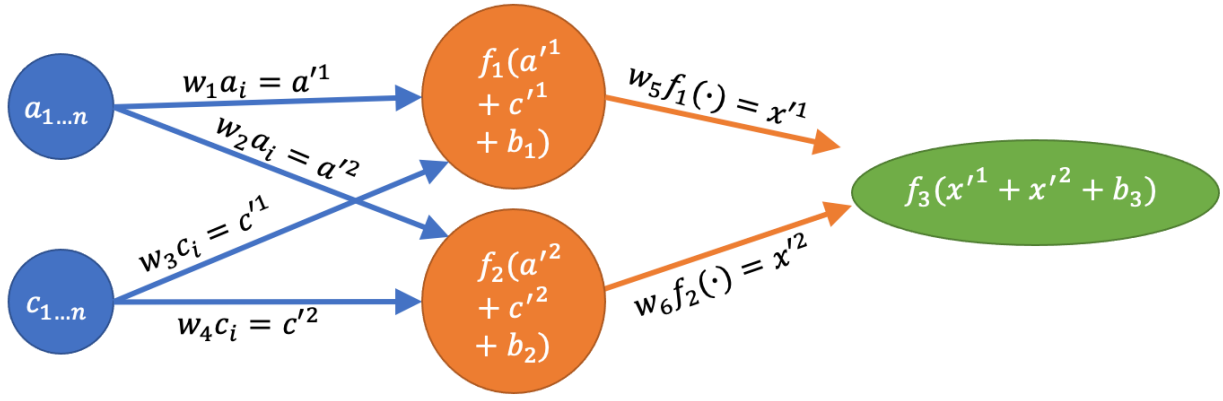
$dis3\_3 < dis1\_2 < dis2\_1 < dis1\_3$

there is a tie when $k=3$

thus, $k = k+1 = 4$

New data is for Class 1

Q3: Given this ANN structure:



And the following parameter/function definitions:

W = [-15, -3, -2, 4, 1, 10] B = [4, 1, -0.5]

$f_1(x) = f_2(x) = \max(0.1x, x)$ $f_3(x) = x^2$

What are the intermediate and/or output values for the following data points?

Q3a (3pt): **Data point:** **a = 0.5, c = 0.5**

a'^1 value: $-15 \times 0.5 = -7.5$ a'^2 value: $-3 \times 0.5 = -1.5$

c'^1 value: $-2 \times 0.5 = -1$

Q3b (3pt): **Data point:** **a = 1, c = 0**

$c'^1 = 0$
$c'^2 = 0$
$a'^1 = -15$
$a'^2 = -3$

$f_1(a'^1) = f_1(-15)$
$f_2(a'^2) = f_2(-3)$

c'^2 value: 0 f_1 value: -1.1

f_2 value: -0.2

Q3c (3pt): **Data point:** **a = 0, c = 1**

$a'^1 = 0$
$a'^2 = 0$
$c'^1 = -2$
$c'^2 = 4$

x'^1 value: 2 x'^2 value: 10

f_3 value: 2652.25

$f_1(x) = f_1(-2+4) = f(2) = 2$
$f_2(x) = f_2(4+1) = f(5) = 5$

$x'^1 = 1 \times 2 = 2$
$x'^2 = 10 \times 5 = 50$

$f_3(x) = f(2 + 50 - 0.5)$
$= 51.5^2$

Extra:

$$P = \frac{TP}{TP+FP}$$

$$R = \frac{TP}{TP+FN}$$

$$TPR = \frac{TP}{TP+FN}$$

$$FPR = \frac{FP}{FP+TN}$$

P O

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

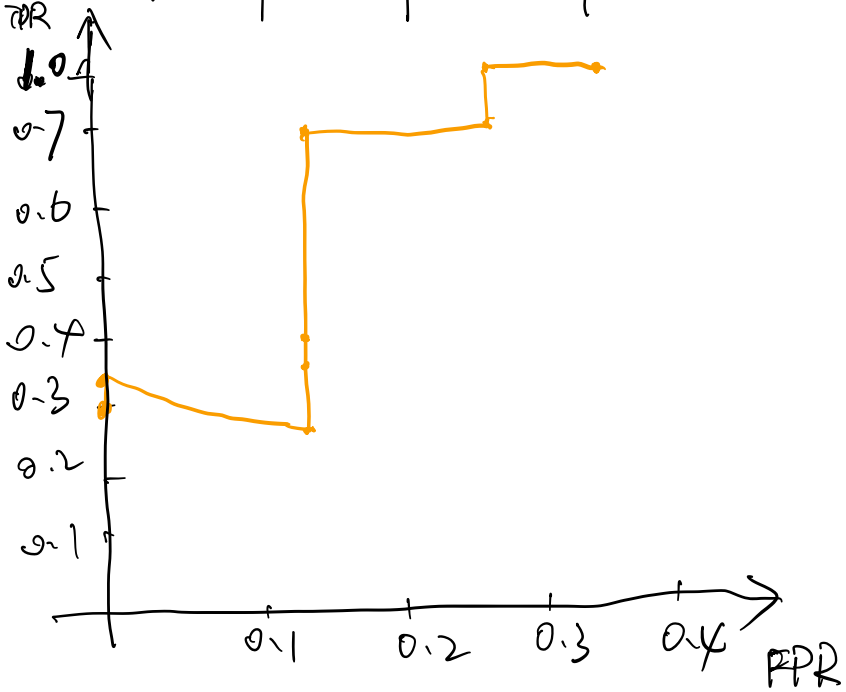
0.9 TP FP 1 5

FN TN 0 14

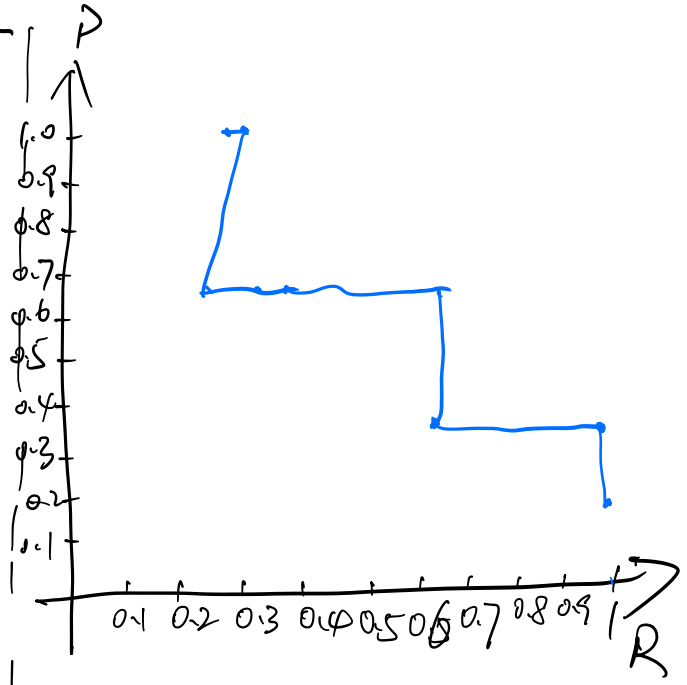
0.3 4 2

11 3

threshold	P	R	TPR	FPR
0.1	1	0.3	0.3	0
0.2	1	0.3 15	0.3 15	0
0.3	0.67	0.27	0.27	0.14
0.4	0.67	0.36	0.36	0.14
0.5	0.67	0.4	0.4	0.14
0.6	0.67	0.67	0.67	0.14
0.7	0.33	0.67	0.67	0.28
0.8	0.33	1	1	0.28
0.9	0.17	1	1	0.36



ROC



PR