CSE514, Fall 2022, HW 2   Name: Ying(Fiona) Xu   Student ID: 510576

Note: This homework is worth a total of 15 points

Q1 (6pts): A friend informs you that a casino is using loaded dice, such that:

$$loaded_k = prob(roll = k) = \begin{cases} 1/_{12} & if\ k \in \{1,2,3\} \\ 1/_{4} & if\ k \in \{4,5,6\} \\ 0 & otherwise \end{cases}$$

Q1a: What is the entropy of a roll at this casino?
Please use log base 2

$$H(loaded) = -\sum p_i \log_2 p_i = -\left[\frac{1}{12} \cdot \log_2\left(\frac{1}{12}\right) \cdot 3 + \frac{1}{4} \cdot \log_2\left(\frac{1}{4}\right) \cdot 3\right]$$

$$= \frac{1}{4}(\log_2 4 + \log_2 3) + \frac{1}{4} \cdot 3 \cdot 2 = \frac{1}{4}(2 + \log_2 3) + \frac{3}{2} = 2 + \frac{1}{4}\log_2 3$$

Q1b: Imagine your friend is right, but you choose to give the casino the benefit of the doubt and assume fair dice. What's the KL divergence of fair dice (Q) from the true distribution (P)? i.e. calculate $D_{KL}(loaded\ dice||fair\ dice)$
Please use log base 2

fair dice : $prob(roll = k) = \frac{1}{6}$ if $k \in \{1,2,3,4,5,6\}$

$$D_{KL}(P||Q) = \sum_i P_i \log_2 \frac{P_i}{Q_i} = \sum_i P_i \log_2 P_i - \sum_i P_i \log_2 Q_i.$$

$$= -\left(2 + \frac{1}{4}\log_2 3\right) - \left[\frac{1}{12}\log_2 \frac{1}{6} \cdot 3 + \frac{1}{4} \cdot \log_2 \frac{1}{6} \cdot 3\right] = \log_2 6 - \frac{1}{4}\log_2 3 - 2$$

$$= \frac{3}{4}\log_2 3 - 1$$

Q1c: Imagine you choose to believe your friend, but it turned out the casino has since switched back to fair dice. What's the KL divergence of the loaded dice (Q) from the true distribution (P)? i.e. calculate $D_{KL}(fair\ dice||loaded\ dice)$
Please use log base 2

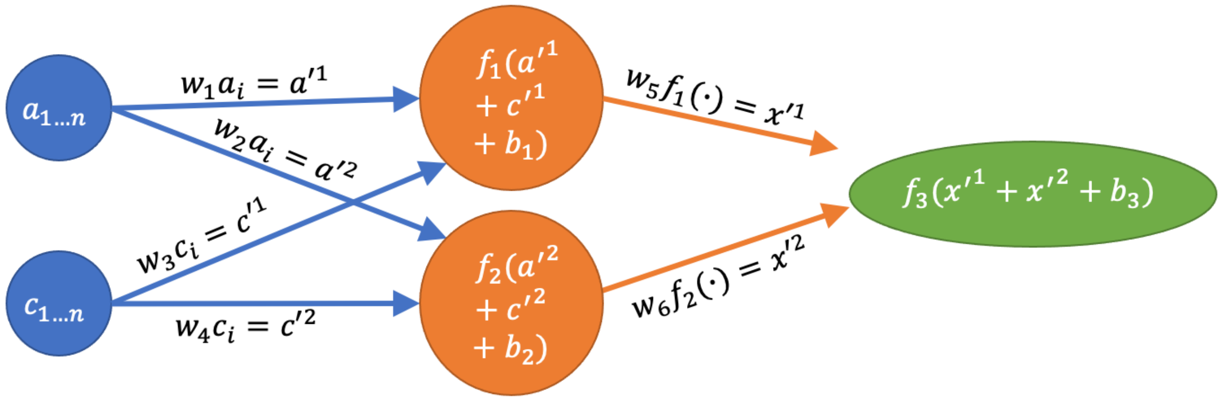
$$D_{KL}(P||Q) = \sum_i P_i \log_2 \frac{P_i}{Q_i} = \sum_i P_i \log_2 P_i - \sum_i P_i \log_2 Q_i$$

$$= \frac{1}{6} \cdot \log_2\left(\frac{1}{6}\right) \cdot 6 - \left[\frac{1}{6} \cdot \log_2 \frac{1}{12} \cdot 3 + \frac{1}{6} \cdot \log_2 \frac{1}{4} \cdot 3\right]$$

$$= -\log_2 6 + \frac{1}{2}\log_2 12 + \frac{1}{2}\log_2 4 = -(1 + \log_2 3) + \frac{1}{2}(2 + \log_2 3) + 1$$

$$= 1 - \frac{1}{2}\log_2 3$$

Q2 (5pts): 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?

Q2a: Data point: a = 0.5, c = 0.5

a'^1 value: -7.5 a'^2 value: -1.5

c'^1 value: -1

Q2b: Data point: a = 1, c = 0

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

f_2 value: -0.2

Q2c: Data point: a = 0, c = 1

x'^1 value: 2 x'^2 value: 50

f_3 value: 2652.25

Q3 (4pts): Given a test data point:

Height = 200

Weight = 200

And the training dataset in the table below, use kNN classification with k=1, k=3, and k=5 to label the test data point. Break ties by increasing k by 1.
Show your work by filling in the table and writing in the model's class label predictions.

Class	Height	Weight	Manhattan Distance from test sample
1	105	114	181
1	92	169	139
1	87	140	173
2	111	109	180
2	79	44	277
2	92	55	253
3	265	331	196
3	330	284	214
3	185	309	124

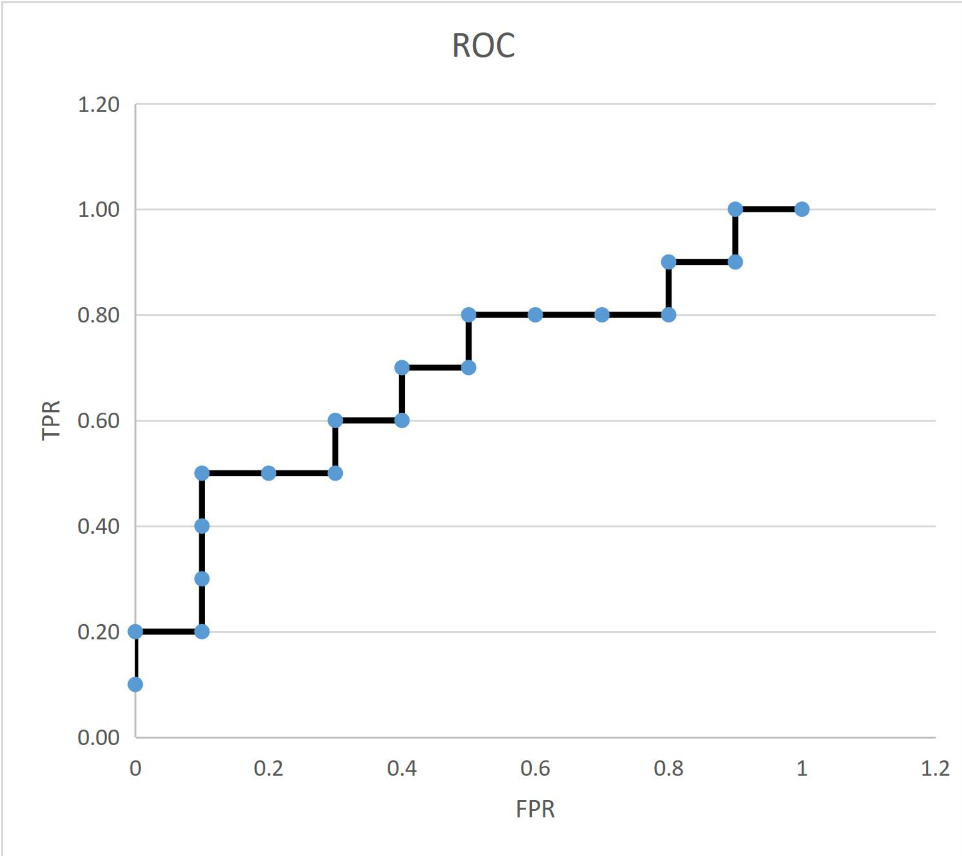
Model predictions for:

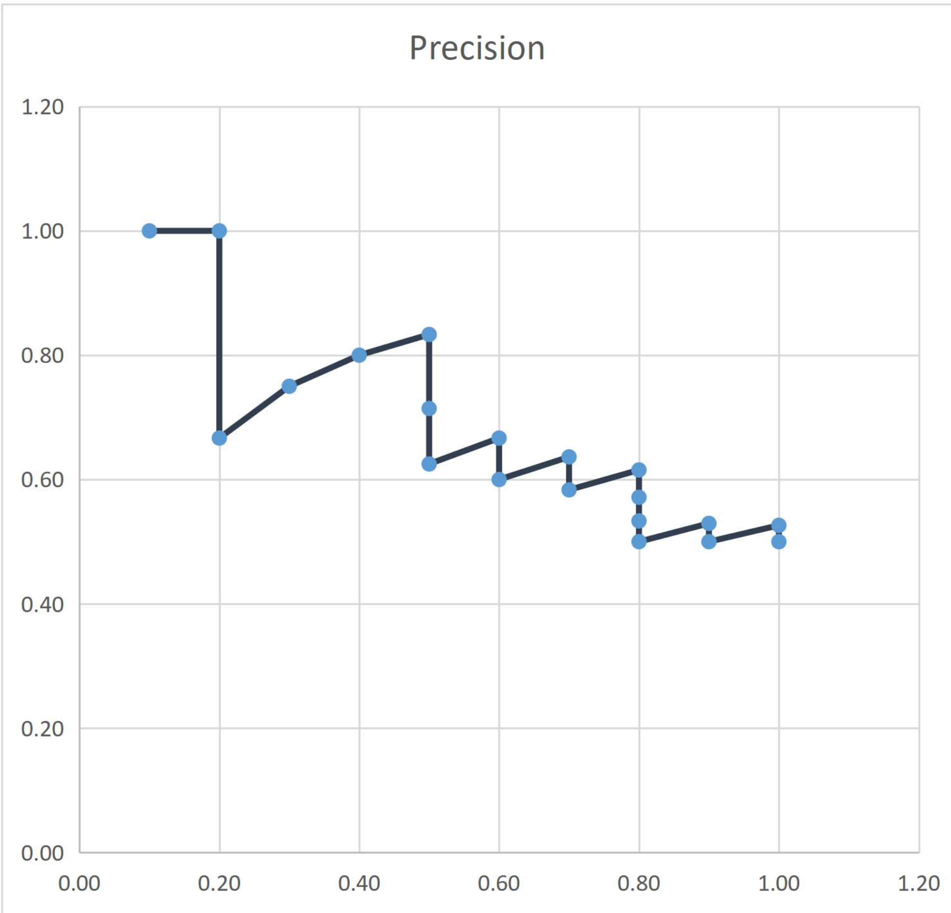
k = 1 ____3____ k = 3 ____1____ k = 5 ____1____

Extra Credit

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

		FP/(FP+TN)	TP/(TP+FN)	TP/(TP+FP)
	Threshold	FPR	TPR(Recall)	Precision
1	0.95	0.00	0.10	1.00
2	0.85	0.00	0.20	1.00
3	0.80	0.10	0.20	0.67
4	0.67	0.10	0.30	0.75
5	0.65	0.10	0.40	0.80
6	0.60	0.10	0.50	0.83
7	0.58	0.20	0.50	0.71
8	0.54	0.30	0.50	0.63
9	0.52	0.30	0.60	0.67
10	0.51	0.40	0.60	0.60
11	0.45	0.40	0.70	0.64
12	0.40	0.50	0.70	0.58
13	0.38	0.50	0.80	0.62
14	0.35	0.60	0.80	0.57
15	0.33	0.70	0.80	0.53
16	0.30	0.80	0.80	0.50
17	0.28	0.80	0.90	0.53
18	0.27	0.90	0.90	0.50
19	0.26	0.90	1.00	0.53
20	0.18	1.00	1.00	0.50


ROC
TPR
FPR
1.20
1.00
0.80
0.60
0.40
0.20
0.00
0
0.2
0.4
0.6
0.8
1
1.2


Precision
1.20
1.00
0.80
0.60
0.40
0.20
0.00
0.00
0.20
0.40
0.60
0.80
1.00
1.20