HW#2, Byeongchan Gwak, 501026, b.wustl@edu

Q1a.

$\boxed{Q_1 a}$ $H(\text{loaded}) = -\sum p_i \log_2 p_i$

$$= -\left(\frac{1}{12} \log_2 \frac{1}{12} \times 3 + \frac{1}{4} \log_2 \frac{1}{4} \times 3 \right)$$

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

$$= -\left(\frac{3}{12} (-\log_2 12) + \frac{3}{4} (-2) \right)$$

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

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

Q1b.

$\boxed{Q.b}$ $D_{KL}(\text{loaded} \,||\, \text{fair})$

$= H(\text{loaded}, \text{fair}) - H(\text{loaded})$

first, we need to get $H(\text{loaded}, \text{fair})$

Using Cross entropy formula

$H(P,Q) = -\sum p_i \log_2(Q_i)$

$$H(\text{loaded}, \text{fair}) = -\sum (\text{loaded}_i) \log_2 (\text{fair}_i)$$

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

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

$$= -\left(-\frac{1}{4} \log_2 6 - \frac{3}{4} \log_2 6 \right)$$

$$= \log_2 6 = \log_2 (2 \times 3) = \log_2 3 + 1$$

$D_{KL}(\text{loaded} \,||\, \text{fair})$

$$= H(\text{loaded}, \text{fair}) - H(\text{loaded})$$

$$= \log_2 3 + 1 - \left(\frac{1}{4} \log_2 3 + 2 \right)$$

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

Q1c.

$\boxed{Q_{1}c}$ $D_{KL}(\text{fair} \| \text{loaded})$

$= H(\text{fair}, \text{loaded}) - H(\text{fair})$

$H(\text{fair}) = -\sum p_i \log_2 p_i$

$= -\left(\frac{1}{6}\log_2 \frac{1}{6} \times 6\right) = -\log_2 \frac{1}{6} = \log_2 6$

$= \log_2(2 \times 3) = \log_2 3 + 1$

$H(\text{fair}, \text{loaded}) = -\sum (\text{fair}_i) \log_2 (\text{loaded}_i)$

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

$= -\frac{1}{2}\left(\log_2 \frac{1}{12} + \log_2 \frac{1}{4}\right)$

$= -\frac{1}{2}\left(\log_2 1 - \log_2 12 + \log_2 1 - \log_2 4\right)$

$= -\frac{1}{2}\left(-\log_2 12 - 2\right) = \frac{1}{2}\log_2 12 + 1$

$= \frac{1}{2}\left(2 + \log_2 3\right) + 1$

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

$D_{KL}(\text{fair} \| \text{loaded})$

$= H(\text{fair}, \text{loaded}) - H(\text{fair})$

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

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

Q2a.

$\boxed{Q2a}$ $a = 0.5 \quad c = 0.5$

$$a'^{1} = W_1 a = -15 \times 0.5 = -7.5$$

$$a'^{2} = W_2 a = -3 \times 0.5 = -1.5$$

$$c'^{1} = W_3 c = -2 \times 0.5 = -1$$

Q2b.

$\boxed{Q2b}$ $a = 1 \quad c = 0$

$$a'^{1} = W_1 a = -15 \times 1 = -15$$

$$a'^{2} = W_2 a = -3 \times 1 = -3$$

$$c'^{1} = W_3 c = -2 \times 0 = 0$$

$$c'^{2} = W_4 c = 4 \times 0 = 0$$

$$f_1 = f_1(a'^{1} + c'^{1} + b_1) = f_1(-15 + 0 + 4) = f_1(-11)$$

$$= \max(0.1 \times -11, -11) = \max(-1.1, -11)$$

$$= -1.1$$

$$f_2 = f_2(a'^{2} + c'^{2} + b_2) = f_2(-3 + 0 + 1) = f_2(-2)$$

$$= \max(0.1 \times -2, -2) = \max(-0.2, -2)$$

$$= -0.2$$

Q2c.

$\boxed{Q_2C}$ $\quad a=0 \quad c=1$

$a'^{1} = W_1 a = -15\times 0 = 0$

$a'^{2} = W_2 a = -3\times 0 = 0$

$c'^{1} = W_3 c = -2\times 1 = -2$

$c'^{2} = W_4 c = 4\times 1 = 4$

$f_1 = f_1(a'^{1} + c'^{1} + b_1) = f_1(0 + -2 + 4) = f_1(2)$

$= \max(0.1\times 2, 2) = \max(0.2, 2) = 2$

$f_2 = f_2(a'^{2} + c'^{2} + b_2) = f_2(0 + 4 + 1) = f_2(5)$

$= \max(0.1\times 5, 5) = \max(0.5, 5) = 5$

$x'^{1} = W_5 f_1 = 1\times 2 = 2$

$x'^{2} = W_6 f_2 = 10\times 5 = 50$

$f_3 = f_3(x'^{1} + x'^{2} + b_3) = f_3(2 + 50 + (-0.5))$

$= f_3(51.5) = (51.5)^2$

Q3.

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.

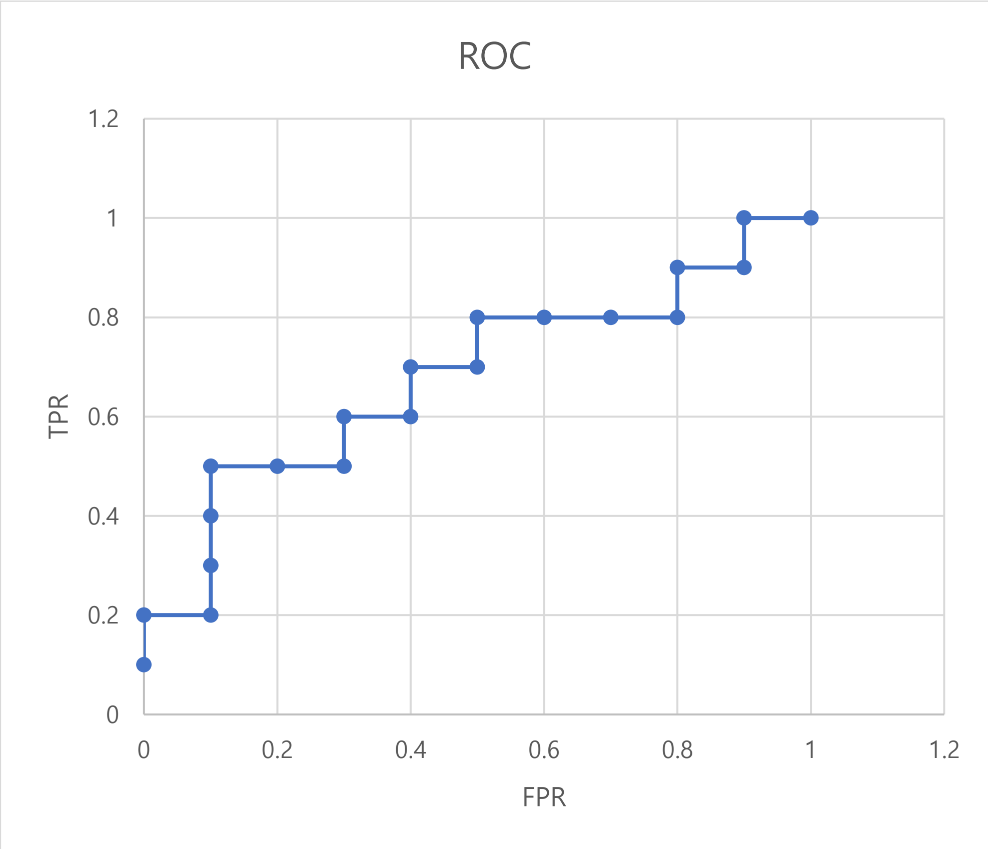
Class	Height	Weight	Manhattan Distance from test sample
1	105	114	95 + 86 = 181 (5)
1	92	169	108 + 31 = 139 (2)
1	87	140	113 + 60 = 173 (3)
2	111	109	89 + 91 = 180 (4)
2	79	44	121 + 156 = 277
2	92	55	108 + 145 = 253
3	265	331	65 + 131 = 196
3	330	284	130 + 84 = 214
3	185	309	15 + 109 = 124 (1)

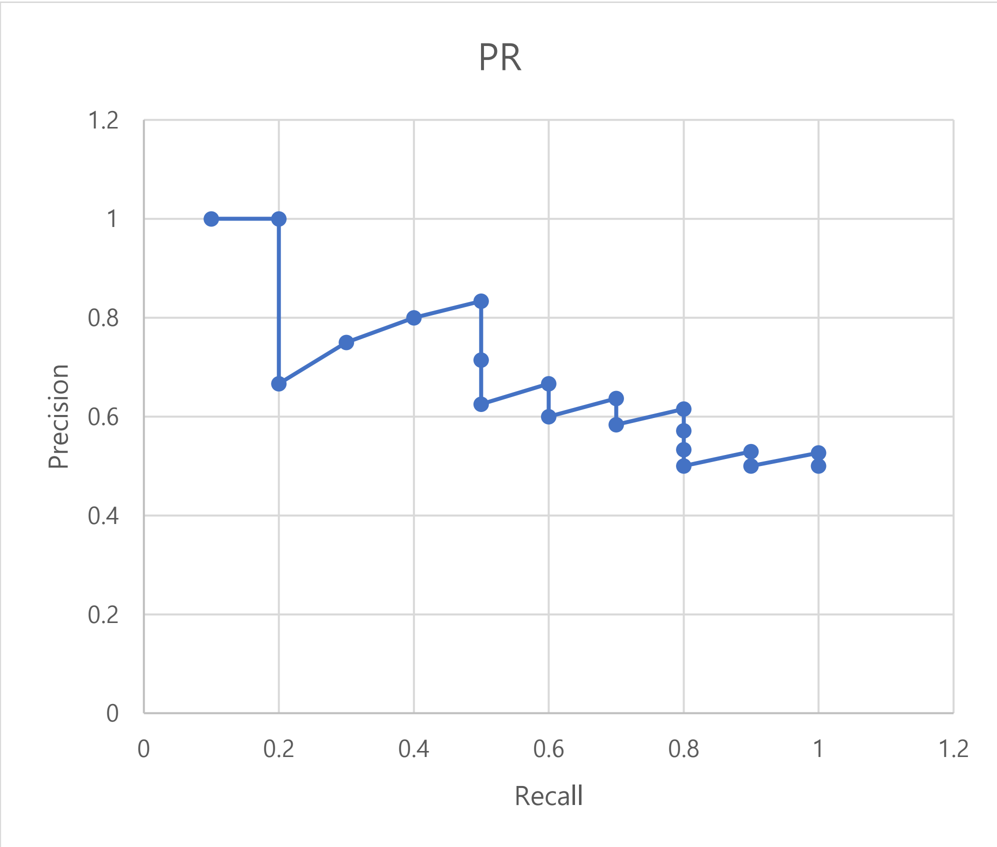
Model predictions for:

k = 1 __3__ k = 3 __1__ k = 5 __1__

Extra.

Threshold	FPR	TPR	Precision
0.95	0	0.1	1
0.85	0	0.2	1
0.8	0.1	0.2	0.666667
0.67	0.1	0.3	0.75
0.65	0.1	0.4	0.8
0.6	0.1	0.5	0.833333
0.58	0.2	0.5	0.714286
0.54	0.3	0.5	0.625
0.52	0.3	0.6	0.666667
0.51	0.4	0.6	0.6
0.45	0.4	0.7	0.636364
0.4	0.5	0.7	0.583333
0.38	0.5	0.8	0.615385
0.35	0.6	0.8	0.571429
0.33	0.7	0.8	0.533333
0.3	0.8	0.8	0.5
0.28	0.8	0.9	0.529412
0.27	0.9	0.9	0.5
0.26	0.9	1	0.526316
0.18	1	1	0.5


ROC
TPR
FPR
0
0.2
0.4
0.6
0.8
1
1.2


PR
Precision
Recall
0
0.2
0.4
0.6
0.8
1
1.2