Problem 1:

a)

b)

P_1	P_2	L1-Norm	L2-Norm
A	A	0	0
A	B	1.5	1.1180
A	C	1.5	1.5
A	D	4.5	3.201
A	E	7.0	5.147
A	F	6.0	4.743
B	A	1.5	1.118
B	B	0	0
B	C	3.0	2.236
B	D	4.0	3.162
B	E	6.5	4.609
B	F	5.5	4.03
C	A	1.5	1.5
C	B	3.0	2.236
C	C	0.0	0
C	D	3.0	2.236
C	E	5.5	4.609
C	F	4.5	4.5
D	A	4.5	3.201
D	B	4.0	3.162
D	C	3.0	2.230
D	D	0.0	0
D	E	2.5	2.5
D	F	3.5	2.692
E	A	7.0	5.147
E	B	6.5	4.609
E	C	5.5	4.609
E	D	2.5	2.5
E	E	0.0	0
E	F	1.0	1.0
F	A	6.0	4.743
F	B	3.5	4.031
F	C	4.5	4.5
F	D	3.5	2.692
F	E	1.0	1.0
F	F	0	0.0

nearest neighbor of	w.r.t	L1-Norm	L2-Norm
A		B & C	B
B		A	A
C		A	A
D		E	C
E		F	F
F		E	E

c) based on this small dataset: the ~~most~~ classifications are the same, except that L2-Norm has 1 wrong prediction for D

Problem 2

a) let $c(x)$ denote the class of x

then $P(c(x_{new}) = A) = \frac{N_A}{N_A+N_B+N_C} = \frac{16}{112}$

$P(c(x_{new}) = B) = \frac{N_B}{N_A+N_B+N_C} = \frac{32}{112}$

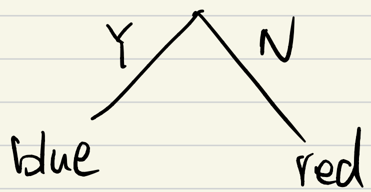
$P(c(x_{new}) = C) = \frac{N_C}{N_A+N_B+N_C} = \frac{64}{112}$

so x_{new} is most likely to be classified as C

b). | th x_{new} is also most likely to be classified as C. since class C has most instances in this dataset, i.e. the 2D region is mostly coverd by class C, which implys that the distance between instances of C and X_{new} is also more likely to be shorter.

Problem 3

there is one : $x_2 > x_1$?



Problem 4

a).

$$iH(y) = -\sum_{c \in C} \pi_c \cdot \log \pi_c \quad \text{where } \pi_c = P(y=c|y)$$

$$= -\left(\frac{2}{5} \cdot \log\frac{2}{5} + \frac{3}{5} \cdot \log\frac{3}{5}\right)$$

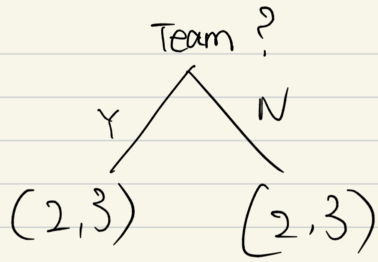
$$= -\frac{1}{5}\left(\log\frac{20}{125} + \log\frac{27}{125}\right)$$

$$= -\frac{1}{5}\left(\log\frac{540}{125^2}\right)$$

$$\approx 0.971$$

$$i_H(\text{root}) = -\left(\tfrac{2}{5}\cdot\log\tfrac{2}{5} + \tfrac{3}{5}\cdot\log\tfrac{3}{5}\right) = A \approx 0.971$$

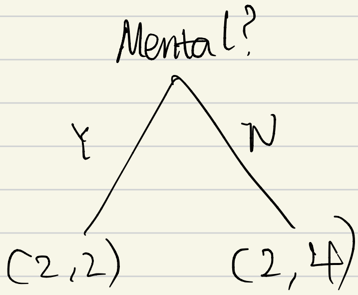
b).



$$i_H(l) = -\left(\tfrac{2}{5}\cdot\log\tfrac{2}{5} + \tfrac{3}{5}\cdot\log\tfrac{3}{5}\right)$$

$$i_H(r) = -\left(\tfrac{2}{5}\cdot\log\tfrac{2}{5} + \tfrac{3}{5}\cdot\log\tfrac{3}{5}\right)$$

$$\Delta = A - \tfrac{1}{2}\cdot A - \tfrac{1}{2}\cdot A = 0$$



$$i_H(l) = -\left(\tfrac{1}{2}\log\tfrac{1}{2} + \tfrac{1}{2}\cdot\log\tfrac{1}{2}\right) = -\log\tfrac{1}{2} = \log 2 = 1$$

$$i_H(r) = -\left(\tfrac{1}{3}\log\tfrac{1}{3} + \tfrac{2}{3}\log\tfrac{2}{3}\right) = -\tfrac{1}{3}\left(\log\tfrac{1}{3} + \log\tfrac{4}{9}\right)$$

$$= -\tfrac{1}{3}\left(\log\tfrac{4}{27}\right)$$

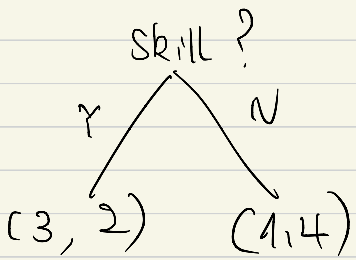
$$= 0.9183$$

$$\Delta = A - \tfrac{2}{5}\cdot 1 - \tfrac{3}{5}\cdot 0.9183$$

$$= 0.971 - 0.4 - 0.6 \times 0.9183$$

$$= 0.971 - 0.95098 = 0.02002$$

$$\approx 0.02$$



$$i_H(l) = -\left(\tfrac{3}{5}\cdot\log\tfrac{3}{5} + \tfrac{2}{5}\cdot\log\tfrac{2}{5}\right) = 0.971$$

$$i_H(r) = -\left(\tfrac{1}{5}\cdot\log\tfrac{1}{5} + \tfrac{4}{5}\cdot\log\tfrac{4}{5}\right) = -\tfrac{1}{5}\left(\log\tfrac{1}{5} + 4\log\tfrac{4}{5}\right)$$

$$= -\tfrac{1}{5}\left(\log\tfrac{1}{5} + \log\tfrac{256}{625}\right)$$

$$= -\tfrac{1}{5}\log\tfrac{256}{3125}$$

$$= 0.722$$

$$\Delta = 0.971 - \tfrac{1}{2}\cdot 0.971 - \tfrac{1}{2}\cdot 0.722$$

$$= 0.1245$$

Since "skill ?" has largest improvement, I'd choose this as the decision.