

Task 1:

a. Let $K_1=80$, $K_2=20$, $K=K_1+K_2=100$

$$\therefore H(A) = H\left(\frac{K_1}{K}, \frac{K_2}{K}\right) = -\frac{K_1}{K} \log \frac{K_1}{K} - \frac{K_2}{K} \log \frac{K_2}{K}$$

$$= -0.8 \cdot \log 0.8 - 0.2 \cdot \log 0.2 = 0.7219$$

b. Let $K_3=20$, $K_4=15$, $K_5=60$, $K_6=5$.

$$\therefore Gain_A = H(A) - \sum_i \frac{K_i}{K} \cdot H(A_i) = H(A) - [\frac{K_3+K_4}{K} \cdot H\left(\frac{K_3}{K_3+K_4}, \frac{K_4}{K_3+K_4}\right) + \frac{K_5+K_6}{K} \cdot H\left(\frac{K_5}{K_5+K_6}, \frac{K_6}{K_5+K_6}\right)]$$

$$= 0.7219 - [\frac{35}{100} \cdot H(20/35, 15/35) + \frac{15}{100} \cdot H(60/65, 5/65)]$$

$$= 0.7219 - [0.35 \cdot 0.9852 + 0.15 \cdot 0.3912]$$

$$= 0.1228$$

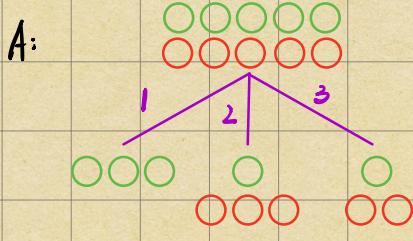
c. Node E is a duplicate node. And people in node E they all those will not eat.

$$Gain_E = H(E) - \sum_i \frac{K_i}{K} H(E_i) = 0 - [0 + 0] = 0$$

d. $A \rightarrow B \rightarrow D$,

The test case will end up in Node D and output: will wait.

Task 2:

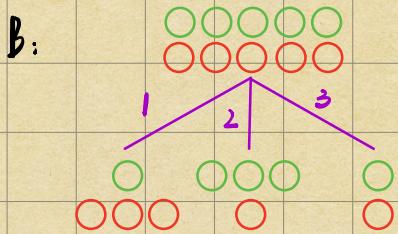


$$Gain_A = H(A) - \sum_i \frac{K_i}{K} H(A_i)$$

$$= 1 - [0 + \frac{4}{10}(-\frac{1}{4} \log \frac{1}{4} - \frac{3}{4} \log \frac{3}{4}) + \frac{3}{10}(-\frac{1}{3} \log \frac{1}{3} - \frac{2}{3} \log \frac{2}{3})]$$

$$= 1 - [\frac{2}{5}(0.5 + 0.3112) + \frac{3}{10}(0.5283 + 0.3900)]$$

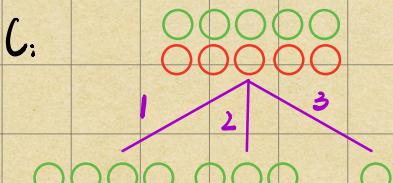
$$= 1 - 0.5997 \approx 0.4$$



$$Gain_B = H(B) - \sum_i \frac{K_i}{K} H(B_i)$$

$$= 1 - [\frac{2}{5}(0.5 + 0.3112) + \frac{2}{5}(0.5 + 0.3112) + \frac{1}{5} \cdot 1]$$

$$= 1 - 0.649 = 0.15104$$



$$Gain_C = H(C) - \sum_i \frac{K_i}{K} H(C_i)$$

$$= 1 - [\frac{1}{2}(-\frac{4}{5} \log \frac{4}{5} - \frac{1}{5} \log \frac{1}{5}) + \frac{2}{5} \cdot 0.8112 + 0]$$

$$= 1 - [\frac{1}{2} \cdot 0.7219 + \frac{2}{5} \cdot 0.8112] = 0.3148$$

○ ○

\therefore Attribute A achieves the highest information gain at the root

Task 3:

a. Lowest entropy value: 0, when all examples are from same class.

$$(A:1000, B:0, C:0, D:0, H = -1 \cdot \log_2 1 - 0 - 0 - 0 = 0)$$

Highest entropy value when 4 classes have equal number ($A:250, B:250, C:250, D:250$)

$$\text{where } H = -\frac{1}{4} \log \frac{1}{4} \cdot 4 = -\log \frac{1}{4} = 2$$

b. Base on the answer of part a:

$$\text{Gain}_{\text{low}} = 0 - 0 - 0 - 0 - 0 = 0$$

$$\text{Gain}_{\text{high}} = 2 - [\frac{1}{4} \cdot 4 \cdot 0] = 2$$

Task 4:

Yes, reverse operation then we can obtain a 72% accuracy classifier.

Task 5:

1. calculate means and variances:

$$\mu_{A_1} = \frac{15+25}{2} = 20 \quad \delta_{A_1}^2 = \frac{1}{2-1} [(15-20)^2 + (25-20)^2] = 50$$

$$\mu_{A_2} = \frac{28+32}{2} = 30 \quad \delta_{A_2}^2 = \frac{1}{2-1} [(30-28)^2 + (32-30)^2] = 8$$

$$\mu_{B_1} = \frac{20+32+25}{3} = 25.67 \quad \delta_{B_1}^2 = \frac{1}{3-1} [(20-25.67)^2 + (32-25.67)^2 + (25-25.67)^2] = 36.33$$

$$\mu_{B_2} = \frac{10+15+15}{3} = 13.33 \quad \delta_{B_2}^2 = \frac{1}{3-1} [(10-13.33)^2 + (15-13.33)^2 + (15-13.33)^2] = 8.33$$

2. Calculate $P(\text{Attribute} | \text{Class})$ by $\frac{1}{\sqrt{2\pi}\delta} \exp\left(-\frac{(x_i - \mu)^2}{2\delta^2}\right)$

$$P(A1|A) = \frac{1}{\sqrt{2\pi \cdot 50}} \cdot \exp\left(-\frac{x-20}{\sqrt{50}}\right)$$

$$P(A2|A) = \frac{1}{\sqrt{2\pi \cdot 8}} \cdot \exp\left(-\frac{x-30}{\sqrt{8}}\right)$$

$$P(A1|B) = \frac{1}{\sqrt{2\pi \cdot 36.33}} \cdot \exp\left(-\frac{x-25.67}{\sqrt{36.33}}\right)$$

$$P(A2|B) = \frac{1}{\sqrt{2\pi \cdot 8.33}} \cdot \exp\left(-\frac{x-13.33}{\sqrt{8.33}}\right)$$

3. Calculate $P(\text{Class})$

$$P(A) = \frac{2}{5} = 0.4$$

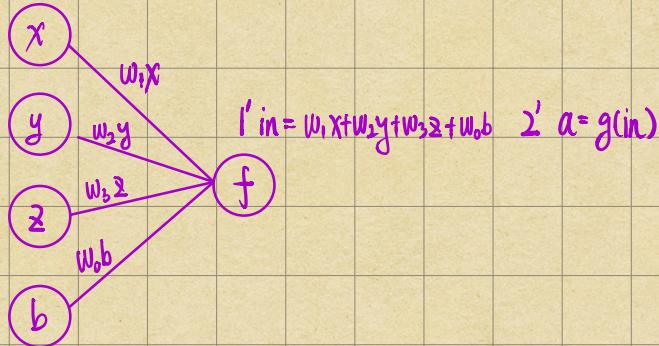
$$P(B) = \frac{3}{5} = 0.6$$

4. Pseudo-Bayes Classifier:

$$\hat{A} = \underset{A}{\operatorname{argmax}} P(A) \cdot P(A_1|A) \cdot P(A_2|A)$$

$$\hat{B} = \underset{B}{\operatorname{argmax}} P(B) \cdot P(A_1|B) \cdot P(A_2|B)$$

Task b:



$\therefore 4x - 7y + 2z = b$, and we want the neural network always return 1.

\therefore We need to make sure $4x - 7y + 2z - b > 0$

0 0 ↗

According to the question, the bias input is +1 and we use "Relu" as activation function

As a result, we can set $w_1=4$, $w_2=-7$, $w_3=2$, $w_0=1$. to simulate $4x - 7y + 2z + 1$, which is definitely larger than 0.