data = [attri class]

1. Probability

(a) L·H·S

$$P(A/B\cap X) = P(A/Q) \text{ where } Q = B\cap X$$

$$= \frac{P(A\cap Q)}{P(Q)} = \frac{P(A\cap B\cap X)}{P(B\cap X)} \quad \text{--- (1)}$$

From R·H·S

$$P(B/A\cap X) = P(B/S) \text{ where } S = A\cap X$$

$$= \frac{P(B\cap S)}{P(S)} = \frac{P(B\cap A\cap X)}{P(A\cap X)}$$

But $P(A\cap B\cap X) = P(B\cap A\cap X) = P(B/A\cap X)\cdot P(A\cap X)$ — (2)

(2) in (1)

$$P(A/B\cap X) = \frac{P(B/A\cap X)\cdot P(A\cap X)}{P(B\cap X)}$$

$$= \frac{P(B/A\cap X)\cdot [P(A/X)\cdot P(\cancel{X})]}{P(B/X)\cdot P(\cancel{X})}$$

$$= \frac{P(A/X)\cdot P(B/A\cap X)}{P(B/X)}$$

(b) $$P(A\cap B/X) = \frac{P(A\cap B\cap X)}{P(X)}$$

$$= \frac{P(A/B\cap X)\cdot P(B\cap X)}{P(X)}$$

$$= \frac{P(A/B\cap X)\cdot P(B/X)\cdot P(\cancel{X})}{P(\cancel{X})}$$

$$= P(A/B\cap X)\cdot P(B/X)$$

(C)

Starting from R.H.S

$$P(A/B,C,D) = \frac{P(A\cap B\cap C\cap D)}{P(B\cap C\cap D)} \quad —①$$

$$P(B/C,D) = \frac{P(B\cap C\cap D)}{P(C\cap D)} \quad —②$$

$$P(C/D) = \frac{P(C\cap D)}{P(D)} \quad —③$$

$$\therefore ① \times ② \times ③ \times P(D) = P(A\cap B\cap C\cap D)$$

②

$$P(C/A,B,D) = \frac{P(C\cap A\cap B\cap D)}{P(A\cap B\cap D)} \quad —①$$

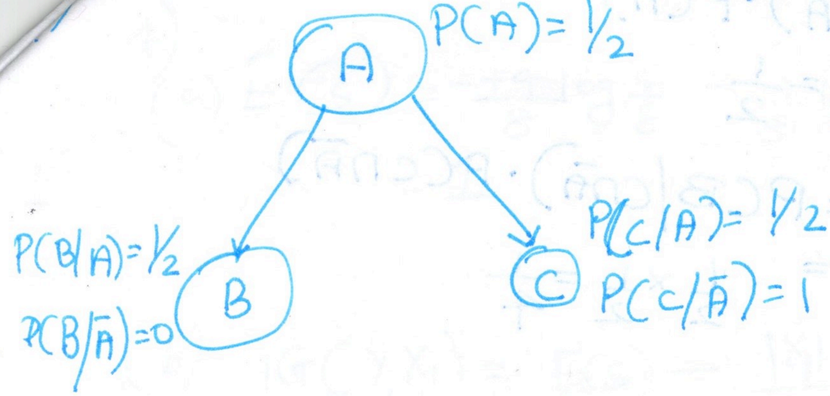
$$P(D/B,A) = \frac{P(D\cap B\cap A)}{P(B\cap A)} \quad —②$$

$$P(A/B) = \frac{P(A\cap B)}{P(B)} \quad —③$$

$$① \times ② \times ③ \times P(B) = \frac{P(C\cap A\cap B\cap D)}{\cancel{P(B)}} \times \cancel{P(B)}$$

$$= P(C\cap A\cap B\cap D)$$

– Both are same.




Let us find the basic for all answers

$$P(B \cap C \cap A) = P(B \cap C / A) \cdot P(A)$$

(or)

$$= P(B / C \cap A) \cdot P(C \cap A)$$

$$= P(B / C \cap A) \cdot P(C/A) \cdot P(A)$$

→ Since B C are conditionally independent given 'A' $= P(B/A) \cdot P(C/A) \cdot P(A)$

$$= \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$$

a) $P(A \cap B) = P(B \cap A) = P(B/A) \cdot P(A) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

b) $P(C) = P(C \cap A) + P(C \cap \bar{A})$

$$= P(C/A) \cdot P(A) + P(C/\bar{A}) \cdot P(\bar{A})$$

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

c) $P(C / B \cap A) = P(C/A)$

$P(B \cap C) / P(C)$

$$\frac{P(C \cap B \cap A)}{P(B \cap A)} = \frac{P(C \cap A)}{P(A)} \qquad \therefore P(C \cap A) \cdot P(B \cap A) = \frac{1}{8} \times \frac{1}{2} = \frac{1}{16}$$

$$\Rightarrow P(C/A) \cdot P(A) \cdot P(B/A) \cdot P(A) = \frac{1}{16}$$

$$\Rightarrow P(C/A) \cdot P(A) = \frac{1}{16} \times 2 \times 2 = \frac{1}{4}$$

~~also we know~~
~~$P(B \cap A \cap C) = P(B / A \cap C) \cdot P(A \cap C)$~~

$$P(B \cap C) = P(B \cap C \cap \bar{A}) + P(B \cap C \cap A) \cancel{P(B/C \cap A)} \cdot \quad (1)$$

$$P(C\cap\bar{A}) = P(C/\bar{A})\cdot P(\bar{A})$$
$$= 1\times\frac{1}{2} = \frac{1}{2}$$

Now $P(B\cap C\cap\bar{A}) = P(B|C\cap\bar{A})\cdot P(C\cap\bar{A})$
$$= \frac{1}{2}\times\frac{1}{2} = \frac{1}{4}$$

$\therefore$ from (1)

$$P(B\cap C) = \frac{1}{4}+\frac{1}{8} = \frac{3}{8}$$

$$\therefore P(B/C) = \frac{P(B\cap C)}{P(C)} = \frac{\frac{3}{8}}{\frac{3}{4}} = \frac{1}{2}.$$

(3) Let A – Appearing Green
B – It is actually Green

$$\therefore P(B/A) = \frac{P(B\cap A)}{P(A)} = \frac{P(B\cap A)}{\text{Prob. of yellow} \times \text{misclassify} + \text{Prob. of green classify}}$$

$$= \frac{P(B\cap A)}{\cdot 9\times\cdot 25+\cdot 1\times\cdot 75} = \frac{P(B\cap A)}{0\cdot 3}$$

$$= \frac{\text{Prob that it is Green on it appears green}}{0\cdot 3}$$

$$= \frac{0\cdot 1\times\cdot 75}{0\cdot 3}$$

$$= 0\cdot 25.$$

4)

(a) $E(S) = -\frac{4}{8}\log\frac{4}{8} - \frac{4}{8}\log\frac{4}{8}$

$= \underline{1}$

(b) $IG(Y, X_1) = E(S) - \frac{|X_1'|}{|S|}E(X_1') - \frac{|X_1''|}{|S|}E(X_2'')$ } $S = Y$

$= 1 - \frac{4}{8}\left[-\frac{2}{4}\log\frac{2}{4} - \frac{2}{4}\log\frac{2}{4}\right] - \frac{4}{8}\left[-\frac{2}{4}\log\frac{2}{4} - \frac{2}{4}\log\frac{2}{4}\right]$

$= 0$

$IG(Y, X_2) = 1 - \frac{4}{8}\left[-\frac{2}{4}\log\frac{2}{4} - \frac{2}{4}\log\frac{2}{4}\right] - \frac{4}{8}\left[-\frac{2}{4}\log\frac{2}{4} - \frac{2}{4}\log\frac{2}{4}\right]$

$= 0$

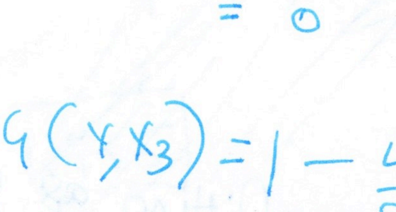
$IG(Y, X_3) = 1 - \frac{4}{8}\left[-\frac{1}{4}\log_2\frac{1}{4} - \frac{3}{4}\log_2\frac{3}{4}\right] - \frac{4}{8}\left[-\frac{1}{4}\log\frac{1}{4} - \frac{3}{4}\log\frac{3}{4}\right]$

$= 0.1887$

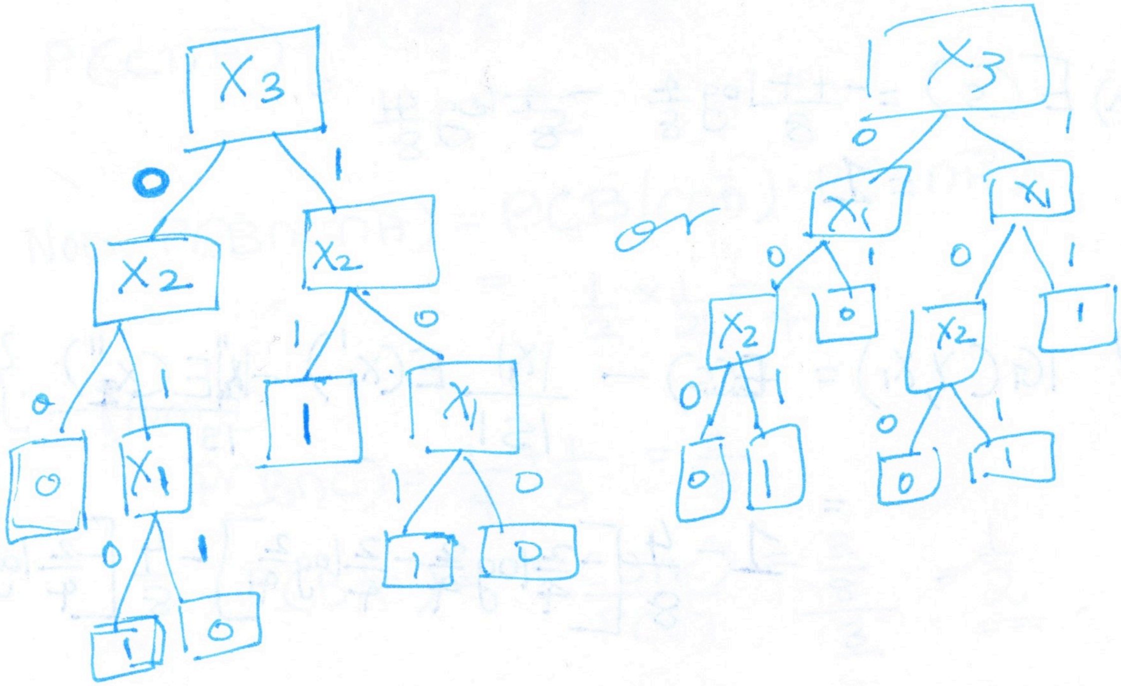
(c) Root should be $\underline{x_3}$ as it has highest infogain

(d) → I would stop when the Entropy = 0 for all the nodes ⇒ No more info to gain

(e)



(e)



(f)

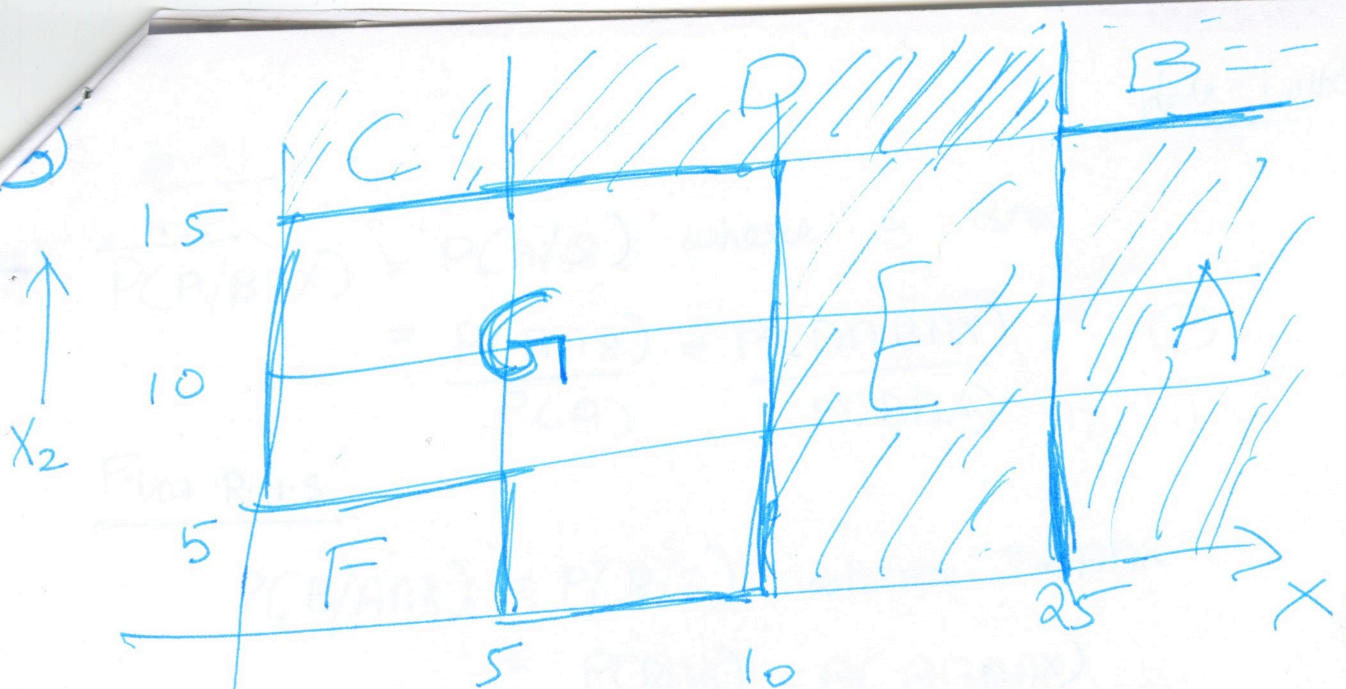
(g) 1

0

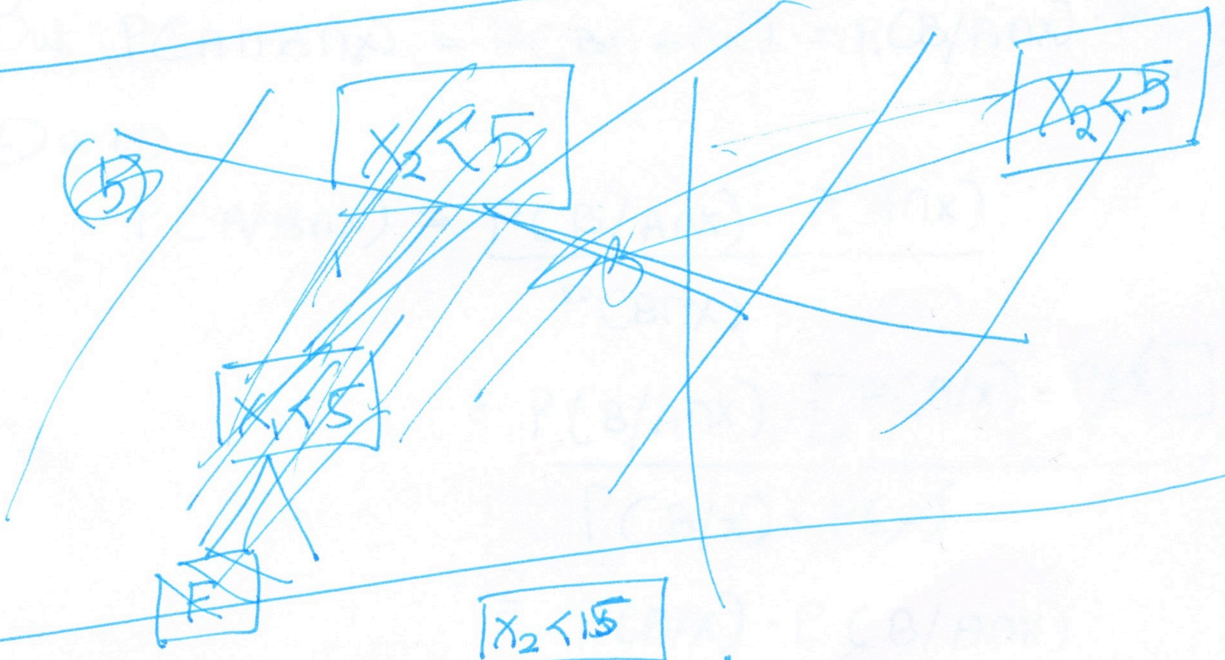
1

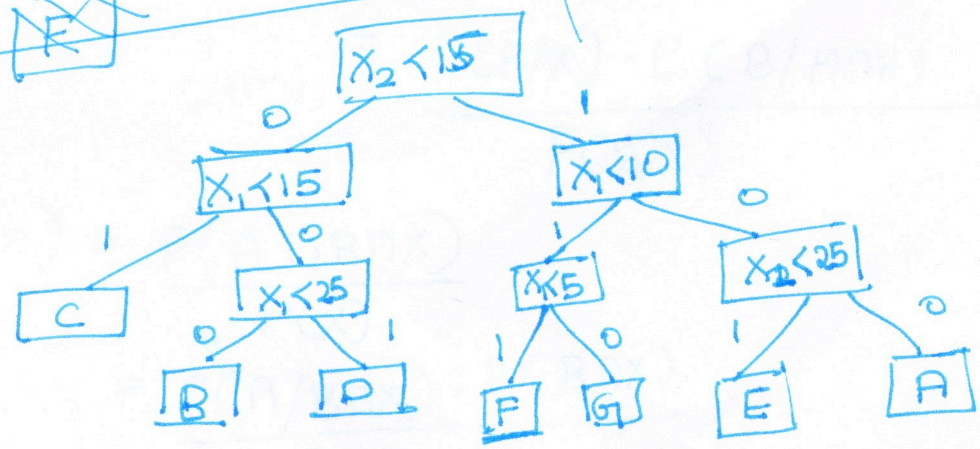
(h) ~~But We need~~ The tree is overfitting as shown in question 'g'. Irrespective of the value of $X_1$ we get the output based only on $X_2$ and $X_3$. We can combat it by creating various ~~dec~~ trees with different number of nodes. Now choosing $\frac{2}{3}$ for training and $1/3$ for testing: we can ~~give~~ determine the best tree that gives best performance.

~~(i)~~ (i)


C
D
B
15
10
5
$X_2$
G
E
A
F
5
10
25
$X_1$

(a)


$X_2 < 5$
$X_1 < 5$
F


$X_2 < 15$
0
1
$X_1 < 15$
$X_1 < 10$
1
0
1
0
C
$X_1 < 25$
$X_1 < 5$
$X_2 < 25$
0
1
1
0
1
0
B
D
F
G
E
A

(b)

(t