1. Gini impurity for the given condition S1 : $x > 10$ to be calculated for all three features for both the classes

(i) Feature a

S1 : $a > 10$

+1	+1	+1	-1	-1	-1	-1	+1
4	5	6	7	8	9	11	12

10

$P_1 = P(\text{class} = +1 \mid a > 10) = \frac{1}{2} = 0.5$

$P_2 = P(\text{class} = -1 \mid a > 10) = \frac{1}{2} = 0.5$

$g(S1) = 1 - P_1^2 - P_2^2 = 1 - 0.25 - 0.25 = 0.5$

(ii) Feature b

S1: $b > 10$

-1	+1	+1	-1	-1	+1	-1
2	3	4	5	6	7	8

+1 (6) 10

$P_1 = P(\text{class} = +1 \mid b > 10) = 0$

$P_2 = P(\text{class} = -1 \mid b > 10) = 0$

$g(S1) = 1 - 0^2 - 0^2 = 1$

(iii) Feature c

S1 : $c > 10$

-1	-1	-1	-1	+1	+1	+1	+1
1	2	3	5	6	7	8	

-1 -1 10

$P_1 = P(\text{class} = +1 \mid c > 10) = 0$

$P_2 = P(\text{class} = -1 \mid c > 10) = 0$

$g(S1) = 1 - 0^2 - 0^2 = 1$

2. Gini impurity for the given condition $x \leq 5$ to be calculated for all three features for both the classes

(i) Feature a

S2 : $a \leq 5$

+1	+1	+1	-1	-1	-1	-1	+1
4	5	6	7	8	9	11	12

$P_1 = P(\text{class} = +1 \mid a \leq 5) = 1$

$P_2 = P(\text{class} = -1 \mid a \leq 5) = 0$

$g(S2) = 1 - P_1^2 - P_2^2 = 1 - 1 - 0 = 0$

(ii) Feature b

S2 : $b \leq 5$

-1	+1	+1	+1	+1 / -1	+1	-1
2	3	4	5	6 / 6	7	8

$P_1 = P(\text{class} = +1 \mid b \leq 5) = \frac{1}{2} = 0.5$

$P_2 = P(\text{class} = -1 \mid b \leq 5) = \frac{1}{2} = 0.5$

$g(S_2) = 1 - P_1^2 - P_2^2 = 1 - 0.25 - 0.25 = 0.5$

(iii) Feature c

S2 : $c \leq 5$

-1 / -1	-1	-1	+1	+1	+1	+1
1 / 1	2	3	5	6	7	8

$P_1 = P(\text{class} = +1 \mid c \leq 5) = \frac{1}{5}$

$P_2 = P(\text{class} = -1 \mid c \leq 5) = \frac{4}{5}$

$g(S_2) = 1 - P_1^2 - P_2^2 = 1 - \frac{1}{25} - \frac{16}{25} = \frac{8}{25} = 0.32$

3. $E(a,5) = P(a>5)\, g(a>5) + P(a\leq 5)\, g(a\leq 5)$

From the dataset, feature a has a value of greater than 5 6 times out of 8

$\therefore E(a,5) = \frac{6}{8}\, g(a>5) + \frac{2}{8}\, g(a\leq 5)$

$g(a>5) = 1 - [P(\text{class} = +1 \mid a>5)]^2 - [P(\text{class} = -1 \mid a>5)]^2$

$= 1 - \left(\frac{2}{6}\right)^2 - \left(\frac{4}{6}\right)^2 = 1 - \left(\frac{1}{3}\right)^2 - \left(\frac{2}{3}\right)^2$

$= 1 - \frac{1}{9} - \frac{4}{9} = 1 - \frac{5}{9} = \frac{4}{9}$

$g(a\leq 5) = 0$ from previous.

$E(a,5) = \frac{6}{8}\left(\frac{4}{9}\right) + \frac{2}{8}(0) = \frac{1}{3}$

$E(b,5) = P(b>5)\, g(b>5) + P(b\leq 5)\, g(b\leq 5)$

$= \frac{4}{8}\, g(b>5) + \frac{4}{8}\, g(b\leq 5)$

$g(b>5) = 1 - [P(\text{class} = +1 \mid b>5)]^2 - [P(\text{class} = -1 \mid b>5)]^2$

$= 1 - \left(\frac{1}{2}\right)^2 - \left(\frac{1}{2}\right)^2 = 1 - 0.25 - 0.25 = 0.5$

$g(b\leq 5) = 0.5$ from previous

$E(b,5) = \frac{1}{2}\left(\frac{1}{2}\right) + \frac{1}{2}\left(\frac{1}{2}\right) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

$$E(c,5) = P(c>5)\, g(c>5) + P(c \leq 5)\, g(c \leq 5)$$

$$= \frac{1}{2}\, g(c>5) + \frac{1}{2}\, g(c \leq 5)$$

$$g(c>5) = 1 - [P(\text{class} = +1 \mid c>5)]^2 - [P(\text{class} = -1 \mid c>5)]^2$$

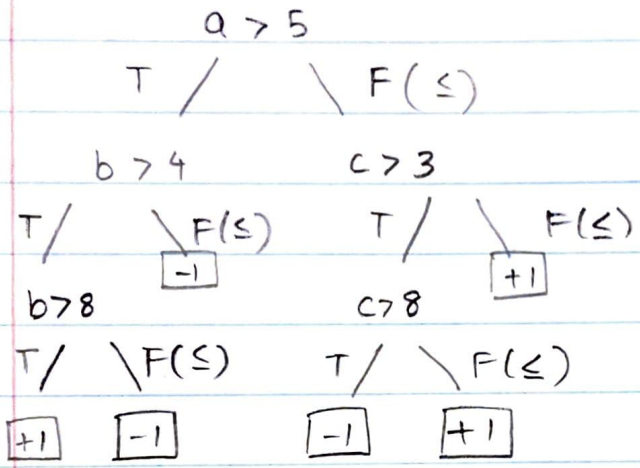
$$= 1 - \left(\frac{1}{3}\right)^2 - (0)^2 = 1 - \frac{1}{9} = \frac{8}{9}$$

$$g(c \leq 5) = \frac{8}{25} \quad \text{from previous}$$

$$E(c,5) = \frac{1}{2}\left(\frac{8}{9}\right) + \frac{1}{2}\left(\frac{8}{25}\right)$$

$$= \frac{4}{9} + \frac{4}{25} = 0.44 + 0.16 = 0.6$$

4.



SAMPLE	PREDICTION
1	-1
2	+1
3	+1
4	-1
5	-1
6	-1
7	-1
8	-1