

1) Calculating the entropy of Target Variable "taste".

$$\begin{aligned} \text{Taste} &= \text{"Meh"} = 5 \\ &= \text{"Yummy"} = 5 \end{aligned}$$

$$\text{Entropy}(S)_{\text{taste}} = -(P_0 \log_2 P_0 + P_0 \log_2 P_0)$$

$$= - \left[\frac{5}{10} \log_2 \frac{5}{10} + \frac{5}{10} \log_2 \frac{5}{10} \right]$$

$$= - \left[\frac{1}{2} \log_2 (0.5) + \frac{1}{2} \log_2 (0.5) \right]$$

$$= - \left[\frac{1}{2} (-1) + \frac{1}{2} (-1) \right]$$

$$= - [-1]$$

$$= 1$$

There is impurity of 100% as it is non-homogenous

2) Calculating the information gain for Visual defects parameter.

We have the Entropy of target variable as '1' for taste variable.
Calculating Entropy for the variable Visual defects.

= Values (Visual defect) = Some, None, Many

$S = 15 \text{ meh}, 5 \text{ yummy} |$

$S_{\text{some}} = 13 \text{ meh}, 0 |$

$S_{\text{none}} = 12 \text{ meh}, 2 \text{ yummy} |$

$S_{\text{many}} = 10, 3 \text{ yummy} |$

$$\text{GAIN}(S, A) = \text{Entropy}(S) - \sum_{v \in \text{values}(A)} \frac{|S_v|}{|S|} \text{Entropy}(S_v)$$

$$\text{Gain}(S, \text{Visual defect}) = \text{Entropy}(S) -$$

$$\left(\frac{3}{10} \text{Entropy}(S_{\text{some}}) + \frac{4}{10} \text{Entropy}(S_{\text{none}}) + \frac{3}{10} \text{Entropy}(S_{\text{many}}) \right)$$

$$\text{Gain}(S, \text{Visual defects}) = 1 - \left(\frac{3}{10} \text{Entropy}(S_{\text{some}}) + \frac{4}{10} \text{Entropy}(S_{\text{none}}) + \frac{3}{10} \text{Entropy}(S_{\text{many}}) \right)$$

$$\text{Entropy}(S_{\text{some}}) = - \left(\frac{3}{3} \log_2 \left(\frac{3}{3} \right) + 0 \right) = 0$$

$$\text{Entropy}(S_{\text{none}}) = - \left(\frac{2}{4} \log_2 \left(\frac{2}{4} \right) + \frac{2}{4} \log_2 \left(\frac{2}{4} \right) \right)$$

$$\text{Entropy}(S_{\text{many}}) = - \left(0 + \frac{3}{3} \log_2 \left(\frac{3}{3} \right) \right)$$

$$\text{Entropy}(S_{\text{some}}) = 0$$

$$\text{Entropy}(S_{\text{none}}) = 1$$

$$\text{Entropy}(S_{\text{many}}) = 0$$

$$\text{Gain}(S, \text{Visual defect}) = \text{Entropy}(S) - \left[\frac{3}{10} \log_2 \left(\frac{3}{10} \right) + \frac{4}{10} \log_2 \left(\frac{4}{10} \right) + \frac{3}{10} \log_2 \left(\frac{3}{10} \right) \right]$$

$$= 1 - \frac{4}{10}$$

$$= \frac{6}{10}$$

$$= 0.60$$

The information gain is 0.6.

3)

3) The value can be =

$$\text{Entropy } H(\text{Taste} | \text{Visual Defect} = \text{some})$$

$$= -\frac{3}{3} \log_2 \left(\frac{3}{3} \right) - 0 \log_2 (0)$$

$$= 0 - 0$$

$$= 0$$

$$\text{Entropy } H(\text{Taste} | \text{Visual Defect} = \text{none})$$

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

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