No. of positive samples in the dataset $(P) = 5$

No. of ~~posit~~ negative samples in the dataset (N) 5

Entropy of the dataset $= -\sum_{i \in c} P_i \log(P_i)$

$$= \left\{ \frac{P}{N+P} \log\left(\frac{P}{P+N}\right) + \frac{N}{P+N} \log\left(\frac{N}{P+N}\right) \right.$$

$$= -\frac{5}{10} \log\left(\frac{5}{10}\right) + \frac{5}{10} \log\left(\frac{5}{10}\right)$$

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

$$= \frac{1}{2} \left\{ \log(2) + \log(2) \right.$$

$$= \frac{2}{2} \log(2) = 1.$$

If Split on SMELL

Woody : Entropy$_W = -\frac{3}{5} \log\left(\frac{3}{5}\right) - \frac{2}{5} \log\left(\frac{2}{5}\right)$

$$= 0.9709$$

Fruity : Entropy$_F = -\frac{3}{5} \log\left(\frac{3}{5}\right) - \frac{2}{5} \log\left(\frac{2}{5}\right)$

$$= 0.9709$$

Information Gain : $1 - \frac{5}{10}(0.9709) - \frac{5}{10}(0.9709)$

$$= 1 - 0.9709$$

$$= 0.029.$$

→ TASTE

Sweet → Entropy S_1 = $\frac{-3}{3} \log\left(\frac{3}{3}\right)$ + entropy

$= + \log 0$

Salty Sour → Entropy S_2 = $\frac{-2}{4} \log\left(\frac{2}{4}\right) - \frac{2}{4} \log\left(\frac{2}{4}\right)$

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

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

Salty Sour → Entropy S_3 = $\frac{-3}{3} \log\left(\frac{3}{3}\right) = 0$

Information Gain = $1 - \frac{3 \times 0}{10} - \frac{3 \times 10}{10} - \frac{4 \times 1}{10}$

$= $ ≈ 0.6

→ PORTION

Large : Entropy $_L$ = $\frac{-4}{5} \log\left(\frac{4}{5}\right) - \frac{1}{5} \log\left(\frac{1}{5}\right)$

$= 0.721$

Small : Entropy $_S$ = $\frac{-4}{5} \log\left(\frac{4}{5}\right) - \frac{1}{5} \log\left(\frac{1}{5}\right)$

$= 0.721$

Entropy Gain = $1 - \frac{5}{10}(0.721) - \frac{5}{10}(721)$

$= + - 0.278$

Taste gets the Maximum ~~Split~~ Information.
Hence the first split is based on Taste

After splitting on Taste.

The Node with all sweet have Negative
Review for all points. Hence that becomes a Leaf Node
with Output Negative

The Node with all Sour have positive review
for all points. Hence that becomes a
Leaf Node with ~~all~~ with positive Outcome

Remaining Data points

P.	Review	Smell	Taste	Texture
	Negative	fruity	Salty	Large
	Negative	fruity	Salty	Large
	Positive	~~Positive~~ fruity	Salty	Small
	Positive	~~Positive~~	Salty	Small

Entropy of the Node. n.

$$Entropy : -\frac{2}{4} \log\left(\frac{2}{4}\right) - \frac{2}{4} \log\left(\frac{2}{4}\right)$$

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

If Split on Smell

$$Fruity : Entropy_f : \frac{2}{4} \log\left(\frac{2}{4}\right) + \frac{2}{4} \log\left(\frac{2}{4}\right)$$
$$\frac{2}{4} \cdot 2 \log\left(\frac{2}{4}\right) = 0 = 1$$

~~S. Weighted. Entropy : log (0)~~

$$Entropy \ gain = 1 - \cancel{} \ 0 = 0$$

If Split on Portion

Large : Entropy$_1$ = $\frac{2}{2}log(^2/_2)$ = 0

Small : Entropy$_2$ = $\frac{2}{2}log(^2/_2)$ = 0

Entropy Gain : 1 - $\frac{2 \times 0}{4}$ - $\frac{2 \times 0}{4}$

$= 1 - 0 = 1$

Hence the Next Split is based on Portion

Final Decision Tree