① Entropy

$$H(Y) = -\sum_{i=1}^{K} P(Y=y_i) \log_2 P(Y=y_i)$$

For variable Taste:

$P(Y = \text{Meh}) = 5/10 = \frac{1}{2}$

$P(Y = \text{Yummy}) = 5/10 = \frac{1}{2}$

$$\rightarrow H(Y) = -\left(\frac{1}{2} \cdot \log_2 \frac{1}{2} + \frac{1}{2} \cdot \log_2 \frac{1}{2}\right) = 1$$

② Information gain is the decrease in entropy after splitting:

$$IG(X) = H(Y) - H(Y|X)$$

$$H(Y|X) = -\frac{3}{10}(1 \log_2 1 + 0 \log_2 0) - \frac{4}{10}\left(\frac{1}{2} \log_2 \frac{1}{2} + \frac{1}{2} \log_2 \frac{1}{2}\right) - \frac{3}{10}(0 \log_2 0 + 1 \log_2 1)$$

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

Visual dog

	some	none	many
Y = Meh:	3	2	0
Y = yummy:	0	2	3

$P(X = \text{some}) = 3/10$

$P(X = \text{none}) = 4/10$

$P(X = \text{many}) = 3/10$

→ $IG(x) = 1 - \frac{4}{10} = \frac{3}{5}$

③ $H(\text{Taste} \mid VD = \text{sonne})$

$= -\frac{3}{10} \cdot (1 \log_2 1 + 0 \log_2 0)$

$= 0$

$H(\text{Taste} \mid VD = \text{none})$

$= -\frac{4}{10} \cdot (\frac{1}{2} \log \frac{1}{2} + \frac{1}{2} \log \frac{1}{2})$

$= \frac{4}{10}$