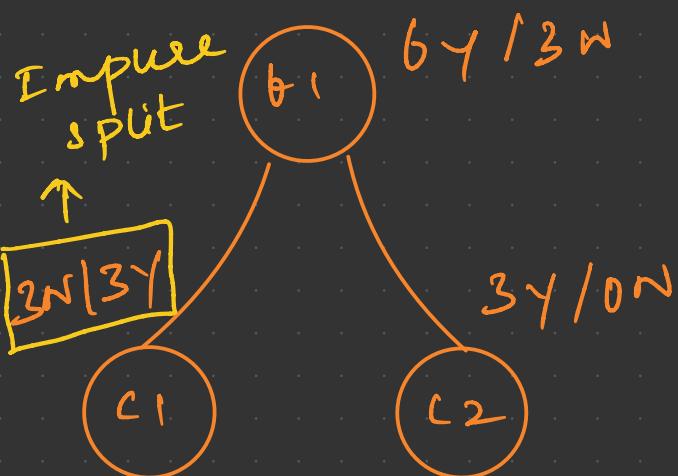


Entropy

$$H(S) = -P_+ \log_2 P_+ - P_- \log_2 P_-$$

P_+ : Probability of one category

P_- : Probability of other category.



$$H(C_1) = -P_+ \log_2 P_+$$

$$- P_- \log_2 P_-$$

$$= -\frac{2}{6} + \log_2 \frac{2}{6} -$$

$$\frac{4}{6} - \log_2 \frac{4}{6}$$

$= 1 \Rightarrow$ Impure split

Gini Impurity

$$G_I = 1 - \sum_{i=1}^n (P_i)^2$$

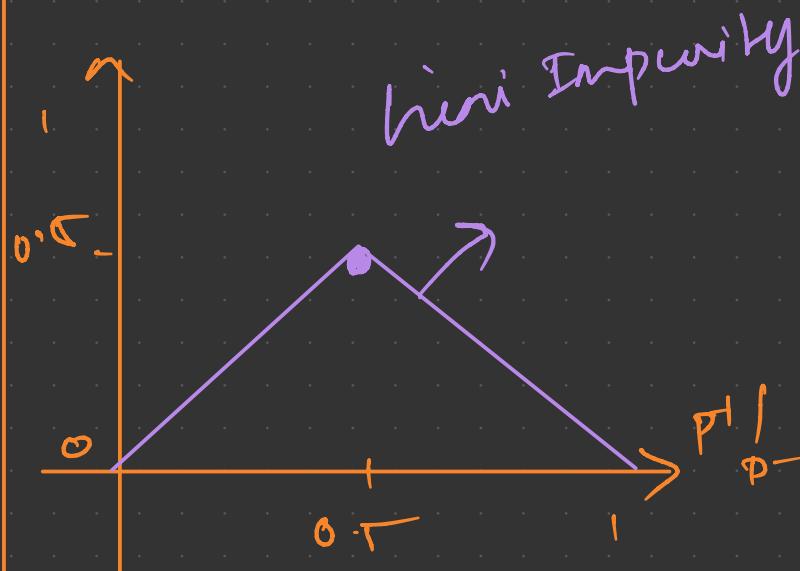
$$= 1 - ((P_+)^2 + (P_-)^2)$$

P_+ : Probability of one category
 P_- : Probability of other category.

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

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

$$= 0.5 \Rightarrow \text{Impure}$$



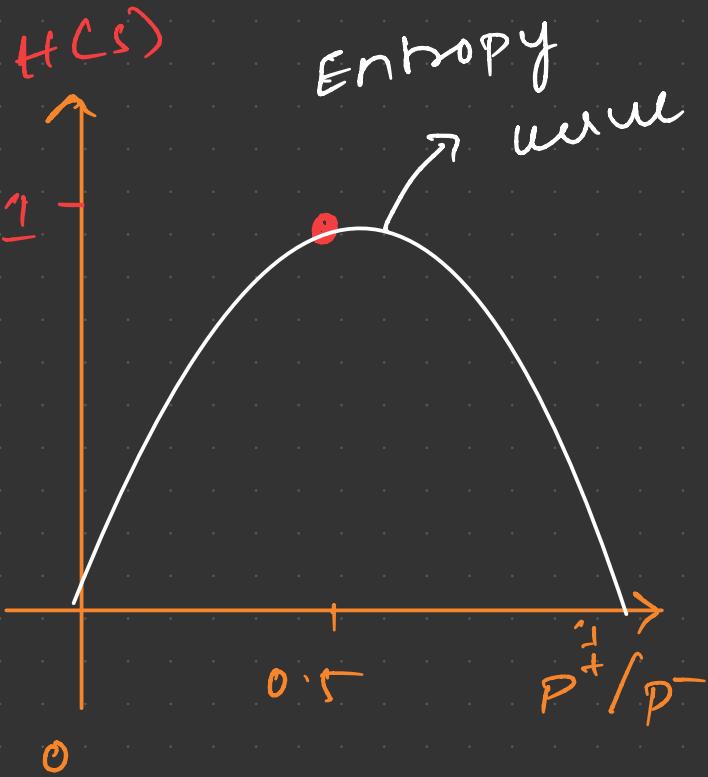
Entropy

$$H(C_2) = -3/3 \log_2$$

$$3/3 - 0 \log_2 0$$

$$= -1 \log_2 1$$

$= 0 \Rightarrow$ pure split



the values will
be always between
0's and 1

Bini Impurity

In case of
bini impurity the
values will be
ranging from 0 to
0.5.