Ex

Day	OUTLOOK	Temperature	Humidity	Wind	Play ball
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	overcast	Cool	Normal	Strong	Yes
8	Sunny	mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	mild	Normal	Weak	Yes
11	Sunny	mild	Normal	Strong	Yes
12	overcast	mild	High	Strong	Yes
13	overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

ID3.

→ ID3 is one of the most common decision tree algo.

→ Dichotomiser means dividing into two completely opposite things

→ Algo iteratively divides attributes into two groups. which are the most dominant attributes & others to construct a tree.

→ Then it calculates the Entropy & Info. Gain of each attribute. In this way the most dominant attribute can be founded.

→ After then, the most dominant one is put on the tree as decision node.

→ Entropy & Gain scores would be calculated again among the other attributes.

→ Procedure continue until reaching a decision for that branch.

— formulas: Calculate the entropy of every attribute using data set (S)

$$\text{Entropy}(S) = \sum - P(I) \cdot \log P_2(I)$$

Split the set S into subset using the attribute for which the resulting Entropy (after splitting is minimum (0). & equivalently info gain (max)

$$\text{Gain}(S, A) = \text{Entropy}(S) - \sum [P(S/A) \cdot \text{Entropy}(S/A)$$

→ make a decision tree node containing that attribute.

→ Recurse on subset using remaining attribute

outlook, Temp. Humidity, Wind are attribute
Play ball is used in decision making

① P = Yes N = NO
P = 09 N = 05

$$I(P,n) = \frac{-P}{P+n}\log_2\left(\frac{P}{P+n}\right) - \frac{n}{P+n}\log_2\left(\frac{n}{P+n}\right)$$

$$= \frac{-9}{14}\log_2\left(\frac{9}{14}\right) - \frac{05}{14}\log_2\left(\frac{05}{14}\right)$$

$$= \frac{-9}{14}\log_2(0.642) - \frac{5}{14}\log_2(0.357)$$

$$\#\frac{\log(0.642)}{\log(2)} = \frac{-9}{14}(-0.639) - \frac{5}{14}(-1.485)$$

$$= 0.941 \simeq 0.940$$

$$I(P,n) = I(9,5) = 0.940$$

Information gain

② Entropy is calcuted for perticular individual Attribute

Entropy of outlook

outlook	Pi	ni	$-I(Pi,ni)$
sunny	2	3	0.970
overcast	4	0	0
Rain	3	2	0.970

PAGE NO.:
DATE : / /

$$I(2,3) = -\frac{2}{5}\log_2\left(\frac{2}{5}\right) - \frac{3}{5}\log_2\left(\frac{3}{5}\right)$$

$$= 0.528 + 0.441$$

$$I(2,3) = 0.970.$$

$$I(4,0) = 0$$

$$I(3,2) = 0.970$$

$$E(\text{Outlook}) = \sum_{i=1}^{V} \frac{p_i + n_i}{p+N} \left(I(p_i, n_i)\right)$$

$$= \left[\frac{2+3}{9+5} + \frac{4+0}{14} + \frac{3+2}{14}\right] \left(I(p_i, n_i)\right)$$

$$= \frac{5}{9+5}(0.970) + \frac{4}{14}(0) + \frac{5}{14}(0.970)$$

$$= 0.692$$

Now - Entropy Temperature =

- Entropy Humidity
- Entropy Wind.