Data Set

Outlook	Temp	humidity	wind	Decision
Sunny	hot	high	weak	No
Sunny	hot	high	strong	No
Overcast	h	h	co	yes
rainfall	mild	h	w	Yes
rain	Cool	h	w	N
rain	cool	normal	S	N
Overcast	Cool	normal	S	Y
Sunny	m	h	w	Y
S	C	n	w	Y
R	m	n	w	n
S	m	n	S	n
O	m	h	S	Y
O	h	n	w	n
R	m	h	S	Y

Root feature → Outlook (7Y|7N)

- Sunny [2Y|3N] → Temp (7Y|7N)
  - Hot [1Y|3N] → Humidity
    - high [5Y|3N]
    - Norm [2Y|4N] → Wind
      - weak [1Y|4N]
      - Strong [3Y|3N]
  - Mild [4Y|2N]
  - Cool [2Y|2N]
- Overcast [3Y|1N]
- rainfall [2Y|3N]

∴ Conclusion :- Temp's IG is higher then rest of the feature.

① entropy [outlook]

Root feature entropy → 0.30

Decision

$$H(S) = -\sum_{i=1}^{n} P_i \log(P_i)$$

$$\Rightarrow -P_Y \log(P_Y) - P_N \log(P_N)$$

$$\Rightarrow -\frac{2}{5}\log(2/5) - \frac{3}{5}\log(3/5)$$ — Sunny

$$\Rightarrow 0.29$$

$$\Rightarrow -\frac{3}{4}\log\left(\frac{3}{4}\right) - \frac{1}{4}\log\left(\frac{1}{4}\right)$$ — Overcast

$$\Rightarrow 0.24$$

$$\Rightarrow -\frac{2}{5}\log 0.29$$ → rainfall

Information Gain [outlook]

$$\hookrightarrow H(S) - \frac{\Sigma |S_v|}{|S_v|} H(S_v)$$

$$\Rightarrow 0.30 - \left[\frac{5}{14} \times 0.29 + \frac{4}{14} \times 0.24 + \frac{5}{14} \times 0.2\right.$$

$$\Rightarrow 0.024$$

② entropy (Temp)

- 0.24 → Hot
- 0.27 → Mild
- 0.30 → Cool
- Info Gain → 0.03

③ entropy (Humidity)

- 0.28 → high
- 0.27 → Normal
- Info Gain → 0.024

④ entropy (wind)

- 1 → weak
- 1 → strong
- Info Gain → 0.024