1 Calculate the overall entropy

Tennis = 17 — Yes (Play Tennis) : 10 — $\frac{10}{17} = 0.588$

No (Don't play) = 7 — $\frac{7}{17} = 0.412$

Entropy Formula

$$= \frac{10}{17} \log_2 \frac{10}{17} - \frac{7}{17} \log_2 \frac{7}{17}$$

$$= 0.977$$

2 Information Gain For Each Attribute

A. Outlook

Sunny, Overcast, Rain

Sunny (5 instances) $= -\frac{2}{7} \log_2 \frac{2}{7} - \frac{5}{7} \log_2 \frac{5}{7}$

$= 0.863$

Overcast (4 instances) = 0 (pure node)

Rain (5 instances) $= -\frac{3}{5} \log_2 \frac{3}{5} - \frac{2}{5} \log_2 \frac{2}{5}$

$= 0.971$

Weighted average entropy:

$E(Outlook) = \frac{7}{17} * 0.863 + \frac{4}{17} * 0 + \frac{5}{17} * 0.971$ $= 0.641$

Information Gain:

Gain (Outlook) = E(S) - E(Outlook) = 0.977 - 0.641 $= 0.336$

B. Temperature Hot, Mild, Cool

Hot (4 instances): Yes : 2
No : 2

$= 1$

Mild (8 instances): Yes : 5
No : 3

$-\frac{5}{8}\log_2\frac{5}{8} - \frac{3}{8}\log_2\frac{3}{8} = 0.954$

$= 0.954$

Cool (5 Instances): Yes : 3
No : 2

$-\frac{3}{5}\cdot\log_2\frac{3}{5} - \frac{2}{5}\log_2\frac{2}{5} = 0.971$

$= 0.971$

Weighted average entropy:

$E(\text{Temp}) = \frac{4}{17} * 1 + \frac{8}{17} * 0.954 + \frac{5}{17} * 0.971 = 0.235 + 0.449 + 0.286$

$= \underline{0.970}$

Information Gain

Gain(Temp) $= 0.977 - 0.970 = 0.007$

c. Humidity High, Normal

High (9 instances): Yes : 4
No : 5

$\frac{-4}{9}\log_2\frac{4}{9} - \frac{5}{9}\log_2\frac{5}{9} = 0.991$

$= 0.991$

Normal (8 instances): Yes: 6
No: 2

$$-\frac{6}{8}\log_2\frac{6}{8} - \frac{2}{8}\log_2\frac{2}{8} = 0.811$$

$= 0.811$

Weighted average entropy:
$E(\text{Humidity}) \quad \frac{9}{17} * 0.991 + \frac{8}{17} * 0.811$

$0.525 + 0.382 \qquad = 0.907$

Information Gain:
Gain (Humidity) = 0.977 - 0.907 = 0.070

$= 0.070$

D. Wind Weak, Strong

Weak (10 instances): Yes: 7
No: 3

$$-\frac{7}{10}\log_2\frac{7}{10} - \frac{3}{10}\log_2\frac{3}{10} = 0.881$$

$= 0.881$

Strong (7 instances) Yes: 3
No: 4

$$-\frac{3}{7}\log_2\frac{3}{7} - \frac{4}{7}\log_2\frac{4}{7} = 0.985$$

$= 0.985$

Weighted average entropy:
$E(\text{Wind}) = \frac{10}{17} * 0.881 + \frac{7}{17} * 0.985 \qquad = 0.924$

3. Information Gain Summary

Outlook : 0.336 , Temperature : 0.007

Humidity : 0.070 , Wind 0.003

4. Decision Tree