

Question 2

1. Outlook.

a) Entropy

$$\text{Sunny} = -\frac{3}{7} \log_2 \frac{3}{7} - \frac{4}{7} \log_2 \frac{4}{7} = 0.9852$$

$$\text{Overcast} = 0$$

$$\text{Rain} = -\frac{2}{5} \log_2 \frac{2}{5} - \frac{3}{5} \log_2 \frac{3}{5} = 0.9710$$

$$\text{Weighted} = \frac{1}{7} \times 0.9852 + \frac{5}{7} \times 0.9710 = 0.6913$$

b) Information Gain

$$0.9774 - 0.6913 = 0.2861$$

2. Temperature.

a) Entropy

$$\text{Hot} = -\frac{2}{4} \log_2 \frac{2}{4} - \frac{2}{4} \log_2 \frac{2}{4} = 1.0$$

$$\text{Mild} = -\frac{6}{9} \log_2 \frac{6}{9} - \frac{3}{9} \log_2 \frac{3}{9} = 0.9183$$

$$\text{Cool} = -\frac{2}{4} \log_2 \frac{2}{4} - \frac{2}{4} \log_2 \frac{2}{4} = 1.0$$

$$\text{Weighted} = \frac{4}{17} \times 1 + \frac{9}{17} \times 0.9183 + \frac{4}{17} \times 1 = 0.9567$$

b) Information Gain

$$0.9774 - 0.9567 = 0.0207$$

4. Humidity.

a) Entropy

$$\text{High} = -\frac{5}{10} \log_2 \frac{5}{10} - \frac{5}{10} \log_2 \frac{5}{10} = 1.0$$

$$\text{Normal} = -\frac{5}{7} \log_2 \frac{5}{7} - \frac{2}{7} \log_2 \frac{2}{7} = 0.8631$$

$$\text{Weighted} = \frac{10}{17} \times 1 + \frac{7}{17} \times 0.8631 = 0.9437$$

b) Information Gain

$$0.9774 - 0.9437 = 0.0337$$

2. Tennis

a) Entropy

$$= -\frac{10}{17} \log_2 \frac{10}{17} - \frac{7}{17} \log_2 \frac{7}{17} = 0.9774$$

~~XXXXXX~~

5 Wind

a) Entropy

$$W_{\text{weak}} = -\frac{7}{10} \log_2 \frac{7}{10} - \frac{3}{10} \log_2 \frac{3}{10} = 0.8813$$

$$S_{\text{strong}} = -\frac{3}{7} \log_2 \frac{3}{7} - \frac{4}{7} \log_2 \frac{4}{7} = 0.9852$$

$$W_{\text{weighted}} = \frac{10}{17} \times 0.8813 + \frac{7}{17} \times 0.9852 \\ = 0.9241$$

b) Information Gain

$$0.9774 - 0.9241 \\ = 0.0533$$

Decision Tree

