COMP9414: Artificial Intelligence
Solutions 6: Machine Learning

1. (i) $Values(Outlook) = \{sunny, overcast, rain\}$
$S = [9+, 5-]$
$S_{sunny} \leftarrow [2+, 3-]$
$S_{overcast} \leftarrow [4+, 0-]$
$S_{rain} \leftarrow [3+, 2-]$
$Gain(S, Outlook) = Entropy(S) - \sum_{v=\{sunny,overcast,rain\}} \frac{|S_v|}{|S|} Entropy(S_v)$
$= Entropy(S) - \frac{5}{14} Entropy(S_{sunny}) - \frac{4}{14} Entropy(S_{overcast}) - \frac{5}{14} Entropy(S_{rain})$
$= 0.940 - \frac{5}{14} \times 0.971 - \frac{4}{14} \times 0 - \frac{5}{14} \times 0.971$
$= 0.247$
$Entropy(S) = Entropy([9+, 5-]) = -\frac{9}{14} log_2 \frac{9}{14} - \frac{5}{14} log_2 \frac{5}{14}$
$= 0.940$
$Entropy(S_{sunny}) = Entropy([2+, 3-]) = -\frac{2}{5} log_2 \frac{2}{5} - \frac{3}{5} log_2 \frac{3}{5} = 0.971$
$Entropy(S_{overcast}) = Entropy([4+, 0-]) = -\frac{4}{4} log_2 \frac{4}{4} - \frac{0}{4} log_2 \frac{0}{4} = 0$
$Entropy(S_{rain}) = Entropy([3+, 2-]) = -\frac{3}{5} log_2 \frac{3}{5} - \frac{2}{5} log_2 \frac{2}{5} = 0.971$

   (ii) $Values(Temperature) = \{hot, mild, cool\}$
$S = [9+, 5-]$
$S_{hot} \leftarrow [2+, 2-]$
$S_{mild} \leftarrow [4+, 2-]$
$S_{cool} \leftarrow [3+, 1-]$
$Gain(S, Temperature) = Entropy(S) - \sum_{v=\{hot,mild,cool\}} \frac{|S_v|}{|S|} Entropy(S_v)$
$= Entropy(S) - \frac{4}{14} Entropy(S_{hot}) - \frac{6}{14} Entropy(S_{mild}) - \frac{4}{14} Entropy(S_{cool})$
$= 0.940 - \frac{4}{14} \times 1.00 - \frac{6}{14} \times 0.918 - \frac{4}{14} \times 0.811$
$= 0.029$
$Entropy(S_{hot}) = Entropy([2+, 2-]) = -\frac{2}{4} log_2 \frac{2}{4} - \frac{2}{4} log_2 \frac{2}{4} = 1.00$
$Entropy(S_{mild}) = Entropy([4+, 2-]) = -\frac{4}{6} log_2 \frac{4}{6} - \frac{2}{6} log_2 \frac{2}{6} = 0.918$
$Entropy(S_{cool}) = Entropy([3+, 1-]) = -\frac{3}{4} log_2 \frac{3}{4} - \frac{1}{4} log_2 \frac{1}{4} = 0.811$

   (iii) $Values(Humidity) = \{high, normal\}$
$S = [9+, 5-]$
$S_{high} \leftarrow [3+, 4-]$
$S_{normal} \leftarrow [6+, 1-]$
$Gain(S, Humidity) = Entropy(S) - \sum_{v=\{high,normal\}} \frac{|S_v|}{|S|} Entropy(S_v)$
$= Entropy(S) - \frac{7}{14} Entropy(S_{high}) - \frac{7}{14} Entropy(S_{normal})$
$= 0.940 - \frac{7}{14} \times 0.985 - \frac{7}{14} \times 0.592$
$= 0.152$
$Entropy(S_{high}) = Entropy([3+, 4-]) = -\frac{3}{7} log_2 \frac{3}{7} - \frac{4}{7} log_2 \frac{4}{7} = 0.985$
$Entropy(S_{mild}) = Entropy([6+, 1-]) = -\frac{6}{7} log_2 \frac{6}{7} - \frac{1}{7} log_2 \frac{1}{7} = 0.592$

   (iv) $Values(Wind) = \{weak, strong\}$
$S = [9+, 5-]$
$S_{weak} \leftarrow [6+, 2-]$
$S_{strong} \leftarrow [3+, 3-]$
$Gain(S, Wind) = Entropy(S) - \sum_{v=\{weak,strong\}} \frac{|S_v|}{|S|} Entropy(S_v)$
$= Entropy(S) - \frac{8}{14} Entropy(S_{weak}) - \frac{6}{14} Entropy(S_{strong})$
$= 0.940 - \frac{8}{14} \times 0.811 - \frac{6}{14} \times 1.00$
$= 0.048$
$Entropy(S_{weak}) = Entropy([6+, 2-]) = -\frac{6}{8} log_2 \frac{6}{8} - \frac{6}{8} log_2 \frac{6}{8} = 0.811$
$Entropy(S_{strong}) = Entropy([3+, 3-]) = -\frac{3}{6} log_2 \frac{3}{6} - \frac{3}{6} log_2 \frac{3}{6} = 1.00$

2. (i) $Values(Temperature) = \{hot, mild, cool\}$
$S_{sunny} = [2+, 3-]$
$S_{sunny,hot} \leftarrow [0+, 2-]$
$S_{sunny,mild} \leftarrow [1+, 1-]$
$S_{sunny,cool} \leftarrow [1+, 0-]$
$Gain(S_{sunny}, Temperature) = Entropy(S_{sunny}) - \sum_{v=\{hot,mild,cool\}} \frac{|S_{sunny,v}|}{|S_{sunny}|} Entropy(S_{sunny,v})$
$= Entropy(S) - \frac{2}{5} Entropy(S_{sunny,hot}) - \frac{2}{5} Entropy(S_{sunny,mild}) - \frac{1}{5} Entropy(S_{sunny,cool})$
$= 0.971 - \frac{2}{5} \times 0.00 - \frac{2}{5} \times 1.00 - \frac{1}{5} \times 0.00$
$= 0.571$
$Entropy(S_{sunny}) = Entropy([2+, 3-]) = -\frac{2}{5} log_2 \frac{2}{5} - \frac{3}{5} log_2 \frac{3}{5} = 0.971$
$Entropy(S_{sunny,hot}) = Entropy([0+, 2-]) = -\frac{0}{2} log_2 \frac{0}{2} - \frac{2}{2} log_2 \frac{2}{2} = 0.00$
$Entropy(S_{sunny,mild}) = Entropy([1+, 1-]) = -\frac{1}{2} log_2 \frac{1}{2} - \frac{1}{2} log_2 \frac{1}{2} = 1.00$
$Entropy(S_{sunny,cool}) = Entropy([1+, 0-]) = -\frac{1}{1} log_2 \frac{1}{1} - \frac{0}{1} log_2 \frac{0}{1} = 0.00$

(ii) $Values(Humidity) = \{high, normal\}$
$S_{sunny} = [2+, 3-]$
$S_{sunny,high} \leftarrow [0+, 3-]$
$S_{sunny,normal} \leftarrow [2+, 0-]$
$Gain(S, Humidity) = Entropy(S_{sunny}) - \sum_{v=\{high,normal\}} \frac{|S_{sunny,v}|}{|S_{sunny}|} Entropy(S_{sunny,v})$
$= Entropy(S_{sunny}) - \frac{3}{5} Entropy(S_{sunny,high}) - \frac{2}{5} Entropy(S_{sunny,normal})$
$= 0.971 - \frac{3}{5} \times 0.00 - \frac{2}{5} \times 0.00$
$= 0.971$
$Entropy(S_{sunny,high}) = Entropy([0+, 3-]) = -\frac{0}{3} log_2 \frac{0}{3} - \frac{3}{3} log_2 \frac{3}{3} = 0.00$
$Entropy(S_{sunny,mild}) = Entropy([2+, 0-]) = -\frac{2}{2} log_2 \frac{2}{2} - \frac{0}{2} log_2 \frac{0}{2} = 0.00$

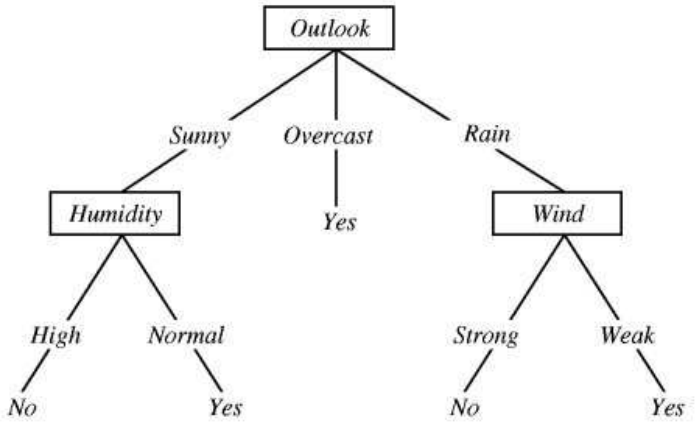
(iii) $Values(Wind) = \{weak, strong\}$
$S_{sunny} = [2+, 3-]$
$S_{weak} \leftarrow [1+, 2-]$
$S_{strong} \leftarrow [1+, 1-]$
$Gain(S, Wind) = Entropy(S_{sunny}) - \sum_{v=\{weak,strong\}} \frac{|S_{sunny,v}|}{|S_{sunny}|} Entropy(S_{sunny,v})$
$= Entropy(S) - \frac{3}{5} Entropy(S_{sunny,weak}) - \frac{2}{5} Entropy(S_{strong})$
$= 0.971 - \frac{3}{5} \times 0.918 - \frac{2}{5} \times 1.00$
$= 0.020$
$Entropy(S_{weak}) = Entropy([1+, 2-]) = -\frac{1}{3} log_2 \frac{1}{3} - \frac{1}{3} log_2 \frac{1}{3} = 0.918$
$Entropy(S_{strong}) = Entropy([1+, 1-]) = -\frac{1}{2} log_2 \frac{1}{2} - \frac{1}{2} log_2 \frac{1}{2} = 1.00$

3. (i) $Values(Temperature) = \{hot, mild, cool\}$
$S_{rain} = [2+, 3-]$
$S_{rain,hot} \leftarrow [0+, 0-]$
$S_{rain,mild} \leftarrow [2+, 1-]$
$S_{rain,cool} \leftarrow [1+, 1-]$
$Gain(S_{rain}, Temperature) = Entropy(S_{rain}) - \sum_{v=\{hot,mild,cool\}} \frac{|S_{rain,v}|}{|S_{rain}|} Entropy(S_{rain,v})$
$= Entropy(S) - \frac{0}{5} Entropy(S_{rain,hot}) - \frac{3}{5} Entropy(S_{rain,mild}) - \frac{2}{5} Entropy(S_{rain,cool})$
$= 0.971 - \frac{0}{5} \times 0.00 - \frac{3}{5} \times 0.918 - \frac{2}{5} \times 1.00$
$= 0.020$
$Entropy(S_{rain}) = Entropy([3+, 2-]) = -\frac{3}{5} log_2 \frac{3}{5} - \frac{2}{5} log_2 \frac{2}{5} = 0.971$
$Entropy(S_{rain,hot}) = Entropy([0+, 0-]) = -\frac{0}{0} log_2 \frac{0}{0} - \frac{0}{0} log_2 \frac{0}{0} = 0.00$
$Entropy(S_{rain,mild}) = Entropy([2+, 1-]) = -\frac{2}{3} log_2 \frac{2}{3} - \frac{1}{3} log_2 \frac{1}{3} = 0.918$
$Entropy(S_{rain,cool}) = Entropy([1+, 1-]) = -\frac{1}{2} log_2 \frac{1}{2} - \frac{1}{2} log_2 \frac{1}{2} = 1.00$

(ii) $Values(Humidity) = \{high, normal\}$
$S_{rain} = [3+, 2-]$

$S_{rain,high} \leftarrow [1+, 1-]$
$S_{rain,normal} \leftarrow [1+, 1-]$
$Gain(S, Humidity) = Entropy(S_{rain}) - \sum_{v=\{high,normal\}} \frac{|S_{rain,v}|}{|S_{rain}|} Entropy(S_{rain,v})$
$= Entropy(S_{rain}) - \frac{2}{5} Entropy(S_{rain,high}) - \frac{3}{5} Entropy(S_{rain,normal})$
$= 0.971 - \frac{2}{5} \times 1.00 - \frac{3}{5} \times 0.551$
$= 0.020$
$Entropy(S_{rain,high}) = Entropy([1+, 1-]) = -\frac{1}{2} log_2 \frac{1}{2} - \frac{1}{2} log_2 \frac{1}{2} = 1.00$
$Entropy(S_{rain,mild}) = Entropy([2+, 1-]) = -\frac{2}{3} log_2 \frac{2}{3} - \frac{1}{3} log_2 \frac{1}{3} = 0.551$

(iii) $Values(Wind) = \{weak, strong\}$
$S_{rain} = [3+, 2-]$
$S_{weak} \leftarrow [3+, 0-]$
$S_{strong} \leftarrow [0+, 2-]$
$Gain(S, Wind) = Entropy(S_{rain}) - \sum_{v=\{weak,strong\}} \frac{|S_{rain,v}|}{|S_{rain}|} Entropy(S_{rain,v})$
$= Entropy(S) - \frac{3}{5} Entropy(S_{rain,weak}) - \frac{2}{5} Entropy(S_{strong})$
$= 0.971 - \frac{3}{5} \times 0.00 - \frac{2}{5} \times 0.00$
$= 0.971$
$Entropy(S_{weak}) = Entropy([3+, 0-]) = -\frac{3}{3} log_2 \frac{3}{3} - \frac{0}{3} log_2 \frac{0}{3} = 0.00$
$Entropy(S_{strong}) = Entropy([0+, 2-]) = -\frac{0}{2} log_2 \frac{0}{2} - \frac{2}{2} log_2 \frac{2}{2} = 0.00$



So the example is assigned the *No* class.

2. $P(Play|Outlook, Temperature, Humidity, Wind) = \frac{P(Outlook, Temperature, Humidity, Wind|Play).P(Play)}{P(Outlook, Temperature, Humidity, Wind)}$

Similarly for $P(\neg Play|Outlook, Temperature, Humidity, Wind)$

$P(Sunny, Hot, High, Weak|Play).P(Play)$
$= P(Sunny|Play).P(Hot|Play).P(High|Play).P(Weak|Play).P(Play)$ by independence
$= 2/9 * 2/9 * 3/9 * 6/9 * 9/14$
$= 0.00705$

$P(Sunny, Hot, High, Weak|\neg Play).P(\neg Play)$
$= P(Sunny|\neg Play).P(Hot|\neg Play).P(High|\neg Play).P(Weak|\neg Play).P(\neg Play)$ by independence
$= 3/5 * 2/5 * 4/5 * 2/5 * 5/14$
$= 0.02743$

Again the example is assigned the *No* class.