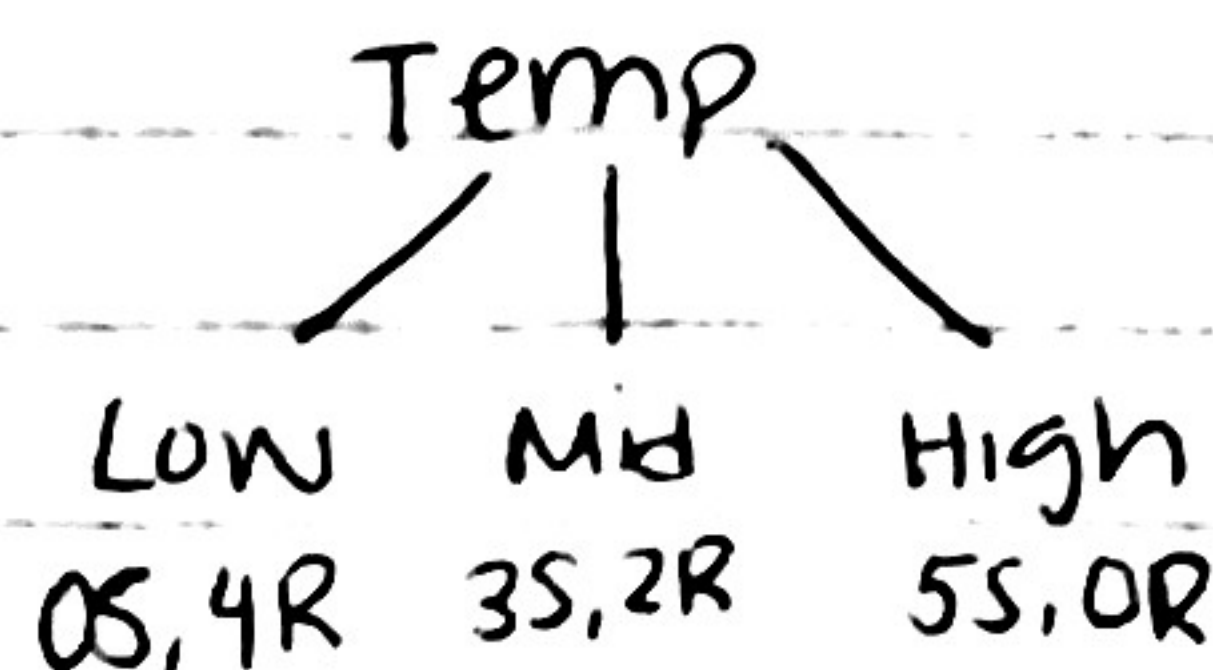


# 1 Split on Temp

High - 5 Sunny, 0 Rainy 5 samples  
 Mid - 3 Sunny, 2 Rainy 5 samples  
 Low - 0 Sunny, 4 Rainy 4 samples



$$\text{High: } H(x) = - \left[ \frac{5}{5} \log_2 \left( \frac{5}{5} \right) + 0 \log_2 (0) \right] = 0 \text{ bit}$$

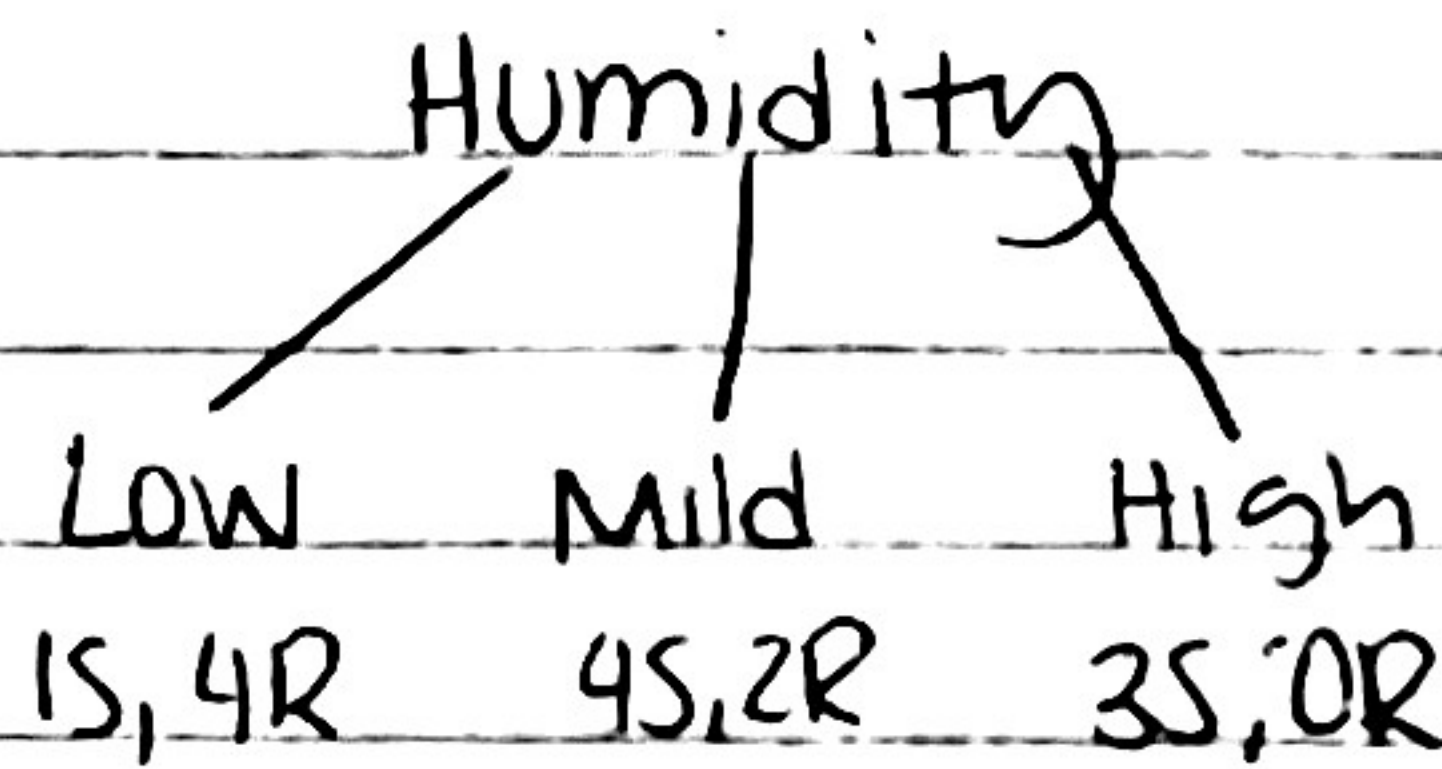
$$\text{Mid: } H(x) = - \left[ \frac{3}{5} \log_2 \left( \frac{3}{5} \right) + \frac{2}{5} \log_2 \left( \frac{2}{5} \right) \right] = 0.97 \text{ bit}$$

$$\text{Low: } H(x) = - \left[ 0 \log_2 (0) + \frac{4}{4} \log_2 \left( \frac{4}{4} \right) \right] = 0 \text{ bit}$$

$$E(H(x)) = \frac{5}{14} \times 0 + 0.97 \times \frac{5}{14} + \frac{4}{14} \times 0 = \boxed{0.34}$$

# Split on Humidity

High - 1 Sunny, 4 Rainy 5 samples  
 Mild - 4 Sunny, 2 Rainy 6 samples  
 Low - 3 Sunny, 0 Rainy 3 samples



$$\text{High: } H(x) = - \left[ \frac{1}{5} \log_2 \left( \frac{1}{5} \right) + \frac{4}{5} \log_2 \left( \frac{4}{5} \right) \right] = 0.72 \text{ bit}$$

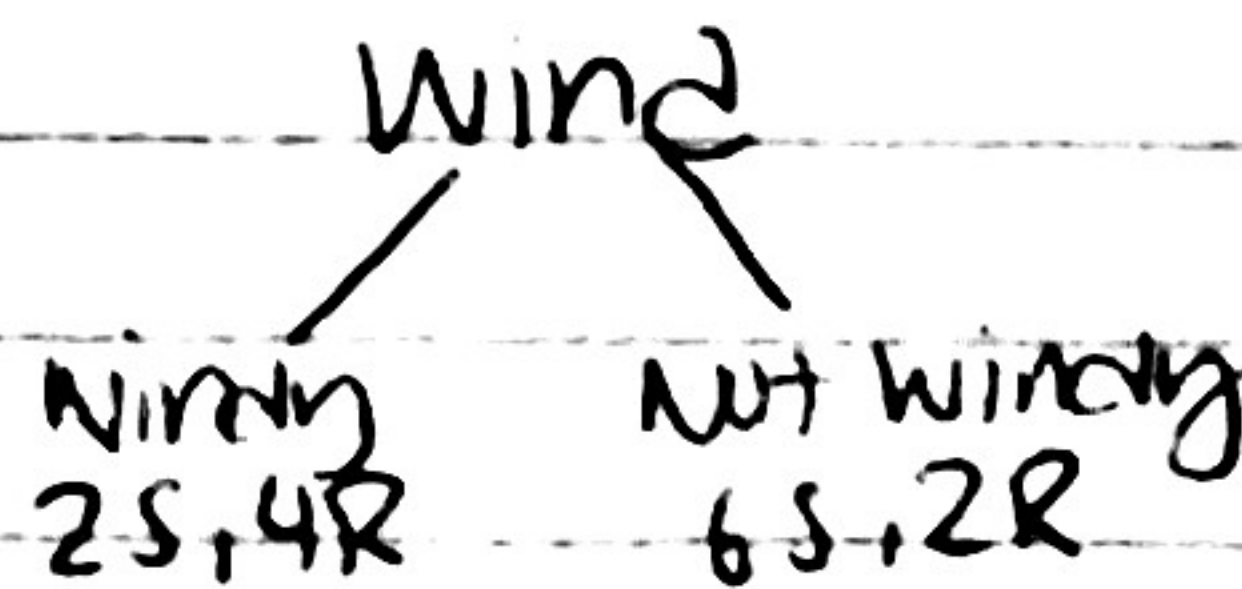
$$\text{Mild: } H(x) = - \left[ \frac{4}{6} \log_2 \left( \frac{4}{6} \right) + \frac{2}{6} \log_2 \left( \frac{2}{6} \right) \right] = 0.91 \text{ bit}$$

$$\text{Low: } H(x) = - \left[ \frac{3}{3} \log_2 \left( \frac{3}{3} \right) + \frac{0}{3} \log_2 \left( \frac{0}{3} \right) \right] = 0 \text{ bit}$$

$$E(H(x)) = 0.72 \times \left( \frac{5}{14} \right) + 0.91 \left( \frac{6}{14} \right) + 0 \left( \frac{3}{14} \right) = \boxed{0.64}$$

# Split on Wind:

Windy - 2 Sunny, 4 Rainy 6 samples  
 Not Windy - 6 Sunny, 2 Rainy 8 samples



$$\text{Windy: } H(x) = - \left[ \frac{2}{6} \log_2 \left( \frac{2}{6} \right) + \frac{4}{6} \log_2 \left( \frac{4}{6} \right) \right] = 0.92 \text{ bit}$$

$$\text{Not Windy: } H(x) = - \left[ \frac{6}{8} \log_2 \left( \frac{6}{8} \right) + \frac{2}{8} \log_2 \left( \frac{2}{8} \right) \right] = 0.81 \text{ bit}$$

$$E(H(x)) = 0.92 \times \left( \frac{6}{14} \right) + 0.81 \times \left( \frac{8}{14} \right) = \boxed{0.85}$$



Total Sunny: 8/14  
Total Rainy: 6/14

Entropy before split:  
 $= - \left[ \frac{8}{14} \cdot \log_2\left(\frac{8}{14}\right) + \frac{6}{14} \log_2\left(\frac{6}{14}\right) \right] = 0.98 \text{ bit}$

Information Gain:

- ① Temp:  $0.98 - 0.34 = 0.64$
- ② Humidity:  $0.98 - 0.64 = 0.34$
- ③ Wind:  $0.98 - 0.85 = 0.13$

Best Feature to split on is Temperature since it minimizes entropy and provides more information gain.