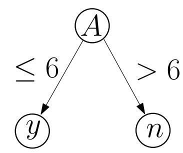
# CMSC5724: Quiz 1

**Name:** **Student ID:**

**Solution to Problem 1.** One possible decision tree is shown below.




**Solution to Problem 2.** Les $S$ be the training set given in Problem 1 and $\mathcal{H}$ be the set of classifiers that can possibly be returned. Denote by $h$ the decision tree we found in Problem 1. As $X$ has two choices ($A$ and $B$) and $v$ has 16 choices, we know $|\mathcal{H}| = 32$. Our decision tree in Problem 1 has empirical error $err_S(h) = 0.125$.

According to the generalization theorem, with probability at least $1 - \delta$, we have

$$\begin{aligned} err_{\mathcal{D}}(h) &\leq err_S(h) + \sqrt{\frac{\ln(1/\delta) + \ln|\mathcal{H}|}{2|S|}} \\ &= 0.125 + \sqrt{\frac{\ln(1/\delta) + \ln 32}{16}}. \end{aligned}$$

By setting $\delta = 0.1$, we know with probability at least 0.9,

$$err_{\mathcal{D}}(h) \leq 0.125 + \sqrt{\frac{\ln(1/0.1) + \ln 32}{16}} \leq 0.73.$$

**Solution to Problem 3.** By Bayes Theorem

$$\mathbf{Pr}[Y = y \mid A = 1, B = 1, C = 0] = \frac{\mathbf{Pr}[A = 1, B = 1, C = 0 \mid Y = y] \cdot \mathbf{Pr}[Y = y]}{\mathbf{Pr}[A = 1, B = 1, C = 0]}$$

and

$$\mathbf{Pr}[Y = n \mid A = 1, B = 1, C = 0] = \frac{\mathbf{Pr}[A = 1, B = 1, C = 0 \mid Y = n] \cdot \mathbf{Pr}[Y = n]}{\mathbf{Pr}[A = 1, B = 1, C = 0]}$$

To know which fraction is bigger, it is sufficient to estimate their numerators:

$$\begin{aligned} & \mathbf{Pr}[A = 1, B = 1, C = 0 \mid Y = y] \cdot \mathbf{Pr}[Y = y] \\ =\; & \mathbf{Pr}[A = 1 \mid Y = y] \cdot \mathbf{Pr}[B = 1 \mid Y = y] \cdot \mathbf{Pr}[C = 0 \mid Y = y] \cdot \mathbf{Pr}[Y = y] \\ \text{(estimate)} \quad =\; & \frac{3}{4} \times \frac{3}{4} \times \frac{1}{4} \times \frac{1}{2} \\ =\; & \frac{9}{128}. \end{aligned}$$

$$\begin{aligned} & \mathbf{Pr}[A = 1, B = 1, C = 0 \mid Y = n] \cdot \mathbf{Pr}[Y = n] \\ =\; & \mathbf{Pr}[A = 1 \mid Y = n] \cdot \mathbf{Pr}[B = 1 \mid Y = n] \cdot \mathbf{Pr}[C = 0 \mid Y = n] \cdot \mathbf{Pr}[Y = n] \\ \text{(estimate)} \quad =\; & \frac{1}{4} \times \frac{3}{4} \times \frac{2}{4} \times \frac{1}{2} \\ =\; & \frac{6}{128}. \end{aligned}$$

We thus conclude that $\mathbf{Pr}[Y = y \mid A = 1, B = 1, C = 0] > \mathbf{Pr}[Y = n \mid A = 1, B = 1, C = 0]$. The predicted label is $y$.