

## Question 1

### Question 2

a) The decision tree in Figure 4 correctly categorizes the provided examples.

$$b) \text{Gain}(A) = I\left(\frac{p}{p+n}, \frac{n}{p+n}\right) - \sum_{i=1}^n \frac{p_i+n_i}{p+n} I\left(\frac{p_i}{p_i+n_i}, \frac{n_i}{p_i+n_i}\right)$$

$$\text{Gain}(GPA) = 1 - \left[ \frac{3}{12} I(1, 0) + \frac{5}{12} I\left(\frac{3}{5}, \frac{2}{5}\right) + \frac{4}{12} I(0, 1) \right] \approx 0.595$$

$$\text{Gain}(University) = 1 - \left[ \frac{5}{12} I\left(\frac{3}{5}, \frac{2}{5}\right) + \frac{3}{12} I\left(\frac{2}{3}, \frac{1}{3}\right) + \frac{4}{12} I\left(\frac{1}{4}, \frac{3}{4}\right) \right] \approx 0.095$$

$$\text{Gain}(Published) = 1 - \left[ \frac{5}{12} I\left(\frac{3}{5}, \frac{2}{5}\right) + \frac{7}{12} I\left(\frac{3}{7}, \frac{4}{7}\right) \right] \approx 0.021$$

$$\text{Gain}(Recommendation) = 1 - \left[ \frac{8}{12} I\left(\frac{5}{8}, \frac{3}{8}\right) + \frac{4}{12} I\left(\frac{1}{4}, \frac{3}{4}\right) \right] \approx 0.093$$

1st level: GPA because the information gain is highest

$$\text{Gain}(University) = I\left(\frac{3}{5}, \frac{2}{5}\right) - \left[ \frac{2}{5} I\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{1}{5} I(1, 0) + \frac{2}{5} I\left(\frac{1}{2}, \frac{1}{2}\right) \right] \approx 0.171$$

$$\text{Gain}(Published) = I\left(\frac{3}{5}, \frac{2}{5}\right) - \left[ \frac{2}{5} I(1, 0) + \frac{3}{5} I\left(\frac{1}{3}, \frac{2}{3}\right) \right] \approx 0.420$$

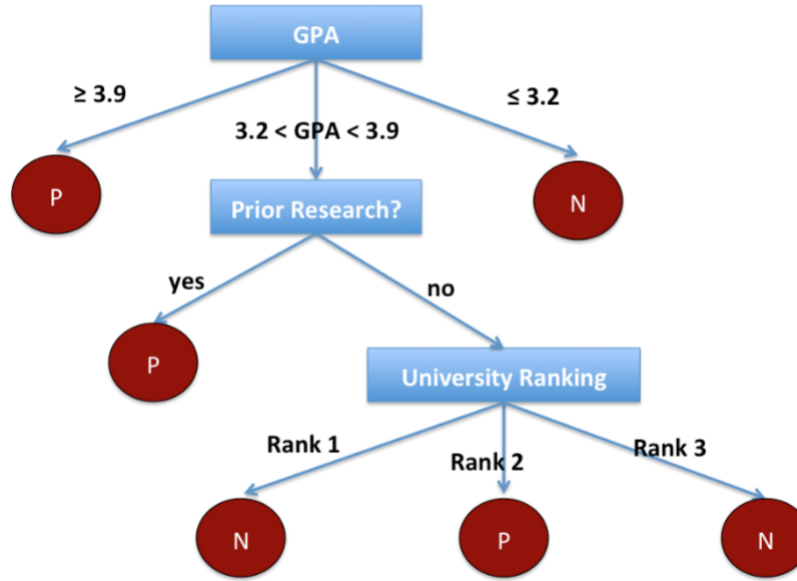
$$\text{Gain}(Recommendation) = I\left(\frac{3}{5}, \frac{2}{5}\right) - [I\left(\frac{3}{5}, \frac{2}{5}\right)] = 0$$

2nd level: Published because the information gain is highest

$$\text{Gain}(University) = I\left(\frac{1}{3}, \frac{2}{3}\right) - \left[ \frac{1}{3} I(0, 1) + \frac{1}{3} I(1, 0) + \frac{1}{3} I(0, 1) \right] \approx 0.918$$

$$\text{Gain}(Recommendation) = I\left(\frac{1}{3}, \frac{2}{3}\right) - [I\left(\frac{1}{3}, \frac{2}{3}\right)] = 0$$

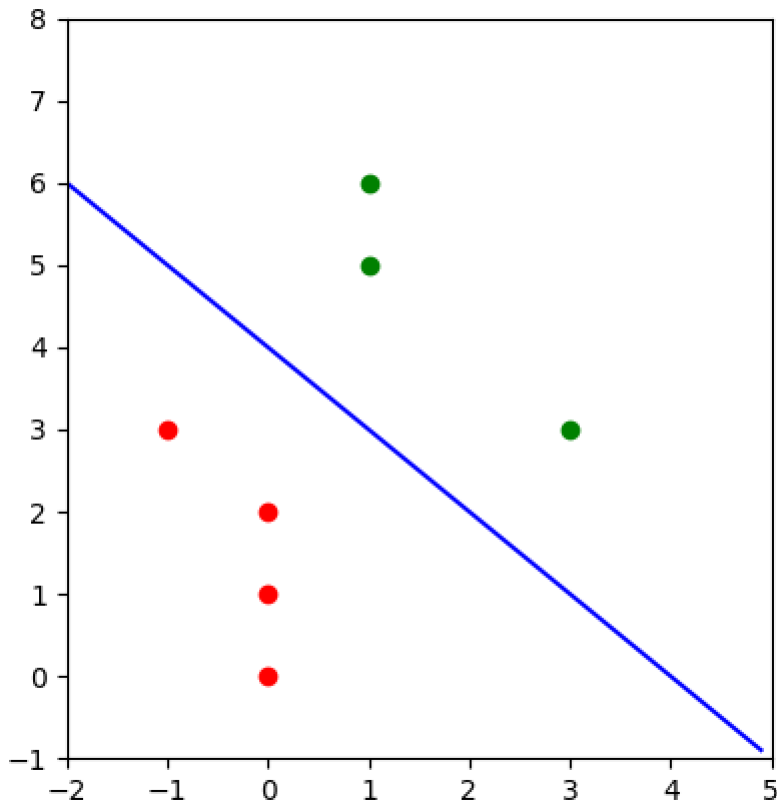
3rd level: University because the information gain is highest



The tree is :

c) The trees are equivalent. This is a coincidence because we were provided with the best decision tree. This data set, to me, did not have an obvious decision tree simply by looking at the table so information gains must be calculated.

### Question 3



a)

b) The decision boundary lies on  $y = -x + 4$ . We have a positive support vector at (0,2) with line equation  $y = -x + 2$  and a negative support vector at (1,5) with line equation  $y = -x + 6$

So  $w = (-\frac{1}{2}, -\frac{1}{2})$  and  $b = 2$

c) The linear SVM is the same as in part b.  $w = (-\frac{1}{2}, -\frac{1}{2})$  and  $b = 2$