Question 2

Part a

1) 4.0 GPA - P, the data match the decision tree
2) 3.9 GPA - no publication - rank 1 - N, the data don't match the decision tree
3) 3.9 GPA - no publication - rank 2 - P, the data match the decision tree

8) GPA 3.4 - No publication - Rank 1 University - N, data matches the tree
9) GPA 3.2 - N, data matches the tree
10) GPA 3.1 - N, data matches the tree
11) GPA 3.1 - N, data matches the tree
12) GPA 3.0 - N, data matches the tree

Part b

Published = Yes : 3 P, 2 N

$$E(Published = Yes) = -\frac{3}{5}\log_2(\frac{3}{5}) - (\frac{2}{5})\log_2(\frac{2}{5}) = .9710 \quad (1)$$

Published = No : 3 P, 4 N

$$E(Published = No) = -\frac{3}{7}\log_2(\frac{3}{7}) - (\frac{4}{7})\log_2(\frac{4}{7}) = .9852 \quad (2)$$

Published Average Entropy

$$I(Published) = \frac{5}{12}(.9710) + \frac{7}{12}(.9852) = .9792 \quad (3)$$

Recommendations = Good : 5 P, 3 N

$$E(Recommendations = Good) = -\frac{5}{8}\log_2(\frac{5}{8}) - (\frac{3}{8})\log_2(\frac{3}{8}) = .9544 \quad (4)$$

Recommendations = Normal : 1 P, 3 N

$$E(Recommendations = Normal) = -\frac{1}{4}\log_2(\frac{1}{4}) - (\frac{3}{4})\log_2(\frac{3}{4}) = .8113 \quad (5)$$

Recommendations Average Entropy

$$I(Recommendations) = \frac{8}{12}(.9544) + \frac{4}{12}(.8113) = .9067 \quad (6)$$