

1.

a.

$$\text{i. } P(C|S) = (P(S|C)P(C))/P(S) = .8(.05)/(.8(.05)+.1(.95)) = 0.296$$

$$\text{ii. } P(C|\neg S) = (P(\neg S|C)P(C))/P(\neg S) = (.85)(.05)/(.85 + (.85)(.05)) = .0476$$

$$\text{iii. } P(\neg C|S) = (P(S|\neg C)P(\neg C))/P(S) = 0.1(.95)/((.15) + 0.1(.95)) = 0.388$$

$$\text{iv. } P(\neg C|\neg S) = (P(\neg S|\neg C)P(\neg C))/P(\neg S) = .85(.95)/(.85 + .85(.95)) = 0.487$$

$$\text{b. } P(C,S) = (P(S|C)P(C))/P(S)$$

$$\text{i. } P(S) =$$

$$\text{ii. } P(C|S) =$$

$$\text{iii. } P(C|\neg S) =$$

c. 1(b) makes more sense. It makes more sense to determine the probability of a cold given that you have a sore throat.

2. You don't want to use the test data as training data because then you are training to the test and so you can't determine the accuracy of the training. You want to set aside some data for testing to be able to accurately evaluate your training process

3.

a.

$$\text{i. } -(3/8)\log_2(3/8) - (5/8)\log_2(5/8) = 0.954434002924965$$

$$\text{ii. } -(2/3)\log_2(2/3) - (1/3)\log_2(1/3) = 0.9182958340544896$$

$$\text{iii. } -(1/5)\log_2(1/5) - (4/5)\log_2(4/5) = 0.7219280948873623$$

$$\text{b. } ((3/8)\log_2(3/8) - (5/8)\log_2(5/8)) - (3/8)((2/3)\log_2(2/3)) - (1/3)\log_2(1/3)) - (5/8)((1/5)\log_2(1/5) - (4/5)\log_2(4/5)) = 0.1588680058499299625$$

4.

a. We building a decision tree, we should choose the node which gives us the most information gain