Homework 6 Primer

Due Wednesday, December 7th at 11:59pm

You are encouraged to discuss the assignment in general with your classmates, and may optionally collaborate with one other student. If you choose to do so, you must indicate with whom you worked. Multiple teams (or non-partnered students) submitting the same solutions will be considered plagiarism.

The goal of this assignment is to deepen your understanding of some of the topics covered in the lectures and readings. We will grade your answers based on whether they demonstrate an understanding of the concepts in each question. We will award partial credit for answers that demonstrate partial understanding, so show your work!

What to Submit

You should submit a file named `homework6primer.pdf`, containing your answers to the questions. You can record your answers on this document (preferred) or create your own.

Edibility (30 points)

I use the decision tree classifier on the right to determine whether I should eat things I find in the back of my refrigerator. Each item is described by three attributes: "Days" (how many days it has been there), "Tasty" (whether I really like it), "Moldy" (whether there is visible mold). The table below the tree shows the labeled data that was used to learn the model.
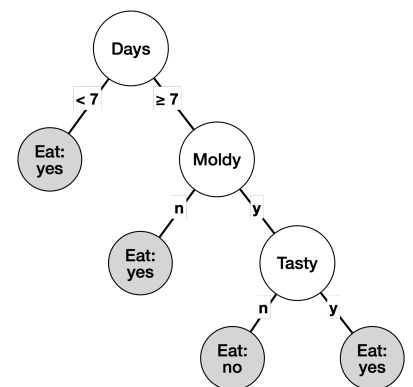
    a. For each of the four leaf nodes, indicate the conditional probability that the node represents, along with the value of that probability.

Training Data

Days	Tasty	Moldy	Eat?
4	y	y	y
2	n	n	y
10	y	n	y
9	n	n	n
5	n	y	n
26	y	y	y
8	y	n	y
11	n	y	n
21	n	n	y
13	y	y	y

Test Data

ID	Days	Tasty	Moldy
A	13	n	y
B	6	n	y
C	147	y	y

b. For each example in the small test set shown on the right, indicate the class value as determined by the decision tree model shown above.

c. If instead I was to train a Naive Bayes Classifier using the training data above (similarly discretizing Days using a threshold of <7), how would it classify the example (Days=4, Tasty=n, Moldy=n)? Show the probability calculation for both classes.