This sounds like a fun project! We'll talk a bit about decision trees in class, and this will likely be a nice extension. Rather than accuracy, you may want to look at F1 score, which is the harmonic mean of precision and recall; spam filtering can look deceptively easy depending on training/test composition, since often spam is a relatively small proportion of the corpus. Examining issues like dealing with continuous attributes and how to decide when to stop building your tree (to avoid overfitting) should make this a good project! Since you're only looking at one classification algorithm, you'll want to be sure to explore variations, such as looking at how held out data performance changes with different criteria around when to stop splitting. One issue that will arise is what precisely to use as features. Please come talk to me if you have questions or run into issues.