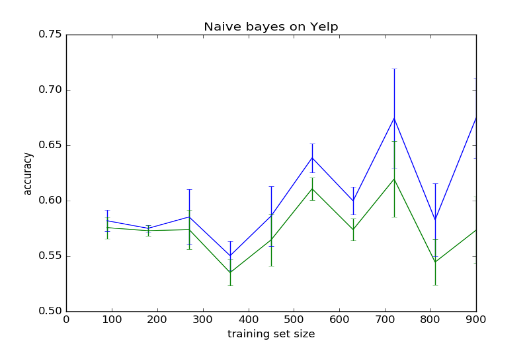
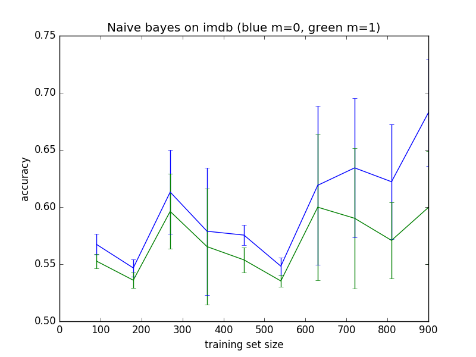
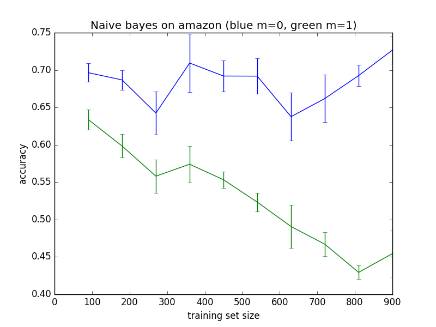
Ethan Hartzell

Below are the results of the first experiment on three different datasets. One thing is clear, add-one smoothing gives lower accuracies in all cases, probably because this smoothing method is taking away too much probability mass from words that may be good determiners of labels. The accuracy increases with training set size, which makes sense because there would be less unknown words. For some reason, this does not appear to be true for the smoothed Amazon data, maybe probably because as the vocabulary size grows, the greater the probability mass loss is of those words that should have high probability mass.



Below are the results of the second experiment on three datasets. The clear trend is that the higher the *m* is for the smoothing method, the lower the accuracy. Even values less than 1 are apparently too big. The larger the value, the less probability mass of words that occurred often.

