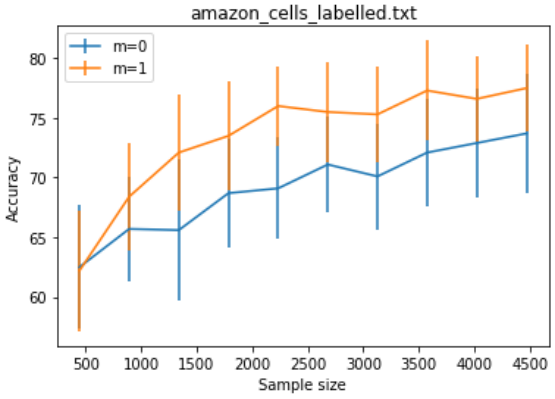
B555: Machine Learning

Programming Project 1

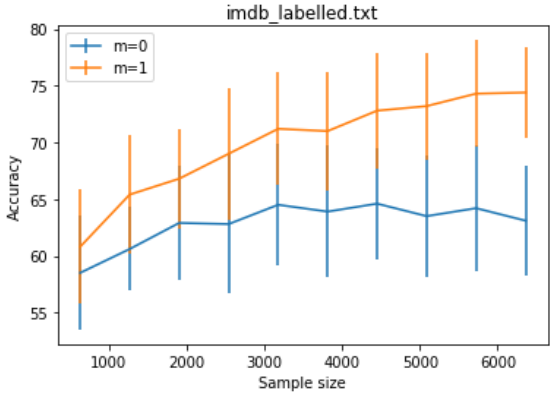
Experiment 1:

Learning curves for each of the three datasets for m=0 and m=1 can be seen below:

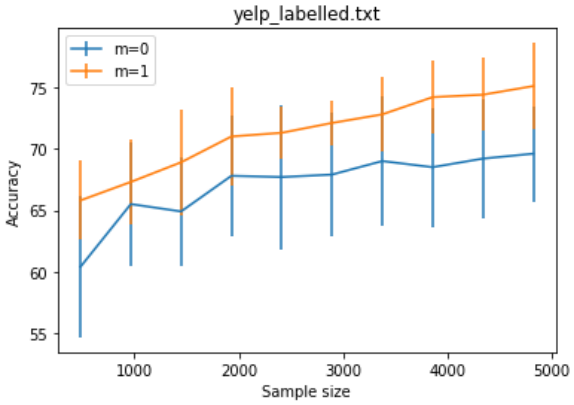
Dataset 1:



Dataset 2:



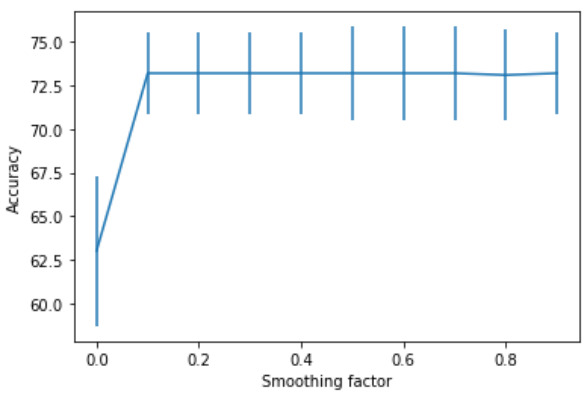
Dataset 3:

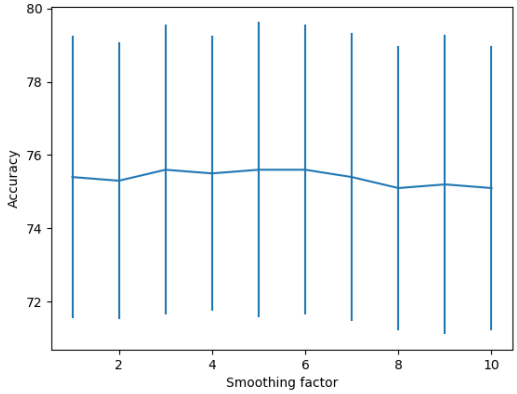


From this experiment, we can make the following observations:

- As the sample size (0.1N, 0.2N, …, N) increases, the average accuracy of the results also increases.
- Similarly, as we increase the smoothing factor m, the accuracy of the results increases. So, for m=1 we get more accurate results than using m=0.

Experiment 2:





From the output of the second experiment, we can observe that:

- There is a drastic change in accuracy when the smoothing factor becomes greater than 0. This is because for m=0, we are using essentially using maximum likelihood estimator and so there are lot of cases with both positive and negative likelihood having value 0.
- When m>0, we are essentially using MAP to predict the outcomes and so the accuracy is increased.
- As per the graph, there is not much difference in the average accuracy when the value of m is increased by a fixed factor.
- The accuracy and standard deviation for higher values of m (i.e., m=1, 2, 3, …, 9) is a little more compared to lower values of m (i.e., m=0.1, 0.2, …., 0.9)