AI MP 4

Matt Mann(mam1010) and Vince Taylor(vt152)

November 2018

Question a.2

The drawback of this formula is that it can become inefficient to compute a large number of values via multiplication. Adding is takes less time than multiplying.

Question a.5

10% = 0.733
20% = 0.772
30% = 0.792
40% = 0.789
50% = 0.800
60% = 0.801
70% = 0.805
80% = 0.815
90% = 0.813
100% = 0.817
A larger training set increases the accuracy of the program, save for a few exceptions, but the payoff for increasing the training set becomes less and less.

Question a.6

We went with k=1. We tested other values and the accuracy of the program was the most accurate when the k value was 1.
Our classifier is reasonably accurate. Its lowest tested accuracy was about 73%. We wouldn't describe this as reasonably good because when a 70%-80% would still feel a bit buggy to someone who is using the software. In other words, we feel it is not up to industry standards.

Question b.1

We used three new sets of features for out advanced features function. We used outside borders only (1), we also used internal hash tags only(2), and finally a

combination of both internal and external which made our set (0,1,2)'s. The border only is to see how effective just calculating the borders would be to maybe save on computation time for larger sets. The use of the internal only was to see if it would benefit the basic implementation, because the borders are not always accurate. Using both allowed us to have a new data set size (one that was not binary) in order to see if it helped increase our accuracy. We have a global variable setfeat that determines which feature set we use and which method of computation to use. 0 = Border+Internal ,1 = basic,2 = internal only, 3 = Border only, 4 = Final Feature Extractor
Advanced feature 1 (Border Only)
10% = 0.64
20% = 0.691
30% = 0.701
40% = 0.715
50% = 0.716
60% = 0.723
70% = 0.726
80% = 0.736
90% = 0.731
100% = 0.731
Advanced feature 2 (Internal Only)
10% = 0.718
20% = 0.755
30% = 0.754
40% = 0.766
50% = 0.774
60% = 0.771
70% = 0.772
80% = 0.783
90% = 0.78
100% = 0.784
Advanced feature 3 (Internal and Border)
10% = 0.696
20% = 0.714
30% = 0.72
40% = 0.724
50% = 0.729
60% = 0.732
70% = 0.732
80% = 0.737
90% = 0.735
100% = 0.748

Question b.2

Advanced feature 1 (Border Only + Basic)
10% = 0.629
20% = 0.629
30% = 0.659
40% = 0.671
50% = 0.67
60% = 0.69
70% = 0.69
80% = 0.71
90% = 0.718
100% = 0.72
Advanced feature 2 (Internal Only + Basic)
10% = 0.631
20% = 0.65
30% = 0.669
40% = 0.679
50% = 0.677
60% = 0.687
70% = 0.695
80% = 0.705
90% = 0.706
100% = 0.71
Advanced feature 3 (Internal and Border + Basic)
10% = 0.733
20% = 0.772
30% = 0.792
40% = 0.789
50% = 0.800
60% = 0.801
70% = 0.805
80% = 0.815
90% = 0.813
100% = 0.817