CMSC 471 Project 3 Report

Sean Murren

May 2015

1 Decision Tree

The algorithm was able to successfully produce the correct results when tested against the easy data set. The accuracy, precision, and recall were all 1.0 when training the 7 training vectors and then testing them. This strongly supports that the program was properly trained. The effectiveness of the program on unclassified data is unkown as there was no availabe data to test it on.

2 Naive Bayes

I had a lot of trouble working out the correct algorithm for this program. The end result was a program that produced good results for the 2-vector test case example provided in the assignment instructions. It was able to classify similar unclassified vectors quite effectively as well. However, the resulting program was not functioning as intended as it was producing probabilities greater than 1.0. I was not able to identify the problems in my implementation within the given time period. When the easy data set was tested on this program it classified every case as 0. The resulting accuracy was $\frac{5}{7}$, precision 0, and recall 0 (assuming true positive is 1, true negative is 0). I am not sure if the issue was regarding accuracy and precision or if the program glitched on the easy data set and arbitrarily returned all 0's.

3 K-nearest Neighbor

The results and conclusions regarding this program are the same as with the Decision Tree. This algorithm was the easiest to implement, but it was the least interesting as well.

4 Conclusion

While the results were promising for the decision tree and k-nearest programs, I was unable to properly test them because I was not sure how to do so with the given information in the assignment. Training sets were provided, but no testing sets. While the naive bayes program produced great results for the simplest of tests, it did not function properly and I was unable to fix the problem in the alloted time. Also, I found the decision tree the most interesting of the algorithms to work with.