COMS 4721 Assignment 3
Boosting

John Fulgoni
Columbia University in the City of New York

31 March 2015

List of Figures

1 Part 1

For each plot in Figure 1, it shows the CDF where the number of samples ranges from 100 to 500, with the title N being 1 to 5. The function that retrieves these values is *pick_vars.m*.

2 Part 2

By running the Bayes Classifier on the Training Set without Boosting, the accuracy on the testing set was **0.8415**.

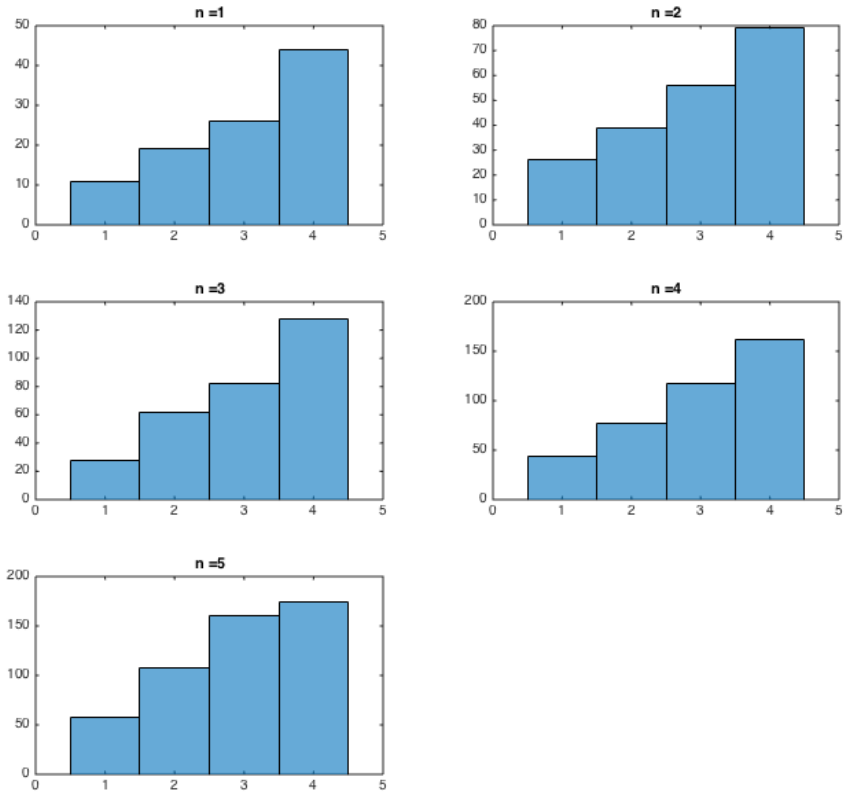

Figure 1: Sample Vector C for Varying N Values

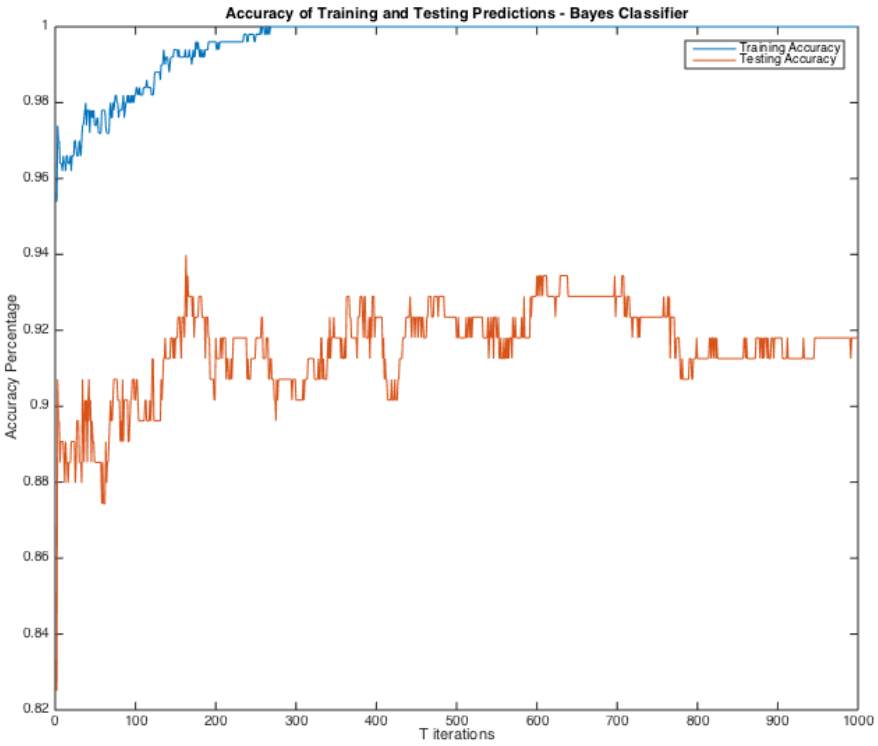

Figure 2: Training and Testing Error for Boosted Bayes Classifier

After 1000 iterations for T, the Training Accuracy of the Boosted Classifier was **1.0**, and the Testing Accuracy was **0.9180**. Figure 2 shows these values over time. Figure 3 shows the Epsilon and Alpha values over 1000 iterations.

For the three points to plot W(i) over T, I chose 168 (blue), 245 (red), and 347 (yellow). The variation in W(i) can be seen in Figure 4.

3 Part 3

By running the Logistic Classifier on the Training Set without Boosting, the accuracy on the testing set was **0.8251**.

After 1000 iterations for T, the Training Accuracy of the Boosted Clas-

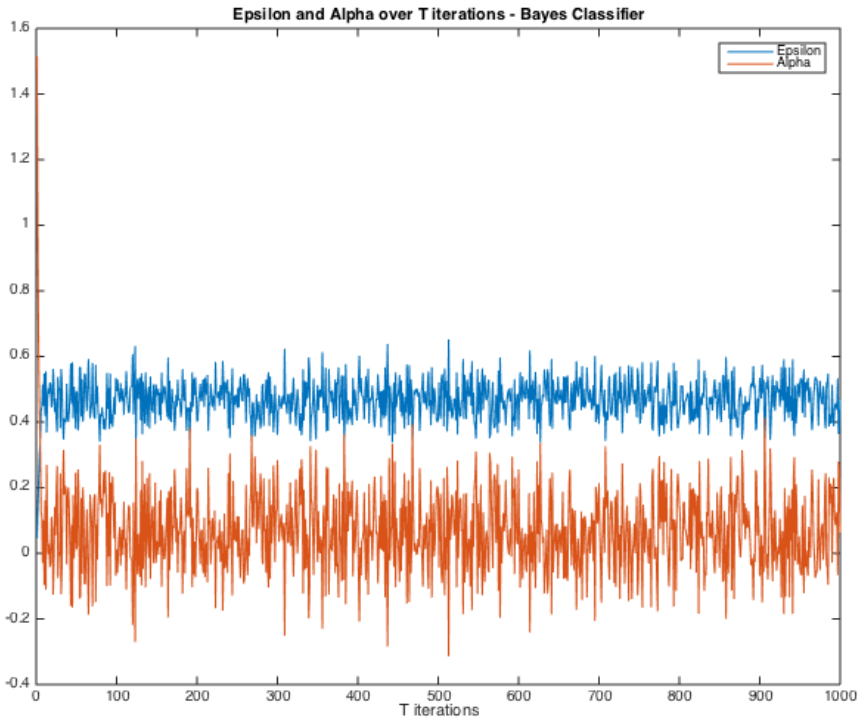

Figure 3: Epsilon and Alpha over T Iterations for Boosted Bayes Classifier

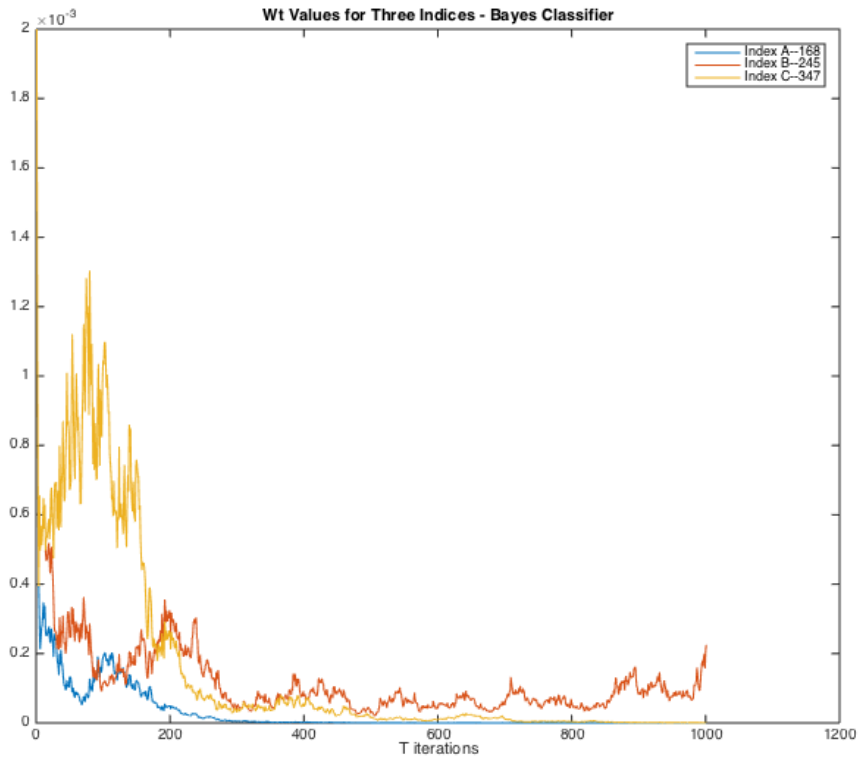

Figure 4: W(i) for points 168, 245, and 347 over T Iterations for Boosted Bayes Classifier

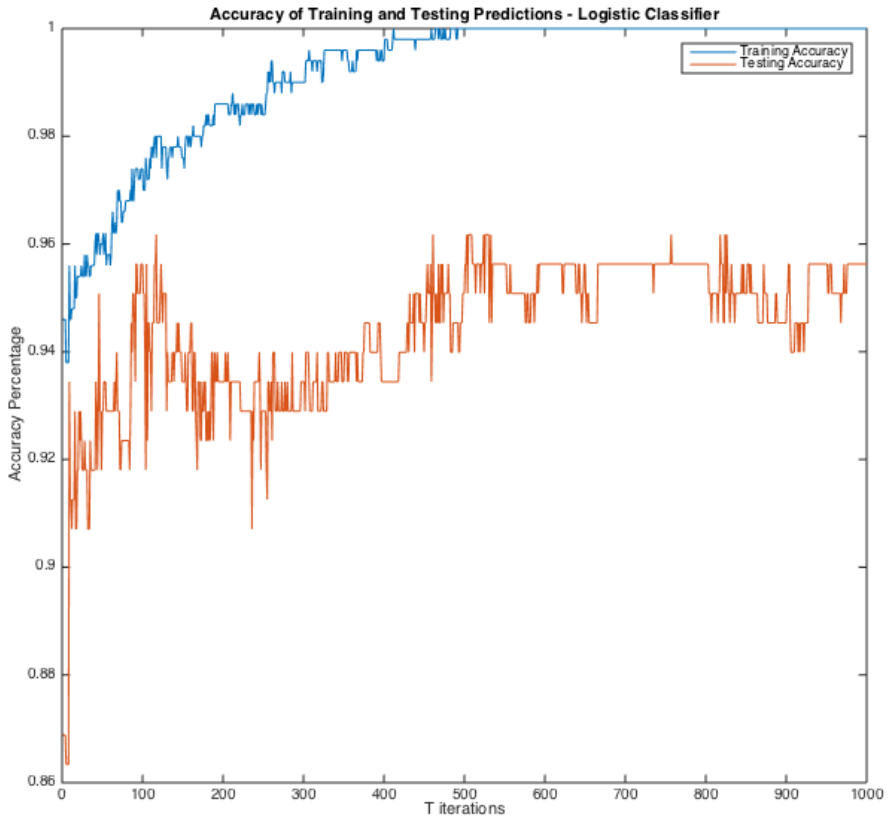

Figure 5: Training and Testing Error for Boosted Logistic Classifier

sifier was **1.0**, and the Testing Accuracy was **0.9563**. Figure 5 shows these values over time. Figure 6 shows the Epsilon and Alpha values over 1000 iterations.

For the three points to plot W(i) over T, I chose 7 (blue), 279(red), and 470 (yellow). The variation in W(i) can be seen in Figure 7. (Note that I messed up the title in Figure 7, but it is still result of the Logistic Classifier)

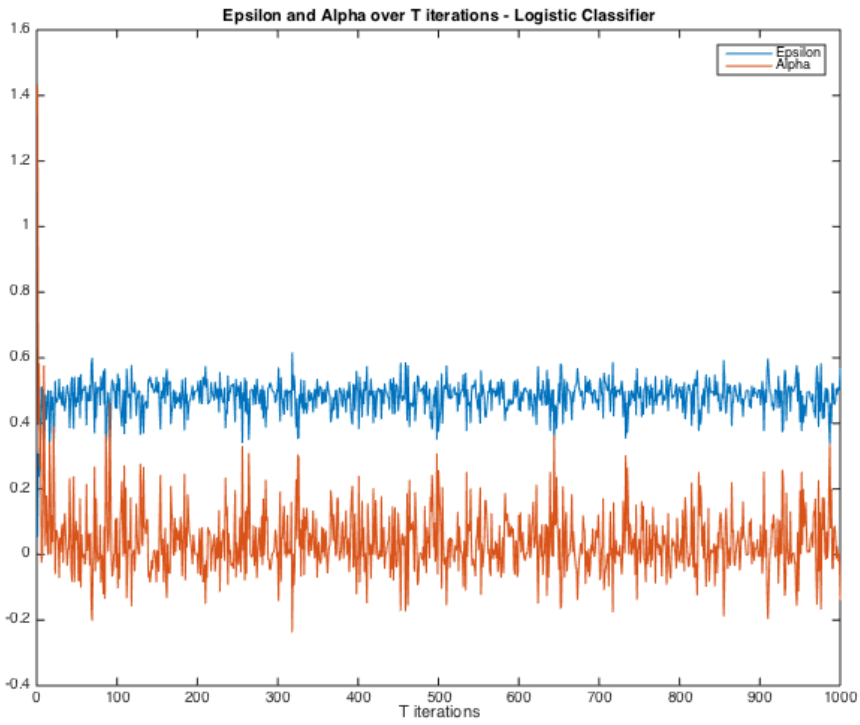

Figure 6: Epsilon and Alpha over T Iterations for Boosted Logistic Classifier

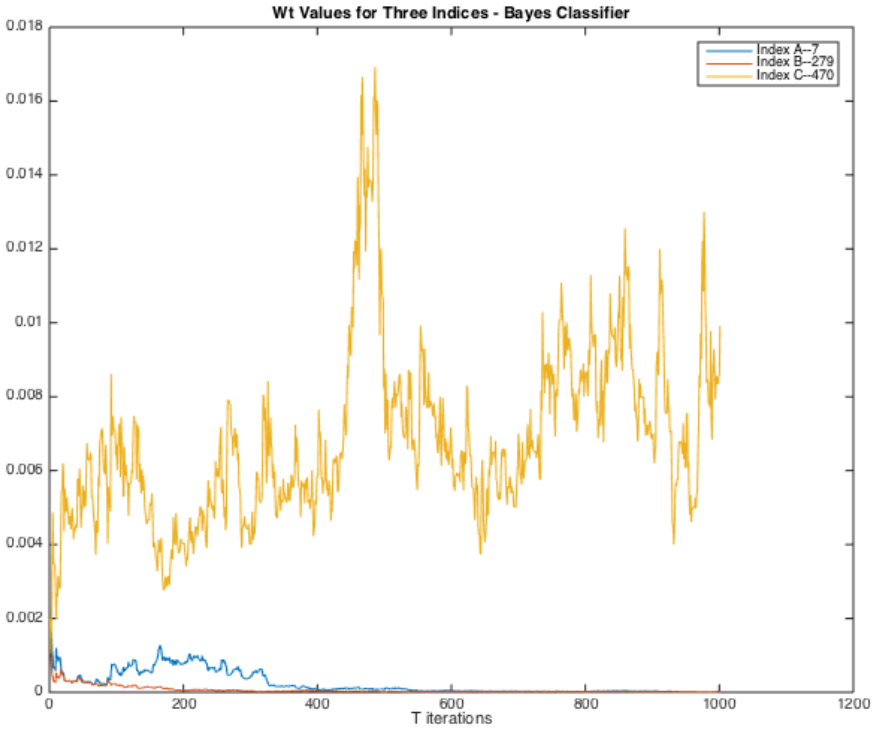

Figure 7: W(i) for points 7, 279, and 470 over T Iterations for Boosted Logistic Classifier