Naive Bayesian Classification
Rubin Stricklin

In this project I created a Naive Bayesian classifier to classify spambase data. I divided the data into training and testing data. Then I got the means and standard deviations for each feature for each class. With these data I used the gaussian Naïve Bayes algorithm to classify the testing data.

Confusion matrix:	Known Negative	Known Positive
Predicted Negative	1012	38
Predicted Positive	382	869

Accuracy = 0.817
Positive precision = 0.694
Negative precision = 0.963
Positive recall = 0.958
Negative recall = 0.725

The naive bayes seems to not be as accurate as the SVM, 82% vs 95%. Because most of the features of this dataset are independent naive bayes does a pretty good job classifying examples. However the top 3 weighted features are related because they all have to do with the number of capital letters in the email, this could be a reason that we saw many more false positives than false negatives.