Project Report Guiding Questions

1. Report the training and testing error of the Naive Bayes classifier.

Training error: 83.54%
Testing error: 84.4%

2. Discuss the results in a short paragraph.

The training error is slightly lower than the testing error. Since the naive bayes classifier is based on counting conditionally independent attributes, the classifier is relatively stable against overfitting because it has low variance. This let's naive bayes perform pretty well on small amounts of training data other classifiers find insufficient. Compared to the log regression on the MNIST dataset from the previous project, this accuracy is much lower because naive bayes is a "naive" classifier and has an inherently biased model.