Hayden Feddock

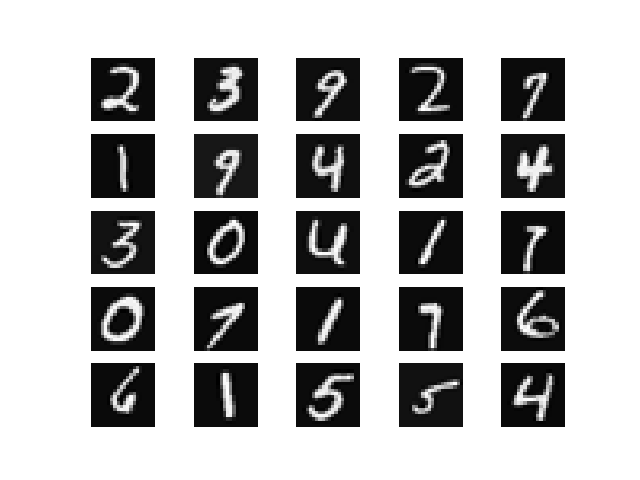
ECE 1395 – Dr. Dallal

Assignment 8

4/14/23

**Question 1**

Part a



Part c

* Subsets saved as .mat files in the input folder

Part d

SVM classification error for training subset 1: 0.01100000000000001

SVM classification error for training subset 2: 0.06000000000000005

SVM classification error for training subset 3: 0.07999999999999996

SVM classification error for training subset 4: 0.07499999999999996

SVM classification error for training subset 5: 0.07099999999999995

SVM classification error for testing data: 0.09199999999999997

Part e

KNN classification error for training subset 1: 0.11699999999999999

KNN classification error for training subset 2: 0.07299999999999995

KNN classification error for training subset 3: 0.135

KNN classification error for training subset 4: 0.124

KNN classification error for training subset 5: 0.10699999999999998

KNN classification error for testing data: 0.14200000000000002

Part f

Logistic Regression classification error for training subset 1: 0.09399999999999997

Logistic Regression classification error for training subset 2: 0.08899999999999997

Logistic Regression classification error for training subset 3: 0.007000000000000006

Logistic Regression classification error for training subset 4: 0.10599999999999998

Logistic Regression classification error for training subset 5: 0.09999999999999998

Logistic Regression classification error for testing data: 0.15600000000000003

Part g

Decision Tree classification error for training subset 1: 0.272

Decision Tree classification error for training subset 2: 0.258

Decision Tree classification error for training subset 3: 0.22099999999999997

Decision Tree classification error for training subset 4: 0.0

Decision Tree classification error for training subset 5: 0.247

Decision Tree classification error for testing data: 0.31999999999999995

Part h

Random forest classification error for training subset 1: 0.09599999999999997

Random forest classification error for training subset 2: 0.06799999999999995

Random forest classification error for training subset 3: 0.08199999999999996

Random forest classification error for training subset 4: 0.08999999999999997

Random forest classification error for training subset 5: 0.0

Random forest classification error for testing data: 0.126

Part i

Majority voting classification error for testing data: 0.09799999999999998

Part j

* Summarize results

|  |  |
| --- | --- |
| **Classifier** | **Classification Error on Testing set** |
| One-vs-All SVM | 0.091 |
| KNN (K = 7) | 0.142 |
| Logistic Regression | 0.156 |
| Decision Tree | 0.319 |
| Random Forest | 0.126 |
| Majority Voting | 0.097 |

* Discussion and comparison of results

According to these results, the One-vs-All SVM classifier performed the best, however it was followed closely by the Majority Voting method. This makes sense because the SVM is very capable of handling a larger number of features, and the majority voting method is very versatile because it combines the predictions from several classifiers which makes it more likely to select the correct one. The poorest classifier was the decision tree with an error of about 30%, this make sense because decision trees can begin to overfit data with lots of features. However, this classifier still contributes to the majority voting classifier because its prediction accuracy is still above 50%. Overall, the process of bagging did help some of the classifiers as well as the majority voting method in order to reduce error in our predictions.