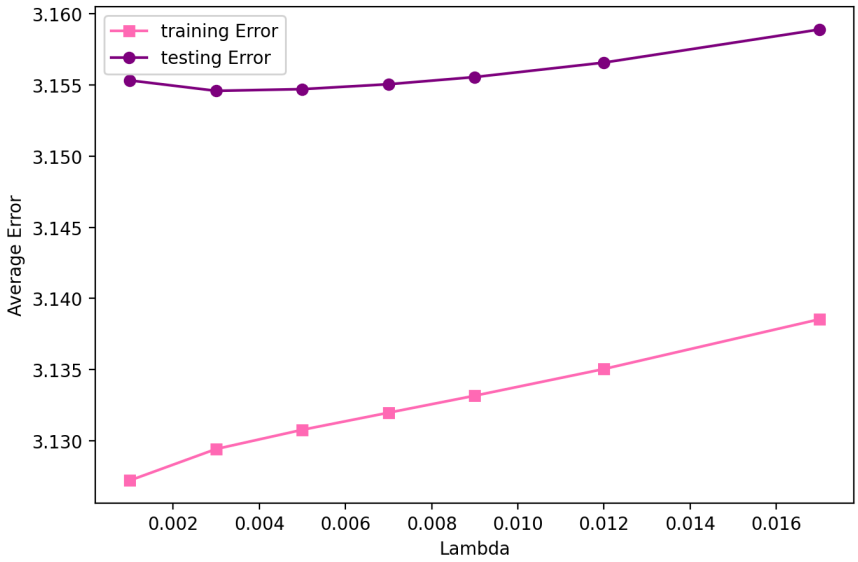
Arissa Buchina
4468171
HW 4

1 - Regularization

b) the size of feature matrix

```
size of the feature matrix X_data:  (1001, 501)
```

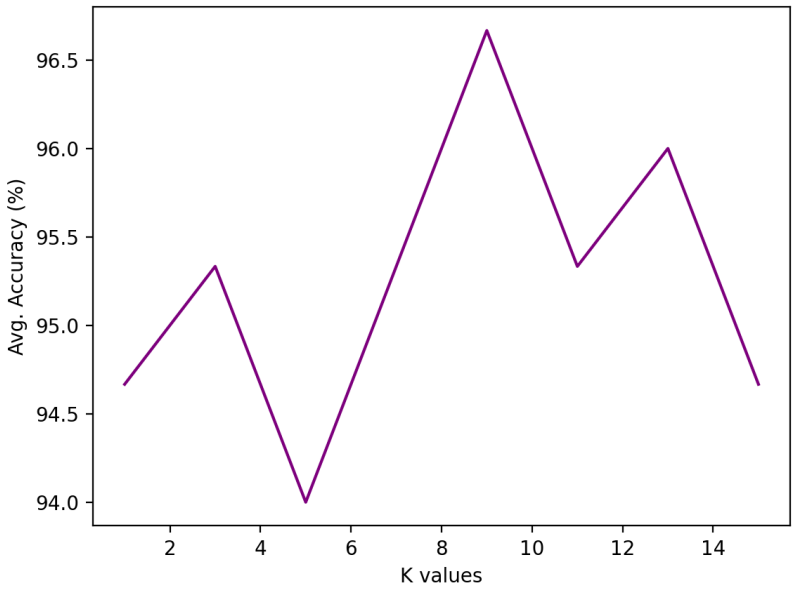
c) For this problem I would suggest using a lambda value of 0.003 since that is where the testing error seems to be the lowest and keeps the training error lower. As seen in the graph, moving up the lambda values from there you start to see an increase in both the training and testing errors. Running the function and creating this graph multiple times however will produce a bit of a difference each time, but most of the outputs I saw around 0.003 lambda would be the best choice to minimize the testing error.



ps4-1-c

2 - KNN: Effect of K

a) For the problem I would suggest using a K value of 9 since it had the greatest accuracy, seen in the graph below. This K value will vary between problems since datasets can vary in size, the number of classes to find distinctions between, and how separated these classes are from each other.



ps4-2-a

3 - One-vs-all

b)

```
one-vs-all y_train accuracy:  93.6 %
one-vs-all y_test accuracy:  84.0 %
```

The accuracy for the training data is 93.6% and the accuracy for the test data is 84.0%. The accuracy for the training data being higher makes sense since the classifiers were trained with that exact data.