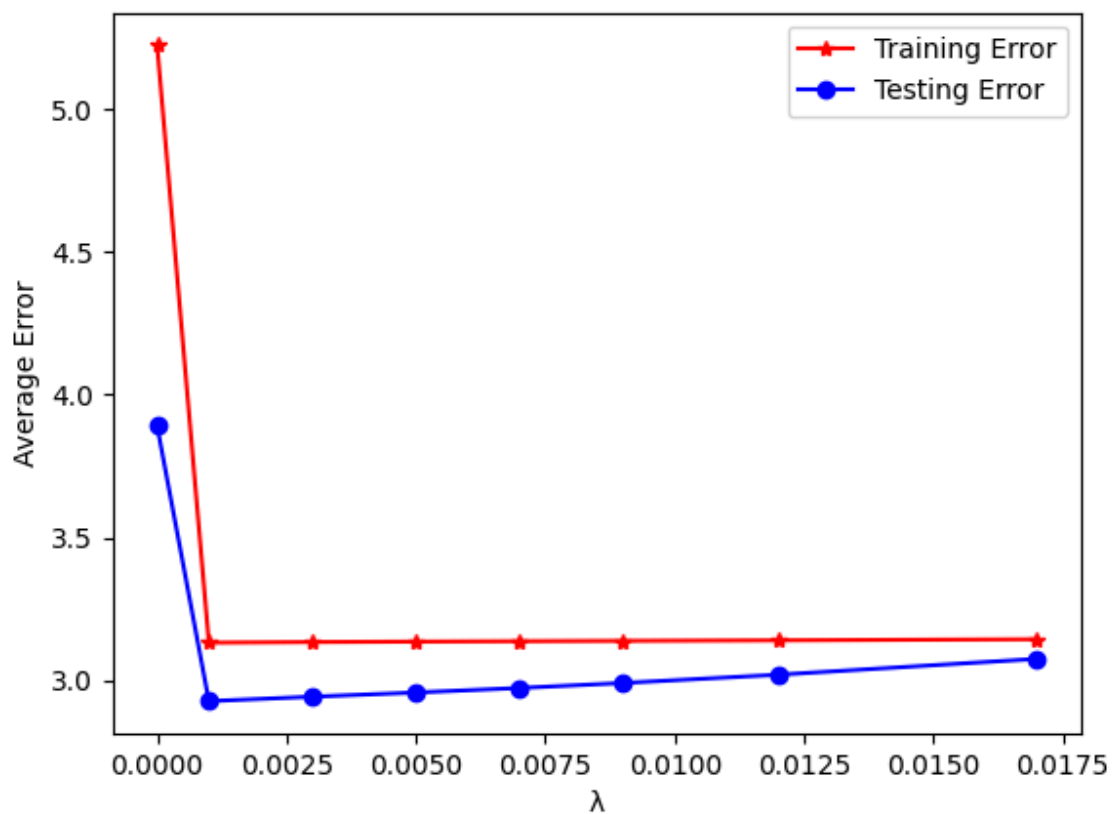


1b.

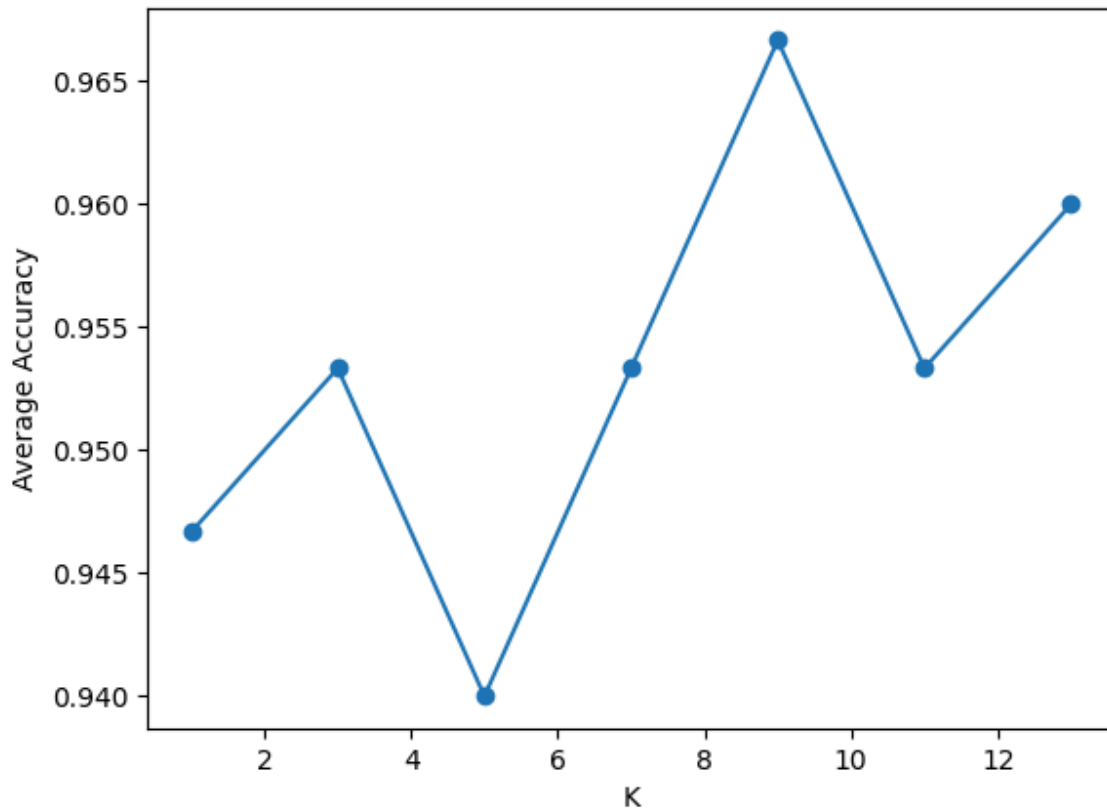
(1b) Feature matrix size: (1001, 501)

1c.



I would suggest a value of 0.001. This is where both the testing and average error are minimized on the figure, and this holds over several trials.

2a.



I would suggest a K value of 9, since this maximizes accuracy for this trial. This value would not necessarily be robust to any other problem since it depends on the data and context of the problem.

2b.

```
dict_keys(['__header__',  
Testing Accuracy: 0.76  
Training Accuracy: 0.8
```

These results suggest that the one-vs-all approach is fairly accurate for the given data set. Other classification approaches may be tested to decide which approach is best, but in this case, OVA provides satisfactory results.