James Schiavo
jcs188
ECE 1395

1b) X shape: 1001 x 501

1c) Lambda = .001 for this particular problem because that is where the difference in cost between the training and testing data is minimized. My figure also shows that there's a high cost when lambda = 0 which means there's no regularization since that term will cancel out, therefore, making regularization for this particular problem necessary due to the amount of features our data has.

2a) I suggest we use k=9 for this particular problem because that is where our accuracy is the highest for our predictions. This value of k is not robust because depending on how far spread out our data is and how well our data is separate this particular value of k might not be ideal. We must always check different values for k to find what makes our model the most accurate because k nearest neighbors algorithm is dependent upon our data.

3b) training accuracy: 0.92
testing accuracy: 0.84
We can see an acceptable accuracy for both my training and testing data sets and an increased accuracy in the training data which is expected. Therefore, the one-vs-all approach for classification is a great way to classify this dataset with multiple classes.