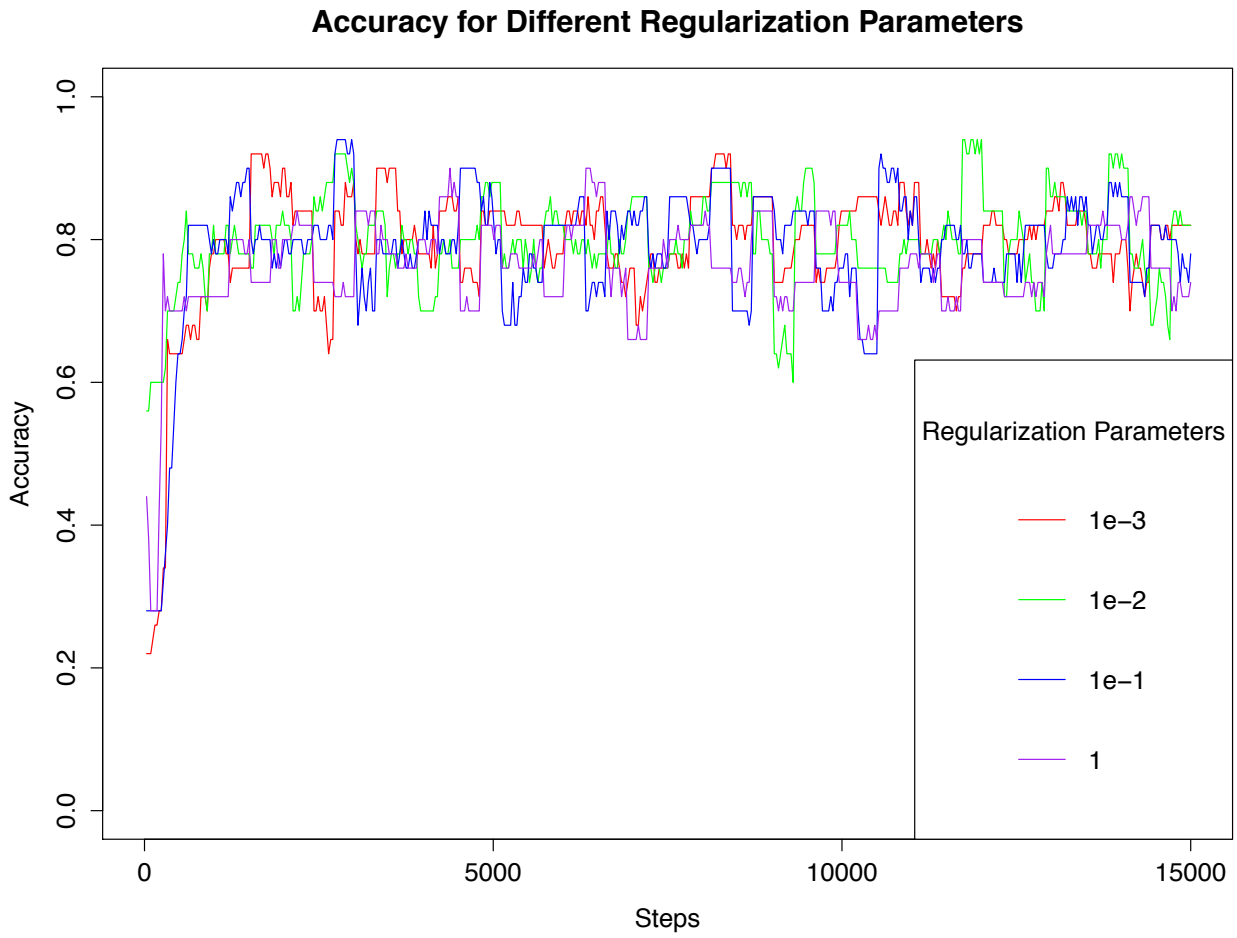
Homework 2 Report

The model used to classify the data was a linear SVM trained using stochastic gradient descent. Step size decreased per epoch by the equation: 1/(p*e+q) where p is 0.1, q is 50, and e is the epoch number.

The best regularization parameter in terms of accuracy on the cross validation data was 0.001. The best classification model did not show significant signs of overfitting on the training data since the difference between the validation accuracy and training accuracy was about 1%. Using this model, the test accuracy ended up being 0.817977067977068, which is consistent with cross validation results.

| Regularization Parameter | Training Accuracy | Cross Validation Accuracy |
|---|---|---|
| 0.001 | 0.810462199928341 | 0.807330057330057 |
| 0.01 | 0.808517172544403 | 0.804668304668305 |
| 0.1 | 0.790013819931412 | 0.780712530712531 |
| 1 | 0.767569227619389 | 0.763104013104013 |


Accuracy for Different Regularization Parameters
Accuracy
0.0
0.2
0.4
0.6
0.8
1.0
0
5000
10000
15000
Steps
Regularization Parameters
1e−3
1e−2
1e−1
1