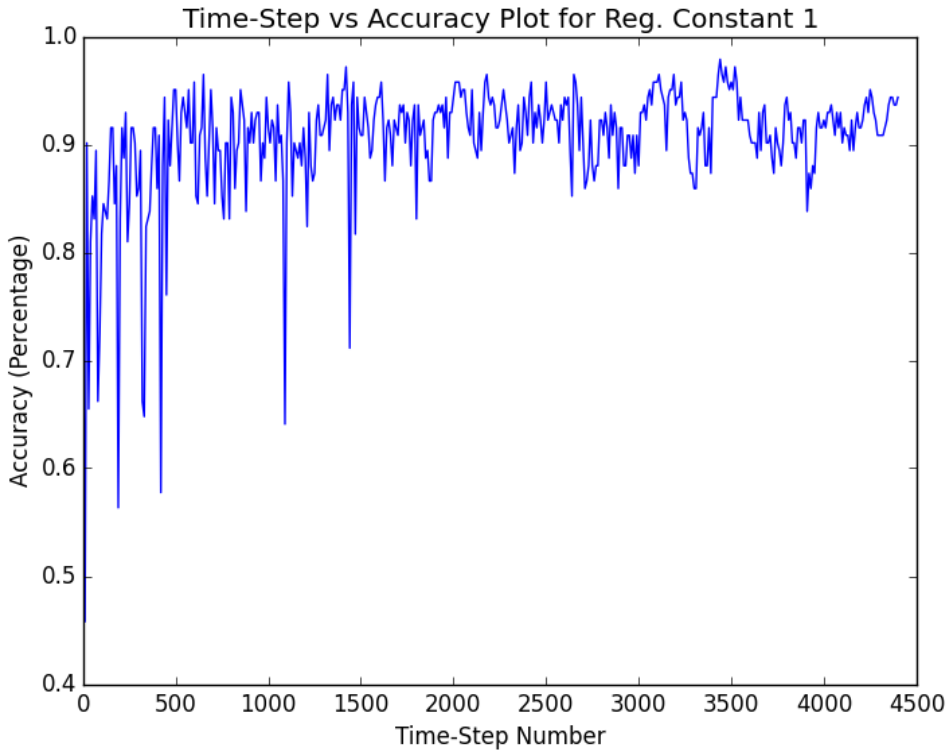
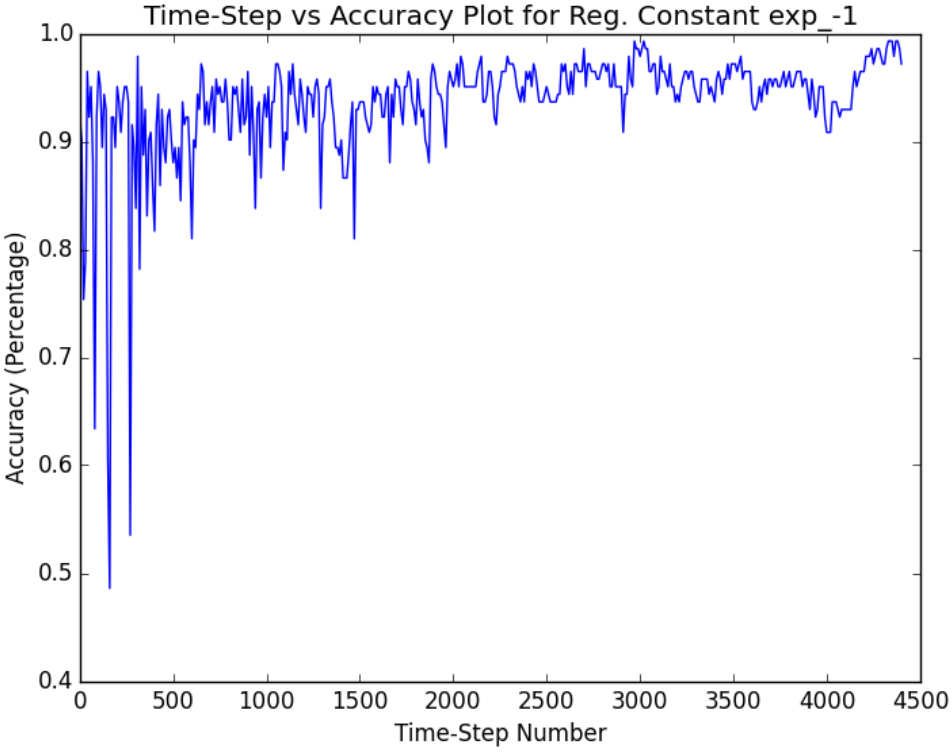
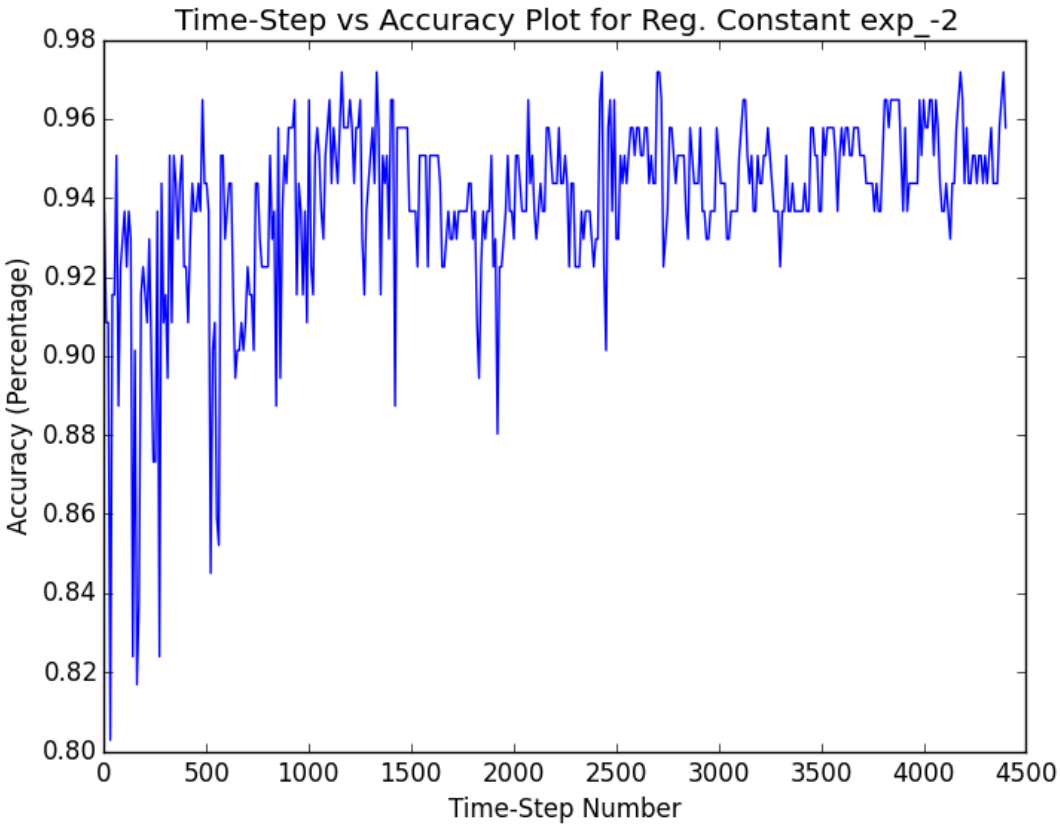
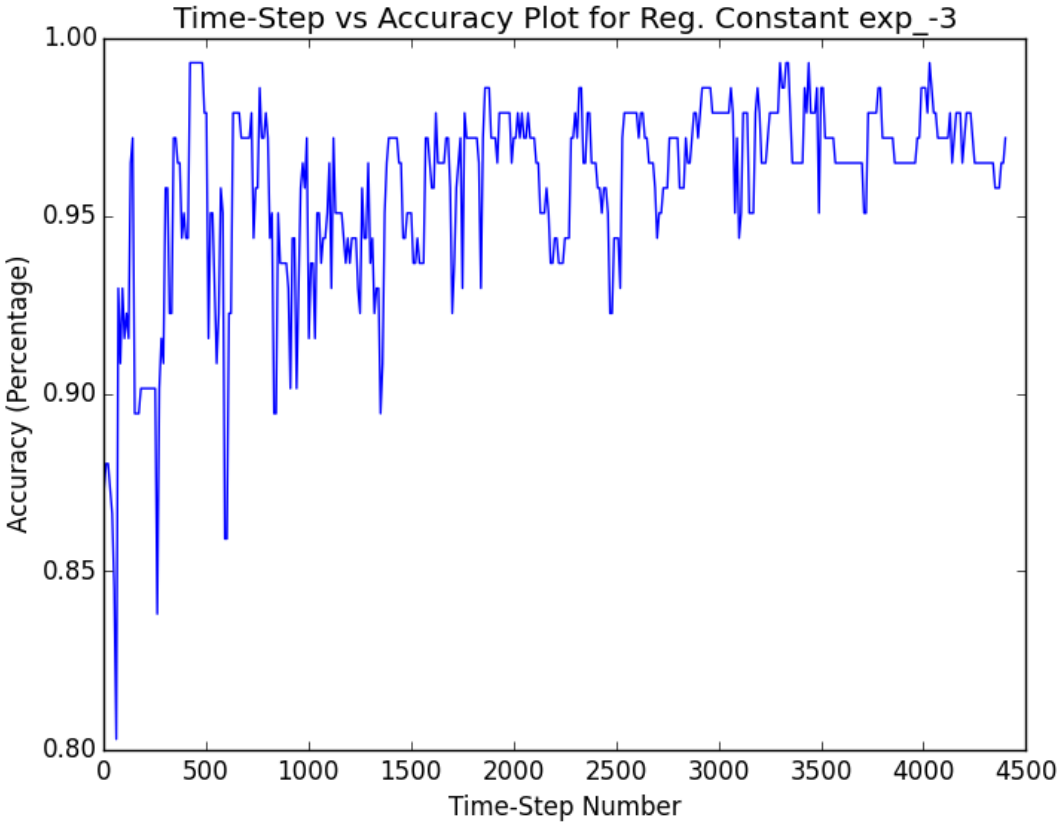
11.1) *Write a program to train a support vector machine on this data using stochastic gradient descent.*

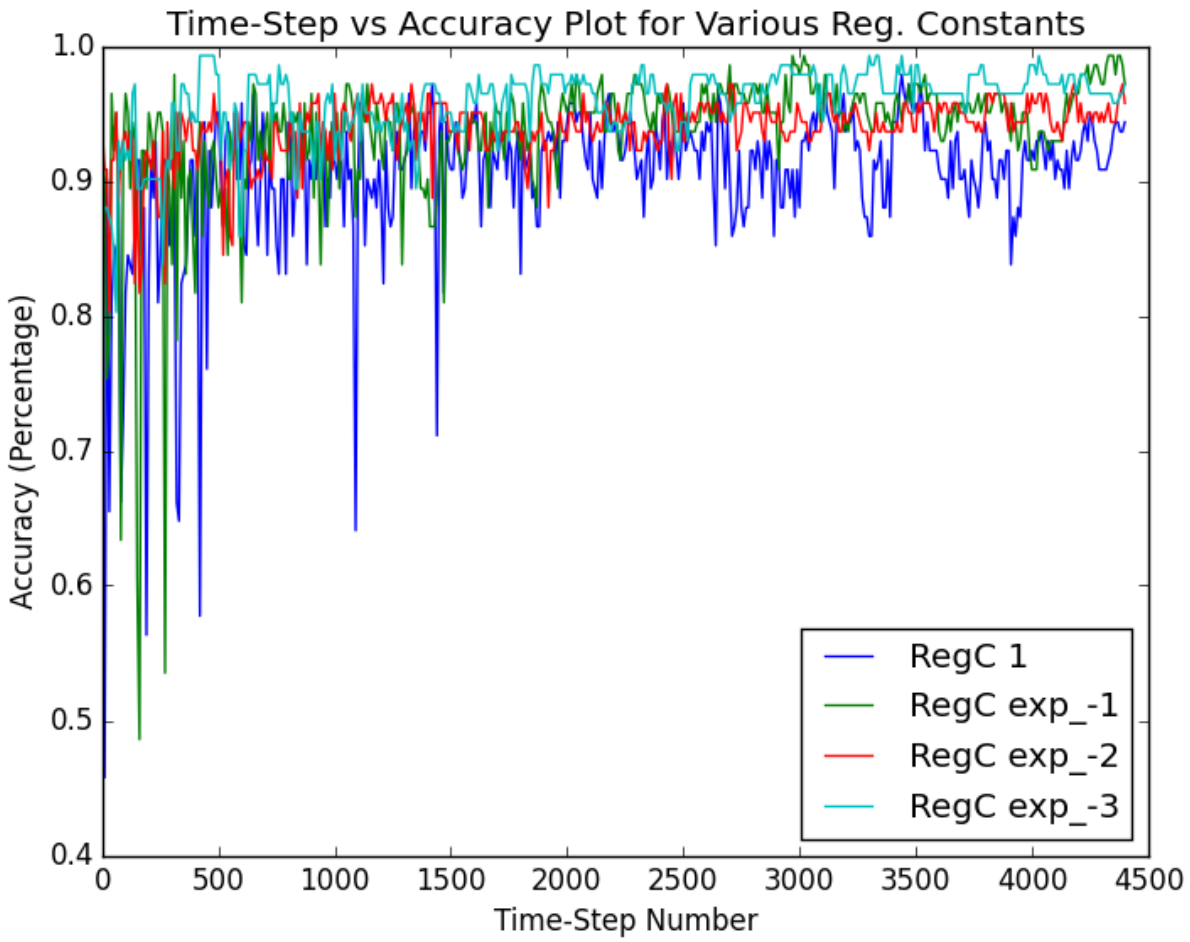
a) *A plot of the accuracy every 10 steps, for each value of the regularization constant.*






Time-Step vs Accuracy Plot for Reg. Constant exp_-2
Accuracy (Percentage)
0.98
0.96
0.94
0.92
0.90
0.88
0.86
0.84
0.82
0.80
0
500
1000
1500
2000
2500
3000
3500
4000
4500
Time-Step Number


Time-Step vs Accuracy Plot for Reg. Constant exp_-3
Accuracy (Percentage)
1.00
0.95
0.90
0.85
0.80
0
500
1000
1500
2000
2500
3000
3500
4000
4500
Time-Step Number



b) *Your estimate of the best value of the regularization constant, together with a brief description of why you believe that is a good value.*

The best estimate of the best value for the regularization constant would be e^{-1}, since there's little variation in its accuracy over time (i.e. it stabilizes fairly quickly) and provides a very-high 97% accuracy by the end of training.

c) *Your estimate of the accuracy of the best classifier on held out data.*

The best stabilized accuracy attained on the comparison between the test data and the validation data was 97.2%, with the biggest peak-accuracy of all steps being 99.3%.