

This homework is group effort of Ling-Hsi Liu (NetId: lhliu2) and Xiaoxia Liao (NetId: xialiao).

- A plot of the accuracy every 30 steps, for each value of the regularization constant

Please see Accuracy.jpg file included in the submission.

- A plot of the magnitude of the coefficient vector every 30 steps, for each value of the regularization constant

Please see the Magnitude.jpg file included in the submission

- Estimate of the best value of the regularization constant and description of why

The best regularization constant is 0.001 among 4 chosen constants: 0.001, 0.01, 0.1 and 1. I chose cross-validation and used all the training set to fit SVM. In the end, 0.001 yields the best accuracy from the validation set, which makes it the best regularization parameter.

- Estimate of the accuracy of the best classifier on the 10% test dataset data

The accuracy of the best classifier is 82.02% on the 10% test dataset data

References:

- data clean at <https://piazza.com/class/jchzguhsowz6n9?cid=220>
- plotting multiple lines at <https://stackoverflow.com/questions/14860078/plot-multiple-lines-data-series-each-with-unique-color-in-r>
- <https://piazza.com/class/jchzguhsowz6n9?cid=415>
- <https://piazza.com/class/jchzguhsowz6n9?cid=332>