CS 498 – Applied Machine Learning

Assignment 2

Problem 1:

* Accuracy of Model Using Different Regularization Constants

A Support Vector Machine classifier was trained on the data using Stochastic Gradient Descent. Several values of regularization factors where used and a plot showing the different resulting accuracies on the held-out data is shown below in figure 1

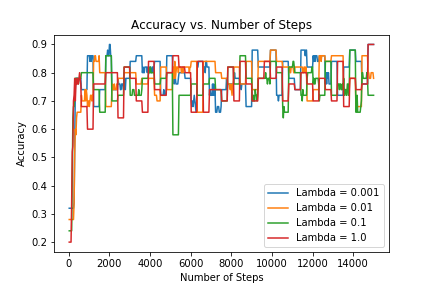


Figure : Accuracy of Held-Out Data vs Number of Steps for Different Regularization Constants

* Magnitude of Coefficient Vector for the Different Regularization Constants

Figure 2 below illustrates the different magnitudes of the coefficient vector achieved with the 4 regularization factors.

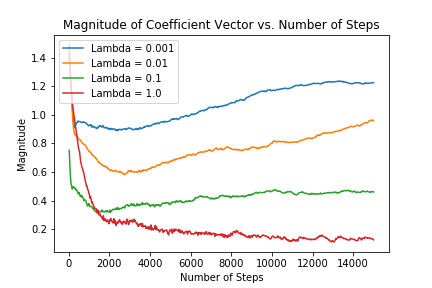


Figure : Magnitude of Coefficient Vector vs Number of Steps for Different Regularization Constants

* Estimate of Best Regularization Constant Value
* Estimate of the Accuracy of the Best Classifier on the Remaining 10% Test Dataset

After choosing the best regularization constant, the model was trained on the original 80% training dataset and then tested on the 10% test dataset.

Accuracy of Model = 0.764