CS 498 – Applied Machine Learning

Homework 2

Training a support vector machine on the Adult Data Set from UCL

using stochastic gradient descent

Larry Poon (lpoon2) & Titus Fong (thfong2)

12th Feburary, 2018

Accuracy against held out set

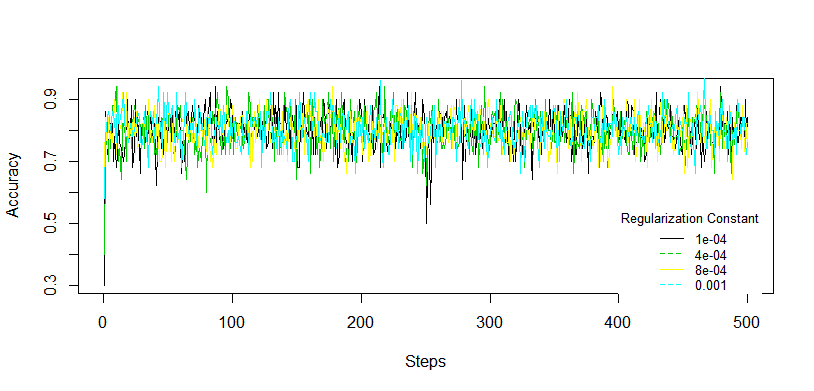


Figure 1. SVM accuracy for every 30 steps per epoch

Magnitude of the coefficient vectors



Figure 2. Magnitude of the coefficient vector every 30 steps

**Estimate of the best regularization constant**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| λ(Regularization const.) | 1e-05 | 1e-04 | 4e-04 | 8e-04 | 0.001 | 0.01 | 1 |
|  | 0.797 | 0.796 | 0.804 | 0.809 | 0.813 | 0.807 | 0.757 |

Table 1. Accuracy on validation set with different regularization constant

The above table shows the accuracy of the support vector machine with different regularization constants. We have tested with constants ranging from 1e-05 to 1. The potential maximum of the accuracy value occurs between regularization constant of 0.0008 and 0.001, since the accuracy drops and increases, respectively, outside of the range. We take the median of the 2 values, **0.009**, as our estimated best regularization constant.

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

For obtaining the estimate of the accuracy, we merged the training and validation sets and retrained our model with our estimated best regularization constant. Then, we test our classifier against the test data. We obtained an accuracy of **0.8081**