CSE 847 Machine Learning
Homework 04 Report
Submitted by: Redwan Sony (sonymd@msu.edu)

Q1. Logistic Regression:
The basic logistic regression is implemented in Python instead of MATLAB. With different number of training samples, the model is trained on and it is tested on the fixed test sample. It is natural that with the increase of number of samples in the training, the accuracy on the test set improved. For all the test cases, I kept the learning rate at 1. The figure 1 below shows different accuracy values for different number of samples in training.

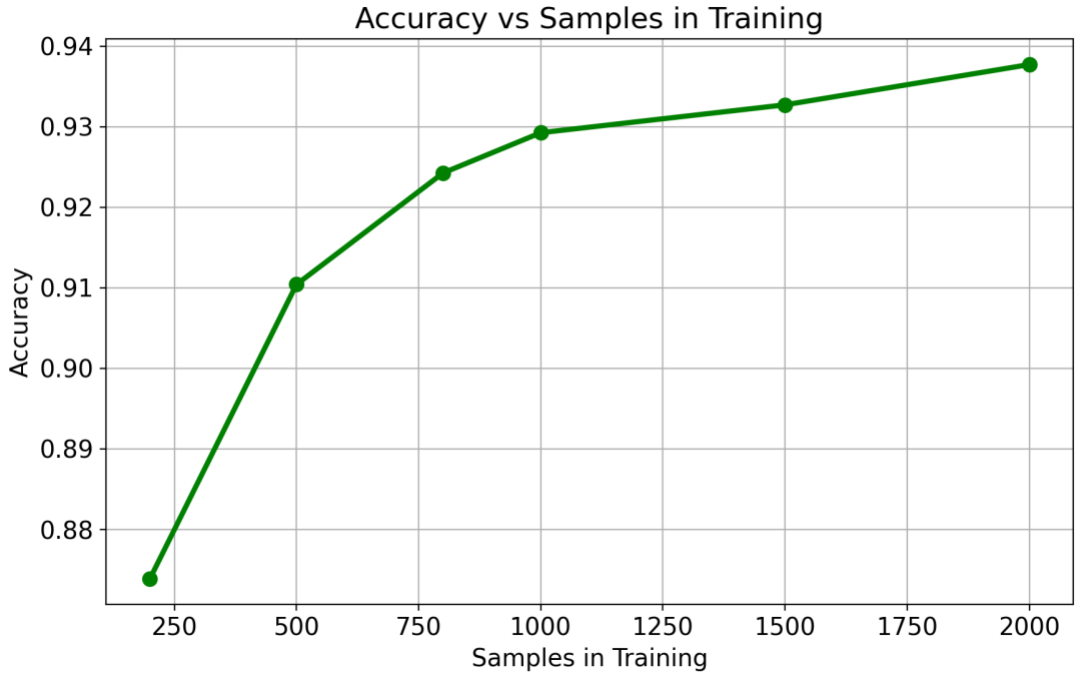

Figure 1

Q2. Sparse Logistic Regression:

Here in this Sparse Logistic Regression process, the regularization parameter is varied to see the effect in the AUC parameter and the number of features selected in the prediction. Figure 2. Shows the change of AUC values with the change of regularization parameter. It is seen that for regularization parameter 0.1, we go the best AUC value.

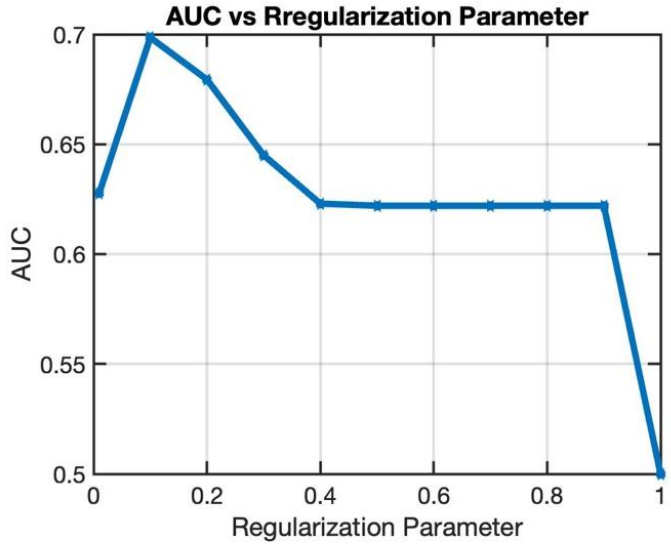

Figure 2

In the following Figure 3, the number of features taken in consideration with the change of regularization parameter is shown. Here for no regularization, all the features are selected. But when regularization parameter is increased, the number of features taken in consideration gets reduced.

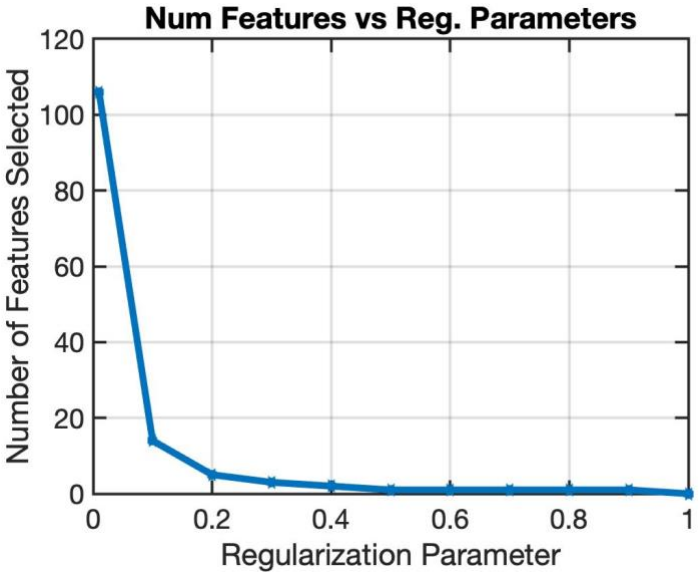

Figure 3

Link of Code Repo: https://github.com/redwankarimsony/cse-847-hw-04