assn4.3

Daniel Loi
dtloi@ucsc.edu 1547401

Kevin Arellano
kcarella@ucsc.edu 1550271

Claudio Sangeroki
csangero@ucsc.edu 1677403

October 2019

1. (a) We are learning 1296 parameters.

   (b) The L2-norm is 34.85.

   (c) The train set accuracy is 97.73%

   (d) For spam class - The Precision is 0.88. The Recall is 0.97. The F1-score is 0.92.

   (e) For ham class - The Precision is 0.99. The Recall is 0.98. The F1-score is 0.99.

   (f) The confusion matrix for the train set is

   ```
   Confusion Matrix
   561     15
   80      3524
   ```

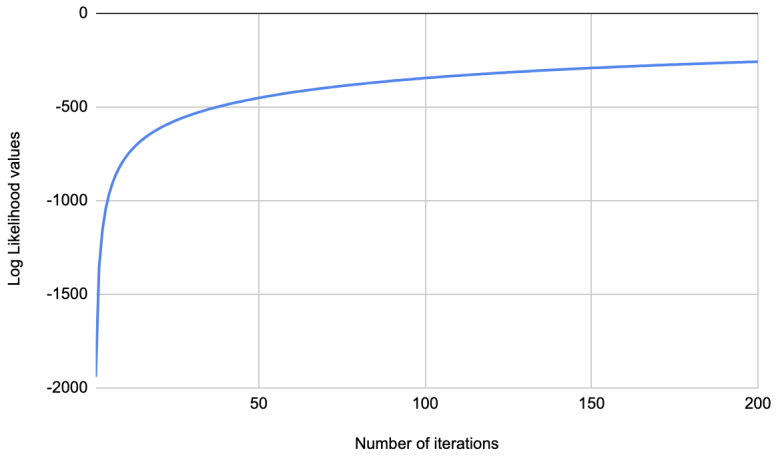
   (g) The test set accuracy is 95.05%

   (h) For positive class, the Precision is 0.73, the Recall is 0.95 and the F1-score is 0.82. For negative class, the Precision is 0.99, the Recall is 0.95 and the F1-score is 0.97.

   (i) The confusion matrix for test set is

   ```
   Confusion Matrix
   162     9
   60      1163
   ```

(j) Graph of log likelihood given below



(k) We can see that it started in a high negative value (around -2000) and started maximizing and converging into a lower negative value (around -300)

(l) Increase number of positive instances in training set

2. (a) We are learning 1297 parameters.

(b) The train set accuracy is 99.67%. The test set accuracy is 98.85%.

(c) The confusion matrix is

```
Confusion Matrix
157     14
2       1221
```

(d) The training and test performance of Logistic Regression with the bias term was better than than the training and test performance of Logistic Regression without the bias term. The addition of a bias term gives the function more leeway when it comes to making predictions.

3. (a) We are learning 1296 parameters.

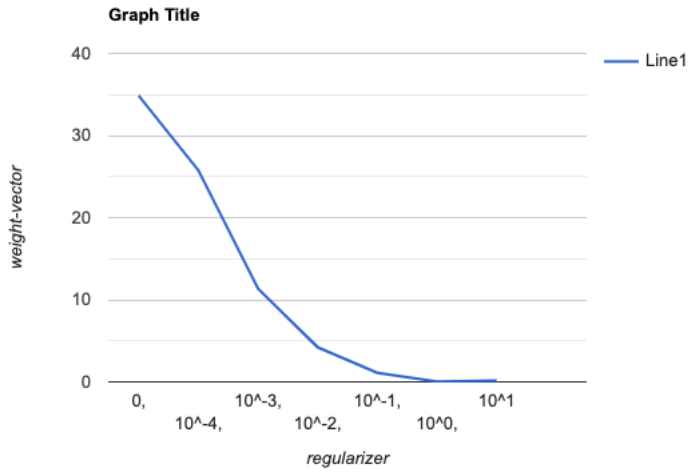
(b) The train set accuracy is 96.36%. The test set accuracy is 95.27%.

(c) The confusion matrix is

```
Confusion Matrix
159     12
54      1169
```

(d) The training set accuracy of the L2-regularized Logistic Regression, at 96.36%, is lower than the accuracy of the vanilla Logistic Regression, at 97.73%. Conversely, the test set accuracy of the L2-regularized Logistic Regression, at 95.27%, is higher than the accuracy of the vanilla Logistic Regression, at 95.05%. We would expect to see this outcome because regularization reduces the accuracy of the training set at hopes of increasing the accuracy of the test set.

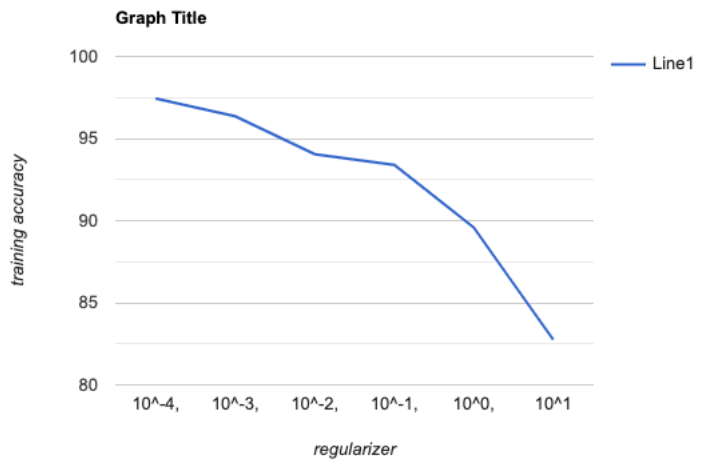
(e) i. Graph below



ii. In this case, there is no difference between the L2]_norm of the weight vector for $= 0$ and L2]_norm of problem 3.1. This can be attributed to the fact that when $= 0$, L2-regularized Logistic Regression is essentially Logistic Regression.

iii. As the value of the regularizer increased, the L2_norm of the weight vector approaches 0. This is because regularization prevents over-fitting by making the weights as small as possible.

(f) i. Graph below



ii. The trend is that training accuracy decreases as the regularizer increases. Regularization is a strategy to prevent overfitting by lowering the training accuracy at hopes of increasing the test accuracy. Because of this, we would expect to see the following trend in the graph.