

CMPS 142 HW3 Part B

Samer Baslan - sbaslan@ucsc.edu - SID: 1367222
Vikram Melkote - vmelkote@ucsc.edu SID: 1562650

June 06 2018

1 Logistic Regression

2.1.1 : There are as many parameters as there are features (1364)

2.1.2 : L2 norm of weight vector: 35.9869

2.1.3 : Train set accuracy: 0.9818

2.1.4 : Train set (positive): P = 0.8903, R = 0.9829, F-1 = 0.9343

2.1.5 : Train set (negative): P = 0.9974, R = 0.9817, F-1 = 0.9895

2.1.6 :

$$\begin{bmatrix} 576 & 10 \\ 71 & 3802 \end{bmatrix}$$

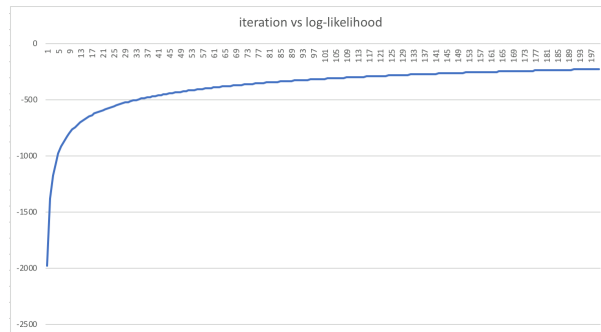
2.1.7 : Test set accuracy: 0.9561

2.1.8 : Test set (positive): P = 0.8111, R = 0.9068, F-1 = 0.8563.

Test set (negative): P = 0.9840, R = 0.9644, F-1 = 0.9741

2.1.9 :

$$\begin{bmatrix} 146 & 15 \\ 34 & 920 \end{bmatrix}$$



2.1.10 :

2.1.11 : Considering that we have significantly more negative instances than positive instances, we can use a dataset with a more equal proportion of "ham:spam". Regularization would help, but we implemented regularization in our code.

2 Logistic Regression with a Bias Term

2.2.1 : We have one more (bias term) parameter than we did in 2.1.1. So, $1364 + 1 = 1365$

2.2.2 :

$$\begin{bmatrix} 144 & 17 \\ 2 & 952 \end{bmatrix}$$

2.2.3 : (Training set without bias), accuracy = 0.9818. (Training set with bias) = 0.9978
(Test set without bias), accuracy = 0.9561. (Test set with bias), accuracy = 0.9830.

In the training set (with bias), the accuracy increased by approximately 0.016%. Yes, including the bias set was useful, because our dataset was biased towards Ham. It shifts our sigmoid function output to account for the bias in our dataset.

3 L2-Regularized Logistic Regression

2.3.1 : Same as the first one, 1364 (since we are not using the bias term here)

2.3.2 :

$$\begin{bmatrix} 147 & 14 \\ 33 & 921 \end{bmatrix}$$

2.3.3 : (Training set without regularization), accuracy = 0.9818. (Training set with regularization), accuracy = 0.9789. In the training set, we can see that we may have been overfitting without the L2 regularization. So after including the regularization term, our performance was still accurate, but eliminating possibilities of overfitting.

(Test set without regularization), accuracy = 0.9561. (Test set with regularization), accuracy = 0.9578. We can see here that the accuracy didn't improve much, reaffirming our claims of overfitting on the training set without the regularization term. This "penalty term" is essentially penalizing high weights.

2.3.4 : L2 Norm (with regularization) = 26.1917. The L2 norm decreased, as expected, due to the penalty term. The regularization term removed the effects of the outlying large weights.