CECS 456: Assignment #1

2/17/2020

Sergio Vasquez
018663448

1 Applying Logistic Regression to Dataset

Algorithm 1 Logistic Regression Algorithm (Batch)

1: **prodedure**(X, y, max_iter, learning_rate) :
2: $w := 0^{1\times d}$
3: For max_iter :
4: $grad = -\frac{1}{N}\sum_{n=1}^{N} y_n \cdot x_n \theta(-y_n w^T x_n)$
5: $\Delta w = -grad$
6: $w := w + learning\_rate \cdot \Delta w$
7: return w

We apply the Logistic Regression Algorithm listed above to the determine how it performs to the dataset. Since the Dataset is not very large we decided to apply the batch variation for the learning algorithm.

```
logistic regression testing...
max iteration testcase0: Train accuracy: 0.834721, Test accuracy: 0.827830
max iteration testcase1: Train accuracy: 0.924407, Test accuracy: 0.900943
max iteration testcase2: Train accuracy: 0.966047, Test accuracy: 0.941038
max iteration testcase3: Train accuracy: 0.973735, Test accuracy: 0.950472
learning rate testcase0: Train accuracy: 0.966047, Test accuracy: 0.941038
learning rate testcase1: Train accuracy: 0.973735, Test accuracy: 0.950472
learning rate testcase2: Train accuracy: 0.978860, Test accuracy: 0.962264
logistic regression test done.
```

Figure 1.1: Train/Test Accuracy for Logistic Regression

Fig [1.1] shows the corresponding training and test accuracy. If we average the results of the Logistic Regression model, the model has an training and test accuracy of 94% and 92% respectively.

2 Feature Transformation

```
3rd order logistic regression testing...
max iteration testcase0: Train accuracy: 0.924407, Test accuracy: 0.898585
max iteration testcase1: Train accuracy: 0.958360, Test accuracy: 0.941038
max iteration testcase2: Train accuracy: 0.970532, Test accuracy: 0.948113
max iteration testcase3: Train accuracy: 0.975016, Test accuracy: 0.955189
learning rate testcase0: Train accuracy: 0.970532, Test accuracy: 0.948113
learning rate testcase1: Train accuracy: 0.975016, Test accuracy: 0.955189
learning rate testcase2: Train accuracy: 0.978219, Test accuracy: 0.964623
3rd order logistic regression test done.
```

Figure 2.1: Train/Test Accuracy for Feature Transformation

We do a third-order polynomial feature transformation to our dataset to test how it does compared to the

previous results

$$\Phi(x) \quad = \quad (1, x_1, x_2, x_1^2, x_1 \cdot x_2, x_2^2, x_1^3, x_1^2 \cdot x_2, x_1 \cdot x_2^2, x_2^3)$$

Fig [2.1] shows the training and test accuracy obtain from applying Feature Transformation. If we average the results of the model with Feature Transformation we obtain a training and test accuracy of 97% and 94% respectively.

3 Conclusion: Which model to choose

Viewing both models we ultimately decided on choosing the linear model with Feature transformation. The linear model with Feature transformation had a better test accuracy which we believe will makes it generalize better for new datasets.