Assignment 2 Report

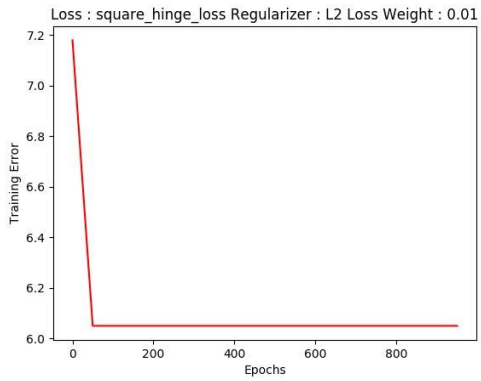
170100035 (Tezan Sahu) & 170040021 (Cherub Kapoor)

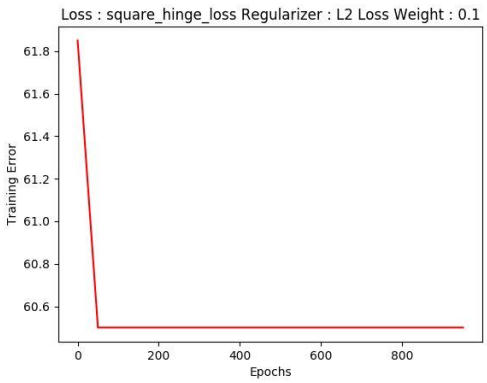
5th September 2018

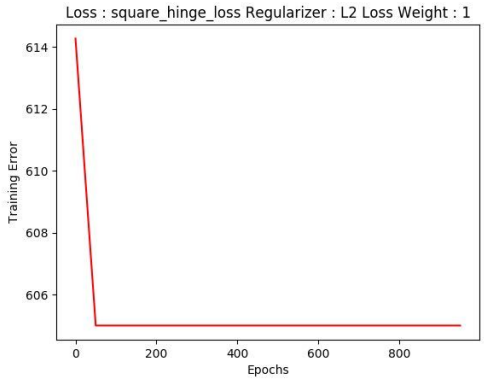
KAGGLE DATASET 1

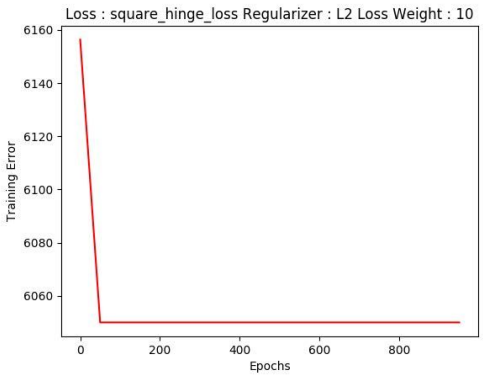
GRAPH PLOTS (REGULARIZED):

- **Hinge Square Loss with L2 Regularizer:**

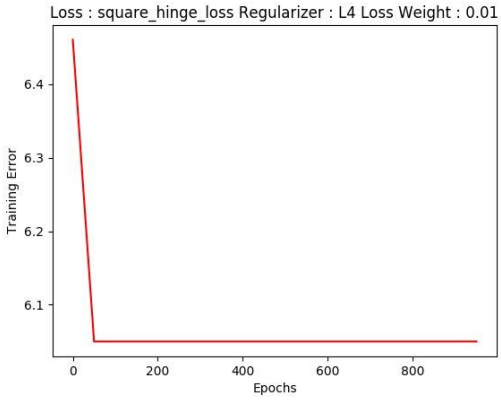


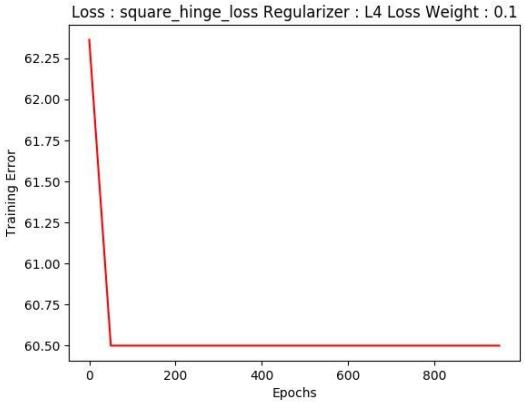


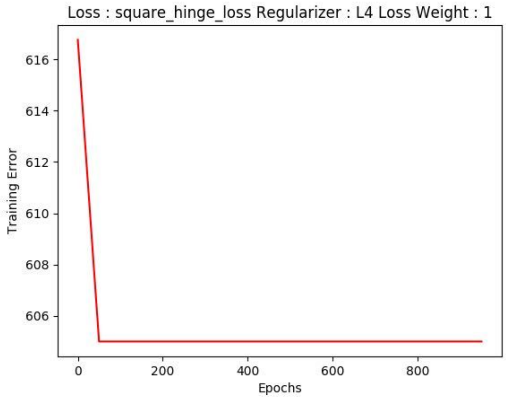


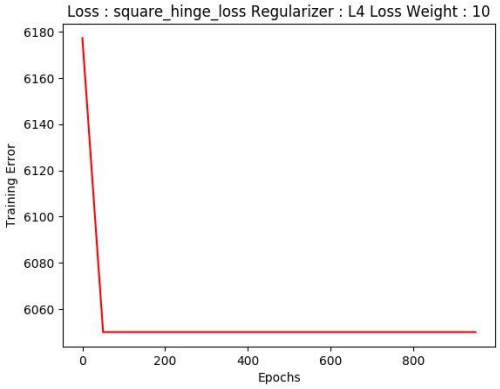


- **Hinge Square Loss with L4 Regularizer**

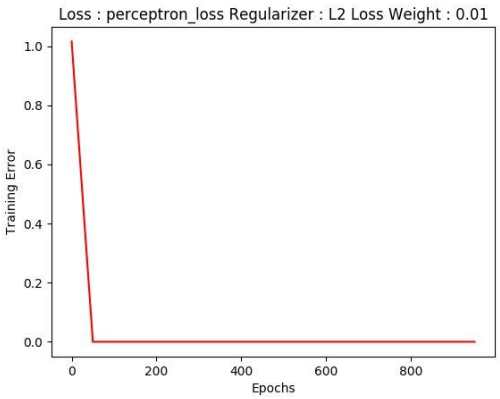


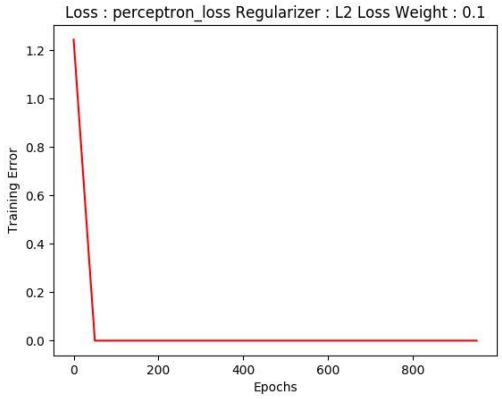


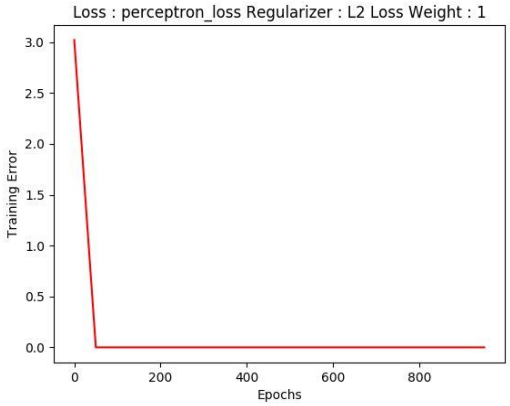


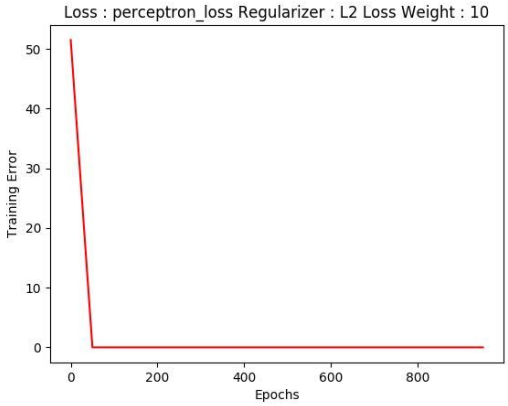


- **Perceptron Loss with L2 Regularizer**

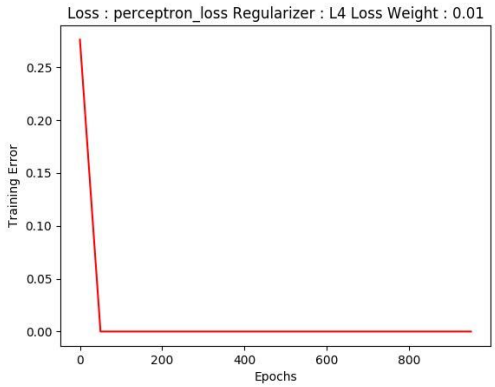


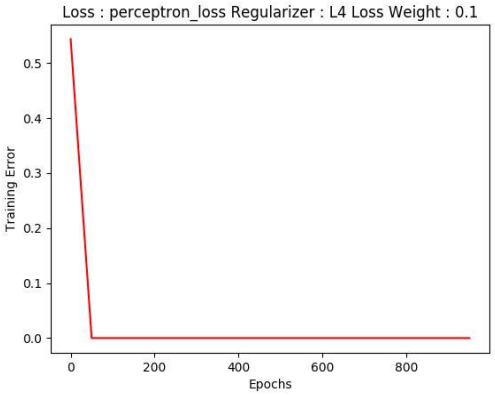


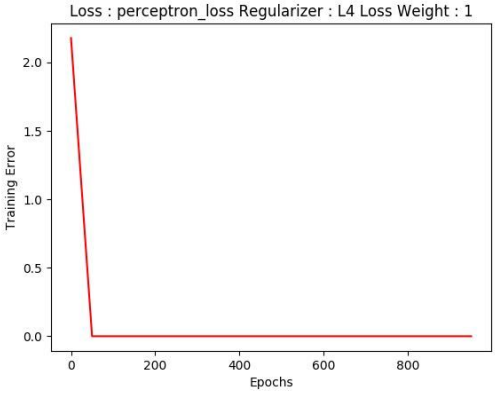


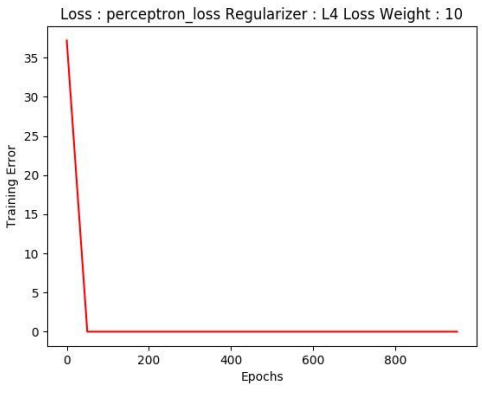


- **Perceptron Loss with L4 Regularizer**

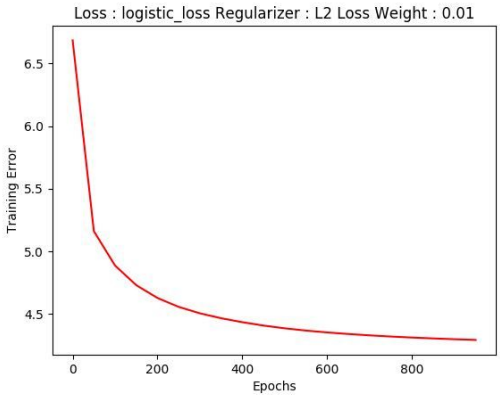


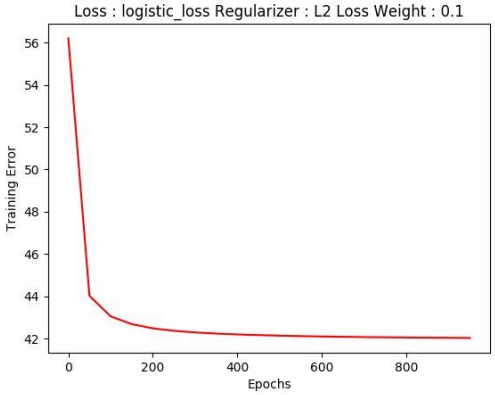


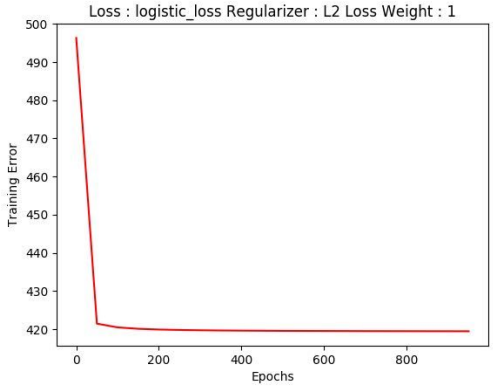


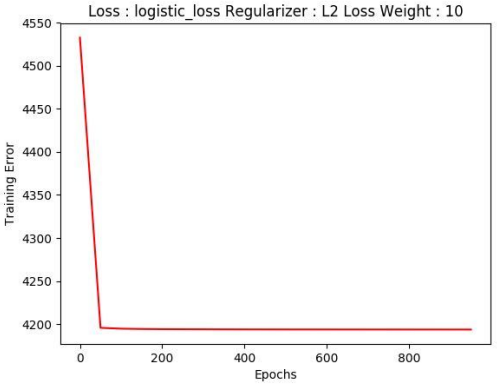


- **Logistic Loss with L2 Regularizer**

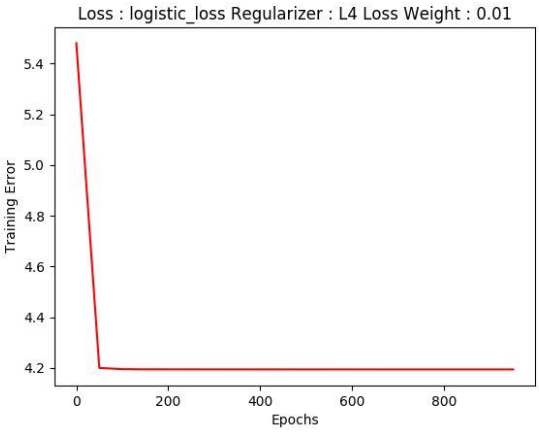


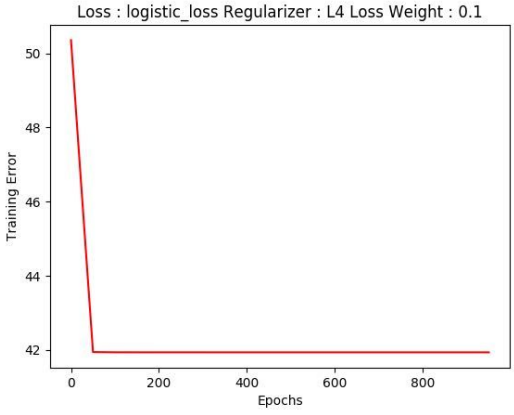


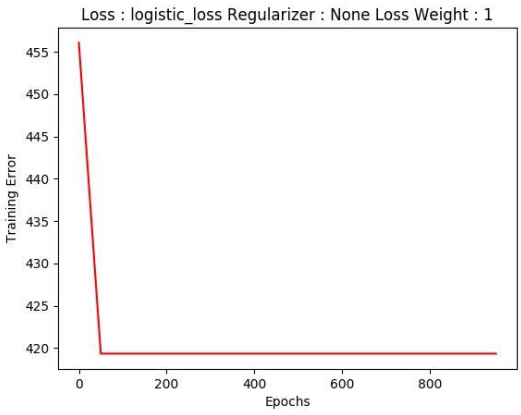


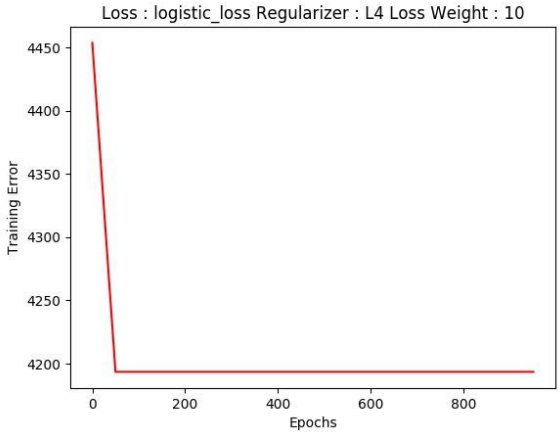


- **Logistic Loss with L4 Regularizer**

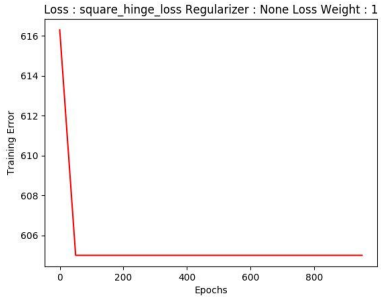


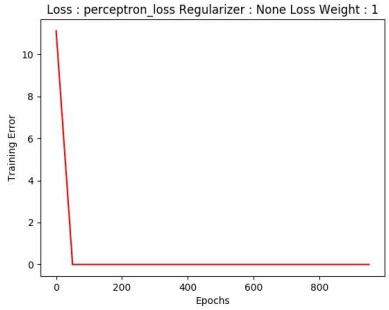


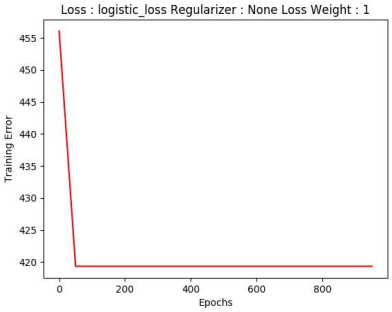




GRAPH PLOTS (NON-REGULARIZED):







DATA PREPROCESSING TO GET RESULT ON KAGGLE:

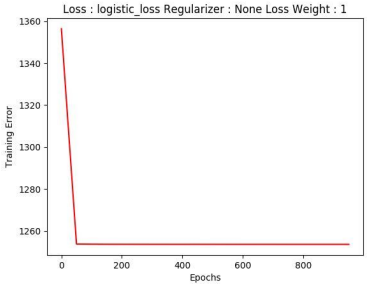
- Used "**Square Hinge Loss**"
- No regularization
- Used **L-BFGS-B** method for optimization
- Batch Size: 64
- No. of epochs: 1000

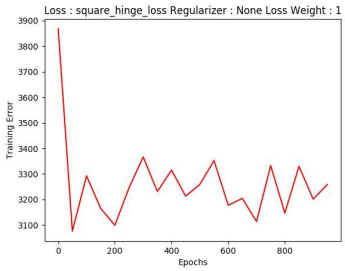
F1 Score obtained: 1.00000

OBSERVATIONS: Almost all the losses (with/without regularizers) lead to similar training losses with increasing epochs. Training converges almost after 50 epochs in all cases.

KAGGLE DATASET 2

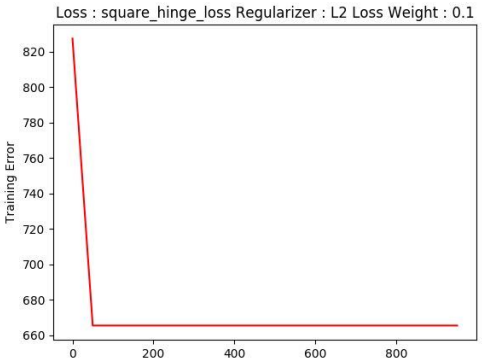
GRAPH PLOTS (NON-REGULARIZED):

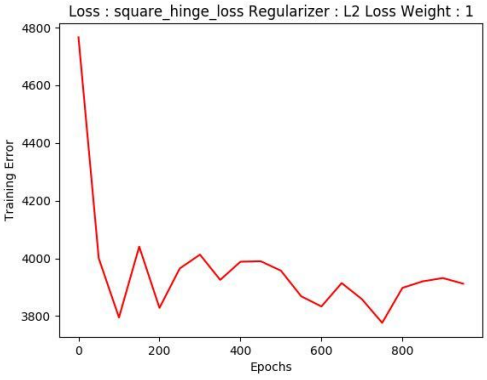


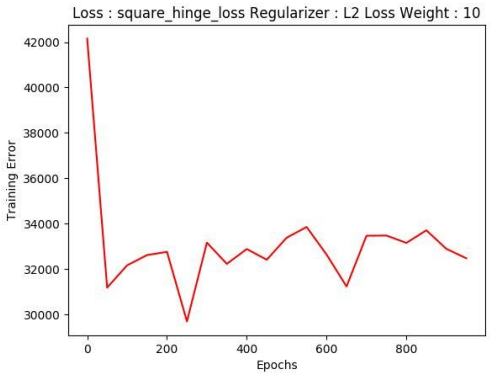


GRAPH PLOTS (REGULARIZED):

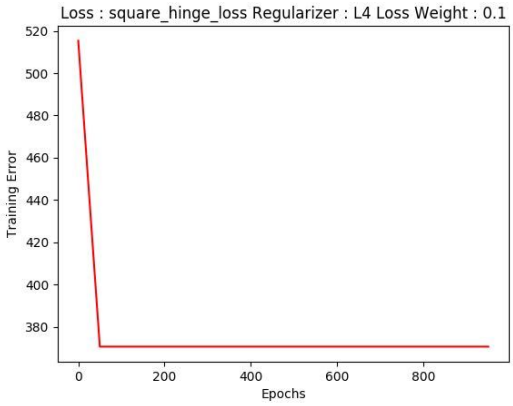
- **Hinge Square Loss with L2 Regularizer:**

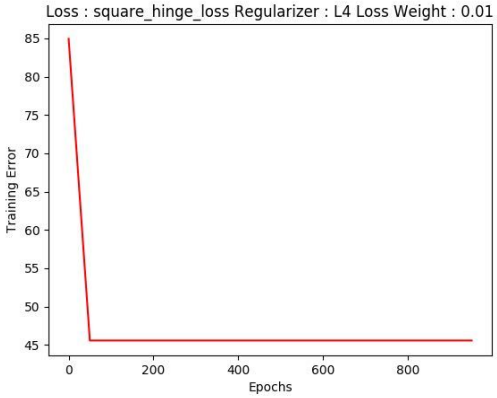


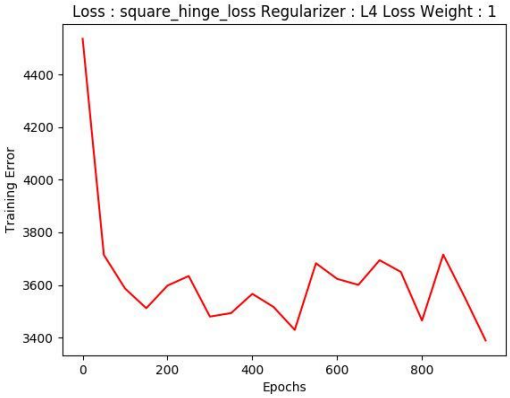


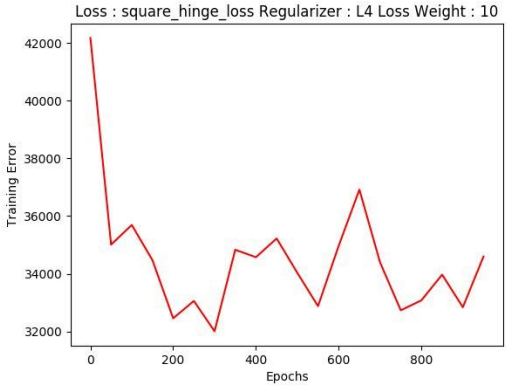


- **Hinge Square Loss with L4 Regularizer:**

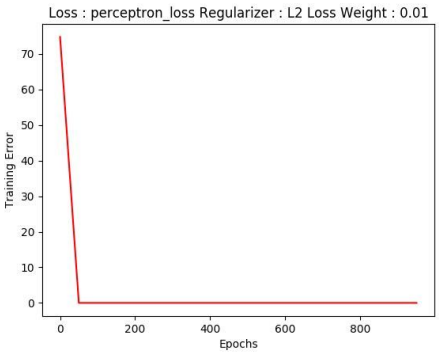


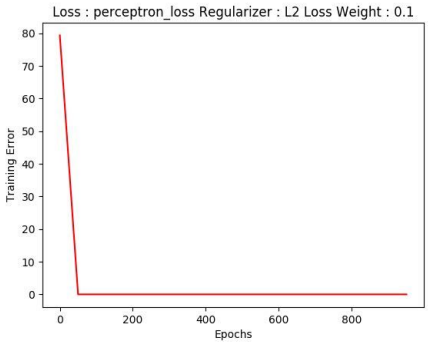


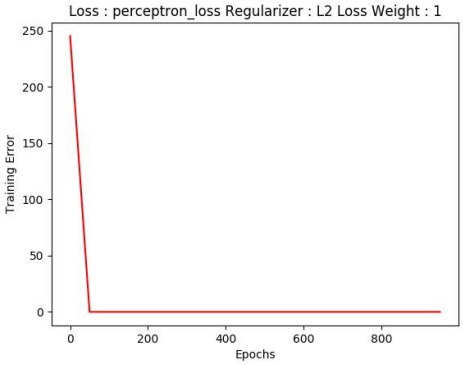


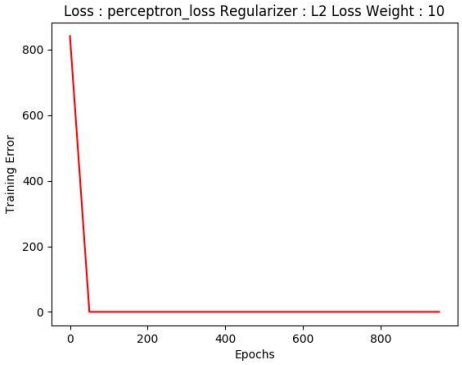


- **Perceptron Loss with L2 Regularizer:**

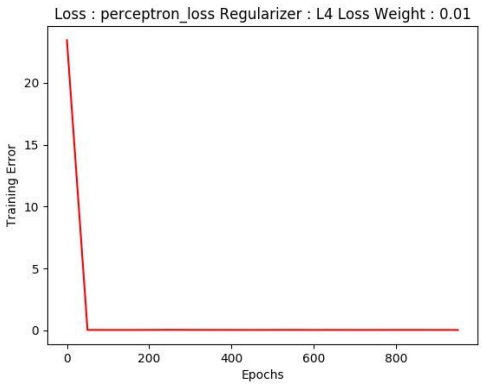


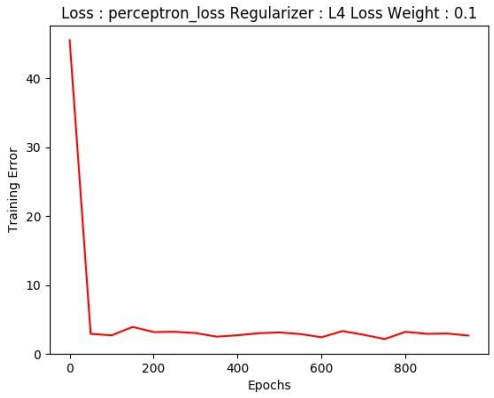


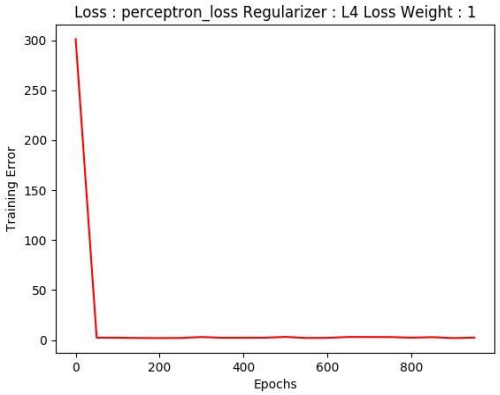


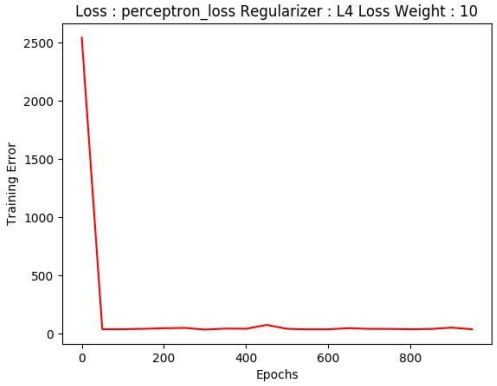


- **Perceptron Loss with L4 Regularizer**

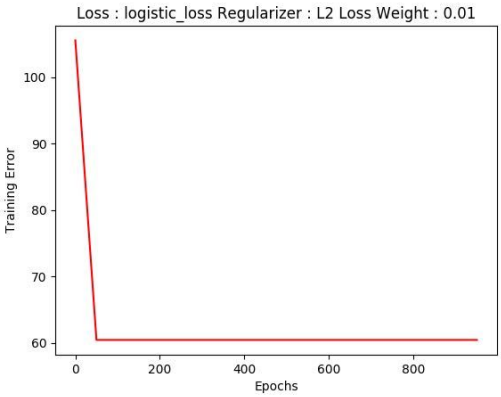


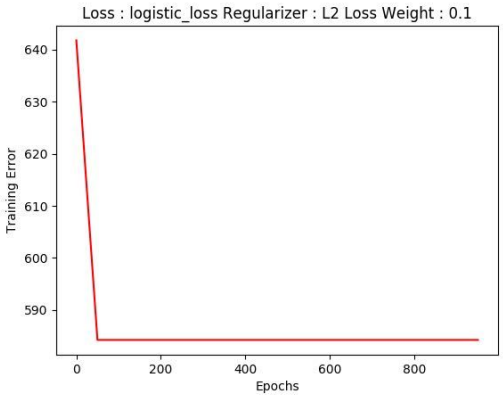


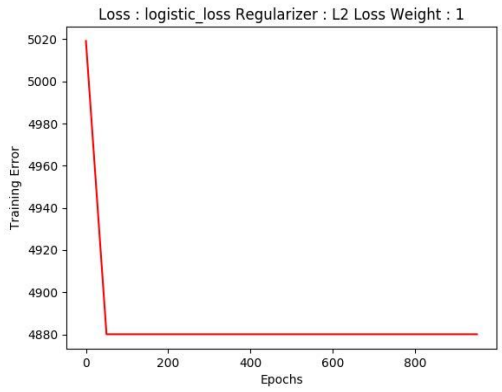


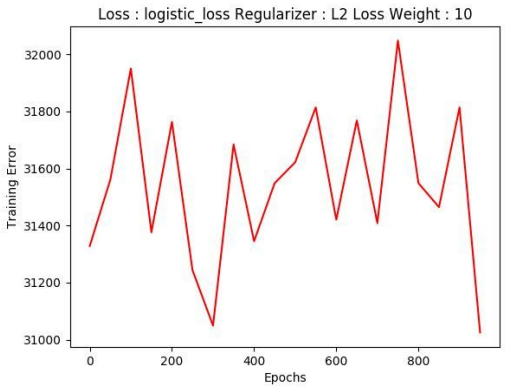


- **Logistic Loss with L2 Regularizer**

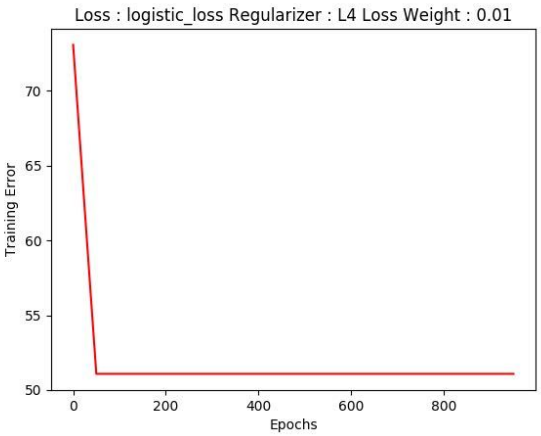


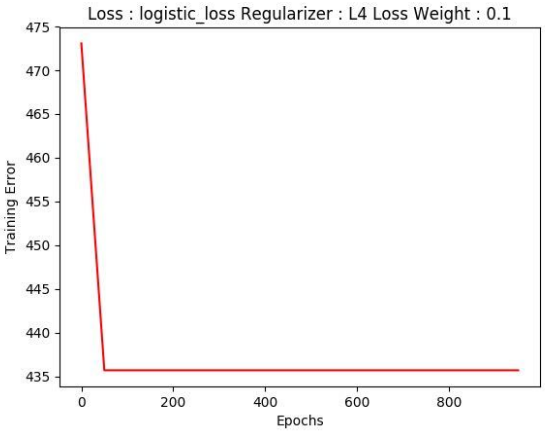


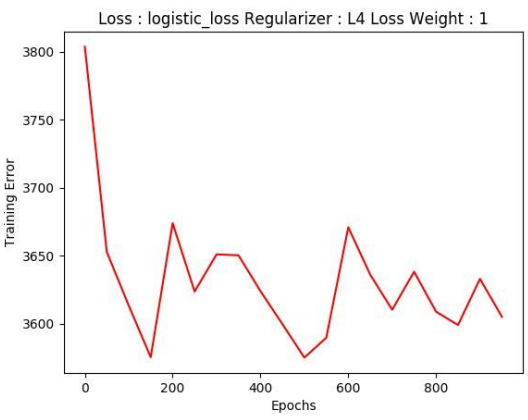


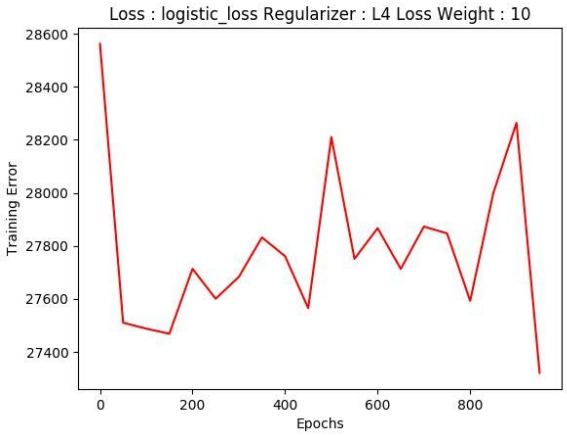


- **Logistic Loss with L4 Regularizer**









DATA PREPROCESSING TO GET RESULT ON KAGGLE:

- Used "**Logistic Loss**"
- No regularization
- Used **L-BFGS-B** method for optimization
- Batch Size: 1000
- No. of epochs:

F1 Score obtained: 0.53996

OBSERVATIONS: Square Hinge, Logistic and perceptron losses converge to an optimum around 50 epochs with L2/L4 regularization and relatively lower value of C. Moreover, due to the use of batches in Stochastic Gradient Descent (done by the optimizer), the training error is not monotonically decreasing with increase in number of epochs.