Assignment-4

Things implemented are:

Multi Layer Perceptron

- Classifier
- Regressor

Functions provided

- Sigmoid
- Tanh
- Relu

Methods Provided

- Stochastic gradient descent
- Mini Batch
- Batch Gradient

Vanishing Gradient problem occurs when the gradient at input layers becomes extremely small such that error becomes constant

It can be resolved using function like Relu

But sometimes Relu runs into Exploding gradient problem which requires gradient normalization and gradient clipping to resolve it.
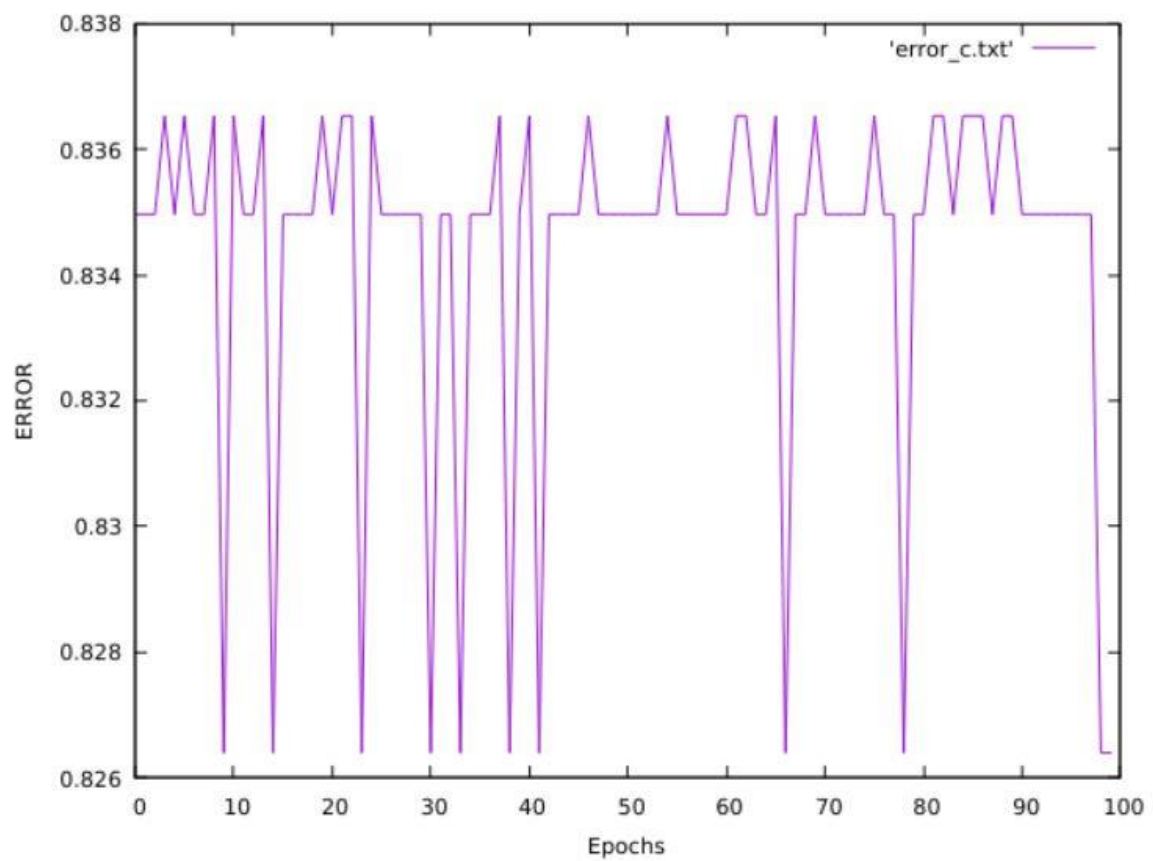
Observations:

Done on 1 Hidden layer and 100 epochs

1. Classifier
    a. ReLU – Accuracy 85.087719
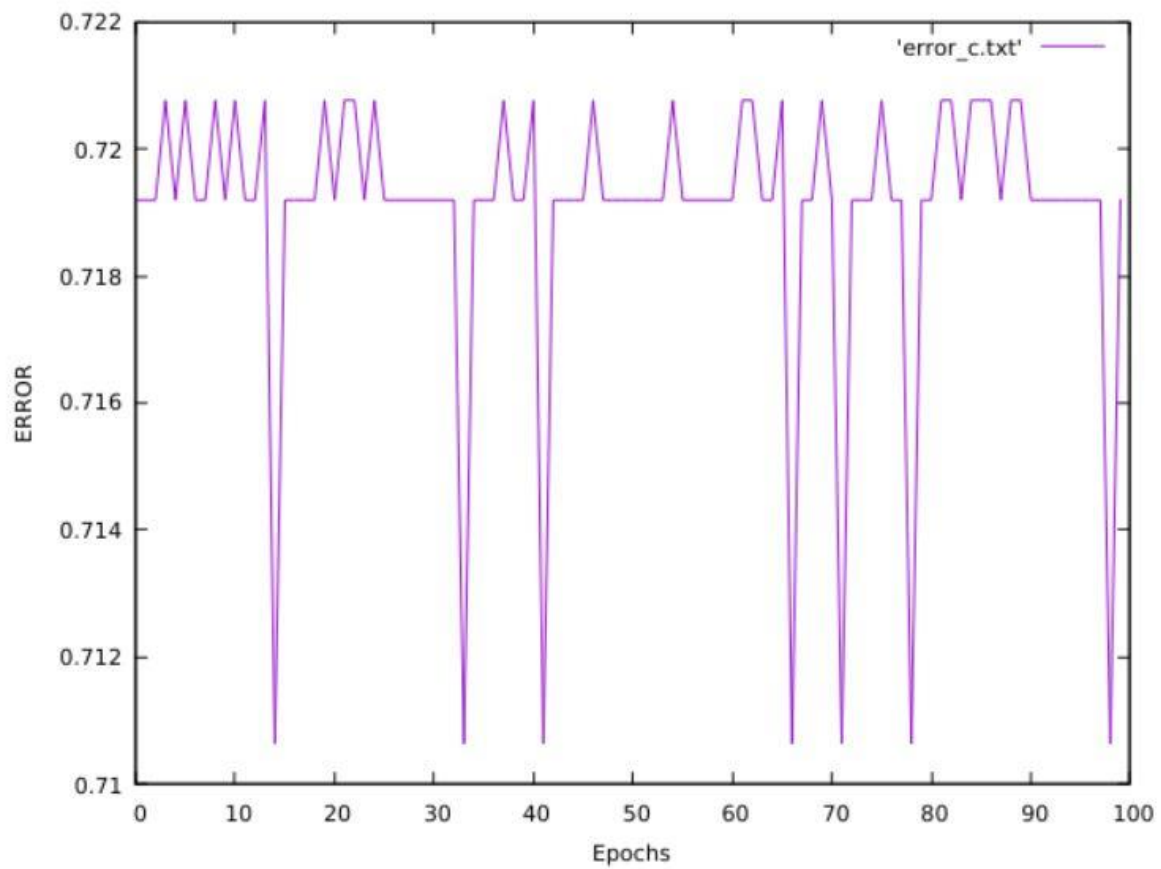   i. Loss function – cross entropy
    b. Tanh – Accuracy 87.719928

i. Loss function – cross entropy
c. Sigmoid Accuracy 90.368972
i. Loss function -cross entropy
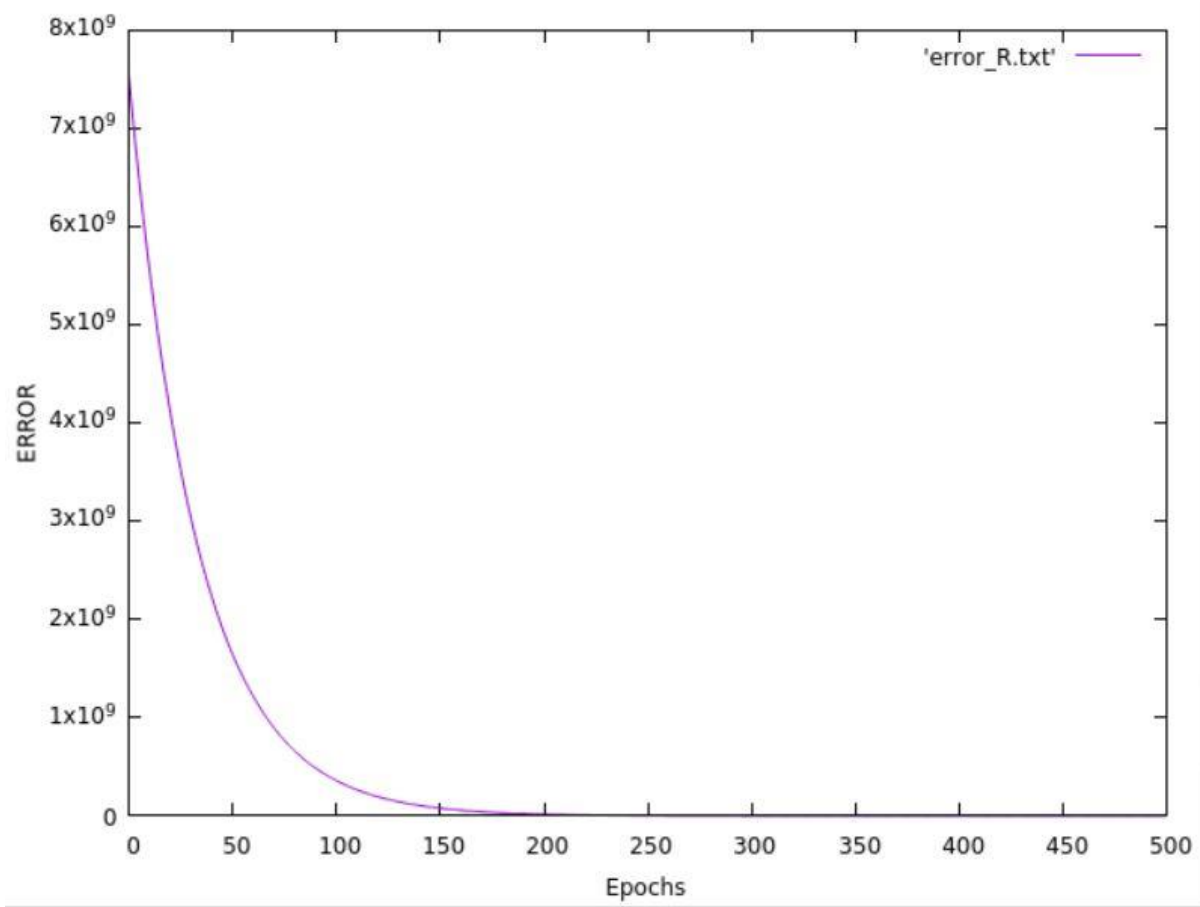
Fast convergence results in local minima

ReLU

SIGMOID

2. Regression Done on 1 Layer and 3 neuron with 500 epochs
    a. Sigmoid cost function MSE
        i. MSE 45.236598
    b. Tanh cost function MSE
        i. MSE 42.912376
    c. ReLU faces Exploding gradient problem

SIGMOID

TANH