CS6910: Programming Assignment 2

Vimarsh Sathia
CS17B046

Indian Institute of Technology Madras — October 31, 2020

General Information

This report only provides metrics about part A of Assignment 2. Part B is on a seperate scanned pdf file(since it's handwritten).

Part-A

Net 4 from the previous assignment was chosen as the model for the regularization experiment. For all the different experiments, the learning rate was fixed at 0.001. Cross entropy loss was considered as the loss function. Each model was trained for 25 epochs, with a batch size of 32.
The loss curves are visualized in figures fig. 1 to fig. 5.

Table 1: Accuracy and loss with different decay values

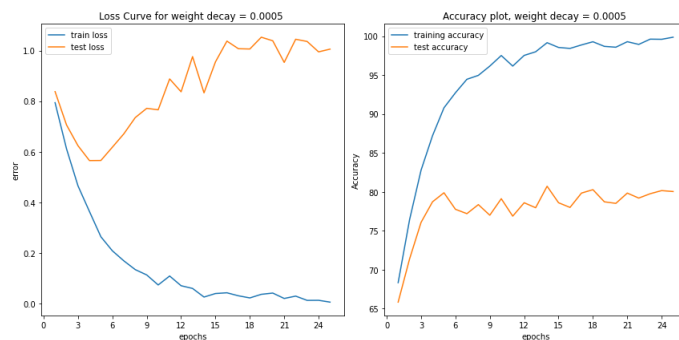| Weight decay	Train loss	Train accuracy(%)	Test accuracy(%)
0.0005	0.00585011	99.856	79.72
0.005	0.00322493	100	82.32
0.05	0.400755	87.14	80.44
0.1	0.750709	70.688	70.24
0.2	1.4262	37.204	37.12

Figure 1: Loss and accuracy curves for weight decay of 5e-4

Inferences

From the loss curves, it is easy to see that the model overfits the dataset when the decay is $5e-4$ and $5e-3$. Best performance(w.r.t) loss history is seen for weight_decay = 0.05. Beyond this value, the model begins underfitting the dataset(due to high bias).
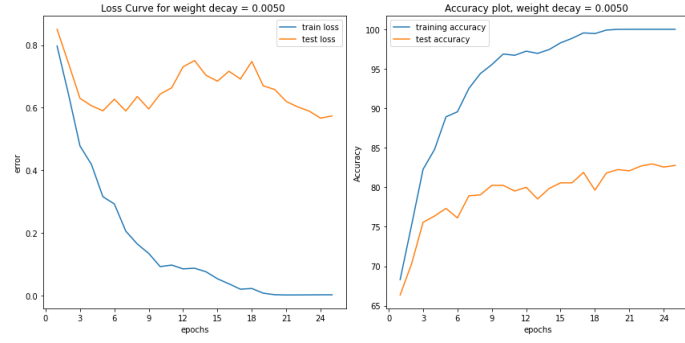
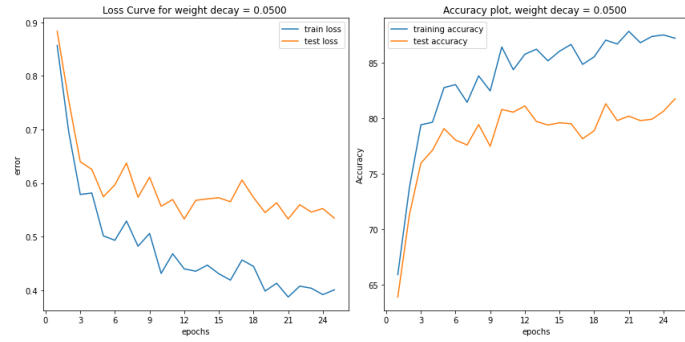Figure 2: Loss and accuracy curves for weight decay of 5e-3

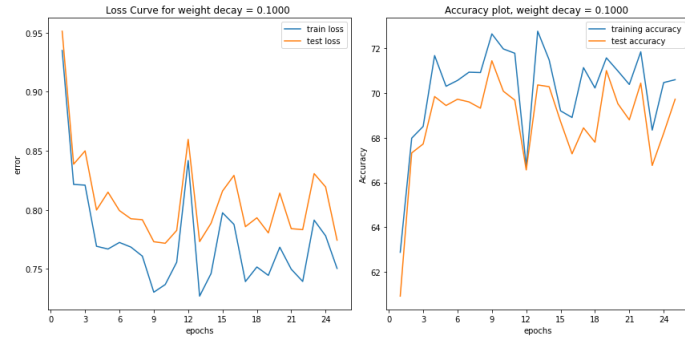Figure 3: Loss and accuracy curves for weight decay of 0.05
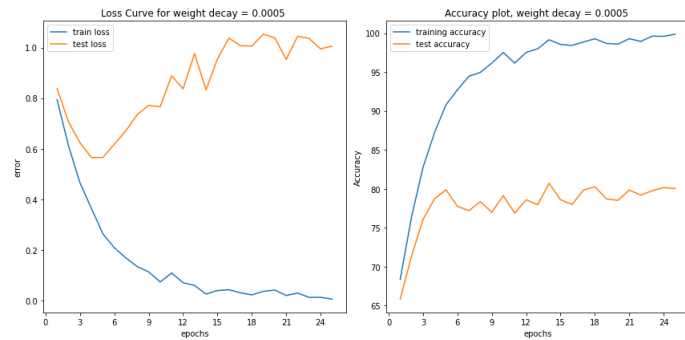

Figure 4: Loss and accuracy curves for weight decay of 0.1

Figure 5: Loss and accuracy curves for weight decay of 0.2