BT3041 – Analysis and Interpretation of Biological Data

Assignment 2

a. Network Architecture

input layer – 784 nodes data image,

one hidden layer – 512 nodes

output layer – 10 nodes for each digit from 0-9

B/w input and hidden layer, **relu** function was used as the activation function

B/w hidden and output layer, **softmax** function was used as the activation function

Error function used for softmax function – **cross entropy** function

b. Activation function

relu function:

$$y(x) = max\,(o, x)$$

softmax function:

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$

σ = softmax

$\vec{z}$ = input vector

e^{z_i} = standard exponential function for input vector

K = number of classes in the multi-class classifier

e^{z_j} = standard exponential function for output vector

e^{z_j} = standard exponential function for output vector

c. Loss function

$$L_{\text{CE}} = -\sum_{i=1}^{n} t_i \log(p_i), \text{ for n classes,}$$

where t_i is the truth label and p_i is the Softmax probability for the i^{th} class.

d. Learning Algorithm

Mini Batch gradient descent

4000 epochs

1 epoch → 40 batches of size 20 each.

e. Learning Rate

constant learning rate of 0.01

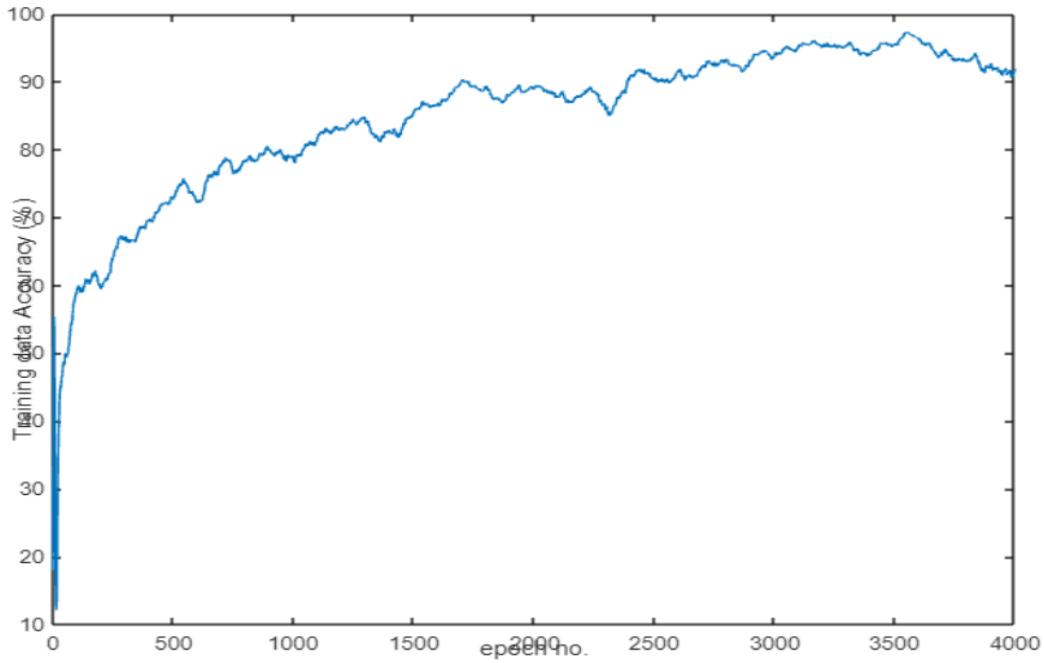
f. Loss and Accuracy of testing data set

Final Accuracy – 71.5%

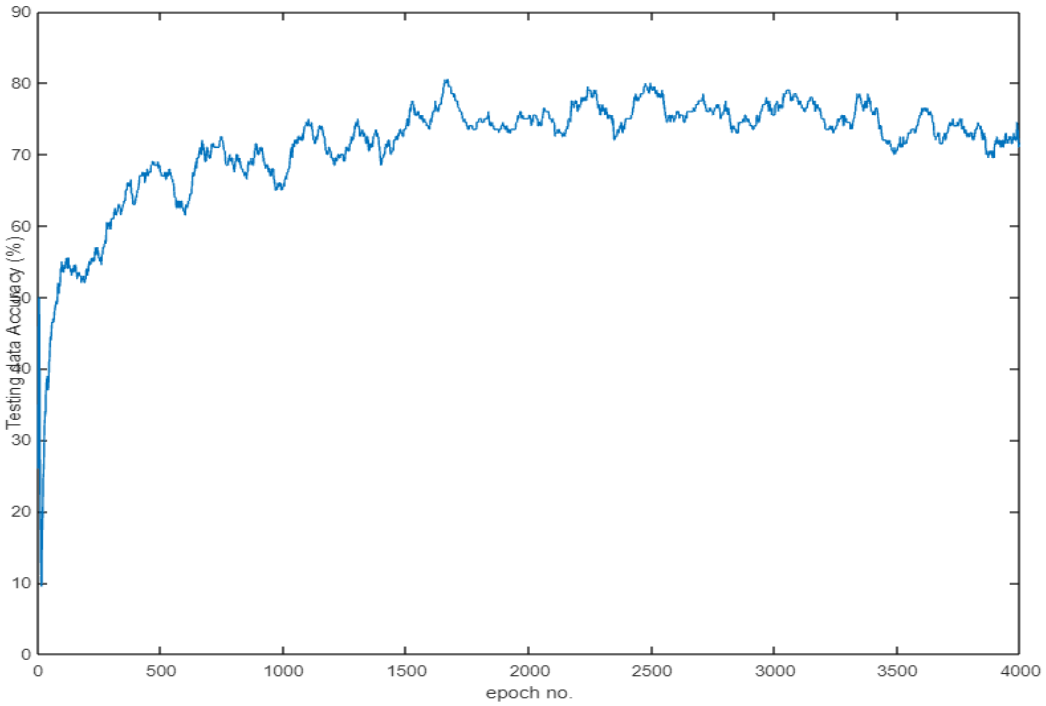
Final average Loss – 0.0018

PLOTS

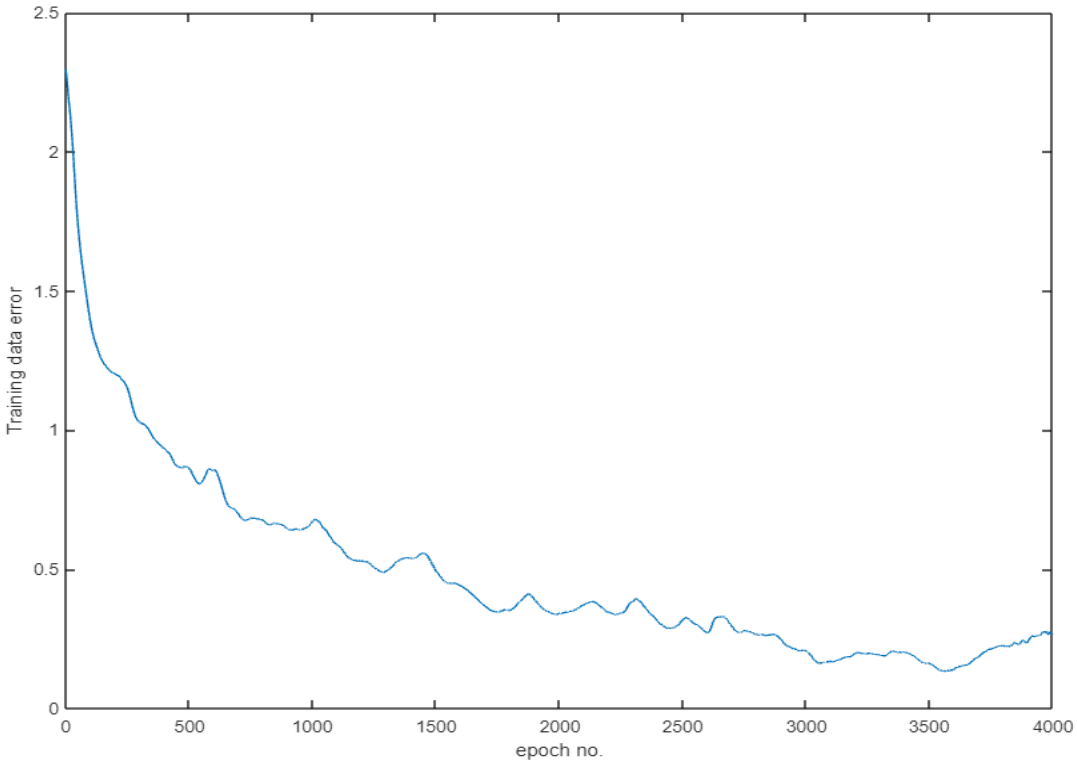
Training data accuracy



Testing data accuracy



Training data Loss



Testing data Loss