Neural Networks

Introduction

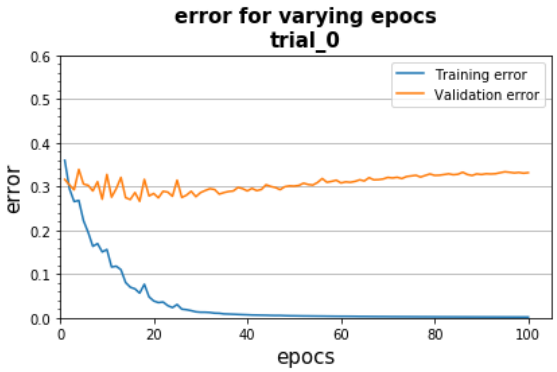
In this Lab I implemented the neural network to classify the handwritten digits for the MNIST database given in the form of 5000 by 400 dataset.

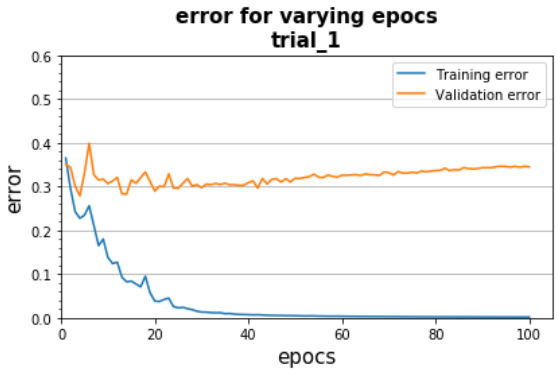
Implementation

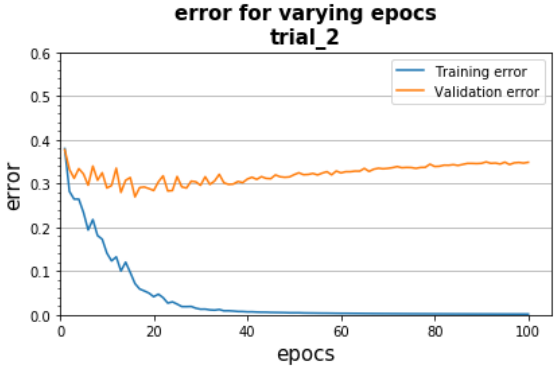
For this lab I used one hidden layer containing 500 nodes. For training, in the forward pass I have taken the tanh function as the activation function between input to hidden layer and hidden layer to output layer. Training of the model is done in the function neuralnetwork whose parameters are training set, number of hidden layers and number of epochs. This function is returning errors on training set and validation set along with the parameters learned. forward propagation and backward propagation are implemented in neuralnetwork(). Error is calculated using the cross entropy function in error() function which takes parameters learned and dataset as its parameter. Graphs are plotted for the training error and validation error for the 5 different trials. A graph is plotted by taking average of all 5 trials as in fig.

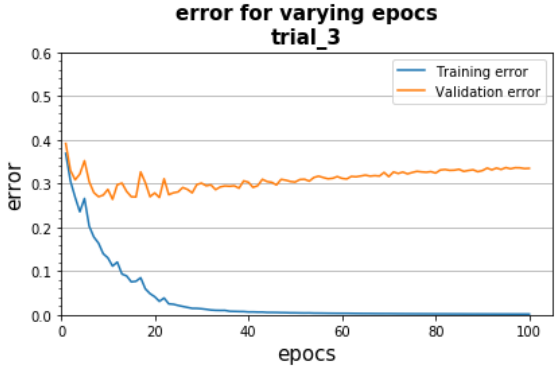
Plots and the observations

Graphs are plotted for for all 5 trials separately as following:









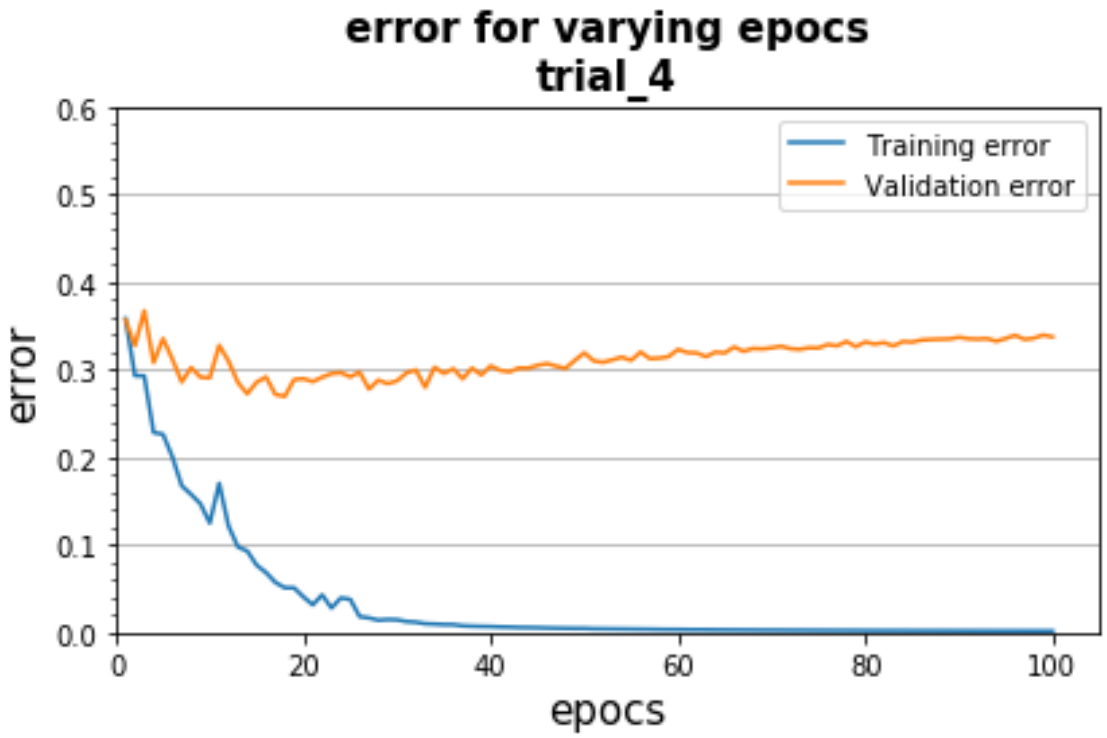


Figure 1: epochs v/s error for different trials

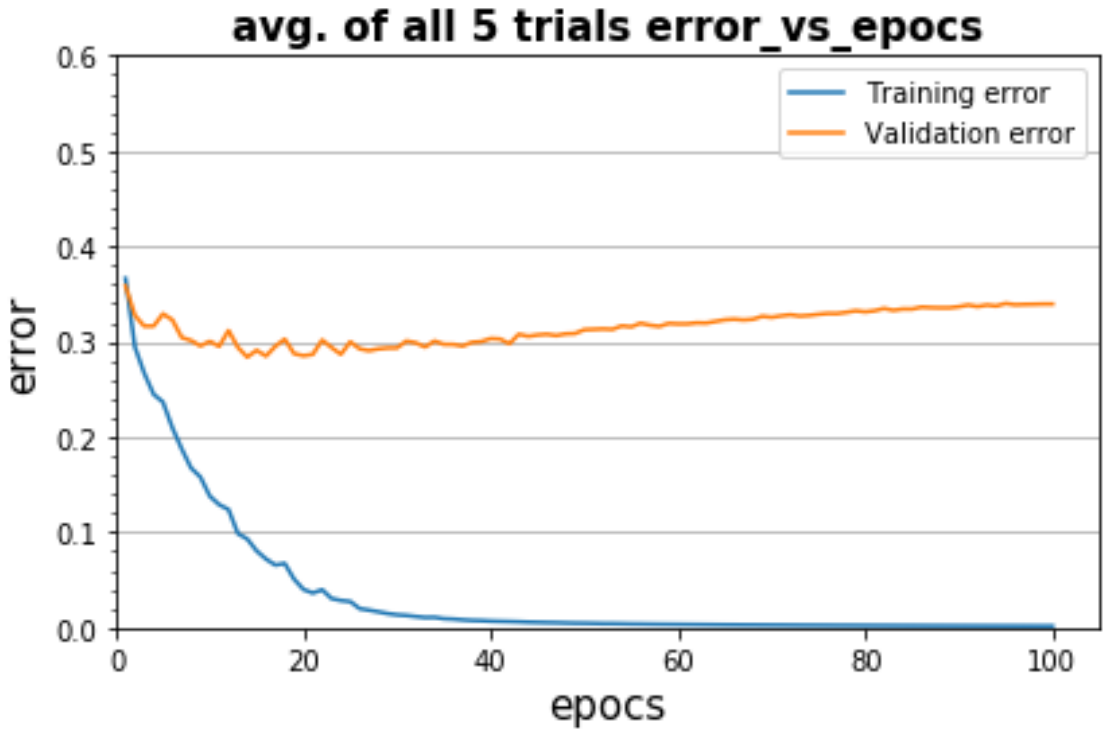


Figure 2: Epochs v/s Avg.error of all trials

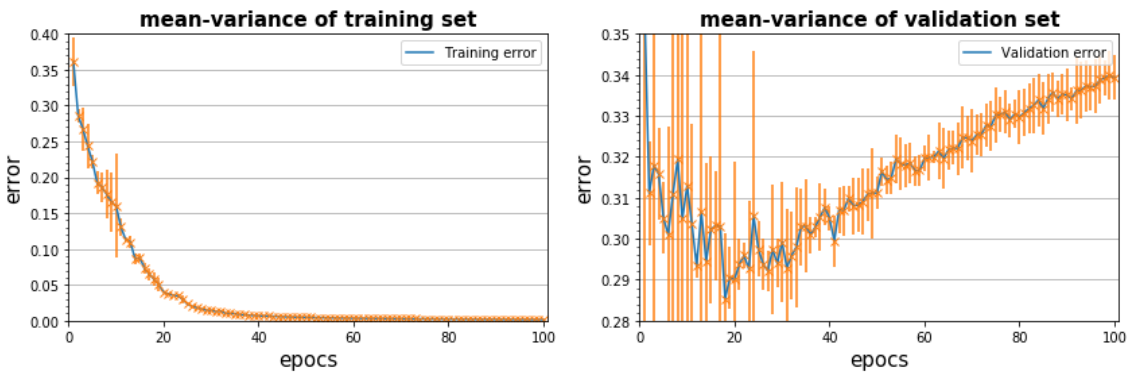


Figure 3: Epochs v/s Avg.error of all trials

Fig.3 shows the mean and variance of error of training and validation dataset. The vertical bar shows the variance at that particular epoch and the actual variance is 100 times small(did 100x for showing on plot) to the depicted in plots.

Explanation:

In all five trials, in fig.1, as more training epochs are made, the error on the training set decreasing,but the error on the validation set starts to increase beyond a certain point. In the starting all weights are close to zero but as the training continues, wights starts getting away from the 0 and thus increasing the complexity of model and leading to the poor generalization. This problem is known as *Overtraining*. And the solution to this problem is that learning should be *stopped early*, in our case early-stopping would be around epoch 30.