Assignment 3
Deep Learning

The whole training data(60000) is divided into a training(50000) and validation set(10,000). The learning rate is kept at 0.001 as it gave a stable training and best validation accuracy. The batch size is chosen as 100 for training. The number of hidden units in 3 hidden layers are taken as 400 as it is almost half the number of input features and found to give best results. Relu activation is used for all the layers. The model is trained for 100 epochs.

The weights corressponding to best validation is saved in the weights folder and further used for testing and training the logistic regression based predictor which takes features from different layers.

Once the modeled is trained testing is done with the testing dataset and an accuracy of 0.8543 is obtained.

Further the features from layers 1, layer 2 and layer 3 are extracted and fitted using a logistic regressor. When the trained regression model is tested for three cases the accuracy obtained are:

1) layer 1 : 0.8587
2) layer 2 : 0.859
3) layer 3 : 0.8593

It can be seen that features extracted from the first layer gave good enough result for the logistic regression based classifier. Moreover, all the accuracies are almost same. So we can infer that for this dataset one hidden unit was enough to give the desired results. The extra 2 hidden units didnot give any extra information about the given input data. Hence we are obtaining testing accuracy similar to the earlier asignment where we have used a single layer.