DEEP-LEARNING ASSIGNMENT-2
MANASVI AGGARWAL (MTECH. (RES.))
SR. NO. 16223

Part1: Train a neural network on Fashion MNIST dataset.

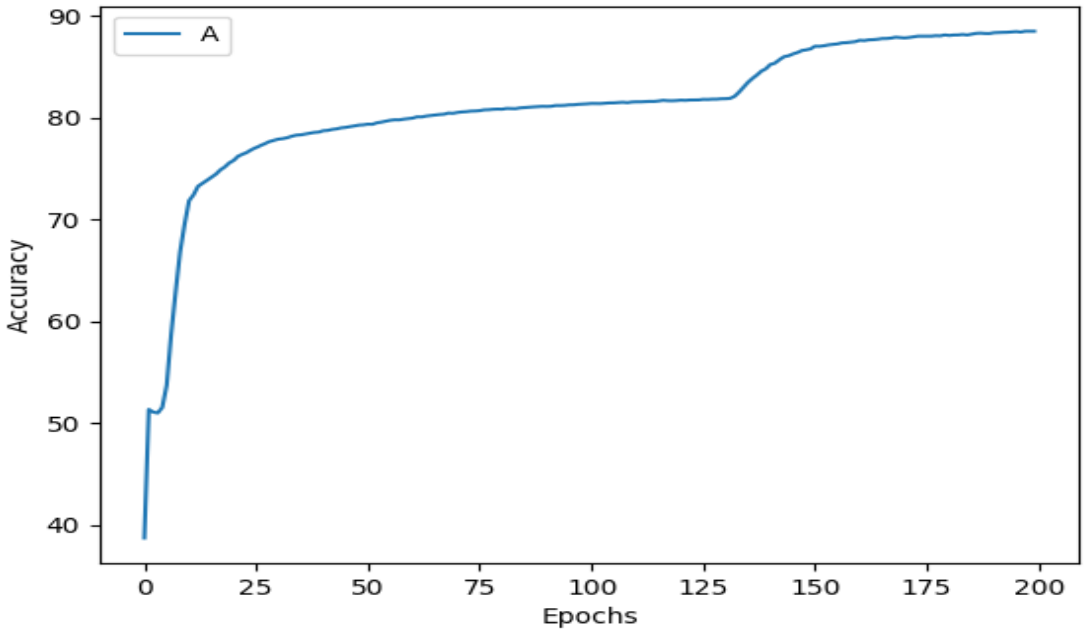
My model specifications are as follows:

#layers 4
#neurons in each layer [512,128,16,10]
activation used in each (relu)
batch size 512
#epochs 100
dropout rate 0.

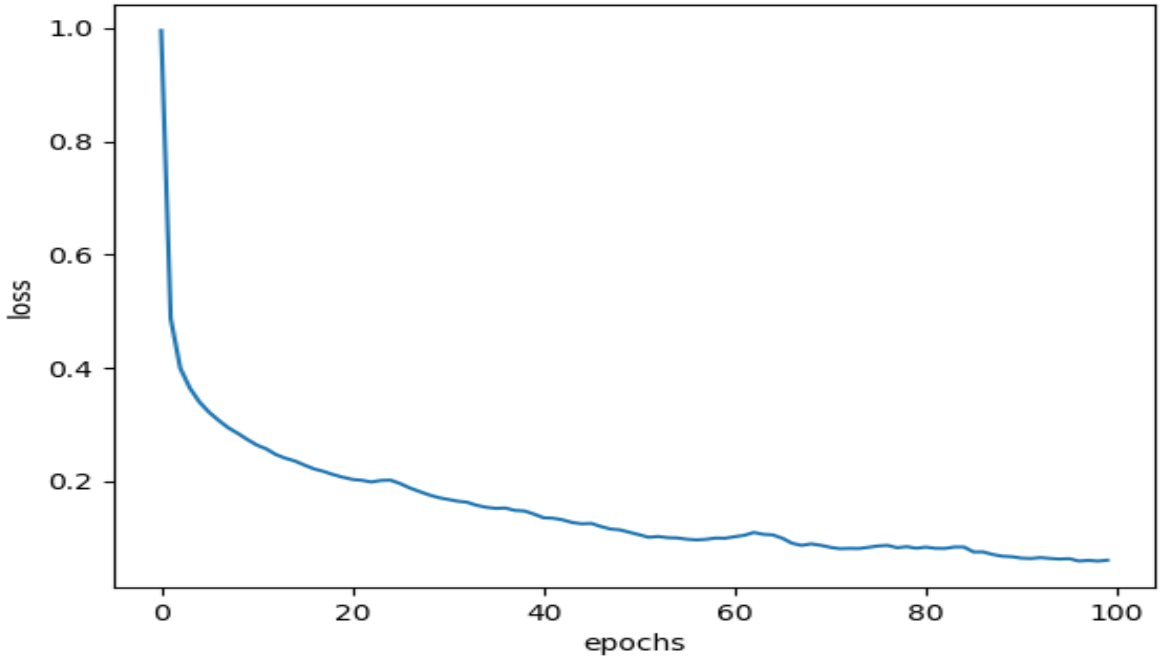
I came to these specifications based on the validation loss and accuracy. And the validation loss was minimum for this configuration so I saved this model. Validation set contains 10% of the train set. 90% of the train set is used for training. At the end I tested my saved model on the test set. I testes various configurations. For #layers I tested for 2,3,4 and for #neurons I tesetd for 16,64,256. Training accuracy of the saved model is 96% and test accuracy is 89.35%. I choose current architecture as validation loss was minimum on this architechture. Also, more complex model requires more data therefore, I didn't choose more than 4 layers. Also, Learning rate is set to 0.001 as higher learning rate will lead to oscillations during training.

Given are some plots for some configurations which I tested: (All plots on y axis denotes the train accuracy/train loss) :

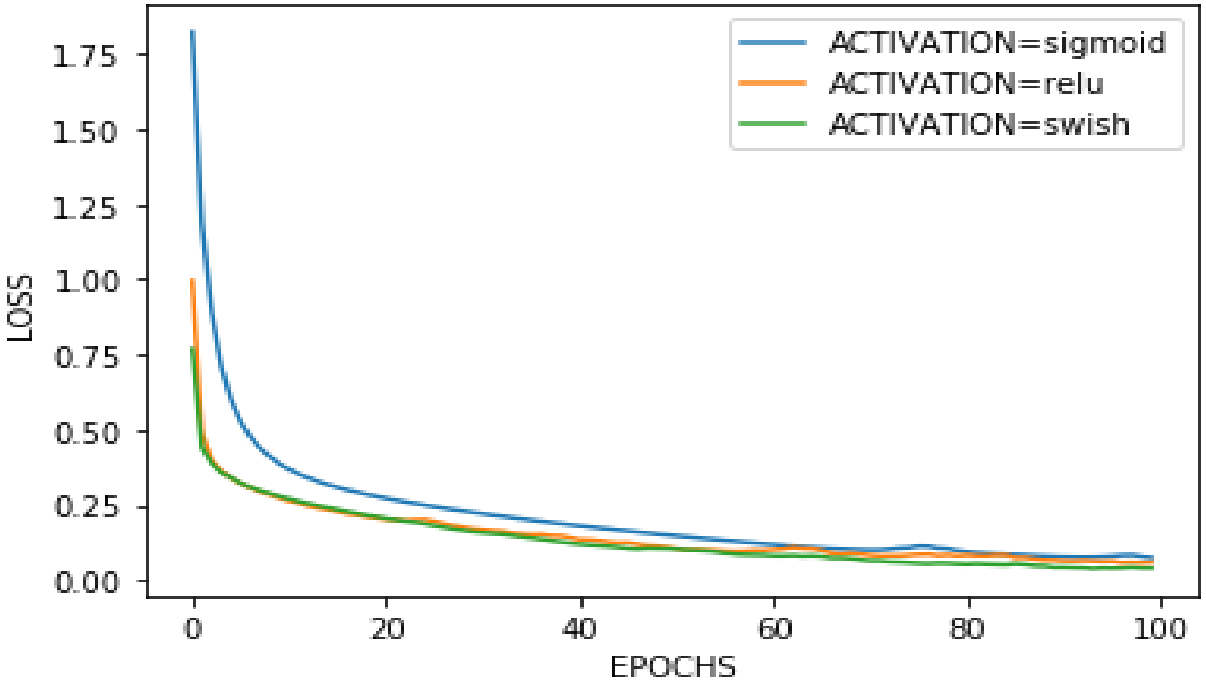
Plot 1: Epochs vs Accuracy



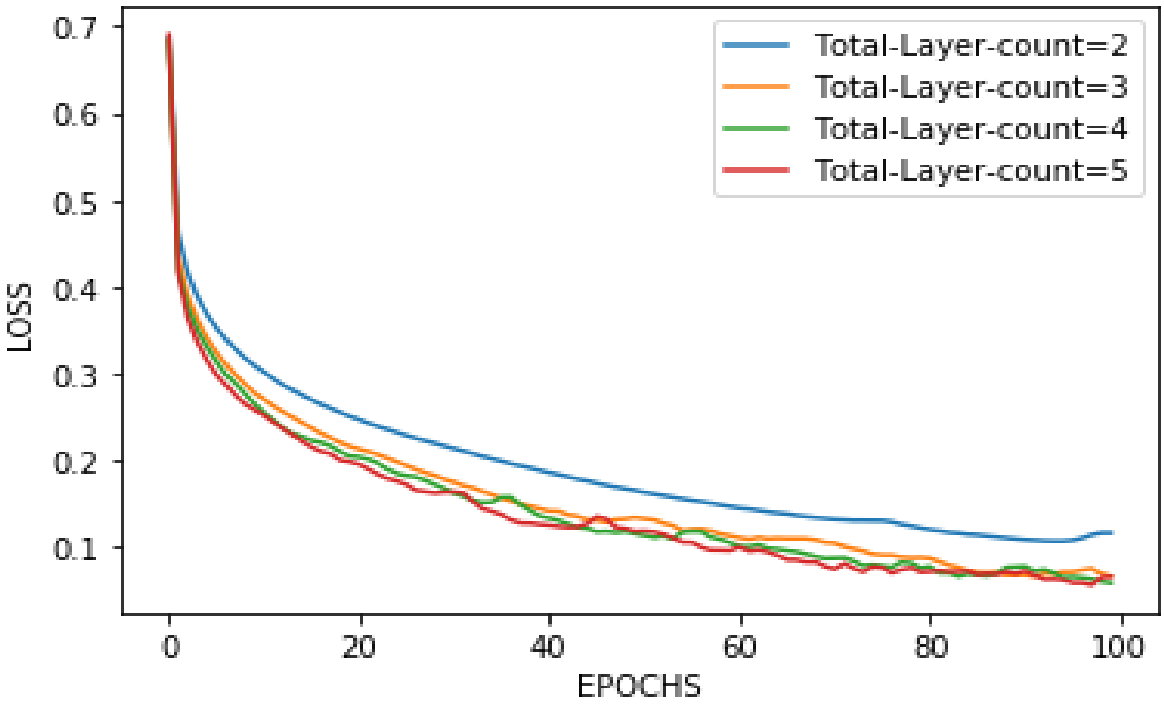
Plot 2: EPOCHS VS LOSS



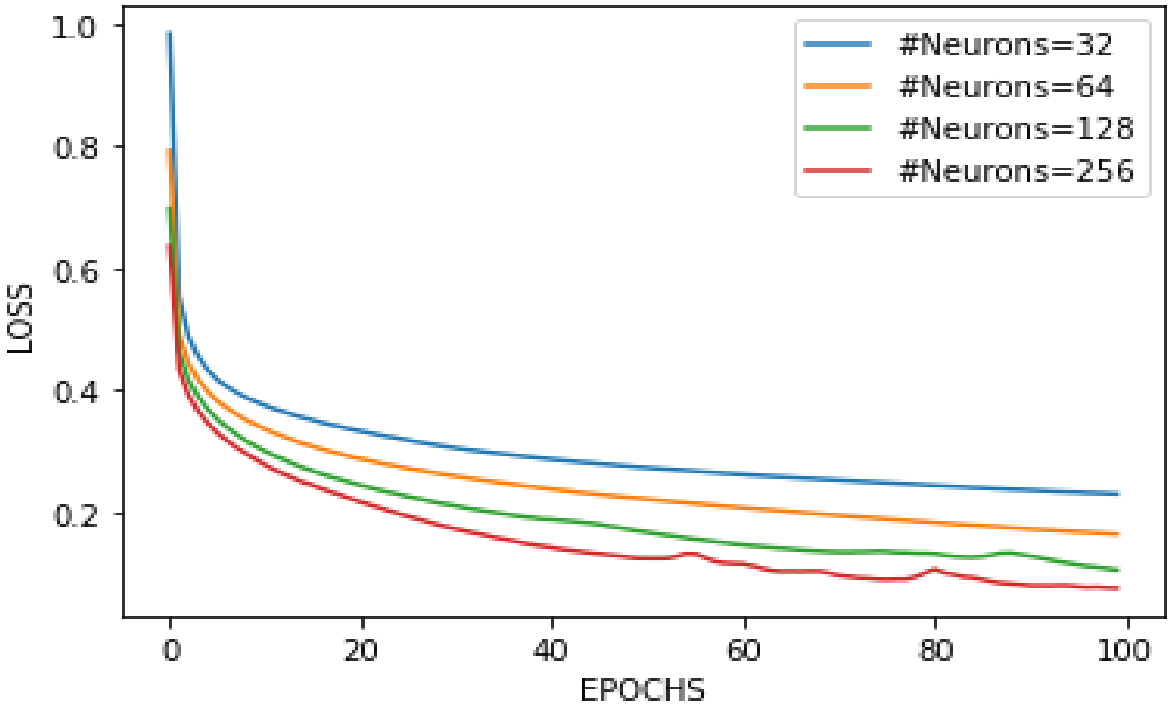
Plot 3: VARYING ACTIVATION FUNCTIONS



Plot 4: DIFFERENT NUMBER OF LAYERS



Plot 5: VARYING NUMBER OF NEURONS



PART 2: Train a CNN on Fashion MNIST dataset.

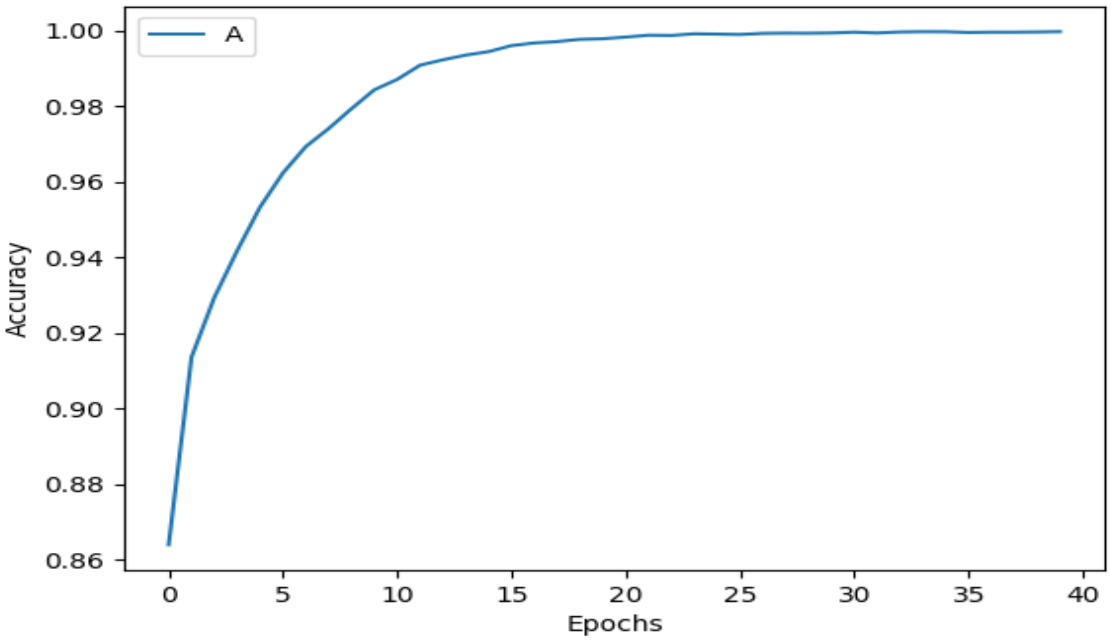
My model specifications are as follows:

#CNN layers 2
Filter size 5,2
activation used in each layer relu
batch size 128
#epochs 25
#filters in each CNN layer 32,64
#FCC layers 3
#neurons in each FCC layer 512,10

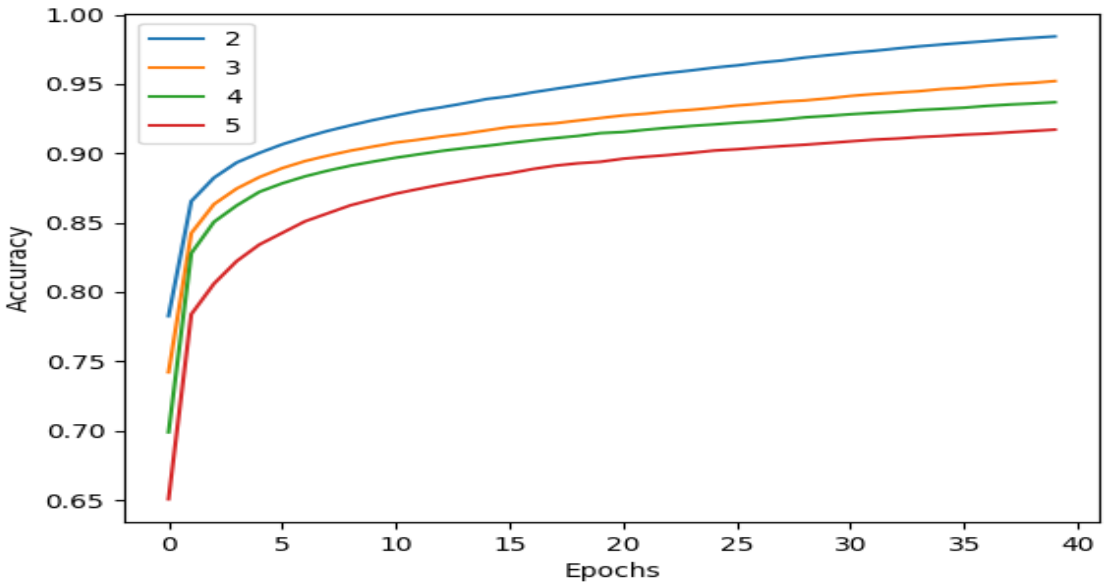
I came to these specifications based on the validation loss and accuracy. I keep doubling the #channels in each CN layer as input to a CNN layer is decreasing after every CNN layer. And the validation loss was minimum for this configuration so I saved this model. Validation set contains 10% of the train set. 90% of the train set is used for training. At the end I tested my saved model on the test set. I testes various configurations. For #CNN/FC layers I tested for 2,3,4 and for #neurons I tested for 16,32,64,128,256 for each FC layer. Training accuracy of the saved model is 95% and test accuracy is 91.55%. More complex model requires more data therefore, I didn't choose more than 2 CNN layers and 2 FCL. Also, model is converging and hence these configurations are used. The early stopping is used while training CNN and when validation loss starts increasing for some epochs I stop the training. Also, Learning rate is set to 0.001 as higher learning rate will lead to oscillations during training.

Given are some plots for some configurations which I tested: (All plots on y axis denotes the train accuracy)

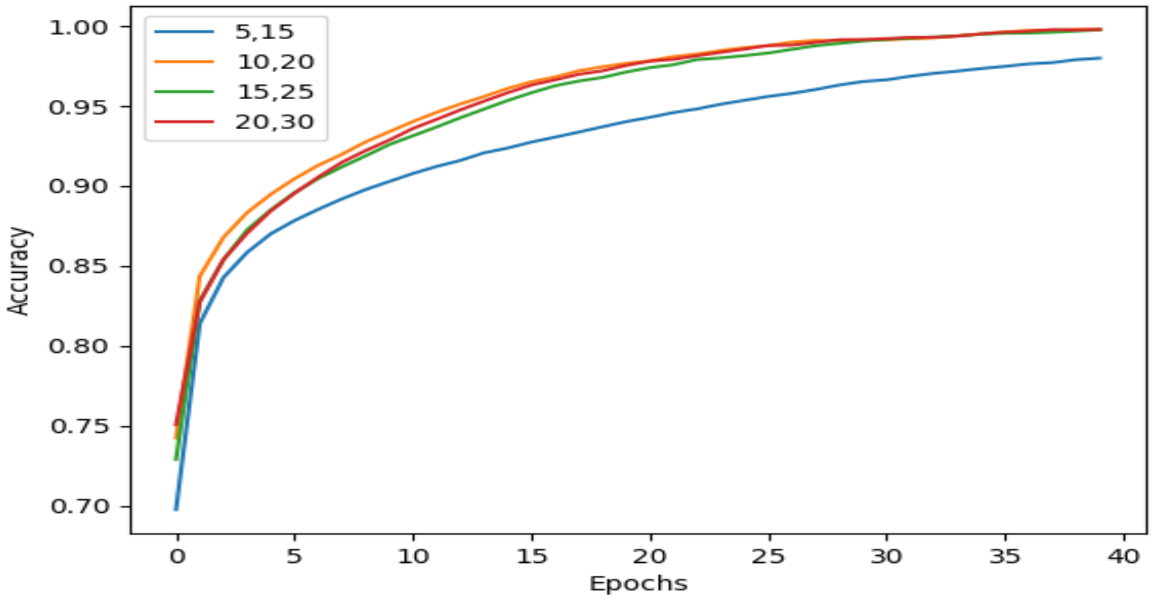
Plot 1: EPOCHS VS ACCURACY



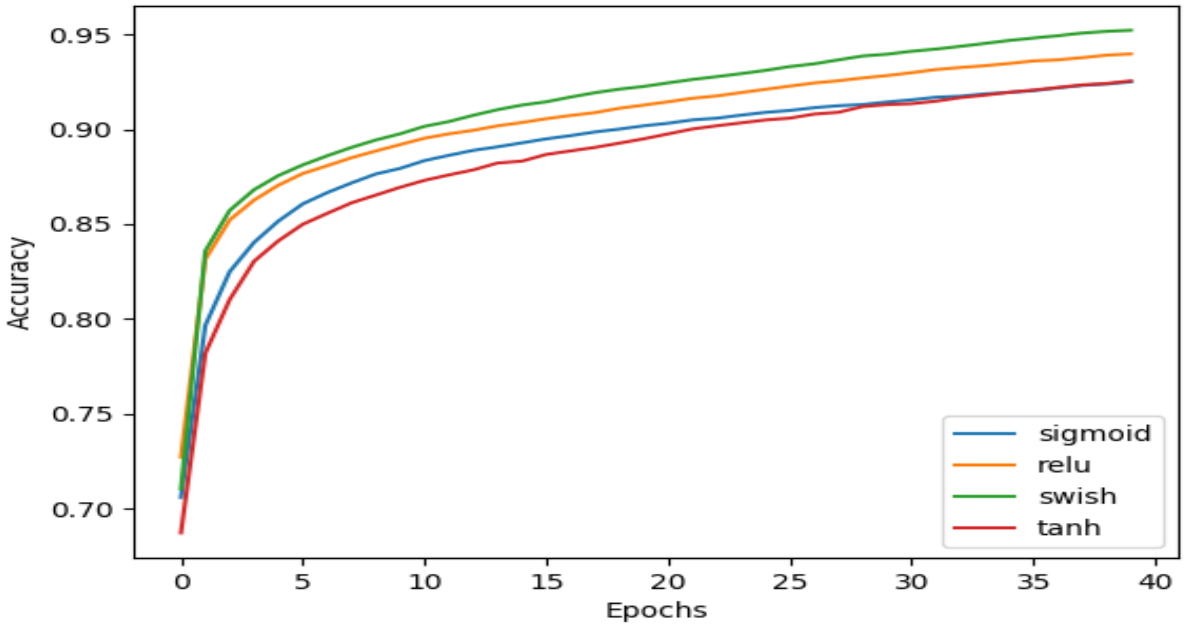
Plot 2: DIFFERENT NUMBER OF CNN LAYERS



Plot 3: VARYING NUMBER OF NEURONS



Plot 3: VARYING ACTIVATION FUNCTIONS



Plot 4: TESTING DIFFERENT INITIALIZATION