COMPARISON OF THE PERFORMANCE OF TWO MULTI-LAYER PERCEPTRONS IN MNIST DATASET

Submitted by,

Sanjay Sunil
S5 CSE
Roll no: 56
TVE19CS058

SOURCE CODE

The Source code used to train the model can be found [here](https://github.com/sanjaysunil34/ML_assignment) (https://github.com/sanjaysunil34/ML_assignment)

I trained the MNIST dataset in two models. The model summary of the two models is as shown in the images below

Model 1

```
Model: "sequential"

 Layer (type)                 Output Shape              Param #
=================================================================
 flatten (Flatten)            (None, 784)               0

 dense (Dense)                (None, 80)                62800

 dense_1 (Dense)              (None, 10)                810

=================================================================
Total params: 63,610
Trainable params: 63,610
Non-trainable params: 0
```

Model 2

```
Model: "sequential_1"

 Layer (type)                 Output Shape              Param #
=================================================================
 flatten_1 (Flatten)          (None, 784)               0

 dense_2 (Dense)              (None, 50)                39250

 dense_3 (Dense)              (None, 128)               6528

 dense_4 (Dense)              (None, 10)                1290

=================================================================
Total params: 47,068
Trainable params: 47,068
Non-trainable params: 0
```

A comparison on both the model training is tabulated below:

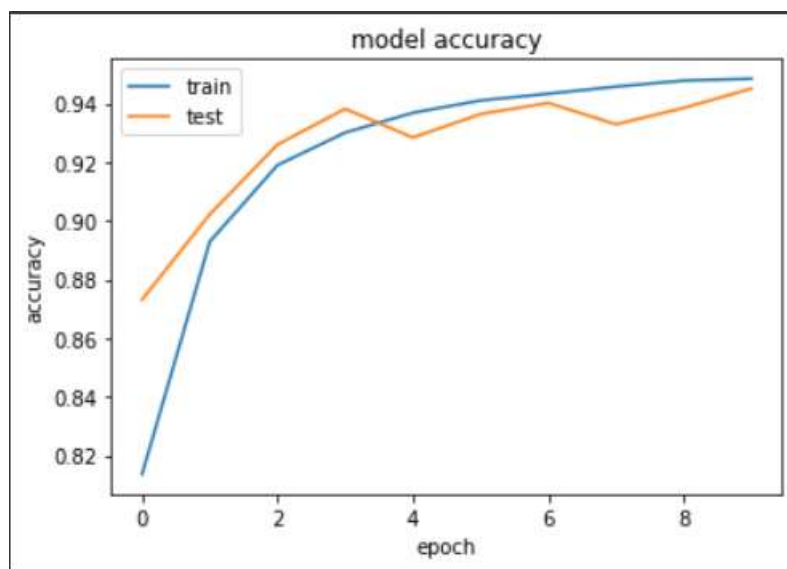| Criteria	Model 1	Model 2
# of Hidden layers	1	2
# of parameters	63,610	47,068
Total # of neurons	90	188

Both the models were trained with same Loss function (cross_entropy), same optimization function(Adam) and the same number of epochs(ten).
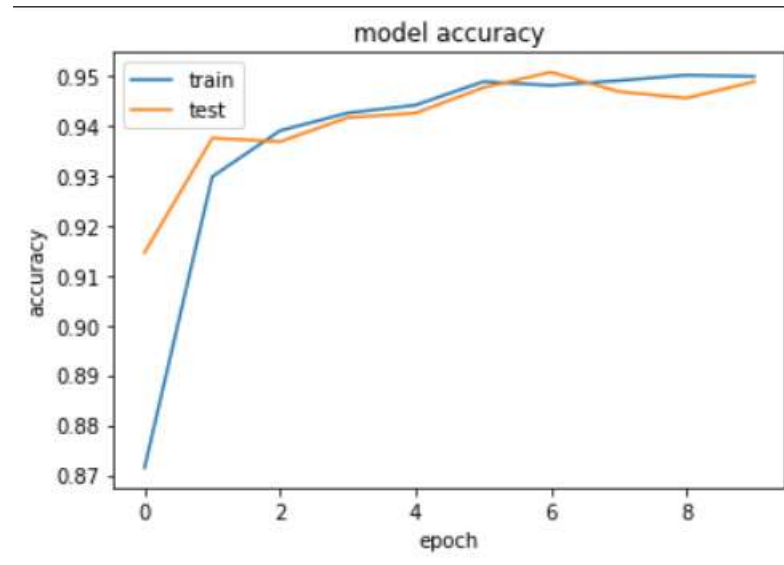
Validation Accuracy of Model 1 = 94.510%
Validation Accuracy of Model 2 =94.870%

Accuracy vs Epoch Plots

<u>Model 1</u>

Model 2

Inference

The 2nd Model is more accurate than the 1st model.

Both models are trained through 10 epochs.

The 1st model has more parameters than the 2nd.

But the 2nd model has more hidden layers and so it is able to give better predictions with better accuracy.