Three of the multiple models trained are present in this folder. The output layer for all three contains 10 neurons, with the activation function as softmax. The loss function used is cross-entropy.
model_1.ipynb

- No. of hidden layers = 1
- Layer 1
  - No. of neurons = 784
  - Activation function - Relu
- Optimizer - SGD, momemtum = 0.9
- Batch size = 100
- Epochs = 10
- Training accuracy = 93.24%
- Testing accuracy = 93.64%

model_2.ipynb

- No. of hidden layers = 2
- Layer 1
  - No. of neurons = 20
  - Activation function - Relu
- Layer 2
  - No. of neurons = 100
  - Activation function - Relu
- Optimizer - SGD, momemtum = 0.9
- Batch size = 20
- Epochs = 10
- Training accuracy = 95.20%
- Testing accuracy = 95.02%

model_3.ipynb

- No. of hidden layers = 2
- Layer 1
  - No. of neurons = 784
  - Activation function - Relu
- Layer 2
  - No. of neurons = 200
  - Activation function - tanh
- Optimizer - adagrad
- Batch size = 100
- Epochs = 10
- Training accuracy = 99.94%
- Testing accuracy = 98.47%

Hence, we can see that model_3.ipynb is the best among the ones trained.