Nyan Tun
UID: 204607621

HW 9

1. The network architecture used is a fully connected neural network. In the input layer, there are 784 neurons. In the first hidden layer, there are 512 neurons. In the second hidden layer, there are 128 neurons. In the output layer, there are 10 neurons.

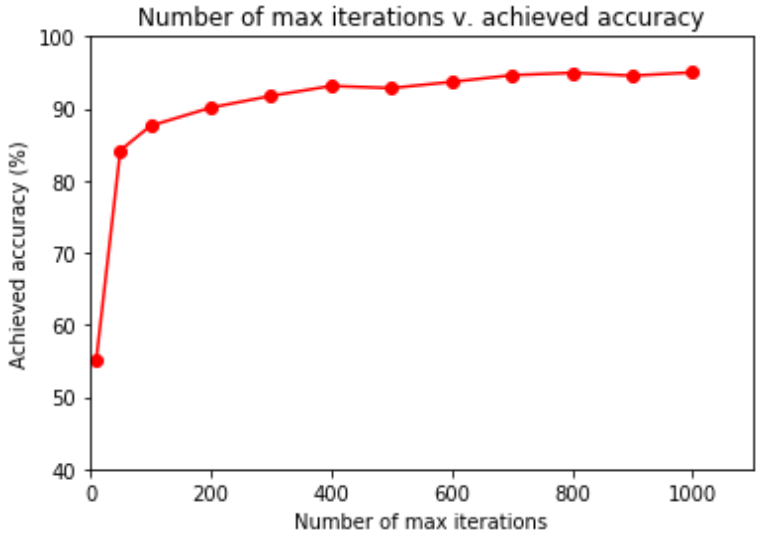
2. The activation function used is rectifier linear unit, and the mathematical function is: $f(x) = x^+ = \max(0, x)$.

3. The accuracy of the network on the 10000 test images is 95 %.

4. The accuracy of the network on the 10000 test images became 89 % after reducing the number of neurons in each hidden layer by a factor of 16.

5. The test accuracy dropped because the model no longer fitted the data as closely as the number of neurons was reduced by a factor of 16 (fewer number of features).

6. Trend of achieved accuracy vs number of iterations:
   It can be observed that as the number of iterations increases, the test accuracy increases significantly, from 10 to 50 max iterations. It continues to increase a lot less past around 400 max iterations, with an increasingly higher chance of overfitting the higher the number of max iterations.



Data points:

```
MAX_ITERS_ARR =
[10, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000]
TEST_ACCURACY_RESULTS =
[55.05, 84.2, 87.6, 90.13, 91.76, 93.14, 92.84, 93.7, 94.63, 94.96, 94.56,
95.02]
```