

Xiaojia Quan
CS161
HW9

1. The network architecture we used is Neural Network. The input layer have $1*28*28$ neurons. The first hidden layer have 512 neurons and the second hidden layer have 128 neurons. The output layer have 10 neurons.
2. The activation function used is called Relu. Formula: Relu: $A(x) = \max(0, x)$
3. Test accuracy: 96%
4. Test accuracy after reduce: 92%
5. The test accuracy reduced as the number of neurons decreases. This is because too less number of neurons would result in underfitting.
6. 10 different max iters values between 10 to 1000

MAX_ITERS	Accuracy
100	88%
200	89%
300	91%
400	92%
500	92%
600	93%
700	94%
800	94%
900	95%
999	96%

The accuracy is proportional to the number of iterations. This is because the more iteration we have, it will train a more adaptive model while it is not overfitting.