

1. These are the layers:

Inputs: 784 inputs

Layer 1: 512 nodes

Layer 2: 128 nodes

Outputs: 10 outputs

2. The activation function is the ReLU (Rectified Linear Units) function.

3. The test accuracy is 96%.

4. The test accuracy is 91%.

5. Yes there is a change. The test accuracy decreased. This is expected because of the fact that we decrease the number of hidden neurons so we are limiting the connections that are formed, the pathways that can be taken, and the overall processing power of the network. Thus, this is analogous to decreasing the depth of the network, and as discussed in class, this means the network will approximate the data with less accuracy.

6.

MAX_ITEERS	Test Accuracy
10	60
25	76
50	82
75	84
100	85
101	88
250	91
500	92
750	94
1000	95

Clearly, as the number of iterations increases, the test accuracy also increases. The reason for this is because with more iterations, the loss is decreased as the network fits the data better, so the accuracy goes up. With each iteration, it self-improves its fit on the data.