Better accuracy was observed With 2 hidden layers: 784, 50, 40, 10 than 1 hidden layer 784 100 10.

Best accuracy was observed for alpha= 0.006 (keeping no iterations = 1000)
Decreasing the alpha to 0.001 required no of iterations to be increased to get same accuracy.
Alpha greater than 0.006 (example 0.01) gave less accuracy.

Optimal no of neuron was found to be 40-80 in hidden layer.