

CS547

HW1

Tianqi Wu

The accuracy on the test set is **0.9745** after 5 epochs of training. The model has one hidden layer and 100 hidden units. The nonlinear function is ReLU. The learning rate is 0.01 without decreasing. The input size is 28×28 and output size is 10. I used Xavier initialization and stochastic gradient descent as optimization method. No special tricks are used.