CS547
HW2
Tianqi Wu

The accuracy on the test set is **0.967** after 2 epochs of training. The model has one convolutional layer and one fully connected layer. There are 4 channels and the filter dimension is 3*3. There are 4*26*26 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 except vectorizing for speeding up the computation.