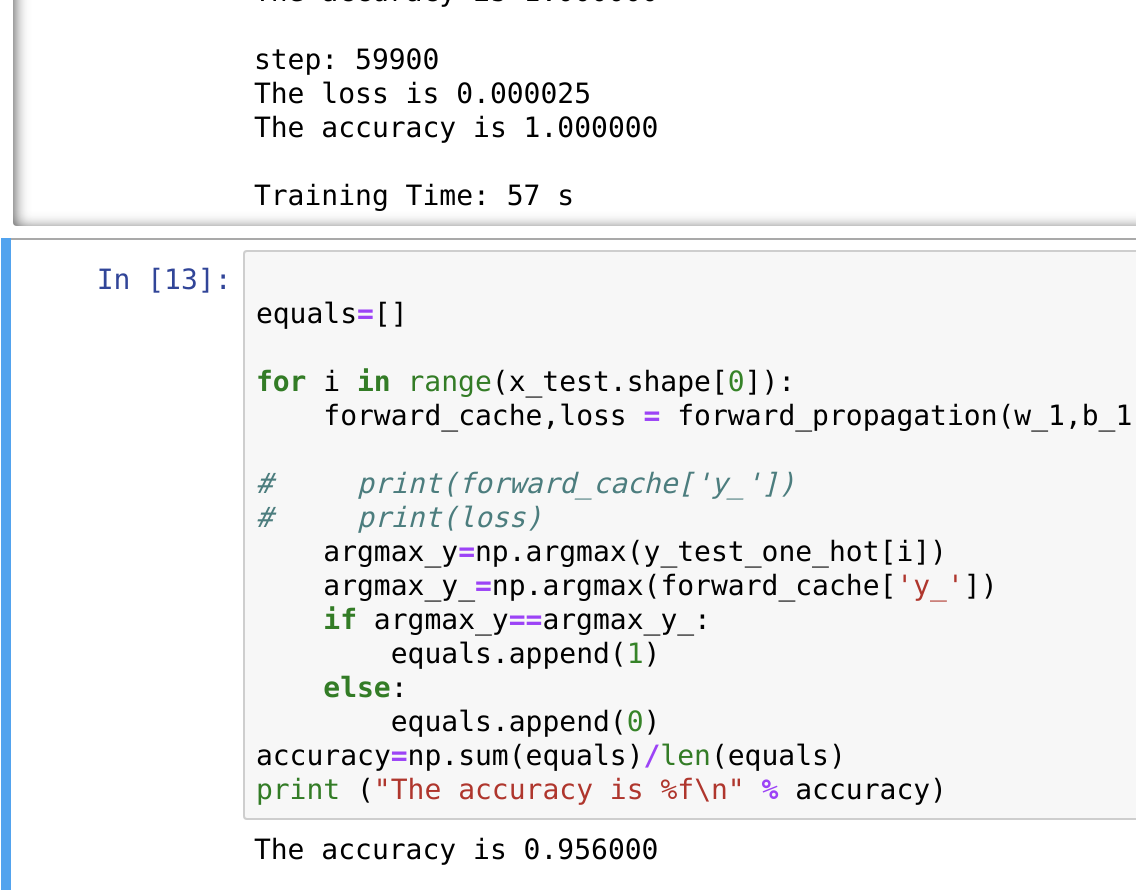
Xin Yang

xy213

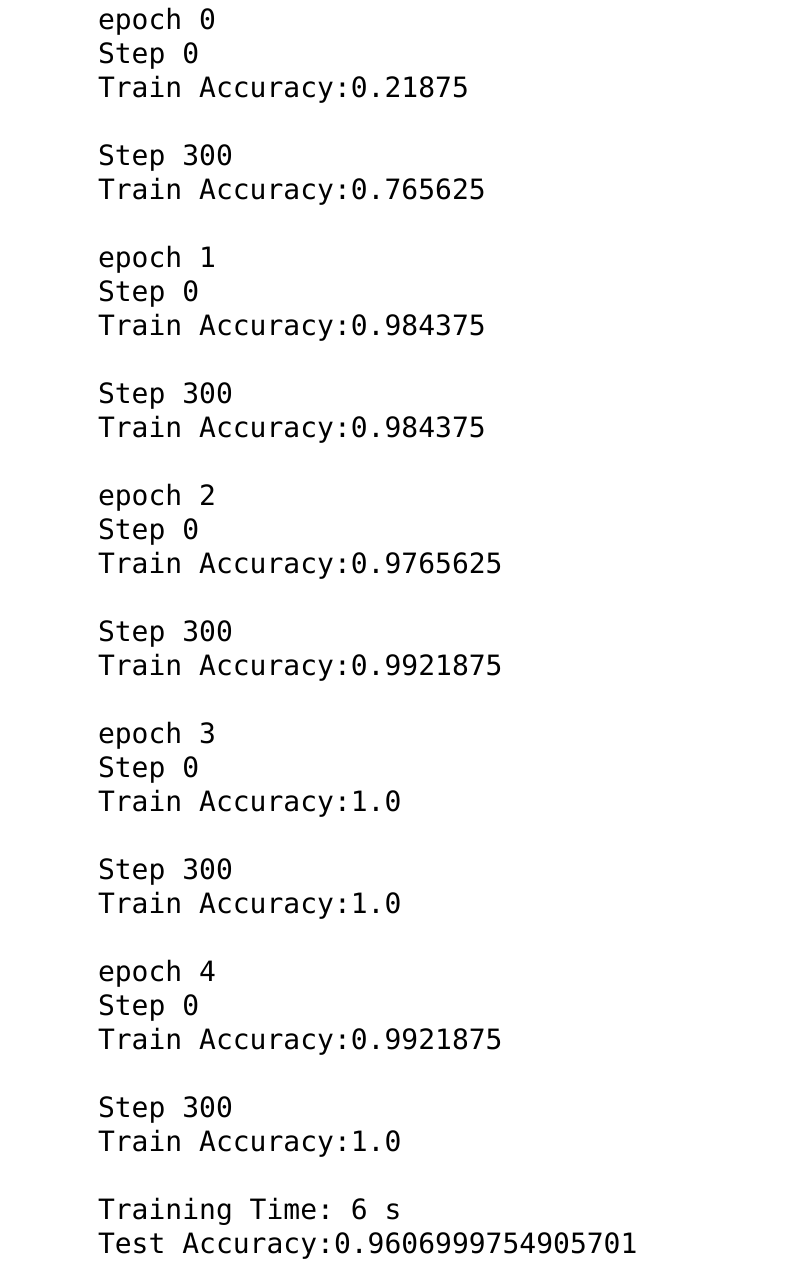
HW3

Q1:

For hand written part, I select the mean square error as the loss function. This program runs on a stochastic gradient descent optimizer. The learning rate is 0.01 and number of epochs is 5. The test set accuracy achieves 95.6%, which takes 57 seconds to complete.



Q2:

For the framework based solution, I utilize the TensorFlow with GPU support. The loss function is the cross entropy. My algorithm runs on a mini-batch stochastic gradient descent optimizer. The learning rate is 0.01, batch size is 128 as suggested. By running 5 epochs, the overall accuracy on the test data set is over 96%, and the training duration takes 6 seconds.