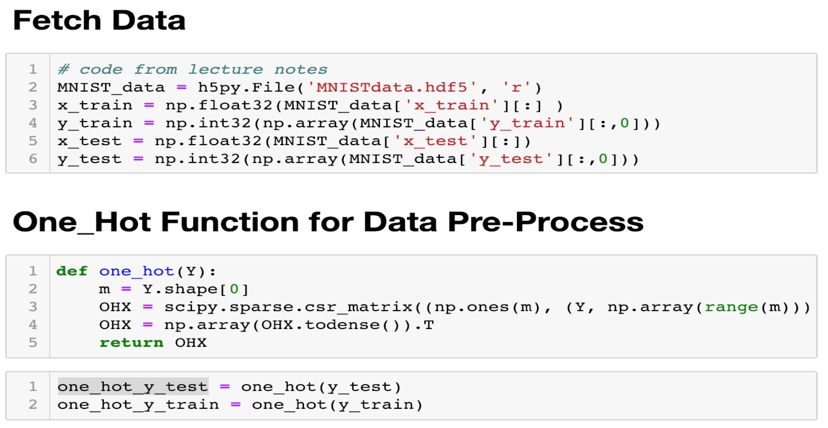
**CS398-Deep Learning**

**Homework 2**

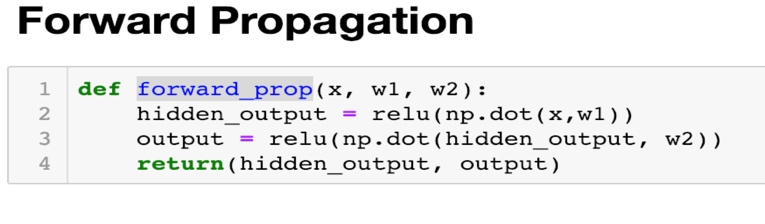
**Lingyi Xu (lingyix2)**

1. **Description of Implementation**

First, I load the MNIST data from the dataset and preprocess the labels to one-hot format.

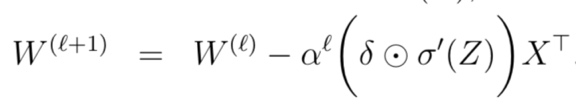
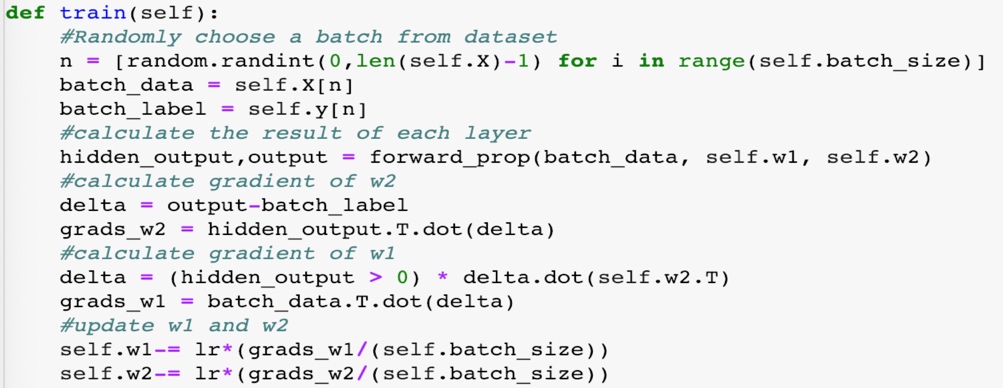
Then I build my mini-batch softmax logistic regression model.

I define the nonlinearities σ(z) (relu) function I will use:

The forward propagation step:

And the gradient I will use for update weight matrix w:

When training, I first initialize w1 and w2 to random value, w1 is the matrix which get x as input and hidden layer as output and w2 is the matrix which get hidden layer as input and y as output. Then I use 14000 iterations to update the W by gradient decent.

For each iteration,we can update the parameter matrix w by

For this particular dataset, I use the following parameters:

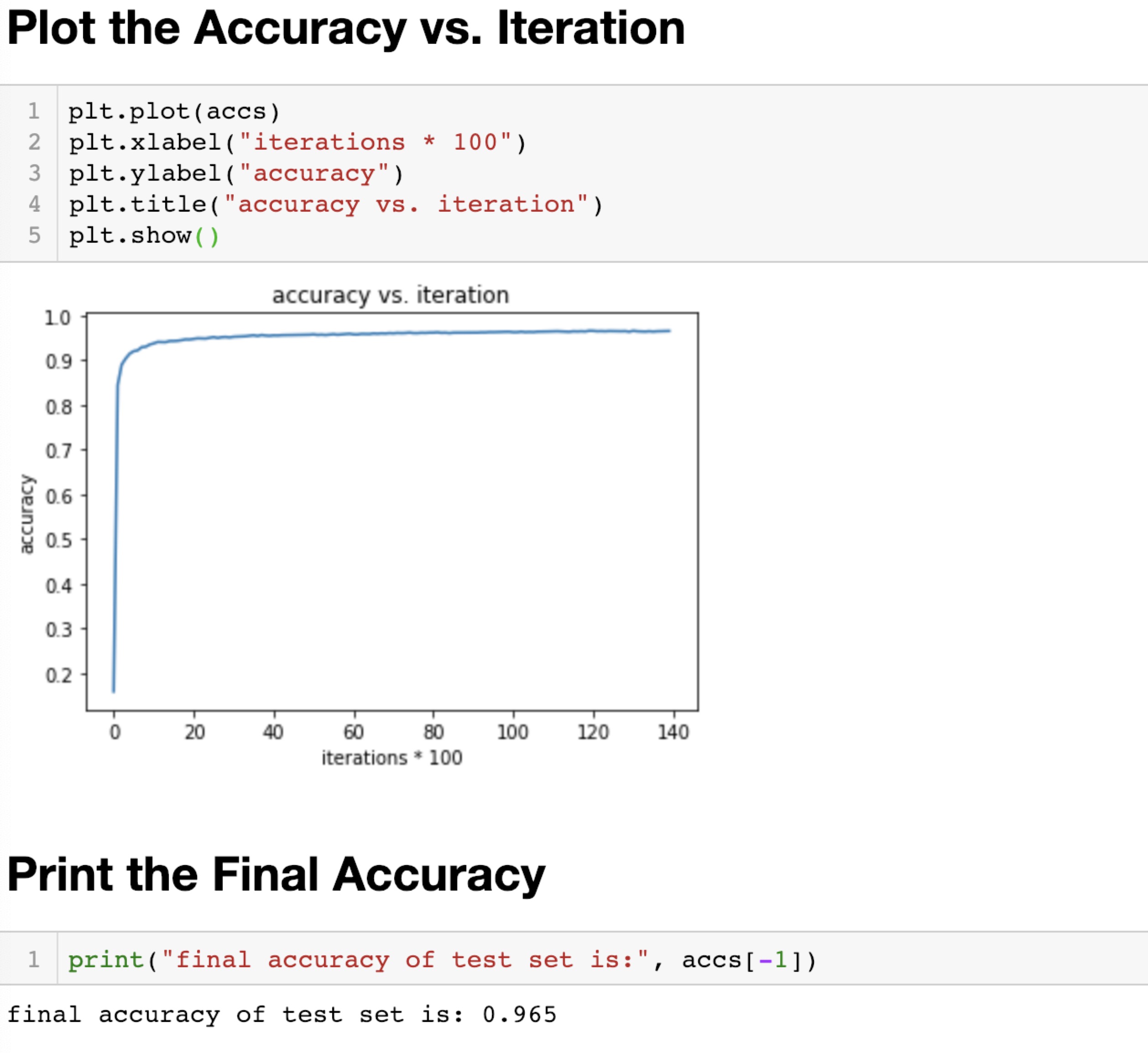
Iteration = 14000

Batch size = 30

Learning rate = 0.25/（0.001\*itr+1）

**2. Final Test Accuracy:**

96.5%

****