作业一

要求：

1. 实现给定要求的全连接神经网络，网络框架要求见图1；
2. 利用1中构建的深度网络实现MNIST数据集识别（数据聚集已经提供，亦可直接下载）；
3. 提交的作业：全部代码+代码说明文档（包括最终实验结果截图，不少于500字）。



图1

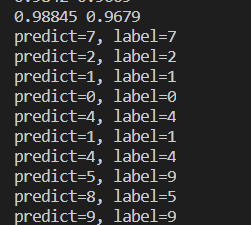
代码说明文档：

该网络包括4个Affine仿射层，3个ReLU激活层，1个Softmax层。仿射层每一层的神经元个数分别为100，100，50，10，前一层的输出数据是后一层的输入数据。采用误差反向传播法计算梯度；前三个Affine仿射层后面接有一个ReLU激活层，最后一个Affine仿射层后面接有一个Softmax层；并采用Cross-Entropy Loss作为损失函数，使用SGD迭代法更新网络参数，采用He初始值进行初始化。

总共训练样本为60000个，测试样本为10000，每次训练的时候从训练样本里选出100个，对网络进行训练。训练的总迭代次数为3000次，每训练300次就测试一下网络的正确率，总共测试10次。

代码中自定义了4个类，分别是Relu类、Affine类、SoftmaxWithLoss类和FourLayerNet类。

最终实验结果截图：



训练数据准确率为98.8%，测试数据准确率为96.8%