# Homework 3

陈春祺 5140379022

## Problem 1

设计前馈神经网络，包括输入层，一层隐藏层，输出层，一共三层。其中，隐藏层节点个数为1024个，训练轮数为5000轮，对tanh和sigmoid函数作为激活函数分别进行了测试，测试结果如下：

可以看出，在只训练5000轮的情况下，batch-size过小，会导致训练不完全，而当随着batch-size增加之后，准确率也随之上升，最后都收敛到较高的准确率。而在两种激活函数的情况下，sigmoid函数在相同的batch-size情况下，最后能得到比tanh更高的准确率。最后选取最好的结果，即以sigmoid函数作为激活函数，batch-size大小为1500，训练轮数5000，准确率0.9723.

代码见feed\_forward.py.

## Problem 2

#### 结构

实现了一个leNet的近似版本，一共有六层，分别为卷积层、池化层、卷积层、池化层、两层全连接层，其中第一层卷积层的大小为5\*5，深度为32，第二层卷积层的大小为5\*5，深度为64，池化层过滤器为2\*2，步长为2，两次全连接层的节点数都为512。

#### 优化

1. 训练时为全连接层添加L2正则化系数，避免过拟合；
2. 训练时使用dropout方法，可以进一步提升可靠性，避免过拟合；
3. 使用滑动平均方法，衰减率为0.99
4. 拟使用学习率衰减函数，初始学习率设为0.8，衰减率为0.99，但在实际训练过程中，似乎会使学习率衰减的过小，导致收敛极慢，故最后采用0.05的固定学习率，也取得了较好的效果。

#### 训练

训练轮数：30000轮

Batch Size：100

#### 验证

在验证集上的准确率为0.993.

#### 比较

下方的表格比较了前馈神经网络和CNN在最好情况下的各项参数的对比

|  |  |  |
| --- | --- | --- |
|  | 前馈神经网络 | CNN |
| Batch Size | 1500 | 100 |
| 训练轮数 | 5000 | 30000 |
| 训练时间 | 329 s | 1952 s |
| 准确率 | 0.9723 | 0.993 |

代码见cnn\_inference.py.