

HW1 report

姓名+学号

刘佳欣-19307130299

代码链接：

<https://github.com/ZnStar/HW1>

由于文件内容并不大，github上已包含全部数据集和模型，可以直接运行

模型链接：

链接：https://pan.baidu.com/s/1dd3aZudBY1eLd2_OF17K1A?pwd=CVHW

提取码：CVHW

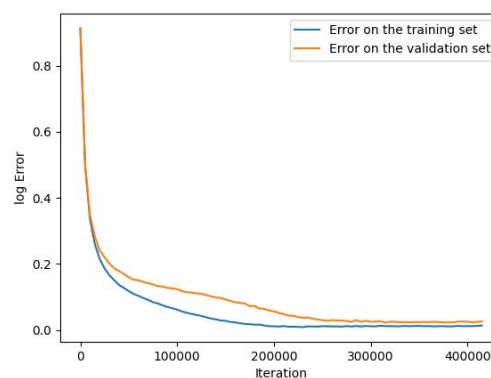
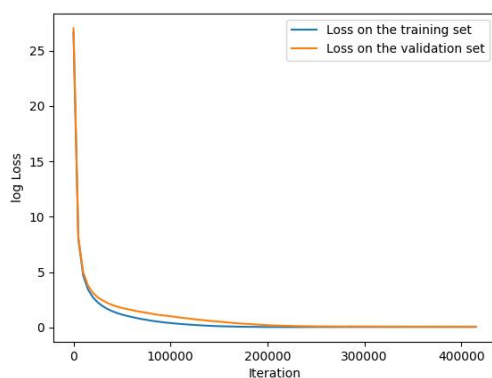


代码文件运行详见ReadMe.md

训练

- 从训练集中抽取10%作为验证集
- 采用Adam算法使用动量作为参数更新方向并自适应调整学习率
- 用tanh作为激活函数
- 小批量梯度下降算法：选择batch size为8
- L2正则化：正则化强度参数
- 提前停止策略：如果累计20次error不下降则中止训练

在实验中得到最优的模型超参数为：300层，学习率为 $1e-4$ ， $\lambda=1e-2$



Error

Loss

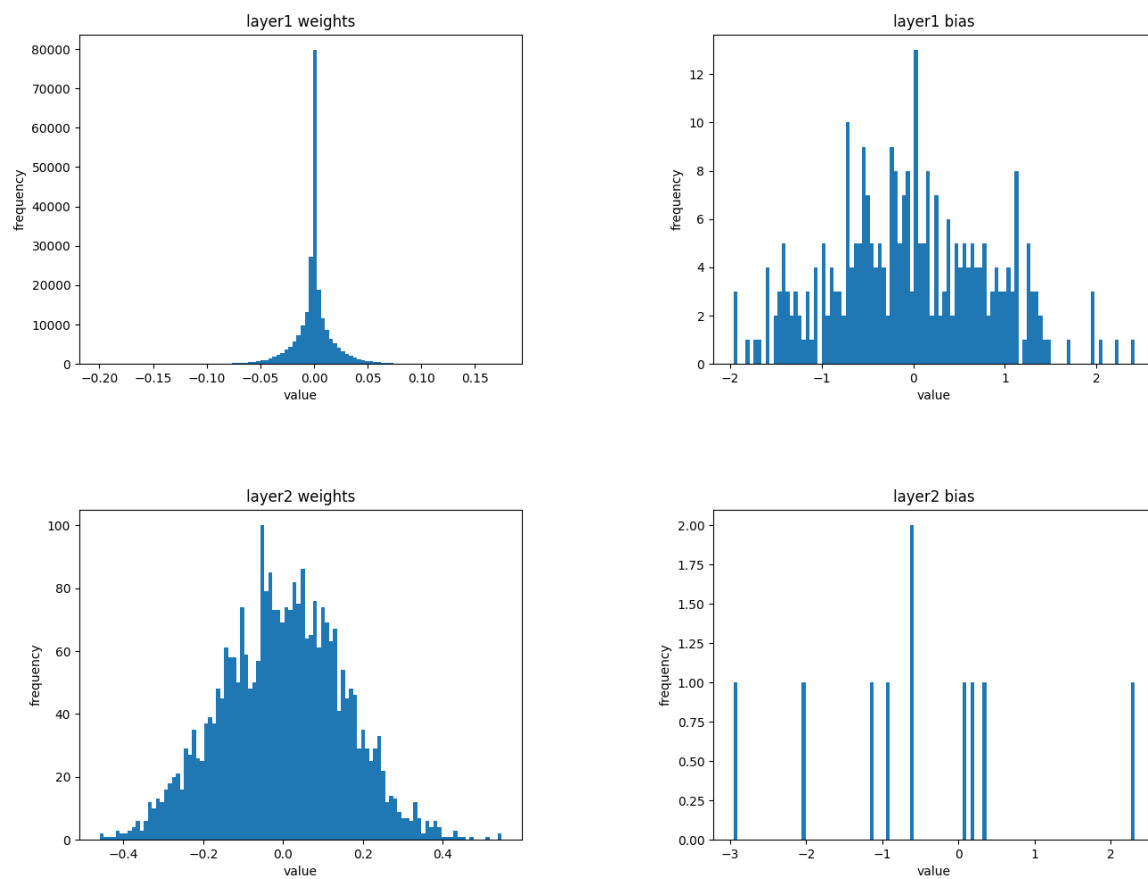
测试

模型在测试集上进行测试，准确率为97.24%

```
Test accuracy with final model = 0.9724
```

网络参数可视化

直方图如下



热力图如下

