

# HW1

姓名：柴劲阳

学号：22210980029

课程代号：DATA620004

课程名：神经网络和深度学习

## 相关链接

github repo 地址：

[https://github.com/brannchai/FUDAN\\_DATA620004.git](https://github.com/brannchai/FUDAN_DATA620004.git)

模型网盘地址：

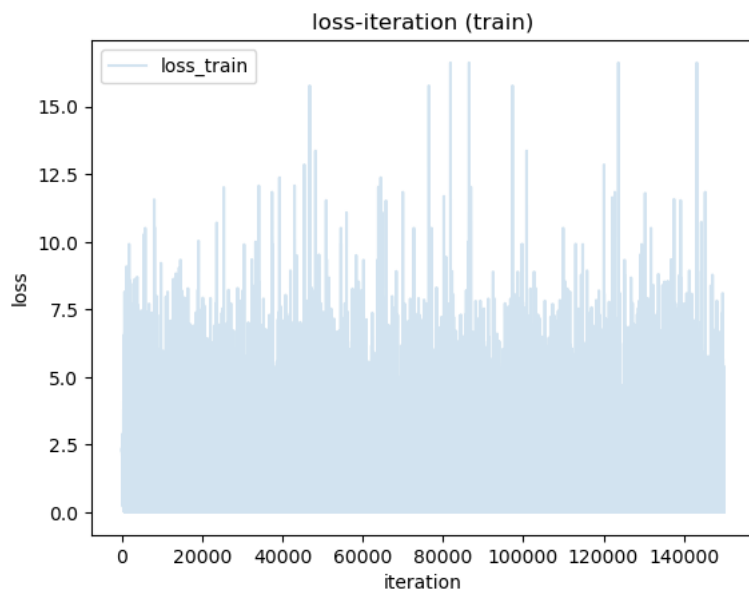
链接：<https://pan.baidu.com/s/1SnBXV0i03HXtCKiUiD321A>

提取码：c4ez

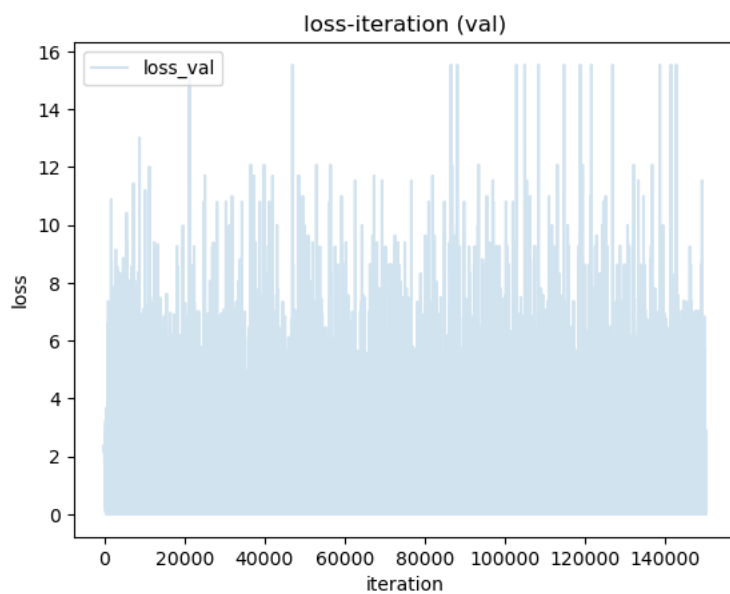
文件说明：1) test model 存储了训练好的模型参数；2) result 存储了模型的结果；3) MNIST 是数据集；4) HW1\_22210980029 是代码文件。

## 相关可视化结果

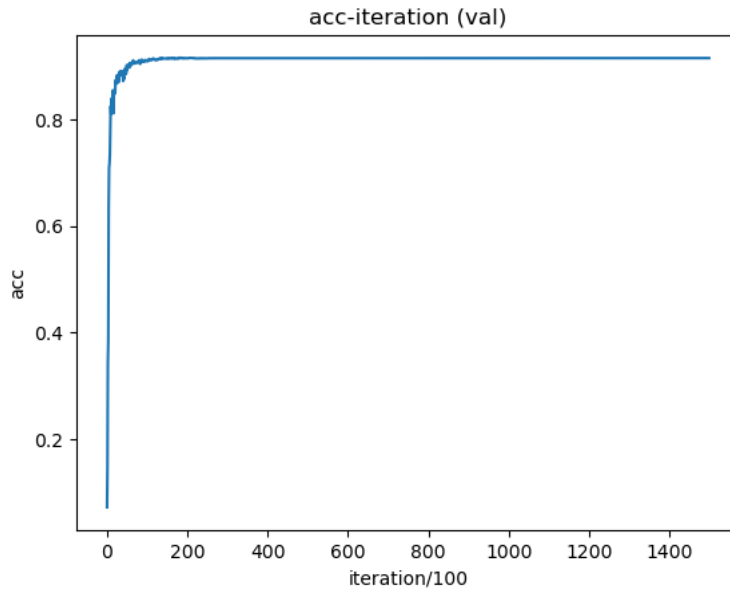
train set 上的 loss 曲线：



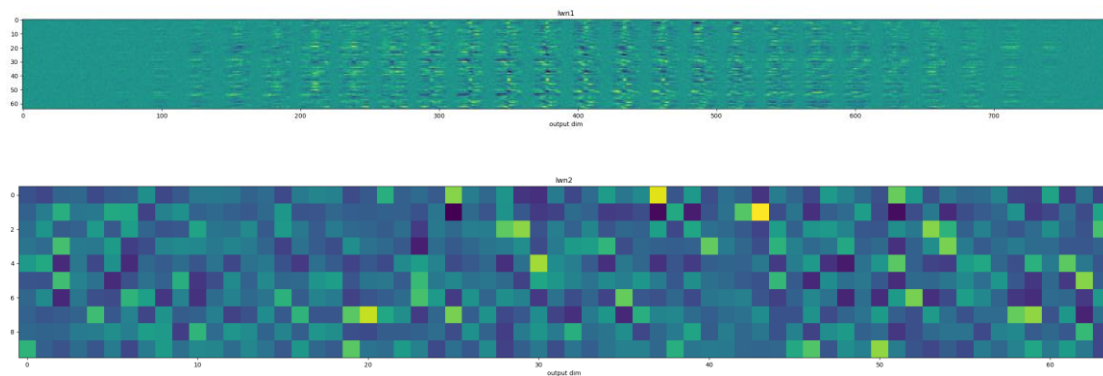
validation set 上的 loss 曲线:



validation set 上的 accuracy 曲线:



模型的参数可视化:



模型预测性能展示:

最终得到的模型在测试集合上的acc为: 0.9135  
最终得到的模型在测试集合上的recall为: 0.9121  
最终得到的模型在测试集合上的precision为: 0.9123  
最终得到的模型在测试集合上的f1为: 0.9120

可以看出模型在测试集上的表现相对不错,但为何 train 和 val 上的损失会不太好看呢? 我个人认为是由于 loss 是基于某一个样本点得到的,可能有较大的方差,即不太能反应模型对整体数据集的拟合情况。我同时尝试了使用透明度较小的散点图进行可视化,可以发现随着 iteration 的增大,loss 是有减小的趋势

的，但也存在着较少部分 iteration 时有较大的 loss，正是这部分 loss 使得 train-set 和 val-set 上的 loss 曲线不那么好看。