

# Homework1

## 1. 复现相关链接及步骤：

- 1) Github 链接: <https://github.com/WTYd2l/DeepLearningHW1.git>
- 2) 百度网盘地址 (模型):  
[https://pan.baidu.com/s/1iiwmVYYBs60HRd\\_gD05z8Q](https://pan.baidu.com/s/1iiwmVYYBs60HRd_gD05z8Q) 提取码: 7XC9
- 3) 运行 main.py 文件 (将注释部分取消注释), 运行得到数据, 运行 draw.py 文件可得到相应的图像。可以自己运行得到模型, 也可以根据百度网盘地址下载模型, 直接进行预测。

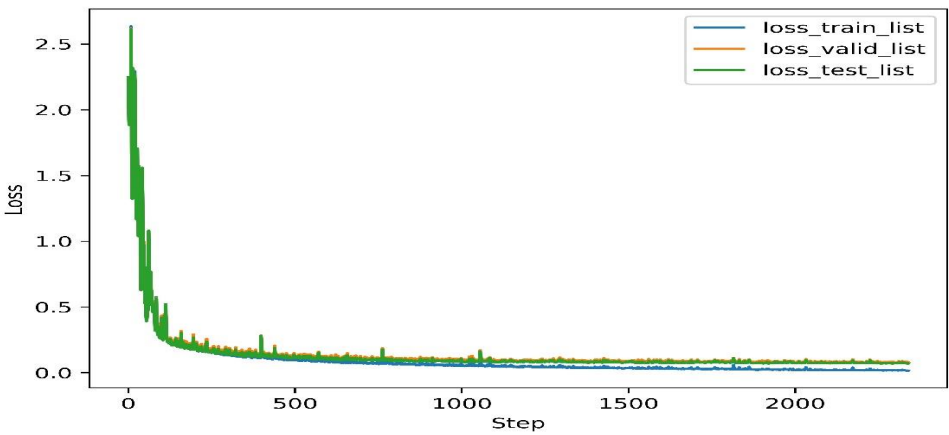
## 2. 结果分析：

在本次实验中, 选用的数据集为 Mnist 数据集, 采用了两层全连接层网络, 激活函数为 Relu, 损失函数为交叉熵。

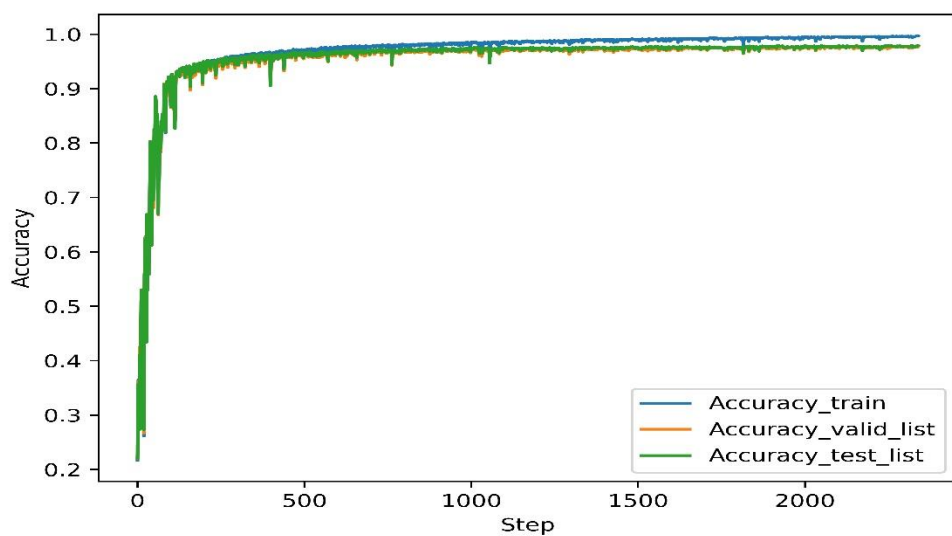
参数查找结果及分析:

| 超参数  | Learning_rate                   | Regular                | Hidden_size   |
|------|---------------------------------|------------------------|---------------|
| 参数空间 | 0.001, 0.005, 0.01, 0.05, 0.075 | 0.005, 0.01, 0.05, 0.1 | 128, 256, 512 |

如上的参数选择主要考虑到 SGD 优化器对学习率比较敏感所以对学习率的选择会更密集一些, 经过训练后得到的参数选择结果如下: 学习率取 0.005, 正则化系数取 0.05, hidden\_size 取 512, 最终得到的准确率为 0.9779。训练曲线和 loss 曲线如下图所示:



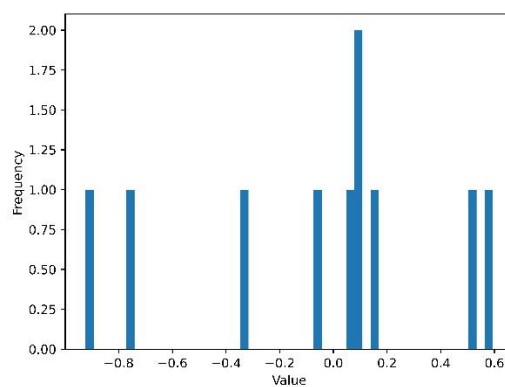
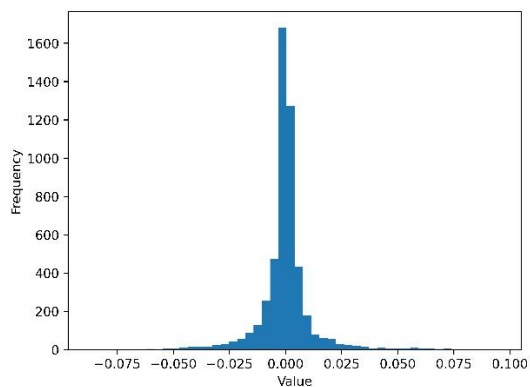
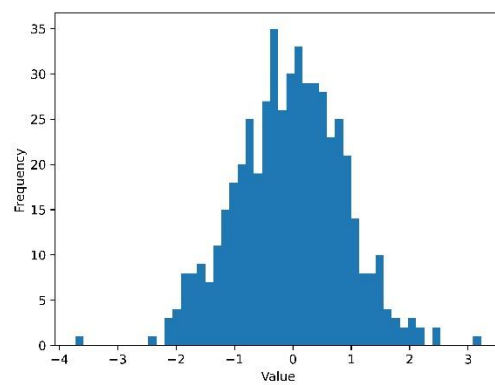
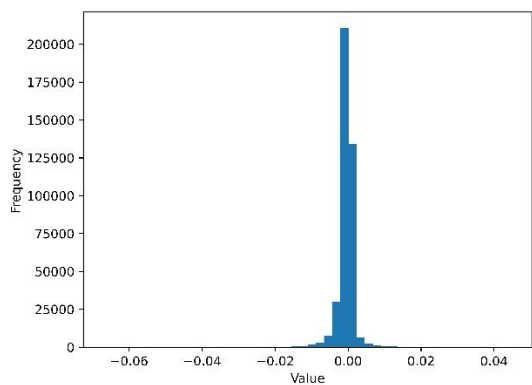
图表 1 Loss\_Curves



图表 2 Accuracy\_Curves

由上述曲线可知，所选用的网络可以较好地拟合数据，并且没有出现严重的过拟合现象，在 900steps 左右可以认为已经收敛。

训练参数可视化结果如下所示，分为两个层的参数分别进行考虑：



上述四张图中，前两张为第一个全连接层的权重系数和偏差，后两张为第二层的对应系数，我们可以看出对于权重系数，两层的分布基本一直，而偏差在两层中出现很明显的差异，第二层的偏差分布明显比较分散。