

研究生《深度学习》课程

实验报告

|  |  |
| --- | --- |
| **实验名称：** | **实验2 前馈神经网络实验** |
| **姓 名：** | **郑楚彬** |
| **学 号：** | **21140129** |
| **上课类型：** | **专业课** |
| **日 期：** | **2021.09.18** |

## 一、实验内容

## 二、实验设计

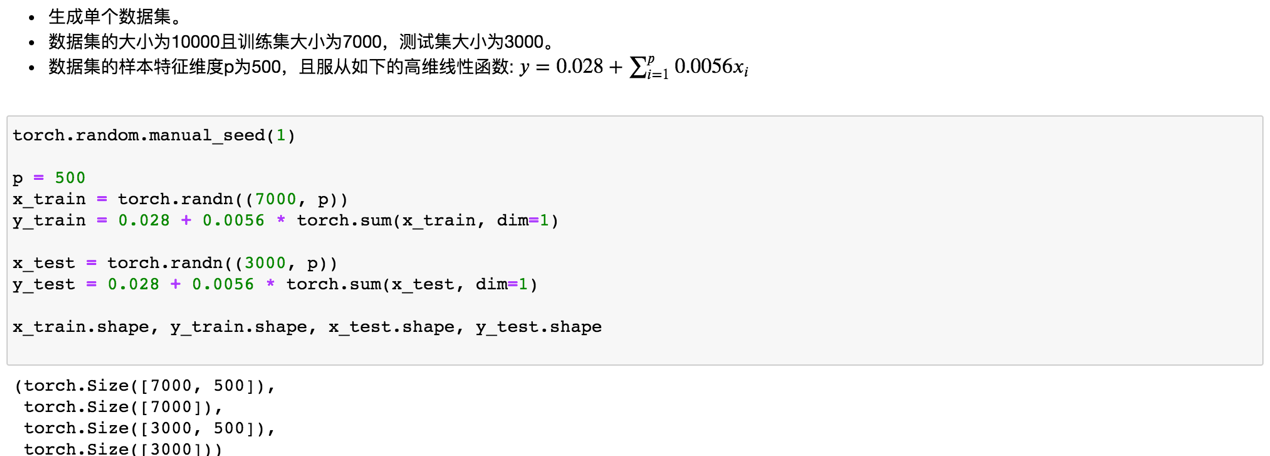
## 三、实验环境及实验数据集

1. 实验环境

Ubuntu 18.04、CPU 4核、32GB、Pytorch 1.6.0、Jupyter Notebook

1. 数据集：

* 回归任务



* 二分类任务



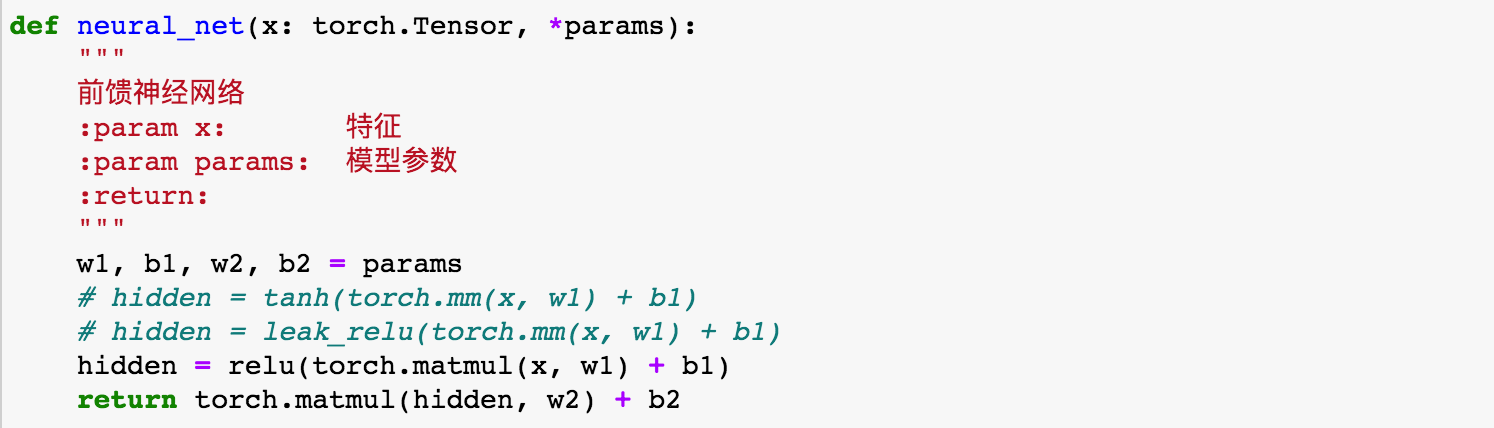
* 多分类任务



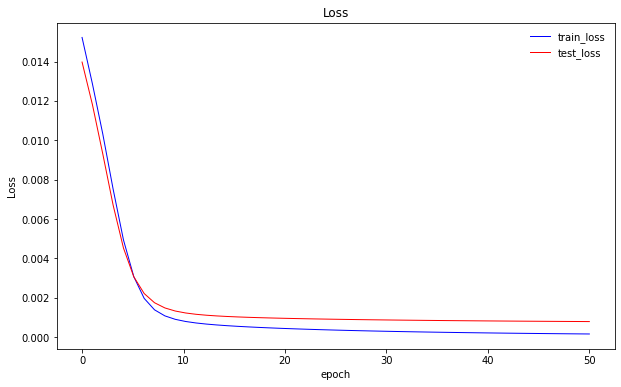
## 四、实验过程

#### 1. 回归任务

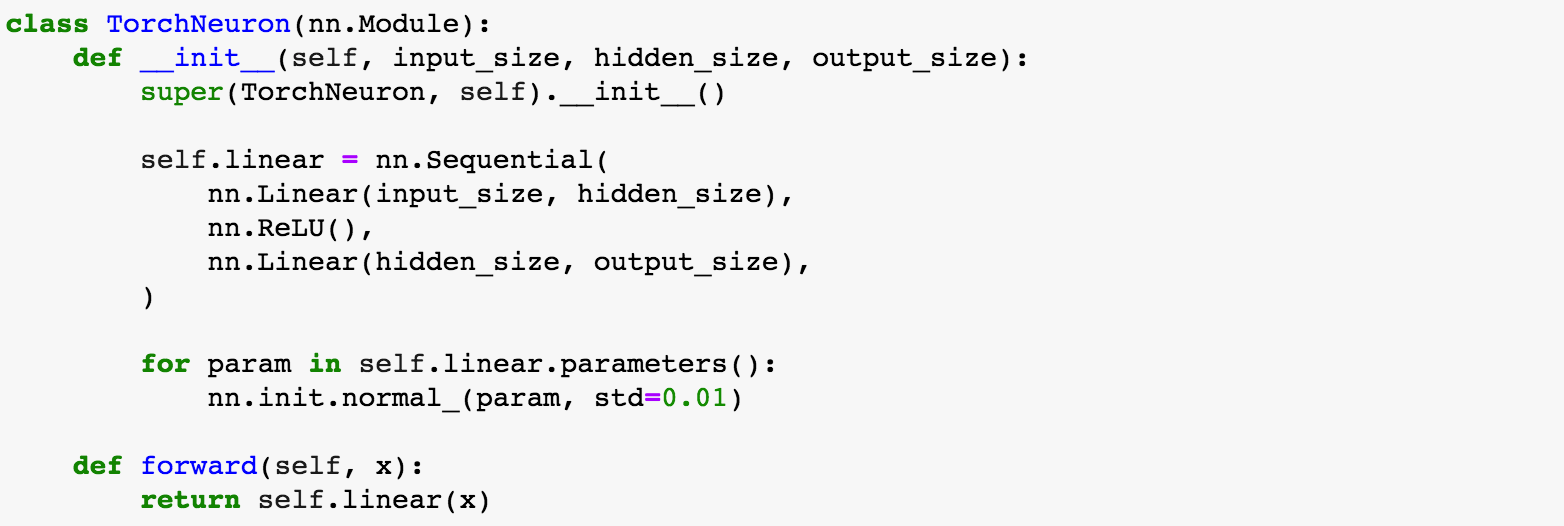
#### 1.1 手动实现

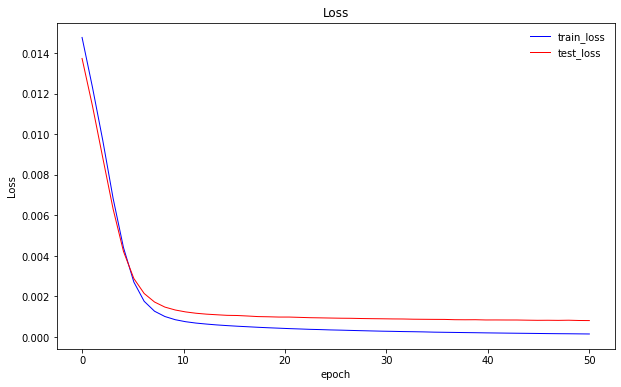


训练集与测试集的loss均接近0，说明模型能很好地学习数据集。



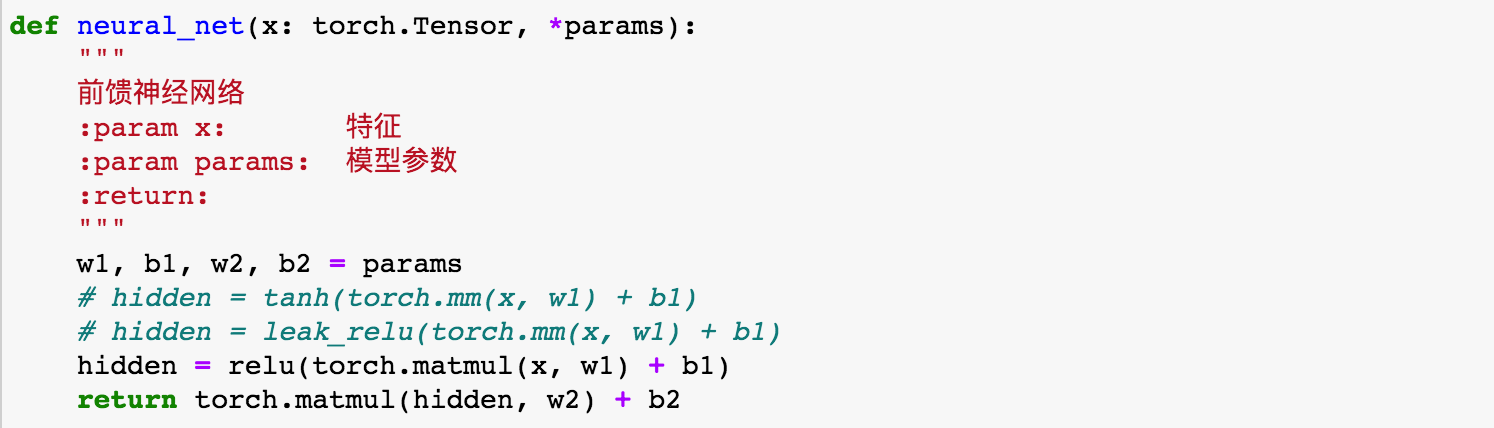
#### 1.2 torch.nn实现



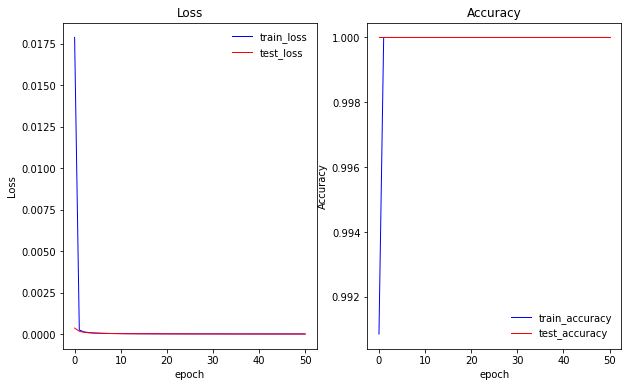


#### 2. 二分类任务

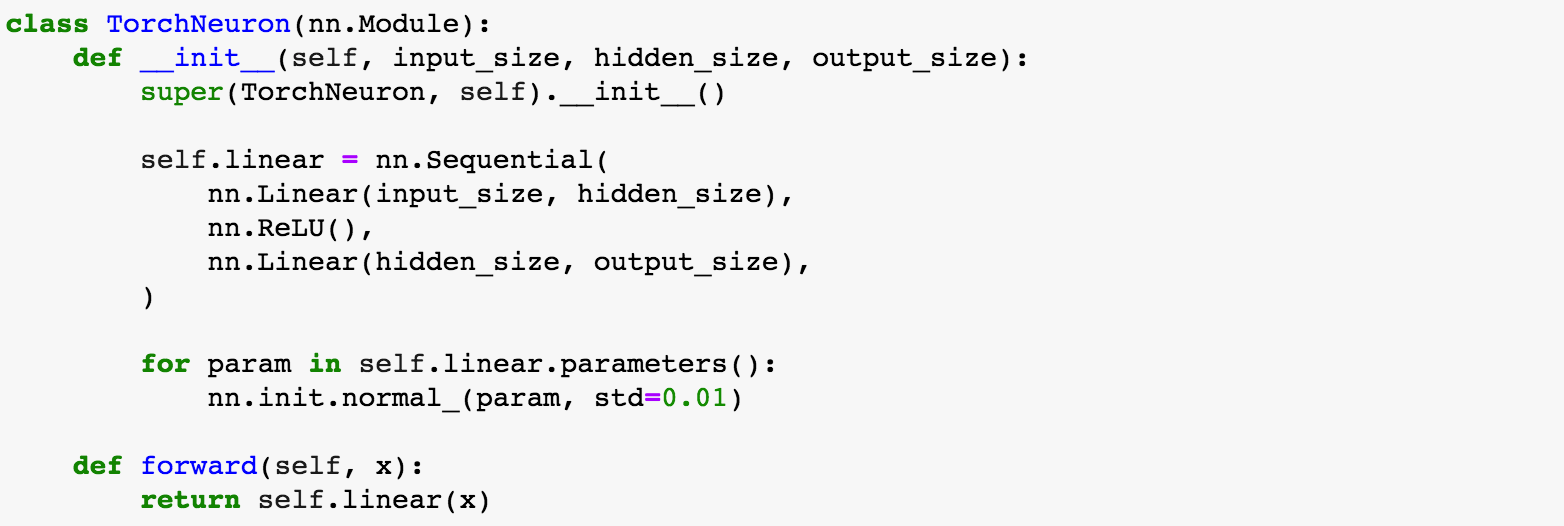
#### 2.1 手动实现

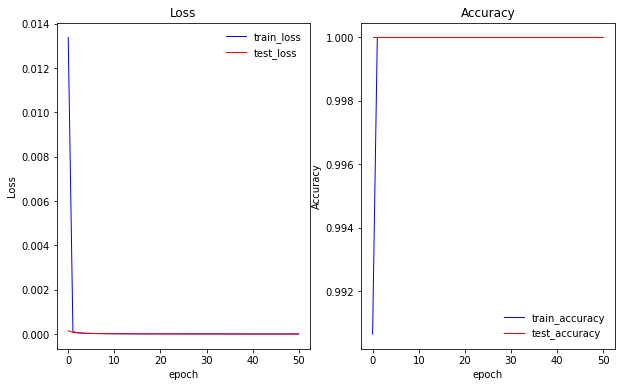


训练集与测试集的loss均接近0，并且他们的accuracy都等于1，说明模型能够完美地学习到数据集的规律。



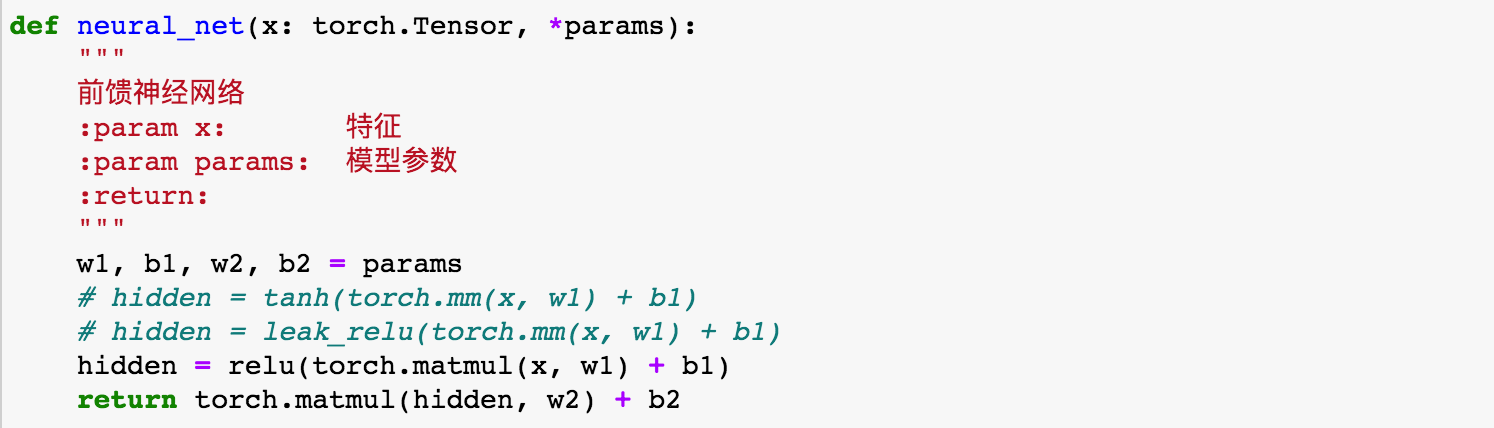
#### 2.2 torch.nn实现



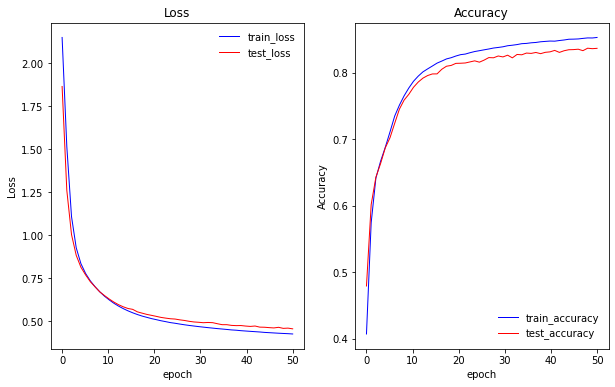


#### 3. 多分类任务

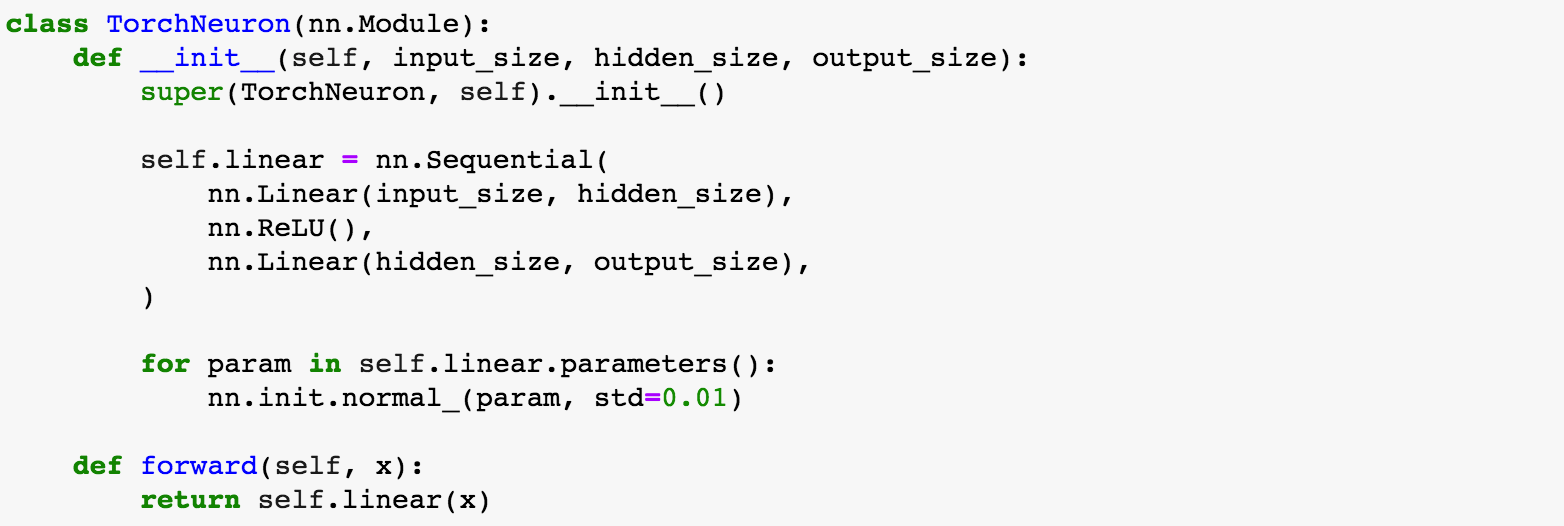
#### 3.1 手动实现

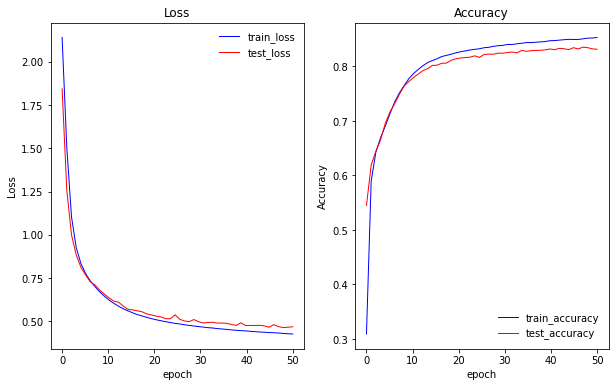


模型在训练集和测试集的表现不错，也没有出现过拟合或欠拟合的情况，准确率在训练集上达到85.2%，在测试集上达到83.65%。



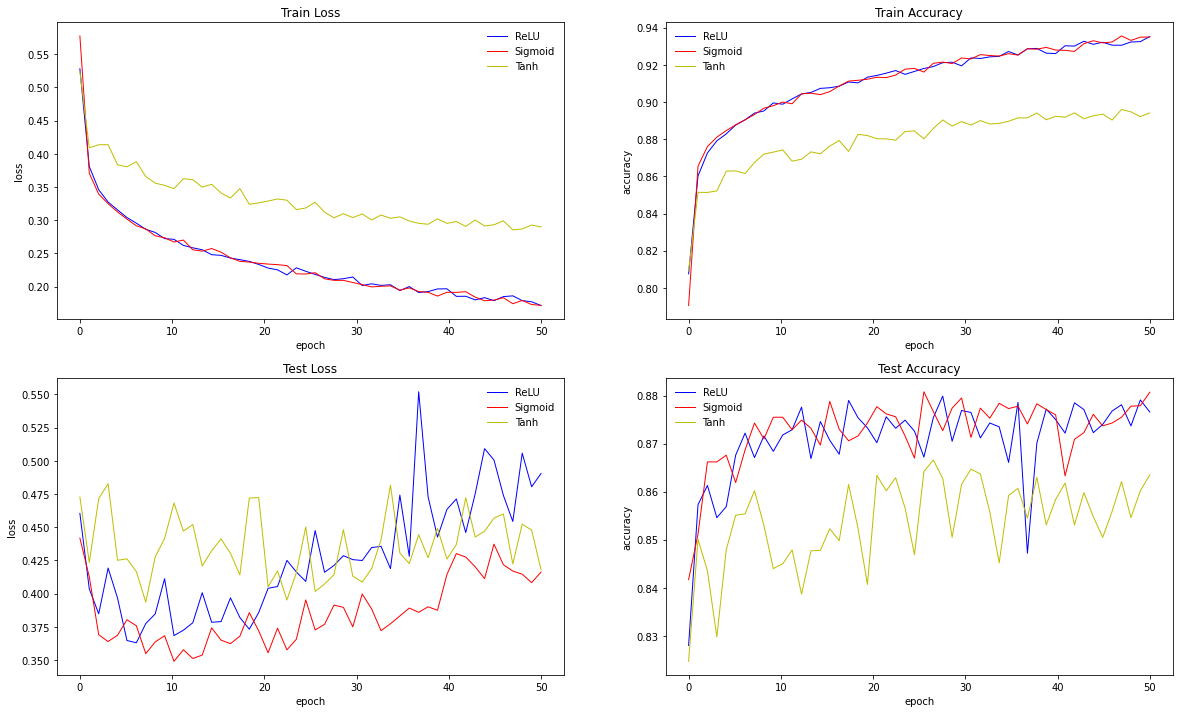
#### 3.2 torch.nn实现





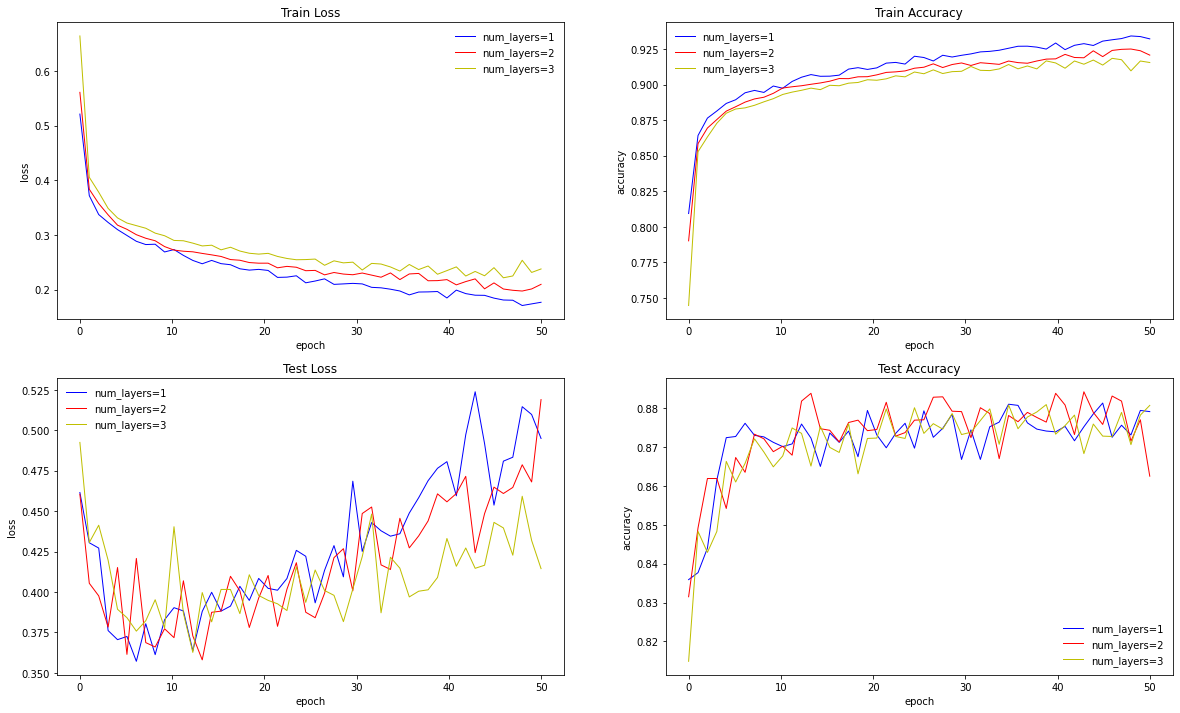
#### 3.3 使用至少3种不同的激活函数

分别使用以下3种激活函数：Relu、Tanh、Sigmoid。结果如下图，可以看出Than函数的效果最差，Relu和Sigmoid相差不大，但激活函数一般不用Sigmoid，而是使用Relu。



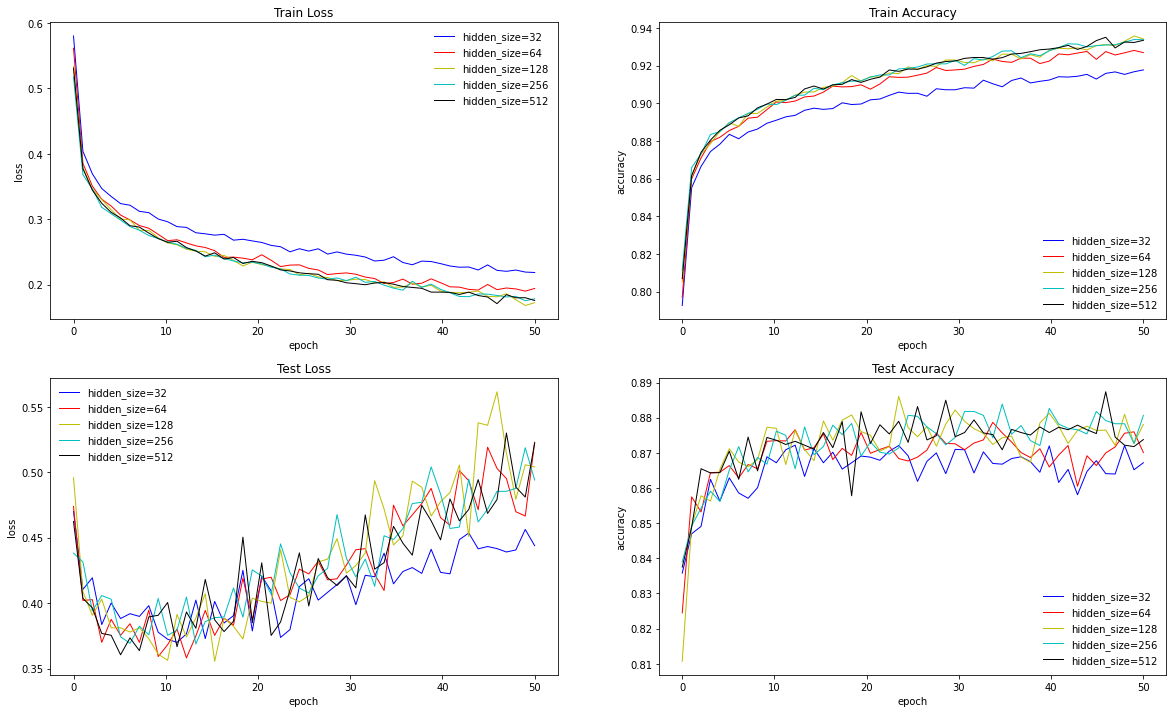
#### 3.4 使用不同的隐藏层层数

分别设置隐藏层层数为1、2、3。结果如下图，层数越多，训练时间越长，但效果提升并不显著，相反，层数达到一定数目后容易过拟合。综合考虑，隐藏层层数=1最佳。

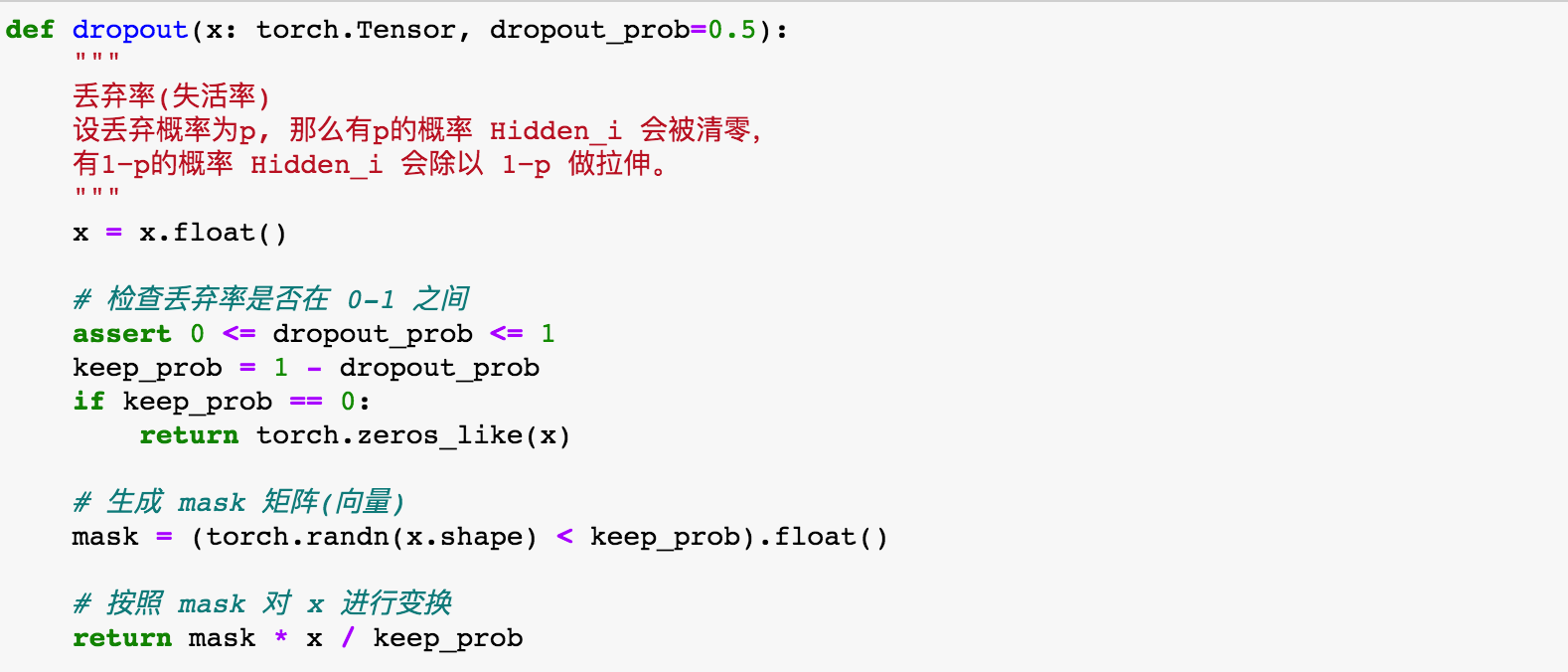


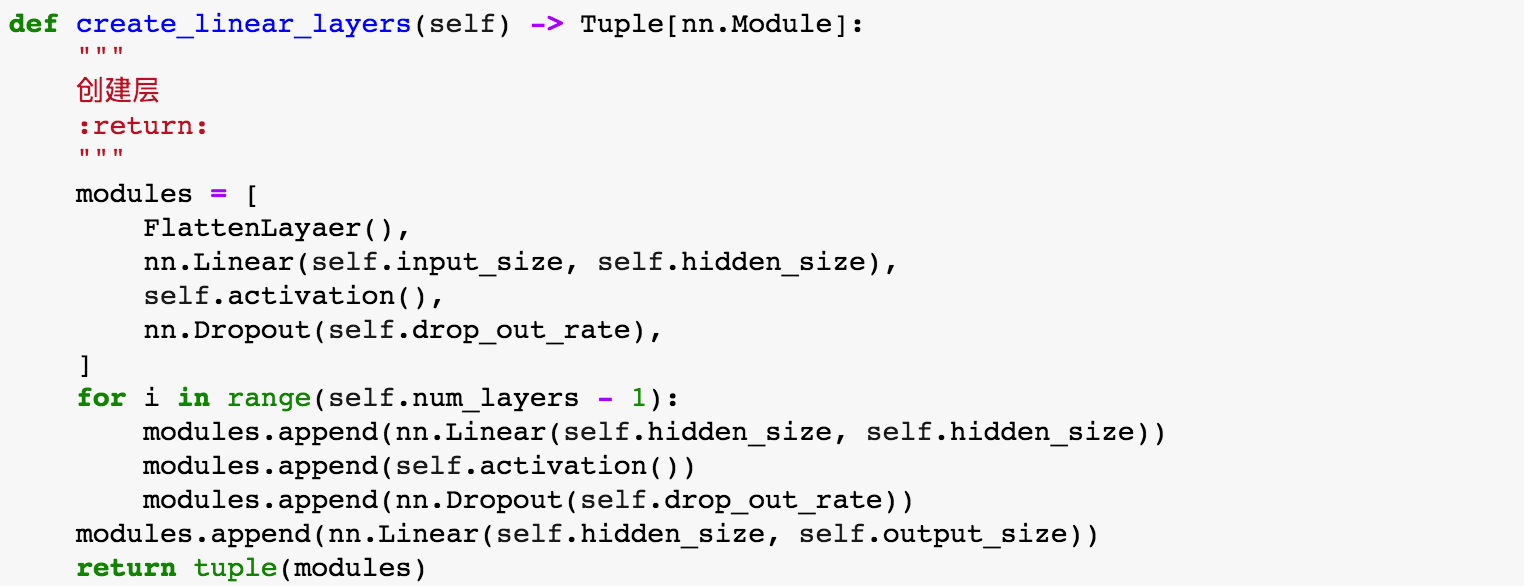
#### 3.5使用不同的隐藏单元个数

分别设置隐藏单元个数为32、64、128、256、512。结果如下图，隐藏单元个数越多，训练时间越长，效果提升也较为明显，但当达到一定的个数，提升不显著，容易过拟合。综合考虑训练时长和效果，隐藏单元个数=128最佳。

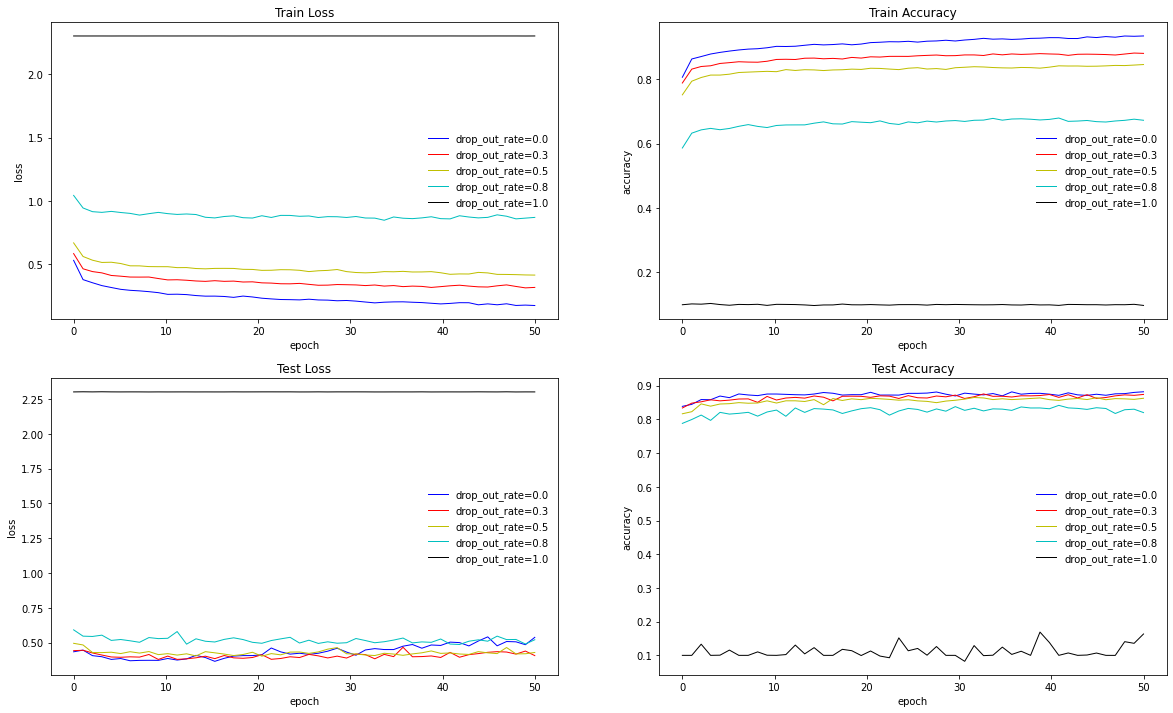


#### 3.6 探究不同丢弃率对实验结果的影响

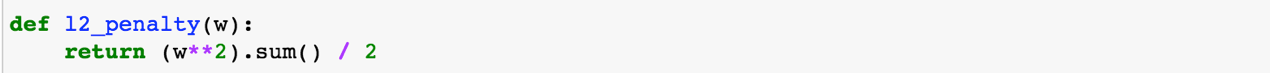


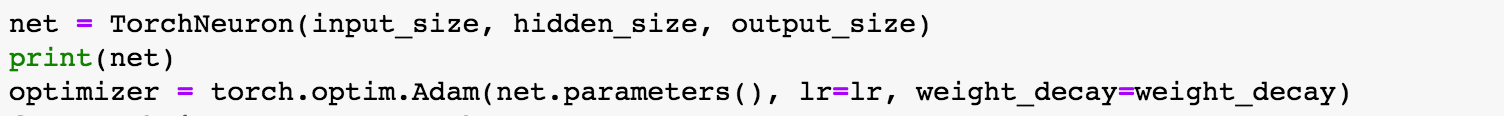


分别设置丢弃率为0.0、0.3、0.5、0.8、1.0。丢弃率偏大或偏小，都无法起到防止模型过拟合的效果。结果显示，丢弃率=0.3最佳。

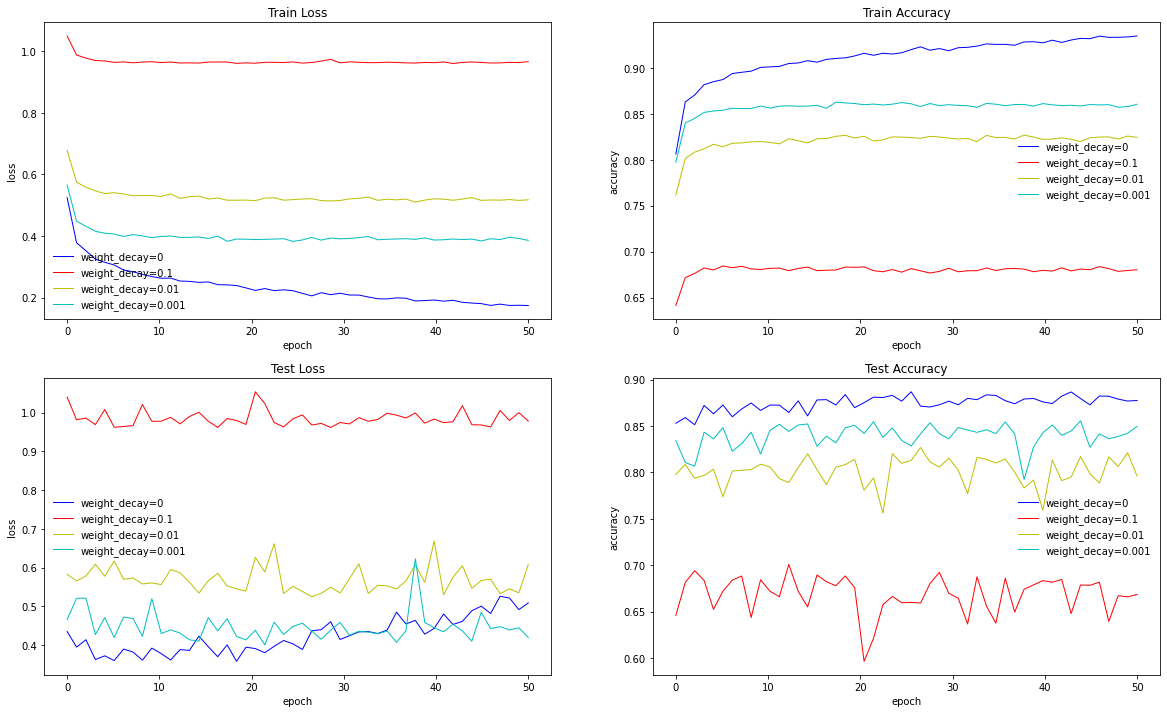


#### 3.7 探究惩罚项的权重对实验结果的影响

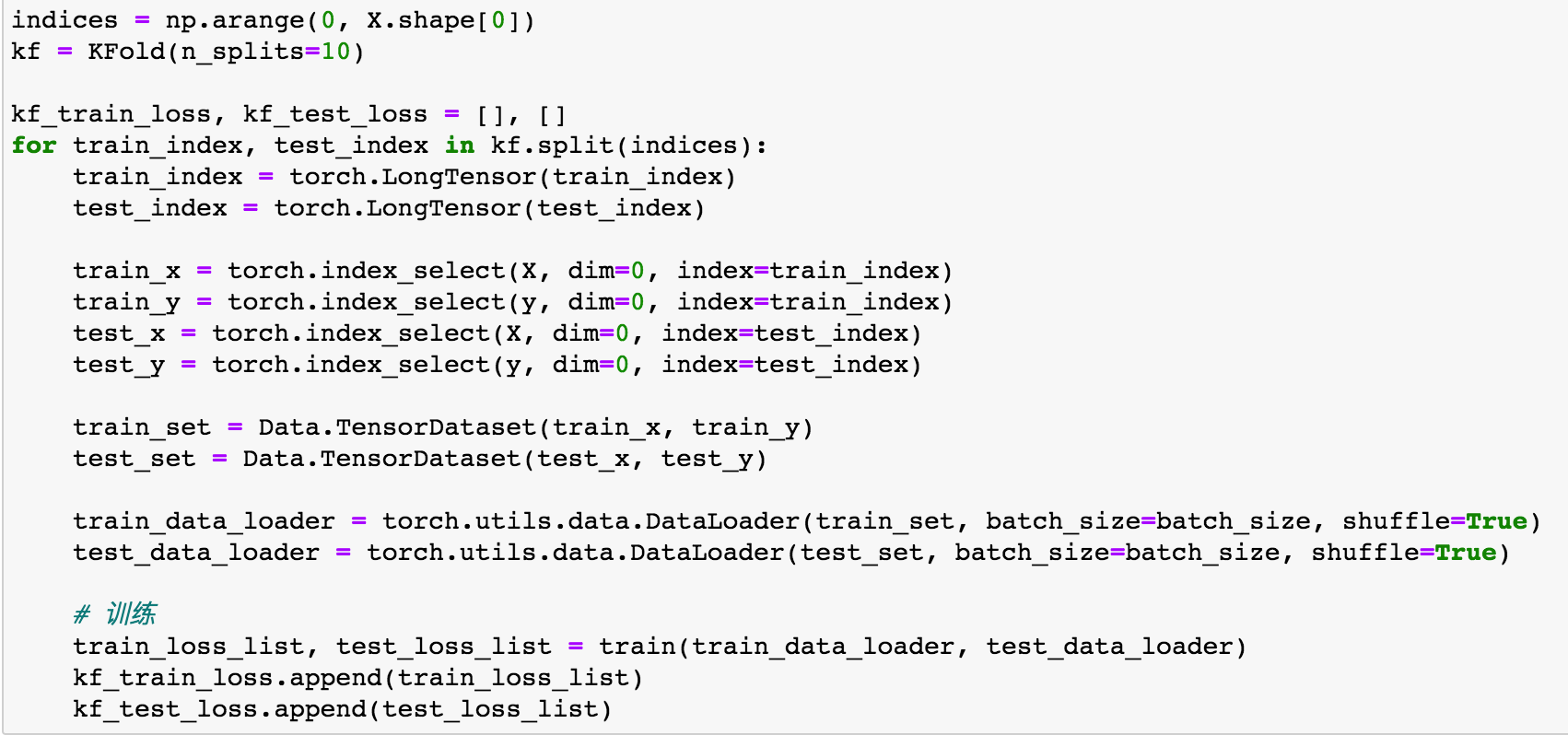




利用torch.optim的weight\_decay参数实现L2范数正则化。分别设置权重衰减系数为0、0.1、0.01、0.001。结果显示，权重衰减系数=0.0最佳。



#### 4. 交叉验证(10折)



根据上述比对实验，按如下表格的值设置超参数。通过10折交叉验证，记录模型在测试集上的损失(LOSS)和ACC(分类任务的准确率)，对每次验证中的所有迭代的损失值、准确率（epoch）分别取均值、最小或最大值列于表格中。

|  |  |  |
| --- | --- | --- |
| 超参数 | 值 | 说明 |
| activation | Relu | 激活函数 |
| num\_layers | 1 | 隐藏层层数 |
| hidden\_size | 128 | 隐藏单元个数 |
| drop\_out\_rate | 0.3 | 丢弃率 |
| weight\_decay | 0.0 | 权重衰减系数(惩罚项) |

#### 4.1 回归任务

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| loss | F1 | F2 | F3 | F4 | F5 |
| 最小 | 0.00028868 | 0.00031434 | 0.00031275 | 0.00030661 | 0.0003032 |
| 均值 | 0.00163858 | 0.00184026 | 0.00177395 | 0.00171711 | 0.00174201 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| loss | F6 | F7 | F8 | F9 | F10 |
| 最小 | 0.0003139 | 0.0003213 | 0.00025909 | 0.00034103 | 0.00032338 |
| 均值 | 0.00167926 | 0.00180045 | 0.00147343 | 0.00179616 | 0.00171343 |

#### 4.2 二分类任务

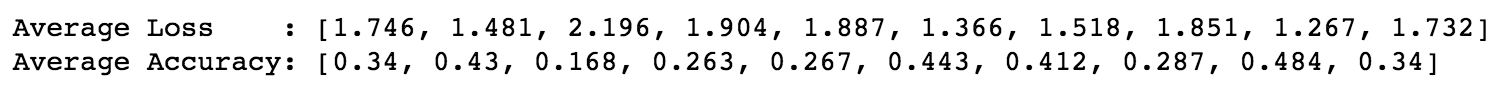
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| loss | F1 | F2 | F3 | F4 | F5 |
| 最小 | 1.55238e-05 | 1.34269e-05 | 1.41567e-05 | 1.3864e-05 | 1.702e-05 |
| 均值 | 3.80892e-05 | 2.95931e-05 | 3.02551e-05 | 3.00e-05 | 3.879e-05 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| loss | F6 | F7 | F8 | F9 | F10 |
| 最小 | 1.67884e-05 | 1.78620e-05 | 1.52344e-05 | 1.2095e-05 | 1.448e-05 |
| 均值 | 3.66381e-05 | 3.89902e-05 | 3.42971e-05 | 2.55e-05 | 3.24e-05 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ACC | F1 | F2 | F3 | F4 | F5 | F6 | F7 | F8 | F9 | F10 |
| 最大 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| 均值 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |

#### 4.3 多分类任务

对多分类任务进行交叉验证时，loss和accuracy都不理想，很纳闷，可能是代码中存在bug，困扰了很久，依然没有解决。



## 五、实验结果

## 六、实验心得体会

通过本次实验，了解到了前馈神经网络的基本实现，以及常用超参数如失活率、惩罚项权重、激活函数等的使用。这让我对整个深度学习有了个初步的认识，通过不断地调整参数、尝试，使得模型的准确率逐提高，虽然过程比较煎熬，但成就感爆棚。

## 七、参考文献

## 八、附录

需要补充说明的内容，如无可略。

**实验报告编写要求**

1. 正文要求小四号宋体，行间距1.5倍；

2. 英文要求小四号Times New Roman；

3. 在实验内容、实验过程、实验结果三部分需要针对当次实验不同的实验内容分别填写（模版以实验一为例），实验设计中如有必要也可以分开填写；

4. 实验报告配图的每幅图应有编号和标题，编号和标题应位于图下方处，居中，中文用五号宋体；

5. 表格应为三线表，每个表格应有编号和标题，编号和标题应写在表格上方正中，距正文段前0.5倍行距。表格中量与单位之间用“/”分隔，编号与标题中的中文用五号宋体；

6. 图、表、公式、算式等，一律用阿拉伯数字分别依序连续编排序号。其标注形式应便于互相区别，可分别为：图1、表2、公式(5)等。