**樊赫铮实验报告1——TASK1.1 Pytorch实现ResNet**

**实验时间：**

6.22-6.23

**Task 1任务描述**

第一个任务是用pytorch实现ResNet，并应用在股票数据上。

我们会提供一份标注股票数据集，你们需要实现模型，并在我们提供的数据集上进行训练和validation。

数据在服务器 /data01/data\_for\_intern/ 路径下。

可用的feature在：factor\_data\_13\_19.csv、 trading\_data\_13\_19.csv、 valuation\_data\_13\_19.csv 三个文件中。

可用的label在：return\_data\_13\_19.csv

将任务分为1.1：实现ResNet并利用MNIST数据集验证网络效果（本次实验报告）；1.2：带入股票数据。

**代码地址**

<http://124.251.74.84:42206/user/fanhezheng/notebooks/Task1.1_Pytorch_ResNet_Implementation_Test_on_MNIST.ipynb#>

**实验目标**

实现ResNet并利用MNIST数据集验证网络效果。

**实验思路**

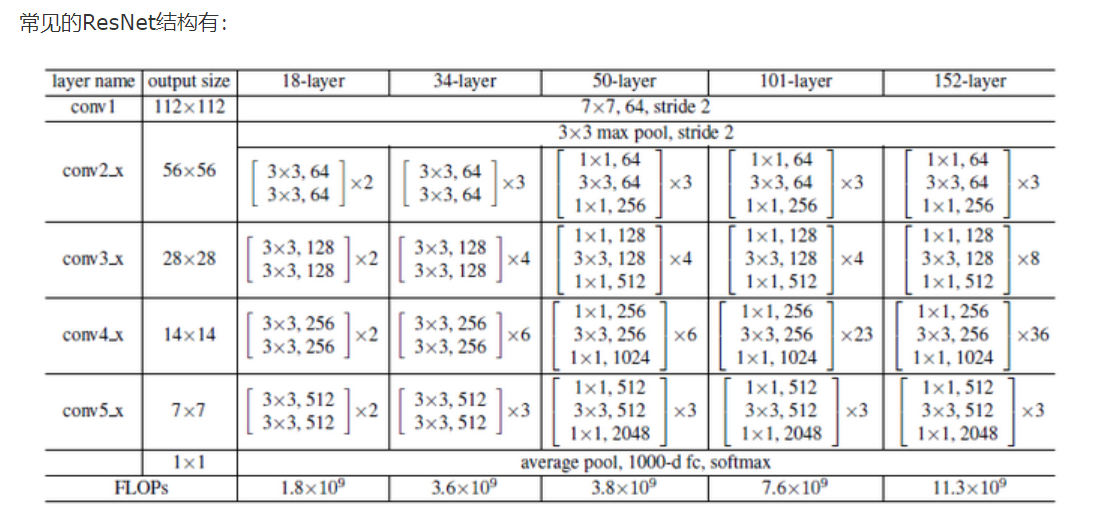
主要以实现resnet34的结构为主，但要保证代码质量使后续调用时可以灵活地go deeper或者调整网络结构。

完成网络部署后利用sklearn的

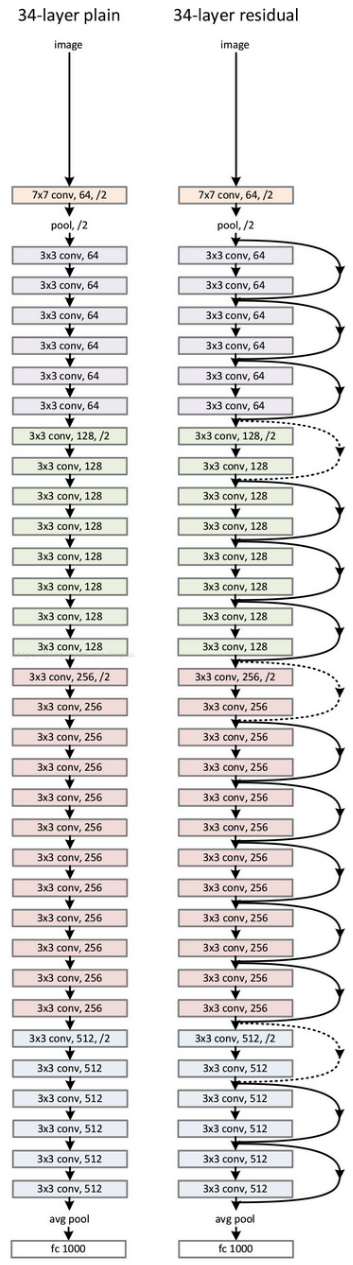
from sklearn import datasets

digits = datasets.load\_digits()

导入MNIST数据集对网络进行训练和测试，检验网络有没有出错的地方。

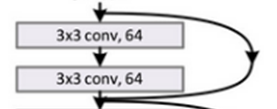
****

Resnet34 结构



**实验内容和过程中需要注意的事项总结**

1.代码中的BasicBlock为ResNet34中主要的block

类似于此的一个结构

创建BasicBlock为一个类，里面两个函数

（1）\_\_init\_\_(self, in\_channel, out\_channel, stride=1, downsample=None)

此函数中主要调用conv等函数实现卷积模块的搭建。需要注意的是在同一个block中每个data conv后要保证size不变（resnet34的要求）

Downsample的左右主要是监督残差需不需要改变size或channel

（2）forward(self, x)

前向传播函数

2. 其他残差会改变的size和channel需要用到的block：Bottleneck

结构与BasicBlock类似

3. 搭建整体网络ResNet

（1）\_\_init\_\_(self, block, blocks\_num, num\_classes=1000, feature\_channel=3,device=torch.device("cpu"))

用来拼装layer；同时在此函数里完成weights的初始化，注意weights一定要to（device）才能使数据类型一致，后续才能用gpu训练。

（2）\_make\_layer(self, block, channel, block\_num, stride=1)

在此函数中处理down\_sample

注意layer搭建的逻辑

(3) forward(self, x)

正向传播；注意fully connection部分

3. 测试

#实现custom pytorch dataset； GetLoader(Dataset)函数

这里实现的主要是custom numpy array的数据，和之后带入股票数据会比较类似。

如果要用已经整理好的数据文件可以选择pd.read\_csv 百度有详细步骤

训练时要注意数据类型（float double long...）不符合的报错之后手动转换

预测函数的写法和训练函数较为类似，但也要注意数据类型问题

**实验表现**

利用以下Hyperparameter进行训练：

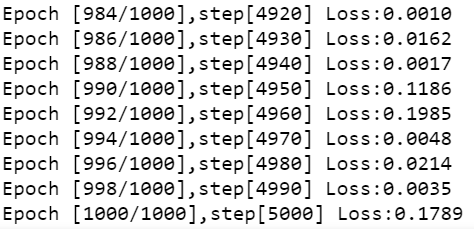
num\_epochs=1000

num\_classes=10#number 0~9

batch\_size=256

learning\_rate=0.001

得：



测试准确率：94.01947148817803%

**实验结论**

实现的ResNet34是成功的

但后续带入股票数据要把代码再整理一下，网络结构也要根据需要变化下（比如卷积的kernel\_size）。不过代码已经有了，应该在网络代码部分不会花太长时间。

下一部分主要的问题是股票数据的逻辑处理 和 有一个可尝试的逻辑之后处理dataset（预计工作量较大）