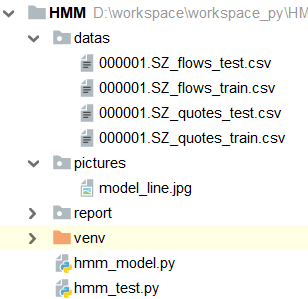
## HMM项目介绍

1. **简介**

HMM项目是基于隐马尔可夫模型的股票预测及回测的系统。

1. **项目结构**

HMM项目的目录结构如下图



datas文件夹保存股票数据的csv文件，csv文件名的末尾train和test分别表示是训练集和测试集的数据。

pictures文件夹保存模型回测结果的曲线图，曲线图中绿线表示股票实际收盘价走势，红线表示模型回测的走势（不含手续费），黄线表示模型回测的走势（含手续费）。

report文件是本报告所在文件夹。

venv文件夹忽略。

hmm\_model.py文件是对隐马尔可夫模型的抽象类，包含模型的训练和回测。

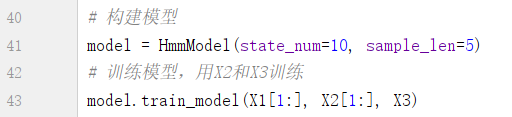
hmm\_test.py文件获取股票的训练集和测试集数据，并调用模型的训练和回测。

1. **代码流程说明**
2. **hmm\_test.py**

1.1获取训练集的数据（9-38行）



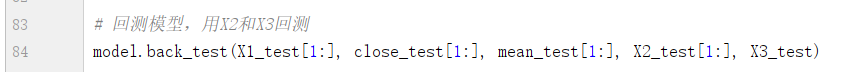
1.2构建并训练模型（40-43行）



1.3获取测试集的数据（48-81行）

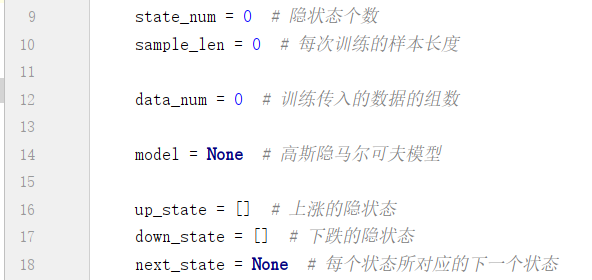


1.4对测试集用训练好的模型进行回测（84行）

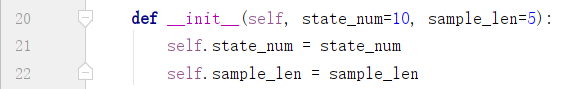


1. **hmm\_model.py 定义模型的类名为HmmModel**

2.1类的属性（9-18行）

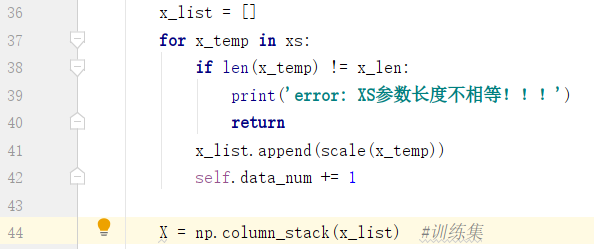


2.2类的构造方法（20-22行）



2.3训练模型的方法（24-85行）

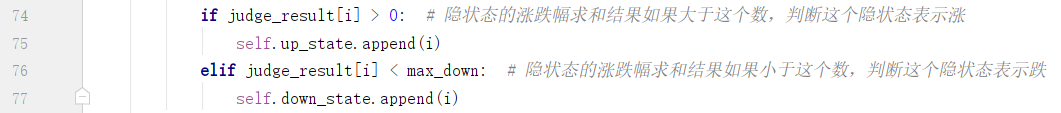
36-44行：对传入的训练的数据做归一化处理并整合成一个二维矩阵



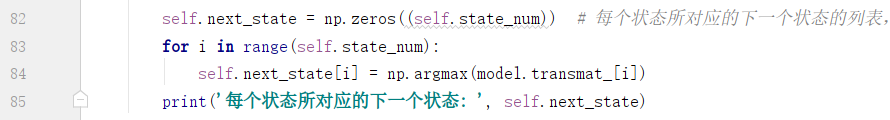
53-55行：用上面整合的二维矩阵训练模型



74-77行：判断每个隐状态表示涨还是跌，如下两种情况之外的求和结果表示震荡的隐状态

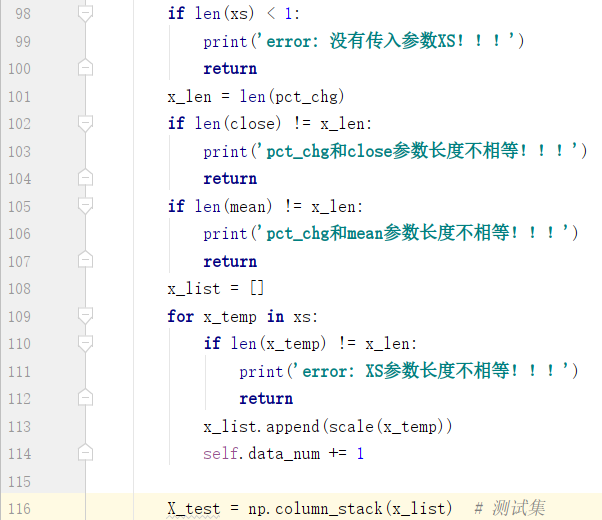


82-85行：获得每个隐状态最可能出现的下一个隐状态

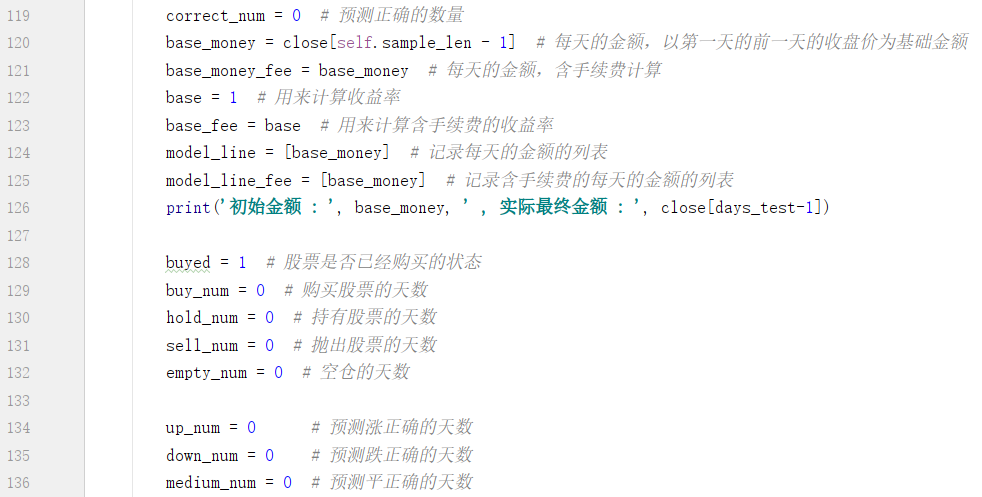


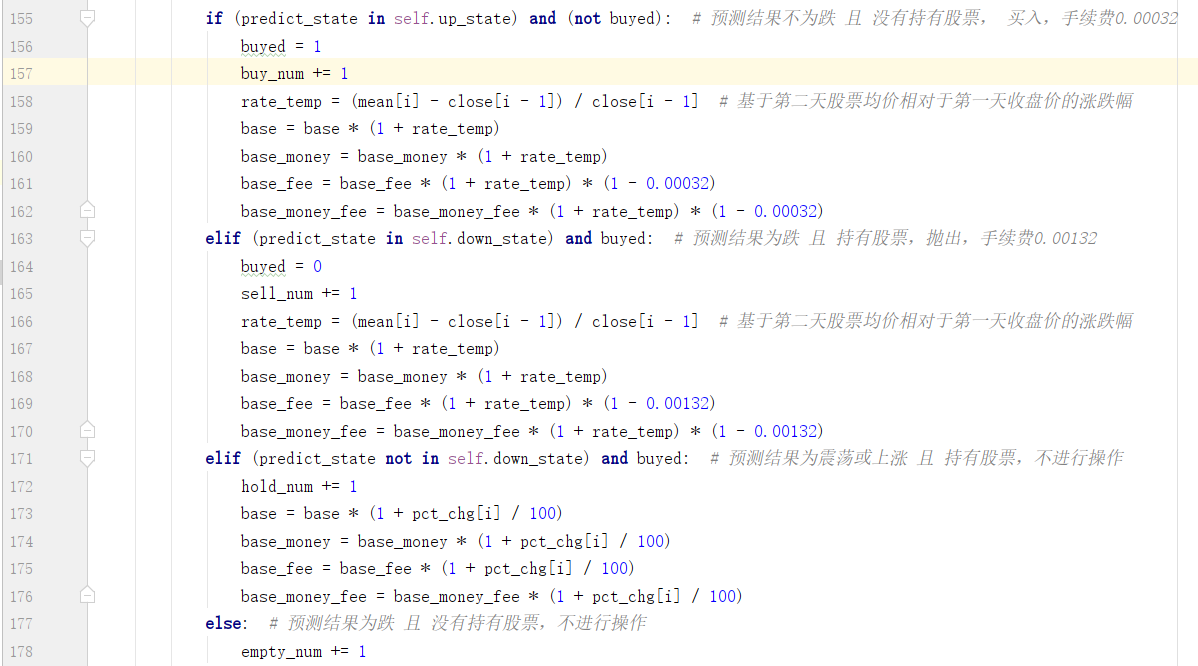
2.4回测方法（87-200行）

98-116行：校验传入数据并做归一化处理，整合成二维矩阵



119-136行：初始化统计的数据



138-180行：根据模型的样本长度sample\_len，对当天的前sample\_len天作隐状态序列的预测，根据序列的最后一个隐状态（即当天的前一个交易日的模型预测的隐状态）以及隐状态转移矩阵，获得当天的隐状态。155-178行是根据预测的当天的隐状态，进行股票的买入卖出并计算收益。（第55行需要说明的是，如果是下跌的趋势，代码可修改为预测状态为涨买入，如果是上涨的趋势，代码可修改为预测状态不为跌，即涨或平，买入）

194-200行，绘制图表

