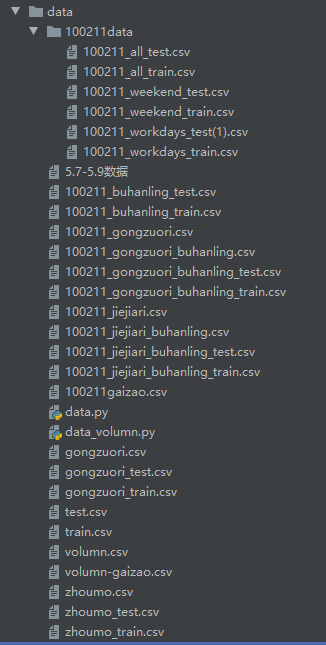
基于模型方式的寻找，首先进行深度学习模型来进行对数据集的训练

环境采用anaconda+pycharm+python3.6

Pandas tensorflow keras numpy matplotlib scikit-learn等库的利用

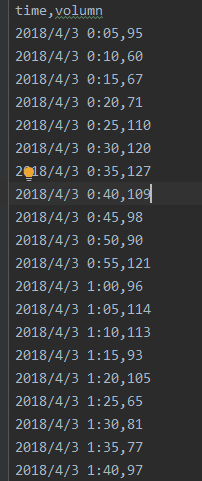
数据集来自于github，以csv文档格式进行存储



存储格式：.csv

Train.py进行对数据集的训练

Test.py进行预测和数据集的测试，输出mape



为年月日 时间（从零点开始） 时间间隔五分钟 车辆数目

首先对数据集进行平滑降噪处理，转换成可利用的数据集

模型的应用：

Lstm gru SAEs

简单介绍：

lstm：  长短期记忆模型（long-short term memory）是一种特殊的RNN模型，是为了解决RNN模型梯度弥散的问题而提出的；在传统的RNN中，训练算法使用的是BPTT，当时间比较长时，需要回传的残差会指数下降，导致网络权重更新缓慢，无法体现出RNN的长期记忆的效果，因此需要一个存储单元来存储记忆，因此LSTM模型被提出。

Gru：GRU即Gated Recurrent Unit。前面说到为了克服RNN无法很好处理远距离依赖而提出了LSTM，而GRU则是LSTM的一个变体，当然LSTM还有有很多其他的变体。GRU保持了LSTM的效果同时又使结构更加简单，所以它也非常流行。

SAEs：利用历史刷卡数据，提出了一种基于深度学习的地铁短时客流量预测方法，基于栈式自编码器构建深度神经网络模型，采用自下而上逐层非监督预训练，在预训练结束之后，采用反向传播ＢP算法自上而下来微调整个网络的参数。

该模型底层由栈式自编码器构成，用于特征提 取；顶层为一个逻辑回归器，用于客流量预测。

**AE** ，自动编码器 AE（AutoEncoder）是一种试图重构原始输入信号的神经网络，  
**SAE**，栈式自编码器SAE由一组自动编码器连接构 成，它将下层自动编码器的输出作为上层自动编码器的输入。L层的SAE,训练集作为第一次的输入，第k（小于L）层的输出作为第k+1的输入。堆栈的方式逐层组合在一起。文中为了预测流量，在编码层添加一个预测器-逻辑回归，

**评价指标** ，MAE，MAPE，RMSE

**数据**，地铁刷卡数据

**结论**，SAE模型比 Ｗａｖｅｌｅｔ －ＮＮ 模型和 ＡＲＩ － ＭＡ模型预测精度更高。

代码参考：<https://github.com/xiaochus/TrafficFlowPrediction>

Model.py作为调用类进行调用

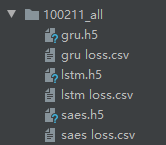
将lstm gru seas 以.h5的形式进行存储然后进行调用，依赖h5py包

并把训练得到的

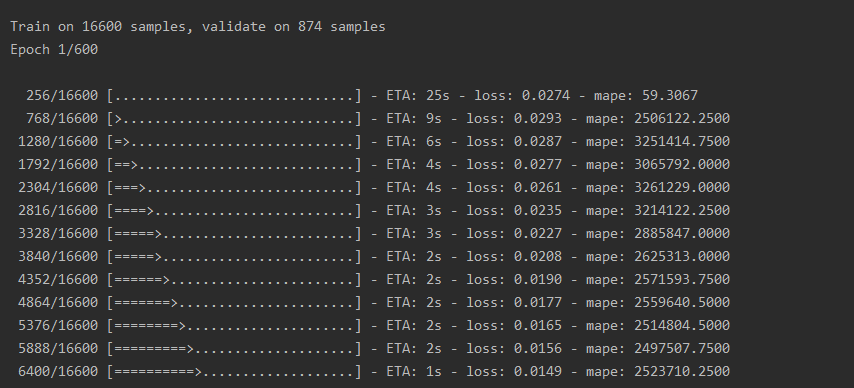
loss,mean\_absolute\_percentage\_error,val\_loss,val\_mean\_absolute\_percentage\_error

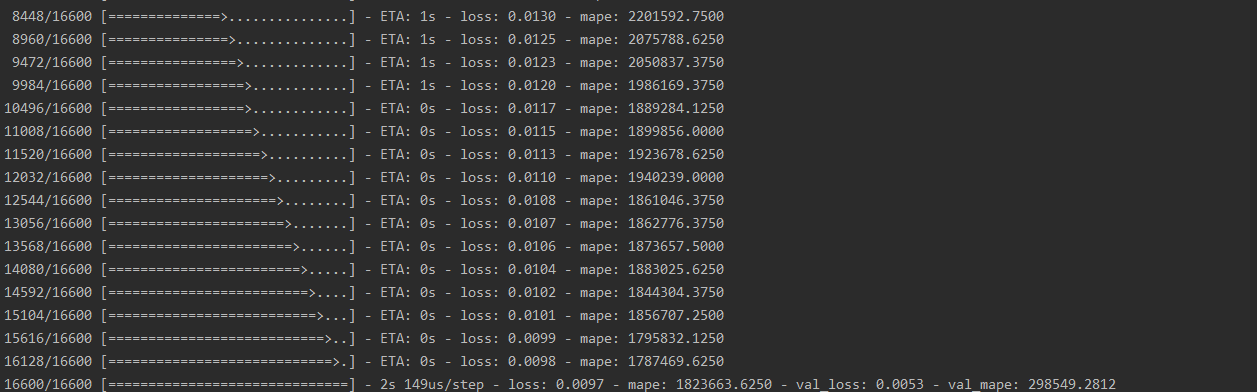
以.csv文件形式进行存储

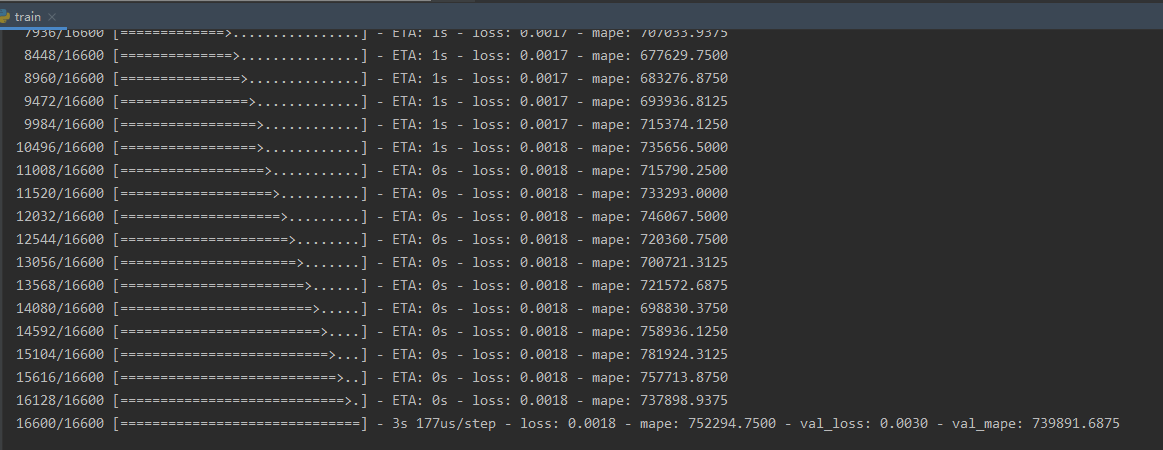
例：



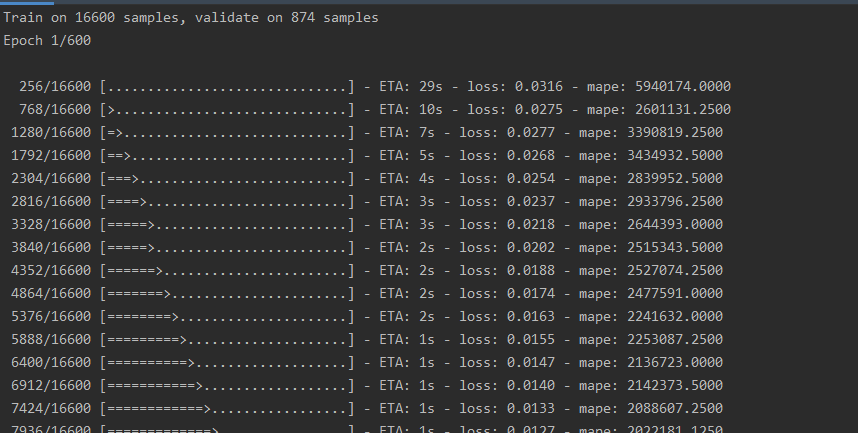
Lstm:

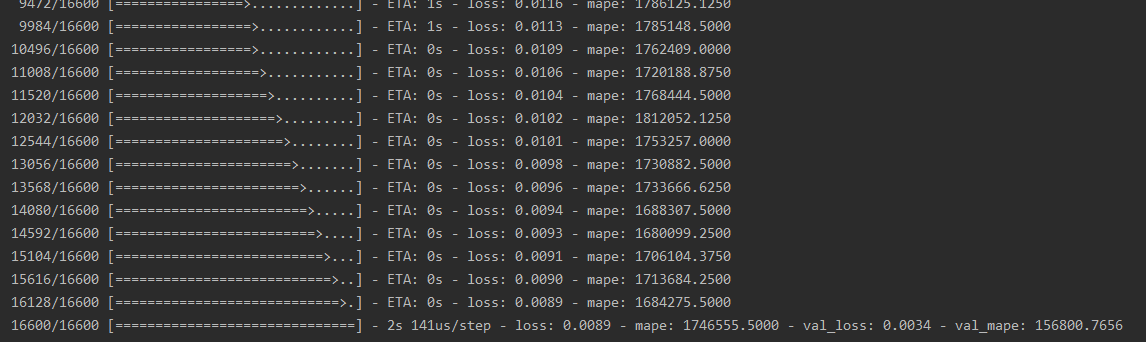


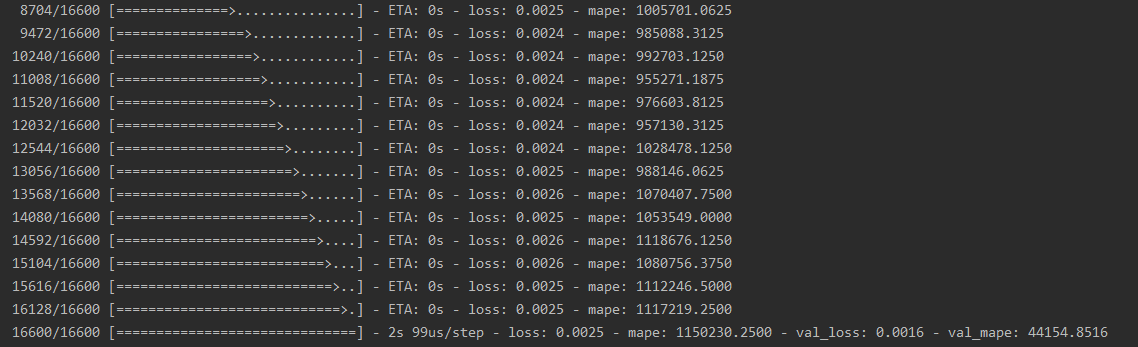




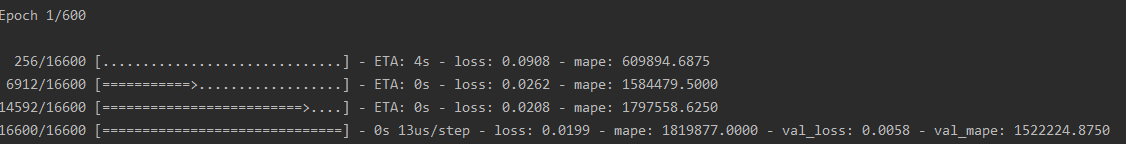
Gru:







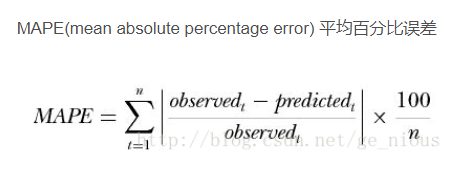
Saes:



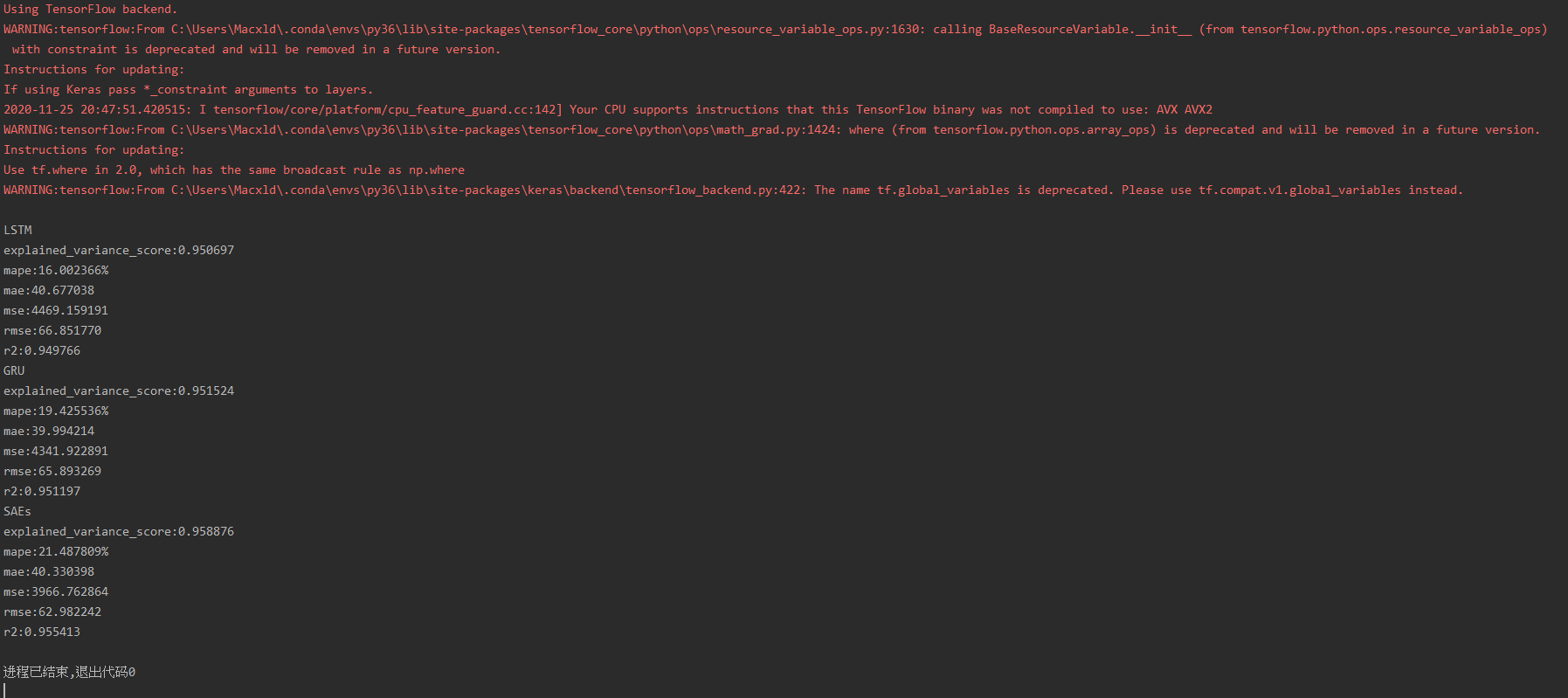
train.py训练600次

Loss为损失函数值

Mape如下：



Test.py的初步测试预测



分别展现出来预测各个指数的值

统筹各个指数

Mape越接近于0%越准确

绘制的图像为：

