团队：

汪凯 221017000261

• 业务目标：加载IMDb的数据集，使用简单的LSTM模型对数据集进行情感分析。

• 处理流程：

o 准备数据集

.使用nltk下载停用词

.在网上下载IMDb数据集

o 模型定义：使用LSTM模型

o 模型训练与调参

o 测试

• 停用词和数据集情况说明：

o 总句子数：25000个句子

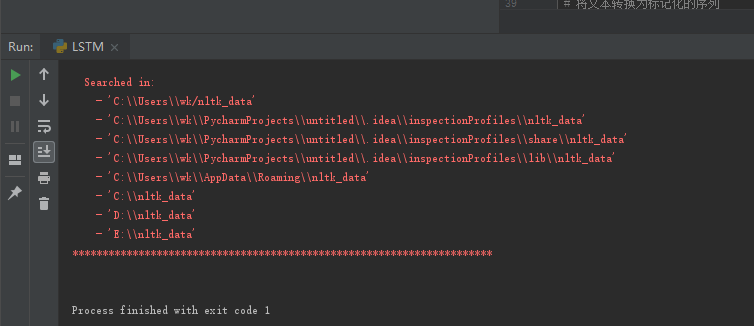
o 英文停用词：179 字符

• 模型训练：

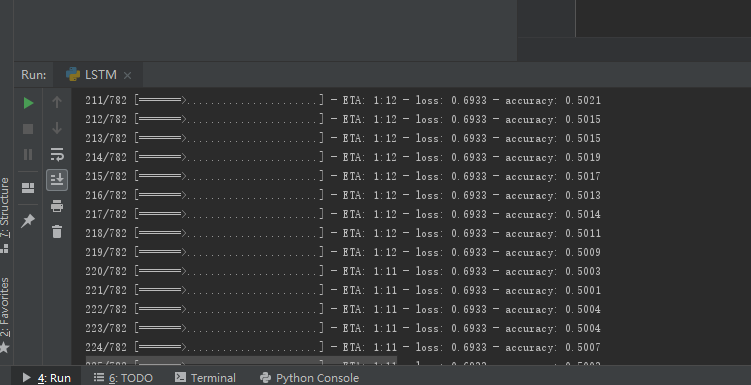
o 训练环境：NVIDIA Geforce RTX 3060

o 在pycharm中执行LSTM.py

• 说明：若服务器网络不好的话， 使用nltk下载停用词失败，所以此次项目已经预先下载好了nltk\_data停用词数据。将数据放在以下路径都可以正常执行。



o 训练过程可以通过pycharm提供的窗口进行实时监视：



• 模型运行结果：

• 测试集前10个样本的预测结果和置信度：

