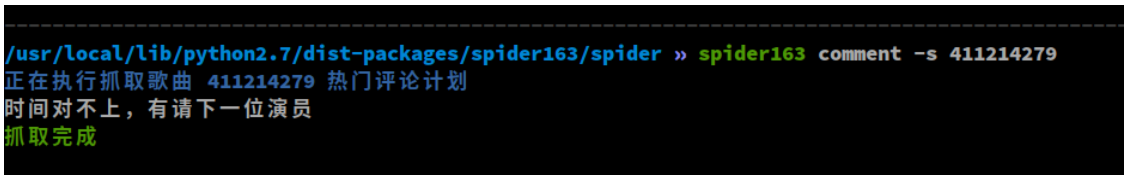
本周进度

- 学习斯坦福深度自然语言处理课程
  - 第一讲：NLP与深度学习入门
- 修改 Spider163 爬虫程序 -- 部分完成
- 论文查找和阅读

1.网易云数据爬取

- 爬取非精彩评论

  目前只修改部分源码，可以按照时间线进行评论的爬取，需要对爬取数据进一步解析，未进行文件保存



2. 问题总结

- 爬取数据中包含所以评论和其他比较乱的信息，需要进一步进行解析处理

下周进度

- 继续学习课程
- 修改爬虫程序
- 同步学习 pytorch 与 transformers