

第三方库

学习目的

掌握如何安装使用第三方库

自学任务

1. 通过第三方中文分词库，优化中文词频统计函数的功能

1. 阅读参考资料1
2. 通过 pip 安装 jieba 第三方库
3. 通过 jieba 精确模式将 stats_word.py 中 stats_text_cn 函数的输入字符串进行分词
4. 统计分词完成之后的词频（中文词只统计长度大于等于2的词）
5. 输出 tang300.json 文件内容中的词频前 20 的词和词频数

2. 提交作业

1. 将本地仓库关于本次作业的变更提交为一个 commit
2. 通过 Github 桌面客户端将本地电脑的变更推送到自己账户下的作业仓库
3. 回到 Github 自己账户下的作业仓库页面，向远程公用作业仓库的 master 分支发起 Pull Request，在提交的 Pull Request 的标题（title）中填写自己的班级编号+作业编号，如示例：19100101 day1，并在评论（comment）中 @自己的教练 提醒他检查作业

参考资料

1. [结巴中文分词](#)

