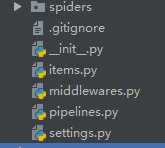
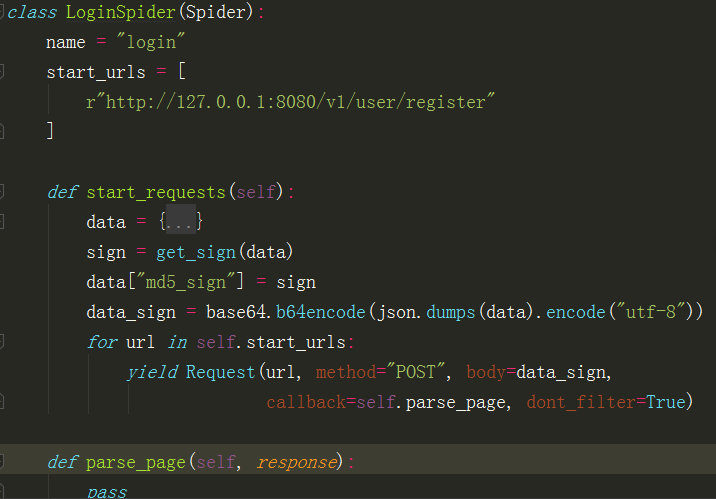
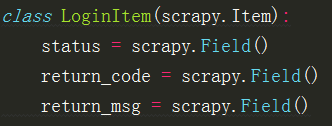
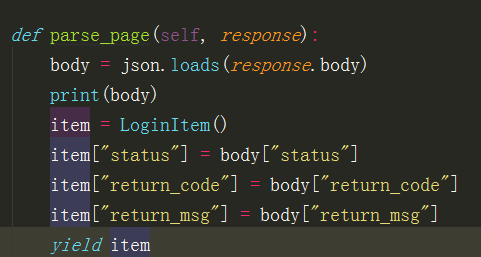
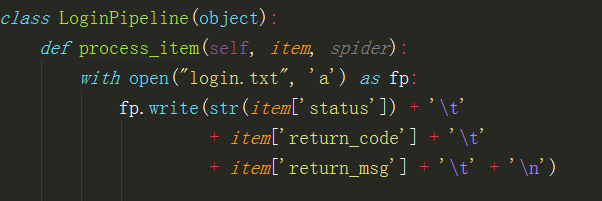
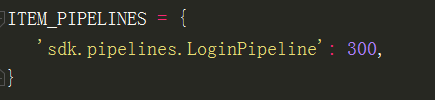
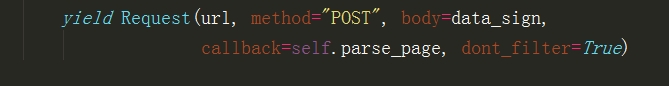
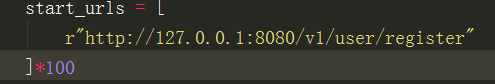
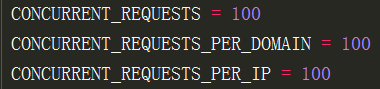
1. Scrapy安装：在windows环境python3.7上进行安装，命令pip install scrapy
2. 如果安装失败提示需要vc14的库，检查twisted是否安装成功，如果因为twisted问题可以在https://www.lfd.uci.edu/~gohlke/pythonlibs/ 上下载对应twisted包进行pip安装
3. 新建scrapy项目：scrapy startproject sdk
4. 新建项目结构如图
5. spiders/是编写爬虫的地方
6. items.py 是数据存储模板，用于结构化数据
7. middleware.py是添加中间件的地方
8. pipelines.py 做数据处理永久化存储
9. settings.py 爬虫设置
10. 在spiders目录下新建一个爬虫文件login.py编写接口测试用例
11. 只针对单个接口进行测试，start\_urls只需要添加测试接口
12. 重写start\_requests方法进行参数构造，使用scrapy.Request发送post请求同时定义回调函数parse\_page
13. 设置数据存储模板items.py
14. 编写parse\_page回调
15. 编写数据处理pipline.py
16. 添加设置settings.py300代表权重，数越小优先级越高
17. 启动爬虫scrapy crawl login
18. 出现错误No module named win32api解决办法：pip install pypiwin32

19.并发设置：1网络请求时dont\_filter=True不过滤相同接口，同时修改start\_urls变量

2

Settings.py设置这三个参数

详细示例见同文件夹