**getnews:** 数据准备、过滤新闻url、保存数据

新闻url列表

页面url

**geturllist:** 获取response、从response提取新闻url列表

main主函数

**SaveNew:** 保存到json

**requests\_get:**

获取response

New对象

response

response

url

url

**getnews:** 获取新闻url

main主函数

**getone:** 多线程提取新闻详细信息

main主函数

**getone:** 初始化、启动多线程detail

新闻url及label

json文件对象

**detail:** 获取response、从response提取新闻详细信息、新闻过滤和清洗、保存新闻