



山东大学

# 豆瓣电影小助手 设计报告

指导教师：李常刚

一组：王朝阳、邓颖岚、耿兆环、黄笑菡、孔湛琦

工作量占比： 王朝阳（40%） 邓颖岚（20%） 黄笑菡（20%）

孔湛琦（12%） 耿兆环（8%）

目录

一、项目介绍 ..... 2

    （一）项目背景 ..... 2

    （二）项目要求 ..... 2

    （三）项目目标 ..... 2

二、程序设计及其原理..... 3

    （一）界面展示 ..... 3

    （二）爬虫实现 ..... 5

    （三）界面 UI ..... 6

    （四）功能拓展构想 ..... 7

三、课程设计总结 ..... 8

# 一、项目介绍

## （一）项目背景

电影，作为现代流行的一种娱乐方式，拥有着巨大的市场需求。在面临电影院排片少，品类单一，票价高昂的现实的时候，人们往往选择线上观影，但线上观影也有不了解电影口碑，资源找不到或者观影要会员的问题。因此我们综合了各大视频网站与网盘搜索引擎，设计了“豆瓣电影小助手”项目。

## （二）项目要求

本项目需要使有线上观影需求的群众拥有独立的资源检索能力，并且给予影片筛选参考。

## （三）项目目标

项目的主要功能是为了帮助用户更为方便的找到查找电影资源，并且为用户提供了不同的观影网站链接以及云盘资源，全面的考虑了用户的不同需求，用户也可以根据自己的喜好来进行选择。此外，本项目的也为用户可以拥有更好的使用体验而增添了许多细节。比如用户在搜索电影时，可以根据排行榜及关键字和不同的电影类型、影片评分及评价人数来选择自己喜爱的电影，并且影片的获取数量也可以自定义。而且为了让使用者可以先提前了解影片的主要情况以免将时间浪费在自己不喜欢的电影上，通过双击该影片在这个程序的右侧就会出现有关该影片的海报、演员表、上映时间等影片详情，帮助用户有效的进行选择。

## 二、程序设计及其原理

### (一) 界面展示

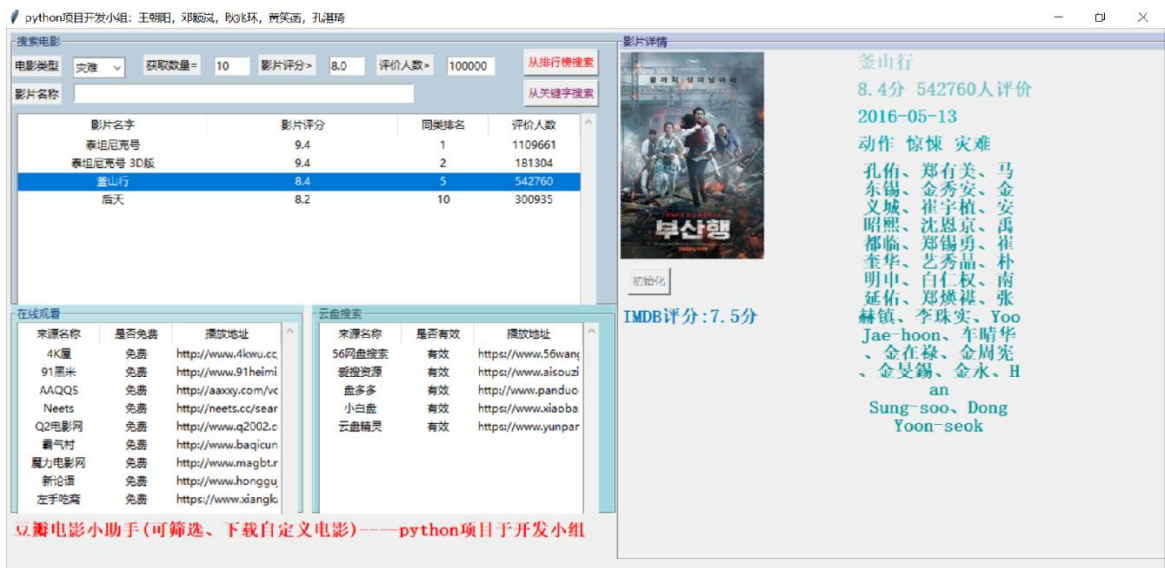


Figure 1 总体界面

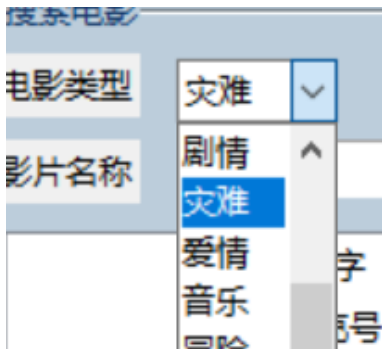


Figure 3 电影类型多样化

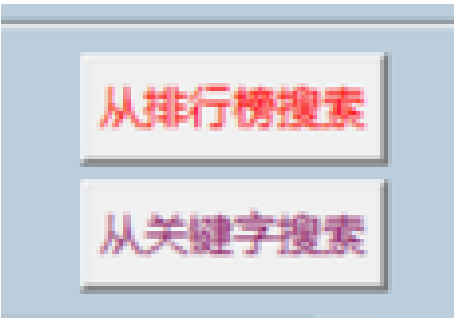


Figure 2 双重搜索方式

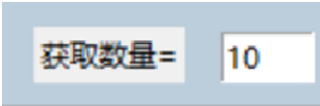


Figure 4 同时筛选多部电影 影片评分随心更改



云盘搜索		
来源名称	是否有效	播放地址
56网盘搜索	有效	https://www.56wan.com
爱搜资源	有效	https://www.aisouzi.com
盘多多	有效	http://www.panduo.com
小白盘	有效	https://www.xiaoba.com
云盘精灵	有效	https://www.yunpar.com

在线观看		
来源名称	是否免费	播放地址
4K屋	免费	http://www.4kwu.cc
91黑米	免费	http://www.91heimi.com
AAQQS	免费	http://aaxy.com/vc
Neets	免费	http://neets.cc/sear
Q2电影网	免费	http://www.q2002.com
霸气村	免费	http://www.baqicun.com
魔力电影网	免费	http://www.magbt.com
新论语	免费	http://www.honggu.com
左手吃藕	免费	https://www.xiangku.com

Figure 5 云盘搜索

Figure 6 在线观看



初始化

IMDB评分

泰坦尼克号 3D版

9.4分 181304人评价

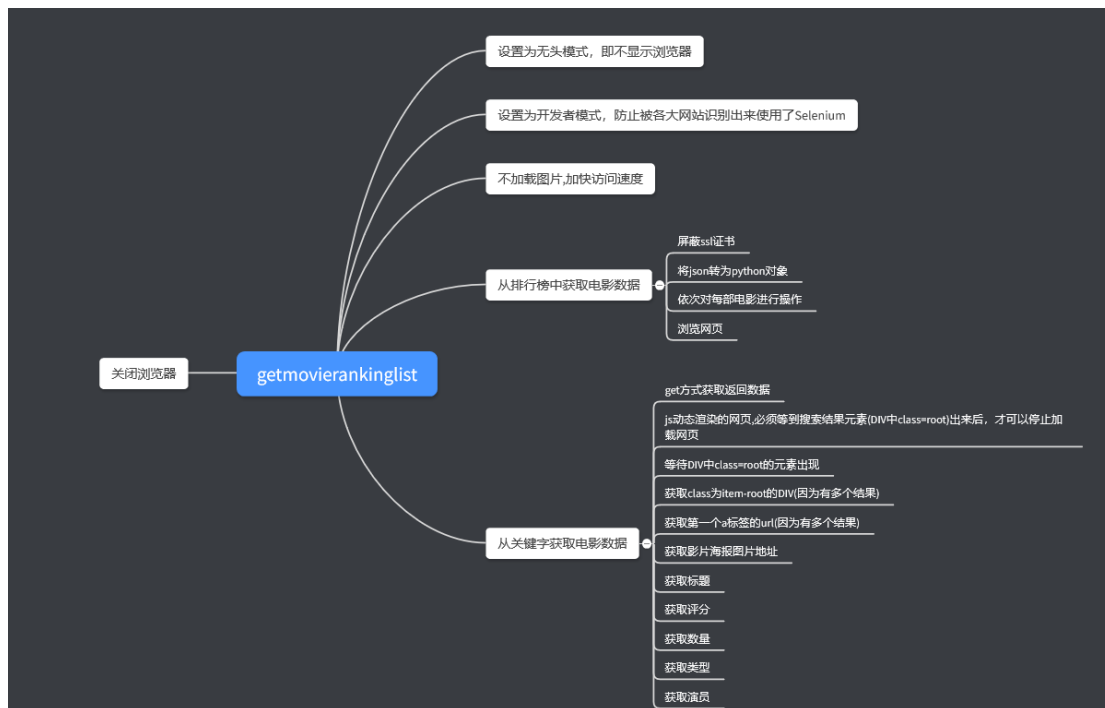
2012-04-10

剧情 爱情 灾难

莱昂纳多·迪卡普里奥、凯特·温丝莱特、比利·赞恩、凯西·贝茨、弗兰西丝·费舍、格劳瑞亚·斯图尔特、比尔·帕克斯顿、伯纳德·希尔、大卫·沃纳、维克多·加博、乔纳森·海德、苏茜·爱米斯、刘易斯·阿姆斯特朗、尼古拉斯·卡斯柯恩、阿那托利·萨加洛维奇、丹尼·努齐、杰森·贝瑞、伊万·斯图尔特、罗恩·多纳基、夏洛特·查顿、博纳德·福克斯、迈克尔·英

Figure 7 影片详情

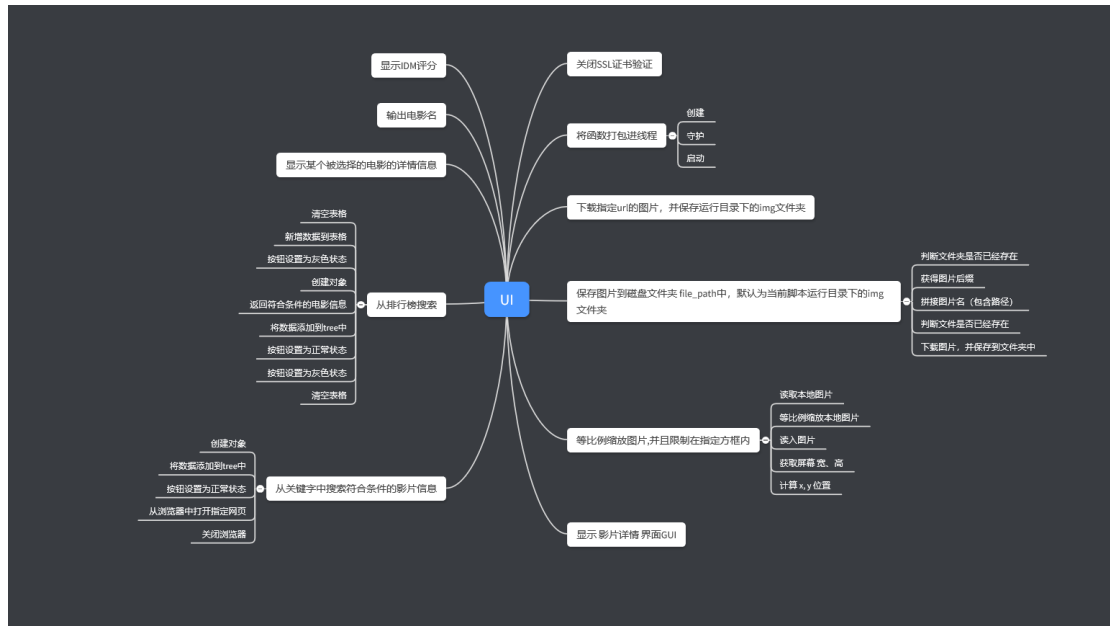
## (二) 爬虫实现



- 设置为无头模式，即不显示浏览器
- 设置为开发者模式，防止被各大网站识别出来使用了 Selenium
- 不加载图片,加快访问速度
- 从排行榜中获取电影数据
  - 屏蔽 ssl 证书
  - 将 json 转为 python 对象
  - 依次对每部电影进行操作
  - 浏览网页
- 从关键字获取电影数据
  - get 方式获取返回数据
  - js 动态渲染的网页,必须等到搜索结果元素(DIV 中 class=root)出来后，才可以停止加载网页
  - 等待 DIV 中 class=root 的元素出现
  - 获取 class 为 item-root 的 DIV(因为有多多个结果)
  - 获取第一个 a 标签的 url(因为有多多个结果)
  - 获取影片海报图片地址
  - 获取标题
  - 获取评分
  - 获取数量

- 获取类型
- 获取演员
- 关闭浏览器

### (三) 界面 UI



- 关闭 SSL 证书验证
- 将函数打包进线程
  - 创建
  - 守护
  - 启动
- 下载指定 url 的图片，并保存运行目录下的 img 文件夹
- 保存图片到磁盘文件夹 file\_path 中，默认为当前脚本运行目录下的 img 文件夹
  - 判断文件夹是否已经存在
  - 获得图片后缀
  - 拼接图片名（包含路径）
  - 判断文件是否已经存在
  - 下载图片，并保存到文件夹中
- 等比例缩放图片,并且限制在指定方框内
  - 读取本地图片
  - 等比例缩放本地图片

- 读入图片
- 获取屏幕 宽、高
- 计算 x,y 位置
- 显示 影片详情 界面 GUI
- 显示 IDM 评分
- 输出电影名
- 显示某个被选择的电影的详情信息
- 从排行榜搜索
  - 清空表格
  - 新增数据到表格
  - 按钮设置为灰色状态
  - 创建对象
  - 返回符合条件的电影信息
  - 将数据添加到 tree 中
  - 按钮设置为正常状态
  - 按钮设置为灰色状态
  - 清空表格
- 从关键字中搜索符合条件的影片信息
  - 创建对象
  - 将数据添加到 tree 中
  - 按钮设置为正常状态
  - 从浏览器中打开指定网页
  - 关闭浏览器

#### (四) 功能拓展构想

- 1.收集客户的观影记录，根据观影类型，口味偏好，影片热度，影片口碑，发布时间等进行个性推荐。
- 2.增加电影爱好者社区，增加观影体会，电影推荐，影片资讯，周边购买等模块。
- 3.增加检索类别，让客户可以根据参演明星，发布时间进行检索。同时，在关键字搜索方式下，根据其他已观看者的观影记录进行“猜你喜欢”推荐。
- 4.增加相关资源的检索，如 OST，高清壁纸等。
- 5.增加社交功能，根据观影口味偏好来推荐好友，先上线基础的聊天功能，之后可以上”一同观影“类别的线上线下的同时互动。



### 三、课程设计总结

通过不到两周的学习，老师确实让我们掌握了 python 的数据结构，函数与控制流，输入输出与模块方面的知识，让我们理解了什么是“面向对象的编程”。在上课的过程中，我们也能感受到了老师的用心。首先，老师对这节课有着明确的定位，了解不到两周时间对没有编程基础的同学们来说最多“入门”，因此教给我们的更多是如何学习一门语言的方法。再者，老师在教学过程中有很多的实践环节，给出的例子也是非常贴合学习内容且可操作性强的。还有课程最后的小组做项目这种形式也是非常不错的，我们既能锻炼自己的编程能力，也能真正体会到写程序，debug，调试程序，后续改进这一系列工作的不容易。相信老师课中的知识在之后的日子中，能成为我们的“利器”，帮助我们去解决未来生活中的一些问题，同时也让我们学会如何像一个程序员一样严谨地思考。

在小组项目过程中，虽然有各种的资料和获得讯息的途径，但是确认项目主题仍是比较艰难。首先，由于我们本身知识的局限，我们需要选择能运用我们学过的，或者短时间内能掌握的知识去做项目；其次，即使确定了项目主题，在短时间内做出来的成果恐怕也有较大的局限性，结果也许不尽人意（因此我们放弃了最开始做好的天猫爬虫）；再者，很多网站（如淘宝）都已经建立了强大的反爬机制，一群初学者要与这些网站“斗智斗勇”也实在不容易，反复的账号冻结很可能会浪费掉很多时间。因此，我们选择了豆瓣电影小助手作为我们的项目主题。

在确认主题后，我们对代码进行了反复的修改，调试，在 Github，CSDN 等网站上多有借鉴，最后界面的颜色我们也参考了设计师配色网站，力求在现有水平上做到更加完善。然而现在的算法并不是很理想，因此关键字搜索运行较慢，如果有更长的时间的话，我们应该会在算法优化上投入更多，这样“小助手”运行的速度将会大大提升。