CUCKOO

HELLO,
I AM CUCKOO...

Poetry in movie , dancing close to me.

中文 | English

中文

目录

项目名称

Cuckoo 电影数据可视化系统

运行环境

1. Python 3.5 及以上
2. Django 2.1.4 Web 应用框架框架
3. Scrapy 1.5.1 Web 抓取框架

文件架构

项目基于 Django 框架，前端网页设计以模板为基础修改、拼接，并实现了后台爬虫、功能实现。
项目结构如图：

前端发送请求的函数写在相应的 js 文件和 html 文件中，较为零散，重要内容参见“软件主要功能说明文档”文件。

附加功能

1. 除登陆界面外，每个界面都可以昼夜模式切换

2. 电影收藏功能：
   a) 收藏夹链接猫眼
   b) 除主页和登陆界面外，每个界面都可见收藏夹

3. 搜索功能：
   a) 模糊搜索
   b) 按照导演、演员、电影名多类型搜索
   c) 搜索结果链接猫眼

详情参见“软件附加功能说明文档”。

网站使用方式

1. 在 manage.py 同级文件夹下命令行输入“python manage.py runserver”。
2. 打开谷歌浏览器进入“http://[ip]:8000/cuckoo/index”，ip 为当前主机的 ip 地址。

HOME	返回主页并选择白天模式/夜晚模式
SERVICES	1. MOVIE TABLE 2. MOVIE TREND 3. MOVIE ACTORS 4. BOX OFFICE SHARE
MOVIE TABLE	按照票房顺序展示所筛选的电影
MOVIE TREND	展示 2015-2018 年内电影票房走势
MOVIE ACTORS	展示参与电影最多的演员或拍摄电影最多的导演
BOX OFFICE SHARE	展示不同类型电影所占的市场票房份额
Favorite	展示已收藏电影，且链接到猫眼
History	展示已打印的报表

测试文件使用方式

在 manage.py 同级文件夹下命令行输入“python tests.py”测试。

如图则为通过自动测试：

爬虫使用方式

1. 运行 webscrapy 下的 main.py 文件。
2. 结束后将新生成数据文件的替换 cuckoo 中的数据文件即可。

各模块实现简介

1. 系统登入功能：
   a) 全部基础功能
   b) 密码默认非明文显示，找回密码时有验证码（加分功能 1）
   c) 密码加密储存（加分功能 2）
   d) 有权限控制，未登录用户不能使用收藏、搜索和打印报表功能（加分功能 3）
   e) 有密码找回功能，用过邮箱验证许可修改为新密码（加分功能 4）
   f) 界面美观（加分功能 5）

2. 爬虫功能：
   a) 全部基本功能
   b) 可以由管理员操作更新数据（加分功能 1）
   c) 采用多线程多 ip 技术加快爬虫速度防止被封（加分功能 3）

3. 电影数据可视化部分：
   a) 全部基本功能
   b) 数据有动态交互能力（加分功能 1）
   c) 用户更换输入后从后端异步获取数据，生成新图并在前端逻辑实现（加分功能 2）
   d) 数据一次性读到全局内存，不需要每次请求都读文件（加分功能 3）
   e) 生成图标迅速（加分功能 4）

4. 电影数据报表部分：
   a) 所有基本功能
   b) 用户可以命名导出的数据报表（加分功能 1）
   c) 后端数据库存储图片 url 使处理高效（加分功能 2）
   d) 拥有历史图表功能，可以记录用户操作以及可以再次下载历史图片（加分功能 3）

English

Contend

Project Name

Cuckoo Movie Visualization System

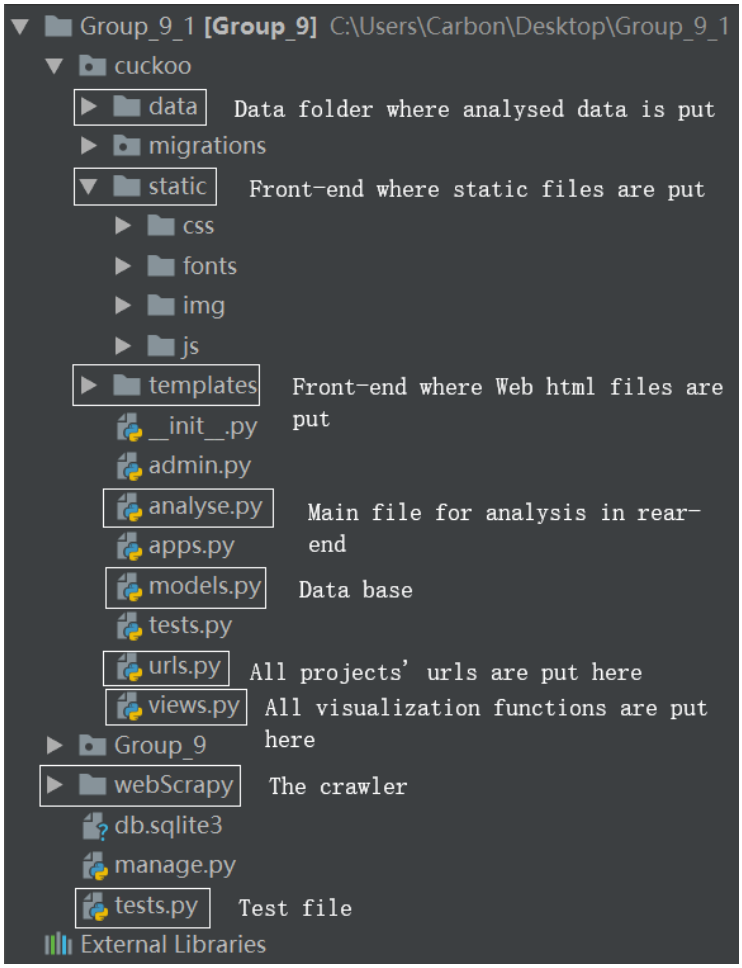
Operating Environment

1. Python 3.5 or above
2. Django 2.1.4 Web Application Framework
3. Scrapy 1.5.1 Web Crawler Framework

File Framework

The project is mainly based on Django Framework. The front-end is based on some Web models and we shaped them. We realized the crawler and data analysis function in the rear-end as well.

File framework is showed below:



The functions used for sending requests in the front-end are scatted in the corresponding "js" and "html" folders. Go to file "软件附加功能说明文档" to see the important details.

Additional Function

1. The switch between day mode and night mode
2. The adding favorite function
3. The fuzzy search function

See "软件附加功能说明文档" for more details.

Website Operation Method

1. Input "python manage.py runserver" in the command line at the same folder as file "manage.py".
2. Use Google Chrome to visit website "http://[ip]:8000/cuckoo/index" where ip is the ip address of your computer.

HOME	Return to home page and choose day mode/night mode
SERVICES	1. MOVIE TABLE 2. MOVIE TREND 3. MOVIE ACTORS 4. BOX OFFICE SHARE
MOVIE TABLE	Show selected movies in the sequence of box office
MOVIE TREND	Show the trend of box office during 2015-2018
MOVIE ACTORS	Show the actors/actresses that participated in the most movies or the directors that shot the most movies
BOX OFFICE SHARE	Show the shares occupied by different type of movies
Favorite	Show your favorite movies linked to Cat Eye
History	Show the report forms that have been printed

Test File Use-Pattern

Input "python tests.pyr" in the command line at the same folder as file "manage.py", if the result is the same as picture below then the automatic test is passed:

Crawler Use-pattern

1. Run the main.py file under the webscrapy folder.
2. Cover the file “cuckoo//data//movie_data” and “cuckoo//data//actor_data” with the new ones after the crawler finishes

Brief Introduction to Module Implementation

1. Log-In Function:
   a) All basic functions
   b) The password will not be displayed in clear text by default, and there is a verification code when the retrieving the password (bonus feature 1)
   c) Cryptographic encryption storage (bonus feature 1)
   d) User access restriction. Unlogged users cannot use the Favorite, Search and Print Function (bonus feature 3)
   e) Password retrieval function. Use mailbox authentication to create a new password (bonus feature 4)
   f) Beautiful interface (bonus feature 5)

2. Crawler Function:
   a) All basic functions
   b) Administrators are able to update the movie data (bonus feature 1)
   c) Using multi-threading and multi-ip technology to speed up the crawler and prevent it from being blocked (bonus feature 3)

3. Visualization Function:
   a) All basic functions
   b) Dynamic interaction of data (bonus feature 1)
   c) Asynchronous data acquisition from the back-end after the user replaces the input, generates new graphs and implements them in the front-end (bonus feature 2)
   d) Data is read into global memory once without requiring being read every time a request is made (bonus feature 3)
   e) Rapid graph generation (bonus feature 4)

4. Film Data Report Section:

a) All basic functions
b) Users are able to name the exported data reports (bonus feature 1)
c) Storing picture URLs in back-end databases to make processing efficiently (bonus feature 2)
d) Graphs are loaded into the History Function, where users can find their graphs and reprint them (bonus feature 3)