

网络空间安全

课程项目设计与实现文档



**课 程 脚本语言程序设计**

**项目题目 四川大学网络空间安全学院本科生登录系统**

**学生姓名 冯孙 简欣娅**

**学 号 2018141531089 2018141531058**

**指导教师 任德斌**

**项目时间 　 2019年秋**

## 项目背景

（选题原因和目的，5分）

Python是-种面向对象、直译式计算机程序设计语言，也是--种功能强大而完善的通用型语言，已经具有十多年的发展历史，成熟且稳定。这种语言具有非常简捷而清晰的语法特点，适合完成各种高层任务，几乎可以在所有的操作系统中运行。目前，基于这种语言的相关技术正在飞速的发展，用户数量急剧扩大，相关的资源非常多。

在本学期的学习过程中，先后学习了python 的基本语句、函数、模块、类以及异常，对python语言有了一个初步的了解，并对面向对象编程有了一个整体的认识。在这里，我们利用课堂上所学习的python知识，编写一个类似于川大本科生登录系统的网站。在**登录进入系统后，可以查看校园新闻信息，以及快速进入四川大学各个子网站**。

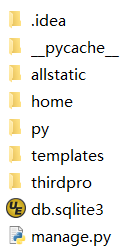
Django是高水准的**Python编程语言驱动**的一个开源模型．视图，控制器风格的Web应用程序框架，它起源于开源社区。使用这种架构，程序员可以方便、快捷地创建高品质、易维护、数据库驱动的应用程序。这也正是OpenStack的Horizon组件采用这种架构进行设计的主要原因。另外，在Dj ango框架中，还包含许多功能强大的第三方插件，使得Django具有较强的可扩展性

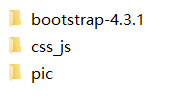
希望通过这个项目的设计和完成，巩固课堂所学知识，查漏补缺发现知识漏洞，将python装备为我们掌握的另一项实用武器。

## 项目的详细设计

（系统的主要结构，主要流程，接口设计，主要功能，核心代码，35分）

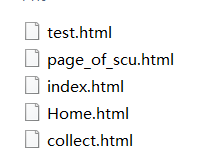
1.主要结构

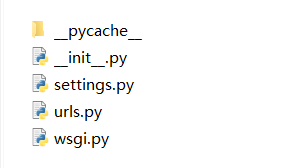


Allstatic是静态文件存放的地方，例如图片，音频，js，css文件等；

Home是一个app，就是一个模块（在大型网站中，很多app组成了整个网站的功能，在这里我们只用了一个app模块）；

Py是放.py文件的地方；

Template是放html文件的地方；

Thirdpro是网站的核心部分

Urls.py是负责url地址的分配

2.主要流程

（1）.利用pycharm生成项目

（2）.在thirdpro的url.py文件进行url地址分配

（3）.在home文件中的views.py进行逻辑处理

（4）.在thirdpro中的modules.py中设定数据接口

3.主要接口

主要是对html页面渲染的时候进行数据传输

和在注册界面对数据进行数据处理和数据储存

4.核心功能

1.模仿校级登录系统，设计一个院级登录系统。

2. 获取四川大学主页的新闻并储存起来，供用户可以查看以前的信息

3.设计注册功能，必须要注册才能登录

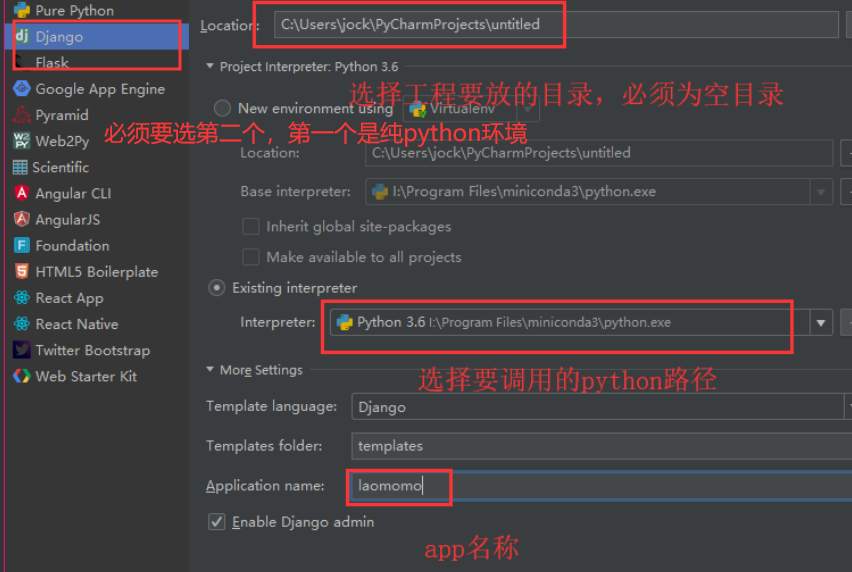
5.主要代码

**一、安装**

首先是利用python安装django模块， python pip install django;

然后安装pycharm profession(专业版)，专业版由于要收费只能参考网上教程进行破解，然后使用。

1.我们打开pycharm新建一个Django工程(也可以直接在terminal下用指令django-admin startproject xxx创建)

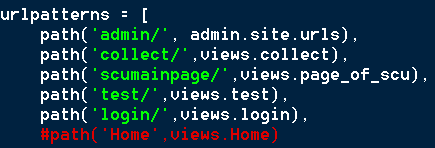




二、编写业务逻辑

新建成功工程后views.py文件里是空的，需要自己编写业务逻辑。

1.首先进行路由处理，这是在thirdpro文件夹下的urls.py处理。在本项目中我们有4个路由；分别是collect/;scumainpage/;test/;login/;



2.然后在app中（也就是home中）views.py进行视图函数处理；

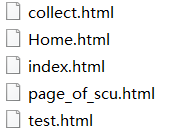
由于有4个路由，我们在views.py中就写了4个函数，分别来处理视图逻辑。



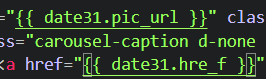
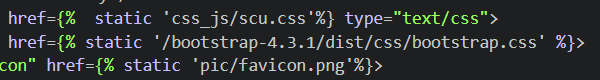
在视图函数中，用request.post[变量名]来获取客户端传输过来的post数据

request.GET.get(变量名)来获取客户端通过get请求传输的数据；然后下函数中编写对应的逻辑处理，最后通过return render（‘文件名’，json格式数据）返回给客户端。

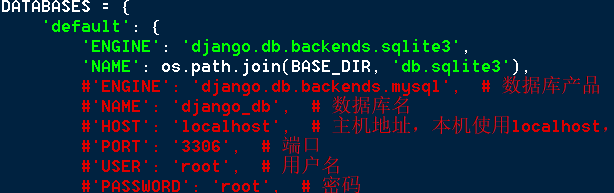
3.编辑html文件，本次项目是利用现成的web课程的html文件，因为我们重点是放在python语言编写上。在templates文件目录下是空的，我们需要在该目录下新建一个html文件来把内容展示到客户端。



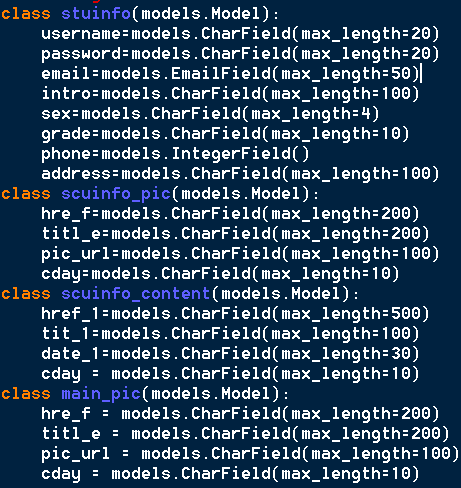
在html中获取服务端需要渲染的数据的格式是{{变量名}}，加载静态文件的格式是{%路径/文件名%}，同时还要设置，否则会找不到文件，



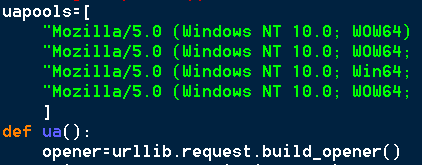
5.数据库处理。首先在thirdpro中的settings.py设置。本项目中只是用到了django自带的数据库sqlite3，如果要转换成mysql也可以。在django中安装相应的模块即可。



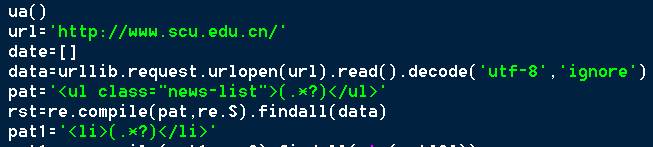
然后在app（home）中的moudles.py中设置相应的表即可，本项目用到了4张表。在class中声明表名，字段，类型等。Stuinfo储存的是学生的注册信息，分别是姓名，密码，邮箱，介绍，性别，年级，电话，地址；其他三张表示川大新闻视图的数据，以及如果需要扩展，直接添加表即可。



6.爬虫处理。在我们项目中需要使用到四川大学主页的新闻数据，于是就利 爬虫爬取了一些数据，用的是urlib包的request发送请求，对爬虫设计了一点简单的优化，那就是使用浏览器代理，利用urllib.request.build\_opener()函数来添加user-agent代理，



在进一步后，我们还打算添加ip代理，这样就能减少被封ip的几率。之后就是对爬取到的数据进行正则处理，由于技术不够成熟，我们进行了几次正则提取，最终才提取到。



然后再视图函数中先调用爬虫导入py文件调用函数，然后再将数据用for循环放到数组中，最后用models.表名.objects.get\_or\_create放入数据库中。









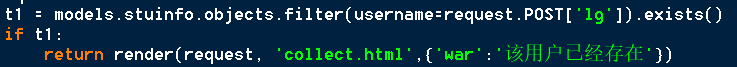
7.用户登录注册处理。登录时，利用request.POST[]得到数据后首先在数据库进行查找判断是否存在用户，不存在就返回“不存在”提示



如果存在就验证密码是否正确。不正确也返回提示。

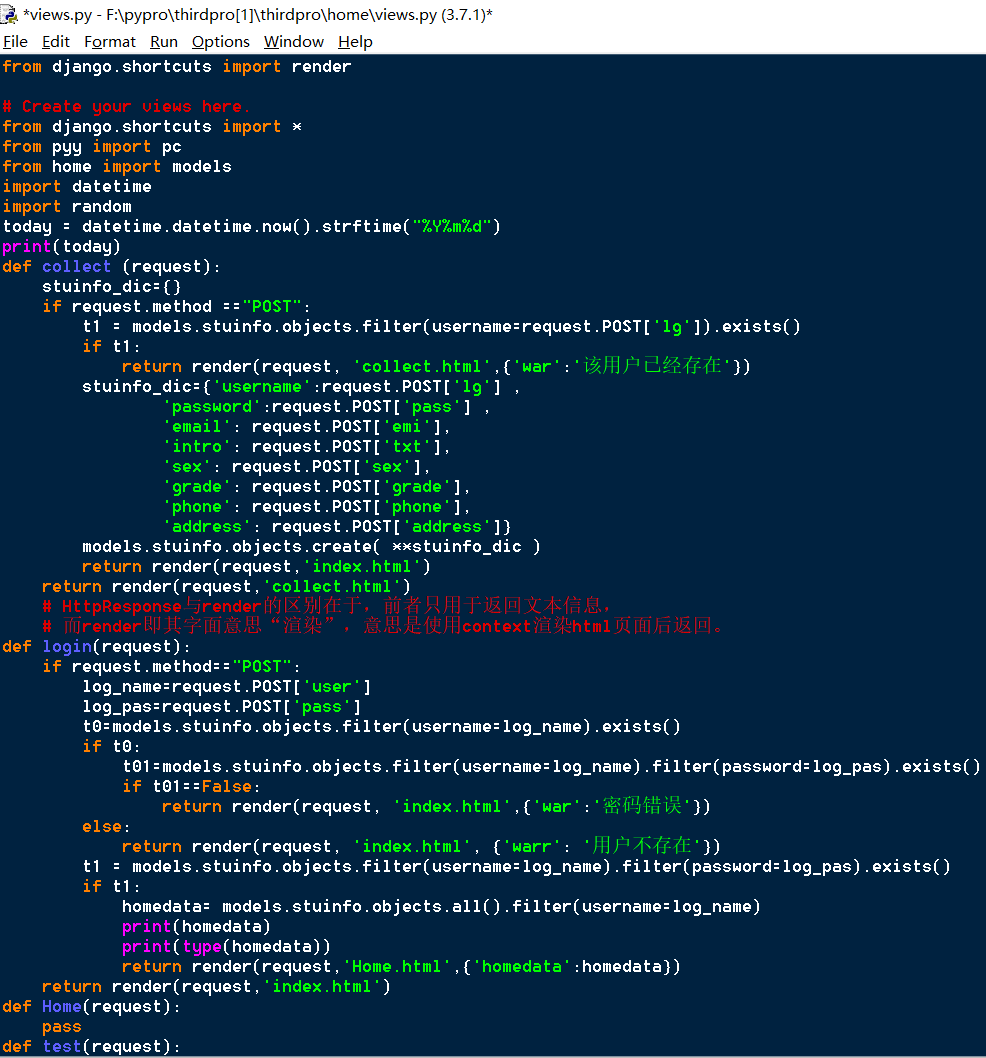


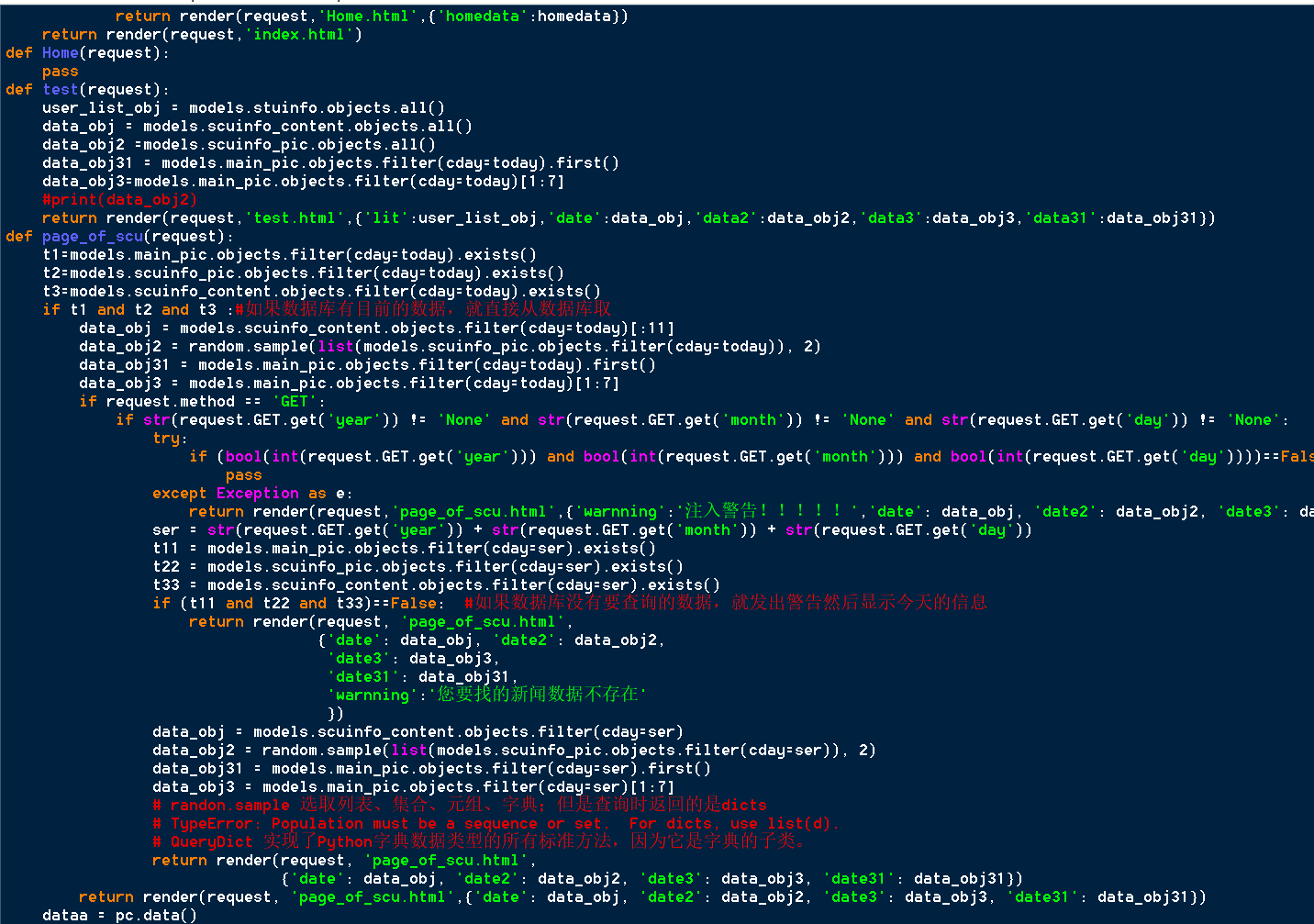
在注册时，先在js进行检验输入合法性，然后再服务端进行存储，存储前会判断是否用户被注册过，如果被注册过就会返回用户已存在。

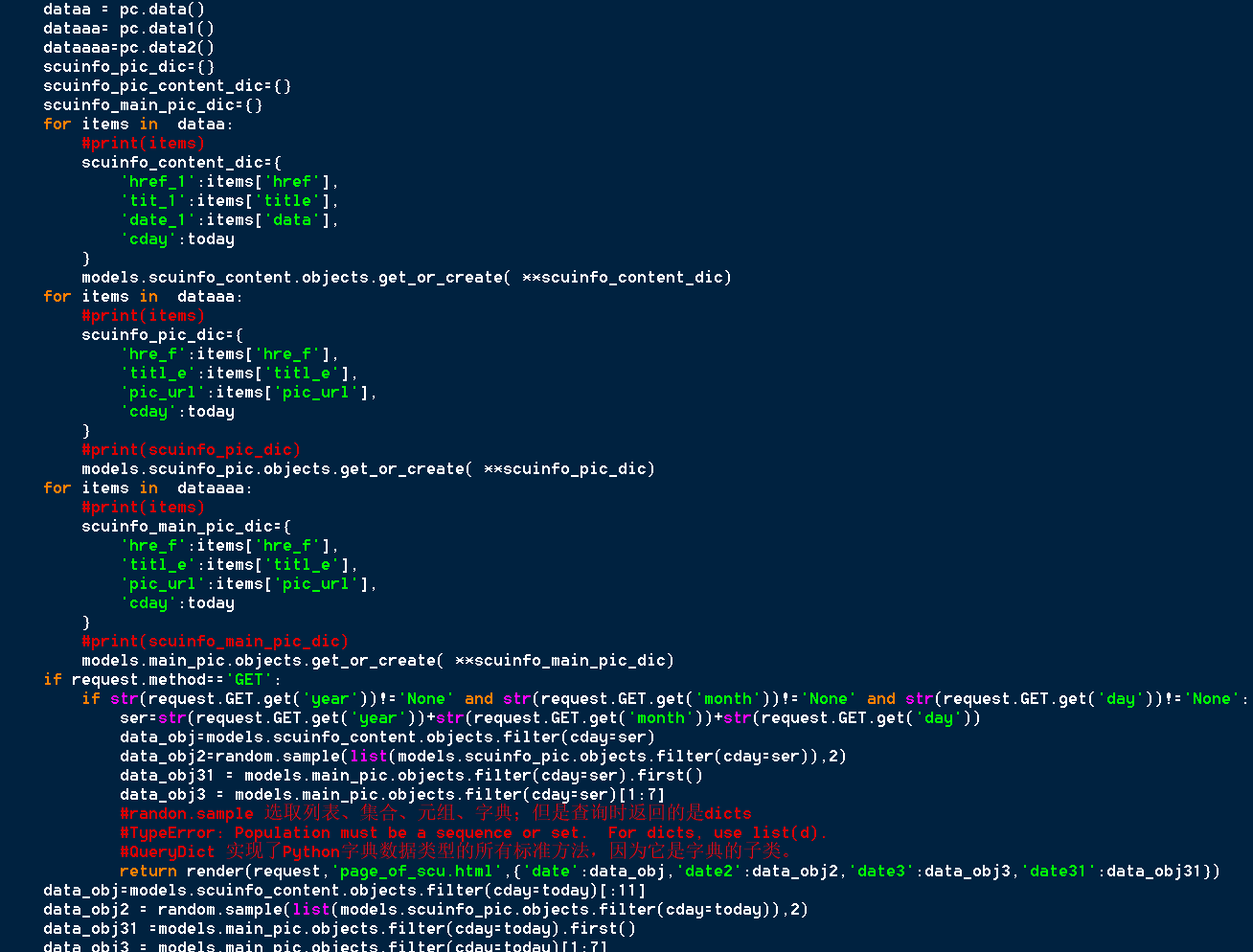


8.附上整个views.py和pc.py的代码（自己一行一行码出来的）

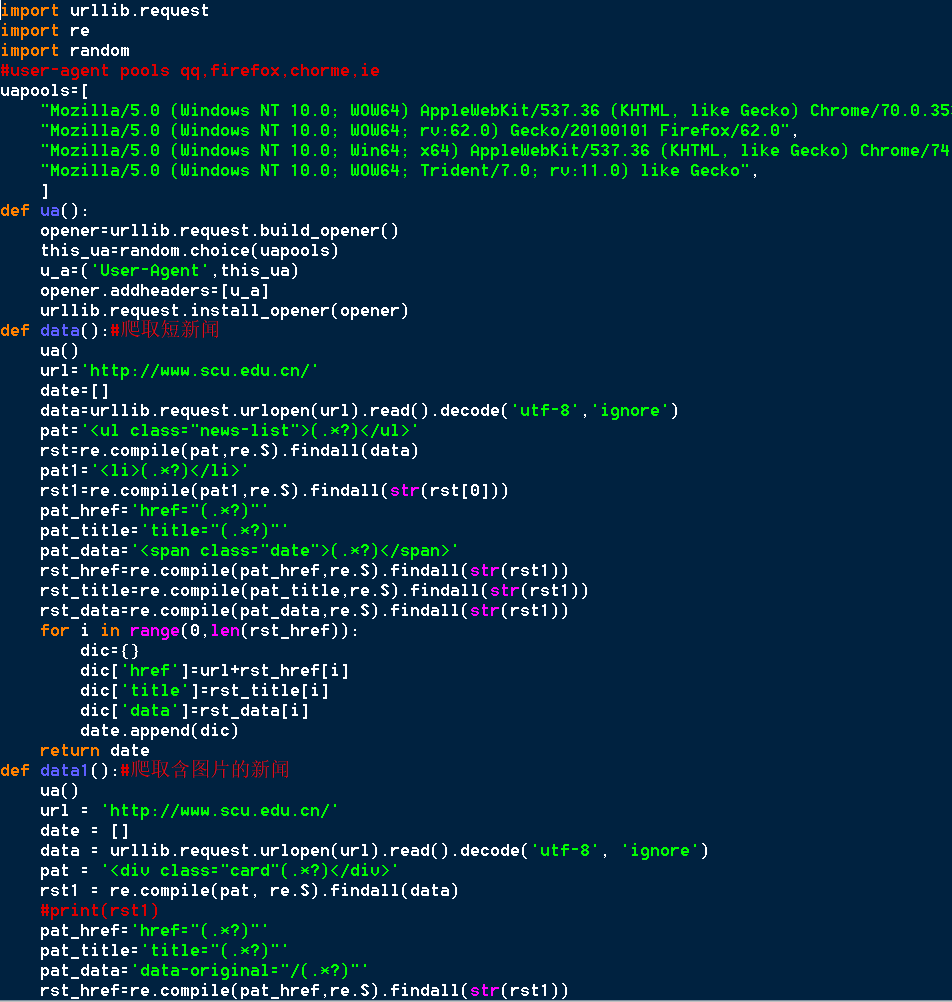
Views.py

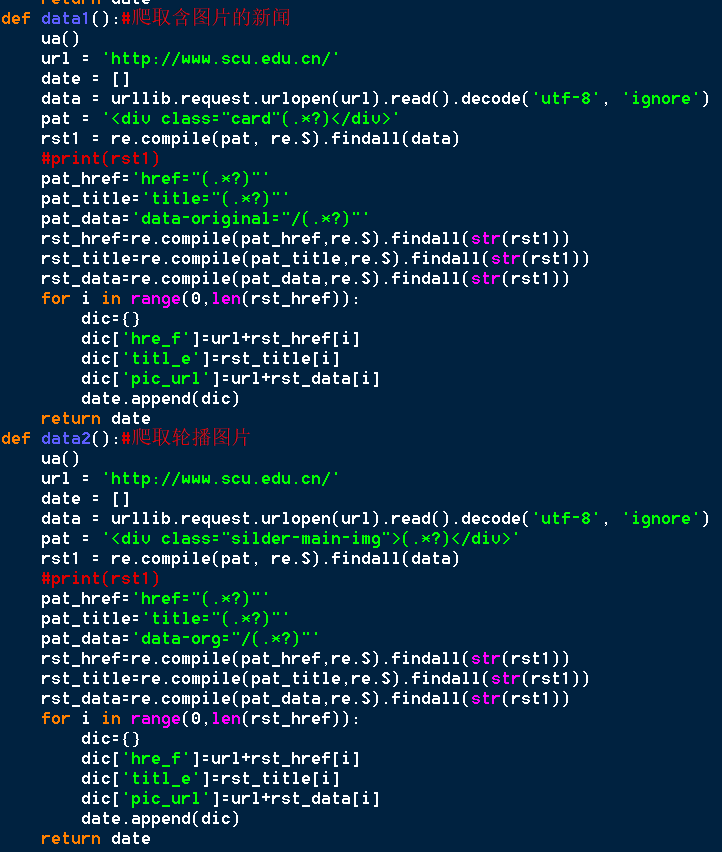






Pc.py





## 项目的分析与测试

（界面，测试用例，测试结果，结果分析，20分）

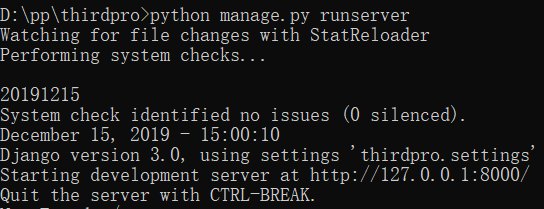


图1-启动服务

浏览器访问http://127.0.0.1:8000/login/



图2-登录界面演示



图3-信息注册页面（正则表达式密集使用区）

以下是一个注册信息的示范，在按照要求正确填写好信息格式之后方可注册成功，注册成功后自动跳回登录界面，现在可以通过刚才的注册信息登录进入系统，在主页显示的信息即为注册时上传的信息。





图4-注册成功演示

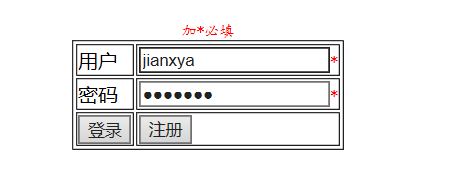




图5-个人主页信息展示

到此注册界面基本功能已展示完毕接下来是python网络爬虫功能的展示。

首先进入到个人主页后点击四川大学新闻概述可进入新闻详情界面，在该页

面可以通过日期查询最近时间段的新闻概况，从后台调取数据并显示在界面上



图6- 新闻详情界面



图7-查询12月6日新闻展示



图8-查询12月15日新闻展示

其他需要说明的事项：

所有后台的数据均存储在Django表单内，可调取查看

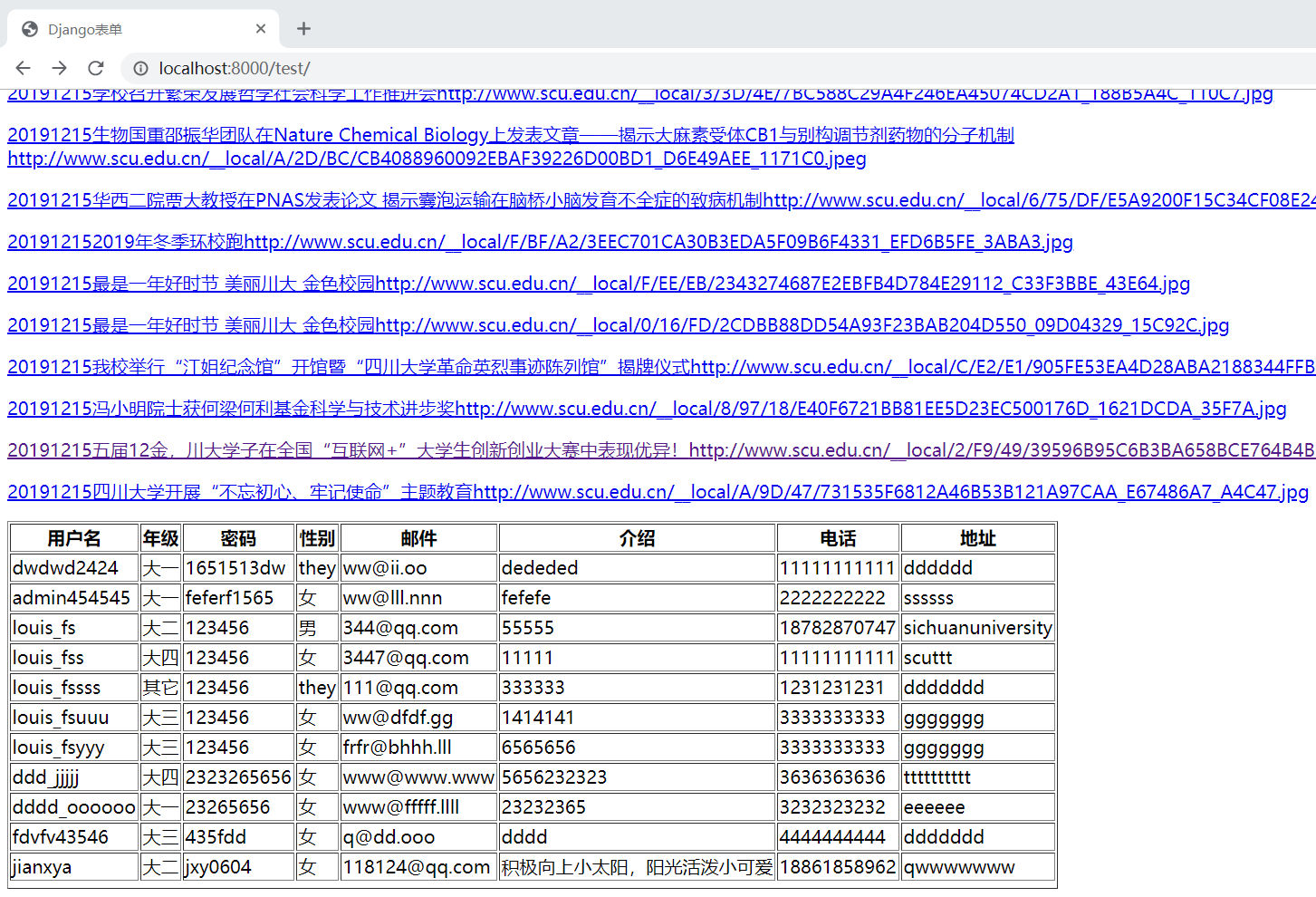


图9-Django表单信息展示

## 总结

（犯了什么错误，如何纠正的，自学了什么书籍或视频课程，是否达到预期目的，收获了什么，是否感受了团队精神等等，20分）

在这次实践中，我们主要碰到的困难有：

1.如今1.x盛行（绝大多数都是用django1.x写的），需要查更多资料，看官方文档

2.urls.py路由路径不熟悉，给加载picture，css等带来困难。主要是2.x和1.x方法不一样，前期看很多的文档才明白路由规则。

3.数据库语法需要逐渐了解，在pycharm中自带有数据库sqlite3，同样在django中数据库数据的增删改查是不一样的。在django中语法都是包装过了的，相比原生的数据库语句更加快捷方便和简单，在项目根目录下的pro文件下，修改moudles.py文件，构造对象即可创建表单。然后再命令行中进行manage.py makmigrations , mangae.py migrate即可。在django中也是可以使用mysql的，只需要在settings.py中修改即可。

4. Django是采用MVT的设计模式进行设计的。区别于MVC模式。MVC中，M指模型，V指视图（主要负责前端页面的处理），C指控制器（负责业务逻辑的处理），比如ThinkPHP采用的就是MVC架构。而在MVT架构中，M指模型，V指视图（负责业务逻辑处理），T指模板（负责前端界面）。因此，在进行实际开发时，应在视图中实现业务逻辑处理的程序，而前端应在模板中实现。

5.在数据传回前端时，遇到的麻烦比较多，主要是不知道数据的类型不清楚，同样在渲染页面时，也不是很容易。同时在提交时，由于网页比较简单，可能会提交空数据，或者反复提交，在限制这些主要是用了在数据库查询限制相同账户注册，同时利用js判断是否没有填入数据和检测是否有攻击手段，在表单提交设置onsubmit，true才允许提交。

6.本项目由于需要使用数据库，所以在四川大学官网首页爬取了一些资料，然后存储在库中，由于django刚好是python的一个模块，所以在每次请求scumainpage页面时，都会调用爬虫，爬取数据，来达到更新数据的目的。在写爬虫的时候首先先看了学习视频，然后参照别人的模板把框架搭好。最后针对川大的页面构造正则表达式，提取所需数据存入库中。

7.在scumainpage页面可以查询以前的新闻数据，就是在库中找数据，每次在存入数据时，我会在每条数据加一个字段date，该date是获取目前的时间，所以就能在确定数据的时间，查询也是根据时间进行的。

整个项目的完成过程中不得不感叹道，在程序员的进阶道路上，百度和谷歌真是小白的好伙伴，有什么不会的都可以通过搜索一些菜鸟教程学习到。我们主要的参考资料如下：



## 教师评语与打分

（课堂报告占20分，报告时间5-10分钟，要求有PPT）

（可以根据需要自行修改模板）

（打印时，请删除模板内所有括号里面的内容，末页至少空半页，以方便教师写评语）