

用户网络行为数据分析系统的设计与实现

🕒 Created @Mar 18, 2021 11:37 PM

开发、运行环境准备

Python + Django + MySQL

1. Python 3.8, 需要安装下列库, 一般直接 `pip install` + 包名即可
 - django 3.1+
 - jieba 中文分词
 - 其他如果有报错说缺什么, 就装什么
2. PyCharm 尽量用 Professional 版本, Community 也行
3. 网页 UI 需要的 css/js 都在项目文件中了, 不需要额外下载
4. MySQL 8.x 版本, 基本不需要对支持中文文本编码做修改, 默认安装即可, 安装路径不要有中文

项目结构

1. `data/sogou.Q.10.reduced` : 数据文件。这只是一个简单示例, 需要自己去 <https://www.sogou.com/labs/resource/q.php> 下载一天数据量的原始数据放在该文件夹下
2. `mysite` 这里面是网页项目运行、启动相关的配置, 一般不需要修改
3. `web_analysis` 中是网页主项目, 主要逻辑在下列文件中
 - `static` : 里面包含网页静态文件, css 样式、js 脚本、img 图片素材等
 - `templates` : 网页组织 HTML 文件模板, 基本一个 HTML 对应一个页面
 - `views.py` : 实现前后端数据交互
4. `data.sql` : 构建数据库的定义, 首先创建一个名为 web 的 schema, 前面几个 `create table` 语句, 定义了表结构, 后面的 `select` 是一些示例查询
5. `db_util.py` : 与数据库交互的相关操作, 数据插入、查询等

6. `manage.py`：网页项目的运行入口

运行方式

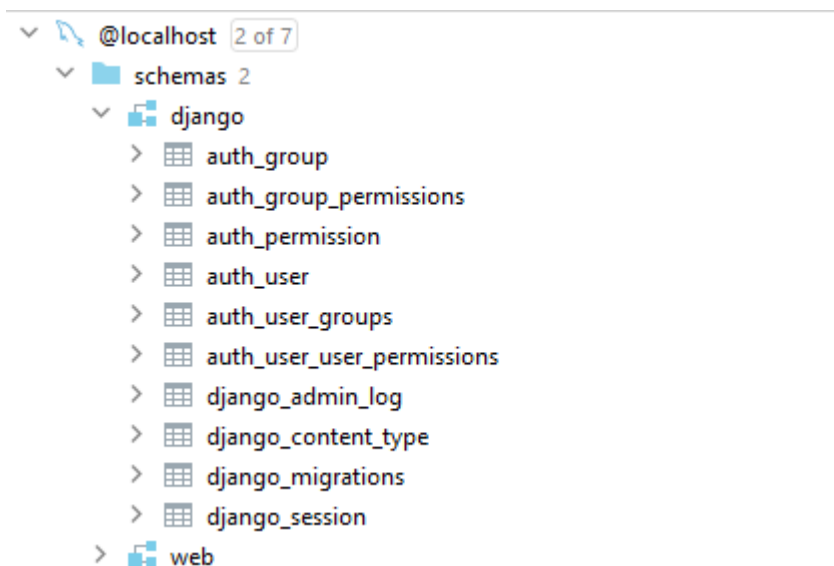
1. 安装好 MySQL，新建一个叫做 web 的 schema，并且使用 `data.sql` 创建好表结构。修改 `db_util.py` 里的 `user`、`password` 参数为对应的用户名和密码

```
def __init__(self):
    self.cnx = mysql.connector.connect(user='root', password='root', # 修改为自己的配置
                                       host='localhost',
                                       database='web')
```

2. 将数据下载、解压后放在 data 目录下，将 `db_util.py` 下面的 `main` 函数改成对应的文件路径和文件名，开始数据解析并保存如数据库的操作，然后运行该项目，直接在 PyCharm 中点击运行或命令行输入 `python db_util.py`，慢慢将原始数据提取出来，并且分别保存到数据库中。该过程较慢，大概要2-3小时。
3. 为用户管理建立用户表。需要在命令行中执行下面两条命令

```
python manage.py makemigrations
python manage.py migrate
```

随后在数据库中，会有一个新的 django schema，里面保存的是用户的用户名、密码等信息



4. 启动 Web 进行数据查询。直接在 PyCharm 中启动，或者命令行 `python manage.py runserver 8000`，看到如下输出表明启动成功，使用浏览器（谷

歌、火狐等，不要用 IE) 访问 <http://127.0.0.1:8000/>

```
Performing system checks...
```

```
Watching for file changes with StatReloader
```

```
System check identified no issues (0 silenced).
```

```
March 18, 2021 - 23:47:26
```

```
Django version 3.1.7, using settings 'mysite.settings'
```

```
Starting development server at http://127.0.0.1:8000/
```

```
Quit the server with CTRL-BREAK.
```