

2022运维寒假学习计划

面向对象

运维21级同学

学习内容

1. python
2. flask

学习目标

1. 学习python基本语法
2. 学习python面向对象编程
3. 学习flask框架
4. 使用python爬取华大新闻第一页所有新闻的标题与链接 (即要求对爬取到的数据进行处理)
5. 使用flask将爬取的结果展示到自己的站点上并对新闻数据实时更新

具体要求

1. python语法是基础要求熟练使用
2. 爬虫不限制过滤html的工具和方式,可以使用beautifulsoup库也可以使用正则,亦或是其他工具
3. flask站点美观程度不影响评分
4. 使用git对项目进行管理,并添加README.md和.gitignore文件,不限制github,gitee或自搭git (但应保证验收时可以访问
5. 使用nginx反向代理flask至{IP或域名}/hdxw
6. 使用nginx配置{IP或域名}/为静态显示这是华大新闻的站点
具体表现为使用静态映射一个index.html文件,并且可以超链接至新闻数据页,同时本目录下应包含一个相同内容的index.md,并配置该文件不可访问
以下为参考html

```
<html lang="en">
<head>
  <meta charset="UTF-8">
  <meta http-equiv="X-UA-Compatible" content="IE=edge">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <title>hdxw</title>
</head>
<body>
  <a href="./hdxw">这是华大新闻的站点</a>
</body>
</html>
```

7. 将以上操作实现,并写成文档
8. 截至日期: 1643544000 这是一个Unix timestamp,届时验收,遇到问题及时沟通

参考链接

1. [python](#)
2. [flask](#)
3. [beautifulsoup](#)
4. [正则](#)
5. [图形化正则](#)
6. [华大新闻](#)