

网络爬虫 实验报告

姓名： 蔡欣彤

日期： 2023. 7. 8

一、 实验目的

1. 熟悉网络爬虫的工作原理
2. 熟悉网络爬虫的基础语法
3. 能够编写一个简单的网络爬虫

二、 实验内容

实验要求：

1. 掌握 requests 模块的安装使用；
2. 掌握 BeautifulSoup 模块的安装使用；
3. 掌握 html 的基本知识；

三、 实验结果

完成实验并截图说明每个步骤

1. 安装 requests、BeautifulSoup4

```
PS C:\Users\cqcxxt\PycharmProjects\Webpachong> pip list
Package                Version
-----
beautifulsoup4         4.12.2
certifi                 2023.5.7
charset-normalizer      3.2.0
idna                    3.4
pip                     23.1.2
requests                2.31.0
setuptools              68.0.0
soupsieve               2.4.1
urllib3                 2.0.3
wheel                   0.40.0
PS C:\Users\cqcxxt\PycharmProjects\Webpachong>
```

代码：（使用 mysql 存储）

```
import psycopg2 as psycopg2
import pymysql as pymysql
import requests
from bs4 import BeautifulSoup
```

```

import sqlite3

def wirte_txt(txt):
    with open('news.txt', "w", encoding='utf-8') as file:
        file.write(txt)

def wirte_sqlite(sql):
    conn = sqlite3.connect('news.db')
    cursor=conn.cursor()
    cursor.execute(sql)
    conn.commit()
    cursor.close()

def res(url):
    headers={
        'user-agent': "Mozilla/5.0 (Windows NT 10.0; Win64; x64)
AppleWebKit/537.36 (KHTML, like Gecko) Chrome/113.0.0.0 Safari/537.36
Edg/113.0.1774.50"
    }
    res = requests.get(url, headers=headers)
    res.encoding = res.apparent_encoding
    return res.text

conn = pymysql.connect(host='localhost', user='root',
password='123456xx', db='news')
#
conn=psycopg2.connect(database="db_2020_01",user="db_user2020_271",pa
ssword="db_user@123",host="116.205.157.173",port=8000)
b = conn.cursor()

url = 'https://news.cqu.edu.cn/newsv2/news-130.html'
txt = res(url)
# print(txt)
soup = BeautifulSoup(txt, 'lxml')
lst = soup.find('div', class_="lists")
itm = lst.find_all('div', class_="item")
# page=soup.find('div', class_="page").find_all('a')[-1]['href']
# print(page[-1]['href'])
# i=itm[0]
num=5
while num!=0:
    # print(0)
    num=num-1
    for i in itm:

```


航空航天学院召开2024届本科生就业动员大会

2023-06-19

6月16日晚，航空航天学院在重庆大学A区研究生楼306召开了2024届本科生就业动员大会。航空航天学院2020级全体本科生参与，辅导员陈勇凡主持。

6月16日晚，航空航天学院在重庆大学A区研究生楼306召开了2024届本科生就业动员大会。航空航天学院2020级全体本科生参与，辅导员陈勇凡主持。大会主要聚焦于“考研”与“就业”两个主题，陈勇凡向同学们抛出两个问题“是否准备考研”与“是否确定目标院校”。结合往届学生的就业情况为同学们的学习规划和职业规划提出宝贵建议。对于“就业”，陈勇凡老师也提出了几点建议。建议计划毕业后参加工作的同学，向心仪单位大胆投递简历，大胆尝试；建议想“考公”的同学，认真复习，明确目标，抱着为国家发展与人民幸福生活服务的理想从事公务员工作。会上，陈勇凡还要求所有同学在规定时间内准备好自己的个人简历，学院还将安排专场讲座介绍简历设计和修改。此次还邀请了航空航天学院18级本科生，现22级研究生吴昱霖为同学们分享考研经验。吴昱霖学长从考研复习的重点出发，从“数学”、“英语”到“政治”、“专业课”，建议同学们认真准备考试内容，特别是数学要重视基础知识的复习。吴昱霖学长也向同学们强调了报考本学院研究生所具有的优势。在本次大会上，各班级的就业委员也正式与同学们见面。就业委员在接下来一年会为班级同学服务，助力同学们顺利毕业、尽早落实去向。

经管学院承办高远杯模拟招聘重庆大学校赛决赛

2023-05-19

5月19日，重庆大学经管学院团委会承办西南财经大学承办的高远杯·四川省第十六届模拟招聘大赛决赛举行。决赛邀请了职就中心市场科科长向章婷，经管学院党委副书记张燕、王毅文和向晨西老师作为评审专家。

5月19日，重庆大学经管学院团委会承办西南财经大学承办的高远杯·四川省第十六届模拟招聘大赛决赛举行。决赛邀请了职就中心市场科科长向章婷，经管学院党委副书记张燕、王毅文和向晨西老师作为评审专家。决赛赛制共分为两轮，第一轮为个人风采展示，第二轮为情景模拟能力测试。在第一环节中，选手按赛前抽签顺序上台进行自我介绍，评委结合选手的个人简历及现场展示进行提问。在第二个环节中，选手进行情景模拟能力测试，现场抽题回答问题，选手们在整个过程中沉着冷静，认真审题，展现出很好的就业素养。经过激烈的角逐，推选出2名选手代表学校参加下周的川渝高校就业模拟总决赛。本届模拟招聘大赛结合数智化时代招聘趋势，通过模拟职场竞聘过程，为参赛者带来直观的求职体验，增强了自立自信、树立正确择业观、提高综合素质和核心能力。通过成渝高校共同打造的就业创业平台，推动两地高校学子更加紧密交流合作。

资源与安全学院开展“访企拓岗促就业”专项活动

2023-05-16

5月11-12日，重庆大学资源与安全学院开展了“访企拓岗促就业”专项活动，走进中国建筑五局华南公司、上海隧道工程有限公司广东分公司。

江林介绍了近期线上校园招聘概况及积极筹备开展线下招聘会事宜。会议邀请沙坪坝区团委书记黄艳宣讲“大学生志愿服务西部计划”；邀请重庆大学党委学生工作部副部长蒋研川结合学校实际案例，解读2020年大学生参军入伍政策。我校2020届毕业生通过腾讯会议在线参与了西部计划和征兵政策的政策讲解。

加州大学戴维斯分校法学院联合培养硕士研究生项目线上招生宣讲会顺利举行

2020-05-16

2020年5月16日上午，我院与UC Davis法学院顺利开展了“法学硕士双学位联合培养项目”线上招生宣讲会。

为了培养经济全球化背景下具有国际视野、通晓国际规则，能够参与国际事务与国际竞争的国际化人才，重庆大学法学院与加州大学戴维斯分校（UC Davis）法学院合作共同联合培养LL.M.硕士研究生项目正在稳步推进中。为集中答疑解惑、加强学子对该项海外留学项目的深入理解，2020年5月16日上午，我院与UC Davis法学院顺利开展了“法学硕士双学位联合培养项目”线上招生宣讲会。UC Davis是位于美国 萨克拉门托西部的一所世界顶尖的研究型大学，隶属于著名的加州大学系统，全球排名前50。宣讲会主要通过LL.M.硕士研究生项目的介绍、已参与该项目的学生分享向学子们展现UC Davis法学院的点点滴滴。宣讲会开幕，学院党委副书记刘淳结合自身国外访学经历的体验，分析了留学项目对学业发展及人生成长的重要意义，也表示学院会为学子们拓展更多的出国交流和学习的渠道。UC Davis法学院国际招生官Adriana Agüena老师详细进行了UC Davis法学院LL.M.联合培养硕士研究生项目介绍，包括录取标准、课程选择、教授师留、社团生活、学院和社区支持、以及奖学金申请等具体情况及注意事项。强调参加联合培养项目的学生，可在攻读本校学位的同时，多获得一个UC Davis LL.M.法学硕士项目，并拥有申请UC Davis奖学金、以及加州政府实习机会。回归留学主体，为使内容充分贴合真实感悟，正在UC Davis的重庆大学法学院学生孙增恒分享了自己的学习与生活体验。她不仅切实描述了在美国的生活的饮食住行、课程安排，还鼓励同学们的 能积极参与到这段充满意义的美国求学旅程中，坚定自我选择，丰富个人经历。UC Davis负责国际合作事务的副院长Beth Greenwood女士则表达了对合作共同联合培养LL.M.硕士研究生项目的支持，并表示UC Davis法学院将怀揣饱满的热情地欢迎中国学生们赴美就读。在全球疫情防控的特殊形势下，学院正稳步推进海外游学与国际合作交流工作，为正在海外或未来有计划赴海外交流的学子做好服务，努力实现法治人才国际化的培养目标。也希望同学们能正确规划自己的未来，勇敢做出选择，利用留学阶段提升眼界，为成为高素质、全方位、国际化的优秀法治人才不断增长学识，奋发向前。

四、实验总结

（在实验中遇到的问题及解决方法、收获是什么）

安装 BeautifulSoup 遇到问题，改为安装 BeautifulSoup4

```
PS C:\Users\cqquxt\PycharmProjects\Webpachong> pip install beautifulsoup
Collecting beautifulsoup
  Using cached BeautifulSoup-3.2.2.tar.gz (32 kB)
  Preparing metadata (setup.py) ... error
error: subprocess-exited-with-error

× python setup.py egg_info did not run successfully.
  | exit code: 1
  |_ [7 lines of output]
Traceback (most recent call last):
  File "<string>", line 2, in <module>
  File "<pip-setuptools-caller>", line 34, in <module>
  File "C:\Temp\pip-install-7nor1x2l\beautifulsoup_b598cc55a8b147949c85ced4d46028eb\setup.py", line 3
    "You're trying to run a very old release of Beautiful Soup under Python 3. This will not work."<>"Please use Beautiful Soup 4, av
, "

SyntaxError: invalid syntax
[end of output]

note: This error originates from a subprocess, and is likely not a problem with pip.
error: metadata-generation-failed

× Encountered error while generating package metadata.
  |_ See above for output.

note: This is an issue with the package mentioned above, not pip.
hint: See above for details.
```