

# 11-让爬虫定时向你回报

## 一.定时爬取

我们可以是用schedule库实现定时功能,schedule,就是用time和datetime来实现的。

因为schedule是第三方库,所以使用之前我们需要使用pip install进行安装

实操代码:

```
1 #导入schedule模块和time模块
2 import schedule, time
3 #定义一个job函数
4 def job():
5     print("Working in progress...")
6 #部署
7 schedule.every(10).minutes.do(job) #部署每10分钟执行一次job()函数的任务
8 schedule.every().hour.do(job) #部署每x小时执行一次job()函数的任务
9 schedule.every().day.at("10:30").do(job) #部署在每天的10:30执行job()函数的任务
10
11 schedule.every().monday.do(job) #部署每个星期一执行job()函数的任务
12
13 schedule.every().wednesday.at("13:15").do(job)#部署每周三的13:15执行函数的任务
14
15 while True:
16     schedule.run_pending()
17     time.sleep(1)
18 #检查部署的情况,如果任务准备就绪,就开始执行任务
```

## 二.邮箱发送

我们可以是用smtplib和email库实现发送邮件功能

代码实操

```

1 import smtplib
2 from email.mime.text import MIMEText
3 from email.header import Header
4 #引入smtplib、MIMEText和Header
5 mailhost='smtp.qq.com'
6
7 #把qq邮箱的服务器地址赋值到变量mailhost上，地址应为字符串格式
8 qqmail = smtplib.SMTP()
9
10 #实例化一个smtplib模块里的SMTP类的对象，这样就可以调用SMTP对象的方法和属性了
11 qqmail.connect(mailhost,25)
12 #连接服务器，第一个参数是服务器地址，第二个参数是SMTP端口号。
13 #以上，皆为连接服务器。
14 sender = input('请输入你的邮箱: ')
15 #获取邮箱账号，为字符串格式
16 password = input('请输入你的密码: ')
17 #获取邮箱密码，为字符串格式
18 qqmail.login(sender,password)
19 #登录邮箱，第一个参数为邮箱账号，第二个参数为邮箱密码
20 #以上，皆为登录邮箱。
21 receiver=input('请输入收件人的邮箱: ')
22 #获取收件人的邮箱。
23 content=input('请输入邮件正文: ')
24 #输入你的邮件正文，为字符串格式
25 message = MIMEText(content, 'plain', 'utf-8')
26 #实例化一个MIMEText邮件对象，该对象需要写进三个参数，分别是邮件正文，文本格式和编码
27 subject = input('请输入你的邮件主题: ')
28 #输入你的邮件主题，为字符串格式
29 message['Subject'] = Header(subject, 'utf-8')
30 #在等号的右边是实例化了一个Header邮件头对象，该对象需要写入两个参数，分别是邮件主题和编码，然后赋值给等号左边的变量message['Subject']。
31 #以上，为填写主题和正文。
32 try:
33     qqmail.sendmail(sender, receiver, message.as_string())
34     print ('邮件发送成功')
35 except:
36     print ('邮件发送失败')
37 qqmail.quit()
38 #以上为发送邮件和退出邮箱

```

### 三.定时发送邮箱

```

1 import requests
2 import smtplib
3 import schedule
4 import time
5 from bs4 import BeautifulSoup
6 from email.mime.text import MIMEText
7 from email.header import Header

```

```

8  #导入我们需要的包
9  #封装获取天气代码
10
11 def get_weather():
12     #设置请求头
13     headers = {'user-agent': 'Mozilla/5.0 (Windows NT 6.1; Win64; x64)
AppleWebKit/537.36 (KHTML, like Gecko) Chrome/78.0.3904.87 Safari/537.36'}
14     #设置url
15     url = 'http://www.weather.com.cn/weather/101020100.shtml'
16     #发送请求
17     res = requests.get(url, headers=headers)
18     #设置编码
19     res.encoding = 'utf-8'
20     #通过BeautifulSoup解析数据
21     bsdata = BeautifulSoup(res.text, 'html.parser')
22     #查找标签为tem的数据
23     data_temperature = bsdata.find(class_='tem')
24     #查找标签为wea的值
25     data_weather = bsdata.find(class_='wea')
26     return data_temperature, data_weather
27
28 #设置函数封装发送邮件
29 def send_email(tem, wea, sender, pwd, receiver):
30     #把qq邮箱的服务器地址赋值到变量mailhost上，地址应为字符串格式
31     mailhost = 'smtp.qq.com'
32     #实例化一个smtplib模块里的SMTP类的对象，这样就可以调用SMTP对象的方法和属性了
33     qqmail = smtplib.SMTP()
34     #连接服务器，第一个参数是服务器地址，第二个参数是SMTP端口号。
35     qqmail.connect(mailhost, 25)
36     #登陆邮箱
37     qqmail.login(sender, pwd)
38     #邮箱正文
39     content = '亲爱的，今天的天气是：'+tem+wea
40     #实例化一个MIMEText邮件对象，该对象需要写进三个参数，分别是邮件正文，文本格式和
编码
41     message = MIMEText(content, 'plain', 'utf-8')
42     subject = '今日天气预报'
43     message['Subject'] = Header(subject, 'utf-8')
44     try:
45         #发送邮件
46         qqmail.sendmail(sender, receiver, message.as_string())
47         return True
48     except:
49         return False
50     #关闭邮箱
51     qqmail.quit()
52
53 #封装获取天气发送
54 def job():
55     print('开始一次发送任务')
56     #获取天气
57     tem, wea = get_weather()

```

```
58     IsSuccess = send_email(tem,wea,sender,pwd,recevier) # 这里需要设置发件人
    的账号密码以及收件人的账号
59     while IsSuccess is False:
60         print("邮件发送失败，正在尝试重新发送")
61         IsSuccess = send_email(tem, wea, sender, pwd, recevier)
62         print('任务完成')
63 #设置每天7点30执行job函数
64 schedule.every().day.at("07:30").do(job)
65 while True:
66     schedule.run_pending()
67     time.sleep(1)
68 #检查部署的情况，如果任务准备就绪，就开始执行任务
```