

[python应用案例] 一.BeautifulSoup爬取天气信息并发送至QQ邮箱

原创 Eastmount 最后发布于2018-05-03 15:28:27 阅读数 4117 ☆ 收藏

展开



Python+TensorFlow人工智能

该专栏为人工智能入门专栏，采用Python3和TensorFlow实现人工智能相关算法。前期介绍安装流程、基础语法...

Eastmount

¥9.90

去订阅

前面作者写了很多Python系列文章，包括：

- Python基础知识系列：[Python基础知识学习与提升](#)
- Python网络爬虫系列：[Python爬虫之Selenium+Phantomjs+CasperJS](#)
- Python数据分析系列：[知识图谱、web数据挖掘及NLP](#)



Python学习系列

文章：16篇
阅读：119908



Python爬虫之Selenium+Phantomjs+CasperJS

文章：33篇
阅读：443874



知识图谱、web数据挖掘及NLP

文章：44篇
阅读：488758

接下来作者将学习并讲解一些Python小应用，它将结合Python爬虫、数据分析、web开发或其他功能进行介绍。一方面希望能提升读者的Python学习兴趣，另一方面也希望读者能学到些知识，做些好玩的应用。本篇文章主要讲解BeautifulSoup爬取每日天气信息，然后将信息发送至QQ邮箱，其难点是如何配置QQ邮箱发送Python邮件。基础性应用文章，希望对您有所帮助，如果文章中出现错误或不足之处，还请海涵~

一. BeautifulSoup爬取天气信息

1.分析网页

中国天气网: <http://www.weather.com.cn/weather/101260101.shtml>

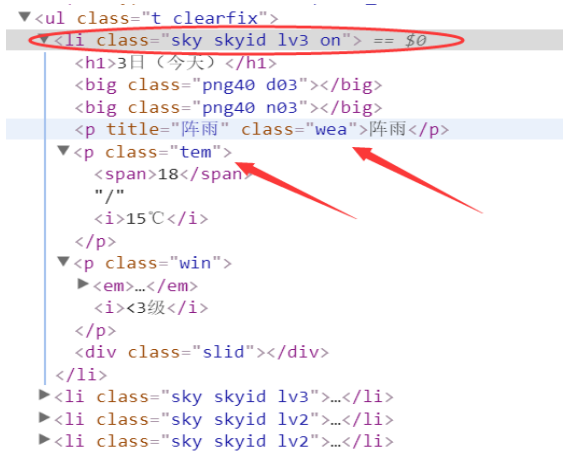
我们需要爬取贵阳市当天的天气信息，比如“5月3日 阵雨 18/15°C”



接下来通过浏览器审查元素，可以看到这些信息位于<li class="sky skyid lv3 on">元素下，接着我定义class为“wea”和“tem”的元

素进行定位，核心代码为：

```
name = soup.find_all(attrs={"class":"sky skyid lv3 on"})
```



2.完整代码

```
# -*- coding: UTF-8 -*-
import urllib
import urllib.request
from bs4 import BeautifulSoup

# 下载数据
url = "http://www.weather.com.cn/weather/101260101.shtml"
content = urllib.request.urlopen(url).read()
soup = BeautifulSoup(content, "html.parser")
#print(soup.title.get_text())

content = ""
name = soup.find_all(attrs={"class":"sky skyid lv3 on"})
for u in name:
    wea = u.find(attrs={"class":"wea"}).get_text()
    tem = u.find(attrs={"class":"tem"}).get_text()
    content = "天气:" + wea + " 温度:" + tem
    content = content.replace("\n", "")
    print(content)
```

输出结果如下图所示：

```
# -*- coding: UTF-8 -*-
import urllib
import urllib.request
from bs4 import BeautifulSoup

# 下载数据
url = "http://www.weather.com.cn/weather/101260101.shtml"
content = urllib.request.urlopen(url).read()
soup = BeautifulSoup(content, "html.parser")
#print(soup.title.get_text())

content = ""
name = soup.find_all(attrs={"class":"sky skyid lv3 on"})
for u in name:
    wea = u.find(attrs={"class":"wea"}).get_text()
    tem = u.find(attrs={"class":"tem"}).get_text()
    content = "天气:" + wea + " 温度:" + tem
    content = content.replace("\n", "")
    print(content)
```

Python 3.4.3 Shell
File Edit Shell Debug Options Window
Python 3.4.3 (v3.4.3:9b73f1c
tel) on win32
Type "copyright", "credits"
>>> =====
>>> 天气:阵雨 温度:18/15°C
>>>

By: Eastmount CSDN

二. QQ邮箱设置SMTP

在使用Python自动发送邮件之前，需要对我们的QQ邮箱进行简单的配置，过程如下：

1.首先登陆QQ邮箱，选择“账户”如下图所示：



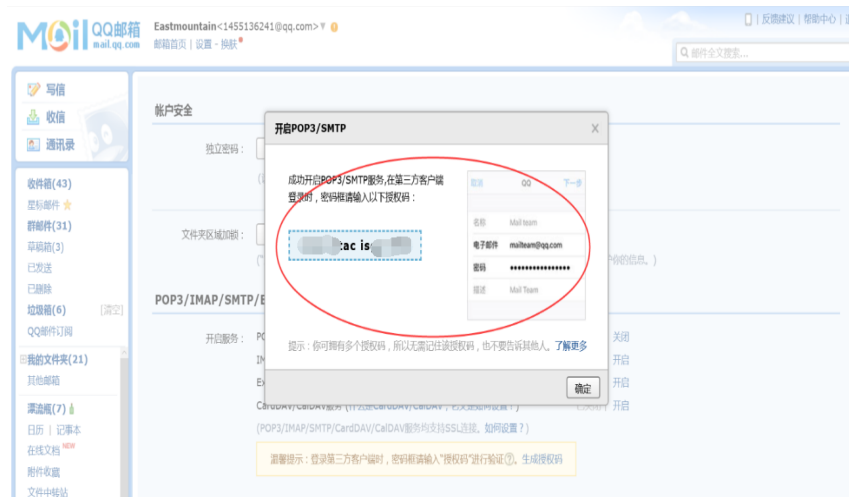
2.在账户页面往下拉，看到“POP3/SMTP”设置，点击开启按钮，如下图所示：



3.弹出如下图所示界面，然后发送这条短信至指定号码，点击“我已发送”按钮。



4.弹出的提示中会显示16位授权码，注意一定要记住这个授权码，后面写Python代码也需要，然后点击“确定”按钮。



5.接下来将收取选项设置为“全部”，并点击“保存”按钮即可。注意端口号如下：



三. Python自动发送邮件

Python 的 email 模块里包含了许多实用的邮件格式设置函数，可以用来创建邮件“包裹”。使用的 MIMEText 对象，为底层的 MIME (Multipurpose Internet MailExtensions, 多用途互联网邮件扩展类型) 协议传输创建了一封空邮件，最后通过高层的SMTP 协议发送出去。MIMEText 对象 msg 包括收发邮箱地址、邮件正文和主题，Python 通过它就可以创建一封格式正确的邮件。smtplib 模块用来设置服务器连接的相关信息。

Python SMTP 对象使用 sendmail 方法发送邮件，语法如下：（参考: [Python SMTP发送邮件](#)）

```
SMTP.sendmail(from_addr, to_addrs, msg[, mail_options, rcpt_options])
```

参数说明：

- from_addr: 邮件发送者地址。
- to_addrs: 字符串列表，邮件发送地址。
- msg: 发送消息

这里要注意一下第三个参数，msg 是字符串，表示邮件。我们知道邮件一般由标题，发信人，收件人，邮件内容，附件等构成，发送邮

件的时候，要注意 msg 的格式。这个格式就是 smtp 协议中定义的格式。

代码如下：

```
# -*- coding: UTF-8 -*-
import smtplib
from email.mime.text import MIMEText
from email.header import Header

#发送邮件
msg_from = "1455136241@qq.com"
#授权码(而不是密码)
EMAIL_HOST_PASSWORD = '****htacisgy****'
#接受邮件
msg_to = "15201615157@163.com"
#主题
subject = "Python测试代码"
#正文
content = "女神，这是我使用python smtplib及email模块发送的邮件。"
print(content)

#MIMEText构建对象 参数分别是：邮件正文、MIMEsubtype中'plain'表示纯文本、utf-8编码
msg = MIMEText(content, 'plain', 'utf-8')
msg['Subject'] = subject
msg['From'] = msg_from
msg['To'] = msg_to

#邮件服务器及端口号
#smtplib.SMTPServerDisconnected: Connection unexpectedly closed
try:
    s = smtplib.SMTP_SSL("smtp.qq.com", 465) #端口号
    s.set_debuglevel(1)
    s.login(msg_from, EMAIL_HOST_PASSWORD)
    s.sendmail(msg_from, msg_to, msg.as_string())
    print("发送成功")
except s.SMTPException as e:
    print("发送失败")
    print(e)
finally:
    s.quit()
```

发送成功之后如下图所示，注意login()输入邮箱名和授权码，而不是密码。



运行过程输出内容如下：

>>>

女神，这是我使用python smtplib及email模块发送的邮件。

```
send: 'ehlo [192.168.0.101]\r\n'
reply: b'250-smtp.qq.com\r\n'
reply: b'250-PIPELINING\r\n'
reply: retcode (250); Msg: b'smtp.qq.com\nPIPELINING\'
send: 'AUTH PLAIN ADE0NTUxMzYyNDFAcXEuY29tAGVveXFodGFjaXNndmlmYmg=\r\n'
reply: b'235 Authentication successful\r\n'
reply: retcode (235); Msg: b'Authentication successful'
send: 'mail FROM:<1455136241@qq.com> size=296\r\n'
reply: b'250 Ok\r\n'
reply: retcode (250); Msg: b'Ok'
send: 'rcpt TO:<15201615157@163.com>\r\n'
reply: b'250 Ok\r\n'
reply: retcode (250); Msg: b'Ok'
send: 'data\r\n'
reply: b'354 End data with <CR><LF>.<CR><LF>\r\n'
reply: retcode (354); Msg: b'End data with <CR><LF>.<CR><LF>'
data: (354, b'End data with <CR><LF>.<CR><LF>')
send: b'Content-Type: text/plain; charset="utf-8"\r\nMIME-Version: 1.0\r\n...r\nTo: 15201615157@163.com\r\n'
reply: b'250 Ok: queued as \r\n'
reply: retcode (250); Msg: b'Ok: queued as'
data: (250, b'Ok: queued as')
发送成功
send: 'quit\r\n'
reply: b'221 Bye\r\n'
reply: retcode (221); Msg: b'Bye'
>>>
```

四. 完整代码实现

完整代码如下所示:

```
# -*- coding: UTF-8 -*-
import urllib
import urllib.request
from bs4 import BeautifulSoup

# 下载数据
url = "http://www.weather.com.cn/weather/101260101.shtml"
content = urllib.request.urlopen(url).read()
soup = BeautifulSoup(content, "html.parser")
content = ""
name = soup.find_all(attrs={"class": "sky skyid lv3 on"})
for u in name:
    wea = u.find(attrs={"class": "wea"}).get_text()
    tem = u.find(attrs={"class": "tem"}).get_text()
    content = "天气:" + wea + " 温度:" + tem
    content = content.replace("\n", "")
    print(content)

# 发送邮件
import smtplib
from email.mime.text import MIMEText
from email.header import Header

msg_from = "1455136241@qq.com"
```

```

EMAIL_HOST_PASSWORD = '****htacisgy****'
msg_to = "15201615157@163.com"

subject = "Python爬取天气"
other = content + "\n这是我使用python smtpplib及email模块发送的邮件。"
print(other)

msg = MIMEText(other, 'plain', 'utf-8')
msg['Subject'] = subject
msg['From'] = msg_from
msg['To'] = msg_to

try:
    s = smtplib.SMTP_SSL("smtp.qq.com", 465)
    s.set_debuglevel(1)
    s.login(msg_from, EMAIL_HOST_PASSWORD)
    s.sendmail(msg_from, msg_to, msg.as_string())
    print("发送成功")
except s.SMTPException.e:
    print("发送失败")
    print(e)
finally:
    s.quit()

```

发送成功之后如下图所示：



需要注意，代码有时能发送成功，有时报错“smtplib.SMTPServerDisconnected: Connection unexpectedly closed”，网上说是设置端口465的原因，但作者已经设置了的，不知道为什么？希望博友帮忙。

(By:Eastmount 2018-5-3 下午4点 <http://blog.csdn.net/eastmount/>)

👍 点赞 2 ☆ 收藏 🔄 分享 ...



Eastmount 博客专家

发布了444 篇原创文章 · 获赞 5939 · 访问量 486万+

他的留言板

关注