

爬虫小案例—带参数的两种方式

爬虫获取百度搜索结果

方法一

百度搜索的搜索关键词是通过链接提交的，比如在百度搜索“开课吧”，答案页的链接为<https://www.baidu.com/s?wd=开课吧>

这样就简单了，我们可以自己拼接url来实现百度搜索关键词提交

```
1 # coding: UTF-8
2 import requests
3 url="http://www.baidu.com/s?wd="
4 wd="开课吧"
5 #拼接url
6 url=url+wd
7 #通过requests发送get请求
8 r=requests.get(url)
9 #打印获取内容的字符串数据
10 print r.text
```

第一行里面# coding: UTF-8是声明编码的，不写的话将不能识别中文。

重点是这句，字符串拼接一下

url="http://www.baidu.com/s?wd="

wd="开课吧"

url=url+wd

这样就可以实现关键词提交

方法二

刚刚已经实现了百度关键词提交。现在讲另一个方法来提交关键词。

```
1 # coding: UTF-8
2 import requests
3 url="http://www.baidu.com/s"
4 kv={"wd":"开课吧"}
5 r=requests.get(url,params=kv)
6 r.encoding="utf-8"
7 #打印字符串的前2000的内容
8 print r.text[:2000]
```

其中,这个网页不设置r.encoding是乱码的,所以设为utf-8

提交搜索词的方法也是构建一个字典键值对

kv={"wd": "开课吧"}

然后通过get方法传进去

r=requests.get(url,params=kv)

r.text[:2000]是取字符串0-2000的部分

可以通过r.request.url来查看生成的链接