scapy学习笔记

Ringo Yungpo Kao

December 12, 2017

# scrapy学习笔记

## 学习环境

|  |  |
| --- | --- |
| 类别 | 环境 |
| 系统 | Windows 10 64bit |
| Shell | PowerShell |

## 环境安装

### Python安装

1. 到官网的[下载链接](https://www.python.org/ftp/python/3.5.3/python-3.5.3-amd64.exe)下载python3.5.3进行python的安装，**勾选Add Python3.5 to PATH**，避免手动添加，然后一路下一步。
2. 在PowerShell中输入python获取版本号无误即安装成功。

PS C:\Users\Ringo> python  
Python 3.5.3 (v3.5.3:1880cb95a742, Jan 16 2017, 16:02:32) [MSC v.1900 64 bit (AMD64)] on win32  
Type "help", "copyright", "credits" or "license" for more information.

### Scrapy安装

1. 在windows下安装Scrapy前必须要安装[pywin32](https://nchc.dl.sourceforge.net/project/pywin32/pywin32/Build%20220/pywin32-220.win-amd64-py3.5.exe)。
2. 下载[Lxml](https://download.lfd.uci.edu/pythonlibs/yjwkc9i2/lxml-4.1.1-cp35-cp35m-win_amd64.whl)，并在相应目录下在PowerShell输入pip install .-4.1.1-cp35-cp35m-win\_amd64.whl
3. 输入pip install Scrapy安装Scrapy。

## Scrapy命令式交互模式

本次现在命令行下面进行对于数据的抓取 抓取的目标网站是[赶集网](http://bj.ganji.com/fang1/chaoyang/) 在PowerShell下输入 >scrapy shell http://bj.ganji.com/fang1/chaoyang/