

# Python 常用的 50 个第 3 方模块及说明

requests	对 HTTP 协议进行高度封装
bs4	解析 HTML 的模块
pymongo	把数据写入 MongoDB
numpy	支持大量的维度数组与矩阵运算，此外也针对数组运算提供大量的数学函数库
pandas	一个强大的分析结构化数据的工具集
scipy	高级科学计算库。包括概率统计、信号、傅里叶等等
sympy	符号计算库
statsmodels	统计&计量库
PIL(Pillow)	图像处理标准库
chardet	字符串编码库
Jinja2	基于 python 的模板引擎
Flask	一个使用 Python 编写的轻量级 Web 应用框架。其 WSGI 工具箱采用 Werkzeug，模板引擎则使用 Jinja2
Django	一个开放源代码的 Web 应用框架，由 Python 写成。采用了 MTV 的框架模式，即模型 M，视图 V 和模版 T
Tornado	一种 Web 服务器软件的开源版本。Tornado 和现在的主流 Web 服务器框架（包括大多数 Python 的框架）有着明显的区别：它是非阻塞式服务器，而且速度相当快
PyQt5	编写 Python 脚本的应用界面
pyecharts	用于生成 Echarts 图表的类库
pygame	专为电子游戏设计的一个模块
jieba	中文分词模块
matplotlib	Python 的绘图库。 它可与 NumPy 一起使用
wordcloud	词云生成器

Scrapy	很强大的爬虫框架，可以满足简单的页面爬取
Crawley	高速爬取对应网站的内容，支持关系和非关系数据库，数据可以导出为 JSON、XML 等
Portia	可视化爬取网页内容
newspaper	提取新闻、文章以及内容分析
selenium	一个调用浏览器的 <b>driver</b> ，通过这个库可以直接调用浏览器完成某些操作，比如输入验证码
cola	分布式爬虫框架
dnspython	DNS 工具包
XlsxWriter	操作 Excel 工作表的文字，数字，公式，图表等
smtplib	发送电子邮件模块
pymysql	操作 MySQL 数据库
NLTK	一个自然语言处理的第三方库，NLP 领域中常用，可建立词袋模型（单词计数），支持词频分析（单词出现次数）、模式识别、关联分析、情感分析（词频分析+度量指标）、可视化（+matplotlib 做分析图）等
sklearn	实现了一些机器学习算法
SQLAlchemy	基于设计模式的对象关系映射 ORM 闻名，它可以把 python 类映射为数据库中表
PonyORM	ORM 程序包，它能自动进行查询、优化，可以通过 python 的生成器表达式来查询数据库
statsmodels	可以配合 numpy、scipy、pandas 完成信号处理，协整、滤波、谱分析等
lxml	xml 和 HTML 读取和解析
win32com	有关 Windows 系统操作、Office（Word、Excel 等）文件读写等的综合应用库
HappyBase	HBase 连接库
OpenCV	图像和视频工作库
TimeSide	能够进行音频分析、成像、转码、流媒体和标签处理的 Python 框架
pydub	支持多种格式声音文件，可进行多种信号处理、信号生成、音效注册、静音处理等
TensorFlow	谷歌的第二代机器学习系统，内建深度学习的扩展支持
theano	深度学习库。它与 Numpy 紧密集成，支持 GPU 计算、单元测试和自我验证
keras	一个用 Python 编写的高级神经网络 API，能够运行在 TensorFlow 或者 Theano 之上，它的开发重点是实现快速实验

Caffe	一个深度学习框架，主要用于计算机视觉，它对图像识别的分类具有很好的应用效果
IPython	一个基于 Python 的交互式 shell，比默认的 Python shell 好用得多，支持变量自动补全、自动缩进、交互式帮助、魔法命令、系统命令等，内置了许多很有用的功能和函数
PTVS	Visual Studio 的 Python 工具
matpython	MATLAB 集成库
PySpark	Spark 提供的 Python API
Plotly	Plotly 提供的图形库可以进行在线 WEB 交互，并提供具有出版品质的图形，支持线图、散点图、区域图、条形图、误差条、框图、直方图、热图、子图、多轴、极坐标图、气泡图、玫瑰图、热力图、漏斗图等众多图形