# 爬虫课程在线安装库

# Ubuntu 16.04教学环境

## 安装pyexecjs模块

sudo pip3 install pyexecjs

## 2、安装 fake\_useragent模块

sudo pip3 install fake\_useragent

## 3、安装机器视觉库tesseract-ocr

sudo apt install tesseract-ocr

sudo pip3 install pytesseract

## 3、安装Scrapy框架

**2.1 安装依赖包**

1、sudo apt-get install libffi-dev

2、sudo apt-get install libssl-dev

3、sudo apt-get install libxml2-dev

4、sudo apt-get install python3-dev

5、sudo apt-get install libxslt1-dev

6、sudo apt-get install zlib1g-dev

7、sudo pip3 install -I -U service\_identity

**2.2 安装Scrapy框架**

1、sudo pip3 install Scrapy **##注意此处S为大写**

2、验证(python交互模式)

import scrapy **##注意此处s为小写**

1. 进入终端，执行如下命令创建scrapy爬虫项目

scrapy startproject Testspider

## 4、安装分布式组件scrapy\_redis

sudo pip3 install scrapy\_redis