

实验一 Python 入门与实践

一、实验目的

1. 熟悉 Python 编程语言，学会安装 Python 开发环境，编写第一个 Python 程序，并学会使用 vscode 或者 IDE（比如 PyCharm）进行 Python 开发；
2. 了解网络爬虫，掌握爬虫的原理；
3. 能够根据需求独立开发爬虫代码，掌握 sql 基础语法，将爬取的数据存储到数据库中；
4. 了解如何管理开发环境，使用 pip 等环境管理工具使用虚拟环境进行工程作业；
5. 了解关于爬虫的最新研究热点技术，比如反爬措施与防反爬措施等。

二、实验原理

1. 实验流程图：
2. 爬虫原理：

三、基础知识

(请详述每部分的内容)

1. 计算机网络(在浏览器中输入一个 url 后回车,后台会发生什么?)
2. HTML/CSS/JS
3. Chrome 浏览器开发者工具的使用
4. Python 基础语法及常用库
5. Mysql 数据库/mongodb 数据库/云数据库(根据自己所选择的实验数据库写一个即可)
6. 开发环境(使用何种方法进行开发环境的管理, pip, Anaconda, docker, 选择自己的使用方法, 详述使用过程和配置过程)

四、实验过程

(给出工程代码的组成思路和详细代码)

五、 小结讨论

(对整个爬虫项目进行小结,就爬虫代码实现过程中遇到的问题进行讨论,比如遇到了哪些反爬虫措施?你使用何种方案进行解决?你还知道哪些公司采取哪些反爬虫措施?或者你还知道哪些应对反爬虫措施的解决方案等)