概要设计说明书

1. 概述

近几年来，计算机和信息技术迅猛发展并且普及应用，行业应用系统的规模迅速扩大，所产生的数据呈爆炸性增长，因此提高对数据的提取、处理和分析能力，通过进一步加工实现数据的“增值”已经成为现在世界的迫切需求。本项目通过编写python网络爬虫程序，使用requests库、bs4库等爬取互联网中近期与计算机相关专业的招聘信息，以及收集各专业学科学习课程内容。通过数据清洗，转化，可用信息提取等技术手段获得可用的原始数据，并进行可视化的展示。利用word2vec、jieba等python库，运用数据分析技术，把项目要求的最终结果通过可视化的手段展现出来。  
2.软件概述

1. 软件概述
   1. 执行背景

信息科技经过 60 余年的发展，已经从高端科技普及到社会运转的每一个角落中。随着信息技术在国家治理、经济运行的方方面面的运转，大量的数据随之产生。而互联网技术的爆发式发展使得近两年来产生的数据总量超过了人类有史以来的数据总和。如何利用这些数据，挖掘数据的深层价值，是未来信息科技发展的趋势，也是大数据技术产生的背景。

* 1. 执行环境

（1）适用于各高校未毕业学生，让其在校学习时了解企业对人才需求的动向，综合开率专业，知识技能，收入，地域等因素。既进行操作实践，又可以对未来的学习有所规划。

（2）适用于各高校应届毕业生，其通过本网站可以轻松查询工作相关信息，可更好的对比自身条件，扬长避短，找到更适合自己的工作。

* 1. 设计约束

后端主要采用了python和Sql语言，前端主要使用HTML5语言,JavaScript，开发工具为pycharm , eclipse。后台存储文件为json格式，数据处理时采用了mysql数据库进行存储以及处理。

1. 整体结构设计
   1. 软件结构

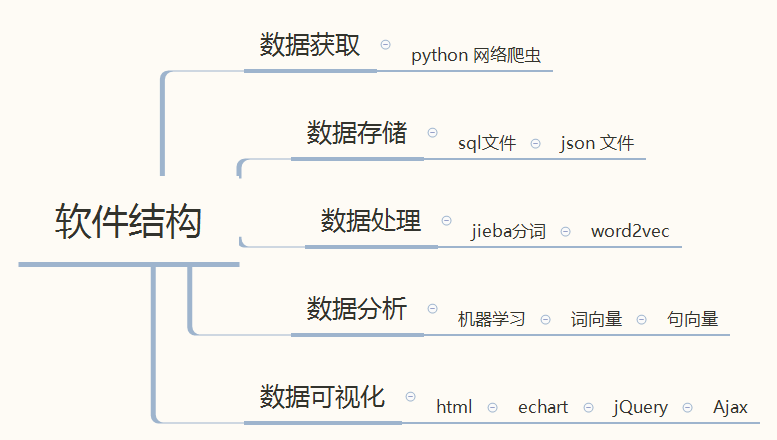


图3-1

* 1. 设计思想

应用大数据挖掘、分析实践，把大数据挖掘、分析工具、语言的熟练应用，理论知识和实践相结合。

采用大数据分析的方法和思路，切实了解社会对大数据技能的需求。

了解企业对人才需求的动向，综合开率专业，知识技能，收s入，地域等因素。

* 1. 模块设计

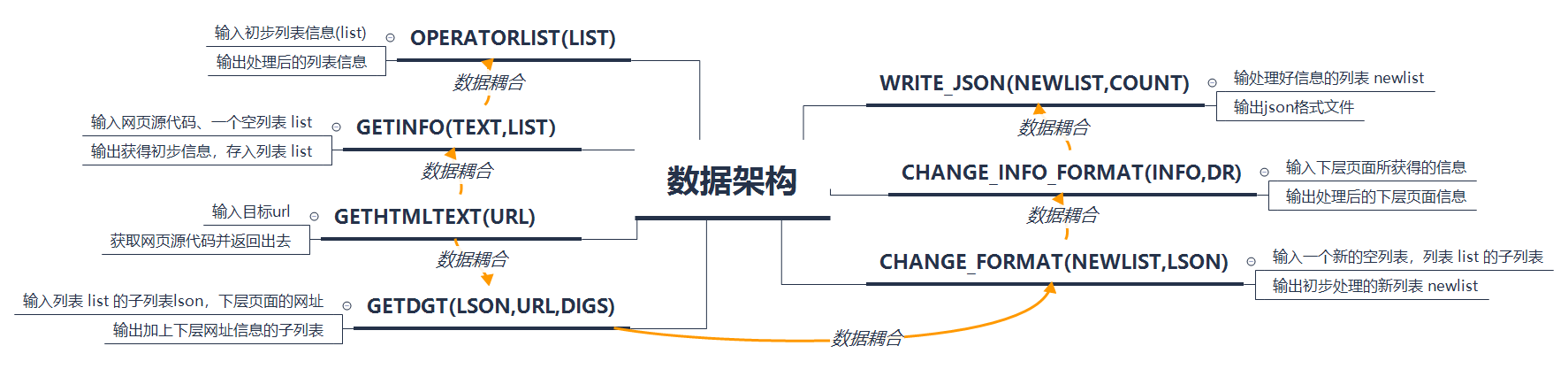


图3-1-1

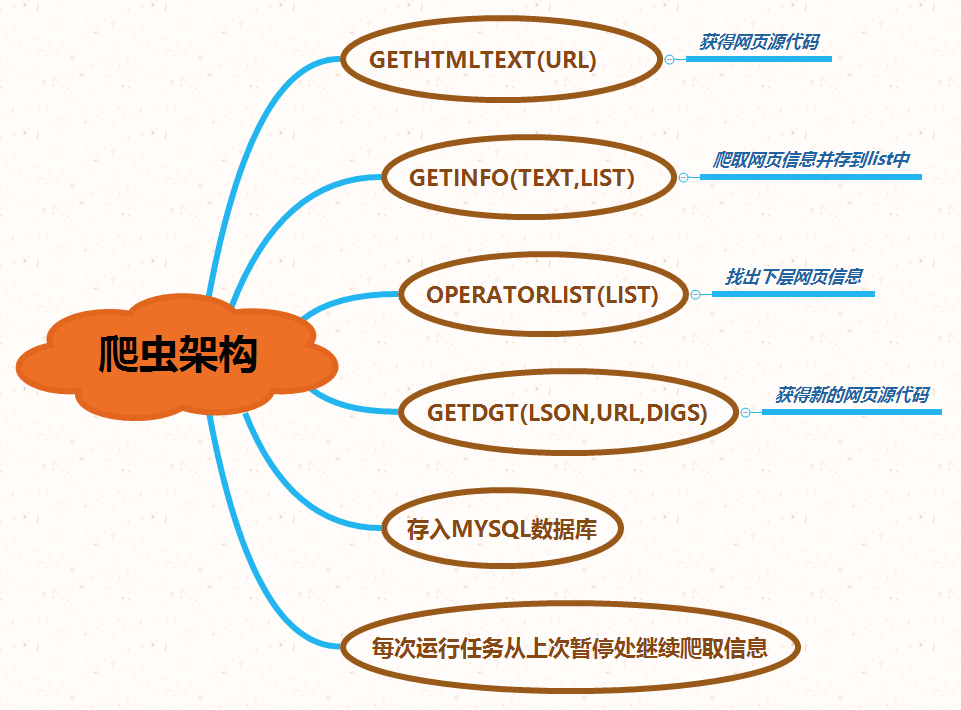


图3-1-2

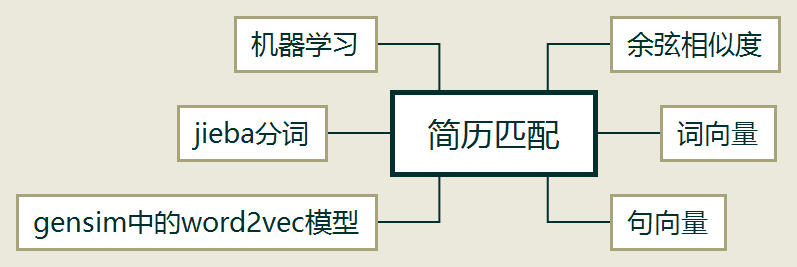


图3-1-3