**数字货币大数据分析平台**

**项目需求分析文档**

**项目名称：数字货币大数据分析平台**

**项目组：数字货币大数据分析平台开发组**

**项目成员：肖文斌（项目经理）、文杰、陈楚凡、景南瑞、吴后望、陈修鸿**

**2021 年 07 月 05 日**



**目 录**

[1. 引言 4](#_Toc9426)

[1.1 文档编写目的 4](#_Toc24322)

[1.2 项目开发背景 4](#_Toc15644)

[1.3 项目开发意义 5](#_Toc22785)

[2. 项目概述 6](#_Toc3622)

[2.1 项目标识 6](#_Toc4588)

[2.2 数据来源 6](#_Toc9195)

[2.3 用户特点 6](#_Toc16684)

[2.4 开发时限 7](#_Toc11810)

[2.5 开发工具 7](#_Toc7620)

[2.6 编程语言 7](#_Toc23875)

[3. 开发配置 8](#_Toc7623)

[3.1 硬件需求 8](#_Toc8358)

[3.2 软件需求 8](#_Toc18038)

[3.3 开发环境 8](#_Toc5802)

[4. 系统需求 9](#_Toc14928)

[4.1 功能需求 9](#_Toc18183)

[4.2 性能需求 10](#_Toc23705)

[4.3 接口需求 10](#_Toc16564)

[*4.3.1 用户界面* 10](#_Toc24096)

[*4.3.2 硬件接口* 10](#_Toc16400)

[*4.3.3 软件接口* 10](#_Toc25945)

[4.4 界面需求 10](#_Toc26405)

[5. 项目开发计划 12](#_Toc21497)

[5.1 需求确定阶段 12](#_Toc8518)

[5.2 数据爬取阶段 12](#_Toc32436)

[5.3 数据分析阶段 12](#_Toc21590)

[5.4 数据可视化阶段 12](#_Toc4062)

[6. 项目开发人员及分工 13](#_Toc21872)

|  |  |
| --- | --- |
| **文档信息** | |
| **标题** | 项目需求分析文档 Project Requirements Analysis Document |
| **作者** | 肖文斌 |
| **创建日期** | 2021.07.05 |
| **更新日期** | 2021.07.19 |
| **版本** | V1.5 |
| **部门名称** | 数字货币大数据分析平台项目开发组-重庆大学实训大数据二组 |

|  |  |  |  |
| --- | --- | --- | --- |
| **文档更新记录** | | | |
| **版本号** | **更新内容** | **操作日期** | **操作人员** |
| **1.0** | 敲确定项目框架、项目需求 | 2021.07.05 | 肖文斌 |
| **1.1** | 确定项目需要实现的功能、开发计划 | 2021.07.07 | 肖文斌 |
| **1.2** | 补充引言、概述、开发背景等信息 | 2021.07.07 | 肖文斌 |
| **1.3** | 确定人员初步分工 | 2021.07.08 | 肖文斌 |
| **1.4** | 修正项目实现功能、人员分工 | 2021.07.19 | 肖文斌 |
| **1.5** | 添加相关性数据集描述，修改一些术语 | 2021.07.19 | 文杰 |

项目需求分析文档 高考志愿查询预测与推荐系统

**项目需求分析文档**

**Project Requirements Analysis Document**

# 引言

## 文档编写目的

本文档是“数字货币大数据分析平台”开发过程中的重要文献资料，是项目工程的起点、软件设计的依据、用户需求的真实反映、开发人员进行系统总体设计的参考以及进行项目具体开发流程的必要参考资料。通过本文档，应能明确软件开发的需求、合理设计项目规划及开发进度、有序组织软件开发及进行相关测试、清晰平台后期运维内容。本文档预期读者为从事“数字货币大数据分析平台”开发的相关人员。

## 项目开发背景

数字货币概念自2008年金融危机中诞生，经过十余年发展，在2015年、2020年迎来爆炸性增长，参与数字货币挖掘的个人、企业人数激增。面对这样一个新兴且热门的事物，掌握其具体的发展数据，对其进行大数据分析是有必要的。本项目开发的软件旨在基于互联网中数字货币交易所数据，对其中热门数字货币的历史数据进行收集，对其中的数字货币进行大数据可视化，并运用股票领域工业模型对数据进行预测，使用机器学习算法对数字货币间的关联性进行挖掘；同时基于数字货币门户网站开发数据，建立实时交易数据展示；爬取互联网中与数字货币市场关联度高的新闻，在本平台进行精简展示。达成以上目标后，用户可以清晰直观获得数字货币市场的历史变化，便捷获取数字货币市场实时交易详情，了解当前全球数字货币市场的舆论情况，帮助用户对数字货币市场做出及时准确的判断。

## 项目开发意义

本项目集爬虫、数据接口、数据存储、数据预处理、数据清洗、数据分析、机器学习、数据展示于一体，涵盖Python、Html、JavaScript、Css、MySQL等程序设计语言，运用Spark数据处理平台、ECharts绘图工具，采用Flask后端开发架构以及Vue、jQuery等前端框架，使用Vscode、PyCharm等集成开发环境以及Git版本控制工具。本次项目旨在帮助计划从事大数据方向的专业学生熟练掌握、运用实训课程的理论知识及相关开发平台、架构，并结合具体应用背景开发一个独立的大数据项目，以帮助学生了解大数据项目的开发流程及具体技术的应用方式，为将来学生的就业打下基础，同时检验学生的学习成果及相关开发能力，锻炼学生的工作能力和团队协作意识。

# 项目概述

## 项目标识

项目全名：数字货币大数据分析平台

## 数据来源

本项目的数据来源为数字货币门户网站<https://coinmarketcap.com/>，所有数据均从该网站中爬取。预计共得到一个数据集：

①该数据集包括：coin.sql数据库表，涵盖100种当前热门的数字货币；

②数据集时间自2013年1月1日（或其发行日）至2021年7月12日，收录了数字货币的名称、简称、当日时间、当日最高价时间、当日最低价时间、当日开盘价、当日最高价、当日最低价、当日收盘价、当日成交量、当日市值等信息；

数据集经清洗、分析后，共得到七个数据集：

①交易量数据集；avg\_volume\_marketcap.sql数据库表，该数据集收录了每日交易量、每日市值、每周平均交易量、每周平均市值、每月平均交易量、每月平均市值；

②交易价格数据集；coin\_price.sql数据库表，该数据集收录了每日开盘价、每日最高价、每日最低价、每日收盘价、每日市值；

③交易涨跌幅数据集；coin\_rise\_fall.sql数据库表，该数据集收录了当日收盘价、昨日收盘价、今日涨跌幅、一周前收盘价、一周内涨跌幅、一月前收盘价、一月内涨跌幅；

④交易市值数据集；marketcap\_percentage.sql数据库表，该数据集收录了当日时间、当日市值、当日市值占比；

⑤数字货币价格预测数据集；coin\_predict.sql数据库表，该数据集收录了10种热门数字货币（分别由市值排名前5的主流币和市值排名前5的空气币组成）的未来五天的价格预测，预测开盘价格、预测最高价格、预测最低价格、预测收盘价格。

⑥热门货币收盘价数据集；corr\_original\_data.sql数据库表，该数据集收录了10种热门货币最近两个月的闭盘价格并修改格式，用于计算该10种货币的相关性矩阵。

⑦热门货币相关性数据集；correlation.sql数据库表，该数据集收录了进过计算分析得到的10种热门货币的相关性矩阵，其本质上为一个10\*10的相关性矩阵。

## 用户特点

1、 对数字货币市场具有兴趣的个人、企业，渴望快速获得数字货币市场的相关数据；

2、 从事数字货币市场投资的相关个人、企业，数字货币市场的监管者；

3、 有基本的网络运用能力，具备基本金融知识、风险意识。

## 开发时限

2021 年 7 月 5 日至 2021 年 7 月 24 日共 19 天。

## 开发工具

PyCharm、VSCode、Spark、MySQL、ECharts、Flask、jQuery、Vue

## 编程语言

Python、JavaScript、Html、Css

# 开发配置

## 硬件需求

本项目可以在单机或集群上完成，没有特殊的硬件配置需求。

## 软件需求

①操作系统：Windows 10专业版

②编程工具：Pycharm、Vscode

③数据处理工具：Apache Spark

④数据库：MySQL

⑤版本控制工具：Git

## 开发环境

本项目在 PC 环境下开发，要求各计算机配备足够的存储空间。

# 系统需求

## 功能需求

* + 1. *注册功能*

用户使用平台需要先行注册，注册需要输入用户名、密码、邮箱地址、邮箱验证码，对用户名、密码、邮箱进行格式验证，对验证码进行正确性验证，所有验证通过后允许用户注册，将用户信息存入数据库中。

* + 1. *登录功能*

用户使用已注册的账户进行登录，输入用户名、密码后，系统与数据库中用户数据进行比对，比对通过后登录成功并进入大数据展示页面，比对失败则登录失败。

* + 1. *重置密码功能*

用户忘记密码时，可通过重置密码进行重置，用户输入已注册用户名、新密码、邮箱验证码，系统验证用户名存在性、密码格式、验证码正确性，所有验证通过后重置密码成功。

* + 1. *周涨跌幅展示功能*

用户可以在大数据展示面板中查看一百种数字货币的周涨跌幅，也可查看具体某种数字货币的周涨跌幅。

* + 1. *开盘价排名展示功能*

用户可以在大数据展示面板中查看一百种数字货币的今日开盘价、今日市值，对开盘价进行排名滚动展示。

* + 1. *市值占比展示功能*

用户可以在大数据展示面板中查看一百种数字货币中市值最高六种，采用南丁格尔玫瑰图展示。

* + 1. *数字货币K线图展示功能*

用户可以在大数据展示面板中查看一百种数字货币的所有K线图数据，使用选择框进行选择，同时展示当前数字货币的当日价格和当日涨跌幅。

* + 1. *数字货币价格预测展示功能*

用户可以在大数据展示面板中查看十种数字货币的后五日价格预测，使用选择框进行选择，可展示预测开盘价格、预测收盘价格、预测最高价格、预测最低价格，采用折线图展示。

* + 1. *数字货币排名展示功能*

用户可以在大数据展示面板中查看当前数字货币市场的主流币排行前五、空气币排行前五及该数字货币的市值、当日交易额，采用排行榜展示。排行榜会在经过固定时间后进行刷新

* + 1. *数字货币关联性展示功能*

用户可以在大数据展示面板中查看当前数字货币市场的主流币市值前五和空气币市值前五的相关性，其中数据大小范围为[-1,1]，负数表示负相关程度，正数表示正相关程度，相关性矩阵采用热图展示。

* + 1. *数字货币实时交易信息展示功能*

用户可以在实时信息展示页面中查看当前时刻数字货币市场所有数字货币的实时交易数据，采用本币/法币的展示样例，采用折线图展示。

* + 1. *数字货币实时舆论信息展示功能*

用户可以在实时信息展示页面中打开某一具体的数字货币交易，可在Latest News展示栏查看实时的互联网数字货币新闻舆论。

* + 1. *平台后端管理-数据管理功能*

平台管理员可登录后台管理页面，查看数据库中存储的原始数字货币，采用选择框进行数字货币选择，可选展示条数。

* + 1. *平台后端管理-用户管理功能*

平台管理员可登录后台管理页面，查看数据库中存储的平台用户信息。

* + 1. *平台后端管理-管理员管理功能*

平台管理员可登录后台管理页面，查看数据库中存储的管理员信息。

## 性能需求

本项目对数据处理能力无要求，数据查询响应速度快即可。

## 接口需求

* + 1. *用户界面*

用户在 Web 界面上点击特定区域时能切换到对应的其他页面。

* + 1. *硬件接口*

无。

* + 1. *软件接口*

通过 Flask 连接数据库与前端，通过jQuery、Vue连接前端与后端。

## 界面需求

①输入设备：鼠标；

②输出设备：PC 显示器；

③分辨率：满足 Web 网页的分辨率；

④提示风格：弹出窗口、跳转其他页面；

⑤其他：页面风格简约美观，能让用户直观地查询和浏览各数据、便捷地使用各项功能。

# 项目开发计划

## 需求确定阶段

对数字货币大数据分析平台进行需求分析，确定该项目需要实现的所有具体功能及其实现方法。随后对整个流程进行分工，撰写项目需求分析报告。

## 数据爬取阶段

根据上一阶段敲定的功能确定需要爬取的数据内容，编写爬虫程序，进入目标网站进行数据爬取。确定获取实时数据的门户网站，申请或解析数据API，尝试进行数据API连接并使用。

## 数据分析阶段

将已爬取的数据进行分析。首先进行数据清洗，清除无关数据与各种符号， 并将互相有关联的数据整合，得到相关数据分析表。采用相关机器学习算法对数据进行预测分析并将预测结果存储。

## 数据可视化阶段

根据具体功能和 Web 页面的布局，将各数据、图像等展示在合适的位置， 满足项目要求中的界面需求，达到良好的用户交互。

# 项目开发人员及分工

**肖文斌：项目经理、爬虫工程师、前端工程师、后端工程师**

负责项目的整体协调工作，项目需求分析文档编写，文档审查与修改，原始数据集的爬取；

负责登录、注册、忘记密码、后台管理等功能的前端、后端功能设计及实现。

**文杰：前端工程师、后端工程师、测试工程师**

负责项目展示大面板的前端、后端功能设计及实现，协助登录、后台管理部分功能实现及Bug修复；

负责平台功能测试工作，编写测试结果文档。

**陈楚凡：爬虫工程师、前端工程师、后端工程师**

负责实时大数据的数据采集、分析、展示，项目系统架构文档编写；

负责实时大数据展示前端、后端功能设计及实现。

**景南瑞：前端工程师、后端工程师、测试工程师**

负责登录、注册、忘记密码等功能的前端、后端功能设计及实现，大数据展示面板透明化设计及实现。

负责项目开发计划文档编写，负责项目测试计划设计及相关文档编写。

**吴后望：前端工程师、后端工程师、测试工程师**

负责项目展示大面板的前端、后端功能设计及实现；

负责项目测试用例设计及相关文档编写，负责平台安全、前端、压力测试工作。

**陈修鸿：数据库工程师、数据分析师、算法工程师**

负责项目所使用的数据库设计及实现，对原始数据进行数据挖掘、分析，编写相关数据库设计文档；

负责项目使用机器学习算法的设计，实现预测、关联数据的计算获取。