# 概要设计说明书

# 1 引言

## 1.1编写目的

本文将给出本系统系统的基本需求和基本的数据流程图，要实现整个系统，需要对用户的需求进行设计，概要设计主要是利用比较抽象爱你个的语言对整个需求进行概括，确定对系统的物理配置，确定整个系统的处理流程和系统的数据结构，接口设计，实现对系统的初步设计。本文的预期读者为对本项目有兴趣的人以及程序设计人员。

## 1.2背景介绍

如今社会十分注重投资，股票投资作为一种经典并且广泛的投资方式被大部分人所接受，其中一部分的股民希望通过更加规律性的事物对于股票的涨跌进行预判，并得到能够长期稳定盈利的方式，而本系统则是为该部分人群制作，希望能够通过本系统的提供的数据，对大部分股票进行正确的预判。另一功能则是为股民提供一些信息的整合与汇总，本功能为次要功能。

|  |  |
| --- | --- |
| 系统名称 |  |
| 系统提出者 |  |
| 开发者 |  |
| 主要用户 |  |
| 硬件平台 |  |

## 1.3定义

Spring Boot：Spring Boot是由Pivotal团队提供的全新框架，其设计目的是用来简化新Spring应用的初始搭建以及开发过程。该框架使用了特定的方式来进行配置，从而使开发人员不再需要定义样板化的配置。通过这种方式，Spring Boot致力于在蓬勃发展的快速应用开发领域(rapid application development)成为领导者。

MongoDB: 是一个介于关系数据库和非关系数据库之间的产品，是非关系数据库当中功能最丰富，最像关系数据库的。他支持的数据结构非常松散，是类似json的bson格式，因此可以存储比较复杂的数据类型。Mongo最大的特点是他支持的查询语言非常强大，其语法有点类似于面向对象的查询语言，几乎可以实现类似关系数据库单表查询的绝大部分功能，而且还支持对数据建立索引。

Ubuntu：Ubuntu（友帮拓、优般图、乌班图）是一个以[桌面](https://baike.baidu.com/item/%E6%A1%8C%E9%9D%A2)应用为主的开源GNU/Linux操作系统，Ubuntu 是基于[Debian](https://baike.baidu.com/item/Debian)GNU/[Linux](https://baike.baidu.com/item/Linux)，支持x86、amd64（即x64）和[ppc](https://baike.baidu.com/item/ppc/150)架构，由全球化的专业开发团队（Canonical Ltd）打造的。

Java：[Java](https://baike.baidu.com/item/Java/85979)是一门[面向对象](https://baike.baidu.com/item/%E9%9D%A2%E5%90%91%E5%AF%B9%E8%B1%A1)编程语言，不仅吸收了[C++](https://baike.baidu.com/item/C%2B%2B)语言的各种优点，还摒弃了C++里难以理解的[多继承](https://baike.baidu.com/item/%E5%A4%9A%E7%BB%A7%E6%89%BF)、[指针](https://baike.baidu.com/item/%E6%8C%87%E9%92%88)等概念，因此Java语言具有功能强大和简单易用两个特征。Java语言作为静态面向对象编程语言的代表，极好地实现了面向对象理论，允许程序员以优雅的思维方式进行复杂的编程。

Docker：Docker 是一个[开源](https://baike.baidu.com/item/%E5%BC%80%E6%BA%90/246339)的应用容器引擎，让开发者可以打包他们的应用以及依赖包到一个可移植的容器中，然后发布到任何流行的 [Linux](https://baike.baidu.com/item/Linux) 机器上，也可以实现[虚拟化](https://baike.baidu.com/item/%E8%99%9A%E6%8B%9F%E5%8C%96)。容器是完全使用[沙箱](https://baike.baidu.com/item/%E6%B2%99%E7%AE%B1/393318)机制，相互之间不会有任何接口。

IDEA编译器：IDEA 全称IntelliJ IDEA，是java语言开发的集成环境，IntelliJ在业界被公认为最好的java开发工具之一， IDEA是JetBrains公司的产品，总部位于捷克共和国的首都布拉格

## 1.4参考资料

# 2平台需求分析

## 2.1社会需求

随着商场经济的逐渐深入，现代人对于资产的投资理念提高到了一个新的高度，而股票作为一种重要的投资手段，被人们所广泛接受和使用，电视节目与网络节目也都有大量的以“通过股票的形式盈利”的内容，这也从侧面反映出了，股民对于股市的期待以及对于能够通过一些途径获得盈利信息的心态。

下图为“上海证券交易所”2017年至2012年概况数据（数据来源为上交所<http://www.sse.com.cn/market/stockdata/overview/yearly/>）

我们可以明显的看到股票市场处2015年牛市以外均保持一定幅度的上升趋势，股票的总价值上升，可以从侧面体现股票市场的盈利存在一定的规律。

而希望知道或找出这个规律的人也自然而然的应运而生，再结合时代互联网经济的特征，我们处理数据的手段也不断提升，使用科学技术对于现有数据进行分析并找出规律，这是一个十分通用的方式。着同样也是本程序被设计的初衷与想法。

## 2.2系统目标与愿景

本系统的主要目的是开发一款能够比较准确预计股票走势，并且提供大部分股票数据的后台项目。

对于用户可以通过调用接口获取部分股票的历史数据，并且获得股票的预测信息。

根据本项目的计划，系统主要分成以下两部分

1. 数据爬取与展现
2. 数据分析与预测

主要目标有以下几点

1. 数据抓取：

股票历史信息

上市公司企业信息

股票指数信息

股票其他必要信息

1. 数据展示

调用接口展示JSON格式数据

1. 数据分析

使用部分技术对于已获得的数据进行分析，分析预测股票走势（调用接口）

## 2.3系统面向用户类别和特性

系统主要面向从事股票交易的个人投资者，即为股民，倾力投资国内股票市场，希望通过股票交易些许赢利以养家糊口，但不得不或长或短地担任股东，在交易大厅中曾经汗流浃背，曾经瑟瑟发抖，被社会另眼视为赌徒，被大股东视为小猪，且屡受绝望煎熬，无力自救，只能寄希望于政府那样的一个弱势人群。

其特点主要为：具有[分散性](https://baike.baidu.com/item/%E5%88%86%E6%95%A3%E6%80%A7/7479876)和流动性的特点。并且不论男女老少，不分文化职业，只要有购买能力都可以参与。因为门槛低，也导致了大部分用户不具备专业的投资水平和更多的投资时间，这也是大部分股民赔钱的根本性原因。所以，要想赚钱必须：1、需要花费更多的时间和精力；2、提高自身的水平；3、寻找外界的帮助。4。懂政策技术和指标。

对于这些人群希望通过所谓的技术手段来进行股票的盈利，然而还没有找到比较正确的途径与方法，本系统将服务该部分人群。

## 2.4运行环境

### 2.4.1硬件平台

|  |  |  |
| --- | --- | --- |
| 客户端 | | |
| 序号 | 硬件 | 相关组成描述 |
| 1 | PC端客户机 | 能够发送http请求，具备一定的json格式解析能力 |

|  |  |  |
| --- | --- | --- |
| 服务器端 | | |
| 序号 | 硬件 | 相关组成描述 |
| 1 | 路由器 | 可以稳定发送信息和连接网络的设备 |
| 2 | 服务器 | 暂时进行考虑中，使用云平台或直接使用家用电脑作为主要平台，最低配置要求为：英特尔至强系列或i3-8100或以上，内存8GB以上，硬盘空闲1GB以上，带宽1Mb/s以上 |

### 2.4.2 软件平台

1、Docker：容器（进阶目标）

2、MongoDB: 数据库

3、JDK1.8或OpenJDK1.8以上

## 2.5外部接口需求

### 2.5.1第三方股票查询接口

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 接口描述 | 所需服务 | 数据和控制信息 | 通信定义 | 需求来源 |
| 网易股票接口 | 查询股票信息 | 所有需要读写的数据 | 第三方接口 | 需要获取股票的历史数据与部分的参数 |

## 2.6功能需求

### 2.6.1查询股票历史数据

返回数据必须包含股票的当日基本信息，能够根据日期进行必要的查询，参考sina股票接口数据。

### 2.6.2查询上市公司的基本信息

返回数据必须包含该上市公司的基本信息，其中包含（行业板块信息，公司所在地等与法律相关信息，以及股票分析所需的基本信息，具体请参见雪球公司信息）

### 2.6.3查询其他技术指数信息

返回数据包含股民希望知道的一些指数技术信息以及一部分的网络信息（评论数，关注度等等）

### 2.6.4预测信息

返回数据需要包括提供的预测信息，和建议购买的股票编号与推荐程度，用户可以通过结合自身的经验与本程序提供的各种建议与推荐程度，进行一个有选择性的投资。

并且对于该接口希望能够获取不同算法下的推荐值，或者根据不同的算法进行的结果整合也是需要的。

### 2.6.5业务场景

假设现在有一名无丰富股票基础知识与投资经验的用户A，一位刚刚进入市场，懂得一些股票知识，但是并不熟悉股票的真正运作方式，并且没有过的多的股票投资经验的用户B，以及一位有多年股票投资经验并且对于股票有很深的了解无论从政策方面还是经济方面对于个股有着自身的理解的用户C。以上三位用户都希望通过股票来进行合理的投资是自己的财富或者生活更加美好或有趣。

对于A用户，一位对股市没有过多了解的人来说，希望得到一款系统，能够直接的向这名用户指明那些股票是值得购买的，而这些股票中最推荐购买的又是哪一种，对于这样的完全新手型的股票购买者来说程序需要极其的简单，简单到可以直接不需要通过任何思考就能获利，并提供一个比较便利的途径来达成目的，他只需要一个接口就是最终结果。

对于B用户，一位对于股票有一定的了解，但是无法总是做出正确的决定的投资者而言，他希望调用一个预测接口，并且对于股票添加一些简单的判断，他会调用多种算法实现的不同的股票预测结果，然后再对于每一个股票调用，指数型信息，来实现自己的简单判断，然后在自己的脑海或其他方式中总结出，哪一支股票可以帮助到他，也就是，他需要的简单指数信息和不同算法下的预测结果信息与推荐程度。

对于C用户而言，作为一个资深的股民，他需要一个程序来帮助他缩小范围，能够更加精准快速的找出具有良好潜力的股票，来节约大量的时间，同时他不需要太多的基础技术型指标来帮助他进行判断，他会打开其他比较熟悉的软件进行额外的判断。

### 2.6.6功能需求

优先级：5为最高，1为最低，数值越大将会越优先完成

|  |  |  |
| --- | --- | --- |
| 用户 | | |
| 需求名称 | 优先级 | 备注 |
| 查询股票历史数据 | 3 |  |
| 查询上市公司相关数据 | 3 |  |
| 查询其他技术指标信息 | 2 |  |
| 查询推荐购买的股票 | 5 | 需要使用多种模型进行推理啊 |

|  |  |  |
| --- | --- | --- |
| 管理员 | | |
| 需求名称 | 优先级 | 备注 |
| 爬去信息接口 | 5 | 需要爬去各种信息，并且完成信息的存储 |

### 2.6.7非功能性需求

#### A 性能需求

1、数据精准度

时间精准到天

由于部分股票价格剥夺比较薄弱，使用股票价格精准至分

2、并发特性

在网络稳定并且服务器达到标准时，可以承受大于1000的并发量并且不产生丢包或数据丢失等情况

3、故障处理特性

保证在每十五秒为周期检查接口活跃度，如果接口为非活跃状态或假死状态则停止服务并自动重新启动

#### B 安全措施需求

1、程序将会拥有一套完整的日志体系，可以打印出所有的错误发生和时间点，并且通过接口活跃度检查，将会进行邮件通知的报警行为。

2、 应用日志必须按日保存，每个文件保存应用一天的完整日志；

3、 历史应用日志必须压缩后按日保存，以减小历史日志文件对硬盘空间的占用；

4、 历史应用日志必须保留一个月（30天），以保证问题出现后，日志可以得到较好的保存。

#### C安全需求

系统要有足够的防御能力，防御普通黑客的攻击，尽量减少各种可能出现的漏洞，防止用户信息被不良分子窃取。

保证用户权限的正确设置，防止出现各种越权行为，保证系统的正常良好运行。  用户的信息保存采用国际流行的MD5加密，保证安全。

管理员接口将进行简单的鉴权（输入一串加密字符串，并进行解析鉴权）

#### D 操作需求

系统操作者多为非专业人士，所以必须提供直观的图形界面，使能够快速上手使用。

#### E 质量需求

易用性优于易学性，或者可移植性优于有效性。

## 2.7软件需求规格表

|  |  |  |  |
| --- | --- | --- | --- |
| 软件需求表 | | | |
| 子系统 | 模块内容 | 对应接口 | 功能说明 |
| 日志与测试系统 | 日志分析  自动化测试 | 日志分析接口  自动化测试检测接口 | 获取与分析日志内容，以最快的速度对日志进行分析，并且对错误日志进行汇总并且提交邮件，同时进行接口活跃度测试，实时监控接口活跃度达到良好的用户体验 |
| 安全系统 | 基本鉴权 | 包装型的鉴权接口 | 对于管理员的请求进行轻量级的鉴权服务，以此来抵挡误输入与部分的攻击行为，并提供给数据以基本的安全体系 |
| 核心算法预测系统 | 股票推荐与分数排名 | 各算法对应的推荐接口 | 本程序的核心算法与核心需求，对于本部分的接口我们使用不同的算法进行分析，以此来保证准确性。 |
| 爬虫系统 | 将存在于网络的上的数据接入本地或远程数据库 | 网易爬虫接口  百度爬虫接口  雪球爬虫接口  东方财富爬虫接口  新浪爬虫接口  官方爬虫接口 | 对于不同的数据我们使用不同的数据源来进行良好的爬去，对于一些特定的数据源我们也将使用特定的爬虫以此获取，这个系统也同样是本程序的核心算法之一，因为所有的数据分析都是建立在已有数据的基础上的 |
| 数据展现系统 | 将数据库包含的数据展示给用户 | 历史记录接口  技术指标接口  公司信息接口 | 主要提供给第三方或股民用户其他的评判条件 |