

武汉大学计算机学院

大型应用软件设计

题 目: Stockalize 证券决策系统

专业(班): 15 卓越工程师班

项目组成员: 组长: 王致远 2015301500132 13026138150

成员 1: 罗格 2015301500123

成员 2: 王一舟 2015301500153

成员 3: 李文洲 2015301500162

成员 4: 王睿达 2015301500082

2018 年 1 月 20 日

成绩及评语

学号	姓名	小组成绩	个人成绩	个人小结	总评成绩
2015301500132	王致远	10	作为组长吗，这次开发尽量要考虑到各个成员的能力水平和负载均衡，以及团队间成员的协调和总进度的把控。此外还有微信服务器的搭建。对我是一个比较大的锻炼。		
2015301500123	罗格		这次我在团队主要负责模型开发，这是一个关键的工作，既要应用数据还要交付给微信。对我也是一个很大的锻炼。		
2015301500153	王一舟		这次的工程项目让我受益良多，不仅是对上学期《软件工程》课程内容的复习和运用，更是让我体验了小组分工合作的应用开发的整个流程。		
2015301500162	李文洲		本次大型软件应用设计项目我负责了前期部分文档的撰写以及股票各类信息的爬取存档相关代码编写调试。通过这次大软实践，我对实际文档编写有了更深入的了解，也在实验过程中对各类数据爬取工具进行了多方面比较，最终确定了python+tushare 的组合，对数据爬取有了更深层次的掌握		
2015301500082	王睿达		这次项目开发相比暑假的工程项目实践来说，老师的指导更少，更需要我们自己讨论分工，交流想法，合作完成任务，在磨合中提高了我们的默契。也让没有选修软件工程的我也在合作开发中大致了解了规范的开发流程。		

评语：

目录

项目计划书	4
可行性研究和软件开发计划.....	5
需求规格说明书.....	40
软件设计说明书.....	59
测试计划	59

Stcoklize 证券分析系统

第一部分

项目任务书

目录

项目背景	5
项目简介	6
项目人员	6
项目交付	6

项目背景

随着证券分析技术和软件技术的发展，就股票投资方法论而言，炒股软件进化出很多功能：技术分析、基本分析、演化分析，以及资讯汇集、智能选股、自动选股、联动委托交易等等，也因此分化出种种不同流派特点的炒股软件产品。

基本分析、技术分析、演化分析三种分析方法，基于完全不同的理论体系和逻辑结构，其主要研究对象，只侧重于市场运作的某一特定方面或者范畴，都有其合理性和局限性，但它们对于探讨股市博弈的方法和对策，又都具有参考价值。它们的理论基础、前提假设、范式定位、基本策略各不相同，在实际应用中它们既相互联系，又有重要区别。

其中，基本分析属于一般经济学范式，技术分析属于数理或牛顿范式，演化分析属于生物学或达尔文范式；基本分析主要应用于投资标的物的选择上，技术分析和演化分析则主要应用于具体操作的时机和空间判断上，作为提高股票投资分析有效性和可靠性的重要手段。具体来说，就是正确理解和综合应用基本分析、技术分析、演化分析三种方法，重点针对企业价值、成长性与安全边际，以及市场、制度、价值演化等方面，进行科学分析，客观判断，谨慎决策。

项目简介

本软件为证券决策系统，名为 Stockalize。依托于微信公众平台提供服务。制作人为武汉大学计算机学院卓工小组。

主要功能为

1. 提供股票历史走势
2. 根据模型预测股票走势
3. 同时给出对历史的预测和真实值的差异，直观提供准确性的比较
4. 后续会推出分层功能

项目人员

卓工实践小组

项目交付

所有源代码及公众号，文档。

第二部分

可行性研究和软件

开发计划

可行性分析(研究)报告(FAR)

说明：

- 1.《可行性分析(研究)报告》(FAR)是项目初期策划的结果，它分析了项目的要求、目标和环境；提出了几种可供选择的方案；并从技术、经济和法律各方面进行了可行性分析。可作为项目决策的依据。
- 2.FAR 也可以作为项目建议书、投标书等文件的基础。

目录

可行性分析(研究)报告(FAR)	Error! Bookmark not defined.
1 引言	Error! Bookmark not defined.
1.1 标识	Error! Bookmark not defined.
1.2 背景	Error! Bookmark not defined.
1.3 项目概述	Error! Bookmark not defined.
1.4 文档概述	Error! Bookmark not defined.
2 引用文件	Error! Bookmark not defined.
3 可行性分析的前提	Error! Bookmark not defined.
3.1 项目的要求	Error! Bookmark not defined.
3.2 项目的目标	Error! Bookmark not defined.
3.3 项目的环境、条件、假定和限制	Error! Bookmark not defined.
3.4 进行可行性分析的方法	Error! Bookmark not defined.
4 可选的方案	Error! Bookmark not defined.
4.1 原有方案的优缺点、局限性及存在的问题	Error! Bookmark not defined.
金融指标	Error! Bookmark not defined.
可视化:	Error! Bookmark not defined.
决策模型:	Error! Bookmark not defined.
4.2 可重用的系统，与要求之间的差距	Error! Bookmark not defined.
4.3 可选择的系统方案 1	Error! Bookmark not defined.
数据获取	Error! Bookmark not defined.
金融指标	Error! Bookmark not defined.
决策模型	Error! Bookmark not defined.
可视化	Error! Bookmark not defined.
4.4 可选择的系统方案 2	Error! Bookmark not defined.
可视化	Error! Bookmark not defined.
数据获取	Error! Bookmark not defined.
决策模型	Error! Bookmark not defined.
4.6 可选择的系统方案 3	Error! Bookmark not defined.
数据获取	Error! Bookmark not defined.
4.7 可选择的系统方案 4	Error! Bookmark not defined.
数据获取	Error! Bookmark not defined.
4.8 选择最终方案的准则	Error! Bookmark not defined.
5 所建议的系统	Error! Bookmark not defined.
5.1 对所建议的系统的说明	Error! Bookmark not defined.
金融指标	Error! Bookmark not defined.
数据获取	Error! Bookmark not defined.
5.2 数据流程和处理流程	Error! Bookmark not defined.
金融指标	Error! Bookmark not defined.
数据获取	Error! Bookmark not defined.
5.3 与原系统的比较(若有原系统)	Error! Bookmark not defined.
5.4 影响(或要求)	Error! Bookmark not defined.
5.4.1 设备	Error! Bookmark not defined.

<u>5.4.2 软件</u>	Error! Bookmark not defined.
<u>5.4.3 运行</u>	Error! Bookmark not defined.
<u>5.4.4 开发</u>	Error! Bookmark not defined.
<u>5.4.5 环境</u>	Error! Bookmark not defined.

1 引言

1.1 标识

Stocalize1.0

1.2 背景

随着金融市场的飞速发展，炒股的股民也越来越多。而炒股软件也就顺应这个时期而腾空而出。现在市场上布满了琳琅满目的股票软件，广告也铺天盖地，都说得神乎其神。既然是这个时期的产物，它就有用的价值。一份成功的股票软件方案是赢得订单的重要砝码，如何写一份令客户心动的股票软件开发方案则需要下苦功夫。

股票软件在市场上出现的价值:内容重点阐述了股票软件能给股民带来什么价值，以及通过何用方法去实现这种价值，从而帮助股民赚钱；另外，还有股民在使用软件的一些想法和顾虑的解决。软件开发的过程，要求和说明:这一项是一个整个软件的开发过程，中间需要各种材料，要求和细节的说明。是一个比较复杂和完整的过程。

从各方面具体的详解软件开发后期的说明:主要是对软件完工之后，双方各执行的要求，和补充说明。使软件在完工的基础山更加完善。股票软件在市场上出现的价值,股票软件带给股民的好处和其价值.

1.3 项目概述

本系统旨在向用户提供一个平台用于存储股票历史数据，随时查阅历史走势和其他一些分析数据，绘制一定时间段内的股票交易数据图表，根据深度学习决策模型预测股票走势，还具有动态添加删除股票等功能。实现了多用户账户以及用户账户登录管理流程。

本系统为初次开发，无历史版本信息。

1.4 文档概述

本文档通过对任务的分解，然后查阅相关资料，对各个模块的各方面可行性进行论证。

2 引用文件

1. Qasem A. Al-Radaideh Adel Abu Assaf and Eman Alnagi, Predicting Stock Prices using Data Mining Techniques, The International Arab Conference on Information Technology (ACIT'2013)

2. Gunter Senyurt, Abdulhamit Subasi, Stock market movement direction prediction using tree algorithms.
3. Sadegh Bafandeh Imandoust, Mohammed Bolandraftar, Forecasting the direction of stock market index movement using three data mining techniques: the case of Tehran Stock Exchange, Int. Journal of Engineering Research and Applications, ISSN : 2248-9622, Vol. 4, Issue 6(Version 2), June 2014, pp.106-117.
4. Rosillo, R.; Giner, J. & and De la Fuente, D. (2014): "Stock Market simulation using support vector machines", Journal of Forecasting, vol. 33, n° 6, (488–500)
5. GB T-8567-2006 计算机软件文档编制规范

3 可行性分析的前提

3.1 项目的要求

1. 金融指标：根据所选取的金融特征 MACD，将其作为指标，在以收集到的是够的历史数据基础上，通过算法预测股票的走势，即涨还是跌、幅度为多少。
2. 数据获取：能得到历史数据和实时数据。
3. 可视化：将所数据得到进行直观地展示
4. 决策模型：能对股票走势作出预测。

3.2 项目的目标

1. 金融指标：在金融指标 MACD 不变的情况下，尽量准确的预测股票走势，控制预测失误率在一定范围内，并且在预测出错时重新分析调整预测方向。
2. 可视化：受众能够通过各种可视化图标直观地得到想要的结果，使本项目能高效传递信息。
3. 决策模型：使用决策模型预测股票走势，准确率达到 70% 左右。

3.3 项目的环境、条件、假定和限制

本项目无资金支持，仅能使用市面上开源的，免费的软件或工具来达成目标。

3.4 进行可行性分析的方法

对市面上已有的工具和代码进行调查，通过分析对比优劣选出我们使用的工具。

4 可选的方案

4.1 原有方案的优缺点、局限性及存在的问题

金融指标

很多方案都采用 KDJ 作为金融指标，KDJ 指标是一种超前指标，运用上多以短线操作为主，KDJ 指标的超前主要体现在对股价的反应速度上，但由于其速度较快而往往造成频繁出现的买入卖出信号失误较多。

KDJ 随机指标是技术分析人员经常使用的一种指标，此种指标的优点在于反应灵敏，能给出非常明显的进货信号和出货信号，使用者易于掌握，只要看信号进出货就可以了。但 KDJ 指标又有非常明显不足，如它的反应过于敏感，使使用者不是进货太早被套牢，就是出货太早被轧空，也就是所谓 KDJ 指标的低位钝化和高位钝化问题。

可视化：

JavaScript 工具集

图表库

- C3 - 以 d3 为基础构建的可重用图表库
- Chart.js - 带有 canvas 标签的图表
- Chartist.js - 具有强大浏览器兼容能力的响应式图表
- Dimple - 适用于业务分析的面向对象的 API
- Dygraphs - 适用于大型数据集的交互式线性图表库
- Echarts - 针对大型数据集的高度定制化交互式图表
- Epoch - 可以完美创建的即时图表
- Highcharts - 基于 SVG 和 VML 呈现的图表库. 免费 (CC BY-NC 可用于非盈利项目)

- MetricsGraphics.js - 能够针对时间序列数据进行优化
- Morris.js - 非常漂亮的时间序列线状图
- NVD3 - 使用 d3.js 实现的可重用性图表库
- Peity - 可创建小型内连 svg 图表
- Plotly.js - 支持 20 种图表类型的强大的声明式库
- TechanJS - 股票以及金融图表

图形 (graphs) 图表库

- Cola.js - 通过基于约束的优化技术创建图表的工具. 和 d3 以及 svg.js 共同发挥作用
- Cytoscape.js - 由 Cytoscape 核心开发人员维护的用于绘制图形的 JavaScript 库.
- Linkurious - 一款基于 Sigma.js 的能加快图形可视化和交互式应用的开发速度的工具箱
- Sigma.js - 致力于图形绘制的 JavaScript 库
- VivaGraph - 针对 JavaScript 的图形绘制库

地图 (Maps)

- CartoDB - CartoDB 是一款开源工具并且其允许对网页上的地理数据进行存储和可视化
- Cesium - WebGL 虚拟地球仪和地图引擎
- Leaflet - 对移动端友好的交互式地图 JavaScript 库

- Leaflet Data Visualization Framework - 使用了 Leaflet 的设计用于简化数据可视化和主题映射的框架
- Mapael - 基于 the.js 的能够展示矢量地图的 jQuery 插件
- Mapsense.js - 将 d3.js 和瓷砖式覆盖地图结合起来
- Modest Maps - 使用 Javascript, 依照 BSD 许可的瓷砖覆盖式为基础展示和交互地图库
- dc.js 是一个多维图表构建工具, 其能够与 crossfilter 完美地配合工作
- angular-dc - dc.js 的 AngularJS 指令
- dc.leaflet.js - 使用了 Leaflet 地图的 dc.js 图表
- ember-dc - dc.js 的 Ember Component Wrappers (组件包装)

杂项

- Chroma.js - 用于处理色彩的小型库文件
- Piecon - 图标上的饼状图绘制工具
- Recline.js - 使用纯 JavaScript 和 HTML 的用于构建数据应用的简单而又强大的库
- Textures.js - 用于创建 SVG 模式的库
- Timeline.js - 创建交互式时间线
- Vega - Vega 是一个用于创建, 保存和分享交互式可视化数据设计的语法和格式
- Vis.js - 一个包括了时间线, 网络和图表 (2D 和 3D) 的动态可视化库

安卓工具

- HelloCharts - 兼容安卓 API 8 及以上版本的图表库
- MPAndroidChart - 一款功能强大而又易于使用的图表库

C++工具

- Visualization Toolkit (VTK) - 用于 3D 图形和图像处理和可视化的开源库

Go 语言工具

- Charts for Go - 基于 Go 的基础图表. 其能够对 ASCII、SVG 和 图像进行渲染
- svgo - 针对 SVG 的 Go 语言库

iOS 工具

- JBChartView - 同时支持线性和条形图的图表库
- PNChart - 使用了 Piner 和 CoinsMan 的简单并且美丽的图表
- ios-charts - MPAndroidChar 的 iOS 端口. 其可以使用非常简单的代码为两个平台创建图表

Python 工具

- bokeh - 用于 Python 的交互式网页绘图工具
- ggplot - 与 ggplot2 面向 R 语言的 API 相同
- glumpy - OpenGL 科学可视化库

- matplotlib - 2D 绘图库
- pygal - 一个动态 SVG 图表库
- PyQtGraph - 交互式和实时的 2D/3D/图像 绘制以及科学/工程工具
- seaborn - 一个能够制作极具吸引力的和展现翔实统计信息数据的图表库
- toyplot - 目标为大型数据图表的小型 Python 数据图表绘制工具
- Vincent - 面向 Vega 翻译器的 Python 工具
- VisPy - 基于 OpenGL 的高效科学可视化工具
- mpld3 - Matplotlib Graphics 的 D3 渲染工具

R 工具

- ggplot2 - 一个基于图形语法的绘图系统
- lattice - R 语言格子图形
- plotly - 交互式图表 (向 ggplot2 的输出中添加了交互性) , 统计图和简单网络图
- rbokeh - 针对 Bokeh 的 R 语言接口
- rgl - 使用了 OpenGL 的 3D 可视化
- shiny - 用于创建交互式应用和可视化的框架
- visNetwork - 交互式网络可视化

Ruby 工具

- Chartkick - 使用 Ruby 的单线创建图表的工具

其他工具

不与特定平台或语言绑定的工具

- Charted - 一个能够从任何数据文件中创建自动化，可分享的图表的工具
- Gephi - 一个用于可视化和制作大型图表的开源平台
- Lightning - 一个提供以 API 为基础的方式获取可再生，网络为基础的交互式可视化图表的数据可视化服务
- RAW - 由 CSV 和 Excel 文件创建的网络可视化工具
- Spark - 命令解释程序 (shell) 走势图. 其包含多种实现语言
- Periscope - 直接由 SQL 查询语句创建图表

决策模型：

在 20 世纪 70 年代后期和 20 年代 80 初期，机器学系研究人员 J.Ross Quinlan 开发了决策树算法，称为迭代的二分器 (ID3)。Quinlan 后来提出了 C4.6，成为新的监督学习算法的性能比较基准。1984 年，对位统计学家出版了著作《Classification and Regression Trees》(Cart)，介绍了二叉决策树的产生。决策树使用类似于流程图的树结构，其中每个内部节点表示在一个属性上的测试，每个分支代表该测试的一个输出，而每个叶子结点存放一个最终的决策。

许多类型的基于规则的决策树模型已被广泛用于预测金融时间序列。Qasem A AL Radaideh, Adel Abu Assaf 和 Eman Alnagi 研究了 ID3 和 C4.5 基于六个响应变量属性的决策树的分类模型的效用；他们将每个属性数值转换为离散值 (Positive / Negative / Equal)。他们分析了安曼证券交易所 (ASE) 三大公司投资者的交易行为。他们的研究结果显示，这些模型的分类器精度在 44%-54% 之间。Gunter Senyurt, Abdulhamit Subasi 研究了决策树模型的影响，并用 ID3, C4.5 和随机森林方法预测伊斯坦布尔证券交易所 10 年期指数 (XU-100) 的走势。他们部署了 10 个技术指标并得出结论，决策树提供了大约 77% ~ 78% 的准确性来正确分类实例。Sadegh Bafandeh Imandoust, Mohammed Bolandraftar 试图开发三种模型，并比较其在德黑兰证券交易所 (TSE) 指数中股价走势方向的预测。他们使用决策树，随机森林和朴素贝叶斯分类器分类技术，采用十个微观经济变量和三个宏观经济变量作为输入。十个微观经济变量是：10 天 SMA, 10 天 WMA, 动量, 随机 %K, 随机 %D, RSI (相对强弱指数), MACD (移动平均收敛差异), Larry William 的 %R, A/D 累积/分配 振荡器和 CCI。由于石油，黄金和美元/INR 在伊朗经济中发挥突出作用，所以考虑了这些变量作为基本指标。实验结果表明，决策树模型 (80.08%) 超过其他两个模型随机森林 (78.8%) 和朴素贝

叶斯分类器（73.8%）。

决策树模型的构造不需要任何领域知识或者参数设置，因此适合与探索式知识的发现。决策树可以处理高位数据。获取的知识用树的形式表示是直观的，并且容易被人理解。决策树归纳的学习和分类步骤是简单快速的。一般而言决策树分类器具有很好的准确率。

决策树的缺陷在于：容易过拟合，泛化能力差，对异常值过于敏感，很容易导致树的结构的巨大的变换。

支持向量机的第一篇论文由 Vladimir Vapnik 和他的同事 Bernhard Boser 以及 Isabelle Guyon 于 1992 年发表。SVM 是一种对线性和非线性数据进行分类的方法，它使用一种非线性映射，那原训练数据映射到较高的维上。在新的维上，它搜索最佳分离超平面。SVM 可以用于数值预测和分类。

最近有报道说使用 SVM 来解决金融预测问题。2003 年，出现了 SVM 金融时间序列预测的两个应用：Cao 和 Tay (2003)，将 SVM 应用于芝加哥商品市场预测若干期货合约的问题，显示出 SVM 优于 BP 神经网络；在 Kim (2003) 中，SVM 用于预测韩国综合股票指数的变化方向，并以 BP 神经网络和 Case Base Reasoning 为基准。实验结果表明，SVM 优于其他方法，应被视为金融时间序列预测的有前途的方法。Huang et al (2005) 的工作中，支持向量机分类器用于预测 Nikkei225 指数的定向运动，具有非常好的结果。Ince 和 Trafalis (2006) 也尝试使用 SVM 来解决组合问题优化问题。

最后，Lee (2009) 解释了基于 SVM 的预测模型，并采用混合特征选择来预测股市走势。表明 SVM 在股票趋势预测问题上优于 BP 神经网络。Dunis 等人 (2013a) 表明，可以使用 SVM 分类器在某些选择的风险规避参数下预测一段时间的 IBEX-35 指数。Dunis 等人 (2013b) 的工作中，遗传算法用于优化输入选择过程和 SVM 模型的参数。该方法适用于使用 FTSE100 和 ASE20 指数的前一天预测和交易问题。新的面向金融的拟合函数加上确认过滤器和杠杆技术来提高整体方法的性能。实验结果表明，该方法优于 MACD，ARMA 模型，贝叶斯预测器和高阶神经网络等经典技术。

SVM 也有不可避免的缺陷：对缺失数据敏感、对非线性问题没有通用解决方案，必须谨慎选择核函数来处理，即使确定核函数以后，在求解问题分类时，要求解函数的二次规划，这就需要大量的存储空间。

4.2 可重用的系统，与要求之间的差距

由于无历史版本，此项为空

4.3 可选择的系统方案 1

数据获取

1. 从第三方软件导出
 - a) 通达信

将股票数据导出的前提必须将数据从服务器下载到本地，首先打开通达信股票软件的[系统]菜单，选择[盘后数据下载]选项。在弹出的盘后数据下载界面中，选择沪深日线页面以便下载日线数据。勾选[日期和实时行情数据]选项可选择下载数据的

时间范围。勾选[下载所有 A 股类品种的日线数据]选项则代表下载所有 A 股的数据，如只想下载特定的品种，可以反选此项并添加您想下载的品种代码即可。

- b) 同花顺
提供实时盯盘信息

金融指标

在应用上，MACD 应先计算出快速（一般选 12 日）移动平均值与慢速（一般 26 日）移动平均值，以这两个数值作为测量两者（快速与慢速线）间的“差离值”依据。所谓差离值（DIF），即 12 日 EMA 数值减去 26 日 EMA 数值。在持续的涨势中，正差离值（+DIF）会愈来愈大；反之，在跌势中，差离值会变负，并且也会越来越大。当正或负差离值缩小到一定程度的时候，便是真正的行情反转信号。

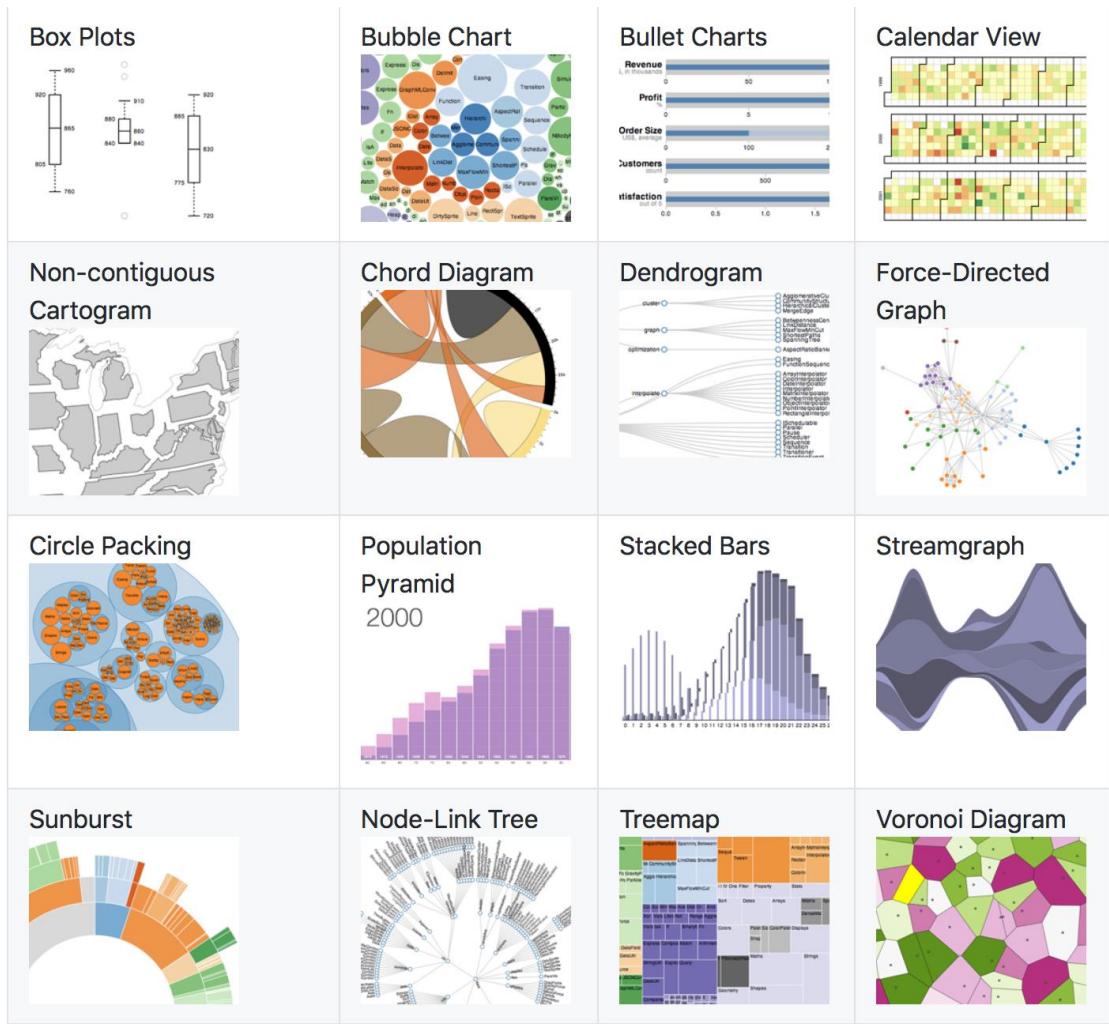
决策模型

将目标作为二分类问题，预测下一个时间段内数据的走向（涨/跌）。将时间序列数据转化为图像（矩阵），使用 CNN 训练数据并预测。

可视化

D3.js

作为多个可视化工具的基石，D3 的全称是 (Data-Driven Documents)，是一个被数据驱动的文档，使用它主要是用来做数据可视化的。D3 提供了各种简单易用的函数，大大简化了 JavaScript 操作数据的难度。由于它本质上是 JavaScript，所以用 JavaScript 也是可以实现所有功能的，但它能大大减小你的工作量，尤其是在数据可视化方面，D3 已经将生成可视化的复杂步骤精简到了几个简单的函数，你只需要输入几个简单数据，就能够转换为各种绚丽的图形。



4.4 可选择的系统方案 2

可视化

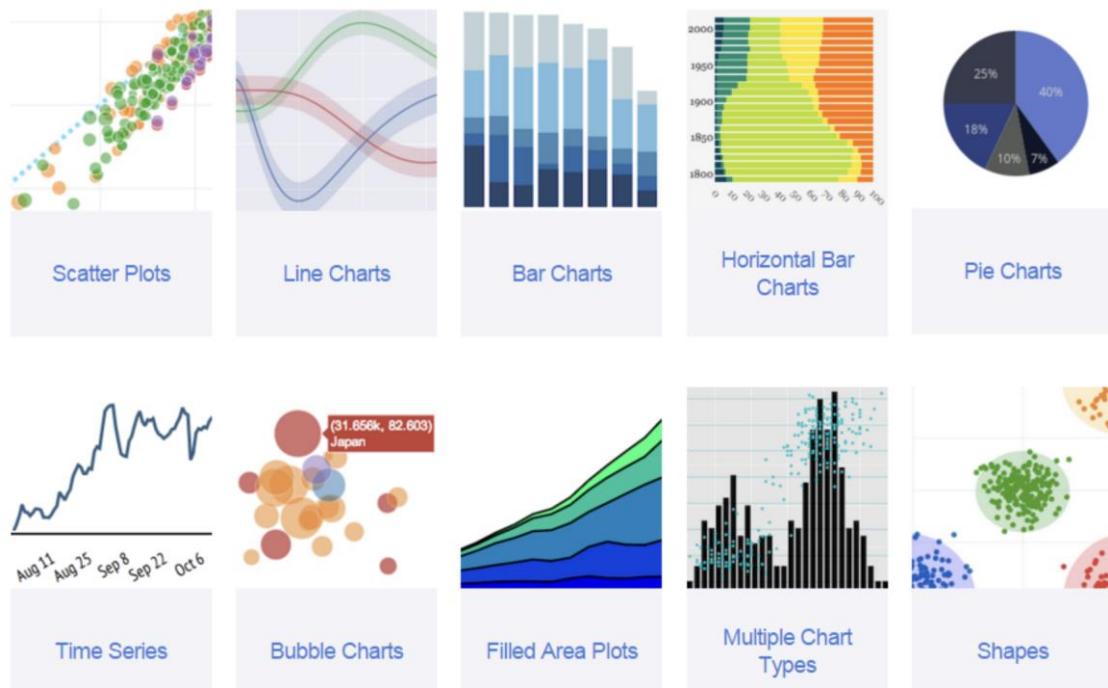
Plotly.js

数据可视化平台 Plotly 开源旗下强大的 JavaScript 图表库，支持三种不同类型的图表，包括地图，箱形图和密度图，以及更常见的产品如，条状和线形图。此前 Plotly 在 R, Python 和 MATLAB 的客户端一直是开源的，但核心图表层 plotly.js 却是闭源的。而现在 Plotly 可以 100% 离线地用于 RStudio, MATLAB 或 Python Jupyter。

Plotly 在声明中称，和许多 JavaScript 数据可视化库不同， Plotly.js 不依赖于 jQuery。这使得 plotly.js 性能显著超过其它竞品。plotly.js 采用 JSON 的图表规范来制做交互可视

化，所以从一种格式转换到另一种更加简单，比如将 CSV 文件转成 Excel 图表，Python 代码或交互图表，还有 R 代码。

Plotly 基于 D3.js、stack.gl（WebGL 组件库）和 SVG，用 JavaScript 在网页上实现了类似 MATLAB 和 Python matplotlib 的图形展示功能，支持 20 种图形，包括 2D 和 3D，交互性很流畅，足以满足一般性科学计算的需要。



数据获取

从 web 页面提取

使用 python，很容易写出多线程的抓取网页程序。

线程库：threading

网页抓取：urllib2

XML 解析：xml.dom

HTML 解析：SGMLParser

a) 东方财富网：<http://quote.eastmoney.com/stocklist.html>

b) yahoo 财经：<https://hk.finance.yahoo.com/>

提供“XML 数据下载”，比分页的 HTML 更方便解析。

c) 新浪股票：<http://finance.sina.com.cn/realstock/company/sz000877/nc.shtml>

d) 百度股市通：<https://gupiao.baidu.com/stock/>

个股信息 <https://gupiao.baidu.com/stock/sz002439.html>

所以只要百度股票网的网址+每只股票的代码即可，而每只股票的代码我们已经有前面的程序 getStockList 从东方财富网解析出来了，因此对 getStockList 函数返

回的列表进行遍历即可，代码如下：

```
for stock in lst: url = stockURL + stock + ".html"
```

决策模型

将目标作为二分类问题，使用 RNN 训练预测下一个数据的趋向（涨/跌）。

4.6 可选择的系统方案 3

数据获取

基于图像识别的方法

a) 针对图片中的文字数据

使用开源的 OCR 工具 ImageMagick、tesseract-ocr

ImageMagick 是一套功能强大、稳定而且开源的工具集和开发包，可以用来读、写和处理超过 89 种基本格式的图片文件，可以自由使用，复制，修改，发布。支持大多数的操作系统。

Tesseract：开源的 OCR 识别引擎，初期 Tesseract 引擎由 HP 实验室研发，后来贡献给了开源软件业，后经由 Google 进行改进，消除 bug，优化，重新发布

mini_magick 是用 ruby 对 ImageMagick 命令的封装

b) 针对股票走势曲线图

首先使用 ImageMagick 将爬到的图片归一化处理

再利用 getdata 或者 origin 的 digitizer 插件，从图表中提取数据。

4.7 可选择的系统方案 4

数据获取

调用第三方 API

a) 阿里云股票查询 API

价格适中，但该接口主要用于分析统计、模拟学习、复盘、资料查阅，不具有实时性，不适用于交易及投资建议。

b) Sina 股票数据 API

不提供历史价格，只有当前价格，实时性强。

c) TuShare

TuShare 是一个免费、开源的 python 财经数据接口包。考虑到 Python pandas 包在金融量化分析中体现出的优势，TuShare 返回的绝大部分的数据格式都是 pandas DataFrame 类型，非常便于用 pandas/NumPy/Matplotlib 进行数据分析和可视化。

d) vn.py - 基于 python 的开源交易平台开发框架

封装了丰富的 Python 交易和数据 API 接口，基本覆盖了国内外常规交易品种（证券、期货、期权、外汇、CFD）

4.8 选择最终方案的准则

可视化工具应具有优秀的绘图响应速度及多种图形支持、具备高度可移植性、跨平台适用性、工具本身开源避免商业纠纷。

金融指标由于 MACD 反应是平均值，是市场的整体趋势，在中短期上的发挥很稳定，不会有大问题发生，所以选择 MACD。

数据获取应确保数据准确，数据具有实时性，可以依据需求选取不同数据输出。

5 所建议的系统

5.1 对所建议的系统的说明

金融指标

MACD 称为指数平滑移动平均线，是从双指数移动平均线发展而来的，由快的指数移动平均线（EMA12）减去慢的指数移动平均线（EMA26）得到快线 DIF，再用 $2 \times$ （快线 DIF-DIF 的 9 日加权移动均线 DEA）得到 MACD 柱。MACD 的意义和双移动平均线基本相同，即由快、慢均线的离散、聚合表征当前的多空状态和股价可能的发展变化趋势，但阅读起来更方便。当 MACD 从负数转向正数，是买的信号。当 MACD 从正数转向负数，是卖的信号。当 MACD 以大角度变化，表示快的移动平均线和慢的移动平均线的差距非常迅速的拉开，代表了一个市场大趋势的转变。

数据获取

使用开源的 TuShare 接口包

5.2 数据流程和处理流程

金融指标

首先计算出快速移动平均线（即 EMA1）和慢速移动平均线（即 EMA2），以此两个数值，来作为测量两者（快慢速线）间的离差值（DIF）的依据，然后再求 DIF 的 N 周期的平

滑移动平均线 DEA (也叫 MACD、DEM) 线。

以 EMA1 的参数为 12 日 EMA2 的参数为 26 日, DIF 的参数为 9 日为例来看看 MACD 的计算过程

1、计算移动平均值 (EMA)

12 日 EMA 的算式为

$EMA(12) = \text{前一日 } EMA(12) \times 11/13 + \text{今日收盘价} \times 2/13$

26 日 EMA 的算式为

$EMA(26) = \text{前一日 } EMA(26) \times 25/27 + \text{今日收盘价} \times 2/27$

2、计算离差值 (DIF)

$DIF = \text{今日 } EMA(12) - \text{今日 } EMA(26)$

3、计算 DIF 的 9 日 EMA

根据离差值计算其 9 日的 EMA, 即离差平均值, 是所求的 MACD 值。为了不与指标原名相混淆, 此值又名

DEA 或 DEM。

今日 $DEA(MACD) = \text{前一日 } DEA \times 8/10 + \text{今日 } DIF \times 2/10$ 计算出的 DIF 和 DEA 的数值均为正值或负值。

用 $(DIF-DEA) \times 2$ 即为 MACD 柱状图。

数据获取

import tushare as ts 后可使用

ts.get_hist_data()

ts.get_stock_basics()

等接口获得数据, 再使用

$df = ts.get_hist_data('000875')$ #直接保存

$df.to_excel('c:/day/000875.xlsx', startrow=2,startcol=5)$ #插入 excel 中

$df.to_json('c:/day/000875.json',orient='records')$ #保存为 json 格式

等接口保存即可。

5.3 与原系统的比较(若有原系统)

无原系统。

5.4 影响(或要求)

无。

5.4.1 设备

PC

5.4.2 软件

Python 2.7 + Cuda 8 + Pytorch

5.4.3 运行

Python 2.7 + Pytorch

5.4.4 开发

Python

5.4.5 环境

Ubuntu 16.04

Pandas 0.21.0

Anaconda 5.0.1

Tushare 0.8.6

5.5 局限性

只能使用市面上非商用的工具。

6 经济可行性(成本----效益分析)

6.1 投资

服务器投资 60 元。

7 技术可行性(技术风险评价)

本团队人员均为武汉大学计算机学院本科生，本项目旨在增强同学们团队合作与开发能力，所以应进行学习与开发。

软件开发计划(SDP)

项目名称: Stockalize(股票分析软件)

组长: 王致远

组员: 罗格, 王一舟, 李文洲, 王睿达

说明:

1. 《软件开发计划》(SDP)描述开发者实施软件开发工作的计划, 本文档中“软件开发”一词涵盖了新开发、修改、重用、再工程、维护和由软件产品引起的其他所有的活动。
2. SDP 是向需求方提供了解和监督软件开发过程、所使用的方法、每项活动的途径、项目的安排、组织及资源的一种手段。
3. 本计划的某些部分可视实际需要单独编制成册, 例如, 软件配置管理计划、软件质量保证计划和文档编制计划等。

目录

<u>软件开发计划(SDP)</u>	26
<u>1 引言</u>	29
<u>1.1 标识</u>	29
<u>1.2 系统概述</u>	29
<u>1.3 文档概述</u>	29
<u>2 引用文件</u>	29
<u>3 交付产品</u>	30
<u>3.1 程序</u>	30
<u>3.2 文档</u>	30
<u>3.3 服务</u>	30
<u>3.4 最后交付期限</u>	30
<u>4 所需工作概述</u>	30
<u>5 实施整个软件开发活动的计划</u>	31
<u>5.1 软件开发过程</u>	31
<u>5.2 软件开发总体计划</u>	31
<u>5.2.1 软件开发方法</u>	31
<u>5.2.2 软件产品标准</u>	31
<u>6 实施详细软件开发活动的计划</u>	32
<u>6.1 项目计划和监督</u>	32
<u>6.1.1 软件开发计划(包括对该计划的更新)</u>	32
<u>6.1.2 CSCI 测试计划</u>	32
<u>6.1.3 系统测试计划</u>	32
<u>6.2 建立软件开发环境</u>	32
<u>6.2.1 软件工程环境</u>	32
<u>6.2.2 软件测试环境</u>	32
<u>6.2.3 软件开发库</u>	33
<u>6.3 系统需求分析</u>	33
<u>6.3.1 用户输入分析</u>	33
<u>6.4 软件设计</u>	33
<u>6.4.1 CSCI 级设计决策</u>	33
<u>6.4.2 CSCI 体系结构设计</u>	33
<u>6.4.3 CSCI 详细设计</u>	33
<u>6.5 软件实现和配置项测试</u>	34
<u>6.5.1 软件实现</u>	34
<u>6.5.2 配置项测试准备</u>	34
<u>6.5.3 配置项测试执行</u>	34
<u>6.5.4 修改和再测试</u>	34
<u>6.5.5 配置项测试结果分析与记录</u>	35
<u>6.6 配置项集成和测试</u>	35
<u>6.6.1 配置项集成和测试准备</u>	35
<u>6.6.2 配置项集成和测试执行</u>	35
<u>6.6.3 修改和再测试</u>	35

<u>6.6.4 配置项集成和测试结果分析与记录</u>	35
<u>6.7 软件移交准备</u>	35
<u>6.7.1 可执行软件的准备</u>	35
<u>6.7.2 源文件准备</u>	35
<u>6.7.3 支持现场的版本说明的准备</u>	35
<u>6.7.4“已完成”的 CSCI 设计和其他的软件支持信息的准备</u>	36
<u>6.7.5 系统设计说明的更新</u>	36
<u>6.7.6 支持手册准备</u>	36
<u>6.7.7 到指定支持现场的移交</u>	36
<u>6.8 软件配置管理</u>	36
<u>6.8.1 配置标识</u>	36
<u>6.8.2 配置状态统计</u>	36
<u>6.9 软件质量保证</u>	37
<u>6.9.1 软件质量保证评估</u>	37
<u>6.10 问题解决过程(更正活动)</u>	37
<u>6.10.1 问题/变更报告</u>	37
<u>6.10.2 更正活动系统</u>	37
<u>7 进度表和活动网络图</u>	38
<u>8 项目组织和资源</u>	39
<u>8.1 项目组织</u>	39
<u>8.2 项目资源</u>	39
<u>9 培训</u>	39
<u>9.1 项目的技术要求</u>	39
<u>9.2 培训计划</u>	39
<u>10 项目估算</u>	40
<u>10.1 规模估算</u>	40
<u>10.2 工作量估算</u>	40
<u>12 支持条件</u>	40
<u>12.1 计算机系统支持。</u>	40
<u>12.2 需要需方承担的工作和提供的条件。</u>	40

1 引言

1.1 标识

软件名字: Stocalize

版本号: 1.0.0

1.2 系统概述

本系统旨在向用户提供一个平台用于存储股票历史数据，随时查阅历史走势和其他一些分析数据，绘制一定时间段内的股票交易数据图表，根据深度学习决策模型预测股票走势，还具有动态添加删除股票等功能。实现了多用户账户以及用户账户登录管理流程。

本系统为初次开发，无历史版本信息。

本系统暂时不适用投资方、需方、用户、开发方和支持机构等信息。

本系统运行现场为个人 PC。

其他相关文档：可行性分析报告。

1.3 文档概述

本文档用于详细描述软件开发计划，以及软件开发计划的具体实施过程。

2 引用文件

6. Qasem A. Al-Radaideh Adel Abu Assaf and Eman Alnagi, Predicting Stock Prices using Data Mining Techniques, The International Arab Conference on Information Technology

(ACIT'2013)

7. Gunter Senyurt, Abdulhamit Subasi, Stock market movement direction prediction using tree algorithms.
8. Sadegh Bafandeh Imandoust, Mohammed Bolandraftar, Forecasting the direction of stock market index movement using three data mining techniques: the case of Tehran Stock Exchange, Int. Journal of Engineering Research and Applications, ISSN : 2248-9622, Vol. 4, Issue 6(Version 2), June 2014, pp.106-117.
9. Rosillo, R.; Giner, J. & and De la Fuente, D. (2014): "Stock Market simulation using support vector machines", Journal of Forecasting, vol. 33, nº 6, (488–500)
10. GB T-8567-2006 计算机软件文档编制规范

3 交付产品

3.1 程序

名为 Stocalize 的 Windows 下 App，该产品为需要开发和维护的产品。

3.2 文档

用户使用手册

3.3 服务

本产品的服务主要为产品的维护服务，包括 bug fix 等

3.4 最后交付期限

2017-2018 年上学期结束前

4 所需工作概述

概述如下：

- A. 该项目开发的需求：能够存储股票历史数据，查阅历史走势和分析数据，绘制股票交易数据图表，预测股票走势，动态添加删除股票，多用户账户以及用户账户登录管理流程。
- B. 对项目文档编制的需求：文档按照 GB+T-8567-2006 计算机软件文档编制规范编写；文档

- 必须保证原始性和真实性，归档文件必须是原件；必须与原工程实际相符。
- C. 开发计划在系统生命周期中为开发阶段。
 - D. 本项目暂无采购策略或选用计划。
 - E. 本项目按照项目开发指导老师的安排来控制进度。资源约束不适用。
 - F. 其他需求和约束暂时不适用。

5 实施整个软件开发活动的计划

5.1 软件开发过程

开发过程包括初期的需求分析和软件开发计划的讨论和编写，中期的软件开发的具体分工实施，后期的软件对接以及调试，最后的软件提交。

5.2 软件开发总体计划

5.2.1 软件开发方法

本软件采用面向对象的开发方式（OMT），采用先自底向上归纳，再自顶向下分解。同时根据需求分析建立对象模型的建立。

本软件也采用可视化开发方法，以事件为驱动，响应函数响应的模式来完成可视化方面的编程。

5.2.2 软件产品标准

- A. 格式标准：缩进按照 `tab=4space`，函数名和变量名需要首字母大写。
- B. 注释标准：不同的数据模块和函数模块都需要在函数名前一行给出相应的注释，复杂的数据模块需要对不同的模块内变量给出注释，函数模块需要多特定的算法代码给出注释。
- C. 其他注释标准：暂无
- D. 命名约定：参数、程序包、过程尽量采用英文全称，首字母大写，变量方面遇到复杂变量名时可以采用缩写
- E. 编程语言或功能的使用限制：使用 `python` 语言，其他限制暂无
- F. 代码聚合复杂性的制约：暂无。

6 实施详细软件开发活动的计划

6.1 项目计划和监督

6.1.1 软件开发计划(包括对该计划的更新)

1. 完成决策模型的实现
2. 使用已有的数据对模型进行训练
3. 编写对金融数据的爬取
4. 绘制并编写 GUI

6.1.2 CSCI 测试计划

- 使用已有数据验证模型的准确度
- 测试 GUI 是否用户友好
- 测试软件的稳定性

6.1.3 系统测试计划

完整功能测试，检查软件是否达到预期的准确度。记录软件一个月在真实数据下的预测数据并与真实结果做对比。

6.2 建立软件开发环境

6.2.1 软件工程环境

软件文档编写工具: Word/WPS

版本控制工具: Git

软件开发集成环境: Pycharm、Webstorm、Android Studio

绘图工具: Microsoft Office Visio

6.2.2 软件测试环境

服务器软件环境: Ubuntu 16.04 + Anaconda 2 + Pytorch + Scrapy + Apache

客户端软件环境: Windows 7 及以上/OS X/Linux/Android

服务器硬件环境: Intel Core i7 + NVIDIA Geforce GTX 970M + 8G RAM

6.2.3 软件开发库

- 决策模型输入输出
- GUI 交互
- 服务器数据提交与结果返回
- 数据爬取模块

6.3 系统需求分析

6.3.1 用户输入分析

用户通过对 GUI 的操作获取某一支股票的预测结果。或者让软件推荐若干走势较好的股票以供选择。

6.4 软件设计

6.4.1 CSCI 级设计决策

软件分为 4 个模块：

决策预测模块、数据爬取模块、后端 API 模块、前端模块。

6.4.2 CSCI 体系结构设计

前端向后端 API 发起请求，数据爬取模块从互联网上爬取相应的数据交由决策预测模块进行分析预测，将结果通过 API 返回至前端。

6.4.3 CSCI 详细设计

前端使用 Web/Android 开发。后端使用 Python 作为服务端应用编写，主要的 API 是通过 GET 方法输入一个股票代码，返回一个预测结果。数据爬取模块输入一个股票代码，向国内的股票网站爬取对应股票一段时间的历史价格数据作为输出。决策预测模块输入一段时间序列数据，经过 CNN/RNN 得到一个未来一段时间内的总体走势（涨/跌）作为输出结果。

6.5 软件实现和配置项测试

6.5.1 软件实现

实现 Stockalize 的基本功能，即预测，实时显示图表等。

6.5.2 配置项测试准备

功能测试是对软件需求规格说明中的功能需求项逐项进行测试，以验证其功能是否满足要求，功能测试一般需进行：

- 1) 每一个软件功能应被一个测试用例或一个认可的异常所覆盖，对大的功能应进一步分解为更细的功能，使测试用例可以直接和功能对应；
- 2) 用基本数据类型和数据值测试；
- 3) 用一系列合理的数据类型和数据值运行，测试超负荷、饱和及其它“最坏情况”的结果；
- 4) 用假想的数据类型和数据值运行，测试排斥不规则输入的能力；
- 5) 每个功能的合法边界和非法边界值都应被作为测试用例；
- 6) 应考虑软件功能对操作模式、运行环境、运行状态、状态转换、运行时间等的覆盖要求；
- 7) 对于在需求规格说明中没有指明，而在用户手册、操作手册中表明出来的每一功能及操作，都应有相应测试用例覆盖。

性能测试

性能测试是对软件需求规格说明中的性能需求逐项进行的测试，以验证其性能是否满足要求。性能测试一般需进行：

- 1) 测试程序在获得定量结果时程序计算的精确性（处理精度）。
- 2) 测试程序在有速度要求时完成功能的时间（响应时间）；
- 3) 测试程序完成功能所能处理的数据量；
- 4) 测试程序各部分的协调性、如高速、低速操作的协调；
 - 5) 测试软/硬件中因素是否限制了程序的性能；
 - 6) 测试程序的负载潜力；
 - 7) 测试程序运行占用的空间。

6.5.3 配置项测试执行

运行软件，查看是否满足要求。并查看文档是否满足规范。

6.5.4 修改和再测试

若发现软件或文档有错误，则需要重新修改并重新进行测试。

6.5.5 配置项测试结果分析与记录

将结果与预期目标对比，并记录在测试文档中。

6.6 配置项集成和测试

6.6.1 配置项集成和测试准备

单个配置项测试完毕后，接着对集成后的配置项进行测试，编写测试模块代码。

6.6.2 配置项集成和测试执行

使用手动和自动方法进行场景测试，功能测试，系统测试。

6.6.3 修改和再测试

如在测试期间发现了错误，则重新修改后重新测试。

6.6.4 配置项集成和测试结果分析与记录

将结果与预期目标对比，并记录在测试文档中。

6.7 软件移交准备

6.7.1 可执行软件的准备

一台 PC 机。

6.7.2 源文件准备

打包好的源程序文件，报告可安装的文件。

6.7.3 支持现场的版本说明的准备

本项目小组提供的最新版本。

6.7.4 “已完成”的 CSCI 设计和其他的软件支持信息的准备

提交给客户设计报告。

6.7.5 系统设计说明的更新

由后期维护阶段具体说明。

6.7.6 支持手册准备

用户手册的准备，本组提供的用户手册。

6.7.7 到指定支持现场的移交

由本小组组长和客户经理双方进行移交。

6.8 软件配置管理

6.8.1 配置标识

Stocalize1.0

6.8.2 配置状态统计

为跟踪工作产品基线，配置管理者需手机下列信息：

- 基线类型
- 工作产品名称
- 配置项名称/标识符
- 版本号
- 更改日期/时间
- 更改请求列表
- 需要更改的配置项
- 当前状态
- 当前状态发生日期

项目组每周提交配置项清单及其当前版本。配置管理人员每半个月提交变更请求的状态统计。

6.9 软件质量保证

项目质量评审报告将根据各阶段质量检查点所应检查的内容作出质量上的评审。对在质量评审工作中发现的问题，将根据其性质、范围划分级别。相关人员以此为依据来确定缺陷修复的优先级，缺陷严重程度高的修复优先级也高。

项目质量评审报告将交由项目管理办公室讨论，质量保证经理与项目管理办公室将对项目质量评审报告中需要明显改善的质量问题作出具体地改进方案，以及时间上和人员上的安排。项目质量评审报告的内容如下，其中问题列表的序号编码方式为 XX-YYY。XX 表示项目质量评审阶段，YYY 代表问题序号。

6.9.1 软件质量保证评估

项目经理和 QA 在项目初期对项目交付文档的模板进行检查和审核，确保项目交付品的架构和内容

大纲的完整性和正确性。在项目进行过程中，QA、项目经理及各项目功能小组组长会分别对交付品进度

6.10 问题解决过程(更正活动)

本条应分成若干分条描述软件更正活动中要遵循的方法.各分条的计划应覆盖合同中论及它的所有条款。

6.10.1 问题/变更报告

它包括要记录的具体条目(可选的条目包括：项目名称，提出者，问题编号，问题名称，受影响的软件元素或文档，发生日期，类别和优先级，描述，指派的该问题的分析者，指派日期，完成日期，分析时间，推荐的解决方案，影响，问题状态，解决方案的批准，随后的动作，更正者，更正日期，被更正的版本.更正时间，已实现的解决方案的描述)。

6.10.2 更正活动系统

系统质量检查中出现问题时，遵循逐级上报的原则：

项目组成员级不能解决的上报到组长，由其协调解决；组长不能解决的上报到项目经理。各子系统开发组组长在每个软件开发生命周期阶段结束后，总结本子系统的软件质量状况和质量目标的实现情况，并填写《质量总结报告》。

项目经理每周整理一份项目状态报告，记录一周来项目的进展情况和出现的问题，如果这些问题还能被解决，将被移交到更高一层的领导者协调处理。

必要时，对出现的问题将予以记录(如评审缺陷清单、遗留问题清单等)，交配置管理工程师备案。质量管理组将对问题追踪进行监督，确保问题关闭。

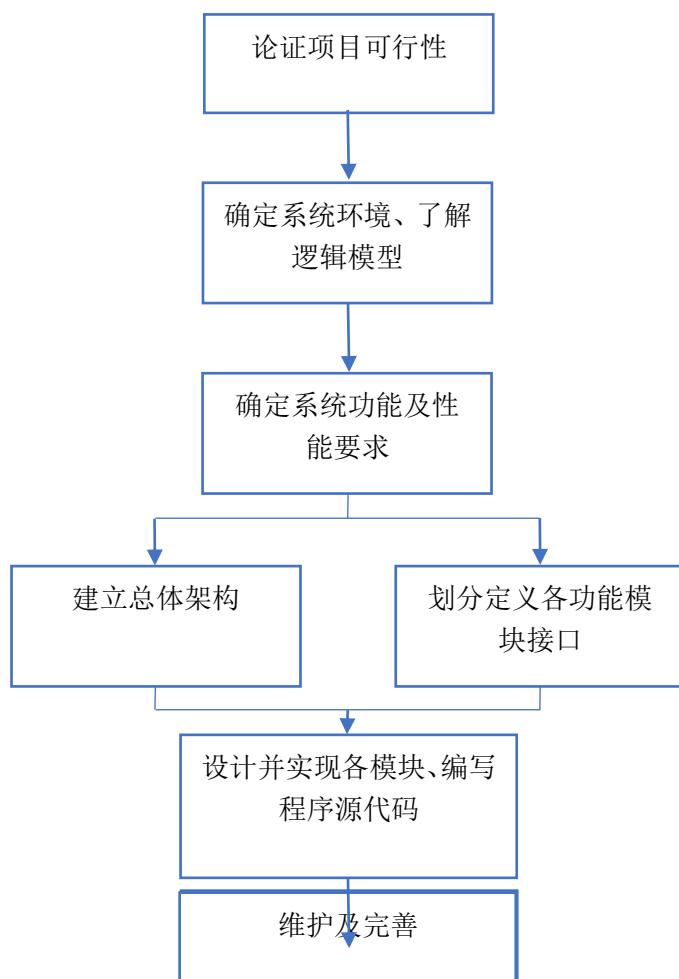
对在质量检查工作中发现的问题，将根据其性质、范围和对软件的影响程度划分级别。相关人员以此为依据来确定缺陷修复的优先级：缺陷严重程度高的修复优先级也高；同理，回归测试的优先级也依此确定。

7 进度表和活动网络图

a. 进度表

计划内容	提交草稿	时间轴（周）							
		1	2	3	4	5	6	7	8
论证项目可行性	项目初步开发计划进度表	■							
确定系统环境、了解逻辑模型	需求规格说明、项目开发计划								
确定系统功能及性能要求	项目开发计划		■	■					
建立系统总体架构、划分定义各功能模块接口	数据库设计说明书			■	■				
设计并实现各模块、编写程序源代码	用户手册			■	■	■	■		
集成测试、验收测试、进行开发总结	开发工作总结							■	■
维护纠正错误、完善应用	修订用户手册							■	■

b. 活动网络图



8 项目组织和资源

8.1 项目组织

本项目由金融工程项目开发团队完成。

8.2 项目资源

1. 本项目将投入人力共有金融工程项目开发团队的 5 名成员；
2. 本项目 5 名成员共同承担以下软件开发过程中：管理，软件工程，软件测试，软件配置管理，软件产品评估，软件质量保证和软件文档编制等；
3. 开发人员使用设施：5 台个人电脑；
4. 执行工作地理位置：湖北省武汉市武汉大学信息学部。

9 培训

9.1 项目的技术要求

开发技术：熟悉 Python、C++、Java 及 Web 前端开发

管理技术：利用软件工程相关知识进行项目管理

9.2 培训计划

本项目需对成员进行金融分析相关培训。拟安排为期 2-3 周金融预测模型相关学习。

10 项目估算

10.1 规模估算

本项目主要实现以下功能点：金融数据获取、金融数据存储、数据量化决策、交易决策、可视化输出、软件界面设计实现。其编码工作量预估占总规模 60%。预计总代码量 15000LOC。

10.2 工作量估算

本项目开发周期预计为 60 天，人均月代码行 $15000/60=250$ LOC/人月

12 支持条件

12.1 计算机系统支持。

本项目预计主要开发环境为 Microsoft Windows，运行环境为 Windows、Ubuntu、Android。
编码、测试及运行环境需求：

- (1) 网站环境
 - a. 处理器型号：
 - i. Intel® Pentium® IV Processor 或同等级及以上桌面级 CPU;
 - ii. Qualcomm® Snapdragon® 200 或同等级及以上移动级 CPU。
 - b. 浏览器内核：
 - i. Chrome 21+
 - ii. Opera 12.1+
 - iii. Firefox 22+
 - iv. Safari 6.1+
 - v. IE 10+
 - c. 输入输出设备：
 - i. 键盘及鼠标或触控设备
 - ii. 显示器及打印机等
- (2) APP 环境
 - Android 4.0.1 及以上

12.2 需要需方承担的工作和提供的条件。

需方应提供使用运行时 Ubuntu 及 Android 环境，运行环境需求：

- (1) 网站运行环境
 - a. 处理器型号：

- i. Intel® Pentium® IV Processor 或同等级及以上桌面级 CPU;
 - ii. Qualcomm® Snapdragon® 200 或同等级及以上移动级 CPU。
- b. 浏览器内核:
 - i. Chrome 21+
 - ii. Opera 12.1+
 - iii. Firefox 22+
 - iv. Safari 6.1+
 - v. IE 10+
 - c. 输入输出设备:
 - i. 键盘及鼠标或触控设备
 - ii. 显示器及打印机等

(2) APP 运行环境

Android 4.0.1 及以上

第三部分

需求规格说明书

软件需求说明书 ERROR! BOOKMARK NOT DEFINED.

<u>1 引言</u>	44
<u>1.1 编写目的</u>	44
<u>1.2 背景</u>	44
<u>1.3 参考资料</u>	44
<u>2 任务概述</u>	45
<u>2.1 目标</u>	45
<u>2.2 用户的特点</u>	46
<u>2.3 假定和约束</u>	46
<u>3 需求规定</u>	46
<u>3.1 对功能的规定</u>	46
<u>3.2 对性能的规定</u>	46
<u>3.2.1 精度</u>	46
<u>3.2.2 时间特性要求</u>	46
<u>3.2.3 灵活性</u>	47
<u>3.3 输入输出要求</u>	47
<u>3.4 数据管理能力要求</u>	47
<u>3.5 故障处理要求</u>	48
<u>3.6 其他专门要求</u>	48
<u>4 运行环境规定</u>	48
<u>4.1 设备</u>	48
<u>4.2 支持软件</u>	49
<u>4.3 接口</u>	49

1 引言

1.1 编写目的

软件需求规格说明描述了“股票分析预测系统（Stocalize）”1.0 版本的软件功能性需求和非功能性需求。这一文档计划由实现系统功能和验证系统功能正确的项目团队成员共同使用。除非有另外说明，这里指定的所有需求都需要在该版本中得到实现。

该软件需求规格说明具体针对开发人员、测试人员以及用户。

1.2 背景

该软件系统的名称为：Stocalize，版本号为：1.0

本项目的任务提出者为小组的指导老师，开发者为 15 级卓工实践小组，用户为我们小组和指导老师，项目暂时没有计算中心或计算机网络。

随着金融市场的飞速发展，炒股的股民也越来越多。而炒股软件也就顺应这个时期而腾空而出。现在市场上布满了琳琅满目的股票软件，广告也铺天盖地，都说得神乎其神。既然是这个时期的产物，它就有用的价值。一份成功的股票软件方案是赢得订单的重要砝码，如何写一份令客户心动的股票软件开发方案则需要下苦功夫。

股票软件在市场上出现的价值:内容重点阐述了股票软件能给股民带来什么价值，以及通过何用方法去实现这种价值，从而帮助股民赚钱；另外，还有股民在使用软件的一些想法和顾虑的解决。软件开发的过程，要求和说明:这一项是一个整个软件的开发过程，中间需要各种材料，要求和细节的说明。是一个比较复杂和完整的过程。

从各方面具体的详解软件开发后期的说明:主要是对软件完工之后，双方各执行的要求，和补充说明。使软件在完工的基础山更加完善。股票软件在市场上出现的价值,股票软件带给股民的好处和其价值。

1.3 参考资料

1. Qasem A. Al-Radaideh Adel Abu Assaf and Eman Alnagi, Predicting Stock Prices using Data Mining Techniques, The International Arab Conference on Information Technology (ACIT'2013)
2. Gunter Senyurt, Abdulhamit Subasi, Stock market movement direction prediction using tree algorithms.
3. Sadegh Bafandeh Imandoust, Mohammed Bolandraftar, Forecasting the direction of stock market index movement using three data mining techniques: the case of Tehran Stock Exchange, Int. Journal of Engineering Research and Applications, ISSN : 2248-9622, Vol. 4, Issue 6(Version 2), June 2014,

- pp.106-117.
4. Rosillo, R.; Giner, J. & and De la Fuente, D. (2014): "Stock Market simulation using support vector machines", Journal of Forecasting, vol. 33, n° 6, (488–500)
 5. GB T-8567-2006 计算机软件文档编制规范
 - a.

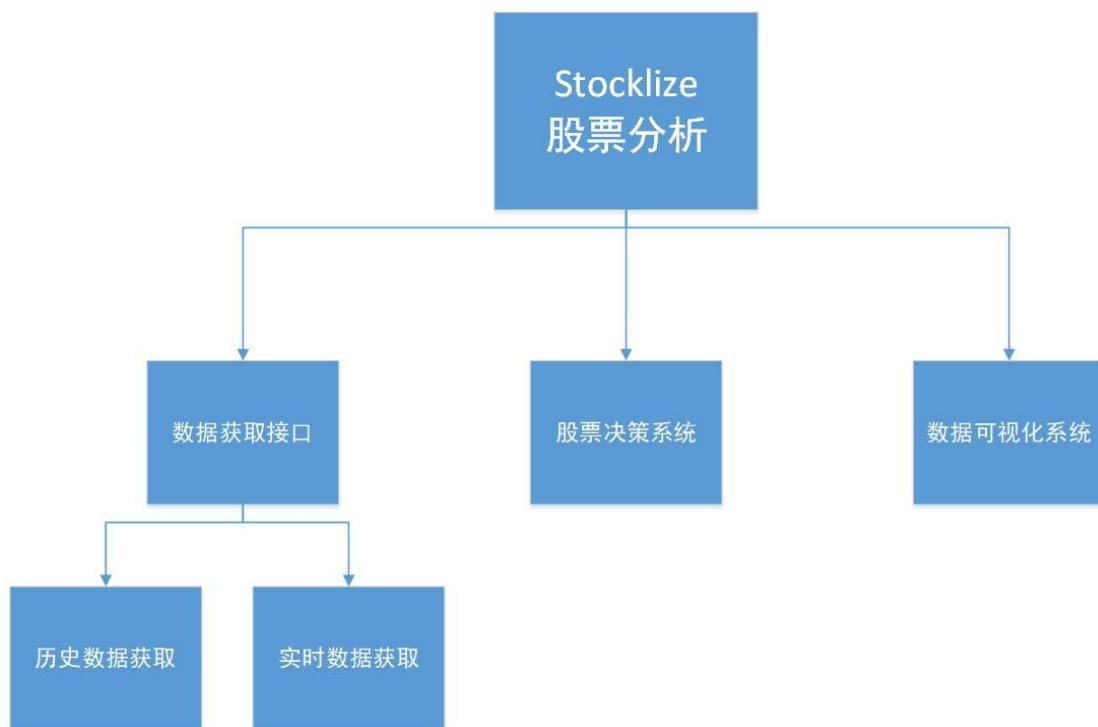
2 任务概述

2.1 目标

叙述该项软件开发的意图、应用目标、作用范围以及其他应向读者说明的有关该软件开发的背景材料。解释被开发软件与其他有关软件之间的关系。如果本软件产品是一项独立的软件，而且全部内容自含，则说明这一点。如果所定义的产品是一个更大的系统的一个组成部分，则应说明本产品与该系统中其他各组成部分之间的关系，为此可使用一张方框图来说明该系统的组成和本产品同其他各部分的联系和接口。|

本股票分析软件旨在为股民提供方便快捷的股票分析情况，在有一定股票知识的前提下，可自由组合相关股票与指标，最终得出股票走势分析与历史走势展示。

本软件前端接入网络开放的实时股票 API，后端接入现有的数据可视化库，中间分析部分自行实现，框图如下。



2.2 用户的特点

本软件使用对象为对股票知识有一定了解的人群，需要能看懂股票走势图与指标分析图。维护人员同样需要了解一些股票知识，以方便了解软件运行错误情况。

2.3 假定和约束

本软件开发无经费，需要在学期末结束前完成，本软件遵从开源软件协议。

3 需求规定

3.1 对功能的规定

功能模块	输入	处理	输出
查阅历史走势	指定股票及查询时间范围	从数据库获取数据	绘制一定时间段内的股票交易数据图表
预测股票走势	指定股票	从数据库获取数据，根据深度学习决策模型预测股票走势	绘制一定时间段内的股票交易数据图表及预测的走势
动态添加删除股票	指定股票与操作	修改用户关注股票列表	新的关注股票列表
修改个人信息	需要修改的内容	在数据库中修改内容	成功或失败的标志

3.2 对性能的规定

3.2.1 精度

股票增幅保留至小数点后三位。

3.2.2 时间特性要求

软件因为是股票软件，所以前端响应速度要求高，股票走势要及时更新最新交易数据，而预测

及推荐系统要求适中

- a. 录入结束响应时间：高
- b. 更新处理时间：中
- c. 数据的转换和传送时间：低
- d. 查询时间：高
- e. 修改时间：低

3.2.3 灵活性

说明对该软件的灵活性的要求，即当需求发生某些变化时，该软件对这些变化的适应能力，如：

- a. 对输入数据的检查，防止用户输入错误；
- b. 图形界面操作，鼠标点击增删股票列表，点击股票查看详情，滚轮缩放时间轴等，方便用户操作；
- c. 记录用户超过有效次数阈值的操作，以便之后升级软件优化相关操作体验；
- d. 依据用户使用记录自动排序推荐显示内容，推荐同类型股票；
- e. 可在设置中由用户选取详细信息列表中显示的数据类型。

3.3 输入输出要求

功能模块	输入	数据类型	输出	数据类型
查阅单只股票	1、指定股票 2、查询时间范围	String	绘制一定时间段内的股票交易数据图表	Image
预测股票走势	指定股票	String	预测的走势图	Image
动态添加删除股票	1、指定股票 2、操作类型	String Int	新的关注股票列表	String
修改个人信息	需要的内容	String	成功或失败的标志	Bool

解释各输入输出数据类型，并逐项说明其媒体、格式、数值范围、精度等。对软件的数据输出及必须标明的控制输出量进行解释并举例，包括对硬拷贝报告（正常结果输出、状态输出及异常输出）以及图形或显示报告的描述。

3.4 数据管理能力要求

需要存储在数据库内的数据有

- 历史数据：经过处理的股票历史数据，数据为 TB 级别
- 实时数据：从网络上获得的免费的股票实时数据，一般为 GB 级别，使用过的实时数据存入历史数据数据库内

3.5 故障处理要求

- 硬件故障：运维团队确定硬件具体故障后，由使用方进行更换
- 股票数据异常：若无法从网络上获取到实时的数据，则检查相关接口是否有问题，若有问题则更换其他可用的接口
- 网络问题：由使用方保障网络通畅
- 其他问题：联系运维团队，由其进行故障检查

3.6 其他专门要求

如用户单位对安全保密的要求，对使用方便的要求，对可维护性、可补充性、易读性、可靠性、运行环境可转换性的特殊要求等。

- 保密性要求：
 1. 若使用数据有保密性要求，则由其进行数据保密
 2. 训练模型的保密性同样由使用方保密
- 可维护性：便于后期升级开发
- 使用方便：对有股票知识的人要便于操作
- 跨平台要求：使用 python 语言，多平台支持

4 运行环境规定

4.1 设备

- a. 处理器型号及内存容量；

最低处理器要求：Intel Pentium 4/AMD Athlon 1.7GHz 或相等级

最低内存要求：4GB

- b. 外存容量、联机或脱机、媒体及其存储格式，设备的型号及数量；

软件磁盘空间占用大小约为 300MB

c. 输入及输出设备的型号和数量，联机或脱机；

输入设备：键盘、鼠标、触摸屏

输出设备：显示器

d. 数据通信设备的型号和数量；

以太网适配器或无线网络适配器

e. 功能键及其他专用硬件

无

4.2 支持软件

服务端操作系统：Ubuntu 14.04 及以上版本

编译/解释程序：python 2.7

4.3 接口

使用 HTTP 协议从金融网站上获取数据

第四部分

软件设计说明书

1 引言	3
1.1 编写目的	3
1.2 背景.....	3
1.3 参考资料	3
2 总体设计	4
2.1 需求规定	4
2.2 运行环境	4
2.3 基本设计概念和处理流程	5
2.4 结构.....	6
2.5 功能需求与程序的关系	6
2.6 人工处理过程	7
2.7 尚未解决的问题	7
3 接口设计	7
3.1 用户接口	7
3.2 外部接口	7
3.3 内部接口	8
4 运行设计	8
4.1 运行模块组合	8
4.2 运行控制	8
4.3 运行时间	8
5 系统出错处理设计.....	9
5.1 出错信息	9
5.2 补救措施	9
5.3 系统维护设计	9

1 引言

1.1 编写目的

软件需求规格说明描述了“股票分析预测系统（Stocalize）”1.0 版本的软件功能性需求和非功能性需求。这一文档计划由实现系统功能和验证系统功能正确的项目团队成员共同使用。除非有另外说明，这里指定的所有需求都需要在该版本中得到实现。

该软件需求规格说明具体针对开发人员、测试人员以及用户。

1.2 背景

该软件系统的名称为：Stocalize，版本号为：1.0

本项目的任务提出者为小组的指导老师，开发者为 15 级卓工实践小组，用户为我们小组和指导老师，项目暂时没有计算中心或计算机网络。

随着金融市场的飞速发展，炒股的股民也越来越多。而炒股软件也就顺应这个时期而腾空而出。现在市场上布满了琳琅满目的股票软件，广告也铺天盖地，都说得神乎其神。既然这是这个时期的产物，它就有用的价值。一份成功的股票软件方案是赢得订单的重要砝码，如何写一份令客户心动的股票软件开发方案则需要下苦功夫。

股票软件在市场上出现的价值:内容重点阐述了股票软件能给股民带来什么价值，以及通过何用方法去实现这种价值，从而帮助股民赚钱；另外，还有股民在使用软件的一些想法和顾虑的解决。软件开发的过程，要求和说明:这一项是一个整个软件的开发过程，中间需要各种材料，要求和细节的说明。是一个比较复杂和完整的过程。

从各方面具体的详解软件开发后期的说明:主要是对软件完工之后，双方各执行的要求，和补充说明。使软件在完工的基础山更加完善。股票软件在市场上出现的价值,股票软件带给股民的好处和其价值。

1.3 参考资料

1. Qasem A. Al-Radaideh Adel Abu Assaf and Eman Alnagi, Predicting Stock Prices using Data

Mining Techniques, The International Arab Conference on Information Technology
(ACIT'2013)

2. Gunter Senyurt, Abdulhamit Subasi, Stock market movement direction prediction using tree algorithms.
3. Sadegh Bafandeh Imandoust, Mohammed Bolandraftar, Forecasting the direction of stock market index movement using three data mining techniques: the case of Tehran Stock Exchange, Int. Journal of Engineering Research and Applications, ISSN : 2248-9622, Vol. 4, Issue 6(Version 2), June 2014, pp.106-117.
4. Rosillo, R.; Giner, J. & and De la Fuente, D. (2014): "Stock Market simulation using support vector machines", Journal of Forecasting, vol. 33, n°6, (488–500)
5. GB T-8567-2006 计算机软件文档编制规范

2 总体设计

2.1 需求规定

输入项目：需要分析的股票代码。

输出项目：需要分析的股票未来趋势图和分析图。

性能要求：1s 内得到需要的结果。

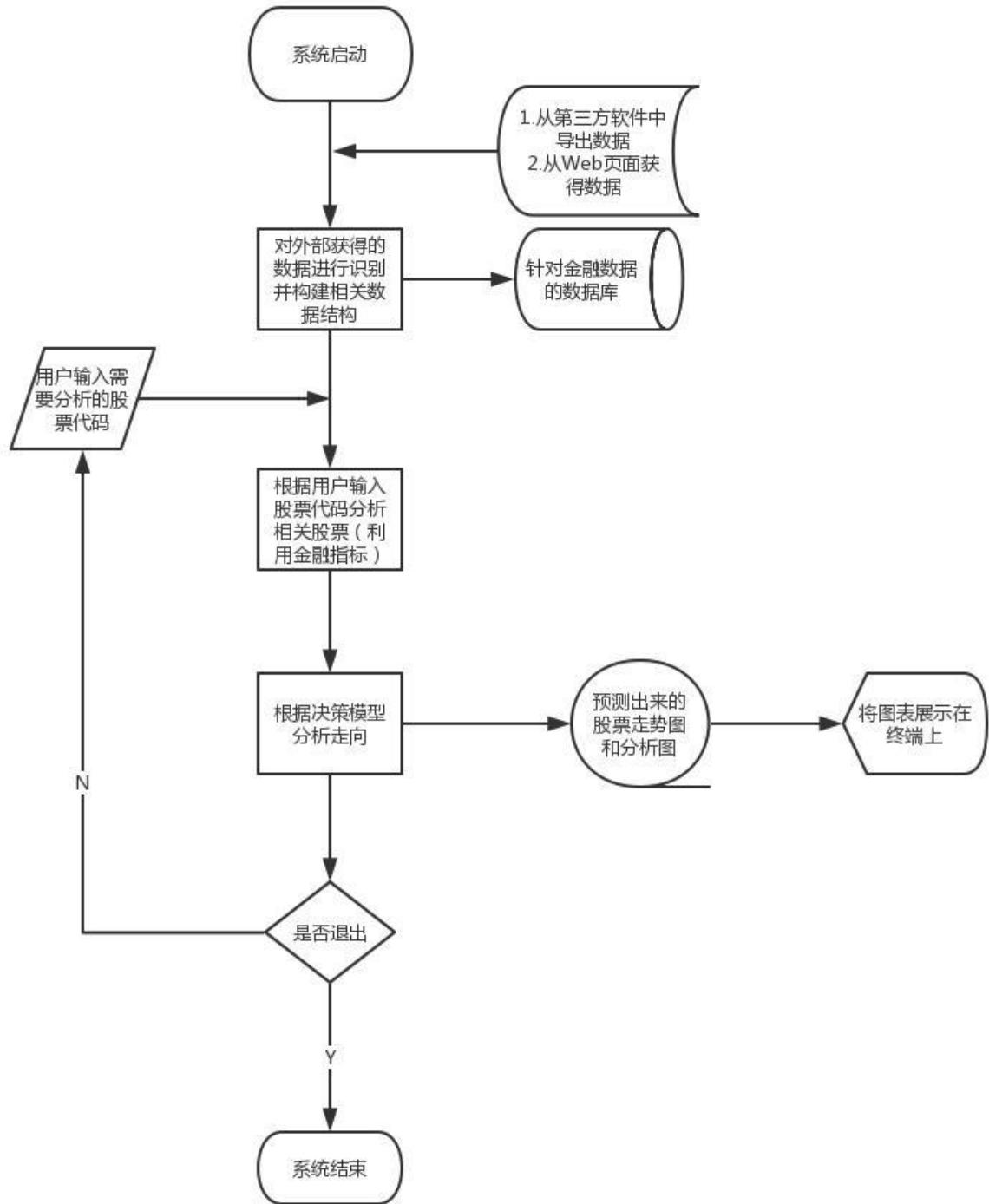
2.2 运行环境

服务器软件环境：Ubuntu 16.04 + Anaconda 2 + Pytorch + Scrapy + Apache

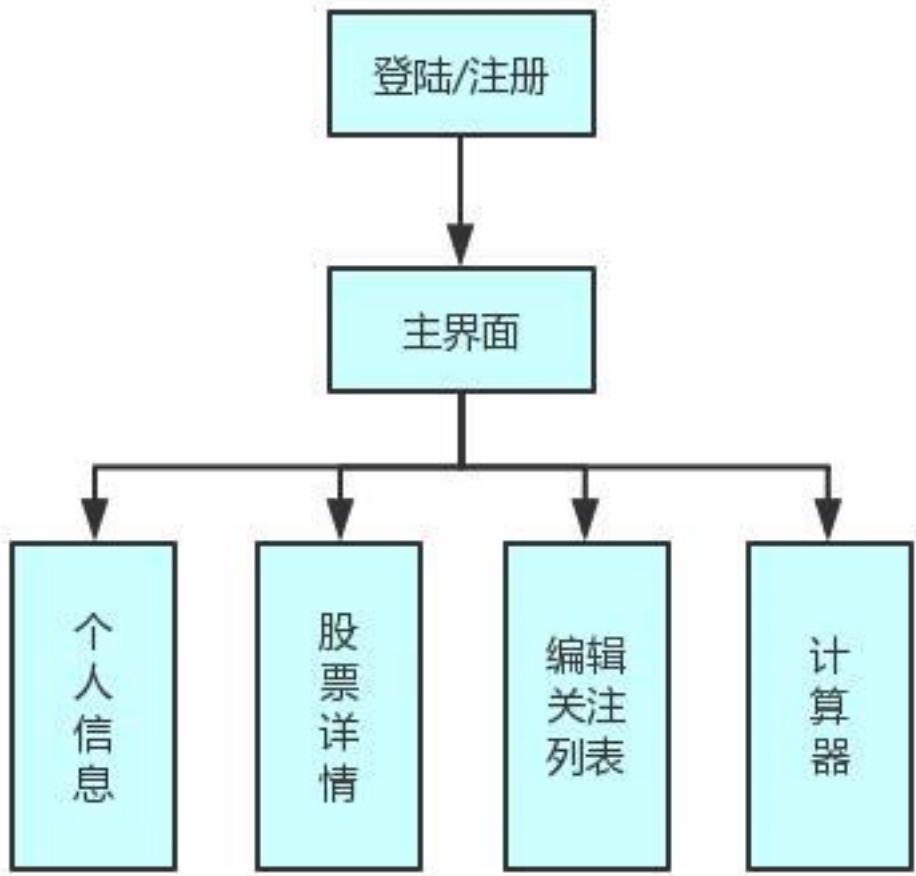
客户端软件环境：Windows 7 及以上/OS X/Linux/Android

服务器硬件环境：Intel Core i7 + NVIDIA Geforce GTX 970M + 8G RAM

2.3 基本设计概念和处理流程



2.4 结构



2.5 功能需求与程序的关系

	界面程序	股票详情	个人信息	计算器	后台数据库
显示股票列表	√				√
股票详细数据查看		√			√
股票走势预测		√			√
添加删除股票			√		√
修改个人信息			√		√

股票数据计算器				√	
---------	--	--	--	---	--

2.6 人工处理过程

导入初始历史数据，手动选择需要分析的股票。

2.7 尚未解决的问题

程序无法运行在非 x8 平台上

3 接口设计

3.1 用户接口

根据需求分析的结果，用户需要一个用户友善界面。在界面设计上，应做到简单明了，易于操作，注意到界面的布局，应突出的显示重要以及出错信息。外观上为做到合理化，考虑到用户多对 Windows、Android 风格较为熟悉，应尽量向其靠拢，故采用 Material Design 的设计风格。

本软件向用户提供如下具体操作接口：

1. 股票查询接口。本软件提供股票走势预测查询，用户输入要查询的股票名称或代码并确认提交后，本软件将进行后台查询并返回该股票的相关基本信息及预测信息以显示。
2. 用户登录接口。本软件提供用户账户功能，用户注册并登录软件后将获取后台存储的用户自选股等相关信息完成多设备同步及恢复。同时，也提供用户密码丢失找回、密码修改等功能。

3.2 外部接口

1. MATLAB 接口：通过调用 MATLAB 完成部分数据分析。主程序向 MATLAB 提供股票金融数据，MATLAB 返回模型下计算出的回归值。
2. 雅虎 YQL 接口：通过使用雅虎 YQL 提供的股票数据服务接口获取过去至现在所有股票的金融数据。主程序向雅虎 YQL 接口发送股票代码及数据日期等请求，雅虎 YQL 接口返回对应的股票金融数据。
3. 数据库接口：主程序可使用数据库提供的接口，进行对数据库的所有访问
4. 网络接口：使用一种无差错的传输协议，采用滑动窗口方式对数据进行网络传输及接受
5. 输入输出硬件接口：对于键盘、鼠标、触摸屏的输入，可使用标准输入/输出接口进行处理；对于显示器的输出，也可以用标准输入/输出接口进行处理。

3.3 内部接口

内部各模块之间采用函数调用、参数传递、返回值的方式进行信息传递。 接口传递的信息将以数据结构封装了的数据，以参数传递或返回值的形式在各个模块间传递。主要有如下数据传输：

1. 数据获取模块向数据分析模块提供获取的金融数据
2. 数据分析模块向显示模块提供要显示的计算结果数据
3. 数据获取模块向显示模块提供股票代码等不变的信息数据
4. 用户登录模块和数据库交互用户名密码及匹配确认具体参数结构将在后文说明。

4 运行设计

4.1 运行模块组合

	历史数据获取模块	实时数据获取模块	股票决策模块	数据可视化模块
查阅历史走势	√			√
预测股票走势	√	√	√	
动态添加删除股票	√	√		
修改个人信息				

4.2 运行控制

功能	控制方法
查阅历史走势	在股票列表中选择股票，以及查询的时间范围。系统从数据库中抽取出对应的数据并绘制成图表。
预测股票走势	在股票列表里选择股票，系统根据决策模型进行分析和预测给出股票一段时间内的走势预测。
动态添加删除股票	用户对关注股票列表进行增加和删除，系统对增加的股票信息进行爬取，在数据库中对删除的股票进行删除。
修改个人信息	用户输入需要修改的个人的信息，系统更新数据库的相关信息。

4.3 运行时间

不可用

5 系统出错处理设计

5.1 出错信息

出错类型	错误提示	处理方法
股票实时数据关闭	无法获取数据	由维护人员更换接口
无法分析数据	数据格式错误	选择适当格式的数据
无法连接网络	网络连接失败	咨询网络服务商
软件不兼容	运行错误 210	更换兼容平台
文件丢失	缺少运行模块	重新安装软件
缺少运行环境	运行错误 501	安装所需环境

5.2 补救措施

- a. 自动验证文件完整性，并把程序运行核心文件自动备份至指定文件夹，若发现文件缺失则去默认路径寻找缺失文件。
- b. 若出现资源不足的情况，则停止复杂模块的运行，简化软件流程，牺牲性能换取程序的正常运行，否则提示资源不足。
- c. 本软件若崩溃，则排查故障后重新启动即可，软件会自动获得最新信息。

5.3 系统维护设计

本软件可分为调试模式和正常模式，使用宏定义，调试模式则使用 `debug` 开关打开调试按钮，然后进行调试，会输出一些对用户无用的调试信息，但对程序员有指导意义。

正常模式会隐藏调试信息，并停止一些测试模块，让其只对用户使用。若出现了问题，则打开调试开关即可排查具体问题。