**《数据库实训》 课程报告**

**(2020-2021 学年第2 学期)**

**题目 投资组合评比器**

**提交日期： 2021 年 6 月 日**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **学 院** | **软件学院** | | **专业班级** | **软件工程19级一班** | |
| **课程名称** | **数据库实训** | | **任课教师** | **曾兵** | |
| **人数** | **3** | |  |  | |
| **学 号** | | **学生姓名** | | | **分 数** |
| **201936380086** | | **陈涵** | | | +1 |
| **201930381317** | | **梁永豪** | | | +1 |
| **201930381355** | | **廖宇延** | | | -2 |
|  | |  | | |  |
|  | |  | | |  |
|  | |  | | |  |
|  | |  | | |  |
| **教师评语：**  **教师签名：** | | | | | |

目录

[正文 2](#_Toc74677815)

[一、需求分析 2](#_Toc74677816)

[二、概要设计 3](#_Toc74677817)

[1. 系统逻辑架构 3](#_Toc74677818)

[2. 系统物理架构 3](#_Toc74677819)

[3. 系统数据模型 3](#_Toc74677820)

[4. 系统开发架构 4](#_Toc74677821)

[三、详细设计 4](#_Toc74677822)

[1. 概述 4](#_Toc74677823)

[2. 术语表 4](#_Toc74677824)

[3. 设计概述 4](#_Toc74677825)

[四、软件测试 5](#_Toc74677826)

[五、软件演示 5](#_Toc74677827)

[六、软件使用教程 5](#_Toc74677828)

[小组讨论纪要 5](#_Toc74677829)

[个人总结 5](#_Toc74677830)

[项目进度安排等项目管理内容 5](#_Toc74677831)

# 正文

## 一、需求分析

1. 实现从蛋卷基金和且慢基金网站上爬取相关基金数据。
2. 将爬取的相关基金数据存储至本地数据库中。
3. 实现图形化界面，可对以下内容进行展示。

a.可以把投资组合放在同一时间段下进行比较，时间段可由用户指定，例如, 20180101-20210401

b.可以对比不同的指标，如：涨幅/年化收益率, 最大回撤, 夏普比, 年化波动率。

c.可以根据上述指标进行排序。

4）可扩展数据源，可以添加且慢和蛋卷平台新的数据源。

## 二、概要设计

### 1. 系统逻辑架构



### 2. 系统物理架构

|  |  |
| --- | --- |
| 建议CPU | Intel® Core™ i7 Gen8 or higher  AMD Ryzen R5 Gen2 or higher |
| 建议内存 | 4GB DDR4 Memory or more |
| 系统 | Windows 10 20H2 or later/MacOS Big Sur 11.4 or later |

### 3. 系统数据模型

E-R图

图标

描述已自动生成

### 4. 系统开发架构

图形化界面-python3.8 with tkinter, matplotlib

数据库-mysql8.0

爬虫-python3.8 with requests, json

## 三、详细设计

### 1. 概述

本系统可用于爬取蛋卷基金、且慢基金网站上的投资组合信息，存储至本地数据库中。可以不断更新新获取的数据。同时能用图形化界面向用户展示基金的各方面信息，便于用户将不同的投资组合放在一起进行比较。

本系统基于python3.8，配合mysql8.0用作数据存储，实现本地爬取，存储，展示数据。

### 2. 术语表

|  |  |
| --- | --- |
| 历史净值 | 该基金某日的净值。 |
| 最大回撤 | 在选定周期内任一历史时点往后推，产品净值走到最低点时的收益率回撤幅度的最大值。 |
| 年化收益率 | 某一时间段内收益率，除以时间长度乘以360天，用来衡量区间内收益。 |
| 无风险化利率 | 指没有风险的理财方式所获得的资金收益比例。专业点说是对无信用风险和市场风险的资产的投资，指到期日期等于投资期的国债的利率。 |
| 年化波动率 | 一年里基金收益的波动程度。年化波动率=一年里日收益的标准差\* |
| 夏普率 | 它代表基民每多承担一分风险，可获得高出无风险回报率几分的回报。夏普率=(年化收益率-无风险利率)/年化波动率。 |

### 3. 设计概述

#### 系统接口设计

##### 用户与系统间交互

用户通过图形化界面与系统进行交互

##### 系统内的交互

1. 主程序直接调用爬虫从网站上爬取数据
2. 主程序通过pymysql库中定义的连接函数与数据库连接，将爬虫爬取的数据存储至数据库
3. 图形界面从数据库中获取数据后，将数据存储在缓存中，根据用户指令对数据进行操作后再进行展示

#### 对象模型设计

屏幕上有字

描述已自动生成

#### 系统用例实现详细设计

#### d)系统非功能设计

1、为了加快爬取速度，增加了多线程爬取。

2、为了加快数据的获取，在数据库中建立了索引。

3、针对数据库操作增加了异常处理。

## 四、软件测试

## 五、软件演示

## 六、软件使用教程

# 小组讨论纪要

|  |  |  |  |
| --- | --- | --- | --- |
| 第一次小组讨论 | | 时间 2021年5月20日 | |
| 参会人员 | 陈涵 | 梁永豪 | 廖宇延 |
| 1. 在Gitee上初始化仓库，填写项目基本信息。 2. 讨论爬虫的实现，且慢基金因为header的反爬校验暂时不能爬取。 3. 讨论前端实现运用的语言。 4. 确认后期工作任务进度。 5. 确认要实现功能的相应模块。 | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 第二次小组讨论 | | 时间 2021年5月30日 | |
| 参会人员 | 陈涵 | 梁永豪 | 廖宇延 |
| 1. 讨论实现前端界面所用到的库。 2. 对前端样式进行初步设定。 3. 讨论前端所要展示的内容以及呈现样式。 4. 对爬取过程进行优化，能够只爬取数据库中没有的新数据。 | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 第三次小组讨论 | | 时间 2021年6月15日 | |
| 参会人员 | 陈涵 | 梁永豪 | 廖宇延 |
| 1. 检查所有已实现的工作。 2. 确定所有必须功能已全部实现。 3. 对软件进行测试。 4. 整理报告以及相应文档。 | | | |

# 个人总结

# 项目进度安排等项目管理内容

|  |  |
| --- | --- |
| 12周 | 完成爬取 |
| 13周 | 设计数据库和对象，完成数据的存储 |
| 14周 | 完成数据的计算和比较 |
| 15周 | 设计图形化展示界面 |
| 16周 | 优化设计 |