**1.系统设计文档**

**1.1系统概述**

本系统是一个基于Python的数据爬取工具，用于从东方财富网（EastMoney）API中爬取基金的历史净值数据，并将其保存为CSV文件。系统通过模拟HTTP请求获取数据，并使用正则表达式和JSON解析提取所需信息，最后通过Pandas库进行数据格式化和存储。

**1.2系统架构**

系统主要由以下模块组成：

1.数据请求模块：负责发送HTTP请求并获取响应数据。

2.数据解析模块：通过正则表达式和JSON解析提取所需数据。

3.数据存储模块：使用Pandas将数据格式化为DataFrame，并保存到CSV文件中。

4.控制模块：管理循环请求和数据合并逻辑。

**1.3技术选型**

Python：编程语言。

requests：用于发送HTTP请求。

re：用于正则表达式解析。

json：用于JSON数据解析。

pandas：用于数据处理和存储。

**1.4数据流程**

1.请求数据：通过requests.get发送HTTP请求。

2.解析数据：通过正则表达式提取JSON数据，并解析为Python对象。

3.格式化数据：将解析后的数据转换为PandasDataFrame。

4.存储数据：将所有页面的数据合并后保存为CSV文件。

**2.说明文档**

**2.1功能描述**

本工具的主要功能是从东方财富网API中爬取指定基金的历史净值数据，并将其保存为CSV文件。工具支持分页爬取，能够处理多页数据并合并。

**2.2输入与输出**

输入：

基金代码（如519185）。

分页参数（如pageIndex和pageSize）。

请求头信息（如Referer和UserAgent）。

输出：

CSV文件，包含爬取的历史净值数据。

**2.3配置参数**

fundCode：目标基金代码。

pageIndex：当前请求的页码。

pageSize：每页返回的数据条数。

startDate和endDate：可选的时间范围参数。

**2.4依赖库**

requests：用于发送HTTP请求。

re：用于正则表达式解析。

json：用于JSON数据解析。

pandas：用于数据处理和存储。

**3.使用文档**

**3.1环境准备**

安装依赖库：pipinstallrequestspandas

**3.2配置代码**

根据需要修改代码中的以下部分：

1.基金代码：将fundCode替换为目标基金代码。

2.分页参数：根据API的分页规则调整pageIndex和pageSize。

3.请求头：根据目标网站的要求调整Referer和UserAgent。

**3.3运行代码**

程序运行完成后，会在当前目录生成一个名为股票数据.csv的文件，包含爬取的历史净值数据。

**3.4注意事项**

1.确保API接口可用，且请求频率符合网站要求。

2.如果API返回数据为空，可能需要检查分页参数或API限制。

3.请遵守相关法律法规和网站的使用条款。

**4.测试文档**

**4.1测试目标**

验证系统是否能够正确爬取基金历史净值数据，并将其保存为CSV文件。

**4.2测试环境**

Python版本：3.8及以上。

依赖库：requests、pandas。

**4.3 测试用例**

4.3.1 测试用例1：正常数据爬取

输入：

基金代码：519185。

分页参数：pageIndex=1，pageSize=20。

预期输出：

CSV文件中包含至少一页的历史净值数据。

数据格式正确，无缺失字段。

4.3.2 测试用例2：多页数据爬取

输入：

基金代码：519185。

分页参数：循环请求多页（如pageIndex=1到51）。

预期输出：

CSV文件中包含多页数据。

数据合并正确，无重复或丢失。

4.3.3 测试用例3：API返回空数据

输入：

基金代码：519185。

分页参数：超出API返回范围的页码。

预期输出：

程序能够正确处理空数据，不报错。

CSV文件中保存已有数据，不包含空行。

4.3.4 测试用例4：请求头错误

输入：

基金代码：519185。

请求头：错误的Referer或User-Agent。

预期输出：

程序能够捕获错误，提示用户检查请求头。

不生成CSV文件。

**4.4 测试结果**

测试用例1：通过，CSV文件中包含一页数据，格式正确。

测试用例2：通过，CSV文件中包含多页数据，合并正确。

测试用例3：通过，程序正确处理空数据，CSV文件正常。

测试用例4：通过，程序提示错误，未生成CSV文件。

**4.5测试结论**

系统能够正确爬取基金历史净值数据，并处理多页数据和异常情况。测试结果符合预期。