

图书借阅推荐系统设计与实现

张博然, 乌 伟

西京学院计算机学院, 陕西 西安

收稿日期: 2025年3月25日; 录用日期: 2025年4月23日; 发布日期: 2025年4月30日

摘 要

随着互联网技术的不断发展, 图书借阅服务也逐渐向线上转移。用户可以通过网络借阅图书, 这样就不再局限于实体书库的开放时间和地点, 同时也更加方便快捷。但是, 随着书籍资源日益丰富、品种日益多样以及借阅总量的扩大, 怎样给用户提供更人性化的推送服务, 怎样针对不同用户的自身偏好推送书籍, 实现用户更方便、更快捷、精准地获得自身所需求的、优质的书籍资源, 成为等待解决的问题。因此, 发展图书推荐服务、满足用户的个性化阅读需求成为了图书借阅服务面临的一个问题。针对上述情况, 该平台基于产品的协同过滤算法, 并利用协同筛选的方法可实现基于用户的和基于产品的图书阅读体系。其前后端设计已完全分离, 前台使用了Vue + Element以及UI进行设计, 而后台则使用了Spring及Boot模式。利用scrapy爬虫框架在豆瓣图书官网爬取网络用户的图书借阅信息, 最终系统设计实现用户登录注册、图书借阅推荐、个性化推荐和用户列表及图书列表展示功能, 解决了用户无法在众多书目中快速找到目标和高质量图书的问题, 提高了查找图书效率, 为用户提供了更多选择。

关键词

协同过滤算法, 推荐系统, MySQL, Spring Boot, Vue + Element UI

The Design and Implementation of a Book Borrowing Recommendation System

Boran Zhang, Wei Wu

School of Computer Science, Xijing University, Xi'an Shaanxi

Received: Mar. 25th, 2025; accepted: Apr. 23rd, 2025; published: Apr. 30th, 2025

Abstract

With the continuous development of Internet technology, the book lending service has gradually transferred to the online. Users can borrow books through the network, so it is no longer limited to the physical library of the opening time and place, but also more convenient and fast. However, with more and more books resources, more and more kinds of books and the increase in the number of

loans, how to provide personalized recommendation service for users, how to recommend books according to different users' own preferences, to achieve users more convenient, faster, accurate access to their own needs, high-quality book resources have become a problem waiting to be solved. Therefore, the development of book recommendation service to meet the user's personalized reading needs has become a book lending service facing a problem. In order to solve these problems, this paper designs a user-based collaborative filtering algorithm, and uses the collaborative filtering recommendation algorithm to realize the user-based and item-based book recommendation system. The front-end and back-end of the system are designed separately. The front-end uses technologies such as Vue + Element UI, and the back-end uses the Spring Boot framework. Scrapy crawler framework is used to crawl the web users' book lending information on Douban Library. Finally, the system design realizes the functions of user login, book borrowing recommendation, personalized recommendation, user list and book list display. It solves the problem that the user can not find the target and the high-quality books quickly in the numerous bibliographies, improves the efficiency of searching books, and provides more choices for the user.

Keywords

Collaborative Filtering Algorithm, Recommendation System, MySQL, Spring Boot, Vue + Element UI

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 绪论

1.1. 选题背景

随着互联网的发展,越来越多的人开始使用在线图书借阅服务。但是,随着信息量庞大在线书籍资料日益充实、品种日益丰富,给用户带来了更多的选择权。提供个性化推荐给用户是一个挑战。传统的分类检索和热门推荐,不能满足读者快速、精准地获取到自己所需要的、高质量的图书资源[1]。因此,发展图书推荐服务、满足用户的个性化阅读需求是在线图书馆的必然趋势和重要目标。

本次设计研究内容如何利用用户和图书多维度信息,提出了基于大数据的图书馆推荐借阅方法,实现用户的图书借阅精准推荐,以提高读者借阅图书时查找效率和图书资源利用率,为大数据推荐系统实际应用提供一定借鉴。

1.2. 国内外研究与发展现状

国外发展现状:

外国高校通常拥有比较大量的固定的技术开发团队来进行技术服务和支持,并且相对于技术软件而言,要比在国内起步早许多,使得国外的系统性能更为稳定,技术也更为完善。而外国的电子图书系统真正蓬勃发展的阶段则是在 20 世纪 90 年代以后[2]。在这一时期,来自微软的以图形界面显示的操作系统也开始问世,新科技的快速发展也促使了书籍管理系统的开发人员们意识到,以前开发的书籍系统早已不再适应了网络时代的发展趋势,于是开始把新科技运用于书籍系统的研发,比如人们在检索书籍信息时,可以利用最先进的自动搜索的成果,并结合数据库技术增强了搜索的高效性能,书籍管理系统也呈现出了多元化发展的态势[3]。

国内发展现状:

纵观目前的研究现状,因为技术起步较晚,在准确性和信息更新性上存在着相当的不足,而且手段

和技术存在相当的分享局限性。国内的数字化书籍管理工作才刚刚兴起，在 20 世纪 70 年代，中国国内就已经开始了引入国外的先进图书管理系统[4]。到了 20 世纪 90 年代末期，中国国内也开始了引入 C/S 管理系统框架下的图书管理工作，如今中国国内的图书管理工作已经开始进入了一个新阶段，也已经开始逐步缩小了与外国图书管理信息系统之间的差距，中国国内的图书管理工作也已经开始逐渐形成了一个拥有知识产权的全产业链体系。

2. 关键技术简介

2.1. 基于近邻的协同过滤算法

关于推荐算法的理论分支很多，按照理论所用的数据源和选择方法可粗略划分为图 1 给出的模型。

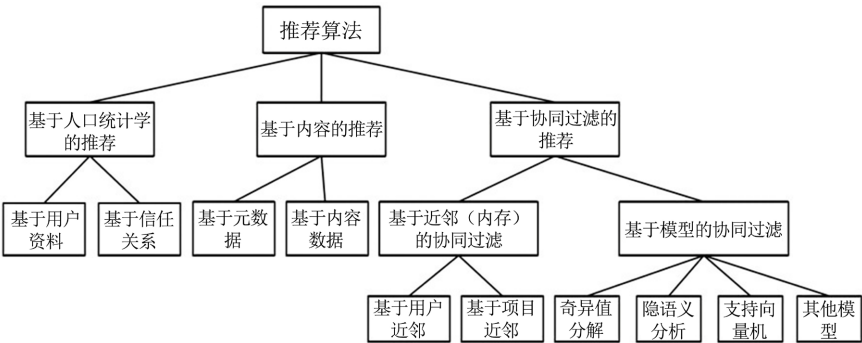


Figure 1. Recommendation algorithm classification
图 1. 推荐算法分类

(一) 算法原理

协同过滤算法(Collaborative Filtering, CF)是一种广泛应用于推荐系统的技术，其核心思想是通过分析用户的行为数据，挖掘用户之间的相似度或物品之间的相似度，从而进行个性化推荐。协同过滤算法主要分为两大类：基于用户的协同过滤(User-Based Collaborative Filtering, UserCF)和基于物品的协同过滤(Item-Based Collaborative Filtering, ItemCF) [5]。

1) 算法基本步骤

数据收集：收集用户的行为数据，如评分、点击、浏览、购买等。

数据预处理：对收集到的数据进行清洗、去重和归一化处理，使其符合后续数学模型的要求。

计算相似度：根据用户或物品的行为数据，计算用户之间的相似度或物品之间的相似度。

生成推荐列表：根据相似度计算结果，为用户推荐可能感兴趣的物品。

2) 公式推导

用户相似度计算

常用的相似度计算方法包括余弦相似度、皮尔逊相关系数等。

余弦相似度：通过计算两个向量在向量空间中的夹角余弦值来评估相似度。

皮尔逊相关系数：通过计算两个用户对同一组物品的评分之间的相关性来衡量相似度。

物品相似度计算

物品相似度的计算方法与用户相似度类似，也是基于用户对物品的评分数据。例如，可以使用余弦相似度或皮尔逊相关系数来计算物品之间的相似度。

预测评分

在计算出用户相似度或物品相似度后，可以使用这些相似度来预测用户对未评分物品的评分。

(二) 参数选择

相似度计算方法的选择：根据具体的应用场景和数据特点选择合适的相似度计算方法。例如，在评分数据较为稠密的情况下，皮尔逊相关系数可能更为合适；而在隐式反馈数据(如点击、浏览)较为稀疏的情况下，余弦相似度可能更为合适。

邻居数量的选择(即 TopN 或 TopK)：邻居数量对推荐结果的准确性和多样性有很大影响。邻居数量过多可能导致推荐结果过于泛化，而邻居数量过少则可能导致推荐结果过于个性化[6]。在实际应用中，需要根据系统的性能和用户的需求进行权衡。

数据预处理参数：如数据归一化的范围、去重和清洗的规则等，这些参数的选择也会影响推荐系统的性能。

(三) 相似度计算方法

除了上述的余弦相似度和皮尔逊相关系数外，还有其他一些常用的相似度计算方法，如杰卡德相似系数、欧几里德距离等。

杰卡德相似系数：主要用于计算符号度量或布尔值度量的个体间的相似度。

其中，A 和 B 是两个集合， $|A \cap B|$ 表示集合 A 和 B 的交集元素个数， $|A \cup B|$ 表示集合 A 和 B 的并集元素个数。

欧几里德距离：用于计算两个向量在欧几里德空间中的距离，距离越小表示相似度越高。

其中， A_i 和 B_i 分别表示向量 A 和 B 的第 i 个分量。为了将距离转化为相似度，可以对距离进行倒数或指数变换。

2.2. 爬虫框架 Scrapy

Scrapy 是一个分布式协作的框架，基于 Twisted，主要适用于使用较 Python 快、高层次的方法实现从屏幕抓取并捕获 web 网页信息，从网页中获取大量结构化数据，主要应用于信息挖掘、数据处理和存储历史数据等的程序中。由于它所采用的框架结构简单明了，数据结构清晰，模块的耦合度非常低，而扩展性又相当高，所以任何人员都可以根据爬取数据的需要而对架构进行改变，并且程序编写相对简单。它还提供了类似 BaseSpider、sitemap 之类的爬虫基类，功能强大，使用面广 Scrapy 框架主要包含了以下六个模块 [7]，分别为：Engine；Internet；Scheduler；Downloader；Spiders；Item Pipeline。Scrapy 框架如图 2 所示。

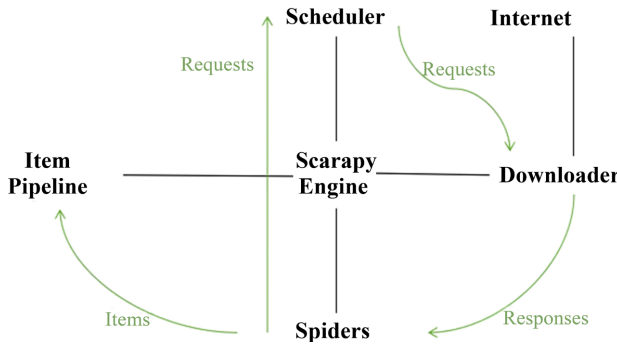


Figure 2. Scrapy framework
图 2. Scrapy 框架

2.3. 前端开发技术 Vue + Element UI

前端 Vue 是一个流行的 JavaScript 框架，它使用了现代化的 Web 开发技术，可以帮助开发者快速构建交互性强、高效、响应式的单页应用程序。Vue 提供了一些便捷的开发工具和可复用的组件，可以大

大减少前端开发的工作量和开发时间，同时也提高了代码的可维护性和可扩展性。Element UI 是 Vue 生态圈中的一个 UI 组件库，它是由饿了么的前同事们共同开发的，展示了大量的精美 UI 控件，包括表格、表单、弹窗、按钮等[8]。Element UI 具有灵活的布局、可定制的主题、良好的文档和社区支持等优点，能够帮助开发者快速构建具有良好用户体验的 Web 应用程序。

3. 图书借阅推荐系统的分析与设计

3.1. 系统总体功能设计

本图书借阅推荐系统要实现的功能主要包括为你推荐和猜你喜欢，用户管理，图书管理，借书管理和还书管理这五大板块。能根据不同用户的借阅浏览记录和评分记录为其实现个性化推荐，以使用户能快速找到并借阅自己感兴趣的图书。

3.1.1. 系统实现流程设计

系统实现流程图如图 3 所示：

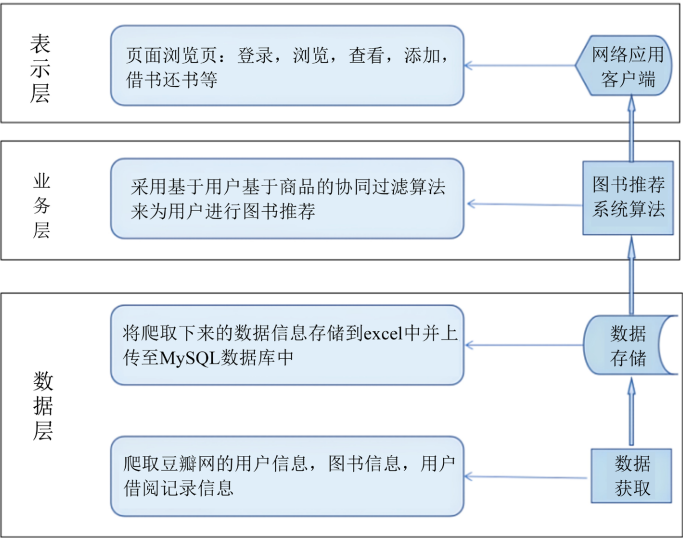


Figure 3. System implementation flowchart
图 3. 系统实现流程图

3.1.2. 系统总体功能框架

本系统要实现的页面的总体功能架构如图 4 所示：

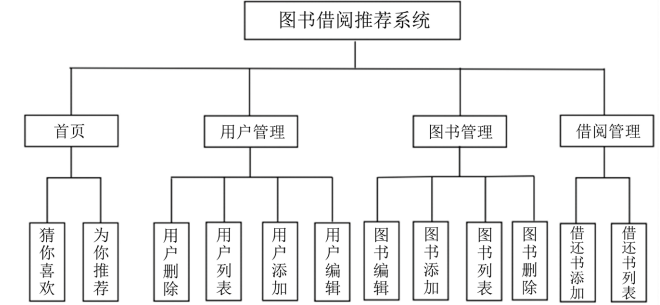


Figure 4. System overall functional architecture diagram
图 4. 系统总体功能架构图

3.2. 数据库设计

本节主要对图书借阅推荐系统进行数据库设计，经过前面的详细分析，明确了本系统需要存储的数据。本系统使用的关系型数据库是 MySQL，字符集使用 utf-8，共有 3 张表。主要的表是：用户表、图书表、借阅信息表。实体 - 关系(E-R)模型如图 5 所示：

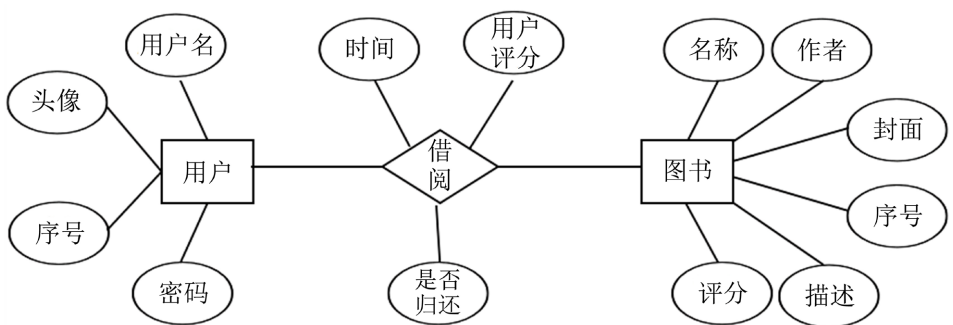


Figure 5. E-R diagram of the database
图 5. 数据库 E-R 图

用户表(user) (表 1)存储用户的基本信息，包括用户的序号、用户名、密码、头像。

Table 1. User table
表 1. 用户表

列名	类型	长度	主键	注释
name	varchar	128	是	用户名
img_url	varchar	1000		用户头像
password	varchar	128		登陆密码
id	bigint	auto increment	是	用户序号

图书表(book) (表 2)存储图书的基本信息，包括图书名、序号、作者类型、封皮、评分、状态等。

Table 2. Book table
表 2. 图书表

列名	类型	长度	主键	注释
name	varchar	128	是	图书名称
author	varchar	128		作者
img_url	varchar	1000		图书封面
star	varchar	10		评分
content	varchar	1000		描述
id	bigint	auto increment	是	图书序号

借阅信息表(borrow_info) (表 3)存储用户借阅记录，包括借阅图书名称，借阅用户名称，借阅时间，用户评分，图书状态。

Table 3. Borrowing information table
表 3. 借阅信息表

列名	类型	长度	主键	注释
user_id	bigint			用户序号
book_id	bigint			图书序号
time	varchar	100		借阅时间
score	varchar	10		用户评分
status	varchar	32		借阅或归还状态

4. 图书借阅推荐系统的开发与实现

4.1. 推荐系统环境部署

开发方法：利用虚拟化的开发工具 VMware Workstation 搭建 Linux 虚拟机，其版本为 VMware workstationpro16；操作系统为 CentOS-7-x86_64-Minimal-一九零八年；研发时采用的开发工具，以及版本为：IntelliJ IDEA 2021.3.1。

开发语言和框架：选择一种适合图书借阅推荐系统的开发语言和框架：Java、Python；Java 版本为 jdk1.8.0_181，python 版本为 Pycharm2021.2.2；并选择适合的框架：Spring，以便于快速构建系统。

数据库和数据存储技术：选择一种适合系统需求的数据库和数据存储技术，使用的数据库是开源的 MySQL 5.7.26，并选择适合的数据库管理工具 MyBatis 等，以便于管理和操作数据。

前端开发技术：选择一种适合系统需求的前端开发技术：Vue，并选择适合的 UI 组件库：Element UI，以便于实现系统的界面和交互[9]。

软件选择具体内容如表 4 所示：

Table 4. Software selection
表 4. 软件选择

软件名称	安装包名称	下载地址
VMware Workstation	VMware Workstation Pro 16	https://www.vmware.com/products/workstation-pro/workstation-pro-evaluation.html
CentOS	CentOS-7-x86_64-Minimal-1908	https://www.centos.org
Java	jdk1.8.0_181	http://www.oracle.com
Hadoop	hadoop-2.7.5	http://hadoop.apache.org
MySQL	MySQL 5.7.26	http://www.MySQL.com

4.2. 数据采集与存储

借阅数据的主要来源是爬取豆瓣网在线借阅平台的借阅信息，主要包括：用户信息，包括用户的用户名、登陆密码和用户序号；图书信息，包括图书名称、作者、图书评分、图书描述和图书序号；借阅信息，包括借书时间、借书用户、借书名称和图书是否归还的状态。使用 scrapy 爬虫框架进行数据获取，并将结果保存至 excel 中。

把 excel 中的数据文件集导入至 MySQL 中，但是因为在 MySQL 中只是存储了一些少量的业务数据，所以我们就将数据也同步至存储数据量最大的资料仓储 hive 中。数据库中用户的列表保存形式为{用户

ID, 用户名, 用户头像, 登陆密码}; 数据库中图书的列表保存形式为{图书 ID, 图书名, 作者, 图书封皮, 评分}; 数据库中图书借阅的列表保存形式为{ID, 用户 ID, 借阅图书 ID, 借阅时间, 用户评分, 图书状态} [10]。

4.3. 系统功能实现与展示

4.3.1. 登录页面的实现与展示

当用户重新登陆计算机或进入系统网址后, 将自动切换至正常登录页面, 一旦其填写了正确的用户名和密码并点击了正确登录按钮将可以进行重新登录, 并由此形成了正确登陆提示, 而一旦填写的所有用户名或者密码都错误了, 则会出现关于用户名和密码的错误提示。用户登录页面。

4.3.2. 猜你喜欢页面的实现与展示

用户在登陆成功后, 会进入到首页, 首先看到的功能为猜你喜欢推荐功能, 本功能的实现就是通过记录用户对图书的浏览行为, 基于用户对图书的偏好找到相似的图书, 然后根据用户的历史偏好, 推荐相似的图书。

4.3.3. 为你推荐页面的实现与展示

“为你推荐”是指, 通过当前用户对物品的偏好搜索邻近用户(相似用户), 从而将邻近用户(相似用户)最喜欢的东西推荐给当前使用者。它通过将每个用户对其他物品的偏好作为一种向量, 计算人之间的相似度。搜索邻居产品之后, 经过对比邻居的产品相似度权重以及他们对产品的偏好等, 选择当前用户最关心的产品之后, 会得到一个排序的产品清单并做出选择。所谓最大相似度, 即是二个向量之间的距离长度。间距越小则最大相似度就越大, 这将有助于实现个性化选择[11]。

在首页还会看到为你推荐功能。根据当前用户对物品的喜好找到类似商品, 进而把类似用户感兴趣的物品介绍给当前用户, 在发现了类似用户后, 可以通过对相似用户的相似率权重和他们对书籍的喜好, 预测当前用户所未提及的书籍。

4.3.4. 用户管理的实现与展示

用户管理功能, 主要包括了用户设置和浏览用户列表功能。在账户增加页面里, 您可以完成新增账户的增加, 选择您要新增的账户后, 选择自拟账户, 增加完成后, 系统会产生增加的提示。在查看用户列表页面下, 我们可以查看所有用户的信息, 包括查看用户序号, 用户名称, 用户头像的功能, 还能对用户信息进行编辑、删除和修改密码的操作[12]。

4.3.5. 图书管理的实现与展示

图书管理页面主要包括图书添加和查看图书列表功能。在图书添加页面下, 我们可以实现新图书的增加, 输入我们想要增加的图书名称、作者、封面, 还能对新增的图书进行评分和描述, 新增成功后, 系统会出现新增成功提示[13]。在查看图书列表页面下, 我们可以查看所有图书的信息, 包括查看图书的编号、名称、作者、封皮、评分、描述的功能, 还能对图书信息进行编辑和删除的操作。

4.3.6. 借书管理的实现与展示

借书管理界面包括输入图书名称和借书账号等信息, 输入自己想借的书便可借书成功。

4.3.7. 还书管理的实现与展示

还书管理页面包括查看还书列表功能, 在查看还书列表页面下, 我们可以查看已还图书的编号、图书名称、还书用户的编号、还书用户的名称、还书时间和图书状态。

5. 结论

系统核心目标达成: 本文成功设计并实现了一个基于数据挖掘和人工智能技术的图书借阅推荐系统。该系统能够自动分析读者的借阅行为, 挖掘其潜在兴趣, 并根据分析结果为读者推荐符合其需求的图书。这一功能的实现, 有效提升了图书馆的服务水平, 满足了读者对个性化服务的需求[14]。

技术实现与创新: 在系统的技术实现过程中, 本文采用了多种先进的技术手段, 包括数据挖掘、机器学习、人工智能等。通过综合运用这些技术, 本文实现了对读者借阅行为的精准分析, 并设计了高效的推荐算法。此外, 本文还采用了 B/S 架构和 MySQL 数据库等成熟的技术方案, 确保了系统的稳定性和可扩展性。

实际应用效果显著: 通过实际测试和应用, 本文设计的图书借阅推荐系统表现出了良好的性能和效果。系统能够准确识别读者的兴趣偏好, 为其推荐符合需求的图书, 显著提高了读者的借阅满意度和图书馆的借阅量。同时, 系统还提供了丰富的管理功能, 方便图书馆管理员对图书和读者信息进行高效管理。

未来展望: 尽管本文设计的图书借阅推荐系统已经取得了一定的成果, 但仍然存在一些改进和优化的空间。例如, 可以进一步丰富数据来源, 引入更多的用户行为数据, 以提高推荐算法的准确性; 可以优化系统界面和交互设计, 提升用户体验; 还可以加强系统的安全性和隐私保护, 确保用户数据的安全性和隐私性。此外, 随着大数据和人工智能技术的不断发展, 未来还可以探索更多创新性的应用场景和技术方案, 以推动图书借阅推荐系统的不断发展和完善[15]。

综上所述, 本文设计的图书借阅推荐系统不仅实现了对读者借阅行为的精准分析和个性化推荐, 还通过实际应用验证了其有效性和可行性。未来, 我们将继续深化研究, 不断优化和完善系统功能, 为图书馆提供更加高效、便捷、个性化的服务。

参考文献

- [1] 崔庆才. Python 3 网络爬虫开发实战[M]. 北京: 人民邮电出版社, 2018.
- [2] 罗刚. 自己动手写网络爬虫[M]. 北京: 清华大学出版社, 2010.
- [3] 韩贝, 马明栋, 王得玉. 基于 Scrapy 框架的爬虫和反爬虫研究[J]. 计算机技术和发展, 2019, 29(2): 139-142.
- [4] 张小秋. 基于 Scrapy 框架的网络爬虫分析与抓取实现[J]. 电脑编程技巧与维护, 2022(2): 18-19, 44.
- [5] 徐鲁辉. Hadoop 大数据原理与应用实验教程[M]. 西安: 西安电子科技大学出版社, 2020.
- [6] 邵方舒. 基于协同过滤及关联规则的个性化图书推荐[D]: [硕士学位论文]. 杭州: 浙江工商大学, 2018.
- [7] 肖仁锋. 基于协同过滤的图书馆个性化推荐方法的研究[D]: [硕士学位论文]. 济南: 山东师范大学, 2017.
- [8] 李越洋. 基于借阅记录的图书个性化推荐方法研究与应用[D]: [硕士学位论文]. 北京: 北方工业大学, 2017.
- [9] 张春. 基于 Hadoop 平台的个性化图书推荐系统的研究[D]: [硕士学位论文]. 马鞍山: 安徽工业大学, 2017.
- [10] Resnick, P. and Varian, H.R. (1997) Recommender systems. *Communications of the ACM*, **40**, 56-58.
<https://doi.org/10.1145/245108.245121>
- [11] 顾倩. 数据挖掘应用于高校图书馆个性化服务的探讨[J]. 图书馆杂志, 2013, 32(8): 63-65.
- [12] 刘军. 数据挖掘在读者阅读需求偏好研究中的应用[J]. 图书馆论坛, 2012, 32(3): 89-93.
- [13] 赵恒. 数据挖掘中聚类若干问题研究[D]: [博士学位论文]. 西安: 西安电子科技大学, 2005.
- [14] 杨江丽, 高凡, 董若剑. 基于数据挖掘的高校图书馆读者行为研究——以西南交通大学图书馆为例[J]. 图书馆研究, 2013, 43(3): 106-110.
- [15] 汪伦, 金朵, 刘明月. 基于用户兴趣度的协同过滤算法[J]. 电脑知识与技术, 2012, 8(19): 4709-4711.