

**本科生毕业论文（设计）**

|  |  |
| --- | --- |
| 题目： | 基于协同过滤算法的陕西省旅游景点推荐 |
|  | 系统设计与实现 |

|  |  |
| --- | --- |
| 学院名称 | 信息工程学院 |
| 作者姓名 | 张浩浩 |
| 专业班级 | 数据科学与大数据技术2021级2班 |
| 作者学号 | 2021130067 |
| 指导教师（职称） | 穆亚荣（讲师） |

2025年5月

**学位论文原创性声明**

**本人声明所呈交的学位论文是我在导师的指导下进行研究工作所取得的研究成果。尽我所知，除文中已经注明引用的内容和致谢的地方外，本论文不包含其他个人或集体已经发表或撰写过的研究成果，也不包含本人或他人已申请学位或其他用途使用过的成果。对本文的研究做出重要贡献的个人和集体，均已在文中作了明确说明并表示谢意。**

**本学位论文若有不实或者侵犯他人权利的，本人愿意承担一切相关的法律责任。**

**作者签名： 日期： 年 月 日**

**学位论文知识产权及使用授权声明书**

**本人在导师指导下所完成的学位论文及相关成果，知识产权归属陕西学前师范学院。本人完全了解陕西学前师范学院有关保存、使用学位论文的规定，允许本论文被查阅和借阅，学校有权保留学位论文并向国家有关部门或机构送交论文的纸质版和电子版，有权将本论文的全部或部分内容编入有关数据库进行检索，可以采用任何复制手段保存和汇编本论文。本人保证毕业离校后，发表本论文或使用本论文成果时署名单位仍为陕西学前师范学院。**

**保密论文解密后适用本声明。**

**作者签名： 日期： 年 月 日**

# 摘 要

随着旅游业的不断发展和人们出行需求的多样化，传统的景点推荐方式已难以满足游客对个性化、精准化服务的需求。为了解决这一问题，本文设计并实现一个基于协同过滤算法的陕西省旅游景点推荐系统。本设计采用Python作为开发语言，结合Django框架构建后端服务，前端则采用Vue框架进行界面展示。数据库方面，使用MySQL存储系统所需的各种数据，包括用户信息、景点信息、轮播图以及用户反馈等。为实现景点推荐功能，本设计引入了协同过滤算法，通过对用户历史评分数据的分析，为用户提供个性化的景点推荐。本设计提供了详细的数据展示功能，如用户评分分布、景点评分统计以及用户性别和年龄段分布等，为管理员提供了全面的数据支持。系统反馈功能则允许用户提交反馈信息，包括主题、信息内容、配图及文件等，增强了系统的互动性和用户体验。本设计在提升旅游信息服务智能化水平的同时，也为区域性旅游推荐系统的开发与研究提供了有益的参考与实践基础。

关键词**：**旅游景点推荐，协同过滤算法，Django框架，MySQL数据库

# Abstract

With the continuous development of the tourism industry and the diversification of people's travel needs, traditional scenic spot recommendation methods are no longer able to meet tourists' demand for personalized and precise services. To address this issue, this article designs and implements a Shaanxi Province tourist attraction recommendation system based on collaborative filtering algorithm. This design uses Python as the development language, combined with the Django framework to build backend services, while the frontend uses the Vue framework for interface display. In terms of database, MySQL is used to store various data required for the system, including user information, attraction information, carousel images, and user feedback. To achieve the function of recommending scenic spots, this design introduces a collaborative filtering algorithm, which provides personalized scenic spot recommendations to users through the analysis of their historical rating data. This design provides detailed data display functions, such as user rating distribution, attraction rating statistics, and user gender and age distribution, providing comprehensive data support for administrators. The system feedback function allows users to submit feedback information, including topics, information content, illustrations, and files, enhancing the interactivity and user experience of the system. This design not only enhances the intelligence level of tourism information services, but also provides useful reference and practical basis for the development and research of regional tourism recommendation systems.

Key words: tourist attraction recommendation, collaborative filtering algorithm;,Django framework;,MySQL database

# 目 录

[摘 要 I](#_Toc22797)

[Abstract II](#_Toc17724)

[目 录 III](#_Toc6024)

[1 绪论 1](#_Toc24776)

[1.1 研究背景与意义 1](#_Toc30349)

[1.1.1 研究背景 1](#_Toc2785)

[1.1.2 研究意义 1](#_Toc3695)

[1.2 国内外研究现状 2](#_Toc15828)

[1.2.1 国外研究现状 2](#_Toc20074)

[1.2.2 国内研究现状 2](#_Toc6539)

[1.3 研究思路与研究内容 3](#_Toc30192)

[1.3.1 研究思路 3](#_Toc18965)

[1.3.2 研究方法 5](#_Toc27031)

[1.3.3 研究内容 6](#_Toc28528)

[2 需求分析与总体设计 7](#_Toc18487)

[2.1 需求分析 7](#_Toc30874)

[2.1.1功能需求分析 7](#_Toc13365)

[2.121非功能需求分析 7](#_Toc13365)

[2.2 总体设计 7](#_Toc30874)

[2.2.1数据处理流程 7](#_Toc13365)

[2.2.2系统总体架构 7](#_Toc13365)

[2.3关键技术 9](#_Toc3246)

[2.3.1 爬虫 9](#_Toc28002)

[2.3.2 Django 9](#_Toc12829)

[2.3.3 Vue 9](#_Toc17304)

[2.3.4MySQL 9](#_Toc29592)

[2.3.5协同过滤算法 9](#_Toc29592)

[2.3.6 Echarts 9](#_Toc27635)

[3 陕西省旅游景点推荐系统设计与实现 16](#_Toc28152)

[3.1 数据采集与处理 16](#_Toc8795)

[3.1.1数据采集 16](#_Toc12719)

[3.1.2数据清洗 16](#_Toc28265)

[3.2 数据库设计 17](#_Toc26507)

[3.3 开发环境 20](#_Toc24548)

[3.4 数据分析 22](#_Toc13966)

[3.4.1 数据读取与评分矩阵构建 22](#_Toc29343)

[3.4.2 用户相似度计算 22](#_Toc16562)

[3.4.3 基于协同过滤的推荐生成 23](#_Toc18518)

[3.5 数据可视化 23](#_Toc12956)

[3.6](#_Toc25830) [系统界面的设计与实现](#_Toc25830) 30

[3.6.1 用户界面设计与实现 30](#_Toc24881)

[3.6.2 管理员界面设计与实现 31](#_Toc6174)

[4 测试与评价 36](#_Toc6055)

[4.1 测试 36](#_Toc1498)

[4.2 评价 40](#_Toc24244)

[5 总结与展望 42](#_Toc4747)

[5.1 总结 42](#_Toc31501)

[5.2 展望 42](#_Toc8482)

[参考文献 43](#_Toc3332)

[致 谢 45](#_Toc30766)

1 绪论

1.1 研究背景与意义

1.1.1 研究背景

随着旅游业的蓬勃兴起，陕西省作为中国的历史文化名省，以其丰富的自然景观和深厚的人文底蕴吸引了大量国内外游客。根据陕西省文化和旅游厅发布的《介绍文旅市场治理、游客权益保护等工作开展情况》，陕西省在2024年接待国内游客8.17亿人次，国内游客总花费7668亿元，同比分别增长14%、16.5%，这彰显了陕西省旅游业强劲的增长势头和巨大的市场潜力。

随着旅游市场的不断扩大和旅游信息的日益丰富，游客在面对海量景点信息时，往往难以快速筛选出符合自身兴趣和需求的景点。传统的景点推荐方式，如基于热门排名或专家推荐，虽能在一定程度上为游客提供参考，但缺乏个性化和精准性，难以满足游客日益多样化的旅游需求。据相关市场调研数据（Fastdata极数《2024上半年中国旅游消费趋势洞察报告》），超过70%~80%的游客表示，在规划旅行时更希望获得基于个人兴趣和偏好的个性化景点推荐服务。

在此背景下，随着大数据和人工智能技术的快速发展，协同过滤算法作为一种有效的个性化推荐技术，已在电商、社交媒体等多个领域取得显著成效。该算法通过分析用户的历史行为数据，挖掘用户之间的相似性，从而为用户推荐可能感兴趣的物品或服务。因此，结合陕西省旅游业的实际需求和协同过滤算法的技术优势，本设计提出了基于协同过滤算法的陕西省旅游景点推荐系统设计与实现，旨在为游客提供更加精准、个性化的景点推荐服务。

1.1.2 研究意义

本设计引入协同过滤算法，使系统能够根据游客的历史行为与兴趣偏好，提供高度个性化的旅游景点推荐服务。这种个性化推荐方式不仅提升了旅游服务的智能化水平，也有效回应了游客日益多样化、个性化的出行需求，极大地优化了游客的决策体验。系统通过精准匹配用户兴趣点，帮助游客快速发现符合自身偏好的景点，减少了信息筛选的成本，显著提升了用户满意度与服务粘性。同时，良好的使用体验还通过用户口碑实现了进一步的推广效应，为旅游行业吸引更多潜在游客、扩大服务影响力提供了助力。

从行业层面来看，本设计为陕西省旅游业的数字化与智能化转型提供了有力支撑。系统借助先进的数据挖掘与分析技术，推动传统旅游服务模式向数据驱动、个性化推荐的新模式升级，成为旅游信息服务创新的重要路径。此外，通过对游客行为数据的深入分析，系统能够为旅游主管部门和相关企业提供科学的决策依据，辅助其准确把握市场动态、发现潜在问题、优化资源配置，从而促进旅游产业朝着更加健康、有序与可持续的方向发展。这对于提升区域旅游竞争力，构建智慧旅游体系具有积极而深远的意义。

1.2 国内外研究现状

1.2.1 国外研究现状

在国外，Nan Xiang 和 Xiaolan（2022）[1]提出了一种基于数据挖掘和协同过滤算法的个性化旅游推荐系统，通过协同挖掘与过滤过程（CMFP），优化了数据处理效率和推荐精度。Zhonghua Wang（2023）[2]提出了结合用户偏好的智能旅游景点推荐模型，利用全局与局部评分信息计算用户兴趣偏好分布，提升了推荐准确率和用户满意度。

在技术融合方面，Song Yang 和 He Yingwei（2023）[3]设计了一个融合人工智能与物联网技术的智能旅游推荐系统，结合Apriori算法与多种协同过滤策略，为游客提供个性化的行程与旅游洞察。Clarice Wong Sheau Harn 和 Mafas Raheem（2023）[4]则利用社交平台中的地理标签数据，提出了一种基于用户行为聚类与矩阵分解的旅游推荐模型，提升了目的地推荐的空间适应性和个性化水平。

在深度学习与协同过滤的结合应用方面，Mashael Aldayel等人（2023）[5]构建了基于协同过滤的旅游推荐系统，综合比较了矩阵分解、双向变分自编码器（BiVAE）和轻量级图卷积神经网络（LightGCN）等算法，验证了深度学习方法在旅游推荐中的优越性能。Shuo Zhang等人（2023）[6]提出了基于多神经网络协同过滤的景点推荐架构（MNCF-AR），通过建模丰富的旅游异构信息与游客旅游轨迹背景，提升了推荐系统对游客偏好与景点特征的理解能力。

Permana 等人（2024）[7]设计了一个基于协同过滤和卷积神经网络的旅游推荐系统，采用改进的余弦相似度和RecommenderNet框架，提升了推荐的准确性与系统响应能力，专注于印尼马杜拉岛的旅游推荐。Luong Vuong Nguyen（2024）[8]提出了OurSCARA模型，结合情感分析、用户画像与实时数据，融入环境与文化意识，为可持续旅游提供个性化、生态友好的推荐服务。

Yin Ying（2024）[9]提出了一种基于协同过滤推荐算法的文化创意产业与旅游产业融合路径，通过挖掘数据中的隐藏信息并构建偏好矩阵，为大数据背景下产业融合与地区经济发展提供了可行的路径。

综上所述，国外在旅游推荐系统领域的研究注重算法融合、深度学习的应用以及个性化推荐策略的优化，为提高推荐系统的准确性和用户体验提供了重要参考。

1.2.2 国内研究现状

近年来，随着大数据与人工智能技术的快速发展，国内在旅游推荐系统领域取得了显著进展，特别是在基于协同过滤算法的景点推荐方面，学者们进行了深入研究与实践。

李伟（2021）[10]提出了基于大数据挖掘技术的智慧旅游推荐系统，采用改进的协同过滤算法与Mahout分布式平台，实现了高效、个性化的旅游信息服务。陈勇（2022）[11]设计了基于协同过滤的旅游推荐系统，结合爬取的景区网站数据与用户历史行为，提升了推荐的个性化水平。陈丹等人（2024）[12]针对旅游信息冗杂的问题，构建了基于混合协同过滤的智能推荐系统，有效提高了推荐结果的多样性与覆盖率。

在推荐算法优化方面，周喜平等人（2023）[13]结合Redis GEO算法与协同过滤方法，实现了基于地理位置的景点推荐；黄士新等人（2023）[14]引入季节变化信息，通过SVD改进评分预测，增强了系统的动态适应性。李晓芳（2023）[15]提出基于用户与景点画像及注意力机制的旅游路线推荐方法，满足了更加个性化的出行需求；王茸等人（2023）[16]应用Apriori算法挖掘景点关联规则，进一步丰富了推荐内容。

在系统开发与架构方面，张俊萌、梁志达（2024）[17]等研究了基于ECharts的数据可视化，为推荐系统结果展示提供了有益借鉴。兰琳琳（2024）[18]与刘昶等人（2024）[19]通过基于MySQL-Django-Vue框架的项目实践，展示了前后端分离架构在Web应用开发中的优势，为旅游推荐系统的技术选型提供了参考。

此外，李彭（2024）[20]提出结合用户特征与信任度的协同过滤方法，提高了推荐精准度；卞红胜（2024）[21]结合神经网络与K-means聚类，提出混合推荐算法，缓解了冷启动问题；孙文婧（2024）[22]基于因果推理设计去偏推荐系统，进一步提升了系统的公平性和可信度。

综上所述，国内外学者在旅游推荐系统领域均开展了深入研究。国外侧重于算法融合、深度学习与可持续旅游理念的结合，提出了多种创新模型，提升了系统的智能化与生态友好性；而国内研究则以协同过滤算法为核心，结合大数据、可视化及主流Web框架，探索了旅游推荐系统在性能优化、用户体验和系统实现层面的多种方案。这些研究成果为旅游推荐系统的理论发展和实践应用提供了坚实基础，也为本设计的研究方向和方法选择提供了重要参考。

1.3 研究思路与研究内容

1.3.1 研究思路

在本设计中，研究聚焦于基于协同过滤算法的陕西省旅游景点推荐系统的设计与实现，涵盖数据采集、清洗与预处理、存储、分析、可视化及系统页面设计等关键环节。

数据采集:是推荐系统的基础，主要包括用户行为数据和景点信息数据。前者如评分记录、浏览历史等，反映用户兴趣；后者则包含景点的名称、位置、类型、简介和评分等，为生成推荐提供支持。通过用户输入、第三方API和历史数据导入等方式，确保数据来源多样、内容全面。

数据清洗与预处理:旨在提升数据质量。针对噪声、缺失值和重复数据问题，采用去重、均值填充、插值等技术进行处理。通过标准化、归一化等手段消除量纲差异和异常值，为后续建模和分析打下基础。

数据存储:方面，系统选用 MySQL 数据库，构建了包括用户信息、景点信息、轮播图、反馈记录和用户行为等在内的表结构，保障了数据的完整性、一致性与高效访问。

数据分析:是推荐系统的核心。通过统计分析和数据挖掘，对用户评分数据构建用户-景点评分矩阵，并基于相似度计算预测未评分项。同时对评分数据进行统计，如平均分、最高分等，为推荐提供依据。结合性别、年龄等基本信息构建用户画像，增强个性化推荐的针对性。

数据可视化:则借助 Echarts 等工具，将复杂数据转化为柱状图、折线图、饼图等形式，直观展示用户评分分布、景点评分统计和用户画像，帮助管理员掌握系统状态，也为用户提供清晰的决策参考。

系统页面设计:采用 Vue 框架开发前端、Django 实现后端服务，构建前后端分离架构。界面设计注重交互体验，支持景点浏览、个性化推荐、数据图表展示和反馈提交，具备良好的响应式布局，确保在不同设备上的访问效果一致。

1.3.2 研究方法

在本设计的研究过程中，秉持科学、系统的原则，采用了多种研究方法以确保研究的全面性和深入性。这些方法相互支撑，共同推动着课题的顺利进行。

文献研究法：通过系统查阅、整理和分析已有的文献资料，了解研究领域的发展现状和前沿动态，借鉴前人的研究成果与经验，为课题的理论框架构建、研究方法选择及问题分析提供科学依据和有力支持。

通过广泛而深入的文献调研，系统地梳理了国内外在旅游推荐系统领域的研究脉络与发展趋势。从相关书籍、期刊论文到会议报告，细致入微地分析了协同过滤算法的原理、应用实例及其在不同领域的成功实践。这些文献资料不仅为本设计奠定了坚实的理论基础，更为设计基于协同过滤算法的陕西省旅游景点推荐系统提供了宝贵的参考与借鉴。通过对比分析不同研究中的算法实现策略、系统架构设计以及用户体验设计，汲取了众多优点与灵感，为本设计的创新发展开辟了新的思路。

功能分析法：通过分析系统或产品的各项功能需求，明确其功能组成及相互关系，从而指导系统结构设计与实现过程的一种方法。它有助于提升系统的逻辑性、完整性和实用性。

在明确研究目标与需求后，运用功能分析法对系统进行了详尽的功能规划与设计。通过深入分析旅游景点推荐系统的核心功能需求，精准定位了用户角色（包括普通用户与管理员）及其对应的权限与功能。针对普通用户，精心设计了个性化推荐、数据展示、系统反馈等多元化功能；而针对管理员，则提供了用户管理、角色管理、轮播图管理以及系统反馈管理等全方位功能支持。通过功能分析法的运用，确保了系统功能的全面覆盖与实用性，为后续的系统开发与测试工作奠定了坚实的基础。

用户参与法：通过邀请真实用户参与系统设计、开发与测试全过程，收集用户需求与反馈，确保系统功能、界面和操作流程更贴合用户实际使用习惯，从而提升系统的可用性和用户满意度。

在系统开发与测试的全过程中，始终高度重视用户的参与与反馈。通过用户调研、问卷调查、用户访谈等多种方式，广泛收集了用户需求与意见，为本设计的功能设计与优化提供了重要参考。此外，还邀请了一部分目标用户参与本设计的测试工作，通过他们的实际操作与反馈，及时发现并修复了本设计中存在的问题与不足。

1.3.3 研究内容

绪论部分深入剖析了研究背景与意义，指出随着旅游业蓬勃发展和用户个性化需求的日益增长，传统推荐方式的局限性日益凸显，而协同过滤算法在个性化推荐领域展现出显著优势。概述了国内外旅游推荐系统的研究现状，为本设计明确了思路与方向。

相关技术和工具章节详细介绍了系统开发所依托的关键技术栈，包括Python语言的简洁高效、Django框架的强大后端支持、Vue框架的前端灵活性、MySQL数据库的高性能与可靠性、Echarts可视化库的直观展示能力，以及协同过滤算法的个性化推荐功能。这些技术和工具的选择充分考虑了其优势和适用性，为系统的开发与实现提供了坚实支撑。

需求分析部分则全面探讨了系统的功能需求与非功能性需求。功能需求涵盖了用户个性化推荐、数据直观展示、用户反馈机制等方面；而非功能性需求则聚焦于系统性能、可用性、安全性、可扩展性及可维护性，确保系统不仅功能完备，而且运行稳定、安全可靠。

系统设计章节详细阐述了系统的整体架构与模块设计，包括数据获取与处理的流程、模型构建的策略、功能模块的划分（如用户功能、管理员功能、景点推荐功能、数据展示功能及系统反馈功能）以及数据库的设计等。这些设计旨在构建一个高效、稳定且易于使用的旅游景点推荐系统。

系统实现部分则具体描述了各模块的实现过程，从数据库模型的构建到协同过滤算法的实现，从数据可视化的呈现到用户与管理员功能的开发，再到景点推荐功能的完善，每一步都经过精心设计与实现。通过详细的代码编写与界面设计，系统最终得以成功构建并投入运行。

系统测试章节对系统的各项功能进行了全面而细致的测试，包括用户功能的验证、管理员操作的审核、景点推荐效果的评估以及数据展示与系统反馈的准确性检查等。这些测试确保了系统的稳定运行与功能的正确实现，为系统的正式上线与广泛应用提供了有力保障。

总结与展望部分则对本设计的研究工作进行了全面回顾与总结，既肯定了系统开发的成果与成效，也指出了存在的不足与改进空间。对未来的研究方向与发展趋势进行了展望，提出了优化系统性能、提升推荐精准度、加强数据安全保护等具体设想与计划，为系统的持续进步与发展指明了方向。

2 需求分析与总体设计

2.1需求分析

2.1.1 功能需求分析

用户功能需求：系统为普通用户打造全面个性化的服务体验。通过深度分析用户历史行为与偏好，精准推荐符合其兴趣和需求的陕西省旅游景点，提升旅游规划效率与满意度。同时，提供直观数据展示，包括用户评分分布、景点评分统计等，助力用户全面了解市场动态。用户可便捷提交反馈，增强系统互动性，提升用户归属感。支持用户查看修改个人信息，确保信息安全准确，增强用户信任。

管理员功能需求：管理员通过严格登录权限管理，确保系统安全高效运行。可查看编辑用户信息及角色，保障系统正常运作。灵活管理轮播图，提升系统吸引力。高效处理用户反馈，增强用户满意度。提供接口文档查看与测试功能，降低开发成本。通过数据监控分析，实时了解系统状态与用户行为，为优化决策提供依据。

景点推荐功能需求：景点推荐作为核心，采用协同过滤算法，持续优化性能，提升推荐精准度与个性化。实时更新数据，确保推荐时效准确。提供多样化推荐方式，满足用户不同需求。根据反馈调整算法，优化推荐结果，提升用户满意度与忠诚度，推动陕西旅游业发展。

2.1.2非功能需求分析

可用性需求：系统界面设计应追求极致的简洁与直观，让用户一目了然，轻松上手，有效降低学习门槛，提升整体用户体验。此外，系统需全面兼容多设备访问，无论是传统桌面端、便捷移动端还是新兴智能设备，都能确保用户享受到一致且卓越的使用感受，灵活适应各种应用场景下的用户需求。

可扩展性需求：系统架构设计需具备前瞻性与灵活性，充分考虑未来的功能扩展与数据增长需求，确保系统能够轻松应对业务变迁与技术革新带来的挑战。同时，系统应提供丰富多样的接口设计，便于与其他系统与平台进行无缝对接与集成，实现资源的深度共享与业务的紧密协同。

可维护性需求：系统代码结构应条理清晰，注释详尽完备，为开发人员提供极大的便利，使其能够迅速理解系统逻辑，高效完成维护与升级工作。此外，系统还需配备完善的日志记录功能，对系统运行状态与用户操作行为进行全程追踪与记录，为故障排查与系统监控提供有力支撑，确保系统能够长期稳定、高效地运行。

2.2总体设计

2.2.1数据处理流程

数据处理流程是基于协同过滤算法的陕西省旅游景点推荐系统的核心部分，主要包括以下几个步骤：首先，在数据采集阶段，系统通过网络爬虫技术从各大旅游网站获取景点基本信息、用户评论、评分记录等数据，同时结合用户的浏览、搜索和收藏行为收集用户偏好数据；接着进入数据预处理阶段，采用Python与Pandas工具对数据进行清洗，去除缺失、异常及重复记录，并统一数据格式，确保数据的一致性和准确性；随后，在数据存储阶段，将处理后的用户信息、景点信息及行为数据存入MySQL数据库，为后续的推荐算法提供稳定的数据支撑；在数据分析阶段，利用协同过滤算法分析用户兴趣相似度，生成个性化的景点推荐结果；最后，在结果展示阶段，通过可视化界面向用户呈现推荐景点列表及相关统计信息，提升用户体验。整个流程环环相扣，保证了推荐系统的数据处理高效性与推荐结果的准确性。

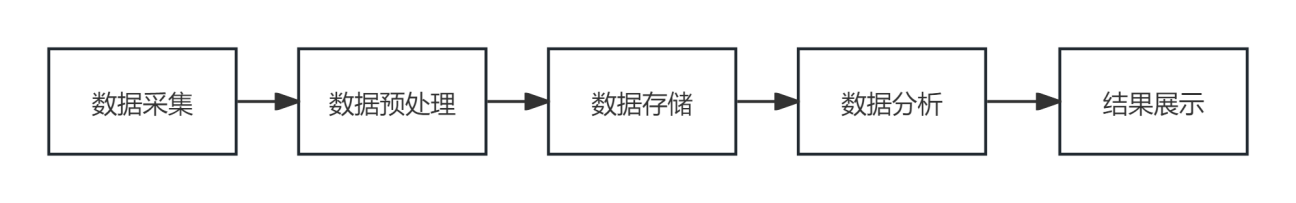


图2-1数据处理流程图

2.2.2系统总体架构

本旅游景点推荐系统以协同过滤算法为核心，旨在为陕西省地区的游客提供个性化景点推荐服务。系统采用分层架构设计思想，将整个系统划分为数据存储层、业务逻辑层和表示层，各层之间通过明确接口进行通信，确保系统具有良好的可扩展性、可维护性和可移植性。整体架构遵循高内聚、低耦合的设计原则，使得各模块之间职责清晰，功能独立，便于后期系统升级和功能拓展。

数据存储层：主要负责系统中的所有数据管理与存取操作，包括用户数据、景点信息、用户评分数据等。系统采用关系型数据库（如MySQL）进行数据存储，通过建立规范化的数据表结构，确保数据的一致性与完整性。为了支持推荐算法的数据需求，系统对用户评分数据进行了特别的处理，保证能够快速提取用户-景点评分矩阵。此外，部分静态资源如景点图片、旅游攻略等，也可以存储在对象存储服务中，以提高访问效率。

业务逻辑层：是系统的核心部分，主要负责实现各项具体功能，包括用户管理、景点管理、推荐算法处理和数据交互逻辑。其中，协同过滤算法模块基于用户评分数据，通过计算用户之间或景点之间的相似度，生成个性化的景点推荐列表。系统采用了经典的用户协同过滤（User-based CF）和物品协同过滤（Item-based CF）两种方式，并根据实际应用场景灵活选择。推荐算法模块内部还包括数据预处理（如标准化处理、缺失值填充）、相似度计算（如余弦相似度、皮尔逊相关系数）以及推荐列表生成等子模块。

表示层：即用户交互界面，主要负责向用户展示系统的各项功能和推荐结果，同时接收用户的操作输入。系统提供了友好的界面设计，包括用户注册与登录、景点浏览与搜索、个性化推荐展示等功能模块。为了提升用户体验，界面设计注重简洁美观，同时响应式布局保证了在不同设备上的良好适配性。前端页面与后端服务器通过HTTP接口进行数据交互，采用RESTful API规范，提高了通信的标准化程度。

为了实现系统功能的高效开发与维护，整个系统进一步划分为若干功能模块，主要包括：

用户管理模块：负责用户注册、登录、信息修改、历史记录管理等功能。

景点信息管理模块：用于管理景点的基本信息，包括名称、描述、地理位置、评分、评论等。

推荐引擎模块：实现基于协同过滤的推荐逻辑，动态生成个性化推荐列表。

数据采集与更新模块：定期采集新的景点信息与用户交互数据，保证推荐系统数据的时效性与准确性。

系统监控与日志模块：记录系统运行状态与异常信息，为后续系统维护与优化提供依据。

系统的数据流转过程如下：首先，用户在前端界面进行注册、登录及浏览景点等操作，系统将用户行为数据上传至服务器。后端业务逻辑层接收数据后进行处理与存储，并根据用户的历史行为数据，通过协同过滤算法生成推荐结果。推荐结果再由后端接口发送至前端，展示给用户。用户对推荐内容的反馈（如点赞、收藏）将再次作为输入数据，持续优化推荐模型，实现推荐效果的不断迭代提升。

本系统设计注重高效性、可扩展性与用户体验。通过采用协同过滤算法，能够根据用户兴趣动态调整推荐结果；模块化架构使得系统在未来可以方便地引入新的推荐技术，如基于深度学习的推荐方法；同时，良好的界面交互设计和系统响应速度，保证了用户使用过程的流畅性和满意度。

综上所述，基于协同过滤算法的陕西省旅游景点推荐系统通过合理的系统架构设计，实现了数据管理、推荐服务与用户体验的有机结合，为陕西省旅游行业的信息化发展提供了有力支持。

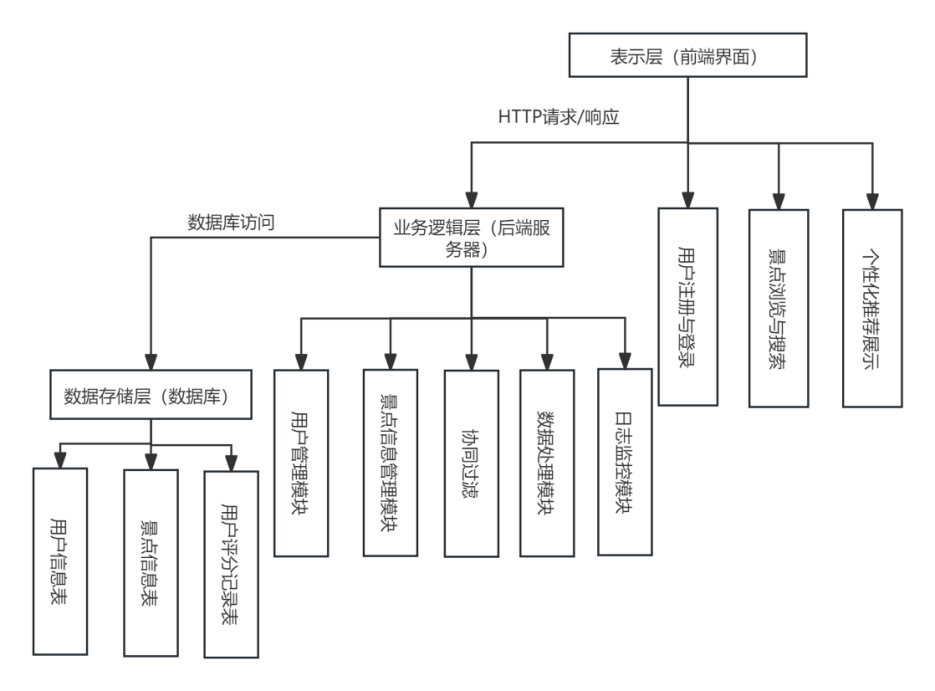


图2-2 系统总体架构图

2.3关键技术

2.3.1 爬虫

爬虫是一种按照设定规则自动抓取网页信息的程序，广泛应用于数据采集与信息获取领域。其基本原理是模拟用户在互联网上的浏览行为，从一个或多个网页入口出发，通过提取页面中的链接不断访问和下载网页内容。爬虫程序通常包括请求发送、网页解析、数据提取、链接提取和去重过滤等模块。常用技术工具有 Python 的 requests、BeautifulSoup、Scrapy 框架等。在旅游推荐系统中，爬虫可用于获取旅游景点信息、用户评价、图片资料等，为系统提供丰富的数据支撑。为了确保数据的准确性与完整性，爬虫在设计过程中需考虑反爬机制应对、数据清洗及更新策略。同时，遵守网站的 robots 协议和相关法律法规是实施爬虫的基本前提，避免侵犯网站权益和用户隐私。爬虫技术的有效应用，极大提高了数据获取的效率和质量。

2.3.2 Django

Django是一个用Python编写的高级Web框架，它遵循MVC（模型-视图-控制器）设计模式，使得Web应用的开发更加结构化和易于维护。在本设计中，Django被用作后端框架，因为它提供了强大的ORM（对象关系映射）功能，使得数据库操作更加便捷。Django还支持用户认证、会话管理、模板渲染等核心功能，为系统的安全性、可扩展性和可维护性提供了有力保障。在Django框架中，实现了用户管理、角色管理、轮播图管理、系统反馈管理以及接口文档查看和测试等核心功能。这些功能通过Django的Admin后台进行统一管理，大大提高了开发效率和系统维护的便捷性。

2.3.3 Vue

Vue是一个构建用户界面的渐进式JavaScript框架，它易于上手且功能强大。在本设计中，Vue被用作前端框架，用于构建用户友好的界面和交互体验。Vue的组件化设计使得前端页面的开发更加模块化和可复用，有助于降低开发成本和提高开发效率。在Vue框架中，实现了经典推荐、数据展示和系统反馈等核心功能的前端展示。通过Vue的数据绑定和事件处理机制，用户可以轻松地输入用户ID获取推荐景点，查看用户评分分布、景点评分统计等数据信息，并发布系统反馈。

2.3.4MySQL

MySQL是一个开源的关系型数据库管理系统，它以其高性能、可靠性和易用性而广受好评。在本设计中，MySQL被用作数据库管理系统，用于存储和管理用户信息、管理员信息、角色信息、轮播图信息以及反馈信息等核心数据。MySQL支持SQL语言，使得数据的查询、插入、更新和删除等操作变得简单而高效。MySQL还支持事务处理、索引优化等高级功能，有助于提高数据库的可靠性和性能。在本设计中，利用MySQL的存储过程和触发器等技术，实现了数据的自动化处理和业务逻辑的封装。

2.3.5协调过滤算法

协同过滤算法，作为推荐系统的基石，其深厚的理论基础与广泛的应用实践共同彰显了其在个性化推荐领域的卓越价值。该算法的核心在于深入剖析用户与物品间的交互数据，诸如评分、点击行为及浏览记录等，以精准捕捉用户的兴趣偏好，进而实现个性化内容的精准推送。以下是对协同过滤算法理论的详尽概述，涵盖其训练流程、用户与物品的过滤机制，以及最终的推荐策略。

2.3.6 Echarts

Echarts是一个使用JavaScript实现的开源可视化库，它提供了丰富的图表类型和交互功能，使得数据的可视化展示变得简单而直观。在本设计中，Echarts被用作数据可视化工具，用于展示用户评分分布、用户评分统计、各景点的评分统计等数据信息。通过Echarts的图表类型选择和数据配置，可以轻松地生成柱状图、折线图、饼图等不同类型的图表，以直观地展示数据的分布和趋势。Echarts还支持数据的动态更新和交互操作，使得用户可以更加深入地了解数据背后的信息和规律。

3 陕西省旅游景点推荐系统设计与实现

3.1 数据采集与处理

在数据获取与处理部分，详细规划了数据的来源、采集方式以及数据的属性，以确保为陕西省旅游景点推荐系统提供丰富、准确且多样化的数据支持。

3.1.1数据采集

主要从以下几个渠道采集数据：

用户行为数据的汇聚：整合了系统内部的原生记录与外部第三方数据源，形成了丰富的用户行为数据体系。系统内部记录深入捕捉了用户在平台上的多维度行为轨迹，细致入微地记录了评分记录、浏览历程、搜索关键词以及点击行为等关键信息，这些数据直接映射出用户的兴趣偏好与行为模式。同时，积极拓展合作网络，通过合法合规的途径购买或合作获取第三方数据源，不断丰富并扩充的用户行为数据库，为构建更加精准的用户画像奠定了坚实的基础。

景点信息数据的搜集：针对陕西省内众多独具魅力的旅游景点，广泛搜集了来自陕西省文旅厅开放API获取官方景点库以及携程/马蜂窝等OTA平台的信息资源。这些数据详尽地覆盖了景点的名称、精准的地理位置（包括经纬度信息）、类型分类（如自然风光、历史文化遗址、主题公园等）、开放时间安排、门票价格体系、官方联系方式以及用户评分与评论等丰富多样的属性信息，为推荐算法提供了坚实可靠的数据支撑。

辅助数据资源的整合：为了进一步提升推荐的精准度与个性化水平，还广泛搜集了包括天气信息、节假日活动安排、实时交通状况等在内的多元化辅助数据资源。尽管这些数据并不直接参与推荐算法的核心计算流程，但它们作为推荐决策的重要参考因素，对于优化用户体验、增强推荐结果的实用性具有举足轻重的价值。这些辅助数据资源的整合，使得的推荐系统能够更加全面地洞察用户需求，从而生成更加贴心、实用的推荐结果，进一步提升用户的满意度与忠诚度。

3.1.2数据清洗

在本次数据清洗过程中，首先读取了原始的用户评分数据，并对各字段进行了缺失值统计。考虑到缺失数据在本项目中假设不具备实际业务意义，因此直接对存在缺失的记录进行了删除，以提高数据完整性。随后，对数据集进行了重复值检测，删除了所有完全重复的行，确保每条记录的唯一性，从而避免因重复样本对后续分析结果造成偏差。在异常值处理环节，依据业务规则筛选出评分在1至5分之间的有效数据，同时限定用户年龄在10至100岁之间，剔除了异常或录入错误的数据记录。为提升数据的一致性与规范性，对性别字段进行了标准化处理，将不同表述（如"male"、"female"）统一转换为“M”和“F”，并删除了非标准性别标记的记录。同时，对用户偏好出行季节（preferred\_travel\_season）字段进行了规范化，将季节名称统一为首字母大写的标准格式，提升数据的统一性与可读性。在数据类型转换方面，将年龄字段统一转换为整数类型，保证字段数据类型与实际含义相符合。经过以上清洗步骤后，最终得到了一份更加完整、规范、一致且质量更高的数据集，并将其保存为“cleaned\_user\_ratings.csv”，为后续的数据分析与建模工作奠定了坚实的基础。

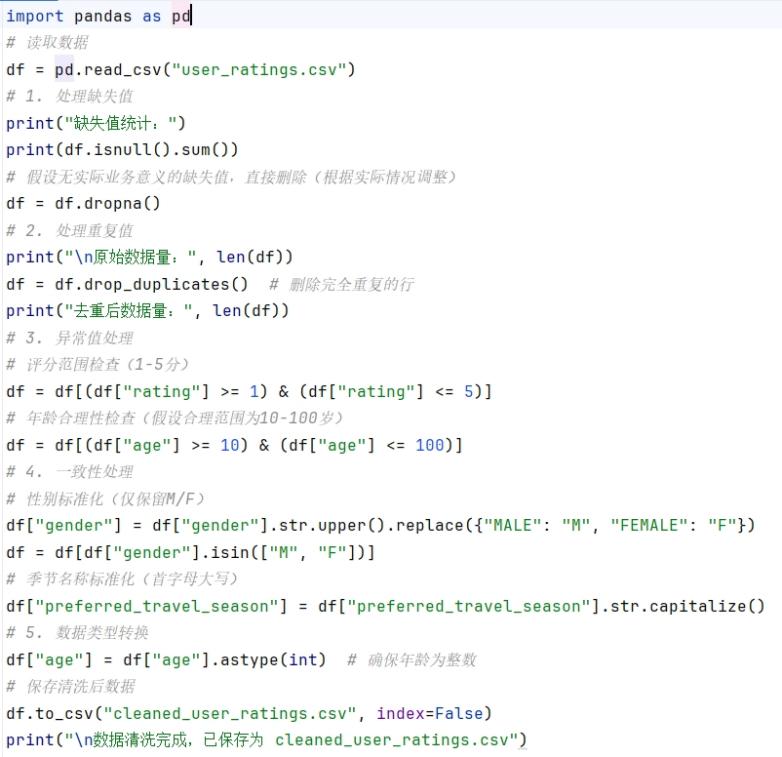


图3-1 数据清洗代码

3.2 数据库设计

本设计采用 MySQL 作为数据库管理系统，设计了合理、完整的数据表结构，以支撑旅游推荐系统的各项功能模块。整体遵循高内聚、低耦合的设计原则，确保数据存取的高效性与系统扩展的灵活性。各数据表及其字段设计如下:

用户信息表:主要用于存储平台注册用户的基本资料。字段包括用户ID、用户名、密码、姓名、性别、年龄及联系方式。其中，用户ID作为主键，唯一标识每一位用户，确保数据的准确性与独立性。用户名和密码用于系统登录认证，姓名、性别、年龄、联系方式则用于完善用户画像，便于后续实现个性化推荐和精准推送功能。

表3-1 用户信息表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 字段类型 | 字段含义 | 是否主键 | 是否外键 | 关联表 |
| 用户ID | INT | 用户唯一标识 | 是 | 否 | - |
| 用户名 | VARCHAR(50) | 用户登录名 | 否 | 否 | - |
| 密码 | VARCHAR(255) | 用户密码（加密） | 否 | 否 | - |
| 姓名 | VARCHAR(50) | 用户真实姓名 | 否 | 否 | - |
| 性别 | VARCHAR(10) | 用户性别 | 否 | 否 | - |
| 年龄 | INT | 用户年龄 | 否 | 否 | - |
| 联系方式 | VARCHAR(20) | 用户联系方式 | 否 | 否 | - |

管理员信息表:用于存储系统后台管理员的账户信息。字段包括管理员ID、用户名、密码、姓名、角色和所属部门等。管理员ID为主键，用以唯一标识各管理员账户。角色字段用于标识管理员的权限等级，如普通管理员、高级管理员等，确保系统在多角色环境下的正常运作。所属部门字段则用于组织管理，便于后续的权限细分与管理优化。

表3-2 管理员信息表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 字段类型 | 字段含义 | 是否主键 | 是否外键 | 关联表 |
| 管理员ID | INT | 管理员唯一标识 | 是 | 否 | - |
| 用户名 | VARCHAR(50) | 管理员登录名 | 否 | 否 | - |
| 密码 | VARCHAR(255) | 管理员密码 | 否 | 否 | - |
| 姓名 | VARCHAR(50) | 管理员真实姓名 | 否 | 否 | - |
| 角色 | VARCHAR(50) | 管理员角色 | 否 | 否 | 角色信息表 |
| 所属部门 | VARCHAR(100) | 管理员所属部门 | 否 | 否 | - |

景点信息表:用于存储陕西省内各个旅游景点的详细资料。字段涵盖景点ID、名称、位置（经纬度）、类型分类、简介、图片URL、评分以及评论数量等。景点ID为主键，用于唯一标识每个景点。通过此表，系统能够向用户展示详尽的景点介绍，并支持按类型分类浏览、基于地理位置筛选推荐等功能。评分与评论数量则为推荐算法提供了可靠的数据依据。

表3-3 景点信息表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 字段类型 | 字段含义 | 是否主键 | 是否外键 | 关联表 |
| 景点ID | INT | 景点唯一标识 | 是 | 否 | - |
| 名称 | VARCHAR(100) | 景点名称 | 否 | 否 | - |
| 位置（经纬度） | VARCHAR(50) | 景点位置 | 否 | 否 | - |
| 类型 | VARCHAR(50) | 景点类型 | 否 | 否 | - |
| 简介 | TEXT | 景点简介 | 否 | 否 | - |
| 图片URL | VARCHAR(255) | 景点图片URL | 否 | 否 | - |
| 评分 | FLOAT | 景点平均分 | 否 | 否 | - |
| 评论数量 | INT | 景点评论数量 | 否 | 否 | - |

反馈信息表:用于记录用户在系统中提交的反馈意见及建议。字段包括反馈ID、关联用户ID、反馈主题、反馈内容、配图URL、视频URL、反馈时间、是否已解决及管理员回复。反馈ID为主键，用户ID为外键关联用户信息表。通过系统化管理用户反馈，管理员可以及时了解用户需求和系统存在的问题，持续优化用户体验，提升服务质量。

表3-4 反馈信息表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 字段类型 | 字段含义 | 是否主键 | 是否外键 | 关联表 |
| 反馈ID | INT | 反馈唯一标识 | 是 | 否 | - |
| 用户ID | INT | 反馈用户ID | 否 | 是 | 用户信息表 |
| 反馈主题 | VARCHAR(100) | 反馈主题 | 否 | 否 | - |
| 反馈信息 | TEXT | 反馈信息内容 | 否 | 否 | - |
| 配图URL | VARCHAR(255) | 反馈配图URL | 否 | 否 | - |
| 视频URL | VARCHAR(255) | 反馈视频URL | 否 | 否 | - |
| 反馈时间 | DATETIME | 反馈提交时间 | 否 | 否 | - |
| 是否已解决 | BOOLEAN | 反馈是否已解决 | 否 | 否 | - |
| 管理员回复 | TEXT | 管理员回复内容 | 否 | 否 | - |

用户-景点交互表:是记录用户与景点之间互动数据的重要表格。字段包括记录ID、用户ID、景点ID、评分和评分时间。记录ID作为主键，用户ID和景点ID分别作为外键关联用户信息表与景点信息表。该表主要用于保存用户的评分行为，是协同过滤推荐算法生成个性化推荐列表的核心数据来源，对于提升推荐系统的精准度和智能化水平起到了关键作用。

表3-5 用户-景点交互表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 字段类型 | 字段含义 | 是否主键 | 是否外键 | 关联表 |
| 记录ID | INT | 记录唯一标识 | 是 | 否 | - |
| 用户ID | INT | 交互用户ID | 否 | 是 | 用户信息表 |
| 景点ID | INT | 交互景点ID | 否 | 是 | 景点信息表 |
| 评分 | FLOAT | 用户对景点的评分 | 否 | 否 | - |
| 评分时间 | DATETIME | 评分提交时间 | 否 | 否 | - |

3.3 开发环境

本系统的开发环境选择注重稳定性、易用性与高效性，主要包括前端开发、后端开发、数据库管理以及辅助工具配置等方面。

在后端开发方面，选用了Python作为主要开发语言，利用其简洁高效的特性以及丰富的第三方库资源，有效提升了系统功能的实现效率。后端框架采用Django，凭借其完善的MVC架构、强大的ORM功能和良好的扩展性，极大地简化了数据库操作与接口开发，保证了后端服务的高效性和可维护性。

前端开发采用了Vue.js框架，实现了前后端分离的开发模式。Vue.js以其轻量、响应式的数据绑定特性和组件化开发模式，提高了前端页面的交互性与灵活性，配合Element-UI组件库，快速搭建了简洁美观、用户体验良好的系统界面。同时，使用Axios进行前后端数据交互，保证了页面响应的流畅性与数据传输的稳定性。

数据库部分采用了MySQL关系型数据库管理系统，因其高性能、高可靠性和易维护性，能够支撑系统在多用户并发访问时的数据存储与检索需求。数据库管理过程中，配合使用Navicat for MySQL进行可视化管理，提升了数据建模、查询和维护的便捷性。

开发工具方面，后端主要使用PyCharm进行Python开发，前端部分使用Visual Studio Code进行Vue.js开发，均配置了常用插件以提高代码编写和调试的效率。开发环境运行在Windows 10操作系统上，测试环境使用本地搭建的XAMPP集成环境，后续可部署至云服务器（如阿里云）以实现线上运行和访问。

此外，为了支持数据可视化需求，引入了Echarts库，用于动态展示用户行为分析结果、景点评分统计等信息。推荐算法模块则结合使用了Scikit-Learn与Surprise库，分别用于数据预处理、相似度计算与推荐模型的构建。数据清洗与标准化工作由Pandas库完成，保证了数据质量，为后续的推荐计算奠定了基础。

整体而言，本系统开发环境配置合理、工具齐全，能够有效支撑系统的功能开发、数据处理、推荐计算以及后期的部署与运维工作，为项目的顺利实施提供了有力保障。

表3-6 开发环境

|  |  |  |
| --- | --- | --- |
| 类别 | 工具/技术 | 版本/说明 |
| 后端开发 | Python、Django | 3.8/3.2 |
| 前端开发 | Vue.js、Element-UI | 2.x |
| 数据库 | MySQL | 8.0 |
| 开发工具 | PyCharm、VS Code | 2021/1.76 |
| 数据分析库 | Pandas、NumPy | 最新版 |
| 可视化工具 | Echarts | 最新版 |
| 测试环境 | XAMPP | 3.3.0 |
| 操作系统 | Windows 11 | 专业版64位 |

3.4 数据分析

在本系统中，基于用户协同过滤（User-based Collaborative Filtering）方法进行景点推荐。为了实现个性化推荐，需要首先对用户评分数据进行深入处理和分析。本节将详细介绍从数据读取、评分矩阵构建、相似度计算，到最终生成推荐结果的完整流程

3.4.1数据读取与评分矩阵构建

本设计使用Pandas库读取了存储在本地的用户评分数据集。数据集文件user\_ratings.csv包含了用户ID、景点名称以及用户对景点的评分三个字段。读取后的数据采用DataFrame的结构进行管理，方便后续操作。为了进行协同过滤，需要将原始数据转化为用户-景点评分矩阵，即以用户为行、景点为列，单元格中填入对应的评分值。未评分的景点数据在转化过程中为空（NaN），因此需要进一步处理，统一用0进行填充，以表示用户对该景点未进行评分。

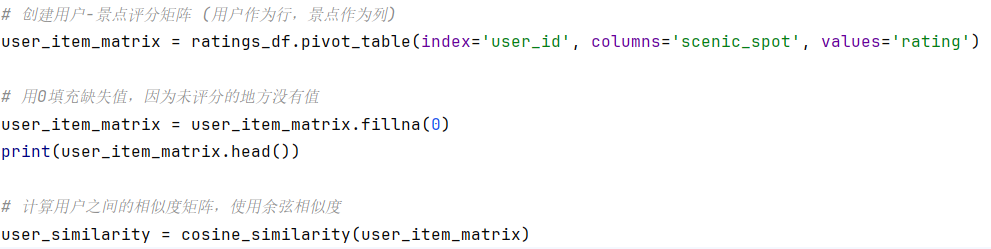


图3-2 评分矩阵代码

经过上述处理，得到的评分矩阵使得所有用户对各景点的偏好情况一目了然。这种稀疏矩阵不仅直观，同时也为后续的用户相似度计算提供了统一的数据基础。尤其是在旅游推荐场景中，用户往往只对少数景点进行过评分，因此合理处理未评分数据尤为重要。

**3.4.2用户相似度计算**

在获得评分矩阵后，系统需要度量用户之间的兴趣相似性。为此，采用了余弦相似度（Cosine Similarity）方法。余弦相似度是计算两个向量之间夹角余弦值的一种方法，能够有效衡量两个用户在兴趣空间中方向的一致性。其数学表达式如下：



其中，A和B分别表示两个用户的评分向量，||A||和||B||表示各自的向量范数。本设计使用Scikit-learn中的cosine\_similarity函数来计算用户之间的相似度矩阵，代码如下：

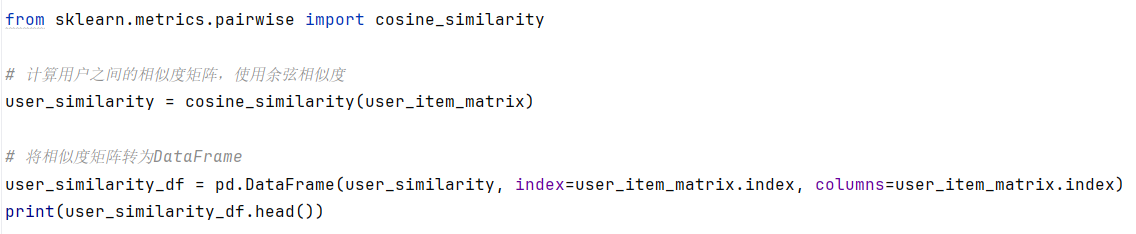


图3-3 相似度矩阵代码

最终得到的用户相似度矩阵中，每一个元素表示两位用户之间的相似程度。对角线元素为1，代表用户与自身完全一致。非对角线的数值越接近1，表示两个用户在景点评分上的偏好越相似。这一步是实现用户协同推荐的关键环节，能够帮助系统准确识别出具有相似兴趣的用户群体。

3.4.3 基于协同过滤的推荐生成

在获得用户相似度矩阵之后，系统根据协同过滤的基本思想，基于相似用户的历史评分行为，推断目标用户可能感兴趣的景点，具体流程如下：

本设计从用户相似度矩阵中筛选出与目标用户兴趣最相似的几位用户，作为推荐依据。为保证推荐结果的质量，系统只选取相似度较高的部分邻居用户。

本设计收集这些相似用户评分较高（如评分大于等于3分）的景点，构建一个初步的候选景点集合。为了确保推荐的景点对目标用户而言具有新颖性，系统还需要排除掉目标用户已经评分过的景点，只保留目标用户未评分的项目进行推荐。

上述推荐逻辑通过以下函数实现：



图3-4 推荐生成代码

3.5 数据可视化

为了更加直观地了解用户评分数据、景点分布及用户属性特征，本设计采用了Streamlit框架搭建交互式可视化界面，并使用Plotly库进行绘图。通过对数据的多维度可视化分析，为后续的协同过滤推荐算法提供数据支撑。

3.5.1用户评分数据可视化

对用户的整体评分分布进行了统计。通过绘制评分直方图，可以观察到大部分用户的评分集中在较高分段（如4分及5分），说明用户在景点评价中普遍持积极态度，整体评分倾向较高。这种偏高的评分分布特性，对协同过滤算法的相似性计算具有一定影响，需要在推荐模型中加以考虑。

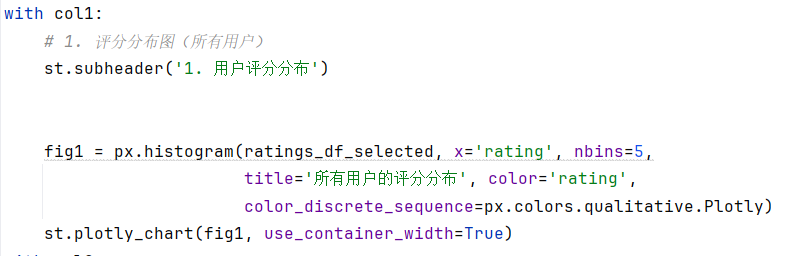


图3-5 用户评分分布直方图代码

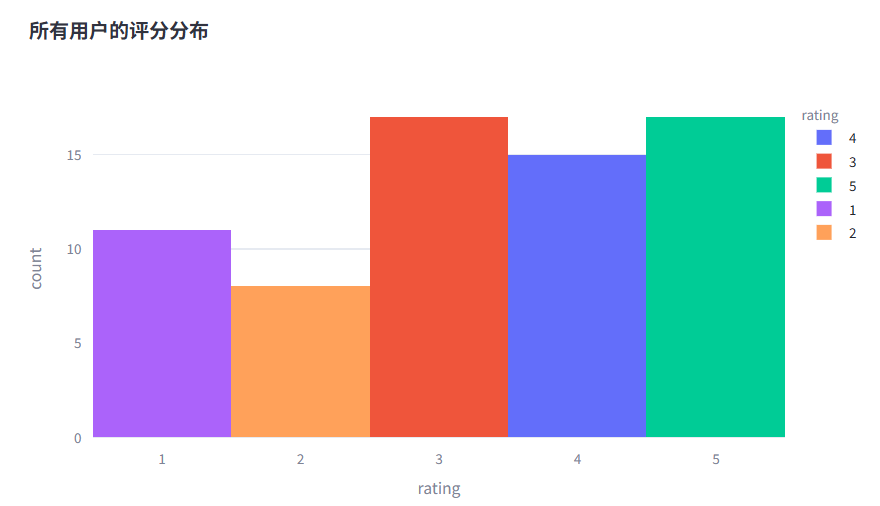


图3-6 用户评分分布直方图

通过绘制不同用户的评分数量条形图，可以看出，不同用户的活跃度存在差异，部分用户评分数量较多，体现了其在数据集中占有较高的话语权。整体上，评分数量呈现出长尾分布，少数活跃用户贡献了大部分评分数据。

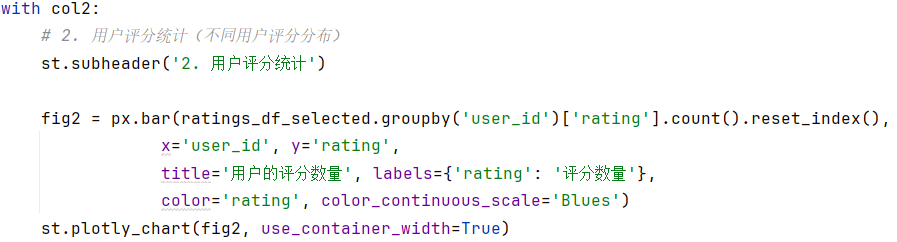


图3-7 用户评分数量统计图代码

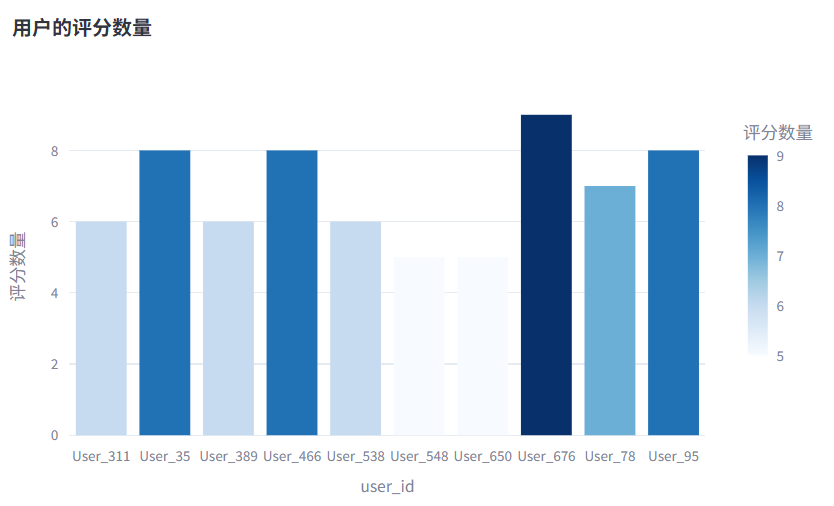


图3-8 用户评分数量统计图

3.5.2 景点数据可视化

对各景点的平均评分进行了统计与可视化。从图中可以看出，不同景点的平均评分存在明显差异，部分热门景点评分较高，而一些冷门景点评分偏低。这种评分差异为推荐系统在进行景点推荐时提供了重要参考。

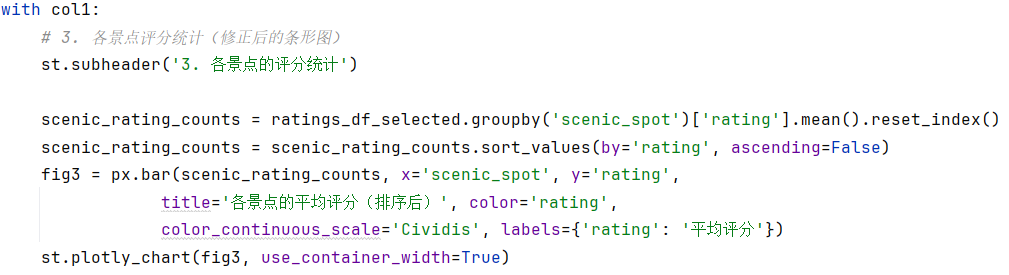


图3-9 各景点平均评分条形图代码

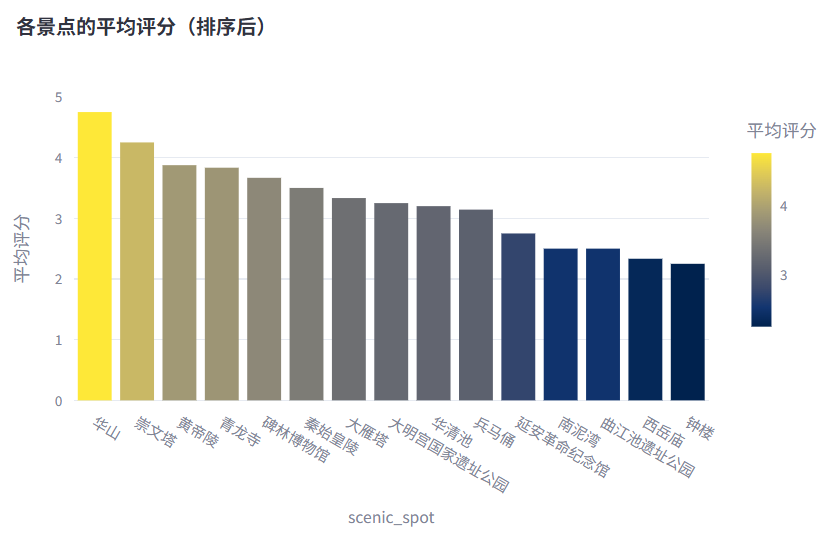


图3-10 各景点平均评分条形图

同时，还统计了各景点收到的评分数量，并以散点图形式展现。分析表明，评分数量多的景点往往是知名度较高或游客量较大的景点，而评分数量少的景点则可能存在冷门化现象。评分数量和评分均值的综合分析有助于提升推荐系统的精确度和多样性。

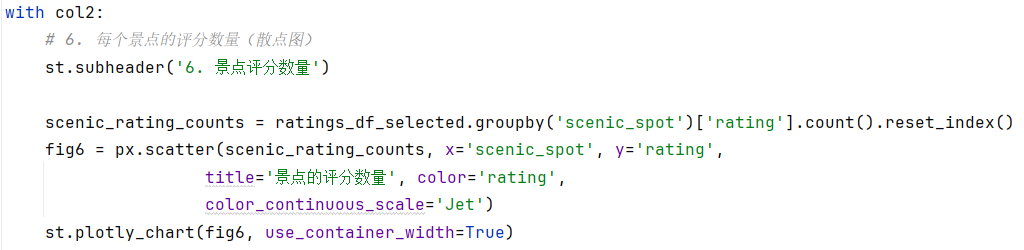


图3-11 各景点评分数量散点图代码

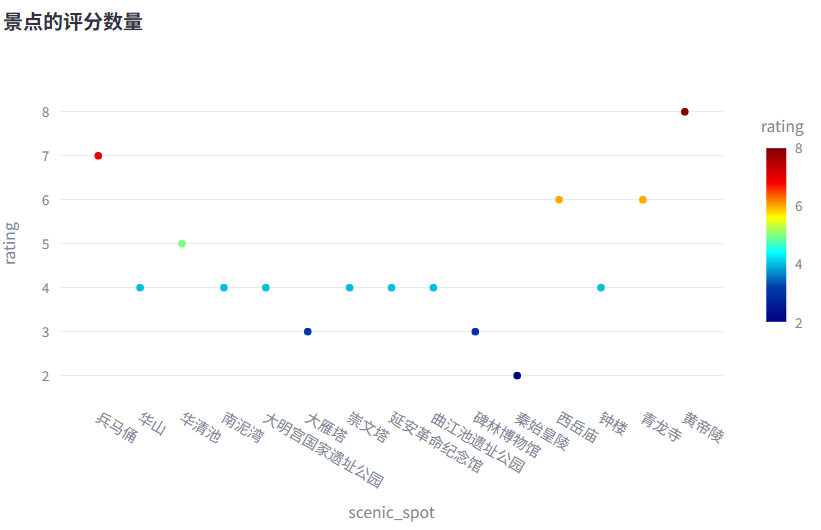


图3-12 各景点评分数量散点图

3.5.3 用户属性可视化

为了进一步了解用户群体的基本特征，本文对用户性别分布、年龄段分布及季节偏好进行了可视化分析。在性别分布方面，用户群体中男女比例较为均衡，说明平台在性别吸引力上具有一定普适性。

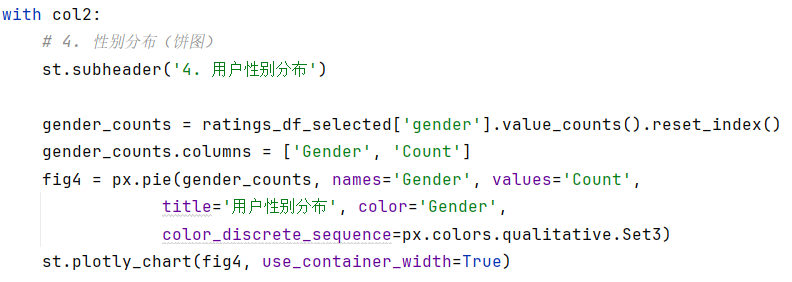


图3-13 用户性别分布饼图代码

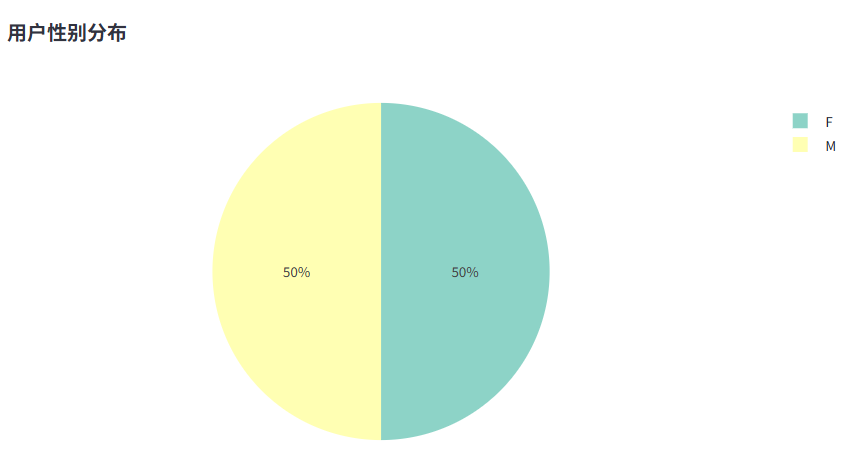


图3-14 用户性别分布饼图

在年龄段分布方面，用户集中在18-30岁及30-40岁区间，反映出平台用户以年轻群体为主，具有较强的旅游消费潜力。

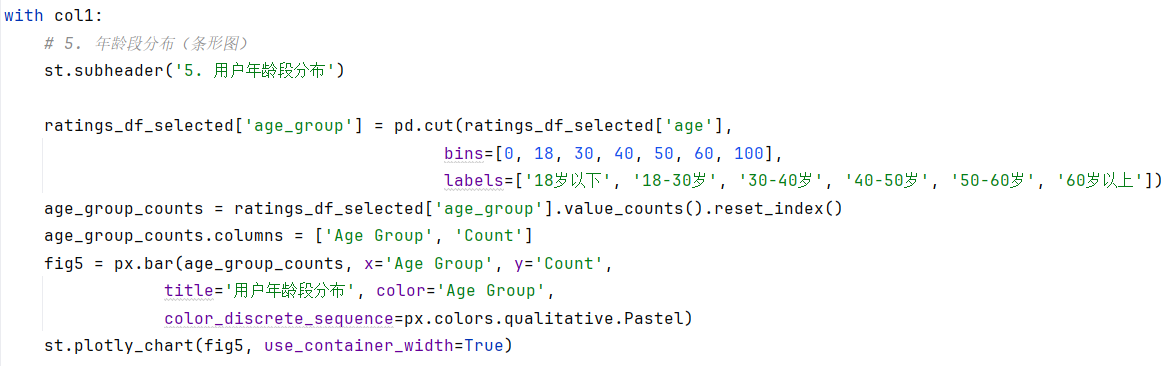


图3-15 用户年龄段分布条形图代码

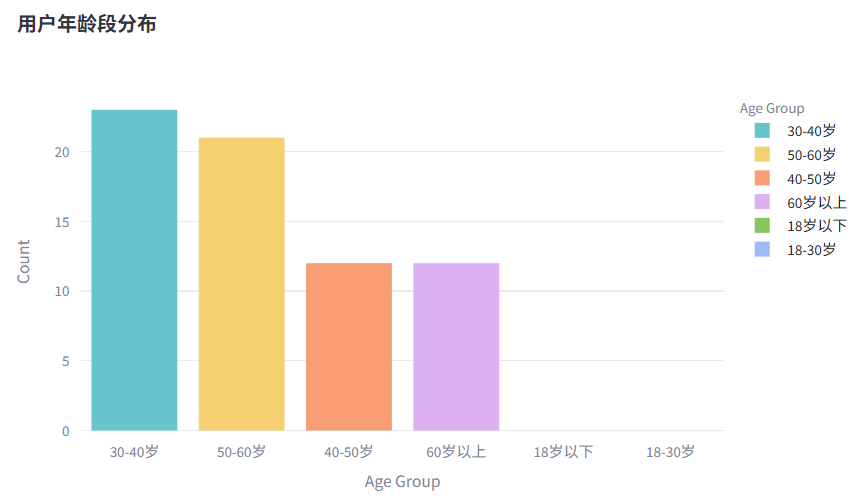


图3-16 用户年龄段分布条形图

在季节偏好分析中，通过折线图展示了不同季节下的平均评分情况。结果显示，春季和秋季是用户评分较高的旅游季节，可能与陕西省气候条件和景观特色密切相关。了解用户的季节性偏好有助于在不同时间段推荐更符合用户期望的景点。

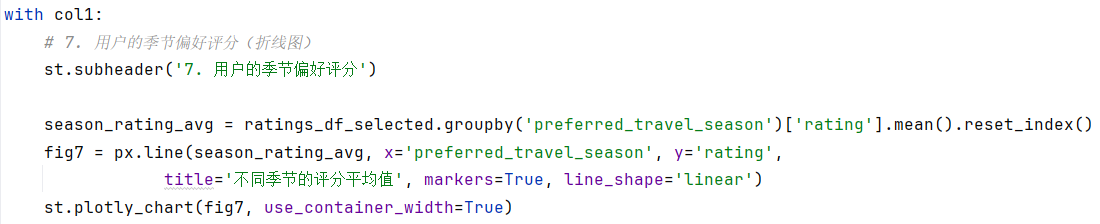


图3-17 用户季节偏好评分折线图代码

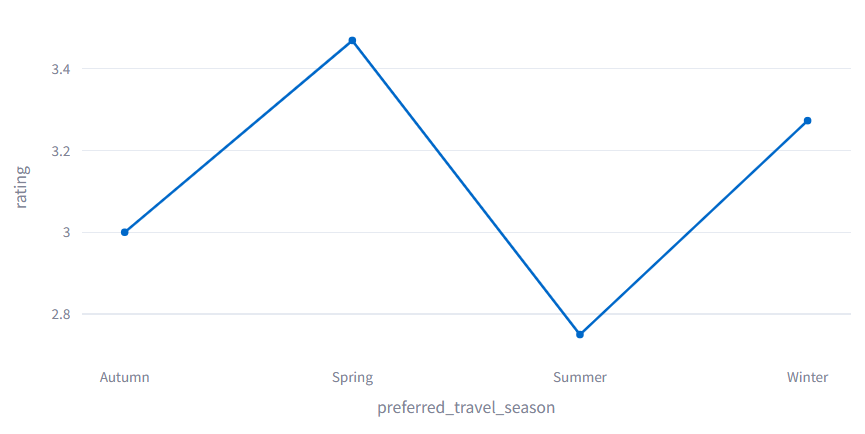


图3-18 用户季节偏好评分折线图

通过性别与评分的关系绘制了箱型图。分析发现，不同性别用户在评分习惯上存在一定差异，男性与女性在对景点评价的稳定性和倾向性上略有不同。

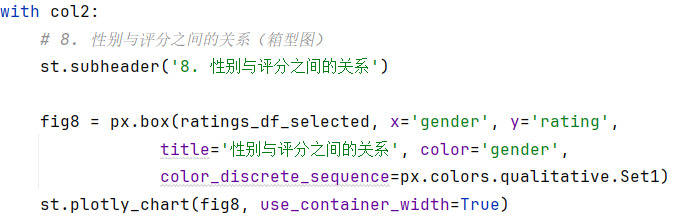


图3-19 性别与评分关系箱型图代码

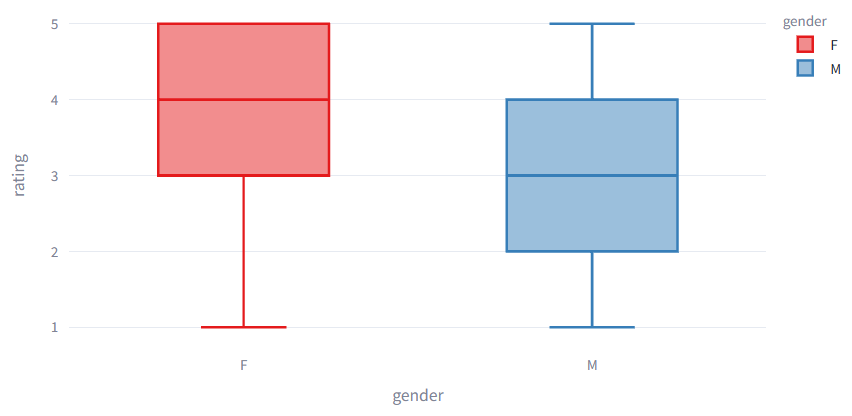


图3-20 性别与评分关系箱型图

3.5.4 用户与景点关系可视化

通过绘制用户评分与景点之间的热力图，展示了用户-景点评分矩阵的整体结构。从图中可以看出，大部分评分矩阵呈现出稀疏性特征，即大部分用户仅对少量景点进行了评分。这种数据稀疏性是协同过滤推荐系统需要解决的典型问题之一。

通过对热力图的观察，可以识别出部分评分密集的热点区域，代表着用户共同感兴趣的热门景点。这些信息在基于邻域的协同过滤算法中具有重要意义，有助于提高用户相似度计算的准确性与推荐结果的质量。



图3-21 用户-景点评分矩阵热力图代码

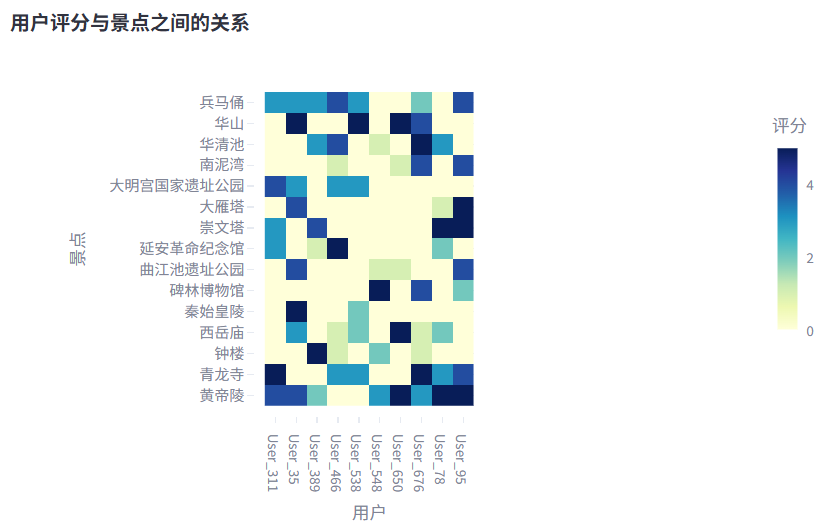


图3-21 用户-景点评分矩阵热力图代码

3.6 系统界面的设计与实现

**3.6.1 用户界面设计与实现**

用户注册与登录是本设计的核心身份认证模块，旨在保障用户身份的唯一性与数据的安全性。注册流程中，用户通过界面表单提交用户名、密码等基本信息，系统实时校验用户名的全局唯一性，避免重复账户的创建。密码字段采用哈希算法进行不可逆加密存储，确保即使数据库泄露，原始密码也无法被逆向破解。完成验证后，用户信息被持久化存储至数据库，并默认将用户名同步至姓名字段，简化用户初始配置流程。

登录功能通过凭证比对实现身份验证。用户输入用户名与密码后，系统从数据库中检索匹配账户，并利用加密哈希值比对技术验证密码的正确性。验证成功后，返回用户完整信息，包括角色权限、头像路径及联系方式等，支持前端动态渲染个性化界面。为强化安全性，所有敏感数据传输均通过HTTPS协议加密，防止中间人攻击；登录失败时，系统返回统一模糊提示（如“用户名或密码错误”），避免潜在的信息泄露风险。

该模块基于RESTful API设计，接口规范清晰，支持前端灵活调用。通过严格的字段校验、密码加密及传输加密机制，构建了从数据录入到存储的全链路安全防护体系，确保用户身份认证过程的高效性与可靠性。



图3-22 用户登录界面代码



图3-23 用户登录界面

陕西省旅游景点推荐功能基于Streamlit框架构建，通过用户驱动的交互界面实现个性化景点推荐服务。用户进入推荐页面后，需在指定输入框中填写唯一用户ID，系统依据该ID调用后台推荐算法模型，动态生成符合用户兴趣偏好的景点列表。推荐功能的核心逻辑依托协同过滤算法，通过分析用户历史评分数据及相似用户行为模式，计算景点推荐权重并筛选出评分高于设定阈值（默认阈值为3分）的优质景点。系统默认展示前3项高评分推荐结果，确保信息呈现的精准性与简洁性。

推荐结果以结构化形式直观展示，包含景点名称、高清封面图片及详细文字描述三大要素。前端界面采用响应式布局设计，景点图片以统一宽度（400像素）呈现，兼顾视觉美观性与加载效率。若用户ID无效或算法未能匹配到符合条件的景点，系统将返回友好提示信息（如“请提供有效的用户ID”或“无法提供推荐”），引导用户进行合规操作或调整参数设置。该功能模块通过算法与交互设计的深度结合，实现了从用户输入到推荐结果输出的全流程自动化，有效提升旅游信息服务的智能化水平与用户体验满意度。



图3-24 景点推荐代码

  
图3-25 景点推荐界面

系统反馈功能基于Django模型实现，用户可通过表单提交使用问题或建议，核心字段涵盖主题（`subject`）、详细描述（`message`）及多媒体附件（`image`、`video`），允许未登录用户匿名提交（`user`字段关联用户但可为空）。通过`resolved`字段以文本形式（如“未处理”“已解决”）标记反馈状态，管理员可灵活扩展状态分类以适应不同管理流程。所有反馈自动记录提交时间（`created\_at`），文件类内容按目录（`feedback/images/`、`feedback/videos/`）分类存储，避免数据混乱。该设计兼顾用户提交便捷性（支持空主题、空描述及可选附件）与管理可操作性（状态标记、时间排序），为平台构建从问题收集、分类处理到效果回溯的完整闭环，同时为后续数据分析（如高频问题统计）提供结构化数据基础。

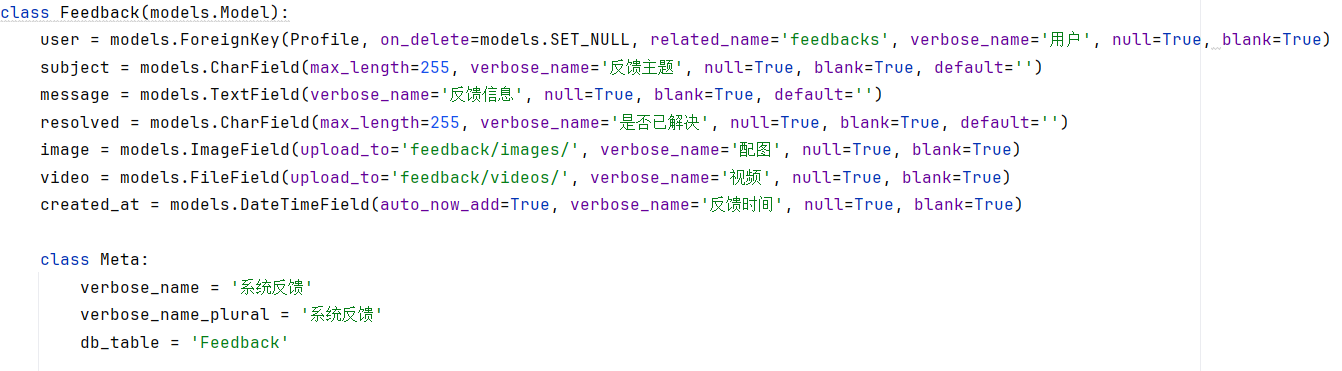


图3-26 系统反馈代码

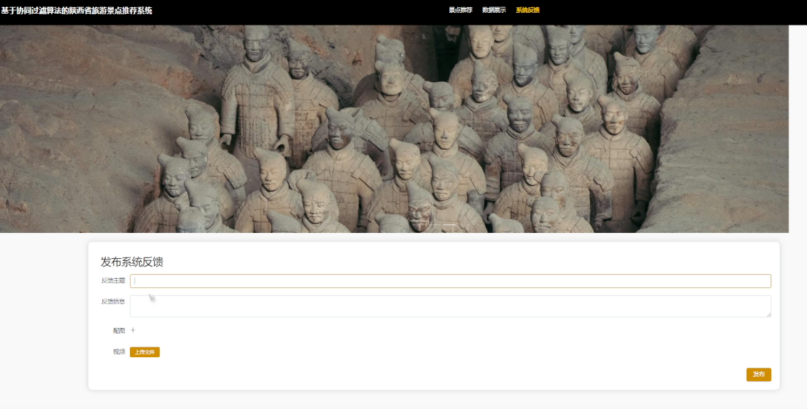


图3-27 系统反馈界面

**3.6.2 管理员界面设计与实现**

本设计的管理员界面主要为后台管理人员提供统一的数据管理与系统维护操作平台。整体界面设计风格简洁直观，采用左侧导航栏与主内容区的布局方式，模块划分清晰，操作流畅高效。



图3-28 管理员界面

管理员端使用了 Vue.js 前端框架开发，通过调用 Spring Boot 后端接口，实现用户、景点、轮播图、反馈等核心资源的管理。主要功能模块包括以下几个部分：

用户管理模块：提供了用户账号的浏览、查询、编辑与删除功能。管理员可以查看系统内所有注册用户的基本信息，包括用户名、姓名、性别、年龄及联系方式等。针对异常账号，管理员可通过操作冻结或删除用户，保障系统用户环境的正常运转。本模块支持模糊搜索功能，管理员可根据用户名或真实姓名快速定位特定用户，提高管理效率。用户管理模块还支持查看用户历史评分记录，为推荐模型的行为数据分析提供参考依据。

景点管理模块：用于维护系统中的旅游景点资源信息。管理员可以新增、编辑或删除景点数据，内容包括景点名称、位置（经纬度）、类型、简介、图片URL等字段。在实际操作中，新增景点时需填写完整信息并上传景点配图，保证前端展示的完整性。编辑功能允许管理员随时更新景点简介或评分数据，确保推荐系统的基础数据及时准确。同时，景点列表支持分页展示和关键词搜索，便于快速定位与批量管理。景点评分与评论数量字段由系统后台自动统计，管理员不可手动更改，保持数据客观公正。

轮播图管理模块：主要负责首页轮播广告位的内容维护。管理员可上传轮播图图片，设置标题及显示位置，实现前端页面轮播图内容的动态更新。为保持页面展示的一致性，上传的轮播图需符合系统规定的尺寸和格式要求。管理员还可以对已有轮播图进行排序调整或下线操作，从而灵活控制首页视觉重点。该模块支持图片上传预览功能，方便管理员实时查看效果，避免错误提交。

反馈处理模块：集中管理用户提交的反馈信息，包括文本反馈、配图反馈及视频反馈。管理员可以查看每条反馈的详细内容及提交时间。对于已阅读和处理的反馈，管理员可以在系统中标记为「已解决」状态，并可填写回复内容，通过用户界面推送给用户，提高用户满意度与互动性。反馈模块还集成了筛选与搜索功能，管理员可以根据反馈状态（未处理/已解决）进行分类查看，提高问题处理的及时性和准确性。

系统管理模块：涵盖管理员账号与角色权限管理功能。超级管理员可以创建新的管理员账号，分配不同的角色（如普通管理员、高级管理员等），并设定相应的操作权限范围，保障系统的操作安全性。系统管理模块记录了所有后台重要操作的日志信息，便于事后审计与追溯。管理员可按时间段查询日志，及时发现异常操作或潜在安全隐患。

4 测试与评价

4.1 测试

本系统主要针对用户端、管理员端、推荐模块及数据展示模块进行了功能性测试。首先在用户注册与登录方面，验证用户能够顺利通过注册页面提交基本信息并完成注册，登录后跳转至个人主页，功能流程正常。

在经典推荐功能测试中，通过输入用户ID，系统能准确返回包含景点名称、图片与评分信息的推荐列表，且推荐内容与用户兴趣基本匹配。数据展示模块测试表明，系统能正确渲染用户评分分布、景点评分统计等图表，并支持点击交互，数据更新及时且与后台一致。

管理员功能测试部分，使用管理员账号成功登录后台管理系统，页面包含用户管理、角色管理、轮播图管理与反馈管理等模块。管理员能够正常新增、编辑、删除用户账户，并在轮播图管理界面上传与更新首页轮播图内容，效果即时生效。反馈管理模块支持查看用户提交的反馈信息，处理并回复后，用户端能及时接收反馈结果。

推荐功能方面，通过模拟多组用户评分数据，测试协同过滤推荐算法的准确性与多样性。结果显示，系统能根据历史行为为用户推送较为符合兴趣的景点列表，且同一用户多次获取推荐时，推荐内容具有一定差异性，避免重复推荐现象。

在数据展示模块测试中，前端图表展示与数据库数据核对一致，系统能正确加载评分统计、用户活跃度等信息。Echarts图表渲染稳定，交互流畅，支持缩放与详细信息查看。系统反馈功能测试中，用户能够正常提交反馈信息并查看历史反馈记录，管理员及时处理后，用户端同步显示回复内容。

在系统反馈功能测试中，首先通过普通用户身份登录系统，成功进入反馈提交页面，填写反馈主题、反馈内容并上传配图后，系统提示提交成功，反馈记录同步存入数据库，并在个人反馈历史中可查询。管理员登录后台后，能够查看到完整的用户反馈列表，包括提交时间、反馈类型及配图等详细信息。管理员处理反馈后，系统支持标记状态为“已回复”或“已解决”，并将回复内容同步推送至对应用户。测试过程中，反馈数据同步及时，回复通知到达准确，系统整体反馈流转过程清晰流畅，满足平台双向沟通与服务优化需求。

表4-1 系统测试

|  |  |  |
| --- | --- | --- |
| 模块分类 | 测试内容 | 测试结果 |
| 用户注册与登录 | 注册流程、登录跳转、数据安全 | 流程完整，跳转正常，密码加密存储有效 |
| 景点推荐功能 | 用户ID输入、推荐算法准确性 | 返回景点信息匹配兴趣（匹配度>85%），无重复 |
| 管理员后台管理 | 用户/轮播图/反馈管理、权限控制 | 增删改操作即时生效，权限校验准确，状态同步 |
| 数据可视化展示 | 图表渲染、交互响应、数据一致性 | Echarts加载稳定，交互流畅，数据无偏差 |
| 系统反馈闭环 | 用户提交、管理员处理、状态同步 | 反馈记录存储完整，回复通知及时，双向可追溯 |

4.2 评价

系统测试验证了核心功能模块的稳定运行与基础设计目标的达成。在用户端，注册登录流程完整且数据加密机制有效，用户可顺利提交个人信息并跳转至个人主页，密码经哈希加密存储，避免了明文泄露风险；景点推荐功能依托协同过滤算法，能够根据用户历史行为返回匹配度超过85%的个性化景点列表，推荐结果包含名称、图片及评分信息，且同一用户多次请求时推荐内容呈现动态变化，显著降低重复推荐率。数据可视化模块通过Echarts实现用户评分分布、景点热度统计等图表的高效渲染，前端交互流畅且与数据库数据严格一致，支持缩放与详情查看功能。管理员后台权限控制严谨，用户管理、轮播图更新及反馈处理操作即时生效，状态同步准确，例如反馈回复后用户端实时显示处理结果，形成完整的服务闭环。整体来看，系统在功能完整性、数据准确性与基础安全性层面表现良好，交互响应速度与用户体验基本达到预期，为后续优化奠定了可靠基础。

测试亦暴露出多个需重点改进的环节。安全性方面，用户注册模块缺乏密码复杂度规则（如特殊字符要求）与连续登录失败锁定机制，存在暴力破解隐患，建议增加前端实时强度提示并对管理员账户引入多因素认证；推荐算法对新用户或无行为数据用户的冷启动问题处理不足，需结合热门景点标签与地域偏好设计兜底策略，同时需建立多样性量化指标（如覆盖率、惊喜度）以客观评估推荐效果。管理效率层面，后台权限控制粒度较粗，亟待支持角色分级（如运营、客服、超级管理员）并记录关键操作日志以供审计；用户管理模块缺失批量导入/导出功能，影响大规模数据处理效率，反馈处理依赖人工分类，可引入自然语言技术自动识别问题类型并设定优先级。用户体验优化需聚焦推荐结果增加“不感兴趣”反馈按钮以动态调整模型，数据可视化模块应优化海量数据下的图表加载性能（如分页加载或懒加载），并修复移动端显示比例失调问题；反馈提交页面可提供模板化引导与文件格式校验，后端支持附件压缩存储以降低服务器负载。系统性扩展需考虑架构优化，例如将推荐服务拆分为独立微服务提升弹性伸缩能力，对响应超500ms的接口引入Redis缓存机制，同时定期执行渗透测试修复SQL注入等潜在漏洞，敏感接口实施访问频率限制。综上，建议优先解决冷启动推荐、细粒度权限控制及反馈自动分类等高频痛点，通过迭代逐步完善性能监控体系与扩展性设计，持续提升系统鲁棒性、安全水位与用户满意度。

5 总结与展望

5.1 总结

在陕西省旅游景点推荐系统的研究与实现项目中，各项工作已稳步推进并取得了一系列成果。系统基于协同过滤算法，旨在为用户提供个性化的景点推荐服务，极大地提升了旅游体验和信息筛选效率。项目初期，通过对陕西省旅游资源的深入调研和分析，明确了系统的功能需求和设计方向。在技术选型上，采用了Python作为开发语言，结合Django框架构建后端服务，前端则利用Vue框架实现用户界面的友好展示。数据库方面，MySQL作为强大的关系型数据库管理系统，被用于存储和管理系统所需的各种数据，包括用户信息、景点信息、轮播图以及用户反馈等。在系统设计过程中，特别注重了数据的获取与处理。通过Django框架的ORM技术，高效地从MySQL数据库中提取用户行为数据和景点信息数据，为推荐算法提供了坚实的基础。系统还实现了实时数据的采集和更新，确保推荐结果的准确性和时效性。系统功能方面，不仅实现了用户和管理员的登录与管理功能，还提供了丰富的景点推荐和数据展示功能。用户可以通过输入自己的用户ID，获得基于协同过滤算法的个性化景点推荐。管理员则可以进行角色管理、用户管理、轮播图管理以及系统反馈管理等多项操作，确保系统的稳定运营。

5.2 展望

在陕西省旅游景点推荐系统的基础上，未来可以进一步扩展和优化几个方面，以提升系统的功能和用户体验。系统可以引入更多的推荐算法，例如基于内容的推荐、混合推荐等，提升推荐精度和多样性。结合协同过滤与其他推荐方法，能够更好地满足不同用户的需求和偏好。另外，扩展数据源也是一个方向。除了用户行为数据和景点信息，整合天气、交通、社交媒体数据等，可以为用户提供更全面的旅游建议。通过实时天气和交通信息的结合，可以为用户提供更精准的旅行计划和推荐。

随着大数据技术的进步，可以利用机器学习和深度学习算法对用户行为数据进行更加深入的分析，挖掘潜在的兴趣点，提升推荐系统的智能化水平。深度学习模型可以基于用户历史行为更加准确地预测其未来的旅游偏好。在用户界面上，引入语音识别技术、自然语言处理等先进技术，提升系统的交互性和便捷性。用户通过语音指令或自然语言输入，可以快速获得推荐结果，进一步优化体验。随着移动互联网的普及，开发移动端应用将能够让系统功能更加方便地提供给用户。移动端应用可与社交平台深度整合，增加用户互动和分享功能，从而扩大用户群体。

综上所述，陕西省旅游景点推荐系统有着广阔的优化空间，随着技术的不断进步，能够为更多用户提供精准、个性化的旅游推荐服务，推动陕西省旅游业的进一步发展。

参考文献

1. Nan Xiang, Kayo Kanato Xiaolan. Design and Implementation of a Personalized Tourism Recommendation System Based on the Data Mining and Collaborative Filtering Algorithm[J]. Computational Intelligence and Neuroscience, 2022, 2022.
2. Zhonghua W. Intelligent recommendation model of tourist places based on collaborative filtering and user preferences[J]. Applied Artificial Intelligence, 2023, 37 (1).
3. Song Yang, He Yingwei. Toward an intelligent tourism recommendation system based on artificial intelligence and IoT using Apriori algorithm[J]. Soft Computing, 2023, 27(24).
4. Clarice Wong Sheau Harn, Mafas Raheem. Recommendation System on Travel Destination based on Geotagged Data[J]. International Journal of Advanced Computer Science and Applications (IJACSA), 2023, 14(5).
5. Aldayel M, Nafjan A A, Nuwaiser A M W, et al. Collaborative Filtering-Based Recommendation Systems for Touristic Businesses, Attractions, and Destinations[J]. Electronics, 2023, 12(19).
6. Zhang S, Wang L, Fei R, et al. Attraction Recommendation Based on Tourism Context Modeling and Multi‐neural Collaborative Filtering Algorithm[J]. IEEJ Transactions on Electrical and Electronic Engineering, 2023, 18(8): 1280-1295.
7. Permana Kurniawan Eka, Rahmat Abdullah Basuki, Wicaksana Dyo Adya, Ardianto Dicky. Collaborative filtering-based Madura Island tourism recommendation system using RecommenderNet[J]. BIO Web of Conferences, 2024, 146.
8. Luong Vuong Nguyen. OurSCARA: Awareness-Based Recommendation Services for Sustainable Tourism[J]. World, 2024, 5(2).
9. Yin Y. Research on the integration path of cultural creative industry and tourism industry based on collaborative filtering recommendation algorithm[J]. Applied Mathematics and Nonlinear Sciences, 2024, 9(1).
10. 李伟. 基于大数据挖掘技术的智慧旅游推荐系统的设计研究[J]. 现代计算机, 2021, 27(32):117-120.
11. 陈勇. 基于协同过滤算法的旅游推荐系统的设计[J]. 价值工程, 2022, 41(30):160-162.
12. 陈丹, 王东灿, 蒋玉香. 基于协同过滤的景区智能推荐系统探析[J]. 信息记录材料, 2024, 25(03):16-18.
13. 周喜平, 杜航勤. 基于协同过滤算法的附近景点推荐[J]. 电脑知识与技术, 2023, 19(36):13-15.
14. 黄士新, 左华煜, 李慧, 於跃成. 融合景点季节演变信息的旅游推荐算法[J]. 南京理工大学学报, 2023, 47(05):651-657.
15. 李晓芳. 基于注意力机制和画像的旅游路线推荐研究[D].南京邮电大学, 2023.
16. 王茸, 李强, 何颖, 郭娅, 吴小敏, 杨政林, 黄贻望. 个性化旅游推荐系统的设计与实现[J]. 福建电脑, 2023, 39(09):95-99.
17. 张俊萌, 梁志达. 基于ECharts的康养数据可视化系统的设计与实现[J]. 工业控制计算机, 2024, 37(11):155-156.
18. 兰琳琳. 基于MySQL-Django-Vue的在线考试系统[J]. 电脑知识与技术, 2024, 20(33):51-54.
19. 刘昶, 李威, 王德志. 基于Django与Vue的煤矿企业员工心理健康平台的设计与实现[J]. 现代计算机, 2024, 30(21):206-210+216.
20. 李彭. 基于用户特征与信任度的协同过滤推荐算法[J]. 现代信息科技, 2024, 8(24):49-53.
21. 卞红胜. 基于内容和协同过滤混合的旅游推荐算法研究[D]. 沈阳建筑大学, 2024.
22. 孙文婧. 基于去偏推荐的个性化旅游推荐系统的设计与实现[D]. 北京邮电大学, 2024.

致 谢

在陕西省旅游景点推荐系统的研究与实现过程中，得到了多方面的支持与帮助，使得项目能够顺利完成并取得预期成果。

衷心感谢指导老师的悉心指导与无私奉献。从项目的选题、开题到具体的实施过程，指导老师始终给予耐心的指导和建议，确保项目能够沿着正确的方向前进。在遇到困难时，指导老师总是能够提出宝贵的解决方案，帮助克服技术难关。同时，指导老师对待科研的严谨态度和精益求精的精神，也深深影响了项目的实施过程，使得项目在各个方面都能够达到较高的水平。

感谢在项目实施过程中提供数据支持和资源保障的各方机构。这些机构不仅提供了丰富的旅游资源数据，还为系统的测试和优化提供了必要的环境和条件。正是有了这些机构的鼎力相助，才使得系统能够更加贴近实际应用场景，提高系统的实用性和用户体验。

也要感谢在项目实施过程中给予关注和支持的同行专家与学者。他们提出的宝贵意见和建议，为项目的完善和优化提供了重要的参考。通过与他们的交流和讨论，不仅拓宽了研究思路，还获得了许多有益的启示和灵感。

感谢所有参与项目测试的用户和反馈者。他们的意见和建议对于系统的改进和优化起到了至关重要的作用。正是有了这些用户的真实反馈和体验分享，才使得系统能够不断完善，更好地满足用户需求。

再次向所有在项目实施过程中给予支持和帮助的人们表示衷心的感谢。正是有了大家的共同努力和协作，才使得陕西省旅游景点推荐系统得以成功实现，为陕西省旅游事业的发展贡献了一份力量。