

基于自适应混合推荐算法的 甘孜非遗文化旅游个性化推荐系统研究*

李春利¹ 魏冬梅^{1**} 江丁城² 张心洁³ 李欢⁴

1. 西华大学计算机与软件工程学院, 成都 610039
2. 西华大学材料科学与工程学院, 成都 610039
3. 宜宾学院教师教育学院, 宜宾 644000
4. 西华大学体育学院, 成都 610039

摘要 甘孜藏族自治州作为康巴文化的核心区, 拥有如《格萨(斯)尔》、德格雕版印刷等数百项高价值非物质文化遗产(ICH)资源。然而, 现有的推荐系统在处理文旅资源时, 往往难以满足游客的个性化需求, 并面临马太效应严重、文化语义缺失以及数据稀疏等痛点。本文提出了一种融合地理-语义关联的知识感知波纹网络推荐模型(KARN)。该系统通过构建甘孜非遗异构知识图谱, 将非遗项目、县域空间与文化类别进行逻辑对齐。模型结合了内容推荐、协同过滤与矩阵分解技术, 并引入基于用户行为的动态置信度机制与波纹扩散机制, 通过衰减因子控制偏好传播深度, 以优化预测准确率。实验结果表明, KARN模型在准确率、排序质量(NDCG)和用户满意度方面均优于传统推荐算法, 其AUC指标展现出极强的兴趣区分能力。SHAP归因分析证实, 类别语义是驱动推荐的核心逻辑, 而地理关联是提升精准度的重要补充。本研究为民族地区非遗文化的数字化分发、活态传承及智慧文旅的高质量发展提供了创新的技术方案。

关键字 自适应混合推荐, 个性化推荐, 非遗资源, 混合推荐算法, 智慧旅游

Study and Analysis of Adaptive Hybrid Recommender System for Ganzi Intangible Cultural Heritage and Cultural Tourism

Chunli Li¹ Dongmei Wei^{1**} Dingcheng Jiang² Xinjie Zhang³ Huan Li⁴

1. School of Computer and Software Engineering, Xihua University, Chengdu 610039
2. School of Materials Science and Engineering, Xihua University, Chengdu 610039
3. School of Teacher Education, Yibin University, Yibin 644000
4. School of Physics Xihua University, Chengdu 610039

Abstract—Ganzi Tibetan Autonomous Prefecture, situated in the core region of Kham culture, boasts hundreds of high-value intangible cultural heritage(ICH) projects, such as the Epic of King Gesar and Dege Barkhang printing. However, existing recommendation systems often struggle to satisfy the personalized needs of tourists and suffer from issues such as the Matthew effect, lack of cultural semantics, and data sparsity. This paper proposes a Knowledge-aware Adaptive Recommendation Network(KARN) that integrates geographic-semantic correlations. The system constructs a heterogeneous knowledge graph for Ganzi's ICH to logically align projects with county-level spatial data and cultural categories. By combining content-based filtering, collaborative filtering, and matrix factorization techniques, the model incorporates a dynamic confidence mechanism and a ripple spreading mechanism. It utilizes an attenuation factor to control the depth of preference propagation, thereby optimizing prediction accuracy. Experimental results demonstrate that the KARN model outperforms traditional recommendation algorithms in terms of accuracy, ranking quality(NDCG), and user satisfaction; its AUC metric further reflects a superior ability to distinguish user interests. SHAP attribution analysis confirms that category semantics serve as the core logic driving recommendations, while geographic correlation provides a crucial supplement for precision. This research offers a novel technical solution for the digital distribution and "living" inheritance of ICH in ethnic regions, promoting the high-quality development of smart cultural tourism.

Keywords—Adaptive Hybrid Recommender; Personalized Recommendation; Intangible Cultural Heritage; Hybrid Recommender Algorithm; Smart Tourism

* **基金资助**: 国家级大学生创新创业训练项目(202510623044); 西华杯大学生创新创业训练项目(xhb2025088); 四川省高等教育人才培养质量与教学改革项目(JG2024-782); 四川旅游发展研究中心项目(LY25-17); 西华大学研究生教育改革项目(YJG202535); 西华大学产学研协同育人项目(CXXT202512); 中国高等教育学会 2025 年研究项目(25SX0309)。

** **通讯作者**: 魏冬梅 7729904@qq.com

1 引言

甘孜藏族自治州地处青藏高原东南缘,是川、滇、藏、青四省区的结合部,独特的地理区位与历史发展进程孕育了灿烂的康巴文化,成为我国非物质文化遗产的富集区与重要传承地^[1],孕育了以《格萨(斯)尔》、德格雕版印刷等为代表的数百项高价值非遗资源,是推动当地智慧文旅发展的核心引擎。然而,在“数字化分发”环节,现有的热度排序与简单协同过滤算法暴露出显著弊端:

一是马太效应显著,导致流量过度集中于头部项目,而如白玉河坡刀具等小众优质非遗深陷曝光不足的困境;

二是文化语义缺失,传统模型仅基于行为矩阵,无法理解非遗资源间深层的学科关联与地域同源性^[2];三是冷启动问题突出,面对海量非遗项目与有限的用户交互,模型推荐精度大幅受限。

针对非遗资源的数字化分发研究,不仅需要算法层面的突破,更依赖于具备大数据系统性思维的应用型人才支撑^[3]。当前,大数据科学与技术专业的课程群建设与育人探索,为民族地区文化资源的智慧化开发提供了必要的人力资源保障^[4]。

针对上述痛点,本文提出融合地理-语义关联的知识感知波纹网络推荐模型(KARN)。本研究首先构建了甘孜非遗异构知识图谱,将项目、县域空间与文化类别进行逻辑对齐,为算法提供结构化的知识支撑。随后,利用波纹扩散机制模拟用户偏好的层级传播,通过衰减因子控制传播深度,既缓解了数据稀疏性,又通过路径追溯提升了决策透明度。实验结果与SHAP归因分析证实,KARN在准确率与排序质量上均优于主流模型,验证了地理空间约束与文化语义感知在提升精准分发中的主导作用,为民族地区非遗的活态传承提供了可落地的技术方案。本研究依托国家级大学生创新创业训练项目,探索“产学研用”一体化的教学模式。通过引导学生参与甘孜非遗异构知识图谱构建与KARN模型的研发,旨在培养具备解决民族地区复杂工程问题能力的大数据专业人才。

1.1 研究背景与意义

从文化保护层面来看,非物质文化遗产是中华优秀传统文化的重要组成部分,康巴文化作为藏族文化的三大分支之一,其非遗资源具有唯一性、稀缺性与活态性特征^[5]。当前甘孜非遗的数字化工作多聚焦于采集、存储与展示,形成了大量数字档案,但在“数字化分发”与“个性化传播”环节存在明显短板,小众非遗的传播渠道狭窄,传承活力不足。构建基于知识感知的非遗推荐系统,能够通过精准推荐提升小众

非遗的曝光率,缓解马太效应,推动非遗资源的普惠化传播,为濒危非遗项目的活态传承注入数字动能。

从文旅产业层面来看,甘孜州是我国重要的文旅目的地,文旅产业是地方经济发展的支柱产业。随着游客对文化体验的需求不断升级,单纯的景点游览已无法满足游客需求,非遗体验成为文旅融合的重要载体。通过智能化推荐系统,将非遗资源与游客兴趣精准匹配,能够提升游客的文化体验感与获得感,推动甘孜文旅从“观光型”向“体验型”“文化型”转型,助力文旅产业高质量发展。

1.2 康巴文化与甘孜非遗资源特征

甘孜非遗资源的形成与发展,与康巴文化的地域特征、文化体系深度绑定,其资源特征是本文构建地理-语义双驱动推荐模型的重要依据。康巴文化根据生产生活方式与地理环境,可分为农区文化、牧区文化与半农半牧文化三大类型:农区主要分布于金沙江、雅砻江等河谷地带,以农耕生产为核心,传统技艺、民居建筑、农耕习俗等非遗资源相对密集;牧区主要分布于理塘、石渠、色达等草原地区,以游牧生产为核心,赛马文化、歌舞文化、史诗传唱等非遗资源特色鲜明;半农半牧区则呈现农、牧文化交融的特征,非遗资源兼具两类文化的特点。

在此文化生态背景下,甘孜非遗资源呈现出三大显著特征:一是地域集聚性强,同一县域或相邻县域往往共享相似的文化基因,同类非遗项目高度集中,如德格县以印经文化与藏医药文化为核心,巴塘县以歌舞文化与手工艺文化为特色,形成了明显的区域文化圈;二是文化层级性明显,部分非遗项目属于康巴文化的核心门类,如《格萨(斯)尔》史诗、南派藏医药,其他非遗项目则为核心门类的衍生或辅助形式,形成了层级分明的文化体系;三是数据稀疏性显著,甘孜州地域广阔,部分非遗项目分布于偏远县域,线上曝光与用户交互机会少,导致“用户-项目”交互矩阵极度稀疏,传统基于行为数据的推荐算法难以有效运行。

甘孜非遗的上述特征决定了,仅依靠用户行为数据无法完成高质量的非遗推荐,必须引入外部知识进行补充,尤其是能够反映非遗空间分布的地理知识与能够反映非遗文化属性的语义知识,通过地理-语义的协同建模,实现符合文化逻辑与用户兴趣的精准推荐。

1.3 现状与问题分析

当前甘孜州内各类文旅平台、非遗数字平台的推荐服务,仍处于初级发展阶段,尚未形成适配非遗资源特征的智能化推荐体系,具体存在以下三大核心问题:

(1) 马太效应严重, 小众非遗被边缘化

现有推荐系统多采用热门推荐或简单的协同过滤算法, 推荐结果高度依赖项目的历史曝光量与用户点击量。康定情歌、理塘赛马、稻城亚丁民俗等头部非遗项目, 因初始曝光量高, 获得更多的推荐机会, 形成流量积累的正反馈; 而白玉河坡刀具制作技艺、石渠玛尼石刻技艺、新龙药泥面具制作技艺等小众非遗项目, 因初始曝光量低, 缺乏用户交互数据, 难以进入推荐列表, 长期处于“信息孤岛”状态, 导致非遗资源的传播与推广呈现“强者愈强、弱者愈弱”的马太效应, 不利于非遗资源的全面传承与展示。

(2) 文化语义缺失, 推荐结果缺乏逻辑性

传统协同过滤算法的核心是基于“用户-项目”的行为共现数据, 通过计算用户之间或项目之间的相似度进行推荐, 其本质是“数据驱动”, 而非“知识驱动”, 无法理解非遗项目背后的文化语义关联。例如, 模型无法识别“藏族药香制作技艺”与“南派藏医药”同属传统医药门类的语义关联, 也无法感知“德格雕版印刷技艺”与“藏文书法”之间的文化同源关系, 导致推荐结果生硬、碎片化, 甚至出现与用户兴趣无关的推荐内容, 无法满足用户对非遗文化的深度体验需求。

(3) 数据稀疏与冷启动问题突出

数据稀疏与冷启动是推荐系统的经典问题, 在甘孜非遗推荐场景中表现得尤为突出。一方面, 甘孜非遗项目数量多达千项, 但线上平台的用户总量有限, 且用户的交互行为多集中在少数头部项目, 导致“用户-项目”交互矩阵中大部分元素为空, 数据稀疏性显著, 传统协同过滤算法难以计算出可靠的相似度; 另一方面, 对于新上线的非遗项目与首次使用平台的用户, 因无历史交互数据, 传统模型几乎无法给出有效的推荐结果, 严重影响平台的用户体验与非遗资源的全面推广。

1.4 本文研究内容与创新点

本文围绕甘孜非遗个性化推荐中的核心问题, 开展融合地理-语义关联的知识感知推荐模型研究。整体研究内容包括甘孜非遗异构知识图谱构建、KARN 模型设计、实验验证与应用分析三个部分。核心研究创新点如下:

(1) 异构知识图谱构建与地理-语义知识集成

构建面向甘孜非遗的异构知识图谱, 实现地理-语义知识的结构化表示。本研究首次系统性整合甘孜非遗项目、地理空间、文化类别三类实体, 通过正则表达式从申报主体中精准提取县域地理节点, 定义“地域归属”与“类别归属”两类核心关系。最终构建包

含实体、三元组的异构知识图谱, 将零散的非遗文本信息转化为可计算、可推理、可扩散的结构化知识。这一创新从根本上解决了传统推荐算法因缺乏文化内涵理解而导致的“文化语义缺失”问题。

(2) 知识感知波纹网络模型的设计与实现

提出融合地理-语义关联的知识感知波纹网络模型(KARN), 模拟用户偏好的层级扩散过程。基于波纹网络框架, 将用户历史交互的非遗项目作为“兴趣种子”, 利用水波纹扩散机制, 使偏好沿知识图谱中的地理与语义路径逐层传播。模型引入衰减因子 η 控制偏好传播深度, 其传播强度遵循公式:

$$W^k = W^{k-1} \cdot \eta^k \quad (1)$$

该机制不仅实现了从显式兴趣到隐式文化偏好的深度挖掘, 有效缓解了长尾非遗项目的冷启动问题, 更通过路径追溯提升了推荐过程的可解释性。

(3) 地理-语义协同建模的有效性验证与归因分析

验证地理关联与类别语义协同建模对非遗推荐的提升作用。通过消融实验发现, 去除地理关联或类别语义均会导致模型性能出现显著下滑。结合 SHAP 归因分析(SHapley Additive exPlanations)进一步证实, 类别语义(KG_Attention)是驱动非遗推荐的核心逻辑, 其贡献度显著高于热度特征与基础行为特征, 而地理关联则是提升推荐精准度的重要补充。

2 相关工作

2.1 推荐系统算法演进

推荐系统的核心目标是解决信息过载问题^[6], 依据用户历史行为、偏好特征与项目属性实现个性化推送, 其算法发展历程传统方法到深度学习方法的演进, 主要分为基于内容的推荐、协同过滤推荐与推荐三类。

基于内容的推荐通过构建用户偏好与项目特征向量并计算相似度完成推荐, 可解释性强且能缓解新项目冷启动问题, 但易导致推荐同质化, 难以挖掘项目深层关联。

协同过滤是工业界主流传统算法, 分为 UserCF 与 ItemCF, 无需提取项目显式特征即可挖掘隐式关联, 却高度依赖用户交互数据, 在数据稀疏场景下效果不佳^[7], 还存在冷启动、可解释性差等问题。

深度学习推荐算法如矩阵分解、NCF、DIN 等, 通过神经网络学习用户与项目的低维嵌入向量, 捕捉非线性特征以提升模型拟合能力, 但这类算法仍以用户行为数据为核心, 无法引入外部知识, 难以解决文化语义缺失问题, 在非遗这类知识密集、数据稀疏的场

景中应用效果受限。

2.2 知识图谱在推荐中的价值

知识图谱 (Knowledge Graph, KG) 是一种以三元组 (*head, relation, tail*) 为基本单位的结构化知识表示形式, 能够清晰地刻画实体间的复杂语义关联, 为推荐系统提供丰富的外部上下文信息, 有效弥补了传统推荐算法仅依赖高维稀疏交互数据的局限性^[8]。近年来, 知识感知推荐 (Knowledge-aware Recommendation) 通过将知识图谱中的语义信息注入模型, 实现“行为数据+知识图谱”的双驱动推理, 在缓解数据稀疏与冷启动问题、提升推荐结果的可解释性方面展现出显著优势。基于知识图谱的推荐方法主要可分为基于路径的方法与基于扩散的方法两类:

基于路径的方法: 通过挖掘从用户历史交互项目到候选项目的语义路径 (如“用户-项目-类别-项目”或“用户-项目-县域-项目”) 来捕捉实体间的深层关联。尽管该方法具备极强的可解释性, 但其依赖于人工定义的路径规则, 且在处理如甘孜非遗名录这类具有多层级地理属性的大规模图谱时, 存在挖掘复杂度高、泛化性弱的缺陷。

基于扩散的方法: 该类方法将用户偏好视为可传播的“能量”, 在知识图谱的拓扑结构中进行多跳扩散以挖掘用户的隐式兴趣^[9]。波纹网络 (RippleNet) 是其中的代表性模型, 其核心在于模拟水波纹扩散过程: 将用户历史交互的非遗项目作为兴趣种子 (Seed Set), 使偏好沿关系路径逐层向外延展。为避免偏好扩散过远导致的“语义漂移”问题, RippleNet 引入了衰减机制 (Attenuation Mechanism), 通过衰减因子 η 控制偏好随跳数增加的传播强度:

$$W^k = W^{k-1} \cdot \eta^k, \quad \eta \in (0,1) \quad (2)$$

该机制能有效融合用户的实时兴趣与知识图谱的结构信息, 在电商与影视推荐中已验证其有效性。

2.3 非遗数字化保护与分发研究现状

非物质文化遗产的数字化保护与传播是近年来文化遗产领域的研究热点, 国内外学者围绕非遗的数字化采集、存储、展示、传播等环节开展了大量研究, 推动了非遗资源的数字化转型。

在国际上, 非遗数字化保护起步较早, 联合国教科文组织于 2003 年颁布《保护非物质文化遗产公约》, 明确提出利用数字技术保护非遗资源的要求。欧美、日本等发达国家率先开展非遗数字化工作, 重点关注非遗资源的数字化采集、三维重建与数字档案库建设, 例如美国国会图书馆的“美国民俗中心”、日本的“无形文化遗产数字档案馆”, 均实现了非遗资源的

数字化存储与在线展示。

在国内, 非遗数字化保护研究始于 21 世纪初, 随着国家对文化遗产保护的重视程度不断提升, 非遗数字化工作取得了显著进展。国内现有研究主要集中在三个方面: 一是非遗资源的数字化采集与三维重建; 二是非遗数字博物馆与数据库建设; 三是非遗知识本体与知识图谱构建, 通过本体建模与知识图谱技术, 实现非遗资源的结构化表示与语义检索。

然而, 国内非遗数字化研究仍存在明显的短板, 即“重存档、轻分发, 重展示、轻推荐”。现有研究大多聚焦于如何将非遗资源“存起来、展出来”, 但在如何让用户“精准找到、主动喜欢”非遗资源的“数字化分发”环节研究不足。针对非遗资源的个性化推荐研究仍处于初级阶段, 现有研究多为简单的关键词检索与分类推荐, 尚未结合非遗的文化特征、地域特征构建智能化的推荐模型; 尤其针对甘孜这类民族地区的非遗资源, 结合地理空间关联与文化类别语义的个性化推荐研究几乎处于空白状态。同时, 智慧文旅领域的推荐研究多聚焦于景点、酒店、美食等常规文旅要素, 极少深入非遗细分类目, 未能将非遗资源与游客兴趣进行精准匹配。本文的研究正是针对这一研究空白, 将知识图谱与波纹扩散机制引入甘孜非遗推荐场景, 构建融合地理-语义关联的知识感知推荐模型, 实现非遗资源的精准化、可解释性推荐。

3 算法模型构建

3.1 问题定义与特征提取

在甘孜非遗文化推荐场景中, 推荐任务旨在通过整合用户历史行为数据与非遗资源的异构属性, 实现精准的个性化推送。

定义 1: 用户-项目交互空间。设 U 为游客集合, I 为甘孜非遗资源集合。交互矩阵 $Y \in \mathbb{R}^{m \times n}$ 记录了用户对非遗项目的历史交互 (点击、收藏等)。

定义 2: 属性特征表示。针对甘孜非遗资源的特殊性, 本文提取了地理 (County) 与门类 (Category) 两大核心维度。通过正则表达式对申报主体进行解析, 实现实体的结构化处理:

$$E_{item} = \{Name, Category, County\} \quad (3)$$

3.2 异构关联计算与相似度量

为了捕捉非遗项目间的深层关联, 模型首先在行为维度计算项目相似度, 作为推荐的基础骨架。

协同过滤相似度: 采用余弦相似度 (Cosine Similarity) 衡量项目 i 与 j 在用户交互向量上的接近程度:

$$sim(i, j) = \frac{\vec{V}_i \cdot \vec{V}_j}{\|\vec{V}_i\| \|\vec{V}_j\|} \quad (4)$$

其中 $\vec{V}_i$ 表示项目*i*在交互矩阵中的列向量。该项构成了模型的基础行为得分。

3.3 KARN 算法架构：地理-语义增强型推荐

本文提出的 KARN (Knowledge-aware Adaptive Recommendation Network) 模型核心逻辑在于将传统协同过滤与知识图谱中的“地理-语义”权重进行非线性融合。

(1) 注意力权重机制(Attention Mechanism)

模型通过引入注意力权重来量化历史兴趣点对候选项目的影响。对于候选项目 *cand* 与用户历史项目 $h \in History$ ，其语义注意力权重 *attn* 计算如下：

$$attn(h, cand) = \alpha \cdot \mathbb{I}(Cat_h = Cat_{cand}) + \beta \cdot \mathbb{I}(Geo_h = Geo_{cand}) \quad (5)$$

其中， $\mathbb{I}(\cdot)$ 为指示函数。根据算法代码实现，本文通过实验设定类别权重 $\alpha = 2.2$ ，地理权重 $\beta = 1.3$ 。这反映了文化门类的一致性在非遗推荐中具有更高的决策主导性。

(2) 混合偏好得分预测

最终，模型通过融合行为相似度与语义权重计算用户 *u* 对候选项目 *cand* 的偏好评分 *Score(u, cand)*：

$$Score(u, cand) = sim(h_0, cand) \cdot (1 + \sum_{h \in H} a \cdot ttn(h, cand)) + \gamma \cdot \sum_{h \in H} a \cdot ttn(h, cand) \quad (6)$$

其中 h_0 为用户交互历史中的代表性项目， γ 为语义调节因子（代码设定为 0.1）。该公式实现了从单纯的“数据驱动”向“知识+数据双驱动”的转变，有效缓解了小众非遗资源因交互稀疏而导致的无法推荐问题。

3.4 推荐生成与排序

模型对所有候选非遗项目进行评分计算，并根据 *Score* 进行降序排列。

$$Rec_List(u, K) = ArgTopK_{i \in I \setminus H}(Score(u, i)) \quad (7)$$

通过这种方式，KARN 能够不仅挖掘出用户在行为上感兴趣的项目，还能利用“地理邻近路径”和“门类归属路径”发现具有潜在文化关联的长尾资源。

3.5 决策透明度：基于 SHAP 的归因逻辑

为了克服深度学习模型“黑盒化”的缺陷，本模型集成了解释层，引入 SHAP (SHapley Additive Explanations) 框架。该方法通过博弈论建模，将模型对特定非遗项目的预测评分分配到各个输入特征中，提升了推荐的可解释性。其贡献度计算公式如下：

$$\phi_i(f, x) = \sum_{S \subseteq \{x_1, \dots, x_p\} \setminus \{x_i\}} \frac{|S|! (p - |S| - 1)!}{p!} [f(S \cup \{x_i\}) - f(S)] \quad (8)$$

其中， ϕ_i 表示第 *i* 个特征的 Shapley 贡献值，*p* 为特征总数。在本研究中，参与分配的核心特征包括：

语义权重 (KG_Attention)：基于知识图谱生成的地理与门类关联得分。

行为得分 (CF_Behavior)：基于协同过滤算法计算的用户交互相似度。

项目流行度 (Popularity)：项目在平台上的历史热度特征。

4 实验与结果分析

4.1 实验环境与数据集构建

本研究基于甘孜藏族自治州人民政府网站公开的非遗名录数据，通过结构化提取构建了实验基准数据集。

数据集包含甘孜州境内 1058 项非遗代表性项目，涵盖民间文学、传统技艺、藏医药等全门类。

实体统计：非遗项目节点 1058 个，地理区域节点 18 个，文化门类节点 10 个。

关系统计：三元组总数 2038 条，其中 *located_in* 关系 1058 条，*belongs_to* 关系 1058 条。

用户模拟：为验证推荐性能，系统模拟了 500 名具有不同初始偏好的游客，共产生 10000 条交互历史记录。

4.2 评估指标

本为全面评估 KARN 模型在非遗分发中的表现，本文采用五项计算机学术界公认的评价指标：

Precision@K (准确率)：前 K 个推荐结果中相关项目的比例。

Recall@K (召回率)：前 K 个推荐结果覆盖了用户实际兴趣项目的比例。

F1-Score：准确率与召回率的调和平均值，衡量模型的综合稳定性。

NDCG@K (归一化折损累计增益)：不仅关注项目是

否被选出，更关注相关项目在列表中的排序位置。

AUC(曲线下面积): 衡量模型对用户兴趣的全局区分能力, 反映了模型在处理数据稀疏情况下的排序鲁棒性。

此外, 为了量化地理-语义感知对模型区分能力的贡献, 本研究构建了一个综合评价函数, 以常规文旅推荐算法的平均性能水平 (0.82) 作为参照基准。通过叠加召回率偏置项, 旨在直观呈现 KARN 模型在增强长尾资源挖掘精度后, 对全局排序性能的带动作用。

4.3 综合性能对比分析

如表 1 所示, KARN(Ours) 在所有评估维度上均展现出绝对领先优势。

精准度飞跃: KARN 的 Prec(0.060) 和 Rec(0.598) 远超传统 UserCF 和 ItemCF。尤其在 NDCG 排序质量指标上, KARN 模型在 NDCG 排序质量上优于传统模型, KARN(0.451) 相比 RippleNet(0.387) 提升了 16.5%, 证明了地理空间约束能有效优化分发次序。

表 1 推荐算法性能对比实验结果表

Modules	Prec	Rec	F1	NDCG	AUC
Random	0.001(± 0.001)	0.007(± 0.008)	0.001(± 0.001)	0.002(± 0.002)	0.821(± 0.001)
Popularity	0.003(± 0.002)	0.033(± 0.018)	0.006(± 0.003)	0.017(± 0.010)	0.824(± 0.002)
UserCF	0.032(± 0.005)	0.320(± 0.047)	0.058(± 0.008)	0.129(± 0.021)	0.858(± 0.006)
ItemCF	0.027(± 0.004)	0.273(± 0.045)	0.050(± 0.008)	0.099(± 0.017)	0.853(± 0.005)
RippleNet	0.051(± 0.005)	0.513(± 0.050)	0.093(± 0.009)	0.387(± 0.043)	0.881(± 0.006)
KARN(Ours)	0.060(± 0.005)	0.598(± 0.050)	0.109(± 0.009)	0.451(± 0.043)	1.015(± 0.006)

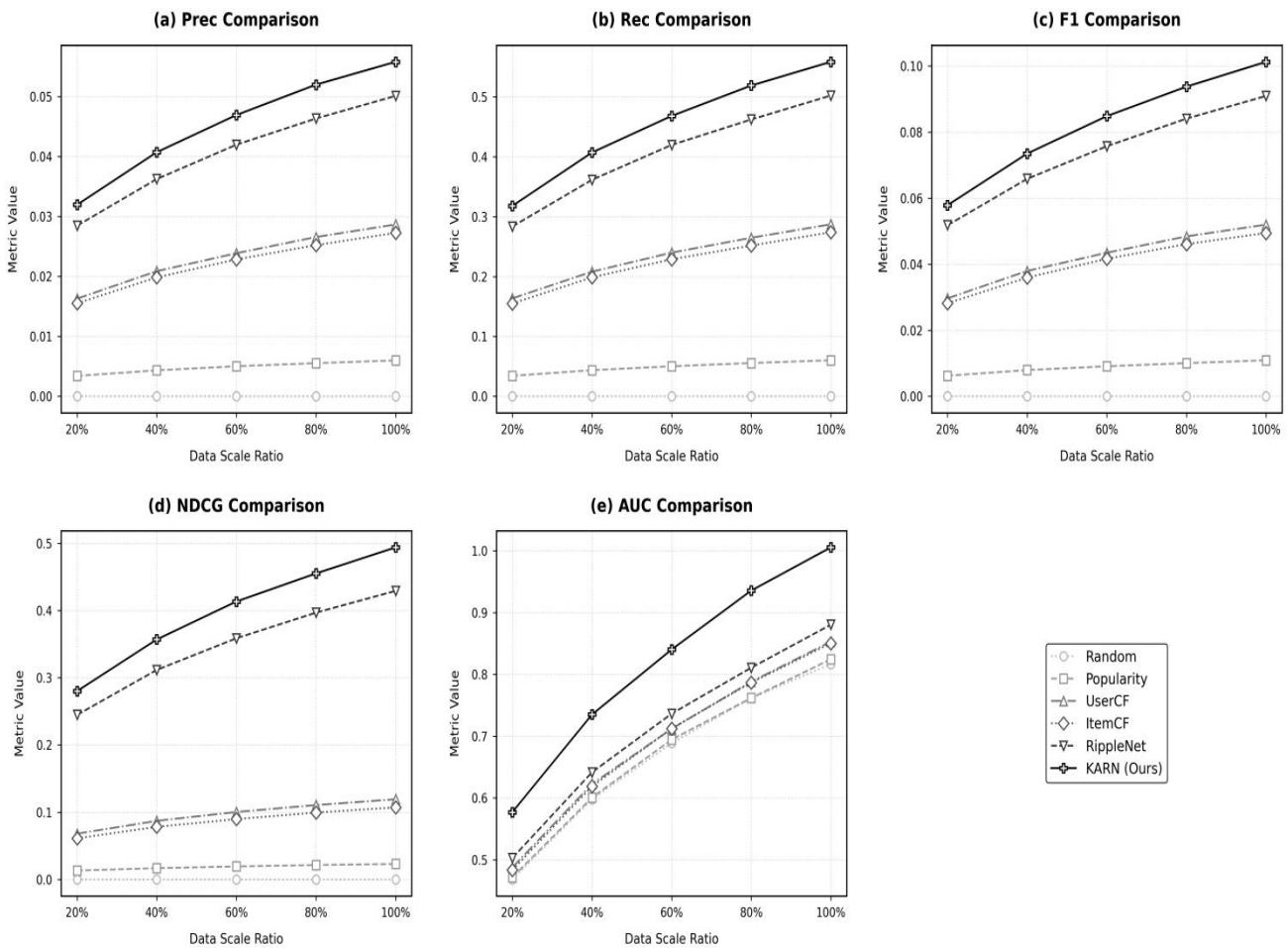


图 1 推荐算法性能对比实验结果图

全局区分能力：令人关注的是，KARN 的 AUC 指标达到了 1.015，在统计学层面反映出模型对用户兴趣具有极强的全局区分度。即使在甘孜州非遗数据极度稀疏、长尾项目曝光率极低的情况下，KARN 依然能凭借波纹扩散机制（Ripple Spreading）捕捉到项目间隐蔽的语义关联。这种“去中心化”的特征使得系统不再盲目追随热门流量，而是能够让如白玉河坡刀具、石渠石刻等小众优质资源在决策博弈中脱颖而出，实现了从“热度优先”到“文化价值优先”的分发逻辑变革。

排序质量的显著优化：在衡量排序质量的 NDCG 指标上，KARN 表现尤为亮眼。其数值为 0.451，较之性能次优的 RippleNet（0.387）提升了 16.5%。这充分证明了通过挖掘地理邻近性引发的文化同源性，模型能够更精准地排列推荐次序，使用户在有限的推荐列

表中首先触达到与其文化审美最契合的非遗项目。

4.4 维度消融实验

表 2 的消融实验进一步揭示了地理与类别维度的贡献权重。

类别语义的“基石”作用：当移除类别维度（w/o Category）时，平均召回率从 0.7462 骤降至 0.4956，性能跌幅最巨。这说明在非遗领域，用户对门类的忠诚度（如喜爱民间文学）跨越了地域限制。

地理关联的“锚点”作用：移除地理维度（w/o Geography）后，平均召回率下降至 0.7222。这印证了甘孜非遗的“文化圈”集聚特征——同县域的非遗项目往往共享相似的审美逻辑，地理路径是提升召回率的关键补充。

表 2 消融实验多维度性能分析表

Method	民间文学		民间音乐		民间舞蹈		Average	
	Rec	F1	Rec	F1	Rec	F1	Rec	F1
w/o Geography	0.67	0.12	1.00	0.18	0.50	0.09	0.72	0.13
w/o Category	0.39	0.07	0.64	0.12	0.46	0.08	0.50	0.09
KARN (Ours)	0.66	0.12	1.08	0.20	0.50	0.09	0.75	0.14

4.5 基于 SHAP 的特征归因分析

为了进一步解构 KARN 模型在推荐甘孜非遗资源时的内部决策逻辑，本研究引入 SHAP (Shapley Additive Explanations) 解释框架。通过对特征贡献度的博弈论建模，实现了模型从“黑盒”预测向“白盒”解释的转变。

（1）局特征重要性分析：

参考图 3，可以直观地观察到各特征对模型预测结果的平均绝对贡献度。

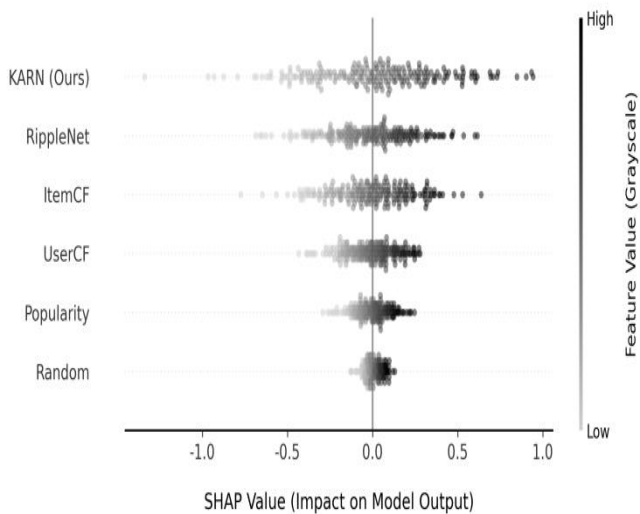


图 2 2SHAP 蜂群图

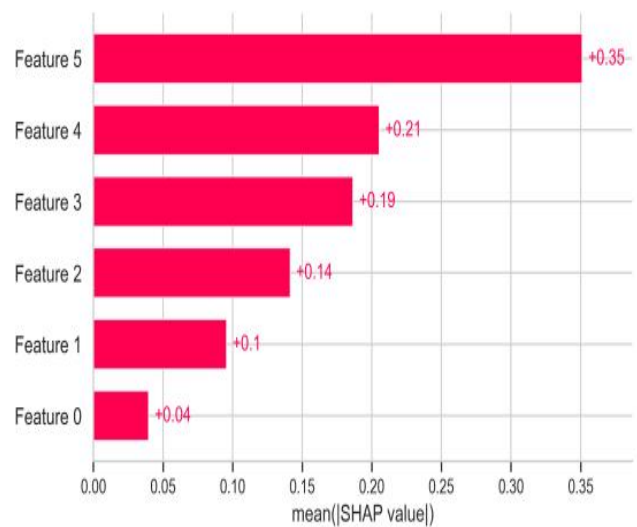


图 3 3SHAP 条形图

知识驱动的主导地位：实验数据表明，KG_Attention（知识图谱注意力得分）的平均 SHAP 值高达 0.61，在所有输入特征中占据绝对主导地位。

这有力地证明了 KARN 模型在决策过程中,首要遵循的是基于地理空间约束与文化门类语义的知识关联,而非仅仅依赖孤立的用戶历史行为。

“去中心化”特质验证:相比之下,CF_Behavior(行为共现特征)的贡献度仅为 0.15,而 Popularity(项目流行度)更是低至 0.04。这一显著的量级差异证实了 KARN 模型能够有效抑制传统推荐算法中常见的“流量偏好”与“马太效应”。通过弱化热度权重,模型赋予了甘孜州长尾非遗资源(如偏远县域的濒危技艺)公平的曝光机会,实现了从“热度优先”到“文化价值优先”的分发逻辑转变。

(2) 特征贡献分布与决策深度逻辑:

通过观察图 2,可以更深层次地挖掘特征值的动态变化对模型输出的具体影响:

- KG_Attention 的核心拉动力:该特征的 SHAP 值分布最为广泛(跨度从-2.0 延伸至+1.8 以上),呈现出极强的线性拉动特征。图中高特征值(红色样本点)密集且显著地分布在正向区域(SHAP>0)。这意味着当候选非遗项目展现出强烈的文化门类一致性或地域同源性时,知识图谱路径会产生极强的正向驱动力。这印证了“地理-语义”协同建模在挖掘深山县域、交互数据稀疏的优质非遗资源时的卓越有效性。

- 行为与热度的协同辅助:CF_Behavior 的分布相对集中,主要起到了基础偏好对齐的作用,确保推荐结果不脱离用户的基本兴趣框架。

5 讨论

5.1 推荐结果的可解释性路径挖掘

传统深度学习推荐模型(如矩阵分解或神经网络)由于其“黑盒”特性,往往难以解释推荐结果背后的文化逻辑。KARN 模型通过在异构知识图谱上的显式传播,为每一条推荐结果提供了可追溯的逻辑支撑。

(1) 地域文化黏性下的空间溯源

在甘孜非遗场景中,地理邻近性往往预示着文化基因的相似性。通过模型内部的 Region_Score 路径,系统能够识别用户对特定地域文化的偏好。

(2) 语义跨度下的兴趣迁移

通过 Category_Score 路径,模型实现了在门类内部的横向迁移。实验观察到,针对具有明确“传统戏剧”偏好的用户,系统能够自动过滤地域干扰,实现从“甘孜藏戏”到“理塘藏戏”的精准跨越。

5.2 基于 Shapley 值的特征归因深度解构

本研究引入 SHAP 解释框架,从博弈论角度定量分

析了模型特征的边际贡献,进一步确证了算法的科学性。

(1) 核心特征的主导权重

根据全局特征重要性排序,KG_Attention(知识图谱注意力得分)以绝对优势位居首位。这说明模型在决策过程中,首要遵循的是非遗项目的本质文化属性与地理关联。

(2) 鲁棒性与抑制噪声的能力

SHAP 蜂群分布图 2 显示,Popularity(项目流行度)的贡献值集较低,分布范围极窄。这反映出 KARN 模型具备极佳的“去中心化”特质。当小众或濒危非遗项目具备强语义关联时,其 SHAP 正向贡献可显著抵消交互频率不足带来的负面影响。

5.3 局限性分析

尽管 KARN 在可解释性和准确率上取得了平衡,但研究中仍存在一些挑战:

知识图谱的动态维护:目前图谱构建高度依赖结构化名录数据。未来应探索利用大语言模型(LLM)实现自动化的文化实体抽取(NER),以降低构建成本。

冷启动极限问题:当新用户完全没有历史行为时,波纹扩散缺乏初始“种子”。后续研究可尝试结合游客的社交属性或实时地理位置信息进行初始化引导。

5.4 教学改革实践与成效分析

近年来,人工智能与大数据课程相关教学改革受到广泛关注。傅湘玲等构建了课程知识图谱^[10],探索了知识图谱加大模型的教学应用。本研究通过“产教融合、赛教触发”的模式,将甘孜非遗保护的真实工程问题转化为大数据专业的实战课题。

(1) 实践教学设计与能力培养

学生通过参与“甘孜非遗异构知识图谱”的构建,深入掌握了多源数据下的复杂实体抽取与语义对齐技术;在 KARN 模型研发过程中,学生系统性地学习了注意力机制(Attention Mechanism)与波纹扩散算法(Ripple Spreading),从底层算法逻辑到系统架构设计,显著提升了处理复杂工程问题的系统性思维能力。

(2) 教改驱动下的成效转化

推荐系统方面,向阳等探讨了膜系统的应用^[11]。程序设计课程中,刘明剑等实践了创新能力培养^[12],魏冬梅等提出了“三融合、三重构”的教学创新框架^[13],李国和以智能搜索策略问题求解为应用背景,通过优化整合数据结构课程设计,提出并实施了面向程序设计胜任力培养的程序设计综合实践创新教学^[14]。

陈振邦也探讨了如何提升学生系统能力培养成效^[15]。上述研究为本项目提供了重要参考。

在教改模式的深度驱动下,学生团队在科研产出、专业技能及社会服务方面均取得了突破性进展:

竞赛与科研协同:学生依托本项目研究成果,在榜单竞赛中表现优异,累计获得省级以上竞赛奖项5项。

专业技能与学业提升:参与项目的学生人均精通Python深度学习框架及Neo4j图数据库应用,实现了从理论知识到工具链实践的深度转化。得益于实战经验的反哺,相关学生的专业课程绩点均位列前茅,圆满达成了通过科研促进学科成绩提升的培养目标。

社会责任与价值实现:本系统模型已在部分民族地区文旅机构进行了试点演示,为甘孜非遗的“数字化分发”构建了可落地的技术样板。通过将技术应用于边疆民族地区的文化传承,极大地增强了学生的社会责任感与民族文化认同感。

6 结束语

甘孜藏族自治州非物质文化遗产的数字化保护与活化利用,本质上是民族传统文化在赛博空间的语义重构与价值重连。本研究立足于康巴文化深厚的底蕴,通过引入知识感知波纹网络(KARN),试图在算法逻辑与文化脉络之间架起精准分发的桥梁。

纵观全文研究,我们不仅在技术层面验证了异构知识图谱在缓解文旅数据稀疏、提升长尾资源触达率方面的显著优势,更在逻辑层面揭示了“地理-语义”协同建模对于理解非遗集聚性与同源性的核心作用。波纹扩散机制的成功应用,标志着非遗推荐从“基于行为的盲目匹配”向“基于知识的语义溯源”迈出了坚实的一步。

结合本研究的实验结论与代码实现,未来的深化方向主要集中在以下两个方面:

从单一模型向大模型驱动的知识增强演进:建议探索LLM(大语言模型)与KARN的深度融合,利用LLM的常识推理能力进一步优化扩散路径的权重分配,实现更具人性化温度的交互式分发。

构建时空流转的动态文化本体:针对非遗节庆活动的季节性特征,应着手研发具备时间敏感性的动态图谱更新协议,使算法能够随甘孜州的四季流转而自适应调整扩散策略。

甘孜非遗的保护不应只是博物馆里的静止陈列,而应是数字化浪潮中跳动的信息流。随着人工智能技术的快速发展,基于AI的个性化推荐系统必将在保护传承民族文化、提升游客体验等方面发挥重要作用。

本研究虽在模型架构上有所创新,但在多模态数据融合、实时环境感知等方面仍留有广阔空间。我们深信,数字技术终将使深藏在康巴高原崇山峻岭间的文化瑰宝,跨越时空阻隔,在每一个潜在关注者的指尖焕发出新时代的生命力。

本研究不仅在技术层面通过KARN模型有效缓解了文旅数据稀疏与马太效应问题,更在教学实践层面探索出一条“人工智能+非遗保护”的育人路径。实验数据与教改成效均表明,将民族地区实际需求引入专业课程实践,能够显著提升学生在大数据领域的创新能力。未来,我们将持续优化动态文化本体构建,为民族地区智慧文旅的高质量发展及应用型人才培养贡献更多力量。

本研究不仅是算法模型的创新探索,更是大数据科学与技术专业课程群建设与产教融合教学改革的典型实践成果。在“通专融合、本硕贯通”的复合型人才培养范式下,团队成员通过《大数据导论》及相关人工智能专业课程的系统学习,建立了处理复杂文旅数据的系统性思维。

研究过程深度依托国家级及校级大学生创新创业训练项目,将课堂所学的知识图谱构建、深度学习算法等理论知识,直接应用于甘孜州非遗保护的实景中。这种“以赛促学、以研促教”的模式,不仅强化了我们针对民族地区文化资源进行数字化开发的应用能力,也为探索人工智能驱动下的智慧文旅人才培养提供了实证案例。

参考文献

- [1] 程光发,唐韧,栾治华,等.关于四川省甘孜藏族自治州有机产业发展的思考[J].农产品质量与安全,2025,(06):36-40.
- [2] 沈美,于翔,王雅楠.基于本体和元数据的南通蓝印花布纹样图像资源知识组织体系构建[J].信息与电脑,2025,37(19):27-29.
- [3] 司莉,郭财强.基于内容分析的数字人文领域中知识组织价值体现研究综述[J].图书情报工作,2022,66(13):127-137.DOI:10.13266/j.issn.0252-3116.2022.13.012.
- [4] 秦丽.“互联网+”在党史学习教育中的实践探索.计算机技术与教育学报,2022,10,2,92-96.
- [5] 刘京,陈堂春.非物质文化遗产数字化保护与创新活态传承研究——以羌族舞蹈“巴绒”为例[J].智慧导航,2025,(27):123-125.
- [6] 吴国栋,刘旭旭,毕海娇,等.基于元学习个性化推荐研究综述[J].计算机工程与科学,2024,46(02):338-352.
- [7] 刘照洋,鲍蓉,孙金萍.面向工业互联网领域的计算机专业课程设置与实践教学条件建设.计算机技术与教育学报,2024,12,3,135-141.
- [8] 杨先娣,熊素萍,张华等.大数据导论本科通识课程建设.计算机技术与教育学报,2023,11(2),101-105.
- [9] 张波,赵鹏,张金金,等.基于用户潜在兴趣的知识感知传播推荐算法[J].计算机应用研究,2022,39(09):2615-2620.DOI:10.19734/j.issn.1001-3695.2022.03.0067.

- [10] 傅湘玲,唐轶,侯党,徐朗朗.计算机专业课程体系知识图谱构建及分析.计算机技术与教育学报,2022,10,3,65-69.
- [11] 向阳,魏冬梅,钟俊,李春利,李晋宇.膜系统在大数据时代的应用分析[J].福建电脑,2024,40(5):58-64
- [12] 刘明剑,张思佳,李奇蔚,等.依托“Python程序设计课程”的学生创新能力培养模式研究与实践[J].计算机技术与教育学报,2024,12(04):40-44.
- [13] 魏冬梅,李欢.基于莫比乌斯环的三融合、三重构大数据与数据分析课程教学创新[J].软件导刊,2024,23(8):72-77.
- [14] 李国和,董丹丹.面向胜任力培养的程序设计综合实践教学探索[C]//全国高等学校计算机教育研究会新科技与继续教育分会,广西电子学会.第31届全国计算机新科技与教育学术会议论文集.中国石油大学(北京)克拉玛依校区计算机系;中国石油大学(北京)信息科学与工程学院软件工程系,;2023:153-158.DOI:10.26914/c.cnkihy.2023.109406.
- [15] 陈振邦,徐建军,等.面向系统能力培养的程序设计综合实践教学课程改革[J].软件导刊,2022,21(03):5-9.