



协同信号增强的大模型用户画像生成与推荐

郭世圆, 汪佳茵, 孙培杰, 张敏

引用本文:

郭世圆, 汪佳茵, 孙培杰, 等. 协同信号增强的大模型用户画像生成与推荐[J]. *智能系统学报*, 2026, 21(2): 487-497.

GUO Shiyuan, WANG Jiayin, SUN Peijie, et al. Collaborative signal enhanced LLM user profiling and recommendation[J]. *CAAI Transactions on Intelligent Systems*, 2026, 21(2): 487-497.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202506031>

您可能感兴趣的其他文章

融入学习者模型在线学习资源协同过滤推荐方法

A collaborative filtering recommendation method for online learning resources incorporating the learner model
智能系统学报. 2021, 16(6): 1117-1125 <https://dx.doi.org/10.11992/tis.202009005>

非结构化文档敏感数据识别与异常行为分析

Unstructured document sensitive data identification and abnormal behavior analysis
智能系统学报. 2021, 16(5): 932-939 <https://dx.doi.org/10.11992/tis.202104028>

基于知识图谱和用户长短期偏好的个性化景点推荐

Personalized attraction recommendation based on the knowledge graph and users' long-term and short-term preferences
智能系统学报. 2020, 15(5): 990-997 <https://dx.doi.org/10.11992/tis.201904064>

基于图游走的并行协同过滤推荐算法

Parallel collaborative filtering recommendation algorithm based on graph walk
智能系统学报. 2019, 14(4): 743-751 <https://dx.doi.org/10.11992/tis.201806002>

旅游知识图谱特征学习的景点推荐

Tourism knowledge-graph feature learning for attraction recommendations
智能系统学报. 2019, 14(3): 430-437 <https://dx.doi.org/10.11992/tis.201810032>

知识图谱的推荐系统综述

Review of recommendation systems based on knowledge graph
智能系统学报. 2019, 14(2): 207-216 <https://dx.doi.org/10.11992/tis.201805001>

DOI: 10.11992/tis.202506031

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20260203.1644.002>

协同信号增强的大模型用户画像生成与推荐

郭世圆, 汪佳茵, 孙培杰, 张敏

(清华大学 计算机科学与技术系, 北京 100084)

摘要: 用户画像的质量直接影响推荐系统的表现。传统的推荐系统通过建模用户与物品间的协同信息来获得用户画像, 无法充分利用用户与物品的文本描述信息。大模型处理文本信息、常识推理的能力及其拥有的世界知识, 为用户画像建模提供了新的机会。二者的结合可充分发挥彼此优点, 共同提高表现。本文提出了将用户的潜在兴趣与协同等级这两个来自推荐系统的协同信号引入大模型, 增强其生成用户画像的方法。用户画像生成通过与大模型进行多次交互的方式, 生成的用户画像进一步转换为特征向量, 通过对比学习与推荐系统内的用户表征相融合, 以增强个性化推荐表现。在两个数据集、多个推荐模型上的实验结果表明, 本文方法能够显著提升推荐模型的表现。本文方法弥合了大模型与推荐系统间的鸿沟, 为后续类似研究工作启发了新思路。**关键词:** 信息检索; 推荐系统; 大语言模型; 用户画像; 用户建模; 对比学习; 协同过滤; 用户表征; 特征向量
中图分类号: TP311.13; TP309 **文献标志码:** A **文章编号:** 1673-4785(2026)02-0487-11

中文引用格式: 郭世圆, 汪佳茵, 孙培杰, 等. 协同信号增强的大模型用户画像生成与推荐 [J]. 智能系统学报, 2026, 21(2): 487-497.

英文引用格式: GUO Shiyuan, WANG Jiayin, SUN Peijie, et al. Collaborative signal enhanced LLM user profiling and recommendation[J]. CAAI transactions on intelligent systems, 2026, 21(2): 487-497.

Collaborative signal enhanced LLM user profiling and recommendation

GUO Shiyuan, WANG Jiayin, SUN Peijie, ZHANG Min

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

Abstract: The quality of the user profile directly affects the performance of the recommender system. User profile in a traditional recommender system can be derived through modeling the collaborative information between users and items, but is unable to fully utilize the text description information of users and items. The textual information processing and commonsense reasoning capabilities of LLMs, combined with their world knowledge, provide new opportunities for user profiling. The combination of a recommender system and LLM can give full play to the advantages of each other, and improve each other's performance mutually. This paper proposes a method to introduce two collaborative signals, named potential interest and collaborative scale, from a recommender system into LLM to further enhance the user profile generation of LLM. The user profile is generated through multiple times of interaction with LLM, and is further transformed into an embedding, fusing with the user representation in the recommender system through contrastive learning to improve recommendation performance. Experimental results on two datasets and multiple recommender models show that the proposed method can significantly improve the performance of the recommender model. The proposed method bridges the gap between LLM and recommender system, and sheds light on further similar research work.
Keywords: information retrieval; recommender system; large language model; user profile; user modeling; contrastive learning; collaborative filtering; user representation; feature embedding

推荐系统根据用户历史行为中的显式反馈和隐式反馈, 推理出用户的偏好, 发现用户可能感兴趣的物品, 进行个性化推荐。它拓宽了人们获取信息的渠道, 改变了人们的生活方式, 一定程

度上缓解了信息过载问题^[1-5]。用户画像作为推荐系统对于用户偏好的推理建模结果, 其质量对推荐系统的表现至关重要。

当前的推荐系统通常从用户的历史行为序列中挖掘其与各个物品间的协同信息, 来对用户进行建模^[6-7]。其中, 用户与物品均由特征向量来表

收稿日期: 2025-06-25. 网络出版日期: 2026-02-04.

通信作者: 张敏. E-mail: z-m@tsinghua.edu.cn.

征,蕴含着用户与物品间的协同信息。然而,传统方法对于用户的建模是不深入的——用户除了非文本类型的隐式反馈外,还可能包含若干文本信息,例如性别、职业、评论文本等;每个物品也包含标题、类别、描述信息等文本信息。这些信息均包含着用户或物品的重要特征,但传统方法无法很好地将其利用。

而将大模型运用于推荐系统的用户建模之中,便可弥补上述不足之处。大模型的总结摘要能力,能够综合文本与非文本类型的用户反馈及信息,总结、提炼用户的偏好;大模型的常识推理能力,能够在已有的信息之外引入外部的世界知识,丰富用户与物品的表示。因此,大模型运用于用户建模具有很大的潜在价值。并且,大模型的上述优势是传统推荐系统所不具备的,传统推荐系统中的协同信息,也是大模型所难以掌握的。若将推荐模型的协同信息以合适的方式引入大模型中,协同信号增强的大模型能够更好地理解用户与物品,从而生成更好的用户画像;而来自大模型的知识进一步引入到推荐模型中,增强的推荐模型也能对用户与物品均有更好的理解与认识,从而相比于之前的模型有更好的表现和可解释性。

因此,本文尝试将大模型与传统推荐系统结合起来,共同参与用户画像的构建,能够发挥二者彼此的长处,实现对于用户更加深入准确的建模。然而,文本模态的大模型与特征向量模态的推荐系统之间存在着天然鸿沟,并且大模型的预训练过程使用的语料是自然语言语料,而非推荐领域的语料与数据。因此,需要在二者之间建立弥合这一天然鸿沟的方法,将彼此的信息引入对方之中,以发挥二者彼此的长处,实现强强联合。

本文主要关注来自推荐系统的协同信号对大模型生成用户画像的增强,进一步对推荐系统表现的提升。其主要贡献包括 3 个方面:

1) 提出了将来自推荐系统的非文本化的协同信号文本化,引入大模型以辅助增强用户画像生成的方法。该方法考虑了两种协同信号的引入:用户的潜在兴趣、用户的协同等级。协同信号与用户当前的画像、物品的文本信息、用户对物品的评分等共同作为大模型的输入,生成更进一步的用户画像。

2) 为使用户的交互历史得到充分运用、使得用户画像更加深入完善,具有良好的可读性,本文设计与大模型进行多次交互,逐步生成用户画像的框架,并在用户画像中的关键词过多时

其进行进一步提炼与概括,以促进对其的语义理解,提高用户画像的深入性与简洁性。生成的用户画像进一步转换为特征向量,通过对比学习的方法,将用户在推荐系统中的表征与之进行对齐,从而将经由大模型得到的外部知识引入推荐模型之中。

3) 样例分析表明,两种协同信号对用户画像的生成均有增强效果。进一步的实验结果表明,将大模型生成的用户画像引入推荐系统之中,能够增强个性化推荐的表现,且协同信号增强的用户画像对推荐系统的表现有进一步提升作用,从而验证了本文方法能够联结大模型与推荐系统,使他们相得益彰,取得双赢。

1 相关工作

1.1 传统用户画像建模方法

传统上,在大模型出现之前,用户画像的建模方法^[8]一般可分为:

1) 基于规则的方法——通过创建一组预先定义的规则或条件来进行用户建模,例如专家系统^[9]、决策树^[10]、人工定义的判别规则等;

2) 基于邻居的方法——通过分析当前用户和与之相似的一组用户的行为或特征的关联,来对当前用户画像进行建模,例如协同过滤^[11-12]、K-最近邻方法^[13]等;

3) 基于统计学习的方法——通过统计学习来对用户数据中的模式进行分析建模,例如贝叶斯分层模型^[14-15]、隐半马尔可夫模型^[16]、条件随机场^[17]、支持向量机^[18]、随机森林方法^[19]等;

4) 基于深度学习的方法——通过神经网络来捕捉用户数据中的复杂模式与特征,从而建模用户画像,例如循环神经网络^[20-22]、长短期记忆神经网络^[23-25]、注意力机制^[26-27]、图神经网络^[28-30]等。

然而,无论采用上述何种方法挖掘用户的特征与偏好、对用户进行建模,得到的用户特征均由特征向量表示,无法很好地利用用户与物品的文本信息。

1.2 使用大模型的用户画像建模方法

目前已有若干工作聚焦于使用大模型来进行用户画像建模。有的研究工作直接将用户全部的交互历史信息一次性输入大模型^[31-37],获得大模型总结、归纳出的用户偏好;与之相对应的,有的工作进行了随机采样^[38-39],将部分交互历史进行输入。有的工作使用了其他的技术与方法来完成用户画像建模,例如使用大模型模拟卷积过程^[40]、使

用分治法对用户交互历史进行分组概括与合并^[41]、将用户交互历史分批输入大模型进行多次更新^[42]等。

然而,上述方法均存在着不足。首先,这些方法均未与传统推荐系统相结合,未能利用传统推荐模型的优势与能力来进行用户画像生成,从而可能导致次优的结果。其次,现有方法的用户画像生成过程也存在着不足。当用户全部的交互历史被一次性输入时,其中每个物品的特征无法得到深入地理解,导致用户建模结果较为粗糙;若将用户交互历史进行采样,可能会遗漏一些出现频率较小,但对用户很重要的关键特征,同样不够准确和细致。最后,已有的将用户交互历史分批输入的方法未在生成过程中对用户画像进行进一步提炼与概括,不能保证生成的用户画像的深入性与可读性。相比之下,本文方法能够将用户全部交互历史纳入考虑,生成深入且简洁的用户画像,能够弥补已有工作的不足之处。

1.3 协同信号增强大模型的方法

目前,已有研究工作大多将用户与物品在推荐系统中的表征映射到大模型的词元表征空间中,融合在提示词中引入大模型,然后通过 LoRA (low-rank adaptation) 等大模型微调技术针

对特定下游任务对大模型进行微调^[43-46]。然而,上述工作主要完成的是有明确标签的监督学习工作,例如评分预测、下一个物品预测等,对于开放式生成任务,例如生成用户画像的效果尚未被探索。并且,对大模型进行微调的时间与资源开销较大,对方法在实际场景中得到大规模应用造成影响。相比之下,本文提出的两种朴素但有效的协同信号引入方式具有普适性与高效性,具有良好的应用前景。

2 协同信号增强的大模型用户画像生成与推荐方法

2.1 方法总体框架

本文方法的整体框架如图 1 所示。其中,传统推荐系统以用户交互历史作为输入进行训练,训练完成后的推荐系统参数中包含着用户与物品的表征,蕴含着用户与物品间的协同信号。获得的协同信号转换为文本的形式,与用户交互历史一同作为大模型的输入,进行协同信号增强的用户画像生成。在生成过程中,大模型根据给定的输入进行归纳总结,同时也会结合其拥有的世界知识。生成的用户画像转换为特征向量,引入推荐系统之中,以增强个性化推荐。

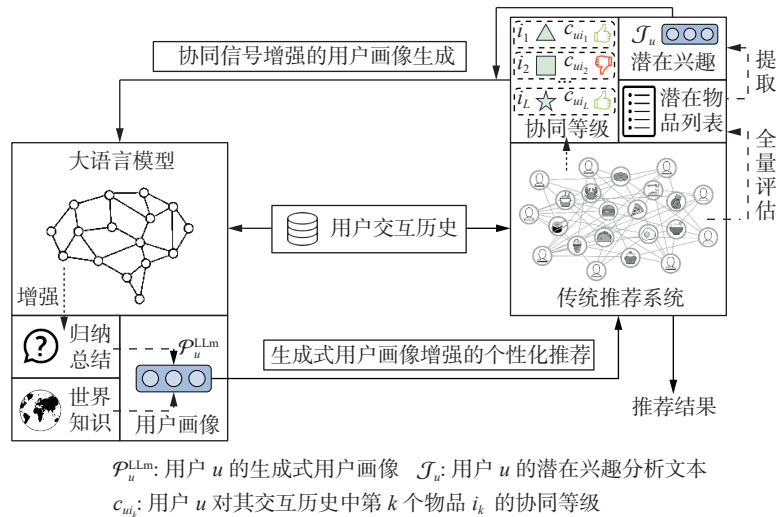


图 1 方法整体框架

Fig. 1 Overall framework of the method

本文考虑了两种协同信号,分别是用户的潜在兴趣和用户的协同等级。下面分别对其引入大模型并增强用户画像生成进行介绍。

2.2 引入用户的潜在兴趣

用户的潜在兴趣,即用户未交互过的物品中蕴含的用户偏好。除用户交互历史外,将用户潜在兴趣中的特征囊括进来,能够获得用户更加深

入的偏好,构建出更加完善的用户画像。这也与现实的推荐系统场景相契合——如果推荐系统仅利用用户的交互历史进行推荐,而不探索用户交互历史之外的潜在兴趣,则可能会导致推荐的物品广度越来越窄,造成用户的审美疲劳,导致信息茧房问题。而引入用户的潜在兴趣能够扩大用户的认知边界,在“利用”与“探索”之间达成平衡。

为引入用户的潜在兴趣,首先需要获取潜在物品列表。对于每一个推荐模型,当其训练完成后,本文对其进行全量评估(full evaluation),即在数据集全集或其较大规模的子集上进行评估,计算每个用户针对除训练集、验证集外其他所有物品(或较大规模的负采样的物品)的表征相似性。所有推荐模型计算得到的相似性取平均并排序,最终取排序列表中前若干个物品,它们可能是用户未交互过的,但协同分数较高的物品,可能是用户潜在感兴趣的,因此将其称为潜在物品列表。对于用户 u 而言,其潜在物品列表可以表示为

$$\mathcal{L}_u = [\mathcal{P}_{u_{i_1}}, \mathcal{P}_{u_{i_2}}, \dots, \mathcal{P}_{u_{i_T}}]$$

式中: $\mathcal{P}_{u_{i_j}}$ 为用户 u 潜在物品列表中第 j 个物品的画像,包括其标题、类别和描述文本等, $j = 1, 2, \dots, T$; T 为潜在物品列表中的物品个数。

获得用户的潜在物品列表后,本文将其中物品的画像输入大模型,总结、概括出用户的潜在兴趣文本。当大模型生成最终的用户画像后,本文又设计了一步推理,将潜在兴趣文本引入大模型,促使大模型结合最终用户画像和潜在兴趣文本,总结出包含用户潜在兴趣的,且保留了原有用户特征的新的用户画像。该过程可以被表示为

$$\mathcal{J}_u = \text{LLM}_1(\mathcal{L}_u)$$

$$\mathcal{P}_u^{\text{LLM+PI}} = \text{LLM}_1(\mathcal{P}_u^{\text{LLM}}, \mathcal{J}_u)$$

式中: \mathcal{J}_u 为用户 u 的潜在兴趣文本, LLM_1 为输入输出均为文本的大语言模型, $\mathcal{P}_u^{\text{LLM}}$ 为大模型生成的最终用户画像, $\mathcal{P}_u^{\text{LLM+PI}}$ 为包含用户潜在兴趣的用户画像。

2.3 引入用户的协同等级

除了提取出的用户潜在兴趣,用户与物品在推荐系统中的表征的相似性也是一个重要的协同信号,它反映着用户与物品的匹配程度,是推荐系统产生推荐结果的直接依据。本文将其称为用户与物品间的协同分数。相比于用户对物品的主观评分,协同分数是来自推荐系统的客观评价,将其同时引入大模型辅助用户画像生成,直觉上也能够促进获得更加准确的用户偏好。

对于训练完成的推荐系统,协同分数保存于用户-物品评分矩阵之中,分数大于 0 代表用户对该物品持正面态度,小于 0 则代表持负面态度。本文获取评分矩阵,以 0 为中心,在负数与正数上分别进行 Min-Max 归一化,分别将其值映射到 $[-1, 0]$ 和 $[0, 1]$ 范围内。形式上,对于一组 j 个实数 $\mathbf{x} = [x_1, x_2, \dots, x_j]$ 和其中的一个实数 $x_i, i \in \{1, 2, \dots, j\}$, 其使用 Min-Max 归一化到区间 $[a, b]$ 后的值为

$$x_i' = a + \frac{(x_i - \min(\mathbf{x}))(b - a)}{\max(\mathbf{x}) - \min(\mathbf{x})}$$

考虑到大语言模型对数字大小的比较能力不如对文本的理解能力,本文定义了由文本短语描述的若干协同等级,将协同分数进行离散化,映射到不同的等级后再输入大模型。如图 1 所示,对于用户 u , 其交互历史中的物品 i_1, i_2, \dots, i_L 对应的协同等级分别表示为 $c_{u_{i_1}}, c_{u_{i_2}}, \dots, c_{u_{i_L}}$ 。

最终,用户与物品的协同等级,将会随用户的交互历史一同,被输入大模型,以增强用户画像的生成。用户画像的生成过程将在 2.4 节进行介绍。

2.4 用户画像生成及对个性化推荐的增强

本文通过与大模型进行多次交互来逐步生成用户画像,以使得用户的交互历史得到更加充分的运用。用户画像由若干关键词组成,每个关键词反映了用户的一个偏好。对于每一次交互,用户当前的画像、交互历史中一批物品的画像、用户对这些物品的评分及协同等级,共同作为大模型的输入,由大模型对其中隐含的用户偏好进行分析推理,生成若干关于用户偏好的新的关键词。该过程可以被形式化为

$$\Delta \mathcal{P}_u = \text{LLM}_1(\mathcal{P}_u, [(\mathcal{P}_{i_1}, r_{u_{i_1}}, c_{u_{i_1}}), \dots, (\mathcal{P}_{i_B}, r_{u_{i_B}}, c_{u_{i_B}})])$$

式中: \mathcal{P}_u 为用户当前的画像, $r_{u_{i_k}}$ 代表用户 u 对本次输入的第 k 个物品的评分, B 为生成的新关键词个数, $\Delta \mathcal{P}_u$ 即为新关键词集合。

对于 $\Delta \mathcal{P}_u$ 中的每个关键词 Δt_u , 本文计算其与 \mathcal{P}_u 中所有关键词的语义相似度,只有全部语义相似度均小于设定的阈值时,才认为 Δt_u 是一个新的用户偏好,将其添加到 \mathcal{P}_u 之中。这一过程可以被形式化为

$$\Delta e_u = \text{LLM}_e(\Delta t_u), e_u^j = \text{LLM}_e(t_u^j),$$

$$s_u^j = \cos(\Delta e_u, e_u^j), j = 1, 2, \dots, k$$

式中: t_u^j 为 \mathcal{P}_u 中的一个关键词, k 为 \mathcal{P}_u 中的关键词个数, LLM_e 为将文本转换为特征向量的大模型, s_u^j 即为语义相似度。经处理后的用户画像使用 \mathcal{P}'_u 来表示。

当用户画像中的关键词数量达到阈值时,本文使用大模型对其进行提炼概括,合并相似含义的关键词,中和相互矛盾的关键词,提高其语义表达能力、可读性与简洁性。这一过程可以被形式化为

$$\mathcal{P}'_u = \text{LLM}_1(\mathcal{P}'_u)$$

在下一交互中, \mathcal{P}'_u 将会替代 \mathcal{P}_u , 与另一批物品的画像及交互信息一同进行输入,直到全部用户交互历史被输入完毕,得到最终的用户画像 $\mathcal{P}_u^{\text{LLM}}$ 。

进一步地, $\mathcal{P}_u^{\text{LLM}}$ 可以被用来增强个性化推荐。本文将 $\mathcal{P}_u^{\text{LLM}}$ 中的每个关键词用逗号串联起来, 整体转换为特征向量, 作为该用户的大模型表征, 本文使用 e_u^{LLM} 来表示。同样地, 由物品画像也可以得到物品的大模型表征。本文使用对比学习来将大模型表征引入推荐系统之中, 其目标是使得同一个用户或物品的推荐系统表征与自身的大模型表征相互接近, 而其他用户或物品的大模型表征相互远离。对比学习损失函数使用 InfoNCE (information noise contrastive estimation)^[47]。以用户为例, 形式上, 上述过程可以被形式化为

$$e_u^{\text{LLM}'} = (\text{LeakyReLU}(e_u^{\text{LLM}} \cdot W_{1_u})) \cdot W_{2_u}$$

$$\mathcal{L}_{\text{CL}_u} = -\frac{1}{U} \sum_{u=1}^U \log \frac{\exp(\cos(e_u^{\text{RS}}, e_u^{\text{LLM}'})/\tau)}{\sum_{u' \in \mathcal{U}_b} \exp(\cos(e_u^{\text{RS}}, e_{u'}^{\text{LLM}'})/\tau)}$$

$$\mathcal{L} = \mathcal{L}_{\text{RS}} + \beta \cdot (\mathcal{L}_{\text{CL}_u} + \mathcal{L}_{\text{CL}_i})$$

式中: W_{1_u} 、 W_{2_u} 为可训练的参数矩阵, e_u^{RS} 为用户 u 的推荐系统表征, $e_u^{\text{LLM}'}$ 是 e_u^{LLM} 经过线性层映射到与 e_u^{RS} 相同的线性空间后的表征, U 为系统中总的用户数量, \mathcal{U}_b 为当前一批的用户集合。 $\mathcal{L}_{\text{CL}_u}$ 为用户的对比学习优化目标, \mathcal{L}_{RS} 为推荐系统本身的优化目标, 总优化目标 \mathcal{L} 即为推荐系统优化目标与对比学习优化目标的加权和, β 与 τ 均为超参数。

通过上述引入方式, 最终推荐系统训练完成后, 其中的用户与物品表征既包含协同知识, 又包含大模型从文本与非文本信息, 及其拥有的世界知识中总结得到的知识, 从而使得用户与物品建模更加精准, 充分发挥大模型与推荐系统彼此的优势, 对于推荐系统表现的提升有潜在可能。

3 实验设定与结果分析

3.1 实验设定与实现细节

本文使用了两个数据集来对所提出的方法进行评测: MovieLens-25m^[48] 和 Amazon-CDs and Vinyl^[49]。经过预处理的两个数据集的统计信息如表 1 所示。

表 1 预处理后的数据集统计信息
Table 1 Dataset statistics after preprocessing

数据集	用户数量	物品数量	交互历史数量	稀疏度/%
MovieLens-25m	5 000	24 140	376 739	99.69
Amazon-CDs and Vinyl	5 000	135 257	375 061	99.94

本文将用户的交互历史按照时间戳顺序倒序排列, 即用户最近交互过的物品排在首位, 最先输入大模型。对于用户画像生成所需的物品类别

和描述文本, Amazon-CDs and Vinyl 数据集中均有包含, 而 MovieLens-25m 数据集中仅有物品类别, 并无显式的描述文本。为此, 本文从 TMDb (the movie database) 资料库获取数据集中全部电影的描述文本, 若 TMDb 网站缺失, 则再查询 IMDb (internet movie database) 资料库进行补充, 以对数据集中各个物品的描述文本进行充分的补足。

本文使用归一化折扣累计增益 (normalized discounted cumulative gain, NDCG)^[50] 和二元偏好 (binary preference, BPREF)^[51] 两个评价指标来评估推荐模型的表现。其中, BPREF 是一个基于偏好的信息检索评价指标, 主要用于评价检索结果中相关文档是否被排在不相关的文档之前, 在推荐系统场景下, 可以衡量正向物品能够尽可能排在负向物品之前的能力。BPREF 指标的计算公式为

$$\mathcal{B} = \frac{1}{R} \sum_r \left(1 - \frac{|比r排序靠前的n|}{\min(R, N)} \right)$$

式中: R 是正向物品数量, r 是其中的一个正向物品, N 是负向物品数量, n 是前 R 个负向物品的其中之一。

对于大模型, 本文使用 Llama 3 同时作为文本大模型和语义向量大模型。一次交互输入大模型的物品个数设为 5, 用户画像的关键词数量阈值设为 32, 关键词添加阈值设为 0.98。推荐模型内的表征维数设为 64, 大模型表征维数为 4 096, 用于映射的线性层中间层维数设为 2 080。所有模型的其他超参数均使用网格搜索的方法进行调优, 对于每组实验都使用 5 个不同随机种子进行初始化, 并报告各指标平均值。

对于两个协同信号, 经过在两个数据集上的参数搜索, 潜在兴趣物品集大小在 MovieLens-25m 数据集上设为 30, 在 Amazon-CDs and Vinyl 数据集上设为 20。协同等级数量设为 3, 分别为“喜欢”“不喜欢”与“中立”, 在中立等级中, 根据用户协同分数的正负性, 又进一步划分为“中立偏向喜欢”和“中立偏向不喜欢”。3 个协同等级的分数比例大约为 3:4:3, 较为均匀。

3.2 应用模型及对比方法

本文在 FM (factorization machines)^[52]、SimpleX^[53]、LightGCN^[54]、XSimGCL^[55] 和 LightGCL^[56] 这 5 个推荐系统研究中常见的模型上应用所提出的方法, 并评估其表现。其中, FM 与 SimpleX 是基于协同过滤的, 非图神经网络相关的基础模型; 而 LightGCN、XSimGCL 和 LightGCL 是图神经网络相关的模型。

对于每个模型, 本文的对比方法如下:

1) Base——不引入大模型生成的用户画像的原始推荐模型的表现;

2) UP——引入生成的用户画像 (user profile, UP), 但不大模型进行协同信号增强时推荐模型的表现;

3) UP+PI——引入大模型生成的用户画像, 且向大模型引入用户的潜在兴趣 (potential interest, PI) 时推荐模型的表现;

4) UP+CS——引入大模型生成的用户画像, 且向大模型引入用户的协同等级 (collaborative scale, CS) 时推荐模型的表现;

5) UP+PI+CS——引入大模型生成的用户画像, 且同时向大模型引入用户的潜在兴趣和协同等级时推荐模型的表现。

3.3 实验结果分析

表 2 反映了 5 个推荐模型在各个方法上的实验结果。其中, “*”代表本文的方法相比于“Base”方法具有显著差异, 本文使用双样本异方差假设 t-检验, P 值设为 0.05。“N”和“B”分别代表 ND-CG 与 BPREF 指标, 其后的数字 K 代表在前 K 个物品的排序列表中计算指标。粗体和下划线分别表示最优结果和次优结果。

表 2 5 个推荐模型应用不同方法时的表现
Table 2 Performances of 5 recommender models applied with different methods

模型	方法	MovieLens-25m						Amazon-CDs and Vinyl							
		N@1	N@5	N@10	N@20	B@5	B@10	B@20	N@1	N@5	N@10	N@20	B@5	B@10	B@20
FM	Base	0.505 6	0.755 5	0.753 4	0.741 1	0.476 9	0.571 4	<u>0.627 9</u>	<u>0.382 8</u>	0.624 5	0.648 7	0.636 6	0.332 3	0.431 0	0.517 3
	UP	<u>0.519 6</u>	0.757 9	0.754 6	0.740 7	0.466 2	0.562 6	0.623 4	0.381 6	0.622 9	0.646 9	0.635 2	0.330 8	0.431 0	0.518 6
	UP+PI	0.514 9	0.757 5	<u>0.754 8</u>	<u>0.741 6</u>	0.472 1	0.569 1	0.627 0	0.381 6	0.623 6	0.647 2	0.635 6	0.331 3	0.430 2	0.518 0
	UP+CS	0.523 4	0.758 5	0.755 8	<u>0.741 6</u>	0.463 3	0.559 9	0.626 7	0.382 4	0.623 6	0.647 5	0.635 8	<u>0.331 8</u>	<u>0.431 5</u>	0.517 8
	UP+PI+CS	0.515 6	<u>0.758 2</u>	0.755 8	0.743 4	<u>0.474 0</u>	<u>0.569 9</u>	0.628 2	0.382 9	<u>0.623 8</u>	<u>0.647 6</u>	<u>0.635 9</u>	0.331 3	0.431 9	<u>0.518 1</u>
SimpleX	Base	0.494 8	0.759 1	0.760 4	0.759 3	0.446 3	0.488 2	0.491 6	0.348 0	0.625 9	0.649 7	0.639 3	0.309 4	0.377 3	0.446 4
	UP	0.482 0	0.754 6	0.757 2	0.756 1	0.438 6	0.481 6	0.485 2	0.365 7*	0.652 5*	0.667 6*	0.651 5*	0.320 6*	0.379 2	0.448 4
	UP+PI	0.488 4	0.756 3	0.758 3	0.757 4	0.439 2	0.483 3	0.487 0	0.364 3*	0.652 9*	0.668 0*	0.651 8*	0.320 6*	0.380 9	0.451 1
	UP+CS	0.479 6	0.753 4	0.756 1	0.755 0	0.435 5	0.478 6	0.482 0	<u>0.371 3*</u>	<u>0.655 0*</u>	<u>0.669 4*</u>	<u>0.653 6*</u>	<u>0.326 6*</u>	<u>0.387 6</u>	<u>0.455 2</u>
	UP+PI+CS	<u>0.490 4</u>	<u>0.756 7</u>	<u>0.759 0</u>	<u>0.758 0</u>	<u>0.441 7</u>	<u>0.484 4</u>	<u>0.487 9</u>	0.372 4*	0.655 5*	0.669 7*	0.654 0*	0.331 1*	0.393 7	0.458 7
LightGCN	Base	0.528 2	0.780 8	0.781 2	0.779 9	0.490 9	0.549 7	0.552 3	0.380 5	0.637 7	0.658 3	0.646 2	0.354 7	<u>0.429 2</u>	<u>0.487 2</u>
	UP	0.534 4*	0.784 2*	0.783 5*	0.781 9*	0.496 0*	0.559 4*	0.561 5*	0.384 0*	0.669 2*	0.682 7*	0.666 1*	0.364 0	0.428 7	0.485 5
	UP+PI	0.534 7	0.783 9	0.783 0	0.781 5	0.495 1	0.556 4	0.558 7	0.415 1*	0.669 4*	0.683 1*	0.666 7*	0.364 5	0.428 3	0.486 6
	UP+CS	0.537 1*	0.785 1*	0.784 2*	0.782 6*	0.497 4*	0.560 6*	0.562 8*	<u>0.421 4*</u>	<u>0.672 1*</u>	<u>0.685 1*</u>	<u>0.668 8*</u>	0.368 0	0.427 6	0.487 9
	UP+PI+CS	<u>0.537 0*</u>	<u>0.784 9*</u>	<u>0.784 0*</u>	<u>0.782 4*</u>	<u>0.496 9*</u>	<u>0.559 6*</u>	<u>0.562 0*</u>	0.422 1*	0.673 8*	0.686 7*	0.669 3*	<u>0.365 9</u>	0.429 3	0.487 0
XSimGCL	Base	0.542 4	<u>0.788 2</u>	0.787 0	0.783 1	0.490 5	0.551 9	0.566 1	0.382 8	0.655 8	0.670 1	0.656 0	0.329 7	0.386 8	0.456 0
	UP	0.532 0	0.781 5	0.781 3	0.779 1	0.485 6	0.540 6	0.545 5	0.392 1*	0.663 3*	0.676 7*	0.661 8*	0.337 5*	0.396 4*	<u>0.462 6*</u>
	UP+PI	<u>0.548 5</u>	0.787 7	<u>0.787 3</u>	<u>0.785 2</u>	<u>0.499 0</u>	<u>0.554 0</u>	0.558 5	<u>0.395 0*</u>	0.664 8*	0.677 9*	0.662 3*	<u>0.340 6*</u>	0.398 9	0.462 8*
	UP+CS	0.541 1	0.785 2	0.785 0	0.782 7	0.494 7	0.549 2	0.553 9	0.394 6*	<u>0.664 4*</u>	0.677 5*	<u>0.662 5*</u>	0.340 1*	0.395 7	0.461 7
	UP+PI+CS	0.550 1	0.789 0	0.788 5	0.786 2	0.501 6	0.558 4	<u>0.563 0</u>	0.397 0*	0.664 8*	<u>0.677 8*</u>	0.662 8*	0.342 4*	<u>0.397 9*</u>	0.461 9
LightGCL	Base	0.521 4	0.778 0	0.778 5	0.777 2	0.483 7	0.542 3	0.545 2	0.381 3	0.631 6	0.653 5	0.641 9	0.343 6	0.417 1	0.483 9
	UP	<u>0.536 5*</u>	0.784 4*	0.784 1*	0.782 4*	0.492 6*	<u>0.556 9*</u>	<u>0.559 7*</u>	0.412 6*	0.668 1*	0.680 3*	0.663 8*	0.358 2*	0.417 2	0.482 9
	UP+PI	0.534 2	0.783 5	0.783 2	0.781 5	0.491 6	0.556 2	0.558 4	<u>0.417 4*</u>	<u>0.670 7*</u>	<u>0.682 3*</u>	<u>0.665 0*</u>	0.363 4*	<u>0.420 1</u>	<u>0.483 7</u>
	UP+CS	0.532 8	0.783 7	0.783 3	0.781 7	<u>0.493 4</u>	0.556 8	<u>0.559 7</u>	0.419 3*	0.671 9*	0.684 0*	0.666 7*	0.368 0*	0.421 7	0.483 0
	UP+PI+CS	0.537 0*	<u>0.784 3*</u>	0.784 1*	0.782 4*	0.494 9*	0.558 2*	0.561 0*	0.417 0*	0.669 2*	0.681 5*	0.664 7*	<u>0.364 6*</u>	0.419 6	0.482 4

注: *表示每种方法相比于相同模型的Base方法的表现, 有P<0.05的显著变化; 加粗表示每个模型各指标各方法中表现最优的结果, 下划线表示表现次优的结果。

1) 总体上, 应用本文提出的方法后, 推荐模型的表现能够在 NDCG 与 BPREF 两个指标上显著优于原始模型, 虽有少量不如原始模型的结果, 但其均不显著。这显示了本方法的有效性——通过与大模型进行多次交互, 本方法能够生成深入且简练的用户画像, 并能同时发挥大模型与推荐系统的优势, 建模出同时包含两者知识的、更加全面准确的用户表征。结果表明, 本方法在推荐质量与正负样本的区分能力的提升上均具有有效性。

2) 对于 $NDCG@K$ 和 $BPREF@K$, 一般 K 值较小时, 本文提出的方法相较于原始模型的提升率较大, 特别是 $NDCG@1$ 和 $BPREF@5$ 。这说明, 本方法对于排名较高的物品特别有效, 即正向物品更有可能排在推荐列表的最前面被推荐, 进一步展现了其有效性。

3) 在两个数据集上, 总体上所提出的两种引入协同信号的方式对推荐系统的表现均有增强效果。这显示了本文所提出的方法的有效性——通过引入用户的潜在兴趣, 用户的偏好得以被扩

展, 增强了推荐场景中“利用”与“探索”的平衡, 使得对于用户的建模更加深入与充分, 提升了推荐系统的表现; 通过引入用户的协同等级, 大模型在用户主观的评分之外, 获得了一个来自推荐系统的客观参考指标, 能够促使其对用户与物品间的情感态度进行更加准确的把握, 使得生成的用户画像更加准确, 进一步提升个性化推荐的表现。将二者同时进行引入, 能够综合二者的优点, 相比于单独引入其中一个协同信号或不引入协同信号, 都能够取得更优的结果。

3.4 用户画像生成样例分析

图 2 给出了当用户的潜在兴趣与最终获得的用户画像列表同时输入大模型, 并进一步完善用户画像的样例。在生成增强的用户画像之前, 大模型首先会根据提示词中的任务要求及输入, 输出它的推理信息, 并基于此生成用户画像。红色文字代表基于用户交互历史得到的用户画像列表中的关键特征, 即用户的真实偏好; 蓝色文字代表潜在兴趣文本中的关键特征, 即用户的潜在兴趣。



图 2 引入用户的潜在兴趣时用户画像生成样例

Fig. 2 Example of user profile generation when introducing user's potential interest

从该样例中可以看出, 大模型能够在推理信息中深入、有效地总结、归纳潜在兴趣分析中用户的潜在偏好, 将其抽象为若干文本短语, 添加到用户画像列表之中。并且, 对于已经存在于用户画像之中的潜在兴趣(潜在兴趣中以红色标记的特征), 大模型并不会将其重复添加至用户画像列表之中; 对于和用户画像相互矛盾的潜在兴

趣, 例如用户画像中记载用户不喜欢另类摇滚, 但另类音乐又是用户的潜在兴趣(在图中以下划线标记), 大模型也会以用户的真实兴趣为准, 不将这一潜在兴趣添加到用户画像列表。最终生成的增强的用户画像能够保留原来用户画像中的全部特征(为了简洁起见, 在图中以“当前的用户画像”表示), 且包含总结出的用户潜在兴趣, 相比

于原有用户画像更加深入、完整，能够实现更完善的用户建模。

图 3 给出了将用户的协同等级引入更新模块之中，以增强用户画像生成的样例。通过这种方式，大模型能够综合考虑用户对物品的评分及协

同等级，更准确地把握用户对每个物品的态度，从而生成更加精确的用户画像。从该样例推理信息中的红色文字可以看出，大模型能够同时关注每个物品的评分与协同等级，并根据已生成的用户画像进行推断，生成新的用户画像。

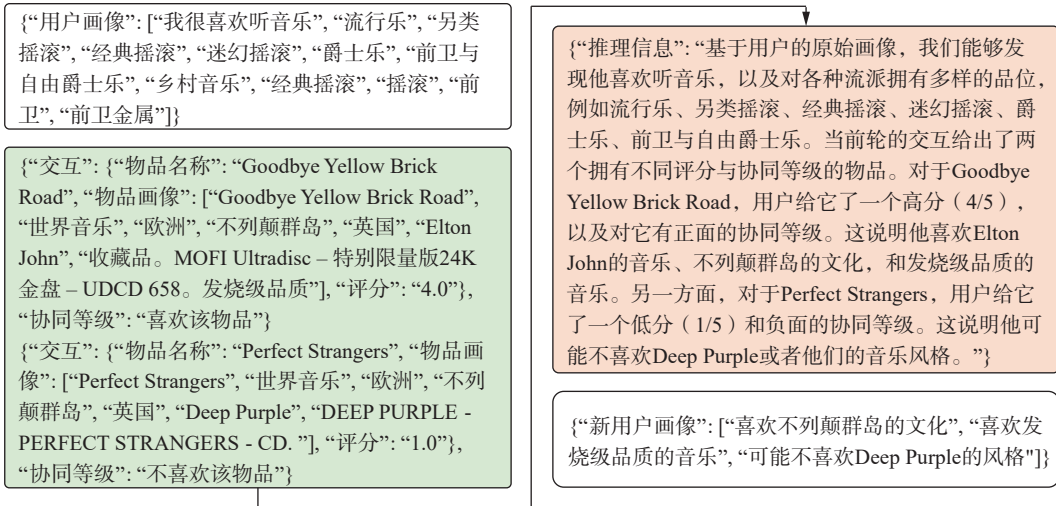


图 3 引入用户的协同等级时用户画像生成样例

Fig. 3 An example of user profile generation when introducing user's collaborative scale

综上，本文提出的两种协同信号增强方式均是有效的，能够提升用户画像的准确性与质量，进而提升推荐系统的表现。

4 结束语

本文提出了将推荐系统的协同信号通过提示词引入大模型，以增强大模型的用户画像生成，并进一步引入推荐系统，以提升个性化推荐表现的方法。其中，通过引入用户的潜在兴趣，用户交互历史外的偏好也被纳入考虑，可以得到更加全面、深入的用户画像，取得“利用”与“探索”之间的平衡；通过引入用户的协同等级，推荐系统中对于用户与物品匹配程度的客观评价同用户的主观评分一同被纳入考虑，可以得到更加综合、准确的用户画像。本文进一步通过与大模型进行多次交互的方式生成用户画像、通过对比学习方式将生成的用户画像引入推荐模型之中，以更加充分地利用用户交互历史，同时综合大模型与推荐系统各自的优势，实现强强联合，获得更加全面的用户建模。在两个数据集、多个推荐模型上的实验结果表明，本文所提出的方法能够显著提升推荐模型的表现，验证了本文方法的有效性。并且，本文方法可拓展性强，能够被应用于任意的推荐系统与大模型之中，且大模型与推荐系统相互关联却又相对独立，因此也非常适合在实际场景中线上部署。本文工作为后续类似工作启发

了新思路，未来可以尝试将用户在推荐系统内的表征融入大模型的输入词元中，研究对于用户画像生成的增强效果。

参考文献：

- [1] CHEN Hao, BEI Yuanchen, SHEN Qijie, et al. Macro graph neural networks for online billion-scale recommender systems[C]//Proceedings of the ACM Web Conference 2024. Singapore: ACM, 2024: 3598–3608.
- [2] FENG Zhichao, XIE Junjie, LI Kaiyuan, et al. Context-based fast recommendation strategy for long user behavior sequence in Meituan Waimai[C]//Companion Proceedings of the ACM Web Conference 2024. Singapore: ACM, 2024: 355–363.
- [3] DONG Zhikang, LIU Xiulong, CHEN Bin, et al. MuseChat: a conversational music recommendation system for videos[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2024: 12775–12785.
- [4] HASAN M R, FERDOUS J. Dominance of AI and machine learning techniques in hybrid movie recommendation system applying text-to-number conversion and cosine similarity approaches[J]. *Journal of computer science and technology studies*, 2024, 6(1): 94–102.
- [5] YANG Xinran, ZHANG Liaoniao, FENG Zixin. Personalized tourism recommendations and the E-tourism user

- experience[J]. *Journal of travel research*, 2024, 63(5): 1183–1200.
- [6] BILLSUS D, PAZZANI M J. Learning collaborative information filters[C]//International Conference on Machine Learning. Madison: PMLR, 1998: 46–54.
- [7] SARWAR B, KARYPIS G, KONSTAN J, et al. Item-based collaborative filtering recommendation algorithms [C]//Proceedings of the 10th International Conference on World Wide Web. Hong Kong: ACM, 2001: 285–295.
- [8] PURIFICATO E, BORATTO L, DE LUCA E W. User modeling and user profiling: a comprehensive survey [EB/OL]. (2024–02–15)[2025–06–25]. <https://arxiv.org/abs/2402.09660>.
- [9] ADOMAVICIUS G, TUZHILIN A. Expert-driven validation of rule-based user models in personalization applications[J]. *Data mining and knowledge discovery*, 2001, 5(1): 33–58.
- [10] CHEUNG CHIU B, WEBB G I. Using decision trees for agent modeling: improving prediction performance[J]. *User modeling and user-adapted interaction*, 1998, 8(1): 131–152.
- [11] KIM H N, ALKHALDI A, EL SADDIK A, et al. Collaborative user modeling with user-generated tags for social recommender systems[J]. *Expert systems with applications*, 2011, 38(7): 8488–8496.
- [12] SÁNCHEZ P, BELLOGÍN A. Building user profiles based on sequences for content and collaborative filtering[J]. *Information processing & management*, 2019, 56(1): 192–211.
- [13] P N, SAITEJA K, RAM K K, et al. University recommender system based on student profile using feature weighted algorithm and KNN[C]//2022 International Conference on Sustainable Computing and Data Communication Systems. Erode: IEEE, 2022: 479–484.
- [14] ZIGORIS P, ZHANG Yi. Bayesian adaptive user profiling with explicit & implicit feedback[C]//Proceedings of the 15th ACM International Conference on Information and Knowledge Management. Arlington: ACM, 2006: 397.
- [15] ZHANG Yi, KOREN J. Efficient Bayesian hierarchical user modeling for recommendation system[C]//Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. Amsterdam: ACM, 2007: 47–54.
- [16] XIE Yi, YU Shunzheng. A large-scale hidden semi-Markov model for anomaly detection on user browsing behaviors[J]. *IEEE/ACM transactions on networking*, 2009, 17(1): 54–65.
- [17] TANG Jie, YAO Limin, ZHANG Duo, et al. A combination approach to web user profiling[J]. *ACM transactions on knowledge discovery from data*, 2010, 5(1): 1–44.
- [18] AHMAD S B S, RAFIE M, GHORABIE S M. Spam detection on Twitter using a support vector machine and users' features by identifying their interactions[J]. *Multimedia tools and applications*, 2021, 80(8): 11583–11605.
- [19] CHEN Tinggui, YIN Xiaohua, PENG Lijuan, et al. Monitoring and recognizing enterprise public opinion from high-risk users based on user portrait and random forest algorithm[J]. *Axioms*, 2021, 10(2): 106.
- [20] TRIPATHI A, T S A, GUDDETI R M R. A reinforcement learning and recurrent neural network based dynamic user modeling system[C]//2018 IEEE 18th International Conference on Advanced Learning Technologies. Mumbai: IEEE, 2018: 411–415.
- [21] YU Zeping, LIAN Jianxun, MAHMOODY A, et al. Adaptive user modeling with long and short-term preferences for personalized recommendation[C]//Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence. Macao: IJCAI, 2019: 4213–4219.
- [22] GE Songwei, DOU Zhicheng, JIANG Zhengbao, et al. Personalizing search results using hierarchical RNN with query-aware attention[C]//Proceedings of the 27th ACM International Conference on Information and Knowledge Management. Torino: ACM, 2018: 347–356.
- [23] ZHU Y, LI Hao, LIAO Yikang, et al. What to do next: modeling user behaviors by time-LSTM[C]//International Joint Conference on Artificial Intelligence. Melbourne: IJCAI, 2017: 3602–3608.
- [24] MA Zhengyi, DOU Zhicheng, BIAN Guanyue, et al. PSTIE: time information enhanced personalized search [C]//Proceedings of the 29th ACM International Conference on Information & Knowledge Management. Virtual Event: ACM, 2020: 1075–1084.
- [25] GU Yulong, DING Zhuoye, WANG Shuaiqiang, et al. Hierarchical user profiling for E-commerce recommender systems[C]//Proceedings of the 13th International Conference on Web Search and Data Mining. Houston: ACM, 2020: 223–231.
- [26] QI Tao, WU Fangzhao, WU Chuhan, et al. News recommendation with candidate-aware user modeling[C]//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. Madrid: ACM, 2022: 1917–1921.
- [27] AVNY BROSH T, LIVNE A, SAR SHALOM O, et al.

- BRUCE: bundle recommendation using contextualized item embeddings[C]//Proceedings of the 16th ACM Conference on Recommender Systems. Seattle: ACM, 2022: 237–245.
- [28] CHENG Zhiyong, HAN Sai, LIU Fan, et al. Multi-behavior recommendation with cascading graph convolution networks[C]//Proceedings of the ACM Web Conference 2023. Austin: ACM, 2023: 1181–1189.
- [29] HAN Jinkun, LI Wei, CAI Zhipeng, et al. Multi-aggregator or time-warping heterogeneous graph neural network for personalized micro-video recommendation[C]//Proceedings of the 31st ACM International Conference on Information & Knowledge Management. Atlanta: ACM, 2022: 676–685.
- [30] WANG Xiaolin, SUN Guohao, FANG Xiu, et al. Modeling spatio-temporal neighbourhood for personalized point-of-interest recommendation[C]//Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence. Vienna: IJCAI, 2022: 3530–3536.
- [31] TANG Xiangru, ZHANG Xingyao, SHAO Yanjun, et al. Step-back profiling: distilling user history for personalized scientific writing[EB/OL]. (2024-06-20)[2025-06-25]. <https://arxiv.org/abs/2406.14275>.
- [32] ZHANG Jiarui. Guided profile generation improves personalization with LLMs[EB/OL]. (2024-09-19)[2025-06-25]. <https://arxiv.org/abs/2409.13093>.
- [33] SUN Zhongxiang, SI Zihua, ZANG Xiaoxue, et al. Large language models enhanced collaborative filtering[C]//Proceedings of the 33rd ACM International Conference on Information and Knowledge Management. Boise: ACM, 2024: 2178–2188.
- [34] ZHOU J, DAI Yijia, JOACHIMS T. Language-based user profiles for recommendation[EB/OL]. (2024-02-23)[2025-06-25]. <https://arxiv.org/abs/2402.15623>.
- [35] JALAN R, PRAKASH T, PEDANEKAR N. LLM-BRec: personalizing session-based social recommendation with LLM-BERT fusion framework[C]//The Second Workshop on Generative Information Retrieval. Washington: ACM, 2024: 1–10.
- [36] XI Yunjia, LIU Weiwen, LIN Jianghao, et al. Towards open-world recommendation with knowledge augmentation from large language models[C]//18th ACM Conference on Recommender Systems. Bari: ACM, 2024: 12–22.
- [37] WEI Wei, REN Xubin, TANG Jiabin, et al. LLMRec: large language models with graph augmentation for recommendation[C]//Proceedings of the 17th ACM International Conference on Web Search and Data Mining. Merida: ACM, 2024: 806–815.
- [38] REN Xubin, WEI Wei, XIA Lianghao, et al. Representation learning with large language models for recommendation[C]//Proceedings of the ACM Web Conference 2024. Singapore: ACM, 2024: 3464–3475.
- [39] TANG Zuoli, HUAN Zhaoxin, LI Zihao, et al. One model for all: large language models are domain-agnostic recommendation systems[J]. *ACM transactions on information systems*, 2025, 43(5): 1–27.
- [40] DU Yingpeng, WANG Ziyang, SUN Zhu, et al. Large language model with graph convolution for recommendation[EB/OL]. (2024-02-14)[2025-06-25]. <https://arxiv.org/abs/2402.08859>.
- [41] ZHENG Zhi, CHAO Wenshuo, QIU Zhaopeng, et al. Harnessing large language models for text-rich sequential recommendation[C]//Proceedings of the ACM Web Conference 2024. Singapore: ACM, 2024: 3207–3216.
- [42] EBRAT D, RUEDA L. Lusifer: LLM-based user simulated feedback environment for online recommender systems[C]//2025 IEEE 4th International Conference on Computing and Machine Intelligence. MI: IEEE, 2025: 1–6.
- [43] ZHANG Yang, FENG Fuli, ZHANG Jizhi, et al. CoLLM: integrating collaborative embeddings into large language models for recommendation[J]. *IEEE transactions on knowledge and data engineering*, 2025, 37(5): 2329–2340.
- [44] KIM S, KANG H, CHOI S, et al. Large language models meet collaborative filtering: an efficient all-round LLM-based recommender system[C]//Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Barcelona: ACM, 2024: 1395–1406.
- [45] LIU Zhongzhou, ZHANG Hao, DONG Kuicai, et al. Collaborative cross-modal fusion with large language model for recommendation[C]//Proceedings of the 33rd ACM International Conference on Information and Knowledge Management. Boise: ACM, 2024: 1565–1574.
- [46] ZHAO Chuang, SU Xing, HE Ming, et al. Collaborative knowledge fusion: a novel method for multi-task recommender systems via LLMs[J]. *IEEE transactions on knowledge and data engineering*, 2025, 37(9): 5017–5033.
- [47] VAN DEN OORD A, LI Yazhe, VINYALS O. Representation learning with contrastive predictive coding[EB/OL]. (2018-07-10)[2025-06-25]. <https://arxiv.org/abs/1807.03748>.
- [48] HARPER F M, KONSTAN J A. The MovieLens datasets: history and context[J]. *ACM transactions on interactive*

intelligent systems, 2016, 5(4): 1–19.

- [49] NI Jianmo, LI Jiacheng, MCAULEY J. Justifying recommendations using distantly-labeled reviews and fine-grained aspects[C]//Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing. Hong Kong: Association for Computational Linguistics, 2019: 188–197.
- [50] JÄRVELIN K, KEKÄLÄINEN J. Cumulated gain-based evaluation of IR techniques[J]. *ACM transactions on information systems*, 2002, 20(4): 422–446.
- [51] BUCKLEY C, VOORHEES E M. Retrieval evaluation with incomplete information[C]//Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. Sheffield: ACM, 2004: 25–32.
- [52] RENDLE S. Factorization machines[C]//2010 IEEE International Conference on Data Mining. Sydney: IEEE, 2011: 995–1000.
- [53] MAO Kelong, ZHU Jieming, WANG Jinpeng, et al. SimpleX: a simple and strong baseline for collaborative filtering[C]//Proceedings of the 30th ACM International Conference on Information & Knowledge Management. Virtual Event: ACM, 2021: 1243–1252.
- [54] HE Xiangnan, DENG Kuan, WANG Xiang, et al. LightGCN: simplifying and powering graph convolution network for recommendation[C]//Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. Virtual Event: ACM, 2020: 639–648.
- [55] YU Junliang, XIA Xin, CHEN Tong, et al. XSimGCL: towards extremely simple graph contrastive learning for re-

commendation[J]. *IEEE transactions on knowledge and data engineering*, 2024, 36(2): 913–926.

- [56] CAI Xuheng, HUANG Chao, XIA Lianghao, et al. LightGCL: simple yet effective graph contrastive learning for recommendation[EB/OL]. (2023–02–16)[2025–06–25]. <https://arxiv.org/abs/2302.08191>.

作者简介:



郭世圆, 硕士研究生, 主要研究方向为推荐与用户建模, 获国家发明专利授权 2 项, 申请软件著作权 1 项, 发表学术论文 2 篇。E-mail: gsy22@tsinghua.org.cn。



孙培杰, 副教授, 主要研究方向为推荐算法、用户行为建模。获 2024 年度“钱伟长中文信息处理科学技术奖”一等奖、自然语言处理实证方法会议 (EMNLP) 资源奖 (Resource Award), 获安徽省 2023 年度优秀博士学位论文, 发表学术论文 31 篇。E-mail: peijiesun@njupt.edu.cn。



张敏, 教授, 国际计算机学会出版委员会 (ACM Publications Board) 理事, 国际计算机学会信息系统期刊 (ACM TOIS) 首位亚洲主编, 担任国际计算机学会 2026 年信息检索大会 (SIGIR)、2025 年多媒体大会 (MM)、2024 年人机信息交互或检索会议 (CHIIR) 等会议主席或程序主席, 主要研究方向为网络检索、个性化推荐与用户建模。曾获评 IBM 全球杰出学者奖 (Global Faculty Award)、中国高校计算机专业优秀教师奖励计划、信息检索大会 (SIGIR) 时间检验奖、北京科学技术奖一等奖、“钱伟长中文信息处理科学技术奖”一等奖等。发表学术论文 300 余篇。E-mail: z-m@tsinghua.edu.cn。