

基于知识图谱和大语言模型增强的推荐系统研究

王敏, 高晓影, 汪诗蕊, 向阳

同济大学计算机科学与技术学院, 上海 200000

摘要

推荐系统的核心是用户和商品, 用户与商品之间的关系可抽象为图结构, 因此图神经网络在推荐领域具有广泛应用。但基于图的推荐交互数据较稀疏, 严重依赖于编号信息和图结构信息, 忽略了与用户和商品相关的有价值的文本信息, 表征信息较少。同时, 隐式反馈数据中存在一定的噪声和偏差, 为推荐系统理解用户行为与偏好带来挑战。为了解决这些问题, 提出了一种基于知识图谱和大语言模型增强的推荐系统。知识图谱可以提供商品的结构化信息, 使模型能学习商品之间的潜在关系, 理解用户行为和偏好。大语言模型具备非常出色的生成与理解能力, 可以通过提示工程技术, 深入分析并挖掘文本信息, 推理获取商品和用户画像特征。所提模型分别将这些辅助信息增强的特征编码, 并对表征进行增强以与图神经网络获得的ID表征对齐, 完成下游推荐任务。实验结果证明, 本文提出的系统可以全面地表征用户和商品, 具有较好的性能。

关键词

大语言模型; 知识图谱; 推荐系统; 表征学习

中图分类号: TP391.1

文献标志码: A

doi: 10.11959/j.issn.2096-0271.2025026

Research on recommendation system based on knowledge graph and large language model enhancement

WANG Min, GAO Xiaoying, WANG Shirui, XIANG Yang

College of Computer Science and Technology, Tongji University, Shanghai 200000, China

Abstract

The core of recommendation system is user and commodity, the relationship between user and commodity can be abstracted as graph structure, so graph neural network has been widely used in recommendation field. However, graph-based recommendation interaction data is sparse, relying heavily on numbering information and graph structure information, ignoring valuable text information related to user and commodity, and less representational information. At the same time, there are some noise and bias in the implicit feedback data, which bring challenges for recommendation system to understand user behavior and preferences. To solve these problems, this paper proposes a recommendation system based on knowledge graph and large language model enhancement. Knowledge graphs can provide structured information about commodities, enabling models to learn potential relationships between commodities and understand user behavior and preferences. Large language models have excellent ability to generate

and understand, and can use prompt engineering techniques to deeply analyze and mine text information, and deduce the features of commodity and user portraits. The features enhanced by these auxiliary information are encoded, and the representations are enhanced to align with the ID representations obtained by the graph neural network to complete the downstream recommendation task. The experimental results show that the proposed system can help to comprehensively characterize users and commodities, and has good performance.

Key words

large language model, knowledge graph, recommendation system, representation learning

0 引言

随着电子商务和网络应用进程的蓬勃发展,推荐系统已渗入人们的日常生活中,成为人们生活的重要组成部分。不管是搜索引擎、娱乐软件、社交媒体,还是购物应用等,都可以看到推荐系统的身影。推荐系统的本质是一种信息过滤系统,能够了解用户偏好,在大量信息中过滤出用户感兴趣的消息,进行精准推送,缓解信息过载状况。

早期对推荐系统的研究,最常见的是使用协同过滤方法计算用户和商品之间的相似度,从表示学习的角度研究并提高推荐系统性能,但该方法较简单,无法处理复杂的用户行为和数据输入。随着神经网络和深度学习技术的发展,深度学习推荐模型逐渐成为推荐系统领域研究的主流,其在一定程度上改善了推荐系统的性能。然而,其忽略了数据中的高阶结构信息,具有一定的局限性。因此,基于图的推荐系统逐渐成为研究的主流与热点,图神经网络利用图结构信息对节点之间的关系进行建模,每个节点都可以获取高阶邻居节点的信息,从而提高推荐系统的性能。

然而,由于用户与商品的交互相对较少,图结构通常较稀疏,表征信息较少,用户商品的关系建模不够充分。而且图推

荐系统的研究多数严重依赖于ID信息和图结构信息,忽略了数据集中丰富的文本信息,限制了推荐系统对用户行为与偏好的深层理解。此外,图结构中的用户商品交互多为隐式反馈(浏览、点击、收藏等),可能会有一定的噪声,如用户误点击、浏览过却不购买等,这些噪声会对推荐系统产生负面影响,使其对用户需求产生错误理解。为缓解图数据稀疏的状况,较经典的方法是引入辅助信息^[1],如商品知识图谱^[2]、数据集中的特征信息与评论文本^[3]等,其可以在一定程度上缓解图数据稀疏性的状况,提高推荐系统的性能。但该方法未能充分理解和挖掘文本信息之间的联系,无法获得较准确的用户和商品画像。其多使用传统的自然语言处理技术,直接根据文本获得嵌入向量来辅助推荐,未能充分挖掘用户和商品多段相关文本之间的联系,仅提供较浅显的表征,对用户和商品偏好的理解不够深刻。

为弥补上述方法的不足,本文提出基于知识图谱和大语言模型增强的推荐(knowledge-graph and large language model enhanced recommendation, KLERec)系统。随着海量数据的积累和深度学习技术的不断发展与进步,大语言模型(large language model, LLM)开始出现并兴起,其展现出了强大的文本理解与推理能力,能够高效挖掘不同文本之间的复杂关联,在自然语言处理领域表现出色。

因此，本文在将知识图谱作为辅助增强信息的基础上，引入LLM对数据集中的辅助文本信息进行处理，挖掘不同文本之间的联系，获取较准确的用户和商品画像；采用合适的方法从引入的辅助信息获取用户和商品表征并增强，通过图神经网络获取用户和商品的ID表征；采用双向对抗和对比学习的方法对这两种表征进行对齐，从而完成下游推荐任务，提高推荐系统的性能和鲁棒性。本文主要贡献如下。

- 本文借助LLM出色的理解和推理能力，设计严谨有效的提示模板和流程，充分理解和挖掘数据集中的文本信息，获取较准确的用户和商品画像。

- 本文对引入的辅助信息和用户商品ID分别编码，采用双向对抗和对比学习两种方法对齐两类表征，最大限度地利用两类表征信息。

- 本文将所提系统与当前较先进的多个图推荐模型相结合，并与基础模型对比，证明了所提系统的有效性。

1 相关工作

1.1 图在推荐系统中的应用

在图推荐系统的研究中，图卷积神经网络具有重要的作用。LightGCN (light graph convolution network)^[4]去除了GCN中的特征变换和非线性激活模块，显著降低了计算成本。LR-GCCF (linear residual graph convolutional collaborative filtering)^[5]去除了GCN中的非线性激活函数，同时提出了一种残差网络结构，缓解图卷积神经网络中的过度平滑状况。NGCF (neural graph collaborative filtering)^[6]在二部图中设计图卷积操作，对用户和商品嵌入进行传播，获得协同信号。

然而，推荐系统中多为隐式反馈数据，数据较稀疏，且含有一定的噪声和偏差，为基于图神经网络的推荐系统研究带来了一定的挑战与影响。为此，研究人员开始使用自监督学习方法作为辅助手段来增强推荐系统的鲁棒性。其中，SGL (self-supervised graph learning)^[7]通过节点丢弃、边缘丢弃和随机漫步3个算子，以不同的方式对图结构进行增强。HCCF (hypergraph contrastive collaborative filtering)^[8]在超图层面对图结构进行增强，通过对比学习捕获局部和全局的图结构协同关系，提升表征质量。文献[9]在图结构和语义空间中引入用户和商品的相邻节点，其中图结构邻居节点用于构建结构对比目标，语义邻居节点用于挖掘潜在的语义关系，并引入对比目标来优化表示学习。

此外，还有一些研究通过引入知识图谱来缓解图推荐系统中的数据稀疏状况，提升推荐系统性能。文献[2]建立了用户-商品交互图和知识图谱混合的协作知识图，并通过注意力机制来区分嵌入传播时邻居节点的重要性。RippleNet^[10]在知识图谱中扩展用户的兴趣路径，逐层探索与用户和商品相关的实体和关系，从而更好地捕捉用户的潜在兴趣。文献[11]主要通过多任务学习结合知识图谱，旨在同时学习多种特征，从而更好地捕捉用户、商品及其关联信息。

1.2 LLM在推荐系统中的应用

当前，LLM在推荐系统中的应用主要集中在数据增强、特征编码阶段以及评分或排序等任务上。借助LLM完成知识增强任务，主要通过提示工程对其进行引导，输出符合用户要求的内容。文献[12]提出

生成式推荐框架，引导LLM生成获取额外的新闻摘要、用户配置文件和合成新闻片段，并使用生成的数据更新当前数据，以供下一轮的提示和生成使用。文献[13]通过提示工程从LLM中获取商品的事实知识、推理用户的偏好信息，将获取的知识作为增强特征，以提升推荐模型的性能。文献[14]利用数据集中丰富的文本信息和提示工程，以多种方式增强交互图，增强对用户-商品交互、商品节点属性和用户节点属性的理解。当使用LLM完成特征编码任务时，通常直接将LLM作为文本特征编码器。文献[15]研究利用BERT (bidirectional encoder representations from transformer) 技术将用户的评论转换成高密度向量形式，优化了用户的嵌入特征表示。此外，标签选择^[16]、新闻推荐^[17]、代码示例推荐^[18]等研究中均使用LLM作为相关任务的辅助编码器，以丰富用户商品的特征表示、改善用户体验。

尽管LLM显示出一定的推荐能力，但其在实际应用中仍会存在一些问题。随着用户数据量的增长，LLM的计算需求及其

推理时间的成本会大幅增加。同时，LLM在生成文本时会出现“幻觉问题”，生成的文本内容缺乏真实性和准确性。

2 算法模型

KLERec系统结构如图1所示，共包含3个部分。

- 特征增强层：引入商品知识图谱获得商品的结构化信息，通过LLM对辅助信息进行推理，获得用户和商品的画像特征，对用户和商品特征进行增强。
- 表征增强层：通过TransR方法对商品知识图谱编码，获得商品实体表征，通过语言模型对LLM推理获得的用户和商品画像特征进行编码，再通过门控机制对两种商品表征进行增强融合，获得增强的商品表征。
- 表征对齐层：通过图神经网络获得用户和商品的ID表征，再利用双向对抗和对比学习两种方法对齐辅助信息，增强获得的用户商品表征和用户商品的ID表征，

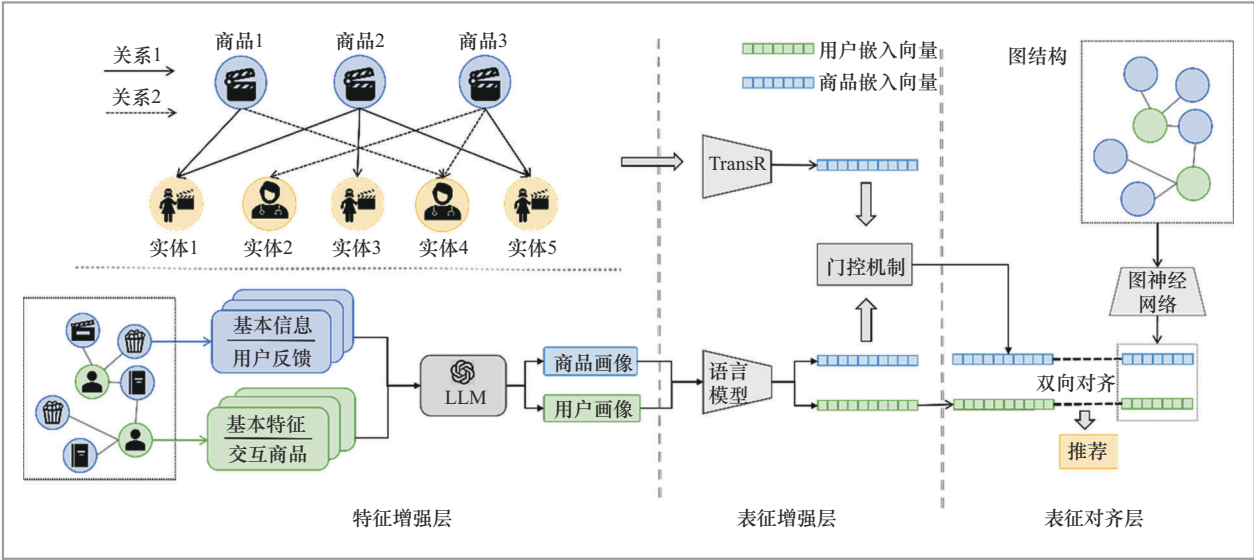


图1 KLERec系统结构

最大限度利用两种表征的信息。最后，模型进行学习预测，获得用户对不同商品的偏好，进行精准推荐。

2.1 特征增强层

特征增强层中的“特征”指推荐系统中的用户和商品特征，KLERec系统通过引入知识图谱和LLM对用户和商品特征进行增强。

2.1.1 知识图谱引入商品特征

许多推荐算法数据集中含有商品特征信息，但这些信息通常较少且不完整，无法全面展示商品的属性。同时，推荐系统难以挖掘并了解不同商品之间的关联，限制了推荐系统对用户需求的理解。

为弥补上述不足，本文引入知识图谱对商品侧特征进行增强。知识图谱系统化地整合了商品的多维特征，引入与商品相关的实体节点，通过“边”的形式表示商品和这些实体之间的关系，不仅弥补了商品特征信息的缺陷，还能通过图谱路径挖掘商品之间的潜在关联。这种结构化信息可以显著提升推荐系统的可解释性，使推荐结果更加准确和个性化。

2.1.2 LLM推理画像特征

在本文提出的系统中，画像特征被定义为对用户和商品的“偏好”刻画，用户的画像特征为该用户喜欢的商品类型，商品的画像特征为该商品吸引的用户类型。为了准确捕获这些偏好特征，本文使用LLM作为辅助工具，其具有强大的语义理解和总结推理能力，能够理解和推理文本中的复杂语义关系。具体地，本文采用提示工程方法借助LLM完成文本信息推理，

获得用户商品增强特征。

(1) 推理用户画像特征

在推荐算法数据集中，通常会包括用户的基本特征和用户与商品的交互记录（商品的名称和用户对不同商品的评论）。本文将这些信息整合送入提示模板中，引导LLM根据这些信息推理用户喜欢的商品类型。在设计提示模板时，让LLM提供得到推理结论所对应的解释性文本，可在一定程度上提高生成文本的质量和可靠性。

$$P_u = \text{LLM}_{\text{prompt}} \left(f_u + \sum_{k=1}^K \text{His}(i_k, u) \right) \quad (1)$$

其中， f_u 为数据集中的用户特征， $\text{His}(i_k, u)$ 为用户 u 与第 k （ $1 \leq k \leq K$ ）个商品的交互记录， $\text{LLM}_{\text{prompt}}()$ 为提示工程方法， P_u 为推理获得的用户画像特征。

(2) 推理商品画像特征

同样地，在数据集中，通常会包含商品的一些基本特征属性和用户与商品的交互数据。为此，本文获取每个商品的基本特征和不同用户的反馈，将这些信息整合送入提示模板中，引导LLM根据这些信息推理商品所吸引的用户类型，并给出原因。

$$P_i = \text{LLM}_{\text{prompt}} \left(f_i + \sum_{k=1}^K \text{Feedback}(u_k, i) \right) \quad (2)$$

其中， f_i 为数据集中商品较重要的属性特征， $\text{Feedback}(u_k, i)$ 为第 k （ $1 \leq k \leq K$ ）个用户对商品 i 的反馈评论， $\text{LLM}_{\text{prompt}}()$ 为提示工程方法， P_i 为推理获得的商品画像特征。

2.2 表征增强层

本文完成特征增强后，需要分别对用户和商品画像特征进行表征编码并处理，捕获数据的深层联系，帮助推荐系统更

好地理解用户行为和偏好，提升推荐性能。

2.2.1 获得用户商品表征

(1) 知识图谱实体表征

本文通过 TransR 方法对知识图谱中的实体和关系进行编码，将实体嵌入实体空间、关系嵌入关系空间，并使用线性映射将实体投影到关系空间中，从而更好地表示不同关系下的实体对。经过 TransR 方法编码后的知识图谱中的商品向量如式 (3) 所示。

$$\mathbf{K}_i = \text{TransR}(\mathbf{K}_g) \quad (3)$$

其中， $\mathbf{K}_g$ 为前文引入的商品知识图谱， $\mathbf{K}_i$ 为经过 TransR 方法编码后的商品表征。

(2) 语言模型文本表征

本文选择一些语言模型对前文推理获得的用户、商品画像特征进行编码，经过语言模型编码后的节点向量如下。

$$\mathbf{L}_u = \text{LM}(P_u) \quad (4)$$

$$\mathbf{L}_i = \text{LM}(P_i) \quad (5)$$

其中， P_u 和 P_i 分别为前文推理获得的用户和商品画像特征， $\text{LM}(\cdot)$ 为语言模型编码方法， $\mathbf{L}_u$ 和 $\mathbf{L}_i$ 为经过语言模型编码后的用户与商品画像语义表征。

2.2.2 增强商品表征

本文对增强特征进行编码后，会获得两组商品表征 $\mathbf{K}_i$ 和 $\mathbf{L}_i$ ，对这两种表征通过门控机制进行融合，获得最终增强的商品表征。

具体地，对于商品表征 $\mathbf{K}_i$ 和 $\mathbf{L}_i$ ，门控机制会计算一个门控向量 $\mathbf{g}$ 。

$$\mathbf{g} = \sigma(\mathbf{W}_g[\mathbf{K}_i; \mathbf{L}_i] + b_g) \quad (6)$$

其中， $\mathbf{W}_g$ 为权重矩阵， b_g 为偏置项， $\sigma(\cdot)$ 为 Sigmoid 函数。

模型利用门控向量 $\mathbf{g}$ 对两组向量进行加权融合，如式 (7) 所示。

$$\mathbf{F}_i = \mathbf{g} \odot \mathbf{K}_i + (1 - \mathbf{g}) \odot \mathbf{L}_i \quad (7)$$

其中， $\odot$ 为逐元素相乘， $\mathbf{F}_i$ 为最终融合获得的商品增强表征。

2.3 表征对齐层

通过第 2.2 节，本文获得了辅助信息增强所得到的用户和商品表征向量。本节先通过图神经网络获取用户和商品的 ID 表征，然后通过两种方法对用户和商品表征进行对齐。

2.3.1 获得用户商品 ID 表征

推荐系统中数据均可抽象为图结构，用户和商品为图中节点，用户和商品之间的交互为图结构中的边。本文选择经典的图神经网络对图中节点进行协同信号编码，经过图神经网络编码后的节点向量如下。

$$\mathbf{G}_u = \text{GNN}(u) \quad (8)$$

$$\mathbf{G}_i = \text{GNN}(i) \quad (9)$$

其中， u 和 i 为图结构中的用户和商品节点， $\text{GNN}(\cdot)$ 为图神经网络编码方法， $\mathbf{G}_u$ 和 $\mathbf{G}_i$ 则为经过图神经网络编码后的用户与商品编号协同表征。

2.3.2 用户商品表征对齐

根据前文所有工作，本文已经获得辅

助信息增强的表征和ID协同表征。由于这两种表征是在完全不同的语义空间获得的,其数据表征方式和分布同样有一定的差异,且都可能含有一定的噪声,需要通过一定的方式来对齐这两种表征信息,以期最大化利用两种表征,减少两种表征的噪声,提高推荐系统性能。

(1) 双向对抗网络

表征对齐的目的是最小化两个表征之间的差异,提高二者分布的相似性。基于此,笔者想到了对抗网络,其思想与表征对齐的期望较吻合。然而,仅使用单向对抗网络的训练过程是辅助增强表征分布不断向ID表征分布靠近的过程,失去了辅助增强表征作为辅助信息的意义。

因此,本文提出使用双向对抗网络的方法,如图2所示。网络中有两个判别器,分别将ID表征和辅助增强表征设为真向量,训练过程中先通过特征映射器将表征映射到与另一个表征相同的维度,判别器使每个表征的分布都不断向另一个表征靠近,最终实现两种表征的对齐,充分利用两种表征的信息与优势。

其中,特征映射器和判别器的结构都

是多层感知器(multilayer perceptron, MLP)。特征映射器由两个全连接层组成:第一个全连接层将输入表征从原始维度映射到128维,并使用LeakyReLU激活函数引入非线性变化;第二个全连接层将128维特征映射到与另一个表征相同的维度,使用Tanh激活函数将输出表征限制在 $[-1, 1]$ 。这样的结构可以学习到从一种表征到另一种维度表征的非线性变换,提取出更有效的特征。判别器同样由两个全连接层构成:第一个全连接层将输入表征从原始维度映射到128维,使用LeakyReLU激活函数增加非线性变化;第二个全连接层将128维特征映射到1维,并使用Sigmoid激活函数获取输入表征为真实表征的概率。

该训练过程的损失函数由判别器1和判别器2的损失组成。

$$L_D = L_{D_1} + L_{D_2} \quad (10)$$

以判别器1为例,其损失函数为:

$$L_{D_1} = -\frac{1}{N} \sum_{i=1}^N [\log(D_1(x_i)) + \log(1 - D_1(M_1(\hat{x}_i)))] \quad (11)$$

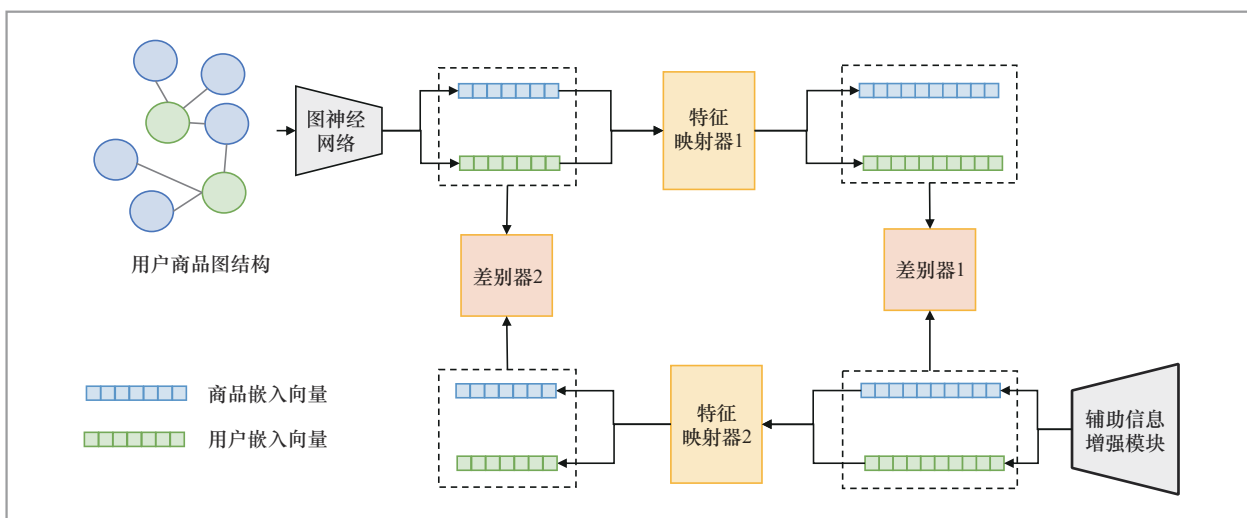


图2 双向对抗网络结构

其中, N 为数据量, D_1 为判别器 1 对应的变换, M_1 为映射器 1 对应的变换, $\mathbf{x}$ 为辅助增强表征, $\hat{\mathbf{x}}$ 为经过图神经网络编码获得的协同表征。

判别器 2 的损失 L_{D_2} 同理, 此处不再赘述。

需要注意的是, 在两种表征的分布不断进行对抗训练的过程中, 均会不断靠近判别器中真向量的分布, 可能会损失较多的原有表征信息。为了确保映射过程中尽可能保留原始关键信息, 本文引入重构模块来尽可能最小化这种信息丢失。以图神经网络获得的协同编码为例, 通过特征映射器 1 将其映射至与辅助增强表征相同的维度进行训练后, 再通过特征映射器 2 将其映射回原表征空间, 获得重构表征向量, 如图 3 所示。

其中, 特征映射器 1 和特征映射器 2 的结构也是 MLP, 具体结构与前文提到的特征映射器相同, 此处不再赘述。

本文分别经过两个特征编码器获得重构向量与原向量之后, 通过余弦相似度来衡量两组向量的相似性, 对应的损失函数定义如下。

$$L_{\text{recon}} = 1 - \frac{\mathbf{E}_o \cdot \mathbf{E}_r}{\|\mathbf{E}_o\| \|\mathbf{E}_r\|} \quad (12)$$

其中, $\mathbf{E}_o$ 为原始向量, $\mathbf{E}_r$ 为重构向量。

(2) 对比学习

对比学习是机器学习领域中的一种自监督学习方法, 其灵感来源于人类大脑神经在学习时对比或对照的思想。当人们面对不同的样本时, 通常通过比较它们之间的相似性或差异性进行区分, 基于这种理念, 对比学习通过比较计算不同样本之间的相似度或差异性, 使模型学习到更好的表征, 如图 4 所示。

辅助增强表征和图协同 ID 表征的维度不同, 本文需要通过特征映射器将辅助增强表征映射至与图协同 ID 表征相同的维度。特征映射器结构是 MLP, 由两个全连接层组成: 第一个全连接层将辅助增强表征从原始维度映射到隐藏维度, 用 LeakyReLU 激活函数引入非线性变化; 第二个全连接层将隐藏表示映射到与图协同 ID 表征相同的维度。

由于每个用户都会与商品有一定的交互, 本文先分别在图结构协同表征和经过特征映射的辅助增强表征中构建正负样本

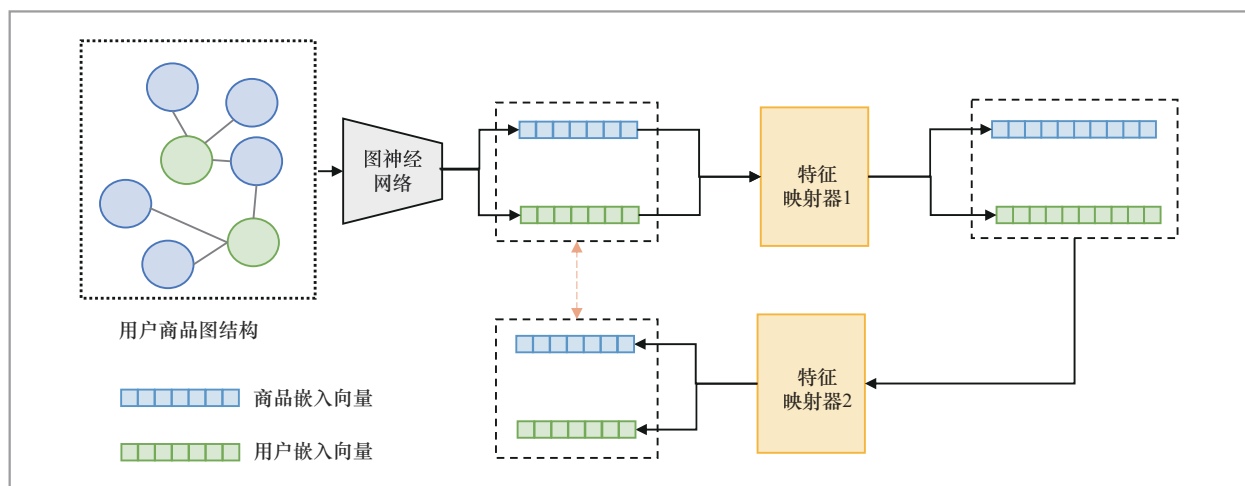


图3 重构部分模型

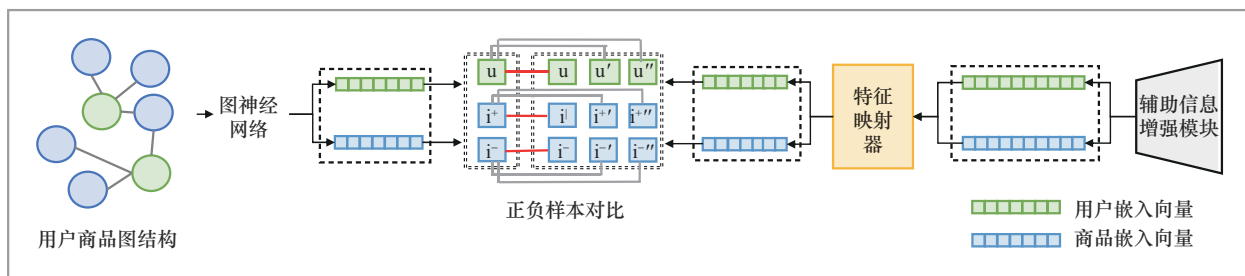


图4 对比学习

对，然后在两个表征之间构建正负样本对。获取对比学习中的正负样本对后，本文使用 InfoNCE 作为损失函数对模型进行训练，其定义如下。

$$L_{\text{InfoNCE}} = -\frac{1}{N} \sum_{i=1}^N \log \left(\frac{\exp\left(\frac{\mathbf{t}_i \cdot \mathbf{s}_{i^+}}{\tau}\right)}{\sum_{j=1}^N \exp\left(\frac{\mathbf{t}_i \cdot \mathbf{s}_j}{\tau}\right)} \right) \quad (13)$$

其中， N 为样本的数量， $\mathbf{t}_i$ 为图协同ID编码中的第 i 个样本， $\mathbf{s}_{i^+}$ 为辅助增强特征编码中与 $\mathbf{t}_i$ 相对应的正样本， $\mathbf{s}_j$ 为辅助增强表征中的第 j 个样本， τ 为温度超参数，用于调节相似度的分布。

2.4 模型学习

2.4.1 推荐预测

在本研究中，推荐系统的最终任务是根据用户与商品之间的隐式反馈（浏览、购买等）对用户进行商品排序推荐。因此，本文采用 BPR（bayesian personalized ranking）算法对模型进行训练，提高推荐的准确性。假设用户对商品的喜好是有排序的，其会对排名更靠前的商品感兴趣。BPR 算法的基本思想为：对于一个用户，如果有两个商品，将用户有过交互的商品定义为正样本，反之定义为负样本，每次训练时都将正样本放在负样本之前，通过

最大化正样本和负样本之间的分差来学习用户的偏好。BPR 算法最终可以获得每个用户对应的商品喜好排序，返回用户最感兴趣的那些商品。

为了最大化正负样本之间的分差，BPR 算法的损失函数定义为负对数似然函数。

$$L_{\text{BPR}} = -\sum_{(u, i^+, i^-) \in D} \ln \sigma(f_u(i^+) - f_u(i^-)) \quad (14)$$

其中， D 为数据集中所有的 (u, i^+, i^-) 三元组， $f_u(\cdot)$ 为用户 u 对不同商品的偏好排序函数。

2.4.2 模型损失函数

综上，模型在训练时，损失函数主要由3个部分组成。

当表征对齐方式为双向对抗时，模型的整体损失函数如下。

$$L_{\text{BiA}} = L_{\text{BPR}} + \lambda_b \cdot (L_D + L_{\text{recon}}) + L_2 \quad (15)$$

当表征对齐方式为双向对抗时，模型的整体损失函数如下。

$$L_{\text{Con}} = L_{\text{BPR}} + \lambda_c \cdot L_{\text{InfoNCE}} + L_2 \quad (16)$$

其中， L_2 为正则惩罚项，用于对模型进行约束，防止出现过拟合，具体如下。

$$L_2 = \lambda \cdot \|\omega\|^2 \quad (17)$$

其中， λ 为正则项超参数，控制正则化的力度与强度， ω 为模型参数。

λ_b 和 λ_c 均为超参数，分别用于控制双向对抗损失和对比损失在整体损失函数中的权重。具体训练时，可通过调整 λ 、 λ_b 和 λ_c 来提高推荐系统的性能。

3 实验

3.1 实验设置

3.1.1 数据集

本文主要选择了3个数据集，分别为 Amazon-Books、MovieLens-1M 和 Yelp 数据集。

这3个数据集中均有显式反馈数据（评分），因此首先需要获得隐式反馈。Amazon-Books 和 Yelp 数据集过于稀疏，因此本文不对其设置正评分阈值，设置 MovieLens-1M 的正评分阈值为4。此外，为了保障数据集的质量，本文执行 K 核过滤，过滤掉不够“核心”的用户和商品。

此外，本文还需要为商品构建知识图谱。对于 Amazon-Books 和 MovieLens-1M 数据集，本文遵循文献[19]中的方法，通过 ID 匹配，将商品映射到 Freebase 实体。对于 Yelp 数据集，本文从推荐算法数据集中提取信息作为知识图谱的数据。为了确

保知识图谱的质量，在预处理时会过滤掉出现次数较少的实体，并保留在三元组中多次出现的关系。

经过上述处理后，3个数据集的用户-商品交互和商品知识图谱属性见表1。

本文将用户与商品之间的交互数据表示为邻接矩阵，该矩阵的行代表用户，列代表商品，矩阵中的每个元素代表用户与商品之间的交互状态。当用户与商品发生交互时，矩阵对应位置的值为1；相反，若用户未与该商品进行交互，则矩阵对应位置的值为0。然后，按3:1:1的比例将数据集划分为训练集、验证集和测试集，分别用于后续推荐模型的训练与预测。

3.1.2 评价指标

本文使用两个广泛采用的基于排名的指标 Recall@ K 和 NDCG@ K 来衡量模型的有效性。

(1) Recall@ K

召回率（recall）是推荐算法在召回阶段常用的评价指标，表示当前给用户的推荐物品列表中有多少用户对其比较感兴趣。

$$\text{Recall}_u = \frac{|R_u \cap T_u|}{|T_u|} \tag{18}$$

其中， u 为用户， R_u 为模型经过训练后获得的该用户的推荐物品列表， K 为推荐物

表1 数据集基本信息

数据集	用户-商品交互			知识图谱		
	用户数/个	商品数/个	交互数/个	实体数/个	关系数/个	三元组数/个
Amazon-Books	13 549	9 569	372 753	10 744	24	272 418
MovieLens	6 034	3 103	572 755	6 220	60	245 461
Yelp	10 604	9 026	661 498	18 527	37	220 173

品列表的长度, T_u 表示测试集中用户真实感兴趣的物品集合。

(2) NDCG@K

归一化折损累计增益 (normalized discounted cumulative gain, NDCG) 在推荐系统中用于对模型返回的推荐物品列表进行评估, 其计算过程如下。

$$CG_k = \sum_{i=1}^k rel_i \quad (19)$$

其中, rel_i 表示第 i 个推荐商品的相关性评分, 累积增益 (CG) 是前 k 个推荐商品的相关性得分的总和。

折损累积增益 (DCG) 在 CG 的基础上引入了折扣因子, 通常是根据位置 i 的对数值进行折扣。位置越靠后, 其相关性得分的影响越小。

$$DCG_k = rel_1 + \sum_{i=2}^k \frac{rel_i}{\log_2(i+1)} \quad (20)$$

$$IDCG_k = \sum_{i=1}^{|REL_k|} \frac{2^{rel_i} - 1}{\log_2(i+1)} \quad (21)$$

其中, $|REL_k|$ 为推荐返回的按商品重要性依次排列的列表, IDCG 表示理想折损累积增益, 是推荐项按相关性评分从高到低排序后计算的 DCG。它代表了在理想情况下, 前 k 个推荐商品的 DCG。

$$NDCG_k = \frac{DCG_k}{IDCG_k} \quad (22)$$

NDCG 表示归一化折损累积增益, 是实际 DCG 与 IDCG 的比值, 其值范围为 0 到 1, 越接近 1 表示推荐系统的效果越好。

3.1.3 实验细节

本文相关实验均在 NVIDIA RTX 6000 Ada 显卡上完成, 其具备强大的计算能力。同时, 为了公平地比较推荐系统性能, 本

文的基线模型均使用笔者发布的开源代码。

在通过提示工程借助 LLM 完成数据增强时, 本文使用 OpenAI 的 “gpt-3.5-turbo” 模型。在进行对比实验时, 本文均选用 “text-embedding-ada-002” 作为文本编码模型来获得增强后数据的向量表示, 再进行文本编码器的相关讨论。

具体训练时, 本文选用的优化器为 Adam, 设置的学习率为 1.0×10^{-3} , 设置的 Batch 大小为 4 096, 每 10 个训练轮次进行一次验证。同时, 本文使用早停策略对模型进行约束, 避免模型过拟合。

3.2 对比试验

3.2.1 基线模型

本文在图推荐系统的基础上, 借助知识图谱和 LLM 为工具, 充分利用辅助信息, 获得增强表征, 并将图协同信号 ID 表征与增强表征对齐, 以提高原有图推荐系统的性能。因此, 本文的基线模型分别在推荐系统中较经典图神经网络的基础上进行改进, 并对实验结果进行评估。对 5 个基线模型的简要介绍如下。

- LightGCN^[4]: 其去除了图卷积神经网络中的特征转换和非线性激活模块, 是一种轻量级的图卷积神经网络。

- LR-GCCF^[5]: 其在图卷积神经网络的基础上去除了非线性激活模块, 使用线性的激活函数, 降低了计算复杂度; 同时提出了基于残差的网络结构, 缓解了 GCN 的过渡平滑状况。

- SGL^[7]: 其通过 3 种算子 (node dropout、edge dropout 和 random walk) 为图结构中的节点创建对比视图, 提高模型对噪声交互的鲁棒性。

- SimGCL^[20]: 其向嵌入空间均匀加入噪声, 创建对比视图。

• AutoCF^[21]：其通过掩码再重建的方法实现自动图增强，通过掩码自编码器对增强后的图结构掩码重构，增强表征能力。

3.2.2 实验结果

为了证明本文所提系统在提高图推荐系统性能方面的有效性，本文参考文献[22]，设计了一系列对比实验，将 KLERec 系统分别与 5 种经典的图神经网络模型相结合，并进行性能对比，评估在不同评估指标下两种对齐方式（Base 为不做任何改进的基础模型，BiA 为双向对抗方法，Con 为对比学习方法）的表现。3 个数据集的实验结果分别见表 2、表 3、表 4。

本文由上述实验结果可知以下结论。

- 本文提出的基于知识图谱和 LLM 的

推荐系统的研究是有效的，其相比于基线模型均有一定的提升，可提高推荐系统的准确率，提升推荐系统性能。

- 在多数情况下，使用对比学习方法进行对齐的效果优于使用双向对抗方式，推测其原因可能是对比学习通过对比损失直接优化表征空间，减小相似对之间的距离，较稳定且直观。而双向对抗方法可能会在训练过程中增加较多复杂性，使模型难以集中于对齐任务本身。

3.3 变化系统

本文为验证引入的不同辅助信息对推荐系统性能的影响，设计了以下 3 个变体系统进行对比消融实验。

- KLERec_0：其不引入任何辅助信息。

表2 Amazon-Books数据集结果

基线模型	表征对齐方法	Recall@K			NDCG@K		
		K=10	K=20	K=40	K=10	K=20	K=40
LightGCN	Base	0.0860	0.1388	0.2169	0.0751	0.0946	0.1205
	KLERec_BiA	0.0878	0.1416	0.2202	0.0762	0.0962	0.1222
	KLERec_Con	0.0900	0.1452	0.2268	0.0789	0.0993	0.1261
LR-GCCF	Base	0.0809	0.1325	0.2077	0.0706	0.0899	0.1148
	KLERec_BiA	0.0839	0.1357	0.2115	0.0729	0.0922	0.1172
	KLERec_Con	0.0843	0.1379	0.2174	0.0739	0.0938	0.1200
SGL	Base	0.0963	0.1526	0.2322	0.0842	0.1051	0.1314
	KLERec_BiA	0.0987	0.1563	0.2366	0.0858	0.1073	0.1338
	KLERec_Con	0.0976	0.1566	0.2386	0.0859	0.1079	0.1350
SimGCL	Base	0.0930	0.1477	0.2300	0.0813	0.1016	0.1287
	KLERec_BiA	0.0950	0.1507	0.2338	0.0826	0.1032	0.1306
	KLERec_Con	0.0942	0.1517	0.2348	0.0820	0.1034	0.1307
AutoCF	Base	0.1027	0.1612	0.2411	0.0906	0.1123	0.1386
	KLERec_BiA	0.1046	0.1628	0.2449	0.0917	0.1133	0.1405
	KLERec_Con	0.1049	0.1660	0.2483	0.0917	0.1145	0.1415

表3 MovieLens-1M数据集实验结果

基线模型	表征对齐方法	Recall@K			NDCG@K		
		K=10	K=20	K=40	K=10	K=20	K=40
LightGCN	Base	0.1380	0.2220	0.3385	0.2246	0.2355	0.2696
	KLERec_BiA	0.1383	0.2223	0.3387	0.2250	0.2359	0.2698
	KLERec_Con	0.1432	0.2309	0.3467	0.2294	0.2417	0.2759
LR-GCCF	Base	0.1362	0.2206	0.3335	0.2156	0.2286	0.2629
	Self_BiA	0.1368	0.2217	0.3365	0.2157	0.2293	0.2643
	KLERec_Con	0.1379	0.2226	0.3376	0.2185	0.2314	0.2663
SGL	Base	0.1278	0.2057	0.3131	0.2067	0.2173	0.2499
	Self_BiA	0.1384	0.2192	0.3281	0.2180	0.2297	0.2628
	KLERec_Con	0.1490	0.2375	0.3569	0.2330	0.2464	0.2828
SimGCL	Base	0.1394	0.2252	0.3415	0.2257	0.2375	0.2721
	KLERec_BiA	0.1485	0.2361	0.3544	0.2295	0.2434	0.2797
	KLERec_Con	0.1490	0.2374	0.3545	0.2334	0.2466	0.2820
AutoCF	Base	0.1083	0.1744	0.2681	0.1925	0.1966	0.2224
	KLERec_BiA	0.1127	0.1803	0.2743	0.2017	0.2058	0.2320
	KLERec_Con	0.1153	0.1827	0.2756	0.2030	0.2070	0.2322

表4 Yelp数据集实验结果

基线模型	表征对齐方法	Recall@K			NDCG@K		
		K=10	K=20	K=40	K=10	K=20	K=40
LightGCN	Base	0.0612	0.1050	0.1742	0.0830	0.0965	0.1247
	KLERec_BiA	0.0633	0.1071	0.1767	0.0855	0.0987	0.1271
	KLERec_Con	0.0664	0.1123	0.1850	0.0894	0.1034	0.1331
LR-GCCF	Base	0.0598	0.1024	0.1697	0.0806	0.0939	0.1214
	Self_BiA	0.0621	0.1071	0.1760	0.0837	0.0978	0.1260
	KLERec_Con	0.0638	0.1089	0.1804	0.0856	0.1000	0.1291
SGL	Base	0.0624	0.1073	0.1772	0.0843	0.0985	0.1270
	Self_BiA	0.0697	0.1183	0.1946	0.0927	0.1083	0.1397
	KLERec_Con	0.0686	0.1185	0.1944	0.0917	0.1081	0.1391
SimGCL	Base	0.0605	0.1028	0.1703	0.0821	0.0950	0.1225
	KLERec_BiA	0.0615	0.1038	0.1723	0.0828	0.0957	0.1237
	KLERec_Con	0.0663	0.1114	0.1840	0.0896	0.1032	0.1328
AutoCF	Base	0.0734	0.1233	0.2032	0.0983	0.1137	0.1462
	KLERec_BiA	0.0743	0.1256	0.2058	0.0988	0.1152	0.1480
	KLERec_Con	0.0742	0.1266	0.2066	0.0988	0.1158	0.1486

- KLERec_kg：其仅引入商品知识图谱作为辅助信息。

- KLERec_llm：其仅引入 LLM 作为特征增强的工具，推理用户商品画像特征。

由前文可知，对比学习在多数情况下的推荐效果要强于双向对抗，因此，表5展示了基线模型为 LightGCN、对齐方式为对比学习时，所设计的变体系统和 KLERec 推荐系统在不同数据集上的实验结果。

将上述 3 个变体系统和 KLERec 系统进行分析对比，可得出如下结论。

- KLERec 系统的推荐效果在 3 个数据集上均优于其他变体系统，表明了引入知识图谱作为辅助信息和使用 LLM 作为工具辅助推理用户商品画像特征在提升推荐系统推荐效果方面的有效性。

- KLERec_llm 系统的推荐效果优于 KLERec_kg 系统，说明使用 LLM 作为工具辅助推理获得用户商品画像特征时，对推荐系统推荐效果的提升要比引入知识图谱作为辅助信息时更加显著。

3.4 实验讨论

3.4.1 表征处理方法讨论

本文将获得的两类表征通过双向对抗和对比学习的方法进行对齐，使模型能充分利用这两种表征的信息，提升推荐系统

的性能。

在本节中，本文探究其余不同的表征处理方法对推荐系统推荐性能的影响，选取了一些较经典的特征处理与融合方法，有拼接、按位加、按位乘、TFN（tensor fusion network）^[23]和 AFF（attentional feature fusion）^[24]。

具体地，本文选择 Amazon-Books 数据集，以 Recall@20 为评价指标，训练并获取系统在不同的特征处理方法下的性能。实验结果见表6。

从表6中可以看出，本文采用的两种对齐方法对表征进行处理后推荐的准确率均优于其他表征处理方法，证明了本文所提方法性能的有效性与优越性。将两种表征进行对齐，不仅提高了推荐系统推荐的准确率，还在多数情况下超过了传统的表征处理方法，显示出巨大的潜力。同时，本文发现在有些情况下，传统的表征处理方法的表现甚至比仅使用图协同 ID 表征还要差。推测其原因，可能为进行特征处理融合后，不仅没有进行有效的特征相互补充，反而可能引入了噪声或冗余信息，对推荐系统捕获用户和商品之间的关系造成了干扰，降低了推荐效果。

3.4.2 推荐系统鲁棒性讨论

同样地，参考文献[22]的实验探究，

表5 消融实验结果

变体系统与 KLERec 系统	Amazon-Books		MovieLens-1M		Yelp	
	Recall@20	NDCG@20	Recall@20	NDCG@20	Recall@20	NDCG@20
KLERec_0	0.1388	0.0946	0.2220	0.2355	0.1050	0.0965
KLERec_kg	0.1418	0.0967	0.2260	0.2374	0.1108	0.1017
KLERec_llm	0.1440	0.0987	0.2269	0.2390	0.1119	0.1029
KLERec	0.1452	0.0993	0.2309	0.2417	0.1123	0.1034

表6 系统在不同特征处理方法下的性能对比

表征处理方法	不同基线模型下的 Recall@20				
	LightGCN	LR-GCCF	SGL	SimGCL	AutoCF
拼接	0.1268	0.1121	0.0607	0.1506	0.0396
按位加	0.1010	0.0767	0.0819	0.1358	0.0445
按位乘	0.1138	0.0745	0.0777	0.0983	0.0328
TFN	0.1193	0.0794	0.0811	0.1231	0.0372
AFF	0.1242	0.1123	0.1084	0.1427	0.0459
KLERec_BiA	0.1416	0.1357	0.1563	0.1507	0.1628
KLERec_Con	0.1452	0.1379	0.1566	0.1517	0.1660

为分析本文提出的推荐系统性能是否始终优于原始图推荐系统，本文向原始交互数据中添加噪声，即添加不存在的用户商品交互（随机将用户商品交互邻接矩阵中的0值置为1），对推荐系统的鲁棒性进行分析测试。具体地，本文向用户商品的原始邻接矩阵中依次加入5%、10%、15%和20%的噪声，比较分析推荐系统在这些扰动下的性能。

此处以 Amazon-Books 数据集为例，向其原始数据中分别加入不同的噪声，分析比较 LightGCN 和本研究所提出系统的 Recall@20 和 NDCG@20 值，实验结果见表7、表8。

由表7、表8可知，随着加入噪声的增多，推荐系统的推荐性能均在逐渐下降。而且加入不同比例的噪声时，本文提出的系统性能始终优于基线模型，表明本文提出的系统在面对噪声数据时具有较强的鲁棒性与稳定性。同时，可以观察到加入噪声后，使用对比学习进行对齐的方法在多数情况下推荐系统的性能优于使用双向对抗方式进行对齐的方法，面对噪声时具有更强的稳定性。推测其原因，可能为

对比学习实现表征对齐的方法为构建正负样本对，即使存在噪声，推荐系统仍然可以学习到相似样本之间的关系，受噪声影响较小。

表7 不同噪声比例下推荐系统的 Recall@20 值

噪声比例	NDCG@20		
	Base	KLERec_BiA	KLERec_Con
5%	0.1254	0.1270	0.1320
10%	0.1144	0.1162	0.1192
15%	0.1043	0.1052	0.1108
20%	0.0961	0.0962	0.1026

表8 不同噪声比例下推荐系统的 NDCG@20 值

噪声比例	NDCG@20		
	Base	KLERec_BiA	KLERec_Con
5%	0.0875	0.0886	0.0915
10%	0.0814	0.0819	0.0841
15%	0.0755	0.0765	0.0797
20%	0.0720	0.0720	0.0757

4 结束语

基于图的推荐系统中用户商品交互较稀疏, 过于依赖ID和图结构信息, 表征信息较少, 忽略了数据集中有价值的文本信息, 同时图中交互数据可能存在一定噪声和偏差, 因此本文引入知识图谱作为商品辅助特征、LLM作为工具, 辅助推理获得商品用户画像特征, 对用户和商品进行特征增强, 充分利用商品结构信息和数据集中的文本信息, 获得全面、丰富与准确的用户商品表征, 提升推荐系统的性能。本文选取了3个数据集, 进行了大量的对比试验与探究, 证明了本研究的有效性。未来, 我们可以进一步引入其余的辅助信息, 对用户和商品进行特征增强, 不断改善用户和商品表征, 进一步改善推荐系统的性能。

参考文献:

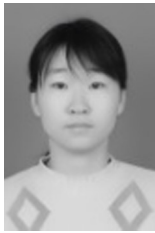
- [1] GUO G B, QIU H H, TAN Z H, et al. Resolving data sparsity by multi-type auxiliary implicit feedback for recommender systems[J]. Knowledge-Based Systems, 2017, 138: 202–207.
- [2] WANG X, HE X N, CAO Y X, et al. KGAT[C]//Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM, 2019: 950–958.
- [3] WANG C Y, GUO Z Q, LI G H, et al. A light heterogeneous graph collaborative filtering model using textual information[J]. Knowledge-Based Systems, 2021, 234: 107602.
- [4] HE X N, DENG K, WANG X, et al. LightGCN[C]//Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2020: 639–648.
- [5] CHEN L, WU L, HONG R C, et al. Re-visiting graph based collaborative filtering: a linear residual graph convolutional network approach[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(1): 27–34.
- [6] WANG X, HE X N, WANG M, et al. Neural graph collaborative filtering[C]//Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2019: 165–174.
- [7] WU J C, WANG X, FENG F L, et al. Self-supervised graph learning for recommendation[C]//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2021: 726–735.
- [8] XIA L H, HUANG C, XU Y, et al. Hypergraph contrastive collaborative filtering[C]//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2022: 70–79.
- [9] LIN Z H, TIAN C X, HOU Y P, et al. Improving graph collaborative filtering with neighborhood-enriched contrastive learning[C]//Proceedings of the ACM Web Conference 2022. New York: ACM, 2022: 2320–2329.
- [10] WANG H W, ZHANG F Z, WANG J L, et al. RippleNet[C]//Proceedings of the 27th

- ACM International Conference on Information and Knowledge Management. New York: ACM, 2018: 417–426.
- [11] WANG H W, ZHANG F Z, ZHAO M, et al. Multi-task feature learning for knowledge graph enhanced recommendation[C]//Proceedings of the World Wide Web Conference. New York: ACM, 2019: 2000–2010.
- [12] LIU Q J, CHEN N, SAKAI T, et al. ONCE: boosting content-based recommendation with both open- and closed-source large language models[C]//Proceedings of the 17th ACM International Conference on Web Search and Data Mining. New York: ACM, 2024: 452–461.
- [13] XI Y J, LIU W W, LIN J H, et al. Towards open-world recommendation with knowledge augmentation from large language models[EB]. arXiv preprint, 2023, arXiv: 2306.10933.
- [14] WEI W, REN X B, TANG J B, et al. LLM-Rec: large language models with graph augmentation for recommendation[C]//Proceedings of the 17th ACM International Conference on Web Search and Data Mining. New York: ACM, 2024: 806–815.
- [15] MAO Z M, WANG H M, DU Y M, et al. UniTRec: a unified text-to-text transformer and joint contrastive learning framework for text-based recommendation[EB]. arXiv preprint, 2023, arXiv: 2305.15756.
- [16] HE J D, XU B W, YANG Z, et al. PTM4Tag: sharpening tag recommendation of stack overflow posts with pre-trained models[EB]. arXiv preprint, 2022, arXiv: 2203.10965.
- [17] ZHANG Y H, DING H, SHUI Z R, et al. Language models as recommender systems: evaluations and limitations[C]//I (Still) Can't Believe It's Not Better! NeurIPS 2021 Workshop.[S.l.: s.n.], 2021.
- [18] RAHMANI S, NAGHSHZAN A, GUERROUJ L. Improving code example recommendations on informal documentation using BERT and query-aware LSH: a comparative study[EB]. arXiv preprint, 2023, arXiv: 2305.03017.
- [19] ZHAO W X, HE G L, YANG K L, et al. KB4Rec: a data set for linking knowledge bases with recommender systems[J]. Data Intelligence, 2019, 1(2): 121–136.
- [20] LIU C, YU C H, GUI N, et al. SimGCL: graph contrastive learning by finding homophily in heterophily[J]. Knowledge and Information Systems, 2024, 66(3): 2089–2114.
- [21] XIA L H, HUANG C, HUANG C Z, et al. Automated self-supervised learning for recommendation[C]//Proceedings of the ACM Web Conference 2023. New York: ACM, 2023: 992–1002.
- [22] REN X, WEI W, XIA L, et al. Representation learning with large language models for recommendation[C]//Proceedings of the ACM Web Conference 2024. 2024: 3464–3475.
- [23] ZADEH A, CHEN M H, PORIA S, et al. Tensor fusion network for multimodal sentiment analysis[EB]. arXiv preprint, 2017, arXiv: 1707.07250.
- [24] DAI Y M, GIESEKE F, OEHMCKE S, et al. Attentional feature fusion[C]//Proceedings of the 2021 IEEE Winter Conference on Applications of Computer Vision (WACV). Waikoloa: IEEE, 2021: 3560–3569.

作者简介



王敏（2002-），女，同济大学计算机科学与技术学院硕士生，主要研究方向为推荐系统、大语言模型、知识图谱。



高晓影（1996-），女，同济大学计算机科学与技术学院博士生，主要研究方向为推荐系统、自然语言处理。



汪诗蕊（2001-），女，同济大学计算机科学与技术学院硕士生，主要研究方向为数据挖掘、知识图谱。



向阳（1962-），男，博士，同济大学计算机科学与技术学院教授，主要研究方向为数据挖掘、自然语言处理、知识图谱。

收稿日期: 2024-09-27
通信作者: 向阳,shxiangyang@tongji.edu.cn
基金项目: 国家自然科学基金项目(No.72071145)
Foundation Item: The National Natural Science Foundation of China (No.72071145)