

基于多视角的自监督推荐方法^{*}

王丹丹¹, 刘海洋¹, 刘 壮², 原继东¹

¹(北京交通大学 计算机与信息技术学院, 北京 100044)

²(北京航空航天大学 计算机学院, 北京 100191)

通信作者: 原继东, E-mail: yuanjd@bjtu.edu.cn



摘 要: 自监督学习可以从原始数据中挖掘自监督信号, 在提高推荐性能方面蕴含着巨大的潜力. 然而, 目前基于自监督学习的推荐方法存在两个关键的挑战. 首先, 大多数自监督推荐模型采用对同一节点随机扰动的方式, 将生成的不同结果作为自监督信号, 然而, 由于推荐系统中存在着广泛的同质性, 这种方式会忽略邻居节点信息, 影响推荐性能. 其次, 用户-物品之间的历史交互信息以及用户与用户之间的社交关系信息是目前基于自监督学习推荐模型关注的焦点, 而忽略了物品之间的内在联系, 同样会导致产生的自监督信号不够充分. 基于这些挑战, 提出一种基于多视角的自监督推荐方法, 分别从偏好视角、用户视角、物品视角考虑, 进而使用多视图共同训练的自监督学习方法, 结合用户之间的社交关系、物品之间的类别关系、用户-物品之间的历史交互信息, 充分挖掘自监督信号. 在 3 个真实的公开数据集上进行实验, 实验结果验证了基于多视角的自监督学习方法在改进推荐性能方面是有效的.

关键词: 自监督学习; 推荐系统; 多视角

中图法分类号: TP18

中文引用格式: 王丹丹, 刘海洋, 刘壮, 原继东. 基于多视角的自监督推荐方法. 软件学报, 2026, 37(2): 684–699. <http://www.jos.org.cn/1000-9825/7419.htm>

英文引用格式: Wang DD, Liu HY, Liu Z, Yuan JD. Self-supervised Recommendation Method Based on Multiple Views. Ruan Jian Xue Bao/Journal of Software, 2026, 37(2): 684–699 (in Chinese). <http://www.jos.org.cn/1000-9825/7419.htm>

Self-supervised Recommendation Method Based on Multiple Views

WANG Dan-Dan¹, LIU Hai-Yang¹, LIU Zhuang², YUAN Ji-Dong¹

¹(School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China)

²(School of Computer Science and Engineering, Beihang University, Beijing 100191, China)

Abstract: Self-supervised learning (SSL) can extract self-supervised signals from raw data, which holds great potential for improving recommendation performance. However, two key challenges remain in current self-supervised learning-based recommendation methods. First, most self-supervised recommendation models apply random perturbations to the same node and use the generated different results as self-supervised signals. However, due to the extensive homogeneity in recommendation systems, this method ignores the information from neighboring nodes, which affects the recommendation performance. Secondly, while historical interaction information between users and items as well as the social relationship information between users are the focus of current self-supervised learning-based recommendation models, the internal relationships between items are often neglected. This also leads to insufficient self-supervised signals. Based on these challenges, a self-supervised recommendation method based on multiple views is proposed. This method considers perspectives from preference, user, and item, and employs a multi-view joint training approach for self-supervised learning. By combining the social relationships between users, the category relationships between items, and the historical interaction information between users and items, self-supervised signals are fully extracted. Experiments conducted on three real public datasets validate that the proposed multi-view-based

* 基金项目: 中央高校基本科研业务费专项资金 (2023JBZY035)

收稿时间: 2023-03-24; 修改时间: 2024-04-07; 采用时间: 2025-02-26; jos 在线出版时间: 2025-11-13

CNKI 网络首发时间: 2025-11-14

self-supervised learning method is effective in improving recommendation performance.

Key words: self-supervised learning; recommender system; multi-view

近年来,随着互联网的快速发展,海量数据开始涌现,进而出现了信息超载问题,推荐系统是解决这一问题最有潜力的方法.它能够根据用户偏好对海量信息进行过滤,为其提供个性化的推荐.然而,用户一般只与自己感兴趣的物品进行交互,导致推荐数据比较稀疏,难以带来理想的推荐效果.为了解决这一问题,常用的方法有两种:一是结合辅助信息,比如社交网络、知识图谱等,来改进推荐性能,提高用户满意度^[1];二是借助于自监督学习^[2],从大量无标签数据中自动产生监督信号,缓解数据稀疏问题.本文主要从自监督学习的角度来研究个性化推荐方法,旨在改善推荐性能.

自监督学习是一种新兴的学习方式,它不需要人工标注标签,可以从原始数据中自动产生监督信号,目前已经广泛应用于众多领域^[3-7],比如计算机视觉、自然语言处理、推荐系统等.在推荐系统中,数据一般是以图的形式存在,比如用户与物品之间的交互二部图、用户之间的社交关系图等,因此,近年来的一些方法致力于利用自监督图表示学习^[8,9]来改进推荐性能,并取得了较好的效果.例如, SGL (self-supervised graph learning)^[10]提出了一种用于推荐任务的自监督图学习框架,将用户与物品之间的历史交互建模为一个二部图,采用随机扰动(随机删除图中的节点/边)的方式来产生监督信号,并结合推荐损失和对比损失来完成推荐任务.基于自监督学习的推荐方法 SGL 模型实际上在用户-物品交互二部图上以自判别的方式来产生自监督信号,即对同一个节点的不同扰动结果(随机删除图中的节点或边)作为对比信号,从而忽略了邻居节点的信息.考虑到删除一些重要节点会大幅度改动图的结构信息, HHGR (hierarchical hypergraph graph recommendation)^[11]提出了基于超图的自监督推荐方法,从粗粒度和细粒度两个层面对节点进行随机删除,以细化用户和群组的表示,缓解数据稀疏问题.然而,这些方法都是以自辨别的方式来产生监督信号,即对同一个节点的不同扰动结果作为对比信号,从而忽略了邻居节点的信息.因此, SEPT (socially-aware self-supervised tri-training)^[12]提出了一种社交感知的邻居判别的自监督推荐方法,它结合社交网络信息并将同质的邻居节点的扰动结果作为监督信号,进一步改进了推荐的性能.但是这种方法仅考虑了用户之间的社交关系,忽略了物品之间的内在联系,导致产生的自监督信号不够充分,无法更好地建模物品的表示.

为了解决以上的问题,我们提出了一种基于多视角的自监督推荐方法,包括偏好视角、用户视角和物品视角.其中,偏好视角旨在使用 BPR (Bayesian personalized ranking)^[13]损失不断迭代更新用户和物品表示,用户视角和物品视角用于产生自监督信号,结合推荐任务和自监督任务,采用多任务联合训练的方式获得用户和物品的表示.相较于同为自监督推荐算法的 SGL 模型,本文提出的方法在生成自监督信号时,既考虑了自身信息也考虑了邻居信息,并且融入了社交关系信息以及物品的类别信息,从而获得的自监督信号更加丰富.为了验证所提出方法的有效性,分别在 3 个公开数据集上进行了实验,实验结果表明,我们提出的方法优于对比的基线方法,并通过消融实验验证了模型中每个模块的必要性.

总的来说,本文的主要贡献如下.

- (1) 使用物品的类别信息构建物品联系图,在物品视角下结合对比学习方法更新用户和物品的表示.
- (2) 提出了一种基于偏好视角、用户视角和物品视角的多视角推荐方法,并且采用多任务联合训练的方式完成推荐任务.
- (3) 在 3 个公开数据集上进行了充分的实验,结果表明所提出的方法优于对比的若干基线方法,提高了模型推荐性能.

1 相关工作

本节主要介绍与我们方法密切相关的工作,主要包括:基于图神经网络的推荐方法、基于自监督学习的推荐方法和融合社交关系的推荐方法.

1.1 基于图神经网络的推荐方法

在推荐系统中,大多数信息具有图结构.例如,用户之间的社交网络、物品之间的知识图谱以及用户和物品之

间的交互关系都可以用图结构来表示,因此,基于图的推荐系统应运而生.目前,图神经网络(graph neural network, GNN)^[14]是最常使用的方法之一,图神经网络能够通过图节点之间的消息传递来捕获图的依赖性,它的主要思想是从邻居节点中迭代地聚合特征信息,并在传播过程中将聚合的信息与当前的中心节点表示相结合.

图神经网络因其在求解图相关推荐任务方面的有效性受到了广泛的关注,尤其是图卷积网络(graph convolutional network, GCN)^[15].它是谱图卷积的一阶近似,许多图神经网络模型都是由它衍生而来的.比如, GCMC (graph convolutional matrix completion)^[16]模型是一种对用户-物品交互矩阵进行补全的自编码框架,其利用用户和物品的交互信息生成用户和物品之间的隐特征,并通过双线性解码器来重建用户和物品之间的连接.但是 GCMC 只考虑一阶邻居信息,不能够很好地利用信息传递来挖掘图结构信息.为了解决此问题, STAR-GCN (stacked and reconstructed graph convolutional network)^[17]模型被提出,其提出了重构机制,旨在从聚合的表示中恢复原始输入节点的表示,并且将恢复得到的表示作为下一个块的输入,并在训练过程中加入了一个基于掩盖节点的重构损失,整个损失函数由预测损失和重构损失两部分组成.但是,无论 GCMC 还是 STAR-GCN,都平等地对待邻居节点的影响,并且其传播机制仅仅依赖于邻居节点,它们都使用 GCN 最后一层的输出作为节点表示来进行评分预测. NGCF (neural graph collaborative filtering)^[18]模型解决了以上方法的缺陷,它设计的传播机制不仅依赖于图结构还加入了图结构与中心节点的亲和度.另外,它融合 GCN 不同层的表示来得到最终的节点表示,尽管 NGCF 模型已经表现出很好的效果,但是它的设计笨重,为了简化模型,轻量级的 LightGCN^[19]模型被提出,它删除了 GCN 模型中的特征变换和非线性激活这两种操作,只保留 GCN 中最关键的操作-邻居聚合.研究发现,在删除它们之后可以显著提高推荐性能.以上基于 GNN 的推荐方法是在用户-物品二部图中进行卷积操作,另外, GNN 也可以用于其他特定图上,比如基于会话图的推荐,常见的方法有 SR-GNN^[20]模型;基于社交网络图的推荐,常见的方法有 DiffNet++^[21]、MHCN (multi-channel hypergraph convolutional network)^[22].值得一提的是, GNN 经常用于社交计算,因为在社交网络中传播的信息可以被其传递的信息很好地捕获.这就是为什么我们求助于社交网络来获取由图神经编码器产生的自我监督信号的原因.

1.2 基于自监督学习的推荐方法

自监督学习作为一种新的学习模式,它可以减少对数据标签的依赖并且可以基于大量无标签数据进行训练.自监督学习的基本思想是可以从大量无标签数据中提取信息并应用到不同的机器学习任务中.正是由于可以处理普遍存在的标签不足的问题,自监督学习已经被广泛地应用于很多领域,包括视觉表示学习、节点/图分类、推荐系统等.自监督学习这种灵活的学习范式,将其应用于图表示学习是目前研究的热点.基于图的自监督模型,主要利用图结构来挖掘自监督信号,这一研究方向的主要机制是图对比学习,它对比了同一图的不同视图,其中,局部视图是通过对原始图进行数据增强而建立,实现方式就是对原始图随机删除边或节点扰动方式生成.

受到图对比学习成功的启发,这种想法也逐渐应用于推荐领域中.比如 SGL 是基于对比学习的自监督推荐模型,首先通过随机删除节点/删除边方式进行数据增强,生成两种不同数据视图,在增强视图上应用共享的 LightGCN 编码器学习节点表示.在负节点采样方面,节点级对比通过优化 InfoNCE^[21]损失函数来生成, BPR 损失函数来做推荐.但是考虑到删除一些重要节点会对局部结构造成高度倾斜, HHGR 提出了双尺度节点 dropout 方法在分组的推荐场景中生成有效的自监督信号,提出的粗粒度的 dropout 方案一定比例地从所有组中删除了用户节点,然而细粒度的删除节点方案只是随机地从精确的分组中删除选中的节点.这种方法通过不同粒度的删除,最大化用户节点之间的互信息.尽管超图在用户建模方面比简单图具有很大的优势,但它在社会推荐方面的优势尚未得到充分的体现,因此通过研究融合超图建模和图卷积网络的潜力来填补这一空白,并提出了一种多通道超图卷积网络 MHCN,它定义了 3 种类型的三角社交关系并且将它们建模为多通道超图编码器,利用高阶用户关系来增强社会推荐.这种多通道和聚合多用户表示学习的方式,可以获得全面的用户表示,被认为是包含多类型高阶关系信息并且使用 item 嵌入来生成更好的推荐结果.以上基于自监督学习的推荐模型都是通过自识别方式来挖掘用户之间的内在联系,从而忽略了有用的邻居节点信息,而 SEPT 采用了多视图共同训练的方式生成自监督信号,它结合社交网络信息并将同质的邻居节点的扰动结果作为监督信号,这种融合社交感知的多视图共同训练的自监

督推荐模型更好地提高了推荐性能, 这一模型被称为自监督图对比学习中的 SOTA 模型. 因此我们模型多视角中的用户视角、物品视角都是通过这种多视图共同训练的方式生成自监督信号.

1.3 融合社交关系的推荐方法

近年来, 图神经网络可以自然地整合节点信息和拓扑结构, 因此其在关于图数据的学习中具有强大的能力. GNN 的这些优点为推进融合社交关系的推荐方法提供了巨大的潜力, 因为社交推荐系统中的数据可以表示为社交关系图 and 用户-物品交互图, 并且从图中学习用户和物品的表示是关键.

推荐算法的核心是学习用户和物品的表示, 通过对用户和物品的表示进行比较, 从而可以对相似的用户或者物品进行推荐. 目前, 协同过滤 (collaborative filtering, CF) 算法^[23]是比较流行的一种算法. CF 算法本质上是对用户和物品的互动关系进行分析, 从而分析出用户和物品之间的偏好关系并完成推荐. 但是 CF 算法并没有较好地解决数据稀疏性问题, 如果对用户-物品二部图邻接矩阵进行计算, 很容易导致过拟合, 这种情况下就会导致推荐效果不理想. 由于数据稀疏性问题, 社交推荐算法应运而生, 社交推荐算法主要是通过用户-用户之间的关系进行分析, 将关系比较亲近的用户之间的偏好关系进行建模, 从而加强了推荐算法的性能并解决了数据稀疏性问题. DiffNet++ 网络对社交网络中的用户之间的关系进行迭代计算, 将用户的偏好影响力进行了 k 轮传播, 最终形成了用户之间互相影响的偏好传播. 由于 DiffNet++ 模型中存在大量的冗余参数操作, 导致推荐性能受限. MHCN 是将社交关系建模为超图, 它的性能要优于 DiffNet++, 这说明超图的表达能力要强于简单图. 此外, 引入社交关系的实验效果要优于单纯的协同过滤方法, 说明了融入社交网络对推荐性能的改善有正向作用.

总体来说, 目前结合社交网络的自监督推荐方法主要存在两个方面的不足: 第一, 大多数方法只考虑用户层面的监督信号, 而忽略了物品层面的信息, 比如物品之间的类别关系等, 进而导致自监督信号挖掘不充分, 推荐性能不尽人意; 第二, 目前的方法将同一个节点的不同扰动结果作为自监督信号, 从而忽略了邻居节点的信息. 为了解决以上的问题, 我们提出融合基于多视角的自监督推荐方法, 它不仅融合了用户层面的社交信息, 而且考虑了物品之间的类别关系, 并结合邻居信息产生自监督信号, 进一步改进了推荐效果.

2 模型框架

本节介绍所提出模型的具体细节. 首先给出本文使用到的数学符号和问题定义, 然后对所提出的模型进行整体概览, 接着详细介绍模型中每个模块的细节, 包括偏好视角、用户视角和物品视角, 最后介绍模型的训练方法.

2.1 问题定义

为了方便理解, 我们首先介绍本文使用到的数学符号. 假设 $U = u_1, u_2, u_3, \dots, u_m$ 表示所有的用户集合, $I = i_1, i_2, i_3, \dots, i_n$ 表示所有的物品集合, m 和 n 分别表示用户数量和物品数量. $\mathbf{P} \in \mathbb{R}^{m \times d}$ 和 $\mathbf{Q} \in \mathbb{R}^{n \times d}$ 分别表示用户和物品的嵌入表示矩阵, d 为嵌入向量的维度. 用户和物品之间的交互关系图定义为 G_r , 它是一个二部图, 对应的交互矩阵为 $\mathbf{R} \in \{0, 1\}^{m \times n}$, 如果用户 u 与物品 i 交互过, $r_{ui} = 1$, 否则 $r_{ui} = 0$. 用户之间的社交网络图定义为 G_u , 对应的社交矩阵为 $\mathbf{S} \in \{0, 1\}^{m \times m}$, $s_{uv} = 1$ 表示用户 u 与用户 v 为朋友关系. 物品之间的联系图定义为 G_i , 这里我们将属于同一类别的物品之间建立一条边, 其对应的邻接矩阵为 $\mathbf{T} \in \{0, 1\}^{n \times n}$, $t_{ij} = 1$ 表示物品 i 和物品 j 属于同一类别. 表 1 总结了本文用到的数学符号.

本文提出了基于多视角的自监督推荐方法, 并且在用户视角以及物品视角下使用多视图共同训练的方式来生成自监督信号, 这里详细阐述视角及视图的定义. 视角是一个宏观层面的概念, 是对同一个问题不同角度的表达, 比如偏好视角旨在根据用户-物品二部图从显式角度挖掘用户偏好, 用户视角和物品视角是从用户和物品的角度分别生成不同的自监督信号, 以更好地得到用户、物品表示, 它们是从隐式角度来挖掘用户偏好. 视图是一个微观层面的概念, 它指的是根据图中节点之间的数据关系从原始图中生成的不同子图, 例如用户共享视图是结合用户之间的社交网络图以及用户-物品二部图而得到的, 朋友关系视图是从用户之间的社交网络图中考虑可靠关系得到的, 它们都是原始图的一个子图.

表 1 本文用到的数学符号

符号	描述	符号	描述
G_r	用户-物品二部图	$\mathbf{P}$	用户嵌入表示矩阵
G_s	社交网络图	$\mathbf{Q}$	物品嵌入表示矩阵
G_t	物品联系图	r_{ui}	用户 u 对物品 i 的评分
$\mathbf{R}$	用户-物品交互矩阵	$\tilde{G}^u$	用户视角下用户-物品二部图联合社交网络图的扰动子图
$\mathbf{S}$	社交矩阵	$\tilde{G}^i$	物品视角下用户-物品二部图联合物品联系图的扰动子图
$\mathbf{T}$	物品关系矩阵	$\tilde{z}_u$	用户视角下包含用户表示的无标签数据集
$\mathbf{U}$	用户集合	$\tilde{z}_i$	物品视角下包含物品表示的无标签数据集
$\mathbf{I}$	物品集合	—	—

基于以上的定义, 我们的任务是: 首先根据物品的所属类别信息构建物品之间的联系图 G_t , 然后结合用户-物品交互图 G_r , 以及用户之间的社交网络图 G_s , 在训练集上通过模型训练, 学习用户和物品的向量表示 $\mathbf{P}$ 和 $\mathbf{Q}$, 通过 $\mathbf{P}$ 和 $\mathbf{Q}$ 预测用户 u 对物品 i 的评分 $\hat{r}_{ui}$, 根据预测评分为用户推荐可能感兴趣的物品.

2.2 模型概览

我们的模型主要分为两大任务, 分别为推荐任务和对比学习任务, 模型框架如图 1 所示. 推荐任务作为模型的主任务, 输入为用户-物品二部图, 利用用户偏好信息初步生成用户和物品的向量表示; 对比学习任务又分为用户视角对比学习和物品视角对比学习, 输入分别为社交网络图和物品类别图, 用于辅助更新推荐任务中学习到的用户和物品的向量表示.

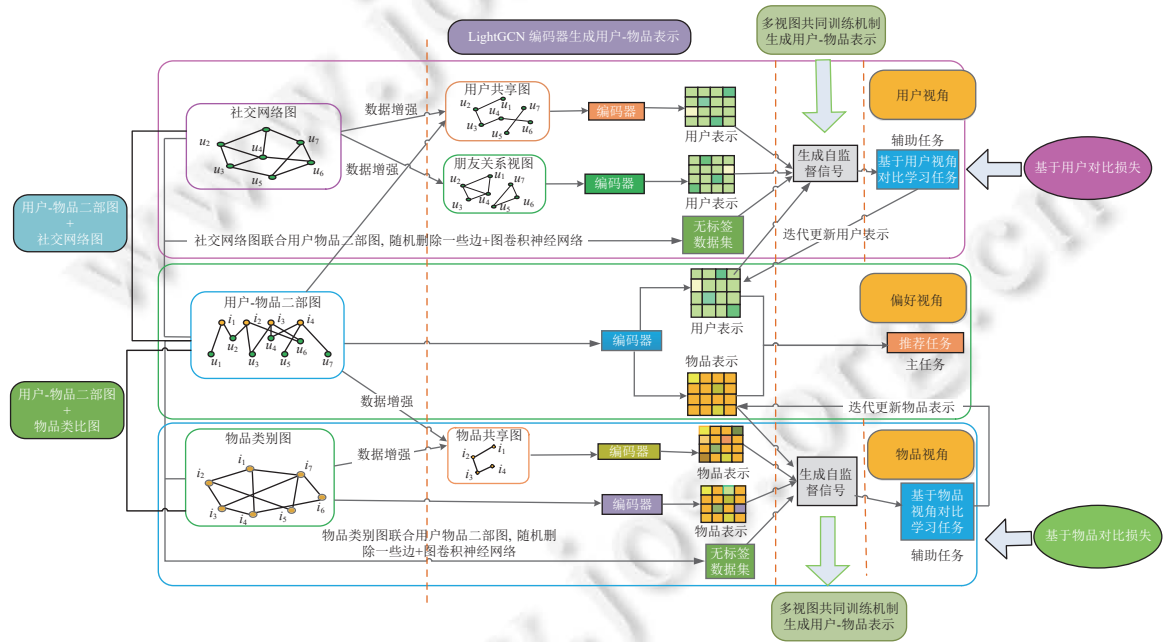


图 1 整体模型框架

2.3 偏好视角

在偏好视角下, 主要目标是根据用户与物品的历史交互数据, 学习用户、物品表示矩阵. 具体来说, 我们将用户-物品之间的交互数据建模为二部图 G_r , 然后使用 LightGCN 编码器编码用户和物品信息, 获得用户和物品的表示, 最后将用户 u 和物品 i 的向量表示做内积得到预测的推荐得分, 即 $\hat{r}_{ui} = \mathbf{P}_u \times \mathbf{Q}_i$. 在偏好视角下, 我们通过降低 BPR 损失, 不断迭代更新用户和物品的表示 $\mathbf{P}$ 和 $\mathbf{Q}$. BPR 损失的基本思想是与用户 u 交互过的物品的得分大于未交互

过的物品的得分, 即使得正负样本之间的预测得分的差值尽可能大, 用公式表示为:

$$\mathcal{L}_r = \sum_{i \in I(u), j \notin I(u)} -\log \delta(\hat{r}_{ui} - \hat{r}_{uj}) + \eta (\|P\|_2^2 + \|Q\|_2^2) \quad (1)$$

其中, i 表示与用户 u 交互过的物品, j 表示与用户 u 未交互过的物品, $\hat{r}_{ui}$ 、 $\hat{r}_{uj}$ 分别表示用户 u 对物品 i 、物品 j 的预测得分, $I(u)$ 表示与用户 u 交互过的物品集合, η 表示正则化系数.

2.4 物品视角

在偏好视角下, 我们初步学习到物品表示, 然而数据稀疏性一直是推荐系统中存在的问题, 而自监督学习正是解决这一问题很好的方式, 它在很多模型中得到了验证. 自监督学习可以从原始数据中自动生成自监督信号, 例如, 基于图对比学习^[9]的方法是通过随机删除原始图中的节点/边, 创建增强视图, 然后最大化同一节点在不同视图下学习到的向量表示之间的一致性. 推荐系统中存在着大量的图结构数据, 但是, 与其他基于图的任务相比, 推荐系统中用户和物品之间有着广泛的同质性^[24], 现有的基于自监督学习的推荐方法, 都是通过自识别的方式生成正负样本, 由于推荐系统中的同质性, 有些节点是假负样本, 如果把这些假负样本识别为正样本, 实际上可以学习到更好的表示. 相反, 如果把它们粗略地归结为负样本会导致推荐性能下降.

为了解决这一问题, 我们提出多视角共同训练与自监督学习相结合的推荐方法, 物品视角下的模型框架如图 1 物品视角标注所示. 考虑到物品之间的同质性, 我们将属于同一类别的物品之间建立一条边, 形成物品类别图 G_r . 物品视角下的数据源包括用户-物品二部图 G_r 、物品类别图 G_i . 从用户-物品交互图中, 我们分析到同一用户交互过不同的物品, 并且这些物品在物品类别图中又属于同一类别, 那么根据用户-物品偏好关系以及物品类别关系, 可以从物品的角度, 给用户推荐其喜爱的物品, 从而提高推荐性能.

基于物品视角, 我们使用多视图共同训练的方式生成自监督信号. 首先根据物品之间的类别关系进行数据增强, 得到增强视图, 即模型框架图 1 中标注的物品共享图, 它是根据用户-物品二部图和物品类别图而得到的, 将属于同一类别并且与相同用户交互过的物品之间建立一条边, 是物品类别图的一个子图. 物品共享图承载了物品类别信息以及用户-物品交互信息, 它表示一个用户购买过的同一种类别的物品. 物品共享图对应的邻接矩阵由如下公式得到:

$$A_i^s = (RR^T) \odot T \quad (2)$$

其中, $T \in \mathbb{R}^{m \times n}$ 代表物品类别图的邻接矩阵, $R \in \mathbb{R}^{m \times n}$ 代表用户-物品交互二部图的邻接矩阵, $\odot T$ 确保 A_i^s 中的关系是 T 中关系的子集. 从图 1 中可以看出, 在物品视角下, 有 3 个视图从不同的角度对物品信息进行描述, 即物品类别图、物品共享图和用户-物品交互图, 我们将利用它们来挖掘具有多视图编码的自监督信号.

首先, 我们在 3 个视图上构建了 3 个非对称编码器, 其中两个仅用于学习物品表示并给出伪标签, 另一个针对用户-物品交互视图完成推荐任务. 这里非对称编码器是指对于构建的每一个视图, 分别使用 LightGCN 编码器进行编码, 得到对应视图下的节点的表示, 由于每个视图的连接关系不同, 因此每个视图上的编码器的参数也不同, 从而使得所得到的节点表示也不同. 这里使用的是多视图共同训练方式, 需要将用户-物品交互图和物品类别图组合成一个新的图, 以概率 p 动态删除原始图中的边, 动态生成无标签数据集, 这个过程可以描述为:

$$\tilde{G}^i = (\mathcal{N}_r \cup \mathcal{N}_i, m \odot (\mathcal{E}_r \cup \mathcal{E}_i)) \quad (3)$$

其中, $\mathcal{N}_r$ 、 $\mathcal{N}_i$ 和 $\mathcal{E}_r$ 、 $\mathcal{E}_i$ 分别为 G_r 、 G_i 中的节点和边; $m \in \{0, 1\}^{|\mathcal{E}_r \cup \mathcal{E}_i|}$, m 为随机删除节点/边的概率. 我们同时扰动物品类别图和用户-物品交互图, 因为对物品的影响包含在上述两个增强视图中.

在以上增强视图和无标签数据集的基础上, 我们使用 LightGCN 编码器对其进行编码, 其定义如下:

$$Z = H(Q, \mathcal{V}) \quad (4)$$

其中, H 为编码器, $Z \in \mathbb{R}^{n \times d}$ 为节点的最终表示, 相同大小的 Q 为 3 个编码器共享的初始节点嵌入, $\mathcal{V} \in \{R, A_i^s, G_i\}$ 为 3 个视图中的任何一个. 通过在 3 个视图上进行图卷积, 编码器学习到了 3 组物品的表示, 由于每个视图都反映了物品的不同方面, 因此从其他两个视图中寻求监督信息, 以改进当前视图的编码器, 给定一个物品, 我们使用来自其他两个视图的物品表示来预测它在无标签数据集中的语义上的正样本. 以偏好视图中的物品 i 为例子, 表述为:

$$y_{i^+}^i = \text{Softmax}(\phi(\tilde{z}_i, z_i^i)), y_{i^+}^s = \text{Softmax}(\phi(\tilde{z}_i, z_i^s)) \quad (5)$$

其中, ϕ 为余弦相似度函数, z_i^i 和 z_i^s 分别为 H_i 和 H_s 学习到的物品表示, 在这里为了更加明确地区分不同视图的编码器, 在物品视角下, 我们将 H_i 表示物品类别视图的编码器, H_s 表示物品共享图的编码器, H_r 定义为用户-物品二部图的编码器. $\tilde{z}_i$ 为通过图卷积得到的无标签样本集中物品的表示, $y_{i^+}^i$ 和 $y_{i^+}^s$ 表示每个物品在相应视图中为物品 i 的语义正样本的预测概率. 为了避免噪声样本, 只有 H_i 和 H_s 都统一将物品 i 标记为正样本, 才能在 H_r 将物品 i 进行标记. 我们将在两个视图上的预测概率取均值, 得到:

$$y_{i^+}^r = \frac{1}{2}(y_{i^+}^i + y_{i^+}^s) \quad (6)$$

选择前 k 个概率最大的正样本, 此过程可表述为:

$$\mathcal{P}_{i^+}^r = \tilde{z}_{ik} | k \in \text{Top-K}(y_{i^+}^r), \tilde{z}_i \sim \tilde{\mathcal{G}}^i \quad (7)$$

在每次迭代中, 通过随机扰动, 不断对 $\tilde{\mathcal{G}}^i$ 进行扰动, 经过图编码器学习后生成不同的物品表示. 通过生成的伪标签, 基于邻居识别对比学习方法生成自监督信号, 对于给定某个物品 i , 最大化它的节点表示和 $\mathcal{P}_{i^+}^r$ 中的有标签物品表示之间的一致性, 并最小化它的节点表示和无标签物品表示之间的一致性. 我们采用在互信息估计中有效的 InfoNCE 损失函数来计算物品视角下的对比损失, 以最大化正样本对之间的一致性, 最小化负样本对之间的相似性:

$$\mathcal{L}_{\text{ssl}}^i = -\mathbb{E} \sum_{v \in [r, s, f]} \left[\log \frac{\sum_{p \in \mathcal{P}_{i^+}^r} \varphi(z_i^v, \tilde{z}_p)}{\sum_{p \in \mathcal{P}_{i^+}^r} \varphi(z_i^v, \tilde{z}_p) + \sum_{j \in U/\mathcal{P}_{i^+}^r} \varphi(z_i^v, \tilde{z}_j)} \right] \quad (8)$$

其中, $\mathbb{E}$ 代表数学期望. $\varphi(z_i^v, \tilde{z}_p) = \exp(\phi(z_i^v \times \tilde{z}_p) / \tau)$, 这是一个判别器函数. 输入两个向量, z_i^v 表示物品 i 在任意一个视图下学习到的物品表示, $\tilde{z}_p$ 表示在多视图训练下, 学习到的物品正样本的表示, $\tilde{z}_j$ 是经过编码器学习后得到的物品负样本的表示, 判别器函数可以预估它们之间的一致性. τ 在损失函数中表示温度参数, 损失函数的一个重要的性质是困难负样本的发现, 这一性质在自监督表示学习中发挥重要的作用. 困难负样本可以提供更大的梯度值, 能够有效促使损失函数向最优方向推进, 而温度参数 τ 在其中就起着调节对困难样本的关注程度的作用, 它控制模型对负样本的区分度.

2.5 用户视角

基于对比学习任务, 我们详细介绍了物品视角下如何更新物品表示矩阵 $\mathbf{Q}$, 下面介绍在用户视角下如何更新用户矩阵 $\mathbf{P}$. 用户视角下的模型框架如图一用户视角标注所示. 用户-物品交互图、社交网络图是从两个不同的角度反映了用户偏好的数据源, 用户-物品交互图表示在历史数据中直接反映用户对物品的偏好, 社交网络图从另一个隐式角度中捕捉用户之间的社交关系, 反映了具有朋友关系的用户可能具有相似的偏好, 根据用户之间的同质性, 我们从 G_r 、 G_s 交互中的三元结构来得到两个增强数据视图, 分别为任意两个用户之间都是朋友关系的视图和共同购买过同一件物品的朋友的视图, 称为可靠朋友关系视图和用户共享图. 朋友关系视图和物品共享图的邻接矩阵由如下公式得到:

$$\mathbf{A}_u^f = (\mathbf{SS}) \odot \mathbf{S}, \mathbf{A}_u^s = (\mathbf{RR}^T) \odot \mathbf{S} \quad (9)$$

其中, $(\mathbf{SS})$ 和 $(\mathbf{RR}^T)$ 计算连接两个用户的路径, 而 $\odot \mathbf{S}$ 使得用户与用户之间形成三角关系并确保 $\mathbf{A}_u^f$ 和 $\mathbf{A}_u^s$ 中的关系是社交网络图中关系的子图. 同样, 我们使用物品视角下多视图共同训练的方法, 得到前 k 个概率最大的正样本. 然后, 采用 InfoNCE 损失函数来计算用户视角下的对比损失, 以最大化正样本对之间的一致性, 最小化负样本对之间的相似性:

$$\mathcal{L}_{\text{ssl}}^u = -\mathbb{E} \sum_{v \in [r, s, f]} \left[\log \frac{\sum_{p \in \mathcal{P}_{u^+}^r} \varphi(z_u^v, \tilde{z}_p)}{\sum_{p \in \mathcal{P}_{u^+}^r} \varphi(z_u^v, \tilde{z}_p) + \sum_{j \in U/\mathcal{P}_{u^+}^r} \varphi(z_u^v, \tilde{z}_j)} \right] \quad (10)$$

其中, z_u^i 表示用户 u 在任意一个视图下学习到的用户表示, z_p 表示在多视图训练下, 学习到的用户正样本, 而 z_j 是经过编码器学习后得到的用户负样本.

2.6 模型训练

本文模型的训练方式采用多任务联合训练, 其中包括推荐任务、对比学习任务, 模型的损失函数包括 3 部分, 推荐损失 ($\mathcal{L}_r$)、用户视角对比损失 ($\mathcal{L}_{ssl}^u$)、物品视角对比损失 ($\mathcal{L}_{ssl}^i$), 整体的损失函数定义如下:

$$\mathcal{L} = \mathcal{L}_r + \lambda(\mathcal{L}_{ssl}^u + \mathcal{L}_{ssl}^i) \quad (11)$$

其中, λ 为对比损失函数权重, 后面会详细介绍其取值对整体模型性能的影响. 模型训练过程中, $\mathbf{P}$ 和 $\mathbf{Q}$ 的更新过程执行如下公式:

$$\mathbf{P} = \mathbf{P} - \mathcal{L}_r \times \frac{\partial \mathcal{L}}{\partial \mathbf{P}}, \mathbf{Q} = \mathbf{Q} - \mathcal{L}_r \times \frac{\partial \mathcal{L}}{\partial \mathbf{Q}} \quad (12)$$

其中, $\mathcal{L}_r$ 为学习率.

模型训练具体算法步骤如算法 1 所示.

算法 1. 基于多视角自监督推荐方法模型框架.

输入: 用户集合 U 、物品集合 I 、用户-物品历史交互数据、用户社交关系以及物品类别信息;

输出: 推荐列表.

- 1) 构建用户-物品二部图 G_r 、社交关系图 G_s 以及物品类别图 G_i
 - 2) 随机初始化用户与物品的表示 $\mathbf{P}$ 和 $\mathbf{Q}$
 - 3) $\mathcal{L}_r = \sum_{i \in I(u), j \notin I(u)} -\log \delta(\hat{r}_{ui} - \hat{r}_{uj}) + \eta(\|\mathbf{P}\|_2^2 + \|\mathbf{Q}\|_2^2)$ //执行 BPR 损失函数进行模型预训练
 - 4) $\mathbf{A}_s = (\mathbf{R}\mathbf{R}^T) \odot \mathbf{T}$, $\mathbf{A}_f = (\mathbf{S}\mathbf{S}) \odot \mathbf{S}$, $\mathbf{A}_s = (\mathbf{R}\mathbf{R}^T) \odot \mathbf{S}$ //基于 G_r 、 G_s 、 G_i 进行数据增强, 得到增强视图
 - 5) for $epoch$ do //迭代次数循环
 - 6) for $\mathcal{N}_r \cup \mathcal{N}_i$, $m \in (\mathcal{E}_r \cup \mathcal{E}_i)$ //构建无标签数据集, m 为随机删除节点/边的概率
 - 7) $\tilde{\mathbf{Z}} = H(\mathbf{Q}, \mathcal{V})$ //LightGCN 编码器学习用户、物品表示 $\tilde{\mathbf{Z}}$
 - 8) for (u, i) in (U, I) do //多视图共同训练
 - 9) $y_+^r = \text{Softmax}(\phi(\tilde{z}, z_i^r))$, $y_+^s = \text{Softmax}(\phi(\tilde{z}, z_i^s))$
 - 10) $y_+^f = \frac{1}{2}(y_+^r + y_+^s)$
 - 11) $\mathcal{P}_+^r = \{\tilde{z}_k | k \in \text{Top-K}(y_+^r), \tilde{z} \sim \tilde{\mathcal{G}}\}$
 - 12) endfor
 - 13) $\mathcal{L} = \mathcal{L}_r + \lambda(\mathcal{L}_{ssl}^u + \mathcal{L}_{ssl}^i)$ //将推荐任务、对比学习任务联合训练
 - 14) $\mathbf{P} = \mathbf{P} - \mathcal{L}_r \times \frac{\partial \mathcal{L}}{\partial \mathbf{P}}$, $\mathbf{Q} = \mathbf{Q} - \mathcal{L}_r \times \frac{\partial \mathcal{L}}{\partial \mathbf{Q}}$ //更新 $\mathbf{P}$ 、 $\mathbf{Q}$
 - 15) endfor
 - 16) endfor
-

3 实验分析

为了验证所提出方法的有效性, 我们在 3 个公开数据集上进行实验. 具体来说, 旨在回答以下 3 个研究问题 (research question, RQ).

RQ1: 我们的方法与其他基线方法 (例如, 基于图神经网络的 LightGCN^[19]、基于社交网络的自监督推荐模型 SEPT^[12]等) 相比, 效果如何?

RQ2: 所提出的方法中每个模块 (例如, 社交网络图、物品联系图等) 对整体的推荐性能有何影响?

RQ3: 模型中的重要超参数 (例如, 损失函数权重 λ 、温度参数 τ) 对推荐性能的灵敏度如何?

3.1 数据集

我们在 3 个公开数据集上进行实验, 包括 Yelp^[18,19,25]、Ciao^[26]、Epinions^[25,27,28]. 表 2 给出了数据集所对应的详细信息.

表 2 数据集统计信息

数据集	Yelp	Epinions	Ciao
用户数量	43 043	18 081	7 317
物品数量	66 576	251 722	104 975
用户/物品交互数量	283 512	751 821	283 320
用户/物品交互密度 (%)	0.009 8	0.015 7	0.036 9
社交关系数量	549 451	590 641	111 781
社交关系密度 (%)	0.029 6	0.180 6	0.208 8
物品所属类别数量	1 336	28	28

Yelp 数据集是从 Yelp 平台上搜集而来, 它是美国最大的点评网站 (<https://www.yelp.com/>). 它包含了社交关系信息和物品所属的类别信息, 并且具有社交关系的用户都具有相同的兴趣.

Epinions 数据集主要记录了来自 Epinions 网站 (<http://www.epinions.com/>) 上用户对不同物品的反馈信息. Epinions 是一家知名的大众消费者点评网站, 成立于 1999 年. 在 Epinions 数据集中, 用户可以阅读有关各种物品的评论, 以帮助他们决定购买行为或者他们可以免费加入并开始撰写评论, 从而获得回报和认可. 该数据集中的社交网络来自朋友关系, 也包含物品的类别信息.

Ciao 数据集主要记录了来自电商平台网站 (<https://www.librec.net/datasets.html>) 上用户对他们购买的物品的评分、用户与用户之间的社交关系以及用户购买的物品之间的类别信息.

实验中选取 Yelp、Ciao、Epinions 这 3 个公开数据集来验证本文提出的基于多视角的自监督推荐模型, 这 3 个数据集分别属于不同规模, 不同领域并且都包含用户对物品的反馈评分, 用户之间的社交关系数据及物品所属类别数据. 为了进行精确的评估, 对所有实验都进行了 5 折交叉验证, 并展示其平均结果.

3.2 基线模型简介

本节将提出的多视角自监督推荐模型与其他 6 个近年来提出的推荐算法进行了比较, 以检验提出的模型的有效性. 近期提出的基于自监督学习的 HHGR 模型是基于超图的自监督推荐方法, 不同于本文构建的简单图, 因此未选取 HHGR 作为基线模型, 本文选取的基线模型分别为 LightGCN^[19]、NeuMF^[29]、BUIR^[30]、MHCN^[22]、DiffNet++^[21]、SGL^[10]、SimGCL (simple graph contrastive learning for recommendation)^[31]、LightGCL (simple yet effective graph contrastive learning for recommendation)^[32]、SEPT^[12].

(1) BUIR 是一个单分类协同过滤模型. 它在训练过程中不需要负样本, 只依靠正样本来学习用户和物品的表示. 它包括两个编码器: 在线编码器和目标编码器. 在线编码器用于预测用户和物品之间的交互, 它的预测目标由目标编码器提供, 二者交替进行.

(2) NeuMF 是目前很流行的基于深度学习的推荐算法模型. 它将传统矩阵分解方法和多层感知机联合, 并且同时提取低维和高维特征, 在推荐质量方面得到了很大的提升.

(3) LightGCN 是基于图卷积神经网络的通用推荐模型. 它利用用户与物品之间的相似性来学习节点表示并完成推荐任务. 该推荐模型简单有效.

(4) DiffNet++是近期提出的一种基于 GCN 的社交推荐方法. 它融合了社交网络, 并利用用户和物品在隐空间中的动态扩散机制来完成推荐任务.

(5) MHCN 是一种最新的基于超图卷积神经网络的社交推荐方法. 它利用超边来建模用户之间的复杂相关性, 以提高推荐性能.

(6) SGL 是一种基于 GCN 的自监督推荐方法. 它在用户-物品交互图上以自判别的方式来产生自监督信号, 即对同一个节点的不同扰动结果 (随机删除图中的节点或边) 作为对比信号.

(7) SimGCL 是一种在原始表示中添加随机均匀噪声来进行数据增强的方法. 它放弃了图增强, 可以平滑地调节均匀性, 并且实现起来比较容易.

(8) LightGCL 是使用奇异值分解方法来进行数据增强, 用于提取用户-物品交互图中的全局协同关系, 可以更好地保留图的语义结构, 从而提高推荐性能.

(9) SEPT 是近期提出的基于自监督的社交推荐模型, 模型算法加入了用户之间的社交关系, 并且使用自监督+Tri-training 框架来实现推荐流程.

3.3 评价指标

为了直观地评价多视角推荐模型与其他模型算法的优势, 实验将使用以下指标进行衡量.

(1) *Precision*: 准确率是准确推荐的物品与推荐列表中的物品总数的比率, 衡量的是推荐系统的查准率. *Precision@10* 是指推荐列表中预测得分最高的前 10 个物品的准确率.

$$Precision = \frac{\sum_{u \in U} |P_u \cap T_u|}{\sum_{u \in U} P_u} \quad (13)$$

其中, u 表示用户, U 表示用户集合, P_u 表示所有被推荐的物品, T_u 表示测试集中与用户 u 交互过的物品.

(2) *Recall*: 准确推荐的物品和测试集中与用户交互过的所有物品的比率, 衡量的是推荐系统的查全率.

$$Recall = \frac{\sum_{u \in U} |P_u \cap T_u|}{\sum_{u \in U} |T_u|} \quad (14)$$

(3) *F1*: 该评价指标是 *Precision* 和 *Recall* 的加权平均值.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (15)$$

(4) *NDCG* (normalized discounted cumulative gain): 归一化折损累计增益, 它考虑到位置因素, 能够对一个用户的推荐列表进行评价, 用公式表示为:

$$NDCG = \frac{DCG}{IDCG} \quad (16)$$

其中, $DCG = DCG_k = \sum_{i=1}^k \frac{2^{rel_i} - 1}{\log_2(i+1)}$, $IDCG = \sum_{i=1}^{|REL|} \frac{2^{rel_i} - 1}{\log_2(i+1)}$, rel_i 表示位置 i 的推荐结果的相关性, k 表示推荐列表的大小.

3.4 实验参数设置

为了进行公平的比较, 我们参考在基线模型原始论文中设置的最佳参数, 然后对其进行了微调, 以确保达到最佳性能. 本文提出的模型, 在训练过程中我们将向量维度设置为 50, 正则化参数 η 设置为 0.001, 损失函数权重系数 λ 设置为 0.002, 批处理大小设置为 2000, 迭代次数 *epoch* 设置为 50, 在第 3.7 节我们将研究该模型的参数灵敏度, 并给出最佳参数设置. 使用 Adam 优化器来优化所有模型, 初始学习率为 0.001, 温度参数为 0.1.

3.5 基线模型对比实验 (RQ1)

为了验证所提出方法的有效性, 我们将该模型与所考虑的 6 种基线模型进行了对比, 实验结果如表 3 和表 4 所示. 其中, 下划线表示基线模型中效果最好的结果, 最后一行给出了我们提出的方法与目前最好的基线模型的结果的相对提升比例, $\uparrow$ 代表相对提升. 根据实验结果, 我们有如下的发现.

(1) BUIR 和 NeuMF 模型的推荐性能明显低于其他基线模型, 这是由于这两种模型都属于协同过滤方法, 没有融入邻居信息以及额外的辅助信息, 仅依靠用户与物品之间的交互历史进行建模, 导致其无法表达用户与物品

之间复杂的高阶关系,进而导致推荐性能下降.

(2) LightGCN 是一种基于图神经网络的协同过滤方法,也是我们提出的模型的基础编码器,它的性能要优于 BUIR 和 NeuMF,这表明考虑用户与物品之间的高阶关系有助于改进推荐性能.

表 3 基线模型对比实验结果 (*Precision@10*, *Recall@10*)

模型	Epinions		Yelp		Ciao	
	<i>Precision@10</i>	<i>Recall@10</i>	<i>Precision@10</i>	<i>Recall@10</i>	<i>Precision@10</i>	<i>Recall@10</i>
BUIR	0.0110	0.0198	0.00714	0.0330	0.0245	0.0375
NeuMF	0.0095	0.0169	0.0036	0.0156	0.0178	0.0278
LightGCN	0.0180	0.0325	0.0066	0.0303	0.0246	0.0405
DiffNet++	0.0132	0.0244	0.0044	0.0199	0.0171	0.0280
MHCN	0.0201	0.0340	0.0072	0.0325	0.0215	0.0312
SGL	0.0216	0.0393	0.0068	0.0227	0.0224	0.0392
SimGCL	0.0213	0.0319	0.0062	0.0325	0.0231	0.0419
LightGCL	0.0221	0.0386	0.0069	0.0328	0.0251	0.0408
SEPT	<u>0.0226</u>	<u>0.0401</u>	<u>0.0077</u>	<u>0.0330</u>	<u>0.0248</u>	0.0398
Ours	0.0230	0.0415	0.0079	0.0350	0.0270	0.0428
提升 (%)	↑1.7	↑3.4	↑2.5	↑6.0	↑8.2	↑7.0

表 4 基线模型对比实验结果 (*F1@10*, *NDCG@10*)

模型	Epinions		Yelp		Ciao	
	<i>F1@10</i>	<i>NDCG@10</i>	<i>F1@10</i>	<i>NDCG@10</i>	<i>F1@10</i>	<i>NDCG@10</i>
BUIR	0.0141	0.0191	0.0117	0.0205	0.0297	0.0395
NeuMF	0.0121	0.0153	0.0059	0.0096	0.0217	0.0271
LightGCN	0.0232	0.0303	0.0108	0.0189	0.0306	0.0394
DiffNet++	0.0171	0.0214	0.0072	0.0118	0.0212	0.0259
MHCN	0.0253	0.0338	0.0119	0.0204	0.0309	0.0417
SGL	0.0245	0.0359	0.0114	0.0196	0.0269	0.0365
SimGCL	0.0236	0.0379	0.0123	0.0208	0.0302	0.0415
LightGCL	0.0279	0.0396	0.00057	0.0203	0.0289	0.0398
SEPT	0.0289	0.0380	0.0125	0.0213	0.0311	0.0420
Ours	0.0294	0.0398	0.0129	0.0220	0.0324	0.0432
提升 (%)	↑1.7	↑4.7	↑3.2	↑3.2	↑4.0	↑2.8

(3) DiffNet++和 MHCN 是两种社交推荐模型, MHCN 是将社交关系建模为超图, 它的性能要优于 DiffNet++, 这说明超图的表达能力要强于简单图. 另外引入社交关系的实验效果要优于单纯的协同过滤方法, 这也说明了融入社交网络对推荐性能的改善有正向作用.

(4) SGL、SimGCL 和 LightGCL 模型都是近期提出的基于自监督学习的推荐方法, 其中, SGL 和 LightGCL 是对用户-物品二部图进行扰动得到的自监督信号, 而 SimGCL 是对用户和物品的表示向量进行扰动, 得到自监督信号, 但是它们都忽略了社交网络信息和物品类别信息, 因此其推荐性能略差于本文所提出的模型.

(5) SEPT 是目前最好的方法, 它不仅考虑了用户之间的社交网络, 而且利用对比学习来增强模型的表达能力, 这使得 SEPT 方法要优于其他基线模型. 但是, 它忽略了物品之间的内在联系, 无法对物品的表示更好地建模, 导致它的性能要差于我们的模型.

(6) 结合表 3、表 4 的实验结果可以发现, 我们所提出的方法一致地优于所有考虑的基线模型, 而且与目前最好的 SEPT 模型相比, 我们的模型在 3 个数据集上以及 4 种不同的评价指标下均有不同程度的提升, 这说明引入物品之间的内在联系可以进一步改善推荐性能, 尤其在 Ciao 数据集上的提升效果更为明显, *Precision@10* 的提升最高可达 8.2%, 这说明融入物品关系图一方面可以改善推荐系统中的数据稀疏问题, 另一方面可以学到更好的物

品表示, 进而提升推荐的性能.

3.6 消融实验 (RQ2)

为了验证所提出方法中每个模块对整体推荐性能的影响, 我们考虑以下的消融实验.

(1) w/o 去掉用户视角: 表示删除模型中的用户视角, 即不考虑用户社交网络, 只将物品之间的内在联系图融入推荐模型中.

(2) w/o 去掉物品视角: 表示删除模型中的物品视角, 即不考虑物品之间的内在联系图, 只将用户之间的社交网络融入推荐模型中.

(3) w/o 去掉对比损失: 表示删除模型中的对比损失, 只使用推荐损失来训练整个模型.

以上变种模型的实验结果如图 2 所示, 从中我们得到以下几点结论.

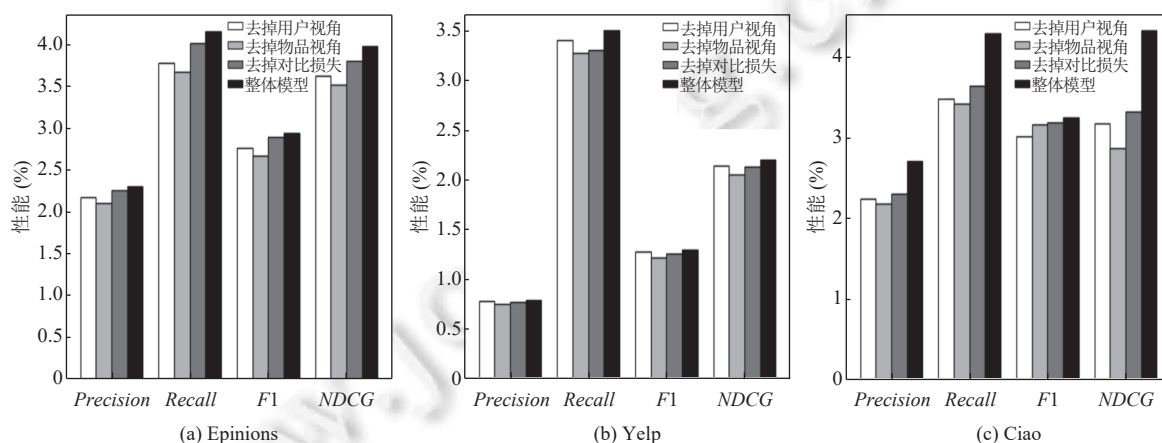


图 2 消融实验结果

(1) 显而易见, 无论去掉哪个模块, 都会导致整体的推荐性能下降, 这说明模型中的每个模块都会从不同的角度来改善用户和物品的表示, 进而提升推荐的性能.

(2) 当去掉用户视角, 即不考虑用户之间的社交网络关系时, 从图 2 中可以看出对 Ciao 数据集的影响最大, 这表明在电商领域中用户的购买偏好受社交关系影响最大, 而 Epinions 数据集的影响要大于 Yelp 数据集, 这是由于 Epinions 数据集是大众消费者点评网站, 用户的反馈程度占比较大, 因此在此数据集上用户的贡献度较大, 在 Epinions 数据集、Yelp 数据集以及 Ciao 数据集下, 当去掉用户视角时, 根据消融实验结果, 各项评价指标都在下降, 在 Epinions 数据集下, Precision、Recall、F1、NDCG 指标相对降低比例分别为 5.7%、7.9%、6.5%、8.4%; 在 Yelp 数据集下, Precision、Recall、F1、NDCG 指标相对降低比例分别为 1.3%、3.0%、1.6%、4.1%; 在 Ciao 数据集下, Precision、Recall、F1、NDCG 指标相对降低比例分别为 17.0%、18.5%、7.1%、26.7%. 因此, 根据实验数据相对降低比例以及图 2, 用户之间的社交关系对推荐结果起着很大的正向作用.

(3) 当去掉物品视角, 即不考虑物品之间的内在联系时, 在 3 个数据集上的推荐性能下降最为明显, 在 Epinions、Yelp 以及 Ciao 数据集下, 当去掉物品视角时, 根据消融实验结果, 各项评价指标都在下降, 在 Epinions 数据集下, Precision、Recall、F1、NDCG 指标相对降低比例分别为 9.2%、11.6%、10.3%、12.5%; 在 Yelp 数据集下, Precision、Recall、F1、NDCG 指标相对降低比例分别为 6.4%、7.1%、6.9%、9.5%; 在 Ciao 数据集下, Precision、Recall、F1、NDCG 指标相对降低比例分别为 19.2%、20.1%、2.5%、34.0%. 这表明, 通过挖掘物品之间的内在联系 (比如类别关系) 一方面可以融入相同物品之间的共性特征, 另一方面可以学到更好的物品表示, 进而显著提升推荐性能.

(4) 在我们提出的模型中, 用户视角和物品视角下的对比学习作为推荐中的辅助任务是为了学到更好的用户

和物品表示,它是通过随机扰动的方式获取自监督信息,并使用对比损失最大化正样本之间的相似度,最小化负样本之间的相似度.从图2的实验结果可以发现,当去掉对比损失时,推荐性能有所下降,在Epinions数据集、Yelp数据集以及Ciao数据集下,当去掉物品视角时,根据消融实验结果,各项评价指标都在下降,在Epinions数据集下,*Precision*、*Recall*、*F1*、*NDCG*指标相对降低比例分别为8.9%、10.5%、9.2%、11.2%;在Yelp数据集下,*Precision*、*Recall*、*F1*、*NDCG*指标相对降低比例分别为5.9%、6.8%、6.2%、8.0%;在Ciao数据集下,*Precision*、*Recall*、*F1*、*NDCG*指标相对降低比例分别为14.8%、15.1%、1.9%、23.1%.这也说明了对比损失对提升推荐性能的有效性.

由图2的实验结果以及各项评价指标的相对降低比例可以看出,无论去掉哪个模块都会导致推荐性能下降,但是从中可以看出去掉物品视角,推荐性能下降最为明显,因此在用户-物品交互图以及社交网络图的基础上加入物品类别关系图,并且使用多视图共同训练的方式生成自监督信号,通过不断降低推荐损失和对比损失从而不断更新用户、物品表示,可以显著提高推荐性能.

3.7 参数灵敏度分析 (RQ3)

在本节中,为了验证模型中重要超参数对推荐性能的灵敏度,我们在一定范围内调整损失函数权重 λ 和温度参数 τ ,观察推荐结果的变化情况.

对于超参数损失函数权重 λ ,根据经验,我们在一个小的取值范围内寻找适当的值,将其取值设置为{0.001, 0.002, 0.003, 0.005, 0.007, 0.008, 0.01}.首先提出的推荐模型中,损失包括3部分:推荐损失、用户视角对比损失、物品视角对比损失,由于损失函数通过梯度更新,反向传播机制可以更新用户、物品的表示,并且对比损失函数在一定程度上可以使得模型的训练趋于平滑,因此对比损失在整体模型中起着重要的作用.实验结果如图3所示.在Epinions数据集上,随着 λ 的增加,该模型的性能逐渐提升,当 λ 取值为0.002时各项评价指标值达到了峰值,说明0.002是在该数据集上的拐点,拐点之后,评价指标先缓缓降低,然后呈现稳步上升趋势,但是我们经过实验验证,总体来说,每项指标值不会大于0.002时的推荐效果,说明当 λ 值为0.002时,对比损失函数对模型训练质量达到最优点.在Yelp数据集上,从指标结果的变化可以看出,当 λ 取值为0.001, 0.002时效果最好,随着 λ 值的增大,每项指标都有下降趋势,当 λ 取值为0.008时,推荐性能达到最低,大于这个值后会有回升趋势,但是随着 λ 值增大整体推荐效果不理想,这说明Yelp数据集中样本的类别分配比较均衡,较小的 λ 值对比损失函数就达到最优.在Ciao数据集上可以很明显地看出,指标结果的变化当 λ 取值为0.003时效果最好,当 λ 取值大于0.003时,各项指标结果逐渐降低,同时表明推荐性能逐渐降低,说明0.003是在该数据集上的拐点,当取此值时,模型推荐性能达到最好.在这3个公开数据集上可以看出,随着 λ 值的增大,训练结果都会出现峰值,然后趋于收敛,因此合理地调整损失函数权重对整体模型训练效果起着重要的作用.

接下来研究温度参数对模型性能的影响,实验结果如图4所示.这里,我们将温度参数 τ 的取值分别设置为{0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.7, 1.0}.从图4可以看出在Epinions数据集上,模型性能变化对参数 τ 比较敏感,性能波动比较大,当取值为0.1时,模型性能最好.随着 τ 的增大,评价指标逐渐降低,意味着在Epinions数据集上 τ 的值越大,对比损失函数对负样本的区分度会变低,从而推荐性能降低.当取值为0.5时,模型性能最低,当大于0.5时,性能有稳步提升,并达到均匀的变化趋势,说明温度参数设置的越大,对比损失对负样本的感知程度发挥不了自身的作用.在Yelp数据集上,可以看出模型在 τ 的取值范围内整体变化比较平稳,当温度参数 τ 取值为0.2时,模型效果达到最好.说明 τ 取值0.2时,在对比损失中对负样本的感知程度最大.随着 τ 值的增大,逐步呈稳定趋势,说明在Yelp数据集上,温度参数 τ 值越大对比损失函数对数据集中的负样本感知程度越低,从而模型的推荐性能越小.在Ciao数据集上,可以看出模型推荐性能在 τ 取值范围内波动比较明显,当温度参数 τ 取值为0.5时,*Precision*、*F1*和*NDCG*指标达到最大,当温度参数 τ 取值为0.4时,*Recall*指标达到最大,因此说明模型性能在温度参数取值为0.4、0.5时,对负样本的感知程度达到最大.而通过在Epinions数据集和Yelp数据集上的实验结果对比,可以看出由于Epinions整体数据集量比Yelp数据集量大,当温度参数取值较小时,模型对负样本的感知程度才会变大,才能达到好的推荐效果,但是温度参数 τ 的值不能设置太小,因为在本文研究中我们本身要根

据用户社交网络关系以及物品类别找到与当前用户(物品)更相似的类别,其相似样本就可能存在于负样本中. 经过充分的实验验证了在 Epinions 数据集、Yelp 数据集以及 Ciao 数据集上,模型达到最优的温度参数分别取值为 0.1、0.2、0.5.

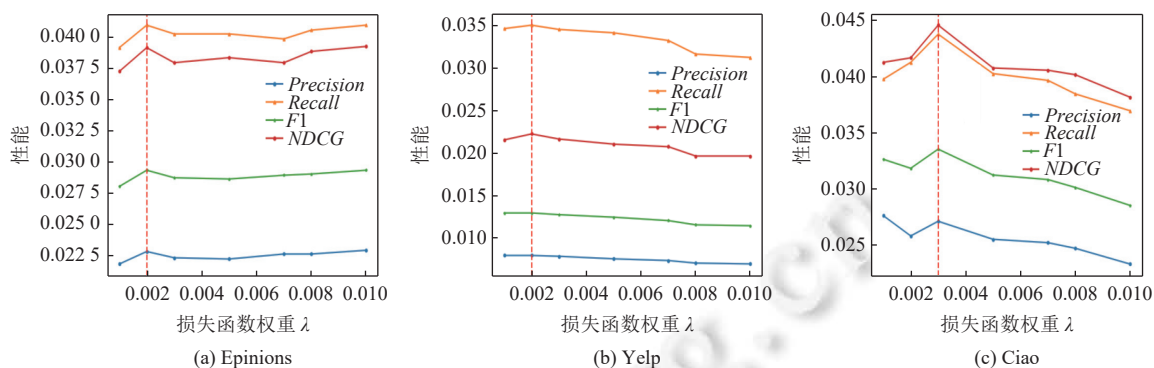


图3 损失函数权重 λ 实验结果

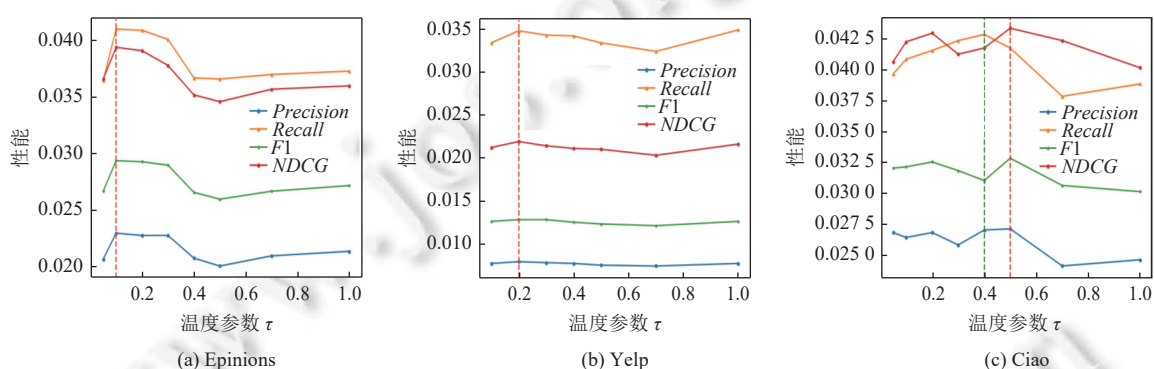


图4 温度参数 τ 实验结果

4 总 结

我们提出了基于多视角的自监督推荐模型,分别从偏好视角、用户视角、物品视角详细介绍了模型细节. 我们采用多视图共同训练的方式生成自监督信号,充分利用用户-物品交互图、社交网络图、物品联系图之间的用户、物品关系,并使用互信息估计中的对比损失函数计算对比损失, BPR 来计算推荐损失. 我们的模型分别在 3 个公开数据集上进行了充分的实验,并做了消融实验、参数灵敏度分析,在消融实验中,可以看到我们去掉任何一个视角都会降低我们提出的整体模型的实验结果,我们分别在损失函数系数和对比损失中的温度系数两个参数下充分进行了实验,对实验结果做了充分的介绍和说明. 实验结果一致地表明我们的模型在提高推荐性能方面得到了验证.

References

- [1] Wu SW, Sun F, Zhang WT, Xie X, Cui B. Graph neural networks in recommender systems: A survey. ACM Computing Surveys, 2023, 55(5): 97. [doi: 10.1145/3535101]
- [2] Liu X, Zhang FJ, Hou ZY, Mian L, Wang ZY, Zhang J, Tang J. Self-supervised learning: Generative or contrastive. IEEE Trans. on Knowledge and Data Engineering, 2023, 35(1): 857–876. [doi: 10.1109/TKDE.2021.3090866]
- [3] Chen T, Kornblith S, Norouzi M, Hinton G. A simple framework for contrastive learning of visual representations. In: Proc. of the 37th Int'l Conf. on Machine Learning. 2020. 1597–1607.
- [4] Han T, Xie W, Zisserman A. Self-supervised co-training for video representation learning. In: Proc. of the 34th Int'l Conf. on Neural

- Information Processing Systems. Vancouver: Curran Associates Inc., 2020. 477.
- [5] Lan ZZ, Chen MD, Goodman S, Gimpel K, Sharma P, Soricut R. ALBERT: A lite BERT for self-supervised learning of language representations. In: Proc. of the 8th Int'l Conf. on Learning Representations. 2020.
 - [6] van den Oord A, Li YZ, Vinyals O. Representation learning with contrastive predictive coding. arXiv:1807.03748, 2018.
 - [7] Qiu JZ, Chen QB, Dong YX, Zhang J, Yang HX, Ding M, Wang KS, Tang J. GCC: Graph contrastive coding for graph neural network pre-training. In: Proc. of the 26th ACM SIGKDD Int'l Conf. on Knowledge Discovery & Data Mining. ACM, 2020. 1150–1160. [doi: [10.1145/3394486.3403168](https://doi.org/10.1145/3394486.3403168)]
 - [8] Jin W, Derr T, Liu HC, Wang YQ, Wang SH, Liu ZT, Tang JL. Self-supervised learning on graphs: Deep insights and new direction. arXiv:2006.10141, 2020.
 - [9] You YN, Chen TL, Sui YD, Chen T, Wang ZY, Shen Y. Graph contrastive learning with augmentations. In: Proc. of the 34th Int'l Conf. on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2020. 488.
 - [10] Wu JC, Wang X, Feng FL, He XN, Chen L, Lian JX, Xie X. Self-supervised graph learning for recommendation. In: Proc. of the 44th Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. ACM, 2021. 726–735. [doi: [10.1145/3404835.3462862](https://doi.org/10.1145/3404835.3462862)]
 - [11] Zhang JW, Gao M, Yu JL, Guo L, Li JD, Yin HZ. Double-scale self-supervised hypergraph learning for group recommendation. In: Proc. of the 30th ACM Int'l Conf. on Information & Knowledge Management. Queensland: ACM, 2021. 2557–2567. [doi: [10.1145/3459637.3482426](https://doi.org/10.1145/3459637.3482426)]
 - [12] Yu JL, Yin HZ, Gao M, Xia X, Zhang XL, Viet Hung NQ. Socially-aware self-supervised tri-training for recommendation. In: Proc. of the 27th ACM SIGKDD Conf. on Knowledge Discovery & Data Mining. ACM, 2021. 2084–2092. [doi: [10.1145/3447548.3467340](https://doi.org/10.1145/3447548.3467340)]
 - [13] Rendle S, Freudenthaler C, Gantner Z, Schmidt-Thieme L. BPR: Bayesian personalized ranking from implicit feedback. arXiv:1205.2618, 2012.
 - [14] Wu ZH, Pan SR, Chen FW, Long GD, Zhang CQ, Yu PS. A comprehensive survey on graph neural networks. IEEE Trans. on Neural Networks and Learning Systems, 2021, 32(1): 4–24. [doi: [10.1109/TNNLS.2020.2978386](https://doi.org/10.1109/TNNLS.2020.2978386)]
 - [15] Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. In: Proc. of the 5th Int'l Conf. on Learning Representations. 2017.
 - [16] van den Berg R, Kipf TN, Welling M. Graph convolutional matrix completion. arXiv:1706.02263, 2017.
 - [17] Zhang JN, Shi XJ, Zhao SL, King I. STAR-GCN: Stacked and reconstructed graph convolutional networks for recommender systems. In: Proc. of the 28th Int'l Joint Conf. on Artificial Intelligence. 2019. 4264–4270. [doi: [10.24963/ijcai.2019/592](https://doi.org/10.24963/ijcai.2019/592)]
 - [18] Wang X, He XN, Wang M, Feng FL, Chua TS. Neural graph collaborative filtering. In: Proc. of the 42nd Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. Paris: ACM, 2019. 165–174. [doi: [10.1145/3331184.3331267](https://doi.org/10.1145/3331184.3331267)]
 - [19] He XN, Deng K, Wang X, Li Y, Zhang YD, Wang M. LightGCN: Simplifying and powering graph convolution network for recommendation. In: Proc. of the 43rd Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. ACM, 2020. 639–648. [doi: [10.1145/3397271.3401063](https://doi.org/10.1145/3397271.3401063)]
 - [20] Wu S, Tang YY, Zhu YQ, Wang L, Xie X, Tan TN. Session-based recommendation with graph neural networks. In: Proc. of the 33rd AAAI Conf. on Artificial Intelligence. Honolulu: AAAI, 2019. 346–353. [doi: [10.1609/aaai.v33i01.3301346](https://doi.org/10.1609/aaai.v33i01.3301346)]
 - [21] Wu L, Li JW, Sun PJ, Hong RC, Ge Y, Wang M. DiffNet++: A neural influence and interest diffusion network for social recommendation. IEEE Trans. on Knowledge and Data Engineering, 2022, 34(10): 4753–4766. [doi: [10.1109/TKDE.2020.3048414](https://doi.org/10.1109/TKDE.2020.3048414)]
 - [22] Yu JL, Yin HZ, Li JD, Wang QY, Hung NQV, Zhang XL. Self-supervised multi-channel hypergraph convolutional network for social recommendation. In: Proc. of the 2021 Web Conf. Ljubljana: ACM, 2021. 413–424. [doi: [10.1145/3442381.3449844](https://doi.org/10.1145/3442381.3449844)]
 - [23] Oldale A, Oldale J, van Reenen J, Campbell M. Collaborative filtering: US, 2004054572(A1). 2004-03-18.
 - [24] Yin HZ, Zhou XF, Cui B, Wang H, Zheng K, Nguyen QVH. Adapting to user interest drift for POI recommendation. IEEE Trans. on Knowledge and Data Engineering, 2016, 28(10): 2566–2581. [doi: [10.1109/TKDE.2016.2580511](https://doi.org/10.1109/TKDE.2016.2580511)]
 - [25] Huang C, Xu HC, Xu Y, Dai P, Xia LH, Lu MY, Bo LF, Xing H, Lai XP, Ye YF. Knowledge-aware coupled graph neural network for social recommendation. In: Proc. of the 35th AAAI Conf. on Artificial Intelligence. AAAI, 2021. 4115–4122. [doi: [10.1609/aaai.v35i5.16533](https://doi.org/10.1609/aaai.v35i5.16533)]
 - [26] Yang LW, Liu ZW, Dou YT, Ma J, Yu PS. ConsisRec: Enhancing GNN for social recommendation via consistent neighbor aggregation. In: Proc. of the 44th Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. ACM, 2021. 2141–2145. [doi: [10.1145/3404835.3463028](https://doi.org/10.1145/3404835.3463028)]
 - [27] Fan WQ, Ma Y, Li Q, He Y, Zhao E, Tang JL, Yin DW. Graph neural networks for social recommendation. In: Proc. of the 2019 World Wide Web Conf. San Francisco: ACM, 2019. 417–426. [doi: [10.1145/3308558.3313488](https://doi.org/10.1145/3308558.3313488)]
 - [28] Song LQ, Bi Y, Yao MQ, Wu ZY, Wang JM, Xiao J. DREAM: A dynamic relation-aware model for social recommendation. In: Proc. of

- the 29th ACM Int'l Conf. on Information & Knowledge Management. ACM, 2020. 2225–2228. [doi: [10.1145/3340531.3412115](https://doi.org/10.1145/3340531.3412115)]
- [29] He XN, Liao LZ, Zhang HW, Nie LQ, Hu X, Chua TS. Neural collaborative filtering. In: Proc. of the 26th Int'l Conf. on World Wide Web. Perth: ACM, 2017. 173–182. [doi: [10.1145/3038912.3052569](https://doi.org/10.1145/3038912.3052569)]
- [30] Lee D, Kang S, Ju H, Park C, Yu H. Bootstrapping user and item representations for one-class collaborative filtering. In: Proc. of the 44th Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. ACM, 2021. 317–326. [doi: [10.1145/3404835.3462935](https://doi.org/10.1145/3404835.3462935)]
- [31] Yu JL, Yin HZ, Xia X, Chen T, Cui LZ, Nguyen QVH. Are graph augmentations necessary? Simple graph contrastive learning for recommendation. In: Proc. of the 45th Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. Madrid: ACM, 2022. 1294–1303. [doi: [10.1145/3477495.3531937](https://doi.org/10.1145/3477495.3531937)]
- [32] Cai XH, Huang C, Xia LH, Ren XB. LightGCL: Simple yet effective graph contrastive learning for recommendation. In: Proc. of the 11th Int'l Conf. on Learning Representations. 2023.

作者简介

王丹丹, 硕士生, 主要研究领域为数据挖掘, 个性化推荐.

刘海洋, 博士, 讲师, CCF 专业会员, 主要研究领域为机器学习, 数据挖掘.

刘壮, 博士, 主要研究领域为个性化推荐, 图网络表示学习.

原继东, 博士, 副教授, CCF 专业会员, 主要研究领域为数据挖掘, 时间序列分类.