

基于可解释图神经网络的可视推荐分析系统

汤颖, 周元博, 孙国道*

(浙江工业大学计算机科学与技术学院 杭州 310000)
(guodao@zjut.edu.cn)

摘要: 针对推荐系统中图神经网络的可解释性进行研究, 从可解释模型出发, 将推荐问题转换为图分类问题, 利用可解释图神经网络对推荐系统进行解释, 突破了以往推荐中解释多为实例级的情况, 从实例级和组群级出发, 探索推荐场景下的多粒度解释. 另外, 为了增强对解释模型提取的图模式的理解, 设计了可视分析系统, 以更好地理解图模式和模型解释过程, 从单用户、用户群和多个用户群 3 个层级展开探索, 便于分析人员探索图神经网络的推荐模式, 从而对解释的可靠性进行验证. 最后, 在豆瓣电影数据集和 Last-FM 这 2 个真实数据集上应用图模式改进调整训练集, 对比实验中推荐评估指标都得到了提升, 从定量角度进一步证明了解释的可靠性和系统的有效性.

关键词: 可解释图神经网络; 推荐系统; 可视分析; 可解释性
中图法分类号: TP391.41 **DOI:** 10.3724/SP.J.1089.2023-00549

A Visual Recommendation Analysis System Based on the Interpretable Graph Neural Network

Tang Ying, Zhou Yuanbo, and Sun Guodao*

(College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310000)

Abstract: A study on the interpretability of graph neural networks in recommendation systems. Starting from the interpretable model, we convert the recommendation problem into the graph classification problem, and use interpretable graph neural networks to explain the recommendation system, breaking through the previous situation in which explanations in recommendations were mostly at the instance level. The study explores multi-granularity explanations in the recommendation scenario from the instance level to the group level. In addition, to enhance the understanding of the graph patterns extracted by the interpretable model, a visual analysis system is designed to better understand the graph patterns and the model interpretation process. The visual system is explored from three levels: single user, user group and multiple user groups, which facilitates users to explore the recommended pattern of graph neural networks, thereby verifying the reliability of the explanation. Finally, the study designs and implements comparative experiments on two real data sets, Douban movie and Last-FM, in which the extracted graph pattern is used to improve and adjust the training sets. The experimental results show that the recommendation evaluation indicators are improved, which further proves the reliability of the explanation from a quantitative perspective.

Key words: interpretable graph neural network; recommendation system; visual analysis; interpretability

收稿日期: 2023-08-17; 修回日期: 2024-02-22. 基金项目: 国家自然科学基金(61972355, 72192820); 浙江省自然科学基金(LZ23F020010). 汤颖(1977—), 女, 博士, 教授, 博士生导师, CCF 会员, 主要研究方向为大数据分析可视化、计算机图形学; 周元博(1998—), 女, 硕士研究生, 主要研究方向为数据分析与可视化; 孙国道(1988—), 男, 博士, 副教授, 博士生导师, 论文通信作者, 主要研究方向为信息可视化与可视分析.

图结构广泛应用于许多领域,如社交网络、交通路线和分子结构等,为了挖掘图结构的信息,传统的深度学习网络如卷积神经网络(convolutional neural network, CNN)已不再适用,于是图神经网络(graph neural network, GNN)被提出.当前,已提出许多成功且应用广泛的 GNN 模型^[1-3],在节点分类、链接预测等任务中取得了显著的成果.同时, GNN 也被广泛应用于推荐系统中,通过将用户与物品、物品与属性之间的关系构建为图结构,迭代地聚合邻域信息,更新节点表示,并基于节点的嵌入表示得到推荐结果.

在推荐系统中,对于推荐结果进行解释是非常必要的.合理有效的解释可以帮助用户理解推荐结果与他们的历史行为之间的关系,进而让用户对推荐模型建立信赖.以往大部分面向推荐的解释工作^[4]包括分析用户的兴趣偏好,并对用户特征进行画像,从而对当前推荐结果进行解释.现有的基于 GNN 的推荐系统是从用户与推荐结果的路径出发,将注意力分数最大路径作为解释.这些工作在一定程度上能够对推荐结果提供有效的解释,但是这些解释没有从模型出发且都针对特定用户,往往具有个体独特性,并不适用于其他用户,所以这样的解释不具有说服力.因此,如何把解释拓展到一个局部范围,给出一个适用于群体的通用解释,是目前需要解决的问题;另外,以往的工作在推荐的可视化设计上较简单,通常只是简单地展示推荐结果的静态信息,缺乏交互设计,无法支持用户对推荐结果进行深入的互动探索.然而,用户通常期望了解推荐背后的推理过程、推荐结果的可靠性,以及对推荐结果进行个性化调整的可能性.

本文利用可解释 GNN 对推荐结果进行解释,但是 GNN 解释模型却无法直接对推荐场景进行解释,主要有以下原因.

(1) 已有的可解释 GNN 模型支撑的下游任务通常为图分类和节点分类,这类任务通常有清晰的评估标准.例如,在图分类场景中使用的 MUTAG 数据集,该数据集用于预测分子图是否导致癌变,解释的评估标准是致癌分子结构是否真实存在.然而,在推荐系统中,由于数据集和应用场景的不同,对于推荐结果生成解释的可靠性缺乏明确的评估标准,因此系统不仅无法直接使用现有解释模型中的目标函数对推荐场景进行解释,同时也很难定义出新的目标函数来优化解释模型.

(2) 已有的基于 GNN 的可解释工作提供实例级(instance-level)、组级(group-level)或类级(class-level)

的解释.实例级方法如 PGExplainer^[5],组级方法如 GNNExplainer^[6],类级方法如 XGNN^[7].与组级和类级方法相比,实例级方法提供的解释更细粒度和更精确,而组级和类级方法提供的解释更通用.然而,目前现有的推荐系统的可解释性工作都集中在特定实例上,对于特定实例的解释往往存在局限性.因此,如何突破实例级的限制,从整体出发对组群的推荐结果进行解释是一个待解决的问题.此外,如何将实例级和组群级结合起来形成对推荐系统中 GNN 的多粒度解释,也是本文需要探究的问题.

另外,解释需向用户展示,可视分析不仅提供了直观展示的方式,还可以提供对解释的可视验证.然而,在真实的推荐场景中,由于没有 GroundTruth 作为衡量基准,解释的好坏难以确定.因此,如何对推荐场景中 GNN 的解释进行可视呈现和可视评估是一个重要的问题.

综上所述,本文的研究重点是如何利用可解释的 GNN 对推荐系统进行解释,并通过可视分析技术对解释进行分析和评估.本文的解决思路是将推荐结果与 GNN 解释模型相结合.解决方法是将用户推荐结果转换为图结构,将推荐得到的嵌入空间分布转换为分类任务,使用 GNN 解释模型提取原始图中重要的子图结构(图模式),通过图模式对推荐结果进行解释.为了理解图模式,本文设计了一个可视分析系统来展示图模式,从单个用户、用户群和多个用户群 3 个层级展开探索,以便分析人员能够进行多粒度的图模式挖掘;同时探索 GNN 的推荐模式;最后通过研究解释图模式与用户偏好模式的一致性,验证解释模型的可靠性.

1 相关工作

1.1 推荐系统中的 GNN

GNN 通常用于处理图结构数据,其核心思想是进行信息聚合,通过沿着图的边关系递归传递信息来更新节点表示.常见的 GNN 结构包括:基于消息传递的模型(如 MPNN^[8]),基于注意力机制的模型(如 GAT^[3]),基于元学习的模型(如 Meta-GNN^[9])等. GNN 通常应用的场景包括推荐系统、链路预测、文本分类和图像视觉等.在推荐系统中, GNN 主要分为 3 类: (1) 基于用户-物品二部图的推荐算法,它利用 GNN 对二部图中的节点进行特征学习,如 GATNE^[10]; (2) 基于序列图的推荐算法,它利用 GNN 对序列图中的节点进行信息聚

合, 以解决用户行为具有时序性的推荐问题, 如 GraphRec^[11]; (3) 基于知识图谱的推荐算法, 它通过引入物品属性, 利用物品间的语义关系构建高阶连接关系, 如 KGAT^[12].

但在推荐系统中, 对于 GNN 所产生的推荐结果缺乏很好的解释. 在现有的推荐可解释工作中, 解释是通过把具有最大注意力分数的路径作为依据进行的, 它停留在注意力层面, 但这样的解释具有局限性, 一个有效的解释应该具有普适性和可靠性. 因此, 本文提出了将可解释 GNN 应用于对推荐结果的解释.

1.2 GNN 的可解释性

目前, 关于 GNN 的可解释性的研究按照解释粒度分为实例级别、组级别和类级别. 实例级别的解释方法是为单个输入样本提供解释, 与输入有关, 已有的解释方法包括 PGExplainer^[5], RCExplainer^[13], ReFine^[14]和 CF2^[15]等. 其中, CF2^[15]框架通过从因果理论中洞察反事实和事实推理来制定优化问题, 提取必要且充分的信息, 其中反事实确保提取到必要的信息, 事实确保提取到充分的信息, 最终得到理想的子图解释. 组级别解释方法是提供了一个原型样本来解释一组样本的结果. 其中, GNNExplainer^[6], 是第 1 个通用且与模型无关的解释方法, 它将一个训练好的 GNN 及其预测结果作为输入, 为一组样本提供解释, 这种解释包括一个小的节点特征子集和一个紧凑的子图结构, 这些节点特征和子图在 GNN 的预测任务中起着关键作用. 类级别解释方法目前只有 XGNN^[7]和 GNNInterpreter^[16] 2 种, 其中, XGNN^[7]通过训练一个图生成器来解释 GNN, 使得图生成器生成的图模式能够最大化模型对某种类别的预测结果, 那么这个图模式就作为该类别的解释. XGNN 将图生成器表述为强化学习任务, 通过预测如何向给定的图中添加一条边, 然后根据训练好的 GNN 对生成的图进行反馈, 用策略梯度^[17]对生成图进行训练. 关于类级别解释方法少的原因在于实例级解释是为单个样本制定的解释, 而类级别给出的解释不依赖于输入, 而是通过研究哪种图模式可以导致某种预测结果, 为某一类别样本提供共有的解释, 具有全局性.

在现有的研究中, 可解释 GNN 广泛应用于节点分类和图分类任务, 但缺乏对推荐任务下 GNN 的解释研究. 此外, 目前推荐中的解释都是基于单个用户的, 而单独用户具有独特性. 因此, 本文将推荐的解释从实例级扩展到组级别, 更深层次地

探索 GNN 在推荐任务中的推荐模式.

1.3 面向模型理解的可视化工作

根据文献[18], 将可视分析相关技术分为模型构建前、模型构建期间和模型构建后的技术. 模型构建前的可视分析技术主要是帮助模型开发人员更好地准备建模所需的数据, Ingram 等^[19]提出的可视化分析系统, 允许用户探索特征及其关系, 并交互式地删除不相关和冗余的特征. 模型构建过程中的可视分析技术能够直观地解释模型的工作机制, 如参数的变化如何影响模型的结果, 以及为什么模型会对特定的输入给出特定的输出. Utopian^[20]的开发者为了细化主题模型的工作机制, 允许用户交互式地合并或分割主题, 并会相应地自动修改其他主题; 构建和部署机器学习模型, 帮助用户以一种直观的方式理解模型输出, 以促进对模型输出的信任. 这类模型构建后的可视分析方法由于面向的对象通常是用户, 因此其重点是直观表示和探索模型输出数据. Görg 等^[21]开发了一个多视图可视化, 包括列表视图、集群视图、单词云图、网格视图和文档视图, 以直观地说明文档摘要、文档聚类、情感分析、实体标识和推荐的分析结果. 通过将交互式可视化与文本挖掘技术相结合, 为用户提供了一个顺畅且丰富的探索系统.

基于上述研究工作, 本文决定利用可解释 GNN 对推荐系统进行解释, 并设计可视分析系统来更好地展示解释过程, 验证解释模型的可靠性; 系统支持推荐场景下的多粒度解释, 帮助分析人员探索推荐模式, 增强对 GNN 的解释.

2 数据准备和工作流程

2.1 数据准备

为了以系统且公平的方式理解 GNN 在推荐系统中的推荐模式, 并为模型和推荐结果提供丰富的上下文信息, 本文分别在 2 个数据集进行实验.

2.1.1 豆瓣电影

本文使用 scrapy 框架从真实的豆瓣电影网站爬取了大量数据, 包括 2 万份豆瓣用户数据, 2 万部豆瓣电影数据, 1 万份导演数据, 1 万份演员数据以及 300 万条用户历史观影记录. 这些电影数据涵盖了上映年份、类型、标签、导演和演员等详细信息, 数据内容如表 1 所示. 爬取内容包含电影信息和用户信息. 通过对电影条目中导演数据进行统计分析, 生成导演特征文件, 包含导演名称和导演 ID; 通过对电影条目中主演数据进行统计分析, 生成演

员特征文件, 包含演员名称和演员 ID; 通过对电影条目中类型数据进行统计分析, 生成类型特征文件, 包含类型名称和类型 ID.

表 1 豆瓣电影网站爬取数据描述

条目	数据内容
电影条目	电影 ID, 电影名称, 导演, 主演, 上映日期, 类型, 标签, 豆瓣评分, 电影海报链接
用户条目	用户 ID, 已观看电影 ID

为确保数据质量, 防止出现无效数据, 以用户的观影记录至少大于 50、电影被有效用户观看次数大于 20 和电影的类型文本完整等条件进行了数据过滤. 此外, 考虑真实场景中推荐任务通常面临着冷启动问题, 保留了 10% 的观影记录小于 50 部电影的用户作为冷启动数据, 以便全面地评估模型性能. 经过数据过滤之后, 构建图数据, 节点类型为用户(U)、电影(M)、导演(D)、演员(A)和电影类型(G)这 5 种, 边类型为用户-电影、电影-导演、电影-演员和电影-类型 4 种. 该数据集最终统计结果如表 2 所示.

表 2 豆瓣电影网站爬取数据集信息

节点类型	节点数量	边关系	边数量
用户(U)	9622	用户-电影	2195882
电影(M)	12385	电影-演员	44331
演员(A)	9598	电影-导演	17987
导演(D)	6287	电影-类型	29951
电影类型(G)	34		

2.1.2 Last-FM

该公开数据集来自 Last-FM 在线音乐系统中的用户收听记录数据, 包括 23 566 个用户, 48 123 个音乐. 该数据集基本信息如表 3 所示.

表 3 数据集信息

节点类型	数量
用户	23 566
音乐	48 123
边关系	3 034 763
实体属性	58 266
关系	9

2.2 特征相关性分数

本文使用特征相关性分数(特征分数)来衡量用户的观影特征. 该方法基于目标用户的历史观影记录中各电影的属性特征进行计算, 包括类型相关性分数、导演相关性分数和演员相关性分数. 因为所

有电影的类型都在 35 种以内, 所以类型集合 $L = \{L_1, L_2, \dots, L_i, \dots, L_{35}\}$. 用户集合 $U = \{U_1, U_2, \dots, U_k, \dots, U_{9622}\}$ 表示 9 622 个用户, 假设 $k \in \{1, 2, \dots, 9622\}$, 用户 U_k 看过 n 部电影, 其集合表示为 $M_k = \{m_1, m_2, \dots, m_j, \dots, m_n\}$. 其中, 电影 m_j 对应的类型集合为 G_j , 使用 R_j^i 表示 m_j 中类型 L_i 的特征分数. 以类型相关性分数为例, 其具体计算公式为

$$R_j^i = \begin{cases} 0, & \text{if } L_i \subseteq G \\ \frac{1}{|G_j|}, & \text{otherwise} \end{cases} \quad (1)$$

针对用户的类型相关性分数, 用户 U_k 看过的电影中关于类型 L_i 的相关性分数为

$$S_{L_i} = \sum_{j=1}^{|M_k|} R_j^i \quad (2)$$

用户群的相关性分数计算是在用户的相关性分数基础上进行求和操作.

2.3 需求分析

本文提出的基于可解释 GNN 的可视推荐分析系统, 目标用户是具有一定机器学习背景的人员(分析人员), 通过与他们讨论后, 发现他们比较关心下述问题: 推荐模型往往生成高维的嵌入表示, 代表用户实体的嵌入表示, 用户的位置和其观影偏好之间是否存在联系; 不只是单个用户, 对于一个用户群来说, 所包含的用户是否有共同的观影偏好, 不同用户群之间是否有差异, 能否对差异和推荐结果进行解释; 能否提供对解释的可视交互和可视验证, 帮助分析人员更快地理解推荐细节.

本文围绕上述问题, 给出系统需要完成的任务.

T1. 支持用户和用户群观影数据的偏好分析与挖掘. 推荐模型通常会生成高维的实体嵌入表示, 其包含了该实体在高阶关系下的语义信息, 然而对于分析人员, 高维数据往往难以理解和解释. 因此, 系统应从高维嵌入结果出发, 通过降维和聚类来获得有意义的散点分布图. 同时, 将空间分布图和特征相关性分数相结合, 对用户和用户群的观影偏好模式进行比较和分析, 从而揭示其异同之处并展开深入的细节挖掘.

T2. 基于可解释 GNN 对推荐结果进行解释. 除了对用户的历史偏好模式进行分析之外, 还需要对用户和用户群的推荐偏好模式进行探索. 利用优化后的用户推荐图, 将推荐问题与图分类问题相融合, 确保不同用户群之间的推荐图有区分

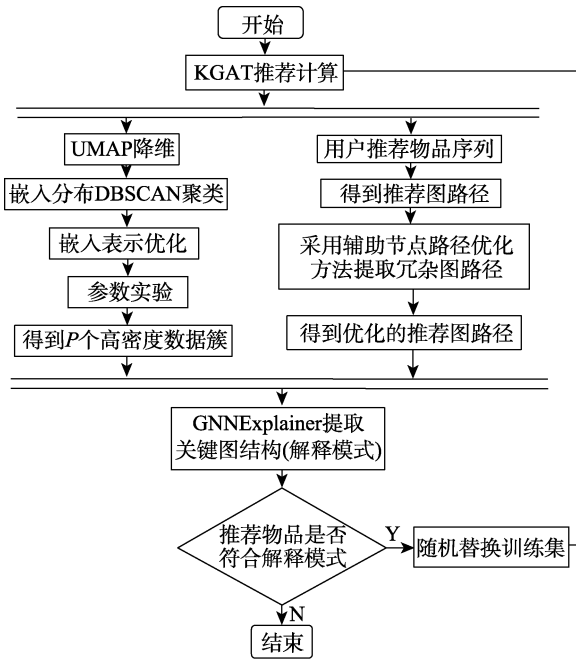


图2 本文方法详细步骤



图3 用户实体原始分布

Step2. 嵌入表示优化. 由于用户群分布存在交叉错乱的情况, 为了能够捕捉到不同用户群的特点, 获得有意义的用户群, 本文对用户的嵌入表示进行更新, 如

$$\mathbf{e}_{u_i} = \sum_{j=1}^N \frac{\mathbf{e}_{u_i} \otimes \mathbf{e}_{u_i, j}}{N} \quad (3)$$

其中, $\mathbf{e}_{u_i}$ 表示用户 i 的嵌入; $\mathbf{e}_{u_i, j}$ 表示推荐模型为用户 i 推荐的第 j 个推荐物品的嵌入; N 表示推荐模型为用户推荐的前 N 个推荐物品, 在本文实验中取 $N=20$, 将推荐的前 20 个物品的嵌入表示与用户原有的嵌入表示相结合, 从而更新用户的嵌入表示. 与 Step1 的聚类参数保持一致, 重新聚类后, 用户嵌入分布如图 4 所示, 可以看到, 在对嵌入表示进行优化处理后, 得到了高密度的用户簇.

Step3. 参数实验. 在前 2 个步骤中, 采用降维方法 t -SNE, 发现其处理速度慢且只保留了数据局部的结构, 为了能更好地保留数据的全局结构, 进而使用基于流形

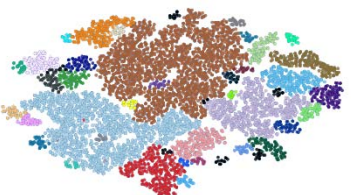


图4 嵌入表示优化

理论的方法降维, 使得降维后的数据点在低维空间中更接近原始数据的全局布局. 随后, 通过多次参数实验, 如图 5 所示, 在 UMAP 算法^[22](最近邻居数量为 100, 相邻点最小距离为 0.15), 无监督 DBSCAN 算法^[23](邻域半径为 3, 最小样本数为 35)的情况下, 获得了分类明显的用户簇.

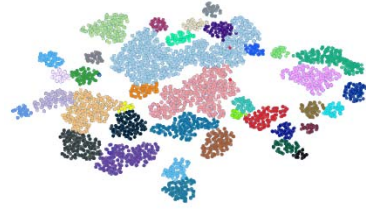


图5 用户群空间分布

最后, 对嵌入优化的合理性进行讨论: 嵌入优化处理后能得到高密度的用户簇. 对优化后的嵌入重新聚类, 可以得到分界清晰的用户群分布, 用户群之间不混杂, 散点位置包含了一定的推荐语义信息. 嵌入优化的目的在于, 每个簇类(用户群)内的用户具有一定程度的相似推荐语义信息; 相反, 空间位置远离的用户群之间必然存在推荐信息差异, 这就构成了分类任务.

3.2 推荐图路径处理

推荐图路径处理的目的是将用户的推荐结果转换为图结构. 该模块包括以下步骤.

Step1. 推荐物品序列. 根据用户 ID, 获得推荐模型为其推荐的 Top 20 个物品 ID. 以用户为起始点、推荐物品为目标点、用户观影信息为中间节点, 构建了如图 6a 所示用户和推荐物品之间的完整路径图.

Step2. 辅助节点路径优化. 采用提出的辅助节点路径优化方法, 实现推荐图路径的优化, 解决图路径冗余, 不易提取信息的缺陷.

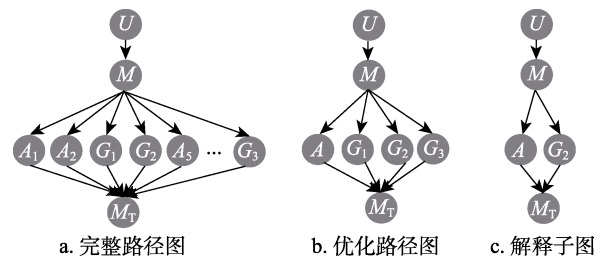


图6 图路径变化顺序示意图

在推荐图结构中, U 表示用户, M 表示用户已观看电影; A_i 表示不同演员, $i \in \{1, 2, \dots, 9598\}$; D_j 表示不同导演, $j \in \{1, 2, \dots, 6287\}$; G_k 表示不同电影类型, $k \in \{1, 2, \dots, 34\}$; U_L 表示相似用户; M_T 表示推荐电影. 如图 6a 所示, 已观看电影 M 和所推荐电影 M_T 之间存在复杂的关系, 通常这种关

联是由共同的演员、电影类型以及导演等节点类型造成的,正是这些共同点促使模型向用户推荐这部电影。

复杂关系导致复杂的路径图,为了解决用户已观看电影繁多使得推荐图路径冗杂,导致不存在共有图模式的问题,本节借鉴数据增广思想^[24],提出了辅助节点路径优化方法。数据增广适用于处理冗杂的图结构,其目的是通过应用一些转换,在不影响语义标签的情况下生成新的图结构。在图数据中,常用的数据增广方法^[25]包括基于节点的方法、基于边扰动的方法、基于属性掩码的方法和基于子图的方法。基于这些方法,在不影响语义的情况下,本文对推荐图路径做数据增广,将繁多的推荐路径图优化为 37 类路径图(其中包括 34 类电影类型,1 类导演,1 类演员,1 类相似用户),得到了如图 6b 所示优化后的推荐图路径。图 6b 中将演员、导演、相似用户各自归为一类节点的处理。以演员举例,如果保留演员节点的特殊性,由于演员众多而无法提取出用户群的共有图模式,因此采用屏蔽演员特殊性,统一归类为演员类别(象征节点“ A ”),简化成了 $U-M-A-M_T$ 这样的路径。对于导演,相似用户也采用了类似的处理方式;但保留类型的原因是因为电影的类型是固定的且小数目(本文的豆瓣数据集中类型为 34 种, $G_1 \sim G_{34}$),不存在路径繁多的情况。因此,通过辅助节点路径优化方法,能够保留路径中类型节点的特殊性;同时屏蔽辅助节点的特殊性,得到优化的推荐图路径。在优化后的推荐图路径中,路径 $U-M-D-M_T$ 表示目标电影与用户观看过的电影由相同导演执导, $U-M-A-M_T$ 表示目标电影与用户观看过的电影由相同演员参演, $U-M-(U_1)-M_T$ 表示目标电影和用户观看过的电影具有相同的观影用户, $U-M-(G_1)-M_T$ 表示目标电影与用户观看过的电影都具有“剧情”类型。

3.3 推荐场景中基于 GNNExplainer 的可解释 GNN 模型

以豆瓣数据作为样例,对推荐场景中的可解释性进行阐述。

在得到用户和推荐物品之间的推荐路径图结构之后,如何从路径图中提取出真正影响推荐结果的子图,即提取重要图模式,是本节要解决的问题。具体地,如图 7 所示,在电影推荐场景中,给用户推荐的电影含有非常多的特征,包含喜剧类型、动作类型、剧情类型、动画类型、导演、演员类型和爱情类型等节点类型,但推荐预测中对用

户起关键作用的往往是一小部分节点类型,如根据喜剧和动作类型去推荐电影。

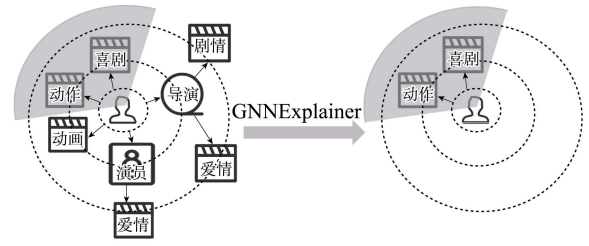


图 7 推荐场景示意图

那么针对图 6b 所示用户群的推荐路径图,如何识别出一个如图 6c 所示紧凑的子图结构,该子图结构在用户群的分布预测中起着至关重要的作用,是本节要解决的问题。GNNExplainer 能够解决该问题,它的核心思想是识别出对预测结果最具有影响力的子图结构和节点特征。GNNExplainer 的工作原理是将一个已经训练好的 GNN 模型和其预测结果作为输入,然后输出一个子图,该子图能够对该 GNN 模型的预测结果产生最大的影响。GNNExplainer 子图提取方法具体步骤如下。

Step1. 初始化。初始化图掩码矩阵和特征掩码矩阵,用于选择重要子图结构和节点特征。

Step2. 输入数据。输入一个已训练好的 GNN 模型和其预测结果。

Step3. 子图生成。通过调整图掩码生成能对 GNN 模型预测结果产生最大影响的子图。

Step4. 特征选择。使用特征掩码选择子图中真正重要的节点特征集。

Step5. 目标定义。给定一个节点 v , 目标是识别出子图 $G_s \subseteq G_c$ (c 为类别)和相关的特征 X_s , 这些特征对 GNN 的预测 $y^{\wedge}$ 很重要。

Step6. 优化函数制定。使用互信息将“重要性”概念形式化,制定优化公式

$$\max_{G_s} MI(Y, (G_s, X_s)) = H(Y) - H(Y | G = G_s, X = X_s) \quad (4)$$

Step7. 解释结果。输出的子图为 GNNExplainer 提供的解释结果,代表 GNN 模型的预测过程中真正对预测结果产生影响的节点子集。

经过以上步骤, GNNExplainer 从推荐路径图提取真正影响推荐结果的子图,这个子图称为解释子图(即图模式),作为对推荐结果的解释。

4 可视分析系统

4.1 设计目标

为了支持第 2.3 节所述的系统分析任务,本文

制定了系统的设计目标。

G1. 从特征和特征组合出发, 分析用户和用户群的偏好模式, 并与空间分布的一致性进行对比。

G2. 基于推荐嵌入结果分布, 提供用户群偏好模式的对比分析, 并支持细节探索。

G3. 基于观影数据和目标电影, 提供推荐结果和解释结果的细节展示, 以及推荐的上下文信息。

G4. 基于解释模型提取图模式, 支持多粒度的图模式挖掘, 并提供图模式的对比探索。

G5. 分析图模式, 评估解释模型的可靠性。

4.2 系统总体流程

综上所述, 本文采用 Vue 框架和 Flask 框架进行前后端开发, 并利用 D3.js 设计视图, 实现了图 8 所示的可视分析系统, 并提供了系统演示视频^①。该系统由特征分数概览视图、历史偏好分析对比视图、解释模式探索分析视图以及解释评估视图构成。特征分数概览视图基于推荐模型的嵌入表示结果, 用来探索用户嵌入分布的特点, 它包括图 8a 所示特征列表视图, 图 8b 所示特征关注视图和图 8c 所示嵌入分布视图; 历史偏好分析对比视图基于用户的观影数据, 展示用户和用户群的观影偏好模式, 它包括图 8d 所示用户群偏好分析对比视图和图 8e 所示用户偏好分析对比视图; 解释模式探索分析视图用来展示解释模型提取到的图模式结果, 并支持图模式的细节展示, 它包括图 8f 所示模式对比探索视图和图 8g 所示推荐结果溯源视

图; 图 8h 所示解释评估视图用于评估解释的可靠性, 并提供可视验证。

分析人员使用该可视化系统进行数据分析步骤如下。

Step1. 在图 8a 中选择感兴趣的特征和特征组合, 这些特征将用于后续的用户群分析。

Step2. 在图 8b 中查看关注所选特征的用户群, 分析人员可以了解与特定特征相关的用户群体。

Step3. 在图 8c 中查看用户群在嵌入空间中的位置分布, 有助于分析人员观察用户群之间的相对位置关系。

Step4. 在图 8d 中, 分析用户群的观影偏好模式, 同时, 该系统还支持对用户的观影偏好模式进行探索。分析人员可以从图 8c 出发, 选择感兴趣的用户群, 通过扩散操作进行更详细的观察, 联合用户关于类型、导演、演员的相关性分数进行分析, 滑动条用来筛选感兴趣用户。

Step5. 在图 8e 中显示用户的观影信息、电影海报以及其他详细文本信息。

Step6. 图 8f 用于对感兴趣用户和用户群的图模式进行解释分析。

Step7. 图 8g 展示推荐结果和解释结果的溯源信息。

Step8. 图 8h 提供对解释的评估验证。

为了更好地理解该可视化系统的分析流程, 图 9 展示了系统的主要流程。

4.3 系统视图

4.3.1 特征分数概览图

根据目标 G1, 该视图由特征列表视图, 特征关注视图、嵌入分布视图组成, 如图 8a~图 8c 所示,

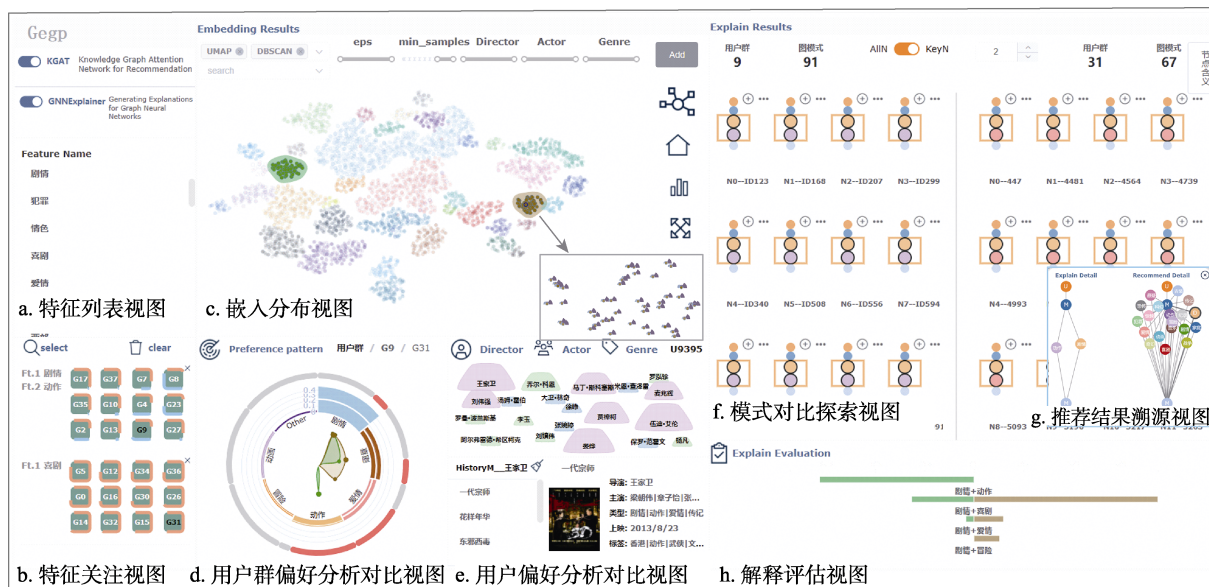


图 8 系统总体视图

① <https://cstr.cn/31253.11.sciencedb.j00188.00038>

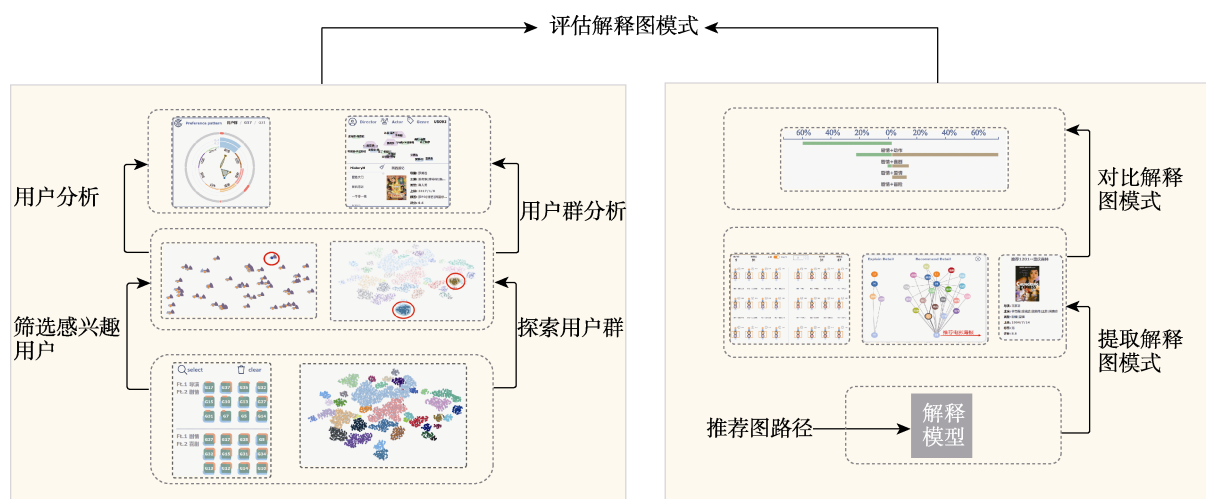


图9 系统分析流程

通过列表图、矩形图、散点图和环形图来实现。分析人员可以在图 8a 中选择感兴趣的特征或联合多个特征,并在图 8b 中查看对该特征组合关注的用户群。在图 8b 中,由于矩形图的区域面积有限,因此为了展示多种特征的相关性重要程度,系统对矩形图的外框进行了设计,将矩形图外框等分为 X 块,其中 X 代表所选特征组合的特征数量,并使用线条颜色编码特征,线条环绕矩形图展示。特征相关性分数的大小通过线条长度进行编码,矩形图内部文字表示用户群编号。

在选择感兴趣的用户群后,将在图 8c 所示嵌入空间分布图中突出显示它们的空间位置。图 8c 是通过推荐的高维嵌入降维聚类得到的,并允许分析人员针对感兴趣的用户群进行缩放观察,查看用户的类型、演员和导演相关性分数,并且还可以只对某种相关性指标进行深入研究。通过拖动滑动条来筛选区间数据,可以更细粒度地探索感兴趣的用户。

4.3.2 偏好分析对比图

根据目标 G2,偏好分析对比图由用户群偏好分析对比视图和用户偏好分析对比视图组成,如图 8d 和图 8e 所示。图 8d 利用条形图、雷达图和环形图实现。条形图展示不同用户群对不同特征的关心程度,高度越高,关心程度越大。雷达图用于提取用户群关注排名前三的特征,雷达图顶点越大,关心程度越大,雷达图可以很好地呈现不同用户群关注特征的覆盖情况。最外圈的环形图,用于突出显示用户群关注特征的差异情况,红色外环占比越多说明差异越大,在每段弧圈中,红色部分在前且红色弧越长,代表前一个用户群对该特征的关心程度更为突出,反之亦然。图 8e 由气泡

图、历史观影信息和电影细节信息组成。分析人员可以通过气泡图对感兴趣用户进行深入研究。气泡图利用多边形编码展示用户观影所关注的导演、演员和类型等节点类型信息,其中多边形的大小反映相关性分数的大小,相当于关注程度的表示;通过点击气泡图中的演员实体,可以呈现用户观看的由该演员参演的电影列表,进而可以在电影细节信息图中查看电影的海报和详细信息。

4.3.3 解释模式探索分析图

解释模式探索分析图由推荐结果溯源图和模式对比探索图组成,如图 8g 和图 8f 所示。

根据目标 G3,推荐结果溯源图利用可伸缩的力导图来展示目标用户与推荐结果间的图路径,节点越大,代表其所包含的实体节点越多。通过点击感兴趣的节点,可以展开该类型节点所包含的详细实体信息。例如,观看了哪部电影,属于哪种类型,由哪位导演和演员参与等更细粒度的信息。

根据目标 G4,模式对比探索图由列向圆组成。分析人员在嵌入分布视图选择感兴趣的用户群,利用解释模型提取用户推荐图中的重要图模式,为了直观地显示不同用户群的图模式差异,系统设计了列向圆视图。该视图可以对提取到的图模式进行整体概览,并可以调整图模式中重要节点的个数。点击“KeyN”按钮,可框选用户群的重要节点,帮助分析人员快速地理解用户群的推荐模式。在列向圆中,节点颜色代表节点类型,在列向圆的右上方,有推荐详细按钮和解释详细按钮。点击推荐详细按钮,可展示推荐电影的细节信息;点击解释详细按钮,可展示推荐结果溯源图。图 8g 为一个面板,该面板被分为左右 2 部分,右边展示的是推荐详细图,即目标用户与推荐结果之间的

完整路径图; 左边展示的是解释详细图, 即通过 GNNExplainer 从推荐详细图中提取到的重要图模式. 采用左右对比的设计能够提高用户对于解释的理解.

4.3.4 解释评估图

根据目标 G5, 解释评估图如图 8h 所示. 为了能够对解释模式进行定量评估, 并验证用户群的解释模式的可靠性, 系统设计了双向条形图来展示提取到的某种解释图模式在所有解释图模式中的定量占比, 同时通过与偏好分析对比图进行联合分析, 可以查看其解释图模式与历史偏好模式的一致性, 从而评估解释图模式的可靠性.

5 案例和结果分析

5.1 案例 1: 面向单个用户解释的探索

案例 1 面向用户群级别, 探究用户解释图模式和观影偏好的一致性.

开发人员在寻找感兴趣的用户时, 往往根据特征进行过滤筛选, 因此案例 1 针对“演员”指标进行研究. 随机选择用户 U_{206} , 在用户偏好细节图 10a 中发现气泡图面积都较小, 即对“演员”这一特征的关注程度都较低, 查看该用户观影中相对最关注的演员“梁朝伟”如图 10b, 发现他参演电影的出现次数也仅为 2, 进一步证明该用户不关注演员, 观影历史中演员的重复率不高.

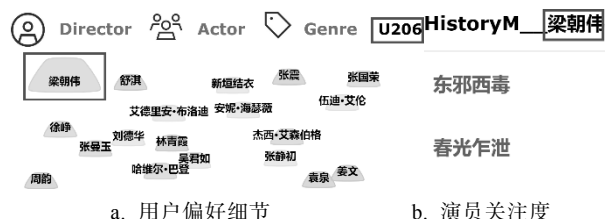


图 10 用户 U_{206} 细节信息展示

接着滑动“演员”相关性过滤条, 随机选择一位演员相关性分数较高的用户 U_{245} . 在图 11a 中发现多个气泡图面积均较大, 证明其对“演员”的关注程度较高, 点击演员“莱昂纳多”, 在图 11b 中发现该用户观看了“莱昂纳多”参演的 8 部电影.

通过以上信息可以看出, 用户 U_{245} 对“演员”的关注程度较高, 那么猜想利用解释模型提取到的解释图模式是否也捕捉到了“演员”这一特征, 为用户进行解释. 接着进行实验来验证这一猜想.

在图 12a 所示的列向圆中, 橙色节点、蓝色节点和浅蓝色节点是固定的, 分别代表用户、已观看

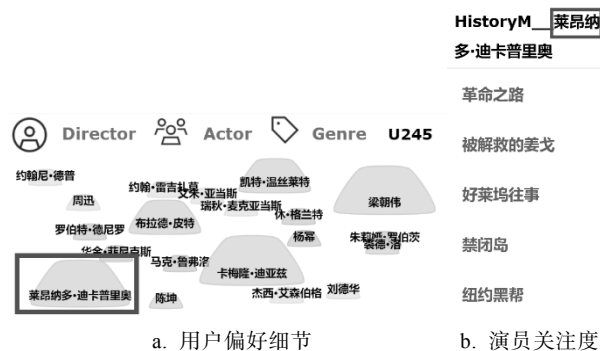


图 11 用户 U_{245} 细节信息展示

电影和推荐电影. 由解释模型提取的解释模式发现, 其为 U_{245} 推荐的 TOP20 部电影, 图模式中关键节点大部分颜色为红色. 为了探究红色节点的具体含义, 可以在图 12a 点击红色框内的“⊕”图标, 点击之后如图 12b 所示面板.

图 12b 所示面板分为左右 2 个部分, 左侧为解释详细图 (Explain Detail), 右侧为推荐详细图 (Recommend Detail). 在解释详细图中发现该红色节点代表 A, 即“演员”.

接着, 通过直方图 12d 对图模式进行统计分析, 在用户 U_{245} 的推荐电影中, 70% 的电影成功地捕捉到了“演员”的重要性, 这验证了猜想的正确性, 推荐模型在为该用户推荐电影时是根据“演员”来推荐的.

接着展开细节探索, 在图 12b 面板右侧的推荐详细图中点击“M”节点, 查看推荐电影, 如图 12c 所示电影“猫鼠游戏”, 包含演员“莱昂纳多”, 这 and 用户观看过多部“莱昂纳多”的电影 (图 11) 保持一致, 验证了用户解释图模式和观影偏好是一致的.

案例 1 证明了解释模型提取到的用户的图模式, 与用户的观影偏好是一致的.

5.2 案例 2: 面向用户群空间分布解释的探索

案例 2 面向用户群级别, 探究用户群的空间分布与用户群的观影偏好是否一致.

图模式的级别不应局限于用户级别, 系统将其扩大到用户群级别, 案例 2 研究用户群空间分布和历史观影偏好是否一致.

选择空间分布靠近的用户群, 如图 13a 所示, 猜想其观影偏好应该是相似的, 接着去验证这一猜想. 在偏好分析对比图 13b 中, 可以看到中间雷达图部分中 2 个用户群对应 2 个雷达图, 而雷达图是完全重叠的, 说明 2 个用户群的观影偏好是一致的; 通过雷达图中顶点的大小, 可以看到用户群关注的前 3 个电影类型分别都是“剧情、爱情和喜剧”.

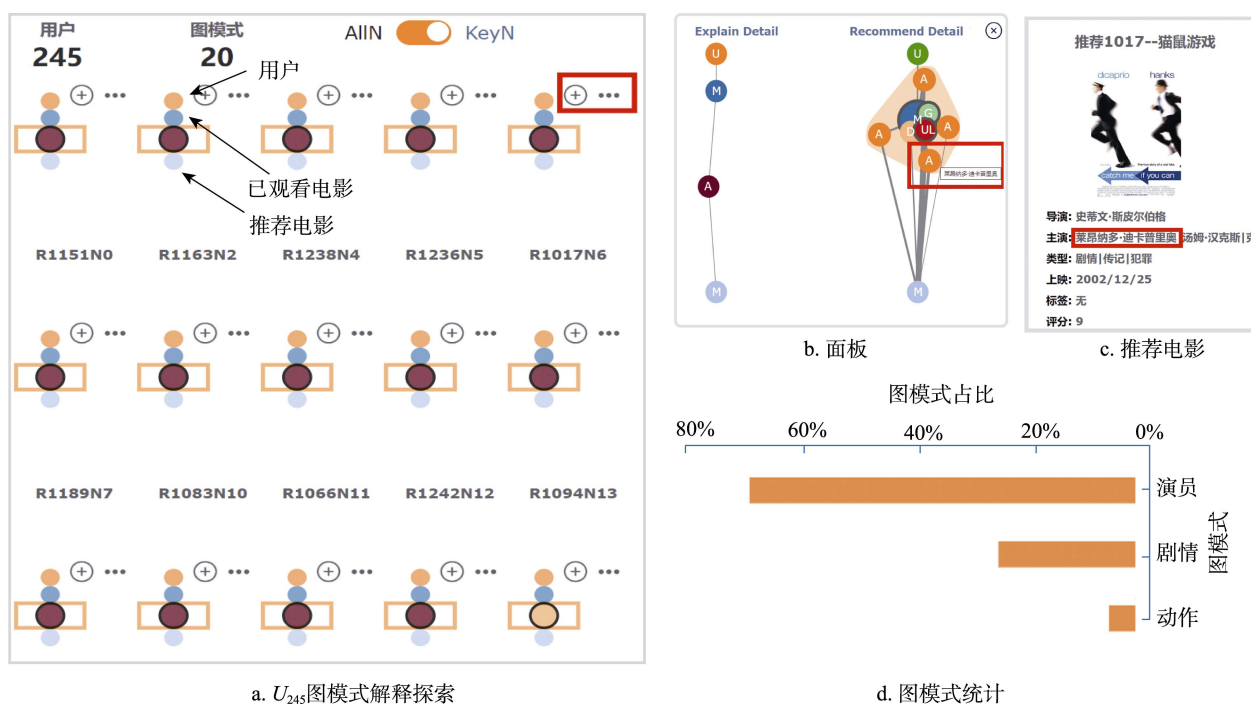
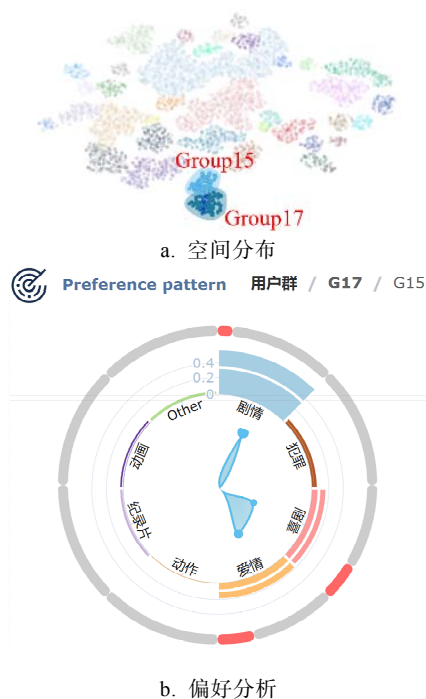
图12 用户 U_{245} 解释信息展示

图13 观影特征相似用户群

由此得到,用户群观影偏好相似对应用户群空间位置分布靠近.接着又探索了一系列空间位置接近的用户群,可视化结果显示这些用户群都具有相似的观影偏好.

另外,系统探索了空间分布远离的用户群,是否在观影偏好上存在较大差异性.选择“动作”特征,选择相关性分数排名最前和最末的2个用户

群,其空间分布如图14a所示,发现用户群空间位置分布原理,猜想其观影偏好应该是有较大差异的,接着在图14b偏好分析对比图中验证这一猜想,可以看到雷达图不重叠,且红色外环部分显示的差异性较大;从雷达图的顶点大小,得知用户群Group8关注“动作”电影,而Group3关注“动画”电影.接着又探索了一系列空间分布远离的用户群进行多次验证,可视化结果显示用户群观影偏好都有较大差异.

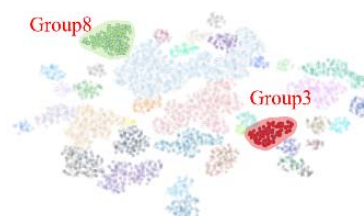
5.3 面向用户群解释的探索

第5.2节的案例分析验证用户群空间分布和观影偏好模式是一致的,本节在用户群级别,通过单个用户群(案例3)、2个用户群(案例4)和多个用户群(案例5),进一步验证解释模型提取到的用户群解释图模式和观影偏好的一致性.

5.3.1 案例3: 单个用户群解释探索

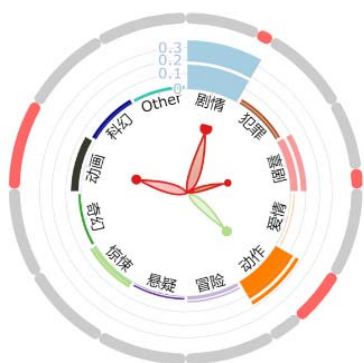
案例3对单个用户群的图模式进行探索.随机选择用户群,从偏好分析图15a中查看该用户群观影中关注的电影类型依次是“剧情、爱情和喜剧”.接着利用图15b所示解释模型提取解释模式,设置关键节点个数为3时,发现大多数用户的解释图模式都为“剧情、爱情和喜剧”,接着又探索了一系列用户群并进行多次验证,均证明了解释图模式和用户群偏好分析是一致的.

案例3证明了解释模型提取到的图模式,能够呈现推荐模型对用户的推荐偏好.



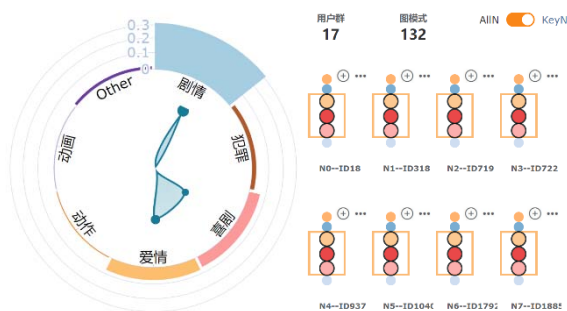
a. 空间分布

Preference pattern 用户群 / 8 / 3

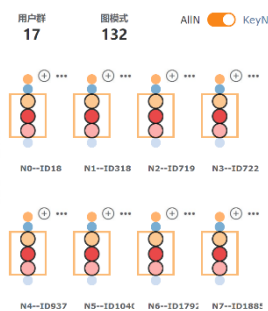


b. 偏好分析

图 14 观影特征差异用户群



a. 偏好分析



b. 图模式解释探索

图 15 单个用户群图模式探索

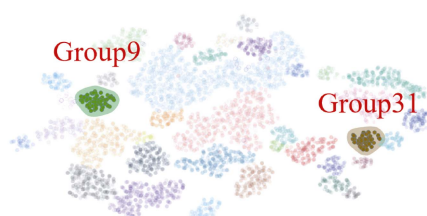
5.3.2 案例 4: 用户群对比探索

在案例 3 的基础上, 通过案例 4 展开用户群图模式对比探索.

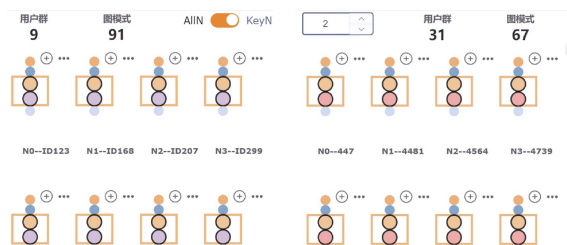
随机选择如图 16a 所示空间位置分布远离的用户群, 并利用图 16b 所示解释模型提取解释模式, 能够直观地发现 2 个用户群图模式中第 2 个关键点的颜色是不同的, 在解释详细图中发现橙色节点为“剧情”类型, 差异性体现在紫色节点代表的“动作”和粉色节点代表的“喜剧”之间.

接着, 通过如图 16c 所示直方图对图模式进行统计分析, 绿色的直方图柱显示用户群 Group9 中 60% 以上的用户关注“剧情和动作”电影, 棕色的直方图柱显示用户群 Group31 中 80% 以上的用户关注“剧情和喜剧”电影.

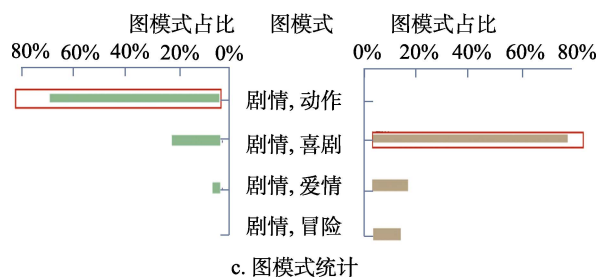
进一步, 对解释模型提取到的图模式可靠性



a. 空间分布



b. 图模式解释探索



c. 图模式统计

图 16 用户群观影模式对比探索

进行验证, 从图 17 偏好分析对比图得知, 用户群 Group9 在观影时关注“剧情和动作”电影, 用户群 Group31 观影时关注“剧情和喜剧”电影, 且雷达图展示用户群在类型“动作”和“爱情”的差异性, 这与通过图模式对比得出的结论是一致的.

Preference pattern 用户群 / 9 / 31

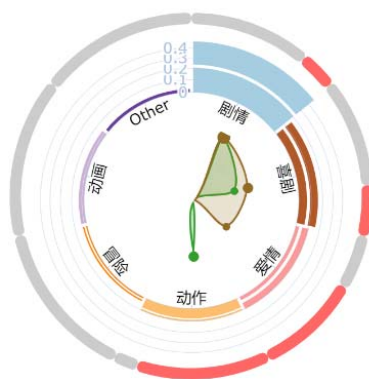


图 17 偏好分析

案例 4 展现了推荐模型在 2 个用户群的推荐偏好上的差异, 且与观影偏好是一致的.

5.3.3 案例 5: 多个用户群探索

在案例 4 的基础上, 通过案例 5 对多个用户群

的图模式进行对比探索。

随机选择分布远离的多个用户群,其空间位置如图 18a 所示;并利用解释模型提取解释模式,如图 18b 所示,能够直观地发现图模式中第 2 个关键点的颜色是不同的,在解释详细图中发现橙色节点为“剧情”类型,差异性体现在粉色节点代表的“喜剧”类型,红色节点代表的“爱情”类型,紫色节点代表的“动作”之间。

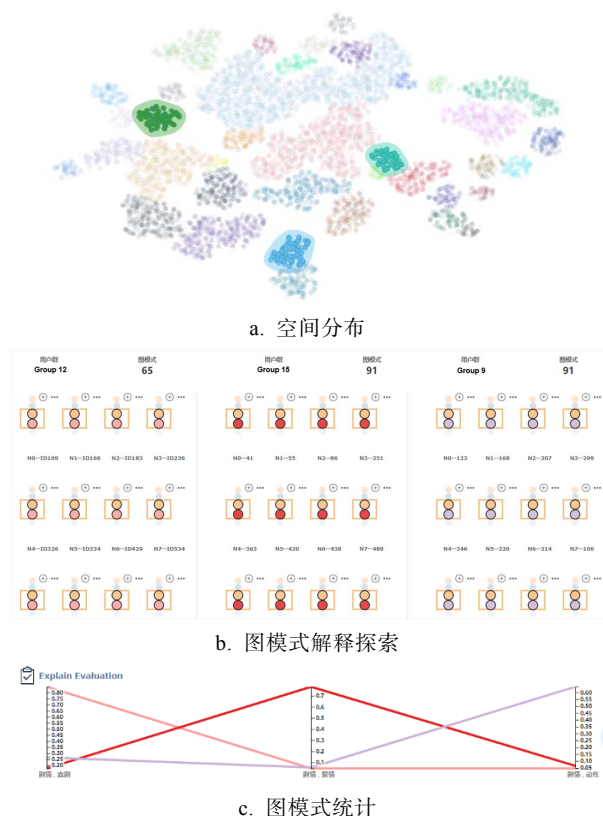


图 18 多个用户群观影模式探索

接着通过图 18c 所示平行坐标图对图模式统计分析,粉色折线表示的用户群 Group12 中,超过 80% 的用户都关注“剧情和喜剧”电影;红色折线表示的用户群 Group15 中,超过 75% 的用户都关注“剧情和爱情”电影;紫色折线表示的用户群 Group9 中,超过 63% 的用户关注“剧情和动作”电影。

进一步,对解释模型提取的图模式可靠性进行验证,从图 19 中得知,用户群 Group12 观影时关注“剧情和喜剧”电影,用户群 Group15 观影时关注“剧情和爱情”电影,用户群 Group9 在观影时关注“剧情和动作”电影,这与通过解释模型提取的解释模式得出的结论是一致的。

案例 5 展现了推荐模型在多个用户群上的推荐差异,且与观影偏好一致。

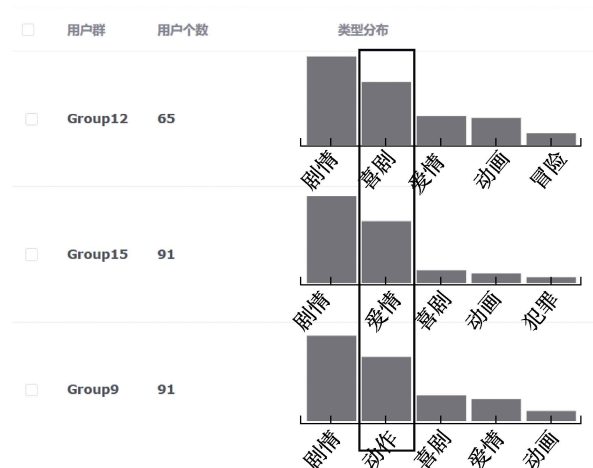


图 19 多个用户群偏好分析

5.4 用户所属用户群解释探索

案例 6 对用户 and 所在用户群之间的关系进行探索. 随机选择用户群, 打开相关性分数指标, 发现该用户群的类型特征分数均较高, 如图 20 所示.

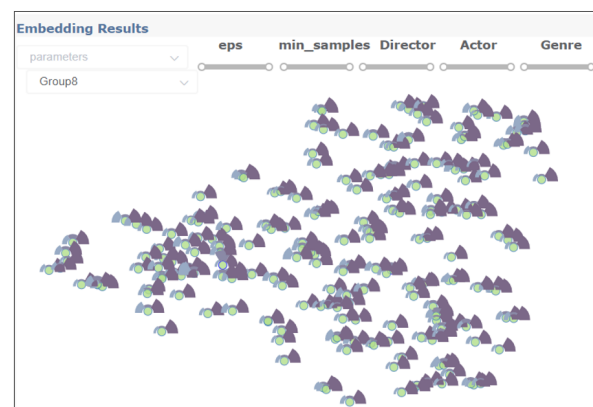


图 20 单个用户群图模式探索

对用户群具体偏好分析进行研究,统计其在观影类型、导演和演员方面的关注分布.从图 21 所示偏好分析发现,该用户群偏好“剧情和动作”电影,“史蒂文·斯皮尔伯格”等导演,“周星驰、刘德华”等演员。

在用户群中对个体用户进行研究,随机选择用户,从如图 22 所示用户偏好细节发现,所选用户的偏好分布和用户群的偏好分布一致,将鼠标悬浮至文本上显示全名,发现其钟爱“动作”类型电影,“杜琪峰、周星驰、韦家辉”等动作片导演,“成龙、李连杰、刘青云”等动作片演员,这证明了个体用户与所在用户群的偏好是一致的。

案例 6 展现了个体用户与所在用户群观影偏好是一致的。

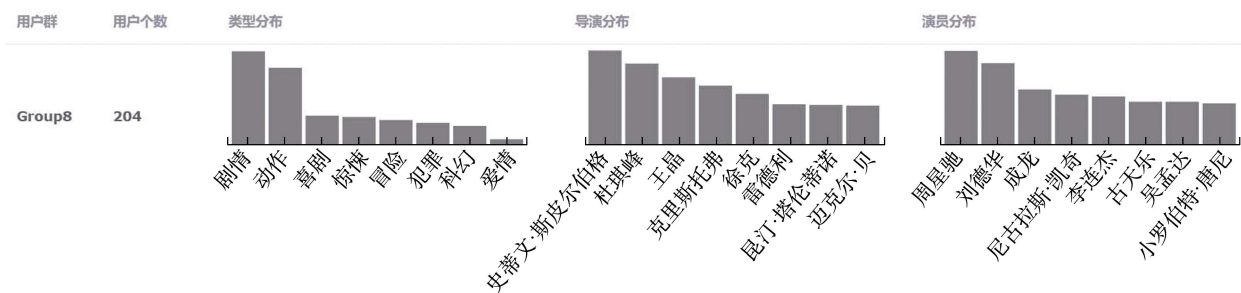


图 21 单个用户群偏好分析

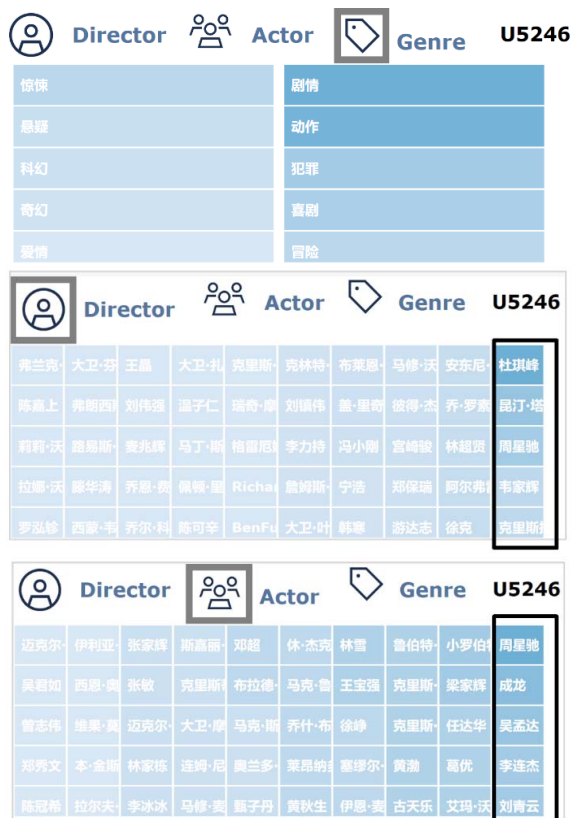


图 22 用户偏好细节

6 用户实验

本次实验招募了 15 名志愿者, 年龄为 20~25 岁, 其中 9 名男性和 6 名女性. 所有志愿者均为符合该系统的目标用户要求, 即有一定机器学习背景的人员. 在此基础上, 其中有 10 位志愿者的研究方向是数据可视化, 对可视分析十分熟悉; 有 2 位志愿者的研究领域是推荐系统; 另外 3 位志愿者了解并使用过静态可视化图表.

每个志愿者的评估平均需要 30 min. 首先, 本文向 15 名志愿者简要介绍了系统的主要目的, 并通过简要讲解和 5 min 的系统演示视频, 让他们了解其中各个视图的内容和视图间的联动; 然后根

据任务要求进行交互并回答问题, 期望他们在任务体验的过程中能够对推荐模型推荐的细节得到深入的理解, 并且能够对推荐结果的差异性原因得到可靠的解释启发. 以下为用户实验的 4 个任务, 并利用 0~5 分来记录他们对于系统对任务的支持程度:

(1) 从特征出发, 分析人员能否分析出关注该特征的用户群, 用户群的偏好模式与嵌入空间分布是否一致, 探索不同用户群偏好模式的相同性和差异性.

(2) 系统是否能够从用户群具体到用户, 提供过滤手段来支持分析人员对感兴趣的用户进行分析, 并提供对不同用户的观影细节分析和对比探索.

(3) 基于用户的观影数据和推荐目标电影, 系统是否支持使用解释模型对推荐结果进行解释, 以及解释的呈现是否多层次且直观易理解.

(4) 系统是否支持多粒度的解释, 从单个用户、用户群和多个用户群的多粒度出发进行解释的探索, 以及解释图模式之间的对比分析.

用户研究的结果如图 23 所示.

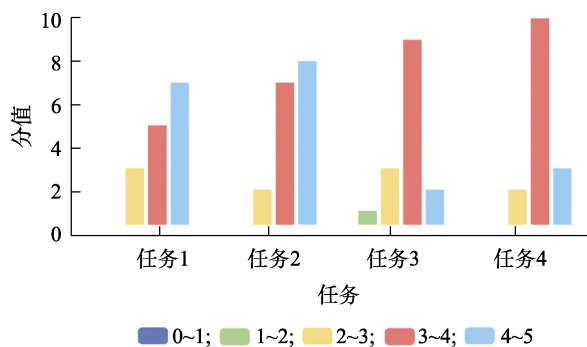


图 23 用户实验结果

可以观察到, 大多数志愿者对该系统的各项任务表现持肯定态度. 与此同时, 本文也对一些评分较低的志愿者进行了访谈, 分析了他们对系统的不满之处, 并且对他们的建议进行了认真考虑, 对系统进行了一些改进.

其中3名志愿者 P_1, P_2, P_3 提到, 在任务1中, 选择特征组合时, 他们需要手动滑到当前最新的选择结果上. 他们希望系统能够在保存历史记录的同时, 提供删除的操作. 本文认为该建议是可取的, 进行了如图24所示的改进.



图24 系统视图建议

P_6 提到, 他期望系统能够对图8g中力导引图节点上的文本进行提示. 虽然文本U很容易理解为用户, 但对于文本UL, 却一时间很难确定其代表的是哪类节点. 因此, 系统增加了提示按钮, 旨在更快地帮助用户理解节点的含义, 如图25所示.



图25 增添节点含义提示信息

7 基于推荐指标的图模式定量评估

本节基于推荐指标来对图分类和解释图模式的可靠性进行定量评估.

本文设计了对比实验来对可靠性进行定量评估. 对比实验的具体步骤如下.

Step1. 利用可视系统提取解释模式.

Step2. 提取符合解释模式的推荐电影.

Step3. 随机替换并调整训练集, 重新进行模型训练.

Step4. 利用新的推荐结果, 计算推荐模型在精确率、召回率、准确率和归一化折损累计增益4种常用推荐系统指标上的性能.

对比实验的整体流程, 以Last-FM数据集为例进行描述. 首先, 在推荐模型KGAT训练下, 得到所有用户的嵌入表示; 然后, 通过推荐音乐的嵌入

表示对用户的嵌入表示进行优化, 得到良好的用户群分类结果; 其次, 利用本文提出的辅助节点优化方法, 得到优化后的用户与推荐音乐之间的推荐图路径. 在本文的可视系统的帮助下, 利用GNNExplainer从推荐图路径中提取重要的子图, 为解释图模式. 过滤出符合解释图模式的推荐电影随机替换并调整训练集, 重新训练推荐模型, 得到新的推荐音乐序列, 计算推荐模型在精确率、召回率、准确率和归一化折损累计增益4种常用推荐系统指标上的分数. 对比实验的结果如表4所示.

表4 不同方法在不同数据集上性能对比实验结果

数据集	推荐指标	文献[12]	本文
豆瓣电影	精确率	0.052 0	0.074 8
	召回率	0.017 0	0.025 5
	准确率	0.293 0	0.362 6
	归一化折损累计增益	0.241 5	0.389 1
Last-FM	精确率	0.105 2	0.192 6
	召回率	0.186 7	0.237 3
	准确率	0.354 4	0.368 4
	归一化折损累计增益	0.305 2	0.483 1

从实验结果中可以清楚地观察到, 在本文开发的系统的帮助下, 推荐模型在4个指标上的推荐性能大多得到了提升. 这证明了解释的可靠性和系统的有效性.

8 结 语

本文将推荐场景与可解释GNN结合起来, 通过将推荐问题转换为图分类问题, 利用可解释GNN对推荐场景进行解释. 此外, 本文提出的可视分析系统, 能够从实例级和组群级出发, 支持推荐场景下的多粒度解释, 以更好地理解GNN的推荐模式; 通过对比实验, 验证了本文方法的可靠性, 证明了系统的有效性.

未来, 将继续扩大系统的延展性, 在更多的数据集上进行探索, 提供更灵活的交互方式支持深入探索解释图模式, 帮助分析人员更直观地理解推荐结果, 增强推荐场景的可解释性. 同时, 期望利用其他的可解释GNN对推荐场景进行解释, 特别是在类别上, 对推荐场景中的GNN进行解释, 增强可理解性. 此外, 如何基于真实数据对可解释模型结果进行改进训练集以外的客观验证, 也是未来将要进行的工作.

参考文献(References):

- [1] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks[OL]. [2023-08-17]. <https://arxiv.org/pdf/1609.02907.pdf>
- [2] Hamilton W L, Ying Z, Leskovec J. Inductive representation learning on large graphs[C] //Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: Curran Associates Inc, 2017: 1025-1035
- [3] Veličković P, Cucurull G, Casanova A, *et al.* Graph attention networks[OL]. [2023-08-17]. <https://arxiv.org/pdf/1710.10903.pdf>
- [4] Sun Kanggao. Research and implementation of visual interactive recommendation algorithm based on model ensemble[D]. Hangzhou: Zhejiang University of Technology, 2018(in Chinese)
(孙康高. 基于模型融合的可视交互推荐方法研究与实现[D]. 杭州: 浙江工业大学, 2018)
- [5] Luo D S, Cheng W, Xu D K, *et al.* Parameterized explainer for graph neural network[C] //Proceedings of the 34th International Conference on Neural Information Processing Systems. New York: Curran Associates Inc, 2020: Article No.1646
- [6] Ying Z, Bourgeois D, You J X, *et al.* GNNExplainer: Generating explanations for graph neural networks[C] //Proceedings of the 33rd International Conference on Neural Information Processing Systems. New York: Curran Associates Inc, 2019: Article No.829
- [7] Yuan H, Tang J L, Hu X, *et al.* XGNN: Towards model-level explanations of graph neural networks[C] //Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM Press, 2020: 430-438
- [8] Gilmer J, Schoenholz S S, Riley P F, *et al.* Neural message passing for quantum chemistry[C] //Proceedings of the 34th International Conference on Machine Learning. New York: JMLR.org, 2017: 1263-1272
- [9] Zhou F, Cao C T, Zhang K P, *et al.* Meta-GNN: on few-shot node classification in graph meta-learning[C] //Proceedings of the 28th ACM International Conference on Information and Knowledge Management. New York: ACM Press, 2019: 2357-2360
- [10] Cen Y K, Zou X, Zhang J W, *et al.* Representation learning for attributed multiplex heterogeneous network[C] //Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM Press, 2019: 1358-1368
- [11] Fan W Q, Ma Y, Li Q, *et al.* Graph neural networks for social recommendation[C] //Proceedings of the World Wide Web Conference. New York: ACM Press, 2019: 417-426
- [12] Wang X, He X N, Cao Y X, *et al.* KGAT: knowledge graph attention network for recommendation[C] //Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM Press, 2019: 950-958
- [13] Bajaj M, Chu L Y, Xue Z Y, *et al.* Robust counterfactual explanations on graph neural networks[C] //Proceedings of the Thirty-Fifth Conference on Neural Information Processing Systems. New York: Curran Associates, Inc, 2021: 5644-5655
- [14] Wang X, Wu Y X, Zhang A, *et al.* Towards multi-grained explainability for graph neural networks[C] //Proceedings of the Thirty-Fifth Conference on Neural Information Processing Systems. New York: Curran Associates Inc, 2021: 18446-18458
- [15] Tan J T, Geng S J, Fu Z H, *et al.* Learning and evaluating graph neural network explanations based on counterfactual and factual reasoning[C] //Proceedings of the 31st ACM Web Conference. New York: ACM Press, 2022: 1018-1027
- [16] Wang X Q, Shen H W. GNNInterpreter: a probabilistic generative model-level explanation for graph neural networks[OL]. [2023-08-17]. <https://arxiv.org/pdf/2209.07924.pdf>
- [17] Sutton R S, McAllester D, Singh S, *et al.* Policy gradient methods for reinforcement learning with function approximation[C] //Proceedings of the 12th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 1999: 1057-1063
- [18] Yuan J, Chen C J, Yang W K, *et al.* A survey of visual analytics techniques for machine learning[J]. Computational Visual Media, 2021, 7(1): 3-36
- [19] Ingram S, Munzner T, Irvine V, *et al.* DimStiller: Workflows for dimensional analysis and reduction[C] //Proceedings of the IEEE Symposium on Visual Analytics Science and Technology. Los Alamitos: IEEE Computer Society Press, 2010: 3-10
- [20] Choo J, Lee C, Reddy C K, *et al.* Utopian: User-driven topic modeling based on interactive nonnegative matrix factorization[J]. IEEE Transactions on Visualization and Computer Graphics, 2013, 19(12): 1992-2001
- [21] Görg C, Liu Z C, Kihm J, *et al.* Combining computational analyses and interactive visualization for document exploration and sensemaking in jigsaw[J]. IEEE Transactions on Visualization and Computer Graphics, 2013, 19(10): 1646-1663
- [22] McInnes L, Healy J, Melville J. UMAP: uniform manifold approximation and projection for dimension reduction[OL]. [2023-08-17]. <https://arxiv.org/pdf/1802.03426.pdf>
- [23] Ester M, Kriegel H P, Sander J, *et al.* A density-based algorithm for discovering clusters in large spatial databases with noise[C] //Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. Palo Alto: AAAI Press, 1996: 226-231
- [24] Herzig R, Bar A, Xu H J, *et al.* Learning canonical representations for scene graph to image generation[C] //Proceedings of the 16th European Conference on Computer Vision. Heidelberg: Springer, 2020: 210-227
- [25] You Y N, Chen T L, Sui Y D, *et al.* Graph contrastive learning with augmentations[C] //Proceedings of the 34th International Conference on Neural Information Processing Systems. New York: Curran Associates Inc, 2020: Article No.488