

DOI: 10.3969/j.issn.1671-0673. 2026. 04. 015

多模态知识图谱注意力网络的实体推荐算法

蔡熠非¹, 瞿小龙², 钱 凯¹

(1.宜春学院, 江西 宜春 336000; 2.北京林业大学, 北京 100083)

摘要: 针对信息检索和推荐任务中存在的数据过载和多模态数据缺失等问题, 提出一种视觉模态生成增强的多模态知识图谱注意力网络(VG-MKGAN)。该模型通过多模态聚合器对文本、图像和结构化三元组信息进行联合建模, 并引入图注意力机制以动态捕捉用户与项目之间最相关的邻域语义信息。同时, 针对部分实体缺乏视觉模态数据的问题, 设计一种基于邻居节点视觉特征的模态生成方法, 实现视觉信息的有效补全。在中医药推荐数据集(TCM-MMKG)、电影数据集(MovieLens)和音乐数据集(Last-FM)上的实验结果表明, VG-MKGAN 较现有最优基线模型 MKGAT 在 TCM-MMKG 和 Last-FM 数据集上的 AUC 指标分别提高 2.7% 和 3.1%, $F1$ 值分别提高 11.2% 和 2.5%。

关键词: 多模态; 知识图谱; 推荐系统; 注意力网络

中图分类号: TP391.3

文献标识码: A

文章编号: 1671-0673(2026)04-0489-06

Entity Recommendation Algorithm Based on Multimodal Knowledge Graph Attention Network

CAI Yifei¹, QU Xiaolong², QIAN Kai¹

(1.Yichun University, Yichun 336000, China; 2.Beijing Forestry University, Beijing 100083, China)

Abstract: To address the issues of data overload and missing multimodal data in information retrieval and recommendation tasks, a visual-modality generation enhanced multimodal knowledge graph attention network (VG-MKGAN) is proposed. Textual, visual and structural triple information are jointly modeled via a multimodal aggregator, and a graph attention mechanism is incorporated to dynamically capture the most relevant neighborhood semantic information between users and items. Furthermore, to tackle the problem that certain entities lack visual modality data, a modality generation strategy based on the visual features of neighboring nodes is designed to effectively complete visual information. Experimental results on the traditional Chinese medicine recommendation dataset (TCM-MMKG), movie dataset (MovieLens) and music dataset (Last-FM) demonstrate that AUC metrics are improved by 2.7% and 3.1% respectively on the TCM-MMKG and Last-FM datasets, and $F1$ scores are increased by 11.2% and 2.5% correspondingly, compared with those of the best existing baseline model MKGAT.

Key words: multimodal; knowledge graph; recommendation system; attention network

传统基于内容的推荐方法^[1]通常将实体视为孤立的个体, 忽略了它们之间的语义关联和上下文关系^[2], 从而无法充分处理语义信息。协同过滤方法^[3-4]由于依赖大量的交互数据而面临冷启动挑

战, 在新用户和数据稀缺的场景下尤为突出。为突破这些限制, 研究者采用知识图谱(Knowledge Graph, KG), 通过结构化三元组来丰富实体特征^[5], 而多模态知识图谱(Multimodal Knowledge Graph,

收稿日期: 2025-08-26; 修回日期: 2026-03-05

基金项目: 江西省教育厅科学技术研究项目(GJJ180259)

作者简介: 蔡熠非(1996—), 男, 硕士, 主要研究方向为软件工程。Email: 568185638@qq.com

MMKG)则通过整合视觉、文本和属性数据^[6]扩展了这一能力。这些框架支持跨模态相关性分析,通过全面的实体表征和对用户-项目交互的情境理解增强推荐的个性化。目前的研究重点是将这些技术与情境信息(如社交网络、属性和图像等^[7-8])有效整合,以克服数据稀疏性并提高系统的推荐准确性。

然而,MMKG面临的一个关键挑战源于现实世界数据的不完整性,尤其是视觉信息缺失^[4]。为这些实体分配准确的视觉模态表征仍然具有挑战性,需要耗费大量的人工整理。现有的推荐系统难以处理多模态数据的不完整性,尤其是视觉模态信息的缺失。当前方法也忽视了多模态数据生成的目标导向性,未能系统地将模态与任务需求联系起来。

为解决这些问题,提出一种视觉模态生成增强的多模态知识图谱注意力网络(Visual-Modality Generation Enhanced Multimodal Knowledge Graph Attention Network, VG-MKGAN)。该模型利用多模态聚合器聚合实体的文本、图像和结构化数据,增强用户和项目的特征表示,并在聚合实体高阶邻居节点时,通过图注意力网络高效地选择对用户和项目关注度最高的邻居实体。此外,提出一种针对缺失视觉模态数据的数据增强方法,通过聚合缺失视觉模态实体周围一阶邻居节点的视觉模态来生成当前实体的视觉模态特征。将VG-MKGAN应用于中医药、电影和音乐推荐3个实际场景,实验结果表明,VG-MKGAN的性能优于当前最先进的同类方法和推荐系统基线模型。

1 问题定义

参考前人工作^[6-9],首先引入多模态知识图谱和用户项目交互矩阵的概念,然后介绍推荐场景中的问题定义。

1.1 多模态知识图谱

对多模态知识图谱进行如下定义:

$$\mathcal{G}_{\text{MMKG}} = \{\mathcal{E}, \mathcal{R}, \mathcal{A}, \mathcal{M}, \mathcal{T}_{\mathcal{R}}, \mathcal{T}_{\mathcal{A}}\} \quad (1)$$

式中: $\mathcal{E}$ 、 $\mathcal{R}$ 、 $\mathcal{A}$ 和 $\mathcal{M}$ 分别表示实体、关系、属性和属性值; $\mathcal{T}_{\mathcal{R}} = \mathcal{E} \times \mathcal{R} \times \mathcal{E}$ 和 $\mathcal{T}_{\mathcal{A}} = \mathcal{E} \times \mathcal{A} \times \mathcal{M}$ 分别表示关系三元组和属性三元组。例如:三元组 $(e_h, r, e_t) \in \mathcal{T}_{\mathcal{R}}$ 表示头节点实体 $e_h \in \mathcal{E}$ 和尾节点实体 $e_t \in \mathcal{E}$ 的关系为 $r \in \mathcal{R}$;三元组 $(e_h, a, m) \in \mathcal{T}_{\mathcal{A}}$ 表示实体 $e_h \in \mathcal{E}$ 具有属性值 $m \in \mathcal{M}$,其中 $a \in \mathcal{A}$ 为属性

类型。实体集 $\mathcal{E}$ 包含项目集 $\mathcal{V}$,这意味着 $\mathcal{V}$ 单映射到 $\mathcal{E}$ 且 $\mathcal{V} \subset \mathcal{E}$ 。

1.2 用户项目交互矩阵

在一个典型的推荐场景中,存在用户 $\mathcal{U} = \{u_1, u_2, \dots, u_M\}$ 和推荐项目 $\mathcal{V} = \{v_1, v_2, \dots, v_N\}$,其中 M 和 N 分别是用户和项目的数量。用户项目交互矩阵 $\mathbf{Y} = \{y_{uv} | u \in \mathcal{U}, v \in \mathcal{V}\}$,大小为 $M \times N$,定义如下:

$$y_{uv} = \begin{cases} 1, & u \text{ 和 } v \text{ 有互动;} \\ 0, & \text{其他.} \end{cases} \quad (2)$$

如果 y_{uv} 等于1,表示用户 u 和项目 v 之间有历史互动,比如点击、浏览和下载等;如果等于0,则表示未发生互动,但这并不意味着用户 u 对项目 v 不感兴趣。

1.3 模型定义

给定一个用户项目交互矩阵 $\mathbf{Y}$ 和一个多模态知识图谱 $\mathcal{G}_{\text{MMKG}}$,尝试预测用户 u 与之前未交互过的项目 v 的交互概率。因此,所提VG-MKGAN的目标是从正样本和生成的负样本中学习一个交互概率预测函数 $\mathcal{F}(\cdot)$,如下所示:

$$\hat{y}_{uv} = \mathcal{F}(u, v | \mathbf{Y}, \mathcal{G}_{\text{MMKG}}, \Theta) \quad (3)$$

式中: $\hat{y}_{uv} \in [0, 1]$ 表示向用户 u 推荐项目 v 的概率; Θ 表示 $\mathcal{F}(\cdot)$ 的模型参数。

2 推荐算法模型

VG-MKGAN框架如图1所示,包含5层。1)知识传播层将项目与MMKG中的实体对齐,并传播高阶邻居实体的知识;2)数据增强层基于视觉模态生成,为缺少视觉模态数据的实体生成视觉模态;3)多模态聚合器层聚合实体的文本、图像和结构表示,以在多模态聚合后获得增强的表示;4)在图注意力网络(Graph Attention Network, GAT)的引导下,GAT聚合层逐步聚合高阶邻域,以改进当前项目表示;5)预测层使用聚合表示来预测用户和项目的推荐概率。

2.1 知识传播层

首先,知识传播层的作用是将项目 v 映射到MMKG中的实体节点 e 。然后,由于知识图谱的连通性,实体 e 可以沿着关系逐层向外传播,从而获得其高阶邻居集合。例如,对于实体 e ,其 l 阶邻居是通过沿着MMKG中的关系边传播 $(l-1)$ 阶邻居得到的。实体 e 的 l 阶邻居定义为 $\mathcal{N}(e)^l$:

$$\mathcal{N}(e)^l = \{e_l | (e_h, r, e_l) \in \mathcal{T}_{\mathcal{R}}, e_h \in \mathcal{N}(e)^{l-1}\} \quad (4)$$

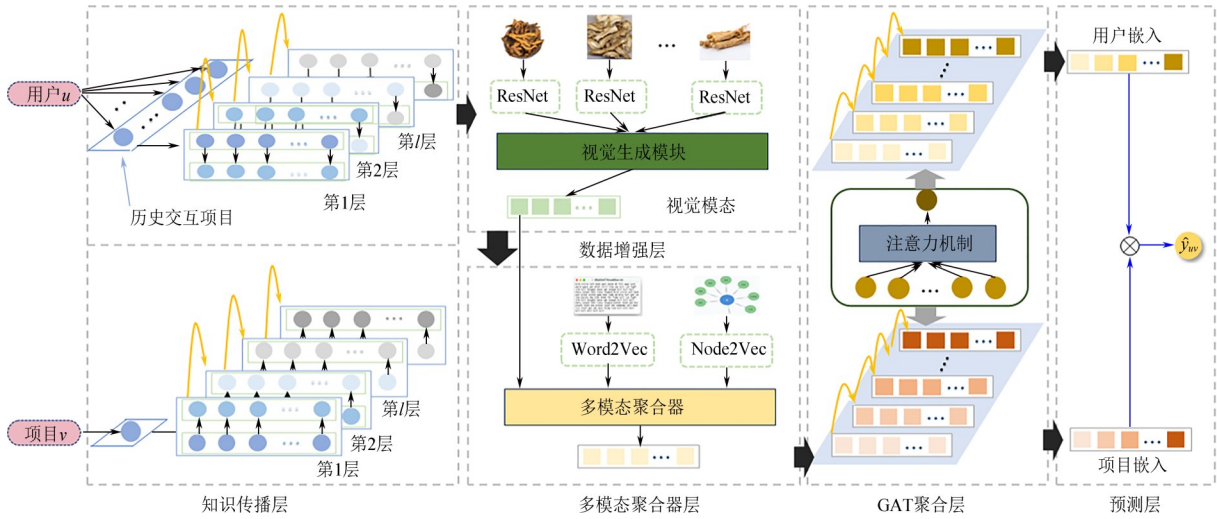


图1 VG-MKGAN框架

对于用户 u , $\mathcal{N}(e)^0$ 是用户 u 的历史交互项目集合, 也记为 H_u ; 对于项目 v , $\mathcal{N}(e)^0$ 是其对齐实体。

2.2 数据增强层

数据增强层的作用是利用视觉模态生成模块, 为缺乏视觉模态的实体生成相应的视觉特征。数据增强层框架如图2所示, 展示了视觉模态生成模块的核心思想。对于缺乏视觉数据的实体, 该方法通过挖掘其关联实体的视觉模态信息来实现

特征生成。具体而言, VG-MKGAN 利用关系网络中目标实体邻居节点所包含的视觉特征, 并通过特征传播和融合机制, 构建能够有效表示该实体的视觉模态。对于缺少图像数据的实体 e , 数据增强层会随机采样 k 个 e 的一阶邻居实体。然后, 将这些实体的视觉模态应用到生成策略中, 得到生成的视觉模态, 将此作为 e 的视觉模态特征。在 VG-MKGAN 中, 实现了均值生成策略。

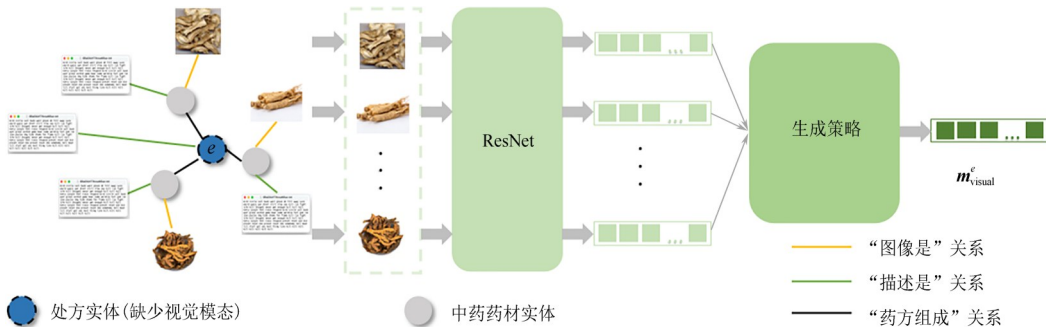


图2 数据增强层框架

均值生成策略将 k 个一阶邻居实体的视觉模态表示相加, 取平均值, 最后进行非线性变换:

$$V_{\text{mean}}^{\text{visual_gen}} = \sigma \left(\mathbf{W}_m \cdot \frac{1}{k} \left(\sum_{i=1}^k \mathbf{m}_{\text{visual}}^{e^i} \right) + \mathbf{b}_m \right) \quad (5)$$

式中: $\mathbf{W}_m$ 和 $\mathbf{b}_m$ 分别是可训练的变换权重和偏置; $\sigma(\cdot)$ 是激活函数 LeakyReLU; $e^i \in \mathcal{N}(e)^1$ 是缺失视觉模态实体 e 的一阶邻居实体; $\mathbf{m}_{\text{visual}}^{e^i}$ 是实体 e^i 的视觉模态表示; k 表示采样邻居的数量。

视觉生成是 VG-MKGAN 中非常重要的组成部分, 实验部分将讨论不同生成策略对视觉生成器 (visual generator) 性能的影响。通过视觉生成, 存在视觉数据缺失的实体 e 可以获得一个视觉模态表征 $\mathbf{m}_{\text{visual}}^e$, 定义如下:

$$\mathbf{m}_{\text{visual}}^e = \text{visual_gen}(e^1, \dots, e^k) \quad (6)$$

2.3 多模态聚合器层

多模态聚合器层的作用是将实体 e 的 3 个模态数据进行融合, 以增强实体 e 的表示建模。由于各模态数据的维度并不一致, 在多模态数据正式进入多模态聚合器之前, 首先进行主成分分析降维, 使各模态数据的维度与 MMKG 中节点的嵌入维度 d 对齐。在 VG-MKGAN 中, 实现了拼接聚合器。

该聚合器在应用非线性变换之前, 连接 (concat) 多模态数据的表示:

$$\mathbf{M}_{\text{concat}}^{\text{multi_agg}} = \sigma \left(\mathbf{W}_m \cdot \text{concat}(\mathbf{m}_{\text{textual}}^e + \mathbf{m}_{\text{structural}}^e + \mathbf{m}_{\text{visual}}^e) + \mathbf{b}_m \right) \quad (7)$$

式中: $\mathbf{m}_{\text{textual}}^e$ 是文本模态表征; $\mathbf{m}_{\text{structural}}^e$ 是实体 e 的节点嵌入。

通过多模态聚合器获取 MMKG 中实体 e 的多模态增强表示 e_m 的具体过程如下:

$$e_m = \text{multi_agg}(m_{\text{textual}}^e + m_{\text{structural}}^e + m_{\text{visual}}^e) \quad (8)$$

2.4 GAT 聚合层

GAT 聚合层为用户和项目建模高阶表示。对于每个实体, VG-MKGAN 从周围采样 k 个邻居。若实体 e 的邻居数量小于 k , 则进行重采样直至邻居数量等于 k ; 若大于 k , 则随机采样 k 个邻居。随后, 聚合器从最外层邻居开始, 逐层计算不同邻居的注意力权重, 然后聚合邻居表示向量, 直至到达第 0 层。最终获得用户和项目的增强嵌入表示向量。

利用多头自注意力机制实现了图注意力网络, 捕捉用户对不同项目的偏好。GAT 聚合器使用自注意力机制处理邻居项目的嵌入表示, 以计算注意力权重和加权嵌入向量:

$$\begin{cases} Q_i = K_i = V_i = e_{m,i}^u, i = 1, \dots, k; \\ z_i = \text{Softmax}\left(\frac{Q_i W_Q (K_i W_K)^T}{\sqrt{d_k}}\right) V_i W_V, i = 1, \dots, k. \end{cases} \quad (9)$$

式中: Q_i 、 K_i 和 V_i 分别表示查询(Query)、键(Key)和值(Value)向量; W_Q 、 W_K 和 $W_V \in \mathbf{R}^{d \times d_k}$ 表示可学习参数矩阵; d_k 表示投影后查询、键和值向量的维度; $z_i \in \mathbf{R}^d$ 表示第 i 个邻居项目的加权嵌入向量。

GAT 聚合器针对从外到内的聚合定义如下:

$$e_m^{l-1} = \frac{1}{k} \sum_{i=1}^k z_i \times e_{m,i}^l \quad (10)$$

最后, 将加权的嵌入向量在第 0 维上进行平均池化, 得到用户或项目的嵌入表示。最终的用户嵌入表示 $u \in \mathbf{R}^d$ 和项目嵌入表示 $v \in \mathbf{R}^d$ 如下:

$$\begin{cases} u = \text{GAT_agg}(e_{m,1}^{1,u}, \dots, e_{m,k}^{1,u}); \\ v = \text{GAT_agg}(e_{m,1}^1, \dots, e_{m,k}^1). \end{cases} \quad (11)$$

式中: $e_{m,i}^{1,u} \in H_u, i = 1, \dots, k$ 表示已从用户 u 的历史交互项目中选择并进行多模态 GAT 聚合的 k 个项目; $e_{m,i}^1 \in \mathcal{N}(e)^1, i = 1, \dots, k$ 是项目 v 的多模态 GAT 聚合一阶邻居集合。

2.5 预测层

在获得用户表征 u 和项目表征 v 之后, 采用预测函数来估计用户 u 和项目 v 之间的交互概率:

$$\hat{y}_{uv} = f(u, v) \quad (12)$$

式中, $f(\cdot)$ 表示向量点积运算。

最后, 采用贝叶斯个性化排名损失作为损失函数, 该损失函数假设观察到的正样本得分高于负样本。损失函数定义如下:

$$\mathcal{L} = \sum_{(u, v_i, v_j) \in O} -\ln(\text{Sigmoid}(\hat{y}(u, v_i) - \hat{y}(u, v_j))) + \lambda \|\Theta\|_2^2 \quad (13)$$

式中: O 表示训练集; (u, v_i) 表示正样本集; (u, v_j) 表示负样本集; λ 为 L2 正则化系数。

3 实验结果及分析

3.1 数据集简介

本文使用 1 个自建中医药数据集 (Traditional Chinese Medicine Multimodal Knowledge Graph, TCM-MMKG) 以及两个公开的数据集——电影数据集 MovieLens 和音乐数据集 Last-FM 进行实验。其中, MovieLens 和 Last-FM 数据集中的图像数据, 使用搜索引擎根据关键词进行爬取。3 个数据集的相关统计如表 1 所示。

表 1 数据集统计

数据集	用户	项目	交互记录	实体	关系类别	三元组数量
TCM-MMKG	5 816	16 163	52 572	18 519	8	107 802
MovieLens	41 849	4 828	1 813 301	65 801	19	145 406
Last-FM	44 548	19 123	3 209 509	23 781	51	152 921

在所有数据集的多模态数据处理中, 主要利用文本描述作为文本模态数据, 图像作为视觉模态数据。对于文本描述, 采用固定的 100 采样长度, 从 token 序列中删除标点符号和空格, 然后使用预训练的 768 维 Word2Vec 模型提取文本特征向量。对于图像数据, 使用预训练的 ResNet 模型提取 1 024 维的特征向量。VG-MKGAN 的一大亮点在于能够为缺失视觉数据的实体生成视觉模态表示。因此, 移除电影数据集 MovieLens 和音乐数据

集 Last-FM 中推荐项目 (分别为电影和音乐类型实体) 的视觉模态数据, 仅保留其他类型实体的视觉模态数据。

3.2 基线模型

将 VG-MKGAN 与先进的基线模型进行比较。这些基线模型可分为 3 类, 均未包含视觉生成模块。SVD^[8] 和 CKE^[10] 为基于协同过滤的方法; RippleNet^[9]、KGCGN^[5] 和 IGAT^[11] 为基于 KG 的方法; MIGCT^[12]、MKGCN^[13] 和 MKGAT^[4] 为基于 MMKG 的方法。

3.3 实验设置

使用PyTorch实现VG-MKGAN模型,并应用网格搜索来优化超参数。具体来说,使用Xavier初始化器初始化VG-MKGAN的参数,并使用Adam优化器优化VG-MKGAN。批量大小、学习率和 λ 系数分别使用以下数值: $\{64, 128, 512, 1024, 2048\}$ 、 $\{0.05, 0.01, 0.005, 0.001\}$ 和 $\{10^{-5}, 10^{-4}, \dots, 10^1, 10^2\}$ 。对于基准模型,根据其各自论文中报告的参数设置最优参数。

使用随机负采样策略生成负样本,并确保正样本与负样本的比例为1:1,然后随机划分80%的交互数据用于训练,剩余20%用于测试。此外,从训练数据集中,进一步留出20%作为验证数据集,这在确定最优超参数时起着至关重要的作用。实验评估VG-MKGAN在点击率(Click-Through Rate, CTR)预测方面的表现,采用曲线下面积(Area Under Curve, AUC)和F1值作为评估指标,以及采用Recall@ k 和NDCG@ k 来评估top- k 推荐和偏好排序

的有效性。 k 值介于 $\{1, 2, 5, 20, 50, 100\}$ 之间。每个实验进行5次以上,结果以平均值形式报告。

3.4 对比实验结果

比较并分析所提出的VG-MKGAN与基线模型的性能。所有方法在CTR预测和top- k 推荐方面的结果分别列于表2、图3和图4。所有方法的实验结果均通过选取最优超参数值获得。

表2 各模型的AUC和F1值结果对比

模型	TCM-MMKG		MovieLens		Last-FM	
	AUC	F1	AUC	F1	AUC	F1
SVD	0.546	0.515	0.875	0.861	0.743	0.639
CKE	0.721	0.637	0.927	0.874	0.792	0.704
RippleNet	0.710	0.640	0.976	0.927	0.796	0.723
KGCN	0.747	0.676	0.977	0.930	0.887	0.799
IGAT	0.786	0.693	0.972	0.925	0.867	0.825
MIGGT	0.817	0.714	0.978	0.926	0.859	0.819
MKGCN	0.832	0.725	0.981	0.935	0.886	0.832
MKGAT	0.839	0.731	0.980	0.931	0.894	0.850
VG-MKGAN	0.862	0.813	0.980	0.932	0.922	0.871

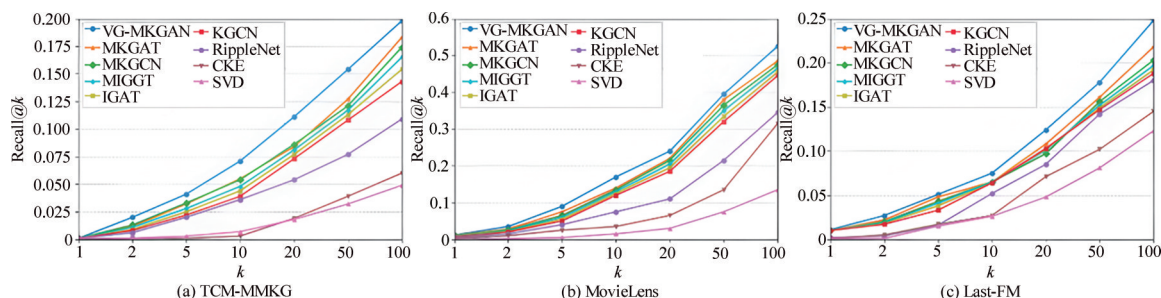


图3 各模型的Recall@ k 结果对比

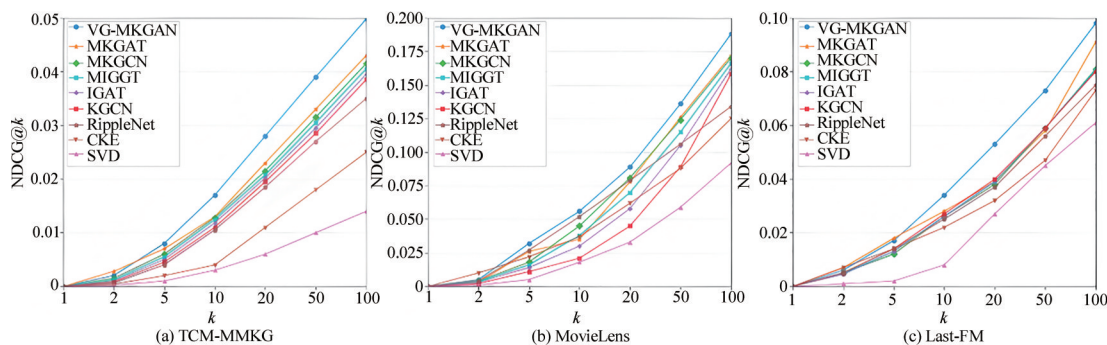


图4 各模型的NDCG@ k 结果对比

实验结果显示,SVD和CKE的性能明显较差,这表明基于协同过滤的方法难以有效利用上下文信息,从而导致推荐准确率较低。RippleNet、KGCN和IGAT都是基于KG的方法,其性能优于协同过滤方法。这凸显了知识图谱在利用上下文信息提升推荐准确率方面的有效性。

MIGGT、MKGCN、MKGAT和VG-MKGAN均是基于MMKG的方法,并且均表现出优于基于KG的

方法和协同过滤方法的性能。这凸显MMKG在提升推荐系统性能方面的重要作用。虽然MKGCN和MKGAT均利用包含视觉和文本信息的多模态数据,但MKGAT的表现优于MKGCN,这可能是由于GAT比GCN更加关注来自不同邻居节点的贡献,从而使实体表示建模更加准确。然而,MKGCN和MKGAT在处理视觉模态缺失方面都面临挑战,导致它们在具有缺失视觉数据的MMKG上表现不

佳。MIGGT虽通过模态独立感受野适配不同模态的传播特性,并借助采样式全局Transformer补充全局信息,但未针对MMKG的模态缺失问题优化,整体性能不及上述三者。VG-MKGAN利用视觉模态生成有效地解决了这个问题,并提升了其在不完整MMKG上的性能。此外,VG-MKGAN在聚合过程中利用图注意力网络来选择重要的邻居节点也是其性能优异的关键因素。

VG-MKGAN在MovieLens数据集上的表现不及MKGCN,这是因为MovieLens数据集中电影项目的邻居节点密集,文本节点信息丰富,因此无需借助视觉模态数据,即可完成对电影项目的精确建模。

4 结束语

针对不完整多模态知识图谱中实体视觉数据缺失的问题,提出一种用于推荐系统的视觉生成增强型多模态图注意力网络(VG-MKGAN)。具体而言,使用图注意力网络自适应地为每个邻居节点分配权重,从而学习并突出重要的邻居节点特征,减少无关节点或噪声节点的干扰。同时,使用多模态聚合器实现实体的结构、文本和视觉模态的聚合,丰富实体的特征表示。对于缺失的视觉模态,基于知识图谱的拓扑结构实现视觉模态生成模块。该模块使用一阶邻居节点为缺失视觉模态的实体生成视觉模态。在自建的中医药数据集TCM-MMKG以及公开的电影数据集MovieLens和音乐数据集Last-FM上进行了大量的实验,实验结果证实了VG-MKGAN的优越性。在未来的工作中,计划深入研究不同模态之间的更细粒度的融合和交互方法,以及进一步设计更高效的视觉生成器,以生成更准确的视觉模态。

参考文献:

- [1] 周春华,郭晓峰,沈建京,等.基于注意力机制的深度学习态势信息推荐模型[J].信息工程大学学报,2019,20(5):597-603.
- [2] 刘广腾,王峰,吴中博.下一个兴趣点推荐算法综述[J].计算机科学与探索,2025,19(7):1747-1770.
- [3] Hao X S, Zhu Y, Appalaraju S, et al. MixGen: a new multi-modal data augmentation [C] //Proceedings of the 2023 IEEE/CVF winter conference on applications of computer vision workshops. Piscataway, USA: IEEE, 2023: 379-389.
- [4] Sun R, Cao X Z, Zhao Y, et al. Multi-modal knowledge graphs for recommender systems [C] //Proceedings of the 29th ACM international conference on information & knowledge management. New York, USA: ACM, 2020: 1405-1414.
- [5] 胡瑞娟,刘海视,葛磊.融合知识图谱的多层级突发事件表示模型[J].信息工程大学学报,2023,24(6): 685-690.
- [6] 崔好丽,赵震.多模态知识图谱补全研究综述[J].计算机技术与发展,2025,35(8):10-18.
- [7] Wang H W, Zhang F Z, Zhang M D, et al. Knowledge-aware graph neural networks with label smoothness regularization for recommender systems [C] //Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining. New York, USA: ACM, 2019:968-977.
- [8] Koren Y. Factorization meets the neighborhood: a multifaceted collaborative filtering model [C] //Proceedings of the 14th ACM SIGKDD international conference on knowledge discovery and data mining. New York, USA: ACM, 2008:426-434.
- [9] Wang H W, Zhang F Z, Wang J L, et al. RippleNet: propagating user preferences on the knowledge graph for recommender systems [C] //Proceedings of the 27th ACM international conference on information and knowledge management. New York, USA: ACM, 2018:417-426.
- [10] Zhang F Z, Yuan N J, Lian D F, et al. Collaborative knowledge base embedding for recommender systems [C] //Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. New York, USA: ACM, 2016:353-362.
- [11] Elahi E, Anwar S, Al-kfairy M, et al. Graph attention-based neural collaborative filtering for item-specific recommendation system using knowledge graph [J]. Expert Systems with Applications, 2025,266:126133.
- [12] Hu J, Hooi B, He B S, et al. Modality-independent graph neural networks with global transformers for multimodal recommendation [C] //Proceedings of the thirty-ninth AAAI conference on artificial intelligence. New York, USA: ACM, 2025:11790-11798.
- [13] Cui X H, Qu X L, Li D M, et al. MKGCN: multi-modal knowledge graph convolutional network for music recommender systems[J]. Electronics, 2023,12(12):2688.

(编辑:李志豪)