

大语言模型增强的知识图谱 推荐算法在电子商务中的 应用与推广

王应文, 张 伟, 李志文

贵州大学大数据与信息工程学院, 贵州 贵阳

收稿日期: 2025年10月14日; 录用日期: 2025年10月29日; 发布日期: 2025年11月27日

摘 要

在电子商务快速发展的背景下, 推荐系统已成为提升用户体验与促进平台增长的关键技术。传统方法面临交互稀疏、冷启动与长尾问题, 限制了推荐性能与商业价值。知识图谱(KG)增强推荐通过建模结构化语义关系, 缓解了部分挑战, 但现有图神经网络方法仍受限于知识缺失、语义不足与计算冗余。本文提出一种融合大语言模型(LLM)与知识图谱的推荐框架, 利用LLM对物品的本地知识子图进行自然语言理解与语义建模, 生成高质量表示, 避免多跳传播带来的效率与泛化瓶颈。该算法在电商场景中可有效捕捉商品属性与用户兴趣之间的深层关系, 提升冷启动与长尾推荐能力, 助力平台实现智能推荐优化与数字经济价值增长。

关键词

推荐算法, 电子商务, 知识图谱, 大语言模型, 语义建模

Applications and Deployment of Large Language Model-Enhanced Knowledge Graph Recommender Algorithm in E-Commerce

Yingwen Wang, Wei Zhang, Zhiwen Li

College of Big Data and Information Engineering, Guizhou University, Guiyang Guizhou

Received: October 14, 2025; accepted: October 29, 2025; published: November 27, 2025

文章引用: 王应文, 张伟, 李志文. 大语言模型增强的知识图谱推荐算法在电子商务中的应用与推广[J]. 电子商务评论, 2025, 14(11): 2640-2651. DOI: 10.12677/ec.2025.14113730

Abstract

Against the backdrop of rapid e-commerce development, recommender systems have become a key technology for enhancing user experience and driving platform growth. Traditional methods suffer from sparse user-item interactions, cold-start issues, and insufficient coverage of long-tail items, thereby constraining recommendation performance and commercial value. Knowledge graph (KG)-enhanced recommendation alleviates some of these challenges by modeling structured semantic relations; however, existing graph neural network approaches remain limited by knowledge incompleteness, coarse-grained semantics, and computational redundancy. This paper proposes a novel recommendation framework that integrates large language models (LLMs) with knowledge graphs. By applying LLMs to the natural-language understanding and semantic modeling of local knowledge subgraphs of items, the framework produces high-quality representations and avoids the efficiency and generalization bottlenecks caused by multi-hop propagation. In e-commerce scenarios, the proposed algorithm effectively captures deep relationships between product attributes and user interests, enhances recommendations for cold-start and long-tail items, and thereby supports intelligent recommendation optimization and the growth of digital economic value for online platforms.

Keywords

Recommender Algorithm, E-Commerce, Knowledge Graph, Large Language Model, Semantic Modeling

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

推荐系统在电子商务领域的迅猛发展已成为推动用户参与度提升与平台收益增长的核心动力[1]。借助对用户浏览行为、购买记录、加购操作及页面停留时间等多维数据的分析,推荐系统能够有效建模用户兴趣偏好并提供个性化商品排序与推荐服务。伴随用户交互行为的持续积累,推荐模型在预测消费需求、提升转化率及挖掘潜在消费潜力方面展现出显著价值。与此同时,商品间的共现关系、品牌偏好路径、功能属性聚合等信息逐渐成为建模用户偏好迁移与商品语义依赖的重要基础,推动了知识图谱(Knowledge Graph, KG)在推荐任务中的广泛应用。

知识图谱通过建模实体(如商品、品牌、品类等)之间的多种语义关系,构建结构化的图数据,为推荐系统提供了丰富的高阶语义关联信息。研究表明,将知识图谱与协同过滤[2] [3]、图神经网络(Graph Neural Network, GNN)等模型结合,可有效提升冷启动和长尾推荐场景下的推荐准确性与系统稳健性。然而,现有的 KG 增强推荐方法仍存在三方面不足:其一,真实世界中的 KG 往往存在实体与关系的缺失,语义结构不完整,影响了图结构学习的质量[4];其二,大多数方法依赖多跳图神经传播机制建模高阶关系,计算成本高昂,且深度传播常出现过度平滑(over-smoothing)问题,导致节点表示区分性下降[5];其三,传统 KG 方法多基于离散 ID 编码与结构路径建模,缺乏对自然语言中隐含语义的深度理解与泛化能力,限制了推荐系统的表达空间[6]。

随着大语言模型(Large Language Models, LLMs)在自然语言理解、知识推理与通用语义建模方面的突破,其在推荐系统领域的潜力也逐步显现[7]。大语言模型具备跨领域文本建模与结构数据语义融合的能力

力,能够将知识图谱中的子图结构、实体描述、关系路径等信息转化为自然语言提示(prompt)输入,实现“图到文”的转换,进一步提升图结构的语义表达能力。在推荐场景中,通过将用户-物品交互[8]、商品 KG 子图等转换为语言上下文输入,LLM 可生成语义丰富、上下文感知的高质量表示[9][10],替代传统图神经网络中的跳数传播与聚合操作,从而缓解 KG 不完备、语义稀疏与计算冗余等问题。

为了解决传统知识图谱推荐中高阶关系建模能力不足、结构信息利用不充分以及语义表达受限等问题,本文提出了一种融合大语言模型与知识图谱的电子商务推荐算法。该方法包括两项关键设计:首先,基于提示构造的图语言建模机制,将知识图谱中与用户或物品相关的子图结构转化为自然语言形式,并输入大语言模型进行统一语义编码,从而实现图结构的语义增强;其次,引入多粒度的语义聚合策略,融合大语言模型生成的实体表示与结构邻居信息,提升对高阶语义路径与潜在关联的建模能力。通过将语言模型的表达能力与知识图谱的结构优势相结合,本文算法不仅有效缓解了图谱稀疏与冷启动问题,也显著提升了推荐系统的泛化性能与语义可解释性。

2. KGLM 模型

本文提出的 KGLM 模型的总体框架如图 1 所示,模型主要由三层组成,三者协同工作以实现语义增强的电子商务推荐。第一层为物品语义建模层,通过提取物品在知识图谱中的局部子图结构,并将其转化为自然语言描述,输入大语言模型以生成语义丰富的物品表示,从而捕捉商品属性间的深层语义联系。第二层为用户语义建模层,结合用户的历史交互记录与关联实体信息,通过大语言模型生成上下文感知的用户语义表示,以刻画其潜在偏好特征。第三层为推荐预测层,综合用户与物品的语义嵌入,通过匹配函数计算相关性得分,实现个性化推荐结果输出。该模型充分融合了知识图谱的结构优势与大语言模型的语义理解能力,在冷启动与长尾商品推荐场景中表现出更强的泛化性与准确性。

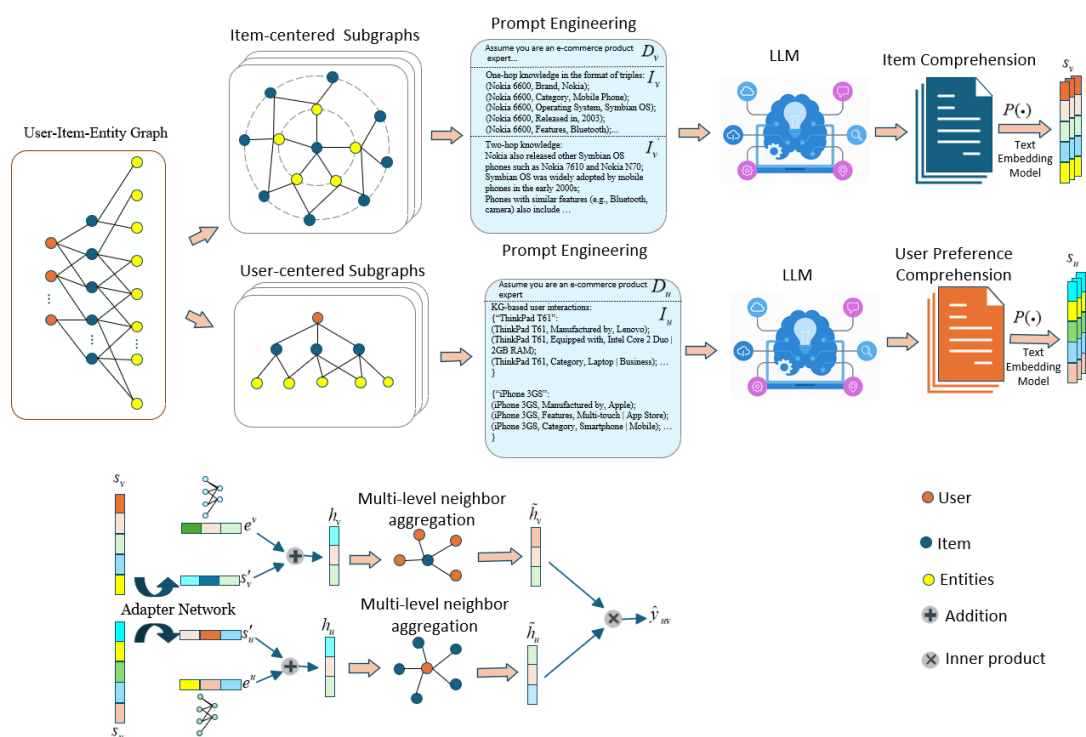


Figure 1. Overall framework of KGLM model
图 1. KGLM 模型总体框架

2.1. 物品语义建模

在电子商务推荐任务中,知识图谱(Knowledge Graph, KG)通过建模商品间的结构化关系,能够有效揭示潜在的语义关联。然而,传统知识图谱多依赖人工构建,存在事实缺失与知识覆盖不足等问题,且其高度结构化的形式难以直接与自然语言描述对齐,限制了语义表达的丰富性。为此,本文引入大语言模型(Large Language Model, LLM)以增强知识图谱的理解与语义建模能力,从而构建更高质量的商品表征。我们选用的大语言模型是 DeepSeek-V2,选择 DeepSeek-V2 的主要原因在于它兼具强通用知识与稳定指令跟随能力,非常契合“以结构化三元组为输入、产出规范化语义文本”的任务设定。电商知识图谱常见事实缺失、同义异名与长程依赖,DeepSeek-V2 在关系推理、缺失补全与一致化重述方面表现稳健,能把一阶和二阶三元组整合成高保真、可编码的说明文本,显著缓解仅靠图传播带来的过度平滑与噪声累积。相较将大模型直接端到端微调,DeepSeek-V2 作为只读“语义理解器”,只负责离线把知识图谱转成高质量文本和语义向量,不参与下游的反向传播与训练,后续的训练只更新自己的参数,有助于控制训练难度。

在该模块中,我们重点介绍如何运用 LLM 对以物品为中心的物品图进行有效的理解及完善,首先针对每个物品

$$T_v = \{(v, r, e) | (v, r, e) \in G_K\} \quad (1)$$

其中, G_K 表示全局知识图谱, r 表示实体间的关系类型, e 为与物品 v 相连接的属性实体。在实际电商场景中,一级邻居节点通常代表商品属性(如“品牌”、“类别”、“产地”等),因此这些节点能够反映物品在语义空间中的局部特征。当部分属性或关系缺失时,我们采用统一占位符“missing”进行补全,以保持结构完整性。考虑到一阶邻居无法充分捕捉高阶语义信息,本文进一步引入二阶邻接结构,以增强子图的上下文表达能力。设

$$E_v = \{e | (v, r, e) \in T_v\} \quad (2)$$

为一阶邻居集合,则对于每个 $v \in E_v$,形成二阶三元组以构造扩展集合 T_e ,定义如下:

$$T_e = \{(e, r, v') | (e, r, v') \in G_K, v' \neq v\} \quad (3)$$

通过该策略,在不显著增加计算复杂度的前提下,模型能够从有限的局部子图中学习到更丰富的上下文依赖。

我们首先把以目标物品 v 为中心的一阶与二阶关系统一表示为三元组。对一阶邻域,将集合 T_v 中的所有 (v, r, e) 串接成一段结构化文本,记为 I_v ;对二阶邻域,则从每个一阶实体的三元组集合中采样,并通过预设模板转写为连贯句子,得到 I'_v ,以便大语言模型理解。随后,我们设计系统指令 D_v 作为生成约束,具体的提示模板设计如图 2 所示。

如图所示将 D_v 、 I_v 与 I'_v 拼接成完整提示,驱动 LLM 对以物品 v 为中心的 KG 子图进行理解,综合考虑结构化与语言化信息,我们将输入构造如下格式:

$$C_v = \text{LLM}(D_v, I_v, I'_v) \quad (4)$$

其中, I_v 与 I'_v 分别表示由一阶与二阶三元组生成的文本描述, C_v 为 LLM 对以物品 v 为中心的知识图谱的语义理解结果。该生成结果包含了知识补全、关系推理与语义重构等多维信息,体现了 LLM 对结构化图谱的自然语言认知能力。LLM 产出的物品语义描述需要被稳定地量化为向量,以便与协同信号共同训练。为此我们选用句向量表现成熟的 sup-SimCSE-RoBERTa-large 作为文本嵌入模型 $P(\cdot)$,该文本嵌入模型其监督式对比学习在句子层面具有更强语义聚散性,能把 LLM 生成的物品/用户描述映射为稳定、可

度量的向量；与协同空间经轻量投影即可对齐融合，计算与时延远低于直接用 LLM 向量，便于离线缓存与冻结，降低过拟合与噪声放大风险。 $P(\cdot)$ 将 C_v 转换为嵌入向量表示，获得物品的语义增强嵌入：

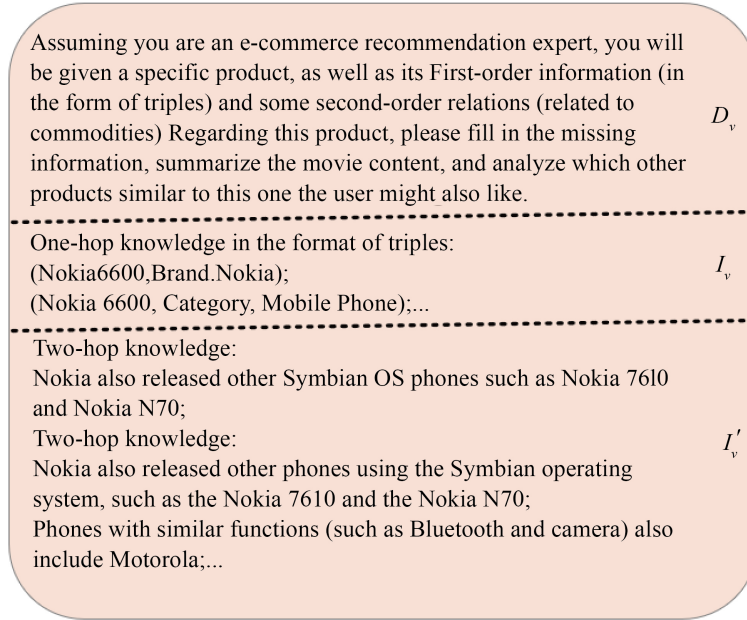


Figure 2. Prompt template for item knowledge graph

图 2. 物品知识图谱的提示模板

$$s_v = P(C_v) \quad (5)$$

所得嵌入 s_v 不仅包含了图结构关系信息，也整合了 LLM 对语言层面的深层语义建模能力，从而在推荐系统中实现了对物品语义特征的精细化表征。该模块有效缓解了传统图模型在知识稀疏、属性缺失及多跳传播冗余方面的不足，为后续用户建模与推荐预测提供了高质量的语义表示基础。

2.2. 用户语义建模

本节基于知识图谱的用户语义建模方法，通过构建大语言模型可理解的用户侧“知识档案”，实现从历史交互行为到语义偏好的高质量建模。在该过程中，我们首先利用知识图谱扩展传统的用户-物品二部图结构，将用户在训练阶段的所有交互记录统一映射为结构化文本表示，再通过文本嵌入函数获取其语义向量表示，用于后续推荐模型的训练与预测。该方法的核心在于借助知识图谱提供的丰富语义上下文，补全交互信号中缺失的高阶关联，从而在更完整的知识空间中刻画用户兴趣。具体而言，首先从用户-物品二部图中，抽取给定用户 u 的交互子图 B_u ，即用户在训练集内曾交互过的所有物品集合。针对其中每一个物品 $v \in B_u$ ，提取以该物品为中心的一阶知识图谱子图，并表示为三元组集合 T_v ，形式为 (v, r, e) 。这些三元组代表该物品与其属性之间的语义关系，我们将 T_v 中所有三元组按序拼接为结构化文本 I_v ，形式化地表示为：

$$I_u = \bigoplus_{v \in I_u} name_v : I_v \quad (6)$$

其中 $\bigoplus$ 表示串接操作， $name_v$ 为物品 v 的名称。

在此基础上，结合一段预设的系统提示指令 D_u ，具体的提示模板设计如图 3 所示：

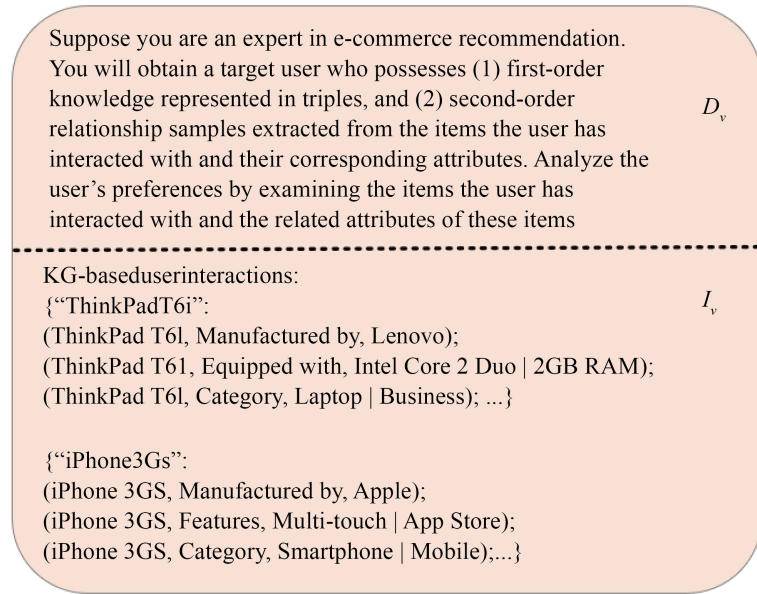


Figure 3. Prompt template for user knowledge graph

图 3. 用户知识图谱的提示模板

将上图提示输入至大语言模型生成用户偏好文本表示 C_u ，即：

$$C_u = \text{LLM}(I_u, D_u) \quad (7)$$

为将上述文本化偏好进一步纳入神经推荐模型中，我们使用一个预训练文本编码模型 $P(\cdot)$ 对其进行 C_v 转换为嵌入向量表示，获得用户的语义增强嵌入，得到最终的用户语义表示 s_u ：

$$s_u = P(C_u) \quad (8)$$

这一过程实现了用户交互历史与其相关知识的统一建模，即将符号交互序列嵌入到统一的语义向量空间中，为后续与物品语义表示的匹配或融合提供了良好的基础。

2.3. 预测模型

在完成物品与用户语义建模后，我们分别获得了语义嵌入 $s_v \in \mathbb{R}^{d_s}$ 和 $s_u \in \mathbb{R}^{d_s}$ ，它们由大语言模型基于知识子图语义理解生成。与此同时，推荐系统中传统的 ID 嵌入 $e_v \in \mathbb{R}^d$ 和 $e_u \in \mathbb{R}^d$ 则直接来源于模型参数。在语义空间与协同空间维度不一致、模态差异显著背景下，亟需引入跨模态对齐机制以实现信息融合。

为此，本文设计了轻量的投影适配器网络(Adapter Network)，用于将 LLM 生成的语义嵌入映射至 ID 嵌入所在的协同空间。该映射过程通过线性变换与非线性激活函数实现：

$$s'_v = \sigma(W_1 s_v), \quad s'_u = \sigma(W_2 s_u) \quad (9)$$

其中， $W_1, W_2 \in \mathbb{R}^{d \times d_s}$ 为可学习的投影矩阵， $\sigma(\cdot)$ 表示激活函数(如 ELU)。在整个训练过程中， s_v 与 s_u 被视为固定特征，只有投影参数 W_1, W_2 与后续推荐模型参数参与训练，以保证 LLM 提供的语义信息不被过拟合所干扰。

投影后，我们采用均值池化策略将投影后的语义表示与原始 ID 嵌入进行融合，构造用户与物品的统一表示：

$$h_v = \frac{1}{2}(e_v + s'_v), \quad h_u = \frac{1}{2}(e_u + s'_u) \quad (10)$$

该融合表征 $h_u h_v \in R^d$ 兼具协同表示的行为信息与 LLM 理解下的语义知识, 有助于增强模型对冷启动与长尾物品的识别能力。

在此基础上, 我们将融合后的表示作为输入, 构建图神经推荐网络以建模用户 - 物品交互关系。本文采用 LightGCN 作为推荐主干模型, 其通过图结构的多层邻居聚合, 捕捉用户 - 物品之间的高阶协同信号。具体地, 第 l 层的嵌入更新规则为:

$$h_u^{(l+1)} = \sum_{v \in M_u} \frac{1}{\sqrt{|M_u| |M_v|}} h_v^{(l)}, \quad h_v^{(l+1)} = \sum_{u \in M_v} \frac{1}{\sqrt{|M_v| |M_u|}} h_u^{(l)} \quad (11)$$

其中, M_u 表示与用户 u 发生交互的物品集合, M_v 为与物品 v 交互过的用户集合; $h_u^{(0)} = h_u, h_v^{(0)} = h_v$ 为前述融合后的初始嵌入。经过 L 层传播后, 我们对每一层的嵌入进行平均以获得最终的用户和物品表示:

$$\tilde{h}_u = \frac{1}{L+1} \sum_{l=0}^L h_u^{(l)}, \quad \tilde{h}_v = \frac{1}{L+1} \sum_{l=0}^L h_v^{(l)} \quad (12)$$

最后, 用户 u 与物品 v 的交互概率预测由内积操作给出:

$$\hat{y}_{uv} = \tilde{h}_u^\top \tilde{h}_v \quad (13)$$

该分数 $\hat{y}_{uv}$ 表示系统预测用户 u 与物品 v 存在交互的可能性。模型在训练阶段采用 Bayesian Personalized Ranking (BPR) 损失函数进行优化, 以强化正样本对负样本的排序优势, 从而提升整体推荐效果。

该推荐预测模块通过跨模态对齐、融合建模与图神经传播三个环节, 成功实现了语义知识与协同信息的深度融合, 在提升推荐性能的同时, 也为冷启动、知识稀疏等问题提供了有效解决路径。

2.4. 训练模型

本模型训练过程可分为两个阶段, 分别对应语义信息建模阶段与推荐模型学习阶段。在第一阶段, 我们利用大语言模型对用户与物品的知识子图进行语义理解, 生成高质量的语义嵌入向量 $s_u \in R^{d_s}$ 和 $s_v \in R^{d_s}$ 。该阶段为无监督过程, 所有语义表征在模型训练期间保持冻结, 避免其受到下游任务梯度干扰, 保持语言知识的一致性与泛化能力。

第二阶段为主推荐模型的训练阶段, 我们将上述语义嵌入通过投影与融合操作, 构建得到融合后的表示 $\tilde{h}_u$ 和 $\tilde{h}_v$, 并作为输入构建基于图神经网络的推荐预测模块。在该阶段, 我们采用广泛使用的 Bayesian Personalized Ranking (BPR) 损失函数作为优化目标, 以提升模型对用户 - 物品交互偏好的排序能力:

$$L = \sum_{(u, v^+, v^-) \in Y} -\ln \sigma(\hat{y}_{uv^+} - \hat{y}_{uv^-}) + \lambda \|\Theta\|_2^2 \quad (14)$$

其中, $Y = \{(u, v^+, v^-)\}$ 为训练三元组集合, (u, v^+) 表示用户 u 的正反馈交互样本, (u, v^-) 为从未观察交互中随机采样的负样本; $\sigma(\cdot)$ 为 sigmoid 函数, 用于衡量两个物品的得分差异; $\hat{y}_{uv}$ 为用户 u 与物品 v 的预测得分; λ 是正则化项系数, 用以缓解模型过拟合风险, Θ 表示所有可训练参数集合。

通过最小化上述目标函数, 模型能够有效学习到用户偏好排序, 并充分融合 LLM 所提供的语义信息与图结构中的协同信号, 从而显著提升推荐系统在冷启动、长尾分布与语义稀疏情形下的表现。

3. 实验

3.1. 实验环境与评价指标

本研究在本地计算环境中完成实验评估, 硬件平台配置为: Intel(R) Core(TM) i5-14600KF 3.50 GHz 处理器与 NVIDIA GeForce RTX 5070 显卡, 操作系统为 Windows 11。所有实验均在该平台上独立完成,

确保实验结果的可复现性。我们选用的实验数据集为 ml-1m，该数据集由美国明尼苏达大学的 GroupLens Research Group 提供，是推荐系统领域中最广泛使用的标准基准数据集之一。该数据集专为推荐任务设计，包含约 100 万条用户评分记录，涉及 6000 多个物品与 6000 多位用户，具备结构清晰、数据密度适中、属性信息丰富等优势，已被广泛应用于协同过滤、图神经网络、知识图谱增强推荐等研究任务。为公平对比，我们严格复现了所有基线方法，均基于其官方开源实现，并统一将推荐模型中的嵌入维度设定为 64。针对训练过程，我们采用 Adam 优化器，学习率设为 0.001，最大训练轮数为 2000。推荐模型的隐藏层数 L 设为 3，并分别对嵌入层与隐藏层施加 Dropout 正则化，Dropout 比例从{0.2, 0.4, 0.6, 0.8} 中选取。在语义建模模块中，我们选用的大语言模型为 DeepSeek-V2，其在多个标准基准任务与开放式生成任务中表现出色。具体而言，我们通过调用 DeepSeek-V2 的开放 API 接口实现知识注释文本与偏好描述的生成。为控制生成稳定性与可重复性，我们将温度参数 temperature 设为 0，采样上界参数 top-p 设为 0.001。在文本语义向量生成阶段，本文采用 sup-simcse-roberta-large 作为文本嵌入模型，并加载其官方预训练权重。对于所有涉及 LLM 或文本嵌入的对比方法，我们均保持相同的 API 调用方式与嵌入模型配置，以确保实验的可比性与公平性。模型具体的超参数设置如表 1 所示：

Table 1. Model hyperparameter setting
表 1. 模型超参数设置

超参数	值
嵌入维度	64
图传播层数	3
批量大小	4096
学习率	0.001
训练轮数	2000

为全面衡量推荐模型在不同场景下的性能表现，本文采用两项被广泛认可的主流评价指标：召回率 (Recall)与归一化折损累计增益(Normalized Discounted Cumulative Gain, NDCG)。这两种指标兼顾了结果的相关性与排序质量，能够有效评估模型在实际推荐中的有效性与用户满意度。具体而言，我们报告了 Recall@k 与 NDCG@k 的结果，其中 $k \in 10, 20$ ，分别衡量模型在前 k 个推荐结果中成功命中用户真实点击物品的比例，以及这些命中结果在推荐列表中的排序合理性。Recall 指标关注是否成功召回用户感兴趣的物品，而 NDCG 进一步对召回结果的位置敏感性建模，给予较高排名的相关项更大权重，因此能够更真实地反映用户对推荐排序的接受程度。在评估协议上，为保证结果的公平性与一致性，我们采用了全排序协议(All-Ranking Protocol)，即对每位用户，将其未交互的所有物品视为候选集，并在此全集上进行打分与排序。该协议比常用的随机负采样方法更具挑战性，也更贴近实际推荐系统部署中的开放式推荐场景，能够更有效衡量模型对全局物品空间的排序能力。总体而言，Recall@k 与 NDCG@k 的结合使得本实验的评估框架既关注命中能力，又兼顾排序质量，为比较不同模型的推荐效果提供了科学、客观的量化依据。

3.2. 实验环境与评价指标

为了全面评估本文所提出的 KGLM 在推荐任务中的性能表现，我们选取了五种具有代表性的基线模型作为对比对象，涵盖了当前知识图谱增强推荐系统的主流技术路线。这些对比方法分别为：BPR-MF [11]、RippleNet [12]、KGAT [13]、CKE [14]与 RLMLRec [15]，它们从不同角度对用户偏好建模与知识引入

方式进行了探索,具有良好的代表性与对比价值。其中, BPR-MF 是一种经典的协同过滤模型,采用贝叶斯个性化排序方法,通过学习用户与物品之间的隐式偏好差异进行推荐,其优势在于优化目标的直接性与实现的简洁性。RippleNet 则是一种基于路径扩散的知识图谱推荐方法,通过在图中模拟用户兴趣的逐层传播,有效捕捉了用户与多跳实体之间的语义关联。KGAT 引入注意力机制对知识感知协同图中的节点信息进行加权聚合,从而增强了模型在多跳结构建模上的灵活性与表达能力。CKE 属于知识图谱嵌入类方法,专注于将知识图谱中的结构信息嵌入至统一的向量空间,与用户行为进行融合建模。RLMRec 是一种融合大语言模型(LLM)能力的开放世界推荐算法,通过自然语言提示(Prompting)获取实体的文本语义表示,从而缓解知识图谱结构不完整或语义不足的问题,代表了大模型引入推荐系统的前沿探索方向。

在实验评估方面,为确保比较的公正性与系统性,我们采用了广泛应用的 NDCG@k 和 Recall@k 两类指标,并在全排序(All-Ranking)协议下对各模型进行一致性测试。实验结果显示, KGLM 在各项指标上均显著优于所有基线模型。在本实验中, KGLM 模型在所有四项指标上均取得最优结果,尤其在与 KGAT 和 RLMRec 两个具代表性的强基线模型进行对比时,展现出明显优势。其结果如图 4 和图 5 所示。具体而言,在 Recall@10 指标上, KGLM 达到 0.1683,较 KGAT (0.1536)与 RLMRec (0.1610)分别提升了 9.5%与 4.5%;在 NDCG@10 上, KGLM 的值为 0.4092,相较于 KGAT (0.3782)与 RLMRec (0.3903)分别提升 8.2%与 4.8%。在更长的推荐列表中, KGLM 同样表现稳定,其 Recall@20 为 0.2609,高于 KGAT (0.2412)与 RLMRec (0.2515),分别提升了 8.2%与 3.7%;在 NDCG@20 上则为 0.3881,分别高出 KGAT 的 6.0%与 RLMRec 的 2.5%。值得关注的是,尽管 RLMRec 同样引入了大语言模型的文本语义增强机制,但由于缺乏结构化知识的系统融合,其整体推荐精度依然低于 KGLM,表明后者在融合结构知识与语义表示之间达到了更优平衡。

综上所述, KGLM 充分发挥了结构化知识与语言语义的协同作用,通过构建统一的语义增强推荐框架,显著提升了模型对用户偏好、高阶关系与长尾物品的感知能力。相比于传统方法, KGLM 不仅在准确性指标上取得了显著优势,更在模型的泛化能力、语义表达精度与实际适用性方面展现出强大潜力,为下一代知识图谱推荐系统提供了具有推广价值的建模范式。

为了进一步验证 KGLM 模型中各关键模块对推荐性能的影响,本文设计了一系列消融实验。通过移除模型的部分核心组件,分析其在语义建模与推荐推理中的作用。具体地,我们构建了三种消融模型与完整的 KGLM 进行对比,实验结果见表 2。

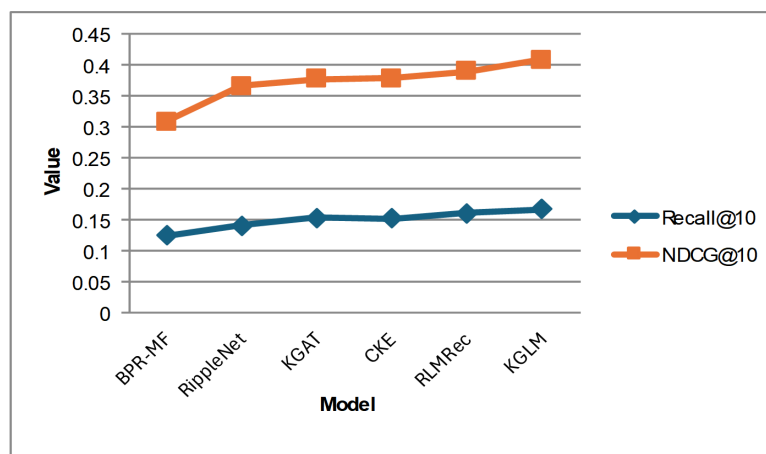


Figure 4. Recall@10、NDCG@10 values of different algorithms on the dataset

图 4. 不同算法在数据集上的 Recall@10、NDCG@10 值

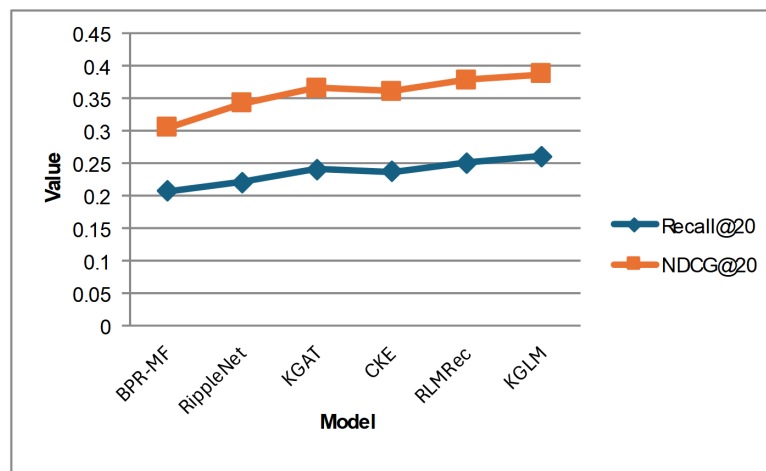


Figure 5. Recall@20、NDCG@20 values of different algorithms on the dataset

图 5. 不同算法在数据集上的 Recall@20、NDCG@20 值

KGLM-v: 在该变体中, 去除了物品侧的语义嵌入, 仅保留了基于用户-物品交互的结构信息, 以验证由大语言模型(LLM)生成的物品语义表征在推荐性能中的贡献。

KGLM-u: 此变体去除了用户侧的语义嵌入, 仅依赖物品表示进行推荐预测。该实验用于评估 LLM 对用户知识子图的语义理解在用户偏好建模中的重要性。

KGLM-d: 在该模型中, 移除了 LLM 提示模板中引入的二阶三元组信息, 使模型仅基于一阶知识结构进行语义建模, 从而评估高阶关系对知识推理与推荐结果的影响。

通过对比可知, 移除任一模块都会导致模型性能下降, 其中物品语义嵌入的影响最为显著, 说明 LLM 从知识图谱中提取的语义知识在提升推荐精度中起到关键作用。同时, 用户语义建模与二阶关系信息亦在捕捉复杂偏好模式方面发挥了重要作用, 进一步证明了 KGLM 在语义增强与结构理解上的有效性。

Table 2. Comparison chart of ablation experiment results

表 2. 消融实验结果对比图

模型	Recall@10	NDCG@10	Recall@20	NDCG@20
KGLM	0.1683	0.4092	0.2609	0.3881
KGLM-v	0.1601	0.3898	0.2515	0.3691
KGLM-u	0.1636	0.3956	0.2584	0.3802
KGLM-d	0.1612	0.3815	0.2572	0.3792

消融实验结果表明, KGLM 的性能提升并非来自单一组件, 而是源于多模块的协同作用。移除物品侧语义嵌入后, 模型在各数据集与不同阶段评估下均出现稳定退化, 说明仅依赖用户-物品交互难以覆盖属性与关系等细粒度信息, 由 LLM 提炼的物品语义是缓解交互稀疏、增强长尾与冷启动推荐的重要来源。进一步去除用户侧语义建模时, 个性化能力明显削弱, 模型更易偏向热门项而难以刻画细微兴趣差异, 体现出 LLM 对用户知识子图的理解能够有效补足偏好表达。若在提示中删除二阶三元组, 模型对跨实体的语义路径与潜在因果联系的把握明显变弱, 仅凭一阶事实会使视野局限于局部邻域, 无法充分利用知识图谱的全局结构, 从而影响排序相关性与稳健性。

总体而言, KGLM 的完整配置之所以表现最佳, 在于同时引入了物品与用户的语义增强, 并通过高

阶关系提示让 LLM 更全面地理解与推理知识图谱结构。该组合既提升了全局语义整合能力，又强化了个性化刻画与泛化鲁棒性，特别对长尾与冷启动场景更为有效，并在保持可解释性的同时稳定提升推荐质量。

4. 总结

本文面向电子商务推荐场景，提出了融合知识图谱(KG)与大语言模型(LLM)的 KGLM 框架。模型以“物品语义建模 - 用户语义建模 - 推荐预测”三层协同为主线：首先由 LLM 对以实体为中心的局部 KG 进行语言化理解，生成高质量的物品与用户语义表示；随后通过轻量投影完成语义与协同空间对齐，并结合基于交互图的预测网络完成打分与排序。离线两阶段训练与全排序评测表明，KGLM 在 Recall@k、NDCG@k 等指标上整体优于多条主流技术路线。对比分析显示，KGLM 相较于以结构传播为核心的 KGAT 和以文本语义为核心的 RLMRec 均取得稳定领先，说明“结构 + 语义”的深度融合比单一路径更能刻画用户偏好与物品关系。消融实验进一步验证：物品语义、用户语义与二阶关系提示缺一不可，高阶语义路径显著提升排序相关性与模型稳健性。

从业务价值看，KGLM 有效缓解交互稀疏、冷启动与长尾覆盖不足问题，提升了推荐的准确性、可解释性与泛化能力，为电商平台的转化率优化与供需匹配提供了可复制的技术方案。受限之处在于：对提示设计与 LLM 成本敏感，KG 噪声与领域迁移仍需系统性处理，离线生成的语义向量与实时兴趣漂移存在时效差。未来可沿三方面拓展：其一，引入噪声感知的端到端/在线学习，学习三元组置信度与置信门控；其二，多模态与外部知识的统一表征(图像、标题、评论、属性)与成本优化；其三，面向开放世界的持续知识增强与安全合规(隐私、偏见与可解释)。总体而言，KGLM 为“LLM × KG”的电商推荐提供了可落地的范式与坚实的实验依据。

参考文献

- [1] Bandyopadhyay, S., Thakur, S.S. and Mandal, J.K. (2020) Product Recommendation for E-Commerce Business by Applying Principal Component Analysis (PCA) and K-Means Clustering: Benefit for the Society. *Innovations in Systems and Software Engineering*, 17, 45-52. <https://doi.org/10.1007/s11334-020-00372-5>
- [2] He, X. and Chua, T.S. (2017) Neural Factorization Machines for Sparse Predictive Analytics. *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Tokyo, 7-11 August 2017, 355-364. <https://doi.org/10.1145/3077136.3080777>
- [3] He, X., Deng, K., Wang, X., Li, Y., Zhang, Y. and Wang, M. (2020) LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation. *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, Xi'an, 25-30 July 2020, 639-648. <https://doi.org/10.1145/3397271.3401063>
- [4] Wang, M., Zeng, D., Xu, Z., Guo, R. and Zhao, X. (2023) Federated Knowledge Graph Completion via Latent Embedding Sharing and Tensor Factorization. *2023 IEEE International Conference on Data Mining (ICDM)*, Shanghai, 1-4 December 2023, 1361-1366. <https://doi.org/10.1109/icdm58522.2023.00175>
- [5] Wang, S., Sui, Y., Wang, C. and Xiong, H. (2024) Unleashing the Power of Knowledge Graph for Recommendation via Invariant Learning. *Proceedings of the ACM Web Conference 2024*, Singapore, 13-17 May 2024, 3745-3755. <https://doi.org/10.1145/3589334.3645576>
- [6] Yang, Y., Huang, C., Xia, L. and Huang, C. (2023) Knowledge Graph Self-Supervised Rationalization for Recommendation. *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, Long Beach, 6-10 August 2023, 3046-3056. <https://doi.org/10.1145/3580305.3599400>
- [7] Fan, W., Zhao, Z., Li, J., Liu, Y. and Mei, X. (2023) Recommender Systems in the Era of Large Language Models (LLMs). *IEEE Transactions on Knowledge and Data Engineering*, 36, 6889-6907. <https://doi.org/10.1109/TKDE.2024.3392335>
<https://ieeexplore.ieee.org/abstract/document/10506571>
- [8] Wang, H., Zhao, M., Xie, X., Li, W. and Guo, M. (2019) Knowledge Graph Convolutional Networks for Recommender Systems. *The World Wide Web Conference*, San Francisco, 13-17 May 2019, 3307-3313. <https://doi.org/10.1145/3308558.3313417>
- [9] Wu, L., Zheng, Z., Qiu, Z., Wang, H., Gu, H., Shen, T., et al. (2023) A Survey on Large Language Models for

- Recommendation. <https://doi.org/10.1007/s11280-024-01291-2>
- [10] Zhao, W.X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., *et al.* (2023) A Survey of Large Language Models. <https://arxiv.org/abs/2303.18223>
- [11] Rendle, S., Freudenthaler, C., Gantner, Z. and Schmidt-Thieme, L. (2012) BPR: Bayesian Personalized Ranking from Implicit Feedback. <https://arxiv.org/abs/1205.2618>
- [12] Wang, H., Zhang, F., Wang, J., Zhao, M., Li, W., Xie, X., *et al.* (2018) RippleNet: Propagating User Preferences on the Knowledge Graph for Recommender Systems. *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, Turin, 22-26 October 2018, 417-426. <https://doi.org/10.1145/3269206.3271739>
- [13] Wang, X., He, X., Cao, Y., Liu, M. and Chua, T. (2019) KGAT: Knowledge Graph Attention Network for Recommendation. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Anchorage, 4-8 August 2019, 950-958. <https://doi.org/10.1145/3292500.3330989>
- [14] Zhang, F., Yuan, N.J., Lian, D., Xie, X. and Ma, W. (2016) Collaborative Knowledge Base Embedding for Recommender Systems. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, 13-17 August 2016, 353-362. <https://doi.org/10.1145/2939672.2939673>
- [15] Ren, X., Wei, W., Xia, L., Su, L., Cheng, S., Wang, J., *et al.* (2024) Representation Learning with Large Language Models for Recommendation. *Proceedings of the ACM Web Conference 2024*, Singapore, 13-17 May 2024, 3307-3313. <https://doi.org/10.1145/3589334.3645458>