



计算机应用研究
Application Research of Computers
ISSN 1001-3695, CN 51-1196/TP

《计算机应用研究》网络首发论文

题目: 基于图神经网络与深度学习的商品推荐算法
作者: 冯兴杰, 生晓宇
DOI: 10.19734/j.issn.1001-3695.2021.05.0183
收稿日期: 2021-05-28
网络首发日期: 2021-07-26
引用格式: 冯兴杰, 生晓宇. 基于图神经网络与深度学习的商品推荐算法[J/OL]. 计算机应用研究. <https://doi.org/10.19734/j.issn.1001-3695.2021.05.0183>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

基于图神经网络与深度学习的商品推荐算法

冯兴杰, 生晓宇[†]

(中国民航大学 计算机科学与技术学院, 天津 300300)

摘要: 基于图神经网络的推荐算法可以提取传统方法无法提取的用户与商品之间的关联关系。目前此类算法大多忽略了用户和商品的评论数据中所存在的一般偏好。针对这一问题, 提出了一种方法, 在利用图神经网络提取关联关系的同时, 利用深度学习提取评论的优势提取用户和商品的一般偏好, 并进行特征融合来提升推荐效果。在4组公共数据集中进行了对比实验, 使用召回率和归一化折损累计增益作为评价指标, 并通过消融实验验证了方法的有效性。实验表明该方法比已有相关算法的效果更好。两种网络的特征融合对推荐效果有提升作用。

关键词: 推荐系统; 图神经网络; 深度学习; 注意力机制

中图分类号: TP391 doi: 10.19734/j.issn.1001-3695.2021.05.0183

Item recommendation algorithm based on GNN and deep learning

Feng Xingjie, Sheng Xiaoyu[†]

(College of Computer Science & Technology, Civil Aviation University of China, Tianjin 300300, China)

Abstract: The recommendation algorithm based on graph neural network can extract the association relationship between users and goods. Traditional methods cannot extract this relationship. At present, most of these algorithms ignore the general preferences in the review data of users and products. In order to solve this problem, this paper proposed a new method. This method uses the graph neural network to extract association relations, and the advantage of deep learning to extract the general preferences, and carries out feature fusion to improve the recommendation effect. This paper conducted comparative experiments and ablation experiments on four sets of public data sets to verify the effectiveness of the proposed method. The evaluation indexes include the recall rate and normalized discounted cumulative gain. Experiments show that this method is more effective than the existing algorithms. The feature fusion of the two networks can improve the recommendation effect.

Key words: recommender system; GNN; deep learning; attention mechanism

0 引言

互联网时代的到来使得人们所能接触到的数据量激增从而产生信息过载问题, 人们难以在琳琅满目的商品中挑选出符合自己需求的物品。推荐系统成为了解决这一问题的的重要手段。其根据用户历史行为分析用户不同偏好, 从而进行信息过滤, 最终给出个性化推荐内容^[1]。目前, 推荐系统在电子商务、社交、娱乐、教育等领域被广泛应用, 诞生了诸如天猫、微信、网易云音乐、今日头条等依托推荐系统的热门应用。

近年来, 深度学习方法在自然语言, 语音、图像视频等领域发展迅猛^[2]。利用深度学习方法从用户ID, 评论文本等数据中提取其中所隐藏的用户物品特征, 根据该特征预测用户对新物品的打分从而给出推荐是传统推荐算法主流做法。图神经网络提出以后, 研究人员发现可以利用图神经网络提取传统推荐算法中无法提取的用户与物品相互之间的连接特性, 使用图神经网络构建推荐算法逐渐成为当今推荐算法的研究热潮。

目前大部分基于图神经网络的推荐算法均使用用户物品的ID信息提取用户物品特征, 忽略了数据中评论文本等其他数据所隐含的用户物品特征, 或使用了辅助信息却使得网络训练更复杂。为此, 本文提出了一种基于图神经网络的融合了传统深度学习的推荐算法, 来融合不同用户物品特征。首先, 利用Light-GCN^[3]基础网络同时增加注意力机制来提取用户物品之间的高阶特性, 作为用户物品特征表示的一部分。其次, 使用传统深度学习网络从评论文本中提取用户物品的特征的另一部分。最终将两特征表示进行融合获得最终

用户物品特征表示, 进而给出推荐。实验结果表明, 该方法可以较好的融合用户物品的高阶特征与一般特性, 使得用户物品特征获得更好的表达。有效提升对用户的推荐效果。

本文的主要工作和贡献如下:

a) 本文提出了一种基于图神经网络的融合了传统深度学习的推荐算法, 通过传统深度学习方法与图神经网络共同提取用户的特征, 从而更好的表达用户的兴趣偏好;

b) 利用注意力机制, 为图神经网络中每一层添加注意力权重, 从而使用户特征表示中更重要的层获得更好的表达;

c) 在四个亚马逊公共数据集上进行了实验, 结果显示, 本文推荐算法的召回率与归一化折损累计增益高于同类算法。

1 相关工作

目前推荐算法主要分为三大类: 基于矩阵分解的推荐算法, 基于深度学习的推荐算法和基于图神经网络的推荐算法。

1.1 基于矩阵分解的推荐算法

传统推荐算法大部分利用矩阵分解 MF(Matrix factorization)将用户和物品映射到 n 维空间, 利用用户与物品的映射来计算用户对物品的偏好, 即进行评分预测任务。MF应用最成功的方法是 LFM 算法^[4], 其利用 MF 的同时引入偏置项, 极大提高了预测准确度。在 SVD++^[5]算法在 LFM 的基础上加入用户对其历史评分商品的隐式反馈向量的计算, 以修正预测结果。PMF^[6]算法将概率分布函数引入到 MF 中, 提供了一种矩阵分解的新思路。RMR^[7]和 TopicMF^[8]算法运用主题模型 LDA 提取文本特征, 同时利用评分数据进一步提升了预测精度。PGMF^[9]算法针对推荐系统中用户数据隐私

收稿日期: 2021-05-28; 修回日期: 2021-07-13 基金项目: 国家自然科学基金青年基金项目(61301245, 61201414)

作者简介: 冯兴杰(1969-), 男, 河北邢台人, 教授, 硕士, 博士, 主要研究方向为推荐系统、数据库及数据仓库、智能信息处理理论与技术; 生晓宇(1996-), 男(通信作者), 山东济南人, 硕士研究生, 主要研究方向为推荐系统(13001258943@163.com)。

保护带来的噪声问题, 将遗传算法引入矩阵分解中, 在为用户数据提供良好保护的同时, 降低了噪声问题, 保证了推荐结果的准确性。

随深度学习的逐渐发展, CNN、RNN 等深度学习方法可以提取矩阵分解技术无法提取的评论、图像等数据中的隐藏信息, 进而提升推荐效果。因此将深度学习方法应用到推荐算法中, 成为了推荐算法研究领域中的热门。

1.2 基于深度学习的推荐算法

基于深度学习的推荐算法, 最早体现在对评论文本的信息提取中, 例如 ConvMF^[10]利用 CNN 从评论中提取用户物品偏好, 相比矩阵分解方法, 算法性能有了很大提升。之后围绕 CNN 出现了很多算法: DeepCoNN^[11]采用并行 CNN 结构同时提取用户物品偏好; LCPMF^[12]将矩阵分解中的 LDA 主题模型与 CNN 结合起来, 从而综合利用评论文档的主题和深层语义信息; DIN^[13]引入注意力机制学习用户物品表达, 取得了不错效果; MAGN^[14]采用多头注意力机制, 在注意力机制的基础上增加了新的注意力, 来获取用户的不同朋友在不同方面对用户所施加的影响; Wide&Deep^[15]同时考虑了低阶特征与高阶特征, 提升了效果, 并利用多线程提高了处理效率; CapIPTV^[16]引入了胶囊网络, 利用胶囊网络的动态路由机制以及聚类特性, 同时利用注意力机制来得出不同用户的不同兴趣偏好。CNN 之后, 利用 RNN, LSTM, Transformer 等新的深度学习方法, 出现了 JNTM^[17]、RRN^[18]、BST^[19]等新的推荐算法。

图神经网络出现之后, 由于其可以学习到深度学习方法无法学习到用户与物品相互之间的关联关系, 基于图神经网络和深度学习方法共同成为目前推荐算法研究中的主流。

1.3 基于图神经网络的推荐算法

2009 年, 图神经网络 GNN^[20]模型被提出, 2017 年 GCN^[21]模型的出现使得图神经网络得到快速发展。GCN(图卷积神经网络)在推荐系统领域得到广泛应用: NGCF^[22]模型将 GCN 引入推荐算法, 建模用户项目的高阶连接性并据此给出推荐; KGCN-PN^[23]模型将图中实体的邻居实体进行排序采样, 利用 GCN 沿关系路径进行信息融合, 得到实体特征向量; KGAT^[24]算法将注意力机制融合进图神经网络中, 取得了良好的效果; BGANR^[25]算法利用注意力机制, 同时加入偏差以更好地捕获节点之间的高阶连接性; NIA-GCN^[26]在 GCN 的基础上进一步考虑的邻居节点之间的相互交互, 可以有效聚合各深度的邻居节点信息; GCN-ONCF^[27]模型将 GCN 设计为一个编码器, 利用外积操作将编码向量变为二维特征矩阵, 经过卷积自解码器实现卷积矩阵分解; Light-GCN 在 NGCF 基础上简化了使用二分图的 GCN 中冗余的部分, 从而提升了模型效率和性能。

尽管 GCN 可以很好的提取用户与项目之间的关联关系, 但大部分基于 GCN 的推荐算法没有考虑到深度学习在文本等其他信息处理中的优势, 因此, 本文首先结合注意力机制, 同时利用 GCN 的优势来提取用户与物品的关联关系, 再通过深度学习的优势来提取用户物品评论中蕴涵的兴趣偏好, 最终通过特征融合的方式来提升模型性能。

2 模型设计

2.1 问题阐述及相关定义

给定一个包含 x 个样本的数据集 D , 其中每个样本 (u, i, r_{ui}) 表示用户 u 对商品 i 写了一条评论 r_{ui} 。如图 1 所示, 本文的核心任务是训练出一个模型根据全部用户与全部商品交互(不包含用户 u 与商品 i 的交互)来学习用户 u 、商品 i 的关联关系表示, 同时根据用户 u 的评论集(用户 u 除 r_{ui} 外的评论)、商品 i 的评论集(商品 i 除 r_{ui} 外的评论)学习用户 u 与商

品 i 的一般偏好表示, 综合两种特征表示预测用户 u 对商品 i 的评分 P_{ui} 。



图 1 模型核心方法

Fig. 1 Model's core method

使用相同方法得到用户 u 对全部商品的评分集合 P_u , 根据该评分集合对用户 u 给出前 K 个商品推荐, 最终任务是使该推荐集合其更接近用户 u 未来的购买行为。表 1 给出了本文使用的相关运算符号及其定义。

表 1 符号定义

Tab. 1 Symbol definitions

符号	定义
\odot	对于两维度相同向量 a 、 b , $a \odot b$ 表示 a 与 b 之间对应元素相乘
\otimes	矩阵乘法
\oplus	两维度为 d 的向量 a 、 b , $a \oplus b$ 结果为二者的 $2 \times d$ 维的拼接向量
$*$	卷积操作

2.2 LGDL 模型框架

以商品 i_i 和用户 u_3 为例, LGDL(LightGCN and Deep Learning)模型详细结构图如图 2 所示, 其中包含三个主要模块: 嵌入层、前向传播层和评分预测层。嵌入层负责利用商品与用户交互信息及商品和用户评论文本对商品和用户建模。前向传播层中, 嵌入层得到的商品与用户的交互信息建模表达 $e_i^{(0)}$ 与 $e_{u_3}^{(0)}$ 传入两个并行 GCN 网络, 来提取商品与用户的高阶信息, 其中上标 0 表示为初始嵌入, 即也就是前向传播层中的第 0 跳表达。然后通过自注意力机制得到最终商品与用户的关联关系的最终表达 e'_i 与 e'_{u_3} ; 嵌入层得到的商品与用户的评论文本信息建模表达 b_{i_i} 与 b_{u_3} 传入两个并行深度学习网络(包含卷积层、最大池化层、Dropout 和全连接层), 提取商品与用户的一般偏好信息, 得到最终商品与用户的一般偏好的最终表达 b'_i 与 b'_{u_3} 。两种表达在评分预测层进行特征融合得出最终商品与用户的表达 I_i 和 U_3 , 并计算得出用户 u_3 对商品 i_i 的预测评分, 根据用户 u_3 对不同商品的预测评分构成集合 P_u , 根据该集合的预测评分排序给出用户 u_3 前 K 个推荐的商品内容。

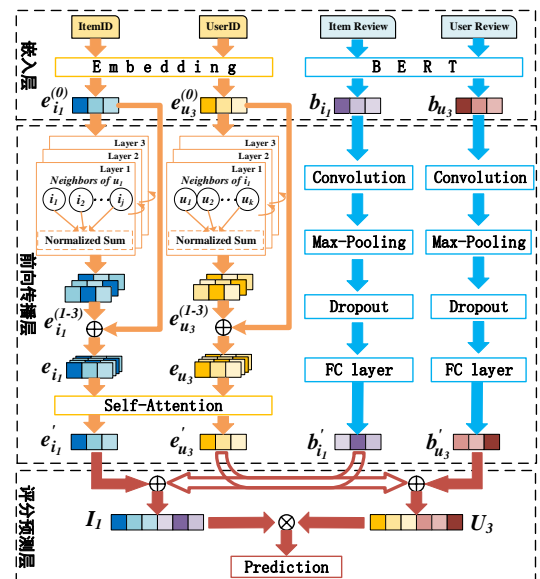


图 2 LGDL 模型结构

Fig. 2 Model structure of LGDL

2.3 嵌入层

嵌入层负责将用户商品 ID 及用户商品评论信息输入模型。

2.3.1 ID 特征嵌入

假定共有 N 个用户, M 个商品, 第 i 个用户的 ID 嵌入向量可以表示为 $e_u^{(0)} \in R^d$, 第 j 个商品的 ID 嵌入向量可以表示为 $e_{i_j}^{(0)} \in R^d$, 其中 d 为嵌入向量的维度, 可以视为可调整的超参数。全部用户的 ID 嵌入向量构成集合 $e_u^{(0)} \in R^{N \times d}$, 即:

$$e_u^{(0)} = [e_{u_1}^{(0)}, \dots, e_{u_N}^{(0)}] \quad (1)$$

同理全部商品的 ID 嵌入向量构成集合 $e_{i_j}^{(0)} \in R^{M \times d}$, 即:

$$e_{i_j}^{(0)} = [e_{i_1}^{(0)}, \dots, e_{i_M}^{(0)}] \quad (2)$$

商品与用户的 ID 嵌入向量均为初始状态, 通过在前向传播层进行传播来进一步细化嵌入, 使 ID 嵌入向量可以更好地表达其内含的关联关系。

2.3.2 评论文本特征嵌入

为简化计算, 假定每个用户与每个商品的评论文本的数量为 RN 条, 每条评论的长度为 RL , BERT 预训练模型的输出维度为 c 维, 第 i 个用户的评论嵌入向量可以表示为 $b_{u_i} \in R^{RN \times RL \times c}$, 第 j 个商品的评论嵌入向量可以表示为 $b_{i_j} \in R^{RN \times RL \times c}$ 。类似 ID 特征嵌入, 共有 N 个用户, M 种商品, 则全部用户的评论嵌入向量构成集合 $b_u \in R^{N \times RN \times RL \times c}$, 全部商品的评论嵌入向量构成集合 $b_{i_j} \in R^{M \times RN \times RL \times c}$ 。 b_u 与 b_{i_j} 中包含了评论文本信息的先验知识, 在前向传播层通过深度学习对其进行特征提取来提取用户与商品的一般偏好。

2.4 前向传播层

前向传播层分为两个并行框架来分别提取用户与商品的关联关系及一般偏好。

2.4.1 关联关系提取

假定全部已知关联关系所构成的二分图为 G 。对于用户与商品的关联关系提取采用类似 LightGCN 的方法。以商品 i_j 和用户 u_i 为例, 用户 u_i 的 ID 嵌入在图卷积网络传播一次(即一跳)的计算规则为

$$e_{u_i}^{(1)} = \sum_{i_j \in N_{u_i}} \frac{1}{\sqrt{|N_{u_i}|} \sqrt{|N_{i_j}|}} e_{i_j}^{(0)} \quad (3)$$

其中, $e_{u_i}^{(1)}$ 表示用户 u_i 在图卷积中第 1 跳的表达, $e_{i_j}^{(0)}$ 表示商品 i_j 在图卷积中第 0 跳的表达, 即嵌入层的 ID 嵌入。 $1/(\sqrt{|N_{u_i}|} \sqrt{|N_{i_j}|})$ 为遵循了 GCN 原始设计中的聚合操作, 其中 N_{u_i} 表示所有包含用户 u_i 的所有邻居节点集合, N_{i_j} 表示所有包含商品 i_j 的邻居节点集合。简言之, 式(3)将用户 u_i 的所有邻居节点中的所有商品节点的初始 ID 嵌入进行了聚合操作, 从而得到了用户 u_i 在图中第一跳的嵌入表达。

同理, 商品 i_j 的 ID 嵌入在网络中传播一次的计算规则为

$$e_{i_j}^{(1)} = \sum_{u_i \in N_{i_j}} \frac{1}{\sqrt{|N_{i_j}|} \sqrt{|N_{u_i}|}} e_{u_i}^{(0)} \quad (4)$$

节点在图卷积网络中的一阶传播(式(3), (4))建模了用户和项目之间的一阶关联关系特征, 利用一阶传播的计算方法, 可以在图卷积网络中, 堆叠多层图卷积来建模用户与项目之间的高阶关联关系特征。可以推出, 用户 u_i 和商品 i_j 在前向传播层中图卷积部分中第 k 跳到第 $k+1$ 跳的前向传播规则(节点聚合规则)定义如下:

$$\begin{cases} e_{u_i}^{(k+1)} = \sum_{i_j \in N_{u_i}} \frac{1}{\sqrt{|N_{u_i}|} \sqrt{|N_{i_j}|}} e_{i_j}^{(k)} \\ e_{i_j}^{(k+1)} = \sum_{u_i \in N_{i_j}} \frac{1}{\sqrt{|N_{i_j}|} \sqrt{|N_{u_i}|}} e_{u_i}^{(k)} \end{cases} \quad (5)$$

其中, $e_{u_i}^{(k+1)}$ 和 $e_{i_j}^{(k+1)}$ 表示用户 u_i 和商品 i_j 在图卷积中第 $k+1$ 跳的嵌入表达, $e_{u_i}^{(k)}$ 和 $e_{i_j}^{(k)}$ 表示用户 u_i 和商品 i_j 在图卷积中第 k 跳的嵌入表达。

最终得到用户 u_i 和商品 i_j 的表达 $e_{u_i}, e_{i_j} \in R^{k \times d}$ 为

$$\begin{cases} e_{u_i} = e_{u_i}^{(0)} \oplus e_{u_i}^{(1)} \oplus \dots \oplus e_{u_i}^{(k)} \\ e_{i_j} = e_{i_j}^{(0)} \oplus e_{i_j}^{(1)} \oplus \dots \oplus e_{i_j}^{(k)} \end{cases} \quad (6)$$

为更好的得到图神经网络的最终表达, 采用自注意力机

制(Self-Attention)来建模每一层嵌入的重要程度。对于用户 u_i , 其图神经网络的注意力 $att_{u_i} \in R^{1 \times k}$ 的计算方式如下:

$$att_{u_i} = \text{softmax}(\omega_1 \otimes \tanh(\omega_2 \otimes e_{u_i}^{(t)})) \quad (7)$$

其中, $\omega_1 \in R^{1 \times t}$, $\omega_2 \in R^{t \times d}$, t 可视为可调整的超参数。 att_{u_i} 中包含了第 0 层到第 k 层的嵌入表达的权重 $att_{u_i}^{(0)} \sim att_{u_i}^{(k)}$, $\text{softmax}()$ 函数用于将 k 层嵌入的权重归一化。

利用注意力向量将每层的嵌入向量加权求和可以得到最终的用户 u_i 的关联关系嵌入表达 $e'_{u_i} \in R^d$ 为

$$e'_{u_i} = \sum_{z=0}^k att_{u_i}^{(z)} e_{u_i}^{(z)} \quad (8)$$

同理, 商品 i_j 的关联关系嵌入最终表达 $e'_{i_j} \in R^d$ 为

$$e'_{i_j} = \sum_{z=0}^k att_{i_j}^{(z)} e_{i_j}^{(z)} \quad (9)$$

2.4.2 偏好特征提取

对于用户 u_i 的评论嵌入向量 $b_{u_i} \in R^{RN \times RL \times c}$, 假定有 m 个卷积核, 每个卷积核 $K_{k_i} \in R^{k_s \times c}$, 其中 k_s 为可调整的卷积核宽度, 对于用户 u_i 的一般偏好的特征提取首先经过卷积操作:

$$z_{u_i} = \text{LeakyReLU}(\text{reshape}(b_{u_i}) * K_{k_i} + \mu_{u_i}) \quad (10)$$

其中, reshape 操作将 b_{u_i} 的维度转换为二阶张量, 使 $b_{u_i} \in R^{(RN \times RL) \times c}$, $*$ 表示卷积操作, μ_{u_i} 为偏差。在完成一次卷积操作后所得的向量 $z \in R^{(RN \times RL) \times k_s + 1}$, 由于存在 m 个卷积核, 卷积操作会产生 m 个向量 z 。激活函数公式如式(11), 其中 a 为可自行设置的函数参数。

$$\text{LeakyReLU} = \begin{cases} x & x \geq 0 \\ ax & x < 0 \end{cases} \quad (11)$$

经过卷积操作后, 对于卷积层输出的每一个向量 z_{u_i} 进行最大池化(Max-Pooling)操作, 来进一步提取特征。具体做法是求 z_{u_i} 的最大值 o_{u_i} , 见式(12), 然后将这些组成池化层的输出 $O \in R^m$, 见式(13)。

$$o_{u_i} = \max\{z_{u_{i1}}, z_{u_{i2}}, \dots, z_{u_{i(RN \times RL) \times k_s + 1}}\} \quad (12)$$

$$O = [o_1, o_2, \dots, o_m] \quad (13)$$

在池化层之后, 将得到的特征 o 输入全连接层(Fully Connected Layer)。首先使用伯努利分布生成随机失活向量 $p_{u_i} \in R^m$, 见式(14), 随机失活比率 $(1-\beta)$ 为可调整超参数。那么, 全连接层的传播计算公式为式(15)。

$$p_{u_i} \sim B(m, \beta) \quad (14)$$

$$b'_{u_i} = \text{LeakyReLU}(\omega_3 \otimes (p_{u_i} \odot O) + g_{u_i}) \quad (15)$$

其中, 权重矩阵 $\omega_3 \in R^{d \times m}$, d 与用户的 ID 嵌入的维度相同, g_{u_i} 为偏差, $b'_{u_i} \in R^d$ 为用户 u_i 的最终一般偏好特征。同理可计算出商品 i_j 的最终一般偏好特征 $b'_{i_j} \in R^d$ 。

2.5 评分预测层

经过前向传播层, 可以得到用户 u_i 和商品 i_j 的关联关系特征向量 $e'_{u_i}, e'_{i_j} \in R^d$ 和一般偏好特征向量 $b'_{u_i}, b'_{i_j} \in R^d$, 本文采用 concat 方法对两特征向量进行特征融合, 从而得到最终用户 u_i 和商品 i_j 的特征表示 $U_i, I_j \in R^{2d}$, 见式(16)。

$$U_i = e'_{u_i} \oplus b'_{u_i}, I_j = e'_{i_j} \oplus b'_{i_j} \quad (16)$$

用户 u_i 对商品 i_j 最终评分预测为

$$P_{ij} = U_i \otimes I_j^T \quad (17)$$

2.6 模型优化

为了学习模型参数, 使模型可以更好的建模用户与商品的特征, 本文采用推荐系统中使用广泛的 BPR[28]损失。该损失基于贝叶斯排序, 考虑到了可观察到和不可观察到的用户与物品交互的相对顺序, 认为观察到的交互项的重要程度比不可观察到的交互项的重要程度更高。BPR 损失的计算方式如下:

$$\text{Loss} = \sum_{(u,a,b) \in D} -\ln \sigma(P_{ua} - P_{ub}) + \lambda \|\Theta\|_2^2 \quad (18)$$

其中, $D = \{(u,a,b) | (u,a) \in R^+, (u,b) \in R^-\}$ 代表成对的训练数据, R^+

代表正样本即观察到的交互项, R 代表负样本即未观测到的交互项, σ 代表 $\text{sigmoid}()$ 函数。 $\lambda\|\Theta\|_F$ 为正则项, 其中 $\Theta = \{E, \{\omega_1^{(i)}, \omega_2^{(i)}, \omega_3^{(i)}\}_{i=1}^L\}$ 表示模型中所有可训练的参数, λ 为正则项的可调整系数, 用来控制模型参数大小防止过拟合。本文模型采用随机梯度下降法训练模型网络, 使用 Adam 优化器来优化模型参数, 最小化损失函数。

3 实验结果与分析

3.1 数据集及评价指标

亚马逊数据集作为推荐系统使用最广泛的数据集之一, 收录了 1996 年至 2018 年的 29 类商品购买记录, 其中包含了商品 ID、用户 ID、评论文本、商品标签等诸多数据信息。本文采用四个 5-core 版本(每个商品/用户至少包含 5 条交互记录)的数据集作为本文模型实验验证的数据集, 包括: Automotive(Auto)、Toys_and_Games(TG)、Tools_and_Home_Improvement(THI)、Sports_and_Outdoors(SO)。本文模型使用了数据集中的商品 ID、用户 ID 及评论文本数据信息。

为了使模型训练的效果更好, 在进行实验之前对原始数据集进行了预处理。首先从总数据集中提取所需数据, 其次由于评论文本数据存在空值, 评论字数过少等情况, 需要对评论文本数据进行数据清洗。由于空评论无法反映用户对商品的偏好, 即使进行随机填充也无法正确表达用户偏好, 因此去除空值评论数据; 对于评论字数过少的数据通过对评论进行复制填充, 对于评论字数过长的数据进行删除(无意义、意义相对较低的词, 符号等), 使评论字数长度统一为 $RL=50$ 。通过对用户 ID 和商品 ID 进行统计分析, 制作 10-core 版本的数据集。预处理后数据集统计信息如表 2 所示。

表 2 数据集统计信息

数据集	用户数	商品数	交互数	稀疏度
Auto	15 280	8 157	226 477	0.001 82
TG	15 780	10 345	267 628	0.001 64
THI	26 988	14 240	417 492	0.001 09
SO	33 816	17 142	533 041	0.000 92

本文采取 TopK 推荐方法进行推荐, 其中 $K=20$, 采用了召回率 Recall, 归一化折损累计增益 NDCG (Normalized Discounted Cumulative Gain) 两种评价指标来评价模型性能。Recall 用来衡量 Top20 推荐中, 推荐列表中用户有过交互行为的商品数目占测试集中用户所有有过交互行为的商品数目的比例, 召回率越高模型效果越好。假定测试集全部用户集合为 U , 对任意用户 $u \in U$, Top20 的推荐列表为 L_u , 测试集中用户 u 的真实交互列表为 L_u^{test} , 则模型召回率的计算公式为

$$\text{recall} @ 20 = \sum_{u \in U} \frac{|L_u \cap L_u^{test}|}{|L_u^{test}|} \quad (19)$$

NDCG 衡量了推荐列表中不同位置推荐结果的相关性得分, 与用户相关性越高的推荐商品的排序越靠前, 其推荐效果越好得分也越高。NDCG 综合评估了全部用户的推荐列表的推荐质量。假定 L_u 为 Top20 推荐列表中第 i 个位置的推荐, $f(x)$ 在 $x > 0$ 时取 1, 否则取 0。假定相关性得分 $rel \in \{0, 1\}$, 则 NGCD 计算公式为

$$\text{NDCG} @ 20 = \sum_{u \in U} \frac{1}{|U|} \frac{\text{DCG}_u @ 20}{\text{IDCG}_u @ 20} = \sum_{u \in U} \frac{1}{|U|} \frac{\sum_{i=1}^{20} \frac{f(L_u \cap L_u^{test})}{\log_2(i+1)}}{\sum_{i=1}^{20} \frac{f(L_u^{test})}{\log_2(i+1)}} \quad (20)$$

3.2 对比模型

本文采用以下对比模型进行实验对比:

a) LFM 是最为经典的矩阵分解算法, 仅利用了用户与商品的交互去学习用户与商品的隐含特征, 以商品评分为预测目标进行学习, 其推荐效果受数据稀疏性影响。

b) NeuMF^[29]是矩阵分解的深度学习版本, 其在用户和商品嵌入在元素级别和连接方式上增加了多个隐层, 来刻画用户与商品的非线性交互。

c) NGCF 遵循标准 GCN 模型, 成功将 GCN 运用在推荐系统中。其成功利用二分图来提取用户和商品之间的嵌入相互作用(即关联关系)来获取评分预测。

d) Light-GCN 在 NGCF 的基础上去掉了冗余的特征变换部分和激活函数部分, 大大提高了模型的运行效率和预测准确度。

3.3 实验方案与模型的超参数调整

为确保实验对比的可行性与公平性, 将数据集划分训练集、验证集和测试集, 比例为 8:1:1。每个用户与每个商品的评论文本的数量蛇者为 $RN=10$ 条。四个对比模型的最终嵌入向量的维度与本文模型 LGDL 的维度相同。所有模型的卷积层卷积核大小均设置为 3, 步长设置为 1, 并且不使用 padding 策略。所有实验均通过深度学习框架 Pytorch 实现。本文模型最终使用的相关参数见表 3。

表 3 模型参数表

参数名称	参数含义	参数值
d	嵌入向量的维度	64
c	BERT 预训练模型的输出维度	128
a	LeakyReLU 函数内部参数	0.01
β	Dropout 中保留比例	0.6
t	注意力机制的维度	64
lr	学习率	5e-5
la	图卷积迭代的次数	2
λ	正则化权重	5e-5
batchsize	批处理大小	128

为使模型效果达到最佳, 本文在实验部分对相关超参数进行了调整, 其中三项超参调整方式如下:

a) 正则化权重 λ 的调整。如表 4 所示, 采用控制变量法对正则化权重进行调整, 其中图卷积迭代次数 $la=3$, 嵌入向量维度 $d=64$, 其余参数设置参照表 3。在表 4 中分析可得在该参数为 $2e-5$ 至 $1e-5$ 时可达模型效果最佳。注意到在数据集 SO 中, 模型召回率效果从 $5e-5$ 变至 $2e-5$ 时发生了降低, 考虑到模型的泛化性能, 正则化权重采用 $5e-5$ 。

表 4 正则化权重调整对比实验

λ	Auto		TG		THI		SO	
	Recall	NDCG	Recall	NDCG	Recall	NDCG	Recall	NDCG
5e-4	9.8241	6.5283	8.8096	5.7302	5.2510	3.4672	6.2045	4.1539
1e-4	10.2942	6.9420	9.1509	6.0912	5.5827	3.9675	6.4925	4.4673
5e-5	10.3401	7.0078	9.4051	6.2600	5.6102	4.0349	6.5177	4.5005
2e-5	10.4013	7.1638	9.4563	6.3137	5.6492	4.1366	6.5169	4.5198
1e-5	10.4209	7.2014	9.4511	6.2848	5.6714	4.0348	6.5188	4.5250

b) 图卷积迭代次数 la 的调整。如表 5 所示, 对图卷积迭代次数调整的对比实验中, 嵌入向量维度 $d=64$, 其余参数设置参照表 3。通过分析表 5 中对比实验的数据, 可以很明显看出在 $la=2$ 时, 模型效果达到最优, 因此最终迭代层数确定为 2 层。考虑到数据集的稀疏性, 出现这种情况可以判断出四个数据集分别所构建的图神经网络中绝大部分节点存在 2 跳至 3 跳路径, 少部分存在 3 跳及以上路径(可通过数据集验证)。这种情况下, 在 $la=2$ 时可以提取到相比 $la=1$ 时更多的相关关系, 模型效果会相对变好; $la=3$ 并不能提取到比 $la=2$ 更多的相关关系, 因此模型效果不会变好, 反而会相对降低。

c) 嵌入向量维度 d 的调整。如表 6 所示对比实验, 只调整了嵌入向量维度, 其余参数设置参照表 3。通过分析可以看出随嵌入向量维度的增加, 模型效果越好。注意到在四个

数据集中 $d=16$ 到 32、64 模型效果的平均增长幅度相比 $d=64$ 入向量所占用空间翻倍增长, 模型效果增长幅度相对未达到到 128 的增长幅度要大, 考虑到 64 维嵌入向量到 128 维嵌入预期涨幅, 故嵌入向量维度最终采用 64 维。

表 5 图卷积迭代次数调整对比实验

Tab. 5 Contrast experiments of the iteration number of GCN

la	Auto		TG		THI		SO	
	Recall/%	NDCG/%	Recall/%	NDCG/%	Recall/%	NDCG/%	Recall/%	NDCG/%
1	9.8397	6.9274	9.1737	6.2981	5.2834	3.9126	6.5456	4.6557
2	10.5950	7.3966	9.8583	6.6669	5.8646	4.2626	6.8748	4.8656
3	10.3401	7.0078	9.4051	6.2600	5.6102	4.0349	6.5177	4.5005
4	9.7383	6.4130	8.7148	5.6458	5.1811	3.6233	5.9357	4.0745
5	9.0009	5.8678	7.9250	5.0405	4.6038	3.1567	5.4012	3.5982

表 6 嵌入向量维度调整对比实验

Tab. 6 Contrast experiments of the dimension of embeddings

d	Auto		TG		THI		SO	
	Recall/%	NDCG/%	Recall/%	NDCG/%	Recall/%	NDCG/%	Recall/%	NDCG/%
16	8.3111	5.4323	7.4561	4.7991	4.0521	2.5438	5.1063	3.3493
32	9.3789	6.3375	8.5416	5.5009	4.8329	3.2502	5.8150	3.9669
64	10.5950	7.3966	9.8583	6.6669	5.8646	4.2626	6.8748	4.8656
128	11.3703	8.2850	10.5248	7.3331	6.5334	4.8986	7.7225	5.5809

3.4 性能对比

本文在四个数据集上对 5 个模型进行了实验对比, 对比结果如表 7 所示, 其中粗体代表模型效果最佳结果。

a)MF 模型在四个数据集的表现均不佳, 由于数据集稀疏度都达到千分之一, 说明矩阵分解在极度稀疏的数据集中并不足以更好的捕捉用户与商品的特性。

b)NeuMF 模型相比 MF 增加了隐层, 从实验结果来看, NeuMF 的模型效果均优于 MF, 平均实验效果增加 3%(四个数据集中 NeuMF 相比 MF 的召回率和归一化折损累计增益提升效果相比 MF 的平均增幅)。实验表明该模型隐层的增加确实可以进一步提取用户与商品之间的非线性特征交互, 但其本质上还是 MF 的一种, 因此模型效果提升并不是特别明显。

c)NGCF 模型将 GCN 引入推荐算法中, 从实验结果来看, 其相比 NeuMF 模型的模型效果提升了约 19%, 提升效果明显, 说明图神经网络对于极度稀疏的数据集具有缓解数据稀疏性的作用, 且对提升推荐效果有相当贡献。

d)Light-GCN 模型相比 NGCF 在图卷积神经网络部分去除了冗余的特征变换及非线性变换部分。从实验结果来看,

其相比 NGCF 的模型效果约提升了 21%, 提升效果明显, 可以看出 Light-GCN 的改进的有效性。

e)本文模型 LGDL 与其他模型对比可以看出, 其在四个数据集的模型表现均优于其他模型, 召回率在数值上至少有 0.15% 的提升, NDCG 在数值上至少有 0.37% 的提升。相比次优模型 Light-GCN 模型效果相对提升了约 5%。证明本文模型设计的方法对于模型效果的提升有相应贡献。

3.5 消融实验

本文针对 LGDL 模型的各个部分设置了四种变体来进行消融实验, 以验证模型的有效性。模型对比结果如图 3 所示, 各模型变体描述如下: a)LGDL-1 表示模型取消注意力机制部分且不使用评论文本嵌入数据, 类似于 Ligh-GCN 模型。以用户为例, 即将式(8)修改为 $e_u = \sum e_u^{(i)}$, 且最终模型的用户特征表达为 $U_i = e_u$; b)LGDL-2 表示模型增加了注意力机制部分, 且不使用评论文本嵌入数据。以用户为例, 即将式(16)修改为 $U_i = e_u$, 其余部分不变。c)LGDL-3 表示模型只使用评论文本嵌入数据进行推荐。以用户为例, 即将式(16)修改为 $U_i = b_u$, 其余部分不变。

表 7 5 种模型在 4 个数据集上的性能比较

Tab. 7 Performance comparison of five models on four datasets

model	Auto		TG		THI		SO	
	Recall/%	NDCG/%	Recall/%	NDCG/%	Recall/%	NDCG/%	Recall/%	NDCG/%
MF	5.9932	4.2271	5.8083	3.8465	3.7922	2.6783	5.4156	3.8993
NeuMF	6.1849	4.3606	5.9604	3.9380	3.9352	2.7445	5.6187	4.0232
NGCF	7.9051	5.0963	7.1550	4.8682	4.5501	3.2923	6.3813	4.7384
Light-GCN	9.8707	6.7869	9.3938	6.2929	5.7166	3.8930	6.8748	4.8656
LGDL	10.5950	7.3966	9.8583	6.6669	5.8646	4.2626	7.3819	5.2968
Improve	0.7243	0.6097	0.4645	0.3740	0.1480	0.3696	0.5071	0.4312

从图 3 中可以看出, LGDL-1 采用简单的 GCN 层嵌入的融合效果, 相比 Light-GCN 效果略低。LGDL-2 相比 LGDL-1 增加了注意力机制模型效果明显增加, 说明注意力机制模块在模型中是有效的。LGDL-3 只采用评论文本数据, 相比前两个变体效果略低, 但结合评论文本及注意力机制的 LGDL 的效果比单使用注意力机制的 LGDL-2 效果要好, 说明评论文本中提取的用户商品一般偏好特征在最终特征表达中作出了贡献, 即添加评论文本数据处理模块对模型效果增加是有贡献的。

3.6 案例分析

为了更好地理解推荐过程, 以数据集 Automotive 中的用户 232 为例, 采用 3 跳传播, top10 推荐进行案例分析, 具体过程如图 4 所示。

首先用户 232 本身构成了第一个嵌入, 即 Embedding_0。其次用户 232 的一阶邻居, 即其直接购买的商品有商品 16, 商品 375, 商品 7296 等, 这些商品共同生成用户 232 的一阶嵌入, 即 Embedding_1。以商品 16 的下一跳为例, 商品 16 曾被用户 40、211、4696 等人购买过, 即为用户 232 的部分

二阶邻居。全部二阶邻居聚合出用户 232 的二阶嵌入 Embedding_2。同理全部三阶邻居聚合出用户 232 的三阶嵌入, 即 Embedding_3。

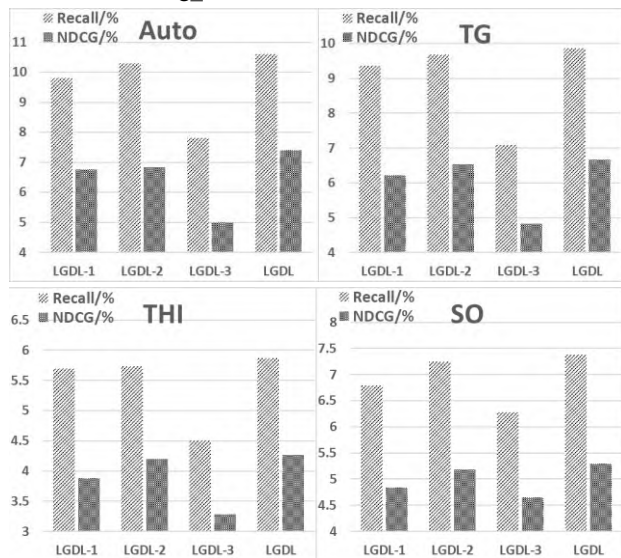


图 3 消融实验对比分析

Fig. 3 Comparative analysis of ablation study

注意力机制为每一阶嵌入分配注意力权重为 0.147、0.203、0.367 和 0.283。可以看出 2 阶嵌入所占比重最大, 其次是三阶嵌入, 说明其在最终用户表达中发挥了更重要的作用, 也同时验证了图卷积中迭代 2 次效果最佳的有效性。

用户 232 的商品评论集合经过 BERT 模型等一系列操作生成了用户一般偏好 Bert_Embedding。该偏好与根据注意力权重得到的用户嵌入共同指导, 给出了用户 232 的商品推荐序列, 即商品 171、6244、325 等。观察到商品推荐序列中存在商品 6244 与商品 41, 这两种商品的推荐与之前图中嵌入生成过程一致, 证明了该模型具有一定的可解释性。

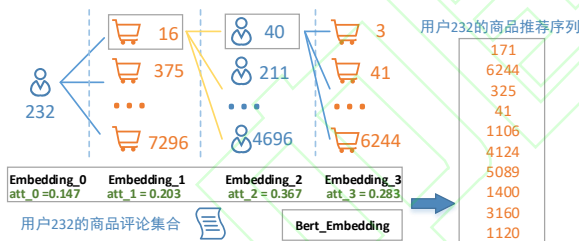


图 4 案例分析

Fig. 4 Case analysis

4 结束语

目前大部分基于图神经网络的推荐算法对评论文本利用不足。本文提出一种利用图神经网络优势提取用户商品相关关系, 同时利用深度学习优势来提取用户商品的一般偏好的混合推荐算法。同时采用自注意力机制优化了图卷积部分嵌入向量的最终表达。在亚马逊四种数据集上进行了实验验证, 并进行了消融实验。结果表明, 本文提出的方法对提高推荐效果有一定的作用。未来将在评论文本处理方面采用更先进的框架, 采用不同的特征融合机制, 探究不同特征融合对推荐系统的影响。

参考文献:

- [1] 姚金平. 个性化推荐系统综述 [J]. 中国集体经济, 2020 (25): 71-72. (Yao Jinping. Survey of personalized recommendation system [J]. China Collective Economy, 2020 (25): 71-72.)
- [2] LeCun Y, Bengio Y, Hinton G. Deep learning [J]. Nature, 2015, 521 (7553): 436-444.

- [3] He Xiangnan, Deng Kuan, Wang Xiang, *et al.* LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation [C]// Proceedings of the 43rd International ACM SIGIR Conference, 2020: 639-648.
- [4] Funk S. Netflix update: Try this at home [EB/OL]. (2006) [2021-5-27] <http://sifter.org/~simon/journal/20061211.html>.
- [5] Koren Y. Factorization meets the neighborhood: a multifaceted collaborative filtering model [C]// Proceedings of International Conference on Knowledge Discovery and Data Mining, 2008: 426-434.
- [6] Salakhutdinov R, Mnih A. Probabilistic matrix factorization [C]// Advances in Neural Information Processing Systems, 2007: 1257-1264.
- [7] Ling Guang, Lyu M R, King I. Ratings meet reviews, a combined approach to recommend [C]// Proceedings of the 8th ACM Conference on Recommender systems. ACM, 2014: 105-112.
- [8] Bao Yang, Fang Hui, Zhang Jie. TopicMF: Simultaneously exploiting ratings and reviews for recommendation [C]// Proceedings of 28th AAAI conference on Artificial Intelligence. ACM, 2014: 2-8.
- [9] 王永, 冉冉, 尹恩民, 等. 满足差分隐私保护的矩阵分解推荐算法 [J]. 电子科技大学学报, 2021, 50 (03): 405-413. (Wang Yong, Ran Xun, Yin Enmin, *et al.* Matrix factorization recommendation algorithm for differential privacy protection [J]. Journal of University of Electronic Science and Technology of China, 2021, 50 (03): 405-413.)
- [10] Kim D, Park C, Oh J, *et al.* Convolutional matrix factorization for document context-aware recommendation [C]// Proceedings of the 10th ACM Conference on Recommender Systems. ACM, 2016: 233-240.
- [11] Zheng Lei, Noroozi V, Yu P S. Joint deep modeling of users and items using reviews for recommendation [C]// Proceedings of the 10th ACM International Conference on Web Search and Data Mining. ACM, 2017: 425-434.
- [12] 田保军, 刘爽, 房建东. 融合主题信息和卷积神经网络的混合推荐算法 [J]. 计算机应用, 2020, 40 (07): 1901-1907. (Tian Baojun, Liu Shuang, Fang Jiandong. Hybrid recommendation algorithm by fusion of topic information and convolution neural network [J]. Journal of Computer Applications, 2020, 40 (07): 1901-1907.)
- [13] Zhou Guorui, Zhu Xiaoqiang, Song Chenru, *et al.* Deep interest network for click-through rate prediction [C]// Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2018: 1059-1068.
- [14] 邱叶, 邵雄凯, 高榕, 等. 基于注意力门控神经网络的社会化推荐算法 [J/OL]. 计算机工程与应用: 1-9 [2021-07-04]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20210602.1413.014.html>. (Qiu Ye, Shao Xiongkai, Gao Rong, *et al.* Social recommendation algorithm based on attention gated neural network [J/OL]. Computer Engineering and Applications: 1-9 [2021-07-04]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20210602.1413.014.html>.)
- [15] Cheng H T, Koc L, Harmsen J, *et al.* Wide & deep learning for recommender systems [C]// Proceedings of the 1st workshop on deep learning for recommender systems. 2016: 7-10.
- [16] 高铭蔚, 桑楠, 杨茂林. 基于胶囊网络的交互式网络电视视频点播推荐 [J/OL]. 计算机应用: 1-9 [2021-07-04]. <http://kns.cnki.net/kcms/detail/51.1307.TP.20210507.1322.006.html>. (Gao Mingwei, Sang Nan, Yang Maolin. IPTV video-on-demand recommendation based on capsule network [J/OL]. Journal of Computer Applications: 1-9 [2021-07-04]. <http://kns.cnki.net/kcms/detail/51.1307.TP.20210507.1322.006.html>.)
- [17] Yang Cheng, Sun Maosong, Zhao Wayne Xin, *et al.* A Neural Network Approach to Joint Modeling Social Networks and Mobile Trajectories [J]. ACM Transactions on Information Systems, 2016, 35 (4): 36.

- [18] Wu Chaoyuan, Ahmed A, Beutel A, *et al.* Recurrent Recommender Networks [C]// 10th ACM International Conference on Web Search & Data Mining. ACM, 2017.
- [19] Chen Qiwei, Zhao Huan, Li Wei, *et al.* Behavior sequence transformer for e-commerce recommendation in alibaba [C]// Proceedings of the 1st International Workshop on Deep Learning Practice for High-Dimensional Sparse Data. 2019: 1-4.
- [20] Scarselli F, Gori M, Tsoi A C, *et al.* The Graph Neural Network Model [J]. IEEE Transactions on Neural Networks, 2009, 20 (1): 61-80.
- [21] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks [J]. arXiv preprint arXiv: 1609. 02907, 2016.
- [22] Wang Xiang, He Xiangnan, Wang Meng, *et al.* Neural Graph Collaborative Filtering [C]// Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, 2019: 165-174.
- [23] 李世宝, 张益维, 刘建航, 等. 基于知识图谱共同邻居排序采样的推荐模型 [J/OL]. 电子与信息学报: 1-8 [2021-07-04]. <http://kns.cnki.net/kcms/detail/11.4494.TN.20210121.1119.002.html>. (Li Shibao, Zhang Yiwei, Liu Jianhang, *et al.* Recommendation model based on public neighbor sorting and sampling of knowledge graph [J/OL]. Journal of Electronics & Information Technology: 1-8 [2021-07-04]. <http://kns.cnki.net/kcms/detail/11.4494.TN.20210121.1119.002.html>.)
- [24] Wang Xiang, He Xiangnan, Cao Yixin, *et al.* KGAT: Knowledge Graph Attention Network for Recommendation [C]// Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019: 950-958.
- [25] 王建芳, 文茜琳, 杨许, 等. 基于偏差的图注意力神经网络推荐算法 [J/OL]. 控制与决策: 1-9 [2021-07-04]. <https://doi.org/10.13195/j.kzyjc.2020.1626>. (Wang Jianfang, Wen Xilin, Yang Xu, *et al.* A bias-based graph attention neural network recommender algorithm [J/OL]. Control and Decision: 1-9 [2021-07-04]. <https://doi.org/10.13195/j.kzyjc.2020.1626>.)
- [26] Sun Jianing, Zhang Yingxue, Guo Wei, *et al.* Neighbor Interaction Aware Graph Convolution Networks for Recommendation [C]// Proceedings of the 43rd International ACM SIGIR Conference, 2020: 1289-1298.
- [27] 苏静, 许天琪, 张贤坤, 等. 基于图卷积与外积的协同过滤推荐模型 [J/OL]. 计算机应用研究: 1-5 [2021-07-04]. <https://doi.org/10.19734/j.issn.1001-3695.2021.02.0053>. (Su Jing, Xu Tianqi, Zhang Xiankun, *et al.* Collaborative filtering recommendation model based on graph convolution and cross product [J/OL]. Application Research of Computers: 1-5 [2021-07-04]. <https://doi.org/10.19734/j.issn.1001-3695.2021.02.0053>.)
- [28] Rendle S, Freudenthaler C, Gantner Z, *et al.* BPR: Bayesian personalized ranking from implicit feedback [C]// Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence, Montreal, QC, Canada, June 18-21, 2009. AUAI Press, 2009.
- [29] He Xiangnan, Liao Lizi, Zhang Hanwang, *et al.* Neural collaborative filtering [C]// Proceedings of the 26th International Conference on World Wide Web. ACM, 2017: 173-18.