

文章编号:1007-5321(2021)05-0088-06

DOI:10.13190/j.jbupt.2021-014

融合文本信息的轻量级图卷积网络推荐模型

张 栋, 陈鸿龙

(中国石油大学(华东)控制科学与工程学院, 青岛 266580)

摘要: 在基于图卷积网络的推荐模型中,图卷积对仅包含编号信息的输入节点进行信息聚合会引发严重的瓶颈问题,影响推荐精度. 为缓解此问题,考虑通过辅助信息丰富节点特征,提出了一种融合文本信息的轻量级图卷积网络推荐模型. 模型把卷积神经网络提取出文本评论特征添加到图的节点嵌入中. 为了简化图卷积网络结构,采用轻量级图卷积框架将融合的特征信息在用户-电影项目图上线性传播来学习用户和电影项目的嵌入,并将所有图卷积子层上特征嵌入的加权总和作为最终特征输出,用于预测评分. 3 个实际数据集上的实验结果表明,该方法可以缓解信息聚合瓶颈问题,提高推荐的准确度,并且模型可以缓解推荐中的冷启动问题.

关 键 词: 推荐模型; 信息聚合; 图卷积网络; 文本信息

中图分类号: TP393

文献标志码: A

A Lightweight Graph Convolutional Network Recommendation Model Incorporating Text Information

ZHANG Dong, CHEN Hong-long

(College of Control Science and Engineering, China University of Petroleum(East China), Qingdao 266580, China)

Abstract: In the recommendation model based on graph convolution network, the graph convolution only aggregates information from the input nodes with identifier information, which will decrease the recommendation precision and, thus, lead to a bottleneck problem. To solve this problem, a lightweight graph convolution network recommendation model based on text information fusion is proposed by considering enriching node features with auxiliary information. The model extracts text comment features from convolution neural network and adds them to the node embedding of graph. To simplify the structure of graph convolution network, the proposed lightweight graph convolution framework is used to transmit the fused feature information linearly on the user-movie item graph to learn the embedding of the user and movie item. The weighted sum of all sub-levels of the graph convolution is used as the final feature output for predicting the rating. Experimental results on three real datasets show that the proposed method can alleviate the bottleneck problem of information aggregation and improve the accuracy of recommendation. The model can also alleviate the cold start problem.

Key words: recommender model; information aggregation; graph convolutional network; text information

收稿日期: 2021-01-27

基金项目: 国家自然科学基金项目(61772551); 国家自然科学基金国际合作与交流项目(62111530052); 中石油重大科技项目(ZD2019-183-003)

作者简介: 张 栋(1995—), 男, 硕士生.

通信作者: 陈鸿龙(1984—), 男, 副教授, 博士生导师, E-mail: chenhl@upc.edu.cn.

近年来,推荐系统已成为解决信息过载问题的最重要工具之一,它通过分析用户的历史行为和个人属性,对用户的兴趣进行建模,从大量项目中(如商品、电影、音乐、图书等)找出与用户兴趣最为匹配的项目,从而实现个性化推荐.推荐系统作为主动型过滤信息的工具,近年来迅速发展,遍布于电商、影视、金融等各种网站,网站中包含大量用户对推荐项目的文本评论信息.为了充分利用这些信息,通过结合当下最流行的图卷积推荐框架,提出了融合文本信息的轻量级图卷积网络推荐模型,该模型在图卷积节点中能有效地融合文本评论特征,缓解图卷积信息聚合时出现的瓶颈,提升推荐质量.

1 相关工作

早期推荐系统的算法模型有2个主要的分支:协同过滤(CF, collaborative filtering)推荐模型^[1]和基于内容(CB, content-based recommendations)^[2]的推荐模型. CF模型通过考虑用户和项目的交互历史数据来预测用户对项目的感兴趣程度,而CB模型利用用户和项目的内容信息进行推荐.

CF通过学习用户和项目的嵌入特征表示,根据嵌入特征来进行预测^[3]. 矩阵分解^[4]是CF中的一种简单有效的模型,它通过奇异值分解,直接将用户和项目的编号信息投影到特征嵌入,以其内积形式完成预测. 近年来用于表示图结构数据的图嵌入技术^[5]迅速发展,为推荐系统的发展带来了新的契机. 作为图嵌入技术的主要方式之一,图卷积网络(GCN, graph convolutional network)^[6-7]通过图嵌入网络结构获得图或图节点的向量表示,大量作用于社交网络,图像识别及自然语言处理等各个领域.

在基于GCN的推荐模型中,为了令交互信息在图卷积的高阶子图邻域中传递,图神经协同过滤(NGCF, neural graph collaborative filtering)模型^[8]利用图卷积中高阶邻域节点的特征,通过叠加若干卷积层实现多阶邻域的信息传递. 为了简化GCN结构, Wu等^[9]将网络中多个权重矩阵分解为一个,提出简易图卷积网络(SGCN, simplifying graph convolutional networks)模型,用于节点分类任务. 而He等^[10]在NGCF模型上进行消融性实验,证明了在节点只包含编号信息时,图卷积中非线性变换和权值矩阵对推荐是非必要的,提出了轻量级图卷积网络推荐(LightGCN, simplifying and powering graph con-

volution network for recommendation)模型. 但随着数据量的增加,节点仅依靠编号信息将不足以区分不同用户和电影项目的特征,此时图卷积对仅包含编号信息的输入节点进行信息聚合会引发严重的瓶颈问题,影响推荐精度,这种问题被称为图卷积信息聚合瓶颈. 为了解决此问题,研究者通常会借助外部信息辅助建模. 图卷积矩阵补全(GCMC, graph convolutional matrix completion)模型^[11]在GCN中添加文本标签信息来丰富嵌入节点嵌入特征的表示,通过GCN的链接预测实现矩阵补全任务,在一定程度上缓解了信息流瓶颈问题,但是该方法中利用的外部信息较为简单,对推荐精度的提升效果有限.

针对图卷积信息聚合时出现的瓶颈问题,提出了一种融合文本信息的轻量级图卷积网络推荐模型(T-LGCN, a lightweight graph convolutional network recommendation model incorporating text information),模型将复杂的文本评论信息集成到轻量级图卷积网络中,以丰富图卷积的节点表示. 为验证T-LGCN模型的有效性,在3个电影数据集上进行了仿真实验. 实验结果显示T-LGCN模型相比与最优对比方案,在3个数据集上的评分预测性能分别提高了3.7%、5.2%和3.8%,证明了T-LGCN模型能够有效地将节点编号信息与文本评论信息结合,缓解图卷积信息聚合的瓶颈问题,提升推荐精度.

2 模型框架

T-LGCN模型如图1所示,以各结构作用的不同可将模型分为文本特征提取层、图卷积聚合层和预测层.

2.1 文本特征提取层

文本特征提取层用于对文本评论进行特征提取,将提取出的文本评论特征作为节点嵌入的辅助信息. 它主要由2个卷积神经网络构成,分别处理基于用户的评论和基于电影项目的评论. 下面以处理基于用户评论的卷积神经网络展开介绍,其结构如图2所示. 该网络以用户评论的词嵌入矩阵作为输入,通过卷积层、最大池化层和映射层提取用户评论中潜在特征.

2.1.1 嵌入层

由于词序对于后续的文本特征提取至关重要,为了词序在后续文本建模中可以被有效利用,T-LGCN通过词嵌入技术^[11]创建词向量矩阵来保留

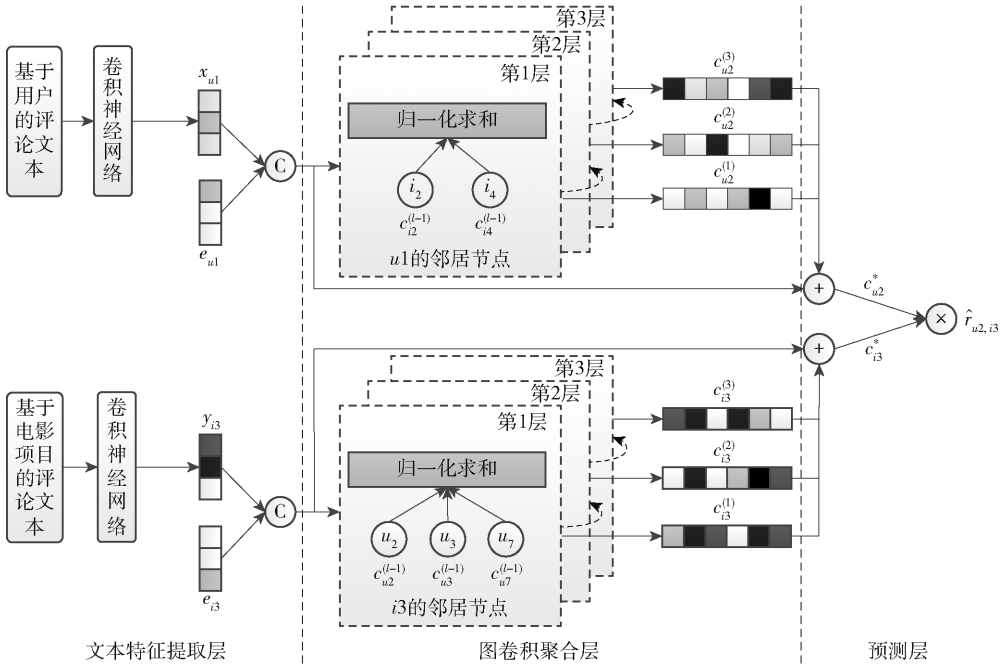


图1 T-LGCN 模型框架

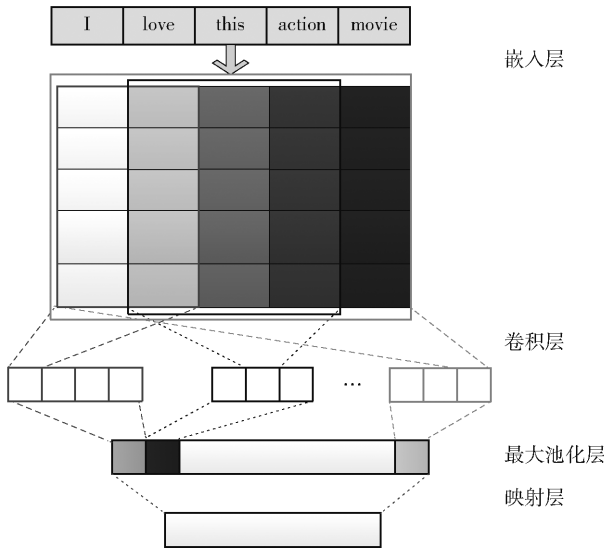


图2 文本特征提取层

词序。首先预训练一个词嵌入的单词表,模型在嵌入层中使用词嵌入函数将单词映射到 b 维向量分布^[12]。对用户 u ,将其对所有电影的评论 d_k^u 拼接成一条评论文档:

$$d^u = d_1^u \parallel d_2^u \parallel \cdots \parallel d_k^u \parallel \cdots \parallel d_K^u \quad (1)$$

其中: \parallel 为拼接操作; $k \in \{1, 2, \dots, K\}$, K 为用户 μ 评论的电影总数。

从预训练词表中查询 d^u 中对应的词向量形成词嵌入矩阵:

$$V_{1:n}^u = \varphi(d^u) =$$

$$\varphi(w_1^u) \parallel \varphi(w_2^u) \parallel \cdots \parallel \varphi(w_n^u) \quad (2)$$

其中: $V_{1:n}^u$ 为由 n 个单词所构成的用户 u 的词嵌入矩阵; w_j^u 为评论文档中第 j 个单词向量; $\varphi(\cdot)$ 为查询函数,可以从预训练词表中返回对应单词 w_j^u 的 b 维词向量。

2.1.2 卷积层

在卷积层中,使用第 h 个窗口尺寸为 $b \times t$ 的卷积核 $W_h \in \mathbf{R}^{b \times t}$ 对词矩阵进行卷积运算,其中 $h \in \{1, 2, \dots, n - t + 1\}$,获得文本特征:

$$z_h = g(W_h * V_{1:n}^u + a_h) \quad (3)$$

其中: $*$ 为卷积运算符; a_h 为 W_h 的卷积层偏置项; $g(\cdot)$ 为非线性激活函数,此处采用 ReLU 函数。

2.1.3 最大池化层

模型利用最大池化层进一步提取卷积层中的文本特征,将最大值作为该卷积核对应的特征。通过池化操作,文本特征向量的维度被减小到一个固定长度:

$$o_m = \max\{z_1, z_2, \dots, z_h\} \quad (4)$$

将不同尺寸窗口的滤波器遍历文档获取各种特征,特征组合 O 作为卷积层的输出向量:

$$O = \{o_1, o_2, \dots, o_m, \dots, o_r\} \quad (5)$$

其中 r 为卷积层中卷积核尺寸的种类数。

2.1.4 映射层

用户 u 的文本评论特征 x_u 最终由 O 映射得到

$$x_u = g(W^{fc}O + v) \quad (6)$$

其中: \mathbf{W}^{lc} 为映射层的权值矩阵, \mathbf{v} 为偏置项。

同样,通过处理基于电影项目评论的卷积神经网络可以提取出电影项目 i 的文本评论的特征 \mathbf{y}_i 。

2.2 图卷积聚合层

将文本评论信息作为属性特征,将其拼接到节点中,得到用户 u 的初始节点特征:

$$\mathbf{c}_u^{(0)} = g(\mathbf{W}^f \mathbf{e}_u + \mathbf{s}) \parallel \mathbf{x}_u \quad (7)$$

其中:输入 \mathbf{W}^f 为权值矩阵; \mathbf{s} 为图卷积层偏置项; 输入特征矩阵 $\mathbf{E} = [\mathbf{e}_1^T, \dots, \mathbf{e}_M^T]$ 为单位矩阵, 它为用户-电影项目交互图中的每个节点提供唯一的独热向量 \mathbf{e}_u , M 为节点的总个数, 即用户和电影项目的总数, \mathbf{e}_u 经映射后与 \mathbf{x}_u 连接作为初始节点特征 $\mathbf{c}_u^{(0)}$ 。

图卷积的核心思想是邻域聚合, 用户节点 u 在第 $(l+1)$ 层图卷积中的特征为

$$\mathbf{c}_u^{(l+1)} = f(\mathbf{c}_u^{(l)}, \{\mathbf{c}_k^{(l)} : k \in N_u\}) \quad (8)$$

其中: N_u 为 u 的一阶邻域; $f(\cdot)$ 为聚合函数, 用于聚合第 l 层目标节点的一阶邻域节点信息以获得目标节点的新表示。T-LGCN 的图卷积公式具体为

$$\mathbf{c}_u^{(l+1)} = \sum_{k \in N_u} \frac{1}{\sqrt{|N_u| |N_k|}} \mathbf{c}_k^{(l)} \quad (9)$$

其中: N_k 为节点 k 的一阶邻域; 系数 $\frac{1}{\sqrt{|N_u| |N_k|}}$ 为拉普拉斯标准化, 表示为一阶邻域节点对目标节点的贡献程度, 邻居节点的一阶邻域越大, 则其对目标节点的贡献度就越小。

由于图卷积层的堆叠, 高层子图节点会包含更多节点信息, 即不同层级子图节点在反映用户偏好方面有不同的贡献, 所以模型对不同层获得的特征进行加权求和得到用户 u 的最终特征表示:

$$\mathbf{c}_u^* = \sum_{l=0}^L \alpha_l \mathbf{c}_u^{(l)} \quad (10)$$

其中: α_l 为第 l 层嵌入的权重, 统一设置为 $1/(l+1)$; L 为图聚合子图的个数。

可以通过同样的操作得到电影项目 i 的特征 \mathbf{c}_i^* 。

在式(9)中, 不同于其他图卷积模型同时聚合

目标节点及其邻域的方式^[13], T-LGCN 的图卷积中只聚合邻域节点, 这是由于式(10)将不同层的特征与加权和相结合, 可以实现图卷积自连接的效果。

2.3 预测层

模型的评分预测以用户和电影项目最终特征输出的内积形式计算:

$$\hat{r}_{u,i} = \mathbf{c}_u^* \mathbf{c}_i^* \quad (11)$$

损失函数为

$$L = \sum_{(u,i) \in T} w_{u,i} (r_{u,i} - \hat{r}_{u,i})^2 + \lambda \|\Theta\|_2^2 = \sum_{(u,i) \in T} w_{u,i} (r_{u,i} - \mathbf{c}_u^* \mathbf{c}_i^*)^2 + \lambda \|\Theta\|_2^2 \quad (12)$$

其中: T 为训练集; $r_{u,i}$ 为真实评分, $\hat{r}_{u,i}$ 为预测评分; λ 和 Θ 分别为正则化系数和模型参数; $w_{u,i}$ 为置信值, 用于为每条评分项赋予不同的信任权重, 因为极端评分比中等评分更能表达用户的偏好, 置信值计算公式为

$$w_{u,i} = 1 + \theta \cdot \left| \frac{r_{\max} - r_{u,i}}{2} \right| \quad (13)$$

其中: 置信度 θ 是控制置信等级大小的超参数, r_{\max} 为数据集中评分的最大值。这种置信机制可以确保极端评分获得更高的权重, 对评分条目的可信度进行区分。

3 实验评估

实验的操作系统环境为 Ubuntu16.04, 程序使用 Python3.7 编写, 在 Nvidia 1080Ti GPU 上训练模型。

3.1 数据集

实验共使用 3 个数据集, 分别缩写为 ML-1m 数据集^[14]、ML-10m 数据集^[14]和 Douban 数据集^[15]。其中, ML-1m 数据集和 ML-10m 数据集均是 MovieLens 电影评分数据集系列中的一种, 包含用户的评分, 电影的属性信息和用户属性信息; Douban 数据集是经过 Monti 等^[15]预处理得到的, 以图结构的形式包含用户和电影项目的节点信息。3 个数据集相关信息如表 1 所示。

表 1 数据集相关信息

数据集	用户数	项目数	评分数	评分标签	评分矩阵稀疏度/%
ML-1m	6 040	3 883	1 000 209	1, ..., 5	95.53
ML-10m	71 567	10 681	10 000 054	0.5, 1, ..., 5	98.69
Douban	2 999	3 000	136 891	1, ..., 5	98.48

由表可知,3 个数据集具有高度稀疏性. 由于数据集本身并不包含文本评论信息,数据准备阶段利用 Web 爬虫技术爬取数据集的用户评论文档,并进行数据预处理工作:① 将原始文本最大长度设置为 500;② 删除停止词;③ 选择前 6 000 个不同的单词作为词汇表,并从原始文档中删除其他单词. 针对每个数据集,随机选取 80% 的数据作为训练集,其余 20% 的数据作为测试集.

3.2 评估指标

实验选择均方根误差 (RMSE, root mean squared error) 对各模型的评分预测的准确度进行评估:

$$S = \sqrt{\frac{\sum_{u=1}^U \sum_{i=1}^I (\hat{r}_{u,i} - r_{u,i})^2}{|N|}} \quad (14)$$

其中: $|N|$ 为测试集中评分的数量; $\hat{r}_{u,i}$ 与 $r_{u,i}$ 分别为预测评分与真实评分, U 和 I 分别为测试集中的用户数量和电影数量. 评分预测的 RMSE 越小,准确度越高.

3.3 实验与分析

3.3.1 推荐模型对比实验

将所提出的 T-LGCN 模型与偏置矩阵分解 (BiasMF, biased matrix factorization) 模型^[4]、多图递归神经网络 (sRGCNN, geometric matrix completion with recurrent multi-graph neural networks) 模型^[15] 和 GCMC 模型^[11] 在 3 个数据集上进行性能对比. 其中 BiasMF 模型通过分解评分矩阵中的已知评分得到用户和项目嵌入,利用嵌入特征的内积预测评分. sRGCNN 模型利用深度学习的图形来实现矩阵补全任务,在用户和项目的最近邻图上使用图卷积,并使用递归神经网络以迭代的方式学习用户和项目的特征嵌入. GCMC 模型是由一个图自动编码器和双线性译码器组成. 模型在嵌入层可以融合简单的文本标签作为辅助信息以缓解图卷积信息聚合时出现的瓶颈问题.

表 2 所示为 T-LGCN 模型与上述 3 种模型的性能对比. 与传统模型 (BiasMF 模型) 相比,基于 GCN 的推荐模型 (GCMC 模型和 T-LGCN 模型) 可以实现较大的改进. T-LGCN 模型相比与表现突出的 GCMC 模型而言,在 ML-1m 数据集、ML-10m 数据集和 Douban 数据集上性能分别提高了 3.7%、5.2% 和 3.8%. 由上可知,将复杂的文本评论信息通过图嵌入的方法集成到轻量级图卷积框架中能够更好地分析用户和电影项目的特征,提升推荐质量.

表 2 T-LGCN 模型与其他模型在 3 个数据集上的 RMSE 对比

模型	ML-1m	ML-10m	Douban
BiasMF	0.847	0.803	—
sRGCNN	0.857	0.806	0.801
GCMC	0.832	0.771	0.733
T-LGCN	0.801	0.731	0.705

3.3.2 用户冷启动实验

为了研究图卷积推荐模型结合文本评论信息对缓解冷启动问题的有效性,在包含一定数量冷启动用户 (评分数量少的用户) 的数据集中进行实验. 首先需要对 ML-1m 数据集进行预处理:① 随机选取 P 名用户作为冷启动用户集 I_p ;② 对 I_p 中的每个用户随机保留 Q 条评分,其他评分删除,作为冷启动用户使用. 预处理后的 ML-1m 数据集简称为 PreML-1m,它包含 P 名冷启动用户和 $6040 - P$ 名正常用户,其中每个冷启动用户只含有 Q 条评分.

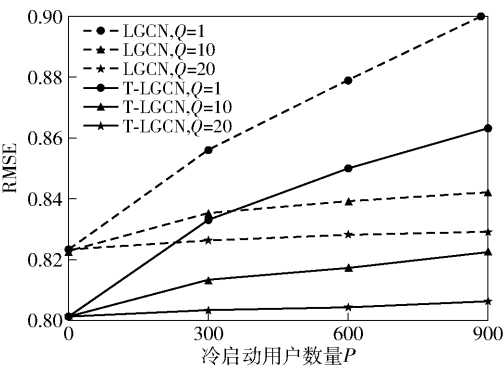


图 3 添加文本评论信息前后的 RMSE 对比

不同冷启动用户数量和评分数据下,添加文本评论信息前后的 RMSE 对比如图 3 所示. $Q \in \{1, 10, 20\}$ 为每个冷启动用户的评分数量,图中 LGCN 模型是指 T-LGCN 在图卷积聚合层未融合文本特征的模型. 由于 PreML-1m 中总用户数量不变,而随着冷启动用户数量上升,正常用户的数量下降,导致评分条目数量减少,因而使模型的训练效果下降,预测时的 RMSE 上升. 由图 3 可知,在 P 与 Q 相同的条件下, T-LGCN 模型的准确度均优于 LGCN 模型,证明了图卷积模型结合文本评论信息有助于推荐性能的提升. 在冷启动用户数量为 900 且只保留 1 个评分的条件下, T-LGCN 的预测误差远远小于 LGCN,这说明 T-LGCN 模型更适用于处理具有较高稀疏性的数据集.

3.3.3 图卷积层组合实验

不同图卷积层数量下,使用层组合策略前后的 RMSE 对比如图 4 所示. 其中未使用层组合的模型只使用最后一层图卷积来预测评分. 实验选用 ML-10m 数据集. 由图 4 可知,随着图卷积层数量的变化,在性能上使用层组合的 T-LGCN 模型始终优于未使用层组合的模型,证明使用层组合能有效地整合不同层的特征,提升推荐质量.

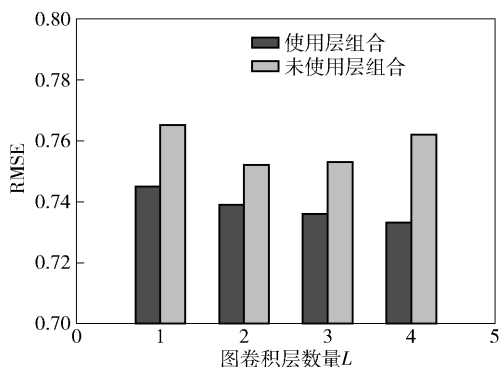


图4 使用层组合前后的 RMSE 对比

4 结束语

为解决图卷积信息聚合时出现的瓶颈问题,提出了 T-LGCN 模型提升推荐质量. 实验结果表明, T-LGCN 模型在 3 个不同数据集上的表现均优于对比模型. 下一步工作可以在 T-LGCN 模型基础上继续丰富辅助信息,增加其他模态的信息(如社交网络、图像、音轨等)来辅助推荐,进一步提升推荐性能.

参考文献:

- [1] Cui Zhihua, Xu Xianghua, Fei Xue, et al. Personalized recommendation system based on collaborative filtering for IoT scenarios[J]. IEEE Transactions on Services Computing, 2020, 13(4): 685-695.
- [2] Pazzani M J, Billsus D. Content-based recommendation systems[M]. Heidelberg: Lecture Notes in Computer Science, 2007: 15-16.
- [3] Cheng Zhiyong, Ding Ying, Zhu Lei, et al. Aspectware latent factor model: rating prediction with ratings and reviews[C]// International Conference on World Wide Web Conference. Lyon: WWW, 2018: 639-648.
- [4] Koren Y, Bell R, Volinsky C. Matrix factorization techniques for recommender systems[J]. Computer, 2009, 42(8): 30-37.
- [5] Yue Xiang, Wang Zhen, Huang Jingong, et al. Graph embedding on biomedical networks: methods, applications and evaluations[J]. Bioinformatics, 2020, 36(4): 1241-1251.
- [6] Defferrard M, Bresson X, Vandergheynst P. Convolutional neural networks on graphs with fast localized spectral filtering[J]. Advances in Neural Information Processing Systems, 2016, 29(1): 3844-3852.
- [7] Zhang Chuxu, Song Dongjin, Huang Chao, et al. Heterogeneous graph neural network[C]// 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Anchorage: ACM, 2019: 793-803.
- [8] Wang Xiang, He Xiangnan, Wang Meng, et al. Neural graph collaborative filtering[C]// ACM SIGIR Conference on Research and Development in Information Retrieval. Paris: ACM, 2019: 165-174.
- [9] Wu F, Souza A, Zhang Tianyi, et al. Simplifying graph convolutional networks[C]// ICML 2019: Proceedings of 2019 International Conference on Machine Learning. California: ICML, 2019: 6861-6871.
- [10] He Xiangnan, Deng Kuan, Wang Xiang, et al. LightGCN: simplifying and powering graph convolution network for recommendation[C]// ACM SIGIR Conference on Research and Development in Information Retrieval. Salamanca: Computer Science, 2020: 639-648.
- [11] Berg R, Kipf T N, Welling M. Graph convolutional matrix completion[C]// KDD'18 Deep Learning Day. London: ACM, 2018: 161-168.
- [12] Mikolov T, Sutskever I, Chen K, et al. Distributed representations of words and phrases and their compositionality[C]// Advances in Neural Information Processing Systems. Nevada: NIPS, 2013: 3111-3119.
- [13] Zhao Cheng, Li Chenliang, Fu Cong. Cross-domain recommendation via preference propagation graphnet[C]// 28th ACM International Conference on Information and Knowledge Management. Beijing: ACM, 2019: 2165-2168.
- [14] Harper F M, Konstan J A. The movielens datasets: history and context[J]. ACM Transactions on Interactive Intelligent Systems (THIS), 2015, 5(4): 1-19.
- [15] Monti F, Bronstein M, Bresson X. Geometric matrix completion with recurrent multi-graph neural networks[C]// Advances in Neural Information Processing Systems. Long Beach: NIPS, 2017: 3697-3707.