

基于图注意力的神经协同过滤社会推荐算法

章琪, 于双元, 尹鸿峰, 徐保民

引用本文

章琪, 于双元, 尹鸿峰, 徐保民. [基于图注意力的神经协同过滤社会推荐算法](#)[J]. 计算机科学, 2023, 50(2): 115-122.

ZHANG Qi, YU Shuangyuan, YIN Hongfeng, XU Baomin. [Neural Collaborative Filtering for Social Recommendation Algorithm Based on Graph Attention](#) [J]. Computer Science, 2023, 50(2): 115-122.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[利用异构图神经网络实现情绪-原因对的有效抽取](#)

Utilizing Heterogeneous Graph Neural Network to Extract Emotion-Cause Pairs Effectively
计算机科学, 2023, 50(1): 205-212. <https://doi.org/10.11896/jsjcx.211100265>

[基于影响力剪枝的图神经网络快速计算图精简](#)

Fast Computation Graph Simplification via Influence-based Pruning for Graph Neural Network
计算机科学, 2023, 50(1): 52-58. <https://doi.org/10.11896/jsjcx.220900032>

[结合全局信息的深度图解耦协同过滤](#)

Deep Disentangled Collaborative Filtering with Graph Global Information
计算机科学, 2023, 50(1): 41-51. <https://doi.org/10.11896/jsjcx.220900255>

[基于图神经网络和依存句法分析的文本分类](#)

Text Classification Based on Graph Neural Networks and Dependency Parsing
计算机科学, 2022, 49(12): 293-300. <https://doi.org/10.11896/jsjcx.220300195>

[融合多层次视觉信息的人物交互动作识别](#)

Human-Object Interaction Recognition Integrating Multi-level Visual Features
计算机科学, 2022, 49(11A): 220700012-8. <https://doi.org/10.11896/jsjcx.220700012>

基于图注意力的神经协同过滤社会推荐算法

章琪¹ 于双元¹ 尹鸿峰² 徐保民¹

1 北京交通大学计算机与信息技术学院 北京 100044

2 沧州交通大学计算机与信息技术学院 河北 沧州 061199

(zqi_97@126.com)

摘要 互联网技术的发展使得信息过载问题日趋严重,为了解决传统推荐技术的数据稀疏和冷启动问题,社会推荐逐渐成为近年来的研究热点。图神经网络(GNNs)作为一种能够自然整合节点信息和拓扑结构的网络,为改进社会推荐提供了巨大的潜力。但基于图神经网络的社会推荐还存在许多挑战,例如,如何从用户项目交互图和社交网络图中学习准确的用户和项目的潜在因子表示;简单映射用户和项目的固有属性来获取嵌入,但用户项目交互的关键协作信号未被学习。为了学习更准确的潜在因子表示,捕获关键的协作信号,提升推荐系统的性能,提出了基于图注意力的神经协同过滤社会推荐模型(AGNN-SR)。该模型基于用户项目交互图和社交网络图,通过多头注意力机制多角度地学习用户和项目的潜在因子;此外,图神经网络利用高阶连通性递归地在图上传播嵌入信息,显式编码协作信号,探索用户和项目之间的深层复杂的交互关系。最后,在3个真实数据集上验证了AGNN-SR模型的有效性。

关键词: 社会推荐;图神经网络;多头注意力;神经协同过滤

中图法分类号 TP301

Neural Collaborative Filtering for Social Recommendation Algorithm Based on Graph Attention

ZHANG Qi¹, YU Shuangyuan¹, YIN Hongfeng² and XU Baomin¹

1 School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China

2 School of Computer and Information Technology, Cangzhou Jiaotong University, Cangzhou, Hebei 061199, China

Abstract The development of Internet technology has made the problem of information overload more and more serious. In order to solve the problems of data sparse and cold start of traditional recommendation technology, social recommendation has gradually become a research hotspot in recent years. As a network, graph neural networks(GNNs)can naturally integrate node information and topology,offer great potential for improving social recommendation. But there are still many challenges for social recommendation based on graph neural network. For example,how to learn accurate latent factor representations of users and items from user-item interaction graphs and social network graphs;Simply mapping of inherent properties of users and items to obtain embeddings,but key collaborative signals of user-item interactions are not learned. In order to learn more accurate latent factor representations,capture key collaborative signals,and improve the performance of recommender systems,a graph attention-based neural collaborative filtering social recommendation model(AGNN-SR) is proposed. The model is based on user-item interaction graphs and social network graphs,and learns latent factors of users and items from multiple perspectives through a multi-head attention mechanism. In addition,graph neural networks utilize higher-order connectivity to recursively propagate embedding information on the graph,explicitly encoding collaborative signaling to explore deep and complex interactions between users and items. Finally,the effectiveness of the AGNN-SR model is verified on three real datasets.

Keywords Social recommendation,Graph neural network,Multi-head attention,Neural collaborative filtering

1 引言

随着互联网与信息技术的快速发展,信息过载问题逐渐变得严峻。推荐系统应运而生,其主要目的是根据用户的历史交互数据为用户提供个性化的推荐服务,从而减少用户检索过滤信息的成本。推荐建议一般是基于用户偏好、项目特征、用户与项目的历史交互以及一些附加信息得出的,其核心

是分析用户的偏好模式以提出潜在目标。推荐系统模型主要分为协同过滤、基于内容和混合推荐系统。其中,协同过滤(Collaborative Filtering,CF)^[1]作为最成功的推荐技术之一,结合不同时期的技术拥有许多变形研究^[2-3],其核心数学问题是矩阵分解。协同过滤在许多应用方面取得了效果,但仍存在着数据稀疏和冷启动问题,这极大地限制了推荐系统性能的提升^[4]。随着社交网络的兴起,人们引入用户的社交关系

到稿日期:2021-12-01 返修日期:2022-04-13

基金项目:沧州市重点研发计划(204102013)

This work was supported by the Key Plan of Research and Development of Cangzhou(204102013).

通信作者:尹鸿峰(hfyin@bjtuhbxy.edu.cn)

作为辅助信息,基于用户会受到社交好友兴趣的影响这一事实,从而有效地缓解了推荐系统存在的数据稀疏和冷启动问题^[5]。

作为全球范围内的研究热点,社交推荐利用用户的社交网络数据有效地解决了数据稀缺性问题,进而提高了推荐效果。Ma 等先后提出的 SoRec^[6]、SoReg^[7] 等模型,均是基于概率矩阵因子化的因子分析方法,通过同时使用用户的社交网络信息和评级记录来解决数据稀疏和预测准确性差的问题,同时也规范化了社会推荐矩阵因子化框架。之后,研究人员基于矩阵因子法陆续提出了一系列的社会推荐系统变形研究,如 Jamali 等的 SocialMF 模型^[8]、Guo 等的 TrustSVD 模型^[9]以及 Yang 等的 TrustMF 模型^[10]。其中,TrustSVD 模型的推荐性能最好,该模型是基于 SVD 的,进一步将受信任用户的显式和隐式反馈影响纳入对活跃用户的项目预测。最近,图卷积网络(Graph Convolutional Network, GCN)利用图结构以及节点特征在图中的信息扩散过程进行建模,显示出了良好的效果。越来越多的研究人员将目光集中在 GCN 上,试图通过引入 GCN,借用其优势来捕捉用户的偏好是如何受到社交网络中社交扩散过程的影响,从而达到提升精度和性能的目的。Wu 等提出的 GCMC 模型^[11]就是一种有效的基于图卷积神经网络的社会推荐模型。一般来说,社交网络中的用户不仅受到其朋友(一阶邻居)的影响,而且还受到朋友的朋友(高阶邻居)的影响。为了捕捉高阶社会关系,Fan 等结合 GCMC 和社交网络提出了 GraphRec 模型^[12];Guo 等提出了一种基于深度图神经网络的社交推荐框架 GNN-SoR 模型^[13];Luo 等提出了 ASR 模型^[14],通过设计 Rec-conv 图神经网络层来提取社会因素;Song 等^[15]则将重点放在社交网络局部内隐影响和广播给用户的项目的全局内隐影响上,通过分别对两种隐式影响进行建模,改进了最新的基于 GNN 的社会推荐方法。上述方法通过堆叠多个 GNN 层获取高阶社交关系,均取得了良好的实验结果。但是,堆叠过多的 GNN 层可能会受到过度平滑问题的影响,进而导致性能显著下降。MHCN 模型^[16]提出用超边缘对高阶社会关系进行建模,它可以连接两个以上的节点,并以自然的方式对高阶关系进行建模。HOSR 模型^[17]使用 GCN 聚合邻居的信息,来捕获社会图中的高阶关系。另外,一些研究^[18-19]引入了更多的附加信息,例如 TGRec 模型^[20]将用户行为的时间信息引入社交推荐中,以进一步加强社会推荐。

可学习的 CF 主要有两个关键部分:1)嵌入,将用户与项目转换为矢量表示;2)交互建模,根据嵌入重构交互历史。但是,现在大部分模型仅使用描述性信息(如 ID 和属性)构建嵌入函数,例如矩阵分解(Matrix Factorization, MF)直接将用户或项目 ID 嵌入为向量,并对用户项目的交互进行内积建模^[21]。但由于上述方法的嵌入函数缺少对关键协作信号的显式编码,而关键协作信号隐藏在用户项目交互中,以揭示用户(或项目)之间的行为相似性。因此,当嵌入不足以捕获 CF 时,则必须依靠交互函数来弥补嵌入的不足。但做好这份工作并非易事,在实际应用中,交互的规模很容易达到数百万甚至更大,因此很难提取所需的协作信号。此时,可以通过探索高阶交互信息,利用高阶连通性递归地在图上传播潜在特征信息,产生更优的嵌入表示,从而应对协作信号提取这一挑战。

综上所述,本文提出了一个新模型:基于图注意力的神经协同过滤社会推荐模型(Attention Graph Neural Network for Social Recommendation, AGNN-SR)。该模型是基于用户项目交互图和社交网络图两个不同的图,分别从中学习建模用户的潜在因子表示和项目的潜在因子表示。其中,用户不仅受到具有评分的连接项目的影响,还受到了与其有关联的其他用户偏好的影响。而为了从多角度融合用户的特征信息,这里引入了多头注意力机制,在每个图上传递消息,从用户项目交互图学习项目潜在因子表示,最终的用户潜在因子表示则通过合并两个图的嵌入向量得到。最后,将用户和项目的潜在因子输入神经协同过滤推荐模块中,在图神经网络上利用图的高阶连通性递归地传播嵌入信息,细化节点的嵌入表示,捕获用户和项目之间的深层复杂交互,从而进行预测推荐。

2 相关工作

2.1 社会推荐

Web2.0 时代,随着 Facebook, Twitter 等在线社交网络的快速发展,传统推荐系统整合用户的社交信息来提高推荐精度的相关研究也越来越多,一些经典研究的实验结果^[22-23]也证明了用户的行为偏好受到与其相关联的其他用户的影响。由于社会推荐系统考虑的用户行为偏好受到的社会影响对学习用户表示很有帮助,使得推荐领域的研究更多地从社会特性出发,达到提升推荐准确性的效果。协同过滤广泛用于构建推荐系统,现有的大多数社会推荐系统都基于协同过滤技术^[24-25]。其中,图的结构特性能最直观、最自然地描述社交网络类的数据,因此现阶段基于图神经网络模型的社交推荐方法被广泛研究。社交网络中的用户和项目被表示为图上的一个个不同的节点,用户间的社交程度以及用户与项目之间的交互使用加权边来表示。而推荐系统的主要目的就是用户和项目的评分图和用户社交图分别获得用户和项目的特征表示,来进行相应的推荐。TidalTrust 模型^[26-27]中,用户对某个项目的喜好值是用户熟悉的所有用户的平均值,该方法是基于信任加权平均值的基准策略。MoleTrust 模型^[28]是基于 TidalTrust 模型进行改进的, MoleTrust 模型的最大路径阈值是预先设置的,整个搜索过程是静态的。此外,不同于 TidalTrust 模型的加权平均法, MoleTrust 模型在预测评分时使用了协同过滤技术。与经典的协同过滤推荐相比, MoleTrust 模型对冷启动用户具有更好的预测效果。常见的基于模型的 CF 技术是矩阵分解。矩阵分解技术可以将用户项目得分矩阵分解为两个或多个低维矩阵项,实现维度的规范化,并利用低维空间中的数据研究高维数据的属性。在社会化推荐系统中,矩阵分解方法分为两类。1)基于矩阵分解的因子学习模型,该模型的目标是将用户项目评分矩阵分解为用户因子矩阵和项目因子矩阵的乘法形式。其优点是能够集成用户和项目的多个社会学属性,并缓解数据稀疏和冷启动问题;其缺点是需要研究的参数较多,推荐效果仅限于历史数据。2)社交化矩阵分解模型,在社交推荐系统中,矩阵分解方法关注用户的社交网络信息对用户潜在特征向量的影响。社交化矩阵分解将用户的各种社交网络的信息集成到矩阵的最优分解过程中,以提取更好的潜在特征向量^[29]。社交化矩阵分解很大程度上缓解了数据稀疏和用户冷启动问题,并且可以

自动生成推荐结果,这是该方法的优点;但社会化矩阵分解也有缺点,该技术存在项目冷启动问题,推荐效果差。

2.2 图神经网络

图表示学习是近年来的一个研究热点,其主要目的是学习节点在图结构数据上的准确嵌入。GraphSAGE^[30]将图卷积网络扩展到归纳学习任务,并推广了未知节点。图形注意力网络(GAT)^[31]将注意机制纳入消息传递步骤,并根据自我注意策略计算每个节点的表示。越来越多的研究者利用图表示学习的优势来提高推荐系统的性能。GCMC^[32]将矩阵完成视为图形上的连接预测问题,并使用图形自动编码器将交互数据与边信息相结合。GraphRec提供了一种原则性的方法来联合捕获用户项目图中的意见交互,并将注意机制引入模型中。

3 AGNN-SR 模型

本节将介绍本文提出的 AGNN-SR 模型的整体框架,以及模型各个组件部分的实现原理,最后讨论了如何学习模型参数。

3.1 模型框架

AGNN-SR 模型的整体框架如图 1 所示。该模型由用户建模、项目建模和评级预测 3 部分组成。第一部分是用户建模,其目的是学习获取用户的潜在因子表示。不同于传统的推荐算法,社会推荐模型相比传统推荐方法新增了用户的社交网络图,即包括用户项目交互图和社交网络图两个不同的图,让算法能从更多不同的角度去多方位地学习用户潜在因子。第二部分是项目建模,这里引入了用户聚合,即在项目建模中汇总用户的意见,目的是从用户项目交互图中学习项目相关的潜在因子表示。第三部分则是评级预测,通过前两部分建模学习得到的用户和项目潜在因子来进行预测训练,学习模型参数。接下来将详细说明每一个模型组件。

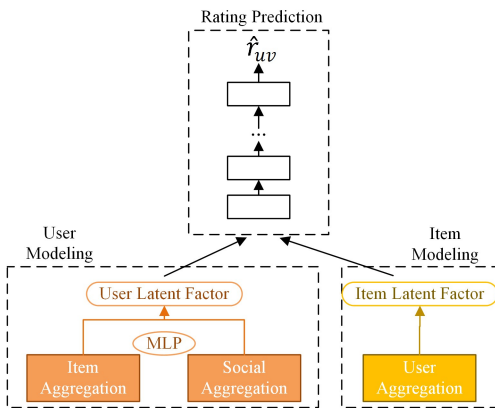


图 1 AGNN-SR 的整体框架

Fig. 1 Architecture of AGNN-SR

3.2 潜在因子学习

3.2.1 用户潜在因子

用户建模旨在了解用户的潜在因子表示 h_i 。为了将用户项目交互图和社交网络图更有效地结合起来,这里使用两种聚合来分别学习两个图上的嵌入向量表示。一个为项目聚合,是从用户项目交互图中学习项目空间的用户潜在因素,记为 h_i^U ,计算过程如式(1)所示;另一个为社交聚合,是从用户

社交图中学习社会空间的用户潜在因素,记为 h_i^S ,计算过程如式(2)所示。最后将这两个用户潜在因素利用多层感知器(Multi-Layer Perception, MLP)进行如式(3)所示的整合计算,得到最终的用户潜在因子表示 h_i 。数学上相关的聚合函数表示如下:

$$h_i^U = \prod_{k=1}^K \sigma \left(\sum_{j \in C(i)} \alpha_{ij}^k \mathbf{W}_u^k q_{jr} \right) \quad (1)$$

$$h_i^S = \prod_{k=1}^K \sigma \left(\sum_{o \in N(i)} \alpha_{io}^k \mathbf{W}_s^k u_o \right) \quad (2)$$

$$h_i = g_{us} (h_i^U \oplus h_i^S) \quad (3)$$

其中,用户对项目的评分或评级 r 表达了相应的偏好。为了更准确地建模用户潜在因子,将每个意见 r 表示为更加稠密的向量表示 $e_r \in \mathbb{R}^d$, d 为嵌入向量的长度,将初始项目嵌入向量 v_j 和评级 e_r 通过多层感知器(MLP)结合,计算可得:

$$q_{jr} = g_u (v_j \oplus e_r) \quad (4)$$

图 2 给出了结合多头注意力与自注意力机制的网络聚合过程。在实验中用两层的多头注意力神经网络来参数化 α_{ij} , 归一化定义如下:

$$\alpha_{ij} = \text{softmax}(e_{ij}) = \frac{\exp(e_{ij})}{\sum_{k \in C(i)} \exp(e_{ik})} \quad (5)$$

$$e_{ij} = \mathbf{W}_2^T \cdot \sigma(\mathbf{W}_1 \cdot (\mathbf{W}_u u_i \oplus \mathbf{W}_s q_{jr}) + b_1) + b_2 \quad (6)$$

其中, α_{ij} 表示交互空间 $C(i)$ 内用户与不同项目交互时的注意力权重; e_{ij} 为注意力系数,体现了 q_{jr} 对用户产生的影响的重要程度,这里使用的是自我关注(Self-Attention)方法。 $\mathbf{W}_u \in \mathbb{R}^{d' \times d}$ 表示线性变换,其中 d' 为输出向量的长度。

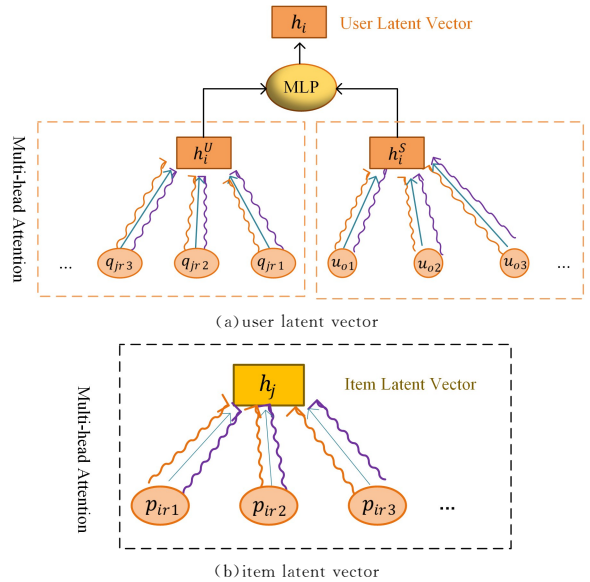


图 2 用户和项目潜在因子

Fig. 2 User & item latent factor

3.2.2 项目潜在因子

图 2 中,在项目建模部分,将从用户项目交互图中学习获取项目潜在因子表示 h_j ,整个过程类似于项目空间中学习用户潜在因子的方法。将与项目产生交互的用户 u_i 定义为空间 $B(j)$,用户聚合过程如下:

$$h_j = \prod_{k=1}^K \sigma \left(\sum_{i \in B(j)} \alpha_{ij}^k \mathbf{W}_i^k p_{ir} \right) \quad (7)$$

$$p_{ir} = g_u (u_i \oplus e_r) \quad (8)$$

$$\alpha_{ij} = \text{softmax}(e_{ij}) = \frac{\exp(e_{ij})}{\sum_{t \in B(j)} \exp(e_{it})} \quad (9)$$

$$e_{ij} = \mathbf{W}_4^T \cdot \sigma(\mathbf{W}_3 \cdot (\mathbf{W}_v v_j \oplus \mathbf{W}_u p_{ir}) + b_3) + b_4 \quad (10)$$

经过以上模块组件,现已获得最终的用户潜在因子和项目潜在因子,接下来将介绍评级预测模块。

3.3 神经协同过滤推荐

受到现有神经协同过滤方法的启发,本文设计了一种嵌入传播层,在嵌入空间构建信息流,通过注入高阶信息,堆叠多个嵌入传播层,以高阶连接捕获协作信号。这样可以更有效且准确地对用户与项目间的内部交互行为进行建模,生成更优的嵌入表示,从而提高预测精度。整个网络的结构如图3所示,主要分为输入层、嵌入传播层、预测层3部分。

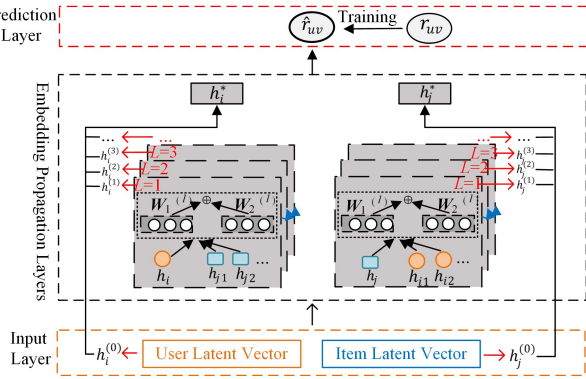


图3 神经协同过滤推荐

Fig. 3 Neural collaborative filtering recommendation

3.3.1 输入层

将之前学习得到的用户潜在因子和项目潜在因子作为初始值,输入模块中。

3.3.2 嵌入传播层

嵌入传播层是该模块的核心部分,以GNN构建消息传递架构,捕获CF信号,并优化用户和项目的嵌入表示。一般来说,与用户交互的项目提供了用户偏好的直接证据;同理可得,对某个项目进行评分或评级的用户可以被视为该项目的特征,并用于度量项目间的协作相似性。在此基础上,定义消息传递和消息聚合两个操作在连接的用户和项目之间执行嵌入传播。

(1) 消息传递

对于连接的用户和项目,有如下定义:

$$m_{u \leftarrow v}^{(l)} = \frac{1}{\sqrt{|\mathcal{N}_u| |\mathcal{N}_v|}} \cdot (\mathbf{W}_1^{(l)} h_j^{(l-1)} \mathbf{W}_2^{(l)} (h_j^{(l-1)} \odot h_i^{(l-1)})) \quad (11)$$

$$m_{u \leftarrow u}^{(l)} = \mathbf{W}_1^{(l)} h_i^{(l-1)} \quad (12)$$

其中, l 是嵌入传播层的层数, $m_{u \leftarrow v}^{(l)}$, $m_{u \leftarrow u}^{(l)}$ 表示消息的传递($u \leftarrow u$ 表示用户到用户, $u \leftarrow v$ 表示项目到用户); $\mathbf{W}_1^{(l)}$, $\mathbf{W}_2^{(l)}$ 是提取传播有用信息的可训练权重矩阵; $\frac{1}{\sqrt{|\mathcal{N}_u| |\mathcal{N}_v|}}$ 为图的拉

普拉斯范数,可看作一个折扣因子,传播的消息应随路径长度的增加而衰减; \mathcal{N}_u 和 \mathcal{N}_v 表示用户 u 和项目 v 的第一邻居,反映了历史项目对用户偏好的贡献程度。与传统图卷积网络只考虑项目 h_j 的贡献不同,这里还将用户 h_i 和项目 h_j 之间的交互协作编码通过消息进行传递。

(2) 消息聚合

在这一阶段,聚集用户 u 邻域传播的消息,以改进用户 u 的嵌入表示。聚合函数的定义如下:

$$h_i^{(l)} = \text{LeakyReLU}(m_{u \leftarrow u}^{(l)} + \sum_{v \in \mathcal{N}_u} m_{u \leftarrow v}^{(l)}) \quad (13)$$

其中,LeakyReLU^[33]的激活功能允许消息对正信号和负信号进行编码。通过叠加更多的嵌入传播层来探索高阶连通性信息,将获取到的编码协作信号用于计算用户对项目的预测评分至关重要。

3.3.3 预测层

在经过 l 层嵌入传播之后,获得了用户 u 的多个表示,即 $\{h_i^1, \dots, h_i^l\}$,代表了每一层不同连接传播的信息,对用户的偏好影响程度也不尽相同。对项目也进行上述类似的操作。将每一层的嵌入向量连接起来,构成用户和项目的最终嵌入表示。

$$h_i^* = h_i^{(0)} \parallel \dots \parallel h_i^{(l)} \quad (14)$$

$$h_j^* = h_j^{(0)} \parallel \dots \parallel h_j^{(l)}$$

最后,计算内积得到预测评级:

$$\hat{r}_{ij} = h_i^* \top h_j^* \quad (15)$$

3.4 模型训练

为了估计AGNN-SR模型参数,同时考虑到模型训练更多的是关注推荐系统中的评级预测任务,合理选取了较为通用的目标函数:

$$\text{Loss} = \frac{1}{|\mathcal{O}|} \sum_{(i,j) \in \mathcal{O}} (\hat{r}_{ij} - r_{ij})^2 \quad (16)$$

其中, r_{ij} 是用户 i 在项目 j 上的真实评级, \hat{r}_{ij} 是模型聚合学习到的预测评级。

4 实验

4.1 实验设置

4.1.1 数据集

本文选择了3个具有代表性的公共数据集:Ciao, Epinions和FilmTrust。这3个数据集分别来自不同的社交网络,且都包含了用户与项目的交互信息和用户的社交好友信息。其中,Ciao与Epinions的评级范围是 $\{1, 2, 3, 4, 5\}$,FilmTrust的评级范围是 $\{0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0\}$;根据不同的评级嵌入向量来随机初始化意见嵌入。这3个数据集的统计数据如表1所列。

表1 Ciao, Epinions和FilmTrust数据集统计

Dataset	Ciao	Epinions	FilmTrust
# of Users	7317	40163	1508
# of Items	104975	139738	2071
# of Ratings	283319	664824	35497
# of Density(Ratings)/%	0.0368	0.0118	1.1400
# of Social Connections	111781	487183	1853
# of Density(Social Relations)/%	0.2087	0.0290	0.4200

4.1.2 评价指标

选择平均绝对误差(MAE)和均方根误差(RMSE)这两个经典指标来对模型进行性能评估。

$$\text{MAE} = \frac{1}{|T|} \sum_{(u,v) \in T} |\hat{r}_{ij} - r_{ij}| \quad (17)$$

$$\text{RMSE} = \sqrt{\frac{1}{|T|} \sum_{(u,v) \in T} (\hat{r}_{ij} - r_{ij})^2} \quad (18)$$

其中,MAE和RMSE的值越小,表示模型的推荐预测准确率

越高;且 MAE 或 RMSE 数值上的微小改变也可能会对前几条建议的质量产生重大影响。

4.1.3 参数相关

将数据集按照训练集、验证集、测试集进行划分,分别对比了训练集为 60%、验证集为 20%、测试集为 20%和训练集为 80%、验证集为 10%、测试集为 10%两种不同情况下的模型性能,根据效果最终选择使用 80%作为训练集来学习模型参数。对于训练批次大小,在[32,64,128,512]范围内进行实验。嵌入大小在范围[8,16,32,64,128,256]内进行选择。经过实际结果的对比,最后将训练批次大小设置为 128,将嵌入大小设置为 64。此外,将学习率设定为 0.001;激活函数为 ReLU。对于整个神经网络,采用高斯分布随机初始化模型参数,其中的平均值和标准偏差分别为 0 和 0.1;对于模型中使用到的多层感知器(Multi-Layer Perception,MLP)均使用隐藏层。在 Pytorch2 基础上来具体实现本文模型。

4.1.4 Baselines

为了更好地体现本文模型的优势以及更客观地进行性能评估,选取了若干传统社会推荐模型和基于深度神经网络的推荐模型,分别在上述 3 个数据集上进行实验,将实验结果与 AGNN-SR 模型进行对比。其中,选择的 Baselines 模型的参数保留了其在文献[9]中的设定,以获得模型的最佳性能。下面将对每个模型进行相应的介绍。

(1)PMF^[34]。PMF 是现代推荐系统的基础算法之一,采用低维向量模型,提出了著名的概率矩阵分解的方法来解决传统协同过滤的限制。概率矩阵分解的思想仅使用与用户相关的评分系数,将用户的偏好建模成一个一系列向量的线性组合。

(2)SoReg。SoReg 利用社会正则化条件约束矩阵分解目标函数,应用于解决信任感知的推荐问题。

(3)NeuMF^[35]。NeuMF 是一个神经协同过滤的矩阵分解框架,解决了内积不足以捕捉到用户交互数据中复杂结构信息的问题,可以探索用户与没有社会关系信息的项目之间的深层次和内部交互作用。

(4)SocialMF。SocailMF 是一种基于信任传播的模型,通过矩阵分解来学习用户和项目的潜在特征向量,其中每个用户的特征向量依赖于其在社交网络中的直接邻居的特征向量,SocailMF 的优势就在于考虑了推荐系统的信任信息和信任信息在矩阵分解模型中的传播。

(5)TrustMF。TrustMF 采用矩阵分解技术,对信任网络进行分解,将用户和项目映射到相同的低维潜在特征空间,生成信任者和受托者模型,分别具有不同的含义。

(6)SoRec。SoRec 作为一种新的社会推荐框架,采用概率矩阵分解法将用户的项目评分矩阵与用户的社交网络相结合。模型的中心思想是基于概率矩阵分解 PMF 的因子分析,学习低阶用户潜在特征空间和项目潜在特征空间,提出了社会化建议。

(7)GraphRec。GraphRec 创建了一个基于图神经网络的社会推荐模型,提供了一种原则性的方法来联合捕获用户项目图中的交互和意见。同时,该模型也通过加入注意力机制

来考虑社会关系的异质性优势,从而区分关系优势。

4.2 实验结果与分析

4.2.1 结果对比

表 2—表 4 列出了 AGNN-SR 模型与 baselines 在 Ciao, Epinions, FilmTrust 这 3 个数据集上的推荐性能对比。

表 2 Ciao 上的性能对比

Metric	MAE	RMSE	
Algorithms	PMF	0.9057	1.1302
	SoReg	0.8702	1.0935
	NeuMF	0.8272	1.0739
	SocialMF	0.8314	1.0651
	TrustMF	0.7950	1.0586
	SoRec	0.8633	1.0711
	GraphRec	0.7432	1.0017
AGNN-SR	0.7301	0.9937	
Improve	1.79%	0.47%	

表 3 Epinions 上的性能对比

Metric	MAE	RMSE	
Algorithms	PMF	1.0099	1.2217
	SoReg	0.9201	1.1835
	NeuMF	0.9147	1.1442
	SocialMF	0.8952	1.1479
	TrustMF	0.8576	1.1402
	SoRec	0.8996	1.1521
	GraphRec	0.8169	1.0791
AGNN-SR	0.7871	1.0547	
Improve	3.78%	2.31%	

表 4 FilmTrust 上的性能对比

Metric	MAE	RMSE	
Algorithms	PMF	0.7995	0.9730
	SoReg	0.6788	0.8933
	NeuMF	0.6641	0.8547
	SocialMF	0.6503	0.8473
	TrustMF	0.7950	0.8211
	SoRec	0.6417	0.8109
	GraphRec	0.6294	0.7994
AGNN-SR	0.5975	0.7742	
Improve	5.34%	3.25%	

从实验结果可以看出,PMF 由于是传统的简单矩阵分解模型,仅仅使用了用户对项目的评级这一数据,没有考虑社交信息,因此在数据集上的表现最差。而 NeuMF 虽然也不包含社交关系数据,但因为该模型引入了深度神经网络,利用神经网络在推荐模型中的优势,相比传统的矩阵分解,更好地模拟了用户与项目之间的交互,因此比 PMF 等传统模型有更好的表现,更好地说明了深度神经网络对提高推荐模型准确度的必要性。SoReg, SocialMF, TrustMF, SoRec 等模型融合了用户对项目的评级信息和用户的社交网络,进一步获取了更全面的用户和项目的向量表示,因此这些模型都比传统的 PMF 拥有更好的推荐效果。而 GraphRec 模型在图神经网络的基础上得到了更进一步的优化,相比其他 baselines,它拥有最佳的实验结果。GraphRec 利用图神经网络学习更加准确的节点嵌入表示,而图数据结构的天然优势,又使得用户项目交互信息和社交网络信息能在图上更好地传播。同时,结合

注意力机制的图神经网络可以很好地区分不同强度的社交关系对用户产生的不同程度的影响与贡献。实验结果表明, 本文的 AGNN-SR 模型在所有实验中表现最好。与 baselines 中效果最好的 GraphRec 相比, AGNN-SR 改进了注意力机制, 在自注意力的基础上利用多头注意力机制从更多不同的角度进行信息聚合, 获取更加准确的用户与项目潜在因子表示; 其次, 模型中的嵌入传播层提取用户与项目交互的协作信号, 注入高阶交互信息, 以产生更优的嵌入表示。接下来将对模型进行更加深入的研究分析, 以更好地了解不同组件对模型提升做出的贡献。

4.2.2 模型分析

(1) 多头注意力机制的影响

现有的研究^[36]将多头注意力机制应用于图神经网络, 并取得了优异的实验结果。自然而然, 尝试将该机制引入社会推荐中, 从多个角度去学习节点嵌入。为了进一步评估多头注意力机制对模型性能的影响, 将 AGNN-SR 与 GNN-SR, GNN-SR- α 两种模型变体进行实验对比。变体模型的定义如下。

GNN-SR: 不使用任何注意力机制, 直接利用基于均值的聚合函数来建模, 以获取用户和项目潜在因子, 模型效果很大程度上受限于均值聚合器。

GNN-SR- α : 使用普通的自注意力机制, 通过调整注意力系数来了解用户和项目的特征表示。

使用不同注意力机制的模型的实验结果如图 4 所示。从图中可以看出, 完全不使用注意力机制的模型的推荐精度明显下降, 这证明了在社交网络中, 不同的社交好友对用户产生的影响是不尽相同的, 因此区分具有异质优势的社交关系非常重要。同时, 实验结果也表明, 相比普通注意力机制为邻居节点分配系数, 多头注意力机制是在自注意力机制的基础上, 为了稳定自注意力的学习过程, 并行执行了 k 个相互独立的注意力机制计算, 然后将它们的特征向量连接起来, 得到最终的潜在特征向量表示的过程。多头注意力机制能多角度地学习更准确的、更高层次的用户和项目潜在因子, 进而获取更优的推荐效果。

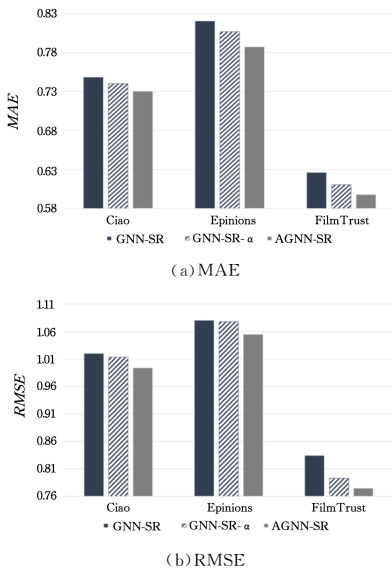


图 4 注意力机制对数据集的影响

Fig. 4 Effect of attention mechanisms on datasets

在多头注意力机制中, 注意头的数量 k 是一个十分重要的参数, 它对模型的性能有着至关重要的影响。因此对比了不同数量的注意力头下的模型效果, 分析了参数 k 对模型的影响。实验结果如图 5 所示。

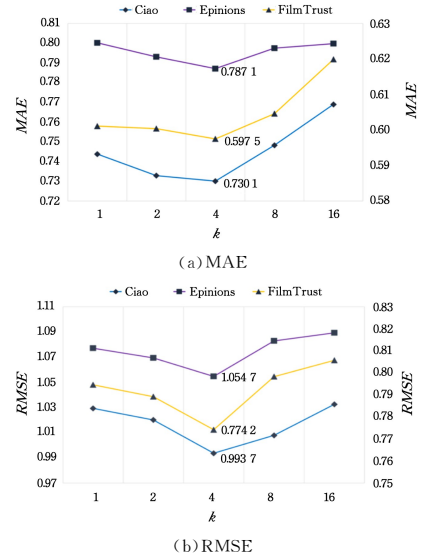


图 5 注意力头个数 k 对数据集的影响

Fig. 5 Effect of the number of attention head on datasets

从图 5 可以看出, 由于不同数量的头部具有不同的注意层维度, 因此当注意力头从 1 增加到 4 时, 推荐精度逐步提高, 但注意力头部个数从 4 继续增加时, 模型性能反而下降了。因此, 注意力头个数为 4 时能很好地平衡性能与复杂度, 需要在注意头的数量和注意层的维度之间找到一个平衡点, 以使模型的推荐性能达到最佳。

(2) 神经协同过滤模块^[37]的影响

为了探索神经协同过滤模块对 AGNN-SR 模型的重要性, 本文设计了变体 AGNN-SR-IP 模型, 与 AGNN-SR 模型进行实验对比。AGNN-SR-IP 对用户和项目的潜在因子应用经典内积方法进行评级预测, 替换 AGNN-SR 模型中的神经协同过滤推荐模块, 内积是推荐系统中矩阵分解的经典方法。

实验结果如图 6 所示。

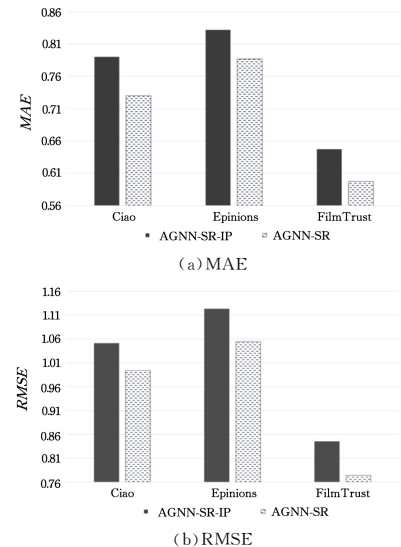


图 6 神经协同过滤对数据集的影响

Fig. 6 Effect of neural collaborative filtering on datasets

从图 6 可以看出,使用经典内积的 AGNN-SR-IP 模型的表现远不如神经协同过滤模型 AGNN-SR,这很好地说明了基于深度神经网络的协同过滤技术对推荐系统的重要性。神经协同过滤能有效地捕获用户和项目间的深层交互,使得推荐模型有更好的效果。

(3) 嵌入传播层层数的影响

为了探究模型从多个嵌入传播层中获益的程度大小,通过调整传播层层数来进行实验对比,层数范围为 1~4,实验结果如表 5 所列。从中可以得出,当逐步增加传播层层数时,模型能有效地学习获取用户和项目的协作信号以及用户间的相似度,推荐精度逐步提升;但是当继续叠加嵌入层数时,模型在数据集 FilmTrust 上会出现过拟合,可以推测这是由于过深的网络结构给学习带来了意外的噪声;同时,综合其他两个数据集的结果,传播层层数为 3 时足以捕获协作信号,盲目地增加传播层层数并不会给模型带来明显的提升。

表 5 嵌入传播层层数的影响

Table 5 Effect of different embedding propagation layers

Layers	Ciao		Epinions		FilmTrust	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
1	0.7330	0.9968	0.7889	1.0563	0.5994	0.7766
2	0.7326	0.9942	0.7885	1.0559	0.5980	0.7758
3	0.7301	0.9937	0.7871	1.0547	0.5975	0.7742
4	0.7299	0.9936	0.7870	1.0545	0.5983	0.7751

结束语 本文提出的 AGNN-SR 模型分别从用户项目交互图和社交图两个图上,利用多头注意力机制,多方位多角度地学习更为准确的用户和项目的潜在向量;同时设计神经协同过滤推荐模块来编码用户和项目间的关键协作信号。在 3 个真实数据集上进行实验对比,实验结果证明了本文 AGNN-SR 模型的有效性,能有效提高推荐质量。通过对模型参数的进一步分析,探索多头注意力机制和嵌入传播层相关参数的贡献。

未来,将深入探索用户和项目交互行为发生的时间以及用户与社交好友产生关系的时间对推荐的影响,通过建模时间信息有效地提升社交推荐系统的性能。

参 考 文 献

[1] ZHAO J Y, ZHUANG F Z, AO X, et al. A Overview of Collaborative Filtering Recommendation System[J]. Journal of Cyber Security, 2021, 6(5): 17-34.

[2] ZHAO W T, ZHANG S. Collaborative Filtering Algorithm Based on User Preference Under Sparse Data[J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2021, 33(4): 669-674.

[3] KANG Y, LI T, LI H, et al. Recommendation Model Fusing with Knowledge Graph and Collaborative Filtering[J]. Computer Engineering, 2020, 46(12): 73-79, 87.

[4] SHEN J, QIAO S J, HAN N, et al. Personalized Recommendation Model with Multi Information[J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2021, 35(3): 128-138.

[5] SHI H Y, NI Y D. Research Progress on Cold Start of Recom-

mendation System[J]. Research on Library Science, 2021(12): 2-10.

[6] MA H, YANG H, LYU M R, et al. Social Recommendation Using Probabilistic Matrix Factorization[C]// Proceedings of the 17th ACM Conference on Information and Knowledge Management. Association for Computing Machinery, 2008: 931-940.

[7] MA H, ZHOU D, LIU C, et al. Recommender Systems with Social Regularization[C]// Proceedings of the Forth International Conference on Web Search and Web Data Mining (WSDM 2011). Association for Computing Machinery, 2011: 287-296.

[8] JAMALI M, ESTER M. A Matrix Factorization Technique with Trust Propagation for Recommendation in Social Networks [C]// Proceedings of the Fourth ACM Conference on Recommender Systems. Association for Computing Machinery, 2010: 135-142.

[9] GUO G, ZHANG J, YORKE-SMITH N. TrustSVD: Collaborative Filtering with Both the Explicit and Implicit Influence of User Trust and of Item Ratings[C]// Twenty-Ninth AAAI Conference on Artificial Intelligence, 2015.

[10] YANG B, LEI Y, LIU J, et al. Social collaborative filtering by trust[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(8): 1633-1647.

[11] WU L, SUN P, HONG R, et al. SocialGCN: An efficient graph convolutional network based model for social recommendation [J]. arXiv:1811.02815, 2018.

[12] FAN W, MA Y, LI Q, et al. Graph Neural Networks for Social Recommendation [C] // The World Wide Web Conference. ACM, 2019: 417-426.

[13] GUO Z, WANG H. A deep graph neural network-based mechanism for social recommendations[J]. IEEE Transactions on Industrial Informatics, 2020, 17(4): 2776-2783.

[14] LUO D, BIAN Y, ZHANG X, et al. Attentive social recommendation: towards user and item diversities[J]. arXiv:2011.04797, 2020.

[15] SONG C, WANG B, JIANG Q, et al. Social Recommendation with Implicit Social Influence [C]// Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. Association for Computing Machinery, 2021: 1788-1792.

[16] YU J, YIN H, LI J, et al. Self-Supervised Multi-Channel Hypergraph Convolutional Network for Social Recommendation [C]// Proceedings of the Web Conference 2021. Association for Computing Machinery, 2021: 413-424.

[17] LIU Y, CHEN L, HE X, et al. Modelling high-order social relations for item recommendation [J]. IEEE Transactions on Knowledge and Data Engineering, 2022, 34(9): 4385-4397.

[18] HUANG C, XU H, XU Y, et al. Knowledge-Aware Coupled Graph Neural Network for Social Recommendation [C]// Proceedings of the AAAI Conference on Artificial Intelligence. AAAI, 2021: 4115-4122.

[19] ZHANG J, GAO C, JIN D P, et al. Group-buying recommendation for social e-commerce[J]. arXiv:2010.06848, 2020.

- [20] BAI T, ZHANG Y, WU B, et al. Temporal Graph Neural Networks for Social Recommendation[C]// Proceedings of the 2020 IEEE International Conference on Big Data (Big Data). IEEE, 2020:898-903.
- [21] KOREN Y, BELL R, VOLINSKY C. Matrix factorization techniques for recommender systems[J]. Computer, 2009, 42(8): 30-37.
- [22] MCPHERSON M, SMITH-LOVIN L, COOK J M. Birds of a feather; homophily in social networks[J/OL]. Annual Review of Sociology, 2001, 27: 415-444. <http://www.jstor.org/stable/2678628>.
- [23] MARSDEN P V, FRIEDKIN N E. Network studies of social influence[J]. Sociological Methods & Research, 1993, 22(1):127-151.
- [24] TANG J, HU X, LIU H. Social recommendation: a review[J]. Social Network Analysis & Mining, 2013, 3(4):1113-1133.
- [25] YANG X, GUO Y, LIU Y, et al. A survey of collaborative filtering based social recommender systems[J]. Computer Communications, 2014, 41(5):1-10.
- [26] GOLBECK J A. Computing and applying trust in web-based social networks[D]. Maryland, USA: University of Maryland, College Park, 2005.
- [27] GOLBECK J, MANNES A. Using Trust and Provenance for Content Filtering on the Semantic Web[C]// Proceedings of the Workshop on Models of Trust for the Web. Edinburgh, UK, 2006:23-35.
- [28] MASSA P, AVESANI P. Trust-aware collaborative filtering for recommender systems[C]// OTM Confederated International Conferences on the Move to Meaningful Internet Systems. Berlin; Springer, 2004:492-508.
- [29] JIANG M, CUI P, LIU R, et al. Social Contextual Recommendation[C]// Proceedings of the 21st ACM International Conference on Information and Knowledge Management. Association for Computing Machinery, 2012:45-54.
- [30] HAMILTON W L, YING R, LESKOVEC J. Inductive Representation Learning on Large Graphs[C]// Proceedings of the 31st International Conference on Neural Information Processing Systems. Curran Associates Inc, 2017:1025-1035.
- [31] VASWANI A, SHAZEER N, PARMAR N, et al. Attention Is All You Need[C]// Advances in Neural Information Processing Systems, 2017:5998-6008.
- [32] BERG R, KIPF T N, WELLING M. Graph convolutional matrix completion[J]. arXiv:1706.02263, 2017.
- [33] MAAS A L, HANNUN A Y, NG A Y. Rectifier Nonlinearities Improve Neural Network Acoustic Models[C]// Proceedings of the ICML Workshop on Deep Learning for Audio, Speech and Language Processing, 2013.
- [34] SALAKHUTDINOV R, MNII A. Probabilistic Matrix Factorization[C]// Proceedings of the 20th International Conference on Neural Information Processing Systems. Curran Associates Inc, 2007:1257-1264.
- [35] HE X, LIAO L, ZHANG H, et al. Neural Collaborative Filtering [C] // Proceedings of the 26th International Conference on World Wide Web. International World Wide Web Conferences Steering Committee, 2017:173-182.
- [36] MARSDEN P V, FRIEDKIN N E. Network studies of social influence[J]. Sociological Methods & Research, 1993, 22(1):127-151.
- [37] WANG X, HE X, WANG M, et al. Neural Graph Collaborative Filtering[C]// Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, 2019.



ZHANG Qi, born in 1997, postgraduate. Her main research interests include recommendation system and so on.



YIN Hongfeng, born in 1975, master, associate professor. His main research interests include machine learning and distributed system.

(责任编辑:喻黎)