

基于记忆的注意力图神经网络专家推荐方法

陈卓,朱淼[†],杜军威,袁玺明
(青岛科技大学信息科学技术学院,山东青岛266061)

摘要:在问答社区专家推荐算法中,图神经网络主要利用问答社区中用户与问题的交互关系建模,其模型性能取决于交互数据的稠密度,难以对无交互信息的用户及问题进行有效表示学习.针对这一问题,提出了一个基于记忆的注意力图神经网络专家推荐方法.该方法首先设计了面向用户多维特征的联合表示子网络,然后构建了一个记忆网络,为每个问题保存用户回答过的与其相似的问题,同时在用户表示与相似问题表示之间引入注意力机制,从不同用户的视角,有针对性的融合相似问题构建新问题的向量表示,最终基于用户和问题的表示为问题推荐专家,有效提高了专家推荐的准确性.在数据集中对本文所提出的方法进行验证,相较于同类其他模型性能均有所提升.

关键词:专家推荐;图神经网络;记忆网络;注意力机制
中图分类号:TP18 **文献标志码:**A

Memory-based Attention Graph Neural Network for Network Expert Recommendation

CHEN Zhuo, ZHU Miao[†], DU Junwei, YUAN Ximing

(School of Information Science and Technology, Qingdao University of Science and Technology, Qingdao 266061, China)

Abstract: In the expert recommendation algorithm of the Q & A community, the graph neural network mainly uses the interactive relationship between users and questions to build a model, and its model performance depends on the density of interactive data. So it is difficult to effectively represent and learn users and questions without interactive information. This paper proposes an attention graph neural network expert recommendation method based on memory. Firstly, a multi-dimensional feature-oriented subnetwork is designed, and then a memory network is constructed to store the similar questions answered by users for each question. At the same time, an attention mechanism is introduced between user representation and similar question representation, and the vector representation of new questions is constructed by fusing similar questions from different users' perspectives. Finally, experts are recommended based on the representation of users and questions, which effectively improves the accuracy of expert recommendation. The proposed method is validated on the Q & A community data set and public data set, and its performance is improved compared with other similar models.

Key words: expert recommendation; graph neural network; memory network; attention mechanism

* 收稿日期:2021-06-29

基金项目:国家自然科学基金资助项目(F030810,6217072142), National Natural Science Foundation of China (F030810, 6217072142);山东省重点研发计划项目资助(2018GGX101052), Shandong Provincial Key Research and Development Plan Project Funding (2018GGX101052)

作者简介:陈卓(1978—),女,山东青岛人,青岛科技大学副教授,博士

[†]通信联系人,E-mail:1264196623@qq.com

在线问答社区是当下万维网普及和社会快速发展的产物,模拟现实中的社区,使人们能够即刻交流和获取知识.而快速、准确、主动地发现社区用户中的领域专家,成为提高用户参与社区互动积极性、提高社区内容质量的关键性技术问题.

然而,目前问答社区的专家推荐策略主要依靠传统的信息检索和排序技术对符合条件的用户进行排序推荐.这类方法没有考虑问题的语义信息,无法满足用户的需求.后来学者将推荐算法应用在专家推荐领域.这类算法通常计算用户偏好特征与问题内容文本特征之间的匹配度为目标问题产生候选专家用户集或者通过相似历史问题表示新问题,未考虑用户特征,因而无法体现现实世界中,用户在教育背景、从业经历等方面的差异.随着深度学习的发展,深度学习的方法被引入专家推荐.这类专家推荐方法从问题文本特征的角度表示问题的内容,通过计算专家得分进行推荐.但未考虑用户的回答能力和社交关系,无法完整的建模用户信息.

针对现有的专家推荐模型的问题,本文在用户记忆力的基础上,提出了一种融合用户、问题、用户与问题交互信息的图神经网络专家推荐模型(Graph recommendation method based on memory neural network, GMNN). 本文的贡献可以总结为以下两点:

1) 给出了基于图的用户联合特征表示学习方法,解决了问答社区中高维系数的用户特征建模难题,增强了用户的表示能力.

2) 利用记忆网络和注意力机制从用户的不同视角表示新问题,解决了数据较稀疏的情况下,传统的图神经网络难以对社区中没有与用户产生交互行为的问题进行表示的难题.

1 相关工作

1.1 专家推荐算法

专家推荐的目的是根据提出的问题内容,在社区中找到具有一定解决或回答问题能力的专家,提高问答效率,优化用户体验.研究前期,专家推荐主要依靠传统的信息检索和排序技术.目前的专家推荐算法通常利用自然语言处理技术进行文本特征的提取,然后挖掘用户和问题之间的相似关系提供推荐服务.如San等^[1]基于问答社区的用户倾向于回答自己专业问题这一基本原理,提出了一种在线CQA社区专家模型RankSLDA算法,利用专业知识

的主题对用户进行建模.Mandal等^[2]提出了一种查询似然语言模型中的主题方法来为社区中新提出的问题推荐潜在的回答者.这类方法忽略了用户与问题交互网络中的拓扑结构,无法捕获用户与问题在网络中的关联信息.图神经网络出现后,因其对图形数据强大的表示能力自然的被应用在推荐系统中.Fan^[3]提出了一个面向社交推荐的图神经网络框架GraphRec,该框架联合用户的社交信息和用户与项目的交互信息对用户进行向量表示学习.王根生^[4]通过计算不同元路径下的用户相似度,得出用户相似度矩阵,然后把用户相似度矩阵融合到矩阵分解推荐算法的目标函数中.

图神经网络能够捕获节点的特征信息与网络的结构信息,但是多数图神经网络无法对没有边的新节点产生向量表示,因此无法解决推荐系统中的冷启动问题.

1.2 记忆网络

记忆网络由Weston^[5]等最先提出,通过构建记忆组件将信息存储起来,以达到长期记忆的效果.在此之前,很多神经网络都有一定的记忆机制,但由于存储能力有限,在训练过程中容易丢失一部分语义信息.如由Hochreiter等^[6]提出的长短期记忆网络(Long Short-Term Memory, LSTM)有效解决了RNN记忆能力弱,无法存储足够信息的问题.但在现实世界任务中,如在数据较稀疏的推荐领域,短期的记忆难以完整的表示项目,需要更加长久的记忆以支撑系统的建模从而提高准确度.Weston等在最初的记忆网络基础上又提出了端到端的记忆网络^[7]以及键值记忆网络^[8],使网络可以更好的存储问答系统所需要的先验知识.目前记忆网络已在语音识别、社交推荐等各领域任务上取得了显著的成功.

1.3 注意力机制

注意力机制最早为解决图像分类问题提出,其目的是将视线聚焦于输入的信息中对当前任务更为重要的信息,减少对其他信息的关注,提高效率.何柔莹^[9]通过利用注意力机制增强卷积神经网络对于问题文本特征的提取能力,使建模更加准确.吕晓琦^[10]结合注意力机制与循环神经网络,除了实现了问题与标签的联合表示,还捕获了动态变化的用户偏好.

注意力机制目前在推荐系统中已有广泛的应用,通过为用户和项目分配注意力权重,以更有效地捕捉与任务相关的长期或短期的用户兴趣^[11].

2 方法

随着问答社区用户规模和问题数目的不断扩大,传统的图神经网络难以对社区中没有与用户产生交互行为的问题进行表示,为了解决这一问题,本文构建了一种基于记忆的注意力图神经网络推荐方法(Memory based attention graph neural network rec-

ommendation method, GMNN)为问答社区中的新问题进行专家推荐,算法模型如图 1 所示,图中显示模型由三部分组成:用户特征联合表示学习模块、新问题表示模块和用户回答能力预测模块.前两个模块分别对用户和问题的向量进行表示学习,然后在用户回答能力预测模块进行拼接,输入多层感知机,输出对用户回答能力的预测结果.

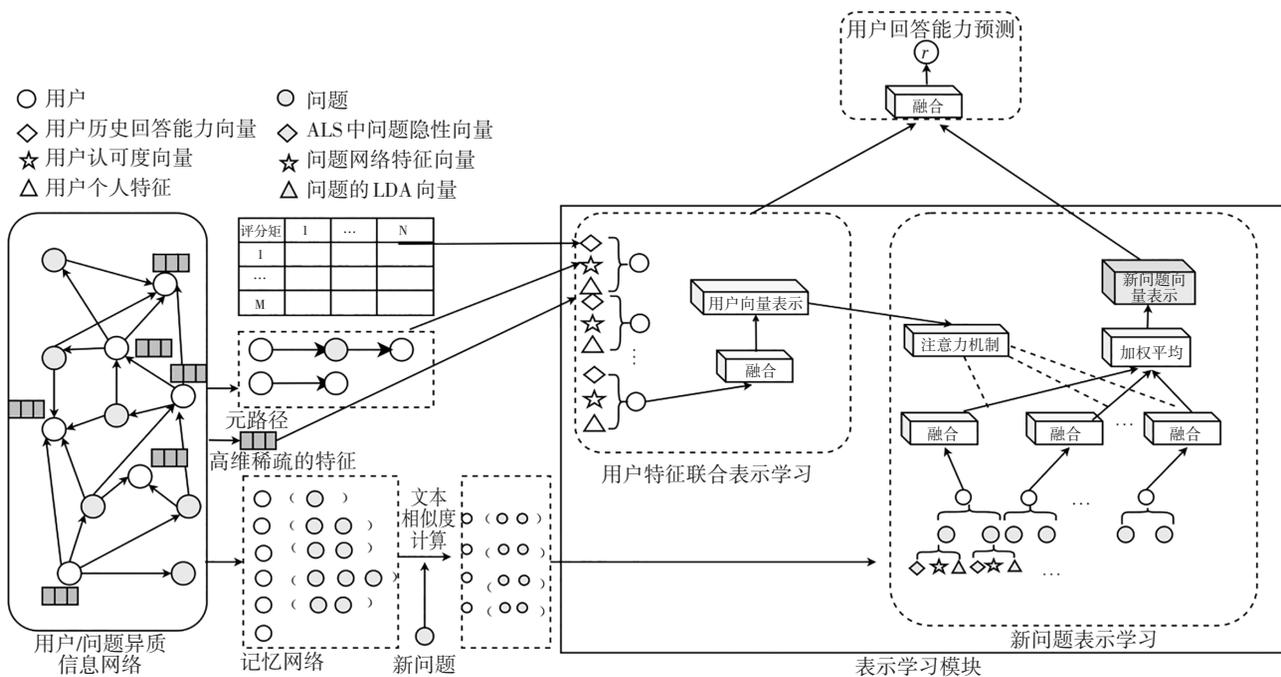


图 1 GMNN 模型图

Fig.1 GMNN model diagram

2.1 用户特征联合表示学习

问答社区通常会记录用户的个人基本信息、用户在社区中提出的问题、回答的答案及由此获得的财富打赏等交互信息.传统的社区用户表示学习方法,基本都是直接从用户个人基本信息、用户与问题交互信息生成用户的向量表示.但随着社区规模的扩大,问答社区平均用户回答问题个数少,用户与问题之间的交互矩阵很稀疏,传统的用户表示学习方法难以准确表示用户的偏好和回答能力等信息,因此本文联合用户多个维度的特征,对用户进行表示学习.本文首先利用矩阵分解中的交替最小二乘法(Alternating Least Squares, ALS)算法对用户的历史回答数据进行分析,计算出能够表示用户历史回答能力的隐性向量矩阵 P ;再利用用户的问答与打赏关系构建异质信息网络,使用深度游走算法对网络中的用户和问题进行网络表示学习,计算出能够表

示用户认可度的特征矩阵 G_1 ;对问答社区中高维稀疏的用户个人特征矩阵自编码,转化为低维稠密的用户个人特征矩阵 I .本文利用多层感知机对上述三个维度的用户特征进行融合,最终预测用户对新问题的回答能力.

2.1.1 用户历史回答能力建模

矩阵分解可从已知评分数据中学习用户的隐性偏好和项目隐性特征.相比于其他矩阵分解算法,ALS更加适合稀疏矩阵的矩阵分解,因此本文使用ALS对用户/问题交互数据进行矩阵分解,产生用户历史回答能力的隐性向量表示.

用户历史回答能力由该用户历史所获得财富打赏值表现出来.本文使用带标签和伪标签用户/问题交互数据构建用户/问题交互矩阵.其中 $U = \{u_1, u_2, \dots, u_M\}$ 和 $Q = \{q_1, q_2, \dots, q_N\}$ 分别表示 M 个用户和 N 个问题,用户参与回答问题所获得的财富值

构成了一个 M 行 N 列的评分矩阵 $R_{M \times N}$, 缺失的财富值由 DeepFM 预测.

在问答社区中, 使用 ALS 将用户和问题的评分矩阵 $R_{M \times N}$ 分解为 $M \times K$ 维的用户隐性矩阵和 $K \times N$ 维的问题隐性矩阵, 超参数 K 为潜在因子空间. 用户/问题评分矩阵可以表示为用户矩阵与问题矩阵的乘积, 计算公式见公式(1).

$$R_{M \times N} = P_{M \times K} D_{K \times N} \quad (1)$$

式中: $P_{M \times K}$ 表示用户隐性向量矩阵, $D_{K \times N}$ 表示问题隐性向量矩阵.

用户 u_i 对问题 q_j 的评分预测值 r'_{ij} 的计算公式见公式(2).

$$\hat{r}_{ij} = p_i d_j \quad (2)$$

式中: p_i 表示用户隐性向量, d_j 表示问题隐性向量. 在计算 $P_{M \times K}$ 与 $D_{K \times N}$ 的乘积与原始矩阵的误差时, 损失函数的计算公式见公式(3).

$$LOSS = \sum_{ij} (r_{ij} - \hat{r}_{ij})^2 + \lambda (\|P\|^2 + \|D\|^2) \quad (3)$$

式中: λ 表示正则化系数. 为了优化损失函数, 使用交替固定 P 和 D 对损失函数求偏导, 降低损失函数.

2.1.2 用户认可度建模

在实际问答过程中, 其他用户对回答者给予的财富值能反映出对该答案的认可程度, 专家等权威用户给出的财富值比普通用户更有价值; 而有权威用户参与的问题往往也具有更高的质量和难度. 用户被认可的程度可以通过他回答过哪些用户的提问、哪些用户与他共同回答过同一问题以及哪些用户打赏过他的回答这三个方面体现.

本文将用户和问题作为节点, 用户与问题之间的问答关系和用户间的打赏关系作为边, 构建图 1 中用户/问题异质信息网络, 并生成元路径, 根据元路径, 对用户进行网络表示学习, 表示用户认可度. 本文将化工问答社区的问答关系和打赏关系整理为异质信息网络 G , G 的定义见公式(4).

$$G = (V, E), E \in (V \times V) \quad (4)$$

式中: V 表示 G 中的顶点集, 顶点分为用户和问题两个类别. E 代表了 G 中顶点与顶点之间连接的边集, 用户顶点与问题顶点之间的边表示用户提出或回答了该问题, 用户顶点与用户顶点之间的边则表示用户之间的打赏关系.

根据化工问答社区用户间的交互形式, 本文为随机游走设计了两种元路径.

元路径 1: 体现用户和问题间的问答关系

用户 $\xrightarrow{\text{提问}}$ 问题 $\xrightarrow{\text{回答}}$ 用户

元路径 2: 体现用户和用户间的打赏关系.

用户 $\xrightarrow{\text{打赏}}$ 用户

对于本文构建的异质信息网络 G 和元路径 p : $A_1 \xrightarrow{R_1} A_2 \xrightarrow{R_2} \dots \xrightarrow{R_i} A_{i+1}$, 式(5)表示分布式生成游走路径 p .

$$p(v_{i+1} | v_i, p) = \begin{cases} \frac{1}{|N^{A_{i+1}}(v)|}, & v \in \mathcal{E}, \phi(v_{i+1}) = A_{i+1} \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

式中: v_i 是路径中的第 i 个节点, v 的类型为 A_i , A_i 共有问题和用户两种类型, $N^{A_{i+1}}$ 是具有 A_{i+1} 类型的节点 v 的一阶邻域集, 游走将重复遵循元路径的模式直到它达到预定义的长度. 为了计算节点 v_i 的向量, 本文选用深度游走的方法根据元路径采样求节点表示, 这种方法不仅能适应网络的变化, 而且对网络中相似的用户, 低维连续的用户向量表示也是相似的. 深度游走使用公式(6)最大化 v_i 到其周围点的平均概率, 生成向量矩阵.

$$\underset{\varphi}{\text{minimize}} - \log \Pr \left(\{v_{i-n}, \dots, v_{i-1}, v_{i+1}, \dots, v_{i+n}\} \mid \varphi(v_i) \right) \quad (6)$$

式中: φ 表示节点 v_i 到对应向量的映射函数, \Pr 表示在随机游走过程中, 出现节点 v_i 时, 出现它的 n 个节点范围内节点的概率.

2.2 新问题表示学习

由于传统的图神经网络无法从交互网络中直接学习新问题的表示, 因此在图神经网络之前加入保存用户历史回答问题集合的记忆网络, 并通过注意力机制实现面向用户视角的新问题表示学习方法.

2.2.1 用户记忆网络的构建

基于内容的推荐思想能够为新问题从社区中找到与之相似的旧问题, 但是这种方法忽略了用户视角对于问题的理解程度以及偏好, 因此本文首先构建用户历史回答问题集合的记忆网络, 然后通过文本相似度的计算为新问题从记忆网络中寻找相似问题序列. 为了更好地计算文本之间的相似程度, 本文通过 BERT (Bidirectional Encoder Representations from Transformers) 模型^[12]进行文本表示后计算问题文本向量之间的余弦相似度.

当新问题出现时, 首先使用 BERT 预训练模型对问题文本生成了向量表示, 然后基于文本向量之间的余弦相似度, 计算出在用户历史回答过的问题中, 与新问题最相似的旧问题, 生成矩阵 S_{ij} , j 为与新问题相似的问题个数, 为超参数, i 为历史回答数大于

相似问题个数的用户人数. 对于任意问题 q , 都可以按照相似度排序找到与文本内容相似的问题序列.

本文为新问题寻找 n 个最相似的问题 ID 序列进行保存, 当需要对新问题进行向量表示时, 先通过记忆网络构建其相似度矩阵, 对这些相似旧问题进行表示学习.

2.2.2 基于注意力的新问题表示学习

对于不同用户来说, 针对每个相似问题的回答能力和偏好是不相同的, 因此本文引入注意力机制, 从不同的用户视角, 为每个相似问题分配一个权重来表示每个相似问题对新问题表示时做出的不同贡献, 注意力网络的计算见公式(7).

$$a^* = W_0^T \cdot \sigma(W_1 \cdot [u_c \oplus p_m] + b_0) + b_1 \quad (7)$$

式中: u_c 为用户向量表示, p_m 为相似问题的向量表示. 使用 softmax 函数获得最终注意力权重, 可以被解释为每个相似问题在表示新问题时对新问题潜在因子的贡献, 每个问题 t 的权重 a_t 计算见公式(8).

$$a_t = \frac{\exp(a_t^*)}{\sum_{i \in T} \exp(a_i^*)} \quad (8)$$

式中: T 表示相似问题集合, 最终通过注意力机制对相似旧问题给予不同权重, 获得新问题的表示 q', q' 的计算见公式(9).

$$q' = \sigma(W \cdot \left\{ \sum_{i \in T} p_i a_i \right\} + b) \quad (9)$$

2.3 基于记忆的图神经网络专家推荐方法

本文将用户向量表示和新问题向量表示两者拼接后输入多层感知机进行深度融合, 最终预测出目标用户 i 对于新问题 j 回答后可能获得的财富值 r'_{ij}, r'_{ij} 的计算公式见公式(10)至公式(13).

$$g_0 = (u_c \oplus q') \quad (10)$$

$$g_1 = \sigma(W_2 \cdot g_0 + b_0) \quad (11)$$

...

$$g_t = \sigma(W_t \cdot g_{t-1} + b_{t-1}) \quad (12)$$

$$r'_{ij} = W^T \cdot g_t \quad (13)$$

计算出 r'_{ij} 后, 为了训练 GMNN 模型的参数, 需要设置一个目标函数进行优化. 由于 GMNN 模型的任务是预测用户回答问题获得的财富值, 因此本文采取 MSE 作为 GMNN 模型的损失函数, MSE 的计算公式为式(14).

$$\text{MSE} = \frac{1}{n} \sum (r_{ij} - r'_{ij})^2 \quad (14)$$

式中: r_{ij} 为用户 i 回答问题 j 所获得的实际财富值, GMNN 模型中的权重以及偏置参数由随机初始化获

得, 通过随机梯度下降的方法更新各个参数的值, 最小化模型的损失, 使 GMNN 模型的预测效果达到最优. 预测出所有用户对新问题回答后可能获得的财富值, 将其进行排序, 将问题推荐给排名靠前即预测回答能力强的用户回答.

3 实验与分析

3.1 实验数据集

作为国内最大的化工领域问答社区, 海川化工问答社区为广大从业人员提供了获取专业知识与分享化工领域知识的平台, 见证并记录了国内化工行业的发展, 据官方统计, 该化工问答社区已拥有约 480 万的用户. 它由第一代问答网站发展而来, 并在问答网站中加入了搜索功能, 是知识社区和社交网络的结合体. 本文爬取了海川化工问答社区中 6 017 名用户的问答数据及获得财富值情况, 在统计时间内, 这 6 017 名用户共对 10 633 个问题做出了 35 773 条回答并获得了财富值. 由于 GMNN 模型着重于新问题的解决, 因此本文对数据集使用基于时序的划分方法, 将 2017 年 11 月 7 日至 2018 年 10 月 7 日的问答数据作为训练集, 2018 年 10 月 8 日至 12 月 7 日的问答数据作为测试集, 划分结果见表 1.

表 1 化工问答社区数据集

Tab. 1 Dataset of the Chemical Q & A community

数据集	问题个数	答案个数	时间跨度
训练集	9 079	31 637	11 个月
测试集	1 554	4 136	2 个月

3.2 实验对比模型

本文针对的问题是在数据稀疏情况下难以对无交互信息的用户及问题进行有效表示, 从而难以对新问题推荐专家回答. 本文选取了以下几种对比模型, 在推荐领域对稀疏数据有很好的表示能力.

Factorization Machine (FM)^[13]: 该模型旨在解决大规模稀疏数据下的特征组合问题, 考虑不同特征之间的相互关系, 被广泛地用于预测和推荐任务上.

Deep Factorization Machine (DeepFM)^[14]: 该模型在新的神经网络中集合了用于推荐的 FM 模型的功能和深度学习的特征学习能力, 可以更快地训练以及更精准的学习并给出推荐.

Wide&Deep^[15]: 该模型被用于推荐系统中, 结合

了神经网络的泛化能力和线性模型的记忆能力,通过对稀疏特征的学习,提高了推荐系统的准确性和多样性.

3.3 评价指标

本文采用均方根误差(Root mean squared error, RMSE)和平均绝对误差(mean absolute error, MAE)、召回率、归一化折损累计增益(Normalized Discounted cumulative Gain, NDCG)和平均倒数排名(Mean Reciprocal Rank, MRR)评价模型回答者推荐效果.

3.4 参数选择

对于本文提出的 GMNN 算法,通过将矩阵分解得到的用户和问题隐性向量、深度游走算法得到的用户和问题的网络表示、问题 LDA 概率向量和用户个人特征进行输入,对选取相似旧问题个数进行选择,进行分组实验,实验结果见图 2.

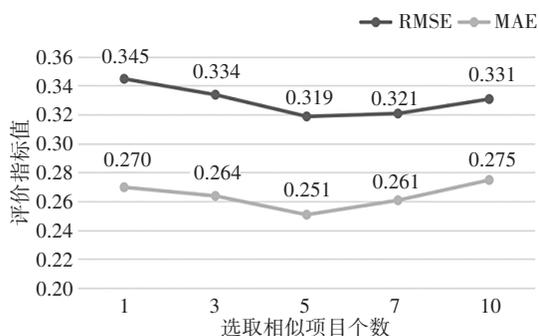


图 2 性能随相似旧问题个数变化趋势图

Fig.2 Performance trend with the number of similar projects

从图 2 可知,参与训练模型的相似旧问题数为 5 时, RMSE 和 MAE 最小,因此在使用 GMNN 时,将相似旧问题数设置为 5.

在推荐回答者时,本文将问题中的回答所获得的财富值按照数量进行了排名统计,并将财富值排名前 N 名的回答占该问题下总体的比例进行了计算,结果见表 2.

表 2 财富值占比情况统计

Tab. 2 Statistics on the proportion of wealth value

Top N	1	2	3	4	5
财富值占比/%	55.5	79.4	84.3	93.1	94.8

从表 2 中可知,财富值排名最高的回答能够获得该问题下 55.5% 的财富值,财富值排名前五名的回答可以获得 94.8% 的财富值. 因此本文在对用户回答问题可能获得财富值进行排序时,选择前 5 名

为问题进行专家推荐.

对 ALS 的隐性向量长度进行分组实验,实验结果见表 3.

表 3 ALS 不同隐性向量长度分组实验结果

Tab.3 Experimental results of ALS different latent lengths

隐性向量长度	RMSE	MAE
20	0.27	0.24
70	0.23	0.18
140	0.24	0.21
200	0.23	0.19

从表 3 可知,隐性向量长度为 70 时, RMSE 和 MAE 最小,因此在使用 ALS 分析用户历史回答数据,对用户和问题隐性向量进行表示时,隐性向量长度设置为 70.

3.5 对比实验

3.5.1 不同模型之间的对比实验

为了验证本文提出 GMNN 算法的有效性,本文训练了 FM、DeepFM、Wide&Deep 模型作为对比模型,在 DeepFM 模型对目标用户的财富值预测时将提问者特征、问题文本 LDA 概率向量表示和回答者特征输入网络进行训练. 不同模型在测试集上的预测效果见表 4.

表 4 不同算法的模型效果对比

Tab. 4 Comparison of model effects under different algorithms

模型	FM	DeepFM	Wide&Deep	GMNN
RMSE	0.346	0.318	0.320	0.319
MAE	0.349	0.265	0.295	0.251
MRR	0.849	0.874	0.856	0.894
NDCG	0.951	0.964	0.959	0.968
RECALL	0.034	0.036	0.036	0.038

通过分析表 4 可知, GMNN 算法由于加入了记忆网络且考虑了用户的社交关系、偏好与能力,因此在化工问答社区数据上的表现优于其他三种流行的算法.

3.5.2 数据集划分对比实验

为了验证 GMNN 算法在海川化工数据集中不同时间数据的泛化性,本文采用交叉验证方法对算法进行验证. 海川化工数据集由化工论坛 2017 年 11 月 7 日至 2018 年 12 月 7 日的问答数据构成,本文分别

将6月~8月、8月~10月的问答数据作为测试集进行实验,实验结果如表5-6所示.

表5 6—8月不同算法的模型效果对比

Tab. 5 Comparison of model effects of different algorithms from June to August

模型	FM	DeepFM	Wide&Deep	GMNN
RMSE	0.332	0.316	0.328	0.320
MAE	0.256	0.235	0.315	0.212
MRR	0.831	0.910	0.849	0.940
NDCG	0.950	0.969	0.960	0.986
RECALL	0.033	0.031	0.034	0.038

表6 8—10月不同算法的模型效果对比

Tab. 6 Comparison of model effects of different algorithms from August to October

模型	FM	DeepFM	Wide&Deep	GMNN
RMSE	0.336	0.329	0.330	0.328
MAE	0.255	0.256	0.312	0.208
MRR	0.945	0.941	0.875	0.956
NDCG	0.971	0.978	0.965	0.980
RECALL	0.030	0.031	0.030	0.036

从表5、表6可以看出,GMNN算法对于海川化工数据集中不同时间的数据依然有很好的表现,说明该算法具有一定的泛化性.

3.6 消融实验

3.6.1 用户特征有效性分析

为了验证本文提取用户三部分特征的有效性,分别将用户的隐形向量特征 P ,用户认可度的特征矩阵 G 和用户个人特征矩阵使用自编码转化为低维稠密的个人特征矩阵 E ,三类特征进行不同的组合,输入GMNN模型进行消融实验,实验结果见图3.

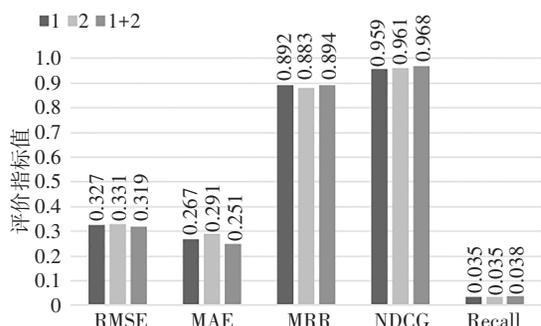


图3 不同用户特征组合下的模型效果对比

Fig.3 Comparison of model effects under different user feature combinations

通过对用户不同维度特征输入模型进行消融实验,由图3可以看出,融合了用户的隐形向量特征 P 、用户认可度的特征矩阵 G 和用户个人特征矩阵 E .三类特征的个人特征模型效果最好,验证了本文设计面向用户多维特征的联合表示子网络的有效性.

3.6.2 用户元路径有效性分析

为了验证本文提取两条元路径的有效性,本文分别将两条元路径生成的用户网络特征表示向量放入基于记忆的图神经网络模型进行消融实验,实验结果见图4.

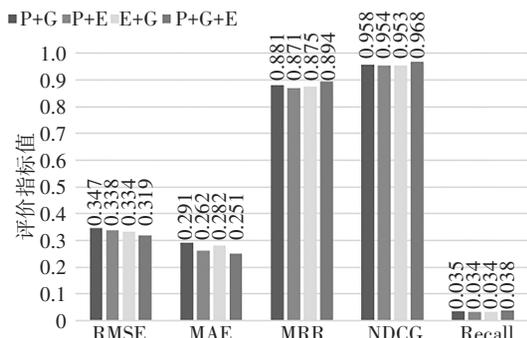


图4 不同元路径组合下的模型效果对比

Fig.4 Comparison of model effects under different meta-path

通过分析图4可知,将用户两条元路径共同生成的用户网络特征表示输入模型进行训练,比其他两条元路径单独生成的用户网络特征表示效果好.

3.6.3 注意力机制有效性分析

为了验证本文采用的注意力机制的有效性,在数据集上进行注意力机制消融实验,实验结果见图5.

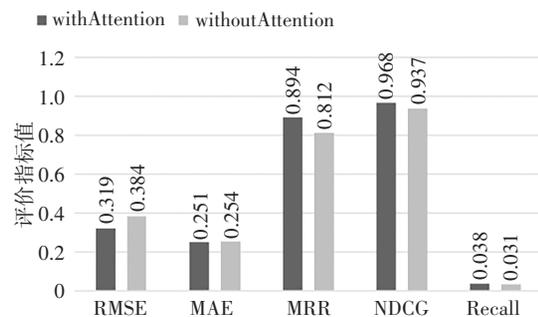


图5 引入注意力机制性能分析

Fig.5 Performance analysis of introducing attention mechanism

由图5可得,引入注意力机制的模型性能在数据集上有所提升.注意力机制可以解释为不同用户对待问题的不同视角,体现出不同用户对同一问题

的理解程度不同.说明了本文提出的基于注意力的新问题图网络表示学习的有效性.

4 结论

数据稀疏和冷启动是推荐系统目前面临的主要挑战,传统的专家推荐算法无法有效建模新问题,本文提出一种基于记忆的注意力图神经网络推荐方法,对用户进行多特征联合表示学习,有效缓解数据稀疏问题,利用记忆网络和注意力机制,从用户的不同视角表示新问题.通过在海川化工问答社区数据集上的相关实验,本模型的表现都优于同类模型,为解决数据稀疏和冷启动问题提供了新的思路.

随着推荐系统的发展,数据呈现多模态的发展趋势,后续研究将重点设计针对不同模态的信息的统一框架使推荐系统建模更加准确.

参考文献

- [1] PEDRO J S, KARATZOGLOU A. Question recommendation for collaborative question answering systems with RankSLDA[C]//RecSys '14: Proceedings of the 8th ACM Conference on Recommender systems. 2014:193-200.
- [2] MANDAL D P, KUNDU D, MAITI S. Finding experts in community question answering services: a theme based query likelihood language approach[C]//2015 International Conference on Advances in Computer Engineering and Applications. March 19-20, 2015, Ghaziabad, India. IEEE, 2015:423-427.
- [3] FAN W Q, MA Y, LI Q, *et al.* Graph neural networks for social recommendation[C]//WWW '19: The World Wide Web Conference. 2019:417-426.
- [4] 王根生,潘方正.融合加权异构信息网络的矩阵分解推荐算法[J].数据分析和知识发现,2020,4(12):76-84.
WANG G S, PAN F Z. Matrix factorization algorithm with weighted heterogeneous information network[J]. Data Analysis and Knowledge Discovery, 2020, 4(12):76-84. (In Chinese)
- [5] WESTON J E, SZLAM A D, FERGUS R D, *et al.* End-to-end memory networks:US10664744[P]. 2020-05-26.
- [6] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. Neural Computation, 1997, 9(8):1735-1780.
- [7] LI S, JOUPPI N P, FARABOSCHI P, *et al.* MEMORY NETWORK:WO 2014178856 A1[P].
- [8] MILLER A H, FISCH A J, DODGE J D, *et al.* Key-value memory networks:US20180357240[P]. 2018-12-13.
- [9] 何柔莹,徐建.基于注意力卷积神经网络的工作票专家推荐方法[J].南京理工大学学报,2019,43(1):13-21.
HE R Y, XU J. Expert recommendation for trouble tickets using attention-based CNN model[J]. Journal of Nanjing University of Science and Technology, 2019, 43(1):13-21. (In Chinese)
- [10] 吕晓琦,纪科,陈贞翔等.结合注意力与循环神经网络的专家推荐算法[J/OL].计算机科学与探索:1-12[2021-04-16].<http://kns.cnki.net/kcms/detail/11.5602.tp.20210324.1045.004.html>.
LV X Q, JI K, CHEN Z X, *et al.* Expert recommendation algorithm by combining attention and recurrent neural network[J/OL]. Journal of Frontiers of Computer Science & Technology:1-12[2021-04-16]. (In Chinese)
- [11] CHAUDHARI S, MITHAL V, POLATKAN G, *et al.* An attentive survey of attention models[J]. ACM Transactions on Intelligent Systems and Technology, 2021, 12(5):1-32.
- [12] DEVLIN J, CHANG M W, LEE K, *et al.* BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Stroudsburg: ACL, 2019: 4171-4186.
- [13] RENDLE S. Factorization machines[C]//2010 IEEE International Conference on Data Mining. December 13-17, 2010, Sydney, NSW, Australia. IEEE, 2010:995-1000.
- [14] GUO H F, TANG R M, YE Y M, *et al.* DeepFM: a factorization-machine based neural network for CTR prediction[C]//Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. August 19-26, 2017. Melbourne, Australia. California: International Joint Conferences on Artificial Intelligence Organization, 2017:1725-1731.
- [15] CHENG H T, KOC L, HARMSEN J, *et al.* Wide & deep learning for recommender systems[C]//DLRS 2016: Proceedings of the 1st Workshop on Deep Learning for Recommender Systems. 2016:7-10.