

文章编号 1000-5269(2023)05-0072-08

DOI:10.15958/j.cnki.gdxbzrb.2023.05.09

基于兴趣和专业度建模的CQA专家发现方法

丁邱^{1,2}, 严馨^{*1,2}, 刘艳超³, 徐广义⁴, 邓忠莹¹

(1. 昆明理工大学信息工程与自动化学院, 云南昆明 650500; 2. 昆明理工大学云南省人工智能重点实验室, 云南昆明 650500; 3. 湖北工程学院信息技术中心, 湖北孝感 432000; 4. 云南南天电子信息产业股份有限公司, 云南昆明 650040)

摘要: 现有问答社区专家发现方法通过学习用户解答的问题序列单向信息建模用户兴趣, 忽略了用户兴趣的波动性, 对于解答过较少问题的用户建模准确度将受到影响, 此外, 未考虑历史回答与问题的语义相关性对评估用户表现的作用。论文提出基于兴趣和专业度建模的CQA专家发现方法, 首先, 使用BERT4Rec学习用户近期解答的问题序列双向信息得到近期动态兴趣表示; 其次, 构建用户社交网络, 使用DeepWalk算法学习网络结构特征, 得到用户长期兴趣表示; 再次, 构建用户专业度评估网络, 依据用户回答与问题的语义相关性及反馈信息计算权重, 对相应问题进行加权, 引入注意力机制, 重点关注用户在与新问题相近问题上的表现, 得到用户专业度表示; 最后, 综合用户近期动态兴趣、长期兴趣和专业度表示与新问题进行匹配打分, 为新问题找出有意愿接受邀请并能提供优质回答的用户。实验表明, 该方法取得了较好表现, 较基线方法在英语、3D打印和天涯问答数据集的MRR评价指标上分别提升了5.2%、2.7%、16.1%。

关键词: 问答社区; 专家发现; 动态兴趣建模; 社交网络; 专业度建模

中图分类号: TP391 **文献标志码:** A

问答社区(question answering community, CQA)专家发现就是找出潜在的能够提供高质量答案的用户, 邀请他们回答问题。缩短提问者等待回复的时间, 促进用户参与度, 更快获取到用户感兴趣并有可能提供答案的问题^[1]。面向问答社区的专家发现方法, 包括基于图、基于内容和基于深度学习的方法。

基于图的方法是将用户之间建立关系图, 把他们关联起来。YANG等^[2]应用社交网络构建任务和用户声望图, 基于图的算法衡量用户的专业知识和任务声望。SUN等^[3]提出应用专业知识增益假设来缓解数据稀疏性问题, 并从用户的历史活动中构建竞争图, 将竞争图的层次结构解释为问题难度和用户专业知识。该类方法本质上是依据用户是否在相似领域活跃过, 并没有系统地对用户专业性和兴趣进行建模。

基于内容的方法是将其看作一个主题建模问题。通过用户、问题等信息, 找到能回答对应主题的专家。YANG等^[4]从用户偏好主题和拥有的专业知识寻找对应专家。ZHAO^[5]提出主题级的专

家学习框架, 综合链接分析与内容分析寻找合适的专家。MUMTAZ等^[6]利用问题相似度找到相关用户, 再利用反馈信息对相关用户进行排序。利用主题建模思想找出对应主题专家的方法, 使用关键词简单表征文本特征, 忽略了问题和答案文本的深层特征。

基于深度学习的方法是引入深度学习提升专家发现方法性能。HE等^[7]设计基于神经网络的协作过滤通用框架, 将矩阵因式分解与前馈神经网络进行联合学习, 建模用户和项目的潜在特征。SEDHAIN等^[8]提出基于自动编码器范式的协同过滤模型, 将传统的线性内积替换为自动编码器中评级矩阵的非线性分解。LIAN等^[9]提出新的压缩交互网络自动学习高阶特征交互, 将其与深度神经网络结合应用于推荐任务。HE等^[10]利用卷积神经网络以分层方式从局部到全局学习嵌入维度之间的高阶相关性, 找出潜在的回答者。LI等^[11]通过异质信息网络学习问题、提出者和回答者表示, 用卷积打分函数找出高分用户作为专家。TANG等^[12]提出基于注意力的因子分解机变体, 不仅模

收稿日期: 2023-01-12

基金项目: 国家自然科学基金资助项目(61562049, 61462055)

作者简介: 丁邱(1997—), 女, 在读硕士, 研究方向: 自然语言处理, E-mail: 1452229049@qq.com.

* 通讯作者: 严馨, E-mail: kg_yanxin@sina.com.

拟个体特征对之间的交互,而且强调关键特征的作用和成对的交互。尤丽等^[13]将答案质量作为推荐因子之一,寻找专家用户。YIN 等^[14]考虑社区动态性,提出适应用户兴趣漂移进行推荐的方法。吕晓琦等^[15]根据用户历史解答序列学习用户动态兴趣,依据用户兴趣推荐专家。HE 等^[16]提出动态用户建模方法,考虑用户兴趣的动态性,利用反馈信息获取用户的专业度。

若想找出有极大意愿接受邀请且拥有相应专业知识的用户,就要综合考虑用户兴趣和专业度,但现有方法要么只考虑了一个方面,要么对兴趣和专业度的建模还不够充分,有待改进。从用户解答过的问题能判断用户对哪些方向领域感兴趣,在问答社区专家发现任务中,往往将用户解答过的问题序列称为用户行为序列。目前,大多数方法利用循环神经网络学习用户兴趣表示,只实现了从左到右学习序列的单向特征,忽略了用户行为序列更丰富的信息。有些用户解答过较少的问题行为序列较为稀疏,导致仅从行为序列对用户进行建模的方法性能受到极大影响,而用户社交网络蕴含丰富信息,将社交网络融入用户建模中,能够聚合与用户有相同兴趣的其他用户特征增强用户兴趣表征,克服数据稀疏造成的影响。现有方法多利用回答收到的反馈信息来判断用户的专业度,有些冷门问题其受众较少,回答收到的反馈信息较少甚至没有,但不能因此就判定用户不够专业。一个好的回答往往与问题有较高的语义相关性,因此,利用回答与问题的语义相关性能判断用户在对应问题上表现是否专业,克服反馈信息缺乏时对用户评估造成的影响。关注用户在与新问题相近那些问题上的表现,能够评估用户在新问题领域的专业度。本文提出基于兴趣和专业度建模的 CQA 专家发现方法主要贡献如下:

使用基于变换网络的双向编码进行推荐的模型^[17](sequential recommendation with bidirectional encoder representations from transformer, BERT4Rec)学习用户近期解答的问题序列双向信息,克服循环神经网络仅学习序列单向信息的局限性,得到更充分全面的用户近期动态兴趣,有利于找出有极大意愿接受邀请的用户。

构建用户社交网络,使用深度游走算法^[18](online learning of social representations, Deepwalk)学习网络结构特征,考虑用户在社区的社交关系,

聚合与用户兴趣相似的邻接用户特征得到用户长期兴趣表示。融合网络表示学习对用户进行建模,有利于解决用户行为数据稀疏问题,即对那些解答过较少问题的用户,也能通过蕴含丰富信息的社交网络对其进行建模。

构建用户专业度评估网络,引入用户回答与问题的语义相关性结合反馈信息判断用户在对应问题上表现是否专业,克服反馈信息缺乏造成的影响。引入注意力机制重点关注用户在与新问题相近问题上的表现情况,得到更为准确的用户专业度表示。

在英语、3D 打印和天涯问答三个数据集上进行实验,本文方法皆取得了较好表现,为新问题找出了有意愿接受邀请的潜在优质回答者。

1 模型构建

1.1 问题分析

使用 $U = \{u_1, u_2, \dots, u_l\}$ 表示用户集,用户 u 解答过的历史问答对序列表示为:

$Q(u) = \{(q_1^u, a_1^u, t_1^u, y_1^u), (q_2^u, a_2^u, t_2^u, y_2^u), \dots, (q_{|Q(u)|}^u, a_{|Q(u)|}^u, t_{|Q(u)|}^u, y_{|Q(u)|}^u)\}$ 。其中, $(q_1^u, a_1^u, t_1^u, y_1^u)$ 表示用户 u 在 t_1 时刻回答了问题 q_1 , 所给答案为 a_1 , y_1 为答案 a_1 收到的反馈信息。问答对序列是按用户回答问题的时间排列的 $t_i < t_j (i < j)$, $|Q(u)|$ 是用户 u 回答过的问题数量。

专家发现任务就是给出一个新问题 q_m , 候选用户集 U 以及每个候选用户 u 解答过的历史问答对序列 $Q(u)$ 。计算每个用户 $u \in U$ 与 q_m 的匹配分数,最高分的用户就是有意愿接受邀请,可以为新问题提供较优答案的专家用户。

1.2 模型框架

模型的总体框架如图 1 所示。用户近期动态兴趣建模是通过用户近期解答的问题序列,捕捉用户动态兴趣,判断用户是否有意愿接受邀请;用户社交网络建模是利用用户间的交互,挖掘用户长期不易发生变化的长期兴趣,提升用户建模准确度;用户专业度建模是综合用户历史表现情况,评估用户专业度,判断用户是否有能力回答问题。综合用户近期动态兴趣、长期兴趣和专业度得到用户最终表示,将用户与新问题进行匹配打分。得分较高的用户将是有意愿接受邀请,能够为新问题提供较优答案的专家用户。

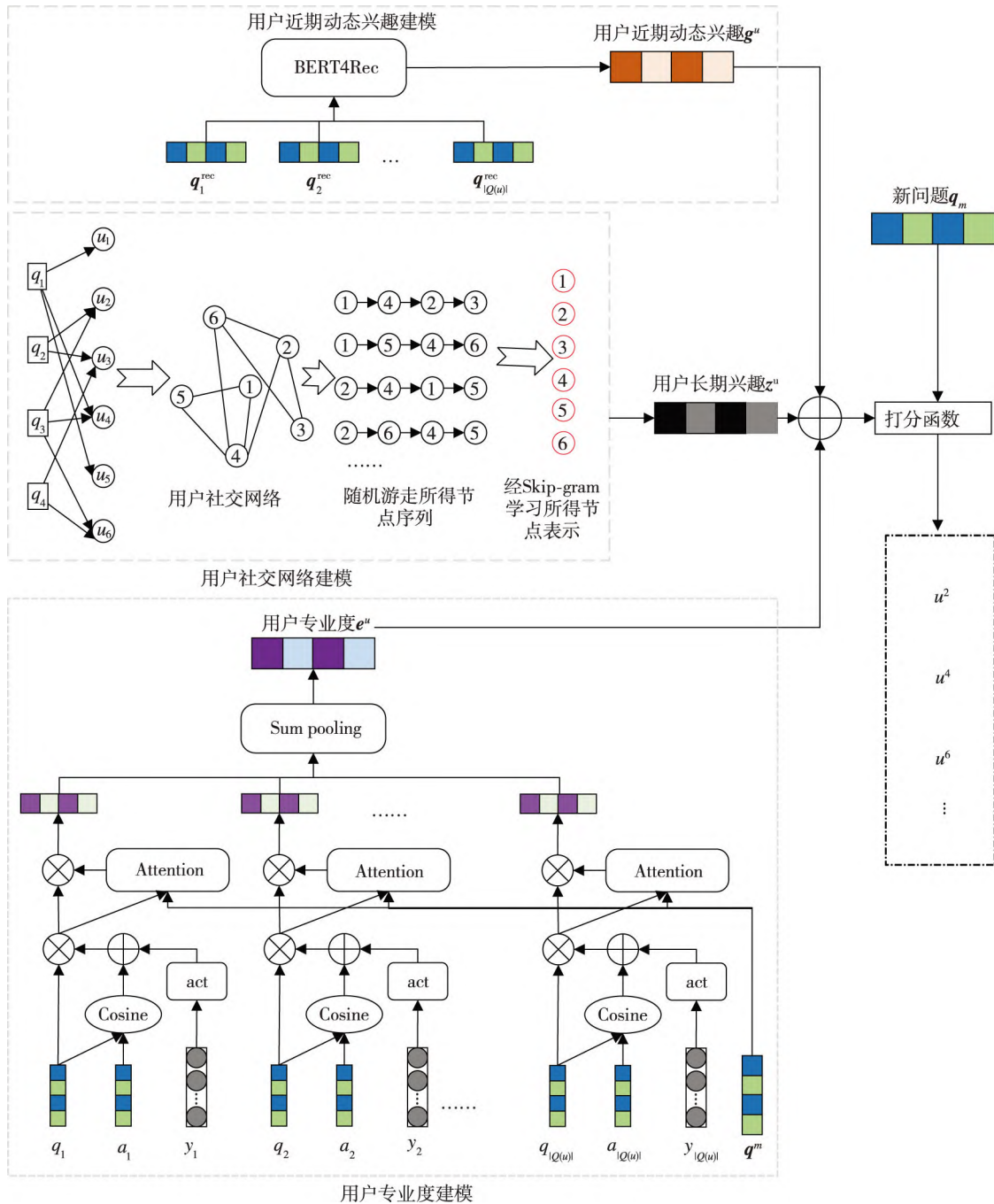


图 1 模型总体框架图

Fig. 1 Overall framework of the model

1.2.1 问题和答案编码

使用预训练的基于变换网络的双向编码^[19] (bidirectional encoder representation from transformer, BERT)模型对问题和答案进行编码,学习问题和答案的深层特征信息,得到问题和答案表示。问答社区中问题文本长度较短,一般就是 0 ~ 100 区间。答案文本长度相对要长一点,但大多数都是 100 ~ 200 区间,符合预训练 BERT 模型能处理的

最大长度限制。个别答案会超出 512 的长度限制,据统计超出限制的答案前 200 个字符已经可以概括这个答案的主要思想,因此,若是遇到超出 BERT 长度限制的答案,就将其做截断处理,保证输入文本符合 BERT 的长度限制。

1.2.2 用户近期动态兴趣建模

用户的近期解答的问题序列能够表征其动态的近期兴趣,这将是决定用户目前是否愿意回答新

问题的一个关键因素。现有专家发现方法广泛用于处理序列数据的循环神经网络,获取动态兴趣表示,但这种方法只能单向的学习序列特征,不能同时学习序列的双向特征。用户兴趣是波动的,并不是单向发展的,需要利用其近期解答的问题序列充分捕捉到兴趣来回波动的情况。受文献[17]启发中,使用 BERT4Rec 代替循环神经网络对用户近期解答的问题序列进行处理,从而克服循环神经网络没能有效利用序列双向信息的缺点。BERT4Rec 允许用户历史行为中的每一项融合来自左侧和右侧的信息,充分学习行为序列的双向信息,获取到更为深层全面的用户动态兴趣表示。将用户 u 在新问题 q_m 提出之前回答的 n 个问题序列 $\{q_1^{\text{rec}}, q_2^{\text{rec}}, \dots, q_n^{\text{rec}}\}$ 作为其近期行为,将编码后的问题序列输入 BERT4Rec,获取用户近期动态兴趣表示。图 2 为使用 BERT4Rec 进行动态兴趣建模的框架图。

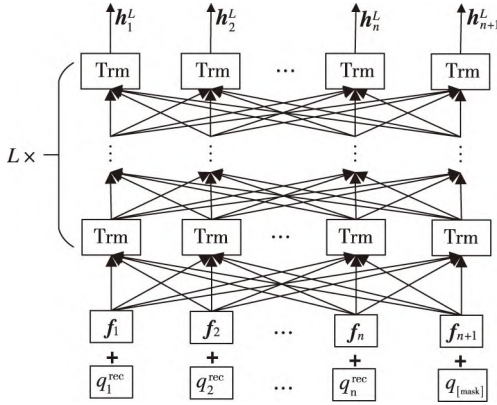


图 2 用户近期动态兴趣建模框架图

Fig. 2 Framework diagram of user's recent dynamic interest modeling

BERT4Rec 本质就是由多个双向 Transformer^[20]层堆叠起来的,为了使用输入的序列信息,需要将位置向量与初始输入求和作为 Transformer 的输入。计算见公式(1)。

$$h_i^0 = q_i^{\text{rec}} + f_i \quad (1)$$

式中, $q_i^{\text{rec}} \in \mathbf{R}^d$ 是输入序列的第 i 项, $f_i \in \mathbf{R}^d$ 是位置索引 d 维的位置向量, BERT4Rec 使用可学习的位置向量。将加入位置信息的输入序列送入多层双向 Transformer 层,迭代计算每一项在每一层的隐藏表示 $h_i^l \in \mathbf{R}^d$, l 指第 l 个 Transformer 层。

Trm 包含多头自注意力和位置前馈网络两个子层,使用这两个子层对输入序列的隐藏矩阵 $H \in \mathbf{R}^{n+1 \times d}$ 进行处理。在预测用户未来对什么感兴趣时, BERT4Rec 在输入序列的末尾附加了

masked,根据该 masked 的最终隐藏表示来预测下一项,也就是输入序列实际长度为 $n+1$, d 为隐藏维度。

$$\begin{aligned} M_{\text{MH}}(H^l) &= [\eta_1; \eta_2; \dots; \eta_h] W^0 \\ \eta_i &= A_{\text{Attention}}(H^l W_i^Q, H^l W_i^K, H^l W_i^V) \\ A_{\text{Attention}}(Q, K, V) &= S_{\text{softmax}}\left(\frac{QK^T}{\sqrt{d/h}}\right)V \end{aligned} \quad (2)$$

式中, H^l 为隐藏矩阵; l 表示第 l 个 Transformer 层; h 为自注意力的头数; $W^0 \in \mathbf{R}^{d \times d}$, $W_i^Q \in \mathbf{R}^{d \times d/h}$, $W_i^K \in \mathbf{R}^{d \times d/h}$ 和 $W_i^V \in \mathbf{R}^{d \times d/h}$ 为可学习参数。

位置前馈网络子层计算如下:

$$\begin{aligned} P_{\text{FFN}}(H^l) &= [F_{\text{FFN}}(h_1^l)^T; \dots; F_{\text{FFN}}(h_n^l)^T]^T \\ F_{\text{FFN}}(x) &= G_{\text{GELU}}(xW^{(1)} + b^{(1)})W^{(2)} + b^{(2)} \\ G_{\text{GELU}}(x) &= x\Phi(x) \end{aligned} \quad (3)$$

其中, $\Phi(x)$ 是标准高斯分布的累积分布函数; $W^{(1)} \in \mathbf{R}^{d \times d}$, $W^{(2)} \in \mathbf{R}^{d \times d}$, $b^{(1)} \in \mathbf{R}^d$, $b^{(2)} \in \mathbf{R}^d$ 是可学习参数。BERT4Rec 使用层规范化函数 L_{LN} 对同一层所有隐藏单元的输入进行规范化,具体计算如下:

$$\begin{aligned} H^l &= T_{\text{Trm}}(H^{l-1}), \forall l \in [1, \dots, L] \\ T_{\text{Trm}}(H^{l-1}) &= L_{\text{LN}}(A^{l-1} + D_{\text{Dropout}}(P_{\text{FFN}}(A^{l-1}))) \\ A^{l-1} &= L_{\text{LN}}(H^{l-1} + D_{\text{Dropout}}(M_{\text{MH}}(H^{l-1}))) \end{aligned} \quad (4)$$

式中, $H^{l-1} \in \mathbf{R}^{n+1 \times d}$ 是第 $l-1$ 层输出的隐藏矩阵,作为第 l 层 Trm 的输入,经过多头自注意力和位置前馈网络两个子层后输出 $H^l \in \mathbf{R}^{n+1 \times d}$ 。将最后一层 Transformer 层的最后一个隐藏表示 $h_{n+1}^L \in \mathbf{R}^d$ 作为用户的近期动态兴趣 $g^u = h_{n+1}^L$ 。

1.2.3 用户社交网络建模

问答社区中回答相同问题的用户往往兴趣相似,若是能够将用户间的社交关系融入用户建模中,将能有效改善行为数据稀疏问题对建模造成的影响,通过挖掘用户间的关系得到更为精准的用户画像。构建用户社交网络 $G = (U, E)$,用户作为网络节点,用户间的关系作为边,两个用户若是回答过相同问题,这两个用户节点之间就有边。 U 是用户节点集, E 是边集。

运用经典网络表示学习方法 DeepWalk^[18]学习用户社交网络节点表示,即通过挖掘用户间的关系,聚合与用户兴趣相似的邻接用户特征增强用户节点表征,得到用户稳定的长期兴趣表示。首先,使用随机游走得到社交网络对应节点序列集;然后,利用 Skip-gram 算法学习得到用户节点表示,采用负采样方法进行优化,目标函数见公式(5)。

$$\log p(j|i) = \log \sigma(\mathbf{u}_i^T \mathbf{u}_j) + \sum_{l=1}^k E_{j' \sim p(j')} [\log \sigma(-\mathbf{u}_i^T \mathbf{u}_{j'})] \quad (5)$$

式中,第一项为正例,即确实互为邻接的节点, $\mathbf{u}_i \in \mathbf{R}^d$ 为节点 i 的中心向量, $\mathbf{u}_j \in \mathbf{R}^d$ 为节点 j 的上下文向量;第二项为负例, k 为负例的数目, j' 是从预先定义的噪声分布 $p(j')$ 采样的负例节点。

最后,将经过优化增强后的用户节点的中心向量作为用户长期兴趣表示 $\mathbf{z}^u \in \mathbf{R}^d$ 。

1.2.4 用户专业度建模

用户所给回答收到的反馈信息是评估用户表现的重要指标,但不同社区拥有的反馈信息会有所不同,甚至有些受众较少的问题没有反馈信息,过度依赖反馈信息会降低模型通用性。为了提升模型通用性和专业度评估准确度,构建用户专业度评估网络,获取用户专业度表示。

受答案选择方法^[21]启发,一个优质回答往往与问题有较高的语义相关性,故使用回答与问题的语义相关性及回答收到的反馈信息计算用户在相应问题上的表现权重,对用户回答的对应问题进行加权,即给那些用户表现较好的问题更高的权重。计算见公式(6)。

$$\mathbf{p}_i = (C_{\cosine}(\mathbf{q}_i, \mathbf{a}_i) + A_{act}(\mathbf{y}_i)) \cdot \mathbf{q}_i \quad (6)$$

其中, $C_{\cosine}()$ 是余弦相似度函数; $\mathbf{q}_i$ 和 $\mathbf{a}_i$ 分别是问题和答案表示; $A_{act}()$ 是一个前馈网络; $\mathbf{y}_i$ 是反馈向量。对于没有反馈信息的回答将其反馈权重设为 0。使用的反馈信息见表 1。

表 1 反馈信息

Tab. 1 Feedback information

统计信息	含义
点赞数	答案收到的点赞数
比率	答案点赞数与对应问题所有答案点赞数总和的比值
最佳答案	答案是否为最佳答案

用户在与新问题相近的那些问题上的表现,最能体现用户在新问题领域专业水平如何。因此,引入注意力机制,给予那些与新问题相近的问题更多关注。注意力权重计算见公式(7)。使用注意力给予用户解答的与新问题相近的问题更高的权重后,对加权后的问题表示序列求和得到用户专业度表示 $\mathbf{e}^u \in \mathbf{R}^d$,计算见公式(8)。

$$\beta_i = \frac{\exp(\mathbf{p}_i^T \mathbf{W}_e \mathbf{q}_m)}{\sum_{j=1}^{|Q(u)|} \exp(\mathbf{p}_j^T \mathbf{W}_e \mathbf{q}_m)} \quad (7)$$

$$\mathbf{e}^u = \sum_{i=1}^{|Q(u)|} \beta_i \mathbf{p}_i \quad (8)$$

式中, $\mathbf{W}_e \in \mathbf{R}^{d \times d}$ 是可训练的参数; $\mathbf{p}_i \in \mathbf{R}^d$ 是问题表示经语义相关性结合反馈权重对其加权后的表示; β_i 为注意力权重。

1.2.5 用户问题匹配

经过用户近期动态兴趣、社交网络建模和专业度建模,获得用户近期动态兴趣表示 $\mathbf{g}^u \in \mathbf{R}^d$ 、长期兴趣表示 $\mathbf{z}^u \in \mathbf{R}^d$ 和专业度表示 $\mathbf{e}^u \in \mathbf{R}^d$ 。将三者级联经过线性变换得到用户最终表示 $\mathbf{u}$,计算见公式(9)。将用户与新问题进行匹配打分,计算见公式(10)。

$$\mathbf{u} = \mathbf{W}_p[\mathbf{g}^u, \mathbf{z}^u, \mathbf{e}^u] + \mathbf{b}_p \quad (9)$$

$$S(\mathbf{q}_m, \mathbf{u}) = \mathbf{q}_m \cdot \mathbf{u} \quad (10)$$

其中, $\mathbf{W}_p \in \mathbf{R}^{(2d+d1) \times d}$; $\mathbf{b}_p \in \mathbf{R}^d$ 是可训练的变换参数。按得分对候选用户进行排序,选择排名最前用户作为推荐人选。使用 hinge 损失函数对模型进行训练。

$$l_{\text{loss}} = \frac{1}{N} \sum_{i=1}^N \max(0, \lambda - S(\mathbf{q}_m, \mathbf{u}_+) + S(\mathbf{q}_m, \mathbf{u}_-)) \quad (11)$$

式中, N 为样本数; λ 是一个常数; $S(\mathbf{q}_m, \mathbf{u}_+)$ 、 $S(\mathbf{q}_m, \mathbf{u}_-)$ 分别表示正例用户和负例用户的得分。正例用户就是能够给问题提供较优答案的用户,反之其他用户则为负例用户。

2 实验与结果分析

2.1 数据集

Stack Exchange 包含许多特定领域的子数据集,如数学、英语和 3D 打印等。本文选用英语和 3D 打印这两个子数据集进行实验。数据集包含 2011 年到 2021 年的问题、答案和用户,以及问题提出时间、回答时间、回答得分、最佳答案等信息。为了验证方法的通用性,我们搜集了天涯问答社区的数据进行实验。天涯问答数据集中的问题和答案,有些是没有点赞量这类反馈信息的,将这部分数据也用于模型测试,进一步验证模型通用性。选取候选用户时,在新问题提出的近期没有活跃的用户则不将其作为候选用户,保证邀请的用户还在活跃。表 2 为这三个数据集预处理后用于实验的概况。

表2 数据集概况

Tab.2 Dataset overview

数据集	问题数/个	答案数/个	用户数/个
英语	11 433	21 674	6 702
3D 打印	3 626	6 706	6 702
天涯问答	14 134	124 042	3 416

2.2 参数设置

实验基于 Pytorch 深度学习框架实现,使用 Adam 优化器,学习率为 0.001。批次大小设置为 32,设置了 8 个 epoch。使用开源 BERT-base 预训练模型对问题和答案编码,编码所得的问题和答案表示维度为 $d = 768$ 。用户近期动态兴趣建模部分将 BERT4Rec 隐藏表示和位置嵌入维度设为 768,使用 2 层 Transformer 层,多头注意力的数量设为 2 头,序列最大长度为 31,即 $n = 30$ 取在新问题提出之前用户最近回答的 30 个问题作为其近期行为。社交网络学习的节点维度设为 $d_1 = 256$ 。hinge 损失函数中的 λ 设置为 0.2。将最佳答案的提供者作为正例用户,其余用户作为负例用户。

2.3 基线

将 EDQAU 与近几年的专家发现方法在英语、3D 打印和天涯问答数据集上进行对比实验:

ConvNCF^[10]:该方法基于矩阵分解的多层神经网络结构。使用卷积神经网络 (Convolutional Neural Networks, CNN) 学习用户和项目嵌入之间的交互, CNN 以多层结构从局部到全局学习嵌入维度之间的高阶相关性。

Expert2Vec^[6]:该方法基于嵌入方法计算问题相似度,利用相似度为给定问题找到相关用户,最后利用社区反馈对相关用户进行排序。

NeRank^[11]:使用异质网络嵌入算法学习问题、提问者 and 回答者三者的嵌入表示,再将三者作

为卷积打分函数的输入,根据打分结果得到推荐列表。

DUM^[16]:考虑社区动态性,将用户兴趣分解为长期兴趣和动态近期兴趣。利用反馈信息获取用户专业度表示,根据用户兴趣和用户专业度对用户进行排序。

2.4 评价指标

采用平均倒数排名 (mean reciprocal rank, MRR)、Hit@5 和 Prec@1 作为评价指标对模型性能进行评估。

MRR 是标准回答者排名倒数的均值。假设用于测试的问题集为 Q ,那么 MRR 的计算为公式 (12)。

$$M_{MRR} = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{\mu_{i,best}} \quad (12)$$

式中, $|Q|$ 表示用于测试的问题数量; $\mu_{i,best}$ 表示第 i 个问题的标准回答者的排序位置。

Hit@ k 主要是看标准回答者是否出现在推荐列表的前 k 位。计算见公式 (13)。

$$H_{Hit@k} = \frac{1}{|Q|} \sum_{i=1}^{|Q|} I(\mu_{i,best} \leq k) \quad (13)$$

式中, $I(\mu_{i,best} \leq k)$ 表示第 i 个问题的标准回答者是否在推荐列表的前 k 位,在则为 1,不在为 0。

Prec@1 则是看标准回答者是否出现在推荐列表的第 1 位。计算见公式 (14)。

$$P_{Prec@1} = \frac{1}{|Q|} \sum_{i=1}^{|Q|} I(\mu_{i,best} = 1) \quad (14)$$

2.5 实验结果分析

从表 3 不同方法的对比实验结果可见,较考虑了用户动态兴趣的基线方法 DUM, EDQAU 方法在英语、3D 打印和天涯问答数据集的 MRR 评价指标上分别提升了 5.2%、2.7%、16.1%,性能和通用性都有所提升。

表3 不同方法性能比较结果

Tab.3 Performance comparison results of different methods

实验模型	英语			3d 打印			天涯问答		
	M_{MRR}	$H_{Hit@5}$	$P_{Prec@1}$	M_{MRR}	$H_{Hit@5}$	$P_{Prec@1}$	M_{MRR}	$H_{Hit@5}$	$P_{Prec@1}$
ConvNCF	0.417	0.658	0.245	0.392	0.635	0.204	0.339	0.581	0.119
Expert2Vec	0.372	0.619	0.202	0.365	0.602	0.188	0.237	0.396	0.105
NeRank	0.516	0.713	0.332	0.497	0.696	0.321	0.462	0.624	0.276
DUM	0.572	0.802	0.392	0.564	0.796	0.382	0.424	0.652	0.267
EDQAU	0.624	0.930	0.410	0.591	0.890	0.390	0.585	0.840	0.430

本文 EDQAU 方法使用 BERT 模型对问题和答案进行编码,所得问题和答案的嵌入表示捕捉了

文本上下文信息以及更深层次的语义特征。ConvNCF 和 Expert2Vec 仅是简单的使用问题 ID 或是

关键字作为嵌入表示,忽略了文本的丰富语义信息和上下文间的关系。NeRank 和 DUM 虽然使用了网络嵌入算法学习文本的嵌入表示,但其方法还是存在局限,没能充分捕获到问题、答案这类短文本的深层语义信息,影响了模型的整体性能。EDQAU 与 ConvNCF、Expert2Vec 和 NeRank 这几个没有考虑用户动态兴趣的方法相比,性能得到较大的提升,因为依据用户近期行为建模用户动态兴趣,一定程度上保证了用户有意愿接受邀请。DUM 使用长短期记忆网络(long short-term memory, LSTM)对用户动态兴趣建模,仅学习了序列的单向信息,信息捕捉不够充分,使用 BERT4Rec 对用户动态兴趣建模克服了 LSTM 这一缺陷。此外,我们构建用户社交网络,通过网络结构利用兴趣相似的邻接用户特征增强目标用户表示,对于行为数据稀疏的用户也能挖掘到其稳定的长期兴趣,依据用户历史表现构建专业度评估网络,评估用户专业度,保证找到的用户有能力回答问题,使模型性能有所改善。

DUM 和 Expert2Vec 在天涯问答数据集上表现较在英语和 3D 打印上表现波动要大于其他方法。因为天涯问答数据集拥有的反馈信息与另外两个数据集存在差异,甚至是有的用户、问题没有反馈信息,这就导致过度依赖反馈信息的方法在该数据集上表现受到极大影响。本文方法在三个数据集上性能波动较小,较基线方法通用性有所提升。

2.6 消融实验

为了验证用户近期动态兴趣建模、社交网络建模和专业度建模的重要性,进行了消融实验。表 4 为消融实验结果,其中 EDQAU-D 表示在本文方法基础上去除用户近期动态兴趣建模;EDQAU-S 表示去除社交网络建模;EDQAU-E 表示在本文方法基础上去除用户专业度建模,上述方法均通过英语数据集进行实验。

表 4 模型简化实验结果

Tab. 4 Experimental results of model simplification

实验模型	M_{MRR}	$H_{Hit@5}$	$P_{Prec@1}$
EDQAU	0.624	0.930	0.410
EDQAU-D	0.538	0.630	0.341
EDQAU-S	0.512	0.607	0.302
EDQAU-E	0.509	0.612	0.313

无论去除哪一个建模部分都对模型最终表现产生了影响,在 MRR 评价指标上分别降低了

8.6%、11.2% 和 11.5%。用户有时对生活常识类感兴趣,有时对旅游类感兴趣。用户最近对哪个方面感兴趣是很重要的,它决定了用户有多大的意愿会接受社区的邀请,为推送给他的新问题提供答案。EDQAU-D 模型去除用户近期动态兴趣建模,导致模型未能察觉到用户兴趣发生了偏移。这些兴趣发生偏移的用户极有可能会拒绝社区邀请,从而对模型的性能造成一定影响。EDQAU-S 模型去除社交网络建模,导致模型忽略了用户社区交互,而社交网络富含丰富信息,有利于解决行为数据稀疏对用户建模造成的影响,通过网络结构充分挖掘用户长期兴趣表示,提升用户建模准确度。EDQAU-E 模型去除用户专业度建模,只考虑用户对新问题是否感兴趣,但有意愿接受邀请的用户不一定可以提供优质的回答,这使得找到的用户有意愿接受邀请,但在新问题领域不一定拥有相应的专业知识,因此,降低了模型性能。

3 总结

针对现有方法往往局限于某个视角对用户进行建模,导致用户建模不够充分、准确的问题,提出基于兴趣和专业度建模的 CQA 专家发现方法。使用 BERT4Rec 克服传统的循环神经网络只能学习序列单向信息的缺陷,对用户近期行为进行建模得到近期动态兴趣表示;构建用户社交网络,使用 DeepWalk 算法学习网络结构特征,融合目标用户邻接用户特征得到用户长期兴趣表示;构建用户专业度评估网络,根据用户在历史问答对序列中的表现得到用户专业度表示。从多个视角对用户进行建模,得到更为精准的用户画像。在三个数据集进行实验表明,本文方法在专家发现任务中取得较好的效果。

在下一步研究工作中,考虑对问题和答案编码进行改良,尝试不同的编码方法,获取问题和答案文本更深层、精准的语义信息。问题往往要比答案短,直接使用相同的编码方式对问题和答案进行编码,求得两者的相似度可能还不够准确。改良编码问题将能进一步提升模型专业度评估的准确度,使模型性能得到进一步提升。

参考文献:

- [1] WANG X, HUANG C, YAO L, et al. A survey on expert recommendation in community question answering [J]. Journal of Computer Science and Technology, 2018, 33(4): 625-653.

- [2] YANG J, ADAMIC L A, ACKERMAN M S. Competing to share expertise: the taskcn knowledge sharing community [C]//ICWSM 2008. Washington: ICWSM, 2008: 1-8.
- [3] SUN J, MOOSAVI S, RAMNATH R, et al. QDEE: question difficulty and expertise estimation in community question answering sites [C]//Twelfth International AAAI Conference on Web and Social Media. California: ICWSM, 2018: 375-382.
- [4] YANG L, QIU M, GOTTIPATI S, et al. Cqarank: jointly model topics and expertise in community question answering [C]//Proceedings of the 22nd ACM international conference on Information & Knowledge Management. San Francisco California: ACM, 2013: 99-108.
- [5] ZHAO T, BIAN N, LI C, et al. Topic-level expert modeling in community question answering [C]//Proceedings of the 2013 SIAM International Conference on Data Mining. Austin Texas: Society for Industrial and Applied Mathematics, 2013: 776-784.
- [6] MUMTAZ S, RODRIGUEZ C, BENATALLAH B. Expert2vec: experts representation in community question answering for question routing [C]//Advanced Information Systems Engineering: 31st International Conference, CAiSE 2019. Rome Italy: Springer International Publishing, 2019: 213-229.
- [7] HE X, LIAO L, ZHANG H, et al. Neural collaborative filtering [C]//Proceedings of the 26th international conference on world wide web. Perth: ACM, 2017: 173-182.
- [8] SEDHAIN S, MENON A K, SANNER S, et al. Autorec: autoencoders meet collaborative filtering [C]//Proceedings of the 24th international conference on World Wide Web. Florence: ACM, 2015: 111-112.
- [9] LIAN J, ZHOU X, ZHANG F, et al. Xdeepfm: combining explicit and implicit feature interactions for recommender systems [C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. London: ACM, 2018: 1754-1763.
- [10] HE X, DU X, WANG X, et al. Outer product-based neural collaborative filtering [C]//Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence Main track. Stockholm: IJCAI, 2018: 2227-2233.
- [11] LI Z, JIANG J Y, SUN Y, et al. Personalized question routing via heterogeneous network embedding [C]//Proceedings of the AAAI Conference on Artificial Intelligence. Honolulu: AAAI, 2019, 33(1): 192-199.
- [12] TANG W, LU T, LI D, et al. Hierarchical attentional factorization machines for expert recommendation in community question answering [J]. IEEE Access, 2020, 8: 35331-35343.
- [13] 尤丽. 融合答案质量评估的问答社区专家用户推荐方法研究 [D]. 上海: 上海财经大学, 2020.
- [14] YIN H, ZHOU X, CUI B, et al. Adapting to user interest drift for poi recommendation [J]. IEEE Transactions on Knowledge and Data Engineering, 2016, 28(10): 2566-2581.
- [15] 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]. 计算机科学与探索, 2022, 16(9): 2068-2077.
- [16] HE T, GUO C, CHU Y, et al. Dynamic user modeling for expert recommendation in community question answering [J]. Journal of Intelligent & Fuzzy Systems, 2020, 39(5): 7281-7292.
- [17] SUN F, LIU J, WU J, et al. BERT4Rec: sequential recommendation with bidirectional encoder representations from transformer [C]//Proceedings of the 28th ACM international conference on information and knowledge management. Beijing: ACM, 2019: 1441-1450.
- [18] PEROZZI B, AL-RFOU R, SKIENA S. Deepwalk: online learning of social representations [C]//Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. New York: ACM, 2014: 701-710.
- [19] DEVLIN J, CHANG M W, LEE K, et al. Bert: pre-training of deep bidirectional transformers for language understanding [C]//Proceedings of naacL-HLT. Minneapolis: NAACL, 2019: 4171-4186.
- [20] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [J]. Advances in neural information processing systems, 2017, 30: 5998-6008.
- [21] FENG M, XIANG B, GLASS M R, et al. Applying deep learning to answer selection: a study and an open task [C]//2015 IEEE workshop on automatic speech recognition and understanding (ASRU). Scottsdale: IEEE, 2015: 813-820.

(责任编辑:于慧梅)

(下转第95页)

Evaluation and Optimization of Landslide Dam Breaking and Flood Peak Parameter Models

JIAO Xu, FENG Zhenyu, LIAO Haimei *

(College of Civil Engineering, Guizhou University, Guiyang 550025, China)

Abstract: Landslide dams are generally caused by the fact that unstable geological bodies like landslide quickly accumulate and block rivers, which pose a great threat to the safety of people's life and property in the downstream area after the breach. Therefore, establishing a fast and accurate prediction model of breach parameters is necessary for hazard prevention and mitigation. In this paper, 75 cases of landslide dam failure including parameters like dam height, reservoir capacity, peak discharge of breach and erosion coefficient of dam body were collected. The mean square error E_{RMS} , correlation coefficient R^2 and percentage deviation P_{BIAS} were used to compare and analyze the prediction effect of representative parameter models of earth-rock dam and landslide dam. Through regression analysis, a new prediction model of dam breach depth and peak discharge is established, and the new model is compared with the existing models. Finally, the applicability of the model is verified in actual cases. The results show that the single parameter, double parameters and earth-rock dam models in the existing models are not effective in predicting the dam break parameters, but the parameter model considering the erosivity of the dam has high accuracy. The prediction effect of the new model established in this paper has been improved with wider application, and the prediction results can provide theoretical reference for disaster prevention and mitigation measures of landslide dams.

Key words: landslide dam; peak discharge; depth of breach; parameter model

(上接第 79 页)

CQA Expert Discovery Method Based on Interest and Expertise Modeling

DING Qiu^{1,2}, YAN Xin^{*1,2}, LIU Yanchao³, XU Guangyi⁴, DENG Zhongying¹

(1. Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China;

2. Yunnan Key Laboratory of Artificial Intelligence, Kunming University of Science and Technology, Kunming 650500, China;

3. The Information Technology Center, HuBei Engineering University, Xiaogan 432000, China;

4. Yunnan Nantian Electronic Information Industry Co., Ltd., Kunming 650040, China)

Abstract: The existing question answering community expert discovery methods model user interest by learning the one-way information of the question sequence answered by users, ignoring the volatility of user interest, which will affect the accuracy of modeling for users who have answered fewer questions. In addition, the role of semantic relevance of historical answers and questions in evaluating user performance is not considered. Therefore, in this research a CQA expert discovery method based on interest and expertise modeling is proposed. First, BERT4Rec is used to learn the two-way information of the recent question sequence answered by users to obtain the recent dynamic interest representation. Secondly, this research builds a user social network, and gets the long-term interest expression of users using DeepWalk algorithm to learn the network structure characteristics. Then, the user professionalism evaluation network is constructed, and weighting corresponding questions is calculated according to the semantic correlation between user answers and questions and feedback information. The attention mechanism is also introduced to focus on the user's performance on issues similar to the new questions, and the user professionalism is expressed. Finally, the user's recent dynamic interest, long-term interest and professional expression are combined to match with new questions for scoring, so as to identify users who are willing to accept the invitation and can provide high-quality answers to new questions. The experiment shows that this method has achieved good performance: compared with the baseline method, the MRR evaluation indexes of English, 3dprinting and Tianya Q&A datasets are improved by 5.2%, 2.7% and 16.1% respectively.

Key words: CQA; expert discovery; dynamic interest modeling; social networking; professional modeling