

基于知识图谱的测井储层推荐算法研究

尚福华, 张月霞, 曹茂俊

(东北石油大学 计算机与信息技术学院, 黑龙江 大庆 163318)

摘要:针对实际应用中测井解释人员由于经验不足以及测井解释模型所选取的参数不当、差错等,所造成的测井处理解释结果的评价精度不准等问题,提出了一种基于知识图谱的测井储层推荐算法。首先,构建测井领域知识图谱对测井解释操作人员和其测井储层特性信息进行统一表示;其次,通过加入注意力机制和 TransR 算法对测井领域知识图谱进行补全,再通过连接机制把推荐算法与知识图谱结合起来;最后,根据潜在特征预测并推荐测井领域中储层参数信息。在测井领域数据集上的实验结果表明,该推荐算法在 Top-K 推荐中 $K=15$ 时,其准确率、召回率和归一化折损累计增益三个指标比同类算法分别提高了 6.21%、8.05%、4.82%,并能够高精度预测和推荐测井领域储层参数信息以及更准确地判断储层油气状况,从而揭示出基于知识图谱构建测井储层推荐算法的研究方案是可行的。

关键词:知识图谱;知识图谱补全;测井储层信息;注意力机制;推荐算法

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2023)04-0132-08

doi:10.3969/j.issn.1673-629X.2023.04.020

Research on Logging Reservoir Recommendation Algorithm Based on Knowledge Graph

SHANG Fu-hua, ZHANG Yue-xia, CAO Mao-jun

(School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China)

Abstract: In order to solve the problems of inaccurate evaluation accuracy of logging interpretation results caused by inexperience of logging interpreters and improper parameters and errors of logging interpretation model in practical application, a reservoir recommendation algorithm based on knowledge graph is proposed. Firstly, the logging domain knowledge map is constructed to represent the logging interpretation operators and their logging reservoir characteristics uniformly. Secondly, attention mechanism and TransR algorithm are added to complete the log domain knowledge map, the recommendation algorithm is combined with the knowledge graph by connecting mechanism. Finally, the reservoir parameter information in logging field is predicted and recommended according to the potential characteristics. Experimental results on logging data sets show that when $K=15$ in Top-K recommendation, the accuracy, recall rate and normalized loss cumulative gain of the proposed algorithm are 6.21%, 8.05% and 4.82% higher than those of similar algorithms respectively. And it can predict and recommend reservoir parameter information in logging field with high precision and judge reservoir oil and gas condition more accurately, thus revealing that the research scheme of constructing logging reservoir recommendation algorithm based on knowledge graph is feasible.

Key words: knowledge graph; completion of knowledge graph; logging reservoir parameters; attention mechanism; recommendation algorithm

0 引言

随着计算机技术的飞速发展,计算机深度学习辅助石油勘探开发在石油工程中显得愈加重要,而且随着测井技术的进步和仪器探测设备的升级更新,包括岩性、岩相、储集物性参数^[1](孔隙度、渗透率、泥质含量、含油气水饱和度^[2]等)和地层压力等可用的测井

信息类型和数量剧增,精度不断提高,测井数据的处理解释越来越精细,应用范围也越来越广。在这种形势下,利用大量有用测井和地质信息去探索从而推荐测井储层参数的新方法,将是测井解释技术发展的主要趋势^[3]。

测井解释模型在油气评价方面无疑是最基础、最

收稿日期:2022-06-27

修回日期:2022-10-27

基金项目:黑龙江省自然科学基金(LH2019F004);东北石油大学优秀中青年科研创新团队(KYCXTD201903)

作者简介:尚福华(1962-),男,博士,教授,研究方向为知识图谱和人工智能油田智能软件;通讯作者:曹茂俊(1978-),男,博士,副教授,研究方向为知识图谱、油气领域人工智能及其应用研究、机器学习。

有价值的重要技术,但在实际的测井储层特性中仍然存在着不少解释不准、判断不灵的情况,严重地影响这些地区油气勘探的效率与成功率。常规测井解释在确定某种解释模式之前,需首先建立岩石体积模型,进而确定测井响应方程和统计经验公式^[4],包括但孔隙度分析处理解释模型、复杂岩性分析处理解释模型、粘土矿物分析解释模型等。目前常规的测井解释分析程序通常都需要提供数十种预置参数,最优化测井解释程序所需预先提供的参数更多。随着测井技术的进步和仪器探测设备的升级更新,包括岩性、岩相、储集物性参数和地层压力等可用的测井信息越来越丰富。使用正确的解释模型,就可用有关的解释方程把测井信息加工成地质信息进而来辅助判断油气状况。因此,为了避免解释人员由于经验不足与测井解释模型中出现的参数选取不当或差错,并且克服参数选取中可能出现的诸多人为因素,推荐算法在地球物理测井储层参数的预测与推荐中显得尤为重要。其中传统的协同过滤推荐算法^[5]以及改进算法,都仅考虑用户少量且单方面的个人偏好,容易忽略用户与其他有效信息的关联关系,例如社交网络或媒体里的信息等,具有较大的局限性,容易出现数据稀疏性和冷启动问题。所以,可在推荐系统中引入知识图谱^[6]、神经网络^[7]、注意力机制^[8]、社交网络等辅助信息来增强测井领域知识上下文的关联性,来解决以上传统推荐算法中存在的两种问题,大大增强了推荐系统的整体性能。其中知识图谱作为一种可解释性非常强的语义网络,可在推荐领域中提供潜在的辅助信息。Gazzotti 等人^[9]提出一种基于知识图谱的算法来解决患者住院预测问题,通过不同的知识图谱中提取信息来丰富 EMR 的向量表示,以及利用其中的特征进行自动选择来区分噪声的特征和有利于决策的特征;Kang Yang 等人^[10]提出基于多模态知识图的推荐系统,在基于用户历史点击视频的基础上,构造相应的多模态知识图,然后利用此推荐系统模型逐层提取该图,最后进行推荐;Ye Qing 等人提出了一种基于知识图和推荐系统的药物靶向相互作用预测框架^[11](KGE_NFM)。该框架首先学习知识图谱中各种实体的低维表示,然后通过神经因子分解机集成多模态信息,最后进行准确的预测。同时,随着深度学习的发展,为解决数据稀疏性问题,神经网络和注意力机制也逐步引入到了推荐系统中。Chen Ming 等人^[12]提出基于神经网络的推荐算法(CoNet),它可有效地对协同过滤(CF)中的共现模式进行建模,进而提取出高度描述性的特征,从而获取更好的性能;Liang Zhang 等人^[13]提出一种集成神经网络的 Top-N 推荐算法,主要通过对互联网信号中 Top-N 推荐算法的干扰,达到保护人们隐私的目的。研究表明,当

信号干扰强度为 5,推荐算法为 F 时,网民的隐私保护效果最好,故集成神经网络的 Top-N 推荐算法在保护人们隐私方面具有重要的潜在价值。Jin Huang 等人^[14]提出一种基于注意机制的知识图谱完成的深度嵌入模型,旨在解决一系列的跨模型和语义匹配模型只关注知识图谱的阴影信息中无法捕捉到知识图谱中隐含的细粒度特征,称为方向多维注意卷积模型,该模型探索了方向信息和三元组固有的深层表达特征。

鉴于以上研究,该文将融合注意力机制的知识图谱及知识图谱补全与推荐算法相结合,主要贡献如下:(1)采用面向测井解释操作人员和某地区储层信息构建了测井领域的知识图谱,通过注意力机制强化地区储层信息和知识补全的测井领域知识图谱,充分挖掘储层信息和测井操作人员的低阶与高阶属性,通过知识图谱得到储层信息和测井操作人员的统一表示;(2)融合深度神经网络协同过滤算法(NeuralCF)和加入注意力机制的知识补全算法(TransR)来提取储层信息和测井解释操作人员的潜在特征,并在此基础上为测井操作人员预测和推荐测井领域储层参数信息。

1 算法基本原理

该文提出了一种基于知识图谱的测井储层特性推荐算法:A-KgNc(Attention-Knowledge Graph Network-Collaborative),测井解释人员进行常规测井处理解释操作(孔隙度分析处理解释模型、复杂岩性分析处理解释模型、粘土矿物分析解释模型)时,根据该算法模型来为其推荐所需要的参数,例如:计算泥质含量时,GCUR(地区经验系数)对第三纪地层为 3.7,对老地层是 2。

1.1 A-KgNc 模型的总体框架

算法模型如下:将构建好的测井领域知识图谱应用到解决测井领域中进行测井解释时的储层参数选取不当或差错的问题中,进而结合注意力机制用于补全并加强测井领域知识图谱中的统一信息,最后通过连接机制将推荐算法与知识图谱两者结合来有效并高精度地给测井解释人员推荐储层参数,对测井操作人员的历史行为以及推荐储层参数结果做出解释,提供相关依据,进而更加准确地判断储层油气状况。

首先,构造测井操作人员与储层信息的交互信息矩阵和已知内容包括的三元组(h, r, t)的测井领域知识图谱 G ;然后,使用知识补全算法(TransR)以及注意力机制对测井领域知识图谱进行补全,同时搭建神经网络协同过滤推荐算法;最后,通过交叉压缩机制融合测井领域知识图谱与推荐算法来完成测井储层参数的推荐。

模型整体框架如图 1 所示。

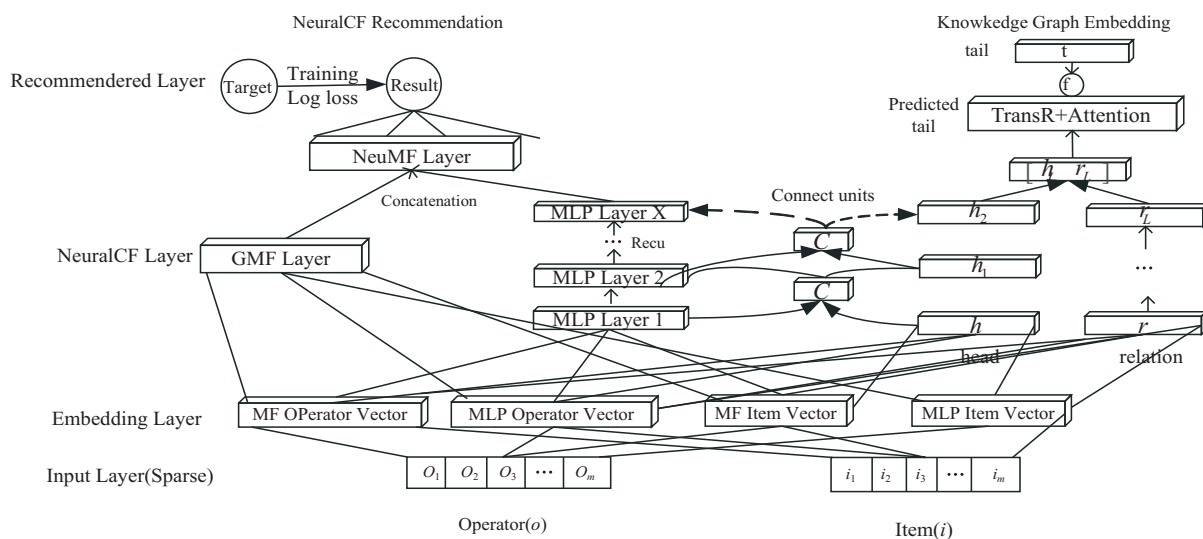


图 1 A-KgNc 模型框架

1.2 测井领域知识图谱的构建

1.2.1 知识图谱构建算法

首先,构造测井操作人员与储层信息的交互信息矩阵和已知内容包含的三元组 (h, r, t) 的测井领域知识图谱 G , 其中 h 表示头实体:包括测井解释操作人员或测井解释分析程序等; t 表示尾实体:包括储层参数数据(孔隙度、渗透率、泥质含量、含油气水饱和度等)、油层数据、结论等; r 表示头尾实体的连接关系:包括测井解释人员操作井数据、储层参数数据、地层参数数据等。例如,三元组(测井操作人员,操作,POR解释分析程序)表示测井操作人员操作测井解释分析程序^[15]。

在许多推荐方案中,储层特性信息可能相关联于知识图谱中的一个或多个实体,比如:“解释程序”这

一实体在知识图谱中就包含多种解释程序类型(孔隙度分析处理解释模型、复杂岩性分析处理解释模型、粘土矿物分析解释模型等)。

本课题构建的知识图谱有多个,包括测井操作人员-类型知识图谱、测井操作人员-储层参数知识图谱、储层参数-油层数据知识图谱等,例如测井操作人员-储层参数知识图谱(22 543 条数据)(一位测井操作人员会操作多个地层参数数据)。

其中测井操作人员-储层参数知识图谱如图 2 所示。

然后是储层参数-油层数据知识图谱(26 839 条数据)(同一地区储层会分为干层、水层、地产油层、油层等)构建的知识图谱,如图 3 所示。

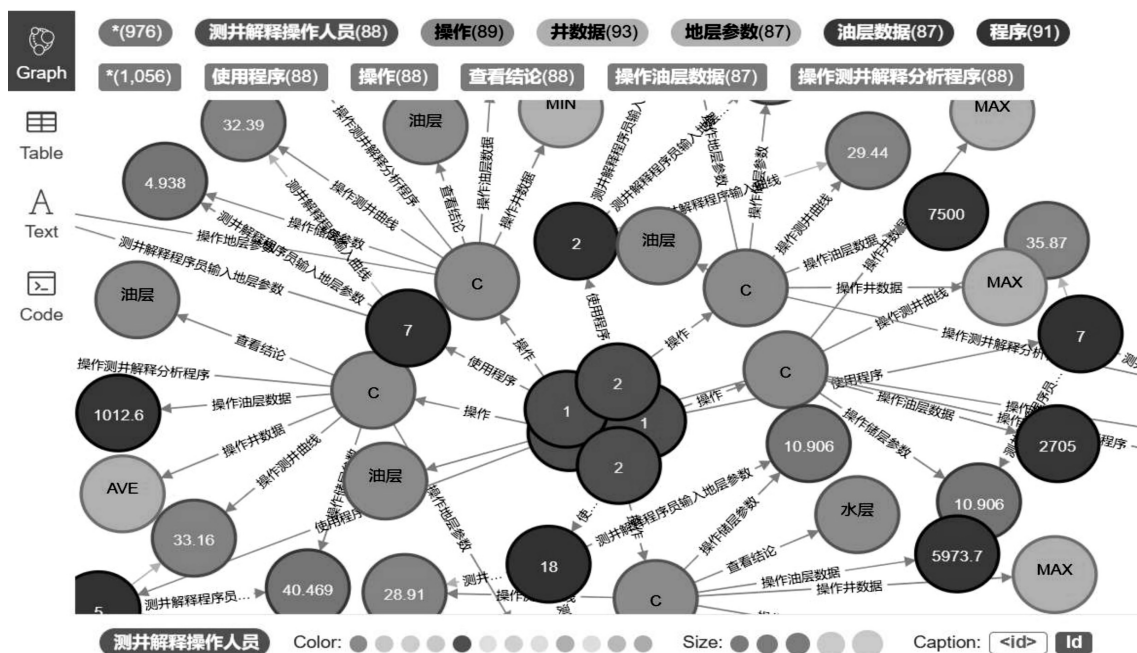


图 2 测井操作人员-储层参数知识图谱

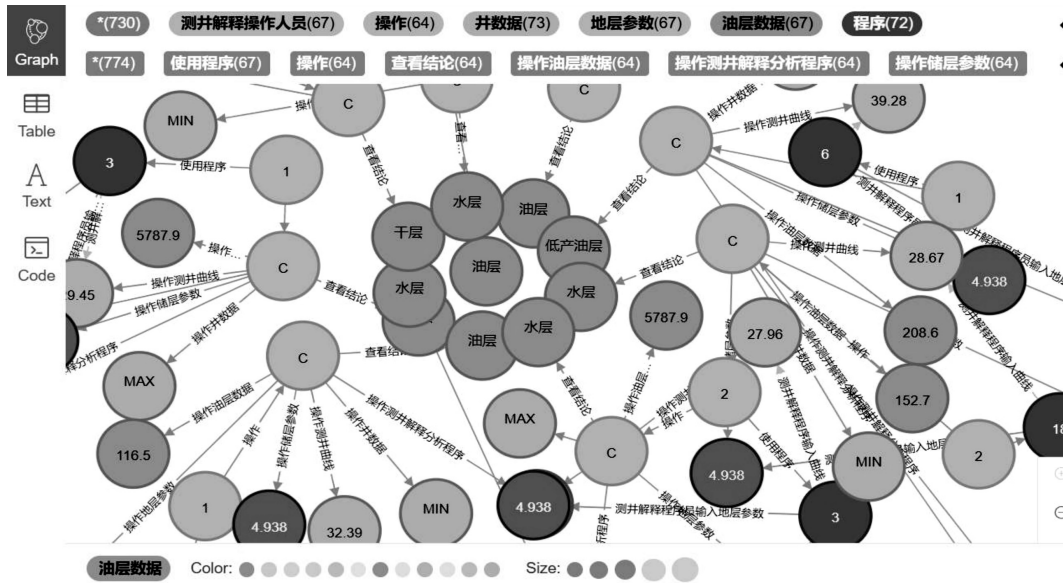


图3 储层参数-油层数据知识图谱

1.2.2 知识图谱补全算法

该文在测井领域储层特性的推荐任务中,将测井领域知识图谱作为辅助信息融入到推荐系统中,现有的大多数知识感知推荐算法都默认知识图谱是完整的,但现实中知识图谱实体间的关系有很多是缺失的,会影响推荐结果的准确性。为丰富测井领域知识图谱,将数据嵌入 TransR 模型中进行知识图谱补全,如图4所示。

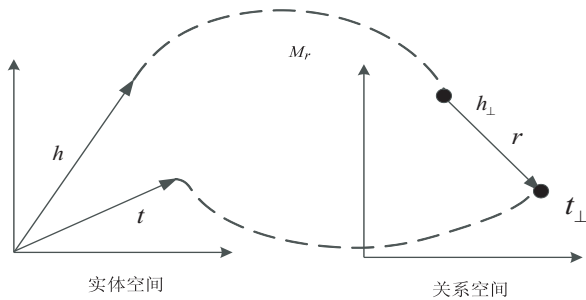


图4 TransR 模型

在 TransR 模型中,分别在测井领域知识图谱中的测井解释操作人员等的实体空间及其相关关系空间中对实体和关系进行建模,进而在关系空间中进行翻译。

其中,三元组 (h, r, t) 中 h 、 t 分别为头实体和尾实体, r 为两者关系,即 $h, t \in R^k$ 。

此模型认为不同的关系有不同的语义空间, h 、 t 根据其关系 r 通过投影矩阵 M_r 映射到关系空间中,即 h_{\perp} 、 t_{\perp} 表示为:

$$t_{\perp} = M_r t \quad (1)$$

然后,将得分函数定义为:

$$f_{\text{TransR}}(h, r, t) = \|h_{\perp} + r - t_{\perp}\|^2 \quad (2)$$

通过在训练集上最小化基于边际的排序损失学习式(1)和式(2)中的嵌入:

$$L_{\text{kgc}} = \sum_{S_p \in S^+} \sum_{S_n \in S^-} \max(0, f_r(h, t) + \gamma - f_r(h', t')) \quad (3)$$

其中, $\max(x, y)$ 的目的是获取 x 和 y 之间的最大值, $S_p \in S^+$ 表示测井领域知识图谱中的关系,补全后的知识图谱即为 S^- :

$$S^- = \{(h', r, t), (h, r, t') \mid h' \in E, t' \in E, (h, r, h) \in S^+, (h', r, t) \in \tilde{S}^+, (h, r, t') \in \tilde{S}^+\} \quad (4)$$

其中, \tilde{S}^+ 是 S^+ 的补集,并且 $\tilde{S}^+ \cup S^+ = G$, $\tilde{S}^+ \cap S^+ = \emptyset$ 。

1.2.3 注意力机制

为了探索测井解释操作人员及其储层特性信息之间的交互信息,该文提出将注意力机制引入知识图谱中,与 TransR 算法模型结合,进而根据权重来补全测井领域知识图谱,获取其潜在信息。

注意力机制的模型如图5所示。

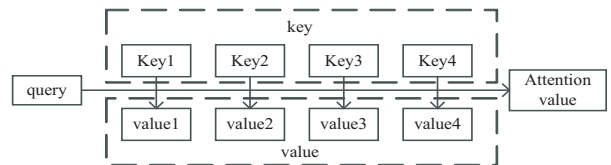


图5 注意力机制模型

相应的表达式如式(5)所示:

$$A(i, j) = \sum_{i=1}^L S(j, k_i) * V_i \quad (5)$$

式中, $A(i, j)$ 是测井储层特性中已知地区储层信息 j 与目标地区储层信息 i 之间的 Attention 值,其中 $S(j, k_i)$ 是已知地区储层信息 j 与目标地区储层信息 i 对应的相似值, V 是目标地区储层信息 i 的值。

结合 TransR 模型,将输入测井操作人员-储层特性信息三元组矩阵 R_u^h 进行降维变成低维向量 R_{im} ,经

过注意力值的运算得到注意力值,通过输出不同储层信息的注意力分值来分配个性化权重,分值越高,其重要性也越高,如式(6)所示:

$$A_u = \sum S(j, R_{im}) * V_i \quad (6)$$

式中, R_{im} 表示测井解释操作人员-储层特性信息集合, $S(j, R_{im})$ 表示候选储层特性信息 j 与测井解释操作人员-储层特性信息集合中的储层信息之间的相似度, V_i 表示储层信息 i 在目标测井解释操作人员-储层项目信息集合中的全值。

1.3 连接机制

为了对测井储层信息和测井操作人员之间的特征交互进行建模,设计了 A-KgNc 框架中的连接单元,如图 6 所示。

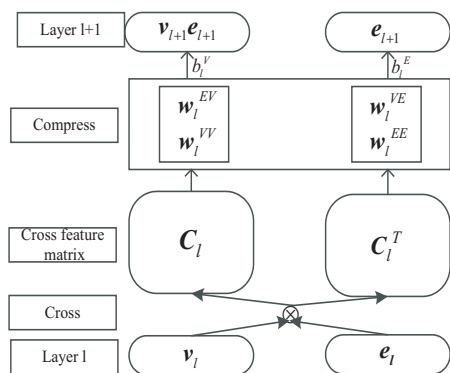


图 6 连接机制

其中,第 v 项储层信息和关联解释人员 e ,首先在层 l 其潜在特征 $v_l \in R^d$ 和 $e_l \in R^d$ 里的 $d \times d$ 对进行相互连接:

$$C_l = v_l e_l^T = \begin{pmatrix} v_l^{(1)} e_l^{(1)} & \cdots & v_l^{(1)} e_l^{(d)} \\ \vdots & & \vdots \\ v_l^{(d)} e_l^{(1)} & \cdots & v_l^{(d)} e_l^{(d)} \end{pmatrix} \quad (7)$$

其中, $C_l \in R^{d \times d}$ 是层 l 的连接矩阵。在连接特征矩阵中明确建模了解释人员 v 及其关联解释人员 e 之间的关系,之后通过把连接特征矩阵投影到其潜在表示空间中,最后输出下一层的测井储层信息和解释人员的特征向量:

$$\begin{aligned} v_{l+1} &= C_l w_l^{VV} + C_l^T w_l^{EV} + b_l^v = v_l e_l^T w_l^{VV} + e_l v_l w_l^{EV} + b_l^v \\ e_{l+1} &= C_l w_l^{VE} + C_l^T w_l^{EE} + b_l^e = v_l e_l^T w_l^{VE} + e_l v_l w_l^{EE} + b_l^e \end{aligned} \quad (8)$$

其中, $w_l^{VV} \in R^d$ 和 $b_l^v \in R^d$ 分别是可训练的权重、偏差向量,权重向量将连接特征矩阵从 $R^{d \times d}$ 空间投影^[16]回特征空间 R^d 。

通过此连接机制, A-KgNc 可以自适应地调整相关实体及关系的权重,并知推荐系统和知识图谱这两个任务之间的联系程度。

1.4 推荐算法设计

近年来,深度学习的应用改进了传统的协同过滤

算法,大大提高了协同过滤算法的泛化能力和拟合能力,因此 A-KgNc 中推荐模块使用的是 NeuralCF(神经网络协同过滤)算法,其输入由原始特征向量 u 和 v 组成,分别描述测井解释操作人员 u 及其项目 v 。

该文使用 L 层建立一个全连接层即嵌入层,与输入层连接,用 MLP(多层感知机)提取其潜在的连接特征:

$$u_L = M(M(\cdots M(u))) = M^L(u) \quad (9)$$

其中, $M(x) = \sigma(W_x + b)$ 是神经网络层,权重 W ,误差 b 和非线性激活函数,对于测井储层信息 v ,使用 L 个连接单元提取其特征^[16]:

$$v_L = E_{e \sim S(v)}[C^L(v, e)[v]] \quad (10)$$

其中, $S(v)$ 是测井项目相关实体的集合 v 。

在获得测井解释操作人员 u 的潜在特征 u_L 及其项目 v 的潜在特征 v_L 之后,通过预测函数 f_R 将这两种途径结合起来。为测井解释操作人员 u 推荐其项目 v 的储层参数最终预测概率:

$$\hat{y}_{uv} = \sigma(f_{RS}(u_L, v_L)) \quad (11)$$

2 实验结果及分析

2.1 矿场资料数据

实验数据来源于大庆油田、新疆油田某工区的三口水平井(Welldatalens)(分别为大庆油田的 A、B 井和新疆油田的 C 井),每口水平井均包含十条测井曲线、储层参数数据、地层参数数据、测井解释分析程序、井数据、油层数据、结论数据等,其中测井曲线为电阻率(RT)、密度(DEN)、中子(CNL)、声波时差(HAC)和井径(CAL),分别选取测井曲线 662 m~930 m 数据段(采样间隔为 0.125 m)。测井解释模型选择的是:孔隙度分析处理解释模型、复杂岩性分析处理解释模型、粘土矿物分析解释模型。同时把以上数据构建为测井领域知识图谱:测井操作人员-类型知识图谱、测井操作人员-储层参数知识图谱、储层参数-油层数据知识图谱等来进行实验。

根据以上数据集,其对应的统计数据如表 1 所示,并且数据集的密度=(测井操作人员-储层参数交互数)/(测井操作人员数-储层数),因此数据集的稀疏度=1-数据集密集度。

表 1 数据集基本统计

数据集	操作 人员数	项目 数	油层评 分数	数据集 密集/%	冷启动 项目
Welldatalens	765	8 957	41 694	0.631	281

2.2 实验对比模型

(1) NFM。

该模型^[17]在传统的 FM 模型基础上,融合 DNN 的

策略,进而引进特征交叉池化层的结构,使得 FM 与 DNN 进行连接,从而不仅吸取了 FM 的建模低阶特征交互能力,还结合了 DNN 学习高阶特征交互能力,提高了推荐的准确率。

(2) DIN。

该模型^[18]将 NLP 中机器翻译的 Attention 机制引入 CTR 预估模型,即在计算用户兴趣向量的时候,根据候选广告的不同从而动态改变用户兴趣向量的值,大大提高了推荐系统的性能。

(3) KGCN。

该模型在传统的推荐模型中引入了知识图谱,结合图卷积网络的方法来辅助获取用户与物品之间的关系,为其推荐的结果提供了可解释性,提高了推荐系统的准确率^[19]。

2.3 实验评价指标

本课题采用以下三个评价指标对模型 TOP-10 推荐性能进行判断。

准确率指标,对模型最终的 TOP-10 推荐进行的一个最直接的展示,如式(12)所示:

$$\text{Precision} = \frac{\sum_o |R(o) \cap T(o)|}{\sum_o |R(o)|} \quad (12)$$

以上的准确率表示的是最终推荐列表中的测井操作人员的历史常用的储层参数占整体推荐的储层参数列表的比例。

召回率指标,如式(13)所示:

$$\text{Recall} = \frac{\sum_o |R(o) \cap T(o)|}{\sum_o |T(o)|} \quad (13)$$

此召回率指标是指最终推荐列表的测井操作人员的推荐储层参数占整体测井操作人员历史交互列表的比例。其中, $R(o)$ 表示最终给测井操作人员 o 的储层参数列表, $T(o)$ 表示测井操作人员在历史交互列表中展现常用的储层参数集合。

归一化折损累积增益,如式(14)所示:

$$\text{NDGC} = \frac{\text{DCG}}{\text{IDCG}} = \frac{\sum_{i=1}^p \frac{\text{rel}_i}{\log_2(i+1)}}{\sum_{i=1}^{\text{rel}} \frac{\text{rel}_i}{\log_2(i+1)}} \quad (14)$$

其中, DCG 是折损累计增益, IDCG 是将 DCG 进行归一化处理之后的指标, rel_i 是候选储层数据与操作人员历史交互记录中储层数据之间的相似度,取值为 $[0, 1]$, 整个指标所表示的是用来衡量整体推荐算法对候选储层数据与操作人员历史交互记录中储层参数信息数据之间的相似度大小的指标。

2.4 实验环境及参数设置

本课题将构建好的测井领域知识图谱,融合注意力机制和 TransR 补全模型(对测井操作人员交互信息

的潜在语义进行辅助与补充及深层次的信息挖掘),嵌入到神经网络协同过滤推荐算法模型中。

鉴于实验结果的可靠性,故需要对测井领域数据集 Welldatales 做预处理,将其 80% 用于训练模型, 20% 用于测试模型,经过多次实验取得平均值。

实验环境的配置如下: Windows10、英特尔-i7-2630QM @ 2Ghz、Visual Studio Code2020、Neo4j-community_win-dows-x64_3_3_1、python-3.8、sklearn-0.24.1、tensorflow-2.5.0。

其中神经网络部分的参数设置如下:关于 NCF 神经网络的过滤器数量设置为 64,嵌入维度设置为 16,批次处理大小设置为 256,权值设置为 16,学习率为 0.01, TOP-K 推荐数为 10。

2.5 结果分析

在所有参数默认设置的情况下,将文中算法与现在主流的推荐算法在测井领域数据集 Welldatales 上进行了相关指标的对比分析,具体数据见表 2。

表 2 算法对比分析

算法	Precision@ 10	Recall@ 10	NDGC@ 10
NFM	0.278 6	0.265 4	0.654 3
DIN	0.284 3	0.273 8	0.668 2
KGCN	0.308 2	0.297 4	0.695 6
A-KgNcW	0.355 3	0.339 5	0.718 3

2.5.1 实验准确性分析

本课题实验结合大庆油田、新疆油田某工区的三口水平井数据的相关性分析结果设计如下实验,在测井操作人员进行操作测井解释模型(孔隙度分析处理解释模型、复杂岩性分析处理解释模型、粘土矿物分析解释模型)时,分别根据大庆两口井(A 和 B)的储层特征以及测井数据的含油状况来估计新疆未知井(C)的含油情况以推荐操作所需要的参数。

在以上模型默认参数的设置下,在测井数据集 Welldatales 上,与同类算法分别在三个指标上(召回率 Recall、准确率 Precision 和归一化折损累积增益指标 NDGC)进行相关测试,具体实验结果如图 7 ~ 图 9 所示。

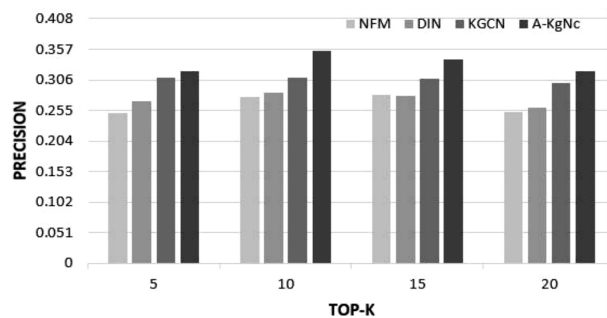


图 7 Precision 指标对比

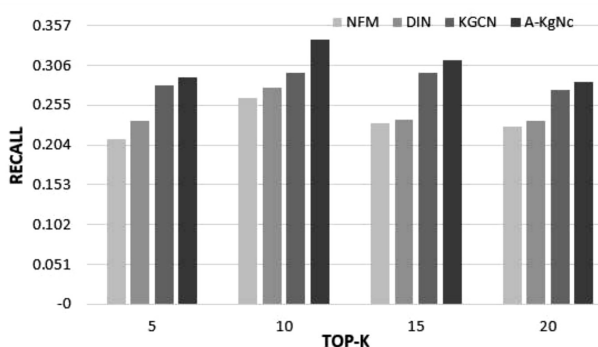


图 8 Recall 指标对比

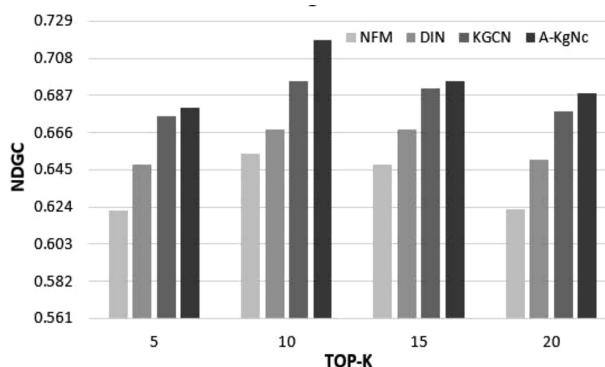


图 9 NDGC 指标对比

经过上面三个图对比分析可知,在测井领域数据集 Welldatales 上,所提 A-KgNc 算法在三个评价指标 Precision、Recall 和 NDGC 上均优于其他算法;另外,还可以看出,在 $K=15$ 时,三个相关性能指标显示最好。

2.5.2 融合注意力机制的知识图谱有效性分析

为了验证本课题模型中融合注意力机制的知识图谱可缓解数据稀疏以及冷启动问题,实验如下:

将 A-KgNc 与 NCF 模型(去掉文中算法中融合注意力机制与知识图谱部分得到的)在数据集 Welldatales 上进行融合注意力机制的知识图谱有效性下的准确率 Precision、召回率 Recall 和归一化折损累积增益指标 NDGC 的验证,其结果分布如图 10~图 12 所示。

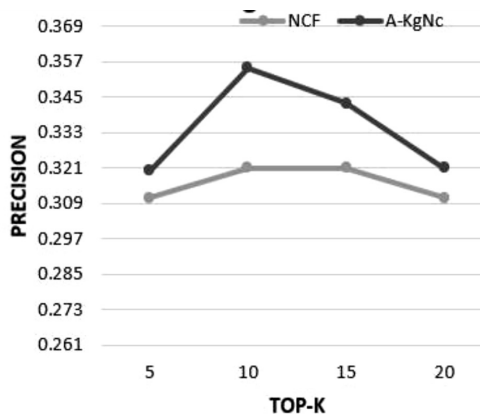
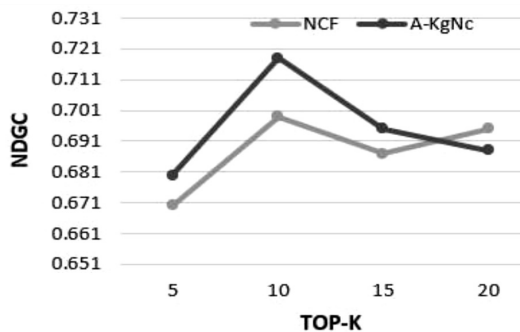
图 10 融合注意力机制的知识图谱有效性验证
Precision 指标分布

图 11 融合注意力机制的知识图谱有效性验证

Recall 指标分布

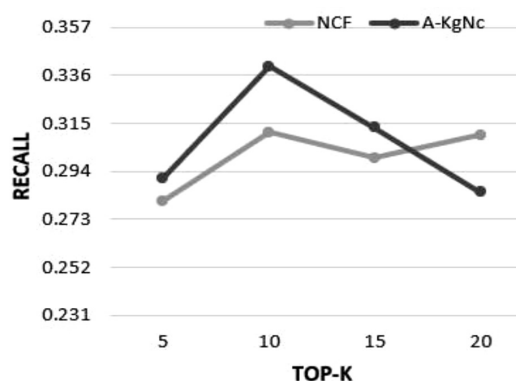


图 12 融合注意力机制的知识图谱有效性验证

TOP = K 指标分布

经过以上三个指标的折线对比图分布可知,文中算法模型在测井领域数据集 Welldatales 的实验下,含注意力机制的知识图谱模型的三个指标均优于 NCF 模型,故可验证融入注意力机制的知识图谱的有效性。

3 结束语

考虑到实际应用中测井解释人员由于经验不足以及测井解释模型所选取的参数不当、差错等,所造成的测井处理解释结果的评价精度不准等问题,在测井储层参数推荐任务上,A-KgNc 算法在补全后的测井领域知识图谱中针对测井储层参数预测问题具有很高的准确率,尤其是在加入了注意力机制的情况下,对比同类 A-KgNc 算法有了更高的准确率,说明了该算法具有较好的预测性能。

对比同类推荐算法的研究结果可知:知识图谱作为一种可解释性非常强的语义网络可在推荐算法领域中提供潜在的辅助信息。该文做了如下补充:(1)对知识图谱做了补全工作,使测井领域知识图谱更加完整;(2)算法中加入了注意力机制,增强了知识图谱与推荐算法的联系,可为测井解释人员提供潜在信息进而更准确的推荐。

A-KgNc 算法将加入注意力机制及补全后的知识图谱作为辅助信息,引入到神经网络推荐算法中,其模型复杂度方面没有得到最优处理。因此,下一步工作

就是将使用的 NCF 网络(神经网络协同过滤)换成 GCN(图卷积神经网络)网络,并且结合专家经验来更加精确地捕捉储层参数信息,从而提高整体推荐算法性能。

参考文献:

- [1] 刘玲,王琳霖,吴疆,等.鄂尔多斯盆地南部彬长区块上石盒子组7段储层识别与预测[J].石油与天然气地质,2021,42(5):1124-1135.
- [2] NOVIKOVDMITRY A. 西西伯利亚盆地北极油气区侏罗系—白垩系储集层气—水系统平衡模拟[J].石油勘探与开发,2022,49(2):319-329.
- [3] 邵蓉波,肖立志,廖广志,等.基于多任务学习的测井储层参数预测方法[J].地球物理学报,2022,65(5):1883-1895.
- [4] 池佳玮.河南省北部C区山西组煤层气储层测井解释方法研究[D].西安:西安石油大学,2020.
- [5] 刘军,杨军,宋姗姗.基于用户购买意愿力的协同过滤推荐算法[J].吉林大学学报:理学版,2021,59(6):1432-1438.
- [6] WU C, LIU S, ZENG Z Y, et al. Knowledge graph based multi-context-aware recommendation algorithm[J]. Information Sciences, 2022, 595(3): 179-194.
- [7] JUAN F, BAOCAI L, MINGXIA G. Collaborative filtering recommendation algorithm based on deep neural network fusion[J]. International Journal of Sensor Networks, 2020, 34(2): 71-80.
- [8] LI J, LI R R, JIA Y J, et al. Prediction of short-term photovoltaic power via codec neural based deep learning approach[J]. Journal of Energy Resources Technology, 2021, 10(6): 1794-1811.
- [9] GAZZOTTI R, FARON C, GANDON F, et al. Extending electronic medical records vector models with knowledge graphs to improve hospitalization prediction[J]. Journal of Biomedical Semantics, 2022, 13(1): 6.
- [10] KANG Y, SHUO W, TONGLE W. Multimodal knowledge graph recommendation system[J]. International Core Journal of Engineering, 2021, 7(11): 1621-1631.
- [11] YE Q, HSIEH C Y, YANG Z Y, et al. A unified drug-target interaction prediction framework based on knowledge graph and recommendation system[J]. Nature Communications, 2021, 12(1): 6775.
- [12] CHEN M, LI Y H, ZHOU X. CoNet: co-occurrence neural networks for recommendation[J]. Future Generation Computer Systems, 2021, 124: 308-314.
- [13] LIANG Z. Top-N recommendation algorithm integrated neural network[J]. Neural Computing and Applications, 2020, 33(9): 1-9.
- [14] HUANG J, ZHANG T, ZHU J, et al. A deep embedding model for knowledge graph completion based on attention mechanism[J]. Neural Computing and Applications, 2021, 33: 9751-9760.
- [15] BAI Z, TAN M J, SHI Y J, et al. Log interpretation method of resistivity low-contrast oil pay-sin Chang 8 tight sandstone of Huanxian area, Ordos Basin by support vector machine[J]. Scientific Reports, 2022, 12(1): 1046.
- [16] WANG H, ZHANG F, ZHAO M, et al. Multi-task feature learning for knowledge graph enhanced recommendation[C]//The world wide web conference. [s. l.]: ACM, 2019: 2000-2010.
- [17] HE X, CHUA T. Neural factorization machines for sparse predictive analytics[C]//Proceedings of the 40th international ACM SIGIR conference. Shinjuku: ACM, 2017: 355-364.
- [18] HE X N, HE Z K, SONG J K, et al. NAIS: neural attentive item similarity model for recommendation[J]. IEEE Transactions on Knowledge and Data Engineering, 2018, 30(12): 2354-2366.
- [19] HONG W, ZHAO M, XIE X, et al. Knowledge graph convolutional networks for recommender systems[C]//The world wide web conference. San Francisco: ACM, 2019: 3307-3313.
- [12] LI T, NI AN, ZHANG C Q, et al. Short-term traffic congestion prediction with Conv-BiLSTM considering spatio-temporal features[J]. IET Intelligent Transport Systems, 2020, 14(14): 1978-1986.
- [13] ZHANG Z, LI Y, SONG H, et al. Multiple dynamic graph based traffic speed prediction method[J]. Neurocomputing, 2021, 461(10): 109-117.
- [14] 陈华伟,邵毅明,敖谷昌,等.面向在线地图的GCN-LSTM神经网络速度预测[J].交通运输工程学报,2021,21(4): 183-196.
- [15] ZHAO L, SONG Y, ZHANG C, et al. T-GCN: a temporal graph convolutional network for traffic prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(9): 3848-3858.
- [16] 黄伟坚,李春贵.基于时空注意力图卷积神经网络的交通速度预测[J].广西科技大学学报,2022,33(1): 54-62.
- [17] LIANG X, WU L, LI J, et al. R-Drop: regularized dropout for neural networks[J]. arXiv: 2106. 14448, 2021.

(上接第125页)