

基于知识图谱的多模态文本信息抽取算法

李 晗¹, 陈 鑫¹, 瞿 涛¹, 赵 海²

(1. 辽宁工业大学 电子与信息工程学院, 辽宁 锦州 121001; 2. 东北大学 计算机科学与工程学院, 沈阳 110819)

摘 要:在多模态文本信息抽取中,文本、音频、图像等多模态的数据在时间与空间上存在对齐不一致现象,使多模态数据的融合性能下降,导致最终抽取的信息覆盖度与 F_1 值下降。为此,研究基于知识图谱的多模态文本信息抽取算法。首先,利用隐马尔可夫模型和卷积神经网络提取多模态文本与图像特征信息。根据知识信息,结合 Neo4j 图形数据库,构建多模态文本知识图谱,将实体作为节点,关系作为连接节点的边,实体的属性作为节点的附加信息,这种直观的表达方式使原本复杂且难以理解的多模态数据之间的联系变得清晰可见,有效解决数据对齐问题;然后,采用深度优先搜索算法深入挖掘知识图谱中的隐含知识,从中提取度中心性、图密度以及介数中心性特征项,构建多模态文本信息抽取特征集;最后,通过建立基于 BiLSTM 的信息抽取模型,并通过梯度下降法对模型展开训练,将获取的多模态文本信息抽取特征信息输入训练好的 BiLSTM 模型中,实现多模态文本信息抽取。实验表明:该算法在抽取不同模态文本信息时, F_1 值达到 0.94 以上。与其他方法相比,在抽取结果的信息覆盖度方面优势明显。

关键词:知识图谱;文本信息;多模态;信息抽取;Neo4j 数据库;BiLSTM 模型

中图分类号:TP18 **文献标志码:**A **文章编号:**1671-5497(2026)06-1687-08

DOI:10.13229/j.cnki.jdxbgxb.20250103

Multimodal text information extraction algorithm based on knowledge graph

LI Han¹, CHEN Xin¹, QU Tao¹, ZHAO Hai²

(1. School of Electronic and Information Engineering, Liaoning University of Technology, Jinzhou 121001, China; 2. School of Computer Science and Engineering, Northeastern University, Shenyang 110819, China)

Abstract: In multimodal text information extraction, there is inconsistency in the temporal and spatial alignment of multimodal data such as text, audio, and images, which leads to a decrease in the fusion performance of multimodal data and a decrease in the final extracted information coverage and F_1 value. To this end, research is conducted on multimodal text information extraction algorithms based on knowledge graphs. Firstly, utilizing Hidden Markov Models and Convolutional Neural Networks to extract multimodal text and image feature information. Based on knowledge information and combined with the

收稿日期:2025-02-11.

基金项目:辽宁省教育厅高等学校基本科研项目(LJ212410154009);辽宁省揭榜挂帅科技攻关专项项目(2022JH1/10400009);国家自然科学基金项目(12371363);辽宁工业大学教学改革研究项目(xjg2022033).

作者简介:李晗(1984-),男,副教授,博士.研究方向:图谱,复杂网络.E-mail:lih_neu@163.com

Neo4j graph database, a multimodal text knowledge graph is constructed, with entities as nodes, relationships as edges connecting nodes, and entity attributes as additional information for nodes. This intuitive representation makes the connections between complex and difficult to understand multimodal data clear and visible, effectively solving the problem of data alignment. Then, the depth first search algorithm is used to deeply explore the implicit knowledge in the knowledge graph, extract degree centrality, graph density, and betweenness centrality feature items, and construct a multimodal text information extraction feature set. Finally, by establishing an information extraction model based on BiLSTM and training the model using gradient descent, the obtained multimodal text information extraction feature information is input into the trained BiLSTM model to achieve multimodal text information extraction. The experiment shows that the algorithm achieves an F1 value of 0.94 or above when extracting text information from different modalities. Compared with other methods, it has significant advantages in terms of information coverage of the extracted results.

Key words: knowledge graph; text information; multimodal; information extraction; Neo4j database; BiLSTM model

0 引言

多模态文本作为一种融合了文字、音频、图像等多种信息模态的数据形式,以其多样性、交互性和丰富性为特点,蕴藏着海量的有价值信息和知识^[1,2]。这些信息和知识若能被有效提取,便可以极大地增强信息的可用性,促进知识的共享与创新。为了高效地利用这些来自不同模态的数据,挖掘其内在价值,提升信息处理的准确性,对多模态文本进行信息抽取显得尤为重要^[3-5]。

关于信息抽取问题,诸多学者进行了研究,如方萍等^[6]通过BERT模型对文本进行双向编码,将句子映射成富含语义信息的向量表示,然后利用图模型技术对这些向量实施结构化处理,剖析句子间的重要性,依据重要性评估结果抽取出关键信息,达成信息抽取的目的,但该方法在处理长文本时存在一定限制,会导致关键信息的丢失,影响信息抽取的完整性;祖弦等^[7]通过随机游走算法,分析文本中各单词与其候选词的关联分值,并选取分值最高的前 N 个词作为关键词,实现关键信息的抽取,但该方法对文本语义关系的挖掘不足,这在一定程度上影响了关键信息抽取的准确性;Smith等^[8]利用结构规则知识对文本展开句法、语义分析,随后,利用大量标注数据训练出的RNN模型学习文本特征与关键信息之间的关联模式,利用该模型提取出文本中的关键信息,但结构规则知识通常依赖于语言学家的专业知识和经验进行制定,这增加了规则制定的主观性和不确定性,进而影响信息抽取的可靠性;王晓宇等^[9]结

合文本关键词的位置以及词与词之间连接的强度,计算每个词语的重要分值,选取得分较高的词语作为文本的关键词,以此实现关键信息的有效抽取,但该方法对于包含大量专业术语、复杂句式或隐晦表达的文本,则难以准确捕捉其关键信息,影响信息抽取准确性。

知识图谱属于图数据结构,其通过描绘实体之间的关系,形成结构化的知识体系,在多个领域得到大力推广^[10,11]。对此,本文研究基于知识图谱的多模态文本信息抽取算法,结合Neo4j图数据库构建多模态文本知识图谱,直观且精准地表示多模态数据中的实体、关系和属性,将实体作为节点,关系作为边,属性作为节点附加信息,使原本复杂的多模态数据之间的联系清晰呈现。这种表示方式不仅有助于理解数据关系,更重要的是有效解决了多模态数据在时间与空间上的对齐问题。结果显示, F_1 值达到0.94以上。与其他方法相比,在抽取结果的信息覆盖度方面优势明显。能有效为多模态文本信息抽取领域提供帮助。

1 多模态文本信息抽取

1.1 构建多模态文本知识图谱

为提升多模态文本信息抽取的效率与精确度,可以先构建多模态文本知识图谱,因为它能融合文本、音频、图像等多种感知数据源,为信息抽取算法提供一个更为广泛且深入的认知框架,从而更有效地挖掘和利用文本中蕴含的多模态信息。多模态文本知识图谱构建框架,见图1。

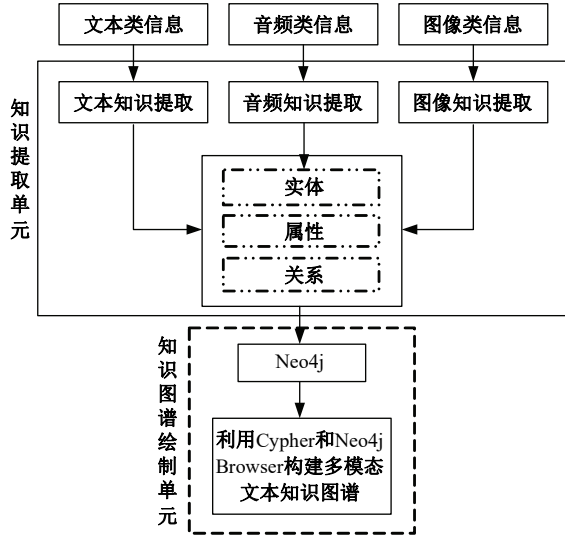


图 1 多模态文本知识图谱构建框架

Fig. 1 Construction framework of multimodal text knowledge graph

由图 1 可知,多模态知识图谱构建框架主要由知识提取单元以及知识图谱绘制单元组成。其中知识提取单元主要是从文本、音频、图像的多种模态数据中获取关于实体、属性以及它们之间关系的知识信息。知识图谱绘制单元则是负责对获取的知识信息进行整合并构建多模态文本知识图谱。

1.1.1 多模态文本知识提取

利用知识提取单元对来自文本、音频、图像等多模态数据进行深入挖掘,获取实体、属性、关系等知识,这些知识是构建知识图谱所必需的要素。采用隐马尔可夫模型对文本数据进行特征提取,采用卷积神经网络对图像数据进行特征提取具体步骤如下:文本数据知识提取本文通过隐马尔可夫模型(HMM)来实现。HMM将文本数据中的知识标签视为隐藏的马尔可夫链所生成的状态序列。这些状态(知识标签)虽然不可直接观测,但每个状态都会对应生成文本中的一个可观测的词汇^[12]。简而言之,HMM通过隐藏的状态标签序列来间接生成可见的词汇序列,从而实现文本数据知识(实体、属性、关系)提取。将HMM定义成五元组模型 $HMM=(S, O, A, C, \pi)$,并将其作为文本数据知识提取模型,其中:

$$S = \{S_1, S_2, \dots, S_n\} \quad (1)$$

$$O = \{O_1, O_2, \dots, O_m\} \quad (2)$$

$$A = \{a_{ij} = P(q_{i+1} = S_j | q_i = S_i)\}, 1 \leq i, j \leq n \quad (3)$$

$$C = \{c_j(O_k) = P(O_k | q_i = S_j)\} \quad (4)$$

$$1 \leq i \leq n, 1 \leq k \leq m$$

$$\pi = \{\pi_i = P(q_1 = S_i)\}, 1 \leq i \leq n \quad (5)$$

式中: S 为隐藏状态集合(文本知识标注集合); n 为状态数量; O 为观测序列(文本中的词汇序列); m 为词汇数量; A 为状态转移概率矩阵; a_{ij} 为从一个知识 q_i 转移到知识 q_{i+1} 的概率; P 为概率值; C 为释放矩阵; $c_j(O_k)$ 为处于状态 c_j 时,释放词汇 O_k 的概率; π 为初始状态分布(初始概率向量)。

隐藏状态集合 S 作为核心,通过初始状态分布 π 确定起始状态,利用状态转移概率矩阵 A 实现状态的动态变化,再借助释放矩阵 C 生成可观测的词汇序列 O 。这种逻辑关系使得隐马尔可夫模型能够通过隐藏的知识标签状态序列间接生成可见的文本词汇序列,进而实现文本数据知识(实体、属性、关系)的提取。

音频数据的知识提取,则先利用语音识别技术将其转换成文本数据,再利用文本数据知识提取方法(HMM模型)实现对其知识的提取。

图像数据知识提取,本文通过卷积神经网络(CNN)实现。CNN可以通过卷积层、池化层和全连接层等结构,逐步抽象出图像中的低级特征到高级特征(如物体的部分、整体以及它们之间的语义关系)^[13]。这些高级特征可以被进一步用于识别图像中的实体(如物体、人物),判断它们的属性(如颜色、形状、大小),以及推断它们之间的关系(如位置关系、交互行为)。本文构建的基于CNN的图像数据知识提取模型,见图2。

模型中,卷积层1对输入层传递过来的图像数据实施卷积操作,捕捉局部特征,并将其转换为一系列特征图后,传送至卷积层2。

卷积层2接收来卷积层1输出的特征图,对其开展更深层次的卷积操作,得到图像信息内容更加复杂和抽象的语义特征。这两层的卷积操作可以描述为:

$$X' = (X * \mu) HMM \quad (6)$$

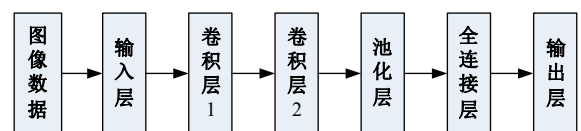


图 2 基于 CNN 的图像数据知识提取模型

Fig. 2 Image data knowledge extraction model based on CNN

式中: X 为上层传递过来的图像或特征图; X' 为当前层输出的特征图; μ 为卷积核。

池化层主要负责对卷积层 2 传递的信息实施下采样, 用于降低数据维度, 同时保留最重要的特征信息, 以提供更加紧密和有效的特征表示。本文采用最大池化法来实现这一环节。

全连接层负责将池化层的输出结果实施线性变换, 并将得到的结果映射至输出空间内。在这一层中, 每个神经元都与池化层输出的 X_i'' 相连接, 利用权重矩阵和偏置矩阵可以实现二者的线性组合, 生成一个线性变换结果 Z , 即全连接层输出, 其公式为:

$$Z = \omega \cdot X_i'' + \varphi X' \quad (7)$$

式中: ω 为权重; φ 为偏置。

输出层负责将全连接层传递的结果, 利用激活函数输出对应的是预定义的类别标签集合, 这些标签代表了图像中的某种物体(实体)、物体的某种特征(属性)或物体之间的关系。

1.1.2 基于 Neo4j 的多模态文本知识图谱绘制

基于上述获取的文本与图像特征, 由于多模态的数据在时间与空间上存在对齐不一致现象, 使多模态数据的融合性能下降, 本文采用知识图谱将其进行整合, 以有效解决数据对齐问题。Neo4j 的图形存储结构能够精准地表示实体的属性及它们之间的关系。在多模态文本知识图谱中, 实体作为节点, 关系作为连接节点的边, 实体的属性作为节点的附加信息。这种直观的方式使得原本复杂且难以理解的多模态数据之间的联系变得清晰可见。

根据知识提取单元提取的多模态文本实体、属性及关系信息, 通过知识图谱绘制单元将这些知识进行整合, 绘制多模态文本知识图谱。这一步骤本文基于 Neo4j 来实现。Neo4j 作为图形数据库, 它凭借丰富的图形查询语言和高效的数据库特性^[14], 能轻松应对多模态文本数据中大量实体及其复杂关系的处理, 其图形存储结构能够精准地表示实体的属性及它们之间的关系, 从而为后续的查询和深入分析提供了极大的便利。基于 Neo4j 的多模态文本知识图谱绘制的具体过程描述如下:

步骤一: 数据导入与知识图谱构建

(1) 数据准备: 根据 2.1.1 小节方法, 从多模态文本中提取实体、关系以及属性知识信息。

(2) 构建知识单元: 将这些实体、关系和属性整理成适合 Neo4j 存储的格式。实体作为节点, 关系作为连接节点的边, 实体的属性则作为节点的附加信息。

(3) 数据导入 Neo4j 数据库: 为了方便导入, 可以将整理好的数据保存为 CSV 文件, 并利用 Neo4j 提供的 API 工具, 将其导入至 Neo4j 数据库内。

步骤二: 知识图谱的绘制

(1) 使用 Neo4j 中的 Cypher 查询语言, 编写 Cypher 语句, 从 Neo4j 数据库中检索出多模态文本知识图谱所需的数据。

(2) 图模式匹配与路径遍历: 利用 Cypher 查询语言的图模式匹配和路径遍历功能, 查询出实体之间的关系路径和属性信息。在此过程中, 需计算实体相似度, 以便进行实体间的链接。本文采用基于余弦相似度的方法来衡量实体之间的相似性, 余弦相似度 $p_{\text{similarity}}$ 公式为:

$$p_{\text{similarity}} = \cos(W, B) = Z \frac{\vec{W} \cdot \vec{B}}{\|\vec{W}\| \|\vec{B}\|} \quad (8)$$

式中: $\vec{W}$ 和 $\vec{B}$ 为实体 W 和 B 的属性向量。

(3) 可视化展示: 将查询结果导入 Neo4j Browser 可视化工具中, 以有向图的形式展示多模态文本知识图谱。在图形中, 节点代表实体, 边代表关系, 节点的属性则以标签或附加信息的形式展示。

(4) 调整图形布局: 在 Neo4j Browser 中, 可以通过拖拽节点和边来调整图形的布局, 使其更加直观易懂。

(5) 迭代更新: 由于知识图谱是一个动态变化的系统, 因此需要对其进行及时更新, 以确保知识图谱的时效性和准确性。

在迭代更新知识图谱时, 增量更新方法是一种常用的技术, 它允许在不重建整个图谱的情况下, 将新的数据纳入知识图谱中。这一过程涉及到新信息筛选以及数值属性(如权重、评分等)的更新。其中更新信息筛选可通过置信度指标评估更新信息的可信度, 置信度 β 的公式为:

$$\beta = (\vec{\lambda} + \vec{h} + d) / 3p_{\text{similarity}} \quad (9)$$

式中: $\vec{\lambda}$ 、 $\vec{h}$ 和 d 分别为词向量、句向量和编辑距离。

数值属性的更新则可通过均值更新公式来实现, 描述为:

$$v'=(v*(n'-1)+\partial)/n'\beta \quad (10)$$

式中: v 、 v' 为更新前、后数值属性均值; ∂ 为文本中新加入的数据点; n' 为文本中数据点数量。

通过置信度公式可以筛选出可靠的更新信息,确保知识图谱的准确性;而均值更新公式则可以高效地更新数值属性,提高知识图谱的时效性。

1.2 多模态文本信息抽取的实现

为了更深层地理解和利用多模态文本知识图谱中的信息,本文首先针对上述获取的多模态文本知识图谱,采用深度优先搜索(DFS)算法来深入挖掘图谱中的隐含知识,特别是实体间的潜在关系、属性的深层次联系等,以此揭示图谱中节点和边的深层特征。然后基于选取的深层特征进行多模态文本信息抽取。具体过程为:

从多模态文本知识图谱中选择一个起始节点 w ;创建一个栈(Stack)用于存储访问过的节点;创建一个集合对已访问的节点实施登记。将 w 放入Stack中,并将其标注成已访问。如果Stack不为空,那么实施下列操作:从栈顶弹出一个节点,作为当前节点;访问当前节点的所有未访问过的邻近节点,并将他们放入Stack内,同时标注为已访问。递归地执行上述过程,直到栈为空。

在搜索过程中,记录每个节点的访问路径和邻接关系。通过分析这些信息,可以发现实体间的潜在关系、属性的深层次联系,根据挖掘出的隐含知识,提取出对后续信息抽取任务有用的特征项。这些特征项分别为:

(1)度中心性。其反映的是某节点在知识图谱中的直接连接数量或重要性,在多模态文本信息抽取任务中,将度中心性作为特征项可以帮助识别那些在知识图谱中具有较高连接度或影响力的节点(实体),这些节点包含更多有用的信息。节点 h 的度中心性 $\rho(h)$ 的公式为:

$$\rho(h)=(\deg(h))/(M-1)v' \quad (11)$$

式中: $\deg(h)$ 为 h 的度数,即与 h 直接相连的节点个数; M 为知识图谱中节点(实体)总个数。

(2)图密度。其反映的是知识图谱中节点之间连接的紧密程度,将其作为文本信息抽取的特征项可以帮助识别文本中实体之间的关系紧密性。图密度 g 公式为:

$$g=(2L)/M(M-1)\rho(h) \quad (12)$$

式中: L 为知识图谱中的边数(关系数)。

(3)介数中心性。它反映的是节点在知识图

谱中的重要性程度,将其作为文本信息抽取的特征项,可以评估文本中关键词或句子在文本信息中的重要性。图谱中节点 h 的介数中心性 $\gamma(h)$ 公式为:

$$\gamma(h)=\sum_{b,v}(\sigma_{bv}(h))/g\sigma_{bv} \quad (13)$$

式中: b 、 v 为多模态知识图谱中的节点,且 $b \neq n \neq v$; σ_{bv} 为 b 到 v 的全部最短路径的总数; $\sigma_{bv}(h)$ 为 b 到 v 的全部最短路径中,经过 h 的路径数。

根据上述特征值,构建多模态文本信息抽取特征集,用于后续的多模态文本信息抽取。Bi-LSTM(双向长短期记忆网络)模型由正、反向的LSTM网络模型组成,能精确捕捉多模态文本中的语义和上下文依赖关系^[15]。其中,利用正向LSTM模型获取文本中上文特征,利用反向LSTM模型获取下文特征。单独一个LSTM模型往往由输入、遗忘、输出三个不同的门,以及记忆单元组建。其中,输入门 u_t 负责筛选新信息(多模态文本特征信息),即控制哪些新信息将被添加到记忆单元中,遗忘门 z_t 负责控制信息的保留与遗忘, u_t 和 z_t 用公式描述为:

$$u_t=\omega_u l_{t-1}+\omega'_u G_t+b_u \gamma(h) \quad (14)$$

$$z_t=\omega_z l_{t-1}+\omega'_z G_t+b_z u_t \quad (15)$$

式中: G_t 为 t 时Bi-LSTM模型输入信息(多模态文本信息抽取特征信息); l_{t-1} 为前一时刻隐藏层的输出状态(上文多模态文本信息内在特征表示); ω_u 、 ω'_u 和 b_u 为 l_{t-1} 和 G_t 的权重、偏置; ω_z 、 ω'_z 和 b_z 为在遗忘门中 l_{t-1} 和 G_t 的权重、偏置。

由此可得出当前单元状态 f_t ,描述为:

$$f_t=z_t f_{t-1}+r_t \dot{f}_t \quad (16)$$

式中: $\dot{f}_t$ 为临时单元状态。

Bi-LSTM模型输出门 o_t 决定了整个模型的输出的隐藏状态 l_t (当前多模态文本信息内在特征表示),描述为:

$$o_t=\omega_o l_{t-1}+\omega'_o G_t+b_o f_t \quad (17)$$

$$l_t=o_t \tanh(f_t) \quad (18)$$

式中: ω_o 、 ω'_o 和 b_o 分别为输出门中 l_{t-1} 和 C_t 的权重和偏置。

将正、反向LSTM模型的输出结果进行融合处理后,可得到结合了前后文信息的内在特征向量。利用SoftMax函数将其转换为概率分布的形式,每一个概率值都对应着一个可能的、从输入的多模态文本中抽取出的信息序列。最终,选择概

率值最高的信息序列作为多模态文本信息抽取的最终结果。

选用梯度下降法作为 Bi-LSTM 模型训练的优化方法,将 2.2.1 小节获取的多模态文本信息抽取特征信息输入至训练好的 Bi-LSTM 模型中,其输出结果即为多模态文本信息抽取结果。

3 实验分析

3.1 实验环境

为了分析本文所提基于知识图谱的多模态文本信息抽取算法的应用效果,实验从某电视台的新闻语料库中收集了相关数据样本。该新闻语料库中包含大量的文本、音频以及图像等多模态文本信息,这些信息内容广泛,涵盖了政治、医学、经济、社会、文化等多个领域。

针对本次实验设置的主要参数,见表 1。实验从该新闻语料库中收集一则关于医学领域某专家的新闻报道,其中包含了文本、音频以及图像等数据。

表 1 实验主要参数	
Table 1 Main parameters of the experiment	
参数名称	选型/数值
操作系统	Ubuntu 16.04 LTS x64
处理器	Pentium3.2 GHz
卷积层 1 卷积核数量/尺寸/步长	128/2/1
卷积层 2 卷积核数量/尺寸/步长	128/2/2
池化层尺寸/步长	2/2
LSTM 单元隐含层	3
隐藏层神经元数量	40
最大迭代数量	80
时间步长	12
批处理大小	64
学习率	0.08

3.2 实验指标

为了衡量本文算法的应用效果,实验利用 F_1 值指标对多模态文本信息抽取结果进行了评估。 F_1 值是精确率与召回率的调和平均数,可以用来评估信息抽取算法的综合性能,取值范围为 $[0,1]$ 。其值越大,说明算法在识别多模态文本中的信息时,既能准确抽取相关信息(高精确率),又能覆盖到尽可能多的实际信息(高召回率);反之,则说明算法在抽取信息时出现了较多的误判(导致精确率下降)或遗漏了较多的实际信息(导致召回率下降)。 F_1 值的公式为:

$$F_1 = 2(\rho_e \times \rho_r) / (\rho_e + \rho_r) \tag{19}$$

式中: ρ_e 为信息抽取的精确率; ρ_r 为信息抽取的召回率。

3.3 知识图谱构建

然后,对数据实施了知识提取,并构建知识图谱,见图 3,可知,应用本文方法构建的知识图谱,可以清晰、直观地将该新闻报道涉及的主要实体(人物、事件)、属性及关系进行了呈现,条理分明,关系明确。由此说明,本文知识图谱构建方法有效可行,能够准确捕捉并展示新闻报道中的核心要素及其相互联系,为理解和分析新闻内容提供了有力的工具。

3.4 实验结果分析

经过测试,得出的结果见图 4,可知,随着数据量的增加,应用本文算法抽取多模态文本信息的 F_1 值,虽有一定程度的下降,但整体仍保持着较高的水平。即使是在数据量为 1 600 的情况下,在抽取不同模态文本信息时, F_1 值也达到 0.94 以

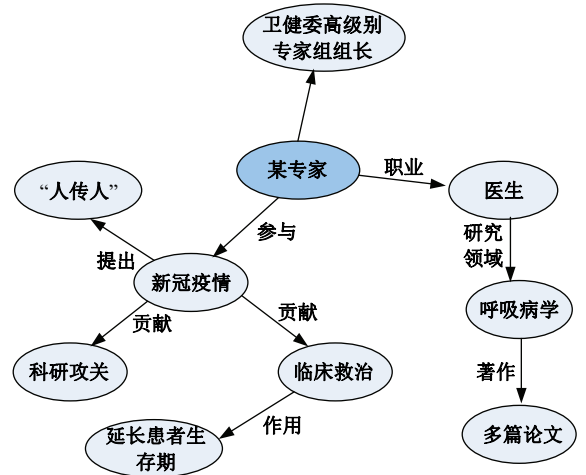


图 3 知识图谱构建
Fig. 3 Knowledge map construction

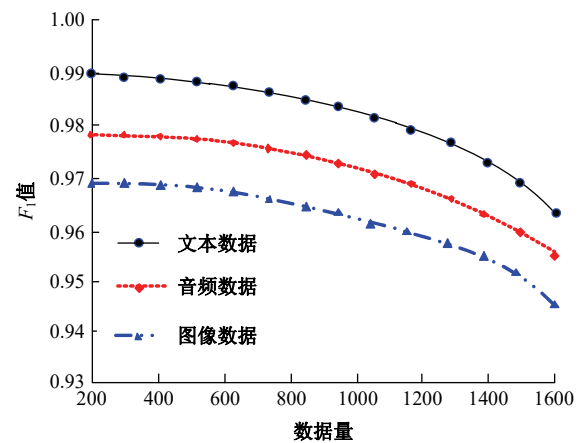


图 4 本文算法 F_1 值情况
Fig. 4 F_1 value of the algorithm in this paper

上。由此说明,本文算法可以准确抽取文本信息,在面对不同模态的文本时,表现出了较强的适应性和泛化能力。

对于信息抽取问题来说,信息覆盖度是一个重要的评价指标。是指从文本中抽取出来的信息相对于原文中实际存在信息的全面性和完整性程度。高信息覆盖度意味着算法能更全面地捉到文本中的关键信息,减少信息的遗漏。为此,实验利用该指标对本文算法的性能进行了评价,并与文献[6]基于BERT双向预训练的信息抽取方法与文献[7]融合词和文档嵌入信息的抽取方法进行对比分析,得出的结果见图5。

由图5可知,本文算法能捕捉到更多的关键信息细节,其信息覆盖度显著高于基于BERT双向预训练的信息抽取方法与融合词和文档嵌入信息的抽取方法。由此说明,本文算法在信息抽取的全面性和完整性方面具有明显优势,有着更强的信息提取能力,减少了信息的遗漏情况,能有效提升信息抽取的准确性和实用性,为相关应用提供了更为可靠的信息支持。

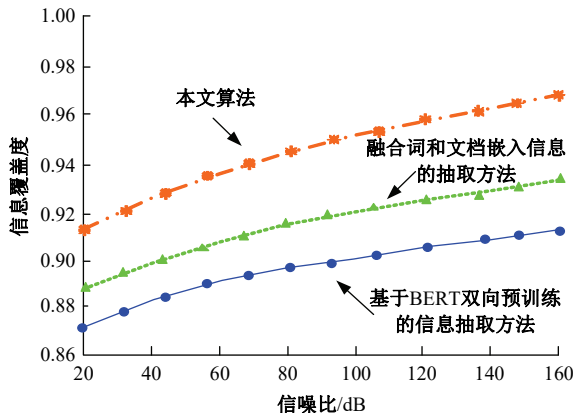


图5 本文算法对比其他算法的信息覆盖度情况

Fig. 5 Information coverage of the proposed algorithm compared with other algorithms

4 结语

本文基于知识图谱的多模态文本信息抽取算法,该算法结合了知识图谱的语义表示能力和多模态信息处理技术的优势,能更全面地理解和抽取文本中的信息,为后续的文本分析、知识推理等任务提供有力支持。实验结果表明:在抽取不同模态文本信息时,本文算法 F_1 值达到0.94以上。信息覆盖度显著高于对比方法。由此说明,本文算法可以准确抽取文本信息,在面对不同模态的

文本时,表现出了较强的适应性与稳定性。

参考文献:

- [1] 胡代旺,焦一源,李雁妮.一种新型高效的文库知识图谱实体关系抽取算法[J].西安电子科技大学学报:自然科学版,2021,48(6):75-83.
Hu Dai-wang, Jiao Yi-yuan, Li Yan-ni. Novel and efficient algorithm for entity relation extraction with the corpus knowledge graph[J]. Journal of Xidian University (Natural Science Edition), 2021, 48(6): 75-83.
- [2] 林泽柠,汪嘉鹏,金连文.视觉信息抽取的深度学习方法综述[J].中国图象图形学报,2023,28(8):2276-2297.
Lin Ze-ning, Wang Jia-peng, Jin Lian-wen. Visual information extraction deep learning method: a critical review[J]. Journal of Image and Graphics, 2023, 28(8): 2276-2297.
- [3] 李玉琦,李龙.基于模式识别算法的网页重复信息抽取仿真[J].计算机仿真,2022,39(3):439-443.
Li Yu-qi, Li Long. Simulation of web repetitive information extraction based on pattern recognition algorithm[J]. Computer Simulation, 2022, 39(3): 439-443.
- [4] 吉旭瑞,魏德健,张俊忠,等.中文电子病历信息提取方法研究综述[J].计算机工程与科学,2024,46(2):325-337.
Ji Xu-rui, Wei De-jian, Zhang Jun-zhong, et al. Research progress on information extraction methods of Chinese electronic medical records[J]. Computer Engineering and Science, 2024, 46(2): 325-337.
- [5] 赖娟,洪艳伟.基于规则约束的深度学习网络用于文本信息抽取[J].计算机工程与设计,2021,42(12):3548-3554.
Lai Juan, Hong Yan-wei. Deep learning network based on rule constraints for text information extraction[J]. Computer Engineering and Design, 2021, 42(12): 3548-3554.
- [6] 方萍,徐宁.基于BERT双向预训练的图模型摘要抽取算法[J].计算机应用研究,2021,38(9):2657-2661.
Fang Ping, Xu Ning. Graph model summary extraction algorithm based on BERT bidirectional pretraining[J]. Application Research of Computers, 2021, 38(9): 2657-2661.
- [7] 祖弦,谢飞,刘啸剑.融合词和文档嵌入的关键词

- 抽取算法[J]. 计算机科学与探索, 2021, 15(2): 294-304.
- Zu Xian, Xie Fei, Liu Xiao-jian. Keyphrase extraction combining word and document embeddings[J]. Journal of Frontiers of Computer Science and Technology, 2021, 15(2): 294-304.
- [8] Smith E, Papadopoulos D, Braschler M S K. Lillie: information extraction and database integration using linguistics and learning-based algorithms[J]. Information Systems, 2022, 105(3): 101938.
- [9] 王晓宇, 王芳. 基于语义文本图的论文摘要关键词抽取算法[J]. 情报学报, 2021, 40(8): 854-868.
- Wang Xiao-yu, Wang Fang. Keyword extraction from a paper's abstract based on semantic text graph [J]. Journal of the China Society for Scientific and Technical Information, 2021, 40(8): 854-868.
- [10] 景丽, 姚克. 融合知识图谱和多模态的文本分类研究[J]. 计算机工程与应用, 2023, 59(2): 102-109.
- Jing Li, Yao Ke. Research on text classification based on knowledge graph and multimodal[J]. Computer Engineering and Applications, 2023, 59(2): 102-109.
- [11] 赵一鸣, 吴林容, 任笑笑. 基于多知识图谱的中文文本语义图构建研究[J]. 情报科学, 2021, 39(4): 23-29.
- Zhao Yi-ming, Wu Lin-rong, Ren Xiao-xiao. Chinese text semantic graph construction based on multiple knowledge graphs[J]. Information Science, 2021, 39(4): 23-29.
- [12] 陈勇, 邢欣, 张锦文. 面向文书的情报关键信息抽取算法[J]. 火力与指挥控制, 2023, 48(1): 142-148.
- Chen Yong, Xing Xin, Zhang Jin-wen. Document-oriented intelligence key information extraction algorithm[J]. Fire Control & Command Control, 2023, 48(1): 142-148.
- [13] 朱兵兵, 罗飞, 罗勇军, 等. 基于子句抽取的文本摘要自动提取算法[J]. 华东理工大学学报:自然科学版, 2024, 50(1): 114-120.
- Zhu Bing-bing, Luo Fei, Luo Yong-jun, et al. An automatic text summarization algorithm based on clause extraction[J]. Journal of East China University of Science and Technology(Natural Science Edition), 2024, 50(1): 114-120.
- [14] 党雪云, 王剑. 基于门控图神经网络的新闻要素信息抽取方法[J]. 电视技术, 2022, 46(5): 24-29.
- Dang Xue-yun, Wang Jian. A method of extracting news element based on gated graph neural network [J]. Video Engineering, 2022, 46(5): 24-29.
- [15] 赵晔辉, 柳林, 王海龙, 等. 知识图谱推荐系统研究综述[J]. 计算机科学与探索, 2023, 17(4): 771-791.
- Zhao Ye-hui, Liu Lin, Wang Hai-long, et al. Survey of knowledge graph recommendation system research[J]. Journal of Frontiers of Computer Science and Technology, 2023, 17(4): 771-791.