

煤矿井下不安全行为的命名实体识别方法

付燕, 刘佩怡, 叶鸥

(西安科技大学计算机科学与技术学院, 陕西 西安 710054)

摘要: 为提高井下安全管理效率, 实现煤矿安全生产, 根据煤矿行业相关标准规范, 并结合井下不安全行为领域知识, 采用 BIO 标注策略构建一个包含 8 类实体类别、2 359 条样本的煤矿井下不安全行为语料库。针对煤矿井下不安全行为命名实体识别任务中存在的语义信息利用不足、实体分布不均衡、实体边界模糊的问题, 提出一种基于 Global Pointer 和对抗训练的煤矿井下不安全行为命名实体识别模型。首先, 采用改进的分层 RoBERTa 模型并利用多层语义信息增强井下不安全行为文本向量化, 结合对抗训练对词嵌入层进行扰动, 缓解数据不平衡问题, 增强模型的鲁棒性; 其次, 在特征提取层采用双向门控循环单元 (BiGRU) 可以更有效地捕获语料的上下文语义特征, 加强文本语义关联; 最后, 在解码层构造 Global Pointer, 获得更准确的实体边界识别结果。为验证提出模型的有效性, 在自建的小样本煤矿井下不安全行为数据集上进行实验, 结果表明, 该模型的精确率、召回率和 F1 值分别为 78.77%、78.20%、78.48%, 相比于 BERT-Global Pointer 模型分别提高了 2.27、0.63、1.45 个百分点, 为构建井下不安全行为知识图谱提供基础。

关键词: 井下不安全行为; 命名实体识别; RoBERTa 模型; 对抗训练; Global Pointer 模型

中图分类号: TP391.1

文献标志码: A

DOI: 10.19678/j.issn.1000-3428.0069917

Named Entity Recognition Method for Unsafe Underground Behaviors in Coal Mines

FU Yan, LIU Peiyi, YE Ou

(College of Computer Science and Technology, Xi'an University of Science and Technology, Xi'an 710054, Shannxi, China)

【Abstract】 A coal mine unsafe behavior corpus containing 8 entity categories and 2 359 samples has been constructed using a BIO labeling strategy to improve the efficiency of underground safety management and realize safe coal mine production, based on the relevant standards and norms of the coal mine industry as well as insights into the field of underground unsafe behavior. Aiming at the problems of insufficient semantic information utilization, unbalanced entity distribution, and fuzzy entity boundary in the named entity recognition task of unsafe behavior in coal mines, this study proposes a named entity recognition model based on Global Pointer and adversarial training. First, the improved hierarchical RoBERTa model is used to make full use of multi-layer semantic information to enhance the text vectorization of underground unsafe behavior, and the word embedding layer is disturbed by adversarial training to alleviate the problem of data imbalance and enhance model robustness. Second, Bidirectional Gated Recurrent Unit (BiGRU) is used in the feature extraction layer to more effectively capture the contextual semantic features of the corpus and enhance the semantic association of the text. Finally, Global Pointer is constructed in the decoding layer to obtain more accurate entity boundary recognition results. The effectiveness of the proposed model is evaluated on a self-built small sample coal mine underground unsafe behavior dataset. The results show that the accuracy, recall, and F1 value of the proposed model are 78.77%, 78.20%, and 78.48%, respectively, which are 2.27, 0.63, and 1.45 percentage points higher than those of the BERT-Global Pointer model. The findings provide a basis for constructing a knowledge graph of unsafe behavior in underground mines.

【Key words】 unsafe underground behavior; named entity recognition; RoBERTa model; adversarial training; Global Pointer model

0 引言

在国家能源战略与安全生产并重的发展格局下, 煤矿行业作为我国能源供应的基石, 面临着既要

保证能源稳定供应又要确保生产安全的双重挑战。煤矿井下作业环境复杂多变, 不安全行为的存在成为制约安全生产的关键因素。据统计结果显示, 大多数煤矿安全事故的发生与作业人员的不安全行为

基金项目: 中国博士后科学基金(2020M673446)。

作者简介: 付燕(CCF 会员), 女, 教授、博士, 主研方向为计算机图形图像处理技术、科学计算及其可视化技术等; 刘佩怡(通信作者), 硕士研究生; 叶鸥, 副教授。

收稿日期: 2024-05-27

修回日期: 2024-07-28

E-mail: 422803396@qq.com

直接相关,如违规操作、忽视安全规程等。因此,如何高效、准确地识别和预防这些不安全行为,成为提升煤矿安全生产管理水平亟待解决的问题。

近年来,随着“智慧矿山”的提出和推进,物联网、云计算、大数据、人工智能等技术逐渐被应用到煤矿安全管理工作中,其中以文本挖掘为代表的自然语言处理(NLP)技术为提升煤矿安全管理水平提供了新思路。对煤矿井下不安全行为知识进行文本挖掘,抽取其中关键信息,可以帮助煤矿安全管理人员掌握不安全行为的诱因、获取事故发生的影响因素以及及时采取预防措施避免发生事故。而在信息抽取过程中,命名实体识别是基础,其识别效果决定着信息抽取的准确性。因此,进行煤矿井下不安全行为的命名实体识别研究,不仅能帮助煤矿安全管理人员迅速锁定不安全行为的具体原因,包括但不限于违规操作、操作失误、忽视警示标识等,还可以驱动煤矿安全管理迈向智能化的新阶段,实现对不安全行为的全面分析及预警。本文针对煤矿井下不安全行为命名实体识别任务中存在的语义信息利用不足、实体分布不均衡、实体边界模糊的问题,提出一种基于 Global Pointer 和对抗训练的煤矿井下不安全行为命名实体识别模型。

1 相关研究

命名实体识别是构建知识图谱的基础任务,指从文本中识别特定实体并划分到相应的类别。命名实体识别方法主要分为三类,即基于词典和规则的方法、基于统计机器学习的方法和基于深度学习神经网络的方法。

命名实体识别早期主要是基于词典和规则的方法。HANISCH 等^[1]基于规则的方法设计了 ProMiner 系统,该系统利用同义词词典自动识别生物医学文本中的蛋白质和基因名称。QUIMBAYA 等^[2]将词典信息和规则相结合,在电子健康记录的命名实体识别任务上取得了较好的效果。但是此类方法依赖于专家制定规则模板,耗时耗力,泛化能力和可移植性差。

随着概率模型的提出,研究人员提出了基于统计机器学习的方法,常用的方法有隐马尔可夫模型^[3]、条件随机场模型^[4]、支持向量机^[5]等。刘杰等^[6]利用隐马尔可夫模型来学习和预测文本中的机构名,提高了识别的准确性。胡文博等^[7]为识别复杂地名和复杂机构名,构建了多层条件随机场模型,在《人民日报》的数据集上取得了不错的效果。卢青华等^[8]通过构建支持向量机分类器识别软件中的标签并分类,实

现软件命名实体识别。

近年来,随着预训练模型的提出,基于深度学习神经网络的方法进行命名实体识别成为了主流。命名实体识别技术在一些通用领域已较为成熟,并且取得了不错的识别效果。余丹丹等^[9]提出了基于 ALBERT 的中文电子简历命名实体识别方法,在 Resume 电子简历数据集中取得了 94.86% 的 F1 值。褚天舒等^[10]提出了基于词汇增强和表格填充的 TLEXNER 模型,该模型在 BERT 的嵌入层中集成了字符与词汇组的相对位置信息,并使用条件层归一化和双仿射模型构造字符对表格,能有效地识别平面实体和嵌套实体。崔少国等^[11]针对中文电子病历数据专业性强与结构复杂的问题,将汉字图像特征、五笔字型编码与医学词典的潜在编码融合,提出一种融合语义及边界信息的命名实体识别模型 WHSemantic+Latice,在 Resume 数据集上的 F1 值达到了 96.06%。林娜等^[12]对机电领域文本进行了多维数据增强,提出了基于改进 loss 的 BERT-BiGRU-CRF 模型,改善了数据不平衡问题。而在煤矿领域,曹现刚等^[13]在煤矿装备维护领域中使用 BERT 模型捕获具有深层语义的词向量,并通过 BiLSTM-CRF 进一步获取上下文特征及解码,改善了一词多义问题。王向前等^[14]针对煤矿事故案例文本的实体识别,将深度学习与传统机器学习算法相结合,提出 ALBERT-BiLSTM-CRF 模型,解决了传统方法对词典和规则依赖性强的问题。刘飞翔等^[15]采用 ERNIE-BiGRU-CRF 模型对煤矿安全隐患文本进行了实体识别,较 BiLSTM-CRF 基线模型效果有一定的提升。刘致豪等^[16]在 BiLSTM-CRF 基础上引入 MLP 进行参数迁移,使得公共数据集 THUCNews 输出维度与煤矿数据输出维度达成统一,使其能够学习更多特征,从而弥补煤矿井下不安全行为数据量不足的问题。

虽然命名实体识别已广泛应用于各个领域,但由于煤矿井下不安全行为文本专业术语多、语义复杂,仍面临以下问题:

- 1) 目前仍缺乏针对煤矿井下不安全行为知识的命名实体识别模型,且缺乏相关标注数据。
- 2) 现有基于 BERT 的深度学习模型词向量化时未充分利用语义信息,造成大量实体和语义丢失。
- 3) 煤矿井下不安全行为文本中实体分布不均衡,实体边界难以判定,进而影响实体识别准确率。

针对上述问题,本文构建了一个煤矿井下不安全行为语料库,并提出了基于 Global Pointer 和对抗训练的煤矿不安全行为命名实体识别方法。首先基于改进的分层 RoBERTa 充分融合多层语义信

息,获得文本的向量表示,并对词向量采用对抗训练加入扰动并生成对抗样本改善实体类别不平衡问题,其次采用 BiGRU 提取井下不安全行为文本的上下文语义信息,最后利用考虑了实体边界的 Global Pointer 进行解码,提高实体识别的效果。

2 煤矿不安全行为语料库构建

2.1 文本获取与预处理

由于煤矿不安全行为领域知识专业术语多,缺乏公开完整的数据集,本文构建了煤矿不安全行为语料库。语料数据来源有两方面,一是《煤矿安全规程》以及各类隐患防治细则,二是通过爬取煤矿安全生产网、安全管理网、煤矿安全监察局网等百科网站和中国知网发表的煤矿不安全行为相关文献。通过

编写正则表达式对语料去除多余的字符、网址、HTML 元素等多源异构数据,转换为格式规范的人工标注语料共计 2 359 条,字符 28 万余个,包含 8 类实体 3 893 个。

2.2 实体标注策略

本文根据煤矿不安全行为领域知识需求确定实体类别和标注策略。

1)实体类别划分:通过参考中国国家标准化管理委员会发布的煤矿科技术语汇总表,并对文献[17-18]中关于不安全行为的研究内容进行分析,将井下不安全行为实体分为煤矿工作人员、疏忽性行为、遗忘性行为、违反性行为、错误性行为、影响因素、导致后果、预防措施 8 类实体,具体定义说明如表 1 所示。

表 1 实体标签说明

Table 1 Entity label description

实体类别	实体数量/个	实体示例	开始标签	中间或者结尾标签
煤矿工作人员(Person)	693	瓦检员、爆破工等	B-Person	I-Person
疏忽性行为(Careless)	419	忽视安全工作、注意力不集中等	B-Careless	I-Careless
遗忘性行为(Forget)	59	遗漏警示标识、忘记关闭风门等	B-Forget	I-Forget
违反性行为(Violate)	691	不按规定设置、违章作业等	B-Violate	I-Violate
错误性行为(Error)	744	操作错误、不懂爆破知识等	B-Error	I-Error
影响因素(Factor)	422	个体因素、年龄、文化程度等	B-Factor	I-Factor
导致后果(Cause)	396	瓦斯爆炸事故、火灾事故等	B-Cause	I-Cause
预防措施(Prevent)	469	提高支护水平、加强技能教育培训等	B-Prevent	I-Prevent

2)标注策略:本文采用命名实体识别任务中常用的序列标注方法 BIO^[19]标注策略逐字进行标注,其中,B(Begin)表示实体的起始位置,I(Inside)表示实体的中间或结尾位置,O(Outside)表示非实体。根据上述实体类别定义,B-Label 表示井下不安全行为实体的第一个字符,I-Label 表示井下不安全行为实体中间或结尾字符,O 表示其他非实体字符。标注工具采用 label-studio,通过人工标注得到井下不安全行为实体识别语料库。

2.3 实体特点分析

煤矿井下不安全行为领域文本中专业词汇多、领域性较强,其命名实体存在实体分布不平衡、实体边界难以区分等特性。

1)实体分布不平衡:井下不安全行为文本中不同类别实体出现的频次不同。如违反性行为 and 错误性行为实体的数量远高于遗忘性行为,而不安全行为往往是由煤矿工作人员在执行生产任务时产生的,因此煤矿工作人员实体出现频次远高于其他实体。

2)实体边界难以区分:井下不安全行为领域中

预防措施类别的实体是多个方面的综合,如“强化职工安全防范意识”和“加强班组和煤矿的安全氛围”实体包含影响因素实体类别里的“安全防范意识”和“安全氛围”,且实体长短不一导致实体边界难以区分。

3 井下不安全行为命名实体识别模型

3.1 模型结构

针对第 2.3 节中所描述的数据特点,本文提出基于 Global Pointer 和对抗训练的命名实体识别模型,整体结构如图 1 所示,主要包括分层 RoBERTa 编码层、BiGRU 语义特征提取层和 Global Pointer 层。模型架构整体处理流程为:文本输入到模型后,编码层采用分层 RoBERTa 对 12 层语义信息进行融合,结合各层语义特点,更好地实现文本向量化。然后对嵌入层通过快速梯度法(FGM)对抗训练算法进行扰动,加入生成的对抗样本,增强模型的鲁棒性。将得到的嵌入向量一同输入到双向门控循环单元(BiGRU)进行特征提取,最后通过 Global Pointer 对输出进行解码,准确地识别出实体。

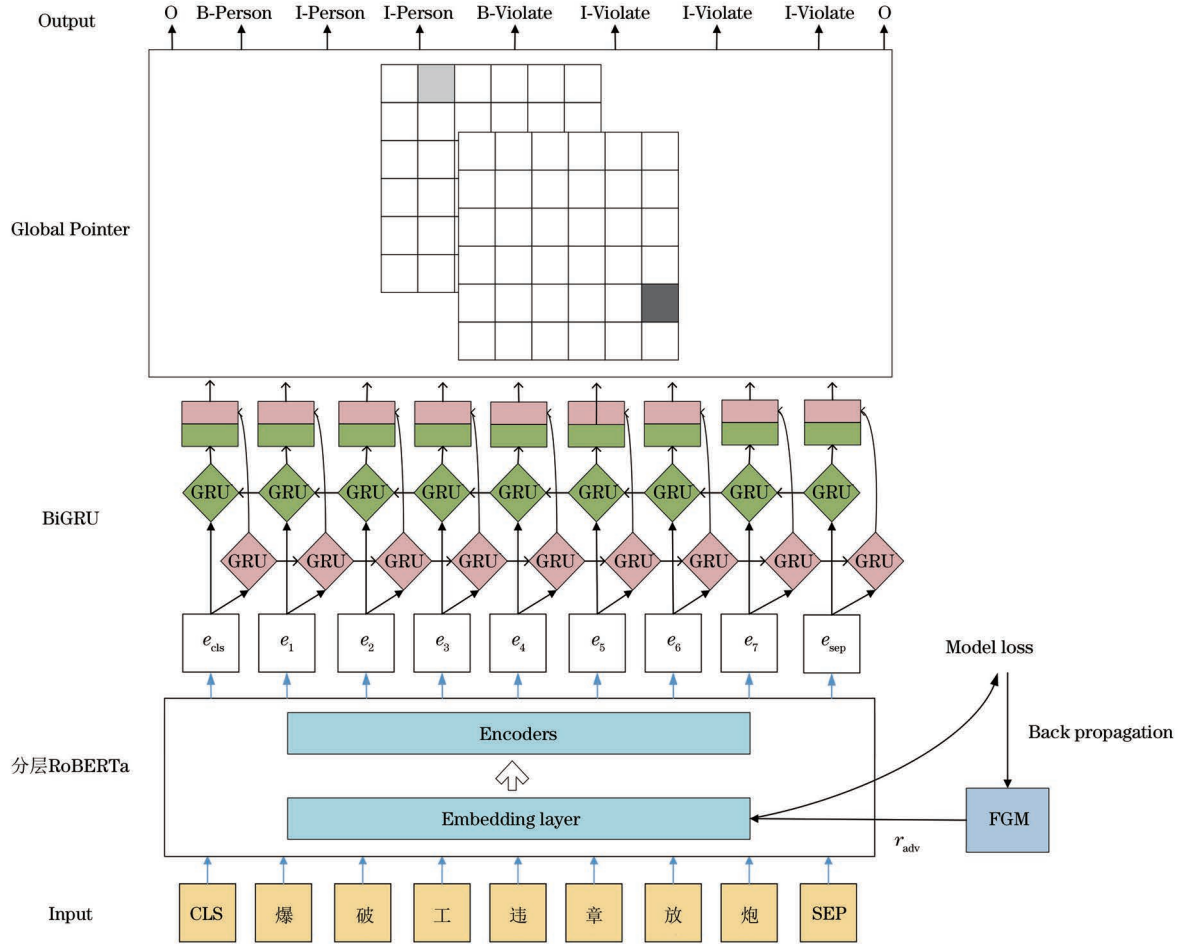


图 1 模型整体结构

Fig. 1 Overall structure of the model

3.2 分层的 RoBERTa 编码层

RoBERT^[20] 是 2019 年提出的基于 Transformer 架构的预训练模型,广泛地应用于 NLP 任务。与传统的 BERT 模型相比,RoBERTa 模型取消了 NSP 任务,集中于学习更好的上下文表示,模型性能得到了提升。RoBERTa 采用动态掩码的方式解决了 BERT 训练时煤矿井下不安全行为文本部分信息丢失的问题,调整了训练策略,在井下不安全行为数据集上具有较高的准确性,因此,本文选择 RoBERTa 作为预训练模型。

JAWAHAR 等^[21] 通过三项探测任务揭示,BERT 编码了丰富的层次信息且每层学习到的表征不同。但现有的研究者仅利用模型最后一层的结果作为输出,忽视了内部其他各层捕捉的特征,造成一定信息的损失。由于 RoBERTa 模型与 BERT 模型网络架构一致,由 12 层 Transformer Encoder 组成,因此本文对 RoBERTa 模型进行分层,整合不同深度层次的信息,以充分挖掘煤矿井下不安全行为文本的语义特征,深化 RoBERTa 对文本的语义理解。如图 2 所示,考虑到 RoBERTa 的 12 层 Transformer

Encoder 在处理信息时的逐步深化特性,本文将其分为 3 个层次,首先将 1~4 层划分为低层网络 low,专注于捕捉文本的表层特征;其次将中间的 4 层构成中层网络 mid,聚焦于文本的句法信息特征;最后 4 层合并为高层网络 high,专注于提取全局上下文信息和深层次语义,理解文本的整体含义和逻辑关系。为了有效整合以上 3 层的输出,本文为每层分配一个可学习的权重参数并初始化,以低层网络 low 为例,计算公式如下:

$$L_{\text{low}}^{\text{Layer}} = \sum_{i=0}^3 L_i^{\text{Layer}} \quad (1)$$

$$\alpha_{\text{low}} = \text{Dense}_{u_{\text{unit}}=1}(L_{\text{low}}^{\text{Layer}}) \quad (2)$$

式中: L_i^{Layer} 表示第 i 层的表征; α_{low} 表示低层网络 low 的权重。中层网络 mid 和高层网络 high 的计算同理。通过训练过程中的反向传播,各层的权重值 α 不断调整优化,从而达到最佳权重分配。最后将低层、中层、高层三层网络的输出加权求和作为模型的最终输出,计算公式如下:

$$o_{\text{output}} = \sum_{i=l_{\text{low}}}^{\text{high}} \alpha_i \cdot L_i^{\text{Layer}} (i \in \{\text{low}, \text{mid}, \text{high}\}) \quad (3)$$

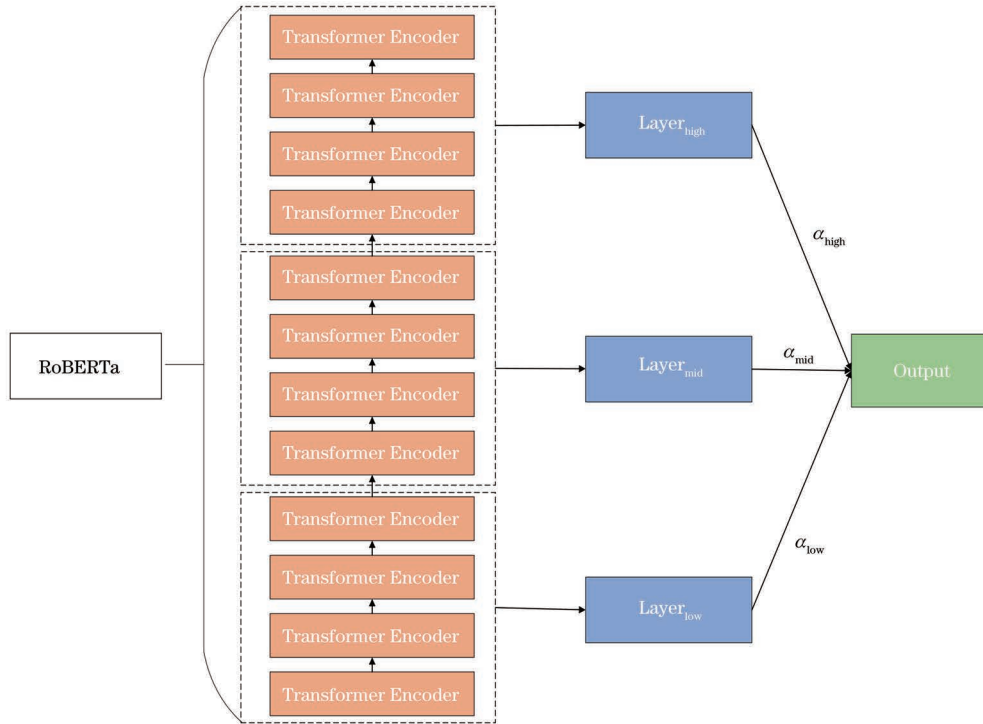


图 2 分层的 RoBERTa 结构

Fig.2 Hierarchical RoBERTa structure

3.3 FGM 对抗训练

由于煤矿井下不安全行为标注的数据样本少,且网络爬取的数据具有随机性,使得实体类别分布不均衡,如违反性行为 and 错误性行为等实体数量远远超过遗忘性行为实体数量。针对上述问题,本文采用 FGM^[22] 对抗训练算法对分层 RoBERTa 词嵌入层的词嵌入向量添加扰动,生成对抗样本,提高模型的鲁棒性和抗干扰能力。通过训练对抗样本,模型能学习到更健壮的特征表示,从而更准确地识别实体,少样本的实体类别也有较好的泛化性。对抗训练公式如下:

$$\min_{\theta} E_{x,y} [\max_{r_{adv} \in S} L(\theta, x + r_{adv}, y)] \quad (4)$$

式中: θ 表示模型的参数; E 表示对抗样本集上的期望损失; D 表示数据集; L 表示单个样本的损失; x 和 y 代表输入数据和对应的标签; r_{adv} 是扰动; S 是扰动空间。对抗训练的基本思想分为两部分,一是最大化内部损失函数,寻找合适的扰动,二是最小化任务判别损失,寻找最优的模型参数。对于 FGM 算法,首先计算嵌入层的损失和梯度,原始输入的梯度表达式如式(5)所示,然后沿着梯度提升的方向构造扰动,扰动计算如式(6)所示,将扰动与原始嵌入相加获得新的损失,梯度回传并累积,最后恢复原始嵌入,循环迭代下一批数据,利用累积后的梯度更新模型参数。

$$g = \nabla_x L(\theta, x, y) \quad (5)$$

$$r_{adv} = \epsilon g / \|g\|_2 \quad (6)$$

式中: g 是 x 的梯度; r_{adv} 是扰动; ϵ 是超参数。

3.4 BiGRU 层

由于煤矿井下不安全行为文本数据较长,本文采用 BiGRU 能够更好地捕获上下文语义信息,提取更完整的文本特征。门控循环单元 (GRU)^[23] 是在长短期记忆 (LSTM) 的基础上简化而来的,由更新门和重置门构成,与 LSTM 相比,GRU 参数少,计算复杂度低,计算速度更快。GRU 单元的计算过程如下:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \quad (7)$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \quad (8)$$

$$\tilde{h}_t = \tanh(W_h \cdot [r_t \cdot h_{t-1}, x_t] + b_h) \quad (9)$$

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t \quad (10)$$

式中:更新门和重置门分别是 z_t 和 r_t ; σ 为 Sigmoid 函数; W_z, W_r, W_h 为权重矩阵; b_z, b_r, b_h 为偏置值; x_t 为 t 时刻的输入; h_{t-1} 和 h_t 分别为上一时刻和当前时刻的隐藏状态; $\tilde{h}_t$ 为候选隐藏状态。

GRU 通过门控机制控制信息流,有助于缓解梯度消失问题,但是它是单向地从前往后传播,只获取到前文信息,不能充分获取上下文信息。所以采用双向 GRU 组成的 BiGRU 模型,从而能够捕捉序列数据的双向依赖关系,具体结构如图 3 所示。

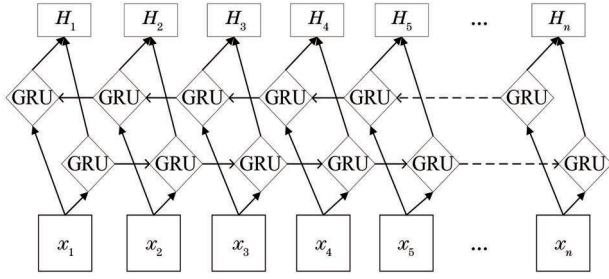


图 3 BiGRU 模型结构

Fig.3 Structure of the BiGRU model

3.5 Global Pointer 层

为了更好地处理煤矿井下不安全行为实体边界模糊问题,本文使用 Global Pointer 识别实体。Global Pointer^[24]是一种新的基于片段的 NER 框架,从全局的视角考虑实体的起始和终止位置,可以更准确地预测实体类型。当起始位置为 i 、终止位置为 j 时,坐标 $[i, j]$ 标记为 1,表示存在实体,其他标记为 0。如图 4 所示,在“爆破工违章爆破”这句话中存在两类实体,所以对应两个头部,每个头部代表一类实体类别。对于煤矿工作人员这一实体类别,实体名称为“爆破工”;而违反性行为这类别的实体是“违章爆破”。与 CRF 相比,Global Pointer 可以解决字符级标签错误,更准确地识别实体的边界。

	爆	破	工	违	章	爆	破
爆	0	0	1	0	0	0	0
破		0	0	0	0	0	0
工			0	0	0	0	0
违				0	0	0	0
章					0	0	0
爆						0	0
破							0

(a) Person

	爆	破	工	违	章	爆	破
爆	0	0	0	0	0	0	0
破		0	0	0	0	0	0
工			0	0	0	0	0
违				0	0	0	1
章					0	0	0
爆						0	0
破							0

(b) Violate

图 4 Global Pointer 识别实体示意图

Fig.4 Schematic diagram of Global Pointer identifying entity

具体来说,假设给定一个文本序列 n ,那么存在 $n(n+1)/2$ 个连续子序列,即候选实体,从候选实体中识别出实体 k ,Global Pointer 将它看作是多标签分类问题,以实体为基本单位进行判别。对于每一类实体类型 α ,分别对输入 n 的序列表征 $H \in \mathbb{R}^{n \times v}$ 变换得到 $q_\alpha = \text{MLP}(H) \in \mathbb{R}^{n \times d}$ 和 $k_\alpha = \text{MLP}(H) \in \mathbb{R}^{n \times d}$,分别表示其作为实体起始和终止位置的表征。因此,对于片段 $s[i:j]$,属于实体类型 α 的得分如下:

$$S_\alpha(i, j) = q_{i, \alpha}^\top k_{j, \alpha} \quad (11)$$

为了充分利用边界信息,Global Pointer 将 ROPE 位置编码应用到实体表示中,满足 $R_i^\top R_j = R_{j-i}$ 。因此片段 $s[i:j]$ 属于 α 的得分为:

$$S_\alpha(i, j) = (R_i q_{i, \alpha})^\top (R_j k_{j, \alpha}) = q_{i, \alpha}^\top R_i^\top R_j k_{j, \alpha} = q_{i, \alpha}^\top R_{j-i} k_{j, \alpha} \quad (12)$$

此外,由于每个句子的实体数不多,会导致类别不均衡,Global Pointer 使用了适用于总类别数多、目标类别数少的多标签分类问题的损失函数,表示如下:

$$\log_a \left(1 + \sum_{(i, j) \in P_\alpha} e^{-s_\alpha(i, j)} \right) + \log_a \left(1 + \sum_{(i, j) \in Q_\alpha} e^{s_\alpha(i, j)} \right) \quad (13)$$

式中: P_α 是所有 α 类型的实体首尾集合; Q_α 是所有非实体或者非 α 类型的实体首尾集合。

4 实验结果与分析

4.1 实验设置

本文使用 Python 3.8 和 PyTorch 1.8.1 作为实验的配置环境,最大训练样本长度为 128。模型采用 Adam 作为优化器,学习率为 2×10^{-5} 。模型训练的 batch 为 32,迭代次数 epoch 为 50。本文共构建了 2 359 条煤矿不安全行为语料进行命名实体识别,考虑到模型是小样本学习,将语料按照 7:2:1 的比例划分为训练集、验证集和测试集,各类实体数量分布如表 2 所示。

表 2 各类实体数量

实体类别	训练集	验证集	测试集
煤矿工作人员-Person	485	139	69
疏忽性行为-Careless	293	84	42
遗忘性行为-Forget	41	12	6
违反性行为-Violate	483	139	69
错误性行为-Error	520	149	75
影响因素-Factor	295	85	42
导致后果-Cause	277	80	39
预防措施-Prevent	328	94	47

4.2 评估指标

本文采用精确率 (Precision, P)、召回率 (Recall, R) 和 F1 值 (F_1 , F_1) 作为本文模型性能评价指标,各指标计算方式如下:

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \times 100\% \quad (14)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}} \times 100\% \quad (15)$$

$$F_1 = \frac{2 \times P \times R}{P + R} \times 100\% \quad (16)$$

式中: N_{TP} 表示实际为正类且被预测为正类的样本数; N_{FP} 表示实际为负类但被预测为正类的样本数; N_{FN} 表示实际为正类但被预测为负类的样本数; P 是指正类占预测结果的比例; R 是正类占实际结果的比例; F_1 是指 P 和 R 的调和平均值。

4.3 结果分析

4.3.1 模型性能对比

为了验证本文方法的有效性以及保证实验的严谨性,设计了对比实验,并且进行 3 次随机实验并取各评价指标平均值,具体结果如表 3 所示。

表 3 模型对比实验结果

Table 3 Model comparison experimental results %

模型	P	R	F1
BiLSTM-CRF ^[25]	62.03	65.25	63.60
BERT-CRF ^[26]	68.54	77.19	72.61
BERT-BiLSTM-CRF ^[27]	71.31	75.68	73.43
BERT-BiGRU-CRF ^[28]	71.54	75.97	73.69
RoBERTa-BiLSTM-CRF ^[29]	73.70	78.60	76.07
RoBERTa-BiLSTM-MLP-CRF ^[16]	74.30	78.41	76.30
BERT-Global Pointer ^[24]	76.50	77.57	77.03
分层 RoBERTa-BiGRU-Global Pointer-FGM	78.77	78.20	78.48

表 4 预测结果举例

Table 4 Examples of predicted results

序号	标注语料	预测结果
1	斜井提升时,严禁蹬钩(Prevent)	斜/O 井/O 提/O 升/O 时/O, /O 严/B-Prevent 禁/I-Prevent 蹬/I-Violate 钩/I-Violate。 /O
2	把钩工(Person)安全意识差(Factor),存在侥幸心理(Factor)	把/B-Person 钩/I-Person 工/I-Person 安/B-Factor 全/I-Factor 意/I-Factor 识/I-Factor 差/I-Factor, /O 存/O 在/O 侥/B-Factor 幸/I-Factor 心/I-Factor 理/I-Factor
3	忽视警告标志(Careless),带电作业(Violate)	忽/B-Careless 视/I-Careless 警/I-Careless 告/I-Careless 标/I-Careless 志/I-Careless, 带 B-Violate 电 I-Violate 作 I-Violate 业 I-Violate
4	电工(Person)在主井电路上操作失误(Error),导致火灾事故(Cause)	电/B-Person 工/I-Person 在/O 主/O 井/O 电/O 路/O 上/O 操/B-Error 作/I-Error 失/I-Error 误/I-Error, /O 导/O 致/O 火/B-Cause 灾/I-Cause 事/I-Cause 故/I-Cause。 /O
5	通过后忘关风门(Forget)	通/O 过/O 后/O 忘/B-Forget 关/I-Forget 风/I-Forget 门/I-Forget

4.3.2 消融实验

为验证本文所提模型的有效性以及各模块对此模型的影响程度,以分层 RoBERTa-BiGRU-Global Pointer-FGM 为基准模型进行消融实验,实验结果如表 5 所示。从表 5 结果可以看出,分层 RoBERTa-BiGRU-Global Pointer-FGM 模型比 RoBERTa-BiGRU-Global Pointer-FGM 模型在精确率和 F1 值的指标上分别提升了 0.28 和 0.22 百分点,表明融合多层语义信息比直接采用 12 层作为输出能更好地增强文本的理解。分层

由表 3 可知,相较于 BiLSTM-CRF 基线模型, BERT-CRF、BERT-BiLSTM-CRF、RoBERTa-BiLSTM-CRF、RoBERTa-BiLSTM-MLP-CRF 4 种模型的 F1 值分别提升了 9.01、9.83、12.47、12.7 百分点,说明 BERT 和 RoBERTa 预训练模型能更好地学习文本的语义关系。基于 BERT 预训练模型, Global Pointer 相较于 CRF 在 F1 值上提升了 4.42 百分点,说明 Global Pointer 能更准确地识别实体边界。分层 RoBERTa-BiGRU-Global Pointer-FGM 模型相较于 BERT-Global Pointer 模型在精确率、召回率、F1 值上分别提升了 2.27、0.63、1.45 百分点,说明融合多层语义信息能提升模型的语义理解能力,对抗训练能进一步优化模型。

测试集的部分实体预测结果如表 4 所示。序号 2、3、4、5 的预测结果与标注结果一致,主要是因为煤矿工作人员、影响因素、导致后果等这几类实体形式较为固定且规范,所以识别效果较好。在序号 1 中,实际标注时将“严禁蹬钩”作为“Prevent”类别标注,而“蹬钩”这一行为有时会在“Violate”实体中出现,比如“把钩工蹬钩拔插销”这句话中的“蹬钩拔插销”、“脚蹬绞车绳”等,实体边界不清晰,所以模型会误识别。后续会增强数据样本量以及修正数据标注精准度,进一步提升实体识别性能。

RoBERTa-Global Pointer-FGM 模型取得了较高的精确率,这是因为 FGM 增强了模型的鲁棒性,提高了预测为正例的准确性,但由于数据中存在一词多义现象,不能准确地识别出真正的正例,召回率较低,引入了 BiGRU 模块后,召回率和 F1 值指标分别提升了 4.06 和 1.29 百分点,同样, BERT-BiGRU-CRF 模型比 BERT-BiLSTM-CRF 模型的 F1 值也有一定的提升,表明 BiGRU 能更好地捕捉上下文语义特征,提高模型的实体识别效果。分层 RoBERTa-BiGRU-Global Pointer-FGM 模型比

层 RoBERTa-BiGRU-Global Pointer 模型精确率提升了 2.46 百分点,召回率提升了 1.06 百分点,F1 值提升了 1.76 百分点,说明对于小样本数据集,加入对抗训练能缓解实体类别的不均衡,增强模型的鲁棒性和泛化能力。

表 5 消融实验结果对比

Table 5 Comparison results of ablation experiments %			
模型	P	R	F1
分层 RoBERTa-BiGRU-Global Pointer-FGM	78.77	78.20	78.48
RoBERTa-BiGRU-Global Pointer-FGM(-分层)	78.49	78.04	78.26
分层 RoBERTa-Global Pointer-FGM(-BiGRU)	80.50	74.14	77.19
分层 RoBERTa-BiGRU-Global Pointer (-FGM)	76.31	77.14	76.72

5 结束语

本文构建了煤矿不安全行为文本语料库,采用 label-studio 标注工具和 BIO 标注方法完成了实体标注,其中包含 8 种实体类别。针对煤矿井下不安全行为命名实体识别任务中存在语义信息利用不足、实体分布不均衡、实体边界模糊的问题,提出一种基于 Global Pointer 和对抗训练的命名实体识别模型。提出了改进的分层 RoBERTa,融合多层的语义信息增强文本的向量化,同时对词嵌入层采用 FGM 对抗训练算法添加扰动进行对抗学习,缓解井下不安全行为语料少以及实体类别不均衡问题,提升了模型的鲁棒性。通过 BiGRU 进一步获取上下文语义特征,最后使用 Global Pointer 更准确地识别出井下不安全行为实体,提高对实体边界的识别能力。本文提出的模型在自建的小样本的煤矿井下不安全行为文本语料库上取得了较好的效果,精确率为 78.77%,召回率为 78.20%,F1 值为 78.45%,较基准模型有一定的提升。实验结果证明本文提出模型的有效性。后续将考虑增加煤矿不安全行为文本标注语料、构建领域词典或者引入外部知识库来进一步提高对不安全行为实体识别的准确性。

参考文献

[1] HANISCH D, FUNDEL K, MEVISSSEN H T, et al. ProMiner: rule-based protein and gene entity recognition[J]. BMC Bioinformatics, 2005, 6(1): 14.

[2] QUIMBAYA A P, MÚNERA A S, RIVERA R A G, et al. Named entity recognition over electronic health records through a combined dictionary-based approach[J]. Procedia

Computer Science, 2016, 100: 55-61.

[3] ZHOU G D, SU J. Named entity recognition using an HMM-based chunk tagger[C]//Proceedings of the 40th Annual Meeting on Association for Computational Linguistics. Morristown, USA: ACL Press, 2001: 473-480.

[4] TSAI R T, HUNG H, SUNG C, et al. On closed task of Chinese word segmentation: an improved CRF model coupled with character clustering and automatically generated template matching[C]//Proceedings of the 5th SIGHAN Workshop on Chinese Language Processing. Sydney, Australia: [s. n.], 2006: 1-8.

[5] ISOZAKI H, KAZAWA H. Efficient support vector classifiers for named entity recognition[C]//Proceedings of the 19th International Conference on Computational Linguistics. Morristown, USA: ACL Press, 2002: 271-278.

[6] 刘杰. 基于改进的隐马尔科夫模型的中文命名实体识别算法[J]. 太原师范学院学报(自然科学版), 2009, 8(1): 80-83, 90.

LIU J. The arithmetic of Chinese named entity recognition based on the improved nidden Markov model[J]. Journal of Taiyuan Normal University(Natural Science Edition), 2009, 8(1): 80-83, 90. (in Chinese)

[7] 胡文博, 都云程, 吕学强, 等. 基于多层条件随机场的中文命名实体识别[J]. 计算机工程与应用, 2009, 45(1): 163-165, 227.

HU W B, DU Y C, LÜ X Q, et al. Study on Chinese named entity recognition based on cascaded conditional random fields[J]. Computer Engineering and Applications, 2009, 45(1): 163-165, 227. (in Chinese)

[8] 卢青华, 袁丽娜. 基于组合神经网络的软件命名实体识别仿真[J]. 计算机仿真, 2023, 40(1): 489-492, 509.

LU Q H, YUAN L N. Software named entity recognition simulation based on combined neural network[J]. Computer Simulation, 2023, 40(1): 489-492, 509. (in Chinese)

[9] 余丹丹, 黄洁, 党同心, 等. 基于 ALBERT 的中文简历命名实体识别[J]. 计算机工程与设计, 2024, 45(1): 261-267.

YU D D, HUANG J, DANG T X, et al. Recognition of named entity in Chinese resume based on ALBERT[J]. Computer Engineering and Design, 2024, 45(1): 261-267. (in Chinese)

[10] 褚天舒, 唐球, 梁军学, 等. 基于词汇增强和表格填充的中文命名实体识别[J]. 电子技术应用, 2024, 50(2): 23-29.

CHU T S, TANG Q, LIANG J X, et al. Chinese named entity recognition based on lexicon enhancement and table filing[J]. Application of Electronic Technique, 2024, 50(2): 23-29. (in Chinese)

[11] 崔少国, 陈俊桦, 李晓虹. 融合语义及边界信息的中文电子病历命名实体识别[J]. 电子科技大学学报, 2022, 51(4): 565-571.

CUI S G, CHEN J H, LI X H. Named entity recognition for Chinese electronic medical record by fusing semantic and boundary information[J]. Journal of University of Electronic Science and Technology of China, 2022, 51(4): 565-571. (in Chinese)

[12] 林娜, 岳希, 唐聘. 基于数据增强和损失平衡的机电领域命名实体识别[J]. 计算机工程与应用, 2025, 61(7): 222-232.

LIN N, YUE X, TANG D. Named entity recognition in electromechanical field based on data enhancement and loss balancing[J]. Computer Engineering and Applications, 2025, 61(7): 222-232. (in Chinese)

[13] 曹现刚, 吴可昕, 张梦园, 等. 基于 BERT 的煤矿装备维护知识命名实体识别研究[J]. 机床与液压, 2023, 51(9): 103-108.

CAO X G, WU K X, ZHANG M Y, et al. Coal mine equipment maintenance knowledge named entity recognition model based on BERT[J]. Machine Tool & Hydraulics, 2023, 51(9): 103-108. (in Chinese)

- [14] 王向前, 李敏敏, 孟祥瑞. 基于 ALBERT-BiLSTM-CRF 的煤矿事故案例文本命名实体识别方法[J]. 阜阳师范大学学报(自然科学版), 2022, 39(3): 56-64.
WANG X Q, LI M M, MENG X R. Named entity recognition method of coal mine accident case text based on ALBERT-BiLSTM-CRF [J]. Journal of Fuyang Normal University(Natural Science), 2022, 39 (3): 56-64. (in Chinese)
- [15] 刘飞翔, 李泽荃, 赵嘉良, 等. 基于 ERNIE-BiGRU-CRF 模型的煤矿安全隐患命名实体智能识别研究[J]. 煤炭工程, 2024, 56(2): 206-212.
LIU F X, LI Z Q, ZHAO J L, et al. Intelligent recognition of named entities of coal mine safety hidden danger based on ERNIE-BiGRU-CRF model [J]. Coal Engineering, 2024, 56(2): 206-212. (in Chinese)
- [16] 付燕, 刘致豪, 叶鸥. 基于煤矿井下不安全行为知识图谱构建方法[J]. 工矿自动化, 2024, 50(1): 88-95.
FU Y, LIU Z H, YE O. A method for constructing a knowledge graph of unsafe behaviors in coal mines [J]. Journal of Mine Automation, 2024, 50 (1): 88-95. (in Chinese)
- [17] 黄辉, 张雪. 煤矿员工不安全行为研究综述[J]. 煤炭工程, 2018, 50(6): 123-127.
HUANG H, ZHANG X. Review of research on unsafe behavior of miners[J]. Coal Engineering, 2018, 50(6): 123-127. (in Chinese)
- [18] 李红霞, 樊欣怡. 人因视角下国内煤矿安全领域研究现状与发展趋势[J]. 煤炭工程, 2022, 54(1): 181-186.
LI H X, FAN X Y. Status and development trend of coal mine safety research from the perspective of human factors [J]. Coal Engineering, 2022, 54(1): 181-186. (in Chinese)
- [19] 隗昊, 刁宏悦, 孔亮宸, 等. 东北亚舆情文本细粒度命名实体识别方法研究[J]. 计算机工程, 2024, 50(5): 354-362.
WEI H, DIAO H Y, KONG L C, et al. Research on fine-grained named-entity-recognition method for public-opinion texts in Northeast Asia [J]. Computer Engineering, 2024, 50(5): 354-362. (in Chinese)
- [20] LIU Y H, OTT M, GOYAL N, et al. RoBERTa: a robustly optimized BERT pretraining approach [EB/OL]. [2024-04-20]. <https://arxiv.org/pdf/1907.11692>.
- [21] JAWAHAR G, SAGOT B, SEDDAH D. What does BERT learn about the structure of language? [C] //Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, USA, ACL Press, 2019: 232-241.
- [22] GOODFELLOW I J, SHLENS J, SZEGEDY C. Explaining and harnessing adversarial examples [EB/OL]. [2024-04-20]. <https://arxiv.org/pdf/1412.6572>.
- [23] CHO K, VAN MERRIENBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation [EB/OL]. [2024-04-20]. <https://arxiv.org/pdf/1406.1078>.
- [24] SU J L, MURTADHA A, PAN S F, et al. Global pointer: novel efficient span-based approach for named entity recognition [EB/OL]. [2024-04-20]. <http://arxiv.org/abs/2208.03054>.
- [25] HUANG Z H, XU W, YU K. Bidirectional LSTM-CRF models for sequence tagging [EB/OL]. [2024-04-20]. <https://arxiv.org/abs/1508.01991>.
- [26] 关斯琪, 董婷婷, 万子敬, 等. 基于 BERT-CRF 模型的火灾事故案例实体识别研究[J]. 消防科学与技术, 2023, 42(11): 1529-1534.
GUAN S Q, DONG T T, WAN Z J, et al. Fire accident case named entity recognition based on BERT-CRF model [J]. Fire Science and Technology, 2023, 42(11): 1529-1534. (in Chinese)
- [27] 谢腾, 杨俊安, 刘辉. 基于 BERT-BiLSTM-CRF 模型的中文实体识别[J]. 计算机系统应用, 2020, 29(7): 48-55.
XIE T, YANG J A, LIU H. Chinese entity recognition based on BERT-BiLSTM-CRF model [J]. Computer Systems & Applications, 2020, 29(7): 48-55. (in Chinese)
- [28] 王叔于, 李振华, 涂志鹏, 等. 基于 BERT-BiGRU-CRF 模型的岩土工程实体识别[J]. 地球科学, 2023, 48(8): 3137-3150.
WANG Q Y, LI Z H, TU Z P, et al. Geotechnical named entity recognition based on BERT-BiGRU-CRF Model [J]. Earth Science, 2023, 48(8): 3137-3150. (in Chinese)
- [29] LOU Q F, WANG S T, CHEN J H, et al. Named entity recognition of traditional Chinese medicine cases based on RoBERTa-BiLSTM-CRF [C] // Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine. Washington D. C., USA: IEEE Press, 2023: 4609-4614.

文字编辑 索书志
栏目编辑 宋 圆