

基于联合学习的电力安全实体抽取及风险预测方法

罗艾珂¹, 喻召杰^{2*}

(1.广东电网有限责任公司东莞供电局, 广东 东莞 523129; 2.南方电网数字平台科技(广东)有限公司, 广东 深圳 518053)

摘要:针对电力安全生产方面存在的知识分散和知识获取效率低的问题,提出基于联合学习的电力安全实体抽取及风险预测方法。采用电力安全知识数据模型与电力安全实体标签体系,对电力作业文本进行有效表示。构建基于双向编码器表征(bidirectional encoder representations from transformers, BERT)、双向长短期记忆网络(bidirectional long short-term memory, BiLSTM)、条件随机场(conditional random fields, CRF)的电力安全生产知识联合抽取模型,自动识别电力作业文本中的关键字实体并预测潜在显性风险。测试结果与样例分析证明电力安全生产知识联合抽取模型在电力实体抽取和风险的有效性,为后续电力安全知识图谱的建立提供可靠的技术支撑。

关键词:电力安全;实体抽取;风险预测;联合学习

中图分类号:TP391;TM73

文献标志码:A

引用格式:罗艾珂,喻召杰.基于联合学习的电力安全实体抽取及风险预测方法[J].山东大学学报(理学版),2026,61(5):27-37.

Methods of electric power safety entity extraction and risk prediction based on the joint learning

LUO Aike¹, YU Zhaojie^{2*}

(1. Dongguan Power Supply Bureau of Guangdong Power Grid Co., Ltd., Dongguan 523129, Guangdong, China; 2. China Southern Power Grid Digital Platform Technology (Guangdong) Co., Ltd., Shenzhen 518053, Guangdong, China)

Abstract: Aiming at the problems of dispersed knowledge and inefficient knowledge acquisition in electric power safety production, a joint learning-based method for electric power safety entity extraction and risk prediction is proposed. An electric power safety knowledge data model and a labeling system for electric power safety entities are adopted to effectively represent the text of electric power operations. A joint extraction model of electric power safety production knowledge based on bidirectional encoder representations from transformers (BERT), bidirectional long short-term memory (BiLSTM) and conditional random fields (CRF) is constructed to automatically identify the keyword entities in the text of electric power operations and predict the potential explicit risks. Test results and sample analysis prove the effectiveness of the joint extraction model for electric power work safety knowledge in electric power entity extraction and risk identification, which providing a robust and reliable technical foundation for the subsequent construction of power safety knowledge graphs.

Key words: electric power safety; entity extraction; risk prediction; joint learning

0 引言

应用先进的数字化技术提升电力行业从业人员的安全生产素质与效率,是保证安全生产的重要手段^[1]。基于人工的风险识别方式效率和精度低^[2-3],存在知识分散和知识获取效率低的问题^[4]。构建电力安全知识图谱,将分散的安全生产知识整合支撑安全知识检索,是电力安全生产面临的重要课题^[5]。知识图谱本质上是用图结构揭示实体间关系的语义网络^[6],应用于医疗^[7]、金融^[8]、法律^[9]等垂直领域。文献

收稿日期:2025-06-05; 网络出版时间:2026-03-18

基金项目:南方电网科技基金资助项目(031900KC23040017(GDKJXM20230401))

第一作者:罗艾珂(1995—),女,硕士,研究方向为安全管理。E-mail:luoaike_gz@163.com

*通信作者:喻召杰(1985—),男,高级工程师,研究方向为安全生产信息化。E-mail:380259399@qq.com

[10] 构建电力领域的中文词典与知识图谱,验证知识图谱技术在电力领域的适用性;文献[11]概述电力系统中知识图谱的知识表示、构建与应用现状,提出电力领域知识图谱应用框架设计。

实体抽取是知识图谱构建的重要内容,其目标是从自然语言文本中识别出具有特定意义的实体,再确定实体的边界与类型^[12]。实体抽取方法有基于规则或词典^[13]、基于统计机器学习的知识驱动方法以及基于深度学习、基于预训练模型的数据驱动方法^[14]。早期,实体抽取采用规则或词典的方法,依赖人工编写的规则与预定义的实体词典,简单高效、可解释性强,但维护成本高昂、泛化能力弱,适用于数据量不大的场景^[15]。统计机器学习使用特征工程训练条件随机场(conditional random fields, CRF)^[16]、支持向量机(support vector machine, SVM)^[17]等分类模型,提高泛化能力与领域适应性,但模型的性能强依赖于特征工程质量。当前,实体抽取多采用深度学习模型与预训练语言模型的组合^[18],包括卷积神经网络(convolutional neural network, CNN)^[19]、双向长短期记忆(bidirectional long short-term memory, BiLSTM)^[20]网络、双向门控循环单元(bidirectional gate recurrent unit, BiGRU)、双向编码器表征法(bidirectional encoder representations from transformers, BERT)^[21]、稳健优化的BERT预训练方法(robustly optimized bert pretraining approach, RoBERTa)^[22]等基础模型。

近年来,电力领域的实体抽取研究取得显著进展。文献[23]采用BiLSTM和CRF对电力缺陷文本进行命名实体识别,提高实体识别准确率;文献[24]采用Transformer^[25]、BiGRU和CRF模型,有效地实现了电力调度领域的命名实体识别;针对流水线方式的误差累积等问题,文献[26]提出面向电力文本的重叠实体关系联合抽取方法,构建电力领域专用的PowerRoberta预训练模型,将知识抽取准确率提升到91.67%;电力安全领域,文献[18]提出基于三元组分类的实体关系联合抽取方法,利用BERT预训练模型对电力安全文本进行向量化,并设计分类器训练模型联合抽取三元组向量, F_1 分数相比传统流水线方法提高了10%。

当前,电力安全领域的实体抽取研究面临诸多问题和挑战。首先,电力领域文本专业性较强,术语繁多,实体类别复杂^[27],通用模型表现不佳。其次,缺乏高质量的电力安全公开数据集,且人工标注成本高,需要结合大量专家知识。此外,电力安全数据具有多源异构性^[28],来源不一、格式各异,整合相对困难。

据此,本文融合BERT、BiLSTM和CRF模型,提出基于联合学习的电力安全实体抽取及风险预测方法。首先,结合电网原始数据与专家知识,设计电力安全知识数据模型,有效表示研究数据;建立电力安全实体标签体系,形式化实体输出。其次,构建安全生产知识联合抽取模型,采用BERT-BiLSTM-CRF模型进行电力文本实体抽取、BERT-BiLSTM-Sigmoid模型进行风险多标签分类。最后,构建安全生产知识训练数据集,训练并测试独立学习与联合学习模型,验证模型应对两个子任务的有效性,为电力安全知识图谱的构建提供技术基础。

1 电力安全实体识别及风险预测数据体系构建

本文旨在从电力安全文本中抽取电力安全实体、识别风险实体。为有效表示研究数据,本文设计电力安全知识数据模型,以电网数据为原始资料,从两票数据、《安全工作规程》文件、事故事件三类原始文本中抽取3类关键字与14项显性风险,数据模型如图1所示。两票数据是对专业人员进行某次作业的相关描述,包含若干属性。在两票数据的构建过程中,本文保留包括“作业计划”在内的4个原始属性,并新增“关键字”和“显性风险”2个属性。其中,关键字表示相应作业主要内容的实体,从“作业计划”文本中抽取得到,设置为电压等级、设备名称、工作内容3个方面,对应于目标输出的电力安全实体的类型。显性风险为相应作业可能产生的显性风险,设置为人身风险、设备风险、电网风险3个方面,细化为坠落、外力外物致伤等14项具体类型,对应于目标输出的预测风险的类型。

为更精准地描述电力安全与预测风险的实体类型、各类型的定义与属性,本文基于电力安全生产的特点,对电网相关数据整理、分析和归纳,抽取置信度较高的概念并将其标准化,形成电力安全实体类型标签体系,涵盖3类关键字与14项显性风险在内的若干实体类型标签,如图2所示。

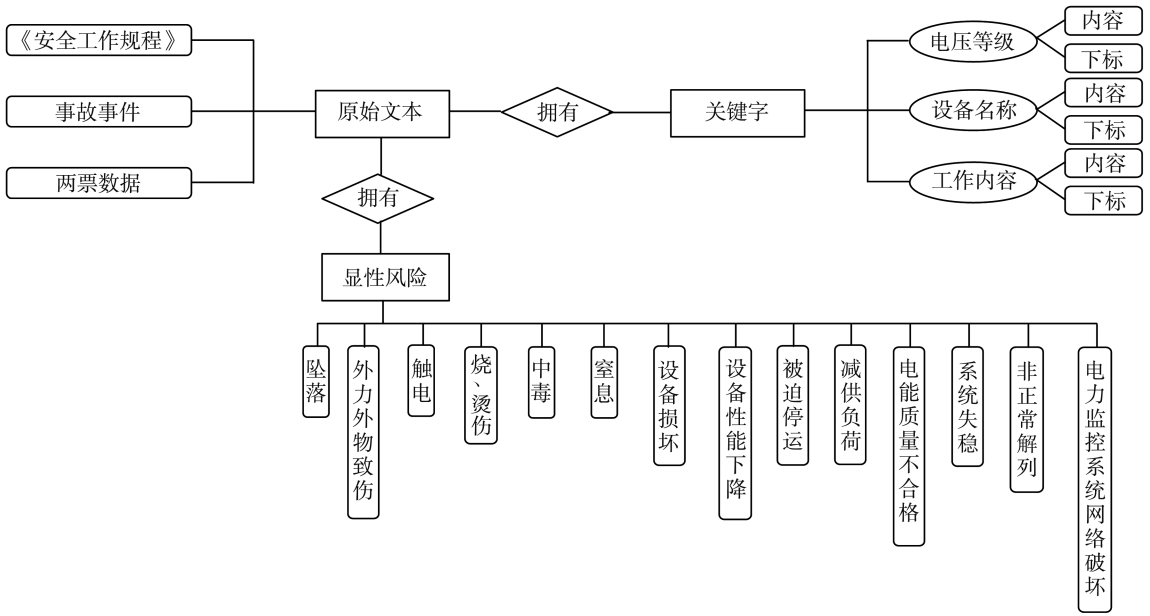


图 1 电力安全知识数据模型

Fig.1 Data model of electric power safety knowledge

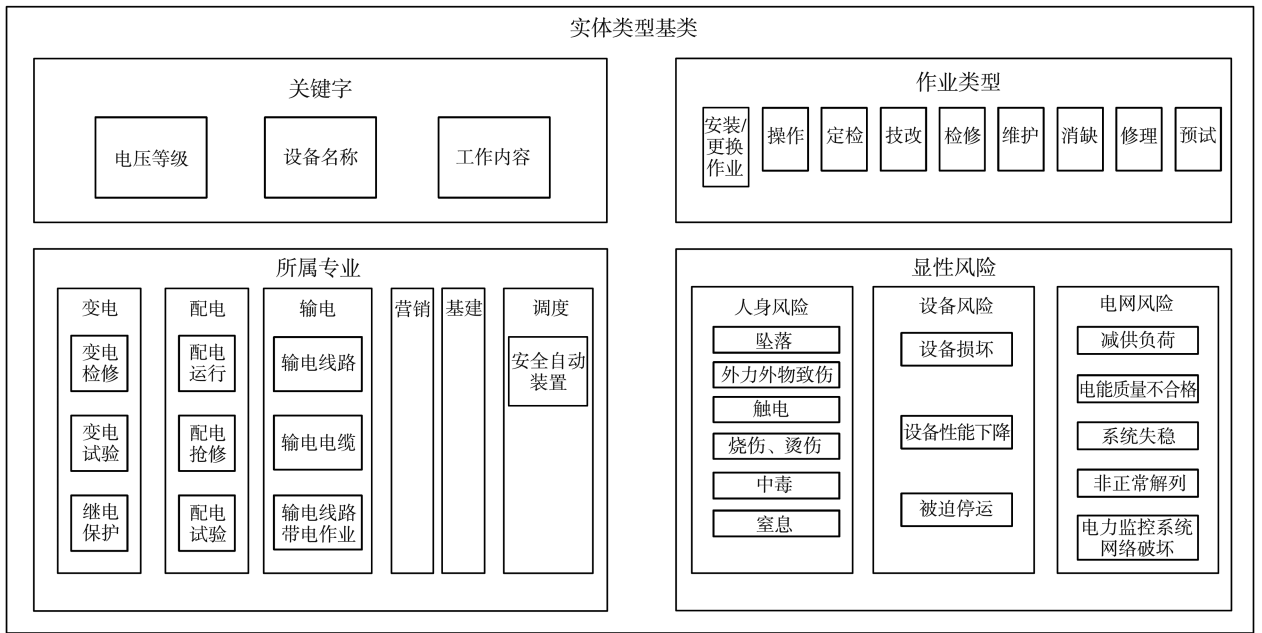


图 2 电力安全实体类型标签体系

Fig.2 Labeling system of electric power safety entities

基于以上构建的电力安全知识数据模型,给定作业计划文本,从中抽取电压等级、设备名称、工作内容 3 类关键字实体的类型与边界,并预测文本是否包含 14 类风险实体。其中,关键字实体从文本中直接提取而得,风险实体通过对文本进行间接预测得到。比如,输入为作业计划文本“500 kV 横沥站 220 kV-母线方式调整”,输出的关键字实体为[电压等级:500 kV][电压等级:220 kV][设备名称:横沥站][设备名称:母线][工作内容:调整],输出的风险标签为“减供负荷”“被迫停运”。

2 电力安全生产知识联合抽取模型

本文采用实体抽取技术识别给定的电力文本中包含的三类关键字。由于电力安全领域重点关注电力作业所产生的风险和危害^[29],因此在实体抽取的基础上加入了多标签预测技术,自动识别电力作业可能产生的显性风险。在电力文本中,抽取的关键字实体和显性风险存在一定的因果关系^[30],因此采用联合学习模型,同时学习关键字实体抽取和风险多标签预测任务。

2.1 BERT-BiLSTM-CRF 模型

本文使用 BERT-BiLSTM-CRF 模型进行电力文本的实体抽取。此部分使用基于中文的 BERT 模型 (BERT-base-chinese)^[31] 对文本进行编码。模型结构如图 3 所示。图 3 中, [CLS] 为输入序列开头的序列语义表示符号, [SEP] 为句子间的分隔符。设序列长度为 n , 嵌入向量维度为 d , 经过线性层降维后的向量维度为 d^* 。首先, 原始输入字序列 $W, W \in \mathbf{R}^{n \times d}$, 经过 BERT 模型层, 被转化为上下文表示向量序列 $H, H \in \mathbf{R}^{n \times d}$, 然后将 H 输入 BiLSTM 模型层, 学习向量的深层特征, 获取序列上下文强化表示向量序列 $H^*, H^* \in \mathbf{R}^{n \times d}$ 。接着, 将序列 H^* 送入线性层, 以缩小向量维度, 得到新的序列 $S, S \in \mathbf{R}^{n \times d^*}$ 。最后, 将序列 S 输入 CRF 模型层, 采用 BIO(begin-inside-outside) 实体标注策略, 得到条件概率最大的标签类别, 从而输出全局最优实体标签序列。

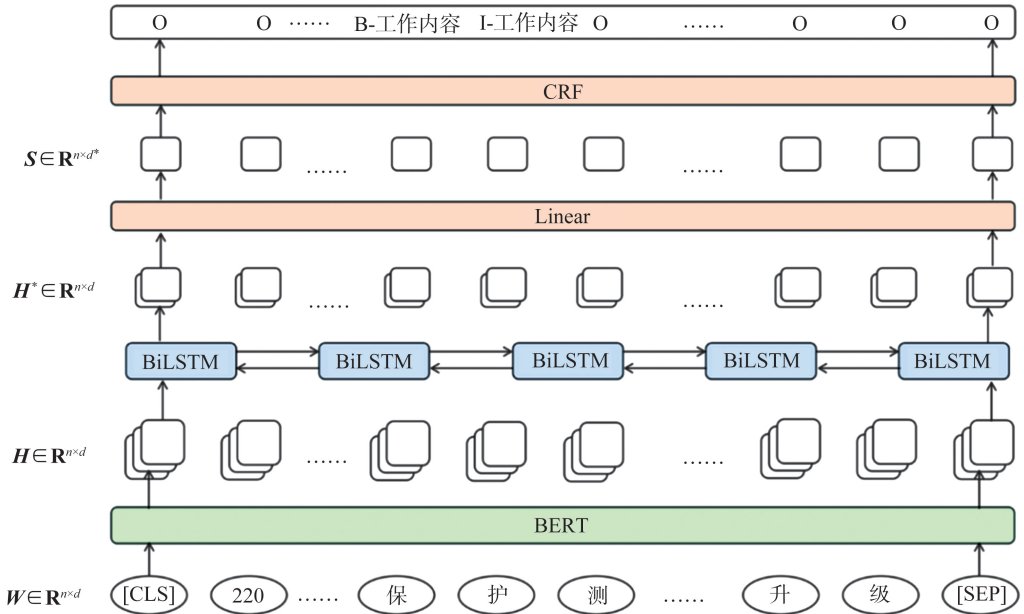


图 3 BERT-BiLSTM-CRF 模型
Fig.3 BERT-BiLSTM-CRF model

2.2 BERT-BiLSTM-Sigmoid 模型

多标签预测的任务是根据“作业计划”文本, 自动预测该作业可能产生的 14 类显性风险。本文使用 BERT-BiLSTM-Sigmoid 模型进行电力文本多标签风险分类。模型结构如图 4 所示。

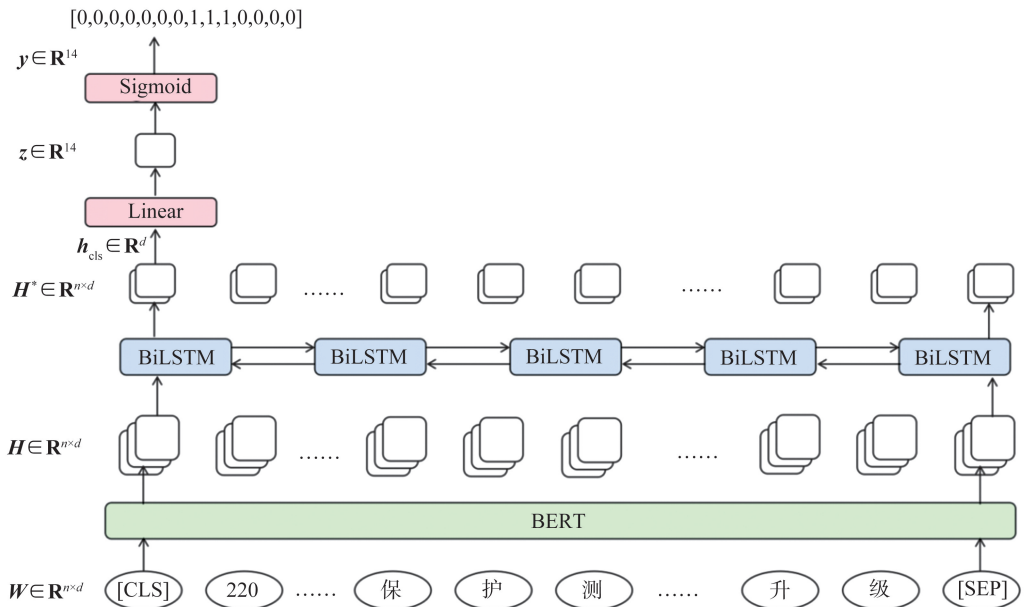


图 4 BERT-BiLSTM-Sigmoid 模型
Fig.4 BERT-BiLSTM-Sigmoid model

BERT 模型层与 BiLSTM 模型层的设计与实体抽取模型类似。在此基础上,修改线性层的设计,引入 Sigmoid 层。对 BiLSTM 层输出的向量序列 H^* ,将 [CLS] 处的向量作为整个句子的向量表示,输入线性层提取特征、缩小向量维度,最后通过 Sigmoid 函数输出概率向量 y ,阈值为 0.5,处理得到结果向量 y^* 。 z 是句子中压缩后的核心向量,即

$$z=L(h_{cls}), \tag{1}$$

式中:核心向量 h_{cls} 为向量序列 H^* 中 [CLS] 处的向量, L 为线性层函数。

概率向量 y 为

$$y=\sigma(z), \tag{2}$$

式中: σ 为 Sigmoid 激活函数。结果向量 y^* 的第 i 个分量为

$$y_i^* = \begin{cases} 1, & y_i > 0.5, \\ 0, & y_i \leq 0.5, \end{cases} \tag{3}$$

式中: y_i 为向量 y 的第 i 个分量, $i=1,2,\dots,n$ 。

2.3 电力安全生产知识联合抽取模型

从任务类型看,实体抽取和多标签预测任务本质上都是分类任务。从任务性质看,由专家经验可知,抽取得到的关键字和预测得到的显性风险之间存在因果关系。因此,在前 2 节的基础上,设计联合学习模型,同时学习关键字实体抽取和风险多标签预测任务,进一步挖掘文本的含义、关键字与显性风险间的关系。

在联合学习模型中,为更充分地利用关键字实体与显性风险间的语义关联,将 BERT-BiLSTM-CRF 模型与 BERT-BiLSTM-Sigmoid 模型设计为共享编码层的联合结构,采用参数共享机制,两者共用 BERT 模型层与 BiLSTM 模型层获得上下文的语义表示与增强表示,在此基础上,分别接入 CRF 模型层、Sigmoid 模型层,实现实体序列标注与多标签分类任务的并行学习。采用前两节所述的方法分别进行概率计算,得到实体抽取和多标签预测最终结果。联合学习的模型结构如图 5 所示。

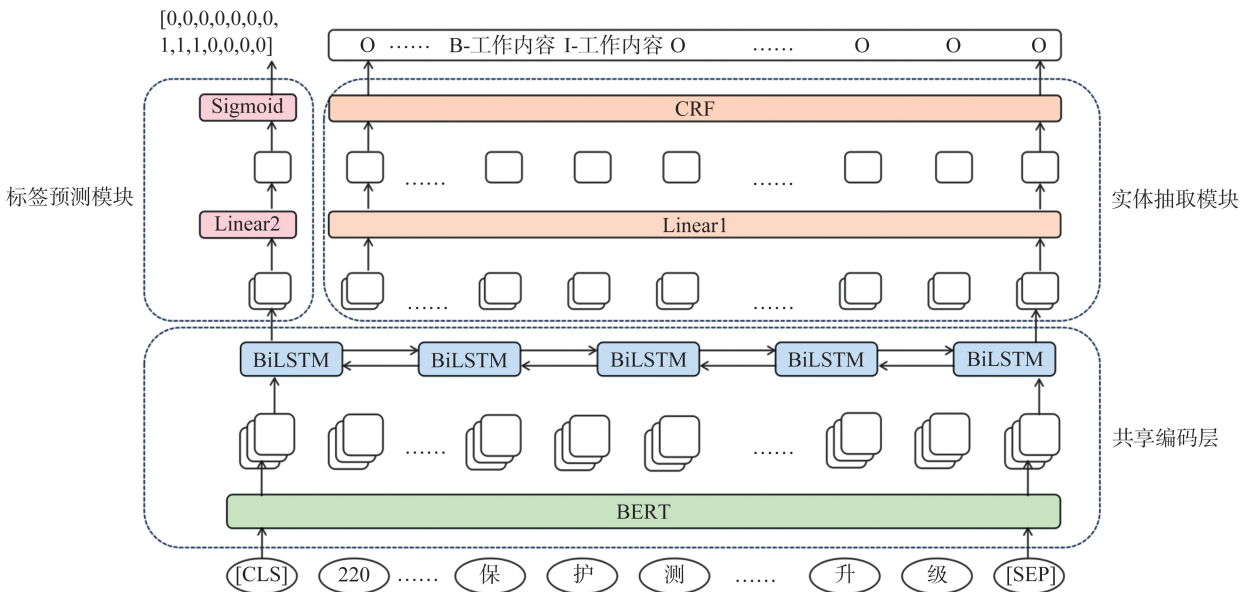


图 5 实体抽取和多标签预测联合学习模型
Fig.5 Joint model of entity and multi-label recognition

3 电力安全生产知识训练数据集的构建

实验训练数据集来自电网系统提供的两票数据。数据集构建流程分为 5 个部分,如图 6 所示。

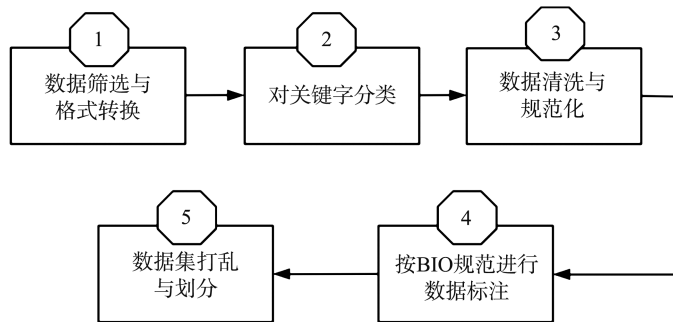


图6 实体抽取数据集构建流程

Fig.6 Construction process of entity extraction dataset

3.1 数据预处理

数据预处理是提升训练数据质量的关键步骤,目的是将原始数据转化为适合模型训练的格式。为此,首先采用正则匹配的方法,从xlsx格式的原始数据中筛选出所需字段,存储为json格式,便于后续处理。然后,根据专家制定的关键字词典,将关键字字段细分为电压等级、设备名称、工作电压3类。最后,对所有字符串类型的文本进行数据清洗和规范化,包括英文字母统一小写转换、换行符删除、空格的删除和替换、数字与单位标准化等。

3.2 数据集标注与划分

针对实体抽取,本文对两票数据采用命名实体识别中的BIO标注策略。在训练过程中,将文本按字划分,将每个字标注为“B-X”“I-X”或“O”,其中X为相应的实体标签。以电压等级为例,如果一个字位于类型为“电压等级”的实体的开头,则将其标记为“B-电压等级”;如果一个字位于类型为“电压等级”的实体的中间或结尾,则将其标记为“I-电压等级”;如果该字不属于任何实体,则标记为“O”。

针对多标签预测,本文采用独热编码形式。用14位的向量对14类显性风险标签编码,若某位为1,则表示该显性风险标签有效。

以训练数据为例:输入的作业计划文本为“#2主变保护测控装置定检”,输出的关键字实体为[工作内容:保护][工作内容:定检],输出的风险标签为“被迫停运”“设备性能下降”“减供负荷”,对应的实体标注结果如表1所示,多标签预测标注结果如表2所示。

表1 实体标注结果

Table 1 Annotation results of entities

文本数据	标注格式	文本数据	标注格式
#	O	测	O
2	O	控	O
主	O	装	O
变	O	置	O
保	B-工作内容	定	B-工作内容
护	I-工作内容	检	I-工作内容

表2 多标签预测结果

Table 2 Annotation results of multi-label prediction

风险标签	对应分量	风险标签	对应分量
坠落	0	设备性能下降	1
外力外物致伤	0	被迫停运	1
触电	0	减供负荷	1
烧、烫伤	0	电能质量不合格	0
中毒	0	系统失稳	0
窒息	0	非正常解列	0
设备损坏	0	电力监控系统网络破坏	0

最后,将处理好的两票数据随机打乱,按照8:1的比例设置训练集和验证集,其余数据作为测试集,作得到电力安全生产知识训练数据集。训练集、验证集、测试集的样本数量分别为10 000、1 250和20 119。

文献[32]表明,实体抽取模型的性能受到训练数据规模的影响,适度扩充实体样本数量和类别数量,有助于提升联合模型在边界识别和复杂语义下的表现。本研究已经基于来自变电专业的两票数据构建规模超过3万条的高质量数据集,未来可能探索基于弱监督或数据增强的方式进一步提升数据多样性与模型稳健性。

4 实体抽取和多标签预测实验

4.1 实验评价指标

实体抽取的评价指标为精确率 P 、召回率 R 和 F_1 分数。精确率 P 为

$$P = \frac{T_p}{T_p + F_p}, \quad (4)$$

式中: T_p 为被正确预测的正例个数, F_p 为原本为反例、被错误预测为正例的个数。

召回率 R 为

$$R = \frac{T_p}{T_p + F_N}, \quad (5)$$

式中: F_N 为原本为正例、被错误预测为反例的个数。

F_1 分数为

$$F_1 = \frac{2 \times P \times R}{P + R}. \quad (6)$$

本文采用的标签预测评价指标为准确率 A 。若一个文本的所有显性风险标签都预测正确,那么认为该文本预测正确,即

$$A = \frac{C}{S} \times 100\%, \quad (7)$$

式中: S 为待预测的文本总数, C 为预测正确的文本条数。

4.2 实体抽取与多标签预测实验结果

训练基于 BERT-BiLSTM-CRF 的电力文本实体抽取模型与基于 BERT-BiLSTM-Sigmoid 的多标签预测模型,在测试集上的独立学习结果如表3所示。结果显示,在20119条作业计划文本中共抽取44721个关键字实体,实体抽取的 F_1 分数为94.19%,在测试集上的多标签分类准确率为86.34%。可见,模型对于电力文本的实体抽取效果与多标签预测能力良好。

表3 独立学习实验结果
Table 3 Experimental results of independent learning

分类	$P/\%$	$R/\%$	$F_1/\%$	$A/\%$	实体数量
电压等级	96.28	96.68	96.48		3 631
设备名称	93.09	93.88	93.48		15 090
工作内容	93.95	94.59	94.27		26 000
总关键字	93.85	94.52	94.19		44 721
显性风险				86.34	

在训练集上训练实体抽取与多标签预测联合学习模型,在测试集上的结果如表4所示。结果显示,在20119条作业计划文本上共抽取44867个关键字实体,测试的 F_1 分数为94.22%;预测出38107个显性风险实体,准确率为81.48%。结果表明,目前联合学习和单独学习的效果相当,在实体抽取上性能提升不明显,在多标签预测上性能有所下降。这可能是由于联合学习模型只对原有的实体抽取和多标签预测模型进行拼接,难以学习更深层的特征。

同时,通过实验样例可知,联合学习在一定程度上修正了独立学习的实验结果。以输入文本“配合班组‘接地变保护执行新定值’相关工作布置安全措施、恢复安全措施及工作验收操作。”为例:在独立学习任务

中,输出的风险预测值为“设备性能下降”“被迫停运”;在联合学习任务中,输出的实体预测值为(设备名称,接地变)、(设备名称,保护)、(工作内容、“执行”)、(“工作内容”“定值”),输出的风险预测值为“设备性能下降”“被迫停运”“设备损坏”。可见,联合学习通过促进实体抽取与多标签预测2个子任务之间的知识共享,能够识别出独立学习中遗漏的风险预测标签。

表4 联合学习实验结果
Table 4 Experimental results of joint learning

分类	$P/\%$	$R/\%$	$F_1/\%$	$A/\%$	实体数量
电压等级	96.02	97.32	96.66		3 665
设备名称	92.52	93.85	93.18		15 178
工作内容	94.11	94.84	94.47		26 024
总关键字	93.73	94.71	94.22		44 867
显性风险				81.48	38 107

4.3 错误结果分析

如表5所示,实体抽取常见错误有两类。一是抽取的实体边界不清晰,如样例1中将“母差失灵保护屏”识别为“母差失灵保护”,样例2中模型将“***站”识别为“***变电站”。这是因为在输入的电力作业文本中,对同一实体存在多种不同表示,模型未能根据不同输入灵活识别。二是未能识别操作主体,如样例3中模型未能识别处理主体“开关”,样例4中模型未能识别处理主体“主变”。这是因为在输入的电力作业文本中,存在多种类型设备,例如在“主变本体瓦斯继电器有渗漏油”中,“主变”为设备主体,“本体”表示设备部件范围,“瓦斯继电器”表示设备保护元件,设备间存在复杂的依存关系,而模型未能很好地理解设备间的关系,导致识别出错。

表5 实体抽取结果样例
Table 5 Instances of entity extraction results

问题	样例	输入	实体抽取真实值	实体抽取预测值
识别边界不清晰	1	220 kV 母差失灵保护屏 2 原屏更换,220 kV 母差失灵保护装置 *** 恢复相关跳闸回路及电流、电压回路、核对定值及投运。	电压等级:220 kV 设备名称:*** 变电站、母差失灵保护屏 工作内容:综合改造	电压等级:220 kV 设备名称:*** 变电站、母差失灵保护 工作内容:综合改造
	2	220 kV *** 站变电站-接地变、站用变停电:*** 站用变及 *** 开关送电及相关操作。	电压等级:220 kV 设备名称:*** 站、接地变 工作内容:送电	电压等级:220 kV 设备名称:*** 站变电站、接地变 工作内容:送电
未能识别操作主体	3	110 kV *** 站:*** 开关 sf6 气体压力检查处理;*** 主变本体瓦斯继电器有渗漏油、*** 散热器上方连接处。	电压等级:110 kV 设备名称:*** 站、开关 工作内容:无	电压等级:110 kV 设备名称:*** 站、瓦斯 工作内容:漏油、处理
	4	110 kV *** 站 *** 主变油位偏低缺陷处理。	电压等级:110 kV 设备名称:*** 站、主变、油位 工作内容:无	电压等级:110 kV 设备名称:*** 站、油位 工作内容:无

注:“***”表示脱敏处理。

根据以上分析,在电力实体抽取的后续研究中可采取以下改进思路:对识别边界不清晰的问题,可以进行数据增强,强化模型对不同实体表示的学习;对未能识别操作主体的问题,可进一步细化“设备名称”这一分类,完善实体抽取的分类体系。

如表6所示,风险预测常见错误有2类。一是作业地点识别不清导致的预测错误,如样例1中“站用变”通常安装在0.5~1.5 m的基础平台上,而样例2中“气体绝缘开关设备穿墙母线筒”通常位于较高位置,可能须搭设脚手架、使用高空作业车或攀爬气体绝缘开关设备平台进行封堵,易造成坠落。模型无法识别作业中包含的隐性地点信息,导致风险预测错误。二是未能准确理解作业内容,如样例3中的作业内容为安全自动装置的专项维护,包括检查装置状态、硬件损坏程度、功能测试等,为常规的检查工作,并不一定表示设备损坏或性能下降。

表 6 风险预测结果样例
Table 6 Instances of risk prediction results

问题	样例	输入	风险真实值	风险预测值
	1	*** 站用变及 514 开关预试。	触电,设备损坏	坠落、触电
作业地点 识别不清	2	110kV *** 站:*** 主变有载开 关气体继电器更换;*** 主变变高 中性点 *** 接地刀闸 B 修;气体绝 缘开关设备靠主变侧穿墙母线筒封 堵改造。	坠落,外力外物致伤,触电	外力外物致伤,触电
未准确理 解作业内容	3	安自装置特维巡视及标签整改。	被迫停运	设备损坏,设备性能下降

注:“***”表示脱敏处理。

根据以上分析,在后续研究中可采取以下改进思路:对作业地点识别不清的问题,可构建电力安全知识图谱,将“气体绝缘开关设备穿墙母线筒”关联到“高空作业”“六氟化硫气体泄漏”等风险,通过知识图谱推理来预测风险;对未准确理解作业内容的问题,可以引入行业规范作为语义解析基准,使模型更好地理解电力知识。

5 结语

本文基于电网相关数据及专家经验,提出基于联合学习的电力安全实体抽取与多标签预测方法。通过系统地整合电力安全生产资料,设计电力安全知识数据模型,构建一套完整的电力安全实体标签体系。应对电力领域标注数据集稀缺的问题,构建全面的电力安全生产知识训练集。基于 BERT-BiLSTM-CRF、BERT-BiLSTM-Sigmoid 模型以及二者的单阶段联合学习模型实现电力文本的实体抽取和风险预测。实验结果表明,联合学习模型中关键字实体抽取 F_1 分数为 94.22%,风险预测准确率为 81.48%,证明模型在电力实体抽取与风险预测任务上的有效性。后续将采用两阶段联合学习模型与数据增强方法,指导模型更准确地抽取电力实体并预测风险。

参考文献:

[1] 黄晓丽. 大数据技术在电力安全生产管理中的风险评估及运用分析[J]. 电气技术与经济, 2023(10):231-234.
HUANG Xiaoli. Risk assessment and application analysis of big data technology in power safety production management[J]. Electrical Equipment and Economy, 2023(10):231-234.

[2] 荆朝霞,陈达鹏,谭慧娟,等. 城市输电网最大供电能力改进 Benders 算法[J]. 中国电机工程学报, 2019, 39(15):4373-4384.
JING Zhaoxia, CHEN Dapeng, TAN Huijuan, et al. Improved Benders algorithm for the maximum power supply capacity of urban transmission grids[J]. Proceedings of the CSEE, 2019, 39(15):4373-4384.

[3] 尹蕊,郭江涛,王晓磊,等. 面向智能电网的关联规则挖掘算法的设计[J]. 微型电脑应用, 2020, 36(5):85-88.
YIN Rui, GUO Jiangtao, WANG Xiaolei, et al. Design of association rules mining algorithm for smart grid [J]. Microcomputer Applications, 2020, 36(5):85-88.

[4] 阮兆文,葛军,孟干,等. 基于知识图谱的电力信息安全威胁可视化技术[J]. 信息技术, 2021(11):173-177.
RUAN Zhaowen, GE Jun, MENG Gan, et al. Visualization technology of power information security threat based on knowledge map[J]. Information Technology, 2021(11):173-177.

[5] 张晓华,刘道伟,李柏青,等. 信息驱动的大电网运行态势知识图谱框架及构建模式研究[J]. 中国电机工程学报, 2024, 44(11):4167-4181.
ZHANG Xiaohua, LIU Daowei, LI Baiqing, et al. Research on information-driven operation situation knowledge graph framework and construction mode of large power grid[J]. Proceedings of the CSEE, 2024, 44(11):4167-4181.

[6] JI S X, PAN S, CAMBRIA E, et al. A survey on knowledge graphs: representation, acquisition, and applications[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 33(2):494-514.

- [7] 何霆,吴雅婷,王华珍,等. 基于 EHR 的医疗知识图谱研究与应用综述[J]. 哈尔滨工业大学学报,2018,50(11):137-144.
HE Ting, WU Yating, WANG Huazhen, et al. A survey of medical knowledge graph based on EHR[J]. Journal of Harbin Institute of Technology, 2018, 50(11):137-144.
- [8] 胡扬,闫宏飞,陈翀. 面向金融知识图谱的实体和关系联合抽取算法[J]. 重庆理工大学学报(自然科学版),2020,34(5):139-149.
HU Yang, YAN Hongfei, CHEN Chong. Joint entity and relation extraction for constructing financial knowledge graph[J]. Journal of Chongqing University of Technology(Natural Science), 2020, 34(5):139-149.
- [9] 唐明伟,赵焯涛,李文雯. 基于词性和语序分析的法律知识图谱自动构建方法[J]. 现代信息科技,2024,8(22):85-91.
TANG Mingwei, ZHAO Huangtao, LI Wenwen. Automatic construction method of legal knowledge graph based on part of speech and word order analysis[J]. Modern Information Technology, 2024, 8(22):85-91.
- [10] 刘津,杜宁,徐菁,等. 知识图谱在电力领域的应用与研究[J]. 电力信息与通信技术,2020,18(1):60-66.
LIU Jin, DU Ning, XU Jing, et al. Application and research of knowledge graph in electric power field[J]. Electric Power Information and Communication Technology, 2020, 18(1):60-66.
- [11] 蒲天骄,谈元鹏,彭国政,等. 电力领域知识图谱的构建与应用[J]. 电网技术,2021,45(6):2080-2091.
PU Tianjiao, TAN Yuanpeng, PENG Guozheng, et al. Construction and application of knowledge graph in the electric power field[J]. Power System Technology, 2021, 45(6):2080-2091.
- [12] RAU L F. Extracting company names from text[C]//Proceedings of the 7th IEEE Conference on Artificial Intelligence Application. Miami Beach: IEEE Computer Society, 1991:29-32.
- [13] GRISHMAN R. Information extraction: techniques and challenges[C]//Information Extraction A Multidisciplinary Approach to an Emerging Information Technology. Berlin:Springer, 1997:10-27.
- [14] 肖蕾,陈镇家. 数据驱动的中文实体抽取方法综述[J]. 计算机工程与应用,2024,60(16):34-48.
XIAO Lei, CHEN Zhenjia. Review of data-driven approaches to Chinese named entity recognition[J]. Computer Engineering and Applications, 2024, 60(16):34-48.
- [15] 刘辉,江千军,桂前进,等. 实体关系抽取技术研究进展综述[J]. 计算机应用研究,2020,37(增刊2):1-5.
LIU Hui, JIANG Qianjun, GUI Qianjin, et al. Review of research progress of entity relationship extraction[J]. Application Research of Computers, 2020, 37(Suppl.2):1-5.
- [16] LAFFERTY J, MCCALLUM A, PEREIRA F. Conditional random fields: probabilistic models for segmenting and labeling sequence data[C]//Proceedings of the Eighteenth International Conference on Machine Learning. San Francisco: Morgan Kaufmann Publishers Incorporation, 2001:282-289.
- [17] CORTES C, VAPNIK V. Support-vector networks[J]. Machine Learning, 1995, 20(3):273-297.
- [18] 张燎原,李英娜. 基于三元组分类器的电力安全领域知识抽取[J]. 电力科学与工程,2024,40(6):11-18.
ZHANG Liaoyuan, LI Yingna. Triple extraction based on classifier of triplet in the field of power safety[J]. Electric Power Science and Engineering, 2024, 40(6):11-18.
- [19] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.
- [20] GRAVES A, SCHMIDHUBER J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures[J]. Neural Networks, 2005, 18(5/6):602-610.
- [21] 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. Minneapolis: Association for Computational Linguistics, 2019:4171-4186.
- [22] LIU Z, LIN W, SHI Y, et al. A robustly optimized BERT pre-training approach with post-training[C]//Proceedings of the 20th Chinese National Conference on Computational Linguistics. Lanzhou: Chinese Information Processing Society of China 2021:1218-1227.
- [23] LI Jianbin, FANG Suwan, REN Yuqi, et al. SWVBIL-CTF: selectable word vectors-based bilstm-crf power defect text named entity recognition[C]//Proceedings of 2020 IEEE International Conference on Big Data. Huhhot: Chinese Information Processing Society of China, 2020:2502-2507.
- [24] 吴超,王汉军. 基于 GRU 的电力调度领域命名实体识别方法[J]. 计算机系统应用,2020,29(8):185-191.
WU Chao, WANG Hanjun. Named entity recognition in electric power dispatching field based on GRU[J]. Computer

Systems & Applications, 2020, 29(8):185-191.

- [25] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]// Proceedings of the 2017 Conference on Advances in Neural Information Processing Systems. Redhook; Curran Associates Incorporation, 2017:6000-6010.
- [26] 束嘉伟,杨挺,耿毅男,等. 面向电力知识图谱构建的重叠实体关系联合抽取方法[J]. 高电压技术, 2024, 50(11):4912-4922.
- SHU Jiawei, YANG Ting, GENG Yinan, et al. Joint extraction method for overlapping entity relationships in the construction of electric power knowledge graph[J]. High Voltage Engineering, 2024, 50(11):4912-4922.
- [27] 高海翔,苗璐,刘嘉宁,等. 知识图谱及其在电力系统中的应用研究综述[J]. 广东电力, 2020, 33(9):66-76.
- GAO Haixiang, MIAO Lu, LIU Jianing, et al. Review on knowledge graph and its application in power systems[J]. Guangdong Electric Power, 2020, 33(9):66-76.
- [28] 徐鸿飞,李英娜. 利用 Bert 模型实现电力安全规程数据中实体关系抽取[J]. 电力科学与工程, 2023, 39(1):44-51.
- XU Hongfei, LI Yingna. Realization of entity relationship extraction from electric power safety regulation data by Bert model[J]. Electric Power Science and Engineering, 2023, 39(1):44-51.
- [29] 秦浩,徐敏,张永梅,等. 基于人工智能技术的电力安全风险控制研究[J]. 自动化技术与应用, 2024, 43(7):84-88.
- QIN Hao, XU Min, ZHANG Yongmei, et al. Research on power safety risk control based on artificial intelligence technology[J]. Control Theory and Applications, 2024, 43(7):84-88.
- [30] 齐冬莲,闫玮丹,闫云凤,等. 面向电力调度的事件知识图谱研究现状及发展[J]. 电子与信息学报, 2024, 46(9):3456-3466.
- QI Donglian, YAN Weidan, YAN Yunfeng, et al. A review of research methods on event knowledge graph for power dispatching[J]. Journal of Electronics & Information Technology, 2024, 46(9):3456-3466.
- [31] CUI Yiming, CHE Wanxiang, LIU Ting, et al. Pre-training with whole word masking for chinese BERT[J]. IEEE-ACM Transactions on Audio Speech and Language Processing, 2021, 29:3504-3514.
- [32] LI Jing, SUN Aixin, HAN Jianglei, et al. A survey on deep learning for named entity recognition[J]. IEEE Transactions on Knowledge and Data Engineering, 2022, 34(1):50-70.

(编辑:陈丽萍)

(上接第 26 页)

- [44] 李洪海. 基于大数据分析的电力行业安全生产隐患识别研究[J]. 现代职业安全, 2025(3):14-17.
- LI Honghai. Research on the identification of safety production hazards in the power industry based on big data analysis[J]. Modern Occupational Safety, 2025(3):14-17.
- [45] 张燎原,李英娜. 基于三元组分类器的电力安全领域知识抽取[J]. 电力科学与工程, 2024, 40(6):11-18.
- ZHANG Liaoyuan, LI Yingna. Knowledge extraction in the field of power safety based on triple classifier[J]. Electric Power Science and Engineering, 2024, 40(6):11-18.
- [46] WANG Yucheng, YU Bowen, ZHANG Yueyang, et al. TPLinker: single-stage joint extraction of entities and relations through token pair linking[C]// Proceedings of the 28th International Conference on Computational Linguistics (COLING 2020), 2020:1572-1582. <https://arxiv.org/abs/2010.13415>
- [47] WEI Zhepei, SU Jianlin, WANG Yue, et al. A novel cascade binary tagging framework for relational triple extraction[C]// Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL 2020), 2020:1476-1488. <https://aclanthology.org/2020.acl-main.136/>
- [48] 李媛,张志荣,杨晶,等. 基于深度学习算法构建电力安全知识学习系统[J]. 微型电脑应用, 2023, 39(4):164-167.
- LI Yuan, ZHANG Zhirong, YANG Jing, et al. A power safety knowledge learning system is constructed based on deep learning algorithms[J]. Microcomputer Applications, 2023, 39(4):164-167.

(编辑:陈丽萍)