

命名实体识别方法及在电力领域的应用

张勇^{1,2}, 纪伟^{1,3}, 钟毅^{1,3}

(1.深圳大学电子与信息工程学院, 广东 深圳 518060; 2.广东省智能信息处理重点实验室, 广东 深圳 518060; 3.南方电网数字平台科技(广东)有限公司, 广东 深圳 518053)

摘要:领域命名实体识别方法在大语言模型技术赋能的背景下,推动了电力领域命名实体识别的发展。在此背景下,本文综述电力领域命名实体识别方法的发展历程,介绍早期基于规则和词典的方法再到统计机器学习的方法。从分布式嵌入层、文本编码层和标签解码层总结基于深度学习方法的模型。本文还探讨大语言模型在命名实体识别任务中的应用及其影响,并且探索当前电力领域命名实体识别存在的问题。

关键词:领域命名实体识别; 电力领域; 大语言模型; 机器学习; 深度学习

中图分类号:TP391; TM73 **文献标志码:**A

引用格式:张勇,纪伟,钟毅. 命名实体识别方法及在电力领域的应用[J]. 山东大学学报(理学版), 2026, 61(5): 1-17.

Methods of named entity recognition and applications in electric power domain

ZHANG Yong^{1,2}, JI Wei^{1,3}, ZHONG YI^{1,3}

(1. Shenzhen University, College of Electronic and Information Engineering, Shenzhen 518060, Guangdong, China; 2. Guangdong Provincial Key Laboratory of Intelligent Information Processing, Shenzhen 518060, Guangdong, China; 3. Southern Power Grid Digital Platform Technology (Guangdong) Co., Ltd., Shenzhen 518053, Guangdong, China)

Abstract: The development of named entity recognition (NER) in the power sector has been propelled by domain-specific NER methods under the empowerment of large language model technology. In this context, the evolutionary journey of NER methods within the power domain is reviewed in this paper, where early approaches based on rules and dictionaries are introduced, followed by statistical machine learning methods. Deep learning-based models are summarized from the perspectives of the distributed embedding layer, the text encoding layer, and the label decoding layer. The application of large language models to NER tasks and their impact are also examined. Furthermore, the existing challenges currently faced by power domain NER are explored. Finally, an outlook on future research directions is presented.

Key words: domain named entity recognition; power domain; large language model; machine learning; deep learning

0 引言

随着信息技术与智能电网的深度融合,电力领域中产生海量的结构化与非结构化的电力数据,包括报告、维护记录、安规、两票等文本信息,这些文本信息蕴涵了大量的电力专业实体,如发电厂名称、作业步骤、电力设备型号、故障类型、电压等级等。这些实体信息构成了电力领域的核心知识基础,对于电力系统的运维管理、故障诊断、任务调度、作业人员安全都有至关重要的作用。然而,人工整理和分析这些庞大而复杂的数据不仅耗时费力,而且容易产生误差。因此,自动化处理这些文本数据的技术需求日益凸显。

收稿日期:2025-06-03; 网络出版时间:2025-12-25

基金项目:广东省重点实验室基金资助项目(2023B1212060076); 深圳市科技计划基金资助项目(KJZD20230923114405012); 南方电网科技基金资助项目(031900KC23040016(GDKJXM20230399), 031900KC23040017(GDKJXM20230401))

第一作者:张勇(1976—),男,教授,博士,研究方向为人工智能技术. E-mail: yzhang@sdu.edu.cn

命名实体识别^[1](named entity recognition, NER)方法作为自然语言处理(natural language processing, NLP)领域的方法之一。自动检测文本中特定的命名实体,归类至预设的实体类别,命名实体识别方法与电力系统、智能问答系统、知识图谱构建等应用有密切联系。电力领域命名实体识别(power domain named entity recognition, PDNER)方法作为一种关键技术手段,用于自动识别电力领域特定实体,实现从原始文本数据中提取关键信息。不同于通用领域的实体识别,电力领域涉及大量的专业实体,如设备名称、变压器型号、输电线路、电压等级等,这些实体具有极强的专业性和精确性。此外,电力数据的来源多样,既有结构化的数据库记录,也有大量非结构化的文档、报告和维护记录,这要求命名实体识别方法不仅能处理长文本,还能理解专业图表和表格中的信息。因此,PDNER方法在电力系统的智能化运维中扮演着核心角色,它能够快速准确地从海量数据中提取关键信息,用于设备状态监测、故障预测、安全合规检查等场景,显著提升电力企业的运营效率和安全性。随着深度学习和大模型技术的成熟,PDNER方法正逐步克服实体识别的准确性、泛化能力和实时性等难题,通过精准识别电力文本中的实体,不仅有助于构建高质量的电力知识图谱,也有利于实现智能的电力信息系统和服务,促进电力行业的数字化转型和智能化升级。

本文总结PDNER在电力领域的主要应用场景和技术挑战,从数据预处理、模型架构优化、多模态信息融合等多个维度梳理当前研究的最新进展。在此基础上,本文进一步分析PDNER方法在实际应用中的关键需求和现存问题,例如领域适应性不足、标注数据稀缺以及推理效率低下等。结合当前的研究趋势和技术发展方向,本文探讨基于预训练大模型的解决方案在PDNER中的应用潜力,以及如何通过技术创新进一步推动PDNER方法向实用化、高效化方向发展,为电力系统的智能化升级提供更加坚实的技术支撑。

1 传统命名实体识别方法

1.1 基于规则和词典的方法

在电力领域,实体类型涵盖从电力设备、技术术语到行业标准等一系列复杂概念。基于规则和词典的方法^[2]是命名实体识别中最早使用的方法,该方法依赖语言学专家和领域专家编制的语法规则和词汇表征,这些规则往往围绕特定领域特有的语法结构和词汇搭配,通过模式匹配的方式辨识不同类型的实体。尽管该方法在初始阶段能快速部署,但局限性在电力文本的复杂性面前显露无疑。李新鹏等^[3]基于D5000系统专家经验抽取智能电网调度控制系统基础平台的实体集合,获取系统业务知识,在处理通用领域内的实体时表现良好,但在电力领域不断演进的技术和设备更新中,规则和词典的维护变得日益困难。王慧芳等^[4]通过构建电力领域本体字典进行文本信息匹配,识别实体信息,然而电力文本的多样性和专业性要求词典必须持续更新,以适应新出现的实体类型和语境变化。实际上,在面对复杂且形态各异的文本时,不同规则体系可能出现相互矛盾的情况^[5],难以覆盖所有的规则。袁金斗等^[6]基于实体词典构建用电安全实体词典、基于规则模板构建实体识别规则库,通过规则模板和词典进行匹配识别,在测试样本较少的情况下,能够识别出大多数命名实体。然而在面对海量的用电安全文本时,可能出现词典信息滞后、规则模板冗余的问题,导致识别准确率不高。显然,基于规则和词典的命名实体识别方法虽然能够在特定的语料上获得不错的识别效果,但是在电力领域中,命名实体识别对规则的精准度和全面性提出了更高要求,而手工规则的构建和维护成本高昂,且难以满足跨系统应用的可移植性。

1.2 基于统计机器学习的方法

21世纪初,基于统计机器学习的NER方法将重心转移到统计概率上,增强模型在不同情境下的泛化能力,提高方法的可移植性。传统的机器学习途径实质上是将统计学习原理融入机器学习框架之内,借助人工精心筛选和构建的特征集刻画每个样本的属性,揭示潜藏在数据背后的共同规律。这种方法的关键在于针对特定领域选择适宜的模型架构,并优化训练参数。相比于基于规则和词典的NER方法,统计机器学习方法摒弃了大量细致繁琐的手工规则编纂过程,用较短时间对人工标注的语料库训练,有效地提升训练效率。对于跨领域的各类问题,只需依据相应领域标记好的训练数据集重新训练模型^[7],实现模型的迁移和适应,序列标注方法是目前最普遍且最有效的方法,经典的序列标注有隐马尔可夫模型(hidden markov model, HMM)^[8]、决策树(decision tree, DT)^[9]模型等。基于统计机器学习的NER方法在电力领域的应用,如刘梓

权等^[10]结合电力设备缺陷记录常用词词典和 HMM 进行分词,对电力专业词典匹配,实现实体识别。孙玉芹等^[11]将 HMM 作为初级学习器,通过该初级学习器剔除测试数据中的冗余信息,提取关键信息,识别电力运检命名实体。潘晖等^[12]建立决策树训练集,对不同电力客户的需求信息关联识别。然而,尽管基于统计机器学习的 NER 方法能显著提升 NER 电力领域的应用效果,但它们依然受限于高质量标注数据的稀缺性和工程特征的复杂性。电力文本特性要求模型具备强大的领域适应性,而收集足够的标注数据以及设计有效的特征对于非技术背景的研究人员来说是一大挑战^[13]。各相关模型的优缺点如表 1 所示。

表 1 传统命名实体识别方法总结

Table 1 Summary of traditional named entity recognition methods

方法	模型	优点	缺点
基于规则和词典的方法	D5000 系统+专家经验 ^[3]	能够进行快速部署,适用于特定领域	D5000 系统的业务种类繁多,实体较多且动态变化,存在处理效率较慢的问题
	构建电力本体词典 ^[4]	实体识别准确率较高	依赖专家知识和经验,对电力本体词典构建质量要求较高、维护成本较高
	词典+规则模板 ^[6]	在少样本情况下,实体识别准确率高	面对海量的用电安全文本时,可能出现词典信息滞后、规则模板冗余的问题
基于统计机器学习的方法	词典+HMM ^[10]	相较于词典和规则的方法具有更高的实体识别精度	随着数据的增加,将出现由词频过低、实例过少导致的数据不平衡问题
	HMM ^[11]	构造 HMM 初级学习器能够剔除冗余信息,从而更好关注关键实体信息	对实体类别的边界识别精准度不高
	决策树 ^[12]	识别电力客户需求准确率较高	无法结合用户标签对变量模糊的地方进行区分

2 基于深度学习命名实体识别方法

相较于传统方法,深度学习能够自动学习并利用复杂的特征组合,显著减少人工特征工程的需求。具体来说,深度学习模型通过自动学习的分布式表示,捕捉电力领域专有名词之间的细微差别及上下文依赖关系,这对于识别电力设备、变电站、输电线路等实体至关重要。

在电力领域中,深度学习框架下的命名实体识别系统通常包含 3 个层次:分布式嵌入层、文本编码层以及标签解码层。分布式嵌入层利用词嵌入技术(如 Word2Vec(word to vector)、GloVe(global vectors for word representation))将电力文本转换为低维稠密向量表示,这些向量能够捕捉到词汇的语义信息以及在上下文中扮演的角色。文本编码层采用先进的编码模型,如长短时记忆网络(long short-term memory, LSTM)、门控循环单元(gated recurrent unit, GRU)、双向 LSTM(BiLSTM)或 Transformer-based 模型整合上下文依赖关系。这个过程对于理解电力领域中实体的复杂性和多变性至关重要。标签解码层使用标签解码器(如条件随机场^[14](conditional random field, CRF)等其他序列标注模型),通过对编码层生成的上下文感知向量序列解码,推断最符合原文语境的命名实体标签序列。

为了进一步提升模型在电力领域的适用性和性能,研究者们还引入了一系列辅助技术和策略。例如,注意力机制能够帮助模型更有效地聚焦与电力实体识别相关的上下文信息,李嘉皓等^[15]引入注意力机制捕获电力设备缺陷实体之间的复杂依赖关系;多任务学习允许模型同时处理多个相关任务,有助于模型共享特征表示,提高整体性能,纪鑫等^[16]利用多任务学习框架结合命名实体识别和关系抽取 2 个任务的学习,实现实体识别和关系抽取的互相补充和互相促进,提高电力故障文本信息抽取的性能。数据增强技术通过增加额外的训练样本(如通过同义词替换、噪声注入等技术生成的新样本),增强模型对电力领域专业术语的识别能力,田雪涵等^[17]利用实体词袋替换的数据增强技术扩大原始数据集,而集成学习则结合多个模型的预测结果,提高最终实体识别的准确性和稳定性,孙玉芹等^[11]采用 3 种模型训练预测,再将预测结果输入 CRF 中训练,经过 2 层模型训练识别电力运检命名实体。

显然,深度学习不仅显著提升了电力领域命名实体识别的性能,还为处理该领域特有的复杂文本提供了强大的工具。这种架构设计不仅提高了实体识别的准确性,还具备较强的灵活性和扩展性,能够随着技术的发展和电力行业需求的变化而不断优化和完善。

2.1 分布式嵌入层

在 NLP 领域中,神经网络无法直接处理原始文本数据,因此要借助分布式嵌入层技术将原始数据转化为低维连续实数向量。在这个经过学习的向量空间里,每个维度都蕴含某种潜在的语义或句法特性。当文本通过分布式嵌入层的映射时,系统自动生成能够体现词汇意义及语法特点的向量表达,使得后续的神经网络模型能够更有效地抽取和理解文本携带的信息特征。以电力领域应用场景为例,假设输入层收到一个简单的示例序列“发电机组”。首先,将这个输入序列拆分为单个字符:发、电、机、和、组。再利用从电力领域语料库中的词汇表,将这些字符映射为向量。通过这种方式,即使是简单的词汇也被转换成包含丰富语义信息的向量,有助于后续模型更好地理解和处理电力领域的专业术语和概念。

文本中的词最初由独热编码表示词向量,编码过程是:根据语料库构建包含所有唯一词的词汇表,每个词对应唯一的索引,对于每一个词,根据其位置信息,分配一个维度与词汇表大小相等的向量,实现文本编码。虽然独热编码的实现简单,并且在处理离散变量时效果好,然而,独热编码的各个词向量之间相互独立,仅由词在词汇表中的唯一索引决定,导致向量稀疏性和离散性,进而忽略词汇之间的语义关联,形成“语义鸿沟”^[18]。因此独热编码无法高效地表示词之间的联系。

然而,预训练词向量模型如 Word2Vec^[19]、Glove^[20] 模型能够弥补这个缺点。Word2Vec 能将海量无标签文本数据转化为低维、密集型的词向量。通过 Word2Vec 模型,词与词之间的语义相似性在向量空间中通过向量间的距离体现,实质上是基于浅层神经网络架构的模型,对自然语言数据特征提取和表示学习。杜修明等^[21]结合 Word2Vec 和负采样技术将电力故障案例句子训练为词向量。蒋逸雯等^[22]采用 Word2Vec 将电力变压器运维过程中形成的两票等文本用低维稠密词向量表示,实现语义信息的自动学习。Glove 模型是对 Word2Vec 的改进和补充,结合全局统计信息和局部上下文窗口的优势,通过学习词共现矩阵得到低维词向量,核心思想是:词之间的共现频率能够反映词之间的语义关系,利用矩阵分解技术挖掘隐藏在共现数据中的潜在语义结构。李强等^[23]在句法特征提取层采用 Glove 模型对配电网运行信息进行扫描,分解成词之间的共现句子,获取相应的词向量。张宇波等^[24]将 Glove 模型作为词向量嵌入模型,选择左右对称窗口训练词向量,将电力设备缺陷文本映射到数值空间表示语义。此外,对比 Word2Vec 和 Glove 词向量映射模型,验证 Glove 模型对不同语义的词汇空间距离更大,能够更有效地区分语义相差较大的专业词汇的结论。Wei 等^[25]提出一种位置和相似性感知的 NER 模型,特别针对电力设备维护工单,嵌入层设计考虑词汇的位置信息和相似性特征,更好地捕捉工单文本中的实体特性。Cong 等^[26]在基于词汇增强和全局指针的电力数据 NER 方法中,强调词汇信息增强在分布式表示中的重要性。

2.2 文本编码层

在文本编码阶段,对分布式嵌入层的向量化,通过文本编码器深层次的处理,能够捕获并整合词汇的上下文关联性,生成具有丰富语义内涵的词向量表达。尽管深度学习模型,在公共数据集的多种 NLP 任务中表现出色,但它们在电力领域的应用效果尚未深入研究。电力数据具有独特的特征和特定领域的细微差别,要求设计和微调模型。

在电力系统的背景下,数据环境复杂,包括来自多源、多模式和多样化的数据格式。这种异构性导致了语义表示的多样性,使得数据的集成、共享和挖掘变得尤为困难。因此,构建适合电力领域的知识图谱显得格外重要,为整合和组织特定领域的知识提供统一的框架。文本编码层中的 NER 模型具备从大量的电力数据中准确提取关键有效信息的能力。根据主体网络结构的不同,将 NER 任务的文本编码层划分为 2 大主流类:一类是基于 LSTM 架构的模型,另一类是基于 Transformer 架构的预训练语言模型。这 2 类模型因其强大的上下文理解能力和预训练得到的丰富知识在 NER 任务中展现出强大的性能。

2.2.1 基于 LSTM 架构的模型

将文本转化为向量之后,应用神经网络抽取向量化文本特征。基于 LSTM 的神经网络模型擅长捕捉和处理序列数据中的长期依赖关系与上下文信息,避免循环神经网络(recurrent neural network, RNN)在捕获远距离信息时产生的梯度消失问题。因此,基于 LSTM 的神经网络模型在诸多 NLP 任务及应用场景中得到了广泛应用。在电力领域的 NER 任务中,双向结构的 LSTM 网络和 GRU 网络能充分利用序列前后方向的信息。双向 LSTM(BiLSTM)网络结合前向与后向传播的 LSTM 结构(如图 1),输出层中每个 x_t 代表时间步 t 的输入数据, y_t 则是双向 LSTM 在时刻 t 对输入 x_t 的编码结果,它综合了前向和后向的语境信息,实现

电力文本在两个时间方向的特征提取,丰富特征空间。

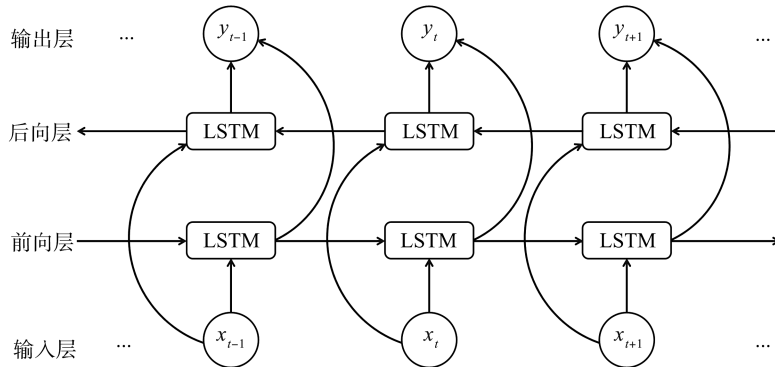


图 1 BiLSTM 结构

Fig.1 BiLSTM structure

文献[27-30]均采用 BiLSTM-CRF 模型,有效识别电力实体,实现较高的识别准确率。肖勇等^[31]结合电力计量词典和中文分词模型对电力百科知识文本和电力计量文本分词,基于字符向量采用 CNN-BiLSTM 联合神经网络提取特征,解决人工构建特征成本高的问题。江叶峰等^[32]有所侧重地提取电网故障处置预案文本中的关键实体,采用 Attention-BiLSTM-CRF 模型准确识别预案文本中细粒度的实体。为了解决电力领域实体类型多、专业程度高、存在大量嵌套实体等问题,徐会芳等^[33]提出 BiLSTM-CNN-CRF 的方法,能够更有效提取字符特征,快速迁移到不同类型的实体标注中。毛宏亮等^[34]提出 BiLSTM-MHATT-CRF 方法,融入多头注意力机制,对 BiLSTM 网络生成的上下文语义特征聚焦,强化对关键信息的捕捉与挖掘,实现电网调度领域实体识别,准确率为 93.62%。Meng 等^[35]提出基于多源嵌入的 NER 模型用于构建知识图谱,并将其应用于电网现场操作违规分析。该模型融合了多种嵌入信息,说明该模型处理复杂电力系统操作文本时的有效性。

相较于 BiLSTM 模型,双向 BiGRU 模型凭借其独特的门控机制,能够高效地提炼输入数据中的全局特征,此外,双向 BiGRU 模型训练效率方面表现出色,得益于其简洁的网络结构设计,因此训练时间较 BiLSTM 模型较少,降低了计算资源的消耗。在参数量方面也有所减少,降低了模型的复杂度。例如,李强等^[22]使用的 BiGRU-GCN 模型,在配电网相关技术文献的任务中,精确度为 93.60%,召回率为 88.40%。而吴超等^[36]的 Transformer-BiGRU-CRF 模型,在电力调度数据上的精确度为 69.71%,召回率为 70.23%,F1 为 69.97%。另外,宋厚岩等^[37]融合 BiGRU 模型和脉冲耦合神经网络(pulse-coupled neural network, PCNN)模型的优越特性,采用 BiGRU-PCNN 模型准确捕获长距离依赖关系,又能提取局部关键特征,实现电力专业知识的有效识别。实验结果证明,混合模型相比于传统的单一模型,训练效果更佳。

2.2.2 基于 Transformer 架构的预训练语言模型

预训练语言模型(pre-trained language model, PLM)在命名实体识别任务中具有优势。这类模型利用大规模无监督数据预先训练通用的语言模型,在面对具体任务时,仅微调参数即可实现高效复用。相较于传统的深度学习模型,PLM 不仅大幅缩短了训练周期,而且在各类下游任务上的表现普遍超越了传统模型。

特别在电力领域,许多术语具有相似的结构,要求利用特征向量准确地识别实体边界。例如,在处理术语“低压电流互感器”时,利用分布式嵌入层,理解该术语为整体的技术含义,而不会误分为“低压电流”和“互感器”两个独立的词汇。这种精确的识别对于正确理解和处理电力领域的专业术语至关重要。

为了使模型能够动态选择词特征向量和字符特征向量的语义信息,利用 PLM 的动态上下文感知能力在此发挥了关键作用。它不仅能够捕获词汇在不同语境下的丰富语义信息,并处理词汇形态变化(如同义词、派生词等)。早期的词嵌入模型如 Word2Vec、Glove 等产生的词向量可视为“静态向量”,无法深入理解词汇在不同上下文中的语义变化、多义性等高级语义特征,缺乏适应文本语境动态调整的能力,在理解文本深层含义时存在缺陷,例如对“transformer”进行编码时,静态向量不能明确区分含义是变形金刚还是变压器。PLM 则从根本上解决了这些问题,为电力领域内的 NLP 应用带来了质的提升。实际上,大部分的 PLM 是基于 Transformer^[38]结构的,自从 Transformer 提出后,就有许多研究人员对 Transformer 进行拆分或者改造,应用到各种任务^[39]中提升性能。

目前基于 Transformer 架构的命名实体识别模型应用在电力领域较多的有 BERT (bidirectional encoder representation from transformer) 模型^[40]、RoBERTa (a robustly optimized bert pretraining approach) 模型^[41]、ERNIE (enhanced representation through knowledge integration) 模型^[42]。

通常电力文本数据相对较复杂,相同实体在不同的位置有不同的含义,例如电力事故案例调查报告中,“停电”这一词汇表示电路状态,也表示检修人员进行的操作,而 BERT 能够缓解电力领域实体识别中一词多义问题。引入掩码语言建模 (masked language modeling, MLM) 技术,随机遮蔽输入文本中 15% 的词并要求模型预测被遮蔽的内容,实现了动态生成字词层面具有深度理解能力的语义向量。此外,BERT 模型还引入 NSP (next sentence prediction),共同优化一个复合损失函数,捕捉句子层面的语义关系。相较于传统语言模式受限于单一方向或浅层拼接的双流向处理存在难以充分捕捉词汇间深层次的语义关联的问题,BERT 模型的性能具有很大的优势。俞阳等^[43]将 BERT 作为嵌入模型,结合神经张量构建关系网络对实体识别,实验结果明显优于其他基线模型。如今科研主流趋势已经广泛采纳预训练语言模型作为核心组件被广泛采纳,将其嵌入到神经网络中的架构中,形成“预训练语言模型驱动”的组合模型范式。应用于电力领域最广泛的模型范式是 BERT-BiLSTM-CRF。文献[40-48]均采用预训练方法构建 BERT-BiLSTM-CRF 实体识别模型,通过 BERT 将字符转化为嵌入向量,利用 BiLSTM 对嵌入向量双向特征提取,最后通过 CRF 预测。实验结果表明,该方法能够有效提升实体识别的准确率。然而,随着技术不断地更新迭代,学者们不断地对主流模型改进,获取更高的模型性能。孙宏云等^[49]认为传统的 BERT-BiLSTM-CRF 模型对 GPU (graphics processing unit) 的利用率不高,提出 BERT-IDCNN-BiLSTM-CRF (bert-iterative dilated convolutional neural network-bidirectional long short-term memory-conditional random field) 方法。其中 IDCNN 模型比 BiLSTM 模型提取更多的配电网数据的局部特征,同时加快 GPU 的并行速度。张智源等^[50]通过构建 BiLSTM-多头注意力机制和多窗口门控 CNN 双通道的特征提取层,捕获句子的全局时序特征和提取电机文本的多粒度局部特征,提取动态字向量中的语义特性。同时经过验证,与传统的 BERT-BiLSTM-CRF 模型相比,该模型具备更强的领域迁移能力。实际上,通过已有研究表明,直接使用通用的 BERT 得到词向量,在处理特定领域的下游任务时会导致不理想的结果。蒋晨等^[51]采用预训练的方式对通用 BERT 调整参数得到 PowerBERT,提高电力检修实体识别的准确率。徐翀等^[52]采用“预训练-精调”的方式得到电力科技领域 PLM。通过大量实验验证,该电力领域 PLM 能够实现电力科技术语实体的识别准确率,提升电力领域知识获取的表现。Zheng 等^[53]采用集成学习的方法,提出 AttCNN-BiGRU-CRF (attention-based convolutional neural network-bidirectional gated recurrent unit-conditional random field) 模型,该模型在分布式嵌入层基于 BERT 结合了字符嵌入和词嵌入,获得更多的语义信息,文本编码层采用结合局部注意机制和 CNN 的方法捕捉局部上下文之间的关系,采用 BiGRU 模型提取电力计量文本的高级特征,同时增加全局多头注意力优化句子级信息的处理,当面对电力文本示例如“用电负荷过大会导致电力计量装置烧坏”,全局多头注意力将会更关注“导致”左右分量,并给予它们更高的权重。

BERT 模型在预训练阶段采用 MLM 技术训练,随机选择一部分单字掩蔽,然而,这种方法并未充分考虑词汇之间的内在关系。例如,在电力领域的电气设备中的“真空断路器”在预训练阶段被拆分为“真”、“空”、“断”、“路”、“器”5 个单字,并且其中的一个或几个字可能会被 [MASK] 符号遮罩。这种方法只能学习到部分字与字之间的关系,而无法准确理解与表示“真空断路器”这个词的整体语义。在电力领域中要深层次语义理解,因此 BERT 模型的局限性导致实体识别性能降低。

RoBERTa 是 BERT 的优化模型,摒弃 BERT 模型中固定的、静态的掩码策略,采用动态掩码策略,在训练过程中,被掩码的 tokens 不断变化,这种设计增加模型对不同遮蔽方式的适应性,提高模型性能。特别是在电力领域,许多术语具有相似的结构,要求模型能够准确地识别实体边界。例如,针对“主变压器”的简称“主变”,RoBERTa 模型能够以字为粒度进行切分,表示为“/主/变/”,或者以词为粒度进行切分,表示为“/主变/”。第 2 种切分形式能够捕获更多词边界信息,对于正确理解和处理电力领域的专业术语至关重要。由于 BERT 模型中的 NSP 任务对某些下游任务的性能提升有限,甚至可能带来负面影响,因此, RoBERTa 移除 NSP 任务,专注于 MLM 的改进,使得模型训练更加聚焦于单个文本片段内的语义理解和上下文建模。扩大训练规模,例如增大 Batch Size,增加训练步数等,促使在各种下游任务的性能持续提升。采用 BPE (byte-pair encoding),从 Byte 的层面分解词,处理未登录词 (out of vocabulary, OOV)。针对电力文

本数据集稀缺、实体类型多样且专业性强的特点,田雪涵等^[17]提出基于增强优化预训练语言模型的电力数据实体识别方法。该方法利用实体词袋替换的数据扩大原始数据集,并采用 RoBERTa 进行动态语义编码,捕捉词汇在不同上下文中的丰富语义信息。此外,还运用 BiLSTM 提取特征,有效整合上下文依赖关系。针对电力领域科技项目申报评审工作中存在的项目与专家精准匹配难的问题,杨政等^[54]采用 RoBERTa 词嵌入,在训练时扩大了训练样本,有效提升电力文本与电力专家匹配任务准确率。为了缓解电力领域经常出现的 OOV 问题,黄源航等^[55]采用 RoBERTa-BiLSTM-CRF 模型,使用电力文本语料对 RoBERTa 预训练,由于预训练阶段采用的语料库与后续电力领域词汇挖掘任务的领域相同,因此能够高效学习该领域的特定属性,较 BERT-BiLSTM-CRF 模型取得更好的效果。RoBERTa 模型的结构如图 2 所示,其中 E_1, E_2, \dots, E_n 为嵌入向量; T_{rm} 为 Transformer 编码器; T_1, T_2, \dots, T_n 为 RoBERTa 输出的动态词向量。

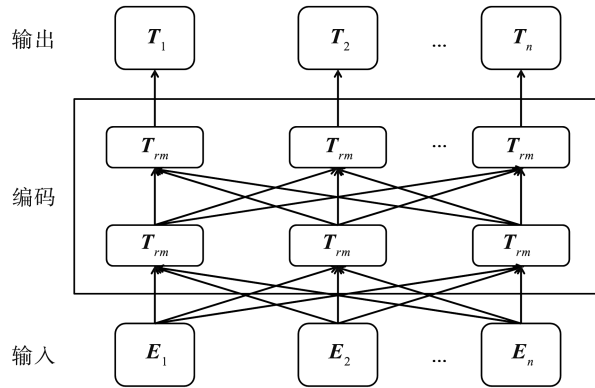


图2 RoBERTa 模型结构

Fig.2 Model of RoBERTa structure

ERNIE 模型是知识增强的语义表示模型,它一种在 BERT 模型基础上改进的预训练语言模型。尽管两者在结构上一致,但 ERNIE 模型对预训练任务设计与随机掩码策略改进,引入基于实体和短语的掩码策略,提升模型对词向量的句法、语法理解能力以及对多维度语义信息的整合能力。张锐等^[56]对比添加 ERNIE 模型与未添加 ERNIE 模型, $F1$ 从 57.14% 提高到 77.26%。王佳琪等^[57]考虑 ERNIE 相较于 BERT 具有更强大的语义表达能力,结合 IDCNN 模型的并行计算能力提出基于 ERNIE-IDCNN-CRF 模型,与能够识别电网调度领域较长的中文实体。Ji 等^[58]提出 ERNIE-BiLSTM-CRF 模型,通过实验验证,与传统的 BERT-BiLSTM-CRF 方法相比,该模型在命名实体识别任务和关系抽取任务具有更大优势。皮俊波等^[59]对 UIE (universal information extraction) 框架改进,接入 ERNIE3.0 模型,在训练时微调参数,有效解决电网故障处置预案逻辑复杂和实体嵌套问题。

对比分析基于 Transformer 方法的预训练模型性能,发现 BERT 模型及其变种 RoBERTa 模型和 ERNIE 模型在电力领域的 NER 任务中表现出了优异的效果。BERT-WWM 模型和 RoBERTa-wwm-large 模型在不同的数据集上的 $F1$ 较大,如李嘉皓等^[14]采用的 BERT-WWM-BiLSTM-Att-CRF 模型在巡检报告数据集上 $F1$ 为 96.81%,而纪鑫等^[16]的 RoBERTa-BiLSTM-CRF 模型在设备故障事件记录文本的 $F1$ 为 90.69%。与此相比, RoBERTa 模型和 ERNIE 模型在电力调度和设备运维的应用场景中,尤其在较为复杂的任务中,仍然表现出较强的适应能力,例如皮俊波等^[59]的 ERNIE3.0 模型通过 UIE 框架和双指针解码在电力调度控制差异化预案中 $F1$ 为 97.91%。虽然 BERT 模型在处理电力文本时较为稳定和高效,但基于 Transformer 的其他如 ERNIE、RoBERTa 等模型,在处理特定任务或更具挑战性的任务时,实体识别精确度更高,尤其在自监督训练和微调的组合下,更好地适应电力行业中不同类型的数据集,提高模型的泛化能力和表现力。

电力领域命名实体识别研究中,模型的适配性优化对识别效果起关键作用,不同场景下的文本特性往往要结合专业知识调整。同时,轻量化模型的研发也是该领域的重要方向,这对于提升实体识别在实际电力系统中的应用效率具有重要意义。Liu 等^[60]探索了基于 ELECTRA (efficiently learning an encoder that classifies token replacements accurately) 的模型用于电力安全命名实体识别,为电力安全文本的实体识别提供有效的预训练模型选择方案。

2.3 标签解码层

标签解码层位于 NER 模型的最后一层,该层的作用是通过标签解码器把依赖于上下文的表示作为输

入,并生成与输入序列对应的标注序列。目前,电力领域中常用的标签解码器有条件随机场、指针网络和胶囊网络等。

2.3.1 条件随机场

在解码阶段,条件随机场^[14]相较于仅依据单个词预测标签的模型,考虑了标签间的依赖关系,即在确定序列标注时,不仅考虑每个词对应的分类标签概率,而且考虑前后相邻标签之间的转移概率。因此,CRF模型并非孤立地为序列中的每个词预测标签,而是输出全局最优的标签序列。目前,在电力领域已有很多研究选择CRF作为标签解码器,将其应用于文本编码器中的神经网络模型之后。

2.3.2 指针网络

指针网络(pointer networks)^[61]本质上是标准Seq2Seq模型的变体^[62],它不依赖所有可能输出结果的加权求和生成最终输出,而是直接利用Softmax层的输出结果指向源序列中特定位置的概率,实现动态检索输入序列中的信息。何俊等^[63]提出基于Seq2Seq的半指针半标注的知识抽取模型,对于实体采用指针网络作为标签解码器,对于实体关系则采用多层指针网络进行解码,其中指针的层数为关系种类的数量,解决输电图文本文实体关系抽取中实体边界模糊的问题。

2.3.3 胶囊网络

胶囊网络^[64](capsnet)凭借向量神经元概念和动态路由算法,将卷积产生的特征图转化为胶囊形式预测分类。当胶囊网络作为标签解码器时,输出胶囊向量,向量模长表示识别标签的概率,方向表示实体的内部信息,如位置属性等,当实体发生变化时,向量的模长不变,而方向发生改变,所以胶囊网络能有效地表征实体的信息。实际上,胶囊网络在电力领域应用极少。为解决电力供应链风险识别准确率低、泛化能力不高且鲁棒性差的问题,冯曙明等^[65]构建供应链胶囊、采购胶囊、生产胶囊、配送与送电胶囊,同时随机生成不确定因素胶囊,构建的CPE模型优于传统机器学习和深度学习模型。

3 基于大语言模型命名实体识别方法

本章聚焦大语言模型(large language models, LLM)在电力领域命名实体识别中的应用。LLM通常指参数量巨大(通常在数十亿至万亿级别)^[66-67]、展现出强大的涌现能力(如上下文学习、指令遵循、复杂推理)的通用语言模型。区别于第2.2.2节讨论的基于Transformer架构的预训练语言模型及其传统的微调应用范式,LLM为NER任务带来了新的可能性,主要体现在利用提示工程、参数高效微调以及内在的少样本/零样本学习能力等方面。

随着模型参数量的不断扩展以及算力资源的持续提升,传统深度学习方法逐渐演进为更加依赖数据驱动的深度学习方式。大型语言模型的出现,使参数规模达到了百亿甚至千亿级别,标志着人工智能进入了大模型时代。伴随多种高性能大模型产品的相继问世,LLM的研究与应用不断迈向新的高度。

早期的大语言模型具有少样本学习能力,这类模型核心训练机制是预测下一个词优化模型,并未充分适应遵循人类指令,有时会产生无效甚至是负面的信息。针对这些挑战,研究者们提出了2种关键的大规模语言模型优化技术路线:指令微调(instruction tuning)^[68]和基于人类反馈的强化学习(reinforcement learning from human feedback, RLHF)^[69]。指令微调利用结构化指令及对应正确答案的训练数据增强模型处理多种通用任务的能力,而RLHF则直观地将人工评估引入大模型的训练循环中,构建反映人类偏好的奖励模型,以此作为指导信号调整并优化大语言模型的行为表现。整个过程从用户通过接口输入信息开始,系统接收到这些输入后,首先进行数据的分析与预处理。数据会经过清洗、筛选和转换,确保其符合后续处理的要求。接着,系统利用特定的算法进行进一步分析和建模,且涉及对数据分类、聚类或其他方式的深度处理。在分析过程中,系统根据不同的决策点做出相应的调整 and 选择,以优化最终结果。最后,系统将处理后的结果反馈给用户,完成整个信息处理的闭环。此外,在大模型的实际运用中,还发展了一系列提示工程技术,例如思维链(chain-of-thoughts, CoT)^[70]和思维树(tree-of-thoughts, ToT)^[71]等方法,这些技术旨在引导大模型揭示其深层次的认知逻辑和推理过程,显著提高解决实际问题时的效能与准确性。

由于LLM具备深入的理解能力和强大的上下文推理能力,它在处理复杂的文本数据时具备更大的优势,为命名实体识别带来革命性的方法。Wei等^[70]将NER任务转换为2阶段的多回合问答问题,第1阶段

目标是将问题和句子结合作为 ChatGPT 的输入找出句子中的实体、关系或事件的现有类型,第2阶段对第1阶段提取的元素类型命名实体识别、关系抽取和事件抽取,进一步提取相关信息,在6个广泛使用的信息抽取数据集上,所提方法平均提高18.98%的性能^[70];受 CoT 的启发,将 NER 任务分解为2个步骤:第1步是输入 Prompt 到 GPT-3.5 确定实体跨度,定位文本实体,提供清晰的边界,第2步是确定实体类型,为每个实体分配准确的类别标签,通过融合实体知识,解决 LLM 缺乏领域知识的问题;Polak 等^[72]提出 ChatExtract,由一组预设的提示构成,分别结合 ChatGPT-3.5、LLaMA2 (large language model meta ai 2)-Chat 和 ChatGPT-4 大语言模型在数据抽取任务中实现高准确率和召回率,采用 ChatGPT-4 在测试集,准确率为90.8%和召回率为87.7%。

在电力领域,Liu 等^[73]微调中文 LLMA (large language model meta ai)-7b 从电力相关科研成果中提取研究问题和技术术语。首先通过规则模板与 GPT-3.5 数据增强相结合的方式构建微调数据集,有效避免手动构建数据集的高昂成本和耗时的问题。其次,采用 LoRA (low-rank adaptation) 算法微调,该算法通过低秩矩阵分解降秩,减少训练参数的数量和 GPU 内存,节省计算资源。最后比较模型 $F1$ 在不同微调数据集提取效果,结果显示,在由规则模板和 GPT-3.5 增强组合构建的数据集上训练的模型对于研究问题和技术术语两类实体有着良好的识别效果, $F1$ 分别为86.75%和86.42%。

实际上,LLM 通常应用于电力知识图谱中,电力知识图谱为 LLM 提供结构化的知识,增强 LLM 在细致理解、情境化响应和特定领域推理方面的能力。Hu 等^[74]提出多模态知识图与 LLM 结合的风电机组装配过程问答系统,采用 Vicuna-7b 模型在各个领域的数据集上预训练,在输入问题的开头添加字符串前缀进行粗粒度解析,LLM 从风电机组装配过程的问题中提取实体以及问题中实体与实体之间的关系,最终输出问题的粗粒度分析结果。由于电力能源知识图谱包括能源知识、电力理论知识、电力设备知识等结构化数据,而结构化数据性质不适合直接用于 LLM 的增量预训练,因此 Yin 等^[75]提出将构建的知识图谱中的实体关系作为关键字,利用 ChatGPT API (application programming interface) 生成大量针对电力能源行业的高质量文本数据。

LLM 的“幻觉”是指模型的输出生成内容缺乏逻辑一致性或事实准确性,并且无法通过输入信息证实。例如,当训练数据集中的输入输出对之间出现差异时,模型在训练过程中容易产生幻觉;当生成的文本超出输入数据的范围时,也可能导致幻觉。为了缓解 LLM 在电力领域出现幻觉及数据泄露问题,Zhao 等^[76]提出自洽性 (self-consistency, SC)、抽取与修正框架,该框架改进了电力运维 (O&M) 领域。其中,将 SC 框架从通用领域转移到了电力 O&M 领域,更适合中文环境。此外,还设计了抽取机制,在电力运维知识图谱上多次随机游走生成证据链,这些证据链有助于生成可靠和准确的回答,为了进一步提升回答的质量,构建的 SCER 框架引入修正机制。该框架整体分为预训练模型构建、下游任务适配以及实体识别3个主要阶段。首先,大模型通过海量语料预训练,获得强大的语义表示能力,为后续任务提供通用特征支持。在此基础上,针对电力领域的需求,进一步构建 NER 模型,即以大模型为基础,结合电力专业语料在实体识别模型训练,从而具备识别电力设备、运维要素等关键实体的能力。同时,框架中还包含句子表示模型的训练环节,利用句向量表征实现电力文本的语义理解与分类任务。随后,这些预训练好的模型参数或语义嵌入将应用到自然语言推理模型中,借助推理机制完成更深层次的语义关系建模,支持复杂文本的分析与理解。最终,该框架能够在电力文本中识别出变压器、绝缘油、变电站、发电站、电压等关键实体,并结合上下文关系准确提取与分类,为电力领域的知识图谱构建、运维管理、调度优化等应用提供智能支撑,实现从通用大模型到电力专业 NER 任务的逐层迁移和优化,体现了预训练—微调—应用的完整链路。

4 电力领域命名实体识别的应用

命名实体识别作为电力文本智能处理的基础环节,提取的关键实体信息是支撑上层应用的核心要素。PDNER 技术在电力领域的多个关键场景中发挥着不可或缺的作用,显著提升了业务处理的智能化水平和效率。

4.1 智能电网运维与管理

PDNER 能够从海量的调度日志、运行报告、检修记录等非结构化文本中自动识别设备名称、设备型号、

缺陷类型、操作指令、状态信息等关键实体。这些信息是构建设备知识图谱、实现设备状态实时监测、故障诊断与预测性维护的核心输入。例如,文献[34]识别电网调度文本中的设备与操作实体,为调度决策提供支持;文献[47]利用识别的输变电设备实体构建运维知识图谱。

4.2 故障诊断与事故分析

在故障报告、事故简报、保护动作信息等文本中,PDNER 精准提取故障设备、故障位置、故障类型、保护装置动作信息、电压电流异常值等实体。结合关系抽取等技术,自动化梳理故障因果链,辅助快速定位故障根源,提升事故分析效率。

4.3 电力安全与合规监管

安规文档、两票、安全检查记录等文本包含大量的安全相关实体,如作业类型、风险点、安全措施、人员角色、作业地点等。PDNER 可自动化抽取这些信息,用于风险自动识别、安全措施完备性检查、违章行为分析等,保障作业安全与合规性。文献[6,32]的工作涉及安全相关文本的实体识别。

4.4 客户服务与用电管理

在客服工单、用户咨询、用电检查报告等文本中,PDNER 可识别用户信息、地址、户号、用电问题类型(如“欠费”、“表计故障”)、诉求内容、设备信息等实体,有助于实现用户问题的快速分类、精准匹配和高效处理,提升客户服务质量。文献[12]探索了客户需求和文本中的实体识别。

4.5 电力知识图谱构建与信息检索

PDNER 是构建高质量电力领域知识图谱的核心前置任务。通过从技术标准、科研文献、设备手册、政策法规等海量文本中抽取电力设备、技术术语、参数、标准规范、机构等实体及其属性,为构建覆盖发电、输电、变电、配电、用电全环节的结构化知识库奠定基础。构建的知识图谱支撑了更精准、更智能的信息检索、问答系统和决策支持。文献[3,16,23,46,56,58]均将 NER 作为构建电力知识图谱的关键步骤。

4.6 调度指令理解与辅助决策

调度指令文本包含复杂的设备操作序列和状态要求。PDNER 能识别指令中的操作对象、操作类型、状态要求、保护投退等实体,辅助调度员理解指令意图,与知识图谱结合进行安全校核和辅助决策生成。实际上,PDNER 技术的应用已渗透到电力系统生产、运维、管理、服务等多个核心环节,是实现电力系统数字化转型和智能化升级的重要技术支撑。Tang 等^[77]利用通用信息抽取框架增强停电调度中的命名实体识别和事件抽取,说明 NER 在电网调度优化中的核心作用。

5 国内主要研究机构技术进展与特点比较

5.1 电力科研院所

这类单位主要服务于国家电网、南方电网等大型央企的核心业务需求,聚焦于行业标准制定、关键技术攻关和大型系统平台,比如智能调度系统、设备状态评估系统的研发。NER 作为知识抽取的基础环节,是构建电力知识图谱、支撑智能应用的核心技术之一。在技术路线上,电力科研院所特别注重技术的实用性、可靠性和与业务系统的集成性,广泛应用规则和词典方法,研究基于深度学习的 NER,特别是预训练语言模型的领域自适应和微调,积极参与国产大模型在电力领域的探索。电力科研院所的优势在于拥有最丰富、最核心的电力生产运行数据,包括调度日志、设备台账、缺陷记录、故障报告、两票等,数据规模庞大且专业性强,但同时也面临数据敏感性高、开放度有限的问题,通常具备强大的专业标注团队或与高校合作标注数据。成果主要表现为大型业务系统/平台的核心模块、行业/企业标准规范、高水平研究报告、专利以及部分开源工具/模型(相对较少),影响力主要体现在行业内部和大型工程项目中。当前面临的主要挑战包括平衡数据安全与开放共享、老旧系统数据格式兼容以及大模型落地成本与效益评估。

5.2 电网/发电企业及其下属科技公司

这类单位直接面向具体业务场景的落地应用和效率提升,需求驱动性强,核心目标聚焦于解决运维、检修、营销、客服等一线业务中的实际问题,例如工单信息自动提取、故障报告分析^[21,51]、调度指令理解^[32,37,57]、客户需求识别^[43,54]等,NER 被视为提升业务流程自动化、智能化水平的关键工具。技术路线追求快速见效、易于部署和维护,广泛采用成熟的深度学习方法,积极探索结合业务规则的后处理或混合模

型^[6,31]或大语言模型,关注提示工程和参数高效微调。在数据资源方面,企业掌握特定业务场景的海量一手数据,这些数据与具体业务流紧密结合、价值密度高,但也存在分散、格式不统一、质量参差不齐的特点,标注工作常依赖业务人员或外包。成果主要是具体的业务应用系统或功能模块、内部技术报告/解决方案以及实用化工具。影响力体现在提升业务指标。面临的主要挑战包括业务需求快速变化、数据孤岛问题、标注成本高且质量保障难、IT 基础设施对模型轻量化的要求以及缺乏核心算法研发深度。

5.3 高等院校

高等院校在电力领域 NER 研究中的核心目标是前沿技术探索、理论方法创新和人才培养,其研究更具前瞻性和普适性,关注点涵盖新模型架构、新学习范式、应用新场景以及大语言模型的基础理论与应用创新^[70-75]。在技术路线上,高等院校处于探索的最前沿,跟进统计机器学习方法,广泛尝试各类网络结构及其组合、注意力机制创新,高等院校是国内电力 NER 研究中预训练语言模型微调、领域自适应、大语言模型提示工程/微调的主力军,研究更具多样性和创新性。虽然数据来源相对多元化,但规模通常小于企业,包括公开的电力文献^[28,44,50]、合作获取的脱敏数据^[17,23-24,33,36-37,49,55]、自行构建的小规模标注数据集^[31,34]以及利用网络爬虫获取的数据^[45],并在数据增强技术^[17]方面有较多探索。成果以高水平学术论文、开源代码/模型、专利和研究生学位论文为主,影响力主要体现在学术界和技术社区。面临的核心挑战是获取高质量、大规模的工业级真实数据困难、研究成果向实际业务场景转化的壁垒以及计算资源限制。

国内电力领域 NER 的研究与应用呈现产学研用紧密结合的特点,不同性质的单位在目标导向、技术路线、数据资源和成果形态上存在显著差异,共同推动着技术的发展与落地。以下对主要类型单位的特点进行比较分析,如表 2 所示。

表 2 国内单位在电力领域 NER 研究与应用特点比较

Table 2 Comparison of research and application characteristics of NER in the power field among domestic organizations

特征维度	电力科研院所	电网/发电企业及科技公司	高等院校
核心目标	行业标准、核心技术、大型系统平台	具体业务场景落地、效率提升、自动化	前沿技术探索、理论方法创新、人才培养
技术路线侧重	可靠性、集成性、领域自适应预训练/微调、大模型落地	深度学习方法、业务规则结合、快速见效、大模型应用探索	模型架构创新、新学习范式、预训练/大模型前沿应用
主要数据来源	核心生产运行数据	特定业务场景数据	公开文献、合作脱敏数据、自建数据集、网络数据(相对分散)
关键技术成果	行业标准、核心系统模块、专利、研究报告	业务应用系统/模块、内部解决方案、实用工具	学术论文、开源模型/代码、专利、学位论文
核心优势	数据深度、行业影响力、系统工程能力	业务理解深度、应用落地能力、一线需求驱动	学术创新力、技术前瞻性、人才储备
主要挑战	数据安全与开放、老旧系统兼容、大模型成本	需求变化快、数据孤岛、标注成本/质量、轻量化部署、算法深度	获取工业级真实数据、成果转化壁垒、计算资源限制

国内电力 NER 生态呈现出清晰的协作链条:高校前沿探索和方法创新,产出理论成果和开源工具,科研院所聚焦核心技术攻关、标准制定和大型平台研发,起桥梁和引领作用,企业则立足具体业务需求,推动技术落地应用和迭代优化。三者相互促进,共同推动中国电力领域 NER 技术的快速发展和应用深化。未来,随着大语言模型等新技术的兴起,产学研用的协同将更加紧密,特别是在数据安全合规共享、大模型行业落地、复合人才培养等方面需要持续加强合作。

6 面临的挑战及展望

6.1 电力领域命名实体识别面临的挑战

电力领域命名实体识别在技术落地过程中面临多重瓶颈,这些挑战既源于电力文本的专业特性,也反映了当前模型在方法论和工程实践中的局限性。

标准化评估体系缺失与数据生态缺陷。当前研究缺乏统一的评估基准,导致模型性能可比性严重不足。

一方面,实体类别定义尚未标准化,部分研究将“电压等级”视为独立实体,而另一些研究将其归为设备属性;另一方面,数据集规模与分布差异显著,公开研究多基于小规模单一场景数据,而工业场景中多模态数据占比超过 50%,且包含超过 20%的噪声信息,这种非标准化导致跨模型对比失真。此外,新兴技术领域(如虚拟电厂、新型储能)的标注样本普遍不足 1 000 条,限制了模型对长尾实体的覆盖能力。

工业场景泛化能力与实时性瓶颈。电子电力文本的长文本嵌套性与领域动态性特点,实验室环境下的高准确率难以复现于真实电力系统。实际上,超过 40%的运维报告包含多层级实体嵌套,而现有模型对嵌套实体的边界识别准确率较单实体低。同时,电网业务的实时性要求模型具备毫秒级响应能力,但主流 Transformer 模型在边缘设备的推理延迟超过 200 ms,无法满足故障预警等实时场景需求。

大语言模型的领域适配性困境。LLM 在电力领域的应用面临“幻觉”与“知识断层”双重挑战。LLM 对电力专业逻辑的理解存在偏差,例如 ChatGPT-3.5 曾将“电抗器匝间短路”错误关联至“输电线路故障”类别,幻觉率较高。现有解决方案多依赖离散提示词,缺乏结构化知识约束,而电力领域特有的设备层级关系、故障因果链尚未有效融入模型训练。

低资源场景与跨领域迁移壁垒。电力系统的垂直细分场景(如输电、变电、配电)存在显著的术语鸿沟,例如输电领域的杆塔型号、线路走廊与配电领域的台区编号、用户负荷几乎无重叠词汇,导致模型跨子领域迁移时 $F1$ 下降。在低资源场景下,传统监督学习模型性能大幅衰减,而现有数据增强方法(如同义词替换)难以捕捉电力术语的构词规律。此外,不同业务系统的数据格式差异进一步加剧模型的跨系统适配难度。Luo 等^[78]提出具有显式关系的联邦 NER 模型用于电网,尝试在保护隐私的同时解决数据孤岛问题并提升模型在分布式环境下的性能,为应对低资源迁移和隐私挑战提供了新的思路。

数据隐私与模型可解释性挑战。电力数据涉及敏感信息,包括用户用电行为、电网运行数据和设备状态等,但现有 NER 模型多采用集中式训练,存在数据泄露风险。联邦学习等隐私保护技术在电力领域的应用尚处探索阶段,并可能因通信开销导致训练效率下降。同时,深度学习模型的黑盒特性阻碍了专业人员对决策逻辑的理解。

综上所述,当前电力领域命名实体识别在数据标准化、模型泛化与实时性、大语言模型适配、低资源迁移以及隐私可解释性等方面均存在显著瓶颈。这些挑战严重制约了 PDNER 技术在复杂电力业务场景中的实际应用效果与推广价值。

6.2 电力领域命名实体识别展望

为有效应对关键挑战,未来研究需围绕标准化-领域知识增强-轻量化部署-隐私安全的技术路线,构建更具鲁棒性与实用性的 PDNER 体系。

构建工业级基准生态与标准化评估体系。标准化是推动 PDNER 技术落地的基础,需联合行业力量制定统一实体分类规范与标注标准,并建立覆盖多模态、多场景的工业级基准数据集 PowerNER,集成文本、表格及噪声数据,设计领域适配性、实时性等复合评估指标,为模型性能对比提供公平基准。在此基础上,通过开源协作推动数据共享,加速技术迭代与产业落地。

知识图谱增强的大语言模型可靠性提升。LLM 的领域可靠性提升需融合知识图谱的动态约束与轻量化适配技术。未来可以通过构建电力知识图谱,采用实体链接与逻辑校验机制抑制幻觉现象,并要求模型输出附带知识来源,降低错误关联风险,同时采用 LoRA 等低秩微调技术,在边缘设备上实现高效领域适配,兼顾性能与资源消耗。

低资源学习与跨领域迁移的技术创新。面对低资源场景与跨领域迁移难题,可探索半监督数据增强技术,利用预训练模型生成伪标注数据并结合专家校验扩充数据集,同时设计跨领域特征解耦架构,通过轻量化适配器实现输电、变电、配电等子领域的快速迁移。此外,未来可以通过探索元学习算法,在少样本场景下快速适应新型电力实体。

隐私保护与实时性的平衡架构。在隐私保护与实时性平衡方面,可集成联邦学习与边缘计算架构,分布式节点上协同训练模型以避免数据泄露,同时通过知识蒸馏压缩模型体积并结合现场可编程门阵列(field-programmable gate array, FPGA)硬件加速,将推理延迟降至毫秒级,满足电网实时监控需求。

增强可解释性。为提升模型可信度,开发注意力机制可视化工具展示关键特征关注区域,并嵌入电力行业规则引擎对解码结果逻辑校验,确保实体识别符合专业规范。

7 结语

命名实体识别是数据挖掘、信息检索及自然语言处理的核心研究内容之一。NER 技术经历了从早期依赖规则和词典的方法到传统机器学习方法的转变,虽然这些方法已不再是主流,但为后续研究积累了宝贵的经验和教训。随着深度学习的不断发展,基于大语言模型的方法在 NER 领域展现出广阔的应用前景,受到了国内外学者的广泛关注,显然,大模型的出现为电力领域命名实体识别开辟了一条新的研究路径。

在电力领域,NER 的重要性尤为突出,它在智能电网管理、故障诊断报告分析等方面发挥着至关重要的作用。例如,准确识别文本中的变电站名称、输电线路编号和电力设备名称等实体,对于实现快速响应电网事件、优化资源配置以及加强电力生产安全具有重要意义。鉴于电力行业的专业特性和术语的独特性,未来的 NER 研究在电力领域更加专注于领域知识图谱的构建与融合,助力电力行业高质量与智能化发展。

参考文献:

- [1] 王颖洁,张程焯,白凤波,等. 中文命名实体识别研究综述[J]. 计算机科学与探索,2023,17(2):324-341.
WANG Yinjie, ZHANG Chengye, BAI Fengbo, et al. Review of Chinese named entity recognition research[J]. Journal of Frontiers of Computer Science and Technology, 2023, 17(2):324-341.
- [2] 李猛,李艳玲,林民. 命名实体识别的迁移学习研究综述[J]. 计算机科学与探索,2021,15(2):206-218.
LI Meng, LI Yanlin, LIN Min. Review of transfer learning for named entity recognition[J]. Journal of Frontiers of Computer Science and Technology, 2021, 15(2):206-218.
- [3] 李新鹏,徐建航,郭子明,等. 调度自动化系统知识图谱的构建与应用[J]. 中国电力,2019,52(2):70-77.
LI Xinpeng, XU Jianhang, GUO Ziming, et al. Construction and application of knowledge graph for dispatching automation system[J]. Electric Power, 2019, 52(2):70-77.
- [4] 王慧芳,曹靖,罗麟. 电力文本数据挖掘现状及挑战[J]. 浙江电力,2019,38(3):1-7.
WANG Huifang, CAO Jing, LUO Lin. Current status and challenges of power text data mining[J]. Zhejiang Electric Power, 2019, 38(3):1-7.
- [5] 冀振燕,孔德焱,刘伟,等. 基于深度学习的命名实体识别研究[J]. 计算机集成制造系统,2022(6):1603-1615.
JI Zhenyan, KONG Deyan, LIU Wei, et al. Research on named entity recognition based on deep learning[J]. Computer Integrated Manufacturing Systems, 2022(6):1603-1615.
- [6] 袁金斗,潘明明,张腾,等. 基于规则和词典的用电安全领域命名实体识别[J]. 电子技术应用,2022,48(12):22-27.
YUAN Jindou, PAN Mingming, ZHANG Teng, et al. Named entity recognition in power safety domain based on rules and dictionaries[J]. Application of Electronic Technique, 2022, 48(12):22-27.
- [7] 徐鹏,龚伟,宋俊典. 基于MRC的设备故障命名实体识别方法[J]. 计算机应用与软件,2024,41(5):171-176.
XU Peng, GONG Wei, SONG Jundian. Named entity identification method of equipment fault based on MRC[J]. Computer Applications and Software, 2024, 41(5):171-176.
- [8] RABINER L R. A tutorial on hidden Markov models and selected applications in speech recognition[J]. Proceedings of the IEEE, 1989, 77(2):257-286.
- [9] HUNT E B, MARIN J, STONE P J. Experiments in induction[M]. New York: Academic Press, 1966:127-156.
- [10] 刘梓权,王慧芳. 基于知识图谱技术的电力设备缺陷记录检索方法[J]. 电力系统自动化,2018,42(14):158-164.
LIU Ziquan, WANG Huifang. Retrieval method for power equipment defect records based on knowledge graph technology[J]. Automation of Electric Power Systems, 2018, 42(14):158-164.
- [11] 孙玉芹,肖静婷,王海超. 基于多模型融合的电力运检命名实体识别[J]. 科学技术与工程,2023,23(36):15545-15552.
SUN Yuqin, XIAO Jingting, WANG Haichao. Power operation and maintenance named entity recognition based on multi-model fusion[J]. Science Technology and Engineering, 2023, 23(36):15545-15552.
- [12] 潘暉,赵岩,李麟,等. 电力客户需求高适配性关联抽取算法[J]. 太赫兹科学与电子信息学报,2023,21(10):1257-1262.
PAN Hui, ZHAO Yan, LI Lin, et al. Highly adaptive association extraction algorithm for power customer needs[J]. Journal of Terahertz Science and Electronic Information Technology, 2023, 21(10):1257-1262.
- [13] 孔静静,于琦,李敬华,等. 实体抽取综述及其在中医药领域的应用[J]. 世界科学技术-中医药现代化,2022,24(8):2957-2963.

- KONG Jingjing, YU Qi, LI Jinghua, et al. Review of entity extraction and its application in traditional Chinese medicine field[J]. World Science and Technology-Modernization of Traditional Chinese Medicine, 2022, 24(8):2957-2963.
- [14] LAFFERTY J, MCCALLUM A, PEREIRA F C N. Conditional random fields: Probabilistic models for segmenting and labeling sequence data[C]//Proceedings of the 18th International Conference on Machine Learning (ICML 2001). Massachusetts: Morgan Kaufmann Publishers, 2001:282-289.
- [15] 李嘉皓,熊威,龚康,等. 融合 BERT-WWM 与注意力机制的电力设备缺陷实体识别研究[J]. 电力学报,2024,39(2):126-135.
- LI Jiahao, XIONG Wei, GONG Kang, et al. Research on entity recognition of power equipment defects integrating BERT-WWM and attention mechanism[J]. Journal of Electric Power, 2024, 39(2):126-135.
- [16] 纪鑫,武同心,余婷,等. 基于多任务学习的电力文本信息抽取[J]. 北京航空航天大学学报,2024,50(8):2461-2469.
- JI Xin, WU Tongxin, YU Ting, et al. Power text information extraction based on multi-task learning[J]. Journal of Beihang University, 2024, 50(8):2461-2469.
- [17] 田雪涵,董坤,赵剑锋,等. 基于增强优化预训练语言模型的电力数据实体识别方法[J]. 智慧电力,2024,52(6):100-107.
- TIAN Xuehan, DONG Kun, ZHAO Jianfeng, et al. Power data entity recognition method based on enhanced optimized pre-trained language model[J]. Smart Power, 2024, 52(6):100-107.
- [18] 吴智妍,金卫,岳路,等. 电子病历命名实体识别技术研究综述[J]. 计算机工程与应用,2022,58(21):13-29.
- WU Zhiyan, JIN Wei, YUE Lu, et al. Review of named entity recognition technology for electronic medical records[J]. Computer Engineering and Applications, 2022, 58(21):13-29.
- [19] CHURCH K W. Word2Vec[J]. Natural Language Engineering, 2017, 23(1):155-162.
- [20] PENNINGTON J, SOCHER R, MANNING C D. Glove: global vectors for word representation[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP 2014). Qatar: Association for Computational Linguistics, 2014:1532-1543.
- [21] 杜修明,秦佳峰,郭诗瑶,等. 电力设备典型故障案例的文本挖掘[J]. 高电压技术,2018,44(4):1078-1084.
- DU Xiuming, QIN Jiafeng, GUO Shiyao, et al. Text mining of typical fault cases of power equipment[J]. High Voltage Engineering, 2018, 44(4):1078-1084.
- [22] 蒋逸雯,李黎,李智威,等. 基于深度语义学习的电力变压器运维文本信息挖掘方法[J]. 中国电机工程学报,2019,39(14):4162-4171.
- JIANG Yiwen, LI Li, LI Zhiwei, et al. Semantic learning-based text mining method for power transformer operation and maintenance[J]. Proceedings of the CSEE, 2019, 39(14):4162-4171.
- [23] 李强,庄莉,赵峰,等. 基于知识增强的配电网运行信息关系抽取方法[J]. 现代电子技术,2024,47(5):171-175.
- LI Qiang, ZHUANG Li, ZHAO Feng, et al. Knowledge-enhanced relationship extraction method for distribution network operation information[J]. Modern Electronics Technique, 2024, 47(5):171-175.
- [24] 张宇波,王有元,梁玄鸿,等. 电力设备缺陷文本的双通道语义增强网络挖掘方法[J]. 高电压技术,2024,50(5):1923-1932.
- ZHANG Yubo, WANG Youyuan, LIANG Xuanhong, et al. Dual-channel semantic enhancement network mining method for defect texts of power equipment[J]. High Voltage Engineering, 2024, 50(5):1923-1932.
- [25] WEI Ziming, QU Shaoheng, ZHAO Li, et al. A position-and similarity-aware named entity recognition model for power equipment maintenance work orders[J]. Sensors, 2025, 25(7):2062.
- [26] CONG Li, CUI Ran, DOU Zeng, et al. Named entity recognition for power data based on lexical enhancement and global pointer[C]//Proceedings of 3rd International Conference on Electronic Information Engineering, Big Data, and Computer Technology (EIBDCT 2024). California: The International Society for Optics and Photonics, 2024, v13181.
- [27] ZHAO Zhenqiang, CHEN Zhenyu, LIU Jinbo, et al. Chinese named entity recognition in power domain based on Bi-LSTM-CRF[C]//Proceedings of the 2nd International Conference on Artificial Intelligence and Pattern Recognition. Beijing: Association for Computing Machinery, 2019:176-180.
- [28] JI Zhixiang, WANG Xiaohui, CAI Changyu, et al. Research on power entity recognition technology based on bidirectional long short-term memory network and conditional random field[J]. Global Energy Interconnection, 2020, 3(2):186-192.
- [29] 杨秋勇,彭泽武,苏华权,等. 基于 Bi-LSTM-CRF 的中文电力实体识别[J]. 信息技术,2021(9):45-50.
- YANG Qiuyong, PENG Zewu, SU Huaquan, et al. Chinese power entity recognition based on Bi-LSTM-CRF[J]. Information Technology, 2021(9):45-50.
- [30] 张大波,郭怀新,储著伟,等. 基于多分类 BiLSTM-CRF 的电网启动方案结构化数据转换模型研究[J]. 电力信息与通信

技术,2023,21(1):54-61.

ZHANG Dabo, GUO Huaixin, CHU Zhuwei, et al. Research on structured data conversion model of power grid startup scheme based on multi-class BiLSTM-CRF[J]. Electric Power Information and Communication Technology, 2023, 21(1): 54-61.

- [31] 肖勇,郑楷洪,王鑫,等. 基于联合神经网络学习的中文电力计量命名实体识别[J]. 浙江大学学报(理学版), 2021, 48(3):321-330.
- XIAO Yong, ZHENG Kaihong, WANG Xin, et al. Chinese power metering named entity recognition based on joint neural network learning[J]. Journal of Zhejiang University (Science Edition), 2021, 48(3):321-330.
- [32] 江叶峰,孙少华,仇晨光,等. 电网故障处置预案文本中的命名实体识别研究[J]. 电力工程技术,2021,40(5):177-183.
- JIANG Yefeng, SUN Shaohua, QIU Chenguang, et al. Research on named entity recognition in power grid fault disposal plans[J]. Power Engineering Technology, 2021, 40(5):177-183.
- [33] 徐会芳,张中浩,谈元鹏,等. 面向电网调度领域的实体识别技术[J]. 电力建设,2021,42(10):71-77.
- XU Huifang, ZHANG Zhonghao, TAN Yuanpeng, et al. Entity recognition technology for power grid dispatching domain[J]. Electric Power Construction, 2021, 42(10):71-77.
- [34] 毛宏亮,艾孜尔古丽,陈德刚. 基于多头注意力的电网调度领域命名实体识别[J]. 计算机技术与发展,2023,33(2):181-186194.
- MAO Hongliang, AIZIERGULI, CHEN Degang. Named entity recognition in power grid dispatching domain based on multi-head attention[J]. Computer Technology and Development, 2023, 33(2):181-186.
- [35] MENG Lingwen, WANG Yulin, BAN Guobang, et al. A multi-source embedding-based named entity recognition model for knowledge graph and its application to on-site operation violations in power grid systems[J]. Electronics 2025, 14(13):2511.
- [36] 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. GRU-based named entity recognition method for power dispatching domain[J]. Computer Systems & Applications, 2020, 29(8):185-191.
- [37] 宋厚岩,王汉军. 基于GRU和PCNN的电力知识抽取[J]. 计算机系统应用,2021,30(9):200-205.
- SONG Houyan, WANG Hanjun. Power knowledge extraction based on GRU and PCNN[J]. Computer Systems & Applications, 2021, 30(9):200-205.
- [38] 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. California: Curran Associates, 2017:5999-6009.
- [39] 张汝佳,代璐,王邦,等. 基于深度学习的中文命名实体识别最新研究进展综述[J]. 中文信息学报,2022,36(6):20-35.
- ZHANG Rujia, DAI Lu, WANG Bang, et al. Review of recent research progress in Chinese named entity recognition based on deep learning[J]. Journal of Chinese Information Processing, 2022, 36(6):20-35.
- [40] DEVLIN J, CHANG M W, LEE K, et al. Bert: pre-training of deep bidirectional transformers for language understanding[C]//Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT 2019), 2019, 1:4171-4186.
- [41] TAN K L, LEE C P, ANBANANTHEN K S M, et al. RoBERTa-LSTM: a hybrid model for sentiment analysis with transformer and recurrent neural network[J]. Ieee Access, 2022, 10:21517-21525.
- [42] SUN Yu, WANG Shuohuan, LI Yukun, et al. ERNIE 2.0: a continual pre-training framework for language Understanding[C]//Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI 2020). New York: Association for the Advancement of Artificial Intelligence, 2020:8968-8975.
- [43] 俞阳,何玮,康雨萌. 一种面向自然语言问题的命名实体识别模型[J]. 电子设计工程,2023,31(14):29-32.
- YU Yang, HE Wei, KANG Yumeng. A named entity recognition model for natural language questions[J]. Electronic Design Engineering, 2023, 31(14):29-32.
- [44] 顾亦然,霍建霖,杨海根,等. 基于BERT的电机领域中文命名实体识别方法[J]. 计算机工程,2021,47(8):78-83.
- GU Yiran, HUO Jianlin, YANG Haigen, et al. BERT-based Chinese named entity recognition method for motor domain[J]. Computer Engineering, 2021, 47(8):78-83.
- [45] 刘斐,文中,吴艺. 基于BERT-BiLSTM-CRF模型的电力行业事故文本智能分析[J]. 中国安全生产科学技术,2023, 19(1):209-215.
- LIU Fei, WEN Zhong, WU Yi. Intelligent analysis of power industry accident texts based on BERT-BiLSTM-CRF model[J]. Journal of Safety Science and Technology, 2023, 19(1):209-215.
- [46] 黄锋,崔志美,黄志都,等. 基于知识图谱的能源互联网输电线路隐患信息检索研究[J]. 电气自动化,2023,45(3):8-10.

- HUANG Feng, CUI Zhimei, HUANG Zhidu, et al. Research on hidden danger information retrieval of energy internet transmission lines based on knowledge graph[J]. *Electrical Automation*, 2023, 45(3):8-10.
- [47] 龚泽威一,肖妮,曹占国,等. 输变电设备运维知识图谱的构建及应用[J]. *电力大数据*, 2023, 26(5):81-89.
GONG Zewei, XIAO Ni, CAO Zhanguo, et al. Construction and application of operation and maintenance knowledge graph for power transmission and transformation equipment[J]. *Power Big Data*, 2023, 26(5):81-89.
- [48] FENG Jun, WANG Hongkai, PENG Liangying, et al. Chinese named entity recognition within the electric power domain[C]// *International Symposium on Emerging Information Security and Applications*. Singapore: Springer Nature Singapore, 2023: 133-146.
- [49] 孙宏云,李喜旺. 面向配电网数据的命名实体识别[J]. *计算机系统应用*, 2023, 32(2):387-393.
SUN Hongyun, LI Xiwan. Named entity recognition for distribution network data[J]. *Computer Systems & Applications*, 2023, 32(2):387-393.
- [50] 张智源,孙水华,徐诗傲,等. 基于BERT和多窗口门控CNN的电机领域命名实体识别[J]. *计算机应用研究*, 2023, 40(1):107-114.
ZHANG Zhiyuan, SUN Shuihua, XU Shiao, et al. BERT and multi-window gated CNN-based named entity recognition for motor domain[J]. *Application Research of Computers*, 2023, 40(1):107-114.
- [51] 蒋晨,王渊,胡俊华,等. 基于深度学习的电力实体信息识别方法[J]. *电网技术*, 2021, 45(6):2141-2149.
JIANG Chen, WANG Yuan, HU Junhua, et al. Deep learning-based power entity information recognition method[J]. *Power System Technology*, 2021, 45(6):2141-2149.
- [52] 徐翀,王其清. 面向知识获取的电力科技领域语言模型研究[J]. *电力信息与通信技术*, 2023, 21(4):31-36.
XU Chong, WANG Qiqing. Research on language model for power science and technology domain oriented to knowledge acquisition[J]. *Electric Power Information and Communication Technology*, 2023, 21(4):31-36.
- [53] ZHENG Kaihong, SUN Lingyun, WANG Xin, et al. Named entity recognition in electric power metering domain based on attention mechanism[J]. *IEEE Access*, 2021, 9:152564-152573.
- [54] 杨政,蔡迪,李慧斌. 基于层次化表示的电力文本命名实体识别和匹配算法[J]. *计算机与现代化*, 2022(5):75-81.
YANG Zheng, CAI Di, LI Huibin. Hierarchical representation-based named entity recognition and matching algorithm for power texts[J]. *Journal of Computers and Modernization*, 2022(5):75-81.
- [55] 黄源航,强梦焯,李涛,等. 基于RoBERTa的电力领域词汇挖掘模型[J]. *电力大数据*, 2022, 25(6):1-8.
HUANG Yuanhang, QIANG Mengye, LI Tao, et al. RoBERTa-based power domain vocabulary mining model[J]. *Power Big Data*, 2022, 25(6):1-8.
- [56] 张锐,刘剑青,张伯远,等. 基于迁移学习的电网故障处置知识图谱构建及实时辅助决策研究[J]. *电力信息与通信技术*, 2022, 20(6):24-34.
ZHANG Rui, LIU Jianqing, ZHANG Boyuan, et al. Construction of power grid fault disposal knowledge graph based on transfer learning and real-time auxiliary decision-making[J]. *Electric Power Information and Communication Technology*, 2022, 20(6):24-34.
- [57] 王佳琪,俞灵,夏文岳,等. 基于ERNIE-IDCNN-CRF模型的电网调度领域命名实体识别方法[J]. *电力信息与通信技术*, 2022, 20(10):1-8.
WANG Jiaqi, YU Ling, XIA Wenyue, et al. ERNIE-IDCNN-CRF model-based named entity recognition method for power grid dispatching domain[J]. *Electric Power Information and Communication Technology*, 2022, 20(10):1-8.
- [58] JI Zhixiang, WANG Xiaohui, ZHANG Jie, et al. Construction and application of power grid dispatching fault handling knowledge graph based on pre-trained model[J]. *Global Energy Interconnection*, 2023, 6(4):493-504.
- [59] 皮俊波,齐世雄,孙文多,等. 基于UIE框架的电网故障处置预案实体和事件识别方法[J]. *中国电力*, 2023, 56(12):138-146.
PI Junbo, QI Shixiong, SUN Wenduo, et al. Entity and event recognition method for power grid fault disposal plans based on UIE framework[J]. *Electric Power*, 2023, 56(12):138-146.
- [60] LIU Peng, SUN Zhenfu, ZHOU Biao. An ELECTRA-based model for power safety named entity recognition[J]. *Applied Sciences*, 2024, 14(20):9410.
- [61] VINYALS O, FORTUNATO M, JAITLY N. Pointer networks[C]// *Proceedings of the 2015 Conference on Advances in Neural Information Processing Systems*. Montréal: Neural Information Processing Systems Foundation, 2015:2692-2700.
- [62] 陈伟,杨燕. 基于指针网络的抽取生成式摘要生成模型[J]. *计算机应用*, 2021, 41(12):3527-3533.
CHEN Wei, YANG Yan. Extractive and generative summarization model based on pointer network[J]. *Journal of Computer*

Applications, 2021, 41(12):3527-3533.

- [63] 何俊,刘鹏,聂勇,等. 基于 Seq2seq 实体关系联合抽取的电力知识图谱构建[J]. 实验室研究与探索,2022,41(7):1-5.
HE Jun, LIU Peng, NIE Yong, et al. Construction of power knowledge map based on joint extraction of Seq2seq entity relationship[J]. Research and Exploration in Laboratory, 2022, 41(7):1-5.
- [64] SABOUR S, FROSST N, HINTON G E. Dynamic routing between capsules[C]//Proceedings of the 2017 Conference on Advances in Neural Information Processing Systems. California: Neural Information Processing Systems Foundation, 2017: 3857-3867.
- [65] 冯曙明,胡天牧,杨永成,等. 基于胶囊网络的电力供应链风险识别模型[J]. 微型电脑应用,2022,38(8):32-34.
FENG Shuming, HU Tianmu, YANG Yongchegn, et al. Capsule network-based risk identification model for power supply chain[J]. Microcomputer Applications, 2022, 38(8):32-34.
- [66] RADFORD A, WU J, CHILD R, et al. Language models are unsupervised multitask learners[J]. OpenAI Blog, 2019, 1(8):9.
- [67] BROWN T, MANN B, RYDER N, et al. Language models are few-shot learners[J]. Advances in Neural Information Processing Systems, 2020, 33:1877-1901.
- [68] WEI J, BOSMA M, ZHAO V Y, et al. Finetuned language models are zero-shot learners[C]//Proceedings of the 10th International Conference on Learning Representations(ICLR 2022). [S.l.]: OpenReview.net, 2022.
- [69] LONG Ouyang, WU Jeffrey, XU Jiang, et al. Training language models to follow instructions with human feedback[C]//Proceedings of the 36th Conference on Neural Information Processing Systems(NeurIPS 2022). NeurIPS: Neural Information Processing Systems Conference, 2022, 35:27730-27744.
- [70] WEI J, WANG X Z, SCHUURMANS D, et al. Chain-of-thought prompting elicits reasoning in large language models[C]//Proceedings of the 36th Conference on Neural Information Processing Systems(NeurIPS 2022). Louisiana: Curran Associates, 2022, 35:24824-24837.
- [71] YAO Shunyu, YU Dian, ZHAO Jeffrey, et al. Tree of thoughts: Deliberate problem solving with large language models[C]//Proceedings of the 37th Conference on Neural Information Processing Systems(NeurIPS 2023). Louisiana: Neural Information Processing Systems Foundation, 2023, 36:11809-11822.
- [72] POLAK M P, MORGAN D. Extracting accurate materials data from research papers with conversational language models and prompt engineering[J]. Nature Communications, 2024, 15(1):1569.
- [73] LIU Wenjing, ZHANG Suxiang, SUN Yang, et al. New energy power domain question-method extraction and soft clustering [C]//Proceedings of the 9th International Conference on Communication and Information Processing (ICCIIP 2023). New York: Association for Computing Machinery, 2023:484-491.
- [74] HU Zhiqiang, LI Xinyu, PAN Xinyu, et al. A question answering system for assembly process of wind turbines based on multi-modal knowledge graph and large language model[J]. Journal of Engineering Design, 2025, 36(7/9):1093-1117.
- [75] YIN Chunlin, DU Kunpeng, NONG Qiong, et al. PowerPulse: power energy chat model with LLaMA model fine-tuned on Chinese and power sector domain knowledge[J]. Expert Systems, 2024, 41(3):1-17.
- [76] ZHAO Jinxiong, MA Zhicheng, ZHAO Hong, et al. Self-consistency, extract and rectify: knowledge graph enhance large language model for electric power question answering[C]//International Conference on Intelligent Computing. Singapore: Springer Nature Singapore, 2024:493-504.
- [77] TANG Wei, ZHANG Yue, MAO Xun, et al. Enhanced named entity recognition and event extraction for power grid outage scheduling using a universal information extraction framework[J]. Energies, 2025, 18(14):3617.
- [78] LUO Jingtang, YAO Shiyang, ZHAO Changming, et al. A federated named entity recognition model with explicit relation for power grid[J]. Computers, Materials and Continua, 2023, 75(2):4207-4216.

(编辑:陈丽萍)