

# 基于NLP的智慧电厂风险数据智能推送研究

郭旺, 黄敏, 任晓敏, 刘海军, 杜杰

(国电内蒙古东胜热电有限公司, 内蒙古 鄂尔多斯 017000)

## Research on NLP-based Intelligent Push Concerning Risk Data of Smart Power Plants

GUO Wang, HUANG Min, REN Xiaomin, LIU Haijun, DU Jie

(Guodian Inner Mongolia Dongsheng Thermal Power Co., Ltd., Ordos 017000)

〔摘要〕随着我国电力行业迅速发展,总装机容量不断增长,但由于受到技术水平、区域环境、机械设备等因素的影响,电厂的安全生产仍面临着诸多风险。基于自然语言处理技术,提出智慧电厂风险数据智能推送方法,以提升风险数据推送的准确率,进一步提高电厂风险智能管控水平。

〔关键词〕智慧电厂;自然语言处理;风险数据;智能推送

**Abstract:** With the rapid development of China's electric power industry, the total installed capacity has been growing up. However, due to the influence of technical levels, regional environment, mechanical equipment and other factors, the safety production of power plants have been still confronted with multiple risks. Based on the natural language processing technology, an intelligent risk data push approach for smart power plants is proposed, which is available to improve the accuracy of risk data push and further improve the intelligent risk control level of power plants.

**Key words:** smart power plant; natural language processing; risk data; intelligent push

中图分类号: TM711 文献标识码: A 文章编号: 1008-6226 (2023) 03-0076-03

## 0 引言

改革开放以来,经济发展带来能源消费不断增长,电力行业保供压力明显增大。太阳能、核能等电力资源日益丰富,电力结构复杂程度及规模日益增加,导致电力系统安全风险显著提升。电力系统作为复杂的人-机-环境交互系统,其安全运行受到多种因素影响。如何全面管控电力系统风险要素成为电力企业关注的焦点。

## 1 电厂风险数据智能推送需求

2022年2月,国家能源局印发《水电站和小散远发电企业安全风险隐患排查整治专项行动方

案》和《电力行业危险化学品安全风险集中治理实施方案》,针对电力行业不同主体、不同风险环节,进一步强化电力安全生产监管工作。电力企业作为电力生产单位,对安全生产过程中的风险预控管理模式也进行了探索和实践,形成了一批典型的经验做法。

风险数据报送作为风险预控管理的首要环节,其数据质量直接影响电力企业风险预控管理水平和效果。但整体来看,我国电力行业风险数据报送水平仍有待提升:一是电力行业风险数据类型差异较大,企业间风险数据库难以统一;二是“两票”工作制度影响风险数据报送的效果,在“两票”作业中,风险数据报送及控制措施依赖工作经验、历史票样,风险数据不全面、不准确以及缺乏客观性等

问题普遍存在。

因此,为实现电力行业风险的精准管控,采用“碎片化处理”模型作为风险智能推送的基础数据库,采用来自变换器的双向编码器表征量 (bidirectional encoder representations from transformers, BERT) 模型对风险数据进行最佳的语义表征,借助 AI 自学习算法整合提炼出新的知识,自动关联用户最终采用的风险数据和已有风险数据,实现系统自我学习,丰富和拓展自身知识库,提高模型推送风险数据的准确率,增强系统的智能化水平。

## 2 智慧电厂风险数据库

采用较为成熟并广泛应用的“4M 风险屏障”理论,通过对数据进行标注,训练基于循环神经网络 (recurrent neural network, RNN) 的命名实体模型,快速对电厂各项风险数据中的设备名、地名等进行“碎片化处理”,建立覆盖运行作业、检修作业、区域环境、设备故障、管理五个方面的全方位电厂风险数据库,为智慧电厂风险智能推送提出基础数据库支撑。

## 3 智慧电厂风险数据智能推送方法

### 3.1 电厂风险通用语言模型

BERT 作为基于深度学习的双向编码表征模型,是近年来自然语言处理 (NLP) 领域公认的里程碑模型。其采用注意力 (Attention) 机制计算输入-输出 (input-output) 之间的关系,当对某个词进行处理时,能同时考虑该词前、后单词的信息,从而获取上下文的语义。该模型已在多个方向的 NLP 领域测试中表现出良好的应用效果。

鉴于电厂风险数据集数据量不大,可采用迁移学习方法,通过对预训练 BERT 模型进行微调 (fine-tuning),进而构建电厂风险通用语言模型。fine-tuning 过程包括遮挡语言模型 (MLM)、上下句预测 (NSP) 两个训练任务,其所需数据均通过随机算法生成。

为保证 fine-tune 后 BERT 模型的泛化能力,在更新参数时,仅对与下一任务最相关的子网络进行调整即采用基于子网络的调整方式

(child-tuning): 第一步,利用费雪信息矩阵 (fisher information matrix, FIM) 估算各参数与下一训练任务的相关性,采用其对角矩阵测算其重要性,选出与下一训练任务最相关最重要的参数,并生成其对应的权重;第二步,在后向传播权重后,仅对子网络中的参数进行调整。

### 3.2 电厂风险文本相似度模型

基于 BERT 构建文本相似度模型主要有无监督学习和监督学习两种渠道:无监督学习将每个文本与所有文本分别通过 BERT 模型计算句向量,利用余弦相似度等方法计算相似度得分;监督学习则是通过将相似度模型转化为分类模型,在提高测算效果的同时,能够有效解决无监督学习带来的向量分布非线性和奇异性。因此,选取监督学习方式构建电厂风险相似度模型。

具体做法为,将相似度任务转化为分类任务,即在句向量输出后添加一定数量的线性 (Linear) 全连接层和 Softmax 层 (对应输入应用归一化指数函数)。训练数据为三元组形式,如:(A 磨煤机检修,检修 A 磨煤机,1)、(A 磨煤机检修锅炉 0 m,检修 A 磨煤机锅炉 0 m,1)、(A 磨煤机检修,检修 B 磨煤机,0)、(A 磨煤机检修,A 磨煤机送电,0),其中前两个维度表示两个文本,第三个维度为 0 或 1 (1 表示同一个文本,0 表示非同一文本)。鉴于此环节采用的 BERT 模型是已使用电厂风险数据进行 fine-tuning 后的模型,因此这部分参数不需要训练,所以文本相似度模型训练速度较快。

### 3.3 电厂风险数据智能推送

系统根据任务描述、任务地点自动推送相关风险数据。

第一步,根据“两票”开票过程中输入的任务描述、任务地点等关键性字段,基于命名实体识别、分词等方法,快速对输入内容中的设备名、地名等进行“碎片化处理”,提取关键信息,并在基础数据库寻找相关风险数据。

第二步,将任务描述、任务地点等拼接为一个文本输入,并与基础数据库的任意一个文本输入构成四个二元组输入,即 (任务描述,任务描述)、(任务描述,任务描述+地点)、(任务描述+地点,任务描述)、(任务描述+地点,任务描述+地点),将所有的二元组输入到文本相似度模型,预测出相似度得分,取高于阈值对应的风险数据。

第三步,将两部分分别推送的数据取并集处理,作为“两票”智能化推送的数据。需要注意,在文本相似度模型中,当基础数据库规模较大时,可以采用多 GPU 进行分布式训练文本相似度模型,或者基于大数据框架进行分布式处理。

## 4 结束语

综上所述,基于人工智能、自然语言处理等技术设计了一套覆盖电厂风险数据库、电厂风险通用语言模型、电厂风险文本相似度模型在内的电厂数据智能推送方法,该方法的应用有效实现了“两票”开票过程中输入的任务描述、任务地点等关键性字段,进行电厂风险数据的智能推送和模型的自学习,为实现智慧电厂风险的精准管控提供了技术支持,为信息技术在电力行业安全生产管理方面的推广应用提供参考。

## 参考文献:

- [1] 李 刚,赵琳颖,吕 进.基于“认知-约束”的电力系统安全风险影响因素关联分析方法研究[J].电力信息与通信技术,2022,20(4):70-78.
- [2] 张成刚.电力企业安全生产的安全风险分级管控[J].电

力安全技术,2020,22(3):17-19.

- [3] 廖书长.基于安全生产风险管理体系的电力企业员工培训需求分析[J].中国电力教育,2021(3):22-24.
- [4] 沈立力,姜 鹏,王 静.基于 BERT 模型的中文期刊文献自动分类实践研究[J].图书馆杂志,2022,41(5):109-118,135.
- [5] 宋张君.风控 4M 屏障理论模型构建及机理研究[J].中国电力企业管理,2021(3):58-60.
- [6] 黄 敏,马 克,黄熙媛,等.基于 4M 屏障理论的智慧火电厂风险数据库管理及应用[J].电子技术与软件工程,2020(22):161-162.

收稿日期:2022-09-15;修回日期:2022-11-07.

作者简介:

郭 旺(1989—),男,工程师,主要从事风险分级管控及隐患排查治理双重预防机制建设工作,email:798406924@qq.com.

黄 敏(1970—),男,工程师,主要从事企业安全、职业健康、环保管理工作。

任晓敏(1985—),男,工程师,主要从事制度建设、环保管理、安全风险预控等工作。

刘海军(1974—),男,高级工程师,主要从事隐患管理工作。

杜 杰(1989—),男,工程师,主要从事生态环保管理工作。

(上接第 75 页)

止“超能力、超强度、超定员”施工作业。

(2) 合理安排人员计划性施工作业,减少人员工作时间过长情况,保证人员有足够的休息时间。

(3) 发挥安全员监督的作用,开展预防过度疲劳的安全教育,加强人员心理疏导、强化安全监督。

(4) 适当提高员工薪酬和福利待遇,阶段性地给予经济或荣誉激励,改善人员生活条件。

## 5.3 自我调节的建议

合理安排工作和生活,适当参加体育锻炼,提高自我对疲劳危险的观念认识,保证充足睡眠和休息时间,并以积极的心态参与到安全生产中,保证工作中拥有充沛的精力和体力,勇于杜绝过度疲劳作业,树立安全意识,保障人身安全。

## 6 结束语

外协施工队伍是电力建设不可或缺的力量。作

为外协人员,在繁忙的工作和生活中,应通过社会、项目部、个人的综合发力,消除过度疲劳,预防人员事故,保障企业和谐稳定和安全水平提升。

## 参考文献:

- [1] 国网湖北省电力有限公司.电力安全心理评估研究[M].北京:中国电力出版社,2020.
- [2] “‘绿十字’安全基础建设新知丛书”编委会.安全心理学运用知识[M].北京:中国劳动社会保障出版社,2014.

收稿日期:2022-06-20;修回日期:2022-08-10.

作者简介:

崔家凯(1973—),男,工程师,主要从事供电企业安全生产工作,email:cuijiakaicjk@163.com.

崔家增(1970—),男,助理工程师,主要从事供电企业安全生产工作。