

# GPT等大规模语言模型在电气设备状态评估领域的应用与挑战

马国明<sup>1</sup>, 秦炜淇<sup>1</sup>, 王渊<sup>1</sup>, 蔺家骏<sup>2</sup>, 胡俊华<sup>2</sup>, 李晨<sup>2</sup>

(1. 华北电力大学新能源电力系统国家重点实验室, 北京 102206; 2. 国网浙江省电力有限公司, 杭州 310014)

**摘要:** 大规模模型对电气工程领域的影响将扩展到产业、教学及科学研究各个领域的多个环节。为促进电气工程领域研究人员对大规模语言模型的理解,文中首先介绍了GPT等大型语言模型的发展历程;其次针对GPT的原理与工作流程进行了阐述;最后讨论了GPT的优势与局限性。然后针对大规模语言模型在电气设备状态评估领域的应用机遇进行了详细分析,并提出了可能存在的问题。最后为了更好的发挥大规模语言模型的作用,提出了需要进一步开展的工作与发展建议。希望为大规模语言模型在电气设备状态评估领域的发展提供参考。

**关键词:** 大规模语言模型; GPT; 机遇与挑战; 应对措施

## Application and Challenges of Large-scale Language Models Such as GPT in the Field of Power Equipment Condition Assessment

MA Guoming<sup>1</sup>, QIN Weiqi<sup>1</sup>, WANG Yuan<sup>1</sup>, LIN Jiajun<sup>2</sup>, HU Junhua<sup>2</sup>, LI Chen<sup>2</sup>

(1. State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China; 2. State Grid Zhejiang Electric Power Co., Ltd., Hangzhou 310014, China)

**Abstract:** The impact of large-scale models on the field of electrical engineering will extend to various aspects of industry, education, and scientific research. To promote a better understanding of large-scale language models among researchers in the electrical engineering field, in this paper the development history of large-scale language models such as GPT is introduced firstly. Then, the principles and workflow of GPT are described. After that, the advantages and limitations of GPT are discussed. Furthermore, the application opportunities of large-scale language models in the field of electrical equipment condition assessment are analyzed in detail, and the potential problems are also proposed. Finally, further work and development suggestions are proposed in order to better utilize large-scale language models. The aim of this paper is to provide a reference for the development of large-scale language models in the field of electrical engineering.

**Key words:** large-scale language models; GPT; opportunities and challenges; response measures

## 0 引言

如何提高电气设备状态评估的准确度与智能化水平一直是本领域的重要研究问题。2022年11月, OpenAI 上线 ChatGPT, 以 GPT(generative pre-trained transformer)为代表的大型语言模型在各个行业引起了广泛讨论。

此类大型语言模型可以追溯到1950年代人工智能的开启。1950年代,第一批人工智能程序被开发出来,例如,1954年乔治城大学和IBM进行了一场翻译实验,展示了机器语言翻译的潜能。由于翻译的句子提前进行了精心挑选,翻译结果比较成功,获得了广泛的关注与报道。1960—1970年代,机器翻译与自然语言处理(natural language processing,

NLP)进步迅速,诞生了第一个基于语法分析的机器翻译系统。1980年代,神经网络被提出,但由于当时计算机算力和数据不足,其发展受到限制。1990—2000年代,自然语音处理主要聚焦在统计模型和规则系统(rule-based systems)。在语料库方面出现了以华尔街时报语料库为代表的大型语料库,促使模型更加准确。2010年以后,深度学习首先在图像处理领域取得了重要应用突破,其后也显著改变了自然语言处理领域的面貌,诞生了一系列强有力的大型语言模型。

这些大型语言模型的影响逐渐扩展到语言翻译以及计算机行业之外。在电气设备状态评估领域,也开始了此类语言模型的研究和探索<sup>[1-5]</sup>。

文[6-8]构建了电气领域的字典,类似规则系统对文本信息进行匹配查询,此类方法对字典的覆盖范围要求较高,对字典中不存在的实体信息无法准确识别,且不能通过上下文语义来准确判断实体信息的种类。此前深度学习在电气工程非结构化文本挖掘方面的应用多集中于文本分类方向。文[9]采用循环卷积神经网络(RNN)对1 000余条变压器故障文本进行分类,按照故障位置、故障现象等信息将设备状态分为“故障、普通、一般”3类。文[10]采用双向长短时记忆与注意力机制(Bi-LSTM+Attention)进行电气设备缺陷文本分类,筛选出对文本分类有影响的电气信息。上述文本分类研究工作分类个数较少的情况下准确率达到90%以上,但缺乏对电气设备文本中具体实体信息识别研究,只能概括地给出整体文本分类结果。文[11]通过使用自注意力机制将电气设备文本的上下文语义信息进行融合,实现了实体信息的高准确率识别。对560篇电气设备文本进行了实验,与传统字典匹配的方法相比 $F_1$ 值最高提升了34.38%,与word2vec方法相比 $F_1$ 值最高提升了11.25%。

尽管上述研究在电气设备缺陷文本识别领域取得了一定的成果,但是随着2023年4月GPT-4的发布,其智能化水平进一步增强,为大规模模型在电气设备状态评估领域的发展提供了新的契机。

文中将首先介绍GPT类大型语言模型的发展历程,其次具体阐述GPT的工作原理,然后讨论GPT的优势与局限性,最后详细分析GPT对电气设备状态评估领域造成的影响与应对措施。希望对于该领域相关人员理解GPT及其影响有所帮助。

## 1 GPT类大型语言模型的发展历程

大型语言模型的本质是将非线性语言投射到

高维线性空间。

1990年代,研究人员首先利用循环神经网络(recurrent neural network,RNN)对文本建模并特征提取。

2013年,谷歌大脑(google brain)团队的Tomas Mikolov等人发表文章,介绍了word2vec预训练方法,它的特点是能够将单词转化为向量,挖掘词之间的联系,进行分布式词语表达(distributed representations),该方法显著降低了计算运算量,加快了训练过程,受到了业界普遍欢迎<sup>[12]</sup>。

2014年左右,研究人员提出应用1997年Hochreiter和Schmidhuber提出的长短时记忆网络(long short term memory,LSTM)<sup>[13]</sup>这类特殊的RNN网络对自然语言进行处理,该方法可以保留重要文本信息并遗忘不重要信息,更加准确的提取长序列特征,特别适合处理长距离文本信息<sup>[14]</sup>。

2017年,谷歌大脑团队的Ashish Vaswani发表了自然语言处理领域的重磅文章“Attention is all you need”,提出了采用注意力机制(attention)的Transformer神经网络模型<sup>[15]</sup>。注意力机制的本质是借鉴了人类的注意力思维方式。通常人类会快速浏览扫描全局文本的图像,得到需要重点关注的区域,然后再进一步对该区域的上下文投入更多的注意(attention)。这种注意力机制提高了人类信息处理的效率和准确性。谷歌大脑团队提出的这种自注意力机制可以按输入文本中各部分内容重要性分配文本权重,为后续大规模语言模型奠定了基础。此外由于Transformer在网络层面抛弃了传统的RNN模型,解决了RNN不能很好并行计算的问题,极大提高了大型语言模型进行并行训练的效率。除了自然语言处理领域,Transformer结构也被应用到了计算机视觉领域,其效果显著优于此前的卷积神经网络模型(CNN)。

此后的大型语言模型多以Transformer为基础,可以分为纯Encoder模型,又被称为自编码 auto-encoding 模型;纯Decoder模型,又被称为自回归 auto-regressive 模型;Encoder-Decoder模型,又被称为 sequence-to-sequence 模型。上述大型语言模型典型代表如下:

GPT系列语言模型(generative pre-trained transformer):由OpenAI公司2018年起开始发布,可以用于语言翻译和机器问答,能够生成篇幅较长且语义连贯的文本。

BERT语言模型(bidirectional encoder representation from transformers),由谷歌公司2018年推出,将多层transformer进行串联组合,根据文本上下文

语义信息动态生成字向量,在自然语言处理各个领域取得出色效果。文中也基于此开展了电气设备实体信息识别与缺陷诊断等领域的研究<sup>[1]</sup>。

RoBERTa 语言模型(robustly optimized BERT approach),由 Facebook 公司 2019 年发布,它通过使用更多的预训练数据和训练时间提高模块智能水平。

T5 语言模型(text-to-text transfer transformer),由谷歌公司 2019 年发布,采用 Transformer 的 encoder-decoder 结构。其基本思想是将每个 NLP 问题都视为“text-to-text”问题,即将文本作为输入并生成新的文本作为输出。

2019 年之后语言模型层出不穷。

2023 年 3 月百度发布“文心一言”,具备跨模态、跨语言的深度语义理解与生成能力,可以实现数理逻辑推算、中文理解、多模态生成等功能。

2023 年 4 月阿里巴巴“通义千问”开始邀请用户测试体验,它支持多轮对话,文案创作,逻辑推理,多模态理解,多语言支持,可以提高多场景下的工作效率。

2023 年 5 月 11 日,谷歌发布通用语言大模型 PaLM 2,其性能得到全方位大幅提升,部分语言能力基准测试结果超过 GPT-4。

2023 年 6 月 27 日,北京市科委发布北京市首批 10 个行业大模型应用案例,国网智能电网研究院联合百度打造的“基于电力行业 NLP 大模型的设备运检知识助手示范应用”上榜,推动了电力设备运检的智能化水平。

近年主要大规模语言模型发展见图 1。

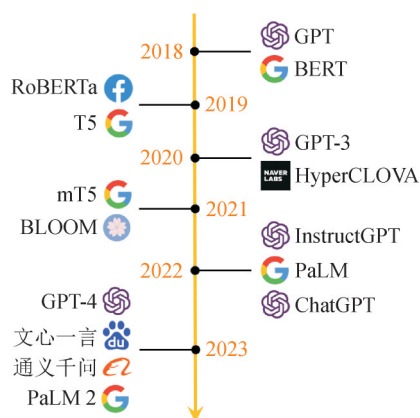


图 1 近年主要大规模语言模型时间轴

Fig. 1 Timeline of major large-scale language models in recent years

## 2 GPT 原理

如前文所述,GPT 是建立在 Transformer 基础上的。2018 年 6 月 OpenAI 发布了其第一版的 GPT,包

含了 1.17 亿个参数,预训练数据量 5 GB。GPT 通过输入海量数据,训练网络模型。2019 年 2 月,GPT-2 发布,参数数量增加到了 15 亿个,预训练数据量 40 GB,大大提高了生成文本的质量和多样性,此时人们已经难以区分文本是模型输出还是人为写作。2020 年 5 月 GPT-3 发布,参数数量进一步增加到了 1 750 亿个,预训练数据量 45 TB,是当时最大的语言模型,它不仅生成自然语言文本,还能进行翻译、问答等多种任务,成为了人工智能领域的里程碑之一。2023 年 3 月,GPT-4 公布,参数数量目前尚未公布。历代 GPT 对比见表 1。

表 1 历代 GPT 对比

项目	初代 GPT	GPT-2	GPT-3
时间	2018-06	2019-02	2020-05
参数量/亿	1.17	15.40	1 750.00
预训练数据量/GB	5	40	45×1 024
序列长度	512	1 024	2 048
Decoder 层数	12	48	96
隐藏层数	768	1 600	12 288

GPT-3 以后没有开源,但其基本原理没有变化,GPT 的核心是生成式预训练与 Transformer 模型。

### 2.1 生成预训练 Generative Pre-trained

GPT 本质是一个大型的神经网络,未经训练的模型参数是随机初始化的,预训练的作用是使得模型能够对自然语言有深入的理解。GPT 中预训练结合了无监督的初步预训练和有监督的微调。

预训练第一步,生成式预训练是指 GPT 有能力产生下一个与输入提示相关、风格相同的文本。与 BERT 不同,OpenAI 坚持用上文进行训练,即通过训练使模型能够根据上文提示,来猜测下一个词。

在模型结构方面,GPT 仅使用了 Transformer 的 Decoder 结构,并对 Transformer Decoder 进行了一些改动。Transformer Decoder 与 GPT Decoder 结构对比见图 2,原本的 Decoder 包含了 MHA、MMHA,而 GPT 只保留了 MMHA,这确保了 GPT 只能关注上文的信息,从而达到单向模型的目的。

这种方法相比于 BERT 根据上下文来填空式的预训练,无疑要难上许多。这使得 GPT 必须模型够大、数据够多才能训练起来。因此随着 GPT-3 参数数量、训练数据显著增大后,GPT 的优势才逐渐显现。GPT-3 接受 token(分词结果)的宽度为 2 048,可以称之为“上下文窗口”。这意味着他有 2 048 个路径,每个 token 都要沿着这些路径进行处理。GPT-3 的这些计算发生在 96 个 Transformer 的 decoder 层

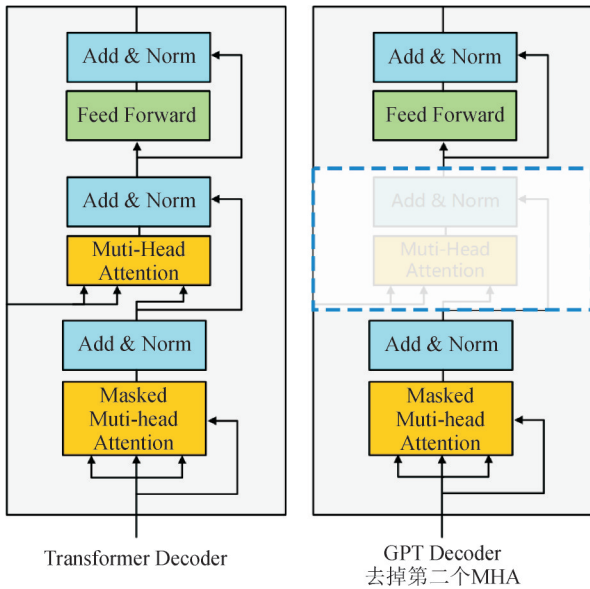


图2 Transformer decoder与GPT decoder结构对比

Fig. 2 Structural comparison of transformer decoder and GPT decoder

中。每个层都有大约单独的18亿参数。GPT-3训练使用的算力为355 GPU年,成本460万美元,GPT-3训练数据集高达3 000亿个文本tokens。

预训练第二步,是通过微调(Fine-tuning)或者针对特定任务调整权重,这一部分有时需要采用有标签的数据完成。针对不同的应用领域应采用不同的微调数据,这样可以确保语言理解的准确性。例如“高压”在健康领域是“高血压”,在油气领域是“高气压”,在电气领域是“高电压”。

### 2.2 Transformer模型

GPT采用了Transformer模型与自注意力机制(attention),采用自注意力机制的作用是使字向量编码过程中有选择地关注周围重点字符而非平均地关注所有字符信息,有助于增加实体信息识别的字符信息权重,减少干扰信息权重,综合考虑一句话中的所有字符与当前字符的关联度,得到最优的当前字符全局表达向量。这部分内容可以参考文[11],这里不再详述。

在实际工作中,人们大量接触的是GPT衍生出的ChatGPT,ChatGPT模型的流程图见图3。

第1步是预处理。首先进行文本数据的收集和清洗;其次将文本按照固定的长度进行分词;最后于分词结果建立词汇表。

第2步是建立大型模型。首先采用多层Transformer Decoder模型;其次,每一层都包括注意力机制与前向神经网络;然后,前一层的输出是下一层的输入。最后,最后一层的输出连接到Softmax函数获得概率分布。

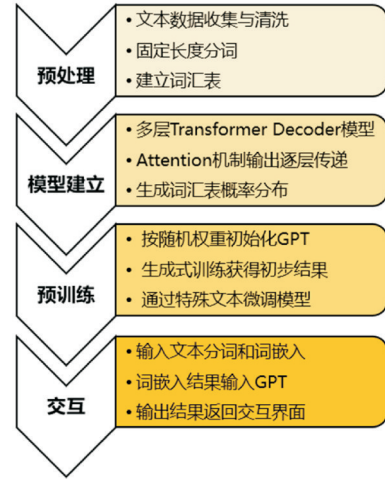


图3 ChatGPT模型流程图

Fig. 3 Model flowchart of ChatGPT

第3步是预训练。首先,给GPT模型赋随机权重。其次,结合有监督学习与无监督学习,通过生成式预训练获得初步训练结果。然后,通过特殊文本内容微调模型模型,进一步提高训练效果。

第4步是交互。首先,将人类输入的句子分词和词嵌入(把词映射为实数域上的向量embedding)。其次,将词嵌入结果输入到GPT获得输出响应。最后,将输出结果返回显示到交互界面。

### 3 GPT模型自身的优缺点

通过使用与调研分析,确定的GPT模型的优点如下:

1)适用性广泛。GPT模型可以满足语言生成,语言翻译,语言分析等多方面的工作需要。在电气设备状态诊断领域资料的翻译与整理等方面应用广泛。

2)输出质量高。相比其他大规模语言模型,GPT输出文字与输入内容的相关性高,输出文字的完整性也优于其他语言模型,更接近人类语言。由于数据量大,GPT在回答电气设备状态诊断领域问题是准确率显著高于其他模型。

3)训练效率高。由于使用了Transformer模型,相比其他大规模语言模型,GPT模型的训练效率更高。未来随着多模态信息的发展,大量电气设备状态诊断领域的图像、声音资料将作为输入,训练效率对于模型收敛至关重要。

4)持续学习能力强。随着输入数据的增加与广泛使用,GPT模型可以持续不断的提高自身的语言处理能力。电气设备状态诊断领域大量专家的交互使用将有助于持续提高模型诊断的准确率。

在具有明显优点的同时,GPT模型目前也还存

在一些不足:

1)需要计算资源多。GPT-3.5的参数数量是700亿个,需要海量的资源进行计算。由于GPT-3.5所需资源未对外公布,参考GPT-3计算资源,训练GPT-3共使用了数千个GPU单元和张量处理单元(tensor processing unit),消耗了3~4 GWh的电能。

2)输入敏感。GPT对输入短语的调整或多次尝试同一提示很敏感。例如,给定一个问题一个短语,模型可以声称不知道答案,但稍微重新措辞,可以获得正确回答。

3)有限的可解释性。首先,GPT的内部架构和训练过程极其复杂。其次,GPT-3.5的参数数量已经达到700亿个,深度神经网络的输入数据和训练过程都是高度非线性的,这使得人类难以对这一过程进行理解。

4)输出偏差。由于输入语料库自身的限制,GPT的输出可能存在与事实不符的情况,特别是对于一些小众的专业知识,ChatGPT的输出有时看似合理但实际不正确或荒谬。

5)道德隐忧。GPT等语言模型功能强大,其潜在的滥用和偏见都可能引发严重的社会和道德问题,必须制定相关政策,以确保安全的使用此类人工智能。

### 4 在电气设备状态诊断领域的应用机遇

自动化和人工智能在电气工程领域已经取得了很多的应用,数字化、大数据、知识图谱等工具的应用都逐渐改变了电力系统的面貌,大规模语言模型的发展将进一步提高电气设备状态诊断的智能化水平与效率。具体的应用场景与机遇如下:

1)有利于基础知识传承,加快青年专家发展。通过针对电气设备状态诊断领域培训知识的微调,可以生成一个适用于青年员工培训的语言模型,回答员工关于基础知识、事故案例等相关问题,加快员工培养速度。

2)有利于海量数据的查询与比对,减少工作时间。经过长年发展,在电气设备状态诊断领域形成了大量标准,而其中一些关键数据可能散布在不同的标准之中,应用大规模语言模型可以方便的查找这些标准,提升工作速度,对于规范的开展工作有巨大助益。

3)有利于历史信息、故障案例的整理。在各电科院、地市公司都存有大量的研究报告、故障分析报告、事故处理报告,这些重要的资源没有得到充

分的利用,通过这些历史信息通过多模态的方式输入,可以传承宝贵的历史经验。

4)有利于事故辅助分析,加快运维策略制定。当电力系统发生事故以后,常规方式是首先收集现场资料,包括但不限于多种传感器测量结果,然后召集专家进行讨论分析,充分发挥专家经验,得出事故原因和改进措施。应用大规模语言模型充分的知识储备,可以对事故原因进行初步分析,提供历史类似案例,加快事故分析速度,辅助专家进行决策。

5)有利于发现家族缺陷,改造升级现有设备。将大量的事故报告和设计经验都输入大型语言模型,可以进一步凝练出事故原因、比例、可能的改进方向等。通过应用大规模语言模型可以总结历史经验,发现家族缺陷,为电力设备结构、材料、工艺等方面的优化改进提供参考。

6)有利于填报日常报告,提高工作效率。日常工作中存在大量相对重复的表格填写工作,由于管理关注点不同,不同部门对格式的要求不同,内容各有侧重。采用大规模语言模型可以方便的生成日常报告,降低日常工作强度。

7)有利于加快文献综述整理,启发科研创新思路。通过使用大规模语言模型可以快速的收集现有知识,并对知识进行整理,形成文献综述。有利于启发研究思路,通过多学科知识的融合,提示研究思路。但目前这种综述仅有知识的罗列,没有深入的思考和见解,仅能为研究人员提供参考。

在实际GPT应用中,首先将现场图像信息、故障事件背景信息、基础数据、基础信息等输入数据库,然后可以要求GPT给出专业知识、整理故障案例、查询对比数据、诊断设备状态、发现家族缺陷、填写日常报告、启发科研思路等(见图4)。

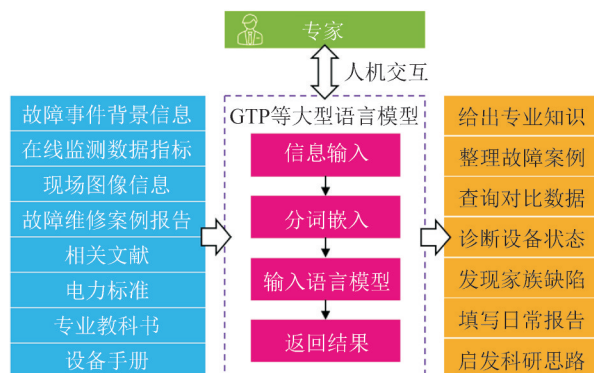


图4 基于GPT模型的电气设备状态诊断框架

Fig. 4 Condition diagnosis of electrical equipment based on GPT model

为了更加清楚地展示GPT在电气设备状态评

估领域的应用方法, 给出一个典型案例, 首先在GPT中输入了事故现场的信息和结构化数据, 然后通过问答方式提炼核心要点, 分析故障原因, 凝练上报报告, 具体如下。

输入: 请根据上述输入内容总结报告核心点。

输出: 根据提供的信息, 以下是对上述事故分析的进一步总结: ①事故经过: 在处理单相接地过程中, 110 kV XX变1号主变后备保护动作跳开10 kV开关, 导致10 kV I段母线失压。事故发生后, 10 kV I段压变烧毁严重, 短路电流达到9 200 A; ②保护SOE故障信息: 通过监控SOE信息可知, 事故发生前, 10 kV I段母线上的间隔保护多次报告了接地报警和线路PT断线的告警信息; ③主变试验数据情况: 主变本体油色谱试验未显示特征气体增长, 绕组频谱变形图谱分析显示中、低压侧有轻度变形, 主变介损和电容量测试结果表明中、低压侧存在绕组变形情况。低电压短路阻抗超过规程要求; ④跳闸后油化验结果: 油中氢气、乙烯、一氧化碳等组分体积分数略有增长, 但未达到异常水平; ⑤事故原因分析: 事故的根本原因是10 kV I段母线发生多次接地故障, 导致谐振过电压和C相压变铁心过饱和, 绝缘性能下降并最终导致短路故障和主变损坏; ⑥临时处理措施: 对10 kV压变柜进行修复, 主变暂时投运, 加强监测和检测工作, 避免再次发生短路故障。故障主变已更换, 并计划将其送厂进行大修。综上所述, 该事故是由10 kV I段母线多次接地故障引起的, 导致了主变损坏。中、低压侧绕组变形严重, 导致短路阻抗偏小, 频谱分析和介损测试也显示了绕组变形的迹象。临时处理措施是修复压变柜并暂时投运主变, 同时加强监测工作。通过更换故障主变来解决问题, 并计划进行大修。

输入: 多次接地故障导致主变损坏的原因是什么?

输出: 多次接地故障导致主变损坏通常是由以下原因引起的: ①连续过电流。多次接地故障可能导致电流通过主变的绕组, 产生过高的电流。这种连续过电流会导致主变内部的绝缘材料受到过热和损坏; ②绝缘破坏。多次接地故障会导致电流通过主变的绝缘系统, 使绝缘受到过电压和过电流的冲击。长时间的过电压和过电流会导致绝缘材料的破坏和老化, 进而导致绝缘性能下降, 最终引发主变损坏; ③机械振动。多次接地故障可能引起主变的机械振动, 这可能导致主变内部的绕组和绝缘件松动或磨损。长期的机械振动会导致绝缘材料破裂或磨损, 增加绝缘失效的风险; ④瞬态电压和电流。多次接地故障会引发电力系统中的瞬态电

压和电流, 这些瞬态电压和电流可能超过主变所能承受的额定电压和电流水平。这样的瞬态过电压和过电流会对主变的绝缘系统造成冲击, 导致绝缘失效和主变损坏。

综上所述, 多次接地故障会导致主变内部发生过电流、绝缘破坏、机械振动和瞬态电压电流等问题, 进而导致主变损坏。为了减少这种损坏的风险, 电力系统运行和维护中需要采取合适的保护措施, 并定期进行绝缘测试和设备检查, 以确保主变的正常运行和安全性能。

输入: 将上述结果凝练成100字报告。

输出: 发生在XX变的事故由10 kV I段母线多次接地故障引起, 导致主变损坏。频谱分析和介损测试显示绕组变形, 短路阻抗偏小。临时措施是修复压变柜, 投运主变并加强监测。故障主变已更换, 计划大修。事故根源解析为谐振过电压和C相压变铁心过饱和。油化试验结果略有异常。需要进一步检查和维护。

## 5 在电气设备状态诊断领域应用中存在的问题

人们往往关注于人工智能的有利影响, 而对于人工智能造成的潜在不利影响缺乏重视, 大规模语言模型造成的潜在不利影响如下。

1) 可能产生知识误导。通过多次的检索尝试发现, ChatGPT在数据和理论方面存在误导, 存在事实偏差, 对于探索性或此前经验较少的问题它可能给出看似合理实则错误的回答。另外ChatGPT研发的主要目的是为了聊天(Chat), 因此很多时候它会按照输入人的思路进行回答, 对于不同人的相同问题, ChatGPT给出的答案可能完全相反。

例如, 当研究人员向ChatGPT提问: “绝缘物体会影响空间磁场分布吗”, ChatGPT会给出回答: “是的, 绝缘物体可以影响空间磁场的分布。虽然绝缘物体本身不导电, 但他们可以改变磁场的路径和分布方式。当磁场遇到绝缘物体时, 绝缘物体会产生磁场的局部扭曲或偏转, 这是由于绝缘物体的磁导率和磁场的相互作用引起的。”然而当研究人员再次提问: “为什么绝缘物体不会影响空间磁场分布”, ChatGPT会根据提问者的态度做出完全相反的回答: “绝缘物体通常不会对空间磁场的分布产生直接的影响。电流是产生磁场的主要源之一, 而磁材料可以在其内部形成磁场。绝缘物体通常不具备这些特性, 因此他们对磁场的分布没有显著影

响。总的来说,绝缘物体对空间磁场的分布影响较小。磁场主要受到电流和磁材料的影响,而绝缘物体通常不具备这些特性。”因此在使用此类大型语言模型时,必须审慎的对待其答复,可以参考,不能盲从。

2)有可能存在数据失真。现有大规模语言模型的答案经常出现不准确或不完整的结果,特别对于时间等数据信息存在较大误差,如果滥用这些基础数据,可能造成严重错误。例如询问“铝的介电常数是多少”答复“铝的介电常数(相对介电常数)约为8.5。”答案显然是错误的。

3)有可能产生依赖。现有大规模语言模型可能使新进研究人员对其产生依赖心理,如果模型认为某个思路不行,或者某种技术无法实现,有可能降低新进研究人员对研究的信心。

4)有可能抑制创新。由于大量使用大规模语言模型,减少了科学研究中收集研究资料的复杂度,但是在收集整理资料过程中形成的其他创新思维的机会也显著降低,探索性思维受到抑制。

## 6 发展建议

为了更好的发挥大规模语言模型的作用,需要进一步开展的工作与发展建议如下:

1)应用电气设备状态诊断领域知识对大规模语言模型进行参数微调。大规模语言模型的核心是神经网络架构与语料库,因此选择适用于电力系统的语料库对语言模型进行训练十分必要,相关研究也表明,使用电力文本信息语料库可以显著提高诊断的准确度<sup>[11]</sup>。目前中国电力科学研究院、国网智能电网研究院以及多个省级电科院已经积累了大量的故障案例,开展了相关探索。在应用电气设备状态诊断领域知识时,应特别注意样本的准确性与正负样本的均衡性。

2)从单模态扩展到多模态、跨模态。目前GPT4已经可以实现多模态的分析。在电气设备诊断领域,存在油色谱、局放检测结果等结构化数据、试验报告等非结构化数据以及红外、可见光成像等图像数据,应用多模态信息的联合分析可以实现交叉印证,提高缺陷诊断的准确性。可以利用百度“文言一心”中相关的电力资源对模型进行优化。

3)进一步提高神经网络参数参量与训练效率。从表1可以看出,GPT模型取得快速的发展与其神经网络参数的数量紧密相关。目前的研究方向是进一步提高神经网络参数参量进而提高诊断的准确性,但这也造成硬件设备成本的进一步提高。

4)持续改进,逐步提高诊断精度。大规模语言模型训练成本大、需要的计算资源多、存在输出偶有偏差等方面的问题,导致一些电气设备诊断领域研究人员和现场运维人员对其实用性存在一定的担心与质疑。但是,人工智能对于所有行业的改变是难以避免的,只有积极接纳先进技术、积极发展模型,才能进一步提高生产效率。随着数据量增大与人们对大规模语言模型理解的深入,诊断精度一定会进一步提高。

5)对于大规模语言模型输出结果应审慎辨别。如前所述,大规模语言模型可能产生严重的知识误导,因此只能作为参考,应特别关注结果是否正确,对于关键问题的答案应特别谨慎的对待,务必多方信息核对,避免造成重大损失。

6)应用大型语言模型对设备状态评价领域的一些经验公式进行可解释性探索。大型语言模型通常被看成黑箱模型,由于其结构复杂,参数数量数以亿记,通常难以理解其内部工作结构。OpenAI最新研究利用GPT-4自动进行大语言模型中神经元行为的解释,并将其应用于GPT-2,公开了GPT-2神经元解释的数据集,加快了人们对大规模语言模型的理解。在电气设备状态诊断领域,目前设备的状态评估、故障诊断等业务中存在大量的专家知识与经验公式,采用大型语言模型与类似方法有望增强对电气设备状态诊断领域中一些经验公式的理解。

7)针对现场应用场合推出轻量化边缘计算模型。如前所述,训练GPT-3共使用了数千个GPU单元和张量处理单元(tensor processing unit),消耗了3~4 GWh的电能,现有语言模型需要极多的软硬件计算资源,只能部署在云端。面对电气设备状态领域应用需求,后续可研制在移动设备端运行的轻量化模型,降低硬件需求,从而提供定制化、低时延、高安全性与隐私性的应用。

## 7 结论

1)总结了GPT类大型语言模型的发展历程,分析了语言模型的演进过程与不同语言模型之间的关联关系。

2)介绍了GPT的工作原理,分析了其与在电气设备状态诊断领域应用较多的BERT模型的异同,讨论了GPT的优势与局限性。

3)讨论了大型语言模型在电气设备状态诊断领域具体的应用场景,分析了其在传承专业知识,整理故障案例,查询对比数据,诊断设备状态、发现家

族缺陷,填写日常报告,启发科研思路等方面的应用机遇。

4)提示了大型语言模型在电气设备状态诊断领域应用时的潜在风险,包括可能产生知识误导、提供的数据失真、产生依赖心理及抑制创新等。

5)给出了大型语言模型在电气设备状态诊断领域需要进一步开展的工作与发展建议,包括应用本领域知识与案例对模型进行参数微调,持续改进模型进而提高诊断精度,扩展模态数量与跨模态,提高神经网络参数参量与训练效率,审慎辨别大规模语言模型输出结果,尝试应用大型语言模型对经验公式进行可解释性探索,针对不同场合推出个性化轻量化边缘计算模型等。

### 参考文献:

- [1] 和敬涵, 罗国敏, 程梦晓, 等. 新一代人工智能在电力系统故障分析及定位中的研究综述[J]. 中国电机工程学报, 2020, 40(17): 5506-5515.  
HE Jinghan, LUO Guomin, CHENG Mengxiao, et al. A research review on application of artificial intelligence in power system fault analysis and location[J]. Proceedings of the CSEE, 2020, 40(17): 5506-5515.
- [2] WANG Jian, WANG Xi, MA Chaoqun, et al. A survey on the development status and application prospects of knowledge graph in smart grids[J]. IET Generation Transmission & Distribution, 2021, 15(3): 383-407.
- [3] MENG Fanqi, YANG Shuaisong, WANG Jingdong, et al. Creating knowledge graph of electric power equipment faults based on BERT-BiLSTM-CRF model[J]. Journal of Electrical Engineering & Technology, 2022, 17(4): 2507-2516.
- [4] 蒲天骄, 乔 骥, 韩 笑, 等. 人工智能技术在电力设备运维检修中的研究及应用[J]. 高电压技术, 2020, 46(2): 369-383.  
PU Tianjiao, QIAO Ji, HAN Xiao, et al. Research and application of artificial intelligence in operation and maintenance for power equipment[J]. High Voltage Engineering, 2020, 46(2): 369-383.
- [5] 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.
- [6] 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 defect records of power equipment based on knowledge graph technology[J]. Automation of Electric Power Systems, 2018, 42(14): 158-164.
- [7] 曹 靖, 陈陆燊, 邱 剑, 等. 基于语义框架的电网缺陷文本挖掘技术及其应用[J]. 电网技术, 2017, 41(2): 637-643.  
CAO Jing, CHEN Lushen, QIU Jian, et al. Semantic framework-based defect text mining technique and application in power grid[J]. Power System Technology, 2017, 41(2): 637-643.
- [8] 贾全焯, 张 强, 宋博川. 一种基于循环神经网络的电网客服语音文本实体识别算法[J]. 供用电, 2020, 37(6): 13-20.  
JIA Quanye, ZHANG Qiang, SONG Bochuan. A text entity recognition algorithm based on recurrent neural network for customer service voice of state grid[J]. Distribution & Utilization, 2020, 37(6): 13-20.
- [9] 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. An information mining method of power transformer operation and maintenance texts based on deep semantic learning[J]. Proceedings of the CSEE, 2019, 39(14): 4162-4171.
- [10] 冯 斌, 张又文, 唐 昕, 等. 基于BiLSTM-Attention神经网络的电力设备缺陷文本挖掘[J]. 中国电机工程学报, 2020, 40(s1): 1-10.  
FENG Bin, ZHANG Youwen, TANG Xin, et al. Power equipment defect record text mining based on BiLSTM - Attention neural network[J]. Proceedings of the CSEE, 2020, 40(s1): 1-10.
- [11] 蒋 晨, 王 渊, 陈 珉, 等. 基于Kbert文本聚类模型的变压器故障识别[J]. 高电压技术, 2022, 48(8): 2991-3000.  
JIANG Chen, WANG Yuan, CHEN Min, et al. Transformer fault recognition based on kbert text clustering model[J]. High Voltage Engineering, 2022, 48(8): 2991-3000.
- [12] MIKOLOV T, CHEN Kai, CORRADO G, et al. Distributed representations of words and phrases and their compositionality[C]// Advances in Neural Information Processing Systems 26(NIPS 2013). USA: NeurIPS, 2013: 3111-3119.
- [13] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [14] SAK H, SENIOR A W, BEAUFAYS F. Long short-term memory recurrent neural network architectures for large scale acoustic modeling[C]// 2014 International Association for Voice Communications Annual Meeting. Singapore: ISCA, 2014: 21437.
- [15] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]// Advances in Neural Information Processing Systems 30 (NeurIPS 2017). USA: NeurIPS Foundation, 2017: 5998-6008.

马国明(1984—),男,教授,研究方向为电力设备先进传感与状态评估(通信作者)(E-mail: ncepumgm@ncepu.edu.cn)。