

机器学习在智能配用电领域中的应用:北美工程实践概述

李亦言^{1,2}, 胡荣兴¹, 宋立冬¹, 贾乾罡², 陆宁¹

(1. 北卡罗来纳州立大学电子与计算机工程系, 罗利市 27695, 美国;

2. 上海交通大学电子信息与电气工程学院, 上海市 200240)

摘要: 机器学习技术是助力能源转型、促进清洁能源消纳的重要工具。近年来,机器学习技术在电力系统中的应用已得到广泛关注。由于机器学习技术的“黑箱”特征,使其在可解释性、鲁棒性等方面仍有待提升,与电力系统高可靠性的运行要求存在一定矛盾,导致其实际工程应用滞后于理论研究。对于机器学习技术的实际应用情况,文中聚焦于北美地区配用电领域,从源、网、荷3个角度梳理了机器学习技术的典型工程实践项目,概述了每个项目的方法、效果以及从中得到的启示。进一步地,将以上项目归纳为态势感知、决策支持2个类别共计5个应用场景,并从工程实践角度分析了下阶段机器学习技术的研究方向。

关键词: 机器学习; 智能配用电; 方法分析; 工程实践

0 引言

能源转型是解决环境问题、实现经济社会可持续发展的必由之路。美国政府宣布将在2050年之前实现100%的清洁能源供应与零排放^[1-2]。中国提出了在2030年前实现“碳达峰”、2060年前实现“碳中和”的双碳战略目标。摆脱对传统化石燃料的依赖,以风能、太阳能等可再生能源为主导,实现清洁、可持续的能源供用模式已经成为各国的共识。

在能源转型背景下,电力系统作为整个能源链中的枢纽以及能源变革的核心环节,其面临的挑战日趋严峻。一方面,风电、光伏等新能源占比快速提升。根据预测,到2060年中国风电、光伏装机容量在电源结构中的比重将达到60%,系统的弱惯量、强随机性、电力电子化特征愈发显著。另一方面,电动汽车、可控负荷、储能等分布式发电资源也将广泛接入电网,系统的运行方式更加灵活,控制的复杂度也随之大幅提升。

机器学习技术是应对以上挑战的强大工具。机器学习是计算科学的一个分支,旨在让计算机通过对已有经验的学习不断地自我提升,并最终实现对人类智慧的模拟^[3-4]。机器学习技术的研究兴起于20世纪50年代,贝叶斯分类器、逻辑回归、感知器模

型等经典的机器学习模型被首次提出,解释大脑神经元工作原理的赫布理论(Hebbian theory)则对后续神经网络模型的发展起到了启蒙作用。此后,机器学习技术依次经历了70年代的寒冬、80年代的再次兴起,并在90年代以支持向量机为代表走向成熟和应用^[5]。2012年,Hinton等人采用深度卷积神经网络在图像分类问题上取得了巨大成功^[6],标志着机器学习技术开始进入深度学习时代,生成对抗网络^[7]、强化学习(RL)^[8]、长短期记忆网络^[9]等模型在数据生成、自然语言处理(natural language processing)、实时决策等多个领域取得了显著的成就。按照学习方式的不同,机器学习技术可分为无监督学习、半监督学习、监督学习以及强化学习四大类,其在电力系统中的应用已经得到广泛的关注。应用场景可分为态势感知、决策支持两大类,包括但不限于电源与负荷预测、系统建模、状态监测、故障检修、系统优化运行等多个方面^[10-11]。虽然已经有大量的研究论文发表,但机器学习技术在电力系统中的实践项目仍旧较为有限,工程实践与理论研究之间存在一定的脱节。这主要是因为机器学习技术的“黑箱”特点,使得其在可解释性、鲁棒性等方面弱于传统物理建模的方式,与电力系统运行中极高的可靠性要求存在一定矛盾。

本文以工程应用为切入点,从源、网、荷3个方面梳理了北美地区机器学习技术在智能配用电领域

收稿日期: 2021-03-20; 修回日期: 2021-06-08。

上网日期: 2021-07-07。

的典型工程实践,旨在介绍机器学习技术在配用电领域的实际应用情况,思考机器学习技术的适用性,并为下一阶段的研究提供借鉴。

1 电源侧

在电源侧,当前工程实践的关注重点在于利用机器学习技术提升对于风电、光伏、储能等分布式电源的态势感知能力,包括状态监测、预测等,降低间歇性电源出力的不确定性。所采用的机器学习算法以分类、回归等监督学习算法为主。

1.1 Watt-Sun 光伏预测

2014年,美国IBM公司得到美国能源部的资助,开展名为Watt-Sun的光伏预测项目,旨在大幅提升光伏预测的准确率^[12]。

1) Watt-Sun 预测方法

相比于传统预测方法,该项目在预测方法上具有以下2个亮点:一是多种数据来源的综合利用与挖掘,包含历史出力数据、历史天气数据、实时量测天气数据、各类光伏电站传感器数据,以及卫星云图、全天空成像仪等图像数据,通过机器学习技术,构建了多源异构数据与光伏出力之间的复杂映射关系,提升了数据的利用效率与预测精度;二是项目涵盖了多种已有的光伏预测模型,并建立了基于深度学习技术的模型选择系统,以决定在何种情况下应更加信赖哪个预测模型。基于该模型选择系统,不同的预测场景可以得到更加适配的预测模型,有助于提升整体的预测精度。

2) Watt-Sun 应用效果

该项目成果在加州电力系统运营商(CAISO)以及佛蒙特州、加利福尼亚州、亚利桑那州等多个公用事业公司共计数十个装设有全天空成像仪的光伏电站得到应用。经过一年的试运行,该系统在未来24 h时间尺度的短期光伏预测方面相较于原有预测模型实现了35%的精度提升。

3) 光伏预测技术分析评价

光伏出力直接受到云层的影响,而云层复杂的生消、运动规律,是导致光伏短期出力波动与预测困难的主要原因。该项目引入卫星云图、全天空成像仪等图像数据以及各类传感器的实时量测数据,采用机器学习技术从中提取有效信息,提升了对云层运行状态的感知能力,从而提升了短期光伏出力的预测精度。对这类多源异构数据的挖掘也是当前光伏预测领域的研究热点之一^[13]。因为不存在一个通用的预测模型能够解决所有光伏预测问题,项目

储备了多个备选预测模型,并通过深度学习技术为不同预测场景匹配最优的预测模型。相比于传统的专家系统,基于机器学习的模型选择系统具有更强的适应性、可扩展性。这种根据问题的不同定制化地选择解决方案的机器学习技术隶属于“元学习(meta learning)^[14]”或“自动化机器学习(AutoML)^[15]”,在配用电领域具有广泛的应用前景。

1.2 风电机组实时状态监测

Turbit公司基于机器学习技术,在2020年开发了风电机组的实时监测与控制系统,实现精准的故障诊断与发电效率的提升^[16]。

1) 风机状态监测方法

项目采用监督式学习技术,输入为数据采集与监控(SCADA)系统量测数据以及温度、风速、湍流强度等天气数据,输出为经过标记的故障信息。训练后的深度神经网络模型可根据实时的量测信息有效辨识出微小的故障,并推荐相应的检修措施。此外,项目构建了机器学习与风电机组控制算法之间的闭环,将机器学习算法的分析/预测结果作为控制算法的输入,以提升控制效果。例如当风电场的一侧检测到阵风时,可根据风向、风速信息快速调节受影响区域的风电机组姿态,以最大化利用风能。

2) 风机状态监测应用效果

该技术成功应用于科罗拉多州的Cedar Creek风电场,发电效率相比应用前提升了5%。

3) 项目难点分析

该项目的挑战之一是大规模量测数据的实时分析与处理。对于每个风电机组而言,SCADA将产生输出功率、叶片角度、齿轮组加速度等500余个量测值。对于这些量测结果,传统行业惯例是每10 min进行数据上传。然而为实现实时的监测与控制,需要将数据上传频率提升至秒级,并要求机器学习算法在几秒内完成分析与计算,这将大幅提升数据处理的难度。Turbit公司采用开源的数据库系统postgresql进行数据管理,并配合Swarm64 DA数据加速器实现数据的压缩、并行处理、接口缩减等,实现秒级以内的处理速度。这种高效的数据管理架构可为同类型项目提供借鉴。

1.3 电池储能管理系统 Autobidder

特斯拉公司融合机器学习、优化算法等技术在2020年开发了电池储能管理软件平台Autobidder^[17],可在市场环境下根据不同用户的需求以及风险偏好优化电池组运行,以最大化用户利益。

1)Autobidder技术特点

在 Autobidder 中,机器学习技术的应用场景主要包括市场价格、负荷、发电能力等关键信息的预测,为电池组的优化调度与智能投标提供支撑。此外,该平台对于不同的用户偏好和应用环境(例如不同的市场环境)具有自适应性,并且可以不断根据新的运行数据对算法进行持续优化。为保证算法执行的时效性与可靠性,平台架设在特斯拉公司的云服务器上,并通过各类应用程序与市场运行人员、系统维护人员以及用户进行交互。

2)Autobidder应用效果

该系统目前在全球范围内超过 1.2 GW 的电池组上成功运行,其中最大的单体项目为南澳大利亚的 Hornsdale 储能项目,规模达到 150 MW。该项目在投运不到 2 年的时间内,通过提供应急调频等服务创造了约 5 000 万美元的收益,同时提升了电网的运行稳定性。

3)Autobidder优势分析

分布式储能资源是重要的系统调节资源,在其开发利用过程中应尊重用户的个性化用能需求,保证用户的利益。Autobidder 项目的一大亮点在于可以适应不同类型用户的用能偏好,并在此基础上自动化地为用户选择最优的电池组盈利方案,综合考虑能量市场、辅助服务市场、系统调节服务、双边合同、针对可再生能源的调节服务等多种盈利场景。特斯拉公司的云计算架构为 Autobidder 的实现提供了硬件支撑。

1.4 电池寿命快速预测

电池的寿命(即有效的充放电循环次数)是电池的核心指标之一。标准的电池寿命测试方法是对电池进行持续的充放电,直至电池衰竭。该过程往往会持续几个月甚至一年时间,导致很高的时间成本。针对该问题,美国麻省理工学院、斯坦福大学以及丰田研究中心在 2019 年联合开展了基于机器学习的电池寿命快速预测项目,旨在通过学习电池早期的充放电特征,实现对电池寿命的准确分类与预测^[18]。

1)电池寿命快速预测方法

项目研究人员建立了 124 个商业电池组的充放电测试数据库。由于产品质量、充放电策略的差异,这些电池的寿命在 150 至 2 300 次循环之间不等(寿命定义为电池容量损耗达到 20% 所需要的充放电循环次数)。针对以上原始数据,研究人员建立了描述充放电过程的特征集,并采用基于弹性网络的线

性模型开展预测工作。

2)电池寿命预测精度

研究人员将 124 个电池组数据划分为训练集(41 个)、验证集(43 个)、测试集(40 个)3 个部分。测试集上的结果表明,仅基于电池前 5 次充放电数据,模型即可实现对电池寿命长/短的有效分类,分类精度达到 95.1%;基于前 100 次充放电数据,可实现对电池寿命的预测,平均预测误差 9.1%,大幅减少了电池寿命的测试时间。

3)经典预测模型价值分析

该项目中所采用的预测模型为简单的线性模型,其主要原因有 2 个:一是电池早期充放电数据较为有限,建立预测模型本质上属于小样本机器学习问题,采用过于复杂的预测模型容易导致过拟合;更重要的是,线性模型具有较好的可解释性,可根据拟合结果反推影响电池寿命的关键因素,从而有助于理解电池的充放电机理。例如在该项目中,研究人员基于拟合结果发现最优的电池寿命预测特征是第 10 次放电循环和第 100 次放电循环之间的电压差。仅基于该单一特征即可实现 85% 以上的预测精度,且该特征与电池寿命末期的容量快速衰减特性高度相关。在配用电领域的工程实践中,与物理机理的分析相结合,这类简单、经典的机器学习模型仍然具有较高实用价值。

1.5 光伏电站无人机自动巡检

无人机巡检技术已在光伏电站广泛应用。对于无人机采集到的图像信息,目前仍主要由人工进行分析和处理。光伏电站的无人机自动巡检技术^[19]在自动巡航收集光伏板图像信息的基础上实现对故障的自动识别,进一步降低了人工成本。

1)无人机自动巡检核心算法

该项技术的核心算法为基于快速区域卷积神经网络(fast region-based convolutional network, Fast R-CNN)的目标检测算法。采集了 300 个标记了缺陷信息的光伏板图像,通过图像旋转实现样本的扩充并在 Xeon 处理器上对模型进行训练,训练过程耗时约 6 h。

2)无人机自动巡检应用效果

训练完成后,每秒钟可辨识 188 张图片,辨识光伏板是否存在缺陷的成功率达到 96.3%,效果如图 1 所示^[19]。基于该方法,可将单台无人机的最大巡检能力由每周 1 次提升至每周 3 次。

3)无人机自动巡检技术分析评价

该项目是图像识别技术在电力领域的早期成功

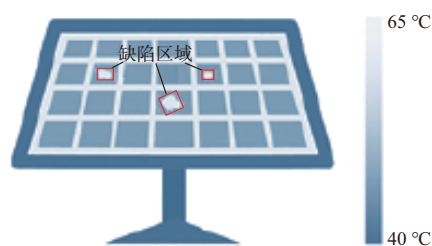


图1 光伏板缺陷辨识结果

Fig. 1 Defect identification results for photovoltaic panel

实践。由于不涉及电力系统的具体模型,其算法的跨领域通用性较强,可直接移植计算机视觉领域的最新算法。最终效果主要取决于算法的有效性以及硬件平台的计算能力。此外,由于故障辨识本质上是监督学习,而故障样本往往数量有限,且故障数据集的标记将耗费较高人力成本。因此,可考虑采用数据增强(data augmentation)方法实现训练样本的扩充。

1.6 Deep-solar 光伏数据库

斯坦福大学在2018年开展了Deep-solar光伏数据库项目,通过对卫星地图的图像识别,辨识出其中的光伏区域,从而以极低的成本实现对大范围光伏安装情况的快速统计^[20]。

1) Deep-solar 核心算法

项目的技术核心为谷歌深度卷积网络Inception V3^[21-22],主要分为图像识别和图像分割2个部分:先辨识出目标图像中是否包含光伏系统,一旦判定为真,则继续进行图像的分割,从中提取出光伏区域,从而实现对其面积(可折算为容量)的统计。

2) Deep-solar 应用效果

项目示例效果详见附录A图A1^[20],其对于光伏系统的辨识准确率可达到90%。基于该数据库,研究人员可结合经济社会信息得到更多结论。例如研究人员发现日均辐照度大于4.5(kW·h)/m²是触发光伏安装的“阈值”;此外,在人口密度为1000人/平方英里(1平方英里约等于2.59 km²),家庭年均收入约15万美元的地区光伏安装密度达到峰值。

3) 基于卫星地图的智能配用电研究

分布式光伏近年来增长迅速,但由于其单个体量小、数量庞大、布局分散,对其装机容量、安装位置等静态信息的统计工作难度较大。而基于Deep-solar所得到的光伏数据库具有很强的实时性,其更新频率可与卫星地图的更新频率一致,对于分析光伏的分布情况、装机密度等信息具有直接帮助,同时

也可为电力系统的规划运行提供参考。此外,卫星地图由于包含了丰富的静态信息,受到了电力系统研究人员的更多关注。基于卫星地图与机器学习结合的配电网规划、模型辨识已有相应的研究成果^[23]。

1.7 美国能源部资助的AI+光伏研究项目

美国能源部下属的太阳能技术办公室(solar energy technology office, SETO)发布了2020年获得资助的研究项目,其中包含10个人工智能技术在太阳能领域的应用,资助总额为730万美元,代表了该方向上美国最新的研究趋势^[24],具体如表1所示。其中采用的核心技术方法均为人工智能与机器学习技术。

通过以上获得资助的项目内容可以看出,当前人工智能在光伏领域的应用关注的核心在于解决光伏出力强波动性与弱量测之间的矛盾及其带来的含光伏系统运行困难问题。为解决该问题,一方面项目充分利用已有的多源异质数据,包括但不限于天气数据、出力数据、传感器量测数据、图像数据等;另一方面,项目积极运用前沿的人工智能技术,例如图神经网络、迁移学习、变分递归网络等,对这些数据进行深入挖掘,并最终提升对含光伏系统的实时感知与预测能力。此外还可以看到,表后光伏及其带来的净负荷分析预测问题成为研究的重点,光热电站的相关运行问题也逐渐受到关注。

2 系统侧

目前配电网的运行方式日趋灵活,运行状态可能在短时间内发生显著变化,需要更加精准、快速的态势感知以及决策支持技术以保障其运行的安全性与经济性。机器学习技术一方面可助力解决配电网的弱量测问题,提升对系统运行的状态感知、故障诊断能力;另一方面基于强化学习可实现更加快速、自适应的运行决策。

2.1 故障风险评估与预防性维护

提前预知配电系统中可能发生故障的设备并对其进行维护或更换,可有效减少故障所带来的损失。美国纽约地区的地下供电系统是世界上最大、最古老的地下供电系统之一,对其进行检修与维护难度较大。2005年起,当地公用事业公司Con Edison与哥伦比亚大学开展合作,尝试利用机器学习技术对系统运行数据进行实时分析,按照故障风险的大小对配电设施进行排序,并对高故障风险的设备进行重点检修^[25-28]。

表1 2020年人工智能在太阳能领域资助项目
Table 1 Funded projects of artificial intelligence in the field of solar energy, 2020

项目主要研究内容	资助单位
光伏电站的异常识别、故障预测以及预防性维护优化	美国亚利桑那州立大学
融合相量量测装置等多种量测数据,提升系统状态估计的准确性,进而促进新能源的消纳	美国亚利桑那州立大学
融合多种量测数据,实现实时的系统状态分析与可视化,提供配电网的精准实时态势感知	美国 Campus Energy 公司
基于图神经网络,提取多个相邻分布式光伏系统之间的时空相关性,以解决数据质量、短期预测以及中长期的评估等问题	美国凯斯西储大学
在大规模表后(behind-the-meter)光伏接入环境下,实现表后光伏的辨识与日前净负荷的概率性预测,其中迁移学习(transfer learning)将被用于解决量测数据缺乏地区的建模与分析问题	美国国家可再生能源实验室(NREL)
在大规模表后光伏/储能系统接入的环境下,实现日前净负荷的预测以及需求响应潜力的评估	美国北卡罗来纳州立大学
通过对大量异质数据的分析,检测出拓扑连接的变化以及故障的产生,并动态更新系统的模型,将采用稀疏估计(sparse estimation)技术	美国东北大学
采用变分递归网络(variational recurrent network),实现高比例光伏接入下的净负荷预测,将综合考虑天气、光照、地理信息、负荷数据,给预测结果一个置信边界	美国太平洋西北国家实验室(PNNL)
基于卫星图像与谷歌街景图,采用图像识别技术辨识其中的光伏系统,建立大范围的高精度光伏数据地图	美国斯坦福大学
综合利用光照、云层以及现场量测数据实现集中式光热电站的建模,为运行决策提供支持,尤其是在多云等光照快速变化的天气下,减少运行的不确定性,提升发电效率	美国威斯康星大学麦迪逊分校

1)设备故障风险排序算法

项目主要包含2个部分:馈线故障排序和电缆井故障排序。馈线故障排序所使用的数据源包括静态数据(如馈线的年龄、拓扑结构等)和动态数据(如负荷序列等)共计150余个属性,并将历史故障数据作为标签。项目采用的机器学习算法为马丁格尔增强(Martingale boosting, MB)算法^[29],可不断迭代筛选出对故障率影响最为显著的属性,具有较好的可解释性。电缆井故障排序所使用的原始数据为自1970年起的电缆井故障记录票(trouble ticket)。由于历史时间长,故障记录票的质量参差不齐,长度从几行到几百行不等,且包含大量的非常规缩写、单词拼写错误、格式变化等,成为该项目的难点之一。为此,研究人员首先通过数据清洗与关键词检索,提取出严重的故障事件(爆炸、起火等)及其对应的特征信息。此后,采用延伸知识发现(extended knowledge discovery, EKD)算法进行监督学习,实现对每个电缆井的风险打分,并按照打分结果排序。

2)故障预测效果

馈线故障预测方面,Con Edison公司利用项目算法对辖区内941个馈线进行了排序。以2005年7月为例,该月共计发生184起馈线故障,算法的排序成功指数达到0.8(1表示完全成功),且排名后500的馈线无一发生故障,表明了该方法的有效性。电缆井故障预测方面,2007年Con Edison公司共计拥有超过5万个电缆井,全年共计发生44起严重事故。该算法风险评分最高的2%的电缆井发生故障的数

量占总故障的11%,表明其具备一定的预测能力。此外,通过特征分析发现,电缆井中电缆的数量和故障记录票中提及的电缆井次数这2个特征对故障的发生具有最强的预测能力。

3)故障排序技术分析评价

该项目开展时间较早,所采用的机器学习算法为经典的排序算法,具有一定的预知能力,且结果具有一定的可解释性,是早期的成功实践之一。对于故障记录票的处理,现阶段可考虑采用自然语言处理技术进行自动化的关键词提取与故障分类。

2.2 系统模型辨识

许多公用事业公司具有悠久的历史,网架结构庞大,在系统建设初期还没有地理信息系统(GIS)、计算机辅助设计(CAD)等软件,甚至还没有计算机。这就导致系统的网架数据源自许多不同的历史层次,包括初级的手绘纸质材料、文件扫描件、以及现代的系统建模工具。在这些异质网架数据融合的过程中,很容易出现各种错误,例如量测不准确、网架结构重叠、数据缺失、设备更新等。若通过现场调研的方式修正网架数据将带来很高的成本。为此,2017年美国通用电气公司开展Mix Core Platform项目的研究,旨在通过机器学习技术恢复准确的中低压配电网结构^[30]。

1)系统拓扑辨识方法

项目通过机器学习技术,从GIS、SCADA、数据库管理系统(DMS)、停电管理系统(OMS)、智能库房管理系统(WMS)等多种不同的数据源中进行特

征提取与关系型建模,建立了配电网设备的语义数据模型(semantic data model)。在此基础上,采用迭代遍历算法推断出设备之间的拓扑连接关系,并最终输出配电网拓扑模型如图2所示^[30]。其中,不同颜色曲线代表不同相的网络拓扑。对于每一个连接关系的推断,系统可同时给出推断的置信度水平。



图2 Mix Core Platform 示意图
Fig. 2 Schematic diagram of Mix Core Platform

2) 系统拓扑辨识效果

该项目成果已经在欧洲、北美等多个地区得到应用。其中代表性的应用案例是将1149个智能表计成功匹配到32个馈线上,匹配成功率超过90%。这32个馈线也成功与其所属的变电站匹配,匹配成功率为100%,且推断的置信度均在90%以上。在这个过程中,系统识别出了原始数据中未标注的地下电缆与节点,修正了部分智能表计与变压器之间的连接关系。

3) 拓扑辨识技术特点分析

系统拓扑模型辨识是典型的“数据驱动-物理模型”相结合的应用场景,要求机器学习的推断结果可解释、可验证。在这个过程中,语义模型是一种值得关注的管理数据与建模工具,用于描述复杂系统中各个组件的属性及其相互之间的连接关系。

2.3 微电网优化运行

2.3.1 BluWave-AI & Sustainable Power Systems

BluWave-AI是致力于采用人工智能技术解决清洁能源问题的公司,在2018年与Sustainable Power Systems联合研制了高比例清洁能源接入下的微电网运行控制模块^[31]。该模块采用人工智能技术实现系统运行状态的实时分析以及负荷、新能源的预测,从而支撑微电网的自主运行(类似于电动汽车的自动驾驶)。该项目的亮点之一是采用“云-边”的系统架构,机器学习算法集中在云平台上运行,而智能表计、恒温控制器、各类物联网传感器可进行数据的采集、上传以及实现部分本地计算功能。

2.3.2 Emera & Sandia

Emera公司与Sandia国家实验室于2020年联合发布了微电网智能控制系统BlockEnergy^[32],整

个系统的部署架构如图3所示。

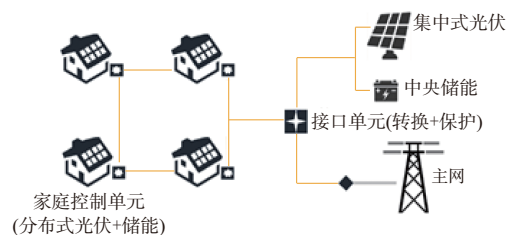


图3 BlockEnergy 系统部署架构
Fig. 3 Deployment architecture of BlockEnergy system

1) BlockEnergy 技术特点

系统除了包含传统的微电网运行控制功能之外,还可实现区域内各个用户之间的能源拍卖交易,并可与电力公司完美对接。而且当中央控制单元发生故障后,其控制功能可以自动转移至区内用户的控制子模块上,使得系统具有更强的韧性。在该系统中,机器学习技术主要用于提供用户侧负荷、分布式光伏等数据的预测,从而为求解优化问题提供支撑。

2) BlockEnergy 应用效果

该产品在美国最大的空军基地之一Kirtland开展了测试,测试系统包含250 kW的负荷、100 kW的屋顶光伏与220 kW·h的储能。测试结果表明,系统在发生故障、突然失去电网供电、所有用户同时开启大功率电器的情境下,仍然能够保持平稳运行。

2.3.3 Holy Cross Energy & NREL

2020年,科罗拉多公用事业公司Holy Cross Energy与NREL合作,开展自治微电网项目的研究^[33]。

1) 自治微电网核心算法

在该项目中,为了应对优化问题规模增长导致的求解时间过长的的问题,研究人员设计了交替方向乘子法(ADMM)与强化学习相结合的分布式优化问题求解架构ADMM-RL,其亮点在于对每个子优化问题采用强化学习进行决策。由于训练完成的强化学习智能体(agent)可以实现快速的决策,大幅减少了优化问题的整体求解时间,可实现秒级以内的实时优化控制。

2) 自治微电网应用效果

项目成果应用于科罗拉多州附近一个包含27个家庭的微电网,每户家庭安装了8 kW的屋顶光伏以及磷酸铁锂电池储能系统。结果表明系统可保持长期稳定运行,且在冬季主网断电、光伏板覆雪的极端情形下仍可维持供电约2 h。此外,微电网内用户的电费相较于本州普通居民降低了85%。

以上工程实践表明,人工智能助力的微电网运行控制技术已经开始进入商业化、产品化阶段。由于微电网可脱网运行,对于停电事故具有更强的韧性。在微电网的优化运行中,机器学习技术的作用主要体现在2个方面:①态势感知,例如负荷、光伏等数据的分析与预测;②计算提速,当微电网内设备数量过多时,可采用强化学习等人工智能技术代替传统的优化问题完成部分决策任务,从而避免维数灾,提升计算效率,满足实时控制的时间要求。

2.4 系统无人机运检与山火预防

加州是山火高风险地区,而电力设备故障是导致山火发生的重要因素。2018年的加州大火即是由输电线路设施老化所导致^[34]。南加州电力公司下辖的区域内有27%的区域是山火高风险地区。为避免山火带来的损失,公司开展了无人机运检与山火预防项目^[35]。

项目通过无人机巡线拍摄电力设备图像,并利用计算机视觉技术自动识别设施中的早期损坏特征与火灾风险,以便检修人员及时维修与更换。进一步地,公司基于所拍摄的线路周边植被图像,辨识出具有造成短路风险的植被信息并对其进行及时处置,可有效降低相应的短路风险与火灾风险。一旦发现起火点,公司将开展火势传播规律的仿真工作,以更好地遏制山火的蔓延。应用结果表明,该技术可将单个设备的平均检修时间由80 min减少至52 min。

2.5 系统故障快速修复

系统发生严重故障后的修复过程,关键在于合理地分配人力、物力等检修资源。该项工作需要考虑诸多因素,例如修复时间要求、成本、设备损坏程度、可用检修资源约束等。此外,随着检修过程的开展,这些因素也在发生动态变化。目前的检修决策仍然主要依靠专家对故障数据的实时分析,该过程往往要持续1 h以上。而当同时发生多个故障时,问题的决策将更加复杂。为此,PG&E公司利用机器学习技术加速故障修复过程^[36]。

1)故障快速恢复方法

项目分为预测、决策、呈现3个阶段。在预测阶段,机器学习引擎读入飓风等灾害预报数据、GIS数据、系统量测数据,预测其所造成的故障类型及所需要的检修资源;决策阶段,系统基于随机优化模型生成最优恢复策略;呈现阶段,系统提取故障状态、可用检修资源等关键信息集中呈献给运行人员。在该项目中,故障事件被建模为多元高斯分布,检修资源的动态变化则由马尔可夫过程描述。机器学习技术

主要用于基于历史数据实现对高斯分布以及状态转移矩阵中参数的估计。

2)故障快速恢复测试效果

该系统在2014至2017年的历史故障数据集上进行了测试,将系统生成的检修方案与实际执行的检修方案进行对比,并由专家对方案进行评估。结果表明系统可在5 min内生成有效的检修方案,相比于原来平均1 h的人工决策时间显著提升了工作效率。

3)故障恢复与机器学习技术

不合理的检修方案将导致故障持续时间的延长以及检修资源的浪费。因此,为避免高昂的试错成本,目前在故障恢复领域机器学习技术主要用于为运行人员提供决策支持,而非替代运行人员进行决策。

3 用户侧

由于智能表计的普及,电力用户侧具有大数据的特征,为机器学习技术的应用提供了良好的环境^[37]。机器学习技术主要用于负荷变化规律的分析与挖掘(建模、预测、模式识别等),为进一步开展负荷管理提供支撑,最终实现用电能效与供电质量的提升。

3.1 全系统负荷预测

为应对多样化的配电网运行要求,PG&E公司在2018年开展了全系统负荷预测项目的研究^[38]。

1)全系统负荷预测方法

项目的技术亮点主要包含2个方面。一是对于数据缺失区域的处理。相较于主网,配电网具有“弱量测”的特征。对于配电网中缺乏高级量测体系(AMI)、SCADA等量测数据的区域,项目基于迁移学习^[39]的思想,在与数据缺失区域特征相似的量测区域内完成模型的训练,并适当调整后应用于数据缺失区域的预测。二是预测模型的混合技术(blending)。项目采用自上而下和自下而上2种预测模型,并对二者结果进行加权得到最终结果。其中的权重由决策树模型确定,将综合考虑数据的可用性等特征。此外,项目可以对预测结果给出置信度,对于可能出现的大预测误差进行预警。

2)全系统负荷预测效果

项目成果在其管辖范围内的317个节点(区域)进行测试。结果表明该项目可实现配电网内任何负荷聚合级别的预测,并具有85%的成功率标识出误差大于50%的极端情景。

3) 复杂预测场景应对方法分析

配电网日益灵活的运行方式,要求适配同样具有灵活性的负荷预测系统,以实现不同区域、负荷聚合水平、负荷类型、时间尺度的预测任务。在该过程中,迁移学习可有效解决部分预测任务数据缺失的问题,并提升模型的训练速度。此外,包含多种预测模型的集成学习(ensemble learning)可以综合不同模型的优点,提升预测系统对于不同任务的应对能力。

3.2 数据中心能耗优化

数据中心的冷却将消耗大量的电能。2016年,Google公司DeepMind团队建立了基于深度学习的数据中心能耗优化模型,以提升能源利用效率^[40]。

1) 数据中心建模方法

数据输入包括现场上千个传感器所收集到的温度、功耗、空气泵速率、空调系统温度设定等数据。系统采用一个深度神经网络模型进行运行决策,以最小化数据中心能效系数(PUE)为目标,PUE越接近1表明数据中心能效越高。另外,系统包含2个深度网络,对未来1h的温度、压力等指标进行预测,以保证决策满足系统的运行约束。

2) 模型应用效果

该模型在实际应用后,可稳定地实现40%左右的冷却负荷减少以及15%左右的能效系数下降,如图4所示^[40]。

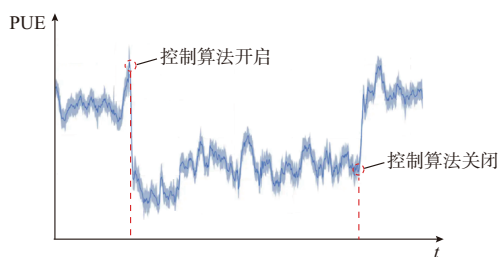


图4 控制算法开启前后PUE对比
Fig. 4 Comparison of power usage effectiveness before and after applying control algorithm

3) 机器学习用于负荷建模的适用性分析

冷却系统的运行优化难度较高,主要包括以下3个原因:①设备与环境之间的热交换过程复杂、非线性,物理建模方法难以对其进行完全描述;②系统具有较强的热惯性,难以适应快速的内部、外部环境变化;③不同的数据中心具有不同的架构和环境特点,导致针对每个数据中心需要进行个性化的建模和维护,工作量较大。而这种复杂非线性、动态、个性化的建模任务,正是机器学习技术的优势所在。

可采用神经网络模型替代物理模型,并不断采用最新的量测结果对模型进行动态更新。

3.3 客户反馈分析

用户的反馈(如用户的评论留言等)是提升公司服务和产品质量的重要信息。这些反馈可能包含语音、文字评论等多种形式,并来自不同的语言,对其进行翻译、整理和理解将耗费大量的人力资源。针对该问题,ABB与IBM合作,采用IBM的Watson Discovery引擎对用户反馈进行自动化整理和分析^[41]。

1) 客户反馈信息处理方法

引擎首先采用自然语言处理将不同的语言评论统一翻译成英文。然后,采用自然语言理解(natural language understanding)从中提取出关键词、概念、实体对象、类别、关系等信息,判断用户的情绪和语气,并构建相应的结构化数据库。最后,针对用户的负面反馈,可通过该数据库自动查询生产环节中可能导致问题的原因,并提供给工作人员进行进一步的分析和改进。

2) 项目应用价值

相比于其他直接作用于电力生产环节的项目,该项目从用户体验触发,通过分析用户的负面反馈反向查找生产环节中的不足,对于配电网运行人员提升供电质量与服务水平具有参考意义。

3.4 窃电监测

窃电是发展中国家所普遍面临的问题。以牙买加为例,2014年有27%的电能损耗是由于窃电等非技术原因所导致。这部分电能损失成本将部分由公用事业公司承担,并最终分摊到用户身上。为解决该问题,牙买加公用事业公司JPS每年花费上千万美元处理窃电问题,但始终无法彻底解决。这是因为其检测方式主要依赖于人工分析,而窃电手段正变得越来越复杂并难以识别。

1) 窃电监测方法

JPS与世界银行下属的能源管理援助计划(energy sector management assistance program, ESMAP)合作,采用机器学习技术实现自动窃电监测^[42]。项目基于用户的历史用电数据、窃电行为审查记录以及所提取的判别特征,采用随机森林模型为每个用户账户进行风险打分,并对高窃电风险用户进行重点检查。

2) 窃电监测效果

该项目执行的第1个月内,检测出窃电行为的

成功率提升了一倍。项目代码已在 Github 向社会公开^[43]。

3.5 居民空调负荷需求响应

在需求响应期间能否实现稳定的负荷削减是衡量需求响应质量的关键,也是需求响应作为产品参与电力市场的前提。对于居民空调负荷而言,传统的需求响应控制方式包括2种。一种方式是周期性地开断空调设备来实现负荷的削减。对于某些用户而言,这种方式会明显损害用户的舒适度,这是因为不同用户的房屋大小、隔热能力各不相同。同样断开0.5 h空调设备,对于部分房屋热容小的用户将导致其室内温度严重越限。另一种方式是直接调节用户的空调温度设定值,并使其保持在一定舒适范围内。该方法虽然可以保证用户的舒适性,但却无法保证负荷削减的稳定性,往往在需求响应初期能实现较好的负荷削减,而在后期由于室内温度越限,空调重新启动导致负荷反弹。

1) 空调用户个性化建模方法

为解决该问题,得州电力可靠性委员会(ERCOT)市场成员采用机器学习技术,通过每个用户的空调使用数据,学习用户的房屋热力学模型以及用户的用能偏好^[44]。在此基础上开展随机优化,以最小化负荷削减量波动为目标实现对空调负荷集群的优化控制。

2) 需求响应执行效果

基于300多个家庭的测试结果显示,相比于传统固定用户空调温度设定的需求响应方式,该方法可降低20%的需求响应控制时间以及25%的用户需求响应退出率。

3) 用户个性化建模价值分析

该项目中,机器学习技术的作用是实现用户对用户用能偏好以及房屋热力学模型的个性化估计,从而避免了传统“一刀切”的控制方式,对每个用户进行个性化调节,最大程度发挥需求响应潜力。例如对于部分房屋热容较大、温度不敏感的用户,可适当降低对其空调设备的控制时间,在不影响需求响应效果的同时提升了用户用能的自由度与舒适度。

4 北美工程实践经验总结

4.1 机器学习特性总结分析

机器学习是一门以计算机为载体、数据驱动的学科,其学习过程本质上是一个由果索因的过程,即从纷繁复杂的数据中提取一般性的规律,从而加深对研究对象的理解,并进一步反作用于研究对象。

在电力系统中,机器学习技术发挥作用的机理主要体现在以下2个方面。

1) 描述复杂规律。电力系统是高度复杂、非线性的系统,其部分运行机理至今尚未得到完全理解,更加难以建立完备的解析化模型进行描述(如3.2节的数据中心建模)。此外,系统产生的海量量测数据也已经超出人类的理解能力(如1.2节的风电机组量测数据)。机器学习技术可以从数据中提取复杂的映射规律,从而帮助运行人员掌握系统的运行状态,理解系统的运行机理。

2) 提升工作效率。一是机器学习模型以计算机算力为支撑,自动化程度高,可以进行大规模部署以替代重复性的人力劳动(如1.5节的光伏电站自动巡检);二是适应能力强,可利用最新的量测数据实现自我更新(如1.1节中光伏预测模型的迭代更新),相比于物理模型维护更加便捷;三是训练完成的机器学习模型在应用过程中,只需简单的数值计算即可得到输出结果,相比于优化等模型具有更高的计算效率,更加适合实时在线应用(如2.3节中基于强化学习的微电网运行决策)。

对系统复杂规律的理解与描述属于电力系统态势感知的范畴;而在理解系统的基础上进一步实现系统的高效优化控制,则属于决策支持的范畴。为此,对第1至3章的工程实践按照态势感知、决策支持归纳为两大类共计5个应用场景,如表2所示。

4.2 态势感知类

电力系统态势感知即在特定的时间和空间下,对电力系统中各元素或对象的察觉、理解,以及对未来状态的预测^[45]。基于机器学习的系统状态监测、预测均可以归类于该范畴。在态势感知方面,机器学习实践经验总结如下。

1) 预测、监测是典型的基于数据驱动的应用,其本质是从大量量测数据中提取出一般性的规律,与电力系统物理模型结合程度较低。因此,其算法的跨领域通用性较强,可直接借鉴于信号处理、计算机视觉、时间序列分析、模式识别等领域的技术与算法,并较为容易地在工程中得到应用。这也是该领域论文发表数量多、软件迭代更新速度快的主要原因。

2) 预测、监测的数据驱动特征使得其对系统量测数据的规模、质量有着较高要求,而这与当前配电网有限量测的状态存在一定矛盾,单一数据源往往无法完全反映出系统的运行状态。因此,在工程应用中应充分挖掘多源异质数据,从不同视角共同提

表2 工程实践案例汇总
Table 2 Summary of engineering practice cases

所属类别	应用场景	典型工程实践案例	数据概况		核心机器学习算法	机器学习应用模式	应用难度	下阶段工程实践要点
			输入	输出				
态势感知	预测	光伏预测 (Watt-Sun, 2014)	天气、传感器、卫星云图、全天空成像仪	光伏出力预测	多模型优选	回归(监督学习)	易(数据驱动、直接应用)	提升数据管理、模型训练的速度,多元异质数据的挖掘,信息安全与异常数据辨识,与运行软件的高效衔接以提升控制效果
		全系统负荷预测 (PG&E, 2018)	历史负荷数据、预测环境特征集	负荷预测	多模型集成学习			
		电池寿命快速预测 (MIT&斯坦福&丰田, 2019)	电池充放电特征集	电池寿命预测	弹性网络			
	监测	风机实时状态监测 (Turbit, 2020)	SCADA 量测、天气	故障辨识结果	深度神经网络	分类、模式识别(监督/非监督学习)	较难(数据-物理模型结合)	
		光伏电站无人机自动巡检 (Intel, 2017)	光伏板图像	故障辨识结果	Fast R-CNN			
		系统无人机巡检与山火预防(Southern California Edison, 2020)	无人机拍摄设备、植被图像	故障,高风险植被辨识结果	卷积神经网络			
		窃电监测(ESMAP & JPS, 2017)	历史用电数据、窃电记录、判别特征	窃电风险评分	随机森林			
	建模	系统模型辨识 (GE digital, 2017)	不同格式的网架数据	语义数据模型	未披露	回归、分类(监督学习)	较难(数据-物理模型结合)	可解释机器学习、用户侧个性化建模、明确模型的有效边界
		居民空调负荷需求响应 (ERCOT, 2018)	用户用能历史数据、天气数据	房屋热力学模型	未披露			
		数据中心能耗优化 (Google, 2016)	温度、功耗、空气泵速率等传感器量测	冷却系统运行策略	深度神经网络			
		Deep-solar 光伏数据库 (斯坦福大学, 2018)	Google 地图	光伏区域图像	Inception V3			
决策支持	资产管理	地下配电系统故障风险评估与预防性维护 (Con Edison, 2005 至 2009)	馈线、电缆井特征集	故障风险评分	MB、EKD	回归、分类(监督学习)	较易(数据驱动、决策支持)	小样本(不均衡样本)机器学习、数据增强、数据的集中管理与分级共享
		故障快速修复 (PG&E, 2019)	灾害预报数据、GIS 数据、系统量测数据	故障概率分布参数估计	多元回归			
		客户反馈分析 (ABB&IBM, 2018)	客户的电话录音、文字评论等	关键词提取结果	NLP、NLU			
	控制	电池储能管理系统 Autobidder (Tesla, 2020)	价格、负荷、发电能力等历史数据	价格、负荷、发电能力预测	多种预测模型库	强化学习	难(数据-物理模型结合)	强化学习,提升模型的适应能力、稳定性、实时性
		微电网优化运行 (BluWave-AI, 2018; Emera&Sandia, 2020; Holy Cross Energy & NREL, 2020)	微电网实时量测	微电网控制策略	ADMM-RL			

升预测、监测工作的精度。

3)提升数据管理与模型训练的效率是发挥机器学习优势的关键。为了满足部分预测、监测工作的实时性要求,在工程应用中一方面应注重提升海量量测数据的传输、处理效率,譬如可采用专业的数据库软件开展数据的压缩、并行处理等工作;另一方面应加速机器学习模型的训练过程,如开展迁移学习、元学习等技术以及云计算架构的应用,以提升模型

的训练速度。

4)随着量测网络的逐步建立以及机器学习技术的普及,系统遭受网络信息攻击(cyber attack)的风险也随之增大。一旦系统中被注入虚假、错误数据,将造成机器学习感知结果的偏差,进而带来系统运行的风险。因此,在工程实践中应注重信息安全问题,提升对于异常数据的甄别能力。

4.3 决策支持类

电力系统运行的核心问题之一是优化决策问题,系统的建模、控制、资产管理均隶属于该范畴。根据机器学习技术与优化问题的结合方式不同,可将其分为4类:边界参数改进、优化参数选择、混合模型、替代模型^[46]。在该领域,机器学习实践经验总结如下。

1)在开展面向用户的建模工作时,应注重模型的个性化。这是因为不同的用户具有不同的负荷构成与用能偏好。传统的需求响应等负荷管理方式往往采用“一刀切”的办法,例如对参与响应的用户统一开断空调、热水器等设备。这种方式简单有效,但势必会影响用户的用能体验,在越来越注重用户体验的今天难以可持续发展。随着量测数据与控制手段的不断丰富,开展更加个性化、精细化的用户用能管理成为可能,在考虑每位用户自身偏好的前提下实现社会效益的最大化^[47]。但同时应注意到,模型精度与个性化程度的提升也增加了用户隐私暴露的风险。如何在挖掘用户数据的同时最大程度保护用户隐私,实现二者之间的平衡,是该领域工程实践中所要面对的问题。

2)在替代传统物理模型时,机器学习模型需要进行更多的实地测试以验证其鲁棒性。这是因为机器学习建模对于数据具有高度依赖性。当采集的训练样本不均衡、不全面时,容易造成机器学习建模结果出现偏差,导致其在部分场景下失效。因此,对机器学习建模结果进行多场景下的实地测试,明确模型的“有效边界”,或根据应用场景的不同适配不同的机器学习模型,将有助于提升机器学习建模的精度。

3)在系统运行过程中,机器学习技术主要为运行提供决策支撑(如2.1节中设备的故障预警、2.5节中提供故障恢复备选方案等),直接进行决策的情形较少。这主要是因为机器学习技术的“黑箱”特征,使得其结果存在解释困难而导致的“信任”问题,与电力系统高可靠性的运行要求相矛盾。随着目前神经网络架构的复杂化,该问题愈发凸显。也正是由于该原因,在近几年的工程实践中一些经典的机器学习技术(例如决策树、回归分析等)仍旧应用广泛。这是因为这些技术的成熟度、稳定性、可解释性较高,相比于“黑箱”性质显著的深度网络而言,对于运行人员更加友好,而这也是工程实践与理论研究的重要差异之一。因此,在工程实践中不应盲目追

求模型的复杂化,应综合考虑实际问题的复杂度、数据的可用性等因素,合理地选择机器学习模型。此外,研究可解释的机器学习(interpretable machine learning)技术是工程应用领域中的重要方向。

4)用于进行实时决策的强化学习目前工程应用案例较少。除了上述提到的可解释性问题外,强化学习技术通过与环境不断交互学习知识,在电力系统中可能带来较大的试错成本。此外,强化学习的模型训练耗时长、调参难度大、模型的泛化能力较弱等问题,均一定程度上制约了其在电力系统中的应用。现阶段强化学习技术主要应用于电网末端容错率较高的场景,如配电网电压控制、2.3节中微电网的运行决策等,即使出现错误的决策也不会造成显著的经济损失。但同时应注意到,相较于传统物理模型,强化学习模型对于环境变化的适应能力更强,在配用电领域仍具有深入研究的价值。

5)电力系统多数情况下运行在稳定状态,故障状态较少,相应的故障样本数据也相对较少。而这些少数故障样本往往更加具有研究价值。反观机器学习技术,通常需要大量的样本数据进行训练以学习其中规律,导致其处理故障状态这些少数“离群点”存在挑战。因此,研究小样本、不平衡样本下的机器学习技术(例如 few-shot learning^[48])、数据增强技术具有重要意义。

6)除了预测技术之外,机器学习技术在配用电规划领域的理论研究和工程应用均相对较少。其中部分原因是规划问题的实时性要求较低,规划人员具有充足的时间开展详尽的仿真分析以及优化模型的求解等工作,对机器学习技术依赖度较低。此外,规划问题要求结果具有强可解释性,这也与机器学习“黑箱”的特征存在一定冲突。但目前已有少量研究开始着眼于通过机器学习技术生成可行的馈线拓扑结构^[49],以提升规划工作的效率。这为规划领域的研究提供了新的思路。

4.4 北美工程实践反思

如上文所述,在北美配用电领域机器学习技术已经有诸多工程实践项目,项目执行主体涵盖政府研究机构、公用事业公司、高校、企业等。然而,现阶段北美地区该领域的机器学习项目资源共享性较差,通过一个项目得到的模型、数据、经验很难被其他项目有效利用^[50],这导致了如下问题。

1)缺乏指导:新项目开展时,难以充分整合利用领域内已有的工程实践经验,在技术路线上容易走

弯路。

2)重复劳动:新项目开展时,往往需要从零开始进行数据搜集、模型训练等工作,造成资源的浪费。

3)横向对比困难:每个实践项目基于自己独有的测试数据集与应用场景进行技术验证,难以横向对比技术的先进性与有效性。

导致以上问题的本质原因是机器学习技术自身“开源、共享”的特征与电力领域强调“保密、专用”之间的矛盾。为解决该问题,可考虑由权威政府机构牵头,联合各类研究机构以及相关电力企业组建机器学习的资源共享平台。具体可包括如下方法。

1)数据资料集中管理:对成员单位的工程实践项目进行统一的跟踪管理,整合项目的数据、模型、

实践经验等各类成果。

2)数据资料分级共享:按照保密程度对数据、模型进行安全分级,根据成员单位的权限分级共享。

3)数据匿名化:对真实数据进行匿名化处理,增强数据的可共享性。

4)建立智囊团:组建专家团队,对成员单位新开展的项目进行技术路线指导。

合理的共享机制有助于提升资源优化配置效率,促进整个领域的快速发展。

4.5 对中国配电网工程实践的启示

供电能力、可靠性以及电能质量是中国配电网工程实践中所关注的核心问题^[51],其中部分问题可由机器学习技术助力解决,梳理如表3所示。

表3 中国配电网工程实践问题及机器学习解决方案
Table 3 Problems of distribution network engineering practice in China and machine learning solutions

核心问题	具体问题	机器学习技术作用	对应北美工程实践案例
供电能力	局部供电能力不足问题	开展精细化负荷预测工作,超前规划系统网架	3.1节
		偏远地区、孤岛灵活采用光伏系统或含光伏微电网供电,由机器学习技术支撑其运行	2.3节
供电可靠性	部分设备老化或存在质量问题,故障率较高,配电自动化覆盖率有限	加强设备状态监测,提前预警设备故障,以及故障后快速修复	1.5节、2.1节、2.4节、2.5节
		挖掘多源量测数据,提升配电网态势感知精度	2.2节
电能质量	新能源、电动汽车等接入导致电压波动	提升风电、光伏等新能源出力预测精度	1.1节、1.2节
		深入挖掘居民侧需求响应潜力	3.5节
		分析用户反馈,及时发现问题	3.3节

综上,机器学习技术作为一种强大的工具,在智能配用电领域有着广阔的下沉空间。但同时应注意到,目前的人工智能机器学习技术仍处于“弱人工智能”阶段,具有诸多局限性。文献[52]指出,现阶段合适的应用场景应满足5个条件:①真实样本充足易获取;②具备一定的容错性;③目标任务单一;④边界确定的封闭环境;⑤任务静态,信息完备且确定。就配用电领域来看,许多应用场景尚无法完全满足以上条件,尤其是在容错性方面。

因此,为促进中国机器学习技术在配用电领域的良性发展,避免一拥而上的过热现象造成资源的浪费,建议在工程实践中关注以下几个方面:一是从实际需求出发,将机器学习技术与配用电领域的物理模型、专家经验深度融合,最终形成“物理+数据”双轮驱动的良好发展局面,避免陷入一味追求机器学习算法先进性的舍本逐末误区;二是如4.4节中提到,由权威政府机构牵头,建立配用电领域机器学习数据、模型、实践经验的分级共享机制,打破数据

壁垒,促进整个行业的快速发展;三是因地制宜循序渐进,对于农村等偏远地区,仍旧以改造基础设施提升供电质量为核心任务,而对于量测体系发达、资源充足的中心城区,可率先开展机器学习技术试点,并从提高自动化水平、态势感知等较为基础的应用场景切入,逐步过渡到决策支持甚至直接参与决策的高级应用场景;四是注重人才的培养与跨领域的交叉合作,确保能够在掌握配用电领域痛点问题的同时,保持对机器学习领域最新进展的动态跟进,实现“问题”与“方法”的最优匹配。

5 结语

本文聚焦机器学习技术在智能配用电领域的应用,从源、网、荷3个方面概述了北美地区的典型工程实践项目,从应用场景、技术路线等角度为国内相关研究与实践工作提供借鉴。在“双碳”目标的驱动下,未来中国新能源发电比重将稳步提升,储能、电动汽车、可控负荷等分布式电源也将广泛参与配电

网运行。在这个过程中,配用电领域的大数据特征将愈发显著,系统运行的复杂性、控制的实时性、建模的精细化程度都将不断提升,机器学习技术也将逐渐发挥更大的作用。

附录见本刊网络版(<http://www.aeps-info.com/aeps/ch/index.aspx>),扫英文摘要后二维码可以阅读网络全文。

参 考 文 献

- [1] Solar Energy Industries Association. Solar market insight report 2020 year in review[EB/OL]. [2020-10-26]. <https://www.seia.org/research-resources/solar-market-insight-report-2020-year-review>.
- [2] The Biden plan for a clean energy revolution and environmental justice[EB/OL]. [2019-01-15]. <https://joebiden.com/climate-plan/>.
- [3] JORDAN M I, MITCHELL T M. Machine learning: trends, perspectives, and prospects[J]. Science, 2015, 349 (6245): 255-260.
- [4] NAQA I E, MURPHY M J. What is machine learning?[M]. New York, USA: Springer, 2015.
- [5] SUYKENS J A K, VANDEWALLE J. Least squares support vector machine classifiers[J]. Neural Processing Letters, 1999, 9(3): 293-300.
- [6] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2012, 25 (2): 1097-1105.
- [7] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. Advances in Neural Information Processing Systems, 2014, 3(11): 1-9.
- [8] SUTTON R, BARTO A. Reinforcement learning: an introduction[M]. Cambridge, USA: MIT Press, 2018.
- [9] TAI K S, SOCHER R, MANNING C D. Improved semantic representations from tree-structured long short-term memory networks[J/OL]. Computer Science [2020-10-08]. <https://arxiv.org/abs/1503.00075v1>.
- [10] 杨挺,赵黎媛,王成山.人工智能在电力系统及综合能源系统中的应用综述[J].电力系统自动化,2019,43(1):2-14.
YANG Ting, ZHAO Liyuan, WANG Chengshan. Review on application of artificial intelligence in power system and integrated energy system[J]. Automation of Electric Power Systems, 2019, 43(1): 2-14.
- [11] 程乐峰,余涛,张孝顺,等.机器学习在能源与电力系统领域的应用和展望[J].电力系统自动化,2019,43(1):15-31.
CHENG Lefeng, YU Tao, ZHANG Xiaoshun, et al. Machine learning for energy and electric power systems: state of the art and prospects[J]. Automation of Electric Power Systems, 2019, 43(1): 15-31.
- [12] Watt-Son project [EB/OL]. [2020-11-03]. <https://www.solarreviews.com/news/sunshot-initiative-ibm-watt-sun-solar-forecasting-050717>.
- [13] AHMED R, SREERAM V, MISHRA Y, et al. A review and evaluation of the state-of-the-art in PV solar power forecasting: techniques and optimization[J]. Renewable and Sustainable Energy Reviews, 2020, 124: 109792.
- [14] LI Y, ZHANG S, HU R, et al. A meta-learning based distribution system load forecasting model selection framework[J]. Applied Energy, 2021, 294: 116991.
- [15] HE X, ZHAO K, CHU X. AutoML: a survey of the state-of-the-art[J]. Knowledge-Based Systems, 2021, 212: 106622.
- [16] Real-time wind turbine monitoring: data challenges, and rewards[EB/OL]. [2020-07-10]. <https://www.powermag.com/real-time-wind-turbine-monitoring-data-challenges-and-rewards/>.
- [17] Autobidder[EB/OL]. [2020-08-30]. <https://www.tesla.com/support/autobidder>.
- [18] SEVERSON K A, ATTIA P M, JIN N, et al. Data-driven prediction of battery cycle life before capacity degradation[J]. Nature Energy, 2019, 4: 383-391.
- [19] Intel automatic defect inspection using deep learning[EB/OL]. [2020-12-08]. <https://software.intel.com/content/www/us/en/develop/articles/automatic-defect-inspection-using-deep-learning-for-solar-farm.html>.
- [20] YU J, WANG Z, MAJUMDAR A, et al. DeepSolar: a machine learning framework to efficiently construct a solar deployment database in the United States[J]. Joule, 2018, 2 (12): 2605-2617.
- [21] Google inception V3[EB/OL]. [2020-12-04]. <https://cloud.google.com/tpu/docs/inception-v3-advanced>.
- [22] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]// IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, USA.
- [23] SAHA S S, SCHWEITZER E, SCAGLIONE A, et al. A framework for generating synthetic distribution feeders using OpenStreetMap[C]// North American Power Symposium, October 13-15, 2019, Wichita, USA.
- [24] SETO 2020-artificial intelligence applications in solar energy[EB/OL]. [2021-01-04]. <https://www.energy.gov/eere/solar/seto-2020-artificial-intelligence-applications-solar-energy>.
- [25] GROSS P, BOULANGER A, ARIAS M, et al. Predicting electricity distribution feeder failures using machine learning susceptibility analysis[C]// AAAI Conference on Artificial Intelligence, July 16, 2006, Boston, USA.
- [26] RADEVA A, RUDIN C, PASSONNEAU R, et al. Report cards for manholes: eliciting expert feedback for a learning task[C]// International Conference on Machine Learning and Applications, December 13, 2009, Miami, USA.
- [27] RUDIN C, PASSONNEAU R, RADEVA A, et al. A

- process for predicting manhole events in Manhattan [J]. Machine Learning, 2009, 80: 1-31.
- [28] RUDIN C, ERTEKIN S, PASSONNEAU R, et al. Analytics for power grid distribution reliability in New York City [J]. Interfaces, 2014, 44(4): 364-383.
- [29] LONG P M, ROCCO A S. Martingale boosting [C]// International Conference on Computational Learning Theory, June 27, 2005, Berlin, Germany.
- [30] GE mix core platform [EB/OL]. [2020-07-26]. https://www.ge.com/digital/sites/default/files/download_assets/improving-confidence-in-energy-distribution-network-connectivity-models-whitepaper.pdf.
- [31] Autonomous AI-controlled microgrids [EB/OL]. [2020-09-15]. <https://eepower.com/news/autonomous-ai-controlled-microgrids-with-renewable-energy/>.
- [32] BlockEnergy [EB/OL]. [2020-10-05]. <https://blockenergy.com/>.
- [33] KROPOSKI B, DALLANESE E, HODGE B M S, et al. Autonomous energy grids [R]. Golden, USA: National Renewable Energy Laboratory, 2017.
- [34] Pacific gas and electric smart grid annual report 2019 [EB/OL]. [2020-11-01]. https://www.pge.com/pge_global/common/pdfs/safety/how-the-system-works/electric-systems/smart-grid/AnnualReport2019.pdf.
- [35] California utilities hope drones AI will lower risk of future wildfires [EB/OL]. [2020-09-11]. <https://www.wsj.com/articles/california-utilities-hope-drones-ai-will-lower-risk-of-future-wildfires-11599816601>.
- [36] PG&E restoration work plan [EB/OL]. [2019-02-22]. https://www.pge.com/pge_global/common/pdfs/about-pge/environment/what-we-are-doing/electric-program-investment-charge/PGE-EPIC-Project-2.10.pdf.
- [37] WANG Y, CHEN Q X, HONG T, et al. Review of smart meter data analytics: applications, methodologies, and challenges [J]. IEEE Transactions on Smart Grid, 2018, 10(3): 3125-3148.
- [38] Real time loading data for distribution operations and planning [EB/OL]. [2010-02-20]. https://www.pge.com/pge_global/common/pdfs/about-pge/environment/what-we-are-doing/electric-program-investment-charge/PGE-EPIC-Project-2.07.pdf.
- [39] PAN S J, YANG Q. A survey on transfer learning [J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10): 1345-1359.
- [40] Deepmind AI reduces Google data center cooling bill by 40% [EB/OL]. [2010-08-05]. <https://deepmind.com/blog/article/deepmind-ai-reduces-google-data-centre-cooling-bill-40>.
- [41] CHEN Y, ELENNE A J D, WEBER G. IBM Watson: how cognitive computing can be applied to big data challenges in life sciences research [J]. Clinical Therapeutics, 2016, 38(4): 688-701.
- [42] Machine learning helps power down electricity theft in Jamaica [EB/OL]. [2019-09-18]. <https://www.esmap.org/multimedia/machine-learning-helps-power-down-electricity-theft-jamaica>.
- [43] Codes of energy theft protection [EB/OL]. [2020-05-21]. <https://github.com/impactlab/jps-handoff>.
- [44] How AI unlocks the full value of demand response [EB/OL]. [2020-04-24]. <https://info.energyhub.com/blog/demand-response-artificial-intelligence>.
- [45] ENSLEY M R. Toward a theory of situation awareness in dynamic systems [M]. London, UK: Routledge, 2017.
- [46] RUAN G, ZHONG H, ZHANG G, et al. Review of learning-assisted power system optimization [J]. CSEE Journal of Power and Energy Systems, 2021, 7(2): 221-231.
- [47] LI Y, YAN Z, CHEN S, et al. Operation strategy of smart thermostats that self-learn user preferences [J]. IEEE Transactions on Smart Grid, 2019, 10(5): 5770-5780.
- [48] WANG Y, YAO Q, KWOK J T, et al. Generalizing from a few examples: a survey on few-shot learning [J]. ACM Computing Surveys, 2020, 53(3): 1-34.
- [49] LIANG M, MENG Y, WANG J, et al. FeederGAN: synthetic feeder generation via deep graph adversarial nets [J]. IEEE Transactions on Smart Grid, 2021, 12(2): 1163-1173.
- [50] SEAB AIML Working Group. Final report of the SEAB to Secretary of Energy Dan Brouillette regarding the Department of Energy and Artificial Intelligence [R/OL]. [2020-08-12]. https://www.energy.gov/sites/default/files/2020/11/f80/SEAB_AIML_Report.pdf.
- [51] 国家能源局. 配电网建设改造行动计划(2015—2020年) [EB/OL]. [2020-02-01]. http://www.gov.cn/xinwen/2016-08/15/content_5099597.htm.
National Energy Administration. Action plan for construction and renovation of distribution network (2015-2020) [EB/OL]. [2020-02-01]. http://www.gov.cn/xinwen/2016-08/15/content_5099597.htm.
- [52] 张钹. 走向真正的人工智能 [C]// 全球人工智能与机器人峰会, 2018年6月29日, 深圳.
ZHANG Bo. Towards a real artificial intelligence [C]// Global AI and Robotics Conference, June 29, 2018, Shenzhen, China.

李亦言(1991—),男,博士,主要研究方向:机器学习与人工智能技术、配电网优化运行。E-mail: yli257@ncsu.edu

胡荣兴(1989—),男,博士研究生,主要研究方向:配电网建模与优化运行。E-mail: rhu5@ncsu.edu

宋立冬(1992—),男,博士研究生,主要研究方向:机器学习与人工智能技术。E-mail: lsong4@ncsu.edu

陆宁(1972—),女,通信作者,博士,教授,主要研究方向:人工智能技术、电力系统建模与优化运行。E-mail: nlu2@ncsu.edu

(编辑 王梦岩)

Application of Machine Learning in Field of Smart Power Distribution and Utilization: Overview of Engineering Practice in North America

LI Yiyan^{1,2}, HU Rongxing¹, SONG Lidong¹, JIA Qiangang², LU Ning¹

(1. Department of Electrical and Computer Engineering, North Carolina State University, Raleigh 27695, USA;

2. School of Electric Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: Machine learning technique is important for assisting the energy transition and promoting the renewable energy consumption. In recent years, the application of machine learning technique in power systems has been widely concerned. Due to the ‘black box’ nature of machine learning technique, its interpretability and robustness are still to be improved. And there is a certain contradiction with the operation requirements of high reliability in the power system, which leads to its practical engineering application lagging behind the theoretical research. In order to introduce the practical application of machine learning technique, this paper focuses on the field of power distribution in North America. The typical engineering practice projects of machine learning technique are summarized from the perspectives of source, network and load, and the method, effect and inspiration of each project are also outlined. Further, the above projects are classified into two categories, i.e., situational awareness and decision support, and a total of five application scenarios. And the research areas of machine learning technique in the next stage is analyzed from the perspective of engineering practice.

Key words: machine learning; smart power distribution; methodology analysis; engineering practice

