

人工智能大数据平台实现生产数据智能分析与决策

郝显阳 张静晨 郝光耀

瀚蓝城市环境服务有限公司

DOI:10.32629/etd.v7i1.22204

[摘要] 生产制造系统在工作过程中不断产生大量、多源、异构的数据资源,传统的基于经验规则以及离线统计的方法已经不能满足当今复杂的动态环境中对于快速响应的需求。而人工智能与大数据平台相结合,可以实现对生产的整个流程进行数据管理、深层次分析以及智能化决策。基于此,在对生产数据的特点以及传统方法存在的问题进行分析之后,提出一种适用于智能分析和决策的人工智能大数据平台分层架构,介绍了数据治理、机器学习、深度模型、知识图谱等相关技术和方法,详细阐述了故障预测、参数优化、质量追溯、风险预警等主要决策手段。该平台可以使数据变成有价值的洞察力,使企业由以往的经验导向型转变为以数据为导向,从而实现具有自我感知、自我学习、自我决策的能力的智能制造。

[关键词] 人工智能; 大数据平台; 生产数据; 智能分析; 决策支持

中图分类号: TP311.13 **文献标识码:** A

Artificial Intelligence Big Data Platform for Intelligent Analysis and Decision-Making of Production Data

Xianyang Hao Jingchen Zhang Guangyao Hao

Hanlan Urban Environmental Services Co., Ltd.

[Abstract] Production and manufacturing systems continuously generate large volumes of data resources that are diverse in source and heterogeneous in structure during operation. Traditional methods based on empirical rules and offline statistics can no longer meet the need for rapid responses in today's complex and dynamic environments. The integration of artificial intelligence with big data platforms enables data management, in-depth analysis, and intelligent decision-making throughout the entire production process. Accordingly, after analyzing the characteristics of production data and the problems associated with traditional methods, this study proposes a layered architecture for an artificial intelligence big data platform suitable for intelligent analysis and decision-making. Relevant technologies and methods, including data governance, machine learning, deep learning models, and knowledge graphs, are introduced, and major decision-making approaches such as fault prediction, parameter optimization, quality traceability, and risk early warning are discussed in detail. The platform can transform data into valuable insights, enabling enterprises to shift from experience-oriented operations to data-driven management and thereby achieve intelligent manufacturing capabilities characterized by self-awareness, self-learning, and self-decision-making.

[Key words] artificial intelligence; big data platform; production data; intelligent analysis; decision support

引言

目前,全世界制造业正在迅速朝着数字化、网络化、智能化的方向发展,在整个生产过程中会产生大量的包括设备、工艺、质量等方面的数据,形成一个生产系统的数字孪生体。但是大多数的企业还是依靠人的经验和传统的统计方法来处理这些海量而且复杂的数据,不能够很好地解决数据是非线性和强关联以及变化快的问题,造成决策迟缓甚至是错误。而人工智能(比如

深度学习、强化学习、知识图谱)在模式识别、预测优化、推理决策方面的进步加上大数据平台提供的大容量存储能力、强大的计算能力和良好的扩展性,使得全过程智能化成为可能。“人工智能大数据平台如何助力生产智能决策”,这是本文的重点所在,主要介绍了它的工作原理、关键技术及核心决策逻辑。

1 生产数据智能分析的现实需求与挑战

1.1 生产数据的类型与特征

生产数据来源丰富,来源于生产设备上的各种传感器(如振动、温度、压力、电流等)、生产工艺过程数据(如速度、流量、浓度等)、产品质量检验数据(如尺寸、性能、缺陷等)、原材料追溯信息(如批次号、库存量、运输情况等)以及工作环境中的相关信息(如粉尘、噪声、气体等)。从数据形式上看,包括结构化的数字型数据、半结构化的日志型数据及非结构化的文本文档或者图片;从时间跨度来看,既有毫秒级的高频率数据也有天周级别的低频率数据;从相互之间的关系看,则是不同工序之间甚至是不同系统之间的互相联系^[1]。总的来说,生产数据有五大特点:数量大(Volume)、种类繁多(Variety)、速度快(Velocity)、价值密度较低(Value)以及真实性问题(Veracity),这些问题都对传统的数据库以及一般的处理方法提出了巨大考验。

1.2 传统数据分析方法的局限

传统的生产数据分析主要是基于统计过程控制、回归分析、假设检验及人的主观经验。而这些方法对于小规模的数据集、明显的线性关系是有一定作用,但在当今复杂的制造业中,由于存在大量的变量、非线性和大量的干扰因素等情况下,上述方法已经不能满足需求。首先,这些方法都是建立在严格的模型之上,在实际不断变化的生产环境中很难适用;其次,特征的选择需要人为进行,无法充分发挥大数据本身所蕴含的价值;第三,得到的结果大多数是总结性的,无法做到及时预警以及提前给出建议;第四,分析的过程和做出决定的过程分离,在分析之后才给出结论,不能形成一个良好的循环。所以需要一种能够自我调节,自我学习并且有很强鲁棒性的智能分析方法来解决这些问题。

1.3 智能分析与决策的变革需求

智能制造的核心要义是利用数据进行准确感知、科学决策以及灵活行动。企业对于生产的分析已经不再满足于仅仅知道发生了什么,而是希望理解造成这种现象的原因是什么、未来会发生什么样的情况以及如何改进。即不仅需要能够及时发现设备状况并发出警报,还需要能够在不断变化的情况下调整生产工艺,还需要可以对未来产出及耗能做出估计从而更好地安排生产计划,还需要能够迅速找出质量问题根源来降低废品率。这都对分析平台提出了很高的要求,它不仅要能够快速处理大量信息,还要能结合各种来源的信息,还要能够发现其中的关系并且自主地进行改善。而这正是人工智能与大数据相结合所能带来的帮助。

2 人工智能大数据平台的架构与关键技术

2.1 平台总体分层架构

最底层是数据采集以及接入层,在工业以太网、5G、OPCUA、MQTT等多种协议的支持下,连接各种生产设备以及控制系统,实现多种来源的数据接入并进行相应的协议转换工作。中间一层是存储及计算层,在分布式文件系统(例如HDFS)、时序数据库(例如InfluxDB)、列式存储的基础上,能够支持高频率写入以及低延迟读取的需求;同时支持批流一体处理能力,可以用于离线模型训练也可以用于在线推理工作。再往上一层是智能化分析

层,集成了各种机器学习库(例如Scikit-learn)、深度学习框架(例如TensorFlow、PyTorch)、以及知识图谱构建工具等,提供了灵活可配置的算法流水线。再往上一层是决策服务层,把上层的结果以标准接口形式或者规则引擎或者是图形化视图展示给用户,对接到制造执行系统、企业资源计划系统中去,发出警告、建议或是控制命令等。最顶层是统一运维管理模块,包括资源分配、监控报警、模型版本控制以及日志审核等功能,保障整个平台平稳高效运转。

2.2 数据治理与预处理关键技术

平台需要有一套完善的数据管理体系,如元数据管理、主数据定义、质量规则检查以及数据血缘追踪等,保证数据的质量可靠以及可追溯。在预处理阶段主要解决的问题有空缺值填补,异常值处理,时间戳对齐以及采样率不同,常用的方法有插值法,根据统计学中的离群值进行剔除,时间对齐和重采样。对于来自多个来源不同类型的数据,则要进行实体对齐及特征级融合形成一个统一的整体^[2]。而特征工程则是利用专业知识设计出有意义的新特征,在此基础上再对特征进行选择与降维以获得最具有代表性的少数几个特征来提升后续机器学习的效果。

2.3 智能分析核心算法体系

对于设备故障以及使用寿命预测问题上,支持向量机、随机森林、梯度提升树等方法适合少量数据分类任务,而卷积神经网络、循环神经网络以及Transformer可以用于从时序信号中提取衰减信息进而实现对设备健康状况进行评价;而在工艺参数调优过程中,通过贝叶斯优化和深度强化学习在连续动作空间寻找符合要求的最佳配置同时兼顾探索与利用;在质量问题根源查找过程里,使用因果推断方法(如Do-calculus)以及图神经网络在工艺流程图上找出主要的影响因素及其相互作用关系;而在排程调度上,则结合启发式方法以及多智能体强化学习来进行多目标、实时的决策制定;此外,知识图谱将设备说明书、工艺标准以及专家知识以三元组的形式表示出来并与基于数据的方法相结合提高最终结果的可解释性和泛化能力。

2.4 决策引擎与交互可视化机制

平台中存在一个决策引擎,用于对接收的模型结果进行规则过滤、置信度评分以及优先级排序,给出备选项供人工审核并最终确定,实现人机结合的方式做决定。此外,决策引擎还有冲突解决以及容错能力,在多个模型给出不同意见时,根据事先设定好的规则来做出判断。可视化交互层有多种类型、可配置的仪表盘,可以查看生产线的数据情况、钻取到具体的异常点分析以及未来的发展趋势等,在界面上拖拉就可以完成相应的设置,减少技术人员的工作负担。而且还可以用自然语言提问以及自动生成报表,让管理者方便快捷地了解分析概况,有利于形成以数据为基础的民主决策氛围。

2.5 云一边一端协同部署架构

为满足生产一线对于低时延以及高可靠性的需求,该平台采用云一边一端三层架构设计。端侧是部署在现场的各类传感器以及设备控制器等,完成原始数据获取及初步的数据判断工

作;边缘节点部署轻量化的推理引擎,在线进行数据分析、异常监测以及快速反馈控制等操作,时延控制在毫秒级别以内;而在云端进行复杂模型训练、整体优化以及大量历史数据保存等工作,并定期将最新的模型权重推送到边缘节点上。这样既能够做到重要决策及时做出,又可以发挥出云侧强大计算能力的作用,同时也降低了对网络的要求,非常适合一些网络环境较差或者对于隐私保护有较高要求场合使用。

3 智能分析与决策的核心机制

3.1 数据驱动的故障预测与健康管理

设备意外停机会给工业企业带来巨大损失,平台不断获取设备工作时的各种参数信息(振动频谱、温度变化趋势、电流谐波等),构建多层次健康的基准线,用自编码器或者孤立森林等方式检测与基准线之间的差异大小,达到提前发现故障的目的,基于深度迁移学习的方法可以在不同的工况以及不同的机器上共享故障的知识,在少量的数据下也可以很好的进行训练^[3]。预测性维护不是定期更换,而是根据设备的实际磨损情况进行安排检修的时间点,可以大大减少不必要的停机时间和过多的维修带来的费用,提高设备的整体性能。并且还可以自动给出维修建议以及备件的需求,帮助管理人员做出更好的决定。

3.2 工艺参数动态优化与自适应调控

在现实中,由于原材料质量、环境温度、湿度以及人为操作等因素变化都会使工艺发生偏移从而影响产品质量的一致性问题,在此平台上可以使用不断产生的实时数据建立工艺参数与产品质量之间的代理模型(例如高斯过程回归或者神经网络)来预测产品质量的变化情况,在发现有重要质量指标发生变化的趋势时,由优化器给出最优参数值并通过控制器进行调整,即实现了一种“感知——分析——决策——执行”的实时反馈循环。这也就使得该平台具有很强的灵活性以应对各种外界干扰并且节省大量的人力资源同时也能够不断自我改进提高自身性能以满足日益增长的需求。

3.3 产品质量追溯与根因诊断

一旦出现质量问题,及时找到根本原因以最小化损失非常重要。平台收集全链条生产过程相关数据,使用因果贝叶斯网络或者结构方程模型,追踪每一个批次在各个工序、每台设备以及每个操作参数上偏离情况,然后用反事实推理计算出不同因素对产品质量造成影响大小。此外,在工艺流程图中利用图神经网络传播信息,找出最有可能引起的问题传播途径^[4]。根因诊断结果以图形化方式展示给工程师,使其快速关注重点部位,节省查

找时间并形成数据库供以后类似问题快速查找到相应解决方法。

3.4 供应链协同与风险预警

生产决策不仅仅是在车间内部进行,还需要与上游原材料供应、中间库存以及下游市场需求密切配合。平台收集供应商供货情况、材料库存量、物流运输进度以及销售预测等信息,用长短时记忆网络和蒙特卡洛方法来预测未来几周可能出现的供货不足以及交货延误问题,在预测结果表明可能存在的短缺大于可接受的程度之后就发出多个级别的警报并且给出一些可供选择的供应商或者是改变生产的建议,这就有助于建立更加有弹性的供应链,减少牛鞭效应的影响,从而在降低库存的同时避免缺货的发生。

4 结语

人工智能大数据平台是进行生产数据智能化分析与决策的技术支撑体系。为了解决传统方法对于多源异构生产数据处理不足的问题,该平台搭建了一套包括采集、存储、计算、分析以及决策的服务化层次结构,并结合数据治理、机器学习、深度模型、知识图谱以及云计算边缘计算协同等技术。基于故障预测、参数优化、质量问题根本原因回溯、供应链风险预警以等方式使原始数据变成有行动力的智能决策。它的主要作用就是把分散的数据提升到整个企业的信息资源甚至是企业的一种新的能力上,使制造企业由依靠经验变为依赖数据,由事后弥补变为防患于未然,由点到面得到质变。而未来随着因果推理、生成式AI以及自主智能体的发展,平台可以自动规划生产工艺以及调度安排,但是要始终贯彻以人为本的原则,在技术上起到辅助的作用,做到让决策公开透明可解释可控,打造一个安全高效的和谐的人机共融智能制造环境。

[参考文献]

- [1]李伯虎,柴旭东,刘阳,等.智能制造中的大数据技术及应用[J].中国科学:信息科学,2022,52(4):605-625.
- [2]姚锡凡,刘敏,张剑铭,等.面向智能制造的人工智能方法研究进展[J].机械工程学报,2022,58(8):1-18.
- [3]杨海成,肖莹莹,徐亮.工业互联网平台关键技术与应用趋势[J].计算机集成制造系统,2023,29(2):345-359.
- [4]陈积明,陈建新,陈虹.工业人工智能:从感知到决策的变革[J].控制与决策,2024,39(1):1-11.

作者简介:

郝显阳(1997--),男,汉族,河南人,本科,助理工程师,研究方向:环境工程。