

批量生产中基于大数据的制造过程异常检测与预警机制研究

孙运宜

烟台科技学院 265600

DOI: 10.32629/ems.v8i2.18501

[摘要] 文章聚焦批量生产领域，系统探究基于大数据技术构建制造过程异常检测与预警机制的路径与方法。契合《“十四五”智能制造发展规划》中“推动大数据、人工智能等新一代信息技术与制造技术深度融合”的战略部署，通过对制造过程产生的海量数据进行深度分析，挖掘数据隐含的规律与特征，实现对制造过程异常情况的精准识别与及时预警。这一机制不仅能够有效提升生产效率、降低生产成本、保障产品质量，更是贯彻落实《工业互联网创新发展行动计划》中“强化工业数据治理能力”的具体实践。明确了相关关键技术与实施方法，为批量生产企业响应国家“数实融合”号召、优化制造过程管理水平提供理论支撑与实践参考，助力制造业高质量发展战略目标的实现。

[关键词] 批量生产；大数据；异常检测；预警机制

引言

在制造业高速发展的当下，批量生产模式已成为众多企业提升生产效率、控制成本的核心手段^[1]。然而，随着生产规模的扩大与生产流程的复杂化，制造过程中异常情况的发生概率显著增加。此类异常不仅可能导致生产延误、产品质量下滑，甚至可能引发安全事故，给企业造成重大经济损失。传统异常检测方法多依赖人工经验与简单统计分析，难以应对海量、复杂的生产数据，且在检测精度与实时性方面存在明显不足。大数据技术的崛起为解决这一难题提供了全新的思路与技术路径。通过对制造过程中产生的海量数据进行采集、存储与深度分析，能够精准识别潜在的异常模式，实现提前预警，进而推动制造过程的智能化管理。因此，开展基于大数据的制造过程异常检测与预警机制研究，具有重要的现实意义与应用价值。

一、批量生产制造过程大数据特征及相关技术

(一) 制造过程大数据特征

1. 数据海量性

在批量生产过程中，从原材料采购、生产设备运行、产品质量检测到成品入库等全流程环节，均会产生大规模数据^[2]。以现代化生产线为例，每分钟可生成数千条设备运行参数记录，单日数据量可达数十万乃至数百万条。这些海量数据涵盖设备状态、生产工艺参数、产品质量指标等多维度信息，为异常检测工作提供了丰富的数据基础。

2. 数据多样性

制造过程数据类型呈现多元化特征，具体包括：结构化数据，如生产计划、设备运行参数、质量检测结果等，可通

过固定表格结构实现存储与处理；半结构化数据，如生产日志、设备维护记录等，具备一定结构特征，但不完全符合关系型数据库的存储模式；非结构化数据，如设备故障报警文本、产品外观图像等。不同类型数据蕴含的信息存在差异，需采用针对性的处理方法开展分析工作。

3. 数据时效性

制造过程数据具有极强的时效性，尤其在高速运转的生产线中，数据的实时性是及时发现异常情况的关键前提^[3]。例如，设备运行参数的瞬时变化可能预示设备即将发生故障，若无法及时捕捉此类变化并采取应对措施，可能引发严重的生产事故。因此，对制造过程数据进行实时采集、传输与分析，成为异常检测与预警工作的核心要求。

(二) 大数据分析技术在制造过程中的应用

1. 数据挖掘技术

数据挖掘技术能够从海量制造数据中提取有价值的信息与模式^[4]。例如，通过关联规则挖掘可发现生产工艺参数间的潜在关联，在汽车制造领域，可识别特定零部件装配顺序与产品最终质量的关联关系，为生产流程优化提供依据；借助聚类分析可对相似生产过程或产品进行分类，助力发现异常生产批次或产品类别。

2. 机器学习算法

机器学习算法在制造过程异常检测中发挥关键作用^[5]。监督学习算法（如支持向量机（SVM）、决策树）可通过对大量已知正常与异常样本的学习，构建分类模型，用于判断新数据是否属于异常情况；无监督学习算法（如主成分分析（PCA）、孤立森林算法）能够在无标注数据的场景下，识别

数据中的异常点。例如, PCA 可对高维制造数据进行降维处理, 提取核心特征, 通过分析特征变化实现异常检测; 孤立森林算法通过构建随机树孤立异常点, 达成异常识别目标。

3. 深度学习技术

深度学习作为机器学习的重要分支, 具备强大的特征学习能力, 尤其适用于处理复杂制造数据。例如, 卷积神经网络 (CNN) 在图像识别领域的应用已较为成熟, 可用于产品外观缺陷检测, 通过对大量正常与缺陷产品图像的学习, 能够精准识别产品表面的划痕、裂纹等缺陷; 循环神经网络 (RNN) 及其变体长短期记忆网络 (LSTM) 适用于处理具有时间序列特征的制造数据 (如设备运行状态时间序列数据), 可捕捉数据中的长期依赖关系, 实现设备故障预测。

(三) 制造过程异常的类型及表现形式

1. 设备故障异常

设备是批量生产的核心载体, 设备故障异常主要表现为设备运行参数超出正常范围, 如温度过高、压力超标、振动异常等。以化工生产为例, 反应釜温度骤升可能是冷却系统故障的信号, 若未及时处理, 可能导致产品质量问题甚至引发安全事故。此外, 设备故障还可能表现为部件损坏 (如机械零件磨损、断裂), 直接影响生产连续性与产品质量。

2. 工艺参数异常

工艺参数异常指生产工艺参数偏离设定标准值。在电子芯片制造过程中, 光刻工艺的曝光时间、蚀刻深度等参数若出现偏差, 可能导致芯片性能下降甚至报废。工艺参数异常的成因包括操作人员失误、设备精度不足、原材料特性变化等。

3. 产品质量异常

产品质量异常表现为产品质量指标不符合规定标准, 如尺寸偏差、性能不达标、外观缺陷等。例如, 汽车零部件尺寸精度不达标会影响整车装配质量与性能; 电子产品电气性能不符合要求则无法正常运行。产品质量异常通常是设备故障、工艺参数异常、原材料质量问题等多因素共同作用的结果。

二、基于大数据的制造过程异常检测与预警机制构建

(一) 数据采集与预处理

1. 数据采集系统设计

构建高效的数据采集系统是获取制造过程数据的基础保障, 该系统需以“全维度、高实时、高可靠”为核心目标, 实现对制造全流程数据的精准捕捉。从技术架构来看, 系统需整合边缘计算节点、工业传感器网络与数据传输协议 (如 MQTT、OPC UA)。其中, 边缘计算节点部署于生产车间本地,

如同数据的“前置处理站”, 可实现对设备数据的实时预处理与初步筛选, 极大减少数据传输压力; 工业传感器网络则根据不同生产场景需求, “因地制宜”地选择适配的传感器类型, 例如在高温生产环境中选用耐高温的热电偶传感器采集温度数据, 在高速旋转设备上安装压电式振动传感器捕捉振动参数, 确保每个关键数据都能被准确采集。

在实际应用中, 系统需覆盖从原材料投入到成品出库的全生产链条数据: 通过在原材料仓储区部署 RFID 识别设备, 实时采集原材料的入库时间、规格型号、批次信息等; 在生产设备上部署多类型传感器, 同步采集设备温度、压力、转速、电流电压等运行参数, 以汽车零部件生产线为例, 可实现每分钟采集 1000 + 条设备运行数据; 在质量检测环节, 利用机器视觉检测设备、高精度测量仪器等, 自动采集产品的尺寸精度、外观特征、性能参数等质量数据, 如同为产品进行全方位的“体检”, 不放过任何细微信息。

2. 数据清洗与转换

采集的原始数据往往存在噪声、缺失值、重复值等问题, 需开展数据清洗工作。针对噪声数据, 可采用滤波算法进行平滑处理; 对于缺失值, 需根据数据特征选择适宜的填充方法 (如均值填充、中位数填充、基于机器学习算法的预测填充)。数据转换环节需将数据转换为适用于分析的形式, 如对分类数据进行编码、对数值数据进行标准化处理, 以消除不同特征间的量纲差异, 提升数据分析准确性。

3. 数据存储与管理

制造过程产生的海量数据需依托高效的存储与管理方案。可采用分布式存储系统 (如 Hadoop 分布式文件系统 (HDFS)), 将数据分散存储于多个节点, 如同将数据分散存放在不同的“仓库”, 提升数据存储的可靠性与扩展性。同时, 结合关系型数据库 (如 MySQL) 与非关系型数据库 (如 MongoDB), 实现不同类型数据的分类存储。

(二) 异常检测模型构建

1. 基于机器学习的异常检测模型选择与训练

需根据制造过程数据特征与异常类型, 选择适配的机器学习算法构建异常检测模型。对于特征明确的异常情况, 可选用监督学习算法开展训练。例如, 针对已知设备故障类型, 收集大量正常与故障状态下的设备运行数据作为训练样本, 采用 SVM 算法构建分类模型, 用于判断设备是否处于故障状态。对于无标签数据或异常模式未知的场景, 可采用无监督学习算法 (如孤立森林算法)。在模型训练过程中, 需合理调整算法参数, 提升模型检测精度与泛化能力; 同时, 采用交叉验证等方法对模型进行评估与优化, 确保模型可靠性。通

过不断优化参数与评估,让模型如同经过严格训练的“数据侦探”,精准识别异常情况。

2. 深度学习模型在异常检测中的应用

对于产品图像、声音信号等复杂制造数据,深度学习模型具备独特优势。以产品外观缺陷检测为例,构建基于 CNN 的深度学习模型的流程如下:首先,收集大量正常与缺陷产品图像数据,对图像进行裁剪、归一化等预处理操作;其次,将预处理后的图像数据输入 CNN 模型进行训练,通过卷积层、池化层、全连接层等网络结构,实现图像特征的自动学习;模型训练完成后,可精准识别产品是否存在缺陷及缺陷类型。在实际应用中,可实时采集产品图像并输入训练完成的模型进行检测,实现产品外观缺陷的快速、精准识别。该过程中,深度学习模型就像拥有“火眼金睛”,能快速且准确地发现产品外观的细微瑕疵。

3. 多模型融合策略

为提升异常检测的准确性与可靠性,可采用多模型融合策略。将不同类型的异常检测模型(如机器学习模型与深度学习模型)进行融合,充分发挥各模型优势。例如,在设备运行数据异常检测中,可同时采用基于机器学习的决策树模型与基于深度学习的 LSTM 模型:决策树模型可快速输出初步检测结果,LSTM 模型可捕捉数据中的时间序列特征,对复杂异常情况进行精准判断。通过加权平均等方法对两个模型的检测结果进行融合,得到最终异常检测结果,进而提升检测精度与稳定性。多模型融合就像不同技能的专家组成团队,各司其职又协同合作,让异常检测更高效准确。

(三) 预警机制设计

1. 预警指标确定

需结合制造过程特征与异常类型,确定科学合理的预警指标。针对设备故障异常,可将设备运行参数变化率、参数偏离正常范围程度等作为预警指标;针对工艺参数异常,可将工艺参数与标准值的偏差、参数波动幅度等作为预警指标;针对产品质量异常,可将产品质量指标不合格率、关键质量参数变化趋势等作为预警指标。通过对预警指标的实时监测,及时识别潜在异常情况。这些预警指标如同生产过程中的“信号灯”,一旦出现异常变化,就能及时发出警示。

2. 预警阈值设定

预警阈值是判断是否触发预警信号的关键参数,其设定需综合考虑生产过程实际情况、历史数据及误报、漏报风险。若阈值设定过高,可能导致漏报,使异常情况无法得到及时处理;若阈值设定过低,可能引发误报,增加不必要的干扰。

可通过对历史数据的统计分析,结合生产经验,确定适宜的预警阈值。例如,通过分析设备运行参数历史数据,计算正常工况下参数的波动范围与分布特征,以此为基础设定合理预警阈值。

3. 预警信息发布与处理

当检测到异常情况且预警指标超出设定阈值时,需及时发布预警信息。预警信息应包含异常类型、异常发生位置、发生时间及相关异常参数等详细内容,便于生产人员快速掌握情况并采取应对措施。预警信息可通过短信、邮件、生产管理系统弹窗提示等多种方式发布。生产人员收到预警信息后,需根据异常严重程度启动相应应急预案:对于轻微异常,可开展进一步观察与分析,判断是否需要调整生产参数或进行设备维护;对于严重异常,需立即停止生产,组织专业人员开展故障排查与修复工作,保障生产过程安全稳定。整个预警信息发布与处理流程,如同一个高效的“应急指挥系统”,确保异常情况能得到及时且妥善的处理。

结束语

综上所述,基于大数据的制造过程异常检测与预警机制,为批量生产企业提供了高效的生产过程管理工具。通过对制造过程大数据的采集、预处理、异常检测模型构建及预警机制设计,能够及时发现生产过程中的异常情况并实现预警,进而提升生产效率、降低生产成本、保障产品质量。然而,该机制在实际应用中仍需不断优化完善,以适应不同生产环境与生产流程的需求。未来,随着大数据技术、人工智能技术的持续发展,基于大数据的制造过程异常检测与预警机制将不断迭代升级,为制造业智能化发展提供更坚实的技术支撑。

[参考文献]

- [1]王泽生,王欢. 生产中小批量调试的解决方案[J]. 汽车工艺师, 2025, (04): 26-28.
- [2]陈玉莹,于惠霞,张甫. 精益管理在小批量生产中的应用[J]. 营销界, 2024, (23): 149-151.
- [3]杨维建,斯兴瑶,王文慧,等. 总装厂整车小批量试生产中静态扭矩管理的研究[J]. 汽车实用技术, 2024, 49 (14): 99-102+130.
- [4]詹璐歆. 批量生产中木质功能板材平衡含水率精确控制技术研究与应用[D]. 贵州大学, 2024. DOI: 10.27047/d.cnki.ggudu.2024.000692.
- [5]吴义彬. 精益生产在多品种小批量电缆组件生产中的应用[J]. 机电元件, 2023, 43 (03): 56-57.