

Prädiktive Prozessüberwachung in der Batterieproduktion

Kompetenzcluster InZePro

Verfasst im Projekt KlproBatt



Impressum

Herausgeber

Universität Bayreuth

Autorinnen und Autoren

Universität Bayreuth, Lehrstuhl für Wirtschaftsinformatik:
Prof. Dr. Maximilian Röglinger, Tobias Fehrer, Tim Meyer-
Hollatz, Carl Luippold.

Bildnachweis

Tara Winstead (Titel) / Flaticon.com

Abbildungen: Universität Bayreuth

Projektförderer

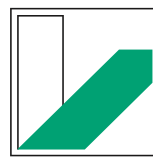
Bundesministerium für Bildung und Forschung
Projektname: KIproBatt - Intelligente Batteriezellfertigung
mit KI-gestütztem Prozessmonitoring auf Basis einer generi-
schen Systemarchitektur
Förderkennzeichen: 03XP0309B

Stand

März 2024

Lizenz

Dieses Werk ist lizenziert unter CC BY 4.0. Um eine Kopie
dieser Lizenz zu sehen, besuchen Sie
<http://creativecommons.org/licenses/by/4.0/>



UNIVERSITÄT
BAYREUTH



Fraunhofer

IGCV



Fraunhofer

ISC



GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung



Editorial

Intelligente Batteriezellfertigung mit KI-gestütztem Prozessmonitoring auf Basis einer generischen Systemarchitektur

Die Batteriezellenfertigung ist ein ressourcenaufwendiger Prozess mit hoher Ausschussquote. Durch die intelligente Verknüpfung von digitaler Abbildung der Produktion, prädiktiver und präskriptiver Vorhersagen und der Einbeziehung nachhaltiger KPIs können Defekte vermindert und die Qualität produzierter Zellen erheblich erhöht werden.

Erste Ansätze zur Verringerung dieser Technologien wurden bereits entwickelt, aber die Implementierung in einer realen Produktionsanlage fehlt noch. Um insbesondere die nationale Zellproduktion zu unterstützen, werden im Rahmen des Projektes "KIproBatt" gezielt Fortschritte in den Bereichen der Digitalisierung und der statistischen Datenauswertung unterstützt.

Mit den im Projekt entwickelten KI-Modellen und -Methoden kann die Zielsetzung einer Produktivitätssteigerung und Erhöhung der Produktqualität erreicht werden. In einem prototypischen Stadium konnten wir zeigen, dass der Einsatz von KI in der Batteriezellenfertigung die Produktqualität deutlich erhöhen kann. Die daraus generierten Erkenntnisse und eine Blaupause zu eigener Implementierung wollen wir im Rahmen dieses Leitfadens mit Ihnen teilen.

Lassen Sie uns gemeinsam dafür sorgen, eine nachhaltigere Batteriezellenfertigung durch den Einsatz von intelligenten Algorithmen zu schaffen und somit die deutsche Batteriezellenproduktion wettbewerbsfähig zu halten.

Ich wünsche Ihnen eine spannende Lektüre.

Prof. Dr. Maximilian Röglinger

Universität Bayreuth

Lehrstuhl für Wirtschaftsinformatik und

Wertorientiertes Prozessmanagement

Leitfaden Prädiktive Prozessüberwachung Ein Ergebnis des KIProBatt Projekts

Mit der zunehmenden Dynamik der Energiewende gewinnt der Ausbau der Batterieindustrie in Europa zunehmend an Bedeutung. Vorausschauende Prozessüberwachungsmethoden können die Optimierung der Batteriezellproduktion unterstützen, indem Effizienzen im Sinne von Zeit, Kosten und Nachhaltigkeit realisiert werden.

Aufgrund seiner Komplexität und seiner enormen wirtschaftlichen und ökologischen Auswirkungen wird der Produktionsprozess von Lithium-Ionen-Batteriezellen kontinuierlich digitalisiert und optimiert, um die Leistung der Zellen zu erhöhen und gleichzeitig den Ressourcenverbrauch und die Produktionskosten zu senken.

In diesem Zusammenhang birgt die künstliche Intelligenz (KI) ein immenses Potenzial für die Nutzung von Produktionsdaten zur Verbesserung des Zellproduktionsprozesses. Ziel des Projektes „KIProBatt - Intelligente Batteriezellfertigung mit KI-gestütztem Prozessmonitoring auf Basis einer generischen Systemarchitektur“ ist es daher, die Zellproduktion durch eine KI-basierte End-to-End-Prozessüberwachung zu verbessern. Zu diesem Zweck wurde eine generische Systemarchitektur als wiederverwendbare Struktur entwickelt, die die Prozessdatenerfassung mit einem ontologiebasierten semantischen Datenraum zu verbinden. Basierend auf dieser Systemarchitektur setzen Ansätze des maschinellen Lernens aus zwei Perspektiven auf: Prädiktive KI, die die Qualität von Batteriezellen während der Produktion abschätzen kann und eine übergreifende End-to-End-Prozessüberwachung mit präskriptiver Entscheidungsunterstützung.

Die Kombination dieser beiden Perspektiven ermöglicht es, Fehler frühzeitig im Produktionsprozess zu erkennen, die Qualitätsleistung zu steigern und flexible Anpassungen der Prozessparameter im Falle von Fehlfunktionen oder Defekten vorzunehmen. Auf diese Weise erwarten wir eine Senkung der Gesamtkosten der Zellproduktion und eine Verbesserung der CO₂-Bilanz durch eine Verringerung des Ressourcen- und Energieverbrauchs.

Einleitung

In der zunehmenden Dynamik der Energiewende gewinnt der Ausbau der Batteriezellindustrie in Europa an Bedeutung [1]. Durch den Einsatz von Digitalisierung und Datenanalyse kann Europa seine Position in der Batteriezellproduktion weiter ausbauen, indem Prozesseffizienz gesteigert und Ausschuss reduziert wird. Gleichzeitig bildet die Verpflichtung zu Nachhaltigkeit und Umweltschutz eine strategische Grundlage, welche europäische Batterieprodukte auf dem Weltmarkt differenziert.

In der Ära von Industrie 5.0 gewinnt die Verbindung von technologischem Fortschritt mit Nachhaltigkeit zunehmend an Bedeutung [2]. Ein Ansatz ist hierbei durch präventive Analytik und Prozessoptimierung im Zero-Defect Manufacturing den Materialausschuss und die Ressourcenverschwendung zu minimieren [3]. Unterstützt wird dieses Vorgehen durch Technologien wie IoT, Big Data, Cloud-Computing und KI [4, 5].

Vor dem Hintergrund der essenziellen Bedeutung von Digitalisierung und Datenanalyse für die Konkurrenzfähigkeit des Batteriesektors zeichnet sich die prädiktive Prozessüberwachung nicht nur als logischer nächster Schritt ab, sondern auch als Schlüsselstrategie zur Förderung einer nachhaltigeren Produktion. Kern dieser Technologie ist die automatisierte Erfassung von Prozessdaten durch eine Vielzahl von Sensoren, welches eine Echtzeitorientierung von Produktionsprozessen erlaubt. Hierzu gehören Kamerasysteme, Geräte zur Messung von Druck und Temperatur sowie Sensoren zur Erfassung von Formdaten. Diese Daten werden in einem semantisch strukturierten Datenraum zusammengeführt, um eine effiziente und effektive Entscheidungsunterstützung zu ermöglichen.

Durch die Verknüpfung dieser diversen Datentypen in einem integrierten System können Muster und Trends erkannt werden, die für die Prozesssteuerung entscheidend sind. Dies ermöglicht es wiederum, potenzielle Probleme frühzeitig zu identifizieren und präventive Maßnahmen zu ergreifen, um die Effizienz und Qualität zu steigern.

Der Einfluss verschiedener Parameter auf die Qualität von Batterien ist bereits Bestand wissenschaftlicher Untersuchungen [6–10]. So wurden unter anderem Techniken aus dem Bereich Computervision eingesetzt, um Defekte sowie aufgequollene Batterien frühzeitig zu identifizieren [11, 12]. Jedoch stellt eine systematische prädiktive Überwachung des Produktionsprozesses von Batterien in Echtzeit ein weitgehend unerschlossenes Feld dar. Ziel des Vorhabens soll es sein, diese Optimierungspotentiale zu erschließen und nutzbar zu machen.

Die Produktion von Batteriezellen kennzeichnet sich durch eine komplexe Prozesskette, in der unbekannte Abhängigkeiten zwischen verschiedenen Faktoren bestehen. Diese unbekannten Abhängigkeiten zwischen Prozessparametern, Zwischenprodukteigenschaften und Qualitätsmerkmalen führen zu hohen Ausschussraten und einem immensen Aufwand für die Qualitätskontrolle [8]. Diese Problematik wird weiter verstärkt durch die langwierigen Formierungs- und Alterungsprozesse, die am Ende der Produktion anfallen und einen signifikanten Beitrag zu den Produktionskosten leisten [13].

Hinzu kommt die hohe Empfindlichkeit des Rohprodukts gegenüber Umgebungsbedingungen [10]. Diese Sensibilität ist ein weiterer Faktor, der die Industrie mit hohen Ausschussraten konfrontiert. Angesichts dieser Herausforderungen sind die Erforschung und Anwendung datengetriebener Ansätze von entscheidender Bedeutung.

Erste Schätzungen zeigen bereits das enorme Potential der prädiktiven Prozessüberwachung von Batteriezellen auf. In einem zukunftsorientierten Optimierungsszenario lässt sich die Ausschussrate von 5% auf 1% senken [14]. Diese Verringerung führe zu einer deutlichen Senkung der Produktionskosten, was insbesondere auf den hohen Kostenanteil des letzten Produktionsschrittes zurückzuführen ist. Durch frühzeitige Identifikation potenzieller Qualitätsprobleme können Abläufe abgebrochen oder angepasst werden. Dadurch wird nicht nur Material gespart, sondern auch die Effizienz des gesamten Produktionsprozesses verbessert.

Grundlagen der Batteriezellserienproduktion

Die Herstellung von Batteriezellen in Serie ist ein hochkomplexer und technologisch fortschrittlicher Prozess. Dieser Prozess umfasst mehrere Schlüsselschritte, die zusammenwirken, um leistungsfähige und zuverlässige Batteriezellen für moderne Anwendungen, insbesondere im Bereich der Elektromobilität, zu produzieren. Von der sorgfältigen Herstellung der Elektroden über die präzise Montage der Zellen bis hin zur entscheidenden Konditionierung der Zellen – jeder Schritt ist entscheidend für die Qualität und Leistung des Endprodukts.

1 | Elektrodenproduktion

Der erste Schritt in der Herstellung von Batteriezellen befasst sich mit der Produktion der Elektroden. Diese Phase hat bereits entscheidenden Einfluss auf den späteren Zustand des Endproduktes, da die Elektroden die zentralen Komponenten jeder Batteriezelle sind. Sie umfasst die Verarbeitung verschiedener Materialien zu einer Schlamm-Mischung, die als aktive Schicht auf Metallträgerfolien aufgetragen wird.

Bei der Elektrodenproduktion werden aktive Materialien (für Anoden meist Graphit, für Kathoden verschiedene Verbindungen wie Lithium-Nickel-Mangan-Cobaltoxid), Bindemittel und leitfähige Zusätze gemischt. Diese Mischung wird in einem mit Lösungsmitteln kombiniert, um eine pastöse Konsistenz zu erreichen. Anschließend wird dieser Schlamm auf die Metallträgerfolien (Kupfer für Anoden, Aluminium für Kathoden) aufgetragen. Die präzise Auftragung und gleichmäßige Dicke der Elektroden-schichten sind für die Leistungsfähigkeit und Lebensdauer der Batteriezellen von entscheidender Bedeutung.

Es folgt mit dem Kalandrieren ein Prozess, bei dem die Dicke der Elektrode eingestellt wird, was wiederum die Zelleigenschaften wie Energie- und Leistungsdichte beeinflusst. Nachdem die Elektroden beim auf die gewünschten Maße zugeschnitten worden sind (**Separation**), wird beim Trocknen die Feuchtigkeit weiter reduziert (**Drying**).

2 | Zellproduktion

Der zweite Schritt im Prozess der Batteriezellproduktion konzentriert sich auf die Zellproduktion selbst. In dieser Phase werden die zuvor hergestellten Anoden und Kathoden zur eigentlichen Batteriezelle integriert.

Anoden und Kathoden werden zunächst in kleine Blätter geschnitten. Darauf wird ein Separator, welcher wie ein „Z“ gefaltet ist, hinzugenommen und die Anoden und Kathoden jeweils im Wechsel in die Zwischenräume eingeführt. Diese Separatoren verhindern den direkten Kontakt zwischen Anode und Kathode, der zu Kurzschlüssen führen könnte, und ermöglichen gleichzeitig den Ionenfluss innerhalb der Zelle. Anschließend werden bei der Kontaktierung interne Kontakte zwischen Anode, Kathode und Separator geschweißt. Ein weiterer wichtiger Bestandteil dieses Prozessschritts (**Stacking**) ist das Schweißen von Kontakten und das Einbringen der zusammengesetzten Zelle in ein Gehäuse. Das präzise Schweißen stellt eine sichere und dauerhafte Verbindung der Zellkomponenten her. Nachdem die Baugruppe in das Zellgehäuse eingesetzt wird, folgt das Befüllen der umschlossenen Baugruppe mit einem Elektrolyten (**Packaging & Filling**). Der Elektrolyt übernimmt in der fertigen Batteriezelle den Transport der Ionen zwischen Anode und Kathode, was im Gegenzug den Elektronen ihre Wanderschaft in Gegenrichtung ermöglicht, die für den Stromfluss sorgt. Der Elektrolyt ist einer der Treiber für die Materialkosten [15].

3 | Zellkonditionierung

Der dritte und abschließende Schritt in der serienmäßigen Herstellung von Batteriezellen, ist die Zellkonditionierung. In der Phase der Zellkonditionierung durchläuft jede Zelle zunächst eine **Formierung**. Dieser Prozess beschreibt das erstmalige Laden und Entladen der Zelle, um die elektrochemischen Eigenschaften zu aktivieren und zu stabilisieren. Darauf folgen weitere Ladezyklen. Dieser Prozess dauert mehrere Tage und trägt

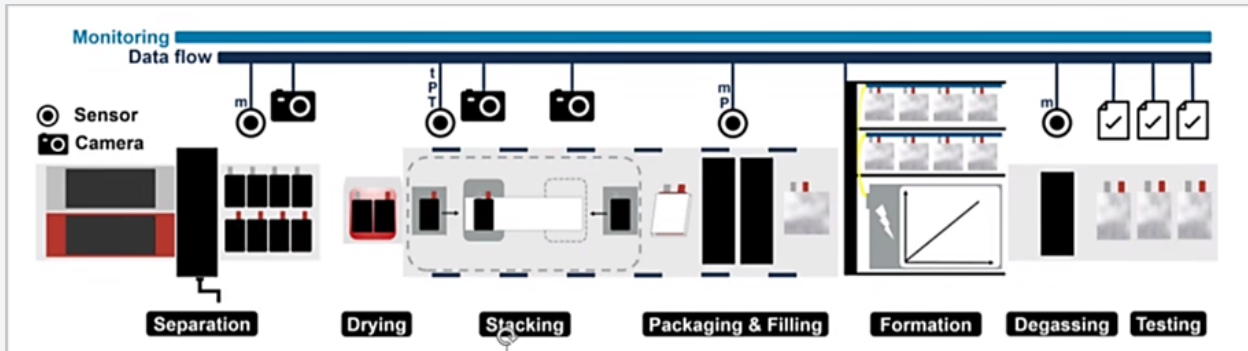


Abb. 1 Schematischer Ablauf der Batteriezellproduktion.

entscheidend zur Leistung und Sicherheit der Zelle bei, da dabei eine Schutzschicht zwischen Anode und Elektrolyt aufgebaut wird. Eine sorgfältige Formation ist essenziell, um die volle Kapazität und Leistung der Batteriezelle zu erreichen und sicherzustellen, dass die Zelle die erwarteten Leistungsstandards erfüllt.

Es folgt die Versiegelung der Zelle. Zuvor muss jedoch das bei der Formierung entstehende Gas durch Druck von aus der Zelle gedrückt werden (**Degassing**). Für Qualitätsmessungen (**Testing**) wird im nächsten Schritt die Alterung der Zellen simuliert, wobei die Zellen unter kontrollierten Bedingungen mehrere Wochen gelagert werden. Hierbei werden Tests durchgeführt, die Defekte wie beispielsweise Kurzschlüsse identifizieren. Vor dem Abschluss folgt eine finale Kontrolle.

Unterstützende Aktivitäten

Unterstützend zu den Schritten der Konditionierung sind verschiedene Prozesse notwendig. Zum einen muss durch Trockenraummanagement darauf geachtet werden, dass eine reine Atmosphäre gewährleistet wird. Zum anderen muss das in der Kathodenproduktion verwendete Lösungsmittel recycelt werden. Hinzu kommen weitere operative Prozesse wie die korrekte Handhabung der Materialien durch sämtliche Produktionsschritte hinweg, sowie Logistik.

Herausforderungen in der Batteriezellproduktion

In der hochmodernen und schnell wachsenden Industrie der Batteriezellfertigung stehen Hersteller vor heute vor mannigfaltigen Herausforderungen, die von der Ressourceneffizienz bis hin zur

Anlagenauslastung reichen. Eine Schlüsselherausforderung besteht darin, die Effizienz des Einsatzes kritischer Materialien zu optimieren und die Recyclingprozesse zu verbessern, um die Nachhaltigkeit der Produktion zu erhöhen. Zeitliche Aspekte sind ebenso entscheidend, da die Reduktion der Herstellungszeit pro Zelle direkt die Fähigkeit beeinflusst, die wachsende Nachfrage zu bedienen und die Wettbewerbsfähigkeit zu sichern.

Die Effizienz der Produktionsanlagen muss durch kontinuierliche technologische Innovationen und Prozessoptimierungen gesteigert werden. Dies umfasst eine effektivere Auslastung der Maschinen und eine fortschrittliche Automatisierung. Das Personal spielt dabei eine zentrale Rolle; ein tiefes Verständnis der komplexen Fertigungsabläufe ist unerlässlich, um Qualität und Effizienz zu garantieren. Energieeffizienz ist ebenfalls ein kritischer Faktor, da die Produktionskosten stark von den Energiepreisen abhängen.

Eine maximale Auslastung der Produktionsanlagen trägt entscheidend zur Kostensenkung bei und erfordert eine ausgeklügelte Logistik und durchdachte Produktionsplanung. Eine integrierte Betrachtung aller Faktoren ist nötig, um eine nachhaltige und wirtschaftlich erfolgreiche Produktion zu gewährleisten.

Eine neue Ära der Industrialisierung

In einer Zeit des schnellen Wandels und der globalen Herausforderungen stellt die fortschreitende Digitalisierung im Rahmen der Industrie 4.0 die Grundlage für eine transformative Entwicklung der europäischen Wirtschaft dar, die fragmentarische Arbeitsschritte als Gesamtprozess besser analysierbar macht. Daten und Informationen werden zum treibenden Herzstück dieser Revolution, wodurch eine neue Ära der Effizienz und Innovation eingeleitet wird. Über die technologischen Fortschritte hinaus rückt die Industrie 5.0 Mensch und Umwelt in den Fokus des Handelns.

Die Wertschätzung und Weiterentwicklung der menschlichen Arbeitskraft, gekoppelt mit einem starken Fokus auf die fortlaufende Schulung und Anpassungsfähigkeit der Belegschaft, kennzeichnen diesen neuen Ansatz. Gleichzeitig wird Nachhaltigkeit nicht länger als bloße Nebenbedingung, sondern als essenzielle Grundlage für zukunftsfähige Geschäftsmodelle verstanden. Diese ganzheitliche Sichtweise, die sowohl die digitale Transformation als auch die sozialen und ökologischen Aspekte umfasst, bildet das Fundament für eine Industrie, die nicht nur wirtschaftlich prosperiert, sondern auch zur Lösung drängender globaler Probleme wie dem Klimawandel beiträgt und somit einen nachhaltigen Wohlstand für die Gesellschaft sichert.

Um die Ziele einer nachhaltigen und intelligenten Industrie zu erreichen, gilt es, vielfältige Herausforderungen zu bewältigen. Diese umfassen die Implementierung von Big Data-getriebenen Systemen und die effiziente Erfassung und Verwaltung industrieller Daten, die Optimierung der Arbeitsteilung zwischen Menschen und intelligenten Maschinen sowie das Adressieren von Sicherheits-, Datenschutz- und ethischen Fragen. Ebenso ist der Aufbau eines umfassenden Systems zur Vernetzung aller relevanten Daten für ein dezentrales Management sowie die Entwicklung einer "intelligenten Kette" zur Verknüpfung von Produktwissen und Entscheidungsfindung erforderlich. Hinzu kommen die Notwendigkeit qualifizierter Arbeitskräfte, hohe Kapitalinvestitionen, der Bedarf an fortschrittlichen Technologien für Datenmanagement und der menschliche Widerstand gegen den Einsatz von Robotern. Weiterhin ist die

Entwicklung effektiver Modelle zur Bewertung von Fertigungselementen, der Aufbau einer umfassenden Plattform für den gesamten Produktlebenszyklus, die Festlegung von Optimierungsrichtlinien und global vereinbarten Nachhaltigkeitsbewertungsmetriken notwendig. Schließlich erfordert die Integration eines nachhaltigen und intelligenten Fertigungssystems die Entwicklung datengesteuerter Algorithmen, eine klare Analyse der Beziehung zwischen Nachhaltigkeitsebenen und Industrie 4.0-Elementen, die Rechtfertigung der Systemkomplexität und die Überwindung von Widerstand gegen drastische Veränderungen für eine soziale Akzeptanz [5].

Ein zentraler Ansatz zur Förderung der Nachhaltigkeit und Effizienz in der modernen Industrie ist die Implementierung von präventiver Analytik und Prozessoptimierung im Rahmen des Zero-Defect Manufacturing. Dieses Konzept zielt darauf ab, den Materialausschuss und die Ressourcenverschwendung zu minimieren, indem Fehler und Mängel im Produktionsprozess frühzeitig erkannt und eliminiert werden. Durch die Nutzung fortschrittlicher Datenanalyse- und Überwachungstechnologien können Unternehmen potenzielle Probleme identifizieren, bevor sie zu kostspieligen Ausschüssen führen. Diese proaktive Herangehensweise trägt nicht nur dazu bei, die Effizienz zu steigern und die Produktionskosten zu senken, sondern unterstützt auch eine nachhaltigere Fertigungspraxis, die den ökologischen Fußabdruck der Industrie deutlich reduziert. Somit stellt die Integration von präventiver Analytik in die Produktionsprozesse einen wichtigen Schritt in Richtung einer umweltbewussteren und ressourcenschonenden Industrieproduktion dar [3].

Fünf Ebenen der Digitalisierung

Um die fortschreitende Digitalisierung im produzierenden Gewerbe im Allgemeinen und der Batteriezellproduktion im Speziellen zu strukturieren, dienen sich fünf Ebenen eines Enterprise Architecture Modells an. Die fünf Ebenen umfassen das Geschäftsmodell, Prozesse, Menschen und Applikationen, Daten und Informationen und die technische Infrastruktur. Für die präskriptive Prozessverbesserung, die Erkenntnisse aus der Kombination von realweltlichen (physischen), abgeleiteten virtuellen (cybervirtuellen) Eindrücken und dem Domänenwissen im Prozess gewinnt, heben wir insbesondere diese drei Teile hervor.

Geschäftsmodell

Unter Berücksichtigung der Turbulenzen von Geschäftsumgebungen stellt es für Unternehmen eine zentrale Herausforderung dar, nachhaltige Chancen, die langfristig erfolgsversprechend sind, von kurzfristigen Trends abzugrenzen. Vor diesem Hintergrund ist das Geschäftsmodell eines Unternehmens von größter Bedeutung, da es das Ausnutzen bestehender Marktpotenziale und die Erschließung neuer Möglichkeiten ermöglicht. Geschäftsmodelle legen Zielmärkte, Betriebsmodelle sowie Kosten- und Umsatzströme fest. Dies beinhaltet auch die Wertversprechen des Unternehmens, die beschreiben, welche Kundenbedürfnisse durch welche Produkt- und Dienstleistungsangebote befriedigt werden. Unter Berücksichtigung von Faktoren wie Effizienz, Umweltauswirkungen, Sicherheit und Innovation, sind auch die Produktionsprozesse in dem Geschäftsmodell eines Unternehmens zu berücksichtigen. Bei der Modellierung eines Geschäftsmodells in der digitalen Ära, spielen diese Aspekte eine bedeutende Rolle, um den langfristigen Erfolg eines Unternehmens sicherzustellen und kurzfristigen Trends zu widerstehen.

Prozess

Um Geschäftsmodelle Wirklichkeit werden zu lassen, sind Unternehmen auf Arbeitsroutinen angewiesen, die funktionsübergreifend strukturiert sind und sich um Prozesse drehen. In der digitalen Zeit

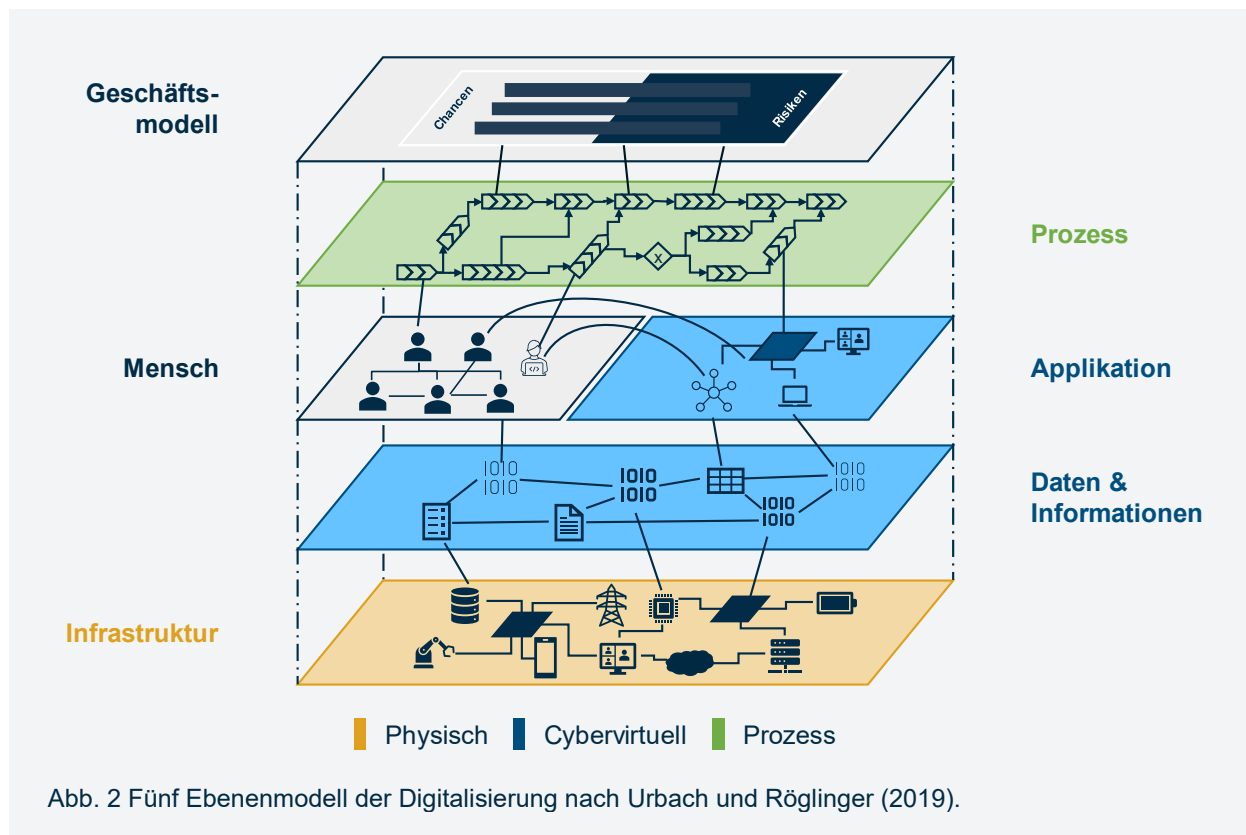
sollte Prozessdenken nicht nur abteilungsübergreifend, sondern auch organisationsübergreifend sein und gesamte Wertschöpfungsketten und Ökosysteme umfassen. Prozesse legen somit fest, welche Aufgaben durchgeführt werden müssen, um bestimmte Ziele zu erreichen oder Produkte zu erstellen. Neuartige Methoden und Kenntnisse aus dem Bereich des Geschäftsprozessmanagements (BPM) können durch steigender Vernetzung, Komplexität und Datenaufkommen zunehmend auf Produktionsprozesse übertragen werden [16].

Menschen und Applikationen

Die in Prozessen enthaltenen Aufgaben können entweder manuell von Mitarbeitern, automatisch von Maschinen und Anwendungssystemen oder in Zusammenarbeit erledigt werden. In diesem Sinne sind Menschen Teil der Organisationsstruktur, die Rollen, Verantwortlichkeiten und Berichtslinien festlegt. Technologien wie Mensch-Maschine-Interaktion, künstliche Intelligenz, Smart Devices und Robotik fördern insbesondere die kollaborative Aufgabenausführung. Da viele dieser Technologien auch die Grenzen der Automatisierung erweitern, können nicht nur gut strukturierte, sondern auch unstrukturierte Aufgaben automatisiert werden.

Daten & Information

Mitarbeiter, Anwendungssysteme und Maschinen erzeugen und bearbeiten Daten und Informationen. Parallel zur steigenden Nutzung digitaler Technologien wächst auch die Menge der verfügbaren Daten rapide und offenbart neues Potenzial für Wissensgewinnung. Strukturierte Daten, wie z.B. Tabellen oder relationale Datenbanken, können nach wie vor durch statistische Analysemethoden ausgewertet werden. Zudem ermöglichen moderne Algorithmen, die auf Fortschritten in der künstlichen Intelligenz basieren (z.B. kognitives Computing oder Deep Learning), eine immer genauere Verarbeitung von unstrukturierten Daten wie Texten, Grafiken, Videos und Audiodateien. Big Data-Analysen erlauben es, große Mengen an Daten aus unterschiedlichen Quellen zu analysieren und zu kombinieren. Dies ermöglicht Organisationen, bessere Entscheidungen zu treffen.



Technische Infrastruktur

Um das mit der Digitalisierung verbundene Potenzial auszuschöpfen, benötigen Unternehmen eine angemessene technologische Infrastruktur. Diese umfasst neben traditionellen Komponenten wie PC, Servern, Netzwerk- und Sicherheitskomponenten auch neuartige Elemente wie cyber-physische vernetzte Produktionsumgebungen und gemeinsam genutzte Ressourcen wie Cloud- und Edgeinfrastrukturen. Im Zeitalter der Digitalisierung wird die herkömmliche Informations- und Kommunikationsinfrastruktur zunehmend mit der Infrastruktur für die Produktion (Operations Technology) integriert, um eine Brücke zwischen der physischen und der digitalen Welt zu bauen.

Disruptionen durch Technologie und das Geschäftsumfeld

Unternehmen sehen sich der Herausforderung gegenüber, rechtzeitig strategische Entscheidungen über den Einsatz *disruptiver Technologien* zu treffen. Im Aktionsfeld der digitalen Disruption geht es darum, aufkommende und sich entwickelnde Technologien zu überwachen und zu analysieren, um Unsicherheiten bei der Technologieauswahl zu verringern.

Ein weiterer wichtiger Aspekt ist das Sammeln von Messwerten der Batteriezeleigenschaften. Dieses Vorgehen dient zwei wesentlichen Zielen: zum einen dem Tracing, also der lückenlosen Dokumentation und Nachverfolgung der einzelnen Zellen durch die Produktionskette, zum anderen der fortwährenden Qualitätsverbesserung. Hierbei geht es nicht nur um die Optimierung der Zellqualität selbst, sondern auch um die Verbesserung der Produktionsprozesse. Die gewonnenen Daten können genutzt werden, um Schwachstellen zu identifizieren und Prozesse effizienter zu gestalten.

Darüber hinaus spielt das kontinuierliche Sammeln und Überwachen von Daten der Produktionsanlagen eine entscheidende Rolle. Hierbei werden Maschinenkennzahlen und der Produktionsprozess als Ganzes erfasst. Dieses Vorgehen ist eng mit dem Konzept der Gesamtanlageneffektivität (Overall Equipment Efficiency, OEE) verknüpft. Durch die kontinuierliche Überwachung und Analyse dieser Daten können prädiktive Wartungsmaßnahmen (Predictive Maintenance) eingeleitet werden, um Ausfallzeiten zu minimieren und die Gesamteffizienz von Anlagen zu maximieren.

Daten und Datentypen

Im Rahmen der Produktionsdatenanalyse spielt die Kategorisierung von Daten eine entscheidende Rolle, um Prozesse zu optimieren, Qualität zu sichern und effiziente Entscheidungen zu treffen. Daten können nach verschiedenen Ansätzen klassifiziert werden, darunter die Art der Daten, ihr Ursprung, der Lebenszyklus und weitere spezifische Kriterien. Diese Kategorisierung ermöglicht es, die Vielfalt an Informationen systematisch zu erfassen, zu analysieren und zu nutzen. Im Folgenden werden drei zentrale Kategorien von Produktionsdaten hervorgehoben, die sich nach ihrem Ursprung und ihrer Dynamik unterscheiden und jeweils spezifische Zwecke innerhalb des Produktionsprozesses erfüllen.

Grundlegende Produktionsparameter (Statische Planungsdaten)

Diese Kategorie beinhaltet detaillierte Informationen über den geplanten Produktionsprozess, bevor dieser startet. Dazu zählen technische Zeichnungen, Materialanforderungen, und Prozessablaufpläne. Diese Daten stammen aus der Planungsphase vor Produktionsbeginn, dienen als Blaupause für die Produktionslinie und legen fest, welche Ressourcen benötigt werden, wie diese zu kombinieren sind und in welcher Reihenfolge die Produktionsschritte durchzuführen sind. Sie sind weitestgehend statisch, da sie als feste Vorgaben für den Produktionsprozess dienen. Sie ermöglichen eine effiziente Vorbereitung und Einrichtung der Produktionsanlagen und stellen sicher, dass alle notwendigen Materialien und Informationen verfügbar sind, um den Produktionsprozess gemäß den Spezifikationen zu starten.

Prozessüberwachungsdaten (Dynamische Produktionsdaten)

Während der Produktion werden kontinuierlich Daten erhoben, um sicherzustellen, dass die Produkte den Qualitätsstandards entsprechen und die Maschinen einwandfrei laufen. Diese umfassen physikalische, chemische oder mechanische Messungen, wie z.B. die Überwachung von Temperaturen, Drücken oder die Prüfung der Materialfestigkeit. Die rechtzeitige Erfassung und Analyse

dieser Daten ermöglicht es, Probleme frühzeitig zu erkennen und Korrekturmaßnahmen einzuleiten, um Ausschuss zu minimieren und die Produktqualität zu sichern. Sie sind hochdynamisch, da sie in Echtzeit während der Produktion erfasst werden und variieren können, abhängig von den aktuellen Produktionsbedingungen.

Entscheidungsunterstützende Daten (Analytische Optimierungsdaten)

Diese Datenkategorie bildet die Grundlage für die Steuerung und kontinuierliche Verbesserung des Produktionsprozesses. Daten entstehen hierbei durch die Analyse und Aggregation von Produktions- und Planungsdaten, einschließlich In-Instanz-Daten und Meta-Produktionsebene-Daten.

In-Instanz-Daten (Operative Optimierungsdaten)

Diese beziehen sich auf Daten, die für sofortige Anpassungen innerhalb einer laufenden Produktionsinstanz genutzt werden. Sie ermöglichen operative Entscheidungen, wie die Feinjustierung von Maschineneinstellungen oder die kurzfristige Umplanung von Produktionsabläufen, um auf aktuelle Herausforderungen zu reagieren.

Meta-Produktionsebene-Daten (Strategische Optimierungsdaten)

Diese Daten werden aus einer Vielzahl von Produktionsprozessen aggregiert und analysiert, um tiefere Einblicke in die Prozesseffizienz, Qualitätsmuster und mögliche Optimierungsbereiche zu gewinnen. Sie ermöglichen systematische Verbesserungen durch die Identifizierung von Trends, die Optimierung von Produktionsabläufen und die strategische Planung zukünftiger Produktionszyklen. Durch das Sammeln und Analysieren dieser Daten über verschiedene Prozesse und Zeiträume hinweg können Unternehmen langfristige Verbesserungen in ihrer Produktionseffizienz und Produktqualität erzielen.

Das Problem fehlender und fehlerhafter Daten

Um eine ganzheitliche Strategie für den Umgang mit fehlenden Daten zu umsetzen zu können, ist es essenziell, sowohl proaktive Maßnahmen zur Vermeidung und Behebung von Datenlücken als auch effektive Data-Imputation-Techniken zu kombinieren. Datenimputation umfasst Methoden, die fehlende Daten auf Basis von bestmöglichen Schätzungen konstruieren.

Die folgenden Beschreibungen geben an, für welche Art von fehlenden Daten sich verschiedene Ansätze zur Datenimputation eignen und skizzieren im nächsten Abschnitt das dahinterliegende Rational zur künstlichen Nacherhebung der Daten. Grundsätzlich lassen sich drei verschiedene Kategorien fehlender Daten nennen: Missing Completely at Random (MCAR), Missing at Random (MAR), Missing Not at Random (MNAR).

Datenverlust aufgrund unvorhersehbarer Ereignisse

Regelmäßig können Daten vollkommen willkürlich fehlen (MCAR). Hierbei stehen die fehlenden Datenpunkte in keiner systematischen Beziehung zu anderen erfassten Variablen. Typische Szenarien sind der Verlust von Daten auf dem Übertragungsweg oder der Ausfall der Datenerfassung. Dies kann beispielsweise durch ein bei der Übertragung verlorengegangenes Datenpaket oder einen durch einen Stromausfall bedingten Sensorausfall geschehen.

Bei MCAR sollte eine einfache Datenimputationsmethode gewählt werden, um die Analyse möglichst nicht zu verfälschen. Fehlende Datenpunkte können durch Lageparameter historischer Datenwerte derselben Variable ersetzt werden. Es bietet sich bei metrischen Variablen der Mittelwert, bei nicht-metrischen der Modus beziehungsweise der Median an.

Fehlende Werte in Abhängigkeit von Beobachtungen

In der Praxis tritt häufiger der Fall auf, dass fehlende Daten dem Muster "Missing at Random" (MAR) folgen. Dabei besteht ein systematischer Zusammenhang zwischen den fehlenden Werten

und den beobachteten Daten, jedoch nicht zu den nicht beobachteten Werten selbst. Ein Anwendungsbeispiel ist die Kapazitätsprüfung von neu gefertigten Batteriezellen: Die Zellen durchlaufen einen Lade- und Entladezyklus, wobei ihre Kapazität erfasst wird. Bei als kritisch eingestuften Zellen mit sehr geringer gemessener Kapazität wird oftmals auf weiterführende Analysen wie die Messung des Innenwiderstands verzichtet. Folglich hängen die fehlenden Innenwiderstandswerte systematisch mit der beobachteten niedrigen Kapazität zusammen, jedoch nicht mit dem nicht gemessenen Innenwiderstand selbst.

Bei einem MAR-Fehler sind einfache Imputationsmethoden wie Mittelwert- oder Modusersetzung nicht ausreichend, da sie die vorhandenen Zusammenhänge zwischen Variablen nicht berücksichtigen. Die Wahl der Imputationsmethode sollte auf dem systematischen Zusammenhang beruhen: Bei linearem Zusammenhang bietet sich die Imputation mittels Linearer Regression an. Zum Beispiel lässt sich annehmen, dass bei steigender Kapazität der Innenwiderstand sinkt. In anderen Fällen kann Hot Deck Imputation oder Multivariate Imputation by Chain Equations sinnvoll sein.

Fehlende Werte aufgrund nicht erfasster Daten

Der dritte Hauptmechanismus für das Auftreten fehlender Werte ist "Missing Not at Random" (MNAR). Dabei hängen die fehlenden Werte nicht nur von den beobachteten Daten ab, sondern auch von den eigentlich fehlenden, unbeobachteten Werten selbst. In einer Batteriezellproduktionslinie werden beispielsweise Daten zu Kapazität, Lade- und Entladerate erhoben sowie der Innenwiderstand gemessen. Allerdings schlagen Innenwiderstandsmessungen häufig bei Zellen mit sehr hohen Innenwiderständen fehl. Die Wahrscheinlichkeit eines fehlenden Innenwiderstandswertes hängt direkt vom eigentlichen, aber fehlenden Innenwiderstandswert selbst ab - je höher der Wert, desto eher fällt die Messung aus.

Um MNAR-Muster zu berücksichtigen, sind oft zusätzliche Modellannahmen oder sogar die Erhebung von Zusatzinformationen über die Ausfallgründe nötig.

Imputation fehlender Datenwerte

Eine Imputation ermöglicht es, fehlende Daten in einer statistischen Erhebung zu ergänzen und somit für Analysen nutzbar zu machen, die vollständige Datensätze erfordern. Allerdings muss dabei berücksichtigt werden, dass Imputationsverfahren zu systematischen Verzerrungen führen können. Daher ist es ratsam, sich zuerst mit den Ursachen für das Fehlen von Daten auseinanderzusetzen und diesen an der Quelle entgegenzuwirken. Dazu gehört auch, zu differenzieren zwischen einem unsystematischen Ausfall von Daten und einem Fehlen, das im Zusammenhang mit anderen erfassten Variablen oder der eigentlich zu messenden Variable selbst steht. Nur wenn die Mechanismen der fehlenden Werte bekannt sind, lassen sich geeignete Imputationsmethoden anwenden und potenzielle Verzerrungen vermeiden oder korrigieren. Eine sorgfältige Analyse der Ausfallgründe ist daher unerlässlich für den angemessenen Umgang mit fehlenden Daten.

Schätzung aus Lageparametern

Die Imputation fehlender Werte durch zentrale Lageparameter der beobachteten Werte derselben Variablen ist ein einfaches und schnell anzuwendendes Verfahren. Allerdings führt diese Methode zu einer Unterschätzung der Varianz und Kovarianz, sodass statistische Tests und Konfidenzintervalle verzerrt werden. Zudem erfordert sie die Annahme eines MCAR-Mechanismus, um unverzerrte Schätzungen zu liefern und kann zu unplausiblen imputierten Werten führen.

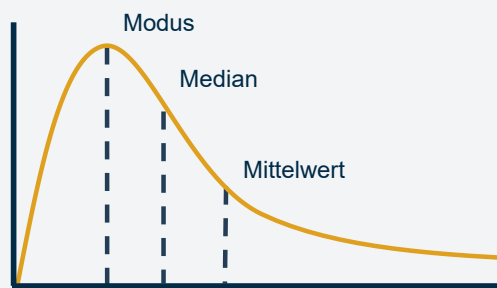


Abb. 3 Imputation durch Schätzung aus Lageparametern.

Imputation mittels Linearer Regression

Die Regressionsimputation nutzt die Zusammenhänge zwischen Variablen, indem zunächst eine Regression der teilweise beobachteten Variable auf vollständig erfasste Variablen für die vorhandenen Datenpunkte geschätzt wird. Anschließend werden die vollständig beobachteten Variablenwerte in das Regressionsmodell eingesetzt, um Schätzwerte für die fehlenden Werte der Zielvariablen zu erhalten. Verglichen mit der Imputation durch Lageparameter, ist für diese Methode lediglich ein MAR-Mechanismus erforderlich, um unverzerrte Schätzungen zu erhalten. Jedoch unterschätzt auch dieses Verfahren die Varianz und Kovarianz der imputierten Werte.

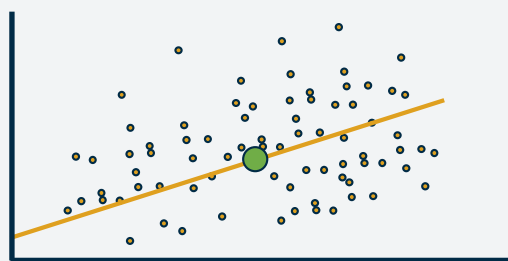


Abb. 4 Imputation via Regression.

Imputation aus ähnlichen Datensätzen

Die Hot Deck Imputation ersetzt fehlende Werte durch beobachtete Werte ähnlicher Fälle oder Individuen. Die Ähnlichkeit wird über Distanzmetriken bemessen. Für einen fehlenden Wert wird der ähnlichste, vollständige Fall ermittelt und dessen Wert imputiert. Der Vorteil ist, dass nur plausible Werte eingesetzt werden und die Methode robust gegenüber Modellfehlspezifikationen ist. Jedoch hängt die Imputationsqualität von der Metrik ab. Varianz sowie Kovarianz werden tendenziell unterschätzt, da zusätzliche Unsicherheit nicht berücksichtigt wird. Die Annahmen für unverzerrte Schätzungen (MCAR/MAR) variieren je nach Verfahrensvariante.

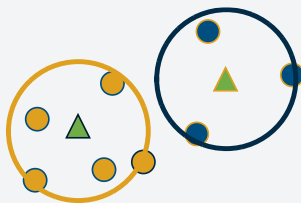


Abb. 5 Imputation aus ähnlichen Datensätzen.

Last-Observation-Carried-Forward (LOCF)

Die Last Observation Carried Forward (LOCF) Methode ist ein Verfahren zum Umgang mit fehlenden Werten in Längsschnittdatensätzen mit mehreren Messzeitpunkten. Dabei werden fehlende Werte eines bestimmten Zeitpunkts durch den letzten vorhandenen Messwert derselben Beobachtungseinheit ersetzt. So werden beispielsweise die fehlenden Daten zum dritten Zeitpunkt mit den Daten des zweiten Zeitpunkts aufgefüllt. Dieses Vorgehen ist nur sinnvoll, wenn über die Zeit keine Veränderung der Werte zu erwarten ist.

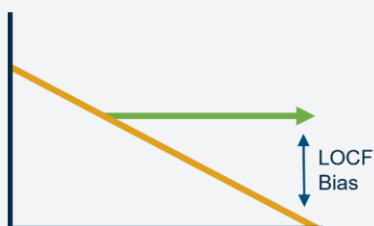


Abb. 6 Imputation mittels LOCF

Multiple Imputation by Chained Equations

Der MICE-Algorithmus ist eine fortschrittliche Technik zur Behandlung fehlender Daten in statistischen Analysen, die durch iteratives Vorgehen charakterisiert wird. Für jede Variable mit fehlenden Informationen wird ein spezifisches Modell auf Basis der bedingten Verteilung dieser fehlenden Daten erstellt, wobei die anderen Variablen berücksichtigt werden. Im Verlauf mehrerer Schritte werden Modellparameter aktualisiert und fehlende Werte durch sorgfältig prognostizierte Werte, unter Einbeziehung eines Zufallselements, ersetzt. Diese Methode ermöglicht es, unterschiedliche Arten von Daten (wie metrische, kategorial und Zähldaten) angemessen zu behandeln und sorgt dafür, dass der Algorithmus nach einer bestimmten Anzahl von Durchläufen (üblicherweise 10 bis 20) stabile und den wahren Daten entsprechende Werte liefert. Ein besonderer Vorteil von MICE liegt in seiner Flexibilität und der Fähigkeit, verschiedene Modellierungsansätze für unterschiedliche Datentypen zu integrieren, obwohl es zu logischen Unstimmigkeiten bei den imputierten Werten kommen kann. Insgesamt bietet MICE einen robusten Rahmen für die multiple Imputation komplexer Datensätze, wodurch umfassendere und genauere statistische Analysen ermöglicht werden.

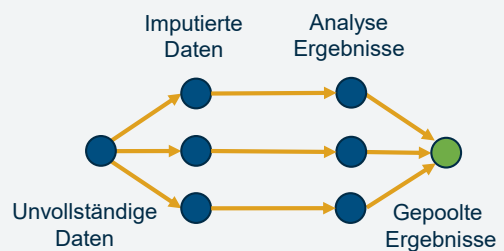


Abb. 7 Imputation durch Rekonstruktion aus teilweise unvollständigen Datensätzen.

Künstliche Intelligenz und maschinelles Lernen in der Batteriezellproduktion

Künstliche Intelligenz (KI) und maschinelles Lernen (ML) sind Technologien, die das Potenzial haben, die Produktion von Batteriezellen wesentlich zu verbessern. Während KI darauf abzielt, Systeme zu entwickeln, die ähnlich wie Menschen denken, lernen und Probleme lösen können, spezialisiert sich ML auf das Entwickeln von Algorithmen, die Computern ermöglichen, aus Daten zu lernen und Entscheidungen zu treffen, ohne dafür explizit programmiert zu sein.

Überwachtes Lernen

Beim überwachten Lernen wird ein Modell mit Daten trainiert, die sowohl die Eingaben (Input) als auch die gewünschten Ausgaben (Output) enthalten. Ziel ist es, ein Modell zu entwickeln, das Muster in diesen Daten erkennt und diese auf neue, unbekannte Daten anwenden kann. Es gibt zwei Hauptkategorien von Problemen, die mit überwachtem Lernen gelöst werden können: Klassifikation und Regression.

Bei der Klassifikation ist das Ziel, Objekte anhand ihrer Merkmale einer vordefinierten Gruppe zuzuordnen. In der Batteriezellenproduktion könnte ein Klassifikationsmodell genutzt werden, um automatisch zu bestimmen, ob eine produzierte Zelle die Qualitätsstandards erfüllt oder nicht. Basierend auf Input-Daten wie Elektrodenbeschaffenheit, Dichte der Elektrolytlösung und Montagegenauigkeit könnte das Modell entscheiden, ob eine Zelle, als "gut" (i.O.) oder "Ausschuss" (n.i.O.) klassifiziert wird.

Die Regression dagegen verfolgt das Ziel einen mathematischen Zusammenhang zwischen Input und Output zu ermitteln, um auf dieser Basis Ereignisse vorherzusagen. Ein Regressionsmodell in der Batteriezellenproduktion könnte dazu dienen, die voraussichtliche Lebensdauer oder Leistung einer Batteriezelle basierend auf verschiedenen Produktionsparametern wie Temperaturbedingungen während der Herstellung oder Ladezyklen zu schätzen.

Zu den wesentlichen Vorteilen gehört die zielgerichtete Analyse. Durch das Training mit spezifischen Produktionsdaten kann ein Modell effektiv für bestimmte Qualitäts- oder Leistungsziele optimiert werden. Dies ermöglicht eine präzise Anpassung der Modelle an die spezifischen Anforderungen der Batteriezellenproduktion. Darüber hinaus erlaubt die klare Ergebnisinterpretation, die sich aus der direkten Beziehung zwischen den Produktionsbedingungen und Qualitätsstandards ergibt, eine einfache Anwendung und Verständlichkeit der Modellergebnisse. Des Weiteren bietet die Möglichkeit, die Effektivität des Modells anhand realer Produktionsdaten zu bewerten und mit anderen Ansätzen zu vergleichen, eine solide Grundlage für die Beurteilung und stetige Verbesserung des Produktionsprozesses.

Eine große Herausforderung ist jedoch die Abhängigkeit von der Qualität und Vielfalt der Trainingsdaten. In der Batteriezellenproduktion bedeutet dies, dass umfangreiche und präzise Daten über den gesamten Herstellungsprozess erforderlich sind, um ein zuverlässiges Modell zu entwickeln. Ein weiteres potenzielles Problem ist das Risiko der Überanpassung (Over Fitting). Wenn ein Modell zu stark auf die spezifischen Bedingungen der Trainingsdaten zugeschnitten ist, könnte es bei

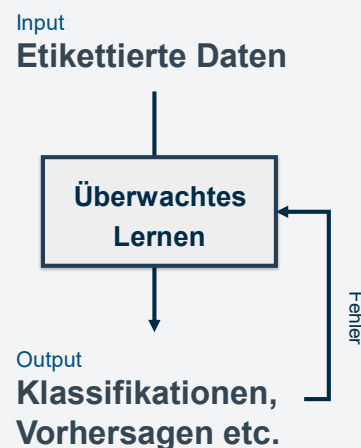


Abb. 8 Überwachtes Lernen.

geringfügigen Abweichungen in der Produktion oder bei neuen Produktionsmethoden unzuverlässige Ergebnisse liefern.

Insgesamt bietet das überwachte Lernen zwar bedeutende Möglichkeiten zur Verbesserung der Effizienz und Qualität in der Batteriezellenproduktion, erfordert jedoch eine sorgfältige Berücksichtigung der Datenvorbereitung und Modellvalidierung, um diese Vorteile voll ausschöpfen zu können.

Unüberwachtes Lernen

Unüberwachtes Lernen findet seine Anwendung in der Entdeckung unbekannter Muster in Daten. Im Gegensatz zum überwachten Lernen, bei dem die Daten mit bekannten Ausgabewerten (Output-Daten) annotiert werden, konzentriert sich das unüberwachte Lernen auf die Analyse der Eingabedaten (Input-Daten) selbst, um Strukturen oder Regelmäßigkeiten zu erkennen. Die drei Hauptanwendungen des unüberwachten Lernens sind Clustering, das Lernen von Assoziationsregeln und Dimensionsreduktion.

Beim Clustering geht es darum, ähnliche Objekte aufgrund ihrer Merkmale in Gruppen zusammenzufassen. In der Batteriezellenfertigung kann diese Methode beispielsweise dazu verwendet werden, Produktionschargen basierend auf bestimmten Qualitätsmerkmalen oder spezifischen Produktionsparametern zu gruppieren. Dies erlaubt es, Muster oder Anomalien in den Produktionsprozessen zu identifizieren, was wiederum dazu beitragen kann, die Qualität der Endprodukte zu verbessern und effizientere Produktionsmethoden zu entwickeln.

Das Lernen von Assoziationsregel zielt darauf ab, Muster, Assoziationen oder Korrelationen zwischen verschiedenen Variablen in den Daten zu finden. In der Batteriezellenproduktion kann diese Technik genutzt werden, um verborgene Beziehungen zwischen verschiedenen Produktionsvariablen zu erkennen. Das Verständnis dieser Beziehungen kann entscheidend sein, um Faktoren zu identifizieren, die sowohl die Qualität als auch die Effizienz der beeinflussen.

Die Dimensionsreduktion, als dritter wesentlicher Ansatz des unüberwachten Lernens, hat das Ziel, die Komplexität der Daten zu verringern, indem die Anzahl der Dimensionen reduziert wird, ohne wesentliche Informationen zu verlieren. Diese Methode ist besonders nützlich, um die kritischsten

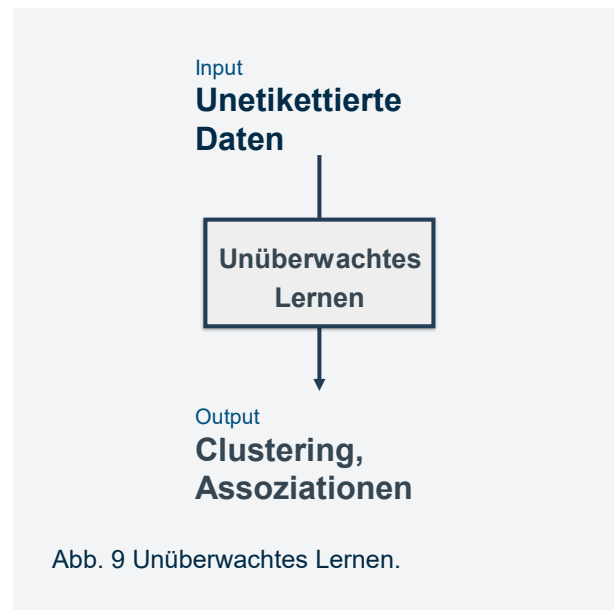


Abb. 9 Unüberwachtes Lernen.

Produktionsparameter zu identifizieren, die die Qualität und Leistungsfähigkeit der Endprodukte maßgeblich beeinflussen. Durch die Fokussierung auf die wichtigsten Faktoren können Hersteller die Produktionseffizienz steigern und gleichzeitig sicherstellen, dass die Qualität der Batteriezellen auf einem hohen Niveau bleibt.

Zu den Vorteilen des unüberwachten Lernens gehört, dass keine vorherige Kennzeichnung von Output-Daten erforderlich ist. Dies erleichtert die Datenvorbereitung, da der oft aufwendige Schritt der manuellen Annotation entfällt. Zudem ermöglicht es die Entdeckung neuer Muster und Zusammenhänge, die zuvor möglicherweise nicht erkannt wurden. Dies kann besonders nützlich sein, um neue Erkenntnisse in komplexen Produktionsumgebungen der Batteriezellenherstellung zu gewinnen.

Die Herausforderungen des unbewachten Lernens liegen vor allem in der Interpretation der Ergebnisse. Da keine definierten Output-Daten zur Validierung vorhanden sind, erfordert die Bewertung der Ergebnisse oft ein tieferes Verständnis der zugrunde liegenden Prozesse und Daten. Dies kann insbesondere in technisch komplexen Bereichen wie der Batteriezellenproduktion eine Herausforderung darstellen, wo Expertenwissen erforderlich ist, um die Ergebnisse des unbewachten Lernens sinnvoll zu interpretieren und anzuwenden.

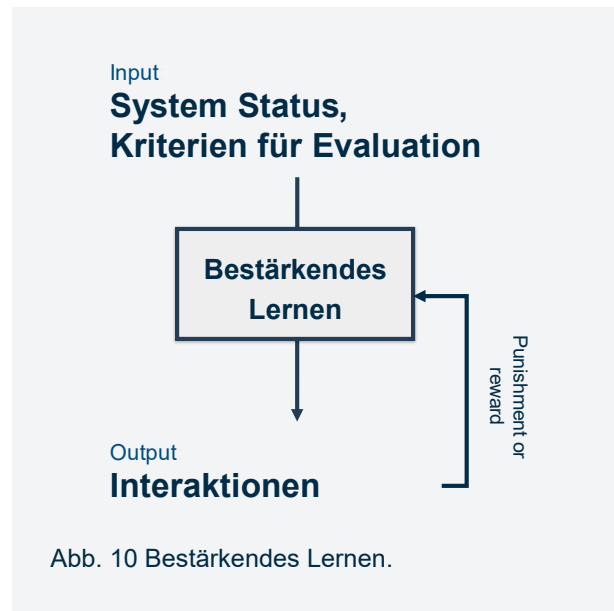
Bestärkendes Lernen

Das bestärkende Lernen, auch bekannt als Reinforcement Learning, stellt einen einzigartigen Ansatz im Bereich des maschinellen Lernens dar, der sich grundlegend von den traditionellen Methoden des überwachten und unbewachten Lernens unterscheidet. Im Gegensatz zu diesen Ansätzen, die auf Input- und Output-Daten angewiesen sind, basiert das bestärkende Lernen auf der Interaktion mit einer Umgebung durch eine Reihe von Aktionen und die daraus resultierende Rückmeldung in Form von Belohnungen.

Bei dieser Methode werden keine spezifischen Input- oder Output-Daten vorgegeben. Stattdessen definiert man eine Menge von möglichen Aktionen, die ein Agent oder Modell in einer gegebenen Umgebung durchführen kann, sowie eine Belohnungsfunktion, die angibt, wie gut eine bestimmte Aktion das Ziel des Agenten erreicht. Diese Belohnungsfunktion dient als Leitfaden für das Modell, um zu lernen, welche Aktionen in verschiedenen Situationen die besten Ergebnisse liefern. Das Modell wird durch Versuch und Irrtum trainiert, wobei es mit jeder Aktion und der daraus resultierenden Belohnung lernt und sich anpasst.

Bestärkendes Lernen ist besonders effektiv für komplexe Probleme, bei denen ein klar definierter Lösungsweg nicht im Voraus bekannt ist. Mit zunehmender Lernzeit ist das Modell in der Lage, immer bessere und manchmal überraschend innovative Lösungen zu entwickeln. In der Batteriezel- lenproduktion könnte bestärkendes Lernen beispielsweise eingesetzt werden, um automatisierte Systeme wie Roboterarme zu trainieren. Diese könnten lernen, spezifische Aufgaben wie die präzise Handhabung von Materialien oder die Optimierung von Fertigungsabläufen durchzuführen, basierend auf der Rückmeldung, die sie durch ihre Leistung erhalten.

Ein Nachteil des bestärkenden Lernens ist jedoch der höhere Rechen- und Speicheraufwand, der mit diesem Ansatz verbunden ist. Daher wird es oft für komplexere Probleme eingesetzt, bei denen die traditionellen Methoden des überwachten oder unüberwachten Lernens nicht ausreichend sind. Trotz des erhöhten Aufwands bietet das bestärkende Lernen ein enormes Potenzial für Anwendungen, in denen ein flexibler, adaptiver Ansatz erforderlich ist, der über die Grenzen herkömmlicher datengetriebener Modelle hinausgeht.



Die Anwendung von maschinellen Lernverfahren in der Datenbasierten Prozesssteuerung

Durch die Vernetzung von Anlagen in den Produktionsstätten steigt die Menge der zur Analyse verfügbaren Daten grundsätzlich an und nimmt schnell schwer überschaubare Ausmaße an. Neben der Aufbewahrung der Daten für die Qualitätssicherung oder aus regulatorischen Gründen, liegen Anwendungsfälle für die Auswertung der Produktionsdaten in der nachgelagerten Fehleranalyse individueller Produktionsprobleme und deutlich seltener in der strukturierten und kontinuierlichen Analyse der laufenden Produktion zur Unterstützung der Entscheidungsfindung über den Produktionsverlauf.

Dabei kann eine kontinuierliche Überwachung von Prozessdaten dabei helfen, frühzeitig relevante Informationen nutzbar zu machen (real-time analysis). Der Rückblick auf bereits produzierte Güter losgelöst vom Tagesgeschäft erlaubt die deskriptive Analyse von Produktionsprozesszusammenhängen und das Training von Modellen, die heute schon eine Indikation auf die Teilequalität von morgen haben (prediction models). Die enge Einbettung dieser Analysen und Modelle in die laufende Produktion erlaubt schließlich das proaktive

und präskriptive Eingreifen in den Produktionsprozess um Entscheidungen auf Basis von erwarteten Produkteigenschaften treffen zu können und diese mit dem zeitlich abgestimmten Kontext der Gesamtproduktion abzuwägen.

Deskriptive datenbasierte Prozessanalyse mit Process Mining

Durch das automatisierte Sammeln von Zustandsdaten direkt aus der Produktion – wie Maschinenzustände, Taktzeiten und Qualitätskennzahlen – wird eine Bewertung von Effizienz und Produktivität auf einer unmittelbaren Datenbasis möglich. Mit Hilfe deskriptiver Analysen können Trends und Muster innerhalb der Produktionsdaten erkannt werden, wodurch sich wiederkehrende Fehlerquellen identifizieren und präventive Maßnahmen ableiten lassen. Dadurch ist die deskriptive datenbasierte Prozessüberwachung ein Werkzeug für die Optimierung von Produktionsprozessen und die Sicherstellung einer gleichbleibend hohen Produktqualität. Process Mining bündelt Methoden und Technologien um realweltliche Prozesse zu analysieren, indem es Informationen aus Ereignisdatenprotokollen (Event Logs) extrahiert [17]. Ursprünglich besonders im Kontext betrieblicher Informationssysteme wie Enterprise Resource

Tab. 1 Exemplarisches Event Log der Batteriezellproduktion.

Zell#	Batch	Zeitstempel	Aktivität	Test	Attribute
238	230306-thizue	2023-02-20T16:13:17	separation	i.O.	Separator=GH2
238	230306-thizue	2023-02-20T16:21:11	drying - start	-	Trocknung_Temp=110 Trocknung_Dauer=4.3h
281	230717-26-jajo	2023-06-12T12:56:52	stacking	n.i.O	kitoolbox_stacking_v=-5.11 kitoolbox_stacking_h=-94.97
281	230717-26-jajo	2023-06-12T13:17:40	filling	n.i.O	vfelyt=1.12 elytistml=1.13 elytistg=1.358 elytsollml=0.939
		...	...		

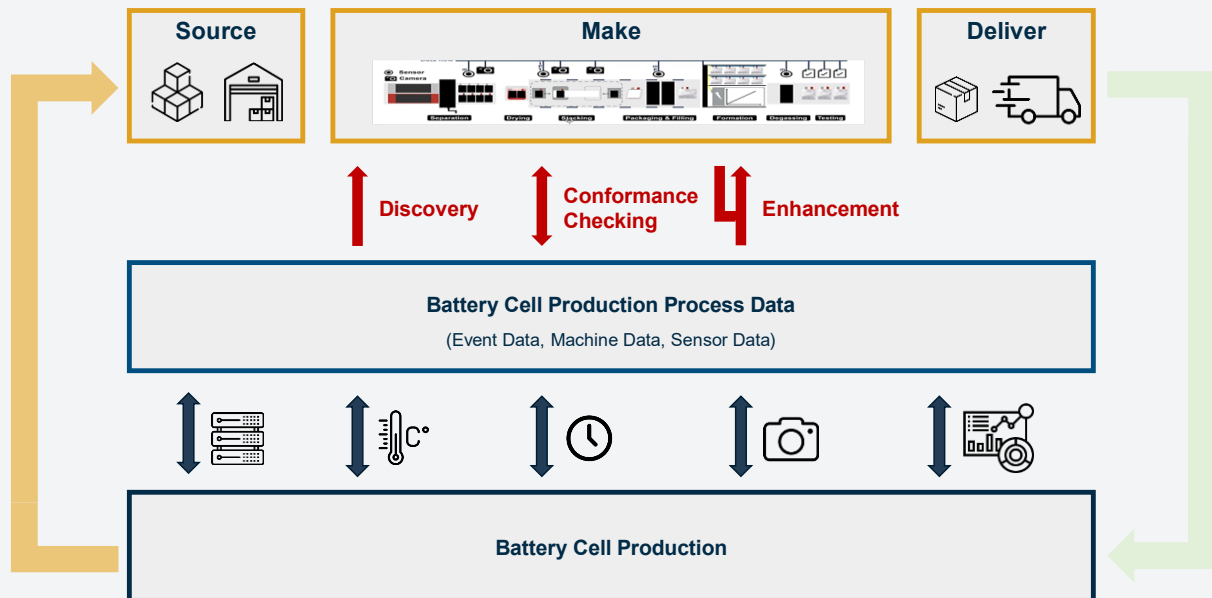


Abb. 11 Drei Hauptanwendungsfälle von Process Mining „discovery“ „conformance checking“ und „enhancement“ im Kontext der Batteriezellproduktion.

Management (ERP) Systeme eingesetzt, bietet der Werkzeugkoffer Process Mining auch im Produktionsumfeld Möglichkeiten zur datenbasierten Prozessanalyse. Im Fokus von Process Mining steht ein gesamter (Geschäfts-)prozess als Summe von Einzelaktivitäten, der wiederholt ausgeführt wird: Für jede Prozessinstanz (jede produzierte Batteriezeile) werden die durchlaufenen Produktionsschritte mit Zeitstempel und relevanten Attributen fortlaufend aufgezeichnet. Tab. 1 zeigt einen stark vereinfachten Ausschnitt eines Event Log aus der Testzellproduktion im Rahmen des Projektes KIProBatt.

Process Mining-Algorithmen können aus Event Logs IST-Modelle der Prozesse konstruieren und ermöglichen Auswertungen, die insbesondere den zeitlichen Verlauf von Prozessinstanzen fokussieren (*Process Discovery*). Der Abgleich mit SOLL-Modellen oder -regeln zeigt Abweichungen von der Norm – sowohl den Ablauf als auch die Leistung des Prozesses betreffend (*Compliance Analysis*). Abb. 11 veranschaulicht, wie Process Mining-Techniken Prozessdaten analysieren, um Verbindungen zwischen dem theoretischen Prozessmodell (in Gelb dargestellt) und dem tatsächlichen Prozessablauf (in Schwarz dargestellt) zu ziehen. Man unterscheidet zwei Arten von

Prozessmodellen: Das deskriptive De-facto-Modell (dargestellt durch den orangenen Pfeil), welches den Ist-Zustand des Prozesses abbildet und direkt aus den realen Abläufen abgeleitet wird, und das normative De-jure-Modell (dargestellt durch den grünen Pfeil), das den idealen Soll-Zustand des Prozesses beschreibt und darauf abzielt, die tatsächlichen Prozesse zu beeinflussen und zu verbessern.

Bei ausreichender und zeitiger Datenverfügbarkeit bieten die aus dem Process Mining gewonnenen Erkenntnisse die Möglichkeit, direkt in umsetzbare Maßnahmen interpretiert zu werden. Die Anwendungen von Process Mining, insbesondere in der Entdeckung, Konformitätsprüfung und Verbesserung von Prozessen, sind von entscheidender Bedeutung für die Optimierung der Produktionsprozesse von Batteriezellen.

Im Kontext der Batteriezellenproduktion können diese Aspekte des Process Mining auf verschiedene Weise eingesetzt werden:

Prädiktives Prozessmonitoring

Prädiktive Prozessüberwachungsmethoden liefert Vorhersagen über den zukünftigen Verlauf einer

bestimmten Prozessausführung [18]. Die mit diesen Methoden erstellten Prognosen können sowohl quantitativ (z. B. Zeit- und Kostenprognosen) als auch kategorial (z. B. Risiko- und Ergebnisprognosen) sein. Dabei ist es möglich, nachfolgende Aktivitäten oder die Abfolge zukünftiger Aktivitäten eines laufenden Prozesses vorherzusagen. Um zu einer Vorhersage zu gelangen, verwenden prädiktive Prozessüberwachungsmethoden Daten aus historischen Prozessausführungen, um ein Vorhersagemodell zu generieren. Dieses Modell kann dann während des Einsatzes mit aktuellen Daten aus der laufenden Prozessausführung gefüttert werden kann.

Abb. 12 zeigt den schematischen Arbeitsablauf der prädiktiven Prozessüberwachung. In der Offline-Phase werden Präfixe aus den historischen Prozessverläufen extrahiert. Präfixe sind die bis zum Prädiktionszeitpunkt abgelaufenen Prozesspfade von strukturanalogen abgeschlossenen Prozessinstanzen (z. B. wurde ein Prozess mit fünfzehn Schritten aufgezeichnet, aber nur die ersten fünf Schritte werden aus dem abgeschlossenen Verlauf extrahiert). Da prädiktive Prozessüberwachungsmethoden darauf abzielen, Vorhersagen zu einem frühen Zeitpunkt in der laufenden Prozessausführung zu treffen, konzentriert sich die Trainingsphase auf Präfixe mit unterschiedlichen Längen. Die Präfixe werden verwendet, um ein Vorhersagemodell zu trainieren. Zur Erstellung eines solchen Modells werden verschiedene Techniken des maschinellen Lernens eingesetzt: Für die Vorhersage eines Prozessergebnisses ist die

Klassifizierung als eine Form des überwachten Lernens am gebräuchlichsten.

Zu den Klassifizierungsalgorithmen, die in vielen Methoden der prädiktiven Prozessüberwachung eingesetzt werden, gehören insbesondere klassische Entscheidungsbäume (z. B. [19, 20]), aber auch Random Forests (z. B. [21]). Während sich die Ergebnisse von Entscheidungsbäumen besser durch den Nutzer interpretieren lassen [22], erzielt der Random Forest häufig eine höhere Vorhersagegenauigkeit. Dies gilt für die Anwendung im Prozesskontext, sowie für Anwendungen mit „flacher“ Datengrundlage.

Für numerische Vorhersagen werden auch in der Prozessvorhersage zumeist Regressionsverfahren verwendet (z. B. [23, 24]). Auch die Verwendung von probabilistischen Modellen oder evolutionären Algorithmen [25] für Prozessvorhersagen ist möglich.

Insgesamt ist jedoch zu beobachten, dass Deep-Learning-Methoden (d.h. tiefe neuronale Netze) bei der Erstellung von Vorhersagen bessere Ergebnisse erzielen. Kratsch et al. (2022) stellten fest, dass viele bestehende Ansätze Techniken des maschinellen Lernens verwenden, obwohl diese im Allgemeinen von Deep-Learning-Techniken übertroffen werden [26]. Die Operationalisierung von Deep-Learning-Techniken in der prädiktiven Prozessüberwachung ist heute möglicherweise nur dessen weniger weit fortgeschritten sein, weil es an einem Verständnis darüber

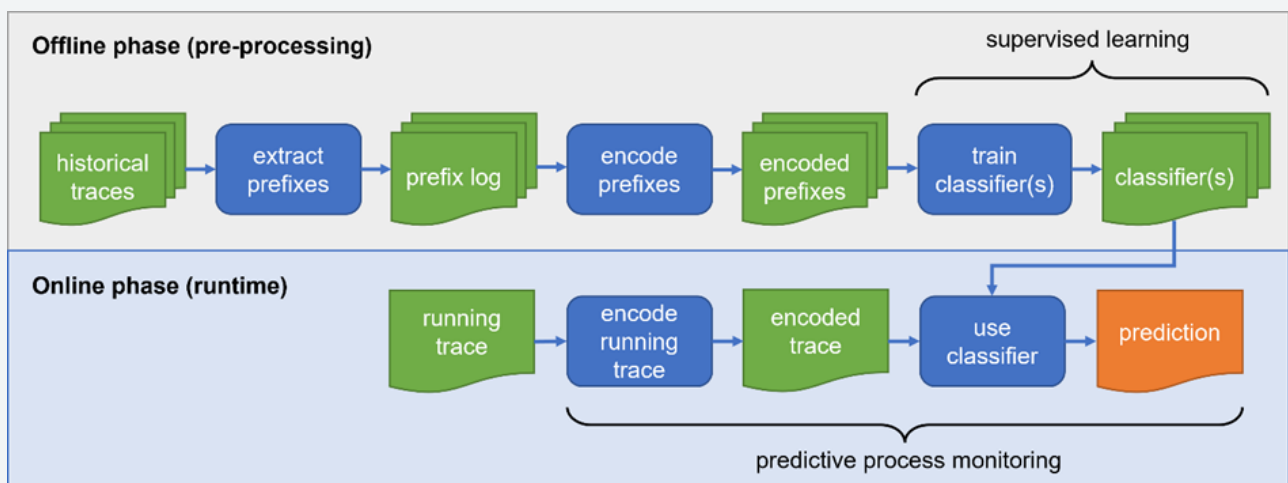


Abb. 12 Arbeitsablauf der prädiktiven Prozessüberwachung.

mangelt, welche Eigenschaften von Ereignisprotokollen den Einsatz von Deep-Learning -Techniken erleichtern oder weil es sich als schwieriger erweist, explizit nachzuvollziehen, wie ein Ergebnis zustande gekommen ist [27]. Gleichzeitig gestaltet in der Regel auch die Erklärbarkeit der von Deep-Learning-Modell schwieriger als bei anderen Ansätzen des maschinellen Lernens.

Wichtige Bewertungsmaßstäbe für prädiktive Prozessüberwachungsmethoden sind die Genauigkeit und die Frühzeitigkeit der Vorhersage. „Eine Vorhersage, die oft ungenau ist, ist eine nutzlose Vorhersage, da man sich bei der Entscheidungsfindung nicht auf sie verlassen kann“ [22]. Eine hohe Genauigkeit ist eine wesentliche Voraussetzung für das Vertrauen in die Vorhersagen und damit für den Erfolg der prädiktiven Prozessüberwachung in der Praxis. Entscheidungen, die auf Vorhersagen beruhen, werden nur dann getroffen, wenn den Vorhersagen vertraut wird. In der Praxis können ungenaue Vorhersagen zu hohen Zusatzkosten führen. Wenn zum Beispiel eine Vorhersage besagt, dass eine produzierte Produktcharge mit hoher Wahrscheinlichkeit so fehlerhaft ist, dass sie nicht verkauft werden kann, wird die Produktion gestoppt und die Charge verworfen. Handelt es sich jedoch um eine Produktcharge, die regulär hätte verkauft werden können, entstehen dem Unternehmen aufgrund einer ungenauen Vorhersage Kosten. Umgekehrt entstehen auch vermeidbare Kosten, wenn eine Produktcharge nach einer positiven Prognose einen teuren, energieintensiven und langwierigen Produktionsprozess durchläuft und sich erst am Ende herausstellt, dass das Endprodukt stark fehlerhaft ist. Die Genauigkeit wird mit Metriken wie Präzision, Recall, F-Score und ROC AUC gemessen [26]. Es ist jedoch zu beachten, dass eine Vorhersage nie völlig fehlerfrei sein kann. Daher sollten Untersuchungen und Berechnungen darüber angestellt werden, wie kostspielig Vorhersagefehler in einem bestimmten Anwendungsfall sind, um das Potenzial der prädiktiven Prozessüberwachung besser einschätzen zu können.

Auch die Frühzeitigkeit der Vorhersage einer Methode ist von großem Interesse. Je früher eine zuverlässige Vorhersage zum Prozessverlauf vorhanden ist, desto kann reagiert werden [22]. Die Frühzeitigkeit der Vorhersage kann gemessen werden, indem getestet wird, wie weit die Länge des Präfixes reduziert werden kann, während ein akzeptables Maß an Genauigkeit erreicht wird.

Prozesserkennung und Konformitätsprüfung in der Materialbeschaffung

In den frühen Phasen der Batteriezellenproduktion, wie bei der Beschaffung, Lagerung und Lieferung von Materialien, spielen Prozesserkennung und Konformitätsprüfung eine wesentliche Rolle. Durch die Analyse von Daten aus Enterprise Resource Planning (ERP)-Systemen kann die Übereinstimmung des tatsächlichen Prozesses mit dem gewünschten Standard überprüft werden. Hierbei lassen sich diverse Einflussfaktoren identifizieren, die zu zusätzlichen Kosten oder Verzögerungen führen können. Beispielsweise könnten fehlgeschlagene Qualitätsprüfungen, die zusätzliche manuelle Tests erfordern, aufgedeckt werden. Ebenso können Muster bei Lieferverzögerungen oder der Qualität von Vorprodukten analysiert werden, um problematische Lieferketten zu identifizieren und die Beschaffungsstrategie zu optimieren. Dies trägt dazu bei, Ausfallzeiten und Störungen in der Fertigung zu reduzieren.

Prozessverbesserung in der Batteriezellenproduktion

Bei der eigentlichen Herstellung der Batteriezellen kann Process Mining zur Analyse von Verzögerungen und Durchlaufzeiten genutzt werden, um den Produktionsablauf zu optimieren. Insbesondere in der komplexen Prozesskette der Batteriezellenproduktion bietet Process Mining die Möglichkeit, ineffiziente Prozessschritte zu identifizieren und entsprechende Optimierungsmaßnahmen zu entwickeln.

Echtzeit-Prozessüberwachung

Im fortgeschrittenen Stadium der Batteriezellenproduktion ermöglicht die Echtzeit-Prozessüberwachung die frühzeitige Erkennung von Abweichungen vom Soll-Zustand. Dies ist besonders wichtig für die Identifizierung von Fehlerquellen und trägt dazu bei, die Ausschussrate zu verringern. Aufgrund der Komplexität des Produktionsprozesses wird dabei häufig auf Methoden der künstlichen Intelligenz zurückgegriffen, um tiefere Einblicke zu gewinnen und präzisere Vorhersagen zu treffen.

Smarte Qualitätsvorhersage und präskriptive Prozesssteuerung mit Virtual Quality Gates

Durch die Kombination von prädiktiven Modellen zur Vorhersage der Batteriezellqualität mit Tools zur prädiktiven Prozesssteuerung lassen sich Entscheidungen für die unmittelbare Prozesssteuerung genauso wie übergreifende Prozessanalysedaten sammeln und aufbereiten. Eine ganzheitliche Strategie für den Prozessdatenlebenszyklus nutzt anfallende Daten weiterhin zum kontinuierlichen Trainieren der prädiktiven Modelle.

Prozesssteuerung in Gateways

Ein zentrales Arbeitsergebnis im Projekt KIProBatt ist die Konzeption von Monitoring und Steuerungsfunktionen im laufenden Produktionsprozess. Konzeptionelle Grundlage stellen hier die Gateways, bekannt aus der Business Process Model and Notation (BPMN). BPMN ist der führende Standard zur Erstellung von Geschäftsprozessmodellen. Gateways sind Entscheidungspunkte, die den Pfad des Flusses unter bestimmten Bedingungen verändern können. Gateways selbst sind keine eigenständigen Aktivitäten, sondern bilden eine Steuerungslogik ab, die den weiteren Fluss einer einzelnen Prozessinstanz regelt. Dafür sind Verzweigungen vorgesehen, die alternativen Prozesspfaden entsprechen. Die BPMN kennt Gateways die einen oder mehrere Pfade gleichzeitig auswählen und dabei einfache oder komplexe Regelwerke nutzen. Bezogen auf die Entscheidungsfindung im Kontext der linearen Batteriezellproduktion müssen komplexe Entscheidungen zu verschiedenen Zeitpunkten

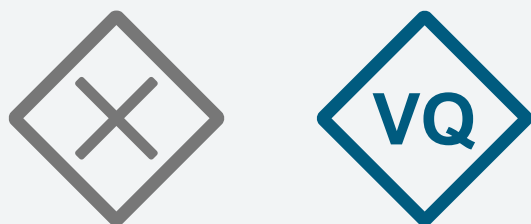


Abb. 13 Gegenüberstellung Exklusives BPMN-Gateways und Virtual Quality Gate.

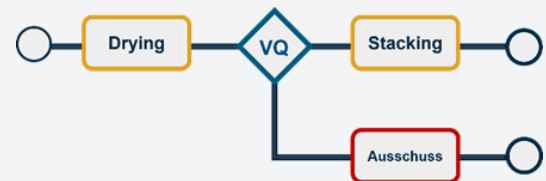


Abb. 14 Vereinfachtes Prozessschaubild mit Virtual Quality Gate.

entlang der Produktion getroffen werden. Entsprechend der multikriteriellen Zielvorgabe einer effizienten, sicheren und ressourcensparenden Produktion müssen vielfältige Aspekte zu verschiedenen Entscheidungspunkten integriert und unter wechselnden Bedingungen betrachtet werden.

Virtual Quality Gates

Die hohen Anforderungen an die Flexibilität bei der Steuerung können erreicht werden, indem an den Entscheidungspunkten eine ausführliche Datengrundlage genutzt werden kann und die Entscheidungsfindung dynamisch gesteuert wird. Um diesen Anforderungen gerecht zu werden konzipieren wir Virtual Quality Gates als Erweiterung zu regulären Gateways der BPMN.

Virtual Quality Gates docken an der physischen Ebene der Produktion an und kombinieren die reale Welt mit Daten der Cyberphysischen Welt und der Prozesswelt. Analog zu klassischen Gateways bilden auch Virtual Quality Gates keine tatsächlichen Aktionen ab, sondern dienen primär der Entscheidungsfindung und der Bestimmung des nächsten Prozesspfads (siehe Abb. 14). In weiterer Funktion ermöglicht es ihre durchgehende Vernetzung in die Prozess-, die Cyberphysische und die Physische Ebene den Datenaustausch mit anderen Virtual Quality Gates als Diagnose- und Monitoringinstrument. Abb. 15 bildet die Konzeptionellen Bestandteile ab.

Input

Virtual Quality Gates können an verschiedenen Stellen im Prozessablauf platziert und konfiguriert

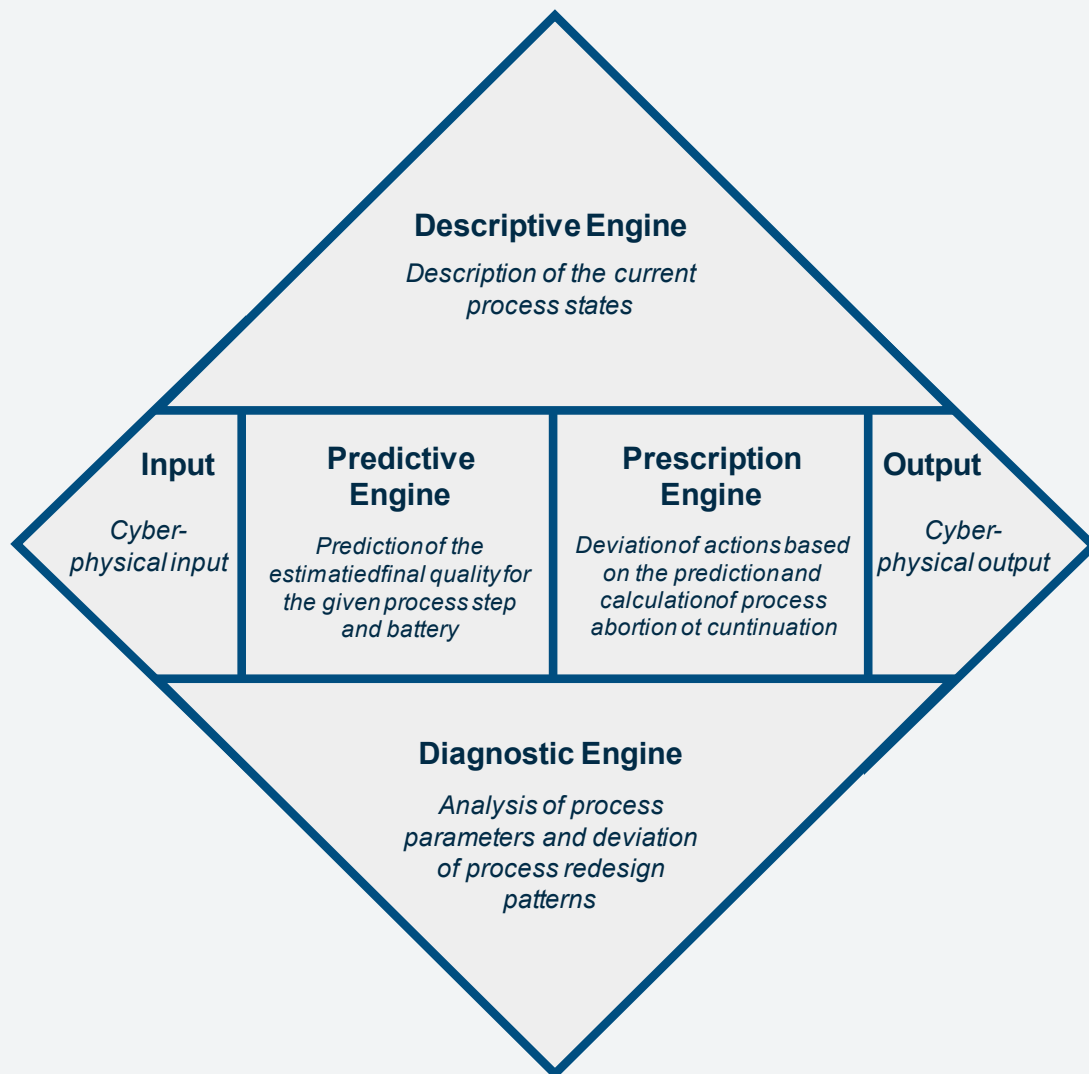


Abb. 15 Illustration des Virtual Quality Gates mit Komponenten zur Diagnostik des IST-Zustands, einer prädiktiven und einer präskriptiven Komponente für die Auswahl des nächsten Prozesspfades.

werden. Für eine konkrete Instanz einer Entscheidungsfindung erhält ein Virtual Quality Gate die aktuellen Produkt- und Produktionsdaten zu einem Teil als direkten Input (In-Instanz-Daten). Für dieses spezifisch vorliegende Teil soll auf Basis der erhobenen Daten eine Entscheidung gefunden werden.

Descriptive Engine

Aus ihrer virtuellen Vernetzung mit anderen Quality Gates verfügen sie über die Gesamtheit der Daten (Grundlegende Produktionsparameter, Prozessüberwachungsdaten, Entscheidungsunterstützende Daten und Meta-Produktionsdaten), die andere Gates, Sensorik und die Konfiguration der Produktionsanlage ihnen bereitstellen. Die Deskriptive Komponente eines Virtual Quality Gates

verknüpft die vorliegenden Daten zu einem ganzheitlichen IST Bild der Prozessinstanz und seines aktuellen Prozesskontextes in der Entscheidungssituation: Welche Prozessschritte wurden gegangen, welche Unterkomponenten bereits verbaut, wie hoch ist die Auslastung der Produktion oder welche Messwerte wurden bisher erhoben. Diese Informationen werden von den weiteren Komponenten für die Entscheidungsfindung genutzt.

Predictive Engine

Die Predictive Engine macht sich Verfahren zu Nutze, die möglichen Zustände des vorliegenden Bauteils unter Annahme verschiedener Szenarien schätzen. Vornean ist das Ziel der Prädiktion die Teilequalität zu schätzen. In der multikriteriellen Bewertung von Prozesskennzahlen sind jedoch

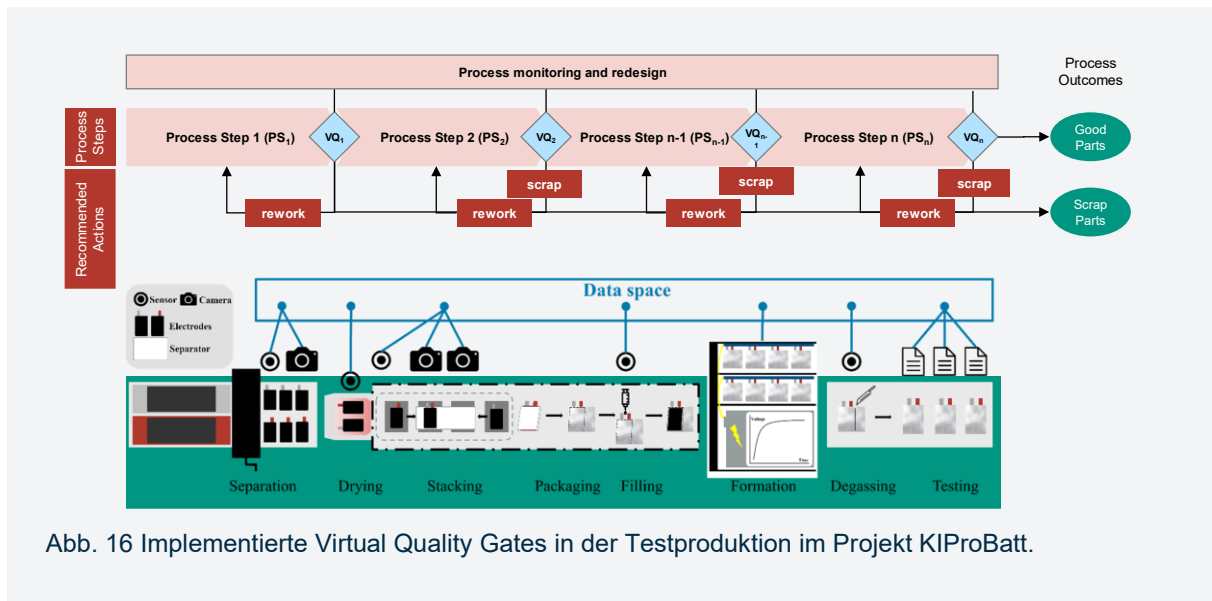


Abb. 16 Implementierte Virtual Quality Gates in der Testproduktion im Projekt KIProBatt.

auch Abschätzungen über das ökonomische Potential, Zeitaufwände, Umwelt- und Sicherheitsbetrachtungen relevante Schätzungen, die den weiteren Prozessverlauf beeinflussen können. Durch die Prädiktion mehrfacher relevanter Größen zu diesem Zeitpunkt kann die weitere Entscheidungsfindung komplexere und bessere informierte Handlungsanweisungen generieren. Je nach Verortung des Virtual Quality Gates im Prozessmodell sind die Verfahren zur Schätzung der Produkteigenschaften basierend auf unterschiedlichen Daten, verwenden unterschiedliche Verfahren und unterliegen schließlich auch unterschiedlichen Sicherheiten.

Prescription Engine

Die präskriptive Komponente hat die tatsächliche Entscheidungsfindung über den weiteren Produktionsverlauf für das vorliegende Bauteil zur Aufgabe. Der Input, Informationen aus der Descriptive Engine und der Predictive Engine bilden die Grundlage für die Entscheidungsfindung. Auf dieser Basis können Entscheidungen sowohl durch Black-Box Verfahren in lernenden Modellen, regelbasierten Entscheidungswerken oder Kombinationen beider Verfahren getroffen werden.

Diagnostic Engine

Die Diagnostische Komponente stellt das Gegenstück zur Deskriptiven Komponente dar. Sie stellt die Inputwerte gemeinsam mit den Ergebnissen aus der Prädiktiven und der Präskriptiven Komponente für die weiteren Virtual Quality Gates zur Verfügung. Auf Basis des Datenlage der Diagnostischen Komponenten kann die Produktion als

Gesamtes Ende zu Ende überwacht werden und Trends im Produktionsverlauf beobachtet werden.

Output

Der Output eines Virtual Quality Gates ist die Entscheidung über die Verzweigung, die das Bauteil im weiteren Produktionsverlauf zunächst nimmt. Durch soft-labeling werden dem Bauteil weitere virtuelle Informationen angeheftet, die die getroffene Entscheidung für weitere Produktionsschritte transparent machen und erklären.

Anwendung im Projekt KIProBatt

Das Konzept der Virtual Quality Gates wurde im Projekt KIProBatt an mehreren Stellen im Batteriezellproduktionsprozess implementiert. Dazu wurden aus einer Testzellproduktion Daten erhoben und für die Operationalisierung aufbereitet. In der Testproduktion wurden mehrere exemplarische Eingriffsfälle und Entscheidungskriterien identifiziert und ausgearbeitet, die auf ein frühestmögliches Ausschleusen von Schlechteilen aus dem kostenintensiven Produktionsprozess fokussieren, ein Nacharbeiten von Teilen anstoßen, oder die Weiterbearbeitung der Zellen unter Risiko abzielen.

Zunächst werden Szenarien erdacht, die alternative Prozesspfade abbilden. Die Standardprozessvariante, der vorliegende Produktionsprozess bleibt davon nicht weiter betroffen. Für die Suche nach alternativen Prozesspfaden können Kreativitätstechniken zusammen mit Domänenexperten genutzt werden, die den Handlungsspielraum für

die Suche aufzeigen: Beispielsweise beschreiben Best Practice Patterns häufig praktizierte Maßnahmen für die Prozessverbesserung als abstrakte Ideen. Das Early Knockout Pattern etwa verfolgt die Idee die Prozesskosten zu reduzieren, in dem potenziell fehlerhafte Prozessinstanzen möglichst früh und möglichst kostengünstig abgebrochen werden. Für KIProBatt wurde eine Möglichkeit für die Anwendung dieser Best Practice nach dem Produktionsschritt Stacking identifiziert: eine Auswertung der Stackinggenauigkeit ermöglicht es, schlecht gestapelte Elektrodenstacks direkt aus der Produktion zu schleußen, bevor Qualitätsprobleme erst am End-of-Line Test sichtbar werden.

Zu jedem Anwendungsfall werden im nächsten Schritt die vorhandenen und erforderlichen Inputdaten für das Quality Gate definiert. Die Stackinggenauigkeit wird nicht direkt gemessen, sondern ist nur unmittelbar aus Fotos der Stacks ermittelbar. In der Implementierung wurde daher ein KI-Modell darauf trainiert aus den Bildern zunächst Abweichungen zu erkennen. Ein weiteres Modell kann dann in der Prädiktion von der Stapelgenauigkeit auf die erwartete Kapazität der Zelle nach der Formierung schließen. Diese prädiktive Angabe der Zellqualität wird in diesem Produktionsschritt für die Entscheidung über die Weiterproduktion oder herangezogen.

Ein virtuelles Abbild der Produktion wurde als Softwaredemonstrator mit den Virtual Quality Gates instanziiert und ist schematisch in Abb. 16 dargestellt.

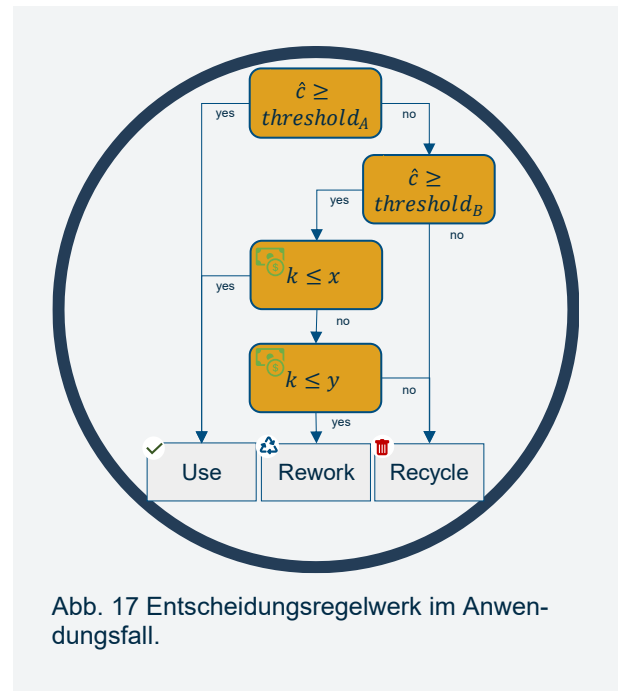


Abb. 17 Entscheidungsregelwerk im Anwendungsfall.

Implementierungshilfe für prädiktive Prozessüberwachung

Der CRISP-DM (Cross-Industry Standard Process for Data Mining) ist ein branchenunabhängiger Standardprozess für Data Mining Projekte. Er beschreibt die Life Cycle einer Anwendung von Data Mining und dient als gängiger Leitfaden in der Praxis. CRISP-DM gliedert ein Data Mining Projekt in sechs miteinander verknüpfte Phasen: Geschäftsverständnis, Datenverständnis, Datenvorbereitung, Modellierung, Evaluierung und Deployment. Diese strukturierte, sequenzielle Vorgehensweise hat sich bei zahlreichen Projekten bewährt und hilft Projektrisiken zu minimieren.

Da vorausschauende Prozessüberwachung auf Data Mining und maschinellem Lernen basiert, eignet sich CRISP-DM als Rahmenwerk für einen systematischen Implementierungsprozess. Die folgenden Implementierungsschritte orientieren sich daher eng an den CRISP-DM Phasen und erweitert sie um spezifische Aspekte für die prädiktive Prozessüberwachung in der Batterieproduktion.

1 | Geschäftsverständnis

In der Phase des Geschäftsverständnisses müssen zunächst die konkreten Geschäftsziele für den Einsatz der prädiktiven Prozessüberwachung in der Batteriezellproduktion definiert werden. Wie im vorherigen Abschnitt erläutert, stehen hier besonders (a) die Senkung der Ausschussrate, (b) die Senkung der Produktionskosten sowie, (c) die Senkung des Energieverbrauchs im Mittelpunkt. Hierbei ist es wichtig, dass alle beteiligten Stakeholder wie Prozessexperten, Produktionsmanager, Mitarbeiter der Qualitätssicherung etc. ihre Anforderungen und Erwartungen einbringen können. Nur wenn die unterschiedlichen Perspektiven vereint werden, lassen sich ganzheitliche Optimierungsziele formulieren.

Es muss in dieser Phase auch geklärt werden, wie die Prozessvorhersagen konkret genutzt werden sollen, um den laufenden Prozess zu verbessern. Dabei sind die kritischen Produktionsschritte und Schlüsselparameter zu identifizieren. Außerdem sollten Kriterien definiert werden, anhand derer

der Erfolg des Projektes gemessen wird, z. B. die Erreichung bestimmter Zielwerte bei den KPIs. Ein umfassendes Geschäftsverständnis bildet die Grundlage für alle weiteren Projektphasen.

2 | Datenverständnis

In der Phase des Datenverständnisses geht es darum, die für die prädiktive Prozessüberwachung relevanten Datenbestände zu identifizieren und zu bewerten. Zunächst müssen die benötigten Daten aus den verschiedenen Quellen in der Produktionsumgebung gesammelt und in eine konsistente Datengrundlage integriert werden. Alle etwaigen Probleme bei der Datenbeschaffung sollten dokumentiert werden. Sollten in einem späteren Schritt systematisch Daten fehlen, lässt sich diese Dokumentation als Hilfe heranziehen.

Anschließend sind die Eigenschaften und Strukturen der gesammelten Daten wie Quantität, Formate, Feldanzahl etc. genau zu beschreiben, um ein Grundverständnis zu schaffen. Durch erste explorative Analysen und Berichte können bereits interessante Muster, Teilmengen oder Hypothesen über die Daten identifiziert werden.

In diesem Schritt ist es wichtig, die Datenqualität kritisch zu bewerten und auf mögliche Mängel wie fehlende Werte zu prüfen. Nur wenn die Datenmenge und -qualität für aussagekräftige Vorhersagemodelle ausreichend ist, sollte die nächste Phase der Datenvorbereitung begonnen werden. Die sorgfältige Untersuchung der Datenbestände stellt sicher, dass die richtigen Datengrundlagen für die Modellierung geschaffen werden.

3 | Datenvorbereitung

In der Phase der Datenvorbereitung gilt es, die finale Datenbasis für die Modellierung der Prozessvorhersagen zu erstellen. Dafür muss zunächst eine gezielte Auswahl der relevanten Datenbestände aus den gesammelten Rohdaten getroffen werden. Diese Auswahl richtet sich nach den zuvor definierten Geschäftszielen und Anforderungen sowie den technischen Voraussetzungen. Man könnte sich beispielsweise auf Sensordaten von bestimmten kritischen Prozessschritten oder

nur Datensätze mit vollständigen Werten beschränken.

Der nächste Teilschritt ist die gründliche Bereinigung der Daten, um Fehler, Inkonsistenzen und Ausreißer zu beseitigen. Ohne saubere Datengrundlage ist keine verlässliche Modellierung möglich. Anschließend werden die Daten transformiert und integriert, indem sie in ein für die Modellierung geeignetes Format gebracht werden. Dazu können Kodierungen, Aggregationen oder Berechnungen von abgeleiteten Merkmalen wie Prozesskennzahlen notwendig sein. Gegebenenfalls müssen auch noch Datentyp-Anpassungen vorgenommen werden. Am Ende dieser Phase steht die bereinigte, transformierte und integrierte Datenbasis, die für effektive Vorhersagemodelle unerlässlich ist.

4 | Datenmodellierung

In der Modellierungsphase müssen zunächst die die am besten geeigneten zu den Voraussetzungen Machine Learning Techniken zur Vorhersage der kritischen Qualitätsparameter in der Batterieproduktion ausgewählt werden.

Je nach Modellierungsansatz wie etwa Klassifikation oder Regression sind unterschiedliche Datenformate erforderlich. Falls die vorbereiteten Daten nicht den Anforderungen genügen, muss in dieser Phase ein Rücksprung in die Datenvorbereitung erfolgen, um die Struktur anzupassen.

Anschließend werden verschiedene Testmodelle erstellt und ihr Vorhersagefehler z.B. bei der Ausschussratschätzung bewertet. Nur die Modelle, die nachweislich in der Lage sind, die definierten Ziele zur Prozessoptimierung wie Senkung der Ausschussquote qualifizieren sich für die nächsten Schritte.

5 | Evaluierung

Vor einem produktiven Einsatz der Modelle müssen diese einer rigorosen Evaluierung unterzogen werden. Dabei ist zu prüfen, ob die Modelle die angestrebten Qualitätsvorgaben wie eine Mindestsenkung der Ausschussrate um einen definierten Prozentsatz erreichen können.

Nur Modelle, die diese kritischen Metriken erfüllen, kommen für die Praxis infrage. Darüber hinaus ist zu bewerten, wie robust die Modelle auf neue, bisher ungesehene Daten aus der Produktion reagieren und wie gut sie auf zukünftige

Szenarien übertragbar sind. Durch Vergleiche unterschiedlicher Modellierungs- und Datenansätze lassen sich die erfolgversprechendsten Kandidaten identifizieren. Werden die Zielvorgaben nicht erreicht, muss der Data Mining Prozess iterativ mit neuen Daten, Parametern oder Techniken wiederholt werden

6 | Deployment

Haben in der Evaluierungsphase belastbare Vorhersagemodelle die Praxistauglichkeit für die Prozessoptimierung unter Beweis gestellt, kann die Produktivsetzung erfolgen. Die Modelle müssen für den robusten Einsatz in der Prozessüberwachungsumgebung vorbereitet und in die bestehenden Systeme der Produktionsüberwachung integriert werden.

Dabei ist ein detaillierter Rollout-Plan zu erstellen, der auch Mitarbeiterschulungen und Durchgänge im Testbetrieb vorsieht. Essenziell für den anhaltenden Erfolg sind außerdem Prozesse für die Überwachung der Modellleistung im laufenden Betrieb und deren regelmäßige Aktualisierung bei Bedarf. Nach erfolgreichem Abschluss des Deployments ist ein prädiktives Prozessmanagement in der Batterieproduktion möglich.

Tabellenverzeichnis

Tab. 1 Exemplarisches Event Log der Batteriezellproduktion	18
--	----

Abbildungsverzeichnis

Abb. 1 Schematischer Ablauf der Batteriezellproduktion.	7
Abb. 2 Fünf Ebenenmodell der Digitalisierung nach Urbach und Röglinger (2019).	10
Abb. 3 Imputation durch Schätzung aus Lageparametern.	13
Abb. 4 Imputation via Regression.	13
Abb. 5 Imputation aus ähnlichen Datensätzen.	14
Abb. 6 Imputation mittels LOCF	14
Abb. 8 Imputation durch Rekonstruktion aus teilweise unvollständigen Datensätzen.	14
Abb. 9 Überwachtes Lernen.	15
Abb. 10 Unüberwachtes Lernen.	16
Abb. 11 Bestärkendes Lernen.	17
Abb. 12 Drei Hauptanwendungsfälle von Process Mining „discovery“ „conformance checking“ und „enhancement“ im Kontext der Batteriezellproduktion.	19
Abb. 13 Arbeitsablauf der prädiktiven Prozessüberwachung.	20
Abb. 14 Gegenüberstellung Exklusives BPMN-Gateways und Virtual Quality Gate.	22
Abb. 15 Vereinfachtes Prozessschaubild mit Virtual Quality Gate.	22
Abb. 16 Illustration des Virtual Quality Gates mit Komponenten zur Diagnostik des IST-Zustands, einer prädiktiven und einer präskriptiven Komponente für die Auswahl des nächsten Prozesspfades.	23
Abb. 17 Implementierte Virtual Quality Gates in der Testproduktion im Projekt KIProBatt.	24
Abb. 18 Entscheidungsregelwerk im Anwendungsfall.	25

Referenzen

- [1] M. Beuse, T. S. Schmidt, and V. Wood, "A "technology-smart" battery policy strategy for Europe," *Science*, vol. 361, no. 6407, pp. 1075–1077, 2018, doi: 10.1126/science.aau2516.
- [2] M. Ghobakhloo, M. Iranmanesh, B. Foroughi, E. Babaee Tirkolaee, S. Asadi, and A. Amran, "Industry 5.0 implications for inclusive sustainable manufacturing: An evidence-knowledge-based strategic roadmap," *Journal of Cleaner Production*, vol. 417, p. 138023, 2023, doi: 10.1016/j.jclepro.2023.138023.
- [3] F. Psarommatis, J. Sousa, J. P. Mendonça, and D. Kiritsis, "Zero-defect manufacturing the approach for higher manufacturing sustainability in the era of industry 4.0: a position paper," *International Journal of Production Research*, vol. 60, no. 1, pp. 73–91, 2022, doi: 10.1080/00207543.2021.1987551.
- [4] F. Mahiri, A. Najoua, and S. B. Souda, "Data-driven sustainable smart manufacturing: A conceptual framework," in *2020 International Conference on Intelligent Systems and Computer Vision (ISCV)*, Fez, Morocco, uuuu-uuuu, pp. 1–7.
- [5] M. Abubakr, A. T. Abbas, I. Tomaz, M. S. Soliman, M. Luqman, and H. Hegab, "Sustainable and Smart Manufacturing: An Integrated Approach," *2071-1050*, vol. 12, no. 6, p. 2280, 2020, doi: 10.3390/su12062280.
- [6] R. P. Cunha, T. Lombardo, E. N. Primo, and A. A. Franco, "Artificial Intelligence Investigation of NMC Cathode Manufacturing Parameters Interdependencies," *Batteries & Supercaps*, vol. 3, no. 1, pp. 60–67, 2020, doi: 10.1002/batt.201900135.
- [7] M. Duquesnoy, T. Lombardo, M. Chouchane, E. N. Primo, and A. A. Franco, "Data-driven assessment of electrode calendaring process by combining experimental results, in silico mesostructures generation and machine learning," *Journal of Power Sources*, vol. 480, p. 229103, 2020, doi: 10.1016/j.jpowsour.2020.229103.
- [8] J. Schnell *et al.*, "Data mining in lithium-ion battery cell production," *Journal of Power Sources*, vol. 413, pp. 360–366, 2019, doi: 10.1016/j.jpowsour.2018.12.062.
- [9] Y. Takagishi, T. Yamanaka, and T. Yamaue, "Machine Learning Approaches for Designing Mesoscale Structure of Li-Ion Battery Electrodes," *Batteries*, vol. 5, no. 3, p. 54, 2019, doi: 10.3390/batteries5030054.
- [10] A. Turetskiy, J. Wessel, C. Herrmann, and S. Thiede, "Battery production design using multi-output machine learning models," *Energy Storage Materials*, vol. 38, pp. 93–112, 2021, doi: 10.1016/j.ensm.2021.03.002.
- [11] O. Badmos, A. Kopp, T. Bernthaler, and G. Schneider, "Image-based defect detection in lithium-ion battery electrode using convolutional neural networks," *J Intell Manuf*, vol. 31, no. 4, pp. 885–897, 2020, doi: 10.1007/s10845-019-01484-x.
- [12] X. Zhan, L. Xu, J. Zhang, and A. Li, "Study on AGVs battery charging strategy for improving utilization," *Procedia CIRP*, vol. 81, pp. 558–563, 2019, doi: 10.1016/j.procir.2019.03.155.
- [13] F. Duffner, L. Mauler, M. Wentker, J. Leker, and M. Winter, "Large-scale automotive battery cell manufacturing: Analyzing strategic and operational effects on manufacturing costs," *International Journal of Production Economics*, vol. 232, p. 107982, 2021, doi: 10.1016/j.ijpe.2020.107982.
- [14] L. Mauler, F. Duffner, W. G. Zeier, and J. Leker, "Battery cost forecasting: a review of methods and results with an outlook to 2050," *Energy Environ. Sci.*, vol. 14, no. 9, pp. 4712–4739, 2021, doi: 10.1039/D1EE01530C.
- [15] Fraunhofer-Institut für System- und Innovationsforschung ISI, *Preisschwankungen bei Batterie-Rohstoffen: Wie die Automobilindustrie reagiert und welche Auswirkungen sie auf die Zellkosten haben*. [Online]. Available: <https://www.isi.fraunhofer.de/de/blog/themen/batterie-update/batterie-rohstoffpreis-schwankungen-wie-reagiert-automobil-industrie-auswirkungen-zellkosten.html> (accessed: Mar. 7 2024).
- [16] N. Urbach and M. Röglinger, Eds., *Digitalization Cases*. Cham: Springer International Publishing, 2019.
- [17] W. van der Aalst, *Process Mining: Data Science in Action*, 2nd ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016. [Online]. Available: <http://nbn-resolving.org/urn:nbn:de:bsz:31-epflicht-1624710>

- [18] C. Di Francescomarino, M. Dumas, F. M. Maggi, and I. Teinemaa, "Clustering-Based Predictive Process Monitoring," *IEEE Trans. Serv. Comput.*, vol. 12, no. 6, pp. 896–909, 2019, doi: 10.1109/TSC.2016.2645153.
- [19] F. M. Maggi, C. Di Francescomarino, M. Dumas, and C. Ghidini, "Predictive Monitoring of Business Processes," in *Lecture Notes in Computer Science*, vol. 8484, *Advanced Information Systems Engineering: 26th International Conference, CAiSE 2014, Thessaloniki, Greece, June 16-20, 2014. Proceedings*, M. Jarke et al., Eds., Cham: Springer International Publishing; Imprint; Springer, 2014, pp. 457–472.
- [20] M. de Leoni, W. M. van der Aalst, and M. Dees, "A general process mining framework for correlating, predicting and clustering dynamic behavior based on event logs," *Information Systems*, vol. 56, pp. 235–257, 2016, doi: 10.1016/j.is.2015.07.003.
- [21] I. Verenich, M. Dumas, M. La Rosa, F. M. Maggi, and C. Di Francescomarino, "Complex Symbolic Sequence Clustering and Multiple Classifiers for Predictive Process Monitoring," in *Lecture notes in business information processing*, vol. 256, *Business process management workshops: BPM 2015, 13th International Workshops, Innsbruck, Austria, August 31-September 3, 2015 : revised papers*, M. Reichert and H. A. Reijers, Eds., Switzerland [Cham]: Springer, 2016, pp. 218–229.
- [22] I. Teinemaa, M. Dumas, M. La Rosa, and F. M. Maggi, "Outcome-Oriented Predictive Process Monitoring," *ACM Trans. Knowl. Discov. Data*, vol. 13, no. 2, pp. 1–57, 2019, doi: 10.1145/3301300.
- [23] I. Verenich, M. Dumas, M. La Rosa, F. M. Maggi, and C. Di Francescomarino, "Minimizing Overprocessing Waste in Business Processes via Predictive Activity Ordering," in *LNCS sublibrary. SL 3, Information systems and applications, incl. Internet/Web, and HCI*, vol. 9694, *Advanced information systems engineering workshops: 28th International Conference, CAiSE 2016, Ljubljana, Slovenia, June 13-17, 2016. Proceedings*, S. Nurcan, P. Soffer, M. Bajec, and J. Eder, Eds., Switzerland: Springer, 2016, pp. 186–202.
- [24] C. Cabanillas, C. Di Ciccio, J. Mendling, and A. Baumgrass, "Predictive Task Monitoring for Business Processes," in *Lecture Notes in Computer Science, Business Process Management*, S. Sadiq, P. Soffer, and H. Völzer, Eds., Cham: Springer Nature, 2014, pp. 424–432.
- [25] A. E. Márquez-Chamorro, M. Resinas, A. Ruiz-Cortés, and M. Toro, "Run-time prediction of business process indicators using evolutionary decision rules," *Expert Systems with Applications*, vol. 87, pp. 1–14, 2017, doi: 10.1016/j.eswa.2017.05.069.
- [26] W. Kratsch, J. Manderscheid, M. Röglinger, and J. Seyfried, "Machine Learning in Business Process Monitoring: A Comparison of Deep Learning and Classical Approaches Used for Outcome Prediction," *Business & Information Systems Engineering*, vol. 63, no. 3, pp. 261–276, 2021, doi: 10.1007/s12599-020-00645-0.
- [27] N. Mehdiyev and P. Fettke, "Local Post-Hoc Explanations for Predictive Process Monitoring in Manufacturing," *ECIS*, 2021. [Online]. Available: <https://arxiv.org/pdf/2009.10513.pdf>

Das Projekt KIProBatt im Kompetenzcluster InZePro

Ziel des Kompetenzclusters Intelligente Batteriezellproduktion (InZePro) ist es, das Produktionssystem der Batteriezellfertigung unter Einsatz von Lösungen der Industrie 4.0 ganzheitlich zu optimieren. Im Projekt KIProBatt haben Projektpartner aus der Forschung an dem Ziel gearbeitet, wie dem Einsatz eines KI-gestützten Prozessmonitorings eine Reduzierung des Ausschusses und eine Steigerung der Zellqualität ermöglicht werden kann.

Übersicht zur Projektseite: <https://inzepro-cluster.de/projekt/kiprobatt/>



Fraunhofer-Gesellschaft
Institut für Gießerei-, Composite- und Verarbeitungstechnik (IGCV)
Am Technologiezentrum 10,
86159 Augsburg
Vertreten durch Dr. Christoph Berger
www.igcv.fraunhofer.de



Fraunhofer-Gesellschaft
Institut für Silicatforschung (ISC)
Neunerplatz 2, 97082 Würzburg
Vertreten durch Dr. Henning Lormann
www.isc.fraunhofer.de



Technische Hochschule Aschaffenburg
Labor für Medizinische IT und Simulation
Würzburger Str. 45,
63743 Aschaffenburg
Vertreten durch
Prof. Dr. Michael Möckel
www.th-ab.de



Universität Bayreuth
für Wirtschaftsinformatik und Wertorientiertes Prozessmanagement
Wittelsbacherring 10, 95444 Bayreuth
Vertreten durch Prof. Dr. Maximilian Röglinger
www.wi.uni-bayreuth.de

