

# VOM DATENSAMMLER ZUM DATENVERSTEHER

**Mit Datenanalyse-Tools können Betreiber verdeckte Zusammenhänge zwischen Prozessparametern aufdecken** – Viele haben vor Data-Analytics-Projekten großen Respekt. Doch man muss nicht immer gleich die ganz großen Räder drehen. Zwei Beispielprojekte zeigen: Wer mit kleinen Schritten startet, schafft eine solide Basis für größere Vorhaben.

STEFAN PAULI\*

Prozessanlagen sind üblicherweise automatisiert und generieren eine Unmenge an Daten über Sensoren (Temperaturen, Durchflüsse, Füllstände, Druck etc.), manuellen Analyse-Eingaben, LIMS usw. Daraus lassen sich mit Trendlinien einfache Zusammenhänge zwischen zwei oder drei Parametern erkennen, eine tiefere Analyse ist jedoch nicht möglich. (Big)Data-Analytics ermöglichen tiefere Einblicke und schaffen neue Zusammenhänge. Die aufgeführten Beispiele zeigen, dass Daten analytische Methoden

teilweise mit geringem Aufwand eine erhebliche Effizienz erreichen können.

## Von Daten zur Analyse bis zu Machine Learning

Dazu müssen in einem ersten Schritt die Daten aus verschiedenen Bereichen einander zugeordnet werden, z.B. via eindeutigen Batch-Nummern und Zeitstempeln. Das können Daten aus Automatisierungstechnisch völlig getrennten Produktionsbereichen sein oder auch aus verschiedenen Automations-Ebenen einer Anlage – SPS, PLS, MES, ERP. Darüber hinaus können zusätzliche Messinstrumente für einige Zeit installiert werden, deren Werte separat

aufgezeichnet werden – heute kabellos sehr einfach möglich. Auch händische Aufzeichnungen eines Betriebshandbuchs, Laboranalysen usw. müssen in eine Datenbank mit einheitlichem Zeitstempel übertragen werden.

Im nächsten Schritt erfolgt die Analyse dieser Daten mit ausgeklügelter Datenbereinigung, modernen Machine-Learning-Algorithmen und aussagekräftigen Visualisierungen. In der Datenbereinigung werden fehlerhafte Messwerte ausgeschlossen, Daten einander zugeordnet, verschiedene Quellen vereint, Datentransformationen vorgenommen (z.B. aus der Produktionszeit errechnen, ob die Tag- oder Nachtschicht produziert). Obwohl moderne Datenana-

Der Anlagenbetrieb liefert Daten in großer Zahl, die Herausforderung ist, die Zusammenhänge zu verstehen.

\* Der Autor ist Senior Data Scientist bei VTU Engineering GmbH, Grambach/Österreich. E-Mail-Kontakt: stefan.pauli@vtu.com



Bild: VTU

lyse-Tools bei dieser Datenbereinigung unterstützende Funktionen bereitstellen, bleibt dies oft der aufwändigste Teil der Analyse. Anschließend sagen die Algorithmen wichtige Zielwerte einer Produktion selbstlernend vorher und alle Nutzer erkennen verdeckte Zusammenhänge zwischen den Zielwerten und den übrigen Produktionsparametern und Daten. Die zielgruppengerechte Präsentation der Analyseresultate ermöglicht einen konstruktiven Austausch zwischen Fach- und Datenspezialisten – ein Erfolgsfaktor für gelungene Datenanalyse-Projekte.

Eine Auswahl von Projekten, welche mit einem Aufwand von weniger als zehn Arbeitstagen bewältigt wurden, verdeutlichen den Nutzen kleinerer (Big)Data-Projekte. Bei den folgenden Projekten wurde auf eine bestehende Datensammlung zugegriffen.

### Schlechte Ausbeuten in der Fermentation

Bei drei neuen baugleichen Fermentern in einem Pharmabetrieb lieferte ein Bioreaktor unter gleichen kontrollierten Bedingungen schlechtere Ausbeuten. Da die Betriebsmannschaft trotz gründlicher Untersuchungen die Ursache nicht finden konnte, wurden die Produktions- und Prozessdaten offline mit dem Machine-Learning-Algorithmus „Random Forest“ ausgewertet: Dieser analysiert gleichzeitig den Einfluss der verschiedenen Produktionsparameter auf die Ausbeute und prognostiziert die Ausbeute der drei Fermenter t (siehe Grafik).

Der deutliche Einfluss des Parameters Substrat-Temperaturwertes (temp\_Sub1) auf die Ausbeute war aus verfahrenstechnischer Sicht unerwartet, da solche Schwankungen nur sehr sporadisch auftraten. Angeregt durch diese neue Information hat ein Verfahreningenieur die Isolierung einer Zulaufleitung genauer untersucht, wobei er einen Isolationsfehler entdeckte, der die Substrat-Temperatur und in der Konsequenz auch die Ausbeute beeinflusste. Mit einer neuen intakten Isolation konnte das Problem einfach behoben werden. Dies ist eines der ein-

drücklichsten Beispiele einer Zusammenarbeit zwischen Fach- und Datenexperten. Denn vor der Analyse hatten Fachexperten erfolglos die Ursache gesucht, noch viel weniger hätten die Datenexperten mit ihren Algorithmen allein diesen Isolationsfehler gefunden. Doch zusammen und durch die Datenanalyse war der Fehler schnell behoben. Mit einem Aufwand von ca. 80 Stunden war dies ein überschaubarer Prozess.

### Predictive Maintenance bei Kompressorschäden

In einer großen Polyolefin-Produktion fiel die Gleitringdichtung des Hauptkompressors in den letzten Jahren mehrfach und unvorhersehbar mit direkten Kosten von über 100 000 Euro aus sowie einigen Tagen Stillstand pro Ausfall.

In nur 40 Stunden wurde eine Offline-Datenanalyse aus vorhandenen Daten von 150 Stell- und Messgrößen entwickelt. Mit dieser Methode wurden Zusammenhänge zwischen Stillständen und einigen der gegebenen Parameter entdeckt. Dank einer Neueinstellung der relevanten Maschinenparameter durch den Hersteller und die Produktion konnten neue Schäden an der Gleitringdichtung bis zum heutigen Zeitpunkt verhindert werden.

### Geringer Aufwand – hoher Nutzen

Die Beispiele zeigen, dass (Big) Data Analytics auch mit geringem Aufwand einen hohen Nutzen bringen. Oft scheuen Betreiber vor datenanalytischen Methoden zurück, weil die Verantwortlichen das Schreckgespenst eines jahrelangen Softwareprojektes vor Augen haben.

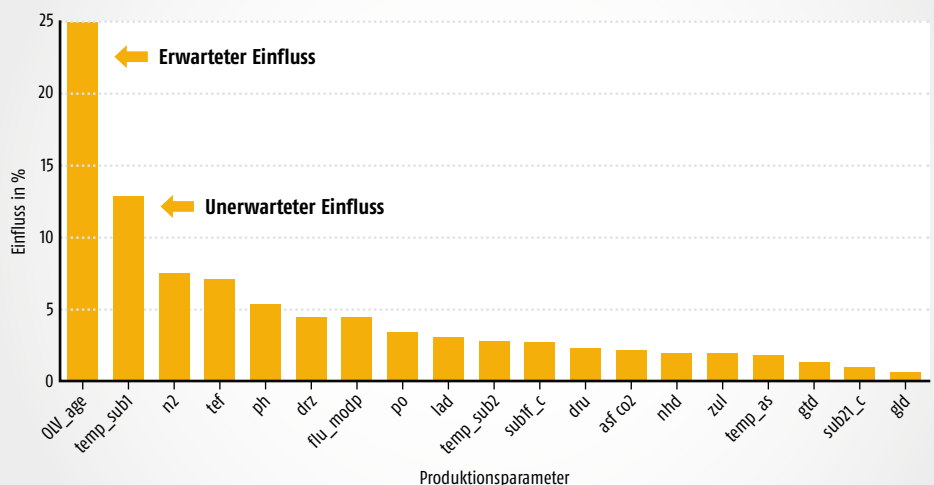
Bei Großprojekten ist ein schrittweises Vorgehen anzuraten. In Piloten, wie den oben beschriebenen, lernt einerseits der Betreiber den Nutzen kennen und andererseits gewinnt der Data Scientist den Überblick über die bestehenden Daten. Die nächsten Schritte zu langfristiger Datennutzung können so fachmännisch aufgesetzt werden.

Weitere Schritte auf Basis von Datenanalysen ergeben sich z.B. aus datengetriebenen Online-Optimierungen der Prozessregelung bis zu ganzen Unternehmensprozessen weit über den Anlagenbetrieb hinaus (Wartung, Lieferketten, Marketing usw.). Die langfristigen Möglichkeiten gehen bis zur selbstlernenden, selbstoptimierenden, selbststeuernden und selbstständig eingreifenden Künstlichen Intelligenz. Doch auch hier sind kleine Schritten ratsam, um später mit dem Erfolg zu wachsen.

#### PROCESS-Tipp

Auf [www.process.de](http://www.process.de) finden Sie eine **ausführliche Version des Beitrags** mit einem weiteren Beispiel.

Einfluss der Produktionsparameter auf die Vorhersage



Der Algorithmus erkennt den Einfluss der verschiedenen Parameter auf seine Vorhersage.

Quelle: VTI