

# Datengetriebene Module für Predictive Maintenance

Betrachtung verschiedener Module für eine datengetriebene, vorausschauende Wartung

Burkhard Hoppenstedt, Rüdiger Pryss, Alexander Treß, Bernd Biechele und Manfred Reichert

Der Wunsch aus Produktionsdaten einen vorausschauenden Wartungszyklus abzuleiten, definiert das Umfeld des Themengebiets Predictive Maintenance. Hierbei unterscheiden sich die jeweiligen Anforderungen an eine Anwendung sehr stark, u. a. aufgrund der zugrunde liegenden Datenbasis, bestehenden Echtzeitanforderungen und vorhandenem Expertenwissen. Im Folgenden wird erörtert, welche Module in den Kreislauf eines Predictive Maintenance Systems einfließen können. Ein Schwerpunkt liegt auf der Datenreduktion mittels der Principal Component Analysis.

Das Themengebiet Predictive Maintenance ist kein neues Thema. Inhaltlich bewegt es sich im Themenumfeld der Automatisierung und kann softwareseitig in ein Manufacturing Execution System integriert werden (vgl. Bild 1). Die Integration ermöglicht es wiederum, dass Ergebnisse aus der Auswertung von Maschinendaten in die Produktionskontrolle miteinbezogen werden können. Im ersten Teil dieses Beitrags stellen wir ein Predictive Maintenance System anhand einer grundlegenden Modulstruktur vor (vgl. Bild 2) und diskutieren ausgewählte Module (vgl. die fettgedruckten Module im Folgenden mit Bild 2). Anschließend stellen wir entlang eines realen Industriedatensatzes eine Methode zur Datenreduktion vor, die einen wichtigen Bestandteil im Kontext von Predictive Maintenance darstellen kann.

## Teil 1: Modulstruktur für Predictive Maintenance Systeme

Aufgrund der umfangreichen Funktionalitäten eines Predictive Maintenance Systems, wurde ein Übersichtsschaubild der Komponenten entwickelt (vgl. Bild 2). Dieses wird im Folgenden erläutert. Zur Einführung eines Predictive Maintenance Systems [1] muss eine Datenbasis des gesamten Maschinenzustandes extrahiert werden. Um dies zu ermöglichen, ist folgende Frage von zentraler Bedeutung: Welche Genauigkeit wird für die Messdaten benötigt? Die Aufgabe des verantwortlichen Sensormoduls liegt darin, (a) einen kontinuierlichen Verlauf der Maschinensignale durch Abtastung zu diskretisieren, (b) die gewünschte Signalaufösung durch Quantisierung festzulegen sowie ggf. (c) eine Metastruktur zur Organisation der Daten zu

identifizieren. Eine Metastruktur könnte entstehen, indem die Maschinensignale mit der Protokollsprache Open Platform Communication Unified Architecture (OPC-UA) softwareseitig wiederum als Maschine abgebildet werden. Um Fehler bei der Übertragung von kontinuierlichen zu diskreten Werten während der Abtastung zu vermeiden, sollte das Nyquist-Shannon-Abtasttheorem [2] beachtet werden, welches eine vollständige Rekonstruktion des Ursprungssignals ermöglicht, indem die Abtastrate größer als die doppelte Ursprungsfrequenz des Datensatzes gewählt wird (d. h.  $f_{\text{abtast}} > 2 \cdot f_{\text{max}(\text{signal})}$ ). Im Bereich von Maschinen mit hochfrequenten Signalen führt dies zu sehr hohen Anforderungen an die Sensorik. Mittels der Quantisierung kann eine verlustbehaftete Kompression durchgeführt werden. Durch das Auf- oder Abrunden des Signalwertes auf die nächste Quantisierungsstufe, welche durch die Schrittweite  $\Delta$  ermittelt wird, ergibt sich der verwendete diskrete Wert. Der Quantisierungsfehler wird durch die mittlere quadratische Abweichung (engl. MSE – Mean Square Error) angegeben. Nach diesen zwei Schritten liegt der gemessene Wert digital im System vor und kann durch Meta-Informationen wie Maschineninformationen (z. B. Laufzeit) ergänzt werden. Um die Sinnhaftigkeit der erhaltenen Werte zu überprüfen, wird optional das Modul Data Quality ausgeführt. Folgende Fragestellungen sind in diesem Kontext relevant:

- **Consistency:** Sind die Daten konsistent?
- **Validity:** Passt die Intervallzuordnung und der Datentyp?
- **Completeness:** Sind alle Sensorwerte gesetzt oder gab es NULL-Einträge?

In diesem Beitrag lesen Sie:

- ✓ aus welchen Modulen ein Predictive Maintenance System aufgebaut sein kann,
- ✓ wie sich diese Module miteinander kombinieren lassen und
- ✓ wie eine Datenreduktion mit der Principal Component Analysis auf Maschinendaten aussehen kann.

**M. Sc. Burkhard Hoppenstedt** ist externer wissenschaftlicher Mitarbeiter am Institut für Datenbanken und Informationssysteme (DBIS) der Universität Ulm. Seine Promotion befasst sich mit Predictive Methoden im Umfeld von Manufacturing Execution Systems.

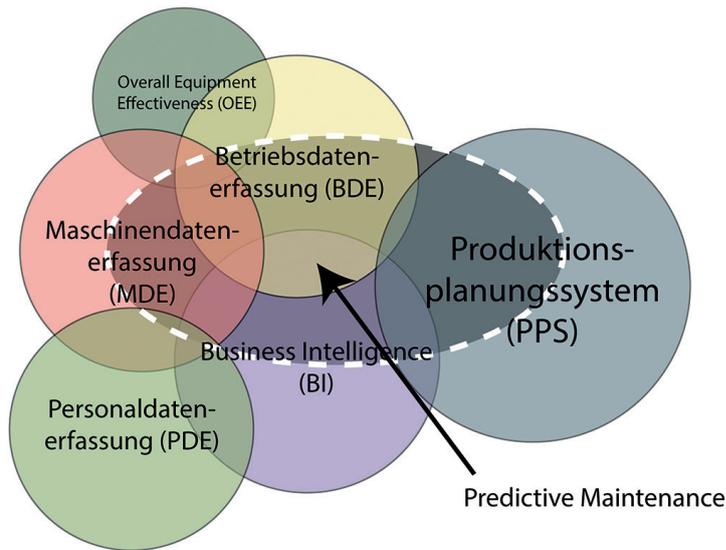
**Dr. Rüdiger Pryss** ist Habilitand am Institut für Datenbanken und Informationssysteme und beschäftigt sich in seiner Forschung mit der innovativen Datensammlung durch den Einsatz mobiler Technologien und ihrer Verknüpfung mit prozessorientierten Informationssystemen.

**Dipl.-Inf. Alexander Treß** ist Geschäftsführer der atr Software GmbH, deren Spezialgebiete die Automatisierungstechnik und Produktionsplanungssysteme sind.

**Prof. Dr. Manfred Reichert** ist Direktor des Instituts für Datenbanken und Informationssysteme an der Universität Ulm. In seiner Forschung entwickelt er adaptive Prozess Management Technologien der nächsten Generation.

**Dipl.-Inf. Bernd Biechele** leitet die Softwareentwicklung der EOS GmbH, die industrielle 3D-Drucker herstellt.

[www.uni-ulm.de/dbis](http://www.uni-ulm.de/dbis)  
[www.atr-software.de](http://www.atr-software.de)



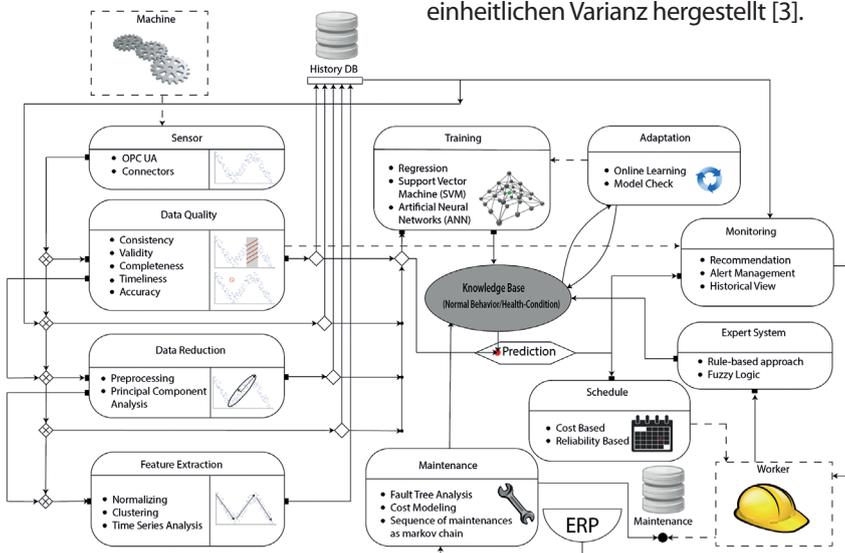
**Bild 1: Predictive Maintenance kann Teil eines MES sein. Jedoch ist es kein isoliertes Modul, sondern es nutzt, erweitert und steuert vorhandene Module.**

- **Timeliness:** Lassen die Zeitstempel Rückschlüsse auf Verzögerungen zu?
- **Accuracy:** Ist die Auflösung gut genug gewählt worden (vgl. MSE)?

Sollte eine dieser Bedingungen verletzt sein, kann ein Datensatz in der Datenbank als korrupt markiert werden und ein Hinweis an das Monitoringmodul gesendet werden. In einem Dashboard werden Verlaufsinformationen, Hinweise und Empfehlungen visualisiert. Eine Verletzung der Completeness-Bedingung könnte beispielsweise durch den Hinweis aufbereitet werden, dass nicht alle benötigten Sensoren angeschlossen sind.

Um Zusammenhänge innerhalb der Signale zu erkennen, muss eine Feature Extraction angewandt werden. Eine Datenkurve wird zuerst in Teilkurven unterteilt und anschließend mit verschiedenen Methoden des Data-Minings oder der Zeitreihenanalyse analysiert. Eine Vergleichbarkeit der Daten wird durch die Normalisierung zu zero mean in Kombination mit einer einheitlichen Varianz hergestellt [3].

**Bild 2: Module eines leistungsfähigen Predictive Maintenance Systems.**



Um ein System „predictive“ zu machen, muss zunächst eine Wissensbasis anhand der vergangenen Daten angeleitet werden, um darauf basierend Prognosen für die weitere Entwicklung aufzustellen. Ein üblicher Ansatz für das Anlernen bzw. Training stellen Neuronale Netze [4] dar, welche eine Inputmenge mittels einer variablen Anzahl von Verbindungsknoten auf eine Ausgabe verrechnen. Die Gewichtung der Zwischenknoten kann über eine Rückwärtsberechnung antrainiert werden. Mittels einer Brute Force-Testmethode kann zudem die Anzahl der Verbindungsknoten ermittelt werden, indem das beste Resultat aus einem getesteten Intervall (vgl. [3]) verwendet wird. An dieser Stelle ist jedoch Vorsicht angebracht, da ein intensiviertes Lernen nicht immer besser ist. So kann es passieren, dass bei verstärktem Lernen viele, auch nebensächliche, Aspekte antrainiert werden. Bei iterativen Lernalgorithmen kann hier die sogenannte Early Stopping Regularization eingesetzt werden, um diesen Effekt zu verhindern. Sequentielles Anlernen ist wesentlich komplexer als Lernen auf einem ganzen Datensatz, weil das Entfernen des Rauschens (Noise) und das nachträgliche Zuordnen zu einem Cluster erschwert werden [5].

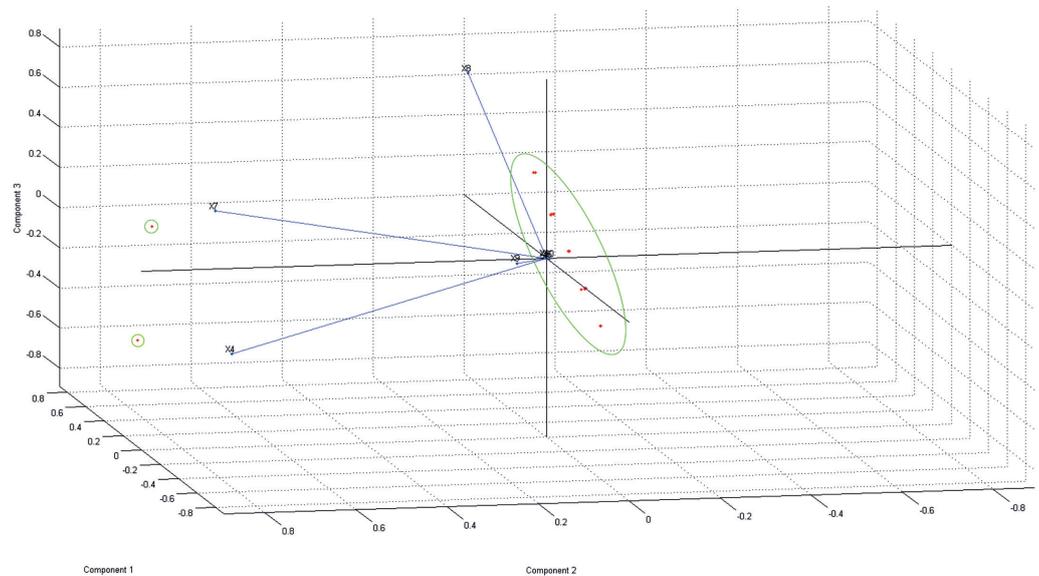
Alle nun berechneten Informationen werden in einer Wissensbasis (Knowledge Base) verwaltet. Der ermittelte Zustand der Maschine und der Abgleich der aktuellen Daten mit einem antrainierten Normalzustand ermöglichen wiederum eine Vorhersage. Es ist an dieser Stelle ratsam, menschliches Expertenwissen mit in die Wissensbasis aufzunehmen. Für die Modellierung von Expertenwissen ist die sogenannte Fuzzy Logic [6] ein sinnvolles Instrument, weil viele Regeln ineinandergreifen und nicht klar gegeneinander abgegrenzt werden können.

Mithilfe der entstandenen Wissensbasis, die sich aus dem Dateninput und dem Expertenwissen speist, wird ein Wartungszyklus aufgebaut. Die Produktionsoptimierung kann entweder eine quantifizierte oder eine qualitative Zielsetzung haben. Während im quantifizierten Ansatz die Maximierung der Verfügbarkeit bei gegebenen Instandhaltungskosten das Ziel ist, wird im qualitativen Ansatz die Reduktion der Instandhaltungskosten bei Erhalt der Verfügbarkeit in den Vordergrund gestellt.

Teil 2: Datenreduktion im konkreten Anwendungsszenario

Die auszuwertenden Daten liefert eine industrielle 3D-Druckmaschine, von welcher zehn messende Sensoren untersucht wurden. Die Sensoren liefern Positions-, Geschwindigkeits-, Temperatur-, Druck- und Gaskonzentrations-

angaben. Ein 3D-Drucker trägt verschiedene Schichten auf und pro Schicht werden die Sensorwerte abgespeichert. Aufgrund der enormen Datenmenge empfiehlt sich eine Datenreduktion, um spätere, rechenaufwändige Berechnungen, wie z. B. das Training eines Neuronales Netzes, möglichst schnell durchzuführen. Eine Transformation bildet die ursprünglichen Datenpunkte in einen subdimensionalen Raum ab (vgl. Bild 3). Diese Methode der Hauptkomponentenanalyse (Principal Component Analysis) [7] bietet zwei Vorteile: Erstens kann eine Dimensionsreduktion der Daten vorgenommen werden (Kompression); Zweitens lassen sich Zusammenhänge innerhalb der einzelnen Eingabedaten feststellen.



**Bild 3:** Die ursprüngliche 10-dimensionale Datenmenge wurde in das neue Koordinatensystem, welches durch die drei blau markierten Hauptkomponenten ( $x_1, x_2, x_3$ ) aufgespannt ist, übertragen. Der Großteil der 34 Datenpunkte bewegt sich entlang der Hauptkomponente  $x_3$ . Zwei Ausreißer (grün eingekreist) sorgen für die große Varianzabdeckung von  $x_1$  und  $x_2$ .

Betrachten wir nun einen exemplarischen Druckauftrag mit 34 Schichten: Dieser liefert die Ausgangsmatrix  $M(34,10)$  mit jedem Zeilen eintrag, definiert durch die zehn Sensorwerte einer Schicht. Als Ergebnis der PCA erhalten wir nun die zehn Hauptkomponenten und eine neue Matrix  $M'$  derselben Dimension wie  $M$  mit den transformierten Datenpunkten. Für jede Hauptkomponente ist ihr Anteil an der Varianz bekannt. Wenn wir die kumulative Summe für den absteigend geordneten Vektor durch die Summe der Varianz teilen, erhalten wir die prozentuale Abdeckung der Gesamtvarianz  $p$  nach  $n$  Komponenten, hier mit folgendem praktischen Ergebnis:  $\{p_1:41.600\%, p_2:72.620\%, p_3:0.993\%, p_4:99.900\%, [...] p_{10}:100.000\%...\}$ . Mit einer Hauptkomponente können wir 41.6% der Ursprungsvarianz abbilden. Ein zufriedenstellendes Ergebnis erhalten wir nach Erreichen der dritten Hauptkomponente.

Somit wissen wir nun, dass bereits mit drei Hauptkomponenten 99% der Varianz abgedeckt werden können. Diese sehr starke Reduktion entstand durch eine hohe Anzahl von Konstanten in den Sensorergebnissen.

Entstandene Vorteile durch die Principal Component Analysis:

- Hochdimensionale Daten können nun mit wenig Informationsverlust im zwei- oder drei-dimensionalen Raum visualisiert werden.
- Durch eine Kompression wird die Anwendung schneller, was im Idealfall sogar Echtzeitanwendungen auf großen Datensätzen ermöglicht.

- Ausreißer in den Datenpunkten (möglicherweise abnormales Verhalten der Maschine) lassen sich leichter durch die vorherige Eliminierung von Daten mit wenig Informationsgehalt ausfindig machen.

### Ausblick

Ein Predictive Maintenance System lässt sich nur sehr schwer generalisieren. Wie in diesem Beitrag gezeigt wurde, kann es aus vielen verschiedenen Modulen bestehen, die miteinander verknüpft werden müssen. Optionale Module, die z.B. die Überprüfung der Datenqualität sicherstellen, bringen neue Aspekte mit sich, wirken sich aber evtl. schlecht auf die Gesamtlaufzeit aus. Ein zukünftiges Ziel sollte es daher sein, in der Produktion die Schnittstellen zu den einzelnen Modulen zu standardisieren, um sie (a) einfacher verknüpfen zu können und (b) mit anderen Anwendungen verwenden zu können, um schließlich die Produktion nachhaltig zu verbessern.

### Schlüsselwörter:

Predictive Maintenance, neuronale Netze, Hauptkomponentenanalyse, Fuzzy Logic

### Literatur:

- [1] Mobley, R.: Introduction to Predictive Maintenance. Woburn 2002.
- [2] Nyquist, H.: Certain Topics in Telegraph Transmission Theory. Oregon 1928.
- [3] Bansal, D.: A real-time predictive maintenance system for machine systems. Birmingham 2004.
- [4] Fausett, L.: Fundamentals of neural networks: architectures, algorithms, and applications, Melbourne 1994.
- [5] Amadou Boubacar, H.: SAKM: Self-adaptive kernel machine A kernel-based algorithm for online clusterin. Douai 2008.
- [6] Simpson, P.: Fuzzy min-max neural networks -- Part 2: Clustering, San Diego 1993.
- [7] Pearson, K.: On lines and planes of closest fit to a system of points in space, London 1901.

### Data-Driven Modules in the Context of Predictive Maintenance

The goal to identify an optimal maintenance schedule defines the requirements for predictive maintenance. The demands for this application type vary significantly due to different databases, real-time requirements and existing expert knowledge levels. In this article, different modules of a Predictive Maintenance Application are explained and the Principal Component Analysis is discussed along industrial data.

### Keywords:

Predictive Maintenance, Neural Networks, Principal Component Analysis, Fuzzy Logic