



Leibniz Universität Hannover
Wirtschaftswissenschaftliche Fakultät
Institut für Wirtschaftsinformatik

Digital Twin zur Ausfallursachenanalyse von Maschinenantriebslösungen

Masterarbeit

zur Erlangung des akademischen Grades „Master of Science (M. Sc.)“ im Studiengang Wirtschaftsingenieur der Fakultät für Elektrotechnik und Informatik, Fakultät für Maschinenbau und der Wirtschaftswissenschaftlichen Fakultät der Leibniz Universität Hannover

vorgelegt von

Vanessa Gurdeep Dass



Prüfer: Prof. Dr. Michael H. Breitner
Betreuer: Daniel Olivotti

Hannover, 03. April 2018

Inhaltsverzeichnis

Abbildungsverzeichnis	i
Tabellenverzeichnis	iii
Formel- und Abkürzungsverzeichnis.....	iv
1 Einleitung.....	1
2 Grundlagen	3
2.1 Digital Twin	3
2.2 Ausfallursachenanalyse	5
2.2.1 Szenarien der Wartung und Instandhaltung	6
2.2.2 Root Cause Analysis.....	8
2.3 Knowledge Discovery in Databases	9
2.3.1 Geschäftsverständnis.....	12
2.3.2 Datenverständnis	13
2.3.3 Datenvorverarbeitung.....	14
2.3.4 Modellierung	15
2.3.5 Evaluierung	16
2.3.6 Bereitstellung und Nutzung	17
2.4 Data Mining.....	18
2.4.1 Aufgaben	18
2.4.2 Interdisziplinarität	23
2.4.3 Statistik	25
2.4.4 Machine Learning.....	27
2.4.5 Datenbanksysteme	31
3 Literaturübersicht.....	38
4 Konzept für die Ausfallursachenanalyse	43
4.1 Geschäftsverständnis	43
4.1.1 Bestimmung der Geschäftsziele.....	43
4.1.2 Bewertung der Situation	43
4.1.3 Bestimmung der Data Mining Ziele	46
4.1.4 Erstellung eines Projektplans	47
4.2 Datenverständnis	57
4.2.1 Datenerfassung.....	58
4.2.2 Datenbeschreibung	61
4.2.3 Untersuchung der Daten	64
4.2.4 Verifizierung der Datenqualität	66
4.3 Datenvorverarbeitung	67
4.3.1 Datenauswahl	67
4.3.2 Datenbereinigung.....	69
4.3.3 Datenkonstruktion	72

4.3.4	Datenintegration.....	72
4.3.5	Datenformatierung	74
4.4	Modellierung	76
4.4.1	Auswahl der Modellierungstechnik	76
4.4.2	Erstellung eines Testdesigns	80
4.4.3	Modellentwicklung.....	83
4.4.4	Modellbewertung.....	87
4.5	Evaluierung.....	91
4.5.1	Auswertung der Ergebnisse	91
4.5.2	Überprüfung des Prozesses und Bestimmung der nächsten Schritte	93
4.6	Bereitstellung und Nutzung	94
4.6.1	Planung der Bereitstellung	94
4.6.2	Planung der Überwachung und Wartung.....	96
4.6.3	Erstellung des Abschlussberichts und Überprüfung des Projekts	98
5	Proof of Concept	99
6	Handlungsempfehlung und Diskussion	108
7	Limitationen	110
8	Fazit und Ausblick.....	112
9	Literaturverzeichnis	I
Anhang	VI
Ehrenwörtliche Erklärung	XIX

1 Einleitung

„In the long run, all machines break down.“
John Maynard Keynes; Ökonom, Politiker und Mathematiker

Dieses berühmten Zitat des britischen Ökonomen John Maynard Keynes, erscheint zunächst sehr dogmatisch, jedoch steckt darin ein Kern der Wahrheit. In einer Studie des Wall Street Journals und Emerson konnte festgestellt werden, dass ungeplante Stillstandzeiten industrieller Hersteller jährlich Kosten in Höhe von etwa 50 Milliarden US-Dollar verursachen. Etwa 42 % der Stillstandzeit kann auf Ausfälle zurückgeführt werden. Zur Vermeidung von Ausfällen ist die Planung einer Wartungs- und Instandhaltungsstrategie folglich von großer Bedeutung.

Sowohl international als auch in der deutschen Wirtschaft führt der digitale Wandel zur Veränderungen der Marktbedingungen. Unternehmen sind dazu gezwungen ihre Geschäftsmodelle anzupassen, um auf dem Markt bestehen zu können. Vor allem im Maschinen- und Anlagenbau wird dieser Wandel deutlich. Im produzierenden Betrieb werden hohe Anforderungen an die Zuverlässigkeit und Effizienz der Maschinen und Anlagen gestellt. Die wichtigsten Bestandteile sind nicht mehr nur die Produktion und der Vertrieb der Produkte, sondern auch die produktergänzende Dienstleistung. Für Maschinen- und Anlagenbauer sowie für Komponentenzulieferer gewinnt die produktergänzende Dienstleistung in Form von Sicherstellung des Betriebs und der Durchführung von Instandhaltung und Wartung immer mehr an Bedeutung. Zulieferer von Komponenten für den Maschinen- und Anlagenbau stehen im Zuge der wachsenden Bedeutung von Industrie 4.0 vor der Herausforderung ihr Geschäftsmodell auf weitreichendere Dienstleistungen anzupassen. Vor allem dann, wenn Stillstände große Einflüsse auf den laufenden Betrieb haben und zu großen Einbußen führen, ist es wichtig Ausfallursachen schnell zu identifizieren oder diesen bestenfalls sogar durch vorausschauende Wartung entgegenzuwirken.

Verwandte Arbeiten, welche sich mit einer ähnlichen Thematik befassen, fokussieren sich auf die Problemlösung mittels der Anwendung verschiedener Analysemethoden. Vorgelagerte Phasen des Verständnisgewinns werden zwar oftmals einbezogen, jedoch meist lediglich oberflächlich angeschnitten, sowie auch die nachgelagerten Phasen, der Bereitstellung und Nutzung. Die Grundlage für die Analyse bilden die verfügbaren Daten. Eine Kombination von Sensordaten und komponentenbezogenen Daten, welche auch als Asset-Daten bezeichnet werden, erfolgt nur in wenigen Fällen. Zudem

fehlt es oftmals an der Übertragbarkeit der Analysemodelle, da diese lediglich für spezifische Maschinen entwickelt sind. Die Forschungsfrage wird daher wie folgt formuliert:

„Wie sieht ein Konzept für ein generisches Modell zur Ausfallursachenanalyse aus, welches sowohl alle Phasen des KDD-Prozesses und zudem neben den Sensordaten domänenspezifisches Wissen aus den Asset-Daten einbezieht?“

Zur Beantwortung werden im zweiten Kapitel der Arbeit zunächst die wesentlichen Begriffe definiert und für das Verständnis relevante Grundlagen vorgestellt. Mittels einer Literaturrecherche nach WEBSTER & WATSON erfolgt in Kapitel 3 eine Sichtung der Literatur, welche sich mit ähnlichen Themenstellungen befasst, um anhand dessen die Forschungslücke zu identifizieren und die Forschungsfrage ableiten zu können. Basierend auf den erarbeiteten Grundlagen, wird im Kapitel 4 ein allgemeines Konzept für die Ausfallursachenanalyse von Maschinenantriebslösungen entwickelt, welches neben den Sensordaten auch Asset-Daten einbezieht und generisch anwendbar ist. Mithilfe des Vorgehensmodells CRISP-DM werden die einzelnen Phasen und generischen Aufgaben für die Erstellung des Konzepts durchlaufen und beispielhaft an einem Versuchsaufbau der Firma Lenze beschrieben. Anschließend wird in Kapitel 5 innerhalb eines Proof of Concept das Konzept anhand von Versuchsdaten, welche in einem Komponententestfeld aufgezeichnet werden, vorgestellt. In Kapitel 6 werden die Ergebnisse kritisch diskutiert, Limitationen werden aufgezeigt und mögliche Handlungsempfehlungen, welche sich aus der Arbeit ableiten lassen, werden gegeben. Zuletzt erfolgt in Kapitel 7 eine Zusammenfassung der Arbeit und ein Ausblick bezüglich weiterer Forschungsansätze.

8 Fazit und Ausblick

Abschließend werden im Folgenden die Ergebnisse der Arbeit zusammengefasst und ein Ausblick für weiterführende Forschungsansätze gegeben.

Im ersten Teil der Arbeit wurden zunächst die wichtigsten Grundlagen erarbeitet. Mittels einer Literaturrecherche nach WEBSTER & WATSON konnte anschließend die Forschungslücke identifiziert werden. Die daraus resultierende Forschungsfrage befasst sich mit der Entwicklung eines allgemeinen Konzepts für die Ausfallursachenanalyse, das neben den Sensordaten auch Asset-Daten, d.h. komponentenbezogene Informationen einbezieht und generisch anwendbar sein soll. Zur Lösung der Forschungsfrage wurde basierend auf dem Verständnis, das durch die Erarbeitung der Grundlagen, geschaffen wurde, im Hauptteil das allgemeine Konzept für die Ausfallursachenanalyse entwickelt. Dafür wurden mittels des Vorgehensmodells CRISP-DM die einzelnen Phasen und generischen Aufgaben durchlaufen und beispielhaft an einem Versuchsaufbau der Firma Lenze beschrieben. Anschließend wurde innerhalb eines Proof of Concepts das Konzept anhand von Versuchsdaten, welche in einem Komponententestfeld aufgezeichnet werden, vorgestellt.

Die Anforderung an den Proof of Concept besteht darin, für eine beschränkte Auswahl an Fehlerfällen eine Klassifizierung vornehmen zu können. Neben dem Normalverlauf wurden drei verschiedene Störfälle aufgezeichnet. Durch die Einstellung des Luftspaltes an der Haltebremse konnte eine leichte Reibung, eine erhöhte Reibung und ein Blockieren des Antriebs simuliert werden. Bei der Auswertung der vier Klassen konnte festgestellt werden, dass vor allem die Klassifizierung des Normalverlaufs eine sehr hohe Genauigkeit liefert. Dies kann mit der Datenmenge, welche für das Training verwendet wurde in Zusammenhang gebracht werden. Ebenso erzielten die Klassifizierung des blockierten Antriebs und der erhöhten Reibung eine sehr gute Genauigkeit. Bei der Klassifizierung der leichten Reibung fällt auf, dass fälschlicherweise häufig die erhöhte Reibung klassifiziert wurde. Um ein besseres Ergebnis zu erzielen, könnte es daher sinnvoll sein, weitere Datensätze dieser beiden Fehlerklassen aufzuzeichnen, um dadurch die Varianz der Verläufe möglichst gut abzudecken und schließlich die Klassifikationsgüte zu erhöhen.

Durch die Festlegung der Verwendung des Vorgehensmodells CRISP-DM wurde von vornherein eine Vielzahl anderer relevanter Ansätze und Techniken ausgeschlossen. Vor allem im Bereich der Ausfallursachenanalyse könnte es jedoch ratsam sein, wei-

tere Ansätze, wie bspw. modellbasierte Ansätze aus dem Bereich der Ingenieurwissenschaften, einzubeziehen. Dies könnte dazu beitragen weitere relevanten Erkenntnisse zu gewinnen.

Für die Entwicklung eines validen Modells, welches eine gute Klassifikationsgüte aufweist, ist es notwendig eine Vielzahl an Daten für das Trainieren, Validieren und Testen bereitzustellen. Für das Komponententestfeld wurde nur eine begrenzte Anzahl an Messwerte für vier Klassen aufgezeichnet. Die Aufzeichnung weiterer Werte könnten zu einer besseren Klassifikationsgenauigkeit und einer verbesserten Robustheit führen. Zudem wurden bei der Konzeptentwicklung nur zwei Fehler mit jeweils fünf Ursachen betrachtet. In der Praxis geht die Anzahl an möglichen Fehlern und Ursachen jedoch meist weit darüber hinaus und es treten oftmals Umstandsänderungen auf, welche eine Anpassung des Modells bedingen. Bei der Durchführung des Konzeptes am beschriebenen Versuchsaufbau sollte in Betracht gezogen werden, weitere signifikante Ausfallursachen zu integrieren.

Für eine zuverlässige Vorhersage der Ausfallursache sollten lediglich störungsbehaftete Sequenzen des Datensatzes für das Training des Modells verwendet werden, da das Modell sonst falsch angelernt wird und somit zur Verfälschung des Klassifizierungsergebnisses führt. Es könnte deshalb vielversprechend sein eine vorgelagerte Identifikation und Kennzeichnung der für den Ausfall verantwortlichen Sequenz vorzunehmen, sodass im Anschluss ein valides Modell zur Klassifizierung gemäß dem Konzept der Ausfallursachenanalyse entwickelt werden kann.