MASTERARBEIT

EINFLUSSFAKTOREN DES DATA MININGS AUF PREDICTIVE MAINTENANCE FÜR EIN WAREHOUSE MANAGEMANT SYSTEM IN DER LOGISTIKBRANCHE

ausgeführt am

Studiengang
Informationstechnologien und Wirtschaftsinformatik

Von: Ing. Stefano Repolust, BSc
Personenkennzeichen: 1910320016

Graz, am 26. März 2021 .....

Unterschrift

EHRENWÖRTLICHE ERKLÄRUNG

Ich erkläre ehrenwörtlich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen nicht benützt und die benutzten Quellen wörtlich zitiert sowie inhaltlich entnommene Stellen als solche kenntlich gemacht habe.

.....

Unterschrift

DANKSAGUNG

Bedanken möchte ich mich bei meiner Freundin und meiner Familie, da sie mich während des berufsbegleitenden Studiums tatkräftig unterstützen und mir auch in schwierigen Zeiten den nötigen Rückhalt gaben.

Großer Dank gilt auch meinem Betreuer, Herrn Dipl.-Ing. Dr. techn. Selver Softic, BSc. Durch seine kurzen Antwortzeiten und sein konstruktives Feedback wurde meine Abschlussarbeit erst möglich.

Des Weiteren möchte ich mich noch bei allen Freunden, Studienkollegen und Studienkolleginnen für eine schöne Studienzeit bedanken.

KURZFASSUNG

Durch Trends wie die Digitalisierung, immer verfügbare Online-Shops und Bestellsysteme, die Nutzung von neuen Technologien und größer werdende Datenmengen, gab es in den letzten Jahren eine massive Veränderung der Logistik. Dem hinzukommend verursachte eine weltweite Pandemie einen immensen Druck auf die Logistikbranche . Szenarien wie Ausfälle der Lieferketten, verursacht durch Maschinen oder Softwaresysteme, müssen vermieden werden . Deswegen versucht die Logistikwirtschaft neue Technologien wie Predictive Maintenance und Data Mining in deren Warehouse Management Systeme einfließen zu lassen, um eine nahezu unterbrechungsfreie Logistiklieferkette gewährleisten zu können.

Aus diesem Grund werden in der vorliegenden Arbeit die Einflussfaktoren und Potentiale des Data Minings auf Predictive Maintenance in einem Warehouse Management System evaluiert.

Um diese Einflussfaktoren des Data Minings zu finden, wurden im ersten Schritt Data Mining Verfahren und deren Einsatzzwecke in der Logistik mit Hilfe einer Literaturrecherche elaboriert. Des Weiteren wurden die Einsatzgebiete von Predictive Maintenance in einem Warehouse Management System seitens Hardware aber auch Software näher betrachtet.

Die im Theorieteil der Arbeit erkannten Einflussfaktoren und Potentiale des Data Minings auf Predictive Maintenance in einem Warehouse Management System wurden mittels Interviews mit Experten und Expertinnen aus der Logistikbranche durchgeführt, ausgewertet und evaluiert. Der zentrale Einflussfaktor ,welcher von allen Experten und Expertinnen genannt wurde, ist das aktive Benutzen und Verwerten des Wissens sowie die gewonnenen Erkenntnisse durch das Data Mining.

Schließlich konnten, durch die Auswertung der Interviews konnten, folgende Einflussfaktoren und Potentiale erarbeitet werden:

- Mittels IOT Geräten, welche an Maschinen und Fördertechnik angebracht werden, können frühzeitig Störungen mittels Data Mining Techniken erkannt und aufgezeigt werden.
- Der Einsatz der Regression auf den Durchsatz eines Systems mittels historischer Daten, um eine Veränderung der Durchlaufzeiten zu erkennen und somit frühzeitig eine Störung oder ein Problem vorauszusagen.

ABSTRACT

Due to trends such as digitalization, increasingly available online stores and ordering systems, use of new technologies and increasing amounts of data, there has been a massive change in logistics in recent years. In addition, a worldwide pandemic put immense pressure on the logistics industry, and scenarios such as supply chain failures caused by machines or software systems were hardly allowed to occur. Therefore, the logistics industry is trying to incorporate new technologies such as predictive maintenance and data mining into their warehouse management systems to ensure an almost uninterrupted logistics supply chain.

For this reason, this thesis evaluates the influencing factors and potentials of data mining on predictive maintenance in a warehouse management system. In order to identify the factors that influence data mining, a literature search was first conducted to investigate data mining methods and their application in logistics. Furthermore, the application areas for predictive maintenance in a warehouse management system were examined more closely in terms of both hardware and software.

The influencing factors and potentials of data mining for predictive maintenance in a warehouse management system identified in the theoretical part of the thesis were then evaluated by means of interviews with experts from the logistics industry. One important finding which all of the experts recognized as an influencing factor is the active utilization of the knowledge and insights gained through data mining.

An evaluation of the interviews yielded the following influencing factors and potentials:

- IOT devices attached to machines and conveyor systems can be used to detect malfunctions at an early stage by means of data mining techniques.
- Use of regression on the throughput of a system using historical data to detect a change in throughput times and thus predict a disruption or problem at an early stage.

INHALTSVERZEICHNIS

1 VERFAHREN DER DATENANALYSE 2

1.1 Zielsetzung 3
1.2 Methode 4
1.3 Vorgehensweise 4
1.4 Aufbau der Arbeit 5

2 GRUNDLAGEN UND BEGRIFFSDEFINITION 8

2.1 Grundlagen des Data Minings 8
2.1.1 Knowledge Discovery in Databases Prozess 9
2.1.2 Cross-Industry Standard Process for Data Mining 11

2.2 Grundlagen Maintenance 13
2.3 Grundlagen Warehouse Management System 16
2.3.1 Basisdaten und Logistische Kennzahlen 17

2.4 Zusammenfassung 18

3 DATA MINING VERFAHREN 20

3.1 Klassifikationsverfahren 20
3.1.1 Entscheidungsbäume 21
3.1.2 Nearest Neighbour Methoden 22

3.2 Regression 22
3.3 Assoziationsverfahren 24
3.4 Clusterverfahren 25
3.5 Zusammenfassung 26

4 PREDICTIVE MAINTENANCE 27

4.1 Verwendung von Predictive Maintenance 28
4.1.1 Maintenance Management Tool 28
4.1.2 Werkzeug zur Systemoptimierung 28
4.1.3 Werkzeug für Verbesserung der Zuverlässigkeit 29

4.2 Predictive Maintenance in Soft- und Hardware 29
4.2.1 Hardware Techniken 30
4.2.2 Software 31

4.3 Zusammenfassung 35

5 ANWENDUNGSGEBIETE 36

5.1 Datenerhebung in der Logistik 36
5.2 Einsatzgebiete in der Logistikbranche 36
5.3 Zusammenfassung 37

6 ZUSAMMENFASSUNG THEORIETEIL 38

7 EINLEITUNG PRAXISTEIL 40

7.1 Merkmale der Logistikbranche 40
7.2 Anforderungen der Logistikbranche 41
7.3 Herausforderungen der Logistikbranche 43
7.4 Einflussfaktoren von Data Mining auf Predictive Maintenance 45
7.5 Hypothesen 46
7.6 Zusammenfassung 46

8 ERSTELLUNG DES INTERVIEWLEITFADENS 48

8.1 Allgemeine Vorgehensweise und Aufbau bei der Erstellung eines Interviewleitfadens für qualitative Interviews 48
8.2 Interviewleitfaden für Interviews mit Experten und Expertinnen 50
8.3 Zusammenfassung 52

9 DURCHFÜHRUNG UND EVALUIERUNG DER INTERVIEWS 54

9.1 Auswahl von Experten und Expertinnen aus der Logistikbranche 54
9.1.1 Interviewpartner 1 54
9.1.2 Interviewpartner 2 54
9.1.3 Interviewpartner 3 55
9.1.4 Interviewpartner 4 55

9.2 Allgemeine Auswertung von qualitativen Interviews 55
9.3 Auswertung der Interviews mit den Logistikexperten 57

9.4 Evaluierung der Interviews mit Experten und Expertinnen aus der Logistikbranche 57
9.5 Zusammenfassung 63

10 ERGEBNISSE 64

10.1 Ergebnisse aus der Literatur 64
10.2 Ergebnisse aus den Interviews 65
10.2.1 Ergebnis der Alternativ- und Nullhypothese aus den Interviews 65
10.2.2 Ergebnis aus der Hypothese 2 65
10.2.3 Ergebnis aus der Hypothese 3 66
10.2.4 Ergebnis aus der Hypothese 4 66

10.3 Fazit und Zusammenfassung 66
10.4 Ausblick 68

ANHANG A - INTERVIEWLEITFADEN FÜR EXPERTENINTERVIEWS AUS DEM LOGISTIKBEREICH 69

ABKÜRZUNGSVERZEICHNIS 74

ABBILDUNGSVERZEICHNIS 75

TABELLENVERZEICHNIS 76

LITERATURVERZEICHNIS 77

1 VERFAHREN DER DATENANALYSE

„In der heutigen Zeit ist die Datenerhebung, Datenspeicherung und Datenverwaltung in jeder erdenklichen Form und Größenordnung dank moderner Informationstechnologien selbstverständlich.“ (Nakhaeizadeh 1998) Diese Aussage von Nakhaeizadeh widerlegte bereits 1998, dass die Datenspeicherung und Datenerhebung vor 22 Jahren bereits zum Standard gehört. Doch der menschliche Aufwand, welcher hierfür betrieben wird, ist immens. Laut Data Center Knowledge (2017) werden bis zu 150 Arbeitskräfte an einem Datenzentrum Standort von Google beschäftigt. Rechnet man auf die Anzahl an Datenzentren hoch, welche von Google weltweit betrieben werden, ist das ein enormes Aufkommen an Arbeitskräften für die Erfassung und Verwaltung von Daten. Dem hinzukommend soll laut Reinsel (2017) das jährliche generierte Datenvolumen bis zum Jahr 2025 auf 163 Zettabyte ansteigen. Von diesen jährlich generierten Daten wurden im Jahr 2012 laut Gantz und Reinsel (2012) jedoch lediglich 0,5 % analysiert, um daraus resultierend Wissen zu generieren.

Durch die ständig steigernden Anforderungen unserer Gesellschaft und der wachsenden Erwartungshaltung von Kunden und Kundinnen von immer verfügbaren Systemen, wie Online-Shops, Suchmaschinen oder Software-Systemen werden verschiedene Technologien wie Industrie 4.0, Internet of Things, Big Data und Predictive Maintenance immer mehr zu Standardtechnologien, welche eingesetzt werden müssen.

“The result of ineffective maintenance management represents a loss of more than $60 billion each year.” (Mobley 2002) Durch die steigenden Anforderungen der Kunden und Kundinnen an ein hochverfügbares System, werden diverse Technologien der Wartungen immer mehr in den Vordergrund gerückt, um die Verfügbarkeit eines Systems in den Vordergrund zu stellen. Wie Mobley (2002) im oben angeführten Zitat zeigt, kann eine ineffektive Wartung einen immensen Verlust für ein Unternehmen darstellen, welches aber durch Wartungstechnologien, wie zum Beispiel Predictive Maintenance entgegengewirkt werden kann. Auch wenn diese Wartungskosten zwischen 15 und 60 Prozent, abhängig von der Branche, der Gesamtkosten des Produktes oder des Systems betragen können, zeigt Mobley (2002) auf, dass diese dem Verlust einer nicht durchgeführten Wartung gegenübergestellt werden muss. (Mobley 2002)

Aus diesem Grund haben ebenso Firmen in den letzten Jahren wiederholt von einer reaktiven Wartung, auf eine präventive Wartung umgestellt. Die reaktive Wartung ist die Wartung bei Auftreten eines Fehlers oder eines Ausfalls, wohingegen die präventive Wartung zu einem fix definierten Zeitpunkt in Intervallen durchgeführt wird. Eine verbesserte Methode der präventiven Wartung ist die vorausschauende Wartung, welche wie zuvor bereits erwähnt, auch als „Predictive Maintenance“ bezeichnet wird. Bei dieser Methode der Wartung wird ein System oder eine Maschine mit Sensoren oder Messinstrumenten ausgestattet, um Daten zu erfassen und diese in Echtzeit zu analysieren, deren Fehler und Zustände zu erkennen und an ein nachgelagertes System zu melden. (Huber und Kaiser 2017) Mit Hilfe der Methode der Predictive Maintenance kann ein ungeplanter Stillstand oder Ausfall eines Systems vermieden werden sowie die Qualität der Produkte und Systeme auf einem hohen Niveau gehalten werden. Des

Weiteren werden die Wartungskosten durch Predictive Maintenance um ein Vielfaches gesenkt. (Schlatt 2014)

Durch Predictive Maintenance und dem Erfassen der Daten von Maschinen und Systemen können somit Fehlverhalten und fehlerhafte Muster im Vorfeld bereits erkannt und vorhergesagt werden. Dies wiederum erhöht die Planbarkeit von Wartungen der Systeme und der Maschinen. (Huber und Kaiser 2017)

Data Mining wiederum wird auch als Schürfen oder vielmehr Graben innerhalb definierter Daten interpretiert. Das Ziel, nach welchem gegraben wird, ist das Generieren von Informationen beziehungsweise das Generieren von Wissen. Heute kann Wissen nämlich dem Gegenwert von Gold gegenübergestellt werden. (Alpar und Niedereichholz 2000) Mittels des generierten Wissens können Predictive Maintenance Methoden angewendet werden, um Wartungen in einem System oder an einer Maschine durchzuführen.

In der vorliegenden Arbeit werden aus diesem Grund die Einflussfaktoren und Optimierungen von Data Mining auf Predictive Maintenance für ein Warehouse Management System evaluiert. Diese Evaluierung erfolgt, aufgeteilt, über mehrere Schritte. Im ersten Schritt werden Hypothesen mittels einer Literaturrecherche aufgestellt. Im zweiten Schritt werden die zuvor definierten Hypothesen durch Expertinnen und Experten der Logistikbranche überprüft. Dadurch ergibt sich folgende Forschungsfrage:

„Welche Einflussfaktoren hat Data Mining auf Predictive Maintenance für ein Warehouse Management System in der Logistikbranche?“

1.1 Zielsetzung

Das Ziel der Arbeit ist zunächst die vorhandenen Data Mining Verfahren, Methoden der Wartung und deren Wechselwirkung im Einsatz bei Softwarewartung im Logistikbereich sowie die Anforderungen der Logistik an eine Wartung zu identifizieren.

Im zweiten Schritt soll eine Evaluierung der Einsatzmöglichkeiten von Predictive Maintenance im Logistikbereich erfolgen. Die beiden aufgezählten Schritte erfolgen durch eine empirische Literaturanalyse.

Im letzten Schritt werden die, mithilfe der Literatur aufgezeigten, Einflussfaktoren durch Experteninterviews erforscht.

Folglich können durch die Evaluierung der erkannten Potenziale, Einflussfaktoren und Einsatzmöglichkeiten mögliche weitere Forschungsprojekte profitieren.

1.2 Methode

Die Arbeit gliedert sich in zwei Teile. Zunächst werden im ersten Schritt Hypothesen erarbeitet Eine empirische Literaturanalyse der vier Themen „Data Mining“, „Predictive Maintenance“, „Wartung von Software“ und „Anforderungen der Logistik“ in einer darauffolgenden Kombination soll helfen, Hypothesen für diese aufgezählten Themen aufzustellen und zu formulieren:

- Alternativhypothese (H1): Durch den Einsatz von Data Mining in der Logistikbranche können Störungen in einem Warehouse Management System mittels Predictive Maintenance früher erkannt werden.
- Nullhypothese (H0): Durch den Einsatz von Data Mining in der Logistikbranche können Störungen in einem Warehouse Management System mittels Predictive Maintenance nicht früher erkannt werden

Im zweiten Schritt sollen durch eine empirische qualitative Befragung die zuvor aufgestellten Hypothesen evaluiert werden. Die Befragungen werden durch Experteninterviews mit Expertinnen und Experten aus der Logistik durchgeführt. Folglich wird eine Stichprobe für eine Gesamtpopulation definiert. Bevor die Interviews mit den Experten und Expertinnen durchgeführt werden, wird ein semistrukturierter Interviewleitfaden erstellt, welcher als Grundlage für die Interviews dient. Zur Auswertung der Interviews wird eine Inhaltsanalyse durchgeführt, wodurch die Forschungsfrage beantwortet wird.

1.3 Vorgehensweise

Die im Folgenden angeführten Punkte beschreiben das Vorgehen in der vorliegenden Arbeit:

- Zunächst werden Begriffe wie „Data Mining“, „Maintenance“ und „Warehouse Management System“ anhand wissenschaftlicher Literatur näher definiert sowie Grundlagenaufgestellt.
- Nachdem eine allgemeine Begriffsdefinition aufgestellt wurde, werden die verschiedenen Verfahren des Data Minings näher erläutert. Dabei wird der Unterschied zwischen „supervised-“ und „unsupervised-learning“ näher betrachtet beziehungsweise die Verfahren der Klassifikation, Regression, Assoziation, des Clusterings und deren Einsatzzwecke aufgezeigt.
- Nachdem eine grundlegende Beschreibung über die verschiedenen Wartungsmethoden erfolgt ist , wird im nächsten Schritt ein Einblick über die Wartungsmethode Predictive Maintenance gegeben. In diesem werden die verschiedenen Verwendungen von Predictive Maintenance sowie Techniken in Hardware und Software näher erläutert.
- Um die Logistik und deren Anwendungsgebiete der zuvor genannten Themen aufzuzeigen, wird die Datenerhebung der Logistikbranche sowie deren Einsatzgebiete der Predictive Maintenance und Data Mining näher beleuchtet.

- Die genannten Abschnitte bilden den Theorieteil der Arbeit. Die ausgearbeiteten Fakten werden schließlich in der Zusammenfassung wiedergegeben.
- Als Einleitung für den praktischen Teil der Arbeit werden die Merkmale der Logistikbranche und deren bisherigen Veränderungen sowie deren Anforderungen an Predictive Maintenance ausgearbeitet.
- Um Hypothesen zu bilden, werden die Einflussfaktoren des Data Minings auf Predictive Maintenance näherbeschrieben. Im Weiteren können daraus die Hypothesen abgeleitet und ein Interviewleitfaden für Interviews ausgearbeitet werden.
- Für diesen Interviewleitfaden bedarf es zunächst jedoch einer Darstellung der allgemeinen Vorgehensweise sowie des Aufbaus bei der Erstellung eines Interviewleitfadens . Im Weiteren kann dann, basierend auf den im Theorieteil erstellten Hypothesen, ein Interviewleitfaden erstellt werden.
- Der zuvor erstellte Interviewleitfaden wird für Interviews mit Expertinnen und Experten aus der Logistikbranche herangezogen. Die Auswahl sowie Begründung, warum diese Expertinnen und Experten für die Interviews herangezogen werden, wird in einem Abschnittnäher beschrieben. Danach werden die Interviews mit den Expertinnen und Experten durchgeführt sowie aufgezeichnet. Die aufgezeichneten Interviews werden auf Basis der zuvor definierten Hypothesen evaluiert und ausgewertet.
- Der letzte Schritt beinhaltet eine Gegenüberstellung der Ergebnisse aus der Theoriemit den Ergebnissen aus den Interviews mit den Expertinnen und Experten. In diesem Teil der Arbeit werden die Ergebnisse zusammengefasst sowie ein Fazit erstellt. Durch die beschriebene Vorgehensweise sollen die Einflussfaktoren von Data Mining auf Predictive Maintenance für ein Warehouse Management System in der Logistikbranche aufgezeigt werden

1.4 Aufbau der Arbeit

Die Arbeit wird in zwei Schritten aufgeteilt, wobei im ersten Schritt Hypothesen erarbeitet werden. Eine empirische Literaturanalyse der vier Themen „Data Mining“, „Predictive Maintenance“, „Wartung von Software“ und „Anforderungen der Logistik“ in einer darauffolgenden Kombination soll helfen, Hypothesen für diese aufgezählten Themen aufzustellen und zu formulieren.

Im zweiten Schritt sollen durch eine empirische Befragung die zuvor aufgestellten Hypothesen evaluiert werden. Die Befragungen werden durch Experteninterviews mit Experten und Expertinnen aus der Logistik aus verschiedenen Branchen durchgeführt. Somit wird eine Stichprobe für eine Gesamtpopulation definiert. Bevor die Interviews mit den Experten und Expertinnen durchgeführt werden, wird ein semistrukturierter Interviewleitfaden erstellt, welcher als Grundlage für die Interviews dient. Zur Auswertung der Interviews wird eine Inhaltsanalyse durchgeführt, wodurch die Forschungsfrage beantwortet wird.

Kapitel 1: Verfahren der Datenanalyse ▪ Zielsetzung ▪ Vorgehensweise ▪ Aufbau der Arbeit		**Kapitel 2: Grundlagen und Begriffsdefinition** ▪ Grundlagen des Data Minings ▪ Grundlagen Maintenance ▪ Grundlagen Warehouse Management System
Kapitel 3: Data Mining Verfahren ▪ Klassifikationsverfahren ▪ Regression ▪ Assoziationsverfahren ▪ Clusterverfahren	**Kapitel 4: Predictive Maintenance** ▪ Verwendung von Predictive Maintenance ▪ Predictive Maintenance in Software und Hardware	**Kapitel 5: Anwendungsgebiete** ▪ Datenerhebung in der Logistik ▪ Einsatzgebiete in der Logistikbranche
Kapitel 6: Zusammenfassung Theorieteil		

Kapitel 7 Einleitung Praxisteil ▪ Merkmale der Logistikbranche ▪ Herausforderungen der Logistikbranche ▪ Einflussfaktoren von Data Mining auf Predictive Maintenance ▪ Hypothesen	**Kapitel 8: Erstellung des Interviewleitfadens** ▪ Allgemeine Vorgehensweise und Aufbau bei der Erstellung eines Interviewleitfadens ▪ Interviewleitfaden für Interviews mit Experten und Expertinnen
Kapitel 9 Durchführung und Evaluierung der Interviews ▪ Auswahl von Experten und Expertinnen aus der Logistikbranche ▪ Allgemeine Auswertung von qualitativen Interviews ▪ Auswertung der Interviews mit den Logistikexperten ▪ Evaluierung der Interviews mit den Experten und Expertinnen aus der Logistikbranche	**Kapitel 10: Ergebnisse** ▪ Ergebnisse aus der Literatur ▪ Ergebnisse aus den Interviews ▪ Fazit und Zusammenfassung ▪ Ausblick

Abbildung 1 Aufbau der Arbeit

2 GRUNDLAGEN UND BEGRIFFSDEFINITION

Im folgenden Kapitel werden die Grundlagen und Begriffsdefinitionen, welche im Zuge dieser Arbeit von Bedeutung sind, erläutert. Diese sind notwendig, um aufgestellten Überlegungen nachvollziehen zu können.

Die folgenden Grundlagen und Begriffsdefinitionen werden in diesem Kapitel beschrieben und erläutert:

- Grundlagen des Data Minings
- Grundlagen Maintenance
- Grundlagen Warehouse Management System

2.1 Grundlagen des Data Minings

Data Mining kann als Schürfen oder Graben innerhalb von Daten interpretiert und folglich auch so übersetzt werden. Wobei das Ziel, nach welchem gegraben wird, Information beziehungsweise das Generieren von Wissen anhand dieser Information selbst ist. Heute kann nämlich Wissen mit dem Gegenwert von Gold, nach dem damals gegraben wurde, gleichgestellt werden. Mittels dieses erlangten Wissens wird es für Unternehmen möglichweitere Umsätze und Gewinne zu generieren. (Alpar und Niedereichholz 2000)

Die Methoden, welche im Data Mining verwendet werden, sind mit Themen der Statistik, künstlicher Intelligenz, maschinellem Lernen und der Mustererkennung eng verknüpft. Ziel hierbei ist es Muster in großen Datenmengen, in einem hohen automatisierten Grad in großen Datenmengen zu erkennen. Diese gefundenen Muster sollen dem Anwender als Wissen präsentiert werden. (Janetzko und Steinhöfel 1997) Die Daten, welche für den Prozess des Data Minings verwendet werden, können aus diversen Quellen kommen. Dies können Daten sein, welche von Menschen eingegeben worden sind. Aber auch Daten des Unternehmens aus SQL (Structured Query Language) oder anderen standardisierten Datenbankformaten sowie automatisch generierte Daten, wie von Maschinen oder Sensoren. (Bramer 2007) Die zuvor gesammelten Daten werden in diversen Schritten bearbeitet und in neuen Mustern gebündelt. Aus diesen Mustern können neue Hypothesen aufgestellt werden. Anhand dieser neuen Hypothesen können wiederum neue Daten und dadurch neue Verfahren aufgestellt werden, die auf Grundlage der in den Daten gefundenen Muster basieren(Bankhofer und Vogel 2008)

Aus diesem Grund wird im nächsten Kapitel auch der Prozess des Knowledge Discovery in Databases und Cross-Industry Standard Process for Data Mining beschrieben, da diese beiden Prozesse mittlerweile implizit mit dem Begriff Data Mining verwendet werden.

2.1.1 Knowledge Discovery in Databases Prozess

Der Prozess des KDD (Knowledge Discovery in Databases) ist eine automatische, explorative Analyse und Modellierung von großen Datenmengen. KDD ist der organisierte Prozess der Identifizierung gültiger, neuartiger, nützlicher und verständlicher Muster aus großen und komplexen Datensätzen. Data Mining ist der Kern des KDD-Prozesses, bei dem Algorithmen hergeleitet werden, die die Daten untersuchen, das Modell entwickeln und bisher unbekannte Muster entdecken. Das Modell wird für das Verständnis von Phänomenen aus den Daten, für die Analyse sowie für die Vorhersage verwendet. (Maimon und Rokach 2010)

Fayyad, Piatetsky-Shapiro und Smyth (1996) unterteilen den KDD Prozess in folgende Schrittebghc:

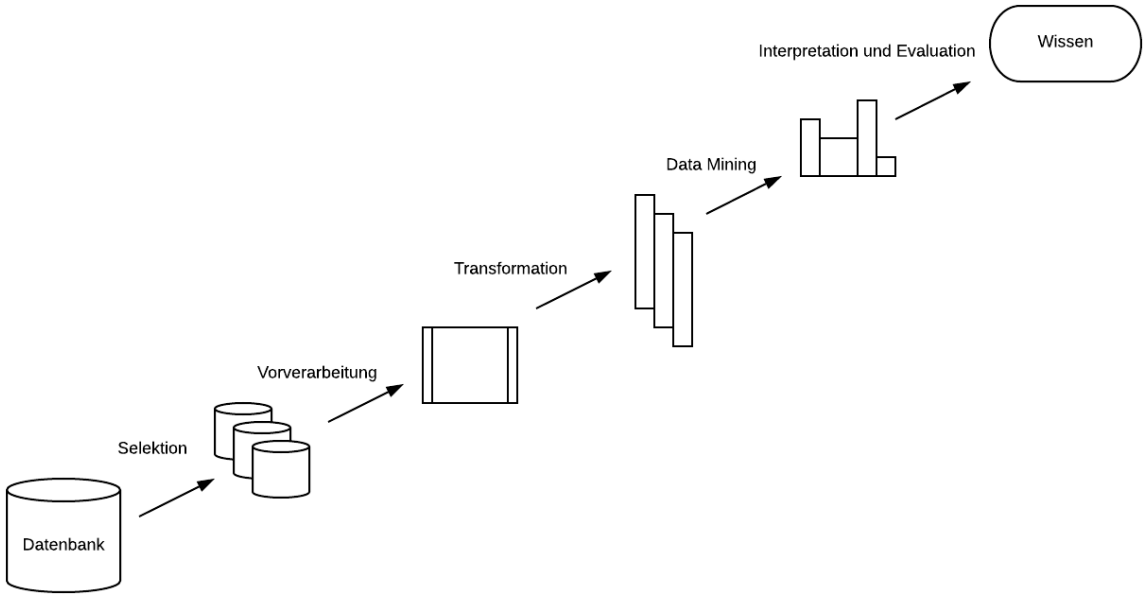

Abbildung 2 Prozess des Knowledge Discovery in Databases (Fayyad et al. 1996)

- **Selektion**

  Der Schritt der Selektion beinhaltet die Festlegung der relevanten Daten für die Fragestellung und bildet somit den Kontext der gesamten Ausarbeitung. In dieser Phase wird nach verfügbaren Daten gesucht sowie Überlegungen, um zusätzliche Daten zu sammeln und alle Daten für eine Wissensermittlung in einem Datensatz zu erheben. Die Selektion selbst ist ein wichtiger Teil im KDD Prozess, da der Mining Algorithmus Daten erkennt und daraus lernen kann. Wenn im Selektionsprozess aber bereits wichtigen Daten fehlen, dann kann die Fragestellung einen anderen Verlauf haben als initial erwünscht.

  Dies ist auch der Grund dafür, dass in diesem Prozessschritt so viele Daten und Attribute erfasst werden sollen wie nur möglich. Auf der anderen Seite sind Operationen, wie das Sammeln, Organisieren und Verarbeiten von riesigen Datenmengen allerdings sehr teuer sowie wird in der Praxis oftmals ein Kompromiss eingegangen. Dieser Kompromiss stellt aber den interaktiven und iterativen Aspekt des KDD dar. Es wird in den meisten Fällen

mit den verfügbarsten Datensätzen begonnen, welche zu einem späteren Zeitpunkt erweitert werden. Die Ergebnisse werden dann im Weiteren beobachtet und verglichen (Maimon und Rokach 2010).

- **Vorverarbeitung**

  Bei der Vorverarbeitung wird die Datensicherheit erhöht. Dies bedeutet, dass in diesem Schritt auch Daten gelöscht, fehlende Werte behandelt werden und das Entfernen von Lärm und Ausreißern in den Datenbeständen bereinigt wird. Durch diese Phase soll die Datenqualität erhöht werden. (Fayyad et al. 1996)

- **Transformation**

  Ziel der Transformation sollen besser vorbereitete Daten für das Data Mining selbst sein. Darunter zählt unter anderen die Methode der Dimensionsreduzierung und Transformation der Attribute. Der Schritt der Transformation ist oft entscheidend für den Erfolg des zugrunde liegenden Projekts. (Lin 2002)

- **Data Mining**

  Im vierten Schritt des KDD Prozesses erfolgt nach den vorangegangenen Schritten der eigentliche Schritt des Data Minings selbst, welcher einen zuvor ausgewählten Algorithmus auf die Daten anwendet. Der Schritt des Data Minings im KDD besteht aus drei Subschritten. (Maimon und Rokach 2010) Diese drei Subschritte beinhalten die Auswahl eines geeigneten Data Mining Verfahrens, die Auswahl eines geeigneten Data Mining Algorithmus und dem Anwenden des Algorithmus auf die Daten selbst.

  Bei der Auswahl des Verfahrens wird zwischen Cluster-, Klassifikations-, Assoziationsverfahren oder der Regression unterschieden. Weitere Details werden im Kapitel Data Mining Verfahren näher erläutert.

  Durch die Auswahl des Verfahrens für die zu lösende Fragestellung ergeben sich dann gewisse Algorithmen, welche auf den Datensatz angewendet werden können, um spezifische Muster in den Daten zu erkennen. Je nach Fragestellung gibt es wieder unterschiedliche Algorithmen, welche ausgewählt werden können. Die Auswahl des Algorithmus hängt oft sehr stark von der Fragestellung selbst ab. Muss ein Algorithmus mehrmals in der Stunde angewendet werden, um eine Frage zu beantworten oder reicht es nur einmal pro Woche?

  Der letzte Subschritt ist das Anwenden des zuvor ausgewählten Algorithmus über mehrere Runden, bis ein zufriedenstellendes Resultat gefunden wurde.

- **Interpretation und Evaluation**

  Der letzte Schritt des KDD-Prozesses beinhaltet die Evaluation und Interpretation der gefundenen Ergebnisse. Die gewonnenen Erkenntnisse werden in Bezug auf, die im ersten Schritt, definierte Fragestellung bewertet. Der Abschnitt der Evaluierung konzentriert sich auf das Verständnis und die Nutzbarkeit des eingeführten Modells. Die in diesem Schritt gefundenen Erkenntnisse werden für die weitere Verwendung dokumentiert, um mögliche Ableitungen vornehmen zu können. Das dadurch gewonnene

Wissen wird aktiv, indem Änderungen am System vorgenommen und die Auswirkungen gemessen werden. Dieser Schritt des KDD-Prozesses bestimmt somit tatsächlich die Wirksamkeit des gesamten KDD-Prozesses. (Maimon und Rokach 2010)

2.1.2 Cross-Industry Standard Process for Data Mining

Der CRSIP-DM (Cross-Industry Standard Process for Data Mining) ist ein Prozess, welcher von mehreren Unternehmen entwickelt wurde. (Chapman et al. 1999) Wie in Abbildung 3 zu erkennen ist, beinhaltet dieser Prozess im Gegensatz zum KDD insgesamt sechs Phasen. Hier ist jedoch zu beachten, dass die Abfolge der einzelnen Phasen nicht starr, sondern dynamisch ist. Der CRSIP-DM ist ein vor- und zurückspringen zwischen einzelnen Phasen, was möglich und teilweise auch nötig ist. (Chapman et al. 1999)

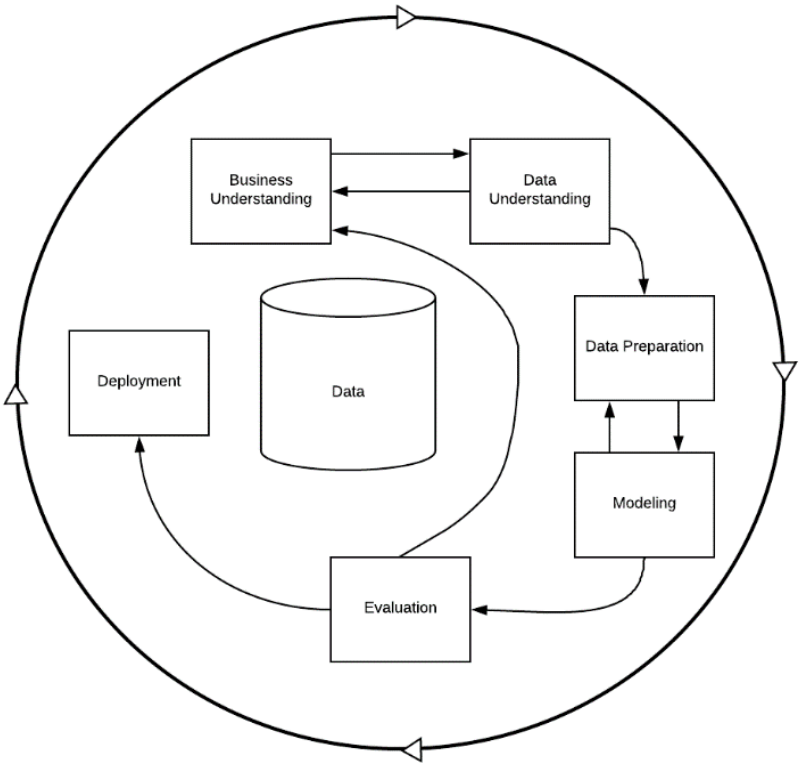

Abbildung 3 CRISP-DM (Chapman et al. 1999, S. 12)

Je nach Ergebnis der Phase wird bestimmt, welche Phase die Nächste sein wird. Die Pfeile in der Abbildung beschreiben die häufigsten und wichtigsten Abhängigkeiten im Prozess selbst. Der äußere Kreis soll die Zirkularität des Data Minings symbolisieren, denn der Prozess des Data Minings endet nicht, sobald eine Lösung verfügbar ist. (Chapman et al. 1999)

Der CRISP-DM ist in folgende Phase eingeteilt:

- **Geschäftsmodell verstehen**

  Diese erste Phase konzentriert sich auf das Verständnis der Projektziele und Anforderungen aus der Geschäftsperspektive heraus. Dabei wird dieses Wissen in eine Data-Mining-Problemdefinition und einen vorläufigen Plan umgewandelt, um die Ziele zu erreichen. (Chapman et al. 1999)

- **Daten verstehen**

  Die Phase „Daten verstehen“ beginnt mit der initialen Erfassung der Daten. Sie umfasst somit die Aktivitäten, mit welchen ein Vertrauen zu den Daten hergestellt wird. In dieser Phase werden erste Erkenntnisse über die Datenqualität und die Zusammenhänge aus Daten über eine visuelle und statistische Aufbereitung aufgezeigt. (Chapman et al. 1999)

- **Daten aufbereiten**

  In der dritten Phase werden die Daten aufbereitet und somit alle Aktivitäten abgedeckt, welche zur Erstellung des endgültigen Datensatzes erforderlich sind. In diesem Schritt werden Daten von Ausreißern bereinigt sowie fehlende Werte und fehlerhafte Daten behandelt. In weiteren Schritten kann eine Konvertierung der Daten stattfinden, indem irrelevante Daten oder sogar stark voneinander anhängige Daten herausgefiltert werden. (Chapman et al. 1999)

- **Modellierung**

  In der Phase der Modellierung erfolgt eine Auswahl des Data Mining Verfahrens sowie dem dazugehörigen passenden Data Mining Algorithmus. Da aber für eine Fragestellung oft mehrere Algorithmen und Verfahren infrage kommen können, treten in dieser Phase oft neue Anforderungen an die Fragestellung selbst auf. Aus diesem Grund gibt es in dieser Phase auch oft Rückkoppelungen in eine vorhergehende Phase. Das schlussendliche Ergebnis dieser Phase ist ein Modell zur Datenanalyse. (Sharafi 2013)

- **Evaluation**

  Bei der Evaluation wird beim zuvor entwickelten Modell vor der Ausführung überprüft, ob die Qualität für das Problem und somit für die Fragestellung geeignet ist und dadurch die Fragestellung ordnungsgemäß beantwortet werden kann. Ein besonderes Augenmerk wird daraufgelegt, dass die Erwartungen an die Fragestellung hinreichend erfüllt wird. Als Ergebnis dieser Phase soll eine Entscheidung über die Verwendung der Data Mining Ergebnisse getroffen werden. (Sharafi 2013)

- **Einsatz**

  In der letzten Phase des CRISP-DM werden die Ergebnisse und somit das erhobene Wissen als Resultat in die Organisation zurückgeführt. Dadurch werden die dargelegten Ergebnisse und Erkenntnisse präsentiert damit diese auch genützt werden können. Dies kann, aufgrund der Fragestellung, entweder nur in Form eines Berichts erfolgen, oder in Form eines neuen Entstehung Prozesses in der Organisation (Sharafi 2013)

2.2 Grundlagen Maintenance

Die Wartungskosten machen einen großen Teil der Gesamtbetriebskosten der gesamten Fertigungs- oder Produktionsanlagen aus. Abhängig von der jeweiligen Branche können die Instandhaltungskosten zwischen 15 und 60 Prozent der Produktionskosten betragen. In der Lebensmittelindustrie beispielsweise machen die durchschnittlichen Wartungskosten etwa 15 Prozent der Kosten der produzierten Güter aus, während die Wartungskosten für Eisen und Stahl, Zellstoff und Papier sowie andere Schwerindustrien bis zu 60 Prozent der gesamten Produktionskosten ausmachen. In den meisten Werken enthalten die angegebenen Wartungskosten viele nicht wartungsbezogene Ausgaben. Beispielsweise beinhalten viele Werke Änderungen an bestehenden Kapitalsystemen, die durch marktbezogene Faktoren, wie neue Produkte angetrieben werden. Diese Ausgaben sind nicht Instandhaltungskosten und sollten auf Nicht-Wartungs-Kostenstellen umgelegt werden. Die tatsächlichen Instandhaltungskosten sind jedoch erheblich und stellen eine kurzfristige Verbesserung dar, die sich direkt auf die Rentabilität der Anlagen auswirken kann. Erhebungen über die Wirksamkeit des Instandhaltungsmanagements zeigen, dass ein Drittel von 33 Cent pro Dollar aller Instandhaltungskosten in den USA als Ergebnis unnötiger oder unsachgemäß durchgeführter Instandhaltung verschwendet wird. Wenn man bedenkt, dass die US-Industrie jedes Jahr mehr als 200 Milliarden Dollar für die Instandhaltung von Anlagen und Einrichtungen ausgibt, wird die Auswirkung auf Produktivität und Gewinn, die der Instandhaltungsbetrieb darstellt, deutlich. (Mobley 2002)

In den USA wird somit jährlich ein Verlust von 60 Milliarden Dollar durch ein ineffektives Instandhaltungsmanagement verursacht. Durch dieses Fehlverhalten der Instandhaltung verlieren Produkte und Services erheblich an Wettbewerbsfähigkeit am Weltmarkt. Um diese Wettbewerbsfähigkeit aufrecht zu halten, gab es erste Überlegungen der Instandhaltungsmanagementphilosophien. (Mobley 2002)

Der Hauptgrund für dieses ineffektive Management ist der Mangel an Sachdaten zur Quantifizierung des tatsächlichen Reparatur- oder Wartungsbedarfs von Maschinen, Anlagen und Systemen. Die Wartungsplanung basierte, und basiert in vielen Fällen immer noch, auf statistischen Trenddaten oder auf dem tatsächlichen Ausfall von Anlagen und Systemen. Bis vor Kurzem hat die mittlere und Unternehmensebene die Auswirkungen der Wartungsarbeiten auf die Produktqualität, die Produktionskosten und, was noch wichtiger ist, auf den Endgewinn ignoriert. Die allgemeine Meinung lautete: "Instandhaltung ist ein notwendiges Übel" oder "Es kann nichts getan werden, um die Instandhaltungskosten zu verbessern". (Mobley 2002).

Diese Aussagen waren vielleicht noch vor 20 oder 25 Jahren zutreffend, aber durch die Entwicklung von Mikroprozessor- und computergestützten Instrumenten, die der Überwachung des Betriebszustandes dienen, ist diese Meinung mittlerweile nicht mehr üblich. Diese Instrumente haben das Mittel zur Verfügung gestellt, um Reparaturen, katastrophale Ausfälle und negative Auswirkungen auf den Betrieb und die daraus resultierenden Wartungsarbeiten zu verringern. (Mobley 2002)

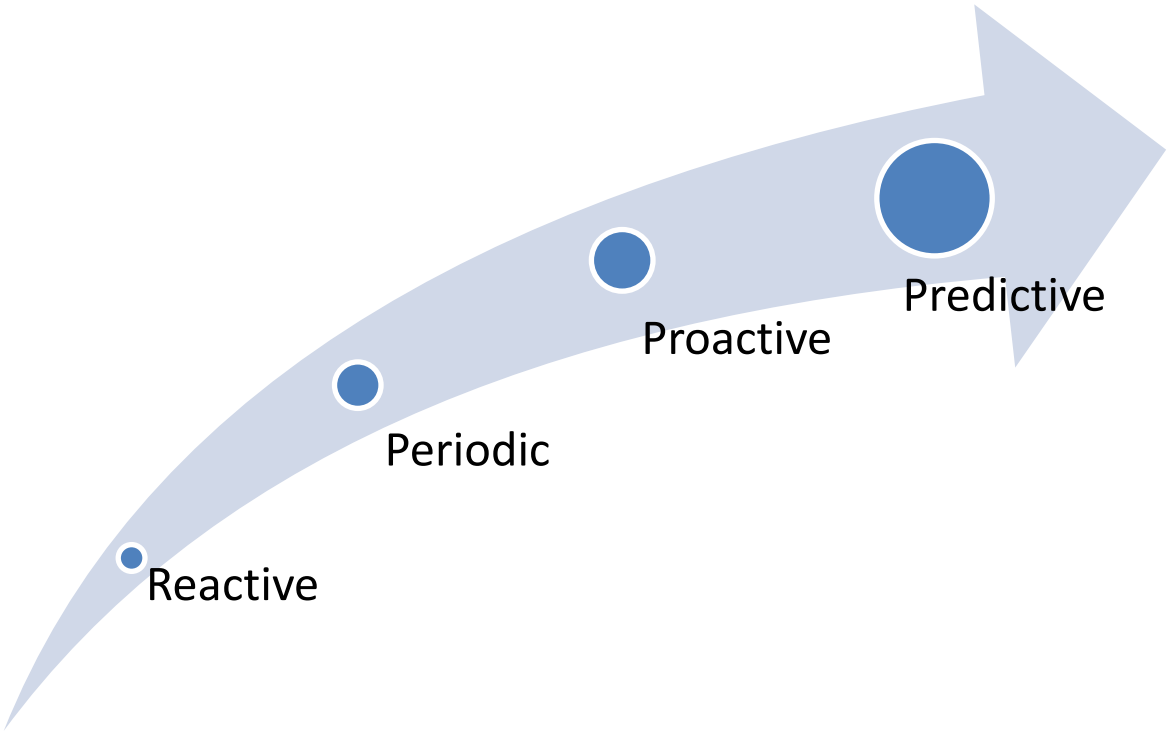

Abbildung 4 Stufen der Wartung (Selcuk 2017)

Wie in Abbildung 4 ersichtlich, gibt es unterschiedliche Stufen und Typen einer Wartung. Um ein Predictive Maintenance System besser zu verstehen, werden nachfolgend die traditionellen Wartungsphilosophien näher erläutert: (Girdhar 2004)

- **Run-to-Failure**

  Der Grundgedanke dieser Wartungsmethode, welche auch als „breakdown" Methode bezeichnet wird, ist es, so lange zu warten, bis es zu einem Problem kommt. Erst wenn eine Komponente einen Fehler aufweist, welcher das System beeinflusst, wird dieser Teil getauscht beziehungsweise gewartet. Die „Run-to-Failure" Methode wird oft dort eingesetzt, wo ein Ausfall keine Kosten verursachen, oder einen Produktionsstillstand hervorrufen würde. Einer der größten Nachteile bei dieser Methode ist, dass bei einem Auftreten eines Fehlers die Wartungsaktivitäten unmittelbar durchgeführt werden soll und ein Ersatzteillager zur Verfügung stehen muss, um die defekten Komponenten auszutauschen. Dadurch, dass eine hohe Anzahl an Ersatzteilen auf Lager gehalten wird, wird versucht, die Einkaufskosten zu reduzieren. Diese Methode ist aber im Grunde die am ineffizienteste der hier Aufgezählten und auch nicht sinnvoll.

- **Preventive Maintenance**

  Preventive Maintenance wird auch als „time-based" Wartungsmethode bezeichnet. Bei dieser Methode werden die Wartungsarbeiten anhand von Zeitintervallen oder Arbeitsstunden an einer Maschine oder einem System durchgeführt. Bei Systemen, welche keine durchgehende und kontinuierliche Benutzung aufweisen, ist diese Methode die sinnvollste. Des Weiteren muss den Bedienern und Bedienerinnen des Systems genügend Wissen und Zeit für die Wartungsarbeiten zur Verfügung stehen, denn es soll

an Tagen, wo das System nicht benutzt wird, eine Wartung durchgeführt werden. Der größte negative Punkt dieser Vorgehensweise bei der Wartung ist, dass die Wartungsarbeiten gegebenenfalls zu früh oder auch zu spät durchgeführt werden. Bei einer zu frühen Wartung könnten Komponenten gewartet oder getauscht werden, welche eine höhere Lebensdauer aufweisen. Bei einer zu späten Wartung könnte es wiederum zu einem unerwarteten Stillstand des Systems kommen.

- **Proactive Maintenance**

  Bei der Methode der Proactive Maintenance oder auch Prevention Maintenance genannt, versucht man, jeden Fehler in einem System auf den Ursachenursprung zu analysieren. Dadurch, dass proaktive Tätigkeiten durchgeführt werden, soll die Ursache des Fehlers sowie ein erneutes Auftreten des Fehlers verhindert werden. Die Proactive Wartungsmethode erweitert den predictive- beziehungsweise condition-based-Ansatz um die Analyse und Beseitigung der Fehler, damit diese nicht erneut auftreten. Diese Methode erfordert jedoch viel Wissen und Kenntnis über das System und dessen Infrastruktur, wodurch es für den Endbenutzer und die Endbenutzerin des Systems nahezu unmöglich ist, diese Methode selbstständig durchzuführen. Aus diesem Grund wird diese Methode oft dem Hersteller und der Herstellerinnen des Systems überlassen. Dem hinzukommend steigen die Wartungskosten bei dieser Methode, da neben dem Wissen auch oft eine spezielle Ausrüstung notwendig ist.

- **Predictive Maintenance**

  Bei der Predictive Maintenance ist eine Methode, bei welcher analytische Werkzeuge eingesetzt werden, um auffällige Muster zu identifizieren sowie um den Zustand eines Systems prognostizieren zu können und einen zukünftigen Fehlerfall oder Ausfall zu verhindern. Dadurch können spezifische Wartungsaktivitäten für das System erkannt, geplant und auch frühzeitig durchgeführt werden. Aus diesem Grund wird diese Methode auch „condition-based“ Wartungsmethode genannt. Durch die permanente Überwachung des Systems und das Erkennen von Mustern können auch Wartungen am System gut geplant und etwaige Ersatzteile frühzeitig auf Lager gekauft werden. Durch diese geplanten Wartungsarbeiten, an einem System, erhöhen sich die Effizienz und auch die Performanceleistung. Der einzige Nachteil an der Methode der predictive Maintenance ist, dass Vorhersagen nicht immer zu 100 Prozent korrekt sein können, womit die Wartungsarbeiten am System wieder zu früh oder zu spät sein können.

2.3 Grundlagen Warehouse Management System

Ein WMS (Warehouse Managementsystem) ist ein Informationssystem zur Steuerung, Kontrolle und Optimierung komplexer Lager- und Distributionssysteme. (Warehouse Logistics 2020) Ein Warehouse Management System besteht aus verschiedenen Teilmodulen, welche im nachfolgenden Abschnitt näher beschrieben sind: (Hompel und Schmidt 2010)

- **Lagerverwaltung**

  Das Modul der Lagerverwaltung ist eines der ureigensten Funktionen eines Warehouse Managementsystem. Einer der Hauptaufgabe ist das Buchführen des vorhandenen Lagerungspotenzials, welches die vorhandenen Lagerplätze des Systems darstellt. Des Weiteren ist ein weiterer Bestandteil der Lagerverwaltung die Verwaltung der gelagerten Einheiten, welches auch als Bestands- und Mengenverwaltung bezeichnet wird. In einem manuellen und somit ohne WMS geführten Lager, arbeiten die Bediener und Bedienerinnen nach ihrem Erfahrungswissen. Anhand ihrer selbstständigen Entscheidungskraft sind sie in der Lage, ein geeignetes Förder- und Lagermittel auszuwählen beziehungsweise auch einzusetzen. Der Vorteil einer systemgeführten Lagerverwaltung besteht darin, dass diese Entscheidungen den Benutzern und den Benutzerinnen abgenommen werden.

  Eine weitere Hauptaufgabe der Lagerverwaltung ist die, wie bereits im Absatz zuvor erwähnte, Bestands- und Mengenverwaltung. Das ist die Verwaltung der Waren nach unterschiedlichen Kriterien, wie den minimalen und maximalen Bestand, um eine Versorgungssicherheit der Waren an gewissen Lagerplätzen zu gewährleisten und Übermengen zu vermeiden. Zu diesen Aufgaben gehört ebenfalls die Überwachung der zulässigen Lagerdauer, damit Artikel im Lager nicht verderben und das Mindesthaltbarkeitsdatum eingehalten wird.

- **Reorganisation**

  Bei der Reorganisation geht es im Wesentlichen darum, dass es in regelmäßigen Abständen erforderlich ist, ein Lagersystem auf dessen Effizienz zu überprüfen und Schritte für eine Optimierung durchzuführen, welche auch Lagerreorganisation genannt wird. Auslöser für eine Reorganisation kann die Änderung des Zugriffverhaltens auf diverse Artikel, wie zum Beispiel das Absinken des Durchsatzes, anlaufende und ablaufende Vertriebsaktionen oder auch die Veränderung des Sortiments sein. Die genannten Gründe können dazu führen, dass gewisse Lagerfächer logisch falsch bestückt sind und somit Transportwege und Raumnutzungsgrade steigen. Bei der Reorganisation werden Artikel wieder auf deren ordnungsgemäße Zugriffsklasse und Lagerzone zugewiesen und dorthin transportiert.

- **Fördermittelverwaltung und Leitsystem**

  Eine weitere Hauptaufgabe von Warehouse Managementsystemen ist die Verwaltung von Fördermitteln und die Aufgabe als Leitsystem. Dies beinhaltet die Steuerung von

automatischen Fördermitteln sowie das Verwalten von manuellen Fördermitteln, wie Stapler und Flurförderfahrzeugen. Darunter zählt das Dispatching und Scheduling von Fördermittel. Beim Dispatching wird für einen aktuell anstehenden Auftrag das geeignete Fördermittel zugewiesen. Eine Zuweisung erfolgt anhand diverser Kriterien und Strategien, wie zum Beispiel das nächste freie Fördermittel oder das räumlich am nächsten befindliche Fördermittel. Das Ziel des Scheduling wiederum ist die Verteilung mehrere Aufträge sowie deren Zuweisung eines idealen Fahrplanes und einer optimalen Auftragsreihenfolge. Eine wesentliche Voraussetzung für die Durchführung von Scheduling ist jedoch das Sammeln von Aufträgen, um in einen Auftragspool diese auch zu verteilen. In der Praxis wird oft eine Mischform dieser beiden Vorgehensweisen eingesetzt.

- **Datenerfassung, -aufbereitung und -visualisierung**

  Ein weiteres Modul ist die Datenerfassung sowie deren Aufbereitung und Visualisierung. Darunter zählt die Leistungserfassung wie Fehlmengendokumentation, Erfassen der Kommissionierfehler, Laufzeiten und Ausfallzeiten. Des Weiteren ist ein wesentlicher Aspekt die Erfassung von personenbezogenen Daten, um eine leistungsbezogene Lohnabrechnung durchführen zu können.

- **Inventur**

  Die Inventur ist eine vom Gesetzesgeber vorgeschriebene Tätigkeit in einem Lagersystem. Dies wird ebenfalls mittels eines WMS abgebildet, indem innerhalb eines Geschäftsjahres eine Zählung inklusive des Lagerorts, Position, Bezeichnung des Gegenstandes und aufgenommener Menge im System festgehalten wird.

2.3.1 Basisdaten und Logistische Kennzahlen

Die Ausgestaltung und Planung von Distributions- und Lagersystemen erfolgt vielschichtig, somit ist eine vollständige Sammlung aller relevanten und möglichen Größen nicht möglich. Ein großer Teil dieser Kenngrößen sind mit einer Aufgabenstellung zu definieren und zu erstellen. Im Folgenden sollen jedoch die elementaren und allgemeingültigen Kenngrößen, welche in einer Vielzahl an Systemen existieren, aufgelistet werden. Basisdaten werden auch als Absolutzahlen bezeichnet und werden durch deren unmittelbare Messung oder Zählung sowie einer Summation oder Differenz von bestimmten Einheiten übernommen oder erfasst. Diese stellen auch die durch ein System zu leistenden Anforderungen und Basisinformationen dar. (Hompel und Schmidt 2010)

Kennzahlen wiederum sind aussagekräftige und verdichtete Informationen, welche zum Vergleich und zur Bewertung der Effizienz von Prozessen und Systemen genommen werden. Hierbei werden Relativzahlen als auch Absolutzahlen eingesetzt, welche in Verhältnis gesetzte Zahlen beziehungsweise Daten zur Anwendung darstellen. (Hompel und Schmidt 2010)

Zu den Basisdaten gehören Stammdaten. Dies sind statische Daten, welche sich über einen längeren Zeitraum nicht verändern. Dies kann unter anderem der Artikelstamm oder auch eine

Auflistung aller eingesetzten Fördermittel im Unternehmen sein. Bestandsdaten wiederum, welche ebenfalls zu den Basisdaten zählen, sind Daten, die nur zu einem Zeitpunkt vorhanden sind. Dies kann zum Beispiel eine Auflistung an aktuellem Bestand in einem WMS sein. (Hompel und Schmidt 2010)

Bei den Kennzahlen in Logistischen Systemen gibt es eine Fülle verschiedener Informationen, welche die Form von Daten annehmen und eine Bewertung oder Optimierung auf Basis dieser Daten komplizieren. Des Weiteren können die Daten irreführend sein, wenn kein entsprechender Kontext hergestellt werden kann. So hat zum Beispiel eine Kennzahl wie die Auftragsanzahl pro Tag recht wenig Aussagekraft, wenn nicht auch die Kennzahl der Anzahl an Positionen in einem Auftrag berücksichtigt wird. (Hompel und Schmidt 2010)

Die nachfolgende Tabelle soll eine Übersicht der häufigsten Kennzahlen in einem Logistischen System darstellen: (Hompel und Schmidt 2010)

Kennzahl	**Definition**	**Zielsetzung/ Fragestellung**	**Einflussgrößen**
Lagerfüllungsgrad	$\frac{\text{Anzahl belegter Fächer (Stk}}{Lagerkapazität\ (Stk)}$	Nutzung der Ressource Lagerkapazität	Bestandsmanagement
Umschlagsgrad mengenbezogen	$\frac{Auslagerungen\ (Stk)}{Lagerkapazität\ (Stk)}$	Dynamik des Lagers	Große Versandeinheiten (durch Kundenaufträge)
Umschlagsgrad wertbezogen	$\frac{Gesamtumsatz\ (€)}{Ø\ Lagerwert\ (€)}$	Bestandskosten	Bestellwesen
Kosten/Lagerplatz	$\frac{Gesamtumsatz\ (€)}{Lagerkapazität\ (Stk)}$	Wahl der Lagertechnik	Lagertechnik
Kommissionierweg/ Position	$\frac{mittlerer\ Kommissionierv}{Ø\ Positionen\ (Stk)}$	Wahl des Kommissionierverfahrens	Ablaufsteuerung, Kommissionierprinzip, Informationsmanagement
Pickdichte	$\frac{Positionenzahl\ (Stk)}{Zugriffsfläche\ (m^2)}$	Mittlere Wegzeiten, Kommissionierleistung	

Tabelle 1 Kennzahlen der Logistik

2.4 Zusammenfassung

In diesem Kapitel wurden die Grundlagen des Data Minings, die Grundlagen zu Maintenance und die Grundlagen über Warehouse Management Systeme näher dargestellt. Im ersten Schritt der Grundlagen des Data Minings wurde Data Mining allgemein beschrieben und die beiden Prozesse KDD und CRISP-DM, welche aus den nachfolgenden Prozessschritten bestehen, erläutert:

KDD

- Selektion
- Vorverarbeitung
- Transformation
- Data Mining
- Interpretation und Evaluation

CRISP-DM

- Geschäftsmodell verstehen
- Daten verstehen
- Daten aufbereiten
- Modellierung
- Evaluation
- Einsatz

Im zweiten Schritt wurden die Grundlagen der Maintenance, welche, die Methoden der Maintenance unter anderem Run-to-Failure, Preventive Maintenance, Proactive Maintenance und Predictive Maintenance bestehen, laut Theorie beschrieben.

Im dritten Schritt wurden die Grundlagen von Warehouse Management Systemen näher betrachtet. Die Aufgaben eines Warehouse Management Systems wurden näher erläutert, welche aus folgenden Themenpunkten besteht:

- Lagerverwaltung
- Reorganisation
- Fördermittelverwaltung und Leitsystem
- Datenerfassung, -aufbereitung und -visualisierung

Als letzter Punkt im dritten Schritt wurden noch die Basisdaten und logistischen Kennzahlen näher dargestellt.

3 DATA MINING VERFAHREN

Wie bereits in den vorangegangenen Kapiteln beschrieben, enthält der KDD Prozess und auch der CRSIP-DM eine Auswahl an Data Mining Verfahren sowie eine Auswahl und Anwendung des geeigneten Algorithmus.

In der Literatur gibt es verschiedene Verfahren von Data Mining selbst. Maimon und Rokach (2010) unterscheiden hier zwischen „supervised“ und „unsupervised learning“. Die Unterscheidung zwischen „supervised“ und „unsupervised“ sind hier jeweils die Eingangsgrößen, mit welchen die Verfahren arbeiten. Beim „supervised learning“ wird die Eingangsgröße zu einem Zielwert identifiziert. Diese Kategorie an Verfahren arbeitet mit bereits bekannten Daten, welche in sogenannten Trainingsrunden zuvor angelernt worden sind. Das „unsupervised learning“ hat hier keine Eingangsgrößen und somit ist der Zielwert auch unbekannt. Es geht ausschließlich um das Auffinden von nicht bekanntem Wissen in den Daten. Das Ziel ist das Auffinden von nützlichen Strukturen und Mustern in den erhobenen Datensätzen.

Eine zweite Kategorisierung laut Literatur ist die Art der zu findenden Muster. Einerseits gibt es Verfahren, welche sich auf die Vorhersage von Daten konzentrieren und auf der anderen Seite gilt es, sich lediglich auf das Finden von nicht bekannten Mustern in den Datensätzen zu fokussieren.

Folgende Verfahren des Data Minings werden in den nachfolgenden Kapiteln detaillierter erläutert:

- **supervised learning**
  - Klassifikationsverfahren
  - Regressionsverfahren
- **unsupervised learning**
  - Assoziationsverfahren
  - Clusterverfahren

3.1 Klassifikationsverfahren

Die Klassifikation zählt zur Klasse des „supervised learning“. Sie wird auch laut Spehling (2007) als prädikative Data Mining Methode bezeichnet. Die Ausgangslage für dieses Data Mining Verfahren sind bereits klassifizierte Objekte in den zu analysierenden Daten. Es wird überprüft, ob ein Modell gefunden werden kann, welches eine allgemeingültige Klassifizierungsvorschrift darstellt. Als Eingangsobjekte werden Daten herangezogen, welche auch als Trainingssatz bezeichnet werden, welche eine Vielzahl von Datensätzen beinhalten, von denen jeder dieser Datensätze mehrere Attribute oder Merkmale enthalten. Jeder dieser Datensätze im Trainingssatz ist mit einer Klassenbezeichnung versehen, welche entweder kategorisch oder

qualitativ ist. Das Klassifikationsverfahren wird oft bei Fragestellungen wie „Welche Kreditnehmer und Kreditnehmerinnen werden einen Kredit zurückzahlen?“, verwendet. (Müller und Lenz 2013)

In den nachfolgenden Kapiteln werden die folgenden zwei Klassifikationsverfahren näher erläutert:

- Entscheidungsbäume
- Nearest Neighbour Methode

3.1.1 Entscheidungsbäume

Da im Data Mining oft Daten aus mehreren Millionen Datensätzen vorkommen, gibt es oft simple Sortierungsmechanismen, welche auch diese großen Datenmengen gut handhaben können. (Mehta et al. 1996) Einer dieser Methoden sind Entscheidungsbäume. Bei Entscheidungsbäumen wird eine rekursive Partitionierung der Daten vorgenommen, bis jede Partition mindestens einen Datensatz einer Klasse besitzt. (Zhong und Zhou 1999)

In diesem Algorithmus wird ein gerichteter Baum aufgestellt, welcher mindestens einen, aber auch mehrere Knoten besitzt. Der Knoten selbst steht als Test auf den Wert eines Attributes eines Datensatzes. Am Ende eines Knotens kann ein Blatt oder ein weiterer Knoten hängen. Ist der zuvor durchgeführte Test auf das Attribut eines Datensatzes erfolgt, wird der Datensatz dem Blatt oder dem Knoten zugeordnet. Dies wird in Abbildung 5 bildlich dargestellt.

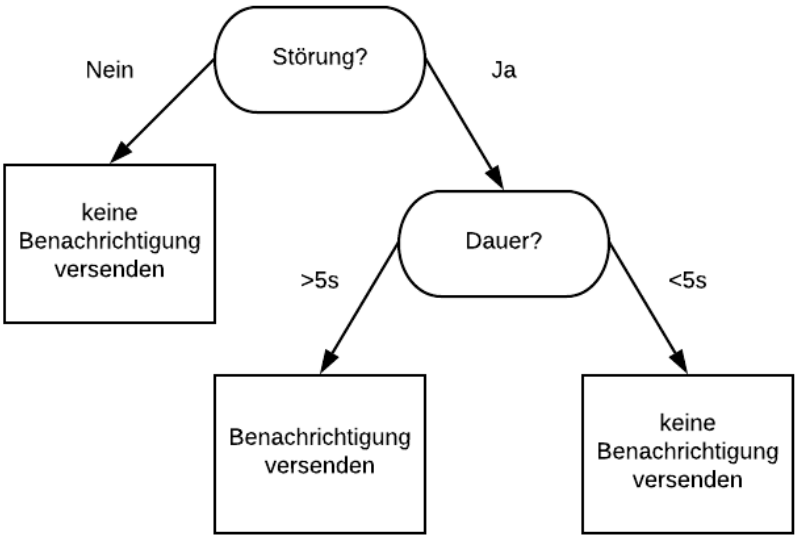

Abbildung 5 Beispiel eines Entscheidungsbaums

Abbildung 5 stellt einen Entscheidungsbaum dar, der die Entscheidung über das Versenden einer Benachrichtigung bei einer Störung prüft. Jede Instanz folg anhand eines Attributwertes eines Datensatzes einem Pfad bis zum Blatt. Das Blatt, bei welchem diese Instanz landet, ist seine jeweilige Klassenzuordnung. Im Beispiel von Abbildung 5 und der Frage, ob eine Störung besteht und dies bejaht wird, folgt darauf die Frage wie lange diese Störung bereits besteht. Wenn diese kürzer als fünf Sekunden besteht, oder auch wenn keine Störung besteht, wird die Dateninstanz jeweils der Klasse „keine Benachrichtigung versenden“ zugeordnet. Sollte die Dauer der Störung

länger als fünf Sekunden sein, wir diese der Klassenzuordnung „Benachrichtigung versenden“ zugeordnet.

3.1.2 Nearest Neighbour Methoden

Die Nearest Neighbour Methode betrachtet für das zu klassifizierende Objekt seine am nächstgelegenen Nachbarn mit einer bekannten Klassenzugehörigkeit. Das zu klassifizierende Objekt bekommt dadurch die Klassenzuordnung seiner Nachbarn. (Müller und Lenz 2013)

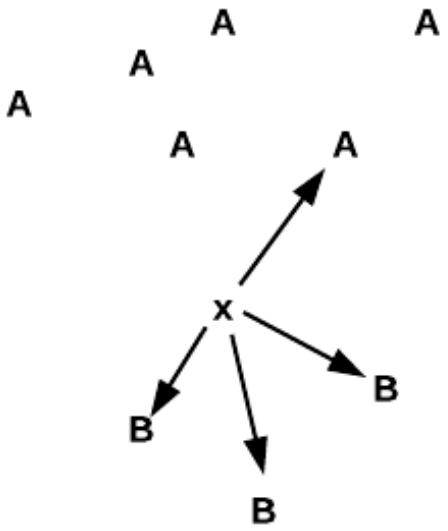

Abbildung 6 Beispiel Nearest Neighbour (Müller und Lenz 2013, S. 106)

Wie in Abbildung 6 zu erkennen, soll für das Objekt X eine Klassifizierung erfolgen. Je geringer der Abstand zu den umliegenden Objekten mit einer bereits zugeordneten Klasse ist, umso höher ist die Wahrscheinlichkeit, dass das Objekt X ebenfalls diese Klassenzuordnung erhalten wird. Auch wenn wie in der Abbildung der Abstand zum Objekt A am geringsten wäre, würde sich der Algorithmus für die Klassenzuordnung B entscheiden, da es in der Umgebung vom Objekt X bereits mehrere Objekte mit der Klassenzuordnung B gibt. Somit wird die Häufigkeit der Klassenzuordnungen ebenfalls im Algorithmus berücksichtig.

3.2 Regression

Die Regression ist ebenfalls ein Data Mining Verfahren, welches dem „supervised learning“ zugeschrieben wird. Es stellt ein Vorhersagemodell für kontinuierliche, numerische und geordnete Werte dar. Durch die Regression wird ein Zusammenhang zwischen einer Zielgröße Y und einer oder mehrere Ausgangs-Variablen $X^{(i)}$ untersucht. (Kronthaler 2014)

Durch die Data Mining Methode der Regression werden in der Praxis oft Fragestellungen der Absatzmenge eines Artikels und dergleichen beantwortet. Ein klassisches Beispiel hierfür wäre,

wie sich die Absatzmenge eines Produktes ändert, wenn sich der dazugehörige Preis dieses Artikels verändert.

Die Regression wird in folgende Arten eingeteilt, welche in weitere Folge erläutert werden:

- **Lineare Regression**

  Die Lineare Regression stellt, wie bereits in diesem Kapitel beschrieben, eine abhängige Variable, Y, und eine unabhängige Variable, X, dar. Diese Darstellung der Variablen ist aber eine rein theoretische Annahme. Es wird angenommen, dass die Variable X einen Einfluss auf Y hat. Es gilt zu beachten, dass diese Annahme die zentrale Annahme für die Regressionsanalyse ist. Wird diese Annahme nicht getroffen, dann sind alle Ergebnisse der Regressionsanalyse wertlos, denn es kann keine Aussage darüber getroffen werden, was mit Y geschieht, wenn sich X verändert. (Kronthaler 2014)

- **Logistische Regression**

  Im Gegensatz zur linearen Regression, welche ein lineares Modell aufstellt, kann mit der logistischen Regression eine Eintrittswahrscheinlichkeit aufgezeigt werden, wohingegen die lineare Regression lediglich die Abhängigkeit der Variablen untereinander aufzeigt. (Tufféry 2011)

- **Multiple Regression**

  Bei der multiplen Regression liegt der Unterschied zur linearen Regression darin, dass die multiple Regression mehrere unabhängige Variablen hat. Diese werden in der Praxis meistens als $X_1, X_2, \dots, X_J$ bezeichnet. Das Ziel der multiplen Regression ist eine Beschreibung, welchen Einfluss die unabhängigen Variablen X_J, auf die abhängige Variable Y haben. Bei der multiplen Regression wird eben durch dieses Set an unabhängigen Variablen gleichzeitig mehrere Variablen simultan untersucht. (Tufféry 2011)

3.3 Assoziationsverfahren

Assoziationsverfahren gehören zur Klasse des „unsupervised learnings“. Die Verfahren des „unsupervised learnings“ besitzen keinen Trainingsdatensatz, mit welchen der Algorithmus angelernt werden kann. Beim Verfahren der Assoziation sollen Algorithmen Assoziationsregeln in bestehenden Daten aufzeigen. Diese Assoziationsregeln können logische Implikationen sein. Ein Beispiel hierfür wäre „Wenn eine Person Chips kauft, dann kauft diese Person wahrscheinlich auch Bier“. Diese Algorithmen werden sehr häufig bei Warenkorb-Analysen von Supermärkten oder Online-Shops angewendet. Mit Hilfe dieser Algorithmen soll aufgezeigt werden, welche Artikel gemeinsam gekauft worden sind. (Müller und Lenz 2013)

Die im Zuge des Assoziationsverfahren aufgestellten Regeln können somit strategische Vorteile für einen Supermarkt oder einen Online-Shop hervorbringen, indem Artikel begünstigt platziert werden, um das Zielpublikum zu beeinflussen. Aus diesem Grund ist das Assoziationsverfahren eines der beliebtesten Data Mining Verfahren, da mit diesem Verfahren sehr leicht Zusammenhänge zwischen Objekten und somit auch Wissen, gemäß dem KDD Prozess, generiert werden kann. (Zhong und Zhou 1999)

Wie bereits im Absatz zuvor erwähnt, werden, um die Korrelation zwischen zwei Artikel aufzuzeigen Assoziationsregeln aufgestellt. Diese werden in folgender Form dargestellt: $A \rightarrow B$ (wenn Artikel-Menge A, dann Artikel-Menge B). Die Menge der Artikel wird als $I = \{I_1, I_2, \dots, I_n\}$ und die relevanten Daten für die Fragestellung als D, dargestellt. Eine Transaktion $T \in D$ unterstützt eine Regel $A \rightarrow B$, wenn folgendes $(A \cup B) \subseteq \mathrm{T}$ gilt.

Die beiden wichtigsten Kenngrößen der Assoziation sind der Support und die Konfidenz. Der Support wird aus der Anzahl der Transaktionen, welche die Regel unterstützen, im Verhältnis zur Gesamtzahl aller Transaktionen, berechnet. (Han und Kamber 2006)

$$Support\ (A \rightarrow B) = \frac{|\{T \in D\ |(A \cup B) \subseteq \mathrm{T}\}|}{|D|}$$

Als Konfidenz wird das Verhältnis zwischen den Transaktionen, welche sowohl Prämisse als auch Konklusion enthalten, sowie den Transaktionen, welche nur Prämisse enthalten, bezeichnet.

$$Konfidenz\ (A \rightarrow B) = \frac{|\{T \in D\ |(A \cup B) \subseteq \mathrm{T}\}|}{|\{T \in D\ |A \subseteq \mathrm{T}\}|} = \frac{support(A \rightarrow B)}{support\ A}$$

Die Schwellwerte für den Support und die Konfidenz, welche nicht unterschritten werden sollten, werden je nach Fragestellung und Ziel festgelegt. (Han und Kamber 2006)

3.4 Clusterverfahren

Das Clusterverfahren gehört ebenfalls, wie das Assoziationsverfahren, zur Klasse der „unsupervised learning“. Das Clusterverfahren gehört, so wie die Klassifizierung, zu den fundamentalen Verfahren des Data Minings. Der Unterschied zwischen Klassifizierung und Clusterverfahren liegt jedoch darin, dass das Clusterverfahren ohne Trainingsdatensätze auskommt. Das Ziel des Clusterverfahrens soll darin bestehen neue Gruppen und Kategorien innerhalb der Daten zu entdecken, welche noch unbekannt sind. Somit steht bei der Clusteranalyse die Beschreibung der Daten im Vordergrund. Ähnliche Objekte sollen einer gleichen Gruppe, die auch als Cluster bezeichnet werden können, zugeordnet werden. Wenn die Objekte unterschiedlich sind, sollen diese einem anderen Cluster zugewiesen werden. Vorrangig behandelt die Clusteranalyse somit die Identifizierung der Gruppen und deren Aufteilung der Objekte in Cluster. (Maimon und Rokach 2010)

Des Weiteren wird das Clusterverfahren auch Segmentierung genannt, weil diverse Gruppen (Cluster) innerhalb der Daten gebildet werden, welche innerhalb der Gruppen so weit wie möglich homogen sein sollen. Mithilfe der Clusterverfahren werden oft Fragestellungen wie „Gruppierung der Kunden und Kundinnen nach demografischen Kriterien“ beantwortet.

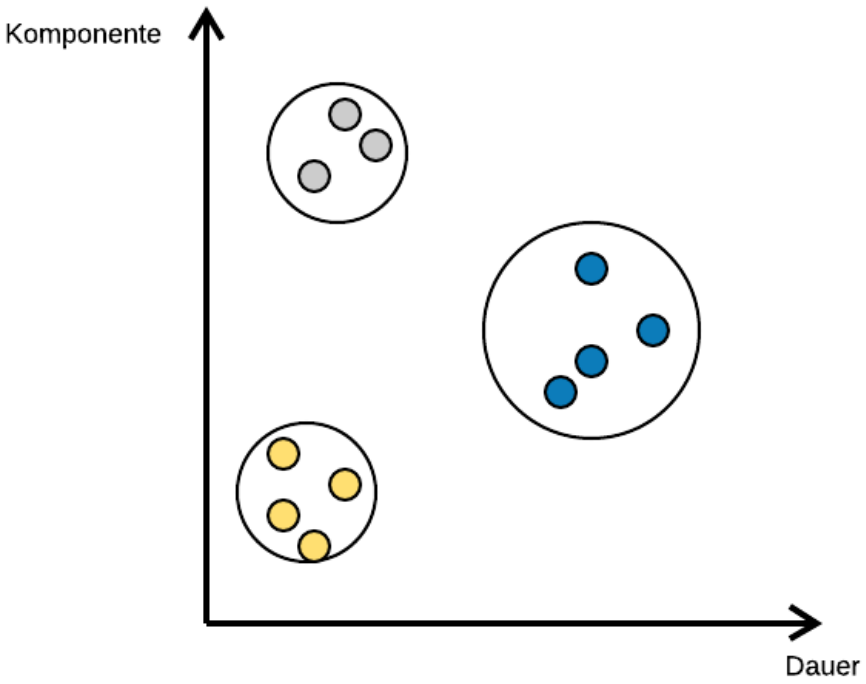

Abbildung 7 Beispiel Clusterverfahren

Abbildung 7 zeigt ein Beispiel, wie das Clusterverfahren in einem zweidimensionalen Raum vorgehen könnte. In der gezeigten Abbildung werden insgesamt drei Cluster identifiziert. Eine Zuordnung eines Datenobjektes erfolgt anhand dessen Abstand zu anderen Objekten. Die abgebildeten drei Cluster könnten zum Beispiel die Dauer und die Komponente einer Störung in einer Software darstellen. Die Störungen im gelben Cluster haben die geringste Dauer, wohingegen die Störungen des blauen Clusters eine längere Dauer haben.

3.5 Zusammenfassung

In diesem Kapitel wurden die folgenden Data Mining Verfahren beschrieben:

- Klassifikationsverfahren
- Regression
- Assoziationsverfahren
- Clusterverfahren

Des Weiteren wurde der Unterschied zwischen „supervised“ und „unsupervised learning“ Methoden beschrieben, welche eine Einteilung der Data Mining Verfahren in Verfahren mit Trainingsdatensätzen und ohne Trainingsdatensätzen darstellt. Bei der Beschreibung der einzelnen Verfahren wurden zudem Beispiele aufgezeigt, wo diese im praktischen Einsatz vorkommen können.

4 PREDICTIVE MAINTENANCE

Wie bereits in den Grundlagen definiert, gibt es verschiedene Methoden der Wartung. Bei der Methode der Predictive Maintenance, oder auch als vorhersagende Wartung bezeichnet, handelt es sich um eine Methode, welche in der Logistik sowie in der Industrie eingesetzt wird, um ungeplante Ausfälle zu verhindern. Bei dieser Methode werden Daten erzeugt, welche wiederum analysiert werden können. Die Erzeugung der Daten erfolgt mittels Technologien, welche nachfolgend erläutert werden. Es werden auch analytische Werkzeuge eingesetzt, um auffällige Muster innerhalb dieser Daten zu identifizieren sowie um ebenso den Zustand des Systems zu prognostizieren. Mit diesen Schritten werden benötigte Aktivitäten für eine Wartung erkannt und können so frühzeitig eingeplant beziehungsweise durchgeführt werden. Die Integration eines Predictive Maintenance Services in ein ERP (Enterprise Ressource Planning) System oder in ein WMS kann große Vorteile mit sich bringen, da durch die Erhöhung der Datenmenge weitere nützliche Zusammenhänge erhoben werden können und folglich wiederum Wissen generiert werden kann. (Huber und Kaiser 2017)

Bei modernen Softwaresystemen werden, durch Technologien, eine hohe Anzahl an Daten, welche Zustände, Durchlaufzeiten, Aufrufzeiten, Objektpositionen oder Bedingungen beinhalten, gespeichert und somit gesammelt. Durch das Erfassen dieser Daten können, wie bereits beschrieben, Auswertungen und Analysen dieser Daten über längere Zeiträume durchgeführt und detaillierte Vorhersagen über diverse Teilkomponenten des Systems getroffen werden. Hierfür werden auffallende Muster, wie eine länger andauernde Abfrage auf eine Datenbank oder ein zu langer Netzwerkzugriff gefiltert, um darauf frühzeitig reagieren zu können. Mit dieser Methode können Wartungseinsätze aufgezeigt, geplant und entsprechende Maßnahmen vorab getroffen werden. Durch die vorhersagende Wartung können längere Ausfälle oder Totallstillstände des Systems vermieden und reduziert werden. Dadurch, dass die Beobachtung der Kunden und der Kundinnen eines Systems meist nicht zu einem gewünschten Ergebnis der Wartungsintervalle und Prognosen führt, ist es sinnvoll, solche Services von dem Hersteller und der Herstellerin des Systems zu kaufen, wenn dies dem Kunden und der Kundin angeboten wird. (Bauernhansl et al. 2014)

Weiteres können sogenannte „Intelligente Sensoren“ eine weitere Verbesserung herbeiführen. Diese Sensoren ermöglichen es, dass Daten über Zustände über einen gewissen Zeitraum aufgezeichnet beziehungsweise gemessen werden. Zudem werden diese Sensoren meist auch mit einem Mikrochip, welcher mit einer entsprechend installierten Software und einer gewissen Intelligenz, das sogenannte Anwendungswissen, ausgestattet. Durch diese installierte Software und der somit ausgestatteten Intelligenz ergeben sich neue Möglichkeiten und Funktionen für die Erstellung von umfangreichen Datenauswertungen sowie Anpassungen an autonome Veränderungen und an der Kommunikation mit einem Netzwerk und anderen Softwaresystemen in einem Unternehmen. Die Logistikbranche legt jedoch den Fokus auf die Informationsgewinnung für die Planung, Optimierung und Steuerung der logistischen Abläufe, Prozesse, Maschinen und Anlagen. Die Einsatzmöglichkeiten solcher Sensoren sind sehr vielfältig. Diese reichen von Lichtschranken und Lichttastern mit geringerer Komplexität bis hin

zu Sensoren mit einer hohen Komplexität wie 2D- und 3D-Kamerasystemen, Laserscannern oder ähnlichen Sensorien. Die zuvor aufgezählten Sensoren sind in nahezu jeder automatisierten Logistikanlage eingebaut, was aufzeigen soll, welches große Potenzial für Predictive Maintenance in der Logistikbranche gegeben ist. (Vogel-Heuser et al. 2017)

4.1 Verwendung von Predictive Maintenance

Bei sachgemäßer Anwendung kann Predictive Maintenance nahezu unbegrenzte Vorteile bieten. Je nach Umfang und Art des Einsatzgebietes von Predictive Maintenance entstehen Vorteile und Verwendungsmöglichkeiten. Typischerweise wird Predictive Maintenance für eine der folgenden Verwendungen eingesetzt: (Mobley 2002)

- Maintenance Management Tool
- Werkzeug der Anlagenoptimierung
- Werkzeug für Verbesserung der Zuverlässigkeit

4.1.1 Maintenance Management Tool

Traditionell wird Predictive Maintenance ausschließlich als Instrument oder Maintenance Management Tool verwendet und eingesetzt. In den meisten Fällen beschränkt sich dieser Einsatz darauf, ungeplante Ausfallzeiten und/oder katastrophale Ausfälle zu verhindern. Obwohl diese Funktion wichtig ist, kann Predictive Maintenance wesentlich mehr Nutzen bringen, indem sie den Umfang oder den Auftrag des Ziels erweitert. Der Schwerpunkt des Tools sollte auf der Eliminierung unnötiger Ausfallzeiten, sowohl geplanter als auch ungeplanter Natur, sein. Des Weiteren sollten durch dies auch die Total-Life-Cycle Kosten reduziert werden. Studien, welche die Anzahl an Unternehmen und Anlagen untersuchten, zeigen auf, dass die Anzahl an Systemen, welche Predictive Maintenance als Wartungswerkzeug nutzen, immer weiter steigen. Einer der Hauptgründe für den Einsatz von Predictive Maintenance, als Werkzeug für die Wartungsplanung, ist, dass Wartungskosten und -zeiten reduziert werden. Weitere Gründe für den Einsatz sind ISO-Zertifizierungen sowie eine Steigerung der Qualität der Produkte. Es überwiegen aber nicht nur die Vorteile, bei dem Einsatz von Predictive Maintenance, sondern es fallen auch Kosten an, welche als Nachteil gesehen werden können. Bei einer Gegenüberstellung der Kosten muss jedoch immer bedacht werden, welchen finanziellen Schaden der Totalausfall eines Systems mit sich bringen würde. (Mobley 2002)

4.1.2 Werkzeug zur Systemoptimierung

Predictive Maintenance kann aber nicht nur als Werkzeug zur Planung von Wartungen eingesetzt werden, sondern die Vorteile von Predictive Maintenance können auch als Instrument zur Optimierung des Systems genutzt und ausgeschöpft werden. Durch den Einsatz von vorausschauender Wartung können Systemteile, welche häufige Probleme aufweisen, analysiert und dementsprechend auch eine Änderung durchgeführt werden. Solche Designveränderungen

des Systems könnten eine Lösung für immer wiederkehrende Probleme sein, welche langfristig die Performance des Systems steigern. Das Werkzeug der Systemoptimierung kann auch für Auswertungen und Vergleiche genutzt werden sowie aufzeigen, ob Systemprobleme zum Beispiel mit der Anzahl an Bewegungen oder der Auslastung eines Systems zusammenhängen. Durch den Erhalt dieser Information, über die Auslastung des Systems, können strategische oder sogar technische Maßnahmen erzeugt werden, welche die Verfügbarkeit des Systems erhöhen können. Zusammengefasst kann durch Predictive Maintenance dem Kunden und der Kundin eine Ursachen-Wirkungs-Analyse angeboten werden, worauf entsprechende Maßnahmen eingeleitet beziehungsweise auch durchgeführt werden können. (Mobley 2002)

4.1.3 Werkzeug für Verbesserung der Zuverlässigkeit

Ein weiterer Vorteil von Predictive Maintenance, gegenüber anderen Technologien ist, dass es als Werkzeug zur Verbesserung der Zuverlässigkeit eines Systems eingesetzt werden kann. Durch die Aufzeichnungen der Fehler und Ausfällen der Vergangenheit bietet es die Möglichkeit, dass negative Veränderungen der normalen Parameter des Betriebs schnell identifiziert werden können und diesen folglich entgegengewirkt werden kann. Durch diese Funktion ist es auch möglich, dass größere Veränderungen der Betriebsparameter vermieden werden können und die damit verbundenen langen Ausfallzeiten verhindert werden. Durch diese Art des Einsatzes von der vorausschauenden Wartung kann die Zuverlässigkeit des Systems langfristig und vor allem nachhaltig gesichert werden. (Mobley 2002)

4.2 Predictive Maintenance in Soft- und Hardware

Bei der vorbeugenden Wartung gibt es diverse Techniken, wie etwa das Messen und Überwachen von Hard- und Software. Ein häufig verwendetes Einsatzgebiet bei Predictive Maintenance im Hardwarebereich ist die Überwachung der Vibrationen eines Systems, die Thermographie oder auch die visuelle Inspektion des Systems. (Girdhar 2004) Bei Software wiederum ist eine Wartung schwieriger, weil mechanische Bauteile nicht vorhanden sind, dementsprechend können Komponenten auch nicht so einfach ausgetauscht oder gewartet werden. Zudem muss jede Software separat betrachtet werden, da stets individuelle Komponenten und Techniken vorhanden sind. Aus diesem Grund werden in den nachfolgenden Kapiteln zuerst die Einsatzgebiete von Predictive Maintenance Techniken in der Hardware, Performance Indikatoren in der Software und darauffolgend die Techniken von Predictive Maintenance in der Software näher erläutert.

4.2.1 Hardware Techniken

Wie bereits zuvor erwähnt, können Techniken, wie die Vibrationsüberwachung, eingesetzt werden, um den Zustand eines Systems zu überwachen. Die Überwachung der Vibration ist dabei die am meisten verwendete Technik, um Maschinen zu überwachen. Durch das Überwachen der Vibration können jedoch nicht alle Informationen bereitgestellt werden, um alle möglichen Fehlermöglichkeiten vorherzusagen. Demzufolge werden andere Techniken in Verbindung mit Predictive Maintenance Systemen eingesetzt. Dabei muss ein Predictive Maintenance System verschiedenen Techniken bei der Überwachung und Diagnose von Systemen einbinden. Die nachfolgend angeführten Techniken werden in der Theorie, wie folgt, beschrieben:

- **Thermographie**

  Durch den Einsatz von Thermographie entsteht die Möglichkeit, nicht nur den Zustand einer Maschine, sondern ebenso den Zustand von gesamten Anlagen oder Systemen zu überwachen, beobachten und zu diagnostizieren. Bei dieser Technik wird Temperatur von Anlagen und Maschinen gemessen. Dies kann auch auf gesamte Hallen oder Teilbereiche ausgeweitet und somit mit einem Standardwert verglichen werden. Mit Hilfe einer Feinanalyse sowie einer speziellen Hardware kann folglich die Ursache der veränderten Temperatur ausfindig gemacht werden. Bei der Technik der Thermographie kommen unter anderem Infrarot Thermometer, Infrarot Bilder und Line Scanner als Hardwareunterstützung zum Einsatz. (Mobley 2002)

- **Vibrationsüberwachung**

  Mit Hilfe der Überwachung, und der darauffolgenden Diagnose von Vibrationsveränderungen an Maschinen und Systemen, kann die aktuelle Bedingung analysiert werden. Ein großer Vorteil der Überwachung von Vibrationen liegt darin, dass Probleme bereits frühzeitig, bevor diese akut das System behindern, erkannt und auch behoben werden können. Der genannte Vorteil entsteht dadurch, dass die Aufzeichnungen der standardmäßigen Vibrationen, den auffälligen Ausreißern gegenübergestellt und dementsprechend die Unterschiede erkannt werden können. Des Weiteren können fehlerhaft durchgeführte Wartungsarbeiten sofort, anhand der Vibrationsänderung, vorzeitig erkannt werden. (Girdhar 2004)

- **Tribologie**

  Die Tribologie steht in der Industrie für die Wissenschaft der Reibung sowie für Verschleiß und Schmierung von gegeneinander bewegten Bauteilen. Im Einsatz mit Predictive Maintenance und Tribologie werden zwei verschiedene Technologien verwendet.

  Die erste Technologie ist die Analyse von Öl. Diese Technologie ist kein Tool, welches den direkten Zustand von Systemen und Maschinen wiedergibt oder ein mögliches bevorstehendes Fehlverhalten von einer Maschine aufzeigt. Indirekt wird aber auch durch diese Technik der Zustand einer Maschine oder eines Systems aufgezeigt, denn durch den Zustand des Schmieröles kann auf einen möglichen Fehlzustand der Maschine geschlossen werden, woraufhin auch mögliche Predictive Maintenance Maßnahmen

getroffen werden können. Das primäre Ziel der Analyse von Schmierölen einer Maschine ist jedoch die Qualitätskontrolle des Öls oder die Reduktion der Schmierölwartung. Durch diese Wartung der Maschine und durch das Erreichen des Zieles, wird die Wartungsqualität erhöht, wodurch sich auch eine erhöhte Lebenszeit des Systems erwartet wird. (Girdhar 2004)

Wie bereits erwähnt, besteht die Tribologie aus zwei Technologien. Die zweite Technologie ist die Analyse der Partikel. Die Analyse erfolgt auf Basis des Schmieröls. Dabei werden jedoch die Partikel im Öl analysiert und ausgewertet. Die Analyse des Öls untersucht somit den Zustand des Schmieröls. Durch die Partikelanalyse können demnach, im Gegensatz zur Schmierölanalyse, die Informationen über den Verschleißzustand eines Systems oder einer Maschine beobachtet werden. Die Partikel im bereits verwendeten Öl können folglich auf Wichtige Informationen des Zustands der Maschine hinweisen. (Girdhar 2004)

- **Visuelle Inspektion**

  Zu einer der ältesten Techniken gehört die visuelle Inspektion des Systems oder der Maschine, durch eine betreuende Person. In diesem Prozess werden keine datenbezogenen Informationen ausgewertet. Die betreuende Person kann durch ihre Erfahrung mögliche Auffälligkeiten frühzeitig erkennen. Die Methode der visuellen Inspektion kann aber auch unabhängig von einem Predictive Maintenance System durchgeführt werden. (Mobley 2002)

- **Ultraschall und Schallemission**

  Mithilfe der Ultraschallmethode kann die Ultraschalldichte gemessen werden, wodurch ein korrosiver auffallender Verschleiß prognostiziert werden kann. (Girdhar 2004) Des Weiteren kann durch den Ultraschall auch eine Schallemission gemessen werden. Durch die höheren Frequenzen können dementsprechend Kompressionen von Luft, Flüssigkeiten oder auch Gas dargestellt werden. Einer der wichtigsten Bestandteile eines Predictive Maintenance Systems, für Hardware, ist die Ultraschallmessung. Diese kann jedoch die Vibrationsmessung nicht ersetzen. (Mobley 2002)

4.2.2 Software

Um Leistung an einem Software System, und die daraus resultierende Leistungsveränderung messen zu können, müssen vorweg KPI (Key Performance Indicator) definiert werden. Mit Hilfe dieser Indikatoren kann ein System überwacht werden, um eine vorausschauende Prognose für die Zukunft zu tätigen und mögliche Änderungen am System durchzuführen. (Cai et al. 2009) Damit aber KPI definiert werden können, muss vorweg der Begriff Performance näher erläutert und betrachtet werden. Der Begriff der Performance stammt aus dem Englischen. Er beschreibt die Leistung beziehungsweise das Abschneiden von immateriellen Gegenständen und Personen. (Razum 2015)

In einer Software kann Performance, wie folgt beschrieben werden: „Performance is the degree to which a software system or component meets its objectives for timeliness" (Smith und Williams 2002) Unter dem Begriff Timeliness, im deutschen Rechtzeitigkeit, kann implizit verstanden werden, dass der Performance-Begriff im Zusammenhang mit Software einen starken Bezug zur Zeit hat. Die Rechtzeitigkeit lässt sich wiederum in die beiden Dimensionen Skalierbarkeit und Antwortverhalten unterteilen, welche im Folgenden näher beschrieben werden. Wird nun ein Informationssystem, wie ein WMS, näher betrachtet, lässt sich in der überwiegenden Mehrzahl ein interaktiver Charakter, sei es durch eine reine Maschine-Maschine-Interaktionen oder Mensch-Maschine-Interaktionen, feststellen. Es gibt also eine Kommunikation mit einem Initiator, der eine Eingabe oder Tätigkeit an ein System stellt und wiederum eine definierte Antwort vom System erwartet. Dieses Antwortverhalten hat dabei, bei dieser Art der Kommunikation mit einem IT-System, zwei Charakteristika. Einerseits die Antwortzeit und zum anderen den Durchsatz. (Smith 2007)

- **Antwortzeit**

  Die Definition der Antwortzeit ist das Intervall zwischen dem Absenden einer Anfrage an ein System und dem Erhalt einer Antwort nach der Bearbeitung des Systems. Die Antwortzeit lässt sich dabei in zwei Bestandteile aufspalten. Dies ist einerseits die Zeit für die Bearbeitung der Anfrage auf der Seite des Empfängers und anderseits die Zeit, welche für die Übertragung benötigt wird, um die Anfrage vom Sender zum Empfänger und wieder zurück zu übertragen. So kann sich die wahrgenommene Antwortzeit eines Systems mitunter signifikant, abhängig vom geografischen Ort des Benutzers und der Benutzerin, ändern. In den meisten Fällen ist die Übertragungszeit in lokalen Netzwerken geringer sowie stabiler als an externen Standorten. (Smith 2007)

- **Durchsatz**

  Der Durchsatz ist definiert durch den Quotienten aus erfolgreich bearbeiteten Anfragen pro Zeiteinheit. Dieser stellt aus Sicht der Benutzer und der Benutzerinnen oftmals eher einen sekundären Aspekt der Performance dar. Dies ist aber abhängig von der Natur der Anwendung des eingesetzten Systems. Betrachtet man zum Beispiel eine klassische Arbeitsplatzumgebung, dann spielt für den einzelnen Benutzer und der einzelnen Benutzerin die Tatsache, dass fünfzig weitere Arbeitskollegen und Arbeitskolleginnen gleichzeitig in einer Anwendung arbeiten können, eine untergeordnete Rolle, wenn die Antwortzeit des Systems nicht den eigenen Anforderungen und Ansprüchen genügt. Für einen Vorgesetzten wiederum kann der Durchsatz einer Applikation einen wesentlichen Faktor darstellen, da dies die Produktivität der vorhin genannten fünfzig Mitarbeiter und Mitarbeiterinnen wesentlich beeinflusst. (Smith 2007)

- **Skalierbarkeit**

  Ein gut skalierbares System besitzt die Fähigkeit, die Zielvorgaben für Antwortzeiten und dessen Durchsatz, auch bei steigenden Anforderungen durch Softwarefunktionen einzuhalten. Systeme, deren Antwortzeit bei linearem Anstieg der Anforderungen ein exponentielles Wachstum besitzen, gelten als nicht gut skalierende Systeme und sind in der Regel auch nicht performant. (Smith 2007)

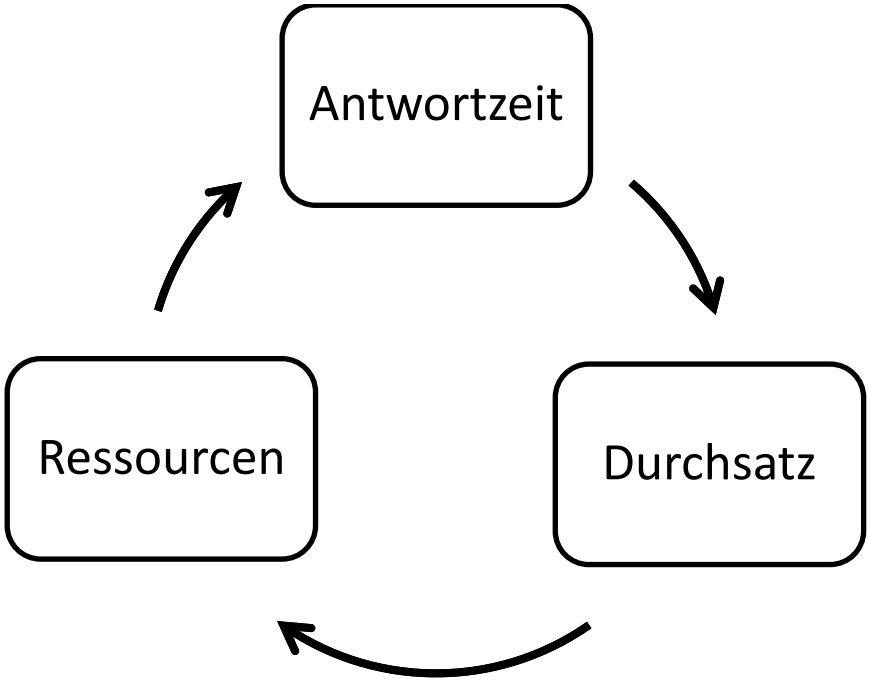

Abbildung 8 Aspekte der Performance

Wie in Abbildung 8 ersichtlich wird, ist aber nicht nur die Antwortzeit und der Durchsatz ein wichtiger Indikator, um eine Aussage über die Performance zu treffen, denn auch die Hardware selbst ist ein Indikator für die Performance eines Systems. Ein Software-System, welches eine simple Anwendung darstellt, aber ressourcentechnisch eine hohe Hardwareanforderung mit High-End Servern hat, wird ebenfalls als nicht performant bezeichnet. (Smith 2007)

Predictive Maintenance kann, wie bereits in den vorangegangenen Kapiteln erläutert, nicht nur in Hardware-, sondern auch in Softwaresysteme eingesetzt werden. Mit Hilfe von Predictive Maintenance können die nachfolgend klassifizierten Änderungsmethoden erkannt und zu einer weiteren Bearbeitung aufgenommen werden: (Grubb und Takang 2003)

- **Korrigierende Änderungen**

  Die korrigierenden Änderungen beziehen sich auf Änderungen, welche durch Fehler in der Software ausgelöst wurden. Ein Fehler kann aus Designfehlern, Logikfehlern und Entwicklungsfehler resultieren.

  All diese Fehler, manchmal auch als "residual errors" oder "Bugs" bezeichnet, verhindern, dass die Software der vereinbarten Spezifikation entspricht. Im Falle eines Systemausfalls, aufgrund eines Fehlers, werden Maßnahmen ergriffen, um den Betrieb des Softwaresystems wiederherzustellen. Unter dem Druck des Managements greift das Wartungspersonal manchmal auf Notfallbehebungen zurück, die als "Patching" bezeichnet werden. Der Ad-hoc-Charakter dieses Ansatzes führt zu einer Reihe von Problemen, zu denen eine erhöhte Programmkomplexität und unvorhergesehene Welleneffekte gehören. (Grubb und Takang 2003)

- **Perfektionierende Änderungen**

Der Begriff der perfektionierenden Änderung wird verwendet, um Änderungen zu beschreiben, die vorgenommen werden, um die bestehenden Anforderungen eines Systems zu erweitern. Eine erfolgreiche Software neigt dazu, einer Reihe von Änderungen unterzogen zu werden, die zu einer Erhöhung seiner Anforderungen und folglich zu einer gesteigerten Performance führen. Dies beruht auf der Prämisse, dass die Benutzer und Benutzerinnen, welche die Software nutzen, mit neuen Anforderungen experimentieren, die über den Rahmen der definierten Anforderungen hinausgehen, für den die Software ursprünglich entwickelt wurde. Die Erweiterung der Anforderungen kann die Form einer Erweiterung der bestehenden Systemfunktionalität annehmen oder zu einer Verbesserung der rechnerischen Effizienz führen, wie zum Beispiel die Bereitstellung eines Management-Informationssystems mit einem Dateneingabemodul oder einer neuen Einrichtung zur Nachrichtenverarbeitung. (Bowen und Hinchey 1999)

- **Präventive Änderung**

  Präventive Änderungen werden vorgenommen, um Fehlfunktionen zu verhindern oder um die Wartbarkeit der Software zu verbessern. Die Änderung ist normalerweise vom Instandhaltungsbetrieb mit der Absicht initiiert, Programme leichter verständlich zu machen und somit die zukünftigen Wartungsarbeiten einfacher zu gestalten. Präventive Änderungen führen in der Regel nicht zu einer wesentlichen Erhöhung der Basisfunktionalität.

  Beispiele für präventive Änderungen sind Code-Restrukturierung, Code-Optimierung und Aktualisierung der Dokumentation. Nach einer Reihe von Schnellkorrekturen an einem Softwaresystem, kann die Komplexität des Quellcodes auf ein unkontrollierbares Niveau ansteigen, was eine vollständige Umstrukturierung des Codes rechtfertigt. Eine Code-Optimierung kann durchgeführt werden, um die Programme schneller laufen zu lassen oder den Speicher effizienter zu nutzen. Die Aktualisierung der Benutzer- und Systemdokumentation wird zwar häufig ignoriert, ist aber oft notwendig, wenn ein Teil der Software geändert wird. (Grubb und Takang 2003)

Die im Kapitel Hardware Techniken und Software beschriebenen Techniken können trotz deren Einsatz zu einen nicht funktionierenden Predictive Maintenance System führen. Der wichtigste Aspekt eines solchen Systems, so wie auch in jedem anderen Informationssystem, ist die durchgängige Unterstützung vom Management bis zum Endbenutzer und zur Endbenutzerin, eine gute Planung sowie gut geschulte Mitarbeiter und Mitarbeiterinnen, welche das System täglich mit voller Überzeugung nutzen. Der Einsatz der verschiedenen Techniken und Methoden ist ebenfalls von der jeweiligen Branche des eingesetzten Systems abhängig. (Girdhar 2004)

4.3 Zusammenfassung

In diesem Kapitel wurden die Grundlagen für Predictive Maintenance definiert und näher erläutert. Im ersten Schritt wurden die Verwendungen von Predictive Maintenance, wie folgt aufgezeigt:

- Maintenance Management Tool
- Werkzeug zur Systemoptimierung
- Werkzeug für Verbesserung der Zuverlässigkeit

Im zweiten Schritt wurden die Techniken in Hardware und Software näher erläutert. Für Techniken der Hardware wurden folgende Bereiche erläutert:

- Thermographie
- Vibrationsüberwachung
- Tribologie
- Visuelle Inspektion
- Ultraschall und Schallemission

Im dritten Schritt wurden vorerst folgende Key Performance Indicators definiert, welche eine Aussage über die Performance von Softwaresystemen liefern:

- Antwortzeit
- Durchsatz
- Skalierbarkeit

Der letzte Schritt beinhaltete Methoden, welche bei einer Softwareänderung eingesetzt werden können:

- Korrigierende Änderung
- Perfektionierende Änderung
- Präventive Änderung

5 ANWENDUNGSGEBIETE

In diesem Kapitel soll aufgezeigt werden, welche Einflussfaktoren durch Data Mining in der Logistikbranche durch Predictive Maintenance für ein Warehouse Management System aufgestellt werden können. Dies beinhaltet die Datenerhebung für das Data Mining, die Anforderungen der Logistik an Predictive Maintenance sowie Beispiele wo, Predictive Maintenance mit den erhobenen Daten angesetzt werden kann.

5.1 Datenerhebung in der Logistik

Die Anzahl an Datenquellen, welche für Data Mining verwendet werden können, ist nahezu unendlich. Aus diesem Grund gibt es eine Kategorisierung von Menschen, Systemen und Maschinen generierten Daten. Die generierten Daten sind wiederum für die im Kapitel 3 erwähnten Data Mining Verfahren wichtig, denn aus diesen können Resultate abgeleitet werden. Die Quellen der Kategorisierungen Mensch, System und Maschine können, wie nachfolgende beschrieben, kombiniert werden. (Dapp 2014, S. 9)

- **Von Menschen generierte Daten:** Unter diesen Daten versteht man all jene Daten, die durch Eingaben von Menschen entstehen. Darunter zählen Formulare, Bilder, Audio, Video und Protokolldaten. In Bezug auf ein WMS wären diese Daten zum Beispiel noch ein per Hand befülltes Formular oder händisch geführte Strichlisten.
- **Von Systemen generierte Daten:** Diese Daten werden während eines Geschäftsprozesses eines Systems oder Informationssystems erzeugt. In einem Warehouse Management System wären dies Daten, welche von einem Benutzer und einer Benutzerin während des Prozesses anfallen. Auch alle Tätigkeiten, welche während der Ausführung des Prozesses digital unterstützt werden, zählen zu den vom System generierten Daten. Dies kann unter anderem Zeiten von Auftragsannahme bis Auftragsabgabe oder ein erfasstes Gewicht sein.
- **Von Maschinen generierte Daten:** All jene Daten, wie Sensor-, Logdaten, Klickstatistiken, aber auch Video-, Sprach- und Audioaufzeichnungen, zählen zu den von Maschinen generierten Daten. In der Logistikbranche könnten dies Sensorbelegungen auf einer Fördertechnik oder eben auch Logdaten von Prozessen eines Warehouse Management Systems sein.

5.2 Einsatzgebiete in der Logistikbranche

Im Folgenden werden die, wie im Kapitel 4.2 beschriebenen, Techniken und Anwendungsgebiete von Predictive Maintenance mit den erfassten Daten in der Logistikbranche aus Kapitel 5.1 verknüpft. Demnach ergeben sich folgende Einsatzgebiete:

- Von Menschen generierte Daten bei der visuellen Inspektion mittels geführter Listen

- Von Maschinen generierte Daten der Tribologie mittels Sensoren auf der Maschine
- Von Maschinen generierte Daten der Thermographie mittels eingebundener Sensoren auf der Maschine
- Von Maschinen generierte Daten der Vibrationsüberwachung mittels eingebundener Sensoren auf der Maschine
- Von Maschinen generierte Daten von Ultraschall und Schallemissionen mittels eingebundener Sensoren auf der Maschine
- Von Systemen generierten Daten der Antwortzeiten mittels Messungen in der Software zwischen Sendezeitpunkt und Empfangszeitpunkt.
- Von Systemen generierte Daten des Durchsatzes mittels Messungen in der Software zwischen Auftragsannahme und Auftragsabschluss

5.3 Zusammenfassung

In diesem Kapitel wurden die Anwendungsgebiete der Logistik und deren Datenerhebung aufgezeigt:

- Von Menschen generierte Daten
- Von Systemen generierte Daten
- Von Maschinen generierte Daten

Nachfolgend wurden die Einsatzgebiete der Logistikbranche, durch eine Verknüpfung der Datenerhebungen und der Techniken, aufgezeigt.

6 ZUSAMMENFASSUNG THEORIETEIL

In Kapitel 2 wurden die Grundlagen und Begriffe näher erläutert, welche in der weiteren Arbeit benötigt werden. Dies beinhaltet die nachfolgenden Themengebiete

- Grundlagen des Data Minings
- Grundlagen Maintenance
- Grundlagen Warehouse Management System

Im Abschnitt der Grundlagen des Data Minings wurden die Prozesse Knowledge Discovery in Databases und Cross-Industry Standard Process for Data Mining näher erläutert. Des Weiteren wurde im Abschnitt Grundlagen Maintenance der Unterschied zwischen Run-To-Failure, Preventive Maintenance, Proactive Maintenance und Predictive Maintenance dargestellt. Im letzten Abschnitt des zweiten Kapitels wurden die Grundlagen des Warehouse Management Systems und was die Kennzahlen eines solchen Systems sind sowie wofür diese eingesetzt werden können, näher erläutert.

Im Kapitel 3 des Theorieteils wurden die folgenden Data Mining Verfahren näher beschrieben beziehungsweise ein Unterschied zwischen supervised learning, welche mit bereits bekannten Daten arbeiten, und dem unsupervised learning, welche keine Eingangsgrößen kennen aufgezeigt:

- Klassifikationsverfahren
- Regression
- Assoziationsverfahren
- Clusterverfahren

Das vierte Kapitel, dieser Arbeit, legt den Fokus auf die Wartungsmethode der Predictive Maintenance. In diesem wird unter anderem die Verwendung von Predictive Maintenance, wie Maintenance Management Tools, Werkzeuge zur Anlagenoptimierung sowie Werkzeuge zur Verbesserung der Zuverlässigkeit dargestellt. Unter anderem werden auch Techniken in Hardware, wie die Thermographie, Vibrationsüberwachung oder die Tribologie näher erläutert. Ein weiterer Abschnitt beschreibt den Einsatz von Predictive Maintenance in einer Software und wie diese überwacht werden kann. Dabei wurden folgende Themengebiete in den Mittelpunkt der Betrachtung gestellt:

- Antwortzeit
- Durchsatz
- Skalierbarkeit

Die angeführten Faktoren können dann wiederum auf die verschiedenen Änderungsmethoden angewendet werden, mit welcher eine Softwarewartung klassifiziert wird:

- Korrigierende Änderung

- Perfektionierende Änderung
- Präventive Änderung

Im Kapitel 5, und somit dem letzten Kapitel des Theorieteils, wurde aufgezeigt, welche Daten es in der Logistik gibt und wie diese erhoben werden können sowie deren Einsatzgebiete.

7 EINLEITUNG PRAXISTEIL

In den nachfolgenden Kapiteln soll aufgezeigt werden, dass die zuvor erarbeitete Theorie eine Relevanz für die Logistik und für andere Unternehmen in dieser Branche schafft. Im ersten Kapitel des Praxisteils wird ein generischer Teil, welcher die Merkmale der Logistikbranche darstellt, aufgezeigt. Des Weiteren werden die Herausforderungen sowie Anforderungen der Logistikbranche näher erläutert und die Einflussfaktoren von Data Mining auf Predictive Maintenance näher beschrieben.

7.1 Merkmale der Logistikbranche

„Einer der bedeutendsten Megatrends unserer Zeit ist ohne Zweifel die Digitalisierung. Ebenso wie beispielsweise die Globalisierung führt sie zu nachhaltigen Veränderungen, die eine Vielzahl von Lebensbereichen durchdringen.“ (Bousonville 2017, S. 1) Wie in diesem Zitat von Bousonville (2017) erkenntlich wird, werden zwei große Veränderungen beschrieben, welche ein Unternehmen einerseits unter Druck setzt, aber anderseits auch Chancen bietet, sich durch die Veränderung der Rahmenbedingungen in Zukunft von dessen Konkurrenz abzuheben. Die Logistik muss auch in Zukunft Transport-, Lager- und Materialflussprozesse bewerkstelligen. Dabei muss sie sich verändern beziehungsweise hat sich in den letzten Jahren bereits sehr stark verändert. Es wurden Technologien und Schlagworte wie autonome Systeme, Big Data, Data Mining oder auch intelligente Objekte, welche sich unter dem Begriff Logistik 4.0 zusammenfassen lassen, den Logistiksystemen hinzugefügt. Diese beeinflussen das wettbewerbsgetriebene Wirtschaftssystem der Logistik zunehmend. (Bousonville 2017)

Zudem steigt die Anzahl an Domänen, in welchen der Begriff „4.0“ adaptiert wird, immer stärker an. Begriffe wie Education 4.0, Health 4.0, Government 4.0 oder auch Logistik 4.0 werden immer präsenter. Diese Begriffe wurden erstmals 2011 in Deutschland vom Begriff Industrie 4.0 abgeleitet, welcher die vierte industrielle Revolution kennzeichnen soll. Die drei vorangegangenen industriellen Revolutionen sind durch die nachfolgenden Veränderungen entstanden: (Bauernhansl et al. 2014)

- Erste industrielle Revolution: Dies war die Einführung der Dampfmaschine und die darauffolgende Mechanisierung von Produktionen sowie die Entwicklung von Eisenbahnen
- Zweite industrielle Revolution: Darunter wird die Nutzbarmachung der Elektrizität und der Einsatz von neuen Organisationsformen zur Massenproduktion verstanden.
- Dritte industrielle Revolution: Beinhaltet die Automatisierung der Produktion durch den immer höher werdenden Einsatz von Elektronik und Computertechnologien in Form von CNC-Maschinen und industriellen Robotern.

Die vierte und aktuell anstehende industrielle Revolution ist durch den Trend der Digitalisierung getrieben. Die treibenden Faktoren und Technologien für die Digitalisierung ist dabei die

Mikroelektronik wie Sensoren, Datenübertragungstechniken, Leistungssteigerungen oder auch die Speicherung von großen digitalen Datenmengen. Ein weiterer Begriff, welcher oft im Zusammenhang mit der vierten Industriellen Revolution genannt wird, ist „Internet der Dinge“. Dieser Begriff beschreibt die hohe Anzahl an digital vernetzen Sensoren, Objekten und Devices. Das Internet der Dinge oder auch IOT, Internet of Things, ist eine Grundlage für die Logistik 4.0, da durch dieses Prinzip eine umfassende Bereitstellung von digitalen Informationen in der Logistikbranche entsteht. Zum aktuellen Zeitpunkt ist die Vernetzung und die daraus resultierenden Informationen der verschiedenen Objekte noch nicht vollständig. Im Endausbau dieses Systems sollte ein Austausch der Informationen durch die Komponenten im System existieren. Diese Informationen der Objekte können verschieden sein, wie die Eigenschaften von Identifikationsnummern oder Größen bis zu Zustandsgrößen wie Orte, Positionen oder Füllgrade. Dadurch, dass die beschriebene Vernetzung immer weiter voranschreitet, werden die Speicherung der zuvor genannten Eigenschaften und Zustandsdaten der Ausstattung der Objekte mit Sensorien in immer mehr autonome Systeme mit intelligenten Geräten eingebunden und entwickelt. Diese Optimierung und dementsprechend die Veränderung durch die Logistik 4.0 ist notwendig, um die aktuell vorherrschenden und zukünftigen Branchenveränderungen positiv zu nutzen und zu überstehen. (Bousonville 2017)

7.2 Anforderungen der Logistikbranche

Ein zentraler und bedeutender Aspekt dieser Arbeit, und der darin enthaltenen Forschungsfragen ist die Logistik. Aus diesem Grund wurde die Logistik schon vorab in der Theorie erläutert. Ein weiteres damit verbundenes Thema ist das Supply Chain Management, aus welchen sich im Anschluss Herausforderungen und Anforderungen der Logistikbranche ableiten lassen.

In der Wirtschaft ist „Logistik“ ein sehr junger Begriff. Erstmals wurde der Logistikbegriff 1590 in den USA und 1970 in Europa verwendet. Seitdem ist die Verbreitung und Verwendung dieser Branche stark wachsend. Die Logistik wurde erstmals im militärischen Bereich im 19. Jahrhundert verwendet. Der Terminus selbst stammt vom französischen Wort „Logis“ ab, welches die Truppenunterkunft bezeichnet. Eine weitere Verbindung zur Herkunft des Wortes stammt aus dem griechischen Begriff „Logos“, von welchem das Wort „Logik“ abgeleitet werden kann. Die beschriebenen Wortherkünfte stellen die Nutzung des Begriffs im militärischen Sprachgebrauch dar. Da unter dieser Definition die Logistik mit der Planung des Nachschubs, der Bewegungen der Truppen und mit der Versorgung der Truppen einhergeht. (Arnold et al. 2008)

Heute wird eine Vielzahl an Definitionen hinsichtlich der Verwendung des Logistikbegriffes eingesetzt. Nachfolgend werden Definitionen der Aufgaben und Anforderungen der Logistik, aus der Theorie, dargestellt:

- „Effizientes Bereitstellen der geforderten Mengen benötigter Objekte in der richtigen Zusammensetzung zur rechten Zeit am richtigen Ort. Logistikobjekte sind Handelswaren, Lebensmittel, Rohstoffe oder Material und vieles mehr. Zu den Logistikleistungen zählen Transport, Umschlagen, Lagern und Kommissionieren.“ (Gudehus 2011, S. 3)

- „Unter (Unternehmens-) Logistik versteht man eine ganzheitliche, die einzelnen Funktionsbereiche im Unternehmen übergreifende Betrachtungsweise. Das Ziel ist die Optimierung der Materialflüsse sowie damit im Zusammenhang stehender Informationsflüsse. Die Logistik ist daher eine typische Querschnittsfunktion, da sie den gesamten Materialfluss des Unternehmens – vom direkten Zulieferer über das Lager und die Produktion bis zum Kunden und der Kundin - betrachtet und optimiert.“ (Benz und Höflinger 2011, S. 9)
- „Logistische Prozesse sind alle Transport- und Lagerungsprozesse sowie das zugehörige Be- und Entladen, Ein- und Auslagern (Umschlag) und das Kommissionieren. Sie lassen sich zusammenfassend dadurch charakterisieren, dass sie auf eine „bedarfsgerechte Verfügbarkeit von Objekten ausgerichtet sind oder abstrakter als Raumüberbrückung (Transport), Zeitüberbrückung (Lagerung) und Veränderung der Anordnung (Kommissionierung) der Objekte.“ (Arnold et al. 2008, S. 3)

Die angeführten Definitionen zeigen das oberste Ziel der Logistik auf, und zwar die Logistikobjekte in richtiger Zusammensetzung, am richtigen Ort und zur richtigen Zeit zu transportieren sowie mit die Durchführung der Raumüberbrückung. Dieses Ziel wird mit Hilfe von Logistikprozessen wie lagern, umschlagen, kommissionieren und transportieren erreicht.

Beim Supply Chain Management wird sogar einen Schritt weiter gegangen, denn dieses behandelt nicht nur die übergreifenden Funktionsbereiche im gleichen Unternehmen. Ein Ziel des Supply Chain Management ist es, eine ganzheitliche und übergreifende Betrachtungen beziehungsweise Prozessverbesserungen der Logistikkette vom Anfang bis zum Ende zu erreichen. (Benz und Höflinger 2011). Einer der Hauptgründe der Entstehung von Supply Chain Management war, und ist, die Globalisierung. Zudem sind die weltweit zunehmenden Verflechtungen der Leistungserstellung, welche mit einer hohen Komplexität der Herstellungsprozesse einhergehen, mitunter ein Grund für die Entstehung. Ein weiterer Aspekt, welcher die Entwicklung des Prozesses des Supply Chain Management vorangetrieben hat, sind die gestiegenen Erwartungen der Kunden und Kundinnen. (Arnold et al. 2008)

Auch der Begriff Supply Chain Management wird in der Theorie auf verschiedene Arten definiert:

- „Supply Chain Management bezeichnet die Koordination von Aktivitäten entlang der Wertschöpfungskette – von der Produktentwicklung des Lieferanten bis zum Beziehungsmanagement des Händlers.“ (Arnold et al. 2008, S. 21)
- „Das Supply Chain Management ist auf die Lieferprozesse von den Lieferanten der Lieferanten durch das eigene Unternehmen bis zu den Kunden und Kundinnen der Kunden und Kundinnen ausgerichtet“ (Gudehus 2011, S. 10)
- Hauptziel des Supply Chain Managements ist die übergreifende Prozessverbesserung der gesamten logistischen Ketten. Der Idealfall bezieht sich auf die Verbesserung vom Lieferanten des Rohstoffes, über die Zulieferer und das Unternehmen bis hin zum Kunden und Kundinnen. Ausgangslage für die Bestrebungen ist das eigene Unternehmen. Dabei sollen durchgängige und übergreifende Prozesse realisiert und gesteuert werden können. (Benz und Höflinger 2011)

Wie anhand der oben angeführten Definitionen zu erkennen ist, ist es von Bedeutung, die gesamte Wertschöpfungskette zu managen, damit Qualität und Kosten optimiert werden können und so die damit verknüpfte Zufriedenheit der Kunden und Kundinnen gesteigert werden kann. Bei Supply Chain Management wird aber nicht immer die gesamte Wertschöpfungskette betrachtet. Ein Grund hierfür ist die nicht vorhandene Transparenz sowie die Macht, über Prozesse und Aktivitäten von Lieferanten und Lieferantinnen und Händler und Händlerinnen, zu beeinflussen. Sehr häufig ist es schon schwierig Aktivitäten und Prozesse von Abteilungen und Bereichen eines Unternehmens anhand des Supply Chain Managements auszurichten. Das ist auch ein Grund warum unternehmensübergreifende Supply Chain Managements häufig nur durch Vertragskonditionen mit einem Hersteller und einer Herstellerin umgesetzt werden kann. (Arnold et al. 2008)

7.3 Herausforderungen der Logistikbranche

Durch den gesellschaftlichen Wandel sowie durch die Globalisierung und deren Einfluss auf die Wirtschaft, werden Konsumverhalten verändert und verstärkt. Der Wunsch des Kunden nach individuelleren Produkten, welche eine hohe Qualität aufweisen und niedrige Kosten verursachen, steigt stetig an. Durch den globalisierten Handel von Produkten und Dienstleistungen, und der im Hintergrund verbundenen Logistik, steigt die Komplexität von Logistiksystem stark. Dadurch wird die Logistik sowie das damit verbundene Supply Chain Management zu einen der wichtigsten Bereiche in einem Unternehmen. Die zuvor erwähnten steigenden Anforderungen und die damit verbundene steigende Komplexität der Systeme in der Logistik, kann nur durch neuartige Technologien der Logistik 4.0 ausgeglichen werden. Dadurch kann die Wettbewerbsfähigkeit gesteigert und weiterhin gewährleistet werden. (Hompel und Kerner 2015)

Infolgedessen gibt es neben der Globalisierung, und dem damit im Zusammenhang stehenden veränderten Konsumverhalten der Kunden und Kundinnen, noch weitere wirtschaftliche Veränderungen. Diese werden als Megatrends bezeichnet, welche bereits zu enormen Veränderungen in der Logistikbranche geführt haben, und auch in Zukunft führen werden: (Lehmacher 2015)

- **Klimawandel**

  Ein immer wichtiger werdendes Thema, auch in der Logistikbranche, ist der Klimawandel und der Anstieg der CO_2 Emissionen. In der Logistik werden zum Beispiel Transportmittel bereits mit Gas oder Strom angetrieben. Zudem werden Themen der Last Mile bereits bearbeitet, indem Elektrofahrzeuge in Stadtgebieten für die Zustellung von Paketen eingesetzt werden.

- **Urbanisierung**

  Die im vorherigen Punkt genannten Probleme der Last Mile, werden durch die Urbanisierung und durch das Anwachsen von Metropolen, immer größer werden. Aber nicht nur der CO_2-Austoß stellt ein großes Problem dar, sondern auch verkehrsbedingte

überfüllte Straßen in Innenstadtgebieten, wodurch eine pünktliche Zustellung über die Last Mile nahezu unmöglich wird. Eine mögliche Entlastung könnten hier Untergrundbahnen und Stadtnetze leisten. Erste Städte und Projekte nutzen bereits umgebaute Straßenbahnen und öffentliche Verkehrsmittel, wie U-Bahnen und Busse, für die Zustellung und den Transport von Paketen und Sendungen.

- **Sicherheit**

  Die Sicherheit ist auch in der Logistik, in der heutigen Zeit, ein brandaktuelles Thema, denn die globale Supply Chain ermöglicht es für Hacker und Angreifer von außen, eine Sabotage am System durchzuführen und somit die Geschäftsfähigkeit des Unternehmens zu beeinträchtigen.

- **E-Commerce**

  Die wesentlichsten und essenziellsten Faktoren des E-Commerce, ist das Internet und die Logistik. Immer mehr Händler im Internet wollen die Transportdienstleistung selbstständig durchführen und nicht an externe Dienstleister vergeben. Durch eine selbst durchgeführte Logistik, erzielen große Onlinehändler heute bereits hohe Gewinne, welche sonst an einem Transportunternehmen weitergegeben werden würden. Durch diesen Trend, kann davon ausgegangen werden, dass in den nächsten Jahren, die bekannten Onlinehändler zu den größten Transportriesen anwachsen und folglich diesen Geschäftsbereich selbständig abdecken werden.

- **Digitalisierung**

  Durch die Digitalisierung wurden Verfahren wie der 3D-Druck immer beliebter. Um lange Wartezeiten und große Transportwege zu vermeiden und zu reduzieren, gibt es bereits Logistikunternehmen, welche den 3D-Druck nutzen, um Ersatzteile für Reparaturen von Maschinen und Geräten selbständig herstellen, beziehungsweise zu drucken. Somit können weite Transportwege von Ersatzteilen, zum Beispiel von Asien nach Europa, vermieden und Kosten verringert werden.

  Zudem werden sich die Prozesse der Logistik, durch die gestiegene Mobilität der Welt, stark verändern. Mit Hilfe von Smartphones können bereits jetzt berufliche und private Tätigkeiten an jedem Ort durchgeführt werden. Aus diesem Grund, können sehr viele Menschen auch nicht mehr voraussagen, wo sie wann sind. Das beeinflusst wiederum den Last Mile Transport und steigert seine Komplexität. Damit gewinnen Zustell- und Abstellboxen an gut zugänglichen Orten oder an Wohnungen immer größerer Beliebtheit. Somit können Zustellungsprozesse auch außerhalb der Öffnungszeiten durchgeführt werden.

Des Weiteren beschreiben Jeske, Grüner & Weiß (2014) ähnliche Herausforderungen der Logistikbranche:

- **Verstädterung**

  Wie in den oberen Abschnitten näher beschrieben, steigt die Anzahl an Metropolen immer weiter an. Durch das erhöhte Verkehrsaufkommen und die erhöhte Lieferdichte steigen

die Wahrscheinlichkeiten für eine Staubildung, welche die Wahrscheinlichkeit einer pünktlichen Lieferung wiederum sinken lässt. Neben den zuvor erwähnten Möglichkeiten des Transportes für die Last Mile in den innerstädtischen Regionen, werden Big Data, Data Mining und Datenanalysen für Logistikunternehmen ein immer wichtiger werdendes Thema.

- **Volatilität**

  Die Planung in der Logistikbranche ist durch die ständig ansteigenden Bedingungen der Wirtschaft, im Vorhinein, kaum einschätzbar. Demzufolge ist die Planung in der Logistik schwer durchführbar. Durch den vorhin bereits erwähnten Einstieg der Onlinehändler in das Logistikgeschäft, könnten sich Logistikunternehmen in Zukunft stark verändern.

- **Erwartungen der Kunden und Kundinnen**

  Der Kunde und die Kundinnen haben eine stetig anwachsende Erwartungshaltung durch die stetig steigenden angebotenen Optionen des Marktes. Durch diese Möglichkeiten werden Services wie Next-Day-Delivery oder flexible Lieferoptionen erst möglich gemacht, welche aber ebenso einen hohen Druck auf Logistikunternehmen erzeugen. Ein negatives Erlebnis mit einem Lieferservice könnte einem Kunden und einer Kundin zu einem Wechsel des Onlineshops verleiten. Durch diesen Qualitätsdruck müssen Technologien wie die Sendungsverfolgung, Kühlkettenmanagement und ähnliche, in den Prozess der Transportdienstleister übernommen werden.

Die angeführten gesellschaftlichen und wirtschaftlichen Veränderungen zeigen die Herausforderungen der Logistikbranche. Der Wandel der Logistik, welcher mit diesen einhergeht, kann von verschiedenen Unternehmen positiv, aber auch negativ in das Geschäftsmodell des Unternehmens übernommen beziehungsweise integriert werden.

7.4 Einflussfaktoren von Data Mining auf Predictive Maintenance

Im folgenden Abschnitt wurde durch eine Kombination von Kapitel 7.2, Kapitel 7.3 und Kapitel 4.2 folgende Kriterien für die Einsatzgebiete von Predictive Maintenance im Logistikbereich ausgearbeitet:

- **Minimierung von Durchlaufzeit**

  Durch den Einsatz von Predictive Maintenance und Data Mining sollen Stillstände vermieden werden, indem Sensoren von Maschinen überwacht und Auffälligkeiten frühzeitig gemeldet werden.

- **Optimierung der Durchlaufzeit**

  Durch die Überwachung des Durchsatzes der letzten Tage, Wochen und Monate des Systems, soll bereits frühzeitig erkannt werden, ob das System performant genug läuft.

- **Optimierung der Materialflüsse**

Bei Störungen an Maschinen oder Softwaresystemen werden Transportwege vom Benutzer und von den Benutzerinnen, und nicht vom System, entschieden. Dadurch können längere Wegstrecken durch Teilausfälle auf Fördertechniken auftreten.

- **Optimierung von Sicherheit:**

  Durch Überwachung von Antwortzeiten im Netzwerk können frühzeitig Sicherheitslücken erkannt werden, um vor dem Auftritt einer potenziellen Sicherheitslücke eine Fehlerbehebung durchzuführen.

- **Minimierung von CO_2 Ausstoß**

  Mit Hilfe der Thermographie und Tribologie sollen Anlagen und deren CO_2 Ausstoß überwacht werden.

7.5 Hypothesen

In diesem Abschnitt werden die Hypothesen, welche als Grundlage für den nachfolgenden Interviewleitfaden dienen, erstellt. Die folgenden Hypothesen wurden auf Basis, der am Anfang dieser Arbeit definierten Alternativ- und Nullhypothese, der Data Mining Methoden und der Einsatzbereiche von Predictive Maintenance Services erarbeitet:

Alternativhypothese (H1): Durch den Einsatz von Data Mining in der Logistikbranche können Störungen in einem Warehouse Management System mittels Predictive Maintenance früher erkannt werden

Nullhypothese (H0): Durch den Einsatz von Data Mining in der Logistikbranche können Störungen in einem Warehouse Management System mittels Predictive Maintenance nicht früher erkannt werden

Hypothese 2 (H2): Wenn die Regression auf den Durchsatz eines Systems über Tage, Monate und Wochen angewendet wird, kann eine Veränderung der Durchlaufzeiten und somit eine frühzeitige Störung oder ein Problem vorausgesagt werden. Mit Hilfe von präventiven Maßnahmen kann dies verhindert werden.

Hypothese 3 (H3): Durch die Klassifikation können Antwortzeiten des Systems kategorisiert und somit Sicherheitslücken frühzeitig erkannt werden und Sicherheit optimiert werden.

Hypothese 4 (H4): Wenn die Regression bei der Thermographie und Tribologie angewendet wird, kann vorausgesagt werden, dass sich der CO_2 Ausstoß bei einer Erhöhung der thermischen Werte auch steigern wird. Durch den Einsatz von vorbeugenden Maßnahmen kann dies unterbunden werden.

7.6 Zusammenfassung

Dieses Kapitel ist die Einleitung für den zweiten und praxisorientierten Teil der Arbeit. Es wurden die Begriffe Logistik sowie Supply Chain Management und deren Unterschied näher erläutert.

Des Weiteren wurden die Herausforderungen der Logistikbranche, abgeleitet von den Anforderungen der Logistikbranche, näher betrachtet, welche folgende Punkte beinhalten:

- Klimawandel
- Urbanisierung
- Sicherheit
- E-Commerce
- Digitalisierung
- Verstädterung
- Volatilität
- Erwartungen der Kunden und Kundinnen

Im letzten Abschnitt dieses Kapitels wurden erste Einflussfaktoren von Data Mining auf Predictive Maintenance aufgestellt, bestehend aus folgenden Punkten:

- Minimierung von Durchlaufzeit
- Optimierung der Durchlaufzeit
- Optimierung der Materialflüsse
- Optimierung von Sicherheit
- Minimierung des CO2 Ausstoßes

Durch die Einflussfaktoren konnten, danach im letzten Abschnitt dieses Kapitels, Hypothesen aufgestellt werden.

8 ERSTELLUNG DES INTERVIEWLEITFADENS

Im folgenden Kapitel wird ein Interviewleitfaden für Interviews mit Experten und Expertinnen aus der Logistikbranche ausgearbeitet. Hierfür wird eine allgemeine Vorgehensweise sowie der Aufbau, für die Erstellung von Interviewleitfäden, von qualitativen Interviews ausgearbeitet und erforscht. Im zweiten Schritt werden die zuvor aufgestellten Einflussfaktoren und Erkenntnisse verwendet, um anhand dieser einen Interviewleitfaden zu erarbeiten. Dieser Interviewleitfaden wird im Anschluss in Interviews mit Experten und Expertinnen aus der Logistikbranche verwendet.

8.1 Allgemeine Vorgehensweise und Aufbau bei der Erstellung eines Interviewleitfadens für qualitative Interviews

In diesem Kapitel wird ein Interviewleitfaden für Interviews mit Experten und Expertinnen aus der Logistikbranche ausgearbeitet. Im ersten Schritt wird die allgemeine Vorgehensweise sowie der Aufbau, für die Erstellung von Interviewleitfäden, von qualitativen Interviews ausgearbeitet und untersucht. Im nächsten Schritt werden, die zuvor erarbeiteten Einflussfaktoren und Erkenntnisse, kombiniert und daraus ein Interviewleitfaden erarbeitet. Der daraus resultierende Interviewleitfaden kommt im Anschluss in Interviews mit Experten und Expertinnen aus der Logistikbranche Zur Anwendung. Dabei dient der Interviewleitfaden als Basis und Richtline und stellt bei allen Interviews die Grundlage dar, um eine gute Vergleichbarkeit der Interviews gewährleisten zu können.

Im ersten Schritt werden Kategorien gebildet anhand welcher, im zweiten Schritt, Fragen zu den Kategorien formuliert werden. (Vogt und Werner 2014) Bei diesen Fragen gibt es eine Unterscheidung der Art der Fragen, die für ExpertInneninterviews sowie für die semi-strukturierte Gestaltung des Interviewleitfadens sinnvoll ist. Anhand dieser Fragearten ergibt sich folglich der Aufbau des Interviewleitfadens, welcher bei einem Interview mit Experten und Expertinnen streng durchgearbeitet werden muss. Mittels der weiter angeführten unterschiedlichen Arten und Schemen der Fragen wird einerseits eine Struktur und anderseits eine offene Diskussion erlangt: (Baur und Blasius 2014)

- Fragen-Antwort-Schema
- Erzählaufforderung-Erzählung-Schema

Dieser Interviewleitfaden kann eine Liste von vorgegebenen und in definierter Reihenfolge zu stellenden Fragen, das Fragen-Antwort-Schema sein. Auch kann dieser aus Erzählaufforderung-Erzählung-Schema Fragen bestehen. Zudem ist eine Kombination dieser beiden Schemen möglich. (Baur und Blasius 2014)

Da Fragen, welche mit einem Fragen-Antwort-Schema definiert wurden, strukturiert sind, ist bei diesen Fragen auch weniger Spielraum für Erzählungen und Diskussionen gegeben. Dadurch können mögliche, zu starke und ungewollte, Abschweifungen von den Fragen vermieden werden.

Das Gegenstück dazu, stellt das Erzählaufforderung-Erzählung-Schema dar. Dieses ist per Definition offener. Dadurch kann wiederum die Expertise und die Erfahrung der Befragten zur Geltung kommen. Werden beide Schemen kombiniert, kann ein Interviewleitfaden, durch die nachfolgenden drei Schritte, definiert werden: (Baur und Blasius 2014)

- **Schritt eins**

  Im ersten Schritt wird dem Befragten oder der Befragten, also dem Experten oder der Expertin die Möglichkeit gegeben, sich so frei, wie nur möglich, zu äußern. Die Erzählaufforderung wird dabei so formuliert, dass bei dem Experten oder der Expertin die Chance besteht, möglichst viele Informationen und interessante Aspekte, welche für die Forschung relevant sind, auch spontan zu erfragen.

- **Schritt zwei**

  Im zweiten Schritt werden nur Aspekte hinterfragt, bei welchen, durch die erste Frage, kein ausreichender Kontext erzeugt werden konnte. Die ersten beiden Schritte können wiederholt eingesetzt werden.

- **Schritt drei**

  Der dritte Schritt beinhaltet, in einem abschließen Teil des Interviews, strukturierte und, in der Formulierung, vorgegebene Fragen.

Zudem können Fragen ausformuliert werden, wobei vorgegeben werden kann, ob und wie flexibel diese Formulierungen übernommen werden können. Die Kombination der offenen Erzählaufforderung, mit der Offenheit der näher zu bestimmenden Nachfrage, hilft, die nachfolgenden drei grundlegenden Anforderungen eines Leitfadens zu erfüllen: (Baur und Blasius 2014)

- **Offenheit als Priorität**

  Eine Intervention als Einschränkung der Äußerungsmöglichkeit, muss so offen wie nur möglich gestaltet werden. Diese müssen ebenso mit dem Forschungsinteresse begründet werden.

- **Übersichtlichkeit**

  Der Leitfaden muss übersichtlich gestaltet sein. Im besten Fall so übersichtlich, dass dieser im Kopf behalten werden kann. Ein Überfluss an Fragen beschränkt die, für das Generieren von Texten, notwendige Erzählzeit.

- **Anschmiegen an den Erzählfluss**

  Das Erstellen eines Leitfadens soll dem Erinnerungs- und Argumentationsfluss folgen. Es soll verhindert werden, dass abrupte Sprünge oder Themenwechsel stattfinden. Die Priorität sollte hier bei spontanen Äußerungen und nicht bei der Einhaltung der Reihenfolge der Fragen liegen. Wenn ein Nachfragen von der interviewten Person angeboten wird, egal ob direkt oder indirekt, dann sollte dies nur aus wichtigen Gründen nicht angenommen werden.

Die Erstellung eines Interviewleitfadens hat, neben dem konkreten Vorgehen, den zweiten Zweck, dass das Forschungsinteresse realisiert und reflektiert werden kann. Somit kann eine Fixierung von Detailfragen gelöst werden, welche eine Haltung der Offenheit blockieren. Die nachfolgenden vier Schritte sind dabei mit der SPSS Formel leicht zu merken: (Baur und Blasius 2014)

- **Sammeln von Fragen („S“)**

  Es sollen möglichst viele Aspekte der Forschung, als Fragen, zusammengetragen werden, ohne etwaige Einschränkungen oder Berücksichtigungen der methodischen Formulierung und Eignung.

- **Prüfen der Fragen („P“)**

  Dieser Punkt beinhaltet die kritische Hinterfragung, der zuvor aufgestellten Liste mit einem Rückbezug auf die Forschung selbst. Dabei kann es sich um Fragestellungen wie „Öffnen die Fragen oder Erzählaufforderungen die Äußerungsmöglichkeiten, sodass Neues und Fremdes zur Sprache kommen kann?“, „Ermöglichen sie die Generierung von Texten?“ handeln. Nach der Überprüfung können auch Fragen gestrichen, oder mit anderen Fragen ersetzt werden.

- **Sortieren („S“)**

  Die verbleibenden Fragen müssen nach zeitlicher Folge inhaltlich zusammengehörig, nach Fragerichtung gebündelt und sortiert werden.

- **Subsumieren („S“)**

  Für die so entstandenen Bündel an Fragen wird ein möglichst erzählgenerierender Impuls gesucht und in einem Einzelaspekt subsumiert. Somit soll die Möglichkeit bestehen, dass einzelne Aspekte spontan angesprochen werden können.

8.2 Interviewleitfaden für Interviews mit Experten und Expertinnen

Die zuvor ausgearbeiteten Erkenntnisse und Ergebnisse, welche bereits in Kapitel 5 ausgearbeitet wurden, werden für den Interviewleitfaden genutzt. Durch die zuvor ausgearbeitete Literatur sowie durch die Kombination des in Kapitel 8.1 erarbeitete Vorgehensweise bei der Erstellung eines qualitativen Interviewleitfadens, wird auf Grundlage dessen ein Leitfaden für Interviews mit Experten und Expertinnen, aus der Logistikbranche, erstellt.

Im Ersten Schritte werden, wie bereits in Kapitel 8.1 beschrieben, Kategorien für die Strukturierung des Interviewleitfadens gebildet. Die Bildung dieser Kategorien erfolgt mit Hilfe einer inhaltlichen (Analysemodell) und formellen (zeitlicher Aspekt) Strukturierung. Dabei wurden die Kategorien aus Kapitel 3 und Kapitel 4 anhand der aufgezeigten theoretischen Grundlagen und Hypothesen erarbeitet.

Im Folgenden erfolgt eine Aufzählung an Kategorien, welche für die Struktur des Interviewleitfadens dienen:

- Kurzfristige Optimierung
- Mittelfristige Optimierung
- Langfristige Optimierung

Im zweiten Schritt werden zu den zuvor aufgestellten Kategorien Fragen erarbeitet. Durch die verschiedenen Frageschemen und- arten, wird eine strukturierte Grundlage für das Interview geschaffen. Zudem wird das Themengebiet, durch diese Vorgehensweise, eingegrenzt. Dabei bleibt jedoch ein Spielraum für Diskussionen und Abschweifungen bewusst bestehen. Dieser Spielraum gilt insbesondere den Erfahrungen und Expertisen der Interviewpartner und Interviewpartnerin. Die Fragen werden, wie zuvor in Kapitel 8.1 beschrieben, nach einem definierten Prozess mit „Erzählaufforderung-Erzählung-Schema“ und „Fragen-Antwort-Schema“ gebildet. Diese müssen jedoch die Prozessschritte nicht strikt einhalten, da die verschiedenen Kategorien und Fragen eine alternative Bedeutung für die Evaluierung und für das Ergebnis haben.

Für die Formulierung der Fragen wird folgendes Vorgehen durchgeführt: Im ersten Schritt werden Aspekte aus den Hypothesen, wie Datenerhebung in der Logistik, Data Mining Verfahren sowie die Verwendung von Predictive Maintenance, mit dem Hauptziel Einflussfaktoren aufzuzeigen, kombiniert. Dafür werden Einsatzgebiete der Data Mining Verfahren sowie Einsatzgebiete von Predictive Maintenance operationalisiert, um in weiterer Folge daraus Fragen formulieren zu können.

Die jeweils erste Frage, pro Hypothese, wird immer eine „Erzählaufforderung-Erzählung-Schema“-Frage sein. Dadurch sollen so viele Inhalte und Aspekte, wie nur möglich, besprochen werden. Auch sollen der Befragte und die Befragte die Option bekommen, sich so frei wie möglich zu äußern. Dabei wird nach den Möglichkeiten gefragt, welche Einflussfaktoren unter der Nutzung der Einsatzgebiete und der Data Mining Verfahren, der jeweiligen Hypothese, es gibt und somit eine Optimierung erzielt werden kann. Aus der folgenden Hypothese wurden die Einsatzgebiete, des Data Mining Verfahrens und des Optimierungspotentials operationalisiert sowie die erste Fragestellung erstellt:

> **Kategorie:** Langfristige Optimierung.
>
> **H1:** Durch den Einsatz von Data Mining in der Logistikbranche können Störungen in einem Warehouse Management System mittels Predictive Maintenance früher erkannt werden.
>
> **Datenquellen:** Von Systemen generierte Daten wie Auftragsannahmen sowie von Maschinen generierte Daten wie Temperaturen.
>
> **Data Mining Verfahren:** Regression.
>
> **Einsatzgebiet**: Sensoren an Maschinen und Historie an vergangenen Bestellungen.
>
> **Einflussfaktoren**: Minimierung von Durchlaufzeiten.
>
> **Erste Frage:** Wie können Ihrer Meinung nach, Sensoren für die Temperaturmessung an Maschinen genutzt werden, um Störungen zu minimieren und somit Durchlaufzeiten zu minimieren?

Im zweiten Schritt der Hypothese wird eine spezifischere Frage formuliert, welche das jeweilige Data Mining Verfahren in die Frage integriert. Die Frage wird als „Erzählaufforderung-Erzählung-Schema“-Frage formuliert und wird dann dem Befragten und der Befragten gestellt, wenn die Inhalte nicht bereits in der ersten Frage angesprochen worden sind. Der Inhalt der zweiten Frage soll die Auswirkung auf die Durchlaufzeit, durch die Nutzung der Datenquellen und der Einsatzgebiete, beinhalten. Die zweite Frage für das zuvor genannte Beispiel der Hypothese 1 lautet wie folgt:

> **Zweite Frage:** Welche Auswirkung auf die Durchlaufzeit einer Bestellung sehen Sie beim Einsatz von Regression und Sensoren an Maschinen, bei den zuvor genannten Datenquellen?

Die Dritte Frage der Hypothese soll im „Fragen-Antwort-Schema“ formuliert sein, um damit den Themenbereich weiter einzugrenzen und folglich die Hypothese abzuschließen. Aus diesem Grund ist die Frage spezifisch in Anlehnung an die Hypothese definiert. Damit eine Überprüfung der Hypothese erfolgen kann, ist kaum Spielraum für Diskussionen oder Abschweifungen gegeben. Zudem erfolgt eine Kombination der gesamten operationalisierten Aspekte, der jeweiligen Hypothese, in der Frage. Dabei wird direkt gefragt, ob nach Meinung des Experten und der Expertin mithilfe der operationalisierten Optimierungspotentialen, durch die Nutzung der jeweiligen Datenquellen, Einsatzgebiete und Data Mining Verfahren, die Durchlaufzeit minimiert werden kann. Die dritte Frage der zuvor genannten Hypothese folglich:

> **Dritte Frage:** Denken Sie, dass durch Temperaturmessungen an Maschinen, in Kombination mit Regression, eine Minimierung der Durchlaufzeit erreicht werden kann, indem Stillstände vermieden werden und der operative Betrieb verbessert wird?

Die oben angeführte Vorgehensweise wird für alle Hypothesen angewendet. Anhand der Antworten kann eine Evaluierung durchgeführt werden, ob die Experten und Expertinnen einen Bezug zwischen den Datenquellen, Einsatzgebieten, Data Mining Verfahren und den, in den Hypothesen formulierten, Einflussfaktoren für möglich halten.

8.3 Zusammenfassung

In diesem Kapitel wurde ein Interviewleitfaden ausgearbeitet. Dieser Interviewleitfaden soll in weiterer Folge für Interviews mit Experten und Expertinnen aus dem Logistikbereich eingesetzt werden. Bei der Erstellung des Interviewleitfadens wurde im ersten Schritt eine allgemeine Vorgehensweise sowie der Aufbau eines Interviewleitfadens für qualitative Interviews aufgelistet. Im Anschluss wurde dies mit den zuvor, aus der Literatur, ausgearbeiteten Erkenntnissen und Ergebnissen kombiniert, wodurch ein Interviewleitfaden für Interviews mit Experten und Expertinnen aus der Logistikbranche entstanden ist.

Im nachfolgenden Kapitel wird die Durchführung und Evaluierung mit den Experten und Expertinnen aus der Logistikbranche erarbeitet. Zunächst erfolgt eine Auswahl und eine Beschreibung der Logistikexperten und Logistikexpertinnen. Im Weiteren werden Interviews, anhand des definierten Interviewleitfaden, durchgeführt und dokumentiert. Als abschließender

Schritt erfolgt eine Auswertung der geführten Interviews. Dadurch soll eine vergleichende Gegenüberstellung der Interviewergebnisse sowie der Ergebnisse aus der vorherigen theoretischen Arbeit ermöglicht werden.

9 DURCHFÜHRUNG UND EVALUIERUNG DER INTERVIEWS

In diesem Kapitel wird die Durchführung und Evaluierung der Interviews beschrieben. Zuerst wird die Auswahl der Experten und Expertinnen, aus dem Logistikbereich, dargestellt, um dadurch deren Expertise im Fachgebiet aufzuzeigen. Danach erfolgt eine Auswertung der geführten Interviews. Zuletzt werden die Ergebnisse der Interviews, mit den aufgestellten Hypothesen gegenübergestellt und schließlich in einer Conclusio zusammengefasst.

9.1 Auswahl von Experten und Expertinnen aus der Logistikbranche

In diesem Abschnitt werden Experten und Expertinnen aus dem Logistikbereich, für die Analyse, bestimmt. Zudem wird deren Tätigkeiten in ihrem Unternehmen dargestellt. Des Weiteren werden deren Erfahrungen in der Logistikbranche sowie in den angeführten Themengebieten beschrieben. Mittels dieser Vorgehensweise soll sichergestellt werden, dass es sich, bei den ausgewählten Interviewpersonen, um Experten und Expertinnen aus dem Logistikbereich handelt. Folgende Kriterien wurden ausgewählt, um einen Status als Experte und Expertin zu erlangen:

- Mehrjährige Erfahrung mit Logistikprojekten
- Erfahrung mit Warehouse Management Systemen
- Erfahrung im Einsatz von Predictive Maintenance
- Kenntnisse von Data Mining und deren Prozesse

9.1.1 Interviewpartner 1

Interviewpartner 1 ist seit vielen Jahren in der Logistik tätig. Sein aktuelles Unternehmen ist einer der größten Schweizer Retailer. Zudem ist Interviewpartner 1 Leiter der Informatik in der Abteilung Prozesse, Warenwirtschaft, Logistik und Supply Chain Management. Diese Abteilung ist in allen WMS Großprojekten involviert und hauptsächlich für die Prozessdefinition und Projektleitung in der Logistik verantwortlich.

Aufgrund seiner Funktion, der Leitung der Informatik, hat dieser Interviewpartner einen umfassenden Einblick in die Prozesse von WMS Systemen und ebenso bereits Erfahrung mit Predictive Maintenance Systemen. Des Weiteren hat Interviewpartner 1 langjährige Erfahrungen als Projektleiter für IT-Projekte gesammelt. Aus diesem Grund eignet sich diese Person besonders gut, um als Experte für das Interview, dieser Arbeit, in Frage zu kommen.

9.1.2 Interviewpartner 2

Dieser Interviewpartner ist Geschäftsführer eines KMU (Klein- und Mittelunternehmen) in Österreich, welche Cloud basierte Lösungen für Logistikunternehmen sowie Versicherungen

anbietet. Diese Lösungen behandeln zum Großteil die Bereiche Predictive Maintenance, Verteilte Systeme beziehungsweise Microservices. Des Weiteren hat Interviewpartner 2 langjährige Erfahrungen als Softwareentwickler für Warehouse Management Systemen und ist somit mit den Prozessen der Logistik sehr gut vertraut.

9.1.3 Interviewpartner 3

Der dritte Interviewpartner ist bei einem der größten österreichischen Logistikdienstleister angestellt und beschäftigt sich mit dem Bereich Data Analytics. In seiner Position als Product Developer und Data Scientist, beschäftigt er sich mit der Aufbereitung von Daten und wie diese für Entscheidungen besser eingesetzt werden können. Dabei ist dieser Interviewpartner nicht nur für interne Prozessverbesserungen zuständig, sondern auch für die Entwicklung von Kundenprodukten, bei welchen unter anderen die Techniken von Data Mining und Predictive Maintenance zum Einsatz kommen.

Durch den Einblick in die Prozesse der Logistik sowie durch die Expertise in der Erfassung und Verarbeitung von Logistikdaten ist Interviewpartner 3 optimal als Experte für ein Interview geeignet.

9.1.4 Interviewpartner 4

Interviewpartnerin 4 arbeitet seit mehr als 20 Jahren in der Logistik und ist aktuell bei einem der größten deutschen Lebensmittelhändler seit über 10 Jahren beschäftigt. In ihrer derzeitigen Position, ist sie für die Einführung von Warehouse Management Systemen sowie für die Konzeptionierung neuer Geschäftsprozesse innerhalb dieser Systeme zuständig. In ihren letzten Projekte wurden teilweise schon Predictive Maintenance Systeme an Warehouse Management Systeme angebunden, weswegen für die Interviewpartnerin 4, die Bereiche der Logistik sowie die Bereiche der Predictive Maintenance und des Data Minings grundlegend bekannt sind.

9.2 Allgemeine Auswertung von qualitativen Interviews

Die Auswertung der Interviews wird mit Hilfe einer zusammenfassenden Inhaltsanalyse, welche eine von drei qualitativen Methoden der Textauswertungen nach Mayring (2015) darstellt, durchgeführt. Die Analyse des Textes wird durch ein von Regeln geleitetes und systematisches Verfahren reduziert. Diese Vorgehensweise der Inhaltsanalyse ist methodisch ähnlich der Zusammenfassung als Interpretationsform, die Menschen auch unbewusst im tagtäglichen Leben anwenden. Im Gegenzug zur wissenschaftlichen Methode, wird mit Hilfe von nachvollziehbaren Regeln dies subjektiv durchgeführt. Die nachfolgend angeführten Schritte, werden bei der zusammenfassenden Inhaltsanalyse nach Mayring (2015) angewendet:

- **Bestimmung des Ausgangsmaterials**

  Beim Schritt der Bestimmung des Ausgangsmaterials wird als Erstes der zu analysierende Text festgestellt. Hierfür wird sehr häufig eine Auswahl von sehr großem Textmaterial

getroffen, wobei die Grundgesamtheit von Aussagen und Meinungen erhalten bleiben soll. Als zweiter Schritt wird beschrieben, von welchen Personen und unter welchen Umständen und Bedingungen das Material produziert wurde. Im dritten und letzten Schritt wird dargestellt, in welcher Form das Material vorliegt. Wenn ein Interview aufgenommen worden ist und erst im weiteren Schritten in eine Textform gebracht wurde, müssen in diesem Abschnitten die Regeln der Transkription dokumentiert werden.

- **Fragestellung der Analyse**

  Im Weiteren werden, bei der Auswertung von qualitativen Interviews, die Fragestellung sowie die Richtung der Analysen aufgezeigt. Bei der Vorgehensweise gibt es verschiedene Möglichkeiten, wie etwa die Auswahl des im Text behandelten Gegenstands, des Autors und der Autorin der Texte, oder der Textwirkung auf eine Zielgruppe. Dabei soll die Fragestellung der Analyse, sprich die Forschungsfrage, durch die Analysen bestätigt werden. Mit Hilfe dieses Abschnittes soll somit eine Eingrenzung erfolgen, um eine Interpretation der Analyse und deren Fragestellung zu erzielen.

- **Ablaufmodell der Analyse**

  Im dritten Schritt nach Mayring (2015), wird eine Technik für die Analyse bestimmt. In der vorliegenden Arbeit ist dies bereits durch die „Zusammenfassende Inhaltsanalyse“ festgelegt. Zudem wird für diese Analysetechnik ein Ablaufmodell beschrieben, welches im ersten Schritt die Analyseeinheiten, welche die Kodiereinheit, die Kontexteinheit sowie die Auswertungseinheit beinhalten, festgelegt. Im zweiten Schritt werden die, zuvor festgelegten, Kodiereinheiten in kurze und auf deren Inhalt beschränkte Form umschrieben. Hierfür werden die Z-Regeln nach Mayring (2015) eingesetzt. Die Z1 Regel wird im ersten Schritt eingesetzt, welche auch Paraphrasierung genannt wird. In den weiteren Schritten werden die Ergebnisse aus der Paraphrasierung verallgemeinert und in Folge mit Hilfe der Z2 Regeln nach Mayring (2015), welche auch als Generalisation bezeichnet werden, ausgewertet. Im dritten Schritt erfolgt eine Minderung, durch Selektion und Streichung, der nicht bedeutungsvollen Paraphrasen. Dies wird mithilfe der Z3 Regel durchgeführt. In einem vierten Schritt werden anhand der Z4 Regel nach Mayring (2015) die Aussagen erneut gebündelt und zusammengefasst. Diese beschriebene Vorgehensweise kann so oft durchgeführt werden, bis ein hohes Abstraktionsniveau entstanden ist und ein gewünschtes Endergebnis erzielt wurde.

- **Interpretation und Überprüfung**

  In einem abschließenden Schritt, der zusammenfassenden Inhaltsanalyse, erfolgt eine Überprüfung der Ergebnisse, durch eine Interpretation.

9.3 Auswertung der Interviews mit den Logistikexperten

Bei der Inhaltsanalyse, in der vorliegenden Arbeit, wird nach der Inhaltsanalyse von Mayring (2015) vorgegangen. Dabei wird in einer reduzierten sowie abgeänderten Vorgehensweise gearbeitet. Das Ausgangsmaterial wurde unter der Berücksichtigung der Auswahl der Interviewpartner und Interviewpartnerinnen und durch die Durchführung mittels Interviewleitfaden gewonnen. In einem weiteren Schritt, welcher die Definition der Fragestellung, der Analyse, beinhaltet, werden die zuvor aufgestellten Hypothesen evaluiert.

In weiterer Folge werden die Antworten der Experten und Expertinnen aus den Interviews reduziert, selektiert und in abstrahierter Form wiedergegeben. Dadurch wird eine Generalisation durchgeführt. Diese ist unter anderem notwendig, um eine Evaluierung der Hypothesen sicherstellen zu können. Die Auflistung der Hypothesen erfolgt anhand der gleichen Kategorien, welche bereits im Interviewleitfaden ausgearbeitet und aufgestellt worden sind:

- Kurzfristige Optimierung
- Mittelfristige Optimierung
- Langfristige Optimierung

9.4 Evaluierung der Interviews mit Experten und Expertinnen aus der Logistikbranche

Im folgenden Abschnitt wird eine Überprüfung der Ergebnisse aus den Interviews, mit den Experten und Expertinnen aus der Logistikbranche, durchgeführt. Dabei werden deren Aussagen reflektiert und kombiniert wiedergegeben, um Erkenntnisse über die jeweilige Hypothese zu erhalten. Dieser Schritt wird mittels der zuvor definierten Kategorien und der dazugehörigen Hypothese durchgeführt. Danach werden in einer Tabelle, pro Hypothese, die Kernaussagen der Interviewpartner und Interviewpartnerinnen zusammengefasst.

Evaluierung der Alternativ- und Nullhypothese

H1: Durch den Einsatz von Data Mining in der Logistikbranche können Störungen in einem Warehouse Management System mittels Predictive Maintenance früher erkannt werden.

H0: Durch den Einsatz von Data Mining in der Logistikbranche können Störungen in einem Warehouse Management System mittels Predictive Maintenance nicht früher erkannt werden

Interviewpartner 1: Interviewpartner 1 sieht ein großes Potential durch den aktiven Einsatz von Data Mining und Predictive Maintenance im Bereich der Logistik. Vor allem durch den Einsatz von modernen IOT Geräten kann kostengünstig ein Sensor an eine Maschine oder an ein System angebracht werden, um verschiedene Informationen zu messen. Das können einerseits wirtschaftliche Faktoren wie die Auslastung und Durchlaufzeiten einer Maschine sein. Anderseits können aber auch Temperaturen, für die Überwachung von Kühlketten oder von Hallen gemessen werden.

Interviewpartner 2: Der zweite Interviewpartner denkt, dass jede Maßnahme, welche eine Störung in einem System erkennen soll, für den Durchsatz einer Maschine oder für ein Systems hilfreich ist. Bei dem Einsatz von Data Mining und Predictive Maintenance sieht Interviewpartner 2, allerdings lediglich eine langfristige Optimierung. Denn der Aufwand, für die Implementierung eines Systems, welches auch mehrere Umgebungsvariablen wie die Auslastung der Mitarbeiter und Mitarbeiterinnen oder die Auslastung des Systems berücksichtigen muss, muss auf einen längeren Betrachtungszeitraum angelegt werden, um richtige Daten zu erhalten. Dabei bedarf es ebenso genügend Personal, welches diese Störungen dann frühzeitig beheben kann.

Interviewpartner 3: Interviewpartner 3 denkt, dass dieses Thema definitiv bei automatisierten, aber auch nicht-automatisierten Systemen eine Daseinsberechtigung hat. Ein Problem, für die Predictive Maintenance, stellen aber die vielen verschiedenen Fehlermeldungen beziehungsweise die Meldungen einer Maschine oder einer Anlage dar. Laut den Erfahrungsberichten, von Interviewpartner 3, gab es in vergangenen Projekten immer wieder das Problem, dass eine Meldung, welche eine Maschine beziehungsweise ein System geworfen hat, immer verschiedene Lösungen beinhaltet hat. Dabei müssen diese Meldungen mit anderen Ereignissen oder Zuständen kombiniert werden, damit eine Fehlerursache ausfindig gemacht werden kann. Sind diese Quellen aber einmal identifiziert und bekannt, dann kann durch den Einsatz von Data Mining, mittels Predictive Maintenance, eine Störung frühzeitig erkannt werden.

Interviewpartner 4: Interviewpartnerin 4 erzählt, dass sie im Zuge ihres letzten Projektes ein Predictive Maintenance System, in Zusammenhang mit einem Warehouse Management System, eingeführt hat. Durch diese Umsetzung, des Projektes, wurden Störungen frühzeitig erkannt beziehungsweise auch Störungen ausfindig gemacht, welche in den letzten Jahren immer wieder vereinzelt vorgekommen sind, aber keine Lösung für das ursächliche Problem gefunden wurde. Durch den Einsatz von Predictive Maintenance und Data Mining konnte bereits frühzeitig ein Problem erkannt werden und spezialisiertes Personal, die Störungen analysieren und somit den langjährigen Fehler des Systems ausfindig machen.

Evaluierung der Alternativ- und Nullhypothese
▪ Großes Potential durch den Einsatz von Data Mining und Predictive Maintenance. ▪ Einsatz von kostengünstigen IOT Geräten für eine langfristige Optimierung und eine hohe Kostenersparnisse. ▪ Langfristiges Projekt, welches nicht sofort umgesetzt werden kann. ▪ Beachtung von vielen verschiedenen Zuständen eines Systems. ▪ Predictive Maintenance kann frühzeitig Fehler aufzeigen, welche noch nicht erkannt worden sind.

Tabelle 2 Evaluierung der Alternativ- und Nullhypothese

Kurzfristige Optimierung

Evaluierung H4: Wenn die Regression bei der Thermographie und Tribologie angewendet wird, kann vorausgesagt werden, dass sich der CO_2 Ausstoß, bei einer Erhöhung der thermischen Werte, auch steigern wird. Durch den Einsatz von vorbeugenden Maßnahmen kann dies unterbunden werden.

Interviewpartner 1: Interviewpartner 1 ist der Meinung, dass beim Einsatz von Sensoren und Regression, im konkreten Beispiel bei den Motoren einer Maschine, frühzeitig erkannt werden kann, ob ein Motor richtig läuft. Dies kann unter anderem durch die Überprüfung der Wärme und somit auch rückwirkend auf den CO_2 Ausstoß erreicht werden. Was wiederum die Wahrscheinlichkeit eines Ausfalls einer Maschine verringert, beziehungsweise eine frühzeitige Meldung ermöglicht, mit welcher einer Wartung angekündigt werden kann.

Interviewpartner 2: Der Interviewpartner 2 denkt, dass durch den Einsatz von Sensoren und Regression, die thermischen Werte überwacht werden können. Somit kann eine Erhöhung der Werte festgestellt und frühzeitig agiert werden. Interviewpartner 2 erzählt auch von bereits umgesetzten Projekten, welche zwar nicht direkt im logistischen Bereich eines Warehouse Management Systems, aber im logistischen Transportprozess bei Diesel-Motoren von Transportschiffen zu finden sind. Hier wurde bereits ein Projekt umgesetzt, welches die Temperatur beziehungsweise auch die CO_2 Werte der Motoren misst, um frühzeitig etwaige Störungen aufzufinden und somit vorbeugend eine Wartung am Motor durchführen zu können.

Interviewpartner 3: Interviewpartner 3 hatte noch keinen direkten Berührungspunkt mit der Thermographie und der Tribologie, im Zuge eines Warehouse Management Systems. Durch seine Tätigkeiten für frühere Arbeitgeber hat dieser jedoch bereits Lösungen implementiert, bei welchen die Temperatur und die Tribologie bei Kränen eingesetzt wurde. Dies wurde aber nicht im Zusammenhang von CO_2 Ausstoß gemessen, sondern im Hinblick auf frühzeitiges Erkennen von Störungen einer Maschine.

Interviewpartner 4: Interviewpartner 4 hatte noch keine Berührungspunkte mit Predictive Maintenance und dem Messen von CO_2 Ausstoß beziehungsweise den veränderten Wärmezuständen einer Maschine oder eines Systems. Die Interviewpartnerin kann sich aber sehr gut vorstellen, dass, wenn eine Regression auf die Thermographie angewendet wird, hier ein Zusammenhang mit dem CO_2 Ausstoß aufgezeigt werden könnte. Laut Interviewpartnerin 4, wäre dies aber lediglich eine kurzfristige Optimierung für ein Warehouse Management System.

Evaluierung Hypothese 4
▪ Der Einsatz auf Motoren ist sinnvoll, um Rückschlüsse auf den CO_2 Ausstoß zu ziehen. ▪ Der CO_2 Ausstoß von Dieselmotoren kann verringert werden. ▪ Es gibt wenig Berührungspunkte von Predictive Maintenance im Zusammenhang mit Data Mining und Warehouse Management Systemen. ▪ Es stellt nur eine kurzfristige Optimierung dar.

Tabelle 3 Evaluierung Hypothese 4

Mittelfristige Optimierung

Evaluierung H3: Durch die Klassifikation können Antwortzeiten des Systems kategorisiert und somit Sicherheitslücken frühzeitig erkannt werden, um somit die Sicherheit zu optimieren.

Interviewpartner 1: Interviewpartner 1 erzählt, dass in seinem Arbeitsumfeld Antwortzeiten und Verfügbarkeitszeiten von verschiedenen Systemen gemessen werden. Des Weiteren werden Performance relevante Anlagenteile oder Softwaresysteme im Umfeld von Interviewpartner 1 alle fünf Minuten automatisiert gemessen. Dies soll gewährleisten, dass bereits frühzeitig erkannt werden kann, ob ein System einen Fehler aufweist oder zukünftig aufweisen wird. Im Zuge dieser Messungen werden etwaige auftretende Fehler aufgezeigt, welche danach kontrolliert werden müssen. Aus diesem Grund sieht Interviewpartner 1 eher den Einsatzzweck vom Messen der Antwortzeiten bei der Performanceauslastung, als bei der Sicherheit von Software und deren umliegenden Systemen.

Interviewpartner 2: Interviewpartner 2 ist deutlich davon überzeugt, dass mittels der Klassifizierung von Antwortzeiten, eines Systems, Anomalien beziehungsweise auch Sicherheitslücken erkannt sowie ein stabileres System geschaffen werden kann. Ein schwieriges Thema ist jedoch, wenn sich zum Beispiel eine Sicherheitslücke über einen Hacker im System aufgetan hat und dieser bereits aktiv Pakete abfängt und mitliest. Dieser Hacker ist dann leider nur ausfindig zu machen, wenn er sich auch anders verhält als das zuvor normal beobachtete Verhalten. Sollte sich vom Verhalten einer dritten Person, zum normalen Verhalten keine großen Abweichungen ergeben, dann ist es auch nur sehr schwer, hier eine Sicherheitslücke durch Automatismen ausfindig zu machen.

Interviewpartner 3: Der dritte Interviewpartner denkt, dass wenn die Antwortzeiten von Softwaresystemen und Maschinen überwacht werden, ein großes Potential besteht Sicherheitslücken frühzeitig zu erkennen. Aus seiner Sicht ist es jedoch schwierig hier Ausreiser von einem wirklichen Angriff zu unterscheiden. Dafür muss entsprechend geschultes Personal vorhanden sein, welches solche Meldungen überprüft. Denn für ein System entsteht kein Nutzen, wenn es eine Software gibt, welches etwas überprüft und Meldungen schreibt, aber diese Meldungen keine sicherheitsrelevanten Themen sind.

Interviewpartner 4: Interviewpartnerin 4 erzählt ebenfalls wieder aus ihrem letzten Projekt, bei welchem ein Predictive Maintenance Systems mittels Data Mining eingesetzt wurde. Bei diesem Projekt wurden teilweise Antwortzeiten von Teilprozessen in einem System gemessen und klassifiziert, um frühzeitig auf potenzielle Störungen aufmerksam gemacht zu werden. Mittels dieser Ausreiser konnte dann ein geschultes Personal sich den Prozess näher ansehen oder gegebenenfalls auch Netzwerkstörungen bereits erkennen, bevor ein Endbenutzer und eine Endbenutzerin der Software diese überhaupt gemeldet hatte. Interviewpartnerin 4 denkt auch, dass dieses System erweitert werden könnte, um potenzielle Sicherheitslücken in einem System frühzeitig zu erkennen. Ihrer Meinung nach sei es jedoch schwierig und auch sehr zeitaufwändig, mit dieser Methode Sicherheitslücken ausfindig zu machen. Hierfür würde es in der Praxis bessere Methoden wie zum Beispiel Honeypots geben.

Evaluierung Hypothese 3

- Eine systematisches Überwachen von Antwortzeiten eines Systems.
- Eher ein Thema von Performanceauslastung, als für eine Sicherheitsüberwachung eines Systems.
- Die Klassifizierung von Antwortzeiten ist sinnvoll. Aber es ist sehr schwierig, allein über diese eine Sicherheitslücke ausfindig zu machen.
- Der Unterschied zwischen gewünschtem Verhalten und einer Sicherheitslücke ist oft schwierig auszumachen.
- Die Klassifizierung von Antwortzeiten ist nicht die beste Methode, um Sicherheitslücken ausfindig zu machen.

Tabelle 4 Evaluierung Hypothese 3

Langfristige Optimierung

Evaluierung H2: Wenn die Regression auf den Durchsatz eines Systems über Tage, Monate und Wochen angewendet wird, kann eine Veränderung der Durchlaufzeiten und somit eine frühzeitige Störung oder ein Problem vorausgesagt werden. Mit Hilfe von präventiven Maßnahmen kann dies verhindert werden.

Interviewpartner 1: Interviewpartner 1 erzählt von Projekten, bei denen bereits jetzt Überwachungen implementiert sind, ob eine Ware innerhalb einer Logistikhalle eine längere Durchlaufzeit als gewöhnlich hat. Dies wird erreicht, indem sogenannte Kreisläufer auf einer Anlage identifiziert und einem Personal gemeldet werden. Kreisläufer sind zum Beispiel Kisten, welche auf einem Anlagenteil nicht ausgeschleust werden können, weil der Barcode der Kiste während des Transportes beschädigt worden ist und ein Sensor diesen nicht mehr lesen kann. Des Weiteren denkt Interviewpartner 1, dass bei ähnlichen Bedingungen der Packstücke, die Durchlaufzeit der Packstücke auch circa gleich sein sollte.

Wenn nun Packstücke eine längere Durchlaufzeit aufweisen, dann kann laut Interviewpartner 1, dies schon ein Indiz dafür sein, dass hier ein Anlagenteil nicht ordnungsgemäß funktioniert. Aus diesem Grund meint Interviewpartner 1, dass die Regression ein gutes Indiz dafür sein kann, dass ein Fehler im System vorhanden sein könnte.

Interviewpartner 2: Aufgrund der historischen und der aktuellen Daten denkt Interviewpartner 2, dass ein Vergleich zwischen diesen aufgestellt werden kann. Somit können ähnliche Daten, welche in der Vergangenheit eine gewisse Zeit gebraucht haben, in der Zukunft optimiert werden, indem zum Beispiel weiteres Personal hinzugezogen wird, um einen höheren Durchsatz und dadurch eine geringere Abarbeitungszeit der bestellten Aufträge gewährleisten zu können.

Dies kann aber nicht nur durch das Hinzuziehen von weiterem Personal erreicht werden. Es könnte auch sein, dass gewisses Personal, bei gewissen Bestellungen, eine bessere

Durchlaufzeit dieser gewährleisten als anderes Personal. Somit kann zum Beispiel durch den gezielten Einsatz von Personal bereits eine schnellere Abarbeitung der Aufträge und Bestellungen gewährleistet werden. Das kann unter anderem durch den Einsatz von Analysen der historischen Daten aufgezeigt werden, um folglich an zukünftigen Daten eine Änderung durchzuführen.

Interviewpartner 3: Interviewpartner 3 denkt, wenn eine Standardisierung der Daten vorhanden ist und Daten miteinander verglichen werden, sehr gut die Durchlaufzeit von Aufträgen verringert werden kann, indem Störungen frühzeitig erkannt werden. Eine längere Durchlaufzeit bei vergleichbaren Bestellungen kann zum Beispiel darauf hinweisen, dass die darunterliegenden Systeme eine Störung haben oder eine Störung bevorstehen wird. Diese Daten können aber leider nicht allein herangezogen werden, um eine Störung vorauszusagen. Dafür bedarf es weitere Faktoren, um diese genauer identifizieren zu können.

Interviewpartner 4: Wie bereits bei der Hypothese 4 erzählt Interviewpartnerin 4, auch bei dieser Fragestellung, vom letzten Projekt, wo nicht nur Antwortzeiten von Systemen und Prozessen gemessen worden sind, sondern auch Regressionen auf den Durchsatz von Daten angewendet wurden. Mit Hilfe der vergangenen Daten konnten Modelle aufgestellt werden, welche die Durchlaufzeiten für zukünftige Bestellungen voraussagen konnten. Wurden diese Durchlaufzeiten im Laufe eines Tages überschritten, wurden immer wieder Nachforschungen angestellt, warum es zu längeren Zeiten kam. An dieser Stelle erklärte Interviewpartnerin 4, dass es im operativen Betrieb zum Beispiel einen Benutzer der Software gab, welcher Probleme mit seinem Stapler hatte, dies aber nicht meldete. Durch die längeren Durchlaufzeiten dieses Benutzers, wurden bestimmte Nachforschungen getätigt und der Staplerfahrer wurde auf seine bereits defekte Maschine hingewiesen, ohne dass dieser einen Defekt bemerkte. Somit sieht Interviewpartnerin 4 hier ein sehr großes Potential und eine langfristige Optimierung beim Einsatz von Regression sowie von vergangenen Bestellungen für Predictive Maintenance.

Evaluierung Hypothese 2

- Die Überwachung der Durchlaufzeiten wird in der Praxis bereits eingesetzt.
- Die Durchlaufzeiten können ein gutes Indiz sein, ob eine Störung vorhanden ist oder auftreten wird.
- Mittels historischer Daten können zukünftige Zeiten, bei ähnlichen Bedingungen, errechnet werden.
- Voraussetzung sind gute historische Daten, welche mit den zukünftigen verglichen werden können.
- Eine langfristige Optimierung ist durch den Einsatz von Data Mining und von historischen Daten möglich.

Tabelle 5 Evaluierung Hypothese 2

9.5 Zusammenfassung

In diesem Kapitel wurde die Durchführung und die Evaluierung der Interviews mit den Experten und Expertinnen der Logistikbranche dargestellt. Im ersten Schritt wurden die Interviewpartner und Interviewpartnerinnen definiert sowie deren Tätigkeitsbereich näher erläutert, um deren Erfahrungswissen im Logistikbereich sowie im Data Mining und Predictive Maintenance aufzuzeigen. In weiteren Schritten wurde eine Evaluierung der Hypothesen durchgeführt.

Im nachfolgenden Kapitel werden die zuvor aus der Theorie aufgestellten Hypothesen aufgelistet, um anhand dieser die Ergebnisse aus der Literatur aufzuzeigen. Im zweiten Schritt werden die Ergebnisse der Interviews mit den Experten und Expertinnen, aus der Logistikbranche, erläutert. Hierfür wird pro Hypothese eine Auswertung durchgeführt. Im dritten und letzten Schritt des nachfolgenden Kapitels wird ein Fazit aus den erarbeiteten Ergebnissen gezogen und beschrieben, womit auch ein abschließender Ausblick über den Nutzen der vorliegenden Arbeit für weitere Forschungen sowie für die Wirtschaft unterstrichen wird.

10 ERGEBNISSE

In diesem Kapitel werden die gewonnenen Ergebnisse dieser Arbeit zusammengefasst dargestellt. Zunächst werden die theoretischen Ergebnisse, anhand der mit Literatur ausgearbeiteten Hypothesen, wiedergegeben. Daraufhin erfolgt eine Auswertung und Zusammenfassung der Interviews mit den Experten und Expertinnen aus der Logistikbranche. Anschließend werden die Ergebnisse, aus der Theorie und Empirie, gegenübergestellt, um einen Vergleich zu ziehen. Infolgedessen soll ein Ausblick über den Nutzen dieser Arbeit für weitere Forschung und für die Wirtschaft aufgezeigt werden.

10.1 Ergebnisse aus der Literatur

Die Ergebnisse aus der Literatur werden durch die nachfolgenden Hypothesen aufgezeigt:

Alternativhypothese (H1): Durch den Einsatz von Data Mining in der Logistikbranche können Störungen in einem Warehouse Management System mittels Predictive Maintenance früher erkannt werden.

Nullhypothese (H0): Durch den Einsatz von Data Mining in der Logistikbranche können Störungen in einem Warehouse Management System mittels Predictive Maintenance nicht früher erkannt werden

Hypothese 2 (H2): Wenn die Regression auf den Durchsatz eines Systems über Tage, Monate und Wochen angewendet wird, kann eine Veränderung der Durchlaufzeiten und somit eine frühzeitige Störung oder ein Problem vorausgesagt werden. Mit Hilfe von präventiven Maßnahmen kann dies verhindert werden.

Hypothese 3 (H3): Durch die Klassifikation können Antwortzeiten des Systems kategorisiert und somit Sicherheitslücken frühzeitig erkannt werden und Sicherheit optimiert werden.

Hypothese 4 (H4): Wenn die Regression bei der Thermographie und Tribologie angewendet wird, kann vorausgesagt werden, dass sich der CO_2 Ausstoß bei einer Erhöhung der thermischen Werte auch steigern wird. Durch den Einsatz von vorbeugenden Maßnahmen kann dies unterbunden werden.

Die Basis der angeführten Hypothesen wurde mittels Theorieteil erarbeitet. Weiterführend wurden diese Hypothesen für die Erstellung des Interviewleitfadens sowie für die Durchführung der Interviews mit den Experten und Expertinnen aus der Logistikbranche eingesetzt. Anhand dieser Vorgehensweise soll eine Überprüfung der Hypothesen erfolgen.

10.2 Ergebnisse aus den Interviews

In diesem Abschnitt werden die zuvor evaluierten Interviews mit den Experten und Expertinnen aus dem Logistikbereich anhand der jeweiligen Hypothese aufgezeigt.

10.2.1 Ergebnis der Alternativ- und Nullhypothese aus den Interviews

Die Interviewpartner und Interviewpartnerinnen sehen durch den Einsatz von Predictive Maintenance und Data Mining in einem Warehouse Management System ein sehr großes Potential. Durch den Einsatz von kostengünstigen IOT Geräten kann eine langfristige Optimierung eines Systems erfolgen, mit welcher eine hohe Kostenersparnis bewirkt werden kann.

Der Einsatz von Data Mining und Predictive Maintenance in einem Warehouse Management System muss jedoch als ein langfristiges Projekt angesehen werden, welches ständig adaptiert und an die Umgebungsvariablen eines Systems angepasst werden muss. Die Interviewpartner und Interviewpartnerinnen sehen aus diesem Grund keine einmalige Installation eines solchen Systems, welches sofort umgesetzt werden kann. Alle Interviewpartner und Interviewpartnerinnen betonen allerdings auch, dass durch Predictive Maintenance ein Fehler frühzeitig aufgezeigt werden kann, welcher in einem System noch nicht erkannt worden ist, und somit eine frühzeitiges Beheben dieses Fehlers möglich wird.

10.2.2 Ergebnis aus der Hypothese 2

Die Interviewpartner und Interviewpartnerinnen setzen teilweise die Überwachung der Durchlaufzeiten bereits in der Praxis ein. Mittels der Durchlaufzeiten kann ein System dahingehend überwacht werden, ob eine Störung am System vorhanden ist und/oder, ob diese wieder auftreten wird. Demzufolge ist die Überwachung der Durchlaufzeiten ein gutes Indiz für die Überwachung eines Systems.

Zudem kann durch historische Daten ein gutes Modell, für die Berechnung von zukünftigen Daten, erstellt werden. Dieses Modell ist jedoch nur dann aussagekräftig, wenn dem Modell ähnliche Bedingungen zugrunde liegen. Folglich ist auch eine der wichtigsten Voraussetzungen, dass die historischen Daten mit den zukünftigen Daten verglichen werden können. Mittels des Einsatzes von Data Mining, historischen Daten und Predictive Maintenance sehen die Interviewpartner und Interviewpartnerinnen eine langfristige Optimierung für ein System und/oder für ein Warehouse Management System.

10.2.3 Ergebnis aus der Hypothese 3

Die Interviewpartner und Interviewpartnerinnen sehen das systematische Überwachen von Antwortzeiten als eine gute Vorgehensweise, um Auffälligkeiten in einem System aufzuspüren. Jedoch wird dies eher im Hinblick auf die Überwachung von Performanceauslastungen als zur Sicherheitsüberwachung eines Systems genutzt. Die Klassifizierung von Antwortzeiten, für diesen Aufgabenbereich ist sinnvoll. Allerdings ist es auch schwierig, allein mittels dieser Daten eine Sicherheitslücke ausfindig zu machen.

Zudem zeigt sich die Komplexität zwischen einem gewünschten Verhalten und einer Sicherheitslücke zu unterscheiden, ohne weitere Metriken zu bestimmen. Aus diesem Grund ist die Klassifizierung von Antwortzeiten nicht die beste Methode, um Sicherheitslücken ausfindig zu machen. Hierfür würde es in der Praxis bessere Methoden und Vorgehen geben.

10.2.4 Ergebnis aus der Hypothese 4

Die Interviewpartner und Interviewpartnerinnen heben den Einsatz von Regression und Predictive Maintenance beim Einsatzgebiet von Motoren, Wärme und CO_2 Ausstoß als sinnvoll hervor. Vor allem kann durch den Einsatz von kostengünstigen IOT Geräten, der CO_2 Ausstoß bei großen Dieselmotoren verringert werden. Die Interviewpartner und Interviewpartnerinnen hatten bis dato allerdings kaum Berührungspunkte in ihren vergangenen Projekten, im Zusammenhang von Warehouse Management Systemen, Predictive Maintenance und Data Mining. Aus diesem Grund sehen die Interviewpartner und Interviewpartnerinnen dies eher als ein kurzfristiges Optimierungspotential.

10.3 Fazit und Zusammenfassung

In diesem Kapitel wurden die Ergebnisse aus der Theorie und aus den Interviews mit den Experten und Expertinnen aus der Logistikbranche und dementsprechend eine Evaluierung der aufgestellten Hypothesen durchgeführt. Die Evaluierung der Ergebnisse ergibt somit folgendes Ergebnis:

Hypothese	Ergebnis
Hypothese 1	Zutreffend
Hypothese 2	Zutreffend
Hypothese 3	Teilweise zutreffend
Hypothese 4	Teilweise zutreffend

Tabelle 6 Darstellung der Ergebnisse

- **H1:** Die Hypothese 1 ist zutreffend. Die Experten und Expertinnen der Logistikbranche sehen ein Optimierungspotential durch den Einsatz von Data Mining in der Logistikbranche, womit Störungen frühzeitig erkannt werden können. Mittels eines langfristig angelegten Projektes können somit Kosten, in einem Warehouse Management System, eingespart werden.
- **H2:** Die Hypothese 2 ist zutreffend. Experten und Expertinnen setzen diese Vorgehensweisen teilweise bereits in Systemen ein und ermitteln mittels der Regression eine Veränderung der Durchlaufzeiten. Dadurch können frühzeitige Störungen und Probleme vorhergesagt werden.
- **H3:** Die Hypothese 3 ist teilweise zutreffend. Das Optimierungspotential wird zwar von den Experten und Expertinnen erkannt, jedoch als nicht als effizient genug angesehen. Hierfür würde es bessere Methoden geben, um das gewünschte Ergebnis zu erzielen.
- **H4:** Die Hypothese 4 ist teilweise zutreffend. Die Expertinnen und Experten können sich zwar vorstellen, dass mittels der Thermographie und Tribologie der CO_2 Ausstoß erkannt werden kann. Bis dato wird das, in realen Projekten, allerdings kaum umgesetzt.

Die Logistikbranche bleibt nicht von der Nutzung von neuen Technologien und Methoden der Digitalisierung verschont. Durch Veränderungen des Kaufverhaltens, bewirkt durch Pandemien, Globalisierung und kürzeren Produktzyklen muss sich auch die Logistikbranche ständig wandeln. Dabei muss sie sich stets an den aktuellen Umständen ihrer Kundinnen und Kunden anpassen und diesen neue Dienstleistungen zur Verfügung stellen. Durch diesen hohen Druck sowie durch die gesteigerten Anforderungen von Kundinnen und Kunden können Techniken wie Predictive Maintenance und Data Mining, die Prozesse effizienter gestalten und Verbesserungsmöglichkeiten aufzeigen.

In der vorliegenden Arbeit wurden Optimierungspotentiale sowie Einflussfaktoren von Data Mining auf Predictive Maintenance für ein Warehouse Management System in der Logistikbranche evaluiert. Im Weiteren wurden die gewonnen Erkenntnisse aus dem Theorieteil sowie die ausgearbeiteten Ergebnisse aus den Interviews mit Expertinnen und Experten gegenübergestellt und anschließend verglichen.

Die nachfolgend angeführten Optimierungen werden von den Expertinnen und Experten als Einflussfaktoren von Data Mining auf Predictive Maintenance für ein Warehouse Management System in der Logistikbranche angesehen:

- Mittels IOT Geräten, welche an Maschinen und Fördertechnik angebracht werden, können frühzeitig Störungen mittels Data Mining Techniken erkannt und aufgezeigt werden.
- Durch den Einsatz von Regression auf den Durchsatz eines Systems anhand historischer Daten, kann eine Veränderung der Durchlaufzeiten erkannt und somit frühzeitig eine Störung oder ein Problem vorausgesagt werden.

10.4 Ausblick

In weiteren Forschungen und Arbeiten können die evaluierten Ergebnisse, welche aus den Interviews mit den Expertinnen und Experten der Logistikbranche gewonnen wurden sowie die zuvor aufgestellten Hypothesen, in praktischen Projekten gemessen und im Detail überprüft werden. Durch den Einsatz der genannten Data Mining Techniken sowie der erwähnten Einsatzgebiete von Predictive Maintenance und Datenquellen der Logistikbranche, kann in nachfolgenden Projekten überprüft werden, inwiefern sich frühzeitige Meldungen einer Wartung durch Predictive Maintenance Techniken auf ein Warehouse Management System auswirken und, ob sich in weiterer Folge die Wartungskosten von Maschinen und Systemen dadurch senken lassen.

ANHANG A - Interviewleitfaden für Experteninterviews aus dem Logistikbereich

Dieses Dokument ist eine Einleitung und Begriffserläuterung für das nachfolgende Interview und wird für die Masterarbeit „Einflussfaktoren des Data Minings auf Predictive Maintenance für ein Warehouse Management System in der Logistikbranche“ im Rahmen des Studiums Informationstechnologien & Wirtschaftsinformatik an der Fachhochschule Campus02 Graz durchgeführt. Im Anschluss an die Einleitung findet sich der Interviewleitfaden für die Interviews mit den Experten aus der Logistikbranche. Dieser Interviewleitfaden dient als Grundlage für die Evaluierung der Hypothesen.

In der angeführten Arbeit soll evaluiert werden, welche Einflussfaktoren Data Mining in der Logistikbranche auf Predictive Maintenance in einem Warehouse Management System hat. Die Evaluierung erfolgt anhand Interviews mit Experten aus der Logistikbranche, weswegen die Hypothesen aus der Arbeit überprüft werden. In der Arbeit wird folgende Forschungsfrage beantwortet:

„Welche Einflussfaktoren hat Data Mining auf Predictive Maintenance für ein Warehouse Management System in der Logistikbranche?“

In der Arbeit wurde beschrieben, welche Datenquellen, welche Data Mining Verfahren, welche Anforderungen die Logistik, und welche Einsatzgebiete es für Predictive Maintenance gibt. Die Vorhersage verwendet Variablen oder Felder in einer Datenbank, um unbekannte oder zukünftige Werte anderer Variablen vorherzusagen. Die Beschreibung konzentriert sich auf das Finden von interpretierbaren Mustern, die die Daten beschreiben.

Für die Vorhersage und Beschreibung gibt es folgende Data Mining Verfahren:

- **Klassifikationsverfahren**

  Beim Verfahren der Klassifikation sollen bereits bekannte Klassen-Zuordnungen aus Trainingsdaten, ein neues Klassifizierungsmodell erstellen beziehungsweise ein Modell erlernt werden. Dieses aufgestellte Modell wird in einem zweiten Schritt auf neue Objekte und Daten ohne zugehörige Klassenzuordnungen angewendet. Ein Einsatzzweck der Klassifikation ist für Banken „Welcher Kreditnehmer und Kreditnehmerin, werden einen Kredit zurückzahlen“

- **Regressionsverfahren**

  Das Regressionsverfahren stellt ein Vorhersagemodell dar. Der Zusammenhang zwischen einer Abhängigen und einer Unabhängigen Variable wird untersucht. Mit Hilfe der Regression können Fragen beantwortet werden, wie sich die Absatzmenge eines Produktes ändert, wenn der Preis eins Produkts geändert wird.

- **Assoziationsverfahren**

  Im Assoziationsverfahren werden Algorithmen angewendet, welche in den bestehenden Daten Verbindungen aufzeigen sollen. Einer der Haupteinsatzzwecke dieses Verfahrens

sind Warenkorbanalysen wie zum Beispiel „Wenn jemand Chips kauft, dann kauf diese Person wahrscheinlich auch Bier“

- **Clusterverfahren**

  Bei diesem Verfahren werden Gruppen, sogenannte Cluster, gebildet, welche im Vergleich zu anderen Gruppen möglichst homogen sein sollen. Eine Gruppierung der Kunden und Kundinnen nach demographischen Kriterien könnte ein Anwendungszweck des Clusterverfahrens sein.

Folgende Predictive Maintenance Einsatzgebiete können in Soft- und Hardware verwendet werden:

Hardware

- **Thermographie**: Bei der Thermographie wird die Temperatur von Anlagen und Maschinen gemessen. Dies können gesamte Hallen oder auch nur Teilbereiche von Systemen sein. Es werden Infrarot Thermometer, Infrarot Bilder und Line Scanner als Hardware eingesetzt.
- **Vibrationsüberwachung**: Es werden die Vibrationen von Anlagen überwacht und somit kann die aktuelle Bedingung von Systemen gut analysiert werden. Es werden bekannte Werte mit Ausreißern gegenübergestellt, um Auffälligkeiten frühzeitig zu erkennen.
- **Tribologie**: Die Tribologie steht in der Industrie für die Wissenschaft der Reibung, Verschleiß und Schmierung von gegeneinander bewegten Bauteilen. Dies umfasst die Analyse von Öl sowie deren Partikel im Öl.
- **Ultraschall und Schallemission:** Mithilfe der Ultraschallmethode kann die Ultraschalldichte gemessen werden, wodurch ein korrosiver auffallender Verschleiß prognostiziert werden kann.

Software

- **Antwortzeit**: Die Definition der Antwortzeit ist das Intervall zwischen dem Absenden einer Anfrage an ein System und dem Erhalt einer Antwort nach der Bearbeitung des Systems.
- **Durchsatz**: Der Durchsatz ist definiert durch den Quotienten aus erfolgreich bearbeiteten Anfragen pro Zeiteinheit. Dieser stellt aus Sicht der Benutzer und Benutzerinnen oftmals eher einen sekundären Aspekt der Performance dar.
- **Skalierbarkeit**: Ein gut skalierbares System besitzt die Fähigkeit die Zielvorgaben für Antwortzeiten und Durchsatz auch bei steigenden Anforderungen durch Softwarefunktionen einzuhalten. Ein System dessen Antwortzeit bei linearem Anstieg der Anforderung, wie zum Beispiel den Anstieg an Benutzer und Benutzerinnen bei einer Server-Anwendung, exponentiell wachsen, gelten als nicht gut skalierende Systeme und sind in der Regel auch nicht performant

Einsatzgebiete in der Logistik:

Zu den oben genannten Einsatzgebieten von Predictive Maintenance können folgende Datenquellen in der Logistikbranche zugrundeliegen:

- Von Maschinen generierte Daten der Tribologie mittels Sensoren auf der Maschine
- Von Maschinen generierte Daten der Thermographie mittels eingebundener Sensoren auf der Maschine
- Von Maschinen generierte Daten der Vibrationsüberwachung mittels eingebundener Sensoren auf der Maschine
- Von Maschinen generierte Daten von Ultraschall und Schallemissionen mittels eingebundener Sensoren auf der Maschine
- Von Systemen generierten Daten der Antwortzeiten mittels Messungen in der Software zwischen Sendezeitpunkt und Empfangszeitpunkt.
- Von Systemen generierte Daten des Durchsatzes mittels Messungen in der Software zwischen Auftragsannahme und Auftragsabschluss

Interviewleitfaden

1. Kurzfristige Optimierung

1.1. **H4:** Wenn die Regression bei der Thermographie und Tribologie angewendet wird, kann vorausgesagt werden, dass sich der CO_2 Ausstoß bei einer Erhöhung der thermischen Werte auch steigern wird. Durch den Einsatz von vorbeugenden Maßnahmen kann dies unterbunden werden.

Datenquellen: Von Maschinen generierte Daten wie Temperatur und Vibrationen

Data Mining Verfahren: Assoziation

Einsatzgebiet: Sensoren an Maschinen welche Vibrationen und Wärme messen

Einflussfaktoren: Minimierung von CO_2 Ausstoß

1.1.1. Wie können Ihrer Meinung nach Ihrer Meinung nach Daten wie Temperatur und Vibration den CO2 Ausstoß verringern?

1.1.2. Wie würden Sie die Auswirkung auf den CO_2 Ausstoß durch die Nutzung von Temperatur und Vibrationssensoren mittels Assoziation einschätzen?

1.1.3. Denken Sie, dass sich durch die Analyse von Temperatur und Vibrationsdaten mittels Assoziation der Daten vorbeugende Maßnahmen und der Ausstoß von CO_2 optimieren lassen?

2. Mittelfristige Optimierung

2.1. **H3:** Durch die Klassifikation können Antwortzeiten des Systems kategorisiert und somit Sicherheitslücken frühzeitig erkannt werden und Sicherheit optimiert werden.

Datenquellen: Von Maschinen und Systemen generierte Daten wie Antwortzeiten

Data Mining Verfahren: Klassifikation

Einsatzgebiet: Historische Antwortzeiten des Systems oder vom Benutzer gesendete Antwortzeiten einer Eingabe am Interface

Einflussfaktoren: Optimierung von Sicherheit

2.1.1. Welche Möglichkeiten gibt es, vom System generierte Daten wie Antwortzeiten, zu kategorisieren und somit Sicherheitslücken aufzufinden?

2.1.2. Welche Auswirkung auf die Sicherheit kann die Klassifikation von Antwortzeiten eines Systems in Hinblick auf die Optimierung mittels Predictive Maintenance haben?

2.1.3. Denken Sie, dass durch die Klassifikation von Antwortzeiten eines Systems und den Einsatz von Predictive Maintenance Techniken die Sicherheit eines Systems optimiert werden kann?

3. Langfristige Optimierung

3.1. **H1:** Durch den Einsatz von Data Mining in der Logistikbranche können Störungen in einem Warehouse Management System mittels Predictive Maintenance früher erkannt werden

Datenquellen: Von Systemen generierte Daten wie Auftragsannahme, Von Maschinen generierte Daten wie Temperaturen

Data Mining Verfahren: Regression

Einsatzgebiet: Sensoren an Maschinen, Historie an vergangenen Bestellungen

Einflussfaktoren: Minimierung von Durchlaufzeit

3.1.1. Wie können Ihrer Meinung nach, Sensoren für die Temperaturmessung an Maschinen genutzt werden, um Störungen zu minimieren und somit Durchlaufzeiten zu minimieren?

3.1.2. Welche Auswirkung auf die Durchlaufzeit einer Bestellung sehen sie beim Einsatz von Regression und Sensoren an Maschinen bei den zuvor genannten Datenquellen?

3.1.3. Denken Sie, dass durch Temperaturmessungen an Maschinen in Kombination mit Regression eine Minimierung der Durchlaufzeit erreicht werden kann indem Stillstände vermieden werden und der operative Betrieb verbessert wird?

3.2. **H2:** Wenn die Regression auf den Durchsatz eines Systems über Tage, Monate und Wochen angewendet wird, kann eine Veränderung der Durchlaufzeiten und somit eine frühzeitige Störung oder ein Problem vorausgesagt werden. Mit Hilfe von präventiven Maßnahmen kann dies verhindert werden.

Datenquellen: Von Systemen generierte Daten wie historische Daten und Zeitstempel

Data Mining Verfahren: Regression

Einsatzgebiet: Historische Auftragsdaten

Einflussfaktoren: Minimierung von Durchlaufzeiten

3.2.1. Wie können Ihrer Meinung nach, historische Auftragsdaten, die Durchlaufzeiten von zukünftigen Aufträgen mittels Predictive Maintenance sinken?

3.2.2. Welche weiteren Möglichkeiten bestehen, durch den Einsatz von Regression und die Analyse von vergangenen Durchlaufzeiten von Auftragsdaten, Ausreißer zu erkennen und somit Fehler im System frühzeitig zu erkennen, um zukünftige Durchlaufzeiten zu sinken.

3.2.3. Denken Sie, dass durch die zuvor genannten Datenquellen und die Regression verknüpft mit Predictive Maintenance Techniken die Durchlaufzeiten von zukünftigen Auslieferungen optimiert werden kann?

ABKÜRZUNGSVERZEICHNIS

SQL	Structured Query Language
KDD	Knowledge Discovery in Databases
CRISP-DM	Cross-Industry Standard Process for Data Mining
ERP	Enterprise Ressource Planning
KPI	Key Performance Indicator
IOT	Internet of Things

ABBILDUNGSVERZEICHNIS

Abbildung 1 Aufbau der Arbeit 7
Abbildung 2 Prozess des Knowledge Discovery in Databases (Fayyad et al. 1996) 9
Abbildung 3 CRISP-DM (Chapman et al. 1999, S. 12) 11
Abbildung 4 Stufen der Wartung (Selcuk 2017) 14
Abbildung 5 Beispiel eines Entscheidungsbaums 21
Abbildung 6 Beispiel Nearest Neighbour (Müller und Lenz 2013, S. 106) 22
Abbildung 7 Beispiel Clusterverfahren 25
Abbildung 8 Aspekte der Performance 33

TABELLENVERZEICHNIS

Tabelle 1 Kennzahlen der Logistik 18
Tabelle 2 Evaluierung der Alternativ- und Nullhypothese 58
Tabelle 3 Evaluierung Hypothese 4 59
Tabelle 4 Evaluierung Hypothese 3 61
Tabelle 5 Evaluierung Hypothese 2 62
Tabelle 6 Darstellung der Ergebnisse 66

LITERATURVERZEICHNIS

Alpar, P. & Niedereichholz, J. (2000). *Data Mining im praktischen Einsatz. Verfahren und Anwendungsfälle für Marketing, Vertrieb, Controlling und Kundenunterstützung* (Business Computing). Wiesbaden: Vieweg+Teubner Verlag.

Arnold, D., Isermann, H., Kuhn, A., Tempelmeier, H. & Furmans, K. (Hrsg.). (2008). *Handbuch Logistik* (VDI-Buch, 3., neu bearb. Aufl.). Berlin: Springer.

Bankhofer, U. & Vogel, J. (2008). *Datenanalyse und Statistik. Eine Einführung für Ökonomen im Bachelor ; [Bachelor geeignet* (Lehrbuch, 1. Aufl.). Wiesbaden: Gabler.

Bauernhansl, T., Hompel, M. ten & Vogel-Heuser, B. (2014). *Industrie 4.0 in Produktion, Automatisierung und Logistik. Anwendung, Technologien, Migration* (SpringerLink). Wiesbaden: Springer Vieweg.

Baur, N. & Blasius, J. (Hrsg.). (2014). *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden: Springer VS.

Benz, J. & Höflinger, M. (2011). *Logistikprozesse mit SAP®. Eine anwendungsbezogene Einführung - Mit durchgehendem Fallbeispiel - Geeignet für SAP Version 4.6A bis ECC 6.0* (3., aktualisierte Auflage). Wiesbaden: Vieweg+Teubner Verlag / Springer Fachmedien Wiesbaden GmbH Wiesbaden.

Bousonville, T. (2017). *Logistik 4.0. Die digitale Transformation der Wertschöpfungskette* (Essentials). Wiesbaden: Springer Gabler.

Bowen, J. P. & Hinchey, M. G. (1999). *High-integrity system specification and design* (Formal approaches to computing and information technology). London: Springer.

Bramer, M. A. (2007). *Principles of data mining* (Undergraduate topics in computer science). London: Springer.

Cai, J., Liu, X., Xiao, Z. & Liu, J. (2009). Improving supply chain performance management: A systematic approach to analyzing iterative KPI accomplishment. *Decision Support Systems 46* (2), 512–521. Decision Support Systems, 46(2), 512-521. doi:10.1016/J.DSS.2008.09.004

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C. & Wirth, R. (1999). CRISP-DM 1.0. Step-by-step data mining guide. https://www.the-modeling-agency.com/crisp-dm.pdf.

Dapp, T. (2014). Frankfurt am Main: Deutsche Bank Research.

Data Center Knowledge. (2017). Google Data Center FAQ. Everything you ever wanted to know (and didn't) about Google data centers but were afraid to ask. https:// www.datacenterknowledge.com/archives/2017/03/16/google-data-center-faq. Zugegriffen: 19. April 2020.

Fayyad, U., Piatetsky-Shpiro, G. & Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *AI Magazine 17* (3), 37–54.

Gantz, J. & Reinsel, D. (IDC IVIEW, Hrsg.). (2012). THE DIGITAL UNIVERSE IN 2020. Big Data, Bigger Digital Shadow s, and Biggest Grow th in the Far East. https://www.emc.com/ collateral/analyst-reports/idc-the-digital-universe-in-2020.pdf.

Girdhar, P. (2004). *Practical machinery vibration analysis and predictive maintenance* (Practical professional books from Elsevier). Oxford: Newnes.

Grubb, P. & Takang, A. A. (2003). *Software maintenance. Concepts and practice* (2nd ed.). River Edge, N.J: World Scientific.

Gudehus, T. (2011). *Logistik. Grundlagen - Strategien - Anwendungen* (4., aktualisierte Aufl 2010). Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.

Han, J. & Kamber, M. (2006). *Data mining. Concepts and techniques* (The Morgan Kaufmann series in data management systems, 2nd ed.). Amsterdam: Elsevier.

Hompel, M. ten & Kerner, S. (2015). Logistik 4.0. *Informatik-Spektrum 38* (3), 176–182. doi:10.1007/s00287-015-0876-y

Hompel, M. ten & Schmidt, T. (2010). *Warehouse Management. Organisation und Steuerung von Lager- und Kommissioniersystemen* (VDI-Buch, 4. Aufl.). s.l.: Springer-Verlag.

Huber, D. & Kaiser, T. (Hrsg.). (2017). *Industrie 4.0. Herausforderungen, Konzepte und Praxisbeispiele* (Edition HMD). Wiesbaden: Springer Vieweg.

Janetzko, D. & Steinhöfel, K. (1997). Lotsen los! Data Mining: Verborgene Zusammenhänge in Datenbanken aufspüren. *c't* (3), 294–300.

Jeske, M., Grüner, M. & Weiß, F. (2014). *White Paper. Big Data in der Logistik*. Frankfurt am Main: T-Systems International GmbH.

Kronthaler, F. (2014). *Statistik angewandt. Datenanalyse ist (k)eine Kunst* (Springer-Lehrbuch). Berlin: Springer Spektrum.

Lehmacher, W. (2015). Wirtschaft, Gesellschaft und Logistik 2050. In P. H. Voß (Hrsg.), *Logistik - eine Industrie, die (sich) bewegt. Strategien und Lösungen entlang der Supply Chain 4.0* (1. Auflage, S. 1–17). Wiesbaden: Springer Gabler.

Lin, T. Y. (2002). Attribute transformations for data mining I: Theoretical explorations. *International Journal of Intelligent Systems 17* (2). doi:10.1002/int.10017

Maimon, O. Z. & Rokach, L. (2010). *Data mining and knowledge discovery handbook* (2. ed.). New York: Springer.

Mayring, P. (2015). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* (Beltz Pädagogik, 12., überarb. Aufl.). Weinheim: Beltz.

Mehta, M., Agrawal, R. & Rissanen, J. (1996). SLIQ: A fast scalable classifier for data mining. In P. Apers (Hrsg.), *Advances in database technology - EDBT '96* (Lecture Notes in Computer Science, Bd. 1057, Bd. 1057, S. 18–32). Berlin [etc.]: Springer.

Mobley, R. K. (2002). *An introduction to predictive maintenance* (Second edition).

Müller, R. M. & Lenz, H.-J. (2013). *Business Intelligence* (eXamen.press). Berlin: Springer Vieweg.

Nakhaeizadeh, G. (1998). *Data Mining. Theoretische Aspekte und Anwendungen* (Beiträge zur Wirtschaftsinformatik, Bd. 27). Heidelberg: Physica-Verlag HD.

Razum, K. K. (Hrsg.). (2015). *Der Duden. In zwölf Bänden; das Standardwerk zur deutschen Sprache* (11., vollst. überarb. und aktualisierte Aufl.). Berlin: Dudenverl.

Reinsel, D., Gantz, J. & Rydning, J. (IDC, Hrsg.). (2017). The Evolution of Data to Life-Critical. Don’t Focus on Big Data; Focus on the Data That’s Big. https://www.seagate.com/www-content/our-story/trends/files/Seagate-WP-DataAge2025-March-2017.pdf. Zugegriffen: 1. Mai 2020.

Schlatt, A. (2014). *Predictive Maintenance. Maschinenfehler und -ausfälle im Vorfeld erkennen*. München: Genios Verl.

Selcuk, S. (2017). Predictive maintenance, its implementation and latest trends. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture 231* (9), 1670–1679. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 231(9), 1670-1679. doi:10.1177/0954405415601640

Sharafi, A. (2013). *Knowledge Discovery in Databases. Eine Analyse des Änderungsmanagements in der Produktentwicklung* (Springer Gabler Research). Wiesbaden: Springer Gabler (Zugl.: München, Techn. Univ., Diss., 2012).

Smith, C. U. (2007). Introduction to Software Performance Engineering: Origins and Outstanding Problems. In M. Bernardo (Hrsg.), *Formal methods for performance evaluation. Advanced lectures* (Lecture Notes in Computer Science, S. 395–428). Berlin, Heidelberg: Springer Berlin Heidelberg.

Smith, C. U. & Williams, L. G. (2002). *Performance solutions. A practical guide to creating responsive, scalable software* (The Addison-Wesley object technology series). Boston, Mass. Addison-Wesley.

Spehling, M. (2007). *Analyse und Erweiterung von Methoden des Data Mining in räumlichen Datenbanken*. Hannover: Leibnitz Universität Hannover.

Tufféry, S. (2011). *Data mining and statistics for decision making* (Wiley series in computational statistics). Chichester: Wiley.

Vogel-Heuser, B., Bauernhansl, T. & Hompel, M. ten (Hrsg.). (2017). *Handbuch Industrie 4.0. Bd. 3: Logistik* (Springer Reference Technik, 2., erweiterte und bearbeitete Auflage). Berlin: Springer Vieweg.

Vogt, S. & Werner, M. (2014). *Forschen mit Leitfadeninterviews und qualitativer Inhaltsanalyse*. Köln: FACHHOCHSCHULE KÖLN.

Warehouse Logistics (http://www.warehouse-logistics.com/, Hrsg.). (2020). Definition WMS / LVS. Zugegriffen: 18. September 2020.

Zhong, N. & Zhou, L. (Hrsg.). (1999). *Methodologies for knowledge discovery and data mining. Third Pacific-Asia Conference, PAKDD-99, Beijing, China, April 26 - 28, 1999 ; proceedings* (Lecture notes in computer science Lecture notes in artificial intelligence, Bd. 1574). Berlin: Springer.