

Madlin Weise

Möglichkeiten und Grenzen BLE-basierter Signal­daten für eine
automatisierte Klassifizierung logistischer Prozesse

Madlin Weise

Möglichkeiten und Grenzen BLE-basierter Signaldata für eine automatisierte Klassifizierung logistischer Prozesse

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der
Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind
im Internet über <http://dnb.d-nb.de> abrufbar.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche
Nationalbibliografie; detailed bibliographic data are available in the
Internet at <http://dnb.d-nb.de>.

ISBN 978-3-95908-798-8

© 2025 Dresden und München
Thelem Universitätsverlag & Buchhandlung GmbH & Co. KG
<http://www.thelem.de>

TUDpress ist ein Imprint von Thelem
Alle Rechte vorbehalten. All rights reserved.
Gesetzt vom Autor.
Printed in Germany.



**Möglichkeiten und Grenzen BLE-basierter Signaldaten für eine
automatisierte Klassifizierung logistischer Prozesse**

Von Madlin Weise
Geboren am 04. März 1988 in Dippoldiswalde

zur
Erlangung des akademischen Grades
Doktor Ingenieur
(Dr.-Ing.)

vorgelegte Dissertation.

Tag der Einreichung: 04. Juni 2024
Tag der Disputation: 05. Dezember 2024

1. Gutachter: Prof. Dr.-Ing. habil. Thorsten Schmidt
2. Gutachter: Prof. Dr.-Ing. Robert Schulz

Vorwort

Die vorliegende Dissertation entstand während meiner Tätigkeit als wissenschaftliche Mitarbeiterin an der Professur für Technische Logistik des Instituts für Technische Logistik und Arbeitssysteme der Technischen Universität Dresden. Mein besonderer Dank gilt Herrn Professor Dr.-Ing. habil. Thorsten Schmidt für das mir entgegengebrachte Vertrauen, die intensive und engagierte Betreuung sowie die wertvollen Anregungen und Hinweise, die die Anfertigung dieser Arbeit optimal unterstützten. Herrn Professor Dr.-Ing. Robert Schulz danke ich herzlich, für die mir entgegengebrachte Bereitschaft zur Begutachtung meiner Dissertation.

Die überaus hilfreichen fachlichen Diskussionen mit Herrn Professor Dr.-Ing. Michael Völker sowie Herrn Dr.-Ing. Mathias Kühn und der daraus entstandene Zugang zu einem schier unerschöpflichen Fachwissens- und Erfahrungsschatz waren mir im Erstellungsprozess dieser Arbeit eine fundamentale fachliche Unterstützung. Dafür spreche ich Herrn Professor Völker sowie Herrn Dr.-Ing. Kühn meinen besonderen Dank aus. Allen ehemaligen Kollegen der Arbeitsgruppe „Fabrikplanung und Fabrikbetrieb“, die mich während meiner Zeit als wissenschaftliche Mitarbeiterin stets hilfreich begleiteten und eine sowohl fachliche als auch menschlich bereichernde Arbeitsatmosphäre schufen, die einen wichtigen Anteil zum erfolgreichen Abschluss dieser Dissertation beitrugen, möchte ich vielmals danken. Meine Dankbarkeit drücke ich zudem den damaligen studentischen Hilfskräften Herrn Peter Klengler und Herrn Martin Volkmann für ihre aktive und konstruktive Zusammenarbeit sowie Unterstützung bei der experimentellen Datenerhebung und deren Vorbereitung.

Meiner Familie, die mich in jeder Phase der Erstellung dieser Dissertation vorbehaltlos unterstützt hat und mich mit motivierenden und fordernden Gesprächen in der finalen Erstellungsphase begleitete, gilt ein ganz besonderer Dank. Allen voran möchte ich meinem Mann, Dr.-Ing. Daniel Weise, danken. Mit seiner Beharrlichkeit, seinem Vertrauen sowie seiner nicht enden wollenden Motivation und Diskussionsbereitschaft hat er v. a. die finale Entstehungsphase der vorliegenden Dissertation überhaupt erst ermöglicht. Seine Bereitschaft den Familienalltag allein zu gestalten und seine liebevolle Fürsorge für unsere Tochter hat mir den entscheidenden Freiraum zur Finalisierung der Dissertation gegeben.

Dresden, im Juni 2024

Madlin Weise

Inhaltsverzeichnis

Vorwort.....	II
Inhaltsverzeichnis	III
Abbildungsverzeichnis	V
Tabellenverzeichnis	IX
Abkürzungsverzeichnis	XI
Symbolverzeichnis	XIII
1 Einleitung	1
1.1 Motivation und Problemstellung	1
1.2 Forschungsgegenstand und Ziele	3
1.2.1 Konkretisierung Forschungsgegenstand	3
1.2.2 Ableitung Zielstellung und Forschungsfragen	5
1.3 Forschungsdesign	6
1.4 Aufbau der Arbeit	7
2 Theoretische Grundlagen – Stand der Technik	8
2.1 Erläuterung zentraler Begrifflichkeiten zum Thema	8
2.2 Funksignale & Beacon-Technologie	8
2.2.1 Bluetooth	8
2.2.2 Bluetooth Low Energy (BLE)	9
2.2.3 Beacon-Technologie	10
2.2.4 Signalausbreitungsmodell und Signaldämpfung	12
2.2.5 Signalglättung	17
2.3 Maschinelles Lernen: Mustererkennung	19
2.3.1 Unüberwachtes Lernen	19
2.3.2 Überwachtes Lernen	20
2.3.3 Bewertungskriterien	22
2.4 Systematische Literaturanalyse	24
3 Vorbetrachtung zum entwickelten Lösungsansatz	26
3.1 Sukzessiver Ablauf des Lösungsansatz	26
3.2 Modulare Systemkomponenten	27
3.3 Generierung Datenbasis & Datengüte	28
3.3.1 Technologieauswahl & Konfiguration	28
3.3.2 Datenaufnahme	32
3.3.3 Herausforderung Datenqualifizierung	36
3.3.4 Zusammenfassende Maßnahmenableitung	43
4 Lösungsansatz zur automatisierten Zuordnung von BLE-Signalstärkedaten zu logistischen Prozessen	44
4.1 Zielstellung & Anforderungen	44

4.2	Definition intralogistische Referenz-Prozesskette	45
4.3	Überblick Klassifizierungsalgorithmen zur Prozessklassifizierung	47
4.3.1	Literaturüberblick zur Algorithmenauswahl	47
4.3.2	Detaillierung ausgewählter Klassifizierungsalgorithmen	51
4.4	Evaluierung & Auswahl ML-Algorithmen	60
4.4.1	Versuchsaufbau	60
4.4.2	Analyse und Bewertung Klassifizierungsalgorithmen	63
4.4.3	Zusammenfassung	70
4.5	Verbesserung Klassifizierungsgenauigkeit	71
4.5.1	Methodenauswahl zur Erhöhung der Klassifikationsgenauigkeit	71
4.5.2	Umsetzung einer automatisierten Hyperparameteroptimierung zur Verbesserung der Klassifizierungsergebnisse	76
4.5.3	Lösungsansatz zur Einbeziehung von Logikbeziehungen	104
4.5.4	Zusammenfassung	111
4.6	Kritische Würdigung des Lösungsansatzes	113
5	Zusammenfassung und Ausblick	117
5.1	Zusammenfassung	117
5.2	Ausblick & weitere Forschungsfelder	122
	Literaturverzeichnis	124
	Anhang	144

Abbildungsverzeichnis

Abbildung 1-1: Anforderungen und Abgrenzung der zu entwickelnden Analysemethodik [Eigene Darstellung]	4
Abbildung 1-2: Aufbau der vorliegenden Arbeit [Eigene Darstellung]	7
Abbildung 2-1: BLE-Architektur [Eigene Darstellung, i.A.a. microchipdeveloper.com 2021]	9
Abbildung 2-2: Sender-Empfänger-Prinzip [Eigene Darstellung, i.A.a. Apple Inc. 2014]	11
Abbildung 2-3: Zusammenhang Distanz-Signalstärke bei Pfadverlust [Eigene Darstellung, i.A.a. Evennou und Marx 2006]	14
Abbildung 2-4: Verhalten elektromagnetischer Wellen bei Reflexion, Brechung und Absorption sowie Streuung [Eigene Darstellung, i.A.a. Demtröder 2017 und Gustrau 2019]	15
Abbildung 2-5: Verhalten elektromagnetischer Wellen bei Beugung [Eigene Darstellung, i.A.a. Demtröder 2017 und Gustrau 2019]	16
Abbildung 2-6: Unschärfe des Rohsignals [Eigene Darstellung]	18
Abbildung 2-7: Ablauf Klassifizierung [Eigene Darstellung; i.A.a. Lechner 2008]	21
Abbildung 2-8: Generalisierungsproblem bei Klassifizierungsalgorithmen [Eigene Darstellung, i.A.a. Nguyen und Zeigermann 2018]	24
Abbildung 3-1: Überblick gesamtheitlicher Lösungsansatz [Eigene Darstellung]	26
Abbildung 3-2: Modulare Systemkomponenten [Eigene Darstellung]	27
Abbildung 3-3: Funktionalitäten App I	30
Abbildung 3-4: Funktionalitäten App II	31
Abbildung 3-5: Versuchsumgebung I [Eigene Darstellung]	33
Abbildung 3-6: Versuchsumgebung II [vgl. Hein 2019]	34
Abbildung 3-7: Versuchsumgebung III	35
Abbildung 3-8: Szenario-basierte Vorgehensweise laut Versuchstemplate [vgl. Koch 2018]	36
Abbildung 3-9: Empfehlung szenario-basierter Datenbedarf im Versuchstemplate [vgl. Koch 2018]	36
Abbildung 3-10: Einfluss Empfängertechnologie [Eigene Darstellung]	37
Abbildung 3-11: Versuchsaufbau Einfluss Signaldämpfung; Aufnahme in Versuchsumgebung I [Eigene Darstellung]	38
Abbildung 3-12: Signaldämpfung durch Störobjekt; RSSI unbereinigt [Eigene Darstellung]	39
Abbildung 3-13: Signaldämpfung durch Störobjekt; RSSI geglättet ($n = 100$) [Eigene Darstellung]	39
Abbildung 3-14: Versuchsaufbau und Signalstärkemessdaten für Störquelle Stahltür [Eigene Darstellung, i.A.a. Hein 2019]	41
Abbildung 3-15: unbereinigte Datenreihe nach Datenaufnahme per App	42

Abbildung 3-16: Gegenüberstellung aufgefüllte RSSI-Rohdaten zu bereinigte Datenreihe mit Gleitlänge $n = 100$ (entsprechend SP).....	43
Abbildung 4-1: Kategorisierung Untersuchungsszenarien und Kennzahlenbedarf [Eigene Darstellung, i.A.a. Koch 2018].....	46
Abbildung 4-2: Ergebnis Literaturanalyse: verwendete ML-Algorithmen zur Klassifizierung [Eigene Darstellung].....	50
Abbildung 4-3: Varianten der Klasseneinteilung; A: nicht-optimale Einteilung, B: optimale Einteilung [Eigene Darstellung, i.A.a. Kecman 2005]	52
Abbildung 4-4: Grundprinzip Kerneltrick: A: nichtlineare trennbare Daten in \mathcal{X} , B: lineare Hyperebene in \mathcal{M} , C: nichtlineare Hyperebene in \mathcal{X} [Eigene Darstellung, i.A.a. Cleve und Lämmel 2016].....	53
Abbildung 4-5: Gegenüberstellung Kernelfunktionen SVM [Eigene Darstellung, i.A.a. scikit-learn 2021d]	54
Abbildung 4-6: Aktivierungsfunktionen KNN: A: lineare Funktionen, B: sigmoide Funktionen [Eigene Darstellung, i.A.a. Hoffmann 2019].....	55
Abbildung 4-7: Schichtaufbau & Neuronenarten KNN [Eigene Darstellung, i.A.a. Cleve und Lämmel 2016].....	55
Abbildung 4-8: A: Trainingsdatenmenge, B: Schema eines DT mit drei Ebenen für das Zweiklassenproblem in A [Eigene Darstellung, i.A.a. Nguyen und Zeigermann 2018].....	58
Abbildung 4-9: Ablauf Versuch I [Eigene Darstellung]	60
Abbildung 4-10: Prozessfolge Versuch I [Eigene Darstellung]	61
Abbildung 4-11: Ablauf Versuch II [Eigene Darstellung].....	61
Abbildung 4-12: Prozessfolge Versuch II [Eigene Darstellung]	62
Abbildung 4-13: Ablauf Versuch III [Eigene Darstellung]	62
Abbildung 4-14: Prozessfolge Versuch III [Eigene Darstellung]	63
Abbildung 4-15: Vergleich Rechenzeit je Klassifizierer für Klassifizierung und CA [Eigene Darstellung]	65
Abbildung 4-16: Verlauf CA-Werte DT über Gleitlänge $n = [1, \dots, 10]$	66
Abbildung 4-17: Versuch I: Ausschnitt Klassenzuordnung mittels DT, $n = 2$	67
Abbildung 4-18: Versuch II: Klassenzuordnung mittels SVM, $n = 9$	68
Abbildung 4-19: Versuch II: Klassenzuordnung mittels RF, $n = 9$	69
Abbildung 4-20: Versuch III: Klassenzuordnung mittels RF, $n = 10$	70
Abbildung 4-21: Einordnung Anwendung Verfahren zur Klassifizierungsverbesserung [Eigene Darstellung].....	76
Abbildung 4-22: HPO: innere und äußere Optimierung [Eigene Darstellung, i.A.a. Koch et al. 2012]	77
Abbildung 4-23: Ergebnis Literaturreview HPO-Verfahren [Eigene Darstellung]	81
Abbildung 4-24: Bestandteile der GA [Eigene Darstellung, i.A.a. Petersohn 2009].....	86

Abbildung 4-25: Evolutionszyklus GA bei HPO [Eigene Darstellung, i.A.a. Petersohn 2009]	87
Abbildung 4-26: PSO-Zyklus zur HPO von Klassifikatoren [Eigene Darstellung, i.A.a. Paulitschke 2016]	88
Abbildung 4-27: vereinfachter BO-Ablauf zur HPO [Eigene Darstellung, i.A.a. Grado et al. 2018]	89
Abbildung 4-28: Verallgemeinerung Ablauf Parametrierungsskripte [Eigene Darstellung]	89
Abbildung 4-29: SVM: Vergleich der Kernelfunktionen anhand CA (links) und Rechenzeit (rechts) [Eigene Darstellung]	91
Abbildung 4-30: Eingrenzung Untermenge des Hyperparameters C der SVM [Eigene Darstellung]	91
Abbildung 4-31: Auswertung PSO, Versuch III; Gegenüberstellung CA und C in Abhängigkeit der Iterationsanzahl [Eigene Darstellung]	94
Abbildung 4-32: Auswertung PSO, Versuch III; Gegenüberstellung CA und γ in Abhängigkeit der Iterationsanzahl [Eigene Darstellung]	94
Abbildung 4-33: Auswertung PSO, Versuch III; Gegenüberstellung CA und $pSVM$ in Abhängigkeit der Iterationsanzahl [Eigene Darstellung]	95
Abbildung 4-34: Auswertung GA, Versuch II; Gegenüberstellung max. CA und $\emptyset CA$ in Abhängigkeit der Iterationsanzahl [Eigene Darstellung]	96
Abbildung 4-35: Auswertung GA, Versuch III; Gegenüberstellung max. CA und $\emptyset CA$ in Abhängigkeit der Iterationsanzahl [Eigene Darstellung]	96
Abbildung 4-36: Auswertung GA, Versuch II; Gegenüberstellung max. CA und $\emptyset CA$ in Abhängigkeit der Iterationsanzahl [Eigene Darstellung]	97
Abbildung 4-37: Auswertung PSO, Versuch II; Gegenüberstellung nhn , nbz und $\alpha Loss$ in Abhängigkeit der Iterationsanzahl [Eigene Darstellung]	98
Abbildung 4-38: Versuch I: Ausschnitt Klassenzuordnung mittels DT und optimierten Hyperparametern [Eigene Darstellung]	100
Abbildung 4-39: Versuch I: Ausschnitt Klassenzuordnung mittels KNN und optimierten Hyperparametern	101
Abbildung 4-40: Versuch II: Ausschnitt Klassenzuordnung mittels SVM und optimierten Hyperparametern	101
Abbildung 4-41: Versuch II: Ausschnitt Klassenzuordnung mittels KNN und optimierten Hyperparametern [Eigene Darstellung]	102
Abbildung 4-42: Versuch III: Ausschnitt Klassenzuordnung mittels RF und optimierten Hyperparametern [Eigene Darstellung]	103
Abbildung 4-43: Versuch III: Ausschnitt Klassenzuordnung mittels KNN und optimierten Hyperparametern [Eigene Darstellung]	104
Abbildung 4-44: Bestandteile erweiterter Logikalgorithmus [Eigene Darstellung]	105

Abbildung 4-45: Logikalgorithmus Stufe I: Einbeziehung der Fenstergröße fg [Eigene Darstellung]	108
Abbildung 4-46: Logikalgorithmus Stufe II: Beheben von Logikfehlern [Eigene Darstellung]	109
Abbildung 4-47: Versuch I: Einfluss Variation Fenstergröße auf CA bei Anwendung KNN [Eigene Darstellung]	110
Abbildung 4-48: Versuch I: Einfluss der Logikprüfung auf Klassifizierungsergebnis [Eigene Darstellung]	110

Tabellenverzeichnis

Tabelle 2-1: Spezifikation Kommunikationsprotokolle für BLE-Beacons	12
Tabelle 2-2: Konfusionsmatrix Klassifizierungsergebnisse [vgl. Brownlee 2014].....	22
Tabelle 2-3: Fragestellung Eingrenzung Literaturrecherchen	25
Tabelle 3-1: Leistungsparameter Sendertechnologie	28
Tabelle 3-2: Anforderungen Signaldatenerfassung	29
Tabelle 3-3: Implementierte Spezifikationen und Einstell- / Inputparameter der entwickelten Android-App	31
Tabelle 3-4: Leistungsszenarien Datengenerierung	32
Tabelle 3-5: Charakteristika Versuchsumgebung I	33
Tabelle 3-6: Charakteristik Versuchsumgebung II.....	34
Tabelle 3-7: Charakteristik Versuchsumgebung III	35
Tabelle 3-8: Parameter Datenaufnahme	37
Tabelle 3-9: Einordnung Dämpfungseigenschaften ausgewählter Störquellen gemäß eigenen Untersuchungen	41
Tabelle 4-1: Zerlegungsvarianten einer Gabelstaplerprozesskette.....	47
Tabelle 4-2: Ergebnis Klassifizierung mit Standardhyperparametern	64
Tabelle 4-3: Ergebnis Klassifizierung mit Standardhyperparametern & optimalen Gleitlängen n	66
Tabelle 4-4: Versuch I: Leistungsparameter des DT, $n = 2$	67
Tabelle 4-5: Versuch II: Leistungsparameter des SVM, $n = 9$	68
Tabelle 4-6: Versuch II: Leistungsparameter des RF, $n = 9$	69
Tabelle 4-7: Versuch III: Leistungsparameter des RF, $n = 10$	70
Tabelle 4-8: Übersicht Suchstring je Verbesserungsmethode.....	72
Tabelle 4-9: Hyperparameter der betrachteten Klassifikatoren.....	79
Tabelle 4-10: Auswertung Vergleichskriterien Literaturreview je HPO-Verfahren	83
Tabelle 4-11: Einordnung HPO zur Leistungsbeurteilung.....	84
Tabelle 4-12: Einfluss Hyperparameter C , γ , $cSVM$ und $pSVM$ der SVM auf CA	91
Tabelle 4-13: Hyperparameter der Klassifikationsalgorithmen mit eingegrenzten Wertebereichen	92
Tabelle 4-14: Gegenüberstellung CA nach HPO für SVM.....	93
Tabelle 4-15: Gegenüberstellung CA nach HPO für DT	95
Tabelle 4-16: Gegenüberstellung CA nach HPO für RF	97
Tabelle 4-17: Gegenüberstellung CA nach HPO für KNN.....	98
Tabelle 4-18: Gegenüberstellung CA und KV nach HPO	99
Tabelle 4-19: Versuch I: Leistungsparameter des DT und optimierten Hyperparametern	100
Tabelle 4-20: Versuch I: Leistungsparameter des KNN mit optimierten Hyperparametern	101

Tabelle 4-21: Versuch II: Leistungsparameter der SVM mit optimierten Hyperparametern	102
Tabelle 4-22: Versuch II: Leistungsparameter der KNN mit und ohne optimierten Hyperparametern	103
Tabelle 4-23: Versuch III: Leistungsparameter der RF mit optimierten Hyperparametern	103
Tabelle 4-24: Versuch III: Leistungsparameter der KNN mit optimierten Hyperparametern	104
Tabelle 4-25: Reihenfolgematrix („x“: zulässige Reihenfolge, „-“: unzulässige Reihenfolge)	105
Tabelle 4-26: Bestandteile Prüfalgorithmus.....	106
Tabelle 4-27: Fallbeispiel Logikprüfung.....	107
Tabelle 5-1: Möglichkeiten und Grenzen des vorgestellten Lösungsansatzes.....	122

Abkürzungsverzeichnis

Adam	Adaptive Moment Estimation
ADL	Activity of Daily Living
BDE	Betriebsdatenerfassung
BLE	Bluetooth Low Energy
BO	Bayes'sche Optimierung
DB	Datenbank (engl. Data Base)
DBIS	Datenbank-Infosystem
DLZ	Durchlaufzeit
DT	Decision Tree
ERP	Enterprise Resource Planning
FN	Falsch Negativ
FP	Fingerprinting
FP	Falsch Positiv
FRF	Fuzzy Random Forest
FTS	Fahrerloses Transportsystem
GA	Genetische Algorithmen
HAR	Human Activity Recognition
HPO	Hyperparameteroptimierung
PPM ²	Industrial Indoor Positioning Maturity Model
ID	Identification
ID3	Iterative Dichotomiser 3
ISM	Industrial, Scientific and Medical
KDD	Knowledge Discovery in Databases
KMU	klein- und mittelständische Unternehmen
KNN	künstlich Neuronale Netze
KV	Kreuzvalidierung
L-BFGS	Limited memory Broyden-Fletcher-Goldfarb-Shanno
L-DPLM	Log-Distance Path Loss Model
LOS	Line Of Sight
LSE	Linear Least Square Estimation
MA	Moving Average
MAC	Media Access Control
MDE	Maschinendatenerfassung
ML	Maschinelles Lernen (engl. Machine Learning)
NB	Naive Bayes
NLOS	Non Line Of Sight
OF	Overfitting
PSO	Partikelschwarmoptimierung
RBF	Radial-Basis-Funktion

RF	Random Forest
RFID	Radio-Frequency Identification
RN	Richtig Negativ
RP	Richtig Positiv
RS	Rastersuche
RSSI	Received Signal Strength Indicator
SAMME	Stagewise Additive Modeling using a Multi-class Exponential Loss Function
SAMME.R	SAMME Real
SGD	Stochastic Gradient Descent
SLUB	Sächsische Landesbibliothek – Staats- und Universitätsbibliothek
SMD	Smart Mobile Device
SQL	Structured Query Language
SVM	Support Vector Machine
TUL	Transport, Umschlag, Lagerung
UF	Underfitting
UUID/UID	Universal Unique Identifier/ Unique Identifier
UWB	Ultra Wide Band
WLAN	Wireless Local Area Network
ZS	Zufällige Suche

Symbolverzeichnis

A_{LSE}	Matrix der Senderkoordinaten bei LSE
A, B, C	SMD-spezifische Konstanten zur Distanzkalkulation nach Android Beacon Library
A_m	festgelegte Anzahl Attribute zur zufälligen Auswahl aus gesamter Attributanzahl des RF
b_d	Vektor mit den euklidischen Distanzen zwischen Sendern und geschätztem Standort sowie den Senderkoordinaten/
b_{SVM}	Parameter der Verschiebung bei der SVM
C_m	Loss-Funktion für Stichprobe m
C_{SVM}	Bestrafungsfaktor bei SVM
c_{NB_i}	zu erkennende Klasse des NB
c_{SVM}	Inhomogenität bei SVM
CA	Klassifizierungsgenauigkeit (engl. Classification Accuracy)
cr_{A_m}	Kriterien des DT ($H(S)$ und $gini(S)$)
∇C_m	Gradient der SGD bei KNN
d	Entfernung zwischen Sender und Empfänger
d_0	bekannte Entfernung bzw. Referenzentfernung
$d_{max_{DT}}$	Maximale Tiefe des DT
$data_{min_k}$	Mindestdatenanzahl je Knoten
$data_{min_l}$	Mindestdatenanzahl je Blatt
fg	Fenstergröße Prüfalgorithmus
$gini(S)$	Informationsgewinn bzw. die Gini-Verunreinigung des DT
$H(S)$	Entropie des DT
hB_t	Basisklassifikatoren des AB, schwache Komponentenklassifikatoren
hS_t	starker Klassifikator des AB
I_{zv}	maximale Anzahl Datenreihen je Liste
i_{zv}	Zählvariable des Prüfalgorithmus für Datensatz/ -liste mit $i_{zv} = [0, 1, \dots, I_{zv}]$, i_{zv} entspricht aktueller Klasse
kl	Klassenlänge, aufeinanderfolgende gleiche Klassen
l	Gewicht des MA mit $l \in \{0, \dots, n\}$
M	Merkmals- oder Attributsvektor des NB
m	Größe der zufälligen Stichprobe aus Trainingsmenge eines KNN zur Berechnung der SGD
m_{zv}	Zählvariable, Vergleichsklasse
n	Gleitlänge bzw. Fenstergröße bei MA
n_{bz}	Losgröße bei Anwendung Optimierungsfunktionen bei KNN
n_{RF}	Anzahl von Datenreihen eines RF

n_{SP}	Umfang Stichprobe aus n_{RF} mit $n_{SP} = n_{RF}$
n_{max_l}	Maximale Blattanzahl
net_j	Netzwert aus der Summe der Eingangssignale eines KNN
$P(c_{NB_i})$	A-Priori-Wahrscheinlichkeit, Wahrscheinlichkeit Klassenzugehörigkeit zu Klasse c_{NB_i} des NB
$P(c_{NB_i} M)$	A-Posteriori-Wahrscheinlichkeit des NB
$P(M)$	Wahrscheinlichkeit der Attribute/ Merkmale M (Merkmalsvektor) des NB
$P(M c_{NB_i})$	Likelihood-Wahrscheinlichkeit des NB
p_{GA}	Anzahl Individuen je Population des GA
p_{SYM}	Polynomgrad der Polynomkernelfunktion der SVM
$RSSI(d)$	empfangene Signalstärke bei der gesuchten Entfernung d
$RSSI(d_0)$	Referenzwert für die empfangene Signalstärke bei bekannter Entfernung d_0
r_i	Kreisradius bei der Lateration mit $i = (1, \dots, n)$
S	Trainingsdaten des DT
S_p	Schwarmgröße mit Anzahl der Partikel je Population der PSO
SI	Sendeintervall Beacons
SP	Scan Periode SMD
U	Untersuchungsraum (Layout)
v	Vektor des geschätzten Standorts bei LSE
v_{KNN}	Parameter SGD
v'_{KNN}	verschobener Parameter v_{KNN} mit Schrittgröße $v_{KNN} \rightarrow v'_{KNN}$
\bar{w}^T	Stützvektor (engl. Support Vector) der SVM
$w_{i,j}$	Gewichtungen der Verbindungen der Neuronen bei KNN
w_T	Gewichtung der klassifizierten Datenreihen des AB
$w_{vh_{i_{zv}}}$	Vorhersagewahrscheinlichkeit je Klasse
w_{vt}	Vertrauenswahrscheinlichkeit Prüfalgorithmus
\vec{x}_i^T	Merkmalsvektor der Trainingsmenge bei der SVM mit $\vec{x}_i^T = (x_1, \dots, x_n)^T$ und $\vec{x}_i^T \in R^n$
$x_{i,j}$	Eingangssignale bei KNN
\bar{x}_i	arithmetischer Mittelwert für Teil eines Datensatzes mit (x_1, x, \dots, x_N)
x_{OG}	Grenzwert (Maximalwert) für Entfernung von Extremwerten
x_{UG}	Grenzwert (Minimalwert) für Entfernung von Extremwerten
(x_{SP}, y_{SP})	Koordinaten des gemeinsamen Schnittpunktes der Sender B bei der Lateration
Y_k	Ausgangssignal bei KNN
y_i	Klasse bei der SVM mit $y_i \in \{1, -1\}$

α	Einfallswinkel Reflexion
α_{loss}	Bestrafungsfaktor für Gewichtung $w_{i,j}$ bei KNN
α_{pfad}	Koeffizient für den Pfadverlust
β	Ausfallwinkel Reflexion
γ	Einflussfaktor der SVM-Kernelfunktionen, Einfluss der Trainingsdaten
η_{AB}	Lernrate des AB
η_{KNN}	Lernrate des KNN bei Anwendung SGD
ξ_i	Schlupfvariable bei der SVM mit $\xi_i \geq 0$
φ	Aktivierungsfunktion eines Neurons bei KNN

1 Einleitung

1.1 Motivation und Problemstellung

Industrie 4.0 – Digitalisierung, Internet der Dinge, Künstliche Intelligenz – unter diesen Trends erlebt die Produktions- und Logistikbranche seit mehr als zehn Jahren eine bedeutungsvolle Veränderung. Durch immer kürzere Prozesslebenszyklen und komplexer werdende Prozesse ist diese Branche gefordert, flexibel und effizient zu agieren und zu reagieren. Dies bedeutet, dass „(...) die Kosten der logistischen Prozesse für die jeweilige Leistung minimal und ihre Leistung bei den jeweiligen Kosten maximal sein sollen“ [Fleischmann 2018]. Um diesen Grundsatz erfüllen zu können, ist das Ableiten von fundiertem Prozesswissen sowie das Erkennen von Wirkzusammenhängen notwendig [vgl. ten Hompel et al. 2018]. Das Prozesswissen ist entscheidend sowohl für die Gestaltung effizienter Arbeitsabläufe als auch für die Durchführung gezielter Prozessanalysen im Sinne der Industrie 4.0. Der Schwerpunkt der Datengenerierung liegt dabei auf Prozessen des Transports, des Umschlags und der Lagerung (TUL) und somit des Materialflusses [vgl. VDI-Fachbereich Technische Logistik 2019].

Die wesentlichen Bestandteile des Materialflusses sind die Materialflussobjekte, die sich in einem Produktions- oder Logistiksystem bewegen oder bewegt werden. Die Generierung von Prozessdaten auf Basis der Überwachung dieser Objekte kann Erkenntnisse über die Materialflussprozesse geben, aber auch hilfreich sein, übergeordnete Fragestellungen der unternehmerischen Leistungsfähigkeit zu beantworten. Gängige Überwachungsverfahren werden unter dem Begriff „Tracking & Tracing“-Systeme, v. a. Indoor-Trackingsysteme, zusammengefasst. Für die Analyse und Verbesserung von Materialflussprozessen können sich bspw. folgende unternehmerische Fragestellungen ergeben:

- Wie hoch ist die Auslastung der Flurförderzeuge?
- Wie hoch ist der Anteil an Leerfahrten?
- Welche Wege werden häufig genutzt?

Zur Beantwortung dieser Fragestellungen lassen sich logistische Kenngrößen ableiten [vgl. VDI-Fachbereich Technische Logistik 2001, 2007b]. Für die Bestimmung der Kenngrößen sind Informationen aus den Prozessdaten notwendig. Das Wissen, welches durch die Kombination der Informationen mit bspw. Erfahrungen entsteht, ist für Prozessanalysen der eigentliche Fokus. Somit setzt die Generierung und Verarbeitung von Prozessdaten eine hohe Datenqualität voraus. Die Qualität von Daten zeichnet sich durch eine Vielzahl an Kriterien aus, wie bspw. der Konsistenz, Vollständigkeit oder Genauigkeit. Jedoch haben Studien vor allem bei produzierenden Unternehmen gezeigt, dass die Qualität von generierten Prozessdaten für die Umsetzung von Daten-basierten Analysen häufig

unzureichend ist. Die Gründe dafür sind entweder fehlende objektive Datenerfassungssysteme (z. B. BDE, MDE, etc.)¹, fehlendes Knowhow zur Verarbeitung großer Datenmengen sowie fehlendes Wissen über die Potentiale eines gezielten Datenmanagements. [vgl. Axmann et al. 2019]

In diesem Zusammenhang sehen sich auch klein- und mittelständische Unternehmen (KMU) diesen Herausforderungen gegenübergestellt [vgl. Groggert et al. 2017]. Die Gründe für die fehlende objektive Datenaufnahme in KMU sind vielseitig. Neben hohen Investitionen, hohem Installationsaufwand sowie laufenden Kosten (z. B. für Wartung, Software-Lizenzgebühren etc.) ist der Nutzen einer permanenten Datenerfassung im Vorfeld einer solchen Investition auf Grund der fehlenden Prozesskenntnis meist nur schwer quantifizierbar [vgl. Köppe 2014]. KMU unterliegen häufig einer hohen Prozessvarianz sowie vielfältig neuen Prozessen aufgrund neuer Produkte bzw. kurzer Produktlebenszyklen (heterogene Arbeitsinhalte sowie unregelmäßige, vielfältige Prozessfolgen), wodurch die Datenerfassung flexibel anpassbar sein muss. Starre Trackingsysteme zur permanenten sind somit wenig effizient und praktikabel [vgl. Girish 2019]. Erfolgt dennoch der Einsatz digitaler Überwachungssysteme, werden diese meist nicht im vollen Funktionsumfang genutzt und unterliegen einem nicht optimalen Aufwand-Nutzen-Verhältnis.

Aus diesen Gründen wird zur Prozessüberwachung und Beantwortung unternehmerischer Fragestellungen in KMU oft auf manuelle Methoden zur Prozessaufnahme zurückgegriffen, wie z. B. Selbstaufschreibung oder Fremdbeobachtung. Die Anwendbarkeit dieser Methoden ist prozessabhängig sowie die Aussagefähigkeit der Analyseergebnisse meist begrenzt, da sich in komplexen Materialflusssystemen diverse Prozesse überlagern und somit zur Generierung einer fundierten Datenbasis ein hoher Ressourcenaufwand (Zeit, Mitarbeiter, Kosten, etc.) notwendig wird. Aufgrund eingeschränkter Personalressourcen werden diese Tätigkeiten somit immer häufiger an externe Beratungsunternehmen ausgelagert. Dies begründet weitere Herausforderungen bei der Anwendung von manuellen Methoden. Ein menschlicher Beobachter kann nur das notieren, was dieser versteht und den richtigen Prozessschritten zuordnet (Interpretation). Ein tiefes Detailwissen ist in der Regel nicht vorhanden, da ein externer Beobachter nicht alle Prozesselemente kennen kann (Abstraktion). Die Informationen sind unvollständig, was zu einer Verfälschung der Daten bzw. zu fehlerhaften Rückschlüssen bei deren Auswertung führen kann. [vgl. Martin 2013; VDI-Fachbereich Technische Logistik 2017]

Es fehlt somit eine Methodik, mit der sowohl KMU selbst als auch beratende Unternehmen befähigt werden, aufwandsarm, flexibel und dennoch objektiv Produktions- und Logistikprozesse messen und analysieren zu können. Daraus ergeben sich folgende allgemeingültige Anforderungen an ein geeignetes Mess- und Analysesystem:

- geringer Initialisierungs-, Installations- und Wartungsaufwand

¹ BDE: Betriebsdatenerfassung; MDE: Maschinendatenerfassung.

- geringer Ressourcenbedarf für Systemanwendung
- autonome Systemfunktionalität
- intuitive Bedienbarkeit
- Eignung für temporären Einsatz, somit wiederverwendbar und an variierende Prozesse anpassbar

Auf der Erkenntnisbasis unterliegt diese Arbeit der übergeordneten Aufgabenstellung, mit Hilfe der Bluetooth-Technologie eine Methodik zur effizienten, temporär durchführbaren Datenerhebung für flexible Produktions- und Logistiksysteme zu entwickeln. Dabei werden mit möglichst geringem Aufwand sowohl eine hinreichend genaue und automatisierte Klassifizierung logistischer Prozesse als auch eine objektive Generierung von analysefähigen Prozessdaten angestrebt.

1.2 Forschungsgegenstand und Ziele

1.2.1 Konkretisierung Forschungsgegenstand

Das übergeordnete Ziel der vorliegenden Arbeit besteht darin, eine praxistaugliche Methodik zur objektiven Generierung logistischer Prozessdaten zu realisieren, die basierend auf der Bluetooth-Technologie die Komponenten der Datengenerierung sowie der Datenaufbereitung und -weiterverarbeitung mit geringem Aufwand und benutzereffizient umsetzen kann. Es folgt auf Basis der genannten Anforderungen eine inhaltliche Konkretisierung.

Die grundlegende Funktionsweise von Indoor-Trackingsystemen² basiert auf einem Sender-Empfänger-Prinzip. Das bedeutet, dass zwischen einem Signalgeber und einem Empfangsgerät Daten übermittelt werden, die je nach Anwendungsfokus entsprechende Daten (Entfernung, Position, Sensorinformationen etc.) verfügbar machen. Beim Einsatz von Indoor-Trackingsystemen ist die wesentliche Herausforderung, die vorhandenen Hindernisse zu berücksichtigen, die die Datenübertragung beeinflussen.

Unter Berücksichtigung der Anforderungen an die zu entwickelnde Methodik ist insbesondere auf einen geringen Ressourcenbedarf zu achten. Die autonome Anwendbarkeit erfordert, dass Analysen der jeweiligen Stoffstromprozesse ausschließlich auf Basis der autonom generierten Prozessdaten möglich sind, ohne dass auf unternehmensinterne Datensysteme zugegriffen werden muss. Die Forderung nach einfacher Bedienbarkeit hängt von der eingesetzten Technik und Technologie ab, insbesondere aber von der für den Anwender sichtbaren Benutzeroberfläche. Diese sollte intuitiv gestaltet sein und dem Anwender eine eigenständige Handhabung bei Installation, Initialisierung und Betrieb des Trackingsystems ermöglichen.

² Nachfolgend werden die Begrifflichkeiten Lokalisierungs-, Tracking- sowie Positionierungssystem synonym verwendet.

Zur Umsetzung der Zielsetzung dieser Arbeit wird auf den Bluetooth-Standard zurückgegriffen, da es sich hierbei um eine seit vielen Jahren weit verbreitete und etablierte Technologie handelt. Durch die Weiterentwicklung des Standards zu Bluetooth Low Energy (BLE) sowie die Eignung für den Indoor-Bereich haben sich die Anwendungsgebiete erheblich erweitert.

Als technisch-technologische Grundlage werden Bluetooth Low Energy (BLE)-Beacons als Signalsender und Smart Mobile Devices (SMD), bspw. Smartphones, als Empfänger eingesetzt. Beacons (siehe Abschnitt 2.2.3) sind einfach in der Handhabung, haben eine integrierte Energieversorgung und sind sowohl mit geringem finanziellem Aufwand zu beschaffen als auch zu installieren. BLE ermöglicht durch einen geringen Energiebedarf bei der Signalübertragung einen vereinfachten autonomen Einsatz bei der Vernetzung von Peripheriegeräten (Sender, Empfänger). Da BLE von gängigen mobilen und stationären Geräten (z. B. Smartphone, Laptops, Raspberry Pis etc.) empfangen und verarbeitet werden kann, sind für die Empfangsgeräte keine aufwendige Technikauswahl und -installation notwendig. Die Wahl der Empfangsgeräte ist zudem neben dem Kostenaspekt in der einfachen Bedienbarkeit und in der breiten Verfügbarkeit begründet. Es steht damit eine flexibel sowie autonom einsetzbare Technologie zur Verfügung.

Die Herausforderung der technisch-technologischen Kombination besteht in der effizienten Analyse bzw. Auswertung der Datenbasis, die ausschließlich auf Received Signal Strength Indicator (RSSI)-Daten basiert. Es werden keine zusätzlichen Signal- oder Sensordaten zur Analyse und Auswertung einbezogen. Dies stellt sicher, dass das generierte Datenvolumen begrenzt wird und somit eine lokale Speicherung der generierten Signaldaten auf den Empfangsgeräten möglich ist. Ein zusätzlicher Bedarf an Zugriffsrechten auf unternehmensinterne Systeme oder einer Internetverbindung zur Datenübertragung ist nicht erforderlich.

Abbildung 1-1 zeigt die Systemanforderungen, deren technisch-technologische Grundlage sowie die Herausforderungen der zu entwickelnden Methodik in komprimierter Form.

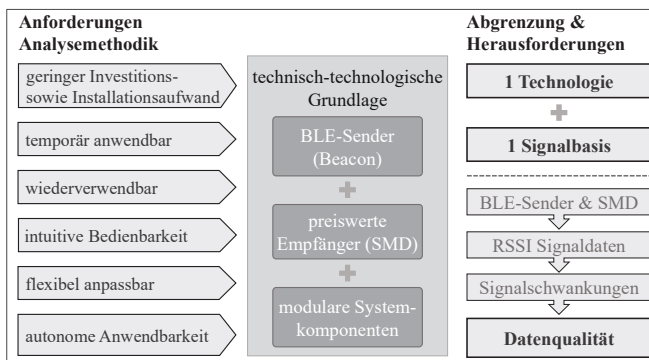


Abbildung 1-1: Anforderungen und Abgrenzung der zu entwickelnden Analysemethodik [Eigene Darstellung]

1.2.2 Ableitung Zielstellung und Forschungsfragen

Für die Gestaltung der effizienten Analysemethodik besteht der primäre Forschungsansatz dieser Dissertation in folgender Zielstellung:

Automatische Erkennung logistischer Prozesse ausschließlich auf Basis der aufgenommenen BLE-Signalstärkedaten mit Hilfe von überwachten Maschinellen Lernalgorithmen mit hinreichend genauen Klassifizierungsergebnis.

Um neben der intuitiven Handhabung bei der Prozessdatengenerierung auch eine gezielte sowie automatisierte Materialflussanalyse zu unterstützen, wird eine autonome Zuordnung vordefinierter logistischer Prozessklassen (z. B. Transport, Leerfahrt etc.) zu den generierten BLE-Prozessdaten umgesetzt. Als Datenbasis dienen die empfangenen RSSI-Daten der sich relativ zueinander bewegenden Materialflussobjekte. Zur Umsetzung der Klassenzuordnung werden Algorithmen aus dem Bereich des Maschinellen Lernens (ML) eingebunden, sogenannte Klassifizierungsalgorithmen. Das Erreichen einer hinreichend genauen Zuordnung der Prozessklassen zur BLE-Signaldatenbasis steht im Entwicklungsfokus. Aufgrund der eingeschränkten Datenbasis wird erwartet, dass die angestrebte Klassifizierungsgenauigkeit nicht durchgehend gegeben ist. Deshalb ist für die erfolgreiche Umsetzung des Forschungsziels die Implementierung bzw. die Entwicklung von Ansätzen zur Erhöhung der Klassifizierungsgenauigkeit vorgesehen.

Hierfür werden die folgenden Forschungsfragen im Rahmen dieser Arbeit beantwortet:

FF1.1 Welche Klassifizierungsalgorithmen eignen sich für die zugrundeliegenden BLE-basierten Signaldaten?

FF1.1.1 Anhand welcher Leistungsparameter kann eine hinreichend genaue Klassenzuordnung gemessen werden?

FF1.1.2 Welche Mindestanforderung muss an die Leistungsparameter gestellt werden, um eine zielführende Klassenzuordnung zu realisieren?

FF1.2 Mit welcher Granularität muss eine logistische Prozesskette in Prozessklassen zerlegt werden, um eine hinreichende Genauigkeit bei der Klassifizierung zu erreichen?

FF1.3 Welche Verfahren zur Verbesserung der Klassenzuordnung sind für den Anwendungsfall geeignet und welche Verbesserungen sind realisierbar?

FF1.4 Sind Verfahren zur Erkennung und zur Korrektur von fehlerhaften Klassenzuordnungen für die BLE-Signaldatenbasis verfügbar und anwendbar?

FF 1.5 Ist BLE als Technologie für die Generierung klassifizierungsfähiger Daten geeignet?

Um die genannten Zielstellungen umsetzen zu können, ist als weitere Herausforderung die Unschärfe der BLE-Signalrohdaten (schwankende RSSI-Werte) gegeben. Die Qualität der Inputdaten ist proportional zu den Ergebnisdaten. Somit ist eine Aufbereitung der unscharfen Signaldaten grundlegend, um die notwendige Datenqualität für die Anwendung der Klassifizierungsalgorithmen bereitzustellen. Es bedarf bspw. sowohl der spezifischen

Konfiguration der Sender- und Empfängertechnik als auch der Auswahl geeigneter Verfahren sowie Filter zur Glättung der gegebenen Signalschwankungen der RSSI-Daten. Ergänzend zur Sicherstellung der Datenqualität ist es für die wissenschaftliche Aufbereitung der folgenden Entwicklungsschritte entscheidend, wie sich der aktuelle Wissensstand in Forschung und Praxis zu den definierten Zielstellungen gestaltet. Die Aufbereitung des Forschungs- und Technikstandes dient dazu, die Relevanz der in dieser Arbeit durchzuführenden Entwicklungen durch vorhandene Forschungslücken zu konkretisieren und zu unterstützen. Zudem werden verfügbare Methoden und Ansätze evaluiert.

1.3 Forschungsdesign

Wird die Themenstellung der vorliegenden Arbeit in die betreffenden Problemdomänen zerlegt, ist der interdisziplinäre Charakter deutlich erkennbar. Erste wesentliche Problemdomäne ist der Bereich der Materialflussanalyse. Die zu entwickelnde Methodik dient vorrangig der Befähigung zur Analyse logistischer und produktiver Prozesse. Die dadurch generierte Unterstützung zur Digitalisierung der Datenaufnahme und -analyse ist in der Problemdomäne der Industrie 4.0 verankert. Die Anwendung von ML-Verfahren zur Datenauswertung wird dem Forschungsbereich des Data Science zugeordnet. Der zunehmende Fokus der Industrie 4.0 auf die Digitalisierung von Produktion und Logistik hat den Bereich des Data Science sukzessive integriert.

Die Interdisziplinarität spiegelt sich auch in den tangierenden Wissenschaftsbereichen wider. Die zu entwickelnde Analysemethodik mit der technisch-technologischen Basis zur Datengenerierung kann den Ingenieurwissenschaften (oder auch Technikwissenschaften) zugewiesen werden [vgl. acatech 2013]. Die darüberhinausgehende softwareseitige Datenaufbereitung sowie die Einbindung von ML-Algorithmen werden der angewandten Informatik zugeordnet.

Die Forschungsmethode der systematischen Literaturrecherche wird als Basis jedes Forschungsziels eingesetzt und fundiert die wissenschaftliche Notwendigkeit der Entwicklungen sowie deren Grundlagen. Für die Zielstellung der automatisierten Prozesszuordnung werden entweder literaturbasierte Algorithmen (Klassifizierungs-, Optimierungsalgorithmen) eingebunden oder diese szenario-spezifisch angepasst. Die Forschungsmethode des deduktiven Schließens wird für die Zielstellung der Verbesserung der Klassifizierungsergebnisse, im Konkreten für die Ableitung eines Regelalgorithmus, angestrebt. Mit Hilfe von wiederholenden Experimenten erfolgt zugleich die Evaluierung der Ergebnisse der einzelnen Entwicklungsschritte bzw. Forschungsziele. [vgl. acatech 2013]

1.4 Aufbau der Arbeit

Die Dissertation gliedert sich im Gesamten in fünf Kapitel (siehe Abbildung 1-2).

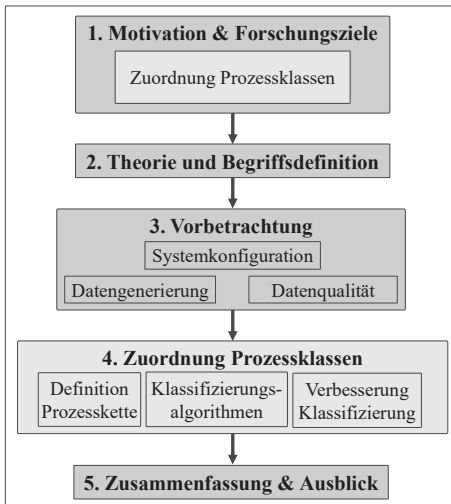


Abbildung 1-2: Aufbau der vorliegenden Arbeit [Eigene Darstellung]

Nach dem einleitenden ersten Kapitel werden in Kapitel 2 wesentliche Begriffe zum grundlegenden Verständnis der Inhalte dieser Arbeit erläutert. Ergänzend dazu werden die theoretischen Grundlagen zu Funksignalen, zur Beacon-Technologie sowie zum Maschinellen Lernen dargelegt. Abschließend wird die Vorgehensweise für die durchzuführenden Literaturrecherchen allgemeingültig erläutert.

In den Abschnitten des dritten Kapitels werden alle Aspekte dargelegt, die notwendig sind, um die nachfolgenden Entwicklungen umsetzen zu können: Auswahl und Konfiguration der technisch-technologischen Basis, Datenaufnahme und -aufbereitung, Einordnung der Entwicklungselemente in den Lösungsansatz sowie grundlegende Erläuterungen zur softwareseitigen Umsetzung der Methodik.

Das Kapitel 4 widmet sich inhaltlich der Umsetzung der automatisierten Zuordnung logistischer Prozessklassen zur Signaldatenbasis und bildet somit den Hauptteil der vorliegenden Arbeit. Dafür werden eingangs die Zielstellung und Anforderungen sowie die Referenzprozesskette definiert. Auf Basis einer Literaturrecherche werden geeignete Verfahren zur Klassifizierung abgeleitet und anhand von Versuchsdaten erprobt. Zur Erreichung einer hohen Genauigkeit der Prozesszuordnung werden Methoden literaturbasiert ausgewählt, umgesetzt und mittels Versuchsdaten auf deren Eignung geprüft. Abschließend werden die Entwicklungsbestandteile und -methoden kritisch betrachtet und eingeordnet.

Die Dissertation wird in dem abschließenden fünften Kapitel zusammengefasst und ein weiterführender Ausblick gegeben.