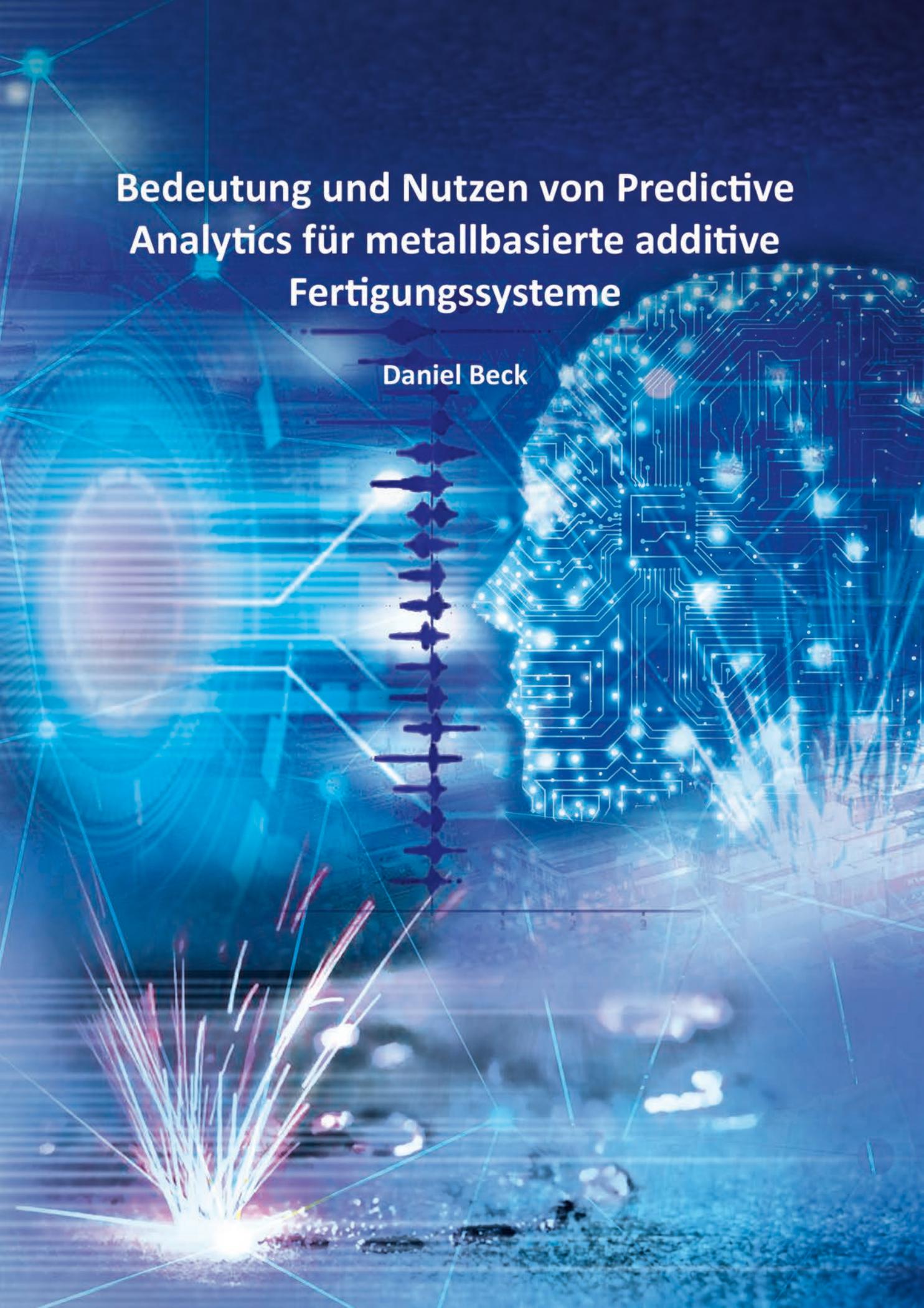


Bedeutung und Nutzen von Predictive Analytics für metallbasierte additive Fertigungssysteme

Daniel Beck



Bedeutung und Nutzen von Predictive Analytics für metallbasierte additive Fertigungssysteme

DISSERTATION
an der
Julius-Maximilians-Universität Würzburg



zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaften

vorgelegt von
Daniel Beck

eingereicht bei
Prof. Dr. Frédéric Thiesse
und
Prof. Dr. Christoph M. Flath

Würzburg, 2023

Berichte aus der Wirtschaftsinformatik

Daniel Beck

**Bedeutung und Nutzen von Predictive Analytics
für metallbasierte additive Fertigungssysteme**

Shaker Verlag
Düren 2023

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.

Zugl.: Würzburg, Univ., Diss., 2023

Copyright Shaker Verlag 2023

Alle Rechte, auch das des auszugsweisen Nachdruckes, der auszugsweisen oder vollständigen Wiedergabe, der Speicherung in Datenverarbeitungsanlagen und der Übersetzung, vorbehalten.

Printed in Germany.

ISBN 978-3-8440-9108-3

ISSN 1438-8081

Shaker Verlag GmbH • Am Langen Graben 15a • 52353 Düren

Telefon: 02421 / 99 0 11 - 0 • Telefax: 02421 / 99 0 11 - 9

Internet: www.shaker.de • E-Mail: info@shaker.de

Danksagung

Mein herzlicher Dank geht an alle Unternehmen, die dieses Projekt mit ihrem Fachwissen und ihren Daten überhaupt erst ermöglicht haben. Des Weiteren danke ich meinen sehr geschätzten Kollegen und Partnern für ihre Bereitschaft, ihre fundierten technischen Erfahrungen sowie kreativen und innovativen Lösungen mit mir zu teilen. Meinem Doktorvater Prof. Dr. Frédéric Thiesse danke ich für die wissenschaftliche und praxisorientierte Zusammenarbeit sowie der umfangreichen fachlichen wie auch strategischen Unterstützung. Einen herzlichen Dank möchte ich ebenfalls an Prof. Dr. Christoph M. Flath für seine zielorientierte und effiziente Betreuung richten. Darüber hinaus möchte ich mich auch bei meinem Softwaredienstleister für die professionelle Unterstützung bei der Verarbeitung der großen Datenmengen bedanken. In ganz besonderem Maße danke ich meiner Frau und meinen Kindern. Vielen Dank, dass ihr mir in den letzten Jahren den Rücken freigehalten habt.

Zusammenfassung

Die vorliegende Arbeit zeigt, dass mit Hilfe von künstlicher Intelligenz ungeplante Ausfälle metallbasierter additiver Fertigungssysteme zu Beginn der Produktionsphase vorhergesagt werden können. Es wurde nachgewiesen, dass die Verwendung von maschinellen Lernverfahren in diesem Bereich zu erheblichen Kosteneinsparungen führt. Das metallbasierte additive Fertigungsverfahren PBF-LB/M (Engl.: Laser-based powder bed fusion of metals, kurz PBF) nimmt im industriellen Umfeld eine starke Position ein. Immer mehr Komponenten in kritischen Anwendungen in der Automobil-, Luftfahrt-, Energie- und Medizintechnikbranche erzeugen hohe Anforderungen an die Bauteilqualität, Prozessstabilität und Maschinenzuverlässigkeit. Aufgrund von Schwankungen in diesen Bereichen steht die additive Fertigung jedoch weiterhin an der Schwelle zum Einsatz in der industriellen Serienproduktion. Bisherige Forschungen auf diesem Gebiet haben in der Regel einen rein materialwissenschaftlichen- bzw. technischen Bezug aus dem Maschinenbau. Die Vermeidung ungeplanter Ausfälle durch künstliche Intelligenz und die damit verbundene Verbesserung der Maschinenzuverlässigkeit wurden aus Sicht der Wirtschaftsinformatik in dieser Form das erste Mal untersucht.

Als Grundlage dienten umfassende Datenbestände mehrerer tausend Baujobs von PBF-Systemen, die Anwendungen aus stark diversifizierten Industriebereichen gefertigt haben. Dies ermöglichte die Erstellung valider Modelle zur Vorhersage ungeplanter Maschinenausfälle. Das Vorgehen beruht auf CRISP-DM, dem Standard für die Planung und Durchführung von Analytics-Projekten. Im Rahmen der Datenextraktion und -aggregation erfolgte eine Zerlegung der Baujobdaten in verschiedene Zeitsegmente. Dies ermöglichte es den frühestmöglichen Vorhersagezeitpunkt zu identifizieren. Um das beste Lernverfahren zu ermitteln, wurde ein Vergleich der ROC-Kurven (Engl.: Receiver Operating Characteristic) als Prognosegütemaß durchgeführt. Das Gradient Boosted Trees Modell zeigte die beste Prognosequalität und wurde anschließend mit XGBoost optimiert. Dies führte zu einer exzellenten Prognosequalität in den ersten 3–6 Stunden der Produktionsphase.

Um die wirtschaftlichen Auswirkungen ungeplanter Ausfälle berechnen zu können, wurden die wesentlichen Faktoren in einem Kostenmodell zusammengefasst. Basierend auf den kostenminimalen Schwellenwerten, mit denen sich das Modell für eine Vorhersage entscheidet, wurde die Berechnung des Einsparpotenzials durch vorbeugende Eingriffe anhand verschiedener Szenarien durchgeführt. Je nach Bauteilvolumen, Produktionszeit und verschiedenen Ausfallraten konnten durchschnittliche Kosteneinsparungen von 10–40 % nachgewiesen werden. Mithilfe von SHAP-Visualisierungen konnte außerdem die Bedeutung einzelner Attribute genutzt werden, um Entscheidungen des verwendeten Black-Box-Modell nachzuvollziehen. Auffällig ist, dass neben den Sensorwerten vor allem die Faktoren der Bauteilgeometrie und des Alters der Maschine eine Rolle bei der Entscheidungsfindung spielen. Des Weiteren ermöglicht SHAP neben einer Plausibilitätsprüfung eine Grundlage für die Erforschung von zusätzlichen präventiven Gegenmaßnahmen.

Abschließend zeigen weiterführende Implikationen mögliche Anwendungsbereiche dank des Datenaustausches zwischen allen Akteuren der additiven Wertschöpfungskette. Neben der Verwendung künstlicher Intelligenz zur Vorhersage von Maschinenstillständen können dadurch weitere Anwendungen entstehen und jeder Marktteilnehmer kann durch zusätzliche Wettbewerbsvorteile profitieren. Zusammenfassend ebnen die vorgestellten Ergebnisse der additiven Fertigung den Weg für einen weiteren Schritt in Richtung industrielle Massenproduktion.

Abstract

This research study demonstrates that unplanned failures of metal-based additive manufacturing systems in the early stages of production can be predicted by using artificial intelligence. It has been demonstrated that the use of machine learning methods in this field results in considerable cost savings. Laser-based powder bed fusion of metals (PBF-LB/M, short PBF) plays an important role in the industry. An increasing number of components in critical applications in the automotive, aerospace, energy and medical technology industries place high demands on component quality, process stability and machine reliability. However, due to fluctuations in these areas, additive manufacturing is still in the process of being introduced into industrial mass production. Previous research in this field has generally had a purely material science or technical background from the field of mechanical engineering. This paper investigates the prevention of unplanned downtime using artificial intelligence and the associated improvement in machine reliability from a business IT perspective in this form for the first time.

This was based on comprehensive data sets of several thousand build jobs from PBF-systems that produced applications from a wide range of industrial sectors, making it possible to create valid models to predict unplanned machine downtime. The approach is based on CRISP-DM, the standard for planning and implementation of analytics projects. In the course of data extraction and aggregation, the build job data was broken down into different time segments, which enabled the earliest possible forecast time to be identified. To identify the best learning process, a comparison of the ROC (Receivers Operating Characteristic) curves was carried out to measure the forecasting quality. The Gradient Boosted Trees model demonstrated the best forecasting quality and was subsequently optimised with XGBoost, resulting in excellent forecasting quality in the first 3-6 hours of the production stage.

The main factors were summarised in a cost model to calculate the economic impact of unplanned downtime. The calculation of the savings potential due to preventive interventions was based on the cost-minimum thresholds at which the

model decides to make a prediction. Depending on component volume, production time and various failure rates, the model showed average cost savings of 10-40 %. Using SHAP visualisations, it was also possible to use the individual attributes to understand the decisions taken by the black box model used. It is noteworthy that, in addition to the sensor values, the factors of the component geometry and the age of the machine play a major role in the decision-making process. Furthermore, SHAP permits a plausibility check as well as providing a basis for the exploration of additional preventive measures.

Finally, further implications show possible areas of application thanks to the exchange of data between all stakeholders in the additive value chain. In addition to the use of artificial intelligence to predict machine downtime, this can lead to other applications and every market player can benefit from increased competitive advantage. In summary, the results of additive manufacturing presented pave the way for a further step towards industrial mass production.

Inhaltsverzeichnis

Danksagung	i
Zusammenfassung	iii
Abstract.....	v
Abbildungsverzeichnis	xi
Tabellenverzeichnis	xvii
Quellcodeverzeichnis.....	xix
Abkürzungsverzeichnis.....	xxi
1 Einleitung und Motivation.....	1
1.1 Predictive Analytics für additive Fertigungstechnologien	5
1.2 Data Mining nach CRISP-DM.....	14
1.3 Aufbau der Arbeit	15
2 Additive Fertigung und PBF-Systemaufbau.....	17
2.1 Prinzip und Verfahrensübersicht	18
2.1.1 Lasersintern (SLS).....	21
2.1.2 Powder Bed Fusion (PBF).....	22
2.1.3 Laserauftragsschweißen (DMD).....	25
2.1.4 Elektronenstrahlschmelzen (EBM).....	25
2.1.5 Material Extrusionsverfahren (FFF).....	26
2.1.6 Material Verdüsung (3DP)	26
2.1.7 Lichtbogen-Drahtschweißen (WAAM).....	26
2.1.8 Multi-Jet Modelling (MJM).....	27
2.1.9 Stereolithografie (SL).....	27
2.1.10 Poly-Jet Modelling (PJM).....	28
2.1.11 Digital Light Processing (DLP).....	28
2.1.12 Aerosolverfahren (AJ)	28
2.1.13 Laser Chemical Vapor Deposition (LCVD).....	29
2.2 PBF-System Aufbau und Funktionsweise	29
2.2.1 Prozesskammer	31

2.2.2	Optikeinheit.....	31
2.2.3	Schutzgassystem.....	32
2.2.4	Prozessrelevante- und komponentenbezogene Sensordaten	33
2.2.5	Einflussfaktoren auf den PBF-Prozess	34
3	Betriebswirtschaftliche Zielsetzung	37
3.1	Ungeplante Ausfälle von PBF-Systemen.....	38
3.1.1	Auswirkungen von Störungen.....	38
3.1.2	Relevanz von Produktionsunterbrechungen.....	40
3.2	Betriebswirtschaftliche Bewertung von ungeplanten Ausfällen	43
3.2.1	Fertigungskosten von PBF-Systemen	43
3.2.2	Kostenmodell für den ungeplanten Produktionsausfall	45
3.3	Betriebswirtschaftliche Ziele.....	52
3.4	Data-Mining-Ziele.....	55
4	Datengrundlage	57
4.1	Datenbeschreibung und -struktur	59
4.1.1	Sensorprotokolldatei.....	59
4.1.2	Ereignisprotokolldatei	63
4.1.3	Konfigurationsprotokolldatei	65
4.2	Zusammenhänge und Abhängigkeiten der Rohdaten.....	65
4.2.1	Abhängigkeiten mit der Zielvariable.....	66
4.2.2	Zusammenhänge einzelner Sensorwerte	68
4.3	Verifikation Datenqualität.....	71
5	Datenvorbereitung	75
5.1	Rohdatenvorbereitung	75
5.2	Datenaggregation	76
5.2.1	Extrema, Mittelwert, Median und Standardabweichung	80
5.2.2	Differenz und Vergleichswerte	83
5.2.3	Wendepunkte.....	84
5.2.4	Wölbung und Schiefe	84
5.2.5	Interquartilsabstand	85
5.2.6	Korrelationskoeffizienten.....	86
5.2.7	Schichtbezogene und allgemeine Aggregation	87
5.3	Attributsselektion	89

6	Modellerstellung	95
6.1	Einordnung Lernverfahren	96
6.2	Modellbildung und Parameterauswahl	98
6.2.1	Rapidminer als Werkzeug	99
6.2.2	Trainings- und Testdaten	99
6.2.3	Generalisierte Lineare Modelle	101
6.2.4	Logistische Regression	102
6.2.5	Entscheidungsbaum	103
6.2.6	Random Forest	106
6.2.7	Gradient Boosted Trees	109
6.2.8	Neuronale Netze	112
6.2.9	Support Vector Machine	115
6.3	Bewertung und Modellvorauswahl	117
6.3.1	Prognosegenauigkeit und Konfusionsmatrix	118
6.3.2	Grenzwertoptimierungskurve (ROC, AUC)	119
6.3.3	Kreuzvalidierung	122
6.3.4	Performance Vergleich der Modelle	123
6.4	Modelloptimierung	125
6.4.1	XGBoost	126
6.4.2	Optimierung Hyperparameter	129
6.4.3	Attributsauswahl	136
6.5	Ergebnisse in Abhängigkeit zum Vorhersagezeitpunkt	139
7	Modellevaluation	143
7.1	Interpretation der Prognosegüte zum Vorhersagezeitpunkt	143
7.2	Blackbox Interpretation der Modellergebnisse	145
7.2.1	SHAP	147
7.2.2	Umsetzung in Python	155
7.2.3	Interpretation von SHAP-Plots	156
7.3	Betriebswirtschaftliche Evaluation	164
7.4	Diskussion und Implikationen	171
7.4.1	Praktische Anwendung	172
7.4.2	Strategische Implikation	173
7.4.3	Additive Wertschöpfungskette	175
8	Fazit und Ausblick	185

9	Literaturverzeichnis.....	189
	Anhang	215
A.1	CRISP-DM.....	215
A.2	Einflussfaktoren und Ablauf additive Fertigung	219
A.3	Attribute und statistische Kennzahlen.....	220
A.4	Rapidminer.....	226
A.5	Python, XGBoost, Code.....	228
A.6	SHAP.....	231
A.7	Betriebswirtschaftliche Aspekte.....	233
	Curriculum Vitae.....	237

Abbildungsverzeichnis

1.1:	Bauteil in der Prozesskammer eines metallbasierten additiven Fertigungssystems.....	2
1.2:	Nutzung additiv gefertigter Bauteile nach Branche.....	2
1.3:	Anzahl Hersteller additiven Fertigungsequipment.....	5
1.4:	Entwicklung der Investitions- und Betriebskosten additiver Fertigungssysteme	7
1.5:	CRISP-DM Ablauf und Zusammenhänge	15
1.6:	Aufbau der Arbeit unter Angabe der Kapitelnummern	16
2.1:	Einordnung Begriff – Additive Fertigung	17
2.2:	Gesamtübersicht additive Fertigungsverfahren	19
2.3:	Gesamtablauf additive Fertigung	20
2.4:	Prinzip PBF-Verfahren	24
2.5:	Übersicht Aufbau, Ablauf und Sensorik eines PBF-Fertigungssystem	30
3.1:	Ergebnisse Experten Einschätzung – Verteilung Fehlerursachen ungeplanter Unterbrechungen von PBF-Systemen.....	41
3.2:	Anteil an Baujobs mit einem ungeplanten vorzeitigem Ausfall in Prozent im Verhältnis zur ursprünglich geplanten Bauzeit.....	42
3.3:	Expertenbeurteilung: Einfluss des Maschinenbedieners auf die Vermeidung eines ungeplanten Ausfalls	42
3.4:	Kostenmodell für lasergenerierte Bauteile	44
3.5:	Typische Prozesskette PBF-Verfahren	45
3.6:	Übersicht Kosten eines Produktionsausfalls.....	46
3.7:	Kostenmodell für den ungeplanten Produktionsausfall von PBF-Systemen.....	47
3.8:	Kostenersparnis für verschiedene Zeitpunkte eines präventiven Eingriffs anhand eines Beispiels – Titan Baujob mit 48 Stunden geplanter Produktionszeit	54
4.1:	Anteil an Baujobs nach Maschinentyp und Material im Gesamtdatenbestand	58

4.2:	Beispiel einer Sensorprotokolldatei mit einer erklärenden Darstellung der Attribute und einem Auszug der aufgezeichneten Werte pro Sekunde	60
4.3:	Aufwärmphase Beispiel Zeitreihenverlauf Prozesskammer Temperatur und -Druck und Scannertemperatur.....	61
4.4:	Aufwärmphase Beispiel Zeitreihenverlauf Kollimator- und Scannertemperatur	62
4.5:	Produktionsphase Beispiel Zeitreihenverlauf Prozesskammer Temperatur und Filterzustand.....	62
4.6:	Produktionsphase Beispiel Zeitreihenverlauf Kollimator- und Scannertemperatur	63
4.7:	Verallgemeinerte Abfolge der Prozessschritte pro Bauschicht	63
4.8:	Vereinfachter Auszug aus einer Ereignisprotokolldatei.....	64
4.9:	Zeitreihe Ereignisprotokoll – Laserbelichtungs-/Abkühlzeit	65
4.10:	Unterschiede zwischen den Maschinentypen oder Basislegierung zur durchschnittlich erreichten Erfolgsquote (Tableau Analyse)	67
4.11:	Abweichung zur durchschnittlichen Erfolgsquote von einzelnen Material-Maschinen-Kombinationen (Tableau Analyse).....	67
4.12:	Abhängigkeit des Maschinenalters zur durchschnittlichen Erfolgsquote (Tableau Analyse)	68
4.13:	Korrelationsmatrix anhand einer beispielhaft ausgewählten Sensorprotokolldatei eines Baujobs	69
4.14:	Nachweis des Bravais-Pearson-Korrelationskoeffizienten-Verlaufs anhand des Vergleichs von jeweils zwei Zeitreihen aus der Sensorprotokolldatei eines Baujobs (Tableau Analyse).....	70
5.1:	Rohdatenaggregation aller drei textbasierten Dateien je Baujob durch Python.....	76
5.2:	Definition von Zeitintervallen zur Aufteilung der Rohdaten	77
5.3:	Gesamtübersicht Zusammensetzung der aggregierten Attribute aus den Rohdaten, Zeitintervallen und statistischen Kennzahlen.....	79
5.4:	Aufbau Bezeichnung Attribute.....	80
5.5:	Beispiel lokales, globales Maximum und Minimum anhand der Prozesskammer-Temperatur.....	81

5.6:	Darstellung Dichtefunktion auf Basis der Häufigkeit der Sensormesswerte der Prozesskammer Temperatur.....	82
5.7:	Zeitabstände mit Wertänderungen am Beispiel von Sauerstoff-, Plattformtemperaturverläufe.....	83
5.8:	Beispiele einiger Wendepunkte am Beispiel des Temperaturverlaufes in der Prozesskammer.....	84
5.9:	Allgemeines Beispiel einer rechtssteilen und linkssteilen Schiefe.....	85
5.10:	Beispiel Berechnung Interquartilabstand anhand des Prozesskammerdruck-Verlaufs.....	86
5.11:	Beispiel Ereignisprotokolldatei mit zeitlich abgrenzbaren Informationen zum Baujobverlauf.....	87
5.12:	Vorgehen Attributsbereinigung im Rahmen der Modellvorauswahl....	91
5.13:	Ergebnisse Attributsauswahl nach Anwendung des Korrelations- und Stabilitätsfilters.....	93
6.1:	Vorgehensweise Modellvorauswahl durch Attributsauswahl, Modellbildung und -vergleich mit anschließender Modelloptimierung.....	95
6.2:	Übersicht häufiger Aufgabentypen aufgeteilt nach Lernform.....	96
6.3:	Kaggle Umfrage – Nutzung Lernverfahren in der Praxis.....	97
6.4:	Beispiel Entscheidungsbaum mit binärer Klassifikation der Zielvariablen.....	104
6.5:	Ablauf Entscheidungsfindung Random Forest.....	107
6.6:	Iterativer Aufbau der Gradient Boosted Tree Berechnung.....	110
6.7:	Neuronales Netz mit vier Schichten. Der Informationsverarbeitung durch Neuronen erfolgt von links nach rechts.....	113
6.8:	SVM-Klassifikation nicht linear trennbarer Trainingsdaten mit Einführung von Schlupfvariablen.....	116
6.9:	ROC-Kurve, AUC-Fläche auf Basis der Modellergebnisse nach dem Gradient Boosted Tree Verfahren.....	121
6.10:	Eigendarstellung des Ablaufes einer Kreuzvalidierung am Beispiel einer $k = 5$ Aufteilung.....	123
6.11:	Modellvergleich ROC-Kurven auf der Datenbasis aller Attribute der ersten 40 Prozent der Produktionsphase.....	124
6.12:	Ablauf Vorgehen der Modelloptimierung mit XGBoost für verschiedene Zeitintervalle der Produktionsphase.....	125

6.13:	Auszug Gradient Boosted Tree – Beispiel Entscheidungspfad und Aufteilungspunkte (gen. Split Points)	128
6.14:	Verallgemeinerte Darstellung der Suche nach der idealen Modellkomplexität mit dem Problem der Über- und Unteranpassung	132
6.15:	XGBoost Ergebnisse AUC-Wert , Prognosegenauigkeit/ACC zur Überprüfung der optimalen Hyperparameterwerte für die Anzahl Bäume (n-estimators)	133
6.16:	XGBoost Ergebnisse AUC-Wert , Prognosegenauigkeit/ACC zur Überprüfung der optimalen Hyperparameterwerte für die Baumtiefe (depths).....	134
6.17:	XGBoost Ergebnisse AUC-Wert , Prognosegenauigkeit/ACC zur Überprüfung der optimalen Hyperparameterwerte für die Lernrate (learning_rate)	135
6.18:	Vorgehen Attributsauswahl mit XGBoost.....	136
6.19:	Ergebnisse der Schleifendurchläufe mit einer Reduzierung/Verdichtung auf weniger Attribute bei steigender Prognosegenauigkeit und AUC-Wert	138
7.1:	Gegenüberstellung Entwicklung der Prognosegüten von relativen und absoluten Zeitsegmenten des Produktionsfortschrittes	144
7.2:	Erklärung SHAP-Werte anhand eines Beispiels für Maschinenbaujob x	151
7.3:	SHAP-Summary-Plot mit den wichtigsten Attributen und deren Einfluss auf die Entscheidungsfindung des Gradient-Boosted-Tree-Verfahrens basierend auf den Daten der ersten 20 % der Produktionsphase.....	156
7.4:	Grafische Darstellung des kostenoptimalen Thresholds bei minimalen Kosten durch präventives Eingreifen	167
7.5:	Überblick Gesamtablauf der Berechnung der Kostenersparnisse und Schwellenwerte für verschiedene Szenarien	169
7.6:	Gegenüberstellung der Entwicklung von Kostenersparnissen zum Schwellenwert (Englisch: threshold) für verschiedene Szenarien	170
7.7:	Implikationen zur Entstehung von Wettbewerbsvorteilen durch Datenaustausch zwischen Marktteilnehmern der additiven Wertschöpfungskette (AM für Additive Manufacturing).....	175

7.8: Konkrete Implikationen für Geschäftsmodellentwicklung verschiedener Marktteilnehmer der additiven Wertschöpfungskette unter Einreihung einer konkreten Anwendung der Erkenntnisse dieser Forschungsarbeit	176
--	-----

Anhang

A.1.1: Zusammenhänge und Entwicklung der wichtigsten Data-Mining- und Data-Science-Modelle und -Methoden	215
A.2.1: Einflussfaktoren und -kategorien des Lasergenerierens	219
A.2.2: Prozesskette des additiven Fertigungsverfahrens Laserstrahlschmelzen/PBF	220
A.3.1: Auswahl Attribute von Untersuchungen des Fraunhofer IPK und der TU Berlin	220
A.4.1: Übersicht genutzte Lernverfahren-Operatoren und sonstige Operatoren in Rapidminer	226
A.4.2: Einfacher Beispielprozess in Rapidminer anhand eines GBT-Operators	226
A.4.3: Optimal Depth Operator mit eingebetteten Folgeoperatoren/Prozessen für die Bestimmung des Hyperparameters „Baumtiefe“ für alle Baumlernverfahren	226
A.4.4: Rapidminer Beispiel Gesamtmodell Generalized Linear Model mit Scoring-, Chartfunktionen	227
A.5.1: Funktionsparameter zur Belegung mit Trainings- und Testdaten für die XGBoost Berechnung der optimalen Hyperparameter	230
A.5.2: Weitere Parameter mit Default-Werten zur Berechnung der optimalen Hyperparameterwerte.....	230
A.6.1: SHAP-Summary-Plot mit den wichtigsten Attributen und deren Einfluss auf die Entscheidungsfindung des Gradient-Boosted-Tree-Verfahrens basierend auf den Daten der ersten 40 % der Produktionsphase	231
A.6.2: SHAP-Summary-Plot mit den wichtigsten Attributen und deren Einfluss auf die Entscheidungsfindung des Gradient-Boosted-Tree-Verfahrens basierend auf den Daten der ersten 180 Minuten der Produktionsphase	232
A.6.3: SHAP-Summary-Plot mit den wichtigsten Attributen und deren Einfluss auf die Entscheidungsfindung des Gradient-Boosted-Tree-Verfahrens	

basierend auf den Daten der ersten 360 Minuten der Produktionsphase	232
--	-----

Tabellenverzeichnis

1.1:	Vor- und Nachteile der additiven Fertigung	4
3.1:	Kriterien zur Berechnung der unkritischen- und kritischen Unterbrechungszeit	40
3.2:	Berechnung direkte Kosten eines ungeplanten Ausfalls.....	52
3.3:	Berechnung indirekte Kosten eines ungeplanten Ausfalls	52
3.4:	Definition eines geeigneten Aufgabenbereichs für die Lösung des Data- Analytic-Problems.	56
4.1:	Übersicht verwendete Metallpulver-Legierungen	58
4.2:	Anzahl Baujob vs. Anzahl Maschinenseriennummern.....	59
4.3:	Beschreibung prozessrelevante Attribute der Sensorprotokolldateien .	60
4.4:	Beschreibung zustandsrelevante Attribute aus der Sensorprotokolldateien	61
5.1:	Auflistung und Beschreibung aller zeitrelevanten Attribute aus der Ereignisprotokolldatei.....	88
5.2:	Indirekt berechenbare Attribute aus den Ereignisprotokolldaten	88
6.1:	Rapidminer Parameter Generalisierte Modelle.....	101
6.2:	Rapidminer Parameter Logistische Regression	103
6.3:	Rapidminer Parameter Entscheidungsbaum	105
6.4:	Rapidminer Parameter Random Forrest	108
6.5:	Rapidminer Parameter Gradient Boosted Trees	111
6.6:	Rapidminer Parameter Neuronale Netze	114
6.7:	Rapidminer Parameter Support Vector Machine.....	117
6.8:	Konfusionsmatrix anhand eines Beispiels mit Angaben zur Sensitivität und Spezifität	118
6.9:	Richtwerte zur Einordnung des AUC-Wertes als Prognosegüte	120
6.10:	Modellvorauswahl durch Vergleich der Prognosegüten auf der Datenbasis aller Attribute der ersten 40 Prozent der Produktionsphase	124
6.11:	Definition der Hyperparameterwerte für die optimale Auswahl durch XGBoost	130

6.12:	Übersicht aller durch XGBoost erzielten Prognoseergebnisse in Abhängigkeit zur jeweiligen Produktionsphase. Mit Auflistung der gewählten Hyperparameter und ausgewählten Anzahl an Attributen .	140
6.13:	Ergebnisse der Prognosegüten auf Basis einer Kreuzvalidierung mit dem Datenbestand der ersten 20 % und ersten 6 Stunden (360 Minuten) der Produktionsphase.....	141
7.1:	Axiome von Shapley-Werten	149
7.2:	Aufteilung Attribute in Gruppen und Rang der SHAP-Werte	157
7.3:	Beispielrechnung Kosten ungeplanter Abbruch in der letzten Sekunde versus Kosten geplanter Abbruch nach 3–6 Std.....	166
7.4:	Variablen und Berechnungsgrundlagen für die Kalkulation der Kostenersparnis durch präventives Eingreifen.....	166
7.5:	Szenarien mit unterschiedlichen Ausfallquoten. Normierung auf 100 Baujobs.....	167
Anhang		
A.3.1:	Gesamtübersicht Kombinationen und Anzahl an Attributen	221
A.3.2:	Bezeichnung und Übersetzung der Kernattribute, statistischen Kennzahlen und sonstigen Attribute	225
A.7.1:	Detaillierte Beispielberechnung der Kosten eines ungeplanten Ausfalls in der letzten Sekunde versus Kosten eines geplanten Abbruch nach 3–6 Std.	235
A.7.2:	Bestimmung des kostenminimalen Thresholds. Anhand des Beispiels: N = 15 (Ausfallquote), Normierung auf 100 Baujobs, Prognosegüte basierend auf Attributen der ersten 6 Std. der Produktionsdauer.....	236

Quellcodeverzeichnis

6.1:	Definition Hyperparameterwerte als Grundlage zur Berechnung einer optimalen Parameterwertekombination	131
6.2:	Bestimmung der optimalen Hyperparameter zur Erreichung einer maximalen Prognosegüte	132
6.3:	Schleifendurchlauf XGBoost Hyperparametersuche	139
6.4:	Kreuzvalidierung $k = 5$	141
7.1:	Umsetzung Python SHAP TreeExplainer und Summary Plots	155
Anhang		
A.5.1:	Auflistung Python – statistische Kennzahlen Teil 1	229
A.5.2:	Auflistung Python – statistische Kennzahlen Teil 2	229

Abkürzungsverzeichnis

3DP	3D-Print Material Verdüsung
ACC	Accuracy, Prognosegenauigkeit
AI	Artificial Intelligence
AJ	Aerosolverfahren
AM	Additive Manufacturing
ASTM	American Society for Testing and Materials
AUC	Area Under the Curve
CAD	Computer-Aided Design
CAE	Computer Aided Engineering
CPU	Central Processing Unit
CRISP-DM	Cross-industry Standard Process for Data Mining
CSV	Comma-separated values
DLP	Digital Light Processing
DMD	Direct Metal Deposition, Laserauftragsschweißen
DSGVO	Datenschutz-Grundverordnung
EBM	Electron Beam Melting, Elektronenstrahlschmelzen
ERP	Enterprise-Resource-Planning
FFF	Fused Filament Fabrication, Material Extrusionsverfahren
FMEA	Failure Mode and Effects Analysis
FN	False Negative
FP	False Positive
GBT	Gradient Boosted Trees
GLM	Generalisierte Lineare Modelle
HACE	Heterogeneous Autonomous Complex Evolving

IDAM	Industrialisierung und Digitalisierung von AM für automobiler Serienprozesse
ILAS	Institut für Laser- und Anlagensystemtechnik
IP	Intellectual Property
IPK	Fraunhofer-Institut für Produktionsanlagen und Konstruktionstechnik
IQA	Interquartilsabstand
IRLSM	Iteratively Reweighted Least Squares Method
ISO	International Organization for Standardization
KDD	Knowledge Discovery in Databases
KI	Künstliche Intelligenz
KIT	Karlsruher Institut für Technologie
KLR	Kernel Logistische Regression
KNN	Künstliche neuronale Netze
LASSO	Least Absolute Shrinkage and Selection Operator
L-BFGS	Limited-memory Broyden-Fletcher-Goldfarb-Shanno Algorithm
LCC	Life Cycle Costing
LCVD	Laser Chemical Vapor Deposition
LIBSVM	Library for Support Vector Machines
LIME	Local Interpretable Model-agnostic Explanations
LIMS	Labor-Information- und Management-System
MES	Manufacturing Execution System
MJM	Multi-Jet Modelling
MTDF	Meantime to detect Failure
OEE	Overall Equipment Efficiency
OEM	Original Equipment Manufacturer
PBF	PBF-LB/M. Laser-based powder bed fusion of metals
PCA	Principal Component Analysis
PJM	Poly-Jet Modelling

PwC	PricewaterhouseCoopers
QMS	Qualitätsmanagementsystem
RAMSYS	Rapid Collaborative Data Mining System
RBF	Radial Basis Function
ReLU	Rectifier Aktivierungsfunktion
ROC	Receiver Operating Characteristic
SEMMA	Sample, Explore, Modify, Model und Assess
SHAP	SHapley Additive exPlanation
SL	Stereolithografie
SLS	Selektives Lasersintern
SPS	Speicherprogrammierbare Steuerung
STL	Standard Triangle Language
SVM	Support Vector Machine
TCO	Total Cost of Ownership
TN	True Negative
TP	True Positive
VDI	Verein Deutscher Ingenieure
VDMA	Verband Deutscher Maschinen- und Anlagenbau
WAAM	Wire Arc Additive Manufacturing, Lichtbogen-Drahtschweißen