forschen.
vernetzen.
anwenden.

Innovationsreport 2024

Industrielle Gemeinschaftsforschung

IGF-Forschungsvorhaben 01IF22312N / 22312 N

Befähigung von kmU zur Nutzung von Potenzialen von Machine Learning in der Produktion und Entwicklung einer Einführungsstrategie (MIready)

Laufzeit:

01.03.2022 - 31.08.2024

Beteiligte Forschungsstelle(n):

IPRI - International Performance Research Institute gGmbH
Stuttgart

IPH - Institut für Integrierte Produktion Hannover gGmbH

Schlussbericht vom 08.01.2025

zu IGF-Vorhaben Nr. 01|F22312N

Thema

MLready - Befähigung von KMU zur Nutzung von Potenzialen von Machine Learning in der Produktion und Entwicklung einer Einführungsstrategie

Berichtszeitraum

01.04.2022 bis 31.08.2024

Forschungsvereinigung

Institut für Energie & Umwelt, Analytik & Technik

Forschungseinrichtung(en)

1: IPRI - International Performance Research Institute gGmbH
2: IPH - Institut für Integrierte Produktion Hannover

Gefördert durch:

Bundesministerium für Wirtschaft und Klimaschutz

aufgrund eines Beschlusses des Deutschen Bundestages

Autoren

Prof. Dr. Mischa Seiter
IPRI – International Performance Research Institute gGmbH

Prof. Dr.-Ing. habil. Peter Nyhuis
IPH – Institut für Integrierte Produktion Hannover gGmbH

Garlef Hupfer
M. Sc.
IPRI – International Performance Research Institute gGmbH

Manuel Savadogo
M. Sc.
IPH – Institut für Integrierte Produktion Hannover gGmbH

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

Zusammenfassung

Das Projekt „MLready“ verfolgte das Ziel, kleine und mittelständische Unternehmen (KMU) zur Nutzung von Machine Learning (ML) in der Produktion zu befähigen und eine Einführungsstrategie zu entwickeln. Ziel war es, den Unternehmen einen Zugang zu den Potenzialen von Machine Learning zu ermöglichen, insbesondere in Bezug auf die Steigerung der Ressourceneffizienz.

Trotz der offensichtlichen Vorteile von ML wird die Technologie in KMU bisher selten eingesetzt. Dies liegt vor allem an Herausforderungen wie unzureichender Datenqualität, mangelndem Wissen über ML und dem Fachkräftemangel im Bereich Data Science. Darüber hinaus fehlt es vielen KMU an einer klaren Strategie zur Einführung von ML, um den technologischen, organisatorischen und menschlichen Anforderungen gerecht zu werden.

Im Rahmen des Projekts wurde eine interaktive Prozesslandkarte erstellt, die Best Practices und potenzielle Anwendungsfälle von ML in der Produktion darstellt. Anwendungsgebiete wie Produktionsplanung, Qualitätsmanagement und Prozessoptimierung wurden identifiziert und visualisiert. Diese Landkarte bietet KMU einen Überblick über mögliche Einsatzgebiete von ML und enthält konkrete Beispiele aus der Praxis.

Eine der zentralen Herausforderungen bei der Einführung von ML ist die Verfügbarkeit und Qualität der Daten. Im Projekt wurde ein Regelwerk zur Messung der Datenqualität entwickelt, das auf Literaturrecherchen und Expertengesprächen basiert. Zudem wurde ein Leitfaden erstellt, der es KMU ermöglicht, relevante Daten zu identifizieren und aufzubereiten.

Um KMU bei der Einführung von ML zu unterstützen, wurde eine umfassende Einführungsstrategie entwickelt. Diese enthält Handlungsempfehlungen, eine Wirtschaftlichkeitsbewertung sowie Kriterien zur Auswahl von ML-Dienstleistern. Zudem wurde eine Roadmap erstellt, die die schrittweise Implementierung von ML in Unternehmen unterstützt.

In Zusammenarbeit mit verschiedenen Unternehmen wurden Fallstudien durchgeführt, um die Praxistauglichkeit der entwickelten Ansätze zu überprüfen. Dabei wurde gezeigt, dass durch den Einsatz von ML die Ressourceneffizienz in der Produktion signifikant gesteigert werden kann.

Das Projekt hat gezeigt, dass ML ein enormes Potenzial zur Steigerung der Ressourceneffizienz in KMU bietet. Die entwickelten Tools, wie die interaktive Prozesslandkarte und der Leitfaden zur Datenaufbereitung, bieten eine wertvolle Unterstützung für KMU, um ML erfolgreich in ihre Produktionsprozesse zu integrieren.

Das Projekt MLready leistet einen wesentlichen Beitrag zur Befähigung von KMU, ML zur Optimierung ihrer Prozesse zu nutzen. Die Ergebnisse und entwickelten Werkzeuge sind praxistauglich und bieten eine solide Grundlage für zukünftige Implementierungen von ML in der Produktion.

1. Wissenschaftlich-technische und wirtschaftliche Problemstellung

Die Einführung von Machine Learning ermöglicht Unternehmen, die Ressourceneffizienz in der Produktion zu steigern. Trotz der evidenten Potenziale werden Machine Learning-Anwendungen in der Produktion bisher kaum genutzt. Bei der Implementierung von Machine Learning-Anwendungen ergeben sich insbesondere für KMU große Herausforderungen (KI-Transfer BW 2022)Klicken Sie hier, um Text einzugeben..

Eine Voraussetzung für erfolgreiche Machine Learning-Anwendungen ist eine hinreichende Datenqualität. Daten bilden die Grundlage von Machine Learning. Je qualitativ hochwertiger die Daten sind, desto besser kann das Erlernte i.d.R. in die Anwendung überführt werden. Die Datenqualität ist jedoch trotz vieler softwaretechnischer Unterstützungsmöglichkeiten oftmals nicht unzureichend (Industry Analytics 2018). Dies spiegelt sich auch in der Selbsteinschätzung bei Unternehmen wieder, wo nur 47 % der Befragten mit der Datenqualität in der Produktion und Logistik zufrieden sind (Deloitte 2014). Damit einhergehend stellt die Erfassung relevanter Daten in der Produktion eine große Herausforderung vor allem für KMU (BMWi und wik 2019) dar. Oftmals ist unklar, wo und in welcher Form Daten entstehen. Eine Befragung der Experian Qas (2013) ergab, dass inkorrekte und veraltete Daten die häufigsten Datenqualitätsdefekte darstellen und zu hohen Budgetüberschreitungen, sinkender Kundenzufriedenheit sowie Kundenbeschwerden und -abwanderungen führen. Bspw. gaben hier zwei Drittel der befragten Organisationen an, signifikante Datenqualitätsprobleme im Customer Relationship Management sowie bei der Analyse beziehungsweise Verwertung von (großen) kundenbezogenen Datenmengen zu haben. Neben Volume, Velocity und Variety wird daher Veracity – also die Qualität der Daten – als vierte zentrale Dimension von Big Data bezeichnet (Turner et al. 2013). Aus diesem Grund müssen Daten eine Vorverarbeitung durchlaufen (IDG 2019). Dieses „Preprocessing“ sowie die Entwicklung und Verwendung der Algorithmen sowie die Interpretation der Ergebnisse erfordert weitreichende Kompetenzen. In der Praxis ist dieses Know-how oftmals nicht vorhanden (VDMA 2018). Zum einen gibt es einen gravierenden Fachkräftemangel insbesondere bei KMU im Bereich Data Science (Bitkom 2020; Winter 2021). Ein Grund hierfür ist, dass Mitarbeiter in diesem Bereich häufig nicht ausreichend weitergebildet werden (Fraunhofer 2018). Zum anderen ist auch unklar, welche Aufgaben bei der Durchführung eines solchen Projekts selbst durchgeführt werden können und welche durch externe Dienstleister realisiert werden sollten.

Eine weitere Herausforderung ergibt sich mit der Auswahl geeigneter Prozesse und Algorithmen (Deloitte 2014). Unternehmen möchten nicht nur auf Preisbasis die Entscheidung für Machine Learning im Unternehmen treffen, sondern vorab Machine Learning verstehen und den Nutzen abschätzen (VDMA 2018). Es ist meist nicht bekannt, in welcher Höhe die Algorithmen zur Steigerung der Ressourceneffizienz in der Produktion beitragen. Jedoch prognostiziert das BMWi,

dass sich die Produktivität bis 2035 durch den Einsatz von Machine Learning bis zu 37% steigert (PAiCE 2018).

Eine weitere Schwierigkeit bei der Auswahl liegt im mangelnden Anwendungsbezug der Forschung (PAiCE 2018). Zudem können sich Unternehmen kaum an Praxisbeispielen orientieren, da es schwierig ist, geeignete Use Cases zu identifizieren (Deloitte 2019). Unternehmen können sich auch nicht an einer einheitlichen Vorgehensweise zur Auswahl geeigneter Prozesse orientieren, da diese bis jetzt nicht wissenschaftlich aufbereitet wurde. Passende Machine Learning-Algorithmen, die mangelnde Nachvollziehbarkeit der Algorithmen, die große Auswahl an Anbietern sowie risikobehaftete Kosten-Nutzenentscheidungen erschweren für KMU die Anwendbarkeit (Bitkom 2017). Die Forschungsfrage des Vorhabens lautet daher:

Wie und mit welcher externen Unterstützung können KMU des Maschinenbaus Machine Learning-Anwendungen in der Produktion nutzen, um die Ressourceneffizienz zu steigern?

Aus der Forschungsfrage ergeben sich mehrere Teilfragen, die eine hohe Relevanz für die zukünftige Nutzung von Machine Learning-Anwendungen in der Produktion haben.

1. Was sind geeignete Anwendungsfälle für Machine Learning zur Steigerung der Ressourceneffizienz in der Produktion?
2. Welche Machine Learning-Algorithmen können für diese Anwendungsfälle verwendet werden?
3. Wie können KMU befähigt werden, die notwendigen Voraussetzungen wie die notwendigen Daten, in der richtigen Qualität und Menge für diese Machine Learning-Algorithmen zu identifizieren und zu erheben?
4. Wie muss ein KMU-gerechtes Machine Learning-Einführungskonzept ausgestaltet sein, das den technologischen, datenbezogenen, organisatorischen sowie menschlichen Anforderungen gerecht wird?
5. Welche Aufgaben können KMU bei der Machine Learning-Implementierung selbst durchführen und welche Aufgaben sollten extern unterstützt werden?

2. Durchgeführte Arbeiten und Ergebnisse im Berichtszeitraum

Auf den folgenden Seiten werden die durchgeführten Arbeitsschritte und die erzielten Ergebnisse dargelegt. Die Ausführungen sind entlang der geplanten Arbeitspakete (AP) strukturiert. Die Kernergebnisse können außerdem auf der Projekthomepage unter https://ipri-institute.com/forschungsprojekte/mlready/ eingesehen werden.

2.1 Arbeitspaket 1: Entwicklung der MLready-Prozesslandkarte Identifikation der Anwendungsfälle, sowie deren Anforderungen und Potentiale

AP 1: ML*ready*-Prozesslandkarte (Identifikation der Anwendungsfälle, deren Anforderungen sowie Potenziale)	
Geplante Arbeiten	**Durchgeführte Arbeiten**
Es werden Best Practices auf Basis einer Literaturrecherche zusammengestellt. Darüber hinaus werden alle Prozessschritte (Ist-Prozess) im Rahmen des Hannoveraner Lieferkettenmodells (HaLiMo, s. https://halimo.education), welche sich direkt auf die Produktion beziehen (Produktionsvorstufe, Zwischenlager, Produktionsendstufe), Methoden der Prozessoptimierung (Reengineering und Kaizen) und der DIN-Norm 9000 für das Qualitätsmanagementwerden auf ihr Verbesserungspotenzial hinsichtlich der Ressourceneffizienz durch den Einsatz von Machine Learning geprüft. Dazu wird eine Prozessanalyse mit Tauglichkeitsbewertung durchgeführt, welche über die Aufarbeitung bisher bekannter Use Cases hinausgeht. Darauf aufbauend wird das Potenzial der Ressourceneinsparungen durch Machine Learning bei mindestens drei Unternehmen (Zusagen von: Hako GmbH, Lauscher GmbH, Hosti GmbH) des PA vor Ort erhoben. Anschließend wird mit den Unternehmen des PA	Es wurden Best Practices auf Basis einer Literaturrecherche zusammengestellt. Darüber hinaus wurden alle Prozessschritte (Ist-Prozess) im Rahmen des **Hannoveraner Lieferkettenmodells** (HaLiMo, s. https://halimo.education), welche sich direkt auf die Produktion beziehen (Produktionsvorstufe, Zwischenlager, Produktionsendstufe), Methoden der Prozessoptimierung (Reengineering und Kaizen) und der DIN-Norm 9000 für das Qualitätsmanagementwerden auf ihr Verbesserungspotenzial hinsichtlich der Ressourceneffizienz durch den Einsatz von Machine Learning geprüft. Dazu wurde eine Prozessanalyse mit Tauglichkeitsbewertung durchgeführt, welche über die Aufarbeitung bisher bekannter Use Cases hinausgeht. Darauf aufbauend wurde das Potenzial der Ressourceneinsparungen durch Machine Learning bei zwei Unternehmen des PA erhoben. Darüber hinaus wurden Ergebnisse der Literaturrecherche verwendet, in denen die Ressourceneinsparung dokumentiert wurde,

ein Workshop zur Identifikation nutzenstiftender Einsatzszenarien von Machine Learning zur Ressourceneffizienz durchgeführt und abschließend mit den bisherigen Ergebnissen abgeglichen.	um weitere beispielhafte Untersuchungen des Potenzials zu dokumentieren. Anschließend wurde mit den Unternehmen des PA sowie im Rahmen einer Vorlesung an der Universität Ulm ein Workshop zur Identifikation nutzenstiftender Einsatzszenarien von Machine Learning zur Ressourceneffizienz durchgeführt und abschließend mit den bisherigen Ergebnissen abgeglichen.
Aktuell verwendete Machine Learning-Ansätze werden auf Grundlage von Recherchen und Expertengesprächen übersichtlich in Form einer Prozesslandkarte zusammengestellt. Diese zeigt Anwendungsfälle in der Produktion in Form von Steckbriefen (Anzahl Stichproben, Einflussparameter, Prozessschritt, Ressourceneinsatz, Machine Learning-Ansatz, etc.). Darüber hinaus wird die Machine Learning Identifikationslogik in einem Leitfaden KMU-gerecht aufbereitet.	Die in der Literaturrecherche und in Expertengesprächen ermittelten Machine Learning-Ansätzen wurden übersichtlich in Form einer Prozesslandkarte dargestellt. Hierzu wurde im ersten Schritt die identifizierten Machine Learning-Ansätzen in die drei Bereiche Produktionsplanung, -optimierung und -überwachung und Qualitätsmanagement unterteilt. Anschließend wurden die einzelnen Anwendungsbeispiele passend zu den Bereichen in Form von Steckbriefen abgebildet. Die Steckbriefe enthalten neben einer kurzen Beschreibung auch eine Übersicht über passenden Machine Learning-Algorithmen, Ziele und erfolgreiche Beispiele. Zur Visualisierung wurde hierbei mit dem Programm Prezi gearbeitet. Mit Hilfe dieses Programms wurde die interaktive Prozesslandkarte mit einer nachvollziehbaren und logisch aufgebauten Struktur erstellt. KMU können sich somit durch die interaktive Prozesslandkarte klicken und sich über die für sie interessante Machine Learning-Ansätze und passende Beispiele informieren. Hiermit kann eine KMU-gerechte Identifikationslogik gewährleistet werden.

Ziel des ersten Arbeitspakets war die Auswahl von Anwendungsgebieten von Machine Learning zur Steigerung der Ressourceneffizienz in der Produktion basierend auf einer Potenzial-Anforderungsanalyse. Hierfür wurden mit einer Literaturrecherche Best Practices zusammengestellt. Weitere Quellen wurden mittels einer Prozessanalyse mittels einer Tauglichkeitsbewertung hinsichtlich ihres Verbesserungspotentials bezüglich der Ressourceneffizienz durch den Einsatz von ML geprüft. Außerdem wurde das Potential der Ressourceneinsparung durch ML bei Unternehmen vor Ort erhoben, bevor mit den Unternehmen des PA, ein Workshop zur Identifikation nutzenstiftender Einsatzszenarien von Machine Learning zur Ressourceneffizienz durchgeführt und mit den bisherigen Ergebnissen abgeglichen wurde. Die Machine Learning-Ansätze wurden auf Grundlage von Recherchen und Expertengesprächen in Form einer Prozesslandkarte zusammengestellt und die Machine Learning Identifikationslogik in einem Leitfaden KMU-gerecht aufbereitet.

2.1.1 Literaturrecherche

Bei der durchgeführten Literaturrecherche wurde gezielt nach Anwendungsfällen von ML in der Produktion und zur Steigerung der Ressourceneffizienz gesucht. Inspiriert durch das Literaturreview von Waltersmann et al. (2021) wurde für die Recherche auf die PRISMA Methode (Page et al. 2021) zurückgegriffen. Die verwendeten Recherchebereiche von Waltersmann et al. wurden abgeändert, um auf die Spezifika von MLready einzugehen. Die Literaturrecherche wurde im Zeitraum von Oktober 2022 bis November 2022 (13.10.22-24.11.2022) durchgeführt. Mit folgendem Suchterm wurde auf der Datenbank Google Scholar gesucht:
(“Machine Learning” OR “Support Vector Machines” OR “Decision Trees” OR “Naive Bayes Classification” OR “K-means” OR “Hierarchical Clustering” OR “Principal Component Analysis” OR “Isolation Forest” OR “Local Outlier Factor” OR “Convolutional Neural Network” OR “Pattern Recognition” OR “Recurrent Neural Networks” OR “Long short-term Memory” OR “Transformer” OR “Markov chain” OR “ State–action–reward– state–action” OR “ Deep Q-Network”, “ Double Deep Q Network” OR “Q-Learning”) AND (“resource efficiency” OR “material efficiency” OR “energy efficiency” OR “water efficiency”) AND abstract:”manufacturing”

Mit dem Suchterm wurden in Google Scholar 108 Ergebnisse gefunden. Auf 3 Artikel war der Zugriff eingeschränkt bzw. nicht möglich, 7 weitere Paper stellten sich als Duplikate heraus. 98 Paper wurden deshalb registriert. Nachdem die Abstracts der gefundenen Paper gescreent wurden, wurden 38 Beiträge aufgrund mangelnder Relevanz für die Forschung ausgeschlossen. Die 50 verbleibenden Paper wurden genauer betrachtet und weitere 12 ausgeschlossen, da sie nur konzeptuelle Inhalte diskutierten und keine genauen ML-Anwendung enthalten. Weitere 8 Paper wurden nach genauem Lesen als irrelevant eingestuft. 30 Paper wurden somit aus der

systematischen Literaturrecherche aufgenommen und als Grundlage für die Identifizierung von ML-Anwendungen verwendet.

Des Weiteren wurden beim PA-Workshop, weitere ML-Anwendungsfelder identifiziert, und für diese nach Literatur gesucht. Auf diese Weise wurden 10 Internetseiten und 30 Paper von Google Scholar aufgenommen. Abbildung 1 zeigt die Suchergebnisse nach PRISMA-Schema.

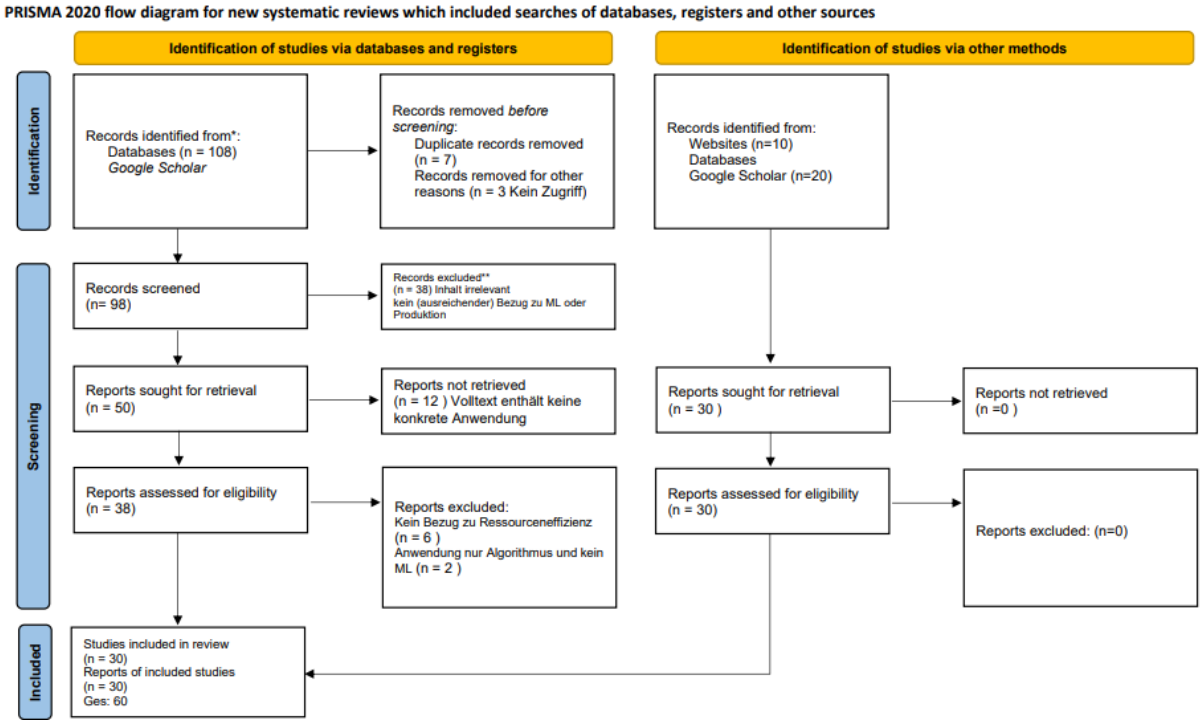


Abbildung 1: Systematische Literaturrecherche nach PRISMA (Page et al. 2021)

2.1.2 Tauglichkeitsbewertung

In einem weiteren Schritt wurden zusätzliche Quellen auf ihr Verbesserungspotenzial hinsichtlich der Ressourceneffizienz durch den Einsatz von Machine Learning geprüft. Dazu wurde eine Prozessanalyse mit Tauglichkeitsbewertung durchgeführt, welche über die Aufarbeitung bisher bekannter Use Cases hinausgeht. Für das Anwendungsgebiet der Prozessoptimierung wurden u. a. die Methoden Kaizen und Reengineering untersucht. Erstere beschreibt eine Methodik, bei der mit Fokus auf den Mitarbeitenden, Prozesse schrittweise verbessert werden, indem Verschwendung, Ungleichmäßigkeiten und Unzweckmäßigkeiten beseitigt werden (Hofmann 2020). Beim Reengineering steht die komplette Neugestaltung von Prozessen im Fokus. Dabei werden die Kerntätigkeiten eines Prozesses unter Beachtung der technologischen und infrastrukturellen Neuerungen innerhalb des Unternehmens optimal neu eingebettet (Hofmann 2020). Diese beiden sowie weitere Methoden der Prozessoptimierung (Six-Sigma, Lean Management, Total Quality

Management etc.) bieten keine direkten Verbesserungsmöglichkeiten durch Machine Learning. Stattdessen können diese Methoden verwendet werden, um Prozesse neu zu evaluieren und mit Machine Learning zu verbessern (Al-Anqoudi et al. 2021). Die Methoden wurden angewandt, um innerhalb der anderen zusätzlichen Quellen (HaLiMo und DIN 9000) neue Anwendungsmöglichkeiten zu finden.

Bei der Tauglichkeitsbewertung des HaLiMo wurden folgende Stellen identifiziert, die sich potenziell durch den Einsatz von ML optimieren lassen:

Tabelle 1: ML-Anwendung aus dem HaLiMo

Anwendungsgebiet im HaLiMo	ML-Anwendung
Auftragsmanagement	
Auftragserklärung	OCR
Grobterminierung der Produktionsaufträge und Sicherheitszeitplanung	Process Mining
Produktionsaufträge realisierbar?	Process Mining, ML-basierte Kapaplanung
finale Wandlung der Kunden- in Produktionsaufträge	ML-basierte Umwandlung der Aufträge, z. B. RPA-basiert
Auftragskoordinierung	Process Mining
Sekundärbedarfsplanung	
Sekundärbedarfsermittlung und Vorlaufverschiebung	Process Mining für stochastische Bedarfsrechnung
Produktionsbedarfsplanung	
Mittelfristige Ressourcengrobplanung	Process Mining
Eigenfertigungsplanung	
Losgrößenrechnung	verschiedene Ansätze der diskreten Optimierung
Durchlaufterminierung	Process Mining
Eigenfertigungssteuerung	
Reihenfolgebildung	Evolutionäres maschinelles Lernen zur dynamischen Anpassung

Produktionsprogrammplanung	
Absatzplanung	Machine Learning zur Absatzprognose
Brutto-Primärbedarfsplanung	Machine Learning zur Absatzprognose
Netto-Primärbedarfsplanung	Machine Learning zur Absatzprognose
Bestandsmanagement und Produktionscontrolling	
Produktionscontrolling	Process Mining Fit-Analysen zur Erkennung von Abweichungen

Bei der Untersuchung der DIN ISO 9000ff. wurden Verbesserungspotentiale durch den Einsatz von ML ermittelt. Diese ergeben sich aus der ISO 9001:2000, wobei die folgenden unterstützt werden können:

- der Aspekt der Infrastruktur als Teil des Managements von Ressourcen,
- die Planung der Realisierungsprozesse, der Beschaffungsprozess sowie die Lenkung der Produktion als Teil der Produktrealisierung, und
- die Datenanalyse als Teil der Messung, Analyse und Verbesserung

Die Elemente des HaLiMos, der DIN EN ISO 9000 und der Methoden der Prozessoptimierung, die als tauglich hinsichtlich des Einsatzes von Machine Learning bewertet wurden, wurden der Liste aller Anwendungsgebiete hinzugefügt.

2.1.3 Ergebnisse

Die Anwendungsgebiete wurden sowohl im Rahmen einer Vorlesung bei der School of Advanced Sciences an der Universität Ulm sowie in einem Workshop während des ersten Treffens mit dem projektbegleitenden Ausschuss validiert. In Abbildung 2 sind die identifizierten Anwendungsfelder übersichtlich dargestellt. Dabei wurden Bezeichnungen teilweise angepasst und die Liste validiert. Anschließend werden die einzelnen ML-Anwendungen in Form von Steckbriefen vorgestellt und mit Beispielen aus der Literaturrecherche abgebildet. In den vorgestellten Steckbriefen sind bereits Ergebnisse aus Arbeitspaket 2 enthalten, in welchem Daten aufbereitet und den Anwendungsfällen zugeordnet wurden. Die ML-Anwendungen sind den drei wesentlichen Produktionsaufgaben (Anwendungsbereiche) zugeteilt, um eine übersichtliche Zusammenstellung zu gewährleisten: Produktionsplanung, Produktionsüberwachung und Qualitätsmanagement & Prozessoptimierung.

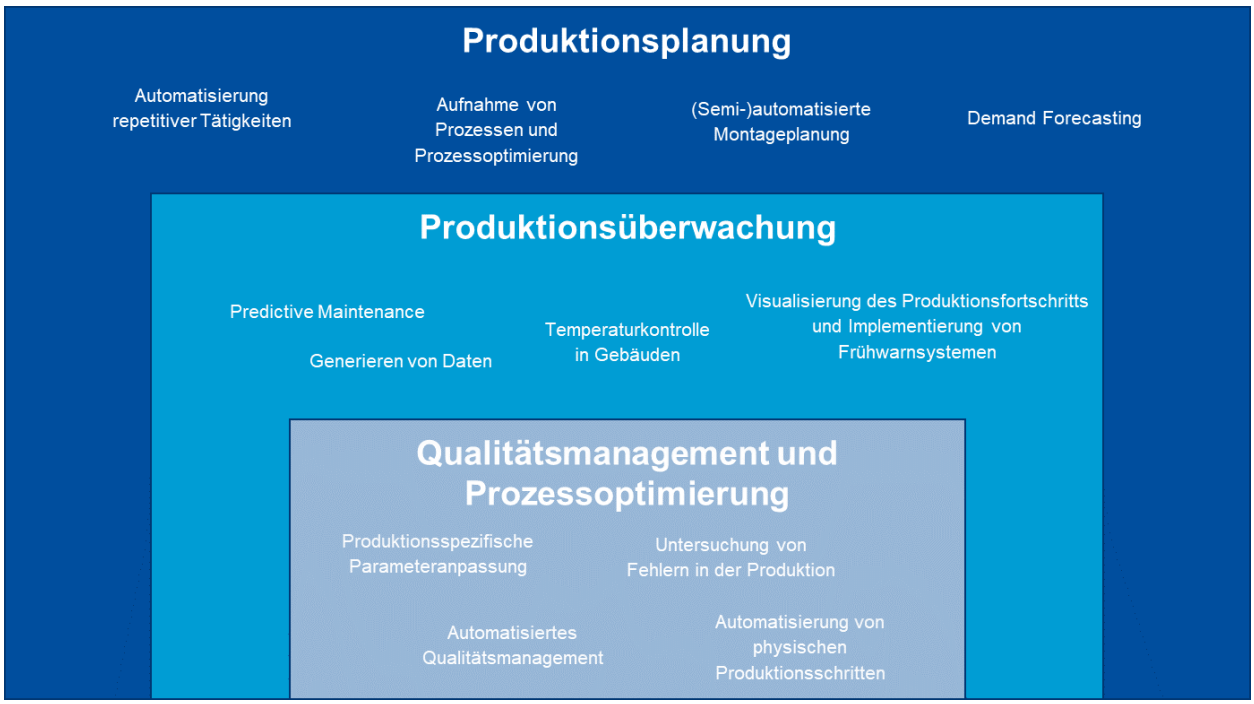


Abbildung 2: Übersicht der Anwendungsgebiete

Alle Anwendungsgebiete können detailliert unter https://prezi.com/view/6074HVoARB98SFgSAxdA/ eingesehen werden.

2.1.4 Potential der Ressourceneinsparung

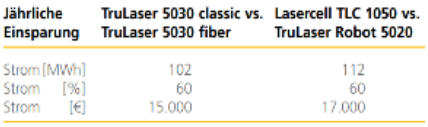
Zur Erhebung des Potenzials der Ressourceneinsparungen wurde zunächst die erfasste Literatur durchsucht, in denen für einige Anwendungsgebiete die Ressourceneinsparung durch die Umsetzung dokumentiert wurde. Das indonesische Unternehmen PT X setzt bspw. Robotic Process Automation für Lieferantenzahlungsprozesse ein. Die Durchlaufzeit der Prozesse kann dadurch bis zu 97,8% und die Personalkosten um bis zu 55% gesenkt werden (Arnaz und Harahap 2020).

Kocsi et al stellen in ihrem wissenschaftlichen Artikel eine Echtzeit Entscheidungsunterstützung für die Produktionsplanung vor, die mit Robotic Process Automation umgesetzt wird. Die Umplanungszeit kann bis zu 50% reduziert, und die Genauigkeit bis zu 90% erhöht werden (Kocsi et al. 2020). Darüber hinaus wurden die Fallbeispiele des Projektes 100 Betriebe für Ressourceneffizienz der Umwelttechnik BW GmbH gesichtet, und relevante Beispiele bezüglich des Potentials der Ressourceneinsparung gesammelt. Die Ergebnisse sind in Tabelle 2 abgebildet und der Inhalt nach Potential und Methodik der Messung unterteilt.

Beim 1. Projekttreffen wurde außerdem im gemeinsamen Austausch deutlich, dass sich das Potential der Ressourceneinsparung für bestimmte ML-Anwendungen nicht pauschalisieren lässt. Die erzielte Ressourceneinsparung unterliegt zu vielen Faktoren, die schlecht bis gar nicht abgewogen werden können, darunter insbesondere die Genauigkeit des angewendeten Machine Learning Modells. Um zu ermitteln, wie viel Ressourcen eingespart werden, eignen sich Erhebungen vor Ort, bei denen die Produktion über einen längeren Zeitraum beobachtet wird und die Aufzeichnungen mit denen, vor Einsatz der ML-Anwendungen verglichen wird. A priori können keine genauen Schätzungen des Potenzials erhoben werden. Daher werden die Erhebungen des Potenzials mit der Umsetzung der Fallstudien in Arbeitspaket 4 präzisiert, um genaue Zahlen zu den Verbesserungsmöglichkeiten angeben zu können. Im Folgenden werden die Ergebnisse der Recherche präsentiert, die durch bereits umgesetzte Methoden des Machine Learning genaue Zahlen berechnen konnten.

Tabelle 2: Beispiele von Machine Learning zur Steigerung der Ressourceneffizienz in der Praxis

Unternehmen & Feld	Potential	Mess-Methodik
Ulrich GmbH & Co.KG Spezialisiert auf die Entwicklung und Herstellung von Implantaten Verfahren: Zerspanung, Fräßen	Vorsimulation des Zerspanungsprozesses mit anschließender Regelung des Vorschubs: Mittels eines **cyber-physischen Systems** werden die Kräfte gemessen, die während der Bearbeitung auf den Fräser wirken und die Prozessparameter der Werkzeugmaschine werden dann entsprechend optimal angepasst ➔ Verringerung des Werkzeugverschleiß um 80% ➔ Monetäre Einsparungen ➔ Einsparung bei der Herstellungszeit	Messung der Herstellungszeit, diese ist je nach Verschleiß des Fräsers pro Implantat unterschiedlich hoch. Beim repräsentativen Musterteil veränderte sich die Herstellzeit von 74 auf durchschnittlich 67 Sekunden. Prozessfähigkeitsuntersuchung→ hier wurde festgestellt, dass die Standzeiten zwischen einzelnen Produktionsschritten mit dem neuen System deutlich höher waren, dadurch wird der Werkzeugverschleiß verringert. - 20h Spaneingriffszeit je Zahn ohne Auswirkungen auf statistische Werte - Zuvor: Nach 2h bereits erster Verschleiß sichtbar. Berechnung der monetären Einsparungen durch die Senkung des Energieverbrauchs und der Senkung der benötigten

		Materialien durch die Verringerung des Werkzeugverschleißes. Je nach Einsatzgebiet lasen sich bis zu 40kg Hartmetall pro Jahr einsparen
Roland Deeg GmbH (Energiewelt-info GmbH) (TRUMPF GmbH + Co. KG) Die Roland Deeg GmbH ist auf Blechbearbeitung spezialisiert und erhält Aufträge aus dem Maschinen- und Anlagebau sowie der Automobilindustrie. Verfahren: Laserschneiden und Laserschweißen	Anschaffung von effizienteren durch ML gestützten Laserbearbeitungsmaschinen (→ TruLaser 5030 von TRUMPF). Die **Robotersysteme** sind **vollautomatisch** und minimieren die Fehlerquote bei Bedienfehlern. Außerdem sind sie mit dem Programm teach line ausgestattet, wodurch die laufenden Prozesse überprüft werden. ➔ Permanente Plausibilisierung der Fehlerquote, sie liegt bei 1 bis max. 1,5% (zuvor war es 2 bis ca. 2,5%) ➔ Materialeinsparung von Stahl- und Edelstahlblech. ➔ Produktivitätssteigerung durch TruLaser 530: 460t mehr Stahl können pro Jahr bearbeitet werden → Steigerung von 40% ➔ Produktivitätssteigerung durch TruLaser Robot 5020: 840t mehr Stahl können pro Jahr bearbeitet werde → Steigerung von 70%	Messung der Fehler und der Ausschussquote, dabei Vergleich zwischen den alten und den neuen Robotern  **Jährliche Einsparung** / **TruLaser 5030 classic vs. TruLaser 5030 fiber** / **Lasercell TLC 1050 vs. TruLaser Robot 5020** Strom [MWh] 102 112 Strom [%] 60 60 Strom [€] 15.000 17.000
Pfizer Manufacturing Deutschland GmbH Pharmakonzern Verfahren: Kontinuierliches Herstellungs-verfahren	Entstehung einer neuen Fabrik. Mittels In-line-Mischung, **Echtzeitanalytik** und Rückkopplung zur Fördertechnik wird die Dosiergenauigkeit gesichert und die Rohstoffmischung wird kontinuierlich auf	• Messung der Durchlaufzeit • Wegfallen von kostenaufwändigen Schritten (traditionelle Kontrollanalyse fällt weg) • Ermittlung der CO2 – Bilanz

	Ihre Zusammensetzung und Qualität geprüft. **Online-Qualitätssicherung** mit Hilfe der Prozess-Analyse-Technik (PAT) ➔ Optimierung der Logistikkette, Dadurch Reduzierung der Fracht und deshalb Umstellung von Luft- auf Seefracht Reduzierung des CO2-Ausstoßes um 33% ➔ Verkürzung der Durchlaufzeit bei der Rohstoffmischung ➔ Verringerung des Risikos des Materialverlustes ➔ Verringerung der Abfallmengen	
SATEMA – Corporate Fashion GmbH Textil und Fashionbranche Verfahren: Veredelung und Konfektion von Textilien	Entwicklung eines **mobilen, selbstlernenden Moduls,** das als Plattform zwischen Planungs-, Produktions- und Steuerungssystem steht. Das Modul wird mit aktuellen Prozessdaten gespeist. ➔ Reduzierung der Fehler- und Ausschuss Quote. Dadurch kann rund 1t an Textilentsorgungen und deren Verpackungen vermieden werden. ➔ Sechsstelliger € Betrag an monetären Einsparungen pro Jahr, errechnet sich aus Materialeinsparung, Fehlerminimierung und Effizienzsteigerung.	Das Einsparpotenzial berechnet sich aus den Posten Materialeinsparung, Fehlerminimierung und Effizienzsteigerung. Die Effizienzsteigerung wiederum resultiert aus mehr Leistungskapazität pro Mitarbeiter.
NOVAPAX Maschinenbau GmbH & Co. KG	Automatisierte Maschinendatenerfassung in Verbindung mit der Optimierung von Produktionsprozessen durch Maschinelles Lernen (Random Forest Algorithmus, künstliche neuronale Netze)	Zur Visualisierung von Verbrauchsdaten bildet man bei NOVAPAX, Energie- und Materialströme in Form von **Sankey-Diagrammen** ab. Bei NOVAPAX hat die Auswertung über KI-Algorithmen in Modellrechnungen ergeben, dass durch Änderung einzelner

	➔ Durch das künstliche neuronale Netz können Handlungsempfehlungen für die betrachteten Maschinen abgeleitet werden, auf diese Weise wurden **theoretische Energieeinsparung von rund 5.000 kWh/Jahr** identifiziert. ➔ Außerdem: Abfallminimierung und Reduzierung des CO2-Ausstoßes	Maschinenparameter theoretische Einsparpotenziale von durchschnittlich bis zu sieben Tonnen CO2 pro Maschine im Jahr möglich sein könnten. Bei 70 Spritzgussmaschinen wären das also rund 500 Tonnen CO2 oder gut 20 Prozent des bisherigen Energieeinsatzes.
Battery LabFactory Braunschweig (BLB)	Ermittlung von Verbesserungspotenzialen auf Grundlage von maschinen- und prozessspezifischen Einflussfaktoren für die Herstellung von Batteriezellen mithilfe von ML. (artifical neural networks & Random Forest) ➔ ANN Modell:	Aus Basis des artificial neural network Modells, beispielhaften Produktionsbedingungen und Faktorvariationen von 5% wurden mögliche CO2-Einsparungen und Kostensenkungen berechnet. Die Energieeinsparungen pro Einflussfaktor betragen bis zu ca. 9%. Die aggregierten Einsparpotenziale von über 30 % sind vielversprechend, können aber in der Realität aufgrund von Abhängigkeiten zwischen den Faktoren nicht vollständig erreicht werden. Mit Einsparpotenzialen im Bereich von 10 bis 30 % (bei Inputschwankungen von nur 5 %) liegen die Ergebnisse jedoch in einer attraktiven Größenordnung.

2.1.5 Integration der Anwendungsgebiete in eine Prozesslandkarte

Zur Integration der Anwendungsgebiete in eine interaktive Prozesslandkarte wurde mit dem Programm Prezi gearbeitet. Die ermittelten Anwendungsgebiete wurden so dargestellt, dass eine übersichtliche und nachvollziehbare Orientierung gewährleistet werden kann (siehe Abbildung 3).

Anschließend wurden die ermittelten Machine Learning-Anwendungen den ermittelten Bereichen zugeordnet. Somit haben KMU die Möglichkeit, sich passend zu den einzelnen Bereichen mögliche Anwendungsbeispiele anzuschauen, um somit einen vertiefenden Einblick und relevante Informationen zu erhalten. Wird einer der einzelne Bereich angeklickt, so öffnet sich eine neue

Oberfläche. Diese bietet eine Übersicht über alle Machine Learning-Ansätze, die dem gewählten Bereich zugeordnet werden können (siehe Abbildung 4).

Abbildung 3: Darstellung der betrachteten Bereiche in der Prozesslandkarte

Abbildung 4: Detaillierte Darstellung der einzelnen Anwendungen

Zunächst wird der ausgewählte Bereich kurz und prägnant beschrieben. Anschließend besteht die Möglichkeit, einen der aufgelisteten Machine Learning-Ansätzen auszuwählen. Wird einer der Machine Learning-Ansätze ausgewählt, öffnet sich wieder eine neue Ansicht, in der der ausgewählte Ansatz näher erläutert. Darüber hinaus werden die passenden Arten von Machine Learning erwähnt. Um KMU einen noch tieferen Einblick in die ausgewählte Machine Learning-Ansätze geben zu können, wurden zusätzlich erfolgreiche Anwendungsbeispiele und Ziele & Nutzen miteingebaut. Somit werden sind die Ansätze für KMU greifbar und besser ersichtlich (siehe Abbildung 5).

Abbildung 5: Beschreibung der Automatisierung repetitiver Tätigkeiten

Im nächsten Schritt des Forschungsprojekt erfolgt die Erstellung eines Leitfadens zur Bewertung der Datenqualität. Die Ergebnisse sollen in die interaktive Prozesslandkarte eingepflegt werden. Um diesen Schritt transparent zu machen, wurde in der Prozesslandkarte bereits der Punkt „notwendige Daten“ eingepflegt, was als Einleitung für die Datenanalyse und den Leitfaden gilt.

2.2 Arbeitspaket 2: Identifikation und Bewertung der erforderlichen Datenquellen

AP 2: Identifikation und Bewertung der erforderlichen Datenquellen	
Geplante Arbeiten	**Durchgeführte Arbeiten**

Basierend auf den Ergebnissen aus AP1 werden mit Hilfe von Experteninterviews sowie Erfahrungen von Einzelfallstudien, die das IPH im Rahmen des „Mittelstand 4.0 Kompetenz-zentrums Hannover“ sowie „Mittelstand Digital Zentrum Hannover“ sammeln konnte, Anforderungen für einen umfassenden Einsatz von Machine Learning evaluiert und anschließend werden Maßnahmen zur Umsetzung (Migrationsschritte zwischen Ist- und Soll-Zustand) in KMU erarbeitet. Im ersten Schritt wird mit den Fallstudienpartnern eine RASCI-Matrix (Responsible-Accountable-Sup-port-Consulted-Informed) zur Zuordnung von Verantwortlichkeiten und Rollen von internen und externen Akteuren entwickelt.	Im Rahmen des Projekts wurden bekannte Bewertungen zur Datenqualität (vgl. z. B. Ishwarappa und Anuradha 2015; Tayi und Ballou 1998; Taleb et al. 2015) analysiert und auf den zugrundeliegenden Fall angewendet. Darüber hinaus wurde eine breitgefächerte Literaturrecherche durchgeführt, um weitere Quellen sowie den aktuellen Stand der Forschung zu Datenqualität aufzunehmen. Damit wurde ein Regelwerk für die Messung der Datenqualität erstellt. Das Ergebnis wurde in Expertengesprächen mit Machine Learning-Dienstleistern sowie Forschern im Bereich der Datenqualität validiert. Zusätzlich wurde das Modell dem projektbegleitenden Ausschuss vorgestellt und dort diskutiert, um die Praxistauglichkeit und Anwendbarkeit sicherzustellen.
Alle verfügbaren qualitativen und quantitativen Unternehmensdaten werden durch Fokusgruppendiskussionen identifiziert, auf ihre Datenqualität untersucht und eingeordnet. Alle relevanten Merkmale und Charakteristika werden in einem morphologischen Kasten zusammengefasst (morphologische Analyse und Fokusgruppendiskussionen). Durch die Analyse wird ein systematischer Überblick über die vorhandene Unternehmensdatenstruktur ermöglicht. Die Ergebnisse werden in die Prozesslandkarte überführt.	Alle verfügbaren qualitativen und quantitativen Unternehmensdaten wurden durch eine eingehende Literaturrecherche und anschließenden Fokusgruppendiskussionen identifiziert, auf ihre Datenqualität untersucht und eingeordnet. Eine umfängliche Tabelle zur Zuordnung der jeweiligen Datensätze zu den in AP1 identifizierten Anwendungsgebieten wurde erstellt. Alle relevanten Merkmale und Charakteristika wurden in einem morphologischen Kasten zusammengefasst (morphologische Analyse und Fokusgruppendiskussionen). Durch die Analyse wurde ein systematischer Überblick über die vorhandene Unternehmensdatenstruktur ermöglicht. Die Ergebnisse wurden in die Prozesslandkarte überführt.

Im dritten Schritt liegt ein besonderer Fokus auf der Gewährleistung einer ausreichenden Datenqualität, indem typische Qualitätsmängel ausgewählter Datentypen identifiziert und Verbesserungsmöglichkeiten gemeinsam mit dem PA erarbeitet werden. Anschließend wird im Rahmen von Fokusgruppendiskussionen ein Leitfaden zur eigenständigen Durchführung des Datenanalyseprozesses für KMU entwickelt, von der Datenerhebung bis hin zur Datenanalyse und -evaluation sowie Datenvisualisierung.	Im dritten Schritt lag ein besonderer Fokus auf der Gewährleistung einer ausreichenden Datenqualität. Typische Qualitätsmängel wurden basierend auf den in den Fallstudien erhaltenen Daten sowie mit Hilfe einer Literaturrecherche identifiziert. Weitere Schwachstellen sowie Verbesserungsmöglichkeiten wurden gemeinsam mit dem PA erarbeitet. Anschließend wurde im Rahmen von Fokusgruppendiskussionen ein Leitfaden zur eigenständigen Durchführung des Datenanalyseprozesses für KMU entwickelt, von der Datenerhebung bis hin zur Datenanalyse und -evaluation. Die Datenvisualisierung ist Teil der Fallstudien und wird in einem späteren Arbeitspaket auf Basis der erhaltenen Daten beschrieben.

Ziel des zweiten AP war die Analyse der vorhandenen Daten zur Anwendung von Machine Learning in den identifizierten Anwendungsfällen. Dies beinhaltet die Bewertung der Daten in Bezug auf die erforderliche Datenqualität sowie die Erarbeitung eines Konzepts zur Steigerung der Datenqualität (datenbasierte Anforderungen). Für die Umsetzung der Datenanalyse wurden zunächst in der Literatur verbreitete Bewertungskonzepte der Datenqualität gesichtet und im Hinblick auf das Projekt analysiert. Auf diese Weise wurde ein Regelwerk für die Messung der Datenqualität erstellt und dieses in Expertengesprächen mit Machine Learning-Dienstleistern validiert. Im nächsten Schritt wurden die verfügbaren Unternehmensdaten identifiziert, bezüglich ihrer Qualität analysiert und eingeordnet. Die relevanten Informationen wurden in einem morphologischen Kasten zusammengefasst. Der Arbeitsschritt liefert somit einen systematischen Überblick über die Unternehmensdatenstrukturen. Die Ergebnisse wurden schließlich in die Prozesslandkarte eingegliedert. Um die Datenqualität weiter zu verbessern, folgten im nächsten Schritt die Identifizierung typischer Datenqualitätsmängel ausgewählter Datentypen sowie die Erarbeitung von Verbesserungsmöglichkeiten in Zusammenarbeit mit dem PA. Zuletzt wurde im Rahmen von Fokusgruppendiskussionen ein Leitfaden zur eigenständigen Durchführung des Datenanalyseprozesses für KMU entwickelt.

Maßgebend für die Arbeiten in diesem Arbeitspaket waren Datensätze, die in Vorbereitung für die Fallstudien in Arbeitspakets bereits eingeholt werden konnten. Eine geplante Fallstudie bei

einem Maschinenbauunternehmen aus Süddeutschland lieferte Bilder von Wendeschneidplatten, bei denen automatisiert die Qualität der einzelnen Schneider erfasst werden soll. Ein Rohrreinigungsunternehmen lieferte Videodateien von Rohranalysen kombiniert mit Daten über Befunde in den Videos, bei denen Befunde durch automatisierte Objekterkennung unterstützt werden sollten. Ein Unternehmen im Bereich der Sicherheitstechnik lieferte Datensätze über alle produzierten Teile bestimmter Produktserien über einen Zeitraum von einem Jahr. Dabei wurden einzelne Produktserien auf Anomalien in der Qualitätsprüfung untersucht. Zusätzlich wurden drei zu kombinierende Datensätze geliefert, wobei der erste Datensatz Komponenten der Endprodukte und deren Messwerte in der Prüfanlage enthielt, der zweite Messwerte und das Endergebnis der Qualitätsprüfung des Endprodukts und der letzte eine Möglichkeit zur Zuordnung der Komponenten zu ihren Endprodukten. Dabei soll untersucht werden, ob die Messwerte der Komponenten in ihrer Prüfanlage Auswirkungen auf die Wahrscheinlichkeit eines positiven Testergebnisses der Endprodukte haben.

2.2.1 Entwicklung eines Regelwerks für die Messung der Datenqualität

Um ein geeignetes Regelwerk für die Messung der Datenqualität zu entwerfen, wurde zunächst eine Literaturrecherche zu bekannten Mess- und Bewertungskonzepten für die Datenqualität durchgeführt. Die gefundenen Modelle (Wand und Wang 1996; English 2002; Naumann 2007; Lee et al. 2002) wurden auf ihre Dimensionen untersucht. Diese wurden extrahiert und in internen Workshops kategorisiert. Anschließend wurden die definierten Dimensionen in drei Überkategorien eingeteilt, welche in weiteren Workshops und Expertengesprächen (u. a. Experten von IT-Dienstleistern und Forschern im Bereich der Datenqualität) validiert wurden. Insgesamt 30 Quellen wurden vertieft untersucht und im MLready-Datenqualitätsmodell vereint. Das Ergebnis ist in Abbildung 6 dargestellt.

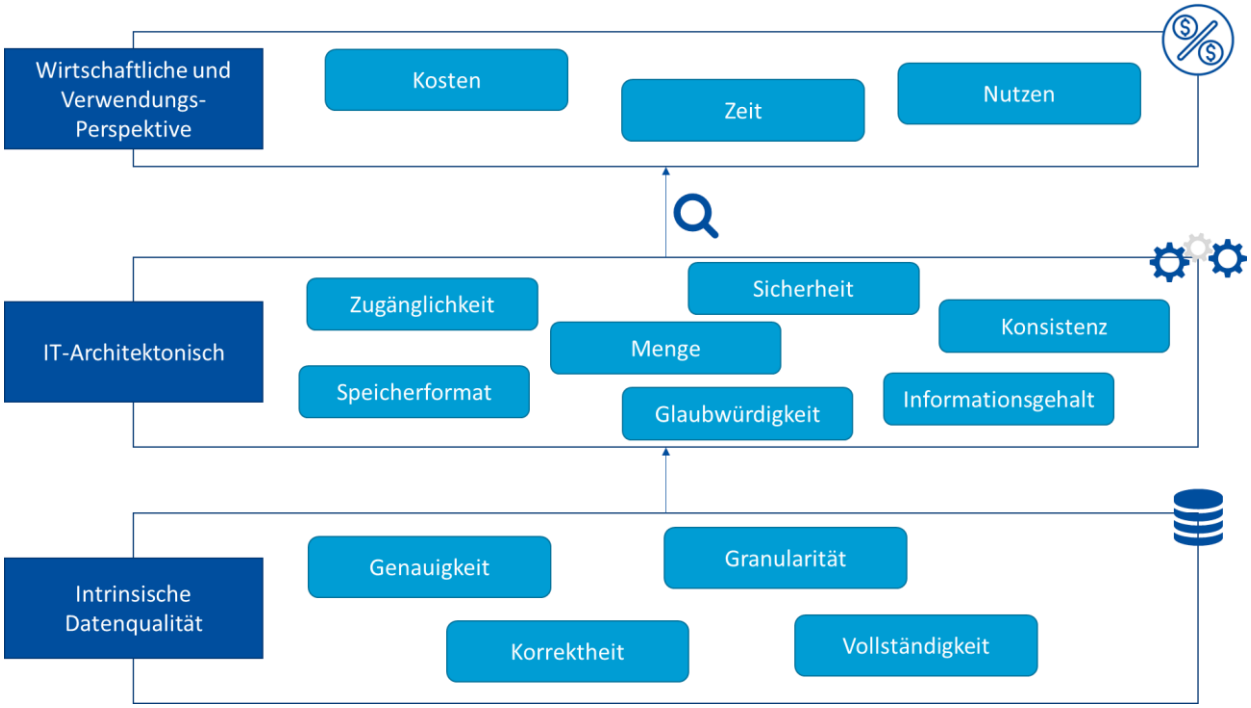


Abbildung 6: MLready-Datenqualitätsmodell

Generell ist anzumerken, dass die Erhebung und Verarbeitung von Daten zu jedem Zeitpunkt legal sein muss. Dies wird im Rahmen des Qualitätsmodells als Grundvoraussetzung angesehen und nicht als Teil bewertet, da die Legalität eine Conditio sine qua non für jede Machine Learning Anwendung darstellt.

Die intrinsische Datenqualität bezeichnet Faktoren, die bemessen, ob und wie die erhobenen Daten die Realität abbilden.

- Die **Genauigkeit** der Daten ist dann gegeben, wenn jede erhobene Information jeweils genau einem Wert in den Daten zugeordnet ist.
- Daten sind **korrekt**, wenn der jeweilige Wert korrekt gemessen und dementsprechend in die Datenbank eingetragen wurde.
- Daten sind **vollständig**, wenn alle für den jeweiligen Nutzen erforderlichen Informationen in die Datenbank aufgenommen wurden und fehlende Werte minimiert werden.
- Die **Granularität** bezeichnet die Kleinteiligkeit der Daten, ob also Informationen auf kleinteiliger Ebene enthalten sind.

IT-architektonische Faktoren bezeichnen solche hinsichtlich der Verwendung und Verwendbarkeit der Daten (ob die Daten für die jeweilige Verwendung geeignet sind).

- Daten sind **zugänglich**, wenn sie für Nutzer möglichst klar und schnell abrufbar sind.

- Das **Speicherformat** bezeichnet die informationstechnische Speicherung der Daten und beinhaltet unter Effizienz im Speichervolumen und die maschinelle Interpretierbarkeit.
- Daten sind erst dann von hoher Qualität, wenn bei ihrer Erhebung, Speicherung und Weiterverarbeitung **Sicherheitsstandards** eingehalten werden. Zum Beispiel müssen die Daten vor Manipulationsversuchen sicher sein und die Informationsweitergabe darf nur an Befugte erfolgen.
- Daten sind **glaubwürdig**, wenn bei der Informationserhebung bestimmte Qualitätsstandards eingehalten werden, nach denen die Nutzer der Korrektheit der Daten vertrauen können.
- **Konsistenz** bedeutet, dass Daten immer in gleicher Form sowie regelmäßig und zeitlich gleichmäßig erhoben und gespeichert werden.
- Das richtige **Datenvolumen** bzw. die geeignete **Menge** der Daten ist dann gegeben, wenn die Quantität der Daten für die jeweilige Nutzung der Daten ausreicht.
- Die Qualitätsanforderungen des **Informationsgehalts** beziehen sich darauf, ob die erhobenen Informationen einen tatsächlichen Nutzen haben, für andere Anwendungen verwertet werden und somit, ob sie nützlich sind.

Die wirtschaftliche und Verwendungsperspektive bezeichnet die Sinnhaftigkeit der Verwendung der Daten (ob der Aufwand der Erhebung, Verarbeitung und Verwendung der Daten den Nutzen rechtfertigt).

- Die Datenspeicherung und Datenpflege kann für das Unternehmen mit hohen finanziellen **Kosten** einhergehen. Bei der Verwendung der Daten muss der Aufwand dem Nutzen angemessen sein.
- Für eine hohe Datenqualität müssen die Daten für den jeweiligen Nutzen **aktuell** sein, das Alter der Daten muss für den jeweiligen Nutzen sinnvoll sein.
- Der **Nutzen** von Daten ergibt sich aus dem umgesetzten Potenzial der Anwendung nach Erhebung, Instandhaltung und Verarbeitung der Daten.

2.2.2 Erarbeitung von Datenquellen

Im zweiten Schritt wurden Datenquellen erarbeitet, die für die Umsetzung von Machine Learning verwendet werden können. Dazu wurde eine systematische Recherche von Ratgebern über Google durchgeführt, die eine breite Basis an Daten empfehlen. Zusätzlich wurden alle Veröffentlichungen untersucht, in denen die Fallstudien aus Arbeitspaket 1 abgeleitet wurden. Wurden in den Papern Daten angegeben, die für die Anwendungsgebiete verwendet wurden, wurden diese der Liste hinzugefügt. Außerdem wurden in den laufenden Fallstudien in Gesprächen mit

den Unternehmen weitere Datenquellen besprochen, die für die zugrundeliegenden Fallstudien von Interesse sein könnten und ggf. ergänzt. Insgesamt 180 verschiedene Beispiele für Daten konnten gefunden werden. Diese wurden in die neun Kategorien Aftersales-Daten, Controlling- & Finanzdaten, Einkaufsdaten, IT-Daten, Logistikdaten, Marketingdaten, Personaldaten, Produktionsdaten und Vertriebsdaten eingeteilt.

In internen Workshops wurden anschließend Merkmale der verschiedenen Daten erhoben. Dazu wurden zunächst wenige Datensätze hinsichtlich ihrer Merkmale beschrieben und sowohl die Dimensionen als auch die Ausprägungen notiert. Anschließend wurden peu à peu weitere Beispiele in das betrachtete Set hinzugefügt und in den entstandenen morphologischen Kasten einsortiert. Neue Ausprägungen wurden ggf. ergänzt. Konnte der morphologische Kasten ein relevantes Merkmal bestimmter Daten nicht aufgreifen, wurde eine neue Dimension entwickelt und dem morphologischen Kasten hinzugefügt. Das Ergebnis ist in Tabelle 3 dargestellt. Der morphologische Kasten umfasst die relevanten Merkmale aller erhobenen Daten und bietet damit die Möglichkeit, weitere Daten für spezifische Anwendungen von Machine Learning zu identifizieren. Durch die Kombination beliebiger Ausprägungen von Merkmalen können Ideen für mögliche Datenquellen gefunden werden. Dabei dient der morphologische Kasten als optionale Erweiterung des normalen Prozesses zur Aufbereitung der Daten, da im in Arbeitsschritt 3 entwickelten Leitfaden bereits eine klare Logik zur Identifikation der für einen Anwendungsfall benötigten Daten definiert ist.

Tabelle 3: Morphologischer Kasten verfügbarer Daten

Merkmal	**Ausprägungen**						
Herkunft Werden Daten intern erhoben oder von extern zugezogen?	Intern				Extern		
Strukturiertheit Liegen die Daten maschinell verwertbar vor?	Unstrukturiert			Semi-strukturiert			Strukturiert
Datenformat In welchem Format liegen die Daten vor?	handschriftlich	Textdateien	Bilddateien	json, xml etc.	csv, xls oder Datenbankformate	Audiodateien	Videodateien
Veränderlichkeit Wie oft werden Daten erhoben?	Sekündlich	Minütlich		Stündlich	Täglich		(>) wöchentlich

Zugänglichkeit Kann direkt auf die Daten zugegriffen werden?	Im eigenen System	Mit Zugriffsbeschränkungen	Frei im Internet zugänglich	Kostenpflichtig	Unzugänglich
Volumen Wie viele Daten sind verfügbar?	Einzelfälle	Regelmäßige Dokumentation	Hochfrequente Dokumentation über längere Zeit		
Abhängigkeiten Können Daten unabhängig verwendet werden?	nur in Verbindung in komplexen Systemen aussagekräftig	in Verbindung mit wenigen anderen Datensätzen aussagekräftig	kann eigenständig verwendet werden		

Anschließend wurden die identifizierten 180 Daten zu den Anwendungsgebieten zugeordnet. Dazu wurde jede der Daten mit jedem der definierten Anwendungsfälle gematcht. Mit Hilfe einer Tauglichkeitsbewertung wurde entschieden, ob Daten für Anwendungsfälle benötigt, unterstützend oder nicht hilfreich sind. Das Ergebnis sind Listen von Daten, die für jeweilige Anwendungsfälle hilfreich sein können. Die aggregierten Arten der Daten (z. B. Produktionsdaten, Controlling- und Finanzdaten) wurden in den Steckbriefen der Anwendungsfälle und der Prozesslandkarte ergänzt. Dadurch wird Unternehmen eine direkte Identifikation von Daten, die aufbereitet werden sollten, ermöglicht.

2.2.3 Leitfaden zur eigenständigen Aufbereitung von Daten

Zunächst wurden Verbesserungsmöglichkeiten für Schwachstellen von Daten erarbeitet. Dazu wurden zunächst gängige Methoden aus der Literatur erarbeitet. Außerdem wurden Aufbereitungsmethoden, die in den Fallstudien angewendet wurden, um die von den Partnern erhaltenen Daten für die Verwendung von Machine Learning vorzubereiten, ergänzt. Die Methoden wurden in Gesprächen mit den Fallstudienpartnern auf ihre Praxistauglichkeit untersucht und validiert. Die Methoden wurden dabei auf einem aggregierten Level gehalten, da genaue Methoden oft zu spezifisch sind und nur für höchstspezielle Anwendungen anwendbar sind. Im Sinne der Übertragbarkeit der Ergebnisse wurden vier Methoden herausgearbeitet, die bei der Verbesserung der Datenqualität unterstützen können.

Geschäftsprozessoptimierung: Die Abläufe in einem Unternehmen werden durch die Geschäftsprozesse festgelegt. Diese Abläufe sind grundlegend und werden aufgrund ihrer Komplexität nur ungern verändert. Dennoch kann es notwendig sein, in die Hauptabläufe des Unternehmens einzugreifen, um die Qualität der Daten sicherzustellen. So kann eine Adaption bestehender Prozesse notwendig sein, um langfristig die Fehler in der Datenqualität zu beheben.

Systemoptimierung: Um die Datenqualität zu steigern, können zudem die verwendeten Informationssysteme optimiert werden. Dies umfasst bspw. die Anpassungen des Datenmodells, die Einführung einheitlicher Datenbezeichnungen, die Einführung von verschiedenen Constraints sowie die Überarbeitung von Geschäftsprozessapplikationen. Dadurch werden Fehler bei der Datenerfassung, -speicherung und -nutzung vermieden.

Datenbereinigung: Die Verbesserung der Datenqualität beinhaltet die Bereinigung von identifizierten Fehlern, entweder automatisiert oder manuell durch Mitarbeitende. Um eine permanente Datenqualität zu wahren ist die Integration von Qualitätsmessungen in die Geschäftsprozesse erforderlich.

Schulung von Mitarbeitenden: Die Schulung der Mitarbeitenden ist ein weiterer Aspekt, welcher zur Verbesserung der Datenqualität beitragen kann. Neben technischen Maßnahmen ist es wichtig, sie durch klare Arbeitsanweisungen zu unterstützen. Schulungen informieren die Mitarbeitende darüber, wie sie mit den Daten umgehen sollen und mit ihrer Arbeit zur unternehmensweiten Datenqualität beitragen können. Mit der Einführung einer sog. Data Culture wird der Umgang mit Daten nachhaltig verändert.

Abschließend wurden die Ergebnisse des Arbeitspakets in einen einheitlichen Leitfaden zur eigenständigen Aufbereitung von Daten gegossen. Der Leitfaden befähigt kleine und mittelständische Unternehmen im produzierenden Gewerbe dazu, eigenständig Datenquellen zu identifizieren und Daten für die Verwendung von Machine Learning vorzubereiten.

Im Leitfaden wird eine Checkliste aufgestellt und einzelne Schritte im Detail beschrieben. Dabei folgt der Leitfaden dem generellen Vorgehensmuster:

1. Anwendungsfall definieren (mit Hilfe der Prozesslandkarte)
2. Datengrundlagen definieren (auf Basis der angegebenen Daten in der Prozesslandkarte und ergänzend mit dem morphologischen Kasten)
3. Datenqualität messen, um eine Abschätzung über die Erfolgswahrscheinlichkeit der Anwendung von Machine Learning zu gewinnen
4. Identifizierte Mängel in den Daten beheben

Der Leitfaden wurde auf der Projekthomepage veröffentlicht und kann dort frei eingesehen werden.

2.3 Arbeitspaket 3: Entwicklung einer KMU-spezifischen Einführungsstrategie für relevante Machine Learning-Ansätze

AP 3: Entwicklung einer KMU-spezifischen Einführungsstrategie für relevante Machine Learning-Ansätze	
Geplante Arbeiten	**Durchgeführte Arbeiten**
Basierend auf den Ergebnissen aus AP1 werden mit Hilfe von Experteninterviews sowie Erfahrungen von Einzelfallstudien, die das IPH im Rahmen des „Mittelstand 4.0 Kompetenzzentrums Hannover“ sowie „Mittelstand Digital Zentrum Hannover“ sammeln konnte, Anforderungen für einen umfassenden Einsatz von Machine Learning evaluiert und anschließend werden Maßnahmen zur Umsetzung (Migrationsschritte zwischen Ist- und Soll-Zustand) in KMU erarbeitet. Im ersten Schritt wird mit den Fallstudienpartnern eine RASCI-Matrix (Responsible-Accountable-Support-Consulted-Informed) zur Zuordnung von Verantwortlichkeiten und Rollen von internen und externen Akteuren entwickelt.	Zur Ermittlung der Anforderungen, die das wesentliche Fundament des umfassenden Einsatzes von Machine Learning darstellen, wurden diverse Experteninterviews mit KI-Experten des Mittelstand Digitalzentrums Hannover durchgeführt. Hierbei wurden relevante Grundvoraussetzungen ermittelt. Darüber hinaus wurde im Austausch mit den Fallstudienpartnern und weiteren Unternehmen aus der Industrie eine Verantwortlichkeitsmatrix entwickelt. Diese enthält zum einen die relevanten Aufgaben, die bei der Einführung von Machine Learning durchgeführt werden müssen. Zum anderen werden die Rollen dargestellt, die für die erfolgreiche Umsetzungen des Projektes notwendig sind.
Im zweiten Schritt folgt die Erarbeitung und Bewertung von Handlungsmaßnahmen zur Umstellung der Prozesse, um Machine Learning-Anwendungen zu ermöglichen. Hierbei wird auf technische, organisatorische und personelle Aspekte eingegangen. Anschließend erfolgt eine Wirtschaftlichkeitsbewertung. Die abgeleiteten Maßnahmen werden hinsichtlich ihres Aufwands und ihres Nutzens, bspw. des Aufwands der Implementierung, bewertet. Zudem werden Maßnahmen zur Erfolgsüberprüfung im Rahmen der Entwicklung von Kennzahlen (bspw. Materialeinsparungen) ergänzt.	Im zweiten Schritt der 3. Arbeitspakets wurden die zum erfolgreichen Einsatz von Machine Learning notwendigen Handlungsmaßnahmen untersucht und identifiziert. Um deren erfolgreiche Umsetzung gewährleisten zu können, wurde vorab eine Reifegraduntersuchung entwickelt. Diese gibt Erkenntnis darüber, ob es sinnvoll ist, sich an diesem Punkt mit den Handlungsmaßnahmen auseinanderzusetzen. Darüber hinaus wurden weitere Vorteile herausgearbeitet, um den Anwendern sowohl monetäre als auch ressourceneffiziente Mehrwerte aufzuzeigen.

Anschließend werden mit dem PA-Fokusgruppendiskussionen zur Fragestellung geführt, wie ein geeigneter Machine Learning-Dienstleister ausgewählt werden kann. Die gesammelten Ergebnisse werden mittels morphologischem Kasten gruppiert. Darauf aufbauend werden Mindestanforderungen für die Dienstleisterauswahl definiert und in einem ML***ready***-Lastenheft zusammengefasst. Die Ergebnisse werden in einer Roadmap KMU-gerecht aufbereitet.	Die Ermittlung der Anforderungen an Machine Learning-Dienstleister ist anschließend durchgeführt worden. Hierbei wurde zuerst eruiert, welche KMU-spezifischen Anforderungen eine hohe Relevanz aufweisen und daraus resultierend für die erfolgreiche Zusammenarbeit mit externen Dienstleistern notwendig sind. Die dabei erzielten Ergebnisse wurde in der PA-Fokusgruppendiskussion besprochen und auf Basis der erhaltenen Rückmeldungen angepasst.
Die Umsetzungsstrategie wird in Form einer Implementierungs-Roadmap (mithilfe der Methode des Roadmappings) und eines Leitfadens zur Implementierung abgebildet, in den die Ergebnisse des Arbeitspakets einfließen. Die allgemeingültige Anwendbarkeit wird in AP4 überprüft.	Im letzten Schritt des dritten Arbeitspakets wurde eine Implementierungs-Roadmap mit Hilfe der Plattform nedyx entwickelt. Hierbei wurden alle ermittelten Ergebnisse eingepflegt, um dem Anwender eine optimale Einführung von Machine Learning-Anwendungen zu ermöglichen.

2.3.1 Evaluation des umfassenden Einsatzes von ML und Entwicklung einer RASCI-Matrix

Da die Grundvoraussetzungen als zentraler Bestandteil der Roadmap deklariert wurde, werden diese im Abschnitt 2.3.4 näher erläutert und daher an dieser Stelle vorerst übersprungen. Die durch den projektbegleitenden Ausschuss als wesentlich identifizierten Aufgaben wurden nach der Durchführung von Experteninterviews in Form einer Verantwortlichkeits- bzw. RASCI-Matrix (Responsible, Accountable, Support, Consulted, Informed) umgesetzt, da diese einen übersichtlichen und leicht verständlichen Aufbau bietet. Des Weiteren erfolgt eine übersichtliche Darstellung der Zuordnung der Verantwortlichkeiten und Rollen, insbesondere der internen Rollen. Die dargestellte Verantwortlichkeitsmatrix bietet KMU die Möglichkeit, eine optimale Aufteilung von Aufgaben und Rollen für KMU zu definieren.

Die zentralen Aufgaben, die sich bei der Transformation des Ist- in den Soll-Prozess ergeben haben, umfassen die Beschaffung und Prüfung der IT-Infrastruktur, die Programmierung und das Training der Machine Learning-Anwendung, die Planung der Schulung für die Mitarbeiter sowie die Definition der Vorgehensweise beim Ausfall oder bei einer Fehlfunktion der Machine Learning-

Anwendung. Diese Auswahl ist bedingt durch die Ermittlung der Grundvoraussetzung, auf die im Kapitel 2.3.4 näher eingegangen wird.

Des Weiteren ist die Rollenverteilung von essentieller Bedeutung für die Einführung eines Projektplans. Daher erfolgte die Ermittlung der relevantesten Rollen in erster Linie durch eine Literaturrecherche. In der Folge wurde eine Diskussion und Validierung mit Experten aus der Industrie und des projektbegleitenden Ausschusses durchgeführt, um eine zielführende Verantwortlichkeitsmatrix zu entwickeln, die neben den relevanten Rollen zentrale Aufgaben erhält (siehe Abbildung 7).

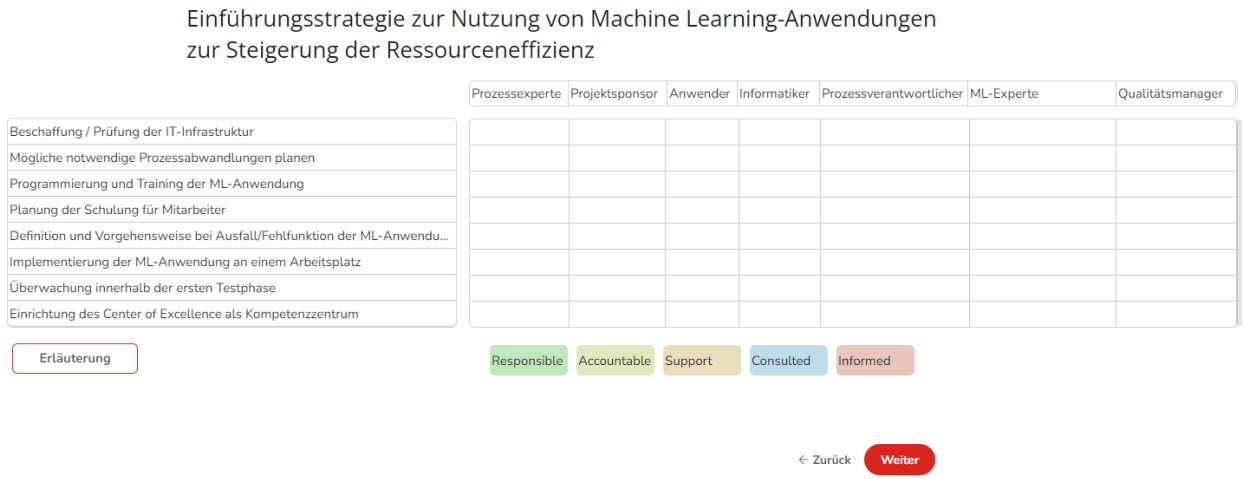


Abbildung 7: Verantwortlichkeits- bzw. RASCI-Matrix

Im Rahmen der Untersuchung wurde ersichtlich, dass die erfolgreiche und nachhaltige Einführung von Machine Learning eine klare Aufgaben- und Rollenteilung erfordert. Die durchgeführten Experteninterviews haben ergeben, dass für eine optimale Projektdurchführung die Beteiligung verschiedener Akteure erforderlich ist. Dies sind zunächst ein Prozessverantwortlicher, ein Prozessexperte, ein Anwender, der Projektsponsor (oder eine Person in leitender Position), ein Machine Learning Experte und ein Qualitätsmanager. Es sei darauf verwiesen, dass die Bereitstellung von sieben Mitarbeitern für ein langfristiges Machine Learning-Projekt für KMU in der Regel nicht realisierbar ist. Daher sei angemerkt, dass die sieben genannten Rollen nicht zwangsläufig von jeweils einer Person wahrgenommen werden müssen. Es besteht die Möglichkeit, dass eine Person mehrere Rollen innerhalb des Projektteams einnimmt. So kann beispielsweise der Prozessexperte zugleich die Funktion des Prozessverantwortlichen oder des Machine Learning-Experten der Informatiker ausüben.

Die entwickelte RASCI-Matrix wurde in die Roadmap integriert, um ihre Effizienz bei der Einführung von Machine Learning zu gewährleisten. In diesem Zusammenhang wurde die Möglichkeit implementiert, dass der Anwender jeder Rolle eine Zuständigkeit, d.h. Responsible, Accountable, Support, Consulted, Informed (vgl. weiter oben in der Erläuterung des Begriffs "RASCI") zuweisen kann. Folglich kann bereits während des Ausfüllens der Verantwortlichkeitsmatrix ein zielführender Projektplan mit der optimalen Aufgaben- und Rollenzuteilung erstellt werden.

2.3.2 Entwicklung von Handlungsempfehlungen und einer Wirtschaftlichkeitsbewertung

Die Entwicklung von organisatorischen, technischen und personellen Handlungsempfehlung zur Umstellung von Prozessen wurde im zweiten Arbeitsschritt untersucht. Hierbei wurde die Ergebnisse aus dem ersten Arbeitsschritt des dritten Arbeitspakets weiter aufgearbeitet und analysiert, wie diese optimal umgesetzt werden können. Die ermittelten Handlungsempfehlungen sind als wesentlicher Bestandteil in die Roadmap eingebaut worden und werden daher in den folgenden Kapiteln erwähnt.

Die Wirtschaftlichkeitsbewertung ist im Rahmen des Forschungsprojekt im Austausch mit dem projektbegleitenden Ausschuss entwickelt und validiert worden. Hierbei wurden mögliche interne Kostenfaktor, wie beispielsweise der Personaleinsatz (Anzahl an Mitarbeitern, Kosten pro Mitarbeiter, notwendige Zeit) und externe Kosten (Kosten für Dienstleister) ermittelt und der Kostenersparnis gegenübergestellt, sodass final ermittelt werden konnte, wann eine Amortisation zu erwarten sei. Jedoch hat sich ebenfalls im Austausch mit dem projektbegleitenden Ausschuss herausgestellt, dass die Kosten je nach vorliegendem Reifegrad (ein Werkzeug zur Reifegradmessung wird in 2.3.4 näher beschrieben), betrachteter Anwendung, Datengrundlage, Zeitkapazität und weitere Faktoren sehr stark variieren, sodass keine verallgemeinerbare Aussage hierzu getroffen werden konnte. Final wurde entschieden, dass die Wirtschaftlichkeitsrechnungen KMU bei Bedarf zugänglich gemacht werden sollen, jedoch keinen festen Bestandteil des Leitfadens darstellen sollen.

2.3.3 Entwicklung von Mindestanforderungen für die Auswahl von Machine Learning Dienstleister

Sofern im Unternehmen kein explizites Know-how im Bereich des maschinellen Lernens vorliegt, ist die Hinzuziehung eines externen Dienstleisters in der Regel unumgänglich. Dies ist darauf zurückzuführen, dass die Schulung des Mitarbeiters mit einem hohen zeitlichen und finanziellen Aufwand verbunden ist. Die Vielzahl an Dienstleistern am Markt führt insbesondere bei KMU zu einer Überforderung, sodass keine weitere Zeit in die Suche nach einem Machine Learning-

Dienstleister investiert wird. Im dritten Arbeitsschritt wurde eine Literatursuche durchgeführt. Im Rahmen dieser Untersuchung wurde analysiert, welche Kriterien KMU bei der Auswahl von Machine Learning-Dienstleistern als maßgeblich erachten.

Die ermittelten Ergebnisse wurden im Anschluss daran im Rahmen eines Experteninterviews mit Machine Learning-Dienstleistern sowie mit den Mitgliedern des projektbegleitenden Ausschusses erörtert (vgl. Abbildung 8). Die Ergebnisse der Untersuchung legen nahe, dass KMU eine aufwandsarme Kommunikation mit Dienstleistern präferieren. Ein weiterer Aspekt, der insbesondere vom projektbegleitenden Ausschuss adressiert wurde, ist die Präferenz von KMU für die Präsenz eines zentralen Ansprechpartners für alle Fragen und Anliegen. Die Bereitschaft von Unternehmen, sich akribisch mit Machine Learning-Dienstleistern auseinanderzusetzen, verringert sich signifikant mit steigendem Aufwand und steigender Anzahl an Ansprechpartnern.

Frage	Trifft zu	Trifft nicht zu
Ist der Dienstleister auf den Anwendungsfall spezialisiert? Kennt er sich beispielsweise mit der bisher verwendeten Software aus und kann Schnittstellen für benötigte Daten schaffen?	☐	☐
Hat der Dienstleister Erfahrungen mit ähnlichen KI-Anwendungen?	☐	☐
Bietet der Dienstleister eine Ansprechperson und kann kurzfristig auftretende Probleme lösen?	☐	☐
Bietet der Dienstleister Cloud-Lösungen an?	☐	☐
• Kann bei der Cloud-Lösung die Sicherheit der unternehmensinternen Daten vor Einsicht dritter Personen gewährleistet werden?	☐	☐
• Können große Datenmengen in der Cloud gespeichert werden?	☐	☐
Passt das Leistungsspektrum des Dienstleisters zu der geplanten Anwendung? Bspw.: Datenerhebung, -aufarbeitung, KI-Implementierung	☐	☐
Besitzt der Dienstleister bereits Erfahrungen in einer ähnlichen Branche, wie der des Unternehmens?	☐	☐
Stimmt das Preis/Leistungsverhältnis?	☐	☐
Passen die Vertragsbedingungen des Dienstleisters zu den Vorstellungen des Unternehmens?	☐	☐
Besitzt der Dienstleister genug personelle Kapazitäten, um das Projekt in einem absehbaren Zeitraum umzusetzen?	☐	☐
Ist der Dienstleister technologisch auf dem aktuellsten Stand und arbeitet mit modernen Methoden?	☐	☐
Ist der Dienstleister in der Lage, die KI-Anwendung unkompliziert in das laufende System zu integrieren?	☐	☐

Abbildung 8: Anforderungsliste an Machine Learning-Dienstleister

Das Preis-/Leistungsverhältnis wurde ebenfalls als wesentlicher Aspekt der Anforderungsliste detektiert. Dies ist immer sehr subjektiv und daher pauschal schwer zu beantworten. Jedoch ist dieser Punkt ein wichtiges Entscheidungskriterium, da zu hohe Kosten dazu führen können, dass sich KMU gegen das Heranziehen eines externen Dienstleisters und möglicherweise final auch gegen die Einführung von Machine Learning in der Produktion entscheiden könnten.

Die Erfahrung des Dienstleisters in Bezug auf Machine Learning-Projekte für KMU ist ein weiteres Kriterium, welches vom projektbegleitenden Ausschuss als relevant eingestuft wurde. Dies ist dadurch bedingt, dass ein hoher Kenntnisstand und viele Erfahrungswerte das Vertrauen von KMU in den Dienstleister steigern. Die Erfahrungen in Machine Learning-Projekten bei KMU sind von besonders großer Bedeutung, da KMU häufig individuellen Herausforderungen unterliegen und somit alternative Herangehensweisen und Lösungsansätze benötigen.

Dass Daten eine fundamentale Rolle bei der Einführung von Machine Learning in der Produktion spielen, wurde bereits ausführlich im zweiten Arbeitspakete analysiert und erläutert. Auch in Bezug auf die Auswahl und Kommunikation mit Machine Learning-Dienstleistern spielen Daten eine zentrale Rolle. Die erfolgreiche Entwicklung und Implementierung von Machine Learning-Anwendungen erfordert eine Datenerhebung und -aufarbeitung. Diese beiden Aufgaben stellen insbesondere KMU vor große Herausforderungen, wie sich in Experteninterviews herausgestellt hat. Insbesondere fehlende oder fehlerhafte Daten und nicht vorhandene Schnittstellen führen dazu, dass KMU sich nicht weiter mit dem Thema Machine Learning auseinandersetzen. Hierzu sind Maßnahmen notwendig, um dem entgegenzuwirken. Eine wesentliche Maßnahme ist, die Datenerhebung und -aufarbeitung durch oder in Zusammenarbeit mit dem externen Dienstleister durchzuführen. Daher ist dieser Aspekt eine der Anforderungen an Machine Learning-Dienstleister. Weist der Dienstleister ebenfalls eine Expertise in diesem Bereich auf, so verringert sich der Aufwand für KMU und die Hemmschwelle, sich akribisch mit der Einführung von Machine Learning-Anwendungen zu beschäftigen, sinkt signifikant.

Schlussendlich lässt sich sagen, dass KMU durch die Anforderungsliste eine optimale Entscheidungsunterstützung für die Auswahl von Machine Learning-Dienstleistern ermöglicht wird. Hierdurch kann der Aufwand bei der Suche nach dem optimalen Machine Learning-Dienstleister reduziert werden, während gleichzeitig eine effiziente Projektdurchführung ermöglicht wird.

2.3.4 Entwicklung einer Roadmap

Im Rahmen der Entwicklung der KMU-spezifischen Einführungsstrategie erfolgte eine Untersuchung bestehender Ansätze und Forschungsergebnisse mittels Literaturrecherche. Im Rahmen der Analyse wurde untersucht, auf welche Weise sich der Reifegrad eines Unternehmens anhand spezifischer Fragen und Antwortmöglichkeiten ermitteln lässt. Des Weiteren wurden sowohl die

entwickelte Prozesslandkarte als auch der Leitfaden zur Datenanalyse in die finale Einführungsstrategie integriert, welche auf der Plattform nedyx implementiert wurde. Ein besonderes Augenmerk galt dabei der Reifegradprüfung, welche KMU als Entscheidungsgrundlage für eine erfolgreiche Einführung von Machine Learning-Anwendungen in der Produktion dienen soll.

Wie bereits dargelegt, erfolgte im ersten Schritt eine Untersuchung bestehender Forschungsansätze und -ergebnisse zur Thematik Einführung von Machine Learning. In dem Forschungsprojekt "ML4P" des Fraunhofer-Instituts wurde festgestellt, dass die Ermittlung des Ist- und des Sollzustandes eine wesentliche Rolle bei der Einführung von Machine Learning spielt. Diesbezüglich wurde seitens des Fraunhofer-Instituts ausgeführt, dass dadurch die Kommunikation und Koordination zwischen Machine Learning- und Prozessexperten optimiert und gleichzeitig erleichtert wird (Fraunhofer 2020). Dieser Aspekt wurde zudem in der Forschung von Bauer et al. bestätigt, in der ein Handlungsleitfaden zur Einführung von KI in der Produktion entworfen wurde. Die Untersuchung des Reifegrads wird hierbei als zentraler Grundbaustein und strategischer Einstieg beschrieben (Pokorni et al. 2021).

Da die Definition und Einordnung von ML nicht immer trivial erscheinen, wurde eine Differenzierung zwischen den Begriffen Künstliche Intelligenz, Machine Learning und Deep Learning vorgenommen und in die Roadmap einfügt (vgl. Abbildung 9).

Abbildung 9: Definition von KI, ML und DL

Im Anschluss daran wurde eine Anleitung erstellt, anhand derer sich die Anwender über den bevorstehenden Ablauf informieren können, um frühzeitig potenzielle Unklarheiten und mögliche Missverständnisse zu identifizieren.

Im ersten Arbeitspaket wurde die Prozesslandkarte entwickelt, um eine optimale Bestimmung des Reifegrads und darauf aufbauend Handlungsmaßnahmen zur Transformation des Ist-Prozesses in den Soll-Prozess gewährleisten zu können. Die Prozesslandkarte bietet kleinen und mittelständischen Unternehmen die Möglichkeit, sich über die identifizierten Machine Learning-Anwendungen zu informieren, Beispielanwendungen einzusehen und dadurch Einsparpotenziale hinsichtlich der benötigten Ressourcen abzuleiten. Um auch weiterhin eine aufwandsarme Bereitstellung der Informationen für KMU gewährleisten zu können, wurden die Inhalte der Prozesslandkarte in die Einführungsstrategie integriert. In einem ersten Schritt wurden, wie in Abbildung 10 dargestellt, alle identifizierten Anwendungen abgebildet.

Abbildung 10: Darstellung der ermittelten Machine Learning-Anwendungen

Hiermit wird den Anwendern die Möglichkeit geboten, sich über die dargestellten Machine Learning-Anwendungen vertiefend zu informieren. Die Informationsfülle kann am Beispiel der Abbildung 11 eingesehen werden. Insbesondere das Ziel und der Nutzen in Form von Ressourceneinsparungen können als Entscheidungsunterstützung verwendet werden.

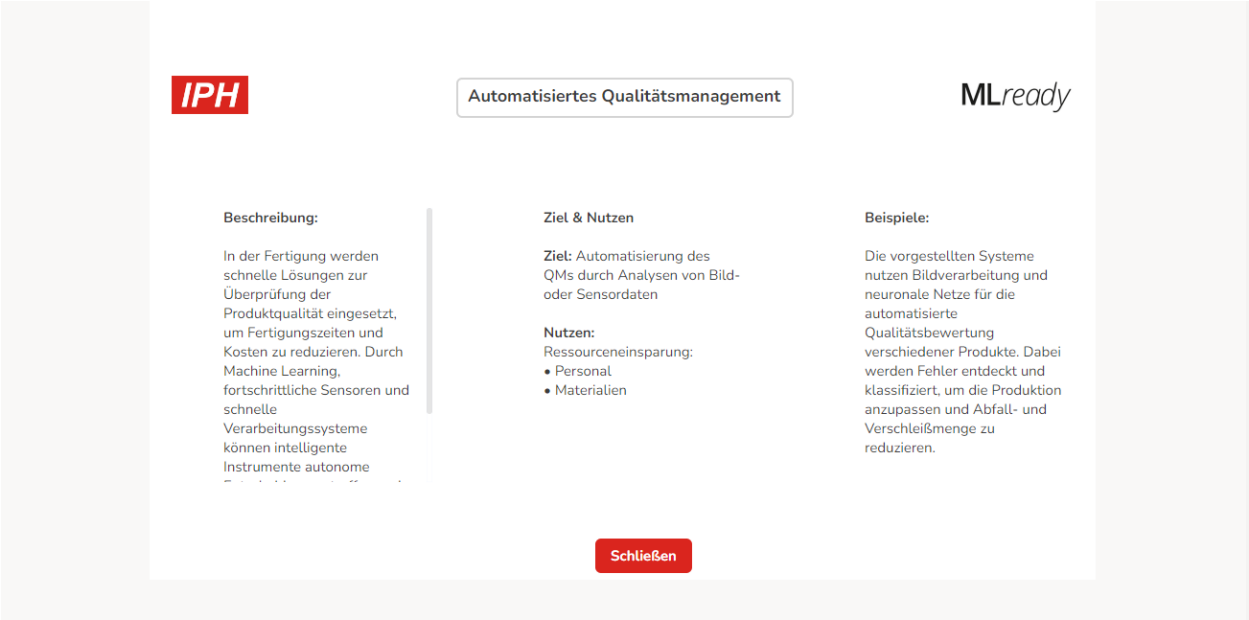


Abbildung 11: Beispielhafte Darstellung einer detaillierten Beschreibung der Machine Learning-Anwendung

Die Auswahl einer optimalen Anwendung von maschinellem Lernen kann Unternehmen vor signifikante Herausforderungen stellen. Dies gilt insbesondere für Unternehmen, die bislang keine nennenswerten Berührungspunkte mit den Machine Learning-Potenzialen in der Produktion hatten. Um dieser Herausforderung zu begegnen, wurde in Kooperation mit produzierenden Unternehmen ein Vorgehen entwickelt, welches die Selektion der geeigneten Machine Learning-Anwendungen erleichtert. Zu diesem Zweck wurden sieben Kriterien definiert, anhand derer die Auswahl derjenigen Machine Learning-Anwendung erfolgen kann, die den zu betrachtenden Prozess am besten beschreibt. Das Vorgehen basiert auf dem Prinzip der Priorisierung, welches sich durch eine einfache Handhabung auszeichnet. Dies impliziert, dass der Anwender die sieben Kriterien, welche in Abbildung 12 dargestellt sind, auf einer Skala von eins bis sieben priorisiert.

Es ist darauf zu achten, dass jede dieser Zahlen ausschließlich ein einziges Mal verwendet wird. Dabei stellt die Zahl "7" die höchste und die Zahl "1" die niedrigste Priorität dar.

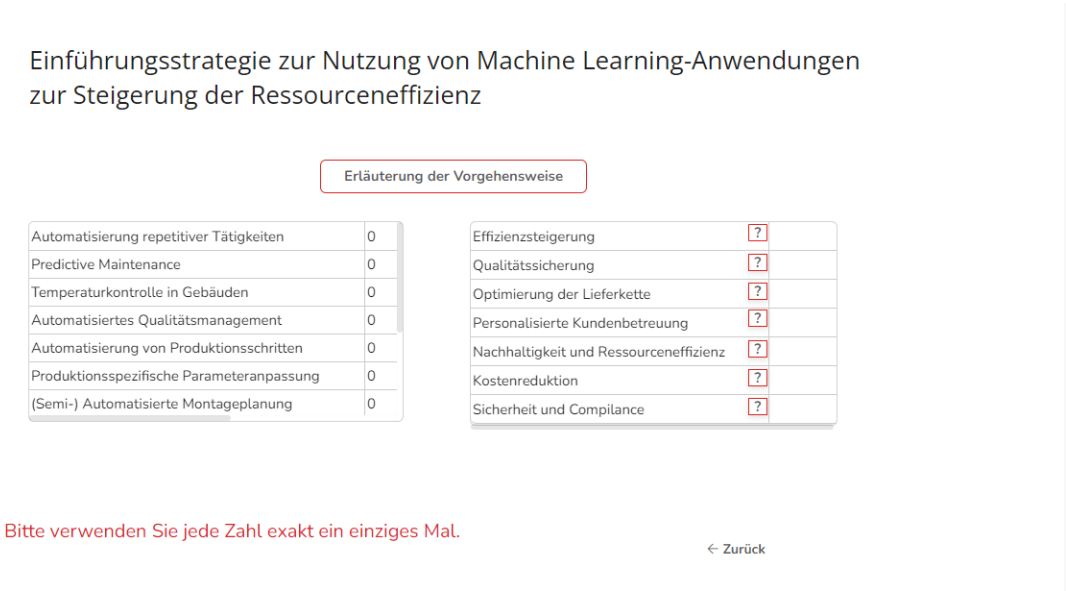


Abbildung 12: Entwickelte Verfahren zur Ermittlung der optimalen Machine Learning-Anwendung

Die Auswertung der genannten Kriterien erfolgte mittels eines darauf basierenden Rankings, welches in Experteninterviews und durch den projektbegleitenden Ausschuss diskutiert und validiert wurde. Die Roadmap wurde so programmiert, dass dem Anwender im Ergebnis die drei Machine Learning-Anwendungen angezeigt werden, dass am ehesten zu getroffen Auswahl bei dir Priorisierung passen. Auch hier wurden die Ergebnisse der Prozesslandkarte eingebaut, um den Zugang zu weiteren, relevanten Informationen zu ermöglichen. Im Anschluss an die Auswahl der Machine Learning-Anwendung wurde die Ermittlung des im Unternehmen vorliegenden Reifegrads entwickelt und eingebaut. Im ersten Schritt wurden die Grundvoraussetzungen ermittelt, die als Fundament für die erfolgreiche Einführung von Machine Learning deklariert wurden (siehe Abbildung 13). Auch diese wurden durch den projektbegleitenden Ausschuss diskutiert und validiert. Das Ziel dieser Untersuchung bestand darin, kleinen und mittelständischen Unternehmen

aufzuzeigen, dass bestimmte Grundvoraussetzungen erforderlich sind, um eine erfolgreiche Projektdurchführung gewährleisten zu können.

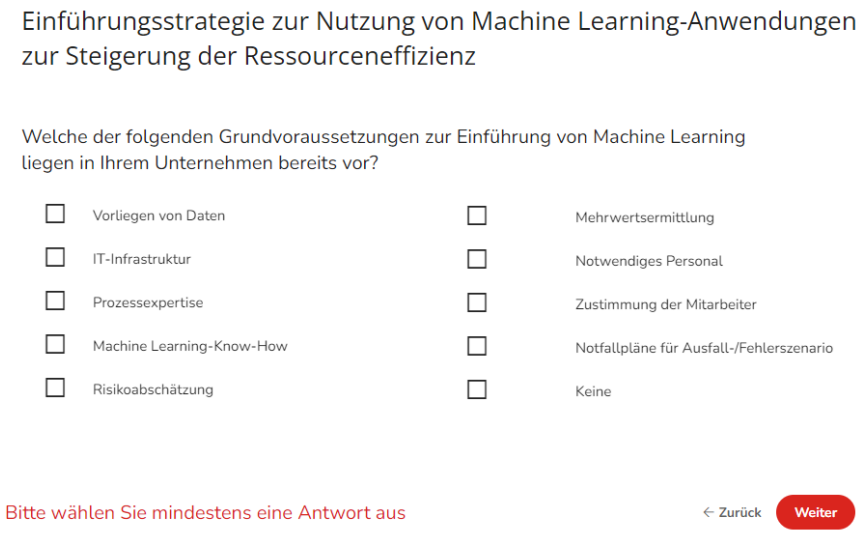


Abbildung 13: Untersuchung der vorliegenden Grundvoraussetzungen

Im Anschluss an die Ermittlung der vorliegenden Grundvoraussetzungen wurde eine Untersuchung des Know-hows des Unternehmens hinsichtlich der Thematik Machine Learning in die Roadmap implementiert. Zu diesem Zweck wurden auf Basis einer umfassenden Literaturrecherche drei Fragen (siehe Frage 1 bis 3) entwickelt, die es ermöglichen, sowohl das Wissen als auch die Erfahrungswerte der Geschäftsführung, der Mitarbeiter und der IT-Verantwortlichen zu ermitteln. Basierend auf bereits durchgeführten Forschungsarbeiten, wie beispielsweise ML4P, wurden insgesamt zehn Fragen entwickelt, mit deren Hilfe die Mitarbeiterakzeptanz, das vorliegende Know-how sowie die Prozesstauglichkeit untersucht werden können (siehe Tabelle 2). Die Fragestellung, ob die Entwicklung und Implementierung der betrachteten ML-Anwendung unternehmensintern auf Basis des vorliegenden Know-hows durchgeführt werden kann oder ob die Hinzuziehung externer Dienstleister empfehlenswert ist, bildet den Hintergrund dieser Untersuchung. Des Weiteren soll eruiert werden können, ob der Prozess je nach ausgewählter Anwendung adäquat ist oder ob er noch gewissen Modifikationen unterliegt.

Tabelle 4: 10 Fragen zur ML-Reifegradermittlung

Frage 1: Wie hoch ist die Kenntnis der Geschäftsführung zum Thema Machine Learning in der Produktion?
Frage 2: Wie hoch ist die aktuelle Kenntnis der Mitarbeiter zum Thema Machine Learning in der Produktion?
Frage 3: Inwieweit sind die IT-Verantwortlichen in der Lage, Machine Learning-Anwendungen in der Produktion zu programmieren und implementieren?

Frage 4: Wie hoch ist die Mitarbeiterakzeptanz gegenüber neuen Technologien?
Frage 5: Wie ist der Prozess aufgebaut?
Frage 6: Wie hochfrequentiert ist der Prozess?
Frage 7: Wie komplex ist der Prozess?
Frage 8: Wie ausgereift ist der Prozess?
Frage 9: Wie schätzen Sie den aktuellen Automatisierungsgrad Ihres Prozesses ein?
Frage 10: Wie fehlerbehaftet ist der Prozess?

Auf Grund der getroffenen Antworten soll ermittelt werden, welcher Gesamtreifegrad zum aktuellen Zeitpunkt im Unternehmen vorliegt. Um diesen zu vervollständigen bzw. zu optimieren, wurde relevante Handlungsmaßnahmen ermittelt (siehe Abbildung 14). Diese sollen bei der Transformation des Ist- Prozesses in den Soll- Prozess unterstützen. Die Handlungsmaßnahmen wurden im Austausch mit produzierenden Unternehmen und in Anlehnung an die ermittelten Grundvoraussetzungen definiert, sodass sich ein logischer Aufbau ergeben hat.

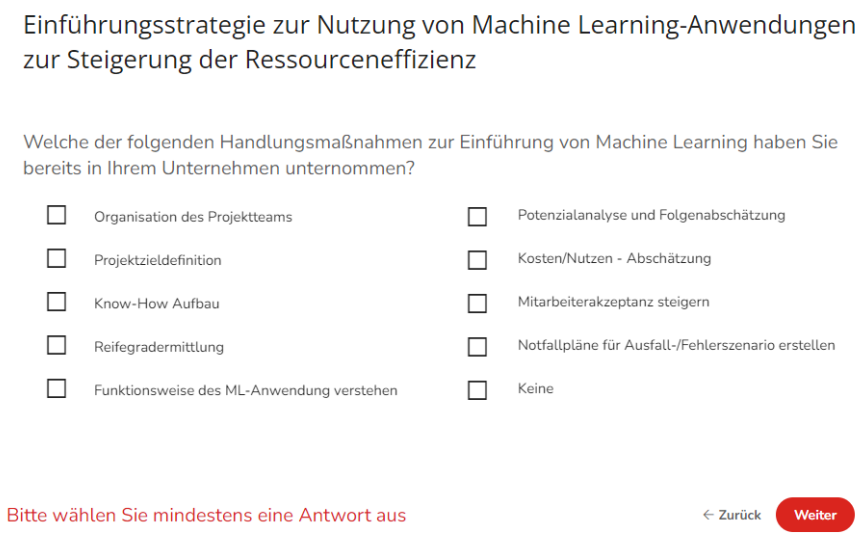


Abbildung 14: Ermittlung der Bereits durchgeführt Handlungsmaßnahmen

Final wurde eine Exportfunktion eingebaut, die die Erstellung eines Dokuments ermöglicht, das sowohl ein Deckblatt als auch alle relevanten Ergebnisse und Handlungsempfehlung enthält. Dieses Dokument bietet dem Anwender eine Grundlage für die erfolgreiche Einführung von der ausgewählten Machine Learning-Anwendung in der Produktion und kann Unternehmen somit einen großen Mehrwert bieten. Dieser Aspekt ist dadurch bedingt, dass dem Anwender neben der Darstellung des aktuell im Unternehmen vorliegenden Reifegrad, Handlungsempfehlungen und eine

Darstellung der nächsten Schritte dargestellt werden. Somit kann einem Projektleiter einen Einblick über bevorstehende Aufgaben gegeben werden (siehe beispielhafte Darstellung in Abbildung 15). Unter dem Link *https://getmlready.com/* besteht darüber hinaus die Möglichkeit, einen Einblick in die entwickelte Einführungsstrategie zu werfen.

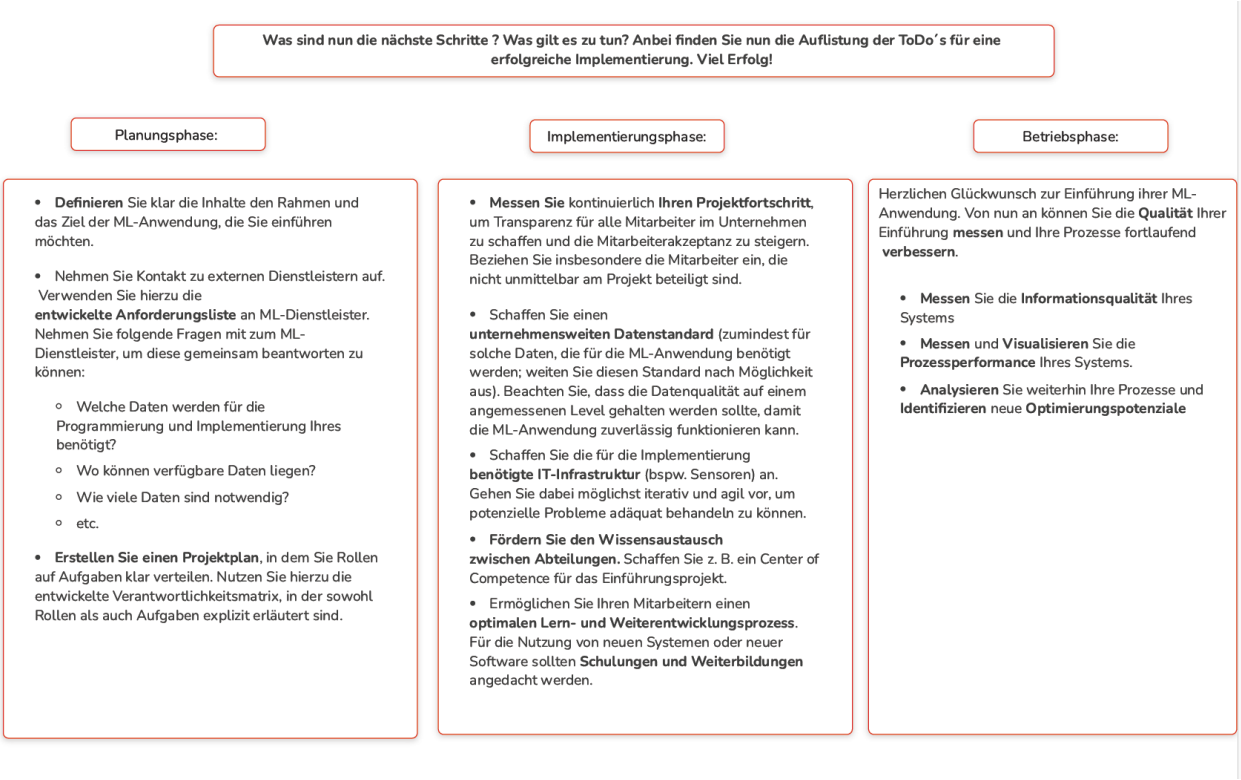


Abbildung 15: Darstellung der Handlungsmaßnahmen aufgeteilt in drei Phasen

2.4 Arbeitspaket 4: Durchführung von Fallstudien

AP 4: Durchführung von Fallstudien	
Geplante Arbeiten	**Durchgeführte Arbeiten**
Im ersten Schritt wird die technische Bereitstellung erforderlicher Daten (bspw. zusätzlicher Hardwarebedarf) analysiert. Anschließend müssen die in AP 3 identifizierten erforderlichen Daten identifiziert und zusammengetragen werden. Anschließend muss ein Preprocessing der Daten durchgeführt werden. Darunter fällt z. B. die Datenreduktion oder Feature Engineering. Aus den Ergebnissen von AP 2 werden geeignete Machine Learning-Ansätze für die	Im ersten Schritt wurden bei den fünf Fallstudien die technische Bereitstellung erforderlicher Daten (bspw. zusätzlicher Hardwarebedarf) analysiert. Anschließend wurden die in AP 3 identifizierten erforderlichen Daten identifiziert und zusammengetragen. Anschließend wurde in jeder Fallstudie ein Preprocessing der Daten durchgeführt. Darunter waren z. B. komplizierte Datenträger- und Dateiauslesungen, Datenreduktion und Bildverarbeitungen.

jeweiligen Fallstudien ausgewählt und getestet. Machine Learning wird abhängig vom Anwendungsfall mit den Paketen scikit-learn oder Keras mit TensorFlow als Backend in Python implementiert. Keras eignet sich besonders für die Implementierung von neuronalen Netzen, wohingegen scikit-learn Implementierungen „simplerer“ Machine Learning-Anwendungen unterstützt. Dabei kann auf bereits zur Verfügung stehende Bibliotheken zurückgegriffen werden, die in Python einfach eingeladen werden können. So kann die Umsetzung der Fallstudien im angegebenen Zeitraum gewährleistet werden.	Aus den Ergebnissen von AP 2 wurden jeweils geeignete Machine Learning-Ansätze ausgewählt und getestet. Machine Learning wurde abhängig vom Anwendungsfall mit den Paketen scikit-learn, pm4py, Keras mit Tensorflow als Backend und yolo umgesetzt. scikit-learn eignete sich besonders für die Implementierung einfacher Machine Learning-Anwendungen wie Entscheidungsbäumen, während pm4py für die Analyse von Log-Daten, Keras für die Implementierung komplizierter neuronaler Netze und yolo für Object Detection in Bilddaten verwendet wurde. Dabei wurde auf bereits zur Verfügung stehende Bibliotheken zurückgegriffen, die in Python einfach eingeladen werden können.
Die in AS 1 umgesetzten Machine Learning-Anwendungen werden bei den Fallstudienpartnern prototypenhaft integriert bzw. auf situative Daten angewandt. Dabei wird stetig dokumentiert, welche Einführungsschritte gut abgelaufen sind und welche Verbesserungspotenziale aufzeigen, um die Einführungsstrategie zu validieren**.** Die validierte Einführungsstrategie wird anschließend zusammen mit der Roadmap aufbereitet, sodass diese anwenderfreundlich zur Verfügung gestellt werden kann.	Die in AS 1 umgesetzten Machine Learning-Anwendungen wurden bei den Fallstudienpartnern prototypenhaft integriert bzw. auf situative Daten angewandt. Dabei wurde stetig dokumentiert, welche Einführungsschritte gut abgelaufen sind und welche Verbesserungspotenziale aufzeigen, um die Einführungsstrategie zu validieren. Die validierte Einführungsstrategie wurde anschließend zusammen mit der Roadmap aufbereitet, sodass diese anwenderfreundlich zur Verfügung gestellt werden kann. Dazu wurde das in AP 3 entwickelte Webtool angepasst und weiterentwickelt.
Im dritten Arbeitsschritt wird ein Experimentaldesign mit zugehörigen Hypothesen entwickelt, in dem der Einfluss der Art und der Qualität zur Verfügung stehenden Daten auf die Akzeptanz von Machine Learning untersucht wird. Die Hypothesen werden hinsichtlich ihrer Praxistauglichkeit mit dem projektbegleitenden	Im dritten Arbeitsschritt war ein Experiment geplant, um den Einfluss der Art und der Qualität der zur Verfügung stehenden Daten auf die Akzeptanz von Machine Learning zu untersuchen. Seit Beantragung des Projekts gab es im Bereich der Algorithmenaversion bedeutende Fortschritte und eine Vielzahl von

Ausschuss validiert. Für das Experiment werden Teilnehmer vor allem unter den Mitarbeitern der Fallstudienpartner akquiriert. Sollten nicht genug Rückläufer generiert werden können, wird auf das Kontaktnetzwerk der durchführenden Forschungseinrichtungen und die berufsbegleitenden Studierenden des Studiengangs Business Analytics zurückgegriffen. Abschließend werden Maßnahmen entwickelt, die zum einen die Akzeptanz von Machine Learning bei den betroffenen Mitarbeitern steigern und zum anderen die Voraussetzungen für den sinnvollen Einsatz von Machine Learning schaffen sollen. Diese werden mit dem projektbegleitenden Ausschuss validiert und in einem Katalog zur direkten Anwendung in KMU festgehalten.	Experimenten, die die Forschungslücken der geplanten Fragestellungen bereits gefüllt haben. Stattdessen wurde auf die Entwicklungen reagiert, die teilweise diffuse und uneinige Ergebnisse hervorgebracht haben. Es wurde eine Meta-Analyse erstellt, die die verschiedenen Ergebnisse untersucht, um ein generelles Level von Algorithmenakzeptanz zu ermittln. Dazu wurde eine systematische Literaturrecherche nach relevanten Papern durchgeführt, die sich auf verschiedenste Anwendungsbereiche erstreckt. Aus den gefundenen Papern wurden alle relevanten Ergebnisse bzgl. der Algorithmenakzeptanz extrahiert, in einer Datenbank zusammengefasst, nach gängigen Standards von Meta-Analysen ggf. umgewandelt und analysiert. Die Ergebnisse zeigen keine Existenz einer generellen Algorithmenaversion. Die Ergebnisse wurden mit dem PA diskutiert, der bereits im Vorfeld keine generelle Algorithmenaversion bei Mitarbeitern beobachten oder problematisieren konnte.

2.4.1 Beschreibung der Fallstudien

Nicht bei allen Fallstudien wurden Freigaben für die Verwendung des Firmennames gegeben. In diesen Fällen wurden generische Namen für die Branche gewählt, z. B. die Rohrreinigung GmbH.

Rohrreinigung GmbH

Die Rohrreinigung GmbH ist ein Unternehmen, das sich auf die Reinigung von Rohren spezialisiert hat. Das Unternehmen ist in Stuttgart und Umgebung tätig und bietet 24-Stunden-Service an. Die Firma hat sich auf die Beseitigung von Verstopfungen und Rückstaus in Abwasserleitungen spezialisiert. Bei der Rohrreinigung GmbH wird bei der Analyse von Rohren eine spezielle Lösung eingesetzt. Dabei fährt ein Fahrzeug mit Kamera in das Rohr und erstellt ein Bild. Der Analyst untersucht das Bild und vermerkt relevante Stellen wie Risse, Löcher oder Fremdmaterial

im Rohr, die im Bericht auftauchen sollten. Um den Analysten zu unterstützen, wurde eine Automatisierung von relevanten Stellen angestrebt.

Datengrundlage und -beschaffung

Bei der Analyse von Rohren werden mit Hilfe des Fahrzeugs und einer Fernsteuerung Videos und Bilder aufgenommen, in einer Analysesoftware verarbeitet und anschließend in einen Bericht exportiert. Alle Dateien werden in einer bestimmten Ordnerstruktur abgespeichert und nach abgeschlossener Analyse auf dem Server des Unternehmens nach Jahr, Einsatznummer und Adresse abgespeichert. Die Dateien enthalten u. a. eine Excel-Datei mit allen Befunden, abgespeicherte Bilder sowie mehrere technische Dateien der Analysesoftware. Für die Vorbereitung der Erhebung der relevanten Daten wurden 5 Fälle via OneDrive mit den Forschungseinrichtungen geteilt. Nachdem das Muster der Speicherung analysiert und ein Programm zur Gewinnung der Dateien geschrieben wurde, wurden bei einem Besuch vor Ort 465 Analyseordner auf eine externe Festplatte kopiert, damit diese in den Forschungseinrichtungen analysiert werden können.

Preprocessing

Bei der Aufbereitung der Daten gab es mehrere Schwierigkeiten.

Für das Training eines Object Detection Netzwerks werden Bilder (immer der gleichen Größe) und sog. Label (Angaben von Boxen, in denen Befunde liegen) benötigt. Die ursprüngliche Sichtung der Daten versprach eine Zuordnung von Fehlerbefunden zu Bildern, in denen über eine Ortsangabe in den Bildern Label erstellt werden sollten (s. Abbildung 16).

Material	Profilart	Profilhöhe (Durchmesser) in mm	Profilbreite in mm	Untersuchungsnummer	Untersuchungsrichtung	Stationierung in m	Steuertext	Schadentext	Langtext	Num. Zusatz	Einheit
STZ	Kreisförmig, Kreisquerschnitt	150	150	1	in	00,0		BCDXP	Anfangsknoten, Inspektionsanfang		
STZ	Kreisförmig, Kreisquerschnitt	150	150			01,0		BAJB	Verschobene Verbindung, radial, 20 mm, 8-2 Uhr	20	mm
STZ	Kreisförmig, Kreisquerschnitt	150	150			01,2		BCAAA	Anschluss, Abzweig, Anschluss offen, 125 mm, 2 Uhr	125	mm
STZ	Kreisförmig, Kreisquerschnitt	150	150			01,2		BDB	Grundstichleitung, Hausseite links		
STZ	Kreisförmig, Kreisquerschnitt	150	150			01,4		BAJB	Verschobene Verbindung, radial, 20 mm, 3-9 Uhr	20	mm
STZ	Kreisförmig, Kreisquerschnitt	150	150			02,3		BABBA	Rissbildung, Riss, in Längsrichtung, 10-2 Uhr		
STZ	Kreisförmig, Kreisquerschnitt	150	150			02,3		BAJB	Verschobene Verbindung, radial, 20 mm, 9-3 Uhr	20	mm

Abbildung 16: Geplante Datengrundlage der Rohrreinigungs GmbH

Bei der Analyse wurde festgestellt, dass einzelne Zeilen in der Excel-Datei nicht eindeutig zu Bildern zugeordnet werden können, da teilweise mehrere Bilder oder kein Bild zu einem Befund gespeichert werden und der Name der Bild-Datei nicht in der Zeile gespeichert wird.

Nach einiger Zeit wurden daher andere Dateien hinsichtlich ihrer Verwendbarkeit untersucht und nach längeren iterativen Versuchen eine sog. Constant Database (.cdb) gefunden, die alle Informationen beinhält. Constant Database ist ein Dateityp, der hauptsächlich auf mobilen Geräten verwendet wird (Android, Nokia Phones etc.) und mit python nicht einfach analysiert werden kann. Daher wurde ein byte-basierter Ausleseansatz gewählt, der die Dateien Zeichen für Zeichen auseinandernimmt und interpretiert. Dabei wurde festgestellt, dass eine bestimmte Zeichenfolge der Beginn einer Feststellung ist und in diesem Abschnitt eindeutig nach dem Namen der Feststellung sowie dem Namen des zugehörigen Bildes gesucht werden kann.

Abbildung 17: Beginn einer Feststellung (dargestellt in einem Editor)

Nach diesen Schwierigkeiten konnten alle Algorithmen angepasst werden. Der Ablauf und die in jedem Schritt gewonnenen Anzahlen von Dateien sind in Tabelle 5 aufgelistet. Damit konnten 2.243 Feststellungen in einer csv gesammelt werden, wobei jeder Zeile eine eindeutige Bilddatei zugeordnet werden kann.

Tabelle 5: Ablauf der Datenextraktion bei der Rohrreinigung GmbH

Vorgang	Ausschluss	Arbeitsdateien	Anmerkung
Ordnerauflistung		465 Ordner	
Existiert Exceldatei?	145 ohne .xls, 2 mit 2 .xls	320 Ordner	Teilweise scheinen zwei Untersuchungsordner in einem zu liegen; können u. U. ergänzt werden
Existieren Viewerdateien?	0	320 Ordner	
Folgt die Exceldatei dem üblichen Format?	1	319 Ordner	

Extraktion aller Einzeluntersuchungen aus den Exceldateien		1289 Untersuchungen	
Zuordnung von notwendigen .cdb-Dateien	125 Untersuchungen ohne .cdb-Datei	1164 Untersuchungen	Das Format, in dem der Name der Untersuchung angegeben war, variiert stark. Mit einigem Aufwand könnte man verfeinert nach noch mehr Formaten suchen. Ausschuss ist aber negierbar.
Zuordnung von Bildern zu den Untersuchungen	Aus 3 .cdb-Dateien konnten keine Bilder extrahiert werden	1161 Untersuchungen	
Extraktion einzelner Feststellungen aus den Untersuchungen		12625 Feststellungen	
Filtern nach Untersuchungen, denen ein Bild zugeordnet werden konnte	7332	5293 Feststellungen	Mit etwas Aufwand könnte über Videozeit untersucht werden, ob ein Bild mehrfach verwendet werden kann.
Extraktion von Untersuchungen mit Langtext	5	5288 Feststellungen	
Extraktion von Untersuchungen, bei denen eine Position vermerkt war	3039	2249 Feststellungen	Man könne auch alle Bilder verwenden und bei Feststellungen ohne

			Positionsvermerk das gesamte Bild labeln.
Suchen der in .cdb-Dateien vermerkten Bildern in den jeweiligen Ordnern	6	2243 Feststellungen	Das Format, in dem der Name des Bildes in .cdb-Dateien angegeben war, variiert stark. Mit einigem Aufwand könnte man verfeinert nach noch mehr Formaten suchen. Ausschuss ist aber negierbar.

Machine Learning Ansatz

Ziel der Anwendung ist die Unterstützung von Analysten, sodass ein Befund automatisch erkannt und nur noch akzeptiert statt selbst erstellt werden muss. Dafür wird ein Object Detection Network trainiert, welches diesen Anwendungsfall zum Ziel hat. Dabei werden von einem YOLO-Netzwerk (You Only Look Once) in einem Bild verschiedene Objekte erkannt, nachdem es auf dem oben beschriebenen Datensatz trainiert wurde.

Ergebnisse und Verwendung

Abbildung 18 zeigt beispielhafte Anwendungen des trainierten Erkennungsnetzwerks bei „neuen“ Bildern (also solche, die das Netz noch nie gesehen hat, sodass es sich auf sie hätte trainieren können). Offensichtlich wird, dass die Erkennung von Problemen generell sehr gut funktioniert, aber die Umrandungsboxen sehr ungenau sind. Das war dadurch zu erwarten, dass die Label aufgrund der ungenauen Angabe von Uhrzeiten nur sehr ungenau gestaltet werden konnten.

Abbildung 18: Beispielanwendung des trainierten Netzwerks

Insgesamt sind die Ergebnisse für die zugrundeliegenden Daten sehr zufriedenstellend. Für alle unterschiedlichen Kategorien liegt R bei 0.312, die mAP50 (mean average precision bei einer Konfidenz von 50) bei 0.376.

Während der Fallstudie hat der Ansprechpartner das Unternehmen verlassen, sodass mit dem Netz nicht weitergearbeitet werden konnte. Mögliche Verbesserungen liegen in dem genaueren Labeling (z. B. durch manuelles Begrenzen) der Probleme sowie in dem Einholen mehrerer Daten.

Reiser AG Maschinenbau

Die Reiser AG hat sich auf die Herstellung von Präzisionsteilen und Baugruppen spezialisiert. Sie bieten Einzelteile aus Metall für verschiedene Branchen, Montage von komplexen Aggregaten sowie 3D-Druck in Metall und Kunststoff. Der Maschinenbauer verwendet Wendeschneidplatten in CNC-Maschinen, die sich über die Zeit hinweg abnutzen. Um den optimalen Zeitpunkt für den Austausch der Verschleißteile zu bestimmen, muss die Qualität der einzelnen Schneiden bemessen werden. Da die Teile sehr klein sind, ist es für Menschen oft schwierig, die Qualität korrekt zu bemessen. Aus diesem Grund soll die Qualität der Schneiden automatisiert erfasst werden.

Datengrundlage und -beschaffung

Die Daten für die Fallstudie bestanden aus Bildern von Wendeschneidplatten. Insgesamt 380 Bilder zu Schneidertyp A, 66 zu Typ B und 84 von Typ C wurden zur Verfügung gestellt (s. Abbildung 19).

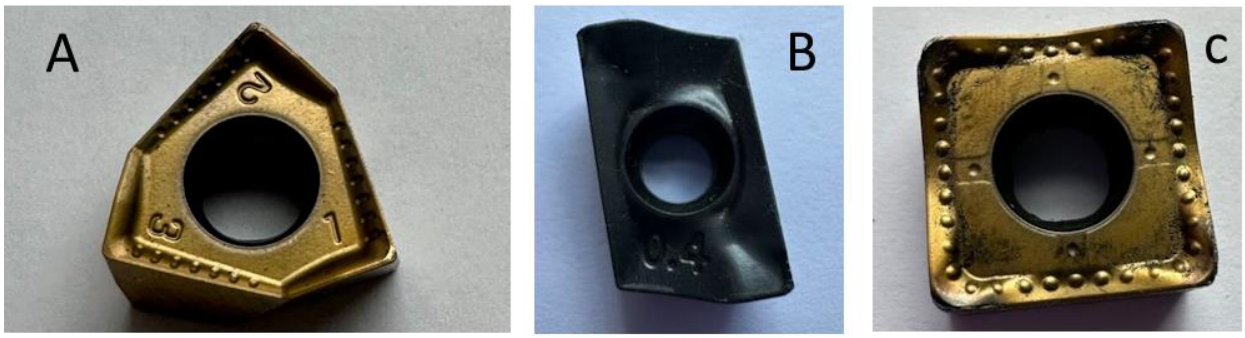


Abbildung 19: Beispielbilder der Wendeschneidplatten A, B und C

Zum Training der Analysenetze sind zusätzlich Qualitätswerte nötig. Diese wurden im ersten Schritt als 1 für neue Teile, 0.5 für Teile, die aktuell in Maschinen in Verwendung waren und 0 für Ausschussteile definiert. Diese Einteilung unterliegt aber mehreren Fehlern:

- Maschinenteile haben in der Regel keine 50 % Restqualität
- Ausschussteile müssen keine 0 % Restqualität haben, sondern können zu früh aussortiert worden sein.

Durch diese Fehlevaluation wurden die Netze falsch trainiert und konnten in der Praxis nicht funktionieren.

Im zweiten Schritt wurden die einzelnen Schneider vom Analysten selbst bewertet, um einen Proof of Concept zu erreichen. Dies führte zu einem besseren Training, war aber durch fehlendes praktisches Wissen vom Analysten noch immer fehlerhaft. Abschließend wurden Werte von Experten in der Fabrikhalle aufgenommen, mit denen die Netze nachtrainiert werden konnten.

Preprocessing

Für den gewählten Machine Learning Ansatz müssen alle Bilder die gleiche Größe haben. Dafür wurden in allen Fällen rechts und links jeweils die Hälfte der Überbreite im Vergleich zur Höhe abgeschnitten und die resultierenden Bilder auf 640x640 Pixel skaliert.

Machine Learning Ansatz

Der Machine Learning Ansatz ist insgesamt dreiteilig, beinhaltet aber in jedem Fall konvolutionelle neuronale Netzwerke. Diese werden verwendet, um mehrdimensionale Input-Daten zu verarbeiten. Im Fall von Bildern sind Inputdaten dreidimensional (640 x 640 x 3 (Farbe eines Pixels in RGB)). Die Einsatzgebiete von konvolutionellen neuronalen Netzen bei Bildern lassen sich in vier Gebiete einteilen:

➔ **Image Classification:** Ein Bild wird einer Klassifizierung (z. B. Hund, Katze, Verkehrsschild) zugeordnet

- ➔ **Object Detection:** Auf einem Bild wird eingegrenzt, wo sich ein bestimmtes Objekt befindet (wenn sich z. B. ein Hund und eine Katze auf dem gleichen Bild befinden)
- ➔ **Object Segmentation:** Jeder Pixel eines Bilds wird eindeutig einem Objekt auf dem Bild zugeordnet
- ➔ **Simple Prediction:** Ein Bild wird auf eine einzelne Kennzahl reduziert.

Im ersten Schritt müssen die drei verschiedenen Typen von Schneidern einem Typ zugeordnet werden. Dafür wurde ein Image Classification Netzwerk trainiert, welches mit einer 100 % Genauigkeit den Typ eines Schneiders erkennen kann.

Im zweiten Schritt müssen die einzelnen Schneider aus dem Gesamtbild extrahiert werden. Dafür wurde ein Object Detection Netzwerk (YOLOv5) trainiert, welches die einzelnen Schneider mit einer 100 % Genauigkeit erkennt und diese sehr gut ausschneidet (s. Abbildung 20).

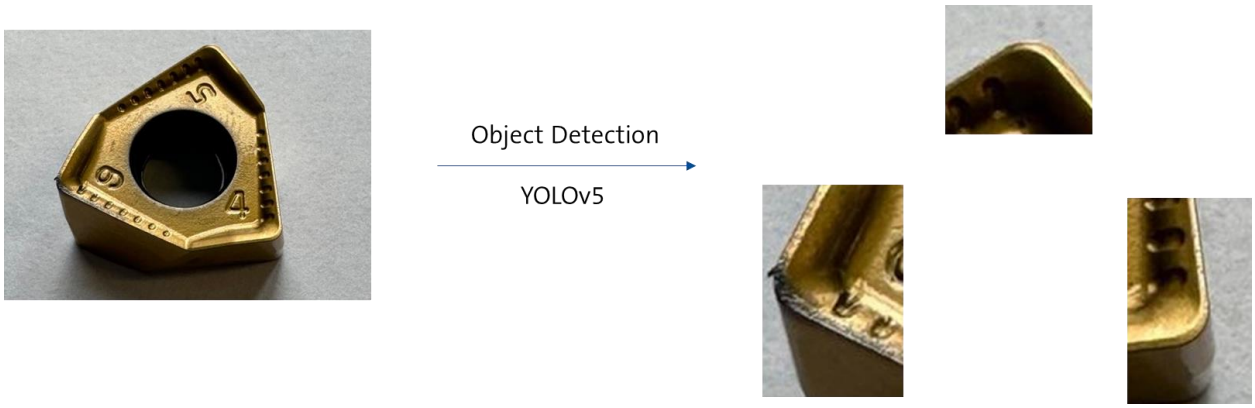


Abbildung 20: Funktion des Object Detection Netzwerks

Anschließend müssen die einzelnen ausgeschnittenen Bilder auf eine einzelne Kennzahl reduziert werden. Dazu musste zunächst wieder sichergestellt werden, dass die Bilder dieselbe Größe haben, da sich ein konvolutionelles neuronales Netzwerk sonst nicht trainieren lässt. In diesem Schritt konnte allerdings nicht beliebig abgeschnitten werden, da sonst relevante Teile des Schneides gelöscht werden könnten. Stattdessen wurden die Bilder quadratisch mit weißer Farbe aufgefüllt und auf 64x64 Pixel skaliert.

Von diesem Punkt an konnte ein Netzwerk zur Simple Prediction trainiert werden. Dafür wurde ein VGG19 mit Hilfe von Transfer Learning auf den Anwendungsfall angepasst. Das trainierte Netz weist dann jedem Einzelbild eine Vorhersage zu (s. Abbildung 21).

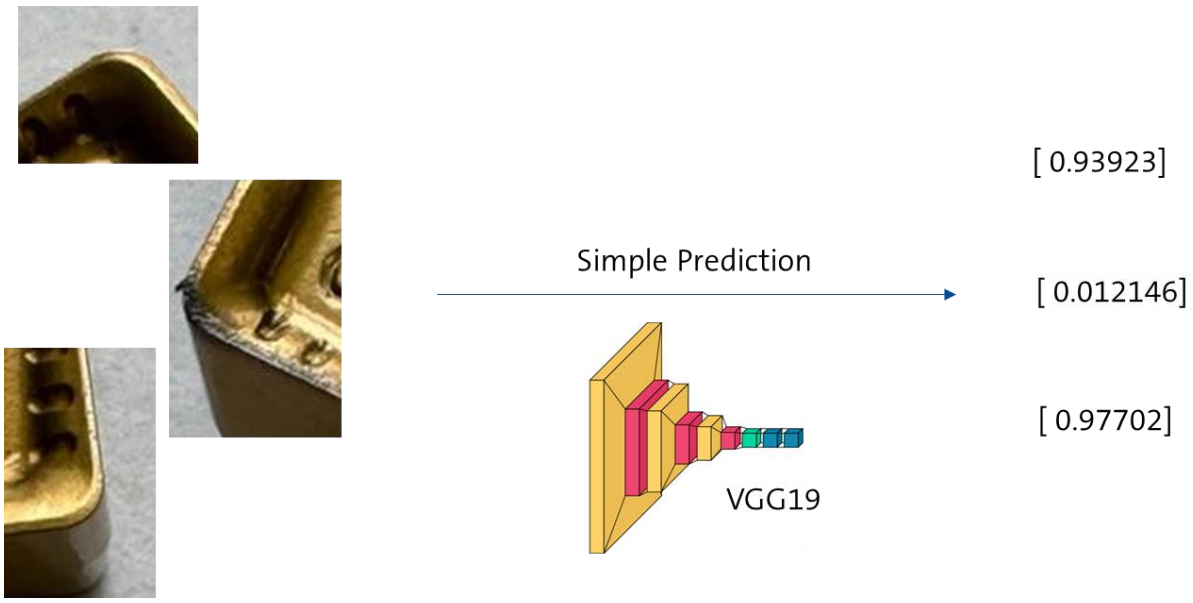


Abbildung 21: Funktion des Simple Prediction Netzwersks

Da sich die Schneider der Typen A und C ähneln, wurde das Simple Prediction Netzwerk für beide Schneider gleichermaßen trainiert. Typ B war durch die andere Form und Farbe zu unterschiedlich, sodass hier ein eigenes Netzwerk trainiert wurde.

Ergebnisse und Verwendung

Alle Netzwerke hatten eine hinreichend hohe Qualität in der Auswertung der Bilder. Mit den Netzen wurde daher eine prototypische Umsetzung bei Reiser vor Ort getestet. Dafür wurde ein python-Skript geschrieben, das den gesamten Analyseablauf beinhaltet, die Ergebnisse in einer kleinen GUI darstellt und Verbesserungsmöglichkeiten zulässt, sodass die Netze mit den Meinungen der Experten in der Fabrikhalle nachtrainiert werden können. Abbildung 22 zeigt die physische Umsetzung. Ein erhöhtes Handy filmt eine befestigte Ablage in einer Fotobox, in der die Wendeschneidplatten eingelegt werden können. Das Handy streamt die Kameraaufnahme via IP Webcam in das Netzwerk vor Ort. Ein von Reiser zur Verfügung gestelltes Notebook verarbeitet dann diese Aufnahme und wendet das python-Skript an. Abbildung 23 zeigt die digitale Funktion. In der linken Bildschirmhälfte wird die IP Webcam angezeigt, sodass evtl. falsch eingelegte Wendeschneidplatten direkt korrigiert werden können. Auf der rechten Bildschirmhälfte läuft das python-Skript in einer von Anaconda verwalteten Virtual Environment. In der Mitte ploppt nach Anwendung des Skripts die oben beschriebene GUI auf.

Die prototypische Umsetzung wurde mit vier Mitarbeitern besprochen und auf Verbesserungsmöglichkeiten getestet. Mit den verbesserten Werten, die die Mitarbeiter eingegeben haben, wurden die Netze ein letztes Mal nachtrainiert, eingespielt und die Fallstudie im Rahmen von MIready damit beendet.

Abbildung 22: Physische Umsetzung

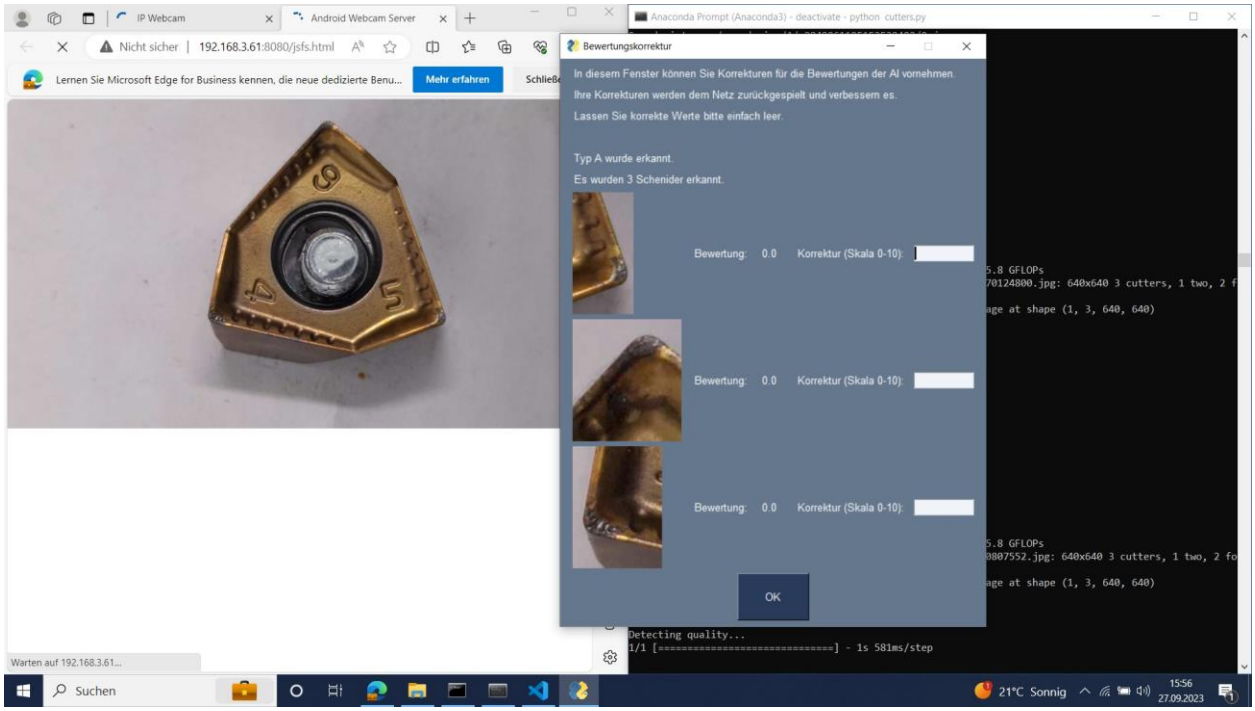


Abbildung 23: Digitale Umsetzung

EUCHNER ist ein Experte für industrielle Sicherheitstechnik – Lösungen und Produkte, die trennende Schutzeinrichtungen an Maschinen und Anlagen absichern. Der Hersteller von Smart Safety-Lösungen testet jedes Teil intensiv, um sicherzustellen, dass alle ausgelieferten Produkte voll funktionsfähig sind. Während der Produktion werden bereits weitere Messwerte von Komponenten aufgenommen. Wenn ein Zusammenhang zwischen Messwerten während der Produktion und Testergebnis des Endprodukts besteht, kann die Verarbeitung bestimmter Teile u. U. vorzeitig abgebrochen werden, um Ressourcen und Zeit zu sparen, die bei der Wiederaufarbeitung verloren gehen. Der Zusammenhang zwischen den Messwerten soll untersucht werden.

Datengrundlage und -beschaffung

Für die Analyse wurden Messdaten von Komponenten mit zugehörigem Testergebnis, Messdaten von Endprodukten mit zugehörigem Testergebnis sowie eine Datei mit Zuordnungen von Komponenten und Endprodukten über einen Zeitraum von einem halben Jahr zur Verfügung gestellt. Die Daten werden vom Fallstudienpartner aufgenommen, zentral gespeichert und konnten so leicht beschaffen und versendet werden.

Preprocessing

Die Daten waren nahezu perfekt, sodass kaum Preprocessing notwendig war. Drei Datensätze mussten so kombiniert werden, dass das Testergebnis eines Endprodukts mit den Messwerten der zugehörigen Komponenten verbunden wird. Beachtet werden musste, dass eine Komponente in mehreren Teilen verbaut sein können (wenn z. B. ein Endprodukt nicht korrekt geprüft und daher auseinandergebaut und neu verwertet wurde) und dass ein Endprodukt mehrere Komponenten desselben Typs beinhalten kann (wenn die Komponente ausgetauscht werden musste). Abgesehen davon konnte der benötigte Datensatz durch einfach merge-Operationen erstellt werden.

Machine Learning Ansatz

Zunächst wurde ein Process Mining-Ansatz gewählt, um die Abläufe besser zu verstehen. Abbildung 24 zeigt den Verlauf der Komponenten und zugehörigen Endprodukte. Einstellvorrichtung 0/1 steht dabei für das (nicht) erfolgreiche Testen der Komponente, Prüfvorrichtung 0/1 für das (nicht) erfolgreiche Testen der Endprodukte. Fälle, die nach der Einstellvorrichtung bzw. nach Prüfvorrichtung 0 enden, könnten solche sein, die zum Ende des Datensatzes hin begonnen und danach nicht weiter verarbeitet wurden, sodass diese nicht unbedingt zu beachten sind.

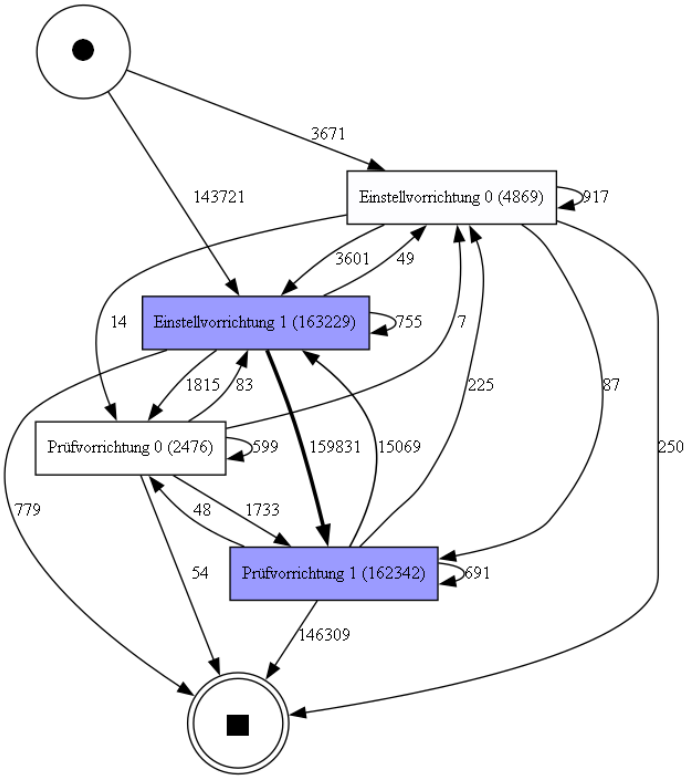


Abbildung 24: Verlauf der Produkte durch Testanlagen

Interessanter sind die Abläufe zwischen Einstellvorrichtung und Prüfvorrichtung, sodass z. B. in 7 Fällen nach der fehlerhaften Prüfung der Endprodukte die Komponente nicht mehr korrekt getestet werden konnte. Mehrere Prüfungen oder der Verlauf von Prüfvorrichtung 1 -> Einstellvorrichtung 0 sind durch die Verzweigungen von mehreren Komponenten auf mehrere Endprodukte zu erklären.

Nach dem Aufbau eines besseren Verständnis der Abläufe wurde ein Entscheidungsbaum trainiert, der basierend auf den Messwerten der Komponenten das Testergebnis des Endprodukts prognostizieren soll. Dazu wurden die Daten typischerweise in Trainings- und Testdaten aufgeteilt (90/10), um das Modell validieren zu können. Außerdem wurde ein neuronales Netzwerk 3 Hidden Layern (30, 10, 10; jeweils Dense mit Aktivierungsfunktion relu) getestet, das allerdings keine besseren Ergebnisse hervorbringen konnte.

Ergebnisse und Verwendung

Abbildung 25 zeigt einen Ausschnitt des Entscheidungsbaums.

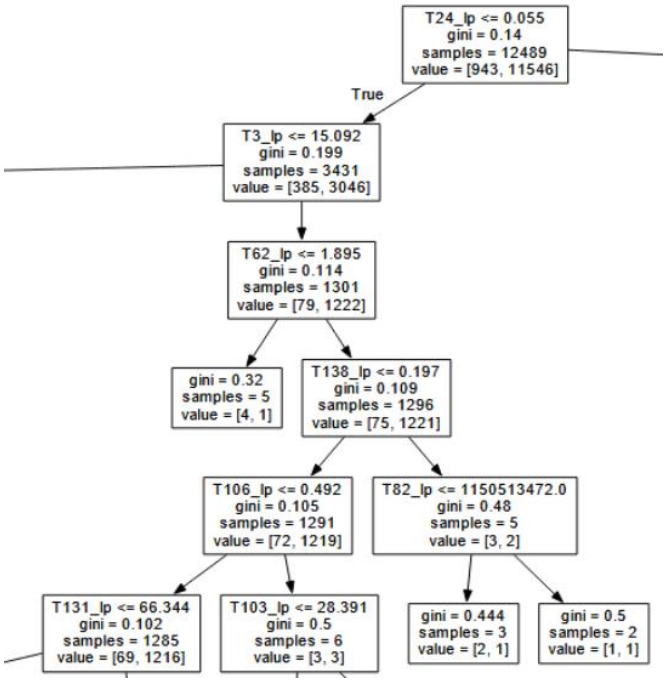


Abbildung 25: Entscheidungsbaum zur Prognose des Testergebnisses des Endprodukts

Auf den Testdaten wurden mit einer Tiefe von 7 folgende Gütewerte erreicht:

	precision	recall	F1-score	support
0	0.46	0.06	0.11	100
1	0.92	0.99	0.96	1149
accuracy			0.92	1249
macro avg	0.69	0.53	0.53	1249
weighted avg	0.89	0.92	0.89	1249

Das R^2 liegt bei knapp 2 %, was nicht verwunderlich ist, wenn man nur eine von 29 Komponenten zum Forecast verwendet.

Das Ergebnis wurde an den Fallstudienpartner zurückgespielt, um eine physikalische Erklärbarkeit zu validieren. Nach längerer Entscheidung kamen die Experten zu der Entscheidung, dass die Variable T24 keinen Zusammenhang mit dem Testergebnis haben könne und dass das Ergebnis zufällig gefunden sein muss. Zum Abgleich wurde ein weiterer Entscheidungsbaum auf allen bis dahin verfügbaren Daten trainiert und auf Daten in den darauffolgenden 3 Monaten angewandt. Tatsächlich konnte die recht hohe Güte nicht repliziert werden, sodass von einem zufälligen Zusammenhang ausgegangen werden muss. Abschließend wurden die einzelnen

Variablen genauer untersucht und Trends, Verläufe und Saisonalität festgestellt (sowohl im Testergebnis als auch in den Messwerten), die die zufällig gefundenen Ergebnisse erklären könnten. Als Ergebnis der Fallstudie bleibt daher die Erkenntnis, dass die Testanlage der Komponenten offenbar fein genug justiert ist, dass keine korrekt geprüfte Komponente einen drastischen Einfluss auf das Endprodukt hat.

Medkomp GmbH

Die Medkomp GmbH bietet eine umfassende Palette an Dienstleistungen rund um Medikamentenverabreichungsgeräte an. In der Produktionshalle kommt es häufig zu Maschinenfehlern. Um diese Fehler frühzeitig zu erkennen und schnell korrigieren zu können, sollen Sensordaten der Maschinen genutzt werden, um vorherzusagen, wann aus welchen Gründen welche Fehler auftreten werden.

Datengrundlage und -beschaffung

Medkomp sammelt Daten von sekündlichen Aufnahmen (Sensordaten, Antriebsdaten, Verbrauchsdaten, Fehlerdaten) in ihrem ERP und hat Daten über 2 Tage hinweg freigegeben.

Preprocessing

Die Daten waren nahezu perfekt. Durch Aufnahme- und Uploadfehler kam es zeitweile zu fehlenden Daten sowie mehreren Dateneinträgen zum selben Zeitstempel. Diese mussten besonders beachtet werden. Doppelte Einträge wurden gelöscht, wobei der Mittelwert der beiden Einträge behalten wurde. Die einzelnen Datensätze (Sensordaten, Antriebsdaten und Verbrauchsdaten) mussten dann via Zeitstempel verbunden werden, um einen einheitlichen Datensatz zu erhalten.

Machine Learning Ansatz

Zunächst wurde die Anzahl der Variablen über einen selbstgeschriebenen recurseve feature elimination (RFE) Algorithmus verringert, da die Korrelation zwischen einigen Variablen sehr hoch war. Abbildung 26 zeigt die Korrelationsmatrix der Variablen vor der Eliminierung, Abbildung 27 danach.

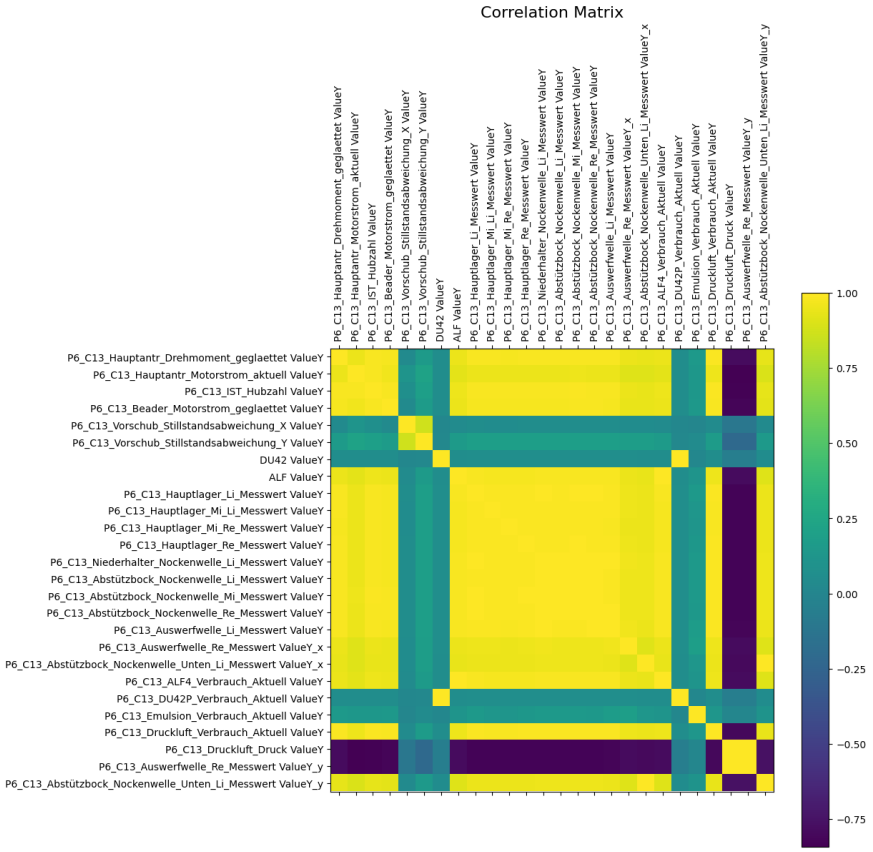


Abbildung 26: Korrelationsmatrix vor RFE

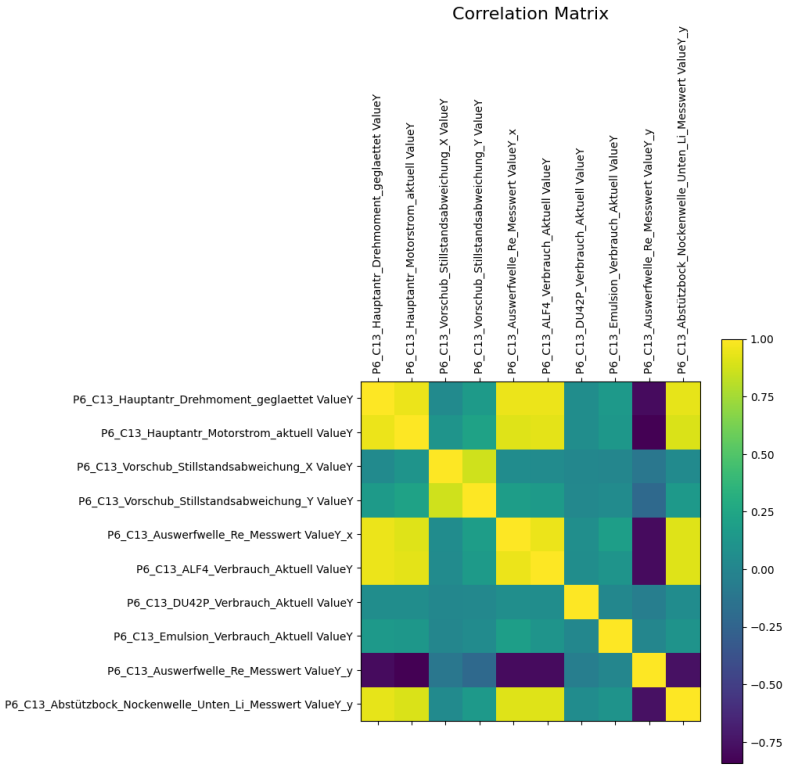


Abbildung 27: Korrelationsmatrix nach RFE

Nach der Eliminierung wurden mit weniger Variablen übersichtlichere Schaubilder über den Verlauf vor und kurz nach einem Fehler erstellt, damit Zusammenhänge besser verstanden werden können. Für den häufigsten Fehler (Kurzschluss Trafo Luftreiniger) sind zwei Verläufe beispielhaft in Abbildung 28 dargestellt. Alle Verläufe enthalten den Anstieg des Motorstroms des Hauptantriebs nach dem Kurzschluss.

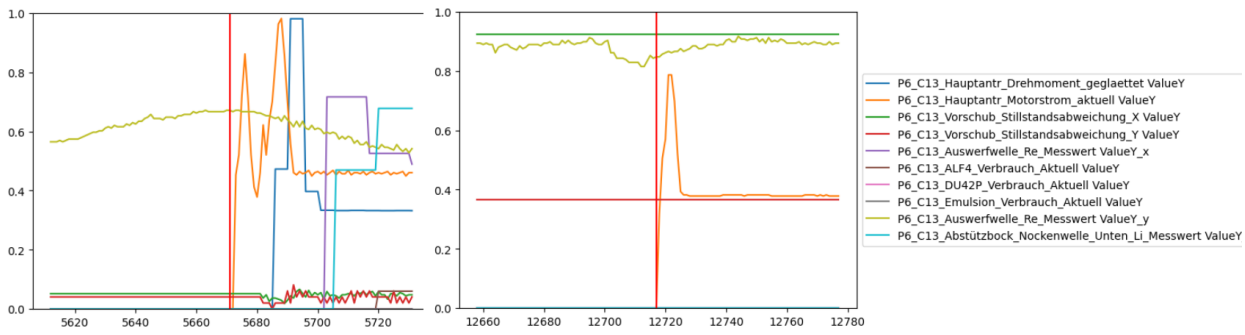


Abbildung 28: Verlauf vor und nach Kurzschlüssen

Mit den Ergebnissen wurde Rücksprache mit Experten aus den Fachabteilungen gehalten, die den Verlauf physikalisch erklären konnten. Durch einen Kurzschluss im Luftreiniger stoppt die Presse, der Hauptantrieb dreht aber kurzfristig durch. Die Experten haben außerdem angemerkt, dass die bisher besprochenen interessanten Probleme u. U. noch nicht in den Störungen abgebildet sind und stattdessen andere Betrachtungen wie physische Elemente in den Anlagen interessanter sein könnten, da z. B. ein Putzlappen im Antrieb in der Vergangenheit zu einem hohen Maschinenschaden geführt haben.

Ergebnisse und Verwendung

Die Einblicke wurden dem Fallstudienpartner zurückgespielt. Eine physikalische Erklärung für den Anstieg des Motorstroms konnte vom Fallstudienpartner nicht at hoc geliefert werden. Ein Kurzschluss im Luftreiniger bedeutet, dass die Walze aufhört zu produzieren. Der Motor allerdings könnte durchdrehen und daher anfänglich der Strom steigen, bevor er sich normalisiert.

Maru Finest Sushi liefert täglich frisches Sushi an verschiedene Filialen in Norddeutschland. Der Nahrungsmittelhersteller benötigt genaue Informationen über die Absatzmenge, da nur kurz haltbare Nahrungsmittel wie Sushi nach kurzer Zeit weggeworfen werden müssen, wenn sie nicht verkauft werden. Daher ist es für den Hersteller von größter Wichtigkeit, die Absatzmenge in jeder Filiale genau zu prognostizieren.

Datengrundlage und -beschaffung

Die Datengrundlage (EAN und Datum) für die Analyse besteht aus Textdateien, die die täglichen Verkaufszahlen pro Produkt im Zeitraum von Mai 2022 bis Januar 2023 enthalten. Die Beschaffung der Daten erfolgte durch den Zugriff auf das Kassensystem, das die täglichen Verkaufszahlen für jedes Produkt erfasst und bereitstellt. Diese Daten dienen als Ausgangspunkt für weitere Analysen, Modellierung und Prognosen im Rahmen des entwickelten Machine-Learning-Ansatzes.

Preprocessing

In einem ersten Schritt wurden die Daten aus mehreren TXT-Dateien im angegebenen Verzeichnis eingelesen und zusammengeführt. Der neu entstandene Datensatz wurde anschließend in ein einheitliches Format überführt und optimiert. Fehlende Daten wurden ergänzt und fehlerhafte Daten gelöscht, so dass ein vollständiger Datensatz mit optimierter Qualität entstand. Im nächsten Schritt wurden Datumseigenschaften wie Wochentag, Monat, Jahr, Wochenend- und Feiertagskennzeichen hinzugefügt. Um den Einfluss des Wetters untersuchen zu können, wurden Wetterdaten aus einer csv-Datei eingelesen.

Machine Learning Ansatz

Zunächst wurden Zeitreihen für jedes Produkt (EAN) separat mit Hilfe von TimeSeriesSplit, einem scikit-learn-Modul, erstellt. Die Zeitreihen wurden für jedes Produkt aufgeteilt, um sicherzustellen, dass das Modell auf den spezifischen Produktverkäufen trainiert wird. Trainings- und Testzeitreihen wurden ebenfalls mit TimesSeriesSplit erstellt, um sicherzustellen, dass die Validierung mit zukünftigen Daten durchgeführt wird. Anschließend wurden die Sequenzen erstellt, um mit ihnen das zu entwickelnde Modell zu trainieren. Das verwendete Modell ist ein LSTM (Long Short-Term Memory) Neuronales Netz. Das LSTM-Modell besteht aus mehreren Schichten, darunter LSTM-

Schichten und Dense-Schichten. Optional kann eine Hyperparameteroptimierung durchgeführt werden, um die besten Werte für die Anzahl der Einheiten, die Dropout-Rate und die Lernrate zu finden. In diesem Fall wurde die Optuna-Bibliothek verwendet. Die EANs wurden vor dem Training und der Vorhersage mit One-Hot-Kodierung kodiert. Nach dem Training konnte das Modell zur Vorhersage zukünftiger Produktverkäufe verwendet werden. Die Vorhersagen wurden durch Umkehrung der zuvor durchgeführten Skalierung gewonnen, um die ursprünglichen Verkaufszahlen zu erhalten. Das trainierte Modell sowie der One-Hot-Kodierer und die Skalierer wurden optional gespeichert, um sie später für Vorhersagen zu verwenden, ohne das Modell erneut trainieren zu müssen.

Ergebnisse und Verwendung

Das entwickelte Modell wurde erfolgreich in eine benutzerfreundliche grafische Benutzeroberfläche (GUI) integriert. Diese GUI ermöglicht es Benutzern, den Prognosezeitraum sowie das spezifische Produkt für die Verkaufsprognose auszuwählen. Die Ergebnisse der Prognose werden in Form eines Diagramms anschaulich präsentiert. Zudem besteht die Möglichkeit, neue Datensätze hochzuladen, die dann für ein erneutes Training des Modells verwendet werden.

Zusätzlich verfügt die GUI über eine Heatmap-Funktionalität, die die Korrelationsmatrix zwischen verschiedenen ausgewählten Merkmalen und den Verkaufszahlen visualisiert. Diese Heatmap bietet einen Einblick in die Beziehungen zwischen verschiedenen Faktoren (z.B., Wochentag, Temperatur, Feiertage) und den Verkaufszahlen (siehe Abbildung 29). Dies kann dazu beitragen, Muster und Zusammenhänge in den Daten zu erkennen.

Abbildung 29: Benutzeroberfläche der Absatzprognose mit Heatmap Funktionen

Leider war die Prognosegenauigkeit des Modells nicht zufriedenstellend. Dies kann hauptsächlich auf die begrenzte Datengrundlage zurückgeführt werden, da lediglich Verkaufsdaten für einen Zeitraum von neun Monaten vorhanden waren. Die Qualität von Zeitreihenprognosen hängt stark von der Verfügbarkeit einer ausreichenden Menge historischer Daten ab. Mit nur neun Monaten historischer Daten könnte das Modell Schwierigkeiten haben, komplexe Muster zu erkennen und genaue Vorhersagen für zukünftige Verkaufszahlen zu generieren. Eine Erweiterung des historischen Datensatzes könnte dazu beitragen, die Prognosegenauigkeit zu verbessern und dem Modell eine bessere Grundlage für das Erlernen von Muster zu bieten.

Im Rahmen der durchgeführten Fallstudien wurde die Anwendbarkeit der entwickelten Roadmap untersucht. Hierbei wurde insbesondere auf die Anwendbarkeit der Reifegraduntersuchung, die Umsetzbarkeit der Handlungsmaßnahmen und das Verständnis der einzelnen Begrifflichkeiten und Erläuterungen analysiert. Hierbei hat sich insbesondere bestätigt, dass eine zu hohe Anzahl an Fragen zur Ermittlung des im Unternehmen vorliegenden Reifegrads zu einer Ablehnung bei produzierenden KMU führt. Bedingt ist dies durch die Tatsache, dass sowohl Zeit als auch Personal knappe Ressourcen sind. Darüber hinaus hat sich gezeigt, dass die Einführung von Machine Learning-Anwendungen stark von der Akzeptanz von Mitarbeitenden abhängt. Im Rahmen der Fallstudie bei der Reiser Maschinenbau AG hat sich gezeigt, dass erhöhte Transparenz, bspw. in Form von Erläuterungen des Mehrwerts und des Nutzens der Machine Learning-Anwendung, dazu führen, dass kritische Mitarbeiter sich der Einführung von ML öffnen und über weitere

Anwendungsmöglichkeiten nachdenken. Dieser Aspekt war auch in den anderen Fallstudien zu erkennen, weswegen sich das nächste Kapitel mit einer Meta-Analyse zur Algorithmenakzeptanz befasst. In dieser sollte untersucht werden, ob eine generelle Algorithmenakzeptanz oder -abneigung existiert.

2.4.2 Meta-Analyse zur Algorithmenakzeptanz

Literaturrecherche nach PRISMA

Mit einer angestrebten Veröffentlichung in einem hochrangigen Management Accounting-Journal wurden Paper zunächst in hochrangen Management Accounting-Journals gesucht. Eine strukturierte Literaturrecherche wurde in folgenden Journals:

- Accounting Review
- Journal of Accounting Research
- Management Science
- Journal of Financial Economics
- Journal of Accounting and Economics
- Academy of Management Journal (AMJ)
- Administrative Science Quarterly (ASQ)
- Academy of Management Review (AMR)
- Contemporary Accounting Research - Recherche Comptable Contemporaine
- Accounting, Organizations and Society
- Organization Science
- Review of Accounting Studies
- European Accounting Review
- Strategic Management Journal (SMJ)
- Management Accounting Research
- Journal of Management Studies (JMS)
- Organization Studies
- Journal of Management (JOM)
- Journal of Financial and Quantitative Analysis (JFQA)
- Management Information Systems Quarterly (MISQ)
- Journal of Management Accounting Research
- Auditing: A Journal of Practice & Theory
- Journal of Economic Behavior and Organization

mit dem Searchstring

(“Decision Support System” OR algorithm) AND (acceptance OR weight OR aversion OR reliance) AND experiment

je nach Möglichkeiten der Searchengines in:

- Abstract, Titel oder Keywords, falls möglich, andernfalls
- Abstract, falls möglich, andernfalls
- Überall

durchgeführt.

Zwecks spezieller Anforderungen wurde außerdem

- In Organization Studies, Management Science und dem Journal of Management Studies ausschließlich in keywords ohne die Anfordeurng experiment gesucht
- in MISQ ohne die Anforderung experiment gesucht
- in JMAR und Auditing ohne experiment und gefundene Paper bereits ohne Aufnahme gemäß PRISMA vorgefiltert

Insgesamt wurden 82 Paper gefunden, von denen 59 beim Screening wegen Irrelevanz ausgeschlossen wurden. Bei intensiver Betrachtung wurden 6 weitere Paper ausgeschlossen, die nur vermeintlich relevant waren, 7 weitere, da sie keine explizite Betrachtung von Algorithmenaversion aufgeführt haben und zwei weitere, die nicht von Algorithmenaversion, sondern der generellen Benutzung von IT-Systemen gesprochen haben. 8 relevante Paper sind aus diesem Schritt entstanden.

Erweiterung der Paper mit modernen Tools

Anschließend wurde die Datenbank der 8 relevanten Papern mit Hilfe von ResearchRabbit erweitert. ResearchRabbit basiert auf einem modernen Konzept von „verbundenen Papern“, die unter anderem durch LLM-basierte Methoden Konzepte in Papern erkennen können oder durch Zitationen verbundene Paper identifizieren. Dabei wurden alle Paper direkt gescreent und ausgeschlossen, wenn sie nicht relevant waren. Alle Untersuchungen vor dem Jahr 2000 wurden ausgeschlossen, da alle Paper vor diesem Datum signifikante Unterschiede zu aktuellen Technologien in Bezug auf den Einsatz von KI und Algorithmen im Management Accounting aufweisen.

Zu den 8 Papern wurden 11 weitere relevante Paper gefunden. Ab diesem Zeitpunkt wurden alle 19 Arbeiten ausgewählt und ähnliche Arbeiten gesichtet. Dabei wurden alle ähnlichen Arbeiten nach 2000 berücksichtigt, die zur Forschungsfrage gepasst haben, da anschließend ein weiteres gründliches Screening durchgeführt wurde. Nach wiederholter Erweiterung wurden insgesamt 74 Paper gefunden, die relevant schienen. Abschließend wurden gefundene Literaturreviews nach evtl. übersehenen Papern durchsucht. Alle relevanten Paper, die in Literaturreviews diskutiert wurden, wurden bereits gefunden. Nach weiterem intensiven Screening wurden insgesamt 29 Studien zur Extraktion herangezogen.

Extraktion der Daten

Aus den 29 Papern wurden anschließend die Experimentaldaten aller Einzelstudien entzogen, die relevant waren. Berücksichtigt wurden ausschließlich Studien, die eine explizite Verwendung

von Algorithmen untersucht haben (Benutzung von Algorithmen im Sinne von Übergabe einer Aufgabe an einen Algorithmus oder Inkorporation der Ergebnisse von Algorithmen im Sinne der Miteinbeziehung z. B. eines Forecasts in den eigenen Forecast), nicht beachtet wurden intentionale Faktoren wie Vertrauen in Algorithmen. Außerdem wurden unterschiedliche Konditionen in Experimenten teilweise aufgeteilt, da keine aggregierten Zahlen genannt wurden oder nur hätten geschätzt werden können. Insgesamt 108 Datenzeilen sind entstanden, die Daten von insgesamt 75.515 Teilnehmern an Experimenten beinhalten.

Nicht berichtete Werte wurden nach Möglichkeit bei den Autoren angefragt, aus den bereitgestellten Daten auf Portalen wie dem Open Science Framework selbst analysiert oder nach gängigen Methoden der Meta-Analyse geschätzt. Z. B. sind die Studien der Algorithmenaversion grundsätzlich in zwei Arten geteilt:

- Entscheidung zwischen einem Menschen und einem Algorithmus; hier gibt der Anteil der Teilnehmer mit Entscheidung für den Algorithmus den Grad der Akzeptanz von Algorithmen wieder
- Gewichtung der Ergebnisse von Algorithmen im Vergleich zu anderen Forecasts; hier gibt der weight on advice den Grad der Akzeptanz wieder, wobei in der Regel der weight on algorithm advice mit weight on expert advice verglichen wird

Zur Umwandlung wurden prozentuale Entscheidungen wie folgt zu einer within-subject Zeile umgewandelt

n_human = n

m_human = 1 – percentage

sd_human = sqrt(m_human (1 – m_human))

n_algorithm = n

m_algorithm = percentage

sd_algorithm = sqrt(n_ algorithm (1 – algorithm))

Bei Studien, die nur ein gesamtes n berichten statt Gruppen-Teilnehmerzahlen, wurden Teilnehmer in ganze Zahlen nahe der Hälfte geteilt, um die beste Annäherung an die wahren Zahlen zu erhalten. Ähnlich wurden andere fehlende Werte geschätzt und angepasste Teilnehmerzahlen für Studien mit within-subject designs berechnet.

Ergebnisse

Abbildung 30 zeigt ein Forest-Plot der Effekte. Die Mittelpunkte stehen für die dokumentierten Effektstärken, die Balken geben ein 95 %-Konfidenzintervall an. Rechts sind die Gewichte der einzelnen Studien basierend auf geschätztem τ^2 und Studienvarianzen sowie die Werte und Konfidenzintervalle in Zahlen. Tatsächlich scheint es in der Literatur keine allgemein dokumentierte Algorithmenaversion zu geben. Vielmehr zeigt die Analyse, dass eine leichte Tendenz hin zu Algorithmen existiert, obwohl kognitive Verzerrungen und generelles Überschätzen der eigenen Fähigkeiten diskutiert wird. Die Ergebnisse wurden mit Vertretern des projektbegleitenden Ausschusses diskutiert, die bereits im Vorfeld Skepsis an einer generellen Algorithmenaversion geäußert haben. Durch die breit gefächerten Anwendungsbereiche der Untersuchungen von medizinischen über militärischen bis hin zu solchen im Rechnungswesen wurde eine hohe Übertragbarkeit sichergestellt. KMU, die Algorithmen zur Steigerung der Ressourceneffizienz in der Produktion einsetzen wollen, wird geraten, den generellen Empfehlungen der MLready-Einführungsstrategie zu folgen, um potenziellen, speziellen Aversionen, die nicht generalistisch betrachtet werden können, zu begegnen. Z. B. sollten Mitarbeiter rechtzeitig informiert und in die Ausgestaltung mit eingebunden werden, um Abneigungen aus Angst vor Verlust der Arbeitsstelle durch Automatisierung vorzubeugen.

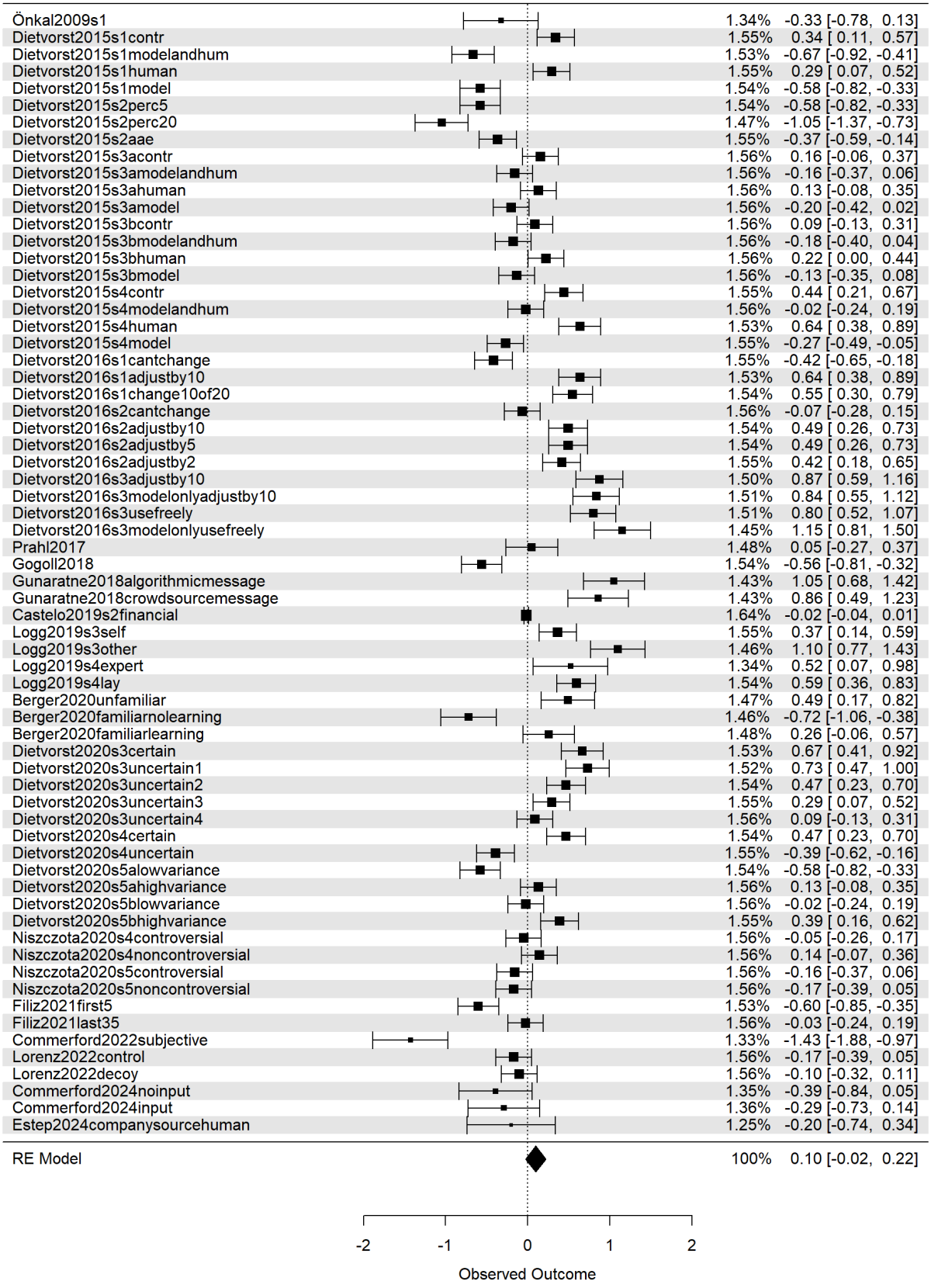


Abbildung 30: Meta-Analyse Algorithmenaversion

2.5 Arbeitspaket 5: Dokumentation, Transfer und Projektmanagement

AP 5:Dokumentation, Transfer und Projektmanagement	
Geplante Arbeiten	**Durchgeführte Arbeiten**
Über die gesamte Projektlaufzeit werden zahlreiche Chancen ergriffen, interessierte Unternehmen über die (Teil-)Ergebnisse des Forschungsprojekts zu informieren. Die Kommunikation der Ergebnisse erfolgt dabei über zahlreiche Kanäle wie das Internet, Printmedien und Vorträge. So wird ein breiter Wissenszugang und -transfer sichergestellt sowie eine umfangreiche Verbreitung gewährleistet. Das Projektmanagement dient darüber hinaus der systematischen und strukturierten Organisation des Forschungsprozesses und stellt somit die Ergebniserarbeitung gemäß Zeit- und Aufwandsplanung sicher.	Über die gesamte Projektlaufzeit wurden zahlreiche Chancen ergriffen, interessierte Unternehmen über die (Teil-)Ergebnisse des Forschungsprojekts zu informieren. Die Kommunikation der Ergebnisse erfolgte dabei über zahlreiche Kanäle wie das Internet, Printmedien und Vorträge. So wurde ein breiter Wissenszugang und -transfer sichergestellt sowie eine umfangreiche Verbreitung gewährleistet. Das Projektmanagement diente der systematischen und strukturierten Organisation des Forschungsprozesses und stellte somit die Ergebniserarbeitung gemäß Zeit- und Aufwandsplanung sicher.

Eine detaillierte Auflistung der Transfermaßnahmen ist in Tabelle 7 und Tabelle 8 dargestellt.

3. Verwendung der Zuwendung

Nachfolgend sind die Angaben zu den aus der Zuwendung finanzierten Ausgaben für Personenmonate des wissenschaftlich-technischen Personals gemäß Beleg über Beschäftigungszeiten (Einzelansatz A.1 des Finanzierungsplans), für Geräte (Einzelansatz B des Finanzierungsplans) und für Leistungen Dritter (Einzelansatz C des Finanzierungsplans) aufgeführt:

- wissenschaftlich-technisches Personal (Einzelansatz A.1 des Finanzierungsplans)

Für die Durchführung der Arbeiten wurde Personal nach A1 (Wissenschaftliche Mitarbeiter) benötigt und eingesetzt:

Tabelle 6: Personaleinsatz der Forschungseinrichtungen

Haushaltsjahr	IPRI	IPH	Gesamt
2022	5,39	2	**7,39**
2023	13,68	12	**25,68**
2024	3,25	10	13,25
Gesamt	22,32	24	**46,32**

- Geräte (Einzelansatz B des Finanzierungsplans)
  - entfällt
- Leistungen Dritter (Einzelansatz C des Finanzierungsplans)
  - entfällt

4. Notwendigkeit und Angemessenheit der geleisteten Arbeit

Die im Forschungsprojekt ML*ready* geleistete Arbeit entspricht in vollem Umfang dem bewilligten Antrag und war daher für die Durchführung des Vorhabens notwendig und angemessen.

Die intensiven Diskussionen im Rahmen der Treffen des projektbegleitenden Ausschusses sowie die intensive Arbeit in den Fallstudien und die Durchführung zahlreicher Expertengespräche haben die im Projektantrag dargestellte Problemstellung, des Fehlens praktikablen Wissens zur Einführung von Machine Learning in KMU für die Steigerung der Ressourceneffizienz bestätigt.

Insbesondere im Bereich der kritischen Phase – der kostenintensiven und risikoreichen Implementierungsphase – wurde bestätigt, dass es den KMU an der notwendigen Expertise für die erfolgreiche Integration von Machine Learning Technologien zur Steigerung der Ressourceneffizienz in der Produktion und den damit verbundenen internen und externen Herausforderungen fehlt. Vor allem bestehende Produktionsprozesse, die derzeit noch nicht in der Lage sind, die durch Machine Learning ermöglichten Effizienzsteigerungen erfolgreich zu realisieren, zeigten die Bandbreite der Handlungsfelder und den Bedarf an intensiver, anwendungsorientierter Forschungsarbeit in diesem Bereich.

Vor diesem Hintergrund erscheint die Förderung des Projektes ML*ready* als angemessen. Die entwickelten Hilfsmittel, wie die erarbeiteten Anwendungsgebiete sowie der Leitfaden zur Identifikation und Aufbereitung von Datengrundlagen bieten den KMU eine große Unterstützung bei der Einführung von Machine Learning. Der praxisorientierte Einführungsleitfaden ergänzt diese Arbeit, um die Voraussetzungen für eine zielgerichtete, erfolgreiche Implementierung von Machine Learning Technologien und die Möglichkeit zur Optimierung und Erweiterung des Wertschöpfungspotenzials von KMU langfristig erfolgreich zu gestalten. Die Erfahrungen im Rahmen des Projektes haben verstärkt aufgezeigt, dass die angestrebte Integration von Machine Learning Technologien besonders für KMU eine große Herausforderung darstellt und nur in Verbindung mit der Identifikation der geeigneten Anwendungsfelder, der Herstellung einer geeigneten Datengrundlage und einem konkreten Plan Erfolg verspricht. Die im Laufe des Projektes gesammelten Erfahrungen haben verstärkt aufgezeigt, dass die angestrebte Effizienzsteigerung besonders für KMU eine große Herausforderung darstellt und die Nutzung strukturierter Hilfsmittel Erfolg verspricht.

Das Erarbeiten der Ergebnisse war für das IPRI und das IPH mit einem hohen personellen Aufwand verbunden. Vor diesem Hintergrund beurteilen die Forschungseinrichtungen die geleistete Arbeit als inhaltlich angemessen und förderungswürdig. Auch die Höhe der Zuwendung erscheint in Anbetracht der erzielten Ergebnisse und des geleisteten Personalaufwandes angemessen.

5. Nutzen, Innovationsbeitrag und Anwendungsmöglichkeiten

Das Ziel des Forschungsvorhabens MLready war es, KMU zur Nutzung von Potenzialen von Machine Learning in der Produktion durch die Entwicklung einer Einführungsstrategie zu befähigen. Hierzu wurden zunächst Anwendungsfälle identifiziert und übersichtlich zusammengestellt. Anschließend wurden bestehende Datengrundlagen identifiziert sowie solche bestimmt, die zur Umsetzung der Anwendungsfälle notwendig oder hilfreich sein können und den Beschreibungen ergänzt. Zusätzlich wurde ein Leitfaden entwickelt, wie KMU Daten für einen Anwendungsfall in ihrem Unternehmen identifizieren und aufbereiten können, um die Grundlage der Anwendung sicherzustellen. Diese Ergebnisse wurden in einer anwenderfreundlichen App zusammengestellt, die den Nutzer durch den gesamten Prozess von Bestimmung der Grundvoraussetzungen über Anwendungsfallidentifikation, Datengewinnung bis hin zu einer Roadmap leitet, die die Umsetzung in KMU bedeutend vereinfachen soll.

Im nächsten Schritt wurde ein Reifegradmodell entwickelt, anhand dessen KMU ihre aktuell vorliegenden “MLreadiness“ prüfen können. Hierzu wurden Fragen zu den Mitarbeitern, dem vorliegenden KnowHow, dem Prozess, der notwendigen Grundvoraussetzungen und den durchzuführenden Handlungsmaßnahmen implementiert. Anschließend wurden die Ergebnisse in eine benutzerfreundliche und intuitive Applikation eingebettet. Diese ermöglicht auf Grund der leicht nachvollziehbaren Logik und hilfreichen Informationen eine aufwandsarme und unkomplizierte Nutzung der Einführungsstrategie. Darüber hinaus wird durch das breite Spektrum an Anwendungsmöglichkeiten eine große Menge an Anwendern sichergestellt.

Durch die Nutzung der Einführungsstrategie erhalten die Anwender sowohl einen Überblick über den aktuell im Unternehmen vorliegenden Zustand als auch über notwendige Schritte, die es zur erfolgreichen Einführung von Machine Learning bedarf. Insbesondere relevante Werkzeuge, wie beispielsweise einen Projektplan zur Rollenverteilung und die Anwendung einer Anforderungsliste für Machine Learning-Dienstleister können KMU unterstützen und dazu beitragen, dass auch diese im Zeitalter der Digitalisierung von den Potenzialen und Vorteilen der künstlichen Intelligenz profizieren.

6. Plan zum Ergebnistransfer und Einschätzung zur Realisierbarkeit des Transferkonzepts

6.1 Einschätzung zur Realisierbarkeit des Transferkonzepts

Um eine umfassende Verbreitung der Ergebnisse des Forschungsvorhabens in der Wirtschaft zu erreichen, wurde eine Reihe von Transfermaßnahmen durchgeführt. Wesentliche Maßnahmen wurden bereits während der Projektlaufzeit umgesetzt, darüber hinaus ist eine Reihe von Maßnahmen nach der Projektlaufzeit geplant und bereits angestoßen. Eine detaillierte Auflistung findet sich in den folgenden Abschnitten. Um die Anwendung der Erkenntnisse über die Projektlaufzeit hinaus sicherzustellen, werden insbesondere die Endergebnisse (Einführungsleitfaden, Leitfaden zur Aufbereitung von Daten, Kompendium von Anwendungsfällen) verfügbar bleiben. Dadurch wird gewährleistet, dass möglichst viele interessierte Parteien (insbes. KMU, Forschungseinrichtungen und Verbände) von den Forschungsergebnissen profitieren können. Die Realisierbarkeit des Ergebnistransfers wird daher als sehr hoch eingeschätzt.

6.2 Spezifische Transfermaßnahmen während der Projektlaufzeit

Tabelle 7: Transfermaßnahmen während der Projektlaufzeit

Maßnahme	Ziel	Ort / Rahmen	Durchgeführte Maßnahmen
Organisation eines „ML*ready-Tages*“	Vorstellung der Projektergebnisse in Vorträgen, Demonstration der Fallstudien sowie Verbreitung der Ergebnisse bei interessierten KMU und Intermediären (IHKs, Verbände, Unis).	IPRI und IPH organisieren ein „ML*ready-Tag*“, an dem der Softwaredemonstrator mit interessierten KMU getestet wird.	Sowohl die durchgeführten Fallstudien als auch die erstellte Einführungsstrategie wurden beim MLready-Tag bei einem Unternehmen bei Kassel von Teilnehmern des projektbegleitenden Ausschusses getestet und validiert.

Maßnahme	Ziel	Ort / Rahmen	Durchgeführte Maßnahmen
Wissenschaftliche und praxisorientierte Veranstaltungen	Sicherstellung der Umsetzbarkeit der Ergebnisse durch Praxisdiskussionen „Tandem-Vorträge“ (Forschung/Unternehmen) Verbreitung von (Teil-) Ergebnissen	7./8. AK 4.0 Symposium in Ulm (Ausrichter: IPRI, Uni Ulm, IHK Ulm) Serviceforum Region Stuttgart 2021 und 2022 (Ausrichter: IPRI, WRS)	Die geplanten Veranstaltungen fanden im Projektzeitraum nicht mehr statt. Stattdessen wurden die Inhalte von MLready fortlaufend auf Besuchen verschiedenster Konferenzen diskutiert: • Event des Network Data Economy am 29.03.2023 • Digitalgipfel 2023 – Wirtschaft 4.0 BW am 22.06.2023 • Workshoptage von nedyx am 17.07.2023
Integration in Industrie-Arbeitskreise	Vorstellung des Projekts und erster Forschungsergebnisse	Schmalenbach Arbeitskreis „Integrationsmanagement für neue Produkte“ Sitzungen IPRI-Arbeitskreis „Industrie 4.0“	Die Arbeitskreise wurden eingestellt. An ihrer Stelle werden aktuell neue Arbeitskreise aufgebaut. Ergebnisse aus MLready wurden teilweise bereits zum Aufbau der Inhalte verwendet.
Integration in Digitalisierungsinitiativen	Verbreitung und Nutzung der Projektergebnisse im Rahmen des Mittelstand Digital Zentrums Hannover	Anwendung der Einführungsstrategie sowie der Roadmap bei Firmen-gesprächen und Digitalisierungsprojekten im Rahmen des KI-Trainers	Die bisher entwickelte Einführungsstrategie wurde im Rahmen des Big Data Workshops des Mittelstand Digital Zentrums Hannover vorgestellt, angewendet und mit den Teilnehmern diskutiert.
Vorstellung auf wissenschaftlichen Konferenzen	Verbreitung und Diskussion der Forschungsergebnisse	Business Research Konferenzen des VHB	Die Ausarbeitung des Papers zur Algorithmenakzeptanz ist im Gange und wird mit dem ersten Diskussionsstand auf wissenschaftliche Konferenzen eingereicht.

Präsenz im Internet	Fortlaufende Information über das Forschungsprojekt und die (Teil-)Ergebnisse	Eigene Webpräsenz für das Forschungsprojekt bzw. Nennung Forschungs-News: breitenwirksame Verteilung auf Plattformen	Website existiert unter https://ipri-institute.com/forschungsprojekte/mlready/ Website existiert unter https://www.iph-hannover.de/de/forschung/forschungsprojekte/?we_objectID=6247 Neues-aus-der-Forschung: https://neues-aus-der-forschung.de/2023/07/10/forschungsprojekt-mlready-steigerung-der-ressourceneffizienz-durch-machine-learning/ **LinkedIn** https://www.linkedin.com/posts/manuel-savadogo-5526001b3_join-conversation-activity-7019683414542802944-kKCX https://www.linkedin.com/posts/iph-ggmbh_mlready-ipri-ggmbh-activity-6993888329226186752-Cxqt https://www.linkedin.com/posts/manuel-savadogo-5526001b3_digitalisierung-industrie-produktion-activity-7029359446598307841-xHyr https://www.linkedin.com/posts/ipri-institute_mlready-machinelearning-bilder-kennung-activity-7119976015472275456-X90J https://www.linkedin.com/posts/ipri-institute_3-projekttreffen-des-projekts-mlready-ausschuss-activity-7136014284404252672-TaSX https://www.linkedin.com/posts/ipri-institute_mlready-machinelearning-ressourceneffizienz-activity-7126168656664780800-1g5J

Maßnahme	Ziel	Ort / Rahmen	Durchgeführte Maßnahmen
			https://www.linkedin.com/posts/ipri-institute_esg-machinelearning-save-activity-7047105552270688256-ef3c
Presse-/ Öffentlichkeitsarbeit	Bekanntmachung des Projekts und weitere Verbreitung der Projektinhalte und -ergebnisse	Informationsdienst der Wissenschaft, Presseverteiler IPH Institutszeitschriften (IPRI-Journal; IPH-Produktionstechnik Hannover informiert)	Pressemitteilung zum Projektstart Pressemitteilung zu Zwischenergebnissen Beitrag im IPRI Jahresbericht 2022 Beitrag im IPRI Journal 2023 Beitrag im IPRI Jahresbericht 2023
Veröffentlichung von Ergebnissen in wissenschaftlichen Medien	Bekanntmachung und Diskussion der Ergebnisse in der Wissenschaft	z. B. Logistics Journal Behavioral Research in Management Accounting	Die Ausarbeitung des Papers zur Algorithmenakzeptanz ist im Gange und wird nach Diskussion auf Konferenzen in einem wissenschaftlichen Journal publiziert.
Veröffentlichung der Projektergebnisse mit Fokus auf die Praxis	Bekanntmachung der Ergebnisse in der Praxis, Aufzeigen von Anwendungsfällen	VDMA Nachrichten; Wissenschaft trifft Praxis; Logistik heute; Controller Magazin Zeitschrift für wirtschaftlichen Fabrikbetrieb	Artikel in der VDI-Z Artikel in der ZWF Artikel in Industry 4.0 Science (im Peer-Review Prozess)

Maßnahme	Ziel	Ort / Rahmen	Durchgeführte Maßnahmen
Treffen des Projektbegleitenden Ausschusses	Validierung der Ergebnisse mit Praxispartnern; Übertragung der Ergebnisse auf praxisrelevante Probleme; Ziel: 4 PA-Treffen	IPH/IPRI Mitglieder des projektbegleitenden Ausschusses	Projekttreffen in Stuttgart (hybrid) am 17.01.2023 Projekttreffen in Hannover (digital) am 20.06.2023 Projekttreffen in Stuttgart (hybrid) am 29.11.2023 Projekttreffen in Marsberg (vor Ort) am 19.06.2024
Integration in die universitäre Lehre	Angebot von studentischen Arbeiten sowie Dissertationsthemen, Integration von ML*ready* in die universitäre Lehre zur nachhaltigen Verbreitung und Weiterentwicklung der Projektergebnisse	Universität Hannover: Integration in die Lehrveranstaltung „Anlagenmanagement“ Universität Ulm: „Business Analytics“ (Modulgruppe: Wirtschaftswissenschaften), Universität Ulm	Vorstellung der Ergebnisse und Workshop in der SAPS am 11.01.2023 sowie am 11.01.2024

6.3 Spezifische Transfermaßnahmen nach Abschluss des Vorhabens

Tabelle 8: Transfermaßnahmen nach der Projektlaufzeit

Maßnahme	Ziel	Ort/Rahmen	Durchgeführte Maßnahmen
Verbreitung des ML*ready*-Leitfadens	Der Leitfaden wird branchenübergreifend an produzierende KMU verbreitet, um nachhaltig auf die Projektergebnisse aufmerksam zu machen	Verbreitung über die Online-Präsenz der Institute sowie auf Präsenz-Veranstaltungen, an denen die Institute teilnehmen (Messen, Workshops etc.)	Der MLready-Leitfaden wurde auf der Hannover Messe, in Workshop „KI in der Produktion“ des Mittelstand Digital-Zentrums Hannover und auf dem Digital Days 2024 in Leverkusen vorgestellt und diskutiert.

Maßnahme	Ziel	Ort/Rahmen	Durchgeführte Maßnahmen
Langfristige Integration in den Arbeitskreis „Industrie 4.0“	Verbreitung der Ergebnisse und deren Überführung in die praktische Anwendung	Arbeitskreis des IPRI (tagt 3x jährlich); Stuttgart, IPRI	Der Arbeitskreis Industrie 4.0 wurde eingestellt. An seiner Stelle werden aktuell neue Arbeitskreise aufgebaut. Ergebnisse aus MLready wurden teilweise bereits zum Aufbau der Inhalte verwendet.
Webinar	Angebot eines Webinars für KMU	IPH und IPRI	Auf der Projekthomepage unter https://ipri-institute.com wird ein Webinar zu den Forschungsergebnissen angeboten.
Angebot von unternehmensspezifischen Transferprojekten	Unterstützung von KMU bei individuellen Problemstellungen durch Beratungsmandate	IPH, IPRI; vor Ort bei den jeweiligen Unternehmen	An Fallstudien interessierte Unternehmen, die nicht im projektbegleitenden Ausschuss tätig waren, wurden Angebote zur Umsetzung von Fallstudien gestellt.

7. Forschungsstellen

7.1 IPRI – International Performance Research Institute gGmbH

Das IPRI ist ein gemeinnütziges Forschungsinstitut und wurde mit der Zielsetzung gegründet, Forschung auf dem Gebiet des Performance Management von Organisationen, Unternehmen und Unternehmensnetzwerken zu betreiben. Unter Leitung von Prof. Dr. Mischa Seiter untersucht das IPRI in Zusammenarbeit mit anderen Forschungseinrichtungen und kleinen und mittelständischen Unternehmen die Wirkungszusammenhänge und Potenziale in den Bereichen Digital Business Models, Sustainability Management, Emerging Technologies und Interorganisational Management.

Tabelle 9: IPRI – International Performance Research Institute gGmbH

Forschungsstelle	International Performance Research Institute gGmbH
Anschrift	Reuchlinstraße 27, 70176 Stuttgart
Leitung der Forschungsstelle	Prof. Dr. Mischa Seiter
Kontakt	info@ipri-institute.com, https://ipri-institute.com

7.2 IPH – Institut für Integrierte Produktion Hannover gGmbH

Das Institut für Integrierte Produktion Hannover (IPH) gGmbH ist eine gemeinnützige GmbH, die sich auf Forschung und Entwicklung im Bereich der Produktionstechnik spezialisiert hat. Sie bietet Beratungsdienste für Industrieunternehmen und bildet den ingenieurwissenschaftlichen Nachwuchs aus. Gegründet wurde das IPH 1988 als Ausgründung der Leibniz Universität Hannover. Das IPH untersucht gemeinsam mit kleinen und mittelständischen Unternehmen die Wirkzusammenhänge und die Potenziale im dynamischen und komplexen Produktionsumfeld.

Tabelle 10: IPH - Institut für integrierte Produktion Hannover gGmbH

Forschungsstelle	Institut für integrierte Produktion gGmbH
Anschrift	Hollerithallee 6, 30419 Hannover
Leitung der Forschungsstelle	Prof. Dr. Peter Nyhuis
Kontakt	info@iph-hannover.de, https://iph-hannover.de

Förderhinweis

Das Projekt „MLready“ [01|F22312N] wurde im Rahmen des Programms „Industrielle Gemeinschaftsforschung (IGF)“ durch das Bundesministerium für Wirtschaft und Klimaschutz aufgrund eines Beschlusses des Deutschen Bundestages gefördert.

Gefördert durch:

Bundesministerium für Wirtschaft und Klimaschutz

aufgrund eines Beschlusses des Deutschen Bundestages

betreut von:

DLR

iuta

Literaturverzeichnis

Al-Anqoudi, Younis; Al-Hamdani, Abdullah; Al-Badawi, Mohamed; Hedjam, Rachid (2021): Using Machine Learning in Business Process Re-Engineering. In: *BDCC* 5 (4), S. 61. DOI: 10.3390/bdcc5040061.

Arnaz, Ricky; Harahap, Muslim Efendi (2020): Analysis of Implementation of Robotic Process Automation: A Case Study in PT X. In: *Proceedings of the Business Innovation and Engineering Conference 2020 (BIEC 2020)*, S. 61–64. DOI: 10.2991/aebmr.k.210727.011.

Bitkom (2017): Künstliche Intelligenz. Wirtschaftliche Bedeutung, gesellschaftliche Herausforderungen, menschliche Verantwortung.

Bitkom (2020): 86.000 offene Stellen für IT-Fachkräfte. Online verfügbar unter https://www.bitkom.org/Presse/Presseinformation/86000-offene-Stellen-fuer-IT-Fachkraefte, zuletzt geprüft am 21.09.2021.

BMWi; wik (2019): Künstliche Intelligenz im Mittelstand. Relevanz, Anwendungen, Transfer.

Deloitte (2014): Data Analytics im Mittelstand. Die Evolution der Entscheidungsfindung. Online verfügbar unter https://www2.deloitte.com/content/dam/Deloitte/de/Documents/Mittelstand/studie-data-analytics-im-mittelstand-deloitte-juni-2014.pdf, zuletzt geprüft am 23.09.2020.

Deloitte (2019): State of AI in the Enterprise. Ergebnisse der Befragung von 100 AI-Experten in deutschen Unternehmen.

English, Larry P. (2002): Total Quality data Management (TQdM). In: Mario G. Piattini, Coral Calero und Marcela F. Genero (Hg.): Information and Database Quality, Bd. 25. 1st ed. 2002. New York, NY: Imprint Springer; Springer US (Advances in Database Systems, 25), S. 85–109.

Experian Qas (2013): The State of data quality. Online verfügbar unter https://www.experian.com/assets/decision-analytics/white-papers/the%20state%20of%20data%20quality.pdf, zuletzt geprüft am 30.06.2020.

Fraunhofer (2018): Maschinelles Lernen. Eine Analyse zu Kompetenzen, Forschung und Anwendung. Online verfügbar unter https://www.bigdata.fraunhofer.de/content/dam/bigdata/de/documents/Publikationen/Fraunhofer_Studie_ML_201809.pdf, zuletzt geprüft am 14.04.2020.

Hofmann, Martin (2020): Prozessoptimierung als ganzheitlicher Ansatz. Mit konkreten Praxisbeispielen für effiziente Arbeitsabläufe. Wiesbaden: Springer Gabler (Fachmedien).

IDG (2019): Studie Machine Learning / Deep Learning.

Industry Analytics (2018): Hürden bei der Nutzung von Big Data in KMU - Industry Analytics. Online verfügbar unter http://www.industry-analytics.de/huerden-bei-der-nutzung-von-big-data-in-kmu/, zuletzt aktualisiert am 20.12.2018, zuletzt geprüft am 23.09.2020.

Ishwarappa; Anuradha, J. (2015): A Brief Introduction on Big Data 5Vs Characteristics and Hadoop Technology. In: *Procedia Computer Science* 48, S. 319–324. DOI: 10.1016/j.procs.2015.04.188.

KI-Transfer BW (2022): Wie gelingt KI im Mittelstand? Erfahrungen und Best Practices aus dem Projekt KI-Transfer BW.

Kocsi, Balázs; Maiko, Michael; László, Matonya; Pusztai, Péter; Budai, István (2020): Real-Time Decision-Support System for High-Mix Low-Volume Production Scheduling in Industry 4.0. Online verfügbar unter https://www.mdpi.com/2227-9717/8/8/912, zuletzt geprüft am 26.01.2023.

Lee, Yang W.; Strong, Diane M.; Kahn, Beverly K.; Wang, Richard Y. (2002): AIMQ: a methodology for information quality assessment. In: *Information & Management* 40 (2), S. 133–146. DOI: 10.1016/S0378-7206(02)00043-5.

Lufthansa Industry Solutions (2020): IDG-Studie 2020 „Machine Learning“.

Naumann, Felix (2007): Datenqualität. In: *Informatik Spektrum* 30 (1), S. 27–31. DOI: 10.1007/s00287-006-0125-5.

Page, Matthew J.; McKenzie, Joanne E.; Bossuyt, Patrick M.; Boutron, Isabelle; Hoffmann, Tammy C.; Mulrow, Cynthia D. et al. (2021): The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. In: *BMJ (Clinical research ed.)* 372, n71. DOI: 10.1136/bmj.n71.

PAiCE (2018): Potenziale der künstlichen Intelligenz im produzierenden Gewerbe in Deutschland.

Taleb, Ikbal; Dssouli, Rachida; Serhani, Mohamed Adel (2015): Big Data Pre-processing: A Quality Framework. In: 2015 IEEE International Congress on Big Data. 2015 IEEE International Congress on Big Data (BigData Congress). New York City, NY, USA, 27.06.2015 - 02.07.2015: IEEE, S. 191–198.

Tayi, Giri Kumar; Ballou, Donald P. (1998): Examining data quality. In: *Commun. ACM* 41 (2), S. 54–57. DOI: 10.1145/269012.269021.

Turner, David; Schroeck, Michael; Shockley, Rebecca (2013): Analytics: The real-world use of big data in financial services. Online verfügbar unter https://www.ibm.com/downloads/cas/E4BWZ1PY, zuletzt geprüft am 03.07.2020.

VDMA (2018): Machine Learning im Maschinen- und Anlagenbau.

Waltersmann, Lara; Kiemel, Steffen; Stuhlsatz, Julian; Sauer, Alexander; Miehe, Robert (2021): Artificial Intelligence Applications for Increasing Resource Efficiency in Manufacturing Companies—A Comprehensive Review. In: *Sustainability* 13 (12), S. 6689. DOI: 10.3390/su13126689.

Wand, Yair; Wang, Richard Y. (1996): Anchoring data quality dimensions in ontological foundations. In: *Communications of the ACM* 39 (11), S. 86–95. Online verfügbar unter https://dl.acm.org/doi/10.1145/240455.240479.

Winter, Cornelia (2021): Mitteilungen der GI im Informatik Spektrum 5/2021. In: *Informatik Spektrum*. DOI: 10.1007/s00287-021-01391-7.