



Digitale Transformation im Retoureneingang: Klassifikationsmodell zur Vorsortierung von Retourenpaketen

Anthony Boyd Stevenson · Julia Rieck

Eingegangen: 15. April 2023 / Angenommen: 1. August 2023 / Online publiziert: 17. August 2023
© The Author(s) 2023

Zusammenfassung Die vorliegende Arbeit präsentiert die Entwicklung eines Klassifikationsmodells und eines Software-Systems zur Vorsortierung von Retourenpaketen bei einem B2C-E-Commerce Versandhändler, dessen Rücksendungen durch die Verwendung von Retourenscheinen bearbeitet werden. Im Rahmen der Klassifikation wird mithilfe vorhandener Daten und unter Verwendung eines dreistufigen Berechnungsschemas die Artikelanzahl in einem Retourenpaket vorhergesagt. Hierdurch wird eine präzisere Klassifikation gegenüber dem ersten Prototyp aus einer früheren Arbeit für die Prozessoptimierung ermöglicht. Eine Evaluierung des Systems zeigt eine signifikante Verbesserung gegenüber dem Prototyp. Das System wurde kostengünstig in die bestehende IT-Infrastruktur integriert und trägt zur Steigerung der Digitalisierung des Retourenlagers bei. Die vorgestellte Case Study belegt den Nutzen des Software-Systems und zeigt, dass Unternehmen, insbesondere KMUs, durch den Einsatz von Logistik-4.0-Technologien ihre Effizienz steigern und wettbewerbsfähig bleiben können. Zudem wird diskutiert, ob durch das Software-System ein Paradigmenwechsel zu Logistik 5.0 im Retoureneingang erreicht wird.

Schlüsselwörter Digitale Transformation · E-Commerce · Retourenmanagement · Case Study · Prozessoptimierung

✉ Anthony Boyd Stevenson · Julia Rieck
Institute for Business Administration and Information Systems, University of Hildesheim,
Universitätsplatz 1, 31141 Hildesheim, Deutschland
E-Mail: stevenson@uni-hildesheim.de

Julia Rieck
E-Mail: rieck@bwl.uni-hildesheim.de

Digital transformation in an incoming returns department: classification model for presorting returns packages

Abstract This paper presents the development of a classification model and a software system for pre-sorting return packages at a B2C e-commerce retailer whose returns are processed through the use of return slips. In the classification process, the number of items in a return package is predicted using existing data and a three-step calculation scheme. This enables a more precise classification compared to the first prototype from an earlier work for process optimization. An evaluation of the system shows a significant improvement over the prototype. The system was cost-effectively integrated into the existing IT infrastructure and helps to increase the digitization of the returns warehouse. The presented case study proves the benefit of the software system and shows that companies, especially SMEs, can increase their efficiency and remain competitive by using logistics 4.0 technologies. In addition, it is discussed whether the software system achieves a paradigm shift to Logistics 5.0 in incoming returns.

Keywords Digital transformation · E-Commerce · Returns management · Case study · Process optimization

1 Einleitung

Die Digitalisierung der Logistik- und Retourenprozesse (digitale Transformation) kann zu einer Steigerung der Effizienz und Wirtschaftlichkeit von Unternehmen führen. Dabei sind allerdings nicht nur reine Technologien einzusetzen, sondern auch die Durchdringung dieser im Unternehmen voranzutreiben (vgl. Günther et al. 2022, S. 1–2). Besonders wichtig ist hierbei die Anpassung der IT-Infrastruktur, die optimal für eine digitale Transformation vorbereitet sein sollte (vgl. Strahinger und Westner 2022, S. 718). Zudem spielen sozio-technische Aspekte eine große Rolle, z.B. die Schulung und Qualifizierung der Mitarbeitenden im Umgang mit neu entwickelten Technologien. Insbesondere der Mittelstand muss sich diesen Herausforderungen stellen, um im globalen Wettbewerb bestehen zu können (vgl. Günther et al. 2022, S. 1).

Die Unterstützung von bereits stark digitalisierten, logistischen Prozessen, wie z.B. dem Routing von Fahrzeugen, mit reaktiven Methoden oder Methoden der künstlichen Intelligenz (KI) ist in einigen Unternehmen bereits etabliert (vgl. Chen et al. 2022; Schlenkrich und Parragh 2023). Im Bereich nicht-digitalisierter Prozesse (z.B. im Retoureneingang) kann ein Anstieg an Arbeitsvolumen mithilfe von Maßnahmen der digitalen Transformation effektiver bewältigt werden. Um dies anhand einer praktischen Anwendung zu belegen, wurde eine Zusammenarbeit mit einem B2C-E-Commerce Versandhändler für Möbel und Wohnaccessoires initiiert. Im Rahmen der Kooperation wurde untersucht, inwieweit die nicht-digitalisierten Retourenprozesse durch einfache und kostengünstige Lösungen verbessert werden können. Hierfür wurde in Stevenson und Rieck (2022) ein Prototyp als erste Digitalisierungsmaßnahme des Retoureneingangs entwickelt. Durch die vielversprechenden

Ergebnisse des Prototyps ergab sich eine weitere Kooperation mit demselben Unternehmen für die vorliegende Arbeit. Dabei sei erwähnt, dass sowohl die erste als auch die vorliegende Arbeit ohne Vorgaben oder Einschränkungen des Kooperationspartners durchgeführt wurden.

Im Folgenden betrachten wir erneut die Prozesse des Retoureneingangs sowie der -bearbeitung des Kooperationspartners. Der erste Prototyp aus Stevenson und Rieck (2022) wird nun erweitert, um (1) eine bessere Klassifikation von Paketen mit „mehr als einem Artikel in den Bestelldaten“ zu erreichen und um (2) eine kostengünstige Überführung des Prototyps als Software-Lösung in die bestehende IT-Infrastruktur zu ermöglichen. In der neuen Klassifikation (1) wird beachtet, dass es bei einer Originalbestellung mit mehreren Artikeln passieren kann, dass nur ein Artikel zurückgesendet wird. Dies ist insbesondere dann der Fall, wenn der Kunde bzw. die Kundin einen Auswahlkauf getätigt hat. Kann ein Paket schon vor dem Öffnen als „Single Return“ (SR, Retoure mit einem Artikel) erkannt werden, ist eine effiziente Vorsortierung und Integration in den Bearbeitungsprozess möglich. Durch die Überführung des Systems in die bestehende IT-Infrastruktur (2) wird die Nutzbarkeit des Systems im Lager signifikant erhöht. Das angestrebte Ziel des in diesem Beitrag vorgestellten Software-Systems ist es, eine Verbesserung bei der Klassifikation von eintreffenden Retouren zu realisieren und die Einbettung des Systems im Unternehmen zu erläutern. Darüber hinaus soll abschließend exemplarisch anhand des Software-Systems die Frage „Logistik 5.0 oder Logistik 4.1: Wo stehen wir im Retoureneingang?“ diskutiert und beantwortet werden.

Das implementierte Software-System sammelt und verarbeitet fortlaufend aktualisierte Bestell- und Kundendaten. Eintreffende Retouren können anhand dieser Daten vorsortiert werden und insbesondere in „SR“ und „Multiple Return“ (MR, Retoure mit mehreren Artikeln) eingeteilt werden. Bei einem „SR“ muss nur ein Artikel dem Paket entnommen und hinsichtlich der Qualität geprüft werden. Bei einem „MR“ sind mehrere Artikel aus dem Paket zu entnehmen, es ist die Qualität zu prüfen, eine Zuordnung der Artikel zu Retourenscheinen durchzuführen und auch die korrekte Bildung von Artikeleinheiten zu kontrollieren. Die Bearbeitung von „MR“ dauert deutlich länger und sollte von erfahrenden Lagermitarbeitenden durchgeführt werden. Insgesamt kann durch ein Klassifikationsmodell zur Vorsortierung eine signifikante Zeitersparnis, eine höhere Arbeitsqualität und Zufriedenheit der Mitarbeitenden sowie eine Steigerung der Wettbewerbsfähigkeit des Unternehmens erreicht werden (vgl. Eierdanz und Blaudszun-Lahm 2020, S. 1).

2 Literaturübersicht

Die digitale Transformation führt in der Regel zu einer hohen Qualität der Logistik, der Lieferfähigkeit, der Kundenzufriedenheit und der Gesamtleistung des Unternehmens (vgl. Grosse 2023, S. 8; Winkelhaus und Grosse 2020, S. 26). Im E-Commerce hat die effiziente Abwicklung von Retouren mittlerweile eine zentrale Rolle eingenommen, da eine Vielzahl an Personen bestellte Ware retournieren und der Retourenprozess somit einen ähnlichen Einfluss auf die Kundenzufriedenheit besitzt, wie der Kauf selbst (vgl. Bold 2021). Im Zuge des Wachstums der einhergehenden

Rücksendungen stehen Versandhändler vor der Herausforderung, Vorhersagemodelle für Retouren und Retourenanzahlen zu entwickeln, die historische Kunden-, Bestell- und Rechnungsdaten nutzen (vgl. Liang et al. 2014). Die Resultate können verwendet werden, um den Bearbeitungsprozess der Retouren zu verbessern oder um geeignete Maßnahmen und Strategien zu entwickeln, um Retouren zu vermeiden (vgl. Trippner 2021). In der wissenschaftlichen Literatur wird im Bereich des Retourenmanagements häufig auch das Management von Beständen oder die Beschaffungs- und Produktionsplanung in den Fokus gerückt. Beispielsweise werden in Toktay et al. (2004) die Vorhersagen von Rücksendungen für operative Entscheidungen in der Bestandskontrolle verwendet. Auf diese Weise kann eine genauere Planung und Steuerung der Lagerbestände durchgeführt werden, um eine optimale Bestandsführung zu gewährleisten. In Krapp et al. (2013) wird bspw. für das Lagerbestandsmanagement ein generischer Ansatz für die Prognose von Produktretouren in geschlossenen Lieferketten bereitgestellt. In Clotey et al. (2012) werden die Vorhersagen als Grundlage für Beschaffungsentscheidungen genutzt und dann auf Basis der Beschaffungen die Produktionsplanung optimiert. Das in Zhou et al. (2016) beschriebene Modell zielt darauf ab, die Menge, Zeit und Wahrscheinlichkeit von Produktretouren zu prognostizieren und darauf aufbauend die wiederverwendbaren Teile, Komponenten und Materialien sowie die erforderlichen Entsorgungsaktivitäten im Zusammenhang mit den Retouren zu identifizieren. Zudem wird in der Arbeit von Wang et al. (2021) untersucht, inwieweit sich die Prozesse des Reverse Logistik und der Retourenbearbeitung outsourcen lassen. Um Retouren vorherzusagen, kommen mathematisch-statistische Methoden, (Meta-)heuristiken oder Methoden der KI zum Einsatz (vgl. Yang et al. 2022; Zhu et al. 2018). In Cui et al. (2020) werden Anzahlen von Produktretouren auf Einzelhändler-, Produkttyp- und Zeitebene vorhergesagt. Durch die Nutzung eines umfangreichen Datensatzes, der detaillierte Informationen für jeden Einzelhändler und jede Produktgruppe enthält, können robuste datengetriebene Modelle entwickelt und getestet werden. Der Umgang mit großen Datenmengen auf Produktebene ist eine Herausforderung und ebenfalls für die vorliegende Arbeit relevant. Vorhersagemodelle im E-Commerce werden nicht nur für die Retouren entwickelt, sondern auch z. B. für die Nachfrageprognose auf Artekebene. Dai et al. (2022) verwenden dabei einen zwei-stufigen Random-Forest-Ansatz, um Features für die Vorhersage auf Artekebene zu ermitteln. Die Betrachtung der Artekebene ist auch bei der Klassifikation in dieser Arbeit von Bedeutung.

Im Rahmen der Literaturrecherche ist aufgefallen, dass die Bearbeitung der Retourenpakete anhand von Retourenscheinen nicht explizit thematisiert wird. Laut Girschner (2022) gilt allerdings der im Paket beigelegte Retourenschein als State-of-the-Art-Lösung für viele mittelständische Onlinehändler. Um die negativen Auswirkungen von Retouren auf Geld, Zeit und Kundenbindung zu minimieren, könnte bspw. auch die Implementierung eines umfassenden Retourenportals vorgesehen werden (vgl. Ager 2021). Angesichts der begrenzten finanziellen Ressourcen mit denen kleine und mittlere Unternehmen (KMU) konfrontiert sind, ist ein Retourenportal häufig keine Option. Auch der Einsatz komplexer und weniger transparenter Methoden, wie z. B. fortschrittliche KI-Algorithmen, ist für KMU ohne eine Data Analytics-Abteilung nicht praktikabel. Stattdessen können einfache statistische

und regelbasierte Modelle verwendet werden, die im Bereich der Retourenprozesse Handlungsempfehlungen aufzeigen und leicht interpretierbare Ergebnisse liefern (vgl. Asdecker und Karl 2018, S. 45). Zudem ist zu sagen, dass der betrachtete Kooperationspartner ein selbst-entwickeltes ERP-System besitzt. Ein Retourenportal müsste daher auch selbst implementiert werden. Aus finanziellen Gründen werden somit weiterhin Retourenscheine im Paket vorgesehen, sodass sowohl der erste Prototyp aus Stevenson und Rieck (2022) und auch das neue Software-System mit dieser Vorgabe umgehen muss. Demzufolge sind vor dem Öffnen eines Retourenpaketes die rückgelieferten Artikel nicht bekannt. Digital vorhanden sind nur die bestellten Artikel der Kund:innen in den Bestelldaten des ERP-Systems. Durch die Verwendung der Bestellhistorie zu den Retourendaten ist eine Vorhersage der retournierten Artikelanzahl dennoch möglich, indem die eingetroffenen Pakete durch die Lagermitarbeitenden gescannt und die Daten im Software-System ausgewertet werden.

3 Logistischer Paradigmenwechsel

Die Europäische Kommission hat im Januar 2021 ein Whitepaper mit dem Titel „Industry 5.0: Towards a sustainable, human-centric and resilient European industry“ veröffentlicht, das eine Entwicklung zum fünften Paradigmenwechsel vorstellt: von Industrie 4.0 zu Industrie 5.0. Das Paper legt den Schwerpunkt auf eine nachhaltige, menschenzentrierte und widerstandsfähige europäische Industrie (vgl. Breque et al. 2021). Durch Übertragung der 4.0-Analogie auf die Logistik entstand das Konzept „Logistik 4.0“, das von vielen Unternehmen als Haupttreiber für den zukünftigen Geschäftserfolg angesehen wird (vgl. Grosse 2023, S. 8). Um exemplarisch die Frage zu beantworten, ob die Entwicklungen und digitalen Anstrengungen im Retoureneingang und bei der Retourenbearbeitung in den Bereich „Logistik 5.0 oder Logistik 4.1“ fallen, werden im Folgenden relevante Begrifflichkeiten beleuchtet.

Im Kern ist die Logistik 4.0 ein Konzept, in dem Informations- und Kommunikationstechnologien genutzt werden, um Prozesse, Objekte und Lieferketten zu verzahnen. Die Integration von Sensoren, Selbststeuerungstechniken, Cloud-Computing, Datenanalyse und KI ermöglicht eine Vernetzung, Kommunikation und Überwachung entlang der vorwärts- und rückwärtsgerichteten Wertschöpfungskette (vgl. Kagermann und Wahlster 2022, S. 5). Zusätzlich kann eine hohe Flexibilität bei individuellen Transporten und Umschlagsprozessen umgesetzt werden. Aufgrund der Integration von KI und selbstlernenden sowie lernfähigen Systemen können Prozesse smarter und autonomer werden (vgl. Bendig et al. 2021, S. 20–22). Die Logistik 5.0 gilt als Weiterentwicklung des Konzepts Logistik 4.0, wobei eine verstärkte Zusammenarbeit zwischen Mensch und Maschine, eine nachhaltige, ressourcenschonende und gesellschaftlich verantwortliche Logistik sowie eine noch stärkere Individualisierung und Flexibilität vorgesehen wird (vgl. Breque et al. 2021, S. 13–24). Logistik 4.0 und Logistik 5.0 setzen die gleichen Kerntechnologien ein, um Prozesse und Objekte effizienter und flexibler zu gestalten. Logistik 5.0 legt allerdings einen stärkeren Fokus auf den menschenzentrierten und nachhaltigen Ansatz. Somit gibt es laut Bendig et al. (2021) im Gegensatz zu früheren industriellen Revolutionen, die

durch die Einführung einer oder mehrerer neuartiger Technologien die Produktion und Logistik stark beeinflussten, keinen technologisch begründbaren Paradigmenwechsel. Logistik 5.0 könnte daher als eine evolutionäre Weiterentwicklung der Logistik 4.0 betrachtet und als „Logistik 4.1“ bezeichnet werden (vgl. Bendig et al. 2021, S. 20–22; Breque et al. 2021, S. 13–24). Da der Retoureneingang unseres Kooperationspartners noch wenig digitalisiert ist und sich das Unternehmen mit dem neu entwickelten Software-System gerade erst im Spannungsfeld der vernetzten Prozesse etabliert, ist die Frage „Logistik 5.0 oder Logistik 4.1: Wo stehen wir im Retoureneingang?“ deutlich mit „Logistik 4.1“ zu beantworten.

4 Erster Prototyp zur Digitalisierung des Retoureneingangs

Der betrachtete Kooperationspartner hat jährlich eine steigende Anzahl an Retourenpaketen zu verzeichnen. Dies liegt an einer stabilen Retourenquote des Unternehmens bei zunehmend höheren Anzahlen an Kundenbestellungen. Insgesamt sind im Jahr 2020 mehr als 73.000, im Jahr 2021 rund 120.000 und im Jahr 2022 über 127.000 Retouren im Bereich der Möbel und Wohnaccessoires im Lager eingegangen. Vor der Einführung des Prototyps erfolgte der Bearbeitungsprozess der Retouren nach dem FIFO-Verfahren. Durch den Prototyp war erstmals eine strukturierte Abarbeitung der Retourenpakete möglich, indem die Pakete nach dem Eintreffen in verschiedene Klassen vorsortiert und erst danach zur weiteren Bearbeitung an die Lagermitarbeitenden übergeben wurden. Durch den Einsatz des Prototyps verringerte sich in einer Evaluierung über mehrere Tage in 2022 die gesamte Bearbeitungsdauer täglich im Schnitt um 1,35 h. Zudem wurden die Pakete im Schnitt 49 Sek. schneller retourniert (vgl. Stevenson und Rieck 2022).

Eine Schwachstelle des ersten Prototypen ist allerdings der Umgang mit Retourenpaketen, die mehrere Artikel in den Daten der Originalbestellung enthalten. Die entsprechenden Retourenpakete werden sofort der Klasse „MR“ zugewiesen. Eine umfangreiche Analyse zeigte allerdings, dass in den meisten Fällen nur ein Artikel zurückgeschickt wird, was eigentlich eine Sortierung in die Klasse „SR“ erforderlich macht (vgl. Stevenson und Rieck 2022). Um dieses Problem zu lösen, wurden zunächst Methoden aus der Clusteranalyse und Klassifikation verwendet, um Features zu entwickeln, die für die Entwicklung einer KI-Lösung erforderlich sind. Leider war der Umfang der verfügbaren Daten im ERP-System so gering, dass es nicht möglich war, ausreichend Features zu extrahieren, beispielsweise für die Entwicklung eines neuronalen Netzes (vgl. Nufer 2022, S. 23–25). Dies haben wir als Anlass genommen, um die Klassifikation des Prototypen durch Anwendung eines regelbasierten Verfahrens zu verbessern und diesen als Software-Lösung kostengünstig in die IT-Infrastruktur einzubetten.

5 Entwicklung des neuen Software-Systems

Basierend auf dem bereits etablierten Prototypen (vgl. Kap. 4) soll der Retoureneingang und auch die -bearbeitung durch eine differenziertere Herangehensweise

bei der Klassifikation von MR-Paketen weiter verbessert werden. Hierbei wird ein Software-System implementiert, das aufgrund vorhandener und fortlaufend aktualisierter Kunden- und Bestelldaten Entscheidungen bezüglich der Vorsortierung trifft. Die Unterteilung eintreffender Retourenpakete in die Klassen „SR“ oder „MR“ ist von großer Wichtigkeit, weil dann die Retourenpakete dem Aufwand nach absteigend abgearbeitet werden können. Zudem können erfahrende Mitarbeitende die MR-Pakete erhalten.

Eine Retoure kann auf Basis eines Einzelkaufs (Bestellung mit = 1 Artikel) oder eines Kombinationskaufs (Bestellung mit > 1 Artikel) entstehen. Da die Klassifikation von Einzelkäufen in die Klasse „SR“ bereits reibungslos erfolgt, soll sie beibehalten werden. Bei den Kombinationskäufen muss die Klassifikation angepasst werden, da eine sofortige Sortierung in die Klasse „MR“ falsch sein könnte. Es ist erwähnenswert, dass aus nachhaltigen Gründen der Kooperationspartner keine Möglichkeit anbietet, mehrere Artikel getrennt zu versenden. Daher werden die zurückgesendeten Artikel typischerweise in einem Paket zusammengefasst. Im Jahr 2022 waren von den ca. 127.000 Retouren knapp 58 % Retouren aus Kombinationskäufen. Von den resultierenden 73.400 Retourenpaketen wurden ca. 45.000 Pakete falsch als „MR“ klassifiziert, obwohl nur ein Artikel enthalten war.

5.1 Problem der Klassifikation von MR-Paketen

Die fehlerhafte Klassifikation entsteht aufgrund des nicht-digitalen Retourenprozesses, bei dem ausgedruckte Retourenscheine verwendet werden (vgl. Kap. 2). Kund:innen geben auf dem Retourenschein die zurückzusendenden Artikel und den Grund für die Retoure der jeweiligen Artikel an. Der Retourenschein wird dann dem Retourenpaket beigelegt. Erst wenn das Paket im Retourenlager geöffnet wird, erfährt das Unternehmen, welche Artikel darin enthalten sind und wieso. Um dieses Problem zu lösen, wurde ein Software-System entwickelt, das auf Basis von fortlaufend aktualisierten Bestell- und Kundendaten sowie mithilfe eines Berechnungsschemas die Anzahl der retournierten Artikel bei Kombinationskäufen vorhersagt. Auf diese Weise wird eine präzisere Klassifikation von MR-Paketen ermöglicht.

5.2 Vorbereitung und Verwendung von Datensätzen

Bei der Entwicklung des Software-Systems wurden ausschließlich Kundenbestellungen mit 2 bis 7 Artikeln berücksichtigt. Bestellungen mit > 7 Artikeln machen weniger als 1,0 % aller Kombinationskäufe aus, weshalb sie zur Vereinfachung vernachlässigt werden. Für eine differenzierte Klassifikation innerhalb der Klasse „MR“ werden zunächst drei einzelne Datensätze durch einen Algorithmus erzeugt. Wie der entwickelte Algorithmus die Daten abfragt und transformiert, wird im Abschn. 5.5 genauer beschrieben. Für den ersten Datensatz (Datensatz Kunde) werden alle Kundennummern erfasst und die entsprechenden Retourenquoten (in %) der Kund:innen gespeichert. Die Berechnung der Retourenquote einer Kundin bzw. eines Kunden erfolgt, indem die Anzahl der Bestellungen durch die Anzahl der Retouren dividiert wird. Auf diese Weise kann das individuelle Retourenverhalten bestimmt werden, z. B. ob oft Auswahlkäufe mit der Absicht einer Retoure getätigt werden oder ob nur

Tab. 1 Datensätze und deren Inhalte

Datensatz Kunde	Datensatz Artikel	Datensatz Bestellung
<i>Gruppirt auf Kundenebene</i>	<i>Gruppirt auf Artikelebene</i>	<i>Gruppirt auf Bestellpositionsebene</i>
Kunden-ID	Artikel-ID	Bestell-ID
Retourenquote pro Kunde	Retourenquote pro Artikel	Kunden-ID
		Bestellpositions-ID
		Artikel-ID

bei defekten Produkten oder bei Produkten, die anders als erwartet ausfallen, retourniert wird. Wir verwenden hier die Informationen aus den vergangenen zwei Jahren, d.h. der Datensatz greift immer auf die letzten 24 Monate zu. Dieses Vorgehen ist legitim, da die gesetzliche Gewährleistung eine Frist von zwei Jahren vorgibt; ältere Artikel können auf Kulanz retourniert werden. Im zweiten Datensatz (Datensatz Artikel) werden sämtliche Artikel gespeichert, welche in den vergangenen zwei Jahren verkauft wurden. Für diese Artikel werden dann die Retourenquoten (in %) zusammengetragen. Aufgrund dessen, dass die Klassifikation von Kombinationskäufen verbessert werden soll, verwenden wir für das Berechnungsschema nur die Retourenquoten der Artikel aus Kombinationskäufen. Somit werden die Daten nicht durch Einzelkäufe verzerrt. Abschließend wird der dritte Datensatz (Datensatz Bestellung) generiert, der sämtliche Bestellungen mit den entsprechenden Artikeln umfasst. In diesem Datensatz werden für jede Bestellung und jede Bestellposition mit individueller Artikelnummer Zeilen in einer Matrix erstellt. Wenn Kund:innen drei verschiedene Artikel bestellt haben, enthält der Datensatz „Bestellung“ drei Zeilen, d.h. für jede entsprechende Bestellposition (1, 2, 3) eine Zeile. Der Inhalt und die Gruppierung der Datensätze kann Tab. 1 entnommen werden.

Mit Hilfe der drei Datensätze kann nun eine Vorhersage der Artikelanzahl in einem MR-Paket durchgeführt werden (vgl. Abschn. 5.3). Da bei den erforderlichen Berechnungen allerdings auch die Retourenquoten mit eingehen, müssen die Daten angepasst werden. Beispielsweise kann es passieren, dass Kund:innen im Vorfeld noch nichts retourniert haben oder ein Artikel erstmals nach Einlistung retourniert wird (Retourenquote jeweils = 0,0 %). Um in einem solchen Fall eine Multiplikation mit 0-Werten zu vermeiden, wird ein Glättungsverfahren eingesetzt. Durch die Anwendung des Glättungsverfahrens von Laplace wird jeder Wert um +1 erhöht (vgl. Jayaswal 2020). Auf diese Weise wird sichergestellt, dass auch bei seltenen Ereignissen aussagekräftige Ergebnisse und Vorhersagen erzielt werden können. Eine Verzerrung der Ergebnisse durch dieses Vorgehen wurde nicht beobachtet.

5.3 Vorhersage der Artikelanzahl in MR-Paketen

Weist eine Bestellung B eines Kunden C „>1 Artikel“ auf, kann ein dreistufiges Modell bzw. regelbasiertes Berechnungsschema durchlaufen werden, um die Anzahl an Artikeln in einem Retourenpaket vorherzusagen. Das Modell wurde in der Programmiersprache „Python“ implementiert. Im Rahmen der ersten Stufe des Modells werden zunächst die Retourenquote R^C des Kunden C mit den Retourenquote R_i der Artikel i (Item) aus der Bestellung durch Multiplikation in Relation gesetzt. Hierbei entsteht für jeden Artikel i ein Score (S_i^C). Die Betrachtung der Bestellung

Tab. 2 Dreistufiges Modell zur Vorhersage der Artikelanzahl in einer Retoure, die zur Bestellung B eines Kunden C gehört

Indizes:	Stufe 1: Berechnung der Scores für alle i	Stufe 2: Bestimmung des Thresholds für B	Stufe 3: Bestimmung der Vorhersage
$i = \text{Item } (i = 1, \dots, \text{Anzahl Bestellpositionen})$	$S_i^C = R^C \cdot R_i$	$T^B = \frac{\max_i(S_i^C)}{2}$	Für alle i : If $(S_i^C > T^B)$ then $\text{Ind}_i = 1$ Else $\text{Ind}_i = 0$ Geschätzte Anzahl Ω^B der Artikel in der Retoure zur Bestellung B: $\Omega^B \hat{=} \sum_i \text{Ind}_i$

findet auf Positionsebene statt, sodass jeder Artikel aus einer Bestellung abgebildet wird. In der zweiten Stufe wird ein im Rahmen einer Sensitivitätsanalyse erprobter Schwellenwert T^B für die Bestellung B berechnet (Threshold). Dabei wird der maximale Score über alle Artikel i von B gebildet und durch 2 geteilt. Der Threshold dient im Folgenden als Abschätzung, ob der Artikel i in dem zugehörigen Retourenpaket vorhanden ist oder nicht.

In der dritten Stufe wird für jeden Artikel ein binärer Indikator Ind_i gesetzt, der vorhersagt, ob Artikel i in der Retoure enthalten ist. Indikator Ind_i erhält den Wert 1, wenn der Score S_i^C von i den Threshold T^B übersteigt. Ist dies nicht der Fall, wird Ind_i gespeichert. Durch die Summe der erhaltenen Indikatoren wird abschließend die Anzahl Ω^B der Artikel in der Retoure vorhergesagt. Der Wert Ω^B für die zur Bestellung B gehörenden Retoure ergibt sich demzufolge durch die Kombination der Ergebnisse aus den Datensätzen „Kunde“, „Artikel“ und „Bestellung“. Insbesondere gehen die Retourenquoten der Kund:innen sowie die Retourenquoten der Artikel aus den Bestellpositionen in die Berechnung ein. Besitzt die Vorhersage einen Wert > 1 ($\Omega^B > 1$), dann wird die Retoure als „MR“ klassifiziert. Bei $\Omega^B \leq 1$ gilt die Retoure als „SR“. Bitte beachten Sie, dass Ω^B auch den Wert Null annehmen kann, wenn insbesondere die Retourenquote sehr klein sind. Wir gehen dann von einem Artikel im Paket aus und klassifizieren es als SR-Paket (Tab. 2).

Zur besseren Verständlichkeit und Nachvollziehbarkeit folgt ein numerisches Beispiel für das Berechnungsschema: Ein Kunde bestellt einen Beistelltisch, eine Vase und eine Lampe. Der Kunde hat eine Retourenquote von 10%, der Beistelltisch von 5%, die Vase von 10% und die Lampe von 30%. In der ersten Stufe werden die Scores berechnet werden: $0,1 \cdot 0,05 = 0,005$ (Beistelltisch), $0,1 \cdot 0,1 = 0,01$ (Vase) und $0,1 \cdot 0,3 = 0,03$ (Lampe). In der zweiten Stufe wird der Threshold bestimmt: $\max(0,005; 0,01; 0,03) \rightarrow 0,03/2 = 0,015$. Abschließend werden die Indikatoren durch die Abfrage $(S_i^C > T^B)$ ermittelt und aufsummiert: Ist $0,005 > 0,015? \rightarrow \text{Ind} = 0$, ist $0,01 > 0,015? \rightarrow \text{Ind} = 0$ und ist $0,03 > 0,015? \rightarrow \text{Ind} = 1$; Summe der Indikatoren = 1. Somit wird das Retourenpaket als „SR“ klassifiziert.

5.4 Evaluation des neuen Software-Systems

Um eine Verbesserung durch das neue Software-System feststellen zu können, wurde im Rahmen der Evaluierung im Data Warehouse (DWH) unseres Kooperationspart-

Tab. 3 Beispiel der Evaluierungstabelle im DWH

Bestell-ID	Anzahl an bestellte Artikel	Retouren-ID	Anzahl an retournierte Artikel	True	Excel	Klassifikation Excel	Python	Klassifikation Python
123456	2	345678	1	SR	MR	–	SR	✓
123457	2	345679	2	MR	MR	✓	MR	✓
123458	3	345680	2	MR	MR	✓	SR	–
123459	7	345681	1	SR	MR	–	MR	–
...

ners eine Tabelle mit 15.148 Retouren und dazugehörigen Bestellungen mit mehreren Artikeln geladen und ausgewertet. Zum Vergleich verwenden wir das Excel-basierte Klassifikationsmodell, welches im Prototyp in Stevenson und Rieck (2022) implementiert wurde und fortan als „Excel“ bezeichnet wird. Das neu entwickelte Modell, das im Folgenden als „Python“ bezeichnet wird, basiert auf dem Berechnungsschema aus Abschn. 5.4. Zusätzlich zur Evaluierung wurden retrospektiv für alle Retourenpakete die tatsächlichen Klassen anhand der bearbeiteten Artikelanzahl der Retourenpakete ermittelt. Diese Ergänzung ermöglicht eine genaue Überprüfung der Klassifikation. Die resultierende Evaluierungstabelle ist in Tab. 3 zu sehen. Für jede Bestellung (mit entsprechender Bestell-ID) wird die Anzahl der bestellten Artikel angegeben und für die zugehörige Retoure (mit entsprechender Retouren-ID) wird die Anzahl der retournierten Artikel ausgewiesen. Anhand dieser Information kann die jeweils korrekte Klassifikation „True“ der eingegangenen Retourenpakete verzeichnet werden. Die Einträge in den weiteren Spalten weisen die Klassifizierung gemäß den entsprechenden Modellen (Excel und Python) aus. Zudem wird eine richtige Klassifikation mit dem Maker „✓“ markiert. Die relative Verteilung der Artikelanzahlen der betrachteten Bestellungen mit mehreren Artikeln waren: 2 Artikel: 43,2%; 3 Artikel: 24,4%; 4 Artikel: 14,4%; 5 Artikel: 9,1%; 6 Artikel: 5,6% und 7 Artikel: 3,1%.

Aufgrund der unterschiedlichen relativen Verteilung der Artikelanzahlen wurden die Klassifikationen der Modelle auf die einzelnen Artikelanzahlen gruppiert und ausgewertet. Zur Berechnung der Fehlerquote einer Bestellung mit $x=2, \dots, 7$ Artikeln wird pro Modell die Anzahl an Fehlern (d.h. die Spalte „Klassifikation“ für das entsprechende Modell enthält keine Markierung) mit der Gesamtanzahl an Zeilen mit x Artikeln ins Verhältnis gesetzt. Auf diese Weise erhalten wir die Werte in Tab. 4 und können ableiten, dass Excel mit 9146 falsch klassifizierten Retourenpaketen fehleranfälliger ist als Python mit nur 8224.

Die Ergebnisse der gruppierten Datensätze zeigen deutlich, dass eine verbesserte Klassifikation mit dem Modell „Python“ erreicht wurde. Insgesamt konnten 923 Pakete zusätzlich richtig nach „SR“ und „MR“ sortiert werden. Unter der Annahme, dass die Ergebnisse auf das Jahr 2022 übertragbar sind, wären mit dem neuen Modell insgesamt zusätzlich ca. 5000 Retouren korrekt klassifiziert worden. Die Verbesserung wäre demnach für das Retourenlager deutlich spürbar und würde insgesamt zu geringeren durchschnittlichen Bearbeitungszeiten der Mitarbeitenden führen.

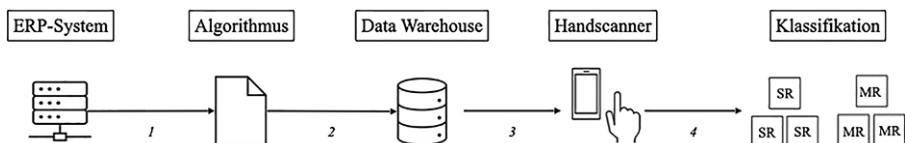
Tab. 4 Evaluierungsergebnis vom Software-System gegenüber Prototypen

Artikelanzahl in Bestellung	Anzahl der Fehler in Excel	Anzahl der Fehler in Python	Verbesserung (in Anzahl Pakete)	Verbesserung (in %)
2	4150	3374	776	18,7
3	2423	2329	94	3,9
4	1299	1271	28	2,2
5	751	732	19	2,5
6	523	518	5	1,0
7	315	314	1	0,3

5.5 Applikation und Integration des Software-Systems

Die in Abb. 1 dargestellte Systemarchitektur des neuen Software-Systems wurde an die Gegebenheiten der IT-Infrastruktur des Kooperationspartners eingepasst. Für die Entwicklung standen die Daten aus dem ERP-System und aus dem DWH zur Verfügung. Die Systemkomponenten wurden nach Erstellung automatisiert, sodass keine manuelle Arbeit am System mehr nötig ist. Auf diese Weise kann das System kontinuierlich und automatisiert die aktuellsten Daten verwenden und anhand derer neue Vorhersagen für die Lagermitarbeitenden ermitteln. Das System zur Klassifikation von MR-Paketen umfasst folgende wesentliche Aufgaben: Datenabfrage, Klassifikation, Datenzugriff und Sortierung.

Die Datenabfrage (1) ist mittels SQL im Algorithmus eingebettet und fragt alle notwendigen Daten zur Generierung der drei Datensätze (vgl. Abschn. 5.2) ab. Die Daten werden direkt aus dem ERP-System entnommen, da das DWH nur einmal täglich beladen wird und daher nicht über alle Bestelldaten verfügt. Hierauf folgt die Klassifikation (2). An dieser Stelle im Algorithmus wird für jede Bestellung mithilfe des dreistufigen Berechnungsschemas (vgl. Abschn. 5.3) eine Vorhersage für die Einordnung in „SR“ und „MR“ durchgeführt. Dies muss für jede Bestellung B passieren, weil im Vorfeld nicht bekannt ist, welche Bestellung retourniert wird. Für die spätere Nutzung werden somit die Ergebnisse der in 5.3 beschriebenen Berechnungen vorberechnet und in eine neue Tabelle ins DWH geschrieben. Auf diese Weise wird dafür gesorgt, dass bei der Verwendung des Systems die Daten nicht ad hoc berechnet werden müssen. Es werden im DWH also die Bestell-IDs dupliziert und um die beiden Informationen ergänzt: (a) vorhergesagte Artikelanzahl, die retourniert wird und (b) resultierende Klasse „SR“ oder „MR“. Durch dieses Vorgehen kann eine Applikation auf den Handscannern über ein einfaches Interface auf die Daten im DWH zugreifen (3). Die Lagermitarbeitenden scannen nun die eingetrof-

**Abb. 1** Infrastruktur des neuen Software-Systems

fenen Retourenpakete und sortieren diese hinsichtlich der ermittelten Klassifikation in „SR“ oder „MR“ (4). Durch dieses Vorgehen werden die bereits in Stevenson und Rieck (2022) beschriebenen Effektivitäts- und Produktivitätssteigerungen nochmals verstärkt, Bearbeitungszeiten reduziert, die Wirtschaftlichkeit des Retourenlagers verbessert und die Digitalisierung des Retourenlagers vorangetrieben.

6 Fazit und Ausblick

Im vorliegenden Beitrag wurde ein Software-System implementiert, das die retournierte Artikelanzahl aus Bestellungen vorhersagt und hierdurch eine differenziertere Vorsortierung in die Klassen „SR“ und „MR“ ermöglicht. Das angestrebte Ziel des Software-Systems war es, eine Verbesserung der Prozesse im Retoureneingang gegenüber dem vorherigen Prototyp zu realisieren und sich dem Konzept Logistik 4.1 anzunähern. Dies kann als gelungen angesehen werden. Durch das Software-System konnten die Prozesse im noch wenig digitalisierten Retourenlager des kooperierenden KMUs gezielt zusammengefasst werden. Auch die Einbettung in die IT-Infrastruktur konnte ohne hohe Mehrkosten durchgeführt werden. Mit Hilfe einer Evaluierung im Rahmen einer Case Study wurde der Erfolg belegt. Als Handlungsempfehlung lässt sich für Onlinehändler, die mit Retourenscheinen arbeiten, formulieren, dass eine FIFO-Bearbeitung von Retourenpaketen hinterfragt werden sollte. Werden eingetroffene Retourenpakete vorsortiert und Retouren, die eine komplexe Bearbeitung erfordern, den erfahrenen Mitarbeitenden zugeordnet, ergeben sich im Schnitt geringere Bearbeitungszeiten.

Für die weitere Forschung besteht die Möglichkeit, den im Berechnungsschema eingeführten Threshold T^B fein zu justieren und weiter zu entwickeln. Zudem könnte bspw. durch die Integration zusätzlicher Datensätze oder die Anwendung von Machine-Learning-Modellen eine verbesserte Klassifikation erreicht werden. Hierfür wäre ein Ausbau der verfügbaren Retourendaten im ERP-System notwendig. Eine weitere Möglichkeit, den Ausbau der Daten vorausgesetzt, wäre die Anwendung von Deep-Learning-Modellen, die in der Lage sind, komplexe Muster und Zusammenhänge in Daten zu erkennen und dadurch eine höhere Vorhersagegenauigkeit zu erzielen. Das vorgestellte Vorgehen des Software-Systems kann außerdem auf andere Unternehmen und Branchen transferiert werden. Hierbei wäre zu prüfen, inwieweit die Ergebnisse aus der Case Study sich reproduzieren lassen. Hierfür könnten Unternehmen mit ähnlichen Prozessen und Datenstrukturen analysiert und ein Leitfaden für die Übertragbarkeit der Vorgehensweise erstellt werden.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung

nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Literatur

- Ager V Wie du mithilfe eines Retourenportals Retouren reduzieren kannst. *epoq.* <https://www.epoq.de/blog/retourenportal-retouren-reduzieren/> (Erstellt: 8. Dez. 2021). Zugegriffen: 22. Juni 2023
- Asdecker B, Karl D (2018) Big data analytics in returns management—Are complex techniques necessary to forecast consumer returns properly? Proceedings of the 2nd International Conference on Advanced Research Methods and Analytics (CARMA 2018). <https://doi.org/10.4995/CARMA2018.2018.8303>
- Bendig D, Lau K, Schulte J, Endriß S (2021) Industrie 5.0 – Die Europäische Kommission auf den Spuren der nächsten industriellen Revolution? *Ind 40 Manag* 2021(6):20–22. https://doi.org/10.30844/I40M_21-6_S20-22
- Bold L QR-Code Retour: Rücksendung mit QR-Code als neue Retourenlösung. *shipcloud.io.* <https://blog.shipcloud.io/ruecksendung-mit-qr-code/> (Erstellt: 30. Aug. 2021). Zugegriffen: 25. Juni 2023
- Breque M, de Nul L, Petridis A (2021) Industry 5.0: towards a sustainable, human-centric and resilient European industry. Publications Office of the European Union
- Chen X, Wang T, Thomas B, Ulmer M (2022) Same-day delivery with fair customer service. *Eur J Oper Res* 308:738–751. <https://doi.org/10.1016/j.ejor.2022.12.009>
- Clotey T, Benton WC Jr., Srivastava R (2012) Forecasting product returns for remanufacturing operations. *Decis Sci* 43(4):589–614. <https://doi.org/10.1111/j.1540-5915.2012.00362.x>
- Cui H, Rajagopalan S, Ward AR (2020) Predicting product return volume using machine learning methods. *Eur J Oper Res* 281(3):612–627. <https://doi.org/10.1016/j.ejor.2019.05.046>
- Dai H, Xiao Q, Yan N, Xu X, Tong T (2022) Item-level forecasting for E-commerce demand with high-dimensional data using a two-stage feature selection algorithm. *J Syst Sci Syst Eng* 31(2):247–264. <https://doi.org/10.1007/s11518-022-5520-1>
- Eierdanz F, Blaudszun-Lahm A (2020) Mitarbeiterzufriedenheit im Rahmen digitaler Transformationsprozesse, S 85–96 https://doi.org/10.1007/978-3-662-61372-6_9
- Girschner S (2022) Retourenportal: 10 Tipps, wie Händler die Prozesse verbessern können. *eCommerce Magazin.* <https://www.e-commerce-magazin.de/retourenportal-10-tipps-wie-haendler-die-prozesse-verbessern-koennen/> (Erstellt: 29. Nov. 2022). Zugegriffen: 22. Juni 2023
- Grosse EH (2023) Logistik 4.0: Stand der Forschung und Praxis. *HMD* 60(1):6–20. <https://doi.org/10.1365/s40702-022-00931-3>
- Günther N, Prell B, Reiff-Stephan J (2022) Industrie 5.0 – Von der Vision des menschenzentrierten Ansatzes zu sozio-cyberphysischen Produktionssystemen für die Praxis (Open Access)
- Jayaswal V (2020) Laplace smoothing in naïve Bayes algorithm—towards data science. <https://towardsdatascience.com/laplace-smoothing-in-na%C3%AFve-bayes-algorithm-9c237a8bdece>. Zugegriffen: 19. Juni 2023
- Kagermann H, Wahlster W (2022) Ten years of Industrie 4.0. *Sci* 4(3):26. <https://doi.org/10.3390/sci4030026>
- Krapp M, Nebel J, Sahamie R (2013) Forecasting product returns in closed-loop supply chains. *Int J Phys Distribution Logist Manag* 43(8):614–637. <https://doi.org/10.1108/IJPDLM-03-2012-0078>
- Liang X, Jin X, Ni J (2014) Forecasting product returns for remanufacturing systems. *J Remanufacturing* 4(1):8. <https://doi.org/10.1186/s13243-014-0008-x>
- Nufer G (2022) Der Einsatz Künstlicher Neuronaler Netze zur Analyse von Markenattributen hinsichtlich der Markenbewertung im Sportmarketing. *Sciamus* <https://doi.org/10.24403/JP.1244733> (dataset)
- Schlenkrich M, Parragh S (2023) Solving large scale industrial production scheduling problems with complex constraints: an overview of the state-of-the-art. *Procedia Comput Sci* 217:1028–1037. <https://doi.org/10.1016/j.procs.2022.12.301>
- Stevenson AB, Rieck J (2022) Optimierung der Prozesse im Retoureneingang: E-Commerce Case Study für den B2C-Bereich. *HMD.* <https://doi.org/10.1365/s40702-022-00934-0>
- Strahinger S, Westner M (2022) Von der IT-Leitung zur (Mit-)Gestaltung der digitalen Transformation: Die sich wandelnde Rolle der Führung in der IT. *HMD* 59(3):717–720. <https://doi.org/10.1365/s40702-022-00875-8>

- Toktay LB, van der Laan EA, de Brito MP (2004) Managing product returns: the role of forecasting. In: Dekker R, Fleischmann M, Inderfurth K, Van Wassenhove LN (Hrsg) Reverse logistics: quantitative models for closed-loop supply chains. Springer, Berlin Heidelberg, S 45–64 https://doi.org/10.1007/978-3-540-24803-3_3
- Trippner C (2021) Wie funktioniert Retouren-vermeidung im Online-Shopping durch KI? Digit Welt 5(3):43–45. <https://doi.org/10.1007/s42354-021-0363-3>
- Wang C-N, Dang T-T, Nguyen N-A (2021) Outsourcing reverse logistics for E-commerce retailers: a two-stage fuzzy optimization approach. Axioms 10(1):Article 1. <https://doi.org/10.3390/axioms10010034>
- Winkelhaus S, Grosse EH (2020) Logistics 4.0: a systematic review towards a new logistics system. Int J Prod Res 58(1):18–43. <https://doi.org/10.1080/00207543.2019.1612964>
- Yang G, Ji G, Tan K (2022) Impact of artificial intelligence adoption on online returns policies. Ann Oper Res. <https://doi.org/10.1007/s10479-020-03602-y>
- Zhou L, Xie J, Gu X, Lin Y, Ieromonachou P, Zhang X (2016) Forecasting return of used products for remanufacturing using Graphical Evaluation and Review Technique (GERT). Int J Prod Econ 181:315–324. <https://doi.org/10.1016/j.ijpe.2016.04.016>
- Zhu Y, Li J, He J, Quanz BL, Deshpande AA (2018) A local algorithm for product return prediction in E-commerce, S 3718–3724