



Identifying Application Areas for Machine Learning in the Retail Sector

A Literature Review and Interview Study

Clemens Brackmann¹ · Marek Hütsch¹ · Tobias Wulfert¹ 

Received: 30 October 2022 / Accepted: 4 May 2023
© The Author(s) 2023

Abstract

Machine learning (ML) has the potential to take on a variety of routine and non-routine tasks in brick-and-mortar retail and e-commerce. Many tasks previously executed manually are amenable to computerization using ML. Although procedure models for the introduction of ML across industries exist, the tasks for which ML can be implemented in retail need to be determined. To identify these application areas, we followed a dual approach. First, we conducted a structured literature review of 225 research papers to identify possible ML application areas in retail, as well as develop the structure of a well-established information systems architecture. Second, we triangulated these preliminary application areas with the analysis of eight expert interviews. In total, we identified 21 application areas for ML in online and offline retail; these application areas mainly address decision-oriented and economic-operative tasks. We organized the application areas in a framework for practitioners and researchers to determine appropriate ML use in retail. As our interviewees provided information at the process level, we also explored the application of ML in two exemplary retail processes. Our analysis further reveals that, while ML applications in offline retail focus on the retail articles, in e-commerce the customer is central to the application areas of ML.

Keywords Retail · Electronic commerce · Machine learning · Application area · Literature review

Introduction

It is estimated that by 2025 more than 150,000 mobile robots will be deployed in retail to execute routine work such as refilling shelves and preparing package dispatch, making use of machine learning (ML) [1, 50]. However, ML cannot only be implemented to overtake cognitive routine tasks involving rule-based tasks [25], but can also spread into domains previously defined as non-routine tasks (e.g., handwriting recognition or car driving) [83]. Furthermore, ML has different applications such as X-ray image analysis [70], micro-organism image analysis [98], histopathology image analysis [13] and cancer image analysis [12]. A study by Frey and Osborne [23] identified that many tasks in online and offline retail are amenable to computerization. Although retail in both offline and online environments involves a variety of

tasks requiring different capabilities, a study by McKinsey revealed a high potential for automatizing in general and ML in retail [55], which is why we focused on application areas for this. The focus in retail is on primary and value-adding tasks that are described by Schütte as “economic-operative tasks” [55, 77].

Brick-and-mortar retailers have faced increasing competition in recent years, especially from online competitors. This process has been intensified in particular by the COVID-19 pandemic [63]. While the use of ML has become strongly established in e-commerce, the pressure on brick-and-mortar retailers is also increasing in this area [27]. While brick-and-mortar retailers manage products, logistics, etc. at store level, online retailers focus on optimizations at the customer level [77]. E-commerce involves much data regarding customers and unstructured data regarding products (product reviews, products created by users on a marketplace). The situation in brick-and-mortar retail is different [17, 33, 65]. The majority of data is present in a structured manner and data about articles are often self-managed by product manufacturers. In addition, there is only a limited amount of data

✉ Clemens Brackmann
clemens.brackmann@icb.uni-due.de

¹ Universitätsstraße 2, 45141 Essen, Germany

about the customers, potentially limiting the introduction of ML to only a few application areas. ML usually refers to either classification or regression, whereas classification is not relevant for our approach, although it can be used in retail, since we focus on revenue which can only be predicted with regression models.

The ML algorithm selection problem—also known as the combined algorithm selection and hyperparameter optimization (CASH) problem—poses a challenge to all ML practitioners [81]. This not only involves the selection of an algorithm that fits the data set and the economic problem of the retailer, but also the model selection and hyperparameter optimization [8]. Due to the abundance of possible combinations of these factors and their influence on the quality of the result, it is essential to choose fitting factors (i.e., algorithm, model, and hyperparameter) that fit the underlying data as best as possible [81]. Recently, however, various automated ML tools have been created for making these decisions, which greatly facilitate the real-world application of ML methods [22]. Thus, even practitioners who are non-experts in the various ML disciplines can achieve good results using the automated ML tools and available procedure models for introducing ML. Although these tools exist, it is still necessary to have experts who determine which tasks can be supported or even automated by ML. Experts also need to be able to create increasingly complex ML models themselves [33]. ML experts can be supported by industry-specific frameworks indicating possible ML applications. Such frameworks can resemble an entry point for ML implementation in online and offline retail. However, reviews in the existing literature on the application of ML in retail are usually limited to one application area and focus only on online or offline retail [10]. Against this backdrop, we propose the following research question:

What tasks in (online and offline) retail can be supported by machine learning?

To address this research question, we provide an overview of multiple application areas for ML in retail using a dual approach. We conducted an application-centered literature review for extracting possible ML applications for retail [88]. We triangulated these preliminary application areas with expert interviews on the usage of ML in retail processes [60]. We made use of an agreed-upon reference model in retail to structure the identified ML application areas, along with retail-specific tasks [77]. This paper contributes to the development of a customized strategy for ML in retail. Hence, we propose decision-making support for retailers on ML usage in specific application areas.

This research paper is an extended version of a paper presented at the 24th International Conference on Enterprise Information Systems [34]. For this special issue of SN Computer Science we modified the introduction, conducted an additional interview study, triangulated the literature-based

application areas with state-of-the-art insights from retail practice, demonstrate our results with selected retail processes, improved the discussion section, and streamlined the concepts used throughout the text.

The remainder of this research paper is structured as follows: in section 2, we elaborate on fundamentals about retail information systems in general and ML in particular. Next, we discuss our structured literature review and expert interview approach. We present the identified application areas of ML in retail in section 4, and then we demonstrate the application of ML in retail by means of two process examples in section 5. We discuss current and future applications of ML in online and offline retail in section 6. The research paper concludes with a brief conclusion and an outlook for future research.

Theoretical Background

In this section, we uncover related literature for retail information systems and machine learning.

Retail Information Systems

Since the advent of the Internet, existing retailers have either extended their traditional in-store business by an additional online channel or new companies that have started up have implemented an e-commerce business model, selling articles online without operating any brick-and-mortar stores [73]. Regardless of the institutional-economic discussion on retail companies, the tasks intended for trading functions are economically necessary [47]. There are three basic functions bridging the gap between manufacturer and customers in the streams of real goods (goods, services, returns), nominal goods (money, credits), and information across space, time, quantity, and quality [5]. Facilitated by the ongoing digitalization of the retail sector, the three basic functions increasingly cope with digital product and price information, payment, logistics, and distribution processes [48].

In e-commerce, trading transactions are carried out digitally to some degree [47]. The degree to which transaction must be digitalized varies in the literature on a continuum between a completely digital transaction and digital functions constituting only a small part of the procurement process [47]. Information systems in retail support the execution of the three trading functions and related tasks. They support the operational-dispositive, the business administration-administrative, and the controlling, as well as corporate planning tasks [6]. They extend the components of merchandise management systems (merchandise planning, logistics, and settlement processes) by using business intelligence and necessary corporate-administrative tasks in an integrative manner to conduct

the business processes of a retail company [6, 77]. The necessary tasks of a retailer and supporting information systems can be depicted in process models and reference models for the retail industry. Reference models comprise a high-level sketch of the system. They can depict the application architecture of a specific company or only part of its application architecture [29]. The current body of literature includes a number of reference models describing retail information systems [3, 6, 76]. In this article, we focus on the shell model, as proposed by Schutte, as it covers the whole value chain, from manufacturer, wholesaler, and retailer to the end consumer, focusing on tasks at the business level, and is designed for both online and offline environments. The shell model is a task-oriented reference model for retail [76]. We use this task-oriented perspective to determine the tasks for which ML is already applied, according to the scientific literature. It follows the principle of the separation of tasks and actors, introduced by Ferstl and Sinz [21]. The shell model copes with the identified shortcomings of the H-model, such as the artificial separation between merchandise management and decision support systems. It consists of four separate but intertwined architectures for each of the aforementioned actors along a value chain. The retail architecture consists of five shells for the main retail tasks (master data, technical tasks, economic-operative tasks, administrative tasks, and decision-oriented tasks) of each actor, and the retailer in particular [77]. Each shell consists of a series of tasks that form the components of the architecture. The tasks relevant to structuring the application areas are described in the analysis section (Sect. 4). The shell model is depicted in Fig. 1.

The master data on the article, supplier, customer, condition, site, and warehouse form the basis for technical tasks, peripheral control, and user access technology. The economic-operative tasks consist of merchandise management, in which the purchasing and sales sides are brought together by uniform marketing. The product range must be defined, the suppliers must be selected, the purchase price is negotiated, and the sales prices and sales performance are determined [92]. Based on the tasks of merchandise management, procurement and distribution processes are controlled with the functions "plan," "provide," and "check invoice." The distribution of goods or the sales process culminates in serving the customer. The cashing and invoicing processes follow serving and lead to the accounts receivable accounting. The goods are then transported to the customer. Additionally, inventory management as part of managing goods refers to the management and optimization of goods and storage, respective the inventory. The administrative tasks consist of the three basic processes of financial accounting, cost and activity accounting, and human resource management. The decision-oriented tasks

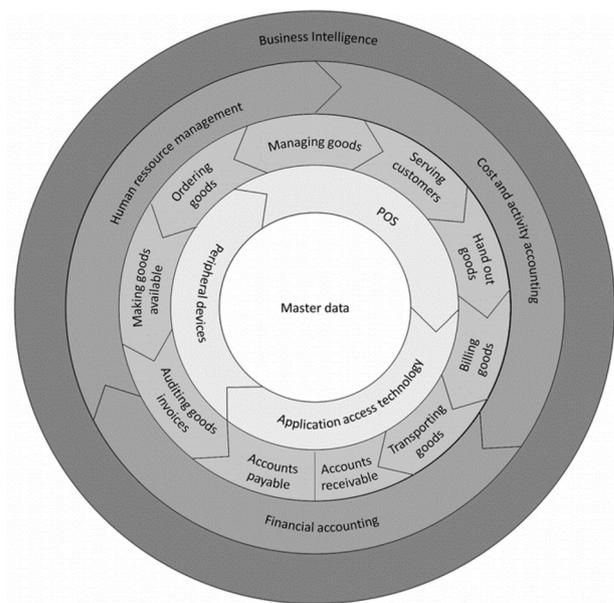


Fig. 1 Retail Information System Architecture [34, 76, 77]

(i.e., business intelligence) provide aggregated information to the management of a retailer [77].

Machine Learning

Historically, computerization has largely been confined to manual and cognitive routine tasks involving explicit rule-based activities [25]. Following recent technological advances, ML is now spreading to domains commonly defined as non-routine [83]. As these non-routine tasks are commonly executed by human employees, this paper reviews research articles to investigate which of these non-routine tasks in offline and online retailing are supported by ML in terms of research and are thus potential candidates for supporting humans with ML in the future.

The definition of ML we use describes ML as follows: "a computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E" [61]. To fulfill this definition, an ML application needs an optimization algorithm, an error function, a model, and a data set. A common way of using ML is illustrated by the framework shown in Fig. 2. This framework consists of six phases, where the first two phases, *Business Insight* and *Data Understanding* focus on the understanding of the context and data where the ML application will be used. The *Data Preparation* phase focuses on the implementation of algorithms for data augmentation and transformation, as well as the preparation for their use in the following phase, the *Modeling Phase*, where the actual implementation and selection of the ML model

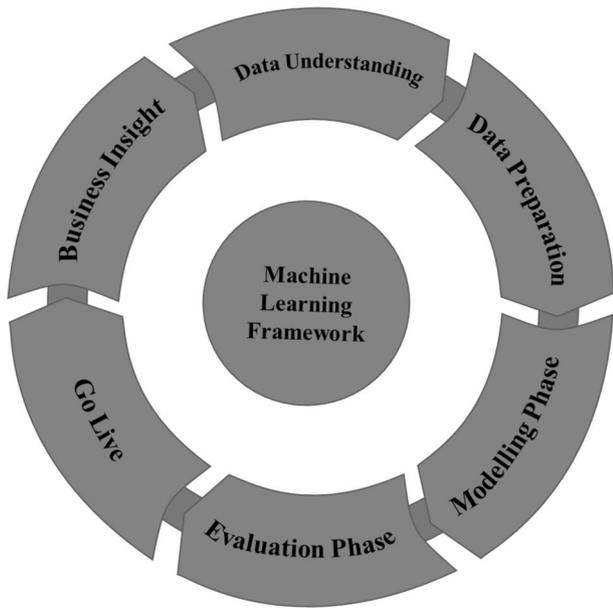


Fig. 2 Typical Machine Learning Framework inspired by Nguyen et al. [62]

and algorithm take place. Next, the Evaluation Phase focuses on testing and refining the previously created ML models, and finally, the *Go Live* phase focuses on the introduction of the ML application in the intended context [62]. In the specific case of retailing, it is the introduction in retail-related business processes with the intention of improving process efficiency and effectiveness.

As we use some ML vocabulary in the remaining paper, we clarify these terms below. A “forecast” is defined by historical time-series data that are used to make assumptions about possible events in the future. A ‘classification algorithm’ makes statements about a data set and assigns these data to a given set of classes/labels. There are many other ML tasks, which are mostly incorporated with a combination of forecasts and classification algorithms, which we mostly refer to as “analysis” in this paper. In a similar work, Weber and Schütte [93] inductively determined that artificial intelligence use cases for brick-and-mortar retail from real-world examples. Although the broader term of artificial intelligence (AI) is used there, application areas of ML are also

included. Just as in this paper, Weber and Schütte [93] used the shell model as a starting point for the literature review (Fig. 1). Both papers complement each other by identifying ML applications from practical contributions [92] and from analyzing scientific contributions. The different approach of this paper makes sense, as the underlying research question is another one. In this paper, the potential application areas of ML in retail are reviewed by analyzing research papers as a framework for practitioners and researchers. Another related work by Bousqaoui et al. [10] focused on the use of ML in the domain of supply chain management.

Scientific Approach

To identify relevant application areas of ML in retail (online and offline), we followed a dual approach. We identified possible application areas of ML in retail from the scientific literature by performing a systematic literature review [88]. In addition, we augmented these theory-driven application areas with eight interviews with domain experts [60].

Systematic Literature Review

To determine the theory-driven application areas of ML in retail, we followed the structured literature review approach proposed by vom Brocke et al. [88]. Following Cooper [15], our literature review can be characterized as follows: We focused on research outcomes (use cases and ML algorithms for retail) and applications of ML in retail. We aimed to integrate the body of literature and generalize application areas from the available research. We objectively represent articles from our exhaustive literature review with citations of selected papers for each application area. In this paper, we collect results from previous works and integrate them into an overarching framework. Our research represents scholars specialized in the use of ML in retail and e-commerce and practitioners implementing ML algorithms.

We identified keywords from the related literature, and an initial literature screening resulted in a generic search query that consisted of a retail and an ML part (see Fig. 3). As we derive existing application areas of ML in retail from relevant journals and conference papers and propose

Fig. 3 Structured literature review design [34]

Research areas and search string		Literature screening and selection					
Machine Learning	Retail	Databases	ACM	AISeL	IEEE	Scopus	WoS
Mitchell et al. (1997), David and Dorn (2013), Goos et al. (2009)	Becker and Schütte (2004), Laudon and Traver (2020)	Initial Sample	84	79	224	4,902	249
		Abstract Check	38	14	30	238	90
		Content Check	11	9	22	156	60
		Duplicate Check	197				
		For-/Backward	28				
		Final Sample	225				

avenues for further ML application in this domain, we chose a broad literature scope. This broad scope was underpinned by the search engines SCOPUS, Web of Science (WoS), IEEE Xplore (IEEE), the AIS electronic library (AISeL), and ACM. We queried their underlying databases, which consist of research from, among others, the economics and computer science literature. An initial query for titles and keywords on 2021-04-15 resulted in an initial sample of 5,538 papers. To ensure an appropriate level of quality, we focused on the scientific literature and added additional quality criteria to the search [71]. We excluded non-English and non-German language articles; panels and commentaries; purely technical articles (e.g., articles that focus exclusively on technological aspects without applying them in an e-commerce context); and articles with a purely e-commerce focus (e.g., articles that focus exclusively on e-commerce or sub-types without the inclusion of conceptual models). In contrast to previous research, such as Weber and Schutte [92], we intentionally excluded white papers and practitioner reports. Based on the title and abstract, considering the quality criteria, we reduced our sample to 410 papers. Further applying the approach of Bandara et al. [4] to the content of the papers, we identified 261 relevant publications, leaving 197 after the exclusion of duplicates. We used a one-time forward and backward search to identify research presented as an alternative scenario, resulting in 28 additional papers and resulting in a final set of 225 papers (see Fig. 3).

The articles in the final set were independently analyzed by full-text screening, and the relevant text passages, including possible application areas, were coded to refine the set of application areas. These application areas were synthesized to ensure the accuracy of our approach. The shell model, with its central master data objects (first shell), technical tasks (second shell), operational tasks (third shell), administrative tasks (fourth shell), and decision-oriented tasks (fifth shell) [76], served as the basis for the coding, which was inductively adapted as needed. The initial set of possible application areas was refined, using our final set of papers and two independent researchers in three coding rounds. After each round, the researchers met via

online communication media to discuss the application area refinements. The framework in Fig. 4 depicts our preliminary results, including 20 theory-driven application areas for ML in retail.

Expert Interviews

The preliminary application areas identified by the literature review were augmented by a series of interviews with domain experts [75]. During the interviews, we addressed the application of ML in retail-specific processes in a semi-structured manner. In total, we conducted eight interviews and discussed a total of seven processes. To create as broad and diverse a knowledge base as possible, we generalized information from a diverse set of interviewees [35]. The experts work at consulting agencies, with online and offline retailers as clients, and in close cooperation with software providers, whose software they implement. The interviewees have an average experience of 18 years as consultants in retail in a variety of process areas (e.g., finance, sales and distribution, forecast, and replenishment). The interviews were conducted online between 2022-06-08 and 2022-06-30 on regular workdays between 12:00 and 17:00, with an average interview duration of 59 min. Descriptive information about the interviews are presented in Table 1.

Our semi-structured interviews were based on a previously known guide to ensure that all conversations covered a similar range of questions regarding the investigated processes while allowing for the greatest possible flexibility in exploring explicit knowledge [58, 60]. In addition to demographic data, we also asked about the role of the interviewees within the companies, the retail sector they work in and their ML knowledge. Then we focused on the processes in which ML can be used. The rest of the interview was conducted through exploratory questions. We defined four leading questions for the interviews: "In which scenario can AI be applied?", "What processes are part of that scenario?", "What are the relevant input data for this scenario?", and "What is the value added through the usage of ML in this scenario?". The semi-structured

Table 1 Overview of interviewees

#	Interviewee's role	Processes	Process area	Domain experience	Date	Time
1	Senior consultant	Audit	Finance	~9 years	22-06-08	12:02–12:59
2	Managing partner	Returns	Sales & distribution	~23 years	22-06-08	13:15–14:19
3	Senior consultant	Procurement	Forecast & replenishment	~6 years	22-06-13	12:32–13:27
4	Senior consultant	Inventory planning	Sales & distribution	~20 years	22-06-16	16:01–16:57
5	Senior consultant	Promotions	Sales & distribution	~27 years	22-06-22	14:36–15:36
6	Solution architect	Distribution	Sales & distribution	~29 years	22-06-23	12:01–12:58
7	Consultant	Contingency	Sales & Distribution	~16 years	22-06-27	13:34–14:39
8	Consultant	Distribution	Sales & distribution	~12 years	22-06-30	15:29–16:25

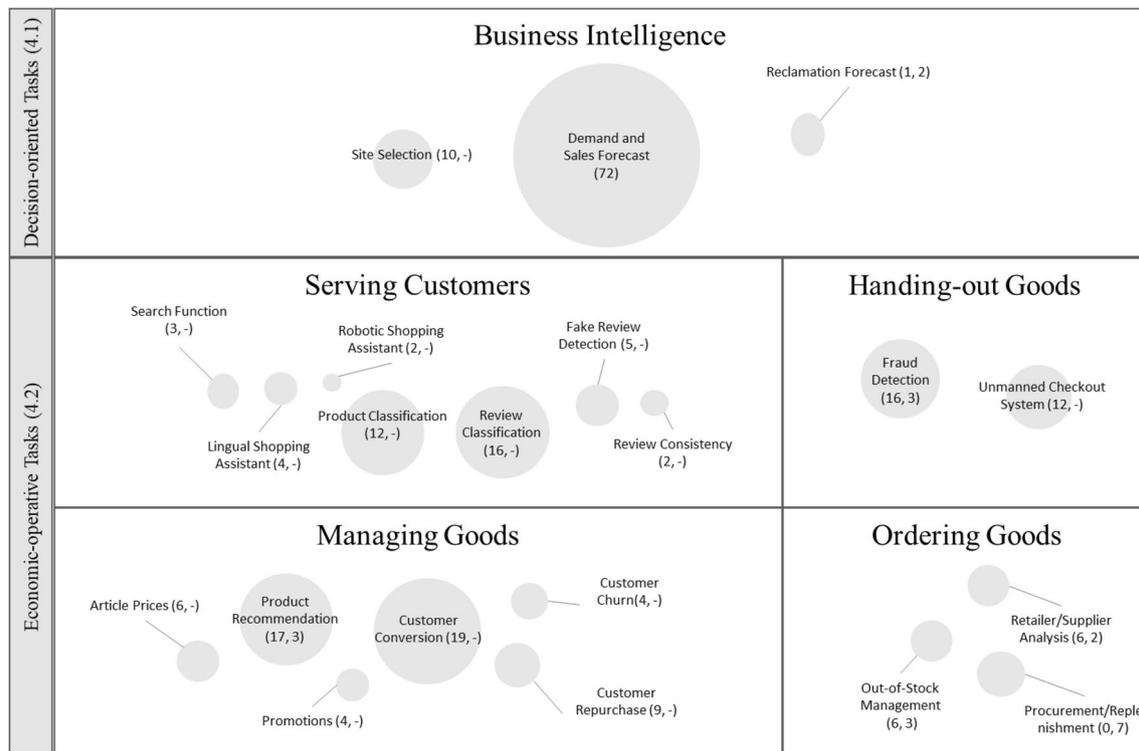


Fig. 4 Framework for Machine Learning Application in Retail [34]. The tuples show the results of the literature review (left) and the interviews (right)

nature of the interviews allowed us to check back certain scenarios and allowed the interviewees to elaborate on selected processes in detail. The processes investigated depended on the process area in focus for each interviewee. The interview guide was created based on our research question, the shell model as reference model, and the preliminary set of application areas identified in the literature review. The interviews were conducted via electronic communication media, such as Zoom and Microsoft Teams.

For the knowledge extraction, the interviews were anonymized and transcribed in order to perform a qualitative content analysis, which aimed at systematically analyzing the object of communication, regardless of its form [75]. Sentences were chosen as coding units, while a paragraph of an interview were regarded as a context unit [56]. Compound sentences and sentences assignable to more than one category were coded multiple times [56, 94]. The first is the smallest possible segment which can be coded, while the latter is the largest codable segment. The interview analysis was supported by MAXQDA [41]. We used our theoretically derived application areas as the initial categorization system and iteratively adjusted this scheme with our eight expert interviewees. In total, we identified 21 application areas for ML in (online and offline) retail, as shown in Fig. 4.

As we not only discussed the use of ML in retail at application area level but also at process level, we demonstrate the use of ML in two selected processes discussed with the domain experts. We followed the approach presented by Weber and Schütte [91]. Consequently, we depict the restocking process and the returns process in the Business Process Model and Notation (BPMN) [67] and indicate activities that can be supported or even replaced by the application of ML. We refer to Becker and Schütte [6] for the reference process activities and events for both processes.

Application Areas of Machine Learning in Retail

This section presents the application areas for ML, which resulted from our structured literature review and the analysis of the interviews with the domain experts. The assignments between business tasks of the retail shell model and ML application areas are illustrated in Fig. 4. The number of identified scientific papers is presented between brackets in Fig. 4 for each application area on the left, and the number of processes discussed in the interviews are presented on the right. Each application area is assigned to a task of the shell model. Each shell of the shell model for retail consists

of a series of tasks that form the components of the architecture. The master data article, supplier, customer, condition, site, and warehouse are the basis for technical tasks, peripheral control, and user access technology. We present each application area in general, and the results of the literature review are described specifically by outlining a sample of the literature review results. We also elaborate on the processes discussed during the expert interviews and the specific application of ML.

Decision-Oriented Tasks

The decision-oriented tasks (i.e., business intelligence) of the shell model provide aggregated information to the management of a retailer [78]. One very important business intelligence ML application area is the sales forecast, which is represented by 72 papers. The sales forecast provides management with information on future customer demands, and therefore management is able to plan operating activities accordingly.

Sales forecasting per article is one of the ML areas in retail that receives the most attention, because it is the basis for numerous advanced algorithms, such as price optimization [11] and promotion optimization [30]. Forecasting techniques are divided into four main groups: qualitative methods which are based on human judgment, time-series methods that use historical data, causal methods that use rule-based forecasts, and simulation methods that simulate the behavior of the customer [14]. If the focus is on-demand forecasting with ML methods, time-series methods are mostly used. The forecast for individual days has the difficulty that rare events are decisive in the prognosis, which are often hardly considered as outliers at weekly or monthly levels [31]. The demand forecast can be extended by features; for example, Verstraete et al. [86] created short-term and long-term forecast models that take weather data into account. Loureiro et al. [52] compared demand forecasts in the context of the fashion industry, where short product life cycles are evident.

The inverse of a sales forecast is the *forecast of returned goods*. For example, when it comes to product recalls, management must factor a high volume of returned products into operational activity planning. In addition, the returns process deals with goods that have been complained about. This can happen in case of wrong deliveries, faulty goods, or even defective goods [6]. Kumar et al. [44] used multiple ML algorithms for the *reclamation forecast*. This forecast is integrated into the supply chain design and planning problems to optimize supply chain tasks. The expert interviews confirmed that ML can be used to predict returns per customer. This involves using historical returns data to evaluate customer profiles and determine the probability of whether a return may occur. As a result, customers are assigned a

score accordingly and prioritized in future ordering processes (Interviewee 2). Furthermore, using ML, so-called returns clustering by the supplier can be performed. In this process, supplier returns are analyzed and evaluated accordingly within the procurement process to make a statement about the delivery reliability of a supplier (Interviewee 3).

A store is an organizational unit used to document the goods in stock for inventory management and to map the associated business processes, such as goods receipt, physical inventory, and goods issue [6]. Part of the process of opening a new store includes the choice of location (*site selection*), which is done by management, where ML algorithms can provide support. For example, Wang et al. [90] made a quantitative site selection, so that a very large number of possible store locations are reduced to a small number, which is then investigated further by qualified employees. Machado et al. [53] followed a different approach, in which a pre-selection had already been made and the remaining five stores were evaluated by a neural network. The site selection was grouped into the shell for master data with the goal of creating a master data record for the store. Another formulation of the site selection problem is when a mall is designed with a given set of sets. For example, Miao [59] used ML algorithms to find a reasonable layout for shopping malls where the relationships of stores to each other were taken into account.

Economic-Operative Tasks

The following subsections address the economic-operative tasks of the shell model, which are the primary activities of retailers. These tasks fulfill the bridging functions and resemble a retailer's value proposition [48]. We have identified ML application areas for the tasks of managing goods, serving customers, handing out goods, and ordering goods.

Managing Goods

A central aspect of the economic-operative tasks is the managing of goods, which is reflected by the large number of papers (59) on this area. The task of managing goods consists of the assortment policy, the conditions policy (purchasing price and selling price), the placement policy, and the promotion policy. In addition, managing goods also includes inventory management, which refers to the management and optimization of goods and storage respective to the inventory. Prices are an important driver that influence how often an article is sold and the resulting profit margin. This also influences inventory, since a higher price leads to lower sales, and vice versa, leading to either empty or full stock, due to varying demand or supply. Therefore retailers can implement *price optimization* so that the demand fits the supply. For example, Chandrashekhara et al. [11] determined

the best prices of smartphones based on different features of the smartphones.

Furthermore, *promotion optimization* can be used to specifically increase sales for a short period of time. The general promotional effect was forecast by Henzel et al. [30]. Retailers implement *product recommendation* systems to enhance the matching of supply- and demand-side participants, which we assigned to promotional tasks and thus the task of managing goods. These systems recommend supply-side products to customers based on previous purchases, ratings, or search behavior [37]. In contrast to the traditional recommendation that suggests popular products to customers, the algorithm of these researchers aims to match customers with appropriate niche products. Thus, online shops can more specifically target the long tail [57] and exploit niche markets to generate additional revenue. To determine the ranking of products that should be recommended to demand-side participants, Liang et al. [51] performed sentiment analyses based on online product reviews from Taobao.com. As personalized product recommendations may be biased by supply-side marketing endeavors, Wan et al. [89] proposed ML algorithms to allow for product recommendations in underrepresented market segments. The authors used two data sets from online shops selling apparel and electronics. Promotional efforts can be performed for a specific group of customers. To determine such a group of customers, ML algorithms are used, which we also categorized to the promotional task. By analyzing the expert interviews, we were able to confirm the use of ML for product recommendation. Three additional processes were identified that can be supported by the use of ML. In e-commerce, ML is used to predict which items a customer will put in the shopping cart. This results in customer-specific recommendations (Interviewee 5). To increase overall revenue, ML is used to initiate cross-selling. For this purpose, a supplementary suggestion is made to the current shopping cart, based on historical data (Interviewee 4). ML is also used in both, brick-and-mortar retail and e-commerce where products are selected for promotions. This is achieved by a sentiment analysis, in which social media data are evaluated to select promotional articles (Interviewee 5).

An important application area of ML in retailing is *customer churning classification* [49]. “Customer churn” describes the loss of customers [39] who generated sales in the past and will no longer generate sales, mostly because of an active decision to generate sales with a competitor. Clustering churners categorizes demand-side participants into two clusters: one with churning customers and the other one with non-churning customers, which results in a binary classification problem [39]. While possible churners can be addressed with special marketing campaigns, the quality of services can be targeted at non-churning customers. Keeping existing customers in the sales funnel is considered less

expensive than acquiring new customers [64, 79]. Customer satisfaction has a positive influence on customer repurchase behavior [36]. To implement a *customer repurchase* analysis, Kumar et al. [43] used a combination of ML techniques including both customer characteristics and e-commerce attributes. This application area allows retailers to predict whether a demand-side participant will rejoin the sales funnel, resulting in additional sales transactions and revenue.

Using an ML algorithm on e-commerce data Singh et al. [79] suggested detecting very loyal customers and providing them with higher-quality services to increase their satisfaction and to retain them in the sales funnel. Jheng et al. [36] mined the transaction logs of an enterprise resource planning system at an online shop to predict customer repurchase behavior. Moreover, Mahaboob et al. [54] develop an ML model to describe the moderating role of customer loyalty on customer retention. Predicting the probability of a lead to convert into a customer and the profitability of this new customer is crucial in retail [17]. The *customer conversion* analysis implemented by D’Haen et al. [17] applied ML algorithms to the sales data of a German business-to-business (B2B) e-commerce company and complementary data from the internet to predict the profitability of a lead. Knowing which leads will be profitable can help retailers to directly address these leads and increase their conversion probability. Niu et al. [65] applied multiple ML algorithms to predict the probability of a customer purchasing a product based on his or her (previous) search behavior. The computational model was applied to e-commerce data from Walmart. Thus, a retailer can predict a customer’s willingness to purchase and pay in an early stage of a sales funnel [9].

Serving Customers

The following application areas are concerned with tasks that serve customers. In brick-and-mortar stores, the products are made available using limited shelf space. The digital equivalent is online stores, which often provide a much larger assortment. Therefore, a product *search function* function is necessary to serve customers based on their demand, which is the reason we categorized the product search function as a business task. Khatwani et al. [38] developed a model for predicting customers’ individual information search preferences using questionnaire data.

To enable a good product search, products must be classified first. E-commerce *product classification* challenging, due to the large scale and complexity of product information and categories. Yu et al. [96] combined multiple ML algorithms to propose an e-commerce text classification. As reviews are the online counterpart to the recommendations of an employee, we categorized review analysis with ML algorithms to the economic-operative task of serving customers. To increase a retailer’s role as a trustee, retailers

often implement systems to review the transaction partner and the products sold. Suppliers and customers of a retailer are aware of reviews, as these often directly impact their business [32]. Vinodhini and Chandrasekaran [87] implemented *review classification* based on customers' opinions (opinion mining). They used publicly available customer reviews and classified these into positive and negative reviews, so that supply- and demand-side participants can preselect reviews possibly more relevant for them.

Opinion spammers and fake reviews exploit customer trust and harm the reputation of e-commerce retailers as a trustee by posting false or deceptive reviews [97]. These reviews are difficult to detect because of complex interactions between several user characteristics [45]. To enable *fake review detection* Hussain et al. [32] implemented a behavioral method that utilizes 13 different behavioral features to calculate a review spam score for each reviewer. As review rankings are an important indicator of the relevance of a reviews to demand-side participants, the *review consistency* between the review ranking and the review summary is crucial. To verify this consistency, Zhang et al. [97] proposed an ML approach based on e-commerce data. This approach enables demand-side participants to identify relevant reviews based on the review rankings. As reviews are often created by other customers, shopping assistants are more responsive to the needs of a special customer and therefore the interaction with an employee is digitally and automatically mapped.

Bertacchini et al. [7] developed a *robotic shopping assistant*. These robotic assistants can promote sales in general and guide sales decisions. A shopping assistant can also be implemented on a customer's smartphone as a companion app [94]. Many online shops offer a 24-hour service using *lingual shopping assistants*. This service is very expensive when done manually. Chatbots can be used as a solution for automatic online shopping, but the chatbot has to be able to give accurate and quick answers. For example, Nursetyo et al. [66] proposed an intelligent chatbot system that can be used as an e-commerce assistant and supports the process of serving customers.

Handing out Goods

The business task of handing out goods includes the transaction in which the product becomes the customer's property. This process can be improved or automated with ML algorithms. Kourouthanassis and Roussos [40] developed an *unmanned and automated checkout system* with a smart shopping cart. Shopping carts are unusual in the fashion industry segment, but, for example, Hauser et al. [28] developed an ML algorithm to determine whether a product had passed the store exit or was only registered because it was placed near the gate-mounted antennas. Suponenkovs et al.

[80] developed a system for *automatic invoice recognition* by analyzing the pixels with respect to their relevance and texture analysis. As an offline scenario, payment for fruit or vegetables in retail stores normally requires them to be manually identified. Rojas et al. [72] and Femling et al. [20] presented image classification methods with the goal of speeding up the checkout process in stores.

Since incorrect operations or failures by employees at points-of-sale (POS) account for 24% of inventory differences in the retail industry [18], ML algorithms are used to reduce these differences. *Fraud detection* is assigned to the business task of handing out goods, since problems with the transaction in which the product becomes the customer's property should be avoided. For interim transaction fraud detection at checkouts, Trinh et al. [82] developed a method that uses image recognition to determine the hand movements of cashiers. It should be recognized by the algorithm whether all articles that are handed out to the customer are also recorded in a sales process, that is, on the receipt. Fraud detection in retrospect must be dealt with in a human resources management process. Pehlivanli et al. [69] pursued a different approach by analyzing transaction data. ML algorithms were trained to classify fraud and non-fraud on the basis of indicators such as profitability, stock turnover, stock cost, and shelf life. This process was supported by our findings from the interview analysis. According to the interviewees, ML can be used to detect anomalies in stock transfers by evaluating order times and quantities, as well as a store's inventory situation, and then making a prediction regarding missing delivery quantities (Interviewee 4). Moreover, ML is implemented to perform an evaluation regarding the reservation of products and their actual collection. This can minimize the risk of blocking inventory (Interviewee 6). Furthermore, ML is used to classify the credit scores of customers in e-commerce. Kulkarni and Dhage [42] used ML and data mining techniques to integrate information obtained from social media to protect retailers from payment defaults. In addition, the interviewees provided insight into the audit process, where ML is used to support retailers in fraud detection. Here, ML is used to check invoices for consistency and correctness and to detect any anomalies in invoices (Interviewee 1).

Ordering Goods

We identified out-of-stock management and supplier analysis as application areas of ML for the ordering of goods tasks. "Out-of-stock" describes shelf space in brick-and-mortar stores that is no longer filled by the intended article and is empty. Customer demands can no longer be fulfilled. As a *detected out-of-stock* triggers the replenishment process, this application area is assigned to the business task of making goods available. For example, Paolanti et al. [68]

implemented a mobile robot using visual and textual analysis to facilitate the automatic detection of shelf out-of-stock situations. The use of ML in out-of-stock management can also be seen in three processes that we were able to extract from the expert interviews. For example, ML can be used to perform zero-stock forecasting, so that goods can be moved and redistributed between various stores depending on actual inventory levels. In distribution, ML can be used to predict and plan inventories. A prognosis of delivery delays is created and then used to predict when an item will be out of stock (Interviewee 6). Another Interviewee explained the usage of ML in out-of-stock management where a forecast of spoilage of goods takes place. For this purpose, historical data are used to minimize depreciation (Interviewee 8). A problem for the retailer is when item inventories are present for more than one day with zero inventory according to the system until the next delivery. These items are out-of-stock items in the store and can be predicted using ML by evaluating the order time and order quantity, to create a risk representation for specific items (Interviewee 1).

Once the replenishment process is triggered, a supplier must be selected. Kuo et al. [46] considered the aspect of the sustainability of suppliers in their *supplier selection* optimization. ML algorithms are used to resolve supplier selection problems regarding operational performance and environmental issues in terms of the criteria of corporate social responsibility, service, cost, environment, quality, and delivery. This was also confirmed by our interview analysis, in which we identified a process of ML use in the course of inventory management. For inventory management, ML can be used to analyze historical point-of-sale data and past inventories to identify anomalies in the movement of goods (Interviewee 4). In inventory planning, ML can be used to forecast potential delivery delays, using historical deliveries for delivery time and amount of goods. This helps in the planning and procurement of goods (Interviewee 3). Another possible application of ML in the context of retailer/supplier analysis is the evaluation of parcel service providers with regard to shipping day, delivery day, and volume, in order to be able to create a forecast for delivery reliability (Interviewee 6). In addition, ML can be used to make a statement about supplier delivery reliability by evaluating the delivery day according to the order and the actual delivery day (Interviewee 4).

In addition to the results of the structured literature research, we were able to identify a further application area for the use of ML in the context of economic-operative tasks, one that is primarily concerned with the procurement and replenishment of goods. Procurement and replenishment is strongly interlinked with the other application areas but does not uniquely belong to one area, thus we chose to see it as a separate application area. It mainly focuses on the timed ordering of goods which can be combined with out-of-stock

management to show only one interlinking. Thus, we identified a total of seven different processes which can be supported by the use of ML. The task of procuring goods deals with the timely reordering of articles in order to avoid a situation of empty shelves. On the other hand, however, only as much should be reordered as can actually be sold, so that there is no overfilling of the shelves, leading to associated higher storage costs. For this purpose, current inventories as well as historical goods movements must be taken into account. In the fashion sector, ML can be used to create seasonal forecasts and thus minimize depreciation. Some of the items will then be carried over to the next season and possibly priced differently (Interviewee 3). Other possible uses of ML in this area are automated stock transfers, for which a forecast of future zero stock must first be created, after which the articles can be distributed evenly across all stores. Another use of ML is to determine a sales forecast per price, first in order to avoid zero stock at a low price or slow-moving items at a high price (Interviewee 7), and second, to create an action forecast based on point-of-sale data to ultimately perform targeted merchandise management (Interviewee 4). In order to examine stocks in more detail, the use of ML can also help. Anomalies within the inventory can be identified in order that the retailer can be informed in which product areas action is required (Interviewee 3). Another option is to analyze the capacity of a company's employees. This can be particularly helpful in workforce planning to determine how many employees can be deployed within a store or warehouse (Interviewee 7). In procurement, ML can also be used to optimize resources. This concerns both own and external capacities. Above all, in production the resources can be optimized in such a way that machines are ideally used up, by, for example, products being pre-produced (Interviewee 4). The last process we were able to extract from the analysis of the interviews concerns contingency. Here, ML can be used at several points within the process, for example, by performing buyer prioritization. This, in turn, is used to allocate inventory to appropriate customers (Interviewee 7).

Demonstration of Machine Learning Application in Selected Retail Processes

This section presents two selected processes based on the reference process description by Becker and Schütte [6], and the use of ML is demonstrated for the corresponding activities extracted from the interviews. We selected a return process and a stock relocation process to demonstrate the application of ML in retail.

The first process deals with *customer returns* as illustrated in Fig. 5. Normally, the return process begins with a call from the recipient to notify distribution about an incorrect,

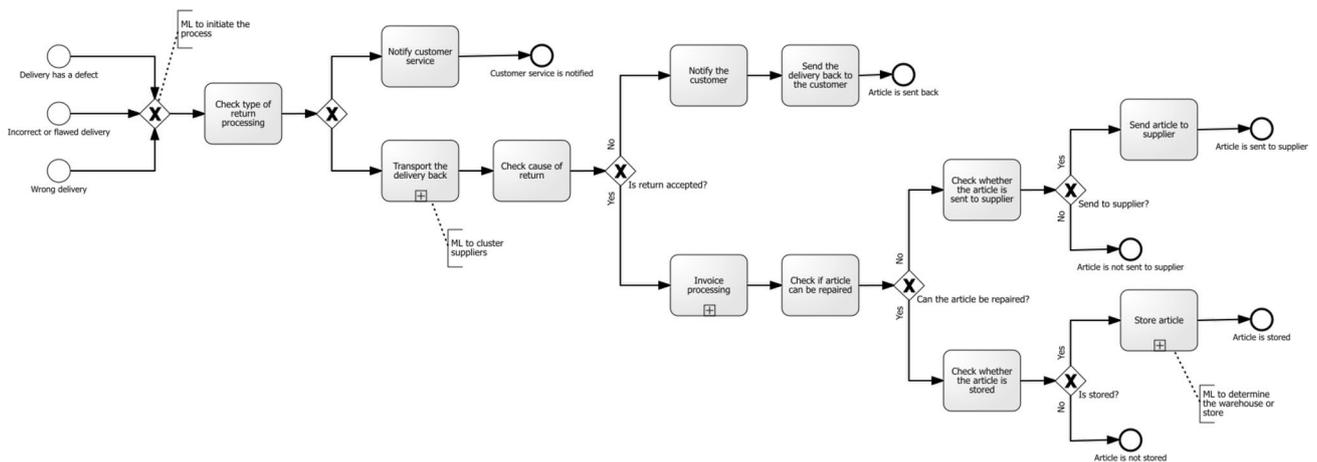


Fig. 5 The Return Process in BPMN

defective, or wrong delivery. At this point, ML can be used to calculate the likelihood of the recipient initiating this call, based on the customer’s historical returns (Interviewee 3). Next, distribution decides how the return will be managed. If it is not possible to fix the problem with the help of customer service, the article of the delivery have to be transported back. Based on an analysis of different suppliers, ML can create a score to cluster suppliers by returns and thus integrate them into the return process (Interviewee 6). The recipient can bring the article back themselves, or the article can be picked up. If the return is then declined, the delivery is sent back to the customer, who gets a notification. In the case of accepting the return, the invoicing process is started, and it is checked whether the article can be repaired or not. If the article cannot be repaired, the article will be sent back to the supplier, or it will be disposed of. If the article can be repaired, it will be fixed and stored in a warehouse. At

this point, ML can be used to determine the warehouse or store where it will be sent by creating a zero-stock forecast (Interviewee 2).

The second process deals with the *stock transfer of goods* as illustrated in Fig. 6. This process can be initiated after analyzing the current stock and calculating the probability that an article will be out of stock (Interviewee 2). This means an article has to be relocated to a store where it may be out of stock in the future. If an article has to be relocated, then it has to be decided whether it is to be relocated to a storage or organizational unit. In the latter case, the relocation method has to be chosen, and often a relocation with invoicing is chosen. The stock transfer process can be automated by creating a zero-stock forecast (Interviewee 2). Next, the supplier and recipient have to be determined, which then leads to several processes involving the identification of these. The supplier could be a storage unit,

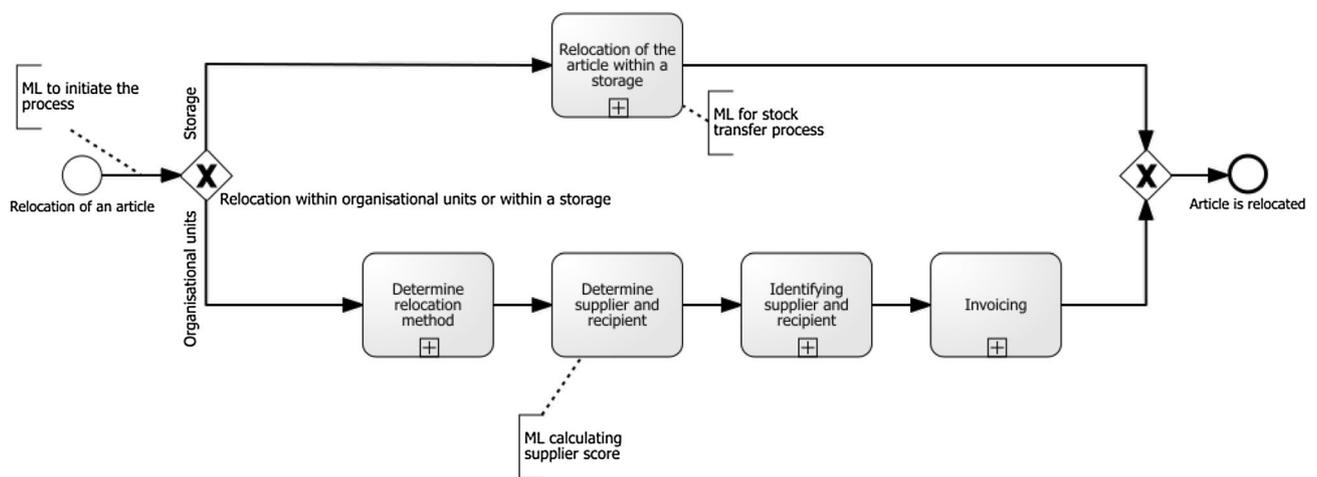


Fig. 6 The Stock Relocation Process in BPMN

customer, or store, while the recipient could also be a storage unit, customer, or store. This involves the use of ML when choosing a customer as supplier. The customer can be given a score based on an evaluation of his or her historical data and, based on this score, the supplier is then determined (Interviewee 4). After this, the invoicing is done. Lastly, the article is relocated, both in the case of storage and organizational units; this is the end of the process.

Discussion

Summary of Key Contributions

Our structured literature review, as well as the analysis of the expert interviews and the subsequent structuring of the ML application areas along the retail-related tasks depicted in the shell model (Fig. 1), led to the development of a framework for determining possible ML applications in retail (Fig. 4). Our structured literature review and interview analysis on ML application areas in retail resulted in three major findings.

First, a closer analysis of the framework reveals that there is a trend to use ML to support decision-oriented and economic-operative tasks in retail [77]. Other crucial tasks in retail, such as master data management as well as technical and administrative tasks, are either not covered by ML applications or underrepresented in our literature sample. On the one hand, the support and automation of economically valuable tasks seem reasonable, as they represent the main bridging functions and the major value proposition, can be used to differentiate a company from its competitors, and are a major source of revenue for retailers [77]. On the other hand, master data and technological facilities consist of the necessary data required for ML algorithms [92]. If these data are inconsistent or not even available, ML models provide false results or can even not be trained.

Second, our integrative approach for the review in brick-and-mortar retail and e-commerce is relatively balanced regarding the application environment, with 107 papers focusing on offline retail and 118 papers focusing on online retail. The papers also reveal that ML applications make use of the peculiarities of each environment. While the selection of an optimal site for a brick-and-mortar store or the support of unmanned checkout systems with ML concern offline retail, ML-based analyses of customers and product reviews are only possible in e-commerce scenarios in which these data exist [32]. As mentioned earlier, e-commerce involves much data regarding its customers and unstructured data regarding products. This process will possibly be driven in future by VR (virtual reality) shopping. Therefore, in addition to click and mouse movements in e-commerce, also hand, eye, head, and leg movements could be tracked [95].

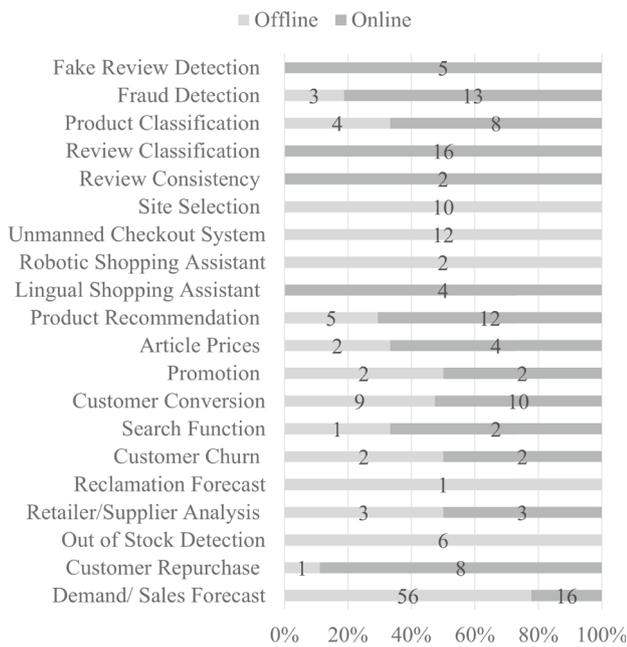
We assume that the existing application areas presented in this work, especially from e-commerce, can be adapted to VR shopping, and future research may find new application areas, especially for VR shopping [95].

However, we also identified application areas that can be applied to ML in both environments, such as customer conversion, promotion optimization, and product classification (Fig. 7a). This trend is also supported by the fact that brick-and-mortar retailers are increasingly integrating technology within their stores to bring in the online convenience and experience [16, 74, 94]. Product placement is a task necessary to optimize revenues in brick-and-mortar stores, but it can also be used in electronic shops to optimize the placement of a product on a website [33]. As these ML models can be used in both environments, the effort of implementation can be balanced by the dual application. Thus, these applications are also useful for retail using a multi- or omnichannel approach [84].

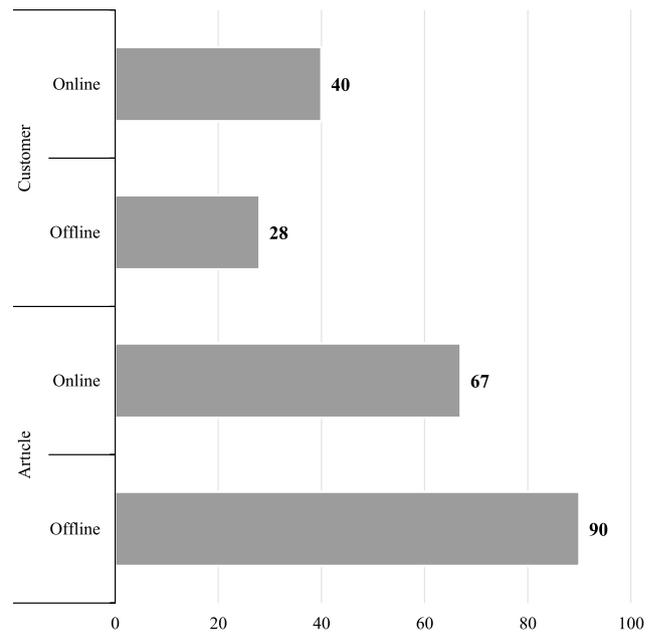
Third, the analysis of the application areas for the object of interest reveals a tendency towards the customer-centered application of ML in e-commerce (40 papers), while the focus in the offline environment is on a single or set of articles (90 papers) (Fig. 7b). Our analysis also confirms the developments described by Grewal et al. [26], towards a more customer-centered focus in brick-and-mortar retail (28 papers). However, we propose to focus even more on the customer for proper ML implementation and address the customer experience for value generation [24, 85]. To select an optimal site, it is necessary to identify a location with an adequate number of potential customers and to select the right articles for the needs of a specific customer milieu [90].

Limitations and Outlook

This research has certain limitations, which at the same time open interesting avenues for future research. As previously mentioned, the proposed framework for ML application areas in retail does not cover master data management or technical and administrative tasks. These should be investigated by future research. We concentrated on retail as the domain of application for ML. Retail can be distinguished from wholesale by small units, trading net prices, and operating using a business-to-customer business model [48]. Although our framework for ML application was developed specifically for retail, we propose that many of the ML applications can also be used in a wholesale context, as the tasks are broadly similar (i.e., bridging functions) when comparing retail and wholesale architectures [77]. We intentionally focused our literature research on ML applications in scientific publications to ensure the maximum level of quality and comprehensibility of our research. However, we are aware that implementations in practice can be ahead of scientific publications [19]. Hence, we conducted interviews with



(a) Offline vs. Online ML Application Areas



(b) Article- vs. Customer-centric ML Applica-

Fig. 7 Application area analysis based on literature analysis [34]

eight domain experts, knowledgeable about retail and enterprise systems. In addition, we aim to integrate practitioner sources, such as company white papers and publications by retail interest groups, into future research. In preparation for the expert interviews, we used the preliminary application areas resulting from our literature review and the tasks summarized in the shell model [77]. This might have led to a potential bias in the selection of the interviewees, although we had the impression that the interviewees reported their experience of the application of ML in retail depending on their domain of interest. For the literature review, we opted for a broad research scope, reflected in our search string and the databases queried. However, we did not provide a more detailed analysis of the algorithms and data used in each ML application. An important avenue for future research would be to detail this broad overview and to provide evidence of proper algorithms to be implemented within each application area. As we focused on identifying ML applications in retail, we used a task-oriented architecture developed by Schütte [76] that focuses on the business layer as a starting point. However, we are aware that there are other reference architectures covering additional application and technical architecture layers in general and integrating ML in particular [3]. Although these architectures include additional layers, they are less detailed with regard to retail-specific tasks. It might, therefore, be worthwhile for future research to extend the existing task-oriented architectures [6, 77] with technical layers integrating ML. Alternatively, an architecture

may profit from providing interfaces to plug in additional services (e.g., ML services), making the architecture more flexible [2].

Conclusion

We have identified 21 application areas of ML in offline and online retail, based on a thorough analysis of the current body of literature and analyzing eight expert interviews. The application areas identified cover decision-oriented (business intelligence) and economic-operative tasks (managing goods, serving customers, handing out goods, ordering goods). Based on our analysis, ML can be implemented to support and automatize structured and unstructured tasks in retail. In addition, current research is equally concerned with the application of ML in offline and online retail. For e-commerce, we identified a tendency for the customer-centric use of ML, while in the brick-and-mortar context, the retail article is more often the object of interest. The contribution of our paper for practitioners and researchers is a general overview of current research on ML applications in retail. For practitioners, the framework for ML application can be used by retailers to determine which tasks ML can support in a company. Thus, the framework can be used as an indicator for the usefulness of an ML application and to determine whether an ML implementation project is feasible and makes sense from an economic perspective. The

algorithms developed and applied in the identified papers can also serve as a starting point for the practical consideration of implementation. In addition, the interview analysis provided insights into where ML is actually used in retail and can be applied to increase the revenue and efficiency of companies, as well as to improve the processes involved in economic-operative and decision-oriented tasks. For researchers, this work provides a retail-specific framework for ML application to foster the development of more holistic and integrated ML models (e.g., promotion-sensitive resource optimization).

Acknowledgements We would like to thank the reviewers of the 24th International Conference on Enterprise Information Systems and the discussants during the conference for their fruitful comments. These comments helped us to improve the line of argumentation of the paper. We acknowledge support by the Open Access Publication Fund of the University of Duisburg-Essen.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability The data that support the findings of this research article contain corporate information and restrictions apply to the availability of these data. Anonymized summaries of the interview data are, however, available from the authors upon request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. ABIresearch: Mobile Robotic Systems: Market Update 2019. Accessed: 15 October 2021. <https://www.abiresearch.com/market-research/product/1034058-mobile-robotic-systems-market-updat-e/%0A>
2. Aulkemeier F, Paramartha MA, Iacob ME, van Hillegersberg J. A pluggable service platform architecture for e-commerce. *IseB*. 2016;14(3):469–89.
3. Aulkemeier F, Schramm M, Iacob ME, van Hillegersberg J. A Service-Oriented E-Commerce Reference Architecture. *J Theor Appl Electron Commer Res*. 2016;11:26–45.
4. Bandara W, Furtmueller E, Gorbacheva E, Miskon S, Beekhuizen J. Achieving rigor in literature reviews: Insights from qualitative data analysis and tool-support. *Commun Assoc Inf Syst*. 2015;37:154–204.
5. Barth, K., Hartmann, M., Schröder, H.: Betriebswirtschaftslehre des Handels, vol. 7., überar. Springer Gabler, Wiesbaden 2015
6. Becker J, Schütte R. Handelsinformationssysteme : domänenorientierte Einführung in die Wirtschaftsinformatik, 2. vollst ed. Frankfurt a. M.: Redline Wirtschaft; 2004.
7. Bertacchini F, Bilotta E, Pantano P. Shopping with a robotic companion. *Comput Hum Behav*. 2017;77:382–95.
8. Biem, A.: A model selection criterion for classification: Application to hmm topology optimization. *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR 2003)*, 2003;104–108
9. Blank S, Dorf B. The startup owner's manual: The step-by-step guide for building a great company. Hoboken, New Jersey: John Wiley & Sons; 2012.
10. Bousqaoui, H., Achchab, S., Tikito, K.: Machine learning applications in supply chains: An emphasis on neural network applications. In: *Proceedings of 2017 International Conference of Cloud Computing Technologies and Applications, CloudTech 2017*, vol. 2018, pp. 1–7. Institute of Electrical and Electronics Engineers Inc. 2018. <https://doi.org/10.1109/CloudTech.2017.8284722>
11. Chandrashekhara, K., Thungamani, M., Babu, C.G., Manjunath, T.: Smartphone price prediction in retail industry using machine learning techniques. In: *Emerging Research in Electronics, Computer Science and Technology*, pp. 363–373. Springer 2019
12. Chen, H., Li, C., Li, X., Rahaman, M.M., Hu, W., Li, Y., Liu, W., Sun, C., Sun, H., Huang, X., Grzegorzec, M.: Il-mcam: An interactive learning and multi-channel attention mechanism-based weakly supervised colorectal histopathology image classification app. *Computers in Biology and Medicine* 2022;143:1–23 . <https://doi.org/10.1016/j.compbiomed.2022.105265>
13. Chen H, Li C, Wang G, Li X, Mamunur Rahaman M, Sun H, Hu W, Li Y, Liu W, Sun C, Ai S, Grzegorzec M. GasHis-Transformer: A multi-scale visual transformer approach for gastric histopathological image detection. *Pattern Recogn*. 2022;130:1–35. <https://doi.org/10.1016/j.patcog.2022.108827>.
14. Chopra S, Meindl P, Kalra DV. Supply chain management: strategy, planning, and operation, vol. 232. MA: Pearson Boston; 2013.
15. Cooper HM. Organizing Knowledge Synthesis: A Taxonomy of Literature Reviews. *Knowledge in Society*. 1988;1(1):104–26. <https://doi.org/10.1007/BF03177550>.
16. Dekimpe MG, Geyskens I, Gielens K. Using technology to bring online convenience to offline shopping. *Mark Lett*. 2020;31(1):25–9. <https://doi.org/10.1007/s11002-019-09508-5>.
17. D'Haen J, Van Den Poel D, Thorleuchter D. Predicting customer profitability during acquisition: Finding the optimal combination of data source and data mining technique. *Expert Syst Appl*. 2012;40(6):2007–12. <https://doi.org/10.1016/j.eswa.2012.10.023>.
18. EHI: Inventurdifferenzen 2020. Accessed: 12 December 2020. <https://www.ehi.org/de/studien/inventurdifferenzen-2020/>
19. Eyes, W.: Welcome to the visual intelligence era 2021. Accessed: 12 December 2020. <https://wideeyes.ai/>
20. Femling, F., Olsson, A., Alonso-Fernandez, F.: Fruit and vegetable identification using machine learning for retail applications. In: *2018 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, pp. 9–15. IEEE 2018
21. Ferstl, O.K., Sinz, E.J.: *Grundlagen der Wirtschaftsinformatik*, vol. 7., aktual. Oldenbourg Verlag, München 2013
22. Feurer, M., Klein, A., Eggensperger, K., Springenberg, J.T., Blum, M., Hutter, F.: *The springer series on challenges in machine learning*. Nips 2019;113–134
23. Frey CB, Osborne MA. THE FUTURE OF EMPLOYMENT: HOW SUSCEPTIBLE ARE JOBS TO COMPUTERISATION? *Technol Forecast Soc Chang*. 2017;114:254–80. <https://doi.org/10.1257/aer.103.5.1553>.

24. Gilly MC, Wolfinbarger M. A comparison of consumer experiences with online and offline shopping. *Consumption Markets & Culture*. 2000;4(2):187–205.
25. Goos M, Manning A, Salomons A. Job Polarization in Europe. *American Economic Review*. 2009;99(2):58–63. <https://doi.org/10.1257/aer.99.2.58>
26. Grewal D, Roggeveen AL, Nordfält J. The Future of Retailing. *J Retail*. 2017;93(1):1–6.
27. Große Holtforth, D.: Machine Learning im E-commerce 2018. Accessed: 16 August 2021. <https://ecommerceinstitut.de/machine-learning-im-e-commerce/>
28. Hauser M, Günther SA, Flath CM, Thiesse F. Towards digital transformation in fashion retailing: A design-oriented research study of automated checkout systems. *Bus Inf Syst Eng*. 2019;61(1):51–66.
29. Heinrich L, Stelzer D. *Informationsmanagement: Grundlagen, Aufgaben*. München: Methoden. Oldenbourg Wissenschaftsverlag; 2009.
30. Henzel, J., Sikora, M.: Gradient boosting and deep learning models approach to forecasting promotions efficiency in fmcg retail. In: *International Conference on Artificial Intelligence and Soft Computing*, pp. 336–345. Springer 2020
31. Huber, J., Stuckenschmidt, H.: Daily retail demand forecasting using machine learning with emphasis on calendric special days. *International Journal of Forecasting* 2020
32. Hussain N, Turab Mirza H, Hussain I, Iqbal F, Memon I. Spam Review Detection Using the Linguistic and Spammer Behavioral Methods. *IEEE Access*. 2020;8:53801–16. <https://doi.org/10.1109/ACCESS.2020.2979226>.
33. Hütsch, M., Wulfert, T.: Implementation of Placement Optimization in Electronic Commerce: Towards a Data Generation Method to Increase Forecast Elasticity. In: *Proceedings of the 25th Pacific Asia Conference on Information Systems (PACIS2021)*, pp. 1–12. Dubai 2021
34. Hütsch, M., Wulfert, T.: A Structured Literature Review on the Application of Machine Learning in Retail. In: *Proceedings of the 24th International Conference on Enterprise Information Systems (ICEIS2022) - Volume 1*, pp. 332–343. Online Event 2022. <https://doi.org/10.5220/0011043200003179>
35. Iivari, J., Rotvit Perlt Hansen, M., Haj-Bolouri, A.: A proposal for minimum reusability evaluation of design principles. *European Journal of Information Systems* 30(3), 1–18 (2021). <https://doi.org/10.1080/0960085X.2020.1793697>
36. Jheng, Y.R., Luo, Z.C.: Customer Retention Prediction with CNN. In: *Proceedings of the 4th International Conference on Data Mining and Big Data (DMBD 2019)*, vol. 1071, pp. 104–113. Springer, Chiang Mai, Thailand (2019)
37. Katarya, R., Verma, O.P.: Effectual recommendations using artificial algae algorithm and fuzzy c-mean. *Swarm and Evolutionary Computation* 36(June 2016), 52–61 (2017). <https://doi.org/10.1016/j.swevo.2017.04.004>
38. Khatwani, G., Srivastava, P.R.: Real-time prediction of information search channel using data mining techniques. In: *Proceedings of the 2015 International Conference on Green Computing and Internet of Things, ICGCIoT 2015*, pp. 924–929 (2016). <https://doi.org/10.1109/ICGCIoT.2015.7380595>
39. Kim K, Lee J. Sequential manifold learning for efficient churn prediction. *Expert Syst Appl*. 2012;39(18):13328–37.
40. Kourouthanassis P, Roussos G. Developing consumer-friendly pervasive retail systems. *IEEE Pervasive Comput*. 2003;2(2):32–9.
41. Krippendorff K. *Content Analysis: An introduction to its methodology*. 3rd ed. Los Angeles, USA: SAGE; 2013.
42. Kulkarni SV, Dhage SN. Advanced credit score calculation using social media and machine learning. *Journal of Intelligent and Fuzzy Systems*. 2019;36(3):2373–80. <https://doi.org/10.3233/JIFS-169948>.
43. Kumar A, Kabra G, Mussada EK, Dash MK, Rana PS. Combined artificial bee colony algorithm and machine learning techniques for prediction of online consumer repurchase intention. *Neural Comput Appl*. 2019;31:877–90.
44. Kumar DT, Soleimani H, Kannan G. Forecasting return products in an integrated forward/reverse supply chain utilizing an anfis. *Int J Appl Math Comput Sci*. 2014;24(3):669–82.
45. Kumar N, Venugopal D, Qiu L, Kumar S. Detecting Review Manipulation on Online Platforms with Hierarchical Supervised Learning. *J Manag Inf Syst*. 2018;35(1):350–80.
46. Kuo RJ, Wang YC, Tien FC. Integration of artificial neural network and mada methods for green supplier selection. *J Clean Prod*. 2010;18(12):1161–70.
47. Laudon, K.C., Traver, C.G.: *E-Commerce 2018: business. technology. society*, vol. 15. Pearson Education, Boston (2020)
48. Levy M, Weitz BA, Grewal D. *Retailing management*, tenth. edit. New York, NY: McGraw-Hill Education; 2019.
49. Li, X., Li, Z.: A hybrid prediction model for e-commerce customer churn based on logistic regression and extreme gradient boosting algorithm. *Ingenierie des Systemes d'Information* 24(5), 525–530 (2019). <https://doi.org/10.18280/isi.240510>
50. Liang, J., Ramanauskas, B., Kurenkov, A.: Job Loss Due To AI - How Bad Is It Going To Be? (2019) Accessed: 15 July 2021. <https://www.skynettoday.com/editorials/ai-automation-job-loss#fn:oxford3>
51. Liang, R., Wang, J.q.: A Linguistic Intuitionistic Cloud Decision Support Model with Sentiment Analysis for Product Selection in E-commerce. *International Journal of Fuzzy Systems* 21(3), 963–977 (2019)
52. Loureiro AL, Miguéis VL, da Silva LF. Exploring the use of deep neural networks for sales forecasting in fashion retail. *Decis Support Syst*. 2018;114:81–93.
53. Machado MAS, Gomes LFAM, Santos DJ, Caldeira AM. Using a bipolar choquet neural network to locate a retail store. *Procedia Computer Science*. 2015;55:741–7.
54. Mahaboob Basha, A.M., Rajaiah, M., Penchalaiah, P., Raja Kamal, C.H., Niranjana Rao, B.: Machine learning-structural equation modeling algorithm: The moderating role of loyalty on customer retention towards online shopping. *International Journal of Emerging Trends in Engineering Research* 8(5), 1578–1585 (2020). <https://doi.org/10.30534/ijeter/2020/17852020>
55. Manyika, J., Chui, M., Miremadi, M., Bughin, J., George, K., Willmot, P., Dewhurst, M.: *A FUTURE THAT WORKS: AUTOMATION, EMPLOYMENT, AND PRODUCTIVITY* (2017). Accessed: 15 August 2021. <https://www.mckinsey.com/~media/mckinsey/featuredinsights/DigitalDisruption/Harnessingautomationforafuturethatworks/MGI-A-future-that-works-Executive-summary.ashx>
56. Mayring, P.: *Qualitative Content Analysis: Theoretical Foundation, Basic Procedures and Software Solution*. Beltz, Klagenfurt (2014). <https://doi.org/10.4135/9781446282243.n12>
57. McAfee, A., Brynjolfsson, E.: *Machine, Platform, Crowd: Harnessing Our Digital Future*. W W Norton & Co Inc, New York City LB - McAfee2017 (2017)
58. Merton RK, Kendall PL. The Focused Interview. *Am J Sociol*. 1946;51(6):541–57.
59. Miao, Y.: A machine-learning based store layout strategy in shopping mall. In: *International Conference on Machine Learning and Big Data Analytics for IoT Security and Privacy*, pp. 170–176. Springer (2020)
60. Misoch, S.: *Qualitative Interviews*. DE GRUYTER, Berlin and München (2015). <https://doi.org/10.1515/9783110354614>

61. Mitchell TM, et al. Machine learning. New York, New York: McGraw-Hill; 1997.
62. Nguyen G, Dlugolinsky S, Bobák M, Tran V, López García Á, Heredia I, Malík P, Hluchý L. Machine learning and deep learning frameworks and libraries for large-scale data mining: a survey. *Artif Intell Rev*. 2019;52:77–124.
63. Nicola M, Alsafi Z, Sohrabi C, Kerwan A, Al-Jabir A, Iosifidis C, Agha M, Agha R. The socio-economic implications of the coronavirus pandemic (COVID-19): A review. *Int J Surg*. 2020;78(1):185–93.
64. Nikulin V. Prediction of the shoppers loyalty with aggregated data streams. *Journal of Artificial Intelligence and Soft Computing Research*. 2016;6(2):69–79. <https://doi.org/10.1515/jaiscr-2016-0007>.
65. Niu, X., Li, C., Yu, X.: Predictive analytics of E-commerce search behavior for conversion. In: Proceedings of the America's Conference on Information Systems (AMCIS 2017), pp. 1–10 (2017)
66. Nursetyo, A., Subhiyakto, E.R., et al.: Smart chatbot system for e-commerce assistance based on aiml. In: 2018 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), pp. 641–645. IEEE (2018)
67. Object Management Group: About the Business Process Model And Notation Specification Version 2.0 (2011). Accessed: 13 March 2020. <https://www.omg.org/spec/BPMN/2.0/About-BPMN/>
68. Paolanti, M., Sturari, M., Mancini, A., Zingaretti, P., Frontoni, E.: Mobile robot for retail surveying and inventory using visual and textual analysis of monocular pictures based on deep learning. In: 2017 European Conference on Mobile Robots (ECMR), pp. 1–6. IEEE (2017)
69. Pehlivanli D, Eken S, Ayan E. Detection of fraud risks in retailing sector using mlp and svm techniques. *TURKISH JOURNAL OF ELECTRICAL ENGINEERING & COMPUTER SCIENCES*. 2019;27(5):3633–47.
70. Rahaman MM, Li C, Yao Y, Kulwa F, Rahman MA, Wang Q, Qi S, Kong F, Zhu X, Zhao X. Identification of COVID-19 samples from chest X-Ray images using deep learning: A comparison of transfer learning approaches. *J Xray Sci Technol*. 2020;28(5):821–39. <https://doi.org/10.3233/XST-200715>.
71. Randolph JJ. A guide to writing the dissertation literature review. *Pract Assess Res Eval*. 2009;14(13):1–13.
72. Rojas-Aranda, J.L., Nunez-Varela, J.L., Cuevas-Tello, J.C., Rangel-Ramirez, G.: Fruit classification for retail stores using deep learning. In: Mexican Conference on Pattern Recognition, pp. 3–13. Springer (2020)
73. Rudolph T, Nagengast L, Melanie B, Bouteiller D. Die Nutzung mobiler Shopping Apps im Kaufprozess. *Mark Rev St Gallen*. 2015;3(32):42–9.
74. Scarpi D, Pizzi G, Visentin M. Shopping for fun or shopping to buy: Is it different online and offline? *J Retail Consum Serv*. 2014;21(3):258–67.
75. Schreier, M.: Qualitative content analysis. In: U. Flick (ed.) *The Sage handbook of qualitative data analysis*, pp. 170–183. SAGE, Los Angeles, Calif. (2014). <https://doi.org/10.4135/9781446282243.n12>
76. Schütte R. Modellierung von handelsinformationssystemen. Münster: Kumulative Habilitationsschrift. Westfälische Wilhelms-Universität Münster; 2011.
77. Schütte R. Information systems for retail companies. In: Dubois E, Pohl K, editors. *International Conference on Advanced Information Systems Engineering*. Essen, Deutschland: Springer; 2017. p. 13–25.
78. Schütte, R., Vetter, T.: Analyse des digitalisierungspotentials von handelsunternehmen. In: *Handel 4.0*, pp. 75–113. Springer, Wiesbaden (2017)
79. Singh P, Agrawal V. A Collaborative Model for Customer Retention on User Service Experience. In: Bhatia SK, Shailesh T, Mishra KK, Tiwari MC, editors. *Advances in Computer Communication and Computational Sciences*. 924th ed. Singapore: Springer Nature; 2019. p. 55–64.
80. Suponenkovs, A., Sisojevs, A., Mosāns, G., Kampars, J., Pinka, K., Grabis, J., Locmelis, A., Taranovs, R.: Application of image recognition and machine learning technologies for payment data processing review and challenges. In: 5th IEEE Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE), pp. 1–6 (2017). <https://doi.org/10.1109/AIEEE.2017.8270536>
81. Thornton C, Hutter F, Hoos HH, Leyton-Brown K. Auto-weka: Combined selection and hyperparameter optimization of classification algorithms. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining Part*. 2013;F1288:847–55.
82. Trinh, H., Fan, Q., Jiyan, P., Gabbur, P., Miyazawa, S., Pankanti, S.: Detecting human activities in retail surveillance using hierarchical finite state machine. In: ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, pp. 1337–1340 (2011)
83. Veres, S.M., Molnar, L., Lincoln, N.K., Morice, C.P.: Autonomous vehicle control systems - A review of decision making. *Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering* **225**(3), 155–195 (2011). <https://doi.org/10.1177/20413041110394727>
84. Verhoef, P.C., Kannan, P.K., Inman, J.J.: From Multi-Channel Retailing to Omni-Channel Retailing. Introduction to the Special Issue on Multi-Channel Retailing. *Journal of Retailing* **91**(2), 174–181 (2015)
85. Verhoef PC, Lemon KN, Parasuraman A, Roggeveen A, Tsirios M, Schlesinger LA. Customer Experience Creation: Determinants, Dynamics and Management Strategies. *J Retail*. 2009;85(1):31–41.
86. Verstraete G, Aghezaf EH, Desmet B. A data-driven framework for predicting weather impact on high-volume low-margin retail products. *J Retail Consum Serv*. 2019;48:169–77.
87. Vinodhini G, Chandrasekaran RM. Opinion mining using principal component analysis based ensemble model for e-commerce application. *CSI Transactions on ICT*. 2014;2(3):169–79. <https://doi.org/10.1007/s40012-014-0055-3>.
88. Vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Clevén, A., Brocke, J.V., Reimer, K.: Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. In: Proceedings of the 17th European Conference on Information Systems (ECIS '17), vol. 9, p. 2206–2217. Verona, Italy (2009). <https://doi.org/10.1108/09600031211269721>
89. Wan, M., Ni, J., Misra, R., McAuley, J.: Addressing marketing bias in product recommendations. In: WSDM 2020 - Proceedings of the 13th International Conference on Web Search and Data Mining, pp. 618–626 (2020). <https://doi.org/10.1145/3336191.3371855>
90. Wang L, Fan H, Wang Y. Site selection of retail shops based on spatial accessibility and hybrid bp neural network. *ISPRS Int J Geo Inf*. 2018;7(6):202.
91. Weber F, Schütte R. Digitalisierung von Handelsunternehmen - diskutiert am Beispiel der Preispolitik. *Wirtschaftsinformatik & Management*. 2021;13(2):144–53. <https://doi.org/10.1365/s35764-020-00312-9>.
92. Weber, F., Schütte, R.: State-of-the-art and adoption of artificial intelligence in retailing big-data-analysen im promotion management view project coming closer to reality-neural correlates of tam view project digital policy, regulation and governance state-of-the-art and adop. [researchgate.net](https://www.researchgate.net) (2019)

93. Weber FD, Schütte R. State-of-the-art and adoption of artificial intelligence in retailing. Regulation and Governance: Digital Policy; 2019.
94. Wulfert, T., Betzing, J.H., Becker, J.: Eliciting Customer Preferences for Shopping Companion Apps: A Service Quality Approach. In: Proceedings of the 14th Internationale Tagung Wirtschaftsinformatik (WI2019), pp. 1220–1234. Siegen, Deutschland (2019)
95. Xi, N., Hamari, J.: Vr shopping: A review of literature. In: AMCIS 2019 Proceedings (2019)
96. Yu, W., Sun, Z., Liu, H., Li, Z., Zheng, Z.: Multi-level deep learning based e-commerce product categorization. In: eCOM@ SIGIR (2018)
97. Zhang D, Zhou L, Kehoe JL, Kilic IY. What Online Reviewer Behaviors Really Matter? Effects of Verbal and Nonverbal Behaviors on Detection of Fake Online Reviews. *J Manag Inf Syst.* 2016;33(2):456–81.
98. Zhang J, Li C, Kosov S, Grzegorzec M, Shirahama K, Jiang T, Sun C, Li Z, Li H. LCU-Net: A novel low-cost U-Net for environmental microorganism image segmentation. *Pattern Recogn.* 2021;115:1–17. <https://doi.org/10.1016/j.patcog.2021.107885>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.