Machine Learning in Business:
A Process and Project Perspective

Dissertation

zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaft

der Rechts- und Wirtschaftswissenschaftlichen Fakultät

der Universität Bayreuth

Vorgelegt

von

Tobias Albrecht

aus

Bayreuth

Dekan:	Prof. Dr. André Meyer
Erstberichterstatter:	Prof. Dr. Maximilian Röglinger
Zweitberichterstatter:	Prof. Dr. Robert Keller
Tag der mündlichen Prüfung:	06.12.2023

To my family

Helga, Jürgen, Carina,

and Carmen

♡

The completion of this dissertation marks the end of an extremely instructive, challenging, and happy phase of my life. I would like to thank everyone who contributed to this experience. In particular, I want to thank my Ph.D. advisor, Max, who has always challenged and encouraged me. I would also like to thank all my co-authors, colleagues, and friends at the FIM Research Center, the Branch Business & Information Systems Engineering of the Fraunhofer FIT, and the University of Bayreuth for making the past years special.

Copyright Statement

The following sections are partly comprised of content from the research papers included in this thesis. To improve the readability of the text, I omit the standard labeling of these citations.

Abstract

At the core of present-day artificial intelligence (AI), machine learning (ML) is a transformative force in the business world. While the tremendous potential of ML is reflected by large-scale investments in industry and research, organizations that are not among the technology leaders are still struggling with effective ML implementation. Accordingly, ML initiatives often fail to realize actual performance gains and drive business value. ML workflows and lifecycle models provide standardized procedures with considerable technical depth. However, understanding and orchestrating the socio-technical process of ML implementation in business remains one of the most critical challenges for scholars and professionals alike. Therefore, theoretical guidance on how to approach different ML methods in the context of diverse industry-specific socio-technical problems is required across all phases of ML implementation (i.e., *business understanding, data understanding, data preparation, modeling, evaluation, and deployment*). In addition, it is essential to progress from a (single-)project perspective to the process level and expand the scope of ML implementation.

This dissertation comprises six research papers that aim to contribute to a better understanding of the socio-technical process of ML implementation in business in two ways: First, this thesis zooms out from a (single-)project perspective to provide guidance on how to examine ML opportunities and business processes in an integrated way (i.e., *process fit*) and how to scale novel ML-enabled processes (i.e., *process ramp-up*) by expanding the scope of previous concepts of ML implementation. Research Paper P1 presents a blueprint for investigating the opportunities of leading-edge digital technologies such as ML, emerging from the dynamic interrelation between technology, processes, and a leading-edge application domain through the theoretical lens of affordance theory. Research Paper P2 introduces business process ramp-up management (BPRUM) as a new business process management (BPM) capability area and presents action-oriented sub-capabilities that provide managerial tools for the effective scaling of novel ML-driven business processes.

Second, this thesis transfers theoretical ML knowledge to applications that solve (domain-)specific problems at the project level. It provides theoretical guidance to practitioners in applying ML methods across all phases of ML implementation and aims to expand the ML knowledge base by gaining practical insights from the application of ML

in leading-edge domains (i.e., mobility and CRM) to advance theory. Research Paper P3 analyzes the role of spatial access in the individual decision-making of carsharing users in small urban areas and presents an XML approach to reveal its influence on ML predictions of user behavior (i.e., *data understanding*). Thereafter, Research Paper P4 explores generative ML for creating synthetic trip data to support carsharing decision-making by overcoming the barrier of limited data access during the introduction and expansion of new services (i.e., *data preparation*). Research Paper P5 and Research Paper P6 examine contact center forecasting. The studies propose a novel forecasting method including contextual information and present an evaluation approach for ML models in this domain (i.e., *modeling* and *evaluation*).

The thesis concludes by pointing to limitations of the presented work as well as directions for future research. Motivated by the need to sustainably realize ML implementation in business, the overall purpose of this dissertation is to guide scholars and practitioners in understanding and performing the operational, methodological, and managerial practices associated with the intricate socio-technical process of effective ML implementation. This thesis strives to contribute to this area of research at both a process and a project level.

Table of Contents

List of Figures

List of Tables

I Introduction

I.1 Motivation

Artificial intelligence (AI) has evolved from a visionary concept in computer science laboratories to a transformative force in the business world. Fueled by broad access to required computing power, ubiquitous data availability, and the practicability of new programming frameworks, machine learning (ML) stands out as one of the main drivers of this development (Brynjolfsson and McAfee 2017; Goodfellow et al. 2016; Janiesch et al. 2021). ML is claimed to have transformational potential across sectors and industries, ranging from mobility (Willing et al. 2017; Zhang et al. 2019) to supply chain management (Shao et al. 2021; Toorajipour et al. 2021) and customer relationship management (CRM) (Kühl et al. 2020; Ma and Sun 2020) to health care (Fechner et al. 2023; Hofmann et al. 2019). Advances in ML have enabled the recent emergence of intelligent systems, providing opportunities to improve business processes (van Dun et al. 2022; vom Brocke et al. 2021), redesign products and services (Haefner et al. 2021; Mariani et al. 2023), and even create new business models (Åström et al. 2022; Sjödin et al. 2021).

ML is considered to be the core of present-day AI (Berente et al. 2021; Jordan and Mitchell 2015). Broadly, AI can be defined as the ability of a machine to perform cognitive functions that reference human intelligence, such as perceiving, reasoning, learning, interacting with the environment, problem solving, decision-making, and even creativity (Berente et al. 2021; Rai et al. 2019; Russell and Norvig 2021). ML models can mimic human cognitive abilities due to their adaptive nature (Janiesch et al. 2021). Hence, ML describes a set of computer-based algorithms capable of iteratively learning from problem-specific training data to improve their task performance without explicitly being programmed (Abdel-Karim et al. 2021; Bishop 2006; Jordan and Mitchell 2015; Kühl et al. 2022). In general, ML can be categorized into three types: supervised learning, unsupervised learning, and reinforcement learning. Supervised ML refers to methods that associate inputs with corresponding target outputs to make predictions or classifications based on labeled training data. Unsupervised ML describes methods that infer underlying patterns in unlabeled data without predefined outcomes, while reinforcement learning includes methods that optimize actions toward a predefined goal based on feedback in the form of penalties and rewards (Goodfellow et al. 2016; Kaplan and Haenlein 2019).

Across all three types of ML, deep learning is a class of ML algorithms based on multi-layer artificial neural networks (NNs) that feature improved learning capabilities for large and high-dimensional data through multiple levels of representation (LeCun et al. 2015; Schmidhuber 2015).

The tremendous potential of ML in business is reflected in the scale of industry attention it has received in recent years. Global spending on AI is predicted to reach $154 billion in 2023, with ML accounting for the majority of it (International Data Corporation 2023; Statista 2023). This represents an increase of nearly 26.9% over the amount spent in 2022, which is in line with the projected compound annual growth rate (CAGR) of 27.0% between 2022 and 2026. Beyond these high projections, scholars and professionals alike continue to seek a better understanding of the underlying concepts, processes, and challenges of implementing the technology (Janiesch et al. 2021). While scholars aim to keep pace with the rapid technological advances of ML research outside of academia (Maslej et al. 2023), organizations that are not among the technology leaders are still struggling to deploy ML effectively (Reis et al. 2020).

Recent research has examined the process of ML deployment with a deep focus on its technical approach (Ashmore et al. 2022; Google 2023; Kühl et al. 2021; Studer et al. 2021). These ML workflows and lifecycle models advance established data science standards such as knowledge discovery in databases (KDD) (Fayyad et al. 1996), the cross-industry standard process for data mining (CRISP-DM) (Chapman et al. 2000), and the team data science process (TDSP) (Microsoft 2022) with ML-specific concepts and practices (e.g., model training and optimization). Shrestha et al. (2021) present an integrated iterative process for ML to solve business problems (see Fig. 1). The process aims to guide the end-to-end realization of *ML projects* in six phases with corresponding sub-activities: (1) The *business understanding* phase includes setting up the project (i.e., specifying the key objectives and identifying relevant data sources). (2) The *data understanding* phase and (3) the *data preparation* phase concern data management (i.e., collecting, analyzing, and preprocessing the data). (4) The *modeling* phase and (5) the *evaluation* phase cover model learning and verification (i.e., building, training, optimizing, assessing, and selecting suitable models). (6) The *deployment* phase describes the operationalization of the model (i.e., integrating and monitoring the model in the productive environment). While valuable, existing process models are by design focused on the technical approach to ML implementation and are often limited to the boundaries of predefined projects.

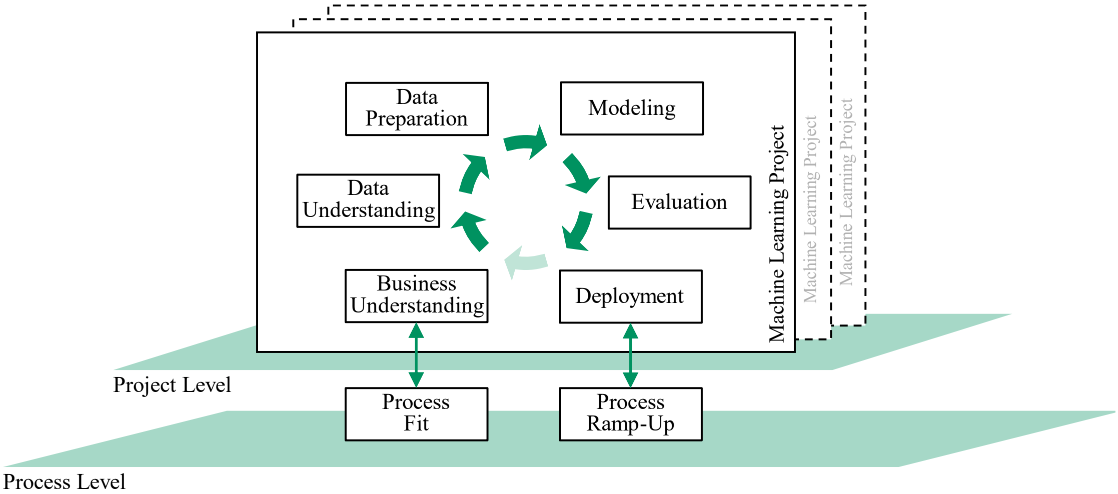

Fig. 1 Focus areas of the dissertation adapted from Shrestha et al. (2021)

Despite the surge in ML interest and industry spending in recent years, few organizations have adopted and deployed ML applications beyond pilot projects (Benbya et al. 2020). Appropriately, a survey among 250 senior managers working on ML projects reports that 47% of professionals find it difficult to integrate ML with existing processes and systems (Deloitte 2017). Consequently, ML initiatives often fail to deliver on their potential to realize performance gains and drive business value, even though they are equipped with the time, effort, and resources to do so (Makarius et al. 2020; Vial et al. 2023). Recent research has started to examine the organizational, technical, and individual challenges that entail the disparity in ML business potential and de facto value creation (Papagiannidis et al. 2023; Ransbotham et al. 2017). Studies on ML capabilities (Lee et al. 2022; Mikalef and Gupta 2021; Wamba-Taguimdje et al. 2020) and organizational readiness (Jöhnk et al. 2021; Pumplun et al. 2019; Uren and Edwards 2023) explore the organizational capacity of leveraging ML-specific resources to support organizational goals. In this context, understanding and orchestrating the socio-technical process through which ML applications are implemented (i.e., planned, developed, and deployed) is an essential factor (Enholm et al. 2022; Weber et al. 2023). While other (technological and organizational) enablers of ML adoption and use (e.g., data, infrastructure, technical skills, top management support) are well understood (Merhi 2023; Zhang et al. 2020b), the socio-technical process of putting ML into action to realize its value-adding potential remains on the research agenda (Mucha et al. 2022).

Existing ML lifecycle models largely focus on the technical workflow during ML implementation, but do not provide guidance on how to sustainably manage and integrate ML projects in organizational practice (i.e., from initiation to target operation) (Vial et al. 2023). To take this next step in how we look at ML implementation, research emphasizes the need of adopting an integrated approach of ML applications and business processes (Reis et al. 2020; Wamba-Taguimdje et al. 2020). In this connection, aligning ML applications with the processes they are supposed to support (i.e., task-technology fit) is essential when planning and initiating ML projects (Jöhnk et al. 2021; Muchenje and Seppänen 2023). The potential of ML is closely related to the purpose it serves within a socio-technical context, which in turn is defined by its actors and processes (Baier et al. 2023; Faulkner and Runde 2019). Thereby, a thorough understanding of the contextual action potential of ML, emerging from its interplay with respective business processes, is needed at the outset to realize meaningful ML projects (Uren and Edwards 2023; Weber et al. 2023).

Realizing ML projects typically involves significant changes to how organizations perform work (e.g., by enhancing human-machine collaboration or ML-enabled process automation) (Brynjolfsson and Mitchell 2017; Teodorescu et al. 2021). Consequently, organizations need to be able to manage the radical process change ML deployment entails to achieve sustainable performance gains (Mikalef and Gupta 2021). This particularly includes implementing and scaling novel ML-driven processes. Thus, theoretical guidance and managerial tools are needed to advance from technical ML deployment to sustainable process change and long-term performance gains (Benbya et al. 2020; Lukyanenko et al. 2019).

In sum, understanding and orchestrating the socio-technical process of ML implementation in business is a critical challenge that engages research and practice in various facets. ML workflows and lifecycle models provide standardized procedures with considerable technical depth (Vial et al. 2023). However, there is a lack of theoretical guidance and actionable insights that facilitate an integrated approach to process-oriented ML implementation (Baier et al. 2019; Maragno et al. 2023; Mucha et al. 2022). To achieve this, more guidelines from theory on how to approach different ML methods in the context of diverse industry-specific socio-technical problems are required across all phases of ML implementation (Abdel-Karim et al. 2021; Maass et al. 2018). In addition, it is essential to zoom out from a (single-)project perspective to the process level and expand the scope of ML implementation (see Fig. 1). In this regard, comprehensive support on how to ensure alignment of ML opportunities and business processes (i.e., *process fit*) at the outset and how to scale novel ML-enabled processes (i.e., *process ramp-up*) after deployment is needed (Benbya et al. 2020; Mikalef and Gupta 2021; Uren and Edwards 2023).

I.2 Research Objectives

Based on the identified research needs, this dissertation aims to contribute to a better understanding of the socio-technical process of ML implementation in business in two ways: First, it is necessary to expand the scope of ML implementation from a (single-)project perspective to the process level to stimulate an integrated approach to process-oriented ML implementation. Therefore, this thesis advances the iterative process for ML to solve business problems presented by Shresta et al. (2021). On the one hand, it adds the notion of *process fit* to the *business understanding* phase. In this connection, this thesis aims to make a methodological contribution by providing a rigorous approach to identifying ML opportunities as an interplay of processes and technology. On the other hand, this work augments the *deployment* phase with *process ramp-up*. Thereby, this dissertation means to support the implementation and scaling of novel ML-enabled business processes by identifying capabilities for effectively managing process change.

Second, theoretical guidance and actionable insights on how to approach ML implementation in the context of diverse industry-specific socio-technical problems are needed. This thesis, therefore, aims to transfer theoretical ML knowledge to applications that solve (domain-)specific problems at the project level, with the goal of supporting practitioners in applying ML methods across all phases of ML implementation (i.e., *data understanding*, *data preparation*, *modeling*, and *evaluation*). Conversely, theoretical insights gained from the practical application of advanced ML methods in leading-edge domains mean to add to the ML knowledge base and provide a basis for scholars to prioritize research foci.

Altogether, this thesis is motivated by the need to sustainably realize ML implementation in business. It strives to guide scholars and practitioners in understanding and performing the operational, methodological, and managerial practices associated with the intricate socio-technical process of effective ML implementation. Its overall purpose is to contribute to this area of research at both a process and a project level.

I.3 Structure of the Thesis and Embedding of the Research Papers

This thesis consists of six research papers that contribute to the stated research objectives. Table 1 provides an overview of the structure of this dissertation and the embedding of the research papers. Collectively, these papers contribute to current research on ML implementation in business at the process and the project level.

I		**Introduction**
II		**Advancing Machine Learning Implementation at the Process Level**
	1	**Leveraging Digital Technologies in Logistics Processes: Systematic Insights from Intra-Logistics** *Albrecht T., Baier M.S., Gimpel H., Meierhöfer S., Röglinger M., Schlüchtermann J., Will L*
	2	**From Zero to Hero: Ramp-Up Management as a New Cross-Cutting Business Process Management Capability** *Albrecht T., Lösser B., Röglinger M.*
III		**Advancing Machine Learning Implementation at the Project Level**
	3	**Are We There Yet? Analyzing the Role of Access Distance in Carsharing in Small Urban Areas** *Albrecht T., Keller R., Röglinger M., Röhrich F.*
	4	**Fake It Till You Make It: Synthetic Data for Emerging Carsharing Programs** *Albrecht T., Keller R., Rebholz D., Röglinger M.*
	5	**Beyond the Beaten Paths of Forecasting Call Center Arrivals: On the Use of Dynamic Harmonic Regression with Predictor Variables** *Rausch T.M., Albrecht T., Baier D.*
	6	**Call Me Maybe: Methods and Practical Implementation of Artificial Intelligence in Call Center Arrivals' Forecasting** *Albrecht T., Rausch T.M., Derra N.D.*
IV		**Conclusion**
V		**References**

Table 1 Structure of this thesis and embedding of the research papers

After motivating the scope of this thesis and defining the research objectives (Section I), Section II (including Research Papers P1 and P2) presents research that advances ML implementation at the process level. Thereby, this dissertation provides guidance on how to examine ML opportunities and business processes in an integrated way (i.e., *process fit*) and how to scale novel ML-enabled processes (i.e., *process ramp-up*) by expanding the previous scope of ML implementation. Research Paper P1 presents a blueprint for investigating the opportunities of leading-edge digital technologies such as ML, emerging from the dynamic interrelation between technology, processes, and the domain-specific environment (i.e., Logistics 4.0), through the theoretical lens of affordance theory. Apart from identifying affordances in a frontline information systems (IS) application domain with high ML potential, the study provides insights on how to methodologically develop a fine-grained understanding of ML action potential in different organizational contexts. Research Paper P2 introduces business process ramp-up management (BPRUM) as a new business process management (BPM) capability area. It presents 40 action-oriented sub-capabilities that provide hands-on knowledge and managerial tools for the effective implementation and scaling of novel ML-driven business processes. These two research papers help organizations to adopt practices that facilitate an integrated approach to process-oriented ML implementation.

Section III (including Research Papers P3, P4, P5, and P6) presents research that advances ML implementation at the project level. Therefore, this thesis contributes theoretical guidance and actionable insights on how to approach ML implementation in the context of diverse industry-specific socio-technical problems. It considers two leading-edge application domains, namely mobility (i.e., carsharing) and CRM (i.e., customer service). Research Paper P3 investigates the role of spatial access in the individual decision-making of carsharing users in small urban areas and presents an explainable ML (XML) approach to analyze its influence on ML predictions of user behavior (i.e., *data understanding*). Research Paper P4 explores generative ML for creating synthetic trip data to support carsharing decision-making by overcoming the barrier of limited data access during the introduction and expansion of new services (i.e., *data preparation*). It presents a blueprint for investigating the use of synthetic mobility data in a broader range of domains. Finally, Research Paper P5 and Research Paper P6 consider the forecasting of customer inquiries in contact centers by proposing a novel forecasting method including contextual information and by presenting an evaluation approach for ML models in this domain (i.e., *modeling* and *evaluation*). These four research papers address domain-specific business problems to derive theoretical guidelines and actionable practical insights.

Section IV concludes this thesis by summarizing the work, outlining limitations, and highlighting avenues for future research. Section VI provides an index of the research papers, my contributions to the papers, and the complete version of each research paper.

II Advancing Machine Learning Implementation at the Process Level

As motivated above, an integrated approach of ML applications and business processes is needed to achieve a holistic understanding of the socio-technical process of ML implementation in business. Therefore, this thesis provides research that enables organizations to identify and systemize the opportunities afforded by ML as an interplay of domain-specific processes and technology (Section II.1; Research Paper P1) and to effectively introduce and scale novel ML-driven business processes (Section II.2; Research Paper P2). Thereby, the two research papers presented in this section directly connect to and expand the initial phase (i.e., *business understanding*) and the final phase (i.e., *deployment*) of ML implementation.

II.1 Business Understanding and Process Fit

Aligning ML applications with the processes they are supposed to support is essential when planning and initiating ML projects (Jöhnk et al. 2021; Muchenje and Seppänen 2023). To do so, organizations need to understand the opportunities brought about by ML as well as their interplay with diverse tasks in domain-specific environments (Borges et al. 2021; Yang et al. 2021). In leading-edge IS application domains, where tasks and processes are transforming rapidly and ML implementation is most pervasive, this is particularly challenging (Kerzel 2021; Toorajipour et al. 2021). Against this backdrop, Research Paper P1 pursues a rigorous approach to identifying and systemizing the opportunities afforded by emerging digital technologies such as ML as an interplay of processes and technology. To this end, it chooses the frontline industrial development of Logistics 4.0 as the application domain (Gupta et al. 2019; Strandhagen et al. 2017), where organizational efforts of digital transformation in general, and ML implementation in particular, are challenged by cyber-physical processes and cross-organizational flows of data and information (Klumpp and Zijm 2019; Sigov et al. 2022).

To achieve the above research objective, Research Paper P1 synthesizes current academic research and advanced industrial insights through the lens of affordance theory (Gibson 1986; Majchrzak and Markus 2013), which has gained momentum in IS research for developing a fine-grained understanding of the action potential of digital technologies in different organizational environments (Islam et al. 2020; Seidel et al. 2013; Wendt et al. 2021). The study follows a two-phase research approach. First, it performs a systematic literature review (Webster and Watson 2002; Wolfswinkel et al. 2013) to rigorously derive a catalog and conceptual framework of digital technology affordances and associated practical manifestations as cues for potential use. Second, a qualitative interview study with ten subject matter experts from academia and industry is conducted to evaluate, expand, and refine the results (Bettis et al. 2015; Goldkuhl 2012).

Fig. 2 summarizes the main results of the study. The conceptual framework illustrates the scope and interrelation of the identified digital technology affordances in four layers (i.e., data, manual tasks, goods and assets, and decisions and management) according to the action potential they comprise and the nature of logistics processes they concern. The framework is complemented by a catalog of the ten affordances that explicates their associated practical manifestations and provides concrete examples for application from literature to foster practical understanding. In addition, a mapping between the digital technologies investigated and the process elements they apply to is provided to allow deeper insights into how (technology perspective) and where (process perspective) the affordances emerge in the present sociotechnical system.

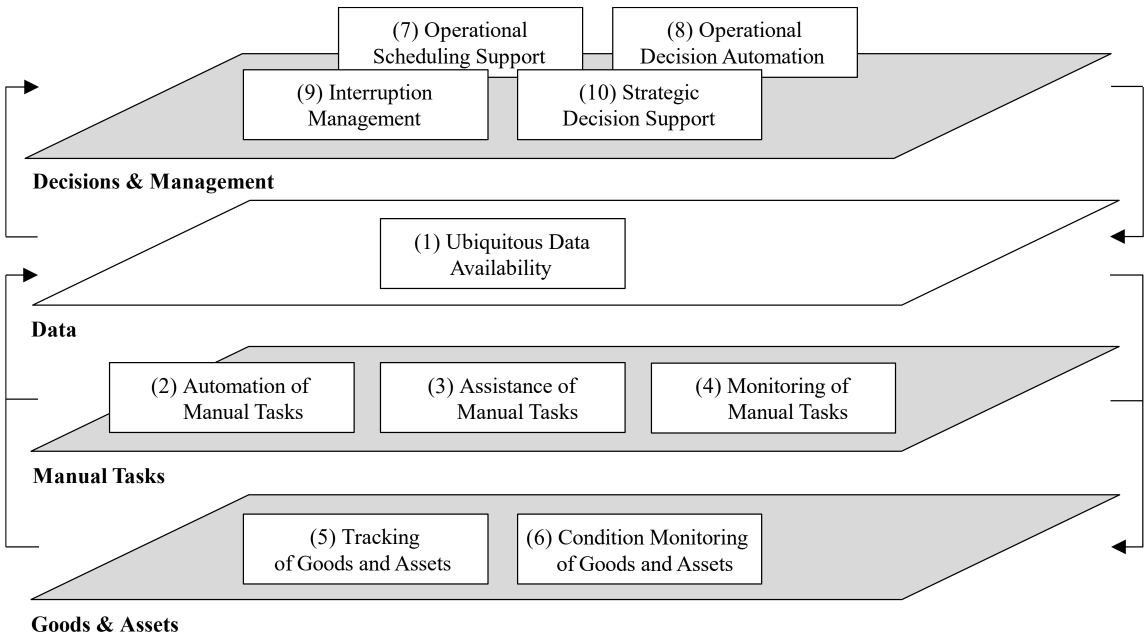

Fig. 2 Conceptual framework of affordance layers for digital technologies such as ML in Logistics 4.0

The results of this study offer valuable implications for IS research in terms of understanding technology affordances in frontline industrial developments. While previous literature provides a useful overview of the predominant technological developments in Logistics 4.0 and detailed insights into individual use cases (Frank et al. 2019; Leofante et al. 2019), it does not offer a comprehensive perspective on the systematic opportunities digital technologies afford in this context. Further, previous studies do not relate their potential to specific processes and tasks (Shao et al. 2021; Winkelhaus and Grosse 2020). Consequently, this study advances the existing body of knowledge by presenting a novel and theoretically well-founded affordance perspective on Logistics 4.0 that is in line with recent IS research drawing on affordance theory (Islam et al. 2020; Seidel et al. 2013; Wendt et al. 2021). The presented catalog and conceptual framework of affordances establish holistic patterns of the action potential that emerges from the interplay between digital technologies with certain features and processes as specific organizational sets of actions. To achieve a close connection between theoretical and application-oriented research, the paper combines theory-based knowledge from the structured literature review with detailed expert insights on real-world examples from an interview study. Thus, the results of this methodological approach purposefully address the need for a comprehensive understanding of how digital technologies such as ML can support organizational processes in a domain-specific context.

Against this background, the primary contribution of Research Paper P1 to the research objectives of this thesis is a methodological one. The study provides a blueprint for systematically assessing the opportunities of ML, emerging from the dynamic interrelation between technology, processes, and domain-specific environments through the theoretical lens of affordance theory. Thus, the study imparts theoretical knowledge on how to methodologically develop a nuanced understanding of ML's action potential in different organizational processes. Adopting this approach supports managers in assessing and actively monitoring the extent to which their own (planned) projects leverage the opportunities of ML in the respective process environment. Based on the results, managers can determine the technological status quo of their business processes by assessing which ML affordances are currently covered and which may advance certain processes. Finally, the process orientation of this study further encourages managers to adopt a task-technology-fit perspective when discussing where to catch up or how and where to proceed when developing their ML portfolio. This helps avoid isolated solutions and offsets potential subjective biases toward certain well-tried methods.

Overall, Research Paper P1 adds the notion of *process fit* to ML implementation and presents a process-oriented approach how to identify and systemize the opportunities afforded by ML as an interplay of processes and technology. Thereby, it employs affordance theory as a theoretical lens to derive patterns that exist through the symbiotic relationship of ML as the technological artifact and an actor's set of actions in leading-edge process environments instead of just focusing on technology functionalities (Majchrzak and Markus 2013).

II.2 Deployment and Process Ramp-Up

The research presented in Research Paper P1 established a process-oriented approach to support ML implementation at the outset. In addition, completing ML projects with lasting results often requires making radical changes to the way work is performed (Brynjolfsson and Mitchell 2017; Gross et al. 2021; Teodorescu et al. 2021). Hence, the successful implementation and scaling of novel ML-driven business processes are crucial for organizations seeking to realize actual performance gains from ML projects (Grisold et al. 2021; Mikalef and Gupta 2021). The associated business process ramp-up (BPRU), i.e., the operational implementation of novel processes following the process design as well as their quantitative and qualitative scaling, often involves radically rethinking existing systems and structures, which can be disruptive and difficult to manage (Gross et al. 2019). Thus, theoretical guidance and managerial tools are needed to advance from technical ML deployment to sustainable process change and long-term performance gains (Benbya et al. 2020; Lukyanenko et al. 2019).

Research Paper P2 takes a BPM perspective to address this research need. As the organizational capability of overseeing how work is performed to ensure consistent outcomes and to take advantage of improvement opportunities (Dumas et al. 2018; van der Aalst 2013), BPM mainly focuses on repetitive transactional process performance rather than rapid transformational change (Bandara et al. 2021; Rosemann 2014). This is also reflected by the pertinent literature structuring BPM through capability frameworks by identifying and categorizing the most relevant capability areas for the successful implementation of process orientation in organizations (de Bruin and Rosemann 2007; Kerpedzhiev et al. 2021; Poeppelbuss et al. 2015; van Looy et al. 2022). Against this backdrop, this study argues that a new BPM capability area dedicated to BPRUM is needed to account for the omnipresent need for organizations to adapt to radical change.

In line with recent research asking to challenge and update existing capability frameworks and their capability areas (Kerpedzhiev et al. 2021; Rosemann 2014; van der Aalst 2013), Research Paper P2 poses the twofold research question as follows: *How does BPRUM add to modern BPM capability areas and what are its sub-capabilities?*

To answer the research question, Research Paper P2 performs an exploratory interview study (Schultze and Avital 2011), conducting 21 in-depth interviews with subject matter experts from different organizations of various sizes (i.e., from start-ups to international enterprises) and industries (e.g., software, technology, and logistics). For data analysis, the study adopts a well-established three-step coding process based on grounded theory analysis techniques (i.e., open, axial, and selective coding) (Corbin and Strauss 1990) to iteratively synthesize observations and develop the results. To evaluate the results, a survey collecting detailed feedback from 10 interview partners (IPs) via an online survey tool (i.e., a response rate of 45.45 percent) is conducted in line with the recommendations on ex-post evaluation of naturalistic research artifacts (Venable et al. 2012). Thereby, the study synthesizes frontline industrial insights following established approaches to BPM capability development (Baumbach et al. 2020; de Bruin and Rosemann 2007).

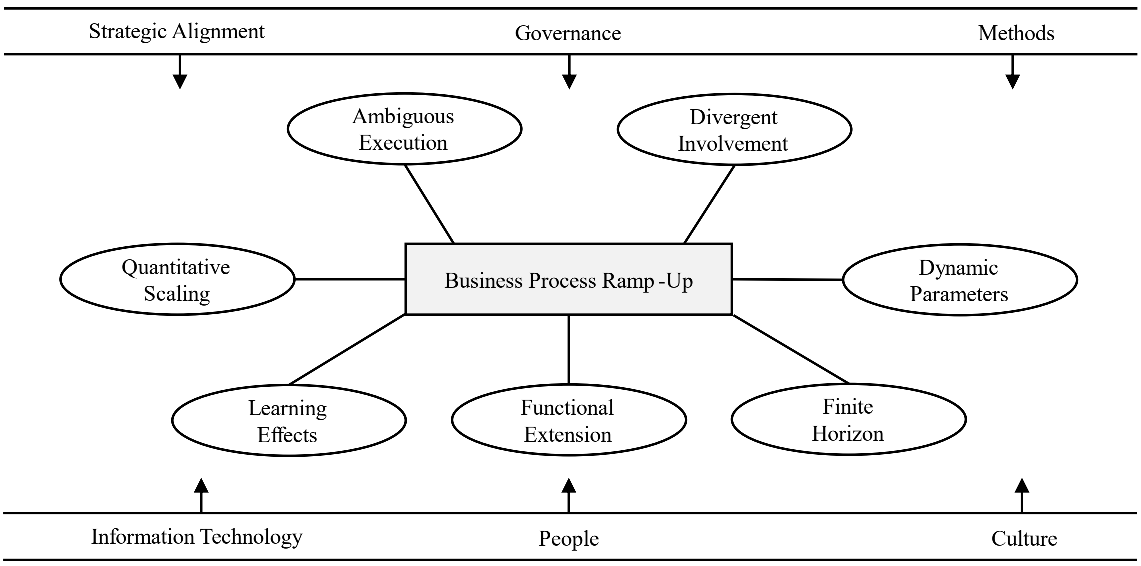

Fig. 3 Overview of BPRU characteristics and management areas

The main results of this research are summarized in Fig. 3 and Table 2. First, seven distinct characteristics outline the nature of BPRU (Fig. 3): *ambiguous execution* (i.e., conflicting activities, interruptions, and errors based on the initial process design), *divergent involvement* (i.e., resistance, euphoria, or indifference of stakeholders), *dynamic param-*

eters (i.e., uncertain and context-dependent planning factors), *finite horizon* (i.e., seamless or abrupt transition to the steady state), *functionality extension* (i.e., additional process paths), *learning effects* (i.e., qualitative improvements in process execution and outcome), and *quantitative scaling* (i.e., erratically increasing process demand). This intermediate result emphasizes the particular challenges of BPRUM compared to the continuous management of business processes in their steady state.

Thereafter, this study presents 40 action-oriented sub-capabilities that provide hands-on knowledge and practical guidance for effective BPRUM (Table 2). Each sub-capability features a description, a unique ID, specific references to the IPs, and is also assigned to the BPRU characteristic it mainly addresses. The results reveal that the implementation and scaling of novel processes require distinct BPM capabilities across all core elements of BPM (i.e., Strategic Alignment, Governance, Methods, IT, People, and Culture) (Rosemann and vom Brocke 2015).

	ID	**Description**	**References**	**Main Characteristic**
Strategic Alignment	Str. Alig-1	Develop and promote a collective process vision to create organizational support for the business process ramp-up.	IP3, IP4, IP15, IP16, IP17, IP21	Divergent Involvement
	Str. Alig-2	Set an ambitious timeline and clear goals for the business process ramp-up in line with the organizational strategy.	IP1, IP2, IP4, IP5, IP6, IP13, IP15, IP16, IP17, IP21	Finite Horizon
	Str. Alig-3	Weigh strategic trade-offs during the business process ramp-up (e.g., simplicity and security) to make conscious design decisions.	IP2, IP3, IP12	Dynamic Parameters
	Str. Alig-4	Ensure management attention to allocate sufficient resources during the business process ramp-up.	IP1, IP3, IP11, IP16	Quantitative Scaling
	Str. Alig-5	Define the initial business process version to be simple and close to proven practices to facilitate stakeholder adoption.	IP3, IP4, IP8, IP9, IP19	Divergent Involvement
	Str. Alig-6	Specify the temporal and functional go-live strategy of the new business process (e.g., smooth or abrupt) to prevent unintentional parallelisms.	IP5, IP12, IP13, IP17	Finite Horizon
	Str. Alig-7	Develop mitigation strategies for potential problems that may arise during the business process ramp-up to handle non-desirable deviations and exceptions.	IP12, IP13, IP15, IP16	Ambiguous Execution
Governance	Gov-1	Establish clear management responsibilities to facilitate fast decision-making during the business process ramp-up.	IP1, IP10, IP18, IP19	Dynamic Parameters
	Gov-2	Repeatedly align organizational structures and process requirements to ensure distinct interfaces during the business process ramp-up.	IP5, IP6, IP8, IP9, IP18, IP20	Ambiguous Execution
	Gov-3	Define clear (non-)financial performance indicators to control goal achievement during the business process ramp-up.	IP1, IP5, IP18	Ambiguous Execution
	Gov-4	Link goal achievement during the business process ramp-up to stakeholder incentives to motivate the completion of non-joyful tasks.	IP5, IP12, IP20	Divergent Involvement

	Gov-5	Manage dependencies to other processes to prevent costly readjustments during the business process ramp-up.	IP2, IP4, IP10, IP15, IP17, IP19, IP21	Dynamic Parameters
	Gov-6	Empower stakeholders to change process roles and responsibilities to gain cross-functional insights during the business process ramp-up.	IP1, IP5, IP9	Learning Effects
	Gov-7	Leverage insights from prior business process ramp-ups (e.g., by involving experienced stakeholders) to anticipate potential pitfalls and adopt good practices.	IP6, IP14, IP21	Learning Effects
	Gov-8	Develop escalation mechanisms to respond to unexpected events during the business process ramp-up.	IP7, IP11, IP12, IP16	Dynamic Parameters
	Gov-9	Leverage workforce flexibility to respond to demand volatility during the business process ramp-up.	IP6, IP21, IP19	Quantitative Scaling
	Gov-10	Assess the behavior of frequent and crucial cases during the business process ramp-up to efficiently manage significant process deviations.	IP1, IP2, IP5, IP12, IP13	Functionality Extension
Methods	Method-1	Initiate early and continuous process improvement to detect and overcome challenges during the business process ramp-up.	IP8, IP15, IP19	Learning Effects
	Method-2	Benchmark the process regularly during the business process ramp-up.	IP5, IP14, IP15	Learning Effects
	Method-3	Share good practices during the business process ramp-up to promote organizational learning and meta-learning.	IP1, IP6, IP15, IP19	Learning Effects
	Method-4	Employ regular performance dialogs, peer reviews, and retrospectives during the business process ramp-up to strengthen personal learning effects.	IP2, IP4, IP7, IP9, IP12, IP15, IP19, IP21	Learning Effects
	Method-5	Test the process hands-on and end-to-end using real cases to ensure viability.	IP1, IP6, IP8, IP11, IP13, IP20	Ambiguous Execution
	Method-6	Start with standard cases of the process to quickly leverage learning curve effects during the business process ramp-up.	IP11, IP12, IP13	Learning Effects
	Method-7	React to problems quickly during the process ramp-up and solve them pragmatically to build trust in the process.	IP1, IP12, IP20	Divergent Involvement
Information Technology	IT-1	Adopt suitable lightweight process technology and digital platforms to support scalability.	IP2, IP5, IP8, IP17	Quantitative Scaling
	IT-2	Align the process technology with existing (IT) infrastructure and interfaces.	IP1, IP2, IP5, IP8, IP21	Quantitative Scaling
	IT-3	Test technical systems early and rigorously during the business process ramp-up to ensure basic functionality.	IP5, IP8, IP11, IP20	Dynamic Parameters
	IT-4	Increase technological capacities to respond to demand volatility during the business process ramp-up.	IP1, IP19, IP21	Quantitative Scaling
	IT-5	Model and monitor the process early during the business process ramp-up to increase its binding character.	IP12, IP14, IP18	Ambiguous Execution
	IT-6	Establish a profound understanding of the process technology among the involved stakeholders to reduce external dependencies during the business process ramp-up.	IP1, IP3, IP12, IP15	Dynamic Parameters
	IT-7	Pre-structure the data generated by the process to identify usage potential and ensure to keep sensitive information private during the business process ramp-up.	IP1, IP6, IP9	Learning Effects

Table 2 Overview of the BPRUM sub-capabilities

ML implementation in business requires prescriptive knowledge of managing the introduction and scaling of novel ML-driven business processes. Research Paper P2 contributes to this need by proposing BPRUM as a new BPM capability area and exploring relevant sub-capabilities. In doing so, it addresses the current challenge of practitioners who are confronted with managing large-scale process change initiatives that increasingly include establishing new processes in response to the socio-technical changes brought about by ML applications. Specifically, the action-oriented sub-capabilities provide low-threshold access to the components and managerial tools of effective BPRUM. Although the novel capability area still needs to be extended towards a maturity model, its hands-on sub-capabilities can be used as guidance for organizations aiming to develop competence in BPRUM to realize sustainable performance gains from ML projects.

Moreover, this research supports practitioners in assessing and actively monitoring the extent to which current initiatives cover the components of effective BPRUM. In this context, the presented sub-capabilities can build a foundation for fit-gap analyses in mature ML projects. Identified gaps can then be addressed by purposefully adopting BPRUM practices of the respective management area or by intensifying cross-functional cooperation with other corporate functions and teams. From a theoretical perspective, the results facilitate further theorizing on the implementation and scaling of novel ML-driven business processes. This study lays the grounds for tailoring the results to the specific challenges of ML implementation by identifying 40 action-oriented sub-capabilities structured according to the established core elements of BPM (Rosemann and vom Brocke 2015). Thereby, it presents distinct categories to be considered to deepen the academic understanding of how to manage the ramp-up of novel business processes to sustainably realize radical process change.

In sum, this section presented two significant contributions that advance ML implementation at the process level. First, a process-oriented approach to identifying and systemizing the opportunities afforded by ML as an interplay of processes and technology was introduced. Second, a novel BPM capability area and action-oriented sub-capabilities for the effective implementation and scaling of novel ML-driven business processes were presented. Together, the two research papers provide guidance on how to examine ML opportunities and business processes in an integrated way (i.e., *process fit*) and how to scale novel ML-enabled processes (i.e., *process ramp-up*) by expanding the previous scope of ML implementation.

III Advancing Machine Learning Implementation at the Project Level

Moving from the process to the project level, theoretical guidance and actionable insights on how to approach ML implementation in the context of diverse industry-specific socio-technical problems are needed. This is particularly true for fast-moving application domains, where organizations require support in applying ML methods and research can gather novel insights on real-world applications to advance theory. This thesis addresses this need by investigating various phases of ML implementation in two leading-edge application domains, namely mobility (i.e., carsharing) and CRM (i.e., customer service). Specifically, this thesis provides research in the carsharing domain that employs XML to analyze user decision-making (Section III.1; Research Paper P3) and that explores generative ML for creating synthetic trip data to support operator decision-making (Section III.1; Research Paper P4) (i.e., *data understanding* and *data preparation*). In addition, this dissertation proposes a novel method in contact center forecasting (Section III.2; Research Paper P5) and presents an evaluation approach for ML models in this domain (Section III.2; Research Paper P6) (i.e., *modeling* and *evaluation*). Thereby, the four research papers presented in this section embrace domain-specific business needs to support practitioners in applying ML methods and derive theoretical insights from ML applications in real-world settings.

III.1 Data Understanding and Preparation

As carsharing has gained traction as a sustainable concept to address the immediate challenges of urban mobility, it has developed into an active field of ML research (Ferrero et al. 2018; Prinz et al. 2022). By providing registered users with short-term access to a fleet of shared vehicles and offering payment tied to the distance traveled and the usage time, carsharing is part of an ongoing shift from vehicle ownership to vehicle access as a service (Schmöller and Bogenberger 2020; Xu and Meng 2019). It has been found to reduce CO_2 emissions, noise pollution, congestion, and parking shortages by decreasing car ownership (Kim et al. 2019; Vélez 2023), lowering fuel consumption (Chen and Kockelman 2016; Meng et al. 2020), and fostering intermodal transportation (Amatuni et al. 2020; Chicco and Diana 2021). Fueled by these benefits, the demand for carsharing services is projected to continue to grow (Mounce and Nelson 2019; Shaheen and Cohen 2020). While carsharing programs were previously the preserve of large operators from the private sector, public institutions, and small operators are increasingly entering the market (Zhang et al. 2020a). In the course of this development, emerging programs are coming up with innovative service offers and tapping into new markets that are no longer limited to metropolitan areas but increasingly include rural and suburban areas (Baumgarte et al. 2022; Illgen and Höck 2020; Lagadic et al. 2019; Rotaris and Danielis 2018).

As the number of carsharing users and the demand for more flexible service offers grow, decision-making for carsharing operators becomes ever more complex (Laporte et al. 2018). Hence, insights on user behavior and predictions of trip characteristics from ML models have become essential for carsharing operators to make informed decisions and optimize their systems (Baumgarte et al. 2022; Cheng et al. 2021). Previous literature highlights the high value of ML for analyzing and predicting, inter alia, trip distances (Baumgarte et al. 2022), user demand (Prinz et al. 2022), and supply imbalances (Willing et al. 2017). Consequently, large carsharing operators increasingly leverage their rich database to generate insights for improving the efficiency and sustainability of their operations, reducing costs, and providing a better service experience for their users (Golalikhani et al. 2021; Yao et al. 2022). In contrast, emerging carsharing programs and programs operating in less densely populated areas are mostly excluded from this frontline industrial development. Besides budget constraints and strict data protection regulations, many small carsharing operators lack technical expertise and struggle with limited data availability or understanding (Lagadic et al. 2019; Vanheusden et al. 2022). Thus, application-oriented research is needed that embraces the business needs of emerging carsharing programs and provides theoretical guidance in implementing ML-driven solutions.

Against this backdrop, Research Paper P3 aims to investigate the role of access distance in carsharing user behavior and to derive its implications for strategic decision-making in the development of carsharing programs in less densely populated areas (e.g., station locations and service areas) as well as for ML predictions of travel behavior as a basis for operational service improvements (e.g., vehicle availability). Therefore, this study raises the following research question: *How does access distance affect carsharing usage behavior in small urban areas?*

To pursue this research goal, this paper draws on a two-phase research approach. It includes both descriptive analysis and predictive modeling approaches while it systematically follows an established data analysis and ML workflow (Chapman et al. 2000; Kühl et al. 2021). First, it analyzes two years' worth of data from a real-world municipal car-sharing program in a small urban area, including 5128 active users and 119,146 trips, using descriptive statistics. It compares and triangulates the results with relevant findings from metropolitan areas. Second, the study expands the descriptive insights by exploring the impact of access distance and other usage factors on the prediction of individual travel behavior. To determine variable importance, the study first uses three different types of ML algorithms (i.e., random forests (RF), gradient boosting machines (GBM), and NN) and one benchmark model (i.e., generalized linear model (GLM)) for the prediction of the usage duration per trip. The selected models reflect the spectrum of commonly used methods in previous research in this domain (Wang et al. 2021; Willing et al. 2017). The study then employs XML in the form of model-agnostic permutation feature importance (Fisher et al. 2019) and accumulated local effects (ALE) plots (Apley and Zhu 2020) that have gained momentum in research to determine the impact of explanatory variables on the prediction performance of diverse ML models (Gao et al. 2021; Severino and Peng 2021).

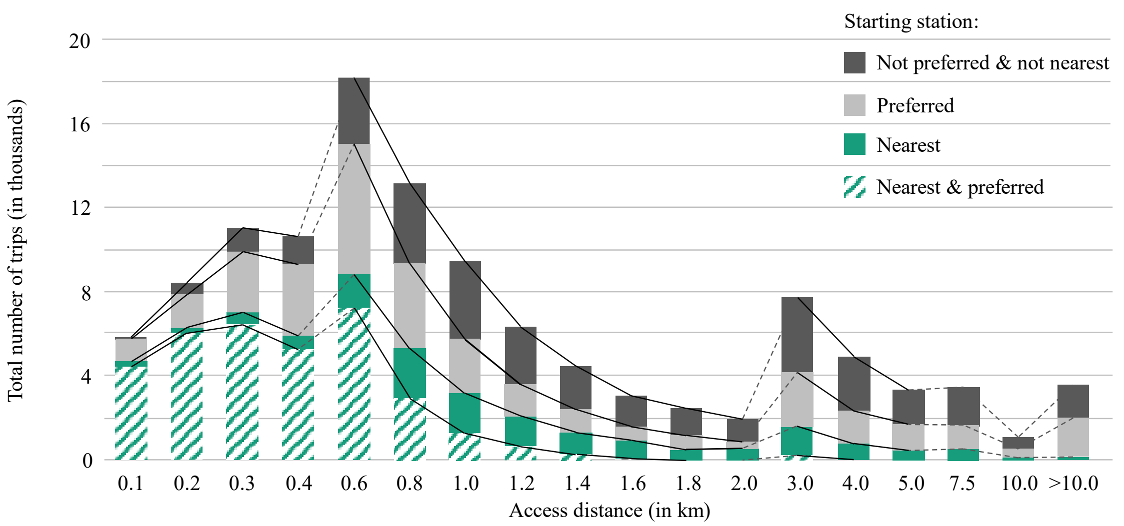

Fig. 4 Overview of trips per access distance (station preferences)

The overall results of the study show that users' access distance is highly relevant for strategic decision-making in the development of carsharing programs in less densely populated areas (e.g., station locations and service areas) as well as for ML predictions of travel behavior as a basis for operational service improvements (e.g., vehicle availability). As part of the descriptive analysis, Fig. 4 exemplarily illustrates the distribution of trips started per access distance of the users as a basis for the analysis of user preferences for stations (i.e., nearest stations and most frequented stations). Thereafter, Fig. 5 shows the main results of the predictive modeling and XML approach. It provides an overview of the 10 explanatory variables with the strongest effect on the prediction performance of each model. The results represent the mean values of the root mean squared error (RMSE) loss function as the measure of permutation feature importance for 50 permutations. They illustrate the actual importance of the variables for the model predictions over the entire set of observations, allowing for comparability across models. Box plots are added to the bars to indicate the distribution of the values of the measure across the permutations.

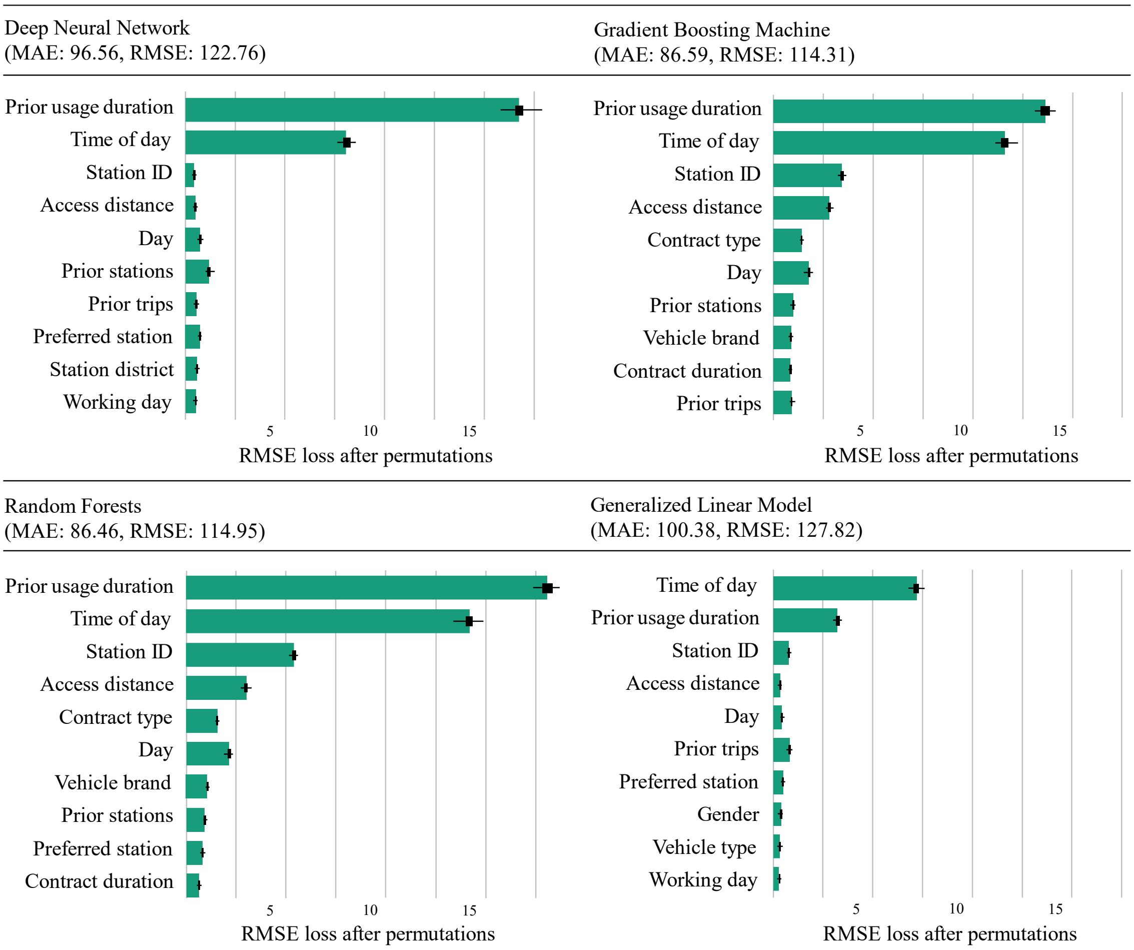

Fig. 5 Permutation feature importance plots

The findings show that the users' mean 'prior usage duration' and the 'time of day' have the strongest effect on the prediction accuracy of all models (i.e., RMSE losses up to 18 for the 'prior usage duration' and up to 14 for the 'time of day'). Taken together, variables related to the users' access to and preference of stations (i.e., 'station ID', 'access distance' and 'nearest station', 'preferred station', and 'prior stations') account for a combined average RMSE loss of 5.95 (or 4.96%). 'Access distance' alone has the fourth strongest (for the tree-based models) to eighth strongest (for the GLM) effect on model predictions (i.e., RMSE loss of 0.5 to 3). Fig. 6 zooms in on how model predictions behave as a function of the access distance using ALE plots. The analysis confirms that a longer ‘access distance’ from the users’ home to the station is related to a higher predicted ‘usage duration’ and that both tree-based models capture a similar non-linear relationship. The ALE plots indicate that a longer ‘access distance’ leads to an increase in the prediction of the ‘usage duration’ up to a certain threshold before the curve begins to flatten out.

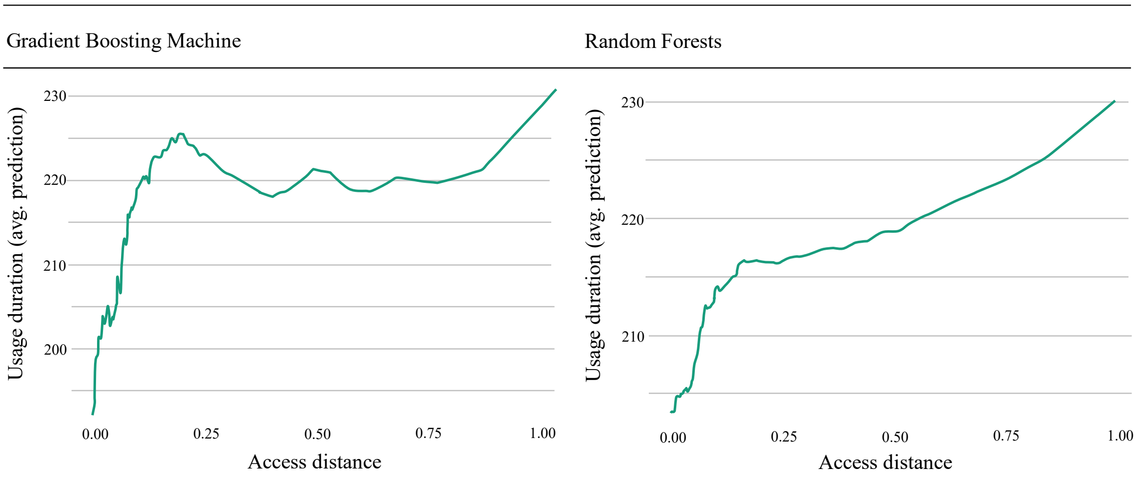

Fig. 6 ALE plots for the ‘access distance’

The development of carsharing programs in less densely populated areas requires a comprehensive data understanding as input for strategic decision-making and predictions of travel behavior. Research Paper P3 addresses this need by descriptively analyzing the role of access distance in carsharing user behavior and by presenting an XML approach to evaluate the influence of explanatory variables on ML predictions of the vehicle usage duration per trip. In this way, the study extents first descriptive research exploring carsharing demand, challenges, and user characteristics in non-urban environments (Illgen and Höck 2020; Leroy et al. 2023; Rotaris and Danielis 2018) and adds to research concerned with the increased need for optimization and decision support in less densely populated areas (Bruglieri and Pisacane 2023; Lagadic et al. 2019). The findings contribute to more accurate and transparent predictions of trip characteristics through a better understanding of the effect of usage factors as input variables to ML models. From a practical perspective, these insights can be leveraged to improve service quality by, inter alia, anticipating vehicle availability and supply imbalances, planning maintenance trips, or developing dynamic pricing schemes (Wang et al. 2021; Willing et al. 2017). While a thorough data understanding is indispensable for employing ML models in carsharing decision support, emerging carsharing programs still face the challenge of restricted access to high-volume usage data that is often limited by privacy restrictions, prohibitively expensive or time-consuming data collection, or lack of data quality and representativeness (Brendel et al. 2017; Gudivada et al. 2017; Lagadic et al. 2019).

Following up on this need for research, Research Paper P4 aims to explore the use of synthetic data to support carsharing decision-making by overcoming the barrier of limited data access during the introduction and expansion of new services. To this end, it means to investigate generative ML models to create synthetic tabular transaction data of carsharing trips for more accurate predictions of user behavior when data access is limited. Thus, this study raises the following research questions: *Can synthetic data created by generative ML models support decision-making in carsharing? And what are appropriate methods?*

To pursue this research objective, the study employs a systematic ML workflow (Kühl et al. 2021; Shrestha et al. 2021) as a rigorous approach to the generation of synthetic data as well as to the evaluation of their fidelity and utility in the context of real-world prediction tasks. Consistent with the application-oriented nature of this work, it considers the case of an emerging carsharing program that is expanding its services to include free-floating electric vehicles (EVs) and aims to obtain more reliable predictions of trip distances and usage times during the service's ramp-up. In this connection, the study analyzes how well synthetic data generated by conditional tabular generative adversarial networks (CTGANs), tabular variational autoencoders (TVAEs) (Lei et al. 2019), and benchmark models (i.e., synthetic minority oversampling technique-nominal continuous (SMOTE-NC) and Gaussian Copula (GC) models) serve the purpose of enhancing the available database to improve the training and performance of prediction models (i.e., NNs, RFs, and extreme gradient boosting (XGBoost)). The study investigates the use of synthetic data for replacing, rebalancing, and augmenting real training data drawing on two evaluation protocols that correspond to the prediction of two different target variables.

The results of the evaluation of the fidelity of the synthetic data sets (i.e., the extent to which they accurately represent the structural characteristics of the real data) indicate that all the models analyzed adequately captured the distribution of the real training data, with some limitations for the GC data. Minor deviations in data quality are indicated by both descriptive statistics and graphical representations. Fig. 7 presents an example of how the evaluation of data fidelity can be supported with graphical representations of the marginal distributions of selected explanatory variables for the real and synthetic data. Thereafter, Table 3 provides a comprehensive overview of the prediction performances (i.e., mean absolute error (MAE) and RMSE) in the two evaluation protocols for data utility. The results show that augmenting real training data with synthetic samples improves the performance of prediction models by up to 4.63% when predicting the usage time and distance of upcoming trips. The findings of the analysis further reveal that the quality of synthetic data (i.e., using generative ML models such as CTGAN and TVAE compared to statistical benchmark models) as well as the ratio of synthetic to real data are critical to maximizing data utility for ML prediction tasks.

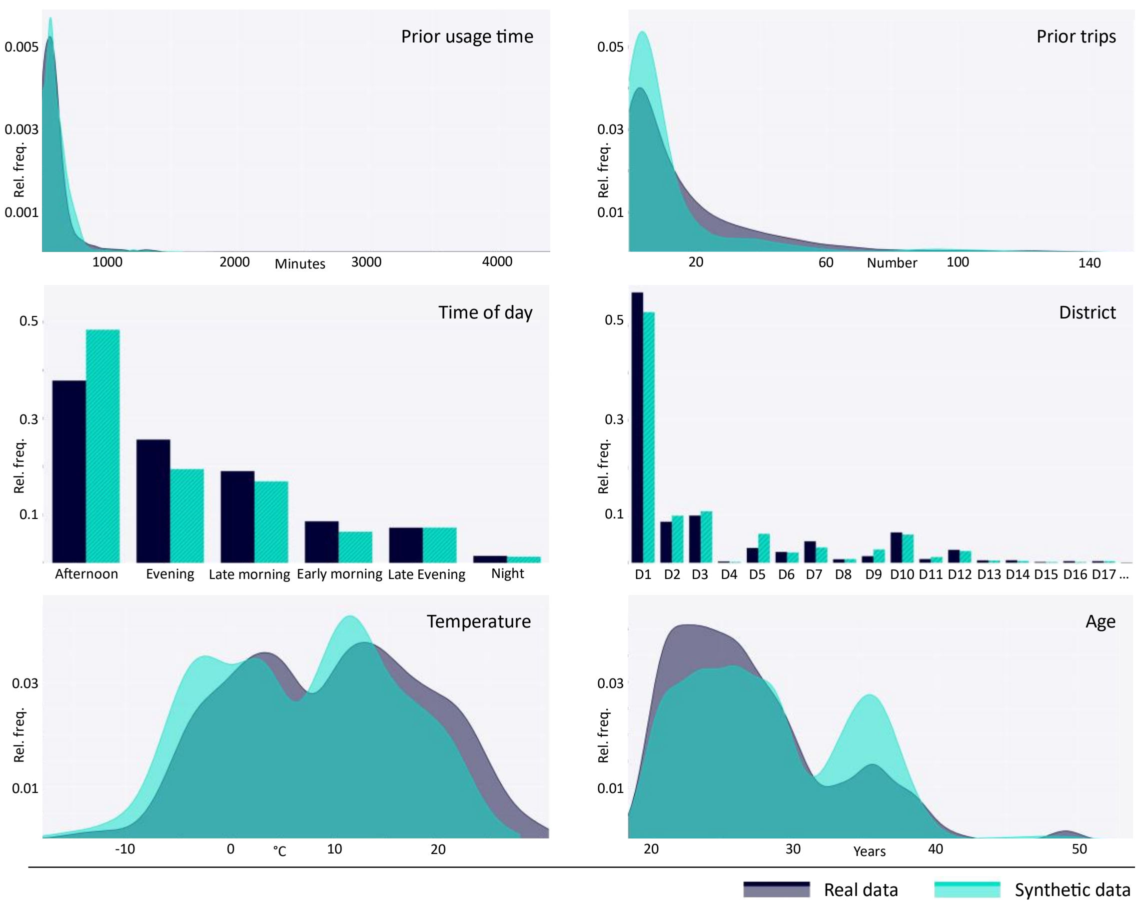

Fig. 7 Marginal distributions of selected explanatory variables (CTGAN)

Target variable	Evaluation mode	CTGAN data			TVAE data		
		NN	RF	XGBoost	NN	RF	XGBoost
Trip distance	replacing	11.3196	11.5520	11.4922	11.8184	11.3487	11.5896
		23.2087	22.6227	22.5921	22.0950	21.8802	21.8726
	rebalancing	11.3523	11.0855	11.1611	11.7991	11.2125	**11.2455**
		21.2471	21.4888	21.2945	21.1358	21.3523	21.0381
	augmenting	**11.0701**	**10.9784**	**10.8630**	**11.2406**	**11.1070**	11.4500
		22.3523	21.4341	22.7119	21.5538	21.2524	21.2246
Usage time	replacing	81.2940	81.9108	80.6387	81.4213	77.6522	**77.4637**
		128.0502	126.5895	127.8963	129.2339	124.1547	124.8255
	rebalancing	80.0625	78.8520	78.1554	81.7888	78.5501	78.5069
		119.9833	120.1922	119.5669	124.1220	118.5909	118.3571
	augmenting	**79.8870**	**77.9111**	**76.4161**	**80.7069**	**76.1038**	77.9207
		121.2118	118.2622	118.7290	123.9520	118.5167	118.8860

Target variable	Evaluation mode	GC data			SMOTE-NC data		
		NN	RF	XGBoost	NN	RF	XGBoost
Trip distance	replacing	12.3650	13.4651	13.4392	11.8364	11.5884	12.0987
		22.5965	22.8592	22.4730	21.8198	22.2567	22.2128
	rebalancing	11.8883	**11.2598**	11.3729	11.6874	11.2754	11.4637
		21.0460	21.3563	21.2242	21.3519	21.4206	21.1773
	augmenting	**11.6539**	11.3464	**11.1518**	**11.6143**	**11.2061**	**11.1693**
		21.6236	21.4064	21.3925	21.4031	21.2714	21.2983
Usage time	replacing	87.6569	93.1882	93.7253	84.2916	80.5134	81.8139
		128.5637	129.1983	129.4977	131.1395	123.2877	124.5412
	rebalancing	84.6442	**79.2452**	**78.6709**	82.6343	79.2314	**79.0084**
		122.7949	119.6962	119.2687	124.0749	119.7012	119.6470
	augmenting	**78.7016**	79.4261	81.0885	**84.0167**	**77.8619**	79.4226
		123.7068	119.0203	120.0561	129.6483	119.4801	120.4374

Target variable	Evaluation mode	Real data		
		NN	RF	XGBoost
Trip distance	replacing	**11.8347**	**11.2619**	11.3898
		20.9877	21.2480	21.0495
	rebalancing	12.7694	11.3134	**11.3543**
		21.3496	21.3229	21.3265
	augmenting	12.5649	11.2968	11.5540
		22.5950	21.3258	21.1755
Usage time	replacing	**82.1344**	79.0867	**78.6856**
		122.8217	119.5922	119.2367
	rebalancing	82.1817	**78.5650**	78.9757
		125.0647	118.2102	119.6662
	augmenting	87.4603	78.7227	80.3747
		130.4195	119.7439	119.4032

Table 3 Prediction accuracy for the synthetic training data sets and the real training data (MAE and RMSE)

The productive implementation of ML to support decision-making in carsharing requires comprehensive data availability. Research Paper P4 addresses this need by presenting a new application-oriented perspective on the use of generative ML in transportation research and stimulating further exploration of leveraging synthetic tabular mobility data. The work contributes to the body of knowledge by being the first to investigate generative ML algorithms able to create synthetic tabular data (i.e., CTGAN and TVAE) to overcome the barriers faced during the introduction and expansion of innovative mobility options, enabling more accurate ML predictions of trip characteristics. Thus, it presents a blueprint to fellow scholars for investigating the use of synthetic mobility data in a broader range of domains and advances previous research building on statistical or simulation approaches to generate synthetic data (Etxandi-Santolaya et al. 2023). From a practical perspective, this research presents a low-threshold way for carsharing operators to evaluate and deploy generative ML models for qualitatively and quantitatively enhancing their database. This helps operators to obtain more accurate prior insights on trip characteristics during the introduction or expansion of their services where data access is a common challenge (Brendel et al. 2017; Lagadic et al. 2019).

Overall, Research Paper P3 and Research Paper P4 presented two notable contributions to the carsharing domain that advance ML implementation at the project level. First, it was investigated how spatial access contributes to individual decision-making of carsharing users and ML predictions of user behavior. Second, the potential of generative ML for creating synthetic trip data was evaluated to improve ML predictions when data access is limited. Together, the two research papers provide guidance on how to employ ML models to solve domain-specific business problems related to *data understanding* and *data preparation* by evaluating and implementing suitable ML solutions.

III.2 Modeling and Evaluation

The research presented in Section III.1 addressed business problems in a relatively new application domain where many data-related issues of ML implementation remain to be solved. In more mature domains, the remaining challenges tend to shift toward methodological advances and practicability (Loureiro et al. 2021). This is the case in customer service, where the advent of ML-based decision support has fundamentally changed the way organizations interact with their customers. (Kaplan and Haenlein 2019). Based on latent characteristics and previous customer behavior, ML models can predict future customer interactions (Wedel and Kannan 2016). One of the most prevalent and data-intensive touchpoints of many service-oriented organizations with their customers are call centers or contact centers, making them a leading-edge ML application domain in CRM (Barrow and Kourentzes 2018; Whiting and Donthu 2006). To constantly provide high service quality and short waiting times at this point, a sufficient number of agents is needed (Atlason et al. 2008). Consequently, predicting the volume of incoming customer inquiries and deciding on the required staffing level is essential. In this connection, ML models promise more flexible and precise predictions and, thus, the possibility of enhanced organizational planning and better customer service. Despite the enormous potential for service improvements and cost savings, a comprehensive understanding of how ML models can add value to contact center forecasting is lacking (Barrow 2016; Ebadi Jalal et al. 2016). In order to gain deeper insights into the performance and practicability of ML models in this context, research taking a methodological perspective and providing an application-oriented angle on model evaluation and selection is needed.

Research Paper P5 contributes to closing this research gap by proposing a new method for call center arrival forecasting capable of both capturing the dynamics of the time series and including predictor variables. To this end, this study extends the established dynamic harmonic regression (DHR) model, which utilizes a sum of sinusoidal terms (i.e., Fourier terms) as predictors to handle periodic seasonality and an autoregressive integrated moving average (ARIMA) error to capture short-term dynamics, by including predictor variables in the considered information space to generate predictions. The study evaluates the predictive potential of the approach using two different real-world call and e-mail arrival series of a leading German online retailer comprising 174 weeks of data. The proposed model is compared to different established time series and ML approaches (e.g., RF and gradient boosting with regularization (GBR)) with time series cross-validation and an expanding rolling window.

Tables 4 and 5 provide insight into the main findings of the study. Comparing the results for the prediction of call arrivals, the proposed DHR model with predictor variables outperforms the remaining models with respect to its MAE results for every considered lead time. Considering 2 and 3 weeks of lead time, RF performs slightly better than the DHR model when considering RMSE as an evaluation metric, highlighting the strong influence of contextual information on prediction performance.

	No lead time	One week	Two weeks	Three weeks
DHR with predictor variables	**11.2336**	**11.2250**	**11.5571**	**11.4648**
DHR	14.3198	14.2243	14.3913	14.5242
STL+ARIMA	17.8254	17.3990	17.7603	18.2596
STL+ETS	17.8028	17.4140	17.8213	18.4370
STL+RWDRIFT	17.7657	17.5317	17.7757	18.2407
TBATS	16.3675	17.1001	17.2664	17.6892
GBR	13.4358	13.7363	13.7492	13.5698
RF	12.0249	12.0516	12.3048	12.3006

The highest forecast accuracy for each lead time is marked in bold.
DHR: Dynamic harmonic regression; GBR: Gradient boosting with regularization; RF: Random forest;
STL + ARIMA: Autoregressive integrated moving average with time series decomposition based on Loess;
STL + ETS: Innovation state space model with time series decomposition based on Loess; STL + RWDRIFT:
Random walk with drift with time series decomposition based on Loess; TBATS: Trigonometric seasonality,
Box-Cox transformation, ARMA errors, trend and seasonal components.

Table 4 MAE results for customer support call arrival forecasts

	No lead time	One week	Two weeks	Three weeks
DHR with predictor variables	**15.7521**	**16.6844**	17.0088	17.1805
DHR	22.4746	22.9245	23.1713	23.4540
STL+ARIMA	30.3195	29.0067	29.3980	30.2608
STL+ETS	30.1632	28.9110	29.4276	30.4809
STL+RWDRIFT	30.1719	28.9619	29.1190	29.8225
TBATS	24.8822	26.0952	26.3939	27.0433
GBR	18.6523	19.2414	19.0341	18.7706
RF	16.6278	16.7476	**16.9844**	**16.9797**

Note: The highest forecast accuracy for each lead time is shown in bold.
DHR: Dynamic harmonic regression, GBR: Gradient boosting with regularization; RF: Random forest;
STL + ARIMA: Autoregressive integrated moving average with time series decomposition based on Loess;
STL + ETS = Innovation state space model with time series decomposition based on Loess; STL + RWDRIFT:
Random walk with drift with time series decomposition based on Loess; TBATS: Trigonometric seasonality,
Box-Cox transformation, ARMA errors, trend, and seasonal components.

Table 5 RMSE results for customer support call arrival forecasts

Similar to the call arrivals analysis, the proposed DHR model with predictor variables outperforms the remaining time series and ML approaches (i.e., MAE and RMSE) in every considered lead time constellation for the customer support e-mail arrival forecasts. However, the performance gap between time series and ML models is not as evident as for the prediction of call arrivals. With a stronger trend in the examined data, time series models operate in a comparable performance range as ML models. As expected, the prediction accuracy mostly decreases with greater lead times while the best performance is obtained without lead time.

This work contributes to the literature by connecting methodological strands of research and introducing a novel method to contact center forecasting. By capturing time series dynamics and including contextual information, the proposed DHR-based approach can increase the accuracy of call and e-mail arrival forecasts compared to established time series models and other ML approaches used in previous research (Andrews and Cunningham 1995; Hyndman et al. 2002; Ibrahim et al. 2016; Rausch and Albrecht 2020). In doing so, it supports practitioners in making more informed staffing decisions. Thus, organizations benefit from the presented approach as they can avoid overstaffing (i.e., keeping operating costs at a minimum) as well as understaffing agents (i.e., maximizing perceived service quality by reducing waiting times). Besides, the results suggest that the individual data environment (e.g., the magnitude of trend in the time series and availability of predictor variables providing contextual information) should be considered during model selection in contact center forecasting. Thus, additional research providing practical guidance on evaluating and selecting suitable ML models for individual business settings is required.

Research Paper P5 addresses this research need by proposing a two-step approach that provides a thorough understanding of the prediction accuracy of ML methods in call arrival forecasting and makes the underlying process of method evaluation and selection feasible for decision-makers in practice. Specifically, this study conducts an in-depth analysis of the forecast accuracy of suitable ML models based on the call arrival data of a real German online retailer. Using two different datasets, i.e., the customer support queue and the customer complaints queue of the corresponding call center, it performs a comprehensive method comparison opposing selected ML models (i.e., gradient boosting with dropout (GBD), GBR, k-nearest neighbor (KNN), RF, and support vector regression (SVR)) to the three most commonly used time series models in this field (i.e., ARIMA, random walk with drift (RWDRIFT), and error, trend, seasonal (innovation state space model) (ETS)). In the second step, this work provides a hands-on walk-through example of the model selection process based on cross-validation with an expanding rolling window. The practical implementation of the process is then illustrated in a programming environment that is accessible to non-ML experts and practitioners.

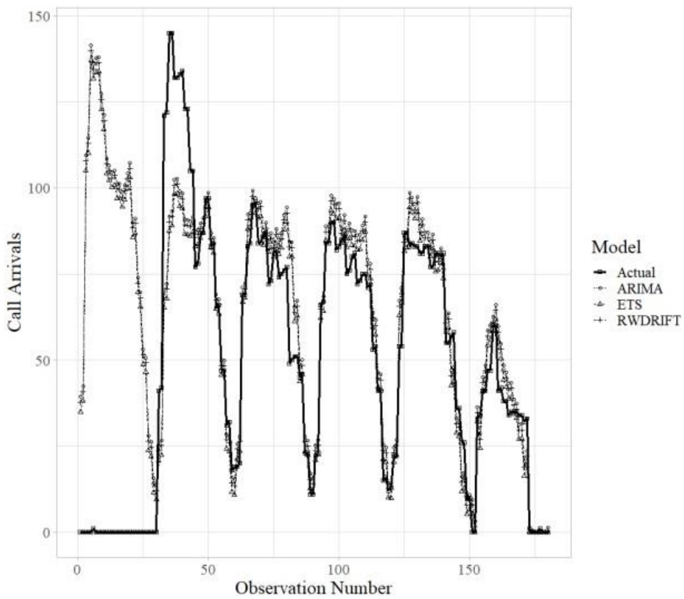

Fig. 8 Last predicted week of the time series models

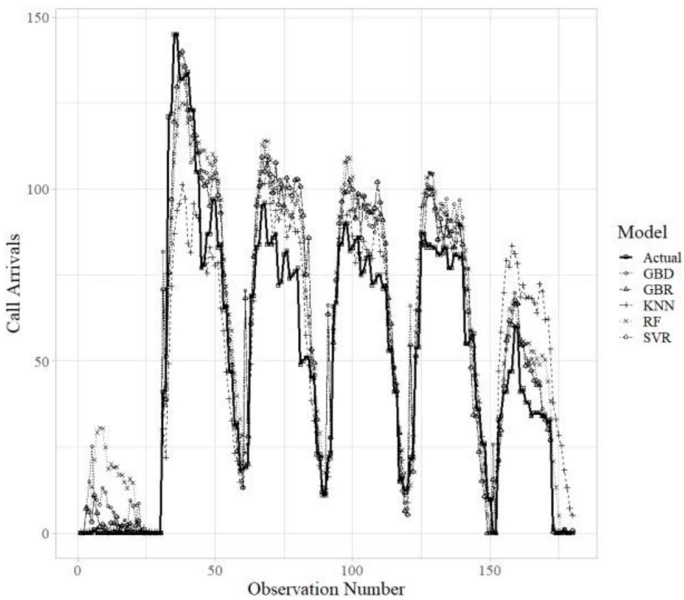

Fig. 9 Last predicted week of the ML models

The results of the model comparison for the customer support queue show that the RF algorithm outperforms the remaining approaches in every lead time constellation: with respect to both MAE and RMSE, the model yields the most accurate forecasts. The GBD, GBR, and SVR models yield comparable results, whereas the KNN approach was the most inaccurate forecasting method. Generally, the considered ML approaches are superior to the benchmark time series models for all lead time constellations (except for the KNN method). Among the time series models, the ETS model is the best-performing approach. In addition, the RF algorithm is found to provide an acceptable trade-off between accuracy and computation time: for the prediction of one iteration k (i.e., of the forecast horizon h = 180 observations), the model takes 24.1 min (using 40 GB RAM). Checking the robustness of these results based on the queue for customer complaints call arrivals, the RF yields the most accurate forecasts compared to the remaining approaches for all considered lead times except for the MAE result with two weeks lead time for which GBR is found to be superior. Figs. 8 and 9 provide more detailed insights into model performances for the last week predicted (i.e., 180 observations) for the customer support queue.

Based on the results of the conducted model comparison, organizations are suspected of benefiting from including ML approaches in their process of evaluating and selecting the most suitable method for forecasting call center arrivals. To make the underlying process of method comparison and selection accessible to decision-makers in practice and to overcome its perceived complexity, this study develops a walk-through tutorial based on cross-validation with an expanding rolling window in the programming environment R. In doing so, it proposes to view method selection in call center forecasting as the overall issue of implementing a forecasting system that includes prediction accuracy as well as practicability.

Fig. 10 illustrates a generic for-loop for the expanding rolling window that can be utilized to identify the most accurate model. Let n be the n^{th} observation (i.e., row) of the dataset, m be the m^{th} variable (i.e., column) of the dataset, and h be the forecast horizon. After analyzing and preprocessing the data, an empty numeric vector is defined, in which the results are stored during the loop. The for-loop itself can be iterated k times: let the forecast horizon h be one week and out-of-sample predictions with cross-validation shall be generated for one year, then k = 52, i.e., 52 weeks. For each iteration k = 1, 2,…, K during the loop, the training and test subsets are defined which roll forward for one unit of the forecast horizon h, i.e., $i * h$. Since $1 * h$ observations are added during the first iteration for syntax reasons, h observations are subtracted from the training and test subsets ($ntrain$ and $ntest$ respectively) to yield the intended initial training and test subsets. After the loop finishes, the loop time is reported to track the models' computation time as a crucial aspect for decision-makers in practice. The MAE and RMSE are both calculated by inserting the vector of actual values as the first argument and the vector of predicted values as the second argument. To test a specific model's predictive ability, it can simply be integrated into the generic for-loop.

```
results <- vector("numeric")
tic("Looptime")
for(i in 1:k){
   train_subset <- data[1:((n_train − h)+(i*h)),]
   test_subset <- data[((n_test − h)+(i*h)):(((n_test − h) − 1)+((i+1)*h)),]
   ## Insert Model here
}
toc(log = TRUE)
timelog <- tic.log(format = TRUE)
results <- pmax(results,0)

mae(data[n_test: (n_test + k * h), m_calls], results)
rmse(data[n_test: (n_test + k * h), m_calls], results)
```

Fig. 10 R Code for rolling expanding window with generic for-loop.
Note: The bold variables have to be replaced depending on the specific dataset.

Responding to the call for practical guidance on the evaluation and selection of suitable ML models for call arrival forecasting in contact centers, this study provides an extensive comparison of suitable ML models and makes the underlying context-agnostic modeling process feasible for practice. Thus, this research presents a starting point for shifting traditional contact center forecasting literature toward a new paradigm drawing on ML methods. It underpins the preliminary findings of previous studies (Barrow 2016; Ebadi Jalal et al. 2016; Rausch and Albrecht 2020) drawing on individual ML models such as NNs and RFs that indicate their high potential in contact center forecasting. In addition, this study is the first to comprehensively include model practicability by simultaneously investigating different lead times (i.e., three weeks, two weeks, one week, and no lead time) as well as the trade-off between complexity (i.e., estimation time and computation effort) and prediction accuracy.

In sum, Research Paper P5 and Research Paper P6 presented two significant contributions to contact center forecasting in CRM that advance ML implementation at the project level. First, a novel method in contact center forecasting was proposed, before a hands-on evaluation approach for ML models in this domain was introduced (i.e., *modeling* and *evaluation*). Together with the contributions presented in Section III.1, this research increases the understanding of how to approach the phases of ML implementation in the context of diverse industry-specific socio-technical problems.

IV Conclusion

IV.1 Summary

At the core of present-day AI, ML offers tremendous business potential that is reflected by large-scale investments in industry and research. Although knowledge in the field of ML has considerably matured, scholars and professionals alike continue to seek a better understanding of the underlying concepts, processes, and challenges of implementing the technology. In this regard, understanding and orchestrating the socio-technical process of ML implementation in business is one of the most critical challenges. Over the past years, ML workflows and lifecycle models have provided standardized procedures with considerable technical depth. Yet, theoretical guidance and actionable insights that facilitate an integrated approach to process-oriented ML are lacking. Consequently, more guidelines from theory on how to approach different ML methods in the context of diverse industry-specific socio-technical problems are required in all phases of ML implementation (i.e., *data understanding*, *data preparation*, *modeling*, and *evaluation*). In addition, comprehensive support on how to ensure alignment of ML opportunities and business processes (i.e., *process fit*) at the outset and how to scale novel ML-enabled processes (i.e., *process ramp-up*) after deployment is vital.

This dissertation and the embedded research articles contribute to the identified gaps and aim to improve the understanding of the socio-technical process of ML implementation in business in two ways: First, this thesis zooms out from a (single-)project perspective to the process level to stimulate an integrated approach to process-oriented ML implementation. Thereby, it introduces the notion of *process fit* to the *business understanding* phase of ML implementation and augments the *deployment* phase with *process ramp-up.* Second, it transfers theoretical ML knowledge to applications that solve (domain-)specific problems at the project level, with the goal of supporting practitioners in applying ML methods across all phases of ML implementation. At the same time, it aims to expand the ML knowledge base by gaining theoretical insights from the practical application of advanced ML methods in leading-edge domains.

Section II presents two research papers that advance ML implementation at the process level by providing guidance on how to examine ML opportunities and business processes in an integrated way (i.e., *process fit*) and how to scale novel ML-enabled processes (i.e., *process ramp-up*). Research Paper P1 introduces a methodological blueprint for investigating the opportunities of leading-edge digital technologies such as ML through the theoretical lens of affordance theory. The study not only identifies affordances in a frontline IS application domain with high ML potential but also presents insights on how to rigorously develop a fine-grained understanding of ML action potential in different organizational contexts. Research Paper P2 introduces BPRUM as a new BPM capability area. It presents hands-on knowledge and managerial tools for the effective implementation and scaling of novel ML-driven business processes through action-oriented sub-capabilities of BPRUM. Together, both research papers help organizations adopt practices that facilitate an integrated approach to process-oriented ML implementation.

Section III includes four research papers that advance ML implementation at the project level by contributing theoretical guidance and actionable insights on how to approach ML implementation in the context of diverse industry-specific socio-technical problems. Therefore, it focuses on two leading-edge application domains, namely mobility (i.e., carsharing) and CRM (i.e., customer service). Research Paper P3 analyzes the role of spatial access in the individual decision-making of carsharing users in small urban areas and presents an XML approach to reveal its influence on ML predictions of user behavior (i.e., *data understanding*). Thereafter, Research Paper P4 delves into generative ML for creating synthetic trip data to support carsharing decision-making by overcoming the barrier of limited data access during the introduction and expansion of new services (i.e., *data preparation*). It introduces a blueprint for investigating the use of synthetic mobility data in a broader range of domains. Lastly, Research Paper P5 and Research Paper P6 examine contact center forecasting. The studies propose a novel forecasting method including contextual information and present an evaluation approach for ML models in this domain (i.e., *modeling* and *evaluation*). Collectively, the four contributions address domain-specific business problems to derive theoretical guidelines and actionable practical insights.

Motivated by the need to sustainably realize ML implementation in business, the overall purpose of this dissertation is to guide scholars and practitioners in understanding and performing the operational, methodological, and managerial practices associated with the intricate socio-technical process of effective ML implementation. This thesis strives to contribute to this area of research at both a process and a project level.

IV.2 Limitations and Future Research

The results of this thesis need to be reflected against limitations that also stimulate future research. While the individual limitations of each research paper can be found in the respective articles (see Sections VI.3 to VI.8), this section provides an aggregated overview of the limitations of this thesis. Thus, the following presents limitations and avenues for future research on the socio-technical process of ML implementation.

First, the results of the research papers advancing ML implementation at the process level (Section II) were established inductively based on systematic literature reviews and qualitative interview studies. Accordingly, some limitations are inherent in the nature of these methodological approaches. To address the potential shortcomings of restricted literature samples and interview expert selection, the development of the results was complemented by extensive evaluations of the completeness, consistency, and applicability of the findings with scholars and practitioners. Nonetheless, future research could expand the scope of data collection to evaluate the presented results in a broader organizational context and in new application domains. In contrast, the technological focus of Research Papers P1 and P2 should be further narrowed down to the specifics of ML. While the theoretical frameworks used in the present studies (i.e., affordance theory and capability frameworks) only constitute one possible way of investigating the process-related opportunities afforded by ML and the ramp-up of ML-driven business processes, the results provide a good starting point for future research to establish an integrated approach to process-oriented ML implementation. From a practical perspective, it is important to note that both affordances and capabilities do not directly generate organizational benefits, but rather serve as a foundation for planning ML initiatives and creating custom capability development plans.

Second, the research papers advancing ML implementation at the project level (Section III) investigate two selected leading-edge application domains (i.e., mobility and CRM). While the presented studies (i.e., Research Papers P3, P4, P5, and P6) lay a strong focus on deriving domain-agnostic methodological guidelines for the addressed phases of ML implementation, the findings need to be validated for other use cases and business environments to account for potential contextual constraints. In addition, the research papers focus on a limited range of ML and benchmark models that were selected based on individual requirements related to the investigated use cases (e.g., accuracy, complexity, and explainability). Consequently, fellow scholars are encouraged to expand the methodological scope of the present research to a broader spectrum of models. As ML is characterized by rapid technological advances, the insights derived in this thesis will likely require repeated updates to reflect the latest developments in the field. Nevertheless, the application-oriented approaches and real-world data included in the studies can be considered a strength of this work, and the selected research designs are supposed to serve as blueprints to address the above limitations in future research.

Third, this thesis proposes a process and project perspective on ML implementation in business, the need for which has been stimulated by previous research and confirmed by the results of this work. As such, it complements existing technical studies (e.g., ML workflows and lifecycle models) and management contributions (e.g., on organizational readiness for ML, ML adoption, and ML capabilities). However, the presented results are expected to overlap with findings from these related disciplines. Accordingly, future research is encouraged to further advance knowledge synthesis between the different research angles on ML implementation. Finally, given the relation of ML to AI and data science, this thesis does not claim the exclusivity of all presented insights to ML. Instead, it aims to encourage to learn from related disciplines and to derive specific interpretations in the context of ML.

Notwithstanding the above limitations, I am confident that this dissertation contributes to the current body of knowledge and will guide scholars and practitioners in understanding and performing the operational, methodological, and managerial practices associated with the intricate socio-technical process of effective ML implementation in business.

V References

Abdel-Karim, B. M., Pfeuffer, N., and Hinz, O. 2021. "Machine learning in information systems - a bibliographic review and open research issues," *Electronic Markets* (31:3), pp. 643-670 (doi: 10.1007/s12525-021-00459-2).

Allen, L., O'Connell, A., and Kiermer, V. 2019. "How Can We Ensure Visibility and Diversity in Research Contributions? How the Contributor Role Taxonomy (CRediT) Is Helping the Shift from Authorship to Contributorship," *Learned Publishing* (32:1), pp. 71–74 (doi: 10.1002/leap.1210).

Amatuni, L., Ottelin, J., Steubing, B., and Mogollón, J. M. 2020. "Does car sharing reduce greenhouse gas emissions? Assessing the modal shift and lifetime shift rebound effects from a life cycle perspective," *Journal of Cleaner Production* (266), p. 121869 (doi: 10.1016/j.jclepro.2020.121869).

Andrews, B. H., and Cunningham, S. M. 1995. "L. L. Bean Improves Call-Center Forecasting," *Interfaces* (25:6), pp. 1-13 (doi: 10.1287/inte.25.6.1).

Apley, D. W., and Zhu, J. 2020. "Visualizing the Effects of Predictor Variables in Black Box Supervised Learning Models," *Journal of the Royal Statistical Society Series B: Statistical Methodology* (82:4), pp. 1059-1086 (doi: 10.1111/rssb.12377).

Ashmore, R., Calinescu, R., and Paterson, C. 2022. "Assuring the Machine Learning Lifecycle," *ACM Computing Surveys* (54:5), pp. 1-39 (doi: 10.1145/3453444).

Åström, J., Reim, W., and Parida, V. 2022. "Value creation and value capture for AI business model innovation: a three-phase process framework," *Review of Managerial Science* (16:7), pp. 2111-2133 (doi: 10.1007/s11846-022-00521-z).

Atlason, J., Epelman, M. A., and Henderson, S. G. 2008. "Optimizing Call Center Staffing Using Simulation and Analytic Center Cutting-Plane Methods," *Management Science* (54:2), pp. 295-309 (doi: 10.1287/mnsc.1070.0774).

Baier, L., Jöhren, F., and Seebacher, S. 2019. "Challenges in the Deployment and Operation of Machine Learning in Practice," *Proceedings of the 27th European Conference on Information Systems (ECIS)*, pp. 8-14.

Baier, M.-S., Berger, S., Kreuzer, T., Oberländer, A. M., and Röglinger, M. 2023. "What makes Digital Technology? A Categorization based on Purpose," *Communications of the Association for Information Systems* (52), pp. 556-609.

Bandara, W., van Looy, A., Rosemann, M., and Meyers, L. 2021. "A call for "holistic" business process management," *Proceedings of the International Workshop on BPM Problems to Solve Before We Die (PROBLEMS 2021)*, pp. 6-10.

Barrow, D., and Kourentzes, N. 2018. "The impact of special days in call arrivals forecasting: A neural network approach to modelling special days," *European Journal of Operational Research* (264:3), pp. 967-977 (doi: 10.1016/j.ejor.2016.07.015).

Barrow, D. K. 2016. "Forecasting intraday call arrivals using the seasonal moving average method," *Journal of Business Research* (69:12), pp. 6088-6096 (doi: 10.1016/j.jbusres.2016.06.016).

Baumbach, S., Oberländer, A. M., Röglinger, M., and Rosemann, M. 2020. "Dynamic capabilities for opportunity exploration: insights from an explorative case study," *International Journal of Entrepreneurial Venturing* (12:6), pp. 575-616 (doi: 10.1504/ijev.2020.10031152).

Baumgarte, F., Keller, R., Röhrich, F., Valett, L., and Zinsbacher, D. 2022. "Revealing influences on carsharing users' trip distance in small urban areas," *Transportation Research Part D: Transport and Environment* (105), p. 103252 (doi: 10.1016/j.trd.2022.103252).

Benbya, H., Davenport, T., and Pachidi, S. 2020. "Artificial intelligence in organizations: current state and future opportunities," *MIS Quarterly Executive* (doi: 10.17863/CAM.63213).

Berente, N., Gu, B., Recker, J., and Santhanam, R. 2021. "Managing Artificial Intelligence," *MIS Quarterly* (45:3), pp. 1433-1450 (doi: 10.25300/MISQ/2021/16274).

Bettis, R. A., Gambardella, A., Helfat, C., and Mitchell, W. 2015. "Qualitative empirical research in strategic management," *Strategic Management Journal* (36:5), pp. 637-639 (doi: 10.1002/smj.2317).

Bishop, C. M. 2006. *Pattern recognition and machine learning*, New Delhi, India, India: Springer (India) Private Limited; Rashtriya Printers.

Borges, A. F., Laurindo, F. J., Spínola, M. M., Gonçalves, R. F., and Mattos, C. A. 2021. "The strategic use of artificial intelligence in the digital era: Systematic literature review and future research directions," *International Journal of Information Management* (57), p. 102225 (doi: 10.1016/j.ijinfomgt.2020.102225).

Brendel, A. B., Rockenkamm, C., and Kolbe, L. M. 2017. "Generating Rental Data for Car Sharing Relocation Simulations on the Example of Station-Based One-Way Car Sharing," *Proceedings of the 50th Hawaii International Conference on System Sciences*, pp. 1554-1563.

Bruglieri, M., and Pisacane, O. 2023. "A survey on emergent trends in the optimization of car-sharing systems," *International Transactions in Operational Research* (30:6), pp. 2867-2908 (doi: 10.1111/itor.13315).

Brynjolfsson, E., and McAfee, A. 2017. "Artificial Intelligence, for real," *Harvard Business Review* (1), pp. 1-31.

Brynjolfsson, E., and Mitchell, T. 2017. "What can machine learning do? Workforce implications," *Science (New York, N.Y.)* (358:6370), pp. 1530-1534 (doi: 10.1126/science.aap8062).

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., and Wirth, R. 2000. "CRISP-DM 1.0: Step-by-step data mining guide," *SPSS inc* (9:13), pp. 1-73.

Chen, T. D., and Kockelman, K. M. 2016. "Carsharing's life-cycle impacts on energy use and greenhouse gas emissions," *Transportation Research Part D: Transport and Environment* (47), pp. 276-284 (doi: 10.1016/j.trd.2016.05.012).

Cheng, J., Chen, X., Ye, J., and Shan, X. 2021. "Flow-based unit is better: exploring factors affecting mid-term OD demand of station-based one-way electric carsharing," *Transportation Research Part D: Transport and Environment* (98), p. 102954 (doi: 10.1016/j.trd.2021.102954).

Chicco, A., and Diana, M. 2021. “Air emissions impacts of modal diversion patterns induced by one-way car sharing: A case study from the city of Turin,” *Transportation Research Part D: Transport and Environment* (91), p. 102685 (doi: 10.1016/j.trd.2020.102685).

Corbin, J. M., and Strauss, A. 1990. “Grounded theory research: Procedures, canons, and evaluative criteria,” *Qualitative Sociology* (13:1), pp. 3-21 (doi: 10.1007/BF00988593).

de Bruin, T., and Rosemann, M. 2007. “Using the Delphi Technique to Identify BPM Capability Areas,” *ACIS 2007 Proceedings* (42).

Deloitte. 2017. “The 2017 Deloitte State of Cognitive Survey,”

Dumas, M., La Rosa, M., Mendling, J., and Reijers, H. A. 2018. *Fundamentals of Business Process Management*, Berlin, Heidelberg: Springer Berlin Heidelberg; Imprint: Springer.

Ebadi Jalal, M., Hosseini, M., and Karlsson, S. 2016. “Forecasting incoming call volumes in call centers with recurrent Neural Networks,” *Journal of Business Research* (69:11), pp. 4811-4814 (doi: 10.1016/j.jbusres.2016.04.035).

Enholm, I. M., Papagiannidis, E., Mikalef, P., and Krogstie, J. 2022. “Artificial Intelligence and Business Value: a Literature Review,” *Information Systems Frontiers* (24:5), pp. 1709-1734 (doi: 10.1007/s10796-021-10186-w).

Etxandi-Santolaya, M., Canals Casals, L., and Corchero, C. 2023. “Estimation of electric vehicle battery capacity requirements based on synthetic cycles,” *Transportation Research Part D: Transport and Environment* (114), p. 103545 (doi: 10.1016/j.trd.2022.103545).

Faulkner, P., and Runde, J. 2019. “Theorizing the Digital Object,” (doi: 10.25300/MISQ/2019/13136).

Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. 1996. “The KDD process for extracting useful knowledge from volumes of data,” *Communications of the ACM* (39:11), pp. 27-34 (doi: 10.1145/240455.240464).

Fechner, P., König, F., Kratsch, W., Lockl, J., and Röglinger, M. 2023. “Near-Infrared Spectroscopy for Bladder Monitoring: A Machine Learning Approach,” *ACM Transactions on Management Information Systems* (14:2), pp. 1-23 (doi: 10.1145/3563779).

Ferrero, F., Perboli, G., Rosano, M., and Vesco, A. 2018. “Car-sharing services: An annotated review,” *Sustainable Cities and Society* (37), pp. 501-518 (doi: 10.1016/j.scs.2017.09.020).

Fisher, A., Rudin, C., and Dominici, F. 2019. “All Models are Wrong, but Many are Useful: Learning a Variable’s Importance by Studying an Entire Class of Prediction Models Simultaneously,” *Journal of Machine Learning Research* (20:177), pp. 1-81.

Frank, A. G., Dalenogare, L. S., and Ayala, N. F. 2019. “Industry 4.0 technologies: Implementation patterns in manufacturing companies,” *International Journal of Production Economics* (210), pp. 15-26 (doi: 10.1016/j.ijpe.2019.01.004).

Gao, L., Lu, P., and Ren, Y. 2021. “A deep learning approach for imbalanced crash data in predicting highway-rail grade crossings accidents,” *Reliability Engineering & System Safety* (216), p. 108019 (doi: 10.1016/j.ress.2021.108019).

Gibson, J. J. 1986. "The Ecological Approach to Visual Perception,"

Golalikhani, M., Oliveira, B. B., Carravilla, M. A., Oliveira, J. F., and Antunes, A. P. 2021. "Carsharing: A review of academic literature and business practices toward an integrated decision-support framework," *Transportation Research Part E: Logistics and Transportation Review* (149), p. 102280 (doi: 10.1016/j.tre.2021.102280).

Goldkuhl, G. 2012. "Pragmatism vs interpretivism in qualitative information systems research," *European Journal of Information Systems* (21:2), pp. 135-146 (doi: 10.1057/ejis.2011.54).

Goodfellow, I., Bengio, Y., and Courville, A. 2016. *Deep learning*, Cambridge, MA: MIT press.

Google. 2023. "Vertex AI and the machine learning (ML) workflow," available at https://cloud.google.com/vertex-ai/docs/start/introduction-unified-platform#ml-workflow, accessed on Aug 8 2023.

Grisold, T., Groß, S., Stelzl, K., vom Brocke, J., Mendling, J., Röglinger, M., and Rosemann, M. 2021. "The Five Diamond Method for Explorative Business Process Management," *Business & Information Systems Engineering*, pp. 1-18 (doi: 10.1007/s12599-021-00703-1).

Gross, S., Malinova, M., and Mendling, J. 2019. "Navigating through the maze of business process change methods," *Proceedings of the 52nd hawaii international conference on system sciences*.

Gross, S., Stelzl, K., Grisold, T., Mendling, J., Röglinger, M., and vom Brocke, J. 2021. "The Business Process Design Space for exploring process redesign alternatives," *Business Process Management Journal* (27:8), pp. 25-56 (doi: 10.1108/BPMJ-03-2020-0116).

Gudivada, V., Apon, A., and Ding, J. 2017. "Data Quality Considerations for Big Data and Machine Learning: Going Beyond Data Cleaning and Transformations," *International Journal on Advances in Software* (10:1), pp. 1-20.

Gupta, S., Drave, V. A., Bag, S., and Luo, Z. 2019. "Leveraging Smart Supply Chain and Information System Agility for Supply Chain Flexibility," *Information Systems Frontiers* (21:3), pp. 547-564 (doi: 10.1007/s10796-019-09901-5).

Haefner, N., Wincent, J., Parida, V., and Gassmann, O. 2021. "Artificial intelligence and innovation management: A review, framework, and research agenda☆," *Technological Forecasting and Social Change* (162), p. 120392 (doi: 10.1016/j.techfore.2020.120392).

Hofmann, P., Oesterle, S., Rust, P., and Urbach, N. 2019. "Machine Learning approaches along the Radiology Value Chain : Rethinking Value Propositions," *Proceedings of the 27th European Conference on Information Systems (ECIS)*.

Hyndman, R. J., Koehler, A. B., Snyder, R. D., and Grose, S. 2002. "A state space framework for automatic forecasting using exponential smoothing methods," *International Journal of Forecasting* (18:3), pp. 439-454 (doi: 10.1016/s0169-2070(01)00110-8).

Ibrahim, R., Ye, H., L'Ecuyer, P., and Shen, H. 2016. "Modeling and forecasting call center arrivals: A literature survey and a case study," *International Journal of Forecasting* (32:3), pp. 865-874 (doi: 10.1016/j.ijforecast.2015.11.012).

Illgen, S., and Höck, M. 2020. “Establishing car sharing services in rural areas: a simulation-based fleet operations analysis,” *Transportation* (47:2), pp. 811-826 (doi: 10.1007/s11116-018-9920-5).

International Data Corporation. 2023. “Worldwide Spending on AI-Centric Systems Forecast to Reach $154 Billion in 2023, According to IDC,” available at https://www.idc.com/getdoc.jsp?containerId=prUS50454123#:~:text=NEEDHAM%2C%20Mass.%2C%20March%207,in%202023%2C%20an%20increase%20of, accessed on Aug 1 2023.

Islam, A. K. M. N., Laato, S., Talukder, S., and Sutinen, E. 2020. “Misinformation sharing and social media fatigue during COVID-19: An affordance and cognitive load perspective,” *Technological Forecasting and Social Change* (159), p. 120201 (doi: 10.1016/j.techfore.2020.120201).

Janiesch, C., Zschech, P., and Heinrich, K. 2021. “Machine learning and deep learning,” *Electronic Markets* (31:3), pp. 685-695 (doi: 10.1007/s12525-021-00475-2).

Jöhnk, J., Weißert, M., and Wyrtki, K. 2021. “Ready or Not, AI Comes— An Interview Study of Organizational AI Readiness Factors,” *Business & Information Systems Engineering* (63:1), pp. 5-20 (doi: 10.1007/s12599-020-00676-7).

Jordan, M. I., and Mitchell, T. M. 2015. “Machine learning: Trends, perspectives, and prospects,” *Science (New York, N.Y.)* (349:6245), pp. 255-260 (doi: 10.1126/science.aaa8415).

Kaplan, A., and Haenlein, M. 2019. “Siri, Siri, in my hand: Who’s the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence,” *Business Horizons* (62:1), pp. 15-25 (doi: 10.1016/j.bushor.2018.08.004).

Kerpedzhiev, G. D., König, U. M., Röglinger, M., and Rosemann, M. 2021. “An Exploration into Future Business Process Management Capabilities in View of Digitalization,” *Business & Information Systems Engineering* (63:2), pp. 83-96 (doi: 10.1007/s12599-020-00637-0).

Kerzel, U. 2021. “Enterprise AI Canvas Integrating Artificial Intelligence into Business,” *Applied Artificial Intelligence* (35:1), pp. 1-12 (doi: 10.1080/08839514.2020.1826146).

Kim, D., Park, Y., and Ko, J. 2019. “Factors underlying vehicle ownership reduction among carsharing users: A repeated cross-sectional analysis,” *Transportation Research Part D: Transport and Environment* (76), pp. 123-137 (doi: 10.1016/j.trd.2019.09.018).

Klumpp, M., and Zijm, H. 2019. “Logistics Innovation and Social Sustainability: How to Prevent an Artificial Divide in Human–Computer Interaction,” *Journal of Business Logistics* (40:3), pp. 265-278 (doi: 10.1111/jbl.12198).

Kühl, N., Hirt, R., Baier, L., Schmitz, B., and Satzger, G. 2021. “How to Conduct Rigorous Supervised Machine Learning in Information Systems Research: The Supervised Machine Learning Report Card,” *Communications of the Association for Information Systems* (48:1), pp. 589-615 (doi: 10.17705/1CAIS.04845).

Kühl, N., Mühlthaler, M., and Goutier, M. 2020. “Supporting customer-oriented marketing with artificial intelligence: automatically quantifying customer needs from social media,” *Electronic Markets* (30:2), pp. 351-367 (doi: 10.1007/s12525-019-00351-0).

Kühl, N., Schemmer, M., Goutier, M., and Satzger, G. 2022. “Artificial intelligence and machine learning,” *Electronic Markets* (32:4), pp. 2235-2244 (doi: 10.1007/s12525-022-00598-0).

Lagadic, M., Verloes, A., and Louvet, N. 2019. “Can carsharing services be profitable? A critical review of established and developing business models,” *Transport Policy* (77), pp. 68-78 (doi: 10.1016/j.tranpol.2019.02.006).

Laporte, G., Meunier, F., and Wolfler Calvo, R. 2018. “Shared mobility systems: an updated survey,” *Annals of Operations Research* (271:1), pp. 105-126 (doi: 10.1007/s10479-018-3076-8).

LeCun, Y., Bengio, Y., and Hinton, G. 2015. “Deep learning,” *Nature* (521:7553), pp. 436-444 (doi: 10.1038/nature14539).

Lee, M. C. M., Scheepers, H., Lui, A. K. H., and Ngai, E. W. T. 2022. “The Role of Artificial Intelligence for Business Value,” *ICIS 2022 Proceedings*.

Lei, X., Skoularidou, M., Cuesta-Infante, A., and Veeramachaneni, K. 2019. “Modeling Tabular data using Conditional GAN,”

Leofante, F., Ábrahám, E., Niemueller, T., Lakemeyer, G., and Tacchella, A. 2019. “Integrated Synthesis and Execution of Optimal Plans for Multi-Robot Systems in Logistics,” *Information Systems Frontiers* (21:1), pp. 87-107 (doi: 10.1007/s10796-018-9858-3).

Leroy, J., Bailly, G., and Billard, G. 2023. “Introducing carsharing schemes in low-density areas: The case of the outskirts of Le Mans (France),” *Regional Science Policy & Practice* (15:2), pp. 239-255 (doi: 10.1111/rsp3.12523).

Loureiro, S. M. C., Guerreiro, J., and Tussyadiah, I. 2021. “Artificial intelligence in business: State of the art and future research agenda,” *Journal of Business Research* (129), pp. 911-926 (doi: 10.1016/j.jbusres.2020.11.001).

Lukyanenko, R., Castellanos, A., Parsons, J., Chiarini Tremblay, M., and Storey, V. C. 2019. “Using Conceptual Modeling to Support Machine Learning,” in *Information Systems Engineering in Responsible Information Systems*, C. Cappiello and M. Ruiz (eds.), Cham: Springer International Publishing, pp. 170-181 (doi: 10.1007/978-3-030-21297-1_15).

Ma, L., and Sun, B. 2020. “Machine learning and AI in marketing – Connecting computing power to human insights,” *International Journal of Research in Marketing* (37:3), pp. 481-504 (doi: 10.1016/j.ijresmar.2020.04.005).

Maass, W., Parsons, J., Purao, S., Storey, V. C., and Woo, C. 2018. “Data-Driven Meets Theory-Driven Research in the Era of Big Data: Opportunities and Challenges for Information Systems Research,” *Journal of the Association for Information Systems*, pp. 1253-1273 (doi: 10.17705/1jais.00526).

Majchrzak, A., and Markus, M. L. 2013. “Technology affordances and constraints in management information systems (MIS),” *Encyclopedia of Management Theory* (1) (doi: 10.4135/9781452276090.n256).

Makarius, E. E., Mukherjee, D., Fox, J. D., and Fox, A. K. 2020. “Rising with the machines: A sociotechnical framework for bringing artificial intelligence into the organization,” *Journal of Business Research* (120), pp. 262-273 (doi: 10.1016/j.jbusres.2020.07.045).

Maragno, G., Tangi, L., Gastaldi, L., and Benedetti, M. 2023. "Exploring the factors, affordances and constraints outlining the implementation of Artificial Intelligence in public sector organizations," *International Journal of Information Management* (73), p. 102686 (doi: 10.1016/j.ijinfomgt.2023.102686).

Mariani, M. M., Machado, I., and Nambisan, S. 2023. "Types of innovation and artificial intelligence: A systematic quantitative literature review and research agenda," *Journal of Business Research* (155), p. 113364 (doi: 10.1016/j.jbusres.2022.113364).

Maslej, N., Fattorini, L., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Ngo, H., Niebles, J. C., Parli, V., Shoham, Y., Wald, R., Clark, J., and Perrault, R. 2023. "The AI Index 2023 Annual Report," AI Index Steering Committee (ed.), Stanford, CA: Institute for Human-Centered AI, Stanford University.

Meng, Z., Li, E. Y., and Qiu, R. 2020. "Environmental sustainability with free-floating carsharing services: An on-demand refueling recommendation system for Car2go in Seattle," *Technological Forecasting and Social Change* (152), p. 119893 (doi: 10.1016/j.techfore.2019.119893).

Merhi, M. I. 2023. "An evaluation of the critical success factors impacting artificial intelligence implementation," *International Journal of Information Management* (69), p. 102545 (doi: 10.1016/j.ijinfomgt.2022.102545).

Microsoft. 2022. "The Team Data Science Process lifecycle," available at https://learn.microsoft.com/en-us/azure/architecture/data-science-process/lifecycle.

Mikalef, P., and Gupta, M. 2021. "Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance," *Information & Management* (58:3), p. 103434 (doi: 10.1016/j.im.2021.103434).

Mounce, R., and Nelson, J. D. 2019. "On the potential for one-way electric vehicle carsharing in future mobility systems," *Transportation Research Part A: Policy and Practice* (120), pp. 17-30 (doi: 10.1016/j.tra.2018.12.003).

Mucha, T. M., Ma, S., and Abhari, K. 2022. "Beyond MLOps: The Lifecycle of Machine Learning- based Solutions," *AMCIS 2022 Proceedings*.

Muchenje, G., and Seppänen, M. 2023. "Unpacking task-technology fit to explore the business value of big data analytics," *International Journal of Information Management* (69), p. 102619 (doi: 10.1016/j.ijinfomgt.2022.102619).

Papagiannidis, E., Enholm, I. M., Dremel, C., Mikalef, P., and Krogstie, J. 2023. "Toward AI Governance: Identifying Best Practices and Potential Barriers and Outcomes," *Information systems frontiers : a journal of research and innovation* (25:1), pp. 123-141 (doi: 10.1007/s10796-022-10251-y).

Poeppelbuss, J., Plattfaut, R., and Niehaves, B. 2015. "How Do We Progress? An Exploration of Alternate Explanations for BPM Capability Development," *Communications of the Association for Information Systems* (36) (doi: 10.17705/1cais.03601).

Prinz, C., Hellmeier, M., Willnat, M., Harnischmacher, C., and Kolbe, L. 2022. "Increasing the Business Value Of Free-Floating Carsharing Fleets By Applying Machine-Learning Based Relocations," *ECIS 2022*.

Pumplun, L., Tauchert, C., and Heidt, M. 2019. “A New Organizational Chassis for Artificial Intelligence - Exploring Organizational Readiness Factors,” *Proceedings of the 27th European Conference on Information Systems (ECIS)*.

Rai, A., Constantinides, P., and Sarker, S. 2019. “Editor’s Comments: Next-Generation Digital Platforms: Toward Human–AI Hybrids,” *MIS Quarterly* (43:1), pp. iii-ix.

Ransbotham, S., Kiron, D., Gerbert, P., and Reeves, M. 2017. “Reshaping business with artificial intelligence: Closing the gap between ambition and action,” *MIT Sloan Management Review* (59:1).

Rausch, T. M., and Albrecht, T. 2020. “The impact of lead time and model selection on the accuracy of call center arrivals’ forecasts,” *Proceedings of the 28th European Conference on Information Systems (ECIS)*, pp. 1-16.

Reis, C., Ruivo, P., Oliveira, T., and Faroleiro, P. 2020. “Assessing the drivers of machine learning business value,” *Journal of Business Research* (117), pp. 232-243 (doi: 10.1016/j.jbusres.2020.05.053).

Rosemann, M. 2014. “Proposals for Future BPM Research Directions,” Springer, Cham, pp. 1-15 (doi: 10.1007/978-3-319-08222-6_1).

Rosemann, M., and vom Brocke, J. 2015. “The Six Core Elements of Business Process Management,” in *Handbook on Business Process Management 1*, Springer, Berlin, Heidelberg, pp. 105-122 (doi: 10.1007/978-3-642-45100-3_5).

Rotaris, L., and Danielis, R. 2018. “The role for carsharing in medium to small-sized towns and in less-densely populated rural areas,” *Transportation Research Part A: Policy and Practice* (115), pp. 49-62 (doi: 10.1016/j.tra.2017.07.006).

Russell, S. J., and Norvig, P. 2021. *Artificial Intelligence: A modern approach*, Essex: Pearson.

Schmidhuber, J. 2015. “Deep learning in neural networks: an overview,” *Neural networks : the official journal of the International Neural Network Society* (61), pp. 85-117 (doi: 10.1016/j.neunet.2014.09.003).

Schmöller, S., and Bogenberger, K. 2020. “Carsharing: An overview on what we know,” in *Demand for Emerging Transportation Systems*, Elsevier, pp. 211-226 (doi: 10.1016/B978-0-12-815018-4.00011-5).

Schultze, U., and Avital, M. 2011. “Designing interviews to generate rich data for information systems research,” *Information and Organization* (21:1), pp. 1-16 (doi: 10.1016/j.infoandorg.2010.11.001).

Seidel, S., Recker, J., and vom Brocke, J. 2013. “Sensemaking and Sustainable Practicing: Functional Affordances of Information Systems in Green Transformations,” *MIS Quarterly* (37:4), pp. 1275-1299 (doi: 10.25300/MISQ/2013/37.4.13).

Severino, M. K., and Peng, Y. 2021. “Machine learning algorithms for fraud prediction in property insurance: Empirical evidence using real-world microdata,” *Machine Learning with Applications* (5), p. 100074 (doi: 10.1016/j.mlwa.2021.100074).

Shaheen, S., and Cohen, A. 2020. “Innovative Mobility: Carsharing Outlook; Carsharing Market Overview, Analysis, and Trends - Spring 2020,” Shao, X.-F., Liu, W., Li, Y., Chaudhry, H. R., and Yue, X.-G. 2021. “Multistage implementation framework for smart supply chain management under industry 4.0,” *Technological Forecasting and Social Change* (162), p. 120354 (doi: 10.1016/j.techfore.2020.120354).

Shrestha, Y. R., Krishna, V., and Krogh, G. von. 2021. “Augmenting organizational decision-making with deep learning algorithms: Principles, promises, and challenges,” *Journal of Business Research* (123), pp. 588-603 (doi: 10.1016/j.jbusres.2020.09.068).

Sigov, A., Ratkin, L., Ivanov, L. A., and Da Xu, L. 2022. “Emerging Enabling Technologies for Industry 4.0 and Beyond,” *Information Systems Frontiers* (doi: 10.1007/s10796-021-10213-w).

Sjödin, D., Parida, V., Palmié, M., and Wincent, J. 2021. “How AI capabilities enable business model innovation: Scaling AI through co-evolutionary processes and feedback loops,” *Journal of Business Research* (134), pp. 574-587 (doi: 10.1016/j.jbusres.2021.05.009).

Statista. 2023. “Machine learning - statistics & facts,” available at https://www.statista.com/topics/9583/machine-learning/#topicOverview, accessed on Aug 1 2023.

Strandhagen, J. O., Vallandingham, L. R., Fragapane, G., Strandhagen, J. W., Stangeland, A. B. H., and Sharma, N. 2017. “Logistics 4.0 and emerging sustainable business models,” *Advances in Manufacturing* (5:4), pp. 359-369 (doi: 10.1007/s40436-017-0198-1).

Studer, S., Bui, T. B., Drescher, C., Hanuschkin, A., Winkler, L., Peters, S., and Müller, K.-R. 2021. “Towards CRISP-ML(Q): A Machine Learning Process Model with Quality Assurance Methodology,” *Machine Learning and Knowledge Extraction* (3:2), pp. 392-413 (doi: 10.3390/make3020020).

Teodorescu, M., Morse, L., Awwad, Y., and Kane, G. 2021. “Failures of Fairness in Automation Require a Deeper Understanding of Human-ML Augmentation,” *MIS Quarterly* (45:3), pp. 1483-1500 (doi: 10.25300/MISQ/2021/16535).

Toorajipour, R., Sohrabpour, V., Nazarpour, A., Oghazi, P., and Fischl, M. 2021. “Artificial intelligence in supply chain management: A systematic literature review,” *Journal of Business Research* (122), pp. 502-517 (doi: 10.1016/j.jbusres.2020.09.009).

Uren, V., and Edwards, J. S. 2023. “Technology readiness and the organizational journey towards AI adoption: An empirical study,” *International Journal of Information Management* (68), p. 102588 (doi: 10.1016/j.ijinfomgt.2022.102588).

van der Aalst, W. M. P. 2013. “Business Process Management: A Comprehensive Survey,” *ISRN Software Engineering* (2013), pp. 1-37 (doi: 10.1155/2013/507984).

van Dun, C., Moder, L., Kratsch, W., and Röglinger, M. 2022. “ProcessGAN: Supporting the creation of business process improvement ideas through generative machine learning,” *Decision Support Systems*, p. 113880 (doi: 10.1016/j.dss.2022.113880).

van Looy, A., Trkman, P., and Clarysse, E. 2022. “A Configuration Taxonomy of Business Process Orientation,” *Business & Information Systems Engineering* (64:2), pp. 133-147 (doi: 10.1007/s12599-021-00700-4).

Vanheusden, W., van Dalen, J., and Mingardo, G. 2022. "Governance and business policy impact on carsharing diffusion in European cities," *Transportation Research Part D: Transport and Environment* (108), p. 103312 (doi: 10.1016/j.trd.2022.103312).

Vélez, A. M. A. 2023. "Economic impacts, carbon footprint and rebound effects of car sharing: Scenario analysis assessing business-to-consumer and peer-to-peer car sharing," *Sustainable Production and Consumption* (35), pp. 238-249 (doi: 10.1016/j.spc.2022.11.004).

Venable, J., Pries-Heje, J., and Baskerville, R. 2012. "A Comprehensive Framework for Evaluation in Design Science Research," Springer, Berlin, Heidelberg, pp. 423-438 (doi: 10.1007/978-3-642-29863-9_31).

Vial, G., Cameron, A.-F., Giannelia, T., and Jiang, J. 2023. "Managing artificial intelligence projects: Key insights from an AI consulting firm," *Information Systems Journal* (33:3), pp. 669-691 (doi: 10.1111/isj.12420).

vom Brocke, J., Jans, M., Mendling, J., and Reijers, H. A. 2021. "A Five-Level Framework for Research on Process Mining," *Business & Information Systems Engineering* (63:5), pp. 483-490 (doi: 10.1007/s12599-021-00718-8).

Wamba-Taguimdje, S.-L., Fosso Wamba, S., Kala Kamdjoug, J. R., and Tchatchouang Wanko, C. E. 2020. "Influence of artificial intelligence (AI) on firm performance: the business value of AI-based transformation projects," *Business Process Management Journal* (26:7), pp. 1893-1924 (doi: 10.1108/BPMJ-10-2019-0411).

Wang, T., Hu, S., and Jiang, Y. 2021. "Predicting shared-car use and examining nonlinear effects using gradient boosting regression trees," *International Journal of Sustainable Transportation* (15:12), pp. 893-907 (doi: 10.1080/15568318.2020.1827316).

Weber, M., Engert, M., Schaffer, N., Weking, J., and Krcmar, H. 2023. "Organizational Capabilities for AI Implementation—Coping with Inscrutability and Data Dependency in AI," *Information Systems Frontiers* (25:4), pp. 1549-1569 (doi: 10.1007/s10796-022-10297-y).

Webster, J., and Watson, R. T. 2002. "Guest Editorial: Analyzing the Past to Prepare for the Future: Writing a Literature Review," *MIS Quarterly*, pp. xiii-xxiii.

Wedel, M., and Kannan, P. K. 2016. "Marketing Analytics for Data-Rich Environments," *Journal of Marketing* (80:6), pp. 97-121 (doi: 10.1509/jm.15.0413).

Wendt, C., Adam, M., Benlian, A., and Kraus, S. 2021. "Let's Connect to Keep the Distance: How SMEs Leverage Information and Communication Technologies to Address the COVID-19 Crisis," *Information Systems Frontiers*, pp. 1-19 (doi: 10.1007/s10796-021-10210-z).

Whiting, A., and Donthu, N. 2006. "Managing Voice-to-Voice Encounters," *Journal of Service Research* (8:3), pp. 234-244 (doi: 10.1177/1094670505281703).

Willing, C., Klemmer, K., Brandt, T., and Neumann, D. 2017. "Moving in time and space – Location intelligence for carsharing decision support," *Decision Support Systems* (99), pp. 75-85 (doi: 10.1016/j.dss.2017.05.005).

Winkelhaus, S., and Grosse, E. H. 2020. "Logistics 4.0: a systematic review towards a new logistics system," *International Journal of Production Research* (58:1), pp. 18-43 (doi: 10.1080/00207543.2019.1612964).

Wolfswinkel, J. F., Furtmueller, E., and Wilderom, C. P. M. 2013. "Using grounded theory as a method for rigorously reviewing literature," *European Journal of Information Systems* (22:1), pp. 45-55 (doi: 10.1057/ejis.2011.51).

Xu, M., and Meng, Q. 2019. "Fleet sizing for one-way electric carsharing services considering dynamic vehicle relocation and nonlinear charging profile," *Transportation Research Part B: Methodological* (128), pp. 23-49 (doi: 10.1016/j.trb.2019.07.016).

Yang, M., Fu, M., and Zhang, Z. 2021. "The adoption of digital technologies in supply chains: Drivers, process and impact," *Technological Forecasting and Social Change* (169), p. 120795 (doi: 10.1016/j.techfore.2021.120795).

Yao, Z., Gendreau, M., Li, M., Ran, L., and Wang, Z. 2022. "Service operations of electric vehicle carsharing systems from the perspectives of supply and demand: A literature review," *Transportation Research Part C: Emerging Technologies* (140), p. 103702 (doi: 10.1016/j.trc.2022.103702).

Zhang, C., Schmöcker, J.-D., Kuwahara, M., Nakamura, T., and Uno, N. 2020a. "A diffusion model for estimating adoption patterns of a one-way carsharing system in its initial years," *Transportation Research Part A: Policy and Practice* (136), pp. 135-150 (doi: 10.1016/j.tra.2020.03.027).

Zhang, K., Jia, N., Zheng, L., and Liu, Z. 2019. "A novel generative adversarial network for estimation of trip travel time distribution with trajectory data," *Transportation Research Part C: Emerging Technologies* (108), pp. 223-244 (doi: 10.1016/j.trc.2019.09.019).

Zhang, Z., Nandhakumar, J., Hummel, J. T., and Waardenburg, L. 2020b. "Addressing the Key Challenges of Developing Machine Learning AI Systems for Knowledge-Intensive Work," *MIS Quarterly Executive*, pp. 221-238 (doi: 10.17705/2msqe.00035).

VI Appendix

VI.1 Index of Research Articles

Research Paper P1: Leveraging Digital Technologies in Logistics Processes: Systematic Insights from Intra-Logistics

Albrecht T., Baier M.S., Gimpel H., Meierhöfer S., Röglinger M., Schlüchtermann J., Will L. (2023). Leveraging Digital Technologies in Logistics Processes: Systematic Insights from Intra-Logistics. In: *Information Systems Frontiers*. DOI: 10.1007/s10796-023-10394-6
(VHB-JQ3: B, IF: 5.261)

Research Paper P2: From Zero to Hero: Ramp-Up Management as a New Cross-Cutting Business Process Management Capability

Albrecht T., Lösser B., Röglinger M. (2023). From Zero to Hero: Ramp-Up Management as a New Cross-Cutting Business Process Management Capability. Accepted (with minor revisions) in: *Information Systems and e-Business Management*.
(VHB-JQ3: C, IF: 2.775)

Research Paper P3: Are We There Yet? Analyzing the Role of Access Distance in Carsharing in Small Urban Areas

Albrecht T., Keller R., Röglinger M., Röhrich F. (2023). Are We There Yet? Analyzing the Role of Access Distance in Carsharing in Small Urban Areas. Submitted to: *Journal of Cleaner Production*.
(VHB-JQ3: B, IF: 11.072)

Research Paper P4: Fake It Till You Make It: Synthetic Data for Emerging Carsharing Programs

Albrecht T., Keller R., Rebholz D., Röglinger M. (2023). Fake It Till You Make It: Synthetic Data for Emerging Carsharing Programs. Submitted to: *Transportation Research Part D: Transport and Environment.*
(VHB-JQ3: B, IF: 7.041)

Research Paper P5: Beyond the Beaten Paths of Forecasting Call Center Arrivals: On the Use of Dynamic Harmonic Regression with Predictor Variables

Rausch T.M., Albrecht T., Baier D. (2021). Beyond the Beaten Paths of Forecasting Call Center Arrivals: On the Use of Dynamic Harmonic Regression with Predictor Variables. In: *Journal of Business Economics*. DOI: 10.1007/s11573-021-01075-4
(VHB-JQ3: B, IF: 3.322)

Research Paper P6: Call Me Maybe: Methods and Practical Implementation of Artificial Intelligence in Call Center Arrivals' Forecasting

Albrecht T., Rausch T.M., Derra N.D. (2021). Call Me Maybe: Methods and Practical Implementation of Artificial Intelligence in Call Center Arrivals' Forecasting. In: *Journal of Business Research*. DOI: 10.1016/j.jbusres.2020.09.033
(VHB-JQ3: B, IF: 10.969)

Throughout the dissertation, I also co-authored the following publications. These research papers are not part of this dissertation.

Albrecht T., Baier D. (2020). Churn Analysis Using Deep Learning: Customer Classification from a Practical Point of View. In: *Archives of Data Science Series A*. DOI: 10.5445/KSP/1000098012/04

Rausch T.M., Albrecht T. (2020). The impact of lead time and model selection on the accuracy of call center arrivals' forecasts. In: *Proceedings of the 28th Conference on European Conference on Information Systems (ECIS)*. URL: https://aisel.aisnet.org/ecis2020_rp/19/

Jöhnk J., Albrecht T., Arnold L., Guggenberger T., Lämmermann L., Schweizer A., Urbach N. (2021). The Rise of the Machines: Conceptualizing the Machine Economy. In: *Proceedings of the 24th Pacific Asia Conference on Information Systems (PACIS)*. URL: https://aisel.aisnet.org/pacis2021/54/

VI.2 Individual Contribution to the Included Research Articles

This cumulative doctoral thesis includes six research papers. All research papers were written in collaboration with multiple co-authors. This section outlines the settings and describes my individual contribution to each of the six papers. The descriptions follow the Contributor Roles Taxonomy (CRediT) by Allen et al. (2019).

Research Paper P1 entitled "*Leveraging Digital Technologies in Logistics Processes: Systematic Insights from Intra-Logistics*" (Albrecht et al. 2023; Section VI.3) was written together with six co-authors. I had a key role in most parts of the research project. I contributed significantly to conceptualizing the presentation of results and their embedding in theory, to visualizing the gained insights, and to coordinating the project up to publication. Additionally, I was responsible for drafting multiple sections of the manuscript and was critically involved in reviewing and editing the final paper.

Research Paper P2 entitled "*From Zero to Hero: Ramp-Up Management as a New Cross-Cutting Business Process Management Capability*" (Albrecht et al. 2023; Section VI.4) was written together with two co-authors. I had a crucial role in all parts of the research project. I contributed significantly to formulating the overarching research goals and designing the research methodology of the project. I also engaged in the synthesis and presentation of the research results as well as in the iterative evaluation process. In addition, I was responsible for drafting multiple sections of the manuscript and was involved in reviewing and editing the entire paper.

Research Paper P3 entitled "*Are We There Yet? Analyzing the Role of Access Distance in Carsharing in Small Urban Areas*" (Albrecht et al. 2023; Section VI.5) was written together with three co-authors. I had a crucial role in all parts of the research project. I contributed significantly to the conceptualization of the project and to the design of the research methodology. In addition, I was responsible for the development of the code and models for the analysis as well as for a substantial part of the evaluation and visualization of the results. In terms of writing, I was responsible for the original drafting of individual sections and for reviewing and editing the entire paper.

Research Paper P4 entitled "*Fake It Till You Make It: Synthetic Data for Emerging Carsharing Programs*" (Albrecht et al. 2023; Section VI.6) was written together with three co-authors. In line with my role as the lead author, I was responsible for the conceptualization and administration of the research project. I designed the research approach and developed its underlying code and models. I also led the iterative evaluation process and the preparation and visualization of results. Finally, I was responsible for the original drafting of the manuscript and was involved in the review and editing of the entire research paper. Although I am the lead author of this paper, the three co-authors were involved in all parts of the project and were instrumental in advancing our contribution.

Research Paper P5 entitled "*Beyond the Beaten Paths of Forecasting Call Center Arrivals: On the Use of Dynamic Harmonic Regression with Predictor Variables*" (Rausch et al. 2021; Section VI.7) was written together with two co-authors. I had a crucial role in all parts of the research project. I contributed significantly to the conceptualization of the project and to the design of the research methodology. In addition, I was responsible for developing the code as well as for a substantial part of the application of the models for analysis and evaluation. In terms of writing, I was responsible for the original drafting of individual sections and was involved in reviewing and editing the entire paper.

Research Paper P6 entitled "*Call Me Maybe: Methods and Practical Implementation of Artificial Intelligence in Call Center Arrivals' Forecasting*" (Albrecht et al. 2021; Section VI.8) was written together with two co-authors. I had a crucial role in all parts of the research project. I contributed significantly to the conceptualization of the project and to the design of the research methodology. In addition, I was responsible for developing the code and its walk-through example as well as for a substantial part of the application of the models for analysis and evaluation. In terms of writing, I was responsible for the original drafting of individual sections and was involved in reviewing and editing the entire paper.

VI.3 Research Paper P1: Leveraging Digital Technologies in Logistics Processes: Systematic Insights from Intra-Logistics

Authors:

Albrecht T. · Baier M.S. · Gimpel H. · Meierhöfer S. · Röglinger M. · Schlüchtermann J. · Will L.

Published in:

Information Systems Frontiers (2023)

Abstract:

Emerging digital technologies are transforming logistics processes on a large scale. Despite a growing body of knowledge on individual use cases ranging from collaborative robots to platform-based planning systems in the frontline industrial development of Logistics 4.0, organizations lack a systematic understanding of the opportunities digital technologies afford for logistics processes. To foster such understanding, this study takes an intra-organizational perspective as a central starting point for digitalization initiatives toward Logistics 4.0. It synthesizes current academic research and industrial insights from a systematic literature review and an expert interview study through an affordance lens. The result is a catalog and conceptual framework of ten digital technology affordances in intralogistics (DTAILs) and 46 practical manifestations. Thereby, this study contributes to understanding and leveraging the opportunities digital technologies afford in a leading-edge information systems application domain. It serves as a foundation for further theorizing on Logistics 4.0 and for structuring strategic discussions among organizational stakeholders.

Keywords:

Affordance theory · Digital technology · Industry 4.0 · Logistics 4.0 · Logistics process · Supply chain management

VI.4 Research Paper P2: From Zero to Hero: Ramp-Up Management as a New Cross-Cutting Business Process Management Capability

Authors:

Albrecht T. · Lösser B. · Röglinger M.

To be published in:

Information Systems and e-Business Management (2023)

Abstract:

Changing business environments challenge and motivate organizations to transform. To remain competitive, organizations need to embrace these dynamics and make radical changes to how work is performed. Business process management (BPM) as a holistic management discipline offers mature methods and end-to-end management activities. However, it is subject to the tension between stability and change. While change through the improvement of existing business processes is well understood, the implementation and scaling of novel business processes have been neglected in BPM research. Hence, this paper proposes business process ramp-up management (BPRUM) as a new cross-cutting capability area for contemporary and future BPM and explores relevant sub-capabilities. Our work synthesizes insights from an exploratory interview study with 21 subject matter experts to advance the understanding of BPM as a corporate capability regarding the implementation and scaling of novel processes. As a result, this study illustrates how BPRUM adds to modern BPM and presents 40 action-oriented sub-capabilities that provide hands-on knowledge and practical guidance for effective BPRUM. Thereby, it serves as a foundation for further theorizing on process ramp-up and for structuring discussions among BPM practitioners.

Keywords:

Business process management · Capability development · Interview study · Organizational change · Process ramp-up

VI.5 Research Paper P3: Are We There Yet? Analyzing the Role of Access Distance in Carsharing in Small Urban Areas

Authors:

Albrecht T. · Keller R. · Röglinger M. · Röhrich F.

Submitted to:

Journal of Cleaner Production

Extended Abstract:

Carsharing is a valuable concept of the sharing economy to address today's mobility challenges such as traffic congestion, noise pollution, and CO2 emissions. Therefore, carsharing is increasingly expanded beyond metropolitan areas. However, lower population densities, higher vehicle ownership, and accessibility concerns lead to lower demand, resulting in an increased need for optimization and municipal support. Business models and service offers need to be closely aligned with the specific demand characteristics outside of the dense metropolitan centers (Lagadic et al. 2019; Narayanan and Antoniou 2022). Thus, a thorough understanding of the usage behavior of non-urban carsharing users is needed to facilitate integrated planning of carsharing development and transport policy (Cohen and Kietzmann 2014; Dowling and Kent 2015). Previous research identifies vehicle accessibility as one of the most critical challenges for carsharing programs in less densely populated areas (De Luca and Di Pace 2015; Illgen and Höck, 2020; Lagadic et al. 2019). Against this backdrop, this work aims to analyze the role of access distance in carsharing in small urban areas based on revealed preferences from trip and user data. This study raises the following research question: *How does access distance affect carsharing usage behavior in small urban areas?*

To pursue this research goal, this paper investigates the case of a municipal carsharing program in a small urban area using two years' worth of data, including 5128 active users and 119,146 trips. To this end, it employs descriptive analyses and predictive modeling complemented by explainable machine learning (XML) approaches in the form of model-agnostic permutation feature importance (Fisher et al. 2019) and accumulated local effects (ALE) plots (Apley and Zhu 2020) that have gained momentum in research to determine the impact of explanatory variables on the prediction performance of diverse machine learning (ML) models (Gao et al. 2021; Severino and Peng 2021).

The results indicate that users' access distance is highly relevant for predictions of travel behavior (i.e., usage duration) and strategic decision-making (e.g., station locations and service areas). More trips than in metropolitan areas are made by users with long access distances (i.e., 3 km or more) to their nearest station, while trips with longer access distances are also made by a higher proportion of public transport subscribers, more often made with special types of vehicles (i.e., vans and nine-seaters), and result in longer average trip distances. The findings have significant implications for the understanding of carsharing user behavior in less densely populated areas as well as for the strategic development and operational improvement of carsharing programs supported by municipal policy and governance.

Keywords:

Carsharing · Explainable machine learning · Green travel · Municipal governance · Spatial access · User behavior

References:

Apley, D. W., and Zhu, J. 2020. "Visualizing the effects of predictor variables in black box supervised learning models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* (82:4), pp. 1059-1086 (doi: 10.1111/rssb.12377).

Cohen, B., and Kietzmann, J. 2014. "Ride On! Mobility Business Models for the Sharing Economy," *Organization & Environment* (27:3), pp. 279-296 (doi: 10.1177/1086026614546199).

De Luca, S., and Di Pace, R. 2015. "Modelling users' behaviour in inter-urban carsharing program: A stated preference approach," *Transportation Research Part A: Policy and Practice* (71), pp. 59-76 (doi: 10.1016/j.tra.2014.11.001).

Dowling, R., and Kent, J. 2015. "Practice and public–private partnerships in sustainable transport governance: The case of car sharing in Sydney, Australia," *Transport Policy* (40), pp. 58-64 (doi: 10.1016/j.tranpol.2015.02.007).

Fisher, A., Rudin, C., and Dominici, F. 2019. "All Models are Wrong, but Many are Useful: Learning a Variable's Importance by Studying an Entire Class of Prediction Models Simultaneously," *Journal of Machine Learning Research* (20:177), pp. 1-81.

Gao, L., Lu, P., and Ren, Y. 2021. "A deep learning approach for imbalanced crash data in predicting highway-rail grade crossings accidents," *Reliability Engineering & System Safety* (216), p. 108019 (doi: 10.1016/j.ress.2021.108019).

Illgen, S., and Höck, M. 2020. "Establishing car sharing services in rural areas: a simulation-based fleet operations analysis," *Transportation* (47:2), pp. 811-826 (doi: 10.1007/s11116-018-9920-5).

Lagadic, M., Verloes, A., and Louvet, N. 2019. "Can carsharing services be profitable? A critical review of established and developing business models," *Transport Policy* (77), pp. 68-78 (doi: 10.1016/j.tranpol.2019.02.006).

Narayanan, S., and Antoniou, C. 2022. “Expansion of a small-scale car-sharing service: A multi-method framework for demand characterization and derivation of policy insights,” *Journal of Transport Geography* (104), p. 103438 (doi: 10.1016/j.jtrangeo.2022.103438).

Severino, M. K., and Peng, Y. 2021. “Machine learning algorithms for fraud prediction in property insurance: Empirical evidence using real-world microdata,” *Machine Learning with Applications* (5), p. 100074 (doi: 10.1016/j.mlwa.2021.100074).

VI.6 Research Paper P4: Fake It Till You Make It: Synthetic Data for Emerging Carsharing Programs

Authors:

Albrecht T. · Keller R. · Rebholz D. · Röglinger M.

Submitted to:

Transportation Research Part D: Transport and Environment

Extended Abstract:

Carsharing is an integral part of the transformation toward flexible and sustainable mobility. New carsharing programs are entering the market to challenge large operators by offering innovative services. The introduction of such new services involves a variety of operational and strategic decisions that benefit from understanding and anticipating user behavior (Golalikhani et al. 2021). Hence, predictions of user behavior and trip characteristics from machine learning (ML) models have become indispensable for carsharing operators to optimize their systems. In this regard, large carsharing operators increasingly leverage their rich database to generate insights for improving the efficiency and sustainability of their operations, reduce costs, and provide a better service experience for their users (Golalikhani et al. 2021; Yao et al. 2022). Emerging carsharing programs are mostly excluded from this frontline industrial development. In addition to budget constraints and lower technological expertise, many small carsharing operators struggle with limited data availability and quality that are prerequisites for the effective use of ML models (Lagadic et al. 2019; Shrestha et al. 2021). In this connection, synthetic data created by generative ML models such as generative adversarial networks (GANs) and variational autoencoders (VAEs) present a promising path to overcome the challenge of restricted access to high-quality and privacy-preserving input for the training of data-intensive ML models.

Against this backdrop, this study aims to explore the use of synthetic data to support carsharing decision-making by overcoming the barrier of limited data access during the introduction and expansion of new services. To this end, it means to investigate the evaluation, selection, and implementation of generative ML models to create synthetic tabular transaction data of carsharing trips for more accurate predictions of user behavior when

data access is limited. Thus, this study raises the following research questions: *Can synthetic data created by generative ML models support decision-making in carsharing? And what are appropriate methods?*

To pursue this research objective, this study employs a systematic ML workflow (Kühl et al. 2021; Shrestha et al. 2021) for analyzing the case of an emerging carsharing program that is expanding its services to include free-floating electric vehicles (EVs) and aims to obtain more reliable predictions of trip distances and usage times during the service's ramp-up. The study investigates how well synthetic data generated by GANs, VAEs, and benchmark models serve the purpose of enhancing the available database to improve the training and performance of prediction models. It evaluates the use of synthetic data for replacing, rebalancing, and augmenting real training data drawing on two evaluation protocols that correspond to the prediction of two different target variables.

The results of this study show that augmenting real training data with synthetic samples improves the performance of prediction models by up to 4.63% when predicting the usage time and distance of upcoming trips. The findings of the analysis further reveal that the quality of synthetic data (i.e., using generative ML models such as CTGAN and TVAE compared to statistical benchmark models) as well as the ratio of synthetic to real data are critical to maximizing data utility for ML prediction tasks. The results present novel insights on the use of generative ML for the creation of synthetic mobility data and help understand the methods and approaches for leveraging synthetic tabular transaction data of carsharing trips for more accurate predictions of user behavior. Carsharing operators can draw on this study to enhance their available database when launching or expanding their services to achieve more accurate predictions of upcoming trips and align their service offers with anticipated user behavior.

Keywords:

Carsharing · Electric vehicle · Generative adversarial network · Machine learning · Shared mobility · Synthetic data

References:

Golalikhani, M., Oliveira, B. B., Carravilla, M. A., Oliveira, J. F., and Antunes, A. P. 2021. "Carsharing: A review of academic literature and business practices toward an integrated decision-support framework," *Transportation Research Part E: Logistics and Transportation Review* (149), p. 102280 (doi: 10.1016/j.tre.2021.102280).

Kühl, N., Hirt, R., Baier, L., Schmitz, B., and Satzger, G. 2021. "How to Conduct Rigorous Supervised Machine Learning in Information Systems Research: The Supervised Machine Learning Report Card," *Communications of the Association for Information Systems* (48:1), pp. 589-615 (doi: 10.17705/1CAIS.04845).

Lagadic, M., Verloes, A., and Louvet, N. 2019. "Can carsharing services be profitable? A critical review of established and developing business models," *Transport Policy* (77), pp. 68-78 (doi: 10.1016/j.tranpol.2019.02.006).

Shrestha, Y. R., Krishna, V., and Krogh, G. von. 2021. "Augmenting organizational decision-making with deep learning algorithms: Principles, promises, and challenges," *Journal of Business Research* (123), pp. 588-603 (doi: 10.1016/j.jbusres.2020.09.068).

Yao, Z., Gendreau, M., Li, M., Ran, L., and Wang, Z. 2022. "Service operations of electric vehicle carsharing systems from the perspectives of supply and demand: A literature review," *Transportation Research Part C: Emerging Technologies* (140), p. 103702 (doi: 10.1016/j.trc.2022.103702).

VI.7 Research Paper P5: Beyond the Beaten Paths of Forecasting Call Center Arrivals: On the Use of Dynamic Harmonic Regression with Predictor Variables

Authors:

Rausch T.M. · Albrecht T. · Baier D.

Published in:

Journal of Business Economics (2021)

Abstract:

Modern call centers require precise forecasts of call and e-mail arrivals to optimize staffing decisions and to ensure high customer satisfaction through short waiting times and the availability of qualified agents. In the dynamic environment of multichannel customer contact, organizational decision-makers often rely on robust but simplistic forecasting methods. Although forecasting literature indicates that incorporating additional information into time series predictions adds value by improving model performance, extant research in the call center domain barely considers the potential of sophisticated multivariate models. Hence, with an extended dynamic harmonic regression (DHR) approach, this study proposes a new reliable method for call center arrivals' forecasting that is able to capture the dynamics of a time series and to include contextual information in form of predictor variables. The study evaluates the predictive potential of the approach on the call and e-mail arrival series of a leading German online retailer comprising 174 weeks of data. The analysis involves time series cross-validation with an expanding rolling window over 52 weeks and comprises established time series as well as machine learning models as benchmarks. The multivariate DHR model outperforms the compared models with regard to forecast accuracy for a broad spectrum of lead times. This study further gives contextual insights into the selection and optimal implementation of marketing-relevant predictor variables such as catalog releases, mail as well as postal reminders, or billing cycles.

Keywords:

Forecasting · Call center arrivals · Dynamic harmonic regression ·
Time series analysis · Machine learning · Customer relationship management

VI.8 Research Paper P6: Call Me Maybe: Methods and Practical Implementation of Artificial Intelligence in Call Center Arrivals' Forecasting

Authors:

Albrecht T. · Rausch T.M. · Derra N.D.

Published in:

Journal of Business Research (2021)

Abstract:

Machine learning (ML) techniques within the artificial intelligence (AI) paradigm are radically transforming organizational decision-making and businesses' interactions with external stakeholders. However, in time series forecasting for call center management, there is a substantial gap between the potential and actual use of AI-driven methods. This study investigates the capabilities of ML models for intra-daily call center arrivals' forecasting with respect to prediction accuracy and practicability. We analyze two datasets of an online retailer's customer support and complaints queue comprising half-hourly observations over 174.5 weeks. We compare practically relevant ML approaches and the most commonly used time series models via cross-validation with an expanding rolling window. Our findings indicate that the random forest (RF) algorithm yields the best prediction performances. Based on these results, a methodological walk-through example of a comprehensive model selection process based on cross-validation with an expanding rolling window is provided to encourage implementation in individual practical settings.

Keywords:

Artificial intelligence · Machine learning · Call center forecasting · Predictive analytics