International Journal of Information Management 60 (2021) 102357

Contents lists available at ScienceDirect

International Journal of Information Management

journal homepage: www.elsevier.com/locate/ijinfomgt

Research Article

Swarm intelligence goal-oriented approach to data-driven innovation in customer churn management

Jan Kozak[a], Krzysztof Kania[b], Przemysław Juszczuk[a,*], Maciej Mitręga[c]

[a] Faculty of Informatics and Communication; Department of Machine Learning, University of Economics, 1 Maja, 40-287 Katowice, Poland
[b] Faculty of Informatics and Communication; Department of Knowledge Engineering, University of Economics, 1 Maja, 40-287 Katowice, Poland
[c] Faculty of Economics; Department of Marketing and Business, VSB – Technical University of Ostrava, Sokolská třída 2416/33, 702 00 Ostrava-Moravská Ostrava, Czech Republic

ARTICLE INFO

Keywords:
Churn management
Data-driven innovation
Machine learning
Decision trees
Classification
Dynamic capabilities

ABSTRACT

One type of data-driven innovations in management is data-driven decision making. Confronted with a big amount of data external and internal to their organization's managers strive for predictive data analysis that enables insight into the future, but even more for prescriptive ones that use algorithms to prepare recommendations for current and future actions. Most of the decision-making techniques use deterministic machine learning (ML) techniques but unfortunately, they do not take into account the variety and volatility of decision-making situations and do not allow for a more flexible approach, i.e., adjusted to changing environmental conditions or changing management priorities. A way to better adapt ML tools to the needs of decision-makers is to use swarm intelligence ML (SIML) methods that provide a set of alternative solutions that allow matching actions with the current decision-making situation. Thus, applying SIML methods in managerial decision-making is conceptualized as a company capability as it allows for systematic alignment of allocating resources decisions vis-à -vis changing decision-making conditions.

The study focuses on the customer churn management as the area of applying SIML techniques to managerial decision-making. The objectives are twofold: to present the specific features and the role of SIML methods in customer churn management and to test if a modified SIML algorithm may increase the effectiveness of churn-related segmentation and improve decision-making process. The empirical study uses publicly available customer data related to digital markets to test if and how SIML methods facilitate managerial decision-making with regard to customers potentially leaving the company in the context of changing conditions. The research results are discussed with regard to prior studies on applying ML techniques to decision-making and customer churn management studies. We also discuss the place of presented analytical approach in the literature on dynamic capabilities, especially big data-driven capabilities.

1. Introduction

Since its creation, big data analytics (BDA) has attracted the attention of academics and practitioners (Elgendy & Elragal, 2014; Chong & Shi, 2015; Mikalef, Pappas, Krogstie, & Giannakos, 2018). BDA is defined in several ways (see Riahi & Riahi, 2018) and has many meanings attached to it. BDA has been identified as a new enabler of competitive advantage (Wamba et al., 2017), as a basis for new business models and a tool for business model innovation as in (Ciampi, Demi, Magrini, Marzi, & Papa, 2021), as a field of competition (Davenport & Harris, 2017) and even as a trigger of the next revolution in management (McAfee, Brynjolfsson, Davenport, Patil, & Barton, 2012). BDA also has a profound and multidirectional impact on innovation (Duan, Cao, & Edwards, 2020), a positive effect on environmental scanning and shifts organizations towards a data-driven culture. Empirical research has shown that companies that make data-driven decisions achieve 5% – 6% more efficiency and productivity than might be expected while taking into account other investments and the use of information technology (Brynjolfsson, Hitt, & Kim, 2011). Additionally, recent studies confirm that BDA has a positive impact on organizational performance through its applications in various areas: Wamba et al. (2017), Ji-fan Ren, Fosso Wamba, Akter, Dubey, and Childe (2017), Raguseo (2018), Akhtar,

* Corresponding author.
E-mail addresses: jan.kozak@ue.katowice.pl (J. Kozak), krzysztof.kania@ue.katowice.pl (K. Kania), przemyslaw.juszczuk@ue.katowice.pl (P. Juszczuk), maciej.mitrega@vsb.cz (M. Mitręga).

https://doi.org/10.1016/j.ijinfomgt.2021.102357
Received 10 November 2020; Received in revised form 16 April 2021; Accepted 16 April 2021
Available online 3 May 2021

Frynas, Mellahi, and Ullah (2019), Yasmin, Tatoglu, Kilic, Zaim, and Delen (2020). A vast analysis of BDA capabilities found in organizations is presented in Conboy, Mikalef, Dennehy, and Krogstie (2020).

The core components required to successfully compete using BDA are found in underlying predictive models (Kraus, Feuerriegel, & Oztekin, 2020), which can be used to undertake appropriate actions for the future. The desired patterns and answers are hidden in the data, and to discover them BDA must be supplemented with Machine Learning (ML) & Artificial Intelligence (AI) methods (Ranjan & Foropon, 2021; Dubey et al., 2020) as they are able to overcome computational and sometimes even human limitations and help analyze enormous volumes of data (Dwivedi et al., 2021). Such an approach is especially important for organizations that operate in dynamic and turbulent environments, where fast but informed decision-making is critical (Fosso Wamba, Akter, Edwards, Chopin, & Gnanzou, 2015). BDA based on AI & ML is also a source of innovative ideas, products, and services (Akter et al., 2019; Wamba et al., 2017). Niebel, Rasel, and Viete (2019) provided empirical evidence for a relationship between BDA and innovation and showed that well-informed decision-making based on advanced data techniques is particularly beneficial for business processes involving uncertainty and risk.

For these reasons, strong investment in AI, analyst teams, new tool adoption, data collection, and data processing has been recorded (Perrault et al., 2019), and with cloud infrastructure, limitations on data volumes and processing speeds have disappeared (Mihet & Philippon, 2019; Liang et al., 2018; Mohanty & Vyas, 2018). Many companies have since offered new products, services, or work according to a new business model. In many works, the possibility of using AI as an inherent element of a company has been widely discussed. For example, Barro and Davenport (2019) stated that the adaptation of intelligent technologies greatly benefits overall company operations. However, crucially, the adaptation of AI solutions will depend on the field involved. In some sectors such as medicine, we do not expect to find any fully automatic AI components being used in the near future. Nevertheless, a broad survey by Kiron and Schrage (2019) suggests that today, the general question is not “if” but rather “how” AI and ML methods can improve overall company performance. The concept of introducing new AI-based solutions to companies is extended even more in Tarafdar, Beath, and Ross (2019), where the authors focus on the support of organizational processes. As a result, more complex tasks could be fully automated and would require only minor assistance from a human decision-maker.

Among the ML and AI methods used in BDA, deep learning techniques related mainly to the use of artificial neural networks (ANNs) are of growing interest (Najafabadi et al., 2015; Kraus et al., 2020). This outcome has resulted from the spectacular successes that ANNs have achieved in solving problems that were previously considered to fall within the exclusive domain of human intelligence, such as those related to speech recognition (Nassif, Shahin, Attili, Azzeh, & Shaalan, 2019), image segmentation (Guo, Liu, Georgiou, & Lew, 2018; Alam, Samad, Vidyaratne, Glandon, & Iftekharuddin, 2020), medical diagnosis (Bakator & Radosav, 2018) and strategy games (Silver et al., 2017). However, in areas where justification for decisions and interpretability are required, such as medicine, social activities, or management, the use of ANNs is associated with serious challenges. ANNs do not provide such explanations, and the means used to arrive at a final result remain hidden (see Montavon, Samek, & Müller, 2018), which among other factors has caused research on the adoption of ANNs in decision-making processes to remain at a very early stage (Kraus et al., 2020).

Despite the undeniably positive changes introduced to companies by AI, some limitations remain. For instance, in Hagendorff and Wezel (2020), 15 challenges of using AI are grouped into three clusters: methodological, societal, and technological. In Canhoto and Clear (2020), consequences of characteristics of AI tools, value creation, and the value destruction potential of AI and implications of using AI for managers are broadly discussed.

At the same time, researchers have noted a number of unresolved issues related to the use of BDA in such a way that has the greatest benefit (Mikalef, Boura, Lekakos, & Krogstie, 2019b; Mikalef, Pappas, Krogstie, & Pavlou, 2020; Mikalef et al., 2018). Although BDA is a key tool for data analysis, it requires the fulfillment of organizational conditions and bridging the human skills gap to transform data into actionable knowledge (Mikalef, Framnes, Danielsen, Krogstie, & Olsen, 2017). The importance of the problem of translating the results of AI and ML algorithms into terms useful to the decision-making process is also emphasized by empirical research presented, among others, in Mikalef and Krogstie (2019) and Mikalef et al. (2019b). For BDA to be of benefit, the knowledge discovered by AI and ML tools has to be applied, and this requires appropriate human skills (Mikalef & Krogstie, 2019) and a connection to decision-making processes (Mikalef, Boura, Lekakos, & Krogstie, 2019a). If decision-makers are not able to interpret data properly or are not able to prepare decisions, big data analytics provides only little value (Janssen, van der Voort, & Wahyudi, 2017). Thus, in addition to product and technological innovations, innovations in decision-making processes are necessary and expected by managers (Davenport & Ronanki, 2018). AI & ML tools hold great potential as decision support tools, but users need to be able to interpret, understand, and apply their results if they are to realize this. The predictive capabilities of AI & ML models are very important, but no less important is the ability to interpret the results they provide and the possibility of using them in subsequent stages of the decision-making process (Duan, Edwards, & Dwivedi, 2019). AI & ML methods do not provide practical knowledge automatically. Most organizations that use BDA report difficulties in achieving the anticipated competitive advantage, mainly due to a failure to act on the knowledge provided (Božič & Dimovski, 2019).

This problem has not been bypassed in the realm of churn management. As noted in Baumann, Coussement, Lessmann, and De Bock (2015), most research and models focus solely on customer churn prediction, and although such a prediction is most desirable and crucial for success, from a managerial point of view, it does not resolve all problems of churn management. Thus, traditional churn prediction models are not fully aligned with business goals (Devriendt, Berrevoets, & Verbeke, 2020) and do not derive recommendations for managers and thus do not meet the needs of decision-makers (Vélez, Ayuso, Perales-González, & Rodríguez, 2020). Managers would like to have access to a tool that enables the simulation and preparation of responses in a systematic and interpretable manner. In other words, more prescriptiveness is needed.

As a step towards resolving this problem, we propose using a modified swarm intelligence algorithm to overcome the limitations of deterministic ML methods and enhance decision-makers' capabilities. Swarm intelligence is an innovative, distributed intelligent paradigm for solving optimization problems through which a finite number of agents (objects in the search space) cooperate and eventually move towards the optimum. This concept initially emerged from nature-inspired concepts related to the swarm, a self-organizing foraging phenomenon found among social insects. There are many examples of techniques based on swarm optimization, such as ant colony optimization (Dorigo, Birattari, & Stützle, 2006; Dorigo & Gambardella, 1997), particle swarm optimization (Eberhart & Kennedy, 1995) and artificial bee colony optimization (Karaboga, 2005).

From an algorithmic point of view, we propose a new, goal-oriented metaheuristic modification of the ant colony decision tree (ACDT) algorithm capable of deriving good classification tools with relatively small computation time when data are very complex or the datasets are large. The proposed approach allows for the construction of very effective classifiers – decision trees evaluated on classification efficiency. The algorithmic novelty of this approach is rooted in its twofold evaluation process using different efficiency measures at the same time.

From a managerial point of view, we provide a set of solutions instead of one solution provided by ML deterministic methods and simulation as a tool of prescriptive analytics for better customer segmentation. Thus, data science (computational intelligence) methods are used to efficiently derive highly effective solutions without additional costs. The proposed methodology could be immediately applied to existing data and used to identify weak areas of company-client relations. The specific goal of the proposed innovation is to improve the efficiency of marketing activities aimed at segmenting and targeting customers by building predictive and prescriptive models to optimize targeting decisions and improve customer retention with limited resources.

In our approach we focus on investigating two issues:

1. How using modified swarm algorithm innovates churn prediction;
2. How optimization of basic prediction measures improves decision making in customer churn management.

The proposed solution also uses implicit/deep analytics by delivering a framework for optimization and predictive modeling. The solution reflects one of the main forces driving the data industry identified by Cao (2017), i.e., data/analytics software, which refers to the new tools and applications that analyze and use data for specific business and scientific purposes and provide quality assurance to support these aspects.

The rest of the article is organized as follows. The next section presents a background of the problem explored and related works. Section 3 describes the churn management dataset used. Section 4 presents the ACDT algorithm and its modifications. Section 5 presents numerical experiments derived based on the dataset described in Section 3, while Section 6 provides a discussion of our research.

2. Theoretical background

2.1. Business analytics

Big data analysis can bring significant value to businesses along the entire value chain through its integration with the decision-making process, which, however, remains a challenge (Akhtar et al., 2019), and prescriptive analytics is one of the tools that can resolve this problem (Lepenioti, Bousdekis, Apostolou, & Mentzas, 2020). Typically, there are three levels of business analytics categorized by the scope of analysis, their goals, and the tools they use (see Cao, 2017, Duan et al., 2020, Lepenioti et al., 2020):

- Descriptive analytics generally looks at past and current data and provides a detailed analysis on them. It is exploratory in nature and is focused on aggregating and visualizing historical and current data. It allows online reporting and drill-down into specific sections of the data, combine and visualize them as dashboards. Some authors highlight another type of diagnostic analytics that allows the development and test of hypotheses for critical reasoning and uses data to establish links between different concepts or events–while others include it in the descriptive one.
- Predictive analytics provides reliable forecasts on future possibilities and trends. It uses statistical techniques and AI & ML methods to synthesize past and ongoing data to build models that can predict the future state of variables the manager may be interested in.
- Prescriptive analytics focuses on decision making and efficiency improvements. By prescriptive analysis, it is possible not only to forecast the future but also to obtain justification and guidance on what should be done to make such a future. By prescriptive analytics system can provide recommendations to the decision-maker answering the questions: “what to do?”, “Why should I do it?” and “how to do it?”. Prescriptive analytics assumes the active role of the manager as it uses optimization, simulation, and other interactive tools that allow the user to enter their preferences and choose from possible future actions.

The relationships between individual types of analyses are still being discussed (see Lepenioti, Bousdekis, Apostolou, & Mentzas, 2019), however, it can be concluded that the main differences between these types of analyses are found in their relation to time, the tools they used, and the relation to the decision-making processes they are to support. Descriptive analytics focuses on the past and present, and predictive analytics is carried out to build a reliable forecast of the future. Prescriptive analytics adds to these mechanisms recommendations for actions to answer the question of ’what is to be done.’ This aim can be achieved by delivering actionable knowledge through optimization tools, heuristics, expert knowledge, and deep insights into business data realized by innovative customized algorithms. In this way, data-driven decision making is directly supported (Cao, 2017; Delen, 2020). On the other hand, the entire process is strongly related to the knowledge and experience of humans. In this way, prescriptive analytics fits within the concept of cognitive processing, also referred to as cognitive intelligence. In essence, the approach goes beyond running algorithms and performing the necessary calculations and takes into account the decision context. These mechanisms enable the decision-maker to interpret and understand the acquired knowledge. In interactive mode, they allow the decision-maker to simulate alternative futures and test the consequences of decisions. The goal of cognitive computing is to build a collaborative mechanism from which human capabilities are combined with the computing power of ML and AI algorithms (Gupta, Kar, Baabdullah, & Al-Khowaiter, 2018).

Prescriptive analyses are most desired by managers because they aim at identifying concrete actions for the future. Thus, in the literature, a growing number of examples of prescriptive analytics can be found, such as in medicine (Lopes, Guimarães, & Santos, 2020; Sadat Mosavi & Filipe Santos, 2020), public transportation infrastructure planning (Brandt, Wagner, & Neumann, 2020), human resource planning (Berk, Bertsimas, Weinstein, & Yan, 2019) and many other areas (Greasley, 2019; Poornima & Pushpalatha, 2020; Olszak, 2020). In the field of churn, predictive and prescriptive analytics serve as a tool for answering customers’ questions and as a decision-making tool for preparing marketing strategies (Lejeune, 2001).

2.2. Churn management

AI offers marketers many tasks and functions (Wedel & Kannan, 2016; Campbell, Sands, Ferraro, (Jody) Tsao, & Mavrommatis, 2020). A necessary condition for effective marketing campaigns is to identify a group of customers with similar preferences and who react similarly to specific marketing activities.

At a certain stage of business development, the number of customers of every business reaches such a level that retaining existing customers becomes more important than winning new ones (Hadden, Tiwari, Roy, & Ruta, 2007). Customer retention is highly profitable for companies. Attracting new clients costs five to six times more than retaining existing customers, long-term customers tend to be more profitable and are more loyal, and losing customers leads to lost income (Verbeke, Martens, Mues, & Baesens, 2011; Devriendt et al., 2020). In a competitive and saturated market, it is becoming increasingly more difficult to retain customers. In the digital market, where customers can switch suppliers with a few clicks, this problem become even more pressing (Gordini & Veglio, 2017). The special importance of churn analysis was emphasized very soon after it was found that retaining existing clients is even more beneficial than acquiring new ones (Goodhue, Wixom, & Watson, 2002).

In a broader context, forecasting customer churn is one of the basic elements of customer relationship management (CRM) (Coussement, Benoit, & Van den Poel, 2010). Risselada, Verhoef, and Bijmolt (2010) indicated that churn management is a part of CRM, while Ngai, Xiu, and Chau (2009) located churn analysis (also called attrition analysis) as one of the dimensions of CRM–customer retention–and called it the central concern of CRM. In turn, Krishna and Ravi (2016) indicate that churn detection is a part of analytical CRM. Verhoef et al. (2010) showed that churn prediction is a part of a broader task of targeting the right customers and refers to activities related to the strategy of maintaining contact with customers. This prediction and quantification of risks of customer loss can be made globally or individually and is mainly used in areas where a product or service is marketed on a subscription basis. The prediction of churn is generally done by studying consumer behavior or by observing individual behavior that indicates a risk of attrition. The approach involves the use of modeling and ML techniques that can sometimes use a considerable amount of data. Churn analysis is a part of a broader set of actions aimed at retaining customers referred to as churn management. The ultimate goal of churn management is to take preventive/retention actions (Hadden et al., 2007).

Churn analysis has a precisely defined goal; it requires processing large datasets, and the models built can be used in many different companies. Due to these characteristics, churn analysis is very suitable for business decision-making support with AI methods (Stone et al., 2020). A wide plethora of data mining and AI & ML techniques are used in churn analysis. Churn analysis is conducted in several areas. Most of these areas, due to ease of data collection, are related to digital markets, including telecommunications (Jain, Khunteta, & Srivastava, 2020; Amin et al., 2019; Ahmad, Jafar, & Aljoumaa, 2019; Olle, 2014), banking (Keramati, Ghaneei, & Mirmohammadi, 2016; Shirazi & Mohammadi, 2019; Nie, Rowe, Zhang, Tian, & Shi, 2011), insurance (Smith, Willis, & Brooks, 2000; Morik & Kopcke, 2004), retail stores (Patel, Struckell, Ojha, & Manikas, 2020; Miguéis, Camanho, Falcão, & Cunha, 2013), and online platforms (cloud computing, music, videos, media, and games) (Kim, Choi, Lee, & Rhee, 2017; Yuan, Bai, Song, & Zhou, 2017; Zdravevski, Lameski, Apanowicz, & Ślęzak, 2020), but also include medicine (Orzol, Hula, & Harrington, 2015) and the automotive industry (Meinzer, Jensen, Thamm, Hornegger, & Eskofier, 2016).

Large sets of data with many features cannot be analyzed other than with AI methods. These AI & ML techniques in churn include:

- clustering (Hung, Yen, & Wang, 2006; Liu & Zhuang, 2015),
- rule discovery (Amin et al., 2016; Verbeke et al., 2011),
- decision trees and random forests (Idris, Rizwan, & Khan, 2012; Höppner, Stripling, Baesens, vanden Broucke, & Verdonck, 2020; Nie et al., 2011),
- genetic algorithms (Stripling, Antonio, Baesens, & Snoeck, 2018),
- deep neural networks (Mena, De Caigny, Coussement, De Bock, & Lessmann, 2019; De Caigny, Coussement, De Bock, & Lessmann, 2019),
- regression models (linear or logistic) (De Caigny, Coussement, & De Bock, 2018),
- Bayesian networks (Kisioglu & Topcu, 2011; Amin et al., 2019),
- SVM (Coussement & Van den Poel, 2008; Xai & Jin, 2008; Yu, Guo, Guo, & Huang, 2011; Gordini & Veglio, 2017).

At the same time, hybrid approaches that combine various AI techniques have been proposed as well as ensemble methods that are based on gathering results of many classifiers and voting procedure to establish a final outcome (Shirazi & Mohammadi, 2019; Chu, Tsai, & Ho, 2007; Kim, Jung, Kim, & Lee, 2012; Huang & Kechadi, 2013; Baran, Kułakowski, & Ligęza, 2014; Keramati et al., 2014; De Caigny et al., 2018). An exhaustive comparative study of ML methods in churn prediction can be found in Vafeiadis, Diamantaras, Sarigiannidis, and Chatzisavvas (2015), Krishna and Ravi (2016), Ahmed, Afzal, Majeed, and Khan (2017), Jain et al. (2020).

Marketing managers try to find useful knowledge from every available source. Most of the research on churn focuses on customer features (such as age, gender, tenure, location, etc. taken from system databases) and customer activities (frequency of activity, time since last contact, etc. taken from system logs); however, other sources such as ads, online articles (Lee, Kim, & Lee, 2017), social networks (Kim, Jun, & Lee, 2014; Verbeke, Martens, & Baesens, 2014; Oskarsdóttir et al., 2017) and complaint data (Hadden et al., 2007) are also taken into consideration. Churn analysis can be related to different types of customers. Most works have been related to individual customers (B2C model); however, the B2B context has also been considered (Tamaddoni Jahromi, Stakhovych, & Ewing, 2014; Gordini & Veglio, 2017; Van Haver, 2017; Barfar, Padmanabhan, & Hevner, 2017). A few frameworks for churn management have been developed (for instance Hadden et al., 2007; Ghorbani & Taghiyareh, 2009; Daramola, Oladipupo, & Musa, 2009). The main problem is that these frameworks are focused on steps related to data collection, preparation, and processing by AI algorithms.

One of the still existing problems of churn management concerns the proper application of the ML methods mentioned above and the translation of analysis results into appropriate actions. According to Akhtar et al. (2019), the final and critical step of the entire BDA process is *action on insight*. Without proper decisions, no one can expect benefits, and the whole analysis remains useless regardless of how valuable the knowledge discovered is. The use of results of BDA in business practice within the scope of churn management is noted in Lee, Lee, Cho, Im, and Kim (2011), where a model of churn prediction is connected to a retention marketing program. Later, this model was developed to directly include business objectives and management costs (Kim, Lee, & Johnson, 2013).

Problems surrounding the lack of connection between the results of data analysis and decisions made regarding a marketing campaign are clearly explored in Baumann et al. (2015), where the authors note that most of the proposed models concern the determination of the probability of customer churn but do not take into account the current situation of the decision-maker (see also Devriendt et al., 2020). In addition, Kim et al. (2013) claim that most prediction models do not consider additional costs of churn management programs (see also Lejeune, 2001). The need to prepare a next step ahead towards decision-making is also discussed in Vélez et al. (2020).

The conducted literature studies allow stating that there are still research gaps and open issues, including:

- limited possibilities of deterministic ML methods in generating various proposed solutions, from which the decision-maker could make a choice,
- a weak connection between the stage of generating solutions by ML and AI tools and the decision-making process itself and limited influence of the manager on the operation of ML and AI algorithms,
- the lack of connection of ML and AI algorithms with tools supporting the decision-making process, such as simulation.

The proposal presented in this article aims to overcome these shortcomings by using swarm intelligence algorithms and expanding the decision-making space (limited to a single solution in deterministic methods), giving the manager the possibility to independently indicate the purpose of the ML algorithm's work and enabling simulation in the solution space.

2.3. Dynamic capabilities in big data analytics

The increased turbulence in the environment observed since the end of the prior century (D'Aveni, Canger, & Doyle, 1995) stimulated the evolution of the resource perspective in strategy research (Amit & Schoemaker, 1993; Barney, 1991) and the emergence of the dynamic capabilities view (DCV) (Teece, Pisano, & Shuen, 1997; Eisenhardt & Martin, 2000). For the DCV, the focus is not on company resources but rather on the processes that transform these resources and operational capabilities in relation to changing conditions in the environment (Teece et al., 1997). These strategic processes are described through an entrepreneurial cycle that involves sensing opportunities, seizing opportunities and reconfiguring company assets to meet changing external conditions and internal limitations (Teece, 2007). The DCV is frequently perceived as a research paradigm with many unsolved issues, such as those related to the appropriate location of dynamic capabilities within an organization and the role of managers (Teece, 2012; Di Stefano, Petraf, & Verona, 2014), the company size that makes developing dynamic capabilities reasonable (Chirico & Nordqvist, 2010; Døving & Gooderham, 2008; Parida, Lahti, & Wincent, 2016) and the time needed to develop dynamic capabilities in radically changing environments (Eisenhardt & Martin, 2000; Vergne & Durand, 2011). Recently, Schilke, Hu, and Helfat (2018) demonstrated that among further research directions proposed in studies published in the DCV, antecedents to the development of dynamic capabilities and underlying processes shaping dynamic capabilities are frequently noted by many scholars. Thus, although it is commonly accepted that flexibly organizing company resources increases chances of company success, our knowledge is limited with regard to the conditions needed in organizations to materialize such a management approach. Importantly, the academic research illustrates that the popularity of the DCV paradigm has resulted in its application to specific business functions such as marketing capabilities (Buccieri, Raj, & Erin, 2020; Mitręga, Spacil, & Pfajfar, 2020), new product development capabilities (Marsh & Stock, 2003; Prieto, Revilla, & Rodríguez-Prado, 2009) alliancing capabilities (Kohtamäki, Rodrigo, & Möller, 2018; Weerawardena, Mort, Liesch, & Knight, 2007), supply management capabilities (Forkmann, Henneberg, Naudé, & Mitręga, 2016; Hong, Yibin, & Minqiu, 2018) and, finally, big data analytics (Ciampi et al., 2021; Conboy et al., 2020; Mikalef et al., 2019b).

The DCV has been used in research on data-driven innovations, demonstrating that data analytics systems may be effectively applied to adjust company strategies to changing environments (Côrte-Real, Tiago, & Ruivo, 2017; Ghasemaghaei, Khaled, & Turel, 2017; Mikalef & Pateli, 2017). The business product that is being modeled and tested in this research stream is company agility understood as the speed of effective company responses to the environment (Katayama & Bennett, 1999). For better competitive results, big data analytics can be combined with knowledge management (Côrte-Real et al., 2017; Sher & Lee, 2004). Building big data analytics capabilities (BDAC) enable various operational and strategic benefits, such as the generation of quick insights about emerging trends and identification of necessary adjustments required to meet future needs (Wang & Hajli, 2017). Ghasemaghaei et al. (2017) provided evidence that using data analytics in the company improves firm agility when it is supported by a high level of fit between the intensity of data analytics use and the use of other strategic resources: analytical tools, people using data and tasks that data should help perform. Most recently, Ciampi et al. (2021) illustrated a direct link between BDAC and business model innovation in companies operating in various industries in the UK.

The link between the dynamic capabilities of organizations and BDAC is usually conceptualized and tested, and BDAC are treated as an antecedent of building organizational dynamic capabilities (Mikalef et al., 2019b; Mikalef et al., 2017; Mikalef et al., 2018; Wamba et al., 2017), which makes these resources similar to operational resources and capabilities built in other functional domains (Schilke et al., 2018). However, some studies propose treating BDAC as a specific form of dynamic capability itself. For example, Ciampi et al. (2021, p. 3) considered "BDAC as lower-order DCs" (dynamic capabilities). Chen and Lin (2020) proposed a model in which the application of artificial intelligence in business is conceptualized as building dynamic capabilities, specifically sensing, transforming and driving capabilities, and the authors also show evidence for interlinks between such capabilities and their sequential impacts on company performance. Interestingly, except for studies that support the general link between company BDAC and company dynamic capabilities, some studies suggest that BDAC as an antecedent for building specific dynamic capabilities. For example, Wamba et al. (2017) illustrates that BDAC leverage process-oriented dynamic capabilities such as by incorporating detailed information into a business process. Although the placement of BDAC within the nomological net of various dynamic capabilities demands further investigation, especially when applying longitudinal research approaches (Conboy et al., 2020), there seems to be a consensus that BDAC are higher-order capabilities involving some lower-order resources and capabilities (Mikalef et al., 2018; Olszak, 2014; Wamba et al., 2017). Thus, BDAC do not manifest in companies possessing single data analytics resources, e.g., appropriate data storage, data analytics software or qualified data analytics personnel, but in various organizational routines through which these resources are utilized in organizations to derive value from changing company environments (Mikalef et al., 2018). As demonstrated in recent research, there is no single way in which BDAC may be applied in business practice; instead, there are various specific processes that companies deploy in various contexts to reorganize their strategic resources based on big data analytics (Conboy et al., 2020; Mikalef & Krogstie, 2020). Thus, there is a need to better understand how specific big data resources may be organized to facilitate various innovation outputs (Mikalef & Krogstie, 2020).

Following the recent call for an improved understanding of context-specific big data resources as foundations of BDAC (Fiorini, Seles, Jabbour, Mariano, & Jabbour, 2018; Mikalef & Krogstie, 2020), in this research, we focus on a specific data-analytics approach well suited to the context of customer churn management (Bi, Cai, Liu, & Li, 2016; Huang, Kechadi, & Buckley, 2012; Neslin, Gupta, Kamakura, Lu, & Mason, 2006). We assume that the proposed SIML approach to churn management may be treated as a key BDAC resource (Mikalef et al., 2018) within this context because the approach enables flexible customer churn prediction vis-a-vis changing decision-making conditions (Teece et al., 1997; Eisenhardt & Martin, 2000). Specifically, the proposed analytical approach allows for flexible and automatized modifications of churn prediction priorities based on the trade-off between two measures, precision and recall without sacrificing the prediction accuracy rate. The traditional approaches to churn prediction were rather static with regard to the importance of these two parameters, so the proposed SIML approach to churn provides novelty in terms of making churn decisions more flexible. In the next sections, we present in detail the SIML algorithm, and we introduce the numerical experiment used to test this algorithm based on data from the telecom industry. In the discussion, we explore how using this algorithm corresponds with prior research on BDAC.

2.4. Evaluation of the quality of classification

Evaluation of the quality of classification is one of the crucial problems of machine learning. It is crucial to evaluate if the given classifier can be considered as good or bad (in sense of the quality). However, there is a lack of classifiers, which could be customized on the basis of the used classification measure, or optimized according to more than a single classification measure at a time. It is important in the case of real-world problems, for which it can be reasonable to balance the importance of two different classification measures. This aspect is not new in literature, and different approaches were used to derive the more broad picture of classification based on more than one measure. The subject is related to the ROC measure presented for example in Fawcett (2006).

Below the most popular measures of the binary classification were presented (for the sets with two different decision classes) - possible to calculate based on the error matrix (Table 1). Thanks to the error matrix it is possible to better evaluate this classification on the basis of the information about the decision class of the analyzed object and the class, for which the object has been classified. Such information allows calculating the accuracy, the recall, the precision, and the F-measure along with the Matthews correlation factor.

One should know, that the accuracy of the classification is the measure describing only the number of objects properly classified. While

Table 1
The error matrix.

	Positive prediction	Negative prediction
Positive cases (*P* class)	True positive (*TP*)	False negative (*FN*)
Negative cases (*N* class)	False positive (*FP*)	True negative (*TN*)

the precision allows us to evaluate, what is the measure of confidence, that the object from the given decision class will be properly classified. It is especially important, where the significance of one decision class over the other is essential for the decision-maker.

Accuracy

Accuracy is one of the most popular measures of the evaluation of classification. However, it should be noted, that it is not a sufficient measure in the case, where a large diversity of objects in the decision classes is observed (Kozak & Boryczka, 2013). Accuracy of the classification describes the ratio of properly classified objects to all classified objects:

$$\mathrm{ev}_{\mathrm{acc}}(T,S) = \frac{(\mathrm{TP}+\mathrm{TN})}{(\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN})}. \tag{1}$$

Recall

Recall is the measure used to evaluate the classifier on the basis of the objects, which were classified incorrectly from N class to P class. In this case, the better option is to skip some objects from the P class, rather than the wrong classification to this class objects from the N class. Recall is calculated on the basis of the ratio of properly classified objects to class P to all objects classified to the P class:

$$\mathrm{ev}_{\mathrm{rec}}(T,S) = \frac{\mathrm{TP}}{(\mathrm{TP}+\mathrm{FN})}. \tag{2}$$

Precision

Precision is calculated on the basis of formulae Eq. (3), which is a ratio of the number of properly classified objects to class P and all objects, which should be classified to this class. It could be said, that in the case of precision, it is better to wrongly assign the object from class N to class P, than omit the object from class P.

$$\mathrm{ev}_{\mathrm{prec}}(T,S) = \frac{\mathrm{TP}}{(\mathrm{TP}+\mathrm{FP})}. \tag{3}$$

When all cases are correctly classified ("false positive" = "false negative" = 0), all these measures take the value 1. For managers, it means complete certainty in making decisions. In such cases, even the use of deterministic algorithms is sufficient. However, such situations are rare and sometimes even suspicious (overfitting of learning algorithm).

3. Data description

The idea of using machine learning and swarm algorithms is based on the properly derived dataset. Such a process in most cases involves data acquisition, data preprocessing, and transformation into a format easily understandable by the algorithm. The Telco Customer Churn dataset available online was selected as a source of data for the proposed method. We selected this particular set due to two crucial factors. The first factor relates to the fact that data are strongly unbalanced, and over $\frac{3}{4}$ of all objects belong to one class, while the remaining objects belong to the second class. We focused on the dataset in which there are only two decision classes because our algorithm is adapted to such data. We should note that it is possible to adapt data with more than two decision classes. However, this would require additional preprocessing leading to modifications of the basic dataset. The second factor is related to a large amount of data describing the single object, which initially could be unrelated to the problem, such as a large amount of demographic information. It was interesting to observe the eventual possible impact of such information on the results.

All available data is presented as the decision table, which can be defined as follows: let there be a set of elements (customers) X, and a set of attributes V. Every attribute V_i includes attribute values $V_i = \{v_{i,1}, v_{i,2}, \ldots, v_{i,n}\}$, where n is overall number of these values for the attribute i. Eventually, the single object x is described by a set of attribute values and it is assigned to one of the decision classes. By a decision class, we understand one of the values of decision attribute.

The overall data description is as follows:

- data about the customers who left/who did not within the last month);
- services that each customer has signed up for (like phone, internet, online security, tech support, and so on);
- customer account information like payment method, monthly charges, total charges;
- general customer information: gender, age range.

Each row represents a customer, each column contains the customer's attributes described on the column Metadata. The raw data contains 7043 rows (customers) and 21 columns. Among these features, we can find 12 String features (descriptive), five boolean attributes, and two integer attributes. The boolean "Churn" attribute is the decision class.

The general goal of using AI & ML methods is to derive a classifier capable of estimating the potential risk of customer churn based on its overall description. Such an estimation could be used to plan more directed contact with customers and involve additional actions leading to the minimization of risks of customer churn.

4. Balanced goal function of the ACDT algorithm

Classical deterministic algorithms have a predetermined goal and focus on obtaining the highest degree of solution accuracy. Such algorithms generate a single result that determines further decisions and actions. From a managerial point of view, this approach is inflexible because the algorithm does not take into account the situation and constraints of the decision-maker in any way. The swarm intelligence algorithms used mostly for optimization allow for a goal-oriented approach. An example of such an algorithm can be found in the adaptive goal function of the ACDT algorithm (Kozak & Boryczka, 2014). In this work, the concept of different classifiers (specifically, the decision trees) oriented on different measures was proposed. Despite the classical accuracy rate, measures such as precision and recall were analyzed as well. Swarm algorithms are inherently stochastic, and as a result of their operation, a set of results is obtained. From the manager's point of view, this approach is slightly better, as it can provide a suboptimal solution but one that is better suited to the current situation.

Our goal is to propose the balanced goal function of the ACDT approach (called BGF-ACDT). In this case, the goal function of the ACDT algorithm is directed towards two measures of classification: precision and recall. This is used to derive the decision-maker's probability of using a control strategy. Obviously, classifiers are still built on the basis of more classical tools such as Classification and Regression Trees (CART) (Breiman, Friedman, Olshen, & Stone, 1984) and thus are related to the accuracy of the classification. However, the modifications introduced into the algorithm are optimized according to the balanced goal function (balanced on the basis of both precision and recall). The possibility of controlling the relative importance of this parameter reflects the novelty of this approach.

Such an approach offers a visible advantage of the ACDT algorithm over different classical approaches, for which there is no opportunity to adapt the classifier to the selected measure, while the split criteria remain unchanged. As an effect, decision trees (classifiers) geared towards precision or recall (depending on the importance of weights) can be derived.

4.1. Ant colony decision tree algorithm

Swarm intelligence is a group of methods based on biological inspirations and is often based on self-organizing swarms of insects. There are many examples of such approaches, including particle swarm optimization, artificial bee colonies, and ant colony optimization. The latter approach is related to artificial systems and inspired by the natural behavior of real ants, specifically based on the pheromone trail ants leave on the ground (Dorigo & Stützle, 2004).

The basic premise of this approach is self-organization. Swarm intelligence systems are complex, which in general means that they consist of a set of simple agents working in parallel and cooperating with each other and with the environment locally. Such an approach allows the production of some behaviors observed globally among the whole population of agents.

The ant colony decision tree algorithm is based on the use of ant algorithms (Dorigo & Stützle, 2004) to construct decision trees. The single run of the algorithm is based on test selection (division test) for every node on the basis of two factors. The first factor is the maximum value compatible with the splitting criterion of the CART algorithm, while additional information is included in the pheromone trail (Kozak, 2019). The second factor is the connection of the pheromone trail to collective intelligence.

The splitting criterion is used to find the best test, which will allow a splitting of data analyzed in the single node into two maximally uniform (as a number of objects in the decision class) parts. All constructed decision trees are binary; thus, the splitting criterion always divides objects into two parts. In the ACDT algorithm, the classical twoing criterion (originally taken from the CART algorithm) was used.

The process of constructing the decision tree with the use of the ACDT algorithm is presented in Algorithm 1. One should note that this approach connects the classical machine learning algorithms (based on the statistics) and in this case, the CART algorithms, whose solutions were used in the heuristic function. The second element is swarm intelligence and in this particular case, the pheromone trail used as a shared swarm memory. Additionally, the concept related to the stochastic approach allows us to include potentially worse solutions that globally could perform better here based on the probability of selecting the test and a roulette wheel.

After setting the initial value of the pheromone trail (responsible for remembering the previously generated solutions), every iteration of agent-ant (number of iterations) leads to creating a new decision tree (as it was shown in rows 3–13). Taking into account the values of the heuristic function (which is the division pointed out by the classical machine learning algorithm) and the pheromone trail (swarm memory), which are derived during the analysis of each node (row 7, where the randomness used in the swarm intelligence algorithms is introduced). The agent-ant which constructed the best decision tree in the iteration (the best in terms of quality like the accuracy of the classification and the size of the decision tree) have an impact on updating the pheromone trail (information used in further iterations) – row 4. As a result of the algorithm, the best decision tree constructed by the agent-ants is derived for the decision-maker.

Algorithm 1. Pseudo-code of the ACDT algorithm

```
1   pheromone = initialization_pheromone_trail();
2   for number_of_iterations do
3     for number_of_ants do
4       //build the decision tree
5       while (stop_condition_is_not_fulfilled)
6         p = calc_the_choosing_probability(pheromone, heuristic); // Eq. (4)
7         //choice the test in the node (roulette wheel)
8         new_tree→test = roulette_wheel(p_{m,m_L(i,j)}(t));
9       endWhile
10      if new_tree is_higher_quality_than best_tree then
11        best_tree = new_tree;
12      endIf
13    endFor
14    update_pheromone_trail(best_tree, pheromone);
15    if best_tree is_higher_quality_than best_constructed_tree then
16      best_constructed_tree = best_tree;
17    endIf
18  endFor
19  result = best_constructed_tree;
```

The value of the heuristic function is calculated based on the split criterion used in the CART algorithm, while the probability of selecting the test in the node is calculated on the basis of the following formula:

$$p_{i,j} = \frac{\tau_{m,m_{L(i,j)}}(t)\cdot\eta_{i,j}^{\beta}}{\sum_{\mathrm{ii}}^{a}\sum_{\mathrm{jj}}^{b_{\mathrm{ii}}}\tau_{m,m_{L(\mathrm{ii},\mathrm{jj})}}(t)\cdot\eta_{\mathrm{ii},\mathrm{jj}}^{\beta}}, \tag{4}$$

where:

$\eta_{i,j}$ – heuristic information coefficient for the test of attribute i (or ii) with the value j (or jj),

$\tau_{m,m_{L(i,j)}}(t)$ – the pheromone trail in time t for the vertex leading from node m to node m_L (for the test of attribute i (or ii) with the value j (or jj)),

β – parameter, for which the value calculated during the experiments is set to 3,

a – the number of attributes,

b_i – the number of values of ii-th attribute.

4.2. Balanced goal function

A new approach related to the BGF-ACDT algorithm is based on the change introduced into the fitness function of agents: ants. This modification is related to a way of calculating the evaluation function of the decision tree (Equation (5)) affecting, for example, the updating of the pheromone trail to the formula Eq. (6). For such an approach, the solution with the higher value of the selected classification measure will be rewarded by the pheromone trail.

Two factors have the most impact on the evaluation of the quality of the decision tree. According to the rule of the minimal description, the size of the decision tree should be minimized while the accuracy of the classification is maximized. The quality of the decision tree under further consideration will be understood according to the following formula:

$$Q(T) = \phi\cdot w(T) + \psi\cdot a(T,S) \tag{5}$$

where:

$w(T)$ – the size of tree T,

$a(T,S)$ – accuracy of the classification of objects from test set S by tree T,

ϕ and ψ – constants describing the relative importance of values $w(T)$ and $a(T,S)$.

In the presented experiments, the following formulation was simplified:

$$Q(T) = \phi\cdot w(T) + \psi\cdot \mathrm{bev}(T,S), \tag{6}$$

where $\mathrm{bev}(T,S)$ denotes the balanced method of the evaluation of classifier T constructed by the agent and on the basis of the dataset S. Value $\mathrm{bev}(T,S$ (defined as Eq. (7)) is calculated depending on the decision-maker's preferences because it is the relative value of recall Eq. (2) and precision Eq. (3).

$$\mathrm{bev}(T,S) = \kappa\cdot \mathrm{ev}_{\mathrm{prec}}(T,S) + (1-\kappa)\cdot \mathrm{ev}_{\mathrm{rec}}(T,S), \tag{7}$$

where parameter κ describes the relative importance of values $\mathrm{ev}_{\mathrm{prec}}(T,S)$ and $\mathrm{ev}_{\mathrm{rec}}(T,S)$. In the case of churn, the parameter may, for example, represent a preference for the scope or nature of a marketing campaign.

5. Numerical experiments

The goal of the conducted experiments was to estimate the impact of the goal function of the BGF-ACDT algorithm on the results obtained with respect to the accuracy rate, precision, and recall. Numerical experiments were conducted to confirm the hypothesis that it is possible to adapt the algorithm according to the decision-maker's preferences and construct classifiers for which the focus is on the precision, recall, or balancing of these two measures. All the above results were obtained with similar accuracy rate values.

5.1. Experiments design

All experiments used the Telco Customer Churn dataset described in Section 3. For validation, the classical "train and test" method was used. The overall dataset was divided into three fragments: training and validation datasets used for algorithm learning and a test dataset used for an independent evaluation of the algorithm. This approach is related to the selected algorithm, where the ant colony's optimization of solutions constructed by the agent ants must be evaluated during the algorithm run (in this case, on the basis of the trained dataset from which the classifier is constructed while the validation dataset is used to evaluate this solution and eventually to leave the pheromone trail). After the algorithm runs, the test dataset is used to evaluate the efficiency of the final classifier derived by the algorithm.

The Telco Customer Churn dataset includes 7043 objects (customers); the dataset is unbalanced where 5174 objects (more than 73%) are in decision class (customer churned) "No" while 1869 are in decision class "Yes." As a result of this division, every set includes $\frac{1}{3}$ of the whole Telco dataset with a similar distribution of decision classes. All presented results reflect a classification of objects from the test dataset by the classifier constructed on the basis of the training and validation datasets ($\frac{2}{3}$ of all cases).

Experiments were performed for the algorithm with the following parameter values:

- ACC – classical ACDT version (only accuracy rate Eq. (1));
- PREC – for $\kappa = 1.0$ with Eq. (7), which denotes optimization towards precision Eq. (3);
- REC – for $\kappa = 0.0$ with Eq. (7), which denotes optimization towards recall Eq. (2);
- BAL – for $\kappa = 0.5$ with Eq. (7), which involves 50% precision and 50% recall;
- BAL25 – for $\kappa = 0.25$ with Eq. (7), which involves 25% precision and 75% recall;
- BAL75 – for $\kappa = 0.75$ with Eq. (7), which involves 75% precision and 25% recall.

Each of these experiments was repeated 30 times, and the presented results are the average values obtained from these runs. Such an approach was used for the statistical confirmation of the obtained results. For every run, 25 iterations and 250 agents were used.

The parameter values employed in ant colony optimization are established in the way first presented by Boryczka and Kozak (2010), Kozak (2019): $q_0 = 0.2$, $\beta = 3.0$, $\gamma = 0.1$. These are the default, experimentally determined values for the ACDT algorithm. The experiments were carried out on an Intel Core i7-6500U CPU 2.50 GHz x 4 Computer with 7.7 GB RAM with the Linux operating system. The algorithm was implemented in C++11, and scripts related to input data preprocessing were included in Python 3.7.2.stem. The algorithm has been implemented in C++11, and scripts related to input data preprocessing in Python 3.7.2.

5.2. Results of experiments

The conducted experiments allowed us to test whether there a decision-maker can control the goal of the algorithm by using a balanced goal function. Charts presenting the ratio of recall to precision for specific classifiers allow us to observe the adaptation of the algorithm to the given task.

In Figs. 1–4, the single dots indicate the quality of the classifier tested on the test dataset. Swarm intelligence algorithms are nondeterministic; thus, 30 algorithm runs were repeated. The solution obtained by the deterministic algorithm would be represented by a single point on the

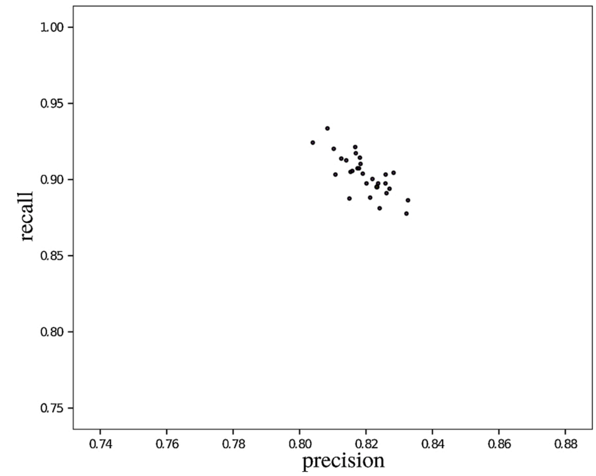

Fig. 1. The ratio of recall and precision for the classical ACDF algorithm.

chart, and therefore, the figure was omitted for the sake of simplicity. In Fig. 1, all classifiers constructed for the classical version of the ACDT algorithm are indicated (which was adapted only to the accuracy rate). The recall and precision ratio for most cases are rather similar. Classical, deterministic algorithms allow us to construct a single classifier with similar results. The possibility of optimization directed for precision or recall is presented in Figs. 2 and 4.

Fig. 2 presents classifiers constructed during the BGF-ACDT (version PREC, REC, and additionally the classical ACDT). The red dots indicate all nondominated classifiers. These classifiers are not comparable to each other; however, all of them dominate the remaining solutions represented by black dots. In general, the nondominated classifiers create a Pareto front, while the remaining solutions could be omitted in the further decision process. It can be observed that the classifiers are arranged on the Pareto front. The control of the goal function of the algorithm results in the construction of classifiers presenting worse results for one measure, while the second measure is visibly improved. However, some gaps between the results for the ACDT and PREC and the REC are observed.

A balanced fitness function allows for any filling of this gap. For the tests we assumed that balance ratios of 50/50, 25/75, and 75/25 would be presented; however, upon setting the κ value, the decision-maker could select any ratio. Fig. 3 presents the results for all of the above-mentioned ratios. The obtained Pareto front is presented without any noticeable gaps, which means that by setting the proper parameter

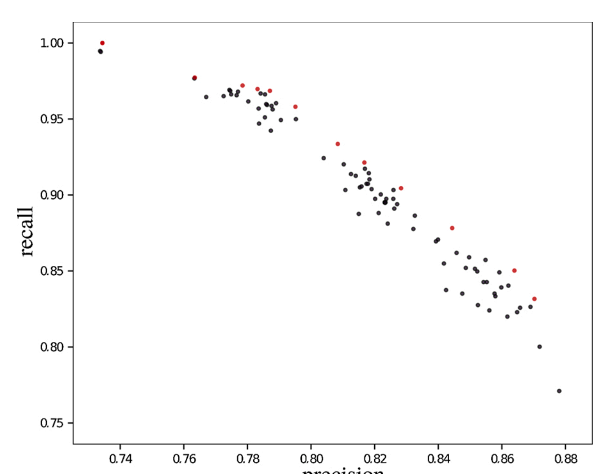

Fig. 2. The ratio of recall for precision for the BGF-ACDF algorithm (PREC, REC).

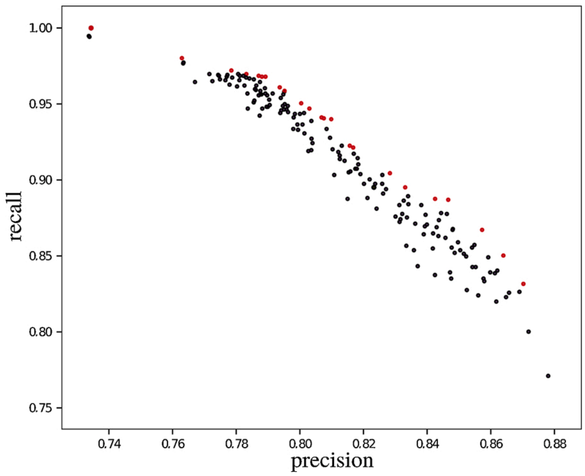

Fig. 3. The ratio of recall for precision for the classical ACDF algorithm (PREC, REC, BAL, BAL25, BAL75).

values, a decision-maker could obtain a classifier compatible with preferences.

The adaptation using the κ parameter is presented in Fig. 4. The obtained Pareto front constructed once again from the classifiers denoted by red dots is presented without any noticeable gaps. All remaining dominated solutions denoted by the black dots are left to show the overall number of solutions derived by the proposed algorithm. By selecting a proper value for the parameter, the emphasis of the selected criterion (or both when κ is set near the 0.5 value) leads to the introduction of a subset of solutions from the Pareto front according to the decision-maker's preferences. Only the use of solutions related to the improvement of the best (in the given time) solutions allows the achievement of a stable classifier presenting good efficiency in the measure selected by the decision-maker.

Exact results for all presented versions of algorithms are given in Table 2. The accuracy rate is similar for every case. The difference between the best (BAL75 and ACC) and worst (REC) observed results is approximately 1.3%., which means that despite the adaptation of the solution to decision-maker preferences, the accuracy level remains relatively high.

Very important observations are found for precision and recall evaluation. As shown for the goal function, oriented towards precision (PREC version), the algorithm achieved over 12% stronger results for this measure than in the classical approach (recall is over 6% lower). At the same time, during the algorithm run oriented towards recall (REC version), the improvement of this measure reaches over 6% (precision is over 4% lower). A further analysis balancing these two measures indicates that fewer changes in accuracy rates, precision, and recall improve the drawback of values according to the selected goal function and the κ parameter. In such a way, for example, version BAL25 (a 25% weight for precision and 75% weight for recall), in comparison to the classical algorithm, allows for a recall improvement of over 5% by lowering the precision value by 3%.

The most important observations related to the conducted experiments concern the values of measures of classification quality related to the algorithm goal function. However, the remaining aspects involved in measuring classification quality should not be omitted. In Table 3, the

Table 2
Value of the measures of the quality of classification depending on the selected approach – average from all algorithm runs (bolded text is the best value).

		Quality of classification		
		acc	prec	rec
Goal function	ACC	0.7827	0.8195	0.9033
	PREC	0.7767	**0.8559**	0.8373
	REC	0.7701	0.7767	**0.9655**
	BAL	0.7828	0.8005	0.9383
	BAL25	0.7780	0.7854	0.9608
	BAL75	**0.7832**	0.8413	0.8689

Abbrev goal function: ACC – accuracy rate; PREC – precision; REC – recall; BAL – 50% precision and 50% recall; BAL25 – 25% precision and 75% recall; BAL75 – 75% precision and 25% recall.
Abbrev quality of classification acc – accuracy rate Eq. (1); prec – precision Eq. (3); rec – recall Eq. (2).

Table 3
Number of nodes of the decision tree and learning time depending on the selected approach – average from all algorithm runs (bolded text indicates the best result).

		Num. of nodes	Time (in seconds)
Goal function	ACC	119.9	240.46
	PREC	104.0	229.58
	REC	**53.7**	225.89
	BAL	88.5	237.57
	BAL25	61.6	**218.33**
	BAL75	108.7	226.37

Abbrev goal function: ACC – accuracy rate; PREC – precision; REC – recall; BAL – 50% precision and 50% recall; BAL25 – 25% precision and 75% recall; BAL75 – 75% precision and 25% recall.

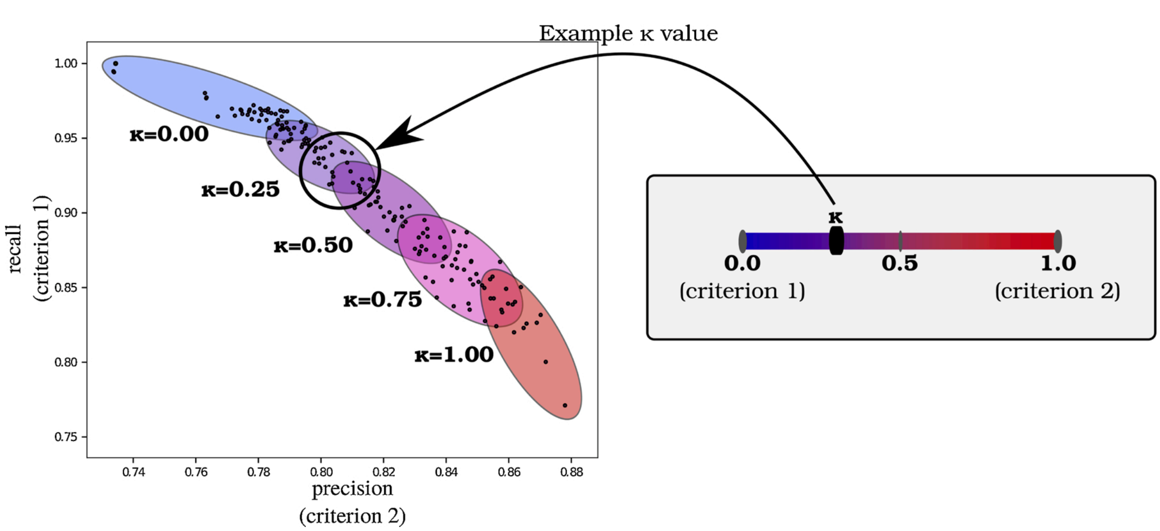

Fig. 4. The ratio of recall for precision for BGF-ACDF algorithm (PREC, REC) – all classifiers.

average numbers of nodes in the decision tree and the time needed to learn the classifier are presented. Lower values are preferred (bolded text indicates the best-achieved results) because every new solution brings improvements related to the number of nodes and learning time relative to the classical ACDT algorithm. The best improvements in the number of nodes were achieved by the REC and BAL25 approaches. The computation time was decreased by a few percent relative to the classical version of the algorithm.

6. Research discussion

Conducted experiments provide evidence that a new balanced goal function of the ACDT algorithm allows for the improvement of the measure selected by the user, while the second measure is at the same time worse. Thus, the decision-maker is capable of controlling the parameter values and indicates which improvement is preferred (assuming that there is a trade-off between these two measures). At the same time, the accuracy rate remains rather unchanged. The learning time of the classifier is not longer, while the number of nodes for the decision tree is decreased for every case of the BGF-ACDT algorithm (in comparison to the classical version of the ACDT). The algorithm is fast enough to enable real-time solution use and customer data stream tracking despite the high complexity of the tasks.

The positive results of the experiment confirm the possibility of overcoming the limitations of deterministic algorithms and extending the manager's decision space. The modified swarm algorithm enhances a set of identified methods of prescriptive analytics (Lepenioti et al., 2020) by using simulation together with evolutionary algorithms where a combination of these two tools is absent. It is worth emphasizing that the proposed modification of the algorithm can be used not only in churn analysis but wherever the decision-maker looks for solutions in multidimensional state space.

The algorithm we present and test here is flexible, and depending on the purpose of the analysis, the decision-maker can select the appropriate direction of optimization. In the churn analysis, the goal is to identify customers who are about to leave, i.e., churners, and in this case, every single point in Fig. 4 represents one solution (separate classification tree) to the problem of classifying customers as churners and nonchurners. Points with higher recall and lower precision (upper left corner of the chart and κ closer to 0) represent solutions involving all possible customers who want to leave but at the expense of precision, i.e., the indicated groups will also include a certain number of nonchurners. If such solutions are selected, the manager must accept that in the campaign target group, there will also be customers who did not intend to leave at all. Thus, these solutions are reasonable when there is resource slack in the company (Boso et al., 2017; Sarkis & Sundarraj, 2003) or if there are no current key customers (i.e., customers with the highest share of turnover or opinion leaders) (Popadiuk & Choo, 2006; Wagner & Hansen, 2004), but they probably exist in the customer database, so losing one strategic customer is too risky, even at the cost of rewarding those customers who would probably stay anyway.

As managers move along the Pareto front towards solutions with greater precision (κ closer to 1), they prioritize finding only actual churners even if some cases of potential churns are to be omitted. Solutions such as this will be more justified if the company has only very limited resources for the campaign and intends to focus only on those customers who are going to leave. In other words, the campaign will reach only those customers who truly want to churn (most precise targeting), but in this case, the risk of missing through the campaign a certain number of customers who also show symptoms of departure is higher. There are many contexts in which there are no slack resources, including socioeconomic crises that become cyclical in the case of small companies or companies in emerging economies. In such cases, targeted marketing with the highest possible precision may be preferred (Jiang, Dan, & Jie, 2019; You et al., 2015).

In the presented proposal, depending on their preferences, decision-makers can control the κ value, so the system indicates the appropriate group of solutions. In this way, our proposition assumes the active role of managers during the simulation, and hence, it combines two managerial elements of AI systems: cognitive insights and cognitive engagement (Davenport & Ronanki, 2018). The results of the classification process, along with the error matrix, can be assigned to predictive analysis, as future unknown cases can be predicted based on historical data. Supplementing this system with simulation builds a basis for the preparation of possible future action scenarios and qualifies the entire solution for prescriptive analysis.

6.1. Discussion and theoretical implications

In marketing science, customer churn/switching behavior has been extensively studied, resulting in the identification of various altitudinal, behavioral, or demographic antecedents of customer churn (Farah, 2017; Ganesh, Arnold, & Reynolds, 2000; Hocutt, 1998; Keaveney & Parthasarathy, 2001; Roos & Gustafsson, 2007; Shin & Kim, 2008). However, prior research does not provide a universal model of customer churn applicable to various research contexts, e.g., various industries and national cultures. Additionally, an important limitation of prior research lies in its generation of churn management variables that are theoretically sound but not managerially implementable, for example, due to the always limited access to customer data, especially attitudinal data (Keaveney & Parthasarathy, 2001). Prior research largely ignores the changing nature of customer perceptions in firm-customer relationships, which create blind spots in customer management (Wägar, Roos, Ravald, & Edvardsson, 2012), and the dynamics of customer experiences that arise beyond the focal customer relationship, e.g., recent campaigns of competitors (Roos & Gustafsson, 2007). With a phenomenological approach, the customer churn decision is strongly dependent on the customer context, so any information base used in churn management is always limited (Wägar et al., 2012). Taking into consideration the limitations of prior research, this study approaches the customer churn problem from a different angle. Specifically, it is assumed that there is always only limited access to altitudinal, behavioral, or demographic features of customers that may be related to an inclination to churn and there is never a stable set of churn determinants, but within such informational boundaries, it is possible to identify such customer features most closely linked to customer churn in the past (i.e., dynamically modifying the past period analysis length using machine-learning methods) and based on that target most likely future churners. In turn, such a target group should be approached with predefined churn handling practices, e.g., dedicated discounts or push messaging (Ascarza, Iyengar, & Schleicher, 2016; Milošević, Živić, & Andjelković, 2017). Although prior research has tested various algorithms that enable the identification of potential churners based on the same assumption, the classical deterministic algorithms used provide the decision-maker with a single result (i.e., only one customer segment with the best accuracy for potential churn), and such an approach cannot be 100% accurate and is inflexible considering the changing contexts of managerial decisions. The prediction approach based on a novel SIML method extends these limitations because instead of providing only one "best solution," it provides the decision-maker with a choice space between two parameters: recall (focusing on all probable churners with a high risk of targeting some potential nonchurners) and precision (most likely focusing on churners only with a high risk of abandoning some potential churners). As "*almost all churn prevention methods have a certain cost to them*" (, p. 331), using the SIML approach programmed for the highest possible precision level may be a means to adjust marketing strategies to the negative COVID-19 economic climate, where marketing budgets are limited and cash flow is the main driver of company behavior (Kang, Diao, & Zanini, 2020).

Our numerical experiment provides evidence that the SIML approach can be effectively applied to select churners in the electronic market because this is a context of a mass-scale global market with fierce competition where targeted marketing should be effective and quick (Wamba et al., 2017; Krishna & Ravi, 2016). The SIML approach presented may be used by the marketing manager to flexibly modify targeted marketing vis-a-vis changing environmental and related budgetary conditions. Considering the volatility of global value chains, which has risen since the economic crisis in 2008, the need for flexible marketing planning is stronger than ever. Electronic markets are very unpredictable because, on the one hand, global turnover is constantly growing in these markets, but on the other hand, local and global crises such as the COVID-19 pandemic discontinue international supply chains and greatly influence companies' cash flows. Therefore, companies functioning in electronic markets may use the data-driven approach we propose here as a way to flexibly react to fast-changing conditions, e.g., prioritizing recall in churn targeting when the economic climate improves and prioritizing precision in difficult times. Notably, our machine-learning approach can be flexibly programmed with regard to the various periods of historical data analysis, which allows for learning mechanisms and for adjusting even to the most recent changes. For example, with such discontinuous changes, there is potential to use historical data related to the emergence of a crisis only for churners targeting subsequent periods.

From the perspective of strategy research, the novelty of the SIML approach proposed in this paper refers to the possibility of predicting important business phenomena (in the context of customer churn) with high accuracy and flexible managerial decision making, which in turn aligns with the principles of the DCV paradigm (Eisenhardt & Martin, 2000; Schilke et al., 2018). The analytical approach that we present and test complements previous studies on innovative managerial cognition and decision making as dynamic capabilities (Helfat & Peteraf, 2015; Peteraf, Di Stefano, & Verona, 2013). Specifically, the balancing of precision and recall as two prediction parameters (embedded in this approach) may be treated as practical means of fast and experiential decision making that are especially important for dynamic capabilities when confronted with rapid technological changes (Eisenhardt & Martin, 2000). Although our study does not provide evidence that managers effectively use this algorithm when making real-life decisions, the numerical experiment demonstrated that such an application affords increased flexibility to prediction. We tested the proposed analytical approach in the context of the electronic market, i.e., the historical data of customers of a telecom company, while such a context facilitates the development of dynamic capabilities (Teece et al., 1997). Following the call for more studies that provide concrete conditions for developing dynamic capabilities (Barreto, 2010; Schilke et al., 2018), we argue that our study proposes one such condition that is fully manageable, especially for companies with access to a large amount of customer data.

Our study complements emerging research on dynamic capabilities applied to big data or big data analytics capabilities (BDAC) (e.g., Côrte-Real et al., 2017; Ghasemaghaei et al., 2017; Mikalef & Pateli, 2017). Prior research provides the conceptualization and measures for such capabilities (Mikalef et al., 2019a; Mikalef et al., 2018; Wamba et al., 2017) as well as the evidence that using such capabilities leverages the development of dynamic capabilities in organizations, especially if the strategic aim is to increase company agility (Côrte-Real et al., 2017; Ghasemaghaei et al., 2017; Mikalef & Pateli, 2017), company innovation (Ciampi et al., 2021; Mikalef et al., 2019b) or company performance in general (Wamba et al., 2017; Chen & Lin, 2020). In contrast to the majority of prior BDAC studies, this study does not focus on the data analytics routines pervading the whole organizational structure to provide decision-making insights for general management; instead, the study proposes a concrete big data prediction approach that has certain features that are well aligned with the principle of flexibly and rapidly adjusting company resources to changing conditions, which is the foundation of the BDAC concept (Mikalef et al., 2018; Wamba et al., 2017) and dynamic capabilities concept in general (Teece et al., 1997; Eisenhardt & Martin, 2000). Specifically, this study proposes a novel prediction approach based on swarm intelligence machine learning (SIML) methods that provides a manager with a set of alternative solutions that allow for the matching actions to the current decision-making situation. The study demonstrates that such a prediction approach is not only novel but also effective at churn prediction based on numerical experiments conducted using real data collected in the telecom industry. However, following the BDAC literature, we do not propose this analytical approach as a substitute for all other important aspects of building and using BDAC (Mikalef et al., 2018; Wamba et al., 2017). Instead, we treat this approach as a complement to the universal conceptualization of BDAC and especially as a complement to resource-building BDAC (Mikalef et al., 2018). We claim that the proposed analytical approach extends beyond tangible technology resources and refers to human skills (Mikalef et al., 2018; Ciampi et al., 2021), which are needed for an appropriate balancing of prediction parameters under changing decision-making conditions. The simulation-based results of our numerical experiments ensure the potential to combine such technological and human resources in churn prediction, while further studies may test the application of our prediction approach to other contexts, including the identification of organizational barriers to applying it. While recent studies follow an explorative approach to the resource manifestations of BDAC in real companies (Conboy et al., 2020; Mikalef & Krogstie, 2020), our study provides a concrete analytical tool and discusses its application as the resource foundation of BDAC.

Although the SIML prediction approach we propose here may potentially be applied to various functional domains of management to automatize the predictions needed for decision-making, in this study, we present an application of this approach to customer churn management, so the study contributes mostly to the emerging research on dynamic capabilities in marketing (Barrales-Molina, Martínez-López, & Gázquez-Abad, 2014; Buccieri et al., 2020; Mitręga, 2019). Specifically, among various marketing resources and capabilities that can be reshaped by dynamic capabilities (Kozlenkova, Samaha, Stephen, & Palmatier, 2014; Morgan, 2012), our study contributes to prior research by providing evidence that the proposed SIML technique allows for flexible customer segmenting adjusted to changing management conditions. Theoretically, the SIML approach proposed here allows for the potential for companies to adjust to changes of various kinds, i.e., under external (e.g., changing economic climate and determinants of customer loyalty) and internal conditions (e.g., changing resource slack). Therefore, companies functioning in electronic markets may use this data-driven approach to flexibly react to fast-changing conditions, e.g., prioritizing recall in churn targeting when the economic climate improves and prioritizing precision in difficult times. The proposed machine-learning approach can be flexibly programmed with regard to the various periods of historical data analysis, which allows for learning mechanisms and adjusting even to the most recent changes in customer tendencies. For example, with discontinuous changes, there is potential to use historical data related to the emergence of a crisis only for churners targeting subsequent periods. Although our study demonstrates the novelty of the proposed SIML approach and its alignment with the BDAC concept (Mikalef et al., 2018; Wamba et al., 2017), the proposed application of such an approach in customer churn prediction remains mainly within conceptual boundaries of sensing opportunities as one of the keystones of BDAC, while BDAC assumes also transforming company resources to seize these opportunities (Conboy et al., 2020). There is no universal combination of BDAC resources that drives positive outcomes (Mikalef & Krogstie, 2020); thus, even if the full exploitation of the proposed prediction approach demands further study, we provide evidence that it may be effectively applied in strategic business functions, which currently involve customer churn management (Bi et al., 2016; Huang et al., 2012; Neslin et al., 2006).

6.2. Practical implications

The SIML approach that we propose in this work was successfully tested with regard to the identification of potential churners in a company functioning in the electronic market, specifically from the historical data of 7043 customers of a telecom company. However, the decision-making flexibility, which is the main contribution of this approach relative to classical deterministic algorithms, i.e., manipulating precision and recall as prediction parameters, may be of more universal use for various microlevel and macrolevel decisions where there is a trade-off between these two parameters. At the microlevel, the SIML approach may be usefully applied in HRM in every organization with access to employees' data. Specifically, the approach may increase the effectiveness of identifying potential churners among employees and targeting them with preventive measures (Hung et al., 2006; Saradhi & Palshikar, 2011). At the macrolevel, i.e., in central government decision making, the same algorithm may be applied for the identification of individuals who may likely be exposed to some environmental threats or engage in dangerous behavior. For example, in the context of the threat of a pandemic, it is possible to analyze the characteristics of inhabitants who died from the disease in the very first period of the pandemic and use this algorithm to prepare screening tests. In turn, the selected individuals could be provided with obligatory tests and early treatment due to their potential health vulnerabilities (Wang et al., 2020). In both of these areas, the decision-maker may flexibly respond to the decision-making context and accordingly balance precision with recall as prediction parameters. For example, if there are strong resource constraints and there is a need for a quick response, maximizing precision may be most reasonable and quickly applied. Although the likelihood of deselecting some individuals exposed would be high, this may be balanced with the application of additional qualitative criteria, e.g., targeting only individuals already in hospitals or targeting only employees in managerial positions in the case of HR managers. However, if such additional criteria cannot be established, the costs of incorrect nonselection are very high, and the decision-maker may flexibly prioritize the recall parameter.

The natural application of the SIML approach involves customer churn management because this marketing function is becoming increasingly important (Bi et al., 2016; Huang et al., 2012; Neslin et al., 2006), and this paper presents numerical experiments related specifically to this area. However, managers should be careful when using the proposed SIML approach to preventing their customers from defecting. Prior research suggests that retention treatment campaigns may have the opposite results, e.g., increased brand switching behavior due to increased awareness that it is easy to change (Ascarza et al., 2016; Bi et al., 2016), which goes beyond churn detection within the churn prevention system (Milošević et al., 2017, p. 330). Therefore, to minimize potential negative effects, we suggest that marketing managers apply a "subtle" treatment of potential churners (e.g., conducting a satisfaction survey to demonstrate customer orientation) rather than "hard" tools (e.g., proposing discounts explicitly related to prolonging the contract). Preventive techniques should always be adjusted to the context of a specific industry. Our numerical experiment was conducted in the mass-market context where simple push notifications sent to individual users may be effective (Milošević et al., 2017), but customer churn management in the context of noncontractual business-to-business settings is less understood (Jahromi, Stakhovych, & Ewing, 2014). Thus, there may be a need to use more complex, multistep prevention techniques, e.g., direct interviews followed by contract renegotiations.

6.3. Limitations and future research

The study presented in this paper is not free of limitations. First, the numerical experiment used here to test the SIML approach uses historical data only. Ideally, such further research would also measure the treatment implemented to prevent a selected customer from defecting. However, customer churn prediction and prevention are distinct aspects of the churn prevention system (Milošević et al., 2017), and in this study, we consciously focused on churn prediction. The other limitation of our numerical experiment relates to its focus on one telecom company dataset that includes only some general customer variables, e.g., monthly charges, customer churn, and demographics. Future research may use data from various industries, including variables related to customer attitudes, e.g., customer satisfaction. While such attitudinal data have been found to be very important in prior research on customer churn management (Hocutt, 1998; Farah, 2017), such data are hardly accessible in scientific research, and it is also not clear to what extent using such data is manageable in business practice (Keaveney & Parthasarathy, 2001). Additionally, while this paper discusses various practical applications of the SIML approach extending beyond customer churn management, i.e., in employee churn management or government detection of risky groups, such applications were not tested here, leaving an avenue for further research.

Finally, although this study conceptualizes utilizing a novel SIML approach as a building block of dynamic capabilities in general (Teece et al., 1997; Eisenhardt & Martin, 2000) and BDAC specifically (Mikalef et al., 2018; Wamba et al., 2017), developing complex BDAC in organizations goes beyond even very sophisticated uses of big data analytics and is based on adjusting organizational routines and organizational cultures to seize big data-related business opportunities (Fiorini et al., 2018; Mikalef et al., 2019b); therefore, future studies may consider organizational and environmental contingencies in applying the proposed SIML approach to various areas of big data innovation. Although our focus mostly on technical aspects of BDAC is a limitation, as our numerical experiment tests the proposed SIML approach not based on the declarations of managers but instead on real conditions and objective customer data, some typical problems related to self-reported bias in business research (Randall & Fernandes, 1991; Fuller, Simmering, Atinc, Atinc, & Babin, 2016) were minimized here.

7. Conclusions

In this paper, we examined the impact of a goal-oriented computational intelligence algorithm for data-driven innovation in customer churn management. The analytical approach we proposed was conceptualized as a building block of dynamic capabilities with regard to big data analytics. The purpose of our analysis was to investigate the innovation provided by the application of the proposed balanced goal function of the ACDT algorithm in churn prediction. In particular, we focused on possibility of application and further extension of selected classifications measure.

Based on the experiments, it can be concluded that by using swarm intelligence that enables focusing on a specific goal, our approach allows the solution to be customized. Depending on the current requirements of the decision-maker, it is possible to steer the construction of a classifier oriented towards recall, precision, as well as establishing the relative importance of these measures. Such an approach allows the decision-maker to focus on solution, which is (from her/his point of the view) the most cost-effective.

In the future, a more universal BGC-ACDT approach can be worked on with more decisions available than just customer churn. In addition, more classification measures can be used as an available goal for the decision-maker to build a classifier.

Author statement

Kozak Jan: conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing – original draft, writing – review & editing, visualization, project administration. Kania Krzysztof: conceptualization, methodology, formal analysis, writing – original draft, writing – review & editing, supervision, project administration. Juszczuk Przemysław: conceptualization, validation, formal analysis, investigation, resources, writing – original draft, writing – review & editing, project administration. Mitręga Maciej:

conceptualization, formal analysis, data curation, writing – original draft, writing – review & editing, supervision, project administration.

References

Ahmad, A. K., Jafar, A., & Aljoumaa, K. (2019). Customer churn prediction in telecom using machine learning in big data platform. *Journal of Big Data, 6*(1). https://doi.org/10.1186/s40537-019-0191-6.

Ahmed, M., Afzal, H., Majeed, A., & Khan, B. (2017). A Survey of evolution in predictive models and impacting factors in customer churn. *Advances in Data Science and Adaptive Analysis, 09*(03), 1750007. https://doi.org/10.1142/S2424922X17500073.

Akhtar, P., Frynas, J. G., Mellahi, K., & Ullah, S. (2019). Big data-savvy teams' skills, big data-driven actions and business performance. *British Journal of Management, 30*(2), 252–271. https://doi.org/10.1111/1467-8551.12333.

Akter, S., Bandara, R., Hani, U., Fosso Wamba, S., Foropon, C., & Papadopoulos, T. (2019). Analytics-based decision-making for service systems: A qualitative study and agenda for future research. *International Journal of Information Management, 48*, 85–95. https://doi.org/10.1016/j.ijinfomgt.2019.01.020.

Alam, M., Samad, M. D., Vidyaratne, L., Glandon, A., & Iftekharuddin, K. M. (2020). Survey on deep neural networks in speech and vision systems. *Neurocomputing, 417*, 302–321. https://doi.org/10.1016/j.neucom.2020.07.053.

Amin, A., Anwar, S., Adnan, A., Nawaz, M., Aloufi, K., Hussain, A., & Huang, K. (2016). Customer churn prediction in the telecommunication sector using a rough set approach. *Neurocomputing, 237*, 242–254. https://doi.org/10.1016/j.neucom.2016.12.009.

Amin, A., Al-Obeidat, F., Shah, Adnan, A., Loo, J., & Anwar, S. (2019). Customer churn prediction in telecommunication industry using data certainty. *Journal of Business Research, 94*, 290–301. https://doi.org/10.1016/j.jbusres.2018.03.003.

Amit, R., & Schoemaker, P. (1993). Strategic assets and organizational rent. *Strategic Management Journal, 14*(1).

Ascarza, E., Iyengar, R., & Schleicher, M. (2016). The perils of proactive churn prevention using plan recommendations: Evidence from a field experiment. *Journal of Marketing Research, 53*(1), 46–60. https://doi.org/10.1509/jmr.13.0483.

Bakator, M., & Radosav, D. (2018). Deep learning and medical diagnosis: A review of literature. *Multimodal Technologies and Interaction, 2*(3). https://doi.org/10.3390/mti2030047.

Baran, M., Kułakowski, K., & Ligęza, A. (2014). A note on machine learning approach to analyze the results of pairwise comparison based parametric evaluation of research units. *Lecture Notes in Artificial Intelligence, 8468*, 27–39.

Barfar, A., Padmanabhan, B., & Hevner, A. (2017). Applying behavioral economics in predictive analytics for B2B churn: Findings from service quality data. *Decision Support Systems, 101*, 115–127. https://doi.org/10.1016/j.dss.2017.06.006.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management.*. https://doi.org/10.1177/014920639101700108.

Barrales-Molina, V., Martínez-López, F. J., & Gázquez-Abad, J. C. (2014). Dynamic marketing capabilities: Toward an integrative framework. *International Journal of Management Reviews., 16*(4), 397–416. https://doi.org/10.1111/ijmr.12026.

Barreto, I. (2010). Dynamic capabilities: A review of past research and an agenda for the future. *Journal of management, 36*(1), 256–280.

Barro, S., & Davenport, T. H. (2019). People and machines: Partners in innovation. *MIT Sloan Management Review, 60*(4), 22–28.

Baumann, A., Coussement, K., Lessmann, S., & De Bock, K. (2015). Maximize what matters: Predicting customer churn with decision-centric ensemble selection. *ECIS, 2015*, 1–16.

Berk, L., Bertsimas, D., Weinstein, A. M., & Yan, J. (2019). Prescriptive analytics for human resource planning in the professional services industry. *European Journal of Operational Research, 272*(2), 636–641. https://doi.org/10.1016/j.ejor.2018.06.035.

Bi, W., Cai, M., Liu, M., & Li, G. (2016). A big data clustering algorithm for mitigating the risk of customer churn. *IEEE Transactions on Industrial Informatics, 12*(3), 1270–1281. https://doi.org/10.1109/TII.2016.2547584.

Boryczka, U., & Kozak, J. (2010). Ant colony decision trees-a new method for constructing decision trees based on ant colony optimization. *International conference on computational collective intelligence*, 373–382.

Boso, N., Danso, A., Leonidou, C., Uddin, M., Adeola, O., & Hultman, M. (2017). Does financial resource slack drive sustainability expenditure in developing economy small and medium-sized enterprises? *Journal of Business Research, 80*, 247–256.

Božič, K., & Dimovski, V. (2019). Business intelligence and analytics use, innovation ambidexterity, and firm performance: A dynamic capabilities perspective. *The Journal of Strategic Information Systems, 28*(4), 101578. https://doi.org/10.1016/j.jsis.2019.101578.

Brandt, T., Wagner, S., & Neumann, D. (2020). Prescriptive analytics in public-sector decision-making: A framework and insights from charging infrastructure planning. *European Journal of Operational Research, 291*(1), 379–393. https://doi.org/10.1016/j.ejor.2020.09.034.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. New York: Chapman & Hall.

Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011). Strength in numbers: How does data-driven decision making affect firm performance? *SSRN Electronic Journal.*. https://doi.org/10.2139/ssrn.1819486.

Buccieri, D., Raj, G. J., & Erin, C. (2020). International new venture performance: Role of international entrepreneurial culture, ambidextrous innovation, and dynamic marketing capabilities. *International Business Review, 29*(2), 101639.

Campbell, C., Sands, S., Ferraro, C., (Jody) Tsao, H. Y., & Mavrommatis, A. (2020). From data to action: How marketers can leverage AI. *Business Horizons, 63*(2), 227–243. https://doi.org/10.1016/j.bushor.2019.12.002.

Canhoto, A. I., & Clear, F. (2020). Artificial intelligence and machine learning as business tools: A framework for diagnosing value destruction potential. *Business Horizons, 63* (2), 183–193. https://doi.org/10.1016/j.bushor.2019.11.003.

Cao, L. (2017). Data science: A comprehensive overview. *ACM Computing Surveys, 50*(3). https://doi.org/10.1145/3076253.

Chen, Y., & Lin, Z. (2020). Business intelligence capabilities and firm performance: A study in china. *International Journal of Information Management*, 102232.

Chirico, F., & Nordqvist, M. (2010). Dynamic capabilities and trans-generational value creation in family firms: The role of organizational culture. *International Small Business Journal, 28*(5), 487–504. https://doi.org/10.1177/0266242610370402.

Chong, D., & Shi, H. (2015). Big data analytics: A literature review. *Journal of Management Analytics, 2*(3), 175–201. https://doi.org/10.1080/23270012.2015.1082449.

Chu, B. H., Tsai, M. S., & Ho, C. S. (2007). Toward a hybrid data mining model for customer retention. *Knowledge-Based Systems, 20*(8), 703–718. https://doi.org/10.1016/j.knosys.2006.10.003.

Ciampi, F., Demi, S., Magrini, A., Marzi, G., & Papa, A. (2021). Exploring the impact of big data analytics capabilities on business model innovation: The mediating role of entrepreneurial orientation. *Journal of Business Research, 123*(October 2020), 1–13. https://doi.org/10.1016/j.jbusres.2020.09.023.

Conboy, K., Mikalef, P., Dennehy, D., & Krogstie, J. (2020). Using business analytics to enhance dynamic capabilities in operations research: A case analysis and research agenda. *European Journal of Operational Research, 281*(3), 656–672. https://doi.org/10.1016/j.ejor.2019.06.051.

Coussement, K., & Van den Poel, D. (2008). Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques. *Expert Systems with Applications, 34*(1), 313–327. https://doi.org/10.1016/j.eswa.2006.09.038.

Coussement, K., Benoit, D. F., & Van den Poel, D. (2010). Improved marketing decision making in a customer churn prediction context using generalized additive models. *Expert Systems with Applications, 37*(3), 2132–2143. https://doi.org/10.1016/j.eswa.2009.07.029.

Côrte-Real, N., Tiago, O., & Ruivo, P. (2017). Assessing business value of big data analytics in european firms. *Journal of Business Research, 70*, 379–390.

D'Aveni, R. A., Canger, J. M., & Doyle, J. J. (1995). Utilizing the new 7S's framework. *The Academy of Management Executive., 9*(3).

Daramola, O., Oladipupo, O., & Musa, G. (2009). A data mining process framework for churn management in mobile telecommunication industry. *International Journal of Computer Sciences and Systems, 1*(1), 34–40.

Davenport, T. H., & Harris, J. G. (2017). *Competing on analytics: The new science of winning (Updated)*. Boston, Massachusetts: Harvard Business Review Press.

Davenport, T. H., & Ronanki, R. (2018). *Artificial intelligence for the real world*. Harvard Business Review.

De Caigny, A., Coussement, K., & De Bock, K. W. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. *European Journal of Operational Research, 269*(2), 760–772. https://doi.org/10.1016/j.ejor.2018.02.009.

De Caigny, A., Coussement, K., De Bock, K. W., & Lessmann, S. (2019). Incorporating textual information in customer churn prediction models based on a convolutional neural network. *International Journal of Forecasting, 36*(4), 1563–1578. https://doi.org/10.1016/j.ijforecast.2019.03.029.

Delen, D. (2020). *Prescriptive analytics: The final frontier for evidence-based management and optimal decision making*. Pearson Education.

Devriendt, F., Berrevoets, J., & Verbeke, W. (2020). Why you should stop predicting customer churn and start using uplift models. *Information Sciences*. https://doi.org/10.1016/j.ins.2019.12.075.

Dorigo, M., & Gambardella, L. M. (1997). Ant colony system: A cooperative learning approach to the traveling salesman problem. *IEEE Transactions on Evolutionary Computation, 1*, 53–66.

Di Stefano, G., Petraf, M. A., & Verona, G. (2014). The organizational drivetrain: A road to integration of dynamic capabilities research. *Academy of Management Perspectives, 28*(4). https://doi.org/10.5465/amp.2013.0100.

Dorigo, M., & Stützle, T. (2004). *Ant colony optimization*. Cambridge: MIT Press.

Dorigo, M., Birattari, M., & Stützle, T. (2006). Ant colony optimization - artificial ants as a computational intelligence technique. *IEEE Computational Intelligence Magazine, 1*, 28–39.

Døving, E., & Gooderham, P. (2008). Dynamic capabilities as antecedents of the scope of related diversification: The case of small firm accountancy practices. *Strategic Management Journal, 29*(8), 841–857. https://doi.org/10.1002/smj.683.

Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data - evolution, challenges and research agenda. *International Journal of Information Management, 48*, 63–71. https://doi.org/10.1016/j.ijinfomgt.2019.01.021.

Duan, Y., Cao, G., & Edwards, J. S. (2020). Understanding the impact of business analytics on innovation. *European Journal of Operational Research, 281*(3), 673–686. https://doi.org/10.1016/j.ejor.2018.06.021.

Dubey, R., Gunasekaran, A., Childe, S. J., Bryde, D. J., Giannakis, M., Foropon, C., & Hazen, B. T. (2020). Big data analytics and artificial intelligence pathway to operational performance under the effects of entrepreneurial orientation and environmental dynamism: A study of manufacturing organisations. *International Journal of Production Economics, 226*(December 2019), 107599. https://doi.org/10.1016/j.ijpe.2019.107599.

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., Medaglia, R., Le Meunier-FitzHugh, K., Le Meunier-FitzHugh, L. C., Misra, S., Mogaji, E., Sharma, S. K., Singh, J. B., Raghavan, V., Raman, R., Rana, N. P., Samothrakis, S., Spencer, J., Tamilmani, K., Tubadji, A., Walton, P., & Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002.

Eberhart, R., & Kennedy, J. (1995). A new optimizer using particle swarm theory, micro machine and human science, 1995. MHS'95. *Proceedings of the sixth international symposium on, IEEE*, 39–43.

Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: What are they? *Strategic Management Journal, 21*(10–11), 1105–1121.

Elgendy, N., & Elragal, A. (2014). *Big data analytics: A literature review paper. in industrial conference on data mining* (pp. 214–227). Springer. LNAI 8557.

Farah, M. F. (2017). Application of the theory of planned behavior to customer switching intentions in the context of bank consolidations. *International Journal of Bank Marketing*. https://doi.org/10.1108/IJBM-01-2016-0003.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010.

Fiorini, P. C., Seles, B. M. R. P., Jabbour, C. J. C., Mariano, B. M., & Jabbour, A. B. L. S. (2018). Management theory and big data literature: From a review to a research agenda. *International Journal of Information Management, 43*, 112–129.

Forkmann, S., Henneberg, S. C., Naudé, P., & Mitręga, M. (2016). Supplier relationship management capability: A qualification and extension. *Industrial Marketing Management, 57*, 185–200.

Fuller, C. M., Simmering, M. J., Atinc, G., Atinc, Y., & Babin, B. J. (2016). Common methods variance detection in business research. *Journal of Business Research, 69*(8), 3192–3198.

Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics, 165*, 234–246. https://doi.org/10.1016/j.ijpe.2014.12.031.

Ganesh, J., Arnold, M. J., & Reynolds, K. E. (2000). Understanding the customer base of service providers: An examination of the differences between switchers and stayers. *Journal of marketing, 64*(3), 65–87. https://doi.org/10.1509/jmkg.64.3.65.18028.

Ghasemaghaei, M., Khaled, H., & Turel, O. (2017). Increasing firm agility through the use of data analytics: The role of fit. *Decision Support Systems, 101*, 95–105. https://doi.org/10.1016/j.dss.2017.06.004.

Ghorbani, A., & Taghiyareh, F. (2009). CMF: A framework to improve the management of customer churn. *2009 IEEE asia-pacific services computing conference APSCC 2009*, 457–462. https://doi.org/10.1109/APSCC.2009.5394085.

Goodhue, D. L., Wixom, D. L., & Watson, H. J. (2002). Realizing business benefits through CRM: Hitting the right target in the right way. *MIS Quarterly Executive, 1*(2), 79–94.

Gordini, N., & Veglio, V. (2017). Customers churn prediction and marketing retention strategies: An application of support vector machines based on the AUC parameter-selection technique in B2B e-commerce industry. *Industrial Marketing Management, 62*, 100–107. https://doi.org/10.1016/j.indmarman.2016.08.003.

Greasley, A. (2019). *Simulating business processes for descriptive, predictive, and prescriptive analytics*. https://doi.org/10.1515/9781547400690.

Guo, Y., Liu, Y., Georgiou, T., & Lew, M. S. (2018). A review of semantic segmentation using deep neural networks. *International Journal of Multimedia Information Retrieval, 7*(2), 87–93. https://doi.org/10.1007/s13735-017-0141-z.

Gupta, S., Kar, A. K., Baabdullah, A., & Al-Khowaiter, W. A. A. (2018). Big data with cognitive computing: A review for the future. *International Journal of Information Management, 42*(April), 78–89. https://doi.org/10.1016/j.ijinfomgt.2018.06.005.

Hadden, J., Tiwari, A., Roy, R., & Ruta, D. (2007). Computer assisted customer churn management: State-of-the-art and future trends. *Computers & Operations Research, 34*(10), 2902–2917. https://doi.org/10.1016/j.cor.2005.11.007.

Hagendorff, T., & Wezel, K. (2020). 15 challenges for AI: Or what AI (currently) can't do. *AI and Society, 35*(2), 355–365. https://doi.org/10.1007/s00146-019-00886-y.

Helfat, C. E., & Peteraf, M. A. (2015). Managerial cognitive capabilities and the microfoundations of dynamic capabilities. *Strategic Management Journal, 36*(6), 831–850.

Hocutt, M. A. (1998). Relationship dissolution model: Antecedents of relationship commitment and the likelihood of dissolving a relationship. *International Journal of service industry management*.

Hong, J., Yibin, Z., & Minqiu, D. (2018). Sustainable supply chain management practices, supply chain dynamic capabilities, and enterprise performance. *Journal of cleaner production, 172*, 3508–3519.

Höppner, S., Stripling, E., Baesens, B., vanden Broucke, S., & Verdonck, T. (2020). Profit driven decision trees for churn prediction. *European Journal of Operational Research, 284*(3), 920–933. https://doi.org/10.1016/j.ejor.2018.11.072.

Huang, B., Kechadi, M. T., & Buckley, B. (2012). Customer churn prediction in telecommunications. *Expert Systems with Applications, 39*(1), 1414–1425. https://doi.org/10.1016/j.eswa.2011.08.024.

Huang, Y., & Kechadi, T. (2013). An effective hybrid learning system for telecommunication churn prediction. *Expert Systems with Applications, 40*(14), 5635–5647. https://doi.org/10.1016/j.eswa.2013.04.020.

Hung, S. Y., Yen, D. C., & Wang, H. Y. (2006). Applying data mining to telecom churn management. *Expert Systems with Applications, 31*(3), 515–524. https://doi.org/10.1016/j.eswa.2005.09.080.

Idris, A., Rizwan, M., & Khan, A. (2012). Churn prediction in telecom using Random Forest and PSO based data balancing in combination with various feature selection strategies. *Computers & Electrical Engineering, 38*(6), 1808–1819. https://doi.org/10.1016/j.compeleceng.2012.09.001.

Jahromi, A. T., Stakhovych, S., & Ewing, M. (2014). Managing B2B customer churn, retention and profitability. *Industrial Marketing Management, 43*(7), 1258–1268. https://doi.org/10.1016/j.indmarman.2014.06.016.

Jain, H., Khunteta, A., & Srivastava, S. (2020). Churn prediction in telecommunication using logistic regression and logit boost. *Procedia Computer Science, 167*, 101–112. https://doi.org/10.1016/j.procs.2020.03.187.

Janssen, M., van der Voort, H., & Wahyudi, A. (2017). Factors influencing big data decision-making quality. *Journal of Business Research, 70*, 338–345. https://doi.org/10.1016/j.jbusres.2016.08.007.

Ji-fan Ren, S., Fosso Wamba, S., Akter, S., Dubey, R., & Childe, S. J. (2017). Modelling quality dynamics, business value and firm performance in a big data analytics environment. *International Journal of Production Research, 55*(17), 5011–5026. https://doi.org/10.1080/00207543.2016.1154209.

Jiang, Z., Dan, W., & Jie, L. (2019). Distinct role of targeting precision of internet-based targeted advertising in duopolistic e-business firms' heterogeneous consumers market. *Electronic Commerce Research*, 1–22.

Kang, J., Diao, Z., & Zanini, M. T. (2020). Business-to-business marketing responses to COVID-19 crisis: A business process perspective. *Marketing Intelligence & Planning.*. https://doi.org/10.1108/MIP-05-2020-0217.

Karaboga, D. (2005). *An idea based on honey bee swarm for numerical optimization*. Erciyes university, engineering faculty, computer engineering department. Technical report-tr06.

Katayama, H., & Bennett, D. (1999). Agility, adaptability and leanness: A comparison of concepts and a study of practice. *International journal of production economics, 60*, 43–51.

Keaveney, S. M., & Parthasarathy, M. (2001). Customer switching behavior in online services: An exploratory study of the role of selected attitudinal, behavioral, and demographic factors. *Journal of the academy of marketing science, 29*(4), 374–390.

Keramati, A., Jafari-Marandi, R., Aliannejadi, M., Ahmadian, I., Mozaffari, M., & Abbasi, U. (2014). Improved churn prediction in telecommunication industry using data mining techniques. *Applied Soft Computing, 24*, 994–1012. https://doi.org/10.1016/j.asoc.2014.08.041.

Keramati, A., Ghaneei, H., & Mirmohammadi, S. M. (2016). Developing a prediction model for customer churn from electronic banking services using data mining. *Financial Innovation, 2*(1). https://doi.org/10.1186/s40854-016-0029-6.

Kim, N., Jung, K. H., Kim, Y. S., & Lee, J. (2012). Uniformly subsampled ensemble (USE) for churn management: Theory and implementation. *Expert Systems with Applications, 39*(15), 11839–11845. https://doi.org/10.1016/j.eswa.2012.01.203.

Kim, Y. S., Lee, H., & Johnson, J. D. (2013). Churn management optimization with controllable marketing variables and associated management costs. *Expert Systems with Applications, 40*(6), 2198–2207. https://doi.org/10.1016/j.eswa.2012.10.043.

Kim, K., Jun, C.-H., & Lee, J. (2014). Improved churn prediction in telecommunication industry by analyzing a large network. *Expert Systems with Applications, 41*(15), 6575–6584. https://doi.org/10.1016/j.eswa.2014.05.014.

Kim, S., Choi, D., Lee, E., & Rhee, W. (2017). Churn prediction of mobile and online casual games using play log data. *PLoS One, 12*(7), 1–19. https://doi.org/10.1371/journal.pone.0180735.

Kiron, D., & Schrage, M. (2019). Strategy for and with AI. *MIT Sloan Management Review, 60*(4).

Kisioglu, P., & Topcu, Y. I. (2011). Applying bayesian belief network approach to customer churn analysis: A case study on the telecom industry of Turkey. *Expert Systems with Applications, 38*(6), 7151–7157. https://doi.org/10.1016/j.eswa.2010.12.045.

Kohtamäki, M., Rodrigo, R., & Möller, K. (2018). Alliance capabilities: A systematic review and future research directions. *Industrial Marketing Management, 68*, 188–201.

Kozak, J., & Boryczka, U. (2013). *Dynamic version of the acdt/acdf algorithm for h-bond data set analysis* (pp. 701–710). In ICCCI.

Kozak, J., & Boryczka, U. (2014). Goal-oriented requirements for ACDT algorithms. *International conference on computational collective intelligence*, 593–602.

Kozak, J. (2019). *Decision tree and ensemble learning based on ant colony optimization*. Springer.

Kozlenkova, I. V., Samaha, Stephen, A., & Palmatier, R. W. (2014). Resource-based theory in marketing. *Journal of the Academy of Marketing Science, 42*(1), 1–21.

Kraus, M., Feuerriegel, S., & Oztekin, A. (2020). Deep learning in business analytics and operations research: Models, applications and managerial implications. *European Journal of Operational Research, 281*(3), 628–641. https://doi.org/10.1016/j.ejor.2019.09.018.

Krishna, G. J., & Ravi, V. (2016). Evolutionary computing applied to customer relationship management: A survey. *Engineering Applications of Artificial Intelligence, 56*, 30–59. https://doi.org/10.1016/j.engappai.2016.08.012.

Lee, H., Lee, Y., Cho, H., Im, K., & Kim, Y. S. (2011). Mining churning behaviors and developing retention strategies based on a partial least squares (PLS) model. *Decision Support Systems, 52*(1), 207–216. https://doi.org/10.1016/j.dss.2011.07.005.

Lee, E.-B., Kim, J., & Lee, S.-G. (2017). Predicting customer churn in mobile industry using data mining technology. *Industrial Management & Data Systems, 117*(1), 90–109. https://doi.org/10.1108/IMDS-12-2015-0509.

Lejeune, M. A. P. M. (2001). Measuring the impact of data mining on churn management. *Internet Research, 11*(5), 375–387. https://doi.org/10.1108/10662240110410183.

Lepenioti, K., Bousdekis, A., Apostolou, D., & Mentzas, G. (2019). Prescriptive analytics: A survey of approaches and methods. *In lecture notes in business information processing*, 339.

Lepenioti, K., Bousdekis, A., Apostolou, D., & Mentzas, G. (2020). Prescriptive analytics: Literature review and research challenges. *International Journal of Information Management, 50*, 57–70. https://doi.org/10.1016/j.ijinfomgt.2019.04.003.

Liang, F., Yu, W., An, D., Yang, Q., Fu, X., & Zhao, W. (2018). A survey on big data market: Pricing. *Trading and Protection. IEEE Access, 6*, 15132–15154. https://doi.org/10.1109/ACCESS.2018.2806881.

Liu, Y., & Zhuang, Y. (2015). Research model of churn prediction based on customer segmentation and misclassification cost in the context of big data. *Journal of Computer and Communications, 03*(06), 87–93. https://doi.org/10.4236/jcc.2015.36009.

Lopes, J., Guimarães, T., & Santos, M. F. (2020). Predictive and prescriptive analytics in healthcare: A survey. *Procedia Computer Science, 170*, 1029–1034. https://doi.org/10.1016/j.procs.2020.03.078.

Marsh, S. J., & Stock, G. N. (2003). Building dynamic capabilities in new product development through intertemporal integration. *Journal of product innovation management, 20*(2), 136–148.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: The management revolution. *Harvard Business Review, 90*(10), 60–68.

Meinzer, S., Jensen, U., Thamm, A., Hornegger, J., & Eskofier, B. M. (2016). Can machine learning techniques predict customer dissatisfaction? A feasibility study for the automotive industry. *Journal of Artificial Intelligence Research, 6*(1), 80. https://doi.org/10.5430/air.v6n1p80.

Mena, C. G., De Caigny, A., Coussement, K., De Bock, K. W., & Lessmann, S. (2019). *Churn prediction with sequential data and deep neural networks. A comparative analysis* (pp. 1–12).

Miguéis, V. L., Camanho, A., Falcão, J., & Cunha, E. (2013). Customer attrition in retailing: An application of multivariate adaptive regression splines. *Expert Systems with Applications, 40*(16), 6225–6232. https://doi.org/10.1016/j.eswa.2013.05.069.

Mihet, R., & Philippon, T. (2019). the economics of big data and artificial intelligence. *International Finance Review, 20*, 29–43.

Mikalef, P., & Pateli, A. (2017). Information technology-enabled dynamic capabilities and their indirect effect on competitive performance: Findings from PLS-SEM and fsQCA. *Journal of Business Research, 70*, 1–16.

Mikalef, P., Framnes, V. A., Danielsen, F., Krogstie, J., & Olsen, D. H. (2017). Big data analytics capability: Antecedents and business value. *Proceedings of the 21st pacific asia conference on information systems: ""Societal transformation through IS/IT"', PACIS 2017"*.

Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018). Big data analytics capabilities: A systematic literature review and research agenda. *Information Systems and E-Business Management, 16*(3), 547–578. https://doi.org/10.1007/s10257-017-0362-y.

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics capabilities and innovation: The mediating role of dynamic capabilities and moderating effect of the environment. *British Journal of Management, 30*(2), 272–298. https://doi.org/10.1111/1467-8551.12343.

Mikalef, P., Pappas, I. O., Krogstie, J., & Pavlou, P. A. (2020). Big data and business analytics: A research agenda for realizing business value. *Information & Management, 57*(1), 103237. https://doi.org/10.1016/j.im.2019.103237.

Mikalef, P., & Krogstie, J. (2019). Investigating the data science skill gap: An empirical analysis. *IEEE global engineering education conference, EDUCON, april-2019(april)*, 1275–1284. https://doi.org/10.1109/EDUCON.2019.8725066.

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research, 98*(May), 261–276. https://doi.org/10.1016/j.jbusres.2019.01.044.

Mikalef, P., & Krogstie, J. (2020). Examining the interplay between big data analytics and contextual factors in driving process innovation capabilities. *European Journal of Information Systems, 29*(3), 260–287.

Milošević, M., Živić, N., & Andjelković, I. (2017). Early churn prediction with personalized targeting in mobile social games. *Expert Systems with Applications, 83*, 326–332. https://doi.org/10.1016/j.eswa.2017.04.056.

Mitręga, M. (2019). Dynamic marketing capability-refining the concept and applying it to company innovations. *Journal of Business & Industrial Marketing, 35*(2), 193–203.

Mitręga, M., Spacil, V., & Pfajfar, G. (2020). Co-creating value in post-communists contexts: Capability perspective. *Journal of Services Marketing.* https://doi.org/10.1108/JSM-03-2019-0114.

Mohanty, S. A., & Vyas, S. (2018). *The economics of artificial intelligence*. Berkeley, CA: Apress.

Montavon, G., Samek, W., & Müller, K. R. (2018). Methods for interpreting and understanding deep neural networks. *Digital Signal Processing: A Review Journal, 73*, 1–15. https://doi.org/10.1016/j.dsp.2017.10.011.

Morgan, N. A. (2012). Marketing and business performance. *Journal of the Academy of Marketing Science, 40*(1), 102–119.

Morik, K., & Kopcke, H. (2004). Analysing customer churn in insurance data - A case study. *Lecture Notes in Computer Science, 3202*, 325–336.

Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. *Journal of Big Data, 2*(1), 1–21. https://doi.org/10.1186/s40537-014-0007-7.

Nassif, A. B., Shahin, I., Attili, I., Azzeh, M., & Shaalan, K. (2019). Speech recognition using deep neural networks: A systematic review. *IEEE Access, 7*, 19143–19165. https://doi.org/10.1109/ACCESS.2019.2896880.

Neslin, S. A., Gupta, S., Kamakura, W., Lu, J., & Mason, C. H. (2006). Defection detection: Measuring and understanding the predictive accuracy of customer churn models. *Journal of marketing research, 43*(2), 204–211. https://doi.org/10.1509/jmkr.43.2.204.

Ngai, E. W. T., Xiu, L., & Chau, D. C. K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications, 36*(2 PART 2), 2592–2602. https://doi.org/10.1016/j.eswa.2008.02.021.

Nie, G., Rowe, W., Zhang, L., Tian, Y., & Shi, Y. (2011). Credit card churn forecasting by logistic regression and decision tree. *Expert Systems with Applications, 38*(12), 15273–15285. https://doi.org/10.1016/j.eswa.2011.06.028.

Niebel, T., Rasel, F., & Viete, S. (2019). BIG data-BIG gains? Understanding the link between big data analytics and innovation. *Economics of Innovation and New Technology, 28*(3), 296–316. https://doi.org/10.1080/10438599.2018.1493075.

Olle, G. (2014). A hybrid churn prediction model in mobile telecommunication industry. *International Journal of e-Education, e-Business, e-Management and e-Learning, 4*(1), 55–62. https://doi.org/10.7763/ijeeee.2014.v4.302.

Olszak, C. M. (2014). Towards an understanding business intelligence. A dynamic capability-based framework for Business Intelligence. *2014 federated conference on computer science and information systems*, 1103–1110.

Olszak, C. (2020). *Business intelligence and big data: Drivers of organizational success*. CRC Press, Tylor & Francis Group. LLC.

Orzol, S. M., Hula, L., & Harrington, M. (2015). Program churning and transfers between medicaid and CHIP. *Academic Pediatrics, 15*(3), S56–S63. https://doi.org/10.1016/j.acap.2015.02.006.

Oskarsdóttir, M., Bravo, C., Verbeke, W., Sarraute, C., Baesens, B., & Vanthienen, J. (2017). Social network analytics for churn prediction in telco: Model building, evaluation and network architecture. *Expert Systems with Applications, 85*, 204–220. https://doi.org/10.1016/j.eswa.2017.05.028.

Parida, V., Lahti, T., & Wincent, J. (2016). Exploration and exploitation and firm performance variability: A study of ambidexterity in entrepreneurial firms. *International Entrepreneurship and Management Journal, 12*(4). https://doi.org/10.1007/s11365-016-0387-6.

Patel, P. C., Struckell, E. M., Ojha, D., & Manikas, A. S. (2020). Retail store churn and performance - The moderating role of sales amplitude and unpredictability. *International Journal of Production Economics, 222*, 107510. https://doi.org/10.1016/j.ijpe.2019.09.031.

Perrault, R., Shoham, Y., Brynjolfsson, E., Clark, J., Etchemendy, J., Grosz, B., … Niebles, J. C. (2019). *Artificial intelligence index 2019 annual report*. Standford University - Human-Centered Artificial Intelligence.

Peteraf, M., Di Stefano, G., & Verona, G. (2013). The elephant in the room of dynamic capabilities: Bringing two diverging conversations together. *Strategic Management Journal., 34*(12), 1389–1410.

Poornima, S., & Pushpalatha, M. (2020). A survey on various applications of prescriptive analytics. *International Journal of Intelligent Networks, 1*(July), 76–84. https://doi.org/10.1016/j.ijin.2020.07.001.

Popadiuk, S., & Choo, C. W. (2006). Innovation and knowledge creation: How are these concepts related? *International journal of information management, 26*(4), 302–312.

Prieto, I. M., Revilla, E., & Rodríguez-Prado, B. (2009). Building dynamic capabilities in product development: How do contextual antecedents matter? *Scandinavian Journal of Management, 25*(3), 313–326.

Raguseo, E. (2018). Big data technologies: An empirical investigation on their adoption, benefits and risks for companies. *International Journal of Information Management, 38*(1), 187–195. https://doi.org/10.1016/j.ijinfomgt.2017.07.008.

Randall, D. M., & Fernandes, M. F. (1991). The social desirability response bias in ethics research. *Journal of business ethics, 10*(11), 805–817.

Ranjan, J., & Foropon, C. (2021). big data analytics in building the competitive intelligence of organizations. *International Journal of Information Management, 56*(February 2020), 102231. https://doi.org/10.1016/j.ijinfomgt.2020.102231.

Riahi, Y., & Riahi, S. (2018). Big data and big data analytics: Concepts, types and technologies. *International Journal of Research and Engineering, 5*(9), 524–528. https://doi.org/10.21276/ijre.2018.5.9.5.

Risselada, H., Verhoef, P. C., & Bijmolt, T. H. A. (2010). Staying power of churn prediction models. *Journal of Interactive Marketing, 24*(3), 198–208. https://doi.org/10.1016/j.intmar.2010.04.002.

Roos, I., & Gustafsson, A. (2007). Understanding frequent switching patterns. *Journal of Service Research, 10*(1), 93–108. https://doi.org/10.1177/1094670507303232.

Sadat Mosavi, N., & Filipe Santos, M. (2020). How prescriptive analytics influences decision making in precision medicine. *Procedia Computer Science, 177*, 528–533. https://doi.org/10.1016/j.procs.2020.10.073.

Saradhi, V. V., & Palshikar, G. K. (2011). Employee churn prediction. *Expert Systems with Applications, 38*(3), 1999–2006. https://doi.org/10.1016/j.eswa.2010.07.134.

Sarkis, J., & Sundarraj, R. P. (2003). Managing large-scale global enterprise resource planning systems: A case study at Texas Instruments. *International Journal of Information Management, 23*(5), 431–442.

Sher, P. J., & Lee, V. C. (2004). Information technology as a facilitator for enhancing dynamic capabilities through knowledge management. *Information & Management, 41*(8), 933–945.

Schilke, O., Hu, S., & Helfat, C. E. (2018). Quo vadis, dynamic capabilities? A content-analytic review of the current state of knowledge and recommendations for future research. *The Academy of Management Annals Forthcoming, 1*. https://doi.org/10.5465/annals.2016.0014.

Shin, D. H., & Kim, W. Y. (2008). Forecasting customer switching intention in mobile service: An exploratory study of predictive factors in mobile number portability. *Technological Forecasting and Social Change, 75*(6), 854–874. https://doi.org/10.1016/j.techfore.2007.05.001.

Shirazi, F., & Mohammadi, M. (2019). A big data analytics model for customer churn prediction in the retiree segment. *International Journal of Information Management, 48*, 238–253. https://doi.org/10.1016/j.ijinfomgt.2018.10.005.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., & Hassabis, D. (2017). Mastering the game of Go without human knowledge. *Nature, 550*(7676), 354–359. https://doi.org/10.1038/nature24270.

Smith, K. A., Willis, R. J., & Brooks, M. (2000). An analysis of customer retention and insurance claim patterns using data mining: A case study. *Journal of the Operational Research Society, 51*(5), 532–541. https://doi.org/10.1057/palgrave.jors.2600941.

Stone, M., Aravopoulou, E., Ekinci, Y., Evans, G., Hobbs, M., Labib, A., … Machtynger, L. (2020). Artificial intelligence (AI) in strategic marketing decision-making: A research agenda. *Bottom Line, 33*(2), 183–200. https://doi.org/10.1108/BL-03-2020-0022.

Stripling, E., Antonio, K., Baesens, B., & Snoeck, M. (2018). Profit maximizing logistic model for customer churn prediction using genetic algorithms. *Swarm and Evolutionary Computation, 40*, 116–130. https://doi.org/10.1016/j.swevo.2017.10.010.

Tamaddoni Jahromi, A., Stakhovych, S., & Ewing, M. (2014). Managing B2B customer churn, retention and profitability. *Industrial Marketing Management, 43*(7), 1258–1268. https://doi.org/10.1016/j.indmarman.2014.06.016.

Tarafdar, M., Beath, C., & Ross, J. (2019). Using AI to enhance business operations. *MIT Sloan Management Review, 60*(4), 37–44.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal, 18*(7), 509–533.

Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal*.

Teece, D. J. (2012). Dynamic capabilities: Routines versus entrepreneurial action. *Journal of Management Studies*.

Wagner, E., & Hansen, E. N. (2004). A method for identifying and assessing key customer group needs. *Industrial Marketing Management, 33*(7), 643–655.

Wägar, K., Roos, I., Ravald, A., & Edvardsson, B. (2012). My customers are in my blind spot: Are they changing and I cannot see it? *Journal of Service Research, 15*(2), 150–165. https://doi.org/10.1177/1094670511435540.

Wamba, S. F., Gunasekaran, A., Akter, S., Ji-fan Ren, S. J., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research, 70*, 356–365. https://doi.org/10.1016/j.jbusres.2016.08.009.

Wang, Y., & Hajli, N. (2017). Exploring the path to big data analytics success in healthcare. *Journal of Business Research, 70*, 287–299.

Wang, L., Li, J., Guo, S., Xie, N., Yao, L., Cao, Y., Day, S. W., Howard, S. C., Graff, C., Gu, T., Ji, J., Gu, W., & Sun, D. (2020). Real-time estimation and prediction of mortality caused by COVID-19 with patient information based algorithm. *Science of the Total Environment*, 138394. https://doi.org/10.1016/j.scitotenv.2020.138394.

Wedel, M., & Kannan, P. K. (2016). marketing analytics for data-rich environments. *Journal of Marketing Research, 80*(6), 97–121. https://doi.org/10.1509/jm.15.0413.

Weerawardena, J., Mort, G. S., Liesch, P. W., & Knight, G. (2007). Conceptualizing accelerated internationalization in the born global firm: A dynamic capabilities perspective. *Journal of world business, 42*(3), 294–306. https://doi.org/10.1016/j.jwb.2007.04.004.

Vafeiadis, T., Diamantaras, K. I., Sarigiannidis, G., & Chatzisavvas, K. C. (2015). A comparison of machine learning techniques for customer churn prediction. *Simulation Modelling Practice and Theory, 55*, 1–9. https://doi.org/10.1016/j.simpat.2015.03.003.

Van Haver, J. (2017). *Benchmarking analytical techniques for churn modelling in a B2B context*. Iniversiteit Gent.

Vélez, D., Ayuso, A., Perales-González, C., & Rodríguez, J. T. (2020). Churn and net promoter score forecasting for business decision-making through a new stepwise regression methodology. *Knowledge-Based Systems, 196*, 105762. https://doi.org/10.1016/j.knosys.2020.105762.

Verbeke, W., Martens, D., Mues, C., & Baesens, B. (2011). Building comprehensible customer churn prediction models with advanced rule induction techniques. *Expert Systems with Applications, 38*(3), 2354–2364. https://doi.org/10.1016/j.eswa.2010.08.023.

Verbeke, W., Martens, D., & Baesens, B. (2014). Social network analysis for customer churn prediction. *Applied Soft Computing, 14*, 431–446. https://doi.org/10.1016/j.asoc.2013.09.017.

Vergne, J. P., & Durand, R. (2011). The path of most persistence: An evolutionary perspective on path dependence and dynamic capabilities. *Organization Studies, 32* (3). https://doi.org/10.1177/0170840610397485.

Verhoef, P. C., Venkatesan, R., McAlister, L., Malthouse, E. C., Krafft, M., & Ganesan, S. (2010). CRM in data-rich multichannel retailing environments: A review and future research directions. *Journal of Interactive Marketing, 24*(2), 121–137. https://doi.org/10.1016/j.intmar.2010.02.009.

Xai, G., & Jin, W. (2008). Model of customer churn prediction on support vector machine. *Systems Engineering - Theory & Practice, 28*(1), 71–77. https://doi.org/10.1016/S1874-8651(09)60003-X.

Yasmin, M., Tatoglu, E., Kilic, H. S., Zaim, S., & Delen, D. (2020). Big data analytics capabilities and firm performance: An integrated MCDM approach. *Journal of Business Research, 114*, 1–15. https://doi.org/10.1016/j.jbusres.2020.03.028.

Yu, X., Guo, S., Guo, J., & Huang, X. (2011). An extended support vector machine forecasting framework for customer churn in e-commerce. *Expert Systems with Applications, 38*(3), 1425–1430. https://doi.org/10.1016/j.eswa.2010.07.049.

You, Z., Si, Y.-W., Zhang, D., Zeng, X., Leung, S. C. H., & Li, T. (2015). A decision-making framework for precision marketing. *Expert Systems with Applications, 42*(7), 3357–3367.

Yuan, S., Bai, S., Song, M., & Zhou, Z. (2017). Customer churn prediction in the online new media platform: A case study on juzi entertainment. *2017 international conference on platform technology and service (PlatCon)*. https://doi.org/10.1109/PlatCon.2017.7883698.

Zdravevski, E., Lameski, P., Apanowicz, C., & Ślęzak, D. (2020). From big data to business analytics: The case study of churn prediction. *Applied Soft Computing, 90*. https://doi.org/10.1016/j.asoc.2020.106164.