Data Mining Approach Improving Decision-Making Competency Along the Business Digital Transformation Journey: A Literature Review

Hyrmet Mydyti(✉) and Arbana Kadriu

CST Faculty, South-East European University, Tetovo, Republic of Macedonia
{hm28315,a.kadriu}@seeu.edu.mk

Abstract. Advanced analytics and artificial intelligence are drivers of deep analysis and change in the perspective of businesses' digital transformation. Data mining, as an essential part of artificial intelligence, is a powerful digital technology, which provides guidance for businesses in terms of analyzing information and predicting in business. The key advantage of the application of the data mining approach in business is the impact by improving customers' experience and decision-making. The aim of this research is to present a theoretical model to understand the researchers' perspectives on data mining application in different business areas and digital transformation, and the discussion of some benefits and challenges of the data mining application in improving decision-making along the digital transformation of businesses. Moreover, this paper analyzes how the implementation of data mining techniques in business can lead to an increased efficiency and business productivity along their digital transformation journey.

Keywords: Advanced analytics · Digital transformation · Data mining in business · Decision-making

1 Introduction

Businesses are striving to adapt their strategies to the digital era, by incorporating novel technologies in their business models, which places more significance on the subject of processes and operations management, and, more essentially assesses their businesses' success of becoming digital [62, 81].

Advanced analytics (AA) and artificial intelligence (AI) are powerful digital technologies, which provide guidance for businesses, guidance on analyzing information, predicting and monitoring processes in business [83]. The analytics systems and intelligent applications used by businesses prove the importance of result delivery in improving decision-making and productivity, efficiency and effectiveness [5].

Data mining (DM), as a confluence of statistics and machine learning (ML), is the process of sorting through a high volume of information stored in repositories, corporate

M. H. Miraz et al. (Eds.): iCETiC 2021, LNICST 395, pp. 129–146, 2021.
https://doi.org/10.1007/978-3-030-90016-8_9

databases, and data warehouses to identify correlations, patterns, and trends and set relationships through data analysis [55, 67]. Through the DM approach, businesses will alleviate the process of reducing costs and enhancing customer experience along their digital transformation (DT) journey [8, 74]. Businesses will become more customer-oriented by advancing their services and saving their customers' time by strengthening their processes [12]. The issue of DM approach delivering linkages between businesses and customers has prompted to be highly significant for this research.

This research aims to analyze how the implementation of DM approach in businesses can lead to an improved decision-making and increased efficiency. The theoretical contribution and practical implication is to understand the researchers' insights and perspectives on DM along the business DT. We place two research questions: (i) RQ1. Which techniques and tools can be helpful if a business thinks about starting to use DM aiming to improve decision-making? (ii) RQ2. Which business areas have used and benefited from DM techniques along the business DT? Both questions have been answered using the literature review approach. After the introduction, the methodology for the analysis of DM research and studies is presented, followed by perspectives on DT, DT technologies, on advanced and big data analytics (BDA) techniques, on DM applications in business areas, benefits and challenges of applying DM, conclusion and recommendations on future work.

2 Methodology

The performed study covers the ensuing steps: perspectives on the DT, identification of DT technologies, perspectives on the DM and other advanced and big data analytics techniques, identification of DM application areas and identification of benefits and challenges of applying DM (Fig. 1).

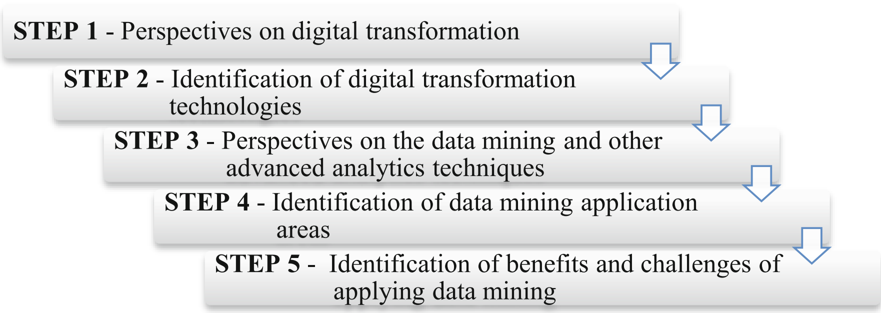

Fig. 1. Methodology workflow

3 Perspectives on Digital Transformation

3.1 Digital Transformation

DT encourages the integration of digital technologies and digital capabilities in innovative business models [68]. Certainly, the role of the digital maturity model (DMM), covering business dimensions, as an approach to empowering DT of businesses by evaluating the situation of the business transformation journey is a very important facet [14]. Agile has become a driving framework in the DT of businesses.

Reis et al. [65] outline DT as the practice of novel digital technologies with the aim of enabling main business enhancements such as improving customer experience, simplifying operations or developing new business models. Accordingly, DT surpasses justly digitizing resources and outcomes in value and incomes being produced.

Deloitte [14] emphasizes that one of the elements keeping the communications industry back from expansive growth in DT is the lack of a clear roadmap. The DMM is a useful means to enable guidance for a clear path throughout the transformation journey. In addition, China Mobile asserts that the DMM will be very helpful and, as a result, strengthens decision-making.

Morakanyane et al. [49] note two fundamental DT drivers such as digital capabilities and digital technologies. At the foundation of all DT attempts, are digital technologies. Digital technologies construct opportunities that businesses leverage. These transformational opportunities have the power to transform business models, operational processes and customer experiences. Thus, businesses benefit from the impact of DT.

Orfanidis [54] hints at the digital capabilities as analytical capabilities, business and IT integration, unified data and processes, and efficient delivery of solutions as the foundations contributing to the formation of digital capabilities in businesses. Morakanyane et al. [49] hint at the digital technologies such as mobile technologies, internet of things (IoT), cloud technologies, BDA, etc. as technologies that businesses adopt to improve their daily operations.

DT is considered as an evolutionary process since it changes with time, and the impacts bring a major change to the business. Similarly, digital technologies as fundamental drivers of DT are evolving. The main impact of DT is value creation. This value is accomplished by both the business and its customers [11, 49].

Assessing Progress of the Business Digital Transformation. The DMM enables decision-makers to give a view on matters of digital strategizing to evaluate and implement the required change to the target areas [4].

The results show that the DMM is to assess the digital capabilities mostly across some common business dimensions, such as customers, strategies, technologies, operations, organization and culture. The consulted papers on identifying most common dimensions of DT have been included above in Table 1.

The strategy, in the perspective of DT concentrates on how the business changes to grow its competitive strength through digital delivering in facility [80]. The customer experience concentrates on the need of tackling customers' demands, and benefits as a foundation for evolving digital service offerings [56]. Technology advances the success of strategy by supporting to collect, protect, analyze and utilize data to respond the

Table 1. Dimensions of digital transformation

Industry/field of study	Strategies	Customers	Technologies	Operations	Organizations and culture
Telecommunications [80]	√	√	√	√	√
SME [42]	√		√		√
Business [21]	√	√	√	√	√
Corporate [19]	√	√	√	√	√
Telecommunications [14]	√	√	√	√	√

demands of clients at low cost and expenses [29]. The operations dimension focuses on the capacities that bolster the service provision. Enhanced maturity within this context proves a better digitized and manageable operation [76]. The organization and culture dimension outlines as an organizational culture with governance and talent processes to bolster improvement along the DMM curve [56].

Driving Framework in Business' Digital Transformation. Agile, the modern approach, has become a driving framework in the DT of businesses. Agile serves as a unique tool to drive DT as it simplifies how technology is used for handling the operations of the business. Agile methodology is the guiding beacon to revolutionize the whole business [1, 77]. Gunasekaran et al. [25] highlight that agile provides stability and flexibility, continuous improvement, risk reduction, great communication and engagement, transparency and high quality.

O'Regan [51] introduces agile as a popular lightweight software development methodology. Agile claims to be more reactive to customer demands than traditional methods and its supporters consider that it leads in higher quality and productivity, and faster time to market and enhanced customer satisfaction.

3.2 The Future of Digital Transformation and Advanced Analytics

Davenport et al. [13] claim that in the future, many segments, such as business models, customer service options, and customer behaviors will be influenced and transformed by AI. AI intelligence will enable e.g., online retailers to predict and know customers' preferences and ship items to customers without an official order; as a result, AI will transform business strategies, business models, and customer behaviors and will be used in areas such as analytics and predictive behaviors.

Makridakis [40] asserts the impact of DT on businesses will be significant, resulting in well interconnected businesses with decision-making according to the exploitation of big data and strengthened, competition among businesses. In general, AI technologies will influence how businesses operate. In addition, DT of businesses has significant impact on different aspects of society, such as our lives and work, our shopping, our entertainment habits and our employment patterns.

McCormick et al. [45] assume how AI technologies will be assimilated into analytics practices, giving business users remarkable access to powerful insights that drive action. Digital technologies, AI, big data, and IoT, will grow businesses' access to data, broaden the variety of data that can be analyzed, and advance the level of sophistication of the resulting insight. An insights-driven business exploits and applies data and analytics at each possibility to distinguish its products and customer experiences.

4 Digital Transformation Technologies

Ziyadin et al. [38], list digital technologies as in Table 2 that change business models. Wiesböck and Hess [84] refer to technologies – social, mobile, analytics, and CC technologies as digital technologies. Hausberg et al. [28] consider the major technological areas, which enable DT. Those areas include, cyber-physical systems, IoT, cloud computing (CC), big data, AI, augmented and virtual reality. Schwertner [69] highlights that maturing digital businesses are focused on integrating digital technologies. The use of analytics empowers business decision-making and the dynamism of decision-making has to adapt due to altering needs and transforming technologies [78]. Telegescu [79] outlines that businesses have possibilities to enfold the advantages that the technologies of the digital economy bring. Finally, the results show that BDA is one of accelerators of DT of businesses. The consulted papers in identifying potential DT technologies have been included in Table 2.

Table 2. Digital transformation accelerators.

Author/s	Big data analytics	IoT	Cloud	Mobile	Social networks
Ziyadin *et al.* [38]	√	√		√	√
Wiesböck and Hess [84]	√		√	√	√
Hausberg *et al.* [28]	√	√	√		
Schwertner [69]	√	√	√	√	√
Telegescu [79]	√	√	√	√	

4.1 Cloud Technologies

Mazumdar and Alharasheh [44] present the key attributes of CC such as on-demand self-service through a secure portal, scalability and elasticity, pay per use, ubiquitous access and location-independent resource pooling. The on-demand self-service attribute is introduced as an independent service, without interacting with service providers in provisioning server, network, and storage capabilities. Scalability and elasticity attributes are introduced as increasing or reducing resources elastically to maintain cost efficiencies.

Neware and Khan [50] define CC as a model for offering adequate, on-demand network connection to a shared pool of configurable resources and introduce CC's three major service models. 1) Infrastructure-as-a-Service is addressed as an on-demand service of virtualization by offering virtual machine, storage infrastructure and network. 2) Platform-as-a-Service builds application or settings to build application by offering virtual machine, operating system, application and development structure. 3) Software-as-a-Service offers on-demand cloud-based foundation for software to the end user, as a whole package.

Quinn et al. [61], through a survey, show how cloud technology contributed to decision-making and improved decision-making in comparison to previous systems. The authors assert that CC provides benefits for decision-making, lowers cost, and alleviates systems administration.

Lastly, Benlian et al. [3] highlight how, on a societal and economic level, CC is an enabler of the DT of industries. CC offers the infrastructure that has driven other key digital trends comprising mobile computing, the IoT, big data, and AI, thus disrupting current business models, and powering DT.

4.2 Internet of Things

Pflaum and Gölzer [57] affirm that smart products, which are at the heart of the IoT, drive the future DT of business and change its business model.

Zeinab and Elmustafa [48] elaborate the IoT as a novel technology, which affords numerous applications to join things to things and humans to things. The applications enabled through the IoT include smart healthcare, homes, energy, cities and environments. The authors consider two major challenges of the IoT, and they are as follows: the coexistence with different networks and big data size of the IoT. The combination of data from several resources with the IoT makes possible the development of applications and advanced services that can combine situation and context awareness into the decision-making components.

Coetzee and Eksteen [10] provide predictions of the future of the IoT, the impact of the IoT on society and an overview of the challenges and highlight the fact that trust and privacy are likely to be the major obstacles in the IoT uptake. The key challenge areas are categorized as follows: 1) privacy, identity management, security and access control; 2) standardization and interoperability; 3) data deluge. Several interesting applications as identified are environment monitoring; intelligent environments; retail, logistics and supply chain management; and healthcare. The IoT is presented as a potential and as a supporter in DT processes.

4.3 Big Data Analytics

Elgendy and Elragal [17] refer to big data as datasets that are not only big, but also high in variety and velocity. More importantly, decision makers can gain valuable insights from such big data, ranging from day-to-day actions, which can be provided using BDA. BDA is the implementation of AA techniques on big data. Several advanced data analytics have been elaborated, such as social media analytics, sentiment analysis, advanced data visualization, etc. BDA provides opportunities in areas, such as fraud detection, customer intelligence, and supply chain management. Moreover, its advantages deliver to various domains.

Memon et al. [46] introduce BDA as a method for looking at big data to reveal hidden patterns, incomprehensible relationship and other important data that can be utilized on enhanced decisions. In addition, Hadoop is elaborated as an open source distributed processing framework that manages data processing and big data storage. Predictive analytics is emphasized as a subsequent operation whereupon it utilizes a range of measurable, displaying, information-mining, and ML strategies and verifiable information, along these lines allowing experts to make forecast customer behavior and other future developments.

Dremel et al. [15] consider that the DT of businesses includes establishing BDA capabilities and that the deployment of the establishment is a challenging process. BDA enable evidence-based decisions for digital business opportunities. Business benefit from the potential of BDA and BDA support evidence based decision-making and enable new digital services.

5 Advanced and Big Data Analytics Techniques Empowering Business Decision-Making

Elgendy and Elragal [18] present how BDA techniques can be implemented, in the context of DT, to enable business transformation and, specifically, to improve decision-making by uncovering hidden insights, and beneficial knowledge. The main aim of the decision makers of the business is to enhance decision-making, insights, and knowledge through the application of BDA techniques.

The authors, in their study [64, 58, 82, 70, 23], provide several advanced analytics (AA) techniques. Finally, the results show that DM, ML, and NLP are the key techniques of big data and AA of DT of the businesses. The consulted papers in identifying potential advanced and BDA techniques have been included in Table 3.

Table 3. Analysis on advanced analytics techniques

Authors	Machine learning	Text mining (NLP)	Data mining	Social network analysis	Visual analytics	Web mining	Statistics
Vivekananth & Baptist [82]	√	√	√	√			
Sadiku *et al.* [70]	√	√	√		√	√	
Galetsia *et al.* [23]	√	√	√	√	√	√	√
Rehman *et al.* [64]	√	√	√				√
Prabhu *et al.* [58]	√	√	√				

5.1 Data Mining

Reddy [63] presents DM as a knowledge discovery process (KDD) by analyzing huge amounts of data from different views and turning them into valuable information affecting different areas of human life, and incorporating business, science, etc.

Siguenza-Guzman et al. [73] categorize functions, or models, of DM based on the task done: association, classification, clustering, and regression. Three DM analysis techniques have been considered: classical statistics, AI, and ML. Unlike other techniques, DM benefits additionally include extracting patterns/trends, and predicting behavior. The authors describe the mining process as an interactive sequence of steps. The first step takes place the integration of data belonging to different sources and formats. In the second step, the cleansing process is applied. Additionally, the transformation of data into an appropriate format is applied. In the third step, knowledge is extracted from the transformed data. Lastly, knowledge is visualized to the user.

Sharma et al. [72] elaborate the implementation of DM projects by pursuing the KDD. The KDD process consists of a number of stages, such as business understanding, data understanding, data preparation, modelling, evaluation and deployment. In the modelling (DM) phase, different techniques are designated and applied. The evaluation phase is described as consisting of thoroughly evaluating the model and reviewing the steps that it accomplishes the business goals. Lastly, the deployment model is described as a formation of the model and usually is not the end of the project.

Data Mining Tools as Modern Solutions. Gergin et al. [24] point out that data science and DM tools are the modern solutions that the businesses are using in this DT age. Businesses improve the cost efficiency of quality control processes with the application of DM methods. Data are processed on DM tools for understanding and inspecting the patterns and relationships. The results on identifying potential DM tools and their advantages have been included in Table 4.

Table 4. Data mining tools [16, 33, 35, 36]

Data mining tool: type	Advantages
Weka (Java): ML	Easy to use – easy user interface, supports numerous DM tasks
Rapid Miner (Java): Statistical Analysis; DM; Predictive Analysis	Visualization - user-friendly GUI, enormous flexibility, offers procedures (such as attribute selection and outlier detection)
R (C, Fortran, R): Statistical Computing	Strong choice for DM tasks, very fast implementation of many machine learning algorithms, better graphics, has specific data types
Orange (C++, Python, C): ML; DM; Data visualization	Better debugger, shortest scripts, poor statistics

5.2 Machine Learning

ML is the study of computer algorithms and statistical models to execute tasks without explicit instructions, such as by using pattern recognition and inference [52]. The implementation of data-intensive ML techniques can be noticed in science, technology and commerce, as a result contributing to more evidence-based decision-making [34].

Simon et al. [75] introduce the classification of ML in two categories, such as supervised learning (SL) and unsupervised learning (UL). Accordingly, the authors tackle deep learning and big data as important fields. “Deep learning algorithms extract complex data patterns, across a hierarchical learning process by analyzing and learning vast reserves of unsupervised data (big data)”.

Bastanlar and Ozuysal [2] introduce the classification of ML tasks when one considers the desired output of a machine-learned system. UL techniques comprise only the input values in the training data, and the learning algorithm comprehends hidden structure in the training data dependent to them. SL techniques demand the value of the output variable for each training sample to be acknowledged.

5.3 Natural Language Processing

NLP is defined as a field of AI, computer science and linguistics preoccupied with the computers - human languages interacts [20].

NLP is vital because it helps us to construct models and processes which take chunks of information as input in the form of voice or text or both and manipulates them as per the algorithm inside the computer. Thus, the output of an NLP system processes speech as well as written text [32]. NLP applications compose numerous areas of studies, such as NLP text processing and summarization, machine translation, speech recognition, AI and expert systems, and so forth [2].

Friedman et al. [22] point out how NLP systems are becoming advanced to ease decision-making, besides supporting information and relations extraction. In addition to NLP techniques, through a method make NLP applications enable decision-making.

6 Data Mining Applications in Different Business Domains

The major goal of the overview of DM applications is to understand the researchers’ perspectives on DM implementation in different domains and DT. DM is implemented in different categories of businesses and impacts driving DT in different businesses [63]. The study of the application of DM is to alleviate the process of comprehending businesses in refining their strategies to a new technology epoch. As a result, a number of research papers are selected to help understand and explain DM in business domains of a) Retail, b) E-commerce, c) Banking and d) Manufacturing.

6.1 Data Mining in Retail

Chen et al. [9] present an analysis to help businesses better understand their customers and as a result convey customer-centric marketing more successfully. The authors demonstrated a case study in how business intelligence for an online retailer is built by tools of

DM techniques to obtain competitive advantages on the market. The study facilitates the process of understanding consumers in the perspective of profitability, and implements appropriate decision-making and marketing strategies, as e result accelerates DT efforts. Association analyses were supportive in the context of establishing customer-buying patterns.

Castelo-Branco et al. [7] build a body of knowledge, so that a project can utilize the techniques related with DM in retail sales, at the same time present concepts as market basket analysis, association rules and cross-selling and up-selling. The DM implementation success stories are also introduced, as follows: 1) The Hewlett-Packard analytics team implemented a manually driven pilot in the online store and call center. Ultimately, the pilot gave importance to the analytics and data-driven decision-making at Hewlett-Packard. 2) The Swedish interior giant IKEA featured image recognition and augmented reality and resulted in increased customer satisfaction and fewer returned products. 3) Macy, the upmarket department store, utilizes big data to provide a more intelligent customer experience. 4) Amazon Go is the technology that is assumed to guide the approach to the future of AI in retail. The essential concept behind is that it resembles a store that blossoms on the principle of no checkout requests. 5) The Starbucks CTO proposed to combine the transaction evidence with other inputs, like weather, promotions, etc. in order to provide better personalized service.

Kaur and Jagdev [37] research the influence of changes brought in the retail sector by big data. It is highlighted that retail is entirely reliant on BDA. The mining of customers analytics is to increase profits, increase growth and be competitive, whether it is in-store or online. Through digital technologies, as a fundamental driver of DT, retailers make well-informed decisions using online data. Next, it is demonstrated how ML takes big data to gather, to utilize and to predict important insight about customers spending patterns and how these patterns experience changes. Big data is intelligent to directly affect sales by collecting data about consumer's exact spending habits. For example, Amazon responds to the competitive market promptly because of its analytical platform, which provides dynamic pricing, and compared to other retailers, it makes this change about every 3 months. Similarly, Metro Group retailers use retail analytics to identify movement of goods within the stores to enable real time information to concerned store personnel and customers for their ease of use. Furthermore, Staples use Hadoop and big data technologies to predict sales by processing around 10 million data transactions per week across 1100 retail outlets in the US.

6.2 Data Mining in e-commerce

Ismail et al. [31] present the DM process for e-commerce. Benefits and challenges of DM are the topic within the paper. The tackled benefits of DM in e-commerce are related to planning, forecasting, basket analysis, segmentation, etc. The tackled e-commerce DM challenges are related to spider identification, data transformations, etc. Herein the concept of DM in e-commerce is elaborated as an integration of statistics, databases and AI with some subjects with the intention of better decision-making. CC, as a key of technology in the age of DT, in e-commerce is considered to effectively utilize resources and reduce costs for companies embracing efficient DM. Additionally, some common

DM tools are described in detail. The end product of DM builds a possibility for decision makers to be capable to pursue their buying customers' patterns, need trends and locations, bringing in the effective way of the strategic decision for the advancement and the revenue of the business.

Zhao [6] studies big data algorithms and their implementations in e-commerce to deliver some recommendations to e-commerce businesses along the implementation of big DM. In the age of DT, big DM perform an important role in the improvement of e-commerce and it is the future of global e-commerce. The challenges of ecommerce applications are presented such as the enhancement of DM algorithms, the enhancement of mobile data and social network mining. Mobile e-commerce is presented as fast developing and changing consumers' behavior and habits. Similarly, social network-related e-commerce is presented as a novel significant e-business model. The presented e-commerce models use big data technology, eliminate risks, analyze market state, make strategies, increase profits, etc. In the paper, eBay is introduced as a model of the largest online business website and DM is the priority. EBay's big data platform comprises of three layers such as the data platform, the data integration and the data access layer. In the end, the presented goals of big DM algorithms in e-commerce are predominantly for: 1) optimization of the platform to advance the customer e-shopping experience; 2) enhancement of the capacity of CRM to improve decision-making; 3) provision of personalized services to advance the e-sales; 4) provision of value-added service.

6.3 Data Mining in Banking

Preethi and Vijayalakshmi [59] study the different DM techniques that can be practiced in the banking area to improve its performance and reduce costs. The authors consider handling enormous transactional data and making decision is an essential task; on the other hand, manual and conservative processing of decision-making is retarding, time consuming and error prone. In this situation, DM techniques offer an effective form of processing and decision-making. In the paper, two areas of banking application tackled are CRM and fraud detection. In addition, DM, by involving processes makes it possible to use the information in applications such as fraud detection, market analysis, production control, science exploration, etc. A CRM system is mainly used by banks to build brand value and find and understand their customers' needs. In this case, DM techniques are used to discover the new customers by using clustering techniques and to retain the customers by using Apriori algorithms. Furthermore, algorithms are another topic that has been tackled. Fraud detection is the next important area elaborated since fraudulent actions are very concerning for banks and most banks are utilizing a hybrid approach to detect fraud patterns. In the end, in the context of DT, it is concluded that by implementing data-mining techniques in data processing and decision-making, banks increase their profit tremendously.

Hasheminejad and Khorrami [27] review DM techniques utilized for analyzing bank customers consequently to designate more effective marketing strategies. Elaboration is performed for certain DM techniques, which are used to support the businesses to make decisions. CRM is considered having the aim to increase the business relationships with customers in this paper. The study of the analysis on customers by implementing DM will be easier. Importance is given to revealing the customers and their demands, as a pillar of DT. Through customer clustering, banks can detect the customers' crucial attributes. Customer clustering is highlighted to be a facilitator on directing the implementation of design marketing strategies for each group of customers. A conduction of research on data sets of recent studies of different researchers is performed. DM techniques are reviewed and summarized in order to identify customers' behavior, which results in the banking industry gaining a competitive advantage.

6.4 Data Mining in Manufacturing

Harding and Shahbaz [26] review applications of DM in manufacturing. This paper has a target to exclusively represent the relation of DM to the manufacturing industry and how DM is becoming important in the face of DT. DM is applied to support decision-making processes. The study in this paper has essentially focused on algorithm applications. An in-depth study has been conducted related to the applications in areas such as manufacturing systems, decision support systems, etc.

Oliff & Liu [53] present the utilization of DM techniques and the advancement of their usage in Industry 4.0 ready factories. Hence, the methodology contoured embraces the principles of DM and utilizes them to help decision-making with regard to quality. Advanced data analytics and ML, as an integral part of DT, are crucial digital technologies and DM and knowledge discovery for utilization of enormous amount of data to comprehend the manufacturing processing are highlighted in the paper. Algorithms are the next subject matter. A study is conducted with the intention to elaborate the best way of utilization of DM to advance assembly and quality control processes and to apply into current systems without drawback. Most of the algorithms are based on Rule Based Learning where those function by means of mathematical relationships.

Dai et al. [11] bring an overview and discuss the demands and challenges of BDA in manufacturing IoT (MIoT). The analytics of MIoT, as a key driver of DT, offers benefits in manufacturing processes. Subsequently, data analytics plays a significant role in extracting information, estimating the upcoming events and predicting the increment of products. Consumer behavior prediction has a vital role in the manufacturing area. The life cycle of BDA for MIoT comprises of the three stages. The challenges of data analytics tackled include: 1) data temporal and spatial correlation, 2) efficient DM schemes and 3) privacy and security. This paper also develops a system prototype in order to prove the feasibility of distributed computing models in MIoT.

7 Benefits of Applying Data Mining

DM adoption is beneficial with regard to driving business digitalizing of processes as DM enables predictive analytics, improves decision-making, increases profitability, enables risk mitigation, and signifies customer behavior [68]. DT makes possible novel business models that depend on smart data analytics, which are applied to obtain new insights and better digitalized decision-making [66]. The study of the decision-making benefit of DM is to ease the process of encouraging businesses in adapting their strategies to a new digital epoch.

7.1 Improvement of Decision-Making

Milovic and Milovic [47] tackle the context of prediction in two phases such as the learning phase and the phase of decision-making. The phase of decision-making by people is usually considered not qualitative when there is a huge amount of data to be classified. As a result, insights obtained with the usage of DM practices can be useful to take effective decisions that will enhance the success of an organization. Pulakkazhy and Balan [60] assume that useful information is hidden within the volume of data, which can be utilized for decision-making processes if they are revealed. Businesses that apply DM techniques immensely benefit since interesting patterns and knowledge can be mined. Ltifi et al. [39], in their paper, propose a model with visual DM methods for supporting dynamic DM. Moreover, the authors note the importance of trends such as visualization, DM and dynamic decision-making. Integration of visual DM techniques in real-time decision-making is crucial for treating compound, complicated and temporal data, and application of DM techniques allows digitalizing the processes. As a result, decision makers are able to visually predict trends in the temporal data and their behaviors. Manita et al. [41] demonstrate the increasing interest of digitalization in businesses and how digital technology affects businesses. Moreover, the paper brings out the significance of implementing digital strategies to provide regulators with the necessary modifications in the context of DT. Data analytics and DM improve the effectiveness of decision-making and enable to predict the performance of businesses with a higher level of confidence.

8 Challenges and Limitations of Applying Data Mining

A few research papers studied will ease the process of revealing the challenges and limitations of applying DM in businesses. The study intends to contribute an approach on how to apply DM by taking into account the limitations coped with.

Zain and Rahman [43] list the five common challenges and disadvantages of applying DM such as technology, skills, inconsistent or missing data as well as privacy and data security. Poor technology utilized in DM prompts the poor handling the data, differently powerful technology prompts the more effective the DM processes as well as can avoid the problems regarding technology. Skills are crucial to handle and manage the vast amount of data, consequently the DM application. Inconsistent data or missing data could give a big and many negative impacts towards businesses. DM is a violation to privacy and a risk to data security due to the fact that it is a threat through the objective of preserving statistics secured and guarding against the intrusion on privacy.

Sharma et al. [71], in their paper list three limitations of DM as follows: security and privacy issues, and misusage of/unreliable information. Businesses collect information about their customers in many ways and the possibility of violating privacy of its users is possible. Security is considered a high-risk challenge since data are collected and the probability of hacking the collected data is a serious matter for businesses. And, lastly, the misuse of information/inaccurate information is considered to be very harmful and cause serious consequences in case it will be used for decision-making by unethical people.

Ikenna et al. [30] list challenges faced with implementing DM in a business as follows: problem of poor data quality, employee empowerment, data integrity and security issues and complexity of integration. DM techniques empower managerial decision-making processes in a business. The authors consider that applying DM techniques to a business problem entails disposal of high quality data, which are usually primary data produced within the business functions. Moreover, finding a DM expert for a DM project within a business can present a challenge. Furthermore, integrity and security are considered crucial challenges with any data collection that is shared and utilized in a business. Finally, complexity of integration is considered a critical issue because the ability to effectively integrate DM projects into the business processes is a difficult task.

9 Conclusions

An inclusive research of data mining has been carried out, and different researchers' perspectives, insights and overviews of domains of application of data mining approach have been provided, together with several important aspects of benefits and challenges of data mining along the business digital transformation journey.

The paper demonstrates advantages and impact of business digital transformation assessing progress - maturity model, and most common digital transformation dimensions. Moreover, the IoT, cloud and big data analytics, as the best suited digital transformation technologies, have been presented as accelerators of digital transformation of businesses. Herewith, data mining, machine learning and natural language processing are the main techniques of big data and advanced analytics of digital transformation of the businesses have been considered. In addition, Weka, RapidMiner, Orange and R have been outlined as important data mining tools.

More importantly, the paper introduces advanced analytics and big data techniques tools by demonstrating their significance and digital transformation as the necessary trend, which influence businesses. The challenges of applying data mining are as follows: technology, skills, problem of poor data quality, misuse of information/inaccurate information, complexity of integration as well as data security, and privacy. The benefits of data mining include facilitating the process of streamlining decision-making, increasing efficiency and business productivity and enhancing customer experience along their digital transformation journey. The data mining technology is being adopted in many domains of businesses for its benefits. The main business domains where data mining is being implemented include retail, e-commerce, banking and manufacturing.

This paper helps businesses in particular, as key beneficiaries, and managers, to understand the positive aspects of data mining along their business digital transformation. Our future research goal is to conduct a case study, based on empirical studies, collect and analyze data, and obtain feedback from a business case study on using the advantages of the data mining techniques and approach in improving decision-making competency along the digital transformation journey.

References

1. Balashova, E., Gromova, E.: Russian industrial sector in the conditions of the 4th industrial revolution. In: IOP Conference Series: Materials Science and Engineering, vol. 404, p. 012014. IOP Publishing (2018)

2. Baştanlar, Y., Özuysal, M.: Introduction to machine learning. In: Yousef, M., Allmer, J. (eds.) miRNomics: MicroRNA Biology and Computational Analysis. MMB, vol. 1107, pp. 105–128. Humana Press, Totowa, NJ (2014). https://doi.org/10.1007/978-1-62703-748-8_7
3. Benlian, A., et al.: The transformative value of cloud computing: a decoupling, platformization, and recombination theoretical framework. J. Manag. Inf. Syst. **35**(3), 719–739 (2018)
4. Boström, E., Celik, C.: Towards a maturity model for digital strategizing - a qualitative study of how an organization can analyze and assess their digital business strategy. Dpt. of informatics, UMEA Universitet. Academic Press (2017)
5. Bughin, J., et al.: Artificial intelligence - The next digital frontier?. https://apo.org.au/node/210501. Accessed 21 May 2021
6. Zhao, X.: A study on the application of big data mining in E-commerce. In: 2018 IEEE 4th ICCC, pp. 1867–1871. IEEE (2018)
7. Castelo-Branco, F., Reis, J.L., Vieira, J.C., Cayolla, R.: Business intelligence and data mining to support sales in retail. In: Rocha, Á., Reis, J.L., Peter, M.K., Bogdanović, Z. (eds.) Marketing and Smart Technologies. SIST, vol. 167, pp. 406–419. Springer, Singapore (2020). https://doi.org/10.1007/978-981-15-1564-4_38
8. Chen, J., et al.: Systems of Insight for Digital Transformation - Using IBM Operational Decision Manager Advanced and Predictive Analytics. IBM Redbooks, USA (2015)
9. Chen, D., et al.: Data mining for the online retail industry - a case study of RFM model-based customer segmentation using data mining. J. Database Mark. Cust. Strategy Manage. **19**(3), 197–208 (2012)
10. Coetzee, L., Eksteen, J.: The Internet of Things-promise for the future - an introduction. In: IST-Africa Conference Proceedings, pp. 1–9. IEEE, Botswana (2011)
11. Dai, et al.: Big data analytics for manufacturing internet of things - opportunities, challenges and enabling technologies. Enterp. Inf. Syst. **14**(9–10), 1279–1303 (2020)
12. Dias, J., et al.: Introducing the next-generation operating model. Introducing the Next-generation Operating Model. McKinsey and Company, New York, 41 (2017)
13. Davenport, T., Guha, A., Grewal, D., Bressgott, T.: How artificial intelligence will change the future of marketing. J. Acad. Mark. Sci. **48**(1), 24–42 (2019). https://doi.org/10.1007/s11747-019-00696-0
14. Deloitte: Digital Maturity Model - Achieving digital maturity to drive growth. https://www.tmforum.org/wp-content/uploads/2018/08/Deloitte-DMM.pdf. Accessed 21 May 2021
15. Dremel, C., et al.: How AUDI AG established big data analytics in Its DT. MIS Q. Exec. **16**(2), 81–100 (2017)
16. Dušanka, D., et al.: A comparison of contemporary data mining tools. In: IS 2017, 4(6). Serbia (2017)
17. Elgendy, N., Elragal, A.: Big data analytics: a literature review paper. In: Perner, P. (ed.) ICDM 2014. LNCS (LNAI), vol. 8557, pp. 214–227. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-08976-8_16
18. Elgendy, N., Elragal, A.: Big data analytics in support of the decision-making process. Procedia Comput. Sci. **100**(2016), 1071–1084 (2016)
19. Eremina, Y., et al.: Digital maturity and corporate performance - the case of the baltic states. JOItmC **5**(3), 54 (2019)
20. Fabian, R., Alexandru-Nicolae, M.: Natural language processing implementation on Romanian ChatBot. In: SMO 2009. WSEAS, Hungary (2009)
21. Felch, V., et al.: Maturity models in the age of Industry 4.0 - Do the available models correspond to the needs of business practice?. In: HICSS, pp. 5165–5174 (2019)
22. Friedman, C., et al.: Natural language processing - state of the art and prospects for significant progress, a workshop sponsored by the national library of medicine. J. Biomed. Inform. **46**(5), 765–773 (2013)

23. Galetsi, P., Katsaliaki, K., Kumar, S.: Big data analytics in health sector - theoretical framework, techniques and prospects. Int. J. Inf. Manage. **50**, 206–216 (2020)
24. Gergin, Z., et al.: Data Mining Approach for Quality Control Process Improvement. (2019)
25. Gunasekaran, A., et al.: Agile manufacturing - an evolutionary review of practices. Int. J. Prod. Res. **57**(15–16), 5154–5174 (2019)
26. Harding, A., Shahbaz, M., Kusiak, A.: Data mining in manufacturing - a review. J. Manuf. Sci. Eng. Acme **128**, 969–976 (2006)
27. Hasheminejad, H., Khorrami, M.: Data mining techniques for analyzing bank customers - a survey. Intell. Decis. Technol. **12**(3), 303–321 (2018)
28. Hausberg, P., et al.: Research streams on digital transformation from a holistic business perspective - a systematic literature review and citation network analysis. J. Bus. Econ. **89**(8–9), 931–963 (2019)
29. Hie, P.: Impact of transforming organizational culture and digital transformation governance toward digital maturity in Indonesian banks. Int. Rev. Manag. Mark. **9**(6), 51–57 (2019)
30. Ikenna, O., et al.: Review of data mining as a tool for organization's growth and productivity. IJCSMC **9**(3), 284–290 (2014)
31. Ismail, M., et al.: Data mining in electronic commerce - benefits and challenges. Int. J. Commun. Netw. Syst. Sci. **8**(12), 501 (2015)
32. Jain, A., et al.: Natural language processing. J. Comput. Sci. Eng. **6**(1), 161–167 (2018)
33. Janošcová, R.: Mining Big Data in WEKA. In: 11th International workshop on Knowledge Management, pp. 29–39. Slovakia (2016)
34. Jordan, I., Mitchell, M.: Machine learning - trends, perspectives, and prospects. Science **349**(6245), 255–260 (2015)
35. Jovic, A., et al.: An overview of free software tools for general data mining. In: 37th MIPRO, pp. 1112–1117. IEEE, Croatia (2014)
36. Kaur, K., Dhiman, S.: Review of data mining with Weka tool. Int. J. Eng. Comput. Sci. **4**(8), 41–44 (2016)
37. Kaur, R., Jagdev, G.: Big data in retail sector - an evolution that turned to a revolution. Int. J. Res. Stud. Comput. Sci. Eng. **4**(4), 43–52 (2017)
38. Ziyadin, S., Suieubayeva, S., Utegenova, A.: Digital transformation in business. In: Ashmarina, S.I., Vochozka, M., Mantulenko, V.V. (eds.) ISCDTE 2019. LNNS, vol. 84, pp. 408–415. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-27015-5_49
39. Ltifi, H., et al.: Enhanced visual data mining process for dynamic decision-making. Knowl. Based Syst. **112**, 166–181 (2016)
40. Makridakis, S.: The forthcoming AI revolution - Its impact on society and firms. Futures **90**, 46–60 (2017)
41. Manita, R., et al.: The digital transformation of external audit and its impact on corporate governance. Technol. Forecast. Soc. Change **150**, 119751 (2020)
42. Williams, et al.: Digital maturity models for small and medium-sized enterprises - a systematic LR. In: ISPIM Conference Proceedings, pp. 1–15. ISPIM, Italy (2019)
43. Zain, M.S.I.M., Rahman, S.A.: Challenges of applying data mining in knowledge management towards organization. Int. J. Acad. Res. Bus. Soc. Sci. **7**(12), 405–412 (2017)
44. Mazumdar, A., Alharahsheh, H.: Insights of trends and developments in cloud computing. SARJET 1(3). South Asian Research Publication (2019)
45. McCormick, J., et al.: https://dmcny.org/wp-content/uploads/attachments/Forrester_Predictions_2017_-Artificial_Intelligence_Will_Drive_The_Insights_Revolution.pdf. Accessed 29 May 2021
46. Memon, A., et al.: Big data analytics and its applications. AETiC **1**(1), 45–54 (2017)
47. Milovic, B., Milovic, M.: Prediction and decision making in health care using data mining. IJPHS **1**(2), 69–78 (2012)

48. Zeinab, M., Elmustafa, A.: Internet of things applications, challenges and related future technologies. World Sci. News **2**(67), 126–148 (2017)
49. Morakanyane, R., et al.: Conceptualizing digital transformation in business organizations - a systematic review of literature. In: 30th Bled eConference, vol. 21, p. 427–444. Slovenia (2017)
50. Neware, R., Khan, A.: Cloud computing digital forensic challenges. In: 2018 Second ICECA, pp. 1090–1092. IEEE (2018)
51. O'regan, G.: Concise Guide to Software Engineering. UTiCS. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-57750-0
52. OECD: Artificial intelligence in Society. https://ec.europa.eu/jrc/communities/sites/jrccties/files/eedfee77-en.pdf/. Accessed 29 May 2021
53. Oliff, H., Liu, Y.: Towards industry 4.0 utilizing data mining techniques - a case study on quality improvement. Procedia CIRP **63**(2017), 167–172 (2017)
54. Orfanidis, P.: Prominence of big data in the digital transformation era. School of Economics, Business Administration and Legal Studies, Thessaloniki – Greece (2018)
55. Palmer, A., et al.: Data mining - machine learning and statistical techniques. In: Funatsu, K. (Ed.) Knowledge-Oriented Appl. in Data Mining, pp. 373–396 (2011)
56. Paschou, T., Rapaccini, M., Peters, C., Adrodegari, F., Saccani, N.: Developing a maturity model for digital servitization in manufacturing firms. In: Anisic, Z., Lalic, B., Gracanin, D. (eds.) IJCIEOM 2019. LNMIE, pp. 413–425. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-43616-2_44
57. Pflaum, A., Gölzer, P.: The IoT and digital transformation - toward the data-driven enterprise. IEEE Pervasive Comput. **17**(1), 87–91 (2018)
58. Prabhu, R., et al.: Big data analytics. In: Big Data Analytics - Systems, Algorithms, Applications, pp. 1-23. Springer, Singapore (2019).https://doi.org/10.1007/978-3-662-43720-9_14
59. Preethi, M., Vijayalakshmi, M.: Data Mining in Banking Sector. Int. J. Adv. Netw. Appl. **8**(5), 1–4 (2017)
60. Pulakkazhy, S., Balan, S.: Data mining in banking & its applications - a review. J. Comput. Sci. **9**(10), 1252–1259 (2013)
61. Quinn, M., et al.: The effects of cloud technology on management accounting and business decision-making. Financ. Manage. **10**(6), 1–12 (2014)
62. Rachinger, et al.: Digitalization and its influence on business model innovation. J. Manuf. Technol. Manag. **30**(8), 1143–1160 (2019)
63. Reddy, C.: A review on data mining from past to the future. Int. J. Comput. Appl. **15**(7), 19–22 (2011)
64. Rehman, H., et al.: The role of big data analytics in industrial Internet of Things. Future Gener. Comput. Syst. **99**, 247–259 (2019)
65. Reis, J. et al.: Digital transformation - a literature review & guidelines for future research. In: WorldCIST'18, pp. 411–421. Springer, Cham. (2018).
66. Roedder, N., et al.: The digital transformation and smart data analytics - an overview of enabling developments and application areas. In: IEEE Big Data 2016, pp. 2795–2802. IEEE, Washington (2016)
67. Rohanizadeh, S., Bameni, M.: A proposed data mining methodology and its application to industrial procedures. J. Ind. Eng. **4**, 37–50 (2009)
68. Saeed, T.: Data mining for small and medium enterprises - a conceptual model for adaptation. Intell. Inf. Manag. **12**(05), 183 (2020)
69. Schwertner, K.: Digital transformation of business. Trakia J. Sci. **15**(1), 388–393 (2017)
70. Sadiku, N., Adebo, O., Musa, M.: Big data in business. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **8**(1), 160–162 (2018)
71. Sharma, B., et al.: Review on data mining - its challenges, issues and applications. Int. J. Curr. Eng. Sci. Res. **3**(2), 695–700 (2013)

72. Sharma, S., Osei-Bryson, K., Kasper, G.: Evaluation of an integrated knowledge discovery and data mining process model. Expert Syst. Appl. **39**(13), 11335–11348 (2012)
73. Siguenza-Guzman, et al.: Literature review of data mining applications in academic libraries. J. Acad. Librariansh. **41**(4), 499–510 (2015)
74. Sima, et al.: Influences of the Industry 40 revolution on the human capital development and consumer behavior - a systematic review. Sustainability, **12**(10), 4035 (2020)
75. Simon, A., et al.: An overview of machine learning and its applications. Int. J. Electr. Sci. Eng. **1**(1), 22–24 (2016)
76. Srivastava, U., Gopalkrishnan, S.: Impact of big data analytics on banking sector - learning for Indian banks. Procedia Comput. Sci. **50**, 643–652 (2015)
77. Betz, C., et al.: The impacts of digital transformation, agile, and DevOps on future IT curricula. In: SIGITE 2016, p. 106. ACM, New York (2016)
78. Teichert, R.: Digital transformation maturity - a systematic review of literature. Acta Univ. Agric. et Silvic. Mendelianae Brun. **67**(6), 1673–1687 (2019)
79. Telegescu, T.: IT in the workspace - the need for digital transformation. In: Proceedings of the International Conference on Business Excellence, vol. 12, pp. 952–965. Sciendo (2018)
80. Valdez-de-Leon, O.: A digital maturity model for telecommunications service providers. Technol. Innov. Manag. Rev. **6**(8), 19–32 (2016)
81. Verhoef, C., et al.: Digital transformation - a multidisciplinary reflection and research agenda. J. Bus. Res. **122**, 889–901 (2019)
82. Vivekananth, P., Baptist, A.: An analysis of big data analytics techniques. Int. J. Eng. Tech. Mgmt. Res. **5**(5), 17–19 (2015)
83. West, M., Allen, R.: https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/. Accessed 29 May 2021
84. Wiesbock, F., Hess, T.: Understanding the Capabilities for Digital Innovations from a Digital Technology Perspective. Arbeitsbericht. WIM, München (2018)