Big Data to Big Impact: Effect of Big Data in Modern Decision Making

R. Akshaya, Varsha Ravi, S. Raja Lakshmi, and R. Aparna(✉)

Department of IT, M.O.P. Vaishnav College for Women, Chennai, India
aparnar.csc@mopvaishnav.ac.in

Abstract. In today's emerging world of computer techniques and technologies of social media and internet, decision making has become a non-trivial task for all major industries and organizations. Due to the surging competition in the business world, a single inaccurate decision can lead to radical setbacks. Data needs to be analyzed in order to arrive at better decisions, thus improving the organization's performance and standards. This is achieved through Big Data which makes use of the process of business intelligence to elevate strategic and operational decision-making. This paper is an attempt to conduct a full-scale review on big data researches and its effects on the decision-making affair in an industry. We plan to consolidate previous findings on big data to provide more functional decision-making solutions and these data are further described using data analytics tools/predictive models to derive conclusions directing towards progressive judgements.

Keywords: Big data · Decision-making · B-DAD · SDN · Filtration

1 Introduction

Big Data can be defined as data sets whose size or type is beyond the ability of traditional relational databases to capture, manage and process the data with low latency [1–3] Characteristics of big data include volume, velocity, variety, veracity, variability, and value. Sources of data are becoming more complex than those for traditional data because they are being driven by artificial intelligence (AI), mobile devices, social media, and the Internet of Things (IoT). For example, the different types of data originate. Decisions that previously were based on guesswork, or on meticulously constructed models of reality, can now be made based on the data itself. Such Big Data analysis now drives nearly every aspect of our modern society, including health care, military, banking, education and more. There is a great scope of using large datasets as an additional input for making decisions. In this paper we consider the consequences of big data in decision making through the analysis of various findings and list our observations in the upcoming sections. The rest of the paper is organized as follows: Literature survey is given in Sect. 2. Benefits of Big data is elaborated in Sect. 3. Techniques and Frameworks are shown in Sect. 4. Application of big data is given in Sect. 5. Section 6 highlights the contributions of this research work. Finally, Sect. 7 concludes this paper.

S. Jayasingh et al. (Eds.): ICETBM 2023, AEBMR 242, pp. 363–373, 2023.
https://doi.org/10.2991/978-94-6463-162-3_32

2 Literature Review

Elisa Bertino of Purdue University and Philip Bernstein [1] described the challenges which includes, issues of scale, diverseness, lack of morphic content, error-handling, privacy & security concerns, promptness, provenance, and visualization, at all stages of the analysis conduit from data acquisition to result inference. They also highlighted how to endorse and coax fundamental research towards addressing these technical challenges. The paper by Samuel Wamba and Edwards [2] states that the review and taxonomy proposed in this study offer a probable useful starting point for the development of improved discernment into these aspects of emerging 'big data' research. This paper presents findings of a systematic review and a longitudinal case study that managers can use to unlock the power of big data along the cross-cutting theme. The findings show that the big data revolution is progressing and organizations should accept it in order to build superior capabilities which can become a decisive competitive advancement. Elgendy and Elragal [3] in their research, state that the theory and to the knowledge base by perusing the literature and accumulate various theories, methods, frameworks, and data analysis techniques from previous research in the knowledge base in order to build the Big-Data Analytics Decision framework (BDAD). Accordingly, it clusters several parts of these features, and integrates and assimilate them into a single framework. Based on the existing definitions of Big Data analyzed in the paper by Isitor Emmanuel and Clare Stanier from Staffordshire University, [4] understood Big Data as: The term Big Data illustrates a data condition in which extensible architectures assist the requisites of analytical and other implementations which process, with high velocity, high volume data which may have a variety of data formats and include high data acquisition. The paper by the Lulea University of Technology, Sweden has surveyed the domain of big data and examines the different approaches utilized for processing and analytics. It discusses the current trends for helping to understand the rapid increase in big data and different processing methods and data analytic techniques. Yin and Kaynak [5] have surveyed an emphasis and versatile topic on big data has the potential to attract an ever-growing attention of both the enterprises communities and the industry related management and finance section. Jeble, Kumari, and Patil, [6] they have presented a structural foundation for developing analytics potentiality on how this emerging knowledge can help small and medium organizations to compete using lesser resources. This foundation can be a starting point for further analysis, amplification, and future research opportunities. In this study by Li, Chen and Shang, [7] they have provided an outlook of industrial big data for intuitive decision-making, then introduce the application of big data-driven technology in intuitive manufacturing. They have also explained the existing problems and challenges in this field. This paper also introduces big data-driven technology into the field of intelligent decision-making in manufacturing. The findings by Quinn and.

Wilson [8] and many other researchers have examined a cross-sectional view of many problems and demonstrated that further research should investigate the impacts of Big Data across a range of organization types and industry sectors. The research by Huang and Wu, [9] stated a conceptual framework for Big-data driven Scientific Data Management (SDM) has been developed, which maps Big-data analytics, tools, and methodologies into the SDM process. The study was conducted based on the Big-Data thinking, and the interpretation is outstanding.

3 Benefits of Bigdata

The concept of Big Data has been around for a while, but it was not until recently that Big Data has revolutionized the economic world [4]. Most organizations have now come in terms with how they can capture the voluminous amount of data that streams into their businesses and apply analytics to transform it into actionable insights. The benefits of big data and analytics have made it a significant element for firms looking to harness their business potential [8].

Big Data has now become a new competitive advantage for companies to outclass their peers. It narrows the segmentation of customers and allows organizations to provide products or services more effectively. Companies need to deduce valuable insights from humungous datasets [2]. Timely analysis of distinct volumes of data together with real-time monitoring and forecasting of events impact the business operations and performance [1]. Big Data insights should integrate structured and unstructured data and analyze it. One real time example is Amazon.com. When you visit the Amazon site, the website immediately shows recommendations of products and deals available [3]. This shows that a very large amount of data is tracked by Amazon each second to show such real-time results as it is shown for each visitor. Today's businesses focus on technology in order to track every aspect of potential consumer's activities. This starts from the browsing pattern, it consists of the user's history, preferences, and almost everything customers are interacting with in the digital world. Big Data helps companies understand how customers regard their products and services. This insight allows companies to adapt finer marketing policies and uncover the customers buying patterns among different demographic and geographic communities [4]. It also provides better customer insight as it helps them to choose what is essential for them by providing clues based on the customer's cart.

3.1 Big Data and Decision-Making

In today's era, each organization is distinctive with respect to its structure and functioning. The organizations from different environments must be managed subsequently [8]. The vital role of the management is to make numerous decisions regarding the organization. The decisions can range from simple to complex and can have either high or low impact [6]. Ordinary to crucial decisions are being made systematically in every organization. High quality and speedy decision-making enhance the performance of an organization. The quality of the decisions must be exceptional for the organization to function in an efficient manner. These Decisions need to be made with extreme caution as a single wrong action can make more damage than envisioned. This includes analyzing tremendous amounts of data namely Big Data [7]. The time taken for analyzing the amount of data using traditional methods is huge. Various researches have come up with various techniques and frameworks to integrate the data from various sources and come up with possible solutions in a relatively faster speed. The frameworks and methods are discussed briefly in the coming sections [8].

3.2 Decision – Making for Next Generation

The potential decision-making process to be gained from the use of big data, as well as the challenges it will pose, will naturally differ from industry to industry [6]. It is expected that computer and electronic products which are used in information sectors, government sectors as well as finance and insurance sectors, helps the organization gain significantly from the use of big data. In general terms, the use of big data can unlock consequential values in such areas as product and market development, operational efficiency, market demand predictions, decision making, and customer relationship management [7]. In a recent study, the results of a survey carried out indicated that the serviceable objectives of the use of big data by the respondents were found to be as follows:

a) Customer-centric outcomes 49%.
b) Operational optimization 18%.
c) Risk/financial management 15%.
d) New business model 14%.
e) Employee collaboration 4%.

It is seen from the results that, for almost half the respondents, the most dominant conclusions expected from the use of big data are the customer-centric ones [7]. They would like to use information accumulated in various ways and forms for customer analytics; to understand the customers' needs and foresee future behaviors thus providing better service to them, which leads to a better decision-making process for the organizations. For instance, the sensors embedded in mobile devices are expected to send back customer insights hence enabling the manufacturer to produce newer models with consumer preferred functionalities. One great example is the apple iPhone, where every new version is more revolutionary and innovative compared to the preceding versions. More after, after sales services acts to take preventive measures before a failure occurs. By this way, big data can be used to improve the development of the next generation of products and services.

4 Techniques and Frameworks

4.1 B-DAD Framework

The B-DAD framework or Big - Data Analytics and Decision framework was developed to map big data tools, architecture, and analytics to different phases of decision making [3]. This framework considers the notion that if the data is big, hence is the analytics and decisions associated with it, thus the word "Big" is hyphenated. This framework consists of 3 phases – intelligence phase, design phase and choice phase. In the first phase, the intelligence phase, the data that is to be analyzed is collected. Mostly data from various distinct platforms that are relevant to the study are collected. Subsequently, in the second phase, the design phase, the model planning takes place [3]. The relationships that need to be identified are defined. In order to provide the relationship between the various attributes in the collected data, several models and analyses are used, some of which are visualization analysis, cluster analysis, correlation and regression analysis, social media

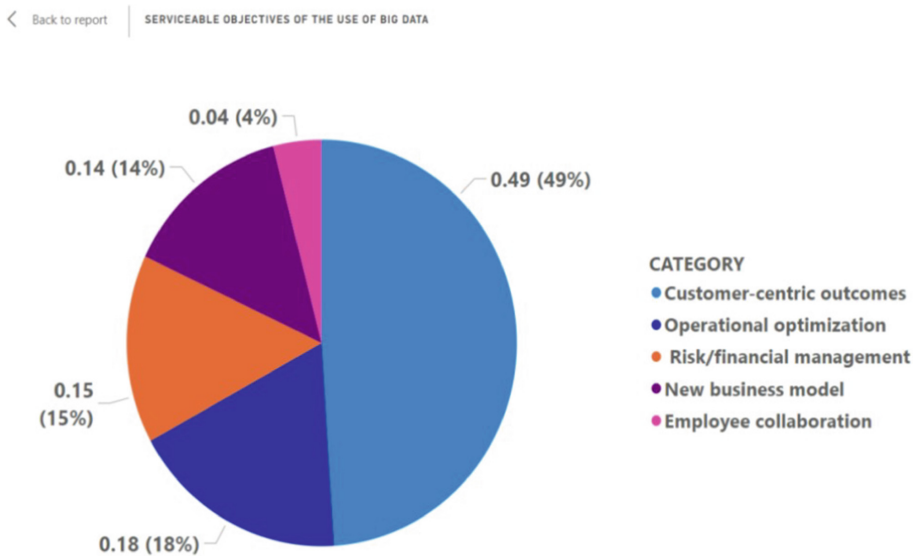

Fig. 1. Illustration of the above Statistics

analysis, text analysis, association analysis, sentiment analysis and decision trees. In the choice phase, a proposed solution from the results of the design phase is chosen. The resulting solution of the choice phase is implemented [3].

This framework is not intended to be a tool for making optimal, structured, and informed decisions; it is rather a conceptualization for mapping the different decision-making tools in the decision-making process and demonstrating how big data analytics can be used in the decision-making process. The B-DAD framework is an iterative process and not a sequential one, hence one should be able to move back and forth between the stages. Figure 1 shows the visualization of statistics.

4.2 SDM Framework

A scientific data management system (SDMS) is a centralized document management system which has the role of collecting, cataloguing, and storing data generated from a scientific lab. It is a system that has long held its position as the "archive and restoration" system. And it is basically designed in order to handle unstructured data generated from data systems like a Laboratory Information Management Systems (LIMS) or an ELN. LIMS has been considered as the ERP of Quality Systems. It helps in collecting thousands of data points month over month and year over year to make or assist in making those quality decisions.

Systems such as Chromatography data systems, generate millions of data points which gravitate to outpour databases overnight. This illustrated an approaching need to offload the data, not alone to generate space for new data, but also to perpetuate the search capability or overall regulation of the data retrieval, storage, and update at an upright elevation. As the vulnerability of the data needs quality domain to store

the data around for years prior to discarding, an application solution is required. Here SDMS plays a major role by serving its purpose, particularly by the inflexibility, ability of registering the data and its rehabilitation. With the help of SDMS, lab researchers, analysts, scientists, or other lab members can preferably stock and explore for lab data and knowledge to speed up workflows and to optimize it. SDMS can be deployed by any kind of laboratory such as clinical, contract, research, or pharmaceutical in order to achieve and ensure data integrity and executive consent and maintain a single source of truth. To get qualified for inclusion in SDMS category, a product should capture, stack, and maintain a variety of unstructured data formats to analyze stored data and to produce reports on lab activities. It supports data import and export and supports interoperability by engaging with lab instruments, systems, and databases to integrate with laboratory information management systems and electronic lab notebooks. SDMS groups data from instruments by computerized upload and reserves it centrally. This functionality is central to the SDMS and here in SDMS, simple file synchronization transfers data from instruments so that the limited storage space of the instrument can be emptied for new data. In other words, SDMS offers solutions to several problems. An SDMS should also be focused on improving research efficiency without relinquishing data sharing and collaboration accomplishments.

4.2.1 Benefits of SDMS Software

a) SDMS solutions provide orderly, standardized formats that make the interpretation of recorded data explicit and crisp. This ability is essential because cells, rows, columns, and other elements of documents and spreadsheets include metadata to have known, pre-defined specifications.
b) With the help of SDMS unstructured data like pdfs, images, videos can be easily rendered in different environments to provide decisions.
c) SDMS assists collaboration between colleagues from any part of the world.
d) SDMS tools have a systematize function, making it painless to approach and allot scientific data from any point while producing a secure and methodical space for storage.
e) It also upholds other computer systems and software in the laboratory.

4.3 Filtration Theory

To work with a big data set, companies should have the efficiency to collect and function large amount of data. In order to accomplish this, the organization should possess a planned IT framework to enable these large volume of data [8]. Organization should therefore have a complete big data architecture to facilitate Big Data analysis. How should organization design and set up their architecture to ease the process of Big Data? And what are the necessities from a storage and functioning perspective? Framework considers the technical proficiencies of Big Data environments. It examines the distinct roles that exist in the Big Data Architecture and looks at the best principles for design [8].

One among the commonly used big data framework is the filtration theory of data. The sections of filtration theory discuss similarities in applying filtration model to big data. The model, however did not develop to include the needed answer. We have modified his hybrid model by including a value step. This step incorporates with the filtration theory to filter data with analytics. Following sections discuss various filters [7].

4.3.1 Filter 1: Domain Related Filter

This filter checks for domain compatibility. CCF method brings both Content- based Filtering and Collaborative Filtering approaches together and to give a recommended suitable decision. HADOOP and Map reduce could also be useful in removing corrupted data using Text analytics, Key word, or Sematic analysis. Domain related filtering alone may not be enough to generate meaningful values or results. Data needs to be further filtered or cleaned to make it meaningful. Next section describes the. "truth" and "fake" data filtration.

4.3.2 Filter 2: Truth and Fake Filter

Once data becomes domain related issues, there is still problem regarding the "truth" and "trustworthiness" of the data. Given the "open" nature of social media it is possible to post fake and untruth data. Purpose of this step is to generate "trust" which is important factor in any data study. User must be assured that data is from trusted sources. There are many generally accepted governments and private trusted sources. However, big data also comes from many open public sources. Dell introduces a structure to identify content on social media and, thus, "predicting" future fake news topics. They use Facebook posts and identify topics that are liable to mislead with 77% accuracy [8]. Pennycook used crowdsourcing to get user's ratings of news outlet reliability. Once data is shown and approved the next step is to analyze it to create alternatives and eventually create "value." The following section defines analytics that can be used in value creation.

4.3.3 Value Creation

This step is to find a potential alternative match for developing a solution or value for the filtration theory. Using Value creation analytics, we can create meaningful intuitions. We discuss several methodologies to achieve more value. Visual analytics can be used to explore intricate contacts or to explain causes of disease spread [9]. Social network analysis can help government agencies to track or trace suspects. Numerous analytical techniques are available and can be used to create values. Table 1 shows the statistical report and Fig. 2 shows the corresponding visualization.

Table 1. Filtration Model Application Using Visual Analytics: Metropolitan Issues of Statistic Report of 2019 Serious Crime

District	AGG.Assault	Burglary	Homicide
Central	2993	450	234
Western	3456	563	345
Northern	4768	279	231
Eastern	2989	563	652
southern	1098	783	129
north-eastern	4800	457	234
north-western	4700	748	655
south-eastern	3560	437	341
South-western	2809	829	252

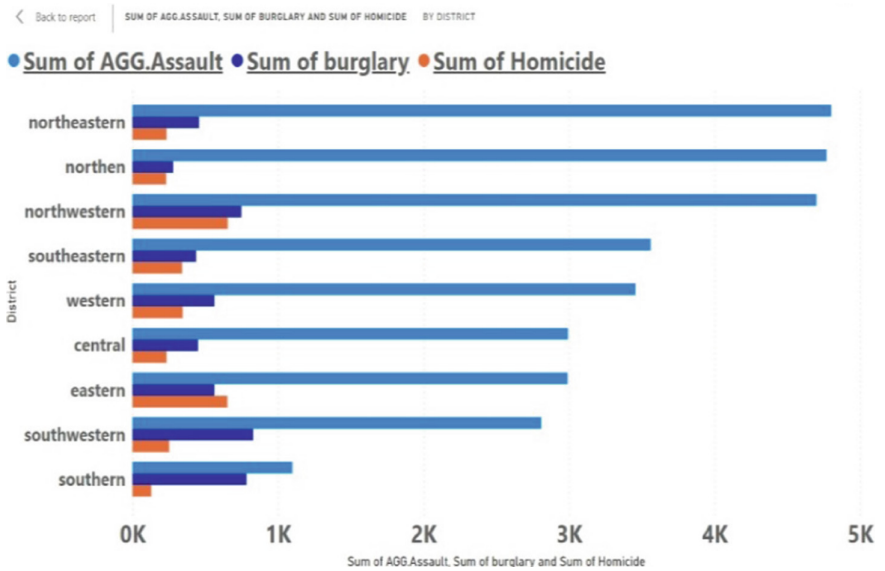

Fig. 2. Visual Analytics on Metropolitan Issues of Statistic Report of 2019 Serious Crime

5 Application of Big Data

In this small world with huge things to explore, Big Data is a real nerve wrecking game changer, The set-up of Big Data is arising and no one can gain expertise in one single sector. Big data analysts are fore front into the journey and even more data centric [1].

5.1 Banking

Western Union is a great example for big data efficiency sector. The company creates a channelized approach that personalize user experience by dealing more than 29 transactions per second and segregating the data in one platform for the purpose of statistical modelling and predictive interpretation [1].

5.2 Education

An example would be that of the University of Florida. The university uses IBM Info Sphere for acquisition, load, and transferring data through various resources and sectors. The university deploys IBM SPSS Modeler in study for predictive analytics and data modelling, and IBM Cognos Analytics for exploratory and projecting the routines of students.

5.3 Media

Big data plays a vital role in transmuting media platforms like Netflix. The search history and viewing history of the user, including the places where the user has paused the video, changes everything from the personalized thumbnails to the shows us watch on the "Popular on Netflix" category.

5.4 Healthcare

Mayo Clinic is a reliable example of Big Data in healthcare. The platform encircles bigdata analytics for helping in identifying various conditions and symptoms of patients and enhancing their life quality. These analytics can also identify at which risk the patients are at and grant them more wide-ranging health control and necessary medical intervention.

5.5 Agriculture

A Bayer Group unit, Bayer Digital Farming, uses machine learning and artificial intelligence for eradicating weed and harmful host growth. Farmers share pictures of weeds in one of the apps which is created, then match them against a large database to detect the species. This app tells the appropriate time, shielding the crops and improving yields.

5.6 Government

The Department of Homeland Security (DHS) is an important example of the government's role in Big Data. For confirming security, the Department of Homeland Security (DHS) uses an interruption-identifying system like sensors that has the ability to investigate internet traffic both in and out of Federal systems.

5.7 Retail

Amazon uses the Data found from their various sources to know about the customer's growth in their approval engine. The more Amazon learns about its customers, the better they can tell what they want to buy and surf [1]. Knowing about customer behavior, we can simplify the process and persuade the consumer to purchase it, like recommending products instead of making the users browse through the entire site catalogue.

6 Contributions of This Research Work

The main contributions of this research work are the discovery and comparison of various framework and analytic techniques concerned with Big Data such as the BDAD framework is a conceptualization for mapping the different decision-making tools. And we also highlight the usage of SDM helps the customers to make decision out of humungous data with the guarantee about privacy. CCF methodology brings out both Content- based Filtering and Collaborative Filtering approaches together and help us to conclude to a better suitable solution. It is recommended to eliminate corrupted data and tells us if the given data or the uploaded data is fake or truthful information. We have also made use of the Microsoft Power BI tool to give a visual representation of the statistics which we have conveyed. The data is collected and analyzed using various models and techniques proposing a variety of solutions to a certain concern. The best solution is chosen and it is converted to actionable decisions in the corporate. The key of this paper lies on the influence of Big Data in the process of Decisionmaking in the modern business world.

7 Conclusion

In this paper, we have studied various researches on big data and provided insights on the role of big data in decision-making. Studies show that emerging analytics in big data is significant on effectiveness and reliable on the decision-making process. There is sufficient supporting evidence to determine that data-driven approaches would be a growing research methodology in business operations. Countless domains can be influenced by big data. Numerous frameworks and analysis show that not only predictive analytics can be used to value and gain understanding but descriptive analytics can be applied as well. It is believed that big data analytics is of great significance in this era of data overflow and can provide unforeseen benefits to decision makers in various areas such as healthcare, military, education, banking and more. If properly exploited and applied, big data analytics has the potential to provide a basis for advancements, on the scientific, technological, and humanitarian levels. In the meantime, it is seen that, big data analytics has changed the requirement in the decisionmaking process to make strategic, tactical, and operational decisions. The role of big data in decision making leads companies into their main- stream business practice that helps them to predict future outcomes, optimize the supply chain and develop real- time decisions.

References

1. Agrawal, Divyakant; Bernstein, Philip; Bertino, Elisa; Davidson, Susan; Dayal, Umeshwas; Franklin, Michael; Gehrke, Johannes; Haas, Laura; Halevy, Alon; Han, Jiawei; Jagadish, H.V.; Labrinidis, Alexandros; Madden, Sam; Papakonstantinou, Yannis; Patel, Jignesh; Ramakrishnan, Raghu; Ross, Kenneth; Shahabi, Cyrus; Suciu, Dan; Vaithyanathan, Shiv; and Widom, Jennifer, "Challenges and Opportunities with Big Data 2011–1" Cyber Center Technical Reports (2011).
2. Samuel Fosso Wamba, Shahriar Akter, Andrew Edwards, Geoffrey Chopin, Denis Gnanzou,
3. How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study, International Journal of Production Economics,Volume 165,2015.
4. Ibrar Yaqoob, Ibrahim Abaker Targio Hashem, Abdullah Gani, Salimah Mokhtar, Ejaz Ahmed, Nor Badrul Anuar, Athanasios V. Vasilakos. Big data: From beginning to future,International Journal of Information Management,Volume 36, Issue 6, Part B,2016. Isitor Emmanuel and Clare Stanier. 2016. Defining Big Data. In Proceedings of the International Conference on Big Data and Advanced Wireless Technologies (BDAW '16). Association for Computing Machinery, New York, NY, USA, Article 5, 1–6. https://doi.org/10.1145/3010089.3010090S. Yin and O. Kaynak, "Big Data for Modern Industry: Challenges and Trends [Point of View]," in Proceedings of the IEEE, vol. 103, no. 2, pp. 143–146, Feb. 2015, https://doi.org/10.1109/JPROC.2015.2388958
5. Couldry, N., & Powell, A. Big Data from the bottom up. Big Data & Society, 1(2) 2014.
6. Jeble, S., & Kumari, S., & Patil, Y. Role of Big Data in Decision Making. Operations and Supply Chain Management: An International Journal, 11(1), 36-44 2017.
7. Chunquan Li, Yaqiong Chen, Yuling Shang,A review of industrial big data for decision making in intelligent manufacturing,Engineering Science and Technology, an International Journal,Volume29,2022,101021,ISSN2215-0986, https://doi.org/10.1016/j.jestch.2021.06.001.
8. Lang Huang, Chao Wu, Bing Wang, Qiumei Ouyang,Big-data-driven safety decisionmaking: A conceptual framework and its influencing factors,Safety Science,Volume 109,2018.
9. Alessandro Merendino, Sally Dibb, Maureen Meadows, Lee Quinn, David Wilson, Lyndon Simkin, Ana Canhoto,Big data, big decisions: The impact of big data on board level decision-making, Journal of Business Research,Volume 93,2018.