

Review Article

Credit Risk Measurement, Decision Analysis, Transformation and Upgrading for Financial Big Data

Wenshuai Wu 

Guangdong Academy of Decision Sciences, Sun Yat-sen University, Guangzhou, Guangdong 510275, China

Correspondence should be addressed to Wenshuai Wu; wuwsh8@163.com

Received 9 December 2021; Revised 12 April 2022; Accepted 15 April 2022; Published 2 May 2022

Academic Editor: Yu Zhou

Copyright © 2022 Wenshuai Wu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

There is no well-built theory on credit risk measurement and decision analysis for financial big data, and an effective and scientific evaluation system for them has not been formed. A review of them can contribute to grasping the abovementioned topics, understanding current issues, analyzing research problems, mastering research challenges, and predicting future research directions. Besides, this paper points out four research directions of credit risk measurement and decision analysis for financial big data. Moreover, this paper can provide some guidance directions and insights for practitioners, researchers, financial institutions, and government departments who have an interest in complex decision-making in big data.

1. Introduction

With the rapid development of Internet finance, cloud computing, the Internet of things, and big data techniques, the data generated by society from structured data, semi-structured data, and unstructured data, such as credit card data, e-commerce data, resume data, picture data, audio data, and video data, are proliferating at an unprecedented speed, which marks the arrival of the big data era [1–4]. In 2012, the United Nations issued a white paper on big data government affairs, which listed big data as a historical opportunity globally. In 2013, the United States proposed to invest USD 200 million to promote the development of big data industries for upgrading the “big data strategy” according to the national will. The strong advantages of big data have attracted more countries and regions to join the process of exploring and applying big data. In China, the 14th five-year development plan for big data has also been issued for emphasizing the need to focus on developing key technologies of big data, innovating big data service mode, cultivating big data product system, and supporting the development of big data industry. Research on financial big data is not only a requirement for the development of society, but it is also conducive to grasping the operational status in order to better promote the transformation and

upgrading of the structure of the financial industry and maintain regional financial stability [2].

The financial field is deeply involved in the calculation of big data events [5]. The financial credit system is not only an important part of the social credit system but also one of the most basic research fields [6]. Financial institutions employ credit scoring models to identify potential borrowers and to determine loan pricing and collateral requirements by measuring credit risk [7–9]. Effective measurement of credit risk is the core of credit risk management. Besides, the measurement of credit risk depends on massive data, which are difficult to improve the accuracy of analysis results [10, 11] because of complex characteristics such as multi-source heterogeneity, timeliness, sociality, and emergence. Therefore, for financial big data, such as bank customer information data, credit management data, and financial transaction information data, the problem of developing the effective technologies of credit risk measurement and decision analysis is a challenging and difficult issue [12].

Although some scholars are engaged in the research of financial big data, most of them have focused on theoretical and policy research [13–15]. For financial big data, there is still a lack of systematic and scientific theories and methods to measure credit risk, in order to explore the evaluation mechanism. Besides, there are also decision-making risks in

the choice of credit risk models in the big data environment. Moreover, big data not only bring unprecedented challenges to the traditional profit model, operation management, and customer service mode of financial institutions but also bring urgent needs to the transformation and upgrading of the financial industry. At present, the key research topics for financial institutions or even government departments are as follows: how to research and develop the effective technologies of credit risk measurement and decision analysis for financial big data to make the technological leap and promote the transformation and upgrading of the financial industry. So, this paper surveys the key technologies of credit risk measurement and decision analysis for financial big data, which are conducive to enhancing the driving force of technological innovation and promoting the upgrading of the financial industry.

The rest of this paper is organized as follows: Section 2 summarizes the related research. Section 3 introduces some research problems. In Section 4, some research challenges are proposed. In Section 5, future research directions regarding credit risk measurement and decision analysis for financial big data are presented. Finally, Section 6 concludes this paper.

2. Related Works

2.1. The Complexity of Big Data. The field of complexity science emerged at the end of the 20th century. However, the following questions persist: What is complexity? What is complexity science? There has always been some controversy surrounding this topic in academia. When the development of complexity science is discussed, big data technology has developed with vigor and vitality. Big data have the following five characteristics, known as the five Vs: (1) [16] Volume refers to a huge amount of data, more than 1000 T or even up to 1 billion T. (2) Variety refers to a wide range of data types, such as pictures, audio, video, network logs, and geographic location information. (3) Value indicates that the data purification value is low; that is, the data value density is relatively low. (4) Velocity refers to strong timeliness requirements, in that the effective information needs to be collected, processed, and fed back in time. (5) Veracity requires high accuracy, which reflects the quality of data [17, 18]. Complexity is one of the key factors restricting the efficiency and quality of big data processing, and it is an important factor for big data to become a scientific problem.

The representative achievement of complexity research is the “open complex giant system” theory, proposed by Qian [19, 20]. The core achievements of this theory are the “comprehensive integration method from qualitative to quantitative” and the “comprehensive integration seminar hall system of the man-machine combination.” After decades of development, the basic technology and method system are formed. In recent years, with the development of big data techniques, the practical process of applying big data has been gradually accelerated. Many countries have launched and implemented national big data strategies. The US has taken big data as a strategic resource and accelerated the sharing, opening, development, and application of data

resources to aid in industrial transformation and upgrading. Big data have brought unprecedented opportunities [13, 15, 21] for industrial transformation and upgrading. However, opportunities and challenges coexist.

At present, big data face three major theoretical problems [22–24]. The first is how to deal with big data effectively, the second is how to describe the complexity and uncertainty of big data, and the third is how to model the big data system. Based on cloud storage and cloud computing, how to use information technology to deal with unstructured and semistructured data effectively has become a cutting-edge scientific research issue of common concern. Besides, how to explore the effective methods and execute the system modeling for big data are important issues for knowledge discovery in big data. The subsequent problem is that of decision heterogeneity and data heterogeneity, the relationship between decision heterogeneity and data heterogeneity for knowledge discovery and management decision-making in big data should be studied in depth. Owing to the complexity of big data, this problem is clearly an important research issue, which presents new challenges to data mining theory and technology [7, 25–27]. At present, analysis of the complexity of big data has mostly focused on specific data sets and algorithm measurement levels. At the system level, the work of analyzing the complexity of big data, exploring its mechanism, and finding the basic methods is of great significance.

2.2. Credit Risk Measurement. Throughout the research process of credit risk measurement, from the credit risk measurement method initially relying on the subjective judgment of experts to the modern credit risk evaluation model based on statistics, operations research, and artificial intelligence, the credit risk measurement model has constantly developed and improved. The research on the credit risk measurement can be roughly divided into five models: experience judgment model, mathematical-statistical model, structural model, internal rating model, and artificial intelligence model.

In the experience judgment model, the credit scoring technology mainly relies on expert experience for scoring. Credit experts score applicants according to the applicant’s materials and some elements of credit analysis. The commonly used methods include the 5C element analysis method, 5P element analysis method, and 5W element analysis method [28]. The characteristics of the experience judgment model mainly adopt the expert analysis method. The credit person in charge makes a subjective judgment on the credit status of the business object and then makes credit decisions, which have strong subjectivity.

In the late 1960s [29], the mathematical-statistical model is gradually developed to avoid the subjective influence of empirical judgment. The representative models include the zeta model [1], discriminant analysis model [30], regression analysis model [31–33], mathematical programming model [34, 35], multiobjective optimization model [36–38], and decision tree model [39].

With the rapid development of the capital market, the nonintermediation of financing and securitization, and the

emergence of many financial innovation tools, the complexity of credit risk became more significant. To solve the development of the capital market and the complexity of credit risk, the structural model based on capital market theory and information science theory is put forward in the 1990s, such as the KMV model [40], credit metrics model [41], credit portfolio view model [42], and risk measurement model based on option pricing technology [43, 44].

The Basel Committee launches the draft of the New Basel Capital Accord in 2001, pointing out that qualified banks should use the internal rating model to measure credit risk. The Office of the Comptroller of the Currency (OCC) defines the key components and characteristics of the internal rating system in 2003.

There are many assumptions in the application of the previous models, but the real data are difficult to fit with the proposed assumptions. With the development of artificial intelligence and computer technology, the artificial intelligence model has begun to enter the field of credit risk measurement. The representative models include the expert system [45], neural network [46], genetic algorithm [47, 48], and support vector machine (SVM) [49, 50].

Research on the credit risk measurement has attracted great attention in academic and application circles, and some research results have been obtained [51–55]. Altman [29], based on Beaver [56], using the Bayesian discriminant idea and linear discriminant technology, established the famous Z-score model to judge whether an enterprise defaults or goes bankrupt. In 1977, the Z-score model was extended by Altman [29] to propose the zeta model. Ohlson [57] took the financial ratio as the index and applied the logistic regression model to predict enterprise default probability. Altman et al. [58] used a neural network to predict bankruptcy and found that the neural network model was superior to the multivariate discriminant analysis method in performance. Jarrow and Turnbull [59] employed an arbitrage-free valuation technique for pricing derivatives on financial securities subject to credit risk. Vapnik [60] applied the SVM method to financial crisis early warnings, because of its strong operability and high prediction accuracy, and recently, the SVM method has been actively researched. Based on the credit conversion matrix, Morgan used the idea of value at risk to calculate the volatility of enterprise value by considering the loss rate of enterprise default loans and proposed the Risk-Metrics model [41]. Yu et al. [61] proposed a modified least-squares SVM classification, called the C-VLSSVM classification model, to evaluate credit risk. Yu et al. [62] applied a multiscale neural network model to address financial crisis events. Li et al. [63] proposed a software process model to measure and manage credit risk, in which the risk management and cost control module help to improve the risk management in the software development process. Based on the theories and methods of multicriteria decision making (MCDM) and data mining, Kou and Wu [64] proposed an analytic hierarchy model (AHM) to solve the model selection problem of credit risk assessment. Florez-Lopez and Ramon-Jeronimo [65] developed a correlated-adjusted decision forest model for ensemble strategy evaluation in credit risk assessment. Sousa

et al. [66] proposed a new dynamic modeling framework to evaluate credit risk. Zhang et al. [24] proposed multiple instance learning models with transaction data for credit risk assessment. Yamanaka [67] presented a random thinning model for top-down-type credit risk assessment. Song et al. [68] indicated that ensemble learning methods perform better in default prediction than statistical techniques and individual classifiers. Huang et al. [69] proposed a GMM-based method to estimate model parameters and test model-implied restrictions for specification analysis of structural credit risk models. Lappas and Yannacopoulos [70] proposed a model combining genetic algorithm and expert knowledge for feature selection in credit risk assessment [71]. In summary, the research on the credit risk measurement model has mainly focused on the improvement and innovation of the above models.

Because most researchers focus on designing new methods and developing new models [10], they are rarely able to address and analyze the mining results in-depth, so many users cannot easily and effectively grasp and use them, resulting in wastage of knowledge and data resources [64]. Besides, although major commercial banks, such as the People's Bank of China, Bank of China, Industrial Commercial Bank of China, and China Construction Bank, have developed some credit scoring models suitable for the banks' business systems, the data information establishing these models lacks accuracy because of the difficulty of sharing data among banks. In addition, these models are customized for their own business systems, so they lack universality. Moreover, under the background of the financial big data era, many studies have not considered the big data environment and its characteristics, and big data have the characteristics of multisource heterogeneity, timeliness, sociality, and emergence, so research faces new and severe challenges, and the measurement of credit risk has become more and more complex. Furthermore, there is a risk of decision-making in the technology selection of big data. Therefore, there is no well-built theory on credit risk measurement and decision analysis for financial big data, and an effective and scientific evaluation system has not been formed.

3. Research Problems

So far, the research on credit risk measurement has attracted extensive discussion and study, and some related results have been achieved [54, 55, 72, 73]. However, most studies on credit risk measurement have mainly focused on the improvement and innovation of the models. At the same time, some researchers have focused on the introduction of model principles, and others have focused on the comparative analysis of various models. Besides, the data between financial institutions are difficult to share, and the reliable data information in the establishment of the model lacks accuracy. Moreover, these methods rarely combine the application domain background, decision-making objectives, background knowledge, and preferences of users.

Although some scholars are engaged in the research of financial big data, most of them have focused on theoretical

and policy research [13–15]. There is still a lack of unified standards for the acquisition, storage, and management of big data [74, 75]. For big data finance, there is also a lack of work to analyze big data using the basic theories and methods of system science, explore its mechanism, and find the basic evaluation system. In addition, there is a risk of decision-making in the model selection of big data. Besides, big data bring unprecedented challenges to the traditional profit model, operational management, and customer service mode of financial institutions. Therefore, this paper aims to excavate complex and heterogeneous massive data for financial institutions to develop key technologies for credit risk measurement and decision analysis to realize technological breakthroughs and improve financial risk prevention and control. Moreover, this paper can provide a decision-making basis for financial institutions and government departments to formulate relevant financial policies and realize the development strategy of the financial industry driven by technological innovation.

4. Research Challenges

In recent years, many large databases have been established in the financial industry, scientific research institutions, and government departments. Considerable multisource heterogeneous data have been accumulated. With the rapid development of computer technology, the era of big data has come. Big data can help to extract the value of data and make better decisions [76]. Big data finance has become an emerging hot research topic in recent years [22, 77]. In February 2011, Science magazine launched its first special issue on data processing: “Dealing with data,” which discussed a variety of problems related to the rapid growth of data. In December of the same year, it published another special issue to discuss the replicability and reproduction of massive data. Ke and Shi [7] pointed out that big data have attracted increasing attention from academia and the industry [23]. As the core of the modern economy, finance is an important force that promotes the economic development of a region and even a country. Financial institutions have more inherent advantages in big data applications: (1) financial enterprises can accumulate a large amount of high-value density data, such as liabilities, assets, capital revenue, customer identity, and payment transactions; (2) financial institutions have a relatively sufficient capital budget, which can not only attract high-end talent in the field of big data but also easily adapt the latest big data technology.

Based on the five Vs of big data and complexity characteristics, big data are mostly stored in nonrelational databases, such as Google’s BigTable, Amazon’s Dynamo, and Hadoop’s HBase, which have a flexible architecture design. In 2011, Ghoting et al. [78] proposed the SystemML and NIMBLE systems to design and implement a series of MapReduce primitives for data mining algorithms. Besides, algorithms, such as data clustering and classification, based on MapReduce have also been implemented and applied in many fields [32, 79]. Recently, the practical process of applying big data has gradually accelerated. From the early scientific research to applications in the fields of biology and

physics, it has begun to spread to more fields, such as the Internet and finance. In the era of big data, the competition among financial institutions is intensifying on the network information platform, where “data are king.”

Big data have brought unprecedented opportunities for the vigorous development of financial institutions [13]. From a marketing perspective, big data can more clearly and accurately obtain the needs and preferences of potential customers to attract customers through precision marketing or personalized intelligent recommendation. From the perspective of risk management, through deep mining of transaction data from multiple angles and channels, the financial industry can monitor financial risks in real-time and give early warnings of potential financial risks in time to reduce the cost of credit risk management. However, any financial activity has risks. For financial institutions, credit risk management is the key to the operation and management of financial institutions and has become the core issue in current social management. At present, although some scholars have engaged in research on financial big data, most of them have focused on theoretical and policy research [14], and there is still a lack of unified standards for the acquisition, storage, and management of big data [74, 75]. Besides, for financial big data, the work of using the basic theories and methods of system science to analyze big data, explore its mechanism, and find the basic evaluation system is also ambiguous. What’s more, how to research and develop the key technologies of credit risk measurement and decision analysis for financial big data to make the technological leap is a key research topic for financial institutions and government departments.

5. Research Directions

McKinsey’s research showed that the financial industry ranked first in the value potential index of big data. Big data have brought new opportunities and development impetus to the financial industry. The great significance of big data should be fully recognized. Future research directions are summarized as follows.

5.1. Theoretical Research on the Complexity Mechanism of Credit Risk Measurement for Financial Big Data. Big data have multisource heterogeneity, timeliness, sociality, and emergence; hence, research on financial credit risk measurement faces new and great challenges. However, most researchers have focused on designing new models or developing new algorithms [10], ignoring the big data environment and its characteristics. In view of this difficult problem, this research direction is to carry out theoretical research on the complexity mechanism of financial credit risk measurement. First, this research is to study the complexity and uncertainty of big data, the complexity of the big data process, and the complexity of the knowledge system involved in processing big data. Second, this research is to discuss and research the basic connotation of data, information, and knowledge to deeply research the characteristics of the financial big data industry and analyze the

challenges of the traditional credit risk measurement model. Finally, this research is to further discuss the essential feature of financial credit risk in the big data environment and carry out theoretical research on the complexity mechanism of credit risk measurement for financial big data.

5.2. Technology Development of Data Mining Models of Credit Risk Measurement. The complexity of big data is a key factor restricting the efficiency and quality of big data processing, and it is also an important factor for big data to become a key scientific problem. Facing the requirements of big data analysis, the traditional data processing, and analysis methods cannot meet the needs of data analysis in the big data environment. Three principles should be followed in big data analysis: all data, hybridity, and correlation. These are also the key characteristics that distinguish big data analysis from traditional data analysis. As an important technology of information processing, data mining aims to mine, extract, and identify valuable patterns, potential knowledge, and laws from massive data, in order to effectively guide business decision-making and scientific theoretical research. In this research direction, for financial big data, first, this research is to study the influencing factors of industrial upgrading based on the collection, processing, storage, and management technology of big data for determining the important factors, key areas, and key directions affecting big data innovation in the financial industry. Second, this research is to effectively screen and match data indicators in the data cloud platform from the three dimensions of personal credit, enterprise credit, and government credit to establish a credit risk measurement index system suitable for the financial industry. Finally, based on data mining models, relying on big data processing platforms (the big data offline processing platform, interactive processing platform, and stream processing platform), combined with Python and the big data ecosystem environment (Hadoop, Spark, HBase, Hive), this research is to develop new data mining models and technologies for credit risk measurement to give the system the advantages of scalability and high availability for seeking a technical breakthrough to solve the problem of credit risk measurement.

5.3. Method Development of Decision Analysis for Secondary Mining and Knowledge Discovery Based on MCDM. In the fields of data mining, machine learning, and big data analysis, many researchers have mostly focused on designing new models or developing new algorithms [10], and they have rarely been able to deeply process, analyze, and display the mined results. Therefore, it is difficult for users to understand the mined results, and the probability of mastering is lower, so the users cannot easily and effectively master and use the results, which causes the wastage of knowledge and data resources. Besides, the rough knowledge, generated by data mining, still needs to be filtered by decision-makers to obtain useful decision-making knowledge. Therefore, how to further integrate domain knowledge and rough knowledge into the decision-making knowledge to guide decision-making is a recognized and challenging scientific problem. If

the process of extracting “rough knowledge” through data mining or big data analysis is called “primary mining,” then the process of combining rough knowledge with quantified domain knowledge (e.g., expert experience, common sense, instinct, situational knowledge, and user preferences) to produce “decision knowledge” is called “secondary mining.” For financial big data, based on the MCDM method [25, 80, 81], this research direction is to carry out secondary mining and knowledge discovery from a multicriteria, multidimensional, multistage, and multistate perspective to develop new decision analysis methods for financial big data. What’s more, this research is committed to transforming the rough knowledge of data mining into the decision-making knowledge required by decision-makers to overcome the previous elite decision-making mode of “intuition and experience” which subverts the traditional research thinking mode and transforms “causal thinking” into “relevant thinking” in the big data era.

5.4. Countermeasure Research for the Transformation and Upgrading of the Financial Industry. With the transformation of economic structure to consumption structure, the original credit market cannot meet the demands of the social economy. In addition, there are obvious deficiencies in the coverage and availability of financial credit investigation data sources. The strong market demand poses a severe challenge to the financial industry. At the same time, in the era of big data, the competition among financial enterprises is more and more intense, and the traditional “intuition and experience” elite decision-making mode has gradually failed. Therefore, in this research direction, first, this research is based on industrial upgrading theory and economic growth theory, integrating cross scientific fields such as data mining, MCDM [82, 83], computer science, operations research, and artificial intelligence to carry out theoretical research on the operational mechanism of financial industry upgrading for financial big data. Second, the research is to combine the development status and characteristics of the financial industry in order to strive to explore the characteristics and operational law of big data in the development and management decision-making of the financial industry. Finally, the research is to further explore the development countermeasure suitable for the transformation and upgrading of the financial industry by providing big data management consulting and decision support services.

6. Conclusion

Financial credit risk analysis is the primary work and key link of risk management of financial institutions, which is related to the survival of financial institutions and social stability. The application of big data in the financial industry has attracted increasing attention. However, for the analysis of the financial credit risk for big data, no scientific evaluation and decision-making system has been formed. Besides, the research in this field has mostly focused on the improvement and innovation of the models and has rarely analyzed the mining results deeply. In addition, the research

has also rarely incorporated the corresponding application field background, decision-making objectives, background knowledge, and preferences of users. With the gradual deepening of economic globalization, the competition among enterprises is becoming fiercer and fiercer, and the traditional elite decision-making method of “intuition and experience” has gradually failed.

With the advent of the big data era, the cost of data acquisition is reducing. However, the amount of data is multiplying exponentially, and the data structure is becoming more and more complex. Human thinking has shifted from “causal thinking” to “relevance thinking” in the big data era. For financial big data, the work on using the basic theories and methods of system science to analyze financial credit risk, exploring its complex mechanism, and finding its evaluation system for financial big data is very meaningful. Therefore, based on the major needs of national economic and social development, this paper surveys theoretical research and technical development on credit risk measurement and decision-making analysis for financial big data. Besides, four research directions are pointed out as follows: first, considering new requirements and new challenges from the complex characteristics of multisource heterogeneity, timeliness, and the five Vs of big data, one direction is to perform theoretical research on the complex mechanism of credit risk measurement for financial big data. Second, because of the complexity of big data, another research direction is to develop new data mining models and technologies for credit risk measurement to achieve a technical breakthrough. Third, the results of mining are hard to understand, so another research direction is to carry out secondary mining and knowledge discovery for reconciling the knowledge gap, based on MCDM methods, in order to overcome the previous elite decision-making mode of “intuition and experience,” which subverts the traditional research thinking mode and transforms “causal thinking” into “relevant thinking” in the big data era. Fourth, given the industrial upgrading theory and economic growth theory, the last direction is to carry out countermeasure research on the transformation and upgrading of the financial industry by providing big data management consulting and decision support services. Moreover, the review can contribute to grasping the previous topics, understanding current issues, analyzing research problems, mastering research challenges, and predicting future research directions. Furthermore, this paper can also provide a decision-making basis for financial institutions and government departments to formulate relevant financial industry policies to promote a great leap forward in the development of the financial industry [84–88].

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no financial conflicts of interest related to the paper.

References

- [1] S. Zhang, W. Xiong, W. Ni, and X. Li, “Value of big data to finance: observations on an internet credit Service Company in China,” *Financial Innovation*, vol. 1, p. 17, 2015.
- [2] T. K. Chen, H. H. Liao, and W. H. Chen, “CEO ability heterogeneity, board’s recruiting ability and credit risk,” *Review of Quantitative Finance and Accounting*, vol. 49, no. 4, pp. 1005–1039, 2017.
- [3] Q. Zha, G. Kou, H. Zhang et al., “Opinion dynamics in finance and business: a literature review and research opportunities,” *Financial Innovation*, vol. 6, p. 44, 2020.
- [4] P. Jovanovic, S. Nadal, O. Romero, A. Abelló, and B. B. Quarry, “Quarry: a user-centered big data integration platform,” *Information Systems Frontiers*, vol. 23, no. 1, pp. 9–33, 2021.
- [5] M. M. Hasan, J. Popp, and J. Oláh, “Current landscape and influence of big data on finance,” *Journal of Big Data*, vol. 7, no. 1, pp. 21–17, 2020.
- [6] T. Kariya, Y. Yamamura, and K. Inui, “Empirical credit risk ratings of individual corporate bonds and derivation of term structures of default probabilities,” *Journal of Risk and Financial Management*, vol. 12, no. 3, pp. 124–129, 2019.
- [7] M. Ke and Y. Shi, “Big data, big change: in the financial management,” *Open Journal of Accounting*, vol. 03, no. 04, pp. 77–82, 2014.
- [8] P. K. Roy and K. Shaw, “A multi-criteria credit scoring model for SMEs using hybrid BWM and TOPSIS,” *Financial Innovation*, vol. 7, no. 1, p. 77, 2021.
- [9] M. Fleckenstein and F. A. Longstaff, “The market risk premium for unsecured consumer credit risk,” *The Review of Financial Studies*, 2022.
- [10] X. Q. Cheng, S. H. Liu, X. Q. Sun et al., “Combating emerging financial risks in the big data era: a perspective review,” *Fundamental Research*, vol. 1, no. 5, pp. 595–606, 2021.
- [11] W. H. Jiang, L. Xu, Z. S. Chen, K. Govindan, and K. S. Chin, “Financing equilibrium in a capital constrained supply Chain: the impact of credit rating,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 157, Article ID 102559, 2022.
- [12] L. Yu, L. Yu, and K. Yu, “A high-dimensionality-trait-driven learning paradigm for high dimensional credit classification,” *Financial Innovation*, vol. 7, no. 1, pp. 32–20, 2021.
- [13] M. E. Mccue and A. M. Mccoy, “The scope of big data in one medicine: unprecedented opportunities and challenges,” *Frontiers in Veterinary Science*, vol. 4, p. 194, 2017.
- [14] B. Tissot, *Financial Big Data and Policy Work: Opportunities and Challenges*, Publications Office of the European Union, Luxembourg, 2019.
- [15] Z. B. Zhou, M. Gao, H. L. Xiao, R. Wang, and W. B. Liu, “Big data and portfolio optimization: a novel approach integrating DEA with multiple data sources,” *Omega*, vol. 104, p. 102479, 2021.
- [16] A. N. Berger and W. S. Frame, “Small business credit scoring and credit availability,” *Journal of Small Business Management*, vol. 45, no. 1, pp. 5–22, 2007.
- [17] D. Howe, M. Costanzo, P. Fey et al., “Big data: the future of biocuration,” *Nature*, vol. 455, pp. 47–50, 2008.
- [18] A. Gandomi and M. Haider, “Beyond the hype: big data concepts, methods, and analytics,” *International Journal of Information Management*, vol. 35, no. 2, pp. 137–144, 2015.
- [19] J. Y. Yu, “Qian Xuesen’s research on open complex giant systems,” *Systems Engineering theory & Practice*, vol. 5, pp. 8–12, 1992.

- [20] N. Wang, "The making of an intellectual hero: Chinese narratives of qian xuesen," *The China Quarterly*, vol. 206, pp. 352–371, 2011.
- [21] K. Fanning and E. Drogat, "Big data: new opportunities for M&A," *Journal of Corporate Accounting & Finance*, vol. 25, no. 2, pp. 27–34, 2014.
- [22] Y. Shi, "Big data and new challenges for science and technology," *Science, Technology and Development*, vol. 10, no. 1, pp. 25–30, 2014a.
- [23] Y. Shi, "Big data history, current status, and challenges going forward," *The Bridge*, vol. 44, no. 4, pp. 4–11, 2014b.
- [24] T. Zhang, W. Zhang, Xu. Wei, and H. Hao, "Multiple instance learning for credit risk assessment with transaction data," *Knowledge-Based Systems*, vol. 161, pp. 65–77, 2018.
- [25] W. Wu, Z. Xu, G. Kou, and Y. Shi, "Decision-making support for the evaluation of clustering algorithms based on MCDM," *Complexity*, vol. 2020, no. 2, pp. 1–17, 2020b.
- [26] H. Y. Qiu, C. Liu, and X. M. Zhang, "Intelligent design of tennis player training schedule based on big data of complexity," *Complexity*, vol. 2021, no. 1, pp. 1–11, 2021.
- [27] M. E. Zmijewski, "Methodological issues related to the estimation of financial distress Prediction models," *Journal of Accounting Research*, vol. 22, pp. 59–86, 1984.
- [28] S. M. Abbadi and S. M. A. Karsh, "Methods of evaluating credit risk used by commercial banks in Palestine," *International Research Journal of Finance and Economics*, vol. 111, pp. 146–159, 2013.
- [29] E. I. Altman, "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy," *The Journal of Finance*, vol. 23, no. 4, pp. 589–609, 1968.
- [30] C. Fraley and A. E. Raftery, "Model-based clustering, discriminant analysis, and density estimation," *Journal of the American Statistical Association*, vol. 97, no. 458, pp. 611–631, 2002.
- [31] J. C. Wiginton, "A note on the comparison of logit and discriminant models of consumer credit behavior," *Journal of Financial and Quantitative Analysis*, vol. 15, no. 3, pp. 757–770, 1980.
- [32] F. Campobasso and A. Fanizzi, "Goodness of fit measures and model selection in a fuzzy least squares regression analysis," *Studies in Computational Intelligence*, vol. 465, pp. 241–257, 2013.
- [33] C. Jiang, Z. Wang, and H. Zhao, "A prediction-driven mixture cure model and its application in credit scoring," *European Journal of Operational Research*, vol. 277, no. 1, pp. 20–31, 2019.
- [34] Y. Peng, G. Kou, Y. Shi, and Z. Chen, "A multi-criteria convex quadratic programming model for credit data analysis," *Decision Support Systems*, vol. 44, no. 4, pp. 1016–1030, 2008.
- [35] Z. Chen, Y. Li, Y. Wu, and J. Luo, "The transition from traditional banking to mobile internet finance: an organizational innovation perspective—a comparative study of Citibank and ICBC," *Financial Innovation*, vol. 3, no. 1, p. 12, 2017.
- [36] G. Kou, Y. Peng, Y. Shi, M. Wise, and W. Xu, "Discovering credit cardholders' behavior by multiple criteria linear programming," *Annals of Operations Research*, vol. 135, no. 1, pp. 261–274, 2005.
- [37] Y. Shi, Y. Peng, G. Kou, and Z. Chen, "Classifying credit card accounts for business intelligence and decision making: a multiple-criteria quadratic programming approach," *International Journal of Information Technology and Decision Making*, vol. 04, no. 04, pp. 581–599, 2005.
- [38] W. Wu and Y. Peng, "Extension of grey relational analysis for facilitating group consensus to oil spill emergency management," *Annals of Operations Research*, vol. 238, no. 1–2, pp. 615–635, 2016.
- [39] N. O. Bakir, "A decision tree model for evaluating countermeasures to secure cargo at United States southwestern ports of entry," *Decision Analysis*, vol. 5, no. 4, pp. 230–248, 2008.
- [40] K. M. V. Corporate, *Credit Monitor Overview*, San Francisco California, 1993.
- [41] J. P. Morgan, "CreditMetrics-technical document," *Jp Morgan*, vol. 5, no. 3, pp. 48–51, 1997.
- [42] T. Wilsom, *Credit Risk Modeling: A New Approach*, Mckinsey Inc, New York, 1997.
- [43] F. A. Longstaff and E. F. Schwartz, "A simple approach to valuing risky fixed and floating rate debt," *The Journal of Finance*, vol. 50, no. 3, pp. 789–819, 1995.
- [44] Z. S. Chen, X. Zhang, K. Govindan, X. J. Wang, and K. S. Chin, "Third-party reverse logistics provider selection: a computational semantic analysis-based multi-perspective multi-attribute decision-making approach," *Expert Systems with Applications*, vol. 166, p. 114051, 2021.
- [45] A. Bahrammirzaee, A. R. Ghatari, P. Ahmadi, and K. Madani, "Hybrid credit ranking intelligent system using expert system and artificial neural networks," *Applied Intelligence*, vol. 34, no. 1, pp. 28–46, 2011.
- [46] J. A. Marshall, "Neural networks for pattern recognition," *Neural Networks*, vol. 8, no. 3, pp. 493–494, 1995.
- [47] J. Galindo and P. Tamayo, "Credit risk assessment using statistical and machine learning: basic methodology and risk modeling applications," *Computational Economics*, vol. 15, no. 1/2, pp. 107–143, 2000.
- [48] P. Pawiak, M. Abdar, J. Pawiak, V. Makarenkov, and U. R. Acharya, "DGHNL: a new deep genetic hierarchical network of learners for prediction of credit scoring," *Information Sciences*, vol. 516, pp. 401–418, 2020.
- [49] Z. Huang, H. Chen, C. J. Hsu, W. H. Chen, and S. Wu, "Credit rating analysis with support vector machines and neural networks: a market comparative study," *Decision Support Systems*, vol. 37, no. 4, pp. 543–558, 2004.
- [50] J. H. Trustorff, P. M. Konrad, and J. Leker, "Credit risk prediction using support vector machines," *Review of Quantitative Finance and Accounting*, vol. 36, no. 4, pp. 565–581, 2011.
- [51] M. Jacobs Jr, "The impact of asset price bubbles on credit risk measures," *Journal of Financial Risk Management*, vol. 04, no. 04, pp. 251–266, 2015.
- [52] W. Wu and D. G. Mcmillan, "Dynamic linkages in credit risk: modeling the time-varying correlation between the money and derivatives markets over the crisis period," *Journal of Risk*, vol. 16, no. 2, pp. 51–59, 2013.
- [53] A. Pichler, "A quantitative comparison of risk measures," *Annals of Operations Research*, vol. 254, no. 1–2, pp. 251–275, 2017.
- [54] R. Chen, Z. Wang, L. Yang, C. T. Ng, and T. C. E. Cheng, "A study on operational risk and credit portfolio risk estimation using data analytics," *Decision Sciences*, vol. 53, no. 1, pp. 84–123, 2020, <https://doi.org/10.1111/deci.12473>.
- [55] M. Mahbobi, S. Kimiagari, and M. Vasudevan, "Credit risk classification: an integrated predictive accuracy algorithm using artificial and deep neural networks," *Annals of Operations Research*, 2021.
- [56] W. H. Beaver, "Beaver financial ratios as predictors of failures," *Journal of Accounting Research*, vol. 4, pp. 71–102, 1966.
- [57] J. A. Ohlson, "Financial ratios and the probabilistic prediction of bankruptcy," *Journal of Accounting Research*, vol. 18, no. 1, pp. 109–131, 1980.

- [58] E. I. Altman, G. Marco, and F. Varetto, "Corporate distress diagnosis: comparisons using linear discriminant analysis and neural networks (the Italian experience)," *Journal of Banking & Finance*, vol. 18, no. 3, pp. 505–529, 1994.
- [59] R. A. Jarrow and S. M. Turnbull, "Pricing derivatives on financial securities subject to credit risk," *The Journal of Finance*, vol. 50, no. 1, pp. 53–85, 1995.
- [60] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, 1995.
- [61] L. Yu, S. Wang, and J. Cao, "A modified least squares support vector machine classifier with application to credit risk analysis," *International Journal of Information Technology and Decision Making*, vol. 08, no. 04, pp. 697–710, 2009.
- [62] L. Yu, S. Wang, K. K. Lai, and F. Wen, "A multi-scale neural network learning paradigm for financial crisis forecasting," *Neurocomputing*, vol. 73, no. 4–6, pp. 716–725, 2010.
- [63] J. Li, M. Li, D. Wu, and H. Song, "An integrated risk measurement and optimization model for trustworthy software process management," *Information Sciences*, vol. 191, no. 1, pp. 47–60, 2012.
- [64] G. Kou and W. Wu, "An analytic hierarchy model for classification algorithms selection in credit risk analysis," *Mathematical Problems in Engineering*, vol. 2014, no. 11–12, pp. 1–7, 2014.
- [65] R. Florez-Lopez and J. M. Ramon-Jeronimo, "Enhancing accuracy and interpretability of ensemble strategies in credit risk assessment. A correlated-adjusted decision forest proposal," *Expert Systems with Applications*, vol. 42, no. 13, pp. 5737–5753, 2015.
- [66] M. R. Sousa, J. Gama, and E. Brandão, "A new dynamic modeling framework for credit risk assessment," *Expert Systems with Applications*, vol. 45, pp. 341–351, 2016.
- [67] S. Yamanaka, "Random thinning model with a truncated credit quality vulnerability factor: application to top-down-type credit risk assessment," *International Journal of Financial Engineering*, vol. 06, no. 03, Article ID 1950024, 2019.
- [68] Y. Song, YY. Wang, X. Ye, D. J. Wang, Y. Q. Yin, and Y. Z. Wang, "Multi-view ensemble learning based on distance-to-model and adaptive clustering for imbalanced credit risk assessment in P2P lending," *Information Sciences*, vol. 525, no. 4, pp. 182–204, 2020.
- [69] J. Z. Huang, Z. Shi, and H. Zhou, "Specification analysis of structural credit risk models," *Review of Finance*, vol. 24, no. 1, pp. 45–98, 2020.
- [70] P. Z. Lappas and A. N. Yannacopoulos, "A machine learning approach combining expert knowledge with genetic algorithms in feature selection for credit risk assessment," *Applied Soft Computing*, vol. 107, Article ID 107391, 2021.
- [71] R. A. Jarrow and D. Lando, "A markov model for the term structure of credit risk spreads," *Review of Financial Studies*, vol. 10, no. 2, pp. 481–523, 1997.
- [72] D. E. Allen, R. J. Powell, and A. K. Singh, "Take it to the limit: innovative CVaR applications to extreme credit risk measurement," *European Journal of Operational Research*, vol. 249, no. 2, pp. 465–475, 2016.
- [73] M. Gubareva and M. R. Borges, "Rethinking economic capital management through the integrated derivative-based treatment of interest rate and credit risk," *Annals of Operations Research*, vol. 266, no. 1–2, pp. 71–100, 2017.
- [74] C. Liu and C. O. Sergeevna, "For intelligent debugging management of offshore oil engineering with big data," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 4, pp. 6091–6101, 2021.
- [75] S. Y. Wang, "An interview with Shouyang Wang: research Frontier of big data-driven economic and financial forecasting," *Data Science and Management*, vol. 1, no. 1, pp. 10–12, 2021.
- [76] A. Pérez-Martín, A. Pérez-Torregrosa, and M. Vaca, "Big Data techniques to measure credit banking risk in home equity loans," *Journal of Business Research*, vol. 89, pp. 448–454, 2018.
- [77] A. A. Bressan, "Big data as a revolutionary tool in finance," *Journal of Financial Innovation*, vol. 1, no. 2, pp. 1–7, 2015.
- [78] A. Ghoting, R. Krishnamurthy, E. Pednault et al., "SystemML: declarative machine learning on MapReduce," in *Proceedings of the 2011 IEEE 27th International Conference on Data Engineering (ICDE)*, pp. 231–242, Hannover, Germany, 2011.
- [79] J. J. Grefenstette, "Credit assignment in rule discovery systems based on genetic algorithms," *Machine Learning*, vol. 3, no. 2–3, pp. 225–245, 1988.
- [80] W. Wu and Z. Xu, "Hybrid TODIM method with crisp number and probability linguistic term set for urban epidemic situation evaluation," *Complexity*, vol. 2020, no. 4, pp. 1–11, 2020.
- [81] W. Wu, Z. Xu, and G. K. Kou, "Evaluation of group decision making based on group preferences under a multi-criteria environment," *Technological and Economic Development of Economy*, vol. 26, no. 6, pp. 1187–1212, 2020a.
- [82] G. Kou, Ö. Olgu Akdeniz, H. Dinçer, and S. Yüksel, "Fintech investments in European banks: a hybrid IT2 fuzzy multi-dimensional decision-making approach," *Financial Innovation*, vol. 7, no. 1, pp. 39–28, 2021a.
- [83] G. Kou, H. Xiao, M. Cao, and L. H. Lee, "Optimal computing budget allocation for the vector evaluated genetic algorithm in multi-objective simulation optimization," *Automatica*, vol. 129, p. 109599, 2021b.
- [84] M. Cao, R. Chychyła, and T. Stewart, "Big data analytics in financial statement audits," *Accounting Horizons*, vol. 29, no. 2, pp. 423–429, 2015.
- [85] W. S. Frame and L. Woosley, "Credit scoring and the availability of small business credit in low- and moderate-income areas," *Financial Review*, vol. 39, no. 1, pp. 35–54, 2004.
- [86] Z. Lin, A. B. Whinston, and S. Fan, "Harnessing Internet finance with innovative cyber credit management," *Financial Innovation*, vol. 1, p. 5, 2015.
- [87] X. Wu, X. Zhu, G. Q. Wu, and W. Ding, "Data mining with big data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 1, pp. 97–107, 2014.
- [88] W. Wu, G. Kou, Y. Peng, and D. Ergu, "Improved AHP-group decision-making for investment strategy selection," *Technological and Economic Development of Economy*, vol. 18, no. 2, pp. 299–316, 2012.