Proceedings of ICFTBA 2025 Symposium: Data-Driven Decision Making in Business and Economics
DOI: 10.54254/2754-1169/2025.BL29249

Identification of Borrower Credit Risk Based on Big Data Technology

Yashu Wang

Beijing Technology and Business University, Beijing, China
2024332002@email.cufe.edu.cn

Abstract. With the development of big data technology, the credit reporting field has ushered in new opportunities, but some problems have also emerged as a result..This study discusses the role and types of current data analysis and the application of big data in the credit reporting field from three aspects: big data technology, the influencing factors of loan risk, and assessment efficiency. The impact of different influencing factors on borrowers' credit scores: including external environments (such as credit scores, employment history and other historical model factors). Emerging influencing factors such as legal and social acceptance, and whether the borrower's own personality is overly optimistic, etc. After taking the above factors into account, this study explored the efficiency improvements of digital Windows compared to traditional Windows from the aspects of efficiency and customer satisfaction. At the same time, inclusive finance should be taken into consideration, and credit reporting standards for less developed countries and regions should be appropriately lowered.

Keywords: Data analysis, Financial credit reporting, Inclusive finance

1. Introduction

In the loan behavior of individual consumers and small and micro enterprises from financial institutions, the problem of bad debts urgently needs to be solved. Information asymmetry makes it impossible for financial institutions to promptly understand the true financial situation of borrowers and determine reliable credit ratings. Moreover, the low entry threshold of the white list makes it difficult for borrowers whose actual credit does not meet the standards to repay on time after successfully borrowing. This led to a further increase in the bad debt rate [1].

Today, with the increasing maturity of big data technology, financial institutions have valuable risk identification tools. However, financial institutions themselves are still in the early stage of technological development and encounter difficulties in systematically and comprehensively identifying and managing risks [2]. The organization, integration and analysis of massive correlated data by big data technology can greatly assist financial institutions in efficiently rating borrowers' credit ratings, serving subsequent decision-making and possible risk hedging strategies [3]. However, due to the problems of big data such as privacy leakage and insufficient quality inspection [4]. In the collection, processing and use of data. The protection of user privacy, as well as the

DOI: 10.54254/2754-1169/2025.BL29249

sources of data, data screening during the data processing and supervision by relevant departments, all need to be carefully considered.

This study aims to explore the role that the high efficiency, high precision and high relevance of big data technology can play in the credit assessment of borrowers and the risk identification brought about by financial institutions based on this assessment. However, given the current emphasis placed by the Internet finance industry on rural revitalization and inclusive finance, highlighting "prosperity" and sharing, financial development also shoulders the mission of providing innovative vitality to underdeveloped regions and enhancing people's well-being [5]. When dealing with borrowers from rural areas, the rating can be appropriately downgraded to protect the rights and interests of financial institutions while maintaining humanistic care.

2. Theoretical basis

Big data technology refers to the collection, cleaning, analysis, storage, etc. of data from different sources, including structured data and unstructured data. It features high speed and high accuracy [6]. Among them, the "analysis" stage is the key to making efficient and correct decisions. In the field of credit reporting, big data is often used to collect borrowers' credit reports, housing provident fund contribution data, etc. Because it can discover other data of borrowers, such as social media and customer transactions, and establish correlations with their credit reports, it has also become an important application to formulate personalized loan plans, segment borrower groups and implement different communication strategies [7]. For small and micro enterprises, debt restructuring and bound operation improvement should be adopted, and an exit mechanism should be designed. For individual consumers, measures such as stress-relieving communication and income reconstruction plans should be adopted. The conclusion of the communication strategy is based on data analysis. The following table 1 shows several data analysis frameworks.

Table 1. Types of data analysis

Data analysis type	Explanation	Example
Descriptive analysis	Provide insights based on past information. What's happening?	It is used for generating standard reports, as well as basic spreadsheet functions such as counting, summing, averaging, percentage changes, and vertical and horizontal analysis of financial statements
Diagnostic analysis	Explore the causes of past results. Why did it happen?	Used for difference analysis and interactive dashboards to explore the causes of past results.
Predictive analysis	By identifying patterns in historical data, it helps understand the future and provides forward-looking insights. What will happen? When did it happen? Why did it happen?	It can be used to predict the accounts receivable balance and collection cycle of each customer, and models with indicators can also be developed to prevent control failures.
Normative analysis	By leveraging optimization techniques and machine learning, help determine the best options to achieve the expected results. What should we do?	It is used to determine the action plan to shorten the accounts receivable collection cycle and optimize the use of accounts payable discounts.

DOI: 10.54254/2754-1169/2025.BL29249

Data analysis is a process that studies structured and unstructured data and generates useful analyses through statistical and quantitative analysis methods, with the aid of inferential models [8]. Through data analysis, financial institutions can assess the repayment risks of borrowers in advance (predictive analysis), infer possible causes (diagnostic analysis), provide risk hedging strategies (normative analysis), generate relevant reports, etc. (descriptive analysis).

Data analysis can monitor in real time and integrate multi-dimensional data to detect financial fraud (such as identity theft, account takeover, etc.) [9]. It can not only help financial institutions identify the risks of borrowers maliciously defaulting on debts and failing to repay on time. For borrowers, it can also help them protect their accounts, financial and identity information from being misused, effectively safeguarding their privacy [10].

3. Factors of loan risk

Credit risk is classified into the following types: default risk (borrowers fail to repay on time), concentration risk (excessive exposure to a single entity or borrower), systemic risk (chain reactions triggered by institutional turmoil or default by financial institutions), and other types of risks. This study mainly focuses on default risk, takes concentration risk into account, and does not involve systemic risk or other types of risk, as can be seen in table 2. After research, the influencing factors affecting borrowers' repayment are as follows [11, 12] .

Table 2. Influencing factors affecting borrower repayment

Influencing factors	Explanation of the cause
Historical model factors (credit scores, income levels and employment history, etc.)	Static, based on past ratings. Unable to cope with the dynamic changes in the financial market and the mutual influence among different factors
Law and Infrastructure	The regulations on the use of social media data are still unclear (for instance, there are no clear standards for including customers' Telegram or Instagram account data in account ratings). Most banks still use traditional servers and have not adopted cloud technology
Social acceptance	Rural residents and elderly residents have a relatively low acceptance of digital scoring
The borrower's own personality	The optimism of some borrowers far exceeds that of others (manifested as overestimating future income, underestimating the completion time of daily tasks, overestimating the probability of survival, etc.)

Summarizing the factors of debt risk, this article conducts a more detailed exploration of social acceptance. The following table 3 shows the acceptance of big data credit investigation by different regions and age groups.

Table 3. The proportion of trust in big data scores among different groups

Category	Group	The proportion of trust in big data technology
region	Urban residents	65%
	Rural residents	35%
age	Young people aged 18 to 20	80%
	People over 50 years old	40%

Past studies have shown that big data can reduce credit analysis errors by 40%. However, the proportion in Uzbekistan is still around 15% to 20%. For developed countries, the accuracy rate of

DOI: 10.54254/2754-1169/2025.BL29249

credit analysis will also lead to an increase in the proportion of trust in numerical scores [8].

4. Efficiency of loan risk assessment

A series of past studies have analyzed the project materials and customer surveys of the Central Bank of Uzbekistan. By using credit risk ratios, artificial intelligence models and SWOT analysis methods, etc., the efficiency (including time, cost, error rate, etc.) of traditional assessment and digital assessment is compared.

Table 4. Comparison of efficiency and customer satisfaction between traditional methods and digital window

Parameter	Traditional method	Digital Window
Credit period restriction	5 to 10 days	1 to 3 days
Error rate	12%	4.1%
Cost (1 loan)	150,000 summs	50,000 soms
Customer satisfaction	62%	89%

(Data sources: Report of the Central Bank of Uzbekistan (2024), data from the Industrial Construction Bank of Uzbekistan)

According to the table 4, it can be seen that the efficiency, fault tolerance rate and cost of digital loans are all significantly lower than those of traditional methods. Moreover, customers are more satisfied with digital loans and call for further development of digital loans. It can be seen that relying on big data technology, the loan services of financial institutions have been greatly improved, which is undoubtedly a positive factor for identifying and further reducing risks.

Research has found that if a central data platform is established, the accuracy rate of scoring will increase by 30% to 35%. Legislative reform can increase customer trust by 20% to 25%. This result strongly proves that a central data platform should be established, with classified and organized data, and relevant data should be connected to obtain more convincing credit scores. At the same time, legislation should be enacted to clarify how social media data is used and to what extent. While fully protecting the borrower's privacy, a more personalized borrowing plan is also specified [8].

Uzbekistan is a less developed country. The overall financial literacy of borrowers is relatively low, and their debt-paying ability is also low. Considering the limitations of less developed countries like Uzbekistan and rural areas, the principle of inclusive finance should also be kept in mind. The dissemination of digital inclusive finance. By investigating the production investment, household income and the strength of social relations (quantifying the number of gifts given in the past) of tens of thousands of rural Chinese families, farmers with strong social relations were identified. Through mutual information exchange, they learned about the benefits of digital finance and enhanced its influence. With the increase in the time spent on mobile financial services, farmers' financial knowledge has improved, and the utilization of digital finance has increased productive investment, leading to an increase in the number of loans [13]. Therefore, through digital inclusive finance, borrowers in rural areas can enhance their financial literacy and contribute to financial development [14].

5. Conclusion

Big data can be classified into four types: descriptive analysis, diagnostic analysis, predictive analysis, and normative analysis. In the field of financial credit information, descriptive analysis is

DOI: 10.54254/2754-1169/2025.BL29249

often used to generate credit reports for customers, diagnostic analysis is employed to explore the reasons why borrowers failed to repay on time in the past, and predictive analysis is utilized to predict the financial situation of each customer and pre-arrange the possibility of repayment, and to formulate hedging strategies. Determine the action plan for communicating with the borrower and whether to grant a loan through normative analysis.

To ensure that big data technology precisely empowers the credit reporting field and effectively protects the legitimate rights and interests of financial institutions, we need to pay attention to the following factors. The factors include historical model factors (such as credit scores, income levels, and employment history, etc.), legal and infrastructure issues, social acceptance, and the borrower's own personality. In many of these developing countries, the laws regarding the use of social media data are still unclear, and the infrastructure is still not well-developed. There is a significant difference in social acceptance among people of different ages and income levels (only data from Uzbekistan is provided, with a comparison of the world average data). Whether a borrower's own personality is overly optimistic or not will affect their judgment of the possibility of timely repayment, and thus influence their credit score.

In regions with different levels of development, there are differences in the efficiency and satisfaction of big data Windows. In developing countries represented by Uzbekistan, the credit period limit, error rate, cost (per loan), and customer satisfaction of the digital window are all far superior to those of traditional methods. There is an urgent call for the promotion of digital financial credit reporting. And considering that the accuracy rate of big data analysis in Uzbekistan is lower than the world average. In developed countries, as the accuracy rate improves, the efficiency and satisfaction of digital Windows will only increase.

To better enable big data technology to serve the credit investigation and financial fields, the following suggestions are hereby put forward. Necessary financial knowledge should be taught to borrowers to ensure they have a correct understanding of sudden income fluctuations and the probability of bankruptcy and report them in a timely manner. Establish a central data platform, carry out legislative reforms on social media, clearly define the boundaries for financial institutions to use borrowers' social media data, and provide personalized borrowing plans while protecting borrowers' privacy and security. Enhance the accuracy and efficiency of big data analysis technology, teach financial knowledge through digital loans, and increase trust in digital finance and credit scoring. Guide small loans, taking into account the characteristics of unstable income and lack of financial knowledge, and flexibly adjust the repayment period for rural areas.

This study only collected partial data and did not conduct a unified statistics on the trust level of people around the world in digital finance. There is a risk of overgeneralization. Therefore, the world average level was compared to enhance credibility and persuasiveness. At the same time, it mainly focuses on the perspective of financial institutions and does not compare the differences among various digital technologies, resulting in omissions of technical issues. In the future, we can focus on the advantages and disadvantages of different digital technologies, and combine the exploration of systemic risks and other risks to provide a new paradigm for breakthroughs in the digital lending field.

References

[1] Fu, Y., & Zhou, X. (2024). Does the big data credit platform reduce corporate credit resource mismatch: Evidence from China. Finance Research Letters, 69, 106133.

[2] Liu, Z., & Wang, L. (2023). Analysis of credit risk control for small and micro enterprises in Chinese commercial banks: From the perspective of big data credit reporting. Financial Engineering and Risk Management, 6(9), 1-15.

DOI: 10.54254/2754-1169/2025.BL29249

[3] VenkateswaraRao, M., Vellela, S., Reddy, V., Vullam, N., & Sk, K. B. (2023, March). Credit investigation and comprehensive risk management system based big data analytics in commercial banking. In 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 2387-2391). IEEE.
[4] Nahar, J., Rahaman, M. A., Alauddin, M., & Rozony, F. Z. (2024). Big data in credit risk management: a systematic review of transformative practices and future directions. International Journal of Management Information Systems and Data Science, 1(04), 68-79.
[5] Zou, J., Yao, L., Wang, B., Zhang, Y., & Deng, X. (2024). How does digital inclusive finance promote the journey of common prosperity in China?. Cities, 150, 105083.
[6] Bose, S., Dey, S. K., & Bhattacharjee, S. (2023). Big data, data analytics and artificial intelligence in accounting: An overview. Handbook of big data research methods, 32-51.
[7] Rane, N. (2023). Enhancing customer loyalty through Artificial Intelligence (AI), Internet of Things (IoT), and Big Data technologies: improving customer satisfaction, engagement, relationship, and experience. Internet of Things (IoT), and Big Data Technologies: Improving Customer Satisfaction, Engagement, Relationship, and Experience (October 13, 2023).
[8] Jumayev, B. (2025). Big data: customer credit analysis using digital banking database. International Journal of Artificial Intelligence, 1(2), 1056-1059.
[9] Udeh, E. O., Amajuoyi, P., Adeusi, K. B., & Scott, A. O. (2024). The role of big data in detecting and preventing financial fraud in digital transactions. World Journal of Advanced Research and Reviews, 22(2), 1746-1760.
[10] Addy, W. A., Ugochukwu, C. E., Oyewole, A. T., Ofodile, O. C., Adeoye, O. B., & Okoye, C. C. (2024). Predictive analytics in credit risk management for banks: A comprehensive review. GSC Advanced Research and Reviews, 18(2), 434-449.
[11] Scott, A. O., Amajuoyi, P., & Adeusi, K. B. (2024). Effective credit risk mitigation strategies: Solutions for reducing exposure in financial institutions. Magna Scientia Advanced Research and Reviews, 11(1), 198-211.
[12] Exler, F., Livshits, I., MacGee, J., & Tertilt, M. (2024). Consumer credit with over-optimistic borrowers. Journal of the European Economic Association, jvae057.
[13] Fu, C., Sun, X., Guo, M., & Yu, C. (2024). Can digital inclusive finance facilitate productive investment in rural households?â An empirical study based on the China Household Finance Survey. Finance Research Letters, 61, 105034.
[14] Maganioti, A.E., Chrissanthi, H.D., Charalabos, P.C., Andreas, R.D., George, P.N. and Christos, C.N. (2010) Cointegration of Event-Related Potential (ERP) Signals in Experiments with Different Electromagnetic Field (EMF) Conditions. Health, 2, 400-406.