

Research on Credit Risk Models of Commercial Banks in the Context of Digital Finance

Jiaqi Ma

*School of International Studies, University of International Business and Economics, Beijing, China
2085170496@qq.com*

Abstract: Digital finance is a crucial part of the “Five Major Articles” in the digital finance era, representing not merely a technology-driven innovative endeavor but also serves as a cornerstone for advancing high-quality economic development and constructing a modern financial architecture. Its evolution integrates cutting-edge technological applications with systemic financial reforms, thereby fostering inclusive growth, optimizing resource allocation efficiency. Credit risk management models for commercial banks are a common focus nowadays for both the academic and practical fields. This paper categorizes the fundamental risk control models and analyzes real-world cases. The findings indicate that although commercial banks actively adopt new models and have proposed new evaluation criteria for the technology sector, these models still face issues such as high training costs, poor interpretability, and inadequate early warning capabilities. Simultaneously, data quality also constitutes a persistent concern. In conclusion, credit risk models still have significant room for optimization in the context of the digital finance era.

Keywords: Commercial Banks, Digital Finance, Credit Risk Models, Fintech

1. Introduction

The State Council of China has issued guiding opinions on improving the “Five Major Articles” in the financial sector, clearly emphasizing the promotion of technological finance, green finance, inclusive finance, pension finance, and digital finance. Among them, digital finance is the deep integration of the internet and the traditional financial industry, driving the digital transformation of traditional finance. Through a series of technological innovations, model innovations, and business innovations, digital finance has profoundly changed the operation mode and service methods of the financial industry. Digital finance offers significant advantages in many aspects. First, in asset supervision, traditional regulatory methods can no longer cope with the rapid innovation in the financial industry. As a result, tools like regulatory technology (RegTech) have emerged, improving compliance efficiency and enabling real-time monitoring and reporting. For instance, physical assets, such as buildings, can be better tracked and controlled, thus reducing credit risk for banks [1]. Second, digitalization has promoted the widespread accessibility of financial services. Mobile payments have enabled more people to enjoy convenient payment services, and online lending platforms (such as Lending Club) have provided financing channels for multiple parties. Third, digital finance has driven data-driven financial innovations. By analyzing big data and applying behavioral finance models, investors’ behavioral data can be analyzed to enhance risk control and optimize investment strategies. Furthermore, big data can be used to analyze customer needs and offer personalized services. Given

the current complex and volatile economic environment, the importance of financial risk assessment has become increasingly prominent. This paper focuses on the different credit risk models adopted by commercial banks in the digital finance era.

The paper is divided into four sections: the first section provides a review of the definitions and characteristics of digital finance and credit risk models; the second section discusses the current application of models by financial institutions; the third section examines the challenges or shortcomings that new models may bring; and the final section offers conclusions and summaries. In conclusion, this paper aims to understand how financial institutions use different models and avoid various risks through research on the digital finance industry, and to analyze the challenges faced by different models in order to offer insights for the adoption of models by financial institutions.

2. Literature review

2.1. Literature review on digital finance

As an important component of financial technology, digital finance leverages the deep integration of digital technologies such as the internet, artificial intelligence, and blockchain to innovate and optimize traditional financial services, offering more efficient, convenient, and inclusive financial products and services. Firstly, digital technology is the core driver of digital finance [2]. Digitalization enables financial services to be delivered through digital platforms like the internet and mobile devices, allowing users to complete transactions such as payments, transfers, and wealth management without relying on physical outlets. For instance, the widespread use of Alipay and WeChat Pay has greatly enhanced the convenience of payments. Secondly, digital finance is driven by big data [3]. By analyzing user behaviors, credit records, and other data, digital finance can provide personalized and precise services. Moreover, the application of digital finance can enhance the financial transparency of enterprises and improve the accuracy of their future earnings forecasts [4]. Thirdly, digital finance is inclusive, reducing service barriers and reaching populations that traditional finance could not, such as those in rural areas and small and micro enterprises. At the same time, the use of the internet has brought financial services to the general public. Research has shown that from 2012 to 2021, citizens' understanding of basic financial knowledge and financial products steadily grew [5]. The fourth point is that digital finance has a decentralized feature. Through digital platforms, networks, and the application of blockchain technology, financial services no longer rely on centralized institutions. This shift has stimulated the growth of the financing market and attracted substantial foreign investment [6]. Additionally, blockchain technology ensures the openness and traceability of transaction data, enhancing users' trust in financial services. Digital finance is highly intelligent, with the application of artificial intelligence and machine learning technologies making financial services more efficient and precise. Intelligent advisory platforms, such as Betterment, provide investment advice through algorithms, while AI-powered customer service offers 24/7 support via chatbots.

2.2. Literature review on credit risk models

Credit risk models are statistical or mathematical models used to assess the likelihood of a borrower or counterparty defaulting. Their primary goal is to quantify credit risk, helping financial institutions make decisions on loans, investments, and risk management. Credit risk models are essential tools for banks to evaluate the likelihood of borrower default and play a crucial role in risk management, especially in the context of internet finance. Different types of credit risk models have their respective advantages and challenges, and financial institutions select appropriate models based on their business characteristics and data conditions for risk management.

According to existing literature, credit risk models can be classified into four categories based on their technical architecture. The first category consists of linear models based on statistics, such as

scoring models, primarily used in heavily regulated scenarios like credit approvals. These models use credit scores or risk levels as the final basis for decision-making. Common examples include the FICO credit score model based on logistic regression and the Z-score model used for corporate credit scoring. The second category includes non-linear models, with typical examples being various neural network models, such as deep neural networks (DNN) and graph neural networks (GNN). These models can automatically extract higher-order features (such as temporal patterns in transaction sequences) and are used for internet fraud detection. The third category is hybrid models, which combine multiple models and weight their predictive results. An example is the KMV model, which combines traditional statistics with financial engineering methods. This model can use market data (such as stock prices) to reflect a company's credit risk in real-time, offering more sensitivity than traditional financial ratio analysis [7]. The fourth category includes models designed to handle unstructured data, such as natural language processing models for text, primarily used for verifying the authenticity of loan materials or monitoring public opinion risks.

3. Current application of models in financial institutions

3.1. Application of core risk control models

In practice, commercial banks do not solely rely on a single model for risk evaluation; instead, they depend on the combined results of multiple models. The following section provides a brief overview of some of the widely used and core risk control models, classified according to the standards mentioned in the previous section.

The scoring card model is a widely used tool for credit risk assessment and decision-making management. In the era of digital finance, compared to traditional scoring cards, modern scoring cards have been extensively improved and optimized by integrating contemporary data science techniques and changes in business requirements. Modern scoring cards incorporate more external and non-traditional data, such as social media data, e-commerce transaction data, and mobile phone usage behavior data. These data sources are integrated using big data technology to enhance the model's predictive power. At the same time, financial institutions primarily use statistical and econometric methods, while also incorporating machine learning models such as random forests, along with artificial intelligence and ensemble learning methods (e.g., machine learning algorithms) to improve prediction accuracy and stability [8]. Furthermore, more advanced binning techniques are applied, such as Monotonic Binning, which better captures the relationships between variables and targets.

Deep learning models also play a significant role in credit risk assessment. These models, based on artificial neural networks, learn complex patterns and features from data through multiple layers of non-linear transformations. In credit risk assessment, deep learning models can handle large-scale, high-dimensional data and capture non-linear relationships within the data. Various neural network models are employed to achieve high-risk prediction accuracy when dealing with large-scale data. For example, Graph Neural Networks (GNN) have widespread applications in the financial industry, particularly in handling complex financial network data, such as transaction networks and corporate relationship networks. GNNs have demonstrated exceptional performance in supply chain finance models, where the relationships between companies in the supply chain are complex and traditional risk assessment methods struggle to capture the transmission of risk. GNNs can build relationship networks between enterprises or individuals, allowing for a more comprehensive analysis and risk assessment. For example, Ant Financial uses GNNs to assess the credit risk of small and micro enterprises. Research has shown that under the supply chain finance model, the accuracy of predicting enterprise default probabilities can be improved, further enhancing China's credit evaluation system and database, thus diversifying the financing channels for small and medium-sized enterprises [9]. By constructing corresponding neural network models, financial institutions can not only help control

bad debt rates on online lending platforms and reduce operational costs, but also enhance investor confidence and improve the platform's public image [10].

In the application of a credit risk model, there are often certain deficiencies. In such cases, most financial institutions adopt hybrid models, which combine multiple models or techniques to leverage their respective advantages and compensate for the shortcomings of a single model. For example, one study proposed an e-commerce credit risk assessment model based on the RB-XGBoost algorithm [11]. The model first uses an adaptive random balancing method to improve data balance. After processing the data, the XGBoost algorithm is applied to construct the model. This hybrid model addresses issues within each individual method, ensuring data balance and improving the accuracy of risk assessments. The advantages of hybrid models include not only improved prediction accuracy and adaptability to complex data, but also enhanced model interpretability and reduced risk of overfitting.

When handling unstructured data, natural language processing (NLP) models provide powerful decision support tools for the financial industry by analyzing text data. The core task of NLP models is to extract useful information from text and convert it into structured data or directly use it for decision-making. In the financial sector, NLP models are widely used in credit risk assessments, market sentiment analysis, and customer relationship management. For example, the BERT model can be used to analyze loan application texts to evaluate borrower credit risk. Sentiment analysis models can predict the impact of news or social media on market sentiment, while text classification models can optimize customer service by identifying frequently asked questions. Bloomberg, for instance, has used NLP techniques to analyze the impact of news on corporate credit.

3.2. Case analysis

In recent years, with the rapid development of technology, a large number of tech-driven innovative enterprises have emerged. In response, the government introduced the "Innovation Credit System" to evaluate tech-based enterprises. This model transforms the innovation data of companies into "financial data" that financial institutions can understand and apply, thereby providing the basis for risk assessment and credit decision-making. The model is essentially a data-driven, quantitative evaluation hybrid model that quantifies a company's innovation capability through a rule engine and statistical or machine learning methods. A series of indicators can help better understand the innovation credit system, as shown in the table below.

Table 1: Innovation credit model

Primary Indicator	Secondary Indicator	Indicator Explanation	Data Period	Startup Period Weight	Growth Period Weight	Mature Period Weight
Technology Innovation Indicators	R&D Expenses (in ten thousand yuan)	The labor and direct input costs of R&D activities	Last two years	0.08	0.08	0.08
	R&D Expense Growth Rate (%)	The proportion of the increase in R&D investment during the reporting period relative to the total R&D investment of the previous year	Current year	0.08	0.08	0.08

Table 1: (continued)

	Enterprise Technology Contract Transaction Volume (in ten thousand yuan)	The transaction amount of technology contracts signed by the enterprise	Last two years	0.08	0.08	0.08
	Revenue from High-tech Products (in ten thousand yuan)	Revenue from the sale of high-tech products by the enterprise	Current year	0.08	0.08	0.08
	Proportion of R&D Personnel (%)	The proportion of R&D personnel to the total number of employees at the end of the period	Current year	0.08	0.08	0.08
	Proportion of Personnel with Graduate Degrees or Above (%)	The proportion of employees with graduate degrees or higher to the total number of employees at the end of the period	Current year	0.05	0.04	0.03
	R&D Tax Deduction and Exemption (in ten thousand yuan)	The tax reduction for research and development expenses that are subject to pre-tax additional deductions as stipulated by policies and tax laws	In those years	0.06	0.06	0.06
Growth and Operational Indicators	Return on Net Assets (%)	$\text{Net profit} \div (\text{Beginning owners' equity} + \text{Ending owners' equity}) \div 2 \times 100\%$	Current year	0.05	0.06	0.07
	Value-added Tax (in ten thousand yuan)	The amount of value-added tax paid by the enterprise	Current year	0.05	0.05	0.05
	Value-added Tax Growth Rate (%)	The proportion of the increase in value-added tax to the total value-added tax of the previous year	Current year	0.05	0.05	0.05
	Corporate Income Tax (in ten thousand yuan)	The amount of corporate income tax paid by the enterprise	Current year	0.05	0.05	0.05
	Number of New College Graduates Hired (persons)	The number of new graduates hired by the enterprise during the year	Current year	0.05	0.05	0.05

Table 1: (continued)

Auxiliary Indicators	Number of New Graduates Hired (persons)	The number of new graduates recruited by the enterprise from various domestic universities during the reporting period	Current year	0.03	0.04	0.05
	Number of Provincial or Above R&D or Innovation Platforms Built (items)	The number of key laboratories, engineering centers, etc., approved for the enterprise in the last two years	Last two years	0.05	0.06	0.07
	Number of Provincial or Above Science and Technology Awards Received (items)	The number of provincial or higher-level science and technology awards received by the enterprise as the lead unit in the last two years	Last two years	0.05	0.06	0.07
	Amount of Venture Capital Received (in ten thousand yuan)	The amount of venture capital received by the enterprise	Current year	0.05	0.05	0.05
	Enterprise Credit Rating	The credit rating of the enterprise	Current year	0.05	0.05	0.05

In addition, small and micro enterprises also play an important role in the economic and social landscape. Credit assessment models have been continuously optimized in the context of the digital era. In order to accurately assess the risks of small and micro enterprises, the application of machine learning technology has become increasingly widespread. One study improved the accuracy of risk assessment by optimizing the scorecard model [12]. This type of model prioritizes the use of the SMOTE algorithm to handle imbalanced data, ensuring a balanced representation of minority class samples. Furthermore, machine learning algorithms such as XGBoost are employed to identify key factors that influence credit. Finally, an XGBoost scorecard model is developed based on these key factors. This model not only leverages the performance advantages of the SMOTE algorithm and the XGBoost model in handling imbalanced data, but also incorporates non-financial information, such as business behavior, as supplementary dimensions. The following figure illustrates the flowchart for credit risk assessment of small and micro enterprises.

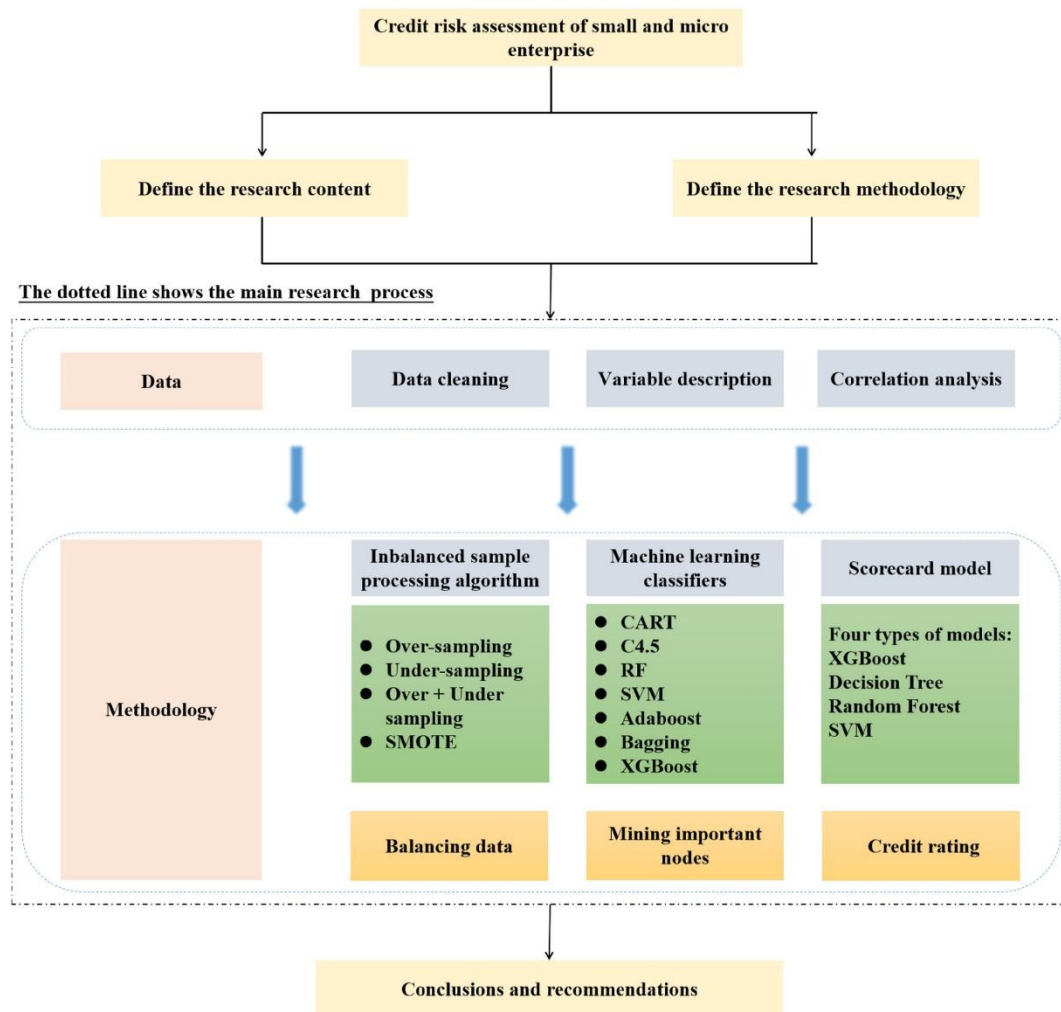


Figure 1: Credit risk assessment flowchart for small and micro enterprises

Overall, risk control models have become a core tool in the risk management of financial institutions. Through technological advancements, major financial institutions have further improved the application scenarios and precision of these models, enabling effective control of credit risk and providing stable support for industry development.

4. Potential risks and challenges in the operation of new models

4.1. High model training costs

Hybrid models typically combine multiple technologies, which brings numerous benefits but also makes the model structure more complex. This complexity means that model development and debugging become more challenging, and maintenance costs rise, particularly in scenarios where the model requires frequent updates. Additionally, hybrid models often require a large amount of high-quality data for training. In practice, hybrid models depend on vast data support [13]. Moreover, the data may come from multiple sources, requiring complex preprocessing and integration. Financial institutions need to ensure data quality and security, optimize computational resources, reduce deployment costs, and strengthen model maintenance and technical updates to ensure the long-term effectiveness of the model.

4.2. Poor model interpretability

Deep learning models are often viewed as “black boxes” because their decision-making processes are difficult to explain. In other words, while neural network models can adjust according to samples to improve predictive accuracy, it is impossible to derive the contribution of input variables from the weights in the network [14]. Models that are difficult to interpret may not meet compliance requirements, leading to legal risks. Deep learning models also typically handle large volumes of sensitive data, which means financial institutions must have extremely high guarantees regarding data privacy and security.

4.3. Data quality issues

The scorecard is one of the most commonly used tools for credit risk assessment in the financial industry. However, the performance of this model is highly dependent on the quality and availability of data. Missing values, noise, or bias in the data can all affect the model’s accuracy. The scorecard model also requires variable selection and binning, which can introduce subjective bias or improper binning strategies, potentially leading to information loss or degraded model performance. Furthermore, scorecard models struggle with processing unstructured data, such as text and images. In complex scenarios, the scorecard model may need to be integrated with other models.

4.4. Insufficient early warning capabilities

In dynamic risk monitoring, credit risk control models have limitations in their dynamic monitoring capabilities, making it difficult to capture changes in risk in real-time. For example, while the market environment and business conditions of enterprises are constantly changing, the speed and adaptability of model updates are insufficient, preventing timely early warnings. Additionally, the predictive capabilities of the models are limited. Existing models in predicting default risks typically rely on historical data and known risk characteristics for assessment, making it hard to capture sudden risks. For example, some enterprises may face sudden market changes or policy adjustments during their operations, which existing models are unlikely to predict in advance.

4.5. Insufficient risk mitigation capabilities

Currently, there is insufficient risk assessment for collateral. Collateral is an important tool for mitigating credit risk, and its value assessment and risk identification capabilities are crucial. However, many financial institutions rely on external agencies for collateral assessment, lacking internal professional capabilities, which leads to inaccurate collateral valuations and an inability to effectively assess risk mitigation potential. The diversity of risk mitigation tools is also low. Existing risk mitigation tools mainly focus on traditional methods such as mortgages and guarantees, lacking diversified risk mitigation strategies. For instance, the application of modern risk mitigation tools like financial derivatives is relatively rare, making it difficult to meet the complex and changing market demands.

5. Conclusion and outlook

In the context of digital finance, the emergence of various types of enterprises has brought new challenges to commercial banks in risk control. Credit risk control has always been an essential component of bank management. This paper first classifies the current models and briefly reviews the core models within each category. It then studies innovative scoring models and the machine learning-based XGBoost scorecard model that have recently emerged. The study found that, although these models have clear advantages in handling new indicators and complex data, they still face

challenges such as high training costs, poor interpretability, insufficient early warning capabilities, inadequate risk mitigation abilities, and data quality issues. The study's results have the limitation of a small sample size, and future research could expand the sample size. Overall, the findings provide valuable insights for the future optimization and improvement of credit risk control models.

References

- [1] Wang, P., & Wang, H. (2025). Can digital financial development improve the accuracy of corporate earnings forecasts? *Finance Research Letters*, 75, 106833.
- [2] Shen, W. (2022). Blockchain financial regulation in the digital economy era: Current status, risks, and responses. *People's Forum: Academic Frontier*, (18), 52-69. <https://doi.org/10.16619/j.cnki.rmltxsqy.2022.18.006>
- [3] Peng, J., Zhu, Y., & Chen, T. (2024). Digital finance, green innovation, and the transformation and upgrading of the manufacturing industry. *Contemporary Economic Science*, 1-18.
- [4] Jiang, M., Zhou, W., & Zong, J. (2025). The role of digital finance in FDI inflow: Facilitator or inhibitor? *Review of World Economics* (prepublish), 1-26.
- [5] Bland, E., Changchit, C., Changchit, C., Cutshall, R., & Pham, L. (2024). Investigating the components of perceived risk factors affecting mobile payment adoption. *Journal of Risk and Financial Management*, 17(6), 216.
- [6] Ratna, S., Saide, S., Mesha Putri, A., Soleha, A., & Andini, P. R. (2024). Financial systems transformation in the digital age: A systematic review and future research directions. *Procedia Computer Science*, 234, 1538-1545.
- [7] Ling, J., & Liu, Y. (2013). Empirical analysis of the credit risk of Chinese commercial banks based on the KMV model: A case study of 10 listed commercial banks. *Journal of South China Normal University (Social Science Edition)*, (05), 142-148 + 209.
- [8] Yan, L., Ma, Y., & Ming, Y. (2024). Research on the construction of personal credit scorecard model based on logistic regression and random forest fusion model. *Scientific Journal of Intelligent Systems Research*, 6(7), 34-39.
- [9] Ma, X., Yuan, Y., & Hu, J. (2023). Research on the enterprise credit risk rating system in supply chain finance based on Jingdong. *Academic Journal of Business & Management*, 5(19).
- [10] Ma, Z., Hou, W., & Zhang, D. (2021). A credit risk assessment model of borrowers in P2P lending based on BP neural network. *PloS One*, 16(8), e0255216.
- [11] Yang, W., & Gao, L. (2021). A study on RB-XGBoost algorithm-based e-commerce credit risk assessment model. *Journal of Sensors*.
- [12] Gu, Z., Lv, J., Wu, B., Hu, Z., & Yu, X. (2024). Credit risk assessment of small and micro enterprises based on machine learning. *Heliyon*, 10(5), e27096.
- [13] Shen, Q., & Zhang, L. (2020). A new method for bank credit risk identification: SVM-KNN combined model. *Financial Supervision Research*, (07), 23-37. <https://doi.org/10.13490/j.cnki.frr.2020.07.002>
- [14] Liu, X., & Wang, W. (2015). A multi-model comparative study of credit risk identification in Chinese commercial banks. *Economic Geography*, 32(06), 132-137. <https://doi.org/10.15931/j.cnki.1006-1096.2015.06.023>