Hindawi
Mobile Information Systems
Volume 2021, Article ID 8189255, 10 pages
https://doi.org/10.1155/2021/8189255

Research Article

Credit Risk Management of Consumer Finance Based on Big Data

Huibo Wang

School of Economics and Management, Beihang University, Beijing 100191, China

Correspondence should be addressed to Huibo Wang; whb1225@buaa.edu.cn

Received 8 June 2021; Revised 26 June 2021; Accepted 13 July 2021; Published 23 July 2021

Academic Editor: Sang-Bing Tsai

In recent years, China's consumer finance has developed rapidly, but the foundation is unstable, and the industry has serious problems of violent competition, excessive credit, and fraud. Therefore, we should attach great importance to the healthy development of consumer finance, especially the management of its credit risk. The application of big data credit investigation can provide early warning of potential risks and prevent the risk of excessive credit investigation. This paper starts with the definition of basic core concepts, such as traditional credit investigation, big data credit investigation, and consumer finance, analyzes the performance and causes of consumer finance credit risk, and combs in detail the relevant theories of the application of big data credit investigation in consumer finance credit risk management. The application of big data credit investigation has optimized the risk management process of consumer financial institutions, deepened the concept of Internet consumer finance, improved the risk management system, created a diversified credit information system, and strengthened the innovation of Internet consumer finance products and services. For example, credit scores provide the most intuitive quantification of consumer credit risk. For consumers with different levels of credit scores, different credit approval processes can be matched. For customers with high scores, the work process can be simplified without affecting the work results. It can reduce the workload of employees by 20% and increase the accuracy of customer credit risk prediction by 16%.

1. Introduction

In an environment where the country has released policy dividends and increased marketization, the consumer finance industry has achieved rapid development. In addition, judging from the rapid development of financial technology technologies such as the Internet, big data, and machine algorithms, as well as the prevailing status of credit behavior, the future development trend of consumer finance must shift from offline to online, and the development of Internet consumer finance the potential is endless. Nowadays, in the context of relatively fierce competition, industry regulatory policies and regulations have emerged one after another in the consumer finance industry, with overcredit and multiple debt problems, and consumer finance credit risks are frequently occurring. From the perspective of the incompleteness of China's Internet credit system, it analyzes the problems of China's Internet consumer finance encouragement; that is, there are defects in online credit evaluation and obstacles to cross-platform cooperation mechanisms, and user privacy protection issues have not been resolved [1]. However, with the gradual improvement of China's credit reporting system, large-scale credit reporting agencies have effectively made up for the shortcomings of traditional credit reporting agencies and have been widely used in the consumer finance industry. The specific application of financial consumption credit risk management is a problem that needs to be solved urgently in the financial theory and practice circles. Therefore, it is necessary to strengthen the application of credit information big data in financial consumption and financial risk management. This article aims to study the specific application and problems of big data credit investigation data in financial consumer credit risk management, improve the application of big data credit investigation data in financial risk management of financial consumers in China, and put forward suggestions.

The widespread application of big data technology and the continuous innovation of the financial industry have made the growth of Internet finance irreversible. The new economic model has improved the efficiency of financial

services and has developed new payment and transaction methods. The traditional credit information system can no longer meet the developmental needs of the financial industry. Liu et al. adopted research methods, including literature research, case analysis, and comparative analysis. They analyzed the development process, current situation, and characteristics of China's personal credit reporting system, as well as the development process and current status of China's personal credit reporting system under Internet finance. They also pointed out the main problems faced by China's personal credit investigation system under Internet finance. However, in view of the specific problems faced by China's personal credit investigation system under Internet finance, they did not make specific suggestions for improvement of China's personal credit investigation [2]. Cullen et al. examine whether bank audit work represented by audit fees is related to asset securitization risk (ASR) and whether the incremental audit work attributed to ASR is related to audit quality. Their sample period included the global financial crisis (GFC) and the introduction of FAS no. 166 and FAS no. 167, which were intended to limit the accounting treatment of asset securitization as sales. Using Bank of America Holdings (BHC) data from 2003 to 2013, Cullen G found a significant positive correlation between ASR and the audit fees of Big N auditors, but not for non-Big N auditors. Audit fees before the global financial crisis are positively correlated with ASR and are even more important for BHCs that report losses. After the implementation of FAS no. 166 and no. 167, this positive correlation continued to exist, mainly driven by BHC's reported losses. Regarding the incremental audit work attributed to ASR, he found that, before the global financial crisis, the incremental audit work of Big N auditors reduced the possibility of subsequent restatements and limited the reporting of income securities. However, his research did not give a detailed explanation on the related categories of capitalization risks [3]. Mohabeer et al.'s study explored how big data analysis can optimize the use of public funds while ensuring that public organizations provide quality services to Mauritian citizens. A political, economic, sociotechnical (PEST) analysis has been performed to scan the environment to identify at least two major policies and initiatives corresponding to big data that will affect the economy of Mauritius in the next 10 years. Subsequently, the layered causal analysis (CLA) has been applied to these two signals, creating room for change in the creation of an alternative future. In fact, the survey results show that the implementation of the open data initiative and the Mauritius e-health project will definitely make a positive contribution to the Mauritius government. However, this study reveals through a matrix of possible futures that the Mauritian government should revise existing conventional laws through reforms and regulations to make full use of big data analysis applications. However, their research is too complicated for specific applications [4].

The innovation of this article is to analyze the risk control of Internet finance from the unique perspective of big data. Based on big data technology, this article collects, mines, analyzes, and extracts valuable data from various platforms and quantifies it. It relies on a unique model to quantify risks and improve the accuracy of risk identification and puts forward the AdaBoost algorithm to achieve overdue prediction. The conclusion method drawn is convincing. Then, through specific operation methods and models, methods and theories can be derived, which can provide better services for the risk control of Internet finance.

2. Implementation Method of Consumer Finance Credit Risk Management Based on Big Data Credit Investigation

2.1. Multipoint Network Organization Consumer Credit Risk Management Algorithm. For borrowers, the purpose is to choose a reputable financial institution to obtain a certain amount of loans at a reasonable interest rate to meet their consumer needs [5]. After this kind of borrower and lender has established a relationship network, the actual operation of the relationship financing is the product of the combination of this network organization and actual needs. At this time, the common goal of the entire network organization is to establish a relationship through much cooperation and obtain a stable and reliable way of borrowing, so as to realize its own profit or consumption needs [6]. Such a network organization actually broadens the width and duration of cooperation between borrowers and lenders and builds a single game on the basis of repeated games, which provides a basis for the incentive effect of reputation. As the mutual understanding between borrowers and lenders deepens and cooperation continues to deepen, network organization theory is used to prevent consumer credit risks and alleviate information asymmetry.

(1) Calculation model of two-node network organization:

This paper uses models and uses the idea of dynamic game to illustrate the role of using relational networks to alleviate the adverse selection phenomenon in consumer credit business. Suppose there are three types of consumer credit customers in the market, good customers (A), general customers (B), and poor customers (D). The difference is that good customers have a high probability of fulfilling their contracts. That is, $R_A > R_D$, respectively. Assuming that the performance probability of the device is R_A, R_B, and R_D, then $R_A > R_B > R_D$. Assuming that the information on the total proportion of various borrowers is public, then anyone can obtain information on the cost of funds. Then, as long as the interest rate satisfies the following formula, a positive return can be obtained:

$$(P_A - \overline{P})R_A - (1 - R_A)(1 + \overline{P}) > 0. \quad (1)$$

$$P_A > \frac{\overline{P} + 1}{R_A} - 1. \quad (2)$$

In the same way, the interest rates provided by the relationship banks to middle- and high-scoring borrowers in the county should meet the following:

$$P_B > \frac{\overline{P} + 1}{P_B} - 1. \tag{3}$$

$$P_D > \frac{\overline{P} + 1}{R_D} - 1. \tag{4}$$

However, for Fei Relations Bank, because it cannot obtain the risk level of each consumer, it can only set a same interest rate for all borrowers based on the proportion of various borrowers, and the interest rate meets the following:

$$\phi_1 \left[(P - \overline{P})R_A - (1 - R_A)(1 + \overline{P})\right] + \phi_2 \left[(P - \overline{P})R_B - (1 - R_B)(1 + \overline{P})\right] + (1 - \phi_1 - \phi_2)\left[(P - \overline{P})R_D - (1 - R_D)(1 + \overline{P})\right] > 0. \tag{5}$$

Simplified:

$$P > \frac{\overline{P} + 1}{\phi_1 R_A + \phi_2 R_B + (1 - \phi_1 - \phi_2)} - 1. \tag{6}$$

Formula (6) represents the proportion of low-risk and medium-risk. At this time, nonrelational banks will also expect that low-risk individuals will not choose to borrow money as long as they have a relational bank. In order to ensure that their income is nonnegative, they will adjust their borrowing interest rate accordingly to satisfy the inequality.

$$P' > \frac{\overline{P} + 1}{(\phi_2/(1 - \phi_1))R_B + ((1 - \phi_1 - \phi_2)/(1 - \phi_1))R_D} - 1. \tag{7}$$

Therefore, once nonrelationship banks make interest rate adjustments to ensure that their returns are nonnegative, they will lose customers with moderate risks. A rational borrower will expect that only high-risk consumers will have bandwidth, which will result in a loan-reluctant behavior.

$$P_0 \geq \frac{R_A(R_A - 1)(2\overline{P} + \lambda) + 2(1 + \overline{P}) + R_A(R_A - 3)}{R_A(R_A + 1)}. \tag{8}$$

Among them, $R_A(R_A - 1) < 0$; therefore, the greater the joint liability, the lower the interest rate required by the bank and the more secure the repayment.

2.2. Data Integration under the Big Data Platform.

The introduction of the experimental data and the classification prediction algorithm model that need to be integrated and processed for the research on the credit risk overdue prediction problem has been completed in the previous article [7, 8]. Credit risk overdue prediction is essentially a binary classification problem on unbalanced data sets [9]. In order to better solve the problem of risk overdue prediction, this chapter integrates the idea of cost-sensitive learning with good performance on unbalanced data sets into AdaBoost. Later, the AdaBoost algorithm was proposed. In order to verify the classification and prediction performance of the new algorithm AdaBoost, this paper first selects the public imbalanced data set to test the algorithm performance and then applies it to the credit risk overdue problem to be studied and analyzes the experimental results.

Among them, TP is the number of samples that are correctly classified, FN is the number of samples that are actually classified as a minority but not as the majority, and FP is the number of samples that are actually classified as a minority. The majority category was mistakenly classified as the minority category. TN represents the number of correctly classified samples that actually belong to the majority category.

According to the confusion matrix, the accuracy of classification can be obtained by the following formula:

$$\mathrm{ACC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{FN} + \mathrm{TN}}. \tag{9}$$

Accuracy emphasizes the accuracy of minority category recognition, compared with the percentage of accuracy on which the prediction result is based.

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}. \tag{10}$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}. \tag{11}$$

The recall rate is the percentage of actual correct predictions in all samples in a few categories starting from the actual data set. High-recall rate classifiers can better identify a small number of samples and rarely misclassify samples that actually belong to the minority category in the majority category. As with earthquake prediction, earthquakes need to be accurately predicted.

$$\mathrm{F1} = \frac{2 \times \mathrm{Recall} \times \mathrm{Precision}}{\mathrm{Recall} + \mathrm{Precision}}. \tag{12}$$

Recall rate and accuracy rate are metrics that render classification results from different angles. The accuracy rate pays more attention to whether the predicted minority category is really a minority category, while the recall rate pays more attention to whether all the minority categories are predicted in the dataset.

The formula for the true rate of AUC and ROC is as follows:

$$\mathrm{TFR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}. \tag{13}$$

The formula for the false positive rate is as follows:

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}. \tag{14}$$

(2) Determination of cost-sensitive factors:

Cost-sensitive learning transforms the problem of minimizing errors into the problem of minimizing the cost of misclassification. The determination of cost-sensitive factors is a crucial issue. For the purpose of minimizing risk (cost), the values of the four cost-sensitive factors are determined.

Given the sample point d, the risk cost of the minority class is predicted:

$$\arg\min_{i=0,1}\{S(T = i|d)\}. \tag{15}$$

That is, the class with the smallest substitution price is the prediction set of sample points.

3. Consumer Finance Credit Risk Management Experiment Based on Big Data Credit Investigation

3.1. Big Data Credit Investigation. Driven by IT technology and financial technology, big data technology has developed rapidly, and big data credit investigation has emerged from the trend, which can be widely used in various industries to play a huge market value. The characteristics of big data credit investigation include timeliness and comprehensiveness [10, 11]. It can achieve precise customer marketing and intelligent credit risk management. For individuals, credit evaluation also needs to consider network data that can represent the individual's daily behavior. In addition to referring to traditional credit investigation data, the basic big data credit investigation database focuses on collecting real-time updated data on the network platform, just as human behavior is always affected by microscopic factors such as environment and psychology. Big data credit investigation details [12, 13] are shown in Figure 1.

Big data provides information query services, including operator data, e-commerce data, student credit data, credit card data, social security data, and travel data and then conduct information verification services and identity transactions, verify academic qualifications, and conduct vehicle inquiries and criminal inquiries [14, 15]. Personal credit investigation services are based on personal credit

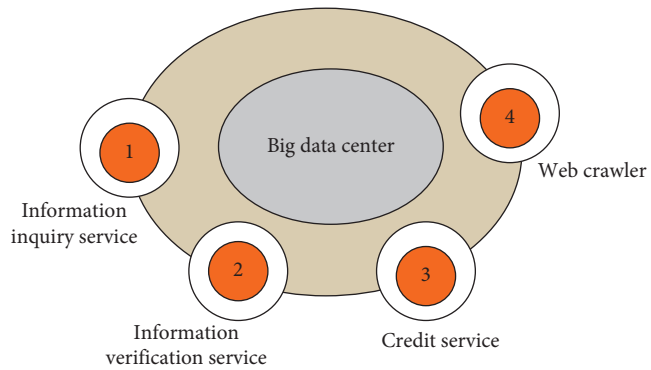

Figure 1: The overall structure of big data credit investigation.

reports, personal risk profiles, bank account verification, and personal analysis and monitoring. Then, the High Court Information Network and the Industrial and Commercial Information Network are used to crawl specific data [16–18].

The difference between big data credit investigation and traditional credit investigation is as follows. The traditional credit investigation is involved with the following: structured data such as credit data, mobile phone bills, and consumer bills, serving the credit market, targeting the assets and debts of large enterprises and personal credit subjects with relatively complete credit records, and circumstances and ability to repay [19, 20]. Big data credit investigation is involved with the following: unstructured data, such as network data focusing on serving inclusive finance, and targeting small and microenterprises and individuals without credit records and paying attention to some social behaviors of credit subjects, such as circle of friends, consumer preferences, and online search records and many more [21, 22].

The perfect integration of big data technology and economy and the transformation and development of the credit investigation industry, which is the foundation of the market economy, cannot be separated from the support of big data technology, opening up a new direction for the development of the market economy [23]. First of all, the prominent "big" characteristics of big data technology make up for the shortcomings of the incomplete database under the traditional credit investigation system, and it is more conducive to the more comprehensive "reflection" of the credit investigation to the information subject. Second, big data technology can conduct in-depth mining and precise analysis of basic information data. With the help of scientific algorithms and model output, the virtual portrait of the information subject can be depicted as truthfully as possible, and it can provide practical life scenarios and expand its application range [24, 25]. The characteristics of big data credit collection, which can collect, analyze, process, and update credit reports at any time and in real time, make it easier for information demanders to grasp the information of creditors in an all-round way. On the one hand, it can effectively make up for the shortcomings of the traditional credit reporting system itself; on the other hand, it can meet the complex and diversified credit reporting service needs of the credit market. At present, big data credit investigation

has been widely used in consumer finance industry, such as Zhima Credit, JD Finance, and other Internet consumer finances [5, 26].

3.2. Wealth Finance

3.2.1. The Concept of Consumer Finance. In China, consumer finance refers to consumer credit products or services provided by financial institutions. Consumer finance in the broadest sense includes mortgage loans, auto consumer loans, personal consumer goods loans, and general personal consumer loans [27]. In the narrow sense, consumer finance refers to consumer loans that do not include housing and auto credit. Internet consumer finance is the application of Internet technology in the business links of traditional consumer finance, such as customer screening, lending, and repayment, which greatly improves the efficiency of traditional consumer finance and makes up for the shortcomings.

3.2.2. Characteristics of Consumer Finance. First, inclusiveness credit, as an intangible capital, plays an increasingly important role in the economic operation. On the one hand, it reduces the cost of credit, and on the other hand, it also improves the efficiency of traditional credit. Credit demanders can obtain credit support in a timely and convenient manner with their own excellent credit capital, which makes up for the inadequacy of traditional credit institutions that cannot fully cover consumers and further accelerates the rapid development of China's inclusive finance. Second, the consumption scene is more diversified and refined. The refinement of consumption scenarios provides opportunities for personalized customization of consumer financial services. Expanding more and more indispensable consumption channels in life, such as education, medical treatment, and tourism, has become a means to increase customer stickiness in the consumer finance industry. On the one hand, consumption scenarios can meet consumer demand in batches. For any consumption scenario, as long as the single payment amount is high, there will be potential credit demand, which also increases the convenience of consumer marketing. On the other hand, the consumption scenario focuses on consumer characteristics, and under the same type of scenario setting, according to the consumption scenario, the characteristics of the customer types obtained in batches are the same, the consumption ability is similar, and the consumption needs are similar. Credit products with similar characteristics are designed and the scene in the vertical direction is refined, allowing to grasp this customer group firmly, gain more customers, and expand more scenarios that require credit services.

3.2.3. Consumer Financial Risk Management. The development of consumer finance business is more based on credit, especially online credit resources. Therefore, credit risk is the most important type of risk considered in the risk management of consumer finance business. Comparing Internet consumer finance and traditional consumer finance, it is pointed out that the advantages of the former are reflected in the ability to connect users' credit evaluation data and behavioral data with big data platforms, reduce the operating costs of consumer credit structures, and improve the efficiency of credit review and issuance. It can more accurately predict the user's loan demand and grasp the user's repayment ability. In terms of consumer finance, research believes that the development of consumer finance can stimulate economic growth, and the scale of industry development is growing rapidly, but there are problems such as insufficient risk prevention, unsound credit system development, and insecure user information. From the perspective of system construction, the countermeasures for the healthy development of consumer finance are proposed. The consumer finance industry with credit resources as the subject of transactions has developed so far. The business model is mainly "online + offline." While each link generates massive amounts of data, it also requires technology to interpret these data more accurately and comprehensively. Big data credit investigation relies on big data technology to collect, mine, and process massive amounts of data and finally generates credit products. Its essence is to solve the problem of information asymmetry and develop. There is a natural internal connection between big data credit investigation and consumer finance. How to accurately monitor customers' financial consumption behavior in real time is a problem that needs to be solved urgently in China credit risk management.

(1) Enrich the theory of consumer finance risk management and broaden the application areas of big data credit investigation. In the context of the Internet, consumer finance business models are mostly "online + offline," data source channels are complex, and risks are deepened. Traditional risk management methods can no longer meet the comprehensive, real-time, accurate, and efficient credit risk management needs of consumer finance. The big data credit investigation that has emerged to meet the needs of Internet finance is also constantly looking for a wider range of development areas. Consumer finance is a financial service based on credit that provides consumers with funds to meet their consumer needs and a good credit environment. It can accelerate its development, so big data credit investigation can expand its application in the consumer finance industry.

(2) Provide decision-making reference for the sustainable development of consumer finance. The development of consumer finance is to comply with the requirements of China's macroeconomic growth in recent years, fully implement the concept of inclusive finance, narrow the gap between the rich and the poor between individuals and regions, and increase the driving effect of consumption on the economy. It has huge development space and development significance. In recent years, China's consumer finance has developed rapidly, but the foundation is unstable. The problems of violent competition, excessive credit and fraud in the industry are serious. The

application of big data credit investigation can provide early warning of potential risks and prevent fraud and fraudulent use. Therefore, we should attach great importance to the healthy development of consumer finance, especially the management of its credit risk.

4. Consumer Finance Credit Risk Management Based on Big Data Credit Investigation

4.1. Consumer Credit Risk Management. Since the beginning of the construction of a bank, bank x, with the help of local market forces, has continued to expand in scale and developed by leaps and bounds. Under the influence of the booming development of the social economy and real estate industry, the scale of personal consumer credit has been qualitatively improved. The loan situation of bank x is shown in Table 1.

It can be seen from Table 1 that, during the period 2016–2019, the total loan amount of bank x has shown a continuous growth trend.

As shown in Table 2, it can be seen that the amount of nonperforming loans of a bank x has continued to decrease, but the speed is relatively slow, and the nonperforming loan ratio of personal loans has a significant gap from 2016 to 2019. The conclusion that can be drawn from this is that, between 2016 and 2019, the nonperforming loan ratio of bank x has shown a downward trend. The important factors are the continuous increase in the total amount of loans and the relatively weak personal loan risk management capabilities.

Figure 2 shows that personal housing credit accounts for an absolute proportion of personal consumption loans. Every year, general consumption and auto loans have increased year by year, and personal consumption credit has become the main force in the credit business.

4.1.1. Market Risk. Market risk refers to the losses caused to banks by market changes. It mainly refers to the risk of market fluctuations, that is, changes in market supply and demand caused by changes in national policies, development plans, and market demand. For bank xx, the market risk it faces is mainly interest rate risk. At present, the bank's loan-to-deposit maturity allocation ratio is not scientific. The deposit cycle is short and the loan cycle is long. There is a negative interest rate sensitivity gap. In the context of changes in interest rates, the bank needs to bear the interest rate risk caused by the decrease in interest income. In recent years, the coexistence of short-term deposits and long-term loans in banks xx is shown in Table 3:

According to Table 3, there is no obvious change in the proportion of time deposits and demand deposits. The amount of medium and long-term loans has shown an increasing trend. Personal housing loans and auto loans are an important part of medium and long-term loans. If interest rates fall, then bank interest rates will rise; if interest rates rise, bank interest rates will fall. The risks will increase correspondingly due to the problems of short-term deposits and long-term loans.

4.2. Current Situation and Problem Analysis of Consumer Financial Risk Management. This article analyzes a subset of the default debt forecast data. Each data is characterized by variables, such as age, income, debt ratio, and default, which are dependent variables by default. The independent variables include user attributes and historical credit records. Table 4 gives a detailed description of the variables.

Numerical variables can be calculated as average values. As shown in Table 4, the customer's age, income, and other basic information and the customer's credit record are used as variables.

Table 5 shows the classification of the credit rating system: excellent price range is 0.6–0.9; good price range 0.5–0.6; very poor price range <0.5. According to this standard, 19 points higher than 18 points are all excellent. However, 20 is good, and the information has been manually verified, which is the same as the result obtained.

4.3. AdaBoost Algorithm Based on Cost-Sensitive Improvement. Regarding the error backpropagation network, it is a typical multilevel unidirectional signal propagation network. The reflection propagation has certain repeatability; if the output of the required signal is not met, the cycle will continue. It is a neural network system that is widely used at this stage. Both the research on its structure and function and the research on its network principles have been quite mature. Compared with other networks of the same type, it has a particularly outstanding performance in nonlinear mapping capabilities.

For a BP network with a single transition layer, almost all functions that appear continuously in the closed interval can be derived. In other words, for a three-layer BP network, if there are only nodes on its transition layer, all the nonlinear derivation functions are simulated. Therefore, any related mapping from N dimension to M dimension can be successfully completed. The process is roughly as shown in Figure 3.

This article will use the test set to calculate the accuracy, recall, accuracy, and F1 measurement values of AdaBoost and other algorithms. In order to avoid the impact of randomness in the calculation process, this paper repeats 1000 simulations for each unbalanced dataset and takes the final average value of the performance evaluation index, as shown in Tables 6 and 7.

The results are shown in Tables 6 and 7 as the performance changes of AdaBoost. At the same time, the commonly used support vector machine and random forest classification methods are compared. Tested on three unbalanced different datasets, the reproducibility of the support vector, random forest, and original AdaBoost is not very good. As the degree of imbalance increases, the recall rate will gradually decrease, and the accuracy rate will also increase.

The debt default dataset has a high degree of imbalance. This chapter mainly focuses on the application analysis of the cost-sensitive AdaBoost algorithm previously proposed, from the classification results of the base classifier to the selection of cost-sensitive factors. Experiments show that the

Table 1: Bank x loan situation.

Years	Loan balance		Loan growth rate	
	Annual loan balance	Personal consumption loan balance	Total loan growth rate (%)	Personal consumption loan growth rate (%)
2016	2547.06	231.33	—	—
2017	2988.73	261.79	17.34	13.17
2018	3397.54	301.52	20.37	15.18
2019	3925.18	368.27	25.42	22.14

Table 2: Bank x nonperforming loan situation table.

Years	Bad debt		Nonperforming loan ratio	
	Total nonperforming loans	Personal consumption nonperforming loans	Total nonperforming loan ratio (%)	NPL ratio for personal consumption (%)
2016	438.15	6.16	17.20	2.66
2017	426.43	5.83	14.27	2.23
2018	389.64	5.35	10.83	1.77
2019	356.26	4.97	7.90	1.35

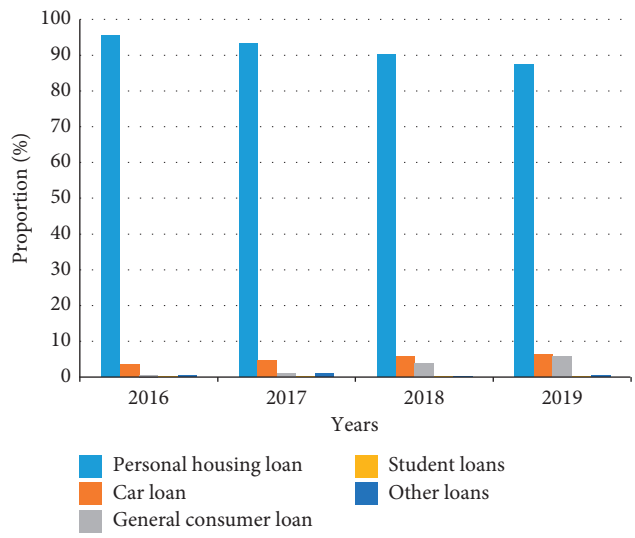

Figure 2: Percentage of various types of consumer credit by banks in each year.

Table 3: The status of short-term deposits and long-term loans.

	2016	2017	2018	2019
Short-term loan	1456.196	1682.534	2037.174	2427.352
Medium and long-term loans	782.484	1124.648	1672.047	2258.876
Proportion of medium and long-term loans	53.73%	66.84%	82.08%	93.06%

Table 4: Statistical table of means of independent variables.

	Independent variable	Overdue within two years	Overdue for more than 90 days	No overdue
Customer information	Age	53.18	46.73	45.52
	Monthly income	6583	5758.14	6815.25
Customer credit history	Overdue times	0.4523	2.1578	0.1964
	Ratio of debt	25.34	23.18	24.76
	Open loans and number of credit loans	8.342	8.539	8.694

Table 5: Credit risk overdue price matrix.

Element	18	19	20
Test value	0.6745	0.8523	0.7693

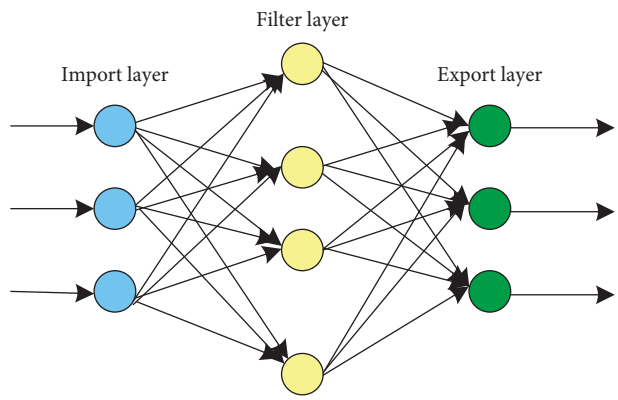

Figure 3: Basic structure of BP neural network.

Table 6: Comparison of the average performance evaluation index of each algorithm (rho = 0.5).

Rho = 0.5	Imbalance ratio	Accuracy	Recall rate	Precision	F1 measurement
SVM	10 : 1	0.8903	0.5234	0.5654	0.1211
Random forest		0.8852	0.4573	0.3891	0.3579
AdaBoost		0.9276	0.4978	0.7677	0.5537
Improve AdaBoost		0.8769	0.7851	0.5318	0.4891
SVM	20 : 1	0.9691	0.0785	1	0.1179
Random forest		0.9452	0.2967	0.5255	0.4123
AdaBoost		0.9783	0.4015	0.7976	0.6053
Improve AdaBoost		0.9216	0.6762	0.4013	0.5457

Table 7: Comparison of the average performance evaluation index of each algorithm (rho = 0.7).

Rho = 0.7	Imbalance ratio	Accuracy	Recall rate	Precision	F1 measurement
SVM	10 : 1	0.8872	0.0735	0.5654	0.5289
Random forest		0.8737	0.2617	0.3891	0.5773
AdaBoost		0.9308	0.2918	0.7677	0.3942
Improve AdaBoost		0.8685	0.5841	0.5318	0.5306
SVM	20 : 1	0.9707	0.5960	1	0.6971
Random forest		0.9419	0.3153	0.1318	0.6891
AdaBoost		0.9693	0.4278	0.7277	0.6368
Improve AdaBoost		0.9178	0.6932	0.3934	0.5571

improved algorithm has improved the classification accuracy of minority classes to a certain extent. Figure 4 shows the comparison of various performance indicators corresponding to different classification algorithms.

Compared with single logistic regression, the AdaBoost algorithm after integrating three weak classifiers has a slight improvement in minority class recognition, and at the same time, there is no major change in the recognition of majority classes, and it still maintains a high accuracy. The cost-sensitive improved AdaBoost algorithm has obvious advantages for the correct recognition of minority classes, but correspondingly, more majority classes are mistakenly classified into minority classes.

As shown in Table 8, the accuracy rate of single logistic regression is the highest, reaching 92.17%. The original AdaBoost uses logistic regression as the base classifier, which has a slight decrease in accuracy, but it is still at a relatively high level and improved. The overall accuracy of AdaBoost is 4.5% lower than the original AdaBoost. However, since the subject of this article is to identify minority samples in an imbalanced dataset, and the dataset used in this article is also highly imbalanced; simply distinguishing the sample as the

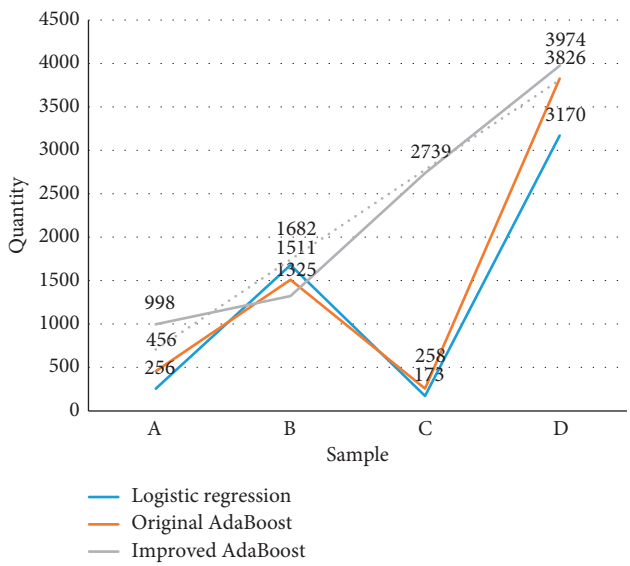

Figure 4: Comparison of logistic regression, original AdaBoost, and improved AdaBoost classification results.

Table 8: Algorithm performance comparison.

	Accuracy	Precision	Recall rate	F1 measurement
Logistic regression	0.9217	0.5257	0.0725	0.1143
Original AdaBoost	0.9103	0.5023	0.1324	0.1719
Improve AdaBoost	0.9093	0.4574	0.3158	0.3237

majority class will result in higher classification accuracy. Therefore, this indicator does not have absolute guiding significance.

5. Conclusion

Based on a large number of studies and sorting out domestic and foreign literature on big data credit investigation and consumer finance, this article researches that big data credit investigation is the future development trend of credit investigation, which can meet the financing needs of small and medium-sized enterprises. To reduce the credit risk of financial institutions and enrich the credit market, based on the practice of big data credit investigation in Internet financial risk control, a plan to use big data credit investigation to promote China's Internet financial risk control is proposed. Big data credit investigation mainly faces risks such as the legal risks, management risks, and network security risks of credit investigation agencies. At the same time, it proposes big data credit risk prevention measures from three aspects: improving governance structure and internal control, strengthening industry self-discipline construction, and strengthening industry supervision. According to the research content, the four aspects are as follows: promulgating and improving big data credit investigation laws and regulations, accelerating the construction of credit information sharing platforms, expanding the application scenarios of big data credit investigation, and strengthening the application research of blockchain technology proposed to improve China's big data investigation. Suggestions on the application of consumer credit in consumer finance credit risk management. The research method of this article is relatively simple, and only relatively simple theoretical analysis can be carried out. The depth of research on big data credit investigation in consumer finance credit risk management needs to be further improved.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The author declares no potential conflicts of interest with respect to the research, author-ship, and/or publication of this article.

References

[1] W. Li and H. Song, "ART: an attack-resistant trust management scheme for securing vehicular ad hoc networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 4, pp. 960–969, 2015.

[2] C.-Y. Liu, L.-J. Chiou, C.-C. Li, and X.-W. Ye, "Analysis of Beijing Tianjin Hebei regional credit system from the perspective of big data credit reporting," *Journal of Visual Communication and Image Representation*, vol. 59, pp. 300–308, 2019.

[3] G. Cullen, D. Gasbarro, G. S. Monroe, G. Shailer, and Y. Zhang, "Bank audit fees and asset securitization risks," *AUDITING: A Journal of Practice & Theory*, vol. 37, no. 1, pp. 21–48, 2018.

[4] P. Mohabeer, M. I. Santally, R. K. Sungkur et al., "An investigation of the potential benefits of big data in the public sector of Mauritius," *Journal of the Knowledge Economy*, vol. 10, no. 3, pp. 1230–1247, 2019.

[5] S. Kiatsupaibul, A. J. Hayter, and S. Somsong, "Confidence sets and confidence bands for a beta distribution with applications to credit risk management," *Insurance: Mathematics and Economics*, vol. 75, no. 1, pp. 98–104, 2017.

[6] J. J. Xiao, *Handbook of Consumer Finance Research Business-Owning Families: Challenges at the Intersection of Business and Family*, Springer, Berlin, Germany, 2016.

[7] E. Raguseo, "Big data technologies: an empirical investigation on their adoption, benefits and risks for companies," *International Journal of Information Management*, vol. 38, no. 1, pp. 187–195, 2018.

[8] Y.-C. Tsao, "Managing default risk under trade credit: who should implement Big-Data analytics in supply chains?" *Transportation Research Part E: Logistics and Transportation Review*, vol. 106, pp. 276–293, 2017.

[9] N. Metawa, M. Elhoseny, M. K. Hassan, and A. E. Hassanien, "Loan portfolio optimization using genetic algorithm: a case of credit constraints," in *Proceedings of 12th International Computer Engineering Conference, ICENCO 2016: Boundless Smart Societies 7856446*, pp. 59–64, Cairo, Egypt, Deember 2016.

[10] Y.-G. Fu and J.-M. Zhu, "The credit game on network supplier and customer based on Big Data," *Electronic Commerce Research*, vol. 18, no. 3, pp. 605–627, 2018.
[11] 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.
[12] N. Porto and J. J. Xiao, *Handbook of Consumer Finance Research*, Springer Publishing, New York, NY, USA, 2nd edition, 2016.
[13] M. Burke and J. Fry, "How easy is it to understand consumer finance?" *Economics Letters*, vol. 177, pp. 1–4, 2019.
[14] E. B. Taylor, "Consumer finance in a mobile age: methods for researching changing user behaviour," *Ethnographic Praxis in Industry Conference Proceedings*, vol. 2017, no. 1, pp. 197–215, 2017.
[15] D. Myles, "Averting a UK consumer credit crisis," *Banker*, vol. 168, no. 1108, p. 16, 2018.
[16] H. Amin, A. R. Abdul Rahman, D. Abdul Razak, and H. Rizal, "Consumer attitude and preference in the Islamic mortgage sector: a study of Malaysian consumers," *Management Research Review*, vol. 40, no. 1, pp. 95–115, 2017.
[17] L. A. Reisch, "Book notes "economics and social sciences" 3/2017," *Journal of Consumer Policy*, vol. 40, no. 3, pp. 413–418, 2017.
[18] Z. Lv, R. Lou, J. Li, A. K. Singh, and H. Song, "Big Data analytics for 6G enabled massive Internet of Things," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5350–5359, 2021.
[19] F. Butaru, Q. Chen, B. Clark, S. Das, A. W. Lo, and A. Siddique, "Risk and risk management in the credit card industry," *Journal of Banking & Finance*, vol. 72, pp. 218–239, 2016.
[20] A. Seetharaman, V. Kumar Sahu, A. Saravanan, J. Rudolph Raj, and I. Niranjan, "The impact of risk management in credit rating agencies," *Risks*, vol. 5, no. 4, p. 52, 2017.
[21] 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, 2018.
[22] T. Zhang, "Research on credit risk management of China's commercial banks," *World Scientific Research Journal*, vol. 5, no. 11, pp. 68–75, 2019.
[23] G. Elliehausen and S. M. Hannon, "The credit card act and consumer finance company lending," *Journal of Financial Intermediation*, vol. 34, pp. 109–119, 2018.
[24] Z. Huang, "Six obstacles to producing reliable," *Big-Data Credit Reports*, vol. 59, no. 50, p. 40, 2016.
[25] H.-D. Wu and L. Han, "A novel reasoning model for credit investigation system based on Fuzzy Bayesian Network," *Procedia Computer Science*, vol. 183, no. 19, pp. 281–287, 2021.
[26] M. Zhu and X. Liu, "Study on Internet finance credit information sharing based on block chain technology," *Asian Social Science*, vol. 14, no. 2, pp. 81–85, 2018.
[27] Z. Liu, L. Ma, Y. Zhu, and W. Ji, "An investigation on responsible innovation in the emerging shared bicycle industry: case study of a Chinese firm," *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 5, no. 3, pp. 42–46, 2019.