

Research Article

Research on Credit Risk Identification of Internet Financial Enterprises Based on Big Data

Hua Peng ^{1,2}

¹Wuyi University, Wuyishan, Nanping 354300, China

²National Changhua University of Education, Changhua 50007, China

Correspondence should be addressed to Hua Peng; penghua@wuyiu.edu.cn

Received 26 August 2021; Accepted 23 October 2021; Published 13 November 2021

Academic Editor: Fazlullah Khan

Copyright © 2021 Hua Peng. 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.

The advent of the era of big data has provided a new way of development for Internet financial credit collection. The traditional methods of credit risk identification of Internet financial enterprises cannot get the characteristics of credit risk zoning, leading to large errors in the results of credit risk identification. Therefore, this paper proposes a new method of credit risk identification based on big data for Internet financial enterprises. According to the big data perspective, the credit risk assessment steps of Internet financial enterprises are analyzed and the weight of assessment indicators is calculated using the improved analytic hierarchy process (AHP), and the linear weighted synthesis method is applied to comprehensively assess the credit of clients. Using the unique characteristics of big data credit risk region division, the big data credit risk is determined by rule-based matching method. The eXtreme Gradient Boosting (XGBoost) machine learning algorithm is used to establish a credit risk identification model of Internet financial enterprises. The kappa coefficient and ROC curve are used to evaluate the performance of the proposed method. Experimental results show that the proposed method can accurately assess the credit risk of Internet financial enterprises.

1. Introduction

Internet finance refers to a new financial form that, based on traditional finance, realizes functions such as payment, financing, and credit intermediaries under a series of Internet technologies such as big data technology and cloud computing technology [1]. The forms of Internet finance are generally third-party payment, financial e-commerce, credit evaluation, Internet money funds, big data index funds, and other models. Internet finance is the assimilation of the Internet field and the financial field. However, it is not the simple integration of the Internet and finance but the innovative transformation of financial business by using a safe and effective Internet network under certain market conditions. It is the fusion of traditional financial business and the Internet spirit. Internet finance is a new thing, its development speed and scale far exceed that of other industries, and its development prospect is great; therefore, the Internet finance industry is a sunrise industry. The risk of Internet finance is a big difficulty to the development of the

industry. The investigation on the risk management of Internet finance has become a hot topic in society and academia [2].

Jiang and Yuan [3] proposed a credit evaluation method in e-commerce transactions based on decision analysis. The credit risk influencing factors of business transactions are divided into several evaluation factors by introducing the Kanno doff integral evaluation method, and the credit risk F-integral evaluation model is constructed to divide the transaction credit grade. The interval distribution of commodity prices is used to add points to successful transactions and deduct points from failed transactions at different levels, to solve the problem of cycle deception in transactions. Based on the credit scoring results, the risk calculation method of e-commerce transactions is designed to evaluate the credit degree of the current transaction. Yang et al. [4] used the enterprise transaction data of an Internet financial platform as the object, analyzed the propagation behavior of overdue loan default, and proposed to build a model to identify the high-risk enterprises of the Internet financial

platform through the propagation characteristics. Based on constructing SIS and SIR models based on threshold propagation and random propagation, the model is transformed into an algorithm that can evaluate the enterprise value at risk and further verified and compared with the actual default data. However, the above two traditional methods cannot obtain the regional division characteristics of credit risk, resulting in large errors in the results of credit risk identification. Gao and Xiao [5] studied the management of credit risk for consumer finance using big data. Their risk management model exhibited good prediction ability, can discriminate between normal loan customers and default loan customers, and is appropriate for practical personal credit risk control business. Wang [6] preprocessed the Internet financial credit data and selected the variables for the active credit tracking model of BP neural network using adaptive genetic algorithm. Liu [7] studied the importance of machine learning and big data as an efficient data exploration approach for insurance risk management using random forest algorithm. Fatao et al. [8] employed supply chain financial credit risk indicators and built an online evaluation index model for supply chain financial credit risk in commercial bank. Shen [9] established effective financial risk early warning systems and techniques and took effective actions to avoid risks and to guarantee the normal operation of Internet banking. The financial risk early warning system based on large data will quickly expand under the financial Internet era background. Lyu and Zhao [10] investigated the use of compressed sensing in risk evaluation of Internet finance based on big data. Yang et al. [11] developed an Internet supply chain financial risk managing model through data science. Zhang [12] constructed a financial investment risk method with the help of an intelligent fuzzy neural network. Similarly, Teles et al. [13] proposed a credit risk prediction model applying artificial neural network (ANN) and Bayesian network models. Xu et al. [14] employed backpropagation neural network (BPNN) and information entropy to recognize and classify the risk of the bank branches. The neural network can solve the nonlinear problem without depending on the function setting to get a more precise simulation effect, so it can measure the estimate effect of the risk early warning model more precisely. In this study, we propose a new credit risk identification method for Internet financial enterprises based on big data. The XGBoost algorithm is employed to develop a credit risk identification model of Internet financial enterprises. The performance of the model is evaluated using the kappa coefficient and ROC curve. Results show that the proposed risk identification model can accurately measure the credit risk of Internet financial enterprises.

The rest of the paper is arranged as follows. In Section 2, the index weight calculation method is discussed. Section 3 illustrates the proposed XGBoost algorithm for the identification of credit risk. The results are given in Section 4, and Section 5 concludes the paper.

2. Credit Risk Detection of Internet Financial Enterprises Based on Big Data

When designing the enterprise credit risk assessment index system, it is essential to consider the defects of the traditional enterprise credit risk assessment index systems [15]. In the development of a credit risk assessment index system, the integration of dynamic data and static data is adopted to complete data mining, analysis, and modeling from big data. Moreover, the enterprise identity data, behavior data, and external data are used to construct an enterprise credit risk assessment index system [16]. It is crucial that the enterprise credit risk evaluation index system shall be mainly designed from the perspectives of income information, loan information, account information, repayment and overdue information, and third-party information, such as multi-end loan information, black and gray list, credit information, and so on. The enterprise credit risk evaluation index system based on big data has the characteristics of rich evaluation data items, a combination of static data and dynamic data, wide data sources, and timeliness. Although the traditional enterprise credit risk assessment index system in the crowd coverage is not ideal, the data are mostly static, and the data authenticity cannot be verified. The data of the big data platform for Internet enterprises are mostly used from multiple channels, such as APP product data of companies, credit information data of the PBC, credit information data of Internet credit information companies, online shopping data of e-commerce companies, data of third-party cooperative enterprises, public data crawled by crawlers, and data published by the Public Security and Inspection Law. The data used in this study are desensitized from many aspects to ensure the safety of the data.

2.1. Credit Evaluation Index Weight Calculation. The improved analytic hierarchy process (AHP) method [17] is used to compute the weight of evaluation indicators. This method uses the concept of optimal transfer matrix to improve AHP, make it naturally meet the consistency requirements, and directly calculate the weight value. The major steps are as follows:

- (i) A credit risk evaluation index system is established based on market transactions and the evaluation index is set according to the customer credit evaluation index system.
- (ii) A judgment matrix is built. After establishing the matrix according to the credit risk evaluation index system, the weight of each level index in the customer credit risk evaluation index system is determined by AHP. By comparing the two elements, the relative importance of each element in the hierarchy relative to a certain factor in the upper hierarchy is determined, and a judgment matrix is created. A comparison matrix of two factors for a given criterion is calculated as follows:

$$Y = (y_{ij})_{\min}, \quad i, j = 1, 2, \dots, n, \quad (1)$$

where y_{ij} is the scale of the importance of factor Y_i and factor Y_j relative to the criterion index.

- (iii) The weight of evaluation indexes at all levels is calculated based on the improved AHP method. There is no need to do a consistency check after calculating the weights of indexes by the improved AHP method. Firstly, the judgment matrix is modified to get the quasi-optimal matrix Y^* , followed by the square root method is used to solve the Y^* eigenvector. Next, the elements of the judgment moment Y^* are multiplied by the lines to get the following expression:

$$M_i = \prod_{j=1}^n y_{ij}, \quad i = 1, 2, 3, \dots, n. \quad (2)$$

By dividing the product into n power roots, an equation $P_i = N_i^{1/n}$ is obtained, and the root vector $P = [P_1, P_2, \dots, P_n]^Q$ is normalized, i.e., the sorting weight vector P can be calculated as

$$P_i = \frac{P_i}{\sum_{j=1}^n P_j}. \quad (3)$$

The improved AHP method is represented in Figure 1.

2.2. Credit Evaluation Based on Linear Weighted Synthesis Method. The linear weighted synthesis method for customer credit evaluation is a kind of comprehensive method to obtain the comprehensive evaluation value by weighted summation of each index value.

The evaluation value of the credit risk of the first j customer can be calculated as

$$F = \frac{\sum_{i=1}^n Y_{ji} \times M_i}{Q_{ji}}, \quad (4)$$

where the index value of the i^{th} index of the j^{th} customer is Q_{ji} , $j = 1, 2, \dots, m$.

The static client credit risk rating is classified into four levels, with grade 1 denoting 89–100 points, grade 2 denoting 75–88 points, grade 3 denoting 59–74 points, and grade 4 denoting ≤ 58 points. The weighted credit scoring method is used in static client credit risk rating. The first step is to score some credit survey indicators as a whole, the second step is to uniformly weighted average, and the last step is to get the method of credit risk score. This can be computed using the following equation:

$$X = \sum_{i=1}^n (b_i \cdot Y_i), \quad (5)$$

where the credit score of a customer is represented by X , The weight of the proposed i^{th} credit investigation index is represented by b_i and set to $\sum_{i=1}^n (b_i \cdot Y_i) = 1$. The evaluation score of the i^{th} evaluation index is shown by Y_i . This method is applied to quantitative analysis and research of customer credit. The theoretical basis of financial enterprise decision making is mainly to obtain the credit status of the enterprise

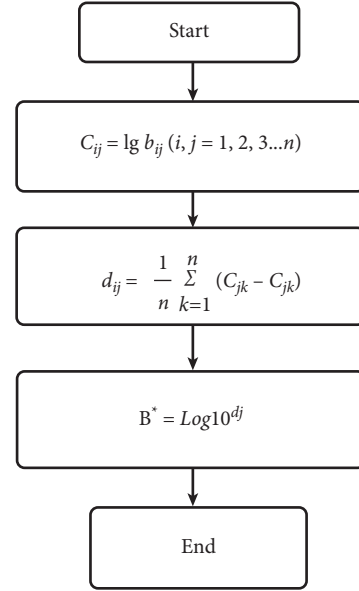


FIGURE 1: Improved AHP.

based on objective data, while the static client credit rating method is a simple and easy method to understand and operate.

2.3. Credit Risk Detection Based on Big Data. According to the regional division characteristics of big data credit risk [18], the credit risk is detected by rule-based matching method. The centralized detection process is shown in Figure 2.

$$\text{Max } f(x) = \frac{1}{\sqrt{\sum_{i=1}^n (b_i(i) - y_i(i))}}, \quad (6)$$

The data are divided into different packets and matched. If the matching process is successful, the output is generated. The credit risk of big data is divided into five areas and the centralized detection problem will be converted to the target maximum value, and read the data of the fitness function before the processing step, where $y_k(x)$ is the actual output of data, $y'_k(x)$ represents the expected output of data, and k shows the total amount of data obtained.

The matching process of packets given in Figure 2 is as follows.

2.3.1. High Credit Risk Data Matching Rules. In the process of matching from right to left, if the character in the string does not match its corresponding string, it shall be handled as follows.

If $x[i]$ data are in a string $S[1, \dots, n]$, the string can be moved to $x[i + 1]$ character alignment, as shown in Figure 3.

$H(x)$ is set to indicate the height at which the string S moves upward, l indicates the length of the string, $R(x)$ indicates that the string S appears at the position closest to the right of the data x , and i indicates that the position below the position does not match (the length from the far left), as shown below:

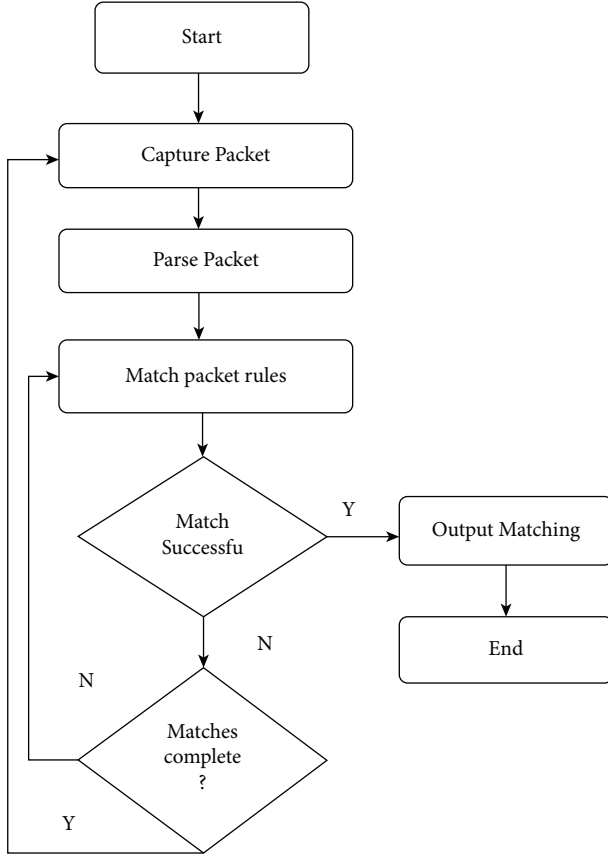


FIGURE 2: Centralized detection process.

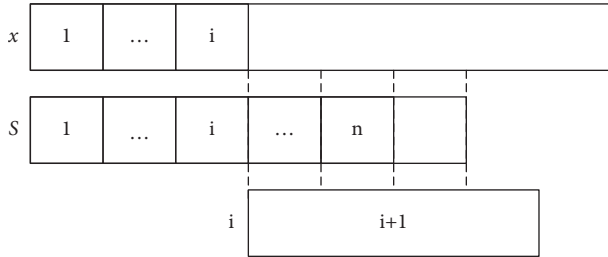


FIGURE 3: Schematic diagram of character jump of high credit risk data.

$$H(x) = \begin{cases} l - i, & \text{the data } X \text{ are not in the string,} \\ i, & \text{the data } X \text{ are in the string.} \end{cases} \quad (7)$$

2.3.2. Matching Rules for Low Credit Risk Data. When different strings are aligned, the moving distance is determined by matching rules of low credit risk data during the matching process from right to left. The specific matching process is as follows.

The string S is entered and the moving distance is initialized. Next, the string is traversed from right to left, and the traversal position i is analyzed as given in the following equation:

$$\begin{cases} i < H(S[i]), & H(S[i]) = l - i, \\ i > H(S[i]), & H(S[i]) \text{ remain the same.} \end{cases} \quad (8)$$

All the strings are aligned and matched one by one from right to left. If it is matched with the leftmost end of the string, it means that the matching is successful. According to the matching process, it can ensure that any distance is a safe match and no omission will occur. Furthermore, it can achieve centralized and precise detection of big data credit risk areas at the fastest matching speed.

3. Credit Risk Identification Model of Internet Financial Enterprises

Based on the XGBoost algorithm [19], the credit risk identification model of Internet financial enterprises is constructed. It is a common and effective open-source implementation of the gradient boosted trees algorithm. The XGBoost algorithm provides better performance because of its vigorous handling of different data types, distributions, relationships, and the variety of hyperparameters that can be fine-tuned [20]. The XGBoost algorithm can be used for regression, binary and multiclass classification, and ranking problems [21]. The basic element of constructing the XGBoost model is a tree set, and the binary tree structure is in the classification and regression tree which can reflect the actual result of the decision tree [22]. There are two branches in the structure of the decision tree, namely, “Yes” and “No,” which correspond to the right and left branches. Each feature variable is partitioned by the binary tree, and the feature space is partitioned to obtain several leaf nodes.

Suppose a set $B = \{(x_i, y_i)\}$, in which there are m variables and n samples. Through K functions, the output \hat{B} of the prediction model can be obtained based on the regression tree integration model:

$$\hat{B} = \sum_{k=1}^T f_k(x_i), \quad f_k \in \Gamma, \quad (9)$$

where $\Gamma = \{f(x) = \omega_{q(x)}\} (q: R^m \rightarrow T, \omega_i \in R^m)$ represents the regression tree space; ω_i shows the score corresponding to the i leaf; T is the number of leaf nodes present in the tree structure; q indicates the tree structure; f_k represents the tree; and x_i is the independent variable corresponding to the i^{th} sample.

The tree model is trained with the objective function u as given in the following equation:

$$u = \sum_i o(\hat{y}_i, y_i) + \sum_k \Omega(f_k), \quad (10)$$

where o is the convex loss function to measure the difference between the real value y_i and the predicted value \hat{y}_i and Ω stands for penalty term, and its expression is as follows:

$$\Omega(\cdot) = \gamma + \frac{1}{2} \|\phi\|, \quad (11)$$

where $\|\phi\|$ represents the regular term and γ stands for leaf node penalty, which is mainly used to avoid overfitting problems.

In the process of credit risk identification of Internet financial enterprises, the European space cannot be directly used to optimize the objective function. Therefore, the credit risk identification model is trained through boosting learning strategy model, and the specific process is as follows:

$$\begin{cases} \widehat{B}_i^{(0)} = 0, \\ \widehat{B}_i^{(1)} = f_1(x_i) = \widehat{B}_i^{(0)} + f_1(x_i), \\ \widehat{B}_i^{(2)} = f_1(x_i) + f_2(x_i), \\ \vdots \\ \widehat{B}_i^{(t)} = \sum_{k=1}^t f_k(x_i) = \widehat{B}_i^{(t-1)} + f_t(x_i), \end{cases} \quad (12)$$

where $\widehat{B}_i^{(t)}$ represents the function newly added to Round t .

Based on the above process, the objective function can be converted as

$$\vartheta = \sum_{i=1}^n l\left(B, \widehat{B}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) + a, \quad (13)$$

where a is a constant.

The fitting results of the model and training data in the identification process can be measured by the loss function $L = \sum_{i=1}^n l(\widehat{B}_i, B)$, in which the logical loss function $l(\widehat{B}_i, B_i) = y_i \ln(1 + e^{-B})$ and the square loss function $l(\widehat{B}_i, B_i) = (B_i - \widehat{B}_i)^2$ are widely used in the identification process. The credit risk identification model of Internet financial enterprises based on the RB-XGBoost algorithm is brought into the square loss function in the objective function and the following equation is obtained.

$$\vartheta = \sum_{i=1}^n l\left(B_i, \left(\widehat{B}_i^{(t-1)} + f_t(x_i)\right)\right)^2 + \Omega(f_t) + a, \quad (14)$$

where $\widehat{B}_i^{(t-1)}$ stands for residual.

The loss function can be approximated by a Taylor expansion to obtain the following expression:

$$f(x + \Delta x) \approx f(x) + f'(x)\Delta x + \frac{1}{2}f''(x)(\Delta x)^2. \quad (15)$$

When the loss function is a square loss during training, the following equation can be computed.

$$\begin{cases} g_i = \partial_{y_i} l\left(B_i, \widehat{B}_i^{(t-1)}\right) = 2\left(\widehat{B}_i^{(t-1)} - B_i\right), \\ h_i = \partial_{y_i}^2 l\left(B_i, \widehat{B}_i^{(t-1)}\right) = \partial_{y_i}^2 \left(\widehat{B}_i^{(t-1)} - B_i\right)^2. \end{cases} \quad (16)$$

Substitution of parameters g_i and h_i in the objective function yields the following expression:

$$\vartheta = \sum_{i=1}^n \left[B_i - \left(\widehat{B}_i^{(t-1)} + f_t(x_i)\right) \right]^2 + \Omega(f) + a, \quad (17)$$

where $\widehat{B}_i^{(t-1)}$ represents the output of the model in Round $t-1$ training and B_i represents the dependent variable in the objective function, and if B_i is known, the above objective function [20] can be simplified to obtain the following expression:

$$\vartheta = \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f) + a, \quad (18)$$

where g_i and h_i are the parameters in the loss function. In different loss functions, the values of the above parameters are different; therefore, the values of g_i h_i are determined by the loss function. Hence, each tree is redefined by the following equation.

$$f_t(x) = \omega_{q(x)}, \quad \omega \in R^T, \quad q: R^d = \{1, 2, \dots, T\}, \quad (19)$$

where ω represents the weights of leaf nodes in the tree structure, $\omega_{q(x)}$ denotes the predicted values obtained by the tree model, and $q: R^d = \{1, 2, \dots, T\}$ C shows the tree structure.

Model complexity includes two parts: L2 regularization of leaf node score and a total number T of leaf nodes. Model complexity $\Omega(f_t)$ can be obtained by tree definition:

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2. \quad (20)$$

The smoothness of leaf nodes can be improved by L2 regularization to solve the overfitting problem. When the complexity of the model increases, there are two different types of accumulation, one of which is $I_j = \{i | q(x_i) = j\}$, where I_j represents the set of samples in a leaf node j . After adding complexity to the objective function, the final objective function, that is, the credit risk identification model of Internet financial enterprises, is obtained:

$$\mu = \sum_{i=1}^n \left[g_i f_t(x_i) + h_i f_t^2(x_i) \right] - \lambda \sum_{j=1}^T \omega_j^2. \quad (21)$$

The credit risk identification model of Internet financial enterprises constructed above is used to complete risk identification.

4. Experimental Results and Analysis

To examine the credit risk evaluation ability of the proposed credit risk identification model of Internet financial enterprises, help the enterprises to avoid the risk of electricity arrears, and provide the basis for urging the payment of electricity fees, five Internet financial enterprises in a certain city are selected as experimental objects, and the annual reports of these five companies in the recent three years are

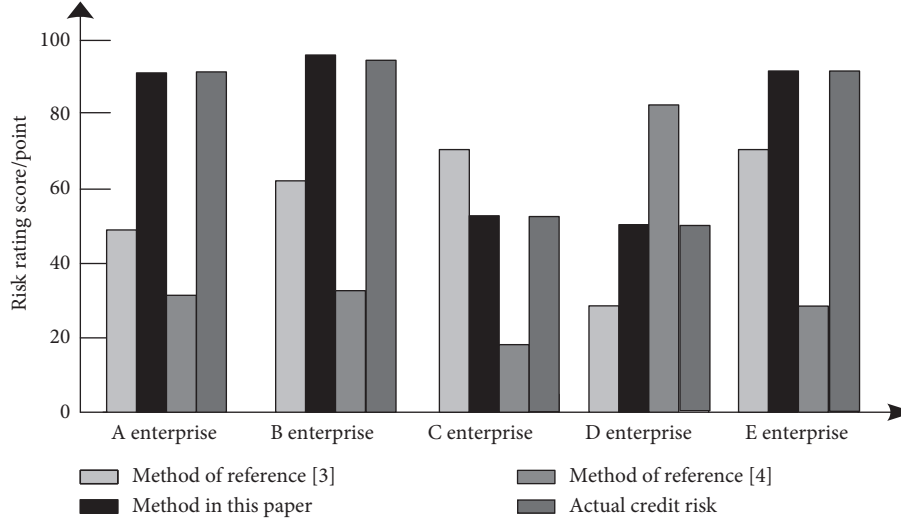


FIGURE 4: Comparison of risk rating scores of different methods.

selected as experimental data samples to carry out the experimental analysis.

4.1. Accuracy Test of Credit Risk Assessment with Different Methods. To evaluate the credit risk of five selected Internet financial enterprises, the proposed method is used along with the methods given in [3, 4], and the results of these methods are compared with the actual credit risk of the 5 companies. The experimental results are shown in Figure 4.

It can be seen that the credit risk grade scores of the five enterprises evaluated by the proposed method are closer to the actual credit risk grade scores of the five enterprises. This is because the proposed method can effectively combine the actual and objective data of the market transaction data of five companies to obtain the credit status of the enterprises. In addition, the proposed method can accurately obtain detailed information of the credit risk of the enterprises and establish an accurate credit risk evaluation index system, which makes the results of credit risk evaluation more scientific and accurate. The comparison results for the credit risk assessment accuracy of the two methods are shown in Figure 5.

By analyzing Figures 4 and 5, we can see that this method can effectively measure and evaluate all credit risk indicators and obtain the credit risk of enterprises. When the number of samples reaches 20000, the credit risk evaluation accuracy of references [3, 4] is 0.50% and 0.30%, respectively, whereas the accuracy of credit risk evaluation of the proposed method reaches 0.92%.

4.2. Kappa Coefficient and ROC Curve Test of Different Methods. To verify the recognition accuracies of all the three methods, the kappa coefficient and ROC curve are used. The kappa coefficient can weigh the difference between the predicted result and the real result. The kappa coefficient K can be computed as

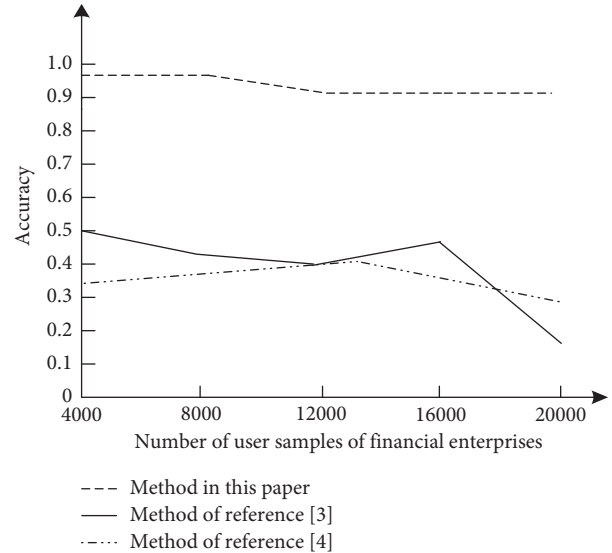


FIGURE 5: Comparison results of credit risk assessment accuracy of the two methods.

$$K = \frac{p_o - p_e}{1 - p_e}, \quad (22)$$

where p_o represents the proportion of correctly identified samples in the total number of samples and p_e is the randomness ratio. The higher the kappa coefficient K is, the more accurate is the recognition result of the method. The kappa coefficients of the proposed method, the method in reference [3], and the method in reference [4] are shown in Table 1.

It is evident that the kappa coefficients of the proposed method are higher than those of reference [3] and reference [4] in multiple iterations, which indicates that the proposed method can accurately identify the credit risks of Internet financial enterprises. This is because the method constructs a risk identification index system based on the data with high

TABLE 1: Kappa coefficients of different methods.

| Number of iterations/time | Kappa coefficient | | |
|---------------------------|-------------------|---------------|---------------|
| | Proposed method | Reference [3] | Reference [4] |
| 1000 | 0.99 | 0.84 | 0.69 |
| 2000 | 0.98 | 0.61 | 0.74 |
| 3000 | 0.97 | 0.66 | 0.66 |
| 4000 | 0.97 | 0.77 | 0.72 |
| 5000 | 0.98 | 0.72 | 0.81 |

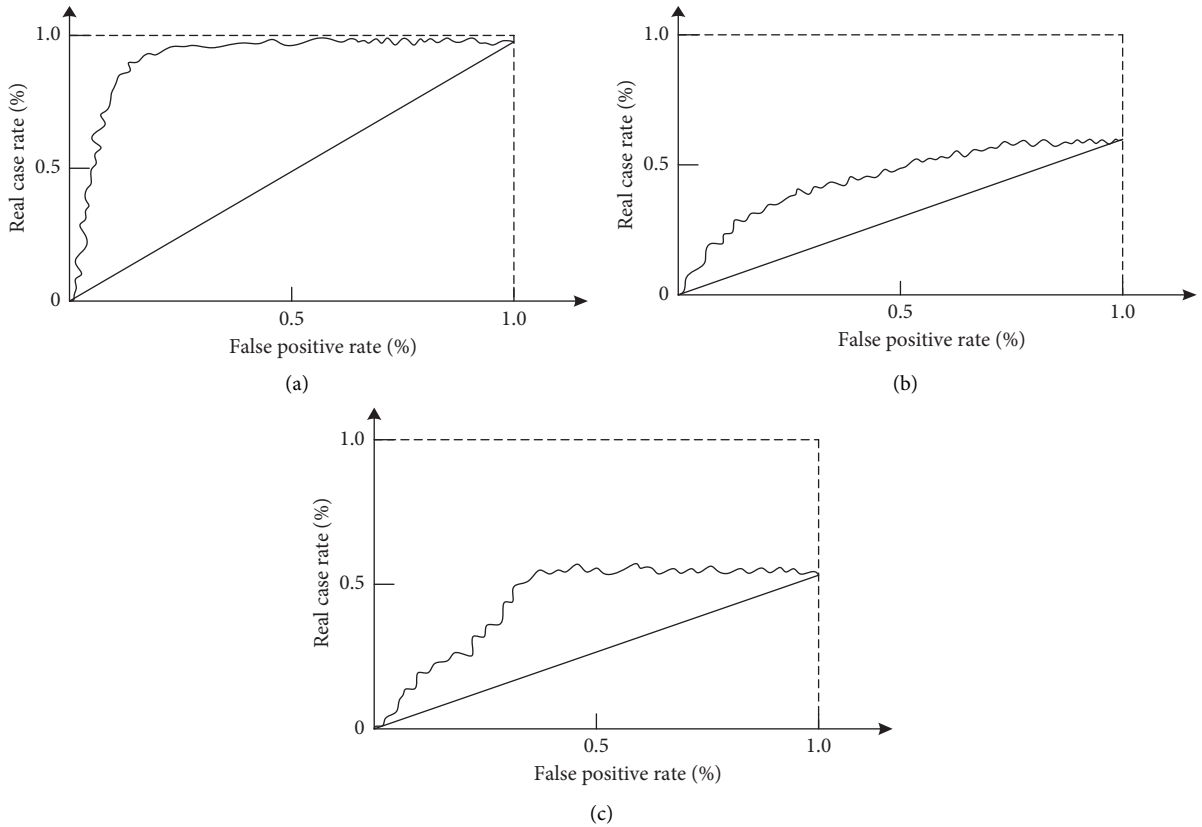


FIGURE 6: ROC curves of different methods. (a) Proposed method. (b) Reference [3]. (c) Reference [4].

balance and completes the identification of credit risks of Internet financial enterprises based on the high-precision risk identification indexes.

In the ROC curve, the abscissa is the real case rate and the ordinate is the false-positive case rate. The larger the area enclosed by the ROC curve and the abscissa is, the higher is the recognition accuracy of the method. The proposed method and the methods given in reference [3] and reference [4] are used to identify the credit risks of different Internet financial enterprises, and the obtained ROC curve is shown in Figure 6.

Figure 6 shows that the area of ROC curves and abscissa obtained by the proposed method is larger than the area of ROC curves and abscissa obtained by the methods proposed in [3, 4], indicating that the accuracy of the proposed method is higher, and the identification of credit risks can be accurately completed in the Internet financial enterprises.

5. Conclusion

The rapid development of Internet finance provides new financing channels for the development of small and microenterprises and individual entrepreneurship. The conventional methods of credit risk prediction of Internet financial enterprises cannot get the characteristics of credit risk zoning, leading to large errors in the results of credit risk identification. In this study, a new method of credit risk identification based on big data for Internet financial enterprises is proposed. The risk evaluation steps of Internet financial enterprises are studied, and the importance of assessment indicators is measured using the improved AHP method. The linear weighted synthesis method is employed to systematically assess the credit of clients. Based on the unique characteristics of big data credit risk region division, the big data credit risk is predicted with the help of rule-based matching method. The XGBoost supervised machine

learning algorithm is used to develop a credit risk prediction model of Internet financial enterprises. The performance of the model is evaluated with the kappa coefficient and ROC curve. Experimental results show that the proposed method can correctly assess the credit risk of Internet financial enterprises.

Data Availability

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

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

- [1] W. Xie, "Spatial aggregation analysis and systemic risk prevention of internet finance based on neural network model," *Journal of Physics: Conference Series*, vol. 1744, no. 4, pp. 176–185, 2021.
- [2] M. A. Ahad and R. Biswas, "Request-based, secured and energy-efficient (RBSEE) architecture for handling IoT big data," *Journal of Information Science*, vol. 45, no. 2, pp. 227–238, 2019.
- [3] Y. H. Jiang and X. Yuan, "Simulation research on accurate evaluation of credit degree in e-commerce transaction," *Computer Simulation*, vol. 35, no. 7, pp. 430–433, 2018.
- [4] X. H. Yang, B. H. Guo, and Z. Mi, "Identification algorithm of high breaching risk member for internet financial platform," *Application Research of Computers*, vol. 36, no. 329(3), pp. 58–67, 2019.
- [5] L. Gao and J. Xiao, "Big data credit report in credit risk management of consumer finance," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 4811086, 7 pages, 2021.
- [6] F.-P. Wang, "Research on application of big data in internet financial credit investigation based on improved GA-BP neural network," *Complexity*, vol. 2018, Article ID 7616537, 16 pages, 2018.
- [7] Q. Liu, "Research on risk management of big data and machine learning insurance based on internet finance," *Journal of Physics: Conference Series*, vol. 1345, Article ID 052076, 2019.
- [8] W. Fatao, L. Ding, H. Yu, and Y. Zhao, "Big data analytics on enterprise credit risk evaluation of e-business platform," *Information Systems and E-Business Management*, vol. 18, pp. 311–350, 2020.
- [9] Y. Shen, "Study on internet financial risk early warning based on big data analysis," *Business Economics and Management*, vol. 4, pp. 92–96, 2014.
- [10] X. Lyu and J. Zhao, "Compressed sensing and its applications in risk assessment for internet supply chain finance under big data," *IEEE Access*, vol. 7, pp. 53182–53187, 2019.
- [11] Q. Yang, Y. Wang, and Y. Ren, "Research on financial risk management model of internet supply chain based on data science," *Cognitive Systems Research*, vol. 56, pp. 50–55, 2019.
- [12] J. Zhang, "Investment risk model based on intelligent fuzzy neural network and VaR," *Journal of Computational and Applied Mathematics*, vol. 371, Article ID 112707, 2020.
- [13] G. Teles, J. J. P. C. Rodrigues, R. A. L. Rabê, and S. A. Kozlov, "Artificial neural network and Bayesian network models for credit risk prediction," *Journal of Artificial Intelligence and Systems*, vol. 2, no. 1, pp. 118–132, 2020.
- [14] M. Xu, D. Xiong, and M. Yang, "Risk recognition and risk classification diagnosis of bank outlets based on information entropy and BP neural network," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 2, pp. 1531–1538, 2020.
- [15] P. Saeidi, S. P. Saeidi, S. Sofian, S. P. Saeidi, M. Nilashi, and A. Mardani, "The impact of enterprise risk management on competitive advantage by moderating role of information technology," *Computer Standards & Interfaces*, vol. 63, pp. 67–82, 2019.
- [16] F. Abastante, S. Corrente, S. Greco, A. Ishizaka, and I. M. Lami, "A new parsimonious AHP methodology: assigning priorities to many objects by comparing pairwise few reference objects," *Expert Systems with Applications*, vol. 127, pp. 109–120, 2019.
- [17] A. Louhichi and Y. Boujelbene, "Credit risk pricing and the rationality of lending decision-making within dual banking systems: a parametric approach," *Economic Systems*, vol. 44, no. 1, Article ID 100740, 2020.
- [18] X. Yu, Q. Peng, L. Xu, F. Jiang, J. Du, and D. Gong, "A selective ensemble learning based two-sided cross-domain collaborative filtering algorithm," *Information Processing & Management*, vol. 58, no. 6, Article ID 102691, 2021.
- [19] R. Geetha, P. dmavathy, and R. Anitha, "Prediction of the academic performance of slow learners using efficient machine learning algorithm," *Advances in Computational Intelligence*, vol. 1, no. 4, pp. 1–12, 2021.
- [20] H. Zhang, D. Qiu, R. Wu, Y. Deng, D. Ji, and T. Li, "Novel framework for image attribute annotation with gene selection XGBoost algorithm and relative attribute model," *Applied Soft Computing*, vol. 80, pp. 57–79, 2019.
- [21] X. Yu, Y. Chu, F. Jiang, Y. Guo, and D. Gong, "SVMs classification based two-side cross domain collaborative filtering by inferring intrinsic user and item features," *Knowledge-Based Systems*, vol. 141, pp. 80–91, 2018.
- [22] M. Yu, T. Quan, Q. Peng, X. Yu, and L. Liu, "A model-based collaborate filtering algorithm based on stacked autoencoder," *Neural Computing & Applications*, vol. 2021, 2021.