Hindawi
Mobile Information Systems
Volume 2022, Article ID 3950210, 12 pages
https://doi.org/10.1155/2022/3950210

*Research Article*

# Credit Risk Assessment Modeling Method Based on Fuzzy Integral and SVM

**Mingyi Zhou**

*School of Science, Yangzhou Polytechnic Institute, Yangzhou, Jiangsu Province 225127, China*

Correspondence should be addressed to Mingyi Zhou; zhoumy@ypi.edu.cn

Received 10 March 2022; Revised 14 April 2022; Accepted 20 April 2022; Published 17 May 2022

Academic Editor: Chia-Huei Wu



With the development of financial globalization and financial market volatility, credit risk has become more prominent and serious, and how to establish an effective enterprise credit evaluation system and bank credit risk evaluation model, provide scientific quantitative decision-making basis for bank decision-making, reduce non-performing loans, and improve the quality of credit assets is a common research topic faced by domestic banks. At present, domestic banks have not been effective to establish risk prevention as the core of credit culture and long-term mechanism, the existence of nonperforming loans is still not fully resolved, new risks continue to appear, and there is a lack of a perfect and effective credit risk evaluation system. With the development of the Internet and financial institutions and the fusion, banks and financial institutions drastically increase the recorded data, and this provides a good prerequisite for the application of intelligent algorithms. In view of the shortcomings of BP neural network in the establishment of credit risk assessment model, such as poor promotion ability and long prediction time, and considering that support vector machine (SVM) can deal with some multi-classification problems, this paper introduces SVM method into the field of bank credit risk assessment and establishes an optimization model of credit risk assessment. This paper discusses the structure and algorithm principle of SVM classification method and proposes an integrated SVM based on fuzzy integral to solve this kind of problem. The results show that the algorithm can effectively improve the prediction accuracy, solve the problem of high computation cost, reduce the occupied memory space, improve the operation efficiency, shorten the training time, and provide a more reliable basis for the rapid and effective evaluation of bank credit risk. On the one hand, the research results expand the application of artificial intelligence technology in the field of economic research; the evaluation model can continuously and accurately measure credit risk is obtained, which provides the necessary basis for upgrading and optimizing credit decision-making, so it has high theoretical value and practical value.

## 1. Introduction

As the "general hub" of the national economy and the financial credit center, banks play an irreplaceable role in financing funds, guiding the flow of funds, and adjusting the balance of social supply and demand. However, banks face various financial risks all the time in the process of operation, including credit risk, interest rate risk, liquidity risk, management risk, capital risk, and policy risk, among which credit risk occupies a special important position and is the most important factor leading to bank bankruptcy. The World Bank's study of global banking crises shows that credit risk is the most common cause of bank failures. The bank's credit evaluation of customers is the core content of commercial loans [1]. Whether the credit evaluation of bank customers is reasonable, scientific, and accurate is related to the risk of bank loans. Bank credit risk refers to the possibility or probability that the borrower cannot repay the principal and interest of the bank loan in accordance with the contract during the operation of the bank, resulting in the loss of the expected income of the credit assets. As the financial globalization process accelerating, increasing complexity of financial market, the importance of bank credit risk and the complexity of the measure also will increase, the traditional credit risk evaluation model has not met the needs of today's financial risk management system, and in-depth research on credit risk evaluation model is of important theoretical significance and practical value. At present, the bank

credit risk assessment method is relatively backward, which is highlighted as follows: the assessment result of borrower credit grade is too simple, and the risk disclosure is seriously insufficient; the basic data of customers is seriously lacking, and the risk model is difficult to establish. The risk rating method is simple, the technology is backward, and the scientific measurement model is lacking [2, 3].

Bank credit risk assessment is a relatively complex system. Simply speaking, the input of the system is the factors affecting credit risk, the output of the system is the final quantitative result of risk assessment, and the intermediate processing module of the system is the optimal mapping relationship found between the input and output. As the input of this complex system, the evaluation index of credit risk plays an important role in the validity and accuracy of evaluation results. Therefore, the primary task for banks in credit risk assessment and management is to establish a credit risk assessment index system guided by the principle of index selection and based on the analysis of factors affecting credit risk, laying a foundation for the risk assessment and management of banks [4]. As the output end of the system, the measurement standard of credit risk not only determines the availability of credit risk assessment results but also affects the selection of credit risk assessment methods and indicators and reflects the purpose of bank credit risk assessment.

At present, enterprise credit risk assessment methods include neural network, support vector machine (SVM), decision tree, logistic regression, Bayesian network, and a series of integration methods. As the sample of loan enterprises is difficult to meet the assumptions of discriminant analysis and statistical analysis, the application of the model is very limited. With the expansion of credit business, the demand for credit risk assessment is increasingly strong. A small improvement in credit assessment accuracy can bring huge benefits to banks. Therefore, people began to actively explore more accurate evaluation methods, and more nonparametric methods and artificial intelligence methods were developed successively, such as decision tree method, neural network method, and SVM. In an increasingly complex market environment, AI black-box simulation has more advantages than statistical analysis [5–7]. The research shows that the integrated algorithm and the combined algorithm have higher prediction accuracy than the single algorithm. However, the application results of neural network integration show that the application of neural network integration method in credit risk assessment is limited, and the prediction accuracy of neural network integration method is not as good as that of single neural network. The main reason is that the neural network method is established on the basis of large samples for model training. However, at present, the data of loan enterprises retained by banks is limited, and the integration algorithm needs to divide the sample data into multiple small samples, which reduces the training sample size, thus affecting the prediction accuracy of the neural network integration model. SVMs have the advantage of small sample learning. Based on this, scholars began to turn to the research of SVM integration methods, including the integration of SVM and statistical methods, and the integration of multiple sub-SVMs. SVM integration has been widely used in many fields to test prediction accuracy. The application results show that the prediction accuracy of integrated SVM is at least as good as that of single SVM. Because SVM has the advantage of small sample learning, under the condition of limited sample size of bank loan enterprises, it is more advantageous to adopt SVM integrated method to evaluate enterprise credit risk. In addition, the majority voting method is generally adopted by scholars for SVM integration, but this method does not consider the important difference of the output results of sub-SVM classification trainers.

Based on this, this paper through the analysis of the characteristics of the loan enterprises is combined with the characteristics of the owners of the loan method as an important factor, to build a comprehensive reflection of the enterprise credit risk evaluation index system. On the premise of considering the importance of learning results of each subtrainer, fuzzy integral and SVM regression integration model is constructed to evaluate enterprise credit risk and measure enterprise credit risk accurately. Put forward a kind of SVM based on fuzzy integral financial credit risk evaluation model of supply chain integration, this model not only considers the objective information of child SVM output but also uses fuzzy density to determine each child SVM to the importance of the output information for the final decision, using the data of a certain commercial bank are empirical study, investigate the classification accuracy of the model. Application results show this method than single SVM method compared with SVM fusion method based on the voting strategy, prediction accuracy is higher, statistical method and the neural network can overcome the loan companies less sample size and the defect of nonnormal distribution, and objective assessment of loan enterprise credit risk exposure break through the traditional pattern of risk assessment classification information reflects the lack of limitations, to meet the risk management requirements of bank loan decision-makers, improve credit risk measurement and management techniques, and establish a scientific credit risk management system. This topic will introduce the research background of this paper, review the existing research literature and prediction and evaluation method based on support vector machine, and elaborate the research background; then establish the purpose of the research; construct the technical route framework of this research; finally, the characteristics and improvement of this study are summarized.

## 2. Integration Model of Fuzzy Integral and SVM

*2.1. Basic Principles of SVM.* At present, the data mining techniques used to establish the bank credit risk assessment model include discriminant analysis, logistic regression, principal component analysis, and other statistical methods, which have the shortcoming of low prediction accuracy. With the increase of credit risk factors, banks have higher requirements for the technology used in credit risk assessment models. Traditional statistical methods can no longer meet the needs of banks [8]. Artificial neural network, expert

system, and other intelligent technologies have been widely used. The artificial neural network is a good approximation capability of nonlinear mapping, and implementation is a major cause of impress the people favor, but with the wide application, more and more researchers have found that neural network because of the lack of theoretical support is hard to overcome some shortcomings. The introduction of kernel function can make the input space map to higher dimensional space to obtain better linear representation and greatly reduce the training time [9, 10]. These advantages of SVM overcome the above shortcomings of neural network, which is the reason for the selection of SVM in this paper.

SVM firstly maps input data to a high-dimensional feature space through mapping function $\varphi(\cdot)$ and finds the optimal separation hyperplane with the maximum classification spacing. The schematic diagram of optimal partition hyperplane is shown in Figure 1.

The optimal hyperplane expression is as follows:

$$y = \omega^{\mathrm{T}}\varphi(x) + b = 0, \tag{1}$$

where $x = (x_1, x_2, \cdots, x_N)^{\mathrm{T}}$ is the number of samples, $N$ is the normal vector of the hyperplane, $\omega$ is the scalar deviation. According to the structural risk minimization principle, classified hyperplane problems can be solved by the following optimization problems:

$$\begin{cases} \text{Minimize } J(\omega, b\,; \xi_i) = \dfrac{1}{2}\omega^{\mathrm{T}}\omega + \gamma\displaystyle\sum_{i=1}^{N}\xi_i, \\ \text{Subject to } \quad \gamma_i\left(\omega^{\mathrm{T}}\varphi(x_i) + b\right) \geq 1 - \xi_i, \quad i = 1, 2, \cdots, N, \\ \xi_i \geq 0, \quad i = 1, 2, \cdots, N, \end{cases} \tag{2}$$

where $\xi_i$ is the nonnegative relaxation variable, $\sum_{i=1}^{N}\xi_i$ is the tolerable classification error rate, and $\gamma$is the regularization parameter. Finding the optimal hyperplane is a quadratic programming problem with high computational complexity, so it is solved based on the least square principle. In the actual credit classification problem, each sample data cannot be completely and clearly divided into a specific class, so the fuzzy membership degree $\mu_i$ is introduced, and the model becomes

$$\begin{cases} \text{Minimize } J(\omega, b\,; \xi_i) = \dfrac{1}{2}\|\omega\|_2^2 + \gamma\displaystyle\sum_{i=1}^{N}\mu_i\xi_i^2, \\ \text{Subject to } \quad \gamma_i\left(\omega^{\mathrm{T}}\varphi(x_i) + b\right) \geq 1 - \xi_i, \quad i = 1, 2, \cdots, N, \\ \xi_i \geq 0, \quad i = 1, 2, \cdots, N. \end{cases} \tag{3}$$

Specifies $\alpha_{ki}$for each element of the training data, and the

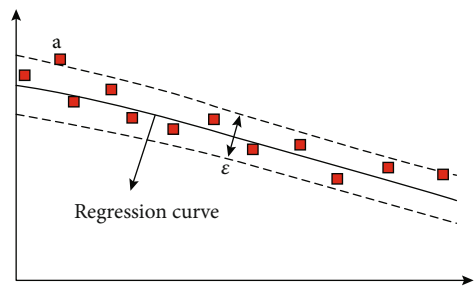



Figure 1: Schematic diagram of optimal partition hyperplane.

corresponding Lagrange function is

$$L_1(\omega, b, \xi, \alpha) = J_1(\omega, b, \xi) - \sum_{i=1}^{M}\sum_{k=1}^{N}\alpha_{ki}\{y_k[\omega(i)\varphi(x_{ki}) + b] - 1 + \xi_k\}. \tag{4}$$

By substituting the necessary conditions of KKT point into the above equation [11], we can obtain

$$y_k\left[\sum_{j=1}^{M}\sum_{i=1n}^{N}\alpha_{ji}y_j\varphi\left(x_{ki}, x_{ji}\right) + b\right] + \frac{\sum_{i=1}^{M}\alpha_{ki}}{\gamma\mu_k}k = 1, 2, \cdots, N. \tag{5}$$

The above equation can be written as

$$\begin{bmatrix} 0 & Y^{\mathrm{T}} \\ Y & Q \end{bmatrix}\begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ \vec{1} \end{bmatrix}, \tag{6}$$

where 0 is a scalar, $Y = [y_1, y_2 \quad ,\cdots, y_N]^{\mathrm{T}}, \vec{1} = [1, \cdots, 1]^{\mathrm{T}}$ $,\alpha = [\alpha_{11}, \cdots, \alpha_{N1}, \alpha_{12}\cdots, \alpha_{N2}, \cdots, \alpha_{1M}, \cdots, \alpha_{NM}]^{\mathrm{T}}$ $,Q = \left[y_1y_1\varphi(x_{1i}, x_{1i}) + (\gamma\mu_1)^{-1} \quad \cdots \quad y_1y_N\varphi \quad (x_{1i}, x_{Ni}) \quad \vdots \quad y_Ny_1\varphi(x_{Ni}, x_{1i})\cdots y_Ny_N\varphi(x_{Ni}, x_{Ni}) + (\gamma\mu_N) - 1\right], i = 1, 2, \cdots$ $, N$.

The classifier expression is

$$f\left(x_j\right) = \operatorname{sign}\left[\sum_{i=1}^{M}\sum_{k=1}^{N}\alpha_{ki}y_k\varphi\left(x_{ki}, x_{ji}\right) + b\right]. \tag{7}$$

Finally, the model becomes

$$\begin{gathered} \min \|\beta\|_1 \\ \text{s.t.} \begin{bmatrix} 0 & Y^{\mathrm{T}} \\ Y & Q \end{bmatrix} \quad \beta = \begin{bmatrix} 0 \\ \vec{1} \end{bmatrix}. \end{gathered} \tag{8}$$

The radial basis kernel (RBF) function was used for nonlinear mapping:

$$K\left(x_i, x_j\right) = \exp\left(\frac{-\left\|x_i - x_j\right\|^2}{\sigma^2}\right). \tag{9}$$

For membership$\mu_i$, select the formula:

$$\mu_k = \frac{s_k - \min\limits_{k=1,\cdots,N} s_k}{\max\limits_{k=1,\cdots,N} s_k - \min\limits_{k=1,\cdots,N} s_k}, \tag{10}$$

where $s_k$ is the preliminary credit score of each input sample obtained by the SVM credit scoring method.

SVM has a strong theoretical basis, which is an important aspect of its superior to other algorithms. It is the support of these theories that SVM overcomes some shortcomings that neural networks cannot overcome [12, 13]. The classification algorithm, regression algorithm, and multiclassification algorithm of SVM are the basis of the algorithm proposed. As can be seen from the above introduction, when SVM deal with multiclassification problems, they are mostly based on binary classifiers and combine multiple binary classifiers to construct multiclass classifiers. In other words, decomposition and reconstruction are regarded as the standard forms of multiclass classification problems, and one-to-one and one-to-many methods are adopted in typical algorithms. It is found that the common problem when using combinatorial binary classifier to solve multiclass problems is that each learner only considers the sample data of two classes. Therefore, neglecting the information of the remaining classes will inevitably lead to the weakening of the relationship between the data and then affect the prediction accuracy of multiclass problems.

*2.2. Basic Principle of Fuzzy Integral.* One shortcoming of SVM algorithm is that it only discriminates once during classification. Meanwhile, the maximum classification algorithm does not consider the importance degree of each output layer, but treats them equally, which will inevitably affect the accuracy of classification. In order to make up for this shortcoming, this paper applies fuzzy mathematics to credit risk assessment based on SVM classification [14]. The fuzzy integral is a nonlinear function, specifically defined as suppose $(X, \Omega)$ is a measure space, $h : X \longrightarrow [0, 1]$ is a measure function, then, the fuzzy integral on the fuzzy measure $g$ about $A(A \subseteq X)$ is defined as

$$\int_X h(x) \circ g(\cdot) = \sup_{E \subseteq X} \left[\min_{X \subseteq E}(\min h(x), g(A \cap E)\right] = \sup_{a \subseteq [0,1]} \left[\min_{X \subseteq E}(\min (a, g(A \cap F_a))\right], \tag{11}$$

where $F_a = \{x : h(x) \geq a\}$. Assuming that the information source set is $X$, let $h : X \longrightarrow [0, 1]$ be the evaluation of the source information$x \in X$, and $g = (\{X\})$ represents the importance of the information source when making object decisions. If you use information sources $A \subseteq X$ to evaluate objects. Consider

$$W(A) = \min_{x \in A} h(x), \tag{12}$$

is the safest decision made and expresses the importance of the information source to this decision. The minimum value is obtained by comparing these two values, thus indicating the consistency between the actual output and the expected value [15]. Therefore, fuzzy integration can describe the maximum possible agreement between object evidence and expected value.

The basic idea of this paper is to select some samples randomly from the original training sample set that has been correctly classified objectively and compose several sample sets. Then, SVM algorithm based on fuzzy integral is used for reclassification, so as to judge the accuracy of the classification method [16]. Credit risk assessment model based on fuzzy integral and SVM regression integration is shown in Figure 2. The specific algorithm of SVM based on fuzzy integral is as follows:

Step 1. For input patterns $x$, each sub-SVM evaluates and outputs membership of $x$ relative to each category.

Step 2. For each category $C_m$, each sub-SVM computes $h_m(\mu_k)$ and $g_h(\mu_k)$ and calculates the fuzzy integral $FI_m$ with respect to the category $C_m$.

Step 3. Make a judgment decision about the class to which the pattern belongs.

## 3. Research on Modeling Methods of Credit Risk Assessment

The possibility that the bank will suffer losses due to the failure of the bank's loan enterprise customers. As the most important risk of banks, credit risk has the following main characteristics: sociality, expansion, periodicity, controllability, and so on [17]. As credit risk comes into being with bank credit business, its existence is objective and cannot be eliminated, which requires banks to carry out scientific and effective management of credit risk. Credit risk management is a system of comprehensive management process and includes credit risk identification, assessment, and control process, such as credit risk assessment is the foundation of credit risk management and the core, the results of the assessment work fit and unfit quality directly affect the subsequent risk management effectively, which deeply affects the sustainable development of the bank. Because credit risk is a nonlinear complex system, modern credit risk assessment needs advanced computer science and technology and mathematical modeling knowledge to deal with the value of many factors interacting with each other [18–20]. It is a research field that requires cross-professional cooperation and continuous development and innovation.

One of the most significant characteristics of credit risk is nonsystematization, that is, it has obvious individual attributes. Unlike other financial risks such as market risk, credit risk is ultimately determined by the payer himself. Although the repayment ability is affected by systemic risks, such as macro policies and economic cycles, strictly speaking, the nonsystematic factors that are individual characteristics of the debtor, such as repayment willingness, operating ability, financial status, industry attributes, and risk preference, are the decisive factors for the borrower's performance. In the establishment of credit risk assessment index system, we should first follow the general principles of index system construction: index content closely linked and strong

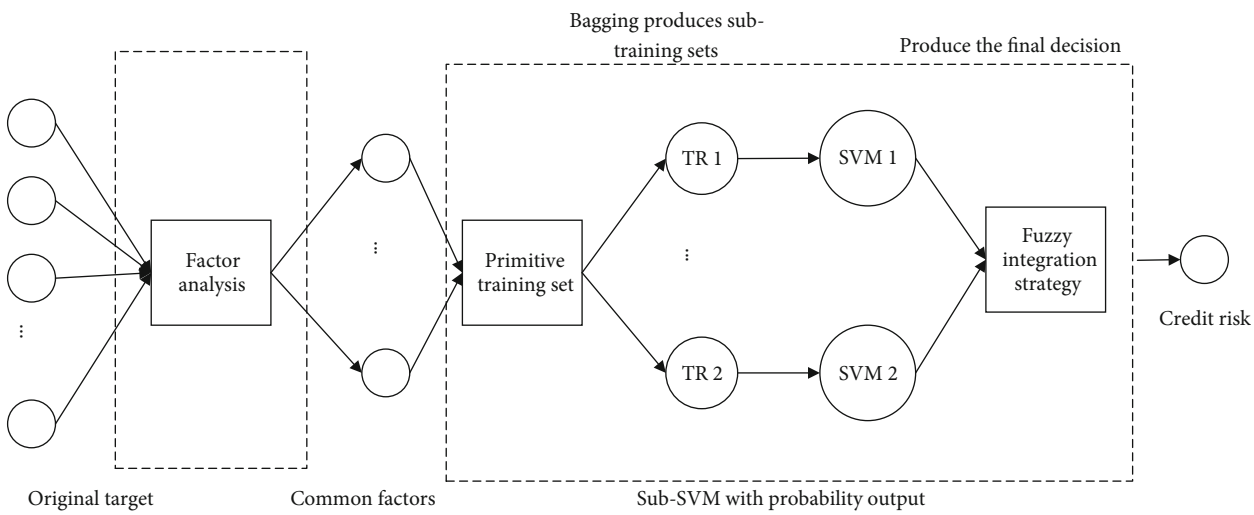



Figure 2: Credit risk assessment model based on fuzzy integral and SVM regression integration.

operability. The index calculation has reliable data sources and is conducive to the use of computer operation. Second, in the selection of indicators, we should also follow the scientific, comprehensive, and hierarchical, feasible, quantifiable four principles, and the selection of bank credit risk assessment indicators mainly focus on the study of enterprise performance ability, that is, the selection of enterprise financial indicators [21]. Considering the important influence of loan method on credit risk, the index system of this paper is also constructed based on four basic factors: solvency, profitability, operating capacity, and loan method.

(1) *Solvency*. Solvency refers to the ability of a business to repay maturing debts. An important factor to measure the security of bank credit funds. Industrial and commercial banks of internal evaluation indicators of solvency including current ratio, quick ratio, operating activities cash inflow/advocate business income and operating cash total debt ratio, the main business revenue profit growth and rate of return on total assets, net assets and the multiple of interest safeguard, the growth rate of net assets, total liabilities/debit, and other long-term solvency evaluation indicators. In view of the limitation of data sources, this paper focuses on extracting financial indicators from short-term debt paying ability [22, 23]. Based on the principle of index selection, the following indicators with highly accessible data and high reliability are selected

(2) *Operating Capacity*. Operating capacity is the ability of enterprises to make profits by using various assets, which is ultimately related to the efficiency of capital utilization and solvency of loan enterprises [24]. Indicators reflecting the operating capacity of an enterprise include accounts receivable turnover, current assets turnover, total assets turnover, and inventory turnover. The following indicators with high importance and strong data reliability are selected in this paper

(3) *Profitability*. Profitability is the ability of enterprises to obtain profits, the level of profitability is related to whether the credit funds can be recovered as scheduled. Banks want to lend to profitable businesses because the more profitable a business is, the more likely it is that credit will be repaid on time [25]. Reflect the corporate profitability indicators including operating margins, cost, profit margin, surplus cash cover, return on total assets, return on equity, and return on capital to six, considering the availability and reliability of the data, and the consideration of whether enterprises have long-term profitability, this paper selects several indexes

(4) *Loan Method*. Although the loan method is not the financial index of enterprises, it has been paid more and more attention by experts and scholars in the field of credit risk research [26]. Loan methods are generally divided into credit, guarantee, mortgage, and pledge, and the bank will set the corresponding risk coefficient for each loan method

(5) In this paper, fuzzy integral integrated SVM regression model is constructed. According to the principles of comprehensiveness, conciseness, scientificity, operability, economy, representativeness, and standardization, an evaluation system consisting of 26 indicators was constructed. Factor analysis was first used to reduce the number of input indicators of SVM, and then the credit risk was output through SVM regression model training [27]. Factor analysis and SVM regression integration are combined in series. The credit risk evaluation model is shown in Figure 3

There is a certain difference between the prediction model of enterprises without credit risk and those with credit risk, and the prediction accuracy of enterprises without credit risk is generally higher than that of enterprises with credit risk. The credit evaluation discussed in this paper is a SVM which divides the credit risk of commercial banks

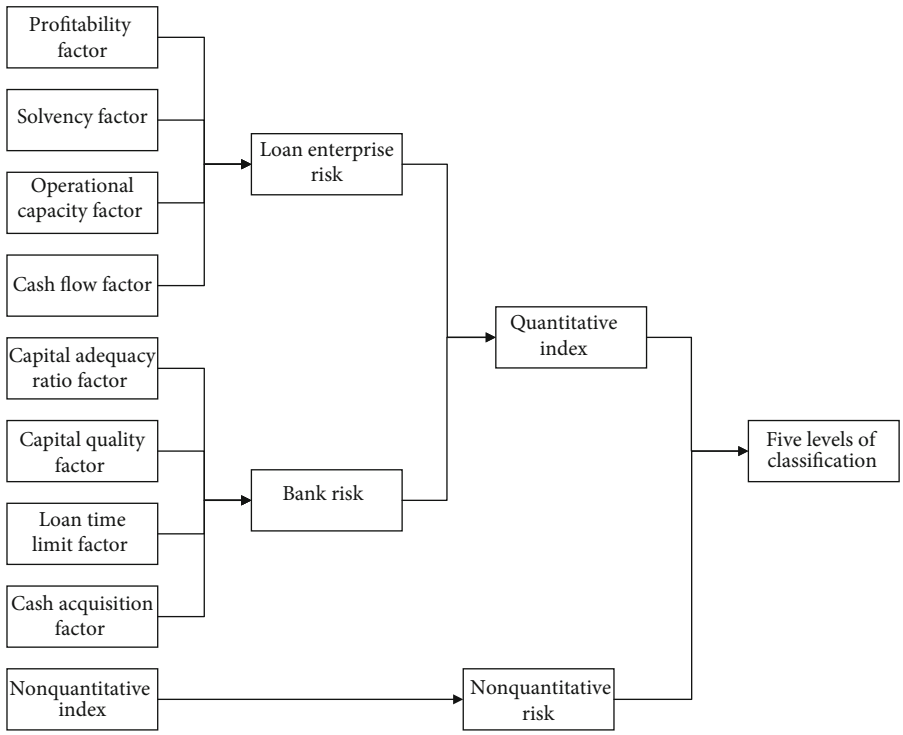



Figure 3: Credit risk evaluation model.

into five types and belongs to multiclassification problem [28]. Given the training set:

$$T = \{(x_1, y_1), \cdots, (x_i, y_i)\} \in (X \times Y)^l, \tag{13}$$

where $x_i \in X = R^n, x_i \in Y = \{1, 2, \cdots, M\}, i = 1, 2, \cdots, l$, finding a decision function $f(x)$: $X = R^n \longrightarrow Y$ to solve the multiclassification problem is actually finding a rule that $M$ can divide the points on $R^n$ into parts.

The usual method of obtaining multiclass classifiers is to construct a series of two-class classifiers, each of which divides one of them from the following classes. The attribution of an input can then be inferred. The "one to many" method is to construct a SVM subclassifier for class problems [29]. This paper studies the problem of accurate evaluation of bank credit risk. Due to the uncertainty of bank credit capital security, there are many evaluation indexes of credit risk, there is a large amount of repeated information among the indexes, and the nonlinear relationship between risk grade and indexes, it is difficult to accurately evaluate the traditional evaluation model, and the evaluation accuracy is not high. Credit risk evaluation model using data mining techniques and statistical analysis method, behavior and credit characteristics, capture the history information and future credit, the relationship between build predictive models that use a credit score to comprehensively evaluate the future credit performance of a consumer or business. Under the condition that the scientific application method of the development process is correct, the credit evaluation model can provide credit managers with a large amount of highly predictive information, help them to formulate effective management strategies, and effectively develop the market, control risks, excavate profits with high accuracy, and achieve high efficiency of credit business [30].

## 4. Analysis and Discussion of Calculation Results

The data in this paper comes from the data of short-term loan enterprises of a bank. The loan issuance date is from March 1, 2020, to March 31, 2021. If the loan is overdue for more than 3 years, it is considered as bad debts. This paper retrieves the loan balance and total loan amount of the sample enterprise on September 22, 2016, to measure the degree of default of the sample enterprise. The financial index, enterprise characteristic index, and enterprise owner characteristic index of the sample enterprises are collected in January of a certain year, which is used as the basis for the credit risk evaluation of the bank when issuing loans. After collection and sorting, this paper obtained 146 samples, involving nearly 600 million yuan of loans. In this paper, two times and three times standard deviation tests were used to eliminate 28 abnormal data (19%) and obtain 118 sample data (81%). In order to eliminate the influence of excessively large index values on the results, this paper adopts the maximum and minimum normalized functions to normalize the sample data, and all index values are normalized to [0, 1]. By random sampling method, 98 (83%) training samples were selected for SVM integration model construction, and 20 (17%) test samples were tested for model generalization ability. The cluster analysis of credit risk evaluation index is shown in Figure 4.

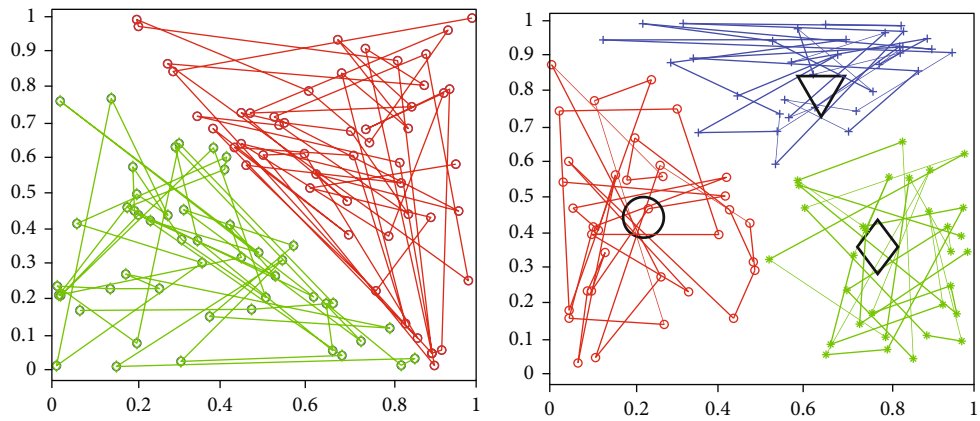



Figure 4: Cluster analysis of credit risk evaluation index.

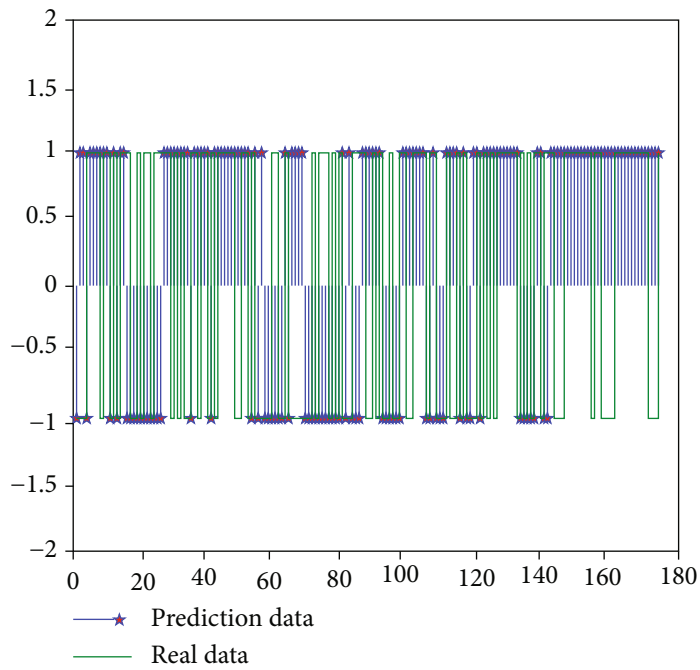



Figure 5: Comparison of predicted and real values.

The prediction accuracy of fuzzy integral SVM ensemble is the highest, which may be because the fuzzy integral ensemble method considers the importance of each subtrainer when performing multiclassifier fusion. However, the fuzzy integral BP neural network integration method has the lowest prediction accuracy, which may be because BP neural network needs to be simulated under large samples to ensure its prediction accuracy. Random sampling under limited samples forms subtraining sample set, which further reduces the training sample size of the subtrainer, but reduces its prediction accuracy. The superiority of small sample training of SVM is further demonstrated. Under the current limited sample of corporate loans, using the fuzzy integral SVM integration method, banks can more accurately assess the risk of corporate loans and timely carry out risk monitoring. The simulation results of test samples show that the model has strong generalization ability. The comparison of predicted and real values is shown in Figure 5. The total precision and the second precision of fuzzy least squares SVM model are higher than other models. The accuracy of the first type is second only to that of the RBF network model and much higher than the average accuracy of other models. This shows that the fuzzy least squares SVM model has a good effect on credit risk assessment.

The five groups of moving averages of different time lengths are shown in Figure 6. In the figure, the thin line represents the change curve of accuracy of logistic regression on the training sample set and test sample set, the dotted line represents the change curve of accuracy of BP neural network on the training sample set and test sample set, and the thick line represents the change curve of accuracy of SVR on the training sample set and test sample set. As can be seen from the figure, the change curve of the logistic

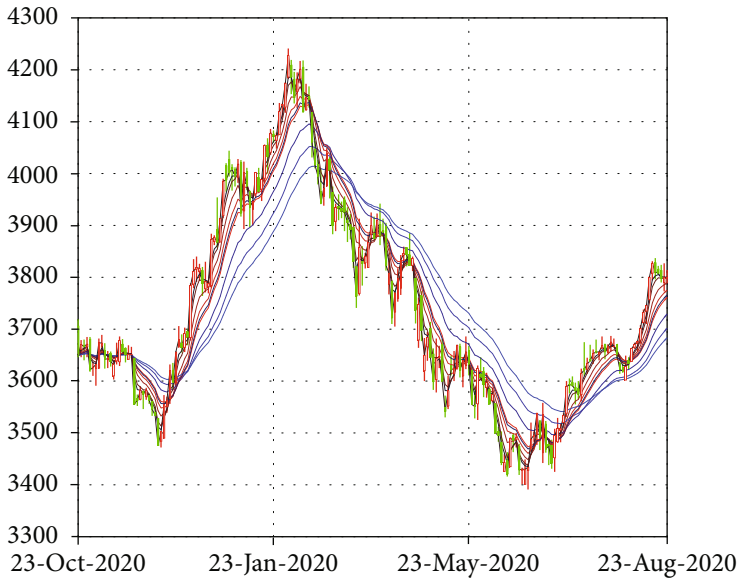



Figure 6: Five groups of moving averages of different time lengths.

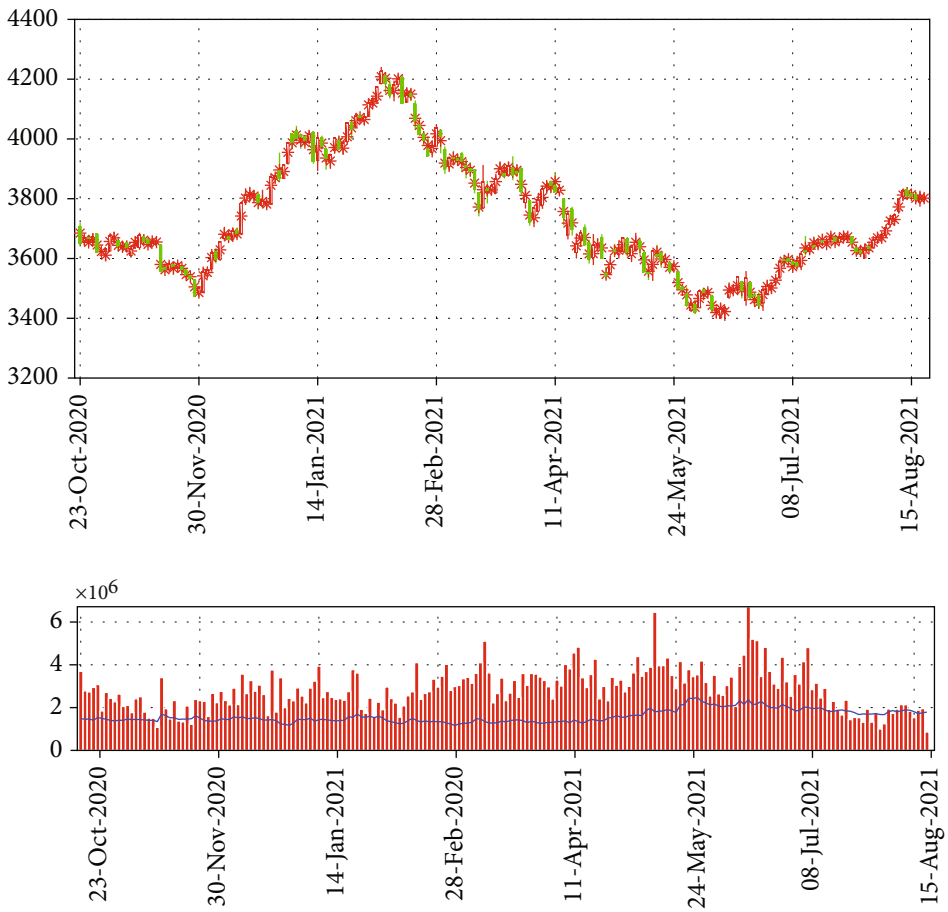



Figure 7: Prediction data and error results.

model did not change much when approaching a straight line, but the accuracy rate was the lowest among the three models. BP neural network has the highest slope of change curve and has the highest accuracy for training samples. However, it has the worst generalization ability because of its low classification accuracy for predicted samples. It does not have the highest accuracy for training samples. But have a rhinoceros high classification accuracy for predicting

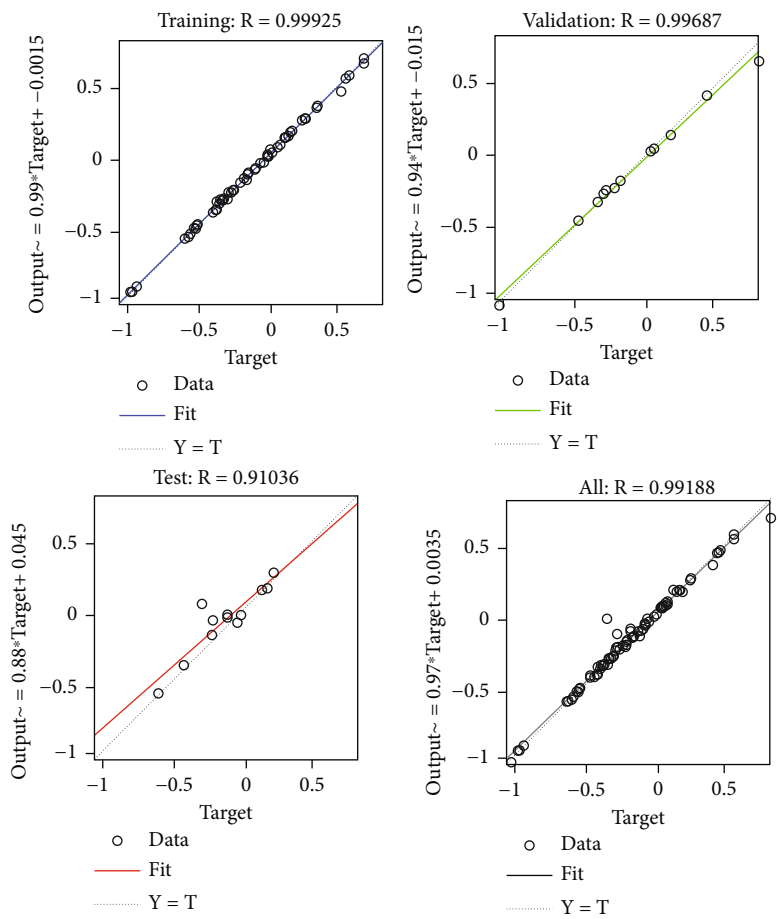
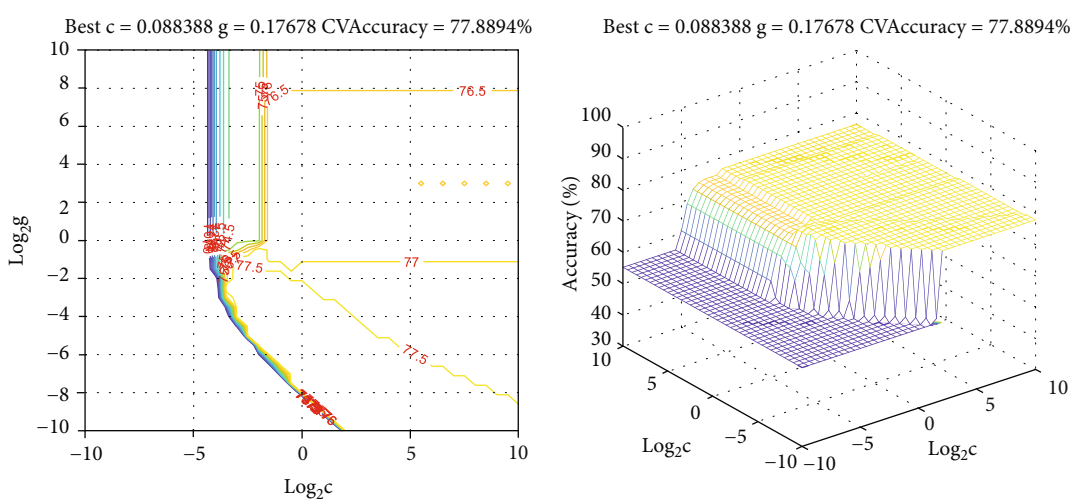



Figure 8: Fuzzy integral SVM training regression state graph.




(a) SVM parameter optimization contour graph

(b) SVM parameter optimization 3D graph

Figure 9: SVM parameter selection result diagram.

sample, further calculation can be obtained, the accuracy of neural network training samples to test sample was reduced by 7%, and the accuracy of SVR from training samples to test sample was reduced by 2%, only demonstrates SVR based on structural risk minimization theory than based on the empirical risk minimum of BP neural network has better generalization ability, this is exactly what the bank credit risk assessment model needs.

The prediction data and error results are shown in Figure 7. It can be seen from the figure that the fuzzy dynamic SVM integration method has the best effect in both training and final prediction. Among them, fuzzy dynamic

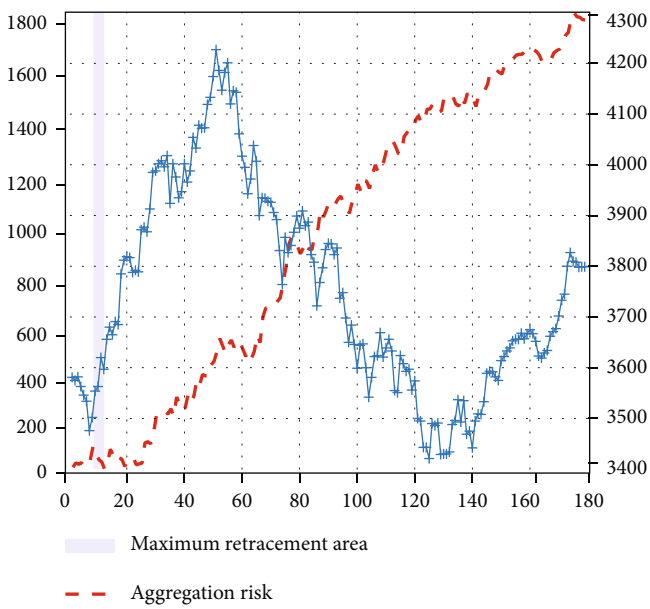



Figure 10: Credit risk trend tracking curve.

SVM integration method > single SVM > BPNN > BPNN integration. Among them, the accuracy of neural network algorithm is significantly lower, which may be mainly due to the limited basic data set studied in this paper, but the neural network method is based on a large sample size. It is confirmed that SVM has small sample prediction as described in other literatures. And in addition, this article is to take more child classified integral integrated way, each individual as a whole system, and considering weight of each to integration, it also improved the degree of accuracy, and reduce the learning time of the whole model, improved the future application in the actual of timeliness, and reduced the time cost algorithm.

The fuzzy integral SVM training regression state graph is shown in Figure 8. As we can see, in the data of different cases, a classification of three values under the condition of the forecasting accuracy is quite, and training time and predict time were significantly less than the average values of three classifier, this is because made simple regression machine classifier to reduce the number of parameters, reduce the occupancy of memory, and improve the efficiency. Its prediction accuracy is nearly 3% higher than that of the combined binary classifier, which is due to the reduction of information loss between classes and improvement of classification accuracy by using regression machine: the training time is about 31% lower than that of the combined binary classifier. The prediction time is reduced by 70% than that of the combined binary classifier, which is obviously smaller than that of the combined binary classifier, because the scale of optimization problem is reduced by the simple classifier of regression machine. The efficiency of the algorithm is improved. It is verified that the multiclassifier construction algorithm based on SVM can provide more accurate judgment basis for banks to predict enterprise credit risk quickly.

Bagging was performed with 50 samples randomly selected from the original training set. Three sub-SVMs were obtained from the three training sets generated by bagging, and they were integrated using fuzzy integral. Radial basis kernel function was used for each sub-SVM, and the corresponding parameters were obtained by 5-fold cross test. Through 10 experiments, the mean classification accuracy of fuzzy integral SVM integration in training sample set and test sample set is shown. Considering the classification of each subclassifier and the importance of each subclassifier relative to the final decision, fuzzy integral SVM integration is better than single SVM, maximum voting SVM integration, and fuzzy neural network. SVM parameter selection result diagram is shown in Figure 9. As can be seen from the figure, we also see the same trend (the error rate of category 1 is significantly lower than that of category 2). We do not have a mechanistic explanation yet, but that is what we are hoping for. If modeling is compared with other new data, the error rate of class 1 of SVM model may be larger. In order to effectively reduce the class 1 error rate, we can impose a large penalty coefficient on the class 1 error when constructing the SVM model, but this may also reduce the overall accuracy of the model.

Fuzzy integral SVMR integration has the highest prediction accuracy, which may be because the fuzzy integral integration method considers the importance of each subtrainer in the fusion of multiple classifiers. However, the fuzzy integral BP neural network integration method has the lowest prediction accuracy, which may be because BP neural network needs to be simulated under large samples to ensure its prediction accuracy. Random sampling under limited samples forms subtraining sample set, which further reduces the training sample size of the subtrainer, but reduces its prediction accuracy. The superiority of small sample training of SVM is further demonstrated. Under the current limited sample of loans to small and micro enterprises, the integrated method of fuzzy integral SVM can be adopted by banks to more accurately assess the risk of loans to small and microenterprises and timely monitor the risk. The simulation results of test samples show that the model has strong generalization ability. The credit risk trend tracking curve is shown in Figure 10.

## 5. Conclusion

On the basis of comprehensive consideration of enterprise characteristics, enterprise financial indicators, enterprise owner characteristics, and loan methods, this paper takes into account the evaluation index system of data availability and uses factor analysis method to extract common factors to reduce the index system. Based on fuzzy integral SVM regression integration method, enterprise credit risk evaluation is carried out. Compared with the methods based on voting integrated SVM, single SVM, and fuzzy neural network, the results show that the classification accuracy of integrated SVM based on fuzzy integral is obviously better than other methods.

(1) Traditional credit risk assessment methods only evaluate the credit risks faced by commercial banks from the perspective of loan enterprises, but ignore the impact of bank's own deposit and loan structure

and risk status on credit risks, resulting in the absence of evaluation subjects. Based on the analysis of the causes of credit risk of commercial banks in China and the implementation of Basel New Capital Accord, this paper analyzes credit risk factors from three aspects: loan enterprise risk factors, bank risk factors, and Basel New Capital Accord risk factors

(2) Established the evaluation index system of commercial bank credit risk from two aspects of qualitative and quantitative analyses in detail the credit risk, finally determine the consists of 23 indicators of credit risk assessment model, which will cash flow analysis for the first time as an added to the index system of influencing factors, and according to the statement of cash flows is increasingly important trends, increase the weight, in order to make the index system more suitable for the reality

(3) Fuzzy integral is proposed to support vector machine integration, which partly solves the problem of parameter selection when using support vector machine to evaluate. The traditional support vector machine integration method cannot reasonably assign different weight to the subsupport vector machine. When fuzzy integral is used for support vector machine integration, it can assign reasonable weight according to the classification result of the subsupport vector machine, which greatly improves the classification effect after integration

(4) The classification results of integrated support vector machine based on fuzzy integral, traditional integrated support vector machine based on voting method, single support vector machine, and fuzzy neural network are compared and analyzed. The classification results of integrated support vector machine based on fuzzy integral are obviously better than the other three methods. It is feasible and effective to evaluate the credit risk of commercial bank with integrated support vector machine based on fuzzy integral

Fuzzy integral SVM and credit risk assessment are two extremely complex and incomplete research fields. Although this paper has made significant improvements and breakthroughs on the basis of existing research, there are still some defects and deficiencies that need further research. Possible research directions in the future include

(1) In the construction of the risk factor system, we should continue to explore the integration of more nonfinancial indicators and build a more complete and comprehensive factor system, which can gradually improve the accuracy of the evaluation

(2) In terms of prediction model, we can continue to consider the integration with other artificial intelligence methods. The combination of different methods may make the model more accurate and feasible

## Data Availability

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

## Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

[1] R. Spicas, R. Kanapickiene, M. Vijunas, and R. Kirka, "Development of enterprise credit risk assessment model for lithuanian credit unions," *IEEE Latin America Transactions*, vol. 17, no. 10, pp. 1733–1740, 2019.

[2] F. Assef, M. T. Steiner, P. J. S. Neto, and D. G. D. Franco, "Classification algorithms in financial application: credit risk analysis on legal entities [J]," *Transformations in Business & Economics*, vol. 17, no. 1, pp. 152–177, 2018.

[3] E. Brons-Piche, G. J. Eckert, and M. Fontana, "Predictive validity of a caries risk assessment model at a dental school," *Journal of Dental Education*, vol. 83, no. 2, pp. 144–150, 2019.

[4] Z. Chaohui, R. Hui, and L. Jiajia, "Research on E-commerce credit evaluation method based on Bayesian and neural network hybrid algorithm," *Information Science*, vol. 38, no. 2, pp. 81–87, 2020.

[5] F. S. Kong, "Development of metric method and framework model of integrated complexity evaluations of production process for ergonomics workstations," *International Journal of Production Research*, vol. 57, no. 8, pp. 2429–2445, 2019.

[6] U. P. Shukla and S. J. Nanda, "Designing of a risk assessment model for issuing credit card using parallel social spider algorithm," *Applied Artificial Intelligence*, vol. 33, no. 3, pp. 191–207, 2019.

[7] B. Song, W. Yan, and T. J. Zhang, "Cross-border e-commerce commodity risk assessment using text mining and fuzzy rule-based reasoning," *Advanced Engineering Informatics*, vol. 40, no. 5, pp. 69–80, 2019.

[8] P. Y. Song and Y. T. Liu, "An xgboost algorithm for predicting purchasing behaviour on e-commerce platforms," *Tehnicki Vjesnik-Technical Gazette*, vol. 27, no. 5, pp. 1467–1471, 2020.

[9] Q. Zhang, J. Wang, A. Lu, S. Wang, and J. Ma, "An improved SMO algorithm for financial credit risk assessment - evidence from China's banking," *Neurocomputing*, vol. 272, no. 8, pp. 314–325, 2018.

[10] X. Ji, J. Wang, Y. Li, Q. Sun, and C. Xu, "Modulation recognition in maritime multipath channels: a blind equalization-aided deep learning approach," *China Communications*, vol. 17, no. 3, pp. 12–25, 2020.

[11] C. Huang, Q. Yang, M. Du, and D. Yang, "Financial distress prediction using SVM ensemble based on earnings manipulation and fuzzy integral," *Intelligent Data Analysis*, vol. 21, no. 3, pp. 617–636, 2017.

[12] N. Bouadjenek, H. Nemmour, and Y. Chibani, "Fuzzy integrals for combining multiple SVM and histogram features for writer's gender prediction," *IET Biometrics*, vol. 6, no. 6, pp. 429–437, 2017.

[13] N. V. Naik, S. P. Singh, and A. K. Panda, "An interval type-2 fuzzy-based DTC of IMD using hybrid duty ratio control,"

*IEEE Transactions on Power Electronics*, vol. 35, no. 8, pp. 8443–8451, 2020.

[14] H. Chaoui, M. Khayamy, and A. A. Aljarboua, "Adaptive interval type-2 fuzzy logic control for PMSM drives with a modified reference frame," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 5, pp. 3786–3797, 2017.

[15] Q. Liu, Y. Shen, L. Wu, J. Li, L. Zhuang, and S. Wang, "A hybrid fcw-emd and kf-ba-svm based model for short-term load forecasting," *CSEE Journal of Power and Energy Systems*, vol. 4, no. 2, pp. 226–237, 2018.

[16] F. Piltan and J. M. Kim, "Bearing fault identification using machine learning and adaptive cascade fault observer," *Applied Sciences-Basel*, vol. 10, no. 17, p. 5827, 2020.

[17] A. Mollajan, H. Memarian, and M. Nabi-Bidhendi, "Fuzzy classifier fusion: an application to reservoir facies identification," *Neural Computing & Applications*, vol. 30, no. 3, pp. 825–834, 2018.

[18] W. Deng, J. Xu, X. Z. Gao, and H. Zhao, "An enhanced MSIQDE algorithm with novel multiple strategies for global optimization problems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 99, pp. 1–10, 2020.

[19] W. Deng, H. Liu, J. Xu, H. Zhao, and Y. Song, "An improved quantum-inspired differential evolution algorithm for deep belief network," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 10, pp. 7319–7327, 2020.

[20] W. Deng, J. Xu, Y. Song, and H. Zhao, "An effective improved co-evolution ant colony optimisation algorithm with multi-strategies and its application," *International Journal of Bio-Inspired Computation*, vol. 16, no. 3, pp. 158–170, 2020.

[21] G. Mingliang and L. Jun, "Support vector machine based on an improved particle swarm optimization algorithm for SEMG signal pattern recognition," *Chinese Journal of Sensors and Actuators*, vol. 20, no. 3, pp. 1459–1464, 2017.

[22] V. Veerasamy, N. I. Abdul Wahab, R. Ramachandran et al., "High-impedance fault detection in medium-voltage distribution network using computational intelligence-based classifiers," *Neural Computing & Applications*, vol. 31, no. 12, pp. 9127–9143, 2019.

[23] T. S. Abdelgayed, W. G. Morsi, and T. S. Sidhu, "A new harmony search approach for optimal wavelets applied to fault classification," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 521–529, 2018.

[24] N. R. Babu and B. J. Mohan, "Fault classification in power systems using EMD and SVM," *Ain Shams Engineering Journal*, vol. 8, no. 2, pp. 1031–1111, 2017.

[25] F. Ben Abid, S. Zgarni, and A. Braham, "Distinct bearing faults detection in induction motor by a hybrid optimized SWPT and aiNet-DAG SVM," *IEEE Transactions on Energy Conversion*, vol. 33, no. 4, pp. 1692–1699, 2018.

[26] K. Guo, R. Chai, H. Candra et al., "A hybrid fuzzy cognitive map/support vector machine approach for eeg-based emotion classification using compressed sensing," *International Journal of Fuzzy Systems*, vol. 21, no. 1, pp. 263–273, 2019.

[27] I. Aydin, "Fuzzy integral and cuckoo search based classifier fusion for human action recognition," *Advances In Electrical and Computer Engineering*, vol. 18, no. 1, pp. 3–10, 2018.

[28] S. Kim, Y. Seo, M. Rezaie-Balf, O. Kisi, M. A. Ghorbani, and V. P. Singh, "Evaluation of daily solar radiation flux using soft computing approaches based on different meteorological information: peninsula vs continent," *Theoretical and Applied Climatology*, vol. 137, no. 1-2, pp. 693–712, 2019.

[29] J. Huang, B. Jiang, C. Xu, and N. Wang, "Slipping detection of electric locomotive based on empirical wavelet transform, fuzzy entropy algorithm and support vector machine," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 8, pp. 7558–7570, 2021.

[30] K. Guo, R. Chai, H. Candra et al., "Dual-tree complex wavelet transform-based features for automated alcoholism identification," *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, vol. 20, no. 4, pp. 1297–1308, 2018.