

Research Article

Application of Data Mining Combined with Fuzzy Multicriteria Decision-Making in Credit Risk Assessment from Legal Service Companies

Jianmiao Hu,¹ Chong Chen ,² and Kongze Zhu³

¹School of Transportation Management, Zhejiang Institute of Communications, Hangzhou, Zhejiang, China

²Zhejiang University of Technology, Hangzhou, Zhejiang, China

³Institute for Advanced Study in the Humanities and Social Sciences, Beihang University, Beijing, China

Correspondence should be addressed to Chong Chen; gdcc@zjut.edu.cn

Received 6 August 2021; Revised 18 September 2021; Accepted 25 September 2021; Published 20 November 2021

Academic Editor: Zaoli Yang

Copyright © 2021 Jianmiao Hu et al. 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 purpose is to avert the systematic financial risks from the Internet financial bubble and improve the efficiency of legal service companies' credit risk assessment ability. Firstly, this study analyzes the commonly used classification model, Support Vector Machine (SVM), and linear regression model, Logistic model, and then puts forward the integrated SVM-Logistic + Fuzzy Multicriteria Decision-Making (FMCDM) to evaluate and analyze the credit risk level of listed companies. In the proposed integrated model, the SVM model classifies the data sample from listed companies, and the Logistic model is used for regression analysis on the credit risk assessment. Based on the credit risk indexes and weight uncertain factors of sample companies, FMCDM based on fuzzy set is applied to obtain the evaluation indexes. Then, the Analytic Hierarchy Process (AHP) is used to obtain the weight of key indexes. Finally, the fit analysis is carried out according to the existing risk status of the sample company and the risk status results of the proposed integrated model. The results show that the integrated SVM-Logistic model is complementary and has high intensive evaluation. According to the fitness value obtained by FMCDM, the company's credit risk status can be accurately evaluated, and the intermediate threshold of corporate credit default risk measurement is 0.56152; if Fit is lower than the threshold, the company's credit is low, and if Fit is higher than the threshold, the company's credit is high. Therefore, the data mining technology based on integrated SVM-Logistic model + FMCDM has high precision and feasible application in the credit risk assessment from legal service companies. This study creates a new method model for legal service companies in the field of corporate credit risk assessment and can provide references and ideas for corporate credit risk assessment.

1. Introduction

Today, the global economy is having an unprecedented tough time due to the ongoing coronavirus disease 2019 (COVID-19) pandemic. China's economic boom is also witnessing a spiking credit crisis yearly; in particular, the growth rate of nonperforming loans remains at a stubbornly high level influenced by various factors, such as the global economic slowdown, the productivity upgradation, the emerging development pattern, the stocks shortage, the intensification of deleveraging, the tax-relief policies, the raising of lending level, and the development of Internet

finance. Under such background, legal cases on credit disputes have shot up, thereby posing both an opportunity and a challenge for legal service companies. On the one hand, legal service companies provide legal advice on credit risk transactions in normalized companies. On the other hand, they help credit companies solve multiple problems involving legal and economic disputes from credit risk. In particular, legal service companies need to innovate and improve the corporate credit risk assessment methods, levels, and the loss assessment on economic disputes legally. Currently, credit risk assessment has become the most difficult business for legal service companies. Mainly, the

research on credit risk assessment of legal service companies is from two aspects: the construction of index system and the improvement of the model.

2. Literature Review

In terms of the latest research, Lei and Zhao (2019) analyzed the correlation between Corporate Social Responsibility (CSR) and corporate economy theoretically, constructed CSR factors from the perspective of relevant stakeholders, and built an innovative Logistic model combined with financial factors by considering the time lag of economic consequences of CSR; based on the data of five strategic emerging industries of China's Shanghai and Shenzhen A-share listed companies from 2013 to 2017, this paper constructed a subindustry logistic regression model. The results showed that the Logistic model had high accuracy and could effectively improve the creditor's risk assessment ability on enterprises to avoid risks [1]. Du and Lu (2018) carried out a case analysis on "microloan network" based on the original credit risk evaluation indexes, international Finance Controlling (FICO) credit scoring methods, and domestic sesame credit scoring methods. Du and Lu (2018) tried to build a credit evaluation index system suitable for domestic P2P online loan platforms to evaluate the borrower's credit. The data of 6,917 borrowers on the "microloan network" platform website were selected, the SVM logistic combination model was adopted, the modified index system was used for credit risk assessment, the test results were compared with the actual results, and the credit risk assessment system was optimized [2]. Zheng and Li (2020) combined traditional financial indexes and added nonfinancial indexes to build a more comprehensive enterprise credit risk assessment index system. Based on sample data from A-share listed companies from 2015 to 2017 as samples, the PCA-SMOTE-GS-SVM model was implemented by combining Principal Component Analysis (PCA), Synthetic Minority Oversampling Technology (SMOTE), and Support Vector Machine (SVM) after parameter optimization through Grid Search (GS) method to evaluate enterprise credit risk. Compared with other models, the results showed that the proposed model had high stability and predictability [3]. Hu et al. (2017) provided a solution for the defects of traditional Backpropagation Neural Network (BPNN) in credit risk assessment application for Micro and Small Enterprises (MSE), such as slow network learning speed, easily falling into local solutions, and large operation result errors caused by random initial weights and thresholds. Based on the Glowworm Swarm Optimization (GSO) algorithm, an Improved Discrete GSO (IDGSO) algorithm is proposed. The simulation results indicated that the model has obvious advantages over the traditional BPNN model, Genetic Algorithm- (GA-) BP model, and continuous GSO-BP model in convergence speed and operation accuracy [4]. Zhang et al. (2018) constructed a Peer-to-Peer (P2P) online loan borrower credit risk assessment model based on Disequilibrium Fuzzy Proximal Support Vector Machine (DFPSVM), proposed the borrower's credit scoring and rating method, and

conducted an empirical analysis using the borrower's credit information on Renren loan platform. The results suggested that the constructed model had better adaptability and higher classification accuracy than other models; it could effectively reduce the impact of sample disequilibrium on classification results and significantly increase the classification accuracy of negative samples [5]. The financial market crash after the Lehman shock in 2008 has created an environment for banks to avoid risks. Such an environment is even more difficult for new small business loans. Since then, P2P lending has become more popular all over the world. However, most studies on the credit risk of P2P lending only consider the event of borrower default, not the amount of loss. From this perspective, Charkaborty et al. (2019) took the Net Rate of Return (NRR) as the standard for forecasting training with labeled data. Additionally, they trained regression models to assess credit risk. The proposed new model predicted the borrower's profit amount. Finally, by using their credit risk assessment model, P2P loan investors could more accurately measure the risk, and the proposed model could also predict the profit amount of the loan [6]. In each organization, performing accurate risk assessment and considering increased accidents is a necessary tool to prevent and reduce the fatal and nonfatal consequences of accidents. One of the most popular risk assessment methods is Failure Mode and Effect Analysis (FMEA), which evaluates the failure modes in the system by using Risk Priority Progression (RPN). These methods have been criticized for several defects, including the influence of personal opinions, the equal importance of factors, and risk rating. Mangeli et al. (2019) used a hybrid method based on SVM and fuzzy inference system to reduce the impact of personal opinions on determining severity and occurrence factors. Meanwhile, logarithmic fuzzy preference programming was used to determine the clear weight of FMEA-dependent factors, and the modified Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) was used to rank risks more accurately. The proposed model could be used to assess the security risks of all organizations. The applicability of the proposed model was verified through demonstration at the copper leaching plant in Kerman, Iran. The results proved that the model could predict the severity and incidence of occupational accidents in 5 years (2012–2017), and the accuracy rates were 87% and 95%, respectively [7]. Through a comparative analysis of the above studies, the following deficiencies are found. Although the Logistic linear model is slightly inferior to the Radial Basis Function (RBF) model in identifying potential defaulters, the accuracy of the Logistic linear model is better than that of the RBF model as a whole. SVM is more effective and superior to BPNN in small samples. However, the simple SVM model has errors in dealing with unbalanced samples.

Given the above analysis, a vector machine data mining technology based on SVM is proposed to classify the credit risk data of sample companies, and the classified variables are simulated again by the Logistic model to continue the credit evaluation of the company. The data mining technology of SVM is widely adopted in model classification and regression. It can classify the corporate subjects of

nonperforming loans and excellent loans in the credit risk assessment of legal service companies [8]. Based on the uncertain factors of the corporate credit risk index and weight, the Fuzzy Multicriteria Decision-Making (FMCDM) based on fuzzy set is applied to obtain fuzzy indexes of credit risk. Then, the Analytic Hierarchy Process (AHP) is adopted to obtain the index weight to evaluate the corporate credit risk. The results show that the integrated SVM-Logistic model combined with the FMCDM has high accuracy and stability. The corporate credit risk status can be accurately evaluated according to the fit value obtained by the proposed model.

3. Materials and Methods

3.1. Analysis of SVM Data Mining Technology. Data mining can be used to excavate useful information from big data and help decision-makers make the decision. Data mining involves supervised learning, unsupervised learning, and association rule mining [9]. This study adopts one of the common data mining technologies, SVM. SVM can transform low-dimensional samples into high-dimensional samples and separate samples, which is a new data mining model [10]. Figure 1 presents the principle diagram of SVM technology.

The calculation process of the SVM model is as follows. Training samples have linear separable characteristics:

$$(X_1, Y_1), \dots, (X_l, Y_l), \quad X_i \in R^d, Y_i \in \{-1, +1\}, i = 1, \dots, l. \quad (1)$$

There must be an interval plane $\omega X + b = 0$ that separates the samples.

$$\begin{aligned} \omega X_1 + b < 0, Y_1 = -1, \\ \omega X_1 + b > 0, Y_1 = +1. \end{aligned} \quad (2)$$

When $\omega \in R^l$, $\text{Margin} = 2/\|\omega\|$ can be obtained. To maximize $\text{Margin} = 2/\|\omega\|$, it is essential to make ω minimum and meet the following condition:

$$Y_1(\omega X_1 + b) - 1 \geq 0. \quad (3)$$

The generated SVM samples must be the samples located on the best segmentation surface and satisfy the equation. If there are linearly nonseparable samples, the kernel function will be used. Common kernel function names and equations are as follows.

The equation of polynomial kernel function reads

$$K(X_i, Y_j) = (X_i, Y_j + 1)^d. \quad (4)$$

The equation of Gaussian kernel function reads

$$K(X_i, Y_j) = \exp(-r|X_i - Y_j|)^d. \quad (5)$$

The equation of linear kernel function reads

$$K(X_i, Y_j) = \tan[b(X_i * Y_j) + c]. \quad (6)$$

3.2. Basic Principle Analysis of Logistic Model. The Logistic model is commonly used in the linear discriminant problems [11]. The dependent variable Y represents the probability of event occurrence. Event Y occurs when the independent variable X falls within the interval. For example, when $X > 0$, $Y = 1$; meanwhile, when $X < 0$, $Y = 0$. $Y = 1$ means that the event occurs, and $Y = 0$ means that the event does not occur. If Y_i^* (dependent variable) and X_i (independent variable) are linearly correlated, $Y_i = \alpha + \beta X_i + \delta$, and there is the following relationship:

$$\begin{aligned} P\left(Y_i^* = \frac{1}{X_i}\right) &= P(\alpha + \beta X_i + \delta_i > 0) \\ &= P(\delta_i > -\alpha - \beta X_i) = P(\delta_i \leq \alpha + \beta X_i) \\ &= F(\alpha + \beta X_i). \end{aligned} \quad (7)$$

$F(\alpha + \beta X_i)$ is the distribution function of δ_i , and the logistic distribution function of δ_i is $P = 11 + e - t$, so $i = F(\alpha + \beta X_i) = 11 + e - (\alpha + \beta X_i)$. The logistic regression model is $\ln P_i/1 - P_i = \alpha + \beta X_i$.

The process of the integrated SVM-Logistic model is as follows.

First, the samples should be classified by SVM technology; then, variables are substituted into the Logistic model, including the variables after the previous classification and other variables. The low credit risk of the evaluated company object is recorded as $Y = 1$; otherwise, $Y = 0$. The equation of the SVM-Logistic model reads

$$\ln \frac{P}{1 - P} = \alpha + \beta X. \quad (8)$$

In (8), P represents the default probability of the evaluated company and $1 - P$ denotes the compliance probability of the evaluated company object. The greater P is, the lower the credit is [12].

3.3. Theoretical Analysis of FMCDM. Fuzzy set theory has first seen its application in MCDM in the 1960s. Since then, fuzzy decision analysis has been widely used to solve uncertainty problems [13]. FMCDM problem has rich content with dynamic criterion weight coefficient information. The information of real numbers, fuzzy numbers, intuitionistic fuzzy sets, and vague sets can be determined. Soon, other FMCDM methods based on the fuzzy number, intuitionistic fuzzy set, and vague set have also been derived [14]. Figure 2 is a summary of the evolution and development of the FMCDM method.

3.4. Counting Influence of Index Weight Determination. AHP is a crucial scientific decision-making model, which objectively quantifies the weight analysis process rather than subjective qualifying [15]. The core of AHP is a structured hierarchy process [16]. AHP can be applied following several steps. The first is to divide the influencing factors to form a hierarchy, the second is to compare the factors with the

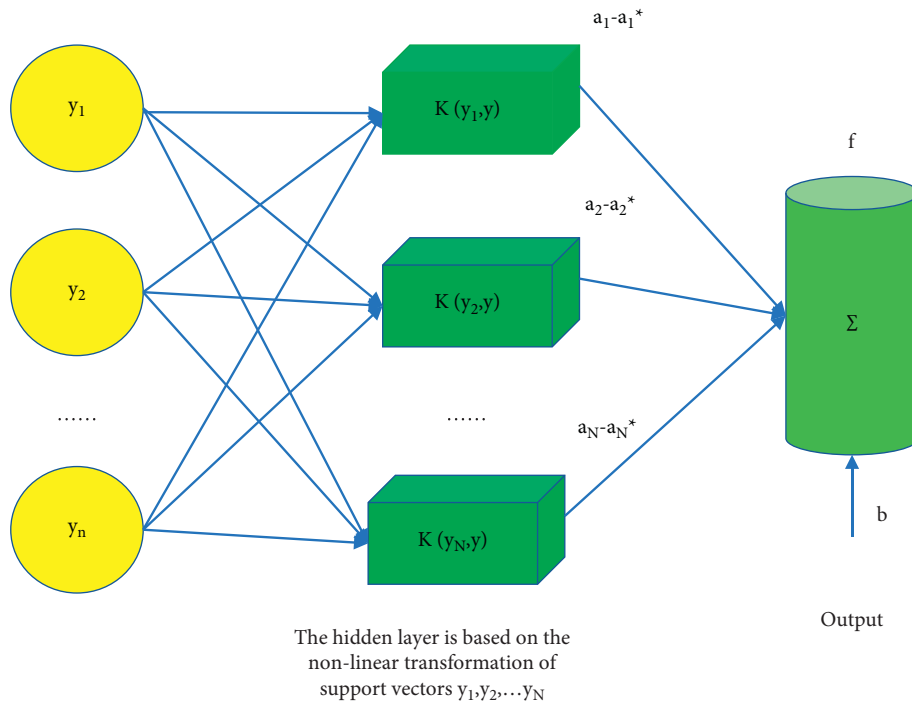


FIGURE 1: Schematic diagram of data mining technology based on SVM.

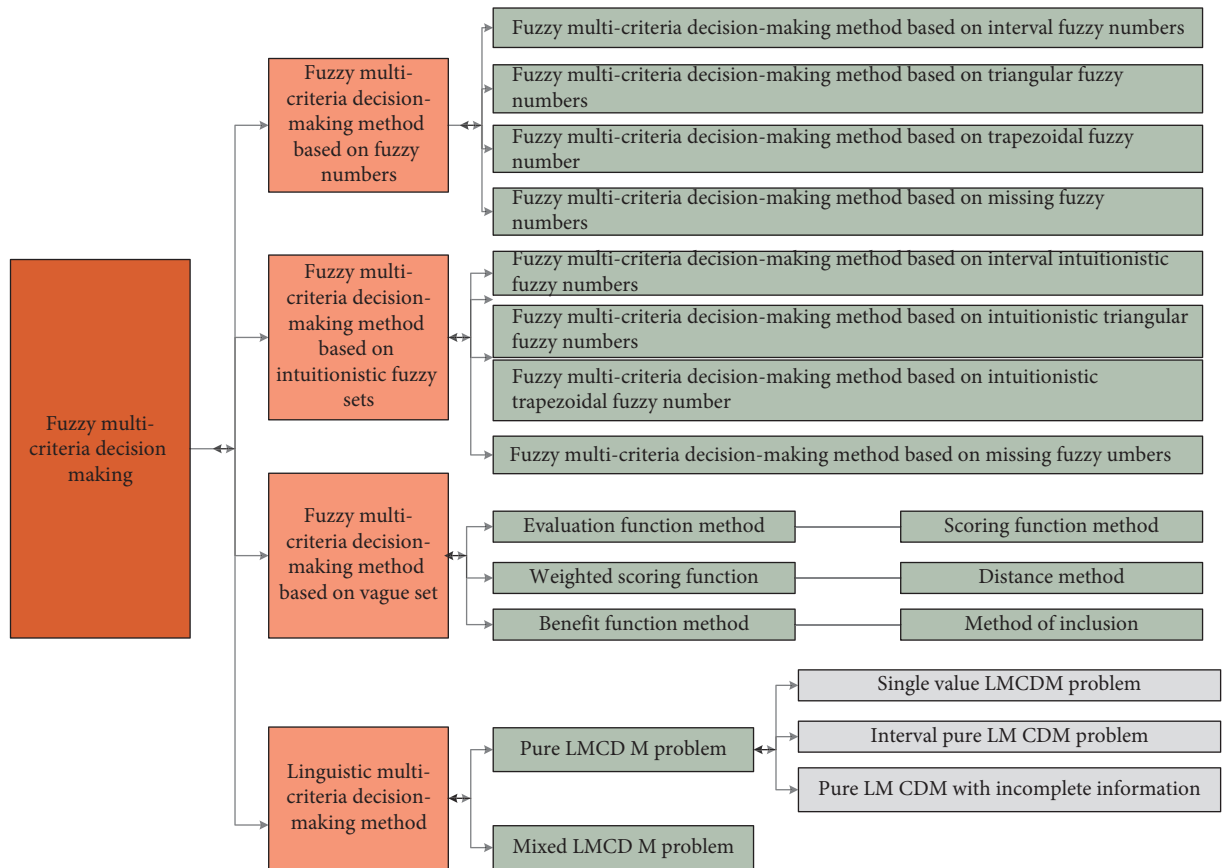


FIGURE 2: Summary diagram of FMCDM.

criteria, and the third is to calculate the relative weight of the factors to the criteria. The calculation basis is the decision-making matrix. Finally, the weight of factors at all levels to the overall decision-making objective is calculated. The comprehensive evaluation and ranking of schemes depend on the above process [17]. Figure 3 displays the mechanism of AHP.

The importance of criteria cannot be quantified in complex MCDM problems, so relative importance between paired criteria is adopted. This method is summarized as the nine-level scaling method. Figure 4 presents the meaning of a specific scale.

Under the criterion C_i , the relative importance relationship of n factors is marked according to the nine-level scaling meaning to form a decision-making matrix: $A = (a_{ij})_{n \times n}$. Under the criterion C_i , the scale of the importance of u_i relative factor u_j is a_{ij} , which is specifically defined as shown in the Figure 4 [18]. Each factor constituting decision-making matrix A must meet the following conditions: $a_{ij} > 0, a_{ji} = 1/a_{ij}, a_{ii} = 1$.

The eigenvector method (EM) proposed by Professor Saaty is usually used in weight calculation [11]. The eigenvalue solution equation of decision-making matrix A reads

$$A\omega = \lambda_{\max}\omega, \quad (9)$$

where λ_{\max} is the largest eigenvalue of A and ω is the corresponding eigenvector. After normalizing the obtained ω , the weight vector can be obtained. The detailed calculation process of EM is as follows:

- (1) The normalized initial vector $\omega = (\omega_1, \omega_2, \dots, \omega_n)^T$ of the same order as the decision-making matrix A is taken, where ω_i satisfies $\omega_i > 0$ and $\sum_{i=1}^n \omega_i = 1$.
- (2) $\bar{\omega}^{q+1} = A\omega^q, q = 0, 1, 2, \dots$ is calculated.
- (3) $\bar{\omega}^{q+1} = \bar{\omega}^{q+1} / \sum_{i=1}^n \bar{\omega}_i^{q+1}$ is normalized.
- (4) For any $\varepsilon > 0$, when $|\bar{\omega}_i^{q+1} - \omega_i^q| < \varepsilon (i \in N)$ is true, $\omega = \bar{\omega}^{q+1}$ is the weight eigenvector corresponding to the maximum eigenroot λ_{\max} of A . Moreover, there is the following expression:

$$\lambda_{\max} = \frac{\sum_{i=1}^n \bar{\omega}_i^{q+1}}{n\omega_i^q}, \quad (10)$$

$$\lambda_{\max} = \frac{\sum_{i=1}^n (A\omega)_i}{n\omega_i} = \frac{1}{n} \sum_{i=1}^n \frac{\sum_{j=1}^n a_{ij}\omega_j}{\omega_i},$$

where $\sum_{i=1}^n \omega_i = 1$.

The vector form of the relative weight $\omega_1, \omega_2, \dots, \omega_n$ of criterion C is $W = (\omega_1, \omega_2, \dots, \omega_n)^T$. ω_i satisfies $\omega_i > 0$, and there is the following expression:

$$\omega_i = \frac{\sum_{j=1}^n a_{ij}}{n \sum_{k=1}^n \sum_{j=1}^n a_{kj}}, \quad i = 1, 2, \dots, n. \quad (11)$$

It is essential to calculate the evaluation weight of each element when there is more than one standard, especially those elements at the lowest level to evaluate the overall goal

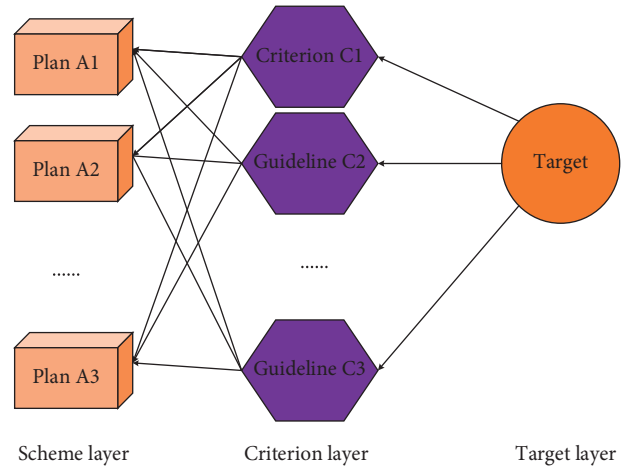


FIGURE 3: AHP diagram.

and select the best and worst solutions. Generally, for such a multilevel network, the weights of different standards at the lower level must be integrated from bottom to top to the upper level, and consistency must be checked at each level [19]. Different decision-making matrix might be constructed because of the complexity of objects and the limitations of subject perception. Consistency means that the decisions must be made logically correct without any contradiction [20], which, however, is almost impossible in MCDM. The contradiction factors are summarized: first, there are too many indexes, so logical contradictions are inevitable in pairwise comparison and evaluation; second, there are too many decision-makers to achieve unified decisions. The consistency of the decision-making table is checked to make the logic unified.

Generally, if all elements constituting decision-making matrix A meet $a_{ij} \bullet a_{kj} = a_{ik}$, A is called consistency matrix, which can also be expressed as that decision-making matrix A is qualified in consistency inspection.

In problem application, the consistency index (CI) is adopted to represent the consistency of the decision-making matrix. The equation reads

$$CI = \frac{\lambda_{\max} - n}{n - 1}, \quad (12)$$

where n represents the order of the matrix and λ_{\max} represents the maximum eigenvalue.

Ideally, when $\lambda_{\max} = n$, that is, $CI = 0$, the decision-making matrix is completely consistent. However, in general, $CI > 0$. Moreover, the larger the CI is, the more difficult it is to achieve the consistency of the decision-making matrix.

The average random consistency index (RI) is introduced to deal with the inconsistency of the discrimination matrix caused by the increase of the dimension of the decision matrix [21]. The consistency rate (CR) can be obtained through CI and RI :

$$CR = \frac{CI}{RI}. \quad (13)$$

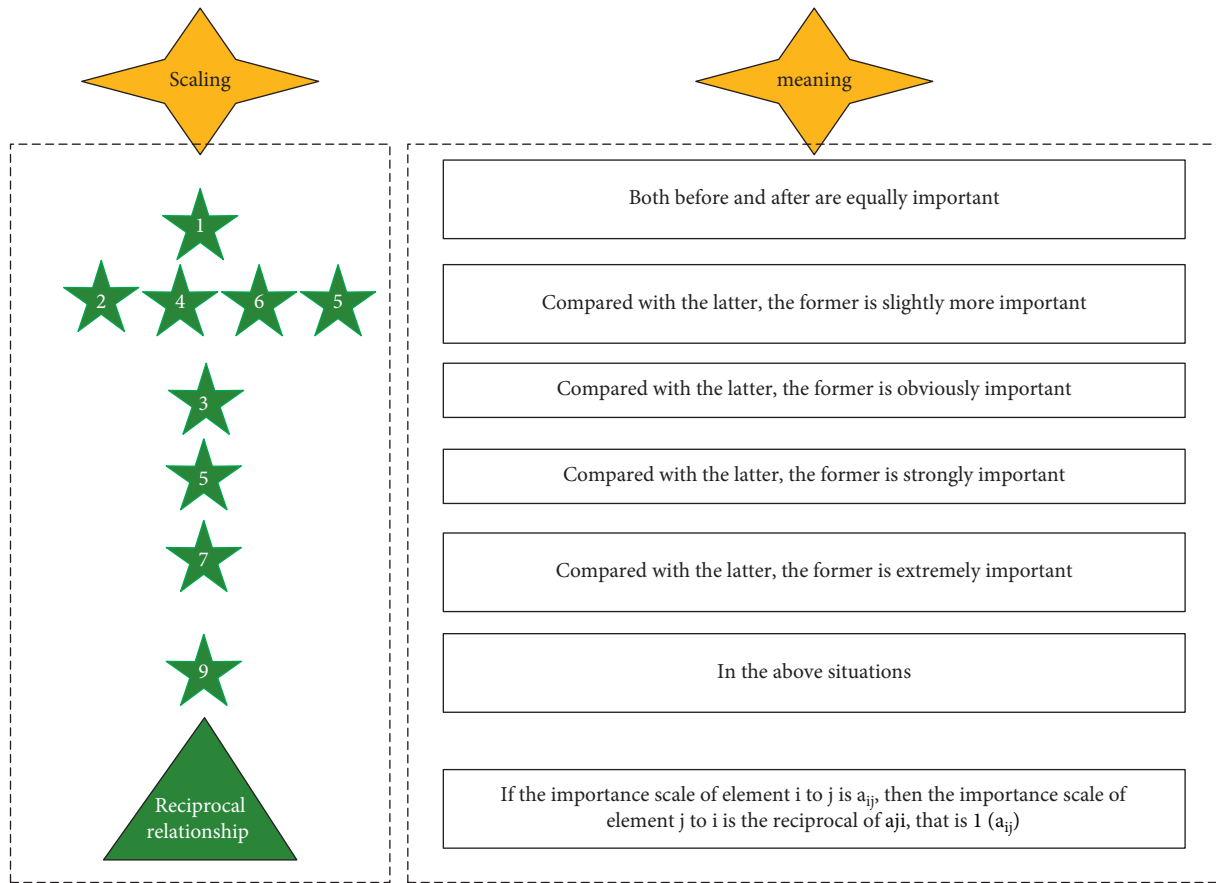


FIGURE 4: Schematic diagram of criterion importance scale.

The consistency of matrix A can be judged by the value of CR. It is concluded that the consistency of matrix A is too poor when $CR \geq 0.1$; the consistency of matrix A passes the consistency test when $CR < 0.1$. Figure 5 shows the RI value of the lower-order matrix.

AHP can be used to solve MCDM problems thanks to the standardized and structured features through which the complex problems can be transformed into simple and separate hierarchical problems. The credit risk assessment problem studied matches the structured feature of the AHP method. The level of corporate credit risk has a target attribute, and each financial impact index of credit risk status has the standard attribute. The evaluated company object can be employed as the final scheme level, so the AHP method is adopted. This study combines the research results of worldwide scholars to establish a hierarchical corporate credit risk assessment index system as shown in Table 1.

4. Results

4.1. Samples Affecting Credit Risk Assessment Results. Data are mined from the financial reports of sample listed companies in China stock exchange provided by GTA Research Service Center [22]. Specifically, 12 sample companies are selected. Table 2 displays the legitimate company

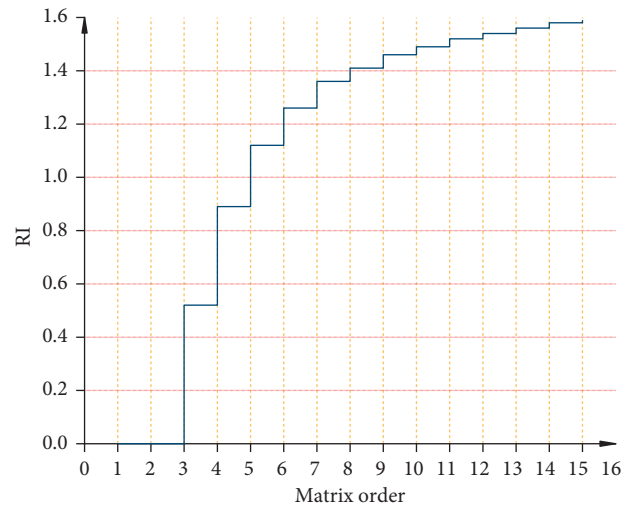


FIGURE 5: RI values of matrices of orders 1 to 15.

stock code and risk, as well as the credit risk status assessment from the market supervision department.

In Table 2, “ST” indicates the company’s credit default, while “SF” stands for the company’s credit security. Because of the timing cumulative characteristics of credit risk and the regularity of data release according to time, the index data of

TABLE 1: AHP indexes of target companies.

Target layer	Criterion layer	Index layer	Indicial notation	
Corporate credit risk	Profitability	Rate of return on fixed assets	Z1	
		Return on equity	Z2	
		Return on invested capital	Z3	
		Long-term return on capital	Z4	
		Business tax rate	Z5	
		Ratio of profits to cost and expense	Z6	
		Charge ratio	Z7	
		Debt coverage ratio	Z8	
		Cash flow capacity	Cash ratio of operating income	Z9
			Cash reinvestment ratio	Z10
	Ratio of accounts receivable to income		Z11	
	Operation capability	Receivables turnover ratio	Z12	
		Long-term asset turnover	Z13	
		Total asset turnover	Z14	
		Rate of capital accumulation	Z15	
	Development capacity	Increasing rate of fixed assets	Z16	
		Total assets growth rate	Z17	
		Current ratio	Z18	
	Short-term liquidity	Conservative quick ratio	Z19	
		Working capital ratio	Z20	
		Cash ratio	Z21	

the sample companies with a sample period of 8 quarters are selected.

The latest default status evaluation results are excluded from sample data to minimize the influence of psychological factors, such as subjective overestimation. The financial information of listed companies in the first two periods (quarters) of ST is adopted to implement a model to predict their default probability. ST is a crucial regulatory system used in the stock market. Abnormal risk state and other abnormal states are two situations under the system. In reality, the abnormal risk state is the normal state, which most companies will be in, so it can be used as a sign of default of listed companies. In other words, for these companies in 2020, the data of eight quarters from September 2017 to September 2019 are selected.

4.2. Application Effect of SVM-Logistic Model in Credit Risk Assessment. The integrated SVM-Logistic model is employed for risk assessment. The specific operations are as follows: in the first step, the appropriate kernel function is selected for credit assessment through the comparison of the classification results from each kernel function, as shown in Figure 6.

Figure 6 reveals that the accuracy of three different kernel functions is different. The accuracy of the polynomial kernel function is the highest, followed by the Gaussian kernel function, and the accuracy of the linear kernel function is the poorest.

In the second step, the polynomial kernel function is employed to classify the samples, and the results are introduced into the Logistic model as an independent variable to participate in the training test. Figures 7 and 8 show the comparison between training results and prediction results.

The training set results in Figures 7 and 8 show that the prediction performance of the Logistic model is poor, while

TABLE 2: Actual credit risk status of sample companies in 2020.

Stock code	Risk status	Stock code	Risk status
600077	ST	000513	SF
600131	ST	000703	ST
600373	ST	600156	SF
002215	SF	002022	SF
002250	SF	600769	ST
002349	SF	600806	SF

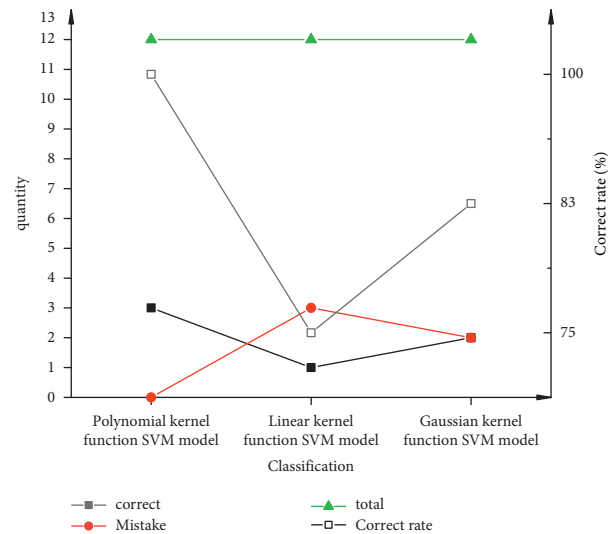


FIGURE 6: Comparison of three kernel functions of SVM.

the prediction results of the SVM-Logistic model and SVM model are basically the same, indicating that the prediction accuracy of the SVM-Logistic model and single model on the training set is high. The test set classification results in

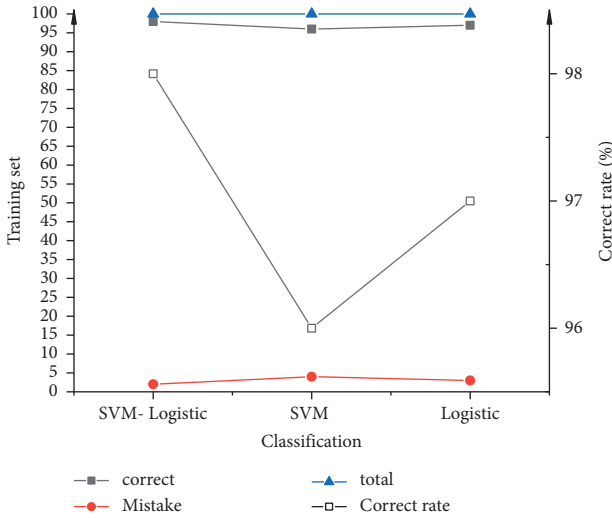


FIGURE 7: Comparison of different models on the training set.

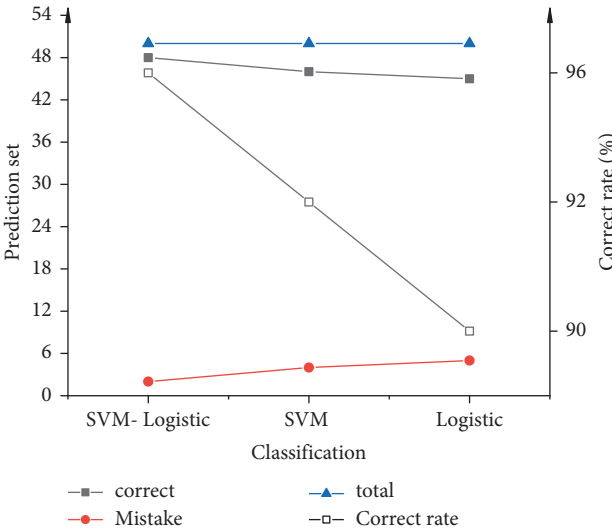


FIGURE 8: Comparison of different models on the prediction set.

Figure 7 show that the accuracy of the SVM-Logistic model is still higher than that of the single model, and the fluctuation of the single model is relatively greater than that of the SVM-Logistic model.

4.3. Application Effect of FMCDM Based on Fuzzy Set. The FMCDM method based on a fuzzy set is adopted to calculate the sufficiency, effectiveness, and total weight of each feature, and a key feature subset can be selected according to the distribution of the total feature weight. Figure 9 illustrates the feature weight calculation.

12 indexes with comprehensive weight no less than 0.3 can be obtained from the attribute comprehensive weight, and then correlation analysis is conducted to get 9 key uncorrelated indexes. AHP method is very commonly used in MCDM. The weight of the above indexes is determined by asking relevant experts to score and calculate. After a series

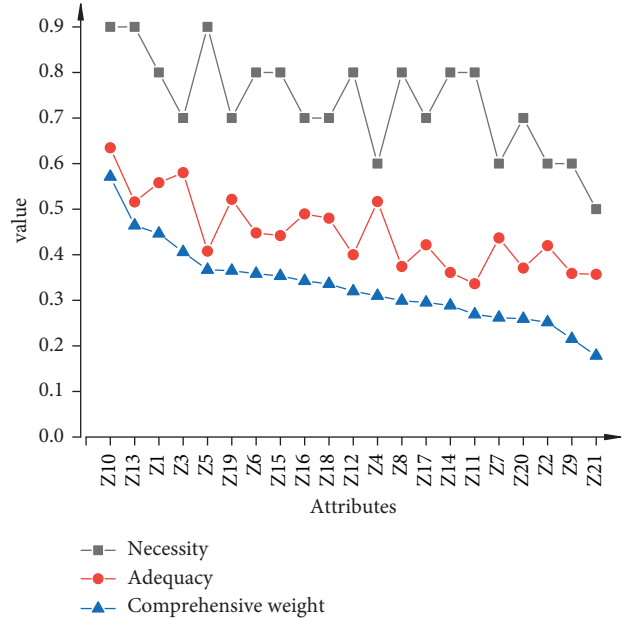


FIGURE 9: Attribute weight result chart.

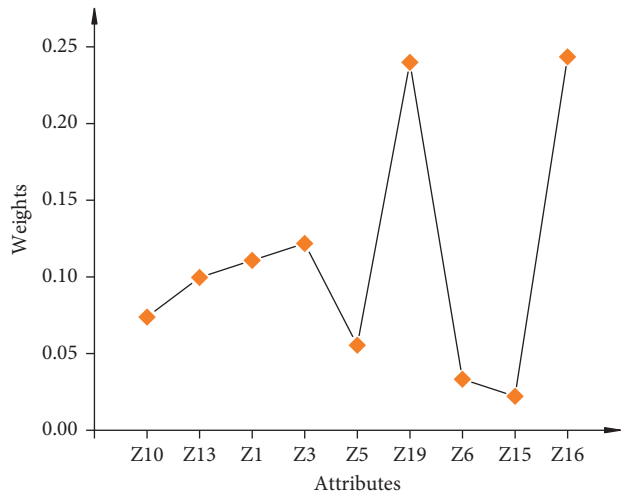


FIGURE 10: Key indexes and weights.

of interviews, consultation, consistency check, and feedback adjustment, the final weight of each index is obtained, as shown in Figure 10.

The credit risk of sample companies is assessed through the integrated SVM model+FMCDM. According to the provisions on sample selection, the fitness of the credit risk assessment results of the selected companies is obtained based on the financial index data of domestic listed companies from June 2017 to June 2019 in 2020. Figure 11 displays the results.

The stock code in the figure represents each sample company. The credit default risk measurement and risk ranking of the company are expressed by Fit. Fit indicates the probability of credit default of the sample company. The lower the Fit is, the higher the probability of the default is.

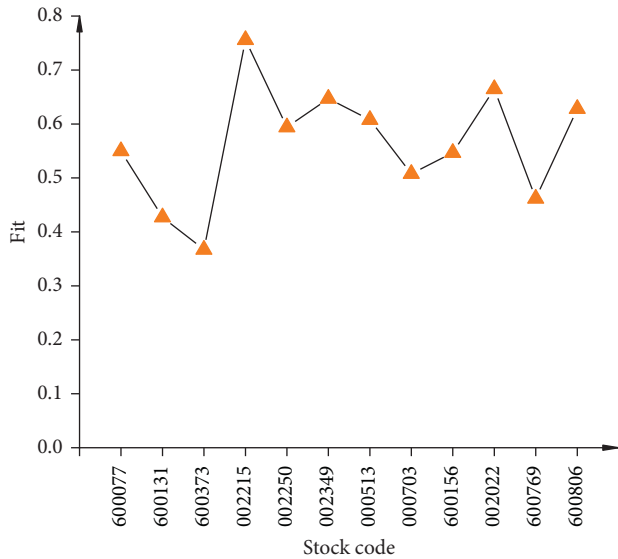


FIGURE 11: Fitness results of credit risk assessment.

Compared with the actual risk status, the companies assessed as ST by the regulatory authorities basically have a low Fit. Besides, ST sample companies and SF sample companies can be matched according to 0.5, and the appropriate threshold can be selected as the standard for judging other sample companies. In this sample, the threshold for company default risk measurement is set as 0.56152. If Fit is lower than the threshold, it indicates that the corporate credit risk is high. Otherwise, corporate credit risk is low.

5. Conclusions

Under the background that the number of corporate economic disputes from credit risk is increasing yearly, there is a rising demand for legal service companies to try and practice new credit risk assessment methods. Here, SVM in data mining technology is selected to classify the sample data concerning corporate credit risk; then it is combined with the logistic regression model to determine the corporate credit risk level, and, further, the integrated SVM-Logistic model is combined with FMCDM to calculate the weights of attributes and key indexes. According to the fit value obtained by the integrated model, the company credit risk status can be accurately evaluated. The accuracies of the training set and the test set of the integrated model are 98% and 96%, respectively. The intermediate threshold of the company's credit default risk measurement and risk fitness (Fit) is obtained as 0.56152: when Fit is lower than the threshold, the company's credit is low; when Fit is higher than the threshold, the company's credit is high. Therefore, the data mining technology based on the integrated SVM-Logistic model + FMCDM has high precision and feasible application in the credit risk assessment service of legal service companies.

This research has created a new method model for legal service companies in the field of corporate credit risk

assessment, improves the credit risk assessment efficiency of legal service companies, and provides a reference for the corporate credit risk assessment. It has made great efforts to resist the systematic financial risks from the Internet financial bubble and has great practical significance. Still, due to the limited amount of sample data, the proposed method needs to be further verified in terms of stability and reliability under the big data environment, as well as to face challenges from possible development retardment. Therefore, the following prospects are put forward. There are various theoretical methods and models to apply data mining technology and FMCDM in corporate credit risk assessment, in which exploration of internal correlation requires the joint exploration and efforts of researchers and practitioners from all walks of life.

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 no conflicts of interest.

References

- [1] H. Lei and S. C. Zhao, "Corporate social responsibility and corporate credit risk assessment-An innovative brief of logistic model based on stakeholder perspective," *Journal of Hunan University:Social Science Edition*, vol. 2019, no. 2, pp. 88–95, 2019.
- [2] H. W. Du and X. W. Lu, "Credit risk assessment of P2P borrowers based on SVM-Logistic combinatorial model-taking microloan.com as an example," *Productivity Research*, vol. 315, no. 10, pp. 37–42, 2018.
- [3] J. Zheng and X. Li, "Research on enterprise credit risk assessment based on SVM model," *Enterprise Technology and Development*, vol. 5, pp. 220–221, 2020.
- [4] X. D. Hu, R. Cao, J. M. Li, S. M. Ruan, and X. Fang, "Research on the construction of IDGSO-BP integrated model for credit risk assessment of micro and small enterprises," *Operations Research and Management*, vol. 26, no. 4, pp. 132–139, 2017.
- [5] W. G. Zhang, Y. Y. Lu, and Y. J. Liu, "Credit risk assessment of P2P loan borrowers based on non-equilibrium fuzzy approximation support vector machine and its application," *Systems Engineering Theory and Practice*, vol. 38, no. 10, pp. 2466–2478, 2018.
- [6] G. Charkabarty, S. F. Chen, L. H. Li, and C. T. Lin, "Credit risk assessment using regression model on P2P lending," *International Journal of Applied Science and Engineering*, vol. 16, no. 2, pp. 149–157, 2019.
- [7] M. Mangeli, A. Shahraki, and F. H. Saljooghi, "Improvement of risk assessment in the FMEA using nonlinear model, revised fuzzy TOPSIS, and support vector machine," *International Journal of Industrial Ergonomics*, vol. 19, no. 69, pp. 209–216, 2019.
- [8] F. R. Jiang, "Research on personal credit evaluation model of online lending based on GA-SVM," *Control Engineering*, vol. 27, no. 186, pp. 109–115, 2020.

- [9] N. L. Ma, Y. Q. Li, Y. Cao, and Q. X. Wu, "Design of mortgage credit risk assessment system based on data mining," *Hubei Agricultural Mechanization*, vol. 238, no. 1, p. 149, 2020.
- [10] S. Y. Wang, X. R. Lan, and L. Y. Liu, "SVM model based on non-convex penalty for credit risk assessment of technology-based small and medium-sized enterprises," *Mathematics in Practice and Knowledge*, vol. 49, no. 3, pp. 307–312, 2019.
- [11] A. Levy and R. Baha, "Credit risk assessment: a comparison of the performances of the linear discriminant analysis and the logistic regression," *International Journal of Entrepreneurship and Small Business*, vol. 42, no. 1/2, pp. 169–186, 2021.
- [12] P. Liu and Z. Zeng, "Theory and application of logistic-SVM two-stage hybrid model," *Journal of Physics: Conference Series*, vol. 1746, no. 1, Article ID 012017, 2021.
- [13] S.-S. Lin, S.-L. Shen, A. Zhou, and Y.-S. Xu, "Risk assessment and management of excavation system based on fuzzy set theory and machine learning methods," *Automation in Construction*, vol. 122, no. 122, Article ID 103490, 2021.
- [14] S. Stoilova and N. Munier, "A novel fuzzy SIMUS multi-criteria decision-making method. An application in railway passenger transport planning," *Symmetry*, vol. 13, no. 3, p. 483, 2021.
- [15] F. Wu, X. Su, Y. S. Ock, and Z. Wang, "Personal credit risk evaluation model of P2P online lending based on AHP," *Symmetry*, vol. 13, no. 1, p. 83, 2021.
- [16] P. K. Roy and K. Shaw, "Modelling a sustainable credit score system (SCSS) using BWM and fuzzy TOPSIS," *The International Journal of Sustainable Development and World Ecology*, vol. 2021, no. 4, pp. 1–14, 2021.
- [17] J. H. Wang and J. A. Jia, "Research on personal credit risk evaluation of P2P online credit based on intuitionistic fuzzy analytic hierarchy process," *Value Engineering*, vol. 038, no. 7, pp. 1–4, 2019.
- [18] X. Huang, J. Sun, and X. Zhao, "Credit risk assessment of supply chain financing with a grey correlation model: an empirical study on China's home appliance industry," *Complexity*, vol. 2021, Article ID 9981019, 12 pages, 2021.
- [19] A. I. Chowdhury, "A hybrid model for credit risk assessment: empirical validation by real-world credit data," *Journal of Risk Model Validation*, vol. 14, no. 4, pp. 1–39, 2021.
- [20] G. Du, Z. Liu, and H. Lu, "Application of innovative risk early warning mode under big data technology in Internet credit financial risk assessment," *Journal of Computational and Applied Mathematics*, vol. 386, no. 12, Article ID 113260, 2021.
- [21] Z. Huang, Z. Yan, Q. Zhao, and K. Ma, "Credit risk assessment in bank based on SMEs using PCA," in *Proceedings of the 2021 4th International Conference on Advanced Algorithms and Control Engineering (ICAACE 2021)*, vol. 1848, Sanya, China, January 2021.
- [22] N. A. Manab, N. Aziz, and D. M. Jadi, "Sustainability risk management: an integrative framework to evaluate emerging risks and other non-quantifiable risks affecting company survival," *World Review of Science, Technology and Sustainable Development*, vol. 16, no. 2, pp. 87–104, 2020.