Eng Int Syst (2025) 2: 213–221

Engineering
Intelligent Systems

The Evaluation of Enterprise Financing Structure Capability Based on RF-CART Integrated Algorithm

Zhaofen Meng[1,*], Min Liu[1] and Bingjie Li[2]

[1] *School ofAccounting, Yanshan College Shandong University of Finance and Economics, Ji'nan, 250000, China*
[2] *School of International Business, Yanshan College Shandong University of Finance and Economics, Ji'nan, 250000, China*

The continuous advancement of information technology and the expansion of application scenarios have highlighted the advantages of data fusion and personalized evaluation of computer technology, which has great potential for application in the financing field of small and medium-sized enterprises. Based on the current information asymmetry and limitations of subjective evaluation, an evaluation model to determine the capability of a financing structure is proposed, based on the RF-CART integrated algorithm. Firstly, the impact indicators of financing structure for small and medium-sized enterprises were collected, a principal component analysis was conducted, and the characteristic variables for the indicators were constructed. Secondly, the CART decision tree was used as a weak learner to construct a credit evaluation model under the RF model, and the classification results processed by multiple decision trees were integrated to obtain the optimal classification result. By using indicator testing and cross validation to analyze the effectiveness of the model algorithm, it was found that there is a negative correlation between the tax effect and financing capacity of small and medium-sized enterprises. The testing and training time of ensemble learning algorithms are both less than 21 seconds, with an average recognition accuracy of over 95%. The accuracy difference of other comparative calculation methods is 4%, and their AUC area for recognition performance is 0.915, indicating good sample discrimination ability. The proposed method can effectively determine the operational status of enterprises and provide indicators and warnings related to financing risks and any decisions pertaining to growth.

Keywords: integrated learning; small and medium-sized enterprises; financing structure; KS; random forest algorithm

1. INTRODUCTION

China's economic system has undergone a number of reforms, and is being increasingly integrated with the global economy. As an important element of China's economic development and social progress, SMEs play a crucial role in fiscal revenue, jobs and employment opportunities. Moreover, SMEs offer unique advantages in terms of forward and backward economic development [1]. SMEs are relatively small in terms of stafng and business scale, so their business mode is mainly direct management and less susceptible to external environmental conditions. However, SMEs are less able to cope with capital risks and changes in the market environment. The development of SMEs is restricted by the differences in their internal control mechanisms, as well as the development of information asymmetry and low liquidity. Also, the impact of changes in the policy environment and the impairment of the capital chain circulation can greatly limit the development of SMEs [2]. The diffculty in financing has become the bottleneck that hinders the further development of SMEs. The main reasons are the diffculty in loans, the imperfect guarantee system and the high interest rate. Due

*Email: mengzhaofen522@163.com

to the short life cycle of SMEs, their own management ability and experience are somewhat inadequate, which makes their operators in efficient in regard to decision-making and implementation [3]. Several scholars have also conducted research on risk assessment, among which Liu J proposed to apply the model chain using the XGBOOST and random forest (RF) fusion algorithm for supply chain risk assessment. The model also analyzes the credit risk and operation status of enterprises. The credit risk classification method based on RF has a 6.39% improvement in assessment accuracy and has good results in empirical application [4]. Uddin MS scholars believe that the RF can perform effect analysis on the credit situation of micro enterprises. The empirical results show that non-traditional variables can effectively help market participants better predict customer credit default [5]. Strengthening the evaluation of the financing capacity of SMEs can assist businesses to formulate decision-making plans and strategies for future development. The differences in regional environment, changes in economic policies and market policies, as well as the enterprise's own industrial structure and innovation ability, financing ability and industry development prospects will all affect the success of its financing structure. Therefore, in this study, the structure of influencing factors os analyzed with the help of an integrated learning algorithm, so as to adjust the financing risk of an SME in order to improve its structural ability and its own anti-risk ability.

2. RELATED WORKS

At present, there are many gaps in strategic innovation research in manufacturing and SMEs. AlQershi N and other academic teams proposed a mechanism for structural capital and strategic innovation in response to these problems. The experimental data were collected by field investigation, and the hypothesis was tested with the help of the PLS-SEM model of the multiplier structure. The results show that strategic innovation plays an important role in the development of enterprises. Also, the adjustment of structural capital can play a significant regulatory role in both, and improve enterprise performance [6]. To better explore the impact of different market entities on the financing situation of SMEs, Van Klyton analyzed experimental data with the help of computational tools. The results show that the inadequate governance structure and the lack of management ability of enterprise managers are the important reasons for the stagnation of SME financing. The results can provide ideas for improvement and research value for the long-term development of enterprises and financing of entrepreneurship [7]. Because the current information architecture is rarely used in card classification data, Paea and Baird propose using a similarity matrix and cluster analysis to achieve multidimensional data architecture. The results show that this method can effectively achieve visual information architecture and centralized data processing [8]. The social media platform provides new channels and ways for the development of SMEs, making their enterprise activities no longer limited to the original single form. The instability of social media platforms in enhancing corporate innovation capabilities has to some extent constrained the improvement of corporate performance. Fan et al. analyzed the impact of entrepreneurial orientation, strategic innovation and other factors on enterprise performance in terms of available resources, and trained the test samples with the PLS algorithm. They found that social media plays an important regulatory role in the strategic relationship of enterprise performance [9]. Tang proposed an accounting selection model based on big data analysis given the large amount of accounting calculation time consumed by SMEs and the difficulty in meeting user needs. The AMBD constraint system is constructed with relevant impact indicators, and the optimal effect of the model is achieved with the help of particle swarm optimization and principal component analysis. The results show that this method can provide an effective accounting audit decision basis for SMEs [10]. It is necessary to strengthen the information business deployment capability of supply chain finance and ensure the effective development of cash flow in logistics finance. Guo L et al. proposed a combination of blockchain technology with the Internet of Things technology to give play to the combined force of decentralization and information connectivity, and improve the transparency of business transaction information through an established information management framework. The results show that this method can achieve the integration of multi-dimensional data flow, and effectively connect capital, information, trade and logistics. This method has wide applicability and proven effectiveness [11].

The lack of transparency of business data makes the credit decision-making process of finance providers more complex and difficult. Malakauskas and Lak š tutien ė scholar combined logical regression with a neural network algorithm, and carried out multi-model analysis on the original prediction factors of a financial cycle. This method can analyze and predict financial risks with a high level of accuracy [12]. Pratono et al. used the least square structural model to explore the relationship between risk taking and enterprise performance to conduct intermediary analysis and adjustment mechanism, and analyzed it using case study data. The risk-taking behavior can bring about a significant change to the performance and development of enterprises during an economic downturn [13]. Liu and Huang proposed applied the support vector machine model to conduct an empirical analysis of supply risk in order to maximize the risk classification performance of a support vector machine. In order to achieve accurate processing of the data, Liu used fuzzy clustering and principal component analysis, and took into account the interference of data noise on prediction error.. This method can screen out the influencing factors of enterprise supply chain financial analysis, and its credit evaluation accuracy is good [14]. Drytakis analyzed the risk situation of SMEs in the epidemic situation with the commercial risk scale, and used t case study data to analyze the cash flow forecast and AI to determine the degree of business risk to an enterprise. Drytakis found that the application of intelligent technology can effectively reduce the business risk of enterprises and help them improve their dynamic planning ability [15]. Most scholars have found that internal and external environmental factors will affect the financing institutions and operations of enterprises, and that integrated algorithms such as stochastic forest algorithm

Table 1 Variable factors of financing structure capacity.

Indicator type		Code
Enterprise size		Q1
Market competitiveness		Q2
Asset value guarantee		Q3
Operational capacity	Inventory turnover rate	Q4
	Debt collection turnover rate	Q5
Profitability	Return on net assets	Q6
	Return on total assets	Q7
Anti-risk capability		Q8
Tax effect		Q9
Internal accumulation level		Q10
Enterprise growth potential		Q11

can effectively evaluate the risk situation. Therefore, in this research, an integrated learning algorithm is applied to evaluate and analyze the financing capacity structure of SMEs, with a view to giving enterprise managers a sound basis for their decision-making.

3. EVALUATION OF SMES' FINANCING STRUCTURE CAPABILITY BASED ON INTEGRATED LEARNING ALGORITHM

3.1 Analysis of Factors Affecting the Capacity of SMEs' Financing Structure

The common evaluation model used to determine the capability of SMEs' financing structure is achieved by constructing a characteristic variable system and the integration of its variables. The selection of characteristic variables should be comprehensive and scientific. This ensures that the characteristic variables can actually reflect the overall development of the enterprise and its credit risks, which is highly explanatory. And the construction of the variable system should also be based on the actual development and management level of different enterprises to help them develop a practical and applicable architecture [16]. Given the diversity and complexity of the data sources of feature variables, it is necessary to first quantify and structure the variables to obtain feature engineering which comprises construction, extraction, selection and evaluation according to its stage dimension. In this study, the variable factors were screened, to differentiate between the financing structure ability of SMEs and their actual operations. Table 1 shows the variable factors.

The indicators in Table 1 are used to analyze the various factors affecting the financing structure capability from the micro perspective of enterprises, aiming to help enterprises better evaluate their overall financing capability. The dimensionality reduction of the data in the indicator system can reduce the volatility of the variable data. The mathematical expression of the indicator data in the observation object is shown in formula (1):

$$X_{(i)} = (X_{i1}, X_{i2}, \ldots, X_{ip}) \tag{1}$$

where i is the number of observation indicators and p is the number of indicators in formula (1). The information extracted by the principal component and its contribution to the variable information are expressed, and calculated with formula (2):

$$\eta_i = \lambda_i / \sum_{k=1}^{p} \lambda_l \tag{2}$$

where η is the main component contribution rate, λ is the indicator characteristic value, and l is the number of indicator components. The cumulative contribution rate of all index components can better represent the ratio difference between the principal component variance and the total variance of the index. The greater the numerical difference, the more information it provides. In classification problems, the IV (Information Value) value is often used to indicate the strength of the prediction ability of the characteristic variables to the results. For instance, take 0.02 and 0.5 as the two interval endpoints of the IV value taking ability, and 0.02 indicates no predictive ability, and 0.50 indicates the best predictive ability. The increasing value within this range can indicate the improvement of characteristic variable prediction ability, and 0.2 is the interval of the range [17–18]. The WOE coding method is used to represent the weight; it is expressed with formula (3):

$$WOE_i = \ln\left(\frac{p_{bad}}{p_{good}}\right) = \ln\left(\frac{B_i}{B} / \frac{G_i}{G}\right) \tag{3}$$

where p_{good}, p_{bad} represent the percentage of the corresponding good and bad samples in the corresponding indicator variables in formula (3). $_{good,bad}$ are the good sample and bad sample respectively. B is the total number of samples. G_i, B_i represent the number of good samples and bad samples in the characteristic variable. Formula (4) is used to calculate corresponding characteristic variable, which can be obtained by calculating the IV value with WOE.

$$IV_i = \left(\frac{B_i}{B} / \frac{G_i}{G}\right) * \ln\left(\frac{B_i}{B} / \frac{G_i}{G}\right) \tag{4}$$

In formula (4), i represents the grouping of the characteristic variables, and by summing up equation (4), equation (5) is obtained:

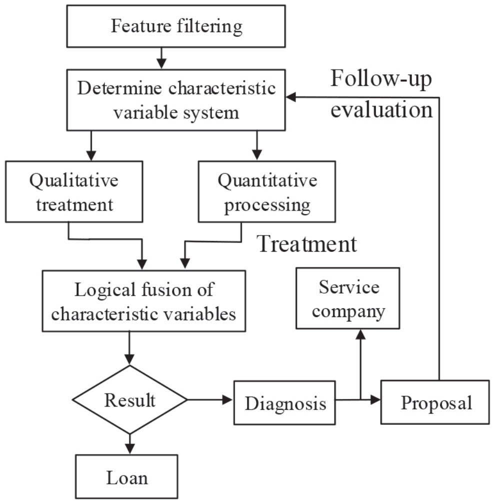

Figure 1 Construction of characteristic variables.

$$IV = \sum_{i}^{n} IV_i = \sum_{i}^{n} \left[\left(\frac{B_i}{B} / \frac{G_i}{G} \right) * \ln \left(\frac{B_i}{B} / \frac{G_i}{G} \right) \right] \quad (5)$$

where n is the number of groups in formula (5). The selection of characteristic variables can effectively reflect the key information of indicator data. Figure 1 shows the process of constructing the characteristic variables.

The indicator system for the selected characteristic variables is constructed, and then quantitative and qualitative processing is carried out on them to achieve the integration of different variable indicators. If the output results can accurately reflect the financing situation of the main body, it can be financed by financial means. If the result of this indicator is unable to reflect the financing situation, it needs to be diagnosed and tracked, evaluated and analyzed according to the actual situation until it can extract the factors and variables that affect the corporate financing structure.

3.2 Enterprise Financing Credit Risk Analysis Based on Integrated Algorithm

In different stages of development, enterprises have different demands for capital and finance, and their financing methods will also vary. In the early and late stages of development, enterprises rely mainly on internal financing and external market financing to develop their business and increase growth. At present, there are problems associated with an inadequate financial credit system and mixed information and data between SMEs and individual consumption in China. Strengthening the extraction and processing of data features has become an important aspect of improving the evaluation of SMEs' financing structure capacity [19]. Therefore, in this paper, the credit problem is analyzed using a combination of big data and small data, and data-oriented credit model is constructed. Based on the evaluation principle of financial engineering credit and the integrated learning idea, the evaluation model of financing structure capacity of SME in line with the development of China's national conditions is constructed. The integrated learning algorithm is able to combine the features of multiple models to achieve comprehensive evaluation. Mathematically, this can be expressed with a linear equation such as formula (6):

$$H(x) = \omega^T h(x) + b \quad (6)$$

$h(x), \omega$ represents the classification result and the corresponding result weight in formula (6). x, b represents the sample point and parameter. The financing structure of enterprises is the amount of funds that the enterprise managers can obtain in the market with the help of different types and different ways of funding debt in the capital financing market. It is important for enterprises to acquire funds and manage their use. Moreover, the quality of credit rating will play an important role in the operation of enterprises, especially for SMEs. In this study, the stochastic forest algorithm is used to build the credit evaluation model, and the classification regression tree (CART) is used to build the sub-model. As a supervised learning algorithm, the CART classification tree can be trained with the help of continuous and discrete features and build the structure comprising complex data. This principle includes the construction and pruning of classification tree. During the generation of the CART classification tree, the characteristic variable with the smallest Gini coefficient is selected as the split node for processing. Equation (7) shows the mathematical expression applied to define the Gini coefficient:

$$Gini(D) = 1 - \sum_{k=1}^{K} \left(\frac{C_k}{D} \right)^2 \quad (7)$$

where D, K represents the data set and classification number. The classification number includes different sample sets $C_k.k$ is the number of sample sets. The coefficient of each characteristic variable is calculated, and the coefficient of the total data set divided by attribute difference can be expressed with formula (8):

$$Gini(D, A) = \frac{|D_1|}{|D|} Gini(D_1) + \frac{|D_2|}{|D|} Gini(D_2) \quad (8)$$

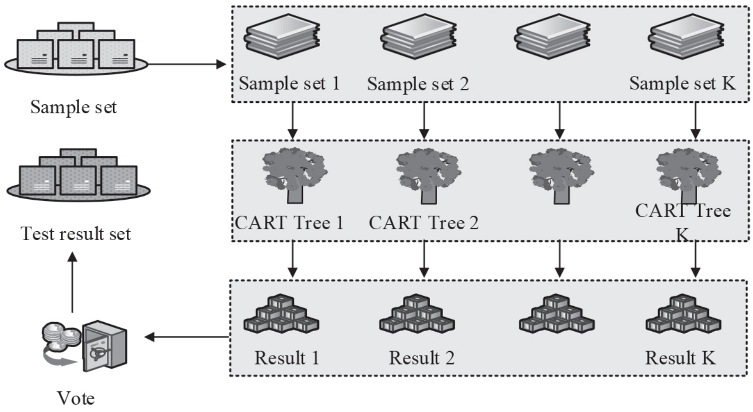

Figure 2 RF flow diagram.

Table 2 IV value of characteristic variable.

Indicator type		**IV value**
Enterprise scale (Q1)		1.441336
Market competitiveness (Q2)		1.606210
Asset value guarantee (Q3)		2.448535
Operational capacity	Inventory turnover rate (Q4)	2.483178
	Debt collection turnover rate (Q5)	3.364865
Profitability	Return on equity (Q6)	3.681499
	Return on total assets (Q7)	3.734281
Anti-risk capability (Q8)		3.968662
Tax effect (Q9)		4.022455
Internal accumulation level (Q10)		4.334627
Enterprise growth potential (Q11)		5.605065

where $D_1 D_2$ means that the data set is divided into two parts according to the same feature A in formula (8). The smallest characteristic variable is taken as the best splitting node to generate the branching result. The node-splitting steps are repeated until the categories they contain are unique. RF (RF) is an algorithm based on decision tree and combining Bagging integration technology and random subspace. It has a simple structure and can effectively and conveniently adjust parameters, and has good scalability and parallelism in high-dimensional data processing [20–21]. And the algorithm is less disturbed by abnormal data values and noise environment during data prediction, and its tolerance and anti-fitting ability are good. RF is a classifier based on CART, and its algorithm flow is shown in Figure 2.

The construction process of the algorithm in Figure 2 is mainly completed in four steps. First, a corresponding number of training sample sets are selected from the original sample size by means of the random sampling method with return. The number of corresponding individual sample sets is the same as the number in the original sample number set. Secondly, the corresponding attribute is selected in the sample attribute set, and the algorithm is trained with the optimal characteristic variable. Each training sample set can generate the corresponding CART. Finally, the classification results processed by multiple decision trees are integrated to select the optimal classification results [22–23]. Bagging's idea is to train the number of samples by putting them back. Some data has not been extracted from the training samples; this data is known as *out of bag*; its corresponding probability is found with formula (9) [24]:

$$P = \left(1 - \frac{1}{N}\right)^N \tag{9}$$

where N represents the corresponding number of samples in the original sample set in formula (9). And the generalization ability of the algorithm can be tested with the help of the unexcited data. At the same time, RF will split nodes to achieve feature extraction when constructing the decision tree, and select the optimal split feature to achieve the construction of the decision tree. Different subsets and their features are independent and random. Therefore, to some extent, the diversity and richness of data types are guaranteed [25].

4. EVALUATION OF THE CAPACITY OF SME FINANCING STRUCTURES

In this study, the online Lending Ciub platform model was used to analyze the empirical results, and the loan data of SMEs was selected to extract the characteristic variables. First, the downloaded data was preprocessed to reduce the volume of the downloaded data; then, the remaining data was preprocessed and the data set was normalized. The prediction accuracy of the characteristic variables proposed in the study were compared. Table 2 presents the results.

The IV values of the selected variable indicators are given in Table 2 above, and suggest that the indicators have good prediction ability. The adjustment of its indicator content will have different impacts on the financing ability of enterprises. It also tests the index system of financing structure capacity proposed by the study, and uses KMO

Table 3 Variable index test.

KMO Measure of sampling Adequacy	**/**	**0.657**
Bartlett sphericity test	Approximate chi square	12838.798
	df	5
	Significance	0.000

Table 4 Regression results for explanatory variable and explained variable.

Indicator variable	Coefficient	Std Error	t−Statistic	Prob.
C	−0.46665	0.109607	−4.170484	0.0000
Q1	0.03739	0.005005	7.913251	0.0000
Q2	0.108617	0.016942	−1.647305	0.0854
Q3	0.142866	0.029903	4.911272	0.0000
Q4	0.00364	0.000035	2.539926	0.0132
Q5	0.00432	−0.00029	0.03647	0.0284
Q6	0.087933	0.068603	−12.36964	0.2105
Q7	2.054848	0.126724	16.24859	0.0000
Q8	3.455761	0.231789	−14.83538	0.0000
Q9	−0.01262	0.000907	−6.895823	0.0000
Q10	1.491919	0.337558	−4.748317	0.0000
Q11	0.239193	0.058707	−3.437315	0.0002

Table 5 Results for Noise Robustness.

Number of experiments	Cross validation error			
	SVM	RF	Research algorithm	P value
1	0.071	0.108	0.003	0.000
2	0.215	0.208	0.031	0.000
3	0.256	0.274	0.028	0.003
4	0.202	0.212	0.012	0.041
5	0.125	0.139	0.004	0.031
6	0.038	0.071	0.036	0.009
7	0.207	0.237	0.021	0.003
8	0.328	0.262	0.081	0.000
9	0.094	0.104	0.043	0.010
10	0.143	0.192	0.022	0.007

(Kaiser-Meyer-Olkin) and sphericity Bartlett test to determine the correlation between different indicators. Table 3 shows the test results.

The sum of squares of the partial correlation coefficients of each variable is less than the value of the correlation coefficients as shown in Table 3. KMO value is 0.657, which is between (0, 1) , indicating that the correlation between different variables is strong and suitable for factor analysis. In Bartlett's sphericity test, the approximate chi-square and df values were 12838.798 and 5 respectively, and the significance of the index test was clear ($P = 0.000 < 0.05$) .

The tax effect of the indicators that affect the ability of enterprises' financing structure has a negative correlation effect as shown in Table 4, while it has a positive correlation with the indicators. The improvement of market competitiveness, guaranteed performance of asset value, operation ability and profitability, anti-risk ability and internal accumulation level, can help SMEs improve their financing ability. Then the performances of the RF models proposed in the study were compared. The results are shown in Table 5.

The error increment of the RF proposed in the study under the number of experiments is significantly smaller than the other two algorithms in Table 5. Its numerical difference is obvious, indicating that the algorithm is less affected by noise. The test performance of different algorithms was analyzed. Results are presented in Table 6.

The performance of different algorithms in the test process is different in Table 6. Specifically, the training time and test time of the support vector machine are basically more than 10s and more than 20s, and its average correct recognition rate is 87.69%. The minimum training time and test time of RF are 10.09s and 13.58s, respectively. Its overall time consumption is 5.40% and 13.29% higher than that of SVM algorithm, and the average accuracy rate of information recognition is 91.19%. The improved RF proposed in the study consumes less time. Its average time in training and testing is 18.15s and 20.51s, which is significantly less than that of SVM algorithm and RS. The average recognition accuracy is 95.50%, and the difference between the accuracy of the other two algorithms is 7.82% and 4.31%. The above results show that the proposed

Table 6 Test performance of different algorithms.

No. of experiments	SVM			RF			Research algorithm		
	Training time(s)	Test time(s)	Correct recognition rate (%)	Training time(s)	Test time(s)	Correct recognition rate (%)	Training time(s)	Test time(s)	Correct recognition rate (%)
1	13.25	20.51	86.58	10.09	13.58	90.1	5.8	8.16	94.41
2	18.27	25.53	88.14	15.11	18.6	91.52	10.82	13.18	95.83
3	26.44	33.7	87.69	23.28	26.77	91.21	18.99	21.35	95.52
4	29.12	36.38	84.46	25.96	29.45	87.98	21.67	24.03	92.29
5	26.54	33.8	85.71	23.38	26.87	89.23	19.09	21.45	93.54
6	28.36	35.62	88.27	25.2	28.69	91.79	20.91	23.27	96.1
7	23.53	30.79	86.71	20.37	23.86	90.23	16.08	18.44	94.54
8	29.66	36.92	90.66	26.5	29.99	94.18	22.21	24.57	98.49
9	30.24	37.5	90.82	27.08	30.57	94.32	22.79	25.15	98.63
10	31.09	38.35	87.85	27.93	31.42	91.37	23.16	25.52	95.68
Average value	25.65	32.91	87.68	22.49	25.98	91.19	18.152	20.51	95.50

Table 7 Results of 50% cross test of integrated algorithm (%).

Times	1	2	3	4	5	6	7	8	9	10
Accuracy	98.28	98.69	98.54	98.35	98.23	98.57	98.52	98.46	98.25	98.61
Times	11	12	13	14	15	16	17	18	19	20
Accuracy	98.64	98.57	98.27	98.36	98.25	98.56	98.37	98.38	97.12	98.57
Times	21	22	23	24	25	26	27	28	29	30
Accuracy	98.56	98.36	98.33	98.41	98.46	98.43	98.54	98.41	98.05	98.64
Times	31	32	33	34	35	36	37	38	39	40
Accuracy	98.45	98.49	98.26	98.67	98.17	98.27	98.27	98.28	98.26	98.23
Times	41	42	43	44	45	46	47	48	49	50
Accuracy	98.37	95.62	98.45	98.25	98.26	98.64	98.35	98.26	98.34	98.16
Mean value	98.32									
Variance	1.94E-06									

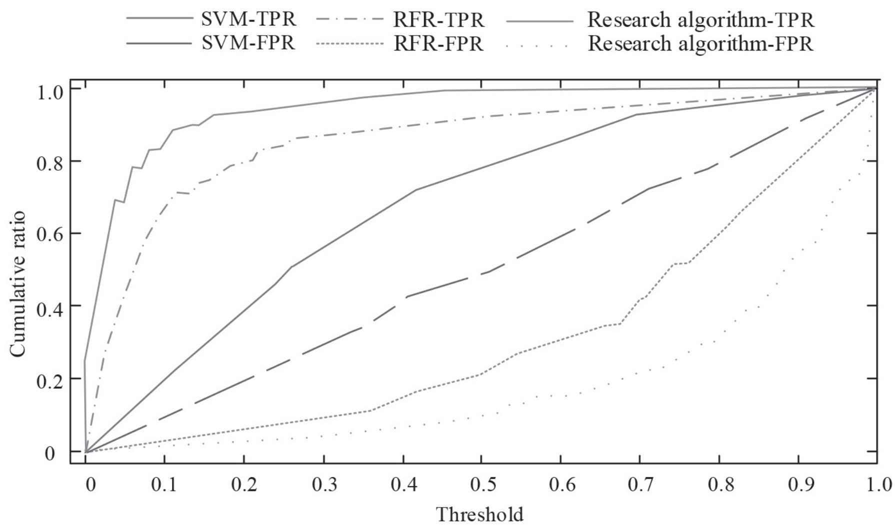

Figure 3 KS results for different algorithms.

algorithm has good performance. The identification result here is to identify the information that affects the enterprise's credit risk indicators. Then the integration algorithm proposed in the study was cross-checked. The results are presented in Table 7.

The cross-validation results show that the proposed method has high test accuracy. The average accuracy rate is 98.32%, and the variance is only 1.94E-06. The application performance of the data model is relatively stable. The risk prediction performance of the algorithm model is shown in Figure 3.

In Figure 3, TPR and FPR represent true rate and false positive rate. The KS value is the threshold value of the maximum distance between two curves. The larger the

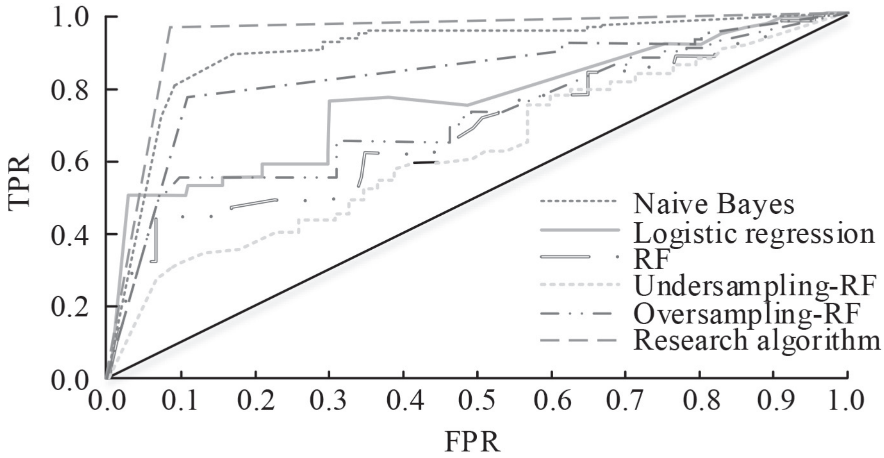

Figure 4 AUC area of different algorithms.

distance, the better the algorithm can distinguish between positive and negative samples. That is to say, the prediction of the sample in the study is better distinguishable. The results presented in Figure 3 show that KS values of SVM, RF and improved RFs are 34%, 67% and 88% respectively. The proposed integration algorithm can reduce prediction errors and achieve good classification results. To further evaluate the performance of the integration algorithm proposed in the study, it is compared with the naive Bayesian algorithm, the logical regression algorithm, as well as the under-sampling and oversampling processing RF. The under-sampling RF is used to reduce the imbalance of the data set by discarding the majority of the sample data. The oversampling RF refers to randomly copying a few samples to reduce the interference of the number of duplicate samples on the accuracy and fitting of the algorithm. Figure 4 shows the experimental comparison results.

As shown in Figure 4, the accuracy of recognition achieved by the various algorithms is different. According to the size of AUC area in the figure, the performance difference of different algorithms is obtained: research algorithm>naive Bayesian algorithm>logical regression algorithm>oversampling RF > RF > undersampling RF. The corresponding AUC area size is 0.915 > 0.866 > 0.817 > 0.779 > 0.534 > 0.487.

In addition, the accuracy of information evaluation of different algorithms is relatively high, and the difference between the integrated algorithm and RF that is proposed to improve the random algorithm is far more than 30%.

5. CONCLUSIONS

SMEs and their development account for a large proportion of the market economy. Inevitably, when an SME is expanding, it faces financing difficulties, and it will be difficult to establish a targeted structural system due to information asymmetry, among other reasons. Therefore, this study determined the factors that affect the financing structure of enterprises by applying integrated algorithms, and proposed a classification model to evaluate credit information. The IV values of the impact indicators are high, and the KMO of the test results for each variable is 0.657, and the indicator has a clear significance. The tax effect has a negative correlation with the financing capacity of enterprises. The effectiveness of the enterprise information classification evaluation model was tested. The information error results produced by the integrated algorithm were significantly smaller than those of the other algorithms being compared. The average time for training and testing was 18.15s and 20.51s, respectively, significantly less than that of the SVM algorithm and the RF. The average recognition accuracy was 95.50%, and the difference between the accuracy of the integrated algorithm and the other two algorithms was 7.82% and 4.31%, respectively. In addition, the average accuracy of the algorithm model proposed in the study in the cross test is 98.32%, and the variance is only 1.94E-06. The performance is relatively stable. Its KS value is 88%, which is higher than 34% and 67% of SVM and RF, and the sample differentiation effect is good. In terms of accurate recognition performance, the AUC area of different algorithms is research algorithm (0.915)> naive Bayesian algorithm (0.866)> logical regression algorithm (0.817)> oversampling RF (0.779) > RF(0.534) > undersampling RF (0.487). The above results show that the proposed method can effectively help SMEs to efficiently evaluate their own financing structure capacity. Enterprises can implement early warning decision analysis at different stages according to their specific conditions. Strengthening the sample size of the data volume and optimizing the training parameters of the model are two aspects that require further research in the future.

REFERENCES

1. Bhatti A, Rehman SU, Rumman JB A. Organizational capabilities mediates between organizational culture, entrepreneurial orientation, and organizational performance of SMEs in Pakistan. Entrepreneurial Business and Economics Review, 2020, 8(4): 85–103.
2. Isichei EE, Emmanuel Agbaeze K, Odiba M O. Entrepreneurial orientation and performance in SMEs: The mediating role of structural infrastructure capability. International Journal of Emerging Markets, 2020, 15(6): 1219–1241.
3. Rishehchi Fayyaz M, Rasouli MR, Amiri B. A data-driven and network-aware approach for credit risk prediction in supply chain finance. Industrial Management& Data Systems, 2021, 121(4): 785–808.

4. Liu J, Liu S, Li J, Li J. Financial credit risk assessment of online supply chain in construction industry with a hybrid model chain. International Journal of Intelligent Systems, 2022, 37(11): 8790–8813.
5. Uddin MS, Chi G Al Janabi M A M, Habib T Leveraging random forest in micro enterprises credit risk modelling for accuracy and interpretability. International Journal of Finance & Economics, 2022, 27(3): 3713–3729.
6. AlQershi N, Abas Z, Mokhtar S. The intervening effect of structural capital on the relationship between strategic innovation and manufacturing SMEs' performance in Yemen. Management Science Letters, 2021, 11(1): 21–30.
7. Van Klyton A, Rutabayiro-Ngoga S. SME finance and the construction of value in Rwanda. Journal of Small Business and Enterprise Development, 2018, 25(4): 628–643.
8. Paea S, Baird R. Information Architecture (IA): Using multi-dimensional scaling (MDS) and K-Means clustering algorithm for analysis of card sorting data. Journal ofUsability Studies, 2018, 13(3): 138–157.
9. Fang GG, Qalati SA, Ostic D, et al. Effects of entrepreneurial orientation, social media, and innovation capabilities on SME performance in emerging countries: a mediated-moderated model. Technology Analysis& Strategic Management, 2022, 34(11): 1326–1338.
10. Tang Z, Srivastava G Liu S. Swarm intelligence and ant colony optimization in accounting model choices. Journal of Intelligent& Fuzzy Systems, 2020, 38(3): 2415–2423.
11. Guo L, Chen J, Li S, et al. A blockchain and IoT-based lightweight fr amework for enabling information transparency in supply chain finance. Digital Communications and Networks, 2022, 8(4): 576–587.
12. Malakauskas A, Lakštutienė A. Financial distress prediction for small and medium enterprises using machine learning techniques. Engineering Economics, 2021, 32(1): 4–14.
13. Pratono A H. Does firm performance increase with risk-taking behavior under information technological turbulence? Empirical evidence from Indonesian SMEs. The Journal of Risk Finance, 2018, 19(4): 361–378.
14. Liu Y Huang L. Supply chain finance credit risk assessment using support vector machine based ensemble improved with noise elimination. International Journal of Distributed Sensor Networks, 2020, 16(1): 1550147720903631.
15. Drydakis N. Artificial Intelligence and reduced SMEs' business risks. A dynamic capabilities analysis during the COVID-19 pandemic. Information Systems Frontiers, 2022, 24(4): 1223–1247.
16. Malakauskas A, Lakštutienė A. Financial distress prediction for small and medium enterprises using machine learning techniques. Engineering Economics, 2021, 32(1): 4–14.
17. Song Y, Wu R. The impact of financial enterprises' excessive financialization risk assessment for risk control based on data mining and machine learning. Computational Economics, 2022, 60(4): 1245–1267.
18. Cao Y, Shao Y, Zhang H. Study on early warning of E-commerce enterprise financial risk based on deep learning algorithm. Electronic Commerce Research, 2022, 22(1): 21–36.
19. Yao L. Financial accounting intelligence management of internet of things enterprises based on data mining algorithm. Journal of Intelligent& Fuzzy Systems, 2019, 37(5): 5915–5923.
20. Papíková L, Papík M. Effects of classification, feature selection, and resampling methods on bankruptcy prediction of small and medium sized enterprises. Intelligent Systems in Accounting, Finance and Management, 2022, 29(4): 254–281.
21. Mao C. Enhanced group decision-making through an intelligent algorithmic approach for multiple-attribute credit evaluation with 2-tuple linguistic neutrosophic sets. International Journal of Knowledge-based and Intelligent Engineering Systems, 2024, 28 (1), 163–177.
22. Zhang X, Li W. A group decision framework for core competitiveness evaluation of small and medium-sized foreign trade enterprises under probabilistic linguistic term sets. International Journal of Knowledge-based and Intelligent Engineering Systems, 2024, 28 (1), 45–58.
23. Peng X, Huang H. Fuzzy decision-making method based on CoCoSo with critic for financial risk evaluation. Technological and Economic Development of Economy, 2020, 26(4): 695–724.
24. Srinivasan S, Kamalakannan T. Multi criteria decision making in financial risk management with a multi-objective genetic algorithm. Computational Economics, 2018, 52(2): 443–457.
25. Li J, Zhu S, Zhang W, et al. Blockchain-driven supply chain finance solution for small and medium enterprises. Frontiers of Engineering Management, 2020, 7(4): 500–511.