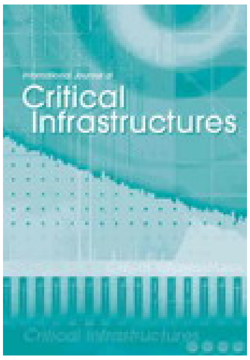


International Journal of Critical Infrastructures

ISSN online: 1741-8038 - ISSN print: 1475-3219
https://www.inderscience.com/ijcis

Design of manufacturing enterprise FEW system based on ML from the perspective of circular economy

Yang Huang, Xinyu Li, Dan Li

DOI: 10.1504/IJCIS.2027.10070396

Article History:

Received:	15 October 2024
Last revised:	10 February 2025
Accepted:	10 February 2025
Published online:	17 September 2025

Design of manufacturing enterprise FEW system based on ML from the perspective of circular economy

Yang Huang and Xinyu Li

Wuhan Maritime Communication Research Institute,
Wuhan 430205, China
Email: majorst@sina.com
Email: ximuwa@sina.com

Dan Li*

Faculty of Management,
Wuhan Donghu College,
Wuhan 430212, China
Email: lidan_403@163.com
*Corresponding author

Abstract: This study designs a financial early warning system for manufacturing enterprises, focusing on machine learning and the circular economy. The random forest model is used as the base model, optimised by the artificial jellyfish algorithm to enhance prediction accuracy. Financial and non-financial indicators are selected through significance testing and feature screening methods. The results show that the optimised model achieves the highest accuracy of 88.42% and AUC of 0.918. Key warning indicators include inventory turnover rate, accounts receivable turnover rate, Herfindahl index, and liquidity ratio. The study highlights the importance of timely warnings for maintaining financial stability in manufacturing enterprises, helping them manage financial crises and supporting sustainable growth. The proposed system provides valuable support for policymakers and industry leaders in managing financial risks and advancing circular economy goals.

Keywords: circular economy; CE; manufacturing enterprises; financial early warning; FEW; random forest; RF; artificial jellyfish algorithm.

Reference to this paper should be made as follows: Huang, Y., Li, X. and Li, D. (2025) 'Design of manufacturing enterprise FEW system based on ML from the perspective of circular economy', *Int. J. Critical Infrastructures*, Vol. 21, No. 9, pp.1–20.

Biographical notes: Yang Huang received her Master of Engineering degree from Huazhong University of Science and Technology, Hubei, China in 2014. Presently, he is working as a senior engineer in the Wuhan Maritime Communication Research Institute, Hubei, China. His areas of interest include information security, software engineering, and enterprise informatisation.

Xinyu Li obtained his Bachelor's in Computer Software from the Computer Science Department of Central China Normal University in Wuhan, China in 2001. During his work, he has accumulated rich experience in design and development as well as project management. Currently, his areas of interest include process management, digitalisation, machine learning, and information security.

Dan Li received her Master of Engineering degree from Huazhong University of Science and Technology, Hubei, China in 2014. Presently, she is working as an Associate Professor in the Faculty of Management, Wuhan Donghu College, Hubei, China. She has published articles in more than 20 internationally reputed peer reviewed journals and conference proceedings. Her areas of interest include financial management and accounting education.

1 Introduction

Presently, China is undergoing a transition from a rapid economy to a circular economy (CE), and the manufacturing enterprise, as the main driving force for economic development, needs to develop more stably. The occurrence of financial crises in enterprises will seriously affect their development and be detrimental to sustainable economic development. So designing a financial crisis system suitable for manufacturing enterprises, and predicting and avoiding financial crises promptly, has crucial influence on the development of the manufacturing enterprise and CE.

It is pivotal to note that the perfection of manufacturing infrastructure is related to the economic resilience of enterprises in response to financial crises. Some scholars have found that economic resilience is reflected in an enterprise's ability to maintain operation, recover quickly and adapt to the new environment in the face of external shocks (Cutrini and Ninivaggi, 2024). Manufacturing infrastructure includes production equipment modernisation level, supply chain resilience, digital technology application, and green technology investment, etc. Its optimisation can not only raise resource utilisation efficiency, but also indirectly reflect the stability of key financial indicators by enhancing production flexibility and risk management ability (Brada et al., 2021). Especially in the CE framework, the green transformation of infrastructure can further reduce the impact of environmental risks on financial health, thus providing a more robust database for financial early warning (FEW) systems. Therefore, the resilience of infrastructure construction is an important support for improving the FEW ability of manufacturing enterprises.

As the technology rapidly develops, machine learning (ML) has become one of the key tools to drive corporate sustainability. Some scholars have found that in the manufacturing industry, ML can help enterprises improve operational efficiency by accurately predicting and optimising resource allocation, so as to achieve a more sustainable production model (Sanusi et al., 2023). At the same time, the CE, as the core concept of economic sustainability, promotes the efficient use of resources and recycling, which provides modern enterprises with new ways to cope with environmental challenges and market pressures. Some studies believe that modern enterprises entering the second era no longer exclusively pursue short-term financial interests, but rather focus more on long-term development and environmental responsibility (Hailemariam and Erdiaw-Kwasie, 2023). In this context, the combination of ML and CE not only promotes the green transformation of the economy, but also brings stronger competitiveness and higher sustainability to enterprises.

The research on FEW mainly includes two aspects: the selection of financial indicators and the construction of FEW models. For the selection of financial indicators, currently the main focus is on selecting financial indicators, with limited selection of non-financial indicators and incomplete qualitative analysis. For the construction of FEW models, ML methods have the best performance at present, among which logical regression (LR) and random forest (RF) models have the best application effect. However, model performance is still unstable. Therefore, this study will focus on the characteristics of manufacturing enterprises, select experimental indicators from two points: financial and non-financial indicators, and use significance testing and packaging methods to screen experimental indicators. Meanwhile, the research will select the RF model as the experimental basis model to design the FEW system, and introduce the artificial jellyfish (AJ) algorithm to optimise the hyperparameter to increase the accuracy and stability of the model. This study is basically concentrated on the following four sections. The first section is a summary and discussion of existing enterprise FEW methods. The second section mainly introduces the selection method of financial indicators and the design of an FEW model. The third section is to conduct comparative experiments between the model in this study and other models, and to organise and analyse the effectiveness advantages of the model. The final section of the article provides a summary of its contents.

2 Related works

As the advancement of CE, it has become very important to provide timely FEW in manufacturing enterprises, and many scholars and scientists have also conducted relevant research. Padhan and Prabheesh (2019) proposed a crisis generative model to capture relevant dynamic information of enterprises to eliminate the problem of improper selection of FEW indicators and low efficiency of FEW models, integrate the historical data of financial crises, select the FEW indicators with strong relevance and high contribution rate, and prove the feasibility of this method through experiments. To select more comprehensive financial risk indicators, the team of Lu and Zhou (2021) proposed an indicator selection method based on fuzzy clustering algorithm. This method used parallel clustering algorithms to perform fuzzy classification, increased the correlation between indicators, and selected actual data of listed companies for experiments to assess the feasibility of this method. Zeng (2022) proposed a method for selecting FEW indicators based on cash flow to analyse the effect of the IoT on corporate financial risks. This method used a backpropagation neural network (BPNN) to process FEW indicator data, and introduced a moving edge algorithm to optimise the model, raise the timeliness of FEW, and experiment verified the feasibility of this method. Regin et al. (2023) raised a risk indicator selection model to address the issue of traditional indicator selection methods not taking into account the internal structure of enterprises. This model introduced BPNN to train algorithms through specific algorithms, established an FEW indicator system, and conducted experiments to assess the feasibility of the algorithm. Arora and Singh (2020), in response to the issue of incomplete bankruptcy prediction data and poor prediction performance, it chose financial ratios as the main prediction indicator and applied ML models for bankruptcy prediction. The experiment findings illustrated that the prediction accuracy of this method was high, and the prediction effect of RF model was the best. To improve the generalisation performance of student support vector machine models and avoid data imbalance, the Huang and Guo (2021) team introduced fuzzy algorithms to optimise the model, and designed a kernel fuzzy twin support vector machine model. Moreover, a comparative experiment was conducted between the model and relevant models, and the findings showed that the algorithm had higher performance and better overcame the imbalance problem. Filippopoulou et al. (2020) proposed a multivariate logistic regression model to address the low accuracy in existing bank FEW models, and used local bank macro databases as experiment data to construct the model. The model's feasibility was experimentally validated, yet the issue of incomplete prediction persisted. Gallagher et al. (2022) designed a blockchain-driven enterprise warning system to address the issue of dispersed small and medium-sized enterprises (SMEs) and low accuracy of FEWs. This system combined the advantages of blockchain and set pair analysis, and set up a distributed consensus mechanism to predict the financial risk situation of SMEs. Liu and Jiang (2021) proposed an improved BPNN model to provide FEW for a large number of enterprises, considering that modern enterprises did not attach importance to their initial financial problems, leading to subsequent severe crises. The improved model has been experimentally verified to have high accuracy. Ding (2021) proposed an intelligent FEW model grounded on fuzzy theory to address the heavy workload and low accuracy in traditional manufacturing prediction methods. At the same time, a valuation method combining AHP and fuzzy evaluation was introduced to solve the effectiveness of the model. Experimental verification demonstrated that this method significantly enhances prediction accuracy and reduces prediction loss.

In conclusion, a significant amount of research has been conducted by numerous scholars on the construction of FEW models. At present, the best performance method is ML method, in which LR and RF model have the best application effect. However, in the FEW of manufacturing enterprises, there are still problems such as low accuracy and unstable performance. Therefore, the research will select the RF model as the experimental basis model to design the FEW system, and introduce the AJ algorithm to optimise the hyperparameter to increase the accuracy and stability of the model.

3 Design of an FEW model for manufacturing enterprises based on ML

As the CE vigorously develops, timely warning and avoidance of financial crises in manufacturing enterprises, and reducing resource waste have become very important. Therefore, this research will design a suitable model to achieve the role of timely warning. The research will select the ML method based on RF to design an FEW model. At the same time, it will select appropriate FEW indicators for experiments according to the characteristics of manufacturing enterprises.

3.1 Selection of FEW Indicators for manufacturing enterprises

When conducting FEW for enterprises, the first step is to select appropriate warning indicators. The selection of FEW indicators must be based on comprehensiveness, industry specificity, predictability, and operability. Only when the FEW indicators are selected appropriately can the accuracy of the model's FEW be guaranteed (Li et al., 2023; Ding, 2021). The selection of indicators needs to consider various factors, such as whether it is in line with the actual situation and application environment, whether it is easy to obtain, and the degree of reasonableness of indicators. This study will integrate the financial management characteristics of manufacturing enterprises and industry development status, and select indicators from two aspects: financial and non-financial indicators. The specific indicator selection situation is expressed in Table 1.

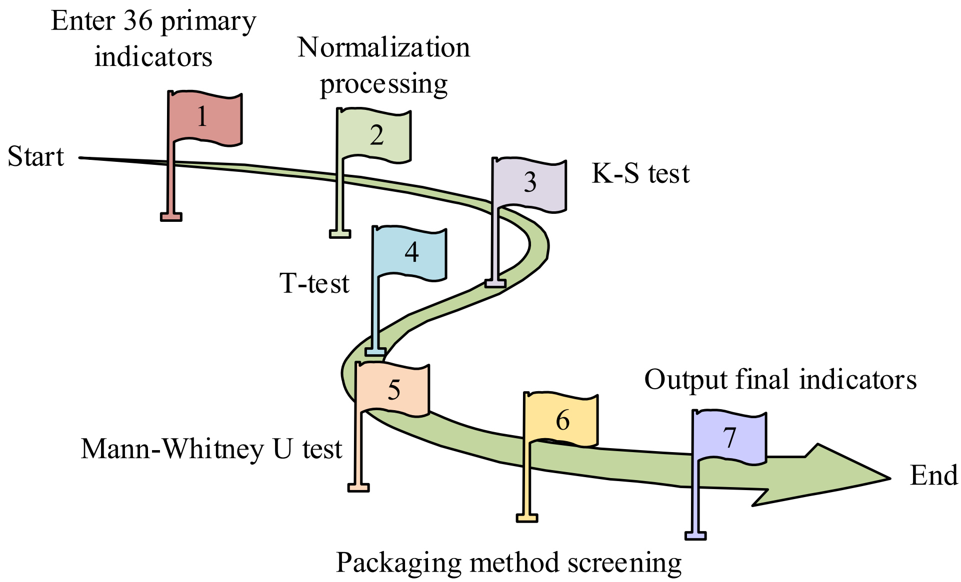
Table 1 Selection results of FEW indicators

Primary indicators	*Secondary indicators*	*Third level indicators*
Financial index	Solvency	Current, quick, asset liability, equity and interest coverage ratios
	Cash flow situation	Operating income cash ratio, total asset cash recovery rate, net profit cash content, net cash flow per share
	Profitability	Net interest rate of sales, net interest rate of total assets, net return on assets, earnings per share, return on assets
	Business Capability	Accounts receivable turnover ratio, inventory turnover ratio, current assets turnover ratio, total assets turnover ratio, guarantee multiple of pre-sale income and turnover rate of accounts payable
	Growth Capacity	Total assets growth rate, net assets growth rate, operating income growth rate, growth rate of operating profit, Net profit
Non-financial indicators	Ownership structure	Equity balance, Z-index, Herfindahl_ 5 Index, management shareholding ratio, and national stock ratio
	Corporate governance	Proportion of independent directors, concurrent positions of chairman and general manager, agency costs for managers, types of audit opinions
	Litigation and arbitration	Is there any major litigation or arbitration

As shown in Table 1, this study selects 36 indicators for the design of an FEW system. The overall indicators are divided into three levels. The first, second, and third level indicators are include 2, 8, and 36, respectively. In FEW indicators, the study starts from five aspects: profitability, solvency, growth ability, cash flow situation, and operating ability, and selects a total of 26 tertiary financial indicators (Wu et al., 2022). In non-financial indicators, considering the actual situation of manufacturing enterprises, a total of 10 tertiary indicators are chosen from three secondary indicators: litigation and arbitration, corporate governance, and equity structure. Among them, corporate governance indicators are important factors that affect corporate performance. Based on the characteristics of China's manufacturing industry, four indicators are selected: the proportion of independent directors, the cost of management agency, audit opinions, and the concurrent appointment of directors. At the same time, indicators such as whether there have been litigation and arbitration are selected to make the FEW indicators more comprehensive.

Due to the large number of 36 evaluation indicators during FEW and the high correlation of some indicators, there may be issues such as data redundancy and reduced efficiency during warning testing. The study will screen the initially selected indicators to obtain the best feature data as the final indicator. The specific screening method is denoted in Figure 1.

Figure 1 Indicator screening flowchart (see online version for colours)



As shown in Figure 1, it first normalises the data of the indicators to avoid misclassification due to inconsistent feature units of the indicators; secondly, the significance test is carried out on indicators to solve the problem of repeated information interaction between indicators. The significance test includes Kolmogorov Smirnov (K-S), student's T (T), and Mann-Whitney U tests (Li and Zhao, 2022; Leohang et al., 2023; Song and Wu, 2022), that is, K-S test is used to determine whether the characteristics conform to the normal distribution, and then T test is conducted on the indicators to determine the sample source. Then, it performs Mann-Whitney U test on the non compliant samples to judge if there is a significant difference in the samples; finally, the packaging method is used to screen the inspected indicators and select the best feature set. The mathematical expression for normalisation treatment is displayed in equation (1).

$$y = (x - x\_\min) / (x\_\max - x\_\min) \tag{1}$$

As shown in equation (1), y means the normalised data; x indicates the input feature data; x_min and x_max express the mini and max values of the input data. The mathematical expression formula of K-S test is shown in equation (2), that is, to test whether the sample distribution $f(x)$ conforms to the equation of normal distribution $g(x)$, assuming that

H1 $f(x)$ does not conform to normal distribution

H0 $f(x)$ conforms to normal distribution.

$$D = \max\left|f(x) - g(x)\right| \tag{2}$$

As shown in equation 2, D infers to the test statistic; $f(x)$ stands for the test sample distribution function; $g(x)$ denotes the normal distribution function. The mathematical expression for T-test is denoted in equation (3).

$$T = (X - \mu) / \left[s / \sqrt{n} \right] \tag{3}$$

In equation (3), X means the mean of the feature samples; T indicates the statistic; μ expresses the population mean; s means the standard deviation of the sample; n is the total capacity of the sample. The confusion matrix is constructed as shown in Table 2.

Table 2 Confusion matrix

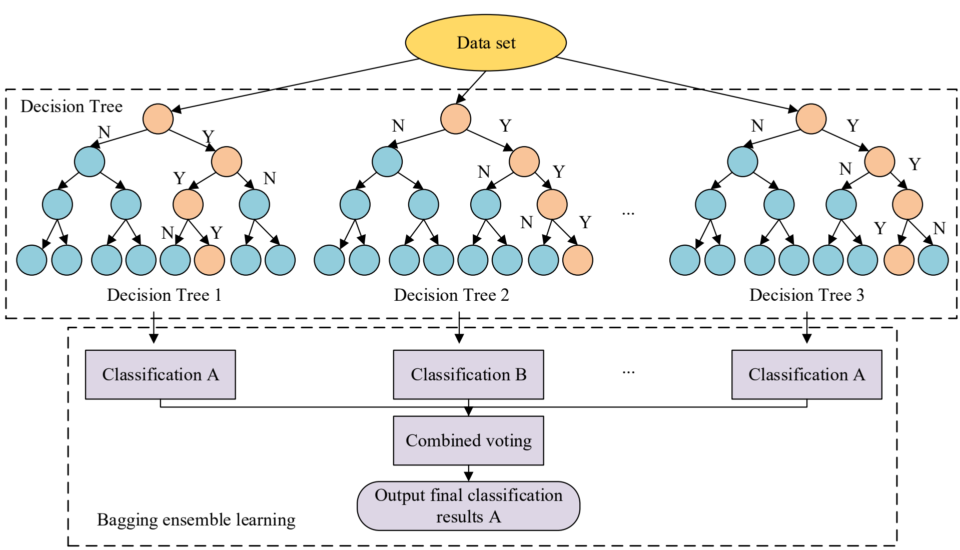
\		*Predicted data*	
		ST(1)	*Non ST(0)*
Real data	ST(1)	TD	FN
	Non ST(0)	FD	TN

As the FEW indicators of manufacturing enterprises are complex and diverse, it is not convincing to use simple accuracy to evaluate the FEW model. Therefore, the study introduces the confusion matrix, as well as the accuracy, recall, and area under the curve (AUC) values to assess the experimental results.

3.2 Design of FEW model based on improved RF algorithm

This research will optimise the FEW model based on RF (Yao and Qin, 2021). That is, the jellyfish search (JS) algorithm is utilised to find the optimal hyperparameter of RF. The hyperparameter of RF includes n_estimators and max_depth of decision trees for the maximum depth of a single decision tree. Using the optimal parameters to build an RF model will make the model performance higher (Guo et al., 2022). The RF algorithm structure is shown in Figure 2.

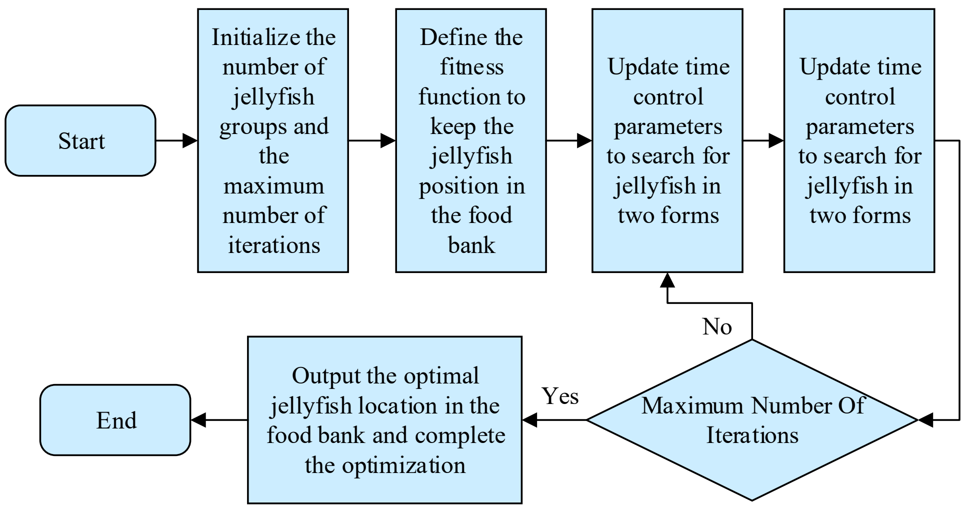
Figure 2 RF algorithm structure (see online version for colours)



In Figure 2, the RF includes the decision tree and bagging integration module. The algorithm mainly performs random sampling classification calculation on the original data features through multiple decision trees, and then uses the BAGAGING integration module to vote the mode of the output results of the decision tree to obtain the final classification outcome. Among them, the decision tree is the basic learner of the RF algorithm, which is used to determine the features. The bagging integration module uses the self-sampling method to train each sample, estimate the sample generalisation error, and vote the classification results of each basic learner to get the final result (Masood et al., 2023). Among them, the decision tree is constructed using the minimum Gini coefficient, and the mathematical calculation for Gini coefficient is shown in equation (4).

$$Gini(p) = 1 - sum\_\{k=1\}\hat{}\{K\}p\_k\hat{}2 \tag{4}$$

Figure 3 AJ search algorithm (see online version for colours)



As shown in equation (4), *Gini* means the Gini coefficient; *p* stands for probability; *K* refers to the number of decision tree classifications; *p_k* represents the probability of the k^{th} classification. A smaller Gini coefficient indicates better classification performance.

The AJ search algorithm simulates jellyfish searching for the optimal location and introduces a time control mechanism to set the algorithm to find the optimal parameters. The JS algorithm is shown in Figure 3.

In Figure 3, the factors affecting the optimal location of jellyfish include current direction, food quantity, fitness function, and time control parameters. The relationship between the direction of ocean currents and the optimal position of jellyfish is shown in equation (5).

$$\begin{cases} \overrightarrow{trend} = \sum \overrightarrow{trend} / n_{pop} \\ \overrightarrow{trend} = X^* - df \\ df = e_c \mu \end{cases} \tag{5}$$

As shown in equation (5), $\overrightarrow{trend}$ refers to the direction of the ocean current; X^* means the optimal jellyfish position; μ indicates the average position of all jellyfish; df denotes the difference between the optimal and average positions; n_{pop} means the size of the jellyfish population, and e_c is the attraction factor. The expression for the attractiveness factor is shown in equation (6).

$$e_c = \beta \times rand(0, 1) \tag{6}$$

As shown in equation (6), β means the distribution coefficient of jellyfish, and $rand(0, 1)$ indicates a random number between 0 and 1. Assuming that the jellyfish position conforms to the normal distribution, the mathematical expression of the optimal position after introducing the jellyfish standard deviation is shown in equation (7).

$$df = \beta \times rand(0, 1)\mu \tag{7}$$

The mathematical expression for the position of jellyfish can be obtained by synthesising equations (4))–(7), as shown in equation (8).

$$\overrightarrow{trend} = X^* - \beta \times rand(0, 1)\mu \tag{8}$$

The expression for updating the position of each jellyfish can be further derived from equation (8), as shown in equation (9).

$$\begin{cases} X_i(t+1) = X_i(t) + rand(0, 1) \times \overrightarrow{trend} \\ X_i(t+1) = X_i(t) + rand(0, 1) \times (X^* - \beta \times rand(0, 1)\mu) \end{cases} \tag{9}$$

In equation (9), X_i means the position of the i^{th} jellyfish, and t indicates the time. Due to the fact that jellyfish motion includes both passive and active motion, the passive motion position update expression is shown in equation (10) after adding relevant parameters to the position expression.

$$X_i(t+1) = X_i(t) + \gamma \times rand(0, 1) \times \left(U_b - L_b\right) \tag{10}$$

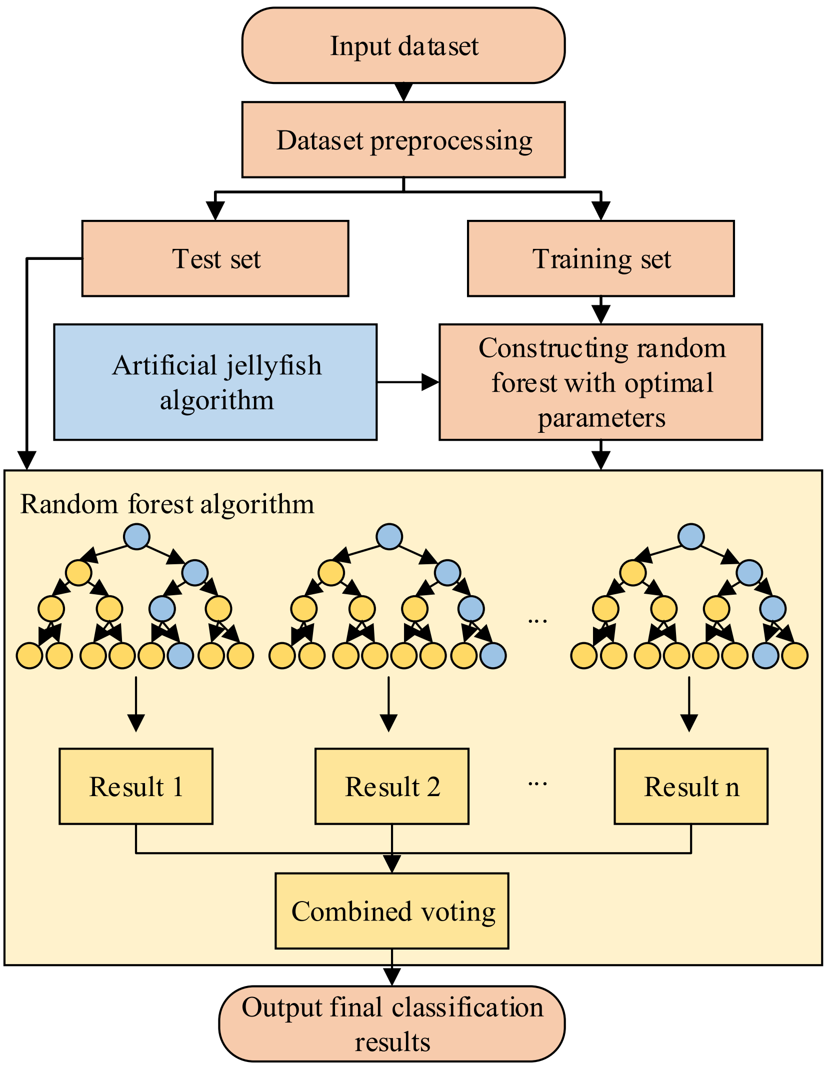
As shown in equation (10), L_b and U_b indicate the mini and max values of the JS space, with γ being the motion coefficient and $\gamma > 0$. The expression for updating the position of active motion is denoted in equation (11).

$$X_i(t+1) = X_i(t) + \overrightarrow{Step} \tag{11}$$

As shown in equation (11), $\overrightarrow{Step}$ means the displacement of jellyfish *i*, and $\overrightarrow{Direction}$ denotes the direction of movement of jellyfish *i*. The specific calculation for $\overrightarrow{Step}$ and $\overrightarrow{Direction}$ are shown in equation (12).

$$\begin{aligned} &\overrightarrow{Step} = rand(0,1) \times \overrightarrow{Direction} \\ &\overrightarrow{Direction} = \begin{cases} X_j(t) - X_i(t), \; f(X_i) \geq f(X_j) \\ X_i(t) - X_j(t), \; f(X_i) \leq f(X_j) \end{cases} \end{aligned} \tag{12}$$

Figure 4 RF-JS model flowchart (see online version for colours)



As shown in equation (12), i and j represent two randomly selected jellyfish, and f is the objective function of position X. The jellyfish algorithm includes a time control mechanism, and the time control function is shown in equation (13).

$$c(t) = \left| \left(1 - \frac{t}{Max_{iter}}\right) \times (2 \times rand(0,1) - 1) \right| \tag{13}$$

As shown in equation (13), $c(t)$ denotes the time control function; t means the number of iterations; Max_{iter} stands for the max amount of iterations. To raise the diversity of the population, the algorithm also introduces a logistic map, whose mathematical expression formula is shown in equation (14).

$$X_{i+1} = \eta X_i (1 - X_i),\ 0 \le X_0 \le 1 \tag{14}$$

In equation (14), X_i means the mapping value of the location of the i^{th} jellyfish; X_0 stands for the mapping value of the initial population, where $X_0 \in (0, 1)$, $X_0 \notin \{0.0, 0.25, 0.5, 0.75, 1.0\}$ and mean chaotic parameters.

$$\begin{cases} X'_{i,d} = \left(X_{i,d} - U_{b,d}\right) + L_b(d), X_{i,d} > U_{b,d} \\ X'_{i,d} = \left(X_{i,d} - L_{b,d}\right) + U_b(d), X_{i,d} < U_{b,d} \end{cases} \tag{15}$$

As shown in equation (15), $X_{i,d}$ denotes the position of the i^{th} jellyfish in the space with dimension d; $X_{i,d}'$ stands for the updated position after checking the boundary conditions; $U_{b,d}$ and $L_{b,d}$ represent the maxi and mini values of the jellyfish in the search space d dimension, respectively.

In this study, the AJ algorithm is applied to RF, and the RF-JS model is constructed to achieve timely FEW for manufacturing enterprises. The flow of the RF-JS model is denoted in Figure 4.

In Figure 4, the general process of the RF-JS model is to first pre-process the original dataset and partition the dataset; Secondly, the training set data are optimised by AJ algorithm, and the optimal hyperparameter extracted from the algorithm are applied to the RF algorithm; Finally, the decision tree is classified and Bagging ensemble is used to obtain the optimal result through voting.

4 Analysis of FEW results for manufacturing enterprises

The main content of the experimental results analysis section is to set experimental parameters to construct an experimental model, and to contrast the effect of the RF-JS model through comparative experiments. The RF-JS model is then applied to actual FEW of manufacturing enterprises for empirical analysis.

4.1 Data source and selection results of FEW Indicators

This study chose MATLAB7.0 as the construction platform for the proposed model. The experiment data was sourced from the public data of Guotai An database. The selected experimental data was the manufacturing company data of A-share listed sectors in two cities of a certain province from 2018 to 2021. 250 companies with similar numbers of warning companies (ST) and non ST were selected as the original dataset for the experiment, including 134 ST samples and 116 non ST samples (Ejegwa and Agbetayo, 2023). The selection results of FEW indicators for manufacturing enterprises are expressed in Table 3.

After screening 36 FEW indicators using K-S test, T-test, Mann-Whitney U test, and packaging method as shown in Section 2.1, 20 experimental indicators were obtained as shown in Table 3. The study used these 20 indicators as the experimental basis for model testing and specific FEW experiments for manufacturing enterprises. At the same time, from Table 3, non-financial factors also had a high impact on FEW in manufacturing enterprises, six non-financial indicators remained after screening. The RF-JS model selected the optimal hyperparameter of RF selected by the AJ algorithm. The selection results of n_Estimators are denoted in Figure 5, where Figure 5(a) shows the influence of decision tree quantity on recall rate, Figure 5(b) represents the impact of the number of decision trees on the AUC.

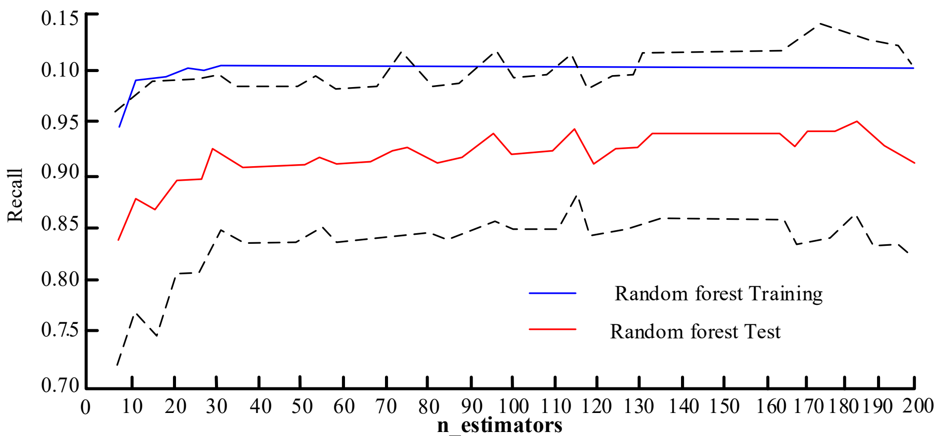
Table 3 Selection results of FEW indicators for manufacturing enterprises

\	*Variable*	*Indicators*	*Variable*	*Indicators*
Financial index	X2	Inventory turnover rate (ITR)	X12	Return on assets
	X3	Fixed asset turnover	X13	Current ratio
	X5	Accounts receivable turnover rate (ARTR)	X15	Asset liability ratio
	X6	Pre-sale income guarantee multiple	X16	Equity ratio
	X8	Net profit from pre-sale	X23	Interest coverage ratio
	X9	Net profit from total assets	X24	Growth rate of net assets
	X10	Roe	X25	Total Assets Growth Rate
Non-financial indicators	X27	Equity balance	X30	National equity ratio
	X28	Z-index	X31	Independent director ratio
	X29	Herfindahl_ 5 Index	X34	Manager agency costs

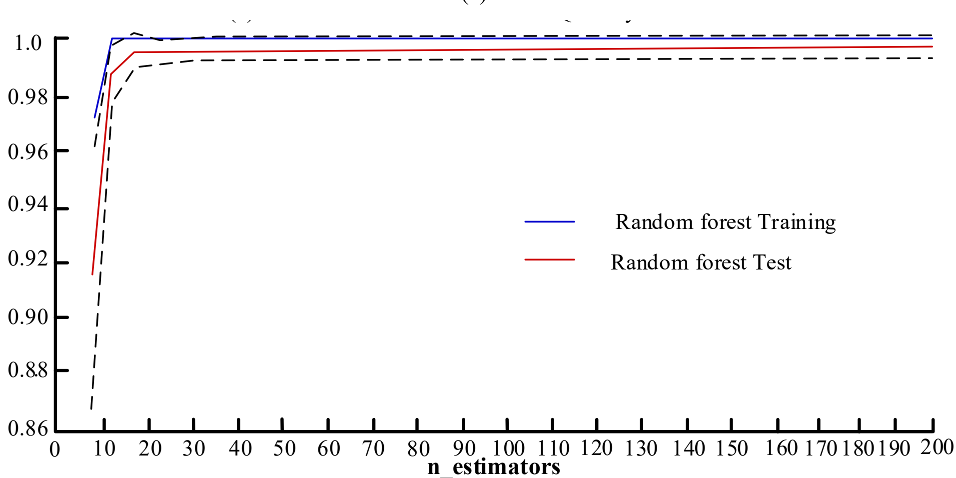
In Figure 5, the solid lines in the figure represent the recall rate and AUC values of the training and testing sets of the RF model, while the dashed lines represent the values of the recall rate and AUC after adding or subtracting the standard deviation. As shown in Figure 5(a), the recall rate of the model tended to stabilise at around 140 to 160 decision trees, and fluctuated significantly under other trees; In Figure 5(b), the AUC of the model reached a stable state when the number of decision trees was 30, indicating that the decision tree had little impact on the AUC value. So, the parameter space of the n_estimators selected in the study was (140–160). To ensure the simplicity and adaptability of the experiment, the values for n_estimators ultimately selected were 140. The selection of max_depth is shown in Figure 6.

Figure 6 is the statistical chart of the decision tree depth of the RF model. As shown in Figure 6, the minimum depth of the decision tree was 4, and the maximum depth was 13, that is, the parameter space range was (4, 13). From Figure 6, the minimum proportion of decision trees with depths of 4 and 12 was 2.5%, while the maximum proportion of decision trees with depths of 9 was 19%. Therefore, considering all factors, the value of max_depth selected for this experiment was 9.

Figure 5 The impact of n_estimators on recall rate (a) the influence of decision tree quantity on recall rate (b) the impact of the number of decision trees on the AUC (see online version for colours)

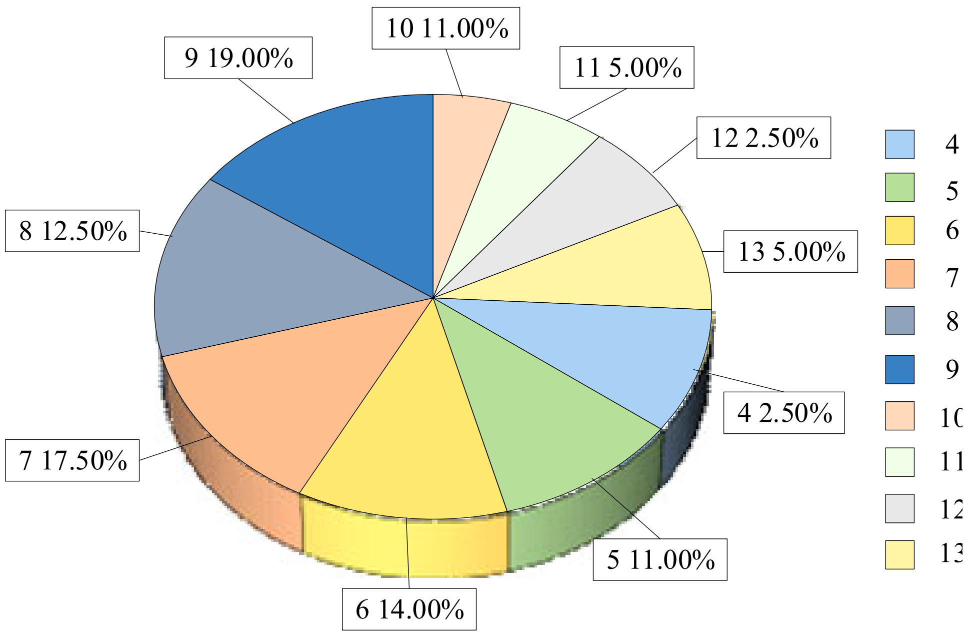


(a)



(b)

Figure 6 Statistics of depth distribution of decision tree (see online version for colours)

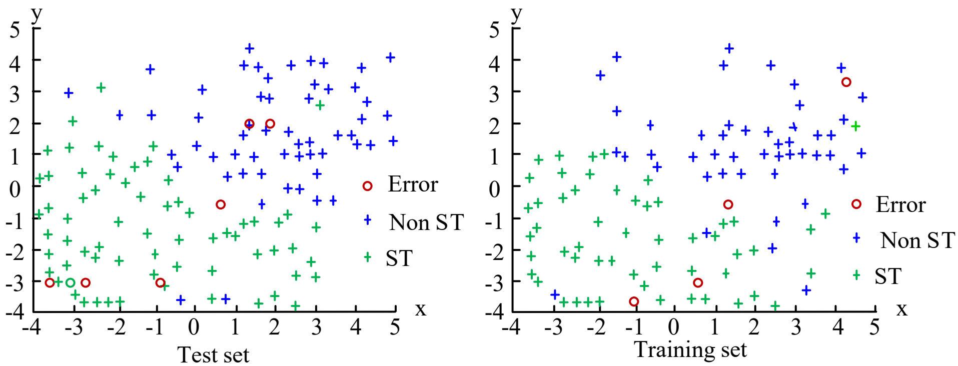


4.2 Analysis of FEW results for manufacturing enterprises

It constructed an RF-JS model for the FEW indicators of 20 manufacturing enterprises selected in the study. 250 manufacturing enterprises selected in Section 3.1 were selected as the research subjects for experiments, with 40% of them as the test set, no AJ algorithm optimisation used, and 60% as the training set for experiments. The laboratory findings demonstrated that the ST classification accuracy of the training set was 92.3%, while the ST classification accuracy of the test set was 88.5%. After visualising the experimental results, it is shown in Figure 7.

As shown in Figure 7, the classification performance of the training and testing sets was relatively good, and there were fewer cases of classification errors. Among the 150 enterprises in the training set, only 4 were misclassified, while in the 100 test sets, 6 enterprises were misclassified, indicating that the accuracy of the model has been improved through AJ algorithm, and the effect of FEW was better. To further assess the function of the model, the experiment carried out experiments on the same dataset with the RF-JS model, the unimproved RF model and the LR model, and introduced the accuracy and recall rates to assess the performance of the model. The experimental outcomes are denoted in Figure 8, where Figure 8(a) represents the accuracy comparison chart and Figure 8(b) represents the comparison chart of recall rate.

Figure 7 Model evaluation results (see online version for colours)



In Figure 8(a), the highest accuracy of the RF-JS model was 88.42%, and it reached a stable state as quickly as possible, reaching a stable state around 50 iterations, indicating that the model had the highest accuracy and best performance. As shown in Figure 8(b), the lowest recall rate of RF was 41.88%, while the recall rate of RF-JS model was the highest and the iteration speed was also the fastest. The recall rate and accuracy rate were 88.42%, which proved that the model had good index data classification balance and good financial crisis prediction effect. To sum up, the overall performance of the RF-JS model was the best, the LR model was good, and the prediction effect of the RF model without optimisation was low. At the same time, to further evaluate the stability of the FEW model, receiver operating characteristic (ROC) and AUC values were introduced for analysis. The ROC curve of the model is expressed in Figure 9.

In Figure 9, the maximum AUC of the RF-JS model reached 0.918, the minimum AUC value of the unimproved RF model was 0.683, and the value of the LR algorithm model was 0.683. At the same time, from the ROC curve in the figure, the RF-JS model had the largest area, smooth curve and the most stable performance, and the RF model had the smallest area, but the curve was also relatively smooth, that is, the accuracy rate was low but the performance was stable, while the curve of the LR algorithm had obvious differences, that is, the stability of the algorithm was not high, and the accuracy rate was good. Overall, the AUC of the RF-JS model was the highest and there was no significant difference, indicating that the algorithm performed the best in FEW for manufacturing enterprises. To better clarify the actual FEW situation of manufacturing enterprises and understand which indicators are key indicators for FEW of manufacturing enterprises, the experiment analysed the contribution of indicators by detecting the warning results of enterprises that have been warned through a model. The comparison of the contribution of the 20 warning indicators choose in this study is displayed in Figure 10.

Figure 8 Comparison chart of model accuracy and recall rate (a) accuracy comparison chart (b) comparison chart of recall rate (see online version for colours)

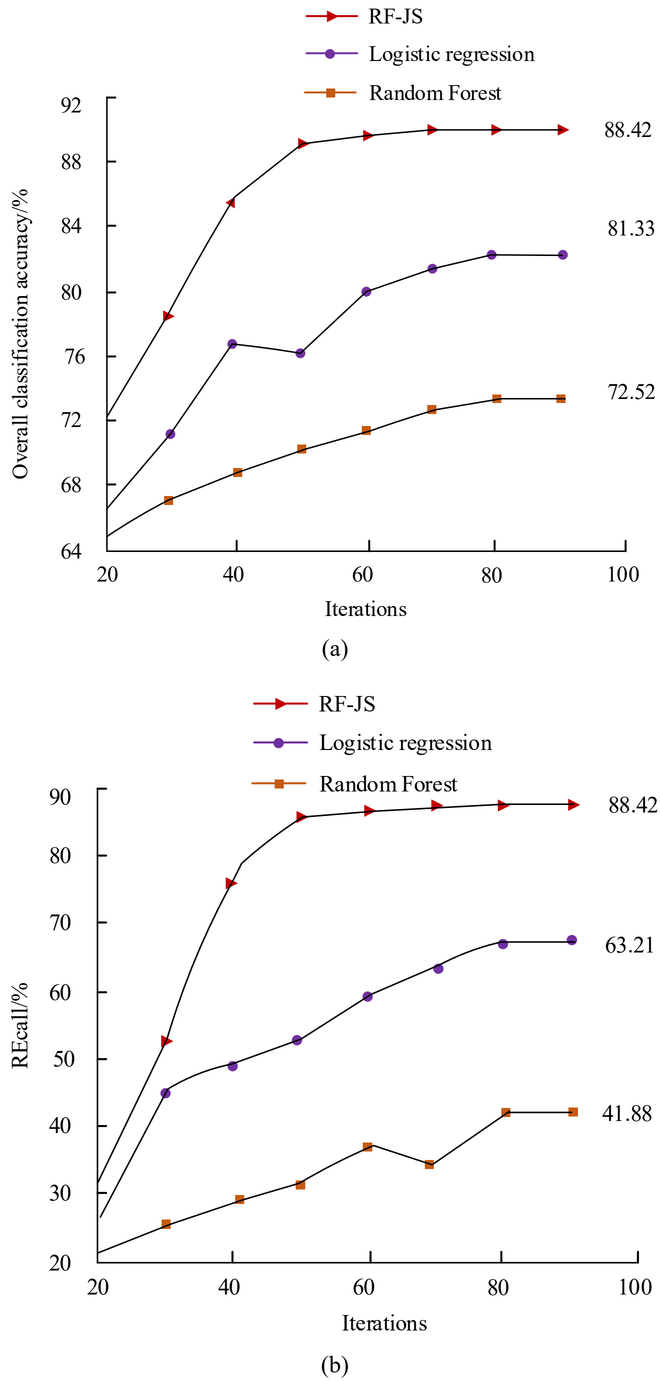


Figure 9 ROC curve comparison chart (see online version for colours)

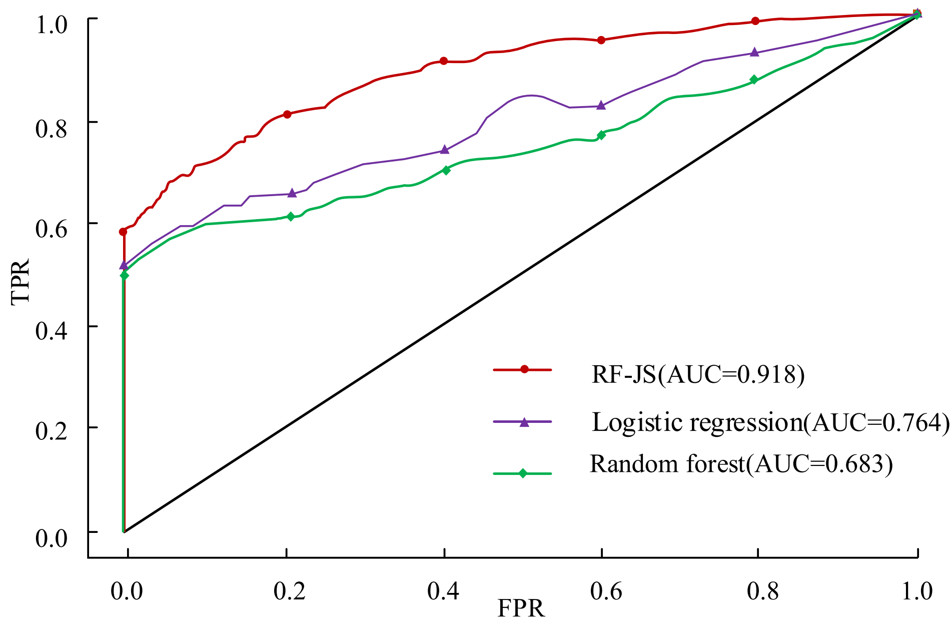
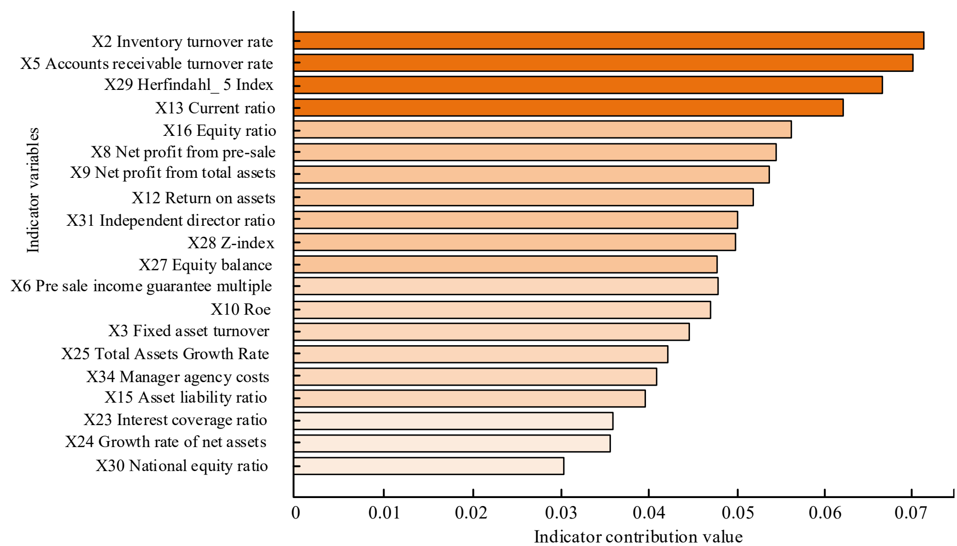


Figure 10 Comparison of contribution of few indicators (see online version for colours)



As shown in Figure 10, among the FEW indicators of the manufacturing enterprise, there were 3 warning indicators with a contribution rate below 0.04, with a relatively low contribution rate. There were 13 indicators with contribution values between 0.04 and 0.06, while inventory, and ARTRs. The warning contribution rate of the Herfindahl_5 index and liquidity ratio was the highest, above 0.06, indicating that these four indicators were key indicators for the occurrence of financial crises in enterprises. Each manufacturing enterprise should prepare contingency plans for these four aspects.

The contribution of the research is not only to raise the ability of manufacturing companies to warn against financial crises, but also to support managers, policy makers and the manufacturing industry to make progress towards achieving the sustainable development goals. Specifically, the study strengthens the economic backbone of the CE by enhancing the economic resilience of critical manufacturing infrastructure to financial crises by optimising FEW systems in the manufacturing sector, in line with several sustainable development goals. For example, decent work and economic growth emphasise inclusive and sustainable economic growth, and explore ways to support the long-term development of the manufacturing sector by reducing financial risks and improving business stability. Industry, innovation and infrastructure emphasise on enhancing the adaptability of infrastructure, and research on optimised FEW models can help manufacturing enterprises remain resilient under financial pressure, and promote smart manufacturing and industrial upgrading. Sustainable cities and communities encourage the sustainability of urban infrastructure and industrial systems, while financially sound manufacturing companies can support urban economic development in terms of supply chains and job security. Responsible consumption and production promote sustainable production patterns, and research helps manufacturing enterprises use resources more efficiently and reduce waste and losses caused by economic fluctuations by combining CE concepts. Therefore, the results of the research not only have practical significance for the financial health of manufacturing enterprises, but also provide strong support for the realisation of sustainable development goals.

5 Conclusions and future research

The development of CE has brought huge challenges to manufacturing enterprises, and designing an FEW system for manufacturing enterprises is very important. This study designed an FEW model based on the method of RF, and used AJ algorithm to optimise the hyperparameters of the RF algorithm. Firstly, the experiment used K-S test, T-test, Mann-Whitney U test, and packaging method to select 20 indicators from 36 original indicators as the manufacturing industry FEW indicators for this study. Secondly, the experiment utilised an AJ search algorithm to optimise the hyperparameters of the RF, and the outcomes demonstrated that the optimised hyperparameters of the RF were n_Estimators = 140, Max_ Depth = 9. Finally, 250 manufacturing enterprise data were selected as the experimental raw data, and evaluation indicators were introduced to test the FEW performance of the manufacturing industry using the RF-JS model, the unimproved RF model, and the logistic regression model. The research findings expressed that the accuracy of the RF-JS model was the highest at 88.42%, and the recall rate and iteration speed were also the fastest. The data balance and accuracy were consistent at 88.42%, and the AUC value reached a maximum of 0.918. There was no significant difference. Overall, the RF-JS model had the best effect on FEW for manufacturing enterprises. Meanwhile, the study conducted indicator contribution analysis on enterprises that were warned, and found that ITR, ARTR, Herfindahl_ 5 index and liquidity ratio were key indicators for the occurrence of financial crises in enterprises.

However, the study still has some limitations. First, despite the introduction of ML methods to optimise the FEW system, the applicability of the model still depends on the quality of the data, especially the distribution of data in different regions or industries may affect the generalisation ability of the model. Second, because the data source is mainly based on listed manufacturing enterprises, it does not adequately cover SMEs, which may affect the applicability of the study results. Finally, although this study combines financial and non-financial indicators, there are still shortcomings in the consideration of factors such as environment, policy impact and market dynamics, which may have an important impact on the financial stability of manufacturing firms. Future research can be expanded from the following aspects: First, the ML model can be further optimised, such as combining deep learning or reinforcement learning technology, to improve the adaptability and generalisation ability of FEW system. Second, the study can be extended to more manufacturing sub-sectors as well as SMEs to verify the applicability of the model. Third, more external variables, such as policy changes, global supply chain fluctuations, and environmental regulations, could be introduced to build a more comprehensive FEW. In addition, combining cutting-edge technologies such as blockchain and big data analysis to enhance the financial transparency of manufacturing enterprises and the real-time warning system is also a direction worth exploring.

References

Arora, I. and Singh, N. (2020) 'Prediction of corporate bankruptcy using financial ratios and news', *International Journal of Engineering and Management Research*, Vol. 10, No. 5, pp.82–87.

Brada, J.C., Gajewski, P. and Kutan, A.M. (2021) 'Economic resiliency and recovery, lessons from the financial crisis for the COVID-19 pandemic: a regional perspective from Central and Eastern Europe', *International Review of Financial Analysis*, Vol. 74, p.101658.

Cutrini, E. and Ninivaggi, F. (2024) 'Pandemic crisis and firm survival: evidence from the Italian manufacturing industry', *Journal of Industrial and Business Economics*, Vol. 52, No. 2, pp.1–33.

Ding, Q. (2021) 'Risk early warning management and intelligent real-time system of financial enterprises based on fuzzy theory', *Journal of Intelligent and Fuzzy Systems*, Vol. 40, No. 4, pp.6017–6027.

Ejegwa, P.A. and Agbetayo, J.M. (2023) 'Similarity-distance decision-making technique and its applications via intuitionistic fuzzy pairs', *Journal of Computational and Cognitive Engineering*, Vol. 2, No. 1, pp.68–74.

Filippopoulou, C., Galariotis, E. and Spyrou, S. (2020) 'An early warning system for predicting systemic banking crises in the Eurozone: a logit regression approach', *Journal of Economic Behavior and Organization*, Vol. 172, pp.344–363.

Gallagher, P., Doherty, T. and Stephens, S. (2022) 'Early warning systems for small business: insights from across Europe', *Journal of Enterprising Culture*, Vol. 30, No. 4, pp.453–470.

Guo, Y., Mustafaoglu, Z. and Koundal, D. (2022) 'Spam detection using bidirectional transformers and machine learning classifier algorithms', *Journal of Computational and Cognitive Engineering*, Vol. 2, No. 1, pp.5–9.

Hailemariam, A. and Erdiaw-Kwasie, M.O. (2023) 'Towards a circular economy: Implications for emission reduction and environmental sustainability', *Business Strategy and the Environment*, Vol. 32, No. 4, pp.1951–1965.

Huang, X. and Guo, F. (2021) 'A kernel fuzzy twin SVM model for early warning systems of extreme financial risks', *International Journal of Finance and Economics*, Vol. 26, No. 1, pp.1459–1468.

Leohang, D.A.K., Morasa, J. and Pinatik, S. (2023) 'Evaluation of credit sales and non-performing loans (NPL) of PT. Bank Sulutgo Ranotana sub-branch office', *Formosa Journal of Multidisciplinary Research*, Vol. 2, No. 1, pp.177–190.

Li, N. and Zhao, X. (2022) 'Credit risk evaluation of SME Loans based on logistic regression model', *Frontiers in Economics and Management*, Vol. 3, No. 10, pp.216–222.

Li, X., Wang, J. and Yang, C. (2023) 'Risk prediction in financial management of listed companies based on optimized BP neural network under digital economy', *Neural Computing and Applications*, Vol. 35, No. 3, pp.2045–2058.

Liu, J. and Jiang, P. (2021) 'A blockchain-driven cyber-credit evaluation approach for establishing reliable cooperation among unauthentic. MSMEs in social manufacturing', *Industrial Management and Data Systems*, Vol. 121, No. 4, pp.724–749.

Lu, S.H.D. and Zhou, Q. (2021) 'Iot big data based on fuzzy association rules mining enterprise financial risk early warning', *Journal of Neural Computing and Applications*, Vol. 33, pp.3901–3909.

Masood, F., Masood, J., Zahir, H. et al. (2023) 'Novel approach to evaluate classification algorithms and feature selection filter algorithms using medical data', *Journal of Computational and Cognitive Engineering*, Vol. 2, No. 1, pp.57–67.

Padhan, R. and Prabheesh, K.P. (2019) 'The effectiveness of the early warning model: a critical review and future direction of the new agenda', *Economic notice Dan Perbankan*, Vol. 22, No. 4, pp.457–484.

Regin, R., Rajest, S.S. and Shynu, T.A. (2023) 'Review of secure neural networks and big data mining applications in financial risk assessment', *Central Asian Journal of Innovations on Tourism Management and Finance*, Vol. 4, No. 2, pp.73–90.

Sanusi, I.T., Oyelere, S.S., Vartiainen, H., Suhonen, J. and Tukiainen, M. (2023) 'A systematic review of teaching and learning machine learning in K-12 education', *Education and Information Technologies*, Vol. 28, No. 5, pp.5967–5997.

Song, Y. and Wu, R. (2022) 'The impact of financial enterprises' excessive financialization risk assessment for risk control based on data mining and machine learning', *Computational Economics*, Vol. 60, No. 4, pp.1245–1267.

Wu, W., Ni, Z., Jin, F. et al. (2022) 'Decision support model with Pythagorean fuzzy preference relations and its application in financial early warnings', *Complex and Intelligent Systems*, Vol. 8, No. 1, pp.443–466.

Yao, F. and Qin, Z. (2021) 'Block chain based supply chain financial risk management research', *Journal of Physics: Conference Series*, IOP Publishing, Vol. 1744, No. 2, pp.22–27.

Zeng, H. (2022) 'Influences of mobile edge computing-based service preloading on the early-warning of financial risks', *The Journal of Supercomputing*, Vol. 78, No. 9, pp.11621–11639.