Journal of Global Information Management
Volume 30 • Issue 7

The Global Organizational Behavior Analysis for Financial Risk Management Utilizing Artificial Intelligence

Shufang Yang, Changshu Institute of Technology, China
Hainan Wu, Anhui University of Finance and Economics, China*

ABSTRACT

Enterprise financial risks are analyzed utilizing the theory of organizational behavior, and a financial risk management system is constructed to improve the design and algorithm of the enterprise risk management system. Based on the CCER (China Center for Economic Research) database, the early warning model for enterprise financial risk management containing five indices is proposed for enterprises. Through logistic regression analysis, the design principle of the financial risk management system based on AI (artificial intelligence) technology is explained. The proposed system innovatively introduces the AI-integrated learning method, optimizes objective function through XGBoost (extreme gradient boosting) algorithm, and trains the model through BP (backpropagation) NN (neural network). Finally, following comparative analysis, the effectiveness of the proposed method is verified.

KEYWORDS

Artificial Intelligence, Backpropagation Neural Network, Financial Risk Management System, ROC Curve Evaluation

INTRODUCTION

Nowadays, with the advancement of the market economy and the rapid change of the business environment, enterprises are faced with higher and higher survival risks, and risks have become unavoidable (Tron et al. 2018). According to statistics, about 80% of enterprises' life span is less than three years, 10% to 20% of enterprises' life span is only eight years, and only about 2% of enterprises' life span can reach more than 40 years. After the global financial risk, countries have given special attention to enterprise financial risks and financial crises (Wei et al. 2017). It is suggested that investigations should be conducted on how companies can warn financial risks and take effective action to prevent enterprise risks at an early stage (Zheng et al. 2018). The huge investment in R & D (Research and Development) and the high uncertainty of the conversion rate of R & D innovation will inevitably multiply the possibility of the financial risks (Tokakis et al. 2019). Therefore, the early investigation and management can improve the enterprises' ability to cope with the financial risk, protect them from the impact of the financial risk in the fierce market competition, reduce the possibility of the financial risk, and promote the stable development of the enterprises.

The concept of risk management is first proposed in the 1960s, which played an important role in decision science. Simola from the United States was the first scholar to manage enterprise risk

DOI: 10.4018/JGIM.292455
*Corresponding Author

and made important contributions to the research of the difference between the effectiveness of corporate leadership behavior and organizational factors (Simola, 2003). Scholars have researched the theory of enterprise risk management to help enterprises adapt to the rapidly changing internal and external environment, thus promoting the flexible management of enterprises (Bundy et al. 2017) through the application of theory into actual enterprise risk management. Thus, a systematic and organized enterprise risk management scheme is formed. Economic risk has become the main threat to the external environment of enterprises. Many enterprises have suffered unprecedented losses and even failures due to economic turbulence (Li et al. (2018). Under such many influential factors, risk management has become a daily work of enterprises (Coombs &Laufer, 2018). Therefore, under the threat of financial risks and other unstable situations, a feasible risk management system is needed to help enterprises prevent risk, respond to risks, and develop after the risks (Gwebu et al. 2018).

Based on the above social background, the organizational behavior theory is introduced first. Then, with extensive domestic and foreign research reviews, the advantages and disadvantages of the existing enterprise financial risk early-warning models are analyzed. Afterward, an optimized AI (Artificial Intelligence) and BPNN (Backpropagation Neural Network) model is constructed based on RF (Random Forest) and decision tree rules. Finally, the constructed model is applied to the design of an enterprise risk management system. The purpose is to seek an appropriate AI method for enterprise risk management and expand the application of integrated learning based on AI in enterprise risk management.

Accordingly, the main structure can be summarized as follows. Section 2 mainly reviews the existing research in related fields. Section 3 elaborates the research methods, data sets, and data analysis methods. Section 4 is the analysis and discussion of the research results and the comparative analysis with the existing research results. Section 5 is the conclusion part, which focuses on the innovation point and clarifies the limitations and the prospects.

RELATED WORK

Overview of Enterprise Risk Management Research

For enterprise risk management, scholars have done a lot of research. Yang et al. (2020) considered that risk management consisted of five parts: risk detection, risk identification, risk analysis, risk assessment, and risk management, and the purpose of risk management is to achieve the maximum utility at the lowest cost (Yang et al. 2020). Williams et al. (2017) referred to enterprise risk management as a dynamic risk management process of risk event detection, prevention. and handling, and adopted scientific thinking method and non-process decision-making model for risk management to minimize the damage caused by enterprise risk (Williams et al. 2017). To analyze enterprise risk management, Almamy et al. (2016) introduced a multivariate statistical method into the enterprise financial risk early-warning model (Almamy et al. 2016). Affes and Hentati Daffel (2019) proposed a logic model and applied it to the prediction of bankruptcy caused by enterprise risk (Affes &Hentati-Kaffel, 2019). Yang et al. (2019) established the supply chain financial risk management model and selected the data science analysis method to analyze the model. The results showed that the model had good subsection fitting, high data evaluation accuracy, and good robustness, which is suitable for the evaluation of enterprise financial risk management (Yang et al. 2019). Correa et al. (2019) analyzed enterprise risk management and provided a reference for enterprise comprehensive management, risk, and strategic coordination by combining risk management with strategic objectives (Corrêa et al. 2019).

The Application of AI Technology in Risk Management

The last decade of the 20th century has witnessed the development of AI technology, and some scholars have applied AI technology to the early warning and analysis of risks. Altman et al. (1994) applied the NN (Neural Network) model to the bankruptcy prediction of an Italian company and found that the performance of the NN model is better than that of the multivariate discriminant analysis model (Altman et al. 1994). Gupta and Sharma (2014) used the SVM (Support Vector Machines) model to forecast and study the financial risk of Korean enterprises and compared the SVM model with the multilinear model, multi judgment model, and NN model. The results show that the prediction accuracy of the SVM model is high (Gupta & Sharma, 2014). SMC et al. (2020) applied AI technology to the risk calculation and prediction of diseases and revealed the effectiveness of AI technology in disease prediction (Smc et al. 2020). Ao et al. (2020) applied the optimized ANN (Artificial Neural Network) based on wavelet transform to the risk assessment of air pollution and found that, compared with the classical NN model, the optimized hybrid ANN model can predict and analyze air pollution more accurately (Ao et al. 2020).

Presently, there are many types of research on enterprise risk management, but most of them are at the quantitative level and are in their infancy. The univariate deterministic model can distinguish the financial status of enterprises but cannot fully reflect their operating status. The single model is unstable and inaccurate. To sum up, AI technology has a wide application in risk management. However, the application of the AI integration method is relatively scarce in the research of early warning models for enterprise financial risk.

METHODS

Index System Construction of Financial Risk Management System

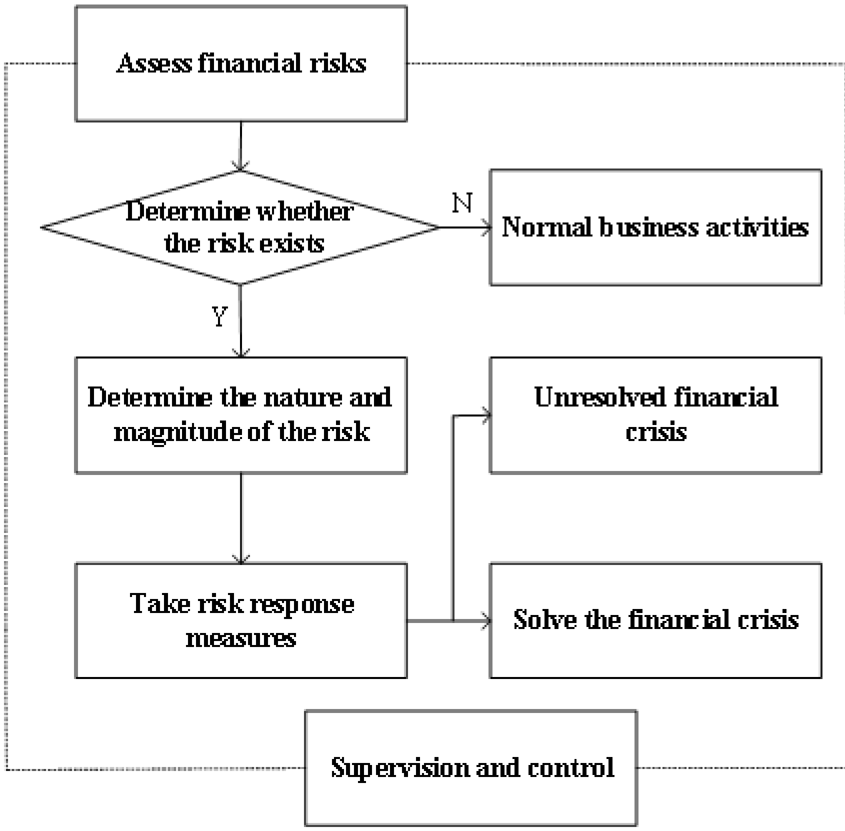
The mutation in the current market, coupled with the impact of management level, has brought a variety of risks to the enterprise. If an enterprise fails to take effective measures, enterprise risks will be inevitable. Financial indices can well reflect the enterprise's profitability, liabilities, and cash flow. If the enterprise's financial indices have early-warning signals, the enterprise needs to manage the risk (Liu et al. 2021). Under a market economy, financial risk is the warning signal of business conditions. Since financial risk covers all aspects of business activities, financial risk research is essential for the development of enterprises (Jin et al. 2018). Enterprises try to obtain the best return under low risk through financial risk management and early risk warning. Enterprise financial risk management process is shown in Fig. 1.

The sample data sets are selected from the CCER (China Center for Economic Research) China Economic and Financial Database from 2015 to 2019. These include healthy enterprises, bankrupt enterprises, and delisted enterprises, specifically, 100 samples of healthy enterprises, 100 samples of ST (Special Treatment) or * ST enterprises, and 100 samples of enterprises, which have declared bankruptcy or delisted. Here, a relatively complete sample financial risk early-warning system is established, the financial situation is considered of each enterprise in the last financial year and the coming financial year. Twenty-five alternative indices are selected for the constructed financial risk early-warning model from five aspects: per-share index, profitability, solvency, cash flow, and growth potential. Table 1 is an index classification system.

Before the establishment of the financial risk early-warning model, the K-S (Kolmogorov-Smirnov) normal distribution test has been performed on 25 selected financial indices through SPSS (Statistical Product and Service Solutions) software, and then key financial indices are significantly analyzed. K-S test can determine whether the sample data follow a specific distribution, so K-S can test whether the sample set conforms to the normal distribution theory or whether there is a significant difference between the two groups of variables (Maingot et al. 2018). Here are the basic steps for K-S testing. First, the assumption should be made based on the sample to be tested. Then, the empirical and theoretical distribution functions are calculated from the index variables to be tested and replaced by the following methods.

$$D_n = \max\{|F_n(x_i) - F_0(x_i)|, |F_n(x_{i-1}) - F_0(x_i)|\} \tag{1}$$

Figure 1. Enterprise financial risk management process



In which, $F_n(x)$ stands for the empirical distribution of the tested variable, and then D_n is compared with the size of the critical value $D_n(\alpha)$. If $D_n > D_n(\alpha)$, the original hypothesis will be rejected. Otherwise, the measured sample data are supposed to satisfy the normal distribution.

Mann-Whitney U test can determine whether there is a significant difference between the average values of the two indices, which is a method for testing the significant difference of the sample data with the unclear distribution. Here, the significance level is set as α=0.05. If the value of P of the index significance level is less than 5%, the null hypothesis is rejected. The analysis of the experimental results shows that among the 25 pre-selected financial indices, the significance level p of the eight indices are all less than 0.05, indicating that there is no significant difference. Therefore, these eight financial indices are excluded. The p values of the other 17 indices are all higher than 0.05, but these are very different and need to be retained.

Design Principle of AI Financial Risk Management System

The development of AI technology promotes the development of all walks of life, including human-computer interaction (Sun et al. 2020) and medical robots (Hu et al. 2019). Meanwhile, advanced AI technology plays a positive role in the development of enterprise management, optimizing various statistical analysis methods. Regression analysis is an optimization model of regression analysis. The maximum likelihood estimation is utilized in the above method to study the distribution of the

Table 1. Index classification system

Category	Variable	Index
Per-share index	X1	Networking capital
	X2	Net profit
	X3	Operating net cash flow per share
	X4	Net assets per share
	X5	Earnings per share
Profitability	X6	ROE (Return On Equity)
	X7	Net profit on the net sale
	X8	Operating profit margin
	X9	Sales cash ratio
	X10	Expenses to sales ratio
Solvency	X11	Assets-liability ratio
	X12	Capital adequacy ratio
	X13	Liquid ratio
	X14	Quick asset ratio
	X15	Capital turnover ratio
	X16	Cash flow to total debt ratio
Cash flow	X17	Rate of stock turnover
	X18	Turnover of current assets
	X19	Fixed assets turnover ratio
	X20	Turnover of total assets
Growth potential	X21	Total assets growth rate
	X22	Operating income growth rate
	X23	Operating profit growth rate
	X24	Net assets growth rate
	X25	Net profit growth rate

critical probability interval of the conditional probability of the selected samples and the extreme value. Consequently, the economic risk probability of samples can be determined (Yu &Zhang, 2017).

In the linear regression model, the hypothesis function of linear regression can be expressed as in equation (2).

$$\hat{y} = \omega^T x \tag{2}$$

In which, $\hat{y}$ denotes the predicted value of the output, ω represents the weight parameter, T stands for the transposition operation, and x represents the input value. Logistic Regression is a generalized linear model. Compared with the ordinary linear regression prediction function, the Logistic Regression prediction function should be converted.

$$y = g(z) \tag{3}$$

$$z = \omega^T x \tag{4}$$

In which, $g(z)$ generally represents a sigmoid function, as shown in equation (5).

$$g(z) = \frac{1}{1 + e^{-z}} \tag{5}$$

Then, a logarithmic form of the concession ratio is defined as $\log it$ and expressed as in equations (6) and (7).

$$\log it(\hat{y}_i) = \log \frac{p(y = 1 \mid x, \omega)}{p(y = 0 \mid x, \omega)} = \log \frac{\hat{y}_i}{1 - \hat{y}_i} = \omega^T x \tag{6}$$

$$\hat{y}_i = \frac{1}{1 + e^{-\omega^T x}} \tag{7}$$

In which, $\log it(p)$ stands for the inverse form of the concession ratio, as expressed in equation (8).

$$\omega^T x = \log \frac{p}{1 - p} \Rightarrow e^{\omega^T x} = \frac{p}{1 - p} \Rightarrow p = \frac{1}{1 + e^{-\omega^T x}} \tag{8}$$

Therefore, the hypothesis function of the linear function is regarded as the independent variable of the sigmoid function, that is, the hypothesis function of Logistic Regression, as expressed in equation (9).

$$\hat{y} = \frac{1}{1 + e^{-\omega^T x}} \tag{9}$$

In general, the hypothesis function of Logistic Regression is further processed, as shown in equation (10).

$$\begin{cases} \hat{y} > 0.5, \omega^T x > 0, y = 1 \\ \hat{y} < 0.5, \omega^T x < 0, y = 0 \end{cases} \tag{10}$$

Suppose that the Logistic Regression output is either category 0 or category 1, the probability is expressed as in equations (11) and (12).

$$p(y = 1 \mid x, \omega) = \pi(x) \tag{11}$$

$$p(y=0 \mid x,\omega) = 1 - \pi(x) \tag{12}$$

Since the value y is either 0 or 1, the probability distribution of y can be expressed as in equation (13).

$$p(y \mid x,\omega) = (\pi(x))^{y}(1-\pi(x))^{1-y} \tag{13}$$

The likelihood function is expressed as in equation (14).

$$L(\omega) = \prod_{i=1}^{m}[\pi(x_i)]^{y_i}[1-\pi(x_i)]^{1-y_i} \tag{14}$$

In which, m denotes the number of samples. The logarithmic likelihood function can be expressed as in equation (15).

$$J(\omega) = -\log L(\omega) = -\sum_{i=1}^{m}[y_i \log \pi(x_i) + (1-y_i)\log(1-\pi(x_i))] \tag{15}$$

The estimated value ω can be obtained through the calculated maximum value of $J(\omega)$.

Regularization can prevent model over-fitting (De Lucia et al. 2020). General regularization includes L_1 regularization and L_2 regularization. The regularizations are implemented after the objective function (or loss function) is obtained, and the multiple paradigms of parameter vectors L_1 and L_2 are added, as shown in equations (16) and (17).

$$J(\omega)_{L_1} = J(\omega) + \lambda \mid\mid \omega \mid\mid_1 \tag{16}$$

$$J(\omega)_{L_2} = J(\omega) + \frac{1}{2}\lambda \mid\mid \omega \mid\mid_2 \tag{17}$$

In which, the sum of absolute values of each parameter in the parameter vector is the norm L_1, and the square sum of each parameter in the parameter vector is the norm L_2. In regularization enhancement, the parameters which carry a small amount of information and have little contribution to the model approximates zero. Each feature makes a small contribution to the model rather than directly being removed (Xu et al. 2012). In general, if only to prevent excessive fitting, L_2 regularization will be chosen. Meanwhile, L_1 regularization should be chosen in the case of L_2 regularization, since the model is still overfitting and the performance of the unknown data set is still very poor.

AI Integrated Learning Method for Financial Risk Management System Design

The most advanced and popular ML (Machine Learning) method in the field of AI today is the integrated learning method (Li et al. 2019). The main idea is to generate multiple classifiers according to specific rules, and then combine different classifiers in a specific way for meta-classification. The integrated learning method has a better regularization than a single classifier and obtains the final output through comprehensive judgment. It avoids the shortcomings of the algorithm based on a single classifier and improves its performance (Ma et al. 2020). A well-known general algorithm in ensemble learning is RF (Random Forest).

To evaluate the importance of features through RF is relatively simple. It averages the contribution of each feature in each tree in RF and compares the contribution of different features. Here, the Gini coefficient can evaluate the contribution. The Gini coefficient index is introduced to measure the uncertainty of the data set. In a classification problem, it is assumed that the category of the sample is class k, and p_i represents the probability that the sample belongs to the i th class. The Gini index is defined as in equation (18).

$$Gini = \sum_{i=1}^{k} p_i(1-p_i) = 1 - \sum_{i=1}^{k} p_i^2 \tag{18}$$

Suppose that there are m features $X_1, X_2, \ldots, X_c$, and the Gini index score VIM_j is calculated for each feature X_j, as shown in equation (19).

$$GI_m = 1 - \sum_{k=1}^{|K|} p_{mk}^2 \tag{19}$$

In which, k represents k categories, and p_{mk} denotes the proportion of k categories in m nodes.

For a certain feature X_j, the change in the importance of the node m, namely, the change in the Gini index before and after the node branching, can be expressed as in equation (20).

$$VIM_{jm}^{gini} = GI_m - GI_l - GI_r \tag{20}$$

In which, GI_l and GI_r represent the value of the Gini index of the two new nodes generated after the m branch of this node.

If the node of feature X_j in a decision tree i falls into the set M, and then the importance of the feature X_j in the decision tree i can be expressed as in equation (21).

$$VIM_{ij}^{gini} = \sum_{m \in M} VIM_{jm}^{gini} \tag{21}$$

Suppose an RF has n trees, and then equation (22) can be obtained.

$$VIM_{j}^{gini} = \sum_{i=1}^{n} VIM_{ij}^{gini} \tag{22}$$

Finally, all importance scores are regularized, as shown in equation (23).

$$VIM_j = \frac{VIM_j}{\sum_{j=1}^{c} VIM_i} \tag{23}$$

XGBoost (eXtreme Gradient Boosting) is an improved algorithm based on GBDT (Gradient Boosting Decision Tree). The XGBoost algorithm can minimize the experience and structural risk loss functions through an optimized objective function to achieve better model generalization ability. The XGBoost algorithm can continuously generate trees, which is realized through the continuous functional division of the growth of each tree (Zhang et al. 2021). The XGBoost model is represented as follows.

$$\hat{y} = \sum_{k=1}^{K} f_k(x_i) \tag{24}$$

$$F = \{f(x) = \omega_{q(x)}\} \quad (25$$

In which, $q : R^m \to T, \omega \in R^T$, $\omega_{q(x)}$ represents the score of the leaf node q, and $f_k(x_i)$ denotes one of the regression trees. The XGBoost objective function can be expressed as in equation (26).

$$Obj = \sum\nolimits_{i=1}^{n} l(y_i, \hat{y}_i) + \sum\nolimits_{k=1}^{K} \Omega(f_k) \tag{26}$$

After t trees are generated, the prediction score can be expressed as equation (27).

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + f_i(x_i) \tag{27}$$

The objective function for the t th iteration can be expressed as in equation (28).

$$Obj^t = \sum\nolimits_{i=1}^{n} l(y_i, \hat{y}_i^{t-1} + f_t(x_i)) + \Omega(f_t) + cons\tan t \tag{28}$$

The next step is to find f_t that minimizes the objective function and the residual fitting error. The objective function can be approximately expressed as in equation (29).

$$Obj^t = \sum\nolimits_{i=1}^{n} [l(y_i, \hat{y}^{t-1} + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) + cons\tan t \tag{29}$$

In which, g_i and h_i denote the first-order derivative and the second-order derivative, respectively, as shown in equations (30) and (31).

$$g_i = \partial_{\hat{y}^{(t-1)}} l(y_i, \hat{y}^{t-1}) \tag{30}$$

$$h_i = \partial^2_{\hat{y}^{(t-1)}} l(y_i, \hat{y}^{t-1}) \tag{31}$$

The objective function can be simplified to equation (32).

$$Obj = \sum_{i=1}^{n}[g_i f_t(x_i) + \frac{1}{2}h_i f_t^2(x_i)] + \Omega(f_t) \tag{32}$$

In the XGBoost algorithm, the regularization term is defined as in equation (33). To optimize the model, the adaptive variational PEE (Partial Differential Equation) is introduced into the model construction, whose effectiveness has been verified by Wei et al. (2019).

$$\Omega(f_t) = \gamma T + \frac{1}{2}\lambda\sum_{j=1}^{T}\omega_j^2 \tag{33}$$

The loss function values of all samples are summed up, and all samples of the same leaf node are reorganized, as shown in equation (34).

$$\begin{aligned} Obj^t &\approx \sum_{i=1}^{n}[g_i\omega_{q(x_i)} + \frac{1}{2}h_i\omega^2{}_{q(x_i)}] + \gamma T + \frac{1}{2}\lambda\sum_{j=1}^{T}\omega_j^2 \\ &= \sum_{j=1}^{T}[(\sum_{i\in I_i} g_i)\omega_j + \frac{1}{2}(\sum_{i\in I_i} h_i + \lambda)\omega_j^2] + \gamma T \end{aligned} \tag{34}$$

Consequently, the optimal ω and objective function optimal values can be obtained, respectively, as in equations (35) and (36).

$$\omega_j^* = -\frac{G_j}{H_j + \lambda} \tag{35}$$

$$Obj^* = -\frac{1}{2}\sum_{j=1}^{T}\frac{G_j^2}{H_j + \lambda} + \gamma T \tag{36}$$

In which, G_j and H_j are defined as in equations (37) and (38), respectively.

$$G_j = \sum_{i\in I_j} g_i \tag{37}$$

$$H_j = \sum_{i\in I_j} h_i \tag{38}$$

$$I_j = \{i \mid q(x_i) = j\} \tag{39}$$

In the above three equations, I_j represents a set of sample labels assigned to the jth leaf node.

Establishment of AI Model for Financial Risk Management System

The ANN (Artificial Neural Network) can learn specific rules continuously through self-training without the mapping relationship between input and output and can obtain the expected output after given the specific input. Hence, the optimized results can be obtained. The BPNN (Back Propagation Neural Network) model is a multilayer feedforward network trained according to the BP algorithm. The basic idea is to minimize the mean square error between the actual output and the expected output of the neural network based on the gradient descent method (Wei, 2019).

Fig.-2 demonstrates the jth basic BPNN node. $x_1, x_2, ..., x_n$ represent the input of the $1, 2, ..., i, ..., n$ th neuron, respectively. $w_{j1}, w_{j2}, ..., w_{ji}, ..., w_{jn}$ represent the connection strength between the jth neuron and the $1, 2, ..., i, ..., n$ th neuron, namely, the weight. b_j represents the threshold, and the transfer function is expressed as $f(\bullet)$. The output of the jth neuron is expressed as y_j. The net input S_j of the jth neuron can be expressed as in equation (40).

$$S_j = \sum_{i=1}^{n} w_{ji} x_i + b_j = W_j X + b_j \tag{40}$$

Figure 2. BP neuron structure

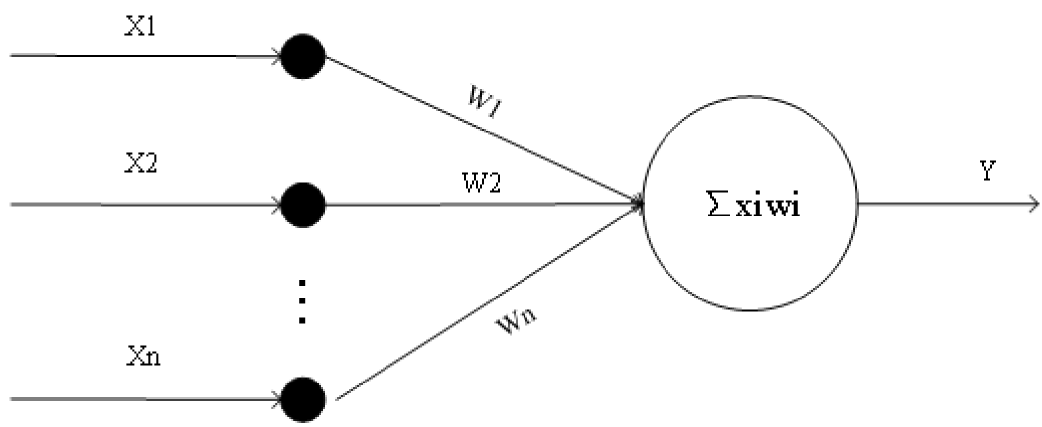


Fig. 2 can be explained through equations (41) and (42).

$$X = [x_1, x_2, ..., x_i, ..., x_n]^T \tag{41}$$

$$W_j = [w_{j1}, w_{j2}, ..., w_{ji}, ..., w_{jn}] \tag{42}$$

If $x_0 = 1, w_{j0} = b_j$, then equations (43) and (44) can be obtained.

$$X = [x_0, x_1, x_2, ..., x_i, ..., x_n]^T \tag{43}$$

$$W_j = [w_{j0}, w_{j1}, w_{j2}, ..., w_{ji}, ..., w_{jn}] \tag{44}$$

The net input S_j can be simplified as in equation (45).

$$S_j = \sum_{i=0}^{n} w_{ji} x_i = W_j X \tag{45}$$

S_j can output y_i through transfer function, as shown in equation (46).

$$y_i = f(s_j) = f(\sum_{i=0}^{n} w_{ji} x_i) = f(W_j X) \tag{46}$$

Here, the transfer function must be a bounded function, usually a monotonically increasing function, and must have a maximum value to ensure that the transmitted signal cannot be increased without restriction. Abnormal data during signal transmission can be input into the model and become abnormal data in the risk early warning model, impacting the survival time of the model. Consequently, the multi sink distributed power control algorithm is introduced into the abnormal data processing. In terms of the effectiveness of the algorithm, Wei et al. (2019) gave a clear explanation and verification in dealing with the abnormal data of the coal mine roadway. The algorithm could reduce the system energy consumption and extend the model lifetime, ensuring the data supervision effectiveness (Wei et al. 2019).

The node numbers of the input layer, the hidden layer, and the output layer of BPNN are set to n, q, m, respectively (Yuan et al. 2019). Then, the weights between the two adjacent layers include the input layer and the hidden layer, and the hidden layer and the output layer are denoted as v_{ki}, w_{jk}, respectively. The corresponding transfer functions of the hidden layer and output layer are $f_1(\bullet), f_2(\bullet)$, respectively. If the threshold is included in the sum, then the output of hidden layer nodes can be expressed as in equation (47).

$$z_k = f_1(\sum_{i=0}^{n} v_{ki} x_i), k = 1, 2, ..., q \tag{47}$$

The output of the output layer node can be expressed as in equation (48).

$$y_j = f_2(\sum_{k=0}^{q} w_{jk} z_k), j = 1, 2, ..., m \tag{48}$$

Thus, the n-dimension space vector of BPNN is mapped to m-dimension space vector approximately. Downtime occurs during model construction due to gradient rise. To deal with downtime, a model navigation method based on continuous information potential field and gradient rise is introduced in the BPNN application. In the existing research, Wei et al. (2017) proved that the model navigation method had good convergence and played a positive role in model optimization, system navigation, and positioning (Wei et al. 2019).

$x_1, x_2, \ldots, x_p$ represents P learning samples, and the output of the pth sample is denoted as y_j^p. t_j^p stands for the expected output, and the error of the pth sample E_p is defined as in equation (49).

$$E_p = \frac{1}{2}\sum_{j=1}^{m}(t_j^p - y_j^p)^2 \tag{49}$$

The global error of P samples can be expressed as in equation (50).

$$E = \frac{1}{2}\sum_{p=1}^{P}\sum_{j=1}^{m}(t_j^p - y_j^p)^2 = \sum_{p=1}^{P}E_p \tag{50}$$

w_{jk} is adjusted through the BP algorithm, and the global error E is reduced. η denotes the learning rate, and then equation (51) can be obtained.

$$\triangle w_{jk} = -\eta\frac{\partial E}{\partial w_{jk}} = -\eta\frac{\partial}{\partial w_{jk}}(\sum_{p=1}^{p}E_p) = \sum_{p=1}^{P}(-\eta\frac{\partial E_p}{\partial w_{jk}}) \tag{51}$$

Error signals can be defined as in equation (52).

$$\delta_{xj} = -\frac{\partial E_p}{\partial S_j} = -\frac{\partial E_p}{\partial y_j}\cdot\frac{\partial y_j}{\partial S_j} \tag{52}$$

In equation (52), the first and second items can be reduced and expressed as in equations (53)-(55).

$$\frac{\partial E_p}{\partial y_j} = \frac{\partial}{\partial y_j}[\frac{1}{2}\sum_{j=1}^{m}(t_j^p - y_j^p)^2] = -\sum_{j=1}^{m}(t_j^p - y_j^p) \tag{53}$$

$$\frac{\partial E_p}{\partial S_j} = f_2^{'}(S_j) \tag{54}$$

$$\delta_{xj} = \sum_{j=1}^{m}(t_j^p - y_j^p)f_2^{'}(S_j) \tag{55}$$

According to the chain theorem, equation (56) can be obtained.

$$\frac{\partial E_p}{\partial w_{jk}} = -\frac{\partial E_p}{\partial S_j} \cdot \frac{\partial S_j}{\partial w_{jk}} = -\delta_{xj} z_k = -\sum_{j=1}^{m} (t_j^p - y_j^p) f_2^{'}(S_j) \cdot z_k \tag{56}$$

In summary, the weight adjustment function can be obtained for each neuron in the output layer, as expressed in equation (57).

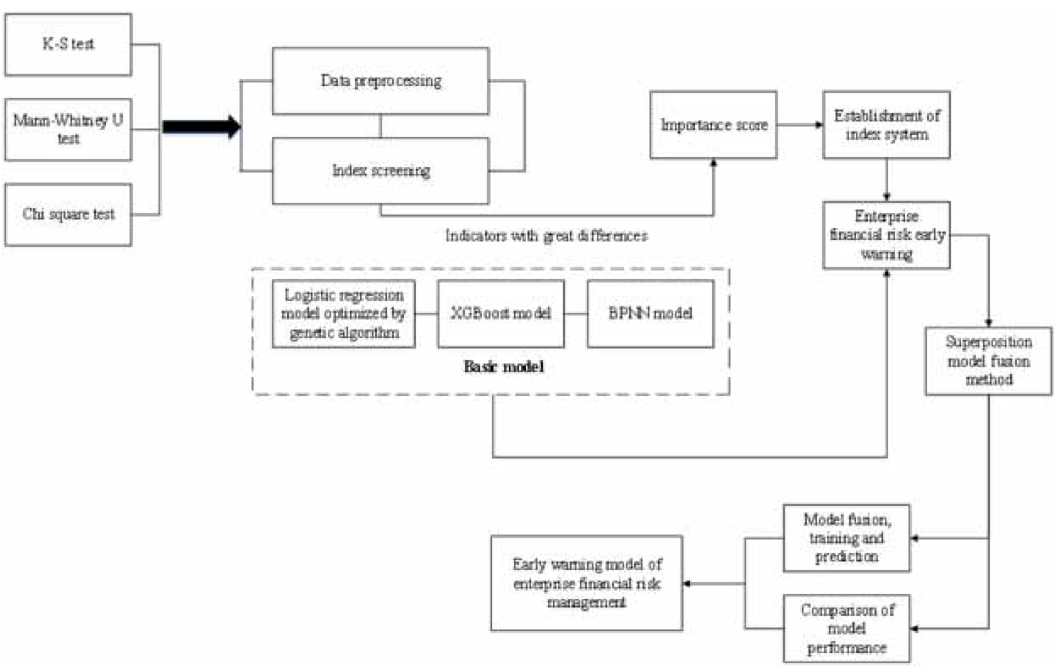
$$\Delta w_{jk} = \sum_{p=1}^{P} \sum_{j=1}^{m} \eta (t_j^p - y_j^p) f_2^{'}(S_j) \cdot z_k \tag{57}$$

The weight adjustment function can be obtained for each neuron in the hidden layer is, as expressed in equation (58).

$$\Delta v_{ki} = \sum_{p=1}^{P} \sum_{j=1}^{m} \eta (t_j^p - y_j^p) f_2^{'}(S_j) w_{jk} f_1^{'}(S_k) x_i \tag{58}$$

In summary, the overall implementation process of the early warning model for enterprise financial risk management based on the integrated learning method of AI is shown in Figure 3.

Figure 3. Overall implementation process for early warning model of enterprise financial risk management



Classification Evaluation of Fusion Models for Enterprise Financial Risk Management

The general evaluation indices for binary classification are introduced to better evaluate the model, including accuracy, precision, recall, F1-Score, AUC (Area Under ROC (Receiver Operating Characteristic Curve) Curve), and ROC. Accuracy rate and recall rate can evaluate binary classification problems (Hou &Chen, 2013). The classification results can be divided into four categories: TP, which represents True Positive, predicts positive categories to be positive. FP denotes False Positive and predicts negative categories to be negative. TN stands for True Negative and is a negative category. FN represents False Negative and predicts positive categories to be negative.

According to the confusion matrix, the precision rate can be defined as in equation (59).

$$precision = \frac{TP}{TP + FP} \tag{59}$$

Recall rate is defined as in equation (60).

$$recall = \frac{TP}{TP + FN} \tag{60}$$

The F1-score is the harmonized average of precision and recall.

Fig. 4 shows the ROC. The vertical axis and horizontal axis of ROC represent TPP (True Positive Rate) and FPR (False Positive Rate), respectively, which are defined as in equations (61) and (62).

$$TPR = \frac{TP}{TP + FN} \tag{61}$$

$$FPR = \frac{FP}{TN + FP} \tag{62}$$

Here, the ROC is introduced to quantify the performance of the classifier. Specifically, AUC can measure the possibility of positive samples being put over the negative samples, while the threshold will not be considered (Metz et al. 2015). The larger the AUC is, the better the classification effect is.

RESULTS AND DISCUSSIONS

Index Screening Results of the Financial Risk Management System

Here, the importance of each function is illustrated, which is evaluated through the algorithm itself. The packaging method trains the evaluator for the initial feature set and deletes the least important feature from the current feature set according to the importance of the feature. Then, the above procedure will be recursed on the trimmed set until the required features are finally selected. The optimal feature subset is selected through the feature engineering packaging method, the RF algorithm is chosen to create the model, and the model is determined according to the learning curve. The feature subset learning curve is demonstrated in Fig. 5.

Figure 4. ROC diagram

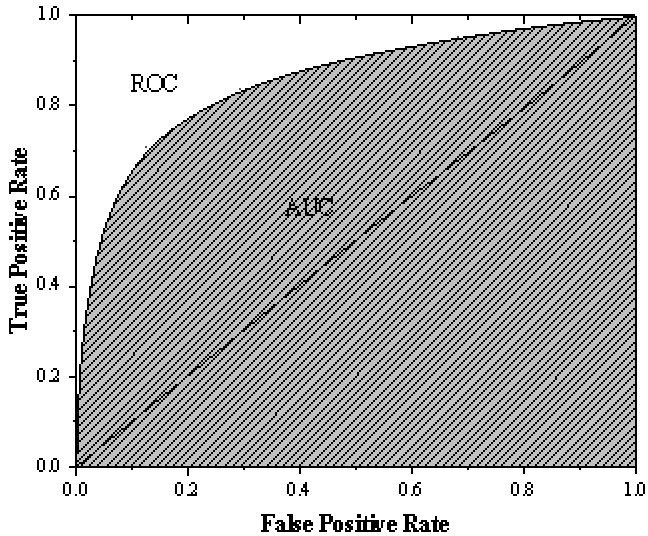


Figure 5. Feature subset learning curve

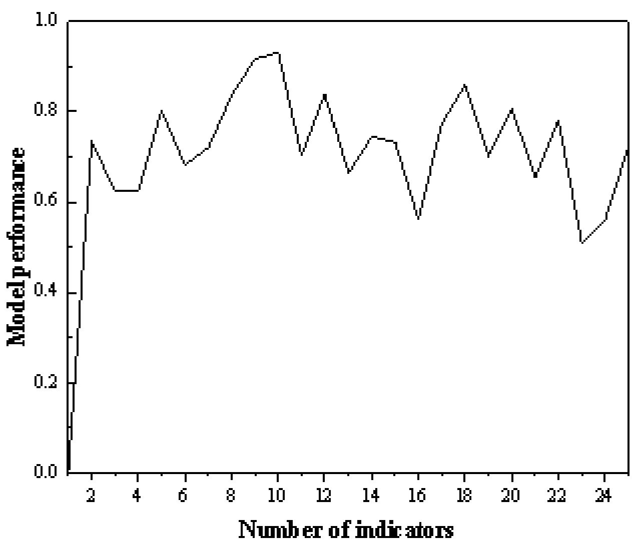
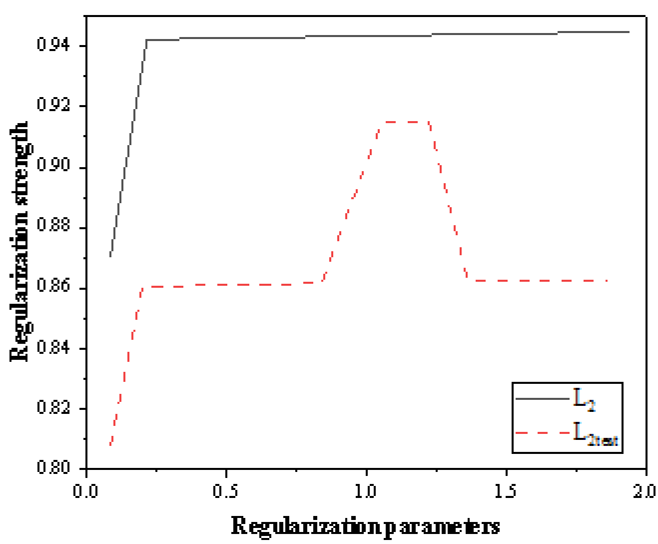


Table 2. Enterprise financial risk early-warning index system

Category	Variable	Index
Per-share index	X5	Earnings per share
Profitability	X6	ROE
	X7	Net profit on the net sale
	X8	Operating profit margin
	X9	Sales cash ratio
	X10	Expenses to sales ratio
Cash flow	X18	Turnover of current assets
	X20	Turnover of total assets
Growth potentialities	X24	Net assets growth rate
	X25	Net profit growth rate

Fig. 5 indicates that the performance of the early-warning model exceeds 93% with 10 features. Meanwhile, the stable points of the early warning imply that the model is most effective with 10 features. To be precise, the current performance has reached 93.2%. Ten indices selected through the packaging method can indicate the financial risks of the enterprise, as illustrated in Table 2.

Figure 6. Regularization parameter learning curve

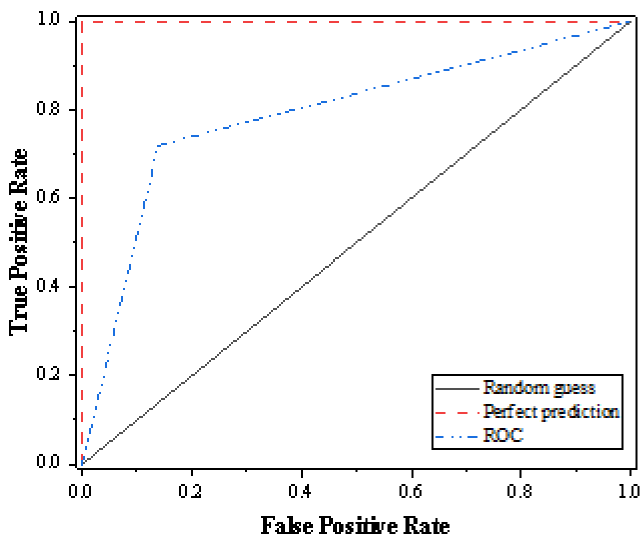


Logistic Regression Model Results of the Financial Risk Management System

To evaluate different models, cross-validation can select and adjust model parameters, and upsampling SMOTE (Synthetic Minority Oversampling Technique) principle can deal with unbalanced data (Liu et al. 2021). Besides, data are divided and retained in advance. Fig. 6 is the regularization parameter learning curve.

Fig. 6 shows that as the regularization parameter C increases, the strength of regularization becomes smaller, and the performance of the early-warning model becomes better regardless of the training set or test set. The performance of the training set keeps improving until C=0.8 while the performance of the unknown data set (test set) begins to decline. There is an overfitting problem in the model. The prediction results of the final model for the test set are shown in Fig. 7.

Figure 7. ROC of Logistic Regression model

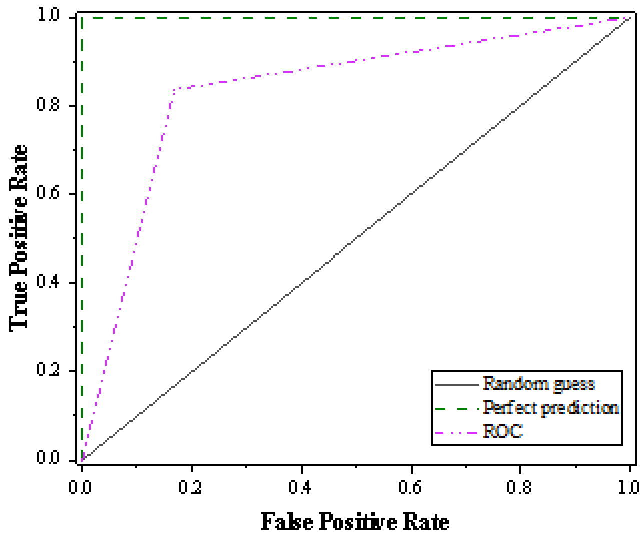


According to the identification and classification of test samples, the model predicts 12 ST companies as non-ST (normal) companies, and the prediction accuracy of ST companies is 73.2%. The model predicts nine non-ST companies as ST companies, the prediction accuracy for non-ST companies is 82%, and the ultimate prediction accuracy of the model is 76.8%. Here, the prediction accuracy of ST is mainly concerned, and the effect of the whole model is well-balanced. As shown in ROC, the AUC value of the curve is 0.793.

Results of XGBoost Model for Financial Risk Management System

To illustrate the trimming process clearly, three groups of curves are utilized for comparison. The first group of the curve can display the original model data, the second group can display previous parameter adjustments, and the last group can show the current model parameter adjustments (Shen et al. 2019). The final prediction results of the model for the test set are shown in Fig. 8.

Figure 8. ROC of XGBoost model



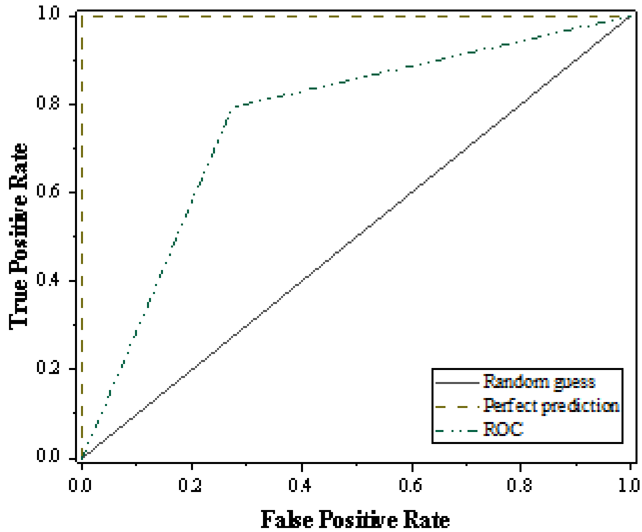
According to the identification and classification of test samples, the model predicts eight ST companies as non-ST companies, and the prediction accuracy of ST companies is 84.3%. The model predicts nine non-ST as ST companies, the prediction accuracy of non-ST companies is 81.7%, and the final prediction accuracy of the model is 82.9%. Compared with the Logistic Regression model, the prediction accuracy of the ST companies is much higher. Meanwhile, the ROC curve shows that the AUC value is 0.836, which is better than the Logistic Regression model. In terms of enterprise risk classification, Wang et al. (2019) discussed the effectiveness of the XGBoost model and compared the XGBoost model with the traditional Logistic regression model. The results showed that the performance of XGBoost was better than that of the Logistic regression model. The ranking of feature importance could improve the explanation of the model, and the XGBoost model is an effective business risk modeling method (Wang &Ni, 2019). Peng et al. (2021) applied the XGBoost model to the prediction of hypertension risk, combined with the feature selection, and found effective results (Peng et al. 2021). To sum up, the validity of the proposed model can be proved based on the above analysis. Meanwhile, the PDE is also introduced here to make the performance more remarkable.

BPNN Model Results of the Financial Risk Management System

After been screened through multiple steps, 10 indices are finally chosen for the financial risk management system with early-warning indices. Hence, the input layer of BPNN has 10 nodes. Here, the number of the hidden layer is determined as six after several tests. If a financial risk occurs, the ST will be output and will be marked as 1. If the enterprise is financially healthy, the non-ST will be output, will be marked as 0, and contains one output layer node. The parameters of the BPNN are set as follows: the training efficiency is set to 0.5, the training target error is set to 0.00001, and the number of network training is set to 10,000. The final prediction results of the model for the test set are shown in Fig. 9.

According to the identification and classification of test samples, this model predicts six ST companies as non-ST companies, and the prediction accuracy of ST companies is 88%. This model predicts seven non-ST companies as ST companies, and the prediction accuracy of non-ST companies

Figure 9. ROC of BPNN model



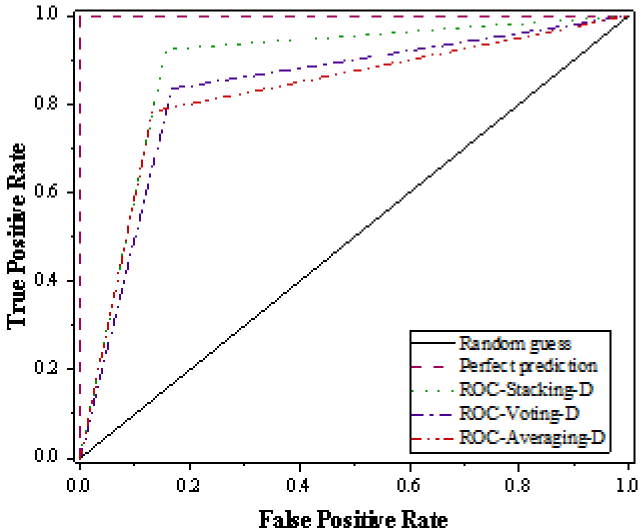
is 86%, and the final prediction accuracy of the model is 87.2%. As shown in the ROC, the AUC value is 0.812. In terms of the BPNN model and financial risk management. Jin et al. (2020) studied the power transmission and transformation investment project and analyzed the cost-effectiveness by constructing the fuzzy optimization BPNN model, finding that the prediction error of the situation based on the optimization model was only 0.3%, and the training speed was very fast, with good results, The expected output was basically consistent with the actual output (Jin et al. 2020). Thus, the applicability of the BPNN model in risk management model training can be further proved. Here, the innovation point is to introduce the model navigation method based on continuous information potential field and gradient rise, so the model can be more targeted for different applications and play an active role in the analysis of abnormal data.

Results for the Fused Model of the Financial Risk Management System

The above three models are different, with low correlation and close model performance. The basic principle of model fusion can be satisfied by these three models. The greater the difference between models is, the better the performance of each model is, and the closer their performances are to the optimum (Chen, 2018). Here, two common model fusion methods are chosen, Voting and Averaging, and are compared with the proposed Stacking fusion model method. The fusion model D discriminant results and ROC comparison are shown in Fig. 10.

Fig. 10 indicates that the performance of the stacking fusion model is better than that of the test set. The accuracy and the AUC value of the model are high, and at the same time, the recall rate is better. Meanwhile, the ROC curves of the other two fusion models are contained within the ROC curve of the Stacking fusion model. Compared with the voting and averaging fusion methods, the performance of the model constructed through the risk management system has been improved. If the determined threshold is low, the performance of the averaging model fusion method is close to that of the stacking fusion model (Chen et al. 2019). However, with the increase of the threshold, the performance of the Stacking model fusion method will be better than that of the other two methods.

Figure 10. D discriminant results and ROC of the fused model



Different from the single model, the integrated fusion model incorporates the excellent performance of the single model, introduces the advanced methods, and finally, presents more excellent performance. Xu et al. (2020) proposed a stacking prediction model based on AI integrated learning and information fusion for the prediction of users' online shopping behavior and verified its effectiveness in pre-distortion applications (Xu et al. 2020). The results strongly support the performance analysis results of the proposed intelligent integrated fusion model. The proposed model integrates various advanced algorithms, so it has more application potential.

CONCLUSION

The descriptive analysis is utilized together with the K-S test, Mann-Whitney U test, and chi-square test for data pre-processing and index screening to obtain significant differences. Functional engineering packaging technology can further screen out very different indices, and according to the importance score, a simple and effective index system is established to warn the financial risk of the enterprise. The Logistic regression model optimized by the genetic algorithm is chosen, and then the XGBoost model and the BPNN models are selected as the basic models. Afterward, the three models are fused, trained, and predicted through the stacking model fusion method. Then, the fused model is compared with the voting and averaging fusion models. Subsequently, an early-warning model for the company's financial risk management is established based on the Logistic Regression model, XGBoost model, and BPNN model. The prediction accuracy of the model is 85.7%, the recall rate is 91.4%, and the AUC value is 0.895, which is better than those of other single models and fusion models.

However, there are still some deficiencies. Here, different algorithms' performances applied in the model are compared and analyzed, while the specific application in enterprise risk management is not involved. In future work, the proposed model will be further optimized, and empirical analysis will be conducted according to the actual situation.

REFERENCES

Tron, A., Valenza, G., & Caputo, A. (2018). Corporate crisis management in Italy: Execution, monitoring and performance analysis of recovery business and financial plans. *International Journal of Foresight and Innovation Policy*, *13*(1-2), 114–135. doi:10.1504/IJFIP.2018.095861

Wei, J., Ouyang, Z., & Chen, H. (2017). Well known or well-liked? The effects of corporate reputation on firm value at the onset of a corporate crisis. *Strategic Management Journal*, *38*(10), 2103–2120. doi:10.1002/smj.2639

Zheng, B., Liu, H., & Davison, R. M. (2018). Exploring the relationship between corporate reputation and the public's crisis communication on social media. *Public Relations Review*, *44*(1), 56–64. doi:10.1016/j.pubrev.2017.12.006

Tokakis, V., Polychroniou, P., & Boustras, G. (2019). Crisis management in public administration: The three phases model for safety incidents. *Safety Science*, *113*(3), 37–43. doi:10.1016/j.ssci.2018.11.013

Simola, S. (2003). Ethics of justice and care in corporate crisis management. *Journal of Business Ethics*, *46*(4), 351–361. doi:10.1023/A:1025607928196

Bundy, J., Pfarrer, M. D., Short, C. E., & Coombs, W. T. (2017). Crises and crisis management: Integration, interpretation, and research development. *Journal of Management*, *43*(6), 1661–1692. doi:10.1177/0149206316680030

Li, Y., Phau, I., Lu, W., & Teah, M. (2018). Crisis management of food security scandals in China: Motivations and solutions towards purchase intention. *Journal of Consumer Behaviour*, *17*(1), 13–24. doi:10.1002/cb.1659

Coombs, W. T., & Laufer, D. (2018). Global crisis management–current research and future directions. *Journal of International Management*, *24*(3), 199–203. doi:10.1016/j.intman.2017.12.003

Gwebu, K. L., Wang, J., & Wang, L. (2018). The role of corporate reputation and crisis response strategies in data breach management. *Journal of Management Information Systems*, *35*(2), 683–714. doi:10.1080/07421222.2018.1451962

Yang, R., Yu, L., Zhao, Y., Yu, H., Xu, G., Wu, Y., & Liu, Z. (2020). Big data analytics for financial market volatility forecast based on support vector machine. *International Journal of Information Management*, *50*(2), 452–462. doi:10.1016/j.ijinfomgt.2019.05.027

Williams, T. A., Gruber, D. A., Sutcliffe, K. M., Shepherd, D. A., & Zhao, E. Y. (2017). Organizational response to adversity: Fusing crisis management and resilience research streams. *The Academy of Management Annals*, *11*(2), 733–769. doi:10.5465/annals.2015.0134

Almamy, J., Aston, J., & Ngwa, L. N. (2016). An evaluation of Altman's Z-score using cash flow ratio to predict corporate failure amid the recent financial crisis: Evidence from the UK. *Journal of Corporate Finance*, *36*(2), 278–285. doi:10.1016/j.jcorpfin.2015.12.009

Affes, Z., & Hentati-Kaffel, R. (2019). Forecast bankruptcy using a blend of clustering and MARS model: The case of US banks. *Annals of Operations Research*, *281*(1), 27–64. doi:10.1007/s10479-018-2845-8

Yang, Q., Wang, Y., & Ren, Y. (2019). Research on financial risk management model of internet supply chain based on data science. *Cognitive Systems Research*, *56*, 50–55. doi:10.1016/j.cogsys.2019.02.001

Corrêa, R. G. F., Neto, F., Souza, J. S., Lampert, V. N., & Barcellos, J. (2019). Enterprise risk management in integrated crop-livestock systems: A method proposition. *The Journal of Agricultural Science*, *156*(10), 1–11.

Altman, E. I., Marco, G., & Varetto, F. (1994). Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience). *Journal of Banking & Finance*, *18*(3), 505–529. doi:10.1016/0378-4266(94)90007-8

Gupta, P., & Sharma, A. M. (2014). A study of the impact of corporate governance practices on firm performance in Indian and South Korean companies. *Procedia: Social and Behavioral Sciences*, *133*(5), 4–11. doi:10.1016/j.sbspro.2014.04.163

Smc, A., Wr, B., Ktw, C., Hf, D., Br, E., & Nh, F. (2020). The ethical, legal and social implications of using artificial intelligence systems in breast cancer care. *The Breast*, *49*, 25–32. doi:10.1016/j.breast.2019.10.001 PMID:31677530

Ao, A., Mmh, A., & Asb, C. (2020). Multi hours ahead prediction of surface ozone gas concentration: Robust artificial intelligence approach. *Atmospheric Pollution Research*, *11*(9), 1572–1587. doi:10.1016/j.apr.2020.06.024

Liu, Z., Lang, L., Hu, B., Shi, L., Huang, B., & Zhao, Y. (2021). Emission reduction decision of agricultural supply chain considering carbon tax and investment cooperation. *Journal of Cleaner Production*, *294*(4), 126305. doi:10.1016/j.jclepro.2021.126305

Jin, M., Wang, Y., & Zeng, Y. (2018). Application of data mining techniques in financial risk analysis. *Wireless Personal Communications*, *102*(4), 3699–3713. doi:10.1007/s11277-018-5402-5

Maingot, M., Quon, T., & Zéghal, D. (2018). Risk management by US and Canadian financial firms during the financial crisis. *International Journal of Disclosure and Governance*, *15*(4), 221–234. doi:10.1057/s41310-018-0049-8

Sun, Y., Xu, C., Li, G., Xu, W., Kong, D., Jiang, B., Tao, B., & Chen, D. (2020). Intelligent human-computer interaction based on nonredundant EMG signal. AEJ -. *Alexandria Engineering Journal*, *59*(3), 1149–1157. doi:10.1016/j.aej.2020.01.015

Hu, J., Sun, Y., Li, G., Jiang, G., & Tao, B. (2019). Probability analysis for grasp planning facing the field of medical robotics. *Measurement*, *141*, 227–234. doi:10.1016/j.measurement.2019.03.010

Yu, Q. W., & Zhang, L. (2017). Financial crisis early-warning model for a listed company in China energy industry based on logistic regression. *Izvestiia po Himiia*, *49*, 31–36.

De Lucia, C., Pazienza, P., & Bartlett, M. (2020). Does good ESG lead to better financial performances by firms? machine learning and logistic regression models of public enterprises in Europe. *Sustainability*, *12*(13), 5317. doi:10.3390/su12135317

Xu, Z., Chang, X., Xu, F., & Zhang, H. (2012). L1/2 regularization: A thresholding representation theory and a fast solver. *IEEE Transactions on Neural Networks and Learning Systems*, *23*(7), 1013–1027. doi:10.1109/TNNLS.2012.2197412 PMID:24807129

Li, G., Li, J., Ju, Z., Sun, Y., & Kong, J. (2019). A novel feature extraction method for machine learning based on surface electromyography from a healthy brain. *Neural Computing & Applications*, *31*(12), 9013–9022. doi:10.1007/s00521-019-04147-3

Ma, C., Wang, K., Chi, Y., & Chen, Y. (2020). Implicit regularization in nonconvex statistical estimation: Gradient descent converges linearly for phase retrieval, matrix completion, and blind deconvolution. *Foundations of Computational Mathematics*, *20*(3), 451–632. doi:10.1007/s10208-019-09429-9

Zhang, Z., Huang, Y., Qin, R., Ren, W., & Wen, G. (2021). XGBoost-based on-line prediction of seam tensile strength for Al-Li alloy in laser welding: Experiment study and modeling. *Journal of Manufacturing Processes*, *64*(5), 30–44. doi:10.1016/j.jmapro.2020.12.004

Wei, W., Zhou, B., Połap, D., & Woźniak, M. (2019). A regional adaptive variational PDE model for computed tomography image reconstruction. *Pattern Recognition*, *92*, 64–81. doi:10.1016/j.patcog.2019.03.009

Yuan, S., Wang, G., Chen, J., & Guo, W. (2019). Assessing the forecasting of comprehensive loss incurred by typhoons: A combined PCA and BP neural network model. *Journal on Artificial Intelligence.*, *1*(2), 69–88. doi:10.32604/jai.2019.06535

Wei, W., Xu, X., Marcin, W., Fan, X., & Ye, L. (2019). Multi-sink distributed power control algorithm for cyber-physical-systems in coal mine tunnels. *Computer Networks*, *161*, 210–219. doi:10.1016/j.comnet.2019.04.017

Lv, S., & Lv, S. (2011). Applying BP neural network model to forecast peak velocity of blasting ground vibration. *Procedia Engineering*, *26*(1), 257–263.

Wei, W., Song, H., Wei, L., Shen, P., & Vasilakos, A. (2017). Gradient-driven parking navigation using a continuous information potential field based on wireless sensor network. *Information Sciences*, *408*, 100–114. doi:10.1016/j.ins.2017.04.042

Hou, W. J., & Chen, H. Y. (2013). Rule extraction in gene-disease relationship discovery. *Gene*, *518*(1), 132–138. doi:10.1016/j.gene.2012.11.060 PMID:23235120

Metz, C. E., Herman, B. A., & Shen, J. H. (2015). Maximum likelihood estimation of receiver operating characteristic (ROC) curves from continuously distributed data. *Statistics in Medicine*, *17*(9), 1033–1053. doi:10.1002/(SICI)1097-0258(19980515)17:9<1033::AID-SIM784>3.0.CO;2-Z PMID:9612889

Liu, Y., Yang, L., & Chen, M. (2021). A new citation concept: Triangular citation in the literature. *Journal of Informetrics*, *15*(2), 101141. doi:10.1016/j.joi.2021.101141

Shen, C., Min, C., & Wang, C. (2019). Analyzing the trend of O2O commerce by bilingual text mining on social media. *Computers in Human Behavior*, *101*(12), 474–483. doi:10.1016/j.chb.2018.09.031

Wang, Y., & Ni, X. (2019). A XGBoost risk model via feature selection and Bayesian hyper-parameter optimization. *International Journal of Database Management Systems.*, *11*(1), 1–17. doi:10.5121/ijdms.2019.11101

Peng, Y., Xu, J., Ma, L., & Wang, J. (2021). Prediction of hypertension risks with feature selection and XGBoost. *Journal of Mechanics in Medicine and Biology*, *21*(05), 2140028. Advance online publication. doi:10.1142/S0219519421400285

Jin, X., Liu, Q., & Long, H. (2020). Impact of cost-benefit analysis on financial benefit evaluation of investment projects under back propagation neural network. *Journal of Computational and Applied Mathematics*, *384*(2), 113172.

Chen, M. (2018). The research of human individual's conformity behavior in emergency situations. *Library Hi Tech*, *38*(3), 593–609. doi:10.1108/LHT-08-2018-0113

Chen, J., Ye, X., Chen, M., & Liang, Y. (2019). Bibliometric analysis of the papers on urban education. *Library Hi Tech*, *37*(4), 894–905. doi:10.1108/LHT-01-2019-0009

Xu, J., Wang, J., Tian, Y., Yan, J., Li, X., & Gao, X. (2020). Se-stacking: Improving user purchase behavior prediction by information fusion and ensemble learning. *PLoS One*, *15*(11), 1–17. doi:10.1371/journal.pone.0242629 PMID:33237926

Hainan Wu is the corresponding author of this paper. He can be contacted at wuhainan9721@126.com.