

## Research Article

# Enterprise Financial Early Warning Based on Improved Whale Optimization Algorithm: Optimize the Perspective with Indicators

Bowei Li,<sup>1</sup> Mengzui Di ,<sup>2</sup> Zikun Wei,<sup>3</sup> Hong Qiao,<sup>1</sup> and Xuzhao Li<sup>4</sup>

<sup>1</sup>School of Economics and Management, Hebei Agricultural University, Baoding 071000, China

<sup>2</sup>School of Mechanical and Electrical Engineering, Hebei Agricultural University, Baoding 071000, China

<sup>3</sup>School of Economics, Nankai University, Tianjin 300071, China

<sup>4</sup>School of Plant Protection, Hebei Agricultural University, Baoding 071000, China

Correspondence should be addressed to Mengzui Di; 202021401061@stu.hebut.edu.cn

Received 7 July 2022; Revised 12 August 2022; Accepted 31 August 2022; Published 16 September 2022

Academic Editor: Juan Vicente Capella Hernandez

Copyright © 2022 Bowei Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The key to solving the problem of redundant financial indicators in addressing financial warning issues is to reduce the dimensionality of the original financial indicators. This paper proposes a model based on the whale optimization algorithm with mixed strategy (IWOA) combined with support vector machine (SVM), namely, the IWOA-SVM early warning model, which simultaneously performs index optimization and dimensionality reduction, and financial risk early warning identification. This paper takes a total of 302 enterprises specially treated in Shanghai and Shenzhen stock exchanges and normal enterprises of the same specification as the research samples to design the model. The results show that the improved whale optimization algorithm has better optimization speed and accuracy and improves the search ability of the original algorithm for the optimal solution. Compared with other dimensionality reduction methods, the IWOA-SVM model has the lowest index dimension after dimensionality reduction and has more excellent recognition effect. The dimensionality reduction results have certain universality for different classifiers, which provides a new idea for the selection of indicators for financial early warning.

## 1. Introduction

Since its inception in 1932, financial early warning models have undergone a transformation from univariate early warning models to multivariate early warning models, to logistic early warning models [1] and modern financial early warning models based on machine learning and integrated learning. When financial early warning models are built, they are prone to redundant financial indicators due to the richness and complexity of financial indicators. The presence of redundant financial indicators not only increases the computational effort of the model but also reduces the accuracy of its identification. Therefore, through the selection and dimensionality reduction of the original financial indicators, screening the optimal financial indicator combination is very important in the construction of the financial early warning model.

Fang and Yang [2] proposed the SGL-SVM method and applied it to the prediction of financial distress, which reduced the original 90-dimensional indicator variables to 24 dimensions, eliminating a large amount of noisy data while obtaining good identification results. Chen [3] and Fang [4] et al. used PCA to reduce the dimensionality of the indicator data and selected the top principal components with high variance contribution instead of the original indicators. In the face of high-dimensional indicators, Huang et al. [5] first performed independence tests on indicators to eliminate insignificant early warning indicators, after which they used random forest and XGBoost to calculate the importance of indicators, eliminated insignificant indicators, and used KPCA to reduce 14-dimensional financial indicators to 7-dimensional to construct a combined KPCA-WLSSVM model with high predictive power and generalization effect. Li et al. [6] selected the top ten features in terms of importance for feature optimization based on the feature

importance evaluation of the random forest algorithm, and the results showed that the overall accuracy, sensitivity, specificity, and AUC of early warning model were all improved after feature optimization. Zhou et al. [7] used grey clustering to select valid variables in their study on early warning of credit risk for listed companies, followed by logistic regression models for prediction. Luo and Wang [8] used an improved MRMR algorithm in constructing a financial early warning model, taking into account both feature relevance and redundancy for feature preference. Xiaoyan et al.[9] used significance tests and normality tests to select features with significant differences as a way to improve the accuracy of financial early warning models.

Feature selection is an important tool for feature dimensionality reduction. When preferential dimensionality reduction is performed based on the correlation information of indicators, features that are favorable to the identification results may be excluded, and the interpretability of the model is relatively poor when the financial early warning model is constructed using factor analysis. Feature preference dimensionality reduction using metaheuristic algorithms has been widely used in areas such as behavioral recognition [10], network intrusion detection [11], and performance prediction [12]. For the field of financial early warning, metaheuristic algorithms are mostly used for hyperparameter optimization of classifiers [13–15] and are less frequently applied in financial indicator preference dimensionality reduction. The whale optimization algorithm (WOA) is a novel metaheuristic algorithm proposed by Mirjalili and Lewis [16], which, like other metaheuristics, suffers from the problem of low accuracy in finding the optimal solution and how to improve the algorithm's ability to search for the optimal solution has received much attention from scientists [17, 18]. Kaur and Arora [19] introduced a chaos mechanism to optimize the initial position of the population. Introducing adaptive weights [20] or improving the convergence factor [21] for the algorithm can balance the search ability between the early and late stages of the algorithm. Therefore, this paper uses improved whale optimization algorithm with mixed strategy (IWOA) and support vector machine (SVM) to preferentially reduce the dimensionality of the original financial indicators and construct an IWOA-SVM financial warning model.

## 2. IWOA-SVM Financial Early Warning Model Construction

**2.1. Whale Optimization Algorithm.** The whale optimization algorithm is a new intelligent optimization algorithm that has been proposed based on the special feeding behavior of humpback whales, which consists of three main parts: encircling prey, bubble-net attack, and searching for prey.

**2.1.1. Encircling Prey.** During the whale feeding process, the location of the prey is first observed and searched to surround it. The optimal solution to the problem is predetermined to be the location of the prey, and during the iteration of the algorithm, the fitness value calculated by the

fitness function is used to evaluate the merit of each group of financial indicators. Therefore, the design of the fitness function is crucial. In order to take into account the smallest possible dimension of the financial indicators while having a high accuracy rate, the fitness function in this paper is

$$\text{Fitn}(IWOA) = 0.99 \cdot (1 - \text{Acc}) + 0.01 \cdot \left( \frac{\text{feasel}}{\text{numfea}} \right). \quad (1)$$

Among them,  $\text{Acc}$  represents the correct rate of five-fold cross-validation of each financial indicator feature combination in the SVM classifier,  $\text{feasel}$  is the dimension of financial indicator features included in randomly selected individuals, and  $\text{numfea}$  represents the total dimension of financial indicator features.

In the iteration process of the algorithm, the position of the individual with the optimal fitness value of the current population is taken as the optimal position, and other individuals are close to the optimal position. The mathematical expression is

$$\begin{aligned} X(t+1) &= X_{\text{best}}(t) - A \cdot D, \\ D &= |CX_{\text{best}}(t) - X(t)|, \end{aligned} \quad (2)$$

where  $X_{\text{best}}(t)$  denotes the position vector of the optimal individual of the current population,  $X(t)$  denotes the position vector of each individual in the current population, and  $t$  denotes the current number of iterations, the expressions of  $A$  and  $C$  are as follows.

$$\begin{aligned} A &= 2 \cdot a \cdot r_1 - a, \\ C &= 2 \cdot r_2, \\ a &= 2(1 - t/T_{\text{max}}), \end{aligned} \quad (3)$$

where  $a$  is the convergence factor,  $r_1$  and  $r_2$  are random numbers within  $[0, 1]$ , and  $T_{\text{max}}$  denotes the maximum number of iterations.

**2.1.2. Bubble-Net Attack.** This phase consists of two types of position update, contraction bracketing, and spiral: first, contraction bracketing, in which the position is updated by adjusting the convergence factor  $a$  in equation (3); and second, spiral position update, which simulates a whale spiraling up to hunt close to its prey, with the mathematical expression:

$$\begin{aligned} X(t+1) &= X_{\text{best}}(t) + D_p \cdot e^{bl} \cos(2\pi l) \\ D_p &= |X_{\text{best}}(t) - X(t)|, \end{aligned} \quad (4)$$

where  $D_p$  denotes the distance between an individual whale and the current optimal solution,  $b$  denotes a constant in the shape of the spiral, and  $l$  denotes a random number in the range  $[-1, 1]$ . At this stage, the individual position of the whale is updated by randomly selecting contraction encirclement and spiral contraction. The mathematical expression is

$$X(t+1) = \begin{cases} X_{\text{best}}(t) - A \cdot D, & p < 0.5, \\ X_{\text{best}}(t) + D_p \cdot e^{bl} \cos(2\pi l), & p \geq 0.5, \end{cases} \quad (5)$$

where  $p$  represents a random number in the range  $[0, 1]$ .

**2.1.3. Searching for Prey.** Individual whales search for prey by randomly swimming away. The random search strategy allows the algorithm to have a good global search performance and randomly selects whale individuals in the population to update their position when  $|A| \geq 1$ . The mathematical expression is

$$\begin{aligned} X(t+1) &= X_{\text{rand}}(t) - A \cdot D, \\ D &= |C \cdot X_{\text{rand}}(t) - X(t)|, \end{aligned} \quad (6)$$

where  $X_{\text{rand}}(t)$  denotes a randomly selected individual whale.

**2.2. Improvements to the Whale Optimization Algorithm.** As the whale optimization algorithm has the problems of poor global search ability, easily falling into local optimal solutions, and poor convergence accuracy in the computation process, this paper uses Gauss mapping to initialize the population to improve population diversity, introduces adaptive weights to balance global and local search ability, and uses a stochastic dimension-by-dimension variation strategy based on Cauchy mutation and reverse learning to improve the ability to jump out of local optima with a hybrid strategy to improve the whale optimization algorithm.

**2.2.1. Gauss Mapping to Initialize the Population.** The original whale optimization algorithm initializes the population in a random way, that is, it randomly generates a combination of financial indicators. In order to expand the population search range and improve the population diversity, this paper adopts Gauss mapping to initialize the population, which is more traversable and more uniformly distributed compared to the original algorithm, and the mathematical expression of Gauss mapping is

$$X_{K+1} = \begin{cases} 0, & X_K = 0, \\ \frac{1}{X_K \bmod(1)}, & X_K \neq 0, \end{cases} \quad (7)$$

$$\frac{1}{X_K \bmod(1)} = \frac{1}{X_K} - \left[ \frac{1}{X_K} \right].$$

Among them,  $\bmod$  represents the remainder function,  $[\ ]$  represents rounding, and  $X_K$  represents the chaotic sequence generated by Gauss mapping.

**2.2.2. Adaptive Weights.** Adaptive weights can effectively balance global and local search capabilities, for which an adaptive weighting factor  $\omega$  is introduced with the following mathematical expression.

$$\omega = \sin\left(\frac{\pi \cdot t}{2 \cdot T_{\text{max}}} + \pi\right) + 1, \quad (8)$$

where  $t$  is the current number of iterations and  $T_{\text{max}}$  is the maximum number of iterations.

The equation for updating the location of individual whales after the introduction of adaptive weights is as follows.

$$X(t+1) = \omega \cdot X_{\text{best}}(t) - A \cdot D, \quad (9)$$

$$X(t+1) = \omega \cdot X_{\text{best}}(t) + D_p \cdot e^{\text{bl}} \cos(2\pi l), \quad (10)$$

$$X(t+1) = \omega \cdot X_{\text{rand}}(t) - A \cdot D. \quad (11)$$

**2.2.3. Stochastic Dimension-by-Dimension Variation Strategy Based on Cauchy Mutation and Reverse Learning Cauchy Mutation.** The Cauchy distribution, similar to the normal distribution, is one of the common distributions in the probability theory and is characterized by a uniform distribution of variances due to its slow decline from peak to zero values, the mathematical expression of which is

$$X'_{\text{best}}(t) = X_{\text{best}}(t) + \text{cauchy}(0, 1) * X_{\text{best}}(t). \quad (12)$$

Reverse learning improves the search performance of an algorithm by solving the current solution backward in the same space and is widely used in various optimization algorithms with the mathematical formulation of

$$X'_{\text{best}}(t) = ub + (lb - X_{\text{best}}(t)). \quad (13)$$

When only one variation strategy is selected in the algorithm, it will lead the algorithm into the local optimum problem, and the traditional variation approach mostly uses random variation or variation in all dimensions, while the dimension-by-dimension variation approach can avoid the influence between dimensions and fully explore the optimal solution; therefore, this paper uses a random dimension-by-dimension variation strategy based on Cauchy mutation and reverse learning to perturb the optimal individual.

Since the new solution generated is not necessarily better than the optimal position, a greedy rule is used to decide whether to adopt the new solution.

$$x' = \begin{cases} x'_{\text{best}}, & f(x'_{\text{best}}) < f(x_{\text{best}}), \\ x_{\text{best}}, & f(x'_{\text{best}}) \geq f(x_{\text{best}}). \end{cases} \quad (14)$$

**2.3. IWOA-SVM Financial Early Warning Model Construction.** Due to a large number of linear and nonlinear complex relationships between the financial indicators of an enterprise, when constructing a financial early warning model, low-quality financial indicator data can lead to feature confounding and high model computation, reducing the financial early warning capability of the model. Therefore, it is necessary to reduce the dimensionality of financial indicators and select the optimal combination of financial indicators with good predictive power. The IWOA-SVM financial early warning model constructed in this paper improves the whale optimization algorithm through mixed strategies and improves the search ability for the optimal

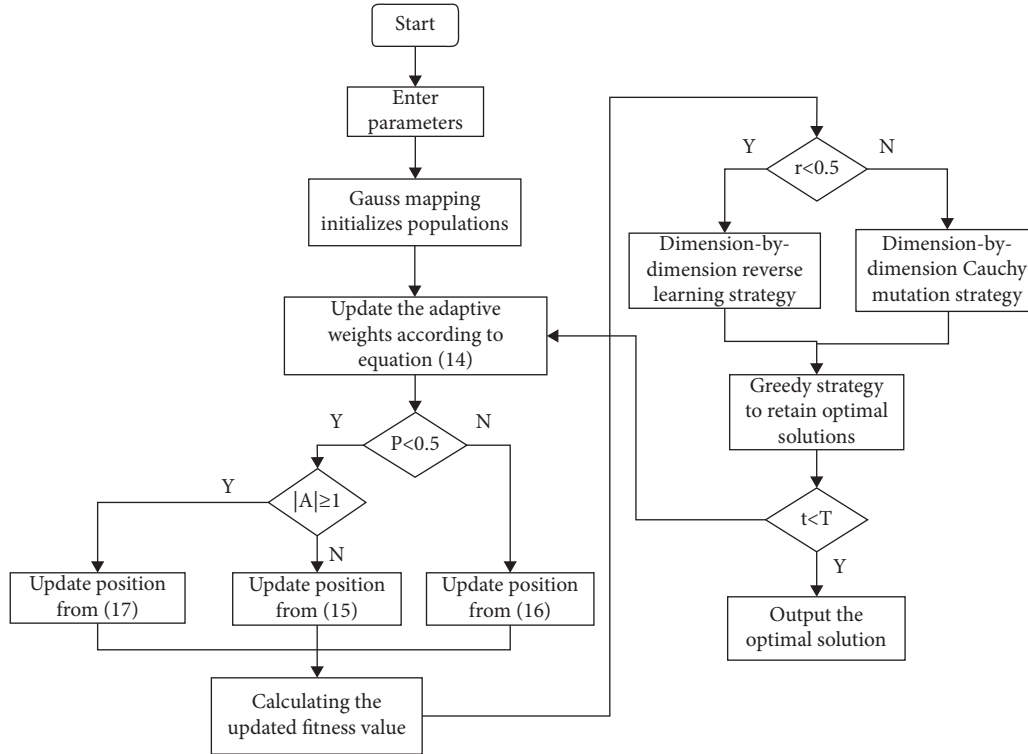


FIGURE 1: IWOA-SVM financial early warning model.

solution, that is, the combination of financial indicators with the best predictive ability, to improve the model recognition effect. The flow chart of the IWOA-SVM model is as follows Figure 1.

The key steps in the IWOA-SVM financial early warning model are as follows:

- (1) Data standardisation: due to the different magnitudes of the financial indicator values and the fact that SVM is typically a classification algorithm based on a distance metric, to avoid impact on the model performance, the financial indicator data is normalized by the formula as follows:

$$X = \frac{x - \mu}{\delta}, \quad (15)$$

where  $\mu$  is the mean and  $\delta$  is the standard deviation of the data.

- (2) Initializing the population. The number of IWOA algorithm populations is set to 20 and the maximum number of iterations is 50, i.e., 20 sets of financial indicator combinations are generated at each iteration using the Gauss mapping strategy.
- (3) Calculating the fitness value. Based on the fitness function, the fitness values of all current combinations of financial indicators are calculated and the individual with the lowest fitness value is taken as the current optimal solution.
- (4) Location update. When the probability  $p < 0.5$  and  $|A| < 1$ , the position is updated according to

equation (9), if  $|A| \geq 1$ , the position is updated according to equation (10); when the probability  $p \geq 0.5$ , the position is updated by equation (11).

- (5) Optimal position perturbation. The optimal solution is optimized using a stochastic dimension-by-dimension variation strategy based on Cauchy mutation and reverse learning. If the fitness value of the new solution is better than the optimal solution before optimization, the current optimal combination of financial indicators and fitness values are retained and discarded if not.
- (6) Iterative operations. Determine whether the maximum number of iterations has been reached, and if the condition is met, output the optimal combination of financial indicators and the optimal fitness value, if not, repeat steps 3–5 to continue the search.
- (7) Output SVM algorithm classification results based on the optimal combination of financial indicators.

### 3. Empirical Analysis

**3.1. Sample Selection and Data Sources.** The article selects A-share listed companies in Shanghai and Shenzhen as the empirical sample and classifies the empirical sample into normal companies and ST companies by taking whether the listed companies are specially treated (ST) during their existence as the criterion. For normal companies, continuous financial data are obtained for the years 2019–2021; for ST companies, the year in which the main body of the ST

company is specially treated is recorded as year  $t$ , and continuous financial data are obtained for the previous two years. In order to eliminate the possible adverse effects of imbalanced categorical data on the model, after excluding companies with serious data deficiencies, the positive and negative samples were matched according to a 1:1 ratio, using the rule that companies belonging to the same or similar industries and the overall size of the companies were comparable. The final selection of 151 ST companies and 151 normal companies yielded a total of 302 valid samples and 906 sets of observations. All data were sourced from the CSMAR database.

**3.2. Selection of Indicators.** When constructing financial early warning models for listed companies, it is common to start with financial indicators that reflect the operating status of the company. In this paper, a total of 32 financial indicators are selected from four aspects: solvency, profitability, operational capability, and development capability, specifically five indicators are selected from solvency. In order, they are current ratio (X1), quick ratio (X2), gearing ratio (X3), equity multiplier (X4), and equity ratio (X5); 12 indicators were selected in terms of profitability, in order, they are return on assets (X6), net profit margin on total assets (X7), net profit margin on current assets (X8), net profit margin on fixed assets (X9), EBITDA (X10), earnings before interest, tax, depreciation and amortization (X11), gross operating margin (X12), operating profit margin (X13), net operating margin (X14), management expense ratio (X15), financial expense ratio (X16), and cost margin (X17); 11 indicators were selected from the operating capacity, in order of accounts receivable to revenue ratio (X18), inventory to revenue ratio (X19), operating cycle (X20), accounts payable turnover (%) (X21), current assets to revenue (X22), current assets turnover (X23), fixed assets to revenue (X24), fixed assets turnover (X25), noncurrent assets turnover (X26), capital intensity (X27), and total assets turnover (X28); and four indicators were selected in terms of growth capacity, in the following order: growth rate of total operating revenue (X29), growth rate of total operating costs (X30), growth rate of administrative expenses (X31), and sustainable growth rate (X32). Considering that nonfinancial indicators are also important in risk identification and early warning, this paper selects two nonfinancial indicators at the level of management's governance capacity, namely, the proportion of management men (X33) and the average age of management (X34) in that order. The finalized corporate financial early warning model contains a total of 34 indicators with a large number of features, which should be subject to dimensionality reduction.

**3.3. IWOA-SVM Model Identification Results.** The article divides the training and test sets in the ratio of 6:4, sets the initial population of the whale optimization algorithm to 20 and the number of iterations to 50, and runs the WOA-SVM before improvement and the IWOA-SVM model after improvement twenty times, respectively, under the same hardware conditions, and the number of financial indicators,

the number of convergence generations, and accuracy curves are shown in Figure 2 after ranking them according to the accuracy rate from smallest to largest, and the distribution is shown in Table 1. After comparison, it can be seen that compared with the WOA-SVM model, the IWOA-SVM model has different effects in terms of convergence speed, convergence accuracy, and reducing the dimension of financial indicators. As can be seen from Figure 2, the IWOA-SVM model largely outperforms the WOA-SVM model in terms of the number of generations of convergence and the number of financial indicators when ranked from small to large in terms of accuracy. As can be seen from Table 1, after the improvement of the mixed strategy, the average number of convergence generations was reduced from 12 to 5.15, a reduction of 6.85; the average number of financial indicator dimensions was reduced from 14.90 to 10.95, a reduction of 3.95; meanwhile, the average accuracy of IWOA-SVM model identification was increased from 84.50% to 86.74%, an increase of 2.24%, and the highest accuracy was increased from 86.23% to 87.60%, an increase of 1.37%. It can be seen that the convergence speed and optimization-seeking ability of the mixed strategy improved whale optimization algorithm proposed in this paper are significantly improved over the original algorithm, which can eliminate redundant indicators while having good accuracy and can meet the financial warning needs of this paper.

The IWOA-SVM model identified the highest accuracy of 87.60%, while the financial indicators with an accuracy of 87.60% were obtained from two groups. The first group includes current ratio (X1), quick ratio (X2), equity multiplier (X4), equity ratio (X5), total assets net profit margin (X7), EBITDA (X11), average age of management (X34), current assets to revenue ratio (X22), fixed assets to revenue ratio (X24), noncurrent assets turnover ratio (X26), and capital intensity (X27), a total of 11 financial indicators; the second group includes current ratio (X1), equity multiplier (X4), total assets net profit margin (X7), EBITDA (X10), EBITDA (X11), cost margin (X17), the average age of management (X34), accounts receivable to revenue (X18), inventories to revenue (X19), fixed assets to revenue (X24), noncurrent asset turnover (X26), total asset turnover (X28), and sustainable growth rate (X32), a total of 13 financial indicators.

The identification accuracy of the two sets of financial indicators, the error rate of the first category, and the error rate of the second category are shown in Table 2. The first type of error rate (error I) is the proportion that the model identifies ST samples as normal samples, and the second type of error rate (error II) is the proportion that the model identifies normal samples as ST samples. In consideration of the principle of accounting prudence, error I should be avoided as far as possible, and a high error rate of St samples will lose the practical significance of financial early warning. Comparing the identification results of the first group and the second group, the second group had 27 ST samples predicted as non-ST, the first group had 25 ST samples predicted as non-ST, and the first group had a lower error I, better early warning effect. Therefore, the combination of financial indicators of the first group is selected as the

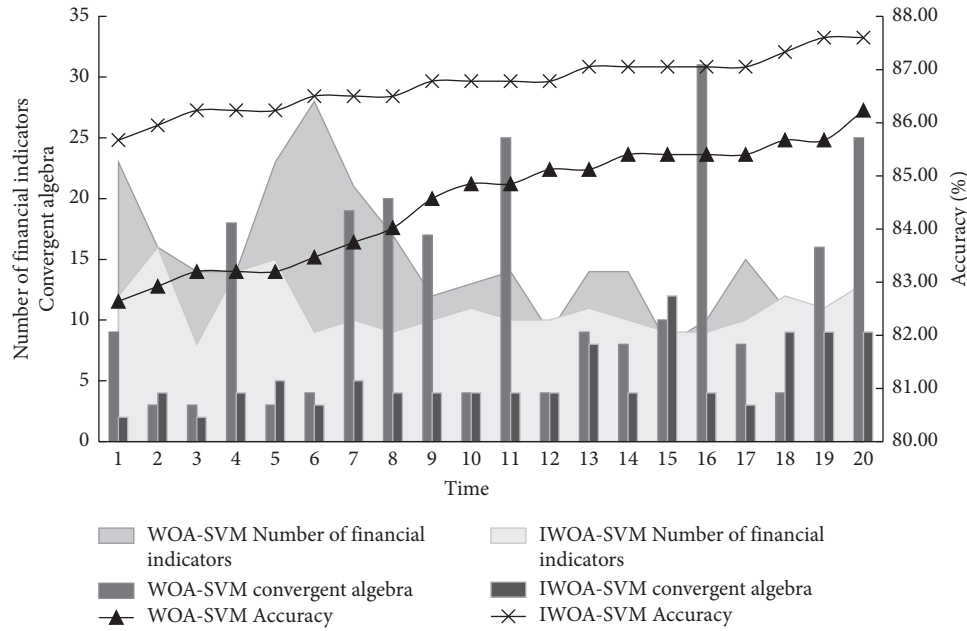


FIGURE 2: Comparison of the results of 20 runs of IWOA-SVM and WOA-SVM.

TABLE 1: Distribution of IWOA-SVM and WOA-SVM model results.

Model	Distribution	Convergence generation	Indicator dimension	Accuracy %
WOA-SVM	Max	31	28	86.23
	Min	3	8	82.64
	Average	12	14.90	84.50
IWOA-SVM	Max	12	16	87.60
	Min	2	8	85.67
	Average	5.15	10.95	86.74

TABLE 2: Comparison of identification results.

Project	True class	Predicted class		Total	Result (%)	
		ST	Non-ST		Error I	Error II
Group one financial indicators	ST	156	25	181	Error I	13.81
	Non-ST	20	162	182	Error II	10.99
	Total	176	187	363	Accuracy	87.6
Group two financial indicators	ST	154	27	181	Error I	14.82
	Non-ST	18	164	182	Error II	9.89
	Total	172	191	363	Accuracy	87.6

optimal combination for financial early warning in this paper. A total of 156 ST samples, 162 non-ST samples were correctly identified, 25 ST samples were predicted as non-ST, and 20 non-ST samples were predicted as ST, with an error rate of 13.81% for the error I and 10.99 for the error II. The meanings of the financial indicators in the first group are shown in Table 3.

3.4. Comparison of Different Dimensionality Reduction Methods. In order to compare the effect of model dimensionality reduction, this paper selected principle component analysis (PCA), mutual information and maximal information coefficient (MIC), recursive feature

elimination (RFE), and XGBoost importance ranking four kinds of dimensionality reduction, with the increase of the selected dimensionality SVM recognition accuracy changes as shown in Figure 3, respectively, and selected the combination of indicators with the highest recognition rate as the corresponding optimal combination; recognition results are shown in Table 4. It can be seen from Table 4 that, among the four dimensionality reduction methods, the XGBoost importance ranking method selects the smallest combination dimension of financial indicators and has the lowest type two error rate, but the type one error rate is also at the highest level. In contrast, the IWOA-SVM model constructed in this paper improves the accuracy by 4.68%, reduces the type one error rate by 7.18%, and reduces the

TABLE 3: Meaning of financial indicators.

Indicator name	Meaning
Current ratio (X1)	Current assets/current liabilities
Quick ratio (X2)	(Current assets–inventories)/current liabilities
Equity multiplier (X4)	Total assets/total owners' equity
Equity ratio (X5)	Total liabilities/total owners' equity
Total assets net profit margin (X7)	Net profit/average balance of total assets
Earnings before interest, tax, depreciation, and amortization (X11)	Net profit + income tax expense + finance costs + depreciation of fixed assets, depreciation of oil and gas assets, depreciation of productive biological assets + amortization of intangible assets + amortization of long-term amortization expenses
Current assets to revenue ratio (X22)	Current assets/operating income
Fixed assets to revenue ratio (X24)	Fixed assets/operating income
Non-current assets turnover ratio (X26)	Operating income/noncurrent assets
Capital intensity (X27)	Total assets/operating income
Average age of management (X34)	The average of the ages of all directors and supervisors of the company (directors and supervisors whose ages are not disclosed are not included in the calculation)

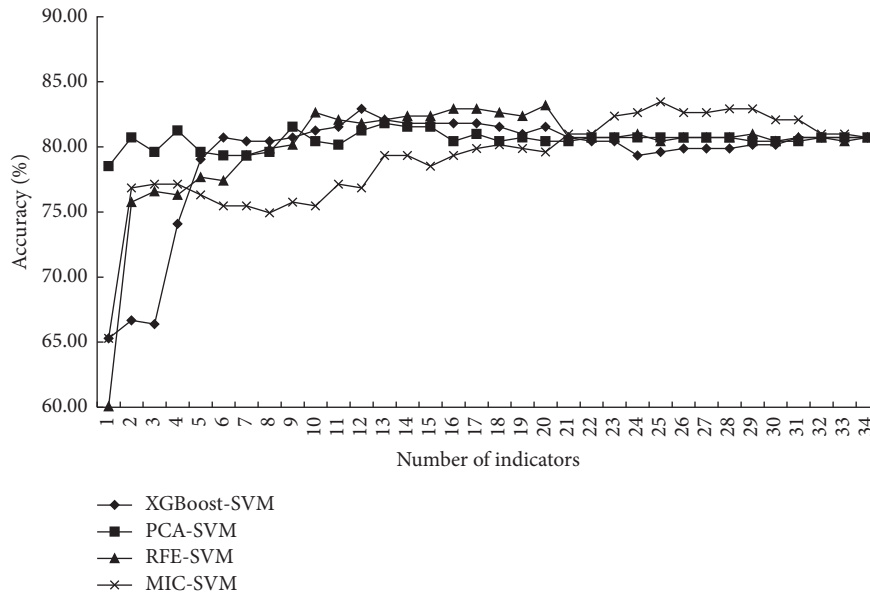


FIGURE 3: Variation in accuracy of the four downscaling methods.

TABLE 4: Identification results of different dimensionality reduction methods.

Method	Accuracy, %	Error I, %	Error II, %	Indicator dimension
PCA-SVM	81.82	20.44	14.93	13
MIC-SVM	83.47	18.78	14.29	25
RFE-SVM	83.20	19.89	13.74	20
XGBoost-SVM	82.92	20.99	13.19	12
IWOA-SVM	87.60	13.81	10.99	11

type two error rate by 2.2% when the index dimension is reduced by 1 dimension. The mic dimensionality reduction model has the highest accuracy and the lowest type one error rate, but the accuracy is 4.13% lower than the IWOA-SVM model, the type one error rate is 4.97% higher, and the type two error rate is 3.3% higher. However, the feature dimension of the model is as high as 25 dimensions, and the dimensionality reduction effect is not obvious. In summary, the IWOA-SVM model constructed in this paper can

achieve optimal identification results with a minimum number of financial indicators.

3.5. Algorithm Performance Comparison. To compare the performance of the IWOA-SVM algorithm, salp swarm algorithm (SSA), particle swarm optimization (PSO), flower pollination algorithm (FPA), and grey wolf optimizer (GWO) were selected. Under the same hardware

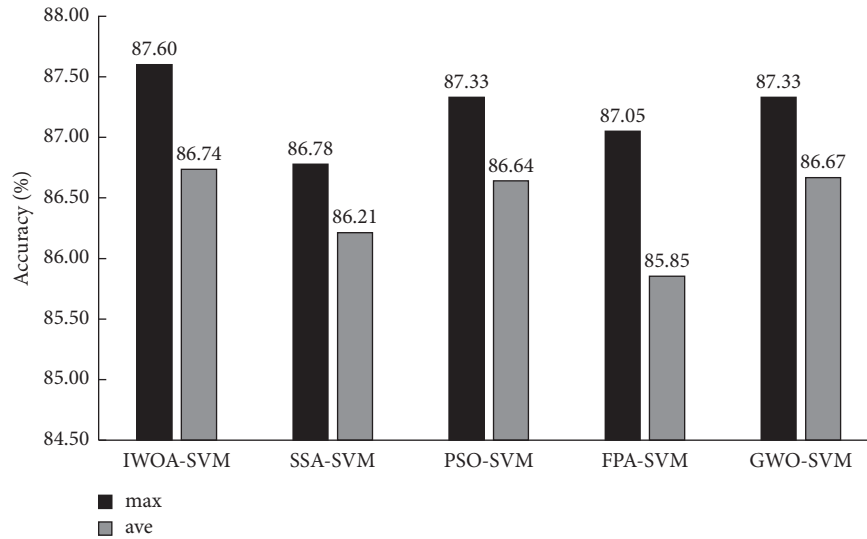


FIGURE 4: Algorithm performance comparison.

TABLE 5: Comparison of recognition results of different classifiers.

Classifier	Before indicator selection			After indicator selection		
	Accuracy, %	Error I, %	Error II, %	Accuracy, %	Error I, %	Error II, %
LG	78.79	20.99	21.43	83.20	19.89	13.74
RF	83.75	18.23	14.29	86.50	16.02	10.99
KNN	77.96	24.31	19.78	82.09	18.23	17.58
DT	79.34	22.10	19.23	81.54	20.44	16.48
SVM	80.72	19.34	19.23	87.60	13.81	10.99

conditions, keep the same parameter settings as IWOA-SVM. After running 20 times, the maximum and average values of the recognition accuracy are shown in Figure 4. It is easy to find that the IWOA-SVM model in this paper has the highest recognition accuracy and average accuracy, and the model performance is better than other metaheuristic algorithms.

**3.6. Comparison of Universality of Dimension Reduction Effect.** In order to verify the algebra liability and identification of the preferred dimensionality reduction effect of the IWOA-SVM financial warning model, five financial warning models based on logistic regression (LG), random forest (RF), K-nearest neighbor (KNN), decision tree (DT), and SVM were constructed simultaneously in this paper. Identify the original financial indicator combination and the optimal financial indicator combination obtained in this paper. The identification results are shown in Table 5. It can be seen from the table that the logistic regression, random forest, K-nearest neighbor, and decision tree models are obtained using this paper. After the optimal combination of financial indicators, the model recognition accuracy increased by 4.41%, 2.75%, 4.13%, and 2.2%, respectively, and the first type error rate decreased by 1.1%, 2.21%, 6.08%, and 1.66%, respectively. Compared with the SVM model for identifying the original data set, the IWOA-SVM model has an accuracy improvement of 6.88% to 87.60%, and the first type error

rate decreased by 5.53% to 13.81%, and the second type error rate decreased by 8.24% to 10.99%. Collectively, it appears that the IWOA-SVM financial warning model constructed in this paper has the highest accuracy rate, the lowest type one error rate, and the best model identification effect.

#### 4. Discussion and Conclusions

To address the issue of indicator selection in the construction of financial early warning models, this paper proposes the IWOA-SVM financial early warning model. A sample of 151 ST companies and 151 normal companies listed on A-shares in Shanghai and Shenzhen in 2019–2021 was used for the empirical analysis, and a mixed strategy is used to improve the original whale optimization algorithm for the problems of poor merit-seeking ability and slow convergence accuracy. The empirical analysis shows that first, the IWOA-SVM algorithm reduced the original 34-dimensional metrics to 11 dimensions, with a model recognition accuracy of 87.60%, an improvement of 6.88%, and a type one error rate of 13.81%, a reduction of 5.53%, the dimensionality reduction and recognition better than the four dimensionality reduction methods such as PCA. Second, compared to the original whale algorithm, the improved whale optimization algorithm with mixed strategy reduced the average financial indicator dimension by 3.95 and the average number of convergence generations by 6.85, the highest accuracy was improved by 1.37%, indicating the



algorithm's convergence speed and optimization-seeking ability was enhanced. Thirdly, using logistic regression, random forest, K-nearest neighbor, and decision tree to identify the combination of indicators before and after dimensional reduction and optimization, the accuracy rate has improved to different degrees, indicating that the method of dimensional reduction and optimization of indicators has certain universality for different classifiers. Taken together, the IWOA-SVM financial early warning model proposed in this paper has the highest accuracy rate and the lowest type I error rate, providing a new way of thinking about the selection of indicators for the construction of financial early warning models, which is an accurate, efficient, and scientific financial early warning model.

## Data Availability

The data presented in this study are available on request from the corresponding author.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Authors' Contributions

B. L., Q. H., and Z. W. conceptualized the study; B. L., M.D., and X. L. contributed to methodology; B. L., M.D., and Z. W. provided software; B. L. and M.D. contributed to data curation. All authors have read and agreed to the published version of the manuscript.

## Acknowledgments

This research was funded by the Hebei Province Social Science Fund Project (grant no. HB20GL039).

## References

- [1] K. Jerzy, K. P. Anna, and L. Wojciech, "Identifying symptoms of bankruptcy risk based on bankruptcy prediction models—a case study of Poland[J]," *Sustainability*, vol. 14, no. 3, 2022.
- [2] K. Fang and Y. Yang, "SGL-SVM with its application in forecasting corporate financial distress[J]," *Statistica*, vol. 35, no. 08, pp. 104–115, 2018.
- [3] Q. Chen, "Research on financial risk warning of listed enterprises based on PCA-NBC algorithm," *Management & Technology of SME*, vol. 27, no. 12, pp. 85–87, 2021.
- [4] F. Fang, "Research on the risk warning of de-listing of listed companies: a supporting vector machine prediction model based on principal component analysis[J]," *Frontiers in Economics and Management*, vol. 2, no. 4, 2021.
- [5] H. Huang, Q. Xu, Y. Zhang, and H. Shi, "Weight-LSSVM financial crisis early warning model based on KPCA dimensionality reduction [J]," *Statistics & Decisions*, vol. 36, no. 20, pp. 180–184, 2020.
- [6] J. Li, S. Ma, M. Jin, and C.-hsien Chu, "Early warning model of crowdfunding defaults based on SA-SVM [J]," *Journal of Statistics and Information*, vol. 33, no. 11, pp. 70–77, 2018.
- [7] Q. Zhou, L. Wang, L. Juan, S. Zhou, L. Li, and J. Ma, "The study on credit risk warning of regional listed companies in China based on logistic model," *Discrete Dynamics in Nature and Society*, vol. 2021, Article ID 6672146, 2028 pages, 2021.
- [8] K. Luo and G. Wang, "The research on financial early warning based on the improved MRMR algorithms and cost sensitive classification[J]," *Journal of Statistics and Information*, vol. 35, no. 03, pp. 77–85, 2020.
- [9] A. Xiaoyan and J. Zhao, "Empirical analysis of financial early warning of resource-based listed companies," *Finance and Accounting Monthly*, vol. 39, no. 02, pp. 79–84, 2018.
- [10] X. Li, *Behavior Recognition Method Based on Acceleration sensor[D]*, Xi'an University of Science and Technology, Xi'an, China, 2017.
- [11] H. Xu, Ce Fang, X. Liu, and Z. Ye, "Network intrusion detection system based on improve moth-flame optimization algorithm [J]," *Journal of Computer Applications*, vol. 38, no. 11, pp. 3231–3235+3240, 2018.
- [12] S. S. Shreem, H. Turabieh, A. A. Sana, and F. Baothman, "Enhanced binary genetic algorithm as a feature selection to predict student performance," *Soft Computing*, vol. 26, 2022.
- [13] Y. Wang, Di Wang, and S. Wang, "Comparison of financial crisis early warning models based on PSO-BP and FOA-BP neural networks [J]," *Statistics & Decisions*, vol. 34, no. 15, pp. 177–179, 2018.
- [14] Z. Zhang, L. Fan, Q. Zhao, and G. Zhang, "Research of financial early-warning model on evolutionary support vector machines based on genetic algorithms[J]," *Discrete Dynamics in Nature and Society*, vol. 2009, no. Pt.3, 2010.
- [15] J. Zhang, C. Fan, and B. Wang, "Early warning study on financial risks of manufacturing companies based on PSO optimization SVM," *Friends of Accounting*, vol. 35, no. 14, pp. 52–56, 2017.
- [16] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Advances in Engineering Software*, vol. 95, pp. 51–67, 2016.
- [17] Z. Geng, Li Mu, S. Cao, and C. Liu, "A whale optimization algorithm based on hybrid revers learning strategy[J]," *Computer Engineering & Science*, vol. 44, no. 02, pp. 355–363, 2022.
- [18] F. Jiang, L. Wang, and L. Bai, "An improved whale algorithm and its application in truss optimization," *Journal of Bionics Engineering*, vol. 18, no. 3, pp. 721–732, 2021.
- [19] G. Kaur and S. Arora, "Chaotic whale optimization algorithm," *Journal of Computational Design and Engineering*, vol. 5, no. 3, pp. 275–284, 2018.
- [20] Y. Zhang, Y. Zhang, G. Wang, and L. Bingshan, "An Improved Hybrid Whale Optimization Algorithm Based on Differential Evolution," in *Proceedings of the 2020 International Conference on Artificial Intelligence and Electromechanical Automation (AIEA)*, pp. 103–107, IEEE, Tianjin, China, June 2020.
- [21] X. Shi and Li Ming, "Whale optimization algorithm improved effectiveness analysis based on compound chaos optimization strategy and dynamic optimization parameters," in *Proceedings of the 2019 International Conference on Virtual Reality and Intelligent Systems (ICVRIS 2019)*, pp. 123–126, Jishou, China, September 2019.