

Research Article

Market Risk Early Warning Based on Deep Learning and Fruit Fly Optimization

Liang Chen¹ and Rui Ma² 

¹School of Economics and Management, Southwest Petroleum University, Chengdu 610500, China

²Health and Rehabilitation College, Chengdu University of Traditional Chinese Medicine, Chengdu 610032, China

Correspondence should be addressed to Rui Ma; marui19840213@cdutcm.edu.cn

Received 4 March 2022; Revised 14 April 2022; Accepted 18 April 2022; Published 9 May 2022

Academic Editor: Man Fai Leung

Copyright © 2022 Liang Chen and Rui Ma. 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.

To improve the ability of market to avoid and prevent credit risk and strengthen the awareness of market risk early warning, SMOTE is used to process the unbalanced sample, and fruit fly optimization algorithm (FOA) is utilized to optimize the parameters of support vector machine (SVM), and thus an improved SVM market risk early warning model is proposed. The simulation results show that the proposed model has excellent stability and generalization ability, and it can predict market credit risk accurately. Compared with the prediction model based on FOA-SMOTE-BP and FOA-SMOTE-Logit, the proposed model performs better on the indicators of G value, F value, and AUC value, which provides a reference for market credit risk prediction.

1. Introduction

Market risk early warning is an important measure to prevent market risk and unknown loss and improve market normalization. In recent years, with the development of intelligent technology, deep learning has been widely used in various fields, including market risk warning. So far, the relevant researchers achieved market risk early warning through the use of deep learning. On the basis of in-depth study of rough set theory (RST), Guan et al. proposed a financial operation risk early warning model based on BP neural network, which effectively realizes the prediction of financial operation risk and profit risk of family farm [1]. In the cross-border environment, scholars put forward the marginal expected gap, delta conditional value at risk, and conditional capital gap to measure the system risk. The feature pair method based on bilateral balance sheet data is different from the paradoxical risk measurement method based on market price. Thus, a systemic risk early warning method based on the network spectral feature pair method analyzing the core global banking system is proposed. The method provides risk early warning for the unstable financial markets based on turning points similar to R numbers in

popular models [2]. Figini et al. improved the sample performance of parametric models and nonparametric models in credit risk estimation, so as to propose a multi-variable outlier detection technology based on local outliers, which can support financial institutions to make decisions and avoid falling into corporate credit risks [3]. In view of the credit risk assessment of the Internet finance industry, Yang and Yuan applied RBF network to analyze the statistical data of online peer-to-peer lending platform and evaluated the credit risk of the platform, and thus a new early warning method of RBF neural network model is proposed, which can reasonably predict the credit risk status of the industry development [4]. Based on the fuzzy theory and related theories of financial risk early warning management, Ding proposed the fuzzy comprehensive evaluation method, which realizes more accurate early warning and assessment of potential and obvious risks of financial enterprises. In addition, the safety of financial enterprise management is greatly improved, and the losses caused by various risks are reduced [5]. Based on BP neural network, Li constructed a risk assessment model of knowledge transfer in transforming enterprises to realize knowledge management risk warning [6]. Dong adopted the improved K-means

algorithm of quantum evolution to divide the risk warning interval by combining the given initial value and the value at risk measured by well-known Chinese online financial companies [7, 8]. Zhang and Chen used the autoregressive conditional Fréchet (ACF) model to predict the tail risk of the capital market, so as to identify major crisis sources [9]. Ouyang et al. applied the deep learning algorithm to the early warning of market risk. The results show that the algorithm has high accuracy compared with the traditional BP and others [10].

Among them, SVM algorithm is widely used in the field of classification because of its nonlinear and small sample advantages, but the parameter optimization of SVM is a research hotspot. For example, Jerlin Rubini and Perumal proposed to optimize the SVM algorithm by using the Drosophila algorithm and applied the optimized algorithm to the classification of chronic kidney disease, showing a good classification effect, indicating that the Drosophila algorithm has great advantages in optimizing SVM [11]. Tian and others used the fruit fly algorithm to optimize the echo state network, which greatly improved the accuracy of prediction [12]. Lu et al. applied the fruit fly algorithm and SVM to the prediction of urban gas load, which greatly improved the accuracy of short-term prediction. It can be seen from the above that the combination of Drosophila algorithm and SVM for classification or prediction has become the focus of current research [13]. The above early warning models based on deep learning realize the early warning of market risks to a certain extent. However, the prediction accuracy needs to be improved. To solve this problem, this paper applies the SVM model with excellent predictive performance and constructs a market risk early warning model by optimizing its parameters and unbalanced samples.

2. Basic Methods

2.1. SVM Model. SVM is a generalized linear classifier, which is proposed based on statistical learning theory and the principle of minimizing structural risk. Its basic principle is to construct an optimal hyperplane to maximize the distance between samples of two different categories, which is shown in Figure 1 [14]. Here, circles and squares represent two different types, respectively, and the optimal hyperplane is to maximize the range between the two dotted lines.

Suppose, dataset = $(x_i, y_i), i = 1, 2, \dots, n, x \in R, y \in \{-1, 1\}$, y is the category number. When $y = 1$, it means that x belongs to the first category. Also, when $y = -1$, it means that x belongs to the second category. Its linear discriminant function is usually expressed as [15]

$$g(x) = w' \cdot x + b, \quad (1)$$

where w' is the inertia weight and b is a constant.

The classification gap is $2/\|w'\|^2$. When $\|w'\|^2$ is minimum, the classification spacing is maximum. The form of standard SVM is [16]

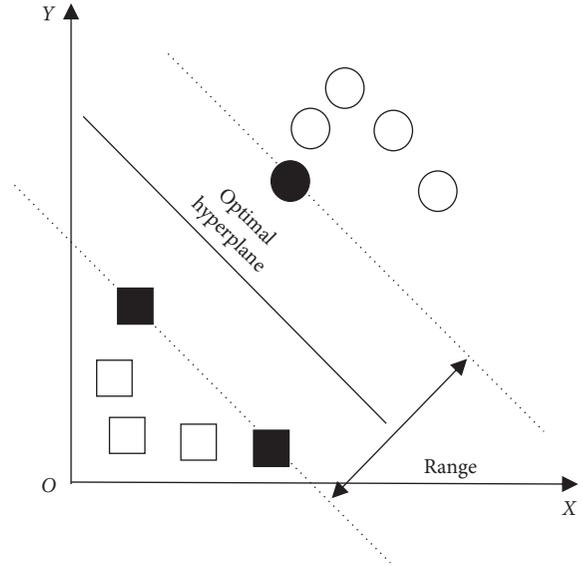


FIGURE 1: SVM for classification data example.

$$\min g(w') = \left(\frac{1}{2} \|w'\|^2 + c \sum_{i=1}^n \varepsilon_i^2 \right), \text{ s.t. } y_i (w'^T x_i + b^T x_i + b) \geq 1 - \varepsilon_i, \quad (2)$$

$$\varepsilon_i \geq 0, i = 1, 2, \dots, n,$$

where c is the penalty function and ε is the slack variable.

SVM is used to perform nonlinear transformation for undivided linear sample data, namely, $\varphi: R^d \rightarrow H$. Thus, the data sample space can be mapped in high-dimensional space. When solving, it should meet the requirement of [17]

$$Q(a) = \sum_{i=1}^n a_i - \frac{1}{2} \sum_{i,j=1}^n a_i a_j y_i y_j K(x_i, x_j), \quad (3)$$

where a_i is the Lagrange multiplier and $\sum_{i=1}^n y_i a_i = 0$. Since the above equation is constrained by the inequality, there is a unique optimal solution corresponding to the Lagrange multiplier a_i^* . The optimal classification discriminant function is [18]

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^n a^* y_i K(x_i, y_i) + b^* \right\}, \quad (4)$$

where b^* is the b value obtained from formula (2) and $K(x_i, x_i)$ is the kernel function. Take the radial basis function (RBF) as an example, which can be expressed as [19]

$$K(x, y) = \exp(-\sigma \|x - y\|^2), \quad (5)$$

where σ is the RBF kernel parameter. According to formulas (4) and (5), the optimal classification discriminant function is

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^n a^* y_i \exp(-\sigma \|x_i - y_i\|^2) + b^* \right\}. \quad (6)$$

As can be seen from the above analysis, the classification effect of the SVM model mainly depends on two aspects: one is whether the number of classification samples is balanced, and the other is whether the kernel parameters and penalty factors of the model are optimal, while the standard SVM model does not consider the situation [20]. Therefore, in order to improve the classification effect of the SVM model, this paper improves the model from the above two aspects.

2.2. SVM Model Improvements

2.2.1. Unbalanced Sample Processing. For the unbalanced classification samples, the Synthetic Minority Oversampling Technique (SMOTE) is used to deal with them from the data level [21]:

- (1) Determine a few sample categories X , calculate the Euclidean distance d between samples in X , and select K samples with the smallest distance d .
- (2) Sample X with the multiplier of $N = \text{minority sample} / \text{majority sample}$, and select $X_i (i = 1, 2, \dots, N)$ from K samples.
- (3) According to formula (7), X_i and X are synthesized into a new sample:

$$X_{\text{new}} = X + \text{rand}(0, 1) \times (X_i - X). \quad (7)$$

- (4) Combine X_{new} and X as a new training set to learn on the SVM model.

2.2.2. Optimization of Model Parameters. To optimize kernel parameters and penalty factors of the SVM model, this paper adopts fruit fly optimization algorithm (FOA) with high searching accuracy to process. Figure 2 shows simulated behavior of fruit fly foraging process [22]. The basic operation is as follows.

- (1) Initialize model maximum iteration, population size, fruit fly population location range (LR), and other parameters. In 2D coordinates (X, Y) , the initial position of each individual fruit fly is

$$\begin{aligned} X_{\text{axis}} &= \text{rand}(LR), \\ Y_{\text{axis}} &= \text{rand}(LR). \end{aligned} \quad (8)$$

- (2) Assign flight direction and distance to all fruit flies and olfactory search is utilized to update [23]

$$\begin{aligned} X_i &= X_{\text{axis}} + \text{rand}(FR), \\ Y_i &= Y_{\text{axis}} + \text{rand}(FR), \end{aligned} \quad (9)$$

where FR represents the single flight range of fruit fly.

- (3) According to formula (12), the distance between the individual position of fruit fly and the origin is calculated [24].

$$\text{DIST}_i = \sqrt{X_i^2 + Y_i^2}. \quad (10)$$

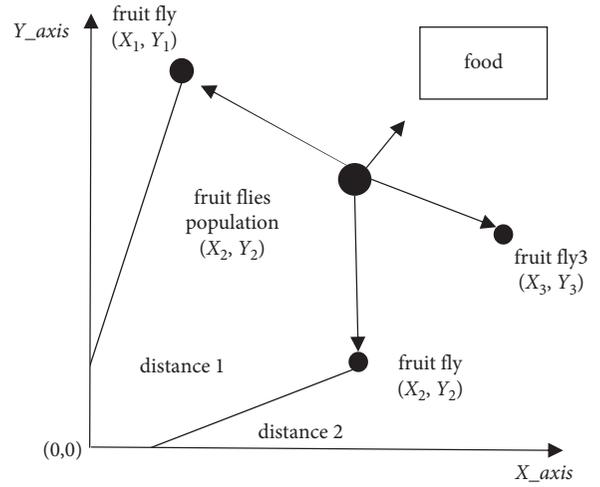


FIGURE 2: Simulation of fruit fly foraging behavior.

- (4) Smell_i and S_i are calculated according to fitness:

$$\text{Smell}_i = \text{fitness}(S_i),$$

$$S_i = \frac{1}{\text{DIST}_i}, \quad (11)$$

where fitness is the discriminant function, smell_i is the flavor concentration value, and S_i is the judgment value of smell.

- (5) Update bestSmell and bestIndex:

$$[\text{bestSmell}, \text{bestIndex}] = \min(\text{Smell}). \quad (12)$$

- (6) Use visual search to make other fruit flies fly to the best position [25]:

$$\begin{aligned} \text{SmellBest} &= \text{BestSmell}, \\ X_{\text{axis}} &= X(\text{bestIndex}), \\ Y_{\text{axis}} &= Y(\text{bestIndex}). \end{aligned} \quad (13)$$

- (7) Repeat steps (2)–(6) until the algorithm iterates to the set number.

FOA guides the search by the current optimal solution and makes the result close to the optimal solution, so as to realize the parameter optimization.

3. Market Credit Risk Early Warning Model Based on Improved SVM

S_{min} and S_{maj} are used to represent the samples of market credit risk and noncredit risk, and S is the set of all samples. Based on the above improvements, the construction process of the market credit risk early warning model is designed as follows:

- (1) Calculate k nearest neighbor points of each sample point $(x_{\text{smi}}, y_{\text{smi}})$ in S_{min} , randomly select a neighbor point $|S_{\text{maj}} - S_{\text{min}}|/2$ to subtract $(x_{\text{smi}}, y_{\text{smi}})$ and multiply it by the random number δ in the

interval $[0, 1]$, and then add (x_{smin}, y_{smin}) to obtain a new credit risk sample x_{new} , and thus there is

$$x_{new} = x_{smin} + (\overline{x_{min}} - x_{smin}) \times \delta. \quad (14)$$

- (2) Repeat the above steps until the number of x_{new} reaches $|S_{maj}-S_{min}|/2$.
- (3) Initialize relevant parameters of SVM and FOA. In this paper, referring to reference [26], the maximum iteration number of FOA is set to 100, and the population size is set to 20.
- (4) Use FOA to optimize the parameters of the SVM model, and the judgment value of flavor concentration is calculated according to $DIST_i = \sqrt{X_i^2 + Y_i^2}$ and $S_i = 1/DIST_i$.
- (5) Continue to iterate until the optimal bestsmell is less than the set value, and then the value is the optimal parameter.
- (6) Plug optimal parameters and x_{new} to construct the improved SVM model and perform prediction.

The above process is illustrated in Figure 3.

4. Simulation Experiment

4.1. Experimental Environment Construction. This experiment is run on 64-bit Windows 7 professional edition system. The CPU is Intel(R)Xeon(R) e5-2620v3 2.40ghz, and the GPU is Tesla K80. In addition, the memory is 16G. The model was built with MATLAB2018a.

4.2. Data Sources and Preprocessing

4.2.1. Data Sources. The financial data of 260 listed manufacturing enterprises in Shenzhen and Shanghai from 2018 to 2020 are selected as the experiment data. Through references [21, 27, 28], there are a total of 20 financial indicators selected as credit risk warning indicators of listed companies, including 6 first-level indicators such as enterprise operation capacity, growth capacity, profitability, and so on, and 20 second-level indicators such as total asset turnover rate, net asset growth rate, return on net asset, and so on. The indicators are listed in detail in Table 1.

4.2.2. Data Preprocessing

(1) *Descriptive Statistics.* Since there are significant differences between the mean of the above indicator variables and standard deviation and maximum and minimum values, descriptive statistics of indicator variables are carried out, and the results are shown in Table 2.

(2) *Normalization.* Considering the dimensional level of index variables, z-score is adopted for normalization, which is shown in the following formula:

$$x_j^* = \frac{x_j - \mu(x_j)}{\sigma(x_j)}, \quad (15)$$

where $\mu(x_j)$ and $\sigma(x_j)$ represent the sample mean and standard deviation corresponding to index j , respectively.

After normalization, descriptive statistics of each indicator variable are shown in Table 3.

(3) *Significance Testing.* For the selection of indicators that can distinguish credit risks and nonuse risks of listed companies, this paper adopts independent sample T -test to test them, and the results are shown in Table 4. The P value of 7 indicators, such as net asset-liability ratio and operating profit growth rate, is more than 10%, which indicates that it is unable to distinguish the credit risk and noncredit risk, so it is deleted in this paper.

4.3. Evaluation Indicators. Set average accuracy (G), F value, and AUC are used to evaluate the prediction performance of model. Confusion matrix is used to represent the dichotomous dataset of credit risk, which is shown in Table 5.

The model sensitivity (SE), specificity (SP), and precision (P) can be calculated as follows:

$$\begin{aligned} SE &= \frac{|TSmin|}{(|TSmin| + |TSmin|)}, \\ SP &= \frac{|TSmin|}{(|TSmin| + |TSmin|)}, \\ P &= \frac{|TSmin|}{(|TSmin| + FSmin)}. \end{aligned} \quad (16)$$

Through the above three indicators, the following can be calculated:

$$G = 0.1044, \quad (17)$$

$$F_measure = 2 * SE * \frac{SP}{(SE + P)}$$

The larger the selected index value is, the better the model performance is.

4.4. Experimental Results

4.4.1. Model Verification. The samples are divided into training sets and testing sets according to different proportions, and experiments are carried out under different kernel functions and different optimal parameter values. The results are shown in Table 6. Under the division condition of different sample proportion, models corresponding to different kernel functions and optimal parameter values perform well in G value, F value, and AUC value, and the differences are small, which indicates that the proposed model has good prediction performance and strong generalization ability [29].

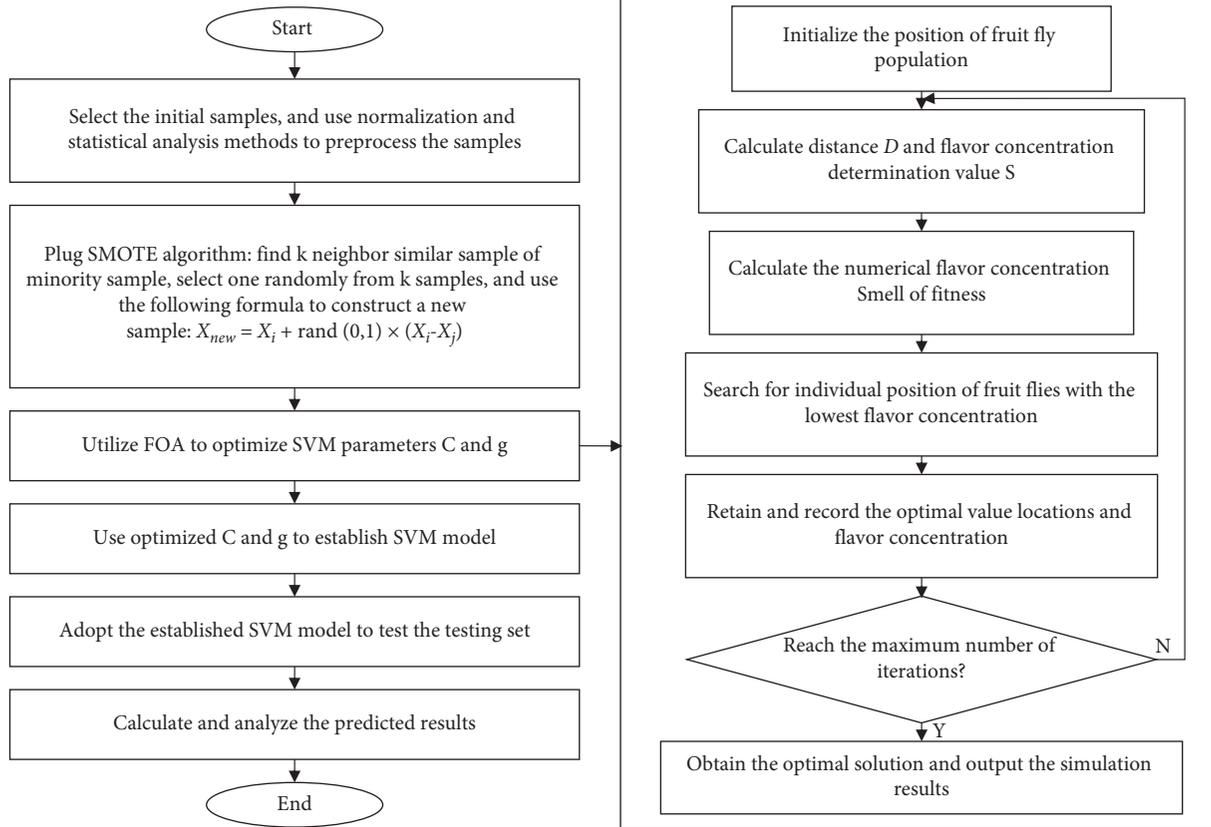


FIGURE 3: Prediction process of market credit risk early warning model based on improved SVM.

TABLE 1: Index selection.

Index type	Index name
Debt paying ability	Current ratio (X_1)
	Quick ratio (X_2)
	Net asset-liability ratio (X_3)
Operation ability	Fixed asset turnover ratio(X_4)
	Total asset turnover ratio (X_5)
	Receivable turnover ratio (X_6)
Profit ability	Return on net assets (X_7)
	Return on total assets (X_8)
	Net interest rate in sale (X_9)
Growth ability	Growth rate of basic earnings per share (X_{10})
	Growth rate of total operating revenue (X_{11})
	Operating profit growth rate (X_{12})
	Growth rate of net assets (X_{13})
Index per share	Earning per share (X_{14})
	Net assets per share (X_{15})
	Retained earnings per share (X_{16})
	Operating revenue per share (X_{17})
Asset structure	Asset-liability ratio (X_{18})
	Rights multiple (X_{19})
	Capital fixed ratio (X_{20})

In order to more intuitively reflect the prediction performance of different kernel functions and optimal parameter values under different training sample proportions, the prediction results in the above table are plotted in

Figures 4–6. Figure 4 shows that under different ratios of training set and testing set, the fluctuation range of G values of different kernel functions is small. Compared with the sigmoid and polynomial models, linear and RBF models have higher G values, which indicates that linear and RBF models have slightly better performance. On the other hand, Figure 5 shows that under different proportions of training sets and testing sets, F values of different kernels fluctuate greatly, but the overall F value is high. There is no trend indicating that the model F value of certain kernels type has the highest value, and thus the proposed model has good generalization ability. As can be seen from Figure 6, the AUC value of the proposed model fluctuates greatly, but the proposed model also achieves good results on this index. To sum up, the model proposed in this paper has good generalization ability and good prediction performance.

Considering that evaluating the model performance only through evaluation indexes lacks a certain scientific character, paired sample T test is adopted to test the prediction performance of different kernel function models, and the results are shown in Table 7. On the G value, the statistics of RBF and polynomial model, linear, and sigmoid model are all less than 10%, which rejects the null hypothesis, indicating that the performance of RBF and polynomial model is significantly different from that of linear and sigmoid model. On the F value, the statistics of all kernel function models are less than 10%, and all accept the null hypothesis, indicating that the performance of different kernel function models is less

TABLE 2: Descriptive statistical results of indicators.

Financial indicators	Mean	Standard deviation	Minimum value	Maximum value
X ₁	2.5217	2.9850	0.15	33.17
X ₂	2.0907	2.7981	0.13	33.17
X ₃	1.4701	1 1.5760	-151.94	231.66
X ₄	3.5915	4.4036	0.17	47.23
X ₅	0.7471	0.3703	0.03	2.64
X ₆	9.1016	9.0790	0.00	108.47
X ₇	2.7914	134.5795	-3106.24	871.51
X ₈	7.89869	7.8307	-60.11	48.03
X ₉	7.8415	15.9206	-222.31	153.66
X ₁₀	-35.1845	549.6023	-6200.00	4550.49
X ₁₁	14.9435	37.1203	-90.05	467.03
X ₁₂	1857.6662	52257.2278	-41692.61	1398352.86
X ₁₃	25.4540	76.0995	-1281.86	691.26
X ₁₄	0.4146	0.6626	-5.75	4.37
X ₁₅	4.7500	3.7401	-3.04	70.53
X ₁₆	1.6847	2.6542	-5.95	51.75
X ₁₇	6.1740	5.9962	0.06	72.55
X ₁₈	38.5987	23.5712	3.03	326.20
X ₁₉	2.6128	9.8620	0.00	232.68
X ₂₀	136 0941	769.6847	-12085.50	13048.27

TABLE 3: Descriptive statistics of index variables after normalization.

Financial indicators	Mean	Standard deviation	Minimum value	Maximum value
X ₁	0	1	-0.7976	10.2645
X ₂	0	1	-0.7032	11.1042
X ₃	0	1	-13.25121	19.8845
X ₄	0	1	-0.7785	9.9074
X ₅	0	1	-1.9700	5.0861
X ₆	0	1	-1.0026	10.9445
X ₇	0	1	-23.1019	6.4551
X ₈	0	1	-8.6845	5.1242
X ₉	0	1	-14.4564	9.1595
X ₁₀	0	1	-11.2170	8.3437
X ₁₁	0	1	-2.8287	12.1791
X ₁₂	0	1	-0.8335	26.7232
X ₁₃	0	1	-17.1797	8.7492
X ₁₄	0	1	-9.2881	5.9572
X ₁₅	0	1	-2.0798	17.5859
X ₁₆	0	1	-2.8729	18 8525
X ₁₇	0	1	-1.0214	11.0679
X ₁₈	0	1	-1.5097	12.2011
X ₁₉	0	1	-0.2650	23.3276
X ₂₀	0	1	-15.8788	16.7760

TABLE 4: Independent sample T-test results of index variables.

Financial indicators	T value	P value	Financial indicators	T value	P value
X ₁	9.015	0.000	X ₁₁	2.163	0.032
X ₂	9.958	0.000	X ₁₂	0.174	0.864
X ₃	0.223	0.827	X ₁₃	2.772	0.014
X ₄	2.118	0.035	X ₁₄	3.862	0.002
X ₅	2.533	0.012	X ₁₅	3.601	0.000
X ₆	2.085	0.038	X ₁₆	4.695	0.000
X ₇	1.711	0.105	X ₁₇	1.246	0.215
X ₈	7.660	0.000	X ₁₈	-4.182	0.002
X ₉	6.117	0.000	X ₁₉	-1.504	0.153
X ₁₀	1.588	0.132	X ₂₀	0.176	0.863

TABLE 5: Confusion matrix of credit risk dataset.

	Judged as noncredit risk sample	Judged as credit risk sample
Actually a noncredit risk sample	$ TS_{maj} $	$ FS_{min} $
Actually a sample of credit risk	$ FS_{maj} $	$ TS_{min} $

TABLE 6: Prediction results.

Kernel function	Proportion of training samples (%)	Proportion of testing samples (%)	Optimal parameter value	G	F	AUC
RBF	90	10	$c = 0.6509, g = 0.6666$	0.6954	0.2023	0.6071
	80	20	$c = 0.0836, g = 0.0806$	0.7659	0.2224	0.7946
	70	30	$c = 0.0908, g = 0.0798$	0.7083	0.4576	0.7338
	60	40	$c = -0.7157, g = -0.6259$	0.7578	0.2888	0.7821
	50	50	$c = 0.1384, g = 0.1355$	0.7151	0.2143	0.7490
Polynomial	90	10	$c = 2.1547, g = 0.7367, d = 6.8425, r = 0.9121$	0.7071	0.2667	0.6843
	80	20	$c = 3.0080, g = 0.8021, d = 2.0911, r = 1.1900$	0.6157	0.2222	0.5429
	70	30	$c = 4.5994, g = 0.4957, d = 0.9812, r = 1.8810$	0.6279	0.4000	0.5514
	60	40	$c = 2.6578, g = 0.7339, d = 1.8111, r = 2.7077$	0.6684	0.0909	0.5786
	50	50	$c = 0.2456, g = 0.0897, d = 0.0835, r = 0.0810$	0.6673	0.2353	0.5677
Linear	90	10	$c = 1.6133$	0.7098	0.2349	0.7465
	80	20	$c = 0.1$	0.7641	0.2311	0.7935
	70	30	$c = 0.0887$	0.7761	0.4295	0.8024
	60	40	$c = 15.6392$	0.6488	0.2127	0.6530
	50	50	$c = 0.1044$	0.6954	0.2110	0.7108
Sigmoid	90	10	$c = 5.7636, g = 0.7344, r = 0.7683$	0.7011	0.2879	0.7243
	80	20	$c = 1.9080, g = 0.6945, r = 0.8052$	0.6970	0.1319	0.7154
	70	30	$c = 3.3143, g = 0.5031, r = 0.8505$	0.7061	0.3002	0.7212
	60	40	$c = 2.4913, g = 0.7543, r = 1.1081$	0.6267	0.2199	0.5436
	50	50	$c = 0.2223, g = 0.0758, r = 0.0757$	0.6443	0.1900	0.5981

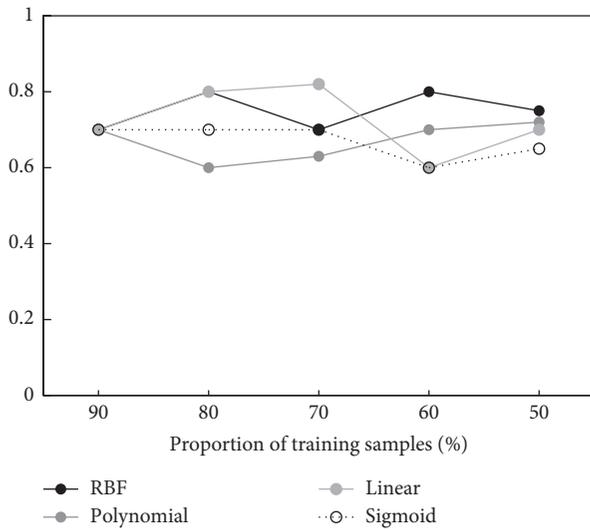


FIGURE 4: G values of the model under different kernel functions.

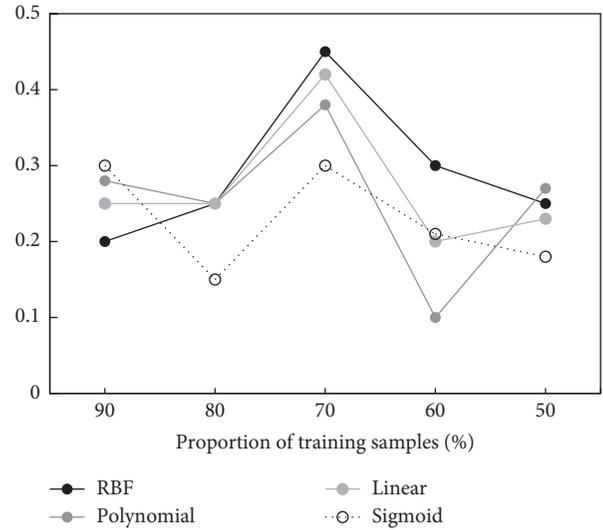


FIGURE 5: F values of the model under different kernel functions.

different. On the AUC value, the statistics of all models are less than 10%, which rejects the null hypothesis, indicating that the performance of linear and polynomial, sigmoid and polynomial, and RBF models is significantly different.

In conclusion, the change of kernel function has little influence on the prediction performance of the proposed model, which means that the prediction performance of the proposed model is relatively stable.

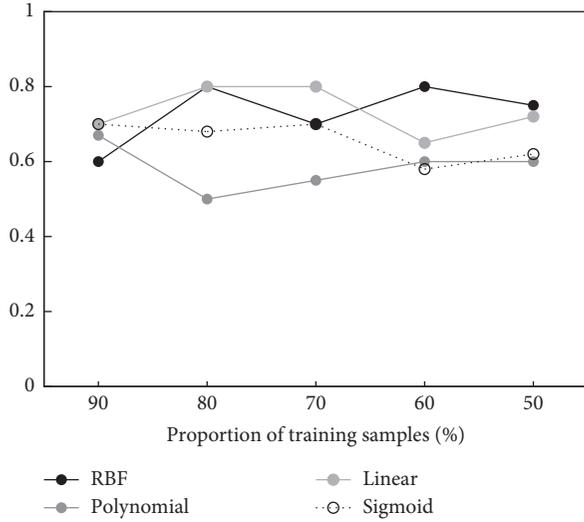


FIGURE 6: AUC values of the model under different kernel functions.

TABLE 7: Paired *T*-test results of model prediction performance under different kernel functions.

Evaluation indicators	Kernel function	Polynomial	RBF	Sigmoid
G	Linear	0.164	0.754	0.023
	Polynomial		0.055	0.531
	RBF			0.101
F	Linear	0.492	0.514	0.324
	Polynomial		0.495	0.708
	RBF			0.274
AUC	Linear	0.019	0.871	0.008
	Polynomial		0.062	0.140
	RBF			0.296

TABLE 8: Comparison of prediction performance of different models.

Model	G	F	AUC
FOA-SMOTE-SVM	0.7083	0.4576	0.7338
FOA-SMOTE-BP	0.6377	0.1760	0.6855
FOA-SMOTE-Logit	0.7026	0.2563	0.7219

4.4.2. Model Comparison. To further verify the validity of the proposed model, the prediction performance of the proposed model is compared with that of other models. The results are shown in Table 8. Compared with FOA-SMOTE-BP and FOA-SMOTE-Logit, the proposed model has the best performance on the indicators of G value, F value, and AUC value, indicating that the model proposed in this paper has the best prediction performance.

In order to observe the prediction performance of different models intuitively, the results of the above table are drawn as shown in Figure 7. It can be seen from the figure that the G value, F value, and AUC value curves of the proposed model are significantly higher than those of the

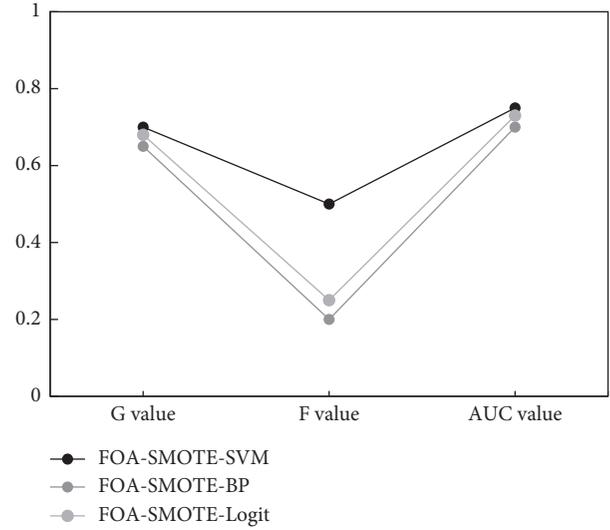


FIGURE 7: Comparison of every indicator of different models.

comparison model, indicating that the proposed model has better prediction performance.

5. Conclusion

In summary, the proposed market risk early warning method based on deep learning takes SVM as basic model and uses SMOTE to deal with unbalanced sample. In addition, minority samples are subjected to oversampling, and FOA is utilized to tune model parameters. Thus, the classification effect of the model is improved. The empirical results show that the proposed model has excellent stability and generalization ability and can accurately predict market credit risk. Compared with FOA-SMOTE-BP and FOA-SMOTE-Logit models, the proposed model performs better on G value, F value, and AUC value indicators and has better prediction performance, which provides a reference for market credit risk prediction research. The contribution of this study is to use a new improved SVM to predict the market risk. It provides a new reference for information management and prevention of the market. However, as the limitation of conditions, there are still some deficiencies to be improved. In the selection of market credit risk indicators, they are selected just through references, without considering the actual situation of China's manufacturing industry, which may affect the final market risk prediction results. To avoid the influence of index selection on prediction accuracy, the next research will try to independently determine the relevant index variables affecting market credit risk.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] Z. Guan, Y. Zhao, and G. Geng, "The risk early-warning model of financial operation in family farms based on back propagation neural network methods," *Computational Economics*, vol. 73, pp. 1–24, 2021.
- [2] S. Markose, S. Giansante, N. A. Eterovic, and M. Gatkowski, "Early warning of systemic risk in global banking: eigen-pair R number for financial contagion and market price-based methods[J]," *Annals of Operations Research*, vol. 163, pp. 1–39, 2021.
- [3] S. Figini, F. Bonelli, and E. Giovannini, "Solvency prediction for small and medium enterprises in banking," *Decision Support Systems*, vol. 102, pp. 91–97, 2017.
- [4] Y. Yang and Y. Yuan, "Research on credit risk early warning based on RBF neural network—taking Internet financial trading platform as an example[J]," *Modeling and Simulation*, vol. 10, no. 02, pp. 257–267, 2021.
- [5] Q. Ding, "Risk early warning management and intelligent real-time system of financial enterprises based on fuzzy theory," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 4, pp. 6017–6027, 2021.
- [6] W. Li, "Research on project knowledge management risk early warning based on b neural network," *Journal of Physics: Conference Series*, vol. 1744, no. 3, pp. 032250–032255, 2021.
- [7] M. S. Dong, "Intelligent early warning of Internet financial risks based on mobile computing," *International Journal of Mobile Computing and Multimedia Communications*, vol. 11, no. 2, pp. 61–78, 2020.
- [8] K. Noureddine, "Using multivariate stochastic dominance to enhance portfolio selection and warn of financial crises[J]," *The Quarterly Review of Economics and Finance*, vol. 80, pp. 480–493, 2021.
- [9] Z. Zhang and Y. Chen, "Tail risk early warning system for capital markets based on machine learning algorithms," *Computational Economics*, pp. 21–23, 2021.
- [10] Z. S. Ouyang, X. T. Yang, and Y. Lai, "Systemic Financial Risk Early Warning of Financial Market in China Using Attention-LSTM model," *North American Journal of Economics and Finance*, vol. 56, Article ID 101383, 2021.
- [11] L. Jerlin Rubini and E. Perumal, "Efficient classification of chronic kidney disease by using multi-kernel support vector machine and fruit fly optimization algorithm," *International Journal of Imaging Systems and Technology*, vol. 30, no. 3, pp. 660–673, 2020.
- [12] Z. Tian, "Echo state network based on improved fruit fly optimization algorithm for chaotic time series prediction[J]," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–20, 2020.
- [13] H. Lu, M. Azimi, and T. Iseley, "Short-term load forecasting of urban gas using a hybrid model based on improved fruit fly optimization algorithm and support vector machine," *Energy Reports*, vol. 5, no. C, pp. 666–677, 2019.
- [14] F. Meng, "Safety warning model of coal face based on fcm fuzzy clustering and GA-BP neural network," *Symmetry*, vol. 13, no. 6, pp. 1082–1087, 2021.
- [15] V. Coudert and J. Idier, "Reducing model risk in early warning systems for banking crises in the euro area," *International Economics*, vol. 156, pp. 98–116, 2018.
- [16] J. Beutel, S. List, and V., "Does machine learning help us predict banking crises?" *Journal of Financial Stability*, vol. 45, no. C, pp. 100693–100698, 2019.
- [17] E. Markus and F. Presbitero Andrea, "Commodity prices and banking crises[J]," *Journal of International Economics*, vol. 131, pp. 103474–103479, 2021.
- [18] T. Jiang, M. Zhou, W. Xuan, S. Wen, T. Shao, and J. Lu, "Dynamics in bank crisis model," *Mathematical Problems in Engineering*, vol. 2015, no. Pt.2, pp. 1–5, 2015.
- [19] I. Ilić, S. Pizarov, and P. S. Schmidt, "Preaching water but drinking wine? Relative performance evaluation in international banking[J]," *Swiss Journal of Economics and Statistics*, vol. 155, no. 1, pp. 1–25, 2019.
- [20] X.-A. Bi and Y. Xie, H. Wu and L. Xu, Identification of differential brain regions in MCI progression via clustering-evolutionary weighted SVM ensemble algorithm," *Frontiers of Computer Science*, vol. 15, no. 6, pp. 156903–156909, 2021.
- [21] J. Fan and G. Huang, M. Chi, Y. Shi, J. Jiang, C. Feng, and Z. Xu, Prediction of chemical reproductive toxicity to aquatic species using a machine learning model: an application in an ecological risk assessment of the Yangtze River, China," *The Science of the Total Environment*, vol. 796, pp. 148901–148908, 2021.
- [22] X. Li, J. Yang, Y. Fan, M. Xie, X. Qian, and H. Li, "Rapid monitoring of heavy metal pollution in lake water using nitrogen and phosphorus nutrients and physicochemical indicators by support vector machine," *Chemosphere*, vol. 280, pp. 130599–130605, 2021.
- [23] S. İlkin, T. H. Gençtürk, and F. Kaya Gülağız, H. Özcan, M. A. Altuncu, and S. Şahin, HybSVM: bacterial colony optimization algorithm based SVM for malignant melanoma detection," *Engineering Science and Technology, an International Journal*, vol. 24, no. 5, pp. 1059–1071, 2021.
- [24] H. Hu and J. Li, "Research on reference crop evapotranspiration forecast based on FOA-GRNN[J]," *International Core Journal of Engineering*, vol. 7, no. 5, pp. 108–116, 2021.
- [25] C. Lan, S. Li, H. Chen, W. Zhang, and H. Li, "Research on running state recognition method of hydro-turbine based on FOA-PNN," *Measurement*, vol. 169, pp. 108498–110852, 2021, (prepublish).
- [26] L. Li, H. Yang, L. Jiang, J. Gu, and Y. Zhang, "Optimal measurement area determination algorithm of articulated arm measuring machine based on improved FOA," *Measurement and Control*, vol. 53, no. 9-10, pp. 2146–2158, 2020.
- [27] V. Jaiswal, A. Negi, and T. Pal, "A review on current advances in machine learning based diabetes prediction," *Primary Care Diabetes*, vol. 15, no. 3, pp. 435–443, 2021.
- [28] Y. Tian, C. Chen, X. Chen, X. Chen, and R. Sun, "Research on real-time analysis technology of urban land use based on support vector machine," *Pattern Recognition Letters*, vol. 133, no. prepublsh, pp. 320–326, 2020.
- [29] D. Sun, B. Wang, X. Hu, and W. Wang, "A fault diagnosis method based on improved pattern spectrum and FOA-SVM," *International Journal of Acoustics and Vibration*, vol. 24, no. 2, pp. 312–319, 2019.