

# Integrating AOA and SVM for Multidimensional Identification and Prediction of Corporate Financial Risk

Fangyuan Ge

Lecturer, Department of Economic Management, Linyi Vocational University of Science and Technology, Linyi, 276000, China, E-mail: FangyuanGee@outlook.com

Project Management

Received February 3, 2026; revised March 25, 2026; accepted September 12, 2026  
Available online September 20, 2026

**Abstract:** The current multidimensional identification and prediction methods for corporate financial risks heavily rely on manual experience, resulting in insufficient reliability of the results. To this end, a multidimensional identification and prediction model for enterprise financial risks is proposed by integrating arithmetic optimization algorithms and Support Vector Machine (SVM) to accurately predict enterprise financial difficulties. This study first uses indicator-screening algorithms to extract key indicators that reflect the financial status of enterprises from large volumes of financial data. Then, an arithmetic optimization algorithm is used to optimize the support vector machine model, which is combined with the extreme gradient boosting model to perform multidimensional identification and prediction of enterprise financial risks. The experimental results show that the model achieves recognition accuracy of 97.8%, F1 score of 92.7%, prediction accuracy of 98.4%, and a missed recognition rate of only 1.62%. In practical applications, the proposed model successfully identified 130 low-risk enterprises out of 135, with an overall prediction accuracy of 90.0%, significantly better than the comparison model. Experimental results show that this model has significant advantages in identifying and predicting financial risks in enterprises and can provide accurate, efficient support for enterprise risk management.

**Keywords:** Arithmetic optimization algorithm; support vector machine (SVM); corporate finance; risk identification; risk prediction; financial risk indicators.

Copyright © Journal of Engineering, Project, and Production Management (EPPM-Journal).  
DOI 10.32738/JEPPM-2026-174

## 1. Introduction

Given the current trend of rapid socio-economic development, identifying and predicting corporate financial risks are among the core safeguards for avoiding capital chain ruptures, ensuring continuous operations, and achieving strategic goals (Xu et al., 2024). However, in traditional methods, extensive human experience and manual data identification and prediction often lead to large gaps between actual and predicted results (Dai et al., 2024). With the emergence of intelligent algorithms, more and more scholars are applying them to the multi-dimensional identification and prediction of corporate financial risks (Anand et al., 2024). Zhuang and Wu (2025) applied ChatGPT to corporate risk identification and prediction, helping enterprises better avoid risks and seize opportunities. Experimental results showed that this method could help enterprises accurately identify and predict crises, reducing risks in investment decision-making (Zhuang and Wu, 2025). Tyagi and Boyang (2023) applied intelligent Internet of Things technology in the fintech sector to identify and predict corporate financial risks, helping enterprises respond more effectively to crises. Experimental results showed that the proposed model outperformed comparative methods in terms of optimal cost, sensitivity, and accuracy. To accurately predict corporate financial distress, Tyagi and Boyang (2023) also proposed an efficient framework for high-dimensional, imbalanced scenarios. Experimental results showed that this method achieved significant gains in the area under the receiver operating characteristic curve and recall, enabling accurate prediction of corporate financial distress. Liu and Yang (2024) proposed a financial risk warning model for listed companies using a backpropagation neural network and rough set theory to improve prediction accuracy. The experimental results show that the model exhibits high accuracy and reliability in predicting financial risks of listed companies. Although the intelligent algorithm has been used in the field of multi-dimensional identification and prediction of enterprise financial risk and has achieved results, there are still some problems, such as difficulty balancing identification accuracy and prediction efficiency, and limited ability to handle high-dimensional, unbalanced financial data.

SVM is a supervised learning model that can analyze data and identify underlying patterns, demonstrating significant advantages in identification and prediction. For instance, Li et al. (2024) proposed an improved sailfish algorithm to optimize

SVM and applied it to predict the temperature of an insulated-gate bipolar transistor. The experimental evaluation showed that the correlation coefficient of the developed model was 99.60%, effectively ensuring the secure and stable operation of the component. The Arithmetic Optimization Algorithm (AOA) is a novel metaheuristic algorithm that can optimize Support Vector Machine (SVM) parameters and significantly improve model performance (Cao et al., 2023). For example, Wei et al. (2023) proposed a non-intrusive online charging detection method based on AOA-SVM to meet the requirements of load monitoring in substation-level power systems. Experimental results showed that this method exhibited strong adaptability and high accuracy under unknown household appliance load conditions. The above research indicates that SVM excels in data classification and pattern recognition, but parameter tuning relies on experience, while AOA excels in global parameter optimization. The two complement each other and can effectively overcome the problem of blind parameter selection in a single SVM model, thereby improving the model's stability and prediction accuracy.

To more effectively identify and predict financial risks for companies, a novel multidimensional model based on AOA and SVM is proposed. It is expected that this model can provide new insights and considerations for predicting corporate financial distress. To enhance the identification and prediction of corporate financial risks, this study proposes a novel multidimensional model that integrates AOA and SVM. The innovations of this model are reflected in two key aspects. First, the integrated application of algorithms such as AOA and SVM. Second, the development of a novel, high-performance multidimensional model for financial risk prediction.

## 2. Methodology

### 2.1. Method for Selecting Corporate Financial Risk Indicators

To conduct multidimensional identification and prediction of corporate financial risks, it is essential to first collect and analyze large volumes of financial data (Yan et al., 2023). Given the high computational cost and quality-assurance challenges associated with raw data, it is necessary to first establish key financial risk indicators to provide a scientific basis for subsequent research (Lyu and Jiao, 2025). Hence, for accurate multidimensional identification and prediction of corporate financial risk, it is necessary to identify key indicators. Principal Component Analysis-Grey Relational Analysis (PCA-GRA) combines two indicator screening methods. It extracts the most representative principal components from numerous financial indicators, effectively reduces data dimensionality, and removes redundant information among indicators. This process allows the model to focus more precisely on key factors that affect corporate financial risk (Eckhart et al., 2023; Zhang and Xu, 2024). Based on this, the study applies PCA-GRA to select corporate financial risk indicators and identify the most representative and critical ones. This step establishes a solid foundation for subsequent multidimensional identification and prediction of financial risk. The basic framework is shown in Fig. 1.

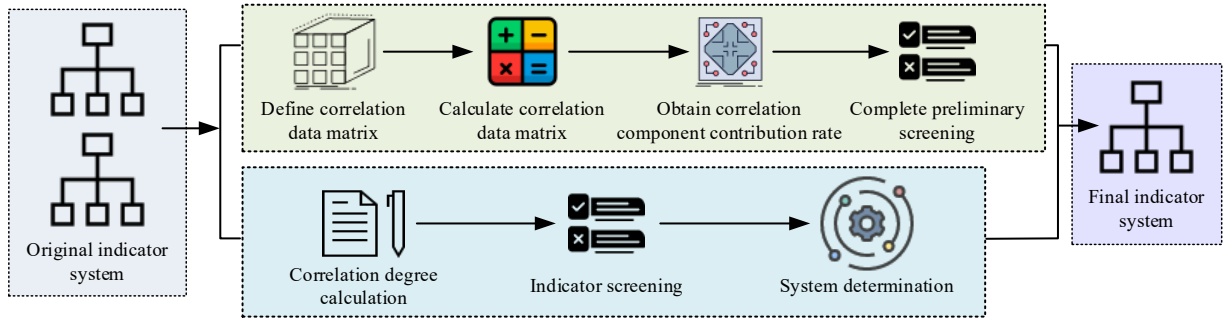


Fig. 1. PCA-GRA framework for corporate financial risk indicator selection

As shown in Fig. 1, this study first inputs the original corporate financial risk indicator system. It then applies PCA to define the corporate financial data matrix. The calculation is shown in Eq. (1).

$$X_{nm} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix} \quad (1)$$

In Eq. (1),  $m$  and  $n$  are the number of samples and the number of variables. This study then calculates the correlation coefficient matrix among financial indicators. The calculation is shown in Eq. (2).

$$W = X^{*T} X^* \quad (2)$$

In Eq. (2),  $X^*$  represents the matrix obtained after data processing. Next, this study calculates the eigenvalues of this matrix to obtain the contribution rates of correlated financial components. The calculation is shown in Eq. (3).

$$P = \sum_{i=1}^m \lambda_i \quad (3)$$

In Eq. (3),  $\lambda$  represents the eigenvalues of the correlation coefficient matrix. Based on the contribution rates, this study selects key components that have a greater impact on the identification and prediction of corporate financial risk. Through empirical analysis, the first three principal components are selected, which together account for 86.2% of the variance, satisfying the basic requirements for information retention during dimension reduction. Based on the three principal components, this study further conducts GRA analysis. Using the enterprise's financial risk status as the reference sequence, it calculates the correlation coefficients between the original indicators and each principal component. This process identifies the most significant indicators of financial risk, thereby establishing the final screening system. The specific calculation formula is presented in Eq. (4).

$$\xi_i(k) = \frac{\min_i \min_k \square_i(k) + \xi \min_i \max_k \square_i(k)}{\square_i(k) + \xi \max_i \max_k \square_i(k)} \quad (4)$$

In Eq. (4),  $\xi$  represents the resolution coefficient, with values in the range [0,1].  $\square_i(k)$  represents the absolute difference of indicator  $k$ . After calculation, financial indicators with lower relational degrees are removed. When it is not possible to determine whether  $\square_i(k)$  shows positive or negative correlation, this study applies the relevant Eq. for judgment. The calculation is shown in Eq. (5).

$$\sigma_i = \sum_{k=1}^n kx_i(k) - \sum_{k=1}^n x_i(k) \sum_{k=1}^n \frac{k}{n}, i = 1, 2, \dots, n \quad (5)$$

In Eq. (5), when  $\text{sign}(\sigma_i) = \text{sign}(\sigma_j)$  holds, the two indicators show positive correlation. When  $\text{sign}(\sigma_i) = -\text{sign}(\sigma_j)$  holds, they show a negative correlation. The corporate financial risk indicator system selected by the algorithm is shown in Table 1. On this basis, this study applies it to the screening of financial risk indicators and obtains a financial risk indicator system. This study selected relevant indicators from three aspects: profitability, debt-paying ability, and operational ability of the enterprise. In terms of profitability, the indicators selected through algorithms include return on equity, return on total assets, gross profit margin, and net profit margin. In terms of debt-paying ability, the algorithm selects four indicators: current ratio, quick ratio, asset-to-liability ratio, and equity ratio. In terms of operational capability, the algorithm selects four indicators: inventory turnover rate, accounts payable turnover rate, current asset turnover rate, and total asset turnover rate. These indicators can reflect the financial status and operating results of enterprises from different dimensions. By selecting and combining these indicators, research can more comprehensively and accurately identify and predict enterprises' financial risks.

## 2.2. Multidimensional Risk Identification Model Based on AOA-SVM

The established enterprise risk indicator system enables efficient identification of financial risks by filtering critical information from large volumes of financial data. However, manual analysis of these indicators is inefficient and inaccurate (Nacpil et al., 2023). To address this, this study develops a machine learning model based on this system, employing the SVM algorithm, renowned for its superior performance in nonlinear classification, to enhance both the precision and efficiency of financial risk identification. In addition, optimizing SVM parameters through optimization algorithms significantly improves its multidimensional identification performance (Khushaba et al., 2022). Sine Cosine Algorithm-AOA (SCA-AOA) is an improved version of AOA. Compared with traditional AOA, it uses the inherent oscillation of sine and cosine functions to search for optimal SVM parameters more effectively (Chen et al., 2024; Qian et al., 2023; Chiong et al., 2023). Therefore, this study applies SCA-AOA to optimize SVM parameters. The optimization process is shown in Fig. 2.

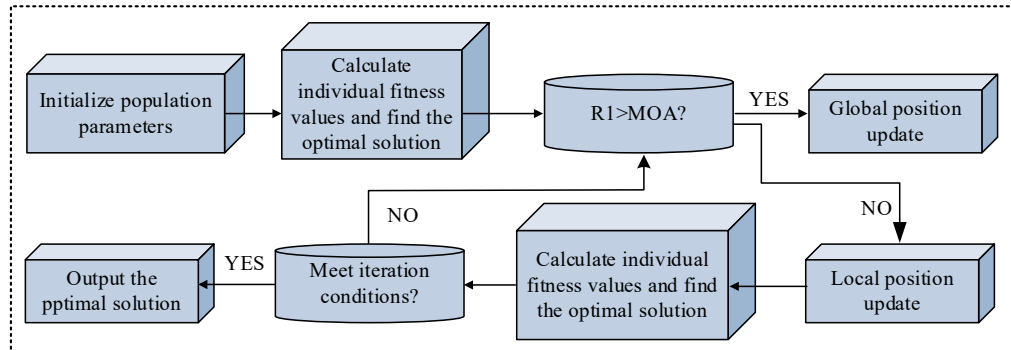


Fig. 2. Optimization process of SCA-AOA algorithm

As shown in Fig. 2, the algorithm first initializes the population and relevant parameters. It then calculates individual fitness values and determines the current optimal solution. Next, the algorithm determines whether to perform a local or global position update based on the judgment condition given by Eq. (6).

$$r_1 > MOA? \quad (6)$$

In Eq. (6),  $MOA$  represents the mathematical accelerator function. When the condition is satisfied, the algorithm performs global position update. The calculation is shown in Eq. (7).

$$x_{i,j}(t+1) = \begin{cases} bestx_j(t) / (MOP + \varepsilon) * ((UB_j - LB_j) * \mu + LB_j), r_2 < 0.5 \\ bestx_j(t) * MOP * ((UB_j - LB_j) * \mu + LB_j), otherwise \end{cases} \quad (7)$$

In Eq. (7),  $\varepsilon$  and  $\mu$  represent constants and control parameters, respectively.  $x_{i,j}(t+1)$  represents the position of the  $j$ -th dimension of the  $i$ -th individual at the  $(t+1)$ -th iteration.  $bestx_j(t)$  represents the position of the optimal solution.  $UB_j$  and  $LB_j$  represent the upper and lower bounds of the  $j$ -th dimension. When the condition is not satisfied, the algorithm performs global position update. The calculation is shown in Eq. (8).

$$x_{i,j}(t+1) = \begin{cases} w * x_{i,j}(t) + MOP * \sin(R_2) * |R_3 best_j(t) - x_{i,j}(t)|, R_4 < 0.5 \\ w * x_{i,j}(t) + MOP * \cos(R_2) * |R_3 best_j(t) - x_{i,j}(t)|, R_4 \geq 0.5 \end{cases} \quad (8)$$

In Eq. (8),  $w$  represents the introduced nonlinear inertia weight factor. The algorithm then recalculates individual fitness values to obtain the current optimal solution. It checks whether the iteration condition is met. If satisfied, the algorithm outputs the optimal SVM parameters. Otherwise, it continues to determine whether updates are global or local. After optimization with SCA-AOA, the SVM obtains optimal kernel parameters, thereby enhancing its ability to identify corporate financial risk. Finally, this study applies the optimized SVM to the multidimensional identification of corporate financial risk, thereby more accurately determining the optimal classification hyperplane and improving identification accuracy. The operational flow of the SCA-AOA-SVM multidimensional corporate financial risk identification model is shown in Fig. 3.

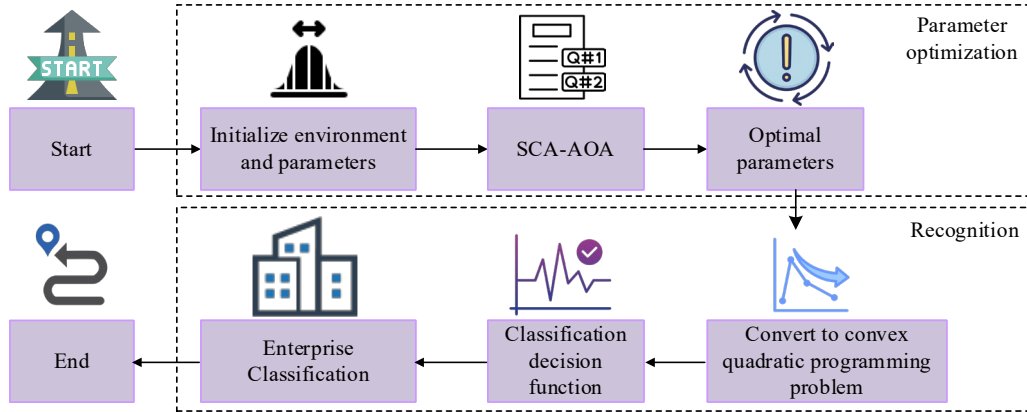


Fig. 3. SCA-AOA-SVM multidimensional corporate financial risk identification model

As shown in Fig. 3, the model first applies SCA-AOA to optimize SVM parameters and select the optimal kernel. This step improves SVM performance. The optimized SVM is then applied to the multidimensional task of corporate financial risk identification. Specifically, the model inputs preselected corporate financial risk indicator data and performs in-depth analysis using SVM's strong nonlinear classification capability. The model first maps financial data into a high-dimensional space, converting the task into a convex quadratic programming formulation. The calculation is shown in Eq. (9).

$$\begin{cases} \min_a \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M a_i a_j y_i y_j k(x_i, x_j) - \sum_{i=1}^M a_i \\ s.t \sum_{i=1}^M a_i y_i = 0 \end{cases} \quad (9)$$

In Eq. (9),  $a_i$  and  $a_j$  represent Lagrange multipliers.  $y_i$  and  $y_j$  represent the risky and non-risky enterprise classes in a binary classification problem. The classification decision function is obtained using the Lagrange multiplier method. The calculation is shown in Eq. (10).

$$f(x) = \text{sign} \left[ \sum_{i=1}^M a_i^* y_i k(x, x_i) + b^* \right] \quad (10)$$

In Eq. (10),  $k(x, x_i)$  is the kernel function, and  $b^*$  is the bias vector. The model ultimately outputs multidimensional identification results of corporate financial risk. It classifies enterprises into risky and non-risky categories and provides timely, effective risk warnings and decision support.

### 2.3. Risk Prediction Model Based on AOA-SVM

The multidimensional corporate financial risk identification model constructed by SCA-AOA-SVM enables preliminary identification and classification of corporate financial risk, as well as an understanding of enterprises' current risk status. However, this model cannot effectively predict future financial risk trends or the changes that might occur. To enhance this model for corporate financial risk forecasting, a forecasting model is necessary to identify potential risks in advance and provide strong support for corporate response strategies (Jahanshahi et al., 2024). Before prediction, historical data are combined to forecast financial risk trends by extracting time-series features from corporate risk data. Attention Mechanism-Long Short-Term Memory (AM-LSTM) integrates AM with LSTM to better capture long-term dependencies in time-series data. Therefore, this study applies AM-LSTM to corporate financial risk analysis to more accurately capture temporal variation trends in corporate financial risk. The AM-LSTM multi-period corporate financial time-series feature extraction model is shown in Fig. 4.

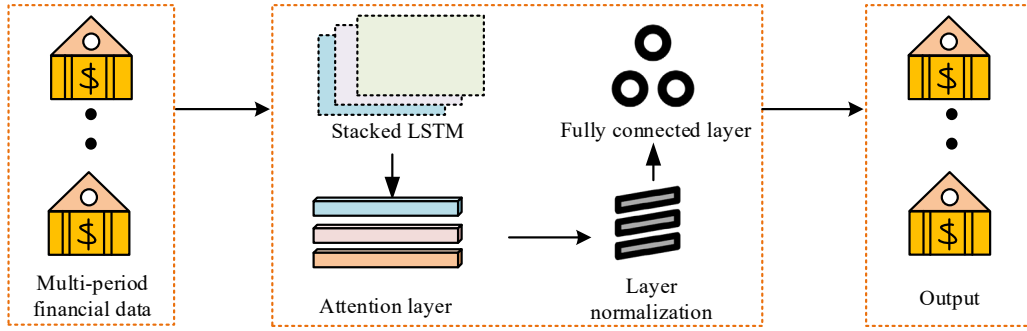


Fig. 4. AM-LSTM multi-period corporate financial time-series feature extraction model

As shown in Fig. 4, the model first takes as input raw multi-period corporate financial data and then undergoes stacked LSTM-based time-series feature extraction. Multi-period financial data first passes through the LSTM's forget gate, which retains key information. The calculation is shown in Eq. (11).

$$f_t = \sigma(w_{L1} [h_{t-1}, x_t] + b_{L1}) \quad (11)$$

In Eq. (11),  $w_{L1}$  represents weights,  $b_{L1}$  represents bias,  $x_t$  is the current financial data input, and  $h_{t-1}$  is historical financial data output from the previous unit. The multi-period data then pass through the input gate to filter key information in the current data, such as sharp declines in net profit or cash flow interruptions. The calculation is shown in Eq. (12).

$$\begin{cases} I = \sigma(w_i \bullet [h_{t-1}, x_t] + b_i) \\ \tilde{C}_t = \tanh(w_c \bullet [h_{t-1}, x_t] + b_c) \end{cases} \quad (12)$$

In Eq. (12),  $\tilde{C}_t$  is the candidate's memory generated for current financial fluctuations, and  $w$  represents the weight matrix. Finally, the output gate integrates processed historical and current information to generate a vector of features, including trends in changes of financial indicators and cyclical fluctuations. The calculation is shown in Eq. (13).

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (13)$$

In Eq. (13),  $W_o$  and  $U_o$  both represent weight matrices, and  $b_o$  represents the bias vector. After LSTM performs initial feature extraction, the model enters the attention layer to focus on key features in financial data. The hidden states of multi-period financial time series from the last LSTM layer are first used as input and transformed into query, key, and value vectors via three projection matrices. The attention mechanism then allocates weights to multi-period financial features. The calculation is shown in Eq. (14).

$$Attention = (Q_i, K_i, V_i) = \text{soft max} \left( \frac{Q_i K_i}{\sqrt{\xi_k}} \right) V_i \quad (14)$$

The outputs of multiple attention heads are then integrated through a linear transformation matrix to form the final multi-period financial feature representation. To improve convergence speed and enhance feature extraction capability, this study introduces residual connections and layer normalization to optimize the integrated feature representation. The calculation is shown in Eq. (15).

$$O = \text{LayerNorm}(A + H) \quad (15)$$

In Eq. (15),  $A$  and  $H$  represent the output of the multi-head attention mechanism and the multi-period financial time-series hidden states output by LSTM, respectively. Using multi-period financial feature extraction with AM-LSTM, the model predicts corporate financial risk. Because eXtreme Gradient Boosting (XGBoost) offers strong nonlinear fitting and high robustness to noise when processing high-dimensional financial features, integrating it with SCA-AOA helps identify optimal parameters that enable high-accuracy prediction of corporate risk status. Therefore, this study combines the AM-LSTM and SCA-AOA-XGBoost algorithms to predict the future financial risk status of enterprises, which is named the ALSAX model, composed of the first letters of the key algorithms in each core technology link. Its specific operational process is illustrated in Fig. 5.

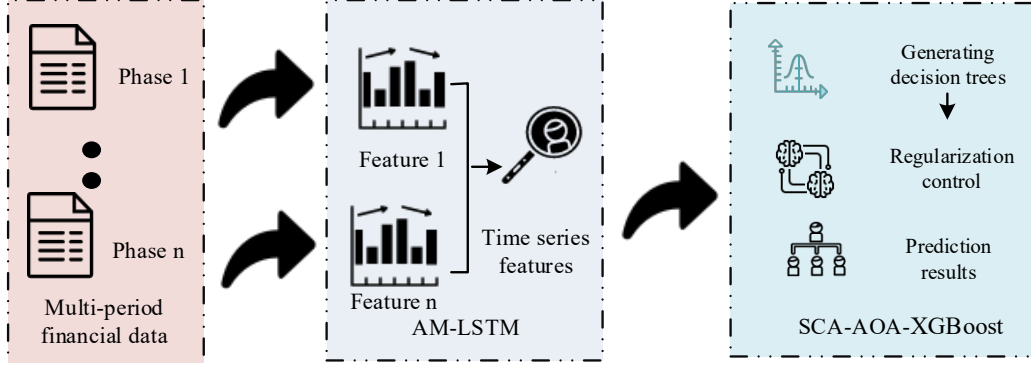


Fig. 5. Operational process of ALSAX corporate financial risk prediction model

In Fig. 5, the model first inputs multi-period corporate financial feature data into the AM-LSTM network. It then applies the strong time-series processing capabilities of LSTM to analyze the data in depth, extract key temporal features that reflect dynamic changes in corporate financial risk, and generate insights. These key temporal features are then input into SCA-AOA-XGBoost for corporate financial risk prediction. In the SCA-AOA-XGBoost model, SCA-AOA first optimizes XGBoost's core parameters. The model then integrates multiple classification and regression trees to perform probabilistic prediction of corporate financial risk status across different future time windows. It outputs risk probability scores and provides strong support for enterprises in formulating response strategies in advance and avoiding potential risks.

#### 2.4. Integration of Corporate Financial Risk Identification and Prediction Models

The designed multidimensional models for corporate risk identification and prediction can identify and predict such risks. Although key indicators are selected from the data, the absence of data or anomalies in the corporate financial statement input data remains unaddressed. Hence, mechanisms for data cleaning and anomaly detection are added just prior to corporate risk identification and prediction using corporate financial statement data. This efficiently resolves data gaps or anomalies in the input data. This study first removes outliers. It applies the Z-score method to map the distance of a financial data point from the mean, in standard deviation units, into a standardized score. Outliers are identified based on this score and removed. The calculation is shown in Eq. (16).

$$Z_A = \frac{(a - \eta'_A)a}{\eta_A} \quad (16)$$

In Eq. (16),  $Z_A$  and  $\bar{\eta}_A$  represent the standardized score of a single financial data point and the standard deviation, respectively.  $a$  represents a data point, and  $\eta'_A$  represents the mean value. After outlier removal, data cleaning is performed to repair incomplete financial data. This study applies the Mean Imputation Method (MIM) to impute missing values. The calculation is shown in Eq. (17).

$$x_k = \frac{1}{n} \sum_{i=1}^n x_i \quad (17)$$

In Eq. (17),  $n$  and  $x_i$  represent the number of non-missing values and observed values, respectively, and  $x_k$  represents the mean value used for imputation. Through data preprocessing, the quality of input data is significantly improved, providing a more reliable foundation for subsequent model identification and prediction. Finally, this study integrates the Z-score-MIM data preprocessing module, the PCA-GRA corporate financial risk indicator selection module, the AM-LSTM feature extraction module, the SCA-AOA-SVM identification module, and the SCA-AOA-XGBoost prediction module to construct an integrated multidimensional corporate financial risk identification and prediction model. This model enables accurate assessment of both current and future corporate risk status. Its operational process is shown in Fig. 6.

In Fig. 6, the model first collects multidimensional data from corporate financial data sources to provide a comprehensive and accurate basis for analysis. PCA-GRA then screens the raw data to remove redundant information and extract key indicators closely related to financial risk. Z-score-MIM preprocesses the data for these key indicators. The processed data

is then entered into the multidimensional corporate financial risk identification model and the prediction model. In the recognition model section, the SCA-AOA-SVM algorithm leverages its powerful nonlinear classification capabilities to analyze the enterprise's input financial risk indicator data in depth. Based on single-time-step features, point prediction is performed, and the enterprise's current risk status recognition result is output. In the prediction model section, the AM-LSTM network's excellent ability to process time-series data is combined with the SCA-AOA XGBoost algorithm. The sequence length is set to 10, and the number of features per time step is 4. The multi-head attention mechanism is used to capture temporal dependencies and achieve multi-domain prediction. The enterprise's multi-period financial feature data is mined and analyzed to obtain predictions for different future time windows. The model integrates current risk identification results and future risk prediction results to produce a comprehensive, multidimensional corporate financial risk assessment report, providing a solid foundation for sustainable corporate development.

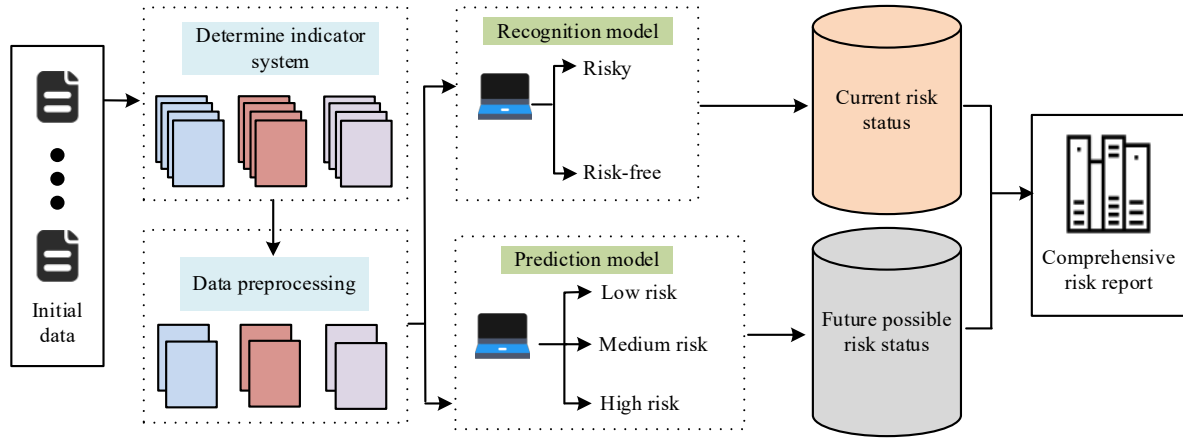


Fig. 6. Z-PASX multidimensional corporate financial risk identification and prediction integration model

### 3. Performance Evaluation

#### 3.1. Identification Performance Validation of SCA-AOA-SVM Model

To validate the proposed SCA-AOA-SVM multidimensional identification model for corporate financial risk, this study compared it with the Artificial Neural Network (ANN) model, the Genetic Algorithm-Optimized Random Forest (GA-RF) model, and the GA-SVM model. The input sequence length is set to 10, and the number of features per time step is 4. The multi-head attention mechanism is used to implement feature weighting, and multiple parallel attention heads capture dependencies in financial time series across different dimensions. The experimental environment was configured as follows: operating system, Windows 10/Linux 5.15.133; deep learning framework, PyTorch; debugger, GDB; programming language, Python 3.8; CPU, Intel Xeon E5-2680 v4; memory, 32 GB. To ensure the authenticity and reliability of this experiment, this study selected financial report data for the top 1,000 listed companies from 2015 to 2024 in the China Stock Market & Accounting Research Database (CSMAR) equity database to create a standardized dataset. This study is based on selecting data from listed companies from 2015 to 2024, which has three limitations: firstly, the sample only covers existing listed companies and does not include delisted or unlisted companies, resulting in sampling bias and survivor bias; second, the sample period covers macroeconomic shocks such as the COVID-19, which may lead to conceptual drift of financial risk characteristics; thirdly, due to database licensing and availability limitations, it is not possible to achieve full data public sharing. First, this study compares the Multi-dimensional Identification Accuracy (MIAcc) of all four models for corporate financial risk classification; the results are presented in Fig. 7.

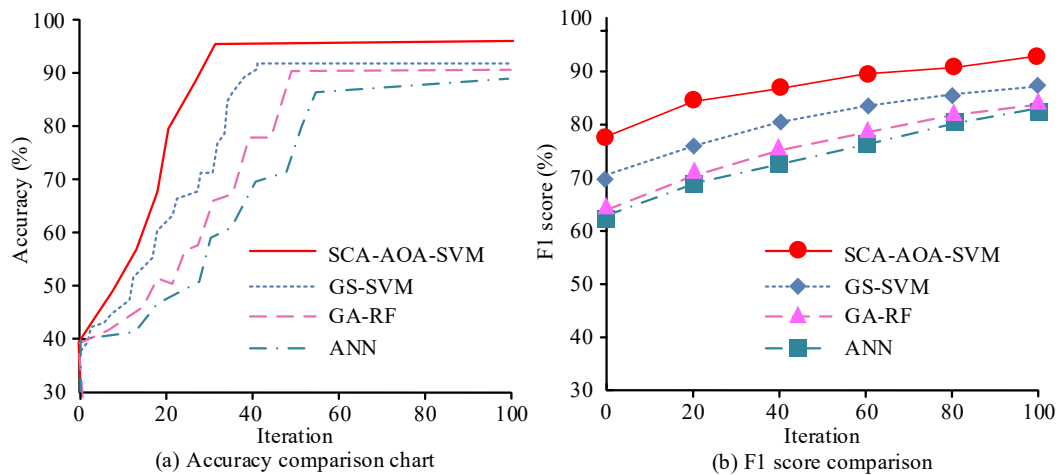


Fig. 7. Comparison of identification accuracy and F1 score results

The MIAcc recognition accuracy measures the proportion of correctly identified samples to the total sample size. The F1 score, the harmonic mean of precision and recall, provides a comprehensive evaluation of the model's classification performance on imbalanced data. As shown in Fig. 7(a), the identification accuracy of the SCA-AOA-SVM model increased rapidly with the number of iterations and reached 97.8% after iteration completion, significantly higher than the 91.2%, 90.1%, and 89.8% of the three comparative models. In Fig. 7(b), the F1 score of the proposed model reached 92.7% after iteration completion, while the F1 scores of the other three comparative models were below 90%, further demonstrating the good classification performance of the proposed model. The results show that the SCA-AOA-SVM model can capture key features in the data more effectively and achieve more accurate classification when addressing the multi-dimensional identification task of enterprise financial risk. The comparison among the four models concerning the iterative curve and the mistaken identification rate is shown in Fig. 8.

In Fig. 8(a), the SCA-AOA-SVM model had a loss of 1.91 after 20 iterations and 0.23 upon completion. The loss values of the three comparative models after iteration completion were all higher than those of the proposed model, especially the ANN model, which reached a loss of 1.21. In Fig. 8(b), the misjudgment rate of the SCA-AOA-SVM model reached a minimum of 2.02% at the end of training and outperformed the comparative algorithms throughout training. In conclusion, the proposed model demonstrates excellent convergence and a low misjudgment rate, providing reliable technical support for multi-dimensional and precise identification of corporate financial risks.

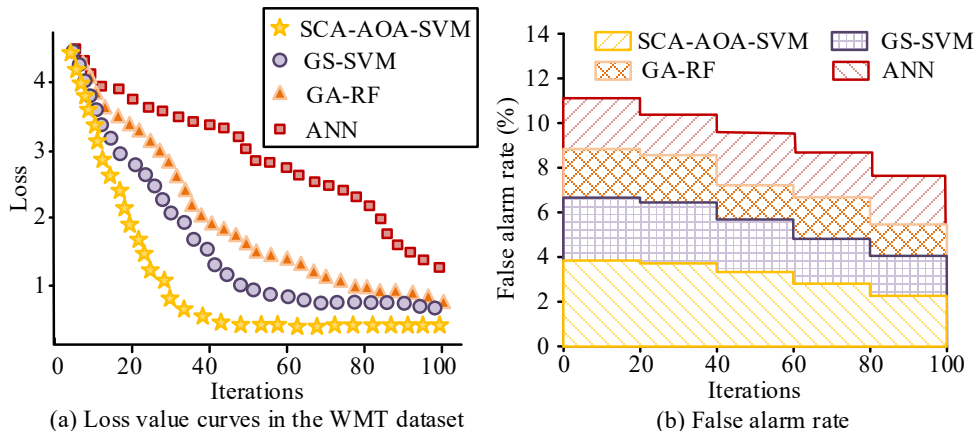


Fig. 8. Comparison of iteration curve and misjudgment rate results

### 3.2. Prediction Performance Validation of ALSAX Model

Based on the validation of the SCA-AOA-SVM model's identification performance, this study continued to validate the performance of the constructed ALSAX corporate financial risk prediction model. The experimental system used Windows 10; the operating system was Linux 5.15.133; the deep learning framework was PyTorch; the debugging tool was GDB; and the programming language was Python 3.8. The CPU was Intel Xeon E5-2680 v4, and the memory was 32 GB. To ensure the authenticity and reliability of the experiment, this study selected financial reports and annual report text information for 1,000 listed companies between 2015 and 2024 from the CSMAR database to construct a standardized dataset. This study compared it with prediction models constructed using three algorithms: ANN, GA-RF, and Logistic. First, this study compared the prediction accuracy and omission rate of the four models for corporate financial risk prediction, and the comparison results are shown in Fig. 9.

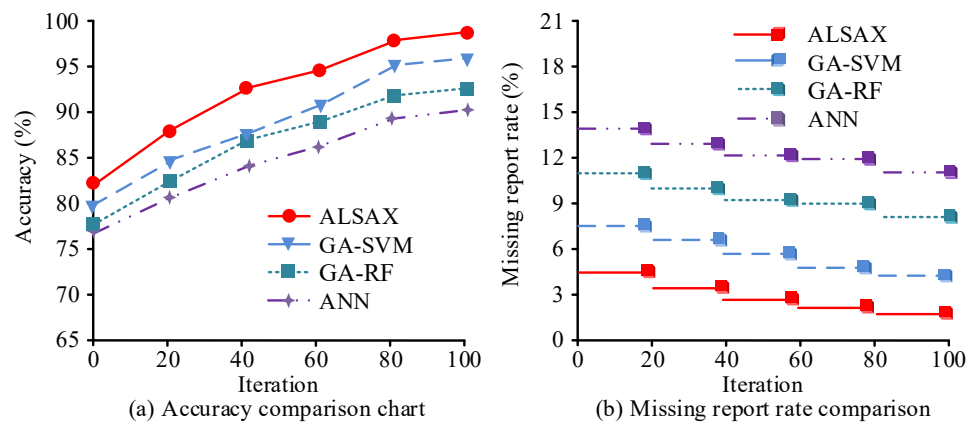
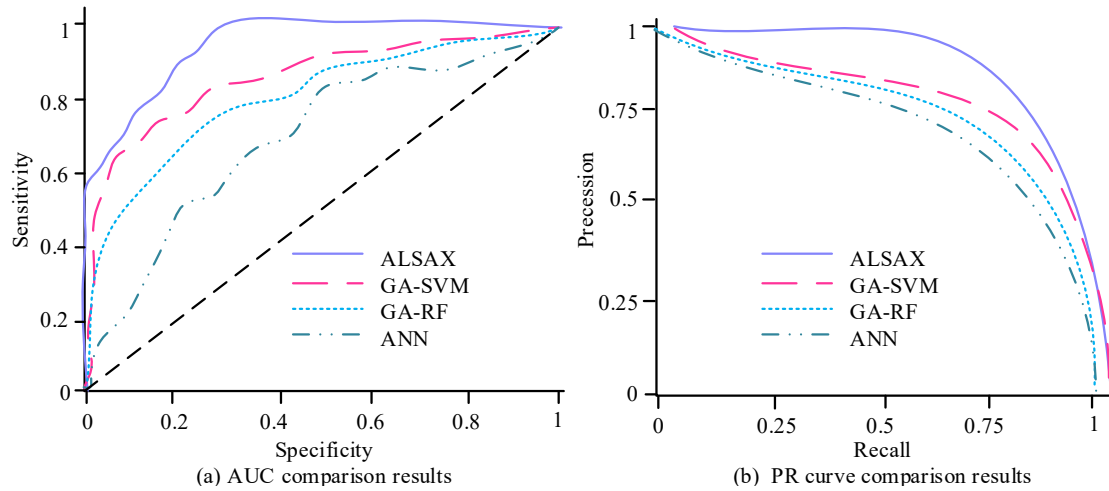


Fig. 9. Comparison of prediction accuracy and omission rate results

In Fig. 9(a), the accuracy of the ALSAX model reached 98.4% upon completion of the iterations, significantly higher than the 95.1%, 92.5%, and 90.1% of the three comparative models, demonstrating the model's strong predictive performance. In Fig. 9(b), the omission identification rate of the ALSAX model remained below 5% throughout and was only 1.62% after

iteration completion, while the three comparative models were far higher than this value. The experimental results indicated that the ALSAX model, with its excellent prediction performance and stability, provided a reliable technical means for corporate financial risk prediction. Next, this study compared the Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curves for the four models for corporate financial risk prediction, and the results are shown in Fig. 10.



**Fig. 10.** Comparison of ROC curve and PR curve results

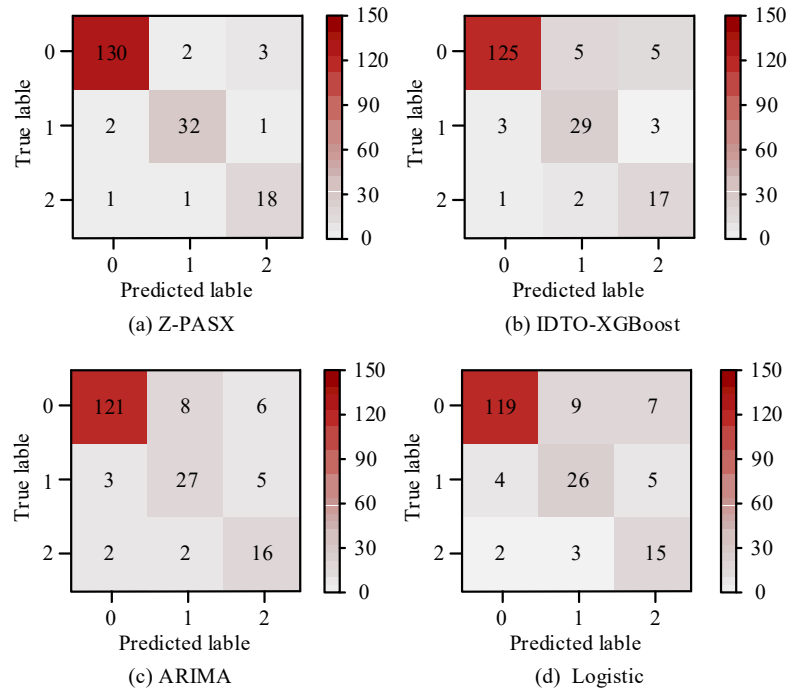
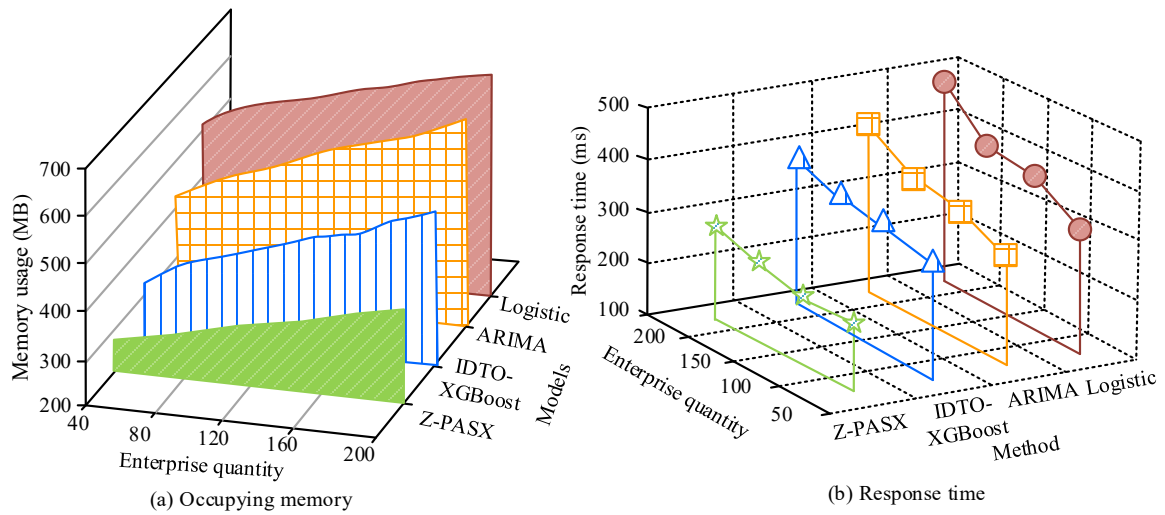
As depicted in Fig. 10(a), the AUC value of the ALSAX model was 0.918, approximating 1 and much higher than 0.825, 0.735, and 0.679 of the other models, demonstrating that the ALSAX model achieved the best performance among all compared models. As shown in Fig. 10(b), the average precision of the ALSAX model was 0.925, whereas that of the other models was below 0.9. In conclusion, the ALSAX model performed best, demonstrating superiority in realistic corporate financial risk prediction tasks, as measured by precision and recall.

### 3.3. Practical Application Performance of the Z-PASX Model

Based on the separate validation of the SCA-AOA-SVM corporate financial risk identification model and the ALSAX prediction model, this study further validated the practical application effectiveness of the constructed Z-PASX model. This study compared it with the Autoregressive Integrated Moving Average Regression Model (ARIMA), the Logistic regression model, and the Improved Tasmanian Devil Algorithm eXtreme Gradient Boosting (IDTO-XGBoost). To verify the authenticity and reliability of the experiment, this study selected financial indicators and annual report texts for 200 listed companies from the Guotai and database from 2015 to 2025, covering the 12 first-level industries designated by the China Securities Regulatory Commission. The data missing rate was 3.2%, and linear interpolation and mode filling were used. According to the "Risk Warning Management Measures for Listed Companies" of the China Securities Regulatory Commission, \* ST/ST companies are defined as high-risk, with ROE  $\geq 6\%$  for three consecutive years and no violations, as low-risk, and the rest as medium risk. First, the 200 companies were classified into three risk categories, such as low, medium, and high, with 135, 35, and 20 companies in each category, respectively. The prediction results of the four models were compared, and the comparison results are shown in Fig. 11.

As presented in Fig. 11(a), for the 135 low-risk firms, the Z-PASX model correctly predicted 130; for the 35 medium-risk firms, the Z-PASX model correctly predicted 32; and for the 20 high-risk firms, the Z-PASX model correctly predicted 18. The Z-PASX model achieved a total predictive accuracy of 90.0%, outperforming the three comparative models with 85.5%, 82.0%, and 80.0%, respectively. The results of this study demonstrate that the Z-PASX model performed better in terms of sensitivity and accuracy for practical corporate financial crisis prediction problems and significantly outperformed the comparative models. Finally, this study compared the response time and memory usage of the four models in corporate risk identification and prediction; the results are shown in Fig. 12.

In Fig. 12(a), as the number of companies to be processed increased, the Z-PASX model had the slowest memory growth. When the number of companies reached 200, the Z-PASX model's memory usage was only 401.2 MB, significantly lower than the 558.4 MB, 649.3 MB, and 692.7 MB of the three comparative models. In Fig. 12(b), the response time of the Z-PASX model was consistently the shortest. When the number of companies reached 200, the Z-PASX model's response time was 252.4ms, while the response times of the three comparative models were all above 300ms. Experimental findings indicated that the Z-PASX model performed with relatively small variations when dealing with large corporate financial datasets, and its data processing capability was more stable and efficient than those of the other algorithms used for comparison. To comprehensively evaluate the decision support value of the model in identifying and predicting dual dimensions, this study introduced three decision-related dimensions of quality, safety, and environment, and quantitatively evaluated the four models. The results are shown in Table 1.

**Fig. 11.** Comparison of confusion matrix for prediction results**Fig. 12.** Comparison of memory usage and response time results**Table 1.** Quantitative evaluation results of four models in dimensions of quality, safety and environment

| Evaluation Dimension | Indicator                                  | Z-PASX | IDTO-XGBoost | ARIMA | Logistic |
|----------------------|--------------------------------------------|--------|--------------|-------|----------|
| Quality              | Identification F1-score                    | 0.92   | 0.89         | 0.88  | 0.86     |
|                      | Prediction F1-score                        | 0.91   | 0.88         | 0.85  | 0.84     |
| Safety               | Identification False Alarm Rate (%)        | 3.0    | 4.6          | 5.6   | 6.2      |
|                      | Prediction Missing Rate (%)                | 2.8    | 4.3          | 4.7   | 5.8      |
| Environment          | Cross-industry Identification Accuracy (%) | 91.2   | 86.5         | 82.3  | 78.6     |
|                      | Cross-industry Prediction Accuracy (%)     | 90.7   | 84.9         | 80.7  | 77.9     |

According to Table 1, in the quality dimension, the Z-PASX model achieves recognition and prediction F1 scores of 0.92 and 0.91, respectively, both higher than those of the three comparison models. In terms of security, the Z-PASX model has

a recognition false alarm rate of 3.0% and a prediction false alarm rate of 2.8%, both lower than those of the other three comparison models. In terms of the environmental dimension, the Z-PASX model achieved cross-industry recognition accuracy of 91.2% and cross-industry prediction accuracy of 90.7%, and performed best among the four models. Overall, the Z-PASX model performs well and provides comprehensive, reliable, and efficient financial risk decision support for enterprises, investors, and regulatory authorities.

#### **4. Further Discussion**

The proposed SCA-AOA-SVM identification model and ALSAX prediction model demonstrated significant advantages in performance comparison. From an identification perspective, the proposed SCA-AOA-SVM model achieved an identification accuracy of 97.8% upon completion of the iterations, with an F1 score of 92.7%, a loss of 0.23, and a misjudgment rate of 2.02%. This was mainly attributed to the model's combination of the parameter optimization capability of the SCA algorithm with the optimization characteristics of the AOA algorithm, further integrated with the nonlinear classification advantages of the SVM algorithm, enabling the model to more precisely capture key features in the data and effectively distinguish companies with different risk states when processing complex corporate financial risk data. This result was consistent with the experimental results of S. Xia et al., who proposed an efficient and robust SVM framework based on granular ball computing, demonstrating the effective performance of the optimized SVM for identification and classification (Xia et al., 2025). In terms of the prediction model, the ALSAX prediction model achieved 98.4% after the end of the iterations with a low omission identification rate of only 1.62%, an AUC of 0.918 (approximating 1), and an average precision of 0.925, which clearly indicates its robustness in the classification prediction of the model. This is because it leveraged the AM-LSTM network's excellent data-processing capability and time-series data, along with the SCA-AOA-XGBoost algorithm, to model potential patterns in multi-period financial feature data for companies. This was supported by the multi-peak load forecasting model proposed by Rubasinghe et al. (2024), which showed that LSTMs-based temporal feature extraction significantly improved its predictive capability.

In terms of practical application effectiveness, the Z-PASX prediction-and-identification fusion model also performed well. Of 135 low-risk companies, 130 were accurately predicted by the Z-PASX model, yielding an accuracy of 96.29%. This was mainly attributed to the Z-PASX model's ability to precisely capture key features in corporate financial risk data while deeply exploring potential patterns in companies' multi-period financial characteristics. Compared to the probabilistic risk assessment and early warning model based on Bayesian deep learning proposed by Liu et al. (2024). This result achieved greater improvement in assessment accuracy and efficiency. Finally, when dealing with a large volume of corporate finance information, the model exhibited low growth in memory consumption and response time, indicating stable and effective memory processing capacity. This result is mainly attributed to the targeted optimization of computational efficiency during model design, ensuring stability and effectiveness when processing large-scale financial data. This result is in line with the approach for handling SVM classification using evolutionary bilevel optimization introduced by Rosales-Pérez et al. (2023). From a management decision-making perspective, the model's practical value is reflected not only in improved accuracy but also in substantial optimization of management processes. The high recognition accuracy and low false alarm rate ensure that financial management personnel can quickly and accurately locate high-risk enterprises during batch screening, effectively reducing the cost of manual secondary screening and significantly improving risk control efficiency. The low forecast omission rate provides a solid guarantee for audit and compliance management, helping avoid major decision-making errors and compliance risks caused by the omission of potentially high-risk enterprises.

#### **5. Conclusion**

This study innovatively proposes the Z-PASX multidimensional identification and prediction model. The model integrates the PCA-GRA algorithm, the AM-LSTM network, the SCA-AOA-SVM model, and the SCA-AOA-XGBoost model to achieve multi-dimensional screening, feature extraction, and accurate risk prediction. Experimental results demonstrate that the SCA-AOA-SVM identification model outperforms the comparative models in accuracy and F1 score, while the ALSAX prediction model excels in predictive accuracy and false-negative rate. In practical applications, the Z-PASX model exhibits higher sensitivity and accuracy in identifying enterprises of different risk levels. When processing large volumes of financial data, it exhibits slow growth in memory consumption, short response times, and stable, efficient processing, effectively overcoming the shortcomings of traditional methods. From a decision-value perspective, the high F1 score, and low misjudgment rate of the recognition model can help the finance department significantly reduce manual screening costs and improve efficiency in quarterly risk assessment. The low false-alarm rate and stable response speed of the predictive model can provide a timely and reliable basis for issuing warnings to audit and risk control departments. However, this study did not classify enterprise types, and the model's generalizability requires further evaluation. Future improvements will focus on multi-industry optimization to enhance adaptability and predictive performance. In future research, financial risk warning thresholds can be set by industry, and monitoring frequencies can be differentiated based on industry characteristics to enhance the model's adaptability and warning sensitivity in practical business. At the same time, to verify the model's generalizability, external validation tests can be conducted across different market environments and time periods in the future to further expand the model's application boundaries.

#### **Funding**

This research received no specific financial support from any funding agency.

#### **Institutional Review Board Statement**

Not applicable.

#### **Declaration of Artificial Intelligence (AI) Tools**

The author used AI tools solely for language editing and to improve readability. The author reviewed and verified all content and takes full responsibility for the accuracy and integrity of the manuscript.

## References

- Xu, C., Li, R., Zhu, L., Shen, X., and Sharif, K. (2024). EWDPs: A Novel Framework for Early Warning and Detection on Ethereum Phishing Scams. *IEEE Internet of Things Journal*, 11(19), 30483-30495, DOI: 10.1109/JIOT.2024.3409742.
- Dai, Y., Noebels, M., Preece, R., Panteli, M., and Dobson, I. (2024). Risk Assessment and Mitigation of Cascading Failures Using Critical Line Sensitivities. *IEEE Transactions on Power Systems*, 39(2), 3937-3948, DOI: 10.1109/TPWRS.2023.3305093.
- Anand, S. C., Teixeira, A. M. H., and Ahlén, A. (2024). Risk Assessment of Stealthy Attacks on Uncertain Control Systems. *IEEE Transactions on Automatic Control*, 69(5), 3214-3221, DOI: 10.1109/TAC.2023.3318194.
- Zhuang, X. and Wu, Y. (2025). Large Language Models in Corporate Investment: An Analysis of ChatGPT's Transformative Applications, Strategic Impacts, and Future Opportunities. *IT Professional*, 27(2), 21-27, DOI: 10.1109/MITP.2025.3543555.
- Tyagi, S. K. S., and Boyang, Q. (2023). An Intelligent Internet-of-Things-Aided Financial Crisis Prediction Model in FinTech. *IEEE Internet of Things Journal*, 10(3), 2183-2193, DOI: 10.1109/JIOT.2021.3088753.
- Chen, Y., Kuang, X., and Guo, J. (2024). LiFoL: An Efficient Framework for Financial Distress Prediction in High-Dimensional Unbalanced Scenario. *IEEE Transactions on Computational Social Systems*, 11(2), 2784-2795, DOI: 10.1109/TCSS.2023.3276059.
- Liu, T. and Yang, L. (2024). Financial Risk Early Warning Model for Listed Companies Using BP Neural Network and Rough Set Theory. *IEEE Access*, 12, 27456-27464, DOI: 10.1109/ACCESS.2024.3367228.
- Li, L., Liu, J., Tseng, M. L., and Lim, M. K. (2024). Accuracy of IGBT Junction Temperature Prediction: An Improved Sailfish Algorithm to Optimize Support Vector Machine. *IEEE Transactions on Power Electronics*, 39(6), 6864-6876, DOI: 10.1109/TPEL.2024.3370690.
- Cao, H., Wu, Y., Bao, Y., Feng, X., Wan, S., and Qian, C. (2023). UTrans-Net: A Model for Short-Term Precipitation Prediction. *AIAA*, 1(2), 106-113, DOI: 10.47852/bonviewAIAA2202337.
- Wei, Z., Lu, H., Cao, N., and Hou, Y. (2023). Non-Invasive Online Abnormal Load Identification Method Based on AOA-SVM. *2023 5th International Conference on Intelligent Control, ICMSP*, Chengdu, China, 965-969, DOI: 10.1109/ICMSP58539.2023.10171070.
- Yan, K., Liu, X., Lu, Y., and Qin, F. (2023). A Cyber-Physical Power System Risk Assessment Model Against Cyberattacks. *IEEE Systems Journal*, 17(2), 2018-2028, DOI: 10.1109/JSYST.2022.3215591.
- Lyu, S., and Jiao, Z. (2025). Optimization of Financial Asset Allocation and Risk Management Strategies Combining Internet of Things and Clustering Algorithms. *IEEE Internet of Things Journal*, 12(4), 3654-3669, DOI: 10.1109/JIOT.2024.348671.
- Eckhart, M., Ekelhart, A., Biffel, S., Lüder, A., and Weippl, E. (2023). QualSec: An Automated Quality-Driven Approach for Security Risk Identification in Cyber-Physical Production Systems. *IEEE Transactions on Industrial Informatics*, 19(4), 5870-5881, DOI: 10.1109/TII.2022.3193119.
- hang, C. and Xu, M. (2024). Time-Segment Photovoltaic Forecasting and Uncertainty Analysis Based on Multi-Objective Slime Mould Algorithm to Improve Support Vector Machine. *IEEE Transactions on Power Systems*, 39(3), 5103-5114, DOI: 10.1109/TPWRS.2023.3333686.
- Nacpil, E. J. C., Wang, Z., Guan, M., Nakano, K., and Jeon, I. (2023). EEG-Based Emergency Braking Prediction Using Data Ablation and SVM Classification. *IEEE Sensors Journal*, 23(14), 16013-16019, DOI: 10.1109/JSEN.2023.3283447.
- Khushaba, R. N., Al-Timemy, A. H., Samuel, O. W., and Scheme, E. J. (2022). Myoelectric Control With Fixed Convolution-Based Time-Domain Feature Extraction: Exploring the Spatio - Temporal Interaction. *IEEE Transactions on Human-Machine Systems*, 52(6), 1247-1257, DOI: 10.1109/THMS.2022.3146053.
- Chen, L., Li, M., Wu, M., Pedrycz, W., and Hirota, K. (2024). Convolutional Features-Based Broad Learning With LSTM for Multidimensional Facial Emotion Recognition in Human - Robot Interaction. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 54(1), 64-75, DOI: 10.1109/TSMC.2023.3301001.
- Qian, Y., Chen, C., Tang, L., Jia, Y., and Cui, G. (2023). Parallel LSTM-CNN Network With Radar Multispectrogram for Human Activity Recognition. *IEEE Sensors Journal*, 23(2), 1308-1317, DOI: 10.1109/JSEN.2022.3224083.
- Chiong, R., Fan, Z., Hu, Z., and Dhakal, S. (2023). A Novel Ensemble Learning Approach for Stock Market Prediction Based on Sentiment Analysis and the Sliding Window Method. *IEEE Transactions on Computational Social Systems*, 10(5), 2613-2623, DOI: 10.1109/TCSS.2022.3182375.
- Jahanshahi, A. A., Cakir, F. S., Kalaycioglu, O., and Adiguzel, Z. (2024). Economic Impacts and Financial Strategies of Turkish Logistics' Companies During the COVID-19 Crisis. *IEEE Engineering Management Review*, 52(2), 69-83, DOI: 10.1109/EMR.2024.3354877.
- Xia, S., Lian, X., Wang, G., Gao, X., Chen, J., and Peng, X. (2025). GBSVM: An Efficient and Robust Support Vector Machine Framework via Granular-Ball Computing. *IEEE Transactions on Neural Networks and Learning Systems*, 36(5), 9253-9267, DOI: 10.1109/TNNLS.2024.3417433.
- Rubasinghe, O., Zhang, X., Chau, T. K., Chow, Y. H., Fernando, T., and Iu, H. H. -C. (2024). A Novel Sequence to Sequence Data Modelling Based CNN-LSTM Algorithm for Three Years Ahead Monthly Peak Load Forecasting. *IEEE Transactions on Power Systems*, 39(1), 1932-1947, DOI: 10.1109/TPWRS.2023.3271325.

- Liu, X., Liu, J., Zhao, Y., Ding, T., Liu, X., and Liu, J. (2024). A Bayesian Deep Learning-Based Probabilistic Risk Assessment and Early-Warning Model for Power Systems Considering Meteorological Conditions. *IEEE Transactions on Industrial Informatics*, 20(2), 1516-1527, DOI: 10.1109/TII.2023.3278873.
- Rosales-Pérez, A., García, S., and Herrera, F. (2023). Handling Imbalanced Classification Problems With Support Vector Machines via Evolutionary Bilevel Optimization. *IEEE Transactions on Cybernetics*, 53(8), 4735-4747, DOI: 10.1109/TCYB.2022.3163974.



Fangyuan Ge was born in June of 1993 and is of Han ethnicity, native to Linyi City, Shandong Province. She graduated from China Agricultural University with a Bachelor of Management in Public Administration in 2015 and obtained a Master of Business Administration degree from Shandong University in 2021. Her research interests include economics, accounting and finance. She has been employed at Linyi Commercial Bank since 2021 to Present, and has served as a full-time teacher at Linyi University of Science and Technology. She has presided over one municipal-level research project, participated in two municipal-level research projects, and obtained one utility model patent.