A Closed-Loop Financial Risk Prediction and Dynamic Control Framework Using Optimized Random Forests

Xuemei Guan, Guixian Tian*
School of Economics and Management, Yancheng Institute of Technology, Yancheng 224051, China
E-mail: guan8803@126.com, 18663510797@163.com

Keywords: enterprise financial risk, random forest, risk early warning, dynamic control, model robustness

Received: October 21, 2025

To address the frequent financial risk (FR) issues exposed by enterprises in the complex economic environment, this paper proposes an optimized model based on random forest (RF), which is used for the early warning and dynamic control of enterprise financial risk (EFR). In existing studies, traditional statistical methods rely excessively on linear assumptions and thus struggle to capture the nonlinear relationships between financial indicators. To tackle these problems, this paper establishes a closed-loop risk management framework of "prediction - classification - control - feedback". Systematic optimization of the RF is further performed, including hyperparameter tuning (e.g., number of trees, maximum depth, minimum samples for splitting, etc.), class weight setting to alleviate the imbalanced sample problem, and indicator selection based on feature importance combined with integration strategy adjustment. This thereby improves the generalization ability and robustness of the model. The optimized model works exceptionally well in terms of prediction accuracy and robustness according to experimental results. For example, on the University of California, Irvine Bankruptcy Prediction (UCI-BP) dataset, its accuracy rate is 0.91, precision rate is 0.88, specificity is 0.90, and Brier score is only 0.11, which is significantly better than the comparison models. Among the robustness indicators, the logarithmic loss is 0.27, Matthews Correlation Coefficient (MCC) is 0.78, Kappa coefficient is 0.75, and stability index is 0.89, all demonstrating stronger robustness. In terms of dynamic control effect, based on the risk probabilities output by the model, this paper classifies enterprises into different risk levels and constructs a strategy mapping rule base of "risk level-control measures". By simulating the implementation of intervention methods such as cash flow optimization and liability structure adjustment, the changes in risk distribution and the transfer rate of high-risk enterprises are evaluated in combination with feedback data. The risk classification accuracy rates of the optimized model on three datasets (UCI-BP, Kaggle Company Bankruptcy Prediction (Kaggle-BP) and Kaggle Corporate Credit Rating (Kaggle-CCR)) are 0.91, 0.89 and 0.88 respectively. Moreover, the transfer rates of high-risk enterprises reach 0.34, 0.32 and 0.30 respectively. According to these findings, the model can successfully assist in the execution of control techniques and precisely identify possible hazards, enabling businesses to take preventative action before crises arise. As a result, this paper offers a fresh approach and useful resource for EFR early warning and control research. It enhances the theoretical framework of risk prediction models and raises the usefulness of risk management in real-world applications.

Povzetek: Članek predstavlja izboljšan model za zgodnje prepoznavanje in obvladovanje finančnih tveganj podjetij, ki omogoča boljše napovedi ter učinkovitejše upravljanje tveganj.

1 Introduction

The problem of corporate financial risk (FR) has steadily grown to be a significant element influencing the survival and long-term growth of businesses against the backdrop of the world economy's growing complexity and degree of uncertainty [1,2]. In recent years, with the frequent occurrence of macroeconomic fluctuations, industrial structure adjustments, and external environmental shocks, problems faced by enterprises, such as increased financing difficulties, heavier debt burdens, and tighter liquidity, have continued to emerge [3]. This directly threatens the enterprises' own debt-servicing capacity and operational stability, and may trigger systemic risks through the industrial chain transmission effect, affecting the overall security of financial market operations and social economic performance. Traditional FR early warning methods are mostly based on linear statistical models. Their simplicity and interpretability are their main advantages, but they clearly fall short when it comes to handling complex interacting relationships and high-dimensional nonlinear features [4]. Machine learning techniques have become increasingly used in the field of FR management as big data and artificial intelligence (AI) have grown. Among these methods, random forest (RF) has attracted attention due to its high accuracy and robustness in classification and prediction tasks. RF may successfully rank the significance of variables and lower the danger of model overfitting by combining different decision trees, offering FR early warning a fresh viewpoint [5,6]. However, "prediction" alone is not sufficient to meet the needs of corporate risk management. Corporate FR management requires early identification of potential risks, and must implement dynamic control and response strategies after risks are identified to reduce the possible losses caused by risks [7,8]. This requires further introducing a "dynamic control mechanism" on the basis of the early warning model, combining risk prediction with risk response, and constructing a closed-loop system covering "identification - early warning - control".

In summary, this paper focuses on the early warning and dynamic control of Enterprise Financial Risk (EFR), with the specific research objectives as follows:

(1) Construct a high-precision EFR prediction model based on RF.

(2) Design and integrate a dynamic control mechanism of "prediction-classification-control-feedback".

(3) Conduct comparison and supplementation of existing research from the technical and application levels, demonstrate the advantages of the proposed model in handling nonlinear relationships, improving interpretability, and enhancing dynamic response capabilities, and provide risk management tools and theoretical references with popularization value for enterprises, financial institutions, and regulatory authorities.

2 Related works

In the research on FR management, existing achievements can be roughly divided into three categories:

(1) FR early warning models based on traditional statistical methods;

(2) Risk prediction methods based on machine learning and AI;

(3) Dynamic control and closed-loop management frameworks oriented to risk response.

Although existing studies have achieved certain progress, they still generally face problems such as overly strong linear assumptions, insufficient model transparency, difficulty in handling data heterogeneity, and lack of dynamic response capabilities, as shown in Table 1.

Therefore, this paper proposes a FR early warning and dynamic control framework based on RF. The nonlinear fitting ability and feature importance mechanism of RF can make up for the problems of overly strong linearity and insufficient interpretability of traditional statistical models. The robustness of RF on medium-sized enterprise data can compensate for the limitations of deep learning models, such as strong dependence on large samples and susceptibility to overfitting. Constructing a dynamic closed-loop of "prediction-classification-control-feedback" based on RF early warning fills the gap in existing research that lacks executable risk response mechanisms.

Table 1: Related work arrangement

Research	Used model	Dataset	Main performance indicators	Limitations
[9]	Z-score and Logit regression	Low dimensional financial data	ACC and F1	Dependent linear hypothesis; Difficult to deal with nonlinearity; Poor performance in dynamic environment
[10]	Multiple discriminant analysis	Enterprise credit data	Classification accuracy	Require the normality and variance homogeneity of variables; The stability of the model is poor; Refractory heterogeneity
[11]	support vector machine	Credit risk data of small and medium-sized enterprises	ACC	Parameter sensitivity; Large-scale data training is inefficient; Limited practicality
[12]	artificial neural network	Enterprise financial data	AUC, ACC	Black box effect is serious; Weak interpretability; It is more difficult to be accepted by the supervision.
[13]	Threshold control method	Liquidity risk scenario	Control success rate	Mostly theoretical research; Lack of real enterprise data verification; The actual operability is weak.

3 Research method

3.1 Analysis of principle and applicability of RF algorithm

RF is a non-parametric classification and regression method based on ensemble learning. Its core principle lies in implementing ensemble prediction using multiple independent decision trees through the "Bagging + random feature selection" approach [14,15]. Compared with traditional statistical models and other machine learning methods, RF has the following advantages in the field of corporate FREW:

(1) Strong nonlinear fitting capability: FRs involve numerous financial and non-financial indicators, and the relationships among these indicators are often nonlinear. RF can flexibly capture the complex interactive relationships between high-dimensional variables without being constrained by linear assumptions [16,17].

(2) Good adaptability to high-dimensional data: Corporate financial data contains a large number of indicators, and these indicators are highly correlated with each other [18]. By means of random feature selection, RF reduces data redundancy, enabling the model to maintain good prediction accuracy even in high-dimensional scenarios.

(3) Strong anti-overfitting ability: RF reduces variance by integrating multiple decision trees, avoiding the overfitting problem of single models. This is particularly effective in scenarios where the volume of financial data samples is limited [19].

(4) Variable importance evaluation: RF can quantify and rank the importance of each variable. This helps identify key FR factors and provides corporate managers with decision-making basis with strong interpretability [20].

(5) High robustness and stability: When data contains noise, missing values, or outliers, RF can still maintain stable prediction performance. It is thus adaptable to the incompleteness that is commonly found in corporate financial data [21].

The application of RF in corporate FREW and dynamic control holds strong practical rationality and academic value. On one hand, corporate financial data includes traditional indicators such as balance sheets, income statements, and cash flow statements, and may involve non-financial indicators related to the market, governance, and industry. RF can integrate multi-source heterogeneous data, thereby improving the comprehensiveness of early warning. On the other hand, by continuously updating samples and features, the RF model can adjust prediction results in real time under a dynamic environment, providing enterprises with phased risk classification early warnings and meeting the needs of dynamic control.

3.2 The construction of EFR early warning model

By analyzing company financial and non-financial data, the corporate FREWM seeks to anticipate possible hazards and provide hierarchical prompts, giving businesses, financial institutions, and regulatory bodies a scientific foundation for decision-making. Taking RF as the core algorithm, this paper combines the financial indicator system and dynamic monitoring mechanism to construct an early warning framework covering "risk feature extraction - model training - result determination - risk classification". This model focuses on prediction accuracy, and emphasizes interpretability and dynamic adaptability, forming a complete closed-loop for risk identification. Its specific framework is shown in Figure 1:

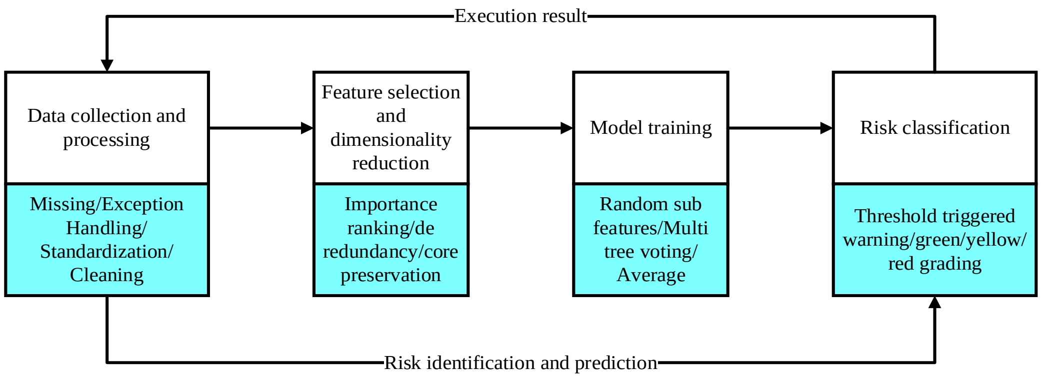


Figure 1: Framework for EFR warning model

The model construction process is as follows:

(1) Data preprocessing: To increase the efficacy of model training, handle outliers, standardize the original data, and imputation of missing values.

(2) Feature selection and dimensionality reduction: Utilize the feature importance evaluation function of RF to screen indicators with higher contribution to FR prediction, eliminate redundant or noisy variables, and optimize the computational efficiency of the model.

(3) Model training: Use RF to train sample data, generate multiple sub-sample sets through Bootstrap sampling, construct multiple decision trees, and perform

random feature selection during node splitting.

(4) Risk identification and prediction: Determine the risk level of newly input enterprise data, output classification results (such as "high risk - medium risk - low risk"), and provide corresponding probability distributions.

(5) Risk classification early warning mechanism: According to the model output results, classify enterprises into different risk levels and trigger corresponding early warning signals. For instance, the system immediately signals "high risk" and suggests appropriate control steps when the estimated likelihood surpasses the predetermined level.

Let the FR label of the ith enterprise be $y_i \in \{0,1\}$, then the risk scoring function of the enterprise (estimated as the probability of high risk) is defined as:

$$s(\mathrm{x}_i) = p(y_i = 1|\mathrm{x}_i) = \frac{1}{T}\sum_{t=1}^{T} h_t(\mathrm{x}_i) \quad (1)$$

T is the total number of decision trees, $h_t(\mathrm{x}_i)$ is the output of t-tree, and $s(\mathrm{x}_i)$ is the score of scoring function. In the model training stage, (x_i, y_i) is taken as the training sample, and the logarithmic loss function with class weighting is taken as the optimization objective to give consideration to the problem of class imbalance:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N} [w_1 y_i \log s(\mathrm{x}_i) + w_0(1-y_i)\log(1 - s(\mathrm{x}_i))] \quad (2)$$

In the equation, N is the number of training samples. w_0 and w_1 are the weights, and $\mathcal{L}$ is the loss function. In the decision-making stage, in order to map the continuous risk score into three types of risk levels, two decision thresholds are introduced:

$$\mathrm{RiskLevel}(\mathrm{x}) = \begin{cases} \text{Low}, & s(\mathrm{x}) < \tau_M \\ \text{Medium}, & \tau_M \le s(\mathrm{x}) < \tau_H \\ \text{High}, & s(\mathrm{x}) \ge \tau_H \end{cases} \quad (3)$$

τ_M and τ_H are obtained by grid search on the verification set, and the optimization objective is to maximize the F1 value or Youden index of high-risk identification to balance the recall rate and false alarm rate.

In summary, the constructed corporate FREWM can identify the potential factors of FRs, and provide hierarchical early warnings for the future risk levels of enterprises, thereby offering a scientific basis for risk prevention and control.

3.3 Design of dynamic control mechanism and model integration

EFRs exhibit multi-dimensional, nonlinear, and time-evolving characteristics. Relying solely on static prediction results is insufficient to support enterprises' continuous risk intervention decisions. To address this, based on the RF-based FR early warning model, this paper introduces a dynamic control mechanism and integrates "risk prediction-risk grading-control strategy-effect feedback-model update" into a closed-loop operational management framework. The overall model consists of three core sub-modules, as detailed in Table 2:

Table 2: Model framework of EFR early warning and dynamic control

Submodel	Main input	Core method and processing logic	Main output
Risk prediction submodel (RF)	Enterprise financial indicators, non-financial information, historical labels, etc.	Based on the integrated learning and probability prediction of RF, the enterprise risk score and risk grade judgment are generated.	Enterprise risk probability and risk level (such as low risk, medium risk and high risk)
Control strategy submodel (dynamic control)	Risk level, enterprise scale, industry attributes, debt structure, etc.	Call the manually designed "risk level-control measures" rule base to match the differentiated intervention combination according to the level.	Collection of control schemes for a single enterprise (credit policy adjustment, monitoring frequency adjustment, etc.)
Feedback and iteration mechanism	Effect data after control implementation (cash flow, default result, rating change, etc.)	Evaluate the control effect, track the risk change trend, and trigger the update when the model performance declines or the environment changes obviously.	Configuration of updated training sample set, model parameters and rule base

In practical implementation, the system classifies enterprises into three levels (low risk, medium risk, and high risk) based on the risk probabilities output by the RF model. The control objectives and typical measures corresponding to different risk levels are shown in Table 3:

Table 3: Model framework of EFR early warning and dynamic control

Risk level/component	Judgment basis (qualitative description)	Main control target	Typical control measures and system behavior
Low risk	The risk probability given by the model is in a low range, and the overall situation of enterprise operation and debt repayment is stable.	Maintain steady operation and avoid excessive intervention and high monitoring cost.	Maintain the normal credit policy; Carry out financial monitoring at regular frequency; Pay attention to industry changes and appropriately optimize the structure of capital use.
Medium risk	The risk probability is at a medium level, and some indicators show a worsening trend or increased fluctuation.	Prevent further risk accumulation and control risk exposure.	Tighten the credit line or raise the entry threshold; Adjust the repayment plan; Improve the frequency of submission and monitoring; Increase management interviews, etc.
High risk	The risk probability is in a high range, the debt repayment pressure is obvious, the liquidity is tight or there is a default signal.	Quickly reduce the probability of default or bankruptcy and protect the safety of funds.	Reduce or freeze the credit line; Asking for additional guarantee or mortgage; Initiate a debt restructuring or bail-out plan; Evaluate the asset disposal or exit plan.
Feedback data collection	Dynamic monitoring after implementation of control measures	Evaluate the effect of the strategy, and provide the basis for the subsequent iteration.	Continuously record changes in cash flow, default, credit rating, etc., and form new training and evaluation samples.
Rules and model updating	According to the feedback results and environmental changes, stage evaluation is carried out.	Maintain the adaptability of early warning and control mechanism to dynamic environment.	When the model performance declines or the risk structure changes obviously, update the training data, retrain the model and adjust the content of the rule base.

It shows that this paper adopts a rule design method based on expert experience and industry norms. First, the risk prediction model provides quantifiable risk levels, and then the rule base offers operable control measures.

4 Experimental design

The datasets selected for the experiment are University of California, Irvine Bankruptcy Prediction (UCI-BP), Kaggle Company Bankruptcy Prediction (Kaggle-BP), and Kaggle Corporate Credit Rating (Kaggle-CCR). UCI-BP is provided by the University of California, Irvine (UCI) Machine Learning Repository, which contains 6 consecutive years of financial data of Taiwanese enterprises. Its goal is to predict whether an enterprise will go bankrupt in the future, making it highly suitable for constructing corporate FREWMs and testing the performance of RF under high-dimensional data. The dataset can be downloaded via https://archive.ics.uci.edu/dataset/572/taiwanese+bankruptcy+prediction. Kaggle-BP is sourced from Kaggle and is based on real corporate financial statements. Its objective is to predict whether an enterprise will go bankrupt, and it is often used in comparative experiments of machine learning models. It can be directly applied to train the RF and evaluate prediction accuracy. The dataset is available for download at https://www.kaggle.com/datasets/fedesoriano/company-bankruptcy-prediction. Kaggle-CCR includes corporate financial indicators and credit ratings, with ratings ranging from AAA to D, which reflect the credit risk level of enterprises. The dataset can be downloaded through https://www.kaggle.com/datasets/agewerc/corporate-credit-rating.

The experiments of this paper are conducted on a well-configured workstation. In terms of hardware: The central processing unit (CPU) is an Intel Core i9-12900K, with a base clock speed of 3.2 GHz and 16 cores/24 threads, which ensures the efficient execution of complex computing tasks. The graphics processing unit (GPU) is an NVIDIA GeForce RTX 3090 with 24 GB of video memory, used to accelerate the model training and prediction processes. The memory is 64 GB DDR5, which guarantees stability in large-scale data processing. The storage device is a 2 TB NVMe solid-state drive (SSD), ensuring high-speed data reading and writing. In terms of software: The operating system is Ubuntu 22.04 LTS. The experiment platform is built on the Python 3.9 environment. The main dependent deep learning and machine learning libraries include scikit-learn 1.2.2, TensorFlow 2.11, and PyTorch 1.13. Meanwhile, NumPy, Pandas, and Matplotlib are used for data processing and visualization. This configuration can provide stable and efficient experimental environment support for the RF model and comparison models in this paper.

The comparison models selected in this paper are Graph Convolutional Network for Financial Distress Prediction (GCN-FD), Bidirectional Long Short-Term Memory with Attention Mechanism (BiLSTM-ATT), and Extreme Gradient Boosting with Dynamic Thresholding (XGBoost-DT). The parameter setting is shown in Appendix file.

5 Analysis of experimental results

5.1 Analysis of experimental results of prediction accuracy and robustness

Two aspects of the experiment are examined: robustness and prediction accuracy. Figure 2 displays the prediction accuracy comparing results.

The proposed optimized RF model outperforms the comparative models overall, according to the data in Figure 2a. In terms of the Accuracy metric: The results of the optimized model on the UCI-BP, Kaggle-BP, and Kaggle-CCR datasets are 0.91, 0.89, and 0.88 respectively, all higher than GCN-FD's 0.87. In Figure 2b. The values of the optimized model on the three datasets are 0.88, 0.86, and 0.85 in sequence, which are significantly better than BiLSTM-ATT's 0.80 on the UCI-BP dataset. In Figure 2c. The results of the optimized model on the UCI-BP, Kaggle-BP, and Kaggle-CCR datasets reach 0.90, 0.87, and 0.86 respectively, higher than XGBoost-DT's 0.82 on the Kaggle-CCR dataset. In Figure 2d. The optimized model achieves the lowest scores, which are 0.11, 0.14, and 0.15 respectively, indicating that its probability prediction accuracy is better than GCN-FD's 0.17. The comparison results of robustness are shown in Figure 3.

In the results of Figure 3a, the optimized model also maintains its advantages. In terms of the Log Loss metric: The results of the optimized model on the three datasets are 0.27, 0.30, and 0.32, which are significantly lower than BiLSTM-ATT's 0.40 on the Kaggle-BP dataset. Figure 3b. The values of the optimized model are 0.78, 0.74, and 0.72 respectively, higher than XGBoost-DT's 0.68 on the Kaggle-BP dataset. Figure 3c. The optimized model achieves 0.75, 0.71, and 0.69 on the three datasets, while GCN-FD only reaches 0.56 on the Kaggle-CCR dataset. In Figure 3d. The optimized model obtains 0.89, 0.86, and 0.85 respectively, all higher than BiLSTM-ATT's 0.80 on the UCI-BP dataset. Overall, the optimized model performs prominently across the eight metrics and three datasets, demonstrating stronger prediction accuracy and robustness. To further verify the reliability of the results, this paper conducts 10 repeated experiments, and calculates the mean ± standard deviation for all prediction indicators and robustness indicators. Meanwhile, to further confirm the stability of the results, 95% confidence intervals are also provided in this paper. All experiments are completed under the same data division ratio and identical training configurations, with the results shown in Table 4.

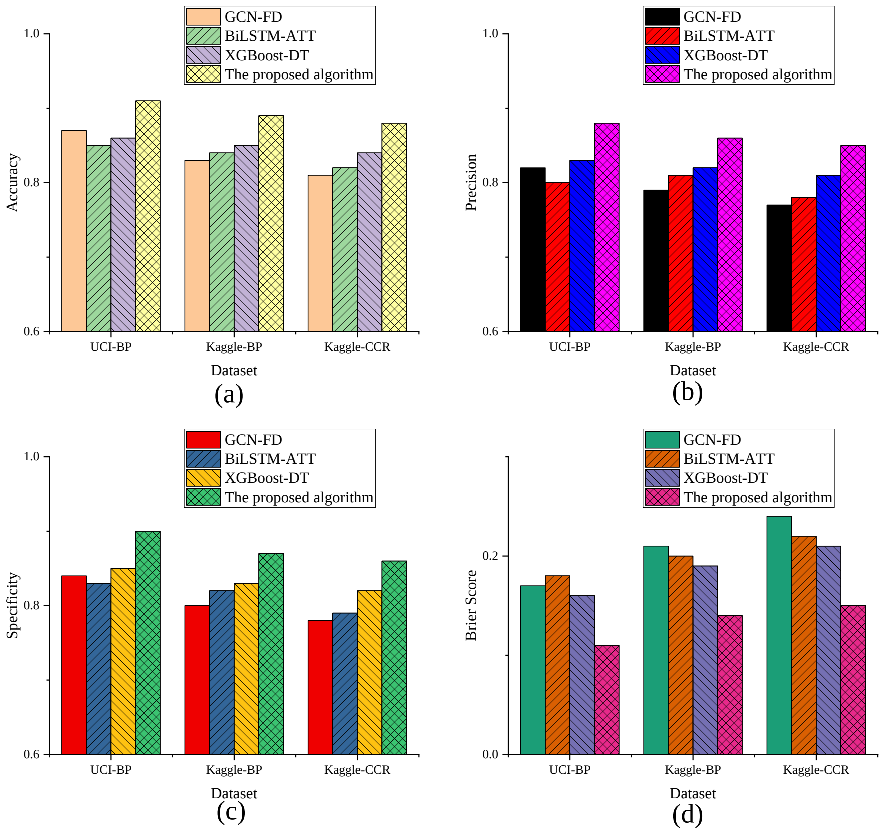


Figure 2: Comparison of experimental results on prediction accuracy(a) Accuracy (b) Precision (c) Specificity (d) Brier Score

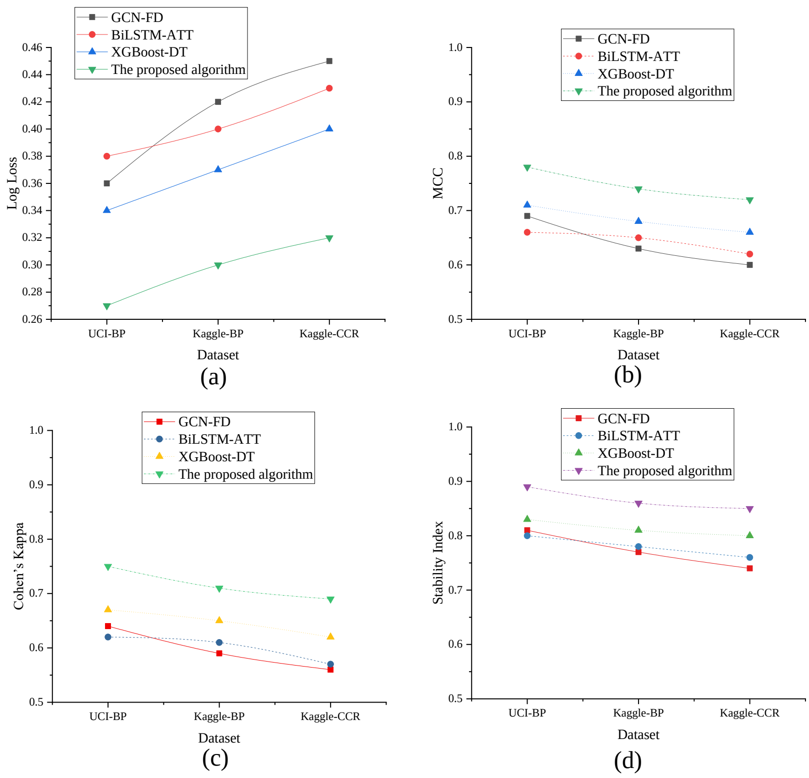


Figure 3: Comparison results of robustness (a) Log Loss (b) Matthews Correlation Coefficient (MCC) (c) Cohen's Kappa (d) Stability Index

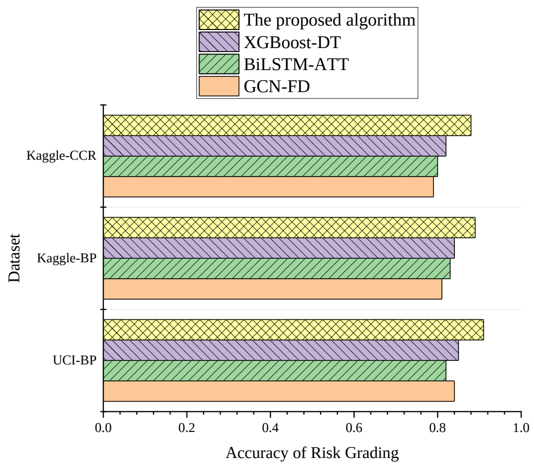
Table 4: Model framework of EFR early warning and dynamic control

Indicator	Dataset	Optimized RF (mean ±SD)	95%CI	Optimal contrast model (mean±SD)
Accuracy	UCI-BP	0.910±0.006	[0.898,0.922]	0.873±0.011 (GCN-FD)
	Kaggle-BP	0.892±0.007	[0.878,0.905]	0.862±0.013 (XGBoost-DT)
	Kaggle-CCR	0.881±0.008	[0.866,0.896]	0.852±0.012 (GCN-FD)
Precision	UCI-BP	0.880±0.007	[0.865,0.894]	0.801±0.015 (BiLSTM-ATT)
	Kaggle-BP	0.861±0.009	[0.842,0.879]	0.824±0.014 (XGBoost-DT)
	Kaggle-CCR	0.847±0.010	[0.827,0.867]	0.818±0.016 (BiLSTM-ATT)
Specificity	UCI-BP	0.899±0.005	[0.889,0.910]	0.861±0.011 (GCN-FD)
	Kaggle-BP	0.874±0.006	[0.862,0.886]	0.832±0.012 (XGBoost-DT)
	Kaggle-CCR	0.862±0.007	[0.849,0.876]	0.822±0.013 (GCN-FD)
Brier Score	UCI-BP	0.112±0.004	[0.105,0.119]	0.171±0.007 (GCN-FD)
	Kaggle-BP	0.141±0.005	[0.131,0.151]	0.182±0.009 (XGBoost-DT)
	Kaggle-CCR	0.148±0.006	[0.136,0.159]	0.191±0.010 (BiLSTM-ATT)
Log Loss	UCI-BP	0.271±0.008	[0.255,0.286]	0.332±0.010 (BiLSTM-ATT)
	Kaggle-BP	0.301±0.009	[0.284,0.318]	0.402±0.012 (BiLSTM-ATT)
	Kaggle-CCR	0.318±0.010	[0.298,0.337]	0.389±0.011 (GCN-FD)
MCC	UCI-BP	0.782±0.006	[0.770,0.794]	0.701±0.013 (XGBoost-DT)
	Kaggle-BP	0.739±0.007	[0.725,0.753]	0.681±0.012 (XGBoost-DT)
	Kaggle-CCR	0.715±0.008	[0.699,0.732]	0.564±0.014 (GCN-FD)
Cohen’s Kappa	UCI-BP	0.748±0.005	[0.739,0.758]	0.692±0.012 (GCN-FD)
	Kaggle-BP	0.706±0.006	[0.693,0.719]	0.658±0.011 (XGBoost-DT)
	Kaggle-CCR	0.684±0.007	[0.671,0.698]	0.561±0.013 (BiLSTM-ATT)
Stability Index	UCI-BP	0.893±0.004	[0.886,0.901]	0.801±0.008 (BiLSTM-ATT)
	Kaggle-BP	0.861±0.005	[0.852,0.870]	0.798±0.010 (XGBoost-DT)
	Kaggle-CCR	0.851±0.006	[0.839,0.863]	0.780±0.009 (GCN-FD)

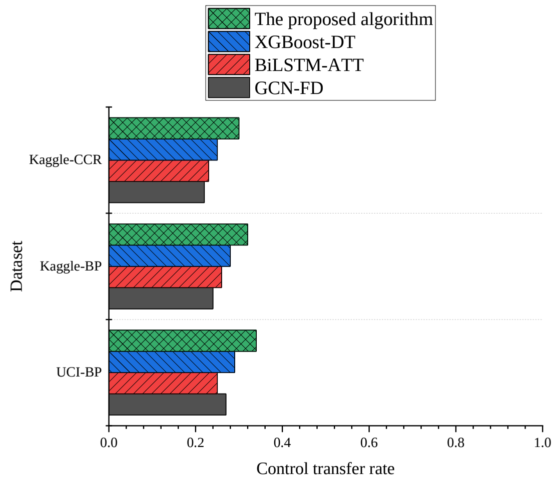
The results in Table 4 suggest that the optimized RF model exhibits low standard deviations (most < 0.01) across all eight indicators on the three types of datasets, indicating strong prediction stability of the model without significant random fluctuations. In addition, the 95% confidence intervals are all relatively narrow, which further confirms the robustness and reproducibility of the model under different training randomness conditions.

5.2 Experiment on dynamic control effect and risk classification

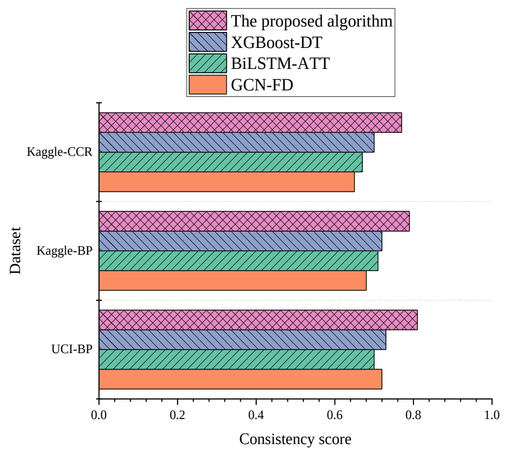
This paper confirms the efficiency of the improved model in dynamic control and risk classification after finishing the FR prediction. The precise procedure is as follows: Businesses are categorized into three risk groups based on the model's risk probability output: low risk (probability ≤ 0.3), medium risk (0.3-0.6), and high risk (probability ≥ 0.6). Under different levels, differentiated control strategies are triggered: Low-risk enterprises conduct regular monitoring. Medium-risk enterprises need to optimize cash flow and debt structure. High-risk enterprises simulate the implementation of measures such as capital injection and debt restructuring. To measure the effect of dynamic control and risk classification, this paper sets three evaluation dimensions, namely risk classification accuracy, the transfer rate of high-risk enterprises under control measures, and the consistency score between predicted distribution and actual distribution. The experimental results are shown in Figure 4:



(a)



(b)



(c)

Figure 4: Experimental results of dynamic control effect and risk grading (a) Accuracy of risk grading (b) Dynamic control transfer rate of high-risk enterprises (c) Reasonable consistency score of risk grading distribution

The results in Figure 4a show the following: In terms of risk classification accuracy, the proposed optimized model achieves results of 0.91, 0.89, and 0.88 on the UCI-BP, Kaggle-BP, and Kaggle-CCR datasets respectively, all higher than those of the comparison models. Specifically, GCN-FD scores 0.84 on UCI-BP, BiLSTM-ATT scores 0.83 on Kaggle-BP, and XGBoost-DT scores 0.82 on Kaggle-CCR, all of which are lower than the optimized model. In Figure 4b, the optimized model obtains values of 0.34, 0.32, and 0.30 on the three datasets, significantly higher than BiLSTM-ATT's 0.26 on Kaggle-BP and GCN-FD's 0.27 on UCI-BP. This indicates that the optimized model performs better in the risk mitigation process. In Figure 4c, the consistency scores of the optimized model on the UCI-BP, Kaggle-BP, and Kaggle-CCR datasets reach 0.81, 0.79, and 0.77 respectively, all outperforming the comparison models. Among the latter, GCN-FD only scores 0.65 on Kaggle-CCR, BiLSTM-ATT scores 0.70 on UCI-BP, and XGBoost-DT scores 0.72 on Kaggle-BP. In terms of risk categorization accuracy, dynamic control efficacy, and risk distribution consistency, the optimized model outperforms the others overall.

6 Discussion

Based on the comprehensive experimental results, the proposed optimized RF model overall outperforms comparative models such as GCN-FD, BiLSTM-ATT, and XGBoost-DT in tasks of EFR early warning and dynamic control, demonstrating favorable empirical effects. However, to truly enhance its applicability in complex financial environments, more in-depth discussions are required from multiple dimensions, including robustness mechanisms, application scenarios, sources of model advantages, and trade-offs.

From the perspective of model robustness, the RF itself, through Bagging and random feature subsampling mechanisms, possesses the effect of resisting uncertainty similar to that in "adaptive robust control" to a certain extent. Most decision trees learn independently on different subsamples and feature subspaces, enabling the model to buffer the impact of single abnormal samples or local structural changes on overall predictions. In the experiments of this paper, the significant decrease in Brier score and logarithmic loss indicates that the model is more stable in probability calibration, which is conducive to providing reliable inputs for risk grading and dynamic intervention. However, it should be emphasized that this study mainly conducts evaluations based on historical samples and normal distribution scenarios, and has not systematically covered extreme or unprecedented financial environments. For "tail risk" scenarios such as severe macroeconomic fluctuations and precipitous industry declines, the model may still face performance degradation. Future research will introduce the ideas of scenario simulation and stress testing. For example, it will construct extreme scenarios that include sudden drops in returns, sharp contractions in liquidity, and violent fluctuations in interest rates or exchange rates. This approach will allow for a systematic evaluation of the model's prediction results, threshold stability, and risk distribution migration. The ultimate goal is to more comprehensively verify the model's stability and reliability under actual market shocks.

The proposed "prediction-classification-control-feedback" closed-loop framework has strong application potential and is expected to be implemented in various financial and practical scenarios. For example, in portfolio management, the model can continuously score the default risk and credit risk of target enterprises, mark high-risk enterprises as "reduced-holding objects" in the portfolio, and track the impact of intervention measures (such as position reduction and hedging tool allocation) on the portfolio loss rate and Sharpe ratio through the feedback loop. In enterprise credit monitoring and supply chain finance scenarios, the framework can be embedded into the risk control systems of banks or factoring institutions to implement hierarchical management of accredited enterprises, triggering credit limit adjustments, additional credit requirements, or early warnings for high-risk enterprises. In terms of ideological path, this framework has certain similarities with the adaptive robust control method in handling uncertainties of nonlinear systems: both emphasize dynamic feedback and strategy adjustment in uncertain environments. However, unlike traditional adaptive robust control which focuses on "stability proof of control laws", this paper focuses more on the comprehensive balance between "prediction accuracy and intervention effectiveness". Through the one-to-one mapping of risk levels and control strategies as well as the effect feedback retraining mechanism, the feedback loop is ultimately transformed into executable decision sequences for enterprises, such as cash flow optimization plans, capital structure adjustments, and asset disposal rhythm arrangements. In the future, further design and deployment plans can be developed at the information system level, such as embedding the model into enterprise ERP/financial sharing platforms or bank credit approval systems to realize automated early warning push and strategy recommendation, thereby enhancing the operability and implementation value of the framework.

In addition, it is necessary to more clearly discuss the potential reasons for the superior performance of the RF model. From the perspective of data characteristics, the UCI-BP, Kaggle-BP, and Kaggle-CCR datasets used in this paper exhibit features such as high dimensionality, significant nonlinear relationships between indicators, and obvious class imbalance. In this context, the RF, by integrating multiple nonlinear decision trees, can effectively characterize complex feature interaction structures and has stronger expressive power in terms of decision boundaries compared to linear logistic regression. Meanwhile, this paper adopts strategies such as class weight balancing and feature importance selection, which suppress the problems of minority class neglect and overfitting caused by redundant features in imbalanced samples and high-dimensional spaces. When compared with deep learning models such as Graph Convolutional Network (GCN) or Bidirectional Long Short-Term Memory (BiLSTM), on the one hand, the RF has relatively lower requirements for sample size and can fully exert its performance on medium-sized datasets. On the other hand, its model structure is relatively simple, and the training and inference processes are more likely to converge, these are also important reasons for its better results under the existing experimental settings. However, the model's superiority does not mean there are no trade-offs and costs. Compared with simpler logistic regression models, the interpretability of the RF is at a "moderate" level. Although it can provide partial explanations for managers through methods such as feature importance and decision path examples, its overall decision-making process still belongs to a "semi-black box", lacking the direct coefficient meaning and marginal effect analysis capability of linear models. For application scenarios that emphasize transparency and auditability (e.g., prudent supervision models in a strong regulatory environment), a balance needs to be struck between prediction performance and interpretability. Meanwhile, compared with deep learning models based on GCN or BiLSTM, the RF usually has advantages in computational cost: the training process can be completed mostly with CPU computing, with low dependence on GPU, and the model scale and inference time are more controllable, making it more suitable for deployment in scenarios with limited resources or the need for rapid iteration. However, deep learning models have inherent advantages in capturing temporal dependencies, network structure relationships, and multimodal information, which is also an important motivation for exploring the "RF + deep learning model fusion" in future research.

Finally, the discussion on robustness improvement needs to be further expanded in dimensions. Based on indicators such as Brier score and logarithmic loss, this paper quantitatively evaluates the stability of the model at the probability prediction level, which mainly reflects the "internal robustness" of the model under a given data distribution. However, in real financial environments, factors such as data drift, structural breakpoints, and institutional changes are likely to cause dynamic evolution of risk feature distributions. To this end, future research will strengthen robustness analysis from three directions: First, introduce concept drift and data drift detection mechanisms to monitor long-term changes in model input distribution and output quality, and promptly trigger model retraining or parameter updates. Second, systematically design model update frequency and rolling training strategies to reduce the risk of model "obsolescence" while balancing stability and adaptability. Third, real-time feedback information, such as cash flow improvements post-intervention or actual default events, is fed into incremental or online learning modules. This allows the "prediction-classification-control-feedback" loop to evolve beyond a static strategy. It achieves genuine dynamic iteration by continuously updating parameters and adjusting its structure. Consequently, the model's long-term robustness and practical utility in realistic, dynamic business environments are significantly enhanced.

In summary, the proposed optimized RF model and its closed-loop framework have demonstrated favorable performance and application prospects in experimental results, but they only represent a starting point for the intelligent management of EFRs. Future work focusing on extreme scenario stress testing, multi-scenario application deployment, deep model fusion, and dynamic robustness enhancement will help further consolidate its theoretical foundation and expand its engineering implementation scope.

7 Conclusion

This paper focuses on corporate FREW and dynamic control, proposing an optimized model based on RF, and conducts empirical research on three public datasets. First, the paper constructs a multi-dimensional indicator system covering debt-servicing capacity, profitability, operational efficiency, cash flow, and non-financial factors. It then uses the RF algorithm to achieve high-precision prediction of corporate FR. Subsequently, combined with a dynamic control mechanism, a closed-loop framework of "prediction - classification - control - feedback" is further established. In terms of prediction accuracy and robustness, the experimental findings demonstrate that the optimized model performs better than the comparison models. It achieves the highest performance in metrics such as Accuracy, Precision, Specificity, Brier Score, Log Loss, MCC, Kappa Coefficient, and Stability Index. Meanwhile, in the experiments on dynamic control effect and risk classification, the optimized model can effectively identify high-risk enterprises, and significantly improve the efficiency of risk mitigation, while maintain a high consistency between the risk distribution and the actual situation. These results fully demonstrate that the proposed optimized model can provide scientific and reliable risk early warning and hierarchical management tools for enterprises, financial institutions, and government regulatory authorities, and has strong theoretical significance and practical value.

Nevertheless, this paper still has several limitations that require cautious consideration. First, the sample sizes of some experimental datasets are limited, and the RF may still have potential overfitting risks under small-sample conditions, so its generalization ability needs to be further verified on larger-scale enterprise financial data. Second, the RF is inherently a static model that lacks the ability to model temporal dependencies, making it difficult to characterize the dynamic evolution process of enterprise financial indicators. In the future, temporal models such as recurrent neural network, LSTM network, or Transformers should be considered for integration. Third, this paper mainly relies on historical financial and credit data for prediction, and does not fully consider exogenous events such as macroeconomic fluctuations, policy changes, and industry shocks, resulting in limited predictive capability for sudden scenarios. In addition, the dynamic control mechanism is constructed based on simulated scenarios, lacking empirical verification from real enterprise implementation cases, and the actual application effect still needs further confirmation. Future research will expand data sources, introduce multimodal information (such as textual public opinion and industry dynamics), and conduct dynamic control experiments in combination with real enterprise scenarios to enhance the robustness and practical applicability of the model.

Data availability statements: All data generated or analysed during this study are included in this published article and its supplementary information files.

References

[1] Wei H, Wang X. Financial risk management early-warning model for Chinese enterprises. Journal of Risk and Financial Management, 2024, 17(7): 255. https://doi.org/10.3390/jrfm17070255

[2] Huang T. Construction and application of financial risk warning system for third-party payment enterprises based on z-score model and efficacy coefficient method. Academic Journal of Business & Management, 2024, 6(12): 121-126. https://doi.org/10.25236/AJBM.2024.061216

[3] Zheng J, Wu X, Tan L, et al. Intelligent financial risk warning for enterprises through knowledge graph-based deep learning. Journal of Circuits, Systems and Computers, 2024, 33(15): 2450262. https://doi.org/10.1142/S0218126624502621

[4] Zhang W, Peng L. Semantic analysis and image processing-jointly driven multimodal deep learning framework for smart warning of enterprise financial risk. Journal of Circuits, Systems and Computers, 2024, 33(17): 2450304. https://doi.org/10.1142/S0218126624503043

[5] Song X, Jing Y, Qin X. BP neural network-based early warning model for financial risk of internet financial companies. Cogent Economics & Finance, 2023, 11(1): 2210362. https://doi.org/10.1080/23322039.2023.2210362

[6] Che F, Jiang Y, Omar K, et al. Enterprise financial risk early warning system based on structural equation model. Applied Mathematics and Nonlinear Sciences, 2023, 8(1): 671-680. https://doi.org/10.2478/amns.2022.2.0053

[7] Min J, Li Z Q, Liu Y, et al. Automotive

manufacturing enterprise financial risk evolution monitoring and early warning simulation: based on the perspective of value chain analysis. Management System Engineering, 2023, 2(1): 10. https://doi.org/10.1007/s44176-023-00021-8

[8] Zhu M. Financial risk control based on optimized z-score financial warning model. Journal of Information Security Research, 2023, 14(4): 101-110. https://doi.org/10.6025/jisr/2023/14/4/101-110

[9] Sánchez-Almeyda C, González-Bueno J, Koval V, et al. Predicting financial distress in the food production sector: a dual-model approach using Z-score and O-score methods. Discover Sustainability, 2025, 6(1): 731. https://doi.org/10.1007/s43621-025-01348-w

[10] Dumitrescu D, Bobitan N, Popa A F, et al. Signaling financial distress through z-scores and corporate governance compliance interplay: a random forest approach. Electronics, 2025, 14(11): 2151. https://doi.org/10.3390/electronics14112151

[11] Hou L, Lu K, Bi G. Predicting the credit risk of small and medium-sized enterprises in supply chain finance using machine learning algorithms. Managerial and Decision Economics, 2024, 45(4): 2393-2414. https://doi.org/10.1002/mde.4130

[12] Gajdosikova D, Michulek J. Artificial Intelligence Models for bankruptcy prediction in agriculture: Comparing the performance of artificial neural networks and decision trees. Agriculture, 2025, 15(10): 1077. https://doi.org/10.3390/agriculture15101077

[13] Dahmen P. Organizational resilience as a key property of enterprise risk management in response to novel and severe crisis events. Risk Management and Insurance Review, 2023, 26(2): 203-245. https://doi.org/10.1111/rmir.12245

[14] Sun J, Zhao M, Lei C. Class-imbalanced dynamic financial distress prediction based on random forest from the perspective of concept drift. Risk Management, 2024, 26(4): 19. https://doi.org/10.1057/s41283-024-00150-8

[15] Wang W, Liang Z. Financial distress early warning for Chinese enterprises from a systemic risk perspective: based on the adaptive weighted XGBoost-bagging model. Systems, 2024, 12(2): 65. https://doi.org/10.3390/systems12020065

[16] Máté D, Raza H, Ahmad I. Comparative analysis of machine learning models for bankruptcy prediction in the context of Pakistani companies. Risks, 2023, 11(10): 176. https://doi.org/10.3390/risks11100176

[17] Samara K, Shinde A. Bankruptcy prediction using machine learning and data preprocessing techniques. Analytics, 2025, 4(3): 22. https://doi.org/10.3390/analytics4030022

[18] Lokanan M E, Ramzan S. Predicting financial distress in TSX-listed firms using machine learning algorithms. Frontiers in Artificial Intelligence, 2024, 7: 1466321. https://doi.org/10.3389/frai.2024.1466321

[19] El Madou K, Marso S, El Kharrim M, et al. Evolutions in machine learning technology for financial distress prediction: A comprehensive review and comparative analysis. Expert Systems, 2024, 41(2): 13485. https://doi.org/10.1111/exsy.13485

[20] Xie X, Jin X, Wei G, et al. Monitoring and early warning of SMEs' shutdown risk under the impact of global pandemic shock. Systems, 2023, 11(5): 260. https://doi.org/10.3390/systems11050260

[21] Liu Y. Research on the inner logic and practical path of empowering enterprise financial intelligence construction with digital economy. International Journal of Social Sciences and Economic Management, 2025, 6 (1), 179. https://doi.org/10.38007/IJSSEM.2025.060115