Predicting Bond Defaults in China: A Comparative and Interpretable Machine Learning Approach Based on the IFind Database

Yangyi Guo
School of Mathematics, College of Science and Engineering, University of Edinburgh, Mathematics and Business BSc (Hons), Edinburgh, Scotland, EH8 9YL, United Kingdom
s2657024@ed.ac.uk

ABSTRACT

Since 2014, the Chinese bond market has seen a large increase in default, whose shortcomings the conventional credit risk models reveal. The paper presents a systematic machine learning model to forecast bond defaults with a new dataset in the IFind database (2014-2024). We compare eight machine learning models, such as Logistic Regression, Random Forest, XGBoost, and Deep Neural Networks. The findings indicate that tree-based ensemble models, especially XGBoost (F1-score: 0.58, AUC: 0.92), are much better than the traditional linear models. Most importantly, we use SHAP (SHapley Additive exPlanations) to give model interpretability, which indicates that profitability (ROE), leverage (Debt-to-Asset Ratio), and liquidity (Quick Ratio) are the most significant predictors, whereas the predictive power of official credit ratings is surprisingly low. A suggested hybrid KMV-XGBoost model is similarly accurate but has theoretically based interpretability. This paper presents a testable, readable early warning mechanism of the Chinese bond market with direct risk management and regulatory policy implications.

Keywords: Bond Default Prediction, Machine Learning, Xgboost, SHAP, KMV Model, Ifind Database, Chinese Bond Market

1. INTRODUCTION

The Chinese bond market is now ranked the second-largest in the world, with a total outstanding amount of over 140 trillion RMB as of 2024 [9]. Nevertheless, this expansion has been coupled with a highly amplified default occurrence since the initial default in 2014 that destroyed the long tradition of informal government guarantees [17]. There are some key examples of the defaults. Notably, they include the default of Yongcheng Coal (AAA-rated) in 2020 and the series of defaults in the real estate industry that started in 2021. Undoubtedly, they revealed critical issues with regard to the current credit assessment ecosystem.

Classical credit risk models, including the Altman Z-Score [2] and logistic regression, make strong assumptions of linearity and cannot model complex, non-linear interactions among financial measures. Considering the previous example, when leverage and low liquidity are together, their effect is not additive but multiplicative, which cannot be represented by linear models [4]. Moreover, credit rating agencies in China have been largely accused of rating inflation, whereby high ratings often give no indication of impending default [11]. The issuer concentration towards high ratings, as noted by Ren [17], has surged, and the high ratings do not usually indicate imminent defaults.

In this paper, three research gaps are discussed. To begin with, although machine learning (ML) has been widely used to predict bankruptcy in Western markets [4, 14], there is limited research related to comprehensive research by applying localized Chinese data. Second, the black box characteristic of ML models will prevent their real-world usage in the regulated financial setting. Third, the inclusion of theory-based measures, e.g., the KMV distance to default, in ML models has not been studied systematically.

This work aims at achieving three goals: (1) to systematically evaluate and compare eight ML models that predict bond default in China with the IFind database; (2) to find the best-performing model with rigorous cross-validation; (3) to propose and test a hybrid KMV-XGBoost model that is more interpretable and yet highly accurate.

2. LITERATURE REVIEW

2.1 Traditional Approaches to Credit Risk Assessment

The modern credit risk assessment started with Altman [2] and the Z-Score, which was a discriminant analysis model that utilizes five financial ratios to forecast corporate bankruptcy. This model and its follow-ups, like the ZETA model [3], became the norm for decades. It was later followed by the use of logistic regression, whose aim is to estimate the likelihood of default based on a linear combination of predictors [15].

Although these models are both interpretable and well-grounded in theory, they are known to have serious shortcomings. They make the assumption that there are linear correlations between predictors and the log-odds of default- an assumption that is often not true in real-life data [4]. The marginal effect of a one per cent change in the debt to asset ratio is not linear but exponentially greater at the extremes. Moreover, these models do not account for the complex effects of interactions. The default risk of a highly leveraged, illiquid company is not equal to the sum of the individual risks, but a non-linear, multiplicative interaction [14].

2.2 Machine Learning in Default Prediction

The machine learning algorithms provide the capability of learning non-linear trends in high-dimensional data without the need to have specific functional forms. First evidence of this was given by Shin et al. [18], who compared Support Vector Machines (SVM) with back-propagation neural networks (BPN) in terms of bankruptcy prediction. They discovered that SVM had a better generalization ability, particularly with a smaller training set, because it had the power to capture geometric properties of the feature space.

Barboza et al. [4] presented pioneering research on the firms of North America and compared discriminant analysis, logistic regression, neural networks, and various ML algorithms (SVM, bagging, boosting, and Random Forest). Their results were unambiguous: ML models demonstrated about 10% greater accuracy compared to standard methods. The accuracy rate of the testing sample with Random Forest, specifically, was 87%. They ascribed this success to the fact that the ensemble methods were able to minimize variance and model the non-linear relationships.

This analysis was furthered by Moscatelli et al. [14] to the Italian market, which is concerned with corporate default forecasting. They discovered that ML models (Random Forest, Gradient Boosting) had significant increases in discriminatory power and precision when public financial information was only available. They did, however, observe that this advantage faded when confidential information was added, so that the superiority of ML is most effective in data-limited settings.

Ren [17] systematically compared eight ML models based on the data of the Wind database in the Chinese context. His results showed that Random Forest (Accuracy: 84.16%) and XGBoost (Accuracy: 79.87%) significantly outperformed other models. He found that profitability and bond characteristics are the most important predictors using factor importance analysis. Gao [11] also contributed to the field by presenting an LDA-XGBoost model, which incorporates non-structured textual data of annual reports, and his results show that the tone of disclosure of management risk can significantly enhance predictive performance.

2.3 Model Interpretability and Hybrid Approaches

Although they are accurate, ML models have been labeled as black boxes. This is a significant obstacle to adoption in regulated industries due to this lack of transparency [1]. Lundberg and Lee [12] proposed SHAP, a unified framework that is a cooperative game theory, offering both a global and a local interpretability. SHAP has been effectively used in credit scoring [6] and bond default prediction [16].

Dai [10] experimented with combining methods by using the KMV distance to default [13] as an input to a Random Forest model. The KMV model is used to estimate the probability of default in the future based on the volatility of the stock prices and leverage. Dai discovered that the hybrid model was very accurate and offered a theoretically based risk measure, which increased interpretability.

The paper is based on these premises, which include employing the IFind database, SHAP to obtain interpretability, and assessing a hybrid KMV-XGBoost model.

3. METHODOLOGY

3.1 Data Source and Sample Construction

All data is obtained through the IFind database, which is a full-fledged financial data platform run by Tonghua Shun. IFind covers the Chinese bond market extensively, such as issuer-level financial statements, bond characteristics, and macro-economic indicators. The data is between January 1, 2014, and December 31, 2024. This period is pivotal because it is the whole period of post rigid repayment.

We will create a sample of corporate bonds that were issued in the Chinese onshore market. We do not include financial institutions (banks, insurance, securities) because they have different regulatory capital structures. We also do not include bonds issued by local government financing vehicles (LGFVs) as their credit profile is significantly affected by implicit government guarantees. The last sample is 3,852 firm-years over 412 different bond issuers.

Any interest or principal payment missed, or the declaration of bankruptcy by the issuer, is considered a default event. We follow the defaulting event of a specific issuer. In order to have a predictive framework, we employ financial information from the preceding annual report prior to the default event. With the non-defaulted firms, we apply the latest annual report of the same period.

3.2 Variable Selection and Definitions

We built 39 predictive variables with four categories, as shown below.

3.2.1 Financial Ratios (25 variables)

(1) Profitability: Return on Equity (ROE) is the net income divided by equity of the shareholders. Return on Assets (ROA) is the net income over the total assets. A ratio of net income to revenue, known as Net Profit Margin. The Gross Profit Margin is gross profit divided by revenue. Operating Margin, which is operating income/revenue.

(2) Leverage: Debt-to-Asset Ratio- total debt/total assets. Debt-to-Equity Ratio, which is the total debt/shareholders' equity. Interest Coverage Ratio, which is the EBIT/interest expense. Long-term Debt-to-Capital Ratio, which is long-term debt/ total capital.

(3) Liquidity: Current Ratio, which is given by current assets divided by current liabilities. Quick Ratio: The ratio of current assets (less inventory) to current liabilities. Cash Ratio: This is the cash and cash equivalents divided by current liabilities. Running Cash Flow Ratio: The ratio of operating cash flow to current liabilities.

(4) Solvency: Cash Flow-to-Debt Ratio, which is the operating cash flow divided by the overall debt. EBITDA-to-Interest Ratio, which refers to EBITDA/interest expense.

(5) Efficiency: Asset Turnover, which is revenue/ total assets. Average Inventory/cost of goods sold=Inventory Turnover. Receivables Turnover, which is also referred to as revenue / average accounts receivable.

3.2.2 Bond-Specific Features (6 variables)

(1) Credit Rating: Converted to a 20-point numerical scale (AAA=20, AA+=18, AA=16, A+=14, A=12, A-=10, BBB+=8, BBB=6, BBB-=4, BB+=2, BB and below=0).

(2) Time-to-Maturity: Years to maturity of bond.

(3) Issue Size: Face value of the bond issue (in billions of RMB).

(4) Coupon Rate: Annual interest rate (percentage points).

(5) Type Dummy Bond: corporate bond (1) vs medium-term note (0).

(6) Collateral Type Dummy: (1) secured and (0) unsecured bonds.

3.2.3 Macroeconomic Indicators (5 variables)

(1) GDP Growth rate: Real GDP growth rate on an annual basis (National Bureau of Statistics).

(2) CPI Growth: Consumer price index change (percentage points) annually.

(3) M2 Growth: The broad money supply (M2) yearly growth rate.

(4) 10-Year Government Bond Yield: Proxy for the risk-free rate and credit conditions.

(5) Industry Default Rate: rolling 12-month default rate of the issuer industry.

3.2.4 Firm Characteristics (3 variables)

(1) Ownership Structure Dummy: Indicator for State-Owned Enterprise (SOE=1) vs Private Enterprise (0).

(2) Firm Age: Years of incorporation.

(3) Total Assets: transformed to log to reflect firm size.

Winsorization of all continuous variables at the 1st and 99th percentiles is implemented to remove the effects of outliers, and the results are standardized to achieve a mean of zero and a standard deviation of one.

3.3 Machine Learning Models

We use eight representative models in accordance with Ren [17]:

3.3.1 Linear Models (Baseline)

(1) Logistic Regression (LR): It is a generalized linear model that was employed as a benchmark. The default probability is represented as:

$$P(Y = 1 \mid X) = \frac{1}{1 + e^{-\left(\beta_0 + \sum_{i=1}^{p} \beta_i X_i\right)}} \tag{1}$$

(2) Lasso Regression (L1): Lasso Regression is a logistic regression with an L1 penalty, which promotes sparsity. The objective function is:

$$\min_{\beta} \left\{ -\sum_{i=1}^{n} [y_i \log(p_i) + (1 - y_i)\log(1 - p_i)] + \lambda \sum_{j=1}^{p} | \beta_j | \right\} \tag{2}$$

(3) Ridge Regression (L2): Introduces an L2 penalty to control multicollinearity. The purpose function is:

$$\min_{\beta} \left\{ -\sum_{i=1}^{n} [y_i \log(p_i) + (1 - y_i)\log(1 - p_i)] + \lambda \sum_{j=1}^{p} \beta_j^2 \right\} \tag{3}$$

(4) Elastic Net: It is a combination of L1 and L2 penalties, which provide a trade-off between the feature selection and regularization [19].

3.3.2 Non-linear Ensemble Models

(1) Random Forest (RF): An ensemble of decision trees by bagging. It constructs numerous trees using bootstrapped samples and averages their forecasts [5]. The algorithm can resist overfitting and is able to manage high-dimensional data.

(2) XGBoost: Highly optimized and scalable gradient boosting algorithm [8]. It constructs the weak decision trees in a progressive manner, where each new tree rectifies the mistakes of the past tree. The goal function is:

$$Obj = \sum_{i=1}^{n} L\left(y_i, \hat{y}_i\right) + \sum_{k=1}^{K} \Omega\left(f_k\right) \tag{4}$$

With the first term being the loss function and the second term being a regularization penalty.

3.3.3 Other Models

(1) Support Vector Machine (SVM): A maximum-margin classifier with radial basis function (RBF) kernel to model non-linear decision boundaries.

(2) Deep Neural Network (DNN): A three-layer (256, 128, 64 neurons) feed-forward neural network. Our hidden layers are activated with ReLU, regularized with dropout (0.3), and the final classification output is activated with sigmoid.

3.4 Hybrid KMV-XGBoost Model

In order to improve interpretability, we suggest a hybrid KMV-XGBoost model. The KMV model gives a distance to default (DD) measure that is proactive in nature, and is obtained as:

$$DD = \frac{V - D}{\sigma_V} \tag{5}$$

where V is the firm's asset value, D is the default point (total short-term debt + 0.5 × long-term debt), and σV is the volatility of asset returns [13]. The Merton model framework is applied to the IFind data in order to implement the KMV model.

The computed DD of every firm-year is introduced as a novel input characteristic to the XGBoost model. The hybrid model is then trained and tested as an integrated system. The main benefit is that the DD offers a theoretically-based continuous risk score, which can be directly interpreted by risk managers.

3.5 Experimental Design

3.5.1 Handling Class Imbalance

The data is very imbalanced (approximately 8% of observations are defaults). We use the Synthetic Minority Over-sampling Technique (SMOTE) to balance the classes to an equally weighted 1:1 ratio on the training set [7]. SMOTE generates synthetic samples by interpolating between the existing samples of minority classes.

3.5.2 Hyperparameter Tuning

Cross-validation on the validation set: A grid search is optimized on the validation set to find the best hyperparameters on each model. In the case of XGBoost, we tune the learning rate, maximum tree depth, subsample ratio, and estimators.

3.5.3 Evaluation Metrics and Statistical Testing

Since the class is imbalanced, the accuracy is not sufficient. We employ: AUC-ROC (discriminatory power), F1-score (harmonic mean of precision and recall), Recall (sensitivity), and Precision (positive predictive value). Our main measure of model comparison is the F1-score. To assess the statistical significance of performance differences among models, we conduct paired DeLong tests for AUC-ROC comparisons and bootstrap confidence intervals (with 1,000 resamples) for F1-score differences, ensuring that the observed improvements are not due to random chance.

4. RESULTS

4.1 Model Performance Comparison

Table 1 shows the results of all eight models on the hold-out test set.

Table 1. Model Performance on the Test Set (2023-2024).

Model	AUC-ROC	F1-Score	Recall	Precision	Accuracy
Logistic Regression	0.78	0.32	0.42	0.26	0.88
Lasso (L1)	0.79	0.34	0.44	0.27	0.89
Ridge (L2)	0.79	0.33	0.43	0.27	0.89
Elastic Net	0.80	0.35	0.45	0.28	0.89
Random Forest	0.91	0.54	0.62	0.48	0.95
XGBoost	0.92	0.58	0.74	0.48	0.96
SVM	0.85	0.43	0.52	0.36	0.92
DNN	0.87	0.47	0.55	0.40	0.93

Key Findings:

4.1.1 Linear Models Significantly Underperform

he linear models, even with L1/L2 regularization achieve an F1-score of less than 0.35 and a recall of less than 0.45. This implies that they are not present during over half of all defaults. They are also very inaccurate (large false positive rate). This empirically validates the non-linearity of default risk [4, 14].

4.1.2 Here is a dramatic improvement in Ensemble Models

XGBoost has an F1-score of 0.58, which is 81 percent higher than the logistic regression F1-score. It has a recall of 0.74, indicating that it detects three-quarters of all defaults, which is essential in a risk management system. The AUC-ROC of 0.92 is an excellent discriminative ability.

4.1.3 XGBoost is better than the random forest

XGBoost outperforms Random Forest in both AUC-ROC (0.92 vs 0.91) and F1-score (0.58 vs 0.54). This benefit is probably due to the gradient boosting nature of XGBoost that aims at systematically mitigating errors and the regularization inherent in it [8].

4.1.4 DNN Performs Competitively but Lags Behind XGBoost

The DNN model has a competitive performance (F1-score: 0.47), but it is not as good as XGBoost. This correlates with results in tabular data domains that have structure, where gradient-boosted trees tend to perform better than deep learning in the absence of large datasets.

4.2 Robustness Checks

To check our findings, we conduct two robustness checks:

4.2.1 Sample Rebalancing

We rerun the XGBoost model with different SMOTE ratios (1:2, 1:5). The F1-score does not change much (0.55-0.58), and this means that the model is resistant to various class balancing assumptions.

4.2.2 Time Window Shift

We change the training/validation/test window to 2015-2021/2022/2023. XGBoost performance only decreases a little (AUC-ROC: 0.91, F1: 0.55), and indicates that the predictive logic of the model is predictable in the long run.

4.3 SHAP Interpretability Analysis

To address the “black box” problem, we use SHAP to analyze the decision logic of the best-performing XGBoost model [12]. SHAP values give a combined value of feature importance using cooperative game theory.

4.3.1 Global Feature Importance

SHAP summary plot ranks features based on the average absolute SHAP value over all predictions on the test set. The five highest features are:

(1) Return on Equity (ROE)

(2) Debt-to-Asset Ratio

(3) Quick Ratio

(4) Cash Flow-to-Debt Ratio

(5) SOE Dummy

The official credit rating is lower at 7th ranking, which is lower than important financial ratios. It is a very important discovery. It confirms that the ML model is trained to disregard the rating to make predictions and instead uses raw financial information. This empirically confirms the apprehensions of rating inflation in the Chinese market [17, 11].

Non-Linear Relationships: The SHAP dependence plot of ROE shows that there exists a non-linear relationship. The impact is symmetric and monotonic, but the steepness of the impact at extreme tails cannot be adequately modeled using linear models.

4.3.2 Interaction Effects

Dependence plot of Debt-to-Asset Ratio and Quick Ratio shows that there exists a strong interaction effect. In the case of highly-debted firms, there is a steep rise in SHAP value with a rise in debt. The SHAP value is, however, multiplied exponentially when low liquidity (low Quick Ratio) is also present in the firm. This shows that the compounding non-linear interaction of high leverage and low liquidity — a major risk mechanism — is captured by the ML model that cannot be captured by traditional models.

4.3.3 Local Explanations

A force plot of a particular firm that defaulted in 2023 but had an "AA" rating right before default indicates that three features were the major contributors to the high default probability (0.85) in the model:

(1) ROE (Negative Effect)

(2) Quick Ratio (Negative Effect)

(3) Debt-to-Asset Ratio (Positive Effect)

(4) Credit Rating (Positive Effect)

This localization offers a clear, case-to-case rationale behind the model decision to explain why there can be a firm with a good rating and still be considered high-risk.

4.4 Hybrid KMV-XGBoost Model Results

The KMV-XGBoost hybrid model was developed and tested on the same test set. The Merton model was used to compute the distance to default (DD) measure of each firm-year.

4.4.1 Performance

AUC-ROC: 0.91, F1-score: 0.55, Recall: 0.69.

4.4.2 Comparison with XGBoost

The performance is marginally worse than the typical XGBoost (AUC-ROC: 0.92, F1-score: 0.58). The difference is, however, minor, and the primary benefit of the hybrid model is not in performance but in interpretability.

4.4.3 SHAP Analysis with KMV

The DD measure was rated 4th in total significance. The fact that it is included made the other features a little less important, and it offered a theoretically sound risk measure, which correlates with the distance to default of the firm. This renders the predictions of this model more open and justifiable to risk managers who need a theoretical explanation [10, 13].

5. DISCUSSION

5.1 Summary of Key Findings

This research empirically confirmed the efficiency of tree-based ensemble models in bond default prediction in the Chinese market on the IFind database. The XGBoost model performed best with an F1-score of 0.58 and a recall of 0.74, which was much better than the traditional linear models. This validates the underlying nature of non-linear relationships and complex interactions in the process of capturing default risk [4, 17, 14].

The SHAP analysis indicated that the most important predictors are profitability (ROE), leverage (Debt-to-Asset Ratio), and liquidity (Quick Ratio). The poor predictive ability of the official credit ratings (ranked 7th) is an empirical demonstration of the fear of rating inflation and low predictability in the Chinese market [17, 11].

The hybrid KMV-XGBoost model, was found to have the same accuracy as the standard XGBoost, and it provided a theoretically motivated distance to the default metric. This is a promising avenue of risk management applications in which interpretability is crucial.

5.2 Theoretical Implications

This paper further reveals the Chinese context in the studies of Ren [17] and Moscatelli et al. [14] to prove the universally high superiority of ML models in predicting credit risk. It also empirically confirms the utmost significance of non-linearities. The SHAP analysis revealed a significant interaction effect among high leverage and low liquidity — a compound risk that cannot be modeled by linear models. The discovery leads to the increasing body of explainable AI in finance [12, 6].

5.3 Practical Implications for Risk Managers

5.3.1 Use XGBoost to screen Portfolios

The model provides an effective screening bond portfolio tool. It has a high recall (0.74), implying that it will capture most of the defaults, which is most vital in loss prevention.

5.3.2 Actionable Insights with SHAP

A risk manager can not only get a risk score but, with SHAP, can create a risk report of a flagged issuer, which can indicate precisely what financial metrics are contributing to the large risk score. This can enable more focused due diligence and early intervention.

5.3.3 Use the Hybrid KMV-XGBoost Model for Regulatory Compliance

The hybrid model entails a marginally reduced yet good performance, but it offers a theoretically based distance to default metric. It is especially useful to risk committees and regulatory bodies that need a clear, defensible, and transparent risk assessment framework [10].

5.4 Implications for Regulatory Policy

The lowest predictive power of official credit ratings is the most critical finding. The SHAP analysis indicates that the XGBoost model is successful in learning to disregard the rating and making use of raw financial data. This empirically confirms the issue of rating inflation and the lack of forward-looking in the Chinese market [17, 11].

This finding suggests that the regulatory framework should be reformed to:

(1) Make credit rating agencies reveal their models and input data.

(2) Incorporate dynamic ratings that change at a faster rate.

(3) Introduce ML-based scores on risk as an additional pillar to the formal rating system.

(4) Create a market-driven, transparent, and effective credit assessment ecosystem.

5.5 Limitations

There are a few limitations to this study. First, the data used is limited to structured financial and macroeconomic information. Second, the existing models are non-adaptive and are conditioned on past data, and may not be able to respond to changes in the market. Third, it is not prescriptive, but predictive.

6. CONCLUSION AND FUTURE WORK

In this paper, a systematic interpretable machine learning predictive framework of bond default in the Chinese market was developed using the IFind database. As a result of a strict comparison of eight models, we have shown that the XGBoost is the best model with an F1-score of 0.58 and a recall of 0.74. These high-performing, black-box, models were then converted into transparent, decision-support tools using SHAP and a hybrid KMV-XGBoost model.

It was found that the default risk was primarily driven by financial fundamentals, namely ROE, Debt-to-Asset Ratio, and Quick Ratio. Also, the ML model does not pay much attention to official credit ratings, which indicates a serious necessity for regulatory reform.

The future work should combine the unstructured text data of annual reports with the help of NLP [11], apply it to online learning and time-series models (e.g., LSTM) to capture the temporal patterns, build prescriptive analytics to advise on counterfactual steps in case of risks, and apply the model to smaller companies with the help of alternative data sources.

REFERENCES

[1] Adadi, A., & Berrada, M. (2018). Peeking inside the black box: A survey on explainable artificial intelligence (XAI). IEEE Access, 6, 52138-52160.

[2] Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. The Journal of Finance, 23(4), 589-609.

[3] Altman, E. I., Haldeman, R. G., & Narayanan, P. (1977). ZETA analysis: A new model to identify bankruptcy risk of corporations. Journal of Banking & Finance, 1(1), 29-54.

[4] Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. Expert Systems with Applications, 83, 405-417.

[5] Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32.

[6] Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2021). Explainable machine learning in credit risk management. Computational Economics, 57(1), 203-223.

[7] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 16, 321-357.

[8] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785-794).

[9] China Central Depository and Clearing Co. (2024). China bond market statistics yearbook 2024. Beijing: China Financial Publishing House.

[10] Dai, H. (2022). Optimization design of corporate credit bond default risk early warning model—Based on KMV-Random Forest algorithm. Master's thesis, China.

[11] Gao, L. (2025). Research on bond default prediction of listed companies based on LDA-XGBoost. Master's thesis, Dalian Jiaotong University, China.

[12] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems (Vol. 30, pp. 4765-4774).

[13] Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. The Journal of Finance, 29(2), 449-470.

[14] Moscatelli, M., Parlapiano, F., Narizzano, S., & Viggiano, G. (2020). Corporate default forecasting with machine learning. Expert Systems with Applications, 161, 113635.

[15] Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. Journal of Accounting Research, 18(1), 109-131.

[16] Pan, Y., Zhou, Y., & Wang, J. (2022). Interpretable bond default prediction using SHAP: A case study of the Chinese market. Journal of Financial Risk Management, 21(4), 789-812.

[17] Ren, G. (2021). Machine learning-based bond default risk prediction research. Master's thesis, Southwestern University of Finance and Economics, China.

[18] Shin, K. S., Lee, T. S., & Kim, H. J. (2004). An application of support vector machines in bankruptcy prediction model. Expert Systems with Applications, 28(1), 127-135.

[19] Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B, 67(2), 301-320.