Applying Machine Learning to Predict Financial Distress: A Case Study of Vietnamese Real Estate Companies

Thi-Thu-Trang Dam[1], Thi-Lich Nghiem[2]

1 Faculty of Mathematical Economics, Thuongmai University, Hanoi, Vietnam
2 Faculty of Economic Information System and E-Commerce, Thuongmai University, Hanoi, Vietnam

DOI: https://doi.org/10.56293/IJMSSSR.2026.6008

IJMSSSR 2026
VOLUME 8
ISSUE 1 JANUARY - FEBRUARY **ISSN: 2582 – 0265**

Abstract: Financial distress prediction plays a crucial role in corporate risk management, investment decision-making, and financial stability, particularly in emerging markets where firms are more vulnerable to economic shocks. The real estate sector in Vietnam is characterized by high leverage, long investment cycles, and strong sensitivity to macroeconomic fluctuations, making early detection of financial distress especially important. In this context, traditional statistical models often face limitations due to their reliance on linear assumptions and their inability to handle complex and imbalanced data. This study aims to enhance financial distress prediction by applying machine learning techniques to listed real estate firms in Vietnam. Using firm-level financial data collected from Vietnamese stock exchanges over the period 2019–2023, the study constructs a binary financial distress variable based on return on assets (ROA). Several widely used machine learning algorithms, including Logistic Regression, Random Forest, Support Vector Machine, and XGBoost, are employed and systematically compared. To address the inherent class imbalance in financial distress data, a data balancing technique is incorporated into the modeling process, and model performance is evaluated before and after data balancing. The proposed research framework integrates data exploration and preprocessing, data balancing, and model development and evaluation. Model performance is assessed using multiple evaluation metrics that are appropriate for imbalanced classification problems. By explicitly examining the impact of data imbalance on prediction performance, this study provides empirical evidence on the importance of data preprocessing in financial distress forecasting. The findings offer practical insights for investors, financial institutions, and policymakers in developing more effective early warning systems for financial distress in the Vietnamese real estate sector.

Keywords: Financial distress prediction, Machine learning, Imbalanced data, Real estate firms

1. Introduction

Financial distress represents a critical deterioration stage in a firm's financial condition and is widely regarded as a precursor to bankruptcy, default, or debt restructuring. The early identification of financial distress plays a crucial role in corporate governance and risk management, as it enables firms to take timely corrective actions while assisting investors, creditors, and regulators in mitigating potential losses. Consequently, financial distress prediction has long been a central topic in corporate finance and accounting research.

Traditionally, financial distress prediction has relied on statistical and econometric models such as Altman's Z-score, discriminant analysis, and logistic regression. Although these methods are relatively transparent and easy to implement, they are based on restrictive assumptions, most notably linear relationships among financial variables and therefore often struggle to capture the complex and non-linear nature of corporate financial dynamics. As business environments become increasingly volatile and firm-level financial data grow in both scale and complexity, the limitations of traditional models have become more apparent (Zhao et al., 2024).

In recent years, the rapid development of machine learning techniques has offered new opportunities for financial risk forecasting. Machine learning models, including tree-based ensembles and kernel-based algorithms, are capable of learning non-linear patterns and intricate interactions among financial indicators without imposing restrictive parametric assumptions. A growing body of literature demonstrates that machine learning models generally outperform conventional statistical approaches in predicting financial distress, particularly in complex

and uncertain environments (Hoang & Wiegratz, 2023). These advantages are especially relevant for emerging and transition economies, where financial data are often noisy, incomplete, and highly imbalanced (Nguyen et al., 2024).

Vietnam provides a compelling context for financial distress research. The real estate sector is one of the key pillars of the Vietnamese economy, contributing significantly to economic growth, employment, and capital market development (K. L. Tran et al., 2022). However, it is also among the most financially vulnerable sectors due to its heavy reliance on leverage, long investment cycles, and sensitivity to macroeconomic shocks. The COVID-19 pandemic and recent tightening of credit conditions have further exacerbated financial risks in the real estate industry, leading to a growing number of firms experiencing financial distress. As a result, accurately predicting financial distress in this sector has become increasingly important for investors, lenders, and policymakers (T. Tran et al., 2023).

Several recent studies have applied machine learning techniques to predict financial distress among Vietnamese listed firms and have reported encouraging results (Dao, 2024; K. L. Tran et al., 2022). Nevertheless, most existing studies primarily focus on comparing different algorithms in terms of overall accuracy, while paying limited attention to a critical challenge inherent in financial distress data named class imbalance. In practice, financially distressed firms typically represent only a small fraction of the overall population. This imbalance can bias classification models toward the majority (non-distressed) class, resulting in deceptively high overall accuracy but poor performance in identifying distressed firms, which are of greatest practical concern.

Despite the growing recognition of data imbalance as a key issue in classification tasks, empirical evidence on its impact in the context of financial distress prediction for Vietnamese real estate firms remains scarce. In particular, few studies systematically examine how data balancing techniques affect the performance of different machine learning models when predicting financial distress in this high-risk sector.

Addressing this research gap, the main contributions of this paper are summarized as follows:

- Evaluate the predictive performance of several widely used machine learning algorithms in forecasting financial distress among listed real estate companies in Vietnam;
- Analyze the impact of data imbalance and data balancing techniques on model performance, with a particular focus on the minority (financially distressed) class; and
- Provide practical implications for investors, financial institutions, and policymakers in designing more effective early warning systems for financial distress.

The remainder of this paper is organized as follows. Section 2 provides a review of prior studies on financial distress prediction and the application of machine learning techniques in this field. Section 3 introduces the proposed research framework, including data exploration, variable construction, and model development. Section 4 reports the experimental design and empirical results. Section 5 discusses the main findings and their practical implications. Finally, Section 6 concludes the paper and suggests directions for future research.

2. Related work

Financial distress prediction has been an active research area in corporate finance and accounting for several decades. Early studies primarily relied on traditional statistical and econometric models, such as discriminant analysis and logistic regression, using accounting-based financial ratios as predictors. Comprehensive surveys indicate that these models have played a foundational role in the development of financial distress research; however, their effectiveness is often constrained by strong assumptions regarding linearity, normality, and variable independence (Zhao et al., 2024). As corporate financial data become increasingly complex and high-dimensional, these limitations have motivated researchers to explore more flexible and data-driven approaches.

In recent years, machine learning techniques have gained substantial attention in financial distress prediction due to their superior ability to capture non-linear relationships and complex interactions among financial variables. Extensive reviews highlight a clear methodological shift from traditional statistical models toward machine learning and deep learning approaches, including decision trees, ensemble methods, support vector machines, and

neural networks (Sethi & Mahadik, 2025; Zhao et al., 2024). Empirical evidence consistently shows that machine learning models tend to outperform conventional approaches in terms of predictive accuracy and robustness, particularly in environments characterized by uncertainty and structural complexity (Hoang & Wiegratz, 2023). More advanced architectures, such as hybrid deep learning models and time-series-based frameworks, have further enhanced predictive performance by incorporating temporal dynamics and optimizing model parameters (Al Ali et al., 2023; Ayuni et al., 2024).

The application of machine learning to financial distress prediction has been especially prominent in emerging and transition economies, where institutional differences, data constraints, and market inefficiencies pose significant challenges to traditional modeling approaches. In (Rahman & Zhu, 2024), using data from China, demonstrate that non-linear machine learning models substantially outperform regression-based methods in forecasting financial distress. Similarly, (Kristanti et al., 2025) provide evidence from emerging markets showing that machine learning–based financial distress models contribute not only to improved prediction accuracy but also to broader financial sustainability objectives. In the context of transition economies, (Nguyen et al., 2024) emphasize that financial distress data are typically scarce and highly imbalanced, and show that incorporating data balancing techniques such as SMOTE can significantly affect model performance, particularly for minority (distressed) firms.

In Vietnam, a growing body of literature has examined financial distress prediction using both traditional and machine learning approaches. (K. L. Tran et al., 2022) apply explainable machine learning techniques to a large sample of Vietnamese listed firms and find that ensemble models, particularly XGBoost and Random Forest, achieve the highest predictive performance. Their study further highlights the importance of model interpretability through SHAP values, providing insights into key financial drivers of distress. Expanding on this line of research, (Dao, 2024) evaluates the predictive ability of various machine learning models for firms listed on the Ho Chi Minh City Stock Exchange and emphasizes the relevance of accounting-based variables in the Vietnamese market context. Complementary evidence using accounting-based distress measures also confirms that the construction and real estate sector exhibits the highest exposure to financial distress risk among Vietnamese industries, especially during periods of macroeconomic shocks such as the COVID-19 pandemic (T. Tran et al., 2023).

A parallel stream of research focuses specifically on the property, real estate, and construction sectors, which are widely recognized as being particularly vulnerable to financial distress due to their capital-intensive nature, high leverage, and sensitivity to business cycles. Empirical studies from emerging markets, including Indonesia and other Asian economies, consistently identify profitability, liquidity, and leverage as key determinants of financial distress in real estate firms (Dewi et al., 2023; Marcella & Utami, 2024; Saraswati, 2025). While many of these studies rely on traditional regression-based methods, recent research has begun to incorporate machine learning techniques in sector-specific contexts. For example, (Ayuni et al., 2022) apply support vector machines to predict financial distress in property and real estate companies, while (Jeong & Kim, 2022) compare multiple machine learning models for medium- to long-term financial distress prediction in the construction industry, demonstrating the advantages of ensemble-based approaches.

3. Proposed method

This study proposes a machine learning–based framework for forecasting financial distress among listed real estate firms in Vietnam. The proposed framework consists of three main stages: (i) Data preparation and variable construction, (ii) Model development and prediction, and (iii) Model evaluation and managerial implications. Figure 1 illustrates the overall research framework employed in this study, which consists of three main stages: data preparation, machine learning modeling, and model evaluation and knowledge presentation.

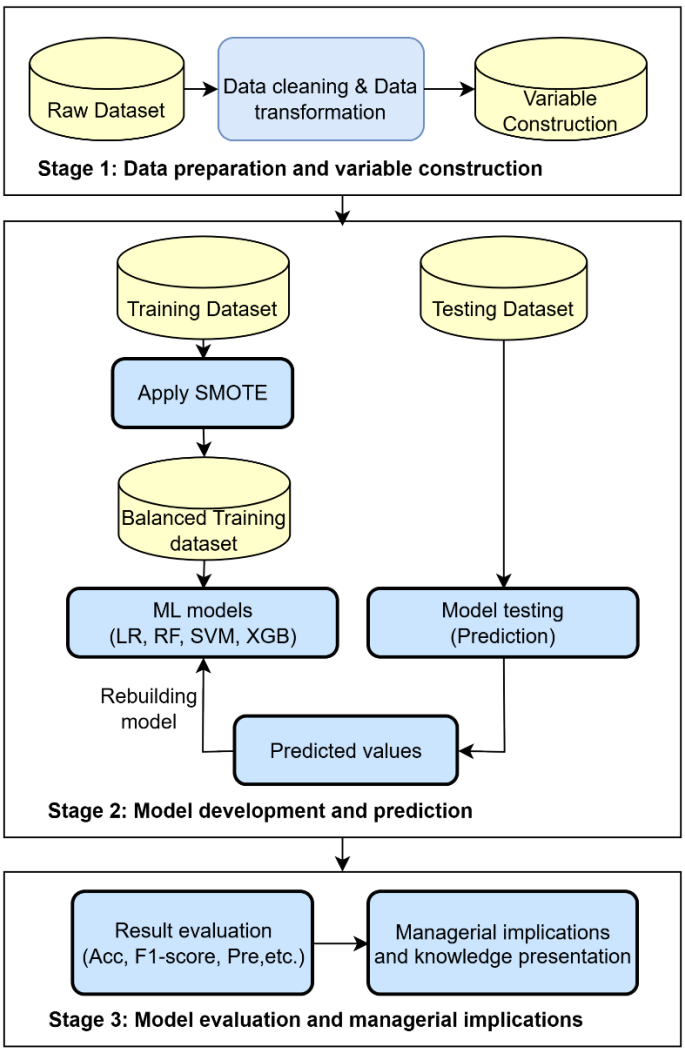

Figure 1. The proposed financial distress prediction framework

(1) Data preparation and variable construction.

In the first stage, raw financial statement data are collected and prepared for analysis. Data cleaning and transformation procedures are conducted to ensure data consistency, including handling missing values, reducing noise, and standardizing variable scales.

Based on the cleaned data, financial variables are constructed to capture key aspects of firms' financial conditions, such as asset structure, leverage, profitability, and operating efficiency. These constructed variables serve as input features for the machine learning models. This stage aims to transform raw accounting data into a structured and informative dataset suitable for financial distress prediction.

(2) Model development and prediction.

In the second stage, the processed dataset is divided into training and testing sets using stratified sampling to preserve the original class distribution. Given the severe class imbalance inherent in financial distress data, the Synthetic Minority Over-sampling Technique (SMOTE) is applied exclusively to the training dataset.

By generating synthetic observations for the minority class, SMOTE helps mitigate bias toward the majority class while preventing information leakage into the testing set. Subsequently, several machine learning models, including Logistic Regression, Random Forest, Support Vector Machine, and XGBoost, are trained using the

balanced training data. The trained models are then employed to generate financial distress predictions for the testing dataset.

(3) Model evaluation and managerial implications.

In the final stage, model performance is evaluated based on predictions obtained from the testing dataset. Multiple evaluation metrics are employed to comprehensively assess classification performance, including accuracy, precision, recall, F1-score, and ROC-AUC.

Special attention is paid to the predictive ability for the minority class, as correctly identifying financially distressed firms is of primary practical importance. Based on the evaluation results, managerial implications are derived to support investors, financial institutions, and policymakers in developing effective early warning systems for financial distress.

4. Experimental Results

4.1. Dataset

The dataset used in this study consists of financial statement data of Vietnamese listed real estate firms over the period 2019–2023. The data are collected from publicly available annual financial reports of firms listed on the Ho Chi Minh Stock Exchange (HOSE), Hanoi Stock Exchange (HNX), and the UPCoM market.

Each observation corresponds to a firm-year record and includes accounting variables reflecting firms' asset structure, leverage, profitability, operating performance, and working capital management. Based on these variables, the dependent variable representing financial distress is constructed using return on assets (ROA), where firms with negative ROA are classified as financially distressed, and those with non-negative ROA are classified as non-distressed. This definition is particularly suitable for emerging markets such as Vietnam, where formal bankruptcy or default records are limited and often delayed.

The independent variables are derived from raw accounting data and grouped into four categories: (i) asset structure, (ii) leverage and liabilities, (iii) operating performance, and (iv) working capital management. These variables include both level-based indicators (e.g., total assets, revenues) and ratio-based measures that reflect firms' financial structure and performance. Prior to model training, continuous variables are scaled using standardization to reduce the influence of magnitude differences and improve model convergence.

The dataset exhibits a high degree of class imbalance, with non-distressed observations accounting for approximately 88.7% of the sample, while distressed observations represent only 11.3%. Such imbalance is typical in financial distress prediction problems and poses a significant challenge for classification models, as it may lead to biased predictions toward the majority class. This characteristic motivates the adoption of data balancing techniques in the subsequent modeling stage.

4.2. Evaluation Metrics

To evaluate the performance of the machine learning models, this study adopts multiple classification metrics that are suitable for imbalanced datasets. Let TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively.

Accuracy measures the overall proportion of correctly classified observations:

$$\text{Accuracy} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}$$

Precision for the distressed class evaluates the proportion of correctly predicted distressed firms among all firms predicted as distressed:

$$\text{Precision} = \frac{TP}{TP + FP}$$

Recall (or sensitivity) measures the ability of the model to correctly identify distressed firms:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

F1-score is the harmonic mean of precision and recall:

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

In addition, the ROC-AUC metric evaluates the model's discriminative ability across all classification thresholds by measuring the area under the Receiver Operating Characteristic curve.

Given the severe class imbalance, particular emphasis is placed on recall and F1-score for the financially distressed class, as failing to identify distressed firms may lead to substantial economic consequences.

4.3. Empirical Results

This section reports and discusses the empirical results obtained from machine learning models under two experimental settings: (i) original imbalanced data and (ii) balanced data using SMOTE. Model performance is evaluated using Accuracy, Precision, Recall, F1-score, and ROC-AUC, with particular attention paid to the minority class (financially distressed firms).

Results on the original imbalanced dataset

Table 1: Performance on the original dataset

MODEL	ACCURACY	F1-SCORE (DISTRESS)	ROC-AUC	PR-AUC
Logistic Regression	0.89.2	0.000	0.507	0.108
Random Forest	0.882	0.154	0.510	0.133
XGBoost	0.882	0.267	0.533	0.155
SVM	0.892	0.000	0.582	0.146

Table 1 illustrates the predictive performance of the four machine learning models trained on the original imbalanced dataset. At first glance, all models achieve relatively high accuracy, ranging from 88.2% to 89.2%. However, this high accuracy is largely misleading due to the severe class imbalance, where financially distressed firms account for only approximately 11% of the observations.

Logistic Regression and SVM completely fail to identify any distressed firms, resulting in an F1-score of zero for the minority class. This outcome clearly illustrates the limitation of accuracy-based evaluation in imbalanced classification problems, as models can achieve high accuracy simply by predicting all firms as financially healthy.

Among the models considered, XGBoost demonstrates the best overall performance, achieving the highest F1-score (0.267) and ROC-AUC (0.533). Nevertheless, the F1-score remains relatively low, indicating that even advanced ensemble models struggle to effectively capture the characteristics of financially distressed firms when trained on imbalanced data. Overall, the results suggest that models trained on the original dataset are strongly biased toward the majority class and are therefore inadequate for practical early warning applications.

Results after data balancing

Table 2: Performance on rebalanced dataset

MODEL	ACCURACY	F1-SCORE (DISTRESS)	ROC-AUC	PR-AUC
Logistic regression	0.419	0.182	0.518	0.113
Random Forest	0.796	0.296	0.563	0.167
XGBoost	0.828	0.333	0.528	0.169
SVM	0.323	0.222	0.525	0.110

Table 2 presents the classification results after applying SMOTE to balance the training dataset. Compared with the imbalanced setting, all models exhibit substantial improvements in their ability to identify financially distressed firms, as reflected by higher F1-scores and PR-AUC values.

XGBoost achieves the best overall performance, with the highest F1-score (0.333) and PR-AUC (0.169), while maintaining a relatively high accuracy of 82.8%. This indicates that XGBoost is able to achieve a better trade-off between detecting distressed firms and preserving overall classification performance.

Random Forest also shows noticeable improvement after balancing, particularly in terms of F1-score and ROC-AUC, suggesting enhanced discriminatory power. In contrast, SVM achieves the highest recall for distressed firms but suffers from a sharp decline in accuracy, indicating a high rate of false positive predictions. Logistic Regression shows only marginal improvement, highlighting its limited flexibility in capturing complex non-linear relationships even after data balancing.

Comparative analysis

A direct comparison between Tables 1 and 2 reveals that data balancing significantly alters model behavior. While accuracy decreases after SMOTE, this decline reflects a shift away from majority-class bias rather than a deterioration in predictive quality. More importantly, key metrics for the minority class—F1-score and PR-AUC—improve consistently across all models.

These findings confirm that data balancing is a crucial step in financial distress prediction, particularly in highly imbalanced contexts such as the Vietnamese real estate sector. Models that perform well on imbalanced data in terms of accuracy may fail entirely to detect financially distressed firms, thereby limiting their practical usefulness. Among the evaluated approaches, XGBoost combined with SMOTE emerges as the most effective model, offering superior performance in identifying financially distressed firms while maintaining reasonable overall accuracy.

5. Discussion

Interpretation of Empirical Results

This study provides empirical evidence on the effectiveness of machine learning models for financial distress prediction in the Vietnamese real estate sector, with particular emphasis on the role of data imbalance. The findings yield several important insights.

First, the results confirm that high predictive accuracy on imbalanced datasets is not a reliable indicator of model effectiveness. When trained on the original dataset, all models achieved high accuracy; however, most failed to identify financially distressed firms, as reflected by near-zero F1-scores for the minority class. This outcome reinforces prior arguments in the literature that accuracy-oriented evaluation can be misleading in financial distress prediction, where the primary objective is to detect rare but critical distress events.

Second, the application of SMOTE significantly improved the minority-class performance across all models. The notable increase in F1-score and PR-AUC indicates that data balancing enables classifiers to better learn the characteristics of financially distressed firms. These findings are consistent with recent studies emphasizing the necessity of addressing class imbalance in emerging markets, where distress events are relatively rare but

economically consequential.

Third, among the evaluated algorithms, XGBoost demonstrates the most robust and stable performance after data balancing. Compared to Logistic Regression and SVM, XGBoost achieves a superior trade-off between overall accuracy and minority-class detection. While SVM exhibits high recall, its excessive false positives limit its practical applicability. This suggests that ensemble tree-based methods are particularly suitable for financial distress forecasting in complex and non-linear environments such as the real estate industry.

From a practical perspective, the results highlight that ignoring data imbalance may lead to ineffective early warning systems, potentially exposing investors and regulators to underestimated risks. By contrast, integrating data balancing techniques with advanced machine learning models can substantially enhance the reliability of financial distress prediction and support more informed decision-making.

Strategic and Managerial Implications

The findings of this study offer several important implications for practitioners and policymakers.

First, for investors and financial institutions, the results emphasize that early warning systems for financial distress should prioritize recall- and precision-based metrics rather than overall accuracy. A model that fails to detect distressed firms may expose investors and lenders to significant financial losses, even if its accuracy appears high. Therefore, machine learning models such as XGBoost, combined with data balancing techniques, provide a more effective tool for screening financially vulnerable real estate firms.

Second, for corporate managers in the real estate sector, the results highlight the importance of monitoring non-linear interactions among financial indicators related to leverage, profitability, and liquidity. Traditional linear models may underestimate financial risk during periods of economic stress, while machine learning-based systems can better capture early signals of financial deterioration. Firms can incorporate such predictive tools into internal risk management and strategic planning processes to enhance financial resilience.

Third, from a regulatory and policy perspective, the evidence suggests that supervisory authorities may benefit from adopting machine learning-driven risk monitoring frameworks, particularly for high-risk sectors such as real estate. Improving the detection of financially distressed firms can support more timely regulatory interventions and contribute to greater stability in capital markets

6. Conclusion

This study investigates financial distress prediction for Vietnamese listed real estate companies using multiple machine learning models and explicitly examines the impact of data imbalance on predictive performance. By comparing model results before and after applying SMOTE, the study provides a clearer understanding of how imbalance influences classification outcomes in practice.

The empirical results demonstrate that models trained on imbalanced data tend to favor non-distressed firms, achieving high accuracy but failing to identify financially distressed companies. After data balancing, all models show improved performance in detecting distress, with XGBoost delivering the best overall results in terms of F1-score and PR-AUC. These findings underscore the importance of incorporating data balancing techniques when developing early warning systems for financial distress.

This research contributes to the existing literature in three key ways. First, it provides industry-specific evidence from the Vietnamese real estate sector, a context that remains underexplored despite its high financial risk. Second, it empirically demonstrates the critical role of data imbalance in financial distress prediction. Third, it offers practical guidance on model selection for stakeholders seeking effective risk monitoring tools.

Several limitations should be acknowledged. The study relies primarily on accounting-based financial indicators and focuses on a single industry and country. Future research could extend this framework by incorporating market-based variables, textual information, or macroeconomic indicators, as well as exploring alternative

resampling techniques and deep learning approaches. Expanding the analysis to other sectors or comparative regional contexts would also enhance the generalizability of the findings.

References

1. Al Ali, A., Khedr, A. M., El Bannany, M., & Kanakkayil, S. (2023). GALSTM-FDP: A Time-Series Modeling Approach Using Hybrid GA and LSTM for Financial Distress Prediction. *International Journal of Financial Studies*, *11*(1), 38. https://doi.org/10.3390/ijfs11010038
2. Ayuni, N. W. D., Lasmini, N. N., & Putrawan, A. A. (2022). Support Vector Machine (SVM) as Financial Distress Model Prediction in Property and Real Estate Companies. In A. Azizah & E. D. Ariyani (Eds.), *Proceedings of the International Conference on Applied Science and Technology on Social Science 2022 (iCAST-SS 2022)* (Vol. 719, pp. 397–402). Atlantis Press SARL. https://doi.org/10.2991/978-2-494069-83-1_72
3. Ayuni, N. W. D., Utthavi, W. H., & Lasmini, N. N. (2024). Artificial Neural Networks: A Deep Learning Approach in Financial Distress Prediction. In A. A. N. G. Sapteka, I. G. L. M. Parwita, I. K. Wiratama, F. Moi, K. W. Widantha, E. Septevany, D. A. I. C. Dewi, W. E. Mariani, & R. N. Fakhrurozi (Eds.), *Proceedings of the International Conference on Sustainable Green Tourism Applied Science-Engineering Applied Science 2024 (ICoSTAS-EAS 2024)* (Vol. 249, pp. 99–108). Atlantis Press International BV. https://doi.org/10.2991/978-94-6463-587-4_12
4. Dao, P.-U. (2024). *Assessing the predictive ability of companies on the Ho Chi Minh City stock exchange to predict financial distress using a machine learning model.* [UEH university]. https://digital.lib.ueh.edu.vn/handle/UEH/74229
5. Dewi, D. N., Murhadi, W. R., & Sutejo, B. S. (2023). Financial Ratios, Corporate Governance, and Macroeconomic Indicators in Predicting Financial Distress. *Journal of Law and Sustainable Development*, *11*(4), e893. https://doi.org/10.55908/sdgs.v11i4.893
6. Hoang, D., & Wiegratz, K. (2023). Machine learning methods in finance: Recent applications and prospects. *European Financial Management*, *29*(5), 1657–1701. https://doi.org/10.1111/eufm.12408
7. Jeong, J., & Kim, C. (2022). Comparison of Machine Learning Approaches for Medium-to-Long-Term Financial Distress Predictions in the Construction Industry. *Buildings*, *12*(10), 1759. https://doi.org/10.3390/buildings12101759
8. Kristanti, F. T., Salim, D. F., & Ihsan, A. F. (2025). Financial distress prediction in emerging markets by using a machine learning approach for financial sustainability. *Discover Sustainability*, *6*(1), 1296. https://doi.org/10.1007/s43621-025-02199-1
9. Marcella, R., & Utami, W. (2024). Determinan Financial Distress: A Study on The Property and Real Estate Sector in The Indonesia Stock Exchange. *Advances In Social Humanities Research*, *2*(11), 1216–1230. https://doi.org/10.46799/adv.v2i11.302
10. Nguyen, M., Nguyen, B., & Liêu, M. (2024). Corporate financial distress prediction in a transition economy. *Journal of Forecasting*, *43*(8), 3128–3160. https://doi.org/10.1002/for.3177
11. Rahman, M. J., & Zhu, H. (2024). Predicting financial distress using machine learning approaches: Evidence China. *Journal of Contemporary Accounting & Economics*, *20*(1), 100403. https://doi.org/10.1016/j.jcae.2024.100403
12. Saraswati, W. (2025). Predicting financial distress in property and real estate companies: Moderation of company size. *JAS (Jurnal Akuntansi Syariah)*, *9*(1), 244–263. https://doi.org/10.46367/jas.v9i1.2460
13. Sethi, S. R., & Mahadik, D. A. (2025). Machine learning methods for financial distress prediction: An analytical overview and implications for sustainability goals. *Journal of Chinese Economic and Business Studies*, 1–52. https://doi.org/10.1080/14765284.2025.2549237
14. Tran, K. L., Le, H. A., Nguyen, T. H., & Nguyen, D. T. (2022). Explainable Machine Learning for Financial Distress Prediction: Evidence from Vietnam. *Data*, *7*(11), 160. https://doi.org/10.3390/data7110160
15. Tran, T., Nguyen, N. H., Le, B. T., Thanh Vu, N., & Vo, D. H. (2023). Examining financial distress of the Vietnamese listed firms using accounting-based models. *PLOS ONE*, *18*(5), e0284451. https://doi.org/10.1371/journal.pone.0284451
16. Zhao, J., Ouenniche, J., & De Smedt, J. (2024). Survey, classification and critical analysis of the literature on corporate bankruptcy and financial distress prediction. *Machine Learning with Applications*, *15*, 100527. https://doi.org/10.1016/j.mlwa.2024.100527