

Performance in Predicting Financial Distress among Technology Companies Listed on the Indonesia Stock Exchange

Tri Agustin¹, Mery Sukartini², Desty Cristya Anisa³

^{1,2,3} Management Study Program, Faculty of Economics and Business, Universitas Muhammadiyah Cirebon, Indonesia

Email: triagustin515@gmail.com, mery.sukartini@umc.ac.id, destycristya.anisa@gmail.com

Abstract:

The growth of Indonesia's technology sector, accompanied by intensifying business competition and economic uncertainty, exposes companies to financial distress risk. Early detection is important to help companies, investors, and stakeholders make appropriate decisions. This study aims to analyze and compare Logistic Regression and Random Forest performance in predicting financial distress among technology companies listed on the Indonesia Stock Exchange (IDX) during 2021–2025. A quantitative approach was employed using secondary data from the annual financial statements of 41 technology companies selected through purposive sampling, yielding 205 observations. The predictors comprise Return on Assets (ROA), Net Profit Margin (NPM), Debt to Equity Ratio (DER), Current Ratio (CR), Interest Coverage Ratio (ICR), Total Asset Turnover (TATO), and Total Debt (TH). Orange Data Mining was used for analysis, with 80% of the data allocated to training and 20% to testing. Model performance was evaluated using accuracy, precision, recall, and F1-score. Random Forest outperformed Logistic Regression, achieving accuracy of 92.7%, precision of 94.5%, recall of 92.7%, and an F1-score of 93.0%. Logistic Regression achieved accuracy of 82.9%, precision of 83.7%, recall of 82.9%, and an F1-score of 83.2%. Feature importance analysis identified Net Profit Margin (NPM) as the most influential predictor of financial distress. Random Forest is therefore considered more effective than Logistic Regression in identifying potential financial distress among technology companies.

Keywords:

Financial Distress; Logistic Regression; Random Forest; Machine Learning; Indonesia Stock Exchange.

I. Introduction

1.1 Background and Context

Digital technology has transformed the global economic landscape, including Indonesia. Digital transformation encourages companies to adopt technology across business activities to improve operational efficiency, accelerate product innovation, and expand market reach. The technology sector is among the fastest-growing sectors, supported by increasing internet penetration, digital infrastructure development, and growing public dependence on technology-based services. These conditions make the sector a driver of national economic growth and an attractive investment opportunity in capital markets (Napitupulu, 2025).

According to Statistics Indonesia (BPS), gross domestic product (GDP) in the information and communication sector increased throughout 2021–2025. Alongside these developments, the Indonesia Stock Exchange (IDX) refined its industry classification through the Indonesia Stock Exchange Industrial Classification (IDX-IC) to accommodate changes in business sectors, including technology. Despite positive growth trends, increasing industry value does not necessarily indicate financially healthy companies. Many technology companies face high

technology development costs, substantial investment requirements, rapid technological change, and global economic uncertainty that can weaken financial performance.

This phenomenon is reflected in several technology companies that have experienced sustained losses, including PT GoTo Gojek Tokopedia Tbk (GOTO), PT Bukalapak.com Tbk (BUKA), and PT Envy Technologies Indonesia Tbk (ENVY). These conditions demonstrate that rapid business growth does not necessarily imply financial stability. High operating costs, research and development investment, and aggressive expansion strategies have placed some companies under financial pressure that may lead to financial distress.

Financial distress occurs when a company's financial capacity deteriorates, preventing it from meeting financial obligations on time. It is often characterized by declining profits, negative cash flows, rising debt burdens, and a reduced ability to maintain business continuity (Hutauruk et al., 2021). If not addressed promptly, financial distress may lead to bankruptcy or delisting from the Indonesia Stock Exchange. Early identification is therefore important for corporate management, investors, creditors, and regulators in making appropriate decisions.

During 2021–2025, the Indonesia Stock Exchange recorded an increase in delistings associated with financial problems, including sustained losses, negative equity, bankruptcy, and failure to meet going-concern requirements. This phenomenon indicates that financial distress is a major threat to corporate continuity in capital markets. Several technology companies also reported negative net income for consecutive years, indicating financial pressure and increasing bankruptcy risk.

In practice, financial distress prediction generally relies on financial ratio analysis and conventional statistical methods. However, these methods have limitations in capturing complex relationships and nonlinear patterns among financial variables. Developments in information technology and artificial intelligence encourage the use of machine learning as an alternative that can generate more accurate predictions by learning from historical data patterns (Iskandar et al., 2021). Compared with traditional statistical approaches, machine learning algorithms can process large datasets, recognize nonlinear relationships, and improve the classification of corporate financial conditions.

Logistic Regression and Random Forest are widely used algorithms for predicting financial distress. Logistic Regression is a classification method valued for its straightforward interpretation and ability to estimate the probability that a company will experience financial distress. Random Forest, by contrast, is an ensemble learning algorithm comprising multiple decision trees, enabling high accuracy while reducing overfitting risk. Random Forest can also identify feature importance, indicating which financial ratios contribute most to financial distress predictions.

Previous studies show that both methods perform well, although findings remain inconsistent. Kristanti and Dhaniswara (2023) report high predictive accuracy for Logistic Regression. Conversely, Febrianta et al. (2024), Mulyanto et al. (2025), and Lazuardy Sidarta (2025) find that Random Forest achieves higher accuracy than other classification methods. These differences indicate a lack of consensus regarding the most effective method, particularly for Indonesian technology companies.

A review of previous research reveals limited direct comparisons of Logistic Regression and Random Forest performance among technology companies listed on the Indonesia Stock Exchange. Machine learning applications in Indonesian financial research also remain relatively

limited compared with other countries. This study therefore aims to analyze and compare the two methods in predicting financial distress among IDX-listed technology companies during 2021–2025. The findings are expected to contribute to financial distress prediction methods and inform investors, corporate management, and academics in selecting the most accurate predictive model.

II. Review of Literature

2.1 Signaling Theory

Signaling Theory, first proposed by Spence (1973), explains that companies communicate signals about their condition and prospects to external parties through published information. For publicly listed companies, financial statements are signals that investors use to assess financial condition. Information showing improvements in profitability, liquidity, or operational efficiency provides positive signals (good news), whereas declining financial performance or sustained losses provide negative signals (bad news) that may reduce investor confidence.

In this study, financial ratios used as predictors—Return on Assets (ROA), Net Profit Margin (NPM), Debt to Equity Ratio (DER), Current Ratio (CR), Interest Coverage Ratio (ICR), Total Asset Turnover (TATO), and Total Debt (TH)—provide signals for identifying potential financial distress. Deteriorating financial ratios indicate a greater likelihood of financial difficulties.

2.2 Financial Distress

Financial distress occurs when a company's financial capacity deteriorates, preventing it from meeting operational obligations and obligations to creditors. Platt and Platt (2002) describe financial distress as a stage of financial decline preceding bankruptcy. Brigham and Ehrhardt (2022) identify its characteristics as declining profitability, an inability to generate adequate cash flows, increasing debt burdens, and difficulty meeting financial obligations.

Financial distress is a major corporate concern because it can undermine investor trust, increase the cost of capital, and ultimately lead to bankruptcy or delisting. Early detection is therefore necessary to enable corrective action before business failure occurs.

In this study, a company is classified as financially distressed if it reports negative net income for three consecutive years, in accordance with the research criteria.

2.3 Logistic Regression

Logistic Regression is a classification method used to predict dichotomous or binary dependent variables, such as whether a company is financially distressed. According to Hosmer, Lemeshow, and Sturdivant (2013), Logistic Regression estimates event probabilities through a logit function, linking independent variables to the probability of class membership.

Logistic Regression offers several advantages, including ease of interpretation, relatively simple computation, and the ability to show each variable's influence on predictions. It therefore remains a widely used statistical approach in financial distress prediction research.

However, Logistic Regression has limitations in capturing complex and nonlinear relationships among variables. Its accuracy on more complex datasets is therefore often lower than that of ensemble machine learning algorithms.

2.4 Random Forest

Random Forest is an ensemble machine learning algorithm introduced by Breiman (2001). It constructs multiple decision trees using bootstrap sampling and random feature selection, then generates the final classification through majority voting.

Random Forest's main advantages include handling high-dimensional data, reducing overfitting risk, and achieving high accuracy. It also calculates feature importance, indicating each variable's contribution to predictions and identifying the most influential predictors of financial distress.

Random Forest is widely used in financial distress prediction because it identifies complex relationships among financial ratios, yielding more accurate predictions than conventional statistical methods.

2.5 Machine Learning in Financial Distress Prediction

Machine learning is a branch of artificial intelligence that enables computers to learn patterns from historical data without being explicitly programmed. According to Goodfellow, Bengio, and Courville (2016), machine learning develops adaptive predictive models by learning from data.

In finance, machine learning is widely used to predict stock prices, credit risk, corporate bankruptcy, and financial distress. Compared with traditional statistical methods, it can process large datasets, recognize nonlinear relationships among variables, and generate higher predictive accuracy.

Advances in machine learning enable companies, investors, and regulators to develop early warning systems that detect potential financial distress more quickly and accurately.

2.6 Previous Research

Studies have compared statistical and machine learning methods for predicting financial distress. Pavličko and Mazanec (2022) find that Logistic Regression provides good classification results with relatively few variables. Kristanti and Dhaniswara (2023) also report high predictive accuracy for Logistic Regression among Indonesian companies.

Conversely, Febrianta et al. (2024), Mulyanto et al. (2025), and Lazuardy Sidarta (2025) show that Random Forest achieves higher accuracy than several other classification algorithms. This suggests that Random Forest more effectively captures complex relationships among financial variables, producing more accurate predictions.

Nevertheless, previous findings differ regarding the most effective financial distress prediction method. Research specifically comparing Logistic Regression and Random Forest among Indonesian technology companies remains relatively limited. This study therefore provides empirical evidence regarding the best-performing method for predicting financial distress among technology companies listed on the Indonesia Stock Exchange.

III. Research Methods

This study employs a quantitative comparative approach to assess Logistic Regression and Random Forest performance in predicting financial distress among technology companies listed on the Indonesia Stock Exchange (IDX). The quantitative approach was selected because the study uses numerical financial ratio data to construct classification models and evaluate each algorithm's accuracy.

3.1 Population and Sample

The population comprises all technology companies listed on the Indonesia Stock Exchange during 2021–2025. Purposive sampling was employed with the following criteria: (1) continuous listing on the IDX throughout the study period; (2) publication of complete annual financial statements; and (3) availability of the data required to calculate all research variables. These criteria yielded 41 companies and 205 observations (41 companies $\times$ 5 years).

3.2 Data Sources and Collection

The study uses secondary data from the annual financial statements of technology companies published on the official Indonesia Stock Exchange website (www.idx.co.id) and in company annual reports. Data were collected through documentation, involving the compilation, recording, and processing of financial information relevant to the research variables.

3.3 Research Variables

The dependent variable is financial distress, classified into two categories: financially distressed and non-distressed companies. Classification is based on companies' net income in accordance with the research criteria.

Table 1. Operationalization of Research Variables

Variable	Indicator
Return on Assets (ROA)	Profitability
Net Profit Margin (NPM)	Profitability
Debt to Equity Ratio (DER)	Solvency
Current Ratio (CR)	Liquidity
Interest Coverage Ratio (ICR)	Interest payment capacity
Total Asset Turnover (TATO)	Activity
Total Debt (TH)	Capital structure

These variables were selected because they represent profitability, liquidity, activity, and solvency conditions that influence the potential for financial distress.

3.4 Data Analysis

Data were analyzed using Orange Data Mining. The analysis comprised the following stages:

- 1) Data collection and preprocessing, including checks for completeness, data cleaning, and normalization where necessary.
- 2) Dataset partitioning using a hold-out method, with 80% allocated to training and 20% to testing, yielding 164 training observations and 41 testing observations.
- 3) Classification model development using Logistic Regression and Random Forest.
- 4) Model performance evaluation using a confusion matrix and four metrics:
- 5) Accuracy: the proportion of correct predictions across all observations.

- 6) Precision: the model's ability to correctly identify positive classifications.
- 7) Recall: the model's ability to detect all observations that actually belong to the positive class.
- 8) F1-score: the harmonic mean of precision and recall.

Model comparison based on accuracy, precision, recall, and F1-score to determine the best-performing algorithm. Feature importance analysis was also conducted for Random Forest to identify the financial variables contributing most to financial distress prediction.

3.5 Research Analytical Framework

The research began with the collection of financial statements from IDX-listed technology companies for 2021–2025. Seven financial ratios were then calculated as predictors: ROA, NPM, DER, CR, ICR, TATO, and TH. Data were processed using Orange Data Mining through preprocessing, training and testing partitioning, Logistic Regression and Random Forest model development, and confusion matrix evaluation. Finally, the algorithms were compared using accuracy, precision, recall, and F1-score to determine the most effective model for predicting financial distress.

IV. Result and Discussion

4.1 Research Data Description

This study uses secondary data from the annual financial statements of technology companies listed on the Indonesia Stock Exchange (IDX) during 2021–2025. Purposive sampling identified 41 eligible companies, yielding 205 observations. All data were processed using Orange Data Mining to develop Logistic Regression and Random Forest predictive models. Before classification, the dataset was partitioned using a hold-out method: 80%, or 164 observations, for training and 20%, or 41 observations, for testing. This partitioning assesses the models' ability to predict previously unseen data.

4.2 Logistic Regression Evaluation Results

The Logistic Regression model was constructed using seven predictors: Return on Assets (ROA), Net Profit Margin (NPM), Debt to Equity Ratio (DER), Current Ratio (CR), Interest Coverage Ratio (ICR), Total Asset Turnover (TATO), and Total Debt (TH). The model was evaluated using a confusion matrix and accuracy, precision, recall, and F1-score.

Table 2. Logistic Regression Evaluation Results

Indicator	Value
Accuracy	82.9%
Precision	83.7%
Recall	82.9%
F1-score	83.2%

Source: Data processing results using Orange Data Mining.

The Logistic Regression model achieved financial distress classification accuracy of 82.9%, indicating that most observations were correctly classified. Precision of 83.7% indicates reasonably good identification of financially distressed companies, while recall of 82.9% indicates that the model recognized most companies actually experiencing financial distress. The F1-score of 83.2% suggests a good balance between precision and recall.

4.3 Random Forest Evaluation Results

The Random Forest model was constructed using the same variables as Logistic Regression and evaluated using the same metrics to allow an objective comparison.

Table 3. Random Forest Evaluation Results

Indicator	Value
Accuracy	92.7%
Precision	94.5%
Recall	92.7%
F1-score	93.0%

Source: Data processing results using Orange Data Mining.

Random Forest achieved accuracy of 92.7%, exceeding Logistic Regression. Precision of 94.5% indicates that most companies predicted as financially distressed actually belonged to that category. Recall of 92.7% indicates a high success rate in detecting financially distressed companies. The F1-score of 93.0% demonstrates very good classification performance.

4.4 Performance Comparison of Logistic Regression and Random Forest

The two algorithms were compared using four model evaluation metrics ownership, and complementary investments in connectivity, literacy, and last-mile delivery. Capacity-building efforts that train producers and public-sector actors, help close the gap between innovation and sustained adoption.

Table 4. Predictive Model Performance Comparison

Indicator	Logistic Regression	Random Forest
Accuracy	82.9%	92.7%
Precision	83.7%	94.5%
Recall	82.9%	92.7%
F1-score	83.2%	93.0%

Source: Data processing results using Orange Data Mining.

Table 4 shows that Random Forest outperforms Logistic Regression across all evaluation metrics. The accuracy difference is almost 10%, indicating that Random Forest more effectively recognizes financial distress patterns among technology companies.

These results suggest that Random Forest's ensemble learning approach reduces classification errors and improves generalization. Conversely, Logistic Regression assumes a linear relationship between independent variables and event probability, making it less effective when financial ratios exhibit complex relationships.

4.5 Feature Importance Analysis

Random Forest also produces feature importance measures indicating each variable's contribution to financial distress predictions.

Table 5. Feature Importance Ranking

Rank	Variable
1	Net Profit Margin (NPM)
2	Return on Assets (ROA)
3	Interest Coverage Ratio (ICR)
4	Debt to Equity Ratio (DER)
5	Total Asset Turnover (TATO)
6	Total Debt (TH)
7	Current Ratio (CR)

Source: Data processing results using Orange Data Mining.

The findings identify Net Profit Margin (NPM) as the most influential variable in determining financial distress, suggesting that net profit generation is a primary indicator of potential financial difficulties. ROA and ICR also make substantial contributions because they reflect companies' ability to generate profits from assets and meet interest expenses. Current Ratio makes the smallest contribution, indicating that short-term liquidity is not a dominant factor in distinguishing financially distressed technology companies.

4.6 Discussion

The findings indicate that Random Forest outperforms Logistic Regression in predicting financial distress among IDX-listed technology companies. This suggests that ensemble learning algorithms capture more complex relationships among financial ratios than conventional statistical methods.

These findings align with Signaling Theory, which explains that financial information provides signals investors use to assess companies. Ratios reflecting profitability, solvency, activity, and interest payment capacity convey information about financial health. When these indicators deteriorate, the market receives negative signals indicating increased financial distress risk.

The study also identifies Net Profit Margin (NPM) as the dominant predictor of financial distress. This suggests that net profit generation is central to maintaining business continuity in technology companies. Companies with low profit margins are more vulnerable to cash flow pressures, difficulties meeting obligations, and bankruptcy risk.

These findings are consistent with Febrianta et al. (2024), Mulyanto et al. (2025), and Lazuardy Sidarta (2025), who report higher accuracy for Random Forest than other classification methods in financial distress prediction. Although Logistic Regression still performs well, its limitations in modeling nonlinear relationships result in lower accuracy than Random Forest.

Practically, these findings suggest that companies, investors, and financial institutions can use Random Forest as an early warning system for potential financial distress. Higher predictive performance enables companies to implement risk mitigation measures sooner, supporting business continuity and minimizing bankruptcy risk.

V. Conclusions

This study compares Logistic Regression and Random Forest performance in predicting financial distress among technology companies listed on the Indonesia Stock Exchange (IDX) during 2021–2025. Analysis of 205 observations from 41 companies yields the following conclusions.

First, Logistic Regression predicts financial distress with accuracy of 82.9%, precision of 83.7%, recall of 82.9%, and an F1-score of 83.2%. These results indicate good classification performance in distinguishing financially distressed from non-distressed companies.

Second, Random Forest outperforms Logistic Regression, achieving accuracy of 92.7%, precision of 94.5%, recall of 92.7%, and an F1-score of 93.0%. These results suggest that Random Forest more effectively recognizes relationships among financial variables, generating more accurate predictions.

Third, feature importance analysis identifies Net Profit Margin (NPM) as the most influential predictor of financial distress, followed by Return on Assets (ROA) and Interest Coverage Ratio (ICR). These findings suggest that profitability and the ability to meet interest obligations are primary indicators of potential financial difficulties among technology companies.

Overall, this study demonstrates that Random Forest is more effective than Logistic Regression in predicting financial distress among IDX-listed technology companies. Random Forest may therefore be considered a more reliable predictive model for early warning systems designed to identify potential corporate financial difficulties.

The study provides practical implications for corporate management, investors, creditors, and regulators seeking to use machine learning to improve financial risk management decisions. Future research should increase the sample size, expand industry coverage, use longer observation periods, and compare Random Forest with other machine learning algorithms, such as Extreme Gradient Boosting (XGBoost), Support Vector Machine (SVM), Artificial Neural Network (ANN), or Light Gradient Boosting Machine (LightGBM), to develop more comprehensive and accurate financial distress prediction models.

References

- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Brigham, E. F., & Ehrhardt, M. C. (2022). *Financial management: Theory and practice* (16th ed.). Cengage Learning.
- Febrianta, R., Pratama, A., & Nugraha, D. (2024). Comparative analysis of machine learning algorithms for financial distress prediction in Indonesian listed companies. *Jurnal Manajemen dan Teknologi Informasi*, 15(2), 120–134.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). John Wiley & Sons.
- Hutauruk, M. R., Octavianus, R. J. N., & Rosidi. (2021). Financial distress prediction using financial ratios: Evidence from Indonesian listed companies. *Jurnal Keuangan dan Perbankan*, 25(2), 329–343.
- Iskandar, D., Hartono, B., & Wibowo, S. (2021). Machine learning approaches for financial distress prediction: A systematic literature review. *Journal of Information Systems*

- Engineering and Business Intelligence, 7(2), 151–164.
- Kristanti, F. T., & Dhaniswara, H. (2023). Financial distress prediction using logistic regression: Evidence from Indonesian public companies. *Jurnal Akuntansi Multiparadigma*, 14(1), 90–104.
- Lazuardy Sidarta, M. (2025). Comparison of machine learning algorithms in predicting financial distress of manufacturing companies. *International Journal of Business Analytics*, 9(1), 15–29.
- Mulyanto, A., Hidayat, R., & Saputra, D. (2025). Performance comparison of Random Forest and Support Vector Machine in financial distress prediction. *Journal of Applied Business Research*, 18(1), 45–58.
- Napitupulu, R. (2025). Digital transformation and financial performance in Indonesia's technology sector. *Jurnal Ekonomi dan Bisnis Indonesia*, 40(1), 55–70.
- Platt, H. D., & Platt, M. B. (2002). Predicting corporate financial distress: Reflections on choice-based sample bias. *Journal of Economics and Finance*, 26(2), 184–199. <https://doi.org/10.1007/BF02755985>
- Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics*, 87(3), 355–374. <https://doi.org/10.2307/1882010>