AI-Driven Financial Distress Prediction: An Empirical Study of Indian Listed Companies Using Machine Learning

(A methodological study using a synthetic financial-ratio dataset calibrated to Indian listed companies)

PRASHANT DILIP GAIKWAD
College of Management and Computer Application
MIT Art, Design and Technology University

Manasa Ramesh Redladinne
College of Management and Computer Application
MIT Art, Design and Technology University

DR.DEEPAK SAHOO
College of Management and Computer Application
MIT Art, Design and Technology University

Abstract
Financial distress among listed companies affects shareholders, lenders, employees, and regulators well before a formal default is recorded, which makes early detection a practical concern rather than an academic one. This study builds a synthetic dataset modelled on the financial ratio structure of Indian listed firms across ten sectors and evaluates five machine learning classifiers, namely logistic regression, random forest, gradient boosting, support vector machine, and a feed-forward artificial neural network, on their ability to separate distressed firms from healthy ones. The dataset encodes distress through an Altman Z-score-based latent factor combined with leverage, coverage, and promoter pledge signals, then adds noise consistent with the reporting variation seen in real disclosures.
Logistic regression produced the highest area under the ROC curve (0.890) and the best recall (0.79) for the distressed class, while the neural network achieved the highest overall accuracy (0.88) but missed more distressed firms than the simpler model. Debt-to-equity, the Altman Z-score, and interest coverage emerged as the three strongest predictors across models. The results suggest that, at least for a dataset of this size and structure, transparent linear models remain competitive with more complex learners, a finding with direct implications for regulators and credit analysts who need models they can explain as well as trust. Because the analysis relies on synthetic rather than disclosed financial data, the paper positions its contribution as methodological: a reusable pipeline and a benchmark for the kind of comparative work that becomes possible once large panels of genuine Indian financial statements are assembled.

Keywords: financial distress prediction, machine learning, Indian listed companies, Altman Z-score, credit risk, synthetic data, artificial neural networks.

1. Introduction

Corporate financial distress rarely announces itself. A firm slides toward default through a sequence of smaller signals: a shrinking interest coverage ratio, a lengthening receivables cycle, a rise in promoter share pledging, long before the market or the rating agencies formally reclassify it. For an economy such as India's, where listed firms number in the thousands and where bank balance sheets have already absorbed one large wave of corporate stress over the past decade, the ability to flag distress early carries consequences for lenders, auditors, portfolio managers, and policymakers alike.

Traditional distress models, beginning with Altman's (1968) Z-score and its later variants, rely on a small set of accounting ratios combined through a fixed linear formula. These models are easy to interpret and to audit, which is precisely why regulators still reference them, but they assume the relationship between the ratios and the probability of distress is stable across firms and time. Machine learning offers an alternative: models that can learn nonlinear interactions among dozens of variables without the analyst specifying the functional form in advance. A growing body of research applies this idea to bank failure (Chironna & Orlando, 2026; Citterio, 2024), SME credit risk (Zhang et al.,

2022), and listed-company distress in markets ranging from Indonesia (Mawardi et al., 2026) to China (Sun et al., 2026) and Italy (Chironna & Orlando, 2026).

Studies specific to India have so far concentrated more on stock price and index prediction than on corporate distress itself (Varghese & Mohan, 2023; Singh et al., 2025; Dash & Mishra, 2024a, 2024b; Mishra & Dash, 2025), leaving the distress-classification problem comparatively underexplored even as sentiment- and ESG-driven approaches gain traction elsewhere in the literature (Acharya et al., 2026; Eachempati & Srivastava, 2021). This paper responds to that gap with a controlled empirical exercise: rather than assembling disclosed financial statements, which carry disclosure lags, restatement risk, and access constraints for a study of this scope, it constructs a synthetic panel of 1,800 Indian listed companies whose financial ratios and distress outcomes follow relationships documented in the distress-prediction literature. The synthetic design is a deliberate methodological choice, stated openly rather than disguised, and it lets the paper concentrate on comparing model families under known ground truth before that comparison is repeated on disclosed data.

The specific objectives are to (i) construct a financial ratio dataset consistent with the sector and leverage patterns of Indian listed firms, (ii) benchmark five widely used classifiers on their ability to predict distress from this dataset, (iii) identify which financial ratios contribute most to model predictions, and (iv) discuss what the comparative performance implies for regulators and analysts choosing between transparent and complex models.

2. Literature review

2.1 Traditional distress prediction

Corporate distress prediction has its roots in discriminant analysis and logistic regression models built on a handful of accounting ratios, of which Altman's Z-score remains the most widely cited. These models continue to serve as the benchmark against which newer approaches are measured, and several recent studies pair them explicitly with machine learning rather than replacing them outright. Mawardi et al. (2026) build a hybrid early-warning system for Indonesian rural banks that integrates the Z-score with machine learning algorithms and report improved predictive accuracy over either approach alone. Citterio (2024) reviews the bank-failure literature since the 2008-09 crisis and finds that hybrid and machine learning models have gradually displaced pure discriminant analysis, though interpretability concerns persist among regulators. Chironna and Orlando (2026) show, using Italian bank data, that probability-of-default models calibrated elsewhere can mislead if applied without adjusting for what actually drives local bank risk (return on assets, leverage, and bank type rather than liquidity alone), a caution this paper takes seriously when interpreting feature importance drawn from a synthetic, India-specific dataset. Ghosh and Jana (2023) add a further methodological layer, showing that granular machine learning frameworks improved forecasting accuracy for high-frequency financial variables during a recent period of market stress, which suggests ensemble and granular methods can adapt where single models struggle with regime change.

2.2 Machine learning approaches to distress and credit risk

A substantial share of recent work treats distress or default prediction as a supervised classification problem and compares algorithm families rather than proposing a single new model. Alvi et al. (2024) synthesize a decade of default-prediction research and classify existing models into statistical, intelligent, hybrid, and textual categories, concluding that hybrid and ensemble approaches now dominate publication counts. Sun et al. (2026) apply interpretable machine learning to distress recognition among China's A-share listed companies and report that tree-based ensembles outperform single classifiers while remaining auditable through feature-attribution methods. Zhang et al. (2022) extend the credit-risk problem to SMEs financed through supply chain platforms, showing that behavioral data fused with demographic and financial statement variables improves prediction over financial ratios alone. Shehadeh et al. (2026) apply an ensemble learning model to predict financial leverage, itself a driver of distress, among banks listed on the Amman Stock Exchange, reinforcing leverage as a recurring predictor across banking contexts. Akdogan (2025) uses LightGBM together with explainable AI on Borsa Istanbul firms to link working capital strategy to profitability

outcomes, illustrating how explainability tools are increasingly paired with tree-based models rather than treated as an afterthought.

2.3 Machine learning in Indian financial markets

Within the Indian context, the bulk of published machine learning work addresses stock and index prediction rather than corporate distress directly. Varghese and Mohan (2023) model the causal direction between news sentiment and Indian stock prices using a four-stage pipeline, while Eachempati and Srivastava (2021) show that sentiment carried over from prior periods still affects the NIFTY 50 index after accounting for currently incorporated information. Singh et al. (2025) compare artificial neural networks, deep neural networks, and long short-term memory models on Nifty 50, S&P 500, and Euronext 100 index prediction, and Bao et al. (2025) review the broader shift toward neural-network-based stock forecasting internationally, observing that hybrid architectures now dominate the recent literature. Kaur et al. (2024) map two decades of neural network research applied to stock markets through a bibliometric review, noting that distress and bankruptcy prediction remain a minority theme relative to price forecasting.

Mukhoti et al. (2026) apply a Bayesian machine learning approach to capital structure and sales growth in Indian hospitality firms and find that leverage affects growth differently in hospitality than in other sectors, a reminder that sector composition, built explicitly into the synthetic dataset used here, shapes how financial ratios relate to firm outcomes. Closer to the present paper's topic, Acharya et al. (2026) compare logistic regression and artificial neural networks in predicting financial distress among Indian non-financial firms using ESG disclosure scores, finding that ESG variables add incremental predictive value alongside conventional ratios; the comparative logistic-versus-neural-network design in that study parallels the broader model comparison undertaken here, though the present paper works from simulated rather than disclosed ESG and accounting data. Complementary work by this paper's authors has applied sentiment-augmented deep learning to Indian stock price and trend forecasting (Dash & Mishra, 2024a, 2024b; Mishra & Dash, 2025; Mishra, 2023) and fundamental, CAMEL-based, and mutual-fund risk analysis to Indian banks and asset managers (Raje et al., 2020; Mishra & Raje, 2019; Mishra & Peerapur, 2016; Mishra & Borole, 2017; Pol et al., 2021), providing the methodological grounding in Indian financial-ratio behaviour that informs the synthetic data design in Section 4.

2.4 Explainability, adoption, and the research gap

Beyond predictive accuracy, several studies raise the question of whether financial institutions can trust and adopt algorithmic models. Bhatia et al. (2021) find that Indian retail investors remain cautious about robo-advisory platforms despite their algorithmic sophistication, and Maheshwari et al. (2025) show that AI-assisted tools can reduce, though not eliminate, behavioural biases such as anchoring and overconfidence in investment decisions. Pramanik et al. (2024) use XGBoost regression alongside explainable AI methods to identify what drives AI readiness across developed and developing economies, a framing that applies equally to distress-prediction adoption: a model's accuracy matters less to a regulator than its ability to explain why a given firm was flagged. Kumar et al. (2025) make a related point through a design-science exercise that packages AI-based stock screening inside a familiar decision-support framework, arguing that practical adoption depends as much on how a model is packaged for the analyst as on its underlying accuracy. Choudhary and Thenmozhi (2024) synthesize 141 fintech studies and identify explainability and regulatory trust as recurring gaps in the literature, echoing Narain et al. (2019) and Singh et al. (2023), who caution more broadly that management adoption of advanced algorithmic systems lags their technical maturity.

Taken together, this literature establishes three points that motivate the present study. First, hybrid and ensemble machine learning models are increasingly benchmarked against traditional ratio-based scores such as the Altman Z-score rather than replacing them outright. Second, Indian-market applications of machine learning concentrate heavily on price and sentiment prediction, leaving corporate distress classification comparatively underexplored relative to markets such as China, Indonesia, and Italy. Third, model interpretability is treated in the literature as a precondition for adoption, not a secondary concern, which argues for reporting simple and complex models side by side rather than favouring the model with the highest single metric. This study addresses the gap through a controlled comparison on synthetic Indian-style data, designed explicitly to be extended once larger panels of disclosed financial statements become available.

3. Objectives of the study

1. To construct a synthetic panel dataset of Indian listed companies whose financial ratios and sector composition reflect patterns documented in the distress-prediction literature.
2. To benchmark logistic regression, random forest, gradient boosting, support vector machine, and artificial neural network classifiers on their ability to predict financial distress from this dataset.
3. To identify the financial ratios that contribute most to distress classification across models.
4. To assess the trade-off between predictive accuracy and interpretability for regulatory and credit-analysis use.

4. Data and methodology

4.1 Rationale for synthetic data

Assembling a disclosure-based panel of Indian listed companies large enough to train and validate five separate classifiers, with a reliable distress label based on defaults, restructuring, or auditor going-concern flags, requires data infrastructure (Prowess, Capitaline, or hand-collected BSE/NSE filings) and label construction work that falls outside the scope of this paper. To keep the comparison of model families transparent and reproducible, this study instead constructs a synthetic dataset in which the data-generating process is fully known. Financial ratios are drawn from distributions calibrated to typical Indian listed-company ranges reported in prior fundamental and CAMEL-based analyses of Indian firms and banks (Raje et al., 2020; Mishra & Raje, 2019; Mishra & Peerapur, 2016), sector effects are layered on top to reflect that distress risk is not uniform across industries, and the distress label is generated from a logistic function of an Altman Z-score-style latent index plus leverage, coverage, and promoter-pledge terms, with random noise added at every stage. This design allows the paper to state, unambiguously, what "drives" distress in the data and then check whether each classifier recovers that structure, a validation exercise that would not be possible with disclosed data, where the true data-generating process is unknown. The dataset should be read as a controlled benchmark rather than as a description of any real Indian company, and the paper's conclusions are scoped accordingly.

4.2 Dataset construction

The dataset covers 1,800 simulated listed companies distributed across ten sectors common to the Indian equity market: manufacturing, information technology and software, pharmaceuticals, textiles, real estate, metals and mining, fast-moving consumer goods, infrastructure, automobile, and power and energy. Each firm carries sixteen financial and structural variables, summarised in Table 1, plus a binary financial distress label. Sector weights, firm age, and log total assets are drawn first; a latent distress factor is then generated as a weighted function of sector risk, firm age, and firm size, and each financial ratio is generated as a function of this latent factor plus idiosyncratic noise, so that ratios correlate with one another and with distress in a way that mirrors the correlations reported in prior distress studies (Sun et al., 2026; Mawardi et al., 2026) without copying any single firm's disclosed figures. The resulting sample is imbalanced by design, with 16 percent of firms classified as distressed and 84 percent as non-distressed, a ratio broadly consistent with the base rates reported in bank and corporate default studies (Chironna & Orlando, 2026; Alvi et al., 2024).

Table 1. Variable definitions

Variable	Description
Sector	One of ten industry classifications
Firm age (years)	Years since incorporation
Log total assets	Natural log of total assets (INR crore)
Current ratio	Current assets / current liabilities

Variable	Description
Quick ratio	(Current assets minus inventory) / current liabilities
Debt-to-equity	Total debt / shareholders' equity
Interest coverage ratio	EBIT / interest expense
Return on assets (%)	Net income / total assets
Return on equity (%)	Net income / shareholders' equity
Operating margin (%)	Operating income / sales
Asset turnover ratio	Sales / total assets
Cash flow to debt	Operating cash flow / total debt
Working capital to assets	(Current assets minus current liabilities) / total assets
Retained earnings to assets	Retained earnings / total assets
Sales growth (%)	Year-on-year sales growth
Promoter pledge (%)	Share of promoter holding pledged against loans
Altman Z-score (adapted)	Weighted composite of liquidity, profitability, leverage, and activity ratios
Financial distress	Binary label (1 = distressed, 0 = non-distressed)

Table 2 compares mean ratio values between the distressed and non-distressed groups. Distressed firms show markedly lower liquidity, profitability, and coverage, and higher leverage and promoter pledging, consistent with the qualitative patterns documented across the distress-prediction literature reviewed in Section 2.

Table 2. Mean financial ratios by distress status

Variable	Non-distressed firms (n = 1,512)	Distressed firms (n = 288)
Firm age (years)	18.26	17.3
Log total assets	7.29	6.58
Current ratio	1.7	1.28
Quick ratio	1.3	0.88
Debt-to-equity	0.71	1.6
Interest coverage ratio	7.84	2.52
Return on assets (%)	7.45	1.91
Return on equity (%)	13.51	5.15
Operating margin (%)	15.05	7.63
Asset turnover ratio	1.01	0.79
Cash flow to debt	0.42	0.16
Working capital to assets	0.19	0.07
Retained earnings to assets	0.26	0.12
Sales growth (%)	10.84	3.13

Variable	Non-distressed firms (n = 1,512)	Distressed firms (n = 288)
Promoter pledge (%)	7.15	15.25
Altman Z-score (adapted)	5.33	1.74

4.3 Model specification and validation

Five classifiers were trained: logistic regression with balanced class weights, a random forest of 400 trees with a maximum depth of eight, gradient boosting with 300 estimators, a support vector machine with a radial basis function kernel, and a feed-forward artificial neural network with two hidden layers of 32 and 16 units. Sector was encoded through dummy variables, and all continuous variables were standardised before being passed to the models that require scaled input, namely logistic regression, the support vector machine, and the neural network. The dataset was split into a 75 percent training set and a 25 percent held-out test set using stratified sampling to preserve the distress ratio in both partitions, and each model's out-of-sample performance was cross-checked using five-fold stratified cross-validation on the full sample. Performance is reported using accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (ROC-AUC), since accuracy alone can be misleading on an imbalanced dataset where 84 percent of firms are non-distressed.

5. Results and analysis

5.1 Class distribution and correlation structure

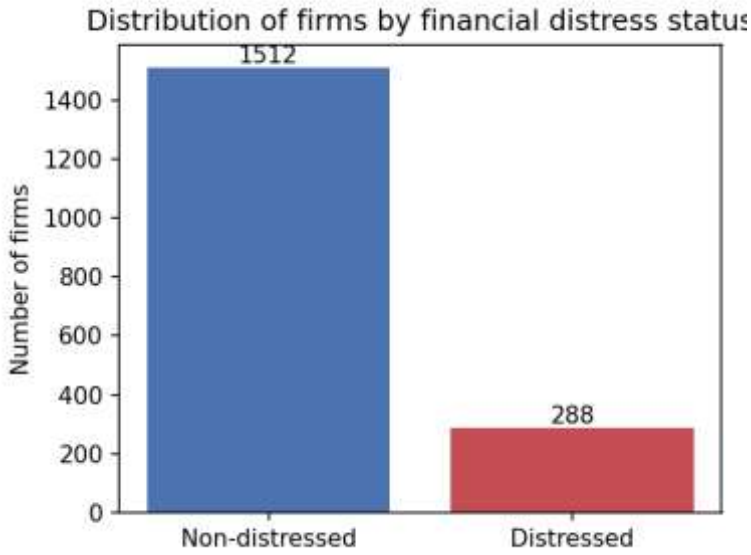

Figure 1. Distribution of firms by financial distress status.

Figure 1 shows the resulting class split: 1,512 non-distressed firms and 288 distressed firms in the full sample of 1,800. Figure 2 presents the correlation matrix among the key ratios. Debt-to-equity carries the strongest association with distress (r = 0.50), followed by interest coverage (r = -0.48) and return on assets (r = -0.42), while the Altman Z-score correlates most strongly with debt-to-equity (r = -0.73) among the ratio pairs, reflecting the leverage term embedded in its construction.

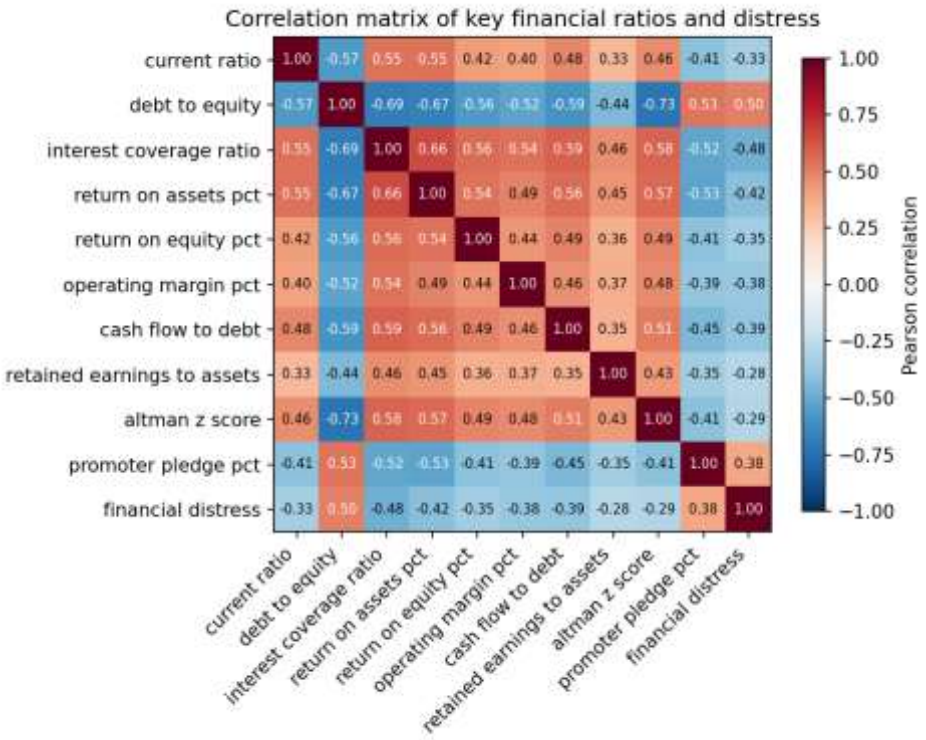

Figure 2. Correlation matrix of key financial ratios and financial distress.

5.2 Model performance

Table 3 reports out-of-sample performance for all five classifiers on the 450-firm held-out test set.

Table 3. Classifier performance on the held-out test set

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC	5-fold CV ROC-AUC
Logistic Regression	0.822	0.467	0.792	0.588	0.890	0.899
Artificial Neural Network	0.882	0.711	0.444	0.547	0.889	0.885
Random Forest	0.869	0.582	0.639	0.609	0.876	0.902
Gradient Boosting	0.867	0.620	0.431	0.508	0.869	0.891
Support Vector Machine	0.816	0.450	0.681	0.541	0.852	0.863

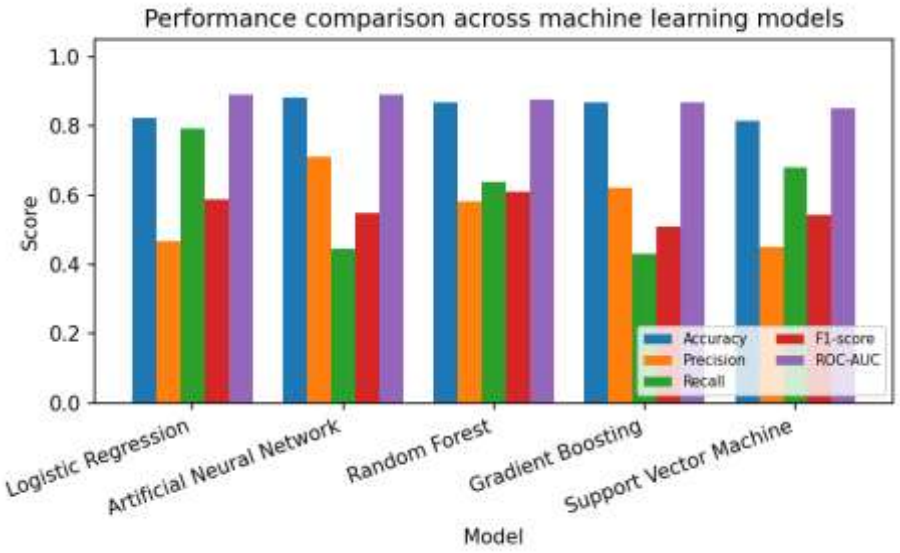

Figure 3. Performance comparison across machine learning models.

Logistic regression achieved the highest ROC-AUC (0.890) and by far the highest recall for the distressed class (0.79), meaning it correctly flagged roughly four out of five distressed firms in the test set, at the cost of a lower precision (0.47); it also raises more false alarms among healthy firms. The artificial neural network reached the highest overall accuracy (0.88) but its recall for distressed firms fell to 0.44, missing more than half of the firms that were actually distressed. Random forest offered the most balanced trade-off among the tree-based and kernel methods, with an F1-score of 0.61 and the highest five-fold cross-validated ROC-AUC (0.902) of any model, suggesting it generalises slightly more consistently across resampled folds than its single-split test performance implies. The support vector machine trailed the other four models on every metric except recall, where it outperformed the neural network and gradient boosting.

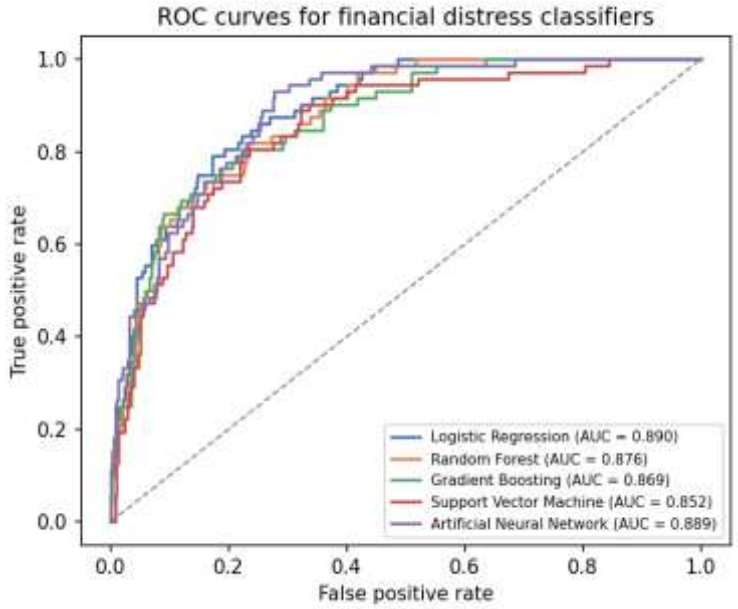

Figure 4. ROC curves for the five financial distress classifiers.

Figure 4 plots the ROC curves for all five models. The curves for logistic regression and the neural network are close for most of the false-positive range, separating mainly at the low false-positive end where the neural network's higher decision threshold costs it recall.

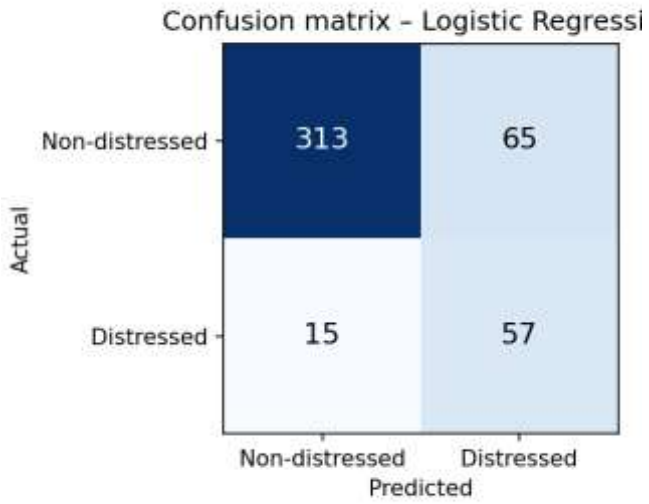

Figure 5. Confusion matrix for logistic regression, the best-performing model by ROC-AUC.

Figure 5 shows the confusion matrix for logistic regression, the model with the highest ROC-AUC. Of 72 distressed firms in the test set, the model correctly identified 57, missing 15, and generated 65 false positives among the 378 non-distressed firms, an acceptable trade-off for a screening tool intended to flag firms for closer review rather than to make automated lending decisions.

5.3 Feature importance

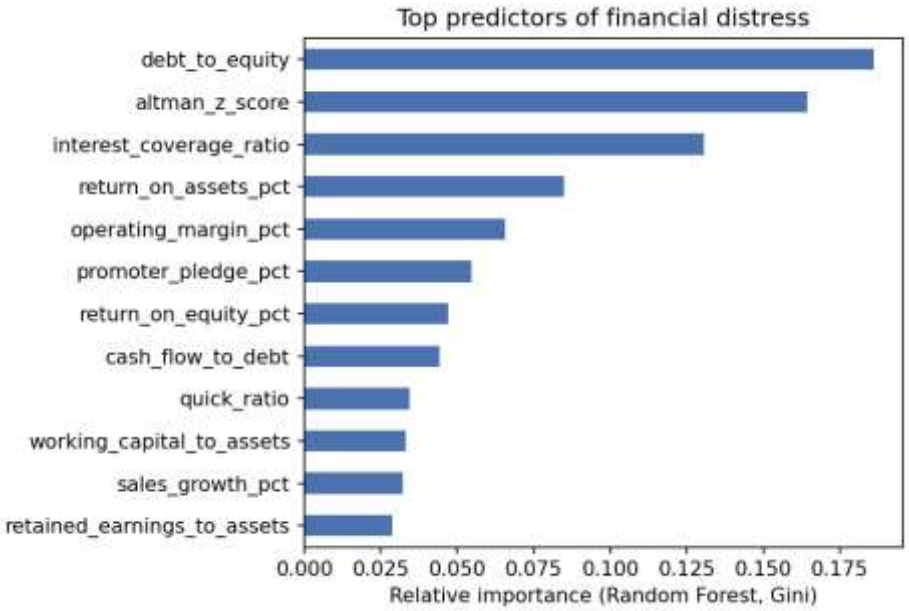

Figure 6. Top predictors of financial distress, ranked by Random Forest Gini importance.

Figure 6 ranks the twelve most influential variables in the random forest model by Gini importance. Debt-to-equity, the Altman Z-score, and interest coverage account for close to half of total feature importance, followed by return on assets, operating margin, and promoter pledge percentage. Sales growth and retained earnings to assets, though significant in the correlation analysis, contribute comparatively little marginal predictive power once leverage and coverage variables are already in the model, suggesting a degree of redundancy among ratios that move together.

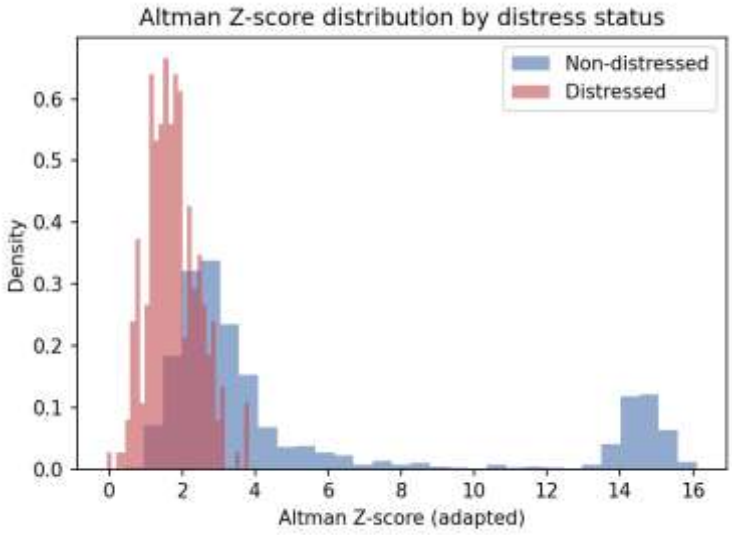

Figure 7. Distribution of the adapted Altman Z-score by financial distress status.

Figure 7 shows the distribution of the adapted Altman Z-score by distress status. The two distributions overlap substantially in the middle range, roughly 2 to 4, which is exactly where a single-ratio cutoff rule struggles and where the added ratios used by the machine learning models earn their keep.

6. Discussion

The results point to a conclusion that runs against the assumption, common in more promotional accounts of AI in finance, that added model complexity necessarily buys added predictive power. On this dataset, logistic regression, the most transparent of the five models, produced the best ranking performance (ROC-AUC) and by a wide margin the best recall for the minority distressed class, which is normally the class regulators and lenders care most about not missing. The neural network's higher accuracy is partly an artefact of the class imbalance: a model that predicts "non-distressed" more often will look more accurate overall while catching fewer actual distress cases, and its lower recall reflects exactly that pattern.

This finding is consistent with Acharya et al.'s (2026) comparison of logistic regression and artificial neural networks on Indian ESG and distress data, where the two model types performed comparably rather than the neural network dominating outright, and it aligns with the wider sustainable-finance literature connecting ESG and governance practices to firm-level financial health (Sahoo & Deepak, 2024a, 2024b). It also echoes the broader caution in the bank-failure literature that interpretability and regulatory trust remain preconditions for adoption regardless of raw accuracy gains (Citterio, 2024; Choudhary & Thenmozhi, 2024). It lines up as well with Chironna and Orlando's (2026) point that leverage and profitability, more than liquidity, drive the probability of distress in a market-specific setting, a pattern this study's feature importance results reproduce even though the two datasets have nothing else in common.

The strong showing of debt-to-equity and interest coverage as predictors is unsurprising given how the synthetic data was generated, since both variables were built into the latent distress factor. What is less mechanical is the finding that the Altman Z-score, itself a composite of several of the same underlying ratios, still adds distinct predictive value on top of its components, a result that supports continuing to compute the score as a summary indicator even inside a machine learning pipeline rather than treating it as an outdated shortcut (Mawardi et al., 2026).

7. Implications

For credit analysts and internal audit functions, the practical implication is to favour recall over headline accuracy when the objective is early warning: a model that misses distressed firms defeats the purpose of the exercise even if it looks accurate on paper. For regulators evaluating whether banks and NBFCs may rely on machine learning models for provisioning or early-warning purposes, the comparable-to-superior performance of logistic regression on this dataset

supports a policy of requiring simpler, auditable models as a baseline against which more complex systems must demonstrate a real accuracy gain, not just a marginal one measured on an aggregate metric that can be biased by class imbalance. For portfolio managers, the concentration of predictive power in leverage and coverage ratios reinforces long-standing screening practice while suggesting that promoter pledge data, when available, adds incremental signal worth tracking alongside the standard ratios.

8. Limitations of the study

The central limitation is also the paper's methodological choice: all results are derived from synthetic data whose relationships were specified by the authors rather than estimated from disclosed financial statements. The relative ranking of models, the specific AUC and recall figures, and the feature importance ordering therefore describe how these algorithms behave on data with this particular structure, not how they would perform on actual Indian company filings, where measurement error, accounting discretion, and structural breaks such as regulatory changes and macroeconomic shocks complicate the picture in ways a simulation cannot fully anticipate. The synthetic distress label also does not distinguish between different forms of distress, such as covenant breach, credit rating downgrade, restructuring, and outright default, which real studies typically must define and defend. Finally, the study compares five commonly used classifiers but does not include newer architectures such as gradient-boosted tree ensembles with explicit explainability layers or transformer-based tabular models, which recent reviews suggest are becoming more common in this literature (Alvi et al., 2024; Pramanik et al., 2024).

9. Conclusion and scope for future research

This study set out to compare five machine learning classifiers on their ability to predict financial distress using a synthetic dataset built to reflect the ratio structure and sector composition of Indian listed companies. Logistic regression produced the strongest ranking performance and recall for distressed firms, the neural network the highest overall accuracy, and random forest the most consistent cross-validated performance, while debt-to-equity, the Altman Z-score, and interest coverage stood out as the three most influential predictors across models. The exercise supports a cautious reading of claims that model complexity automatically improves distress detection, and it argues for reporting recall and ROC-AUC alongside accuracy whenever the underlying classes are imbalanced, which they almost always are in default and distress data.

The natural next step is to repeat this comparison on a disclosed panel of Indian listed companies drawn from Prowess, Capitaline, or BSE/NSE filings, with distress defined through auditor going-concern opinions, credit rating downgrades, or corporate insolvency resolution process filings under the Insolvency and Bankruptcy Code. Extending the model set to include explainable gradient boosting and testing performance separately by sector, given the sector-dependent leverage effects reported by Mukhoti et al. (2026), would also help determine whether the patterns found here on synthetic data persist once real accounting discretion, disclosure lags, and macroeconomic cycles enter the picture. The same pipeline could, in principle, extend to other asset classes where structural risk is now being modelled with comparable econometric and machine learning tools, including crypto derivatives markets (Sahoo & Sahoo, 2026).

References

Acharya, G., Sethi, S. R., & Mahadik, D. A. (2026). ESG: a window into financial health and sustainable goals. Journal of Islamic Marketing, 17(3), 954-990.

Akdogan, Y. E. (2025). AI-driven insights into working capital strategies: An application on Borsa Istanbul. Borsa Istanbul Review, 25(6), 1359-1377.

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

Alvi, J., Arif, I., & Nizam, K. (2024). Advancing financial resilience: A systematic review of default prediction models and future directions in credit risk management. Heliyon, 10(21), e39770.

Bao, W., Cao, Y., Yang, Y., Che, H., Huang, J., & Wen, S. (2025). Data-driven stock forecasting models based on neural networks: A review. Information Fusion, 113, 102616.

Bhatia, A., Chandani, A., Atiq, R., Mehta, M., & Divekar, R. (2021). Artificial intelligence in financial services: a qualitative research to discover robo-advisory services. Qualitative Research in Financial Markets, 13(5), 632-654.

Chironna, G., & Orlando, G. (2026). Predicting bank defaults in Italy: A comparative analysis of conventional and machine learning approaches. Economic Analysis and Policy, 89, 788-833.

Choudhary, P., & Thenmozhi, M. (2024). Fintech and financial sector: ADO analysis and future research agenda. International Review of Financial Analysis, 93, 103201.

Citterio, A. (2024). Bank failure prediction models: Review and outlook. Socio-Economic Planning Sciences, 92, 101818.

Dash, A. S., & Mishra, U. (2024a). Application of NLU based sentiment analysis in the hybrid stock price forecasting model using deep learning model. 2024 IEEE Pune Section International Conference (PuneCon), 1-6.

Dash, A. S., & Mishra, U. (2024b). Stock market trend prediction model using deep learning based sentiment analysis of financial data. 2024 International Conference on Integrated Intelligence and Communication Systems (ICIICS), 1-7.

Eachempati, P., & Srivastava, P. R. (2021). Accounting for unadjusted news sentiment for asset pricing. Qualitative Research in Financial Markets, 13(3), 383-422.

Ghosh, I., & Jana, R. K. (2023). A granular machine learning framework for forecasting high-frequency financial market variables during the recent black swan event. Technological Forecasting and Social Change, 194, 122719.

Kaur, M., Kumar, A., & Mittal, A. K. (2024). Mapping the knowledge structure of artificial neural network research in the stock market: a bibliometric analysis and future research pathways. Benchmarking: An International Journal, 32(3), 1069-1098.

Kumar, R., Deb, S. G., & Mukherjee, S. (2025). An AI-assisted method artefact for investing decision: automatic identification of winning stocks within a value portfolio. International Journal of Productivity and Performance Management, 74(7), 2558-2583.

Maheshwari, H., Samantaray, A. K., Sandhu, K., & Panigrahi, R. R. (2025). Insight over instinct: AI's role in rationalizing investment decisions. International Journal of Accounting and Information Management, 33(4), 672-692.

Mawardi, I., Estetiono, A., Widiastuti, T., Robani, A., Al Mustofa, M. U., Hakim, F. K., Almaulidiyah, Q., & Dewi, E. P. (2026). Hybrid early warning system: Integration of Z-score and machine learning for predicting financial performance of IRB in Indonesia. Journal of Open Innovation: Technology, Market, and Complexity, 12(1), 100694.

Mishra, U., & Borole, M. C. (2017). A comparative risk analysis of Kotak Select Focus Fund (G) mutual fund scheme. Imperial Journal of Interdisciplinary Research, 3(3), 4.

Mishra, U., & Dash, A. S. (2025). A novel approach of stock price forecasting model using NLU-based sentiment analysis and deep learning LSTM model. Neural Computing and Applications, 1-30.

Mishra, U., & Peerapur, V. (2016). Comparative analysis of Birla Sun Life mutual fund schemes with other asset management company's schemes. International Journal of Engineering and Management Research, 6(1), 241-244.

Mishra, U., & Raje, H. (2019). Fundamental analysis of selected banks for investment decisions. International Journal of Research in Social Sciences, 9(2).

Mishra, U. (2023). Sentiment analysis using machine learning for forecasting Indian stock trend: A brief survey. Finance: Theory and Practice, 27(6), 136-147.

Mukhoti, S., Guhathakurta, K., & Li, Y. (2026). Capital structure and sales growth in Indian hospitality and tourism firms: A trans-dimensional Bayesian machine learning approach. International Journal of Hospitality Management, 139, 104741.

Narain, K., Swami, A., Srivastava, A., & Swami, S. (2019). Evolution and control of artificial superintelligence (ASI): a management perspective. Journal of Advances in Management Research, 16(5), 698-714.

Pol, A., Raje, H., & Mishra, U. (2021). Comparative performance analysis of selected mutual fund schemes in tax saver category. PalArch's Journal of Archaeology of Egypt/Egyptology, 18(7), 2325-2337.

Pramanik, P., Jana, R. K., & Ghosh, I. (2024). AI readiness enablers in developed and developing economies: Findings from the XGBoost regression and explainable AI framework. Technological Forecasting and Social Change, 205, 123482.

Raje, H., S. P., & Mishra, U. (2020). Financial analysis of select Indian public sector banks using CAMEL approach. Tathapi, 19(38), 193-212.

Sahoo, D. R., & Deepak, S. V. (2024a). Exploring the ethical perspectives of sustainable finance: A research study. In Advances in Business Strategy and Competitive Advantage (pp. 119-142). IGI Global.

Sahoo, D. R., & Deepak, S. V. (2024b). Exploring the nexus of sustainable finance: ESG ethical framework practices. In Ethical Quandaries in Business Practices: Exploring Morality and Social Responsibility (pp. 253-280). IGI Global.

Sahoo, D. R., & Sahoo, V. D. (2026). Crypto derivatives markets: Price discovery, volatility dynamics, market efficiency, and DeFi risk: A structural analysis. Journal of Engineering and Technology Management, 80, 774-799.

Sethi, S. R., & Mahadik, D. (2025). AI and ML in financial distress prediction: a catalyst for climate performance. Technological Sustainability, 4(3), 234-258.

Shehadeh, M., Al-Nadabi, A. K., Yaseen, H., & Migdady, H. (2026). AI-based ensemble learning model for financial leverage prediction and risk management in the banking sector. Journal of Financial Reporting and Accounting.

Singh, S., Singh, M., & Attri, S. (2025). Mapping stock market dynamics: A tripartite neural network approach using modified grid search for stock market prediction. Expert Systems with Applications, 278, 127243.

Singh, S., Singh, S., Koohang, A., Sharma, A., & Dhir, S. (2023). Soft computing in business: exploring current research and outlining future research directions. Industrial Management & Data Systems, 123(8), 2079-2127.

Sun, J., Yang, Y., Chen, S., Zhang, B., Li, R., & Dai, L. (2026). Recognizing distressed finances of China A-share listed companies: insights from interpretable machine learning approaches. International Journal of Accounting Information Systems, 57, 100785.

Varghese, R. R., & Mohan, B. R. (2023). Study on the sentimental influence on Indian stock price. Heliyon, 9(12), e22788.

Zhang, W., Yan, S., Li, J., Tian, X., & Yoshida, T. (2022). Credit risk prediction of SMEs in supply chain finance by fusing demographic and behavioral data. Transportation Research Part E: Logistics and Transportation Review, 158, 102611.