Unveiling the Anatomy of Financial Distress Prediction in Iranian SMEs: A Hybrid Framework Coupling PLS-SEM with Explainable Machine Learning

Iliya Ardeshir Larijani

Department of Finance, Faculty of Management and Economics, Islamic Azad University, Science and Research Branch, Tehran, Iran

Mohammad Molani Aghdam

Innovation and Management Research Center, Am. C., Islamic Azad University, Amol, Iran

Miroslav Čavlin

Faculty of Economics and Engineering Management in Novi Sad, University Business Academy, Cvećarska 2, 21000 Novi Sad, Republic of Serbia

Abstract

This study proposes an empirical framework for predicting financial distress in manufacturing small and medium-sized enterprises (SMEs) by combining Partial Least Squares Structural Equation Modeling (PLS-SEM) with supervised machine learning techniques. Based on a sample of 138 manufacturing SMEs in Iran, PLS-SEM was initially applied to validate the measurement structure and evaluate the relationships between multi-dimensional risk indicators and perceived financial distress. Five key predictors, comprising Debt Ratio, Market Risk, Technology Risk, Resource Risk, and Project Risk, were then supplied to three algorithms: Artificial Neural Networks (ANN), Random Forest (RF), and Gradient Boosting (GB). The predictive workflows were optimized and evaluated through a nested cross-validation design using Optuna. Model evaluation across standard regression metrics showed that ANN and RF demonstrated closely matched predictive accuracy, with ANN yielding the lowest root mean squared error (0.561) and RF showing strong performance in mean absolute error (0.446) and coefficient of determination (0.739). To address the black-box nature of the models, post-hoc interpretability was conducted using permutation feature importance and SHAP, which indicated that Debt Ratio and Market Risk provided the largest predictive contribution to distress variance. The resulting framework delivers a transparent forecasting approach that can assist enterprise managers, lending institutions, and industrial policymakers in recognizing distress signals and establishing timely risk mitigation measures.

Uncorrected proof

Keywords: Perceived financial distress, Small and medium-sized enterprises, PLS-SEM, Machine learning, SHAP

1- **Introduction**

Manufacturing small and medium enterprises (SMEs) face distinct operating and financial challenges compared to larger corporations. In emerging and volatile economies such as Iran, these businesses operate under severe structural pressures; empirical research indicates that inadequate access to institutional credit, compounding inflation, and adverse macroeconomic conditions represent the primary catalysts of business distress and failure (Arasti et al., 2014). Recent increases in business bankruptcies and other financial problems harm the world economy and affect many stakeholders, such as the government, investors, managers, customers, and participants in the same business. (Sethi & Mahadik,2026).

Given that manufacturing SMEs usually have low levels of internal financial resilience, any operational and financial risk could result in bankruptcy. Hence, developing an empirical framework for predicting incipient perceived financial distress in Iranian manufacturing SMEs is of paramount practical and theoretical importance.

Financial distress in companies can be assessed using various indicators and ratios, but individual measures are insufficient to capture overall financial risk and distress severity in isolation (Colak, 2021). Recent research emphasizes the importance of integrating financial, non-financial, and macroeconomic indicators for robust corporate risk analysis and bankruptcy prediction. Even though traditional financial ratios remain influential, management quality, governance, and economic factors now complement standard models for more nuanced predictions (Sharma et al., 2024). Non-financial risks are increasingly recognized as crucial in risk management systems and as a result, advanced nonlinear methods have been proposed to analyze hidden relationships among the financial and nonfinancial risk factors (Pertseva et al., 2021). Thus, combining quantitative financial indicators with managerial variables, industry factors and corporate nonfinancial standards in this study provides a suitable risk prediction model.

The primary challenge for scholars in predicting financial distress lies in selecting appropriate predictor variables (qualitative or quantitative) and determining the most effective modeling approach. Subsequently, the technological revolution has accelerated the adoption of artificial intelligence (AI) across corporate finance disciplines. Recent studies have explored the application of artificial intelligence (AI) in forecasting financial distress and corporate failure.

uncorrected proof

Artificial Neural Networks (ANN) demonstrated high performance in predicting financial distress in manufacturing industries (Dube et al., 2023). Random Forest outperformed several conventional machine learning algorithms in predicting financial distress among Chinese manufacturing firms (Zhong & Wang, 2022). Ensemble learning methods, particularly Random Forest and XGBoost, achieved superior predictive performance in corporate financial distress prediction (Kristanti et al., 2024). These findings suggest that ensemble and neural networks learning algorithms are particularly effective in modeling the complex and nonlinear relationships that characterize financial distress prediction.

Despite these advances, existing research presents two critical gaps. First, each dimension of strategic risk including, market risk, resource constraints, technology risk, and project implementation failure has been examined in isolation. Yet, no study has combined these four constructs as a unified set for predicting financial distress.

Second, although recent bankruptcy and distress studies have increasingly adopted SHAP to enhance model interpretability (Fasano et al., 2025; Papíková & Papík, 2026), its integration within a structural equation modeling (SEM) framework to capture both linear structural relationships and nonlinear feature importance across latent and observable predictors has not been explored. To bridge these gaps, this study develops a hybrid PLS-SEM and machine-learning framework combined with SHAP interpretability to evaluate and predict perceived financial distress in Iranian manufacturing SMEs.

The present study develops a predictive framework for perceived financial distress in Iranian manufacturing SMEs by integrating Structural Equation Modeling (SEM) with three complementary machine learning algorithms, namely Artificial Neural Networks (ANN), Random Forest (RF), and Gradient Boosting (GB). First, SEM is employed to validate the proposed conceptual framework and examine the relationships among financial and nonfinancial risk factors. Subsequently, the predictive performance of the three machine learning models is evaluated using nested cross-validation. Model selection is primarily based on RMSE, while MAE and R^2 are reported as complementary performance measures. These findings contribute to the growing literature on SME financial distress prediction, providing practical implications for managers, investors, financial institutions, and policymakers in designing more effective predictive frameworks.

2. Literature Review

2.1 Conceptual Definition of Financial Distress

uncorrected proof

Financial distress is fundamentally distinct from legal bankruptcy and outright business failure, although these conditions are closely interrelated. Prior literature broadly defines this state as a company's inability to service its debt obligations as they come due (Yazdanfar & Öhman, 2020). Rather than an instantaneous collapse, distress unfolds as an evolving process that can ultimately lead to either operational turnaround or formal liquidation. Companies experiencing this condition often exhibit noticeable indicators such as reduced dividend payouts, sustained negative earnings, and compromised interest coverage capacity (Geoffrey, 2020). While both internal operating vulnerabilities and external market pressures have received substantial empirical scrutiny, researchers have reported mixed findings regarding how qualitative operational factors interact with quantitative financial indicators (Crovini et al., 2019). Consequently, evaluating these diverse risk dimensions within an integrated predictive structure remains essential for sound managerial choices and corporate financial oversight. These mixed findings indicate that the relative importance of financial and operational indicators may vary across empirical contexts. Accordingly, examining these dimensions jointly may provide a more context-sensitive assessment of perceived financial distress among SMEs.

uncorrected proof

2.2 Conceptual Definition of Risk

Risk is traditionally defined as exposure to adversity or negative outcomes (Moosa, 2007). However, recent research suggests a broader conceptualization. Aven and Renn (2009) defined risk as uncertainty about and severity of consequences regarding something humans value, encompassing both positive and negative outcomes. This aligns with the view of Kenett and Tapiero (2009) that quality and risk management converge, with both addressing consequential effects. As argued, risk and quality are measured as consequences of events with specific statistical characteristics. Furthermore, distinguishing between risk factors and their consequences implies the probability of poor outcomes rather than their immediate manifestation. This distinction is crucial for effective assessment and intervention practices. Recently, a notable shift has occurred towards understanding risk as a multifaceted concept including both potential threats and opportunities.

Recent studies highlight the importance of integrating both financial and nonfinancial risk factors in predicting financial distress among SMEs. Combining financial variables with nonfinancial governance factors improved prediction accuracy in Egyptian SMEs (Ragab &

Saleh, 2021) and Malaysian SMEs (Abdullah et al., 2014). Environmental conditions and strategic variables, such as technology level and quality certification, were associated with financial distress in Spanish manufacturing SMEs. For Indonesian SMEs, profitability measures like operating income to total assets and retained earnings to total assets served as significant predictors of financial distress (Kristanti et al., 2019). These studies emphasize the value of incorporating diverse factors, including governance structures, industry characteristics, and operational efficiency, alongside traditional financial metrics when designing predictive frameworks for SME financial distress. This integrated approach provides more accurate predictions and valuable insights for stakeholders, including financial institutions and SME managers.

Beyond these empirical foundations, the specific risk constructs underlying the proposed framework can be examined, beginning with technological uncertainty. Perceived technological uncertainty refers to the difficulty in understanding or predicting certain aspects of technology that affect decisions during new product development (Song & Montoya Weiss, 2001). For instance, an NPD project manager may perceive a product technology to be well understood, highly developed and thus straightforward in application. Alternatively, the product technology may be perceived as undeveloped and unknown and hence requiring trial and error research. In addition, a project manager may perceive the rate and direction of change in the product technology to be more or less predictable.

While technological uncertainty concerns the firm's internal innovation processes, exposure to external market conditions constitutes an equally significant source of vulnerability. Market and financial risks are regarded as significant factors affecting firm performance and sustainability. Market risk, measured by conditional value at risk, negatively impacts firm performance (Vo, 2023). Corporate risk, particularly unsystematic risk, positively influences financial distress, whereas systematic risk (market risk) exhibits an insignificant positive effect (Gichaiya et al., 2019). The inclusion of market indicators alongside financial variables improved the explanatory power of hazard models predicting financial distress (Deng & Wang, 2008). However, in some markets, financial distress risk and book-to-market equity effects were statistically insignificant in explaining stock returns of distressed firms due to market inefficiency, while size effect remained essential (Idrees & Qayyum, 2018). However, the evidence is not unequivocal: while some studies report that market-based indicators improve the explanation or prediction of financial distress, others find statistically insignificant effects. This inconsistency may reflect differences in market efficiency and in

uncorrected proof

the operationalization of market risk, particularly the distinction between objective market indicators and managerial perceptions. The present study adopts a perceptual operationalization, which is appropriate for SMEs where detailed market-exposure data are often unavailable.

Research indicates that market risk is a significant concern for small and medium enterprises (SMEs), with perceived market risk averaging over 50% in some regions (Belas et al., 2015). This risk manifests as uncertainty in demand and sales volume adequacy (Dvorský et al., 2020). External factors, such as high competition and buyer bargaining power, were associated with increased financial distress, particularly in low technology SMEs (Madrid Guijarro et al., 2011). Supply chain dependence also impacts credit risk, with both customer and supplier dependence being positively related to credit risk and negatively to credit ratings for SMEs (Liu & Park, 2020; Fang et al., 2022). Internationalization presents opportunities for SME growth, although it is often accompanied by significant operational challenges. The perception of market risk when expanding internationally varies among SMEs (Virglerová et al., 2021). Managing these risks is crucial for SME success, with operational risk management focusing on resource utilization, product quality, and supplier diversification (Dvorský et al., 2020). Understanding and addressing these risk factors is essential for developing effective SME support strategies and policies.

In addition to these external market pressures, the availability and efficient deployment of internal resources represent a further critical dimension of SME financial health. Studies have demonstrated that financial distress in SMEs is influenced by various internal resource-based factors, including innovative and competitive abilities (Civelek et al., 2023). However, while innovativeness may reduce bankruptcy risk, it does not necessarily improve financial performance (Civelek et al., 2023). Entrepreneurial orientation, encompassing innovativeness, proactiveness, and risk taking, positively affected SMEs' financial performance in Iran (Jalali et al., 2013). Financial distress risk can lead to lower financial reporting quality as companies attempt to mislead stakeholders regarding their actual performance (Tarighi et al., 2022). Likewise, related party transactions, particularly loans, were associated with profit manipulation and reduced financial reporting quality in Iranian firms (Tarighi et al., 2022). These findings underscore the complex interplay of internal and external factors impacting SMEs' financial health. Financial distress in Iranian SMEs is often attributed to structural constraints, including outdated manufacturing equipment and limited access to modern machinery.

uncorrected proof

Beyond the existing resource base, the manner in which discrete projects and initiatives are managed introduces an additional layer of operational uncertainty. Among the four main phases of project risk management, the phases of risk analysis and risk monitoring and control are the most critical for maximizing benefits and minimizing costs in project implementation. However, these two phases have received the least attention in small and medium enterprises. Accordingly, emphasis on accurate project planning and scheduling as part of risk analysis and continuous project monitoring and control as part of the fourth phase becomes essential (Lima & Verbano, 2021).

A study on lean manufacturing implementation in Iranian SMEs identified financial capability as a major barrier, reflecting limited resources for modernizing equipment (Moradlou & Perera, 2017). Research on financial distress prediction in Iranian companies highlighted the importance of updating resources to avoid crises (Moradi et al., 2013). Additionally, workforce agility was identified as crucial for promoting external manufacturing flexibility (Alavi, 2016), which is essential in dynamic environments. This indicates that the inability to adapt to technological changes contributes to financial distress. These studies demonstrate that inadequate equipment and manufacturing processes significantly impact the financial health of Iranian SMEs.

uncorrected proof

2.3 **Methods of Distress Prediction**

Over time, researchers have proposed various methods to address the shortcomings of traditional financial distress prediction (FDP) models. One of the early benchmarks was the Z-score developed by Altman (1968), which applied Multivariate Discriminant Analysis (MDA). By utilizing multiple financial ratios as predictors and achieving high classification accuracy, the model dominated the field for decades. While Kidane (2004) questioned the effectiveness of its classification, subsequent studies (Rado, 2013; Ngwenya, 2018; Sfakianakis, 2021) confirmed its continued practical utility.

However, the application of MDA models is constrained by strict statistical assumptions that often prove problematic in real-world financial applications (Premachandra et al., 2009). First, predictor variables must be linearly separable and follow normal distributions. Second, the variance-covariance matrices must be equal across both solvent and bankrupt firms. Third, the dataset must be free of outliers. Empirical evidence indicates that these requirements are frequently violated in practice, as financial data typically exhibit non-linear

patterns. Recognizing these limitations, financial researchers turned to Logistic Regression Analysis (LRA) for distress prediction. Multiple studies, including Bešlić Obradović et al. (2018) and Durica et al. (2019), demonstrated LRA's superior predictive performance in this domain. The primary advantage of the logit model lies in its ability to handle non-linear relationships without requiring strict distributional assumptions (Breen et al., 2018).

Nonetheless, both LRA and MDA models remain susceptible to multicollinearity and outlier sensitivity, which can distort coefficient estimates and inflate distress probabilities (Alimi et al., 2023). Recent studies demonstrate that machine learning and penalized regression methods offer greater robustness against these issues (Lokanan & Ramzan, 2024). Although LRA has been widely utilized in FDP, recent advances in artificial intelligence have accelerated the adoption of machine learning techniques due to their superior capability in capturing complex, non-linear interactions among financial and non-financial variables.

Accordingly, this study evaluates three distinct machine learning algorithms: Artificial Neural Networks (ANN), Random Forest (RF), and Gradient Boosting (GB). ANN represents a non-linear learning architecture capable of detecting complex dependencies among variables through adaptive weight optimization. Random Forest is a bagging-based ensemble algorithm that constructs multiple decision trees to aggregate predictions, thereby enhancing model stability and mitigating overfitting. Gradient Boosting is a boosting-based ensemble technique that sequentially trains decision trees, where each subsequent estimator minimizes the residual errors of preceding models.

uncorrected proof

2.4 **Research Hypotheses**

H1: Higher levels of market risk are positively associated with perceived financial distress in SMEs.

H2: Higher levels of resource risk are positively associated with perceived financial distress in SMEs.

H3: Higher levels of technology risk are positively associated with perceived financial distress in SMEs.

H4: Higher levels of project risk are positively associated with perceived financial distress in SMEs.

H5: Higher leverage (Debt Ratio) is positively associated with perceived financial distress in SMEs.

3. Research Methodology

3.1 Data Collection

A mixed dataset was used combining primary and secondary data from small and medium-sized manufacturing enterprises (SMEs) in Iran. The data were collected from 138 small and medium-sized enterprises by the financial managers of the companies, which were of two types: qualitative data collected through a Likert questionnaire and numerical data for financial risk collected through the debt ratio. The questionnaire was made up of a 28-point Likert-like scale with 1 (no agreement) to 5 (high agreement). Data were collected from a diverse range of Iranian manufacturing sectors to ensure model robustness.

A two-stage purposive-random sampling technique was employed. First, active manufacturing SMEs (<250 employees) were screened based on data availability and the participation of their financial managers as key informants. Next, simple random sampling was conducted among eligible enterprises to select the final sample of 138 firms, consistent with Bartlett et al.'s (2001) sample size framework

In this study, Small and Medium Enterprises (SMEs) were defined based on employee count following the official criteria of the Iran Small Industries and Industrial Parks Organization (ISIPO), covering manufacturing firms with fewer than 250 employees (small: < 50; medium: 50–249). Between March and May 2024, 200 questionnaires were distributed to financial managers of eligible manufacturing SMEs. A total of 155 responses were received; 17 were discarded due to missing answers or lack of corresponding debt ratios. The final sample comprises 138 usable responses, giving an effective response rate of 69%.

Although n = 138 may appear modest by big-data standards, the observation-to-predictor ratio (27.6:1) exceeds the (10:1)benchmark commonly recommended for stable supervised learning. Combined with nested 5×4-fold cross-validation, Optuna-based tuning restricted to inner folds, and hyperparameter search spaces that bounded model capacity, this design mitigated overfitting and supported reliable estimation.

3.2 Measurement Instrument and Construct Operationalization

Uncorrected proof

The survey items were either adopted directly from existing validated literature or adapted and modified to fit the local economic and operational context of Iranian manufacturing SMEs. To ensure content and face validity, the finalized questionnaire was reviewed and confirmed by a panel of academic and industry experts.

- Technology Risk was operationalized based on perceived technological uncertainty, adapted from Song and Montoya-Weiss (2001), capturing dimensions of technological novelty, trial-and-error dependency, and rapid technological shifts.
- Market Risk was adapted from (Belas et al. 2015) and Duc Hong Vo (2023), evaluating market volatility, demand unpredictability, and customer/supplier dependency.
- Resource Risk items were grounded in (Civelek et al. 2023) and (Moradlou & Perera, 2017), measuring constraints regarding specialized human capital, financial liquidity, and physical asset acquisition.
- Project Risk was adapted from (Lima and Verbano 2021), addressing project planning inefficiencies, execution bottlenecks, and milestone delays.
- The perceived financial distress scale captures managerial assessment across four critical operational distress dimensions: consecutive revenue contraction (Item 25), sustained operational losses (Item 26), debt-servicing strain (Item 27), and working capital lockup in receivables (Item 28), adapted from corporate distress frameworks. The perceptual approach to measuring financial distress was conceptually drawn from the questionnaire-based dependent-variable operationalization in Sudaryo et al(2025) .
- Debt Ratio was gathered objectively as a single manifest financial indicator (Total Debt / Total Assets) derived directly from the audited financial records of the surveyed firms

uncorrected proof

In the context of unlisted Iranian manufacturing SMEs, assessing financial distress through managerial perceptions is necessary due to the lack of publicly accessible bankruptcy records and formal default filings.

3.3 Data Preprocessing and Reproducibility Settings

Following the initial data screening protocol where 17 incomplete questionnaires were removed via listwise deletion, yielding a final analytical dataset of (N = 138), no missing values remained across any predictor items or financial metrics, obviating the need for

statistical imputation. To prevent data leakage, all predictor features were standardized via fold-specific z-score normalization ($\mu = 0$, $\sigma = 1$) within a stratified cross-validation scheme, with scaling parameters estimated solely on each fold's training split and applied to the corresponding validation split. All stochastic operations (nested CV splitting, Optuna optimization, and weight initializations) were governed by a fixed seed (random_state=42) to ensure computational reproducibility.

3.4 Structural Equation Modeling (SEM)

Structural equation modeling (PLS-SEM) was conducted using SmartPLS 3. The measurement model was evaluated using indicator reliability, composite reliability (CR), average variance extracted (AVE), the Fornell–Larcker criterion, HTMT, and cross-loadings. The structural model was assessed using path coefficients, R^2, SRMR, and NFI. The statistical significance of the estimated paths was evaluated using a nonparametric bootstrapping procedure with 5,000 BCa bootstrap resamples and a two-tailed significance level of 0.05.

PLS-SEM was selected over covariance-based structural equation modeling (CB-SEM) for methodological and substantive reasons. First, the objective of this study was primarily prediction and explanation of perceived financial distress rather than strict confirmation of a fully established covariance structure. Second, the model included several perceptual constructs measured through multiple indicators and is examined in the relatively small sample of 138 SMEs. PLS-SEM was considered appropriate in this context because it is a composite-based estimation approach that is less dependent on multivariate normality and generally imposes fewer restrictions on sample size than CB-SEM.

3.5 Machine Learning Models

Artificial Neural Networks (ANN), Random Forest (RF), and Gradient Boosting (GB) employ different learning mechanisms for perceived financial distress prediction. ANN learns complex nonlinear relationships through interconnected neurons and RF improves prediction stability by aggregating multiple decision trees, whereas GB sequentially enhances model

uncorrected proof

performance by correcting the errors of preceding trees. These methodological differences allow the models to capture different characteristics of financial and nonfinancial risk data, providing a comprehensive basis for predictive modeling and comparison.

3.6 Hyperparameter Tuning

Model development involved identifying the optimal hyperparameter configuration for each machine learning algorithm. Hyperparameter optimization was performed using the Tree-structured Parzen Estimator (TPE) algorithm implemented in Optuna. Predefined search spaces were specified for each algorithm, and the optimization procedure efficiently explored candidate hyperparameter combinations to identify those minimizing the inner cross-validation RMSE.

Table1: ANN Hyperparameter

Hyperparameter	**Search space or setting**
Number of hidden layers	1, 2, or 3
Neurons per hidden layer	4, 6, 8, 12, 16, 24, or 32
Hidden activation	ReLU or tanh
Optimizer	Adam, AdamW, or RMSprop
Learning rate	0.0001 to 0.005; log scale
Batch size	8, 16, or 32
Dropout	0.00 to 0.30; step 0.05
L2 weight decay	1e-6 to 1e-2; log scale
Early-stopping patience	20, 30, or 40 epochs
Minimum validation improvement	0.0001, 0.0005, or 0.001
Maximum epochs	1000
Initialization	He/Kaiming for ReLU; Xavier/Glorot for tanh
Output activation	Identity/linear
Loss function	Mean squared error
Hyperparameter selection criterion	Minimum mean inner-fold RMSE

Uncorrected proof

Table 2: Random Forest Hyperparameter

Parameter	Setting / Search Space
Algorithm	Random Forest Regressor
Number of trees	200–1000 (step = 50)
Split criterion	squared_error, friedman_mse, absolute_error
Maximum depth	None, 3, 4, 5, 6, 8, 10, 12, 15, 20
Minimum samples split	2–20
Minimum samples leaf	1–10
Maximum features per split	sqrt, log2, 0.4, 0.6, 0.8, 1.0
Bootstrap sampling	True / False
Maximum samples per tree	0.6–1.0 (when bootstrap = True)
Maximum leaf nodes	None, 5, 10, 20, 40, 80
Minimum impurity decrease	0.00–0.02
Cost-complexity pruning (ccp_alpha)	0.00–0.02
Hyperparameter optimization	Optuna (Tree-structured Parzen Estimator)
Selection criterion	Minimum inner-fold RMSE
Feature scaling	Not required
Random seed	Fixed seed for reproducibility
Tree initialization	Bootstrap sampling and random feature selection
Stopping criterion	Determined by the optimized number of trees and tree-growth constraints
Regularization	Tree depth, minimum node size, feature sampling, row sampling, impurity threshold, and pruning

uncorrected proof

Table 3: Gradient boosting hyperparameter

Hyperparameter	Search Space
Number of boosting stages (`n_estimators`)	100–1000 (step = 25)
Learning rate	0.005–0.20 (log scale)
Loss function	`squared_error`, `absolute_error`, `huber`
Maximum tree depth	1–5
Minimum samples split	2–20
Minimum samples leaf	1–12
Maximum features	None, `sqrt`, `log2`, 0.4, 0.6, 0.8, 1.0

Subsample	0.50–1.00
Maximum leaf nodes	None, 3, 5, 10, 20, 40
Criterion	`friedman_mse`, `squared_error`
Minimum impurity decrease	0.00–0.02
Cost-complexity pruning (`ccp_alpha`)	0.00–0.02
Validation fraction	0.10, 0.15, 0.20
Early stopping (`n_iter_no_change`)	10, 20, 30
Early stopping tolerance (`tol`)	1e−5–1e−3 (log scale)
Huber alpha	0.80–0.95 (only when Huber loss was selected)

uncorrected proof

3.7 Model Validation

Model validation was conducted using a nested cross-validation strategy consisting of five outer folds and four inner folds. In each outer fold, one subset was reserved for out-of-sample testing while the remaining data were used for model development. Within the training data of each outer fold, a 4-fold inner cross-validation procedure was performed to optimize the ANN hyperparameters using Optuna. This process was repeated across all five outer folds, and the pooled out-of-sample predictions were used to evaluate the final model performance.

3.8 Performance Evaluation

Model performance was evaluated using multiple complementary metrics, including RMSE, MAE, MedianAE, R^2, explained variance, and Pearson's correlation coefficient. RMSE was treated as the primary error metric because it penalizes larger prediction errors; however, model selection was not based on RMSE alone. The relative performance of the models was interpreted jointly across all reported metrics, and no claim of model superiority was made in the absence of formal statistical testing of predictive performance differences.

4. Results

4.1. Structural Model Results

Table 4 presented the structural model results. All structural path coefficients were positive and statistically significant ($p < 0.05$), indicating significant associations between the predictor constructs and Financial Distress. Among the predictors, Market Risk exhibited the strongest association with Financial Distress ($\beta = 0.417$, $t = 7.234$), followed by Resource Risk ($\beta = 0.316$, $t = 5.582$), Technology Risk ($\beta = 0.197$, $t = 3.629$), and Project Risk ($\beta = 0.163$, $t = 2.590$).

Table 4: Structural Model Results

	Original sample (O)	Standard deviation (STDEV)	T statistics (O/ STDEV)	P values
Market Risk → Financial Distress	0.417	0.058	7.234	0.000
Project Risk → Financial Distress	0.163	0.063	2.590	0.010
Resource Risk → Financial Distress	0.316	0.057	5.582	0.000
Technology Risk → Financial Distress	0.197	0.054	3.629	0.000

uncorrected proof

Table 5: R square

Table showed that the structural model explained a substantial proportion of the variance in financial distress ($R^2 = 0.667$; Adjusted $R^2 = 0.657$), indicating that the selected risk indicators provide strong explanatory power for predicting financial distress.

	R square	R square adjusted
Financial distress	0.667	0.657

Table 6: Fornel-larcker criterion

Table 6 presented the Fornell–Larcker criterion for discriminant validity. The diagonal values, representing the square root of AVE, were higher than all corresponding inter-construct correlations. The highest correlation was observed between Financial Distress and Market Risk (r = 0.674), which remained below the lowest diagonal value (0.801). Therefore, the results supported adequate discriminant validity among the constructs

	Financial Distress	Market Risk	Project Risk	Resource Risk	Technology Risk
Financial Distress	0.846				
Market Risk	0.674	0.906			
Project Risk	0.525	0.366	0.817		
Resource Risk	0.602	0.360	0.397	0.840	
Technology Risk	0.556	0.420	0.425	0.360	0.801

Table 7 :Heterotrait-Monotrait Ratio (HTMT)

Table 7 presented the Heterotrait–Monotrait (HTMT) ratios used to rigorously evaluate discriminant validity. All pairwise ratios ranged from 0.380 to 0.738, remaining well below the conservative cut-off threshold of 0.85 (and the more liberal benchmark of 0.90). The highest value was observed between Financial Distress and Market Risk (0.738). These results demonstrated that the constructs were empirically distinct, confirming discriminant validity.

uncorrected proof

	Financial Distress	Market Risk	Project Risk	Resource Risk	Technology Risk
Financial Distress					
Market Risk	0.738				
Project Risk	0.586	0.390			
Resource Risk	0.673	0.380	0.430		
Technology Risk	0.627	0.454	0.469	0.389	

Table 8: Fit summary:

The SRMR value of 0.054 indicates a good model fit. Regarding the NFI (0.840), it should be noted that this threshold was established for factor-based models; for composite models, which characterize the PLS-SEM approach used in this study, thresholds for the NFI are still to be determined (Henseler et al, 2016). Given this, along with the strong reliability and validity results reported above, the overall model fit is considered adequate.

	Saturated Model	Estimated Model
SRMR	0.054	0.054
d_ULS	1.183	1.183
d_G	0.639	0.639
Chi-Square	504.797	504.797
NFI	0.840	0.840

Table 9: Outer loading

Table 9 presented the outer loadings of the measurement model. All indicator loadings exceeded the recommended threshold of 0.70, confirming satisfactory indicator reliability.

Construct	Item Code	Outer Loading
Technology Risk	V1	0.823
Technology Risk	V2	0.725
Technology Risk	V3	0.826
Technology Risk	V4	0.854
Technology Risk	V5	0.818
Technology Risk	V6	0.755
Market Risk	V7	0.899
Market Risk	V8	0.901
Market Risk	V9	0.914
Market Risk	V10	0.897
Market Risk	V11	0.904
Market Risk	V12	0.923
Project Risk	V13	0.803
Project Risk	V14	0.803
Project Risk	V15	0.813
Project Risk	V16	0.840
Project Risk	V17	0.822
Project Risk	V18	0.821
Resource Risk	V19	0.857
Resource Risk	V20	0.850
Resource Risk	V21	0.799
Resource Risk	V22	0.883
Resource Risk	V23	0.859
Resource Risk	V24	0.790
Financial Distress	V25	0.870

uncorrected proof

Financial Distress	V26	0.825
Financial Distress	V27	0.858
Financial Distress	V28	0.830

4.3. Reliability(Composit reliability) & Validity(Convergent validity (AVE))

A Cronbach's alpha value higher than 0.7 indicates acceptable reliability. However, some researchers have introduced a value of 0.6 as the limit of Cronbach's alpha coefficient for variables with a small number of questions. Cronbach's alpha criterion is a traditional one for determining the reliability of structures, but the PLS method uses a more modern criterion called composite reliability. Its advantage over Cronbach's alpha is that the reliability of structures is not calculated in absolute terms but according to the correlation of their structures with each other. As a result, both of these criteria are used to better measure reliability in the PLS method. According to Table 10, the values of Cronbach's alpha and composite reliability obtained were higher than 0.7, and, overall, the reliability of the research was acceptable. In case of validity, convergent validity values above 0.5 indicated acceptable convergent validity, and as the results showed, the conceptual model of the current research provided acceptable convergent validity.

Uncorrected proof

.

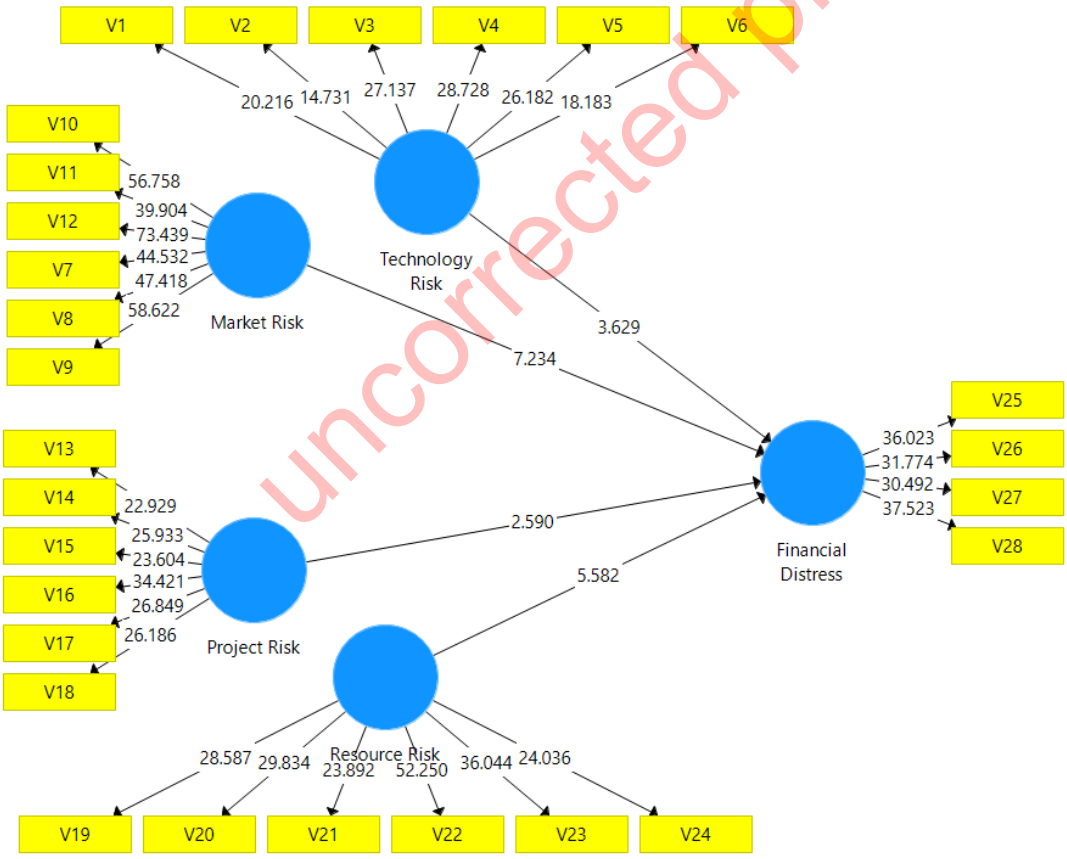
Table 10 : Reliability & Validity of the construct

	Cronbach's Alpha	rho_A	Composite Reliability	Average Variance Extracted (AVE)
Financial Distress	0.867	0.868	0.910	0.716
Market Risk	0.957	0.957	0.965	0.822
Project Risk	0.901	0.906	0.923	0.668
Resource Risk	0.916	0.921	0.935	0.706

Technology Risk	0.888	0.894	0.915	0.642

4.4. Structural model of research

Figure 1 illustrated the bootstrapping results of the PLS-SEM model. The t-values associated with the indicator loadings and structural paths were all greater than the critical value of 1.96, indicating statistical significance at the 5% level. The structural relationships between Market Risk, Technology Risk, Project Risk, Resource Risk, and Financial Distress were statistically significant, confirming the robustness of the estimated model.



uncorrected proof

Figure1: SEM Model

4.5 Machine learning results

The analysis was implemented in Python 3.14.3 on a Windows 10 operating system. Deep learning models were developed using the PyTorch library, while data preprocessing and machine learning utilities were provided by scikit-learn. Hyperparameter optimization was performed using Optuna. The pandas library was used for data handling and preparation, whereas NumPy and SciPy were employed for numerical and scientific computations.

4.5.1 ANN result:

Based on the pooled out-of-fold predictions, the proposed ANN model demonstrated strong predictive performance, achieving an RMSE of 0.561, an MAE of 0.457, an R^2 of 0.732, and a Pearson correlation coefficient of 0.856. Given that Financial Distress was measured on a continuous 1–5 scale, the RMSE value indicates that the model's predictions deviated from the observed financial distress scores by approximately 0.56 units on average, suggesting a relatively low prediction error and good out-of-sample performance. The optimal network architecture consisted of a single hidden layer with eight neurons using the ReLU activation function and the RMSprop optimizer.

uncorrected proof

Table 11: Predictive Performance and Optimal Architecture of the Proposed ANN Model

Metric	Value
RMSE	0.561
MAE	0.457
R^2	0.732
Pearson correlation	0.856
Hidden layers	1
Neurons	8
Activation function	ReLU
Optimizer	RMSprop
Learning rate	0.004833
Batch size	8
Dropout	0.15
Weight decay	1.02×10^{-5}
Selected epochs	82

uncorrected proof

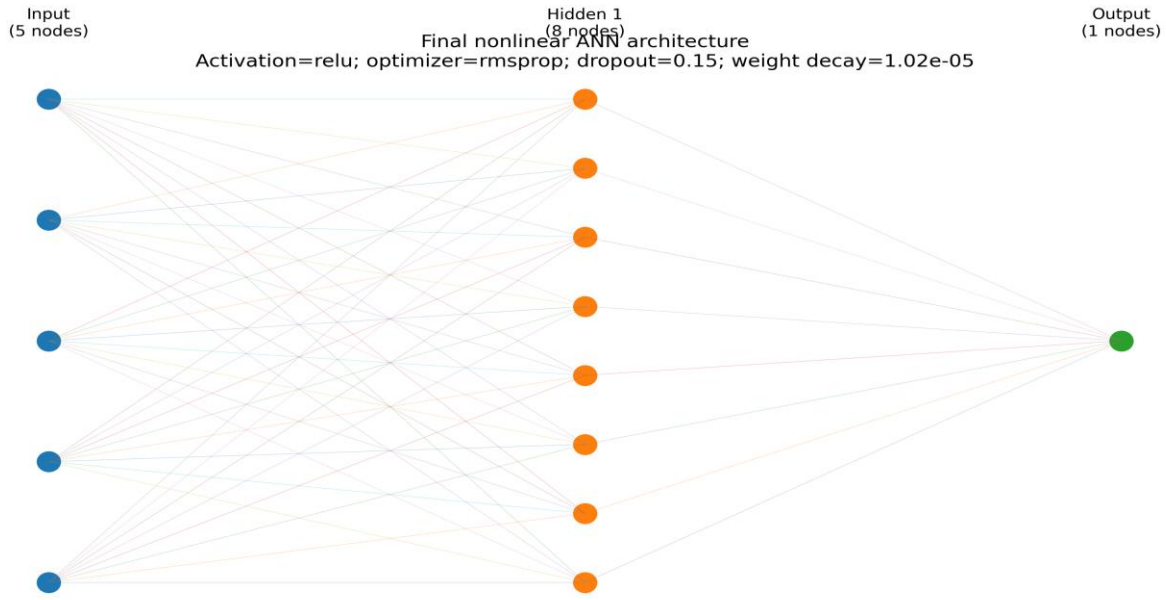


Figure 2 :Final architecture of ANN

uncorrected proof

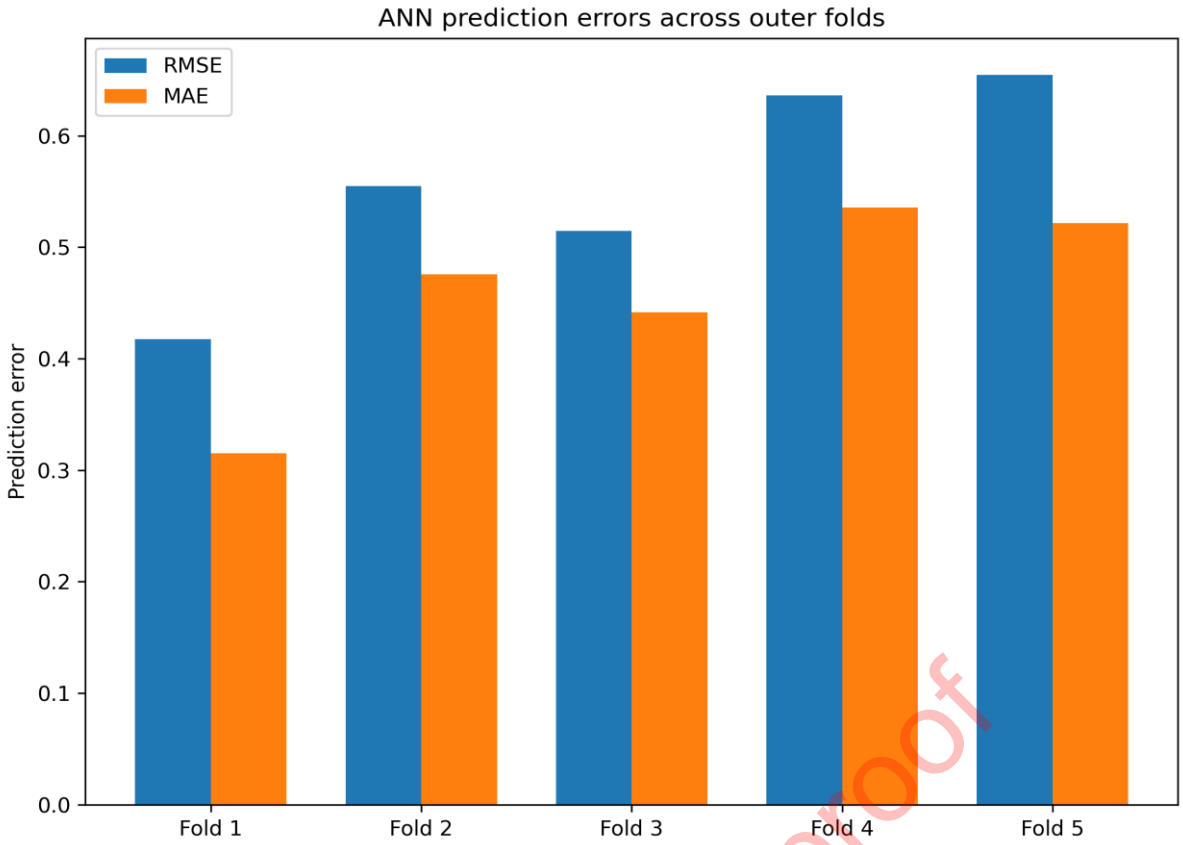


Figure 3: ANN prediction error across Folds

4.5.2 Random Forest result

Based on the pooled out-of-fold predictions, the proposed Random Forest model demonstrated strong predictive performance, achieving an RMSE of 0.562, an MAE of 0.445, an R^2 of 0.739, and a Pearson correlation coefficient of 0.867. Given that Financial Distress was measured on a continuous 1–5 scale, the RMSE indicates that the model's predictions deviated from the observed financial distress scores by approximately 0.55 units on average, suggesting relatively low prediction error and good out-of-sample performance. The optimal Random Forest architecture consisted of 300 regression trees with a maximum depth of 6, and the remaining hyperparameters were optimized using Optuna (Table 12). The fitted trees achieved a mean realized depth of 5.21 and an average of 11.36 terminal leaves.

uncorrected proof

Table 12: **Predictive Performance and Optimal Architecture of the Proposed Random Forest**

Parameter	Value
Performance Metrics	
RMSE	0.562
MAE	0.445
R^2	0.739
Pearson correlation ®	0.867
Optimal Hyperparameters	
Number of trees (`n_estimators`)	300
Criterion	`absolute_error`
Maximum depth (`max_depth`)	6
Minimum samples split (`min_samples_split`)	4
Minimum samples leaf (`min_samples_leaf`)	2
Maximum features (`max_features`)	0.6 (3 candidate predictors per split)
Bootstrap sampling (`bootstrap`)	True
Maximum samples (`max_samples`)	0.9
Maximum leaf nodes (`max_leaf_nodes`)	20
Minimum impurity decrease (`min_impurity_decrease`)	0.011892
Cost-complexity pruning (`ccp_alpha`)	0.013087
Mean realized tree depth	5.21
Mean terminal leaves per tree	11.36

uncorrected proof

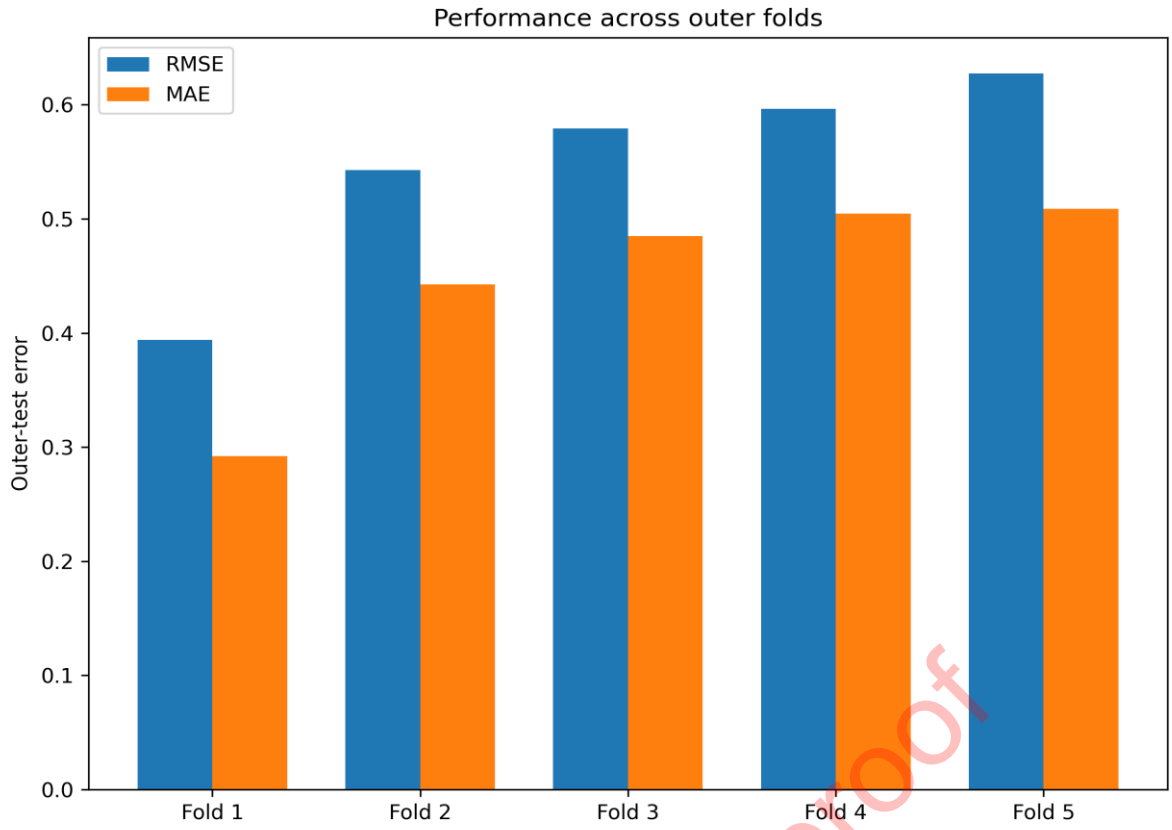


Figure 4: Random Forest prediction error across Folds

uncorrected proof

4.5.3. Gradient boosting result:

Based on the pooled out-of-fold predictions, the proposed Gradient Boosting model achieved an RMSE of 0.566, an MAE of 0.448, an R^2 of 0.727, and a Pearson correlation coefficient of 0.856, indicating good out-of-sample predictive performance. The optimal model was configured with a maximum of 975 boosting stages, of which 72 stages were fitted through early stopping. The optimized hyperparameters included a learning rate of 0.09371, maximum tree depth of 2, 40 maximum leaf nodes, a subsample ratio of 0.75, and the absolute error loss function. The remaining hyperparameters were optimized using Optuna (Table 13).

Table 13: **Predictive Performance and Optimal Architecture of the Proposed Gradient boosting**

Parameter	Value
Performance Metrics	

RMSE	0.566
MAE	0.448
R^2	0.727
Pearson's *r*	0.856
Hyperparameter Configuration	
Number of estimators	975 (72 fitted stages after early stopping)
Learning rate	0.09371
Maximum depth	2
Minimum samples split	3
Minimum samples leaf	8
Maximum leaf nodes	40
Subsample	0.75
Maximum features	log2
Loss function	Absolute Error
Split criterion	Friedman MSE
Validation fraction	0.10
Early stopping (*n_iter_no_change*)	30
Tolerance	1.7×10^{-5}
Cost-complexity pruning (*ccp_alpha*)	0.01635

uncorrected proof

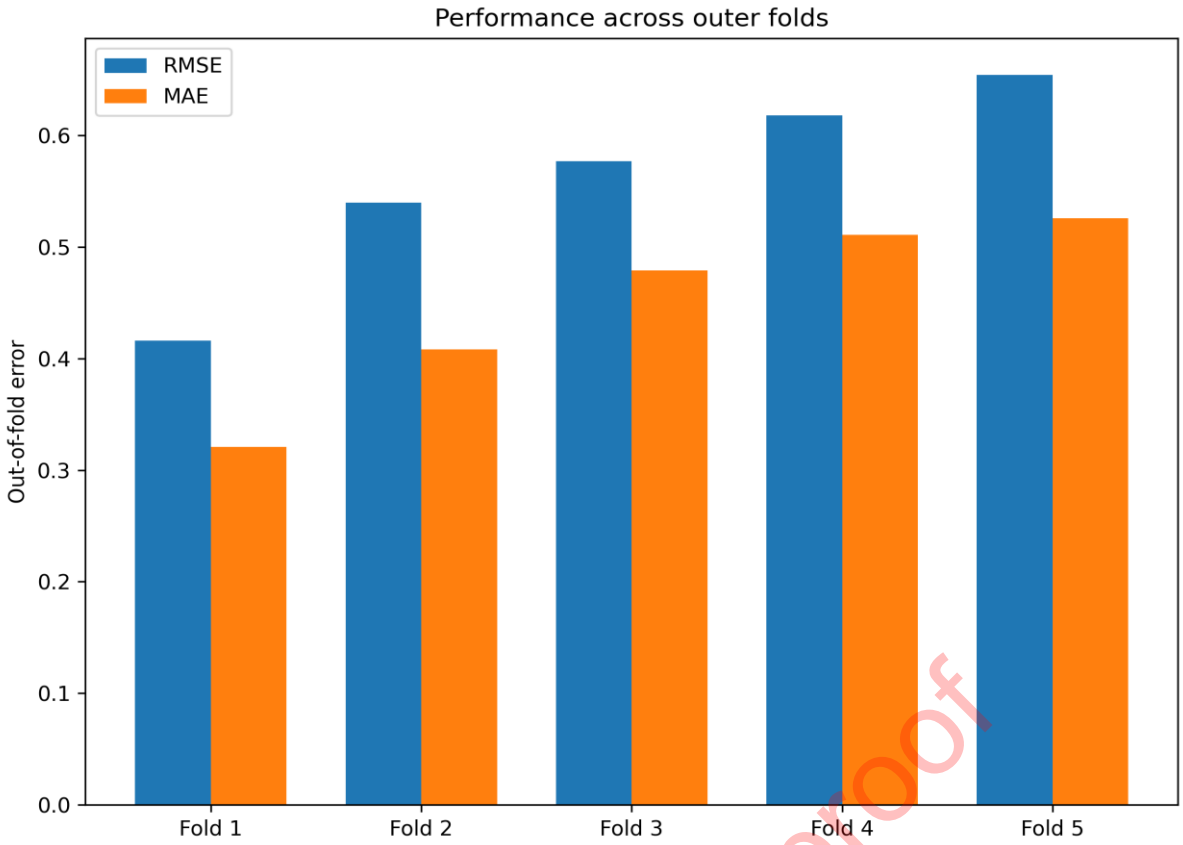


Figure 5: Gradient Boosting prediction error across Folds

uncorrected proof

4.6Comparison of ANN, Random Forest, Gradient Boosting:

Although multiple evaluation metrics (RMSE, MAE, R^2, and Pearson's *r*) were reported to provide a comprehensive assessment of model performance, RMSE was considered the primary criterion for model selection because it directly quantifies prediction error and assigns greater penalties to larger deviations between predicted and observed values. Values reported in brackets represent the 95% bootstrap confidence intervals.

Table 14: comparison

Performance Metric	Artificial Neural Network (ANN)	Random Forest (RF)	Gradient Boosting (GB)
RMSE	0.561 [0.503,0.613]	0.562 [0.499,0.625]	0.566 [0.497,0.631]
MAE	0.457 [0.404,0.512]	0.446 [0.394,0.501]	0.448 [0.390,0.508]
R^2	0.732 [0.659,0.790]	0.739 [0.668,0.795]	0.727 [0.651,0.785]
Pearson's *r*	0.856 [0.816,0.891]	0.868 [0.827,0.904]	0.856 [0.810,0.896]

4.7 Explainability Analysis Based on the ANN Model

Because ANN yielded the lowest RMSE, it was retained for the subsequent SHAP analysis. However, this selection should not be interpreted as evidence that ANN was statistically superior to Random Forest. The two models showed closely matched overall performance, and Random Forest achieved better MAE, R^2, and Pearson correlation. ANN was selected for interpretability primarily because its continuous non-linear function approximation was compatible with the intended explanatory analysis.

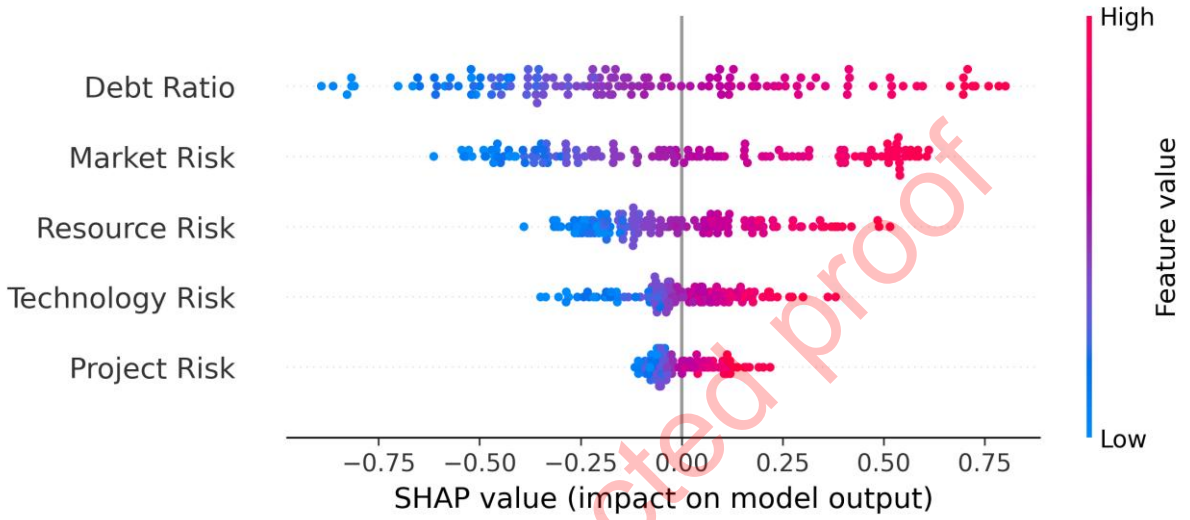


Figure 6: SHAP summary

Figure 6 depicted the summary plot of SHAP. As shown in the summary plot, the plot illustrates the contribution of the predictors of financial distress pertaining to the ultimate ANN model where Debt Ratio was the most important factor among the others. Consequently, Market Risk, Resource Risk, Technology Risk, and Project Risk are respectively ranked among the second and fifth place in terms of importance. Overall, as shown in the plot, higher values of Debt Ratio, Market Risk, Resource Risk, and Technology Risk are associated with higher predicted financial distress scores.

uncorrected proof

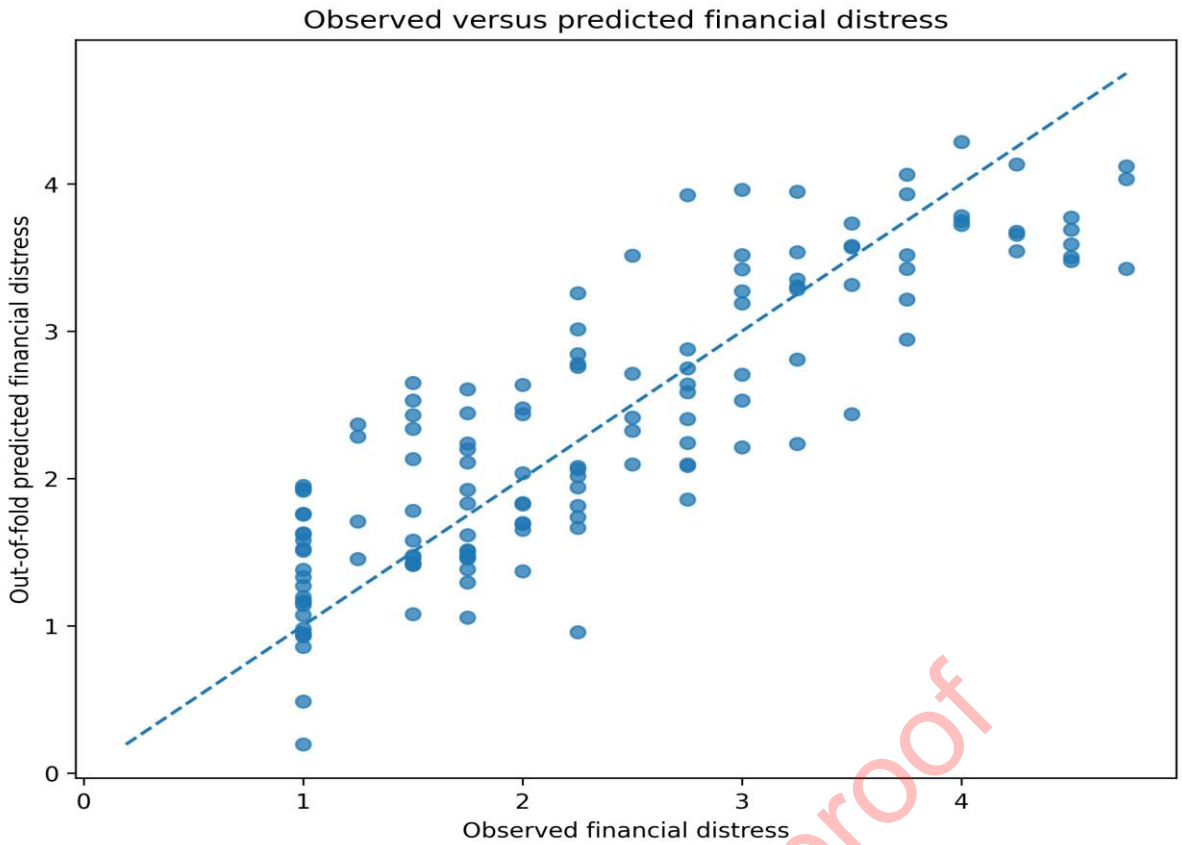


Figure 7: Observes vs Predicted

The correlation between the obtained financial distress scores and those predicted by the out-of-fold approach is captured in Figure 7. As observed predicted results correspond to each of those predicted by the model and the majority of them are very close to the 45° reference line indicating that the model was able to predict values successfully. While some deviations have been also observed, particularly for higher scoring instances, overall, the model was .relatively successful in making out-of-sample data predictions

uncorrected proof

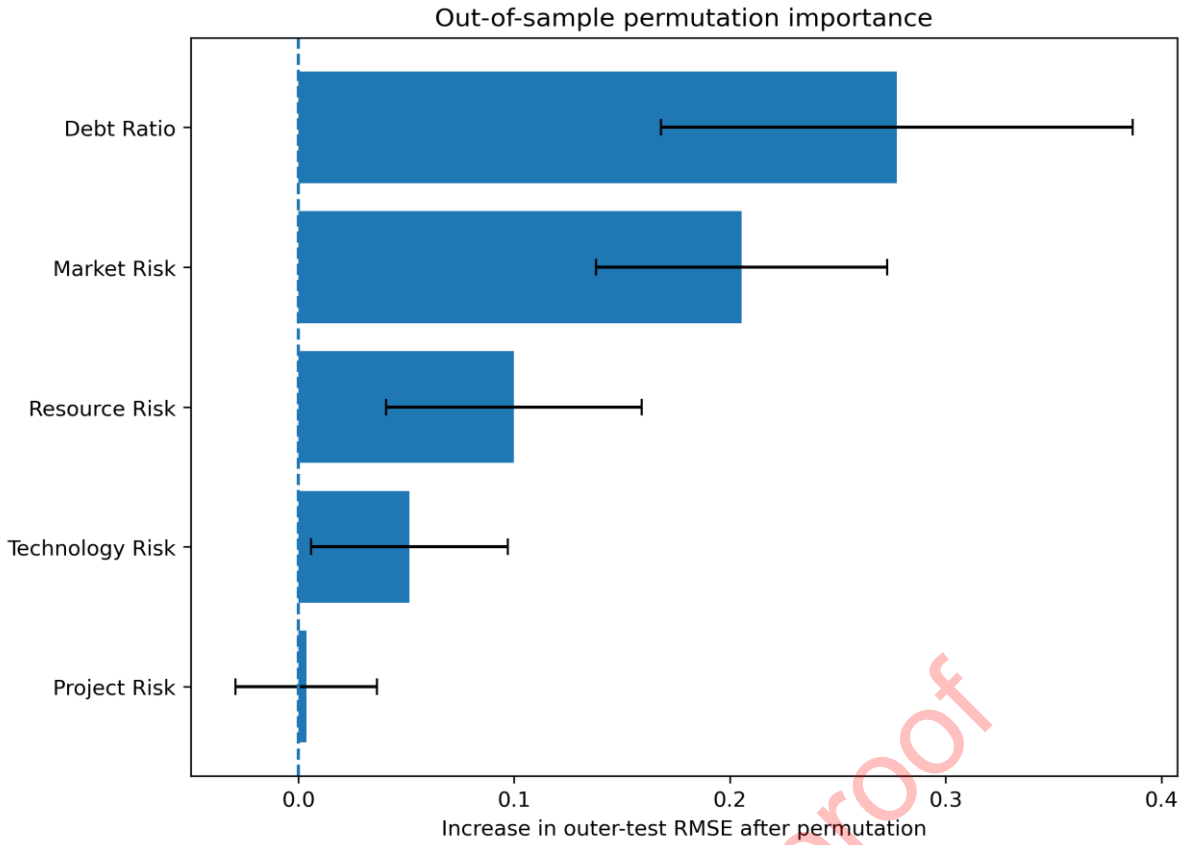


Figure 8: Permutation Importance

Figure 8 displayed the out-of-sample permutation importance of the individual predictors. Debt Ratio made the greatest predictive contribution to the model's function, followed by Market Risk and Resource Risk. Technology Risk exhibited a lower predictive contribution than both of these predictors, while Project Risk showed the least overall significance. Collectively, these results indicated that financially related and market-related factors were the most influential predictors in financial distress prediction.

uncorrected proof

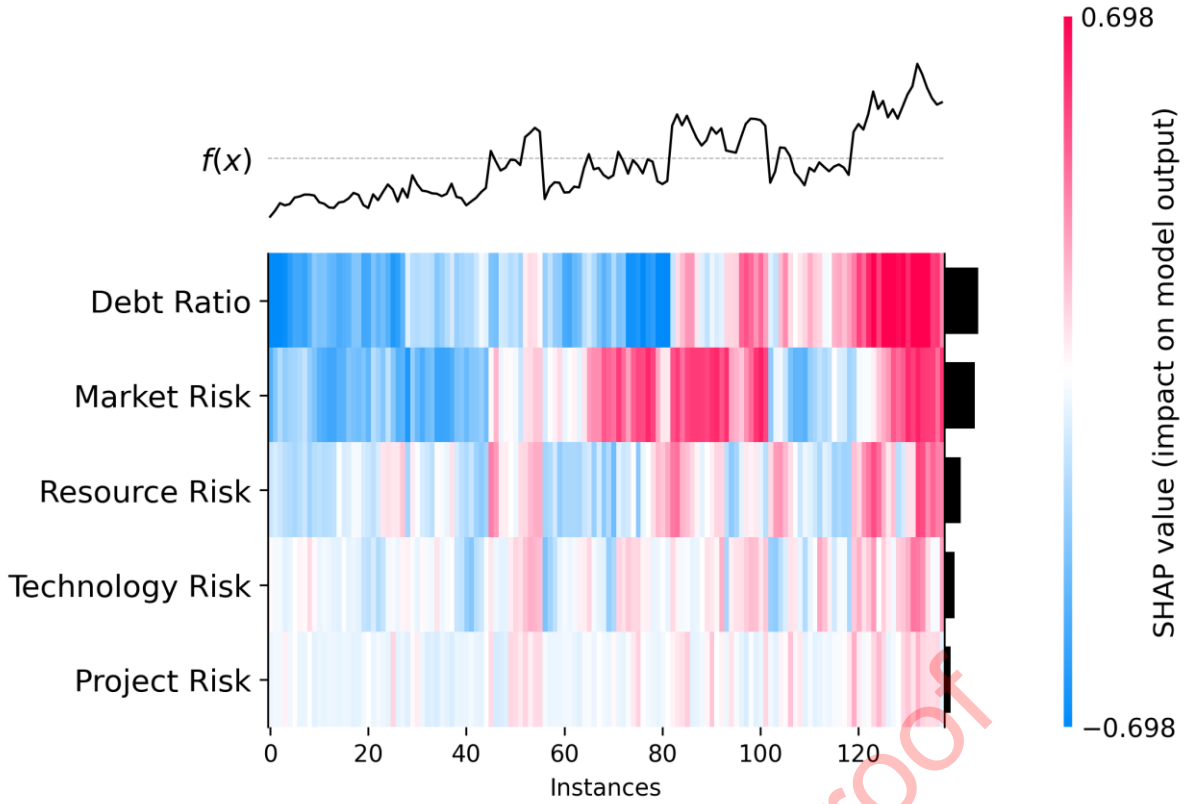


Figure 9: SHAP Heatmap

Figure 9 presented the SHAP heatmap for the selected ANN model. Each column represents an observation, while each row corresponds to one predictor. Red and blue colors indicate positive and negative contributions to the predicted financial distress score, respectively the heatmap showed that Debt Ratio and Market Risk have the largest and most consistent contribution to the model predictions across observations, whereas Resource Risk, Technology Risk, and Project Risk contribute to a lesser extent.

uncorrected proof

5. **Discussion & conclusion**

In the present study, three machine learning algorithms, namely Artificial Neural Networks (ANN), Random Forest (RF), and Gradient Boosting (GB), were developed and compared to evaluate their predictive performance within a financial distress prediction framework for Iranian manufacturing small and medium-sized enterprises (SMEs).

Recent advances in financial distress forecasting have emphasized the importance of combining financial and non-financial indicators to improve the accuracy of FDP models, especially for medium-sized companies. These studies have used advanced techniques such

as data mining, principal component analysis, and logistic regression to examine a wide range of variables, from traditional financial ratios to technological innovation and human capital indicators (Ma & Li, 2013).

The emergence of big data enabled the inclusion of non-financial variables and significantly improved the forecasting performance of models (Lai, 2022). For example, models combining non-financial indicators with traditional financial measures showed high performance in predicting financial crises for IT companies (Lu et al., 2019).

The competitive predictive performance observed across the evaluated models with ANN achieving a marginally lower RMSE (0.561) and Random Forest demonstrating superior MAE, R^2, and Pearson correlation highlights the capacity of non-linear algorithms to model complex interactions among financial and non-financial risk factors. This interpretation is aligned with the multidimensional nature of financial distress, where leverage, market conditions, scarcity of resources, technological problems, and project-related issues interact continuously rather than acting independently. While market uncertainty can plausibly depress sales, strain operating cash flows, and heighten insolvency risk, our cross-sectional model identifies predictive relevance rather than testing this sequential transmission channel directly. Non-linear models like neural networks and tree ensembles are well suited for capturing such complex interactions among risk factors

The predictive relevance of the Debt Ratio can be clarified through the trade-off theory of capital structure. According to this approach, while debt may create tax benefits, high levels of leverage increase fixed repayment commitments as well as expected costs of financial distress (Kraus & Litzenberger, 1973). In the case of SMEs, those pressures may be especially sharp due to the limited possibilities of accessing external financing and lower capacity to absorb temporary cash-flow difficulties. Thus, the predictive role of Debt Ratio in the present model corresponds to the idea that high levels of leverage diminish financial flexibility and create more vulnerability in case of operating cash flow decline.

The predictive contribution of Resource and Project Risks is consistent with the resource-based view of the firm. This perspective suggests that differences in valuable, scarce, and difficult-to-imitate resources and capabilities can underlie firms' ability to maintain competitive performance (Barney, 1991). In this context, heightened perceived risks may reflect perceived vulnerabilities in these internal capabilities rather than a direct measurement of objective firm resources.

Uncorrected proof

The findings of the present study support the view that non-financial risk indicators possess predictive potential independent of traditional financial information and serve as valuable indicators for financial distress (Wang & Liang, 2024).

The major implication of the research is that it is necessary to understand financial distress from the interrelations between financial, operation, technological and market risks rather than treating it solely as a purely financial problem. A key contribution of this study is that it combines PLS-SEM and machine learning techniques, where PLS-SEM evaluates structural construct relationships and machine learning provides robust predictive modeling.

The proposed prediction framework provides an effective decision-support tool that allows investors and other stakeholders to understand the general condition of the company.

6.Limitations and Future Research

This study is subject to several limitations. One of the main limitations relates to the measurement of financial distress. Owing to the limited availability of reliable financial information from Iranian SMEs, financial distress was evaluated using a five-point perceptual Likert scale instead of objective measures derived from audited financial statements. In practice, many SMEs were either unwilling to disclose sensitive financial information or did not maintain standardized financial reporting, particularly in the case of unlisted firms. As a result, the debt ratio was the only objective financial indicator that could be incorporated into the proposed model. Furthermore, due to data constraints in unlisted SMEs, firm-level controls (e.g., firm age and size) and broader financial indicators such as profitability (ROA, ROE), liquidity (current ratio), and cash flow measures were omitted. Although the debt ratio captures baseline leverage, future studies should integrate these additional metrics to mitigate omitted-variable bias and enhance model robustness. Additionally, while the first-stage PLS-SEM structural estimation provided a baseline linear reference ($R^2=0.667$), the scope of this study was centered on integrating PLS-SEM with representative machine learning architectures (ANN, RF, GB) and SHAP explainability. Future research could expand this benchmark set by explicitly incorporating regularized linear baselines (e.g., Elastic Net) and gradient-boosted variants (e.g., XGBoost) combined with formal pairwise statistical tests.

Additionally, the model is constrained by its sample size (N=138) and the potential for self-reporting bias in management-assessed risk factors. Although Nested Cross-Validation was employed to mitigate overfitting risks, future research should prioritize larger, longitudinal,

uncorrected proof

and purely objective datasets to enhance generalizability and reduce reliance on subjective assessments.

Finally, given the cross-sectional survey design, potential endogeneity and reverse causality cannot be entirely ruled out. Thus, the findings and β coefficients reflect predictive associations rather than direct causal effects.

Conflicts of Interest:

The authors declare no conflicts of interest related to this work.

References

Uncorrected proof

Alavi, S. (2016). The influence of workforce agility on external manufacturing flexibility of Iranian SMEs. *International Journal of Technological Learning, Innovation and Development*, 8(1), 111–127. https://doi.org/10.1504/IJTLID.2016.075185

Alimi, A. F., & Malik, do, S. (2023). Handling multicollinearity and outliers in logistic regression using the robust Kibria–Lukman estimator. *Mathematics*, 14(1), 19. https://doi.org/10.3390/math14010019

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 23(4), 589–609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

Arasti, Z., Zandi, F., & Bahmani, N. (2014). Business failure factors in Iranian SMEs: Do successful and unsuccessful entrepreneurs have different viewpoints? *Journal of Global Entrepreneurship Research, 4*(1), 1–14. https://doi.org/10.1186/s40497-014-0010-7

Aven, T., & Renn, O. (2009). On risk defined as an event where the outcome is uncertain. *Journal of Risk Research*, 12(1), 1–11. https://doi.org/10.1080/13669870802488883

Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99–120. https://doi.org/10.1177/014920639101700108

Bartlett, J. E., Kotrlik, J. W., & Higgins, C. C. (2001). Organizational research: determining the appropriate sample size in survey research. *Information Technology, Learning, and Performance Journal*, 19(1), 43–50.

Belas, J., Bilan, Y., Ključnikov, A., Vincúrová, Z., & Macháček, J. (2015). Actual problems of business risk in SME segment. Case study from Slovakia. *International Journal of Entrepreneurial Knowledge*, 3(1). https://doi.org/10.37335/ijek.v3i1.26

Bešlić Obradović, D., Jakšić, D., Bešlić Rupić, I., & Andrić, M. (2018). Insolvency prediction model of the company: The case of the Republic of Serbia. *Economic Research-Ekonomska Istraživanja*, 31(1), 139–157. https://doi.org/10.1080/1331677X.2017.1421990

Breen, R., Karlson, K. B., & Holm, A. (2018). Interpreting and understanding logits, probits, and other nonlinear probability models. *Annual Review of Sociology*, 44, 39–54. https://doi.org/10.1146/annurev-soc-073117-041429

Civelek, M., Krajcik, V., & Fialova, V. (2023). The impacts of innovative and competitive abilities of SMEs on their different financial risk concerns: System approach. *Oeconomia Copernicana*, 14(2), 327–354. https://doi.org/10.24136/oc.2023.009

Colak, M. S. (2021). A new multivariate approach for assessing corporate financial risk using balance sheets. *Borsa İstanbul Review*, 21(3), 239–255. https://doi.org/10.1016/j.bir.2020.10.007

Crovini, C. (2019). *Risk management in small and medium enterprises*. Routledge. https://doi.org/10.4324/9780429276316

Deng, S., Li, Y., Zhu, Y., Wang, B., Ning, H., Yi, S., & Shimada, T. (2025). Financial distress warning and risk path analysis for Chinese listed companies: An interpretable machine learning approach. *Economic Modelling*, 152, 107288. https://doi.org/10.1016/j.econmod.2025.107288

Deng, X.-L., & Wang, T. (2008). Stock market factors and risk of financial distress: An empirical analysis using Cox proportional hazard model. *2008 4th International Conference on Wireless Communications, Networking and Mobile Computing*, 1–4. https://doi.org/10.1109/WICOM.2008.2420

Dube, F., Nzimande, N., & Muzindutsi, P.-F. (2023). Application of artificial neural networks in predicting financial distress in the JSE financial services and manufacturing companies. *Journal of Sustainable Finance & Investment*, 13(1), 723–743. https://doi.org/10.1080/20430795.2021.2017257

Durica, M., Valaskova, K., & Janoskova, K. (2019). Logit business failure prediction in V4 countries. *Engineering Management in Production and Services*, 11(4), 54–64. https://doi.org/10.2478/emj-2019-0033

Dvorský, J., Belas, J., Gavurová, B., & Brabenec, T. (2020). Business risk management in the context of small and medium-sized enterprises. *Economic Research-Ekonomska Istraživanja*, 34(1), 1690–1708. https://doi.org/10.1080/1331677X.2020.1844588

Dvorský, J., Petráková, Z., & Fialová, V. (2020). Perception of business risk by entrepreneurs according to experience with the business failure. *International Journal of Entrepreneurial Knowledge*, 8(1), 76–88. https://doi.org/10.37335/ijek.v8i1.104

Fang, M., Liu, F., & Park, K. (2022). Is inventory performance helping to improve SME credit ratings? The moderating role of supply chain concentration. *Applied Economics Letters*. https://doi.org/10.1080/13504851.2022.2156455

Fasano, F., Adornetto, C., Zahid, I., La Rocca, M., Montaleone, L., Greco, G., & Cariola, A. (2025). The dilemma of accuracy in bankruptcy prediction: a new approach using explainable

Uncorrected proof

AI techniques to predict corporate crises. *European Journal of Innovation Management*, 28(11), 1–22. https://doi.org/10.1108/EJIM-06-2024-0633

Ferreira de Araújo Lima, P., Marcelino-Sádaba, S., & Verbano, C. (2021). Successful implementation of project risk management in small and medium enterprises: A cross-case analysis. *International Journal of Managing Projects in Business*, 14(4), 1023–1045. https://doi.org/10.1108/IJMPB-06-2020-0203

Geoffrey, M. O. (2020). Corporate governance, board independence and financial distress: A critical review of literature. *The International Journal of Business & Management*, 8(1). https://doi.org/10.24940/theijbm/2020/v8/i1/bm2001-037

Gichaiya, M. W., Muchina, S., & Macharia, S. (2019). Corporate Risk, Firm Size and Financial Distress: Evidence from Non-Financial Firms Listed In Kenya. *IOSR Journal of Economics and Finance*. https://karuspace.karu.ac.ke/handle/20.500.12092/2393

Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS path modeling in new technology research: Updated guidelines. *Industrial Management & Data Systems*, 116(1), 2–20.

Idrees, S., & Qayyum, A. (2018). The impact of financial distress risk on equity returns: A case study of non-financial firms of Pakistan Stock Exchange (MPRA Paper No. 85346). *University Library of Munich*. https://mpra.ub.uni-muenchen.de/85346/

Jalali, A., Jaafar, M., & Ramayah, T. (2013). Influence of entrepreneurial orientation on the financial performance: Evidence from SMEs in Iran. *Middle East Journal of Management*, 1(2), 168–185. https://doi.org/10.1504/MEJM.2013.057263

Kenett, R. S., & Tapiero, C. S. (2009). Quality, risk and the Taleb quadrants. *Risk and Decision Analysis*, 1(4), 231–246. https://doi.org/10.3233/RDA-2009-0020

Kidane, H. W. (2004). *Predicting financial distress in IT and services companies in South Africa* (Doctoral dissertation). University of the Free State, Bloemfontein, South Africa.

Kraus, A., & Litzenberger, R. H. (1973). A state-preference model of optimal financial leverage. *The Journal of Finance*, 28(4), 911–922. https://doi.org/10.1111/j.1540-6261.1973.tb01415.x

Kristanti, F. T., Febrianta, M. Y., Salim, D. F., Riyadh, H. A., Sagama, Y., & Beshr, B. A. H. (2024). Advancing financial analytics: Integrating XGBoost, LSTM, and Random Forest algorithms for precision forecasting of corporate financial distress. *Journal of Infrastructure, Policy and Development*, 8(8), 4972. https://doi.org/10.24294/jipd.v8i8.4972

Kristanti, F. T., Rahayu, S., & Isynuwardhana, D. (2019). Integrating capital structure, financial and non-financial performance: Distress prediction of SMEs. *GATR Accounting and Finance Review*, 4(2), 56–62. https://doi.org/10.35609/afr.2019.4.2(4)

Lai, M. (2022). Analysis of financial risk early warning systems of high-tech enterprises under big data framework. *Scientific Programming*, 2022, Article ID 9055294. https://doi.org/10.1155/2022/9055294

Uncorrected Proof

Liu, F., & Park, K. (2020). Managing firm risk through supply chain dependence: An SME perspective. *Journal of Business & Industrial Marketing*, 36(12), 2231–2242. https://doi.org/10.1108/JBIM-05-2019-0229

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

Ma, H., & Li, F. (2013). Empirical study on early-warning of financial crisis for listed manufacturing companies in Yangzi River Delta. In *Proceedings of 2013 IEEE International Conference on Grey Systems and Intelligent Services (GSIS)* (pp. 296–299). IEEE. https://doi.org/10.1109/GSIS.2013.6714803

M Ma'aji, M., Abdullah, N. A. H., & Khaw, K. (2019). FINANCIAL DISTRESS AMONG SMEs IN MALAYSIA: AN EARLY WARNING SIGNAL. *International Journal of Business and Society*, 20, 775–792.

Madrid-Guijarro, A., García-Pérez-de-Lema, D., & van Auken, H. (2011). An analysis of non-financial factors associated with financial distress. *Entrepreneurship & Regional Development*, 23(3–4), 159–186. https://doi.org/10.1080/08985620903233911

Moosa, I. A. (2007). The science of risk management. In *Operational Risk Management*. Finance and Capital Markets Series. Palgrave Macmillan, London. https://doi.org/10.1057/9780230591486_1

Moradi, M., Salehi, M., Ghorgani, M. E., & Sadoghi Yazdi, H. (2013). Financial distress prediction of Iranian companies by using data mining techniques. *Organizacija*, 46(1), 20–27. https://doi.org/10.2478/orga-2013-0003

Moradlou, H., & Perera, T. (2017). Identification of the barriers in implementation of lean principles in Iranian SMEs: Case study approach. *Global Journal of Management and Business Research*, 17(1), 33–41. https://journalofbusiness.org/index.php/GJMBR/article/view/2252

Ngwenya, S. (2018). Assessing the State of Financial Distress of Listed Gold and Platinum Mining Companies in South Africa. *Acta Universitatis Danubius. OEconomica*, 14(4), 655–677.

Papíková, L., & Papík, M. (2026). Similarity failure proximity score: A network-based metric for bankruptcy prediction. *Finance Research Letters*, 94, 109631. https://doi.org/10.1016/j.frl.2026.109631

Pertseva, E. Y., Skobarev, V. Y., & Telenkov, E. E. (2021). Sustainable development and risk management. *Issues of Risk Analysis*, 18(4), 16–27. https://doi.org/10.32686/1812-5220-2021-18-4-16-27

Premachandra, I. M., Bhabra, G. S., & Sueyoshi, T. (2009). DEA as a tool for bankruptcy assessment: A comparative study with logistic regression technique. *European Journal of Operational Research*, 193(2), 412–424. https://doi.org/10.1016/j.ejor.2007.11.036

Rado, M. (2013). *Testing and calibrating the Altman Z-score for the U.K.* (Bachelor's thesis). Erasmus University Rotterdam. https://thesis.eur.nl/pub/13993/Rado-M.-344734.pdf

Uncorrected proof

Ragab, Y. M., & Saleh, M. A. (2021). Non-financial variables related to governance and financial distress prediction in SMEs: Evidence from Egypt. *Journal of Applied Accounting Research*, 23(3), 604–627. https://doi.org/10.1108/JAAR-02-2021-0025

Sethi, S. R., & Mahadik, D. A. (2026). Machine learning methods for financial distress prediction: an analytical overview and implications for sustainability goals. *Journal of Chinese Economic and Business Studies*, 24(1), 1–52. https://doi.org/10.1080/14765284.2025.2549237

Sfakianakis, E. (2021). Bankruptcy prediction model for listed companies in Greece. *Investment Management and Financial Innovations*, 18(2), 166–180. https://doi.org/10.21511/imfi.18(2).2021.14

Sharma, S., & Mittal, A. K. (2024). Evolution of financial, non-financial, and macroeconomic predictors in corporate bankruptcy: A comprehensive review. *Economic Sciences*, 20(2), 70–83. https://doi.org/10.69889/0vxzzm07

Song, M., & Montoya-Weiss, M. M. (2001). The effect of perceived technological uncertainty on Japanese new product development. *Academy of Management Journal*, 44(1), 61–80. https://doi.org/10.2307/3069337

Sudaryo, Y., Hamdani, D., Sofiati, N. A., Sipahutar, D. H. N., & Sutisna. (2025). Assessing the drivers of financial distress in Indonesian rattan SMEs through digital and financial perspectives. *Aptisi Transactions on Technopreneurship (ATT)*, 7(3), 904–913. https://doi.org/10.34306/att.v7i3.534

Tarighi, H., Hosseiny, Z. N., Abbaszadeh, M. R., Zimon, G., & Haghighat, D. (2022). How do financial distress risk and related party transactions affect financial reporting quality? Empirical evidence from Iran. *Risks*, 10(3), 46. https://doi.org/10.3390/risks10030046

Virglerova, Z., Ivanova, E., Dvorský, J., Belas, J., & Krulický, T. (2021). Selected factors of internationalisation and their impact on the SME perception of the market risk. *Oeconomia Copernicana*, 12(4), 1011–1032. https://doi.org/10.24136/oc.2021.033

Vo, D. H. (2023). Market risk, financial distress and firm performance in Vietnam. *PLoS ONE*, 18(7), e0288621. https://doi.org/10.1371/journal.pone.0288621

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

Xu, L., Qi, Q., & Sun, P. (2019). Early-warning model of financial crisis: An empirical study based on listed companies of information technology industry in China. *Emerging Markets Finance and Trade*, 56(7), 1601–1614. https://doi.org/10.1080/1540496X.2019.1703104

Yazdanfar, D., & Öhman, P. (2020). Financial distress determinants among SMEs: Empirical evidence from Sweden. *Journal of Economic Studies*, 47(3), 547–560. https://doi.org/10.1108/JES-01-2019-0030

Zhong, J., & Wang, Z. (2022). Artificial intelligence techniques for financial distress prediction. *AIMS Mathematics*, 7(12), 20891–20908. https://doi.org/10.3934/math.20221145

Uncorrected proof

Appendix A – Questionnaire Items

Construct	number	Item
Technology Risk	**1**	The technologies used in our industrial projects are based on a well-developed body of scientific knowledge
Technology Risk	**2**	The rate (speed and complexity) of technological change used in our industry projects is highly unpredictable
Technology Risk	**3**	The technology used in our products is changing rapidly
Technology Risk	**4**	Technological changes in research and development projects are often unpredictable
Technology Risk	**5**	The technology was not well understood, the phenomena were not well defined and it was very difficult to predict the future state
Technology Risk	**6**	It is very difficult to predict where the technology used in our products will be in the next 2 or 3 years
Market Risk	**7**	Competition is fierce
Market Risk	**8**	It is difficult to predict the performance of competitors
Market Risk	**9**	Severe fluctuations in product sales prices are one of our main concerns
Market Risk	**10**	Our customers' preferences and demands are constantly changing
Market Risk	**11**	There are requests for our products from customers who have not been our customers before
Market Risk	**12**	The needs of new customers are different from those of existing customers
Project Risk	**13**	Projects are completed at the planned cost
Project Risk	**14**	Projects are completed according to the expected timeline
Project Risk	**15**	Task scheduling is inappropriate
Project Risk	**16**	Project completion planning is inappropriate
Project Risk	**17**	The estimated time to complete the project is unrealistic
Project Risk	**18**	Project progress is not monitored adequately
Resource Risk	**19**	Your organization has sufficient human resources to complete the project
Resource Risk	**20**	The staff is highly experienced
Resource Risk	**21**	There are outdated production facilities
Resource Risk	**22**	Financial resources are insufficient to complete projects
Resource Risk	**23**	Key resources that impact the project are sometimes lost
Resource Risk	**24**	There is a lack of cooperation and coordination between team members
Financial Distress	**25**	Our sales have declined in the last two years
Financial Distress	**26**	Our company usually makes a loss
Financial Distress	**27**	The company has a high amount of loan-related debt
Financial Distress	**28**	The company has a high volume of receivables from customers

Uncorrected proof

Appendix B_ Cross loading

	Financial Distress	Market Risk	Project Risk	Resource Risk	Technology Risk
V1	0.484	0.343	0.376	0.286	0.823
V10	0.622	0.897	0.315	0.272	0.370
V11	0.580	0.904	0.314	0.325	0.367
V12	0.631	0.923	0.327	0.338	0.409
V13	0.397	0.266	0.803	0.322	0.420
V14	0.514	0.324	0.803	0.395	0.327
V15	0.359	0.256	0.813	0.217	0.273
V16	0.433	0.337	0.840	0.355	0.393
V17	0.443	0.325	0.822	0.292	0.334
V18	0.393	0.269	0.821	0.333	0.332
V19	0.499	0.289	0.268	0.857	0.294
V2	0.354	0.316	0.246	0.203	0.725
V20	0.520	0.298	0.357	0.850	0.342
V21	0.446	0.253	0.309	0.799	0.263
V22	0.559	0.351	0.356	0.883	0.359
V23	0.535	0.367	0.345	0.859	0.329
V24	0.464	0.240	0.366	0.790	0.210
V25	0.870	0.543	0.513	0.494	0.510
V26	0.825	0.533	0.424	0.512	0.483
V27	0.858	0.587	0.437	0.521	0.416
V28	0.830	0.615	0.403	0.511	0.471
V3	0.405	0.294	0.334	0.235	0.826
V4	0.468	0.329	0.380	0.297	0.854
V5	0.432	0.366	0.342	0.345	0.818
V6	0.498	0.365	0.342	0.339	0.755
V7	0.589	0.899	0.350	0.324	0.370
V8	0.592	0.901	0.314	0.304	0.373
V9	0.645	0.914	0.369	0.389	0.394

uncorrected proof

Appendix C -VIF

Input Feature	VIF	Threshold (< 3.33)
Debt Ratio	1.4941	Satisfied
Technology Risk	1.4221	Satisfied
Project Risk	1.4153	Satisfied
Market Risk	1.3889	Satisfied
Resource Risk	1.3635	Satisfied

uncorrected proof