Journal of Computer Science

Research Article

Forecasting Corporate Financial Distress Using Explainable Machine Learning: Implications for Risk Management Technology

Kanjana Hinthaw and Warawut Narkbunnum

Mahasarakham Business School, Mahasarakham University, Thailand

Article history
Received: 03-07-2025
Revised: 20-04-2026
Accepted: 12-06-2026

Corresponding Author:
Warawut Narkbunnum
Mahasarakham Business School,
Mahasarakham University,
Thailand
Email: warawut.n@acc.msu.ac.th

Abstract: Financial distress prediction is a critical task in financial risk management, yet it remains challenging due to class imbalance, complex financial structures, and limited model interpretability. This study develops an integrated machine learning framework that combines feature selection, class imbalance handling, threshold optimization, and Explainable Artificial Intelligence (XAI) to improve both predictive performance and decision-making reliability. Using a dataset of 6,819 firms, three models, Logistic Regression, Random Forest, and XGBoost, are evaluated under a consistent experimental setting. Feature selection is conducted using mutual information, while the Synthetic Minority Over-Sampling Technique (SMOTE) is applied to address class imbalance. Model performance is assessed using precision, recall, F1-score, and ROC–AUC. The results show that ensemble-based models outperform the linear baseline. XGBoost achieves the highest ROC–AUC of 0.935, while Random Forest provides more balanced performance with an F1-score of 0.421 and a recall of 0.545 for the minority class. The application of SMOTE significantly improves recall across all models, with Logistic Regression reaching a recall of 0.818. In addition, threshold optimization further enhances classification performance by improving the precision–recall trade-off. SHAP-based analysis identifies profitability and leverage-related variables as the most influential predictors, confirming consistency with financial distress theory. Overall, the findings demonstrate that integrating machine learning, handling data imbalance, and XAI leads to more reliable and interpretable financial distress prediction. The proposed framework shifts the focus from model-centric evaluation to decision-oriented financial risk assessment, providing practical value for real-world applications.

Keywords: Financial Distress Prediction, Machine Learning, Class Imbalance, Explainable Artificial Intelligence, SHAP

Introduction

Financial distress prediction has long been a fundamental topic in financial economics and risk management, with the aim of identifying firms likely to experience financial distress. Early studies laid the foundation for this field by demonstrating that financial ratios can effectively predict corporate distress (Altman, 1968; Beaver, 1966). Subsequent developments introduced probabilistic frameworks, such as the Ohlson model, which further enhanced predictive capability by incorporating statistical inference techniques (Ohlson, 1980). These traditional approaches provided strong theoretical grounding; however, they are limited in their ability to capture the complex nonlinear relationships inherent in modern financial data.

In recent years, the accessibility of large-Compared with traditional statistical models, machine learning algorithms such as Random Forest, Support Vector Machines, and gradient boosting methods can model nonlinear interactions and handle high-dimensional data more effectively (Breiman, 2001; Chen and Guestrin,

2016; Gupta, 2025; Kristanti et al., 2025). Empirical data consistently shows that machine learning models surpass traditional statistical methods in predictive accuracy and robustness. Barboza et al. (2017); Lessmann et al. (2015), Recent studies further confirm that ensemble-based methods are effective in predicting financial distress Abrahamsen et al. (2024); Wang and Chi (2024).

Despite these advancements, several methodological challenges remain unresolved. A key challenge is class imbalance, since financially distressed firms usually form a minority in most datasets. This imbalance can cause models to favor the majority class, reducing distress detection accuracy. To mitigate this issue, methods like Synthetic Minority Over-sampling Technique (SMOTE) and cost-sensitive learning have been suggested Chawla et al. (2002); He and Garcia (2009). However, selecting an appropriate imbalance-handling strategy remains context-dependent and is an active area of research.

Another important challenge lies in feature selection and model complexity. Financial datasets often contain many correlated variables, which can introduce noise and reduce model interpretability. Feature selection methods have therefore been widely applied to improve model performance and stability (Guyon and Elisseeff, 2003). Simultaneously, the rise in complex machine learning models has led to concerns about transparency and interpretability, especially in high-stakes financial decisions.

To tackle these issues, Explainable Artificial Intelligence (XAI) has become an essential area of research. Methods like SHAP and LIME help understand model behavior by measuring how much each feature influences the prediction results (Lundberg & Lee, 2017; Ribeiro et al., 2016). Furthermore, recent research highlights that interpretability should be considered a core aspect of predictive model design, not just an afterthought, especially in areas where decisions can lead to substantial financial consequences (Rudin, 2019).

In addition to these challenges, most existing studies focus on structured financial data, often neglecting other relevant sources, such as macroeconomic indicators and behavioral signals. Furthermore, many models are developed using static datasets, which restricts their capacity to reflect temporal changes and shifting financial environments. These constraints emphasize the importance of developing more versatile and complete modeling frameworks.

Considering these gaps, this study seeks to create a machine-learning framework for predicting financial distress that includes feature selection, addresses class imbalance, and incorporates XAI methods. By blending accuracy with interpretability, this approach aims to offer a more reliable and practical tool for financial risk evaluation.

This study makes three main contributions. First, it compares various machine learning algorithms used for predicting financial distress. Second, it utilizes feature selection and data preprocessing methods to improve model performance and stability. Third, it employs Explainable AI (XAI) techniques to interpret the model's predictions and pinpoint key financial indicators associated with distress. These contributions seek to enhance both predictive accuracy and interpretability, fostering better decision-making in financial risk management.

Literature Review

Financial Distress Theory

Financial distress signifies a pivotal transitional phase in a company's lifecycle, marked by declining financial health and a growing failure to fulfill contractual commitments. Unlike bankruptcy, which is a formal legal condition, financial distress is an economic state that precedes failure, often manifested in declining profitability, liquidity constraints, excessive leverage, and operational inefficiencies. Consequently, financial distress is widely regarded as an early warning signal of potential firm failure and has been extensively examined in finance and accounting literature (Altman, 1968; Beaver, 1966; Habib et al., 2020).

The core literature focuses on ratio-based analysis. Beaver (1966) demonstrated that financial ratios, particularly cash flow and profitability measures, are highly predictive through univariate analysis. Expanding on this work, Altman (1968) developed the Z-score model, a multivariate discriminant approach that combines various financial indicators into one predictive score. Despite its straightforward design, the Z-score model is still popular because of its ease of interpretation and practical application (Altman, 2018).

However, traditional statistical models face several limitations. They usually depend on restrictive assumptions such as linearity, normal distribution, and independence among predictors, which are seldom met in real-world financial data. Additionally, they often overlook nonlinear relationships and complex interactions between financial variables, which hampers their predictive accuracy in dynamic settings (Ohlson, 1980; Zhao et al., 2024a).

To overcome these limitations, probabilistic models like logistic regression were introduced. Ohlson (1980) developed a logit-based model that predicts the likelihood of financial distress based on firm-specific features. This method relaxes many assumptions of discriminant analysis and yields probabilistic results, making it more appropriate for risk assessment and decision-making.

Theoretically, financial distress can be understood using capital structure and behavioral frameworks. According to the trade-off theory, firms weigh the advantages of debt against the likely costs associated with

financial distress (Kraus and Litzenberger, 1973). Meanwhile, agency theory highlights conflicts among stakeholders, where financial distress exacerbates issues such as risk shifting and underinvestment (Jensen and Meckling, 1976). In addition, signaling theory explains how distressed firms convey negative information to the market, thereby increasing information asymmetry and reducing investor confidence (Spence, 1973).

Recent research expands the scope of financial distress prediction by including a wider range of financial indicators and utilizing advanced analytical methods. Evidence shows that combining various data sources and using adaptable modeling strategies can greatly improve prediction accuracy (Abrahamsen et al., 2024; Zhao et al., 2024b). These advancements encourage moving toward machine learning methods for predicting financial distress.

Machine Learning in Financial Distress Prediction

Machine learning techniques for predicting financial distress have gained more attention due to increased computational power and data availability. Unlike traditional statistical approaches, these algorithms can detect nonlinear patterns and complex feature interactions without relying on strict assumptions, resulting in improved predictive accuracy (El Madou et al., 2024; Lessmann et al., 2015).

Early applications of machine learning encompassed Support Vector Machines (SVM) and Artificial Neural Networks (ANNs). SVMs are especially efficient in high-dimensional spaces and excel at classification problems with complicated decision boundaries (Min and Lee, 2005). Similarly, ANNs can capture nonlinear relationships using layered structures. However, both methods are limited in interpretability, restricting their application in financial decision-making.

Recently, ensemble learning methods have become the leading approach. Algorithms like Random Forest and XGBoost offer high accuracy, robustness, and scalability. Random Forest decreases variance using bagging, whereas boosting models enhance performance by sequentially concentrating on misclassified cases (Breiman, 2001; Chen and Guestrin, 2016). Recent empirical studies further confirm the effectiveness of ensemble-based approaches for financial distress prediction (Song et al., 2024; Wang and Chi, 2024).

Empirical evidence consistently demonstrates that ensemble methods outperform both traditional statistical models and earlier machine learning techniques. Benchmarking studies confirm that these models achieve superior performance across multiple datasets (El Madou et al., 2024; Lessmann et al., 2015). Furthermore, gradient boosting methods consistently demonstrate strong predictive performance, often achieving high ROC–AUC values in financial risk classification tasks, reflecting their effectiveness in capturing complex nonlinear relationships (Chen and Guestrin, 2016; Wang and Chi, 2024).

Despite these advantages, challenges remain, particularly in feature selection, model interpretability, and class imbalance. These issues highlight the need for integrated frameworks that combine predictive performance with interpretability and robustness.

Imbalanced Classification in Financial Distress Prediction

Imbalanced classification poses a significant challenge in predicting financial distress, since distressed firms usually make up only a small part of the dataset. This imbalance causes predictive models to favor the majority class, greatly decreasing their effectiveness in correctly identifying distressed firms (He and Garcia, 2009; Song et al., 2024).

To tackle this problem, various data-level and algorithm-level strategies have been suggested. Data-level methods, such as resampling, include techniques like SMOTE (Synthetic Minority Over-sampling Technique which create synthetic minority-class samples to balance the dataset (Chawla et al., 2002). While effective, these methods may introduce noise and increase the risk of overfitting, particularly when the minority class is highly sparse (Zhao et al., 2024a).

Algorithm-level methods, like cost-sensitive learning, increase misclassification costs for minority-class instances. These strategies boost recall and help the model better identify financially distressed firms, making them especially useful for risk-sensitive applications (Kuizinienė and Krilavičius, 2024).

Besides model design, evaluation metrics are essential in imbalanced classification. Accuracy alone is inadequate because it can hide weak performance on the minority class. More meaningful metrics include recall, precision, F1-score, and ROC-AUC, with recall especially crucial in predicting financial distress (Saito and Rehmsmeier, 2015). Recent studies further emphasize the importance of precision–recall-based evaluation in highly imbalanced financial datasets (Song et al., 2024).

Evaluation Metrics and Explainable Artificial Intelligence

In financial distress prediction, model evaluation extends beyond conventional accuracy measures to encompass reliability, robustness, and interpretability. Given the high-stakes nature of financial decision-making, predictive models must provide not only accurate classifications but also reliable probability estimates to support risk-sensitive decisions.

In situations with imbalanced datasets where financially distressed companies are a minority, accuracy is generally considered an inadequate metric because it can hide poor detection rates for the minority class. Metrics like recall,

precision, and F1-score are better suited for assessing classification performance. Recall is especially important in predicting financial distress since missing distressed firms can lead to substantial financial losses for stakeholders (Das et al., 2024; Gaudreault et al., 2021).

The trade-off between recall and precision is often summarized by the F1-score, offering a balanced assessment of false positives and negatives. Moreover, threshold-independent metrics like the area under the ROC curve (ROC–AUC) are commonly utilized to evaluate a model's ability to distinguish between classes.

However, recent studies indicate that in highly imbalanced situations, Precision Recall (PR) curves offer more meaningful insights because they better represent the performance of the minority class (Gaudreault et al., 2021).

Beyond classification performance, probability calibration has emerged as a crucial component of model evaluation. Calibration refers to the extent to which predicted probabilities correspond to actual observed outcomes. In financial applications, well-calibrated models are essential because probability estimates are directly used in risk assessment, credit scoring, and decision-making processes. Poorly calibrated models may lead to overconfident or underconfident predictions, resulting in suboptimal financial decisions (Ojeda et al., 2023; Wallace and Dahabreh, 2014).

Various post-hoc calibration methods have been suggested to tackle this problem, such as Platt scaling, isotonic regression, and Bayesian calibration techniques. These approaches adjust model outputs to improve the alignment between predicted and observed probabilities, thereby enhancing the reliability of machine learning models in financial risk contexts (Huang et al., 2020; Naeini et al., 2015). Furthermore, calibration becomes particularly challenging in imbalanced datasets, where minority-class probabilities are often systematically underestimated, necessitating specialized calibration strategies (Wallace and Dahabreh, 2012).

Although modern machine learning models, especially ensemble and boosting algorithms, show strong predictive accuracy, they are often criticized for being opaque. These "black-box" models hide how decisions are made, which limits their use in financial settings that require accountability and regulatory compliance.

To overcome this limitation, XAI has become a significant focus of research. These techniques aim to offer clear insights into how models behave, helping stakeholders comprehend how input features affect predictions. Notably, SHapley Additive exPlanations (SHAP) has become well-known for its strong theoretical basis and its capacity to deliver both overall and detailed interpretability (Černevičienė and Kabašinskas, 2024). SHAP allocates contribution values to each feature using cooperative game theory, guaranteeing consistent and fair feature attribution.

Besides SHAP, model-agnostic methods like Local Interpretable Model-agnostic Explanations (LIME) and Partial Dependence Plots (PDP) are also commonly used to improve interpretability. These methods offer complementary perspectives, allowing both instance-level and global-level explanations of model behavior. However, challenges remain in balancing interpretability and predictive performance, particularly on imbalanced datasets, where model outputs may be influenced by resampling or cost-sensitive strategies.

Overall, accurate financial distress prediction depends on a thorough evaluation framework that combines classification metrics, probability calibration, and explainability. This approach guarantees not just high predictive accuracy but also reliability and transparency, improving the practical use of machine learning models in financial decision-making.

Research Gap and Contribution

Although there have been significant progress in predicting financial distress, some important research gaps still exist.

First, prior studies predominantly focus on isolated aspects of model development, such as predictive accuracy, class-imbalance handling, or interpretability, without integrating these components into a unified analytical framework. This fragmented method reduces the usefulness of models in real-world financial decisions, where accuracy, reliability, and transparency need to be ensured all at once.

Second, many existing studies rely on limited datasets and lack rigorous validation across different modeling conditions. As a result, the generalizability and robustness of proposed models remain uncertain, particularly when applied to imbalanced financial datasets.

Third, evaluation practices in financial distress prediction are still inconsistent. Although metrics like accuracy and ROC–AUC are commonly reported, recent research indicates that these measures may not be enough in imbalanced scenarios, where recall, precision–recall balance, and probability calibration give more relevant evaluations of model effectiveness.

Although XAI techniques have received growing interest, their incorporation into predictive modeling frameworks is still limited. Only a few studies systematically merge high-performing machine learning models with interpretable methods to offer both global and local insights into financial distress prediction.

To overcome these limitations, this study introduces a comprehensive predictive framework that simultaneously includes:

1) Advanced machine learning algorithms
2) Resampling methods to handle class imbalance
3) SHAP-based explainability to enhance interpretability. Unlike prior studies, the proposed

framework evaluates multiple models within a unified experimental setting, enabling consistent performance comparisons across algorithms and data preprocessing strategies

Additionally, this research enhances the literature by showing that combining predictive accuracy, thorough evaluation, and explainability can strengthen the robustness and real-world applicability of financial distress prediction models. The results offer empirical support for creating more dependable and understandable decision-support tools for financial risk management.

This study's empirical results show that the integrated framework not only boosts predictive accuracy but also improves the reliability of detecting minority classes and makes model outputs more understandable. Specifically, models with resampling techniques achieve better recall for financially distressed firms, while SHAP-based analysis offers both overall and detailed explanations of key financial factors. These results address previous limitations, where accuracy, imbalance handling, and interpretability were often considered separately, highlighting the practical usefulness of this framework for financial risk decision-making.

Methods

Research Design

This study employs a quantitative research approach using supervised machine learning methods to create a predictive framework for financial distress. It is organized as a comparative experimental study, focusing on three main areas: Model comparison, class imbalance management, and Explainable AI (XAI) implementation.

The main goal is to assess the performance of chosen machine learning models under uniform experimental conditions, while tackling the challenges posed by imbalanced financial data. Beyond just accuracy, the study highlights the importance of model interpretability to ensure that the findings are both statistically sound and practically valuable for financial decisions.

Data Description

This study utilizes the publicly accessible Company Bankruptcy Prediction dataset sourced from Kaggle. It comprises 6,819 firm-level observations and 96 financial variables that capture multiple dimensions of corporate performance, including profitability, liquidity, leverage, and operational efficiency (FEDESORIANO, 2020).

The target variable is binary, called “Bankrupt?", with 1 signifying a financially distressed firm and 0 indicating a non-distressed firm. The dataset shows substantial class imbalance, as distressed firms are a minority. This mirrors real-world financial situations and makes it suitable for testing imbalance-aware machine learning methods.

Data Preprocessing

Data preprocessing aims to improve data quality and compatibility with machine learning algorithms. Missing values are addressed with median imputation, minimizing the effect of outliers and preserving dataset integrity. Then, feature scaling via standardization is used to normalize the data, helping models converge more effectively, especially those sensitive to feature scales.

All preprocessing steps are applied consistently across the dataset to ensure methodological reproducibility.

Feature Selection

A multi-stage feature selection process is used to cut down dimensionality and enhance model efficiency. It starts with a variance-threshold method that removes features with almost zero variance, since these variables offer little discriminative ability.

Next, mutual information evaluates how relevant each feature is to the target variable. This method captures both linear and nonlinear dependencies, allowing for the identification of the most informative predictors. The top-ranked features are retained for further analysis.

Finally, multicollinearity is managed through correlation-based filtering. Highly correlated feature pairs are detected, and redundant variables are discarded based on their importance. This step decreases redundancy and improves interpretability.

As a result, a final subset of 17 features is obtained, representing key financial indicators grounded in financial distress theory. These features are consistently used across all stages of modeling and interpretation.

Data Splitting and Validation Strategy

To ensure an unbiased assessment of the model's performance, the dataset is split into training and testing sets through stratified sampling. This method maintains the original class distribution in both subsets, which is essential for imbalanced classification tasks. A 70:30 split is used, with 70% of the data for training and 30% for testing, and a fixed random seed is applied to guarantee reproducibility.

Handling Class Imbalance

Due to the dataset's imbalance, SMOTE is used on the training data. It creates synthetic minority class samples by interpolating between existing observations, helping to better represent distressed firms. To avoid data leakage, SMOTE is only applied to the training set, leaving the testing set unaltered. This approach ensures that model evaluation accurately reflects real-world performance.

Machine Learning Models

This study uses three machine learning models to offer a detailed comparison across various levels of complexity and interpretability.

Logistic Regression serves as a baseline model because of its simplicity and capacity to provide probabilistic interpretations. It predicts the chance of financial distress using a linear combination of input features.

Random Forest is an ensemble learning technique that builds several decision trees and combines their predictions. It is especially good at modeling nonlinear relationships and minimizing overfitting.

XGBoost is an advanced gradient boosting algorithm that constructs models sequentially to reduce prediction errors. It is well-known for its excellent predictive accuracy and computational speed when working with structured data.

Model Evaluation

Model performance is assessed through a multi-metric approach to address the dataset's imbalanced nature. While accuracy serves as a general performance metric, more focus is placed on recall and F1-score, as these metrics better suit imbalanced classification scenarios.

Recall measures the model's capacity to accurately detect distressed firms and is crucial in financial risk scenarios where missing positives can be costly. Precision indicates how trustworthy positive predictions are, while the F1-score balances both precision and recall for a comprehensive evaluation. Moreover, the ROC-AUC metric assesses the model's ability to distinguish between classes at various thresholds.

Explainable Artificial Intelligence (XAI)

This study uses SHAP (SHapley Additive exPlanations), a technique based on cooperative game theory, to quantify how each feature affects the model's predictions, thereby improving interpretability.

SHAP analysis is conducted on the XGBoost model, which shows the best predictive performance among the tested algorithms. The TreeExplainer method is chosen for its efficiency and appropriateness for tree-based models.

To achieve a balance between computational efficiency and representativeness, a subset of 500 samples from the training dataset is selected as the background dataset for SHAP value estimation. This analysis offers both global and local interpretability through feature importance rankings, summary plots, dependence plots, and instance-level insights explanations.

Experimental Procedure

The experimental procedure follows a structured pipeline that begins with data collection and preprocessing, then proceeds to feature selection and dataset splitting. Class imbalance is addressed by applying SMOTE to the training data. Machine learning models are trained and assessed with various performance metrics. In the end, SHAP is used to interpret the model's predictions and pinpoint the main financial indicators linked to financial distress.

This systematic workflow ensures consistency, transparency, and reproducibility throughout the study.

Results

Overview of Data

This study employs a structured financial dataset aimed at predicting corporate financial distress. The dataset includes firm-level indicators that measure various aspects of financial health, such as profitability, liquidity, leverage, and operational efficiency. These features offer a well-rounded view of company behavior and are common in research on financial distress prediction.

Table 1 summarizes the key characteristics of the dataset used in this study to offer a clear overview.

According to Table 1, the dataset comprises 6,819 observations and 96 financial variables, all continuous and representing financial ratios. The absence of significant missing values indicates that the dataset is suitable for machine learning modeling without requiring extensive imputation. However, a notable characteristic of the dataset is class imbalance, which has important implications for predictive modeling and will be addressed in subsequent sections.

The study's dependent variable is a binary label, "Bankrupt?", with 1 representing a financially distressed firm and 0 indicating a non-distressed firm. Comprehending the distribution of this target variable is essential, as it directly influences the process of model training and evaluation.

To illustrate this distribution, Table 2 presents the dataset's class composition.

Table 2 illustrates that the dataset is highly imbalanced, with about 96.77% of observations belonging to non-distressed firms and just 3.23% to distressed ones. This stark imbalance poses a major challenge for classification models, which may tend to favor the majority class and thus struggle to accurately identify financially distressed firms.

From a modeling standpoint, this imbalance suggests that relying solely on accuracy is inadequate. A model might attain high accuracy by always predicting the majority class yet still miss minority-class cases. Consequently, specialized methods such as SMOTE are necessary to enhance the model's ability to detect financially distressed firms, as explained in Section 3.6 and tested in subsequent sections.

Table 1: Dataset Overview

Item	Description
Total observations	6,819
Number of features	96
Target variable	Bankrupt?
Type of variables	Continuous
Data type	Financial ratios
Missing values	None / Minimal
Class imbalance	Yes

Table 2: Distribution of the Target Variable Classes

Class	Count	Percentage
Non-bankrupt (0)	6,599	96.77
Bankrupt (1)	220	3.23

In addition to class imbalance, the dataset exhibits high dimensionality, with many financial variables relative to the number of observations. These variables often demonstrate non-normal distributions, including skewness and extreme values, which are typical characteristics of financial data. These properties emphasize the limitations of traditional statistical models and advocate for employing machine learning techniques capable of capturing nonlinear relationships.

Overall, the dataset offers a detailed and realistic view of financial conditions across different firms. Its high-dimensional nature, imbalance, and variety of financial indicators make it especially useful for testing advanced machine learning models and explainable AI methods in predicting financial distress.

Descriptive Statistics Analysis

To better understand the dataset's underlying structure, descriptive statistics are conducted on the selected financial variables. This analysis provides insights into the data's distributional properties, which are essential for selecting appropriate modeling techniques.

The dataset includes financial indicators representing profitability, liquidity, leverage, and operational efficiency. These variables are expected to exhibit heterogeneous distributions due to differences in firm characteristics and financial conditions. Table 3 summarizes the key descriptive statistics for the variables used in the model, highlighting their statistical properties.

The results in Table 3 reveal several important characteristics of the dataset. First, most financial variables exhibit noticeable deviations from normality, as indicated by high skewness and kurtosis values. Several variables, particularly those related to profitability and income measures, show strong positive skewness, suggesting that while most firms operate within a moderate performance range, a smaller subset experiences extreme financial conditions.

In addition, high kurtosis values indicate heavy-tailed distributions, implying that extreme observations, or outliers, are prevalent in the dataset. Such patterns are typical in financial data, where firm performance can vary significantly due to differences in size, leverage, and operational efficiency.

These findings have important implications for model selection. Traditional statistical models that depend on assumptions of normality and linearity may struggle in these situations. Conversely, machine learning algorithms, especially tree-based approaches, are more resilient to skewed data and outliers, thus better suited for predicting financial distress.

To further illustrate the distributional characteristics of key financial variables, Fig. 1 presents distribution plots of selected features.

Fig. 1 confirms the presence of non-normal distributions across multiple variables. Many features exhibit right-skewed patterns, with a large proportion of observations concentrated in the lower ranges and a long tail extending toward higher values. This indicates that only a small number of firms exhibit extreme financial performance, either positively or negatively.

Furthermore, the observed variability across variables suggests that financial distress is influenced by multiple factors rather than a single dominant indicator. This highlights the importance of developing models that can effectively represent complex interactions between variables.

Along with distributional analysis, it is essential to comprehend the relationships between variables to evaluate potential redundancy and multicollinearity. Fig. 2 shows the correlation heatmap of the chosen financial variables.

As shown in Fig. 2, most variables exhibit low to moderate correlations, indicating that multicollinearity is not severe across the dataset. This suggests that each feature contributes unique information, which is beneficial for machine learning modeling.

However, some clusters of moderately correlated variables are evident, particularly among profitability and leverage indicators. These relationships are theoretically consistent, as financial ratios often capture overlapping aspects of firm performance. Although these correlations are not a major concern for tree-based models, they emphasize the importance of selecting features carefully to minimize redundancy and enhance interpretability.

Finally, Fig. 3 shows the distribution of the target variable to highlight the imbalance issue.

Table 3: Overview of Key Financial Variables Using Descriptive Statistics

Variable	Mean	Std	Skew
Interest rate (after tax)	0.781	0.013	-53.20
Borrowing dependency	0.375	0.016	20.84
ROA	0.559	0.066	-1.03
Debt ratio	0.113	0.054	0.98
Persistent EPS	0.229	0.033	5.14
...	...	...	...

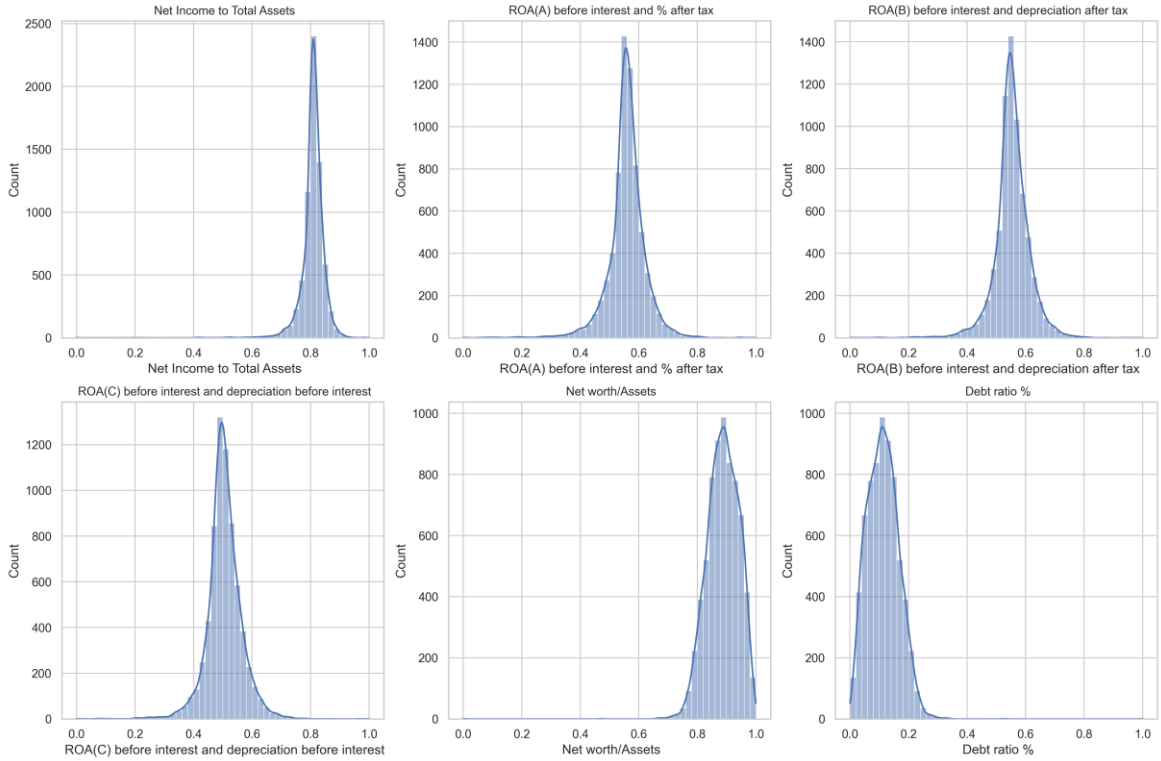


Fig. 1: Distribution of Selected Financial Variables

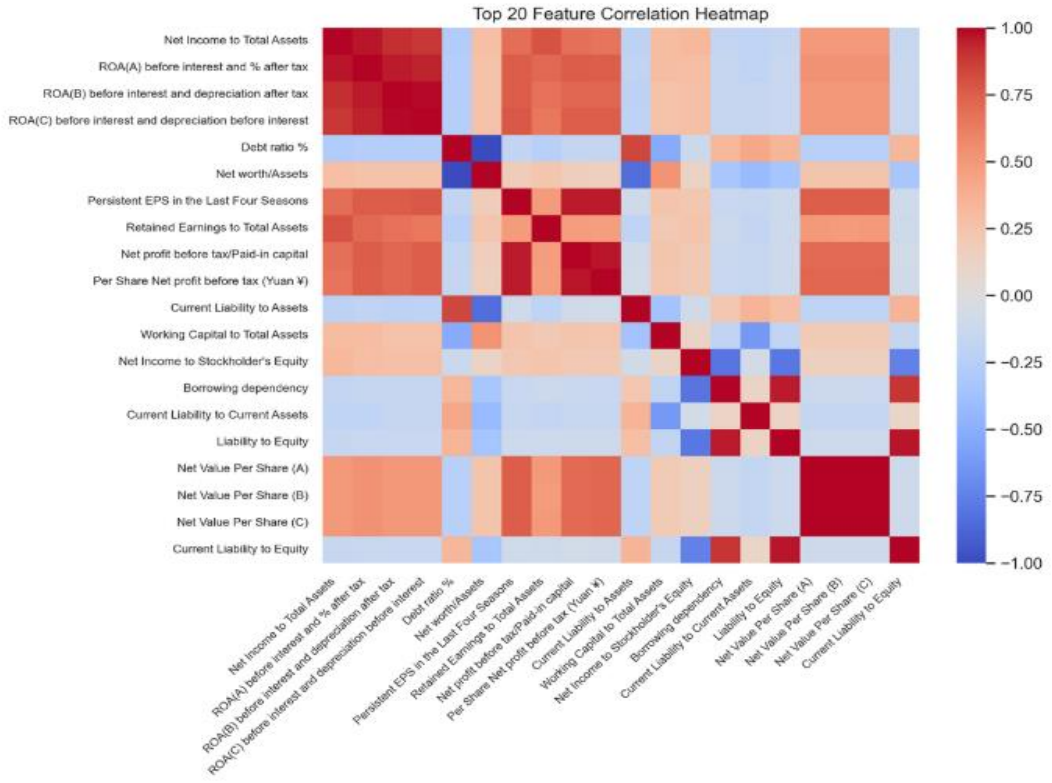


Fig. 2: Correlation Heatmap of Key Financial Variables

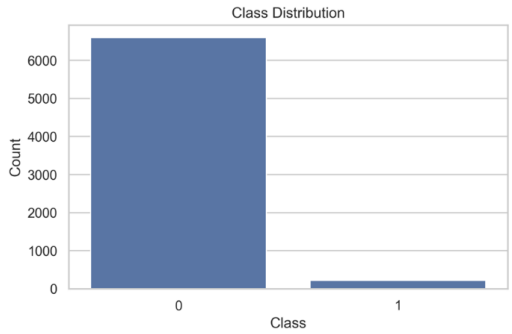


Fig. 3: Distribution of the Target Variable Classes

Fig. 3 clearly shows a highly imbalanced class distribution, with non-distressed firms making up the majority. This imbalance presents a significant challenge for predictive models, which may become biased toward the dominant class and thus struggle to accurately identify financially distressed firms.

From a practical perspective, this issue is particularly important because the minority class represents the primary focus of financial risk analysis. Therefore, addressing class imbalance is essential to ensure meaningful and reliable model performance. This challenge motivates the application of imbalance-handling techniques, such as SMOTE, which will be evaluated in subsequent sections.

Overall, the descriptive analysis reveals three main features of the dataset:

1) Non-normal, skewed distributions
2) Moderate correlations among features
3) Significant class imbalance. These insights justify the use of advanced machine learning methods and underpin the modeling and evaluation discussed in the upcoming section

Feature Selection Results

After the descriptive analysis, feature selection is performed to pinpoint the most informative variables for predicting financial distress. This process is crucial for decreasing dimensionality, removing redundant features, and enhancing both the model's predictive accuracy and interpretability.

Mutual information is used as the main criterion to assess feature relevance. Unlike linear correlation, it detects both linear and nonlinear relationships between input features and the target, which is especially useful for complex financial datasets.

Fig. 4 shows the most important features based on mutual information scores.

As shown in Fig. 4, the most influential features are primarily associated with profitability, earnings stability, and financial structure. Variables like Persistent EPS over the Last Four Seasons, Net Income to Stockholders' Equity, and Debt Ratio have the highest importance scores.

These findings suggest that financial distress is driven by a combination of operational performance and capital structure. Firms that have unstable earnings and carry high leverage are more prone to financial troubles.. This observation is consistent with established theories of financial distress, in which profitability and debt exposure are key determinants of firm solvency.

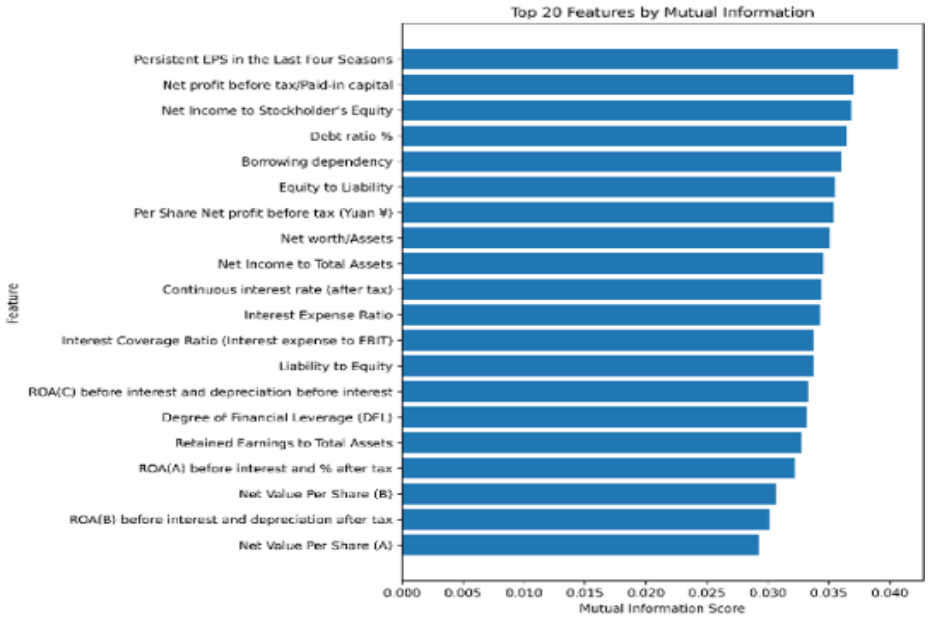


Fig. 4: Top Features Ranked by Mutual Information

In addition to feature relevance, it is important to assess potential redundancy among selected variables. Therefore, a correlation-based filtering process is applied to reduce multicollinearity and improve model stability.

To visualize the relationships among the selected features, Fig. 5 presents the correlation heatmap of the final feature set.

Fig. 5 shows that most feature pairs exhibit low-to-moderate correlations, indicating that redundancy has been effectively reduced through the feature selection process. Although some clusters of moderately correlated variables remain, particularly among profitability-related indicators, these relationships are theoretically expected and do not pose a significant concern for tree-based models.

The absence of widespread high correlation (e.g., above 0.90) suggests that each selected feature contributes distinct information to the predictive models. This variety improves the robustness of the modeling process and decreases the likelihood of overfitting.

Based on the multi-stage feature selection framework, a final subset of 17 features is retained for model development. These features capture key financial dimensions, including profitability, leverage, liquidity, and cost of capital.

Table 4 displays the final feature set and their importance scores to clearly summarize the selected variables.

Table 4 demonstrates that the chosen features are consistent with established theories in financial distress research. Profitability indicators, such as Net Income to Total Assets and Return on Assets, reflect a firm's operational efficiency, while leverage-related variables, including Debt Ratio and Borrowing Dependency, capture financial risk exposure.

The inclusion of both profitability and leverage indicators highlights the multidimensional nature of financial distress. Rather than being driven by a single factor, financial distress emerges from the interaction between declining performance and increasing financial obligations.

From a modeling perspective, reducing the feature set from 96 to 17 variables significantly improves computational efficiency and reduces model complexity. At the same time, retaining only the most informative features enhances interpretability, particularly in the subsequent SHAP analysis.

Overall, the feature selection results demonstrate that the proposed framework successfully identifies a compact and informative set of predictors. This enhanced feature set offers a solid basis for model development and helps ensure that future analysis concentrates on variables that are economically significant.

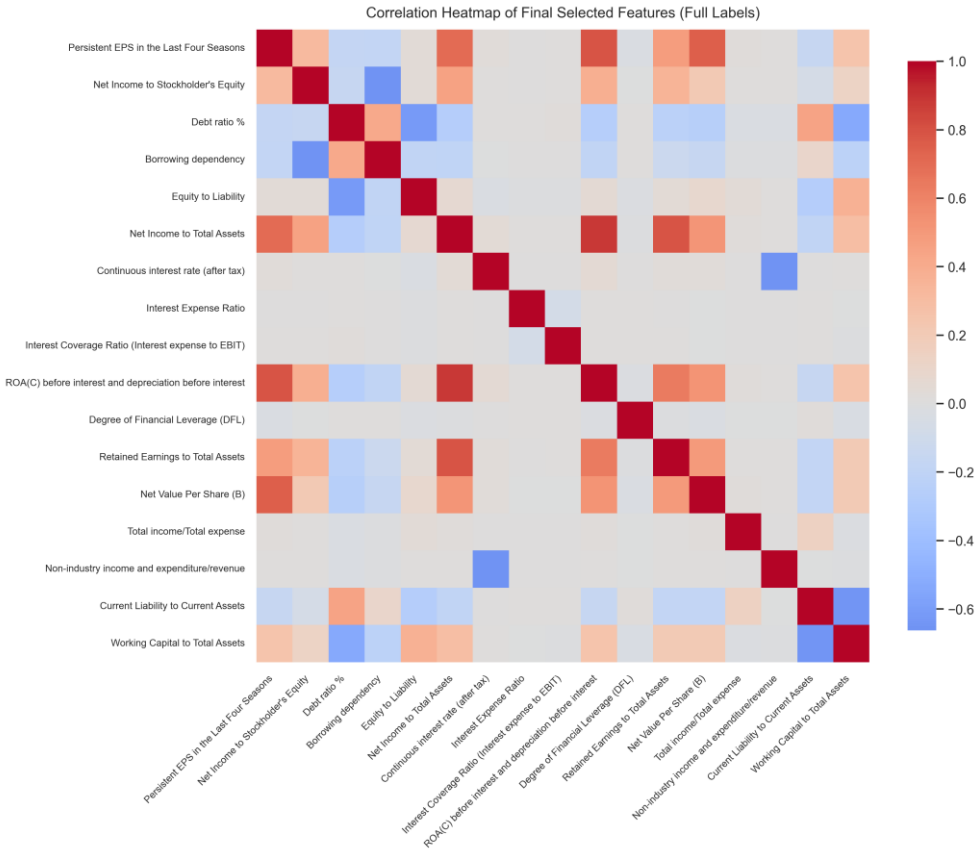


Fig. 5: Correlation Heatmap of Selected Features

Table 4: Selected Features and Their Importance Scores

Rank	Selected Feature	MI rank	MI Score
1	Persistent EPS in the Last Four Seasons	1	0.041
2	Net Income to Stockholder's Equity	3	0.037
3	Debt ratio %	4	0.036
4	Borrowing dependency	5	0.036
5	Equity to Liability	6	0.035
6	Net Income to Total Assets	9	0.035
7	Continuous interest rate (after tax)	10	0.034
8	Interest Expense Ratio	11	0.034
9	Interest Coverage Ratio (Interest expense to EBIT)	12	0.034
10	ROA(C) before interest and depreciation before interest	14	0.033
11	Degree of Financial Leverage (DFL)	15	0.033
12	Retained Earnings to Total Assets	16	0.033
13	Net Value Per Share (B)	18	0.031
14	Total income/Total expense	23	0.027
15	Non-industry income and expenditure/revenue	25	0.024
16	Current Liability to Current Assets	26	0.023
17	Working Capital to Total Assets	29	0.022

Model Performance

To assess the predictive effectiveness of the proposed models, a comparative analysis is performed utilizing various performance metrics such as accuracy, precision, recall, F1 Score, and ROC AUC. These metrics offer a thorough evaluation of model performance, especially when dealing with class imbalance.

Table 5 provides a summary of the overall classification performance by presenting the results of three machine learning models: Logistic Regression, Random Forest, and XGBoost.

As demonstrated in Table 5, all models demonstrate high accuracy and impressive ROC-AUC scores, reflecting solid overall classification effectiveness. Of these, XGBoost records the highest ROC-AUC, indicating its superior ability to differentiate between distressed and non-distressed firms.

However, accuracy alone is insufficient for evaluating imbalanced datasets. Consequently, recall and F1-score are also analyzed to measure the model's effectiveness in identifying financially distressed firms.

The results show that Random Forest offers a more balanced performance in terms of precision and recall, leading to a higher F1 score compared to the other models. This indicates that Random Forest provides a more dependable balance between detecting distressed firms and reducing false positives.

Fig. 6 shows the ROC curves for the three models to further demonstrate their performance at various classification thresholds.

Fig. 6 demonstrates that all models have strong discriminative power, with their curves nearing the top-left corner of the plot. XGBoost has the highest ROC curve, with Random Forest just behind, while Logistic Regression lags slightly.

Although ROC curves are useful, they can sometimes overestimate model performance in datasets with high class imbalance. Hence, Precision–Recall (PR) curves are also analyzed for a more accurate assessment.

As shown in Fig. 7, Random Forest achieves the best precision-recall balance, indicating superior performance in identifying minority-class instances. In contrast, although XGBoost performs well in ROC analysis, its performance in the PR space is slightly lower, indicating a trade-off between general discrimination and the ability to identify minority classes.

This finding highlights an important distinction between evaluation metrics. While XGBoost excels at overall classification performance, Random Forest is more effective at detecting financially distressed firms, a more essential goal in managing financial risk.

Overall, the results indicate that ensemble models surpass the linear baseline, highlighting the significance of nonlinear approaches in predicting financial distress. In particular, Random Forest is the most balanced model, while XGBoost has the highest overall discriminative power.

These findings provide a strong foundation for further analysis of class imbalance handling, which is examined in the next section.

Impact of Class Imbalance Handling

This section examines how applying the Synthetic Minority Over-sampling Technique (SMOTE) affects model performance, given the dataset's significant imbalance. The goal is to determine if addressing the imbalance enhances the model's accuracy in identifying financially distressed firms.

Table 5: Comparison of Machine Learning Model Performance

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC
XGBoost	0.941	0.304	0.636	0.412	0.935
Random Forest	0.952	0.343	0.545	0.421	0.935
Logistic Regression	0.869	0.174	0.818	0.286	0.913

To provide a baseline comparison, model performance without SMOTE is first examined. Table 6 summarizes the classification results of the three models before applying SMOTE.

As presented in Table 6, all models demonstrate high accuracy and robust ROC-AUC scores. Nonetheless, recall is consistently very low across every model, highlighting a significant bias toward the majority class. This imbalance leads to poor identification of financially distressed firms.

This limitation is practically significant because the main goal of predicting financial distress is to accurately identify high-risk firms, not just to attain high overall accuracy.

To further illustrate the models' behavior under class imbalance, Fig. 8 presents the confusion matrices for the models without SMOTE.

Fig. 8 clearly indicates that most distressed firms are misclassified as non-distressed. The number of false negatives far exceeds the number of true positives, indicating that the models struggle to identify minority-class cases.

To overcome this limitation, SMOTE is used on the training data. The models' performance following SMOTE application is summarized in Table 7.

Table 7 shows a significant enhancement in recall for all models. Logistic Regression shows the most significant increase, while Random Forest and XGBoost also demonstrate notable improvements.

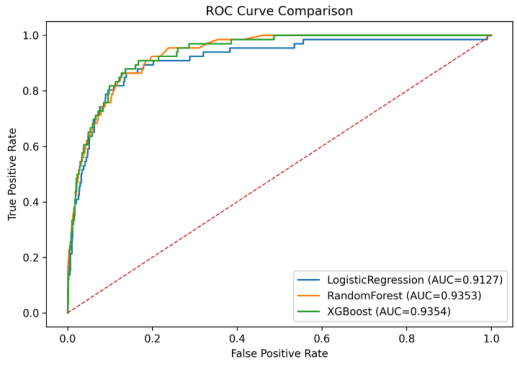


Fig. 6: ROC Curve Comparison of Machine Learning Models

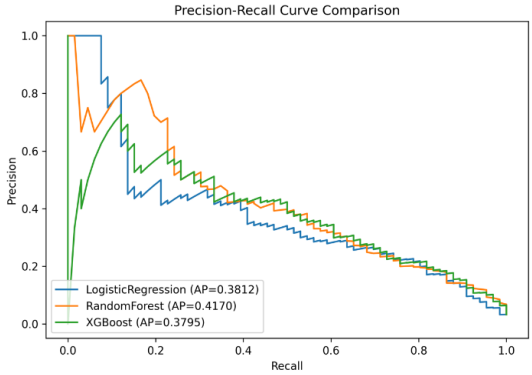
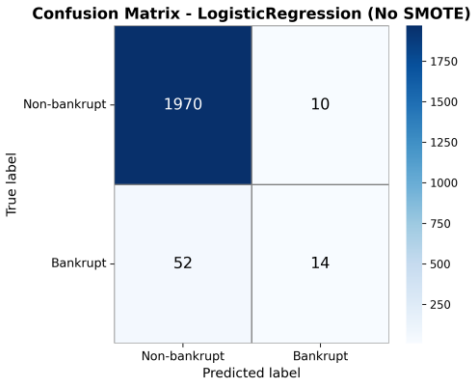
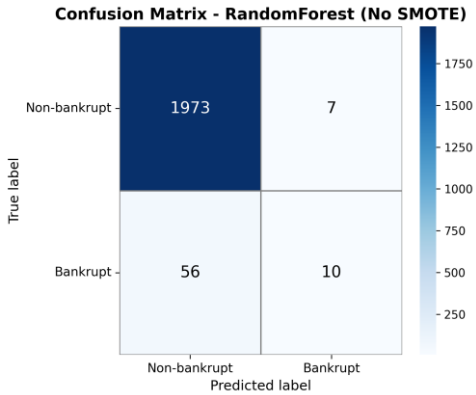


Fig. 7: Precision–Recall Curve Comparison

Table 6: Model Performance Without SMOTE

model	Accuracy	Precision	Recall	F1-score	ROC-AUC
XGBoost	0.968	0.500	0.167	0.250	0.940
Random Forest	0.969	0.588	0.152	0.241	0.921
Logistic Regression	0.970	0.583	0.212	0.311	0.918





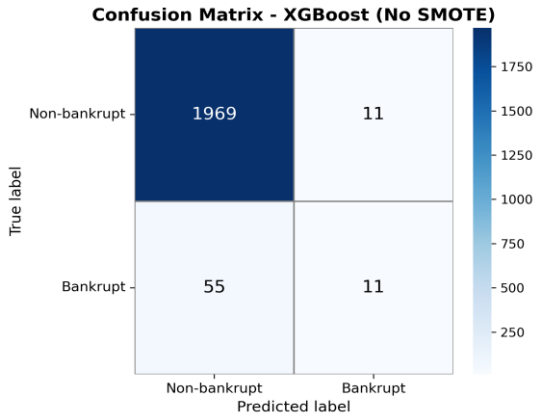


Fig. 8: Confusion Matrices Without SMOTE

Table 7: Model Performance With SMOTE

model	Accuracy	Precision	Recall	F1-score	ROC-AUC
XGBoost	0.941	0.304	0.636	0.412	0.935
Random Forest	0.952	0.343	0.545	0.421	0.935
Logistic Regression	0.869	0.174	0.818	0.286	0.913

Although the precision drops slightly because of more false positives, the overall F1-score improves, indicating a better balance between precision and recall. This trade-off is anticipated and acceptable in financial risk prediction, since missing a distressed firm is more costly than false alarms.

To further illustrate these changes, Fig. 9 presents the confusion matrices after applying SMOTE.

As shown in Fig. 9, the number of correctly identified distressed firms increases significantly after applying SMOTE. The decrease in false negatives highlights how effectively imbalance management enhances the model's sensitivity.

Overall, the comparison between Tables 6-7 and Figs. 8 and 9 confirms that SMOTE significantly improves model performance on imbalanced classification problems. The results show that without addressing imbalance, models might attain artificially high accuracy but still miss minority-class instances.

These findings underscore the importance of incorporating imbalance-aware techniques into financial distress prediction and provide a strong justification for using SMOTE in the proposed framework.

Threshold Optimization and Decision Effectiveness

Although Section 4.5 shows that SMOTE greatly enhances the model's capacity to identify financially distressed companies, the classification results still hinge on the decision threshold that translates predicted probabilities into class labels. The standard threshold of 0.50 might not be the best choice, especially in imbalanced classification scenarios.

To address this issue, threshold optimization is performed using the Precision–Recall (PR) curve. The optimal threshold is chosen by maximizing the F1-score, ensuring a balanced trade-off between precision and recall.

To summarize the optimal thresholds for each model, Table 8 presents the selected threshold values along with the corresponding performance metrics.

As shown in Table 8, the optimal thresholds differ across models, reflecting variations in probability calibration. Logistic Regression tends to require a lower threshold to achieve higher recall, whereas tree-based models tend to exhibit more balanced thresholds.

These results indicate that a fixed threshold (e.g., 0.50) may not be appropriate for all models. Instead, model-specific threshold tuning is necessary to achieve optimal classification performance.

To further illustrate the impact of threshold optimization, Fig. 10 presents the confusion matrices of the models under optimal thresholds.

Fig. 10 illustrates a more balanced distribution between true positives and false positives compared to the default threshold. Notably, the number of correctly identified distressed firms rises without significantly increasing false positives.

Compared to the results, threshold optimization further refines model performance by aligning predictions with the specific objective of financial distress detection. While SMOTE boosts the model's capacity to learn from imbalanced data, threshold tuning refines decision-making by choosing the optimal cutoff point.

Overall, integrating SMOTE with threshold optimization offers a more efficient approach to predicting financial distress. SMOTE enhances the

model's sensitivity, and threshold optimization ensures that predictions are suitable for real-world decision-making.

These findings show that a model's performance depends not only on the algorithm selected but also on the decision strategy used for interpreting its outputs.

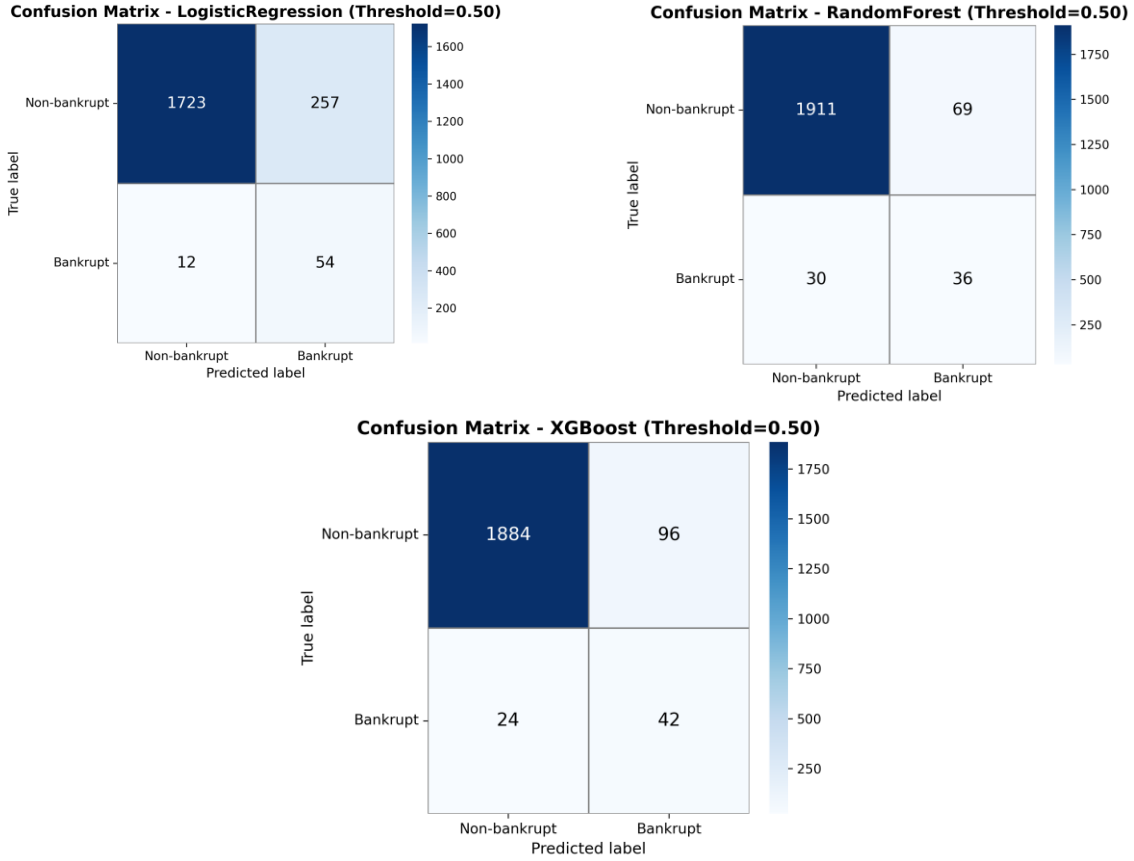
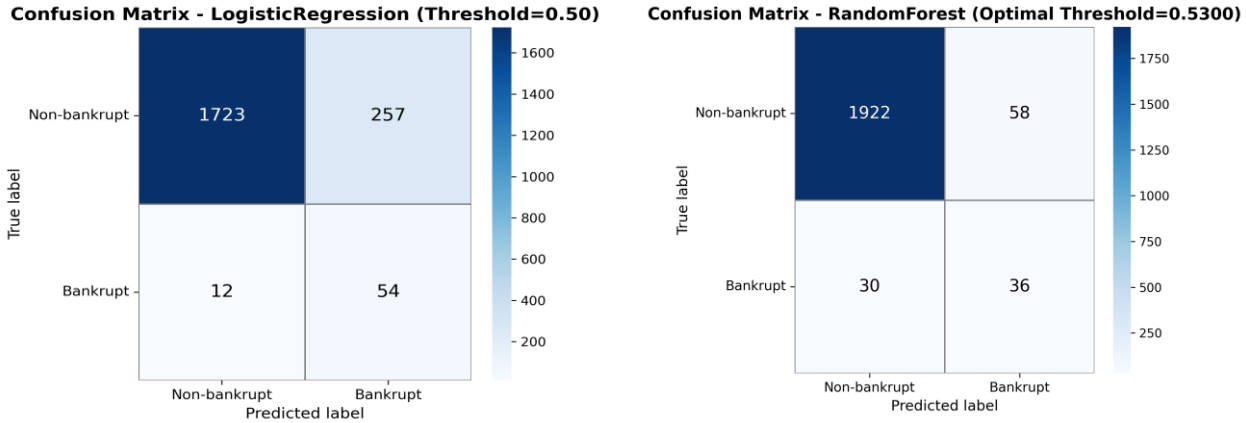


Fig. 9: Confusion Matrices With SMOTE

Table 8: Optimal Classification Thresholds and Performance Metrics

Model	Optimal Threshold	Accuracy	Precision	Recall	F1-score	ROC-AUC
Logistic Regression	0.94	0.963	0.426	0.394	0.409	0.913
Random Forest	0.53	0.957	0.383	0.545	0.45	0.935
XGBoost	0.786	0.962	0.423	0.5	0.458	0.935



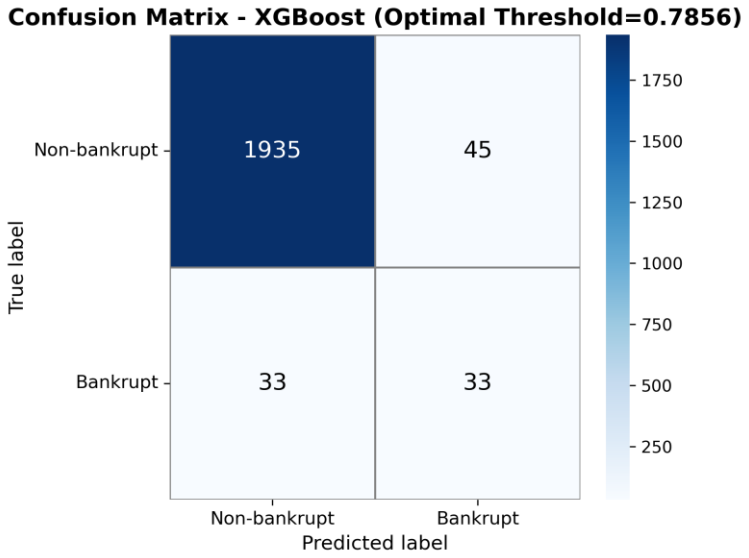


Fig. 10: Confusion Matrices Under Optimal Thresholds

Explainable Artificial Intelligence (SHAP Analysis)

Although earlier sections show that machine learning models can accurately forecast financial distress, understanding the underlying decision-making process remains essential for practical application. In financial contexts, model interpretability is particularly important, as stakeholders require transparency in identifying the key drivers of risk.

To address this need, SHapley Additive exPlanations (SHAP) is employed to interpret model predictions. SHAP offers both overall and detailed explanations by measuring how much each feature influences the prediction result.

Due to its superior discriminative performance, XGBoost is chosen for SHAP analysis based on the results in Section 4.4. The TreeExplainer method efficiently computes SHAP values for tree-based models.

Fig. 11 presents the SHAP summary plot, illustrating the significance of features based on their average impact on the model's predictions.

Fig. 11 shows that features related to profitability and financial structure have the greatest influence on the prediction of financial distress. Variables such as Persistent EPS (4Q), Net Income to Stockholders' Equity, and Debt Ratio consistently exhibit high SHAP values.

The distribution of SHAP values also reveals the direction of impact. Higher values of profitability-related features tend to reduce the probability of financial distress, while higher leverage-related variables increase the likelihood of distress. This pattern aligns with financial theory, in which declining profitability and rising debt burden are key indicators of financial risk.

To further examine the relationship between individual features and model output, Fig. 12 presents SHAP dependence plots for selected variables.

Fig. 12 illustrates how changes in individual feature values affect model predictions. For example, as Debt Ratio increases, the SHAP value also increases, indicating a higher probability of financial distress. Conversely, higher values of profitability indicators are associated with lower SHAP values, reflecting reduced financial risk.

These relationships highlight the nonlinear nature of financial distress prediction, where the impact of variables may vary across different ranges. This highlights the importance of employing machine learning models that can identify complex interactions.

In addition to global interpretation, SHAP can provide local explanations for individual predictions. Fig. 13 presents a representative SHAP waterfall plot for a single observation.

Fig. 13 illustrates how each feature influences a particular prediction: Positive SHAP values steer the prediction toward financial distress, whereas negative values diminish the risk. This level of interpretability is particularly useful for decision-making, as it allows analysts to understand the specific factors driving a model's output.

Overall, the SHAP analysis offers a detailed understanding of the model's behavior, connecting predictive performance to significant financial insights. The alignment between SHAP results and feature selection in Section 4.3 further confirms the robustness of the proposed framework. These findings emphasize that combining machine learning and XAI not only boosts predictive accuracy but also increases transparency and trust in financial distress prediction models.

DOI: 10.3844/jcssp.2026.2291.2311

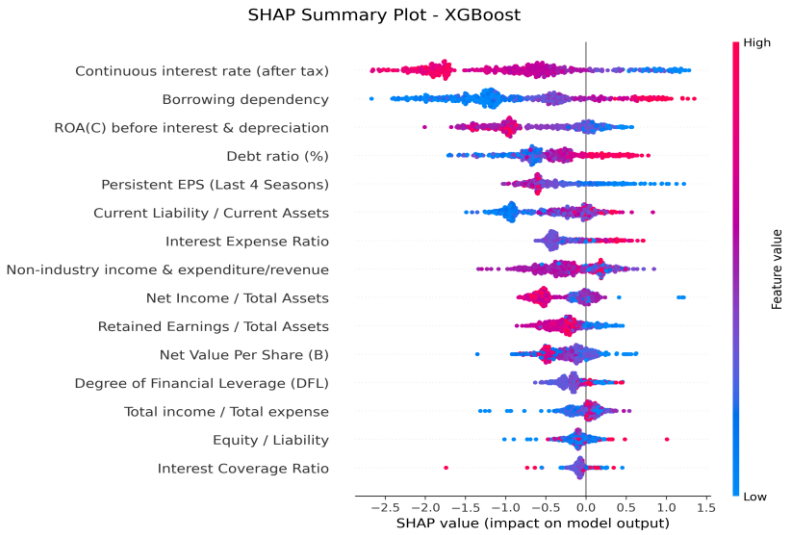


Fig. 11: SHAP Summary Plot Showing Feature Importance

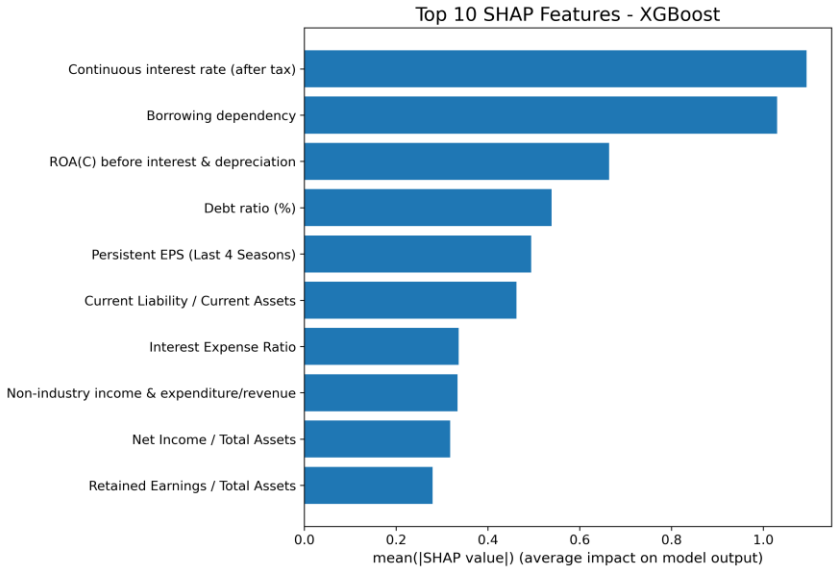


Fig. 12: SHAP Dependence Plots for Key Features

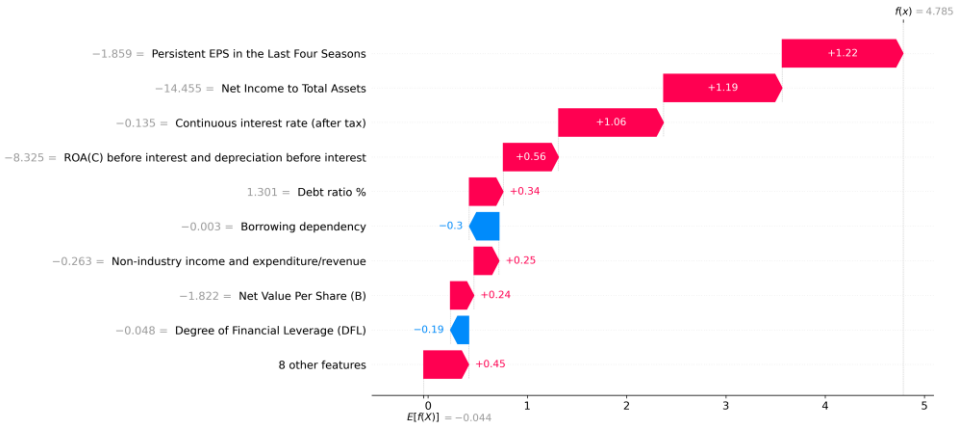


Fig. 13: SHAP Waterfall Plot for Individual Prediction

Discussion

Interpretation of Key Findings

This study offers a thorough assessment of machine learning methods for predicting financial distress, combining class imbalance solutions and explainable AI within a single analytical framework. The findings offer important theoretical and practical insights into the design and application of predictive models in risk-sensitive financial contexts.

First, the comparative analysis of model performance indicates that ensemble-based methods consistently outperform the linear baseline. XGBoost attains the highest ROC–AUC score, reflecting its strong discriminative ability. Nonetheless, under the Precision–Recall evaluation, Random Forest offers a more balanced performance, especially in detecting financially distressed companies. This divergence suggests that model superiority depends on the evaluation perspective, reinforcing prior evidence that ROC–AUC may overestimate performance on imbalanced classification problems (Davis and Goadrich, 2006; Saito and Rehmsmeier, 2015). Therefore, the results highlight the need to choose evaluation metrics that match the research goal, especially when correctly identifying minority-class instances is essential. Although recall has improved, it still plays a crucial role in predicting financial distress because missing distressed firms can cause significant financial and operational risks in practical scenarios.

Second, the results highlight the significant impact of class imbalance on predictive performance. Models trained without resampling methods show high overall accuracy but have very low recall for the minority class, meaning they often miss financially distressed firms. After applying SMOTE, recall improves substantially across all models, confirming that handling class imbalance is essential for meaningful financial distress prediction. This finding aligns with the broader literature, which shows that imbalanced datasets can lead to biased model training and skewed evaluation results. (Buda et al., 2018; He and Garcia, 2009). More importantly, the findings build on prior research by empirically demonstrating that improving the detection of minority classes enhances the real-world application of predictive models in managing financial risk, where misclassification costs are inherently unequal. While SMOTE helps improve minority-class detection, the synthetic samples it generates may not fully capture real-world financial distress scenarios, potentially introducing distributional bias into the model.

Third, the findings related to threshold optimization demonstrate that classification performance is highly sensitive to decision thresholds. The commonly used default threshold of 0.50 is shown to be suboptimal in imbalanced settings. By optimizing thresholds based on the F1 Score, the models achieve a better balance between precision and recall. This result highlights that predictive success depends not just on choosing the right algorithm but also on effective decision-level calibration. The study adds to existing research by showing how incorporating threshold optimization into predictive models emphasizes decision-making performance over solely model accuracy (Provost and Fawcett, 2013).

Ultimately, the explainability analysis using SHAP offers valuable insights into the key factors driving financial distress. The results consistently identify profitability and financial leverage indicators as the most influential predictors. Variables such as Persistent EPS, Net Income to Stockholders' Equity, and Debt Ratio contribute strongly across models. These findings are strongly aligned with financial theory, which posits that declining profitability and increasing leverage are key determinants of financial instability (Altman, 1968; Ohlson, 1980). Beyond confirming theoretical expectations, SHAP values enable both global and instance-level interpretations of model outputs (Lundberg and Lee, 2017), thereby enhancing transparency and supporting actionable decision-making. This integration of predictive performance and interpretability addresses a key limitation in prior studies, where high-performing models often lack sufficient explanatory power.

Taken together, these findings demonstrate that an integrated approach combining machine learning, class imbalance handling, threshold optimization, and XAI can significantly enhance both the accuracy and understanding of financial distress prediction models. This study, therefore, advances the literature by moving beyond isolated methodological improvements toward a holistic, decision-support-oriented framework that is more closely aligned with practical financial risk management applications.

Theoretical Implications

This study's findings enhance the financial distress prediction literature by improving both theoretical and methodological aspects of machine learning in financial analysis.

First, the study reinforces and extends the relevance of financial theory within data-driven predictive frameworks. The consistency between SHAP-based feature importance and traditional financial indicators, particularly profitability and leverage measures, demonstrates that machine learning models can capture economically meaningful relationships rather than serving as purely algorithmic "black boxes." This finding supports the theoretical validity of integrating statistical learning approaches with established financial distress theories, such as those grounded in insolvency risk and firm performance deterioration (Altman, 1968; Ohlson, 1980). More importantly, the findings indicate that XAI

techniques can connect predictive modeling with economic interpretability, thereby enhancing the theoretical basis for machine learning in finance.

Secondly, this research enhances the literature by redefining how model evaluation is approached within the context of imbalanced financial datasets. While prior research has predominantly relied on accuracy and ROC–AUC as primary evaluation metrics, the findings demonstrate that these measures may provide an incomplete or even misleading assessment of model effectiveness in distress prediction tasks. By emphasizing Precision Recall analysis and F1-score, this study aligns model evaluation with the asymmetric cost structure inherent in financial risk prediction. This perspective extends existing theoretical discussions on classification performance by highlighting the importance of context-specific evaluation frameworks, particularly in high-stakes decision environments (He and Garcia, 2009; Saito and Rehmsmeier, 2015).

Third, the study advances the methodological literature by proposing an integrated framework that combines feature selection, class imbalance handling, threshold optimization, and XAI within a unified modeling pipeline. Unlike prior studies that address these components in isolation, this research demonstrates that their joint application leads to more robust and interpretable predictive outcomes. From a theoretical standpoint, this integration shifts the focus from model-centric optimization to system-level design, in which predictive performance, interpretability, and decision effectiveness are jointly considered. This perspective contributes to the evolving paradigm of decision-oriented analytics, in which predictive models are evaluated not only on statistical performance but also on their ability to support actionable, reliable decision-making (Provost and Fawcett, 2013).

Finally, the findings contribute to the emerging discourse on XAI in financial applications. This study shows that SHAP offers both overall and detailed interpretability without sacrificing predictive accuracy. It reinforces the idea that transparency and accuracy can coexist in machine learning models. (Lundberg and Lee, 2017). This insight is especially crucial in financial settings, where understanding the model is key for meeting regulations, gaining stakeholder trust, and ensuring practical use.

Taken together, these contributions extend the theoretical landscape of financial distress prediction by integrating financial theory, machine learning, and XAI into a cohesive analytical framework. The study establishes a basis for future research to investigate theory-driven, interpretable, and decision-focused methods in financial risk modeling.

Practical Implications

Practically, this study offers useful insights for financial analysts, risk managers, and decision-makers working on financial risk assessment and early warning systems.

Initially, the findings indicate that organizations ought to implement machine learning models, especially ensemble methods like XGBoost and Random Forest, to improve the accuracy of predicting financial distress. Nevertheless, the results clearly show that choosing a model should not be based only on overall performance metrics like accuracy or ROC–AUC. Instead, practitioners should prioritize models that demonstrate strong performance in detecting minority-class instances, particularly financially distressed firms. In high-risk environments, the ability to identify potential cases of distress is more critical than achieving high overall classification accuracy.

Second, the study highlights the necessity of incorporating techniques to handle class imbalance into predictive modeling. Methods such as SMOTE significantly improve recall for distressed firms, thereby reducing the likelihood of false negatives. From a risk management perspective, failing to detect financially distressed entities can lead to substantial financial losses and misallocation of resources. Therefore, integrating resampling techniques into model development should be considered a standard practice in financial distress prediction systems.

Third, the findings underscore the importance of optimizing thresholds as a key decision-making component. Rather than relying on the default classification threshold, organizations should calibrate thresholds to their specific operational objectives and risk tolerance. For example, in conservative risk management, a lower threshold might be set to better detect possible warning signs, even if this results in more false alarms. This underscores the importance of aligning predictive models with organizational decision-making strategies rather than viewing them solely as technical instruments.

Fourth, incorporating XAI methods like SHAP is crucial for improving the transparency and user-friendliness of predictive models. Being able to interpret overall trends as well as specific predictions helps analysts grasp the main factors behind financial distress. This aids in making better-informed decisions and enhances communication with stakeholders such as management, auditors, and regulators. In practice, interpretability is vital for building trust and encouraging the acceptance of machine learning systems in finance.

The study emphasizes the need for a comprehensive, system-level strategy in predicting financial distress. Instead of using isolated methods, organizations should combine predictive modeling, imbalance correction, threshold tuning, and explainability into a single decision-support system. This integrated approach improves both prediction accuracy and operational efficiency, allowing organizations to proactively address financial risks. In practice, predictive models should be periodically retrained to account for changes in economic conditions and potential model drift, ensuring sustained

performance over time.

Integration of Findings

The findings of this study collectively suggest that an effective financial distress prediction framework requires integrating multiple complementary components. Machine learning models provide strong predictive performance, class-imbalance handling improves sensitivity to minority-class instances, threshold optimization enhances decision effectiveness, and XAI ensures interpretability and transparency.

Importantly, the results demonstrate that these components should not be treated independently. Instead, their combined application produces a more robust, reliable, and practically applicable predictive system. This integrated perspective shifts the focus from isolated model performance to comprehensive decision-support design, in which predictive accuracy, interpretability, and operational relevance are jointly optimized.

From a broader perspective, this integrated framework is a key contribution of the study, aligning methodological advancements with real-world decision-making requirements. It lays the groundwork for creating next-generation financial distress prediction systems that are both accurate and easy to interpret and act upon. As a result, the study helps connect advanced analytical methods with real-world financial risk management applications.

Although the dataset used in this study provides a realistic benchmark for financial distress prediction, it is based on Taiwanese firms, which may reflect country-specific financial structures, regulatory environments, and economic conditions. These contextual factors may influence both model performance and feature importance. Therefore, the findings should be interpreted with caution when applied to other regions or institutional settings.

Conclusion

This study creates and assesses an integrated machine-learning framework for predicting financial distress, combining feature selection, class imbalance management, threshold tuning, and explainable AI within a single analytical process. Using a structured financial dataset, it systematically compares Logistic Regression, Random Forest, and XGBoost models under uniform experimental conditions.

The findings demonstrate that ensemble-based models outperform the linear baseline in predictive performance. While XGBoost achieves the highest discriminative performance, Random Forest provides a more balanced and practically effective performance in identifying financially distressed firms. Importantly, the results confirm that class imbalance substantially affects model performance. Using SMOTE greatly enhances the identification of minority-class instances, highlighting the importance of handling data imbalance in financial risk prediction. In addition, threshold optimization is shown to play a critical role in aligning model outputs with decision-making objectives, emphasizing that predictive performance depends not only on model selection but also on decision-level calibration.

Beyond predictive accuracy, this study underscores the importance of interpretability in financial applications. The SHAP analysis consistently identifies key financial indicators, particularly those related to profitability and leverage, as major drivers of financial distress. This finding confirms that machine learning models can capture economically meaningful relationships while maintaining transparency. Consequently, the proposed framework demonstrates that predictive performance and interpretability can be achieved simultaneously, addressing a key limitation in prior research.

Overall, this study enhances the literature by developing an integrated, decision-focused framework for predicting financial distress. Unlike traditional approaches that focus on isolated methodological improvements, this research demonstrates that effective prediction requires aligning model performance, data characteristics, and decision strategies. This holistic approach shifts the emphasis from solely evaluating models to creating practical, interpretable decision-support tools for managing financial risk.

Although these contributions are valuable, certain limitations must be recognized. Primarily, the dataset originates from Taiwanese firms, potentially capturing country-specific economic conditions, regulatory environments, and accounting standards. Consequently, the applicability of these findings to other regions might be restricted. Additionally, employing cross-sectional data limits the capacity to observe how financial distress develops over time. Lastly, the study mainly emphasizes structured numerical data and overlooks qualitative or macroeconomic factors that could also impact financial results.

Future research should build on this by including datasets from multiple countries and industries to improve external validity. In addition, integrating longitudinal and time-series data would enable modeling temporal dependencies in financial distress processes. The inclusion of alternative data sources, such as Including textual disclosures and macroeconomic indicators could enhance both predictive accuracy and contextual relevance. Moreover, advanced modeling approaches, including deep learning and sequential models, offer promising avenues for capturing complex patterns in financial risk dynamics. Finally, the development of real-time, explainable decision-support systems represents an important direction for translating predictive analytics into actionable financial strategies.

Acknowledgment

The authors wish to convey their sincere appreciation to the Mahasarakham Business School, Mahasarakham University, for supporting the research infrastructure and technical resources necessary to complete this study.

Funding Information

This Research was Financially Supported by Mahasarakham Business School, Mahasarakham University.

Author's Contributions

Kanjana Hinthaw: Completed the literature review, gathered and prepared the dataset, developed machine learning models, and drafted the first version of the manuscript.

Warawut Narkbunnum: Designed the research framework, supervised model development and SHAP analysis, performed result interpretation, and revised the manuscript critically to ensure academic quality and integrity.

Ethics

The authors state that this study does not involve human participants, personal data, or unethical issues. They report no conflicts of interest. All data utilized are publicly accessible and properly cited.

References

Abrahamsen, N.-G. B., Nylén-Forthun, E., Møller, M., de Lange, P. E., & Risstad, M. (2024). Financial Distress Prediction in the Nordics: Early Warnings from Machine Learning Models. *Journal of Risk and Financial Management*, *17*(10), 432. https://doi.org/10.3390/jrfm17100432

Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, *23*(4), 589–609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

Altman, E. I. (2018). Applications of Distress Prediction Models: What Have We Learned After 50 Years from the Z-Score Models? *International Journal of Financial Studies*, *6*(3), 70. https://doi.org/10.3390/ijfs6030070

Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications*, *83*, 405–417. https://doi.org/10.1016/j.eswa.2017.04.006

Beaver, W. H. (1966). Financial Ratios As Predictors of Failure. *Journal of Accounting Research*, *4*, 71–111. https://doi.org/10.2307/2490171

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Buda, M., Maki, A., & Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks*, *106*, 249–259. https://doi.org/10.1016/j.neunet.2018.07.011

Černevičienė, J., & Kabašinskas, A. (2024). Explainable artificial intelligence (XAI) in finance: a systematic literature review. *Artificial Intelligence Review*, *57*(8), 216. https://doi.org/10.1007/s10462-024-10854-8

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, *16*, 321–357. https://doi.org/10.1613/jair.953

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Das, S., Nayak, S. P., Sahoo, B., & Nayak, S. C. (2024). Evaluating Ensemble Models on Imbalanced Data Sets: A Comparative Study across Varied Minority Class Ratios. *Proceedings of the 2024 International Conference on Emerging Systems and Intelligent Computing (ESIC)*, 774–779. https://doi.org/10.1109/esic60604.2024.10481583

Davis, J., & Goadrich, M. (2006). The relationship between Precision-Recall and ROC curves. *Proceedings of the 23rd International Conference on Machine Learning - ICML '06*, 233–240. https://doi.org/10.1145/1143844.1143874

El Madou, K., Marso, S., El Kharrim, M., & El Merouani, M. (2024). Evolutions in machine learning technology for financial distress prediction: A comprehensive review and comparative analysis. *Expert Systems*, *41*(2), e13485. https://doi.org/10.1111/exsy.13485

FEDESORIANO. (2020). Company Bankruptcy Prediction. *Kaggle Datasets*.

Gaudreault, J.-G., Branco, P., & Gama, J. (2021). An Analysis of Performance Metrics for Imbalanced Classification. *Discovery Science*, *12986*, 67–77. https://doi.org/10.1007/978-3-030-88942-5_6

Gupta, B. (2025). Ensemble Learning for Accurate Prediction of Corporate Financial Distress: A Feature Reduction Perspective. *Innovative Computing and Communications*, *1435*, 27–38. https://doi.org/10.1007/978-981-96-6906-6_3

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *The Journal of Machine Learning Research*, *3*, 1157–1182.

Habib, A., Costa, M. D., Huang, H. J., Bhuiyan, Md. B. U., & Sun, L. (2020). Determinants and consequences of financial distress: review of the empirical literature. *Accounting & Finance*, *60*(S1), 1023–1075. https://doi.org/10.1111/acfi.12400

He, H., & Garcia, E. A. (2009). Learning from Imbalanced Data. *IEEE Transactions on Knowledge and Data Engineering*, *21*(9), 1263–1284. https://doi.org/10.1109/tkde.2008.239

Huang, L., Zhao, J., Zhu, B., Chen, H., & Broucke, S. V. (2020). An Experimental Investigation of Calibration Techniques for Imbalanced Data. *IEEE Access*, *8*, 127343–127352. https://doi.org/10.1109/access.2020.3008150

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, *3*(4), 305–360. https://doi.org/10.1016/0304-405x(76)90026-x

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., Salim, D. F., & Ihsan, A. F. (2025). Financial distress prediction in emerging markets by using a machine learning approach for financial sustainability. *Discover Sustainability*, *6*(1), 1296. https://doi.org/10.1007/s43621-025-02199-1

Kuizinienė, D., & Krilavičius, T. (2024). Balancing Techniques for Advanced Financial Distress Detection Using Artificial Intelligence. *Electronics*, *13*(8), 1596. https://doi.org/10.3390/electronics13081596

Lessmann, S., Baesens, B., Seow, H.-V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research*, *247*(1), 124–136. https://doi.org/10.1016/j.ejor.2015.05.030

Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems*, 4766–4775. https://doi.org/10.48550/arXiv.1705.07874

Min, J. H., & Young-Chan, L. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Systems with Applications*, *28*(4), 603–614. https://doi.org/10.1016/j.eswa.2004.12.008

Naeini, M. P., Cooper, G. F., & Hauskrecht, M. (2015). Binary Classifier Calibration Using a Bayesian Non-Parametric Approach. *Proceedings of the 2015 SIAM International Conference on Data Mining*, 208–216. https://doi.org/10.1137/1.9781611974010.24

Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, *18*(1), 109. https://doi.org/10.2307/2490395

Ojeda, F. M., Jansen, M. L., Thiéry, A., Blankenberg, S., Weimar, C., Schmid, M., & Ziegler, A. (2023). Calibrating machine learning approaches for probability estimation: A comprehensive comparison. *Statistics in Medicine*, *42*(29), 5451–5478. https://doi.org/10.1002/sim.9921

Provost, Foster, & Fawcett, T. (2013). *Data Science for Business. What You Need to Know About Data Mining and Data-Analytic Thinking*.

Ribeiro, M., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?": Explaining the Predictions of Any Classifier. *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*, 97–101. https://doi.org/10.18653/v1/n16-3020

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, *1*(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

Saito, T., & Rehmsmeier, M. (2015). The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PLOS ONE*, *10*(3), e0118432. https://doi.org/10.1371/journal.pone.0118432

Song, Y., Jiang, M., Li, S., & Zhao, S. (2024). Class-imbalanced financial distress prediction with machine learning: Incorporating financial, management, textual, and social responsibility features into index system. *Journal of Forecasting*, *43*(3), 593–614. https://doi.org/10.1002/for.3050

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374. https://doi.org/10.1016/B978-0-12-214850-7.50025-5

Wallace, B. C., & Dahabreh, I. J. (2012). Class Probability Estimates are Unreliable for Imbalanced Data (and How to Fix Them). *Proceedings of the 2012 IEEE 12th International Conference on Data Mining*, 695–704. https://doi.org/10.1109/icdm.2012.115

Wallace, B. C., & Dahabreh, I. J. (2014). Improving class probability estimates for imbalanced data. *Knowledge and Information Systems*, *41*(1), 33–52. https://doi.org/10.1007/s10115-013-0670-6

Wang, S., & Chi, G. (2024). Cost-sensitive stacking ensemble learning for company financial distress prediction. *Expert Systems with Applications*, *255*, 124525. https://doi.org/10.1016/j.eswa.2024.124525

Zhao, J., Ouenniche, J., & De Smedt, J. (2024a). Survey, classification and critical analysis of the literature on corporate bankruptcy and financial distress prediction. *Machine Learning with Applications*, *15*, 100527. https://doi.org/10.1016/j.mlwa.2024.100527

Zhao, Z., Cui, T., Ding, S., Li, J., & Bellotti, A. G. (2024b). Resampling Techniques Study on Class Imbalance Problem in Credit Risk Prediction. *Mathematics*, *12*(5), 701. https://doi.org/10.3390/math12050701