Journal of Process Management and New Technologies
Vol. 14, Issue 1-2, 2026, pp. 82-103

A DECISION-SUPPORT FRAMEWORK FOR BANKRUPTCY PREDICTION USING EXPLAINABLE ENSEMBLE LEARNING

Zineb REDOUANE ALI[1*], Mohammed DEHANE[2]

[1]*Abdelhamid Mehri University, Constantine, Algeria, redouaneali.zineb@univ-constantine2.dz*
https://orcid.org/0009-0002-4003-4902

[2] *Abdelhamid Mehri University, Constantine, Algeria, mohammed.dehane@univ-constantine2.dz*
https://orcid.org/0000-0002-2252-6965

Abstract: *Bankruptcy prediction is a challenging problem in the field of risk management of financial assets due to the rarity of bankruptcy events, the class imbalance that results from such rarity, and the regulatory requirements regarding the interpretability of AI-based decision systems. Given the gradual development of bankruptcy, it is necessary to use AI-based models that can capture non-linear relationships among financial metrics and detect early signs of issues in a company's financial health. These requirements suggest using AI models beyond linear models and financial metrics alone. In this article, a Stacking Ensemble model is developed with both linear and non-linear models in order to investigate the ability of the models to predict bankruptcy with an emphasis on analyzing prediction trade-offs under severe class imbalance, as well as utilizing methods from the field of Explainable Artificial Intelligence (XAI) to investigate the models within the ensemble framework to determine the reasons for the ensemble's performance on the evaluation metrics. Results indicate that the model has good discriminatory power, but is conservative in its decisions to recognize financial distress within companies. However, the requirement for model interpretability is still met, and the model's performance across different evaluation thresholds is considered in the article.*

Keywords: *Bankruptcy prediction, Explainable Artificial Intelligence (XAI), Financial risk modeling, Imbalanced data, SHAP, Stacking Ensemble.*

Original scientific paper
Received: 06.04.2026
Accepted: 06.06.2026
Available online: 14.06.2026
DOI: 10.5937/jpmnt14-67029

1. Introduction

Bankruptcy prediction is vital for the effective management of credit risks and managerial decisions. Interest in bankruptcy prediction research has increased significantly over the past two decades following several major financial crises, such as the 2008 global financial crisis. The complexity of modern financial systems has also revealed the limitations of using several traditional statistical models for financial monitoring and managerial decision-making. Financial distress in a given corporation results from several causes and impacts a company's

---

[*]Corresponding author

financial performance over time. Thus, a bankruptcy prediction system should support managerial decision-making in monitoring financial risks and improving financial performance (Zhao et al., 2024; Gabrielli et al., 2026).

Bankruptcy prediction using machine learning involves a significant class imbalance between companies that will go through bankruptcy and those that will not. Bankruptcy can be represented as a binary problem: a company will either go through bankruptcy or not. However, the complex decline in a company's financial performance before it goes bankrupt can be challenging to represent as a binary outcome. In the case of a failing company, it might experience declining profits, increasing total debt and decreasing liquidity and operational efficiency (Papík & Papíková, 2025). The rarity of companies going through bankruptcy can result in an imbalanced dataset, with significantly fewer bankrupt companies than non-bankrupt ones.

Although many methods have been used to successfully predict bankruptcy using financial ratios, not all are equally effective at capturing the non-linear relationships among financial ratios that may indicate financial distress in a company. Financial ratios provide a good measure of a company's financial health. However, the non-linear relationships among the various ratios indicate the complex nature of a company's financial performance. Thus, other methods that can outperform financial ratio analysis for predicting a company's financial distress have been studied in the past, including machine learning and ensemble learning methods (Zhao et al., 2024; Gabrielli et al., 2026).

Stacking ensemble learning is one such method that has been studied for its applicability to bankruptcy prediction. In this method, a series of different classifiers are trained in an ensemble to learn the distinguishing features of financially distressed companies. The predictions of these classifiers are used as input to another classifier that learns from them to predict whether a company is experiencing financial distress. This method has proven effective for financial data containing several types of linear and non-linear relationships among a company's features (Muslim et al., 2024; Cao et al., 2026). However, the predictions that are obtained from these models must be understandable to financial decision-makers for them to be of any practical benefit.

Artificial intelligence-assisted methods for financial risk management have significantly developed in recent years. These methods aim to produce financial risk prediction models that are understandable and accountable for their decisions. For risk managers, the outputs of a bankruptcy prediction model must be interpretable so they can understand the reasons behind the company's risk classification. Explainable AI techniques, such as SHAP-based interpretations, can significantly aid risk managers in making such predictions. Such explanations of the model outputs are vital for financial governance and AI-assisted management of financial risks (Akter, 2026).

Despite the substantial literature on using machine learning methods to predict bankruptcy, three crucial research gaps remain. First, most existing studies focus on the accuracy score of the bankruptcy prediction models. However, due to the class imbalance in the bankruptcy prediction problem, a model's accuracy could be high yet still fail to identify the number of financially distressed companies. Second, although ensemble models have been shown to have high discriminative power in classifying financially distressed companies, there remains a need to study the usefulness of these models for financial decision-makers. Finally, most studies separate the issue of model-based bankruptcy prediction from model explainability. Financial decision-makers need a model that can both detect signals of financial distress in a company and explain them. These research gaps indicate a need for AI-assisted methods for bankruptcy prediction that address class imbalance in the dataset to improve model performance, go beyond accuracy to evaluate model performance, and use explainable AI techniques to improve the transparency of the model's predictions.

The purpose of this paper is to address these research gaps. Specifically, this paper proposes a decision-support framework for bankruptcy prediction using explainable ensemble learning models. More specifically, a stacking-based ensemble model will be developed using linear and non-linear classifiers to learn the financial signals of companies that will go bankrupt. Furthermore, the evaluation of the developed model will go beyond accuracy to consider other metrics that provide a better understanding of the model's financial prediction performance for decision-makers. Finally, the SHAP-based explainability framework will be used to provide financial decision-makers with an understanding of the reasons behind the model's predictions.

The contribution of this research is threefold. First, an explainable ensemble learning model that can perform bankruptcy prediction in an imbalanced dataset will be developed. Second, the model's performance will be evaluated in light of the limitations of the accuracy score and the importance of other performance metrics for financial decision-makers. Third, the link between explainable AI techniques and financial risk governance will be made explicit in the study's findings.

This paper is organized as follows: following the Introduction, Section 2 presents the conceptual framework for the bankruptcy prediction problem in the context of class imbalance. Section 3 presents the dataset, its preprocessing, the models to be compared, the ensemble learning model framework, and the models' performance and explainability measures. Section 4 presents the model results and their analysis. Section 5 presents the results of the explainability framework analysis. Section 6 discusses the implications of the findings. Finally, Section 7 presents the study's conclusions and suggests avenues for future research on the topic.

2. Conceptual and analytical framework

2.1. Bankruptcy prediction under data imbalance

Bankruptcy prediction is inherently characterized by severe class imbalance in financial datasets, reflecting the rarity of corporate failure events in real-world economic systems. As a result, the minority class of bankrupt firms is inherently rarer than the majority class of surviving firms (Wang & Liu, 2021). Furthermore, many features of failing firms are similar to those of viable firms experiencing financial distress. Thus, not only are there few instances of bankruptcy, but those that do occur are often challenging to model because they are similar to other non-bankrupt firms (Ainan et al., 2024; Zhao et al., 2024). As a result, it is not a simple task to classify firms into two groups based on whether they will experience bankruptcy, but rather to recognize deviations within a space of mostly viable firms.

Traditional statistical models may struggle to adequately capture the complexity of bankruptcy risk, especially given the severe class imbalance in firms that experience bankruptcy. For example, linear models may fail to provide an adequate representation of bankruptcy risk, which is unevenly distributed across firms, and may also fail to account for interactions among those firms' financial metrics. While nonlinear models may provide a better understanding of firms' complex behavior, a trade-off must be made between increased sensitivity to bankruptcies and a decrease in the number of false alarms. Due to the class imbalance problem, it is important to evaluate the bankruptcy prediction model using several different performance measures (recall, precision, F1-score, AUC, specificity, and class-sensitive evaluation procedures) to provide a proper understanding of its performance (Gabrielli et al., 2026; Papík & Papíková, 2025; Zhao et al., 2024). Accuracy alone, as a performance measure, may indicate that the model can appropriately categorize firms as non-bankrupt, but it provides no indication of the model's performance in detecting firms that may go bankrupt.

The balance between avoiding false alarms and missed bankruptcies is inherently a decision problem, as are most financial risk problems. The costs of false-negative and false-positive results for a bankruptcy prediction model could be significant for financial managers. Consequently, the value of a bankruptcy prediction model depends not only on its statistical performance but also on how its errors can be minimized and how its outputs can be incorporated into a financial risk monitoring process. Thus, an appropriate bankruptcy prediction model should take into account the rarity of bankruptcies, the model's performance measures, and how the model's outputs may be most useful to managers making financial risk-monitoring decisions.

2.2. Financial ratios as a cumulative diagnostic tool

Financial ratios are used in bankruptcy estimation because they allow individuals to compare the financial metrics of companies of different sizes. Financial ratios allow individuals to assess a company’s finances by examining its internal financial relationships. Various models based on financial ratios, such as the Z-score model, use different financial indicators to represent a company's economic conditions (Altman et al., 2019; Garcia, 2022). Financial ratios and accounting-based indicators are commonly employed in machine learning frameworks for predicting bankruptcy because these metrics convey information about a company's financial health and stability (Zhao et al., 2024; Gabrielli et al., 2026).Companies typically experience a gradual deterioration in their financial condition: a slow decline in their profit margins, an increase in their leverage, a decrease in their liquidity, and a gradual decline in their operational efficiency.

The financial ratios used in this analysis will be organized into five economic dimensions of the company that indicate its potential for financial imbalance. These five dimensions include the company’s liquidity, leverage, profitability, operational efficiency, and cost structure. Each of these dimensions relies on the others to indicate potential financial imbalances for the company. For instance, a company’s level of financial imbalance will have different implications depending on its profitability. An increase in a company’s leverage may be sustainable if it has strong cash flow. These financial ratios are inherently linked to one another. Thus, none of these ratios can be represented by a linear model. The use of machine learning models with nonlinear representations is therefore helpful in predicting a company's bankruptcy due to their ability to account for relationships among financial ratios, especially in the case of imbalanced data (Zhao et al., 2024; Gabrielli et al., 2026).

Finally, using financial ratios to indicate potential imbalances in a company is also useful, as these ratios are easily understood by both company personnel and regulators. The financial ratios will therefore be used to create a model for detecting financial imbalances and to explain it. Thus, the creation of such a model and an explanation of its implications for companies with financial imbalances will help develop a method for detecting these imbalances and explaining their causes in the subsequent sections of this paper.

3. Methodological framework and data preparation

3.1. Data description and statistical properties of financial ratios

3.1.1. Data source and sample construction

The dataset used in this paper is the American Companies Bankruptcy Prediction Dataset, which consists of financial statement information for U.S. publicly listed companies (NYSE and NASDAQ) from 1999 to 2018. This data set contains 78,682 observations from 8,262 companies. Only financial statements that were complete and contained data for all financial statement ratios were included. The financial statement ratios considered in this paper are 18. The bankruptcy status of these companies was determined from United States bankruptcy court filings under Chapter 11 or Chapter 7. The year a company files for bankruptcy is referred to as the year of bankruptcy; however, to avoid potential distortions from that filing, that year is excluded from the data set. The year prior to the company filing for bankruptcy is labeled as the year the company went bankrupt, and all other years with no bankruptcy court filings are labeled as non-bankrupt years.

The descriptive statistics of the financial ratios exhibit skewness and outliers in their distributions (see Table 1). The correlations among financial ratios differ across the various groups of financial statement ratios. For instance, the ratios of similar financial statement categories exhibit higher correlations with each other than those of different financial statement categories. Additionally, financial leverage ratios exhibit a higher correlation with the ratios of company financial performance than other categories of financial statement ratios.

Table 1. Descriptive statistical analysis of financial ratios

	Count	Mean	Std	Min	25%	50%	75%	max
current_ratio	78682.0	3.491615	88.748344	-12.721311	1.158053	1.897068	3.193725	24107.975892
quick_ratio	78682.0	2.988059	88.742275	-12.983607	0.828205	1.385923	2.523794	24107.975892
wc_to_assets	78682.0	-0.991275	41.529737	-6322.993677	0.036494	0.209823	0.421787	0.996997
debt_to_assets	78682.0	1.923343	46.441877	0.000041	0.302422	0.507032	0.710560	6323.993676
ltd_to_assets	78682.0	0.266361	11.829234	-0.034901	0.000000	0.101972	0.289877	3286.996713
cl_to_assets	78682.0	1.504376	41.537449	0.000041	0.134449	0.218226	0.355895	6323.993676
net_margin	78682.0	-11.483459	279.731844	-32544.967455	-0.174473	0.015439	0.070352	1542.564491
gross_margin	78682.0	-6.320311	208.696160	-29325.697067	0.201613	0.349857	0.539486	358.210545
ebitda_margin	78682.0	-8.676248	214.708004	-29325.697067	-0.044176	0.081687	0.171586	358.210545
ebit_margin	78682.0	-9.049965	219.541027	-30175.696982	-0.108286	0.042613	0.116947	394.473705
roa	78682.0	-0.690219	13.653509	-1514.998485	-0.140486	0.017422	0.066491	244.833313
ebitda_to_assets	78682.0	-0.377226	9.166990	-1223.998776	-0.038916	0.087668	0.149726	35.916664
asset_turnover	78682.0	1.177673	3.240846	-31.586933	0.477908	0.889543	1.458643	419.999580
revenue_to_asset s	78682.0	1.177673	3.240846	-31.586933	0.477908	0.889543	1.458643	419.999580
inventory_turno ver	58981.0	32.857318	470.769402	-53.368996	3.072675	5.475836	13.296107	66510.933489
receivables_turn over	76477.0	21.557225	261.393910	-159.374696	4.805600	6.718381	10.475412	47245.952754
opex_ratio	78682.0	9.676250	214.708004	-357.210545	0.828418	0.918313	1.044176	29326.697067
cogs_ratio	78682.0	7.320312	208.696159	-357.210545	0.460514	0.650143	0.798398	29326.697067

Source: Prepared by the Authors

In the context of financial analysis, multicollinearity is not regarded as a statistical defect to be eliminated prior to modeling. The effects of multicollinearity can be assessed within the modeling framework itself, thereby avoiding the information loss that would result from dimensionality-reduction techniques. Beyond their potential statistical role, financial ratios also contribute to the model's interpretability. Practitioners' familiarity with the ratios enables the model's results to be translated into rationales for the failures of the firms analyzed. This element of interpretability is especially important in the analysis of groups of companies within different sectors of the economy. Finally, because financial ratios can be understood as indicators of changes in a firm's risk from one year to the next, each ratio can be used to create features for modeling risk (see Figure 1).

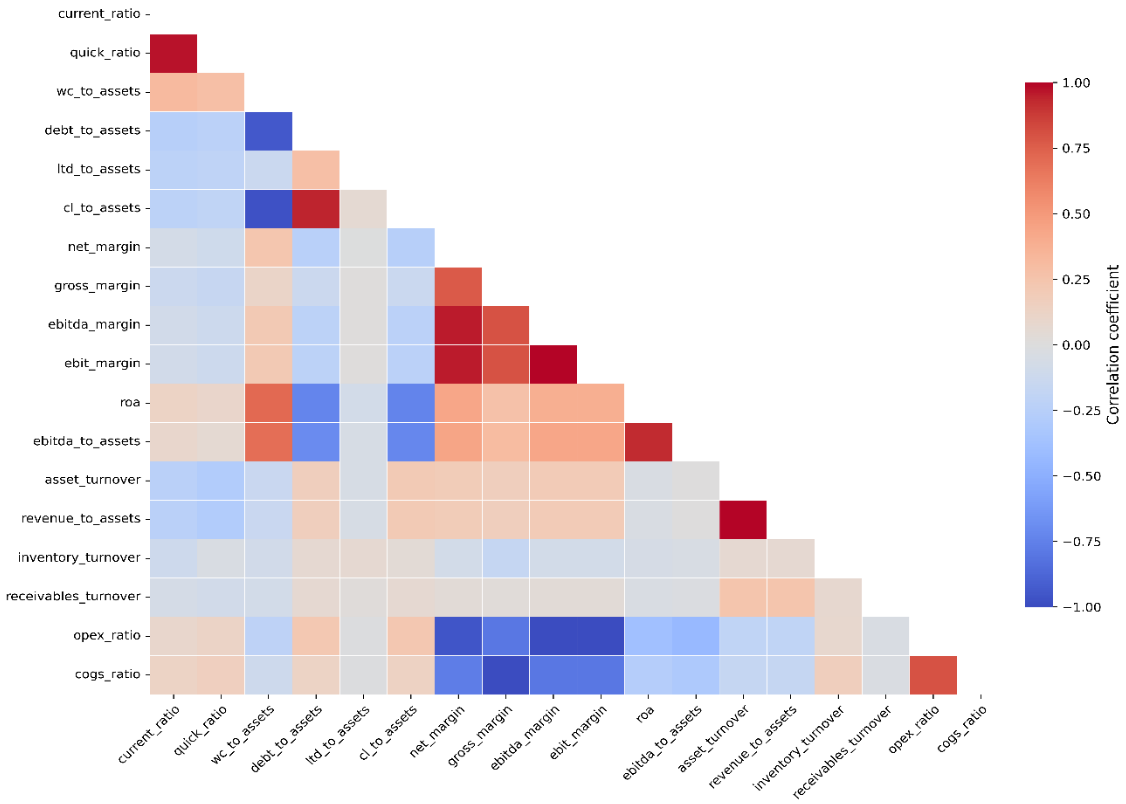


Figure 1. Correlation coefficients among financial ratios
Source: Prepared by the Authors

From the recent literature on the subject, it is evident that financial ratio analysis remains effective even within the context of machine learning models (Park et al., 2021).

3.1.2. Experimental design and data preprocessing

Data preprocessing extends beyond ensuring statistical coherence. The presence of missing data or outliers can introduce bias into the models, making the results less credible and less interpretable. Thus, preprocessing should be performed in a way that specifically considers the properties of the bankruptcy data. Such a strategy includes preprocessing for missing data and outliers, and splitting the data by time to create training and testing sets that mirror how the model would be used in the real world.

3.1.3. Missing values

A descriptive analysis of the financial ratio data indicates a high degree of data completeness. 16 of the 18 financial ratios contained no missing values in the data set. Table 2 reveals that the two ratio variables with missing observations were inventory and receivables turnover. These two ratio variables are not relevant to all firms, especially those in the service and technology industries. Thus, treating these missing values as a distinct outcome rather than as incomplete data points is more appropriate than imputing them.

Table 2. Missing values in financial ratios

current_ratio	0
quick_ratio	0
wc_to_assets	0
debt_to_assets	0
ltd_to_assets	0
cl_to_assets	0
net_margin	0
gross_margin	0
ebitda_margin	0
ebit_margin	0
roa	0
ebitda_to_assets	0
asset_turnover	0
revenue_to_assets	0
inventory_turnover	19701
receivables_turnover	2205
opex_ratio	0
cogs_ratio	0
dtype: int64	

Source: Prepared by the Authors

3.1.4. Outliers and distributional heterogeneity

As presented in Table 3, the descriptive statistics of the financial ratios exhibit pronounced skewness and long-tailed distributions. Prior to data preprocessing, the standard deviations of various financial ratios were 470.77 for inventory turnover, 279.73 for net margin, and 261.39 for receivables turnover. These high standard deviations indicate a high level of dispersion in these data sets, resulting from the diversity of companies in the sample and their financial conditions. The extreme observations within these data sets could have potentially led to misleading analyses of these financial ratios. Thus, winsorization was applied to these datasets to reduce the influence of extreme observations. Following winsorization, the standard deviations of inventory turnover, net margin, and receivables turnover were reduced to 46.02, 13.97, and 29.45, respectively. Furthermore, the standard deviations of EBIT margin and gross margin decreased from 214.51 and 208.70 to 12.14 and 5.98, respectively (see Table 3). These changes in the standard deviations of the financial ratios indicate that the data were not preprocessed to smooth them or to create an even distribution. Instead, the goal of preprocessing this data set was to mitigate the influence of extreme observations on the analyses of these financial ratios.

Table 3. Treatment of outliers

	Before_std	**After_std**
Inventory_turnover	470.769402	46.015648
Net_margin	279.731844	13.972159
Receivables_turnover	261.393910	29.450676
Ebit_margin	219.541027	12.137874
Ebitda_margin	214.708004	11.627784
Opex_ratio	214.708004	11.627783
Gross_margin	208.696160	5.976846
Cogs_ratio	208.696159	5.976845
Current_ratio	88.748344	3.017509
Quick_ratio	88.742275	2.847616

Source: Prepared by the Authors

3.1.5. Temporal splitting and the realism of the evaluation environment

In this study, the data is partitioned more explicitly by time. Table 4 shows that the period between 1999 and 2011 was used to train the model (55,927 observations), the years between 2012 and 2014 were used to validate the model (10,473 observations), and the remaining years between 2015 and 2018 were used to test the model out of sample (12,282 observations). This type of partitioning helps to avoid look-ahead bias and mimics the conditions under which the model would be used in practice.

Table 4. Temporal partitioning of the dataset

Train years	1999 - 2011
Val years	2012 - 2014
Test years	2015 - 2018
Shapes	
X_train	(55927, 18) y_train: (55927,)
X_val	(10473, 18) y_val : (10473,)
X_test	(12282, 18) y_test : (12282,)

Source: Prepared by the Authors

3.1.6. Class imbalance and temporal variation in bankruptcy rates

The results confirm the rarity of bankruptcy cases across all time periods, consistent with the inherent class imbalance in financial distress datasets, yet also demonstrate differences in bankruptcy rates among some subsets of the dataset. The bankruptcy rate is approximately 7.94% in the training set, decreases to 4.69% in the validation set, and further decreases to approximately 2.34% in the out-of-sample test set. These variations may reflect temporal or sampling-related fluctuations in bankruptcy occurrences; however, they do not provide sufficient evidence to establish cyclical behavior in bankruptcy rates conclusively. (see Table 5).

Table 5. Class imbalance across temporal subsets

Train bankruptcy rate	0.07942496468610867
Val bankruptcy rate	0.046882459658168625
Test bankruptcy rate	0.023367529718286924

Source: Prepared by the Authors

3.2. Benchmark models

In order to evaluate the gains of each of these stacking methods against meaningful benchmarks, the following classifiers were utilized in the study:

3.2.1. Logistic regression (lr)

LR is included as one of the baseline methods that have historically been used for credit risk and failure prediction, mainly owing to its interpretability and the structure of its model probabilities. The LR model can provide a standard against which the improvements of the non-linear models can be assessed.

3.2.2. Random forest (RF)

RF, based on the bagging idea, captures non-linearities and higher-order interactions while controlling model variance. RF provides a robust non-linear baseline to the modeling space, acting as a foundation for increasingly complex ensemble methods (Hamdi et al., 2024).

3.2.3. Extreme gradient boosting (XGBoost)

XGBoost is another model that has performed well in various benchmarks and is thus included in this study (Ben Jabeur et al., 2022). XGBoost is one of the strongest models among the stacking models.

All of the models were trained using the same temporal split and pre-processing methods. The models were also tuned during the validation period to identify the best hyperparameters within a specified search space, thereby helping to avoid overfitting.

3.3. Stacking ensemble architecture

The ensemble to be employed is a stacking ensemble that integrates diverse base learners (with varying abilities) under conditions of class imbalance and feature overlap. The base models will generate predicted probabilities for each target class during training; these probabilities will serve as input to a second-layer model. The second-layer model will be a parsimonious logistic regression model. This model is intentionally not complex; its role is to combine predictions from the various base models rather than to introduce complex modeling. Thus, the model reshapes the individual models' decisions by incorporating their diverse strengths, rather than simply maximizing classification accuracy.

3.4. Evaluation metrics and error structure

Given the rare-event nature of bankruptcy, the paper presents standard measures of model performance for completeness. However, these measures alone are insufficient to provide a measure of the model’s performance on the task of bankruptcy prediction. Beyond accuracy, the evaluation framework incorporates metrics that are more sensitive to class imbalance and more aligned with financial risk detection objectives:

- The precision-recall curve and the area under the precision-recall curve (AUPRC)
- The recall and sensitivity of the model for the bankruptcy class
- A measure of the type of failures that the model makes through the use of confusion matrices (specifically false positives and false negatives)

While the evaluation of models usually maximizes accuracy, in the specific case of bankruptcy prediction, the type of error in the characterization of probabilities might be more relevant to the bank and its managers than accuracy itself

3.5. Explainability framework (SHAP)

Explainability is incorporated to ensure that the predictive performance of non-linear and ensemble models remains interpretable within a financial decision-support and risk governance context. SHAP is used to analyze the model at both global and local levels to determine the relative importance of features in predicting bankruptcy at the model- and instance-specific levels, respectively. As the ensemble model is composite, SHAP analysis will be applied to the dominant non-linear model (XGBoost) to explain its predictions. In contrast, the behavior of the stacking ensemble will be analyzed separately. This strategy avoids the ambiguity of applying SHAP analysis to an ensemble model while still allowing interpretability in the financial context. The interpretability of the model is assessed by examining the consistency of the most influential features with established financial theory on bankruptcy risk and financial distress.

4. Empirical results

4.1. Evaluation framework under class imbalance

Given the class imbalance and cost asymmetry inherent in bankruptcy prediction, overall accuracy and other summary performance metrics can mask critical distinctions between models. In response, the evaluation of the models incorporates a variety of performance measures and metrics related to the models' failure to identify firms that will subsequently go bankrupt. Results for the three benchmark models (logistic regression, Random Forest, and XGBoost) are reported in Table 6. Each of these metrics is examined in this paper in relation to the implications of the model's decisions, rather than in relation to one another, and not as a simple ranking of the models according to their performance on each measure.

Table 6. Comparative performance results across models

	Accuracy	Precision	Recall	F1	AUC
Logistic Regression	0.526706	0.034378	0.710801	0.065584	0.678421
Random Forest	0.869891	0.070164	0.372822	0.118102	0.750550
XGBoost	0.719997	0.052527	0.644599	0.097138	0.730466
Stacking Ensemble	0.972806	0.216867	0.062718	0.097297	0.799421

Source: Prepared by the Authors

4.2. Benchmark model performance

4.2.1. Overall accuracy

As shown in Table 6, there are substantial differences in model accuracy. The Stacking Ensemble model exhibits the highest accuracy at 0.972806, followed by the Random Forest model at 0.869891, the XGBoost model at 0.719997, and the Logistic Regression model at 0.526706. At first glance, the Stacking Ensemble model seems to be the best for predicting bankruptcy. However, given the nature of the data (bankruptcy cases being the minority class), accuracy alone is not a sufficient measure of model effectiveness. The high accuracy of the Stacking Ensemble model is likely due to its ability to correctly classify most non-bankrupt companies. Accuracy alone, therefore, is not a sufficient statistic for measuring model performance; additional metrics must also be calculated and considered when evaluating these models.

4.2.2. Sensitivity to bankruptcy events (recall)

Recall provides a more direct indication of the model's ability to detect actual bankruptcy cases. Logistic Regression has the highest value of recall at 0.710801, followed by XGBoost at 0.644599 and Random Forest at 0.372822. Logistic Regression and XGBoost models indicate that they are more sensitive to the cases of companies that will become bankrupt.

The Stacking Ensemble model, however, has a recall value of 0.062718. This indicates that the model detects only a small proportion of actual bankruptcy cases at the default threshold. Thus, despite its high accuracy value, the ensemble model is very conservative in its predictions of bankruptcy cases.

4.2.3. Warning reliability (precision)

Precision reflects the reliability of the bankruptcy warnings generated by each model. The Stacking Ensemble model yields the highest precision at 0.216867, followed by the Random Forest model at 0.070164, the XGBoost model at 0.052527, and the Logistic Regression model at 0.034378. Thus, the Stacking Ensemble model produces relatively more reliable warnings than the other models.

Despite the higher precision of the ensemble methods, the precision of all models remains relatively low. Thus, the rarity of bankruptcy cases presents a challenge for predicting bankruptcy. Furthermore, while the Logistic Regression and XGBoost models exhibit higher recall, their low precision indicates that many of their predictions also generate false-positive warnings about potential bankruptcies.

4.2.4. Composite and ranking metrics

The F1-score, which balances recall and precision, yields the highest value for the Random Forest model among the benchmark models, at 0.118102, followed by XGBoost at 0.097138. In contrast, Logistic Regression records the lowest F1-score, at 0.065584. The Stacking Ensemble records an F1-score of 0.097297, which is very close to XGBoost but lower than Random Forest. This indicates that the ensemble model's higher precision is offset by its very low recall.

Furthermore, the ROC-AUC values indicate differences in the models' ranking capacity. The Stacking Ensemble achieves the highest AUC value, at 0.799421, followed by Random Forest at 0.750550, XGBoost at 0.730466, and Logistic Regression at 0.678421. Thus, the nonlinear and ensemble-based models outperform Logistic Regression in ranking firms by their probability of bankruptcy. However, these results also confirm that stronger ranking capacity does not necessarily imply better detection of bankruptcy cases at the default decision threshold.

The results of each model indicate that no single model is best for identifying firms at risk of bankruptcy. Logistic Regression and XGBoost are more sensitive to bankruptcy events, Random Forest provides a better balance between precision and recall, and the Stacking Ensemble offers stronger ranking capacity and more reliable warnings, but with very limited recall. Therefore, each model has strengths and weaknesses depending on the decision context in which it is used.

4.3. Performance of the stacking ensemble

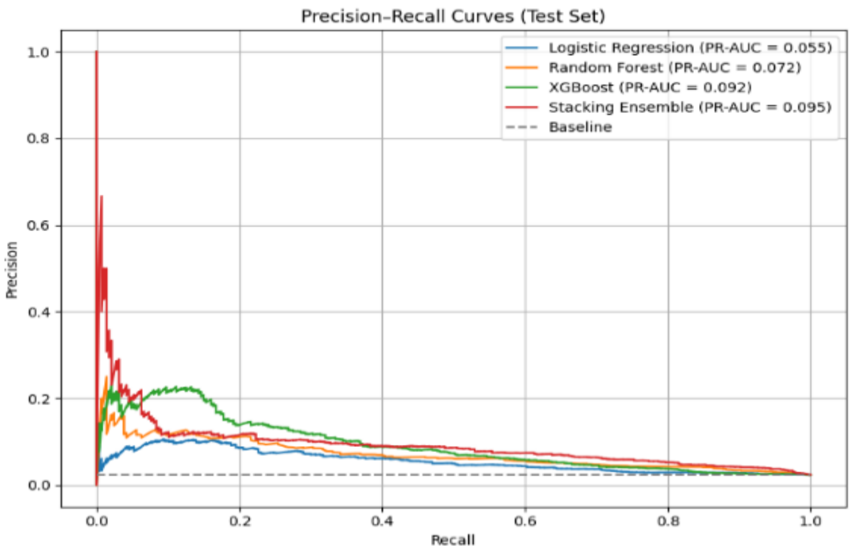
Table 6 presents the out-of-sample performance results for the Stacking Ensemble. The model achieves the highest accuracy of 0.972806 and the highest ROC-AUC of 0.799421. However, the model also exhibits a low recall of 0.062718 under severe class imbalance. While the model's high accuracy suggests it is generally good at predicting bankruptcy, its very low recall indicates it also fails to recognize most companies that go bankrupt. The model does, however, have the highest precision among the tested models, indicating that when it determines that a company is going to go bankrupt, it is correct the most often among all the models. However, the model's high precision at the cost of low recall suggests it is generally not very useful for detecting bankrupt companies.

Therefore, the Stacking Ensemble is not the best model to use for bankruptcy prediction. Its advantages are its high ranking and warning reliability, while its disadvantages include weak bankruptcy detection capability at the default threshold. Thus, threshold calibration is necessary before using the ensemble model to determine if a company is likely to go bankrupt.

4.4. Error trade-offs and threshold behavior

The precision-recall curve, however, provides more insight into each model's performance under class imbalance. The ensemble model achieves the highest area under the precision-recall curve, at 0.095, compared with the XGBoost model (0.092), the Random Forest model (0.072), and the logistic regression model (0.055) (see Figure 2).

Figure 2. Precision–Recall

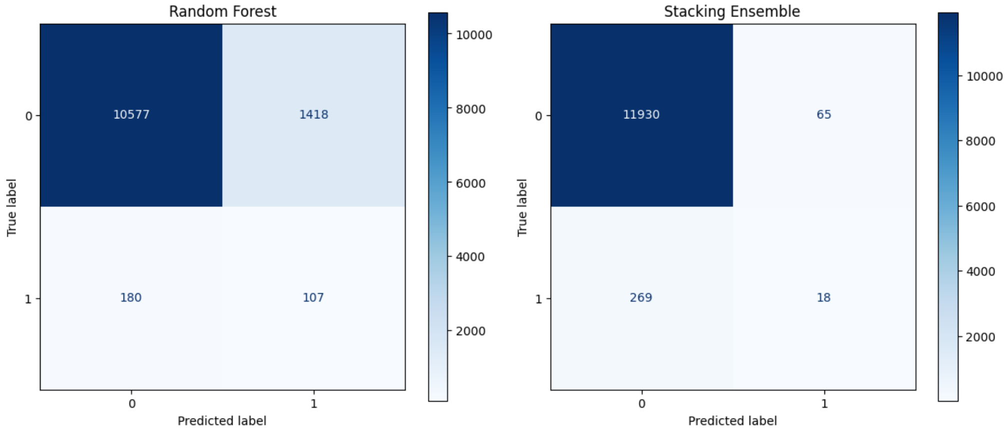


Source: Prepared by the Authors

4.5. Confusion matrix analysis

The confusion matrices further clarify the differences between the two models. The Random Forest correctly identifies 107 bankruptcies, misses 180 bankrupt firms, and incorrectly flags 1,418 non-bankrupt firms as bankrupt. By contrast, the Stacking Ensemble correctly identifies only 18 bankruptcies, misses 269 bankrupt firms, and incorrectly flags 65 non-bankrupt firms as bankrupt. Thus, the ensemble model is less likely to indicate bankruptcy for companies that do not face bankruptcy as a threat. However, it also catches fewer instances of bankruptcy than the Random Forest model (see Figure 3).

Figure 3. Confusion matrices for the Random Forest and Stacking Ensemble models

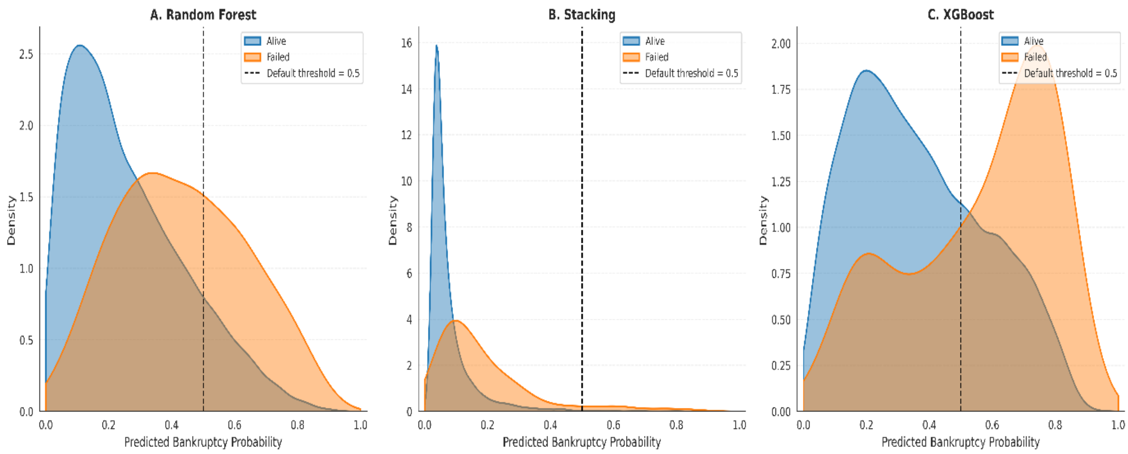


Source: Prepared by the Authors

4.6. Threshold sensitivity and statistical significance

Threshold sensitivity analysis shows that the model's conservative behavior is not intrinsic but depends on the chosen decision rule. Lowering the threshold for the model's predictions indicates increases in the model's recall and F1-score. Furthermore, when examining the predicted probabilities that each instance is classified as positive (see Figure 4), the ensemble tends to produce low positive-class probabilities for most instances.

Figure 4. Predicted probability distributions at a decision threshold of 0.5



Source: Prepared by the Authors

The differences between the predictions of the Random Forest and Stacking ensemble methods are statistically significant (McNemar test) (see Table 7).

Table 7. Statistical significance of model differences

McNemar statistic	1073.4650067294751
P-value	1.932760367626285e-235

Source: Prepared by the Authors

4.7. Summary of empirical findings

The results of the experiments enable comparison of the various bankruptcy prediction methods in terms of their trade-offs. The logistic regression model is shown to be more sensitive in detecting bankruptcy cases within the minority class; the tree methods provide a better balance between sensitivity and specificity; and the ensemble methods achieve higher overall accuracy in identifying bankrupt companies, but at a lower rate for those within the minority class. Overall, each of these methods has distinct strengths and weaknesses relative to the others, providing a solid foundation for analyzing their interpretability in the following section.

5. Empirical interpretability results

The following section presents the results of an interpretability analysis of the models. As presented in the previous section of this report, each of the examined models exhibits distinct levels of predictive performance and error structures. Each of these factors, however, does not help to explain the actual methods by which the models are able to recognize the indicators of potential financial distress for those companies, or how the results of those models can be meaningfully applied to the review of those financial reports by the company's managers. Thus, interpretability is an essential component of any financial decision-making process and its associated models (Babaei & Giudici, 2025; Akter, 2026).

These aspects of interpretability become even more important in predicting bankruptcy for those companies. In general, Explainable AI systems are often associated with improved financial transparency, accountability, compliance with financial regulations, and governance of those models and decisions (Akter, 2026). Thus, the interpretability of the AI system used in this analysis is based on both the use of financial indicators and ratios and the application of explainable AI techniques to explain those indicators and their contributions to bankruptcy risk predictions for those companies.

To meet these interpretability requirements, this report employs various explainable artificial intelligence techniques. More specifically, SHAP analyses will be performed to determine how each financial ratio contributed to bankruptcy predictions for those companies. SHAP analyses are often applied to bankruptcy and financial risk problems, as they help explain how each financial ratio increases or decreases the overall risk of bankruptcy for those companies (Ye et al., 2025). The ensemble model will be analyzed for interpretability, in part because of its many components. More specifically, an analysis of the XGBoost component of the ensemble will help to improve the explainability of the ensemble model (Balasubramaniam et al., 2023) (see Figure 5).

The ensemble model will be analyzed at both global and local levels. Analysis of the model at the global level will help determine which financial ratios contribute most to predicting bankruptcy for companies in the analyzed dataset. Local model analysis will help determine the contribution of each financial indicator to the risk of bankruptcy for one of the companies with a high predicted risk of bankruptcy. Thus, the ensemble model will help reveal not just the model's transparency but also the usefulness of its outputs for companies' financial monitoring, managerial review, and risk governance processes.

5.1. Global interpretative analysis (Global SHAP)

5.1.1. Long-term debt to assets (ltd_to_assets)

The Mean (|SHAP|) results show that the ltd_to_assets variable has the strongest influence on the model's outputs, with values nearly double those of the other variables in the model. Furthermore, the beeswarm plot shows that high values of the ltd_to_assets variable are associated with positive SHAP values, with some of the highest values approaching +1.0. These results show that ltd_to_assets is one of the main factors the model considers when predicting the likelihood of bankruptcy for a given firm; high levels of this ratio indicate that the firm is weak in meeting its financial obligations.

5.1.2. Profitability as a balancing mechanism: gross_margin, ROA, and EBIT_margin

The profitability indicators, such as gross_margin, ROA, and EBIT_margin, exhibit opposing effects. The higher the values of these variables, the lower their negative SHAP values. The importance of this result lies in the balancing effect seen in these profitability variables. While a company's leverage is an important factor in assessing the potential for bankruptcy, its profitability is also a key factor in assessing the likelihood of going out of business. The fact that these two variables interact to influence the company's likelihood of bankruptcy underscores the multidimensional nature of bankruptcy and the value of AI-assisted financial analysis for managers seeking to understand such processes.

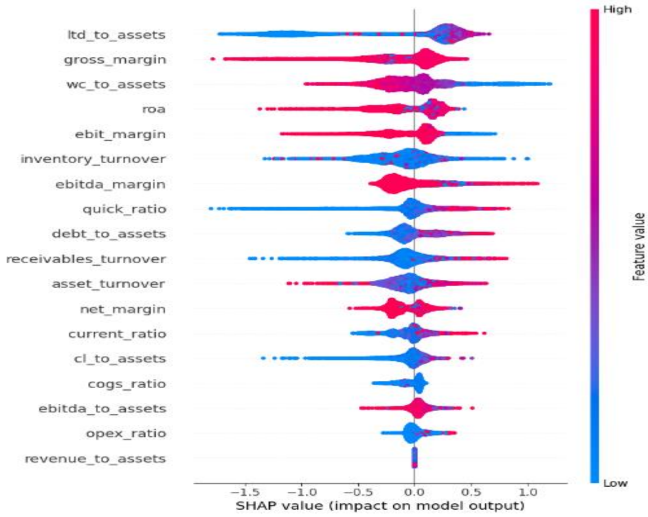
5.1.3. Liquidity and working capital

The SHAP results for the liquidity ratio variables (wc_to_assets, quick_ratio) indicate that they play an intermediate role in predicting bankruptcy. While negative values for these variables are associated with positive SHAP values between +0.3 and +0.5, their contribution to the model remains weaker than that of variables such as leverage and profitability. Thus, the liquidity variables have a major impact on the development of bankruptcy only if both leverage and profitability variables also contribute negatively to the outcome, a pattern immediately reflected in the SHAP values for companies that experienced bankruptcy.

5.1.4. Operational efficiency and high correlation

The correlation matrix shows that several profitability ratios are highly correlated, with coefficients above 0.8 and 0.9. In a linear model, this kind of correlation between variables can lead to inflated model coefficients. However, the SHAP values distribute the importance of these correlated variables according to their contribution to the model decision, rather than according to the inflation of their coefficients in a linear model. Thus, the SHAP values provide more reliable interpretations of the model coefficients (see Figure 5).

Figure 5. SHAP beeswarm plots

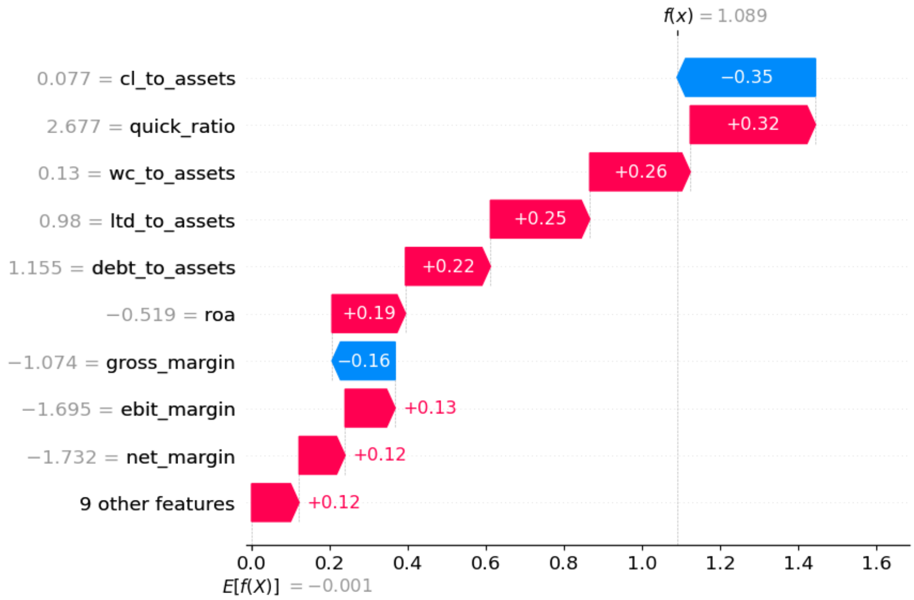


Source: Prepared by the Authors

5.2. Local interpretative analysis (Local SHAP)

The SHAP waterfall plot illustrates how the model moves from the near-neutral baseline model output of E[f(x)]≈–0.001 to the model's final output of f(x)≈1.089 as a result of the contributions of each of the model's features, as indicated in the figure. Furthermore, the plot can be interpreted as a visualization of the balance between the forces of the features that push the model's outcome towards bankruptcy (positive contributions) and those that tend to mitigate it (negative contributions). Thus, the analysis of this specific plot indicates that: (see Figure 6).

Figure 6. SHAP waterfall plot



Source: Prepared by the Authors

5.2.1. Leverage and solvency axis

The largest and most coherent contribution originates from the indicators related to leverage, most notably:

- debt_to_assets has a value of 1.155 and contributes 0.22 to the model
- ltd_to_assets has a value of 0.98 and contributes 0.25 to the model

Thus, the contribution of leverage-related variables to the model is 0.47. This is a critical value for the model, as it indicates that one of the main reasons for bankruptcy is the fragility of the company's financial structure. If the company's long-term debt relative to its assets is high, it may indicate that the company could experience distress if an adverse event negatively impacts its profitability.

5.2.2. Profitability and operating capacity axis

The next profitability indicator to consider is the return on assets, which contributes the most to the model. Its value of –0.519 contributes +0.19. This indicates that the return on the firm's assets is very weak, adding another layer to the firm's financial fragility. The combination of high leverage and low return on assets creates a fragile financial position for the firm.

5.2.3. Liquidity and working capital axis

Local SHAP values reveal how the model contextualizes liquidity in this particular observation:

- quick_ratio = 2.677, contributing +0.32
- wc_to_assets = 0.13, contributing +0.26

The interaction between these variables contributes to the model's score of +0.58, a substantial effect. At first glance, this effect seems counterintuitive. However, in the context of the local, non-linear model, high liquidity can indicate non-monotonicity in some cases. High levels of liquidity may indicate that a company is hoarding cash defensively due to a downturn in investing activity. A lack of investment activity can indicate weak sales or inventory levels. In these cases, the financial ratios tend to improve (due to non-monotonicity), but the company's financial health declines. High liquidity ratios may also result from a company's short-term attempts to improve its financial ratios without addressing underlying issues. This is precisely where the SHAP values become essential to the present study, as they reveal the model's understanding of factors that are not immediately obvious from the linear model alone.

5.2.4. Mitigating factors

Despite the marked upturn of bankruptcy pressure, two factors mitigate that outcome:

The factor of cl_to_assets had a contribution value of 0.077 to the bankruptcy prediction, contributing to the outcome of the case (−0.35) - the strongest mitigating factor for bankruptcy within this case. Lower short-term liabilities are considered a factor that makes it harder for a company to default in the short term.

The gross_margin factor had a contribution value of −1.074 to the bankruptcy prediction, which contributed to the outcome of the case (−0.16). While the factor is inherently positive with respect to bankruptcy risk, the contribution value shows how this factor works against the company in this example.

The mitigating factors contribute to the outcome of approximately −0.51 (−0.35 − 0.16). The factors that contribute to a positive outcome, however, include the contribution values of 0.58 (liquidity), 0.47 (leveraging), 0.19 (ROA), 0.13 (EBIT margin), 0.12 (net margin), and a few other variables with smaller contribution values. Thus, there is no reliance upon any single factor to indicate bankruptcy risk; rather, a combination of factors contributes to that outcome.

These individual explanations of the factors contributing to this company's financial distress should be evaluated in the context of the company's financial governance and risk monitoring in this case study, which informs the bankruptcy risk prediction model's decision for this company. Each of these factors is critical to the understanding of the risk of the company's financial distress from a governance perspective. Thus, while factors related to the company's liquidity and working capital provide some protection against bankruptcy, they are offset by the structure of the company's financials. Thus, although liquidity-related factors provide some protection against the company's potential bankruptcy, they are insufficient given the company's financial structure. These findings from this case study indicate that XAI techniques can be useful in providing explanations of the bankruptcy risk prediction model's decisions for high-risk cases in a financially responsible way. Furthermore, these analyses indicate that each high-performing bankruptcy-prediction model recognizes the same structure of financial risk in companies that experience bankruptcy. This factor can be used to monitor financial institutions for potential risks. In addition to these model findings, these explanations of the factors that contributed to the evaluated company's bankruptcy are significant, as they help reveal the performance of the various models used to assess the company's bankruptcy risk. The Stacking Ensemble model exhibited the highest AUC among the evaluated models. Thus, the XGBoost model's explanations of the factors contributing to bankruptcy for the evaluated company align with the financial structures of companies likely to experience bankruptcy, as identified by the high-performing models. While these individual factors do not indicate any causal relationship between them and the company's bankruptcy, they are sufficient to provide a level of transparency and accountability in reporting the risks those companies face, both to their internal stakeholders and to external financial regulatory bodies.

6. Discussion

The findings of this research study support recent literature suggesting that bankruptcy prediction is not merely a prediction problem but a decision-support problem. Financial analysts and decision-makers need to understand the bankruptcy prediction model's suggested values to support their institution's implementation of risk mitigation strategies. Furthermore, many existing studies into the topic of bankruptcy prediction have revealed that using common evaluation metrics for assessing the performance of these models can fail to accurately reflect the value of the model within financial institutions, especially in cases where there is a class imbalance within the company populations of predicted bankruptcies due to the rarity of those outcomes (Ainan et al., 2024). Thus, these findings further support the suggestion that using common evaluation metrics in the presence of class imbalance can lead to misinterpretations of the model's value.

6.1. Limitations of accuracy

As evidenced by these results, using accuracy to evaluate the models has led to potentially misleading conclusions about each model's ability to detect financial risk and to make sound decisions about which of the two models is performing better on this given data set. Each model has high accuracy. However, those accuracy figures only reflect the models' ability to identify companies outside of the bankrupt companies correctly - the majority class - rather than their ability to accurately recognize the bankrupt companies. These inaccuracies in each model's accuracy figures align with previous studies that utilized the same data set, wherein accuracy scores were similarly found to bias the models towards the majority class of the data set (Billios et al., 2024). Thus, accuracy is a metric that describes how closely the models align with the majority class within the data set, but introduces a gap between such models and their potential economic consequences. Overall, then, these findings support the arguments of the previous studies to utilize different metrics to evaluate the performance of each of these models - particularly those metrics that consider the potential impact of incorrectly predicting which companies will become bankrupt versus incorrectly predicting which companies will not become bankrupt (Ainan et al., 2024).

6.2. Error structures and model-specific decision logic

An examination of the error structures indicates differences in the models' decision logic. Such differences can be discussed in terms of their implications for financial decision-support systems. The logistic regression model has the highest recall but the lowest precision. This reaffirms the previous findings that logistic regression models are more precautionary in their decision-making. For this reason, logistic regression models are best applied to problems where interpretability and early detection of issues are of major concern (Billios et al., 2024). The models based on tree structures exhibit a more balanced approach to precision and recall. Both the random forest and XGBoost models have higher precision and ROC scores than the logistic regression model. The higher precision scores indicate that these models are better at identifying observations that exhibit a specific outcome than the logistic regression model. Higher ROC scores indicate that the random forest and XGBoost models have a stronger ability to discriminate between observations with and without the specific outcome of interest. This results in better early identification of financial distress in financial decision-making contexts (Förch Brenes et al., 2022). Furthermore, the improvement in ROC scores translates into more effective early-warning capability for identifying firms at risk of financial distress within credit monitoring systems.

6.3. Conservative decision behavior in ensemble models

The model with the most conservative decision behavior is the Stacking Ensemble model, which has very low recall values but the highest precision value at the default threshold. Models exhibiting such behavior are common in many recent ensemble and hybrid model frameworks for anomaly detection (Muslim et al., 2024). For instance, the confusion matrices for the Stacking Ensemble model indicate that it has fewer false-positive alerts than any other model. Furthermore, while such conservative behavior of the model in detecting early signs of financial distress within companies is not inherently undesirable from a statistical standpoint, from the standpoint of financial risk management, it indicates the model's lack of effectiveness as a stand-alone risk management system. The system's low recall rate suggests the model may fail to identify a significant number of financially distressed firms. For instance, financial institutions have been reported to prefer risk management models that produce fewer false-positive alerts than those with high rates of false positives (Hao et al., 2025). Thus, there is a need to balance a model's false-positive and false-negative rates according to the risk management objectives of the institutions employing it. Thus, these findings support the interpretation that the Stacking Ensemble model employs a deliberate decision-making process to produce alerts, rather than indicating a failure in the model's learning process.

6.4. Threshold sensitivity and the role of PR–AUC

The results of the sensitivity analysis further confirm that the model's conservative behavior is not structural but conditional. The model’s performance can be adapted to different strategies and thresholds for use within the financial monitoring system. By adjusting the decision threshold, the model can achieve significant gains in recall and F1-score at the expense of a small loss in precision. Furthermore, the model achieves high PR-AUC values. In cases of severe class imbalance, the precision-recall curve is a more informative evaluation metric for the model’s performance. The precision-recall measure evaluates the model’s performance by capturing the trade-off between precision and recall (Ainan et al., 2024). These results are consistent with recent literature on model evaluation in cases of severe class imbalance. Thus, the precision-recall area under the curve is a relevant evaluation measure for monitoring systems in banking applications.

6.5. Interpretability and financial coherence of model decisions

From a financial risk governance perspective, interpretability is critical to ensuring transparency and accountability in automated decision systems. The results of the interpretability analysis also provide further support for the narrative presented through the empirical analysis. The SHAP values for the model indicate that the primary structural contributor to a company's bankruptcy riskis its long-term leverage, though indicators such as profitability and liquidity act as modifiers to that main factor. These factors align with findings from previous studies suggesting that bankruptcy stems from weaknesses in a company’s financing and operating structure rather than from its financial strength (Gunonu et al., 2024). Furthermore, the local SHAP values help indicate the reasons for the failure of individual companies within the bank that are classified as high-risk for bankruptcy. Each of these reasons is defensible and helps justify the bank's decisions to classify individual companies as high-risk for bankruptcy. These findings align with previous studies suggesting that explainable artificial intelligence may assist in making decisions that are both transparent, defensible, and compliant with regulatory expectations in financial institutions. (Hao et al., 2025).

7. Conclusion

In the present study, bankruptcy prediction in the presence of rare events and class imbalance requires decisions about which models to employ in the context of financial governance and decision-making. The findings of the study indicate that conventional model evaluation metrics are insufficient for assessing bankruptcy prediction models in the presence of class imbalance. Furthermore, both linear and nonlinear models exhibited benefits and drawbacks to their decision behavior in the context of financial risk governance; in general, ensemble models that combined linear and nonlinear models exhibited improved performance in evaluating bankruptcy prediction models. Additionally, each of the models based on SHAP explanations of the bankruptcy prediction problems demonstrated an understanding of the financial structures of companies likely to experience bankruptcy, enabling them to provide explanations for their predictions. Thus, the study's findings indicate that approaches to bankruptcy prediction models should consider factors beyond model accuracy.

In addition to providing recommendations for financial regulatory authorities on the types of bankruptcy prediction models to employ in risk-based decision-making, these findings also have implications for explainable artificial intelligence (XAI) and AI-assisted risk monitoring systems. As with the financial regulators discussed above, other institutions that make decisions based on the predictions of AI-based bankruptcy models should also consider the decision-making behavior each model exhibits. Furthermore, any automated decision-making model needs to provide explanations for its predictions. Thus, the integration of XAI methods, such as SHAP, into bankruptcy prediction models makes them interpretable and defensible to the human overseers of those automated models.

Despite the paper's contributions, the study also has limitations. For instance, the data used in this report is drawn from historical financial data of publicly listed companies in the United States, which limits the generalizability of the findings to other markets. Furthermore, the model's explanations of the importance of each financial ratio are not based on established causality but on statistical relationships between the features and the target variable. Lastly, the model was evaluated at thresholds used in operational financial environments, which may not align with those used by the financial institutions deploying the model to determine whether a company is likely to default on its loans.

The study also suggests future research into incorporating temporal models or causal models, alongside model interpretability. Additionally, it is recommended to test the model across various markets and regulatory environments to assess its external validity and identify market-specific governance and regulatory requirements.

The value of the proposed framework is that it can be used to make decisions regarding bankruptcy predictions despite the severe class imbalance in the data. The study demonstrates the potential of explainable ensemble learning methods to assist financial decision-makers who rely on AI systems for financial analysis.

References

Ainan, U. H., Por, L. Y., Chen, Y.-L., Yang, J., & Ku, C. S. (2024). Advancing bankruptcy forecasting with hybrid machine learning techniques: Insights from an unbalanced Polish dataset. IEEE Access, 12, 9369–9381. https://doi.org/10.1109/ACCESS.2024.3354173

Akter, R. (2026). Evolution and emerging frontiers of explainable artificial intelligence (XAI) in financial risk management: A bibliometric analysis. Strategic Business Research, 2, 100118. https://doi.org/10.1016/j.sbr.2026.100118

Altman, E. I., & Hotchkiss, E. (2019). Corporate financial distress, restructuring, and bankruptcy: Analyze leveraged finance, distressed debt, and bankruptcy. John Wiley & Sons.

Babaei, G., & Giudici, P. (2025). Explainable artificial intelligence (XAI) in investment decision-making. Academia AI and Applications, 1(2), AcadAI8017. https://doi.org/10.20935/AcadAI8017

Balasubramaniam, N., Kauppinen, M., Rannisto, A., Hiekkanen, K., & Kujala, S. (2023). Transparency and explainability of AI systems: From ethical guidelines to requirements. Information and Software Technology, 159, 107197. https://doi.org/10.1016/j.infsof.2023.107197

Ben Jabeur, S., Stef, N., & Carmona, P. (2022). Bankruptcy prediction using the XGBoost algorithm and variable importance feature engineering. Computational Economics, 61(2), 715–741. https://doi.org/10.1007/s10614-021-10227-1. https://doi.org/10.1007/s10614-021-10227-1

Billios, D., Seretidou, D., & Stavropoulos, A. (2024). The Power of Numerical Indicators in Predicting Bankruptcy: A Systematic Review. Journal of Risk and Financial Management, 17(10), 433. https://doi.org/10.3390/jrfm17100433

Cao, Y., Luo, Y., Wei, P., Zhai, J., & Shi, S. (2026). Bankruptcy forecasting—Market information with ensemble model. The British Accounting Review, 58(3), 101530. https://doi.org/10.1016/j.bar.2024.101530

Förch Brenes, R., Johannssen, A., & Chukhrova, N. (2022). An intelligent bankruptcy prediction model using a multilayer perceptron. Intelligent Systems with Applications, 16, 200136. https://doi.org/10.1016/j.iswa.2022.200136

Gabrielli G, Melioli A, Bertini F (2026), Corporate financial distress prediction: a machine learning approach in the era of big data. Journal of Accounting & Organizational Change, 22(7), 31–65. https://doi.org/10.1108/JAOC-05-2025-0166

Garcia, J. (2022). Bankruptcy prediction using synthetic sampling. Machine Learning with Applications, 9, 100343. https://doi.org/10.1016/j.mlwa.2022.100343

Gunonu, S., Altun, G., & Cavus, M. (2026). Explainable bank failure prediction models: Counterfactual explanations to reduce the failure risk. Computational Economics. https://doi.org/10.1007/s10614-026-11353-4

Hamdi, M., Mestiri, S., & Arbi, A. (2024). Artificial Intelligence Techniques for Bankruptcy Prediction of Tunisian Companies: An Application of Machine Learning and Deep Learning-Based Models. Journal of Risk and Financial Management, 17(4), 132. https://doi.org/10.3390/jrfm17040132

Hao, Y., Chen, T.-K., & Lin, Y.-C. (2025). Bankruptcy prediction using the text-based communicative value of earnings call transcripts. Review of Quantitative Finance and Accounting. https://doi.org/10.1007/s11156-025-01465-7

Muslim, M. A., Dasril, Y., Javed, H., Alamsyah, Jumanto, Abror, W. F., Pertiwi, D. A. A., & Mustaqim, T. (2024). An ensemble stacking algorithm to improve model accuracy in bankruptcy prediction. Journal of Data Science and Intelligent Systems, 2(2), 79–86. https://doi.org/10.47852/bonviewJDSIS3202655

Papík, M., & Papíková , L. (2025). The possibilities of using AutoML in bankruptcy prediction: Case of Slovakia. Technological Forecasting and Social Change, 215, 124098. https://doi.org/10.1016/j.techfore.2025.124098

Park, M. S., Son, H., Hyun, C., & Hwang, H. J. (2021). Explainability of machine learning models for bankruptcy prediction. IEEE Access, 9, 124887–124899. https://doi.org/10.1109/ACCESS.2021.3110270

Wang, H., & Liu, X. (2021). Undersampling bankruptcy prediction: Taiwan bankruptcy data. PLOS ONE, 16(7), e0254030. https://doi.org/10.1371/journal.pone.0254030

Ye, S., Khishe, M., Ibrahim, B. F., & Smerat, A. (2025). Advanced financial risk forecasting using enhanced kernel-based extreme learning machines: Tackling challenges in bankruptcy problem. Ain Shams Engineering Journal, 16(9), 103518. https://doi.org/10.1016/j.asej.2025.103518

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