



Advanced machine learning approaches for bankruptcy prediction in the shipping industry: a comparative analysis across time horizons

Minsu Kwon¹ · Saeyeon Roh² · Minchul Sohn²

Accepted: 26 January 2026

© The Author(s), under exclusive licence to Springer Nature Limited 2026

Abstract

This study evaluates machine learning methods for predicting bankruptcies in the Korean shipping industry over 1-, 3-, and 5-year horizons. Using an imbalanced dataset, preprocessing techniques such as outlier capping, skewness correction, and SMOTE-based oversampling are applied. Two feature sets—financial variables and those enriched with macroeconomic and shipping indicators—are analyzed. Results demonstrate that ensemble methods like LightGBM, CatBoost, XGBoost outperform linear models and deep neural networks across all horizons. Short-term predictions are driven by liquidity factors, while long-term forecasts benefit from shipping-specific indices like freight rates. SHAP-based analysis reveals that internal cash-flow as critical for one-year predictions, whereas prolonged freight rate declines and financing pressures dominate longer horizons. These findings highlight the importance of horizon-specific modeling, support the adoption of advanced machine learning in maritime risk assessment, and encourage further exploration of sector-specific features to improve predictive accuracy and resilience in the shipping industries.

Keywords Machine learning · Bankruptcy prediction · Shipping industry · Ensemble methods

✉ Saeyeon Roh
saeyeon.roh@plymouth.ac.uk

¹ Department of International Trade and Commerce, Soonchunhyang University, Asansi, South Korea, Korea

² Plymouth Business School, University of Plymouth, Plymouth PL4 8AA, UK



1 Introduction

The shipping industry is essential to global trade but remains vulnerable to economic fluctuations due to its cyclical nature and high capital requirements (Stopford 2013). High-profile collapses, such as Hanjin Shipping's insolvency, illustrate how external shocks and internal mismanagement can quickly destabilize large firms (Kavussanos and Tsouknidis 2016). Accurate default prediction is therefore crucial for a range of stakeholders, yet forecasting in this sector is challenging because it depends on an interplay of financial, operational, and macroeconomic factors (Song et al. 2019).

Data constraints in the shipping industry complicate further these efforts, as the minority of bankruptcy observations is often small and highly imbalanced (Yeo 2016). This imbalance skews many conventional classifiers, limiting their ability to detect early warning signals (Park et al. 2021). In addition, the industry's global reach amplifies the impact of macroeconomic indicators, including trade indices and interest rates, although the degree of their influence varies by forecasting horizon (Yu et al. 2020).

Focusing on the Korean shipping sector offers a compelling case study because it features a limited number of listed firms and pronounced cyclical volatility, as exemplified by Hanjin's collapse. This setting provides both a challenge and an opportunity to test machine learning models that are designed to handle rare default events. Traditional statistical methods, such as logistic regression and discriminant analysis, are widely used (Le 2022) but often rely on linear assumptions and fail to account for complex interactions, particularly under severe imbalance (Zoričák et al. 2020). More recent techniques leverage ensemble learning and deep neural networks (Alam et al. 2021), yet their performance may vary substantially based on whether the forecast window is one, three, or five years (Jones 2017).

Our study systematically compares a broad range of machine learning models across the above three horizons, employing a data preprocessing framework that addresses outliers, skewness, and class imbalance. The analysis assesses two sets of input features: purely financial data and combined financial-plus-macroeconomic variables, including freight indices. Standard evaluation metrics such as the AUC-ROC and a Shapley Additive exPlanations (SHAP)-based interpretability framework highlight how global conditions matter more in extended horizons, while deep learning approaches do not always surpass advanced ensemble methods in this imbalanced context. Although the primary focus is on Korea's shipping industry, the findings hold implications for similarly capital-intensive, cyclical sectors, where robust default predictions are vital for effective risk management.

2 Literature review

2.1 Traditional and advanced approaches to bankruptcy prediction

Bankruptcy prediction research has seen remarkable evolution since the introduction of ratio-based methods. Early linear models, such as Altman's Z-score, fundamentally relied on discriminant analysis to combine financial metrics (Altman et al.



2020). Despite providing interpretable criteria for distress classification, these linear techniques often fail to capture the complexities of modern industries, especially those characterized by high capital requirements and cyclical market demands (Geng et al. 2015). Over time, analysts recognized that firms operating in volatile industrial sectors, especially shipping, exhibit intricate, non-linear patterns in their balance sheets, rendering basic linear methods inadequate (Clintworth et al. 2023).

To address these complexities, researchers turned to machine learning algorithms capable of uncovering deeper relationships. Support Vector Machines (SVM) leverage kernel functions to separate classes, thereby handling non-linear financial data more effectively than linear discriminant analyses (Gogas et al. 2018; Ciampi and Gordini 2013). Artificial Neural Networks (ANN) emerged as another approach, learning multi-layered representations of input features (Zhang et al. 1999). Yet, in practical implementation, both SVM and ANN can become susceptible to overfitting if hyperparameters are not carefully tuned, especially when working with relatively small datasets (Le et al. 2018).

Building on these approaches, ensemble methods have gained prominence by combining multiple learners—often decision trees—into a more robust predictor (Kumar and Ravi 2007). Random Forest, for instance, uses bagging and random feature selection to reduce variance and overfitting risk, while Gradient Boosting Machines (GBMs) iteratively update weak learners by re-focusing on previously misclassified instances (Zhang et al. 2021). The synergy of multiple models allows ensembles to address a variety of hidden patterns in financial data that single classifiers might miss.

In the last decade, attention has shifted toward Extreme Gradient Boosting (XGBoost) (Chen and Guestrin 2016), a highly optimized form of gradient boosting well-suited to tabular data. XGBoost's computational efficiency and flexible regularization have led to consistent success in numerous predictive tasks, including bankruptcy detection. Meanwhile, Recurrent Neural Network (RNN) architectures, particularly Long Short-Term Memory (LSTM) networks, excel at capturing sequential dependencies (Hochreiter and Schmidhuber 1997). LSTM's gating mechanisms mitigate the “vanishing gradient” problem found in basic RNNs, rendering it useful for multi-year predictions in contexts where historical data trends matter. Some studies report that LSTM-based models can outperform traditional techniques when large, high-frequency datasets are available, and when capturing time-dependent behavior is critical (Molina and Preve 2012). Nevertheless, practical challenges—particularly in data-scarce, high-volatility sectors—may curtail deep learning's advantages (Yuen and Ko 2018). Consequently, the application of these advanced models must be carefully adapted, balancing their theoretical strengths against real-world data constraints.

2.2 The role of imbalanced data in shipping

One persistent barrier to effective bankruptcy prediction in shipping is data imbalance, where the proportion of bankrupt firms is overshadowed by the majority of solvent companies (Sun 2007; Zoričák et al. 2020). This skewed distribution can lead standard classification models to favor predicting solvency, thereby overlooking subtle signals of impending distress. The cost of false negatives failing to flag an actual bankruptcy can be disastrous to financial stakeholders, reinforcing the need



for specialized rebalancing strategies. Thus, to train a robust classifier, the dataset must include both solvent and insolvent firms. From a machine learning perspective, exposure to both classes is essential for the algorithm to establish a valid decision boundary and identify the distinguishing features of financial distress (Bishop 2006).

Researchers have responded by developing various data-balancing techniques. Among the most widely adopted is the Synthetic Minority Over-sampling Technique (SMOTE), which synthesizes new minority instances rather than simply duplicating existing ones (Chawla et al. 2002). SMOTE helps classifiers learn richer patterns that might exist in the minority class, improving metrics like recall and F1-score (Zhou and Lai 2017). Others have proposed cost-sensitive learning, where classification errors on the minority class incur heavier penalties than those on the majority, prompting models to pay closer attention to potentially bankrupt firms (Sun et al. 2020). While effective, these methods can also introduce complications: synthetic data may not always accurately reflect real-world nuances, and cost-sensitive tweaks may overcorrect, leading to more false positives (Le et al. 2018). However, relying solely on rare real world bankruptcy cases may lead to overfitting. To mitigate this, we employ SMOTE to generate synthetic minority instances. This approach aligns with recent methodologies such as Sizan et al. (2025) and Lin et al. (2025), where synthetic data serves to enrich the minority class feature space, ensuring the model is not overwhelmed by the majority solvent firms.

Combining SMOTE with ensemble methods, such as XGBoost or Random Forest, has shown promise in shipping. These algorithms can adaptively identify borderline minority observations that might otherwise be overlooked, refining each subsequent iteration of learning (Zhou and Lai 2017). However, shipping is prone to sudden macroeconomic shifts, meaning any synthetic oversampling approach can risk overfitting if market conditions change abruptly. A balanced dataset is essential for model stability, yet it must be complemented by robust validation protocols—like cross-validation and out-of-time tests—to ensure genuine generalizability to evolving market conditions.

2.3 Significance of prediction horizons

The forecasting horizon, whether one year, three years, or five years before bankruptcy strongly influences model design and performance. Short-term predictions can rely heavily on liquidity indicators, such as current ratios or sudden profit declines, which tend to be immediate red flags for default (Zhang et al. 2021). Their limited time window often boosts classifier accuracy, because fewer macroeconomic or strategic variables shift drastically over a single fiscal year. Many studies using logistic regression or discriminant analysis have found acceptable results in short-term forecasting, but these methods may prove inadequate for capturing cyclical or structural factors that take years to unravel (Ghosh and Kapil 2022).

By contrast, long-term predictions (e.g., three or five years out) require the model to incorporate extended debt obligations, capital investments, and global market dynamics. In shipping, multi-year lags in vessel orders, freight-rate fluctuations, and capital refinancing cycles mean that early problems can remain latent for a significant period (Li and Miu 2010). LSTM networks often excel here, leveraging sequential



data to track gradually emerging patterns (Cha and Kang 2018). Nevertheless, their success depends on whether enough timely, longitudinal data are available to capture multi-year evolutions. Moreover, ensemble methods, especially gradient boosting, can be adapted for longer-term contexts by including strategic and macroeconomic features, yet their performance can vary depending on how effectively each round of boosting weighs long-horizon signals. There is little consensus on which algorithm is consistently superior for each timeframe, demonstrating the need for systematic comparisons in sectors like shipping, where both short-run and long-run volatilities are significant.

2.4 Industry-specific challenges in shipping bankruptcy prediction

The shipping industry poses unique difficulties for bankruptcy prediction due to its capital intensity, cyclical nature, and global exposure. Significant capital is tied up for extended periods due to fleet modernization and new vessel acquisitions. This long-term asset rigidity heightens the industry's vulnerability to sudden shifts in freight rates and commodity prices (Haider et al. 2019). A single downturn can trigger liquidity pressures that extend over multiple years, a phenomenon not always visible through basic financial ratios alone. Additionally, shipping companies operate in a fiercely competitive global environment; external shocks, such as volatile fuel prices or deteriorating trade volumes, can significantly distort financial models in short order (Lozinskaia et al. 2017; Mitroussi et al. 2016).

Volatility is further exacerbated by cyclical freight markets. Freight rates may rise or fall rapidly, yet a shipping firm's capital structure is often locked in for longer periods, creating a misalignment between short-term revenues and long-term obligations (Park et al. 2021). Traditional regression-based models struggle to incorporate these abrupt changes, whereas ensemble learning can reweight updated data as new patterns emerge. Nevertheless, the persistent class imbalance remains a substantial barrier with bankruptcies relatively rare, many models risk focusing on the majority solvent class, increasing the risk of failing to detect genuine distress signals (He and Garcia 2009; Wang et al. 2021; Shrivastava et al. 2020).

Balancing short-term and long-term perspectives adds yet another layer of complexity. Short-range forecasts emphasize immediate solvency alerts such as escalating short-term debt or steep declines in Earnings Before Interest, Taxes, Depreciation, and Amortisation (EBITDA), yet they may underestimate structural or macro-level risk factors that only manifest over several business cycles (Perboli and Arabnezhad 2021). Meanwhile, modeling five-year horizons allows analysts to observe how cyclical dips or strategic missteps accumulate, but these extended forecasts can miss more immediate indicators and are prone to larger uncertainty if market conditions shift drastically. Effective bankruptcy prediction models in shipping must thus simultaneously handle cyclical downturns, manage data imbalance, and track both immediate financial signals and multi-year strategic moves.

Given these industry-specific complexities, advanced machine learning offers a path forward but demands careful tailoring of algorithms and feature sets. Accurate shipping default prediction requires (1) robust data-balancing methods to preserve minority-class bankruptcy signals; (2) thoughtful horizon selection to account for



short-term liquidity crises versus long-term strategic vulnerabilities; and (3) integration of macroeconomic and sector-specific features such as freight rates, trade volume, and interest rates to fully capture the external shocks that ripple through highly leveraged shipping firms. In this context, ensemble methods (Random Forest, gradient boosting, XGBoost) and certain specialized neural architectures such as LSTM appear well-positioned to produce actionable insights, provided they are deployed with rigorous validation protocols and an awareness of shipping's cyclical market structure.

3 Methodology

3.1 Data collection and characteristics

The research framework for this study is illustrated in Fig. 1. To investigate how various predictive models and preprocessing techniques perform with limited, imbalanced data, the study compiled firm-level observations from Korean shipping companies spanning 2001 to 2023. Companies were classified as either bankrupt or non-bankrupt. Bankrupt firms were identified through formal court proceedings, restructuring announcements, recognized signals of insolvency, or delisting from the

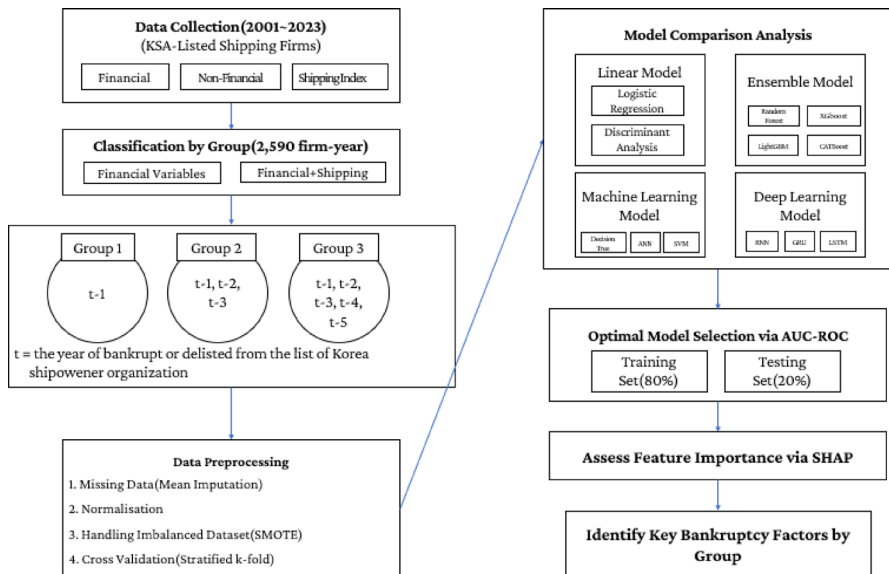


Fig. 1 Research framework for bankruptcy prediction



Korea Shipowners' Association.¹ Because bankruptcies are relatively infrequent in the Korean shipping sector, the dataset is both small and highly imbalanced.

As shown in Table 1, only 2.37% of observations represent bankrupt companies at the 1-year horizon, compared to 97.63% non-bankrupt. The proportion of bankrupt observations increases at longer horizons in 7.19% at 3 years and 11.59% at 5 years, but remains substantially lower than non-bankrupt observations. This severe imbalance can bias conventional predictive models toward the majority class (Haixiang et al. 2017). Furthermore, the limited number of bankruptcy cases constrains the capacity of models to generalize, demanding careful treatment of data imbalance and sample scarcity to ensure robust predictive performance.

The dataset includes a diverse range of variables categorized into financial, operational, and macroeconomic factors. Specifically, 80 variables were selected based on their explanatory power as identified in previous research on bankruptcy prediction, financial distress, and shipping industry performance shown in Table 2. These variables encompass firm-level balance-sheet ratios, profitability indicators, shipping-industry metrics, and macroeconomic measures such as trade volume, Gross Domestic Product (GDP) growth, and oil prices. The large number of input variables is justified by the need to account for the complex and multifactorial nature of bankruptcy prediction, particularly in capital-intensive and cyclically sensitive industries such as shipping. The inclusion of variables with demonstrated explanatory ability ensures the robustness of the models and allows for the evaluation of variable importance using interpretable machine learning method.

By capturing these variables, the study assesses the predictive accuracy of various models over multiple horizons in 1 year, 3 years, and 5 years prior to bankruptcy. This approach allows for a more comprehensive understanding of early warning signals for financial distress and the influence of macroeconomic factors, which prior studies suggest are particularly significant at longer horizons.

3.2 Data preprocessing

Despite efforts to compile a comprehensive dataset, the total number of observations featuring bankrupt firms is relatively limited, leading to a highly imbalanced classification scenario that can bias model training. To address this challenge, the study applies three key data-processing techniques. First, Winsorization at the 0.1% and 99.9% levels is used to cap extreme outliers, a measure deemed suitable for shipping

Table 1 Distribution of active and failed firm-year observations

Years prior to bankruptcy	Bankrupt		Active	
	No	Percentage	No	Percentage
5 years	303	11.59	2310	88.41
3 years	188	7.19	2425	92.81
1 years	62	2.37	2551	97.63

¹ The Korea Shipowner's Association (KSA), which is the prominent organization representing shipowners in Korea, requires its member companies to meet minimum operational and financial standards. When a company is unlisted from KSA, it indicates failure to satisfy these standards which often signal insolvency or an inability to continue operations.



Table 2 List of financial indicators

Category	Variable	Description	Category	Variable	Description
Leverage					
A1	Return on total assets	$\frac{EBITDA}{Total Assets}$	Profitability		
A2	Retained Earnings to Total Assets	$\frac{Retained Earnings}{Total Assets}$	C1	Return on Assets	$\frac{Net Income}{Total Assets}$
A3	Leverage	$\frac{Total Debt}{Total Assets}$	C2	Return on Equity	$\frac{Net Income}{Shareholder's Equity}$
A4	Liabilities to Total Assets	$\frac{Total Liabilities}{Total Assets}$	C3	Return on Sales	$\frac{Net Income}{Sales}$
A5	Equity to Assets	$\frac{Shareholder's Equity}{Total Assets}$	C4	EBITDA to Sales	$\frac{EBITDA}{Sales}$
A6	Debt to Equity	$\frac{Total Debt}{Shareholder's Equity}$	C5	Sales to Current Assets	$\frac{Sales}{Current Assets}$
A7	Liabilities to Equity	$\frac{Total Liabilities}{Shareholder's Equity}$	C6	Sales	$\frac{Sales}{Total Sales}$
A8	Gearing ratio	$\frac{Shareholder's Equity}{Total Debt}$	C7	EBITDA to Interest Coverage	$\frac{EBITDA}{Interest expenses}$
A9	Equity to liabilities	$\frac{Shareholder's Equity}{Total Liabilities}$	C8	Gross Profit to Assets	$\frac{Gross Profit}{Total Assets}$
A10	Liabilities	$\frac{Total Liabilities}{Current Liabilities}$	C9	Net interest margin	$\frac{Interest returns - Interest paid}{Average Assets}$
A11	Current Liabilities to Equity	$\frac{Current Liabilities}{Equity}$	C10	Sales to equity	$\frac{Total Sales}{Shareholder's Equity}$
A12	EBITDA to Liabilities	$\frac{EBITDA}{Total Liabilities}$	C11	Net Income	$\frac{Total Revenue - Total Expenses}{Total Sales}$
A13	EBITDA/EV	$\frac{EBITDA}{Equity}$	C12	Operating Return on Assets	$\frac{Operating Income}{Total Assets}$
A14	EBITDA	$\frac{Net Income + Interest + Taxes + Depreciation + Amortization}{liabilities - current liabilities}$	C13	Sales to Liabilities	$\frac{Total Sales}{Total Liabilities}$
A15	Long-term Liabilities to Assets	$\frac{liabilities - current liabilities}{Total Assets}$	C14	Gross Profit to Current Liabilities	$\frac{Gross Profit}{Current Liabilities}$
A16	Net Income to Liabilities	$\frac{Net Income}{Total Liabilities}$	C15	Gross Profit to Liabilities	$\frac{Gross Profit}{Total Liabilities}$
A17	Working Capital	Current Assets - Current Liabilities	C16	Gross Profit to Sales	$\frac{Gross Profit}{Total Sales}$
A18	Current Assets to Liabilities	$\frac{Current Assets}{Total Liabilities}$	C17	ΔSales	$\frac{Total sales_t - Total Sales_{t-1}}{Total Sales_{t-1}}$
			C18	Sales to Current Liabilities	$\frac{Total Sales}{Current Liabilities}$



Table 2 (continued)

Category	Variable	Description	Category	Variable	Description
Leverage					
A19	EBITDA to Current Liabilities	$\frac{EBITDA}{Current Liabilities}$	Profitability		
A20	Long-term Liabilities to Equity	$\frac{liabilities-current liabilities}{Shareholder's Equity}$	Efficiency		
Liquidity					
B1	Current Ratio	$\frac{Current Assets}{Current Liabilities}$	D1	Asset turnover	$\frac{Total Sales}{Total Assets}$
B2	Working Capital to Assets	$\frac{Current Assets - Current Liabilities}{Total Assets}$	D2	Working Capital to Sales	$\frac{Working Capital}{Total Sales}$
B3	Cash Assets Ratio	$\frac{Cash}{Total Assets}$	D3	Current Assets to Sales	$\frac{Current Assets}{Total Sales}$
B4	Cash Ratio	$\frac{Cash}{Current Liabilities}$	D4	Inventory to Assets	$\frac{Inventory}{Total Assets}$
B5	Current Assets to Assets	$\frac{Current Assets}{Total Assets}$	D5	Current Liabilities to Sales	$\frac{Current Liabilities}{Total Sales}$
B6	Current Liabilities to Assets	$\frac{Current Liabilities}{Total Assets}$	D6	Inventory turnover	$\frac{Average Inventories}{Total Sales}$
B7	Quick Ratio	$\frac{Cash+accounts receivable}{Current liabilities}$	D7	Operating Margin	$\frac{Operating Income}{Sales}$
B8	Cash to Debt Ratio	$\frac{Cash}{Total Debt}$	Non-financial		
B9	Shareholder's Equity	$Total Assets - Total Liabilities$	E1	Size	Proxied by $Log(Total Assets)$
B10	Current Liabilities Ratio	$\frac{Current Liabilities}{Total Liabilities}$	E2	Age	Age of firm
B11	Current Assets	$Cash + Accounts Receivable + Inventory$	E3	Type of Operator	Dummy value of Type: 1 (Container), 2 (Tanker), 3 (Bulk)
B12	Cash Sales Ratio	$\frac{Cash}{Total Sales}$	E4	GDP	$\frac{GDP_t - GDP_{t-1}}{GDP_{t-1}}$
B13	Current Liabilities to Current Assets	$\frac{Current Liabilities}{Current Assets}$	E5	Oil Price	Current US\$ price of Brent oil
B14	Working Capital to Current Assets	$\frac{Working Capital}{Current Assets}$	E6	Currency	Annual exchange rate of KRW to USD
B15	Quick Assets to Total Assets	$\frac{Cash+accounts receivable}{Total Assets}$	Shipping Index		Annual average of Libor interest rates
			F1	LIBOR	
			F2	Growth rate of Global Seaborne Trade Volume	$\frac{Trade volume_t - Trade volume_{t-1}}{Trade volume_{t-1}}$



Table 2 (continued)

Category	Variable	Description	Category	Variable	Description
Leverage					
B16	Working Capital to Equity	$\frac{\text{Current Assets} - \text{Current Liabilities}}{\text{Shareholder's Equity}}$	Profitability		
B17	Working Capital to Current Liabilities	$\frac{\text{Current Assets} - \text{Current Liabilities}}{\text{Current Liabilities}}$	F3	Growth rate of Korean sea-borne Trade volume	$\frac{\text{Trade volume}_t - \text{Trade volume}_{t-1}}{\text{Trade volume}_{t-1}}$
B18	Cash to Current Assets	$\frac{\text{Cash}}{\text{Current Assets}}$	F4	Container Freight Index Rate	Shanghai Containerized Freight Index
B19	Retained Earnings to Current Liabilities	$\frac{\text{Cash} + \text{Retained Earnings}}{\text{Current Liabilities}}$	F5	Baltic Dry Index (BDI),	Baltic dry index per month
B20	Δ Total Asset	$\frac{\text{Total Assets}_t - \text{Total Assets}_{t-1}}{\text{Total Assets}_{t-1}}$	F6	IRONSTEEL	Dow Jones U.S. Iron & Steel Index
			F7	Number of vessels	Number of owned vessels



companies because of the sector's pronounced variability in revenues, costs, and asset valuations (García et al. 2016). Second, a Box-Cox transformation is implemented to correct skewness and facilitate more normal-like distributions for variables such as Liabilities, Income, or Assets. Although many machine learning algorithms do not require normally distributed data, pilot tests showed that modeling outcomes were more stable when skewed features were transformed (Atkinson et al. 2021). Third, SMOTE is applied to manage the severe class imbalance by synthesizing minority-class observations, thereby allowing the model to learn bankruptcy patterns more effectively (Fernández et al. 2018). These three steps (Winsorization, Box-Cox transformation, and SMOTE) are carried out consistently across all time horizons (1, 3, 5 years before bankruptcy) to ensure that each model benefits from an appropriately balanced and standardized training dataset.

The study includes 80 variables selected from prior research for their demonstrated explanatory power in predicting financial distress, particularly in the shipping industry. Feature selection techniques, such as Variance Inflation Factor (VIF) and sensitivity analysis, were not applied for two reasons. First, VIF is commonly used to address multicollinearity by excluding highly correlated variables (Kutner et al. 2005), and sensitivity analysis identifies influential features by assessing their impact on model outputs. However, excluding variables in this study's relatively small dataset risked discarding potentially important predictors. Second, one objective of this study was to evaluate the contribution of all variables using interpretable machine learning techniques (e.g., SHAP), making feature elimination counterproductive (Kou et al. 2021).

Following data preprocessing, the modeling phase is carried out separately for two categories of input features. The first category, referred to as "financial only," includes exclusively accounting-based and firm-specific measures (leverage ratios, liquidity ratios, profitability ratios and efficiency rates). The second, labeled financial plus macroeconomic, augments these internal indicators with external variables such as global and Korean trade volume growth rates, interest rates, and oil prices, among others that may plausibly affect shipping demand or cost structures. This division allows for a straightforward comparison of how much additional predictive power macroeconomic variables can contribute, especially as the forecast horizon grows. We acknowledge that internal financial ratios are influenced by the macro environment. However, we separate these sets to isolate the incremental predictive value of external data. While financial ratios capture firm-specific health, explicit macroeconomic indicators serve as leading signals of systemic risk that are not immediately reflected in a firm's balance sheet (Kheneifar and Amiri 2025).

3.3 Model selection

The models selected for evaluation encompass four broad classes. First, traditional statistical methods (Multivariate Discriminant Analysis and Logistic Regression) serve as baselines, given their historical prominence in bankruptcy prediction studies (Altman 1968; Ohlson 1980). Although these techniques are parsimonious and interpretable, they assume mostly linear relationships that may not adequately capture the nonlinearities in shipping-related data.



Second, classical machine learning methods include Decision Tree, Random Forest, SVM, and ANN. These can model more complex relationships but can become prone to overfitting if hyperparameters are not carefully tuned (Breiman 2001; Rumelhart et al. 1986; Vapnik and Cortes 1995).

Third, ensemble models, including Gradient Boosting Machines, XGBoost, Adaptive Boosting (AdaBoost), Light Gradient Boosting Machine (LightGBM), and Categorical Boosting (CatBoost) have recently achieved strong results in tabular and imbalanced datasets (Friedman 2001; Chen and Guestrin 2016; Ke et al. 2017; Prokhorenkova et al. 2018). Because gradient boosting methods iteratively focus on misclassified observations, they tend to excel in distinguishing minority classes.

Fourth, deep learning approaches such as RNN, LSTM, and Gated Recurrent Unit (GRU) are tested to investigate whether architectures designed for temporal sequences and long-term dependencies can outperform simpler boosting algorithms, even when data are limited and imbalanced interpret (Dey and Salem 2017; Gers et al. 2000; Hochreiter and Schmidhuber 1997).

Each model is trained and validated under a consistent procedure: a five-fold cross-validation approach is used, whereby the dataset is partitioned into five subsets, and each subset is sequentially reserved for testing while the remaining four subsets train the model. Given the limited number of bankruptcy observations, a fivefold approach was chosen over tenfold to ensure that each validation fold contains a statistically sufficient number of minority class samples for stable evaluation. This ensures that every observation serves as a test instance exactly once, mitigating overfitting risks and providing a more realistic gauge of predictive accuracy. Hyperparameter tuning for complex models (XGBoost, CatBoost, LightGBM, deep neural nets) is accomplished through a combination of grid search and random search: the former systematically explores a predefined space of parameters, whereas the latter draws parameter combinations randomly from larger ranges, a procedure designed to handle high-dimensional or less structured hyperparameter configurations (Snoek et al. 2012).

3.4 Evaluation metrics

Model performance was evaluated using both the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC-ROC). The ROC curve provides a visual assessment of the trade-off between true positive and false positive rates across different classification thresholds, while the AUC-ROC serves as a summary measure of the model's discriminative ability (Table 3 and Fig. 2). These metrics are

Table 3 Confusion matrix for binary classification

		Predicted value	
		0 (non-bankruptcy)	1 (bankruptcy)
Actual Value	0 (non-bankruptcy)	True Positive (TP)	False Negative (FN)
	1 (Bankruptcy)	False Negative (FP)	True Negative (TN)

Source: Hastie et al. (2009)



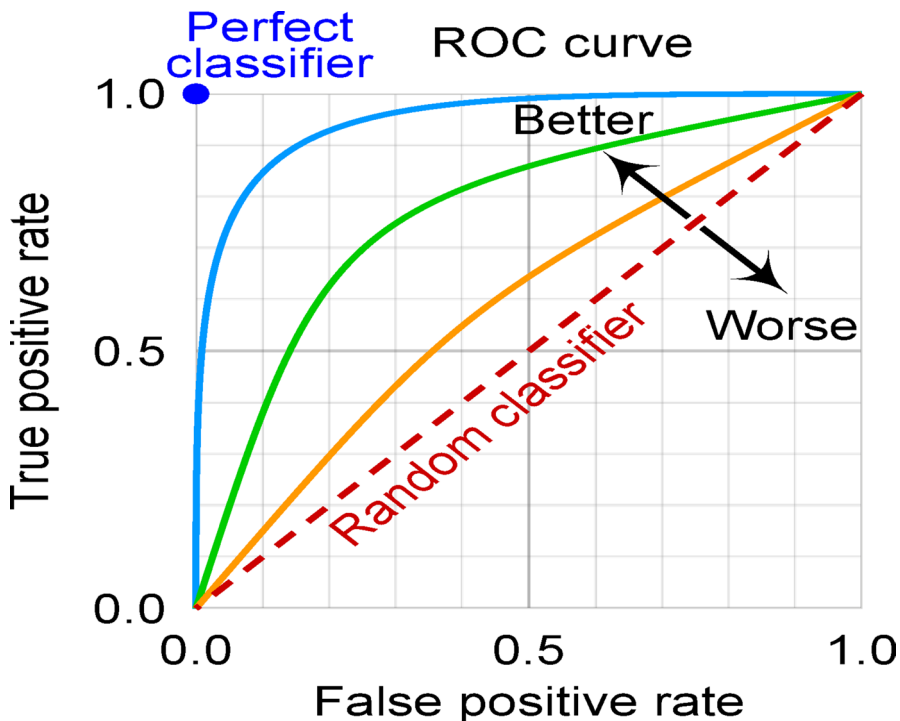


Fig. 2 Receiver operating characteristic (ROC) curve. Source: Cook (2007)

particularly suitable for evaluating models on imbalanced datasets, as they remain robust despite uneven class distributions (Fawcett 2006).

$$Receiver\ Operating\ Characteristics = \int_0^1 TPR(FPR^{-1}(x)) dx \quad (1)$$

For additional interpretability, the study employs the SHAP framework to rank features by their overall contribution to the predictive outcome, thereby identifying which variables between financial and macroeconomic are most decisive. Because shipping bankruptcy often arise from multifaceted interactions such as rising leverage in the face of reduced freight demand, such interpretability is valuable for industry practitioners, allowing them to pinpoint specific reasons behind an elevated bankruptcy risk. SHAP values are computed based on Shapley values from cooperative game theory, providing a unified approach to explain the output of machine learning models. SHAP values decompose a model’s prediction $f(x)$ for an instance x into contributions from each feature i :

$$f(x) = \varphi_0 + \sum_{i=1}^M \varphi_i \quad (2)$$



where φ_0 : The baseline value, representing the average model output over the training dataset; φ_i : The SHAP value feature i , representing its contribution to the deviation from baseline value; M : The total number of features.

By combining these methodological components, the study aims to provide comprehensive insights into how advanced machine learning, under rigorous data preprocessing, can enhance or potentially replace traditional methods of financial distress detection in the shipping sector. The comparison of one-year, three-year, and five-year horizons is critical in clarifying the time-dependent relevance of different predictors. Short-term bankruptcy might hinge on abrupt liquidity crunches, whereas extended predictions are more likely to be influenced by macroeconomic cycles and capital-structure evolutions over multiple years. Similarly, distinguishing between financial-only data and data that incorporate macroeconomic variables allows for a better understanding of when external indicators like trade volumes, oil prices, or GDP movements carry enough weight to shift default probabilities meaningfully.

4 Results

4.1 One-year horizon

When forecasting bankruptcy one year ahead, the dominant signals typically arise from abrupt liquidity crises and short-term fluctuations in revenues or costs (Table 4). In the financial-only setting, ensemble algorithms, especially XGBoost, LightGBM, and CatBoost consistently registered AUC-ROC values of around 0.90–0.92, outpacing traditional baselines such as Logistic Regression and Latent Dirichlet Allocation (LDA), which hovered near 0.75–0.80. A key reason behind the ensembles' success is their ability to model non-linear relationships among near-term ratios (e.g., current liabilities to assets, working capital to sales) while simultaneously handling class imbalance when coupled with SMOTE. Deep learning models including RNN, LSTM, and GRU delivered moderate performance in the 0.70–0.80 range but struggled to match the precision and recall consistently achieved by boosting methods, likely due to the limited size of the bankruptcy sample and the short-term nature of the predictive signals.

Introducing macroeconomic variables including shipping indices provided only minor gains in AUC-ROC, typically around 0.01–0.02 for top-performing ensemble models (Fig. 3). While these external measures occasionally picked up sudden freight-rate collapses or short-term macroeconomic shocks, their overall added value paled in comparison to firm-level financial red flags. In practical terms, a near-term default often stems from acute operational or liquidity strains: rapidly mounting obligations, depleted cash flows, or an inability to refinance short-term debt—all of which overshadow the influence of broader shipping-market conditions. For instance, LightGBM might rise from an AUC-ROC of 0.92 (financial only) to 0.93 (financial+shipping), a notably smaller improvement than observed at longer horizons. Consequently, for institutions requiring an annual outlook, short-term internal financial metrics remain the primary predictive drivers, whereas sector-level indicators have comparatively less impact.



Table 4 Comparison of bankruptcy prediction models (1-year horizon)

Model	LDA	LR	ANN	DT	SVM	RF	ADA	GB	XGB	LGBM	CAT	RNN	GRU	LSTM
Financial only	0.6616	0.6592	0.6272	0.6256	0.6437	0.7355	0.7223	0.6819	0.7775	0.8001	0.8073	0.6390	0.6032	0.5940
Financial + Shipping	0.7589	0.7757	0.7889	0.6610	0.6302	0.8347	0.7486	0.8182	0.8453	0.8512	0.8225	0.8027	0.7197	0.7504



SHAP analysis at the one-year mark supports the conclusion that internal liquidity metrics overwhelmingly determine imminent bankruptcy outcomes, irrespective of whether shipping indices are included (Table 5 and Fig. 4). In the financial-only scenario, variables like current liabilities to sales, net interest margin, and working capital to assets consistently appeared in the top three or four SHAP-ranked features across ensemble models. These ratios point to heightened vulnerability: if a shipping firm experiences sudden revenue contraction or a spike in current liabilities (e.g., interest or short-term borrowing expenses), it may be unable to cover obligations in the immediate future. In practice, this implies that one-year predictions heavily rely on micro level financial statements. While this result aligns with traditional financial theory and is to be expected, it serves as an important validation of the model fundamental soundness, confirming its ability to detect immediate liquidity crises before addressing more complex, long term structural risks.

Adding shipping-related features such as short-term freight index trend, oil price, or trade volume growth did occasionally shift minor SHAP values upward for ensembles. However, these sectoral indicators rarely surpassed the critical liquidity signals in global SHAP rankings. The small improvements observed reflect that a firm's near-term fortunes can, in isolated cases, be further compromised by abrupt freight-rate swings, yet the overwhelming determinant remains its internal capacity to meet short-term debts. In practice, this implies that one-year predictions heavily rely on micro-level financial statements, while macro-level signals serve as secondary confirmations or refinements.

4.2 Three-year horizon

Shifting to a three-year horizon broadens the scope of relevant factors, moving beyond acute liquidity stresses toward medium-range capital-structure decisions, evolving profitability trends, and moderate macroeconomic influences (Table 6). Under financial-only data, ensemble methods particularly LightGBM, CatBoost, and XGBoost consistently achieved AUC-ROC scores between 0.88 and 0.90 (Table 6). In contrast, traditional linear models remained closer to 0.77–0.79, suggesting that forecasting three years ahead calls for algorithms capable of reweighting multiple mid-range signals such as incremental debt growth or gradual declines in net income over sequential fiscal periods. Deep learning models (RNN, GRU, LSTM) performed modestly in comparison, generally failing to surpass the top ensemble scores, although they still benefited from additional macroeconomic inputs.

Once shipping indices were incorporated, performance improvements became more substantial than at the one-year mark. LightGBM and CatBoost typically rose by 0.03–0.05 in AUC-ROC, surpassing 0.93 in some cross-validation folds, and showed parallel gains in average precision (Fig. 5). This finding indicates that cyclical trade-market downturns or modestly rising interest rates over a two-to-three-year window can pressure an already leveraged shipping firm into distress. By capturing these sectoral movements, ensemble methods refined their classification thresholds for mid-term defaults. XGBoost and Random Forest also benefited reaching AUC-ROC values around 0.90–0.92 when shipping factors were added, confirming that



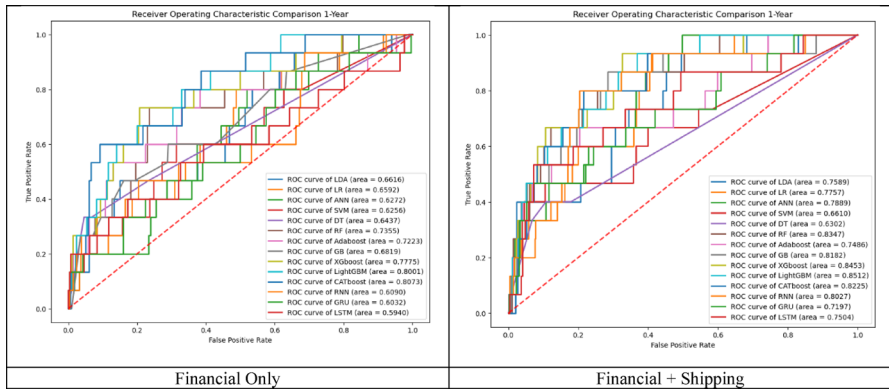


Fig. 3 ROC curve comparison of bankruptcy prediction models (1-year horizon). Source: Authors

Table 5 Top 10 SHAP feature importance (1-year horizon)

Rank	Financial only		Financial + Shipping	
	Variable	SHAP value	Variable	SHAP value
1	Current Liabilities to Sales	1.2039	Global Trade Volume Growth Rate	0.7768
2	EBITDA	0.7359	Leverage	0.7395
3	EBITDA to Interest Coverage Ratio	0.4378	Company Size	0.7080
4	Leverage	0.3412	EBITDA	0.6455
5	Current Liabilities to Total Sales	0.3313	Type of Operator	0.6311
6	Total Sales Growth Rate	0.2601	Korean Trade Volume Growth Rate	0.5942
7	Working Capital to Assets	0.2423	Gross Tonnage	0.5448
8	Total Sales	0.2355	Number of Ships Owned	0.3397
9	Net Interest Margin	0.2298	Net Income	0.3072
10	Current Assets	0.2216	Current Assets	0.3034

industrial variables play a meaningful role once the forecast horizon extends beyond a single year.

The corresponding SHAP analysis for three-year predictions revealed a more balanced interplay between firm-specific ratios and sector-wide indicators in Table 7. In the financial-only model, EBITDA emerged as the top-ranked predictor, with a particularly high average SHAP value. Other internal variables such as Current Assets, Total Sales Growth Rate, and EBITDA to Interest Coverage Ratio also featured prominently. Taken together, these highlight how firms that operate on thin margins or accumulate too much debt over multiple cycles become prime bankruptcy candidates.



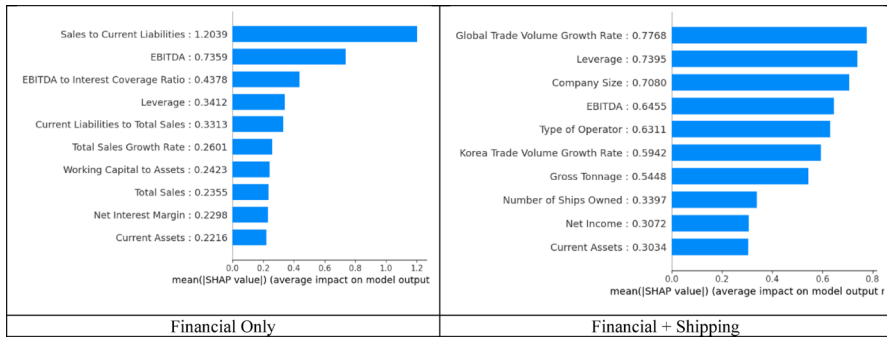


Fig. 4 Top 10 SHAP feature importance (1-year horizon). *Source:* Authors

After adding shipping indices, SHAP importance shifted noticeably toward external cyclical factors (Table 7 and Fig. 6). Company Size assumed the top position, followed closely by EBITDA, Korea Trade Volume Growth Rate, and Global Trade Volume Growth Rate. These results underscore that a mid-term bankruptcy scenario often develops when moderate financial fragility (e.g., leveraged capital structures or eroding net margins) intersects with a persistent or worsening freight market. For instance, a firm might handle rising debts for a year or two, but prolonged freight-rate declines can erode profitability faster than the company can adjust, significantly boosting default risk. Overall, the mid-term SHAP findings highlight the dual necessity of maintaining robust internal fundamentals (steady EBITDA, modest leverage) and operating in a supportive sectoral environment (stable freight demand, manageable interest rates) to avert bankruptcy over a multi-year span.

4.3 Five-year horizon

In the five-year horizon (Table 8), prolonged capital commitments, sustained freight-market cycles, and extended macroeconomic influences become critical, which making external shipping indices far more influential than in either the one-year or three-year contexts. Under financial-only conditions, gradient boosting methods (LightGBM, CatBoost) continued to outperform linear and simpler tree-based models, typically reaching AUC-ROC values of 0.88–0.90. Traditional classifiers tended to hover around 0.75–0.80 at this horizon, reflecting their reduced capacity to capture long-term, non-linear interactions among leverage, earnings, and extended financing costs. Although internal indicators such as high debt service costs, prolonged erosion in profitability can reveal structural weaknesses even five years out, they offer an incomplete view of how macro-level shipping demand and interest rates may evolve to stress capital rollovers and refinancing options over multiple years. It is important to note that high debt levels remain a fundamental vulnerability. However, in the five-year horizon, the relative importance of shipping indices surges because these external factors determine a firm capacity to refinance and service that debt. As noted by Ivanov (2021), long term viability is often dictated by adaptation to external shocks, explaining why macro indicators gain predictive weight over extended periods.



Table 6 Comparison of bankruptcy prediction models (3-year horizon)

Model	LDA	LR	ANN	DT	SVM	RF	ADA	GB	XGB	LGBM	CAT	RNN	GRU	LSTM
Financial Only	0.7762	0.7810	0.7639	0.7579	0.5944	0.9135	0.8180	0.8714	0.9142	0.8842	0.8906	0.7744	0.7479	0.7202
Financial+Shipping	0.8460	0.8596	0.8728	0.8704	0.7932	0.9422	0.8700	0.9397	0.9527	0.9567	0.9513	0.8941	0.8752	0.8326



Once shipping indices were added, performance gaps became the largest observed across all horizons (Fig. 7). LightGBM and CatBoost frequently achieved AUC-ROC scores above 0.92, with some folds reaching 0.94 or higher. This 0.03–0.06 improvement over financial-only baselines underscores how protracted global trade market downturns, vessel oversupply, or climbing long-term interest rates can eventually tip a firm into default—especially if its capital structure lacks resilience. XGBoost remained competitive (0.90–0.92) but did not consistently match the highest CatBoost or LightGBM scores, potentially due to differences in handling categorical features or default hyperparameters. Random Forest also improved upon its financial-only baseline but typically lagged the near-peak metrics of specialized gradient boosting frameworks.

Five-year SHAP analysis revealed a notable shift in top features toward macro- and shipping-specific factors, alongside a few enduring internal stability indicators as shown in Table 9 and Fig. 8 (e.g., Equity to Assets, EBITDA to Interest Coverage). Under financial-only models, these internal metrics remained consistently among the highest-ranked SHAP features, confirming the critical role of steady cash flows and prudent leverage. However, after incorporating shipping indices, external measures such as LIBOR, and operator type appeared in the top five for prediction model, sometimes overtaking purely internal signals in their aggregated SHAP values.

This pattern reflects how cyclical or structural shifts in the shipping sector accumulate over multiple years. If freight indices remain depressed year after year, even a firm with initially healthy financials may face mounting pressure to refinance debt or sustain cash flows, eventually leading to insolvency. Conversely, strong internal ratios may help a company to recover from the short-term slump, but persistent macroeconomic downturns or adverse interest-rate environments can erode solvency over longer periods. Consequently, five-year SHAP insights reinforce the idea that sustained macroeconomic pressures and sector-specific instabilities become decisive in determining whether a shipping enterprise transitions from moderate financial strain to outright bankruptcy.

5 Discussion

5.1 Variation by forecasting horizon

The findings illustrate how different forecasting horizons, specifically one year, three years, and five years, demand distinct modeling strategies. In the one-year context, shipping companies typically fail due to sudden liquidity constraints and abrupt revenue shortfalls. These developments overshadow any incremental effect of macroeconomic or shipping-specific factors. Consequently, models that focus on short-term ratios, including current liabilities to sales or net interest margin, tend to dominate one-year prediction.

In contrast, three-year horizons encompass medium-range vulnerabilities that accumulate over several fiscal periods. These include steadily increasing leverage, long-term profitability declines, and moderate freight-rate fluctuations. Ensemble models, such as CatBoost, LightGBM, and XGBoost, gain accuracy once they inte-



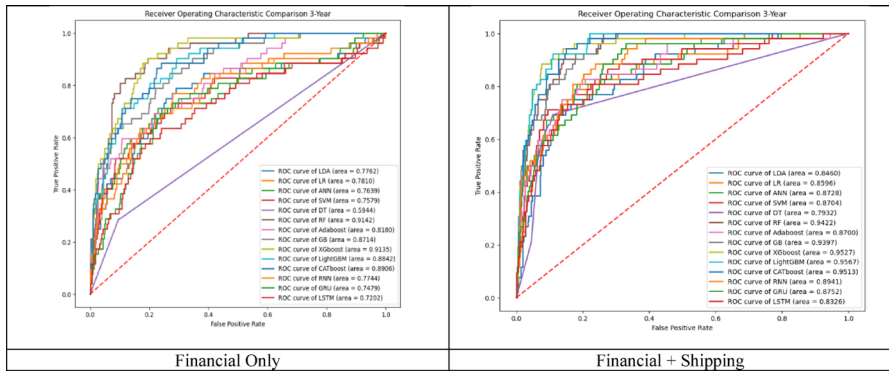


Fig. 5 ROC curve comparison of bankruptcy prediction models (3-year horizon). Source: Authors

Table 7 Top 10 SHAP feature importance (3-year horizon)

Model Rank	Financial only		Financial + Shipping	
	Variable	SHAP Value	Variable	SHAP Value
1	EBITDA	3.6848	Company Size	2.4987
2	Current Assets	1.3605	EBITDA	1.8071
3	Total Sales Growth Rate	1.3319	Korea Trade Volume Growth Rate	1.5832
4	Total Sales	0.9954	Total Sales Growth Rate	1.4813
5	EBITDA to Interest Coverage Ratio	0.9923	Global Trade Volume Growth Rate	1.4347
6	Total Assets Growth Rate	0.8564	Type of Operator	1.2840
7	Current Liabilities to Sales	0.7234	Net Interest Margin	1.0699
8	Gross Profit to Total Sales	0.6475	Total Sales	1.0487
9	Shareholders' Equity	0.5462	Current Assets	1.0047
10	Net Interest Margin	0.4896	Gross Profit to Total Sales	0.9319

grate shipping indices or trade-volume growth rates, confirming that moderate external shifts become significant predictors when observed over multiple years.

By five years, persistent market cycles and extended refinancing risks emerge in full. Models that include shipping-specific indicators capture protracted freight-rate weakness, rising interest rates, and other procyclical factors that may eventually erode capital structures. Traditional methods based on linear assumptions are less capable of representing these extended nonlinear relationships. Instead, gradient boosting models leverage both internal and external data to identify slow-building signs of distress that can lead to eventual default.



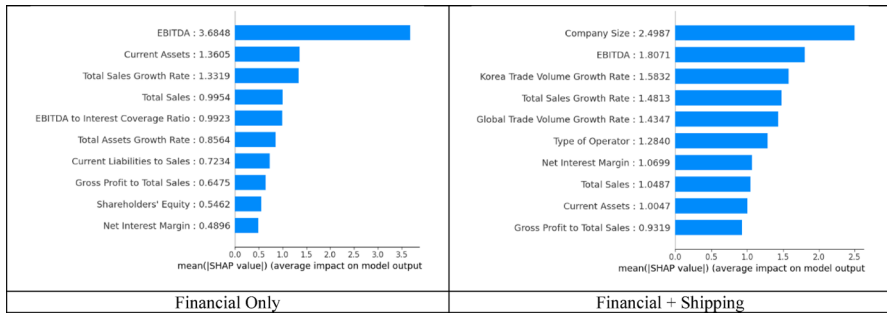


Fig. 6 Top 10 SHAP feature importance (3-year horizon). *Source:* Authors

It is important to note that deep learning architectures, including RNN, GRU, and LSTM, did not provide significantly better results than boosting models in this study. One plausible reason is the relatively small and imbalanced dataset, which may prevent deep networks from learning sufficiently complex patterns. Furthermore, the inclusion of data through 2023 ensures that the models have learned from recent extreme volatility, such as the COVID 19 pandemic, reinforcing the importance of supply chain resilience in the face of global disruptions (Chowdhury et al. 2021). The cyclical nature of shipping may also present irregular temporal shifts that limit the advantages of sequence-based neural models. As a result, advanced ensemble methods exhibit robust performance across time horizons, while deep learning approaches remain more sensitive to data scale and cyclical volatility.

5.2 Role of macroeconomic variables and shipping indices

Although internal financial indicators such as liquidity and leverage remain central at all horizons, macroeconomic and shipping-specific factors become more influential after the first year. In the one-year forecast, imminent liquidity crises—including surges in short-term debt or sudden revenue declines—tend to overshadow moderate variations in global trade or freight rates. Consequently, macro-level measures offer only minor predictive gains for near-term defaults.

By contrast, three-year and five-year forecasts show a clearer impact from shipping-market cycles and interest-rate fluctuations, which can erode profitability over multiple periods. In particular, long-term financing pressures such as rising interest rates or currency shifts frequently appear among the top SHAP-ranked variables, underscoring that moderately leveraged firms face compounding stress over multiple fiscal years. Surprisingly, freight-rate indices (for example, the Baltic Dry Index or container freight benchmarks) do not consistently rank among the highest predictors, even when external variables become more influential. Instead, indicators like overall trade volume growth, operator type, or interest rates often exert a stronger effect on long-horizon insolvency risk. This suggests that, although freight rates reflect sector cycles, other macroeconomic and shipping-specific measures can overshadow them in capturing the extended financial distress trajectories evident in the dataset.



Table 8 Comparison of bankruptcy prediction models (3-year horizon)

Model	LDA	LR	ANN	DT	SVM	RF	ADA	GB	XGB	LGBM	CAT	RNN	GRU	LSTM
Financial only	0.7599	0.7416	0.6913	0.7147	0.7623	0.8787	0.7806	0.8482	0.8806	0.8892	0.8665	0.7772	0.7726	0.7494
Financial+Shipping	0.8320	0.8453	0.8548	0.8567	0.7487	0.9044	0.8641	0.8949	0.9184	0.9265	0.9048	0.8610	0.8773	0.8778



5.3 Practical implications

These findings underscore the need to align analytical focus with the specific forecasting window. In the one-year context, abrupt liquidity signals usually dictate default risk, while three- and five-year forecasts benefit substantially from incorporating shipping and interest-rate indicators. Industry practitioners, such as creditors, investors, and policymakers, can use these insights to prioritize the data most relevant to their time horizon. For example, lenders worried about near-term repayments may concentrate on current liabilities to assets or net interest margin, whereas investors contemplating multi-year commitments should emphasize shipping indices, trade volumes, and possible increases in refinancing costs.

The study also reveals the consistent superiority of advanced ensemble methods in imbalanced shipping data. LightGBM, CatBoost, and XGBoost frequently outperform traditional linear models, which are prone to underestimating complex interactions, as well as deep neural networks, which are sensitive to data scale and cyclical volatility. Furthermore, SHAP analysis clarifies why each feature matters and how its importance shifts over time. This transparency supports targeted interventions. For instance, a firm facing large short-term liabilities might restructure high-cost debt immediately, while another with moderate liquidity but high long-term leverage might focus on hedging strategies against prolonged freight-market weakness.

Overall, these insights align with the capital-intensive and cyclical nature of the shipping industry. Firms that maintain conservative capital structures and track sector-level trends, such as freight demand and interest rates, are better positioned to manage periodic downturns. The lessons learned here may extend to other industries with similar characteristics, including airlines and energy, where significant capital requirements and shifting demand cycles make accurate multi-horizon bankruptcy forecasts essential for effective risk management.

6 Conclusion

This study compared a comprehensive range of machine learning models, including traditional linear approaches, advanced ensemble techniques, and deep neural architectures, for predicting bankruptcy in the shipping industry over one-year, three-year, and five-year horizons. Despite the challenges posed by a relatively small and highly imbalanced dataset, careful data processing, including the use of SMOTE, enabled a reliable evaluation of model performance.

The results indicate that ensemble learners such as LightGBM, CatBoost, and XGBoost generally outperform linear models, which struggle to capture the complex and often nonlinear dynamics that characterize shipping firms. Although deep learning architectures, including RNN, GRU, and LSTM, demonstrate moderate predictive power, they do not exceed the accuracy of top-performing boosting methods. The cyclical nature of shipping, together with limited sample size, may reduce the potential benefits of these more complex neural approaches.

In shorter horizons of 1-year, abrupt liquidity problems and sudden cost or revenue fluctuations emerge as primary triggers of insolvency. Incorporating macro-



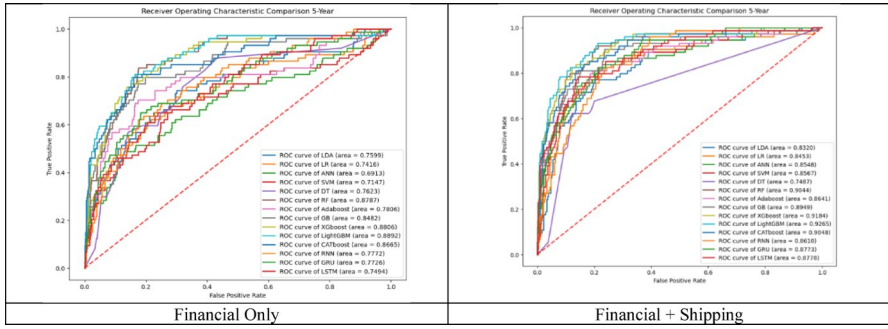


Fig. 7 ROC curve comparison of bankruptcy prediction models (3-year horizon). Source: Authors

Table 9 Top 10 SHAP feature importance (5-year horizon)

Rank	Financial only		Financial + Shipping	
	Variable	SHAP value	Variable	SHAP value
1	EBITDA	1.7841	Company Size	3.4431
2	Current Assets	1.5858	LIBOR	1.8254
3	Equity to Assets	1.1438	Type of Operator	1.6482
4	Gross Profit to Total Sales	1.1132	Gross Profit to Total Sales	1.2321
5	Total Sales Growth Rate	1.1015	EBITDA	1.1106
6	EBITDA to Interest Coverage Ratio	1.0931	GDP	1.0178
7	Shareholders' Equity	1.0871	Shareholders' Equity	0.8706
8	Total Sales	1.0667	Total Sales Growth Rate	0.8578
9	Current Liabilities to Sales	1.0229	Current Assets	0.8403
10	Net Income	1.0110	Net Income	0.8203

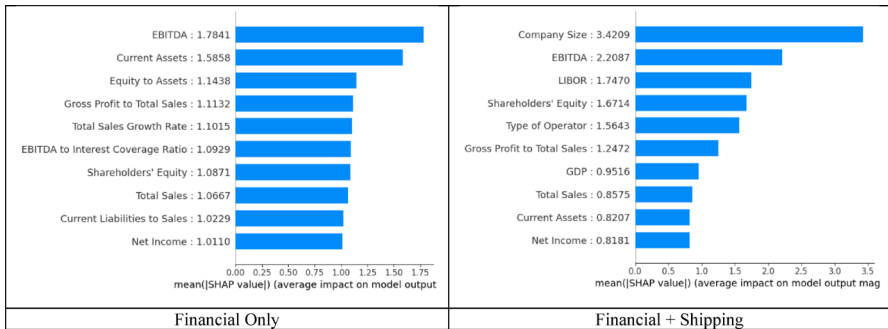


Fig. 8 Bar chart of top 10 SHAP feature importance (5-year horizon). Source: Authors



economic and shipping-related variables into these short-term models produces only slight improvements in predictive accuracy, likely because external sector changes have little time to exert a cumulative impact. However, as the forecast window extends to three and five years, indicators of global shipping demand and interest-rate movements become increasingly influential. The inclusion of shipping indices substantially raises classification performance under longer horizons, indicating that sustained market downturns or shifting financing conditions can critically affect firms with leveraged capital structures.

From a practical perspective, these findings emphasize the need to tailor bankruptcy prediction methodologies to the specific forecast horizon. Companies, creditors, and policymakers concerned with imminent default should focus on financial ratios tied to short-term liquidity, whereas those interested in medium- or long-term solvency risk should prioritize macroeconomic and shipping-specific features. Researchers can extend these conclusions by enlarging datasets, testing more advanced hybrid algorithms, and applying these methods to other capital-intensive and cyclical sectors, such as airlines or energy.

References

- Alam, N., J. Gao, and S. Jones. 2021. Corporate failure prediction: An evaluation of deep learning vs discrete hazard models. *Journal of International Financial Markets, Institutions and Money* 75 : 101455.
- Altman, E. I. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance* 23(4): 589–609.
- Altman, E. I., M. Iwanicz-Drozdzowska, E. K. Laitinen, and A. Suvas. 2020. A race for long horizon bankruptcy prediction. *Applied Economics* 52(37): 4092–4111.
- Atkinson, A. C., M. Riani, A. Corbellini. 2021. The box–cox transformation: Review and extensions.
- Bishop, C. M. 2006. *Pattern recognition and machine learning*. Springer.
- Breiman, L. 2001. Random forests. *Machine Learning* 45:5–32.
- Cha, S., and J. Kang. 2018. Corporate default prediction model using deep learning time series algorithm, RNN and LSTM. *Journal of Intelligence and Information Systems* 24(4): 1–32.
- Chawla, N. V., K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. 2002. SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research* 16:321–357.
- Chen, T., C. Guestrin. 2016. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* 785–794.
- Chowdhury, P., S. K. Paul, S. Kaisar, and M. A. Moktadir. 2021. COVID-19 pandemic related supply chain studies: A systematic review. *Transportation Research Part e: Logistics and Transportation Review* 148 : 102271.
- Ciampi, F., and N. Gordini. 2013. Small enterprise default prediction modeling through artificial neural networks: An empirical analysis of Italian small enterprises. *Journal of Small Business Management* 51(1): 23–45.
- Clintworth, M., D. Lyridis, and E. Boulougouris. 2023. Financial risk assessment in shipping: A holistic machine learning based methodology. *Maritime Economics & Logistics* 25(1): 90–121.
- Dey, R., & F. M. Salem. 2017. Gate-variants of gated recurrent unit (GRU) neural networks. In *2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS)* 1597–1600. IEEE.
- Fawcett, T. 2006. An introduction to ROC analysis. *Pattern Recognition Letters* 27(8): 861–874.
- Fernández, A., S. Garcia, F. Herrera, and N. V. Chawla. 2018. SMOTE for learning from imbalanced data: Progress and challenges, marking the 15-year anniversary. *Journal of Artificial Intelligence Research* 61:863–905.
- Friedman, J. H. 2001. Greedy function approximation: A gradient boosting machine. *Annals of Statistics*. <https://doi.org/10.1214/aos/1013203451>.



- García, S., S. Ramírez-Gallego, J. Luengo, J. M. Benítez, and F. Herrera. 2016. Big data preprocessing: Methods and prospects. *Big Data Analytics* 1:1–22.
- García, V., A. I. Marqués, and J. S. Sánchez. 2019. Exploring the synergetic effects of sample types on the performance of ensembles for credit risk and corporate bankruptcy prediction. *Information Fusion* 47:88–101.
- Geng, R., I. Bose, and X. Chen. 2015. Prediction of financial distress: An empirical study of listed Chinese companies using data mining. *European Journal of Operational Research* 241(1): 236–247.
- Gers, F. A., J. Schmidhuber, and F. Cummins. 2000. Learning to forget: Continual prediction with LSTM. *Neural Computation* 12(10): 2451–2471.
- Ghosh, A., and S. Kapil. 2022. Is Altman's model efficient in predicting bankruptcy? A comparison among the Altman Z-score, DEA, and ANN models. *Journal of Information and Optimization Sciences* 43(6): 1191–1207.
- Gogas, P., T. Papadimitriou, and A. Agrapetidou. 2018. Forecasting bank failures and stress testing: A machine learning approach. *International Journal of Forecasting* 34(3): 440–455.
- Haider, J., Z. Ou, and S. Pettit. 2019. Predicting corporate failure for listed shipping companies. *Maritime Economics & Logistics* 21:415–438.
- Haixiang, G., L. Yijing, J. Shang, G. Mingyun, H. Yuanyue, and G. Bing. 2017. Learning from class-imbalanced data: Review of methods and applications. *Expert Systems with Applications* 73:220–239.
- He, H., and E. A. García. 2009. Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering* 21(9): 1263–1284.
- Ivanov, D. 2021. Supply chain viability and the COVID-19 pandemic: A conceptual and formal generalisation of four major adaptation strategies. *International Journal of Production Research* 59(12): 3535–3552.
- Jones, S. 2017. Corporate bankruptcy prediction: A high dimensional analysis. *Review of Accounting Studies* 22:1366–1422.
- Kavussanos, M. G., and D. A. Tsouknidis. 2016. Default risk drivers in shipping bank loans. *Transportation Research Part e: Logistics and Transportation Review* 94:71–94.
- Ke, G., Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, and T. Y. Liu. 2017. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems* 30:65.
- Kheneifar, M. A., and B. Amiri. 2025. A novel hybrid model for loan default prediction in maritime finance based on topological data analysis and machine learning. *IEEE Access* 13:1–15.
- Kou, G., Y. Xu, Y. Peng, F. Shen, Y. Chen, K. Chang, and S. Kou. 2021. Bankruptcy prediction for SMEs using transactional data and two-stage multiobjective feature selection. *Decision Support Systems* 140 : 113429.
- Kumar, P. R., and V. Ravi. 2007. Bankruptcy prediction in banks and firms via statistical and intelligent techniques—A review. *European Journal of Operational Research* 180(1): 1–28.
- Kutner, M. H., C. J. Nachtsheim, J. Neter, and W. Li. 2005. *Applied linear statistical models*. McGraw-hill.
- Le, T. 2022. A comprehensive survey of imbalanced learning methods for bankruptcy prediction. *IET Communications* 16(5): 433–441.
- Le, T., L. Hoang Son, M. T. Vo, M. Y. Lee, and S. W. Baik. 2018. A cluster-based boosting algorithm for bankruptcy prediction in a highly imbalanced dataset. *Symmetry* 10(7) : 250.
- Li, M. Y. L., and P. Miu. 2010. A hybrid bankruptcy prediction model with dynamic loadings on accounting-ratio-based and market-based information: A binary quantile regression approach. *Journal of Empirical Finance* 17(4): 818–833.
- Lin, Y. C., R. Padliansyah, Y. H. Lu, and W. R. Liu. 2025. Bankruptcy prediction: Integration of convolutional neural networks and explainable artificial intelligence techniques. *International Journal of Accounting Information Systems* 56 : 100744.
- Lozinskaia, A., A. Merikas, A. Merika, and H. Penikas. 2017. Determinants of the probability of default: The case of the internationally listed shipping corporations. *Maritime Policy & Management* 44(7): 837–858.
- Mitroussi, K., W. Abouarghoub, J. J. Haider, S. J. Pettit, and N. Tigka. 2016. Performance drivers of shipping loans: An empirical investigation. *International Journal of Production Economics* 171:438–452.
- Molina, C. A., and L. A. Preve. 2012. An empirical analysis of the effect of financial distress on trade credit. *Financial Management* 41(1): 187–205.
- Ohlson, J. A. 1980. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*. <https://doi.org/10.2307/2490395>.
- Park, S., H. Kim, J. Kwon, and T. Kim. 2021. Empirics of Korean shipping companies' default predictions. *Risks* 9(9): 159.



- Perboli, G., and E. Arabnezhad. 2021. A machine learning-based DSS for mid and long-term company crisis prediction. *Expert Systems with Applications* 174 : 114758.
- Prokhorenkova, L., G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin. 2018. CatBoost: Unbiased boosting with categorical features. *Advances in Neural Information Processing Systems* 31:65.
- Rumelhart, D. E., G. E. Hinton, and R. J. Williams. 1986. Learning representations by back-propagating errors. *Nature* 323(6088): 533–536.
- Saito, T., and M. Rehmsmeier. 2015. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS ONE* 10(3) : e0118432.
- Saltelli, A., S. Tarantola, F. Campolongo, and M. Ratto. 2004. *Sensitivity analysis in practice: A guide to assessing scientific models*, vol. 1. New York: Wiley.
- Schmidhuber, J., and S. Hochreiter. 1997. Long short-term memory. *Neural Computation* 9(8): 1735–1780.
- Shrivastava, S., P. M. Jeyanthi, and S. Singh. 2020. Failure prediction of Indian Banks using SMOTE, Lasso regression, bagging and boosting. *Cogent Economics & Finance* 8(1): 1729569.
- Sizan, M. M. H., et al. 2025. Bankruptcy prediction for US businesses: Leveraging machine learning for financial stability. *Journal of Business and Management Studies* 7(1): 1–12.
- Snoek, J., H. Larochelle, and R. P. Adams. 2012. Practical bayesian optimization of machine learning algorithms. *Advances in Neural Information Processing Systems* 25:865.
- Song, D. W., Y. J. Seo, and D. W. Kwak. 2019. Learning from Hanjin Shipping’s failure: A holistic interpretation on its causes and reasons. *Transport Policy* 82:77–87.
- Stopford, M. 2013. Shipping market cycles. In *The handbook of maritime economics and business* 265–288. Informa Law from Routledge.
- Sun, J., H. Li, H. Fujita, B. Fu, and W. Ai. 2020. Class-imbalanced dynamic financial distress prediction based on Adaboost-SVM ensemble combined with SMOTE and time weighting. *Information Fusion* 54:128–144.
- Sun, L. 2007. A re-evaluation of auditors’ opinions versus statistical models in bankruptcy prediction. *Review of Quantitative Finance and Accounting* 28:55–78.
- Tsai, C. F., Y. F. Hsu, and D. C. Yen. 2014. A comparative study of classifier ensembles for bankruptcy prediction. *Applied Soft Computing* 24:977–984.
- Vapnik, V. 1995. Support-vector networks. *Machine Learning* 20:273–297.
- Wang, T., S. Zhao, G. Zhu, and H. Zheng. 2021. A machine learning-based early warning system for systemic banking crises. *Applied Economics* 53(26): 2974–2992.
- Yeo, H. 2016. Solvency and liquidity in shipping companies. *The Asian Journal of Shipping and Logistics* 32(4): 235–241.
- Yu, B., Y. Li, and Y. Shi. 2020. Cooperation mechanism of the shipping SMEs based on field theory. *Journal of CoaStal Research* 112(SI): 472–476.
- Yuen, S. S., and A. H. Ko. 2018. Corporate governance in shipping logistics industry: A practitioner’s perspectives. *Journal of Business & Economic Policy* 5(4): 63.
- Zhang, G., M. Y. Hu, B. E. Patuwo, and D. C. Indro. 1999. Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis. *European Journal of Operational Research* 116(1): 16–32.
- Zhang, Y., R. Liu, A. A. Heidari, X. Wang, Y. Chen, M. Wang, and H. Chen. 2021. Towards augmented kernel extreme learning models for bankruptcy prediction: Algorithmic behavior and comprehensive analysis. *Neurocomputing* 430:185–212.
- Zhou, L., and K. K. Lai. 2017. AdaBoost models for corporate bankruptcy prediction with missing data. *Computational Economics* 50:69–94.
- Zoričák, M., P. Gnip, P. Drotár, and V. Gazda. 2020. Bankruptcy prediction for small-and medium-sized companies using severely imbalanced datasets. *Economic Modelling* 84:165–176.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

