International Journal of Social Science and Economic Research

ISSN: 2455-8834

Volume:10, Issue:06 "June 2025"

Applying Data Analytics to Financial Risk Management: A Study of Listed Firms on Vietnam's Stock Market

Dinh The Hung[1], Nguyen Quynh Anh[2] and Nguyen Minh Thu[3]

[1]National Economics University, Vietnam

[2]Huynh Thuc Khang High School, Vietnam

[3]Foreign Language Specialized school, Vietnam

DOI: 10.46609/IJSSER.2025.v10i06.016 URL: https://doi.org/10.46609/IJSSER.2025.v10i06.016

Received: 3 June 2025 / Accepted: 24 June 2025 / Published: 30 June 2025

ABSTRACT

This study identifies key factors significantly impacting financial risk in listed manufacturing enterprises on the Vietnamese stock exchange, using the Least Absolute Shrinkage and Selection Operator (LASSO) variable selection method and constructing a machine learning model to predict financial risk. The research results show that the LASSO model identified five financial indicators influencing the financial risk faced by enterprises over five years (2020-2024), including: Short-term solvency ratio (Current Assets/Short-term Liabilities); Total asset turnover ratio (Revenue/Total Assets); ROA (Net Profit/Total Assets); Long-term Debt-to-Total Assets ratio; and Short-term Debt-to-Total Liabilities ratio. The Artificial Neural Network (ANN) model, when combined with the LASSO method for selecting appropriate predictive variables, enhances the performance of forecasting models compared to not performing variable selection for financial risk assessment. Based on these findings, the research team proposed several solutions and policies for businesses.

Keywords: Financial risk, Financial risk management, Machine learning, Data analysis, Predictive analysis

1. Introduction

Unpredictable global economic fluctuations, coupled with high inflation following prolonged monetary easing and geopolitical instability from conflicts such as the war in Ukraine, have increased bankruptcy risks for businesses since the COVID-19 pandemic. Operating in a fiercely competitive and volatile international environment, manufacturing enterprises face numerous risks, particularly financial risks. Although financial risk management in Vietnamese businesses

ISSN: 2455-8834
Volume:10, Issue:06 "June 2025"

has seen positive developments, these improvements remain superficial rather than deep-rooted, which could lead to negative repercussions for the economy and society due to spillover effects on supply chain partners or declining incomes for workers. Therefore, accurate bankruptcy risk prediction has attracted significant attention from organizations, businesses, and policymakers to monitor and provide early warnings about corporate financial health, enabling better management, investment, and regulatory decisions.

In the Industry 4.0 era, information and data are considered valuable resources, making them crucial for effective risk management (Vuong, 2023). By providing accurate and comprehensive data, businesses can utilize information technology and modern risk management models to better identify and mitigate risks.

Corporate bankruptcy prediction models, with the pioneering work of Beaver (1966) and Altman (1968), have developed from a very early age. These traditional models relied on discriminant analysis or non-linear logistic regression, using financial indicators as predictors (Jones & Hensher, 2004; Tian et al, 2015). The advancement of intelligent computational models and technologies with enhanced capabilities for processing complex algorithms has led to the development and application of machine learning-based techniques in bankruptcy prediction (Le & Viviani, 2018; Chen et al, 2019). Compared to traditional statistical methods, machine learning models have demonstrated superior performance, enabling effective analysis of nonlinear relationships as well as highly complex problems without imposing stringent data requirements.

This study evaluates the application of LASSO-based variable selection to enhance bankruptcy risk prediction for Vietnamese firms. This study evaluates the application of LASSO-based variable selection to enhance bankruptcy risk prediction for Vietnamese firms. Experimental results from 284 listed companies on the Vietnamese stock market (2020-2024) demonstrate that integrating LASSO for choosing suitable predicted variables improves predictive model performance compared to non-selection approaches.

2. Literature Review

Intelligent models emerged relatively early in predictive analytics, with neural networks (NN) being the first to be developed and dominating the field throughout the 1990s (Wilson & Sharda, 1994). NNs gained widespread adoption due to their flexibility - they impose no statistical assumptions and can effectively process nonlinear relationships. This capability proves particularly valuable in bankruptcy prediction, as the relationship between insolvency risk and explanatory variables is often inherently nonlinear (Barboza et al, 2017).

ISSN: 2455-8834

Volume:10, Issue:06 "June 2025"

In their comparative study of traditional and intelligent models, Zhao and colleagues (2009) evaluated the predictive performance of logit models, neural networks (NN), and K-nearest neighbors (KNN) using financial ratios for corporate bankruptcy forecasting. Their experimental results demonstrated NN's superior predictive accuracy. Similarly, the comparative analysis of Barboza et al. (2017) revealed that random forest and AdaBoost models significantly outperformed both logistic regression and discriminant analysis in prediction reliability. These conclusions align with subsequent research by Heo and Yang (2014), Xiao et al (2016).

In the Vietnamese context, Dao Trong Thinh and Doan Van Toan (2016) successfully implemented neural networks for credit scoring classification, demonstrating their practical utility in enhancing credit risk management for bank lending operations. Complementing this research, Doan Khanh Hung and Tran Thi Hien (2019) established that financial structure significantly impacts financial risk exposure, with optimal capital allocation serving as an effective risk mitigation strategy for enterprises.

Huang et al (2021) used Random Forest (RF), Support Vector Machine (SVM), and AdaBoost to analyze business risk. The investigation’s results show that 3 types of machine learning algorithms prove to be effective in evaluating risk and have high veracity in judging a business’s financial risk. Murugan (2023) took analytics and used data on a large scale, as well as used machine learning strategies: K-nearest neighbor (KNN), logistic regression (LR), and XGBoost to predict the risk of default and related incidents in lending. This study also mentioned the combination of new technologies (IOT) and machine learning to deploy risk management solutions. This leads to the potential of new technologies in financial risk management and prediction.

The study of Le Hai Trung and Truong Thi Thuy Duong (2023) applied the LASSO model to select financial indexes that have a direct impact on that firm’s probability of bankruptcy. The results show that the LASSO model improves bankruptcy prediction of listed companies on the Vietnamese stock market.

Nguyen Thi Thanh Loan et al (2024) applied the LASSO model to select financial indexes that have a direct impact on that firm’s financial risk management. The results show that the LASSO model manages financial risk of listed companies on the Vietnamese stock market.

The study of Orcun Sarioguz, Evin Miser (2024) focuses on applying machine learning models in risk management and the way these factors affect the predictability and follow relevant rules. Many financial organizations use some advanced AI technologies in their decision-making ability so a clear understanding of their effectiveness and the significance of legal compliance has become essential for their development. The study presents a comprehensive assessment

ISSN: 2455-8834

Volume:10, Issue:06 "June 2025"

about traditional risk management approaches compared to new AI-based methods by thoroughly assessing standard hard metrics, including accuracy and recall ability. The study shows that applying AI methods improves the effectiveness of prediction in a high level and the major urgent legal concerns faced by organizations. Moreover, studies show fields have an advantage to apply machine learning in operation risk management and provide AI user guidance. To improve and reinforce risk management methods and ensure strict adherence to both current and future implementation regulations, this study provides information relevant to current discourse about the future of finance in the context of increasing technological advancements.

Iswarya Konasani (2025) has demonstrated that the integration of artificial intelligence into financial risk management represents a transformative leap for the industry, revolutionizing conventional approaches to credit scoring, risk assessment, and due diligence. The study considered the implementation of AI-powered systems across financial institutions, revealing substantial improvements in operational efficiency, decision-making accuracy, and risk prediction capabilities. This study explores the development of machine learning models, advanced analytics, and automated systems that have enhanced portfolio management effectiveness, regulatory compliance processes and accessibility of inclusive financial services. Through detailed analysis of technical implementations, the study addresses current operational challenges, emerging technologies (quantum computing, NLP applications) and practical deployment considerations. The findings conclusively demonstrate how AI-driven solutions are: reshaping risk management frameworks, creating novel possibilities for financial decision-making, establishing new benchmarks for predictive performance.

3. Application of the LASSO method in improving the efficiency of prediction models

The predictive ability of both traditional and intelligent machine learning models depends on the selection of explanatory variables to incorporate into the model. Previous studies used to select variables based on authors' subjective judgments without establishing a standardized set of predictors for specific datasets or enterprise groups (Tian et al, 2015). To address this limitation, recent research has adopted data-driven variable selection techniques to identify statistically significant predictors and enhance forecasting accuracy. Among these methods, LASSO is a regression that applies a penalty function to automatically eliminate irrelevant variables by shrinking their regression coefficients to zero (Tibshirani, 1996).

In corporate bankruptcy prediction, several studies have demonstrated that combining LASSO with traditional and intelligent forecasting models improves predictive accuracy while reducing error rates. One of the pioneering studies was conducted by Tian et al (2015). Analyzing a large dataset of over 17,000 companies from 1980 to 2009 with nearly 40 explanatory variables, the authors proved that LASSO-optimized models significantly outperformed conventional discriminant analysis and logistic regression in predictive performance.

Then, Shrivastava et al (2020) demonstrated the effectiveness of LASSO when integrated with intelligent models for bankruptcy prediction in Indian firms. Similarly, by applying attribute selection methods to bankruptcy forecasting using data from all Norwegian small and medium-sized enterprises (2006-2017), Paraschiv et al (2021) demonstrated that LASSO-selected variables substantially enhanced model performance.

4. Building research models and methods.

4.1 Building research models.

This study uses the Z-Credit Scoring Model to evaluate the bankruptcy possibility of firms and then build dependent variable classifying businesses with two attributes (risk and no risk). After that, using LASSO model with the aim of selecting 5 important independent variables has a big effect on dependent variable out of the total of 35 independent variables proposed by our group from the collection of the dataset and artificial neural network to forecast and categorize enterprises as either risky or non-risky and then use confusion matrix to evaluate the performance of predictive model.

4.1.1 Z-score model build dependent variable.

Z-score presents the bankruptcy possibility of a company in finance. The Z-score value is a numerical measure used in statistics to determine the relationship between a data point's value and the dataset's mean, expressed in terms of standard deviations from the mean. Indicators employed in the computational formula collected easily from the business's annual financial statement. The Z-score model was first introduced by Edward Altman. Then, Altman developed variants of this model for other fields.

The Z-score model has the form:

$$\mathbf{Z = 1{,}2*X1 + 1{,}4*X2 + 3{,}3*X3 + 0{,}6*X4 + 1{,}0*X5}$$

With: X1: Working Capital to Total Assets Ratio (Working Capitals/Total Assets);

X2: Retained Earnings Ratio (Retained Earnings/Total Assets);

X3: Earnings Before Interest and Taxes (EBIT) to Total Assets Ratio (EBIT/Total Assets);

X4: Market Value of Total Equity/Book Values of Total Liabilities

ISSN: 2455-8834

Volume:10, Issue:06 "June 2025"

X5: Total Asset Turnover Ratio (Sales/Total Assets).

The classifications of businesses:

+ Z >= 1,81: The company is in the safe zone and is not at risk of bankruptcy.

+ Z < 1,81: The company is in the danger zone with high bankruptcy risk.

4.1.2 The LASSO model selects independent variables.

Tibshirani (1996) introduced and developed the LASSO model to select explanatory variables that have a high correlation with explanatory variables described in the prediction model.

The LASSO model has the form:

$$\sum_{i=1}^{n}\left(\left(-Y_{i,t}\left(\beta_0+\sum_{k=1}^{p}x_{i,t-1,k}\beta_k\right)\right)+\log(1+\ e^{\beta_0+\sum_{k=1}^{p}x_{i,t-1,k}\beta_k}\right)$$

With condition:

$$\sum_{k=1}^{p}\omega_k\,|\beta_k|<\lambda$$

That is: $y_{i,t}$ is a binary variable that represents the status of business i at time t; $x_{i,t-1,k}$ are the finance credits number k of business i at time t-1; n is the number of companies in the data.

The conditional function of LASSO has the form:

$$\sum_{k=1}^{p}\omega_k\,|\beta_k|<\ \lambda$$

Represents this function under the condition that the estimated value parameters are limited by the conversion ratio λ. As the value (λ) gets smaller, the number of explanatory variables

ISSN: 2455-8834

Volume:10, Issue:06 "June 2025"

retained in the prediction model will decrease (Le Hai Trung et al, 2023). This function is also called "L1". LASSO will automatically give the estimated value of insignificant variables to 0 and shrink the estimated value of lower-meaning explanatory variables to a smaller value. Another benefit of LASSO is to deal with the problem of multicollinearity between explanatory variables. This will largely support the prediction model because financial risk forecasting usually uses many highly correlated financial variables (Tian et al, 2015).

4.1.3 Artificial Neural Network forecasts businesses having financial risk

Artificial Neural Network (ANN) is a calculation model that simulates human neural network function. Each neuron is a unit that processes information and is a basic factor of a neural network. Foundational factors of a neural network include:

- Set of inputs: input signals of a neuron - these signals are often put in the form of vectors with m directions.

- Set of links: Each link is represented by a weighting factor called a synaptic weight. Synaptic weight between signal number j with neural k is usually symbolized as w_{jk}. Often, these synaptic weights are randomly initialized at the time when the network is initialized, and are continuously updated during the process of developing the neural network.

- Summing function: Usually used to calculate the sum of inputs' product with their synaptic weight.

- Threshold (Also called bias): Quota is usually represented as a factor of the transfer function.

- Transfer function – also called activation function: This function is usually used to limit the boundary's output of each neuron. Its input is the result of the given summing function and the threshold. Usually, the boundary of each neuron is limited in the sector [0,1] or [-1, 1]. Transfer function may be a linear function or a non-linear function. Choosing which transfer function will depend on each math problem and the experience of the network's designer.

- Output: is the output signal of a neuron, with each neuron having a maximum output of 1.

Artificial neural also the same as biological neural networks, receiving the inputs, analyzing by multiplying these signals with synaptic weights, calculating the products, then sending the result to the transfer function, the outcome of the transfer function is the output signal.

$$Y = (X_1 w_1 + b_1) + (X_2 w_2 + b_2) + \ldots + (X_n w_n + b_n)$$

ISSN: 2455-8834

Volume:10, Issue:06 "June 2025"

Each neuron can process specific information, but the ability to calculate a neuron mostly determined by its structural connection with neurons that can process complex calculations and give an accurate result.

4.1.4 Confusion matrix for evaluating forecast performance

In this study, we have used the comparison method between financial risk prediction models via a Confusion Matrix. This is a method to evaluate the efficiency of differentiating observations into two sets, risky or not risky, because of its veracity and levels of coverage of the classification. A risky set will result in a value of 1, and a no-risk set will be valued at 0.

Confusion matrix will includes these indexes (Table 1):

- TP (true positive) is the index predict the positiveness, which means the number of predicted financial risking businesses that correctly forecast as involved in financial risk;
- TN (true negative) is the number of businesses not facing financial risk predicted as not getting involved in financial risk;
- FP (false positive) is the number of businesses not getting in financial risks but predicted as facing financial risk;
- FN (false negative) is the number of businesses facing financial risk but predicted as no financial risk.

Table 1: Confusion Matrix

	True value		
		1	0
Predicted result	1	TP	FP
	0	FN	TN

Source: Shrivastava et al (2020)

The accuracy of the model is the ratio of accurate predictions, calculated via this formula:

Accuracy = (TP + TN) / (TP + FP + TN + FN)

Therefore, to enhance the prediction efficiency of the model, besides the accuracy of the model, two criteria are used to evaluate the effectiveness of the prediction, Precision và Recall:

$$\text{Precision} = TP / (TP + FP)$$

ISSN: 2455-8834

Volume:10, Issue:06 "June 2025"

$$\text{Recall} = TP / (TP + FN)$$

Precision can give the ratio of predicted financial risk businesses that have faced financial risk, and Recall gives the ratio of accurately predicted financial risk businesses to the total of businesses facing financial risk.

4.2 Data and variable

Data: Businesses in the study include manufacturing enterprises listed on the Vietnamese stock market. Financial quotas used for financial risk prediction of these businesses are calculated from indicators in audited public financial statements (balance sheet, cash flow statement, income statement) at the end of year of 248 manufacturing enterprises listed on the Vietnamese stock market from 2020 to the end of 2024, have a sum of 1.420 observations taking from Fiingroup's data resources. Businesses that were included in the study need to guarantee having a financial statement published in FiinGroup.com's data resources during the investigation time. The study group has used 35 financial criteria, calculated from financial statements, consulted from the study of Tian and Yu (2017), Tian et al (2015), and presented in Table 2.

Dependent Variable: The study has used the z-score (Altman, 1968) to classify the business as risky or not risky. Dependent variable (y) will be valued as 1 if Z-score < 1,81 (the business engages in financial risk), variable y will be valued as 0 if Z-score ≥ 1,81 (the business secures its finance). Classification is also suitable for the given results. The result of classifying 1420 observations has 372 observations belonging to groups that have financial risk and 1048 observations that have no financial risk.

Independent variable: According to available sources and references from the study of Zeytinoglu et al (2013), Valaskova et al (2018)…, and from real businesses' activities, the study has built, selected 35 variables (Table 2).

Table 2: Variables that have an impact on financial risk

Variables	Descriptions	Variables	Descriptions
X1	Current Assets/Current Liabilities	X19	Profit before tax/Share capital
X2	(Current assets – inventories)/ Short-term debt	X20	Market capitalization
X3	Revenue/Inventory	X21	Inventory/Sales

X4	Receivables/(Revenue/365)	X22	Inventory Turnover Ratio
X5	Revenue/Fixed Assets	X23	Profit after tax/Net revenue
X6	Revenue/Total Assets	X24	Cost of goods sold/Accounts receivable
X7	Total debts/Total Assets	X25	Fixed assets/total assets
X8	Owner's Equity/Total Assets	X26	Short-term Debt /Total Assets
X9	Total debts/Share capital	X27	Current Liabilities/Accounts Payable
X10	ROS	X28	Log (Total Assets)
X11	ROA	X29	Working Capital/Total Assets
X12	ROE	X30	log (sales)
X13	Cash and cash equivalents/ Current liabilities	X31	Net income/total assets
X14	(Current assets – Current liabilities)/ Total assets	X32	Net income/ (market equity + total liabilities)
X15	Short-term receivables/ Current assets	X33	Net income/revenue
X16	Revenue/Equity	X34	Total liabilities/total assets
X17	Current Liabilities/Total Assets	X35	Short-term debt/revenue
X18	Long-term debt/Total assets		

Source: Summary and recommendations of the authors

5. Study result

International Journal of Social Science and Economic Research

ISSN: 2455-8834

Volume:10, Issue:06 "June 2025"

In forecasting financial risks for manufacturing enterprises listed on the stock exchange in Vietnam, we have divided the data set to 2 sets: training set, includes 4 years in the period 2020-2023, and testing set, is the data in 2024, where the training set is used for machine learning and the test set is used to test the machine's learning ability. In some cases, to avoid the data to be overfitting, we have used a selective method that depends on the LASSO model to select variables that affect a business's finance and can provoke high risk in the business's finance. Specifically, we have divided the process into three sessions:

Session 1: Apply the LASSO model

The group used LASSO regression to find important explanatory variables. After using LASSO regression, the group obtained the results of the selected variables as X1, X6, X11, X18, and X26. The variables characterised for Current Assets/Current Liabilities (X1), Revenue/Total Assets (X6), ROA (Net Profit/Total Assets) (X11), Long-term debt/Total assets (X18), Short-term Debt /Total Assets (X26). The results show that the variables: solvency, financial efficiency, profitability, debt ratio, and financial leverage of the enterprise help predict, as well as distinguish enterprises with the ability to face high or low financial risks.

Session 2: Select a machine learning model

Our team re-ran the forecasting of financially distressed firms with machine learning models using financial variables filtered out by the results of LASSO regression. The performance of these models is shown in Table 3.

Table 3: Forecasting results with variables from LASSO model of machine learning models

Model	accuracy_score	f1_score	precision_score	recall_score
Logistic Regression	94,34%	88,15%	88,18%	88,13%
Decision Tree	91,46%	82,82%	80,22%	84,58%
Random Forest	95,50%	89,37%	90,59%	88,12%
KNN	93,18%	86,13%	91,14%	81,54%
SVM	95,76%	90,94%	89,77%	92,13%
ANN	96,26%	92,03%	93,21%	90,74%

Source: Authors calculated from collected data

The result (Table 3) show that, the business classification forecasting on the test file shows that all models give correct forecasting ability over 92,03% with the Artificial Neural Network model (ANN) gives the highest prediction result's accuracy with the number approximately 96,26%, higher than smart machine learning models such as Random Forest or SVM. Surprisingly, the traditional Logistics regression model gives the accuracy rate of 94,34%, only ranked 4th after other smart models.

Recall_score high will also mean that the rate of businesses at risk of being overlooked is lower, and high precision_score reports that the accuracy of the prediction is also high. The best model is the one that, in addition to high accuracy, also has both high Recall_score and Precision_score (> 90%). In these two indicators, the ANN model continues to give the highest forecasting performance. The Decision Tree model has indicators showing a higher rate of businesses being overlooked and lower veracity in prediction (Recall_score and Precision_score both < 85%), resulting in the stability of the Decision Tree model not being large.

From the comparison result above, the study group has decided to choose the ANN model to continue evaluating and predicting, identifying financial risks in listed companies on the Vietnamese stock exchange.

Session 3: Training and experiment data through the ANN model

An ANN model is a popular form of neural network in machine learning, in which there is at least one hidden layer between the input layer and the output layer.

Retrain on data set.

After choosing ANN as our main model, our group has retrained the model based on data set of businesses listed on the Vietnamese stock market from 2020 to 2023 with variables X1, X6, X11, X18, X26 (Table 2), selecting based on experimental results of the LASSO regression model. Our group continues to randomly divide the training dataset from 2020 to 2023, an ANN model with 70% training data and 30% testing data, then run the neural network model. The accuracy of the model on the 'train' set is 95%, and the accuracy on the 'test' set is 96%.

Model accuracy is used to evaluate the performance of a neural network model. For classification ability, accuracy is often an important indicator to show the accuracy level after being trained on the test set. However, to see the reliability of the results achieved by the model, as well as detect the errors encountered by the model, thereby improving and refining the model to improve the prediction performance and minimize errors during the training and testing of the model, our team used the Confusion Matrix to help evaluate and understand the performance of a classification model, and support decisions in improving the model.

ISSN: 2455-8834

Volume:10, Issue:06 "June 2025"

Result collected from the Confusion Matrix between real results and prediction results from the models:

- Actual: 0 is the real result of the business has no risk.

- Actual: 1 is the real result of the business has risk.

- Predicted: 0 is the predicted result that the business has no risk.

- Predicted: 1 is the predicted result that the business has risk.

Table 4: Confusion Matrix Result on the dataset of 2020-2023

Model	TP	FP	FN	TN
ANN	201	7	6	70

Source: Authors calculated from collected data

From the results of the Confusion Matrix (Table 4), we can see that:

- Actual: 0, Predicted: 0 - Actual business has no risk, forecast model has no risk: 201 businesses.

- Actual 0, Predicted: 1 - Actual business has no risk, the model predicts that the business has risk: 7 businesses.

- Actual: 1, Predicted: 0 - Actual businesses have risks, forecast model has no risks: 6 businesses.

- Actual: 1, Predicted: 1 - Actual business risks, forecasting model of business risks: 70 businesses.

As mentioned in the previous section, a high recall score means that the rate of businesses at risk of being missed is lower, and a high precision score reflects high forecast accuracy. After retraining on the dataset, the recall score and precision score achieved were 92.1% and 90.9%, respectively.

Retry on real data of the year 2024

To determine if the model still works well and does not overfit, the authors tested the model on the 2024 dataset. The model's accuracy on the dataset of listed companies on the stock market from 2020 to 2023 is 96%, so how accurate will the model be on data that has not been learned and tested?

ISSN: 2455-8834

Volume:10, Issue:06 "June 2025"

To verify the performance of the entire model, the authors continued to run the model on a real dataset of listed companies on the stock market in 2024 with variables X1, X6, X11, X18, and X26. The results obtained an accuracy of up to 95%.

Table 5: Result of Confusion Matrix on the dataset of 2024

Model	TP	FP	FN	TN
ANN	204	5	8	67

Source: Authors calculated from collected data

The results when using the confusion matrix (Table 5) show:

- Actual: 0, Predicted: 0 - In reality, the business has no risk; the business forecast model has no risk: 204 businesses.

- Actual: 0, Predicted: 1 - In reality, the business has no risk; the business forecast model has risk: 5 businesses.

- Actual: 1, Predicted: 0 - In reality, the business has risk; the business forecast model has no risk: 8 businesses.

- Actual: 1, Predicted: 1 - In reality, the business has risk; the business forecast model has risk: 67 businesses.

Thus, the machine learning model that the team built using the Neural Network model combined with financial indicators from the LASSO regression model gave an accuracy of up to 95% for real data, and an error rate of 5%. Of which, 2.03% predicted that the business was at financial risk but the business was not at financial risk, which did not pose a danger to the business and only 2.97% predicted that the business was not at financial risk, but the business was at financial risk, which posed a danger to the businesses in this group. Therefore, the model that the team built can be considered a highly accurate model and useful for business risk management.

6. Conclusion and recommendations

6.1. Conclusion

Based on data set of 284 manufacturing enterprises listed on Vietnam stock market from 2020 to 2024, with 35 indicators affecting financial risks, the research results show that from these 35 indicators, through the LASSO model, 5 indicators on: Solvency, Operating efficiency, Profitability, Debt structure and Ratio of assets to long-term debt of the enterprise have the most

ISSN: 2455-8834

Volume:10, Issue:06 "June 2025"

impact on the financial risk of the enterprise in 5 years (2020-2024). This is consistent with reality when these indicators are used by enterprises to consider and evaluate the enterprise's finances.

In the study, the group of authors also compares prediction models and points out that the ANN is the best model to predict whether a business will encounter financial risk. This is consistent with the research results of Dao Trong Thinh and Doan Van Toan (2016) and also consistent with the research results of Qin (2022). The artificial neural network model has strong adaptability when facing different types of data and strong performance. The model built by the research team has 96% accuracy on the training data set and 95% accurate prediction of whether the business is at financial risk based on the real data set, the model also works well with machine learning model evaluation indexes, such as: recall 92.1%, precision is 90.9%.

The result of the study can support businesses in detecting financial risk and help investors and business owners in making their decisions. Besides, this result may help as a reference for later studies.

6.2. Recommendations

Financial risk management is an important element of risk management in enterprises to ensure successful operations, including activities such as identifying possible risks, assessing the impact on business operations, and preparing plans to deal with adverse events. However, financial risk management for enterprises requires a lot of resources and time. Enterprises need to establish a financial risk management department, follow the financial risk management process, use many risk identification techniques, find funding sources for financial risk management activities, etc. Applying data analysis in financial risk management can help enterprises assess in more detail and comprehensively the current state of the enterprise. This helps to maximize the efficiency of resources, save time while still producing accurate results, which is very helpful for the development and survival of enterprises. Businesses need to focus more on human resources and resources in developing data analysis, and should combine experience and data understanding to make the right decisions because data is the factor that most truthfully reflects all business issues.

References

1. Altman, E.I (1968), Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, *Journal of Finance*, 23, 589-609.
2. Altman, E.I., Iwanicz-Drozdowska, M., Laitinen, E.K. and Suvas, A., (2017). Financial distress prediction in an international context: A review and empirical analysis of

ISSN: 2455-8834

Volume:10, Issue:06 "June 2025"

Altman's Z-score model. *Journal of International Financial Management & Accounting*, 28(2), pp.131-171.

3. Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications*, 83, 405-417.

4. Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of accounting research*, 71-111.

5. Campbell, J. Y., Hilscher, J., & Szilagyi, J. (2008). In search of distress risk. *The Journal of Finance*, 63(6), 2899-2939.

6. Chandra, D. K., Ravi, V., & Bose, I. (2009). Failure predict of dotcomtcom companies using hybrid intelligent techniques. *Expert Systems with Applications*,36, 4830–4837.

7. Chen, M. A., Wu, Q., & Yang, B. (2019). How valuais fintechtech innovation?, *The Review of Financial Studies*,32(5),2062–2106.

8. Dao Trong Thinh, Doan Van Toan (2016), Applying artificial neural network in managing credit risks at commercial banks in Bien Hoa City, *Journal of Science of Lac Hong University*, Vol. 5, pp. 31-35.

9. Doan Khanh Hung, Tran Thi Hien (2019), Tác động của cấu trúc tài chính đến rủi ro tài chính của các doanh nghiệp dịch vụ du lịch niêm yết tại Việt Nam, *Tạp chí Khoa học Đại học Huế: Kinh tế và Phát triển*, 128(5A), 93-104.

10. Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of statistics*, 1189-1232.

11. Geomean-based boosting algorithm with over-sampling to resolve the data imbalance problem for bankruptcy prediction. *Expert Systems with Applications*, 42(3), 1074-1082.

12. Heo, J., & Yang, J. Y. (201AdaBoost-basedased bankruptcy forecasting of Korean construction companies. *Applied soft computing*, 24, 494-499.

13. Huang, B., Wei, J., Tang, Y., and Liu, C. (2021), Enterprise risk assessment based on machine learning, *Computational Intelligence and Neuroscience*.

14. Iswarya KonasaniInt. (2025), AI-Driven Risk Management: Transforming Financial Decision-Making, *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 11(): 3574-3582.

ISSN: 2455-8834

Volume:10, Issue:06 "June 2025"

15. Jones, S., Hensher, D. A. (2004). Predicting firm finance distress: a mixed logit model. *Accounting Review*, 79(4), 1011 – 1038.

16. Le Hai Trung, Truong Thi Thuy Duong (2023), Corporate bankruptcy risk prediction in Vietnam: An application with LASSO method, *Journal of Banking Scientific & Trainingning*, 250.

17. Le, H. H., Viviani, J. L. (2018). Predicting bank failure: An improvement in implementing a machine-learning approach to classic financial ratios. *Research in International Business and Finance*, 44, 16 – 25.

18. Lin, T. H. (2009). A cross model study of corporate financial distress prediction in Taiwan: Multiple discriminant analysis, logit, prob, it, and neural networks models. *Neurocomputing*, 72(16-18), 3507-3516.

19. Murugan, M. S. (2023), Large-scale data-driven financial risk management & analysis using machine learning strategies, *Measurement: Sensors*, 27.

20. Nguyen Thi Thanh Loan et al (2024), Nghiên cứu ứng dụng phân tích dữ liệu trong quản trị rủi ro tài chính tại các công ty niêm yết trên sàn chứng khoán Việt Nam, *Economy & Forecast Review*, https://kinhtevadubao.vn/nghien-cuu-ung-dung-phan-tich-du-lieu-trong-quan-tri-rui-ro-tai-chinh-tai-cac-cong-ty-niem-yet-tren-san-chung-khoan-viet-nam-28875.html.

21. Olson, D. L., Delen, D., and Meng, Y. (2012). Comparative analysis of data mining methods for bankruptcy prediction. *Decision Support Systems*, 52(2):464–473.

22. Orcun Sarioguz, Evin Miser (2024), Integrating AI in financial risk management: Evaluating the effects of machine learning algorithms on predictive accuracy and regulatory compliance, *International Journal of Science and Research Archive*, 13(02), 789-811.

23. Paraschiv, F., Schmid, M., & Wahlstrøm, R. R. (2021). Bankruptcy prediction of privately held SMEs using feature selection methods. Available at SSRN 3911490

24. Qin, W. (2022), Research on financial risk forecast model of listed companies based on convolutional neural network, *Scientific Programming*, 1-10, DOI 10.1155/2022/3652931.

25. Serrano-Cinca, C., & GutiéRrez-Nieto, B. (2013). Partial least squares discriminant analysis for bankruptcy prediction. *Decision Support Systems*, 54(3), 1245–1255.

ISSN: 2455-8834

Volume:10, Issue:06 "June 2025"

26. Shrivastava, S., Jeyanthi, P. M., and Singh, S. (2020), Failure prediction of Indian Banks using SMOTE, LASSO regression, bagging and boosting, *Cogent Economics & Finance*, 8(1).

27. Tian, S., Yu, Y., & Guo, H. (2015). Variable selection and corporate bankruptcy forecasts. *Journal of Banking & Finance*, 52, 89-100.

28. Tian, S., & Yu, Y. (2017). Financial ratios and bankruptcy predictions: An international evidence. *International Review of Economics & Finance*, 51, 510-526.

29. Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *Journal of Royal Statistical Society*, B, 58, 267-288.

30. Valaskova, K., Kliestik, T., and Kovacova, M. (2018), Management of financial risks in Slovak enterprises using regression analysis, *Oeconomia copernicana*, 9(1), 105-121.

31. Valaskova, K., Kliestik, T., Svabova, L., and Adamko, P. (2018), Financial risk measurement and prediction modelling for sustainable development of business entities using regression analysis, *Sustainability*, 10(7).

32. Vuong, Q. H. (2023). Mindsponge Theory. *Walter de Gruyter GmbH*.

33. Zeytinoglu, E., and Akarim, Y. D. (2013), Financial failure prediction using financial ratios: An empirical application on Istanbul Stock Exchange, *Journal of Applied Finance and Banking*, 3(3).

