Early Warning of Liquidity Risk in Commercial Banks In Vietnam: Application of Artificial Intelligence Models

Hồ Nguyễn Phúc Đức, Trương Thị Vân Khánh, Đỗ Huyền Linh

Banking academy, Vietnam
**Correspongding author. Email: hoduc1709@gmail.com*

ABSTRACT

In the context of globalization and the rapid evolution of financial markets, liquidity risk remains one of the major challenges for commercial banks. To address this issue, Artificial Intelligence (AI) and Machine Learning (ML) have been increasingly applied in the banking sector thanks to the capability of analyzing large-scale data, processing non-linear variables, and effectively identifying hidden patterns. Our research contributes to the domain of banking and finance by developing AI models in an early warning system for liquidity risk prediction in commercial banks in Vietnam. By identifying key factors of liquidity crises, these models enable banks and financial regulatory authorities to implement timely preventive measures and risk management strategies. Ultimately, this initiative aims to enhance the safety and stability of the banking and finance system.

Keywords: *banking crisis, early forecasting models, liquidity risk, machine learing.*

1. INTRODUCTION

Liquidity risk is an important factor affecting the stability of commercial banks, especially in the context of financial integration. Liquidity risk is not only the lack of immediate payment ability (Drehmann, 2010) but also the balance of assets and short-term obligations (Yao & Song, 2021). Artificial Intelligence (AI) and Machine Learning (ML) are being widely applied in liquidity risk prediction thanks to the capability of processing big data and detecting hidden patterns. (Mashrur et al., 2020) emphasizes the role of ML in financial forecasting, while (Yao & Song, 2021) demonstrate the effectiveness of Recurrent Neural Networks (RNN) in predicting liquidity fluctuations. In Vietnam, (Nguyễn, 2023) proposed ML applications including Random Forest and XGBoost to optimize liquidity risk management. Although the application of AI in the early warning system of liquidity risk requires careful preparation of data and technology (Mashrur et al., 2020), it is considered as a great innovation because of optimizing processes, increasing calculation accuracy and monitoring the impact of market fluctuations. Therefore, commercial banks in Vietnam need to invest in data infrastructure to maximize the efficiency of AI in liquidity management.

2. RESEARCH METHOD

2.1. Research model

2.1.1 LASSO model

LASSO (Least Absolute Shrinkage and Selection Operator), introduced by Robert Tibshirani in 1996, is a key regression method in statistics and ML, particularly useful for feature selection and financial forecasting. By applying L1 regularization, LASSO eliminates insignificant variables by shrinking some regression coefficients to zero, reducing data dimensionality and addressing multicollinearity (Liu & Yu, 2022). LASSO has great potential in early warning systems for liquidity risk due to its ability to filter out data noise and improve forecasting accuracy (Liu & Yu, 2022). When combined with strong nonlinear models like Random Forest or XGBoost, it helps refine data features, enhancing overall model performance (Murugan & T, 2023). According to (Tian et al., 2020), LASSO removes variables with no explanatory significance, mitigating overfitting and strengthening predictions. In banking, where data is complex and nonlinear, LASSO helps identify key factors influencing liquidity risk (Shrivastava et al., 2020)

2.1.2. Random Forest model

Random Forest (RF), developed by Breiman (2001), enhances decision trees by using Bootstrap Aggregation (Bagging) to reduce overfitting and improve accuracy. It builds multiple decision trees on randomly sampled data subsets and aggregates their predictions (Breiman, 2001; Beutel, 2019).

RF excels in handling nonlinear, noisy financial data and identifying complex relationships that linear models miss (Liu & Yu, 2022). It supports strategic decision-making by assessing variable importance, aiding liquidity risk management (Noura et al., 2023). Though computationally intensive and less interpretable, RF remains a powerful tool for financial risk forecasting, improving early warning models for Vietnamese banks (Drudi & Nobili, 2021).

2.1.3. XGBoost model

XGBoost, introduced by Chen & Guestrin, enhances Gradient Boosting for faster training and higher accuracy. It is widely used in finance for liquidity risk forecasting, credit analysis, and portfolio management, outperforming Random Forest and LASSO in handling nonlinear data (Liu & Yu, 2022). XGBoost builds trees sequentially, correcting errors to improve predictions but risks overfitting if not well-tuned (Drudi & Nobili, 2021). It achieves 10–20% higher accuracy in risk detection and enables early warning of liquidity risks 3–6 months ahead of traditional methods (Beutel et al., 2019); (Drudi & Nobili, 2021). Combining it with other models further improves forecasting (Murugan & T, 2023). Though computationally demanding, XGBoost's predictive power makes it a strong tool for liquidity risk management in Vietnamese banks, enhancing early warnings and financial stability.

2.2. Data and variables

This study uses a data set of 239 observations, collected from financial statements and public data sources of 24 commercial banks operating in Vietnam in the period from December 31, 2014 to December 31, 2023 with an annual frequency, reflecting the fluctuations of the banking system in a challenging economic context. To ensure security and consistency in analysis, banks are encrypted by stock code (CK Code). In addition, macroeconomic data is collected from the World Bank, ensuring transparency and high reliability in analysis.

The dependent variable (Default) is used to measure the liquidity risk of commercial banks in Vietnam, determined based on research by Yan & Song (2022). This variable is built based on three financial indicators including Loan-to-Deposit Ratio (LDR), Short-term Liquidity Coverage Ratio (LCR) and Net Stable Funding Ratio (NSFR). Calculated from financial reports of commercial banks and complies with Basel III standards. From there, the study proposes two hypotheses as follows:

H1: Default = 0: The bank does not have liquidity risk.

H2: Default = 1: The bank has liquidity risk.

Independent variables: includes 11 variables used to early forecast liquidity risks of commercial banks including: CAR, LCR, GGDP, INF, ROA, UR, NPL, LDR, SIZE, NSFR, NIM. The variables are described specifically in Table 1. These variables are popular micro and macro financial indicators that have been widely used in previous research.

3. RESULTS AND DISCUSSIONS

3.1. Descriptive statistics

Table 1 Descriptive statistics of variables in the model

Variables	Average	Standard deviation	Min	Max	Observations
CAR	0.106430554	0.003223385	0	0.2453	238
LDR	0.75223142	0.007039792	0.318118951	1.057114104	238
LCR	2.10084E+11	93156214587	0	1E+13	238
NSFR	126.5592256	1.274033565	95.36381715	236.2278558	238
ROA	0.010033217	0.000475653	0	0.03652643	238
NIM	0.031788988	0.000866283	0	0.093705802	238
SIZE	32.86806238	0.072467563	30.3925069	35.37206315	238
NPL	0.01773701	0.000661553	0	0.069120824	238

Variables	Average	Standard deviation	Min	Max	Observations
GGDP	6.047908418	0.119495493	2.553728526	8.123514468	238
INF	3.436622882	0.072672158	1.834715548	6.31200905	238
UR	1.73192437	0.022710127	1.161	2.385	238
Default	0.68487395	0.030176808	0	1	238

Source: Compiled by the authors

Data from the table shows that the average capital adequacy ratio (CAR) reached 10.65%, while the loan-to-deposit ratio (LDR) was 75.20%. The bad debt ratio (NPL) is low, averaging 1.78%, and real GDP growth (GGDP) reached 6.04%. The average inflation rate (INF) is 3.44%, the unemployment rate (UR) remains low at 1.73%. In terms of operating efficiency, the average return on assets (ROA) reaches 1%. The default variable (Default) shows that 68.6% of cases belong to the default group. The data shows relative stability, but there are some outliers to watch out for.

3.2. Discuss research results

Table 2. Results of machine learning models

Model	R2
Lasso	0.8723
XGBoost	0.9787
Random Forest	0.998

Source: Compiled by the authors

The article emphasizes that Random Forest (RF) is the most accurate model for predicting liquidity risks in commercial banks, achieving an accuracy of up to 99.8%. This model significantly outperforms traditional methods, providing better predictions of a bank's solvency and contributing to financial stability. RF's key advantage is its ability to handle complex relationships between financial variables while avoiding overfitting. It efficiently processes large and complex datasets, enabling banks to gain a clearer view of liquidity risks and implement more effective risk management strategies. This helps minimize potential financial instability and ensures long-term resilience. Additionally, the article proposes further research on enhancing liquidity risk prediction models by integrating machine learning with macroeconomic data and global financial trends. This approach would allow banks to better understand external factors influencing liquidity risks and improve long-term forecasting accuracy, making their risk management strategies more reliable and adaptable.

4. CONCLUSION

Our study shows the influence of liquidity risk on not only commercial banks but also the entire banking and finance system in Vietnam. Moreover, the accuracy of Artificial Intelligence (AI) models in early warning of liquidity risk helps these commercial banks to minimize potential financial risks. Specifically, the Random Forest model is the most highly rated in forecasting liquidity risk because of good generalization ability and high accuracy in data classification. The application of AI in liquidity risk management contributes to improving the operational efficiency of the banking system, ensuring financial stability in the context of a volatile economy.

REFERENCES

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

Beutel, J., List, S., & Von Schweinitz, G. (2019). Does machine learning help us predict banking crises? Journal of Financial Stability, 45, 100693. https://doi.org/10.1016/j.jfs.2019.100693

Drehmann, M., & Nikolaou, K. (2010). Funding liquidity risk: Definition and measurement. BIS Working Papers, No. 316. Retrieved from
https://www.bis.org/publ/work316.pdf

Liu, T., & Yu, Z. (2022). The analysis of financial market risk based on machine learning and particle swarm optimization algorithm. EURASIP Journal on Wireless Communications and Networking, 2022(1), 31. https://doi.org/10.1186/s13638-022-02117-3

Mashrur, A., Luo, W., Zaidi, N. A., & Robles-Kelly, A. (2020). Machine Learning for Financial Risk Management: A Survey. IEEE Access, 8, 203203–203223. https://doi.org/10.1109/ACCESS.2020.3036322

Metawa, N., Hassan, M. K., & Metawa, S. (2022). Artificial Intelligence and Big Data for Financial Risk Management: Intelligent Applications (1st ed.). Routledge. https://doi.org/10.4324/9781003144410

Murugan, M. S., & T, S. K. (2023). Large-scale data-driven financial risk management & analysis using machine learning strategies. Measurement: Sensors, 27, 100756. https://doi.org/10.1016/j.measen.2023.100756

Nguyễn, Đ. K., Phạm, M. P., & Lê, Q. T. (2023). Nghiên cứu tác động của rủi ro thanh khoản đến hiệu quả hoạt động kinh doanh của các ngân hàng thương mại tại Việt Nam. Tạp chí Ngân hàng, (15), 24–28. https://www.researchgate.net/publication/374754346_nghien_cuu_tac_dong_cua_rui_ro_thanh_khoan_den_hieu_qua_hoat_dong_cua_cac_ngan_hang_thuong_mai_viet_nam

Nobili, S., & Drudi, M. L. (2021). A Liquidity Risk Early Warning Indicator for Italian Banks: A Machine Learning Approach. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3891566

Tian, Z., Xiao, J., Feng, H., & Wei, Y. (2020). Credit Risk Assessment based on Gradient Boosting Decision Tree. Procedia Computer Science, 174, 150–160. https://doi.org/10.1016/j.procs.2020.06.070

Yao, T., & Song, L. (2021). Examining the differences in the impact of Fintech on the economic capital of commercial banks' market risk: Evidence from a panel system GMM analysis. Applied Economics, 53(23), 2647–2660. https://doi.org/10.1080/00036846.2020.1864275