e-ISSN: 3047-857X
Proceedings of the International Symposium on Management (Volume 22, 2025)

Application of AI in Big Data Analytics to Financial Risk Forecasting

Hanh Dao Thi

Faculty of Accounting – Finance, Dong Nai Technology University, Bien Hoa City, Vietnam
Corresponding author. Email: daothihanh@dntu.edu.vn

ABSTRACT

Artificial intelligence (AI) is increasingly asserting its important role in many fields, especially healthcare and finance, thanks to its ability to process and analyze large amounts of diverse data, detecting complex relationships that are difficult for traditional methods to recognize. In the medical field, AI has proven to be outstandingly effective in diagnosis and treatment support. Similarly, in finance, AI is being strongly applied to risk forecasting, especially credit risk and market risk. Machine learning and deep learning models help increase the accuracy of predictions, processing both traditional and unstructured data. However, the implementation of AI also comes with challenges such as ethics, privacy, and transparency in the model. This study aims to overview the applications of AI in financial risk forecasting, and analyze the potential and challenges to overcome to optimize risk management efficiency in the context of big data.

Keywords: *Artificial Intelligence (AI); Financial risk forecasting; Credit risk; Market risk*

1. INTRODUCTION TO FINANCIAL RISK FORECASTING

Artificial intelligence (AI) is capable of processing and analyzing vast amounts of data, allowing it to detect complex relationships that are difficult for traditional methods to recognize. In the medical field, AI has demonstrated the ability to capture sophisticated connections from multiple multi-parameter datasets, allowing for the discovery of hidden relationships in data and the handling of more complex tasks than traditional methods (Cau et al., 2024). For example, in cardiovascular disease diagnosis, AI can integrate and analyze various types of data such as personal and family histories, physical examinations, electrocardiograms, and subclinical tests. It's worth noting that AI isn't just limited to analyzing financial data. In the medical field, AI has been widely applied to detect and diagnose various diseases, from breast cancer (Bahl, 2020) to cardiovascular disease (Almansouri et al., 2024; Cau et al., 2024) and many other fields. This shows the immense potential of AI in processing and analyzing complex data in a variety of fields. AI's ability to process large amounts of data and detect complex relationships offers many benefits in many fields, not just limited to finance. In healthcare, AI has proven its ability to improve diagnosis, support decision-making, and optimize treatment strategies. However, the adoption of AI also poses ethical, privacy, and regulatory compliance challenges, requiring close oversight and continuous improvement of models (Cheungpasitporn et al., 2024).

Financial risk forecasting is an important area of corporate financial management, which uses various methods and models to assess and predict potential risks. Traditional methods often focus on analyzing financial statements, but may be limited due to delays and inaccuracies in information (Bi et al., 2022). To overcome this, recent studies have applied more advanced machine learning and artificial intelligence techniques. Some of the new methods include the use of neural networks and deep learning to build optimal risk forecasting models (Gu, 2023), as well as the application of generative AI in financial risk analysis (Joshi, 2025). These methods allow for more complex data processing and provide more accurate forecasting. However, human monitoring is still important to mitigate potential errors from fully automated models (Joshi, 2025). Financial risk forecasting is evolving rapidly with a combination of traditional and modern techniques. While the new methods offer many benefits, their adoption needs to be carefully considered to ensure reliability and regulatory compliance. A combination of advanced forecasting models and expert assessments can deliver the best results in corporate financial risk management

2. CREDIT RISK AND MARKET RISK

Financial institutions use a variety of methods to predict credit risk, going far beyond relying solely on traditional factors such as credit scores. Research shows that new online lending markets can predict a borrower's likelihood of defaulting 45% more accurately than their accurate credit score (Iyer et al., 2009). Screening through soft or non-standard information is especially important when evaluating lower-quality borrowers. Some other interesting findings include: providing borrowers with personalized information about their FICO scores can significantly reduce late payments and improve credit scores (Homonoff et al., 2021). In addition, a higher proportion of women among loan officers may increase loan portfolio risk, but this relationship is reduced by the gender proximity between female loan officers and female borrowers (Blanco-Oliver et al., 2021). Predicting credit risk requires a holistic approach, combining both traditional financial data and non-traditional factors such as soft information, ESG awareness, and behavioral dynamics. Financial institutions need to continuously improve their risk assessment models to capture these emerging factors and enhance their ability to predict accurately.

Market risk is closely related to macroeconomic factors such as exchange rates, interest rates, and stock prices. Research shows that there is a causal relationship between exchange rates and stock prices in some countries such as India and Sri Lanka (Smyth & Nandha, 2003). In Vietnam, exchange rates, the DJIA index, and world oil prices have a positive impact on stock returns (Rachmawati & Fadila, 2024). US macroeconomic factors also significantly affect Vietnamese stock prices (Hussainey & Khanh Ngoc, 2009). However, there are some conflicting findings. In Cyprus, although there is evidence of the ability to predict stock returns, it cannot be taken as evidence of market inefficiency (Tsoukalas, 2003). Research in the U.S. shows that dividend changes contribute very little to explaining historical stock price movements (Shiller, 1987). When the economy is strong, the stock market reacts negatively to news of higher economic activity (Mcqueen & Roley, 1993). In conclusion, forecasting market risks requires a comprehensive analysis of macroeconomic factors and financial indicators. In Pakistan, market returns have a strong influence on the sensitivity of bank stock returns to interest rates and exchange rates (Mohsin et al., 2020). In Greece, macroeconomic activity and changes in foreign stock markets only partially explain stock market volatility (Hondroyiannis & Papapetrou, 2001). Therefore, technical managers need to consider macroeconomic factors such as the USD/rupiah exchange rate, interest rates, and inflation when planning budgets (Luwihono et al., 2021).

AI is capable of processing and analyzing large amounts of data from various sources in the financial sector. Global banks process billions of international payments every day, while stock exchanges process trillions of orders and billions of transactions (Veloso et al., 2021). AI can analyze unstructured data from sources such as social media, news, and other external sources of information (Punia et al., 2021; Veloso et al., 2021). Technologies such as big data analytics, process automation, unstructured data analysis, and predictive modeling are used to improve the tax payment monitoring process (Zaqeeba et al., 2024). However, transforming unstructured data into structured data or meaningful information is a very difficult task (Punia et al., 2021). Various machine learning techniques such as decision trees, neural networks, assisted vector machines, and other algorithms are used to transform unstructured data into structured data (Punia et al., 2021). In conclusion, AI has immense potential for processing and analyzing financial data from a variety of sources. However, handling unstructured data is still a challenge. Advanced AI technologies such as deep learning and reinforcement learning are being developed to address these challenges and optimize the use of AI in the financial sector (Veloso et al., 2021).

AI models are capable of providing accurate forecasts of customer behavior and market volatility, helping financial institutions make timely and effective decisions. By analyzing large amounts of data in real-time, AI can predict trends and potential risks, allowing companies to proactively adjust their strategies (Abousaber & Abdalla, 2023; Patil et al., 2024). For example, AI can forecast a customer's solvency based on transaction history and other factors, helping banks make appropriate lending decisions. However, the use of AI in forecasting and decision-making also poses some challenges. It is necessary to ensure that the input data is diverse and unbiased, and that the results are regularly monitored and interpreted to avoid algorithmic biases (Sáez-Ortuño et al., 2023). In addition, ethical issues related to the use of personal data should also be carefully considered. AI is bringing significant improvements in the forecasting and decision-making capabilities of financial institutions. By combining the power of predictive analytics and machine learning, AI helps companies seize business opportunities, manage risk more effectively, and optimize operational performance (Pendyala & Lakkamraju, 2024; Vashishth et al., 2024). However, the implementation of AI needs to be done responsibly, ensuring transparency and fairness in the decision-making process.

Table 1. Comparison of Traditional Models and AI Models in Financial Risk Forecasting

Criteria	Traditional Model	AI (Machine Learning) Models
Model characteristics	Based on statistical models such as Regressions, ANOVA, Logistic Regression.	Based on machine learning algorithms such as Random Forest, XGBoost, Neural Networks, Deep Learning.
Request Data	Often requires less data, with well-defined elements.	Able to process large volumes and unstructured data (such as text, photos, signals from social networks).

Ability to process complex data	Often limited in handling complex relationships between factors.	It is possible to detect and learn complex relationships in data.
Accuracy	The accuracy may not be high in complex situations or rapidly changing data.	Greater accuracy in forecasting, especially in situations with complex or rapidly changing data.
Ability to learn and improve	Traditional models don't automatically improve as new data becomes available.	AI models can improve over time through learning from new data (reinforcement learning, supervised learning).
Flexibility	Less flexibility when applied to different situations.	It is very flexible and can be applied to a variety of data types and situations.
Big Data Analytics	Often not suitable for big data or complex data.	Very powerful in analyzing big and complex data, which can be analyzed from millions of financial transactions.

3. AI IS POPULAR IN FINANCIAL RISK FORECASTING

Artificial neural networks (ANNs) are a machine learning method that simulates how the human brain works, capable of modeling complex nonlinear relationships (Zou et al., 2008). ANNs can learn from data and apply in many areas such as pattern recognition, classification, prediction, and process control (Hwang & Ding, 1997). In the financial sector, the ANN has been effectively used to forecast exchange rates (Yu et al., 2007) and stock market fluctuations (Donaldson & Kamstra, 1996). It is worth noting that ANNs often outperform traditional methods in combining forecasts. Research on stock market volatility forecasting shows that ANNs often produce forecasts that are superior to traditional linear combination methods (Donaldson & Kamstra, 1996). However, the construction of the forecast interval for the ANN can be difficult due to the parameters that cannot be defined, although this problem can be solved in the forecast (Hwang & Ding, 1997). In conclusion, an ANN is a powerful tool for analyzing big data and modeling complex relationships in finance. The ability to learn from data and process non-linear relationships makes ANNs an attractive option for forecasting credit risk and market risk. However, it should be noted that the performance of an ANN depends on many factors such as network architecture, training data, and algorithms (Aggarwal & Song, 1998).

Random Forest is a powerful machine learning algorithm based on a set of multiple decision trees, capable of processing complex data and performing accurate classification (Che et al., 2023; Qi, 2012). It is widely applied in many fields such as text classification, breast cancer prediction, toxic domain name detection, and forest community classification (Aslam et al., 2022; Kabiraj et al., 2020; Li et al., 2013; Sun et al., 2020). Several improvements have been proposed to enhance the performance of Random Forest. For example, the use of a weighted voting mechanism to improve the quality of decision trees (Sun et al., 2020), or in combination with the Lasso method to dynamically select the optimal number of decision trees (Wang & Wang, 2020). Random Forest also has the ability to determine the relative importance of characteristics and measure the similarity between compounds, which is useful in the field of informatics chemistry (Svetnik et al., 2003). Random Forest is a powerful and versatile tool that can efficiently handle complex datasets with small sample sizes and high-dimensional feature spaces. It provides high prediction accuracy, the ability to interpret results, and can detect important factors affecting risk in a variety of applications (Qi, 2012; Xu et al., 2017).

XGBoost (eXtreme Gradient Boosting) is a powerful machine learning algorithm that is widely used in many fields, especially in financial risk forecasting. Numerous studies have demonstrated XGBoost's superior effectiveness compared to other algorithms in detecting financial reporting fraud (Ali et al., 2023), assessing credit risk for personal auto loans (Rao et al., 2022), and predicting home prices (Li, 2023; Sibindi et al., 2022). One of the outstanding advantages of XGBoost is its ability to process incomplete data without preprocessing. Research on insurance risk prediction shows that XGBoost can handle missing values and give accurate results comparable to models that have been processed with prior data (Rusdah & Murfi, 2020). In addition, XGBoost has the ability to learn from complex data patterns, as demonstrated in applications for predicting surface subsidence due to mining (Gu et al., 2024) and predicting neurological outcomes in patients with cervical spinal cord injury (Inoue et al., 2020). XGBoost is a powerful and flexible tool in data analysis and forecasting, especially effective in the financial and medical sectors. The ability to process incomplete data and learn from complex patterns makes XGBoost the first choice for many real-world risk forecasting and decision-making problems.

4. CONCLUSION

e-ISSN: 3047-857X

Artificial intelligence (AI) is reshaping the way financial risk is forecasted, bringing breakthroughs in both accuracy, processing speed, and the ability to adapt to complex, ever-changing data. In the context of an increasingly unstable global economy and increasingly diverse financial data, the application of AI along with big data has opened up a more effective approach, far beyond traditional methods that rely heavily on periodic financial index analysis and linear assumptions.

AI models such as artificial neural networks (ANNs), Random Forest, and XGBoost have proven their superiority in detecting latent behavioral patterns, modeling offline relationships, and integrating multi-source data to provide risk forecasts flexibly and accurately. These technologies not only support the prediction of credit risk and market risk, but also contribute to improving the efficiency of strategic decision-making in corporate financial management and financial institutions.

However, the powerful potential of AI also comes with many challenges that need to be identified and addressed. Issues related to algorithm transparency, data input reliability, bias in models, and liability in automated decisions are factors that require strict oversight and appropriate legal frameworks. At the same time, the role of humans remains central in controlling, calibrating and interpreting the results from the AI system, in order to ensure accuracy, fairness and ethics in the application process.

Overall, AI not only provides a powerful tool to improve financial risk forecasting capabilities, but also opens up opportunities to restructure risk management thinking in the digital age. The harmonious combination of advanced technology and extensive human expertise is a solid foundation for the development of effective, flexible and sustainable risk forecasting systems in the long term.

REFERENCES

Aggarwal, R., & Song, Y. (1998). Artificial neural networks in electrical systems. Part 3: Examples of applications in electrical systems. *Journal of Electrical Engineering*, *12*(6), 279–287. https://doi.org/10.1049/pe:19980609

Bahl, M. (2020). Artificial Intelligence: Bait for mammography radiologists. *Journal of Breast Imaging, 2*(4), 304–314. https://doi.org/10.1093/jbi/wbaa033

Bi, W., Wang, Z., Xu, B., Shen, H., Sun, X., & Cheng, X. (2022). *Company as Tribe: Assess the company's financial risk on a Tribe-style Graph with a hierarchical graph neural network*. *4*, 2712–2720. https://doi.org/10.1145/3534678.3539129

Blanco-Oliver, A., Reguera-Alvarado, N., & Veronesi, G. (2021). Credit risk in the microfinance industry: The role of gender relations. *Journal of Small Business Management*, *59*(2), 280–311. https://doi.org/10.1080/00472778.2020.1844487

Cau, R., Pisu, F., Suri, J. S., Gong, X., Montisci, R., Mannelli, L., Saba, L., & Gatti, M. (2024). Artificial intelligence in the diagnosis of differentiating cardiomyopathy phenotypes. *Diagnosis*, *14*(2), 156. https://doi.org/10.3390/diagnostics14020156

Cheungpasitporn, W., Kashani, KB, & Thongprayoon, C. (2024). The role of artificial intelligence and machine learning in acute kidney injury associated with sepsis. *Renal Research and Clinical Practice*, *43*(4), 417–432. https://doi.org/10.23876/j.krcp.23.298

Donaldson, R. G., & Kamstra, M. (1996). Forecasting in conjunction with neural networks. *Journal of Forecasting*, *15*(1), 49–61. https://doi.org/10.1002/(sici)1099-131x(199601)15:1<49::aid-for604>3.0.co;2-2

Gu, L. (2023). Reverse propagation neural networks are optimized to predict risks in corporate financial management. *Scientific Report*, *13*(1). https://doi.org/10.1038/s41598-023-46528-8

Homonoff, T., Sussman, A. B., & O'Brien, R. (2021). Does knowing your FICO score change financial behavior? Evidence from a field experiment with student borrowers. *Journal of Economics and Statistics*, *103*(2), 236–250. https://doi.org/10.1162/rest_a_00888

Hondroyiannis, G., & Papapetrou, E. (2001). Macroeconomic effects on the stock market. *Journal of Economics and Finance*, *25*(1), 33–49. https://doi.org/10.1007/bf02759685

Hussainey, K., & Khanh Ngoc, L. (2009). The impact of macroeconomic indicators on Vietnamese stock prices. *Journal of Risk Finance*, *10*(4), 321–332. https://doi.org/10.1108/15265940910980632

Hwang, JTG, & Ding, AA (1997). Prediction intervals for artificial neural networks. *Journal of the American Statistical Association*, *92*(438), 748–757. https://doi.org/10.1080/01621459.1997.10474027

Iyer, R., Luttmer, E.F.P., Shue, K., & Khwaja, A.I. (2009). Gently screen colleagues: Deduce the quality of microborrowers. *Management Science*, *62*(6), 1554–1577. https://doi.org/10.1287/mnsc.2015.2181

Joshi, S. (2025). Evaluation of Gene AI Models for Financial Risk Management. *International Journal of Computer Science, Engineering and Information Technology Research*, *11*(1), 709–723. https://doi.org/10.32628/cseit2511114

Luwihono, A., Asih, P., Mahfud, Z., Taryana, T., Saputro, R., Suprihartini, Y., Kalbuana, N., Suherman, B., Rusdiyanto, R., Prasetyo, B., Sembiring, D., & Rasyid, S. (2021). Macroeconomic Influence on Stock Prices: Evidence from Indonesia. *Accounting*, *7*(5), 1189–1202. https://doi.org/10.5267/j.ac.2021.2.019

Mcqueen, G., & Roley, V. V. (1993). Stock price, news, and trading conditions. *Financial Research Review*, *6*(3), 683–707. https://doi.org/10.1093/rfs/5.3.683

Mohsin, M., Zia-your-Rehman, M., Naseem, S., Naiwen, L., & Baig, S.A. (2020). Bank stock price volatility and macroeconomic fundamentals in the Pakistani context: an application of the GARCH and EGARCH models. *Oeconomia Copernicana*, *11*(4), 609–636. https://doi.org/10.24136/oc.2020.025

Punia, S. K., Patan, R., Stephan, T., Kumar, M., & Deverajan, G. G. (2021). Analyze the performance of machine learning algorithms to classify big data. *International Journal of Electronic Health and Health Communication*, *12*(4), 60–75. https://doi.org/10.4018/ijehmc.20210701.oa4

Rachmawati, R., & Fadila, A. (2024). Dynamics of Idxenergy stock returns: A comprehensive analysis of GDP, exchange rates, DJIA index, and oil prices. *Formosa Journal of Sustainability Studies*, *3*(2), 213–230. https://doi.org/10.55927/fjsr.v3i2.8058

Sáez-Ortuño, L., Forgas-Coll, S., Puertas-Prat, E., Sanchez-Garcia, J., & Huertas-García, R. (2023). *The impact of artificial intelligence on marketing research* (pp. 18–42). Global IGI. https://doi.org/10.4018/978-1-6684-9591-9.ch002

Shiller, R.J. (1987). Fluctuations in stock market prices. *Science*, *235* (4784), 33–37. https://doi.org/10.1126/science.235.4784.33

Smyth, R., & Nandha, M. (2003). Binary causal relationship between exchange rates and stock prices in South Asia. *Applied Economics Letters*, *10*(11), 699–704. https://doi.org/10.1080/1350485032000133282

Sun, Y., Zeng, Q., Bian, Y., & Li, Y. (2020). *Applied research on text classification based on random forest algorithms*. 370–374. https://doi.org/10.1109/aemcse50948.2020.00086

Svetnik, V., Sheridan, R. P., Liaw, A., Tong, C., Culberson, J. C., & Feuston, B. P. (2003). Random forests: a classification and regression tool for compound classification and QSAR models. *Journal of Chemical Information and Computer Science*, *43*(6), 1947–1958. https://doi.org/10.1021/ci034160g

Tsoukalas, D. (2003). Macroeconomic factors and share prices on the emerging Cyprus stock market. *Management Finance*, *29*(4), 87–92. https://doi.org/10.1108/03074350310768300

Veloso, M., Shah, S., Reddy, P., Balch, T., & Borrajo, D. (2021). Research on artificial intelligence in finance: discussion and examples. *Oxford Journal of Economic Policy*, *37* (3), 564–584. https://doi.org/10.1093/oxrep/grab019

Wang, H., & Wang, G. (2020). Improvement of the Random Forest Algorithm Using the Lasso Method. *Journal of Statistical Calculations and* Simulations, *91*(2), 353–367. https://doi.org/10.1080/00949655.2020.1814776

Yu, L., Lai, K.K., & Wang, S. (2007). *Forecast exchange rates using artificial neural networks*. Springer us. https://doi.org/10.1007/978-0-387-71720-3

Zaqeeba, N., Lutfi, A., Alrawad, M., Almaiah, M.A., Alshirah, A. F., & Alqudah, H. (2024). The impact of using various types of artificial intelligence technologies in monitoring tax payments. *International Journal of Data and Network Science*, *8*(3), 1577–1586. https://doi.org/10.5267/j.ijdns.2024.3.009

Zou, J., So, S.-S., & Han, Y. (2008). *Overview of Artificial Neural Networks* (Volume 458, pages 14–22). Humanitarian. https://doi.org/10.1007/978-1-60327-101-1_2