Research on the Application Logic and Ethical Boundaries of Artificial Intelligence in Systemic Financial Risk Early Warning

Zhenglin Chen[1], Zhendong Ma[2], Chengkai Kang[3]*

[1]Swinburne College of Shandong University of Science and Technology, Jinan, Chian

[2]Engineering college of China University of Petroleum-Beijing at Karamay, Karamay, China

[3]Trine Engineering Institute, Shaanxi University of Technology, Shaanxi, China

Received: 01 Nov 2025
Revised: 05 Nov 2025
Accepted: 08 Nov 2025
Published: 16 Nov 2025

Abstract: With the deep penetration of artificial intelligence (AI) technology into the field of financial risk control, its application in systemic risk early warning is attracting widespread attention from academia and regulatory agencies. This study focuses on the dual characteristics of AI technology in financial risk early warning: on the one hand, through multi-source data fusion, complex algorithm models, and intelligent decision support, AI technology has significantly improved its ability to identify high-dimensional nonlinear risks, realizing a paradigm shift from static assessment to dynamic early warning; on the other hand, ethical issues such as algorithmic bias, black-box decision-making, and data privacy violations are becoming increasingly prominent, and may even give rise to new systemic risks due to model homogeneity. This paper constructs a three-dimensional analysis system of "technology application-ethical boundaries-governance framework," and uses case analysis and comparative research methods to demonstrate the necessity of establishing a responsible AI early warning mechanism. The study finds that effective risk governance requires the organic combination of technological empowerment and ethical regulation. By developing explainable artificial intelligence, constructing an algorithm audit system, and improving cross-departmental collaborative supervision, while leveraging the early warning advantages of AI technology, it is essential to ensure that its application complies with the fundamental requirements of financial stability and fairness and justice. This provides an important reference for building a new paradigm of intelligent financial supervision in the digital age.

Keywords: Artificial intelligence; Systemic financial risk; Risk warning; Machine learning; Ethical boundaries

1 INTRODUCTION

The contemporary global financial system is facing increasingly severe systemic risk challenges. These risks are characterized by their inherent complexity, suddenness, and strong contagiousness. Traditional early warning mechanisms, relying mainly on linear models and limited macroeconomic indicators, often fall short in addressing the high-dimensional, nonlinear nature of modern financial risks, exhibiting significant early warning flags and blind

spots. Meanwhile, AI technology has made breakthrough progress in recent years and has begun to deeply penetrate all levels of financial operations. Its powerful data processing, pattern recognition, and predictive capabilities provide new technological possibilities for revolutionizing traditional risk control models [1]. Against this backdrop, a core question emerges: how can we effectively leverage the technological advantages of AI to improve early warning efficiency while prudently defining the ethical boundaries of its application, thus constructing a new paradigm for responsible systemic risk early warning? This question has both significant theoretical value and practical implications.

At the theoretical level, by deeply analyzing the internal logic and externalities of AI in risk early warning, we can enrich the theoretical connotations of these two interdisciplinary fields. At the practical level, the research results are expected to provide financial regulatory agencies with policy references that are both technically feasible and ethically reasonable, helping them to improve the predictability and effectiveness of macro-prudential supervision in the digital age [2]. A review of existing literature reveals that the academic community has accumulated considerable achievements in the evolution of systemic risk early warning models, the specific application of AI in financial risk monitoring, and branches such as AI ethics and compliance. However, there is still a lack of research that systematically integrates the logic of technological application with ethical regulation. This forms the starting point of this study.

Therefore, this study will follow the overall approach of "theoretical foundation - application analysis-ethical review-governance construction," comprehensively utilizing qualitative analysis, case studies, and necessary model construction methods to develop the discussion step by step. The full text first solidifies the theoretical foundation of artificial intelligence and systemic risk early warning, then systematically elaborates on its application logic and effectiveness, then focuses on analyzing the ethical challenges it raises, and finally proposes a specific path to build a responsible governance system, forming a logically rigorous and structurally complete analytical framework.

2 THE THEORETICAL BASIS OF ARTIFICIAL INTELLIGENCE AND SYSTEMIC FINANCIAL RISK EARLY WARNING

Section Two aims to build the theoretical foundation of the research, which first requires an in-depth analysis of the theoretical core of systemic financial risk. Systemic financial risk is not the collapse of a single institution, but rather the risk that the entire financial system may become dysfunctional due to damage to some links, ultimately seriously harming the real economy. Its core characteristics are contagion, externality, and procyclicality [3]. The formation mechanism of risk is complex, mainly due to the increasingly close network connection between financial institutions. The crisis of an important institution can be rapidly transmitted through channels such as asset sales and credit contraction; at the same time, the common risk exposure of financial institutions (such as holding similar asset portfolios) will resonate under negative shocks; and the inherent procyclicality of the financial system will amplify economic fluctuations. Such as during an economic downturn, the shrinkage of collateral value will trigger a chain of deleveraging behaviors, exacerbating the crisis.

To deal with such complex risks, artificial intelligence technology provides a series of powerful analytical tools. At the core technology level, machine learning, especially deep learning and unsupervised learning, is good at identifying potential risk patterns from massive amounts of data, even if these patterns have not yet been defined by traditional theories. Natural language processing technology can break through the limitations of traditional

quantitative data, extracting market sentiment and forward-looking risk information from news, social media reports, and regulatory documents, thus achieving efficient utilization of alternative data [4]. Furthermore, complex network analysis can intuitively map the risk contagion chain among financial institutions, while advanced time-series forecasting models can more accurately predict the evolution trajectory of risk factors. These technologies together form the cornerstone of AI early warning capabilities.

There is a high degree of inherent compatibility between artificial intelligence and systemic risk early warning. Traditional econometric models often struggle to handle the high-dimensional, nonlinear relationships common in financial markets, while AI models demonstrate significant advantages in this regard. Their powerful computing capabilities support real-time market monitoring, enabling a leap from static assessment to dynamic early warning. More importantly, the application of AI brings about a paradigm shift in methodology: it does not entirely rely on the rigorous pursuit of causality in traditional economics but rather focuses more on discovering strong correlations between variables through data mining. While this "association discovery" paradigm poses challenges to interpretability, it can provide more timely and sensitive risk signals when dealing with rapidly changing systemic risks whose mechanisms are not yet fully clear, thereby greatly improving the foresight of the early warning system.

3 APPLICATION LOGIC OF ARTIFICIAL INTELLIGENCE IN SYSTEMIC FINANCIAL RISK EARLY WARNING

After establishing a solid theoretical foundation, the specific application logic of artificial intelligence in systemic financial risk early warning can be clearly unfolded. Its core lies in constructing an intelligent framework from data to decision-making. This framework begins with a deep fusion layer of multi-source heterogeneous data, which transcends traditional market transaction and financial statement data, widely incorporating alternative data such as news sentiment, social media emotions, and even search engine trends, aiming to form a holographic risk perception network. Based on this, the AI core algorithm model layer plays the role of the "brain," using complex machine learning models for risk identification and correlation measurement, and dynamically assessing the system's vulnerability by simulating stress scenarios and risk contagion paths [5]. Finally, in the early warning signal output and decision support layer, the system transforms the complex indicators output by the model into operable early warning levels and visual reports, providing regulators with intuitive decision-making basis.

This framework has demonstrated strong potential in several typical scenarios. Such as based on complex network analysis, AI can accurately characterize the asset-liability relationships and risk exposure networks between financial institutions, thereby identifying key nodes in the system that have a "ripple effect" and simulating the contagion path of risk between these nodes, providing targets for precise intervention [6]. In market monitoring, deep learning models can learn patterns of extreme market volatility and asset bubble formation from historical high-dimensional data, providing earlier warnings of nonlinear abrupt changes such as sudden liquidity shortages. Simultaneously, utilizing natural language processing technology, the system can analyze macroeconomic policy statements and market sentiment in real time, quantifying market confidence levels and effectively monitoring confidence crises caused by reversals in expectations.

While its application prospects are broad, we must carefully assess its effectiveness and limitations. Undoubtedly, the introduction of artificial intelligence significantly improves the accuracy and timeliness of early warnings; its ability to handle complex nonlinear relationships allows it to capture weak signals that traditional models struggle to detect in the early stages of a crisis. However, its effectiveness is highly dependent on data quality and completeness,

posing a risk of "garbage in, garbage out." More importantly, the models themselves also carry risks. Such as their "black box" nature makes decision-making logic difficult to trace and verify, and biases learned from training data can lead to misjudgments. Models who are overly reliant on historical data may also fail to anticipate unprecedented new risks. These limitations warn us that while technology empowers, corresponding checks and balances must be established.

4 ETHICAL BOUNDARIES AND RISKS OF ARTIFICIAL INTELLIGENCE APPLICATIONS

Having witnessed the powerful capabilities of AI in risk warning, we must calmly examine the deep ethical boundaries it touches and the new risks it brings. This technology does not operate in a vacuum; its introduction will inevitably create complex frictions and clashes with the existing financial ecosystem, social norms, and regulatory framework.

First, the impartiality of the algorithm faces severe challenges. Its decisions are not always objective and neutral. If the training data contains historical biases, such as uneven data sampling for past risks of certain types of financial institutions or market sectors, it may lead to systematic misjudgments or discrimination against these objects [7]. More problematic is that, because the internal operating mechanisms of the model are often as difficult to interpret as a "black box," such potential biases are not only difficult to detect and correct in a timely manner, but may even self-reinforce during iterative learning, thus silently solidifying and amplifying existing injustices in the financial system, which runs counter to the original intention of maintaining financial stability and fairness.

Following this is the accountability dilemma caused by the lack of transparency. When regulatory agencies make major decisions based on an inexplicable AI warning, the legitimacy and credibility of their actions will be challenged. If the basis for decision-making cannot be clearly explained to the market and the public, the foundation of trust will be shaken [8]. At the operational level, once a warning signal is issued, if the source and transmission path of risk cannot be clearly traced and explained due to the complexity of the model, regulators will find it difficult to formulate precise and effective intervention measures, potentially leading to inappropriate responses and missing the best opportunity to mitigate risks.

Furthermore, the data-driven model has raised serious concerns about privacy and information security. To construct a holistic risk view, the system needs to collect and analyze a vast amount of sensitive information involving individuals and institutions, a process that can easily cross the boundaries of privacy protection [9]. Once this highly sensitive financial data is leaked due to security vulnerabilities or misused by insiders, the consequences will far exceed the individual level, potentially directly impact market confidence and even become a trigger for systemic risks.

Finally, from a macro perspective, the widespread application of AI itself may generate new systemic risks. If most market participants or regulatory agencies adopt similar or homogeneous warning models, it may lead to a high degree of consistency in risk perception and action, thus forming an algorithm-driven "herd effect," which in turn exacerbates market volatility. Even more paradoxically, a highly sensitive early warning system itself can become a source of market disturbance-if its signals are over-interpreted by the market, they may trigger irrational collective panic and sell-offs, thus transforming from a tool for warning of risks into a source of risk creation.

5 ETHICAL BOUNDARY GOVERNANCE: BUILDING A RESPONSIBLE AI EARLY WARNING SYSTEM

Faced with the complex ethical challenges posed by artificial intelligence in financial risk

early warning, building a responsible governance system is imperative. The cornerstone of this system lies in establishing clear core governance principles. We must adhere to a people-centered approach, ensuring that technological applications always serve financial stability and public welfare, and preventing algorithmic decisions from harming market fairness. The principles of transparency and explainability require that the logic of early warning models be understandable and scrutinized, which is a prerequisite for establishing regulatory credibility [10]. At the same time, the robustness and security of the system need to be fully guaranteed, capable of resisting potential attacks and maintaining stable operation. The principle of accountability and supervision throughout requires establishing a clear chain of responsibility for every algorithmic decision.

At the technical level, it is necessary to actively explore effective governance paths. The primary task is to vigorously develop explainable artificial intelligence technology, using methods such as model visualization and attribution analysis to unveil the mystery of the "black box," making the decision-making process traceable and understandable [11]. Simultaneously, a rigorous data quality assessment framework must be established to identify and correct potential biases during data collection and preprocessing, ensuring the fairness of algorithms from the source. Furthermore, routine model audits and stress tests should be implemented to regularly assess the performance, robustness, and potential biases of AI systems under different market environments, ensuring their continued reliable operation.

The realization of this technological path ultimately depends on robust regulation and institutional construction. Regulatory agencies should proactively embed AI ethical requirements into their regulatory technology systems and utilize innovative tools such as regulatory sandboxes to test and refine ethical governance solutions in a controlled environment. Legislation or industry standards must clearly define the boundaries of rights and responsibilities among AI system developers, users, and regulators, establishing a reasonable responsibility-sharing mechanism. Finally, given the complexity of the issue, a collaborative governance framework integrating finance, computer science, law, and ethics is needed. Through close cross-departmental cooperation, we can jointly address the new challenges brought by AI, ensuring that this powerful technology benefits the financial system within a safe framework.

6 CONCLUSIONS AND OUTLOOK

This study systematically explores the application of artificial intelligence (AI) in systemic financial risk early warning and its ethical boundaries. Analysis reveals that AI technology, through multi-source data fusion, complex algorithm models, and intelligent decision support, achieves efficient identification and early warning of high-dimensional nonlinear risks, significantly improving the initiative and accuracy of risk management. However, this technological empowerment process is also accompanied by serious ethical challenges such as algorithmic bias, lack of transparency, data privacy violations, and the potential to trigger new systemic risks. Therefore, building a responsible AI governance system is crucial. This requires establishing core principles of human-centeredness, fairness and transparency, and robustness and credibility. Ethical considerations should be embedded throughout the entire process of technology development and application through pathways such as developing explainable AI technologies, implementing routine model audits, and improving cross-departmental collaborative regulatory frameworks.

Based on these conclusions, this study proposes three policy recommendations. For regulatory agencies, they should accelerate the integration of ethical norms into the regulatory technology system, utilize regulatory sandboxes to test governance tools, and clarify accountability mechanisms for AI decision-making. Financial institutions, when introducing AI early warning systems, must establish internal ethics review committees, focusing on data

quality and model bias to ensure that the application of the technology complies with prudential regulatory requirements. For technology providers, it is necessary to proactively adhere to the principles of "ethical design," embedding interpretability and fairness modules into algorithm development and sharing responsibility for model risk management with financial institutions.

Of course, this research still has limitations. Such as the depth of discussion on specific technical paths for AI models is limited, and it is primarily based on theoretical deduction. Future research prospects are broad: on the one hand, with the maturation of privacy-preserving computing technologies such as federated learning, how to achieve effective risk warnings while keeping data "usable but not visible" will become an important direction; on the other hand, the development of cutting-edge technologies such as quantum computing may fundamentally reshape the capabilities of AI models, and their potential impact on financial stability requires forward-looking assessment. Ultimately, building a financial risk warning system that is both intelligent and robust requires continuous dialogue among technology experts, regulators, and ethicists to jointly explore the dynamic balance between development and security.

REFERENCES

[1] Chohan, M. A., Li, T., Ramakrishnan, S., & Sheraz, M. (2025). Artificial Intelligence in Financial Risk Early Warning Systems: A Bibliometric and Thematic Analysis of Emergin g Trends and Insights. *International Journal of Advanced Computer Science & Application s*, *16*(1). DOI: https://doi.org/10.14569/IJACSA.2025.01601127

[2] O'Halloran, S., & Nowaczyk, N. (2019). An artificial intelligence approach to regulating s ystemic risk. *Frontiers in Artificial Intelligence*, *2*, 7. DOI: https://doi.org/10.3389/frai.2019. 00007

[3] Andrae, S. (2025). Artificial Intelligence and Financial Stability: A Systemic Risk Approac h. In *Economic and Political Consequences of AI: Managing Creative Destruction* (pp. 87-110). IGI Global Scientific Publishing. DOI: https://doi.org/10.1016/j.jbankfin.2021.106290

[4] Kamruzzaman, M. M., Alruwaili, O., & Aldaghmani, D. (2024). RETRACTED: Measuring systemic and systematic risk in the financial markets using artificial intelligence. *Expert Systems*, *41*(5), e12971. DOI: https://doi.org/10.1111/exsy.12971

[5] Tang, P., Tang, T., & Lu, C. (2024). Predicting systemic financial risk with interpretable machine learning. *The North American Journal of Economics and Finance*, *71*, 102088. D OI: https://doi.org/10.1016/j.najef.2024.102088Get rights and content

[6] Olanrewaju, A. G. (2025). Artificial Intelligence in Financial Markets: Optimizing Risk M anagement, Portfolio Allocation, and Algorithmic Trading. *International Journal of Researc h Publication and Reviews*, *6*, 8855-8870. DOI: https://doi.org/10.55248/gengpi.6.0325.1218 5

[7] Giudici, P., Centurelli, M., & Turchetta, S. (2024). Artificial Intelligence risk measuremen t. *Expert Systems with Applications*, *235*, 121220. DOI: https://doi.org/10.1016/j.eswa.2023. 121220

[8] Dugbartey, A. N. (2025). Systemic financial risks in an era of geopolitical tensions, clima te change, and technological disruptions: Predictive analytics, stress testing and crisis respo nse strategies. *International Journal of Science and Research Archive*, *14*(02), 1428-1448. DOI: https://doi.org/10.1016/j.eswa.2023.121220

[9] Guo, X. (2024, December). Research on Systemic Financial Risk Early Warning Based on Integrated Classification Algorithm. In *2024 IEEE 2nd International Conference on Electr*

ical, Automation and Computer Engineering (ICEACE) (pp. 1586-1591). IEEE. DOI: https://doi.org/10.1109/ICEACE63551.2024.10898790

[10] Keller, A., Pereira, C. M., & Pires, M. L. (2024). The European Union's approach to artificial intelligence and the challenge of financial systemic risk. *Multidisciplinary Perspectives on Artificial Intelligence and the Law*, *415*. DOI: https://doi.org/10.1007/978-3-031-41264-6_22

[11] Abdulla, Y. Y., & Al-Alawi, A. I. (2024, January). Advances in Machine Learning for Financial Risk Management: A Systematic Literature Review. In *2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS)* (pp. 531-535). IEEE. DOI: https://doi.org/10.1109/ICETSIS61505.2024.10459536