WEP Warwick Evans Publishing

International Journal of Global Economics and Management

ISSN: 3005-9690 (Print), ISSN: 3005-8090 (Online) | Volume 10, Number 2, Year 2026
DOI: https://doi.org/10.62051/ijgem.v10n2.05
Journal homepage: https://ijgem.org

Artificial Intelligence in Securities Risk Management: Applications, Vulnerabilities, and Governance Frameworks

Lu Lu

College of Economics, GuangXi University, No. 100, Daxue East Road, Xixiangtang District, Nanning, Guangxi, China

ABSTRACT

The rapid development of artificial intelligence (AI) technologies has brought substantial convenience to the conduct of various business activities in securities firms, particularly in the field of risk management, where multiple large-scale models can be employed to address complex real-world problems. The application of AI techniques such as ensemble learning, deep learning and unsupervised learning in the risk management practices of securities firms makes it possible to process and analyze data rapidly, improve the efficiency and quality of decision making, and reduce labor costs, thereby enhancing the overall level of financial risk management. However, the adoption of these new technologies entails not only advantages but also notable drawbacks. The "black box" nature of many AI models weakens their interpretability, and the associated ethical issues, including data leakage and privacy protection, remain to be effectively resolved. These problems can in turn increase the difficulty and cost of regulation and may even trigger systemic risk. Consequently, AI technologies need to be properly aligned with and adapted to the financial system. In addition, all participants in financial markets should develop an appropriate understanding of AI and apply it in a prudent manner, so as to enable financial technology to better serve the functioning and development of financial markets.

KEYWORDS

Artificial Intelligence; Securities companies; Risk management

1. INTRODUCTION

In the financial industry, technological innovation has consistently served as the core driver for the evolution of business models and management methodologies. With the continuous advancement of artificial intelligence (AI) technologies, financial institutions represented by securities firms have achieved unprecedented improvements in the conduct of their business activities, including risk management. The integration of AI has rendered data analysis more precise and rapid, facilitating the analysis of market information, the identification of potential market risks, the optimization of decision-making processes, and the strengthening of responsiveness to unforeseen events. Compared to traditional models, the scope of data analysis has expanded from conventional statistical data to large-scale, multi-dimensional, and heterogeneous big data. This shift has enabled securities companies to achieve significant advancements in the core functional area of risk management.

Historically, the financial sector has relied on rigorous risk measurement methods to manage market volatility and uncertainty. Traditional risk measurement approaches, particularly those rooted in statistical modeling, constitute the theoretical foundation of this discipline. Classical statistical models, such as the mean variance model, the value at risk (VaR) model and the generalized autoregressive conditional heteroskedasticity (GARCH) model, draw on regression analysis of

historical data, volatility forecasting and related techniques to help financial institutions assess potential risks and implement risk control. These methods, predicated on assumed linear relationships and normal distribution patterns of market fluctuations, have played a pivotal role during specific historical periods and remain the cornerstone of financial risk management.

However, as financial market environments evolve and data volumes surge, the limitations of traditional statistical models in addressing complex, non-linear risk patterns have become increasingly evident. Concurrently, the rise of AI technology has provided the financial industry with a new suite of risk measurement tools. Artificial intelligence, specifically deep learning and machine learning, possesses the capability to uncover latent non-linear patterns and intricate interrelationships within massive datasets, transcending the constraints imposed by traditional models under specific hypothetical frameworks. Consequently, while traditional econometric models provide the preliminary framework for risk management, the introduction of AI technology undoubtedly offers more efficient and accurate solutions.

Despite AI's outstanding performance in enhancing efficiency and reducing operational costs, its implementation introduces a variety of challenges. Particularly in the field of risk management, issues such as the "black box" nature of AI, latent data security threats, and potential systemic vulnerabilities have emerged as critical subjects that financial institutions must urgently address.

2. LITERATURE REVIEW

2.1. Application of Artificial Intelligence in Financial Risk Management

Artificial intelligence technologies such as Artificial Neural Networks (ANN), Expert Systems (ES) and Hybrid Intelligent Systems (HIS) have been applied in many areas of the financial sector, including portfolio management, fraud detection, bankruptcy, stock management and risk management [1]. Each of these methods has its own characteristics, and their common features of high efficiency and autonomous learning enable them to handle a greater variety of complex problems. The Back Propagation (BP) neural network approach has been used to conduct empirical research on credit ratings in the Taiwanese and United States markets, thereby verifying the feasibility of this model for cross-market credit analysis [2]. On the basis of neural networks, an early warning system for logistics risks in small and medium-sized enterprises has been constructed, providing an effective tool for reducing the rate of corporate operational failure [3]. Neural networks have also been applied to the evaluation of credit loan defaults among Romanian enterprises, demonstrating their practical utility in credit risk prediction [4].

A rule based expert system has been developed to assist financial institutions in making scientifically grounded decisions during the process of corporate credit loan approval [5]. An expert system that combines a domain knowledge base with an operational knowledge base has been designed to evaluate the credit ratings of listed companies in Taiwan [6]. A further study has proposed a risk management expert system intended to provide data security and risk analysis support for SMEs with limited capital resources [7]. With respect to Support Vector Machines (SVM), SVM models have been used to model and forecast the default risks of German enterprises [8].

2.2. Risks Associated with the Application of Artificial Intelligence Technology

From the perspective of ethical risk, large-scale artificial intelligence models such as ChatGPT exhibit algorithmic bias, which gives rise to academic ethical risks including crises of trust in academic authority, unclear attribution of knowledge ownership, diffusion of contractual responsibility, and a lack of objectivity in academic evaluation [9]. The use of generative AI models may deconstruct the underlying concept of news values, lead to frequent occurrences of news-related infringements, and weaken both the value and the supervisory role of journalistic professionalism [10]. The causes of

these ethical risks are likely to be multifaceted, and it has been emphasized that they should be analyzed along three intrinsic dimensions, namely the actor, ideology and technology [11].

From the perspective of privacy risk, empirical observations indicate that, in concrete practical applications, the DeepSeek large model may collect users' input data and use it for model optimization; however, the lack of transparency in its algorithms may result in data misuse or the storage of data without consent [12]. In examining the data risks arising from the application of generative AI technologies, risks such as the unlawful acquisition of corpora, data leakage and bias in training data tend to emerge concurrently [13].

3. APPLICATION FRAMEWORK OF AI TECHNOLOGIES IN SECURITIES RISK MANAGEMENT

The application of AI in securities risk management is progressively evolving from experimental trials toward systematic integration. According to existing literature and industry practices, ensemble learning models, deep learning-based time-series prediction methods, and unsupervised learning-based anomaly detection algorithms constitute the three most widely implemented technologies. These methods focus on structured data prediction, dynamic temporal risk analysis, and the identification of unknown risks, respectively, forming a relatively comprehensive technical framework for risk management.

3.1. Risk Prediction Models Based on Ensemble Learning

Ensemble Learning represents a category of machine learning methodologies that combine multiple weak learners to achieve superior predictive performance, which is particularly prevalent in the risk management of securities firms. Its core philosophy is the "wisdom of the crowd", means leveraging the complementarity of different models to enhance overall accuracy and robustness. Common ensemble techniques include Bagging (Bootstrap Aggregating), Boosting, and Stacking, with Random Forest and Gradient Boosting Decision Trees (GBDT) being the most frequently utilized representations.

The advantages of ensemble learning models are primarily reflected in their ability to effectively mitigate the overfitting risks of single models and improve generalization capabilities. These models exhibit strong robustness against outliers and data noise, adapting well to complex financial market environments. Furthermore, they enhance interpretability through feature importance analysis, assisting in meeting regulatory compliance requirements. However, the disadvantages lie in their structural complexity, high computational overhead, and the tendency to generate redundant features in high-dimensional spaces, necessitating additional feature selection mechanisms. In practice, securities companies utilize ensemble learning for customer default risk scoring, expected loss forecasting for investment portfolios, and market volatility risk warnings.

3.2. Time-Series Risk Prediction Based on Deep Learning

Deep learning offers significant advantages in processing financial data characterized by non-linear features and complex temporal dependencies. Specifically, Recurrent Neural Networks (RNN) and their refined architecture, Long Short-Term Memory (LSTM) networks, are widely applied. LSTM models effectively overcome the gradient vanishing problem through their unique gating mechanism, enabling the retention of critical information over long durations. These models capture the dynamic patterns of variables such as transaction prices, volume, and volatility, thereby achieving predictions of future risk levels. In the practical risk management process, deep learning-based time-series prediction is employed to monitor market fluctuations, provide early warnings for significant asset price volatility, and forecast the future drawdown magnitude of investment portfolios.

3.3. Anomaly Detection Based on Unsupervised Learning

Anomaly detection identifies samples that significantly deviate from the majority of data, playing a pivotal role in securities operations such as anti-money laundering (AML), transaction fraud monitoring, and internal compliance risk identification. The advantage of unsupervised learning in this field lies in its independence from labeled data, enabling the identification of unknown risk events even in the absence of known fraud samples.

The strengths of unsupervised anomaly detection include its lack of requirement for labeled samples, making it suitable for addressing novel or rare risk events. Moreover, it can automatically discover hidden patterns within large-scale, multi-dimensional data, and its model update frequency is flexible, allowing for rapid adaptation to changes in data distribution. However, its disadvantages are also prominent, including a relatively high false-positive rate, heavy dependence on feature engineering, and insufficient interpretability of results. In securities application scenarios, unsupervised anomaly detection is utilized to identify abnormal large-value transactions, unusually frequent trading, and suspicious capital inflows and outflows.

4. IDENTIFYING POTENTIAL RISKS IN AI APPLICATIONS

4.1. Weak Robustness

Robustness refers to the ability of a model to maintain stable and accurate outputs when confronted with anomalies, noise, or unexpected inputs. When a model exhibits insufficient robustness, it becomes highly susceptible to anomalous inputs or adversarial attacks, leading to erroneous or unreliable decisions. As artificial intelligence technology is currently in a stage of incomplete maturity, it remains vulnerable to complex and volatile operational environments or malicious interference and induction. This vulnerability can result in performance degradation and decision-making errors. If AI fails to demonstrate adequate stability during major market events, it is likely to provide misleading guidance to enterprises, thereby exacerbating their operational burdens.

The currently emerging generative AI models can generate new data on the basis of patterns learned from existing information, which may include text, images, audio, or other types of media. While this ability to create coherent outputs opens various possibilities for different industries, it may also generate "hallucinations" during the output process. Specifically, the model may generate content that appears plausible but is factually incorrect, leading to knowledge bias and misinformation. Because these models struggle to distinguish between semantic correctness and factual accuracy, the inherent semantic coherence of the generated content makes such "hallucinations" difficult to detect and correct on time.

4.2. Insufficient Interpretability

In the field of artificial intelligence, particularly concerning complex large-scale models such as the GPT series, users face a unique challenge: the internal working mechanisms of these models are often perceived as a "black box." Although the underlying code, parameters, and training methodologies may be transparent, it remains exceedingly difficult to track and comprehend how specific inputs are transformed into specific outputs. This complexity arises from multi-layered non-linear data processing and the intricate interaction of a massive number of parameters, making the tracing and understanding of decision paths a significant challenge.

In economic and financial sectors, great emphasis is placed on logical rationality and coherence. Understanding how input information derives a specific output is a critical analytical component for decision-making. The existence of the "black box" implies that the derivation process of results obtained from AI models cannot be traced. This lack of interpretability makes it difficult to effectively

evaluate the reliability, accuracy, and fairness of the models. Consequently, when certain decisions made by securities companies perplex investors, financial institutions may face significant pressure to explain the decision-making process, the underlying logic, and the specific data utilized to the parties involved.

4.3. Privacy Leakage Issues

The leakage of training data represents a core risk that may lead to the theft of customers' personal information. The formidable capabilities of AI models are built upon massive datasets that often contain extensive personal identifiers, medical records, financial data, and even biometric information. Attackers may employ specific "reverse engineering" techniques or "data extraction attacks" to retrieve original sensitive data from pre-trained models. Furthermore, if strict access control and encryption measures are lacking during the collection and storage of training data, the data may be compromised before it is even utilized by the model.

Risks during the deployment and application phases of AI models are equally prominent. Currently, it is not uncommon for securities firms to use AI to handle certain basic services for clients. If some AI assistants or agents are granted excessive privileges and allowed to access sensitive data beyond the scope of their designated functions, any malicious exploitation of such access may result in serious internal data breaches.

5. RISK MANAGEMENT STRATEGIES FOR ARTIFICIAL INTELLIGENCE

5.1. Enhancing Algorithmic Interpretability

At the technical and operational levels, the core challenges reside in the autonomous decision-making capabilities and the "black box" nature of AI systems. To effectively manage these risks, it is imperative to ensure human control over AI during both the system design and operational phases. To guarantee algorithmic fairness and non-discrimination, technical standards must be established, and the compliance of AI systems should be continuously evaluated through independent auditing mechanisms. In addition, the deployment of security safeguards, including alert mechanisms embedded within algorithms and the use of data encryption technologies, constitutes a necessary measure to prevent misuse and cyberattacks.

Technical researchers must also strive to develop more transparent algorithms and models. Through research into Explainable Artificial Intelligence (XAI), efforts can be made to enable AI systems to explain their own decision-making processes. This may involve refining algorithmic structures and developing new interpretative tools to help users better understand how an AI model reaches a specific conclusion. Additionally, exploring novel algorithmic architectures and training methods can reduce structural complexity and opacity. Hybrid approaches that combine rule-based algorithms with machine learning can allow AI systems to maintain operational flexibility while ensuring a requisite level of interpretability.

5.2. Managing Ethical Risks

The widespread application of AI introduces risks at both the organizational and societal levels. Within institutions, maintaining transparent AI workflows and establishing clear user interfaces for AI control help prevent knowledge loss during functional transitions and serve as traceable evidence of AI activities. Moreover, institutions should position AI as an auxiliary tool designed to enhance human capabilities and optimize workflows, rather than as a categorical replacement for human employees. This transition requires effective change management to synchronize the impacts of AI implementation with all stakeholders and provide employees with opportunities for retraining and career transition.

At the societal level, institutions and governments should establish a set of binding ethical codes of conduct to ensure that AI decision-making processes respect fundamental human values and human rights. To this end, big data ethical principles should be formulated to regulate the rational use of personal data. Simultaneously, the right to human oversight over AI activities must be ensured to avoid unfair or discriminatory autonomous decisions. Furthermore, public education regarding AI should be strengthened to help investors understand the fundamental principles, application scope, and potential risks of AI, thereby increasing public awareness and understanding of the "black box" issue. This objective can be achieved through various channels, including investor education programs and media publicity.

6. CONCLUSION

Artificial intelligence (AI) has become increasingly integrated into various facets of social production and daily life, and its application within the risk management of securities institutions has emerged as a significant development. The diverse modeling methodologies inherent in AI ensure that risk identification is no longer confined to the linear frameworks of traditional statistical models; instead, it enables the capture of complex market dynamics and non-linear relationships across higher dimensions. Leveraging its capabilities for the efficient processing and real-time analysis of massive, heterogeneous datasets, AI has markedly enhanced the sensitivity and timeliness of risk monitoring. This provides securities companies with more effective tools to address sudden market fluctuations and mitigate potential systemic risks.

However, the application of any technology entails inherent risks, and its utilization within the financial domain, in particular, necessitates rigorous risk management measures. Algorithmic opacity may lead to a lack of interpretability in risk assessment, while issues regarding data quality and privacy protection may compromise model stability and regulatory compliance. Furthermore, an excessive reliance on automation may undermine the role of human prudent judgment. If these challenges are neglected, the technological advantages of AI may not only fail to be fully realized but could also introduce novel systemic risks. Consequently, while continuously advancing algorithmic innovation, it is essential to strengthen industrial regulatory frameworks and compliance mechanisms, while placing a strategic emphasis on the cultivation of interdisciplinary talent. Such efforts will facilitate a benign interaction between artificial intelligence and risk management, ultimately promoting the long-term, robust development of risk governance within the securities industry.

REFERENCES

[1] Swapnaja Gadre-Patwardhan,Vivek V. Katdare,Manish R. Joshi. A Review of Artificially Intelligent Applications in the Financial Domain [J]. Artificial Intelligence in Financial Markets, 2016: 3-44.

[2] Huang Z., Chen H., Hsu C.J, Chen W.H., Wu S..Credit Rating Analysis with Support Vector Machines and Neural Network:A Market Comparative Study [J]. Decision Support Systems, 2004, 37 (4) .

[3] Xie, KF (Xie, Kefan), et al. Early-warning management of inner logistics risk in SMEs based on label-card system [J]. Production Planning & Control, 2009, Vol.20(4): 306-319.

[4] Dima, Alina Mihaela, Vasilache, et al. Credit Risk modeling for Companies Default Prediction using Neural Networks [J]. Romanian Journal of Economic Forecasting, 2016, Vol.19(3): 127-143.

[5] Luke Hodgkinson, Ellen Walker. An expert system for credit evaluation and explanation [J]. Journal of Computing Sciences in Colleges, 2003, Vol.19(1): 62-72.

[6] Shue L.Y., Chen C.W. Shiue W. The Development of An Ontology-based Expert System for Corporate financial Rating [J]. Expert Systems with Applications, 2009, 36 (2).

[7] Janulevicius J, Goranin N..Expert System for Data Security Risk Management for SMEs [J]. Science-Future of Lithuania, 2013, 5 (2).

[8] Chen, SY (Chen, Shiyi) et al. Modeling default risk with support vector machines [J]. Quantitative Finance, 2011, Vol.11(1): 135-154.

[9] Xiao, H. J., and Z. Yang. Ethical regulation of digital technology: International Experience and reference from China. [J]. Journal of Dongbei University of Finance and Economics, 2023(5):47-61.

[10] Xue, J.H. Strategies for the Development of Artificial Intelligence Industry in the United States and Their Enlightenment for Building China's ArtificialIntelligence Ecosystem [J]. Science and Technology Management Research, 2022, 42(6):85-93.

[11] Tian, X.M. The Main Ethical Risks in Digital Society and Their Countermeasures [J]. Academic Journal of Zhongzhou, 2022(2):87-93.

[12] Jia, W.J. Cognitive Labor in Al Product Production in the Era of ArtificialIntelligence [J]. Academic Journal of Xi'an Jiaotong University, 2022, 42(6):85-93.

[13] JI SHI, MINWOO LEE, V. G. GIRISH, et al. Embracing the ChatGPT revolution: unlocking new horizons for tourism [J]. Journal of hospitality and tourism technology, 2024, 15(3): 433-448.