

Model Risk Management in the Era of Generative AI: Challenges, Opportunities, and Future Directions

Joshi, Satyadhar (2025): *Model Risk Management in the Era of Generative AI: Challenges, Opportunities, and Future Directions*. Published in: International Journal of Scientific and Research Publications , Vol. 5, No. 15 (20 May 2025): pp. 299-309.



PDF
MPRA_paper_125023.pdf
[Download \(274kB\)](#) | [Preview](#)

Abstract

The rapid adoption of generative AI in various sectors, particularly in finance, has introduced new challenges and opportunities for model risk management (MRM). This paper provides a comprehensive review of the current state of MRM in the context of generative AI, focusing on the risks, regulatory frameworks, and mitigation strategies. We explore the implications of generative AI on financial institutions, the evolving regulatory landscape, and the role of advanced MRM frameworks in ensuring compliance and mitigating risks. By synthesizing insights from 50+ recent articles, this paper aims to provide a roadmap for future research and practical applications of MRM in the generative AI era. It examines the key risks associated with these models, including bias, lack of transparency, and potential for misuse, and explores the regulatory frameworks and best practices being developed to mitigate these risks. We delve into the specific challenges faced by financial institutions in adapting their MRM strategies to encompass generative AI, and highlight the emerging tools and technologies that can support effective risk management.

This paper also discusses quantitative methods for risk quantification, such as probabilistic frameworks, Monte Carlo simulations, and adversarial risk metrics, which are essential for assessing the reliability and robustness of generative AI models. Foundational metrics, including fairness measures like demographic parity and equalized odds, are explored to address bias and ensure ethical AI deployment. Additionally, the paper presents pseudocode for key algorithms, such as risk quantification and adversarial risk calculation, to provide a practical understanding of these methods. A detailed gap analysis identifies critical shortcomings in current MRM frameworks, such as the lack of standardized validation methods and inadequate handling of adversarial robustness. Based on these gaps, the paper proposes solutions, including the development of advanced validation frameworks, integration of fairness metrics, and alignment with regulatory standards. These findings and proposals aim to guide financial institutions in adopting generative AI responsibly while addressing the unique risks it poses. This paper serves as a valuable resource for professionals and researchers seeking to understand and navigate the complexities of MRM in the age of generative AI.

Item Type: MPRA Paper

Original Title: Model Risk Management in the Era of Generative AI: Challenges, Opportunities, and Future Directions

Language: English

Keywords: Model Risk Management, Generative AI, Financial Institutions, Regulatory Compliance, Risk Mitigation, AI Governance.

Subjects: [C - Mathematical and Quantitative Methods](#) > [C5 - Econometric Modeling](#) > [C52 - Model Evaluation, Validation, and Selection](#)

Item ID: 125221

Depositing User: [Mr Satyadhar Joshi](#)

Date Deposited: 16 Oct 2025 07:18

Last Modified: 16 Oct 2025 08:02

References:

- [1] "AI model governance: What it is and why it's important Colibra." Accessed: Mar. 18, 2025. [Online]. Available: <https://www.colibra.com/blog/ai-model-governance-what-it-is-and-why-its-important>
- [2] "AI Use by Financial Institutions OSFI and FCAC Recommendations for Sound Risk Management McMillan LLP." Accessed: Mar. 18, 2025. [Online]. Available: <https://mcmillan.ca/insights/ai-use-by-financial-institutions-osfi-and-fcac-recommendations-for-sound-risk-management/>
- [3] J. C. Crisanto, C. B. Leuterio, J. Prenio, and J. Yong, "Regulating AI in the financial sector: Recent developments and main challenges," Dec. 2024, Accessed: Mar. 18, 2025. [Online]. Available: <https://www.bis.org/fsi/publ/insights63.htm>