Financial Risk and Management:
An International Journal
Volume 3, Number 1, 2024

Big Data Analytics for Fraud Detection and Risk Mitigation in Financial Services

Yoga Dwi Satria[1]

[1] Universitas Sarjanawiyata Tamansiswa, Yogyakarta, Indonesia

Abstract

Article history:
Received: January 11, 2024
Revised: February 25, 2024
Accepted: April 6, 2024
Published: June 30, 2024

Keywords:
Big Data Analytics,
Financial Fraud Detection,
Machine Learning,
Risk Governance,
Systematic Literature Review.

Identifier:
Zera Open
Page: 14-26
https://zeraopen.com/journal/frmij

This study examines how big data analytics is transforming fraud detection and risk mitigation in financial services. The rapid expansion of digital finance has amplified cybercrime, money laundering, and complex cross channel fraud, exposing the limitations of traditional rule based and manual controls. Using a systematic literature review approach, this paper synthesizes evidence on the deployment of big data platforms, machine learning, and deep learning across banking, insurance, and capital markets. The findings show that analytics driven systems, particularly those combining supervised and unsupervised techniques, enhance the detection of anomalous behavior, capture rare and non linear fraud patterns, and support continuous, real time surveillance of transaction populations. Beyond transaction level monitoring, big data analytics is increasingly embedded in early warning systems, credit scoring, stress testing, and portfolio monitoring, thereby strengthening credit, operational, and reputational risk management. However, persistent challenges related to data quality, legacy infrastructures, model transparency, regulatory and ethical concerns, and shortages of analytics skills constrain the full realization of these benefits. Overall, the review underscores that big data analytics delivers the greatest value when integrated into institution wide risk governance frameworks that align technology, data governance, regulation, and human expertise.

1. Introduction

Big data analytics has become central to how financial institutions respond to escalating fraud and risk in an increasingly digital financial ecosystem. The rapid growth of online and mobile banking, instant payments, and platform based finance has expanded the attack surface for cybercrime, money laundering, and opportunistic behavior, while also increasing the volume, velocity, and variety of data generated by financial transactions. Recent reviews show that traditional rule based systems and manual investigations struggle to keep up with complex, evolving fraud patterns and cross channel schemes, prompting financial institutions to adopt data-driven and analytics intensive approaches to fraud detection and compliance (Goecks et al., 2022; Mahya et al., 2023).

Big data analytics (BDA) leverages large, heterogeneous datasets and advanced analytical techniques to uncover hidden patterns, anomalies, and network relationships that signal fraudulent behavior or emerging risk exposures. In banking, BDA is increasingly embedded into risk management architectures, supporting early warning systems, credit scoring, stress testing, and portfolio monitoring (Dicuonzo et al., 2019). Studies highlight that, when combined with scalable data infrastructures, BDA can significantly enhance the precision and timeliness of risk assessment, making risk management more forward looking and dynamic compared with traditional reporting based approaches (Pillai, 2023).

In the specific domain of fraud detection, financial institutions are integrating big data platforms with machine learning and deep learning models to move from sample based checks to continuous monitoring of entire transaction populations.

Empirical evidence shows that big data enabled fraud detection systems can capture non linear relationships, rare events, and cross product linkages, improving detection rates and reducing false positives across credit card fraud, insurance fraud, and financial statement manipulation (Elkington et al., 2021; Gomes et al., 2021; Jan, 2021). Bibliometric and systematic reviews further document a growing research focus on big data analytics for financial fraud, spanning supervised classification, unsupervised anomaly detection, and graph based methods that exploit relational information among customers, accounts, and counterparties (Abrol & Gupta, 2023; Al-Hashedi & Magalingam, 2021; Mahya et al., 2023).

Beyond directly flagging suspicious transactions, big data analytics also contributes to broader risk mitigation in financial services. Real time transaction scoring, behavior profiling, and network analytics can reduce credit, operational, and reputational risk by enabling earlier intervention, targeted customer due diligence, and more robust internal controls (Dicuonzo et al., 2019). From a governance and assurance perspective, auditors and regulators are beginning to use BDA to strengthen fraud risk assessment and oversight, employing predictive modeling and visual analytics to prioritize high-risk entities and automate parts of the investigative process (Hussin & Musyaffi, 2022; Goecks et al., 2022).

However, the literature also notes important challenges and unresolved issues. These include data quality and integration problems, model transparency and explainability, regulatory and ethical concerns related to privacy and bias, and capability gaps in analytics skills and organizational readiness (Elkington et al., 2021; Abrol & Gupta, 2023). Existing studies are often fragmented, focusing either on

algorithmic performance in specific fraud us cases or on conceptual discussions of risk management 4.0 in banks, with relatively few works providing an integrated view of how big data analytics simultaneously supports fraud detection and enterprise wide risk mitigation. Addressing this gap, the present study synthesizes recent evidence on analytics driven fraud detection and risk control in finance, and highlights key technological, organizational, and regulatory factors that determine the effectiveness of contemporary risk management frameworks.

2. Literature Review

Big data analytics has been widely positioned as a core technological foundation for modern fraud detection in financial services. Systematic reviews on financial fraud detection highlight that machine learning based models, when trained on large, high dimensional datasets, consistently outperform traditional statistical techniques in identifying anomalous transactions and complex fraud patterns across credit cards, online banking, and other digital channels (Ali et al., 2022). These models leverage supervised, unsupervised, and hybrid approaches to capture subtle irregularities in customer behavior, transaction timing, and network relationships that are difficult to detect with rule based systems alone.

Building on this, more recent work emphasizes the growing role of unsupervised learning in big-data fraud analytics. A systematic literature review on machine learning for financial fraud detection finds that clustering and anomaly detection techniques are increasingly used to exploit unlabeled transaction data, uncover hidden structures, and detect previously unknown fraud schemes in banking

and capital markets (Husnaningtyas & Dewayanto, 2023). This stream of research stresses that big data environments characterized by large volumes and velocity of streaming transactions are particularly well suited to unsupervised models that can adapt to evolving fraud tactics without relying solely on historical labels.

Within the audit and assurance domain, big data analytics is also recognized as a key enabler of more effective fraud risk assessment. A systematic literature review on the impact of big data analytics adoption in audit shows that integrating large, heterogeneous datasets into audit procedures enhances the identification of red flags, supports continuous auditing, and allows auditors to focus on high risk items rather than relying on small samples (Saragih, 2023). By combining structured accounting data with unstructured sources such as emails or logs, audit analytics can reveal unusual relationships and transaction paths, thereby strengthening the fraud detection capabilities of internal and external auditors in financial institutions.

In parallel, the literature links big data analytics to broader risk mitigation in banking. A review of big data applications in banking and financial services argues that banks increasingly use big data architectures to reinforce risk management frameworks, especially for credit, market, and operational risk, by integrating data from offline branches, digital channels, and social platforms into unified analytics pipelines (Chang, 2020). These infrastructures support more granular risk profiling, near real time monitoring, and scenario analysis, allowing institutions to detect emerging vulnerabilities and adjust risk limits or pricing more proactively than under traditional reporting cycles.

Finally, recent studies on technology enabled financial risk management underscore that big data analytics, often combined with artificial intelligence, improves the accuracy of risk prediction and enhances the timeliness of management responses to both fraud and non fraud risks. Evidence from the digital age risk management literature shows that advanced analytics can reduce unexpected losses, strengthen early warning systems, and support more resilient risk governance structures when properly integrated into policies, processes, and human expertise (Dianti, 2023). At the same time, these studies caution that data quality, privacy, infrastructure gaps, and shortages of analytics skills remain critical constraints, suggesting that the effectiveness of analytics driven approaches to managing fraud and broader financial risks ultimately depends on organizational readiness and the alignment between technology and the overall risk management strategy.

3. Methods

The study employs a systematic literature review (SLR) approach to synthesize existing evidence on the use of big data analytics for fraud detection and risk mitigation in financial services. The review begins with the formulation of research questions focusing on how big data analytics is applied to detect fraudulent activities and to support broader risk management in banking and other financial institutions. Based on these questions, a search strategy is developed using combinations of keywords such as “big data analytics,” “financial fraud detection,” “banking,” “risk management,” “machine learning,” and “financial services.” These keywords are applied across major academic databases (for example, multidisciplinary and

business/finance oriented databases) and complemented with backward and forward citation tracking to capture additional relevant studies. The initial search results are screened at the title and abstract level to remove duplicates and clearly irrelevant publications, followed by a full text assessment using predefined inclusion and exclusion criteria.

Studies are included if they explicitly examine big data or analytics driven techniques in the context of financial fraud, financial crime, or risk management within financial institutions, and if they provide conceptual frameworks, empirical findings, or structured reviews relevant to the research questions. Publications that focus solely on algorithm development without a clear financial or risk management context, as well as studies outside the financial services domain, are excluded. For all included articles, key information is systematically extracted, including research objectives, methodological approaches, data sources, types of analytics and models used, fraud or risk categories addressed, and main findings. The extracted data are then analyzed using thematic synthesis to identify recurring patterns and gaps across the literature, allowing the review to map how big data analytics is being deployed along the fraud detection and risk management cycle, and to highlight technological, organizational, and regulatory factors that shape its effectiveness.

4. Results and Discussion

The systematic review shows a clear and consistent pattern: financial institutions are moving from rule-based, sample oriented fraud controls toward analytics driven, real time monitoring architectures. Across the surveyed studies, big

data analytics is primarily deployed to address the limitations of traditional fraud detection systems that struggle with high transaction volumes, cross channel schemes, and rapidly evolving attack patterns (Goecks et al., 2022; Mahya et al., 2023). Machine learning models trained on large, high dimensional datasets are repeatedly found to outperform classical statistical or rule based approaches in detecting anomalous transactions, especially in credit card fraud and online banking contexts (Gomes et al., 2021; Ali et al., 2022). This evidence supports the contention in the introduction that big data platforms, when combined with advanced analytics, enable financial institutions to move from ex post investigations toward continuous, proactive fraud surveillance.

The results also highlight important differences in how supervised and unsupervised learning techniques are used. Supervised models, such as classification algorithms and deep learning architectures, dominate in contexts where labeled fraud data are available, delivering higher detection rates and lower false positives compared with legacy systems (Elkington et al., 2021; Jan, 2021). In contrast, unsupervised techniques clustering, anomaly detection, and graph based methods are increasingly used to exploit unlabeled transaction streams and relational information among customers, accounts, and counterparties, allowing the discovery of previously unknown fraud patterns (Al-Hashedi & Magalingam, 2021; Abrol & Gupta, 2023; Husnaningtyas & Dewayanto, 2023). This division of roles suggests that effective fraud analytics architectures tend to combine multiple model families: supervised models for known fraud signatures and unsupervised or hybrid approaches for emerging, stealthier schemes.

Beyond transaction level detection, the review indicates that big data analytics is being embedded into broader risk management and governance frameworks. In banking, architectures that integrate data from core systems, digital channels, and external sources are used to support early warning systems, credit scoring, stress testing, and portfolio monitoring (Dicuonzo et al., 2019; Chang, 2020). These infrastructures enable more granular risk profiling and near real time scenario analysis, making risk management more forward looking than traditional, report based approaches (Pillai, 2023; Dianti, 2023). From an assurance perspective, the adoption of analytics in audit functions allows internal and external auditors to move toward continuous auditing, focusing on high-risk items and leveraging both structured accounting data and unstructured evidence such as logs and communications (Hussin & Musyaffi, 2022; Saragih, 2023). Taken together, these findings support the view that big data analytics not only enhances fraud detection but also strengthens credit, operational, and reputational risk mitigation across the financial services value chain.

At the same time, the synthesized literature underscores a set of persistent constraints and implementation challenges. Several reviews emphasize that data quality and integration issues, legacy IT infrastructures, and fragmented data governance can undermine the potential of analytics initiatives, especially in institutions that have not fully modernized their information systems (Dicuonzo et al., 2019; Chang, 2020). Concerns about model transparency, explainability, and algorithmic bias are particularly salient in high stakes areas such as fraud flagging, customer due diligence, and credit decisions, where opaque models may raise

regulatory and ethical questions (Elkington et al., 2021; Abrol & Gupta, 2023). In addition, capability gaps in analytics skills and organizational readiness are repeatedly noted as critical bottlenecks, with evidence that advanced tools yield limited risk benefits when not accompanied by appropriate human expertise, governance structures, and alignment with existing risk management policies (Dianti, 2023; Mahya et al., 2023).

Overall, the results suggest that big data analytics delivers the strongest fraud detection and risk mitigation benefits when deployed as part of an integrated, institution-wide framework that combines multiple analytical techniques, robust data infrastructures, and supportive governance. The discussion across the reviewed studies converges on a common implication: technological sophistication alone is not sufficient. To fully realize the promise of analytics driven fraud and risk management, financial institutions must invest in data governance, model risk management, explainability, and human capital, while regulators and auditors continue to refine supervisory expectations around the use of advanced analytics in combating financial crime and safeguarding financial stability.

5. Conclusion

This study concludes that analytics driven approaches built on big data have fundamentally reshaped how financial institutions detect fraud and manage risk. Across the reviewed literature, big data platforms combined with machine learning and deep learning consistently outperform traditional rule based systems in identifying complex, cross channel fraud patterns and anomalous behaviors. The

evidence shows a clear shift from ex post, sample based checks toward continuous, real time surveillance that spans credit card transactions, online banking, insurance claims, and financial reporting. At the same time, the integration of analytics into early warning systems, credit scoring, stress testing, portfolio monitoring, and audit functions demonstrates that these technologies are not confined to fraud detection alone but underpin a broader strengthening of credit, operational, and reputational risk management in financial services.

However, the synthesis also highlights that technology by itself is not enough to guarantee better outcomes. Persistent challenges such as fragmented and poor quality data, legacy IT infrastructures, limited model transparency and explainability, concerns over privacy and bias, and shortages of analytics skills can significantly dilute the benefits of even the most sophisticated tools. The overall implication is that the greatest fraud and risk mitigation gains arise when big data analytics is embedded within an integrated, institution wide framework that combines robust data governance, model risk management, clear regulatory expectations, and strong human expertise. Future research and practice therefore need to focus less on isolated algorithmic performance and more on how technological, organizational, and regulatory dimensions can be aligned to support resilient, trustworthy, and accountable risk management in the digital financial ecosystem.

References

Abrol, S., & Gupta, M. (2023). Role of big data analytics in financial fraud detection: A bibliometric analysis. *Corporate Governance Insight, 5*(1).

Ali, A., Abd Razak, S., Othman, S. H., Eisa, T. A. E., Al-Dhaqm, A., Nasser, M., & Saif, A. (2022). Financial fraud detection based on machine learning: A systematic literature review. *Applied Sciences, 12*(19), 9637.

Al-Hashedi, K. G., & Magalingam, P. (2021). Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. *Computer Science Review, 40*, 100402.

Chang, V., Xiao, L., Xu, Q., & Arami, M. (2020, May). A review paper on the application of big data by banking institutions and related ethical issues and responses. In *2nd International Conference on Finance, Economics, Management and IT Business* (pp. 115–121). SciTePress.

Dianti, A. R. (2023). Enhancing financial risk management in the digital age: A systematic review. *Arthatama: Journal of Business Management and Accounting, 7*(2), 79–91.

Dicuonzo, G., Galeone, G., Zappimbulso, E., & Dell'Atti, V. (2019). Risk management 4.0: The role of big data analytics in the bank sector. *International Journal of Economics and Financial Issues, 9*(6), 40–47.

Elkington, J., Klitner, A., & Karim, K. (2021). Big data analytics and fraud detection: A study of Australian companies. *International Journal of Accounting Information Systems, 41*, 100562.

Goecks, L. S., Korzenowski, A. L., Gonçalves Terra Neto, P., de Souza, D. L., & Mareth, T. (2022). Anti-money laundering and financial fraud detection: A systematic literature review. *Intelligent Systems in Accounting, Finance and Management, 29*(2), 71–85.

Gomes, C., Jin, Z., & Yang, H. (2021). Insurance fraud detection with unsupervised deep learning. *Journal of Risk and Insurance, 88*(3), 591–624.

Husnaningtyas, N., & Dewayanto, T. (2023). Financial fraud detection and machine learning algorithm (unsupervised learning): Systematic literature review. *Jurnal Riset Akuntansi dan Bisnis Airlangga (JRABA), 8*(2).

Jan, C. L. (2021). Detection of financial statement fraud using deep learning for sustainable development of capital markets under information asymmetry. *Sustainability, 13*(17), 9879.

Mahya, L., Tarjo, T., Sanusi, Z. M., & Kurniawan, F. A. (2023). Intelligent automation of fraud detection and investigation: A bibliometric analysis approach. *Jurnal Reviu Akuntansi dan Keuangan, 13*(3), 588–613.

Pillai, V. (2023). Integrating AI-driven techniques in big data analytics: Enhancing decision-making in financial markets. *International Journal of Engineering and Computer Science, 12*(7), 18535.

Rosnidah, I., Johari, R. J., Hairudin, N. A. M., Hussin, S. A. H. S., & Musyaffi, A. M. (2022). Detecting and preventing fraud with big data analytics: Auditing perspective. *Journal of Governance and Regulation, 11*(4).

Saragih, A. D., & Dewayanto, T. (2023). Systematic literature review: Dampak teknologi big data analytics dalam mendeteksi fraud pada bidang audit. *Diponegoro Journal of Accounting, 12*(3).