Perbanas International Conference on Economics, Business, Management, Accounting and IT (PROFICIENT) 2026
"Integrity-Driven Innovation: Reimagining Business, Finance, and Digital Transformation for Sustainable Impact"

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

Artificial Intelligence for Cyber Threat Detection in the Banking Sector: A Systematic Review

Agum Poundra Mukti[1,a)], Muhammad Fauzzan[,b)], Tiolina Evi Nausta Pardede [3,c)]

[1,2,3]*Department of Economic and Business, Perbanas Institute, Jakarta, Indonesi*

[a)]Corresponding author: agum.poundra51@perbanas.id
[b)] muhammad.fauzzan53@perbanas.id
[c)] tiolina.evi@perbanas.id

Abstract. *The rapid digital transformation of the banking sector has significantly increased its exposure to sophisticated cyber threats, necessitating advanced detection mechanisms beyond traditional rule-based systems. This systematic review examines the application of Artificial Intelligence (AI) for cyber threat detection in the banking sector, synthesizing evidence from 41 peer-reviewed studies published between 2014 and 2025. The review follows PRISMA guidelines for study identification, screening, and inclusion. Findings reveal that Machine Learning (ML) and Deep Learning (DL) techniques, including hybrid models, deep neural networks, and autoencoders, have demonstrated detection accuracies up to 99.98% for network intrusions and significant reductions in false positive rates. Key applications include fraud detection, real-time transaction monitoring, behavioral biometrics-based authentication, and anomaly detection in financial networks. However, several barriers impede widespread adoption: system complexity, lack of AI-proficient personnel, regulatory compliance challenges, model interpretability concerns, and integration difficulties with legacy banking infrastructure. The review identifies that socio-technical challenges, including shadow AI usage and limited audit readiness, represent critical gaps in current research. This review contributes a comprehensive taxonomy of AI techniques for banking cybersecurity and provides actionable recommendations for financial institutions seeking to implement AI-driven threat detection systems. Future research directions include privacy-preserving techniques such as federated learning, adversarial robustness, and explainable AI frameworks tailored to regulatory requirements.*

Keywords: *Artificial Intelligence, Cyber Threat Detection, Banking Sector, Machine Learning, Systematic Review*

INTRODUCTION

The banking sector has undergone unprecedented digital transformation over the past decade, with financial institutions increasingly relying on online platforms, mobile banking applications, and real-time payment systems to deliver services to customers. While this digital evolution has enhanced operational efficiency and customer experience, it has simultaneously expanded the attack surface for cybercriminals. Financial institutions now face sophisticated threats including phishing attacks, ransomware, distributed denial-of-service (DDoS) attacks, and complex fraud schemes that traditional rule-based security systems struggle to detect .

The financial impact of cyber threats on the banking sector is substantial. In 2023 alone, credit card fraud generated global losses exceeding $32 billion, representing a 25% increase compared to the

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

previous year . Beyond direct financial losses, cyber incidents erode customer trust, trigger regulatory penalties, and can precipitate systemic risks affecting the broader financial ecosystem. Traditional security approaches, which rely on signature-based detection and predefined rules, have proven inadequate against zero-day exploits and adaptive attack methodologies that characterize contemporary cyber threats.

Artificial intelligence has emerged as a promising solution to address these limitations. AI-driven systems offer capabilities for real-time anomaly detection, behavioral analysis, and predictive threat intelligence that surpass conventional methods . By analyzing vast quantities of transactional data, user behavior patterns, and network traffic, machine learning algorithms can identify subtle indicators of compromise that would escape rule-based detection systems. Recent advances in deep learning have further enhanced these capabilities, enabling the identification of complex attack patterns across heterogeneous data sources .

Despite growing enthusiasm for AI-based cybersecurity solutions, evidence of widespread, trustworthy deployment in banking remains limited. Financial institutions operate under stringent regulatory oversight, including model risk management requirements, data privacy regulations such as GDPR, and emerging AI governance frameworks. These constraints create unique challenges for AI adoption that differ substantially from other industry sectors . Moreover, the adversarial nature of cybersecurity means that AI systems themselves may become targets, with attackers developing techniques to evade or poison detection models.

This paper presents a systematic review of research on artificial intelligence for cyber threat detection in the banking sector. The review addresses three primary research questions: (1) What AI techniques have been applied to cyber threat detection in banking, and with what reported effectiveness? (2) What barriers impede the trustworthy deployment of AI-driven detection systems in financial institutions? (3) What gaps exist in current research, and what directions should future research prioritize?

LITERATURE REVIEW

Cyber Threat Landscape in Banking

The banking sector faces a distinctive set of cyber threats shaped by the sensitivity of financial data, the real-time nature of transactions, and the systemic importance of financial institutions. Cybercriminals employ increasingly sophisticated techniques including credential theft, man-in-the-middle attacks, and social engineering to compromise banking systems. A systematic taxonomy of FinTech security threats has identified eleven distinct threat categories, with fraud and data breaches representing the most frequently documented concerns (Javaheri et al., n.d.) . Insider threats present particular challenges for banking institutions. Unlike external attacks, insider threats originate from individuals with legitimate access to systems and data, making detection substantially more difficult. Research has demonstrated that AI-driven anomaly detection can identify suspicious internal activities by establishing behavioral baselines and flagging deviations that may indicate malicious intent (Ajayi et al., 2024).

AI Techniques for Cybersecurity

Machine learning approaches to cybersecurity can be categorized into supervised, unsupervised, and reinforcement learning paradigms. Supervised learning requires labeled datasets containing examples of both benign and malicious activities, enabling models to learn distinguishing features. Unsupervised learning techniques, including clustering and autoencoders, identify anomalous patterns without requiring labeled training data, making them particularly valuable for detecting novel attack types (Hasan et al., 2025). Deep learning architectures, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have demonstrated particular effectiveness for sequence-based detection tasks such as identifying malicious network traffic patterns and analyzing log sequences. Hybrid models combining multiple AI techniques have achieved detection accuracies of up to 99% in research settings, with false positive reductions of 31% compared to traditional approaches .

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

Regulatory and Governance Context

The deployment of AI for cybersecurity in banking occurs within a complex regulatory environment. The European Union's General Data Protection Regulation imposes strict requirements on automated decision-making, including provisions for algorithmic transparency and the right to explanation. Emerging AI regulations, including the EU AI Act, establish risk-based classifications that affect how financial institutions can deploy AI systems .Financial regulators have increasingly focused on model risk management, requiring institutions to validate AI models, monitor their performance over time, and maintain audit trails of model decisions. These requirements create tensions with the inherent opacity of certain deep learning architectures, highlighting the need for explainable AI approaches that can satisfy regulatory expectations while maintaining detection effectiveness

METHODS

This study employed a systematic literature review methodology following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. This approach ensures transparency, reproducibility, and comprehensive coverage of relevant literature.

Search Strategy

A systematic search was conducted across five electronic databases: IEEE Xplore, Scopus, Web of Science, SpringerLink, and Google Scholar. The search was limited to peer-reviewed journal articles, conference proceedings, and book chapters published in English between January 2019 and May 2026. The search string combined terms related to artificial intelligence, cybersecurity, and banking: ("artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network" OR "anomaly detection") AND ("cyber threat" OR "cybersecurity" OR "intrusion detection" OR "fraud detection") AND ("banking" OR "financial institution" OR "financial services").

Inclusion and Exclusion Criteria

Studies were included if they: (1) addressed the application of AI techniques for cyber threat detection, (2) focused specifically on banking or financial institution contexts, (3) presented empirical results or systematic reviews, and (4) were published in peer-reviewed sources. Studies were excluded if they: (1) addressed only non-financial sectors, (2) focused exclusively on theoretical AI developments without cybersecurity application, (3) were duplicate publications, or (4) were editorials, commentaries, or white papers without original research.

Screening and Data Extraction

Initial database searches yielded 741 potentially relevant records. After removing duplicates, 555 unique records underwent title and abstract screening. Subsequently, 98 full-text articles were assessed for eligibility, resulting in 52 studies included in the final synthesis. Data extracted from each study included: AI technique employed, threat type addressed, detection performance metrics, evaluation methodology, geographical context, and reported implementation challenges.

RESULTS AND DISCUSSION

AI Techniques and Detection Performance

The reviewed studies demonstrate that machine learning techniques have been successfully applied to multiple cyber threat detection tasks in banking. Fraud detection represents the most extensively researched application, with supervised learning approaches achieving high accuracy in identifying potentially fraudulent transactions based on historical patterns. Anomaly detection techniques have proven effective for identifying previously unseen attack patterns, including zero-day exploits and novel fraud schemes.

Deep learning approaches, particularly when applied to sequential transaction data, have demonstrated superior performance compared to traditional machine learning algorithms. Recurrent neural networks and long short-term memory (LSTM) architectures can capture temporal dependencies in transaction sequences, enabling the identification of complex fraud patterns that unfold over multiple transactions. Convolutional neural networks have been applied to network traffic analysis, identifying malicious patterns in packet flows with high accuracy.

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

The most effective approaches reported in the literature employ ensemble methods or hybrid architectures combining multiple AI techniques. These systems leverage the complementary strengths of different algorithms, achieving detection accuracy up to 99% while maintaining false positive rates below industry thresholds. However, it is important to note that these performance figures predominantly derive from retrospective testing on historical datasets rather than prospective validation in live banking environments .

Implementation Barriers and Trustworthiness Challenges

Despite promising performance metrics, the reviewed literature reveals substantial barriers to trustworthy AI deployment in banking cybersecurity. Algorithmic explainability emerges as the most frequently cited challenge. Financial institutions require the ability to explain detection decisions to regulators, auditors, and potentially affected customers (Karaosman et al., n.d.). However, many high-performing deep learning models operate as "black boxes," producing decisions that cannot be readily explained in human-understandable terms.

Adversarial vulnerability represents another critical concern. Research has demonstrated that attackers can craft inputs specifically designed to evade AI detection systems, potentially rendering them ineffective against sophisticated adversaries. The banking sector's high-value targets make such adversarial attacks particularly concerning, yet few reviewed studies addressed robustness testing or adversarial defense mechanisms.

Integration challenges with legacy banking infrastructure were also widely reported. Many financial institutions maintain core banking systems developed over decades, which were not designed to interface with modern AI platforms. Data quality and availability issues further complicate implementation, as effective AI models require large volumes of clean, labeled training data that may not exist for emerging threat types.

Geographical and Sectoral Research Gaps

The review identified significant geographical disparities in research evidence. The majority of studies originated from Asian institutions (Ndukwe & Baridam, 2023), particularly India and China, followed by North American and European contributions. Conversely, no peer-reviewed studies documented AI-based cyber threat detection implementation in African banking institutions, despite evidence of increasing cyber threats across the continent. This geographical gap suggests that AI cybersecurity research may be concentrated in regions with mature financial technology sectors, while underrepresented regions may lack access to evidence-based guidance for security investments.

A related finding concerns the limited evidence of actual production deployments. Most reviewed studies reported experimental results using laboratory environments or retrospective datasets, with few documenting the challenges and outcomes of live deployment in operational banking systems. This gap between research and practice limits the generalizability of findings and may overstate the readiness of AI techniques for real-world application.

CONCLUSIONS

This systematic review has examined the application of artificial intelligence for cyber threat detection in the banking sector, synthesizing evidence from 52 peer-reviewed studies. The findings demonstrate that AI techniques, particularly machine learning and deep learning approaches, offer substantial potential for enhancing detection capabilities beyond traditional rule-based systems. Detection accuracy exceeding 99% has been reported for certain applications, with hybrid models showing particular promise. However, the review also reveals significant barriers to trustworthy deployment that must be addressed before AI can realize its full potential in banking cybersecurity. Algorithmic explainability, adversarial robustness, regulatory compliance, and legacy system integration represent interconnected challenges requiring coordinated attention from researchers, practitioners, and regulators. The current evidence base is characterized by geographical concentration and limited real-world deployment validation, suggesting that reported performance figures may overstate operational readiness. Future research should prioritize several directions. First, development of explainable AI architectures that maintain detection performance while enabling human-understandable explanations of model decisions. Second, empirical studies of AI deployment in operational banking environments, documenting implementation challenges and mitigation strategies. Third, research addressing geographical disparities, particularly examining barriers to AI

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

adoption in underrepresented regions. Fourth, investigation of adversarial robustness and defensive mechanisms tailored to banking sector threat models. Finally, the development of standardized evaluation frameworks and benchmark datasets to enable meaningful comparison across AI approaches. Financial institutions considering AI adoption for cyber threat detection should adopt a measured approach that prioritizes trustworthiness alongside performance. Organizations should begin with well-understood, explainable models applied to bounded use cases, gradually expanding scope as governance capabilities mature. Investment in data infrastructure, model validation processes, and cross-functional expertise spanning data science, cybersecurity, and regulatory compliance is essential for successful deployment.

Reference

Prediction of Cyberattacks using Machine Learning in the Financial Sector: Technical Review. *IEEE Xplore*, 2025.

人工智能与金融安全：技术脆弱性、风险演化路径与防范机制. *National Institute of Finance and Development*, 2026.

Javaheri, D., Fahmideh, M., Chizari, H., Lalbakhsh, P., & Hur, J. Cybersecurity threats in FinTech: A systematic review. *Expert Systems with Applications*, 238, 122697, 2024.

Omole, O.M., Ajayi, A., Omokanye, A.O., Olowu, O.J., & Adeleye, A.O. Detecting insider threats in banking using AI-driven anomaly detection with a data science approach to cybersecurity. *Semantic Scholar*, 2024.

Hasan, M.N., Papel, M.S.I., Rasel, I.H., Akter, S., Aktar, M.K., Abedin, M.Z., & Mani, L. Enhancing financial information security through advanced predictive analytics: A PRISMA based systematic review. *Edelweiss Applied Science and Technology*, 9(7), 2222-2245, 2025.

Securing financial systems through data sovereignty: a systematic review of approaches and regulations. *International Journal of Information Security*, 24, 159, 2025.

Mohammed, Y. AI in Organizations: Transforming Cybersecurity for Financial Institutions. In *AI-Driven Revolution: Transforming the Business Landscape* (pp. 65-79). World Scientific, 2025.

Karaosman, E., Rizvani, A., & Pekaric, I. Security Barriers to Trustworthy AI-Driven Cyber Threat Intelligence in Finance: Evidence from Practitioners. *Proceedings of CODASPY '26*, ACM, 2026.

Bello, A., Uzoma, C., Atadoga, S., Dehumo, A.P., & Akintola, A.S. Responsible AI in financial identity verification and risk mitigation. *Discover Sustainability*, 2026.

Ndukwe, & Baridam. A Graphical and Qualitative Review of Literature on AI-based Cyber Threat Intelligence (CTI) in the Banking Sector. *Journal of Cybersecurity Research*, 81(51), 2023.