



**AI-BASED BEHAVIOURAL ANOMALY DETECTION FOR CYBERSECURITY IN
DIGITAL BANKING**

Snega. R. P¹, M. Revathi², N. Selvaganesh³, T. Duraisamy⁴

¹Department of Information Technology
Sri Venkateswara College of Engineering
Sriperumbudur
snegacsengg@gmail.com

²Department of Computer Science and Engineering
Dhanalakshmi college of Engineering
Manimangalam
mdrevathi93@gmail.com

³Department of Information Technology
Sri Venkateswara College of Engineering
Sriperumbudur
selva492@gmail.com

⁴Artificial Intelligence and Data Science
A.K.T Memorial College of Engineering and Technology
Kallakurichi
duraisamycse@gmail.com

Abstract

India's banking system has moved rapidly from branch-based services toward mobile banking, internet banking, instant payments and data-intensive financial services. This change has created a large security surface in which fraud, phishing, malware, credential theft, account takeover, mule accounts and attacks on digital payment channels can occur at high speed. Artificial intelligence (AI), especially machine learning, can help banks examine large volumes of transactions, identify unusual behaviour, rank alerts, detect suspicious networks and support faster cybersecurity response. This research paper examines the role of AI in cybersecurity in India's banking sector using a documentary and secondary-data approach. The study uses Reserve Bank of India (RBI) evidence on reported bank frauds for 2022-23 to 2024-25, RBI material on artificial intelligence and supervisory analytics, The Indian Computer Emergency Response Team -CERT-In cyber incident requirements, and established academic literature on anomaly detection, explainable AI, adversarial machine learning and cyber threat detection. The RBI data show that reported fraud cases rose sharply in 2023-24 and remained substantial in 2024-25, while the distribution of cases and amounts differed across bank groups. In 2024-25, private sector banks reported 14,233 fraud cases and public sector banks reported 6,935; however, the public sector group accounted for the largest reported amount among the bank groups shown in the RBI table. The paper explains why AI



should be treated as a security capability rather than as a replacement for human investigation. It proposes a practical architecture combining data quality, behavioural analytics, graph analysis, threat intelligence, explainability, human review, model monitoring, privacy protection and incident response. The study concludes that responsible AI can strengthen banking cybersecurity when it is governed through clear accountability, independent validation, secure data practices and continuous testing.

Keywords: Artificial intelligence; cybersecurity; Indian banking; machine learning; fraud detection; anomaly detection; digital payments; RBI; CERT-In; responsible AI.

1. Introduction

Banking in India is now deeply connected with digital technology. Customers can open accounts, transfer money, pay merchants, apply for loans, receive alerts and manage investments through digital channels. The same digital environment also gives attackers more opportunities. A fraudster may use a stolen password, a fake application, social engineering, a malicious link, a compromised device or a network of mule accounts. The speed of digital transactions means that a suspicious event may develop before a traditional manual review is completed.

The RBI has treated cyber resilience as an important banking responsibility for many years. Its 2016 Cyber Security Framework in Banks called for board-approved cyber-security policy, adaptive incident response and recovery, and continuous improvement as threats change [1]. The later regulatory environment has added stronger fraud-risk management, early-warning systems, cyber incident monitoring and technology-enabled supervision.

Table 1. Selected 2024–25 fraud indicators from RBI [2, Table VI.3].

Area of operation	Cases	Amount (₹ crore)	Share of total cases
Advances	7,950	33,148	33.2%
Card/Internet	13,516	520	56.4%
Deposits	1,208	527	5.0%
Others	800	1,392	3.3%
Cash	306	39	1.3%

The RBI 2024–25 data indicate that Card/Internet frauds recorded the highest number of cases, with 13,516 cases representing 56.4% of the total, showing the importance of strengthening digital banking cybersecurity. However, Advances recorded the highest financial amount involved, at ₹33,148 crore, despite accounting for 7,950 cases or 33.2% of cases. Deposits accounted for 1,208 cases and ₹527 crore, while other areas recorded 800 cases involving ₹1,392 crore. Cash-related frauds were comparatively fewer, with 306 cases and ₹39 crore. The findings highlight the need for AI-based monitoring across both digital transactions and high-value banking activities.

AI-Driven Banking Security Analysis

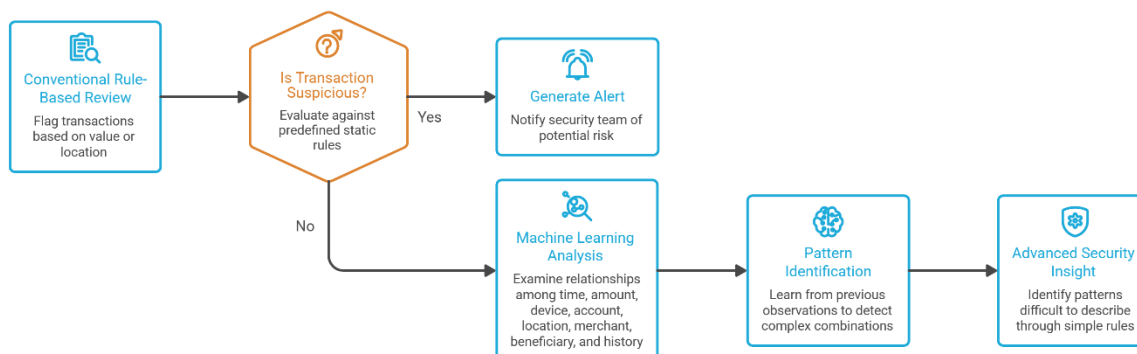


Fig:1 AI Driven Banking Security Analysis

Fig.1 explains AI changes the way security teams can work with banking data. A conventional rule may say that a transaction above a particular value or a transaction from an unusual location should be reviewed. Such rules remain useful, but they can generate too many alerts and may not identify new patterns. Machine-learning models can examine relationships among time, amount, device, account, location, merchant, beneficiary and transaction history. They can learn patterns from previous observations and identify combinations that are difficult to describe through simple rules [3], [4].

The Indian context is particularly important because the country combines very high digital-payment activity with a large and diverse banking population. RBI reported that its Digital Payments Index reached 465.33 in September 2024, up from 445.50 in March 2024 [5]. RBI has also reported phishing, vishing, remote access and other forms of digital-payment fraud [6]. This environment creates a strong case for examining AI as one component of a wider cyber-defence system.

This paper studies AI in banking cybersecurity from four connected angles: the security problems that AI can address; the Indian banking data and regulatory environment; the risks created by AI itself; and a practical model for responsible adoption. The purpose is not to claim that AI automatically reduces fraud. Instead, the paper asks how AI can improve detection and response when it is properly governed.

2. Research Objectives

1. To examine the major cybersecurity and fraud risks faced by digital banking in India.
2. To explain the applications of AI and machine learning in fraud detection, anomaly detection, identity protection, threat intelligence and incident response.
3. To analyse RBI secondary data on reported bank fraud cases and amounts for 2022-23 to 2024-25.
4. To identify the operational, ethical, privacy and security risks associated with AI-based cybersecurity systems.
5. To propose a practical and responsible AI framework for cybersecurity in Indian banks.

3. Research Questions

1. RQ1: What major cybersecurity problems in Indian banking can be addressed through AI?
2. RQ2: How can machine learning improve the detection and prioritisation of suspicious banking activity?
3. RQ3: What do recent RBI fraud data indicate about the security environment in Indian banking?
4. RQ4: What risks can arise when AI systems themselves become targets of cyberattacks?
5. RQ5: What governance and human-oversight measures are necessary for responsible AI-based banking cybersecurity?

4. Review of Literature

4.1 Artificial Intelligence and Cybersecurity

AI includes machine learning, neural networks, natural-language processing and other computational methods. In cybersecurity (Fig.2), its value comes from the ability to process large amounts of information and identify patterns. Bishop [7] explains machine learning as a framework for learning from data, while Goodfellow et al. [8] show how deep learning can represent complex patterns. In banking, these capabilities can be applied to transaction streams, login events, network traffic, device signals and customer behaviour.

Anomaly detection is particularly relevant. Chandola et al. [4] describe anomaly detection as the identification of observations that differ substantially from expected behaviour. This is useful when a bank does not know every possible attack in advance. Sommer and Paxson [9], however, caution that machine learning in computer security has practical difficulties because attackers adapt, data can be incomplete, and security environments are different from ordinary prediction problems. AI should therefore be treated as part of an operational security process.

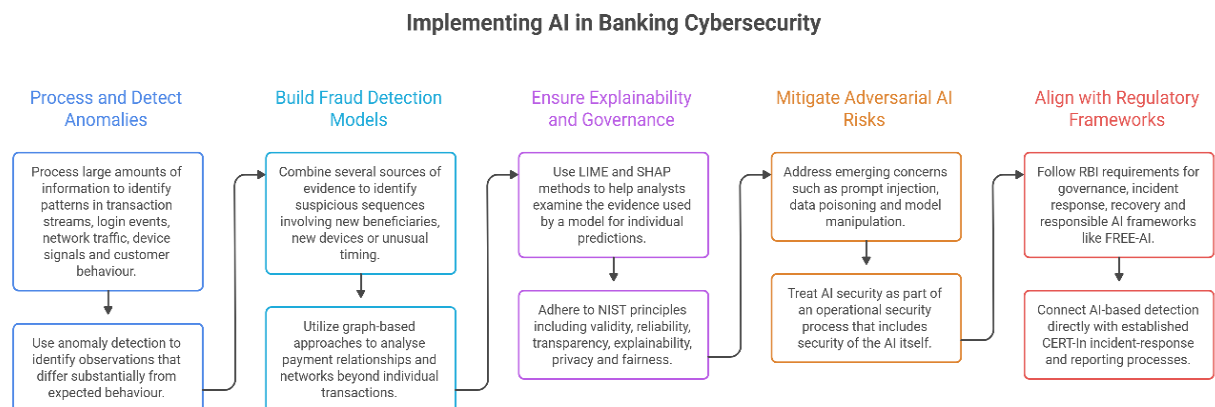


Fig.2 Implementing AI in Banking Cyber Security

4.2 Machine Learning for Fraud and Intrusion Detection

Fraud detection often involves classification, anomaly detection and risk scoring. Breiman's [10] random-forest approach is useful where many variables interact, while Buczak and



Guven [3] review machine-learning methods for intrusion detection and show the importance of selecting suitable algorithms, features and evaluation measures. Shaukat et al. [11] similarly describe the growing role of machine learning in cybersecurity, including intrusion detection, malware analysis and anomaly identification.

The most useful banking models are likely to combine several sources of evidence. A single transaction may look normal, but a sequence involving a newly created beneficiary, a new device, unusual timing and a network of related accounts may be suspicious. Graph-based approaches can therefore complement transaction-level models. Recent financial-sector research also points to the value of analysing payment relationships and networks rather than only individual transactions [12].

4.3 Explainable and Auditable AI

A banking cybersecurity model should not merely produce a score. Security officers need to understand why an alert was raised, especially when a customer account may be blocked or a transaction delayed. Ribeiro et al. [13] introduced LIME for explaining individual predictions, while Lundberg and Lee [14] proposed SHAP as a unified approach to interpreting model outputs. These methods can help analysts examine the evidence used by a model.

Explainability is also connected with governance. The National Institute of Standards and Technology [15] emphasises trustworthy AI characteristics such as validity, reliability, transparency, explainability, privacy and fairness. Banking cybersecurity models require these principles because they operate with sensitive financial and personal information.

4.4 Adversarial AI and New Cyber Risks

AI is not only a defensive tool. Attackers can target the model, the training data or the data pipeline. Papernot et al. [16] demonstrated adversarial examples that can cause machine-learning systems to make incorrect predictions, while Carlini and Wagner [17] showed optimisation-based attacks against neural networks. In banking, an attack on a fraud model could create false negatives, allowing suspicious transactions to pass, or false positives, creating unnecessary customer disruption.

Generative AI adds another layer. Aldasoro et al. [12] note that generative AI may improve cyber defence but can also increase social-engineering risks and data-disclosure risks. BIS research also identifies prompt injection, data poisoning and model manipulation as emerging concerns for AI-enabled financial systems [18]. These risks mean that AI security must include security of the AI itself.

5. Indian Banking and Cybersecurity Context

The Indian banking sector operates within a detailed regulatory environment. RBI's Cyber Security Framework in Banks requires banks to maintain a structured approach to cyber threats, including governance, incident response and recovery [1]. RBI's later supervisory work has expanded into AI and machine learning. In 2023-24, RBI reported work on a fraud vulnerability index using machine learning and the strengthening of early-warning and red-flag systems [19]. In 2024-25, RBI reported that its Advanced Supervisory Analytics Group



had developed models including a fraud vulnerability index, borrowers' vulnerability model and asset-quality prediction model [2].

RBI has also moved toward responsible AI. Its 2024-25 Annual Report states that an external committee was constituted in December 2024 to recommend a Framework for Responsible and Ethical Enablement of AI in the financial sector, known as FREE-AI [2]. This is important because AI in banking is not only a technical issue. It involves governance, customer rights, data protection, model risk, accountability and third-party technology providers.

CERT-In provides another important layer. Its 2022 directions identify incidents such as unauthorised access, phishing, ransomware, data breaches, identity theft, denial-of-service attacks, digital-payment attacks and malicious mobile applications as reportable categories. Covered organisations are required to report specified cyber incidents within the prescribed time after noticing them [20], [21]. For banks, AI-based detection is therefore most useful when it connects directly with established incident-response and reporting processes.

6. Research Methodology

The study adopts a descriptive and analytical documentary research design. It does not claim to measure the performance of one particular AI system used by an individual Indian bank. Instead, it combines official secondary data, regulatory documents and established academic literature to build a sector-level analysis.

6.1 Data Sources

The principal quantitative dataset is the RBI bank-group fraud table covering 2022-23, 2023-24 and 2024-25. The table contains the number of reported fraud cases and the amount involved for public sector banks, private sector banks, foreign banks, financial institutions, small finance banks, payments banks and local area banks. The paper focuses its 3D visualisation on five groups with clearly reported values across the three years. RBI notes that the figures are subject to revisions, that frauds reported in a year may have occurred earlier, and that the amount involved is not the same as the final loss incurred [2].

6.2 Analytical Approach

1. Descriptive comparison of fraud cases across bank groups and years.
2. Comparison of reported fraud amount in ₹ crore.
3. Interpretation of card/internet fraud as a digital-security indicator without treating it as a complete measure of cybercrime.
4. Thematic analysis of AI applications in fraud detection, threat detection, identity security, AML/KYC support and incident response.
5. Risk analysis covering data quality, adversarial attacks, explainability, privacy, model drift, bias, third-party dependence and human oversight.

6.3 Limitation

The RBI fraud dataset is a banking-fraud dataset rather than an AI-performance dataset. It cannot by itself show that AI caused fraud to rise or fall. The paper therefore avoids claiming a direct causal effect. The 3D graph is a visual representation of reported fraud cases, while



the discussion of AI describes how AI can be used to address patterns associated with those risks. This distinction is essential for responsible research.

7. RBI Secondary Data: Reported Fraud Cases and Amounts

Table 2 presents selected RBI bank-group data for three reporting years. The table is deliberately presented as a security-context dataset rather than as evidence of AI success or failure. The number of cases indicates reporting volume, whereas the amount involved indicates the financial magnitude recorded by the reporting system.

Table.2. RBI reported fraud cases and amount involved by bank group

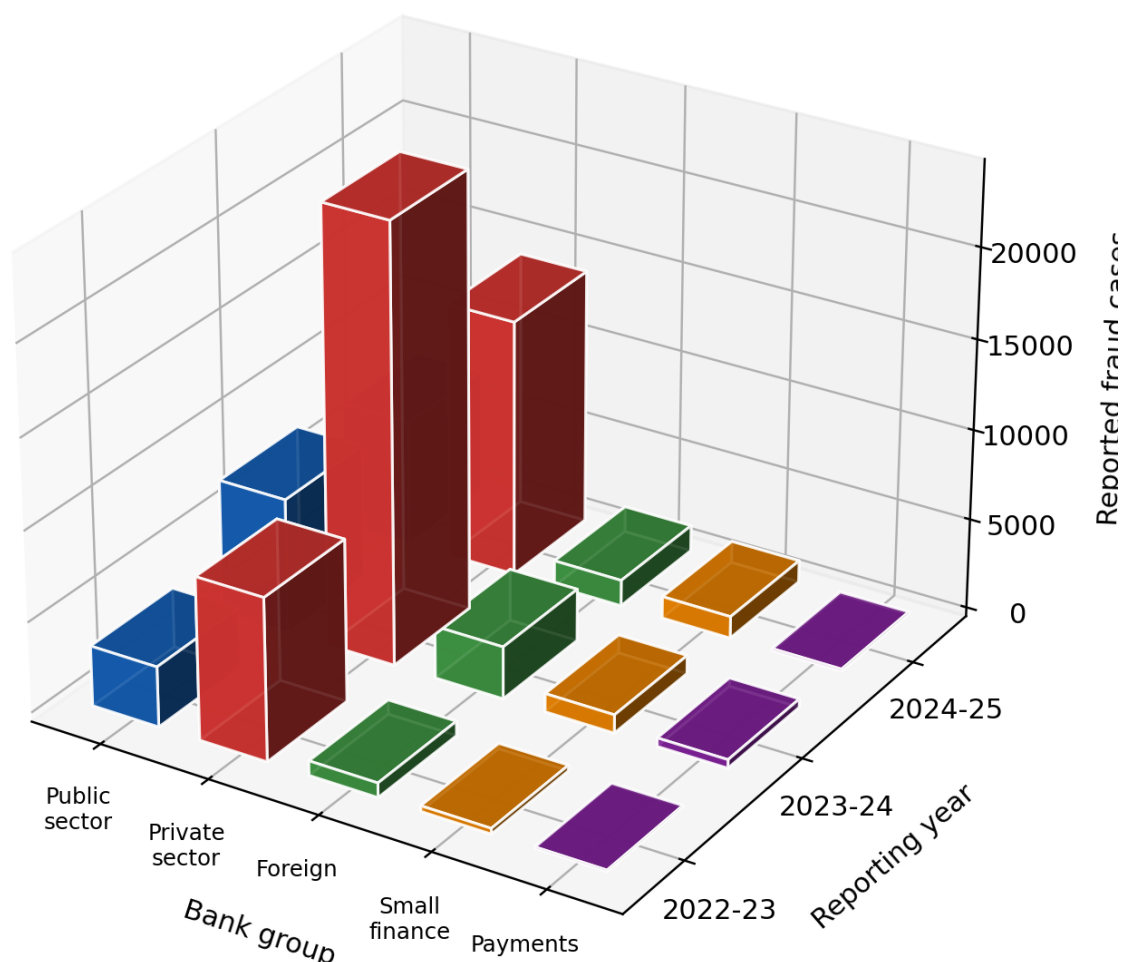
Bank group	2022-23 Cases	2022-23 ₹ crore	2023-24 Cases	2023-24 ₹ crore	2024-25 Cases	2024-25 ₹ crore
Public Sector Banks	3331	12557	7460	9254	6935	25667
Private Sector Banks	8971	5206	24207	2722	14233	10088
Foreign Banks	804	292	2899	154	1448	181
Small Finance Banks	311	31	1019	64	1217	58
Payments Banks	68	7	472	35	113	6
Total (all bank groups)	13,494	18,981	36,060	12,230	23,953	36,014

Source: [2, Table VI.2]. RBI states that the figures cover frauds of ₹1 lakh and above reported during the period, may be revised, and may include frauds that occurred in earlier years. The amount involved should not be read as identical to the final loss [2].

8. Three-Dimensional Visualization of Fraud Cases



3D comparison of reported bank fraud cases in India



Source: RBI Annual Report 2024-25, Table VI.2. Reported fraud cases are not a direct measure of AI effectiveness.

Figure 3. 3D comparison of RBI reported fraud cases and amount involved by bank

The 3D graph (Fig.3) makes one pattern easy to see: reported fraud cases were highest in the private sector bank group in each of the three years shown, while public sector banks also recorded a large number of cases. The 2023-24 jump in private sector cases is especially visible. In 2024-25, private sector cases declined from 24,207 to 14,233, while public sector cases declined from 7,460 to 6,935. Small finance bank cases increased from 1,019 to 1,217, whereas payments-bank cases fell from 472 to 113. These changes describe reporting patterns; they should not be interpreted as a direct measure of the quality of a bank's cybersecurity or its use of AI.



The second dimension of the RBI data is financial amount. In 2024-25, public sector banks reported ₹25,667 crore involved in frauds, compared with ₹10,088 crore for private sector banks. This difference illustrates why case counts alone are not enough for cyber-risk management. A bank may have many small incidents or fewer high-value incidents. AI systems should therefore combine transaction frequency, value, customer context, network relationships and other risk signals rather than relying on a single indicator.

9. Applications of AI in Banking Cybersecurity

AI Applications and Risks in Banking Cybersecurity



Fig.4 Applications of AI in Banking Cybersecurity

9.1 Real-Time Transaction Anomaly Detection

Fig.4 explains, AI can examine transactions as they occur and compare them with a customer's normal behavioural pattern. Useful features may include transaction amount, frequency, time of day, device identity, location, beneficiary history, merchant category and recent account changes. An anomaly score can then be calculated. A high score does not prove fraud; it signals that additional checks may be appropriate. This approach can reduce dependence on fixed rules and help analysts focus on unusual combinations of signals.



9.2 Behavioural Biometrics and Account Takeover Detection

A genuine customer often has a recognisable interaction pattern. Typing rhythm, navigation sequence, device behaviour, session duration and login characteristics can be used as supplementary signals. If an account suddenly behaves in a very different way, the system can increase risk scoring. Such techniques must be designed carefully because behavioural data can be sensitive and legitimate users may behave differently when travelling, changing devices or facing accessibility needs.

9.3 Mule Account Identification

Mule accounts can receive or transfer money on behalf of fraud networks. AI can search for network patterns that are difficult to detect through one-account analysis. Many accounts may receive small amounts and rapidly transfer funds through common beneficiaries. Graph algorithms can identify clusters, repeated paths and unusual relationships. RBI reported that its Innovation Hub developed MuleHunter, an AI tool for detecting mule accounts, showing that this use case has moved from theory toward practical supervisory and banking applications [19].

9.4 Phishing and Social-Engineering Detection

Natural-language models can classify suspicious emails, messages and web content. A security system can examine sender patterns, domain reputation, wording, links, attachments and historical behaviour. AI can also help create simulated phishing exercises for employee training. However, generative AI may also help attackers create more convincing scams. Defensive language models must therefore be protected against malicious inputs and sensitive-data leakage.

9.5 Malware and Endpoint Security

Traditional antivirus tools often rely on known signatures. Machine-learning systems can add behavioural analysis by looking at process activity, unusual network connections, file changes and execution patterns. This can help identify previously unseen or modified malware. Yet model accuracy can change as attackers alter their behaviour, making continuous testing essential.

9.6 AI for Security Operations Centres

A modern bank may receive thousands of security alerts. AI can group related events, remove duplicate alerts, rank high-risk incidents and summarise evidence for security analysts. A human analyst can then investigate the most important cases. This human-machine arrangement is often more practical than asking AI to make every security decision by itself.

9.7 AI for KYC, AML and Network Analysis

AI can help examine customer relationships, transaction networks and unusual flows. BIS research indicates that machine learning can improve pattern recognition in anti-money-laundering and KYC processes, especially when relationship data are available [12]. In India, this capability must operate within applicable privacy, banking and regulatory requirements.

10. Risks of Using AI for Cybersecurity

10.1 Data Quality



An AI model cannot compensate for poor data. Missing device identifiers, inconsistent customer records, delayed labels or incorrect fraud classifications can reduce model quality. Banks should establish data-quality checks before training or deploying a model.

10.2 Model Drift

Customer behaviour and fraud methods change. A model trained on last year's patterns may perform poorly when attackers change tactics. Model monitoring should measure false positives, false negatives, detection latency and population changes. Retraining should follow evidence rather than an arbitrary calendar.

10.3 Adversarial Attacks

Attackers may deliberately manipulate inputs so that an AI model misses suspicious behaviour. Training data may also be poisoned. Banks should conduct adversarial testing and maintain a fallback security process that does not depend on one model.

10.4 Explainability and Customer Impact

A fraud model may block a legitimate payment. If the bank cannot explain why an action occurred, customer trust may be damaged. Explainability does not require exposing every technical detail to the customer, but it does require internal evidence, audit trails and a process for human review.

10.5 Privacy

Banking AI can use highly sensitive data. Data minimisation, access control, encryption, retention limits and secure model-development environments are essential. The ISO/IEC 27001 information-security management approach provides a useful governance foundation, while AI-specific controls should be added according to the use case [22].

10.6 Third-Party and Cloud Risk

Banks may purchase AI models, cloud services, fraud platforms or security products from external providers. This creates dependency risk. A failure, compromise or data leak at a provider can affect many institutions. BIS work on digitalisation and AI stresses governance over third-party technology and operational resilience [23], [18].

11. Proposed Responsible AI Cybersecurity Framework for Indian Banks

Based on the evidence reviewed, this paper proposes an eight-layer framework. The purpose is to connect AI with ordinary security controls instead of treating AI as a separate project.

Layer 1 – Secure data foundation: Create accurate, well-labelled, access-controlled datasets for transactions, devices, identities, alerts and incidents.

Layer 2 – Rules plus machine learning: Keep established security rules and combine them with anomaly detection and supervised models.

Layer 3 – Behavioural intelligence: Build customer, device, account and beneficiary behaviour profiles while respecting privacy.

Layer 4 – Graph intelligence: Map relationships among accounts, devices, beneficiaries, merchants and transactions to identify suspicious networks.

Layer 5 – Explainable risk scoring: Provide analysts with the main signals behind each high-risk alert and preserve an audit trail.



Layer 6 – Human investigation: Use trained investigators for important decisions, appeals, unusual cases and model exceptions.

Layer 7 – AI security testing: Test for adversarial inputs, data poisoning, model drift, prompt injection and leakage before and after deployment.

Layer 8 – Governance and response: Connect AI alerts with SOC procedures, incident reporting, board oversight, independent validation and regulatory requirements.

12. Discussion

The evidence suggests that AI is becoming an important part of banking cybersecurity in India, but its role should be understood carefully. RBI's supervisory use of AI/ML, its fraud vulnerability index, the MuleHunter initiative and the proposed responsible-AI framework indicate institutional movement toward data-driven risk management. At the same time, RBI's fraud figures show that the security environment remains complex. The 2024-25 data record 23,953 reported fraud cases involving ₹36,014 crore across the bank groups in the table. These figures are not proof that banking security is deteriorating or improving in a simple way, because reporting practices, detection, classification and historical fraud discovery affect annual totals.

One important finding is the difference between case volume and financial value. In 2024-25, private sector banks accounted for 59.4% of the reported cases in the RBI bank-group table, while public sector banks accounted for 71.3% of the reported amount involved. Cybersecurity measurement should therefore not rely on a single number. AI dashboards should show incident count, financial exposure, time to detection, time to containment, false-positive rate and recovery separately.

The data also demonstrate why digital fraud and cyber risk require different analytical layers. RBI reported that card/internet frauds represented 56.5% of reported fraud cases in 2024-25, while advances accounted for 92.1% of the amount involved. Therefore, a bank security programme must protect digital channels while also connecting cybersecurity signals to broader fraud-risk management. AI can contribute to both areas, but models and controls should be matched to the risk being addressed.

A second major finding concerns human oversight. AI can identify patterns faster than a person can manually examine millions of transactions. It cannot, however, independently establish the legal meaning of every suspicious event. Human investigators understand context, customer explanations, operational exceptions and evidence quality. The practical model is a human-AI partnership in which machines perform large-scale screening and humans retain accountability for consequential decisions.

A third finding is that AI security must be recursive. If a bank protects its network with AI, the AI system itself becomes part of the attack surface. Model files, training datasets, APIs, prompts, feature stores and decision engines must therefore receive cybersecurity protection. This is consistent with recent BIS analysis that identifies model manipulation, data poisoning and prompt-related risks as emerging concerns in financial AI [12], [18].

13. Recommendations



1. Create a board-level AI and cyber-risk governance structure with clearly assigned accountability.
2. Use AI to complement, not replace, rule-based security controls and human investigation.
3. Build real-time transaction monitoring that combines behavioural, device, network and transaction signals.
4. Use graph analytics for mule-account networks and coordinated fraud patterns.
5. Establish independent validation for high-impact AI models before deployment.
6. Maintain model cards, data lineage, version control, performance reports and audit logs.
7. Test models against adversarial examples, data poisoning, evasion and distribution shift.
8. Protect sensitive banking data through minimisation, encryption, access control and controlled AI-development environments.
9. Measure false positives and false negatives because excessive blocking can harm legitimate customers.
10. Integrate AI alerts with SOC incident-response procedures and CERT-In reporting requirements.
11. Strengthen staff training so cyber analysts understand both AI outputs and AI limitations.
12. Assess third-party AI vendors for security, data use, resilience, explainability and exit arrangements.
13. Develop sector-wide information sharing on emerging fraud patterns while respecting legal and privacy constraints.
14. Evaluate AI programmes through measurable outcomes such as detection time, containment time, recovery, alert quality and customer impact.

14. Conclusion

Artificial intelligence can become a powerful component of cybersecurity in India's banking sector because banks generate large amounts of transactional, behavioural and network data. Machine learning can help identify anomalies, detect suspicious relationships, prioritise security alerts and support faster responses. RBI's own supervisory initiatives show that AI and machine learning are increasingly being used for fraud vulnerability analysis and other supervisory tasks. At the same time, the RBI fraud data demonstrate that the banking risk environment remains complex and that the number of reported cases and the financial amount involved do not always move in the same direction.

The central conclusion of this study is that AI should be understood as a capability inside a wider security system. Its effectiveness depends on good data, suitable models, explainability, secure infrastructure, human expertise, continuous testing and clear governance. AI can detect a pattern, but the bank must decide how that signal should be investigated and acted upon. Similarly, an AI system can reduce repetitive work, but it cannot remove institutional responsibility.



For Indian banks, the next stage should therefore be responsible and measurable adoption. Transaction analytics, graph intelligence, behavioural monitoring, threat detection and AI-assisted security operations can be connected to RBI and The Indian Computer Emergency Response Team -CERT-In requirements. At the same time, banks should prepare for adversarial AI, model drift, privacy concerns and third-party dependencies. A secure banking future will not come from AI alone. It will come from the combination of AI, cybersecurity engineering, human judgment, strong governance and continuous institutional learning.

References

1. Reserve Bank of India, "Cyber Security Framework in Banks," RBI/2015-16/418, DBS.CO/CSITE/BC.11/33.01.001/2015-16, 2016.
2. Reserve Bank of India, *Annual Report 2024–25: Regulation, Supervision and Financial Stability*. RBI, 2025.
3. A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1153–1176, 2016, doi: 10.1109/COMST.2015.2494502.
4. V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys*, vol. 41, no. 3, Art. no. 15, 2009, doi: 10.1145/1541880.1541882.
5. Reserve Bank of India, "RBI-Digital Payments Index for September 2024," RBI Press Release, 2025.
6. Reserve Bank of India, *Report on Currency and Finance 2023–24: Digitalisation and Payment Revolution in India*. RBI, 2024.
7. C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
8. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
9. R. Sommer and V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection," in *Proc. IEEE Symp. Security and Privacy*, 2010, pp. 305–316, doi: 10.1109/SP.2010.25.
10. L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.
11. K. Shaukat, T. M. Alam, I. A. Hameed, W. A. Khan, S. Shabbir, and S. Luo, "A review of machine learning techniques for cyber security in the last decade," *IEEE Access*, vol. 8, pp. 222620–222641, 2020.
12. I. Aldasoro, L. Gambacorta, A. Korinek, V. Shreeti, and M. Stein, "Artificial intelligence and the economy: Implications for central banks," Bank for International Settlements, 2024.
13. M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you? Explaining the predictions of any classifier," in *Proc. ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2016, pp. 1135–1144, doi: 10.1145/2939672.2939778.
14. S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, vol. 30, 2017.



15. National Institute of Standards and Technology, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. U.S. Department of Commerce, 2023, doi: 10.6028/NIST.AI.100-1.
16. N. Papernot, P. McDaniel, S. Jha, S. Fredrikson, Z. B. Celik, and A. Swami, “The limitations of deep learning in adversarial settings,” in *Proc. IEEE European Symp. Security and Privacy*, 2016, pp. 372–387, doi: 10.1109/EuroSP.2016.36.
17. N. Carlini and D. Wagner, “Towards evaluating the robustness of neural networks,” in *Proc. IEEE Symp. Security and Privacy*, 2017, pp. 39–57, doi: 10.1109/SP.2017.49.
18. J. C. Crisanto, C. B. Leuterio, J. Prenio, and J. Yong, “Regulating AI in the financial sector: Recent developments and main challenges,” Bank for International Settlements, Financial Stability Institute, 2024.
19. Reserve Bank of India, *Annual Report 2023–24: Regulation, Supervision and Financial Stability*. RBI, 2024.
20. Central Computer Emergency Response Team, “Directions under sub-section (6) of Section 70B of the Information Technology Act, 2000,” CERT-In, Ministry of Electronics and Information Technology, 2022.
21. Central Computer Emergency Response Team, “Frequently Asked Questions on Cyber Security Directions of 28.04.2022,” CERT-In, Ministry of Electronics and Information Technology, 2022.
22. International Organization for Standardization, *ISO/IEC 27001:2022 Information Security, Cybersecurity and Privacy Protection—Information Security Management Systems—Requirements*. ISO, 2022.
23. Basel Committee on Banking Supervision, *Digitalisation of Finance*. Bank for International Settlements, 2024.