World Journal of Advanced Research and Reviews

eISSN: 2581-9615 CODEN (USA): WJARAI

Cross Ref DOI: 10.30574/wjarr

Journal homepage: https://wjarr.com/

(Research Article)

Artificial Intelligence and Fraud Detection in US Commercial Banks: Opportunities and Challenges

Bridget Nnenna Chukwu [1, *] and Chidozie Ebube Ebenmelu [2]

[1] *Department of Agribusiness and Applied Economics, North Dakota State University, Fargo, ND, USA.*

[2] *Department of Agricultural and Consumer Economics, University of Illinois, Urbana-Champaign, IL, USA.*

World Journal of Advanced Research and Reviews, 2025, 27(03), 1083-1091

Publication history: Received on 10 August 2025; revised on 14 September 2025; accepted on 18 September 2025

Article DOI: https://doi.org/10.30574/wjarr.2025.27.3.3259

Abstract

Fraud detection in commercial banks in the U.S. is being transformed by artificial intelligence (AI), which identifies suspicious activity faster and more accurately, and minimizes financial losses. The conventional rule-based approaches are becoming insufficient to deal with the sophistication of the current fraudulent patterns, so AI-based methods like machine learning, deep learning, and natural language processing form a vital part of effective fraud prevention. This study evaluates the prospects and issues related to the use of AI in banking fraud detection. It examines the role of AI in the real-time detection of anomalies, in improving predictive analytics, and in improving adherence to regulatory frameworks. At the same time, it critically assesses issues such as algorithmic bias, data security, and the financial and operational implications of integrating AI systems into legacy infrastructures. By drawing on contemporary research and industry case studies, the paper contributes to a deeper understanding of how AI can be deployed responsibly to strengthen fraud prevention efforts while addressing the technical, ethical, and governance risks it presents.

Keywords: Artificial Intelligence; Fraud Detection; Commercial Banking; Machine Learning; Deep Learning

1. Introduction

Fraud detection has been a major issue of concern to the banking industry, especially in commercial banks, where large transactions are more prone to being fraudulent (Awosika, Shukla, and Prang Gono, 2023; Nguyen, Pham, and Le, 2022). Financial fraud has become a key issue for the stability of financial systems around the world, and the United States is no exception because of the emergence of cybercrime, insider fraud, and identity theft (Association of Certified Fraud Examiners [ACFE], 2022). Conventional fraud detection techniques, which are based on rule-based systems and are manually reviewed, are becoming insufficient in detecting the sophisticated, dynamic, and changing methods used by fraudsters. This deficiency has created a pressing need for new and automated systems capable of processing large and high-frequency data in real-time (Nguyen et al., 2022). Artificial intelligence (AI) is now a disruptive technology that can address this problem with machine learning algorithms, deep learning models, and natural language processing, all of which can identify fraudulent behavior with greater accuracy. The trend of increasing the use of AI in the banking industry has transformed it into a vital tool that financial institutions can utilize to minimize losses incurred due to fraud, maintain customer trust, and comply with regulatory requirements.

The development of technologies has helped to expand the range of sophistication of fraud techniques, the most active agents of which are digital banking, mobile payments, and instant payment systems, which have raised the range of areas where such techniques can be used (PwC Report, 2023). The Federal Reserve states that U.S. financial institutions are losing billions of dollars a year to scams, which can be categorized as phishing-related and account-takeover schemes, insider collusion, and synthetic identity fraud (Federal Reserve, 2023). The implications of these losses are

[*] Corresponding author: Bridget Nnenna Chukwu

far-reaching in the profitability of the bank, customer confidence, and stability of the financial sector at large. It is against this background that AI is a paradigm shift, as it allows banks to go beyond fixed rules and dynamic self-learning systems that constantly evolve to keep pace with changing fraud trends (Ghosh et al., 2021). As an example, machine learning algorithms can detect small anomalies in customer behavior by examining large volumes of transaction data, and can be used in real time to detect possible fraud with only a small number of false positives. In the same way, natural language processing may be used to scan communications in case of insider trading risks or detect malicious intent in unstructured text messages (Patil and Bansal, 2022). These abilities have seen AI not only as a defensive tool, but as a proactive tool in the management of fraud risks.

Nevertheless, AI usage in fraud detection is not free of problems. The process of introducing AI systems into the current banking software is linked to colossal expenditures in data organization, computer power, and manpower (Zhang and Wei, 2022). Moreover, the quality of AI systems is limited to the quality of the data that they are trained on, and this leads to the issue of data privacy, data quality, and algorithmic fairness. This would lead to discrimination, such as the unfair characterization of certain demographics as high-risk, and leave banks vulnerable to ethical and legal repercussions (Mehrabi et al., 2021). Moreover, the black box nature of some AI models, particularly deep learning, creates interpretability challenges, and it is difficult to justify the compliance officers and regulators why a particular transaction will be considered a fraud (Doshi-Velez and Kim, 2018). All these problems demonstrate the necessity to design transparent and comprehensible AI frameworks and to introduce governance frameworks that would lead to the decrease of fraud, accountability, and fairness.

This study is significant because it is one of the contributions to a growing number of publications that strive to understand the opportunities and traps of AI in the financial sector. This study offers a subtle insight into how AI can be utilized to enhance the fight against fraud by looking at the application of AI in real-time anomaly detection, predictive analytics, and regulatory compliance. Simultaneously, it critically assesses the threats of algorithm bias, data safety, and costs of operation, providing the perspectives of how the challenges may be addressed with the help of responsible AI use. The analysis not only applies in the case of commercial banks but also to regulators, technology providers, and policymakers who are all shaping the future of AI-based financial crime prevention. Having a middle ground that will not only harness the full capabilities of AI but will also minimize its risks will be central to the sustainability and stability of commercial banking systems in the United States in the light of the ever-evolving financial landscape (See: Basel Committee on Banking Supervision, 2023).

2. Literature Review and Theoretical perspectives on AI and fraud detection

Fraud detection is based on the theory of statistical learning, anomaly detection, and adversarial learning. The underlying assumption of these classical theories is that the task of fraud detection is essentially a classification or outlier detection problem: a model of normal behavior has to be modeled, aberrations have to be detected, trade-offs between false positive and false negative rates have to be addressed, and the model must have the ability to generalize under adversarial conditions. In the case of Akhlaq et al. (2024) in the article titled AI-Powered Fraud Detection and Prevention in Banking, LSTM networks with SHAP are used to identify the patterns of fraud, as well as to present the decision process in a way that is understandable, which is aligned with the theories of explainable AI (XAI) that view explainable interpretability as a component of trust and governance. Other theoretical developments encompass the use of Generative Adversarial Networks (GANs) to create synthetic data to work with imbalanced fraud samples, as in the hybrid architecture of a recent study that combines GANs with recurrent neural networks (RNNs) to better capture the changing pattern of fraud and increase detection abilities in cases of data scarcity (GAN + RNN hybrid, 2024). The theoretical basis of these works focuses on the problem of stationarity (patterns of fraud change with time), the problem of disproportionate representation in a class, adversarial adaptation, the complexity of feature representations (sequence, graph, behavioral metadata), and regulatory/ethical limits (privacy, transparency, fairness).

There is empirical evidence of the great potential of these AI solutions, as well as concerns regarding the limitations. In the article Unmasking Banking Fraud: Unleashing the Power of Machine Learning and Explainable AI (XAI) on Imbalanced Data (Nobel et al., 2024), four algorithms (Support Vector Machine, XGBoost, Decision Tree, and Logistic Regression) were evaluated using SMOTE in order to resolve the issue of class imbalance; XGBoost got the highest result, which is approximately 99.88 percent, and with the help of SHAP and LIME, one could interpret the main features of the model. Empirically, this demonstrates that imbalance treatment in combination with XAI has very high detection rates but reduces the opacities. The other empirical study, A Hybrid Deep Learning Approach with Generative Adversarial Network to detect credit card fraud (Technologies, 2024), combines GAN to generate real fraud samples with RNN-based discriminators: this type of hybrid model is more effective than the baseline ML models in high imbalance (precision/recall). Another empirical contribution includes A Comparison Study of Credit Card Fraud Detection: Supervised versus Unsupervised (Niu, Wang, Yang, 2019), which concludes that supervised models (Random Forest, Extreme Gradient Boosting) have slightly better results compared to unsupervised models, but that the latter are still promising- particularly in cases where there are few labels. These works demonstrate in practice that ML / DL solutions can substantially decrease missed fraud. Still, the results vary greatly based on the balance of the datasets, feature engineering, and the conditions of implementation.

In addition to the rates of detection, the empirical work also addresses operational, governance, and risk issues. The research article Transparency and Privacy: The Role of Explainable AI and Federated Learning in Financial Fraud Detection (Awosika, Shukla, and Pranggono, 2023) presents federated learning to enable more than one institution to enjoy improvements in the shared model without having to share any information on the transactions they have conducted- and pairs it with explainable AI to enable interpretability. They find that such models can sustain high performance even when privacy is limited. The other empirical example is the article by Wedge, Kanter, Moral Rubio, Iglesias Perez, and Veeramachaneni (2017, which used automated behavioral feature synthesis (Deep Feature Synthesis) and a Random Forest classifier with the data of a large multinational bank and obtained a similar reduction in the number of false positives on unseen transactions (about 54%). The results of the empirical analysis are indicative of the value of reducing false positives (customer friction, cost of investigation) nearly as much as increasing detection. In addition, the case study, "Fraud Detection on the Swipe to a Major US Bank" (SingleStore case study), illustrates the transition from legacy batch detection to real-time swipe detection with real-time scoring and streaming infrastructure, indicating that response latency can be significantly minimized in US commercial banking settings.

Both the theoretical and empirical approaches converge on a number of challenges. First, adversarial/dynamic behavior in fraud: fraudsters evolve in response to detection; non-adaptive models have diminishing performance. The theoretical presentation of this is in adversarial attack studies (such as… Foe for Fraud: Transferable Adversarial Attacks in Credit Card Fraud Detection, 2025) that demonstrate that credit card fraud detection models are sensitive to small perturbations, both white-box and black-box, which can decrease their performance. Second, quality of data and sampling: lag in labelling fraud, dataset bias (investigated vs all cases), class imbalance, privacy (e.g., sharing data across banks). Empirical studies, such as the federated learning + XAI one, demonstrate that privacy limitations can be overcome, although usually at the expense of model complexity or latency. Third, regulatory, interpretability, and trust concerns: models should be interpretable to meet regulatory requirements and internal audit, and this forces institutions to use XAI procedures (SHAP, LIME, surrogate models), as is the case in most empirical papers. Lastly, deployment issues: real-time requirements (latency, streaming data), infrastructure maturity (real-time scoring, data pipelines), compatibility with human investigator processes, cost of false positives, versus the complexity/maintainability of models, and governance/model risk management.

3. Overview of Fraud Trends in U.S. Commercial Banking

The U.S. commercial banking fraud keeps on developing at an alarming rate, with check fraud, card fraud, real-time payments fraud, and identity-related fraud being some of the most threatening to the sector. The fraud of checks specifically has been on the rise: Suspicious Activity Reports (SARS) of check fraud at deposit institutions increased more than three times between 2018 and 2022, with it becoming one of the most rapidly increasing types of fraud in the industry (SandP Global Report, 2023). According to many banks, check fraud constitutes over a third of the total fraud (with mortgage fraud excluded) now, which represents an outrageous change in the fraud environment (SandP Global Report, 2023). Credit card fraud is also problematic, as more than half of surveyed banking and payments executives mention it among their most significant fraud problems, right after check fraud (American Banker, 2024). Zelle and FedNow are real-time payment systems that fraudsters have targeted, as the systems allow authorized push payment scams and social engineering attacks to occur because of speedy transfers (American Banker, 2024).

Identity-facilitated fraud has also emerged as a core issue, including synthetic identity fraud, business email compromise (BEC), phishing, and AI-based impersonation to a growing degree. Synthetic identity fraud, as a type of fraud that implies combining authentic and fictitious data to establish false identities, has been steadily increasing and is paradoxically difficult to restructure because of loopholes in identity verification at the time of opening an account (BankersHub, 2025). BEC is one of the fraud cases that has been reported to be the most expensive to date, as criminals persuade businesses to send money to fake accounts (U.S. Bank, 2024). The generative AI is also being applied by fraudsters to create deepfakes and synthetic voices to pose as bank employees or clients, and this presents additional threats to authentication and verification procedures (Business Insider, 2025). Identity theft, therefore, is a major pathway to facilitate most forms of banking fraud, and its development is an indication that requires more sophisticated monitoring and detection systems (Arthur State Bank, 2025). The increasing rate and price of fraud incidents is apparent in the banking industry of the U.S., and especially to business customers. According to surveys, most companies have at least one attempted payment fraud per year, with checks being the most prevalent means of attack, followed by ACH, wires, and cards (SouthState Bank, 2024). Financial effects are also noticeable, and banks have recorded hundreds of millions of dollars in losses due to check fraud on an annual basis (SandP Global, 2023). SMEs are especially susceptible, and they are becoming increasingly worried about unauthorized transfers, identity theft, and phishing attacks (AP News, 2024). The following table gives a glimpse of some of the important statistics that would help summarize the prevailing situation of fraud in U.S. commercial banking.

Table 1 Major Financial Fraud Pattern in the U.S Banks

Fraud Trend / Metric	Statistic / Finding	Source
Increase in check fraud SARs	+201.2% (2018–2022)	San DP Global, 2023
Share of banks citing card fraud as top concern	54%	American Banker, 2024
Share citing check fraud as top concern	50%	American Banker, 2024
Businesses targeted by payment fraud attempts	96% (2023)	South State Bank, 2024
Most common channel of fraud	Checks (65% of cases)	South State Bank, 2024
Other payment fraud channels	ACH debits (33%), wires (24%), cards (20%), ACH credits (19%)	South State Bank, 2024
Businesses worried about unauthorized transfers	44%	AP News, 2024
Businesses worried about identity theft	37%	AP News, 2024

Author's Online Review 2025

The table has raised the complexity and magnitude of fraud in commercial banking in the United States, with check fraud taking the top position as a threat. At 201.2 percent, the 2018-2022 growth in SARS of check fraud shows that it has reappeared despite being transferred to digital payments. Other areas of concern include fraud and real-time payment scam fraud, since more than 50 percent of banking leaders are placing them as their priority. The statistics also reveal that the attempts of fraud are almost omnipresent, with 96 percent of the businesses targeted, and the check being the most commonly used mode of attack, with 65 percent of the fraud transactions. The debiting of ACH and wire transfer is also rather risky, and it shows that the fraudsters use different payment rails. The statistics on business sentiment show that unauthorized transfers and identity theft are becoming increasingly important to people, which is also in line with the tendencies in synthetic identity fraud and phishing attacks. The statistics indicate that the fraud is becoming more and more common, and more concentrated and demanding more intensive detection.

4. AI Techniques for Fraud Detection

Artificial Intelligence (AI) has taken center stage in the process of fraud detection in the finance, insurance, and e-commerce industries. One of the most common techniques is traditional Machine Learning (ML), especially supervised learning algorithms, which are trained on past occurrences of fraud and non-fraud to categorize future transactions or behaviors. Random Forest, Support Vector Machines (SVM), logistic regression, decision trees, and ensemble models like XGBoost have proved to be very effective in detecting credit card fraud and financial statement fraud cases when quality labelled data is available (Hernandez-Aros et al., 2024). The advantage of these models is that they manage the imbalance in the classes, feature selection, and normalization, as the fraudulent transactions usually make up a small portion of the entire data, and they can be overlooked easily if the imbalance is ignored (Hafez et al., 2025). The importance of feature engineering is that carefully preprocessing the data is necessary since the ML models distinguish subtle anomalies in transactional data and normal variation (Hernandez-Aros et al., 2024). Studies also note that semi-supervised and unsupervised machine learning algorithms, including clustering, isolation forests, and one-class SVMs, can be useful when the rates of fraudulent labels are low because they do not need balanced training data to identify anomalies (Hafez et al., 2025).

Although ML technologies are effective with structured numerical data, Natural Language Processing (NLP) is being applied to identify fraud due to unstructured text, including corporate financial reports, emails, regulatory filings, and customer complaints. The fraud can be detected in narrative disclosures, the management discussion section, or in minor language discrepancies. NLP can be used to analyze large text collections and find lexical, syntactic, and semantic peculiarities that could reveal the fraudulent motive through the use of suspicious patterns of language, sentiment change, and uncharacteristic word usage (Li, 2023). State-of-the-art NLP models now train deep neural networks, such as embedding layers, recurrent neural networks (RNNs), gated recurrent units (GRUs), and long-short-term memory (LSTM) networks, to learn contextual dependencies between sentences and paragraphs (Li, 2023).

This is essential in fraud detection because most fraudulent disclosures consist of context-dependent language cues that cannot be detected by mere keyword matching. NLP techniques allow regulatory bodies and auditors to prioritize high-risk documents to investigate them further, which is crucial for minimizing manual work by automating the classification of huge volumes of reports (Li, 2023).

Deep Learning (DL) has more features, as it learns the complex representations of data and provides the capability to model the sequential or temporal patterns. Deep neural networks (DNNs), convolutional neural networks (CNNs), RNNs, autoencoders, and transformer-based models, among others, have been used to detect fraud with some level of success (Branco et al., 2020). An example is that autoencoders are trained to reconstruct normal data and detect anomalies by high reconstruction error, which is especially useful in situations in which labelled fraud data is in short supply (Chen et al., 2023). The models that are based on RNN are applied to model transaction sequences and identify the abnormal temporal patterns in the user behavior, whereas transformer-based models are better in working with large and complex sets of features because of the attention mechanisms (Branco et al., 2020). Hybrid models that integrate gradient boosting decision trees and deep neural networks have been promising, as they combine interpretability with the representational capabilities of DL and would overcome one of the most important criticisms of black-box models (Chen et al., 2023). But the DL methods also have their problems, such as the computational cost of the training, the possibility of overfitting, and the need to provide explanations to meet regulatory compliance (Hernandez-Aros et al., 2024). Empirical evidence suggests that although deep models often outperform simpler ML algorithms on large and complex datasets, their performance can degrade if not continuously updated to detect newly emerging fraud patterns (Hafez et al., 2025).

5. Opportunities and Challenges of Artificial Intelligence and Fraud Detection

Artificial intelligence (AI) driven real-time detection is changing the way organizations react to threats and anomalies by creating a major drop in the time between when an event happens and when it is detected. AI systems can analyze streams of transactional or behavioral information within milliseconds in other domains, such as finance, insurance, cybersecurity, and e-commerce. An example is an anomaly detection and an ensemble model-based fraud detection framework that was nearly 99.7 percent accurate and minimized tens of millions of dollars of losses annually by detecting suspicious transactions prior to their occurrence (Hopsworks, 2025). This is the capability to detect and deal with issues as they arise and not when they have been spread out, and this has enabled one to act promptly and reduce the damages caused by fraud, intrusion, or misconduct. Real-time identification also aids in augmenting situational awareness since the organizations will be able to monitor the fluctuating threats or trends that would have been too sensitive or fast to be detected by human operators. In general, the real-time detection increases the speed and accuracy of response, which is translated into organizational resilience (Fang et al., 2025).

The real-time detection is closely related to risk mitigation but is broader in the scope of possible strategies that may be applied with the help of AI. In case a warning is picked up early enough, then remedial measures can be employed to avoid the adverse impacts. AI models can forecast the risk trajectories, enabling decision-makers to understand which vulnerabilities are to be prioritized and which resources are to be focused. As an example, predictive analytics applied to the banking industry can help a bank predict potential risks linked to default, trends, or failures in processes and reduce exposure and improve compliance (Visure Solutions, 2024). Moreover, the real-time detection feeds continuous monitors, i.e., mitigation is not predefined, but adaptive, i.e., the AI improves risk scoring with time and changes the thresholds or models, i.e., mitigation is adaptive (Allana et al., 2025). This ability facilitates dynamism in risk management that is especially applicable in situations that are unpredictable or where the opponent is quick to change his tactics.

The cost-effectiveness becomes one of the strongest benefits of integrating real-time detection and risk mitigation with the help of AI. Time-consuming and repetitive tasks, including the review of flagged transactions or other routine compliance checks, can be automated to save a lot of labor (Hopsworks, 2025). It saves the organization not only on human expenditure, but also on losses through undetected fraud, penalties due to regulatory non-compliance, downtime, or reputation losses. One of the documented examples in the banking industry revealed that AI significantly decreased the number of manual reviews and false positives, allowing employees more time to work on valuable activities (Lunartech, 2025). Moreover, proper risk forecasting may minimize capital held against risk, optimize resource use, and reduce insurance or security expenses. A recent pilot program involving AI forecasting tools and foreign exchange risk hedging saved more than 30% on the cost of hedging (Reuters, 2025). All these efficiencies combine to give defensive as well as competitive advantages, which allow firms to reinvest the savings in innovation and expansion.

One of the most significant challenges in the deployment of AI systems on a large scale for detecting risks and mitigating them in real time is data privacy. The training and inference of AIs can also demand substantial amounts of personal or sensitive data, which can lead to unauthorized access, data leakage, re-identification of individuals, or data misuse (Lee et al., 2023). Recent research has demonstrated that AI not only reproduces familiar privacy challenges but also creates other issues, including membership inference attacks, data poisoning, and deepfake generation (Fang et al., 2025).

The regulations are legal limits on the kind of data that may be gathered, kept, or exchanged, and how consent to any of these activities is handled, such as the General Data Protection Regulation (GDPR) in Europe and the Health Insurance Portability and Accountability Act (HIPAA) in the U.S. (Allana et al., 2025). The organization should develop privacy-sensitive data pipelines, take into account anonymization, or differential privacy, and re-identification risk, which is a challenging task when real-time performance and high data granularity are demanded.

The other significant challenge is algorithmic bias. Still, despite big data and sophisticated algorithms, AI systems have the potential to reproduce or even enhance inequity in training data or the deployment context (Barocas et al., 2021). Bias may be a result of unrepresentative sampling, measurement error, or mislabeling datasets, but may also be introduced during feature selection or thresholding decisions (Fang et al., 2025). In practice, such bias can lead to unfair treatment, e.g. in welfare fraud detection systems, the algorithm has found hundreds of thousands of claims to be fraudulent, many of which ended up being legitimate, and vulnerable population groups were overrepresented (The Guardian, 2024). Bias is hard to quantify and remedy, as in some situations, it might require access to sensitive demographic information, which in its turn, can be restricted by privacy regulations, presenting a dilemma between justice and rule-breaking (Allana et al., 2025).

The other obstacle is model interpretability, especially in high-stakes decisions where AI systems are used. Complex algorithms such as deep neural networks can be treated as black boxes, and there is little transparency on how they make their decisions (IBM, 2024). Stakeholders, including regulators and those impacted by a decision, might demand reasons as to why a decision has been taken, e.g., why a transaction has been blocked or a claim has been rejected. The lack of interpretability makes the errors more difficult to debug, and the accountability is undermined (Wharton AI Risk Governance, 2024). Such a transparency deficiency may compromise trust to the extent of becoming regulatory non-compliant. Although interpretability can be enhanced with SHAP values, LIME, and counterfactual explanations, these approaches typically have trade-offs with model performance; thus, it is challenging to balance the accuracy and explainability of mission-critical systems (Allana et al., 2025).

6. Future Directions and Recommendations

Going forward, the future of AI in commercial bank fraud detection in the U.S. will be determined by increased integration of new technologies, collaboration of regulatory bodies, and ethical AI practices. The implementation of explainable AI (XAI) is one of the opportunities that may help to make machine learning models more transparent and explainable to compliance officers and regulators (Doshi-Velez and Kim, 2018). This will assist in minimizing the black box issue and also enhance trust and accountability in the fraud detection mechanism. In addition, the development of federated learning could also enable banks to collaborate and share information on fraud without revealing the personal data of their clients and increase the overall security of the industry against fraudsters (Nguyen et al., 2022). Live analytics and adaptive systems will be relevant in staying updated with fraud schemes that are increasingly becoming sophisticated. Financial institutions should also look into hybrid systems in which machine learning is utilized with traditional rule-based approaches in search of a compromise between precision and interpretability.

On the operational level, commercial banks should pay attention to the upskilling and AI literacy of their fraud analysts and compliance teams. Interpretation of alerts, false positives reduction, and continual enhancement of fraud detection models will be hugely applied to the collaboration between humans and AI (Ghosh et al., 2021). Also, data quality management should be made a strategic priority, in which data sets to be used in training should be representative, unbiased, and within the laws of privacy. With efficient model monitoring and auditing systems, the banks will find themselves in a position to detect model drift and re-train systems with a change in fraud patterns. Banks should also establish AI ethics committees/governance boards that will be used to examine potential risks of discrimination, privacy loss, and unwanted financial exclusion. This will help make sure that the use of AI does not contradict the regulatory requirements, such as the Basel Committee on Banking Supervision (2023) requirements, and prevent the loss of reputation by the institutions.

Finally, there will be a need to cooperate with regulators, banks, and technology providers to develop a unified system of AI-based fraud detection. Policymakers must think about the creation of industry-wide standards of AI model validation, explainability, and data governance to provide fairness and interoperability. Probably, through privacy-preserving mechanisms, the encouragement of cross-industry data sharing initiatives would also improve the early-warning systems and strengthen the fraud detection on a systemic scale (PwC Report, 2023). Future studies are supposed to create lightweight and affordable AI models that can be implemented even by small- and mid-sized banks and decrease the technology adoption gap in the industry. The prospect of AI maximizing the potential to secure the systems and ensure the trust of the population can be completely used by the U.S. commercial banks by synchronizing the technological innovation with good governance and cooperation.

7. Conclusion

This paper has demonstrated that artificial intelligence (AI) offers potent potential to enhance fraud detection in U.S. commercial banks, as it can be used to detect anomalies in real-time, predictive analytics, and adaptive risk management. Although machine learning, deep learning, and natural language processing represent the most efficient AI technologies, their combination brings up some serious questions, such as the problem of data privacy, bias in algorithms, the cost of their implementation, and the problem of interpretability. The banks should strive to ensure that technological innovation and governance, transparency, and ethical standards are balanced to maximize the benefits. The future of fraud prevention lies in the joint structures that entail financial institutions, regulators, and technology providers to create just, dependable, and viable AI systems.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Allana, S., Kankanhalli, M., and Dara, R. (2025). Privacy Risks and Preservation Methods in Explainable Artificial Intelligence: A Scoping Review. arXiv preprint arXiv:2505.02828. Retrieved from https://arxiv.org/abs/2505.02828

[2] American Banker. (2024). Card and check schemes dominate bankers' fraud woes: Report. Retrieved from https://www.americanbanker.com/news/card-and-check-schemes-dominate-bankers-fraud-woes-report

[3] AP News. (2024). Business owners increasingly worry about payment fraud, survey finds. Retrieved from https://apnews.com/article/ff4f397e4be076216a28f1a75affab7a

[4] Arthur State Bank. (2025). The latest banking fraud trends in 2025: What you need to know. Retrieved from https://www.arthurstatebank.com/blog/the-latest-banking-fraud-trends-in-2025-what-you-need-to-know

[5] Association of Certified Fraud Examiners. (2022). Report to the nations: Global study on occupational fraud and abuse. ACFE.

[6] Awosika, T., Shukla, R. M., and Pranggono, B. (2023). Transparency and privacy: The role of explainable AI and federated learning in financial fraud detection. arXiv preprint arXiv:2312.13334. https://arxiv.org/abs/2312.13334

[7] BankersHub. (2025). Top 5 banking fraud trends in 2025 and how to prevent them with modern employee training. Retrieved from https://www.bankershub.com/blogs/blog/top-5-banking-fraud-trends-in-2025-and-how-to-prevent-them-with-modern-employee-training

[8] Barocas, S., Hardt, M., and Narayanan, A. (2021). Fairness and Machine Learning: Limitations and Opportunities. Patterns, 2(5), 100234. Retrieved from https://www.sciencedirect.com/science/article/pii/S2666389921000611

[9] Basel Committee on Banking Supervision. (2023). Sound practices: The use of artificial intelligence and machine learning in banking. Bank for International Settlements.

[10] Branco, B., Abreu, P., Gomes, A. S., Almeida, M. S. C., Ascensão, J. T., and Bizarro, P. (2020). Interleaved sequence RNNs for fraud detection. arXiv preprint arXiv:2002.05988. https://arxiv.org/abs/2002.05988

[11] Business Insider. (2025). The clever new scam your bank can't stop (deepfake fraud). Retrieved from https://www.businessinsider.com/bank-account-scam-deepfakes-ai-voice-generator-crime-fraud-2025-5

[12] Chen, Y., Xu, Z., Huang, T., and Zhang, W. (2023). Efficient fraud detection using deep boosting decision trees. arXiv preprint arXiv:2302.05918. https://arxiv.org/abs/2302.05918

[13] Doshi-Velez, F., and Kim, B. (2018). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608.

[14] Fang, X., Li, J., Mulchandani, V., and Kim, J.-E. (2025). Trustworthy AI on Safety, Bias, and Privacy: A Survey. arXiv preprint arXiv:2502.10450. Retrieved from https://arxiv.org/abs/2502.10450

[15] Federal Reserve. (2023). Fraud and risk management report 2023. Board of Governors of the Federal Reserve System.

[16] Ghosh, S., Dey, N., and Nath, S. (2021). Machine learning in banking risk management: Applications and challenges. Journal of Risk and Financial Management, 14(3), 112.

[17] Hafez, I. Y., Alam, M., and Idris, M. (2025). A systematic review of AI-enhanced techniques in credit card fraud detection. Journal of Big Data, 12(1), 1-23. https://journalofbigdata.springeropen.com/articles/10.1186/s40537-024-01048-8

[18] Hernandez-Aros, L., Aros, A., and Caballero, A. (2024). Financial fraud detection through the application of machine learning: A systematic review. Humanities and Social Sciences Communications, 11(1), 1-15. https://www.nature.com/articles/s41599-024-03606-0

[19] Hopsworks. (2025). Real-Time Fraud Detection Use Case. Retrieved from https://www.hopsworks.ai/use-case/realtime-fraud-detection

[20] Hybrid Deep Learning Approach with Generative Adversarial Network for Credit Card Fraud Detection. (2024). Technologies, 12(10), 186. https://doi.org/10.3390/technologies12100186

[21] IBM. (2024). What Is AI Interpretability? Retrieved from https://www.ibm.com/think/topics/interpretability

[22] Lee, H.-P., Yang, Y.-J., von Davier, T., and Das, S. (2023). A Taxonomy of AI Privacy Risks. arXiv preprint arXiv:2310.07879. Retrieved from https://arxiv.org/abs/2310.07879

[23] Li, Q. (2023). Textual data mining for financial fraud detection: A deep learning approach. arXiv preprint arXiv:2308.03800. https://arxiv.org/abs/2308.03800

[24] Lunartech. (2025). Fraud Detection with AI Case Study. Retrieved from https://technologies.lunartech.ai/case-studies/fraud-detection-with-ai

[25] Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., and Galstyan, A. (2021). A survey on bias and fairness in machine learning. ACM Computing Surveys, 54(6), 1–35.

[26] Nguyen, T., Pham, T., and Le, Q. (2022). Artificial intelligence applications in financial fraud detection: A review. Expert Systems with Applications, 194, 116537.

[27] Niu, X., Wang, L., and Yang, X. (2019). A comparison study of credit card fraud detection: Supervised versus unsupervised. arXiv preprint arXiv:1904.10604. https://arxiv.org/abs/1904.10604

[28] Nobel, S. M. N., Sultana, S., Singha, S. P., Chaki, S., Mahi, M. J. N., Jan, T., Barros, A., and Whaiduzzaman, M. (2024). Unmasking banking fraud: Unleashing the power of machine learning and explainable AI (XAI) on imbalanced data. Information, 15(6), 298. https://doi.org/10.3390/info15060298

[29] Patil, M., and Bansal, A. (2022). Natural language processing for financial fraud detection: A survey. Procedia Computer Science, 209, 1125–1133.

[30] PwC Report. (2023). Global economic crime and fraud survey 2023. PricewaterhouseCoopers.

[31] Reuters. (2025, July 18). Citi, Ant International pilot AI-powered FX tool to cut hedging costs. Retrieved from https://www.reuters.com/business/finance/citi-ant-international-pilot-ai-powered-fx-tool-clients-help-cut-hedging-costs-2025-07-18

[32] SandP Global. (2023). US banks increasingly battle costly check fraud schemes. Retrieved from https://www.spglobal.com/market-intelligence/en/news-insights/articles/2023/12/us-banks-increasingly-battle-costly-check-fraud-schemes-79416205

[33] SingleStore. (n.d.). Fraud detection “on the swipe” for a major US bank [Case study]. Retrieved September 17, 2025, from https://www.singlestore.com/blog/case-study-fraud-detection-on-the-swipe/

[34] SouthState Bank. (2024). Payment fraud by the numbers. Retrieved from https://www.southstatebank.com/commercial/stories-and-insights/payment-fraud-by-the-numbers

[35] The Guardian. (2024, June 23). DWP algorithm wrongly flags 200,000 people as possible fraud cases. Retrieved from https://www.theguardian.com/society/article/2024/jun/23/dwp-algorithm-wrongly-flags-200000-people-possible-fraud-error

[36] U.S. Bank Report. (2024). Fight the battle against payments fraud. Retrieved from https://www.usbank.com/financialiq/improve-your-operations/minimize-risk/fight-the-battle-against-payments-fraud.html

[37] Visure Solutions. (2024). AI and Machine Learning for Risk Management. Retrieved from https://visuresolutions.com/blog/ai-and-machine-learning-for-risk-management

[38] Wedge, R., Kanter, J. M., Moral Rubio, S., Iglesias Perez, S., and Veeramachaneni, K. (2017). Solving the “false positives” problem in fraud prediction. arXiv preprint arXiv:1710.07709. https://arxiv.org/abs/1710.07709

[39] Wharton AI Risk Governance. (2024). Artificial Intelligence Risk Governance. Retrieved from https://ai.wharton.upenn.edu/white-paper/artificial-intelligence-risk-governance

[40] Zhang, L., and Wei, Z. (2022). Challenges of AI integration in financial services. Journal of Banking and Finance Technology, 6(2), 205–220.

[41] Zhang, Y., and colleagues. (2025). Foe for fraud: Transferable adversarial attacks in credit card fraud detection. arXiv preprint arXiv:2508.14699. https://arxiv.org/abs/2508.14699