Revista científica Sociedad & Tecnología

ISSN: 2773-7349

fecha de presentación: 28/07/2025, fecha de aceptación: 30/08/2025, fecha de publicación: 01/10/2025

Rahmani Siham
E-mail: siham.rahmani@univ-relizane.dz
Orcid: https://orcid.org/0009-0005-1108-2317

University of Relizane, Algeria **-** Laboratory of Financial Market Management by the Application of Mathematics and Computer Science (GMFAMI) Faculty of Economic, Commercial and Management Sciences.

Siham, R. (2025). AI's dual-use paradox in financial fraud: navigating offense and defense in the digital era. *Revista Sociedad & Tecnología*, 8(4), 673-692. DOI: https://doi.org/10.51247/st.v8i4.666.

==== o ====

AI's dual-use paradox in financial fraud: navigating offense and defense in the digital era

ABSTRACT

This study addresses the dual-use paradox of artificial intelligence (AI) in digital financial fraud, illustrating how the same technologies can both strengthen defensive systems and enable increasingly adaptive attacks. The research adopts a qualitative–conceptual methodology, aimed at developing a theoretical framework that captures this paradox and highlights the AI asymmetry gap between offensive and defensive capabilities. The evidence indicates that this gap is widening due to reactive detection models, data-access constraints, and fragmented intelligence sharing. Although applied solutions (e.g., Mastercard Decision Intelligence Pro; SymphonyAI Sensa Copilot) report improvements in detection accuracy and investigative efficiency, their effectiveness appears constrained by regulatory, institutional, and data limitations. The study proposes a framework that integrates hybrid AI architectures, cross-institutional intelligence networks, adversarial-AI research, and AI-specific regulatory measures.

Keywords: Artificial Intelligence (AI); Dual-use Paradox; Financial Fraud; Fraud Detection; Offensive AI; Defensive AI

===∘===

La IA y la paradoja de su uso dual en el fraude financiero: navegando entre la ofensiva y la defensa en la era digital

RESUMEN

Este estudio aborda la paradoja de uso dual de la inteligencia artificial (IA) en el fraude financiero digital, mostrando cómo las mismas tecnologías pueden tanto fortalecer los sistemas defensivos como habilitar ataques cada vez más adaptativos. La investigación adopta una metodología cualitativa–conceptual, orientada al desarrollo de un marco teórico que capture esta paradoja y resalte la brecha de asimetría de la IA entre las capacidades ofensivas y defensivas. La evidencia indica que esta brecha se está ampliando debido a modelos de detección reactivos, limitaciones en el acceso a los datos y una fragmentación en el intercambio de inteligencia. Aunque las soluciones aplicadas (p. ej., Mastercard Decision Intelligence Pro; SymphonyAI Sensa Copilot) reportan mejoras en la precisión de detección y en la eficiencia investigativa, su efectividad parece estar restringida por limitaciones regulatorias, institucionales y de datos. El estudio propone un marco que integre arquitecturas híbridas de

IA, redes de inteligencia interinstitucionales, investigación en IA adversaria y medidas regulatorias específicas para la IA.

Palabras clave: Inteligencia Artificial (IA); Paradoja de uso dual; Fraude financiero; Detección de fraude; IA defensiva; IA ofensiva

===∘===

A IA e o paradoxo do seu uso dual na fraude financeira: navegando entre a ofensiva e a defesa na era digital

RESUMO

Este estudo aborda o paradoxo do uso dual da inteligência artificial (IA) na fraude financeira digital, mostrando como as mesmas tecnologias podem tanto fortalecer os sistemas defensivos quanto possibilitar ataques cada vez mais adaptativos. A pesquisa adota uma metodologia qualitativa–conceitual, voltada para o desenvolvimento de um arcabouço teórico que capture esse paradoxo e destaque a lacuna de assimetria da IA entre as capacidades ofensivas e defensivas. As evidências indicam que essa lacuna está se ampliando devido a modelos de detecção reativos, restrições de acesso a dados e fragmentação no compartilhamento de inteligência. Embora soluções aplicadas (por exemplo, Mastercard Decision Intelligence Pro; SymphonyAI Sensa Copilot) relatem melhorias na precisão da detecção e na eficiência investigativa, sua efetividade parece estar limitada por restrições regulatórias, institucionais e de dados. O estudo propõe um arcabouço que integre arquiteturas híbridas de IA, redes de inteligência interinstitucionais, pesquisa em IA adversária e medidas regulatórias específicas para a IA.

Palavras-chave: Inteligência Artificial (IA); Paradoxo do uso dual; Fraude financeira; Deteção de fraude; IA defensiva; IA ofensiva

===∘===

INTRODUCTION

In the modern digital economy, trust serves as the invisible currency underpinning every financial transaction. Yet, amid the accelerated digital transformation reshaping global financial infrastructures, this trust faces an unprecedented test. Digital financial fraud has evolved from isolated incidents into a structural threat capable of destabilizing markets, undermining institutional credibility, and eroding public confidence in the systems that sustain commercial activity. The increasing interdependence on interconnected technologies creates fertile ground for the emergence of new forms of fraud that extend far beyond traditional techniques, leveraging advanced tools deployed with precision to infiltrate and manipulate financial systems.

Within this rapidly shifting landscape, artificial intelligence (AI) occupies a paradoxical and central role, as both a force multiplier for offense and a cornerstone for defense. In the hands of malicious actors, AI enables adaptive and covert attack strategies, from deepfakes that impersonate trusted identities to generative models that craft fraud patterns tailored to specific targets. Conversely, when harnessed by defenders, AI powers advanced capabilities including anomaly detection, behavioral analytics, biometric verification, and predictive modeling, offering institutions the means to identify and intercept threats in real time. Yet these defensive advantages are constrained by inherent limitations that challenge their sustained effectiveness.

The dual-use nature of AI transforms digital financial fraud into a dynamic and highly adaptive adversary. However, the rapid advancement of offensive capabilities has begun to outpace the adaptive capacity of defensive systems, widening a technological and cognitive

gap between attackers and defenders. Traditional detection mechanisms, once effective, can no longer keep pace with the qualitative leap in attack sophistication. Closing this gap demands the adoption of intelligent and flexible response models, continuous learning systems, and cross-sector collaborative approaches, underpinned by proactive regulatory frameworks tailored to the realities of AI-driven financial crime. Through such an integrated approach, innovation can be transformed from a source of systemic risk into a strategic pillar for safeguarding and stabilizing the global digital financial ecosystem.

In light of these developments, this paper is positioned as a conceptual/theoretical contribution rather than an empirical study. Building on recent scholarly literature, technical reports, and documented fraud cases, it introduces an analytical construct to describe the observed imbalance between offensive innovation and defensive adaptation in AI-enabled fraud.

Problem Statement

AI strengthens detection and response across financial infrastructures, yet it also lowers the barrier for adaptive, scalable fraud. The core question, therefore, is how AI-based defenses can remain effective and anticipatory when adversaries exploit the very same technologies to outpace them.

Research Methodology

This study adopts a qualitative–conceptual methodology grounded in inductive and interpretive analysis. The approach is designed to develop a theoretical framework that captures the dual-use paradox of artificial intelligence in digital financial fraud and highlights the persistent imbalance between offensive and defensive capabilities. Building on secondary sources, including scholarly literature, technical reports, and documented cases, the analysis led to the formulation of the AI Asymmetry Gap framework, which illustrates how offensive innovations often outpace adaptive defensive mechanisms. These insights were further connected to regulatory and institutional constraints that hinder the effectiveness of defensive tools. Overall, the methodology provides both theoretical and practical contributions to understanding the dynamics of digital financial fraud, while establishing a foundation for future empirical investigations employing quantitative or qualitative methods.

LITERATURE REVIEW

Recent scholarship has increasingly explored the intersection of artificial intelligence and financial fraud prevention, with a growing emphasis on the dual-use nature of advanced technologies. Igba et al. (2025) investigated the use of generative AI to produce synthetic datasets aimed at combating identity fraud and strengthening global financial cybersecurity frameworks. Their work highlighted how synthetic data can improve fraud detection model robustness while mitigating privacy concerns. Building on the threat dimension, Kaushik et al. (2025) examined the malicious application of deepfake technology in business and finance, illustrating how hyper-realistic fabricated media could manipulate investment decisions and undermine trust in financial institutions. Complementing these perspectives, Pazouki et al. (2025) discussed the integration of big data analytics into FinTech services, emphasizing its potential to enhance fraud detection capabilities through real-time, data-driven insights.

Bello and Komolafe (2024) offered a comprehensive overview of AI's role in fraud prevention, mapping key techniques, emerging applications, and the challenges of large-scale deployment. Similarly, Pan (2024) explored the application of machine learning to detect and prevent fraudulent financial transactions, stressing the importance of adaptive algorithms that respond to evolving fraud tactics. Onesi-Ozigagun et al. (2024) investigated AI-driven biometric authentication within FinTech, demonstrating how multi-modal biometric systems can bolster user trust and transaction security. In the cybersecurity arena, Khan et

al. (2024) reviewed AI-driven threat detection methods, focusing on their applicability to financial infrastructures under increasing attack sophistication.

Aros et al. (2024) synthesized the literature on machine learning for financial fraud detection, identifying factors such as data quality and update frequency as critical to model performance and calling for adaptive systems to address dynamic fraud landscapes. Masood et al. (2023) contributed from a technical security perspective by presenting a state-of-the-art review of deepfake generation and detection, including the open challenges in deploying detection mechanisms in high-risk sectors like finance. Finally, Javaheri et al. (2023) conducted a systematic review of cybersecurity threats in FinTech, underscoring the interplay between AI-based defense systems and the evolving sophistication of cyberattacks.

While these studies collectively underscore AI's transformative role in both offensive and defensive dimensions of financial fraud, few have explicitly addressed the strategic balance between leveraging AI for innovation and mitigating its misuse in real-world financial ecosystems. This gap calls for integrated frameworks that anticipate dual-use risks while optimizing AI deployment for sustainable financial security.

AI-DRIVEN FINANCIAL FRAUD

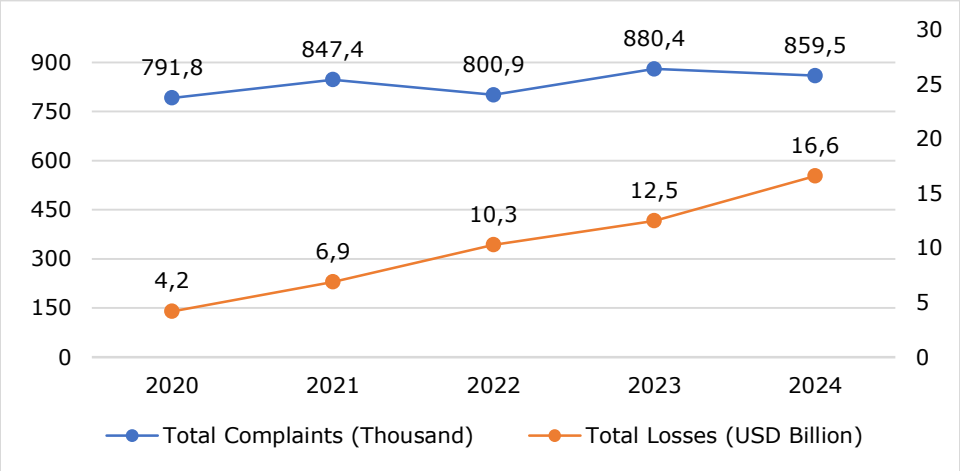
Financial fraud is among the most serious economic crimes threatening the stability of the global financial system, owing to the substantial losses it causes and its direct impact on institutional trust and financial inclusion. It is defined as the use of deception or unlawful means to obtain undue financial benefits (Directive [EU] 2017/1371, 2017). Fraud entails acts such as lying, concealment, or manipulation, intended to mislead a victim into making decisions detrimental to their interests (International Auditing and Assurance Standards Board [IAASB], 2025, para. 18[a]). This may involve deliberate falsification of information, manipulation of accounting data, or presenting false representations for illicit gain. Legally, fraud is considered a defect in consent, as it is grounded in deceit that results in decisions lacking genuine volition (Association of Certified Fraud Examiners [ACFE], 2025). This persistent threat to financial markets underscores the necessity of effective preventive tools for its detection and deterrence.

With the rapid advancement of communication and information technologies, new patterns of fraud have emerged in the digital sphere, commonly referred to as digital financial fraud. This type of fraud involves the use of digital technologies, such as computers and the internet, to target financial assets or information through activities like system intrusions, data manipulation, or identity theft. Digital fraud blends technical tools with traditional methods of deception, increasing its complexity (Pearson & Singleton, 2008). Academic and legal perspectives vary in defining this concept; while some focus on technological means and others on the fraudulent method, a comprehensive definition includes any unlawful activity conducted through digital tools to gain financial benefit by means of deception, concealment, or impersonation (Aros et al., 2024).

The scope of this phenomenon has expanded with the rise of financial technology (FinTech), which leverages technological innovation to deliver financial services. FinTech, short for financial technology, refers to the integration of internet-related technologies with business activities in the financial services industry (Liu et al., 2024). Within this evolving ecosystem, a specific type of fraud known as FinTech fraud has emerged. This includes exploiting digital lending platforms, e-wallets, and encryption systems to steal funds, manipulate data, or compromise systems. Such fraud relies on an advanced mix of technologies, including artificial intelligence, blockchain, big data, and biometric systems, making it more elusive and challenging to detect through conventional means (Javaheri et al., 2023). It surpasses traditional forms of disguise by leveraging in-depth knowledge of digital infrastructures and exploiting systemic vulnerabilities. The constantly evolving nature of this environment further complicates detection, undermining the effectiveness of traditional protective mechanisms

(Ng & Kwok, 2017). Consequently, advanced and adaptive security solutions are essential to keep pace with the evolution of digital crime tools.

Figure 1. Evolution of Internet and Technology-Enabled Financial Fraud: Complaints and Losses, 2020–2024



Source: Compiled by the author using data from the Federal Bureau of Investigation (2023, 2024), Internet Crime Report, Internet Crime Complaint Center

The data reveals a sharp escalation in both the frequency and financial severity of digital fraud incidents between 2020 and 2024. Total complaints rose from 791.8 thousand in 2020 to a peak of 859.5 thousand in 2024, reflecting an overall growth of approximately 8.6%. The initial surge between 2019 and 2020 can be attributed to the accelerated adoption of digital channels following the COVID-19 pandemic, which expanded the attack surface for cybercriminals. Although complaint volumes stabilized after 2021, remaining above 800,000 annually, the persistently high levels indicate a sustained prevalence of fraudulent activity in the digital financial ecosystem.

In contrast, total financial losses increased at a much steeper rate, rising from USD 4.2 billion in 2020 to USD 16.6 billion in 2024, an increase of approximately 295%. This consistent upward trajectory, even in years when the number of complaints plateaued or slightly declined, indicates a qualitative transformation in fraud methodologies. Attacks have shifted from low-value, high-volume scams to more targeted, high-value operations, often carried out using advanced technologies such as AI-driven reconnaissance, social engineering, and deepfake-enabled impersonation.

This sharp shift underscores the importance of understanding artificial intelligence itself as the driving force behind such transformations. Artificial intelligence (AI), as a branch of computer science, is concerned with developing systems capable of simulating human cognitive functions, including learning, analysis, inference, and decision-making (Russell & Norvig, 1995). Such systems rely on advanced algorithms to interact with data, recognize patterns, and adapt to contextual or informational changes. Extending beyond simple automation, AI enables predictive modeling and semi-autonomous decision-making within programmed parameters (Boden, 1996). Together, these features make AI a transformative force in the financial sector.

Indeed, AI's analytical capabilities have also driven notable progress in data management, risk assessment, and service delivery. Through machine learning algorithms, institutions can construct more accurate credit models and predictive tools, anticipating market fluctuations

from subtle indicators that traditional approaches often fail to capture. At the same time, these capabilities enhance customer experience by powering chatbots, enabling personalized offerings, and supporting real-time transaction processing (Hilpisch, 2020).

While these advancements reshape the financial sector in legitimate ways, they also provide fertile ground for exploitation by malicious actors, who employ them in increasingly sophisticated ways to perpetrate financial fraud. Criminals exploit advanced language models to craft convincing phishing messages that closely mimic human communication. Biometric identifiers such as facial recognition and voiceprints are increasingly vulnerable to deepfake technologies, enabling identity impersonation and the bypassing of authentication systems. Documented cases reveal AI being used to clone the voices of financial executives to issue fraudulent instructions with alarming precision. AI is also deployed to probe banking systems, identify security gaps, and develop adaptive malware capable of evading traditional defenses.

Yet, AI’s dual nature means it also stands as the first line of defense against such threats. Its ability to continuously learn and process real-time data supports advanced surveillance systems that differentiate between normal and anomalous behavior, offering immediate responses to prevent escalation. Beyond detection, AI contributes to comprehensive prevention strategies, enhancing credit scoring models, predicting anomalies indicative of fraud, and responding rapidly to unusual user activity. AI-enabled monitoring ensures that irregularities in financial transactions are addressed promptly, strengthening institutional resilience and safeguarding assets, reputation, and customer trust.

AI-POWERED OFFENSIVE PATTERNS IN FINANCIAL FRAUD

The deployment of artificial intelligence (AI) in fraudulent operations represents a significant paradigm shift in the nature and execution of cyberattacks. Unlike traditional methods, AI introduces adaptive and interactive intelligence, enabling fraudulent techniques to evolve dynamically based on feedback and the analysis of human behavior without requiring constant direct intervention by the perpetrator. The primary danger lies in its capacity to conceal malicious activity within familiar data patterns, thereby producing sophisticated threats that surpass conventional notions of system breaches.

The most notable forms of AI-powered offensive applications in digital financial fraud are as follows:

Deepfake:

Deepfake refers to highly sophisticated digital forgeries generated through the integration of deep learning techniques and AI to depict fabricated events and scenarios with exceptional realism. This technology emulates human cognitive processes by feeding machine learning algorithms, such as neural networks, with extensive datasets. These systems learn distinctive facial and vocal features and subsequently create synthetic content that convincingly mimics reality. For example, in face-swapping processes, the facial features of a targeted individual are analyzed to generate a new face that seamlessly replaces theirs in a video (Westerlund, 2019).

The growing threat posed by deepfakes is primarily due to the difficulty of detecting manipulated content, given the high precision of both visual and auditory simulations. Their credibility lies in synchronized movement, consistent expressions, and accurate speech patterns, making the differentiation between authentic and fabricated content increasingly challenging.

The risks associated with this technology are evident in its ability to generate three major types of synthetic content:

Table 1. AI Modalities for Fraudulent Content Creation

Multimodal	Voice Generation	Visual Generation
Full-body animation (puppeteering)	Text-to-speech conversion	Face-swapping
Facial synthesis with voice	Voice conversion and imitation	Lip-syncing
Motion-to-sound synchronization	Voice cloning	Facial attribute manipulation

Source: Compiled by the author based on Masood, M., et al. (2023). Deepfakes generation and detection. Applied intelligence, 53(4).

This technique exploits public trust in audiovisual media, transforming it into a powerful tool for identity theft, financial data manipulation, and influencing investment decisions. The risks are amplified by the rapid progress in AI and machine learning, which now enables the creation of synthetic content with unprecedented levels of realism and persuasiveness (Kaushik et al., 2025).

Synthetic Data Fraud

Synthetic data produced via advanced statistical and generative processes was originally developed to facilitate privacy-preserving data sharing and to train robust machine learning models. Such datasets replicate the statistical properties of real-world data while minimizing the exposure of sensitive personal information (Assefa et al., 2020). However, these same technologies are increasingly exploited for malicious purposes.

Criminals use generative AI to create synthetic identities, highly convincing combinations of AI-generated names, addresses, and personal identifiers that closely resemble those of real individuals. These fabricated identities can bypass conventional verification systems and enable a range of illicit activities, such as account takeovers, fraudulent loan applications, and unauthorized transactions. The blurred boundary between ethically generated synthetic datasets and malicious synthetic identities presents a severe challenge to the financial sector (Igba et al., 2025). Consequently, the adoption of advanced AI-driven fraud detection measures, including biometric verification and adversarial AI techniques, has become essential to identify and counter such threats.

Social Engineering and Deep Phishing

Smart social engineering represents an evolution of traditional manipulation tactics, enhanced by the capabilities of AI, particularly generative models. This approach exploits human psychology, trust, and behavior, not through direct system intrusion, but by crafting highly realistic and contextually tailored narratives. Generative AI facilitates the creation of text, voice, and even video content that closely mirrors authentic human communication patterns (Schmitt & Flechais, 2024).

In financial fraud contexts, such tactics are especially dangerous. Attackers may use personalized phishing, voice impersonation, or deepfake videos to convincingly pose as colleagues, executives, or bank representatives (Timoney, 2025). By analyzing the digital footprint of a target, AI tools can construct scenarios aligned with the victim's professional or social environment, thereby increasing the probability of deception. This precision elevates social engineering into a persistent and advanced threat, capable of circumventing traditional security mechanisms.

AI-Based CAPTCHA Breaking

CAPTCHA systems (Completely Automated Public Turing Test to Tell Computers and Humans Apart) have long been a primary defense against automated system abuse. However, advancements in AI, particularly in computer vision and human interaction modeling, now

enable bots to solve CAPTCHA challenges with high accuracy, undermining their effectiveness (Mohitkar et al., 2025).

In financial platforms such as online banking, trading systems, and e-wallet applications, AI-driven CAPTCHA bypassing facilitates unauthorized access, automated fraudulent transactions, identity theft, and mass data scraping. This escalating risk highlights the urgent need for adaptive, behavior-based authentication methods.

Malicious AI-Driven Tools and Services

A new generation of malicious AI tools has emerged on the dark web, lowering the barrier for conducting sophisticated fraud. FraudGPT, a subscription-based tool, enables the automated generation of phishing content and malware that can evade conventional detection systems. WormGPT, based on the GPT-J model, is designed for business email compromise schemes by generating precise and convincing messages. Other notable tools include VALL-E, capable of cloning human voices, and image-generation models such as DALL-E 2 and Stable Diffusion, which can be used for disinformation and reputational damage (Falade, 2023). Additionally, services such as OnlyFake produce AI-generated fake identities capable of bypassing digital verification protocols (Lemonnie, 2025).

These developments intensify the psychological dimension of fraud, enabling attacks that exploit urgency, trust, and emotional triggers, such as fabricated voice calls requesting emergency fund transfers.

Algorithmic Manipulation in Financial Trading Markets

AI-powered trading algorithms have transformed financial market operations, enhancing execution speed and analytical capacity. However, these systems can also be exploited for market manipulation, such as spoofing or generating false liquidity signals to mislead investors (Olanrewaju, 2025). The opacity of "black-box" AI models further complicates detection, as their decision-making processes remain difficult to interpret.

Fraudsters exploit such vulnerabilities to influence market dynamics, inject falsified data, or induce irrational trading decisions, thereby undermining investor confidence. Addressing these risks requires the implementation of advanced regulatory mechanisms capable of monitoring AI decision-making, enforcing transparency, and ensuring accountability in automated trading environments.

AI-DRIVEN FINANCIAL FRAUD DETECTION AND PREVENTION MECHANISMS

By integrating an advanced set of technologies to strengthen financial security, generative artificial intelligence (AI) has undergone rapid evolution, delivering groundbreaking applications in both cybersecurity and financial protection. Leveraging training on massive datasets, these models can autonomously generate text, images, and code, creating both innovative opportunities and significant security challenges. In the security domain, generative AI enhances threat detection capabilities by identifying patterns of malicious activity and enabling the development of predictive models to anticipate and mitigate risks (Igba et al., 2025). For instance, it can simulate diverse attack scenarios to proactively test defensive measures or generate synthetic datasets that replicate real-world environments without exposing sensitive information, an essential resource for training fraud detection systems capable of adapting to emerging threats. In the financial sector, generative AI transforms security threats into opportunities to enhance institutional resilience by integrating.

Machine Learning–Based Detection Models

Machine Learning (ML) represents a foundational approach to improving fraud detection and prevention in the financial sector. By analyzing vast transaction datasets, ML uncovers subtle and complex fraud patterns that traditional rule-based systems cannot detect (Bello & Komolafe, 2024). It encompasses a range of techniques, including supervised learning, un-

supervised learning, reinforcement learning, deep learning, and natural language processing, each contributing distinct capabilities to anomaly detection in transaction frequency, value, location, and device usage in real time.

Supervised learning, which relies on pre-labeled datasets, enables the development of models capable of distinguishing between legitimate and fraudulent activities. Techniques such as Decision Trees, Random Forests, and Neural Networks allow financial institutions to capture both simple and complex patterns in customer and transaction data, thereby identifying cases such as repeated withdrawals across multiple ATMs within short time intervals or transactions initiated from high-risk regions (Pan, 2024; Hernandez Aros et al., 2024). In contrast, unsupervised learning operates without pre-labeled data, making it particularly effective in detecting irregularities in new transaction types or emerging markets. Methods like k-means[1] clustering, which segments customers by spending profiles, and Isolation Forests[2], which flag rare or atypical behaviors, are often used to detect sudden spikes in spending or geographically inconsistent purchases (Pan, 2024; Hernandez Aros et al., 2024).

Complementing these approaches, reinforcement learning leverages a feedback mechanism to enhance detection strategies iteratively. For instance, banking systems may progressively learn to block large transfers from newly opened accounts or to request additional verification for anomalous spending patterns (Pan, 2024). Deep learning further strengthens these capabilities by employing multi-layer neural networks. Convolutional Neural Networks (CNNs) are effective in identifying visual patterns in financial documents, such as forged signatures, whereas Recurrent Neural Networks (RNNs) excel at analyzing sequential data to uncover coordinated fraudulent behavior (Xu Zhang, 2024). When used together, these models enable near real-time fraud detection without compromising user privacy.

Finally, Natural Language Processing (NLP) expands the scope of fraud detection by allowing systems to interpret and analyze text from diverse sources, including emails, social media, customer service logs, and dark web marketplaces. In financial contexts, NLP can detect linguistic markers of phishing attempts, scam investment proposals, and even identify stolen credit card information. By integrating sentiment analysis, keyword monitoring, and contextual interpretation, NLP provides a proactive layer of defense, allowing financial institutions to anticipate and neutralize fraudulent schemes before they directly impact clients (Khan et al., 2024).

Powered Data Analytics in Financial Fraud Detection

Big data is defined as a technological phenomenon representing a vast, diverse, and continuously flowing stream of information from multiple sources, surpassing the capabilities of traditional data management systems in processing and analyzing it efficiently (De Mauro et al., 2016). In practice, this phenomenon is enabled by modern big-data tools and platforms for storage, distributed processing, and streaming, that operationalize large-scale analytics in real environments (Mazumder, 2016). Big data is typically characterized by volume, velocity, and variety, the amount, speed, and heterogeneity of information such as texts, images, and videos (Pazouki et al., 2025).

[1]. **K-means Clustering Algorithm**: The K-means algorithm functions as an intelligent data organizer similar to sorting a group of clients into similar clusters. The process begins by selecting a predefined number of "centroid points" known as cluster centers. The algorithm then classifies each financial transaction by assigning it to the nearest centroid based on shared characteristics. Source: https://www.simplilearn.com/tutorials/machine-learning-tutorial/k-means-clustering-algorithm

[2]. **Isolation Forest**: Isolation Forest is an intelligent unsupervised algorithm used to detect anomalies. It operates similarly to real forests, which consist of many trees. The algorithm relies on a collection of decision trees that work together to reach a final outcome. What distinguishes this algorithm from others is that it directly searches for abnormal and anomalous data, instead of first defining what is normal and then classifying deviations.Source:https://cloud.rproject.org/web/packages/isotree/vignettes/An_Introduction_to_Isolation_Forests.html

In the context of financial fraud prevention, advanced big data analytics have emerged as a powerful approach for leveraging both structured and unstructured datasets to detect and prevent fraudulent activities. These analytics apply a range of cutting-edge techniques, including data mining, which facilitates the identification of fraud patterns such as stolen credit card usage or rapid sequences of small, suspicious transactions; machine learning, which enables the development of intelligent models capable of automatically detecting unusual activities; and anomaly detection algorithms, which employ AI to uncover breach attempts by malicious bots, human fraudsters, or other abusive entities (Sharma et al., 2016). The effectiveness of these techniques is amplified by modern big-data platforms and tools that allow large-scale, distributed, and real-time analysis of financial transactions (Mazumder, 2016).

Building on these capabilities, financial institutions increasingly implement integrated defense systems that combine multiple analytical mechanisms to identify and respond to fraud early and accurately. Key components of such systems include real-time transaction analysis using streaming technologies like Apache Flink [3]and Apache Kafka[4] to flag unusual geographic usage or atypical transaction amounts immediately; comprehensive user behavior analysis to detect deviations from established behavioral patterns; integration with external data sources, including public records and social media, to uncover inconsistencies in client activity; and predictive modeling that employs supervised, unsupervised, and anomaly detection techniques to continuously adapt to evolving fraud strategies (Udeh et al., 2024).

Together, these elements create a highly efficient digital shield, reducing financial losses and enhancing the resilience of the financial system against increasingly sophisticated and persistent cyber threats

Biometric AI Tools in Financial Fraud Detection

Biometric technologies constitute one of the most robust defense mechanisms in today's digital landscape, with their effectiveness significantly enhanced when integrated with artificial intelligence (AI) algorithms. This integration harnesses the uniqueness of individual physiological and behavioral characteristics, such as fingerprints, facial geometry, and voice patterns, within intelligent systems capable of precise data matching, anomaly detection, and the identification of forgery or impersonation attempts. As a result, biometrics have evolved beyond their traditional role as tools for identity verification to become fundamental components of advanced authentication frameworks designed to safeguard sensitive information, financial assets, and critical infrastructures.

Biometric systems function by automatically identifying or verifying individuals based on distinctive physiological traits, including fingerprints, hand geometry, facial structure, voice, iris, and retinal patterns, or behavioral markers such as signature dynamics and keystroke patterns (Vacca, 2007). These systems capture and process a user's unique biometric data, subsequently authenticating their identity by comparing it against securely stored reference templates collected during the enrollment process.

Within the financial technology (FinTech) ecosystem, biometric authentication has gained significant traction as a countermeasure against fraud. It is widely deployed to secure digital banking operations, protect high-value assets and sensitive data, and ensure the integrity of automated financial processes, while simultaneously reinforcing broader cybersecurity strategies. By leveraging the inherent uniqueness of biological and behavioral traits, biometric authentication addresses the vulnerabilities of traditional methods, such as passwords, PIN codes, ID cards, and credit cards, which remain susceptible to theft, duplication, and unau-

[3]. **Apache Flink**: A distributed data processing engine designed for fast and scalable computations on streaming data, whether bounded or unbounded. See: https://flink.apache.org/

[4]. **Apache Kafka**: A platform for fast and reliable data streaming, enabling real-time analytics and data integration across systems. See: https://kafka.apache.org/

thorized use. In contemporary authentication frameworks, security factors are typically divided into three categories: knowledge-based (e.g., passwords), possession-based (e.g., ID cards), and biometric-based (e.g., fingerprint or facial recognition) (Schiavone et al., 2019).

The integration of AI into biometric authentication systems significantly strengthens fraud prevention by enhancing accuracy, reducing false acceptance and rejection rates, and enabling continuous adaptation to evolving user behaviors. AI-enabled biometric systems employ modalities such as facial scans, iris recognition, and voice authentication to provide higher verification accuracy than conventional approaches, while simultaneously addressing privacy and trust concerns. Beyond identity validation, AI-driven biometrics contribute to fraud detection by analyzing behavioral patterns and identifying anomalies that may indicate malicious intent (Onesi-Ozigagun et al., 2024; Kuraku et al., 2020).

In practice, deviations in biometric input from an established behavioral baseline can trigger secondary authentication protocols, thereby creating a proactive security barrier. This intelligent security framework is built upon three interdependent pillars: (1) biometric enrollment, involving the secure capture, encryption, and storage of unique user identifiers; (2) dual verification, which combines one-to-one and one-to-many biometric matching with supplementary factors such as passwords; and (3) real-time transaction monitoring, in which attributes such as geolocation, transaction amount, and frequency are analyzed against historical behavioral patterns to promptly detect and block irregular activities (Alay & Al-Baity, 2020). By integrating biometric technologies, AI algorithms, and big data analytics, financial institutions establish an adaptive digital shield that enhances resilience, fortifies customer trust, and mitigates financial losses arising from fraudulent activities.

APPLIED AI MODELS FOR COMBATING FINANCIAL FRAUD

The deployment of advanced artificial intelligence (AI) models in financial fraud prevention represents a paradigm shift toward next-generation, intelligent financial security solutions. Among the most prominent implementations are the following:

Decision Intelligence Pro (DI Pro)

Mastercard has developed Decision Intelligence Pro (DI Pro), an advanced AI-powered platform aimed at enabling banks and financial institutions to detect and prevent fraudulent activities with unprecedented precision. Utilizing sophisticated AI algorithms, DI Pro evaluates the interrelationships among multiple entities within a transaction, processing up to one trillion data points to assess fraud risk in real time. This risk evaluation is completed in under 50 milliseconds. Early performance assessments demonstrate that DI Pro can increase fraud detection rates by an average of 20%, and by up to 300% in certain contexts, while simultaneously reducing false positives by more than 85%. Building upon Mastercard's existing Decision Intelligence infrastructure, which already analyzes over 143 billion transactions annually, DI Pro provides financial institutions with a highly effective defense against increasingly sophisticated fraud schemes (Mastercard, 2024). This model exemplifies how real-time, high-volume data analysis can be operationalized for proactive fraud detection. Its ability to assess vast transactional relationships within milliseconds indicates a shift toward predictive, behavior-based security. However, such models also raise concerns regarding algorithmic transparency and potential over-reliance on proprietary infrastructures.

Sensa Copilot Service

Sensa Copilot, developed by SymphonyAI, is an investigative assistant powered by both generative and predictive AI, designed to enhance the productivity and decision-making capabilities of financial crime investigators. The system integrates advanced functionalities, including AI-generated investigative recommendations, intelligent web-based searches, and the extraction of actionable risk insights from interconnected datasets. It further supports relationship mapping to reveal hidden links within complex data environments and offers prioritized task lists alongside suggested investigative steps. Interactive queries, AI-ranked

search summaries, and structured workflow guidance contribute to higher-quality and more consistent investigative outcomes. According to operational evaluations, the integration of generative AI within Sensa Copilot has reduced investigation times by up to 70% while improving overall detection accuracy and consistency (SymphonyAI, 2024). Sensa Copilot reflects the strategic evolution of AI from passive detection to investigative assistance. By integrating generative AI with workflow intelligence, it enhances human decision-making rather than replacing it. Still, its effectiveness may depend heavily on user training, contextual understanding, and ethical handling of generative outputs in sensitive investigations.

These two applied AI models, DI Pro and Sensa Copilot, illustrate the growing maturity of defense-focused AI innovations in the financial sector. While DI Pro emphasizes transaction-level risk scoring through vast data analysis in milliseconds, Sensa Copilot shifts the focus toward investigative augmentation and intelligent decision support. Both systems reflect a crucial trend: the transition from static detection models to real-time, context-aware platforms.

COMPARATIVE ANALYSIS

This section develops a comparative analysis of artificial intelligence (AI) in financial fraud, focusing on two complementary dimensions. Section 5.1 examines the attacker–defender asymmetry, analyzing how AI reshapes the balance between offensive and defensive techniques. Section 5.2 examines the distinction between the direct and supporting roles of AI in fraud, distinguishing between categories where AI generates fraudulent content or actions and those where AI serves as an enabler by amplifying scale, personalization, and automation.
Together, these perspectives provide a structured lens for understanding the systemic implications of AI in fraud ecosystems, setting the stage for the empirical evidence presented in subsequent tables and figures.

AI Asymmetry Gap in Financial Fraud

The concept of asymmetry in cybersecurity refers to the inherent imbalance between attackers and defenders: attackers need to succeed only once, while defenders must succeed every time (Guo et al. 2025). This imbalance is further amplified in the context of artificial intelligence, where offensive actors can rapidly deploy adaptive, automated, and scalable fraud techniques, while defensive systems remain constrained by slow, resource-intensive adaptation, regulatory lag, and fragmented intelligence sharing.

In financial fraud ecosystems, attackers leverage automation, scalability, and stealth, whereas defenders face regulatory, institutional, and data-access constraints. This imbalance renders AI a potential systemic vulnerability unless addressed through integrated technical, organizational, and regulatory strategies.

Table 2. Comparative Summary of Offensive vs. Defensive AI Techniques In Financial Fraud

Dimension	Offensive AI Techniques	Defensive AI Techniques
Purpose	Exploiting system vulnerabilities for financial gain	Detecting, preventing, and mitigating fraud
Key Tools/Applications	Deepfakes, FraudGPT, WormGPT, Synthetic identities, AI CAPTCHA bypass	Machine learning, biometric AI, anomaly detection, NLP, DI Pro
Adaptability	Highly adaptive to defenses and context	Reactive, needs continuous updates
Scalability	Easily scalable via automation and AI agents	Limited by data privacy and infrastructure
Data Utilization	Uses stolen/leaked data, behavioral mimicry	Relies on structured and labeled datasets
Stealth Techniques	Mimics legitimate behavior, social engineering, and deep impersonation	Behavioral baselines, anomaly scoring, and AI pattern learning
Limitations	Legal exposure, traceability in some cases	Data access limitations, fragmented intelligence sharing
Examples	CEO voice cloning, AI-generated fake IDs	Mastercard DI Pro, SymphonyAI Sensa Copilot

Source: Compiled by the author using literature and case analysis

↓ **Creates Imbalance (Gap)**
[The AI Asymmetry Gap]

- Offense evolves faster than defense
- Defense reactive, limited by data/privacy
- Fragmented intelligence sharing
- Regulatory lag

Table 2 presents a comparative overview of offensive versus defensive AI applications in financial fraud, contrasting their purposes, tools, adaptability, scalability, data use, stealth techniques, and limitations. The evidence highlights a widening "AI Asymmetry Gap": offensive AI evolves faster than defenses constrained by data privacy, infrastructure, and coordination. The consequences include market instability, institutional risk, and erosion of trust, underscoring the need for hybrid AI defenses, intelligence-sharing networks, adversarial AI research, and proactive regulatory frameworks.

Having outlined offensive and defensive AI mechanisms and the AI Asymmetry Gap, we now turn to category-level evidence.

AI Enablement in Financial Fraud: Direct vs. Supporting Roles

To move from how mechanisms operate to where they appear in the data, Table (3) reports 2020–2024 complaints and losses along two dimensions: Direct categories (AI-enabled), where AI produces the fraudulent content or executes the deceptive step (e.g., fake investment bots, audio/video impersonation in BEC), and Supporting/enabling categories, where AI expands the attack surface through automation, personalization, and use of leaked data. In the former, AI facilitates the creation of convincing platforms, executive-style messages, or persuasive tech-support scripts, thereby increasing losses per incident. In the latter, AI enables large-scale personalization of phishing and the construction of synthetic identities from breached data for downstream fraud. Thus, AI not only amplifies the reach and persuasiveness of financial fraud but also creates structural linkages between high-volume enabling categories and high-loss direct attacks.

The observed divergence between direct and supporting categories can be better understood by examining the mechanisms through which AI amplifies these schemes. In investment

This distinction lets categories with prevailing attack patterns and helps explain why the AI Asymmetry Gap widens: direct pathways tend to raise loss per incident, whereas supporting pathways increase incident volume and adaptation speed.

Table 3. Financial fraud complaints and losses by category (2020–2024) with AI Enablement (Author's Classification)[5]

Type	Complaints (Thousand)					Loss (USD Billion)					AI Enablement (examples)	AI Relevance
	2024	2023	2022	2021	2020	2024	2023	2022	2021	2020		
Direct (AI-enabled)												
Investment	47,9	39,6	30,5	20,6	8,8	6,57	4,57	3,31	1,46	0,34	AI chatbots and generated trading sites mimicking real platforms	High (core AI-enabled)
Business Email Compromise	21,4	21,5	21,8	20	19,4	2,77	2,95	2,74	2,4	1,87	Deepfake voice/video of executives; AI-crafted emails	High (core AI-enabled)
Tech Support	36	37,6	32,5	23,9	15,4	1,46	0,92	0,81	0,35	0,15	Cloned voices and AI chatbots posing as support agents	High (core AI-enabled)
Confidence Fraud/Romance	17,9	17,8	19	24,3	23,8	0,67	0,65	0,74	0,96	0,6	LLM-driven chats + AI-generated photos/videos for fake personas	High (core AI-enabled)
Government Impersonation	17,4	14,2	11,6	11,3	12,8	0,41	0,39	0,24	0,14	0,11	Deepfake voice of officials; generated official-looking docs	High (core AI-enabled)
Credit Card/Check Fraud	12,9	13,7	23	16,8	17,6	0,2	0,17	0,26	0,17	0,13	Automated pattern mining; AI voice to bypass KBA	Medium (AI-augmented)
Supporting(enabler)												
Phishing/ Spoofing	193,4	298,9	321,1	342,5	269,6	0,07	0,02	0,16	0,13	0,27	Personalized spear-phishing at scale; synthetic voice calls	High (core AI-enabled)
Personal Data Breach	64,9	55,9	58,9	51,8	45,3	1,45	0,74	0,74	0,52	0,19	AI storefronts; persuasive copy generation; automated dispute/ chargeback manipulation.	High (core AI-enabled)
Data Breach	3,2	3,7	2,8	1,3	2,8	0,36	0,53	0,46	0,15	0,13	Automated data exfiltration/analysis for follow-on fraud	High (core AI-enabled)
Identity Theft	21,4	19,8	27,9	51,6	43,3	0,17	0,13	0,19	0,28	0,22	Synthetic identities; AI-aided document forgery	High (core AI-enabled)
Botnet	0,6	0,5	0,6			0,01	0,02	0,02			Automated infrastructure for mass phishing/payload delivery	High (core AI-enabled)
Malware	0,4	0,7	0,8	0,8	1,4	0,001	0,001	0,01	0,01	0,01	AI-assisted obfuscation; LLM-guided droppers	High (core AI-enabled)

Source: Compiled by the author using data from the Federal Bureau of Investigation. (2022–2024). Internet Crime Reports. Internet Crime Complaint Center (IC3)

[5]. Note: The AI enablement classifications (Direct / Supporting) reflect the author's judgment based on the literature and documented cases.

While the table (3) provides detailed values for complaints and losses by type and year, supported by the AI Enablement (Author's Classification) field, the figure (3) illustrates their trajectories over time and the points of increase and decrease.

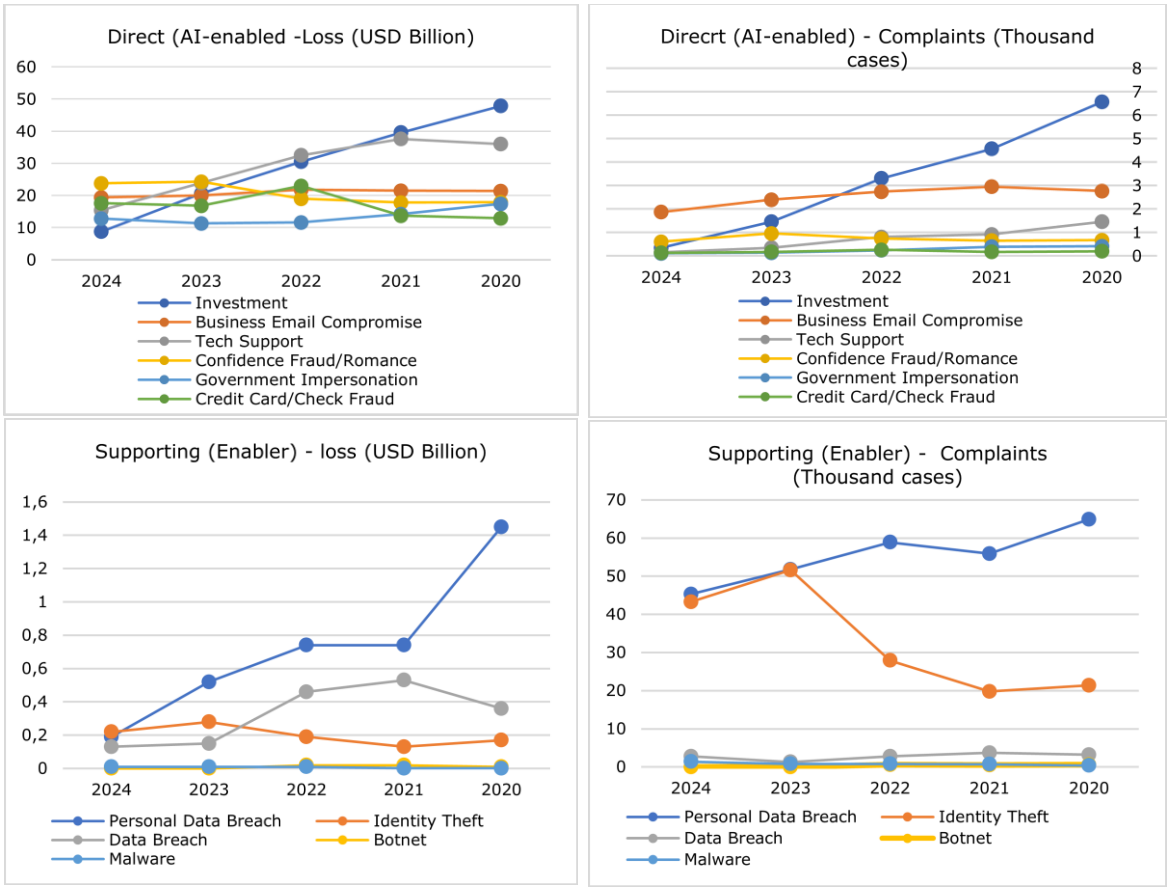


Figure 3. Complaints and Losses Over Time for AI-Driven Financial Fraud (Direct vs. Supporting), 2020–2024

Source: Compiled by the author using Table 3

Table 3, alongside Figure 3, reveals substantial differences between direct and supporting categories of AI-driven financial fraud during the period 2020–2024.

For the direct categories, investment fraud registered a dramatic escalation in losses, rising from approximately USD 0.34 billion in 2020 to more than USD 6.5 billion in 2024, alongside an increase in complaints from 20.6 thousand to 47.9 thousand—corresponding to a compound annual growth rate exceeding 80%. The average loss per incident in this category is about USD 0.14 million (≈137 thousand), the highest among all categories.

Business Email Compromise (BEC) showed relatively stable complaint volumes (around 20–22 thousand annually) but rising losses from USD 1.87 billion in 2020 to USD 2.77 billion in 2024, yielding an average of roughly USD 0.13 million (≈129 thousand) per incident.

Tech Support fraud also displayed striking growth, with losses increasing from USD 0.15 billion in 2020 to USD 1.46 billion in 2024 and complaints from 23.9 thousand to 36 thousand

over the same period, culminating in an average of approximately USD 0.04 million (≈40 thousand) per incident in 2024.

The supporting/enabling categories present a contrasting picture. Phishing/Spoofing remained the largest in volume, exceeding 300 thousand annual complaints in 2021–2022 before declining to 193 thousand in 2024, while losses grew from USD 0.27 billion to USD 0.70 billion, with an average loss of only about USD 0.004 million (≈3.6 thousand) per incident.

Identity Theft declined both in scale (from 51 thousand complaints in 2021 to 21 thousand in 2024) and in losses (from USD 0.28 billion to USD 0.14 billion), reflecting its reduced relative impact.

By contrast, Personal Data Breach exhibited steady growth, with complaints increasing from 45.3 thousand to 64.9 thousand and losses surging from USD 0.19 billion to USD 1.45 billion, yielding an average of USD 0.022 million (≈22 thousand) per incident in 2024, confirming its role as a critical feeder into downstream attacks such as account takeover and synthetic identity construction.

Aggregating the 2024 data highlights the structural divergence: direct categories accounted for approximately 151 thousand complaints with total losses of USD 12.8 billion, averaging USD 0.085 million (≈85 thousand) per incident. Supporting categories, in contrast, registered about 325 thousand complaints but only USD 2.3 billion in losses, with a much lower average of USD 0.007 million (≈7.2 thousand) per incident. In other words, despite their smaller volume, direct categories produced more than five times the financial damage of supporting categories, while the latter generated more than twice the number of complaints but substantially lower per-incident losses.

These results underscore a pronounced bifurcation: high-loss direct categories (Investment, BEC, Tech Support) generate fewer but more damaging incidents, whereas high-volume supporting categories (Phishing, Identity Theft) scale frequency without a proportional increase in harm. Personal Data Breach, moreover, emerges as a structural enabler that broadens the threat surface and fuels downstream fraud schemes, even without recording the highest per-incident averages on its own. This contrast aligns with the asymmetric gap framework and justifies the subsequent shift in discussion toward the trade-off of adaptive capabilities between attackers and defenders.

FINDINGS AND DISCUSSION

By synthesizing recent scholarship, technical reports, case evidence, and the comparative tables, three interrelated patterns emerge. First, offensive AI practices—including deepfakes, synthetic identities, AI-enhanced social engineering, CAPTCHA circumvention, malicious generative tools, and algorithmic market manipulation- exhibit high adaptability and scalability, often leveraging stolen or synthetic data and behavioral mimicry to blend with legitimate activity while expanding systemic exposure.

Second, defensive AI applications such as ML-based anomaly detection, biometric verification, big-data analytics, and applied decision-support platforms (e.g., Mastercard Decision Intelligence Pro; SymphonyAI Sensa Copilot) show measurable gains in accuracy, context awareness, and response time. However, these improvements remain constrained by data-access restrictions, privacy concerns, infrastructural heterogeneity, and the absence of cross-institutional intelligence sharing.

Third, the divergence between offensive acceleration and defensive reactivity constitutes a persistent imbalance, conceptualized in this study as the "AI Asymmetry Gap".Table 3 illustrates this imbalance: direct categories such as Investment Fraud, Business Email Compromise (BEC), and Tech Support generate disproportionately high per-incident losses often exceeding USD 100,000—despite comparatively lower complaint volumes. By contrast, enabling categories such as Phishing/Spoofing and Identity Theft operate at a much larger

scale but yield far lower per-case losses, while Data Breaches function primarily as structural enablers, feeding downstream fraud without a proportional escalation in per-incident costs.

Building on this synthesis, we outline a unified roadmap that integrates technical, institutional, and regulatory measures. This roadmap aims to assist both scholars and practitioners in narrowing the AI Asymmetry Gap by fostering adaptive defenses, strengthening cross-institutional collaboration, and aligning governance with the evolving dynamics of AI-driven financial fraud.

1. Technical Measures
- Develop hybrid AI architectures integrating supervised, unsupervised, and reinforcement learning to address both established and emerging fraud patterns.
- Establish cross-institutional AI threat intelligence networks to enable real-time sharing of fraud indicators among banks, fintech companies, and regulators.
- Invest in adversarial AI research to anticipate and counter evolving generative attack models before they reach operational maturity.
2. Institutional Measures
- Form multidisciplinary fraud response teams comprising data scientists, cybersecurity experts, behavioral analysts, and legal specialists.
- Conduct periodic AI capability audits to ensure detection models remain updated, resilient, and stress-tested against simulated attack scenarios.
- Implement AI literacy training for personnel at all operational levels to strengthen organizational readiness against AI-enabled threats.
2. Regulatory and Policy Measures
- Introduce AI-specific compliance frameworks mandating transparency and explainability in algorithmic decision-making for financial security systems.
- Require mandatory reporting of AI-enabled fraud incidents to centralized regulatory databases to support pattern recognition and coordinated responses.
- Promote international cooperation agreements to facilitate cross-border sharing of AI-related threat intelligence, given the inherently global nature of digital financial fraud.

By embedding these recommendations within unified, adaptive, and collaborative defense strategies, stakeholders can narrow the AI Asymmetry Gap and help shift AI from a systemic vulnerability toward a cornerstone of trust, resilience, and long-term stability in the global digital economy.

CONCLUSION

In conclusion, this study highlights the dual-use paradox of artificial intelligence (AI) in financial fraud and advances the AI Asymmetry Gap as a conceptual lens to understand the imbalance between rapidly evolving offensive innovation and comparatively slower, more constrained defensive adaptation. Through a critical synthesis of recent scholarship, technical reports, and documented cases, we mapped offensive and defensive applications and clarified why the gap tends to widen due to reactive detection models, data-access constraints, fragmented intelligence sharing, and regulatory lag.

The primary contribution is theoretical: AI-enabled fraud is reframed as a moving target that co-evolves with institutional defenses rather than as a static security problem. The framework organizes technical, institutional, and regulatory factors without prescribing specific interventions, providing an orientation for subsequent inquiry. Consistent with this positioning, future research should develop operational metrics for the gap, pursue longitudinal designs linking incident trends to capability shifts, and employ simulation or controlled experiments to evaluate detection under adaptive adversaries across differing environments.

LIMITATIONS

This paper is conceptual and relies on publicly available secondary sources; some findings may be time-sensitive given the pace of AI. Moreover, the conceptual nature of the study limits the generalizability of findings to dynamic real-world contexts. Future research should incorporate broader datasets, multi-jurisdictional evidence, and practitioner input.

CONFLICT OF INTEREST

The author declares no conflicts of interest.

REFERENCES

Alay, N., & Al-Baity, H. H. (2020). Deep learning approach for a multimodal biometric recognition system based on fusion of iris, face, and finger vein traits. Sensors, 20(19), 5523. https://doi.org/10.3390/s20195523

Apache Flink. (n.d.). *Apache Flink: Stateful computations over data streams*. The Apache Software Foundation. Retrieved from https://flink.apache.org/

Apache Kafka. (n.d.). *Apache Kafka: A distributed streaming platform*. The Apache Software Foundation. Retrieved from https://kafka.apache.org/

Aros, C., Astudillo, P., Riffo, V., Vásquez, R., & Díaz, J. (2024). Financial fraud detection through the application of machine learning techniques: A literature review. Humanities and Social Sciences Communications, 11, Article 312. https://doi.org /10.1057/s41599-024-03606-0

Assefa, S. A., Dervovic, D., Mahfouz, M., Tillman, R. E., Reddy, P., & Veloso, M. (2020). Generating synthetic data in finance: Opportunities, challenges, and pitfalls. Proceedings of the First ACM International Conference on AI in Finance, 1–8. https://doi.org/10.1145/3383455.3422559

Association of Certified Fraud Examiners. (2025). What is fraud? Retrieved May 15, 2025, from https://www.acfe. com/fraud-resources/-what-is-fraud

Bello, O. A., & Komolafe, O. (2024). Artificial intelligence in fraud prevention: Exploring techniques and applications, challenges, and opportunities. Computer Science & IT Research Journal, 5(6), 1505–1520. Retrieved from https://fepbl.com/index .php/csitrj/article/view/1252

Boden, M. A. (Ed.). (1996). Artificial intelligence. Elsevier.

De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. Library Review, 65(3), 122–135. https://doi.org/10.1108/LR-06-2015-0061

Directive (EU) 2017/1371 of the European Parliament and of the Council of 5 July 2017 on the fight against fraud to the Union’s financial interests by means of criminal law, Article 3(2), Journal officiel de l'Union européenne, p. 34.

Falade, P. V. (2023). Decoding the threat landscape: ChatGPT, FraudGPT, and WormGPT in social engineering attacks. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 87–190. arXiv:2310.05595. https://doi.org/10.48550/arXiv.2310.05595

Federal Bureau of Investigation, Internet Crime Complaint Center. (2022). *Internet Crime Report 2022.* https://www.ic3.gov/AnnualReport/Reports/2022_ic3report.pdf.

Federal Bureau of Investigation. (2024). Internet Crime Report 2023. Internet Crime Complaint Center (IC3). Retrieved from https://www.ic3.gov/Media/PDF/Annual Report/2023_IC3Report.pdf

Federal Bureau of Investigation. (2025). Internet Crime Report 2024. Internet Crime Complaint Center (IC3). Retrieved from https://www.ic3.gov/Media/PDF/Annual Report/2024_IC3Report.pdf

Hilpisch, Y. (2020). Artificial intelligence in finance. O’Reilly Media. https://doi.org/10.22215/timreview/1282

Igba, E., Salam Olarinoye, H., Ezeh Nwakaego, V., Batur Sehemba, D., Shade Oluhaiyero, Y., & Okika, N. (2025). Synthetic Data Generation Using Generative AI to Combat Identity Fraud and Enhance Global Financial Cybersecurity Frameworks. International Journal of Scientific Research and Modern Technology, 4(2), 1–19. https://doi.org/10 .5281/zenodo.14928919

International Federation of Accountants. (n.d.). Auditor's responsibility to consider fraud in an audit of financial statements: ISA (240) (p. 179). Retrieved from http://www.ifac .org

Javaheri, A., Alimohammadi, S., Yazdinejad, A., Parizi, R. M., Dehghantanha, A., & Karimipour, H. (2023). Cybersecurity threats in FinTech: A systematic review. arXiv.https://doi.org/10.48550/arXiv.2312.01752

Kaushik, P., Garg, V., Priya, A., & Kant, S. (2025). Financial fraud and manipulation: The malicious use of deepfakes in business. In Deepfakes and their impact on business (pp. 173–196). IGI Global. https://doi.org/10.4018/979-8-3693-6890-9.ch008

Khan, M. I., Arif, A., & Khan, A. R. A. (2024). AI-Driven Threat Detection: A Brief Overview of AI Techniques in Cybersecurity. BIN: Bulletin of Informatics, 2(2), 248–261. Retrieved from https://ojs.jurnalmahasiswa.com/ojs/index.php/bin/article/view/357

Kuraku, C., Gollangi, H. K., & Sunkara, J. R. (2020). Biometric Authentication In Digital Payments: Utilizing AI And Big Data For Real-Time Security And Efficiency. Educational Administration: Theory and Practice, 26(4), 954–964. https://doi.org/10.53555 /kuey.v26i4.759

Lemonnie, J. (2025, May 12). The truth about OnlyFake and generative AI fraud. Resistant.AI. https://resistant.ai/blog/onlyfake-generative-ai-fraud

Liu, Z., Chan, K. C., & Chimhundu, R. (2024). Mapping the landscape of financial technology (FinTech) research: A systematic review and bibliometric analysis. Financial Innovation, 10(1), 1–26. https://doi.org/10.1186/s40854-023-00524-z

Masood, M., Nawaz, M., Javed, A., Irtaza, A., & Mahmood, M. T. (2023). Deepfakes generation and detection: State-of-the-art, open challenges, countermeasures, and way forward. Applied Intelligence, 53(4), 3740–3768. https://doi.org/10.1007/ s10489-022-03832-x

Mastercard. (2024a, February 13). Mastercard supercharges consumer protection with Gen AI.https://www.mastercard.com/news/press/2024/february/mastercard-supercharges-consumer-protection-with-gen-ai/

Mohitkar, C., Kharat, A., Nashirkar, N., Pardeshi, T., & Gaikwad, P. S. (2025). Passive Captcha: AI-driven bot detection for seamless user experience. International Journal of Environmental Sciences, 11(1), 362–370. Retrieved from https://www.cibtech.org/ J-ENV-SCI/PUBLICATIONS /2025/Vol_11_No_1/50-JES-014-CHIRAG-MOHITKAR.pdf

Ng, A. W., & Kwok, B. K. B. (2017). Emergence of Fintech and cybersecurity in a global financial centre: Strategic approach by a regulator. Journal of Financial Regulation and Compliance, 25(4), 422–434. https://doi.org/10.1108/JFRC-01-2017-0001

Ogunjide, J. O. (2025). Biometric Authentication and Fraud Detection in Fintech. International Journal of Economics, Business and Management, 11(4), 183–189.https://doi.org/10.56201/ijebm. vol.11.no4.2025 .

Olanrewaju, A. G. (2025). Artificial intelligence in financial markets: Optimizing risk management, portfolio allocation, and algorithmic trading. International Journal of Research Publication and Reviews, 6, 8855–8870.https://doi.org/10.55248/gengpi .2025.6.4.2710

Onesi-Ozigagun, O., Ololade, Y. J., Eyo-Udo, N. L., & Ogundipe, D. O. (2024). AI-driven biometrics for secure fintech: Pioneering safety and trust. International Journal of Engineering Research Updates, 6(2), 1–12.https://doi.org/10.53430/ijeru.2024.6.2 .0023

Pan, E. (2024). Machine learning in financial transaction fraud detection and prevention. Transactions on Economics, Business and Management Research, 5, 243–249.https://doi.org/10.62051/16r3aa10

Pazouki, S., Jamshidi, M. B., Jalali, M., & Tafreshi, A. (2025). The integration of big data in FinTech: Review of enhancing financial services through advanced technologies. World Journal of Advanced Research and Reviews, 25(1), 546–556. https://doi.org/10 .30574/wjarr.2025.25.1.0060

Pearson, T. A., & Singleton, T. W. (2008). Fraud and forensic accounting in the digital environment. Issues in Accounting Education, 23(4), 545–559. https://doi.org/10. 2308/iace.2008.23.4.545

R Project. (n.d.). *An introduction to isolation forests (isotree package vignette)*. The R Project for Statistical Computing. Retrieved from https://cloud.r-project.org/web/ packages/isotree/vignettes/An_Introduction_to_Isolation_Forests.html

Russell, S. J., & Norvig, P. (1995). Artificial intelligence: A modern approach (3rd ed.). Pearson.

Schiavone, E., Ceccarelli, A., Carvalho, A., & Bondavalli, A. (2019). Design, implementation, and assessment of a usable multi-biometric continuous authentication system. International Journal of Critical Computer-Based Systems, 9(3), 215–247. https://doi.org/10.1504/IJCCBS.2019.104490

Schmitt, M., & Flechais, I. (2024). Digital deception: Generative artificial intelligence in social engineering and phishing. Artificial Intelligence Review, 57(12), 324. https://doi.org/10.1007/s 10462-024-10735-5

Sharma, V., Pandey, B., & Kumar, V. (2016). Importance of big data in financial fraud detection. International Journal of Automation and Logistics, 2(4), 332–348. https://doi.org/10.1504/IJAL. 2016.080339

Simplilearn. (n.d.). *K-means clustering algorithm*. Simplilearn. Retrieved from https:// www.simplilearn.com/tutorials/machine-learning-tutorial/k-means-clustering-

SymphonyAI. (2024a). Sensa Copilot. https://www.symphonyai.com/glossary/financial-services/sensa-copilot/

Timoney, M. (2025, April 17). Synthetic identity fraud continues to expand—and Gen AI is making it worse. Federal Reserve Bank of Boston. Retrieved from https://www .bostonfed.org/news-and-events/news/2025/04/synthetic-identity-fraud-financial-fraud-expanding-because-of-generative-artificial-intelligence. aspx

Udeh, E. O., Amajuoyi, P., Adeusi, K. B., & Scott, A. O. (2024). The role of big data in detecting and preventing financial fraud in digital transactions. World Journal of Advanced Research and Reviews, 22(2), 1746–1760. https://doi.org/10.30574 /wjarr.2024.22.2.1575

Vacca, John R. Biometric technologies and verification systems. Elsevier, 2007.

Wenbo Guo, Yujin Potter, Tianneng Shi, Zhun Wang, Andy Zhang, and Dawn Song, *SoK: Frontier AI’s Impact on the Cybersecurity Landscape*, arXiv preprint arXiv:2504.05408 (2025), https://arxiv.org/abs/2504.05408.

Westerlund, M. (2019). The emergence of deepfake technology: A review. Technology Innovation Management Review, 9(11), 39–52.

Zhang, X. (2024). Machine Learning Insights into Digital Payment Behaviors and Fraud Prediction. Applied and Computational Engineering, 77(1), 203–209. https://doi.org/ 10.54254/2755-2721/77/2024MA0066