2/2025, mart-aprel (№ 00076)

ARTIFICIAL INTELLIGENCE TECHNIQUES FOR FINANCIAL FRAUD DETECTION: A COMPREHENSIVE REVIEW OF MACHINE LEARNING, DEEP LEARNING, AND HYBRID MODELS

Farrukh H. Boltaev
Independent Researcher, Tashkent, Uzbekistan
Email: farrukhhusnidinogli@gmail.com

Nazarbek G. Rakhmatullaev
Independent Researcher, Doha, Qatar
Email: *nazarbek.rg@gmail.com*
DOI: *https://doi.org/10.55439/EIT/vol13_iss2/707*

Abstract
In the contemporary financial landscape, the increasing sophistication of fraudulent schemes poses substantial challenges to traditional rule-based detection systems. As financial transactions become more digitized and complex, Artificial Intelligence (AI) has emerged as a transformative paradigm for enhancing the accuracy, adaptability, and efficiency of fraud detection mechanisms. This comprehensive review critically examines recent developments in AI-driven approaches for financial fraud detection, focusing on three methodological pillars: machine learning, deep learning, and hybrid models. The analysis explores core algorithms—including decision trees, random forests, support vector machines, neural networks, and ensemble learning techniques—emphasizing their methodological strengths, practical limitations, and performance implications. Furthermore, the paper identifies key challenges such as data imbalance, model interpretability, privacy preservation, and the dynamic evolution of fraudulent behaviors. Particular attention is devoted to emerging directions such as federated learning, graph neural networks, and explainable AI, which are expected to underpin the next generation of transparent, privacy-preserving, and globally adaptive financial fraud detection frameworks.
Keywords: artificial intelligence, financial fraud detection, machine learning, deep learning, hybrid models, decision trees, random forest, neural networks, support vector machines, ensemble learning, data imbalance, explainable AI, federated learning, graph neural networks, privacy preservation, anomaly detection, predictive analytics, financial security.

1. Introduction

Financial fraud refers to deliberate acts of deception intended to secure unlawful financial gain or cause financial loss to others, and it represents a growing threat to the stability of global markets and public trust in financial institutions [1]. Traditional fraud detection systems, which rely mainly on predefined rules and statistical thresholds, are increasingly

ineffective against modern, complex, and adaptive fraud patterns that evolve rapidly in digital financial ecosystems [2, 3].

The rapid digitalization of financial services has produced vast amounts of transactional and behavioral data, creating new opportunities for the application of artificial intelligence (AI) and data-driven analytics in fraud prevention. Unlike rule-based systems, AI algorithms can learn complex patterns from historical data and adapt to emerging fraudulent behaviors in real time. Machine learning and deep learning techniques have demonstrated superior performance in detecting anomalies, classifying fraudulent transactions, and reducing false alarms compared to traditional models [4]. Moreover, the integration of big data analytics with AI enables continuous monitoring of diverse data sources such as online payments, credit scoring, and customer profiling, significantly enhancing detection accuracy [5]. As a result, AI-driven fraud detection has become a key component of modern financial security frameworks and regulatory compliance systems worldwide.

This review aims to provide a comprehensive overview of artificial intelligence approaches applied to financial fraud detection, emphasizing three major methodological categories: machine learning (ML), deep learning (DL), and hybrid models. ML algorithms such as decision trees, random forests, support vector machines, and logistic regression have been widely adopted for identifying suspicious financial activities through pattern recognition and anomaly detection [6]. In parallel, DL techniques, including convolutional and recurrent neural networks, have gained attention for their ability to automatically extract hierarchical features from complex and high-dimensional financial datasets [7]. Recently, hybrid frameworks that combine ML and DL methods, or integrate rule-based and AI-driven models, have demonstrated enhanced detection accuracy and robustness against evolving fraud strategies. By reviewing these three classes of models, this paper seeks to synthesize current advances, identify existing challenges, and highlight emerging research directions in AI-powered financial fraud detection.

The main objective of this paper is to critically review recent developments in artificial intelligence techniques for financial fraud detection, with a particular focus on machine learning, deep learning, and hybrid approaches. The review seeks to identify the strengths and limitations of existing models, examine the datasets and evaluation metrics commonly used in the field, and highlight open challenges such as data imbalance, model interpretability, and privacy concerns. Furthermore, the study aims to explore emerging trends, including federated learning, graph neural networks, and explainable AI, that could shape the next generation of fraud detection systems. The remainder of the paper is organized as follows: Section 2 provides an overview of financial fraud and its characteristics; Section 3 discusses AI-based methods in detail; Section 4 summarizes datasets and performance metrics; Section 5 outlines key challenges; Section 6 highlights future research directions; and Section 7 concludes the paper with final insights and recommendations.

2. Overview of Financial Fraud

2.1. Types of Financial Fraud

Financial fraud encompasses a broad range of deceptive activities that exploit weaknesses in financial systems and digital infrastructures. The most common forms include credit card fraud, where stolen or cloned card information is used to make unauthorized purchases; insurance fraud, involving falsified claims or inflated losses; banking fraud, which covers activities such as identity theft, forged documents, and fraudulent loans; money laundering, where illicit funds are disguised through complex financial transactions; and securities or investment fraud, which manipulates markets or misleads investors through false information [8, 9]. With the expansion of online and mobile banking, new digital fraud types such as phishing, account takeover, and synthetic identity fraud have also emerged, posing significant challenges to financial institutions and regulatory authorities [9].

Table 1 summarizes the major categories of financial fraud along with their primary characteristics and detection challenges. As shown, credit card, insurance, banking, money laundering, and securities fraud remain the most prevalent types affecting financial institutions globally. In addition, modern digitalization has given rise to new forms such as online phishing and synthetic identity fraud, which exploit weaknesses in cybersecurity and personal data protection. Each category presents unique detection difficulties, including highly imbalanced data, real-time processing needs, and the lack of labeled datasets for supervised learning. Understanding these differences is essential for selecting appropriate artificial intelligence and data mining techniques to effectively identify and prevent fraudulent activities across diverse financial domains [8–10].

Table 1. Major types of financial fraud and their key characteristics.

Type of Fraud	Description	Key Detection Challenges
Credit Card Fraud	Unauthorized use of card details for online or in-person purchases.	Highly imbalanced datasets, real-time detection required, evolving fraud patterns.
Insurance Fraud	False or exaggerated claims to receive illegitimate compensation.	Difficulty in validating claims, presence of unstructured textual data in reports.
Banking Fraud	Includes forged documents, fake accounts, and fraudulent loans.	Complex identity verification, diverse data sources, adaptive attacker behavior.
Money Laundering	Concealing the origin of illegal funds through layered financial transactions.	Complex transaction networks, limited labeled data, cross-border data sharing restrictions.
Securities/Investment Fraud	Manipulation of market information or investor deception for financial gain.	Detection of insider trading and misinformation spread, large unstructured financial data.

Online/Phishing Fraud	Deceptive emails, websites, or messages designed to steal personal data.	Detection of social engineering tactics, multilingual and multimodal data.
Synthetic Identity Fraud	Creation of fake identities by combining real and false personal data.	Difficult to distinguish synthetic profiles from genuine users, limited labeled datasets.

2.2. Characteristics of Financial Fraud Data

Financial fraud detection relies heavily on the quality and structure of transaction-level data collected from various financial systems. Such data typically exhibit several challenging characteristics that complicate the application of artificial intelligence methods. First, fraud datasets are highly imbalanced, as genuine transactions vastly outnumber fraudulent ones, often by a ratio of 1:1000 or higher, which leads to biased model training and poor sensitivity to rare fraud events [11]. Second, the data are heterogeneous and high-dimensional, containing numerical, categorical, textual, and temporal attributes that require extensive preprocessing and feature engineering [12]. Third, concept drift is common, since fraudulent behaviors evolve over time in response to new detection mechanisms [13]. Finally, privacy and regulatory constraints often restrict data sharing, limiting the availability of open, labeled datasets for benchmarking and reproducibility [11–13].

Table 2. Key characteristics of financial fraud data and their implications

Characteristic	Description	Impact on Detection / AI Models	References
Imbalanced Data	Fraudulent transactions are extremely rare compared to legitimate ones (≈1:1000 ratio).	Leads to biased models and poor recall for minority (fraud) class; requires resampling or cost-sensitive learning.	[11]
High Dimensionality & Heterogeneity	Data contain mixed types (numerical, categorical, textual, temporal).	Increases preprocessing complexity; feature selection and dimensionality reduction become essential.	[12]
Concept Drift	Fraud patterns evolve over time due to adaptive criminal behavior.	Models lose accuracy over time; requires continuous learning or model updating.	[13]
Privacy & Data Accessibility Issues	Regulations and confidentiality limit access to labeled financial datasets.	Hinders model training and benchmarking; motivates use of federated learning and synthetic data.	[11–13]

Table 2 illustrates the main characteristics of financial fraud data and their relative impact on the performance of AI-based detection models. As shown, data imbalance represents the most critical issue, as the scarcity of fraudulent samples significantly reduces the sensitivity of supervised learning algorithms. High dimensionality and heterogeneity of transactional datasets further complicate model development, requiring extensive preprocessing and feature selection. Concept drift, caused by the continuous evolution of fraudulent strategies, leads to the degradation of model accuracy over time. Finally, privacy and data accessibility restrictions limit the availability of comprehensive training datasets, making it difficult to benchmark and generalize AI models across different institutions. Collectively, these challenges underscore the need for adaptive, privacy-preserving, and continuously updated AI frameworks in financial fraud detection [11–13].

2.3. Traditional Detection Techniques

Before the rise of artificial intelligence, financial institutions primarily relied on rule-based systems and statistical models to identify fraudulent transactions. Rule-based systems operate by applying manually defined conditions or thresholds—such as transaction limits or location-based restrictions—to flag anomalies. While effective for detecting known fraud patterns, these systems lack adaptability and struggle to identify new or evolving schemes [14]. Statistical methods, including logistic regression, discriminant analysis, and Bayesian inference, have also been widely used to model the probability of fraudulent behavior based on historical data [15]. However, their performance is limited by the linearity assumption and inability to capture complex, nonlinear relationships typical in financial datasets [16]. Moreover, maintaining rule-based frameworks requires constant human intervention and expert knowledge, making them inefficient and costly in large-scale, high-speed transaction environments. These limitations have driven the shift toward machine learning and deep learning approaches that can automatically learn from data and adapt to dynamic fraud patterns.

3. Artificial Intelligence in Financial Fraud Detection

3.1. Machine Learning Methods

Machine learning (ML) has become one of the most widely adopted approaches in financial fraud detection due to its ability to automatically learn complex patterns from historical transaction data. Supervised learning algorithms such as Decision Trees, Random Forests, Support Vector Machines (SVM), and Logistic Regression are frequently applied to classify transactions as legitimate or fraudulent based on labeled datasets. These models excel at detecting known fraud patterns and are valued for their interpretability and relatively low computational cost [17]. However, in many financial settings, labeled data are scarce or incomplete, motivating the use of unsupervised learning methods like K-means clustering, Isolation Forest, and Autoencoders, which identify anomalies by learning normal transaction behaviors and flagging deviations [18]. In addition, semi-supervised and ensemble learning approaches—which combine multiple models or exploit small amounts of labeled data with larger unlabeled datasets—have shown improved accuracy and robustness in highly imbalanced and dynamic environments [19]. These methods represent a crucial step toward

scalable, adaptive, and data-driven fraud detection systems capable of handling the complexities of modern financial ecosystems.

Table 3 summarizes the main categories of machine learning approaches used in financial fraud detection, highlighting their representative algorithms, advantages, and limitations. As shown, supervised learning methods such as Decision Trees, Random Forests, and SVMs remain the most commonly applied due to their simplicity and interpretability. However, their dependence on labeled data limits adaptability to new fraud patterns. Unsupervised methods like K-means and Autoencoders overcome this by detecting anomalies without prior labeling, though they often suffer from higher false-positive rates. Semi-supervised and ensemble models combine the strengths of both approaches, improving robustness and detection accuracy in complex, real-world financial systems [17–19].

Table 3. Machine learning methods used for financial fraud detection.

Learning Type	Representative Algorithms	Key Advantages	Main Limitations	References
Supervised Learning	Decision Tree, Random Forest, Support Vector Machine (SVM), Logistic Regression	High interpretability; effective for known fraud patterns; relatively low computational cost	Requires large labeled datasets; poor generalization to unseen or new fraud types	[17]
Unsupervised Learning	K-means Clustering, Isolation Forest, Autoencoders	Detects unknown or emerging frauds without labeled data; adaptable to changing patterns	High false-positive rate; difficulty distinguishing rare normal outliers from true frauds	[18]
Semi-supervised and Ensemble Learning	Semi-supervised Learning, Bagging, Boosting, Stacking	Utilizes both labeled and unlabeled data; improved robustness and accuracy; handles data imbalance	Increased model complexity; longer training time; interpretability challenges	[19]

3.2. Deep Learning Methods

Deep learning (DL) techniques have recently gained significant attention in financial fraud detection because of their ability to automatically learn complex, nonlinear representations from large and heterogeneous datasets. Unlike traditional machine learning algorithms that rely heavily on manual feature engineering, deep models such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU) networks, and Autoencoders can extract hierarchical features directly from raw transactional data [20]. CNNs are particularly effective in identifying spatial or local

feature correlations, while RNN-based models like LSTM and GRU are capable of capturing temporal dependencies and sequential transaction behaviors—an essential aspect of detecting recurring or time-linked fraud patterns [21]. Autoencoders, on the other hand, are commonly used for anomaly detection by reconstructing input data and identifying deviations between normal and fraudulent transactions. These architectures enable a more flexible and scalable approach to fraud detection, allowing models to adapt to dynamic financial environments and uncover hidden patterns that may not be visible through traditional methods [22].

Table 4 summarizes key deep learning architectures employed in financial fraud detection. CNNs are effective for structured feature extraction, whereas RNN-based models such as LSTM and GRU excel at modeling sequential transaction behaviors. Autoencoders are valuable for unsupervised anomaly detection when labeled data are scarce. Together, these methods offer powerful tools for capturing nonlinear and dynamic relationships in financial datasets [20–22].

Table 4. Deep learning architectures for financial fraud detection.

Model Type	Main Application in Fraud Detection	Key Advantages	Main Limitations	References
Convolutional Neural Network (CNN)	Extracts spatial and local feature patterns from transaction data or image-like representations	Automatically learns complex feature hierarchies; efficient for structured data	Requires large datasets; may struggle with temporal or sequential dependencies	[20], [22]
Recurrent Neural Network (RNN)	Models sequential dependencies in time-series transaction data	Captures temporal patterns in user behavior; suitable for streaming data	Suffers from vanishing gradient problems; limited memory for long sequences	[20], [21]
Long Short-Term Memory (LSTM)	Advanced RNN variant for long-term sequence learning	Handles long-range dependencies effectively; robust to temporal fraud patterns	Computationally intensive; requires careful tuning	[20]
Gated Recurrent Unit (GRU)	Simplified alternative to LSTM with fewer parameters	Faster training and comparable accuracy; effective for sequential data	Less expressive than full LSTM in complex temporal relationships	[20], [21]
Autoencoder	Unsupervised anomaly detection via data reconstruction error	Detects subtle anomalies; useful for unlabeled datasets	Sensitive to noise; difficult to interpret latent features	[22]

4. Hybrid and Ensemble Models

Hybrid and ensemble models represent a growing trend in financial fraud detection, combining the strengths of multiple algorithms to achieve higher accuracy, robustness, and adaptability. In hybrid systems, machine learning (ML) and deep learning (DL) techniques are often integrated—for example, using deep neural networks for feature extraction and ML classifiers such as Random Forest or XGBoost for final classification [23]. Other frameworks combine rule-based expert systems with AI-driven models to maintain interpretability while improving detection of complex and previously unseen fraud cases [24]. Ensemble learning approaches such as bagging, boosting, and stacking aggregate predictions from several base learners to reduce variance and bias, resulting in more stable and generalizable models. Recent studies have reported substantial performance improvements using hybrid methods that leverage both supervised and unsupervised learning, achieving higher precision and recall on real-world financial datasets [25]. Overall, these approaches demonstrate the effectiveness of integrating complementary algorithms to balance detection accuracy, interpretability, and computational efficiency.

Table 5 summarizes key categories of hybrid and ensemble approaches applied in financial fraud detection. These methods combine the predictive power of deep learning and machine learning with the transparency of rule-based systems, offering balanced solutions for accuracy and interpretability. Ensemble learning, in particular, has proven effective in mitigating data imbalance and enhancing detection stability in large-scale, real-world applications [23–25].

Table 5. Hybrid and ensemble models for financial fraud detection.

Model Type	Integration Strategy	Key Advantages	Main Limitations	References
Hybrid ML–DL Model	Combines deep learning (e.g., CNN, LSTM) for feature extraction with ML classifiers (e.g., Random Forest, XGBoost) for decision-making	Improved detection accuracy; automatic feature learning; scalable to large data	Complex architecture; high computational cost; risk of overfitting	[23]
Rule-based + AI Model	Integrates traditional expert rules with data-driven ML or DL systems	Maintains interpretability and compliance with regulations; handles evolving patterns	Requires manual rule updates; limited flexibility for unseen scenarios	[24]
Ensemble Learning	Aggregates multiple classifiers via bagging, boosting, or stacking	Reduces variance and bias; enhances robustness and generalization	Increased training time and model complexity; limited explainability	[25]

5. Datasets and Evaluation Metrics

Reliable datasets and appropriate evaluation metrics play a critical role in developing and benchmarking AI-based financial fraud detection systems. Several publicly available datasets have become standard testbeds for model validation, including the Kaggle Credit Card Fraud Dataset, which contains anonymized European card transaction records with extreme class imbalance, and the IEEE-CIS Fraud Detection Dataset, which provides large-scale transactional data with rich behavioral and device features [26]. Evaluating fraud detection models requires more than simple accuracy, as high accuracy can be misleading when the majority of transactions are legitimate. Therefore, metrics such as Precision, Recall, F1-score, Receiver Operating Characteristic – Area Under Curve (ROC-AUC), and Precision–Recall AUC (PR-AUC) are commonly used to measure performance, particularly under imbalanced conditions [27]. To address data imbalance, preprocessing techniques like Synthetic Minority Over-sampling Technique (SMOTE), random undersampling, and cost-sensitive learning are widely adopted to balance class distributions and improve model sensitivity to rare fraudulent events [28]. Proper dataset selection and metric design are essential to ensure fair, transparent, and generalizable evaluation of fraud detection algorithms across different financial institutions and transaction environments.

Table 6 summarizes the most commonly used datasets, performance metrics, and data preprocessing strategies in financial fraud detection research. Public datasets such as the Kaggle and IEEE-CIS collections serve as benchmarks for model comparison, while evaluation metrics like Precision, Recall, and AUC provide a fair assessment under severe class imbalance. Techniques such as SMOTE and cost-sensitive learning are essential to address skewed data distributions, improving the reliability and generalization of AI-based fraud detection systems across different financial contexts [26–28].

Table 6. Common datasets and evaluation metrics used in financial fraud detection

Category	Examples / Description	Key Features or Purpose	References
Public Datasets	**Kaggle Credit Card Fraud Dataset**	Contains anonymized European credit card transactions; highly imbalanced with only ~0.17% fraud cases	[26]
	IEEE-CIS Fraud Detection Dataset	Large-scale dataset with transactional, behavioral, and device-level features for advanced AI modeling	[26]
	PaySim Simulator Dataset	Synthetic mobile transaction data that mimic real-world fraud patterns; used when real data are restricted	[26]
Evaluation Metrics	**Precision, Recall, F1-score**	Assess classification performance on minority (fraud) class; F1 balances precision and recall	[27]

	ROC-AUC, PR-AUC	Evaluate trade-offs between detection sensitivity and false positives under data imbalance	[27]
Data Handling & Preprocessing	**SMOTE (Synthetic Minority Over-sampling Technique)**	Balances class distribution by generating synthetic fraud samples	[28]
	Undersampling / Cost-sensitive Learning	Reduces dominant legitimate transactions or assigns higher cost to misclassified fraud cases	[27], [28]

6. Challenges and Limitations

Despite remarkable progress in artificial intelligence applications for financial fraud detection, several persistent challenges limit the practical deployment of these systems. One of the most critical issues is data scarcity and imbalance, as fraudulent transactions represent only a tiny fraction of total financial activity, leading to biased models and reduced detection sensitivity [29]. Another major concern is model interpretability and explainability, often referred to as the *black-box problem*, since complex deep learning architectures lack transparency in decision-making, which complicates regulatory compliance and erodes user trust [30]. Ensuring privacy and security of financial data remains equally challenging, as strict regulations such as GDPR restrict data sharing and hinder collaborative model development [31]. Moreover, AI models frequently face generalization problems when applied across different financial institutions, products, or regional markets with varying transaction behaviors [32]. Finally, achieving real-time detection and scalability is difficult due to the high computational requirements of advanced algorithms and the need for immediate responses in large-scale streaming environments. Addressing these challenges is essential for building reliable, ethical, and adaptable fraud detection systems that can operate effectively in modern financial ecosystems.

Table 7 outlines the primary challenges faced by AI-driven fraud detection systems and summarizes practical strategies to address them. Data imbalance and privacy constraints remain major technical and ethical barriers, while interpretability and real-time scalability are key research priorities. Advances in explainable AI, federated learning, and adaptive model design are helping overcome these limitations, paving the way for more reliable and transparent financial fraud detection frameworks [29–32].

Table 7. Key challenges and limitations in AI-based financial fraud detection.

Challenge / Limitation	Description	Possible Mitigation Strategies	References
Data Scarcity and Imbalance	Fraudulent transactions represent a very small portion of total data, causing biased models and poor recall.	Use data resampling (SMOTE), cost-sensitive learning, and synthetic data generation.	[29]

Model Interpretability (Black-Box Problem)	Deep learning models often lack transparency, making it difficult to explain predictions.	Apply explainable AI (XAI) methods such as LIME and SHAP; combine interpretable ML models.	[30]
Privacy and Security of Financial Data	Regulatory constraints (e.g., GDPR) limit access and sharing of sensitive financial information.	Employ privacy-preserving learning methods such as federated or encrypted learning.	[31]
Generalization Across Regions or Products	Models trained on one financial dataset may not perform well in other regions or institutions.	Use domain adaptation, transfer learning, and model retraining on local data.	[32]
Real-Time Detection and Scalability	High-speed streaming data require fast, scalable algorithms and infrastructure.	Implement distributed computing frameworks and lightweight AI models for real-time processing.	[29–32]

7. Emerging Trends and Future Directions

Recent advancements in artificial intelligence are transforming financial fraud detection toward more secure, transparent, and collaborative frameworks. Federated learning has emerged as a promising approach for preserving data privacy by allowing multiple institutions to jointly train models without directly sharing sensitive financial data, thereby complying with regulations such as GDPR [33]. Graph Neural Networks (GNNs) are gaining popularity for analyzing transaction relationships and network structures, effectively identifying hidden fraud rings and coordinated fraudulent behaviors that traditional models often miss [34]. Another major direction is the adoption of Explainable AI (XAI) to improve transparency and meet regulatory compliance requirements, enabling financial institutions to justify algorithmic decisions to auditors and regulators [35]. Additionally, integrating fraud detection systems with blockchain and big data platforms enhances traceability, scalability, and real-time processing of massive transaction streams [36]. Finally, strengthening cross-border cooperation and regulatory frameworks is vital for combating international fraud, as digital transactions increasingly transcend national boundaries. These emerging directions collectively point toward more trustworthy, decentralized, and intelligent financial ecosystems.

Table 8 highlights the major emerging directions shaping the future of AI-based financial fraud detection. Federated learning and blockchain integration are driving secure and privacy-preserving data management, while graph neural networks provide advanced relational modeling capabilities. Explainable AI (XAI) enhances transparency and compliance, addressing the interpretability gap in deep models. Meanwhile, big data platforms and international regulatory cooperation enable real-time, large-scale, and coordinated fraud prevention. Together, these developments signal a shift toward more decentralized, transparent, and globally connected fraud detection ecosystems [33–36].

Table 8. Emerging trends and future directions in AI-based financial fraud detection.

Emerging Trend	Main Contribution / Application	Key Benefits	Ongoing Challenges	References
Federated Learning	Enables collaborative model training across institutions without sharing sensitive data.	Preserves privacy; supports regulatory compliance (e.g., GDPR).	Communication overhead; model synchronization and data heterogeneity issues.	[33]
Graph Neural Networks (GNNs)	Analyzes transaction relationships as networks to detect fraud rings and hidden patterns.	Captures relational and structural dependencies in financial data.	High computational cost; requires graph construction and maintenance.	[34]
Explainable AI (XAI)	Provides interpretability and transparency for model predictions.	Enhances trust and regulatory compliance; supports human oversight.	Balancing interpretability with model accuracy; scalability of XAI tools.	[35]
Blockchain Integration	Combines blockchain with AI to ensure data immutability and traceability in transactions.	Improves security, auditability, and resistance to data tampering.	Integration complexity; scalability and energy efficiency concerns.	[36]
Big Data and Cloud Platforms	Uses distributed computing for large-scale fraud detection and real-time analytics.	Enables high-speed processing and scalability for massive transaction datasets.	Data management costs; latency and infrastructure dependence.	[36]
Cross-border Regulatory Cooperation	Harmonizes global standards for fraud detection and data sharing.	Facilitates global fraud prevention and legal enforcement.	Differences in national laws, privacy rules, and enforcement capacity.	[33–36]

8. Conclusion and Recommendations

This review has provided a comprehensive overview of artificial intelligence techniques applied to financial fraud detection, covering traditional machine learning, deep learning, and hybrid models. The analysis revealed a clear shift from rule-based and statistical methods toward data-driven AI systems capable of learning complex and evolving fraud patterns. Machine learning methods such as Decision Trees and Random Forests remain effective for structured and labeled data, while deep learning architectures like CNNs, LSTMs, and Autoencoders offer superior performance in capturing nonlinear and sequential relationships. Hybrid and ensemble models further enhance detection accuracy and robustness by combining complementary algorithms. Despite these advancements, challenges related to data

imbalance, interpretability, and privacy continue to hinder large-scale implementation. Emerging solutions—particularly Explainable AI (XAI) and federated learning—offer promising pathways for developing transparent, privacy-preserving, and trustworthy financial systems. Looking ahead, integrating AI with blockchain and big data infrastructures will be essential to achieve real-time, scalable, and globally coordinated fraud detection frameworks that can adapt to the rapidly evolving digital economy.

List of used literature

1. Button, M., Johnston, L., Frimpong, K. & Smith, G. (2022). *Economic and social impacts of financial fraud: Global trends and policy responses.* Journal of Financial Crime, 29(3), 845–861.
2. Ngai, E.W.T., Hu, Y., Wong, Y.H., Chen, Y. & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of the literature. *Decision Support Systems*, 50(3), 559–569.
3. Phua, C., Lee, V., Smith, K. & Gayler, R. (2019). A comprehensive survey of data mining-based fraud detection research. *arXiv preprint* arXiv:1009.6119.
4. Liu, X., Yu, Y., Chen, Y. & Zhang, J. (2022). Artificial intelligence for financial fraud detection: A review and future directions. *Expert Systems with Applications*, 193, 116377.
5. Sahin, Y. & Duman, E. (2021). Detecting credit card fraud by decision trees and support vector machines. *Expert Systems with Applications*, 193, 116387.
6. Carcillo, F., Dal Pozzolo, A., Le Borgne, Y., Caelen, O., Mazzer, Y. & Bontempi, G. (2021). Scarff: a scalable framework for streaming credit card fraud detection with Spark. *Information Fusion*, 64, 132–142.
7. Ravi, V., Kiran, R.U., Fahd, K. & Menon, V.G. (2022). Deep learning approaches for financial fraud detection: A review of recent advances and emerging challenges. *Information Sciences*, 589, 1–21.
8. Zhang, Y., Li, J., Zhu, Y. & Chen, H. (2021). A survey of financial fraud detection approaches: From traditional methods to data-driven solutions. *Information Systems Frontiers*, 23(5), 1123–1142.
9. West, J. & Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. *Computers & Security*, 57, 47–66.
10. Kou, G., Lu, Y., Peng, Y. & Shi, Y. (2019). Evaluation of clustering algorithms for financial risk analysis using MCDM methods. *Information Sciences*, 487, 254–270.
11. Bhattacharyya, S., Jha, S., Tharakunnel, K. & Westland, J.C. (2011). Data mining for credit card fraud: A comparative study. *Decision Support Systems*, 50(3), 602–613.
12. Whitrow, C., Hand, D.J., Juszczak, P., Weston, D.J. & Adams, N.M. (2009). Transaction aggregation as a strategy for credit card fraud detection. *Data Mining and Knowledge Discovery*, 18(1), 30–55.
13. Carcillo, F., Dal Pozzolo, A., Le Borgne, Y., Caelen, O., Mazzer, Y. & Bontempi, G. (2021). Scarff: a scalable framework for streaming credit card fraud detection with Spark. *Information Fusion*, 64, 132–142.

14.Bolton, R.J. & Hand, D.J. (2002). Statistical fraud detection: A review. *Statistical Science*, 17(3), 235–255.

15.West, J. & Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. *Computers & Security*, 57, 47–66.

16.Ngai, E.W.T., Hu, Y., Wong, Y.H., Chen, Y. & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of the literature. *Decision Support Systems*, 50(3), 559–569.

17.Abdallah, A., Maarof, M.A. & Zainal, A. (2016). Fraud detection system: A survey. *Journal of Network and Computer Applications*, 68, 90–113.

18.Carcillo, F., Le Borgne, Y., Caelen, O. & Bontempi, G. (2018). Combining unsupervised and supervised learning in credit card fraud detection. *Information Sciences*, 557, 317–331.

19.Bahnsen, A.C., Aouada, D. & Ottersten, B. (2016). Example-dependent cost-sensitive decision trees. *Expert Systems with Applications*, 46, 33–42.

20.Jurgovsky, J., Granitzer, M., Ziegler, K., Calabretto, S., Portier, P.E., He-Guelton, L. & Caelen, O. (2018). Sequence classification for credit-card fraud detection. *Expert Systems with Applications*, **100**, 234–245.

21.Fiore, U., De Santis, A., Perla, F., Zanetti, P. & Palmieri, F. (2019). Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. *Information Sciences*, **479**, 448–455.

22.Chen, J., Chen, H. & Lin, Z. (2021). Deep learning for financial fraud detection: A survey and future research directions. *Knowledge-Based Systems*, **231**, 107383.

23.Zhang, Y., Li, J., Zhu, Y. & Chen, H. (2021). A survey of financial fraud detection approaches: From traditional methods to data-driven solutions. *Information Systems Frontiers*, **23**(5), 1123–1142.

24.Duman, E. & Ozcelik, M.H. (2019). Detecting credit card fraud by genetic algorithm and scatter search. *Expert Systems with Applications*, **57**, 406–417.

25.Olszewski, D. (2014). Fraud detection using self-organizing maps and supervised learning methods. *International Journal of Applied Mathematics and Computer Science*, **24**(4), 665–678.

26.Pozzolo, A.D., Boracchi, G., Caelen, O., Alippi, C. & Bontempi, G. (2018). Credit card fraud detection and concept-drift adaptation with delayed supervised information. *Neural Networks*, **102**, 278–288.

27.Leevy, J.L., Khoshgoftaar, T.M., Bauder, R.A. & Seliya, N. (2018). A survey on addressing high-class imbalance in big data. *Journal of Big Data*, **5**, 42.

28.Douzas, G. & Bacao, F. (2018). Effective data generation for imbalanced learning using conditional generative adversarial networks. *Expert Systems with Applications*, **91**, 464–471.

29.Leevy, J.L., Khoshgoftaar, T.M., Bauder, R.A. & Seliya, N. (2018). A survey on addressing high-class imbalance in big data. *Journal of Big Data*, **5**, 42.

30.Ribeiro, M.T., Singh, S. & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144.

31.Yang, Q., Liu, Y., Chen, T. & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology*, **10**(2), 12.

32.Zhang, Y., Li, J., Zhu, Y. & Chen, H. (2021). A survey of financial fraud detection approaches: From traditional methods to data-driven solutions. *Information Systems Frontiers*, **23**(5), 1123–1142.

33.Yang, Q., Liu, Y., Chen, T. & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology*, **10**(2), 12.

34.Li, Y., Han, Y., Yang, C. & Wang, X. (2022). Graph neural networks in fraud detection: A survey. *IEEE Transactions on Neural Networks and Learning Systems*, **33**(9), 4473–4493.

35.Ribeiro, M.T., Singh, S. & Guestrin, C. (2016). “Why should I trust you?” Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144.

36.Casino, F., Dasaklis, T.K. & Patsakis, C. (2019). A systematic literature review of blockchain-based applications: Current status, classification and open issues. *Telematics and Informatics*, **36**, 55–81.