# A Review on Cryptocurrency Fraud Detection Techniques: Challenges, Solutions, and Perspectives

No Author Given

No Institute Given

**Abstract.** The rapid expansion of fraudulent behavior concerning the cryptocurrency ecosystem has underscored the necessity for intelligent detection frameworks. This paper focuses on a comparative review of modern approaches organized into three methodological categories: cutting-edge AI frameworks, traditional ML methods, and non-ML or heuristic approaches. In analyzing sixteen contributions, we assess the multidisciplinary gaps and challenges of each class in terms of model interpretability, data imbalance, high computational cost, overfitting, insufficient empirical evaluation, and generalizability to novel instances of fraud. To overcome these challenges, we propose the use of explainable AI tools such as SHAP and GNNExplainer, along with hybrid detection pipelines—combining deep learning, traditional models, and heuristic knowledge—data generation through transfer learning, and self-supervised frameworks. We also highlight the need for modular designs focused on system scalability and evolvability, as well as the integration of on-chain and off-chain data to improve situational awareness. Given the complexity of cryptocurrency fraud, we advocate for strategically adaptable and explainable detection frameworks that enhance performance and interpretability in the ever-changing environment of decentralized blockchain systems.

**Keywords:** Cryptocurrency Fraud · Artificial Intelligence · Machine Learning · Deep Learning · Graph Neural Networks · Hybrid Methodologies · Explainable AI · Heuristic Approaches.

## 1 Introduction

The cryptocurrency world has extremely enhanced and changed financial systems on a global scale. Bitcoin and Ethereum passed conventional banking systems, allowing new forms of online business to flourish by providing decentralization, transparency, and transactions across borders. However, pseudonymity, lack of central supervision, and immutability have also made room for many illegitimate activities [1]. Cryptocurrency-based fraud is on the rise and becoming increasingly sophisticated with money laundering, ponzi schemes, pump-and-dump schemes, phishing, smart contract hacking, and transaction hiding. The enormous amount of available information, the borderless nature of transactions, the ever-changing user behaviors, and the immense scale of operations make detecting and preventing these fraudulent activities extremely challenging [2]. To address these issues, many initiatives have started using artificial intelligence (AI) and machine learning (ML) technologies that attempt to identify indicators of potentially fraudulent activities. Recent works have looked at a wide range of approaches—classical supervised learning models like decision trees and support vector machines to more sophisticated architectures like deep neural networks, graph neural networks, and hybrid ensembles [3]. On the other hand, several heuristic or conceptual models have also been proposed, often based on domain-specific rules or indicators derived from the network. Although the results in this field are promising, there are still many gaps, such as the lack of standardized datasets, low model explainability, poor real-time detection, and inconsistent evaluation metrics.

The goal of this paper is to systematically evaluate and analyze *sixteen selected and impactful research publications* directed towards fraud detection within cryptocurrency using AI and its associated technologies. This paper is unique in its *categorization of crime studies into three primary groups* of detection approaches: advanced AI-based methods, classical machine learning techniques, and non-ML or heuristic frameworks, with less attention paid to the specific analytical models employed. Each selected article is analyzed along *nine key criteria*: (1) the type of cryptocurrency studied, (2) the type of fraud addressed, (3) the datasets used, (4) the methodological approach, (5) the AI or non-AI algorithms employed, (6) the performance metrics used, (7) real-time detection capabilities, (8) explainability of the models, and (9) the main limitations and challenges identified.

Our comparative analysis reveals several key findings. First, while advanced AI techniques often outperform classical models in terms of accuracy, they frequently lack transparency and are difficult to interpret, which poses a barrier to their adoption in regulatory contexts. Second, real-time fraud detection remains underdeveloped in most studies, despite being critical in practical applications. Third, there is a clear shortage of benchmark datasets and reproducible experimental settings, making it difficult to compare models fairly. Finally, hybrid and ensemble approaches appear to be a promising direction, balancing performance and generalization, but they often come with increased computational complexity. The remainder of this article is structured as follows. Section 2 describes the theoretical foundations of cryptocurrencies, fraud typologies, and AI-based detection. Section 3 presents the proposed classification framework and categorizes detection approaches identified in the literature. Section 4 compares the selected papers based on nine evaluation criteria, highlighting key methodological trends. Section 5.1discusses major challenges and 5.2 potential solution. Finally, Section 6 concludes with directions for future research.

## 2 Background on Crypto & Fraud Detection

Cryptocurrencies, such as Bitcoin and Ethereum, are digital currencies secured through cryptography. Satoshi Nakamoto released Bitcoin in 2009, whose value has increased through the years. Recognizable names for other cryptocurrencies include Ripple, Litecoin, and many others identified as altcoins [1]. Cryptocurrencies employ blockchains, which are decentralized digital ledgers that record transactions made from multiple computers. Blockchains ensure all transactions are safeguarded against modification or deletion, as well as remove the need for a trusted intermediary to oversee exchanges. The decentralized nature of cryptocurrencies and the absence of a central regulatory authority increase their vulnerability to illicit activities, including fraud. Their intrinsic characteristics, such as user anonymity and the irreversibility of transactions, also make them a fertile ground for malicious behavior [2], mainly:

- *Ponzi Schemes*: Profitable opportunities that are bound to fail. Returns stem from withdrawing new investors' contributions.
- *Pump-and-Dump Schemes*: Scams meant to drastically increase and decrease share prices of low-activity securities.
- *Phishing and Social Engineering*: The act of obtaining confidential information to access an account.
- *Money Laundering and Illicit Transactions*: Using crypto to disguise the origin of illegal funds.
- *Smart Contract Exploits*: Exploiting poorly written or malicious contracts in platforms like Ethereum.

Artificial intelligence (AI), particularly machine learning (ML), is proving invaluable in identifying cryptocurrency fraud. However, the speed and volatility of blockchain data make traditional systems, based on predefined rules, unsuitable. These AI technologies [4] are capable of:

- Detection of unusual and potentially suspicious transactions.
- Behavioral analysis to anticipate fraudulent activity.
- Adaptation to emerging and evolving fraud techniques.
- Interpretable justification of decisions using Explainable AI (XAI).

A systematic review of sixteen academic studies on AI-powered cryptocurrency fraud detection was conducted. Selection criteria included demonstrated relevance to cryptocurrency fraud, application of AI or machine learning methods, and transparent reporting of datasets, models, and evaluations. The selected articles covered different years, regions, and fraud types, ensuring a diverse and representative basis for analysis. In order to better understand the technological maturity and operational efficiency of automated fraud detection systems, we propose to classify the selected articles into three categories according to the nature and complexity of the techniques used and of their integration with AI:

- Class 1—Advanced AI-based approaches
- Class 2—Classical machine learning models
- Class 3—Conceptual, rule-based, or heuristic approaches

This classification provides a structured framework to evaluate the academic significance, technical advancement, and practical relevance of each approach. By distinguishing between advanced, classical, and non-AI-based methods, it reflects the progression of fraud detection techniques in terms of algorithmic complexity and AI integration. It also helps identify current limitations and guide the selection of approaches best suited to emerging challenges, such as real-time detection, adaptability to evolving fraud strategies, and interpretability.

## 3 Classification & Analysis Of Cryptocurrency Fraud

The specific and subtle differences between the three categories of methods identified for detecting cryptocurrency fraud lie in their algorithmic complexity, flexibility, and overall practicality.

### Class 1—Advanced AI-Based Approaches

The first class exploited advanced AI techniques, including deep learning, graph models, and ensemble methods, generally incorporated with explainability tools to address the complexity and dynamism of blockchain data. Kamisetty et al. [5] applied CNNs and RNNs to model cryptocurrency transactions, while Shayegan et al. [6] used graph-based anomaly detection to capture fraud via relationships between nodes. Elmougy and Liu [7] enhanced graphical models with XAI to increase transparency. Ensemble methods, as employed byNayyer et al. [8], combined multiple classifiers and SHAP to obtain robust and interpretable predictions. Finally, Walavalkar et al. [9] integrated token-pattern mining and isolation forests with classical machine learning to target Ethereum-specific anomalies. These studies presented the latest developments and techniques used to detect fraud in cryptocurrency and explained how the combination of advanced AI algorithms and interpretability could produce effective and innovative solutions. Table 1 summarizes these contributions and the specific techniques applied in each category.

| Year | Article | Techniques Used |
|---|---|---|
| 2021 | Kamisetty et al. [5] | Deep Learning (CNN, RNN) |
| 2022 | Shayegan et al. [6] | Graph-based anomaly detection |
| 2023 | Elmougy and Liu [7] | Graph ML, Explainability (XAI) |
| 2023 | Nayyer et al. [8] | Ensemble stacking, SHAP |
| 2024 | Walavalkar et al. [9] | ML, Token patterns, Isolation Forest |

**Table 1.** Deep Learning and Graph AI-Based Techniques in Class 1.

### Class 2—Classical Machine Learning Approaches

The second class focused on traditional machine learning approaches, which included algorithms such as random forest (RF), support vector machines (SVM), k-nearest neighbors (KNN), and boosting methods like XGBoost and AdaBoost. These techniques were computationally effective and comparatively interpretable, making them useful for practical applications. Bartoletti et al. [10] employed RF and KNN algorithms to detect Ponzi schemes on the Ethereum blockchain, while Ostapowicz and Żbikowski [11] applied RF and SVM to classify fraudulent behaviors based on transaction patterns. Boosting techniques were also utilized in Ashfaq et al. [12] andTripathy et al. [4], showing improved accuracy on imbalanced datasets. In the more recent study, Kumari [13] combined RF, SVM, and decision trees to identify Bitcoin fraud. Although effective on structured datasets, these models often struggled to generalize emerging fraud patterns due to their dependence on static features. Table 2 presents the main studies and techniques associated with this class.

| Year | Article | Techniques Used |
|---|---|---|
| 2018 | Bartoletti et al. [10] | Random Forest (RF), K-Nearest Neighbors (KNN) |
| 2019 | Ostapowicz and Żbikowski [11] | RF, Support Vector Machine (SVM) |
| 2022 | Ashfaq et al. [12] | XGBoost, RF |
| 2022 | Anthony et al. [14] | RF, KNN, Stochastic Gradient Descent |
| 2024 | Tripathy et al. [4] | AdaBoost, XGBoost |
| 2025 | Kumari [13] | RF, SVM, Decision Tree |

**Table 2.** Standard Machine Learning Techniques used in Class 2.

### Class 3—Non-ML, Conceptual, or Heuristic Approaches

The third class involved approaches that did not rely primarily on machine learning. They relied on rule-based systems, heuristics, simulation, or theoretical models often grounded in domain expertise. Chen et al. [15] proposed a rule-based classifier for identifying Ponzi schemes using predefined transaction patterns while La Morgia et al. [16] designed a heuristic algorithm to detect pump-and-dump activities in real time. Likewise, Aponte-Novoa et al. [17] employed game-theoretical simulations to analyze 51% attacks and Bello et al. [18] introduced a conceptual framework for fraud detection without relying on data-driven modeling. Though Bartoletti et al. [2] included partial ML elements, their contribution remained largely taxonomic and conceptual. While such models offered valuable insights and interpretable frameworks, they lacked scalability, adaptability, and predictive accuracy, especially in dynamic environments like cryptocurrency markets. Table 3 summarizes this class studies and techniques.

| Year | Article | Techniques used |
|---|---|---|
| 2018 | Chen et al. [15] | Rule-based classification |
| 2020 | La Morgia et al. [16] | Real-time detection algorithm heuristic |
| 2021 | Bartoletti et al. [2] | Taxonomy + partial ML |
| 2021 | Aponte-Novoa et al. [17] | Game theory (51% attack) |
| 2024 | Bello et al. [18] | Conceptual framework |

**Table 3.** Rule-Based and Simulation Techniques in Class 3.

This new framework provides a useful map for exploring the complex field of cryptocurrency fraud detection. By classifying the papers into three broad categories, we were able to highlight the strengths and weaknesses of each approach. Advanced AI methods typically adapt quickly and deliver impressive performance, but they're more like black boxes. Traditional machine learning falls somewhere in between, offering good speed and clearer logic, while heuristic methods remain simple and apply clear rules, even if they rarely evolve. The following sections build on this classification by exploring the most difficult challenges and unexplored questions still facing scientists and practitioners.

## 4 Comparison & Discussion

Now, we conduct a comparative analysis of sixteen selected studies on AI-driven cryptocurrency fraud detection. This comparison sheds light on recurring challenges and emerging solutions across the studied literature.

### 4.1 Comparison

First, this section presents a comparative overview of the selected studies on cryptocurrency fraud detection, focusing on their research scope and practical contributions. A comparison is given in Table 4 below, considering nine criteria: *Type of Fraud, Cryptocurrency, Dataset Used, Methodologies, Algorithms Applied, Evaluation Metrics, Real-Time Detection, Explainability*, and *Limitations Identified.*

- *Type of Fraud:* What type of fraud is analyzed? (Ponzi, pump-and-dump, ransomware, phishing, etc.)
- *Cryptocurrency:* Which cryptocurrency ecosystem is used? (Bitcoin, Ethereum, others crypto, smart contracts, etc.)
- *Dataset Used:* real, simulated, labeled, or enriched dataset used? Is it public?
- *Methodologies:* What methods are applied? (Machine learning, graph-based models, blockchain integration, boosting, etc.)
- *Algorithms Applied:* Which ML algorithms are employed? (Random Forest, SVM, XGBoost, LSTM, etc.)
- *Evaluation Metrics:* Which metrics are reported? (Accuracy, Precision, Recall, F1-score, ROC-AUC, etc.)
- *Real-Time Detection:* Does the approach allow real-time fraud detection? (Yes/No).
- *Explainability:* Are the model results interpretable or explained? (explainable AI techniques like feature importance and visualization).
- *Limitations:* What limitations are acknowledged? (dataset bias, generalization issues, real-time constraints, etc.)

Building on this comparison, we open a reflective discussion of broader implications and methodological perspectives.

### 4.2 Discussion

Second, the following section examines the results in light of the defined evaluation criteria. The objective is to highlight the main trends, strengths, and limitations of the literature, thus providing an overview of the current landscape and prospects of AI-based approaches to cryptocurrency fraud detection.

**Cryptocurrency Focus** The analyzed studies cover a wide range of blockchain platforms. Bitcoin remains the most frequently investigated cryptocurrency, as (Bartoletti et al. [10], Kamisetty et al. [5], Elmougy and Liu [7]), owing to its

**Table 4.** A Comparison of the Studied Cryptocurrency Fraud Detection Techniques

| **Article** | **Year** | **Crypto** | **Fraud Type** | **Data Type** | **Methodologies** | **Algorithms** | **Metrics** | **Real-Time** | **Explainability** | **Limitations** |
|---|---|---|---|---|---|---|---|---|---|---|
| Chen et al. [15] | 2018 | Ethereum Smart Contract | Ponzi schemes via smart contracts | Labeled smart contracts | Supervised ML; On-chain behavior analysis | Rule-based classification | Detection rate; False positive rate | No | Partially (Feature selection) | Scalability limits; lack learning component; static and offline |
| Bartoletti et al. [10] | 2018 | Bitcoin | Ponzi schemes | Labeled dataset of Ponzi | Supervised ML; On-chain behavior analysis (features) | Decision Tree, Random Forest, k-Nearest Neighbors | Precision, Recall, F1-score | No | Partially (Feature importance analysis) | Limited generalization; dataset limited; no real-time capability |
| Ostapowicz and Żbikowski [11] | 2019 | Ethereum | Account-level fraud (scams, phishing, Ponzi) | Labeled dataset of 300K+ Ethereum addresses | Supervised ML, Feature engineering from on-chain data | Random Forest, XGBoost, SVM, k-NN, MLP | Accuracy, Precision, Recall, F1-score | No | Partially (SHAP, Feature importance) | Class imbalance; no real-time adaptation; no off-chain data; potential label noise |
| La Morgia et al. [16] | 2020 | Various (Binance, YoBit ) | Pump and dump schemes | Real-time market data (buy orders) | Real-time detection algorithm based on market behavior | Custom real-time detection algorithm (non-ML) | Detection accuracy, F1-score, Detection time | Yes | Partially (behavioral signal detection) | Only observed pumps; no ML generalization; no explainability |
| Bartoletti et al. [2] | 2021 | General cryptocurrencies | Multi-type scams (phishing, Ponzi, ransomware, hybrid) | Address-reported scams, URL-reported scams from public sources | Literature review; Taxonomy creation; Supervised ML classification | Multi-label classifier (unspecified algorithm) | Precision, Recall, F1-score | No | Partially (taxonomy-based classification) | Incomplete data, no standard taxonomy, poor real-time use |
| Kamisetty et al. [5] | 2021 | Bitcoin | Various types (double-spending, phishing, laundering) | General Bitcoin transaction data | Deep learning models; feature extraction | ANN, CNN, RNN, Autoencoders | Precision, Recall, F1-score | Partially (real-time detection potential) | No (focus on model performance, limited explainability) | No real data, weak validation, no interpretability, no implementation |
| Aponte-Novoa et al. [17] | 2023 | Bitcoin | 51% attack (selfish mining) | Simulated data on mining behavior | Theoretical modeling, Simulation-based study of mining strategies | Game theory models, mining strategy analysis (selfish) | Theoretical profit ratio, attack success rate (simulation) | No | Lacks theory No XAI tools | Idealized setup; incomplete network modeling; no real-world tests |
| Anthony et al. [14] | 2022 | Ethereum | Anomaly detection (general fraud detection) | Ethereum transaction records (Blockchair dataset) | Supervised ML-based anomaly detection | Random Forest, KNN, SGD, GaussianNB | Accuracy, Precision, Recall, F1-score | No | Partially (feature-based decision boundaries) | Offline only; no real-time or cross-chain validation |
| Ashfaq et al. [12] | 2022 | Bitcoin | Transactions fraud (double-spending, Sybil attacks) | Synthetic, legitimate and fraudulent patterns | Supervised ML + Blockchain integration (smart contracts) | XGBoost, Random Forest | Accuracy, False Positive Rate, AUC | Yes | Partially (via smart contract security analysis) | Computational complexity; no real-time detection; Synthetic data |
| Shayegan et al. [6] | 2022 | Bitcoin | Crypto wallet fraud (collective anomaly detection) | Bitcoin transaction graph (blockchain data) | Graph-based anomaly detection; unsupervised clustering | DBSCAN, K-means | Precision, Recall, Detection rate | No | Limited (no XAI, black-box clustering results) | No labeled validation, poor scalability, no real-time deployment. |
| Elmougy and Liu [7] | 2023 | Bitcoin | Financial fraud, Anti-money laundering (AML) behaviors | Elliptic++, 203K transactions | Graph-based ML, Explainable AI | Random Forest, XGBoost, LSTM, Graph Analytics | Precision, F1-score, Recall | No | Yes (via graph structure and feature attribution) | Class imbalance; lack of real-time detection; limited off-chain context |
| Nayyer et al. [8] | 2023 | Bitcoin | Fraudulent Transactions | Labeled dataset of Bitcoin transactions | Supervised ML, Stacking, ADASYN, Tuning | Decision Tree, RF, Naive Bayes, KNN, Logistic Regression | Accuracy, Recall, F1-score, AUC-ROC | No | Partially (via SHAP feature explanations) | Labeled datasets; no real-time validation; limited scalability |
| Tripathy et al. [4] | 2024 | Ethereum | Transaction fraud (money laundering, illicit services) | Labeled dataset of Ethereum transactions | Supervised ML with feature engineering | Logistic Regression, Random Forest, KNN, AdaBoost, XGBoost | Accuracy, Precision, Recall, F1-score | No | No | no real-time detection; lacks off-chain context; overfitting; false positives |
| Walavalkar et al. [9] | 2024 | Ethereum | Token-based fraud (scams, rug pulls, illicit transfers) | Ethereum token (Etherscan API) | Token behavior analysis, ML classification, Graph-based anomaly detection | Random Forest, XGBoost, DBSCAN, Isolation Forest | Precision, Recall, F1-score, Accuracy | No | Partial (feature importance, token pattern insights) | Sparse labels; scaling issues; ERC-20 scope; adversarial risk |
| Bello et al. [18] | 2024 | General Cryptocurrencies | General financial fraud (unspecified) | Conceptual model (no empirical dataset) | Conceptual framework combining ML and blockchain | Logistic Regression, Neural Networks, Clustering | Theoretical only (no empirical metrics) | Yes (conceptual) | Framework only no validation | No real data, no experiments, no scalability evaluation. |
| Kumari [13] | 2025 | Bitcoin | Scam detection, Market manipulation | Dataset of Bitcoin wallet transactions in USA | Scam detection via supervised ML and behavior modeling | Random Forest, SVM, Decision Tree, Logistic Regression | Accuracy, Precision, Recall, F1-score | No | Partial (feature importance, behavior-based) | Data-scarce; not real-time; No generalization to trends |

widespread adoption and the public availability of its transaction data. Ethereum is also prominent in fraud studies, especially concerning smart contract frauds and token-based schemes. Fewer studies adopt a cross-chain perspective or address generic blockchain environments like (Bartoletti et al. [2], Bello et al. [18]), indicating the need for more generalized, blockchain-agnostic detection techniques.

**Type of Fraud.** There is a great variation in the typologies of fraud within the literature. There is a heavy concentration of Ponzi and phishing scams in (Chen et al. [15], Ostapowicz and Żbikowski [11], Bartoletti et al. [10]), as well as laundering, market manipulation, pump-and-dump, and double-spending attacks in (La Morgia et al. [16], Ashfaq et al. [12]). Emerging attack vectors like token scams, in Walavalkar et al. [9], and mining-based frauds such as selfish mining, in Aponte-Novoa et al. [17], has increased the sophistication of the threats. A few works, like (Bartoletti et al. [2], Elmougy and Liu [7]), have adopted a multi-fraud view with the goal to consolidate detection frameworks across different types of fraud.

**Datasets Used.** The choice of datasets significantly impacts the robustness and generalizability of the results. Studies utilizing labeled datasets, notably (Nayyer et al. [8], Ostapowicz and Żbikowski [11]), offer strong benchmarking potential, though class imbalance and label noise remain concerns. Others rely on synthetic data (Ashfaq et al. [12]), or simulated data (Aponte-Novoa et al. [17]), limiting their real-world applicability. Public blockchain data is widely used due to its availability, but few studies integrate off-chain data, which could improve contextual understanding of fraudulent behavior.

**Methodologies.** Supervised machine learning dominates the methodological landscape, particularly in fraud classification tasks, notably (Tripathy et al. [4], Kumari [13], Ostapowicz and Żbikowski [11], Anthony et al. [14]). Graph-based methods are increasingly prevalent in studies focusing on transaction networks and wallet behaviors, such as (Shayegan et al. [6], Elmougy and Liu [7]). A minority of works, including (Walavalkar et al. [9]), use unsupervised learning for anomaly detection or hybrid frameworks associating machine learning with domain-specific heuristics or blockchain analytics. Only one study, Bello et al. [18], proposes a conceptual model integrating AI and blockchain without empirical implementation.

**Algorithms Employed.** XGBoost, Support Vector Machines, and Random Forest are the most commonly used algorithms, all of which offer a good balance between interpretability and performance. Capturing temporal and nonlinear fraud patterns requires the use of deep learning models such as Convolutional Neural Network (CNNs), Recurrent Neural Network (RNNs), and autoencoders, as demonstrated by (Kamisetty et al. [5], Nayyer et al. [8]). When analyzing portfolio-level behavior, graph clustering algorithms such as DBSCAN and K-means are common, among them (Bello et al. [18], Shayegan et al. [6]). There is some variety in algorithms used, but ensemble models combined with hyperparameter optimization remain relatively underutilized in the current literature.

**Evaluation Metrics.** Most papers, including (Anthony et al. [14], Bartoletti et al. [10], Kamisetty et al. [5]), use basic classification evaluation criteria, such as AUC, F1-score, precision, recall, and accuracy. La Morgia et al. [16] investigate real-time detection and estimate the latency required for identifying fraudulent events. However, the lack of standardized metric reporting across studies complicates efforts to compare their results and assess relative performance. Very few works discuss model robustness or statistical significance, which are critical to assessing capability for generalization.

**Real-Time Capabilities.** A limited number of works, illustrated by (La Morgia et al. [16], Ashfaq et al. [12], Bello et al. [18]), support or simulate real-time fraud detection. Most frameworks focus on analyzing historical data, leaving real-time detection a major unsolved problem due to latency, throughput, and scaling challenges in blockchain networks. On the other hand, simulations and conceptual models demonstrate promising potential, although few have been operationalized in real-world or production environments.

**Explainability.** Explainable AI (XAI) remains underexplored, with only partial implementations via feature importance rankings or SHAP values. Many studies, notably (Ostapowicz and Żbikowski [11], Elmougy and Liu [7]), prioritize detection accuracy over interpretability, which poses barriers to regulatory adoption and user confidence. This is especially problematic in financial applications, where decisions often need to be auditable and legally defensible.

**Limitations Identified.** Some of the common limitations include class imbalance, as in (Elmougy and Liu [7]), lack of real-time implementation and deployment, as in (Chen et al. [15], Bartoletti et al. [10]), use of synthetic or limited datasets, as in (Ashfaq et al. [12]), and insufficient validation. Many models remain vulnerable to evolving fraud strategies because they rely on static features or patterns derived from outdated behaviors captured in historical data. Additionally, most studies ignore off-chain data sources like social media or regulatory data, which are increasingly important in capturing the context of fraud. The identified limitations reflect the current progress and ongoing challenges in AI-based fraud detection for cryptocurrency networks. Although supervised learning remains the dominant approach, there is a growing need for non-blockchain solutions, real-time interpretability, and scalable frameworks. These emerging efforts also include the integration of multimodal datasets—capturing on-chain and off-chain information—and the development of adaptive learning models. Addressing these gaps will not only advance academic research but also contribute to building more robust and resilient crypto-financial infrastructures.

## 5 Challenges & Solutions

By categorizing and comparing the latest studies on cryptocurrency fraud detection using AI, we highlight common weaknesses as well as new avenues for researchers to explore. These findings are structured into two overarching categories: key challenges and proposed solutions.

### 5.1 Challenges

Each of the three methodological classes presents distinct challenges and limitations in the context of crypto-fraud detection, as shown in Figure 1.

– Class 1 (Advanced AI-Based Methods): Although models such as graph neural networks and deep learning architectures, used notably by (Elmougy and

Liu [7], Shayegan et al. [6]) demonstrate high detection accuracy and adaptability, they often suffer from interpretability and computational complexity issues.Nayyer et al. [8] exploit ensemble learning and SHAP for explainability, but highlight the challenge of scaling to larger datasets. Additionally, Kamisetty et al. [5] highlight the need for large annotated datasets, which are often not available in real-world blockchain settings.

- Class 2 (Classical Machine Learning Approaches): Models such as Random Forest or XGBoost, employed in (Bartoletti et al. [10], Tripathy et al. [4]) provide more interpretable and computationally efficient solutions but suffer from overfitting and poor generalization. Ashfaq et al. [12] acknowledge that despite achieving strong performance on labeled data, the approach faces challenges with unseen fraud types due to the imbalance of the dataset and the lack of contextual information. Finally,Kumari [13] note that his model's performance deteriorates in highly volatile market conditions.
- Class 3 (Heuristic, Rule-Based, and Conceptual Methods): Works such as (Aponte-Novoa et al. [17], Chen et al. [15]) provide valuable theoretical insights but face limited adaptability and lack of automation. The La Morgia et al. [16] paper relies on hand-crafted heuristics to instantly detect pumping and drainage patterns but acknowledges the difficulty of generalizing them to new treatment models. Similarly, Bello et al. [18] proposes a conceptual framework without experimental validation, making its practical utility uncertain.

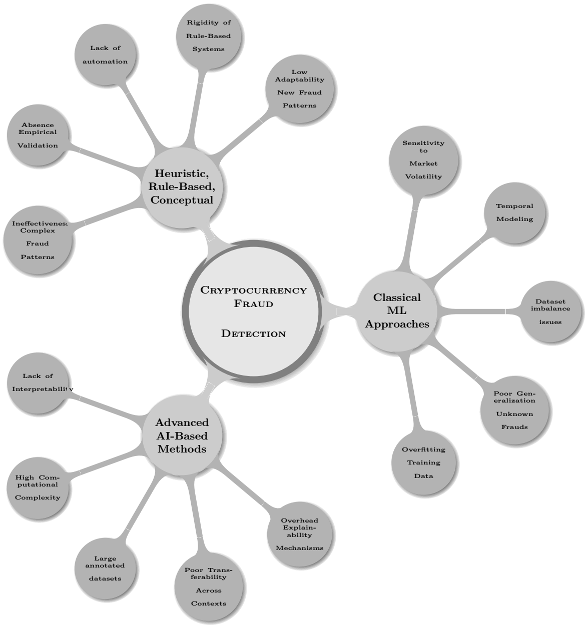
Lack of automation
Rigidity of Rule-Based Systems
Low Adaptability New Fraud Patterns
Absence Empirical Validation
Heuristic, Rule-Based, Conceptual
Ineffectiveness Complex Fraud Patterns
Sensitivity to Market Volatility
Temporal Modeling
Cryptocurrency Fraud Detection
Classical ML Approaches
Dataset imbalance issues
Lack of Interpretability
Advanced AI-Based Methods
Poor Generalization Unknown Frauds
Overfitting Training Data
High Computational Complexity
Overhead Explainability Mechanisms
Large annotated datasets
Poor Transferability Across Contexts

**Fig. 1.** Challenges by methodological class in cryptocurrency fraud detection.

Overall, while each class has unique strengths and forces, they also have critical limitations that must be addressed to enable robust and scalable fraud detection in dynamic cryptocurrency systems.

### 5.2 Solutions

In the field of crypto-fraud detection, each methodological class faces specific technical and practical challenges. Below is a synthesis of the main issues and appropriate solution strategies based on recent initiatives and practices.

**Class 1—Advanced AI-Based Approaches**

- Lack of interpretability: integrate explainability tools like SHAP, LIME, or GNNExplainer to reveal model decision-making processes. In addition, incorporating attention mechanisms can also enhance the model's interpretability.
- High computational complexity: apply model compression techniques such as pruning, knowledge distillation, or quantization. Also take advantage of high-performance architectures and GPU/TPU-accelerated batch processing.
- Requirement for large annotated datasets: leverage semi-supervised learning or self-supervised techniques. Additionally, generate synthetic data or labels via simulation or expert-driven labeling.
- Poor transferability across contexts: use transfer learning or domain adaptation to improve generalization to unseen contexts or different platforms, for example, transferring from Ethereum to Bitcoin.
- Overhead from explainability mechanisms: favor lightweight or integrated explainability methods like attention-based networks or simplified surrogate models.

**Class 2—Classical Machine Learning Approaches**

- Overfitting to training data: use cross-validation, regularization, and ensemble methods, and expand the dataset with synthetic fraud instances.
- Poor generalization to unknown frauds: combine with anomaly detection techniques or integrate unsupervised learning to capture unknown patterns.
- Data imbalance issues: use resampling strategies like SMOTE (Synthetic Minority Oversampling) or undersampling. Alternatively, use Generative Adversarial Networks (GANs) to produce synthetic fraudulent samples.
- Lack of relational or temporal modeling: enrich feature sets with temporal sequences or topological data from transaction graphs, combining time-series preprocessing with graph-based learning methods.
- Sensitivity to market volatility: integrate contextual features, such as volatility indicators, and periodically retrain models using sliding time windows.

**Class 3—Non-ML, Heuristic or Conceptual Approaches**

- Rigidity of rule-based systems: implement rule-learning systems or adaptive logic-based engines. Use symbolic AI or evolutionary algorithms to automate rule refinement.
- Low adaptability to new fraud patterns: develop self-updating rule bases using user feedback loops or active learning strategies for semi-automatic updates.

- Lack of automation: deploy real-time detection systems using event-driven frameworks, such as Apache Kafka, real-time APIs, and smart contracts.
- Absence of empirical validation: encourage open benchmarking using public or synthetic datasets, while validating models through reproducible experiments and cross-comparisons.
- Ineffectiveness on complex fraud patterns: combine rule-based systems with machine learning or anomaly detection in a hybrid detection pipeline to balance interpretability and adaptability.

Hybrid approaches combining the interpretability of inference, the efficiency of traditional machine learning, and the advanced adaptability of artificial intelligence (AI) appear promising. Moreover, the integration of self-supervised learning, transfer learning, and advanced artificial intelligence (AAI) could improve scalability and reliability in dynamic and data-limited blockchain environments.

# 6 Research directions

This section presents research trends in AI-driven cryptocurrency fraud detection, focusing on addressing current gaps through hybrid and scalable approaches. Based on a three-category analysis, it addresses key limitations, such as data sparsity, market volatility, and the evolving complexity of fraud. The following directions are categorized into three thematic areas:

## 6.1 Class 1-Advancing Interpretable and Scalable Deep Learning Models

Advanced AI techniques, such as neural networks and large-scale neural networks, are increasingly used to monitor complex blockchain environments. However, their effectiveness depends on their interpretability, scalability, and practical deployment. To address these challenges, future research should focus on:

- *Improving interpretability and performance:*
  - Design attention-based neural architectures to enhance model transparency and focus on key transactional features.
  - Combine symbolic and subsymbolic reasoning within hybrid frameworks to improve interpretability and logic-based traceability.
- *Reducing computational burden through optimization:*
  - Implement model compression techniques—like pruning, parameter sharing, and knowledge distillation—to reduce computational overhead.
  - Ensure a trade-off between model compactness and predictive accuracy for effective deployment in production environments.
- *Addressing data scarcity with modern training strategies:*
  - Leverage self-supervised and semi-supervised learning paradigms to utilize large volumes of unlabeled blockchain data.
  - Refine smaller annotated subsets to improve accuracy.
- *Enhancing generalizability with transfer learning:*
  - Apply domain adaptation techniques to generalize models across various blockchain infrastructures and cryptocurrency environments.
  - Enhance robustness by transferring learned representations from adjacent domains of financial or cyber fraud detection.

### 6.2 Class 2-Improving Robustness and Adaptability in Classical ML

Traditional machine learning techniques remain attractive due to their simplicity, interpretability, and computational efficiency. However, they face several challenges related to their adaptability and robustness. To strengthen their reliability, future research should focus on:

- *Enhancing model resilience and generalization:*
  - Mitigate overfitting and improve adaptability to evolving fraud patterns.
  - Incorporate temporal and relational features to reflect blockchain dynamics.
  - Develop ensemble frameworks combining accurate classifiers and anomaly detectors.
- *Addressing data imbalance in fraud detection:*
  - Complement oversampling techniques (e.g., SMOTE) with adversarial data augmentation.
  - Generate synthetic fraudulent transactions using GANs to enrich minority classes.
- *Improving feature engineering and temporal modeling:*
  - Combine manual feature extraction with graph-based and temporal representations.
  - Encode transaction sequences and wallet interaction patterns to capture fraud dynamics.
  - Include contextual features such as market volatility or social sentiment.
- *Extending detection capabilities with unsupervised learning:*
  - Integrate unsupervised or semi-supervised modules to detect novel static patterns.
  - Expand coverage from static signatures to dynamic behavioral patterns.

### 6.3 Class 3-Bringing Automation and Flexibility to Heuristic Methods

Heuristic and rule-based approaches offer transparency, scalability, and consistency in the domain. Often, they lack adaptability and empirical robustness. Future work should be considered to modernize these systems and increase their relevance.

- *Automating rule generation and evolution:*
  - Use genetic programming and symbolic AI to derive detection rules from historical data or expert knowledge.
  - Implement feedback loops and active learning to refine rules based on model validation and detected errors.
- *Enabling real-time detection through dynamic systems:*
  - Leverage event-driven systems (e.g., Kafka, smart contracts) for live detection pipelines.
  - Transform static rules into adaptive mechanisms capable of reacting to live fraud events.
- *Ensuring validation and reproducibility:*
  - Evaluate systems in sandboxed blockchain environments.

    - Compare results using datasets from public cryptocurrency ledgers to ensure generalizability.
- *Exploring hybrid models for persistent threats:*
    - Combine heuristics with machine learning or anomaly detection to balance interpretability and adaptability.
    - Enhance adaptability while maintaining explainability through hybrid inference frameworks.

## 7 Conclusion

Several cross-cutting themes emerge across all methodological classes. First, hybrid frameworks integrating deep learning, classical ML and heuristic knowledge provide a powerful path forward, especially when combined with modular architectures that allow of independent component updating or replacement. Second, the emergence of on-chain and off-chain data integration, such as combining blockchain data with social media or regulatory information, opens up new possibilities for context-rich models. Third, there is a growing need for standards, metrics, and standardized datasets to accurately evaluate and compare fraud detection models. Establishing such shared and open empirical frameworks would accelerate innovation and replicability. Finally, regulatory compliance and ethical issues related to fairness, transparency, and privacy should be a focus of future research to make fraud detection systems reliable, verifiable, and compliant with financial governance principles. In short, future research should not only improve methodological performance but also enhance interpretability, adaptability, and accountability. Only these diverse strategies will enable the construction of robust fraud detection systems that protect rapidly evolving cryptocurrency ecosystems.

## References

[1] David S Kerr, Karen A Loveland, Katherine Taken Smith, and Lawrence Murphy Smith. Cryptocurrency risks, fraud cases, and financial performance. risks 11 (3): 51, 2023.

[2] Massimo Bartoletti, Stefano Lande, Andrea Loddo, Livio Pompianu, and Sergio Serusi. Cryptocurrency scams: analysis and perspectives. Ieee Access, 9:148353–148373, 2021.

[3] Yisong Chen, Chuqing Zhao, Yixin Xu, and Chuanhao Nie. Year-over-year developments in financial fraud detection via deep learning: A systematic literature review. January 2025.

[4] Nrusingha Tripathy, Sidhanta Kumar Balabantaray, Surabi Parida, and Subrat Kumar Nayak. Cryptocurrency fraud detection through classification techniques. International Journal of Electrical and Computer Engineering (IJECE), 14(3):2918–2926, 2024.

[5] Arjun Kamisetty, Abhishake Reddy Onteddu, RR Kundavaram, JCS Gummadi, S Kothapalli, and M Nizamuddin. Deep learning for fraud detection in bitcoin transactions: An artificial intelligence-based strategy. NEXG AI Review of America, 2(1):32–46, 2021.

[6] Mohammad Javad Shayegan, Hamid Reza Sabor, Mueen Uddin, and Chin-Ling Chen. A collective anomaly detection technique to detect crypto wallet frauds on bitcoin network. Symmetry, 14(2):328, 2022.

[7] Youssef Elmougy and Ling Liu. Demystifying fraudulent transactions and illicit nodes in the bitcoin network for financial forensics. In Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pages 3979–3990, 2023.

[8] Noor Nayyer, Nadeem Javaid, Mariam Akbar, Abdulaziz Aldegheishem, Nabil Alrajeh, and Mohsin Jamil. A new framework for fraud detection in bitcoin transactions through ensemble stacking model in smart cities. IEEE Access, 11:90916–90938, 2023.

[9] Praniket Walavalkar, Ansh Dasrapuria, Meghna Sarda, and Lynette D'mello. A token-based approach to detect fraud in ethereum transactions,'. International Journal for, 2024.

[10] Massimo Bartoletti, Barbara Pes, and Sergio Serusi. Data mining for detecting bitcoin ponzi schemes. pages 75–84, 2018.

[11] Michał Ostapowicz and Kamil Żbikowski. Detecting fraudulent accounts on blockchain: A supervised approach. pages 18–31, 2019.

[12] Tehreem Ashfaq, Rabiya Khalid, Adamu Sani Yahaya, Sheraz Aslam, Ahmad Taher Azar, Safa Alsafari, and Ibrahim A Hameed. A machine learning and blockchain based efficient fraud detection mechanism. Sensors, 22(19): 7162, 2022.

[13] Saru Kumari. Machine learning applications in cryptocurrency: Detection, prediction, and behavioral analysis of bitcoin market and scam activities in the usa. International journal of Sustainable Science and Technology, 1(1), 2025.

[14] Njoku ThankGod Anthony, Mahmoud Shafik, Fatih Kurugollu, and Hany F Atlam. Anomaly detection system for ethereum blockchain using machine learning. pages 311–316, 2022.

[15] Weili Chen, Zibin Zheng, Jiahui Cui, Edith Ngai, Peilin Zheng, and Yuren Zhou. Detecting ponzi schemes on ethereum: Towards healthier blockchain technology. In Proceedings of the 2018 world wide web conference, pages 1409–1418, 2018.

[16] Massimo La Morgia, Alessandro Mei, Francesco Sassi, and Julinda Stefa. Pump and dumps in the bitcoin era: Real time detection of cryptocurrency market manipulations. In 2020 29th international conference on computer communications and networks (ICCCN), pages 1–9. IEEE, 2020.

[17] Fredy Andres Aponte-Novoa, Ana Lucila Sandoval Orozco, Ricardo Villanueva-Polanco, and Pedro Wightman. The 51% attack on blockchains: A mining behavior study. IEEE access, 9:140549–140564, 2021.

[18] Halima Oluwabunmi Bello, Courage Idemudia, and Toluwalase Vanessa Iyelolu. Integrating machine learning and blockchain: Conceptual frameworks for real-time fraud detection and prevention. World Journal of Advanced Research and Reviews, 23(1):056–068, 2024.