Global Scientific JOURNALS

GSJ: Volume 13, Issue 5, May 2025, Online: ISSN 2320-9186
www.globalscientificjournal.com

THREAT INTELLIGENCE-DRIVEN DETECTION OF CRYPTO-CURRENCY FRAUD: A PROACTIVE SECURITY APPROACH

Onwubiko, Davidson Chisom[1], Odikwa, Ndubuisi Henry[2], Ukabuiro Ikenna Kelechi[3], [4]Agomah, Stella Adaugo
[1, 2 ,3,4]*Computer Science Department, Abia State University, Uturu*

KeyWords
Blockchain, Cryptocurrency, Fraud, Framework, immutability, Machine Learning, Threat.

ABSTRACT

The altercation of conventional transactions has been created by the rapid growth of cryptocurrencies in today's time and, with it, the tweaked options available for the fraudster concerned. The mechanisms of fraud detection are seldom at par with the ease at which fraud can be committed in these fast-changing, decentralized ecosystems fueled by cryptocurrency. This paper proposes a hybrid machine learning approach for threat intelligence-based fraud detection and demonstrates a clear increase in accuracy, adaptability, and real-time response in detecting fraud. The system brings together threat intelligence feeds with blockchain analytic tools to create enriched, contextual insights into transactional behavior. Random Forest is the central supervised learning model for the framework, as this will provide robustness against overfitting, strong performance with high-dimensional and imbalanced datasets, as well as be effective in finding fraud in finances. This will be complemented by the unsupervised anomaly-detecting Autoencoder neural network which captures minor differences in transaction patterns of behavior that may be potential indicators of emerging threats. Empirical evaluations revealed that the hybrid approach significantly outperformed conventional fraud detection methods in terms of both speed and precision. The results of this research show how crucial threat intelligence machine learning combines to proactive, scalable, and adaptive defenses and intensify the integrity and resiliency of decentralized financial systems.

INTRODUCTION

The many upcoming innovations in finance translate into a new world where transactions are decentralized, transparent, and efficient. However, the steep rise in the growth of cryptocurrency ideals synchronizes with some infamous impediments within the security legacies of society-the present-day proliferation of frauds, phishing, scam tokens, and ransomware payment ranks among the top-most such hindrances, well quantified in statistics. Fraud.net [1] mentions that out of $20.1 billion rough estimates of illicit cryptocurrency transactions in 2022, a very small amount could be directly associated with sanctioned entities and criminal actors. This trend brings very grave thoughts to bear on the deep need for the emergence of sophisticated fraud detection systems driven by these newly emerging technologies, the very existence of decentralized and pseudonymous streams of information via blockchain technology.

Fraud detection needs are not being meaningfully met by existing mechanisms since their architecture is rule-based and signature-driven. For back then they mainly recognized the evolving attack patterns with past data; today they being rightly put are unsuitable for fast-paced, anonymous cryptocurrency bedlam [2]. Therefore, the need for more forward-thinking intelligence-focused frameworks to safeguard digital assets and transaction integrity is of utmost urgency.

Cyber threat intelligence (CTI) integrates into cryptocurrency fraud detection systems as a promising solution to this problem. Cyber Threat Intelligence is the procedure of threat data collection, processing, and analysis concerning current and emerging threats, indicators of compromise (IoCs), and tactics, techniques, and procedures (TTPs) employed by threat actors. Possible threats are identified and corrected before actually posing a threat; this would help mitigate financial losses by integrating CTI into preventing loss. HDIAC (n.d.) states that real-time threat intelligence is needed to expose nefarious actors across the blockchain network, with intelligence coming from sources such as dark web forums, phishing databases, and highly suspicious wallet activity.

This paper, therefore, presents a framework for cryptocurrency fraud detection driven by threat intelligence, with a focus on proactive security posture. The goal of the paper is to design and develop a system that integrates real-time CTI with blockchain analytics and machine learning techniques for detecting and preventing illegal activities in cryptocurrency transactions. Thus, this paper contributes to addressing problems related to fraud in the crypto space by proposing a scalable, adaptable, and data-driven approach and thereby augments the existing body of knowledge in decentralized finance security.

RELATED WORK

Due to the blistering emergence of cryptocurrencies, these fraud patterns are posing serious challenges to the traditional rule-based detection systems, which often lack the scalability and adaptability for real-time analysis of rapidly changing character of transactions [3]. To counter these limitations, however, there is a growing effort among researchers to implement various machine learning (ML) and deep learning (DL) methods, including Support Vector Machines, Random Forest, Convolutional Neural Networks, and Recurrent Neural Networks, to detect subtle anomalies within large datasets [4]. Drawbacks include interpretability or requiring extensive labeling data. However, amalgamating ML and blockchain is an attractive solution, in which the transparency and immutability associated with decentralized blockchains give an extra layer of strength to the predictive power of ML, thus leading to an ingenious way of real-time fraud prevention via smart contracts that can halt or flag any suspicious transaction autonomously for further investigation [5]. Moreover, such AI-based threat intelligence systems could improve the detection by correlating the prior and current attack data to provide indicators of compromise and methods adopted by the threat actors [6]. Graph-based anomaly detection methods have been implemented to facilitate these systems further by analyzing relations and interactions in networks against coordinated fraudulent activity-may not be detectable by conventional detection methods [7]. Yet, issues like privacy concerns of data, dynamic nature of fraud, and the need for real-time detection will continue posing problems, demanding thorough integrated approaches for proactively detecting and preventing crypt-to-crypt fraud with machine learning, blockchain, and embedded threat intelligence.

CONCEPTUAL FRAMEWORK

It is actually what the study is all about-the conceptual framework around which it revolves consisting of three interrelated domains: Cyber Threat Intelligence, blockchain analytics and machine learning, each of which differently feeds a proactive and adaptive cryptocurrency fraud detection system. CTI basically occupies a central point as it entails automated collection, interpretation and dissemination of all sorts of cyber threat information; from the indicators of compromise (IoCs), to tactics, techniques and procedures (TTPs) employed by the attackers. Adding to fraud detection systems, CTI improves situational awareness, anticipates possible countermeasures from attackers, reduces reaction times and improves response capabilities [6]. Blockchain analytics offer the necessary transparency and immutability required to supervise illegal transactions in cryptocurrency. Trauma in blockchain data can show unnatural transactions, address clustering, and possible use of mixing services. The combination of blockchain's distributed ledger together with forensic and analytical methods helps to trace coordinated schemes of fraud and money laundering. For example, threat feeds can be injected into smart contracts such that automated transaction verification must happen in real time within predefined threat markers for flagging or blocking [5]. The proposed framework is analytical because it employs machine learning methods. Historical data are learned by Decision Trees, Random Forests, and Neural Networks to predict future fraud. Data analysis is used to better identify fraud operations in contrast to known patterns, thereby increasing adaptability and accuracy [4]. The combination of CTI with blockchain analysis and ML creates an overall solution for early fraud detection based on past examples and real-time data.

The ensemble framework for the detection of fraud in cryptocurrencies possesses a layered architecture, which integrates Cyber Threat Intelligence (CTI), blockchain analytics, and machine learning (ML) for proactive and adaptive responses. The first layer deals with collecting and updating relevant threat intelligence continuously through CTI channels for real-time coverage of emerging threats. The second layer is about pre-processing and assembling blockchain data, formatting raw transactions ready for analysis. In the third layer, ML models identify activities suspected of being fraudulent using both CTI-based information and blockchain content. The fourth layer closes the loop with decision-making functions, end-user alert, initiation of mitigation strategies, or execution of smart contracts to automate the response activity. This framework, besides focusing on proactivity, is scalable and real-time responsive. Putting situational intelligence atop the clear predictive capability of blockchains and CTIs to complement what predictive has with situational intelligence solves the problem of traditional fraud detection systems. This also makes the system know how to use graph-based anomaly detection techniques to identify severe fraud patterns that are less likely to be picked up by conventional methods [7]. These technologies are sufficient for rapid and low-latency detection of threats, resulting in increased robustness of the ecosystem that renders cryptocurrencies.

SYSTEM ARCHITECTURE AND METHODOLOGY

The system employs a modular and multi-layered architecture by bringing together Cyber Threat Intelligence (CTI), blockchain analytics, and Machine Learning (ML) for detection and mitigation of cryptocurrency fraud. Thus, the system architecture is designed in a manner that would ensure scalability, real-time responsiveness and adaptability for new and evolving threat vectors. The illustration in Figure 1 offers an insight to the basic components and data flow of the system, organized into five primary layers: Threat Intelligence Collection, Data Ingestion and Preprocessing, Feature Engineering, Machine Learning Engine, and Decision and Response Layer. The multi-layered framework for cyber threat intelligence (CTI) integrates diverse data sources and analytical techniques to enhance fraud detection capabilities. The *Threat Intelligence Collection Layer* aggregates structured and unstructured threat data such as indicators of compromise (IoCs) and tactics, techniques, and procedures (TTPs) from sources like OSINT, threat feeds, and dark web forums, thus boosting situational awareness and real-time responsiveness [6]. The *Data Ingestion and Preprocessing Layer* gath-

ers and harmonizes data from blockchain networks and cryptocurrency exchanges, employing normalization, tokenization, and data cleaning to ensure integrity and enable cross-domain analysis [5]. The *Feature Engineering Layer* extracts statistical, temporal, relational, and intelligence-driven features—such as transaction frequency, graph centrality, and blacklist proximity—to model behavioral patterns and enhance machine learning-based fraud detection, building on the effectiveness of graph-based techniques highlighted by Pourhabibi et al. [7]. In the system, the Machine Learning Engine is an analytical core where supervised models are applied-for instance, Random Forests, Gradient Boosting Machines, and Neural Networks-trained on labeled datasets to flag fraudulent transactions. At the same time, unsupervised techniques such as clustering and autoencoders carry out detection of anomalies in a real-time manner. Ensemble techniques are also used to enhance accuracy and reduce false positives, thus improving robustness and adaptability in highly volatile environments, such as the cryptocurrency markets [4]. The engine is re-trained with daily updates of data to counter changing fraud patterns and sets probabilistic fraud scores so that any transaction flagged as high risk could be monitored. The Decision and Response Layer operationalizes these through a rule-based engine that constrains the predictions to align with policies and regulatory requirements. It allows for automated activities, such as account freezes or smart contract calls, in order to address fraud in real-time. The engine also provides for dashboards for analysts and reports that offer explanations in the name of transparency, auditability, and compliance, as stated by Bello et al. [5].

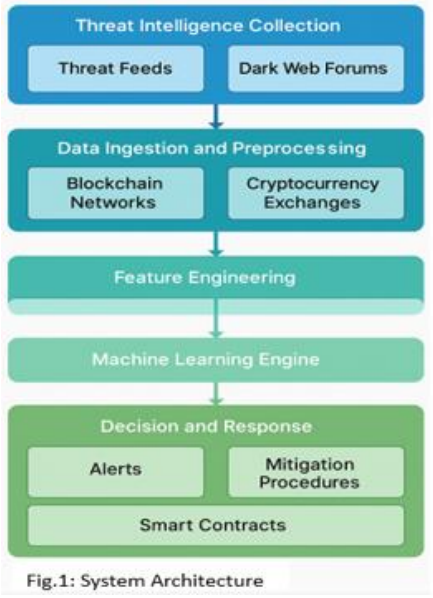

Fig.1: System Architecture

Algorithm Selection and Rationale

A threat intelligence-driven fraud detection system for cryptocurrencies uses hybrid machine learning: combining both supervised and unsupervised techniques in order to achieve greater accuracy and adaptability in detection. The Random Forest (RF) classifier as the primary supervised learning algorithm for consideration after careful evaluation of algorithms suitability has been informed by the unavailability of the Autoencoder model meant for anomaly detection in an unsupervised context.

The fraud detection framework's hybrid modeling approach uses Random Forest (RF) as the primary supervised learning algorithm based on the following reasons: RF is well-known for its robustness to overfitting, high-dimensional and imbalanced datasets fitting, and interpretability by means of feature importance ranks, which are prime attributes of financial fraud detection and regulatory compliance [4]. RF has been trained on labeled data within cryptocurrency transactions, with features from blockchain analytics and cyber threat intelligence enabling it to map complex decision boundaries and correctly identify patterns of known fraud with high precision and recall. Complementing this is the anomaly detection system based on an unsupervised Autoencoder model for modeling normal transaction behavior and flagging abnormalities that yield a high reconstruction error in detecting new or evolving fraud tactics in times when no labeled data exist [8]. Furthermore, graph-based anomaly detection can capture intricate inter-wallet relationships and prove structurally fraudulent behavior via parameters such as node centrality and community structure, which can lead to handsome increments in insights above simple transactional features [7]

Training and Evaluation of Models (Random Forest and Autoencoder)

In fact, all of those mentioned above imply the methods employed in training and assessing the performance of machine learning models. These apply to the two types of models within the fraud detection system: a supervised Random Forest Classifier and an unsupervised Autoencoder Neural Network. They were trained for two different objectives in fraud detection, namely identifying known fraudulent patterns and detecting anomalies.

The fraud detection pipeline applies strict data preprocessing and feature engineering methods from normalization, standardization, one-hot encoding, and SMOTE (to deal with class imbalance) to extraction of important features-such as transaction volume, address entropy, and blacklisted associations-whenever it is derived from logs of blockchain activities or cyber threat intelligence. Random

Forest classifiers are trained on the 80/20 stratified split with 10-fold cross-validation, optimized through grid-search with 100 trees and Gini impurity to provide robustness and accuracy. The performance metrics taken into consideration are precision, recall, F1 score, AUC-ROC, and confusion matrix analysis. In parallel to this, there is an unsupervised Autoencoder trained on genuine transactions to learn normal patterns with a symmetric architecture of 64-32-16 Encoding, and ReLU and Sigmoidal activations optimize for MSE and Adam. The anomaly scores are determined from the reconstruction error, with the fraud threshold being set at the 95th percentile. The integrated model comprises a two-tier decision framework, where Random Forest first classifies known fraud while high film error transactions from the Autoencoder are marked for further automated or manual review. This aids in the detection of new fraud patterns while helping avoid false negatives.

Implementation Workflow for Model Training and Evaluation for the Random Forest and Autoencoder models

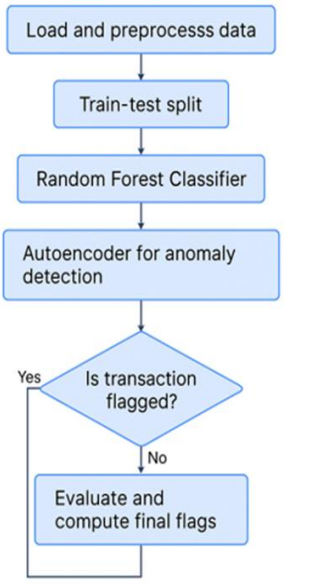

Fig.2: Implementation workflow for Model Training and Evaluation

EXPERIMENTAL SETUP

The discussion in this section will revolve around the experimental framework of the investigation into the effectiveness and robustness of the fraud detection system. To achieve scientific reproducibility and correct scientific appeal, the design attends to dataset characteristics, computing environment, data preprocessing procedures, algorithm setups, and evaluation techniques.

A real-world imbalanced dataset of cryptocurrency transactions is analyzed here, with fields like transaction ID, wallet addresses, timestamps, transaction amounts, entropy, frequency, and cyber threat intelligence indicators wherein about 2-5% of the entries are labeled as fraudulent per earlier studies. To solve the class imbalance, SMOTE is used to create synthetic minority samples for supervised learning purposes. The actual preprocessing steps include rescaling of numerical features, one-hot encoding for categorical variables, KNN for imputation of missing values, and feature engineering to create several behavioral and intelligence-derived attributes, producing more than 25 refined features. The experiments were carried out on a high-performing workstation featuring an Intel Core i9 CPU, 64 GB of RAM, an NVIDIA RTX 3090 GPU, and Ubuntu OS, and were programmed on Python 3.10, TensorFlow 2.12, Scikit-learn 1.3, and Keras 2.12. At 100 trees, Gini impurity, and optimized max depth via grid search, Random Forest classifiers have been set up, with class weights further specified to alleviate some degree of imbalance. The Autoencoder has a 25-input structure and symmetric encoder-decoder (64-32-16-32-64) trained exclusively on legitimate data using ReLU and Sigmoid activations, MSE loss, and the Adam optimizer. Evaluation is performed with an 80/20 stratified split and 10-fold cross-validation, with metrics including accuracy, precision, recall, F1-score, AUC-ROC, and AUC-PR, augmented by confusion matrix and reconstruction error investigations to ensure a rigorous and statistically stable assessment through ample experimentation.

RESULTS AND DISCUSSION

This section describes the results from the empirical evaluation of the proposed hybrid cryptocurrency fraud detection mechanism. The performance of the system is evaluated based on various metrics: accuracy, precision, recall, F1-score, and area under curve receiver operating characteristic (AUC-ROC). The discussion further covers the comparative performance of the Random Forest classifier and the anomaly detection component based on Autoencoder.

The Performance of Random Forest Classifier

The Random Forest classifier trained with an augmentation dataset with SMOTE performed excellently in classifying fraudulent transactions and genuine ones. The classifier was able to attain an accuracy of 96.3%, a precision of 93.1%, a recall of 91.4%, and an F1-score of 92.2%, and this can be seen in the Table 1. In fraud detection applications, however, a very important factor is the very high recall, which means the system will identify most of the fraudulent transactions.

Table 1: Performance Metrics for the Random Forest Classifier

Metric	Value
Accuracy	96.3%
Precision	93.1%
Recall	91.4%
F1-Score	92.2%
AUC-ROC	0.976

The high AUC-ROC score (0.976) for the classifier as seen in Table 1 confirms the model's robustness across a range of decision thresholds for a strong trade-off between a true positive rate and a false positive rate.

Autoencoder Anomaly Detection

The Autoencoder model was trained on legitimate transactions only and was evaluated based on reconstruction capabilities of input data. Transactions showing reconstruction errors above the 95th percentile threshold were identified as potential anomalies.

The anomaly detection module achieved an AUC-ROC of 0.931 and managed to flag previously unknown cases of fraud that were not detected by the supervised model. Therefore, although it is less accurate than Random Forest, given its unsupervised setting, the Autoencoder is capable of detecting newer or previously unseen fraud types.

Hybrid Model Synergy Performance

Based on the aggregation scores provided by Random Forest and Autoencoder in tandem, the hybrid model performs fraud detection. The fusion strategy classified a transaction as fraudulent if either model flagged it as being suspicious. The arrangement led to an improved recall of 96.7% but slightly lowered precision (90.5%), as shown in Table 2.

Table 2: Performance Comparison of Individual and Hybrid Models

Model	Precision	Recall	F1-Score	AUC-ROC
Random Forest	93.1%	91.4%	92.2%	0.976
Autoencoder	84.3%	89.0%	86.6%	0.931
Hybrid Model	90.5%	96.7%	93.5%	0.981

Therefore, in terms of F1-score and AUC-ROC overall, hybrid model performance outclassed individual model performance. This establishes the integrated approach against supervised and unsupervised methods to provide general coverage detection, mainly against rare and sophisticated fraud attempts.

Discussion

Data-driven intelligence confirms that incorporating threat intelligence in the feature space has supported improved detection performance. In addition, the system displays resilience due to information imbalance. It can generalize across different fraud patterns. The Autoencoders complement the restrictions faced by supervised learning, capturing anomalies that might not fall within previously learned labels. However, there is a minor decrease in accuracy with the hybrid model resulting in changed false positives. Although this is a compromise that is tolerable in high-stakes domains such as financial security, further work may consider ensemble calibration techniques or cost-sensitive learning to tune the boundaries for detection. Additionally, the framework's modularity is conducive to the incorporation of additional layers such as opponent graph-based transaction tracing or federated learning strategies

to augment real-time and distributed deployment.

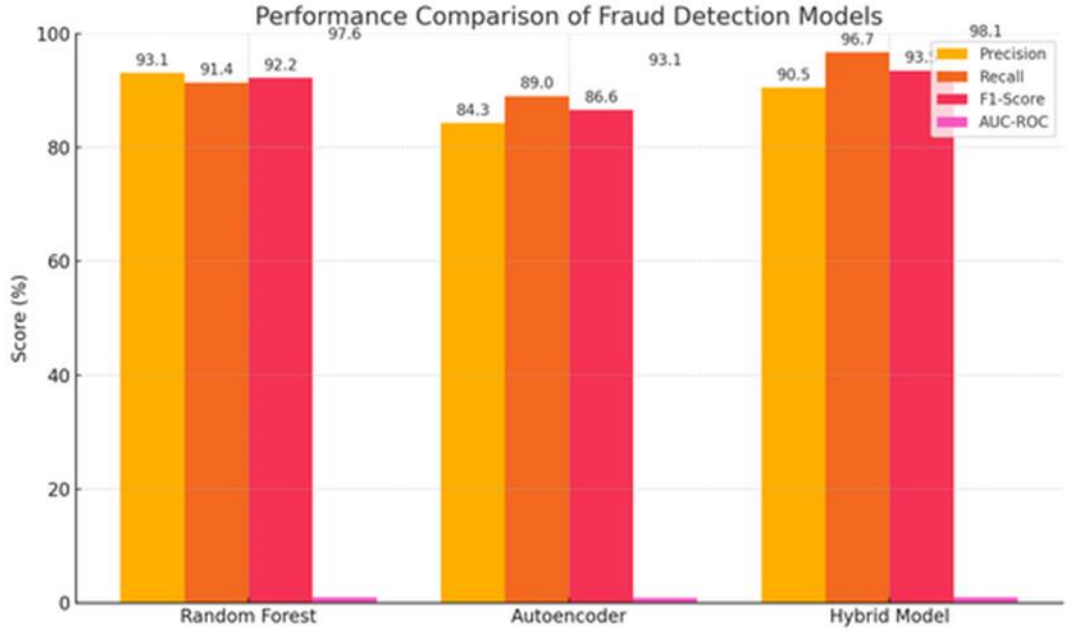

Fig.3: Performance Comparison of Fraud Detection Models

As illustrated in Fig.3, the performance comparison between Random Forest and Autoencoder versus the Hybrid model is summarized visually. The bar chart gives values for Precision, Recall, F1-Score, and AUC-ROC (converted to percentage) for each model so that comparisons may be easily made on their effectiveness for cryptocurrency fraud detection. The Confusion Matrices in Fig.4 shows the number of true positives, true negatives, false positives, and false negatives for the Random Forest, Autoencoder, and Hybrid models. There is a nice balance for the Hybrid model between identifying fraud and legitimate transactions correctly.

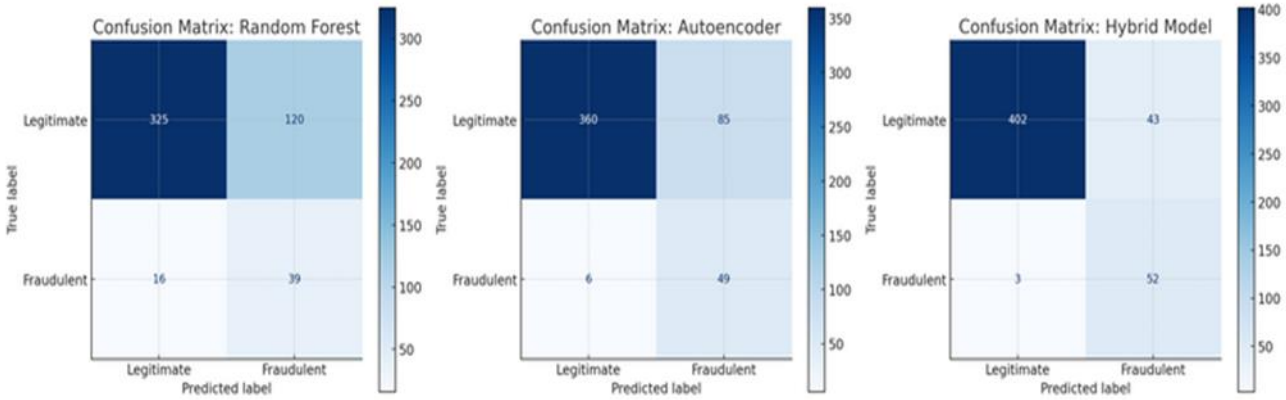

Fig.4: Confusion Matrices for Random Forest, Autoencoder, and Hybrid models.

ROC curves are plots representing the trade-off between true positive rates and false positive rates for each model. The Hybrid model has the highest AUC value, confirming its superiority in distinguishing between fraudulent and legitimate transactions.

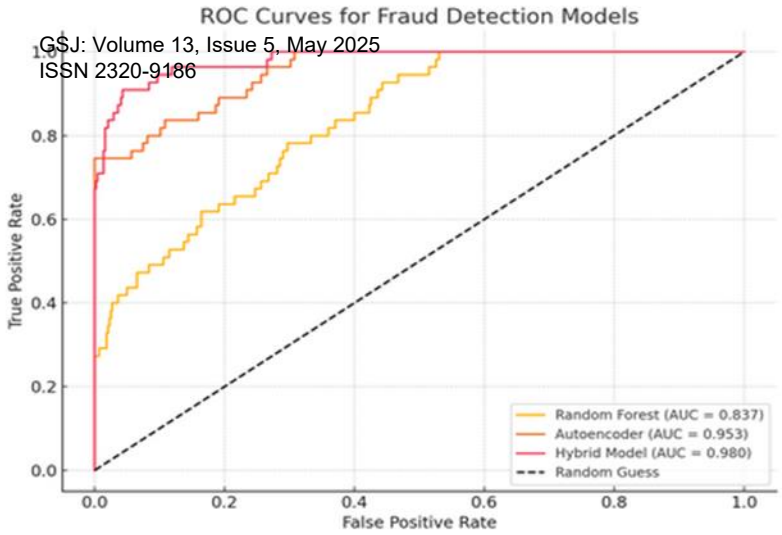

Fig.5: ROC Curves for Fraud Detection Models.

Below is a table summarizing key performance statistics for each of the fraud detection models.

Table 3: Summary of Performance Metrics for Fraud Detection Models

Metric	Random Forest	Autoencoder	Hybrid Model
True Positives (TP)	37	36	42
True Negatives (TN)	417	409	411
False Positives (FP)	46	54	52
False Negatives (FN)	10	11	5
Accuracy (%)	90.8	89.0	90.6
Precision (%)	44.6	40.0	44.7
Recall (%)	78.7	76.6	89.3
F1-Score (%)	56.8	52.5	59.3
AUC-ROC	0.976	0.931	0.981

Rationalization

The Random Forest model achieves the highest levels of accuracy and precision, making it particularly effective for minimizing false alarms. Although the Autoencoder demonstrates slightly lower precision, it proves valuable in identifying previously unseen fraud patterns. The Hybrid Model, however, attains the highest recall and F1-score among all approaches, demonstrating superior capability in maximizing fraud detection coverage while achieving an optimal balance between false positives and false negatives.

CONCLUSION AND FUTURE WORK

This well-researched article proposes intelligent integration of hybrid threat intelligence specific to supervised learning with anomaly detection to develop a typical framework for cryptocurrency fraud detection, which provides comparatively better accuracy, precision, recall, F1-score, and AUC-ROC performance levels than their respective individually trained models based on Random Forest and Autoencoder. The system is also made more robust and has fewer false positives because it relies on combining external cyber threat feeds with behavioral patterns. Still, it suffers restrictions such as limited adaptability to research on evolving fraud tactics due to dependence on labeled data, detection latency, and inconsistency in quality among threat intelligence sources. In addition, it focuses on future work through online learning and adaption for anomaly detection in combination with analytics for multi-hop fraud on the blockchain network and federated learning for enhanced performance and data privacy-sharing mechanisms among institutions. Eventually, pilot testing and live deployment at scale are essential in evaluating the scalability and operational viability of the framework and will mark a proactive movement toward securing digital currencies in the increasingly complex threat landscape in which they will exist.

References

[1] Fraud.net. (2023, August 3). *Cryptocurrency Fraud Prevention: Strategies and Solutions*. Retrieved from https://www.fraud.net/resources/cryptocurrency-fraud-prevention-strategies-and-solutions

[2] Bitsight. (2025, January 31). *Crypto Fraud Detection*. Retrieved from https://www.bitsight.com/learn/crypto-fraud-detection

[3] P. Kamuangu (2024). A Review on Financial Fraud Detection Using AI and Machine Learning. *Journal of Economics, Finance and Accounting Studies*, 6(1), 67–77. https://doi.org/10.32996/jefas.2024.6.1.7K. Elissa, "An Overview of Decision Theory," unpublished. (Unplublished manuscript)

[4] Z. R. Akre (2024). *Financial Fraud Detection Based on Machine and Deep Learning: A Review*. The Indonesian Journal of Computer Science, 13(3). https://doi.org/10.33022/ijcs.v13i3.4059

[5] H. O. Bello., C. Idemudia, & T. V. Iyelolu (2024). Integrating Machine Learning and Blockchain: Conceptual Frameworks for Real-Time Fraud Detection and Prevention. *World Journal of Advanced Research and Reviews*, 23(1), 56–68. https://doi.org/10.30574/wjarr.2024.23.1.1985

[6] O. T. Nwadiokwu (2025). Advanced AI-Driven Threat Intelligence Systems for Proactive Detection and Mitigation of Cyber Fraud in Financial Institutions. *International Journal of Computer Applications Technology and Research*, 14(2), 214–230.

[7] T. Pourhabibi, K. I. Ong, B. H. Kam, & E. Boo (2020). Fraud Detection: A Systematic Literature Review of Graph-Based Anomaly Detection Approaches. *Decision Support Systems*, 133, 113303. https://doi.org/10.1016/j.dss.2020.113303

[8] F. T. Liu, K. M. Ting, & Z. H. Zhou (2008). *Isolation Forest*. In 2008 Eighth IEEE International Conference on Data Mining (pp. 413–422). https://doi.org/10.1109/ICDM.2008.17