

UDC 336.7:343:5:004.8
JEL: C45; G21; G28; O33
DOI: <https://doi.org/10.32983/2222-4459-2026-8-96-106>

MACHINE LEARNING FOR FRAUD DETECTION IN DIGITAL BANKING AND DeFi: TOWARDS SUSTAINABLE FINANCIAL SYSTEMS

© 2026 CAPRIAN IU., CAPRIAN IU., LOM IO. V.

UDC 336.7:343:5:004.8
JEL: C45; G21; G28; O33

Caprian Iu., Caprian Iu., Lom Io. V. Machine Learning for Fraud Detection in Digital Banking and DeFi: Towards Sustainable Financial Systems

The rapid development of digital banking and decentralized finance (DeFi) has increased the complexity of financial fraud detection. Although machine learning methods are increasingly applied to fraud detection, the literature remains fragmented regarding the comparative performance of different algorithms under highly imbalanced conditions and the implications of their application for decentralized and sustainable financial systems. The aim of this study is to comparatively evaluate selected supervised machine learning models for financial fraud detection and to examine the implications of their application for digital banking, DeFi, and sustainable financial systems. The study uses the publicly available Credit Card Fraud Detection dataset containing anonymized financial transactions with a highly imbalanced distribution between legitimate and fraudulent observations. Three supervised machine learning algorithms were comparatively evaluated: Logistic Regression, Random Forest, and XGBoost. Data preparation and model evaluation were implemented in Python using the Scikit-learn and XGBoost libraries. Model performance was assessed using precision, recall, F1-score, and ROC-AUC. The comparative results indicate that the ensemble-based models, particularly XGBoost, achieved stronger overall performance than Logistic Regression across the selected evaluation metrics. Random Forest also demonstrated competitive performance, supporting the applicability of ensemble approaches to fraud detection in complex and highly imbalanced financial data. The scientific contribution of the study consists in the comparative assessment of three widely used machine learning approaches within a unified fraud detection framework and in linking the obtained findings to the emerging challenges of DeFi and sustainable financial systems. The study also discusses the role of explainable artificial intelligence (XAI) in improving transparency and interpretability in financial machine learning applications. The findings may support financial institutions and digital financial service providers in selecting appropriate machine learning approaches for fraud detection and in balancing predictive performance with transparency and operational requirements. The discussion of DeFi highlights the need for adaptive analytical approaches capable of addressing the specific risks of decentralized financial ecosystems. The study concludes that machine learning, particularly ensemble-based approaches, has significant potential to strengthen fraud detection and contribute to safer, more resilient, and sustainable digital financial systems.

Keywords: machine learning; fraud detection; FinTech; DeFi; blockchain; explainable artificial intelligence; sustainability.

Fig.: 2. **Tabl.:** 2. **Bibl.:** 28.

Caprian Iurie – Postgraduate Student of the Department of Business and Administration, State University of Moldova (60 Alexei Mateevici Str., Kishinev, MD-2009, Moldova)

E-mail: iuriecapiian@gmail.com

ORCID: <https://orcid.org/0000-0001-5484-3087>

Caprian Iulia – PhD, Associate Professor of the Faculty of Economics and Legal Sciences, State University of Moldova (60 Alexei Mateevici Str., Kishinev, MD-2009, Moldova)

E-mail: kaprianusm@gmail.com

ORCID: <https://orcid.org/0000-0001-7701-5356>

Lom Ionut Valentin – Research Associate of the Department of Finance and Banking, National Institute for Aerospace Research “Elie Carafoli” (INCAS) (sector 6, 220 Iuliu Maniu Blvd., Bucharest, 061126, Romania)

E-mail: lom.ionut@incas.ro

УДК 336.7:343:5:004.8
JEL: C45; G21; G28; O33

Капріан І., Капріан І., Лом І. В. Машинне навчання для виявлення шахрайства в цифровому банкінгу та DeFi: прямування до стійких фінансових систем

Швидкий розвиток цифрового банкінгу та децентралізованих фінансів (DeFi) ускладнив виявлення фінансового шахрайства. Хоча методи машинного навчання все частіше застосовуються для такого виявлення, наявна література є фрагментарною щодо порівняльної ефективності різних алгоритмів за умов сильної дисбалансованості, а також наслідків їх використання для децентралізованих і стійких фінансових систем. Метою дослідження була порівняльна оцінка обраних методів контрольованого машинного навчання для виявлення фінансового шахрайства та розгляд наслідків їх застосування для цифрового банкінгу, DeFi та стійких фінансових систем. У дослідженні використовувався загальнодоступний набір даних Credit Card Fraud Detection, що містить анонімізовані фінансові транзакції зі сильно дисбалансованим співвідношенням між достовірними та маніпулятивними спостереженнями. Було порівняно три алгоритми керованого машинного навчання: Logistic Regression, Random Forest та XGBoost. Підготовку даних та оцінку моделей реалізовано на мові Python за допомогою бібліотек Scikit-learn та XGBoost. Продуктивність моделей оцінювалась за точністю, повнотою, показником F1-score та метрикою ROC-AUC. Порівняльні результати показали, що моделі на основі ансамблів, особливо XGBoost, досягли кращих загальних результатів у порівнянні з Logistic Regression за обраними метриками оцінки. Random Forest також продемонстрував конкурентоспроможну продуктивність, що підтверджує застосовність ансамблевих підходів для виявлення шахрайства щодо складних і сильно незбалансованих фінансових даних. Науковий внесок цього дослідження полягає у порівняльній оцінці трьох широко використовуваних підходів машинного навчання в межах єдиної системи виявлення шахрайства та у пов'язуванні отри-

маних результатів із актуальними викликами DeFi та стійким фінансовим системам. Дослідження також розглядає роль пояснювального штучного інтелекту у підвищенні прозорості та зрозумілості у застосуванні машинного навчання до фінансових завдань. Отримані результати можуть бути корисними фінансовим установам та провайдером цифрових фінансових послуг при обранні відповідних підходів машинного навчання для виявлення шахрайства та знаходженні балансу між передбачуваною ефективністю, прозорістю та операційними вимогами. Стосовно DeFi, підкреслюється необхідність адаптивних аналітичних підходів, здатних враховувати специфічні ризики децентралізованих фінансових екосистем. За результатами дослідження зроблено висновок, що машинне навчання, особливо підходи на основі ансамблів, має значний потенціал для поліпшеного виявлення шахрайства та для сприяння створенню безпечніших, більш стійких і резильєнтних цифрових фінансових систем.

Ключові слова: машинне навчання; виявлення шахрайства; FinTech; DeFi; блокчейн; пояснювальний штучний інтелект; стійкість.

Рис.: 2. Табл.: 2. Бібл.: 28.

Капріан Іуріє – аспірант кафедри бізнесу та адміністрування, Державний університет Молдови (вул. Олексія Матеевича, 60, Кишинів, MD-2009, Молдова)

E-mail: iuriecpran@gmail.com

ORCID: <https://orcid.org/0000-0001-5484-3087>

Капріан Іулія – доктор філософії, доцент факультету економіки та правових наук, Державний університет Молдови (вул. Олексія Матеевича, 60, Кишинів, MD-2009, Молдова)

E-mail: kaprianusm@gmail.com

ORCID: <https://orcid.org/0000-0001-7701-5356>

Лом Іонут Валентін – науковий співробітник відділу фінансів і банківської справи, Національний інститут аерокосмічних досліджень «Еліє Карафолі» (INCAS) (бульв. Юліу Маніу, 220, сектор 6, Бухарест, 061126, Румунія)

E-mail: lom.ionut@incas.ro

INTRODUCTION

The accelerated digital transformation of financial systems has profoundly reshaped the structure and functioning of global financial markets. The widespread adoption of online banking, mobile payment systems, and DeFi platforms has improved the efficiency and accessibility of financial services.

These technologies enable real-time transactions, reduce operational costs, and expand access to financial services.

However, “this transformation has introduced new systemic vulnerabilities, in particular in the form of increasingly sophisticated financial fraud” [1; 4]. In traditional banking systems, fraud detection mechanisms have mainly relied on fixed rules and manual verification processes. These systems use predefined thresholds and expert-defined rules, such as transaction value limits or geographical inconsistencies. Although efficient in identifying known fraud patterns, they have limited adaptability and reduced ability to respond to emerging fraud strategies. As fraud techniques continue to evolve, including identity theft, phishing attacks, and transaction laundering, traditional rule-based systems become progressively less efficient.

The exponential increase in the volume, speed, and diversity of financial data further complicates the fraud detection process. Modern financial systems generate massive streams of transactional data, making manual analysis impractical and requiring automated, scalable solutions.

In this context, “machine learning (ML) has emerged as an efficient approach for identifying ab-

normal patterns and detecting fraudulent behavior in complex datasets” [2; 5].

Unlike traditional methods, machine learning models can learn from historical data, adapt to new fraud patterns, and continuously improve their predictive performance.

Supervised learning techniques, including Logistic Regression, Random Forest, and XGBoost, “have demonstrated strong performance in fraud detection, particularly when applied to highly imbalanced datasets” [28].

We highlight that one of the major challenges in detecting fraud is the significant imbalance between legitimate and fraudulent transactions, where fraud cases represent a very small proportion of the total observations. The analysis highlights that this imbalance can lead to models biased towards the majority class, reducing the performance of detection systems. For this reason, appropriate machine learning models and evaluation metrics are required to ensure reliable fraud detection in highly imbalanced datasets.

The development of decentralized finance (DeFi) introduces a new paradigm in the financial ecosystem. DeFi platforms are built on blockchain technology and enable peer-to-peer transactions without centralized intermediaries. The analysis highlights that while this architecture increases transparency and accessibility, it also generates new fraud risks, including smart contract vulnerabilities, flash loan attacks, and market manipulation.

The pseudonymous nature of blockchain transactions further complicates fraud detection, “requiring analytical approaches that differ from those used in traditional financial systems” [14; 15].

A key aspect of modern fraud detection systems is model interpretability. As machine learning algorithms become more complex, understanding how models reach their decisions becomes increasingly important for transparency, accountability, and regulatory compliance. “Explainable artificial intelligence (XAI) provides a general framework for improving the interpretability of machine learning models in financial applications” [3].

Beyond the technical aspects, fraud detection has important implications for financial sustainability. Fraudulent activities generate significant economic losses, reduce trust in financial institutions, and limit the development of inclusive financial ecosystems. Efficient fraud detection systems contribute to financial resilience by reducing losses, improving security, and strengthening trust between users and financial institutions.

- ✦ The analysis highlights that efficient detection systems help to strengthen financial resilience by reducing losses, increasing security and strengthening trust between users and institutions.
- ✦ This study examines the role of machine learning in improving fraud detection in digital banking systems and decentralized finance ecosystems.
- ✦ More specifically, the paper aims to make the following contributions:
- ✦ Benchmarking of several machine learning models to detect fraud in highly unbalanced datasets;
- ✦ Discuss the role of explainable artificial intelligence in improving the transparency and interpretability of fraud detection models.
- ✦ Analyze the implications of these models in the context of decentralized finance (DeFi) and financial sustainability.

The scientific contribution of this study lies in the comparative evaluation of widely used machine learning algorithms for fraud detection in both traditional digital banking and emerging DeFi environments. In addition, the study discusses the role of explainable artificial intelligence (XAI) in improving the transparency of fraud detection models and examines the practical challenges associated with applying these techniques in decentralized financial ecosystems.

The structure of the paper is as follows: introduction, relevant literature on fraud detection and machine learning methods, description of the methodology, including the dataset, pre-processing steps and implementation of models, followed by results of the comparative evaluation, presentation of the implications of the results for financial systems and DeFi, at

the end the conclusions and future research directions are offered.

The aim of this study is to comparatively evaluate selected supervised machine learning models for financial fraud detection and to examine the implications of their application for digital banking, decentralized finance (DeFi), and sustainable financial systems.

The specific objectives are to: (1) compare the performance of Logistic Regression, Random Forest, and XGBoost in fraud detection; (2) assess their performance using precision, recall, F1-score, and ROC-AUC; (3) discuss the role of explainable artificial intelligence in improving transparency and interpretability in financial machine learning applications; and (4) examine the implications of machine learning-based fraud detection for DeFi and sustainable financial systems.

LITERATURE REVIEW

Financial fraud detection has been a central area of both academic research and industrial applications for several decades. With the accelerated digitalization of financial systems, fraud detection approaches have evolved from traditional statistical methods to advanced machine learning and artificial intelligence techniques. This section reviews the existing literature, with particular emphasis on machine learning, explainable artificial intelligence, and fraud detection in decentralized finance (DeFi) ecosystems.

Traditional approaches to fraud detection.

The first fraud detection systems were primarily based on statistical methods and rule-based approaches. These systems relied on predefined rules derived from expert knowledge, such as transaction thresholds, frequency limits, and geographical inconsistencies. Statistical techniques, including Logistic Regression and Bayesian models, “were widely used to identify deviations from normal transaction behavior” [1,4].

Although these methods are efficient in detecting known fraud patterns, they have important limitations. Rule-based systems require continuous manual updates and have limited ability to adapt to new fraud strategies. In addition, the increasing volume and complexity of financial transactions reduce the scalability of manually defined detection rules.

Machine learning in fraud detection.

The limitations of traditional approaches have encouraged the widespread adoption of machine learning techniques for fraud detection. These models are capable of learning complex patterns from historical data and adapting to evolving fraudulent behavior.

“Supervised learning methods, including Logistic Regression, Support Vector Machines (SVM),

Decision Trees, Random Forest, and Gradient Boosting algorithms, have been extensively applied in fraud detection research” [2,5]. Ensemble methods generally achieve higher predictive performance because they capture nonlinear relationships and complex interactions among variables.

Extrême Gradient Boosting (XGBoost) has become one of the most widely used algorithms for fraud detection because of “its computational efficiency and strong predictive performance on imbalanced classification problems” [2]. Several studies report that XGBoost frequently outperforms traditional machine learning models, particularly when appropriate feature engineering and data preprocessing techniques are applied.

A major challenge discussed in the literature is the severe class imbalance characteristic of fraud detection datasets, where fraudulent transactions represent only a small proportion of all observations. To address this issue, “previous studies have proposed techniques such as SMOTE (Synthetic Minority Over-sampling Technique), undersampling, and cost-sensitive learning” [5].

Unsupervised and semi-supervised learning approaches have also been examined for detecting previously unknown fraud patterns. “Clustering algorithms, anomaly detection methods, and autoencoders can identify deviations from normal behavior without requiring labeled data” [12]. However, these approaches often suffer from limited interpretability and higher false positive rates.

Explainable Artificial Intelligence in Financial Systems.

As machine learning models become more complex, a lack of decision-making transparency is becoming a critical issue, especially in regulated areas such as finance. The “black-box” models, although performing, offer limited visibility into how they arrive at decisions.

Explainable artificial intelligence (XAI) aims to improve the transparency of machine learning models by providing interpretable explanations of model predictions. One of the most widely used techniques is SHAP (SHapley Additive Explanations), based on co-

operative game theory, which assigns an importance value to each input variable.

In the context of fraud detection, XAI can support:

- ✦ understanding the factors influencing model predictions;
- ✦ ensuring compliance with regulatory requirements;
- ✦ increasing the level of trust in automated decision-making systems.

Recent studies highlight the need to integrate explainability into “fraud detection systems, especially in contexts where transparency and accountability are key” [3]. However, there is a trade-off between performance and interpretability, as simpler models are easier to interpret but often less performing.

Fraud detection in blockchain and DeFi.

The emergence of blockchain technology and decentralized finance (DeFi) has introduced new challenges and opportunities for fraud detection. “Unlike traditional financial systems, DeFi platforms operate without centralized intermediaries, relying on blockchain technology and smart contracts to execute and validate transactions” [28].

Although blockchain provides transparency through publicly accessible transaction records, it also introduces pseudonymity, making the identification of malicious actors more challenging. Consequently, fraud detection in DeFi environments requires advanced analytical methods. Figure 1 illustrates the main types of fraud encountered in decentralized financial ecosystems.

The analysis highlights that recent research has explored “the use of machine learning for analyzing blockchain data and identifying fraudulent behaviors” [13; 14]. In particular, graph-based methods have shown promise in modeling transaction networks and detecting abnormal structures by representing transactions as nodes and edges in a graph.

However, fraud detection in DeFi remains an emerging research area, largely due to the limited availability of labeled data and the rapid evolution of the ecosystem, which requires adaptive analytical models.

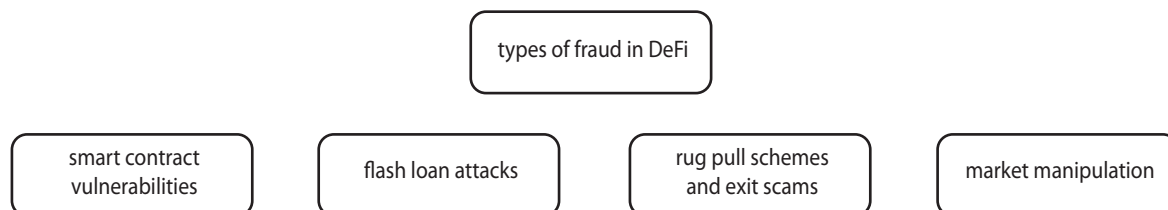


Fig. 1. Fraud types and vulnerabilities in DeFi

Source: Authors' elaboration based on [13; 14].

Gaps in research.

Despite significant advances in the literature, several important research gaps remain. First, most studies focus on traditional financial systems while paying limited attention to decentralized financial ecosystems. Second, the practical integration of explainable artificial intelligence into fraud detection systems remains relatively limited.

In addition, relatively few studies simultaneously address:

- ✦ Model performance on highly unbalanced datasets;
- ✦ interpretability and transparency of predictions;
- ✦ the implications of fraud detection in the context of DeFi and financial sustainability.

In this context, the present study aims to cover these gaps through a comparative analysis of machine learning models for fraud detection, the role of explainable artificial intelligence in improving model transparency and the assessment of their implications in decentralized and sustainable financial ecosystems.

RESEARCH METHODS

This section describes the dataset used, the preprocessing steps, the machine learning models, and the experimental assessment framework. The methodology was designed to provide a structured framework for the comparative evaluation of the selected machine learning models.

DATASET DESCRIPTION

This study uses the publicly available Credit Card Fraud Detection dataset from Kaggle, which contains anonymized credit card transactions made by European cardholders in September 2013. Each transaction is labeled as either legitimate (Class = 0) or fraudulent (Class = 1), enabling supervised binary classification.

The dataset is highly imbalanced, with fraudulent transactions representing less than 1% of all observations. This distribution reflects the characteristics of real-world financial transaction data and represents a major challenge for fraud detection models.

To protect the privacy of cardholders, most input features were transformed using Principal Component Analysis (PCA). As a result, the dataset contains anonymized variables (V1–V28), while only the Time and Amount features remain in their original form. The set includes:

- ✦ Transaction Time
- ✦ Transaction amount
- ✦ a set of anonymized numerical variables (V1–V28)

The target variable is binary, indicating whether a transaction is fraudulent (1) or legitimate (0).

Data preprocessing. Data preprocessing is an essential stage in the development of machine learning models for fraud detection, ensuring that the available data are suitable for subsequent analysis. Prior to model development, the dataset was reviewed to verify its consistency and to prepare the input variables for comparative evaluation.

Data cleaning.

The dataset was examined for missing values and potential data inconsistencies before model development. No missing values were identified, and no additional data cleaning procedures were required.

Feature scaling.

The transaction amount variable was standardized to reduce the influence of differences in scale among the input features. Standardization improves data consistency and facilitates the comparative evaluation of machine learning models.

Class imbalance management.

Fraud detection datasets are typically characterized by a highly imbalanced distribution between legitimate and fraudulent transactions. Consequently, model evaluation focused on complementary performance metrics, including precision, recall, F1-score and ROC-AUC, which provide a more informative assessment than accuracy alone in imbalanced classification problems.

Dataset preparation.

The dataset was organized to enable the comparative evaluation of the selected machine learning models. The available observations were prepared for model development and subsequent performance assessment using the evaluation metrics described below. The methodological framework was designed to provide a consistent comparison of the selected algorithms using the publicly available dataset under the same evaluation conditions.

Machine learning models.

Three supervised machine learning algorithms were selected for comparative evaluation based on their widespread use in fraud detection research and their complementary classification characteristics.

Logistic Regression.

Logistic Regression was selected as a baseline supervised learning algorithm for binary classification. Although relatively simple, it provides a high level of interpretability and serves as an appropriate benchmark for comparing more advanced machine learning models.

Random Forest.

Random Forest is an ensemble learning algorithm that combines multiple decision trees to improve classification performance and reduce the risk of overfit-

ting. Its ability to capture nonlinear relationships and estimate feature importance has made it widely applicable in fraud detection studies.

XGBoost.

XGBoost is a gradient boosting algorithm designed to build predictive models through the sequential optimization of decision trees. Owing to its regularization mechanisms and computational efficiency, it has been widely adopted for fraud detection and other classification tasks involving complex and imbalanced datasets.

Hyperparameter optimization.

The performance of machine learning models is influenced by the selection of model parameters. In the present study, Logistic Regression, Random Forest, and XGBoost were compared using their standard model configurations to provide a consistent comparative evaluation of the selected algorithms. The primary objective of this study was to compare the relative performance of widely used machine learning approaches for fraud detection rather than to optimize the performance of individual models. Future research may extend this work by incorporating systematic hyperparameter optimization and more comprehensive validation strategies.

Given the severe class imbalance, model performance was evaluated using complementary metrics, including precision, recall, F1-score, and ROC-AUC. These metrics provide a more informative assessment than accuracy alone for highly imbalanced fraud detection datasets. Among them, recall is particularly important because it reflects the model's ability to correctly identify fraudulent transactions, thereby reducing the number of undetected fraud cases.

Experimental framework.

The comparative analysis was implemented in Python using the Scikit-learn and XGBoost libraries. The methodological workflow, including data preparation, model implementation, and evaluation metrics, is described to facilitate the interpretation of the comparative results. Future studies may further enhance reproducibility by providing the complete implementation workflow, computational scripts, and model configuration details.

RESULTS

This section presents the results of the comparative evaluation of three supervised machine learning models for fraud detection: Logistic Regression, Random Forest, and XGBoost.

Model performance is assessed using standard classification metrics, followed by an analysis of their practical implications for fraud detection in digital financial systems.

Model performance comparison. Table 1 summarizes the performance of the three evaluated machine learning models.

The results indicate that the ensemble-based models outperformed Logistic Regression across all evaluation metrics. Among the evaluated models, XGBoost showed the strongest overall performance, followed by Random Forest and Logistic Regression. Random Forest also achieved strong and balanced results across the evaluated metrics. Although Logistic Regression achieved lower overall performance, it remains a valuable baseline model due to its simplicity and high interpretability.

Table 1 shows that XGBoost achieved the highest values of precision, F1-score, and ROC-AUC among the evaluated models, while XGBoost and Random Forest achieved the same recall value of 0.89.

Precision–recall trade-off. In fraud detection, an appropriate balance between precision and recall is essential. While high recall reduces the number of undetected fraudulent transactions, high precision helps limit false alerts and unnecessary operational costs. The results presented in Table 1 indicate that XGBoost provides the most balanced performance among the evaluated models, making it a suitable choice for fraud detection applications.

Implications for fraud detection systems. The comparative evaluation suggests that ensemble-based models, particularly XGBoost, can improve the identification of fraudulent transactions in digital financial systems.

Their ability to achieve a balanced performance across multiple evaluation metrics makes them suitable for supporting fraud detection processes. In addition, explainable artificial intelligence represents a promising direction for improving the transparency

Table 1

Performance comparison of the evaluated machine learning models

Model	Precision	Recall	F1-Score	ROC-AUC
Logistic regression	0.84	0.84	0.84	0.84
Random Forest	0.89	0.89	0.89	0.89
XGBoost	0.92	0.89	0.90	0.94

Source: Authors' own elaboration.

and interpretability of machine learning models in practical financial applications.

DISCUSSION

The results demonstrate the effectiveness of machine learning models, particularly ensemble-based approaches, for detecting fraudulent financial transactions. Beyond quantitative performance metrics, these findings should be interpreted within the broader context of financial systems, operational constraints, and ongoing technological developments.

The strong performance of XGBoost can be attributed to its ability to capture complex non-linear relationships and iteratively reduce classification errors. In fraud detection, recall is particularly important because false negative predictions represent fraudulent transactions that remain undetected and may result in substantial financial losses. From a practical standpoint, these findings support the use of models that prioritize the identification of fraudulent transactions, even at the cost of a moderate increase in false positive predictions.

The analysis highlights that Random Forest also offers high performance, especially in terms of stability and robustness. Its ensemble structure helps reduce variance and the risk of overfitting, making it suitable for dynamic financial environments. From the authors' perspective, logistic regression, although inferior in terms of predictive performance, remains relevant due to the high interpretability and simplicity of the model, essential aspects in regulated contexts where decision-making transparency is mandatory.

An important aspect highlighted by the results is the need to balance predictive performance with model interpretability.

The analysis highlights that while complex algorithms such as XGBoost offer superior accuracy, their decision-making processes are less transparent.

Although explainable artificial intelligence was not implemented in the present study, it represents a promising direction for improving the transparency and interpretability of machine learning models in financial applications.

This dimension is particularly important in the financial field, where automated decisions need to be justifiable and transparent to regulators and stakeholders.

DeFi and blockchain integration: fraud detection in decentralized systems

The emergence of decentralized finance (DeFi) has transformed the financial landscape by creating new opportunities for digital financial services while introducing novel fraud risks that differ substantially from those encountered in traditional banking sys-

tems. From the authors' perspective, unlike traditional financial systems, DeFi platforms operate without centralized intermediaries, relying on blockchain technology and smart contracts to execute and validate transactions.

Although blockchain technology provides a high level of transaction transparency through publicly accessible distributed ledgers, it also preserves user pseudonymity by relying on cryptographic addresses rather than verified identities. As a result, traditional identity-based fraud detection methods become less efficient, requiring analytical approaches based on transaction patterns, network structures, and behavioral characteristics.

According to the authors, fraud in DeFi ecosystems manifests itself in various forms, including:

- ✦ Vulnerabilities and exploits of smart contracts;
- ✦ Flash loan attacks, in which attackers manipulate market conditions in a single atomic transaction;
- ✦ Pull-outs, in which developers abandon projects after attracting capital from investors;
- ✦ Market manipulation schemes, including wash trading and pump-and-dump mechanisms.

In this context, machine learning offers promising solutions for detecting fraudulent behavior. Graph-based methods are particularly suitable for modeling blockchain transaction networks, where nodes represent addresses and edges correspond to transactions. Analyzing the structure and dynamics of these networks may facilitate the identification of abnormal patterns associated with fraudulent activities.

The analysis further suggests that real-time fraud detection in DeFi ecosystems requires models capable of processing large volumes of transactions in continuous flow and quickly adapting to ever-evolving attack strategies. The analysis highlights that this requires the integration of machine learning techniques with real-time data processing systems and blockchain analysis tools.

From a practical perspective, machine learning models can be integrated into blockchain monitoring platforms to identify suspicious transaction patterns, detect anomalous wallet behavior, and generate early warnings for potential fraudulent activities. Such solutions may support exchanges, DeFi protocols, and regulatory authorities by improving transaction monitoring and strengthening risk management without compromising the decentralized nature of blockchain networks.

From the authors' perspective, several major challenges remain. The lack of labeled data in DeFi environments is found to limit the applicability of su-

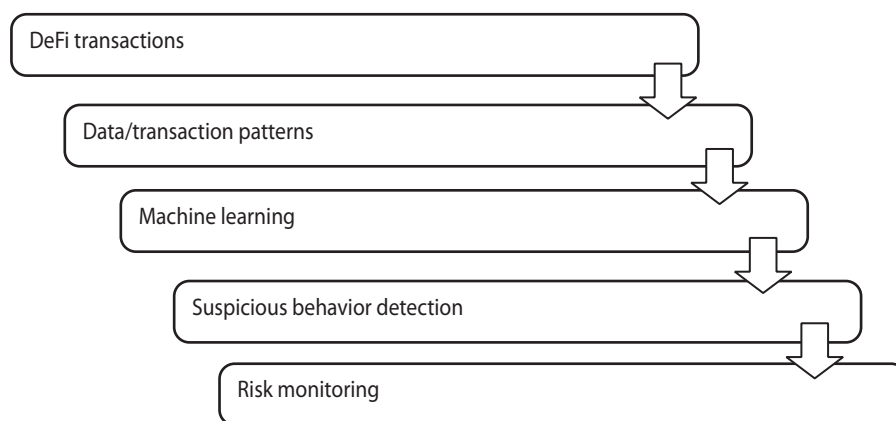


Fig. 2. Conceptual framework for machine learning-based fraud detection in DeFi

Source: Authors' conceptual framework based on [13; 14; 26; 27].

pervised learning methods, while the rapid evolution of smart contracts requires frequent model updates. The analysis highlights that the decentralized nature of these systems raises additional issues related to ac-

countability, governance and the application of the legal framework. Table 2 summarizes the main challenges and potential machine learning applications in DeFi fraud detection.

Table 2

Machine Learning Applications and Challenges in DeFi Fraud Detection

DeFi challenge	Machine learning application	Expected contribution
Pseudonymous transactions	Transaction-pattern analysis	Identification of suspicious behavior
Evolving fraud patterns	Adaptive machine learning models	Earlier detection of emerging risks
Smart-contract vulnerabilities	Anomaly detection	Identification of unusual transaction patterns
Large transaction networks	Graph-based analysis	Detection of abnormal network structures
Limited labeled data	Semi-supervised and unsupervised methods	Improved detection with limited labeled observations

Source: Authors' own elaboration based on [13; 14; 26; 27].

Implications for sustainable financial systems.

The findings of this study suggest that the results have relevant implications for the development of safe, transparent and sustainable financial systems. Financial fraud not only generates direct economic losses but also undermines trust in financial institutions and digital infrastructures. From this perspective, the authors consider that efficient fraud detection mechanisms are an essential element for maintaining the integrity and stability of financial markets.

From a technological point of view, the results indicate that the adoption of machine learning models can significantly improve the ability to identify fraudulent behaviors. The analysis highlights that by leveraging big data and advanced analytical methods, financial institutions can achieve more accurate and efficient detection compared to traditional rules-based systems.

Explainable artificial intelligence represents an important direction for improving transparency and accountability in automated decision-making pro-

cesses. Current compliance requirements place an increasing emphasis on transparency and accountability in automated decision-making processes. Explainable models allow algorithmic decisions to be justified, facilitating regulatory compliance and building user trust.

In the context of decentralized finance (DeFi), the analysis highlights that the implications are even more significant. As DeFi ecosystems continue to expand, ensuring their security becomes a major challenge. From this perspective, the authors believe that fraud detection methods based on machine learning can help reduce risks and support responsible technological development.

From a sustainability perspective, reducing fraudulent activities contributes to strengthening the resilience of financial systems. The analysis highlights that safer infrastructures encourage user participation, stimulate innovation and support long-term economic development. In addition, the integration

between artificial intelligence and blockchain technologies has the potential to support more transparent, efficient, and inclusive financial ecosystems.

LIMITATIONS AND FUTURE RESEARCH

Despite its contributions, the present study has some limitations. First, the analysis is based on a public dataset, which may not fully reflect the complexity and dynamics of financial systems in the real environment. Secondly, the results obtained are influenced by the controlled experimental conditions, which may differ from the operational scenarios encountered in practice.

Future research directions should, from the authors' perspective, aim to:

- ✦ applying machine learning models to real-time financial data;
- ✦ exploring graph-based and deep learning approaches to fraud detection;
- ✦ the development of adaptive models capable of responding dynamically to the evolution of fraudulent strategies;
- ✦ extending analytics to real-world datasets in DeFi ecosystems.

“Combining machine learning techniques with blockchain infrastructures can strengthen the security and sustainability of digital financial systems.” [28]

CONCLUSIONS

This study examined the application of machine learning techniques for fraud detection in digital banking and decentralized finance (DeFi) environments. The comparative evaluation indicates that ensemble-based models, particularly XGBoost, achieved stronger overall performance than Logistic Regression when applied to highly imbalanced financial transaction data.

An important implication of the study is the role of recall in fraud detection. High recall contributes to reducing the number of undetected fraudulent transactions, while an appropriate balance between recall and precision remains necessary to limit false positive alerts and maintain operational efficiency.

Explainable artificial intelligence (XAI) represents an important direction for improving the transparency and interpretability of machine learning models in financial applications. This aspect is particularly relevant in regulated environments, where automated decisions should be explainable, auditable, and consistent with regulatory requirements. Previous research also emphasizes the potential of XAI to strengthen trust in automated decision-making and support regulatory compliance [28].

The findings also highlight the relevance of machine learning for emerging DeFi ecosystems, where decentralized architectures, pseudonymous transactions, and evolving fraud patterns create additional challenges for financial monitoring. In this context, machine learning combined with blockchain analytics may contribute to the development of more adaptive approaches to fraud detection [26; 27].

From a broader perspective, improving fraud detection can contribute to the resilience, security, and sustainability of digital financial systems by reducing potential financial losses and strengthening confidence in digital financial infrastructures. However, the present study does not experimentally evaluate blockchain-based fraud detection or XAI methods; these aspects are considered as directions for extending the current research.

Future research should therefore focus on real-time financial data, graph-based and deep learning methods, PR-AUC and other metrics suitable for highly imbalanced datasets, as well as real-world DeFi transaction data. Further examination of the relationship between predictive performance, model interpretability, and regulatory compliance may also strengthen the practical applicability of machine learning-based fraud detection systems. ■

BIBLIOGRAPHY

1. Bhattacharyya S., Jha S., Tharakunnel K., Westland J. C. Data Mining for Credit Card Fraud: A Comparative Study. *Decision Support Systems*. 2011. Vol. 50. No. 3. P. 602–613. DOI: <https://doi.org/10.1016/j.dss.2010.08.008>
2. Chen T., Guestrin C. XGBoost: A scalable tree boosting system. *Y: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco, CA, 2016. P. 785–794. DOI: <https://doi.org/10.1145/2939672.2939785>
3. Lundberg S. M., Lee S. I. A Unified Approach to Interpreting Model Predictions. *Y: Advances in Neural Information Processing Systems (NeurIPS)*. 2017. P. 4765–4774. URL: <https://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions>
4. Ngai E. W. T., Hu Y., Wong Y. H., Chen Y., Sun X. The Application of Data Mining Techniques in Financial Fraud Detection: A Classification Framework and Academic Review of Literature. *Decision Support Systems*. 2011. Vol. 50. No. 3. P. 559–569. DOI: <https://doi.org/10.1016/j.dss.2010.08.006> URL: <https://www.sciencedirect.com/science/article/pii/S0167923610001302>
5. Dal Pozzolo A., Boracchi G., Caelen O., Alippi C., Bontempì G. Credit Card Fraud Detection: A Realistic

- Modeling and a New Learning Strategy. *IEEE Transactions on Neural Networks and Learning Systems*. 2018. Vol. 29. No. 8. P. 3784–3797.
DOI: <https://doi.org/10.1109/TNNLS.2017.2736643>
6. Carcillo F., Le Borgne Y. A., Caelen O., Bontempi G. Combining unsupervised and supervised learning in credit card fraud detection. *Information Sciences*. 2021. Vol. 557. P. 317–331.
DOI: <https://doi.org/10.1016/j.ins.2019.05.042>
7. Nakamoto S. Bitcoin: A Peer-to-Peer Electronic Cash System. 2008. URL: <https://bitcoin.org/bitcoin.pdf>
8. Goodfellow I., Bengio Y., Courville A. *Deep Learning*. Cambridge, MA : MIT Press, 2016. URL: <https://www.deeplearningbook.org>
9. Fiore U., De Santis A., Perla F., Zanetti P., Palmieri F. Using Generative Adversarial Networks to Improve Classification Effectiveness in Credit Card Fraud Detection. *Information Sciences*. 2019. Vol. 479. P. 448–455.
DOI: <https://doi.org/10.1016/j.ins.2017.12.030>
10. Kourtney Y., Lu C. T., Sirwongwattana S., Huang Y. P. Survey on Fraud Detection Techniques. Y: *Proceedings of the IEEE International Conference on Networking, Sensing and Control*. 2004. P. 749–754.
DOI: <https://doi.org/10.1109/ICNSC.2004.1297040>
11. Phua C., Lee V., Smith K., Gayler R. A Comprehensive Survey of Data Mining-based Fraud Detection Research. *arXiv preprint*. arXiv:1009.6119, 2010. URL: <https://arxiv.org/abs/1009.6119>
12. Ahmed M., Mahmood A. N., Hu J. A Survey of Network Anomaly Detection Techniques. *Journal of Network and Computer Applications*. 2016. Vol. 60. P. 19–31.
DOI: <https://doi.org/10.1016/j.jnca.2015.11.016>
13. Weber M., Domeniconi G., Chen J., Weidele D. K., Bellei C., Robinson T., Leiserson C. E. Anti-Money Laundering in Bitcoin: Experimenting with Graph Convolutional Networks. *arXiv preprint*. arXiv:1908.02591, 2019. URL: <https://arxiv.org/abs/1908.02591>
14. Kshetri N. Blockchain's Roles in Strengthening Cybersecurity and Protecting Privacy. *Telecommunications Policy*. 2017. Vol. 41. No. 10. P. 1027–1038.
DOI: <https://doi.org/10.1016/j.telpol.2017.09.003>
15. Schär F. Decentralized Finance: On Blockchain- and Smart Contract-Based Financial Markets. *Federal Reserve Bank of St. Louis Review*. 2021. Vol. 103. No. 2. P. 153–174.
DOI: <https://doi.org/10.20955/r.103.153-74>
16. Arner D. W., Barberis J., Buckley R. P. FinTech and RegTech: Impact on Regulators and Banks. *Journal of Banking Regulation*. 2018. Vol. 19. No. 4. P. 329–344.
DOI: <https://doi.org/10.1057/s41261-017-0037-1>
17. Esteva A., Kuprel B., Novoa R. A., et al. Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks. *Nature*. 2017. Vol. 542. P. 115–118.
DOI: <https://doi.org/10.1038/nature21056>
18. Lundberg S. M., Erion G., Lee S. I. Consistent Individualized Feature Attribution for Tree Ensembles. *arXiv preprint*. arXiv:1802.03888, 2018. URL: <https://arxiv.org/abs/1802.03888>
19. Aggarwal C. C. *Outlier Analysis*. 2nd ed. Cham : Springer, 2017.
DOI: <https://doi.org/10.1007/978-3-319-47578-3>
20. Kingma D. P., Welling M. Auto-Encoding Variational Bayes. *arXiv preprint*. arXiv:1312.6114, 2013. URL: <https://arxiv.org/abs/1312.6114>
21. Brown T., Mann B., Ryder N., et al. Language Models are Few-Shot Learners. Y: *Advances in Neural Information Processing Systems*. 2020.
DOI: <https://doi.org/10.48550/arXiv.2005.14165>
22. Zohar A. Bitcoin: Under the Hood. *Communications of the ACM*. 2015. Vol. 58. P. 104–113.
DOI: <https://doi.org/10.1145/2701411>
23. Cong L. W., He Z. Blockchain Disruption and Smart Contracts. *Review of Financial Studies*. 2019. Vol. 32. No. 5. P. 1754–1797.
DOI: <https://doi.org/10.1093/rfs/hhz007>
24. Allen F., Gu X., Jagtiani J. A Survey of FinTech Research and Policy Discussion. *Review of Corporate Finance*. 2021. Vol. 1. No. 3–4. P. 259–339.
DOI: <https://doi.org/10.1561/114.000000008>
25. McNally S., Roche J., Caton S. Predicting the Price of Bitcoin Using Machine Learning. *26th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP)*. 2018.
DOI: <https://doi.org/10.1109/PDP2018.2018.00060>
26. Ertam F. Detecting Ethereum Fraud in Near Real-Time Using Explainable Artificial Intelligence in Blockchain Networks. *Applied Sciences*. 2025. Vol. 15. Art. 10841.
DOI: <https://doi.org/10.3390/app152110841>
27. Ghnemat R., Mosa H. Blockchain-Based Fraud Detection: A Systematic Review of Ethereum Network Applications. *Cluster Computing*. 2025. Vol. 28. Art. 1080.
DOI: <https://doi.org/10.1007/s10586-025-04962-9>
28. Caprian I. A Hybrid Modular Architecture for Fraud Detection Using Offline and Online Machine Learning Models. *The Problems of Economy*. 2025. No. 3. P. 312–320.
DOI: <https://doi.org/10.32983/2222-0712-2025-3-312-320>

References

- Aggarwal C. C. (2017). *Outlier Analysis*. 2nd ed. Cham: Springer. <https://doi.org/10.1007/978-3-319-47578-3>
- Ahmed M., Mahmood A. N. & Hu J. (2016). A Survey of Network Anomaly Detection Techniques. *Journal of Network and Computer Applications*, 60, 19–31. <https://doi.org/10.1016/j.jnca.2015.11.016>
- Allen F., Gu X. & Jagtiani J. (2021). A Survey of FinTech Research and Policy Discussion. *Review of Corporate Finance*, 3–4(1), 259–339. <https://doi.org/10.1561/114.000000008>
- Arner D. W., Barberis J. & Buckley R. P. (2018). FinTech and RegTech: Impact on Regulators and Banks. *Journal of Banking Regulation*, 4(19), 329–344. <https://doi.org/10.1057/s41261-017-0037-1>

- Bhattacharyya S., Jha S., Tharakunnel K. & Westland J. C. (2011). Data Mining for Credit Card Fraud: A Comparative Study. *Decision Support Systems*, 3(50), 602–613. <https://doi.org/10.1016/j.dss.2010.08.008>
- Brown T., Mann B., Ryder N. & et al. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*. <https://doi.org/10.48550/arXiv.2005.14165>
- Caprian I. (2025). A Hybrid Modular Architecture for Fraud Detection Using Offline and Online Machine Learning Models. *The Problems of Economy*, 3, 312–320. <https://doi.org/10.32983/2222-0712-2025-3-312-320>
- Carrillo F., Le Borgne Y. A., Caelen O. & Bontempi G. (2021). Combining unsupervised and supervised learning in credit card fraud detection. *Information Sciences*, 557, 317–331. <https://doi.org/10.1016/j.ins.2019.05.042>
- Chen T. & Guestrin C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). <https://doi.org/10.1145/2939672.2939785>
- Cong L. W. & He Z. (2019). Blockchain Disruption and Smart Contracts. *Review of Financial Studies*, 5(32), 1754–1797. <https://doi.org/10.1093/rfs/hhz007>
- Dal Pozzolo A., Boracchi G., Caelen O., Alippi C. & Bontempi G. (2018). Credit Card Fraud Detection: A Realistic Modeling and a New Learning Strategy. *IEEE Transactions on Neural Networks and Learning Systems*, 8(29), 3784–3797. <https://doi.org/10.1109/TNNLS.2017.2736643>
- Ertam F. (2025). Detecting Ethereum Fraud in Near Real-Time Using Explainable Artificial Intelligence in Blockchain Networks. *Applied Sciences*, 15, Art. 10841. <https://doi.org/10.3390/app152110841>
- Esteva A., Kuprel B., Novoa R. A. & et al. (2017). Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks. *Nature*, 542, 115–118. <https://doi.org/10.1038/nature21056>
- Fiore U., De Santis A., Perla F., Zanetti P. & Palmieri F. (2019). Using Generative Adversarial Networks to Improve Classification Effectiveness in Credit Card Fraud Detection. *Information Sciences*, 479, 448–455. <https://doi.org/10.1016/j.ins.2017.12.030>
- Ghnemat R. & Mosa H. (2025). Blockchain-Based Fraud Detection: A Systematic Review of Ethereum Network Applications. *Cluster Computing*, 28, Art. 1080. <https://doi.org/10.1007/s10586-025-04962-9>
- Goodfellow I., Bengio Y. & Courville A. (2016). *Deep Learning*. Cambridge, MA: MIT Press. <https://www.deeplearningbook.org>
- Kingma D. P. & Welling M. (2013). Auto-Encoding Variational Bayes. *arXiv preprint. arXiv:1312.6114*. <https://arxiv.org/abs/1312.6114>
- Kourtney Y., Lu C. T., Sirwongwattana S. & Huang Y. P. (2004). Survey on Fraud Detection Techniques. *Proceedings of the IEEE International Conference on Networking, Sensing and Control* (pp. 749–754). <https://doi.org/10.1109/ICNSC.2004.1297040>
- Kshetri N. (2017). Blockchain's Roles in Strengthening Cybersecurity and Protecting Privacy. *Telecommunications Policy*, 10(41), 1027–1038. <https://doi.org/10.1016/j.telpol.2017.09.003>
- Lundberg S. M., Erion G. & Lee S. I. (2018). Consistent Individualized Feature Attribution for Tree Ensembles. *arXiv preprint. arXiv:1802.03888*. <https://arxiv.org/abs/1802.03888>
- Lundberg S. M. & Lee S. I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems (NeurIPS)* (pp. 4765–4774). <https://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions>
- McNally S., Roche J. & Caton S. (2018). Predicting the Price of Bitcoin Using Machine Learning. *26th Euro-micro International Conference on Parallel, Distributed and Network-Based Processing (PDP)*. <https://doi.org/10.1109/PDP2018.2018.00060>
- Nakamoto S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. <https://bitcoin.org/bitcoin.pdf>
- 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 Academic Review of Literature. *Decision Support Systems*, 3(50), 559–569. <https://doi.org/10.1016/j.dss.2010.08.006>
- Phua C., Lee V., Smith K. & Gayler R. (2010). A Comprehensive Survey of Data Mining-based Fraud Detection Research. *arXiv preprint. arXiv:1009.6119*. <https://arxiv.org/abs/1009.6119>
- Schär F. (2021). Decentralized Finance: On Blockchain and Smart Contract-Based Financial Markets. *Federal Reserve Bank of St. Louis Review*, 2(103), 153–174. <https://doi.org/10.20955/r.103.153-74>
- Weber M., Domeniconi G., Chen J., Weidele D. K., Bellei C., Robinson T. & Leiserson C. E. (2019). Anti-Money Laundering in Bitcoin: Experimenting with Graph Convolutional Networks. *arXiv preprint. arXiv:1908.02591*. <https://arxiv.org/abs/1908.02591>
- Zohar A. (2015). Bitcoin: Under the Hood. *Communications of the ACM*, 58, 104–113. <https://doi.org/10.1145/2701411>
- This is the standard phrasing used in academic funding-disclosure statements.
- Authors' contributions:
- Caprian Iurie: conceptualization, methodology, software development, data analysis, drafting of the initial version.
- Caprian Iulia: supervision, validation, critical review and editing, project administration.
- Lom Ionuț Valentin – visualization, technical review, validation and editing.
- This matches the standard wording used in journal AI-disclosure statements
- The authors declare that there is no conflict of interest.
- Стаття надійшла до редакції / Received: 03.08.2026
- Статтю прийнято до публікації / Accepted: 18.08.2026
- Оприлюднено / Published: 08.10.2026