Magna Scientia Advanced Research and Reviews

eISSN: 2582-9394

Cross Ref DOI: 10.30574/msarr

Journal homepage: https://magnascientiapub.com/journals/msarr/

(REVIEW ARTICLE)

Check for updates

From detection to prevention: A review of end-to-end machine learning pipelines for fraud management in credit union operations

Evelyn Agyei [1,*] and Mary Magdalene Yeboah [2]

[1] *Connecticut, USA.*
[2] *Texas, USA.*

Magna Scientia Advanced Research and Reviews, 2026, 17(01), 417-427

Publication history: Received on 15 May 2026; revised on 25 June 2026; accepted on 27 June 2026

Article DOI: https://doi.org/10.30574/msarr.2026.17.1.0113

Abstract

Machine learning has emerged as a transformative force in financial security, offering institutions the ability to anticipate and neutralize threats before they materialize. Against this backdrop, this study reviews the transition from traditional fraud detection to prevention-oriented strategies in credit union operations through the use of end-to-end machine learning pipelines. It reviews literature on the main stages of fraud management pipelines, including data collection, preprocessing, feature engineering, model training, validation, deployment, and continuous monitoring. Key machine learning techniques for fraud detection, including supervised, unsupervised, ensemble, hybrid, and anomaly detection models, are also covered in the paper, along with an assessment of their relevance to credit union operations. The findings show that machine learning can improve fraud detection accuracy, support real-time decision-making, and strengthen proactive risk management, but its effectiveness depends on data quality, scalable infrastructure, governance, and explainability. Key problems with false positives, model drift, privacy, and regulatory compliance are further highlighted by the study. Overall, the paper makes the case that effective fraud management in credit unions necessitates not just sophisticated analytical tools but also rigorous ethical protections, institutional preparedness, and ongoing adaptability to evolving fraud behaviors.

Keywords: Fraud; Credit Union; Detection; Prevention; Machine Learning (ML)

1. Introduction

Financial fraud's growing complexity and scale have driven financial institutions to explore advanced technological solutions, particularly artificial intelligence, to strengthen their fraud prevention and detection capabilities (Dawood, 2025). Fraud in financial institutions, especially credit unions, poses a significant and evolving threat to the stability and trustworthiness of the financial system. It involves many activities, including investment fraud, financial statement manipulation, internal misconduct, accounting fraud, credit card fraud, insurance fraud, loan fraud, and cyber-enabled crimes (Patel et al., 2018). The rapid increase of fraud attacks on banking systems, financial institutions, and even credit card holders demonstrate the high demand for enhanced fraud detection (FD) systems for these attacks. And there is the need for continuous development of AI models that could be alert to the latest fraudulent activities (Hafez et al., 2025; Amoako et al., 2025). Hence, the digitalization of their systems.

Digital transformation refers to the process by which organizations use technology to enhance their business performance. Since the 1990s, there has been rapid growth in the transformation and computerization of processes, a trend that continues to this day (Schneider et al., 2021). This transformation is constantly evolving due to the introduction of the Internet of Things (IoTs) which provides a new way for people to interact with technology driven by artificial intelligence. The financial or commercial environments stand out having undergone changes in the way they

* Corresponding author: Evelyn Agyei.

interact with money, which has transitioned from being exclusively physical to virtual, relying on continuous technological advancements to maintain operations. These changes have led to socioeconomic shifts and transformed the relationship between customers and financial institutions. Consequently, there has been a significant increase in the use of digital means for financial solutions such as payments, transfers, and even fund deposits (Lana et al., 2025). Investments in technology have proven to be highly necessary and effective in making credit unions more competitive in the market to improve relationships with members, reducing costs, and increasing operating results.

However, the perpetrators have embraced the use of this technology to commit fraud. Through emails, social networking sites, false websites and data breaches, offenders are too easily and conveniently defraud victims globally. The evolution of the online environment has significantly changed the nature of fraud as the shift has multiplied the attack surface available to fraudsters, victims and third-party perspectives (Cross et al., 2022).

Credit unions, like other financial institutions, face a wide spectrum of fraudulent activities that threaten both their financial stability and member trust. Credit unions are susceptible to various types of fraud including payment fraud, loan fraud, identity theft and synthetic identity fraud and insider fraud. Credit unions still rely on traditional rule-based fraud systems that are limited in integrating into sophisticated fraud techniques, which mostly results in high false positives and delayed responses. These systems operate largely within a reactive detection framework, identifying fraud only after it has occurred rather than preventing it in real time (Abdallah et al., 2016). Several studies have underscored the use of machine learning for fraud detection to its prevention. However, to the best of our knowledge, no studies have provided a comprehensive review of end-to-end machine learning pipelines for fraud management, specifically within the context of credit unions. Therefore, this review aims to examine end-to-end ML pipelines in fraud management, explore the shift from detection-oriented approaches to proactive prevention, and identify the specific challenges faced by credit unions.

2. Machine learning pipelines for fraud detection

2.1. Architecture of End-to-End Fraud Detection Pipelines

The rapid expansion of digital payments, e-commerce, and connected systems has intensified the risk of cyber fraud, spanning phishing, account takeover, financial fraud, and Internet of Things (IoT)/Industrial Control Systems (ICS) manipulation. Prevention focused approaches, including privacy preserving learning, cyber threat intelligence pipelines and hybrid human artificial intelligence (AI) frameworks, remain underexplored despite their potential to reduce fraud incidence and financial loss. In the financial domain, comparative and hybrid machine learning (ML) have made gains from ensembles and feature rich pipelines for credit card and transaction fraud, alongside calls for real time monitoring and scalable deployment (Mabitsela et al., 2025).

Credit unions traditionally rely on reactive detection systems to identify fraud and credit risk events after suspicious activity has already occurred. These methods are primarily rule-based or driven by a threshold, which relies on preset factors such as unusual transaction size, geographic deviation, or prior delinquency history. Though rule-based monitoring is operationally paramount due to interpretability and regulatory transparency, reactive systems frequently struggle to modify rapidly changing fraud classification and complex behavioral models. Recent surveys show that static rule engines generate high false-positive rates and are less effective in detecting sophisticated or previously unseen fraud schemes (Li et al., 2021). As such, a purely reactive framework may safeguard institutional stability, but at the expense of delayed maintenance and consumer discomfort. The End-to-end machine learning (ML) fraud detection pipelines offers a structured framework for converting raw, unprocessed transactional data into relevant fraud alerts. Data collection, preprocessing, feature engineering, model training, validation, deployment, and ongoing monitoring are stages integrated into these pipelines. Ensuring the accuracy, scalability, and adaptation of fraud detection systems to evolving fraud patterns is the ultimate. Recent experimental and system-oriented research studies on real-time fraud detection and ML operational pipelines have shown that the efficiency of the whole pipeline architecture is equally important for successful fraud detection as sophisticated algorithms.

2.2. End-to-End Fraud Detection Pipelines Framework

2.2.1. Data Collection

The first stage involves gathering data from sources such as payment transactions, merchant details, cardholder/account profiles, device fingerprints, geolocation, login/activity logs, and external risk indicators or threat intelligence feeds. Modern fraud detection systems often collect both structured and unstructured data to provide a broad view of user behavior. Since fraud detection models mainly depend on data quality and coverage, Kunder et al.

(2026) emphasized the importance of effective data collection techniques. Real-time streaming technologies are increasingly used to support instant fraud assessment during transaction processing.

2.2.2. Data Preprocessing

Although fraudulent transactions are rare, unprocessed financial data frequently includes missing values, duplicate records, mismatched formats, and a significant class imbalance. By cleaning, standardizing, encoding categorized variables, and balancing datasets using procedures like oversampling or synthetic data generation, data preparation tackles these issues. (Talukder et al., 2024) emphasizes that preprocessing is essential for reducing noise and improving model performance, especially in large-scale financial settings where data heterogeneity is typical.

2.2.3. Feature Engineering

Feature engineering converts unprocessed transactional data into useful indicators of fraud. Common features include transaction frequency, expenditure trends, geographic irregularities, device fingerprints, velocity indicators, and customer behavioral metrics Effective feature engineering according to (Alsagri, 2025), significantly enhances identifying anomalies by enabling models to distinguish between suspicious and legitimate activity. Feature selection methods are also to decrease dimensionality and boost computational effectiveness.

2.2.4. Model Training

Machine learning techniques are trained to detect transactions as authentic or fraudulent after features have been developed. Traditional techniques include Logistic Regression, Decision Trees, Random Forests, and Support Vector Machines, while more modern systems increasingly employ ensemble and deep learning techniques to enhance detection accuracy and minimize false positives, (Talukder et al., 2024) recommend a multistage ensemble design in which multiple models work collaboratively. Such architecture tackles the complex and dynamic nature of fraud by utilizing the strengths of multiple algorithms.

2.2.5. Model Validation

Validation guarantees that models that have been trained can effectively adapt to new data. Accuracy, precision, recall, F1-score, Area Under the Recall and precision are frequently regarded as more useful than overall accuracy because fraud datasets are frequently extremely unbalanced. (Kunder et al., 2026) emphasizes on the importance of cross-validation and temporal validation techniques to evaluate model robustness under evolving fraud types.

2.2.6. Deployment

With post validation, models are placed into operation in real-time or batch transaction analysis in production scenarios. Deployment architectures often include APIs, cloud-based services, and streaming platforms capable of handling high transaction volumes. To safeguard against interfering with legitimate client activity, production systems must strike a balance between detection accuracy and low latency. Multiple-stage deployment systems enable businesses to incorporate swift initial screening with more thorough secondary analysis for high-risk transactions (Talukder et al., 2024).

2.2.7. Monitoring and Feedback Loops

Fraud patterns continuously evolve, making ongoing monitoring essential. Data drift, concept drift, and performance degradation are all observed in production models. The system reintegrates consumer complaints, verified fraud incidents, and feedback from fraud analysts to accelerate model retraining and improvement. (Sohel et al., 2024) demonstrate that real-time monitoring systems substantially boost the performance of long-term detection by enabling quick adjustments to emerging fraud schemes. In dynamic financial contexts, ongoing feedback loops guarantee the effectiveness of fraud detection pipelines.

2.3. Machine Learning Models Used in Fraud Detection

2.3.1. Supervised learning

Supervised learning (SL) is a machine learning approach that leverages labelled data to educate a system in forecasting outcomes based on its training. It closely mimics the process of human learning under the guidance of an instructor, employing specific instances to deduce overarching principles (Shetty et al., 2022). SL is typically divided into two main categories. That is the regression and classification supervised learning. Regression is a term used in statistics, which is a type of statistical analysis that aims to understand the relationship between a dependent variable (response variable)

and one or more independent variables (predictors) such as in market trends or weather forecasting. The most common type is linear regression. Classification is a SL technique that involves categorizing data into distinct classes. It is a recursive process that recognizes and groups data objects into pre-defined categories or labels. This technique is used to predict the outcome of a given problem based on input features. It can be applied to structured or unstructured data, and the classes are commonly known as target, label, or categories. The aim of classification is to assign an unknown pattern to a known class. For example, classifying emails as "spam" or "not spam" is a common application of classification. Both the Classification and Regression algorithms can be used for forecasting in machine learning and operate with the labelled datasets. But the distinction between classification vs regression is how they are used on particular machine learning problems. (Albaldawi et al., 2024)

2.3.2. Unsupervised learning

Unsupervised learning involves training the machine without knowing the output, using only input samples or labels. It discovers patterns in data and creates its own data clusters. There are two main types and most common unsupervised learning algorithms, clustering and association analysis. This technique is useful for identifying unknown patterns in data, such as recommendation engines on online stores that rely on unsupervised machine learning, specifically a technique called clustering (Samreen Naeem et al., 2023).

2.3.3. Ensemble learning

Ensemble learning techniques have achieved state-of-the-art performance in diverse machine learning applications by combining the predictions from two or more base models. Ensemble learning models and techniques play a crucial role in credit card fraud detection. These approaches involve combining multiple individual models to create a more robust and accurate fraud detection system. Ensemble methods, such as bagging, boosting, and stacking, are commonly employed to address challenges like imbalanced datasets and to enhance overall predictive power. By leveraging the strengths of diverse base models, ensemble techniques contribute to improved fraud detection performance, reducing the risk of false positives and false negatives. The adaptability and effectiveness of ensemble learning make it a valuable strategy in the continuous battle against credit card fraud. (Khalid et al., 2024)

2.3.4. Hybrid models and Anomaly detection systems

Semi-Supervised Learning (SSL) is a powerful technique that combines supervised and unsupervised learning to improve learning accuracy. It addresses the challenge of limited labelled data by leveraging the abundance of unlabeled data. SSL finds applications in diverse fields like speech analysis, web content classification, protein sequence classification, and text document classifiers. (Albaldawi et al., 2024). An anomaly is a data pattern that does not conform to expected behavior and may indicate faults, attacks, fraud, or rare events. Anomaly detection systems aim to automatically find unusual patterns in data that deviate from what is considered normal. These systems are now central in security, industrial monitoring, finance and healthcare, especially as data volumes grow and real-time responses are needed.

2.4. Feature Engineering and Data Challenges

2.4.1. Class imbalance and imbalanced fraud datasets

Fraud detection systems operate in highly complex and adversarial environments. These models must endure rare fraud occurrences, noisy transactional data, and complicated or complex behavior patterns. Consequently, datasets used for model development are typically characterized by severe class imbalance, high dimensionality, noisy records, temporal dependencies, and evolving fraud patterns (Vallarino, 2025; Mienye & Jere, 2024). The utilization of data and machine learning algorithms has become increasingly prevalent in contemporary society, playing a vital role in the extraction of insights and predictions from large and complex datasets (Rinku et al., 2024). However, a common challenge encountered in the field of machine learning is the phenomenon of class imbalance, defined as a disparity in the number of samples pertaining to different classes (Siam et al., 2025). In view of the prevalence of error rate as the primary metric which guides classifier optimization, the presence of class imbalance can inadvertently favor the majority class, thereby diminishing the effectiveness of classification algorithms (Vasconcelos & Cavique, 2026). To address this issue, several methodologies have been developed and can be broadly classified into four groups: (1) data-level approaches, (2) algorithmic-level approaches, (3) cost-sensitive approaches, and (4) ensemble learning approaches (Compagnino et al., 2025; Hernandez Aros et al., 2024). Data-level approaches involve modifying the data distribution to amplify the relative importance of the minority class during model training. This is achieved through random or informative resampling techniques, such as undersampling or oversampling. Such approaches are widely employed due to their versatility, allowing for their application irrespective of the specific learning model or classification context (Kennedy et al., 2024). In contrast, algorithmic-level approaches focus on adapting existing learning algorithms to mitigate the bias associated

with the majority class, without altering the training set. Strategies encompass adjusting the decision threshold, training class-specific classifiers, weighting-based approaches, among others. While algorithmic-level approaches offer enhanced model interpretability by incorporating user goals directly into the classification algorithms, they demand a comprehensive theoretical understanding of learning algorithms and optimization mechanisms. Moreover, they are constrained to algorithms explicitly adapted for this purpose (Mienye & Jere, 2024). On the contrary, cost-sensitive approaches address the issue of class imbalance by introducing asymmetric classification costs between classes into the modelling process. These costs, typically represented as a cost matrix, aim to prioritize the correct classification of minority instances during model training. However, an inherent challenge associated with this approach is the reliable determination of the cost matrix, which may evolve over time and whose context-dependent misclassification values are very often unknown or difficult to ascertain (Rashid et al., 2025).

2.4.2. Feature selection

Feature selection (FS) is an essential phase in the implementation of machine learning techniques, as it enables the extraction of the most relevant attributes for accurate classification. This is mainly due to the data set employed during the training and testing phases potentially possessing a vast feature space, which may negatively affect the models' overall performance

2.4.3. Data quality

Fraud detection models are only as good as the data they rely on. Research shows that poor data quality directly reduces fraud detection accuracy, increases false alarms, false-positive rates and slows investigations. Fraudulent transactions make up very little of total observations in financial transaction data sets, which frequently have missing values, duplicate records, inconsistent formats, noise, and severely skewed class distributions. To address these challenges, data preprocessing is a critical step before model development. Common techniques in ensuring data quality include data cleaning by handling missing values, eliminating duplicates and rectifying inconsistences. The data can also be transformed to improve model performance by scaling and normalizing numerical variables. Another technique is feature encoding which transforms categorical variables into machine-decoded formats. Paramount to these techniques is class balancing which can be done by utilizing techniques such as cost-sensitive learning, undersampling, Adaptive synthetic sampling (ADASYN), and Synthetic Minority Oversampling Technique (SMOTE). The outlier treatment can also be used in identifying and handling exceptional data that may skew normal supervised learning. Research studies highlight the criticality of feature optimization and extraction during preprocessing (Siam et al., 2025).

With the objective to increase the significance of fraud indicators, (Siam et al., 2025) proposed a hybrid feature-selection framework that combines statistical filtering and machine learning techniques. By choosing complementing feature subsets from several methods, (Talukder et al., 2024) constructed an ensemble feature-optimization strategy that improves predictive performance. In a similar vein, (Alsagri, 2025) created a multi-stage feature-extraction framework that may produce more in-depth transaction representations for the purpose of categorizing fraud.

2.4.4. Transaction velocity features

Transaction velocity features capture the frequency and intensity of customer activities within a specified time window. Velocity features may include number of transactions within the last hour, day or week, total transaction amount over a defined period, average transaction value within a window, number of failed payment attempts, number of transactions across multiple merchants, geographic movement speed between transaction locations and sudden increases in transaction frequency compared with historical patterns. These features assist to detect account takeovers, card-testing attacks, money laundering activities, and other forms of financial fraud characterized by abnormal transaction surge.

2.4.5. Behavioral analytics

Behavioral analytics is understanding and modeling normal customer behavior to identify deviations that may indicate fraudulent activity. Some behavioral indicators include typical spending amounts and frequencies, preferred merchants and transaction categories, usual transaction times and days, device usage, login characteristics and historical risk scores and customer profiles. Behavioral analytics create personalized customer profiles and continuously track variations in transaction patterns, in contrast to rule-based systems.

3. From detection to prevention

3.1. Real-Time Fraud Detection and Preventive Analytics

Fraud systems are transitioning from detecting fraud after it occurs to detecting and preventing it beforehand. Richer data, real-time analytics, and AI-driven behavioral models integrated with transaction flows are all vital to this evolution. The limitations of conventional rule-based systems with post-event inspections led to the development of reactive detection. These transaction monitoring applies statics rules and analyzes past transactions, often detecting fraud only after losses occur (Akinyemi et al., 2025). Also, traditional systems had high false positives and poor adaptability due to fixed thresholds and generic scenarios which often than not miss new patterns and generate many unnecessary alerts. (Prakash et al., 2026) Furthermore, assessment in traditional systems is mostly retrospective and has limited real-time intervention ability because it is done post-event and on a batch basis. Thus, preventive analytics and real-time fraud recognition were introduced. Streaming data, machine learning, and behavioral modeling are used in real-time fraud and preventive analytics to identify and halt suspicious transactions as they transpire. In order to prevent real consumers from being barred, research focuses on scaling to high-velocity streams, identifying both known and new fraud, and lowering false positives.

3.1.1. Elements of Proactive prevention

There are various elements that constitute proactive prevention to help detection fraud in real-time and prevent losses to the organization. Models based on machine learning such as autoencoders, Convolutional Neural Networks (CNNs), Long Short-Term Memory networks (LSTMs), and anomaly detectors are used in preventive behavioral analytics to build behavioral profiles and identify deviations in real time, allowing for action before money is lost. By dynamically profiling each user, behavioral analytics enhance the early detection of identity theft and abnormal behaviors. (Tiwari et al., 2025). Massive transaction streams are processed by real-time risk engines with latency ranging from microseconds to milliseconds, facilitating quick decisions (block, step-up auth, review). Composite risk scores target high-risk entities for quick response by combining anomaly detectors, velocity attributes, and behavioral indicators. (Udeh et al., 2024).

Automated alert systems enable swift response to unusual activities by notifying fraud analysts or triggering predetermined security actions. By enabling prompt action, automated monitoring frameworks minimize response times and increase the effectiveness of fraud mitigation initiatives. (Sohel et al., 2024) Monitoring the real-time feed improves early warning and adaptive controls. Before evident fraud is detected, early warning systems leverage temporal pattern analysis to predict suspect changes in consumer behavior and initiate increased surveillance or controls. (Bello et al., 2024) Adaptive authentication and dynamic security measures adjust friction based on current risk score. (Roy et al., 2024) *3.1.2 AI-Driven Risk Management in Credit Unions*

Artificial intelligence (AI) has transformed risk management practices in credit unions through the development of a more comprehensive and proactive approach to identifying, assessing and mitigating financial risks. Credit unions can develop unified risk management systems that enhance operational effectiveness, member safety, and regulatory compliance thanks to the integration of fraud detection and credit risk management capabilities in modern AI-powered systems (Amoako et al., 2025). By evaluating vast amounts of transactional, behavioral, and financial data, machine learning algorithms can help with both fraud detection and credit risk management. This approach enables organizations to identify high-risk borrowers while also identifying any suspicious activities that may indicate fraud. (Surana, 2024)

In credit unions, predictive models support strategic planning by identifying members, loans, or accounts that exhibit elevated risk levels before adverse events occur. According to Agyei and Lettu (2026), predictive analytics enables credit unions to move from reactive risk management toward proactive risk prevention by providing early warning indicators that support timely intervention. AI-driven member profiling builds detailed risk profiles based on demographic, transactional, behavioral and financial features. Credit unions can better analyze member behavior and improve risk management measures with the aid of these profiles. By identifying subtle fraud signs that conventional rule-based systems can miss, machine learning algorithms may accelerate the approval of legitimate loans while minimizing financial losses.

3.2. Operational Risk Reduction

Technological disruptions, cybersecurity threats, human error, and internal procedure failures all result in operational risks. AI decreases operational risk through intelligent anomaly recognition, automation, and continuous monitoring. Vanini et al. (2023) emphasize that integrating AI-powered fraud risk management within enterprise governance

frameworks strengthens organizational resilience and supports effective risk oversight. The integration of AI technologies into credit union operations provides several strategic benefits, including improved accuracy of credit and fraud risk evaluations, faster identification of suspicious activities, and heightened member protection. Credit unions may construct an integrated AI-driven risk ecosystem that can handle changing financial threats by integrating predictive analytics, member profiling, loan fraud prevention, and operational risk management.

3.3. Blockchain, Privacy, and Decentralized Fraud Prevention

As financial institutions increasingly rely on data-driven fraud detection systems, concerns regarding data privacy, security, and regulatory compliance have become more prevalent. Traditional centralized machine learning techniques sometimes include gathering and storing private client data in a single repository, creating challenges with cybersecurity threats, privacy protection, and legal obligations. Innovative approaches that facilitate collaborative fraud detection while protecting data confidentiality and governance requirements are provided by emerging technologies like blockchain, federated learning, and privacy-preserving artificial intelligence (AI). The goal of decentralized fraud prevention is to identify fraud spanning institutions while protecting sensitive financial information locally and in accordance with privacy laws. In order to facilitate cross-entity collaboration, explainability, and regulatory defensibility, recent work incorporates federated learning, blockchain, and privacy-preserving AI.

3.3.1. Federated Learning

Federated learning is a decentralized machine learning technique that permits credit unions and other financial institutions to collaboratively train models without sharing raw data; only model variables or updates are exchanged. Keeping data local and respecting privacy laws. Online banking, credit card fraud, financial crime in general, DeFi, and cryptocurrency fraud are all included in this framework (AbouGrad & Sankuru, 2025). Local nodes that preprocess data and train local models on private transactions may be part of a federated learning architecture for credit unions. The blockchain layer then assigns new parameters and combines encrypted or masked updates into a global model. Finally, customers can use smart environments or devices for running lightweight models for real-time screening (Koroma et al., 2025)

According to Bello & Iyelolu (2024), integrating blockchain with machine learning enhances fraud detection by ensuring data integrity while providing trustworthy information for predictive analytics models. Blockchain-based systems can also improve accountability and reduce opportunities for data manipulation as its distributed ledger architecture enhances transparency, traceability, and trust among participating entities.

3.3.2. Privacy-Preserving Artificial intelligence

Algorithms that enable businesses to use machine learning while minimizing the exposure of sensitive data are incorporated in privacy-preserving AI. These methods address growing concerns regarding data privacy, cybersecurity, and ethical AI deployment. Federated learning, differential privacy, safe multiparty computation, and homomorphic encryption are a few mechanisms that can preserve privacy. AbouGrad & Sankuru (2025) emphasize that privacy-focused machine learning architectures can help financial institutions meet regulatory obligations while maintaining effective fraud detection capabilities. For fraud management systems to guarantee the quality, security, accessibility, and moral use of data, effective data governance is vital. Organizations must set up governance frameworks that specify roles, guidelines, and data management controls as decentralized technologies proliferate. Decentralized fraud protection solutions are kept reliable, safe, and in line with company goals and legal requirements thanks to strong governance frameworks.

4. Implementation challenges, ethics, and operational considerations

4.1. Technical and Operational Challenges

Machine-learning fraud systems face not only modeling issues but also deep technical and operational constraints that affect scalability, accuracy, cost, and maintainability.

4.1.1. Infrastructure and Scalable Limitations

One crucial technical and operational challenge in the implementation of artificial intelligence is its infrastructure and scalability limits. High-frequency trading (HFT) and online payments generate massive, high-velocity streams; deep and ML models need substantial computational resources and sophisticated infrastructure to keep up in real time (Vanini et al., 2023). Also, real-time deep learning in HFT must balance accuracy with low latency, requiring efficient architecture and often hybrid methods to stay scalable (Oluwabunmi et al., 2024). Further, cloud-based

experimentation exposes practical limits, compute caps, instability, and long training/testing times that complicate moving toward real-time deployment (Khalid et al., 2024). These real-time deep learning also drive-up cost. High data volume and millisecond decision requirements drive significant computing. Storage, and engineering costs for deep models in HFT fraud (Oluwabunmi et al., 2024). Ensemble models achieving high accuracy must still be optimized for training/testing time and memory to be viable for real-time fraud screening

4.1.2. False positives, Model Drift and Evolving behavior

Effective fraud systems must detect as many frauds as possible "without producing too many false alarms. False positives are a central challenge in online payment fraud (Vanini et al., 2023). A payment-fraud ML model coupled with economic optimization achieved a very low false positive rate of 0.4%, highlighting how tightly FP rates are tied to operational viability (Vanini et al., 2023). Ensemble approaches explicitly aim to mitigate false positives and false negatives but still face trade-offs with computational efficiency. Fraud in payments and trading is dynamic, with "ever-improving attacks of fraudsters" and dynamic fraud behavior that require adaptable models. Reviews identify concept drift and evolving fraud tactics as persistent challenges that demand continual retraining and adaptive strategies (Khan et al., 2025).

4.2. Ethical, Regulatory, and Explainability Concerns

Implementing ML in financial fraud prevention must handle integration with existing systems, regulatory constraints, and operational processes (Pan, 2024). Centralized, black-box ML is criticized as rigid, semi-effective, and poorly suited for complex financial environments, corresponding integration and trust (Khan et al., 2025). To guarantee responsible deployment, financial institutions, particularly credit unions, must address significant ethical, regulatory, and explainability issues posed by the use of artificial intelligence in fraud management. (Kumar et al., 2024; Veldurthi, 2025). One major ethical issue is algorithmic bias. When fraud detection models are trained on past transaction data, they may duplicate preexisting institutional or societal biases, particularly if the data demonstrate skewed member behavior patterns, missing records, or unequal treatment. Unfair repercussions may result from this, such as the disproportionate flagging of specific categories of customers, incorrect transaction declines, or more inspection of members who are valid. (Kumar et al., 2024). According to recent research on AI in banking, fairness, accountability, and transparency are necessary for responsible AI adoption since biased or opaque systems can erode public confidence and put organizations at risk for litigation and reputational loss. (Veldurthi, 2025; Aljunaid et al., 2025). In parallel, regulators and institutional governance frameworks are paying greater focus to how AI models are managed in financial services. Research indicates that risk management, consumer protection, privacy, and model oversight processes should be used to regulate AI, with higher standards for documentation, validation, monitoring, and human oversight. (Pan, 2024; Vanini et al., 2023). This is especially crucial for credit unions since they need to be aware of the dangers of using third-party AI services, model drift, poor data quality, and automated decision-making. (Pan, 2024). Sensitive financial and personal data handling is another area of regulatory concern. Organizations has to ensure legitimate data collection, purpose limitation, access control, and safe processing because fraud systems rely on huge quantities of transactional, behavioral, and identity-related data. (AbouGrad & Sankuru, 2025; Koroma, 2025).

In order to enhance fraud detection while reducing avoidable member data exposure, privacy-preserving techniques like federated learning, differential privacy, and secure model-sharing frameworks are being suggested more frequently. (AbouGrad and Sankuru, 2025; Koroma, 2025; Aljunaid et al., 2025). Explainability is equally critical because fraud analysts, auditors, and affected customers need understandable reasons for why a transaction was flagged or blocked. In regulated scenarios, black-box models are more difficult to analyze, challenge, and defend despite their exceptional predictive accuracy. (Aljunaid et al., 2025; Vanini et al., 2023). Therefore, explainable artificial intelligence (AI) technologies are essential for converting model outputs into human-understandable explanations, assisting organizations in supporting analyst decision-making, spotting model flaws, and boosting customer trust. (Aljunaid et al., 2025). Explainability should not be treated as a purely technical add-on; rather, it is a governance requirement that connects model performance with accountability, fairness, and compliance (Kumar et al., 2024; Veldurthi, 2025). In actuality, credit unions require a comprehensive approach that incorporates human inspection, transparent governance, accurate fraud detection, and ethical standards. Regular bias testing, model audits, transparent escalation protocols for disputed decisions, and documentation that shows how AI systems are created, verified, and tracked over time are typical instances of this. (Pan, 2024; Vanini et al., 2023).

5. Recommendations and future directions

5.1. Strategic Recommendations for Credit Unions

Based on the findings of this review, credit unions should adopt a phased and strategic approach based on artificial intelligence fraud management. Since accurate fraud detection is determined by complete, timely, and well-managed data, they should first strengthen data governance by improving data quality, integration, and security. Second, in order to address changing fraud strategies, credit unions should make investments in scalable machine learning pipelines that enable real-time monitoring, feedback loops, and frequent model upgrades. Third, in order to properly assess high-risk alerts and minimize false positives without compromising member safety, institutions should integrate automated detection with human surveillance. Fourth, explainability and fairness should be incorporated into the creation and application of models through frequent bias testing, open documentation, and explicit escalation protocols for choices that are contested. Fifth, because many credit unions have limited technological assets, staff training programs, shared service platforms, and collaborations with reliable technology vendors can help increase internal capabilities while preserving costs under control. Lastly, rather than being independent technical activity, fraud prevention should be viewed as an enterprise-wide approach involving compliance, operations, cybersecurity, and member service.

5.2. Future Research Directions

Future studies should go beyond model comparison studies and concentrate more on the efficacy of fraud prevention systems in actual credit union environments. Since a substantial portion of the literature presently in publication focuses on large commercial banks, general payment systems, or credit card datasets that could not accurately reflect the operational reality of member-owned organizations, additional empirical research on credit unions in particular is required. Additionally, more study is required on decentralization and safeguarding privacy strategies, such as enabling blockchain-based cooperation and federated learning, particularly in situations when organizations need to exchange intelligence without disclosing private member data. In addition, future studies should examine how explainable AI can be implemented in practice to support auditors, fraud analysts, regulators, and customers. Analyzing the operational cost of false positives, analyzing the long-term implications of model drift, and determining how AI-based fraud systems might be linked with more comprehensive credit risk and business risk management frameworks are other relevant ways to proceed. Overall, future scholarships should emphasize practical deployment, governance, and institutional impact so that fraud prevention research can better support sustainable innovation in credit unions.

6. Summary of key findings

This review has shown that fraud management in credit union operations is shifting from traditional reactive detection toward more proactive and preventive approaches built on end-to-end machine learning pipelines. Data collection, preprocessing, feature engineering, model training, confirmation deployment, and ongoing monitoring are all important phases of these pipelines, according to the literature. When contrasted to traditional rule-based systems, the findings show that machine learning models, including supervised, unsupervised, ensemble, hybrid, and anomaly detection approaches, can enhance fraud detection accuracy and facilitate real-time decision-making. The study also emphasizes that the value of machine learning depends not just on model performance but also on data quality, pipeline design, and institutions' capacity to adjust to emerging fraud trends. The paper also highlights the specific operational challenges that credit unions confront, such as insufficient facilities, outdated systems, false positives, model drift, and compliance demands. Ethical and governance concerns, especially around bias, privacy, transparency, and explainability, remain central to responsible AI adoption. Overall, the evidence suggests that successful fraud management in credit unions requires a balanced integration of technical capability, institutional governance, and member-centered safeguards

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

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

[2] AbouGrad, H., & Sankuru, L. (2025). Online banking fraud detection model: Decentralized machine learning framework to enhance effectiveness and compliance with data privacy regulations. *Mathematics, 13*(13), 2110.

[3] Agyei, E., & Lettu, E. (2026). Machine learning in credit unions: A review of predictive analytics for fraud detection and loan risk assessment.

[4] Akinyemi, T. (2025). Employing real-time analytics in fraud detection and prevention mechanisms for financial transactions.

[5] Aljunaid, S., Almheiri, S., Dawood, H., & Khan, M. (2025). Secure and transparent banking: Explainable AI-driven federated learning model for financial fraud detection.

[6] Alnuaimi, A., & Albaldawi, T. (2024). An overview of machine learning classification techniques.

[7] Alsagri, H. (2025). Hybrid machine learning-based multi-stage framework for detection of credit card anomalies and fraud.

[8] Bello, H., & Iyelolu, T. (2024). Integrating machine learning and blockchain: Conceptual frameworks for real-time fraud detection and prevention.

[9] Bello, H., Idemudia, C., & Iyelolu, T. (2024). Implementing machine learning algorithms to detect and prevent financial fraud in real time.

[10] Compagnino, G., et al. (2025). A reinforcement learning framework for fraud detection in highly imbalanced financial data. *Applied Sciences, 16*(1), 252. https://doi.org/10.3390/app16010252

[11] Cross, C. (2022). Meeting the challenges of fraud in a digital world.

[12] Hafez, I. Y., Hafez, A. Y., Saleh, A., Abd El-Mageed, A. A., & Abohany, A. A. (2025). A systematic review of AI-enhanced techniques in credit card fraud detection.

[13] Hernandez Aros, L., et al. (2024). Year-over-year developments in financial fraud detection via deep learning: A systematic literature review. *arXiv preprint arXiv:2502.00201*.

[14] HM, D. (2025). Strategic adoption of AI for fraud prevention in financial institutions: A systematic literature review. *SSRN Electronic Journal*. https://ssrn.com/abstract=5262760

[15] Kennedy, R. K. L., Villanustre, F., Khoshgoftaar, T. M., & Salekshahrezaee, Z. (2024). Synthesizing class labels for highly imbalanced credit card fraud detection data. *Journal of Big Data, 11*(1), 1–22. https://doi.org/10.1186/s40537-024-00897-7

[16] Khalid, A. R., Owoh, N., Uthmani, O., Ashawa, M., Osamor, J., & Adejoh, J. (2024). Enhancing credit card fraud detection: An ensemble machine learning approach.

[17] Koroma, A. (2025). An integrated privacy-preserving AI framework for real-time financial fraud detection.

[18] Kumar, N., Sharma, V., & Sharma, A. (2024). Artificial intelligence in fraud prevention: Exploring techniques, applications, challenges, and opportunities.

[19] Kunder, H., Shetty, R., Salian, A., Dilip, H., & S, V. (2026). End-to-end fraud detection pipeline: Leveraging supervised, ensemble and unsupervised learning for large-scale financial security.

[20] Lana, M. I., Bressan, V. G. F., & Francisco, J. R. D. S. (2025). The influence of digital technologies on the economic and financial performance of Brazilian credit unions.

[21] Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2021). A survey of convolutional neural networks: Analysis, applications, and prospects.

[22] Mabitsela, I., & Matome, M. J. (2025). Defending against fraud: Cyber fraud detection and prevention techniques.

[23] Mienye, I. D., & Jere, N. (2024). Deep learning for credit card fraud detection: A review of algorithms, challenges, and solutions. *IEEE Access, 12*, 96893–96910. https://doi.org/10.1109/ACCESS.2024.XXXXX

[24] Naeem, S., Ali, A., Anam, S., & Ahmed, M. (2023). Unsupervised machine learning algorithms: A comprehensive review. *International Journal of Computing and Digital Systems*. https://doi.org/10.12785/ijcds/130172

[25] Oluwabunmi, H., Bello, H., Bolatito, A., & Ameyaw, M. (2024). Deep learning in high-frequency trading: Conceptual challenges and solutions for real-time fraud detection.

[26] Pan, E. (2024). Machine learning in financial transaction fraud detection and prevention.

[27] Patel, H., Parikh, S., Patel, A., & Parikh, A. (2018). An application of ensemble random forest classifier for detecting financial statement manipulation of Indian listed companies.

[28] Prakash, S. (2026). Predictive behavioral analytics: A machine learning framework for proactive fraud mitigation in digital payments.

[29] Rashid, M. T., et al. (2025). Evaluation of cost-sensitive learning models in forecasting business failure of capital market firms. *Mathematics, 13*(3), 368. https://doi.org/10.3390/math13030368

[30] Rinku, A. K. D., Narang, S. K., & Kishore, N. (2024). A machine learning-based approach for fraud detection in imbalanced credit card transaction dataset. *SSRG International Journal of Electronics and Communication Engineering, 11*(8), 244–259. https://doi.org/10.14445/23488549/IJECE-V11I8P124

[31] Roy, N., & P, S. (2024). Proactive cyber fraud response: A comprehensive framework from detection to mitigation in banks.

[32] Schneider, S., & Kokshagina, O. (2021). Digital transformation: What we have learned (thus far) and what is next.

[33] Shetty, S., Shetty, S., Singh, C., & Rao, A. S. (2022). Supervised machine learning: Algorithms and applications. In *Fundamentals and Methods of Machine and Deep Learning*. https://doi.org/10.1002/9781119821908.ch1

[34] Siam, A. I., Bhowmik, P., & Uddin, M. (2025). Hybrid feature selection framework for enhanced credit card fraud detection using machine learning models.

[35] Siam, A. I., et al. (2025). An introduction to machine learning methods for fraud detection. *Applied Sciences, 15*, 11787.

[36] Sohel, A., Alam, M., Waliullah, M., Siddiki, A., & Uddin, M. (2024). Fraud detection in financial transactions through data science for real-time monitoring and prevention.

[37] Surana, P. (2024). Credit risk assessment and fraud detection in financial transactions using machine learning.

[38] Talukder, M., Khalid, M., & Uddin, M. (2024). An integrated multistage ensemble machine learning model for fraudulent transaction detection.

[39] Tiwari, S. (2025). Enhancing financial crime detection through data science-driven transaction monitoring: A comprehensive framework for modern financial institutions.

[40] Udeh, E., Amajuoyi, P., Adeusi, K., & Scott, A. (2024). The role of big data in detecting and preventing financial fraud in digital transactions.

[41] Vallarino, D. (2025). Detecting financial fraud with hybrid deep learning: A mix-of-experts approach to sequential and anomalous patterns. *arXiv preprint arXiv:2504.03750*.

[42] Vanini, P., Rossi, S., Zvizdic, E., & Domenig, T. (2023). Online payment fraud: From anomaly detection to risk management.

[43] Vasconcelos, M. O., & Cavique, L. (2026). Imbalanced learning in corruption detection: Results explanations with SHAP. *International Journal of Data Science and Analytics, 21*, 22. https://doi.org/10.1007/s41060-025-00889-7

[44] Veldurthi, A. (2025). The role of AI and machine learning in fraud detection for financial services.

[45] Amoako, E. K., Okeke, P. T., Boateng, V., Bonsu, M. A., & Oman-Amoako, M. (2025). Enhancing risk management and fraud detection in the U.S. financial industry through machine learning algorithms: Applications, challenges, and future directions. *World Journal of Advanced Research and Reviews, 28*(3), 134–144.

[46] Amoako, E. K. W., Boateng, V., Ajay, O., Adukpo, T. K., & Mensah, N. (2025). Exploring the role of machine learning and deep learning in anti-money laundering (AML) strategies within US financial industry: A systematic review of implementation, effectiveness, and challenges. *Finance & Accounting Research Journal*, *7*(1), 22-36.