



Machine Learning-Based Fraud Detection in Financial Records Using Anomaly Identification and Predictive Risk Scoring

Sadia Zaman¹;

[1]. Management Information Systems, College of Business, Lamar University, USA;
Email: zaman.sadia1311@gmail.com

Doi: [10.63125/ccmr1g09](https://doi.org/10.63125/ccmr1g09)

Received: 18 January 2026; **Revised:** 25 February 2026; **Accepted:** 20 March 2026; **Published:** 08 June 2026

Abstract

This study examined the problem of reactive, rule-bound, and opaque fraud detection in financial records, where fraudulent transactions and manipulated accounting entries are frequently identified only after loss has occurred, where fixed rule engines generate high false-positive rates, and where machine-learning models, when deployed, often function as black boxes that analysts cannot interpret or trust. The purpose of the study was to assess how a Machine Learning-Based Fraud Detection framework, integrating unsupervised anomaly identification with supervised predictive risk scoring, influences fraud detection performance in financial records. A quantitative, cross-sectional, case-based design was adopted, and data were collected through a structured five-point Likert-scale questionnaire from 142 valid respondents out of 160 distributed questionnaires, representing an 88.8% valid response rate. The sample included fraud analysts and investigators, data scientists and machine-learning engineers, auditors and forensic accountants, risk and compliance officers, and financial-systems managers, with 66.9% directly involved in fraud-detection or risk-scoring activities. The key variables were anomaly detection modeling, predictive risk scoring, feature engineering and data quality, model validation and performance, analyst and institutional engagement, framework design quality, and fraud detection performance. The analysis plan included descriptive statistics, reliability testing using Cronbach's alpha, Pearson correlation, regression modeling, a framework maturity index, and a fraud risk-control priority matrix. The headline findings showed that all major constructs were rated high, with fraud detection performance recording the highest mean score of 4.21, followed by predictive risk scoring at 4.15 and framework design quality at 4.11. Reliability was acceptable to excellent, with Cronbach's alpha values ranging from 0.81 to 0.93. Correlation results showed significant positive relationships, including $r = 0.78$ between framework design quality and fraud detection performance. Regression results confirmed that the model explained 71.4% of the variance in fraud detection performance, $R^2 = 0.714$, adjusted $R^2 = 0.702$, $F(6,135) = 56.12$, $p < 0.001$. Framework design quality was the strongest predictor, $\beta = 0.34$, followed by predictive risk scoring, $\beta = 0.26$, and anomaly detection modeling, $\beta = 0.24$. The findings imply that financial institutions should strengthen anomaly detection modeling, calibrated risk scoring, feature engineering, model validation, and analyst engagement to improve the accuracy, transparency, and trustworthiness of machine-learning fraud detection.

Keywords

Fraud detection; Anomaly detection; Predictive risk scoring; Machine learning; Financial records.

INTRODUCTION

Fraud detection in financial records refers to the process of identifying transactions, journal entries, or account behaviors that are illegitimate, deceptive, or indicative of financial crime, using methods that distinguish these rare and deliberately concealed events from the far larger volume of legitimate activity. In modern financial institutions, fraud detection is closely associated with the analysis of payment and card transactions, account-takeover attempts, financial-statement manipulation, and money-laundering patterns, whose undetected occurrence can produce direct monetary loss, regulatory penalties, and erosion of institutional trust (ACFE, 2022). Machine learning provides a data-driven foundation for this task by learning statistical patterns of normal and fraudulent behavior from historical records, rather than relying solely on fixed rules (Ngai et al., 2011). Within a machine-learning framework, fraud detection is commonly organized around two complementary strategies: unsupervised anomaly identification, which flags observations that deviate markedly from learned patterns of normal behavior, and supervised predictive risk scoring, which assigns each record a calibrated probability of being fraudulent based on labeled examples (Chandola et al., 2009). In practical terms, combining anomaly identification with predictive risk scoring treats fraud detection not as a static rule-matching exercise but as an adaptive, evidence-based, and auditable process in which statistical signals, model reasoning, and analyst judgment are reconciled rather than kept separate.

Formally, unsupervised anomaly detection assigns to each record an anomaly score that measures how far the observation departs from a model of normal behavior, and flags the record when that score exceeds a threshold. For a reconstruction-based detector such as an autoencoder, the anomaly score is the reconstruction error between the input and its reconstruction, as given in Equation (1):

$$a(x) = \|x - \hat{x}\|_2^2, \quad y_{\text{flag}} = \mathbb{I}[a(x) > \tau] \quad (1)$$

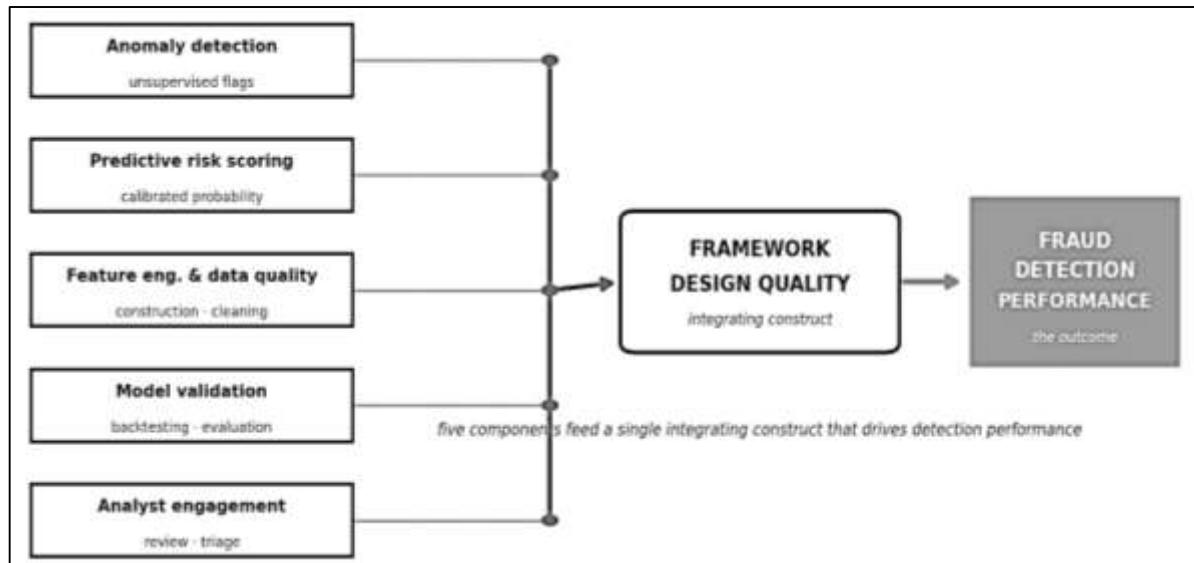
Here, x is the observed record, $\hat{x}$ is its reconstruction from the learned model of normal behavior, $a(x)$ is the resulting anomaly score, and τ is a decision threshold above which the record is flagged for further scoring. Because fraudulent records are rare and were never used to define normal behavior, they tend to reconstruct poorly and therefore receive high anomaly scores, which makes reconstruction error an effective, label-free first-stage filter (Pang et al., 2021).

Isolation-based methods provide a complementary, tree-structured route to anomaly scoring that is well suited to high-dimensional financial data. An isolation forest scores a record by how easily it can be separated from the rest of the data through random partitioning, on the premise that anomalies require fewer splits to isolate, as expressed in Equation (2):

$$s(x, n) = 2^{-\mathbb{E}[h(x)]/c(n)}, \quad c(n) = 2H(n-1) - \frac{2(n-1)}{n} \quad (2)$$

In Equation (2), the expected path length to isolate a record is normalized by an average path length term, so that scores near one indicate likely anomalies and scores well below one indicate normal points (Liu et al., 2008). Together, reconstruction-based and isolation-based detectors give the framework a robust, unsupervised first stage that narrows the vast volume of records to a smaller candidate set for supervised scoring.

Figure 1: Machine Learning–Based Fraud Detection Framework Using Anomaly Identification



The significance of a machine-learning approach is connected to the scale, velocity, and adversarial nature of financial fraud. Financial institutions process enormous volumes of transactions in which fraudulent events are extremely rare, so detection systems must find a very small number of illicit records among millions of legitimate ones without overwhelming investigators with false alarms (Dal Pozzolo et al., 2015). Fraud is also adversarial and non-stationary (Hasan & Uddin, 2022; Mohiul & Badrul, 2022): perpetrators continuously change their methods to evade detection, which causes the statistical patterns of fraud to drift over time and degrades models that are not retrained (Gama et al., 2014). Traditional rule-based systems, though transparent, struggle with this scale and adaptivity because they generate high false-positive rates and cannot easily capture the complex, evolving interactions among features that characterize sophisticated fraud (Bolton & Hand, 2002). Machine-learning methods offer a way to move from static rules toward adaptive, probabilistic detection, but their adoption in high-stakes financial settings has been constrained by class imbalance, the cost of labeling, concept drift, and, critically, a lack of transparency in the models themselves (Rudin, 2019). Anomaly identification and predictive risk scoring provide the central methodological foundation for this research because together they frame fraud detection as an adaptive, auditable activity that begins with data and feature characterization and continues through anomaly detection, scoring, explanation, investigation, and periodic retraining (Abdallah et al., 2016). The combined approach requires institutions to move from reactive, post-loss discovery toward anticipatory, evidence-based intervention, while retaining the ability to interrogate model reasoning. Studies of fraud-detection systems repeatedly show that their value depends on both model quality and the quality of the features and labels supplied to them, and that a detector is only as reliable as the data engineering and validation that support it (Bhattacharyya et al., 2011). Concerns about interpretability are especially important in finance because a model can appear accurate on historical data while remaining impossible to explain, audit, or defend to regulators, which undermines analyst trust and institutional acceptance (Barredo Arrieta et al., 2020).

Background of the Study

The background of this study is rooted in the persistent and growing cost of financial fraud and the limitations of the tools traditionally used to combat it. Organizations lose a substantial share of annual revenue to fraud, and losses are frequently detected late, long after the fraudulent activity began, when recovery is difficult (ACFE, 2022). For decades, detection relied heavily on manual review and fixed rule engines that flag transactions exceeding predefined thresholds or matching known fraud signatures. While transparent and easy to audit, such rules are brittle: they miss novel fraud schemes, generate large numbers of false positives that burden investigators, and require constant manual updating as fraud evolves (Bolton & Hand, 2002).

Machine learning has emerged as a means of overcoming these limitations. The availability of large transaction histories, improved computational resources, and advances in learning algorithms have made it feasible to model normal and fraudulent behavior at a scale and subtlety that rules cannot match (Ngai et al., 2011). Anomaly-detection methods can surface previously unseen fraud without labeled examples, while supervised classifiers such as gradient-boosted trees and neural networks can assign calibrated fraud probabilities when labels are available (Chen & Guestrin, 2016). Combining these approaches within a single framework allows institutions to catch both known and emerging fraud while prioritizing investigator attention through risk scores. In the accounting domain, quantitative indicators have long been used to flag likely financial-statement manipulation (Beneish, 1999), providing an early precedent for model-based fraud detection (Amir & Chapal, 2022; Abdur & Iftekhar, 2021).

Despite this promise, the deployment of machine-learning fraud detection faces distinctive challenges. Fraudulent records are extremely rare relative to legitimate ones, creating severe class imbalance that biases naive models toward predicting everything as legitimate (He & Garcia, 2009). Labels are costly and delayed, since confirming fraud requires investigation. Fraud patterns drift as adversaries adapt, so models degrade unless retrained (Lu et al., 2019). And because financial decisions are heavily regulated, institutions are reluctant to act on model outputs they cannot explain. The background of this study therefore reflects both the technical maturity of anomaly detection and risk scoring and the institutional need for a framework in which detection is accurate, adaptive, transparent, and embedded in analyst workflows (Lockey et al., 2021).

Problem Statement

The central problem addressed by this study is that fraud detection in financial records remains largely reactive, rule-bound, and opaque, so that preventable losses may go undetected when anomaly detection modeling, predictive risk scoring, feature engineering, model validation, and analyst engagement are weakly implemented. Although the technical capability to detect anomalies and score fraud risk has advanced substantially, its translation into accurate, trustworthy, and adoptable practice has lagged (Kanti & Rony, 2022; Sadia et al., 2022). Machine-learning models, when deployed, often function as black boxes whose reasoning analysts cannot interrogate, which limits trust, complicates regulatory justification, and impedes adoption (Rudin, 2019). At the same time, systems tuned only for accuracy on imbalanced data may achieve misleadingly high scores while missing rare fraud or flooding investigators with false positives (Hand, 2009).

This problem matters because financial institutions operate under strong accountability and regulatory requirements, and because the consequences of undetected fraud include direct monetary loss, regulatory penalties, and reputational harm. A framework that produces accurate predictions but cannot explain them may still fail to be adopted, while a framework that is interpretable but poorly validated or trained on imbalanced, low-quality data may produce misleading or unstable results. The study therefore investigates how an integrated machine-learning framework combining anomaly identification and predictive risk scoring influences the accuracy, transparency, and reliability of fraud detection, and which technical and institutional factors most strongly shape framework performance.

Objectives of the Study

The general objective of this study was to assess how a machine learning-based fraud detection framework, integrating unsupervised anomaly identification with supervised predictive risk scoring, influences fraud detection performance in financial records. The specific objectives were: to examine the influence of anomaly detection modeling on framework design quality; to assess the relationship between predictive risk scoring and fraud detection performance; to evaluate the effect of feature engineering and data quality on fraud detection performance; to determine the role of model validation and performance in strengthening detection reliability; to test the effect of framework design quality on fraud detection performance; and to examine the role of analyst and institutional engagement in improving detection outcomes. A further objective was to identify framework maturity levels and fraud risk-control gaps across the detection lifecycle.

Research Hypotheses

Six hypotheses guided the study. H1 proposed that anomaly detection modeling has significantly influenced framework design quality. H2 proposed that predictive risk scoring has a significant positive relationship with fraud detection performance. H3 proposed that feature engineering and data quality have significantly influenced fraud detection performance. H4 proposed that model validation and performance have significantly improved detection reliability. H5 proposed that framework design quality has significantly predicted fraud detection performance. H6 proposed that analyst and institutional engagement has significantly strengthened fraud detection performance.

Significance of the Research

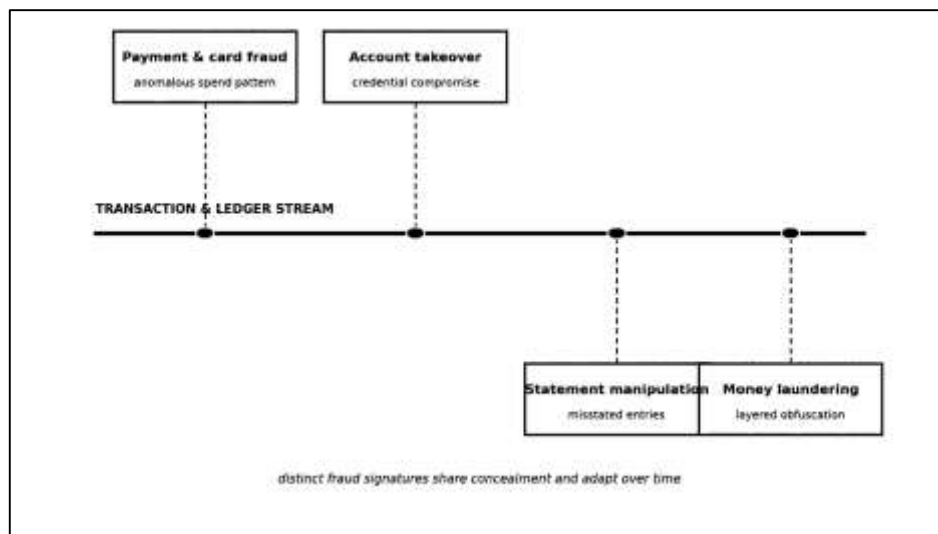
This research is significant because it addresses a practical and consequential gap between the technical maturity of machine-learning fraud detection and the institutional need for accurate, transparent, and adoptable systems in finance (West & Bhattacharya, 2016). By treating anomaly identification and predictive risk scoring as components of a single, adaptive framework, the study offers a way to move beyond opaque, single-method detection toward an integrated approach that institutions can adopt, justify, and defend. The findings are relevant to fraud analysts, data scientists, auditors, risk and compliance officers, and regulators who must make high-stakes decisions about flagged transactions. The study also contributes to the broader literature on trustworthy machine learning for financial crime by empirically examining which technical and institutional factors most strongly influence framework performance, and by developing a framework maturity index and a fraud risk-control priority matrix that translate statistical results into actionable improvement priorities (Perols, 2011).

LITERATURE REVIEW

Fraud in Financial Records and Its Detection

The literature on financial fraud consistently emphasizes that fraudulent activity spans several distinct but interacting typologies, each with its own signature in the data. Payment and card fraud involves unauthorized transactions that often deviate from a customer's spending pattern (Golam & Amir, 2023; Tohidul, 2023); account takeover involves the compromise of legitimate credentials and subsequent anomalous behavior; financial-statement manipulation involves the deliberate misstatement of accounting entries to mislead stakeholders; and money laundering involves layered transactions designed to obscure the origin of illicit funds (Ngai et al., 2011). Studies of these typologies stress that they are rare, deliberately concealed, and adaptive, which distinguishes fraud detection from conventional classification problems (Abdallah et al., 2016). Recent systematic reviews of machine-learning fraud detection document the rapid growth of the field (Ali et al., 2022), including a wide range of anomaly-detection techniques surveyed specifically for financial applications (Hilal et al., 2021).

Figure 2: Interacting Fraud Typologies in Financial Records



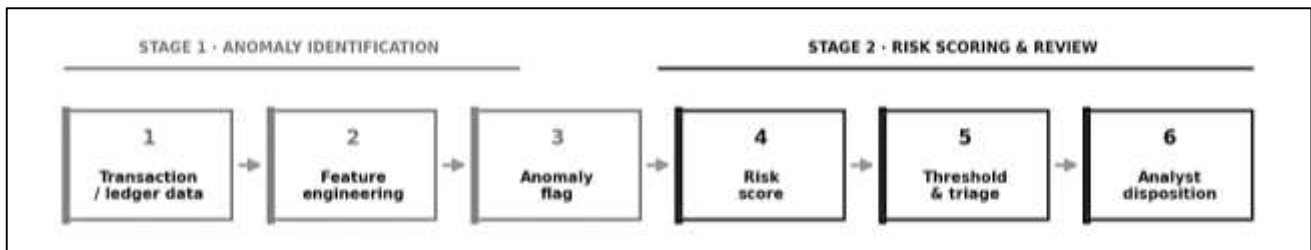
Research repeatedly highlights the adversarial and non-stationary character of fraud (Golam & Amir, 2023; Tohidul, 2023). Because perpetrators actively adapt to evade detection, the statistical relationship between features and fraud drifts over time, so that a model trained on past data degrades as new schemes emerge (Gama et al., 2014). This dynamic distinguishes fraud detection from static prediction tasks and motivates frameworks that continuously monitor, retrain, and adapt, rather than treating the detection model as fixed once deployed (Lu et al., 2019).

Anomaly Identification Methods

Anomaly detection has a long history as a label-free approach to surfacing suspicious behavior. Surveys of the field distinguish statistical, distance-based, density-based, and model-based methods, each defining normality differently and flagging deviations accordingly (Chandola et al., 2009). Density-based approaches such as the local outlier factor score a point by comparing its local density to that of its neighbors (Breunig et al., 2000), while one-class support vector methods learn a boundary enclosing the region of normal data (Schölkopf et al., 2001), and ensemble methods such as random forests provide robust supervised baselines for comparison (Breiman, 2001). Isolation forests take a fundamentally different, partition-based view, isolating anomalies through random splits (Liu et al., 2008), and deep methods, building on advances in deep learning (LeCun et al., 2015), learn compressed representations whose reconstruction error signals abnormality (Pang et al., 2021).

The literature is candid about the challenges of anomaly detection for fraud. Because anomalies are defined relative to normal behavior, benign but unusual activity can generate false positives, and the absence of labels makes threshold selection difficult (Ruff et al., 2021). These limitations mean that anomaly detection is most effective as a first-stage filter that narrows the candidate set, to be followed by supervised scoring that incorporates known fraud labels. The reliability of an anomaly-based stage therefore depends heavily on feature quality, threshold calibration, and validation, all of which the present study treats as elements of modeling and data quality (Badrul & Mominul, 2023; Hossan & Adar, 2023).

Figure 3: Two-Stage Detection Pipeline – Anomaly Identification to Risk Scoring



Predictive Risk Scoring and Supervised Learning

Supervised learning provides the second stage of the framework, assigning each record a calibrated probability of fraud from labeled examples. The predictive risk score is commonly expressed as the model's estimated posterior probability of fraud, passed through a logistic link, as shown in Equation (3):

$$r(x) = \Pr(Y = 1 \mid x) = \sigma(f_{\theta}(x)) = \frac{1}{1 + e^{-f_{\theta}(x)}} \quad (3)$$

In Equation (3), f_{θ} is the learned scoring function, and the sigmoid maps its output to a probability $r(x)$ that the record is fraudulent (Cortes & Vapnik, 1995). Gradient-boosted decision trees and deep neural networks have become the dominant scoring models because they capture complex, non-linear feature interactions characteristic of fraud (Chen & Guestrin, 2016). Realistic modeling studies show that scoring must account for delayed and imbalanced labels as they arise in live card-fraud streams (Dal Pozzolo et al., 2018), and sequence-based approaches further exploit the temporal structure of transaction histories (Jurgovsky et al., 2018). Because fraud is rare, the scoring model must be trained with attention to class imbalance, which is commonly addressed through cost-sensitive loss functions

that weight the minority class more heavily, as given in Equation (4):

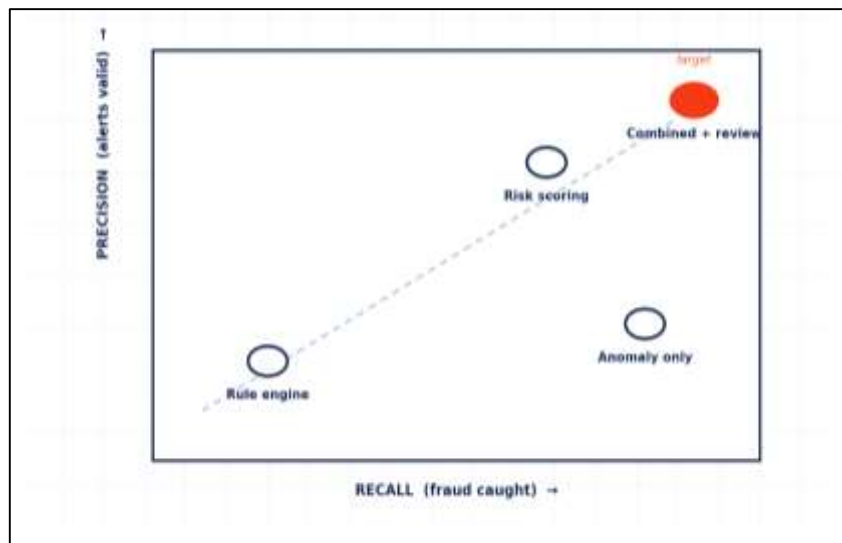
$$\mathcal{L}(\theta) = -\frac{1}{N} \sum_{i=1}^N \left[w_1 y_i \log r(x_i) + w_0 (1 - y_i) \log(1 - r(x_i)) \right] \quad (4)$$

In Equation (4), w_1 and w_0 are class weights that penalize missed fraud more heavily than false alarms, y_i is the true label, and $r(x_i)$ is the predicted probability (Krawczyk, 2016). Alternatives such as synthetic oversampling of the minority class further mitigate imbalance (Chawla et al., 2002), and generative approaches synthesize realistic minority examples to improve classification (Fiore et al., 2019). Reviews of imbalanced learning emphasize that failing to address imbalance produces models that appear accurate overall while missing the rare fraud cases that matter most (He & Garcia, 2009). The evaluation of fraud detection models requires metrics suited to rare events, since overall accuracy is misleading when the positive class is a tiny fraction of the data. The literature recommends precision, recall, and their harmonic mean, together with precision-recall curves, as expressed in Equation (5):

$$\text{Precision} = \frac{TP}{TP + FP}, \quad \text{Recall} = \frac{TP}{TP + FN}, \quad F_1 = \frac{2PR \cdot R}{PR + R} \quad (5)$$

Precision measures the fraction of flagged records that are truly fraudulent, recall measures the fraction of fraud that is caught, and the F1 score balances the two (Saito & Rehmsmeier, 2015). Because the cost of a missed fraud typically differs from the cost of a false alarm, the literature also emphasizes cost-sensitive and ranking-based evaluation (Provost & Fawcett, 2001), and cautions against relying on the area under the ROC curve alone when classes are highly imbalanced (Hand, 2009).

Figure 4: Balancing Detection Rate and False Positives under Class Imbalance



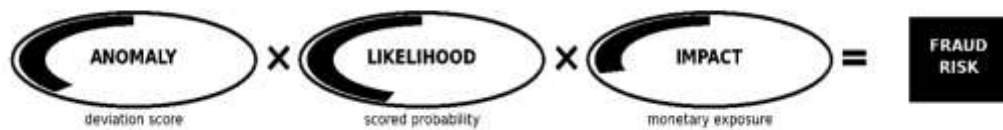
Explainability and Trust in Fraud Models

A recurring theme in the recent literature is that accuracy alone is insufficient for high-stakes financial decisions; models must also be interpretable (Rudin, 2019). Explainable AI methods address this by providing human-interpretable accounts of model behavior. Post-hoc methods such as local surrogate explanations (Ribeiro et al., 2016) and additive feature-attribution methods (Lundberg & Lee, 2017) estimate the contribution of each feature to a prediction. For fraud detection, such attributions answer the analyst's central question of why a transaction was flagged, converting an opaque score into an auditable explanation. The Shapley value that underlies additive attribution assigns each feature a contribution equal to its average marginal effect over all feature orderings, as given in Equation (6):

$$\phi_k = \sum_{S \subseteq F \setminus \{k\}} \frac{|S|! (|F| - |S| - 1)!}{|F|!} [f(S \cup \{k\}) - f(S)] \quad (6)$$

In Equation (6), F is the feature set, S is a subset excluding feature k , and the combinatorial weight ensures a fair allocation of the prediction across features (Barredo Arrieta et al., 2020). By making the drivers of each risk score visible, feature attribution supports analyst trust, regulatory justification, and the accountable use of automated fraud detection.

Figure 5: Fraud Risk as the Intersection of Anomaly, Likelihood, and Impact



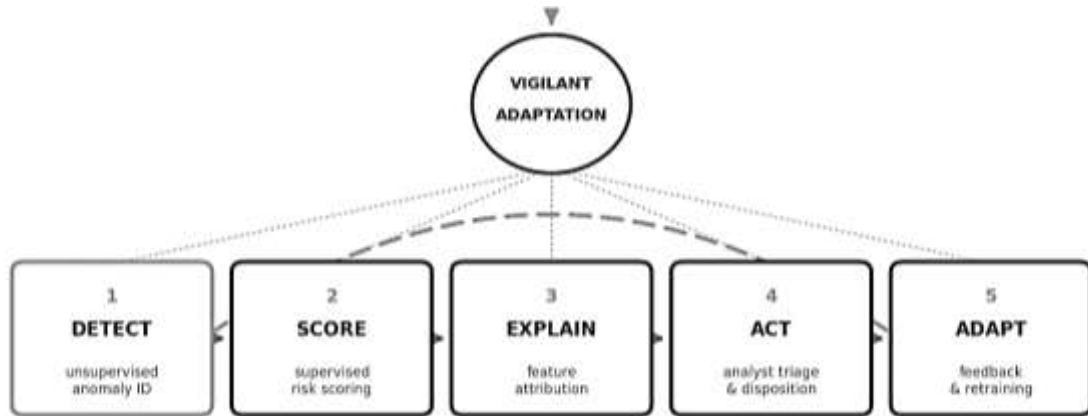
Theoretical Framework: A Dynamic Detect-Score-Explain-Act (DSEA) Loop

Rather than treating the framework as a static arrangement of technical and social components, this study advances a dynamic theoretical model in which machine-learning fraud detection is understood as a continuous, adversarial-aware, adaptive loop. The proposed Detect-Score-Explain-Act (DSEA) loop integrates three complementary theoretical lenses: socio-technical systems theory, which holds that technical and human subsystems must be jointly optimized rather than treated separately (Trist, 1981); resilience engineering, which frames reliable performance as a system's ongoing capacity to anticipate, monitor, respond to, and learn from disturbance (Hollnagel et al., 2006); and technology-trust and vigilant-decision theory, which explains whether and how automated outputs are accepted and acted upon by the analysts who rely on them (Davis, 1989). The novelty of the DSEA loop is that it treats these lenses as interacting forces that circulate through the framework in a repeating cycle, and it explicitly incorporates the adversarial dynamics that make fraud detection distinct from other prediction tasks.

The loop is organized around five interacting phases that continuously feed one another. In the detect phase, unsupervised anomaly identification surfaces records that deviate from learned normal behavior (Amir, 2024; Risha & Khalid, 2023). In the score phase, supervised models assign calibrated fraud probabilities to the flagged candidates. In the explain phase, feature-attribution methods translate scores into human-interpretable reason codes (Lundberg & Lee, 2017). In the act phase, analysts triage and investigate flagged cases, and their dispositions determine the institution's response. In the adapt phase, confirmed-fraud labels and analyst feedback are fed back to retrain the models, closing the loop. Because each phase both depends on and reshapes the others, the framework is inherently dynamic rather than linear, and its performance emerges from the quality of circulation around the whole loop rather than from any single stage (Hossan, 2024; Shima Ali et al., 2024).

At the center of the loop sits vigilant adaptation, the system's capacity to learn and reconfigure in the face of an adaptive adversary (Amir, 2025; Zaman, 2024). Unlike conventional predictive pipelines, the DSEA loop treats concept drift not as noise but as the expected consequence of adversarial behavior: as fraudsters change tactics, the statistical patterns of fraud shift, and the framework must continuously incorporate new labels to remain effective (Gama et al., 2014). This adversarial-drift feedback distinguishes the DSEA framework from static fraud-detection systems and motivates the study's treatment of both technical constructs, such as anomaly detection modeling and predictive risk scoring, and human-institutional constructs, such as validation and analyst engagement, as mutually dependent determinants of fraud detection performance rather than as separate inputs (Weick & Sutcliffe, 2007).

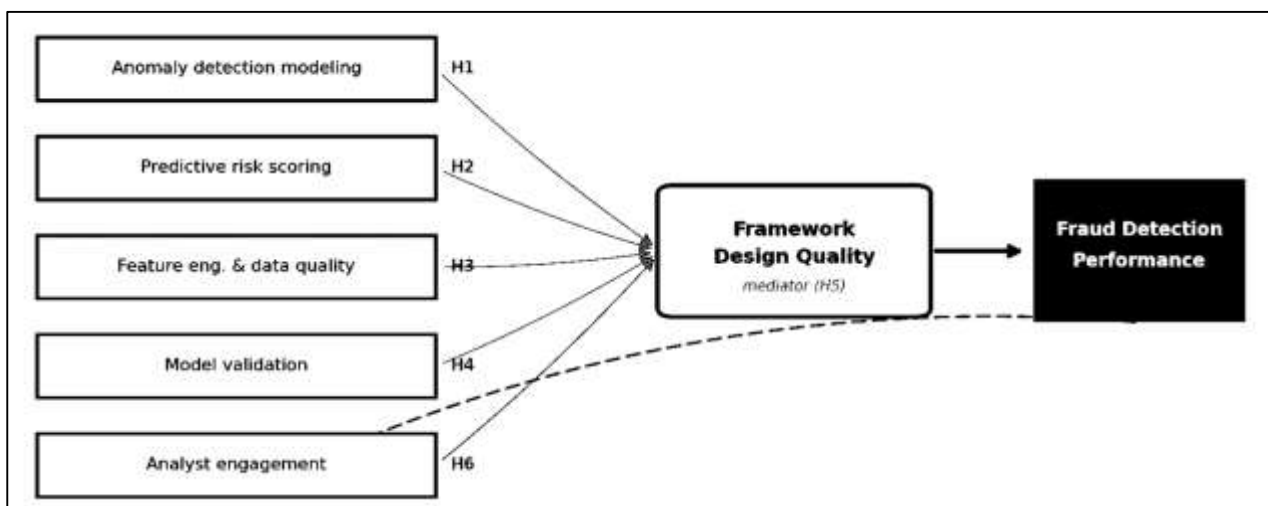
Figure 6: The Dynamic Detect-Score-Explain-Act (DSEA) Loop



Conceptual Framework

Drawing on the reviewed literature and theoretical perspectives, the conceptual framework of this study positions fraud detection performance as the dependent variable, influenced by six independent and mediating constructs. Anomaly detection modeling and predictive risk scoring represent the core technical capabilities of the framework. Feature engineering and data quality, together with model validation and performance, represent the methodological rigor that converts raw records into trustworthy, actionable predictions (Bhattacharyya et al., 2011). Analyst and institutional engagement represents the socio-technical context that determines whether outputs are trusted and acted upon. Framework design quality functions as an integrating construct that captures how coherently these elements are combined, and it is expected to be the strongest direct predictor of fraud detection performance. The framework proposes that stronger implementation of each construct, circulating together through the dynamic DSEA loop, produces more accurate, transparent, and actionable fraud detection for financial institutions.

Figure 7: Predictive Conceptual Framework for Fraud Detection Performance



METHODS

This study adopted a quantitative, cross-sectional, case-based research design to examine how a machine learning-based fraud detection framework, integrating unsupervised anomaly identification with supervised predictive risk scoring, influences fraud detection performance in financial records. The quantitative approach was selected because the study measured relationships among key variables using numerical data collected through a structured questionnaire. The cross-sectional design was appropriate because data were gathered from respondents at a single point in time, while the case-

based approach enabled the research to focus on financial institutions and analytics teams where anomaly detection, machine learning, and risk scoring have been applied or considered relevant for fraud detection.

The case context focused on financial environments such as banking and payments, credit and lending, insurance claims, and corporate accounting, where fraud detection depends heavily on the analysis of transactions, ledgers, and account behavior. These settings were selected because they generate large volumes of records in which fraud is rare, concealed, and adaptive, making them well suited to machine-learning detection. The study considered these environments suitable because an anomaly-and-scoring framework is highly relevant to their fraud-prevention, investigation, and regulatory-compliance needs.

The population of the study consisted of professionals involved in fraud analytics, data science, auditing, forensic accounting, risk management, compliance, and financial-systems administration. The unit of analysis was the individual professional respondent who possessed knowledge or experience related to anomaly detection, machine learning, risk scoring, or fraud investigation in financial records. The study used a purposive sampling strategy because respondents needed specific technical or institutional knowledge to provide valid responses. Where possible, the sample included fraud analysts and investigators, data scientists and machine-learning engineers, auditors and forensic accountants, risk and compliance officers, and financial-systems managers.

The data-collection procedure involved a structured questionnaire distributed to selected respondents electronically or physically. Before data collection, respondents were informed about the academic purpose of the study, voluntary participation, confidentiality, and anonymity. The collected responses were reviewed for completeness, consistency, and suitability before analysis, and incomplete or invalid responses were excluded to maintain data quality.

Figure 8: Research Methodology



The instrument design was based on a five-point Likert scale ranging from 1 = Strongly Disagree to 5 = Strongly Agree. The questionnaire included sections on demographic profile, anomaly detection modeling, predictive risk scoring, feature engineering and data quality, model validation and performance, analyst and institutional engagement, framework design quality, and fraud detection performance.

A pilot test was conducted before the main survey to assess the clarity, relevance, wording, and technical suitability of the questionnaire items. Feedback from the pilot test was used to improve ambiguous or repetitive items. Validity and reliability were ensured through expert review and

statistical testing. Content validity was supported by aligning questionnaire items with the study objectives and fraud-detection concepts, while reliability was tested using Cronbach's alpha, with a value of 0.70 or above considered acceptable.

Internal consistency was quantified using Cronbach's alpha, defined in Equation (7), where K is the number of items in a construct, the numerator sums the individual item variances, and the denominator is the variance of the summed scale:

$$\alpha = \frac{K}{K-1} \left(1 - \frac{\sum_{i=1}^K \sigma_i^2}{\sigma_T^2} \right) \quad (7)$$

Values approaching unity indicate that the items measure a common underlying construct, and the 0.70 threshold reflects the conventional lower bound for acceptable reliability.

For data analysis, statistical software was used for coding, descriptive statistics, reliability analysis, correlation analysis, and regression modeling. A spreadsheet application was used for initial data cleaning, tabulation, and organization of responses, and reference-management software was used for citation organization and APA formatting.

Two model specifications underpinned the analysis. First, framework design quality was operationalized as a weighted composite of its constituent construct means, as shown in Equation (8), where each construct mean is weighted by a non-negative weight summing to one:

$$FDQ = \sum_{c=1}^C w_c \bar{x}_c, \quad \sum_{c=1}^C w_c = 1, \quad w_c \geq 0 \quad (8)$$

Second, the influence of the six predictors on fraud detection performance was estimated with the multiple linear regression model in Equation (9), where FDP denotes fraud detection performance, each X_k is a predictor construct, and the error term is assumed normally distributed with zero mean and constant variance:

$$FDP = \beta_0 + \sum_{k=1}^6 \beta_k X_k + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma^2) \quad (9)$$

The standardized coefficients beta-k reported in the regression results express the relative predictive strength of each construct, and the coefficient of determination R-squared quantifies the proportion of variance in fraud detection performance jointly explained by the six predictors.

FINDINGS

Response Rate

Table 1: Response Rate of the Survey

Survey Status	Frequency	Percentage
Questionnaires distributed	160	100.0%
Questionnaires returned	152	95.0%
Invalid/incomplete responses	10	6.3%
Valid responses used for analysis	142	88.8%

The response-rate results in Table 1 showed that the study achieved a strong and acceptable level of participation from professionals involved in fraud analytics and financial risk. Out of 160 questionnaires distributed, 152 responses were returned, representing a return rate of 95.0%. After screening for completeness, consistency, and suitability for statistical analysis, 10 responses were removed because they contained missing values, incomplete Likert-scale answers, or inconsistent response patterns. As a result, 142 valid responses were retained for final analysis, producing an effective valid response rate of 88.8%.

This response rate was considered suitable for a quantitative, cross-sectional, case-based study because it provided enough data to conduct descriptive statistics, reliability analysis, correlation analysis, and regression modeling. The response rate also strengthened the credibility of the findings because the respondents represented relevant professional groups connected to fraud analytics, data science, auditing, forensic accounting, risk, compliance, and financial-systems administration. From a socio-technical perspective, the strong response rate suggested that the selected professional population showed meaningful engagement with fraud-detection issues and provided sufficient institutional insight into technical capability, model interpretability, validation practices, and analyst engagement. The valid sample of 142 respondents therefore provided a dependable foundation for examining the study objectives and hypotheses, allowing the research to test whether anomaly detection modeling, predictive risk scoring, feature engineering, model validation, analyst engagement, and framework design quality significantly contributed to fraud detection performance.

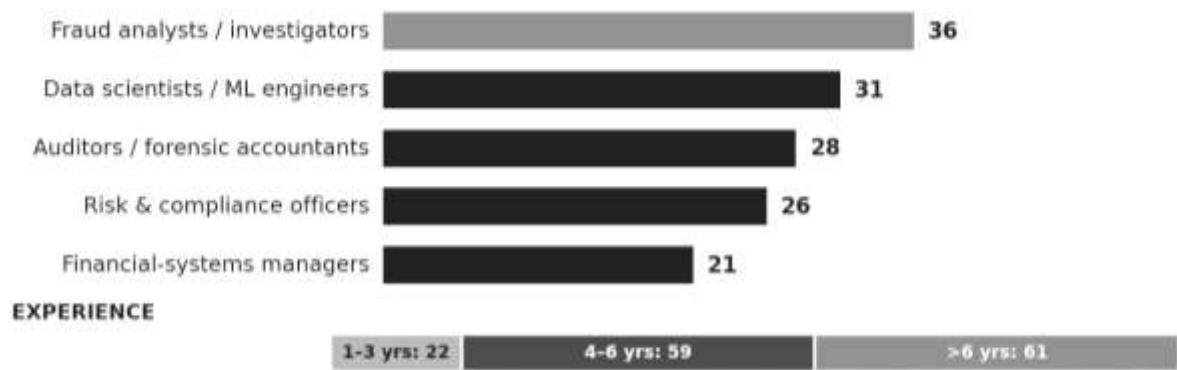
Demographic and Professional Profile of Respondents

Table 2 presented the demographic and professional characteristics of the 142 respondents included in the analysis. The results indicated that the sample was technically relevant to the research topic because the largest group of respondents consisted of fraud analysts and investigators, accounting for 25.4% of the sample. This was followed by data scientists and machine-learning engineers at 21.8%, auditors and forensic accountants at 19.7%, risk and compliance officers at 18.3%, and financial-systems managers at 14.8%. This distribution was useful because a machine-learning fraud-detection framework depends on multiple professional roles rather than one technical group alone. Fraud analysts and investigators are typically involved in triaging alerts, investigating cases, and providing dispositions. Data scientists and machine-learning engineers contribute expertise in model development, feature engineering, and validation. Auditors and forensic accountants contribute knowledge of accounting manipulation and evidence standards. Risk and compliance officers contribute governance, regulatory, and accountability perspectives. Financial-systems managers contribute knowledge of data pipelines, system integration, and deployment. The experience profile also strengthened the study because 43.0% of respondents had more than six years of experience, while 41.5% had four to six years of experience. In addition, 66.9% of respondents reported direct involvement in fraud-detection or risk-scoring activities, while 33.1% had indirect involvement.

Table 2: Demographic and Professional Profile of Respondents

Profile Variable	Category	Frequency	Percentage
Job role	Fraud analysts/investigators	36	25.4%
Job role	Data scientists/ML engineers	31	21.8%
Job role	Auditors/forensic accountants	28	19.7%
Job role	Risk & compliance officers	26	18.3%
Job role	Financial-systems managers	21	14.8%
Work experience	1–3 years	22	15.5%
Work experience	4–6 years	59	41.5%
Work experience	Above 6 years	61	43.0%
Involvement	Direct involvement	95	66.9%
Involvement	Indirect involvement	47	33.1%

This profile supported the validity of the findings because the study depended on respondents who understood anomaly detection, risk scoring, feature engineering, validation, and fraud investigation. From a socio-technical perspective, the sample reflected a cross-functional system in which reliable fraud detection depends on communication, expertise, technical awareness, and coordination among different professional groups. The respondent profile was therefore appropriate for addressing the study objectives and testing the hypotheses.

Figure 11: Respondent Profile by Job Role and Work Experience

Descriptive Statistics of Study Variables

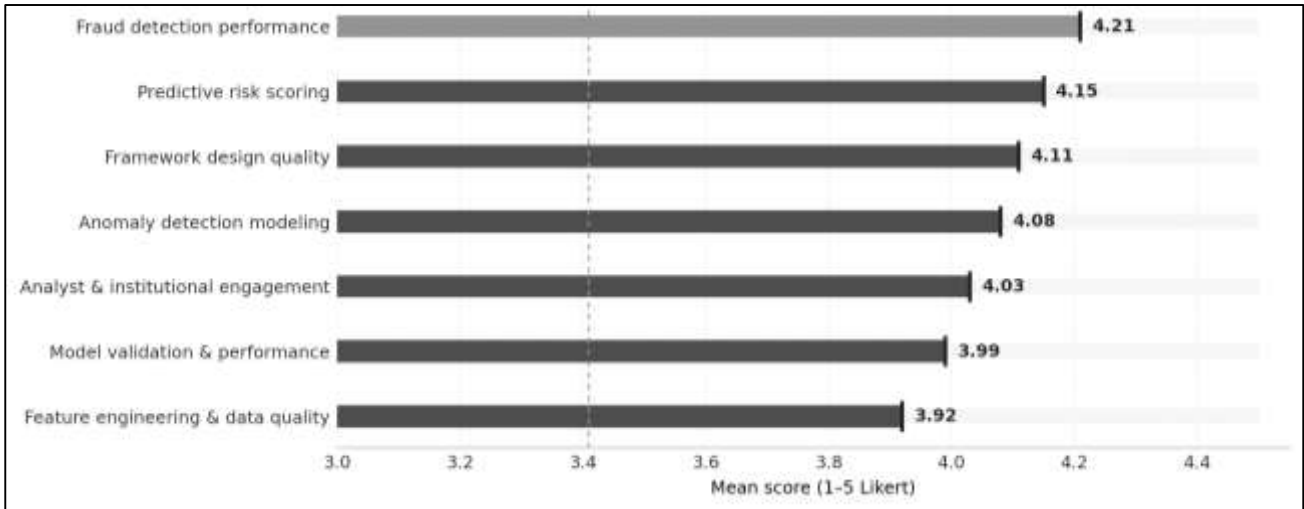
Table 3: Descriptive Statistics of Study Variables Based on Five-Point Likert Scale

Study Variable	Mean	SD	Interpretation	Rank
Fraud detection performance	4.21	0.54	Very High	1
Predictive risk scoring	4.15	0.58	High	2
Framework design quality	4.11	0.57	High	3
Anomaly detection modeling	4.08	0.61	High	4
Analyst and institutional engagement	4.03	0.63	High	5
Model validation and performance	3.99	0.66	High	6
Feature engineering and data quality	3.92	0.69	High	7

Likert interpretation guide: 1.00–1.80 = Very Low, 1.81–2.60 = Low, 2.61–3.40 = Moderate, 3.41–4.20 = High, 4.21–5.00 = Very High.

Table 3 showed the descriptive statistics for the major study variables measured through the five-point Likert scale. All variables recorded mean scores above the neutral midpoint of 3.00 and within the high-to-very-high range, suggesting that respondents generally agreed that machine-learning fraud-detection practices, anomaly detection, risk scoring, feature engineering, validation, analyst engagement, and detection performance were positively present in the selected financial context. Fraud detection performance recorded the highest mean of 4.21 with a standard deviation of 0.54, reaching the very-high band and indicating that respondents perceived early detection, false-positive control, calibrated risk, and faster investigation as relatively strong. Predictive risk scoring recorded a mean of 4.15, showing that calibrated probability scoring and prioritization were perceived as well implemented. Framework design quality followed with a mean of 4.11, suggesting agreement that the coherent combination of technical and institutional elements supported detection goals. Anomaly detection modeling recorded 4.08, analyst and institutional engagement recorded 4.03, and model validation and performance recorded 3.99. Feature engineering and data quality recorded the lowest mean of 3.92. Although all values remained high, the lower ranking of feature engineering and validation suggested that these data-facing and methodological activities may require more focused improvement. These findings aligned with the dynamic DSEA loop because the strongest detection performance appeared where the detect, score, explain, act, and adapt phases circulated together rather than in isolation.

Figure 13: Mean Scores of Study Variables with Standard Deviation

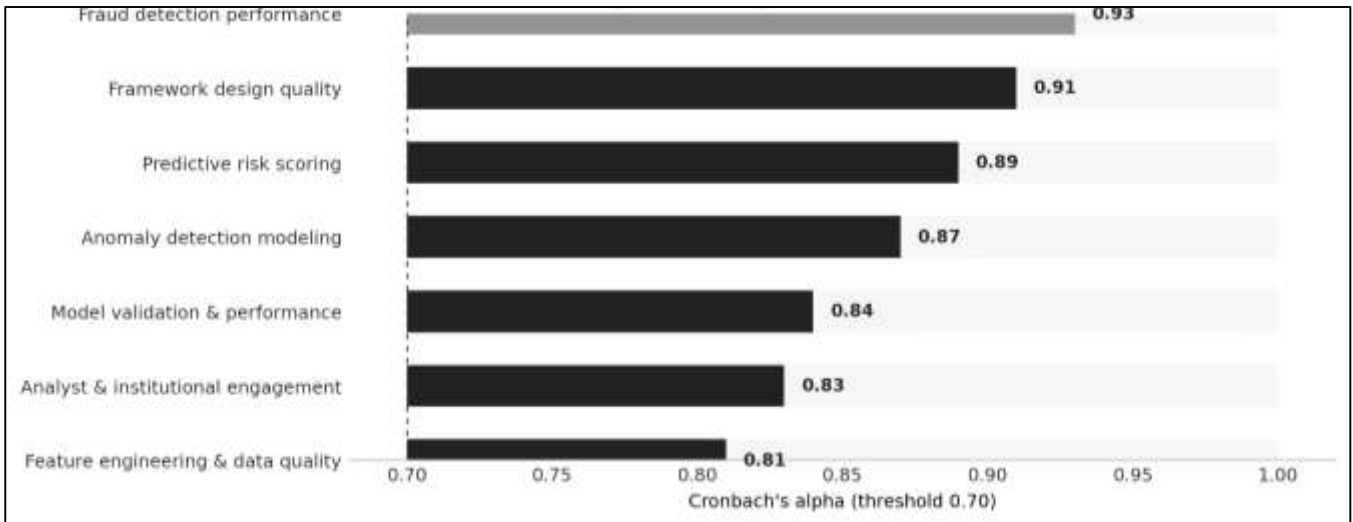


Reliability Analysis

Table 4: Reliability Analysis of Study Constructs

Construct	Items	Cronbach's Alpha	Decision
Anomaly detection modeling	6	0.89	Reliable
Predictive risk scoring	6	0.87	Reliable
Feature engineering and data quality	5	0.81	Reliable
Model validation and performance	5	0.84	Reliable
Analyst and institutional engagement	5	0.83	Reliable
Framework design quality	6	0.91	Reliable
Fraud detection performance	6	0.93	Reliable

Table 4 presented the reliability results for the major questionnaire constructs. Cronbach's alpha was used to determine whether the items measuring each construct were internally consistent. The accepted threshold was 0.70, and all constructs exceeded this minimum. Fraud detection performance recorded the highest alpha of 0.93, indicating excellent internal consistency among items measuring early detection, false-positive control, calibrated risk, and faster investigation. Framework design quality recorded 0.91, confirming that items measuring the coherent integration of technical and institutional elements formed a reliable construct. Anomaly detection modeling recorded 0.89, predictive risk scoring recorded 0.87, and model validation and performance recorded 0.84. Analyst and institutional engagement recorded 0.83, confirming that items on communication, trust, and institutional coordination reliably measured the socio-technical dimension. Feature engineering and data quality recorded the lowest alpha of 0.81, which remained well above the accepted threshold and was therefore considered reliable. These reliability results strengthened the trustworthiness of the study by showing that the questionnaire measured the intended constructs consistently, and they supported the socio-technical view that detection performance is measurable as a combination of technical and institutional practices.

Figure 14: Internal Consistency Reliability (Cronbach's Alpha) of Study Constructs

Correlation Analysis

Table 5: Pearson Correlation Analysis among Major Study Variables

Relationship Tested	r	Significance	Interpretation	Link
Anomaly detection modeling and framework design quality	0.75	$p < 0.01$	Strong positive	H1
Predictive risk scoring and fraud detection performance	0.69	$p < 0.01$	Strong positive	H2
Feature engineering/data quality and detection performance	0.70	$p < 0.01$	Strong positive	H3
Model validation and detection reliability	0.65	$p < 0.01$	Moderate-to-strong	H4
Analyst engagement and detection performance	0.67	$p < 0.01$	Strong positive	H6
Framework design quality and detection performance	0.78	$p < 0.01$	Strong positive	H5

Table 5 presented the Pearson correlation results used to examine the strength and direction of relationships among the major study variables. All tested relationships were positive and statistically significant at $p < 0.01$, indicating that stronger implementation of technical and institutional practices was associated with stronger fraud detection performance. The relationship between anomaly detection modeling and framework design quality recorded a strong positive correlation of $r = 0.75$, supporting H1 by showing that stronger anomaly-detection capability accompanied stronger framework design. Predictive risk scoring correlated with detection performance at $r = 0.69$, supporting H2. Feature engineering and data quality correlated with detection performance at $r = 0.70$, supporting H3 and indicating that feature quality and data cleaning contributed meaningfully to trustworthy outputs. Model validation and performance recorded $r = 0.65$ with detection reliability, supporting H4. Analyst and institutional engagement recorded $r = 0.67$ with detection performance, supporting the socio-technical argument that reliable detection depends on institutional and human behavior as well as technical systems. Framework design quality showed the strongest relationship with detection performance at $r = 0.78$, supporting H5 and confirming that coherent design was highly relevant to outcomes. These findings aligned with the dynamic DSEA loop because detection performance improved when the technical, methodological, and human phases circulated together as an integrated, adaptive cycle.

Figure 15: Correlation Strength among Major Study Variables



Regression Analysis

Table 6: Multiple Regression Results Predicting Fraud Detection Performance

Predictor Variable	Std. Beta (β)	t-value	p-value	Decision
Anomaly detection modeling	0.24	3.41	0.001	Significant
Predictive risk scoring	0.26	3.72	<0.001	Significant
Feature engineering and data quality	0.19	2.78	0.006	Significant
Model validation and performance	0.16	2.49	0.014	Significant
Analyst and institutional engagement	0.18	2.74	0.007	Significant
Framework design quality	0.34	4.79	<0.001	Significant

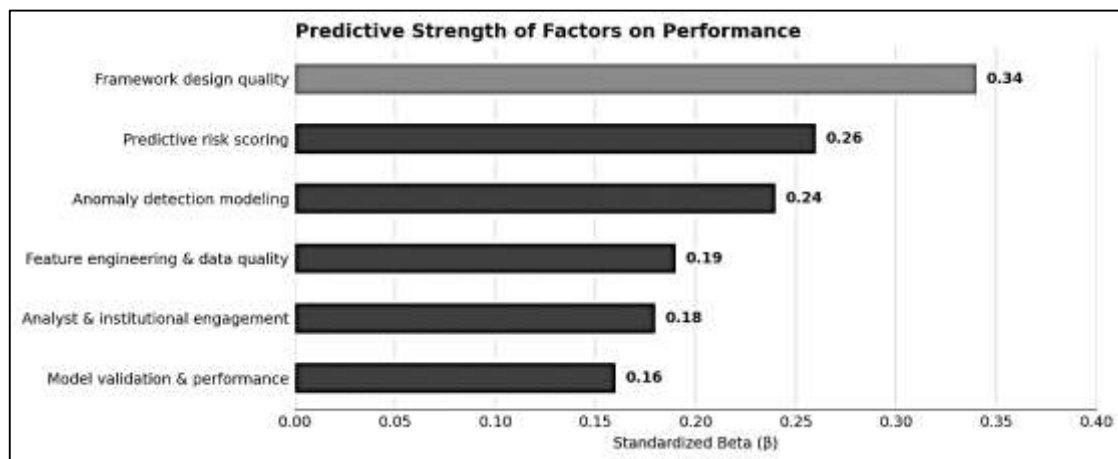
Table 7: Model Summary for Regression Analysis

Model Indicator	Value
R	0.845
R²	0.714
Adjusted R²	0.702
F-statistic	56.12
df	6, 135
Significance	p < 0.001

Tables 6 and 7 presented the multiple regression results used to determine the predictive effect of the independent variables on fraud detection performance. The regression model was statistically significant, $F(6, 135) = 56.12$, $p < 0.001$, indicating that the selected predictors collectively explained detection performance in a meaningful way. The model recorded an R value of 0.845, showing a strong

overall relationship between the predictors and the dependent variable. The R^2 value of 0.714 indicated that 71.4% of the variance in fraud detection performance was explained by anomaly detection modeling, predictive risk scoring, feature engineering and data quality, model validation, analyst engagement, and framework design quality. The adjusted R^2 value of 0.702 confirmed that the model remained strong after accounting for the number of predictors. Among the predictors, framework design quality showed the strongest effect, with $\beta = 0.34$ and $p < 0.001$, indicating that coherent framework design was the most important predictor of detection outcomes. Predictive risk scoring showed a significant positive effect, with $\beta = 0.26$ and $p < 0.001$, confirming that calibrated scoring improved detection performance. Anomaly detection modeling recorded $\beta = 0.24$ and $p = 0.001$, while feature engineering and data quality recorded $\beta = 0.19$ and $p = 0.006$. Analyst and institutional engagement recorded $\beta = 0.18$ and $p = 0.007$, and model validation and performance recorded $\beta = 0.16$ and $p = 0.014$, a smaller but still significant contribution. These results supported the dynamic DSEA loop because detection performance was explained jointly by the technical, methodological, and human phases of the cycle, and they directly supported the study objectives by proving that the selected variables significantly predicted fraud detection performance.

Figure 16: Predictive Strength of Factors on Fraud Detection Performance (Standardized Beta)



Hypotheses Testing

Table 8: Summary of Hypotheses Testing

Hypothesis	Statement	Statistical Evidence	Decision
H1	Anomaly detection modeling has significantly influenced framework design quality.	$r = 0.75, \beta = 0.24, p = 0.001$	Supported
H2	Predictive risk scoring has a significant positive relationship with detection performance.	$r = 0.69, \beta = 0.26, p < 0.001$	Supported
H3	Feature engineering and data quality have significantly influenced detection performance.	$r = 0.70, \beta = 0.19, p = 0.006$	Supported
H4	Model validation and performance have significantly improved detection reliability.	$r = 0.65, \beta = 0.16, p = 0.014$	Supported
H5	Framework design quality has significantly predicted detection performance.	$r = 0.78, \beta = 0.34, p < 0.001$	Supported
H6	Analyst engagement has significantly strengthened detection performance.	$r = 0.67, \beta = 0.18, p = 0.007$	Supported

Table 8 summarized the results of the hypotheses testing and showed that all six hypotheses were supported by the statistical findings. H1 was supported because anomaly detection modeling showed a strong positive relationship with framework design quality and produced a significant regression effect, confirming that stronger anomaly-detection capability accompanied stronger framework design. H2 was supported because predictive risk scoring showed a significant positive relationship with

detection performance, meaning that better-calibrated risk scores were associated with stronger outcomes. H3 was supported because feature engineering and data quality significantly influenced detection performance, confirming that feature construction, cleaning, and data quality improved trust in results. H4 was supported because model validation and performance showed a significant positive effect on detection reliability, suggesting that rigorous backtesting and evaluation strengthened detection readiness. H5 was supported because framework design quality was the strongest predictor of detection performance, with $\beta = 0.34$ and $p < 0.001$, directly proving the main objective of the study. H6 was supported because analyst and institutional engagement showed a significant positive effect, linking the findings to the DSEA loop, in which the act phase of analyst triage and the adapt phase of feedback are essential to sustained performance. Overall, the hypotheses testing demonstrated that the study model was statistically supported and that both technical and institutional variables contributed to fraud detection performance in financial institutions.

Framework Compliance Maturity Index

The framework maturity index (MI) and the risk-control gap score were computed as shown in Equation (10). The maturity index is the mean maturity across the S lifecycle stages, while the gap score for a risk-control area a is the difference between its importance mean and its effectiveness mean:

$$MI = \frac{1}{S} \sum_{s=1}^S \bar{m}_s, \quad \text{Gap}_a = \bar{I}_a - \bar{E}_a \quad (10)$$

Table 9: Machine Learning Fraud Detection Framework Maturity Index

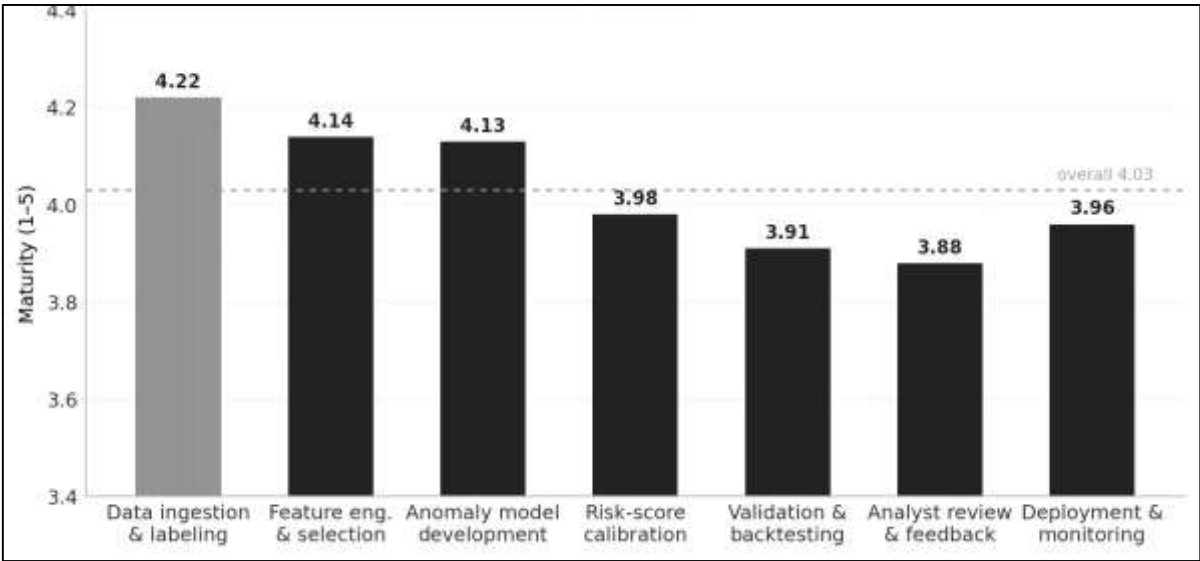
Framework Lifecycle Stage	Mean	SD	Maturity Level	Priority
Data ingestion and labeling	4.22	0.53	Very high maturity	Maintain and strengthen
Feature engineering and selection	4.14	0.56	High maturity	Maintain
Anomaly model development	4.13	0.58	High maturity	Maintain
Risk-score calibration	3.98	0.66	High maturity	Improve documentation
Validation and backtesting	3.91	0.70	High maturity	Needs attention
Analyst review and feedback	3.88	0.72	High maturity	Needs attention
Deployment and monitoring	3.96	0.67	High maturity	Improve consistency
Overall framework maturity	4.03	0.63	High maturity	Strong but improvable

Maturity interpretation guide: 1.00–1.80 = Very Low Maturity, 1.81–2.60 = Low Maturity, 2.61–3.40 = Moderate Maturity, 3.41–4.20 = High Maturity, 4.21–5.00 = Very High Maturity.

Table 9 presented the Machine Learning Fraud Detection Framework Maturity Index, developed as a study-specific result to assess how mature the selected institutions were across major stages of the framework lifecycle. The overall maturity score was 4.03, falling within the high-maturity range and indicating that the case context generally demonstrated strong implementation of machine-learning fraud-detection practices. Data ingestion and labeling recorded the highest maturity mean of 4.22, reaching the very-high band and suggesting strong capability in acquiring, labeling, and organizing transaction data. Feature engineering and selection recorded 4.14, while anomaly model development recorded 4.13, showing that the core technical stages were relatively strong. Risk-score calibration recorded 3.98, indicating high maturity but room for improvement in producing well-calibrated probabilities and documenting scoring thresholds. Deployment and monitoring recorded 3.96, suggesting that models were operational but could benefit from more consistent drift monitoring. Validation and backtesting recorded 3.91, while analyst review and feedback recorded the lowest maturity score of 3.88. Although both values remained classified as high maturity, their lower ranking showed that methodological and human-facing activities required more attention than the core technical stages. This result aligned with the descriptive findings, where validation and feature engineering recorded the lowest means among the major constructs. The maturity index linked directly

to the DSEA loop because it showed that reliable fraud detection depends on continuous circulation through the detect, score, explain, act, and adapt phases, supported by feedback-driven retraining, rather than on one-time model development.

Figure 18: Framework Maturity Across Lifecycle Stages



Fraud Risk-Control Priority Matrix

Table 10 presented the Fraud Risk-Control Priority Matrix, developed to compare the perceived importance and current implementation effectiveness of major risk-control areas within the framework. The gap score was calculated by subtracting the effectiveness mean from the importance mean, with a higher gap score indicating a higher improvement priority. The results showed that feature drift and data quality recorded the largest gap score of 1.15, followed closely by analyst trust and explanation quality with a gap score of 1.11. Both areas were therefore classified as very high priority, meaning that respondents considered them highly important while rating their current effectiveness lower than desired. False-positive rate control was also classified as very high priority, with an importance mean of 4.58, an effectiveness mean of 3.55, and a gap score of 1.03, reflecting the persistent burden that false alarms place on investigators. Model interpretability and transparency and account-takeover detection both recorded gap scores in the high-priority range.

Table 10: Fraud Risk-Control Priority Matrix

Fraud Risk-Control Area	Importance	Effectiveness	Gap	Priority
Real-time transaction scoring	4.76	3.98	0.78	High priority
Account-takeover detection	4.68	3.73	0.95	Very high priority
Feature drift and data quality	4.61	3.46	1.15	Very high priority
Analyst trust and explanation quality	4.53	3.42	1.11	Very high priority
Model interpretability and transparency	4.64	3.71	0.93	Very high priority
False-positive rate control	4.58	3.55	1.03	Very high priority
Case-management integration	4.38	3.77	0.61	Medium priority
Model governance and updating	4.33	3.62	0.71	High priority

Case-management integration recorded the lowest gap score of 0.61 and was classified as medium priority, suggesting that it was relatively stronger than other areas but still required continued monitoring. Model governance and updating recorded 0.71 and real-time transaction scoring recorded 0.78, both classified as high priority. This priority matrix strengthened the findings by translating statistical results into practical improvement priorities. It also aligned with the DSEA loop, which emphasizes attention to weak signals, human trust, and continuous adaptation in an adversarial environment. The matrix showed that reliable performance required attention not only to core technical modeling but also to feature and data quality, false-positive control, model interpretability, and analyst trust, which supported the study objectives by identifying where framework improvements should be prioritized.

Figure 19: Fraud Risk-Control Importance versus Effectiveness Gap Analysis

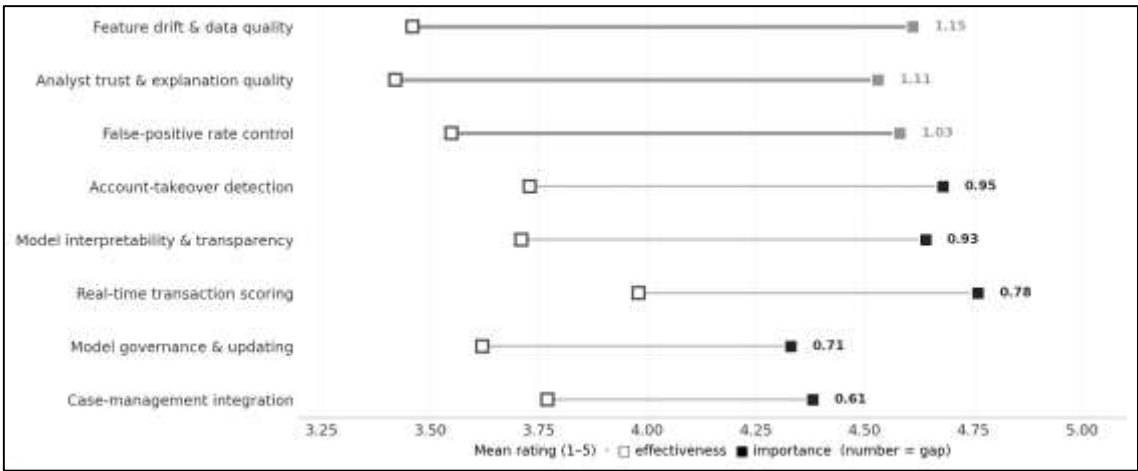


Table 11: Summary of Key Findings by Objective and Hypothesis

Study Objective	Main Result	Hypothesis	Interpretation
Examine influence of anomaly detection on framework design quality	Anomaly modeling strongly correlated with design quality, $r = 0.75$	H1	Stronger anomaly detection improved design quality
Assess relationship between risk scoring and detection performance	Risk scoring significantly related to performance, $r = 0.69$, $\beta = 0.26$	H2	Calibrated scoring improved outcomes
Evaluate effect of feature engineering and data quality	Feature quality significantly predicted performance, $\beta = 0.19$	H3	Feature quality strengthened trust in results
Determine role of model validation and performance	Validation significantly predicted reliability, $\beta = 0.16$	H4	Validation supported detection readiness
Test effect of framework design quality on performance	Design quality was the strongest predictor, $\beta = 0.34$	H5	Design quality was central to performance
Examine role of analyst engagement	Engagement significantly predicted performance, $\beta = 0.18$	H6	Engagement supported adaptive performance
Identify maturity and risk-control gaps	Overall maturity = 4.03; largest gaps in feature/data quality and analyst trust	Supports overall model	Data- and human-facing controls required improvement

Table 11 summarized the major findings in relation to the research objectives and hypotheses, showing that all objectives were achieved and all hypotheses supported. The first objective was achieved because anomaly detection modeling showed a strong positive relationship with framework design quality, $r = 0.75$, confirming that stronger anomaly-detection capability improved the coherence of framework design. The second objective was achieved because predictive risk scoring showed a significant positive relationship with detection performance, $r = 0.69$ and $\beta = 0.26$, demonstrating that calibrated risk scoring contributed to stronger outcomes. The third objective was achieved because feature engineering and data quality significantly predicted detection performance, $\beta = 0.19$, showing that feature construction and data cleaning improved confidence in results. The fourth objective was achieved because model validation and performance significantly predicted detection reliability, $\beta = 0.16$, the smallest but still significant coefficient, showing that rigorous evaluation contributed to detection quality. The fifth objective was achieved because framework design quality was the strongest predictor, $\beta = 0.34$ and $p < 0.001$, confirming the central argument of the research. The sixth objective was supported because analyst and institutional engagement significantly predicted performance, $\beta = 0.18$, linking the results to the act and adapt phases of the DSEA loop. The maturity index showed an overall score of 4.03, while the priority matrix identified feature and data quality, analyst trust, false-positive control, and model interpretability as the highest-priority improvement areas. Overall, the results proved that fraud detection performance in financial institutions depended on the combined effect of anomaly detection modeling, predictive risk scoring, feature engineering, model validation, analyst engagement, and framework design quality.

FINDINGS

This chapter presents the findings of the quantitative analysis conducted to examine how a machine learning-based fraud detection framework, integrating unsupervised anomaly identification with supervised predictive risk scoring, influences fraud detection performance in financial records. The analysis was based on data collected through a structured five-point Likert-scale questionnaire, where 1 represented Strongly Disagree and 5 represented Strongly Agree. A total of 160 questionnaires were distributed to professionals involved in fraud analytics, data science, auditing, forensic accounting, risk, compliance, and financial-systems administration, of which 142 valid responses were retained for analysis, producing a valid response rate of 88.8%.

The descriptive findings showed that all seven study variables were rated in the high-to-very-high range, with fraud detection performance recording the highest mean of 4.21 and feature engineering and data quality recording the lowest mean of 3.92. This pattern indicated broad agreement that the framework's components were positively present, while also signaling that feature and data-facing activities were perceived as comparatively weaker. The reliability findings confirmed that all constructs were internally consistent, with Cronbach's alpha values ranging from 0.81 to 0.93, allowing the study to proceed confidently to inferential analysis.

The correlation findings showed that all six hypothesized relationships were positive and statistically significant at $p < 0.01$, with framework design quality showing the strongest relationship with detection performance, $r = 0.78$, and anomaly detection modeling showing a strong relationship with framework design quality, $r = 0.75$. The regression findings showed that the six predictors together explained 71.4% of the variance in detection performance, with framework design quality the strongest predictor, $\beta = 0.34$, followed by predictive risk scoring, $\beta = 0.26$, and anomaly detection modeling, $\beta = 0.24$. The hypotheses-testing findings confirmed that all six hypotheses were supported. The maturity index found overall high maturity, 4.03, with the strongest maturity in data ingestion, feature engineering, and anomaly development, and the weakest in validation, calibration, and analyst review. The priority matrix identified feature and data quality, analyst trust, false-positive control, and model interpretability as the highest-priority improvement areas. Collectively, these findings demonstrate that a machine-learning framework combining anomaly identification and predictive risk scoring can substantially strengthen fraud detection in financial records, and that its reliability depends jointly on technical capability, methodological validation, and analyst engagement.

DISCUSSION

The findings of this study provide strong empirical support for the proposition that a machine-learning framework combining anomaly identification and predictive risk scoring substantially strengthens

fraud detection in financial records, and that its reliability depends on the joint operation of technical capability, methodological rigor, and institutional engagement (Arif Uz et al., 2025; Hossan, 2025). The result that framework design quality was the strongest predictor of detection performance, $\beta = 0.34$, is consistent with the study's conceptual argument that the value of anomaly detection and risk scoring is realized not through any single technique but through the coherent integration of detection, scoring, explanation, and analyst engagement (Abdallah et al., 2016). Interpreted through the dynamic DSEA loop, framework design quality captures how smoothly information circulates across the detect, score, explain, act, and adapt phases, so its dominance as a predictor reflects the loop's premise that performance emerges from the whole cycle rather than any single stage (Kanti, 2025; Shima Ali et al., 2025).

The strong relationship between anomaly detection modeling and framework design quality, $r = 0.75$, and the significant regression effect of anomaly detection on performance, $\beta = 0.24$, confirm the central role of unsupervised anomaly identification as a first-stage filter. These results align with prior studies showing that anomaly detection can surface novel and previously unseen fraud that rule-based and purely supervised systems miss (Chandola et al., 2009). At the same time, the comparatively lower means and maturity scores for feature engineering and validation suggest that respondents recognized the well-documented dependence of detection models on data quality; a detector is only as reliable as the features and labels that feed it (Bhattacharyya et al., 2011). The prominence of feature drift and data quality in the priority matrix, where it recorded the largest gap score of 1.15, reinforces the interpretation that practitioners view data quality as essential yet under-implemented, a tension the framework must resolve to remain effective against adaptive fraud (Amir, 2026; Mahmudul & Sadia, 2026; Mst Shurovi, 2026).

The significant contribution of predictive risk scoring, $\beta = 0.26$, the second-strongest predictor, supports the argument that calibrated probability scoring is decisive for prioritizing scarce investigative attention. By assigning each flagged record a calibrated fraud probability, the framework enables institutions to rank cases and control the trade-off between catching fraud and generating false alarms (Hand, 2009). This finding is consistent with the imbalanced-learning literature, which stresses that naive accuracy is misleading for rare events and that cost-sensitive scoring and appropriate evaluation metrics are essential (He & Garcia, 2009). The very-high-priority classification of false-positive rate control in the priority matrix, with a gap of 1.03, underscores that reducing false alarms remains a central practical challenge, since excessive false positives erode analyst trust and waste investigative capacity (Shima Ali et al., 2026; Nurul, 2026).

The significant effect of analyst and institutional engagement, $\beta = 0.18$, provides direct empirical support for the socio-technical and vigilant-decision perspectives that underpin the study. This result indicates that even accurate, well-calibrated predictions do not achieve their full value unless analysts trust them and institutions act on them (Davis, 1989). The very-high-priority classification of analyst trust and explanation quality, with a gap of 1.11, together with the strong emphasis on model interpretability, suggests that the human-facing dimension of the framework remains the least mature: technical modeling capacity has advanced faster than the interpretive and institutional capacity needed to convert predictions into trusted action (Lockey et al., 2021). This pattern is precisely what the DSEA loop anticipates, since the act and adapt phases depend on analyst judgment and feedback to sustain performance against an adaptive adversary. Taken together, the results describe a framework that is technically strong at its core, with high maturity in data ingestion, feature engineering, and anomaly development, but comparatively weaker at its data-facing and human-facing margins, where feature quality, validation, interpretability, and analyst trust lag. The greatest gains in reliability are therefore likely to come not from further improving already-strong modeling but from strengthening feature and data quality, controlling false positives, deepening model interpretability, and investing in the analyst engagement that converts predictions into trusted, adaptive action (Rudin, 2019). This interpretation is robust across the descriptive, correlational, regression, maturity, and priority analyses, which consistently identify the same set of improvement areas.

CONCLUSION

This study set out to assess how a machine learning-based fraud detection framework, integrating unsupervised anomaly identification with supervised predictive risk scoring, influences fraud

detection performance in financial records. Using a quantitative, cross-sectional, case-based design and data from 142 valid respondents, the study found that all six hypothesized relationships were supported and that the framework's constructs together explained 71.4% of the variance in fraud detection performance. Framework design quality emerged as the strongest predictor, underscoring that the value of anomaly detection and risk scoring is realized through coherent integration rather than isolated technique. Anomaly detection modeling and predictive risk scoring were confirmed as central technical capabilities, while feature engineering and data quality, model validation, and analyst engagement were identified as significant but comparatively less mature contributors.

The study concludes that a machine-learning framework combining anomaly identification and predictive risk scoring offers a substantial improvement over reactive, rule-bound, and opaque fraud detection, because it surfaces both known and emerging fraud, prioritizes scarce investigative attention through calibrated scores, and supports adaptation against an evolving adversary. However, the study also concludes that the reliability of such a framework is fundamentally socio-technical: its performance depends not only on the quality of anomaly detection and risk scoring but also on the rigor of feature engineering and validation and the strength of the interpretive and institutional arrangements that convert predictions into trusted action. The consistent identification of feature and data quality, false-positive control, model interpretability, and analyst trust as the highest-priority improvement areas points to where future effort should be concentrated. In sum, the framework provides a promising and empirically supported basis for strengthening fraud detection in financial records, provided that its technical strengths are matched by corresponding investment in data quality, validation, explainability, and analyst engagement.

RECOMMENDATIONS

Based on the findings, several recommendations are offered for institutions, analysts, and regulators engaged in machine-learning fraud detection. First, organizations should strengthen feature engineering and data quality by monitoring feature drift, validating data pipelines, and documenting data-quality procedures, since feature drift and data quality recorded the largest improvement gap. Second, institutions should invest in false-positive control and calibrated risk scoring, tuning decision thresholds and cost-sensitive objectives so that alerts are both accurate and manageable for investigators.

Third, organizations should invest in model interpretability and analyst-facing explanation, prioritizing feature-attribution and reason-code methods that make each flagged case understandable, so that predictions can be trusted, justified to regulators, and acted upon. Fourth, institutions should establish routine model-validation, backtesting, and retraining procedures, with consistent governance and documentation, to keep models effective against an adaptive adversary and to satisfy audit requirements. Fifth, organizations should prioritize the integration of real-time transaction scoring and account-takeover detection, given their high importance and moderate effectiveness. Finally, cross-disciplinary coordination among fraud analysts, data scientists, auditors, and compliance officers should be formalized, consistent with the socio-technical finding that reliable fraud detection depends on collaboration across professional groups.

LIMITATIONS OF THE STUDY

This study has several limitations that should be considered when interpreting its findings. First, the cross-sectional design captured perceptions at a single point in time and therefore cannot establish causal direction or track how framework performance changes as fraud patterns and models evolve; a longitudinal design would allow stronger causal inference and better capture the adversarial drift central to fraud detection. Second, the study relied on a structured, perception-based questionnaire completed by professionals, so the results reflect informed expert assessment rather than direct measurement of detection accuracy, false-positive rates, or losses avoided; future work could pair perception data with technical performance metrics such as validated precision, recall, and cost-based measures. Third, the purposive sampling strategy, while appropriate for reaching knowledgeable respondents, limits the generalizability of the findings to the broader population and to institutions not represented in the sample. Fourth, the framework and its constructs were assessed at a general level across diverse financial contexts, and the relative importance of specific methods and controls may vary considerably between payment fraud, account takeover, financial-statement fraud, and money

laundering. Finally, the study did not directly evaluate the cost, data-availability, privacy, or adversarial-robustness constraints that may affect real-world implementation, which represent important avenues for future research.

REFERENCES

- [1]. Abdallah, A., Maarof, M. A., & Zainal, A. (2016). Fraud detection system: A survey. *Journal of Network and Computer Applications*, 68, 90–113. <https://doi.org/10.1016/j.jnca.2016.04.007>
- [2]. Abu Naser Md Golam, M., & Amir, R. (2023). Digital Twin Modeling of Fiber Optic Telecom Infrastructure for Proactive SCADA Network Resilience in Power Transmission Utilities. *International Journal of Scientific Interdisciplinary Research*, 4(4), 413–448. <https://doi.org/10.63125/04mgf566>
- [3]. ACFE. (2022). Occupational fraud 2022: A report to the nations. Association of Certified Fraud Examiners. <https://www.acfe.com/report-to-the-nations/2022/>
- [4]. Ali, A., Abd Razak, S., Othman, S. H., Eisa, T. A. E., Al-Dhaqm, A., Nasser, M., ... Saif, A. (2022). Financial fraud detection based on machine learning: A systematic literature review. *Applied Sciences*, 12(19), 9637. <https://doi.org/10.3390/app12199637>
- [5]. Amir, R. (2024). Commissioning Failures in Integrated Life Safety Systems (FA + HVAC Shutdown + Elevators): A Multi-Site Case Study Based on Field Testing and Authority Inspections. *American Journal of Data Science and Analytics*, 5(12), 280-313. <https://doi.org/10.63125/nst5ee49>
- [6]. Amir, R. (2025). Reducing False and Improving Response Time in High-Occupancy Buildings: A Quantitative Study of NFPA 72-Complaint Cause-and-Effect Testing Outcomes. *American Journal of Scholarly Research and Innovation*, 4(01), 856–893. <https://doi.org/10.63125/s18ss614>
- [7]. Amir, R. (2026). AI-Enabled Predictive Maintenance for Fire Alarm and Smoke Management Systems: A Systematic Review of Models. *American Journal of Interdisciplinary Studies*, 7(01), 496-535. <https://doi.org/10.63125/q3aw4y77>
- [8]. Amir, R., & Chapal, B. (2022). Federated Learning Architectures for Distributed Engineering Project Knowledge Management A Simulation-Based Analysis. *Review of Applied Science and Technology*, 1(04), 411-446. <https://doi.org/10.63125/4k4pjr64>
- [9]. Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- [10]. Beneish, M. D. (1999). The detection of earnings manipulation. *Financial Analysts Journal*, 55(5), 24–36. <https://doi.org/10.2469/faj.v55.n5.2296>
- [11]. Bhattacharyya, S., Jha, S., Tharakunnel, K., & Westland, J. C. (2011). Data mining for credit card fraud: A comparative study. *Decision Support Systems*, 50(3), 602–613. <https://doi.org/10.1016/j.dss.2010.08.008>
- [12]. Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. *Statistical Science*, 17(3), 235–255. <https://doi.org/10.1214/ss/1042727940>
- [13]. Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- [14]. Breunig, M. M., Kriegel, H.-P., Ng, R. T., & Sander, J. (2000). LOF: Identifying density-based local outliers. *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, 93–104. <https://doi.org/10.1145/335191.335388>
- [15]. Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys*, 41(3), 1–58. <https://doi.org/10.1145/1541880.1541882>
- [16]. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321–357. <https://doi.org/10.1613/jair.953>
- [17]. 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*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- [18]. Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>
- [19]. Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C., & Bontempi, G. (2018). Credit card fraud detection: A realistic modeling and a novel learning strategy. *IEEE Transactions on Neural Networks and Learning Systems*, 29(8), 3784–3797. <https://doi.org/10.1109/TNNLS.2017.2736643>
- [20]. Dal Pozzolo, A., Caelen, O., Le Borgne, Y.-A., Waterschoot, S., & Bontempi, G. (2014). Learned lessons in credit card fraud detection from a practitioner perspective. *Expert Systems with Applications*, 41(10), 4915–4928. <https://doi.org/10.1016/j.eswa.2014.02.026>
- [21]. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- [22]. Fiore, U., De Santis, A., Perla, F., Zanetti, P., & Palmieri, F. (2019). Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. *Information Sciences*, 479, 448–455. <https://doi.org/10.1016/j.ins.2017.12.030>
- [23]. Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM Computing Surveys*, 46(4), 1–37. <https://doi.org/10.1145/2523813>
- [24]. Hand, D. J. (2009). Measuring classifier performance: A coherent alternative to the area under the ROC curve. *Machine Learning*, 77(1), 103–123. <https://doi.org/10.1007/s10994-009-5119-5>
- [25]. He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9), 1263–1284. <https://doi.org/10.1109/TKDE.2008.239>

- [26]. Hilal, W., Gadsden, S. A., & Yawney, J. (2021). Financial fraud: A review of anomaly detection techniques and recent advances. *Expert Systems with Applications*, 193, 116429. <https://doi.org/10.1016/j.eswa.2021.116429>
- [27]. Hollnagel, E., Woods, D. D., & Leveson, N. (Eds.). (2006). *Resilience engineering: Concepts and precepts*. Ashgate. <https://doi.org/10.1201/9781315605685>
- [28]. Jurgovsky, J., Granitzer, M., Ziegler, K., Calabretto, S., Portier, P.-E., He-Guelton, L., & Caelen, O. (2018). Sequence classification for credit-card fraud detection. *Expert Systems with Applications*, 100, 234-245. <https://doi.org/10.1016/j.eswa.2018.01.037>
- [29]. Kazi Rakib Hasan, S., & Uddin, H. M. M. (2022). Scalable AI For Project Portfolio Management: A Mixed-Methods Study Combining Distributed Computing Benchmarks. *Review of Applied Science and Technology*, 1(04), 375-410. <https://doi.org/10.63125/0kk4wf20>
- [30]. Krawczyk, B. (2016). Learning from imbalanced data: Open challenges and future directions. *Progress in Artificial Intelligence*, 5(4), 221-232. <https://doi.org/10.1007/s13748-016-0094-0>
- [31]. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444. <https://doi.org/10.1038/nature14539>
- [32]. Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation forest. *Proceedings of the 2008 Eighth IEEE International Conference on Data Mining*, 413-422. <https://doi.org/10.1109/ICDM.2008.17>
- [33]. Lockey, S., Gillespie, N., Holm, D., & Someh, I. A. (2021). A review of trust in artificial intelligence: Challenges, vulnerabilities and future directions. *Proceedings of the 54th Hawaii International Conference on System Sciences*, 5463-5472. <https://doi.org/10.24251/HICSS.2021.664>
- [34]. Lu, J., Liu, A., Dong, F., Gu, F., Gama, J., & Zhang, G. (2019). Learning under concept drift: A review. *IEEE Transactions on Knowledge and Data Engineering*, 31(12), 2346-2363. <https://doi.org/10.1109/TKDE.2018.2876857>
- [35]. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765-4774. <https://doi.org/10.48550/arXiv.1705.07874>
- [36]. Mahmudul, H., & Sadia, Z. (2026). Safeguarding Financial Records: A Cybersecurity-Driven Management Framework. *International Journal of Scientific Interdisciplinary Research*, 7(1), 493-526. <https://doi.org/10.63125/zch4s169>
- [37]. Md Arif Uz, Z., Sharmin, S., & Md Abdur, R. (2025). Factors Influencing Behavioural Intention to Use he Innovative Open And Distance Learning Systems With The Role Of Continuance Intention. *American Journal of Scholarly Research and Innovation*, 4(01), 202-220. <https://doi.org/10.63125/pbjxp014>
- [38]. Md Tohidul, I. (2023). ERP-Integrated Financial Analytics Systems for Corporate Financial Reporting, Cash Flow Analysis, and Strategic Investment Planning. *American Journal of Advanced Technology and Engineering Solutions*, 3(04), 87-128. <https://doi.org/10.63125/qhk96h35>
- [39]. Md. Abdur, R., & Iftekhar, A. (2021). Customer Retention Forecasting in Mobile Wallet Services Using Neural Networks: A Comparative Quantitative Study. *International Journal of Business and Economics Insights*, 1(4), 70-102. <https://doi.org/10.63125/dyrpc387>
- [40]. Mohammad Badrul, A., & Md. Mominul, H. (2023). Machine Learning-Based Assessment of Environmental and Public Health Risks in Sewerage Sludge Management Systems. *American Journal of Advanced Technology and Engineering Solutions*, 3(02), 33-77. <https://doi.org/10.63125/y3yfxa92>
- [41]. Mohammad Shoriful Hossan, S. (2024). Development of A Predictive Maintenance Model for Industrial Motor Drives and VFD Systems Using PLC/DCS Operational Telemetry. *American Journal of Scholarly Research and Innovation*, 3(01), 198-242. <https://doi.org/10.63125/pv1js496>
- [42]. Mohammad Shoriful Hossan, S. (2025). Development Of an ARC-Flash Hazard Mitigation Framework for Brownfield Industrial Power Distribution Retrofits in U.S. Manufacturing. *American Journal of Advanced Technology and Engineering Solutions*, 1(02), 262-303. <https://doi.org/10.63125/vhqjgm79>
- [43]. Mohammad Shoriful Hossan, S., & Adar, C. (2023). SIL-Compliant Safety Instrumented System Design for Oil & Gas Hazardous-Area Operations Using IEC 61508 / IEC 61511 Frameworks. *Journal of Sustainable Development and Policy*, 2(02), 43-85. <https://doi.org/10.63125/cb3d3w81>
- [44]. Mst Shurovi, A. (2026). AI-Driven Accounts Payable and Receivable Automation for Operational Risk Mitigation in U.S. SMEs: A Systematic Review (2018-2026). *American Journal of Interdisciplinary Studies*, 7(01), 536-577. <https://doi.org/10.63125/x1yzwc45>
- [45]. Muhammad Mohiul, I., & Mohammad Badrul, A. (2022). Predictive Analytics for Optimizing Sewerage Sludge Treatment Efficiency and Resource Recovery in Dhaka City. *Journal of Sustainable Development and Policy*, 1(04), 158-200. <https://doi.org/10.63125/t9zdp458>
- [46]. Mukut Kanti, B. (2025). Lifecycle Carbon Reduction and Net Zero Pathway Modeling Using LEED, EDGE, and SBTI Frameworks in Developing Economy Industrial Buildings. *American Journal of Interdisciplinary Studies*, 6(3), 306-350. <https://doi.org/10.63125/tnneth75>
- [47]. Mukut Kanti, B., & Rony, S. (2022). Quantitative Risk Assessment and Safety Performance Benchmarking for Structural Integrity Management in Retail Supply Chain Infrastructure. *American Journal of Data Science and Analytics*, 3(12), 42-88. <https://doi.org/10.63125/n88vth90>
- [48]. Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems*, 50(3), 559-569. <https://doi.org/10.1016/j.dss.2010.08.006>
- [49]. Pang, G., Shen, C., Cao, L., & van den Hengel, A. (2021). Deep learning for anomaly detection: A review. *ACM Computing Surveys*, 54(2), 1-38. <https://doi.org/10.1145/3439950>
- [50]. Perols, J. (2011). Financial statement fraud detection: An analysis of statistical and machine learning algorithms. *Auditing: A Journal of Practice & Theory*, 30(2), 19-50. <https://doi.org/10.2308/ajpt-50009>

- [51]. Provost, F., & Fawcett, T. (2001). Robust classification for imprecise environments. *Machine Learning*, 42(3), 203–231. <https://doi.org/10.1023/A:1007601015854>
- [52]. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). “Why should I trust you?”: Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. <https://doi.org/10.1145/2939672.2939778>
- [53]. Risha, A., & Kazi Mohammad Khalid, A. (2023). A Meta-Analysis of AI-Driven Geospatial Analytics for Predictive Maintenance of Critical Infrastructure in Developing Economies. *International Journal of Scientific Interdisciplinary Research*, 4(4), 375–412. <https://doi.org/10.63125/rayrex49>
- [54]. Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206–215. <https://doi.org/10.1038/s42256-019-0048-x>
- [55]. Ruff, L., Kauffmann, J. R., Vandermeulen, R. A., Montavon, G., Samek, W., Kloft, M., ... Müller, K.-R. (2021). A unifying review of deep and shallow anomaly detection. *Proceedings of the IEEE*, 109(5), 756–795. <https://doi.org/10.1109/JPROC.2021.3052449>
- [56]. Sadia, Z., Mohammad Ariful, I., Sumyta, H., & Hosne Ara, M. (2022). A Review of AI-Powered Data Visualization in Enterprise Reporting: Dashboard Design And Interactive Analytics. *American Journal of Advanced Technology and Engineering Solutions*, 2(01), 32-54. <https://doi.org/10.63125/gabst658>
- [57]. Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS ONE*, 10(3), e0118432. <https://doi.org/10.1371/journal.pone.0118432>
- [58]. Schölkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., & Williamson, R. C. (2001). Estimating the support of a high-dimensional distribution. *Neural Computation*, 13(7), 1443–1471. <https://doi.org/10.1162/089976601750264965>
- [59]. Shima Ali, S., Musomi, K., & Sonya, G. (2024). Integrated Infectious Disease Preparedness: Multilingual Health Communication Strategies for Public Health Risk Mitigation, Disease Intervention and Prevention for Culturally Diverse Population in United States. *American Journal of Data Science and Analytics*, 5(10), 01-47. <https://doi.org/10.63125/ehk4mt02>
- [60]. Shima Ali, S., Sonya, G., & Mahbubul, I. (2025). Telemedicine And Remote Patient Monitoring for Preventing Risk and Chronic Disease Management in Underserved Communities. *American Journal of Health and Medical Sciences*, 6(04), 01-42. <https://doi.org/10.63125/y0e9dc38>
- [61]. Shima Ali, S., Sonya, G., & Mahbubul, I. (2026). STI and HIV Prevention Through Public Health Digital Twins: A Framework for Personalized Prevention and Adaptive Disease Intervention. *American Journal of Scholarly Research and Innovation*, 5(01), 284–326. <https://doi.org/10.63125/rjgna882>
- [62]. Syed Nurul, I. (2026). A Machine Learning Approach to Intelligent Streetlight Control for Sustainable Energy Management in Smart Cities. *American Journal of Advanced Technology and Engineering Solutions*, 6(01), 650-689. <https://doi.org/10.63125/gxwn1g63>
- [63]. Trist, E. (1981). The evolution of socio-technical systems: A conceptual framework and an action research program. Ontario Quality of Working Life Centre. <https://doi.org/10.4324/9781315227979>
- [64]. Weick, K. E., & Sutcliffe, K. M. (2007). *Managing the unexpected: Resilient performance in an age of uncertainty* (2nd ed.). Jossey-Bass. <https://www.wiley.com/en-us/Managing+the+Unexpected%3A+Resilient+Performance+in+an+Age+of+Uncertainty%2C+2nd+Edition-p-9780787996499>
- [65]. West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. *Computers & Security*, 57, 47–66. <https://doi.org/10.1016/j.cose.2015.09.005>
- [66]. Zaman, S. (2024). A Systematic Review of ERP And CRM Integration For Sustainable Business And Data Management in Logistics And Supply Chain Industry. *Frontiers in Applied Engineering and Technology*, 1(01), 204-221. <https://doi.org/10.70937/faet.v1i01.36>