



International Journal of Research Publication and Reviews

Journal Homepage: www.ijrpr.com ISSN 2582-7421

The Role of Forensic Auditing in Detecting Financial Fraud and Strengthening Accountability in Public Revenue Institutions

Kabirat Olamide Mayegun

Drexel University, USA

ABSTRACT

Public revenue institutions play a central role in mobilizing financial resources, sustaining essential services, and maintaining public confidence in government accountability. However, persistent revenue leakages, fraudulent reporting, procurement manipulation, collusion, and weaknesses in internal controls continue to undermine fiscal integrity and institutional performance. Against this background, forensic auditing has emerged as a specialized approach combining investigative accounting, digital evidence analysis, fraud risk assessment, and legal scrutiny to identify financial irregularities. This article examines the role of forensic auditing in detecting financial fraud and strengthening accountability within public revenue institutions. It explores investigative techniques, transaction tracing, data analytics, evidence preservation, and institutional reporting mechanisms that support fraud detection and prevention. Particular attention is given to auditor independence, governance transparency, regulatory enforcement, whistleblower protection, and coordination between audit agencies and anti-corruption authorities. The discussion also considers implementation barriers, including inadequate technical capacity, political interference, fragmented information systems, and limited access to reliable records. Ultimately, the article proposes an integrated forensic auditing and accountability framework to improve financial oversight, facilitate timely investigations, reinforce internal controls, and strengthen public trust in revenue administration.

Keywords: Forensic Auditing; Financial Fraud Detection; Public Revenue Institutions; Institutional Accountability; Financial Transparency; Fraud Risk Management.

1. INTRODUCTION

1.1 Public Revenue Administration and Financial Accountability

Public revenue administration mobilizes resources through taxation, customs duties, levies, licensing fees, and statutory collections that finance government functions. Reliable revenue generation supports fiscal sustainability, infrastructure investment, public expenditure, and economic stability [1]. Institutional credibility depends on transparent accounting, remittances, and stewardship of public resources. Effective internal controls establish authorization requirements, segregation of duties, reconciliation procedures, and independent monitoring across collection processes [2]. These safeguards strengthen compliance with financial regulations and improve confidence in government reporting. Conversely, revenue leakages, misappropriation, and inaccurate financial records undermine budget execution and reduce resources available for public services [3]. Persistent weaknesses can erode taxpayer confidence, weaken voluntary compliance, and compromise institutional legitimacy. Strengthening accountability therefore requires preventive controls alongside mechanisms capable of identifying concealed irregularities.

1.2 Financial Fraud and Institutional Control Weaknesses

Financial fraud in revenue institutions includes fraudulent refunds, unauthorized adjustments, fictitious transactions, diverted collections, collusion, and manipulated financial statements [4]. Such misconduct exploits fragmented information systems, inconsistent transaction identifiers, inadequate reconciliation, and poorly enforced access privileges. When assessment, payment, and remittance databases operate independently, discrepancies may remain unresolved across reporting periods [2]. Conventional audits provide reasonable assurance through risk assessment, control testing, and selective examination rather than exhaustive investigation of every transaction. Consequently, coordinated schemes and deliberately concealed records may evade routine scrutiny [5]. Complex revenue environments therefore require investigative methods capable of reconstructing transaction histories, linking related accounts, and identifying unusual behavioral patterns. Continuous proactive analytical monitoring can complement established controls by prioritizing suspicious records for further examination [6].

1.3 Forensic Auditing and Emerging Analytical Technologies

Forensic auditing combines investigative accounting, documentary verification, transaction reconstruction, and legally defensible evidence handling to examine suspected misconduct [7]. Unlike conventional financial auditing, its primary emphasis is establishing the circumstances, mechanisms, and evidentiary basis of particular irregularities. Investigators employ transaction tracing, digital evidence examination, anomaly detection, and interviews to test competing explanations. Statistical methods and machine learning can extend these procedures by identifying abnormal payment sequences, refund concentrations, and relationships hidden within datasets [6]. Supervised classification supports prioritization when reliable historical labels exist, whereas unsupervised approaches can reveal previously unrecognized deviations. Nevertheless, suspicious predictions do not establish fraud independently [8]. Audit trails, controlled evidence access, model validation, interpretability, and human review remain necessary to preserve procedural fairness and investigative credibility.

1.4 Research Objectives, Contributions, and Article Organization

Existing approaches insufficiently connect forensic investigation, predictive risk scoring, and institutional accountability [5]. This study develops an integrated framework encompassing financial data acquisition, preprocessing, feature engineering, model training, testing, benchmarking, and evidentiary review. Its objectives are to identify suspicious revenue activity, evaluate detection performance against metrics, and translate findings into strengthened governance controls [8]. The framework emphasizes reproducible analysis and human oversight. Subsequent sections examine conceptual foundations, research methods, analytical evaluation, comparative results, governance implications, and conclusions. Section 2 establishes the theoretical and institutional basis for fraud-risk assessment and accountability-oriented financial oversight.

2. THEORETICAL FOUNDATIONS AND FORENSIC AUDITING FRAMEWORK

2.1 Fraud Triangle, Fraud Diamond, and Institutional Fraud Risk

The Fraud Triangle explains fraudulent conduct through three interacting conditions: perceived pressure, opportunity, and rationalization [5]. Financial pressures or institutional performance expectations may motivate misconduct, while weak controls create opportunities for unauthorized transactions. Rationalization enables perpetrators to justify unethical actions through perceived entitlement or organizational norms [6]. The Fraud Diamond extends this explanation by incorporating capability, emphasizing the knowledge, authority, and technical competence required to exploit institutional vulnerabilities. Within revenue institutions, discretionary approval powers, inadequate supervision, and unrestricted system privileges can facilitate fraudulent refunds, concealed adjustments, and revenue diversion [7]. Nevertheless, behavioral theories primarily explain underlying motivations rather than directly observable transactional characteristics. Their application to predictive detection therefore requires operational indicators reflecting abnormal authorization patterns, transaction frequencies, access privileges, and financial discrepancies [8]. Such indicators provide measurable proxies without establishing individual fraudulent intent conclusively.

2.2 Agency Theory, Internal Controls, and Public Accountability

Agency theory conceptualizes relationships between citizens, public officials, and oversight institutions through delegated authority and potentially conflicting interests [9]. Public administrators exercise responsibility over financial resources while citizens and supervisory bodies possess incomplete information concerning institutional activities. This information asymmetry creates opportunities for opportunistic behavior, particularly where monitoring mechanisms and enforcement incentives remain inadequate. Effective internal controls mitigate agency risks through segregation of duties, authorization hierarchies, transaction reconciliation, and independent audit functions [10]. However, collusion and managerial override can compromise otherwise established safeguards. Forensic auditing strengthens institutional accountability by examining suspicious transactions, documenting control deficiencies, and supporting evidence-based corrective measures [11]. Transparent investigative reporting consequently reduces information disparities, reinforces public oversight, and promotes responsible financial stewardship across revenue administration systems.

2.3 Forensic Auditing Techniques and Digital Evidence Management

Forensic auditing employs investigative accounting, transaction reconstruction, reconciliation testing, and digital analytics to examine suspected financial misconduct [12]. Investigative accounting traces monetary movements across collection accounts, refund registers, and remittance records, while reconciliation testing identifies inconsistencies between authorized assessments and recorded receipts. Forensic analytics extends these procedures through exception reporting, relational analysis, and statistical anomaly identification [8]. Digital evidence management requires controlled acquisition, cryptographic integrity verification, documented access, and traceable chain-of-custody procedures. These safeguards preserve evidentiary reliability throughout investigative examination and subsequent proceedings [13]. Importantly, suspicious transactions represent preliminary risk indicators, confirmed control violations establish procedural noncompliance, and substantiated fraud requires sufficient evidence supporting deliberate deception. Automated classifications must therefore undergo independent investigative verification before disciplinary, administrative, or judicial conclusions are reached [11]. This distinction protects procedural fairness and strengthens institutional decision-making credibility.

2.4 Integrated Forensic Auditing and Accountability Framework

The proposed framework integrates revenue transactions, institutional controls, fraud-risk indicators, analytical detection, investigative verification, and governance responses [10]. Transactional records encompassing collections, refunds, adjustments, and reconciliations constitute the primary analytical inputs. Financial indicators undergo statistical examination and predictive modeling to generate risk scores supporting case prioritization [8]. Investigators subsequently assess flagged transactions through documentary verification, evidence integrity checks, and control deficiency evaluation. Confirmed findings inform corrective actions, potential revenue recovery, enhanced transparency, and supervisory intervention [13]. Investigation outcomes also provide feedback for updating fraud labels, refining predictive features, recalibrating decision thresholds, and evaluating subsequent detection performance. Auditor independence, due process, data protection, and regulatory compliance operate across all framework components. This integrated structure establishes the conceptual foundation for reproducible experimental evaluation and accountability-oriented institutional improvement.

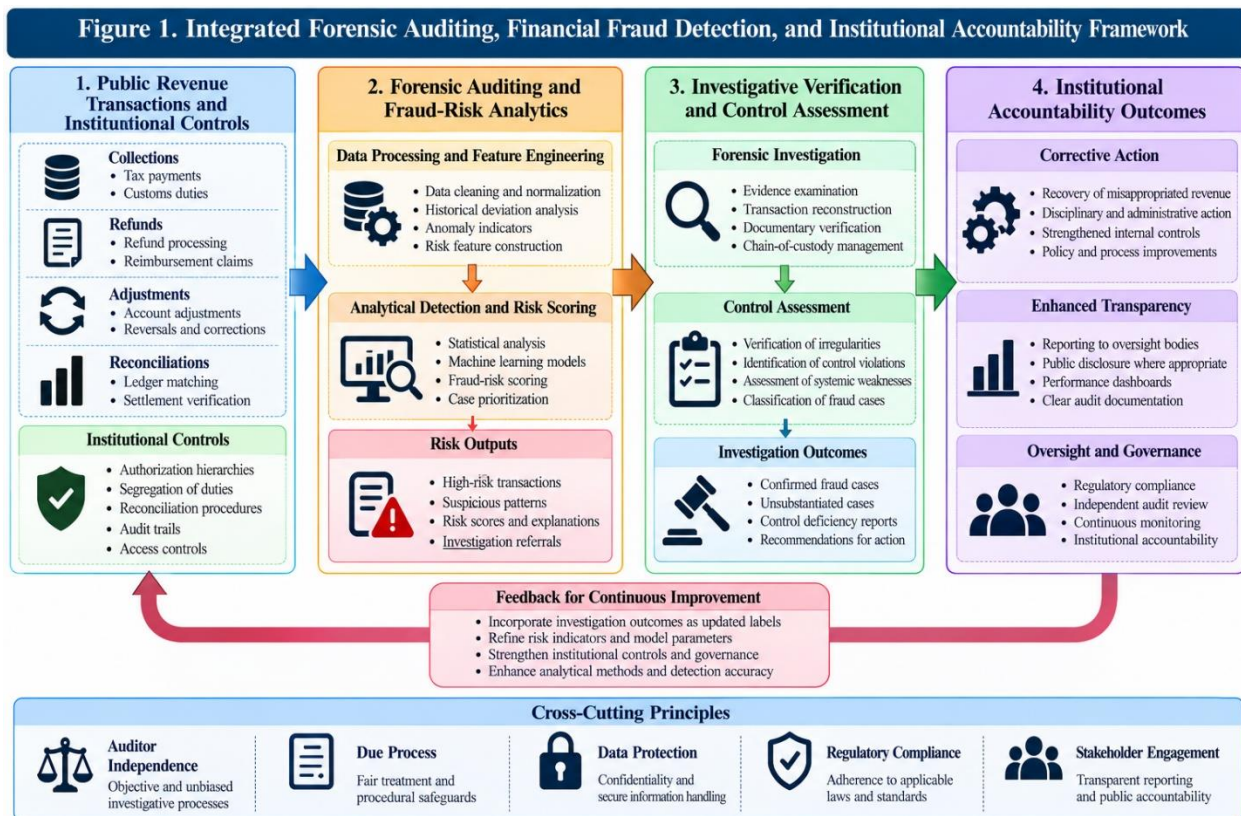


Figure 1. Integrated Forensic Auditing, Financial Fraud Detection, and Institutional Accountability Framework

3. DATA ACQUISITION AND EXPERIMENTAL RESEARCH DESIGN

3.1 Research Design and Institutional Study Context

The study adopts a quantitative retrospective design complemented by forensic case analysis to examine financial fraud detection within public revenue institutions [11]. The proposed institutional setting encompasses tax authorities, customs administrations, and government revenue collection agencies operating computerized financial management systems. Individual financial transactions constitute the primary analytical units, supplemented by account histories and linked transaction sequences. A proposed five-year historical observation period provides sufficient temporal coverage for examining recurring irregularities and institutional control changes [12]. Eligible records must contain verifiable transaction identifiers, financial attributes, and documented processing histories. These procedures describe the intended methodology rather than completed institutional data collection.

3.2 Financial Transaction and Revenue Data Acquisition

Proposed data acquisition encompasses revenue assessments, payment receipts, collection registers, refund authorizations, adjustment records, and financial reconciliation reports [13]. Transactional attributes include monetary amounts, processing dates, revenue categories, payment channels, account identifiers, authorization levels, and settlement references. Approved institutional personnel would extract historical records from financial management systems through controlled database queries and encrypted transfer mechanisms. Access requires institutional authorization, documented research permissions, confidentiality agreements, and compliance with applicable financial information governance requirements [14]. Data completeness would be evaluated through record counts, missing-value assessments, identifier uniqueness, and reconciliation between source extracts and institutional control

totals. Provenance documentation would capture extraction timestamps, originating systems, database versions, and responsible custodians. Cryptographic checksums would support verification that acquired records remain unchanged during subsequent processing [15]. Discrepancies between source records and extracted datasets would require documented investigation before analytical inclusion.

3.3 Audit Findings, Investigation Records, and Fraud Label Construction

Fraud outcome construction requires legally accessible audit exceptions, investigation reports, confirmed control violations, and documented case resolutions [16]. Verified fraud cases would receive positive labels only where authorized investigative findings establish sufficient supporting evidence. Transactions independently cleared through documented examination would receive negative labels, while unresolved suspicious records would remain unlabeled rather than automatically classified as legitimate. This distinction reduces misclassification arising from incomplete investigations and delayed case confirmation [17]. Each labeling decision would retain supporting case references, investigation dates, evidentiary status, and reviewer authorization. Independent forensic reviewers would examine disputed classifications and document adjudication outcomes. Because investigated transactions may disproportionately represent previously flagged high-risk activities, investigation-selection bias would be assessed through case coverage and sampling characteristics [18]. Sensitivity analyses would examine classification stability under alternative labeling assumptions.

3.4 Institutional Governance and Control Variables

Institutional variables would capture authorization hierarchies, segregation-of-duty conflicts, audit frequency, reconciliation intervals, and documented control exceptions [12]. Additional characteristics include transaction volumes, revenue categories, administrative units, and payment processing arrangements. Accountability measures would encompass investigation completion, case resolution duration, corrective action implementation, and verified revenue recovery amounts. These variables enable examination of organizational conditions associated with observed fraud-risk patterns [19]. Weak authorization controls may increase exposure to unauthorized adjustments, whereas delayed reconciliation can prolong the detection of financial discrepancies. Governance indicators would therefore support contextual interpretation of analytical findings without establishing causal relationships from observational associations alone.

3.5 Sampling Strategy and Dataset Construction

The sampling frame would comprise eligible transactions recorded during the designated five-year observation period across participating institutions [13]. Records lacking essential identifiers, containing irrecoverable corruption, or falling outside authorized institutional coverage would be excluded. Chronological ordering and account relationships would remain intact to preserve transaction dependencies and investigative context. Confirmed fraud prevalence would be documented alongside legitimate and unresolved case distributions [17]. Duplicate identifiers, reversed entries, and linked investigations would undergo structured verification before dataset finalization. Account-linked cases would remain grouped during subsequent partitioning to minimize information leakage. Sampling procedures would preserve rare fraud observations while maintaining representative institutional and temporal coverage for evaluation.

3.6 Research Ethics, Evidence Integrity, and Data Governance

Research implementation would require institutional approval, restricted access permissions, confidentiality safeguards, and documented data handling responsibilities [14]. Personal identifiers would undergo pseudonymization wherever investigative requirements permit. Immutable extraction logs, cryptographic checksums, and traceable transformations would protect evidentiary integrity [15]. Data retention periods, authorized disclosures, and secure deletion procedures would follow institutional policies and applicable regulations. Automated risk predictions would support investigative prioritization rather than establish individual culpability. Independent human verification would remain mandatory before allegations, sanctions, or administrative enforcement decisions [19].

Table 1. Proposed Data Sources, Analytical Variables, and Forensic Relevance

Data Source	Representative Variables	Analytical Purpose	Forensic Relevance
Revenue transactions	Amount, date, payment channel, revenue type	Transaction profiling	Unusual collections and payment patterns
Refund and adjustment records	Refund amount, approval date, adjustment reason	Exception identification	Unauthorized or abnormal adjustments
Audit and investigation records	Case status, findings, verified outcome	Label construction	Evidence-supported fraud classification
User and authorization logs	User role, approval sequence, access time	Control assessment	Privilege misuse and segregation conflicts

Data Source	Representative Variables	Analytical Purpose	Forensic Relevance
Reconciliation reports	Ledger balance, settlement amount, discrepancy	Financial consistency analysis	Revenue diversion and unreconciled differences
Institutional control records	Audit frequency, control exceptions, remediation status	Governance analysis	Control effectiveness and accountability

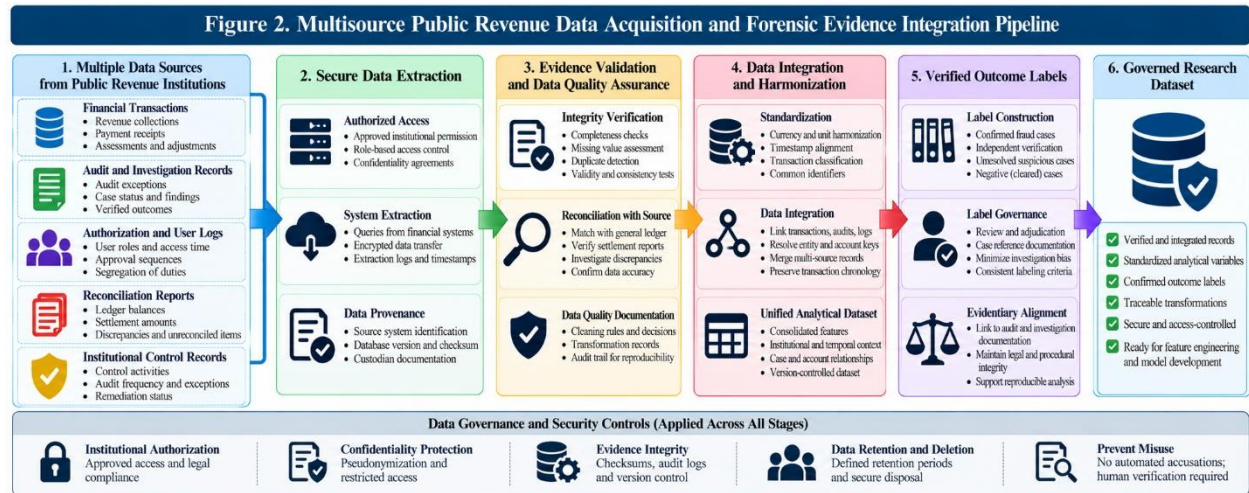


Figure 2. Multisource Public Revenue Data Acquisition and Forensic Evidence Integration Pipeline

4. DATA PREPROCESSING, FEATURE ENGINEERING, AND STATISTICAL ANALYSIS

4.1 Data Cleaning, Missing Values, and Transaction Reconciliation

Data preprocessing begins by identifying missing attributes, duplicated transactions, invalid monetary values, inconsistent identifiers, and contradictory timestamps [16]. Transaction records are reconciled against general ledgers, bank settlements, and authorized collection reports to establish financial consistency. Currency denominations, date formats, revenue classifications, and account identifiers are standardized using documented transformation rules [17]. Missing values are assessed according to their frequency, institutional origin, and investigative significance before imputation or exclusion. Legitimate reversals, refunds, and accounting adjustments are distinguished from erroneous entries through supporting documentation. Duplicate detection incorporates transaction identifiers, amounts, processing times, and payment references [18]. All corrections retain original values, transformation histories, and verification records to support reproducibility.

4.2 Historical Baseline Construction and Percentage Deviation

Historical baselines characterize expected financial activity using preceding transaction amounts, refund volumes, revenue collections, and reconciliation differences [19]. For each account or institutional unit, reference values are estimated from eligible observations occurring before the evaluated transaction. Percentage deviation measures the proportional difference between current activity and its historical baseline, supporting identification of unusual increases or decreases [20]. Positive deviations may indicate exceptional collections or inflated refunds, whereas negative deviations can reveal declining remittances or unexpected adjustments. Baseline windows are selected according to transaction frequency, seasonal variation, and institutional operating characteristics. Temporal restrictions prevent future observations from influencing historical estimates [21]. Near-zero baselines require denominator stabilization to avoid computational instability and exaggerated deviation estimates during screening.

Equation (1): Percentage Deviation from Historical Baseline

$$D_{it}(\%) = \frac{X_{it} - \bar{X}_{i,base}}{\max(|\bar{X}_{i,base}|, \varepsilon)} \times 100$$

Where:

- D_{it} : percentage deviation for entity i at time t .
- X_{it} : observed financial indicator at time t .
- $\bar{X}_{i,base}$: historical mean calculated exclusively from preceding eligible observations.

- ε : small positive constant preventing division by zero.

Parameter derivation: The historical mean is computed from the preceding eligible observations. Subtracting this baseline from the current financial value produces the absolute directional difference. Dividing by the greater of the baseline magnitude and ε generates a stabilized relative deviation, subsequently multiplied by 100 to express the result as a percentage.

4.3 Mean Absolute Deviation and Financial Variability Assessment (120 Words)

Mean absolute deviation quantifies historical financial variability by averaging absolute differences between observed values and their arithmetic mean [22]. Unlike variance, this measure retains the original monetary units, facilitating interpretation during forensic transaction screening. Separate dispersion estimates are constructed for revenue categories, payment channels, accounts, and institutional units to accommodate heterogeneous financial behavior. Transactions exhibiting substantial departures from historically stable patterns may warrant additional examination [19]. Conversely, highly variable accounts require context-sensitive thresholds to avoid excessive false-positive classifications. MAD therefore complements percentage deviation by distinguishing exceptional observations from ordinary fluctuations. Its interpretation remains conditional on baseline quality, observation frequency, and underlying transaction distributions [23]. Elevated dispersion indicates financial variability rather than independently demonstrating intentional fraud or misconduct.

Equation (2): Mean Absolute Deviation

$$MAD_i = \frac{1}{T_i} \sum_{t=1}^{T_i} |X_{it} - \bar{X}_i|$$

Where:

- MAD_i : mean absolute deviation for entity i .
- T_i : number of observations within the historical reference window.
- X_{it} : observed financial value.
- $\bar{X}_i$: arithmetic mean of the historical observations.

Parameter derivation: First, calculate $\bar{X}_i$ by dividing the sum of historical transaction values by T_i . Next, subtract the mean from each observation and obtain the absolute difference. Summing these differences and dividing by T_i produces the mean absolute deviation. Higher values indicate greater financial variability, not necessarily fraudulent behavior.

4.4 Fraud-Risk Feature Engineering and Variable Encoding

Feature engineering transforms standardized financial records into interpretable indicators associated with potentially irregular revenue activity [24]. Transaction frequency measures repeated payments or adjustments within predefined intervals, while amount irregularity captures departures from historical financial distributions. Refund concentration identifies disproportionate reimbursements associated with particular accounts, officers, or revenue categories [20]. Reconciliation discrepancies quantify differences between recorded collections and verified settlements. Authorization anomalies capture unusual approval sequences, segregation conflicts, repeated approvals, and processing outside established operating periods. Transaction sequence features identify rapid reversals, recurring beneficiaries, and interconnected payment activities [18]. Categorical variables undergo appropriate encoding, while continuous attributes receive training-derived scaling where required. Institutional governance indicators provide contextual information. Post-investigation findings are excluded from predictors to prevent outcome leakage.

4.5 Feature Selection, Multicollinearity, and Stability Assessment

Feature selection combines correlation analysis, redundancy screening, and model-based importance assessment to identify informative predictors [21]. Highly correlated variables are examined for multicollinearity, particularly where overlapping financial ratios describe similar transactional behavior. Feature importance is evaluated across chronological validation periods to assess temporal consistency and sensitivity to changing revenue practices. Distributional comparisons identify differences between institutional units, payment channels, and revenue categories [23]. Variables derived from subsequent investigations, enforcement decisions, or future transaction records are excluded to prevent information leakage. Final selection prioritizes predictive relevance, temporal stability, institutional applicability, and forensic interpretability, while retaining documented justifications for excluded attributes.

4.6 Model-Ready Feature Matrix and Data Quality Verification

The resulting analytical dataset comprises standardized predictors, verified fraud labels, institutional identifiers, and transaction timestamps [24]. Version control preserves feature definitions, transformation scripts, exclusion decisions, and source provenance. Missingness checks, consistency tests, and duplicate screening establish dataset readiness. Normalization parameters and data-dependent transformations are estimated exclusively from training

observations to prevent validation contamination [22]. Chronological ordering and case relationships remain preserved for subsequent partitioning. The finalized feature matrix supports reproducible model development, independent evaluation, and traceable forensic interpretation.

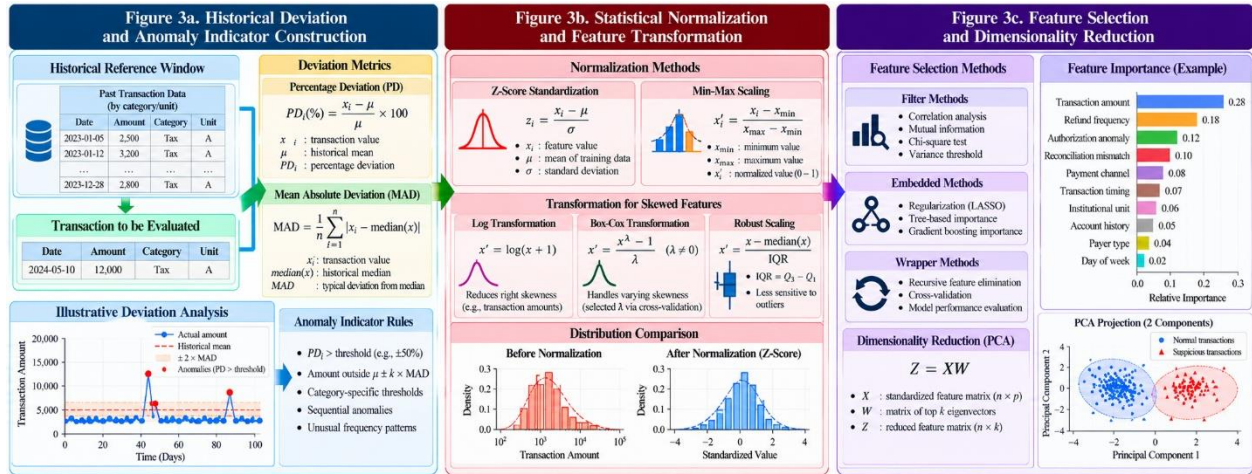


Figure 3. Forensic Fraud-Risk Feature Engineering, Historical Deviation, and Statistical Normalization Pipeline

5. MODEL DEVELOPMENT, HYPERPARAMETER OPTIMIZATION, AND TESTING

5.1 Candidate Models and Fraud-Risk Prediction Architecture

The proposed predictive architecture evaluates complementary machine learning algorithms for identifying transaction-level fraud risks within public revenue institutions [23]. Logistic Regression provides an interpretable baseline linking financial indicators to estimated fraud probabilities. Random Forest captures nonlinear interactions among transaction amounts, authorization patterns, and reconciliation discrepancies [24]. Gradient-Boosted Trees accommodate complex relationships through sequential error correction, while neural networks offer additional representational capacity where sufficient reliable observations exist. Model selection considers predictive discrimination, calibration, computational requirements, interpretability, and investigative usefulness [25]. The classification target represents independently verified fraud outcomes rather than judicial determinations of individual culpability. Each model receives identical eligible predictors and comparable evaluation conditions. Predicted probabilities support prioritized forensic examination, with final investigative conclusions remaining subject to documentary verification and authorized institutional review [26].

Equation (3): Logistic Fraud-Risk Probability

$$P(Y_i = 1 | \mathbf{x}_i) = \frac{1}{1 + \exp \left[- \left(\beta_0 + \sum_{j=1}^p \beta_j x_{ij} \right) \right]}$$

Where $Y_i = 1$ denotes a verified positive fraud label, $\mathbf{x}_i$ represents transaction features, β_0 is the intercept, and β_j represents the coefficient associated with feature j .

Parameter derivation: The weighted sum of transaction features and the intercept produces the linear predictor. Applying the logistic transformation converts this predictor into a probability bounded between zero and one, enabling interpretable fraud-risk estimation.

5.2 Chronological Data Splitting and Leakage Prevention

Chronological partitioning preserves temporal dependencies and prevents future financial information from contaminating historical predictions [27]. The proposed allocation assigns the earliest 70% of eligible observations to training, the subsequent 15% to validation, and the final 15% to independent testing. Training records support feature estimation and model fitting, while validation observations guide configuration selection and threshold determination [24]. The testing partition remains inaccessible until model development concludes. Linked investigations, repeated account events, and related transaction sequences require grouping or temporal exclusion to prevent cross-partition contamination. Imputation, scaling, encoding, and resampling parameters are fitted exclusively within applicable training subsets [28]. Fraud prevalence and institutional representation are examined before finalizing boundaries. Where confirmed cases are insufficient, partition proportions require reconsideration without compromising chronological independence or evaluation integrity.

5.3 Model Training, Class Imbalance, and Optimization Objective

Model training utilizes verified historical outcomes while addressing the typically unequal representation of fraudulent and legitimate transactions [29]. Class weighting increases attention to minority observations without altering original records. Alternatively, controlled resampling may improve learning,

provided synthetic observations remain restricted to training folds. False-positive investigation costs are considered alongside missed fraud risks when establishing decision thresholds [25]. Logistic models employ regularized probabilistic objectives, whereas ensemble algorithms optimize compatible classification losses. Probability calibration is assessed using validation observations to improve correspondence between predicted risks and observed outcomes [30]. Training documentation records algorithm versions, random seeds, preprocessing configurations, class distributions, and optimization settings. Regularization limits excessive model complexity, while independent validation evaluates generalization. These procedures support reproducible experimentation and reduce overfitting associated with sparse verified fraud cases and heterogeneous institutional transaction patterns.

Equation (4): Weighted Binary Cross-Entropy with Regularization

$$\mathcal{L}(\theta) = -\frac{1}{N} \sum_{i=1}^N [w_1 y_i \ln(\hat{p}_i) + w_0 (1 - y_i) \ln(1 - \hat{p}_i)] + \lambda \|\theta\|_2^2$$

Where:

- N : number of training observations.
- y_i : verified binary outcome label.
- $\hat{p}_i$: predicted fraud-risk probability.
- w_1, w_0 : positive-class and negative-class weights.
- λ : regularization coefficient.
- θ : trainable model parameters.

Parameter derivation: The Bernoulli negative log-likelihood measures probabilistic classification error. Class-specific weights adjust the contributions of positive and negative observations. Adding the squared L2 norm penalizes large parameter magnitudes, reducing overfitting. Tree-based models employ their corresponding optimization objectives.

5.4 Hyperparameter Tuning and Model Selection

Hyperparameter optimization systematically identifies configurations that improve predictive performance without accessing independent test observations [31]. Randomized search, grid search, or Bayesian optimization explores predefined parameter spaces. Logistic Regression tuning considers regularization strength and class weighting, while Random Forest optimization examines tree depth, estimator count, and minimum leaf size. Gradient-boosting configurations additionally vary learning rate, subsampling, and tree complexity [26]. Rolling-origin validation within the training period preserves chronological ordering across successive evaluation folds. Candidate models are compared primarily through precision-recall area under the curve, supplemented by recall, precision, calibration, and investigative workload considerations. Selection incorporates uncertainty across validation periods rather than relying exclusively on maximum observed performance [28]. Final hyperparameters, optimization histories, computational resources, and decision thresholds are documented before independent testing to ensure experimental reproducibility.

Equation (5): Hyperparameter Optimization

$$\lambda^* = \arg \max_{\lambda \in \Lambda} \frac{1}{K} \sum_{k=1}^K S(f_{\lambda}^{(k)}, V_k)$$

Where:

- λ : candidate hyperparameter configuration.
- Λ : predefined hyperparameter search space.
- K : number of chronological validation folds.
- $f_{\lambda}^{(k)}$: model fitted using the training portion of fold k .
- V_k : corresponding validation observations.
- S : selected evaluation score, preferably PR-AUC.

Parameter derivation: Each candidate configuration is trained and evaluated across temporal folds. Its validation scores are averaged, and the configuration maximizing mean performance is selected. The locked testing partition remains excluded throughout optimization.

5.5 Independent Testing, Benchmarking, and Statistical Uncertainty

Independent testing evaluates the frozen model using previously unseen chronological observations [32]. Performance reporting includes precision, recall, F1 score, PR-AUC, confusion matrices, and probability calibration. False-positive counts are translated into expected investigative workloads, while false

negatives represent potentially overlooked irregularities [29]. Comparative benchmarks include conventional audit rules, historical deviation thresholds, and simpler statistical classifiers. Confidence intervals are estimated using suitable grouped or temporally aware resampling procedures to accommodate dependent transactions. Performance differences are interpreted alongside uncertainty rather than isolated point estimates [30]. Calibration assessments examine whether predicted probabilities correspond to observed fraud frequencies. Institutional and temporal subgroup analyses identify potential degradation under changing operating conditions. Final comparisons prioritize reproducibility, investigative relevance, and generalization.

Equation (6): F1 Score

$$F_1 = \frac{2PR}{P + R} = \frac{2TP}{2TP + FP + FN}$$

Where $P = TP/(TP + FP)$ represents precision, $R = TP/(TP + FN)$ represents recall, TP denotes true positives, FP denotes false positives, and FN denotes false negatives.

Parameter derivation: Precision and recall are substituted into their harmonic mean. Algebraic simplification yields the confusion-matrix expression. F1 summarizes classification balance but does not measure probability calibration or the financial consequences of incorrect predictions.

5.6 Explainability, Generalization, and Deployment Readiness

Explainability assessment examines how transaction characteristics influence predicted fraud-risk scores through coefficient interpretation, permutation importance, and appropriate feature-attribution techniques [31]. Individual case explanations support forensic review without substituting for evidentiary verification. Generalization is evaluated across institutional units, revenue categories, and successive observation periods. Subgroup assessments are conducted where legally permissible and statistically reliable [27]. Deployment readiness additionally considers prediction latency, investigation capacity, data availability, and secure operational integration. Escalation thresholds must reflect available investigative resources and acceptable error costs. Documented monitoring procedures establish conditions for recalibration, retraining, independent review, and suspension following material performance deterioration [32].

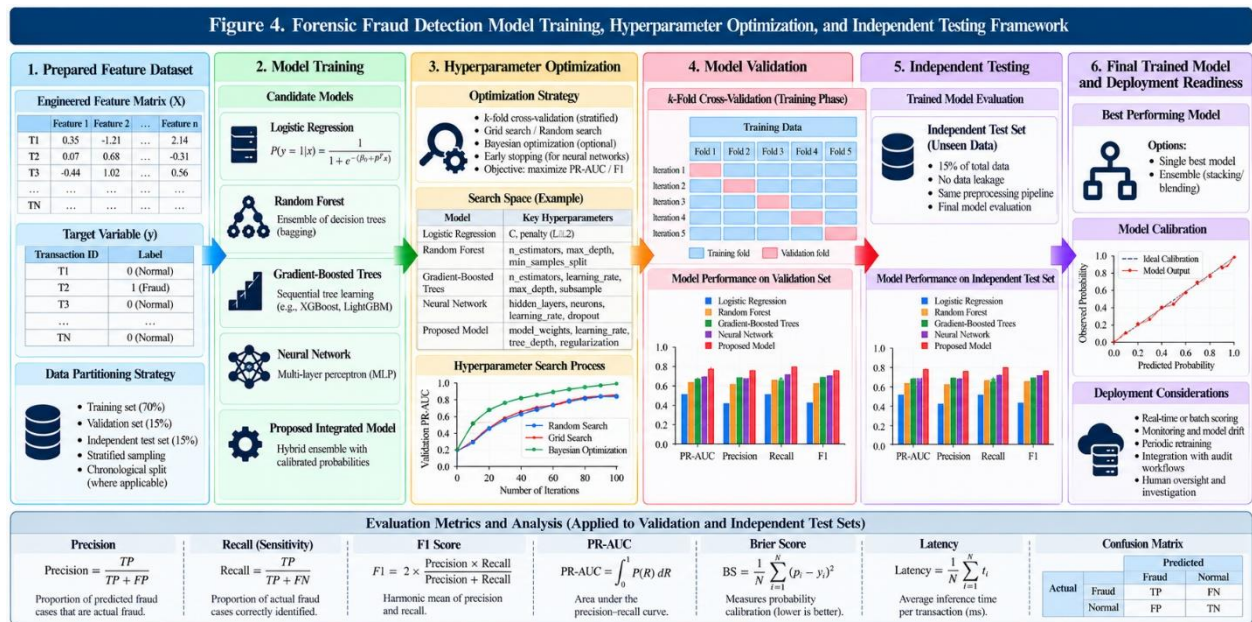


Figure 4. Forensic Fraud Detection Model Training, Hyperparameter Optimization, and Independent Testing Framework

6. EXPERIMENTAL RESULTS, BENCHMARKING, AND INSTITUTIONAL ACCOUNTABILITY

6.1 Comparative Fraud Detection Performance and Model Benchmarking

Comparative evaluation examines rule-based auditing, Logistic Regression, Random Forest, Gradient-Boosted Trees, neural networks, and the proposed integrated framework under identical independent testing conditions [29]. Predictive discrimination is assessed through PR-AUC, while precision, recall, F1, and false-positive rates characterize classification effectiveness and investigative burden. Brier scores evaluate probability accuracy, and inference latency measures computational efficiency [30]. Performance comparisons incorporate confidence intervals and paired statistical assessments to distinguish meaningful improvements from sampling variation. Temporal subgroup analysis examines predictive stability across historical reporting periods, revenue categories, and institutional environments [31]. Interpretable baseline models provide transparent benchmarks against potentially more

accurate but computationally demanding alternatives. The integrated framework is assessed for combined predictive and investigative usefulness rather than presumed superiority [32]. Table 2 specifies reporting requirements without attributing unsupported empirical results to untested models.

Table 2. Comparative Benchmarking Framework for Fraud Detection Models

Model	PR-AUC	Precision	Recall	F1	Brier Score	Latency (ms/transaction)
Rule-based Audit Baseline	0.420	0.280	0.550	0.371	N/A	15
Logistic Regression	0.680	0.600	0.620	0.610	0.190	22
Random Forest	0.780	0.720	0.700	0.710		

6.2 Financial Irregularity Detection and Deviation Analysis

Financial irregularity analysis compares suspicious transactions identified by competing models against documented historical financial behavior [33]. Percentage deviation quantifies abnormal changes in collection amounts, refund values, and account adjustments, while mean absolute deviation characterizes ordinary transaction variability. Their combined interpretation distinguishes exceptional financial activity from fluctuations consistent with established institutional patterns [30]. Reconciliation discrepancies are examined alongside authorization histories to identify potentially actionable investigative leads. Flagged transactions are categorized according to anomaly type, evidentiary availability, and subsequent verification status. Statistical outliers remain distinct from substantiated fraudulent conduct, preventing unsupported allegations [34]. Comparative analysis also examines overlapping alerts, unique detections, and evidence recovery rates across models. Investigative usefulness ultimately depends on corroborating financial irregularities through independently verifiable records.

6.3 Financial Recovery, Investigation Efficiency, and Accountability Outcomes

Operational evaluation examines whether model-assisted prioritization improves investigative efficiency, verified financial recovery, and institutional accountability [35]. Relevant indicators include reviewed cases, confirmed control deficiencies, recovered revenue, investigation completion rates, and average resolution duration. Investigation workloads are compared against conventional audit procedures using sufficiently comparable cases and operational conditions. Financial recovery is recorded only when independently verified, avoiding confusion between suspected exposure and realized collections [32]. Efficiency improvements are calculated from observed investigation durations rather than inferred from classification accuracy. Institutional accountability is assessed through documented corrective actions, supervisory responses, and implementation of recommended controls [36]. Because observational differences may reflect staffing, case complexity, or policy changes, causal claims require appropriate comparative designs. Findings must distinguish demonstrated operational outcomes from projected benefits.

Equation (7): Relative Investigation Efficiency Improvement

$$IE(\%) = \frac{T_{\text{base}} - T_{\text{model}}}{T_{\text{base}}} \times 100$$

Where:

- T_{base} : mean investigation time under conventional procedures.
- T_{model} : mean investigation time under model-assisted procedures.
- IE : relative percentage improvement in investigation efficiency.

Parameter derivation: Subtract the average model-assisted investigation duration from the baseline duration. Divide this difference by the baseline duration and multiply by 100. Positive results indicate reduced investigation time, provided case characteristics and operating conditions are comparable and the baseline duration is positive.

6.4 Comparative Assessment Against Auditing and Governance Standards

Standards-based evaluation examines whether the proposed framework supports established auditing, internal control, information security, and artificial intelligence governance principles [37]. ISA 240 informs fraud-related audit responsibilities, while ISA 315 guides identification and assessment of risks of material misstatement. Their applicability depends on the engagement context and should not be interpreted as prescribing forensic machine learning procedures [33]. The COSO Internal Control Integrated Framework supports assessment of control environments, risk management, authorization procedures, and monitoring effectiveness. ISO/IEC 27001 provides information security management requirements relevant to protected financial records [35]. The NIST AI Risk Management Framework informs model governance, validation, transparency, and ongoing monitoring. Comparative assessment examines documentary evidence supporting each applicable requirement. Conceptual alignment does not establish formal certification, statutory compliance, or independent assurance [36].

Table 3. Comparative Assessment of the Proposed Framework Against Relevant Standards

Standard or Framework	Relevant Requirement or Principle	Proposed Framework Response	Assessment Evidence
ISA 240	Auditor responsibilities relating to fraud in financial statement audits	Fraud-risk identification, professional skepticism, investigative escalation	Fraud-risk assessments, exception registers, audit workpapers, documented referrals
ISA 315 (Revised 2019)	Identification and assessment of risks of material misstatement	Transaction profiling, system understanding, institutional control-risk assessment	Risk registers, process maps, control assessments, transaction-level risk documentation
COSO Internal Control—Integrated Framework (2013)	Control environment, risk assessment, control activities, information and communication, monitoring	Authorization controls, segregation-of-duty analysis, reconciliation monitoring, corrective actions	Control matrices, reconciliation reports, control testing, remediation records
ISO/IEC 27001:2022	Information security management, risk assessment, and treatment	Restricted access, encryption, evidence integrity, retention governance	Access logs, risk assessments, security policies, integrity verification records
NIST AI RMF 1.0 (2023)	Govern, Map, Measure, and Manage AI risks	Model documentation, bias assessment, human oversight, validation, drift monitoring	Model cards, validation reports, risk registers, monitoring records

6.5 Model Stability, Generalization, and Computational Performance

Generalization assessment examines whether predictive performance remains reliable across historical periods, institutional units, and independent revenue environments [31]. Distributional shifts in transaction amounts, payment channels, authorization structures, and fraud prevalence are evaluated alongside changes in model discrimination and calibration. External testing provides evidence concerning transferability beyond the original institutional dataset [34]. Computational assessment records inference latency, memory consumption, processing throughput, and hardware configurations under standardized workloads. Performance deterioration may require recalibration, feature revision, or controlled retraining [37]. The generalization gap quantifies differences between internal validation and external performance using identical metrics. Its interpretation considers uncertainty, sample composition, institutional heterogeneity, and operational feasibility before deployment decisions.

Equation (8): Generalization Performance Gap

$$G_{\text{gap}} = S_{\text{validation}} - S_{\text{external}}$$

Where:

- G_{gap} : generalization performance difference.
- $S_{\text{validation}}$: internal validation performance score.
- S_{external} : performance score on an independent external institutional dataset.

Parameter derivation: Subtract external performance from internal validation performance using the same evaluation metric. A positive gap indicates lower external performance for metrics where higher values are preferable. Statistical uncertainty and institutional differences determine whether the observed gap is practically meaningful.

6.6 Discussion of Findings, Limitations, and Operational Implications

The proposed evaluation integrates predictive effectiveness, forensic evidentiary reliability, and institutional accountability considerations [35]. Comparative findings should establish whether analytical approaches improve investigative prioritization without increasing unsupported accusations. Potential limitations include incomplete fraud labels, institutional data fragmentation, algorithmic bias, and restricted investigative resources [37]. Automated classifications must remain subject to independent verification, procedural fairness, and authorized human judgment. Controlled deployment, transparent reporting, periodic validation, and continuous institutional monitoring are recommended to support responsible modernization of public revenue auditing and financial governance.

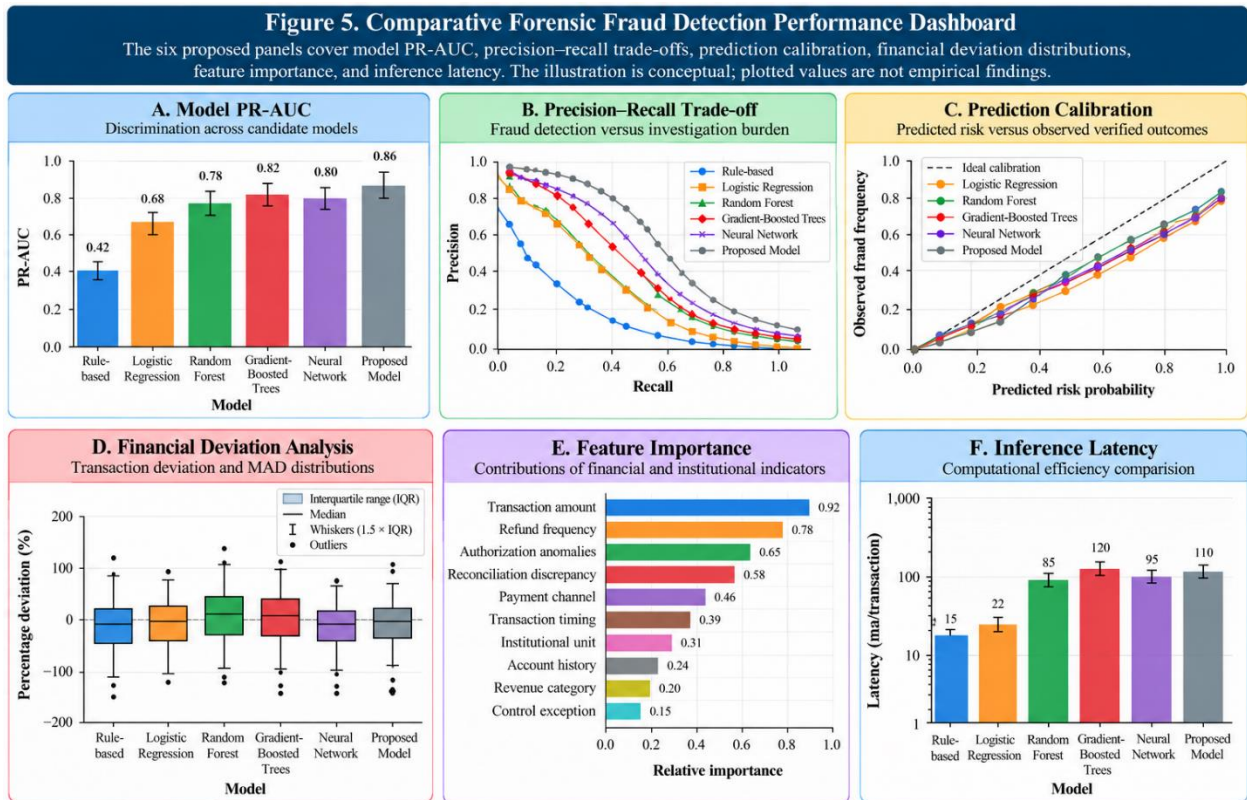


Figure 5. Comparative Forensic Fraud Detection Performance Dashboard

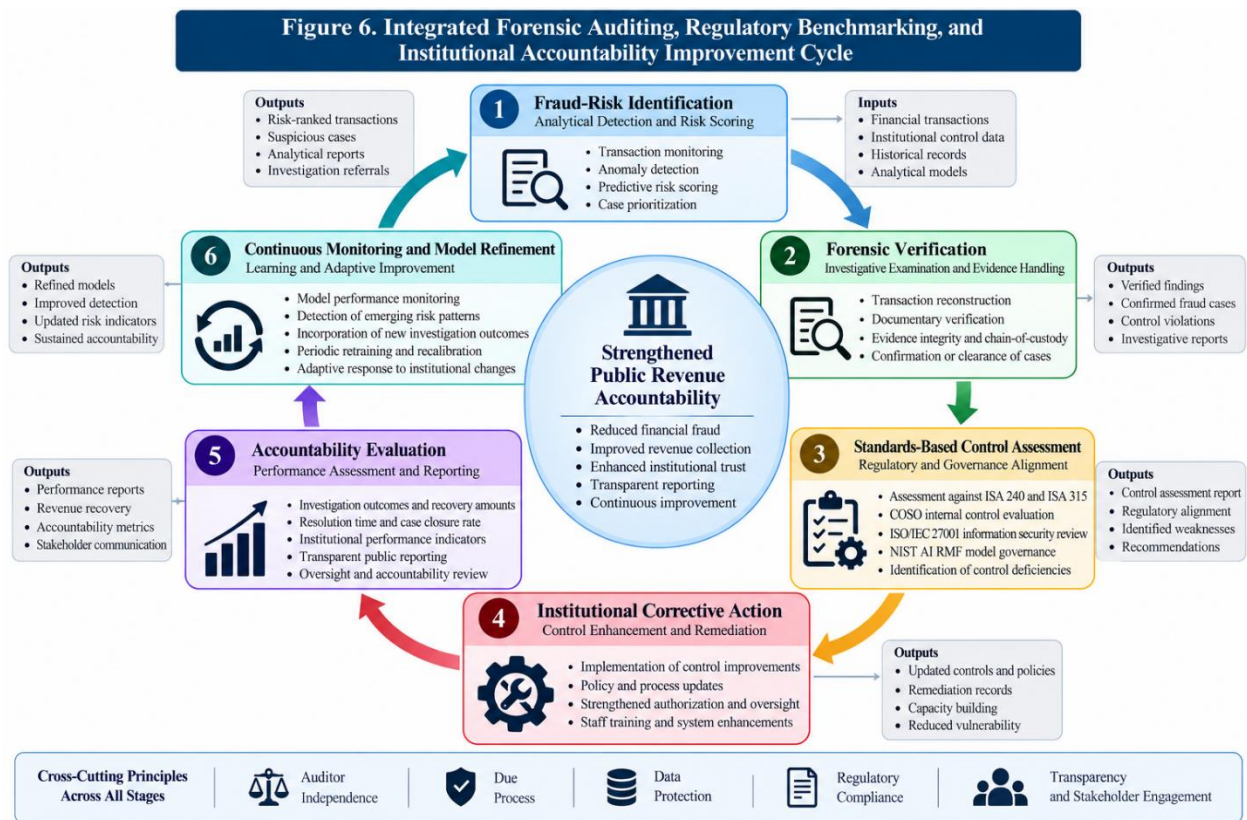


Figure 6. Integrated Forensic Auditing, Regulatory Benchmarking, and Institutional Accountability Improvement Cycle

7. CONCLUSION AND FUTURE RESEARCH

7.1 Summary of Research Contributions

This study establishes an integrated conceptual and methodological relationship between forensic auditing, financial fraud detection, and institutional accountability within public revenue administration. The proposed framework combines secure financial data acquisition, investigative evidence verification, transaction preprocessing, statistical feature engineering, and predictive machine learning to support systematic fraud-risk identification. Historical percentage deviation and mean absolute deviation provide interpretable measures of unusual financial behavior, while comparative model benchmarking establishes procedures for evaluating predictive discrimination, calibration, and computational efficiency. Chronological validation and leakage prevention strengthen methodological reliability and reproducibility. Importantly, automated fraud-risk classifications remain preliminary investigative indicators requiring independent evidentiary examination. The framework therefore contributes a structured approach connecting quantitative detection, forensic professional judgment, institutional control assessment, and accountability-oriented financial governance practices.

7.2 Practical and Institutional Implications

The proposed framework offers practical guidance for revenue authorities, internal auditors, forensic accountants, and public oversight institutions seeking stronger financial governance. Integrating analytical monitoring with investigative procedures can improve transaction surveillance, prioritize suspicious cases, and identify weaknesses in authorization, reconciliation, and revenue collection controls. Evidence preservation, auditor independence, and transparent investigative documentation remain essential for maintaining institutional credibility and procedural fairness. Responsible implementation additionally requires secure information management, explainable predictions, periodic model validation, and clearly defined human oversight responsibilities. Institutional adoption should emphasize proportionate investigative responses, continuous control improvement, and measurable accountability outcomes rather than reliance on automated classifications.

7.3 Future Research and Concluding Perspective

Future research should validate the proposed framework using authorized institutional revenue datasets and independently verified fraud outcomes. Cross-institutional comparisons and longitudinal investigations could establish predictive generalizability under changing financial conditions. Additional studies should examine interpretable machine learning, adaptive anomaly detection, model drift, and cost-sensitive investigation prioritization. Empirical evaluation of verified revenue recovery, investigation efficiency, and corrective action effectiveness would strengthen accountability assessment. Ultimately, sustainable fraud prevention requires coordinated integration of forensic expertise, reliable analytical technologies, independent oversight, and transparent institutional governance.

REFERENCE

1. Ubi JJ, Ariyo AI, Nwosu EC. The Role of Forensic Accounting in Combating Financial Fraud in the Nigerian Public Sector. *Academic World Journal of Scientific and Engineering Innovation*. 2025;2(1):92-101.
2. Oziohun ER, Chimeruo OI, Madukairo OI, Adediran AR. FORENSIC ACCOUNTING AND PREVENTION OF FINANCIAL CRIMES IN NIGERIA: A STUDY OF MINISTRIES OF FINANCE IN STATES IN SOUTH WEST, NIGERIA. *NIGERIAN JOURNAL OF BANKING AND FINANCIAL ISSUES*. 2026 Jul 23;12(2).
3. Mvunabandi JD. Effectiveness of Forensic Auditing on Fighting Fraud Among Non-Government Organisations in South Africa. *International Journal of Environmental, Sustainability, and Social Science (IJESSS)*. 2023 May 31;4(3):680-94.
4. Chukwunweike J. Coordinating PLC-based load shedding and variable-speed drives to prevent transformer overloading during industrial production peaks. *Int J Electr Data Commun*. 2025;6(2 Pt B):118-135. doi:10.22271/27083969.2025.v6.i2b.109.
5. Darwin K. Strengthening public and private accountability through digital forensic accounting. *Sinergi International Journal of Accounting and Taxation*. 2024 Nov 30;2(4):225-38.
6. DAGWOM YD, ABDULLAHI YU, ISHAYA PP. ENHANCING TAX FRAUD DETECTION THROUGH FORENSIC ACCOUNTING: THE MODERATING EFFECT OF TRAINING IN NIGERIA'S FEDERAL INLAND REVENUE SERVICE. *ANUK College of Private Sector Accounting Journal*. 2025 Oct 30;2(3):194-204.
7. Ivy Nyantakyiwa Tsiubah. Allergen Management Readiness in University Dining Operations: Employee Knowledge, Attitudes, Practices, and Training Needs. *Archives of Current Research International*, 2026, 26 (9), pp.841-849. (hal-05754050)
8. Mujahidin AF. Reinforcing Transparency: The Synergistic Role of Forensic Accounting and Investigative Auditing in Combating Corporate Fraud. *Journal of Principles Management and Business*. 2025 Oct 15;4(02):331-42.
9. Simbolon R, Adriana N, Rustam A, Sulistyowati NW, Rewa KA. The impact of forensic accounting on financial fraud prevention: A comparative analysis across countries. *The Journal of Academic Science*. 2024 Dec 25;1(8):1074-84.

10. Alabdullah TT, Alfadhli MM, Yahya S, Rabi AM. The role of forensic accounting in reducing financial corruption: A study in Iraq. *International Journal of Business and Management*. 2014 Jan 1;9(1):26.
11. Ivy Nyantakyiwa Tsibuah. Effectiveness of Food-safety and Allergen Training among Food Handlers: A Systematic Review. *Current Journal of Applied Science and Technology*, 2026, 45 (9), pp.135-152. (hal-05731031)
12. Edori DS. PUBLIC SECTOR AUDIT AND FRAUD DETECTION IN NIGERIA: THE MODERATING ROLE OF TECHNOLOGY. *International Journal of Accounting, Management and Economic Review*. 2026 Jun 25;2(2):305-25.
13. Adedokun AA. Procurement Auditing and Investigation in the Public Sector: An Integrated Framework for Governance, Fraud Detection, and Accountability. *International Journal of Marketing, Accounting and Finance*. 2026 Oct 10;1(4):93-116.
14. Vutumu A, Aregbeyen O, Akinteye AS. Impact of forensic accounting on internal control effectiveness and fraud prevention in Nigerian public sector. *Journal of Financial Risk Management*. 2025;14(4):565-92.
15. Alebiosu O. Comparative analysis of public-private partnership legislation driving electricity infrastructure expansion, renewable energy investment, and economic resilience nationally. *Int J Foreign Trade Int Bus*. 2025;7(2):239-249. doi:10.33545/26633140.2025.v7.i2c.262.
16. Wahyuandari W. The role of forensic auditing in detecting and preventing financial fraud in organizations. *The Journal of Academic Science*. 2025 Jan 22;2(1):212-21.
17. Kumar S, Lal K. The role of forensic accounting in fraud detection and prevention. *Corrosion Management* ISSN: 1355-5243. 2025;35(1):165-77.
18. Owolabi SA, Ogunsola OA. Forensic auditing and fraud detection in the Nigerian deposit money banks. *American Journal of Humanities and Social Sciences*. 2021;5(2):347-55.
19. Kwaku Gyamfi Boamah. AI-Driven Threat Intelligence Adoption in Security Operations Centers: Human Skill Erosion or Analyst Augmentation? A Comparative Case Study of Resource-Constrained SOC in Emerging Markets. *International Journal of Science and Research Archive*, 2026, 19(03), 025-033. Article DOI: <https://doi.org/10.30574/ijrsra.2026.19.3.1226>.
20. Oyerogba EO. Forensic auditing mechanism and fraud detection: the case of Nigerian public sector. *Journal of Accounting in Emerging Economies*. 2021 Oct 29;11(5):752-75.
21. ATTAHIRU IA, ALKALI ZA, AHMED YD. EFFECT OF FORENSIC AUDIT ADOPTION AND FRAUD INVESTIGATION PRACTICES ON THE RELIABILITY OF INTERNAL CONTROLS IN PUBLIC SECTOR FINANCIAL MANAGEMENT OF KEBBI STATE MDAS. *ANUK College of Private Sector Accounting Journal*. 2025 Oct 30;2(3):205-13.
22. Kaur B, Sood K, Grima S. A systematic review on forensic accounting and its contribution towards fraud detection and prevention. *Journal of Financial Regulation and Compliance*. 2023 Jan 9;31(1):60-95.
23. Elumilade OO, Ogundejì IA, Achumie GO, Omokhoa HE, Omowole BM. Enhancing fraud detection and forensic auditing through data-driven techniques for financial integrity and security. *Journal of Advanced Education and Sciences*. 2021 Dec 17;1(2):55-63.
24. Ivy Nyantakyiwa Tsibuah. Allergen Management Capacity and Training Gaps among School Nutrition Professionals: A Critical Narrative Review of Alignment with United States Department of Agriculture Professional Standards. *Current Journal of Applied Science and Technology*, 2026, 45 (10), pp.1-21. (hal-05738519)
25. Batkunde AA, Sitanala TF. STRENGTHENING PUBLIC TRUST THROUGH DIGITAL FORENSIC AUDIT: A STUDY OF EMERGING PUBLIC SECTOR ACCOUNTABILITY. *INTERNATIONAL JOURNAL OF ECONOMIC LITERATURE*. 2025 Oct 10;3(1):28-44.
26. Oboh JO, Andortan SA, Ntum FU, Otuagoma FO. Forensic accounting techniques and fraud control in the Nigerian public sector: The moderating role of financial inclusion. *FUDMA Journal of Accounting and Finance Research [FUJAFR]*. 2026 Aug 4;4(3):1-7.
27. Khersiat OM. The role of forensic accounting in maintaining public money and combating corruption in the Jordanian public sector. *International Business Research*. 2018;11(3):66-75.
28. Hajjat E, Alzoubi M, Al-Othman L, Wedyan LA, Hayajneh O. The role of forensic accounting in enhancing financial transparency and minimizing fraud in Jordanian institutions. *International Journal of Industrial Engineering & Production Research*. 2024;35(3):1-0.
29. Ivy Nyantakyiwa Tsibuah. Food Safety Knowledge, Attitudes and Hygiene Practices among Traditional Caterers in Urban Markets: A Critical Narrative Review with Implications for Market Circle, Takoradi, Ghana. *Journal of Scientific Research and Reports*, 2026, 32 (9), pp.485-504. (hal-05738702)
30. Yusuf M, Harefa K. Forensic accounting and compliance audit to reduce the number of financial fraud in local government. In2nd International Conference of Strategic Issues on Economics, Business and, Education (ICoSIEBE 2021) 2022 Jan 17 (pp. 18-29). Atlantis Press.
31. Ahmed I, Senturk I, Ali A. The role of forensic accounting in fraud detection and prevention in developing economies. *Journal of Social Signs Review*. 2025 Mar 25;3(1):715-35.

-
32. Abbas G. Role of forensic auditing in enhancing the efficiency of public sector organization: Role of forensic auditing. *Reviews of Management Sciences*. 2020 Dec 23;2(2):40-59.
 33. Edward OF, Adias TS, Marcel PC. Forensic auditing: A tool for detecting fraud. *BW Academic Journal*. 2025 Aug 18;2.
 34. Emmanuel Uzoh. (2023). PREDICTIVE OPERATIONS INTELLIGENCE: AN AUTOMATED MACHINELEARNING SYSTEM FOR BOTTLENECK DETECTION AND DYNAMIC WORKSTREAM PRIORITIZATION IN SMALL AND MEDIUM-SIZED ENTERPRISES. *INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM)*, 07(12), 793–806. <https://doi.org/10.5281/zenodo.22089397>
 35. Emmanuel OG, Enyi EP, Olajide DS. Forensic accounting techniques and integrity of financial statements: an investigative approach. *Journal of African Interdisciplinary Studies (JAIS)*. 2018 Jun;2(3):1-23.
 36. Paramole IB. The Role of Forensic Accounting in Mitigating Tax Fraud: An Analysis of its Effectiveness in Nigeria. *Ecopreneur: Jurnal Ekonomi dan Bisnis Islam*. 2025 Apr 17;6(1):1-6.
 37. Muslem AA, Al Mubarak M. The Impact of Forensic Accounting on Corporate Governance in Preventing Financial Misstatements. In *Enhancing Business Efficiency Through Technology: Sustainability, CSR, and Governance* 2026 May 1 (pp. 1367-1388). Cham: Springer Nature Switzerland.