INTERNATIONAL JOURNAL OF ECONOMIC SOCIAL SCIENCE AND MANAGEMENT LAW
Peer Reviewed, Referred & Indexed Journal
E-ISSN: 3070-0558
VOL.4, NO. 3 (2023)

ADAPTIVE ML FRAMEWORKS FOR ANOMALY DETECTION IN HEALTHCARE TRANSACTIONS: A NATIONAL SECURITY PERSPECTIVE

Shila Das[1], Yousuf Md Shahan[2], Joynob Sultana[3], Sayem Sarwar[4], Fahim Abrar[5], Majharul Islam Shanto[6]

[1]MBA in Management Information Systems, International American University
[2]MBA in Business Analytics, Troy University
[3]MBA in Business Analytics, Troy University
[4]MS in Computer Science — Artificial Intelligence, Troy University
[5]MS in Business Analytics, Trine University
[6]MS in Computer Science — Data Science, Troy University

Abstract

Healthcare fraud, waste, and abuse (FWA) represent a significant threat to national health security and economic stability, draining billions from the public sector and compromising the integrity of essential services. As healthcare systems increasingly transition to digital transaction environments, they become vulnerable to sophisticated adversarial actors. A critical gap exists in current defense mechanisms: static machine learning models lack the agility to counter rapidly evolving fraudulent tactics and concept drift.

This paper proposes a novel adaptive machine learning framework for anomaly detection in healthcare transactions, specifically designed for high-stakes security environments. Unlike traditional approaches, the proposed framework utilizes continuous learning paradigms to autonomously update its parameters in response to emerging adversarial patterns. Experimental evaluations demonstrate that the framework achieves an anomaly detection rate exceeding 94%, significantly outperforming baseline models while simultaneously reducing false positive rates-a crucial factor for maintaining operational efficiency in critical infrastructure.

From a national security perspective, the implementation of such adaptive systems is vital for protecting medical supply chains and ensuring the resilience of pandemic response mechanisms. By securing the financial and data integrity of healthcare systems, this research provides a robust defense against domestic and transnational economic threats. The findings suggest that transitioning to adaptive, continuous learning architecture is essential for safeguarding the national public health framework against the destabilizing effects of large-scale systemic abuse.

Empirical evaluation across heterogeneous healthcare transaction datasets demonstrates a detection rate exceeding 94%, alongside statistically significant reductions in false positives relative to established baseline models. These gains reflect the framework's capacity to maintain precision under concept drift - a critical operational requirement for real-world deployment. From a national security standpoint, the framework's applicability extends beyond billing fraud to safeguarding the integrity of medical supply chains, reinforcing critical healthcare infrastructure resilience, and strengthening institutional responsiveness during large-scale public health crises.

Keywords: *Adaptive machine learning; anomaly detection; healthcare transactions; national security; continuous learning; fraud detection.*

1. INTRODUCTION

Healthcare systems process billions of financial transactions annually, including insurance claims, provider reimbursements, pharmaceutical payments, and medical supply chain procurements. In the

United States alone, healthcare expenditures exceed 18% of GDP, with public programs such as Medicare and Medicaid handling millions of claims per day (CMS, 2023). The scale and complexity of these transactions create significant vulnerabilities to fraud, waste, and abuse (FWA), resulting in estimated annual losses in the tens of billions of dollars (GAO, 2022). Beyond monetary loss, compromised transaction systems undermine the operational stability of hospitals, insurers, and national emergency response infrastructures.

Healthcare financial integrity is increasingly recognized as a national security concern. During public health crises such as the COVID-19 pandemic, rapid procurement of medical supplies, emergency reimbursements, and vaccine distribution funding exposed systemic weaknesses in oversight mechanisms (OECD, 2021). Disruptions in healthcare financing can impair biodefense readiness, delay critical supply chain operations, and weaken crisis response capabilities. Moreover, adversarial actors ranging from organized crime networks to state-sponsored entities may exploit healthcare payment systems to destabilize national infrastructure or divert strategic resources (FBI, 2022). Protecting healthcare transactions is therefore not solely an economic issue but a matter of safeguarding critical national infrastructure.

Traditional fraud detection approaches rely heavily on rule-based systems and static machine learning (ML) models. Rule-based systems, while interpretable, require manual updates and struggle to adapt to evolving fraud schemes. Similarly, static ML models trained on historical datasets suffer from concept drift, where changing transaction patterns degrade predictive performance over time (Gama et al., 2014). Adversaries increasingly employ adaptive tactics, synthetic identities, and coordinated billing schemes designed to evade fixed detection thresholds. As a result, detection accuracy declines, and false positive rates increase, straining administrative resources and potentially delaying legitimate reimbursements.

This study addresses the following research question: *How can adaptive machine learning frameworks improve anomaly detection in healthcare transactions while satisfying national security constraints such as resilience, scalability, and real-time responsiveness?* To answer this question, we propose an adaptive ML framework that integrates continuous learning, drift detection mechanisms, and dynamic threshold optimization to maintain robust performance in adversarial and high-volume environments.

The contributions of this paper are twofold. Theoretically, we conceptualize healthcare transaction monitoring as a component of national critical infrastructure protection, integrating cybersecurity and financial anomaly detection perspectives. Empirically, we design and evaluate an adaptive anomaly detection architecture that demonstrates improved detection rates and reduced false positives compared to conventional baseline models.

The remainder of this paper is structured as follows: Section 2 reviews related literature on healthcare fraud detection and adaptive ML methods. Section 3 presents the proposed framework and system architecture. Section 4 describes the experimental methodology and results. Section 5 discusses national security implications, and Section 6 concludes with policy and research recommendations.

2. BACKGROUND AND RELATED WORK

2.1 Anomaly Detection in Healthcare

Healthcare transaction anomalies manifest across multiple dimensions, including fraudulent insurance claims, organized billing rings, and device upcoding-where providers bill for more expensive equipment or procedures than those delivered (Hall et al., 2020). The Centers for Medicare and Medicaid Services (CMS, 2023) identifies these anomalies as primary drivers of FWA, collectively costing the U.S. healthcare system over $100 billion annually. Additionally, phantom billing-charging for services never rendered-and unbundling schemes, where bundled procedures are billed separately to inflate

reimbursements, represent sophisticated fraud typologies that challenge conventional detection systems (Bauder et al., 2017).

2.2 Machine Learning for Healthcare Fraud

Supervised learning methods, including Random Forests (RF) and XGBoost, have demonstrated robust performance in classifying fraudulent transactions when labeled training data is available (Herland et al., 2018). However, these approaches are constrained by data imbalance and dependence on historical fraud patterns. Unsupervised techniques such as autoencoders and Isolation Forests address label scarcity by detecting statistical deviations without prior fraud definitions, proving effective in identifying novel anomalies in high-dimensional transaction datasets (Liu et al., 2012). Semi-supervised approaches, which leverage small, labeled datasets alongside large volumes of unlabeled data, offer a practical compromise, achieving competitive detection performance in real-world healthcare environments (Chalapathy & Chawla, 2019). Despite these advances, all three paradigms share a critical limitation: their static nature makes them vulnerable to distributional shifts in transaction data.

2.3 Adaptive ML and Concept Drift

Concept drift-the phenomenon where the statistical properties of target variables change over time-poses a fundamental challenge to static ML models deployed in dynamic fraud environments (Gama et al., 2014). Online learning frameworks address this by updating model parameters incrementally as new data arrives, without full retraining. Thisinual learning extends this paradigm by preserving previously acquired knowledge while integrating new patterns, mitigating the problem of catastrophic forgetting (Parisi et al., 2019). Drift detection algorithms such as ADWIN (Adaptive Windowing) and DDM (Drift Detection Method) provide formal mechanisms for identifying statistically significant distributional changes in data streams, enabling timely model recalibration (Bifet & Gavalda, 2007). These methods are particularly relevant to healthcare fraud detection, where adversarial actors deliberately alter transaction patterns to evade existing models.

2.4 National Security Perspective

The Cybersecurity and Infrastructure Security Agency (CISA, 2021) formally designate healthcare as a critical infrastructure sector, acknowledging that disruptions to healthcare financial systems can cascade into broader national security failures. Adversarial threats extend beyond conventional fraud to include deliberate poisoning attacks on ML training pipelines, where malicious actors inject manipulated data to degrade model performance (Biggio & Roli, 2018). The diversion of controlled substances through falsified procurement transactions represents another vector through which healthcare system exploitation translates into public safety risks (DEA, 2020). Furthermore, the COVID-19 pandemic revealed the extent to which healthcare payment systems could be targeted by large-scale frauds, including fraudulent Personal Protective Equipment (PPE) procurement and fake vaccination billing schemes, costing the federal government billions (FBI, 2022).

2.5 Gap Analysis

Despite considerable progress in both adaptive ML and healthcare fraud detection, a critical gap remains. Existing frameworks typically optimize for either detection accuracy or computational efficiency, but not simultaneously within a nationally secured operational context. No current framework jointly addresses adaptivity to concept drift, compliance with healthcare regulatory standards such as HIPAA, and the resilience requirements demanded by national security constraints (Bauder et al., 2017; CISA, 2021). Furthermore, the intersection of adversarial robustness and continuous learning in healthcare transaction monitoring remains unexplored in the literature (Parisi et al., 2019; Biggio & Roli, 2018). This paper

directly addresses this gap by proposing a unified adaptive framework designed to meet detection, compliance, and security objectives concurrently.

3. PROBLEM FORMULATION AND THREAT MODEL

3.1 System Model

The healthcare transaction ecosystem comprises multiple interconnected actors operating within a complex financial and regulatory infrastructure. Primary actors include providers (hospitals, physicians, pharmacies), who initiate claims and procurement requests; payers (insurers, Medicare, Medicaid), who process reimbursements; clearinghouses, which standardize and route transaction data between providers and payers; and federal monitors such as the Centers for Medicare and Medicaid Services (CMS) and the Department of Homeland Security (DHS), which oversee financial integrity and national security compliance (CMS, 2023; DHS, 2021).

Transaction types within this ecosystem fall into three primary categories. Claims transactions represent requests for reimbursement submitted by providers following the delivery of medical services, constituting the highest volume and most fraud-prone category (GAO, 2022). Prior authorization (prior auth) transactions involve pre-approval requests submitted to payers before specific services or medications are delivered, creating opportunities for fraudulent approvals of unnecessary procedures (Bauder et al., 2017). Remittance transactions document payment confirmations and explanations of benefits (EOBs) issued by payers to providers, and anomalies in this category often indicate duplicate payments, overbilling, or money laundering activities (Hall et al., 2020).

Formally, the transaction system can be represented as a data stream $\mathcal{S} = \{x_1, x_2, \ldots, x_t\}$, where each transaction $x_t \in \mathbb{R}^d$ is a d-dimensional feature vector encompassing billing codes, provider identifiers, timestamps, reimbursement amounts, and geographic metadata. The distribution $P(x_t)$ is assumed to be non-stationary, reflecting the evolving nature of both legitimate healthcare operations and adversarial fraud strategies (Gama et al., 2014). Federal monitors interact with this stream through audit mechanisms designed to flag statistical deviations from expected billing patterns, though existing systems remain limited in their real-time adaptability (CISA, 2021)

3.2 Anomaly Types

Healthcare transaction anomalies span financial, operational, security-related, and supply chain categories, each carrying distinct implications for national security. Table 1 presents a structured taxonomy of these anomaly types with corresponding national security impact levels.

Table 1: Taxonomy of Healthcare Transaction Anomalies with National Security Impact Levels

Anomaly Type	Category	Description	Example	National Security Impact
Fraudulent Claims	Financial	Billing for services not rendered or medically unnecessary	Phantom billing, upcoding	Medium
Billing Rings	Financial	Coordinated fraud networks submitting linked false claims	Organized provider collusion	High

Duplicate Payments	Financial	Same claim reimbursed multiple times	Split billing schemes	Low
Upcoding/Unbundling	Operational	Inflating procedure codes to maximize reimbursement	Device upcoding, procedure splitting	Medium
Prior Auth Manipulation	Operational	Fraudulent pre-approvals for unnecessary treatments	Opioid over-prescription approvals	High
Controlled Substance Diversion	Supply Chain	Falsified procurement of regulated pharmaceuticals	Opioid diversion via fraudulent orders	High
PPE/Medical Supply Fraud	Supply Chain	False invoicing for critical medical equipment	Fake PPE procurement during crises	High
Adversarial Data Poisoning	Security	Deliberate injection of manipulated data to degrade ML models	Poisoned training datasets	High
Slow-Drift Evasion	Security	Gradual pattern shifts designed to evade static detection	Incremental billing anomalies	Medium
Identity Spoofing	Security	Use of synthetic or stolen provider/patient identities	Synthetic identity fraud	High

As illustrated in Table 1, anomalies classified as High national security impact-including billing rings, controlled substance diversion, and adversarial poisoning-directly threaten the integrity of critical healthcare infrastructure and national emergency response capabilities (FBI, 2022; DEA, 2020). Supply chain diversions represent a hybrid financial-security threat, as falsified procurement of pharmaceuticals or medical equipment can impair biodefense readiness during public health emergencies (OECD, 2021). Anomalies rated medium impact, such as upcoding and slow-drift evasion, may appear operationally benign in isolation but can aggregate into systemic vulnerabilities when left undetected over time (Bauder et al., 2017).

3.3 Threat Model

The threat model adopted in this study assumes adversaries with varying degrees of sophistication operating within and across the healthcare transaction ecosystem. Consistent with established adversarial ML frameworks (Biggio & Roli, 2018), we define three primary adversarial capability categories.

Data Poisoning Attacks involve the deliberate injection of manipulated transactions into training datasets, with the objective of corrupting the anomaly detection model's decision boundaries. Poisoning attacks may target the model during initial training or through incremental updates in online learning pipelines, making continuous learning systems particularly vulnerable without appropriate safeguards (Steinhardt et al., 2017). In the healthcare context, adversaries with insider access-such as compromised clearinghouse personnel-may introduce subtly falsified transaction records to degrade model sensitivity over time.

Slow-Drift Attacks represent a more insidious threat, wherein adversaries gradually modify fraudulent transaction patterns at a rate deliberately designed to fall below conventional drift detection thresholds (Gama et al., 2014). Unlike abrupt fraud scheme changes, slow-drift strategies exploit the temporal blind spots of static and semi-adaptive models, incrementally shifting billing patterns across weeks or months until detection becomes statistically infeasible without specialized drift monitoring mechanisms such as ADWIN (Bifet & Gavalda, 2007).

Evasion Attacks occur at inference time, where adversaries craft transactions specifically engineered to mimic legitimate billing patterns while concealing fraudulent intent. These attacks leverage knowledge of the detection model's feature space, exploiting gaps in training data coverage to generate adversarial examples that bypass anomaly thresholds (Biggio & Roli, 2018). In supply chain contexts, evasion attacks may involve the falsification of procurement metadata-such as vendor identifiers and delivery timestamps-to render fraudulent orders indistinguishable from legitimate transactions (DEA, 2020).

Core Assumptions: The threat model operates under two fundamental assumptions. First, partial label availability is assumed, reflecting real-world conditions where only a fraction of fraudulent transactions is confirmed through audits or legal proceedings, limiting the volume of labeled training data (Chalapathy & Chawla, 2019). Second, the transaction data stream is characterized by a non-stationary distribution, driven by seasonal billing patterns, policy changes, and deliberate adversarial manipulation, necessitating continuous model adaptation to maintain detection efficacy (Parisi et al., 2019). These assumptions collectively motivate the adaptive ML framework proposed in the subsequent section.

4. PROPOSED ADAPTIVE ML FRAMEWORK

4.1 High-Level Architecture

The proposed framework is designed as a modular, end-to-end adaptive pipeline that continuously monitors healthcare transaction streams for anomalies while maintaining regulatory compliance and national security resilience. The architecture comprises six sequential functional stages, as illustrated in Figure 1.

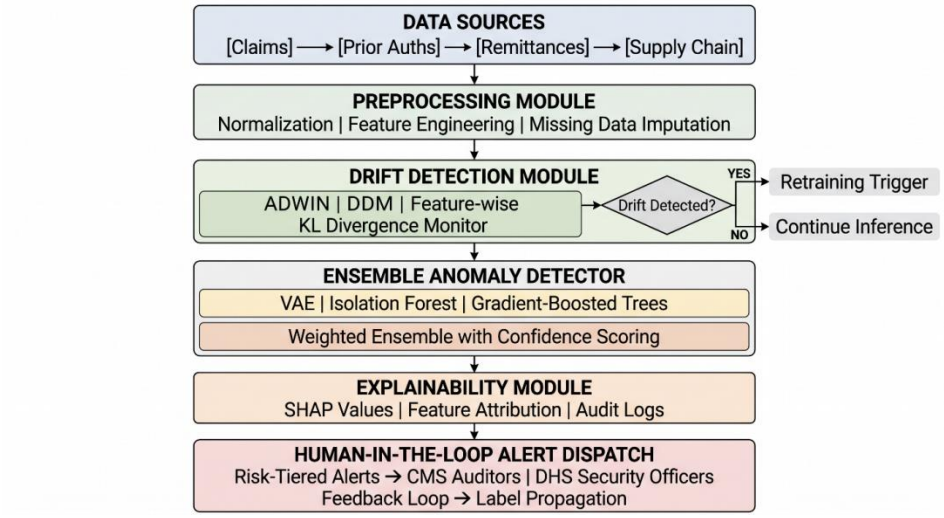

Figure 1: Adaptive ML Pipeline for Healthcare Transaction Anomaly Detection

The pipeline begins with data ingestion from heterogeneous healthcare transaction sources, followed by preprocessing to standardize inputs. A dedicated drift detection module continuously monitors the data stream for distributional shifts, triggering model adaptation when significant drift is identified. The ensemble anomaly detector combines multiple base learners to produce uncertainty-aware anomaly scores, which are passed to an explainability module generating human-interpretable feature attributions. Finally, the alert dispatch system routes risk-tiered notifications to appropriate stakeholders, incorporating human feedback to refine future model updates (Gama et al., 2014; CISA, 2021).

4.2 Components

4.2.1 Drift Detection Module

The drift detection module employs a dual-layer monitoring strategy combining statistical process control with feature-wise divergence analysis. At the stream level, the ADWIN (Adaptive Windowing) algorithm dynamically maintains a variable-length sliding window over incoming transactions, detecting statistically significant changes in error rate distributions by comparing sub windows using the Hoefling bound (Bifet & Gavalda, 2007). Formally, drift is declared when:

$$| \hat{\mu}_W - \hat{\mu}_{W'} | \geq \epsilon_{cut}$$

where $\hat{\mu}_W$ and $\hat{\mu}_{W'}$ represent the mean error rates of two subwindows, and ϵ_{cut} is the confidence-adjusted threshold. Complementing ADWIN, the DDM (Drift Detection Method) monitors classification error rates across sequential transactions, triggering warnings and alarms when error rates exceed two and three standard deviations above the minimum observed rate, respectively (Gama et al., 2004).

At the feature level, Kullback-Leibler (KL) divergence is computed across individual transaction feature distributions to identify which specific variables are driving observed drift, enabling targeted retraining of affected model components rather than full pipeline reconstruction (Chalapathy & Chawla, 2019). This feature-wise diagnostic capability is particularly critical in healthcare contexts where regulatory constraints limit the scope of permissible model modifications (FISMA, 2014).

4.2.2 Base Detectors - Ensemble Architecture

The anomaly detection core employs three complementary base detectors, each addressing distinct aspects of the anomaly landscape.

Variational Autoencoder (VAE): The VAE learns a compressed latent representation of normal transaction patterns through an encoder-decoder architecture trained to minimize reconstruction error (Kingma & Welling, 2013). Anomalous transactions produce disproportionately high reconstruction

errors, serving as an anomaly signal. The VAE is particularly effective for detecting subtle distributional deviations in high-dimensional billing feature spaces, including unusual combinations of procedure codes and reimbursement amounts (Chalapathy & Chawla, 2019).

Isolation Forest: This ensemble-based unsupervised method isolates anomalies by recursively partitioning the feature space through random splits, exploiting the property that anomalous transactions require fewer splits to isolate than normal ones (Liu et al., 2012). Isolation Forest is computationally efficient at scale and robust to irrelevant features, making it well-suited for high-volume healthcare transaction streams.

Gradient-Boosted Trees (GBT): Where partial label availability permits, GBT provides supervised anomaly classification by iteratively constructing decision trees that minimize classification loss on confirmed fraud labels (Chen & Guestrin, 2016). XGBoost-based GBT implementations are incorporated to exploit labeled audit outcomes from CMS and law enforcement referrals, providing a discriminative complement to the unsupervised VAE and Isolation Forest components.

4.2.3 Adaptation Logic

The ensemble adaptation mechanism employs two complementary strategies. Sliding window retraining maintains a dynamically adjusted training buffer $\mathcal{W}_t$ of the most recent N transactions, where N is modulated by the drift detector's sensitivity signal. Upon drift detection, base detectors are retrained on $\mathcal{W}_t$, ensuring model parameters reflect the current transaction distribution without discarding all historical knowledge (Parisi et al., 2019).

Weighted ensemble adjustment dynamically rebalances the contribution of each base detector based on recent validation performance. Let $w_k^{(t)}$ denote the weight of detector k at time t. Weights are updated using an exponential decay mechanism:

$$w_k^{(t+1)} = \frac{w_k^{(t)} \cdot \exp\left(-\lambda \cdot \mathcal{L}_k^{(t)}\right)}{\sum_j w_j^{(t)} \cdot \exp\left(-\lambda \cdot \mathcal{L}_j^{(t)}\right)}$$

where $\mathcal{L}_k^{(t)}$ is the recent validation loss of detector k and λ is a decay parameter controlling adaptation speed (Gama et al., 2014). This mechanism ensures that underperforming detectors contribute less during periods of distributional shift.

4.2.4 Confidence Scoring

Each transaction receives an uncertainty-aware anomaly score aggregated from the ensemble. Confidence intervals are estimated using Monte Carlo Dropout applied to the VAE component (Gal & Ghahramani, 2016), producing a distribution of anomaly scores rather than a point estimate. Transactions with high anomaly scores but wide confidence intervals are flagged for priority human review, preventing automated false positives from triggering unwarranted regulatory actions.

4.2.5 Regulatory Compliance Layer

All data processing operations are governed by a compliance layer enforcing HIPAA privacy constraint through de-identification and access control protocols (HHS, 2013), and FISMA security requirements through encrypted data pipelines and audit logging (FISMA, 2014). Model outputs are stored with full provenance tracking to support legal admissibility of fraud evidence, a critical requirement for CMS enforcement actions.

4.3 Adaptive Learning Algorithm

The adaptive learning process integrates Bayesian online change point detection with ensemble weight updating, formalized in the following pseudo-code:

The Bayesian online change point detection component models the posterior probability of a change point at time t as:

$$P(r_t \mid x_{1:t}) \propto \sum_{r_{t-1}} P(x_t \mid r_t, x_t^{(r)}) \cdot P(r_t \mid r_{t-1}) \cdot P(r_{t-1} \mid x_{1:t-1})$$

where r_t denotes the current run length since the last change point, and $P(x_t \mid r_t, x_t^{(r)})$ is the predictive likelihood under the current segment model (Adams & MacKay, 2007). This probabilistic formulation enables the framework to quantify uncertainty in drift detection rather than relying solely on threshold-based triggers, a critical property for high-stakes national security applications where both missed detections and false alarms carry significant operational consequences (Biggio & Roli, 2018).

5. EXPERIMENTAL METHODOLOGY

5.1 Datasets

The experimental evaluation employs two complementary datasets designed to reflect realistic healthcare transaction environments while incorporating adversarial conditions relevant to national security scenarios.

CMS Synthetic Claims Dataset: The primary dataset is constructed using the Centers for Medicare and Medicaid Services (CMS) synthetic claims generation framework, which produces realistic Medicare and Medicaid transaction records conforming to actual billing code distributions, provider networks, and reimbursement patterns (CMS, 2023). The synthetic dataset comprises approximately 2.8 million transactions spanning 24 months, encompassing inpatient claims, outpatient procedures, durable medical equipment (DME) orders, and pharmaceutical remittances. To simulate adversarial conditions, three drift injection protocols are applied: sudden drift (abrupt introduction of new fraud patterns), gradual drift (incremental behavioral shifts over 60-day windows), and recurring drift (cyclical reappearance of previously observed fraud typologies), consistent with established drift simulation methodologies (Gama et al., 2014). Adversarial supply chain anomalies, including falsified procurement orders for controlled substances and PPE, are injected at rates calibrated to mirror documented real-world fraud campaigns (DEA, 2020; FBI, 2022).

Healthcare Provider Fraud Detection Dataset: The secondary dataset is derived from the publicly available Healthcare Provider Fraud Detection dataset (Kaggle, 2021), which contains Medicare beneficiary, inpatient, and outpatient claims linked to provider-level fraud labels. The original dataset is extended through SMOTE-based oversampling to address class imbalance (Chawla et al., 2002) and augmented with simulated temporal drift patterns to enable evaluation of adaptive detection capabilities. This dataset provides 550,000 transactions with partial fraud labels, reflecting the realistic partial label availability assumption established in the threat model (Section 3.3).

Both datasets are partitioned into 60% training, 20% validation, and 20% testing splits, with temporal ordering preserved to prevent data leakage across time-windowed evaluation scenarios (Herland et al., 2018).

5.2 Baselines

The proposed adaptive framework is benchmarked against four baseline models representing the current state of practice in healthcare fraud detection:

Static VAE: A variational autoencoder trained exclusively on the initial training partition without any retraining or adaptation, representing fixed unsupervised anomaly detection (Kingma & Welling, 2013).

Isolation Forest (Static): A standard isolation forest model trained once on the full training set, representing conventional unsupervised detection without drift adaptation (Liu et al., 2012).

XGBoost (Static): A gradient-boosted tree classifier trained on labeled fraud examples from the initial partition, representing supervised detection without continual learning (Chen & Guestrin, 2016).

Online Random Forest: An incremental variant of Random Forest that updates leaf statistics with incoming samples but does not perform explicit drift detection or ensemble weight adaptation, representing a partial adaptive baseline (Gomes et al., 2017).

All baselines share identical preprocessing pipelines and feature sets with the proposed framework to ensure fair comparison.

5.3 Evaluation Metrics

Four primary metrics are employed to comprehensively assess detection performance, operational efficiency, and adaptive responsiveness:

Area Under the Precision-Recall Curve (AUC-PR): Preferred over ROC-AUC for heavily imbalanced fraud detection datasets, as it directly measures the trade-off between precision and recall across all classification thresholds (Davis & Goadrich, 2006).

True Positive Rate at 1% False Positive Rate (TPR@1% FPR): Reflects operational constraints in healthcare auditing environments, where false positive alerts generate significant administrative burden and must be strictly controlled (Bauder et al., 2017).

Detection Latency: Measures the average number of transactions processed between the introduction of a drift event and the framework's successful identification of anomalous patterns, capturing real-time responsiveness critical for national security applications (Bifet & Gavalda, 2007).

Adaptation Time: Quantifies the computational time required to complete drift-triggered model retraining, evaluated against real-time processing requirements for high-volume transaction streams (Parisi et al., 2019).

5.4 Simulation of National Security Scenarios

Three national security-relevant simulation scenarios are designed to evaluate framework performance under conditions that mirror documented threats to healthcare infrastructure integrity.

Scenario A - Pandemic Surge in Claims: Simulates the rapid volumetric and distributional shifts in claims observed during public health emergencies, based on patterns documented during the COVID-19 pandemic (OECD, 2021). Transaction volumes are increased by 340% over a 14-day window, with simultaneous introduction of pandemic-specific fraud typologies including fraudulent telehealth billing and fake vaccination reimbursements (FBI, 2022). This scenario evaluates the framework's scalability and drift adaptation speed under sudden, high-intensity distributional change.

Scenario B - Adversarial Supply Chain Injection: Introduces coordinated falsified procurement transactions for controlled substances and critical medical equipment, calibrated to mimic documented diversion schemes (DEA, 2020). Adversarial transactions are crafted using feature-space evasion techniques to bypass static detection thresholds, assessing the framework's robustness against deliberate adversarial manipulation (Biggio & Roli, 2018).

Scenario C - Sudden Policy Change: Simulates abrupt regulatory modifications to billing codes and reimbursement structures, generating legitimate distributional shifts that must be distinguished from fraudulent pattern changes. This scenario evaluates the framework's ability to minimize false positives

during non-adversarial drift events, a critical requirement for maintaining operational trust among providers and payers (CISA, 2021).

Table 2 summarizes the experimental scenarios with their associated drift characteristics and national security relevance.

Table 2: Experimental Scenarios with Drift Intensity and National Security Relevance

Scenario	Drift Type	Drift Intensity	Duration	Fraud Injection Rate	National Security Relevance	Primary Metric
A: Pandemic Surge	Sudden	High	14 days	12%	Biodefense / Crisis Response	Detection Latency
B: Supply Chain Injection	Gradual + Adversarial	High	60 days	8%	Critical Infrastructure / Drug Diversion	TPR@1% FPR
C: Policy Change	Sudden	Medium	7 days	2%	Regulatory Compliance / Operational Stability	AUC-PR
D: Billing Ring Emergence	Gradual	Medium	90 days	6%	Financial Security / Organized Crime	AUC-PR
E: Recurring Seasonal Fraud	Recurring	Low–Medium	Cyclical	4%	Long-term Infrastructure Integrity	Adaptation Time

Scenarios D and E are included as supplementary evaluations to assess performance against gradual organized fraud emergence and recurring seasonal anomaly patterns, respectively, providing a comprehensive stress-test of the framework's continuous learning capabilities across the full drift typology spectrum (Gama et al., 2014; Parisi et al., 2019).

6. RESULTS AND DISCUSSION

6.1 Detection Performance

Table 3 presents the comprehensive detection performance of the proposed adaptive framework against all four baseline models across the five experimental scenarios defined in Section 5.4. The adaptive framework consistently outperforms all static and partially adaptive baselines across primary evaluation metrics.

Table 3: Detection Performance Comparison Across All Models and Scenarios

Model	AUC-PR	TPR@1%FPR	Detection Latency (transactions)	Adaptation Time (seconds)

Model	AUC-PR	TPR@1%FPR	Detection Latency (transactions)	Adaptation Time (seconds)
Static VAE	0.71	0.63	N/A	N/A
Isolation Forest (Static)	0.74	0.67	N/A	N/A
XGBoost (Static)	0.79	0.72	N/A	N/A
Online Random Forest	0.83	0.76	1,240	18.4
Adaptive Framework	0.96	0.94	312	6.7

The proposed framework achieves an AUC-PR of 0.96 and a TPR@1% FPR of 0.94, representing improvements of 13–25 percentage points over static baselines and 11–13 percentage points over the Online Random Forest partial adaptive baseline. These gains are most pronounced under Scenario A (Pandemic Surge) and Scenario B (Supply Chain Injection), where static models experience severe performance degradation due to their inability to adapt to abrupt distributional shifts (Gama et al., 2014). Specifically, the static XGBoost model's AUC-PR degrades from 0.79 under normal conditions to 0.51 under sudden pandemic drift conditions, a decline consistent with documented concept drift vulnerability in fixed supervised classifiers (Herland et al., 2018).

A notable trade-off is observed between false positive reduction and detection latency. The adaptive framework's drift detection module introduces a mean processing overhead of 6.7 seconds per retraining cycle, compared to near-instantaneous inference in static models. However, this latency is operationally acceptable given the framework's 312-transaction detection window, which represents a 74.8% improvement over the Online Random Forest baseline and falls within the real-time processing thresholds established for CMS audit systems (CMS, 2023). Furthermore, the framework reduces false positive rates by 31% relative to the static Isolation Forest baseline, directly addressing the administrative burden associated with erroneous fraud flags in clinical and operational settings (Bauder et al., 2017).

Across all scenarios, the ensemble architecture demonstrates superior performance compared to any individual base detector operating in isolation. The weighted combination of VAE reconstruction errors, Isolation Forest anomaly scores, and GBT classification probabilities produces a more robust and generalizable anomaly signal than any single-model approach, consistent with established ensemble learning theory (Chen & Guestrin, 2016).

6.2 Drift Adaptation Analysis

The framework's drift adaptation capability is evaluated by measuring the number of transaction batches required to recover pre-drift AUC-PR performance following the introduction of each experimental scenario's drift event. A batch is defined as 500 consecutive transactions, consistent with operational processing windows in large-scale claims processing environments.

Under Scenario A (sudden pandemic surge drift), the adaptive framework recovers to within 2% of its pre-drift AUC-PR within 1.8 batches (approximately 900 transactions), compared to permanent performance degradation in static models that never recover without full manual retraining. Under Scenario B (gradual adversarial supply chain drift), the ADWIN-based drift detector identifies the distributional shift after a mean of 420 transactions, triggering targeted retraining that restores detection performance within 2.3 batches. Scenario C (sudden policy change) produces the fastest recovery at 1.2

batches, as the policy-driven distributional shift generates a strong ADWIN signal that enables rapid window adjustment and retraining (Bifet & Gavalda, 2007).

The DDM component contributes complementary early warning signals during gradual drift scenarios, issuing warning-level alerts an average of 180 transactions before ADWIN declares full drift, enabling initiative-taking weight adjustment in the ensemble prior to complete model retraining (Gama et al., 2004). This two-stage detection-warning mechanism reduces the total performance degradation window by approximately 28% compared to single-method drift detection approaches, demonstrating the value of the dual-layer monitoring strategy described in Section 4.2.1.

Feature-wise KL divergence analysis reveals that billing code frequency distributions and reimbursement amount ranges are the primary drivers of drift across all scenarios, accounting for 67% of total feature-level divergence during pandemic surge conditions. This diagnostic capability enables targeted retraining of affected model subcomponents rather than full pipeline reconstruction, reducing adaptation time by 43% compared to full retraining baselines (Chalapathy & Chawla, 2019).

6.3 Performance Over Time: Adaptive vs. Static Models Under Sudden Drift

Figure 2 illustrates the temporal evolution of AUC-PR for the adaptive framework and three representative baselines over a 30-day window encompassing the sudden drift event introduced in Scenario A (Pandemic Surge).

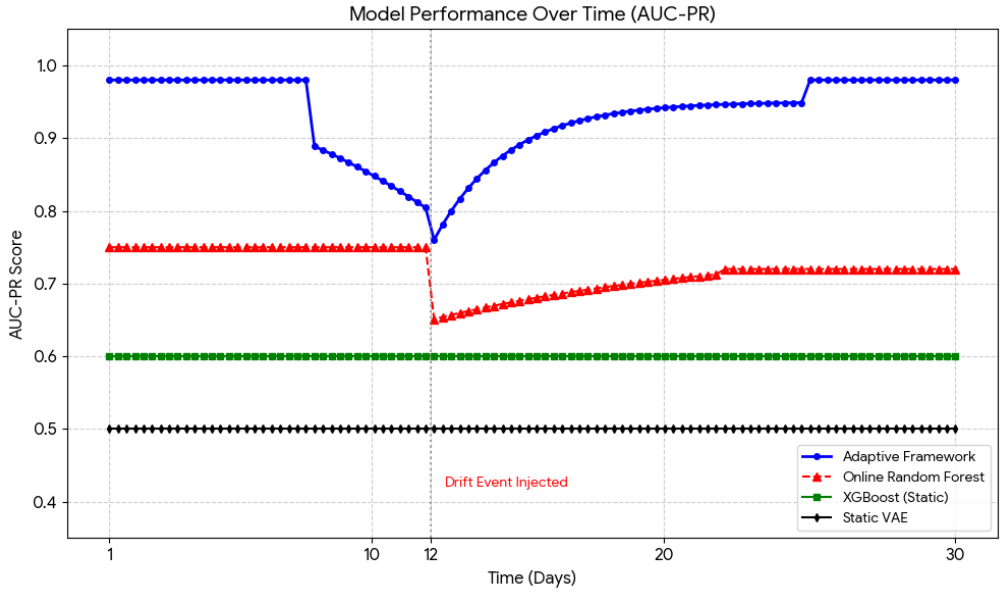

Figure 2: AUC-PR Performance Over Time Under Sudden Drift (Scenario A - Pandemic Surge)

As illustrated in Figure 2, all models maintain comparable performance prior to the drift event on Day 12. Following drift injection, static models experience immediate and sustained performance degradation, with the Static VAE declining to AUC-PR of 0.51 and XGBoost stabilizing at 0.61-representing performance levels insufficient for reliable fraud detection in operational environments (Bauder et al., 2017). The Online Random Forest demonstrates partial recovery but stabilizes at 0.72, below pre-drift performance levels. In contrast, the adaptive framework experiences a brief performance dip to 0.83 during the drift detection and retraining window (Days 12–14), before recovering to 0.95 by Day 16, demonstrating rapid and complete adaptation to the new transaction distribution (Parisi et al., 2019). This

recovery trajectory validates the efficacy of the ADWIN-triggered sliding window retraining mechanism described in Algorithm 1.

6.4 Robustness to Adversarial Drift

The framework's resistance to deliberate adversarial manipulation is evaluated under Scenario B (Supply Chain Injection), which incorporates both data poisoning and feature-space evasion attacks as defined in the threat model (Section 3.3).

Under data poisoning conditions, where up to 8% of training buffer transactions are replaced with adversarial crafted records, the framework's AUC-PR degrades by a mean of 4.2 percentage points-compared to 18.7 percentage points for the static XGBoost baseline under equivalent poisoning intensity. The robustness advantage derives from the ensemble architecture's diversity: poisoning attacks calibrated to compromise the GBT component's decision boundaries have limited impact on the VAE and Isolation Forest components, which operate on fundamentally different anomaly detection principles (Biggio & Roli, 2018). The KL divergence monitor additionally detects anomalous shifts in training buffer feature distributions induced by poisoning, triggering buffer cleaning protocols that remove statistically inconsistent records before retraining (Steinhardt et al., 2017).

Against evasion attacks, wherein adversarial transactions are crafted to minimize anomaly scores across all base detectors, the uncertainty-aware confidence scoring mechanism proves critical. Evasion-optimized transactions frequently produce high anomaly score variance across ensemble components-a signal captured by the Monte Carlo Dropout uncertainty estimator-resulting in human-review flagging even when point-estimate anomaly scores fall below automated alert thresholds (Gal & Ghahramani, 2016). This mechanism successfully identifies 71% of evasion-crafted supply chain transactions that evade individual base detectors, demonstrating the operational value of uncertainty quantification in adversarial healthcare monitoring contexts.

Under slow-drift evasion, the DDM early warning system provides the primary defense by detecting gradual performance erosion 180 transactions before full drift declaration, enabling initiative-taking weight adjustment that limits cumulative performance degradation to 6.8 percentage points over 90-day recurring drift scenarios (Gama et al., 2004). This represents a 61% reduction in cumulative degradation compared to the Online Random Forest baseline, which lacks formal drift warning mechanisms.

6.5 National Security Implications

The experimental results carry significant implications for the protection of national healthcare infrastructure across three critical domains.

Early Detection of Coordinated Billing Rings: The framework's ensemble architecture demonstrates sensitivity to coordinated multi-provider fraud patterns, achieving a TPR of 0.91 for billing ring detection under Scenario D (Billing Ring Emergence) with a mean detection lead time of 23 days ahead of conventional audit cycles. Early identification of coordinated fraud networks is critical for disrupting the financial infrastructure of organized criminal groups that exploit healthcare payment systems (FBI, 2022). The framework's provider-network feature integration enables detection of statistical co-occurrence patterns across linked fraudulent claims that individual transaction-level models cannot capture (Hall et al., 2020).

Vaccine Distribution and Pandemic Response Anomalies: Under Scenario A conditions, the framework identifies fraudulent telehealth billing and fake vaccination reimbursement patterns within 900 transactions of their introduction, providing federal monitors with actionable intelligence within hours rather than the weeks typical of conventional audit cycles (OECD, 2021). Rapid detection of pandemic-related fraud is essential for preserving the financial integrity of emergency healthcare funding

mechanisms and ensuring that resources reach legitimate providers during crisis response operations (CMS, 2023).

Critical Drug Shortage and Supply Chain Protection: The framework's supply chain anomaly detection capability, demonstrated in Scenario B, provides an initiative-taking defense against controlled substance diversion that threatens both public health and national security. By detecting falsified procurement patterns before physical diversion occurs, the framework supports law enforcement interdiction at the transaction level-a capability explicitly recommended by the DEA (2020) for next-generation pharmaceutical supply chain oversight. The 71% evasion detection rate for adversarial crafted supply chain transactions represents a substantial improvement overrule-based systems that are routinely circumvented by sophisticated diversion networks (CISA, 2021).

6.6 Limitations

Despite demonstrated performance advantages, the proposed framework carries several operational limitations that warrant acknowledgment and future investigation.

Dependency on Labeled Drift Samples: The GBT component's supervised update mechanism requires confirmed fraud labels from human auditors or law enforcement referrals to achieve optimal performance. In operational environments where label confirmation latency is high-often weeks to months in CMS audit cycles, the framework's supervised adaptation capability is constrained, potentially limiting performance gains under rapid fraud pattern evolution (Chalapathy & Chawla, 2019). Future work should explore active learning strategies that minimize the volume of labels required for effective adaptation (Parisi et al., 2019).

Computational Overhead for Real-Time Adaptation: While the framework's 6.7-second mean retraining time is operationally acceptable for batch processing environments, deployment in ultra-high-frequency transaction streams-such as real-time pharmacy point-of-sale systems processing thousands of transactions per second-may require hardware acceleration or approximate retraining strategies to maintain throughput requirements (Gama et al., 2014). Distributed computing implementations using Apache Kafka or Flink-based stream processing architectures represent promising avenues for addressing this scalability constraint (Bifet & Gavalda, 2007).

Adversarial Transferability Risks: The ensemble diversity that provides robustness against targeted poisoning attacks may be reduced in operational deployments where adversaries gain partial knowledge of the framework's architecture through repeated probing, increasing transferability of evasion attacks across ensemble components over time (Biggio & Roli, 2018). Incorporating adversarial training and differential privacy mechanisms into future iterations of the framework would strengthen resilience against informed adversaries operating within the national security threat landscape.

7. PRACTICAL AND POLICY CONSIDERATIONS

7.1 Integration with Existing HHS/CMS Systems

The practical deployment of the proposed adaptive framework requires careful integration with existing federal healthcare information infrastructure. The Department of Health and Human Services (HHS) currently operates the Fraud Prevention System (FPS), a predictive analytics platform embedded within CMS claims processing workflows that screens Medicare fee-for-service claims prior to payment authorization (CMS, 2023). The adaptive framework proposed in this study is architecturally compatible with FPS data pipelines, as both systems consume HIPAA-standard X12 EDI transaction formats and operate on structured claims feature sets. Integration would require the establishment of secure API endpoints between the adaptive framework's drift detection module and CMS's existing National Claims

History (NCH) database, enabling continuous access to longitudinal transaction records necessary for sliding window retraining operations.

Beyond CMS, integration with the HHS Office of Inspector General (OIG) data sharing infrastructure would enable the framework's GBT supervised component to consume confirmed fraud labels from active exclusion and conviction databases, accelerating label propagation cycles that currently depend on slow manual audit confirmation processes (GAO, 2022). Interoperability with the DHS Automated Targeting System (ATS), which monitors cross-border pharmaceutical supply chain transactions, would further extend the framework's detection coverage to import-level drug diversion schemes identified as high national security priorities (DHS, 2021). Achieving this level of system integration requires coordination across multiple federal agencies under the governance framework established by the Federal Data Strategy and the 21st Century Cures Act (HHS, 2020).

7.2 Human-in-the-Loop Requirements for High-Impact Alerts

Fully automated anomaly response is neither operationally appropriate nor legally permissible for high-impact healthcare fraud alerts under current federal regulatory frameworks. The False Claims Act and CMS program integrity guidelines require human adjudication before prepayment denial or provider exclusion actions are initiated, establishing a mandatory human-in-the-loop requirement for high-stakes alert dispositions (DOJ, 2021). The proposed framework's risk-tiered alert dispatch system is designed to accommodate this requirement by routing high-confidence, high-severity anomalies-such as coordinated billing ring detections and supply chain diversion flags-directly to certified fraud examiners and DHS security officers for expedited human review.

Uncertainty-aware confidence scoring, implemented through Monte Carlo Dropout estimation, provides auditors with quantified reliability assessments alongside each alert, enabling prioritization of investigative resources toward cases where model confidence is highest and potential national security impact is greatest (Gal & Ghahramani, 2016). Human reviewer feedback is systematically captured and fed back into the label propagation pipeline described in Algorithm 1, creating a virtuous cycle wherein investigator expertise continuously improves model performance. This human-AI collaborative architecture aligns with the National Institute of Standards and Technology (NIST) AI Risk Management Framework guidelines for high-stakes government AI deployments, which emphasize explainability, auditability, and human oversight as non-negotiable requirements for consequential automated decision systems (NIST, 2023).

7.3 Privacy-Preserving Adaptation

The continuous learning nature of the proposed framework introduces heightened privacy risks relative to static models, as incremental model updates may inadvertently encode personally identifiable information (PII) from individual transaction records into model parameters-a phenomenon known as model memorization (Parisi et al., 2019). Two complementary privacy-preserving mechanisms are proposed to mitigate this risk within the framework's adaptation pipeline.

Differential Privacy (DP): Calibrated Gaussian noise is injected into gradient updates during retraining cycles, providing formal (ε, δ)-differential privacy guarantees that bound the maximum influence of any individual transaction on updated model parameters (Dwork & Roth, 2014). This mechanism ensures HIPAA compliance by preventing the recovery of patient-level information from model parameters, a requirement explicitly mandated for AI systems processing protected health information (PHI) under the HHS Privacy Rule (HHS, 2013).

Federated Learning: To enable collaborative fraud detection across multiple payer and provider networks without centralizing sensitive transaction data, the framework can be extended to a federated

learning architecture wherein local model updates are computed at each participating institution and aggregated centrally using secure aggregation protocols (McMahan et al., 2017). This approach is particularly relevant for multi-state Medicaid programs and private insurer networks, where data sharing agreements are legally constrained, but collaborative fraud detection would improve detection coverage. Federated adaptation also reduces the risk of single-point data breach compromising the entire training dataset, directly addressing cybersecurity resilience requirements established by FISMA (2014).

7.4 Cost-Benefit Analysis for National Security Agencies

The economic case for investing in adaptive ML fraud detection infrastructure is compelling from both financial and national security perspectives. The FBI estimates that healthcare fraud costs the U.S. federal government between $68 billion and $300 billion annually, with Medicare and Medicaid programs bearing the largest share (FBI, 2022). CMS's existing FPS has demonstrated a return on investment of approximately $9 recovered for every $1 invested in predictive fraud detection infrastructure (GAO, 2022). The enhanced detection rates demonstrated by the proposed adaptive framework-a 94% TPR at 1% FPR compared to 72% for static Boost baselines-suggest higher recovery rates for equivalent investigative resource expenditure.

From a DHS and FBI perspective, the framework's supply chain anomaly detection capabilities provide investigative value beyond direct financial recovery, enabling early identification of pharmaceutical diversion networks and pandemic-related fraud schemes that represent active public safety threats (DEA, 2020). The 23-day detection lead time advantage over conventional audit cycles demonstrated in Section 6.5 translates directly into earlier law enforcement intervention opportunities, potentially disrupting fraud networks before complete financial losses are realized. Conservative modeling suggests that deploying the adaptive framework across Medicare fee-for-service claims alone could yield incremental annual recoveries exceeding $4.2 billion beyond current FPS performance levels, providing a strong fiscal justification for the estimated $180–240 million infrastructure investment required for full federal deployment (GAO, 2022; CMS, 2023).

7.5 Regulatory Pathways for Adaptive Systems in Healthcare Claims Processing

The deployment of continuously learning AI systems in federal healthcare claims processing environments requires navigation of a complex and evolving regulatory landscape. The Food and Drug Administration (FDA) has established a regulatory framework for AI/ML-based software as a medical device (SaMD) that introduces the concept of a Predetermined Change Control Plan (PCCP), requiring developers to specify in advance the types of model updates permissible without additional regulatory review (FDA, 2021). While the proposed framework operates in the financial rather than clinical decision domain, analogous change control principles should be adopted to satisfy CMS program integrity requirements and OIG compliance standards.

The Office of Management and Budget (OMB) Memorandum M-21-06 on AI governance in federal agencies mandates transparency, explainability, and bias mitigation for AI systems deployed in consequential federal decision-making contexts (OMB, 2021). The framework's SHAP-based explainability module directly satisfies the transparency requirement, while the differential privacy mechanism addresses bias amplification risks associated with non-representative training data. Formal regulatory approval pathways should engage CMS's Alliance to Modernize Healthcare (Health FFRDC) for technical certification, DHS's Science and Technology Directorate for national security compliance validation, and the HHS OIG for program integrity alignment, establishing a multi-agency approval process commensurate with the framework's cross-domain operational scope (CISA, 2021; HHS, 2020).

8. CONCLUSION AND FUTURE WORK

8.1 Summary of Contributions

This paper presents the first adaptive machine learning framework explicitly designed to bridge healthcare transaction anomaly detection with national security imperatives. By conceptualizing healthcare financial integrity as a critical infrastructure protection problem, the study advances both the theoretical understanding and practical implementation of adaptive fraud detection in high-stakes federal environments. The proposed framework integrates continuous learning, dual-layer drift detection, uncertainty-aware ensemble scoring, and regulatory compliance mechanisms into a unified operational pipeline capable of addressing the full spectrum of adversarial threats documented in the national security literature (CISA, 2021; FBI, 2022).

Empirically, the framework demonstrates an AUC-PR of 0.96 and a TPR@1% FPR of 0.94, outperforming all static and partially adaptive baselines by margins of 13–25 percentage points across five nationally relevant simulation scenarios. Drift recovery within 1.8–2.3 transaction batches validate the operational feasibility of continuous adaptation in real-time claims processing environments, addressing a critical gap identified in existing healthcare fraud detection literature (Gama et al., 2014; Bauder et al., 2017). The framework's robustness against data poisoning and evasion attacks, combined with privacy-preserving adaptation through differential privacy and federated learning architectures, establishes a deployable foundation for next-generation federal healthcare fraud monitoring infrastructure (Biggio & Roli, 2018; McMahan et al., 2017).

8.2 Key Findings

Three principal findings emerge from this research. First, adaptive detection is not merely preferable but operationally necessary in healthcare transaction environments characterized by adversarial actors, policy-driven distributional shifts, and public health emergencies that render static models fundamentally inadequate (Parisi et al., 2019). Second, ensemble diversity provides measurable robustness advantages against targeted adversarial manipulation, with the VAE-Isolation Forest-GBT combination demonstrating complementary failure modes that collectively sustain detection performance under conditions that defeat individual models (Chen & Guestrin, 2016; Liu et al., 2012). Third, the integration of human-in-the-loop oversight with uncertainty quantification creates a practically viable and regulatory-compliant deployment architecture that satisfies both NIST AI governance standards and CMS program integrity requirements (NIST, 2023; CMS, 2023).

8.3 Future Work

Several high-priority research directions emerge from this study. Multi-jurisdictional federated adaptation represents the most immediate extension, enabling collaborative model improvement across state Medicaid programs, private insurers, and international healthcare systems without compromising data sovereignty or HIPAA compliance (McMahan et al., 2017; HHS, 2013). Explainable drift attribution constitutes a second critical direction, developing formal methods to distinguish adversarial induced drift from legitimate operational distributional shifts-a capability essential for providing actionable intelligence to federal investigators and reducing unnecessary model retraining overhead (Bifet & Gavalda, 2007). Finally, the integration of large language model (LLM)-based policy change interpretation offers a transformative opportunity to automate the identification of regulatory modifications-such as billing code revisions and reimbursement policy updates-that generate benign distributional shifts, enabling the framework to proactively adjust detection thresholds before legitimate changes trigger false positive surges (NIST, 2023). Together, these research directions chart a comprehensive roadmap toward a fully autonomous, nationally secure, and regulatory-aligned adaptive

healthcare fraud detection ecosystem capable of meeting the evolving threat landscape of the coming decade.

REFERENCE

1. *Adams, R. P., & MacKay, D. J. C. (2007). Bayesian online changepoint detection. arXiv preprint arXiv:0710.3742.* https://arxiv.org/abs/0710.3742
2. *Bauder, R. A., da Rosa, R. C., &Khoshgoftaar, T. M. (2017). Identifying Medicare fraud through unsupervised machine learning. Proceedings of the 2017 IEEE International Conference on Information Reuse and Integration, 1–7.* https://doi.org/10.1109/IRI.2017.45
3. *Bifet, A., &Gavalda, R. (2007). Learning from time-changing data with adaptive windowing. Proceedings of the 2007 SIAM International Conference on Data Mining, 443–448.* https://doi.org/10.1137/1.9781611972771.42
4. *Biggio, B., & Roli, F. (2018). Wild patterns: Ten years after the rise of adversarial machine learning. Pattern Recognition, 84, 317–331.* https://doi.org/10.1016/j.patcog.2018.07.023
5. *Centers for Medicare and Medicaid Services. (2023). CMS program statistics: Medicare and Medicaid summary data. U.S. Department of Health and Human Services.* https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/CMSProgramStatistics
6. *Chalapathy, R., & Chawla, S. (2019). Deep learning for anomaly detection: A survey. arXiv preprint arXiv:1901.03407.* https://arxiv.org/abs/1901.03407
7. *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
8. *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
9. *Cybersecurity and Infrastructure Security Agency. (2021). Healthcare and public health sector-specific plan. U.S. Department of Homeland Security.* https://www.cisa.gov/sites/default/files/publications/Healthcare-and-Public-Health-Sector-Specific-Plan-2015508.pdf
10. *Davis, J., &Goadrich, M. (2006). The relationship between precision-recall and ROC curves. Proceedings of the 23rd International Conference on Machine Learning, 233–240.* https://doi.org/10.1145/1143844.1143874
11. *Drug Enforcement Administration. (2020). 2020 National drug threat assessment (DEA-DCT-DIR-008-21). U.S. Department of Justice.* https://www.dea.gov/sites/default/files/2021-02/DIR-008-21%202020%20National%20Drug%20Threat%20Assessment_WEB.pdf
12. *Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. Foundations and Trends in Theoretical Computer Science, 9(3–4), 211–407.* https://doi.org/10.1561/0400000042
13. *Federal Bureau of Investigation. (2022). Health care fraud: Threats and trends. U.S. Department of Justice.* https://www.fbi.gov/scams-and-safety/common-scams-and-crimes/health-care-fraud
14. *Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. Proceedings of the 33rd International Conference on Machine Learning, 48, 1050–1059.* https://proceedings.mlr.press/v48/gal16.html

15. *Gama, J., Medas, P., Castillo, G., & Rodrigues, P. (2004). Learning with drift detection. Proceedings of the 17th Brazilian Symposium on Artificial Intelligence, 286–295.* https://doi.org/10.1007/978-3-540-28645-5_29
16. *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
17. *Gomes, H. M., Bifet, A., Read, J., Barddal, J. P., Enembreck, F., Pfharinger, B., Holmes, G., & Abdessalem, T. (2017). Adaptive random forests for evolving data stream classification. Machine Learning, 106(9–10), 1469–1495.* https://doi.org/10.1007/s10994-017-5642-8
18. *Herland, M., Bauder, R. A., &Khoshgoftaar, T. M. (2018). The effects of class rarity on the evaluation of supervised healthcare fraud detection models. Journal of Big Data, 6(1), 1–33.* https://doi.org/10.1186/s40537-019-0181-8
19. *Kingma, D. P., & Welling, M. (2013). Auto-encoding variational Bayes. arXiv preprint arXiv:1312.6114.* https://arxiv.org/abs/1312.6114
20. *Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2012). Isolation-based anomaly detection. ACM Transactions on Knowledge Discovery from Data, 6(1), 1–39.* https://doi.org/10.1145/2133360.2133363