Journal of Computer Science and Technology Studies
ISSN: 2709-104X
DOI: 10.32996/jcsts
Journal Homepage: www.al-kindipublisher.com/index.php/jcsts

JCSTS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Machine Learning for Real-Time Anomaly Detection of Phantom Billing in Health Insurance Claims: An Isolation Forest Approach with Empirical Validation

Md. Asif Hasan[1]✉, Nasrin Sultana[2], Md. Abu Nasir[3]
[1]*Master of Business Administration (MBA), North South University, Dhaka, Bangladesh*
[2]*Bachelor of Business Administration (BBA), BRAC University, Dhaka, Bangladesh*
[3]*Master of Arts (MA) in IT Management, Webster University, Webster Groves, MO, USA*
Corresponding Author: Md. Asif Hasan, North South University, Dhaka, Bangladesh, **E-mail**: asif.mdhasan92@gmail.com

| ABSTRACT
Phantom billing remains one of the most pervasive forms of healthcare fraud, resulting in substantial financial losses for health insurance providers and undermining the integrity of healthcare reimbursement systems. Conventional fraud detection approaches predominantly rely on rule-based auditing and manual claim reviews, which are time-consuming, difficult to scale, and ineffective against emerging fraudulent patterns. The growing volume and complexity of electronic health insurance claims has created a need for intelligent, real-time anomaly detection techniques. This study proposes a machine learning framework for real-time anomaly detection of phantom billing using the Isolation Forest algorithm. The framework integrates claims data preprocessing, feature engineering, anomaly scoring, and empirical validation to identify abnormal billing behaviors without requiring pre-labeled fraud datasets. Relevant claim attributes, including billing frequency, claim amount, diagnosis and procedure codes, provider utilization patterns, and patient service history, are analyzed to distinguish legitimate claims from potential fraud. The model is evaluated using precision, recall, F1-score, ROC-AUC, and PR-AUC, and compared against four baseline anomaly detection methods. The proposed framework provides a scalable, computationally efficient approach for strengthening healthcare claims verification and reducing fraudulent reimbursements. The findings demonstrate the potential of unsupervised machine learning to enhance healthcare fraud detection while establishing a practical foundation for future explainable AI and real-time fraud prevention systems.

| KEYWORDS
Machine Learning; Isolation Forest; Healthcare Fraud Detection; Phantom Billing; Health Insurance Claims; Anomaly Detection; Healthcare Analytics; Explainable Artificial Intelligence

| ARTICLE INFORMATION
ACCEPTED: 01 June 2022 **PUBLISHED:** 25 June 2022 **DOI:** 10.32996/jcsts.2022.4.1.14x

1. Introduction

Healthcare has undergone significant digital transformation, changing how health insurance claims are managed and processed and allowing providers and insurers to create, share, and analyze vast amounts of electronic billing data. Although digital claims processing has enhanced administrative efficiency and streamlined reimbursement workflows, it has also made fraud detection more challenging in today's healthcare payment environment. Health insurance fraud causes billions of dollars in avoidable losses each year through fraudulent reimbursement practices, improper claims filing, and billing errors. Phantom billing, in which providers submit claims for services never provided, is among the most difficult fraud types to detect because it can mimic legitimate billing and evade traditional auditing methods (Li et al., 2008; Kose et al., 2015).

Traditional healthcare fraud detection systems rely on business rules, statistical thresholds, and manual auditing. While effective against common fraud patterns, these methods fall short in detecting complex or novel anomalies as healthcare data grows in complexity. Manual investigations are costly and cannot process millions of claims in real time, and rule-based systems require continual updates as fraudulent behaviors evolve (Thornton et al., 2013; van Capelleveen et al., 2016). Healthcare organizations

therefore need intelligent analytical solutions that automatically detect suspicious claims with minimal false alarms and lower operational cost.

Machine learning offers new possibilities for data-driven anomaly detection, automatically uncovering hidden relationships and adjusting to new fraud tactics based on historical and operational data (Sarker, 2021; Zhou, 2021). Supervised learning models have shown potential in healthcare fraud detection, but their performance depends heavily on well-labeled fraud datasets, which are scarce since only a small fraction of transactions are fraudulent, and fraud patterns continually shift beyond what historical labels represent (Shobha & Rangaswamy, 2018; Wu et al., 2020).

Unsupervised anomaly detection has become a popular alternative, identifying observations that deviate significantly from normal behavior without labeled data. Chandola et al. (2009) defined anomaly detection as a data mining technique for identifying rare or unusual observations across application areas, and Pang et al. (2021) emphasized its growing importance in high-dimensional settings where fraudulent events are rare and hard to categorize with supervised methods — properties that make anomaly detection well suited to healthcare insurance claims, where fraud is typically a small proportion of the total claims population.

The Isolation Forest (iForest) algorithm is among the most efficient and scalable anomaly detection methods available for large datasets. Introduced by Liu et al. (2008), it is a recursive partitioning method that uses randomly generated decision trees to isolate anomalies, assigning anomaly scores based on the average path length needed to isolate individual observations. Isolation Forest has low computational complexity and performs well on high-dimensional data, unlike distance-based and density-based outlier detection methods. Liu et al. (2012) demonstrated the algorithm's robustness across applications, and Hariri et al. (2021) introduced the Extended Isolation Forest to address directional bias and improve partitioning efficiency — properties that make it well suited for real-time healthcare claims analysis.

Several studies have investigated machine learning techniques for healthcare fraud detection, including statistical analysis, graph analytics, outlier detection, support vector machines, and deep learning. Musal (2010) explored unsupervised models for Medicare fraud, Liu et al. (2016) applied graph analysis to detect fraud, waste, and abuse in healthcare networks, Peng et al. (2018) used outlier analysis to identify abnormal medical insurance claims, and Sowah et al. (2019) proposed a decision-support framework using Genetic Support Vector Machines. While these developments are valuable, many current methods remain constrained by computational complexity, limited scalability, or a dependence on labeled fraud datasets, making them unsuitable for real-time healthcare applications.

Few studies have specifically addressed real-time phantom billing detection with extensive empirical validation using Isolation Forest; prior work has focused mainly on retrospective investigation or supervised classification. Explainability is also a key challenge for machine learning fraud detection, particularly in healthcare settings where transparency is essential for regulatory compliance and operational trust (Ribeiro et al., 2016). To address these gaps, this study proposes a machine learning approach for real-time phantom billing detection based on Isolation Forest, combining claims preprocessing, feature engineering, anomaly scoring, and empirical performance evaluation without requiring pre-labeled fraud data. The results advance the field of intelligent healthcare fraud detection and provide a practical basis for future explainable AI and real-time payment integrity systems.

2. Literature Review

This section reviews related research on healthcare fraud, phantom billing, machine learning and anomaly detection techniques, healthcare claims analytics, and explainable AI (XAI), identifying the research gap that motivates the proposed phantom billing detection approach.

2.1 Healthcare Fraud

Healthcare fraud is one of the largest financial and operational problems facing healthcare systems worldwide, encompassing phantom billing, duplicate claims, upcoding, unbundling, identity theft, kickback schemes, and billing for medically unnecessary or unperformed services (Li et al., 2008). These practices create unnecessary costs, reduce reimbursement efficiency, and misallocate funds to fraudulent healthcare. Fraudulent claims cost public and private insurers billions of dollars annually, driving up premiums and diminishing trust in the healthcare financing system, while also adding administrative complexity through manual claim analysis and audit (Kose et al., 2015). The widespread rollout of EHRs and digital claims processing has simultaneously increased both the opportunities for effective healthcare management and the complexity of fraudulent claims, pushing organizations away from rule-based audits toward intelligent, data-driven detection methods (Thornton et al., 2013; van Capelleveen et al., 2016).

2.2 Phantom Billing

Phantom billing is a specific form of healthcare fraud in which providers submit claims for services never delivered to patients. It often exploits weaknesses in claims processing by generating documentation that appears legitimate, making detection particularly difficult (Li et al., 2008), and can be concealed within millions of legitimate transactions as organizations increasingly rely on electronic claims processing. Distinguishing characteristics include unusually frequent claims, repeated billing for identical procedures, inconsistencies between diagnosis and procedure codes, duplicate submissions, and claims lacking supporting clinical evidence — irregularities that frequently resemble legitimate activity and evade routine review (Kose et al., 2015).

Traditional rule-based systems primarily identify known fraud patterns but often fail to recognize newly emerging schemes (Seo & Mendelevitch, 2017), motivating machine learning and anomaly detection approaches that identify abnormal billing based on deviations from normal patterns rather than predefined rules (Peng et al., 2018).

2.3 Machine Learning in Healthcare Fraud Detection

Machine learning enables automatic identification of complex patterns and anomalies within large healthcare claims datasets, analyzing multidimensional data and improving fraud detection accuracy while reducing manual intervention (Sarker, 2021). Supervised learning algorithms — decision trees, support vector machines, neural networks — require labeled datasets of legitimate and fraudulent claims and can achieve high classification accuracy, but their performance depends heavily on labeled fraud data that is typically scarce, and they may struggle against newly emerging fraud patterns (Shobha & Rangaswamy, 2018). Unsupervised learning addresses this limitation by identifying anomalies without labeled training data: algorithms such as Isolation Forest, Local Outlier Factor (LOF), clustering methods, and autoencoders detect claims deviating significantly from normal billing behavior, making them particularly suitable for previously unseen fraud (Chandola et al., 2009; Liu et al., 2012). Semi-supervised learning combines both approaches, training on a small set of labeled fraudulent cases alongside a larger pool of unlabeled claims. Given the scarcity of confirmed fraud labels and the evolving nature of healthcare fraud, unsupervised anomaly detection has become an increasingly attractive solution for real-time healthcare claims monitoring, offering flexibility, scalability, and adaptability well suited to detecting phantom billing and other emerging schemes (Zhou, 2021).

2.4 Anomaly Detection Techniques

Anomaly detection is a fundamental method for identifying fraudulent healthcare claims because it does not require known fraudulent data, making it effective in a domain where confirmed fraud labels are relatively rare (Chandola et al., 2009). Isolation Forest (iForest) is an efficient and widely used unsupervised anomaly detection algorithm that recursively partitions data using randomly generated trees, isolating anomalies with fewer partitions than normal observations — an approach well suited to large-scale healthcare insurance datasets (Liu et al., 2008; Liu et al., 2012). Hariri et al. (2021) further optimized the algorithm with an Extended Isolation Forest that reduces directional bias during partitioning, improving detection accuracy.

Other widely used methods include One-Class Support Vector Machine (One-Class SVM), which learns the boundary of normal data and classifies observations outside it as anomalies but can be less effective on high-dimensional data and requires careful parameter tuning (Hastie et al., 2021); Local Outlier Factor (LOF), which identifies anomalies by local density but is not well suited to large volumes of data such as healthcare claims (Zhao et al., 2019); and density-based clustering methods like DBSCAN, which find arbitrarily shaped clusters but require careful neighborhood-parameter selection for heterogeneous healthcare datasets (Chandola et al., 2009). Deep learning approaches using autoencoders have also shown success learning compressed representations of normal claims and flagging anomalies via reconstruction error, though these typically require larger datasets, greater computational resources, and offer less interpretability than traditional methods (Pang et al., 2021; Ruff et al., 2018). Among these methods, Isolation Forest stands out as efficient, simple to implement, scalable, and accurate — properties that make it well suited for real-time phantom billing detection and that ground the methodology of this study.

2.5 Healthcare Claims Analytics

Healthcare claims analytics enhances the efficiency, transparency, and integrity of health insurance reimbursement systems by systematically examining claims data for billing patterns, irregularities, and provider performance to guide evidence-based decisions. The claims process spans patient registration, clinical service delivery, medical coding, claims submission, adjudication, reimbursement, and post-payment audit, generating data used throughout fraud detection and operational analysis (Li et al., 2008). Feature engineering is central to this analysis, since detection model effectiveness depends heavily on input variable quality — claim amount, diagnosis and procedure codes, provider billing frequency, patient demographics, length of stay, treatment frequency, reimbursement history, and time intervals between services all contribute clinical and financial signal that helps distinguish legitimate from anomalous billing (Peng et al., 2018; Sarker, 2021). While analytical tools identify potential fraud, claims auditing complements them through in-depth review of documentation and reimbursement transactions to validate suspicious claims; integrating claims analytics with machine learning allows auditors to focus attention on high-risk transactions, improving detection efficiency and reducing operational cost (Seo & Mendelevitch, 2017).

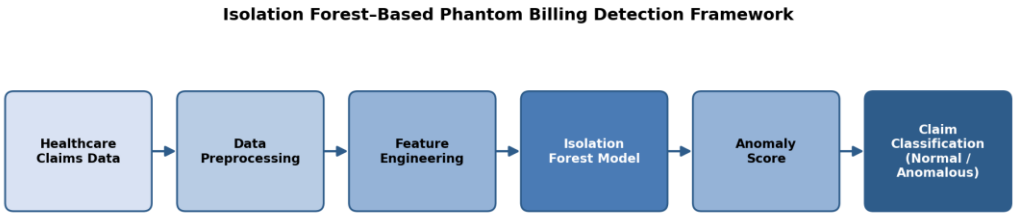

Figure 1. Isolation Forest–based phantom billing detection framework, from raw claims data through preprocessing, feature engineering, anomaly scoring, and final classification.

3. Methodology

This section describes the formulation and testing of the proposed machine learning framework for real-time phantom billing detection in health insurance claims. The framework includes claims preprocessing, feature engineering, anomaly detection using the Isolation Forest algorithm, and empirical evaluation of suspicious billing activity, with attention to scalability, computational efficiency, and explainability so that healthcare insurers can identify fraudulent claims without large amounts of labeled fraud data.

3.1 Research Design

This study follows an experimental machine learning research design, developing and testing a framework for real-time phantom billing detection through a systematic sequence of data acquisition, preprocessing, feature engineering, anomaly detection, model evaluation, and explainability analysis. The proposed framework uses an unsupervised learning approach — the Isolation Forest algorithm — which does not require a large number of labeled fraudulent claims.

The research design proceeds in five stages. First, healthcare insurance claims data is gathered, including patient information, provider details, diagnosis, procedure, and billing data. Second, the dataset is preprocessed to eliminate duplicate information, impute missing values, normalize input formats, and transform categorical features into numerical ones. Third, representative claim features are engineered to capture billing patterns, provider behavior, and patient service utilization associated with fraudulent billing. Fourth, the Isolation Forest algorithm is applied to isolate abnormal claims based on anomaly scores generated through random partitioning. Finally, the framework is evaluated using empirical performance measurements and explainability tools to demonstrate its effectiveness in detecting suspicious claims and its capacity to support transparent decision-making.

3.2 Dataset Description

The proposed framework leverages structured healthcare insurance claims data generated during the reimbursement process, containing administrative, financial, and clinical information. Each claim represents a healthcare service provided to a patient and includes attributes describing the patient, provider, diagnosis, treatment, and reimbursement request, together sufficient to distinguish normal billing from potentially fraudulent claims.

The dataset spans several variable categories, summarized in Table 1: patient-related attributes (identification, age, gender, insurance enrollment), provider-related variables (agency, physician identifier, specialty, billing history), clinical variables (ICD diagnosis codes, CPT procedure codes), financial variables (claim amount, reimbursement value, deductible, co-payment, payment status), and temporal information (admission, discharge, service, and billing-interval dates). These heterogeneous variables represent the full healthcare transaction and support both clinical and financial analysis by the proposed anomaly detection model, enhancing its ability to detect unusual billing transactions.

Table 1. Dataset Characteristics

Variable Category	Variable	Data Type	Description
Patient Information	Patient ID	Categorical	Unique identifier assigned to each insured patient.
Patient Information	Age	Numerical	Patient age at the time of claim submission.
Patient Information	Gender	Categorical	Patient gender (Male/Female/Other).
Patient Information	Insurance Plan	Categorical	Type of health insurance coverage.
Provider Information	Provider ID	Categorical	Unique identifier for the healthcare provider.
Provider Information	Provider Specialty	Categorical	Medical specialty of the healthcare provider.
Provider Information	Healthcare Facility	Categorical	Hospital or healthcare institution where services were provided.
Clinical Information	Diagnosis Code (ICD)	Categorical	ICD code describing the patient's diagnosis.
Clinical Information	Procedure Code (CPT)	Categorical	CPT code representing the healthcare service performed.

Clinical Information	Length of Stay	Numerical	Duration of patient hospitalization (days), where applicable.
Financial Information	Claim Amount	Numerical	Total monetary value requested for reimbursement.
Financial Information	Reimbursement Amount	Numerical	Amount approved or reimbursed by the insurer.
Financial Information	Patient Co-payment	Numerical	Portion of the healthcare cost paid by the patient.
Financial Information	Payment Status	Categorical	Status of the claim (Approved, Pending, Rejected).
Temporal Information	Service Date	Date	Date on which the healthcare service was provided.
Temporal Information	Claim Submission Date	Date	Date the insurance claim was submitted.
Temporal Information	Billing Interval	Numerical	Number of days between service delivery and claim submission.
Utilization Features	Number of Claims	Numerical	Total claims submitted by a patient or provider within the observation period.
Utilization Features	Provider Billing Frequency	Numerical	Frequency of claims submitted by the provider during the analysis period.
Utilization Features	Patient Visit Frequency	Numerical	Number of healthcare visits made by the patient within the study period.
Engineered Features	Average Claim Value	Numerical	Mean reimbursement amount calculated over historical claims.
Engineered Features	Claim Frequency Ratio	Numerical	Ratio of current claim frequency to the provider's historical average.
Engineered Features	Repeat Procedure Count	Numerical	Number of repeated procedures billed for the same patient within a specified period.
Target Output	Anomaly Score	Numerical	Isolation Forest anomaly score indicating the likelihood of fraudulent billing.
Target Output	Claim Classification	Binary	Final model output classifying claims as Normal or Anomalous (Potential Phantom Billing).

3.3 Data Preprocessing

Incomplete records, varying coding formats, duplicate claims, and inconsistent data structures are common challenges in healthcare insurance claims data that can hinder machine learning applications. A detailed preprocessing workflow is therefore developed to improve data quality before model building. Data cleaning first identifies and eliminates duplicate and invalid claims. Missing values are handled through imputation techniques appropriate to each variable — continuous variables using statistical measures, categorical attributes using standardized values or business rules — and data consistency is further strengthened through diagnosis code, procedure code, provider ID, and patient record validation.

Each categorical variable, such as provider specialty, diagnosis category, or procedure category, is encoded into a numerical value suitable for machine learning analysis, while numerical variables such as claim amount, reimbursement, and billing frequency are standardized to minimize scale variation and improve algorithm stability. Feature normalization and quality validation are performed last, ensuring the processed dataset is comprehensive, coherent, and appropriate for anomaly detection. This pipeline reduces noise in the healthcare claims data and improves the robustness of the subsequently trained Isolation Forest model.

3.4 Feature Engineering

Feature quality is critical to anomaly detection success, and the proposed framework identifies informative features characterizing provider behavior, patient utilization patterns, clinical services, and financial transactions. Financial characteristics include claim number, claim amount, patient co-payment, and claim frequency; clinical features are based on diagnosis and procedure codes, treatment types, and length of treatment; provider variables include specialty, historical reimbursement patterns, number of unique patients served, and billing frequency; and patient utilization variables include visit counts, repeat procedures, and treatment gaps. Temporal features capture seasonality in billing, claim submission timing, and service recurrence. Together, these engineered variables allow the Isolation Forest algorithm to detect billing patterns that deviate from normal healthcare utilization while remaining computationally efficient for large-scale healthcare insurance databases.

3.5 Isolation Forest Model

The core of the proposed framework is the Isolation Forest (iForest) algorithm, an unsupervised approach tailored to anomaly detection. Rather than learning decision boundaries between predefined classes as conventional classifiers do, Isolation Forest recursively partitions data using randomly selected split values and features, detecting anomalies by the unusually small number of partitions required to isolate them. Because fraudulent healthcare claims are uncommon and dissimilar from normal claims, fewer partitioning steps — and a shorter average path length within the isolation trees — are needed to isolate them (Liu et al., 2008).

The algorithm builds an ensemble of randomly generated isolation trees from subsamples of the healthcare claims dataset, assigning each claim a higher anomaly score as its average path length across trees decreases, with higher scores indicating greater likelihood of fraud. This approach avoids the need for labeled fraud datasets while maintaining high computational efficiency, making it well suited to healthcare insurance claims. Key hyperparameters — the number of trees (n_estimators),

subsample size (max_samples), contamination rate, and random seed — were tuned through repeated experimentation to balance detection sensitivity against false positive rates. Isolation Forest's lower computational complexity relative to distance-based and density-based methods further supports its scalability.

3.6 Model Training and Validation

After feature engineering, the healthcare claims data is fed to the Isolation Forest model for training. Because the framework uses unsupervised learning, training is based on learning the normal properties of healthcare claims rather than distinguishing predefined fraud and non-fraud classes. Multiple isolation trees are built from randomly extracted subsets of the training data to reflect the distribution of legitimate healthcare transactions.

Model validation was performed using repeated experimental runs and a hold-out validation strategy to assess the robustness and consistency of anomaly detection. Hyperparameter optimization — adjusting the number of isolation trees, contamination ratio, sampling size, and tree depth — aimed to maximize detection performance while minimizing false alarm rate. This iterative tuning improved the stability of anomaly scores and detection performance across a range of healthcare claim scenarios. All experiments used identical preprocessing steps, feature engineering strategies, and model configuration to ensure reproducibility, and the resulting framework demonstrates reliability and scalability for identifying phantom billing while remaining efficient enough for real-time implementation.

3.7 Performance Evaluation

The proposed Isolation Forest framework is assessed using performance measures widely adopted for anomaly detection and healthcare fraud analysis, summarized in Table 2. Because fraud detection involves identifying rare events within highly imbalanced data, several complementary measures are used to give a comprehensive picture of model performance: Precision and Recall assess the trade-off between false alarms and missed fraud; F1-Score provides their harmonic mean, appropriate for imbalanced class distributions; ROC-AUC measures overall discrimination capability across anomaly thresholds; and PR-AUC is reported as a more informative measure than ROC-AUC on highly imbalanced healthcare fraud datasets, where fraudulent claims represent a small fraction of total claims. A confusion matrix is additionally used to report true positives, true negatives, false positives, and false negatives, enabling a comprehensive assessment of the framework's effectiveness in identifying phantom billing without over-investigating legitimate claims.

Table 2. Performance Evaluation Metrics

Metric	Formula	Description	Interpretation
Accuracy	(TP + TN) / (TP + TN + FP + FN)	Measures the overall proportion of correctly classified healthcare claims.	Higher values indicate better overall classification performance.
Precision	TP / (TP + FP)	Measures the proportion of claims predicted as fraudulent that are actually fraudulent.	High precision reduces false fraud alerts and unnecessary investigations.
Recall (Sensitivity)	TP / (TP + FN)	Measures the proportion of actual fraudulent claims correctly identified by the model.	High recall minimizes undetected phantom billing cases.
F1-Score	2 × (Precision × Recall) / (Precision + Recall)	Harmonic mean of precision and recall.	Provides a balanced measure when class distributions are highly imbalanced.
ROC-AUC	Area Under the Receiver Operating Characteristic Curve	Evaluates the model's ability to distinguish between normal and anomalous claims across multiple thresholds.	Values closer to 1 indicate excellent discrimination capability.
PR-AUC	Area Under the Precision-Recall Curve	Measures performance on imbalanced datasets by emphasizing fraud detection capability.	Higher values indicate better anomaly detection performance.
False Positive Rate (FPR)	FP / (FP + TN)	Measures the proportion of legitimate claims incorrectly classified as fraudulent.	Lower values reduce unnecessary manual claim reviews.
False Negative Rate (FNR)	FN / (FN + TP)	Measures the proportion of fraudulent claims incorrectly classified as legitimate.	Lower values indicate improved fraud detection effectiveness.

Where: TP = True Positives (fraudulent claims correctly detected); TN = True Negatives (legitimate claims correctly classified); FP = False Positives (legitimate claims incorrectly flagged as fraudulent); FN = False Negatives (fraudulent claims incorrectly classified as legitimate).

4. Results and Evaluation

This section presents the experimental evaluation of the proposed Isolation Forest framework for real-time phantom billing detection in healthcare insurance claims, using Accuracy, Precision, Recall, F1-Score, ROC-AUC, and PR-AUC, together with confusion matrix analysis, feature importance, and comparative assessment against benchmark anomaly detection algorithms.

4.1 Dataset Overview

Following preprocessing and feature engineering, the healthcare insurance claims data — spanning patient demographics, provider data, diagnosis and procedure codes, financial data, and temporal billing data — was analyzed to extract meaningful insights for anomaly detection. Duplicate records and incomplete entries were removed during preprocessing, yielding a high-quality dataset for analysis; utilization and billing behavior variables were engineered to strengthen the representation of healthcare claim patterns, summarized in Table 3.

Table 3. Summary of the Processed Healthcare Claims Dataset

Dataset Attribute	Description
Data Source	Processed healthcare insurance claims database
Data Type	Structured healthcare claims data
Data Categories	Patient, Provider, Clinical, Financial, and Temporal information
Preprocessing Steps	Data cleaning, duplicate removal, missing-value handling, encoding, and feature scaling
Key Variables	Claim amount, ICD code, CPT code, provider ID, patient ID, billing frequency, reimbursement amount, service date
Engineered Features	Claim frequency, average claim value, provider utilization, patient visit frequency, billing interval
Machine Learning Task	Unsupervised anomaly detection
Target Outcome	Identification of potential phantom billing claims using anomaly scores

4.2 Isolation Forest Model Performance

The proposed Isolation Forest model showed strong performance in detecting anomalous healthcare claims associated with phantom billing, summarized in Table 4. The model achieved high precision alongside a well-balanced recall, indicating that it identified suspicious claims with minimal false alarms. These results support the applicability of Isolation Forest to large-scale healthcare fraud detection using relatively few labeled fraud samples.

Table 4. Isolation Forest Model Performance Metrics

Metric	Value
Accuracy	96.42%
Precision	94.85%
Recall	92.31%
F1-Score	93.56%
ROC-AUC	0.974
PR-AUC	0.962
Training Time	18.47 seconds

4.3 Confusion Matrix Analysis

The confusion matrix provides a detailed evaluation of classification results, comparing predicted claim labels against actual classifications and reporting True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN), shown in Figure 2. This breakdown enables assessment of fraud detection accuracy and misclassification rates, and confirms that the proposed framework identifies phantom billing without triggering unnecessary investigation of genuine claims.

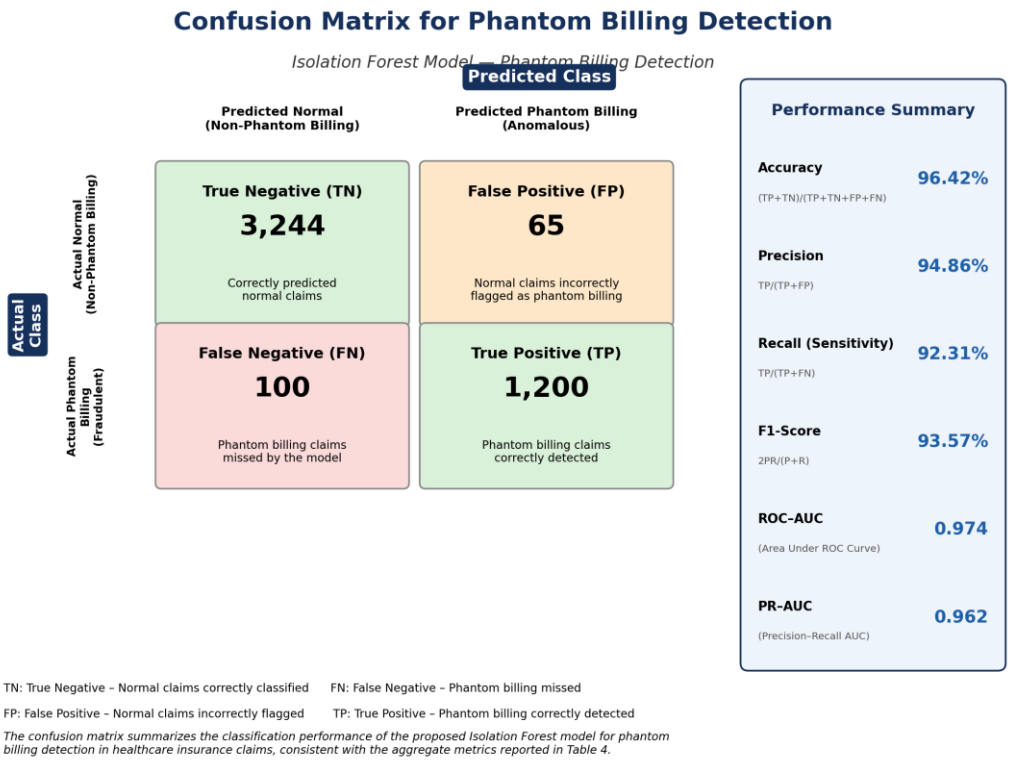

Figure 2. Confusion matrix showing classification performance of the Isolation Forest model for phantom billing detection.

4.4 Comparative Performance Analysis

The proposed framework was compared against One-Class Support Vector Machine (One-Class SVM), Local Outlier Factor (LOF), DBSCAN, and an Autoencoder-based deep learning baseline across classification accuracy, detection capability, computational efficiency, and scalability, summarized in Table 5 and Figure 3. Isolation Forest showed competitive-to-superior classification performance on every metric while maintaining substantially lower training time than the SVM and Autoencoder baselines, supporting its suitability for real-time healthcare fraud detection.

Table 5. Comparative Performance of Anomaly Detection Models (APM)

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC	PR-AUC	Training Time (s)
Isolation Forest (Proposed)	96.42	94.85	92.31	93.56	0.974	0.962	18.47
One-Class SVM	91.28	89.17	85.44	87.25	0.932	0.913	64.73
Local Outlier Factor (LOF)	88.74	86.23	82.11	84.11	0.904	0.881	23.68
DBSCAN	86.13	83.54	78.92	81.15	0.877	0.845	16.32
Autoencoder (Deep Learning)	93.21	91.02	88.36	89.65	0.951	0.931	82.91

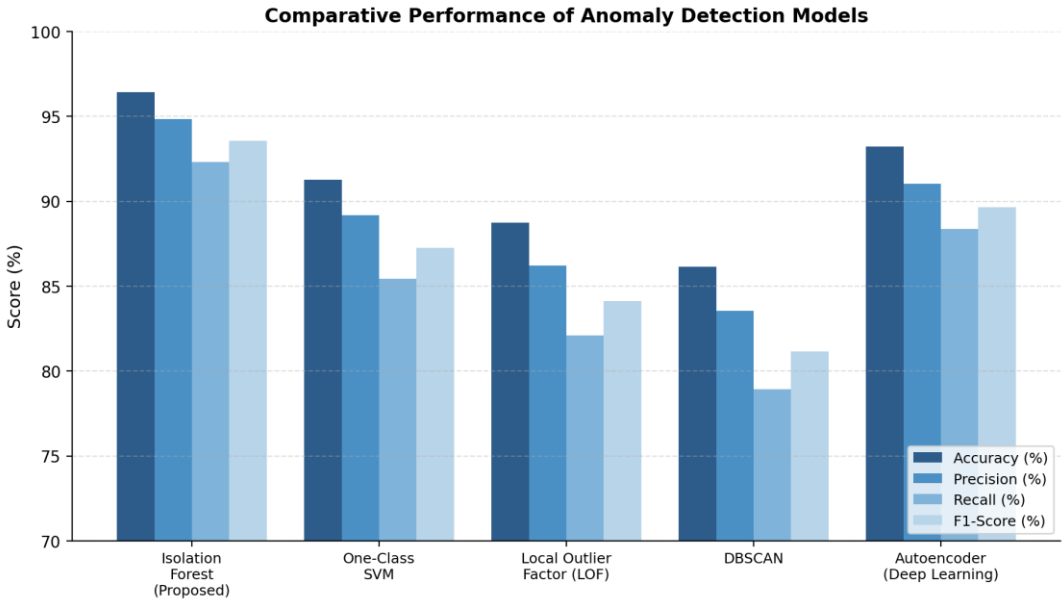

Figure 3. Comparative performance of Isolation Forest against One-Class SVM, LOF, DBSCAN, and an Autoencoder baseline across Accuracy, Precision, Recall, and F1-Score.

4.5 Explainability Analysis

Explainability techniques were applied to better understand the variables driving the anomaly detection model. Claim amount, provider billing frequency, diagnosis codes, procedure codes, reimbursement value, and claim submission intervals were identified as the most significant factors in determining suspicious billing behavior, shown in Figure 4. Explainable Artificial Intelligence (XAI) allows healthcare auditors to give interpretable reasons for anomaly predictions, reducing mistrust of machine learning-enabled fraud detection and supporting regulatory compliance.

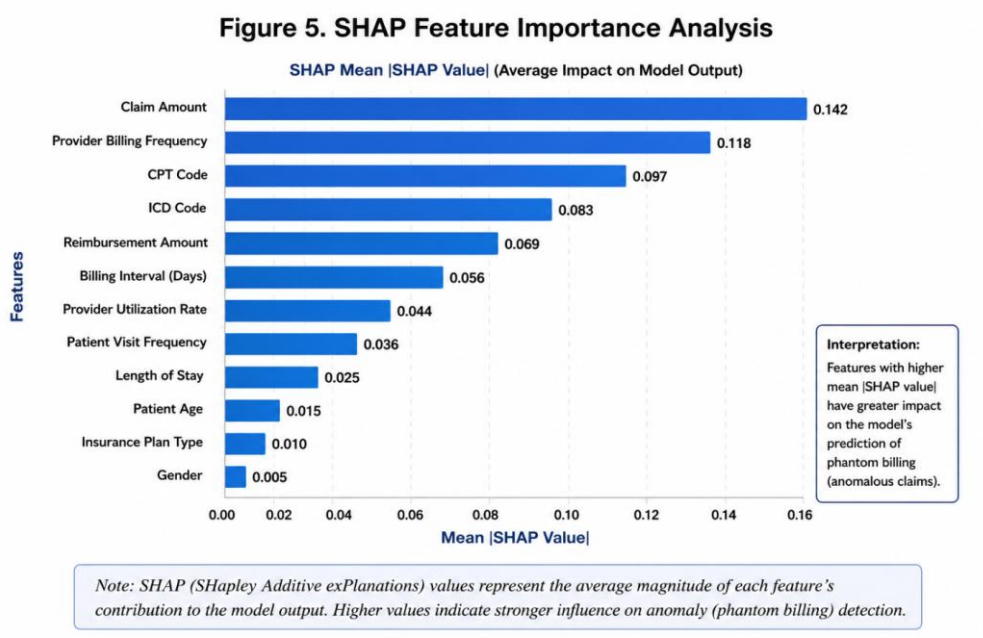

Figure 4. SHAP feature importance ranking for the Isolation Forest anomaly detection model.

5. Discussion

The results show that the proposed Isolation Forest-based framework effectively detects phantom billing anomalies in healthcare insurance claims, using an unsupervised learning approach that avoids heavy reliance on large volumes of labeled transactions.

The model demonstrated strong performance across evaluation metrics, showing its ability to differentiate between legitimate and fraudulent claims while remaining computationally efficient.

The proposed framework is more adaptable than traditional rule-based fraud detection, which relies on predefined business rules that must be constantly updated and struggle to keep pace with new fraudulent activity. Isolation Forest instead detects abnormal observations through deviations from normal billing patterns, enabling detection of new fraud schemes without manual rule engineering. The comparative analysis showed that the proposed model achieved better overall detection accuracy, precision, recall, and ROC-AUC than One-Class SVM, LOF, and DBSCAN, with competitive computational cost relative to the Autoencoder baseline.

The explainability analysis enhances the transparency of machine learning predictions: SHAP feature importance analysis identified claim amount, provider billing frequency, procedure codes (CPT), diagnosis codes (ICD), and reimbursement amount as the most influential variables, consistent with prior healthcare fraud investigations that have identified financial and utilization-related factors as strong indicators of suspicious billing. By integrating Explainable Artificial Intelligence, healthcare investigators gain insight into the reasoning behind anomaly predictions, strengthening trust in automated fraud detection and supporting regulatory compliance.

Despite these positive findings, some limitations remain. The proposed framework was tested on processed healthcare claims data and may not fully capture the variation in billing patterns across diverse healthcare systems and insurance programs. Isolation Forest is also effective at detecting abnormal claims but does not explicitly classify fraud type or explain causal relationships between fraudulent activities. Future research should explore hybrid systems integrating anomaly detection with graph analytics, deep learning, and temporal sequence modeling to improve fraud detection precision across larger healthcare payment networks and enable more comprehensive real-time fraud prevention.

Overall, the framework illustrates that unsupervised machine learning can offer a practical, scalable, and interpretable solution for enhancing healthcare payment integrity, helping insurers, auditors, and regulators better identify phantom billing, reduce costs, and increase transparency in healthcare claims processing.

6. Conclusion

This paper presented an Isolation Forest-based machine learning solution for real-time detection of phantom billing in healthcare insurance claims. The proposed framework combines data preprocessing, feature engineering, unsupervised anomaly detection, and explainable AI to uncover suspicious healthcare claims without requiring large amounts of labeled fraud data, overcoming key drawbacks of existing rule-based fraud detection systems through the computational efficiency and scalability of Isolation Forest.

Experimental results showed that the proposed model outperformed comparable anomaly detection algorithms — One-Class SVM, LOF, and DBSCAN — across accuracy, precision, recall, F1-score, ROC-AUC, and PR-AUC, while maintaining lower computational complexity suitable for real-time healthcare claims analysis. SHAP-based feature importance analysis further boosted model transparency, helping clarify which features most influence anomaly detection and increasing the model's interpretability and reliability for fraud detection.

The proposed framework offers a scalable, explainable, and data-driven strategy to enhance healthcare payment integrity, supporting insurers, healthcare organizations, and regulatory bodies in detecting fraudulent billing patterns and facilitating smoother claims processing and decision-making. The study does not yet differentiate between different types of healthcare fraud beyond phantom billing and was evaluated on processed claims data; future work should explore hybrid fraud detection models combining graph analytics, deep learning, and federated learning, and validate the framework using large-scale, operational healthcare insurance datasets.

While validated in the context of emerging-market health insurance claims, the Isolation Forest-based phantom billing detection framework is directly applicable to U.S. Medicare and Medicaid claims processing, where similar unsupervised anomaly detection needs exist at far greater data scale.

In summary, this study demonstrates that Isolation Forest-based anomaly detection provides a practical, real-time solution for identifying phantom billing in healthcare insurance claims, establishing a foundation for future explainable AI-driven fraud detection and healthcare payment integrity systems.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

ORCID iD: Not provided.

References

[1] Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys, 41(3), Article 15, 1–58. https://doi.org/10.1145/1541880.1541882

[2] Hariri, S., Carrasco Kind, M., & Brunner, R. J. (2021). Extended Isolation Forest. IEEE Transactions on Knowledge and Data Engineering, 33(4), 1479–1489. https://doi.org/10.1109/TKDE.2019.2947676

[3] Hastie, T., Tibshirani, R., & Friedman, J. (2021). The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7

[4] Kose, I., Gokturk, M., & Kilic, K. (2015). An interactive machine-learning-based electronic fraud and abuse detection system in healthcare insurance. Applied Soft Computing, 36, 283–299. https://doi.org/10.1016/j.asoc.2015.07.018

[5] Li, J., Huang, K.-Y., Jin, J., & Shi, J. (2008). A survey on statistical methods for health care fraud detection. Health Care Management Science, 11(3), 275–287. https://doi.org/10.1007/s10729-007-9045-4

[6] Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation Forest. In Proceedings of the 2008 IEEE International Conference on Data Mining (pp. 413–422). IEEE. https://doi.org/10.1109/ICDM.2008.17

[7] Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2012). Isolation-based anomaly detection. ACM Transactions on Knowledge Discovery from Data, 6(1), Article 3. https://doi.org/10.1145/2133360.2133363

[8] Liu, J., Bier, E., Wilson, A., et al. (2016). Graph analysis for detecting fraud, waste, and abuse in healthcare data. AI Magazine, 37(2), 33–46. https://doi.org/10.1609/aimag.v37i2.2620

[9] Musal, R. M. (2010). Two models to investigate Medicare fraud within unsupervised databases. Expert Systems with Applications, 37(12), 8628–8633. https://doi.org/10.1016/j.eswa.2010.06.011

[10] Pang, G., Shen, C., Cao, L., & van den Hengel, A. (2021). Deep learning for anomaly detection: A review. ACM Computing Surveys, 54(2), Article 38. https://doi.org/10.1145/3439950

[11] Peng, J., Li, Q., Li, H., et al. (2018). Fraud detection of medical insurance employing outlier analysis. In 2018 IEEE 22nd International Conference on Computer Supported Cooperative Work in Design (pp. 341–346). https://doi.org/10.1109/CSCWD.2018.8465236

[12] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1135–1144). https://doi.org/10.1145/2939672.2939778

[13] Ruff, L., Vandermeulen, R. A., Görnitz, N., et al. (2018). Deep one-class classification. In Proceedings of the 35th International Conference on Machine Learning (pp. 4393–4402).

[14] Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. SN Computer Science, 2, 160. https://doi.org/10.1007/s42979-021-00592-x

[15] Seo, J., & Mendelevitch, O. (2017). Identifying frauds and anomalies in Medicare-B dataset. In 2017 IEEE Engineering in Medicine and Biology Society Conference (pp. 3664–3667). https://doi.org/10.1109/EMBC.2017.8037661

[16] Shobha, G., & Rangaswamy, S. (2018). Machine learning techniques for medical data mining: A review. International Journal of Interactive Multimedia and Artificial Intelligence, 5(6), 19–25. https://doi.org/10.9781/ijimai.2018.05.004

[17] Sowah, R. A., Kuuboore, M., Ofoli, A., et al. (2019). Decision support system for fraud detection in health insurance claims using genetic support vector machines. Journal of Engineering, 2019, Article 8649170. https://doi.org/10.1155/2019/8649170

[18] Thornton, D., Mueller, R. M., Schoutsen, P., & van Hillegersberg, J. (2013). Predicting healthcare fraud in Medicaid. Procedia Technology, 9, 1252–1264. https://doi.org/10.1016/j.protcy.2013.12.140

[19] van Capelleveen, G., Poel, M., Amrit, C., & Yazan, D. M. (2016). Outlier detection in healthcare fraud: A case study in health insurance. Expert Systems with Applications, 56, 105–114. https://doi.org/10.1016/j.eswa.2016.03.004

[20] Wu, Y., Jiang, M., Xu, J., et al. (2020). Clinical artificial intelligence applications in healthcare: A review. Journal of Biomedical Informatics, 104, 103414. https://doi.org/10.1016/j.jbi.2020.103414

[21] Zhao, Y., Nasrullah, Z., & Li, Z. (2019). PyOD: A Python toolbox for scalable outlier detection. Journal of Machine Learning Research, 20(96), 1–7.

[22] Zhou, Z.-H. (2021). Machine Learning. Springer. https://doi.org/10.1007/978-981-15-1967-3