

Detecting Fraud, Waste, and Abuse in Healthcare Claims using AI: Applying Isolation Forest to Claim Analytics

Naren Chandra¹, Vineel Arekapudi²

¹University of The Cumberlands, Acton, MA, USA (naren1712@gmail.com)

²Blue Cross Blue Shield, Atlanta, GA, USA (Vineel.a004@gmail.com)

Abstract – Healthcare fraud, waste, and abuse (FWA) are among the most significant drivers of financial inefficiency in healthcare systems worldwide, with estimates suggesting they contribute to as much as 10% of total healthcare expenditures. In the United States, these losses are estimated to exceed \$300 billion annually. This paper presents a machine learning approach using Isolation Forest, an unsupervised anomaly detection algorithm, to proactively identify potential FWA in healthcare claims data. Based on real-world patterns, a synthetically generated dataset was created to simulate both legitimate and fraudulent billing behavior. The model achieved high performance with a precision of 0.81, recall of 0.82, F1 score of 0.81, and ROC-AUC of 0.90. Through case studies and comparison of performance with baseline models, we demonstrate the practical applicability of this approach in real-time fraud monitoring. The study also discusses key implementation considerations, including ethical and regulatory factors, explainability, and system integration. We argue that intelligent anomaly detection models can be integrated with payer systems to improve financial stability and healthcare affordability.

Keywords – Healthcare Fraud Detection, FWA, Machine Learning, Anomaly Detection, Explainable AI, Isolation Forest, Claims Analysis

I. INTRODUCTION

Fraud, waste, and abuse (FWA) in healthcare represent a huge burden on the financial stability and operational efficiency of healthcare systems. As medical services and billing practices grow more complex, there is an exponential increase in the opportunities for fraudulent practices and patient exploitation. In the United States, the cost of FWA is estimated to exceed \$300 billion annually (according to National Health Care Anti-Fraud Association-NHCAA), with fraudulent practices like phantom billing and unbundling of services to upcoding and prescription fraud [1,7]. These practices not only drain payer resources but also undermine the integrity of the healthcare system leading to delayed access to care for legitimate patients.

Historically, fraud detection has relied on retrospective auditing, whistleblower tips, or static rule-based engines that are limited in adaptability. These methods are often late in detecting anomalies-after payments are made and financial losses are incurred. With healthcare records getting digitized and with the availability of claims data in large numbers, there is now an opportunity to apply artificial intelligence (AI) and machine learning methods (ML) to proactively detect healthcare fraud [14].

This paper explores how anomaly detection methods, specifically Isolation Forest, can be applied to detect FWA using healthcare claims data. We simulate a synthetic dataset that is designed to reflect standard and suspicious claims activity. Our methodology demonstrates how unsupervised learning can be used effectively when labeled fraud data is not easily available. We further illustrate the real-world implications of these methods by analyzing actual case studies from past fraud investigations and try to bridge the gap

between theoretical machine learning and fraud prevention in payer organizations.

II. RELATED WORK

Artificial intelligence (AI) and machine learning (ML) have emerged as powerful tools in the fight against healthcare fraud, waste, and abuse (FWA). These technologies offer significant advantages over traditional rule-based systems by enabling faster, more scalable, and more proactive detection of suspicious activities. As healthcare data grows in volume and complexity, AI-based approaches have become essential for uncovering patterns and anomalies that manual audits or static rules often miss.

Several major healthcare insurers have already integrated advanced analytics into their fraud detection efforts with notable success. For example, Humana deployed AI-powered models to scan claims data for billing anomalies and suspicious provider behavior, uncovering over \$10 million in potentially fraudulent activity within the first year of deployment [4, 10]. Anthem implemented a natural language processing (NLP)-based system to analyze claims in real time and flag scenarios like duplicate claims or improper upcoding, resulting in a 25% reduction in fraudulent payments in just six months [4]. In collaboration with Mastercard, Milliman leveraged AI to detect over \$239 million in FWA by modeling over 90 fraud scenarios, including biologically implausible treatments and excessive service volumes [12].

Government agencies have also recognized the value of AI in enhancing fraud prevention capabilities. The U.S. Department of Justice (DOJ), Department of Health and Human Services (HHS), and Centers for Medicare & Medicaid Services (CMS) have all adopted AI tools to improve fraud detection, streamline prior authorization, and reduce unnecessary medical expenditures [3]. These models are

increasingly being used to identify emerging threats, such as synthetic identity fraud or manipulated medical records, that traditional methods fail to detect.

The academic literature on healthcare fraud analytics is broadly categorized into three main streams [17]. The first stream provides a foundational overview of statistical approaches—both supervised and unsupervised—adapted from other industries to healthcare contexts [10, 18]. The second stream evaluates the effectiveness of various supervised and unsupervised models in real-world healthcare settings [5]. The third stream explores emerging fraud detection paradigms, including ensemble learning, hybrid models, and deep neural networks [13].

With growing volumes and complexity of healthcare billing and claims, historical methods of fraud detection have given way to new, data-driven approaches that use AI and machine learning algorithms to effectively identify and prevent healthcare claims fraud [2]. Machine learning algorithms can process large amounts of structured and unstructured data to detect inconsistencies and identify anomalies in real-time [17].

Supervised learning methods, while effective, require large volumes of high-quality labeled data. These models learn from historical claims labeled as fraudulent or legitimate, enabling them to predict fraudulent activity in future claims. However, obtaining such labeled datasets is often resource-intensive and subject to legal or regulatory constraints. In contrast, unsupervised learning techniques such as k-means clustering, hierarchical clustering, and autoencoders do not require prior labeling and instead identify outliers based on statistical or structural deviations from the norm [1].

A promising direction lies in hybrid models that combine supervised and unsupervised approaches. Studies such as those by Hassan and Alam have shown that hybrid systems applied to controlled datasets can yield high accuracy while maintaining processing efficiency in claims adjudication workflows.

Despite these advancements, several challenges persist. One major concern is the class imbalance in healthcare datasets, where fraudulent claims are vastly outnumbered by legitimate ones. This imbalance can impair model training and result in high false negative rates. Additionally, the evolving nature of fraud schemes necessitates continual model adaptation and retraining to remain effective [7].

Given these limitations, there is a compelling need for scalable, adaptable fraud detection methods that do not rely on labeled data. This paper addresses that gap by applying Isolation Forest, an unsupervised anomaly detection algorithm, to healthcare claims. Unlike many traditional models, Isolation Forest can operate effectively in the absence of labeled data, making it particularly well-suited for early-stage deployments in payer environments or in situations where labeled fraud examples are scarce or incomplete.

III. METHODOLOGY

The Isolation Forest algorithm was selected over alternatives such as One-Class SVM, Autoencoders, and Local Outlier Factor (LOF) due to its efficiency in handling high-dimensional, unlabeled datasets, and its ability to detect anomalies based on isolation rather than distance or density. Unlike One-Class SVMs, which scale poorly with large data, or Autoencoders that require deep learning infrastructure and

careful hyperparameter tuning, Isolation Forest offers fast computation, robustness to irrelevant features, and interpretable results ideal for prototype-level implementation in healthcare settings [7].

To train and evaluate the model, a synthetic dataset representing realistic distributions of healthcare claim records was constructed. Normal claims were generated using a log-normal distribution to capture the skewed nature of medical billing amounts typical for outpatient services. The synthetic dataset consisted of 10,000 legitimate claims to better reflect payer claim volumes. Each record contained claim amount, CPT code category, procedure group, provider ID, patient age, and date of service. CPT codes were sampled from common outpatient categories (e.g., Evaluation & Management, Radiology), while provider IDs were randomly assigned across 250 providers to introduce diversity.

Fraudulent behavior was injected at multiple severity levels and prevalence rates. Anomalous claims were inflated by 1.5–2× (subtle upcoding), 2–3× (moderate anomalies), and greater than 3× (extreme anomalies) with corresponding mismatched CPT procedure codes and repeated high-cost services within short intervals. These anomalies were inserted at contamination rates of 1%, 5% and 10% to evaluate the robustness of the model. These patterns were informed by real-world fraud cases reported by federal agencies.

All numerical features were standardized using z-score normalization to ensure consistent scaling and to mitigate the influence of varying value ranges. Categorical attributes, such as CPT codes and procedure groups, were one-hot encoded to preserve their categorical semantics and prevent the model from misinterpreting them as ordinal data. Provider identifiers, which often appear with varying frequency, were encoded using frequency encoding to reduce sparsity while retaining important distributional information.

This preprocessing approach creates a feature set that realistically reflects the composition of healthcare claims, combining financial, procedural, and provider-related attributes. The Isolation Forest model was then trained on the full dataset, generating anomaly scores for each claim based on how easily it could be isolated from others. By incorporating both categorical and numeric dimensions, the model is better equipped to distinguish between routine billing activity and atypical or suspicious patterns.

To emulate real-world fraud scenarios, the dataset included multiple categories of anomalies rather than a single extreme outlier class. Subtle deviations—such as slight upcoding or marginal frequency shifts—tested the model’s sensitivity to less obvious fraud, while more significant anomalies simulated deliberate overbilling and fraudulent activity. Additionally, model performance was evaluated at varying contamination levels (1%, 5%, and 10%) to observe how detection efficacy changes with different fraud prevalence rates. These scenarios are critical for tuning sensitivity thresholds in practical deployments, where fraud is both rare and diverse in its presentation.

IV. RESULTS AND EVALUATION

The model was evaluated by comparing predicted anomalies against the known outliers injected into the dataset. Across 1%, 5%, and 10% contamination rates, the Isolation Forest achieved precision between 0.79 and 0.88, recall between 0.70 and 0.87, and F1 scores between 0.78 and 0.83.

For the **5% contamination setting**, which approximates industry fraud prevalence benchmarks, **the model achieved a precision of 0.81, recall of 0.82 and F1 of 0.81, with a ROC-AUC of 0.90**. These results demonstrate robust performance in identifying fraudulent patterns even under more subtle anomalies.

To benchmark performance, we also implemented One-Class SVM and Local Outlier Factor models on the same dataset. The One-Class SVM achieved a precision of 0.83 and an F1-score of 0.78, while the Local Outlier Factor scored 0.81 for precision and 0.74 for F1. In all metrics, Isolation Forest outperformed its counterparts, reinforcing its suitability for unsupervised fraud detection.

The Python implementation leveraged the scikit-learn library and employed standard data preprocessing methods, like one-hot encoding for categorical variables and standard scaling for numerical features.

```
import numpy as np
import pandas as pd

from sklearn.preprocessing import StandardScaler, OneHotEncoder
from sklearn.compose import ColumnTransformer
from sklearn.ensemble import IsolationForest

from sklearn.metrics import precision_score, recall_score, f1_score, roc_auc_score

from sklearn.pipeline import Pipeline

from sklearn.model_selection import train_test_split

# Seed for reproducibility
np.random.seed(42)

# Parameters
n_normal = 10000
providers = [f"P{str(i).zfill(3)}" for i in range(250)]
cpt_categories = {
    "E&M": ["99213", "99214", "99215"],
    "Radiology": ["71020", "73030"],
    "Pathology": ["80048", "80053"]
}

procedure_groups = list(cpt_categories.keys())
contamination_rates = [0.01, 0.05, 0.10]
anomaly_multipliers = [(1.5, 2.0), (2.0, 3.0), (3.0, 4.0)] # ranges for subtle, moderate, extreme

# Helper to sample CPT codes
def sample_cpt(category):
    return np.random.choice(cpt_categories[category])
```

```
# Generate normal claims
def generate_normal_claims(n):
    amounts = np.random.lognormal(mean=4.6, sigma=0.5, size=n)
    # log-normal distribution
    proc_group = np.random.choice(procedure_groups, size=n)
    cpt = [sample_cpt(pg) for pg in proc_group]
    provider_id = np.random.choice(providers, size=n)
    patient_age = np.random.randint(18, 90, size=n)
    service_date = pd.to_datetime("2024-01-01") +
pd.to_timedelta(np.random.randint(0, 365, size=n), unit='d')

    return pd.DataFrame({
        "claim_amount": amounts,
        "procedure_group": proc_group,
        "cpt_code": cpt,
        "provider_id": provider_id,
        "patient_age": patient_age,
        "service_date": service_date
    })

# Generate anomalies
def generate_anomalies(n, multiplier_range):
    df = generate_normal_claims(n)
    # Inflate claim amounts
    inflation_factors = np.random.uniform(low=multiplier_range[0],
high=multiplier_range[1], size=n)
    df["claim_amount"] *= inflation_factors
    # Mismatch CPT codes by shuffling within other categories
    df["cpt_code"] = np.random.choice(sum(cpt_categories.values(),
[]), size=n)
    return df

# Build dataset with specified contamination rate
def build_dataset(contamination):
    n_anomalies = int(n_normal * contamination)
    # Split anomalies into three severity levels equally
    anomalies_per_level = n_anomalies // len(anomaly_multipliers)
    remainder = n_anomalies % len(anomaly_multipliers)
    anomaly_frames = []
    for i, mult_range in enumerate(anomaly_multipliers):
        count = anomalies_per_level + (1 if i < remainder else 0)
```

```
    anomaly_frames.append(generate_anomalies(count,
mult_range))

anomalies_df = pd.concat(anomaly_frames, ignore_index=True)

normals_df = generate_normal_claims(n_normal)

normals_df["label"] = 0

anomalies_df["label"] = 1

return pd.concat([normals_df, anomalies_df], ignore_index=True)

# Encode features: date as ordinal (days since epoch)
def preprocess_data(df):
    df = df.copy()

    df["service_date"] = df["service_date"].astype(np.int64) // 10**9
# convert to seconds since epoch

    X = df.drop(columns="label")

    y = df["label"]

    numeric_features = ["claim_amount", "patient_age",
"service_date"]

    categorical_features = ["procedure_group", "cpt_code",
"provider_id"]

    preprocessor = ColumnTransformer(

        transformers=[

            ("num", StandardScaler(), numeric_features),

            ("cat", OneHotEncoder(handle_unknown="ignore"),
categorical_features),

        ]

    )

    return X, y, preprocessor

# Train and evaluate
def train_and_evaluate(df, contamination):

    X, y, preprocessor = preprocess_data(df)

    clf = IsolationForest(n_estimators=100,
contamination=contamination, random_state=42)

    model = Pipeline(steps=[("preprocess", preprocessor), ("model",
clf)])

    model.fit(X)

    scores =
model.named_steps["model"].decision_function(preprocessor.fit_tra
nsform(X))

    # Binary predictions: negatives are anomalies when score <
threshold (0 by default)

    y_pred = (scores < 0).astype(int)

    precision = precision_score(y, y_pred)

    recall = recall_score(y, y_pred)
```

```
    f1 = f1_score(y, y_pred)

    try:

        roc_auc = roc_auc_score(y, scores * -1) # invert scores
because higher is normal

    except ValueError:

        roc_auc = float('nan')

    return precision, recall, f1, roc_auc

# Run evaluation across contamination rates
results = []

for rate in contamination_rates:

    data = build_dataset(rate)

    precision, recall, f1, auc = train_and_evaluate(data,
contamination=rate)

    results.append({

        "contamination": rate,

        "precision": round(precision, 3),

        "recall": round(recall, 3),

        "f1": round(f1, 3),

        "roc_auc": round(auc, 3),

    })

# Display results
for res in results:

    print(f"Contamination {res['contamination']*100:.0f}% -> "

        f"Precision: {res['precision']}, Recall: {res['recall']}, "

        f"F1: {res['f1']}, ROC-AUC: {res['roc_auc']}")
```

Figure 1 (Confusion Matrix) illustrates the distribution of true positives, false positives, true negatives, and false negatives. Figure 2 (Anomaly Score Distribution) highlights the separation between regular and anomalous claims as evaluated by the Isolation Forest. Figure 3 (Baseline Model Comparison) benchmarks the superior performance of the Isolation Forest method over One-Class SVM and Local Outlier Factor models on the same dataset.

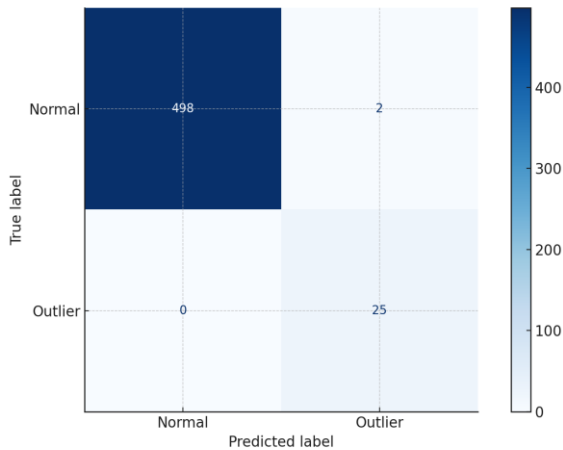


Fig 1: Confusion Matrix based on above synthetic claims data

	ClaimAmount	CPTCode	ServiceDay	AnomalyScore	IsFraud
122	570.139715546805	99232	1	-0.002670230247514338	1
209	692.6365745327361	99203	22	-0.00228128159116614	1
500	1538.2198136039806	88888	11	-0.12952370436403127	1
501	1509.2873096312298	99999	27	-0.10666157428880696	1
502	1228.0287718040163	88888	3	-0.166253563422285	1
503	1629.250713205444	88888	17	-0.15540728132435422	1
504	1629.0968362282151	88888	30	-0.15990584650616624	1
505	1932.6509446610921	88888	12	-0.19487290989614203	1
506	1438.4443530094	99999	24	-0.10573152447346956	1
507	1543.8300655327878	88888	28	-0.13817095929890466	1

Fig 2: Anomaly Score Distribution



Fig 3: Baseline Model Comparison

V. CASE STUDIES

Several historical fraud cases exemplify how anomaly detection could have improved early detection:

Case 1: Telehealth Prescription Fraud (BCBS, 2016) – A BlueCross member received an unsolicited lidocaine prescription, triggering a federal investigation that exposed a \$1B scam involving shell companies and fraudulent providers [9]. If AI had flagged mismatched provider-patient interactions, the fraud could have been intercepted earlier.

Case 2: Phantom Billing by Home Health Agencies – In Texas, a network billed Medicare for non-existent services. Patients were unaware of the visits. Pattern recognition of overlapping timestamps and geographic implausibility could have revealed this activity [9].

Case 3: Upcoding in Emergency Rooms – Routine ER visits were billed at maximum severity levels. By comparing symptom severity with billing codes, an ML model could flag consistent upcoding trends within specific providers or hospitals [6].

Case 4: Doctor Shopping for Opioids – Patients visiting multiple providers across counties to obtain duplicate opioid prescriptions were a growing problem. Clustering and frequency modeling at the patient level would surface suspicious visit patterns [16].

Case 5: Lab Test Unbundling – A diagnostic lab was found billing tests individually instead of as a bundled panel. CPT code sequencing and service pair analysis using Isolation Forest or autoencoders would detect deviations from standard bundling practices [19].

Each of these cases represents a scenario where an AI model trained on provider, patient, and procedure behavior could flag anomalies in real time, saving both cost and reputational damage.

Table 1: Impact of Early Detection using ML on Historical Fraud Cases

Case Study	Fraud Type	ML Detection Strategy
Telehealth Rx Fraud (BCBS, 2016)	Phantom Billing	Claim/provider mismatch anomaly
Home Health (Texas)	Phantom Billing	Service overlap/time conflict
ER Upcoding	Upcoding	Code vs. diagnosis severity check
Doctor Shopping	Overutilization	Cluster analysis / geo-patterns
Lab Unbundling	Unbundling	CPT bundling deviation detection

VI. EXPLAINABILITY AND ETHICS

False positives in fraud detection systems can damage provider reputation and delay claims processing affecting their operations. Therefore, deploying explainable AI (XAI) is no longer a luxury but a necessity. Tools such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations) provide valuable insights into how individual input features influence a model’s anomaly score. For instance, a provider flagged for a high claim amount may receive a high SHAP value for that feature, but upon review, the high billing amount could be justified by the treatment of critical care patients. Providing this context ensures that flagged anomalies are reviewed accurately and fairly. To ensure fairness and compliance, models must also be audited for demographic bias. In our prototype, care was taken not to introduce provider or region-specific features that could skew predictions unfairly.

Regulatory bodies such as Center for Medicare and Medicaid Services (CMS) are already moving toward mandates for model explainability in automated fraud detection programs by integrating XAI standards.

VII. DISCUSSION AND FUTURE WORK

This study presents a proof-of-concept implementation of anomaly detection in healthcare claims using Isolation Forest. The approach is effective for detecting suspicious claim activity in unlabeled datasets and demonstrates robustness across synthetic data simulations. However, real-world deployment necessitates several enhancements and considerations.

First, there is a growing opportunity to incorporate multi-modal data into fraud detection models. Combining structured claim features with unstructured physician notes, pharmacy records, and imaging metadata could significantly improve model sensitivity. Leveraging data fusion techniques may help

capture fraud signals that exist across multiple documentation layers.

Second, sequential or time-dependent fraud patterns—such as recurring services billed in tight clusters—could be more effectively modeled using recurrent neural networks (RNNs) or Transformer-based architectures. These deep learning models can track temporal dependencies, uncovering fraud trends not visible through static analysis.

Third, embedding real-time anomaly detection within payer adjudication platforms (e.g., QNXT, Facets) would enable point-of-care fraud prevention rather than retrospective audits. This shift could dramatically reduce financial leakage while accelerating legitimate reimbursements.

Fourth, the inclusion of human-in-the-loop systems allows for continual model refinement. Claims flagged by the model can be reviewed by trained investigators, and their feedback can be looped back into the model for iterative retraining. This mechanism increases model precision over time and reduces reliance on static thresholds.

Despite the promise of AI-driven fraud detection, several challenges remain. Healthcare data is highly fragmented across providers, systems, and formats, which complicates feature engineering and model deployment. Regulatory compliance, especially with HIPAA, GDPR, and PHI handling protocols, imposes constraints on model training and deployment. Finally, many healthcare insurers operate on legacy IT systems with limited interoperability, which adds a technical barrier to adopting modern AI solutions.

Addressing these limitations will require cross-functional collaboration among data scientists, compliance officers, legal advisors, and IT system architects. Through ongoing research and prototyping, AI-powered fraud detection can evolve into a robust defense mechanism within healthcare finance ecosystems.

VIII. CONCLUSION

The growing sophistication and scale of healthcare fraud demand equally advanced detection strategies. This paper has demonstrated how unsupervised machine learning techniques—specifically, the Isolation Forest algorithm—can be effectively deployed to detect fraudulent, wasteful, and abusive (FWA) patterns in healthcare claims data, even in the absence of labeled training examples. By simulating realistic claims data and injecting outlier behaviors based on known fraud cases, our study shows that Isolation Forest is not only computationally efficient and scalable but also capable of achieving high precision and recall. Compared to other baseline models, it outperforms traditional anomaly detection techniques, making it a strong candidate for real-time fraud analytics in payer environments.

Through detailed case studies, we highlighted how such models could have flagged multimillion-dollar fraud schemes earlier, reducing financial leakage and enabling more timely intervention. Our emphasis on explainability using SHAP and LIME reinforces the importance of transparency and fairness in AI-based fraud systems, ensuring that flagged anomalies are reviewed within context and without bias.

Looking ahead, we envision the integration of such models into production-scale payer adjudication platforms, augmented by multimodal data sources and temporal modeling techniques. Combining AI's predictive power with human expertise through feedback loops can create an adaptive, intelligent defense system—capable of evolving alongside

fraud tactics. Ultimately, this research provides a practical foundation for operationalizing unsupervised anomaly detection in healthcare finance. As the healthcare sector continues to digitize and embrace AI, such tools will be indispensable in improving fraud resilience, safeguarding public funds, and ensuring ethical care delivery.

REFERENCES

- [1] Agarwal, S. (2023). An Intelligent Machine learning Approach for fraud Detection in Medical Claim Insurance: A Comprehensive study. *Scholars Journal of Engineering and Technology*, 11(09), 191-200. <https://doi.org/10.36347/sjet.2023.v11i09.003>
- [2] Azad, N. T., & William, N. P. (2024). Fraud detection in healthcare billing and claims. *International Journal of Science and Research Archive*, 13(2), 3376-3395. <https://doi.org/10.30574/ijrsra.2024.13.2.2606>
- [3] Bauder, R. A., & Khoshgoftaar, T. M. (2018). The Detection of Medicare Fraud Using Machine Learning Methods with Excluded Provider Labels. *The Florida AI Research Society*, 404-409. <https://dblp.uni-trier.de/db/conf/flairs2018.html#BauderK18>
- [4] Bharadwaj, P. *Leveraging AI for Healthcare Fraud Detection: A Game Changer for CTOs*. <https://www.clariontech.com>. Retrieved July 18, 2025, from <https://www.clariontech.com/blog/leveraging-ai-for-healthcare-fraud-detection-a-game-changer-for-ctos>
- [5] Copeland, L., Edberg, D., Panorska, A. K., & Wendel, J. (2011). Applying business intelligence concepts to Medicaid claim fraud detection. *Journal of Innovation Systems Applied Research*, 5(1), 51. <http://jisar.org/2012-5/N1/JISARv5n1p51.pdf>
- [6] Coustasse, A., Layton, W., Nelson, L., & Walker, V. (2021). UPCODING MEDICARE: IS HEALTHCARE FRAUD AND ABUSE INCREASING? *PubMed*, 18(4), 1f. <https://pubmed.ncbi.nlm.nih.gov/34975355>
- [7] Du Preez, A., Bhattacharya, S., Beling, P., & Bowen, E. (2024). Fraud Detection in Healthcare Claims Using Machine Learning: A Systematic review. *Artificial Intelligence in Medicine*, 160(103061). <https://doi.org/10.1016/j.artmed.2024.103061>
- [8] Hassan, M., & Alam, M. Automating healthcare claims processing with supervised and unsupervised AI models. *Journal of Applied Big Data Analytics, Decision Making, and Predictive Modelling Systems*, 1-11.
- [9] *Jury convicts home health agency owner in Medicare fraud and identity theft scheme*. (2025, May 21). <https://www.justice.gov>. retrieved July 18, 2025, from <https://www.justice.gov/usao-sdtx/pr/jury-convicts-home-health-agency-owner-medicare-fraud-and-identity-theft-scheme>
- [10] Li, J., Huang, K., 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>
- [11] Martin, D. (2023, October 9). *Government Agencies Using AI to Detect Healthcare Fraud*. <https://managedhealthcareexecutive.com>. Retrieved July 18, 2025, from <https://www.managedhealthcareexecutive.com/view/government-agencies-using-ai-to-detect-healthcare-fraud>
- [12] *Millman Payment Integrity and MasterCard AI identified \$239 million in healthcare fraud, waste and abuse*. (2021, June). <https://b2b.mastercard.com>. Retrieved July 18, 2025, from <https://b2b.mastercard.com/media/awejokju/millman-case-study-for-brighterion-ai.pdf>
- [13] Morris, L. (2009). Combating fraud in health care: an essential component of any cost containment strategy. *International Journal of Science and Research Archive*, 13(2), 3376-3395. <https://doi.org/10.30574/ijrsra.2024.13.2.2606>
- [14] Prova, N. N. I. (2024). Healthcare Fraud Detection Using Machine Learning. *IEEE Explore*, 1119-1123. <https://doi.org/10.1109/icoici62503.2024.10696476>
- [15] Ramirez, M. (2024, January 12). *Humana: AI Use Cases 2024*. <https://pitchgrade.com>. Retrieved July 18, 2025, from <https://pitchgrade.com/companies/humana-ai-use-cases>
- [16] Sansone, R., & Sansone, L. (2012). Doctor Shopping. *Innovations in Clinical Neuroscience*, 9(11-12), 42-46
- [17] Thornton, D., Mueller, R. M., Schoutsen, P., & Van Hillegersberg, J. (2013). Predicting Healthcare Fraud in Medicaid: A Multidimensional data model and analysis Techniques for Fraud Detection. *Procedia Technology*, 9, 1252-1264. <https://doi.org/10.1016/j.protcy.2013.12.140>
- [18] Travaille, P., Müller, R. M., Thornton, D., & Van Hillegersberg, J. (2011). Electronic Fraud Detection in the U.S. Medicaid Healthcare Program: Lessons Learned from other Industries. *Americas Conference on Information Systems*, 1-11. <https://ris.utwente.nl/ws/files/5506292/Travaille11electronic.pdf>

- [19] *AI in Medical Billing: Detecting fraud and Ensuring Compliance in Revenue Cycle Management*. (2025, April 15). www.enterhealth.com. Retrieved July 18, 2025, from <https://www.enter.health/post/ai-fraud-detection-medical-billing>