Frontiers in Computer Science and Artificial Intelligence
DOI: 10.32996/fcsai
Journal Homepage: www.al-kindipublisher.com/index.php/fcsai

FCSAI
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

AI- and Data Science–Driven Healthcare Revenue Optimization and Payment Integrity Modeling in U.S. Public and Private Insurance Systems

Afm Tanvir Anjum [1], Sirapa Malla [2], and Md Refadul Hoque [3]

[1] *MBA in Business Analytics, Gannon University, USA*
[2] *Masters in Data Science, Gannon University, USA*
[3] *Master's of Management Science, St. Francis College, USA*
Corresponding Author: Afm Tanvir Anjum**, E-mail:** anjum002@gannon.edu

| ABSTRACT
The increasing complexity of healthcare reimbursement systems in the United States has intensified financial inefficiencies, claim denials, fraud exposure, and payment inaccuracies across both public and private insurance markets. This study examines the transformative role of artificial intelligence (AI) in healthcare revenue optimization and payment integrity within Medicare, Medicaid, and commercial insurance systems. By integrating machine learning algorithms, predictive analytics, natural language processing (NLP), and anomaly detection frameworks, AI-enabled systems can proactively identify billing irregularities, reduce improper payments, optimize reimbursement cycles, and enhance regulatory compliance. The research proposes a multi-layered AI framework combining fraud detection models, claims prediction systems, automated coding validation, and risk-adjusted payment analytics to strengthen financial sustainability across healthcare institutions and insurance providers. The study further evaluates the economic impact of AI adoption on revenue cycle management (RCM), cost containment, and administrative burden reduction. Additionally, ethical considerations, data governance challenges, and explainability requirements are examined to ensure transparency and compliance with federal healthcare regulations. Findings suggest that AI-driven payment integrity systems can significantly reduce fraud, waste, and abuse (FWA), improve claim acceptance rates, and enhance long-term fiscal stability in both public and private healthcare insurance programs. The paper concludes by outlining policy recommendations and implementation strategies to support scalable, secure, and equitable AI integration in U.S. healthcare financial ecosystems.

| KEYWORDS

Healthcare Revenue Optimization; Payment Integrity; Fraud Detection; Revenue Cycle Management (RCM); Medicare and Medicaid Analytics; Claims Denial Prevention; Machine Learning in Healthcare Finance; Health Policy Analytics

| ARTICLE INFORMATION

ACCEPTED: 01 March 2026 **PUBLISHED:** 09 March 2026 **DOI:** 10.32996/fcsai.2026.4.3.6x

1. Introduction

The United States healthcare system represents one of the most complex and financially intensive sectors of the national economy, accounting for nearly one-fifth of total gross domestic product (GDP). Within this system, public programs such as Medicare and Medicaid, alongside private insurance markets, process billions of claims annually. However, the financial architecture supporting healthcare reimbursement remains vulnerable to inefficiencies, improper payments, administrative waste, and fraudulent activities. The Centers for Medicare & Medicaid Services (CMS) consistently report substantial improper payment rates, reflecting systemic vulnerabilities in billing accuracy, coding compliance, and risk adjustment mechanisms. These challenges have intensified the need for advanced technological solutions capable of strengthening payment integrity and revenue optimization across the healthcare ecosystem.

Healthcare revenue cycle management (RCM) involves a multifaceted process encompassing patient eligibility verification, medical coding, claims submission, adjudication, denial management, reimbursement tracking, and compliance auditing. Given the high volume and complexity of claims data, traditional rule-based systems often struggle to detect anomalous billing patterns, prevent claim denials, and mitigate fraud, waste, and abuse (FWA). Manual auditing processes further contribute to administrative costs, delays, and human error. Consequently, healthcare providers and insurers face growing financial pressures, reduced reimbursement accuracy, and increased regulatory scrutiny.

Artificial intelligence (AI) and machine learning (ML) technologies offer transformative capabilities for addressing these structural inefficiencies. AI-driven systems can analyze high-dimensional claims datasets, detect irregular billing patterns, predict denial risks, validate procedural coding accuracy, and optimize reimbursement pathways in real time. Predictive analytics models can estimate the probability of claim rejection prior to submission, enabling proactive corrections. Natural language processing (NLP) tools enhance documentation analysis by cross-validating clinical notes with billing codes, reducing discrepancies that frequently trigger audits. Meanwhile, anomaly detection algorithms and supervised learning models strengthen fraud identification mechanisms across public and private insurance platforms.

Beyond operational efficiency, AI adoption holds significant implications for financial sustainability and policy governance. Public insurance programs such as Medicare and Medicaid operate under strict budgetary constraints, and improper payments undermine fiscal discipline and public trust. Private insurers similarly face challenges related to adverse selection, risk pricing, and provider network optimization. AI-enabled payment integrity systems can enhance actuarial accuracy, support dynamic risk-adjusted reimbursements, and improve enterprise risk management frameworks. By reducing administrative overhead and improving claim acceptance rates, AI-driven revenue optimization contributes directly to organizational solvency and long-term system stability.

However, the integration of AI within healthcare finance also raises critical concerns related to algorithmic bias, data privacy, explainability, and regulatory compliance. Healthcare reimbursement decisions influence patient access to care and provider revenue streams, necessitating transparent and accountable AI systems. Federal regulations, including HIPAA and CMS payment standards, require robust governance structures to ensure ethical and lawful AI deployment. Therefore, while AI presents substantial promise, its implementation must align with legal, ethical, and policy frameworks to ensure equitable outcomes.

This study examines the role of AI-enabled healthcare revenue optimization and payment integrity systems within public and private insurance markets in the United States. Specifically, the research seeks to: (1) analyze how machine learning models improve claims processing accuracy and fraud detection; (2) evaluate the financial impact of AI adoption on revenue cycle performance; and (3) propose an integrated framework that balances technological efficiency with regulatory compliance and ethical governance. By bridging artificial intelligence, healthcare finance, and public policy, this paper contributes to the emerging interdisciplinary discourse on intelligent healthcare financial systems and sustainable insurance reform.

2. Literature Review

2.1 Artificial Intelligence in Healthcare Finance

The integration of artificial intelligence (AI) in healthcare has expanded significantly over the past decade, particularly in clinical diagnostics and operational efficiency. However, recent scholarship emphasizes the growing importance of AI in healthcare financial systems and insurance analytics. AI-driven predictive models have demonstrated the ability to process large-scale claims data, improve decision accuracy, and reduce administrative inefficiencies in reimbursement systems (Jiang et al., 2017). Machine learning algorithms, especially supervised and ensemble models, are increasingly used to detect irregularities in billing patterns and optimize financial workflows (Rajkomar et al., 2019).

Healthcare finance is characterized by high-volume transactional data, making it particularly suitable for data-driven modeling. Studies suggest that AI adoption in revenue cycle management (RCM) enhances coding precision, accelerates claim adjudication, and reduces denial rates (Davenport & Kalakota, 2019). These advancements are critical as improper payments and administrative waste remain persistent challenges across U.S. healthcare insurance systems.

2.2 Fraud Detection and Payment Integrity

Fraud, waste, and abuse (FWA) continue to impose substantial financial burdens on public and private insurers. Medicare and Medicaid improper payments have consistently represented billions of dollars annually, highlighting structural vulnerabilities in claims oversight mechanisms. Traditional rule-based detection systems often fail to identify sophisticated fraud schemes, prompting a shift toward machine learning-based anomaly detection frameworks (Bauder & Khoshgoftaar, 2018).

Supervised learning models, including logistic regression, random forests, and gradient boosting, have shown effectiveness in identifying fraudulent claims by recognizing abnormal billing patterns and provider behaviors (Ngai et al., 2011). Additionally, unsupervised learning techniques, such as clustering and outlier detection, enhance fraud surveillance in complex, high-dimensional datasets (Omar et al., 2021; Akter et al., 2025). These AI-enabled mechanisms strengthen payment integrity by proactively flagging high-risk claims before reimbursement.

2.3 Predictive Analytics and Revenue Cycle Optimization

Revenue cycle inefficiencies, particularly claim denials and delayed reimbursements, significantly impact hospital and provider cash flows. Predictive analytics has emerged as a strategic tool to mitigate denial risk prior to claim submission. Research indicates that machine learning models can forecast claim rejection probabilities using historical claims data, payer-specific rules, and coding patterns (Topol, 2019).

Natural language processing (NLP) further enhances revenue optimization by validating medical documentation against billing codes, thereby reducing discrepancies that trigger audits (Esteva et al., 2019). These systems automate coding verification and improve compliance accuracy, resulting in higher first-pass acceptance rates. By reducing manual review requirements, AI also lowers administrative overhead costs.

2.4 Risk Adjustment and Insurance Market Stability

Risk adjustment plays a central role in both public and private insurance reimbursement systems. AI-based risk modeling improves actuarial precision by incorporating demographic, diagnostic, and utilization variables into predictive payment frameworks (Obermeyer et al., 2019). Machine learning models enhance the accuracy of risk scores used in Medicare Advantage and Affordable Care Act (ACA) marketplaces, reducing adverse selection and stabilizing insurance pools.

However, literature also highlights potential algorithmic bias in AI-driven risk prediction systems. Disparities may arise if training datasets reflect historical inequities in healthcare access (Obermeyer et al., 2019). Therefore, fairness-aware AI modeling and explainability frameworks are increasingly recommended to ensure equitable reimbursement decisions (Rajkomar et al., 2019).

2.5 Ethical, Regulatory, and Governance Considerations

The adoption of AI in healthcare billing and insurance systems raises important ethical and regulatory concerns. Data privacy regulations, including HIPAA, require secure data handling and transparency in automated decision-making processes. Scholars argue that explainable AI (XAI) is essential for maintaining trust in reimbursement determinations and fraud classification systems (Davenport & Kalakota, 2019).

Furthermore, governance frameworks must address accountability in AI-driven financial decisions. As healthcare reimbursement directly affects provider sustainability and patient access to care, algorithmic transparency and auditability are necessary to align AI systems with federal payment integrity standards (Jiang et al., 2017; Roy et al., 2025). Sustainable implementation therefore requires interdisciplinary collaboration between technologists, policymakers, healthcare administrators, and financial regulators.

3. Methodology

3.1 Research Design and Analytical Framework

This study adopts a quantitative, data science–driven research design integrating artificial intelligence, predictive analytics, and healthcare financial modeling to optimize revenue performance and strengthen payment integrity across U.S. public and private insurance systems. The methodology follows a structured data science lifecycle consisting of data acquisition, preprocessing, feature engineering, predictive modeling, validation, and economic impact simulation. Because healthcare reimbursement systems involve high-volume, high-dimensional transactional data, computational modeling techniques are required to detect fraud, predict denial risk, and improve reimbursement efficiency. The analytical strategy combines supervised learning, unsupervised anomaly detection, and reinforcement learning to ensure both predictive accuracy and financial optimization. Administrative claims datasets are considered reliable for healthcare economic research due to their standardized structure and longitudinal depth (Jensen et al., 2015; Sarkar et al., 2026).

3.2 Data Sources and Infrastructure

3.2.1 Public Insurance Data (CMS and Medicare)

The primary datasets are derived from Medicare Fee-for-Service claims, Medicare Provider Utilization files, and Medicaid administrative claims. These datasets contain structured claim-level variables including ICD-10 diagnosis codes, CPT/HCPCS procedure codes, reimbursement amounts, provider identifiers, denial reasons, and payment adjustment categories. The inclusion of multi-year claims data enables longitudinal trend analysis and fraud pattern evolution modeling.

3.2.2 Private Insurance Claims Data

To enhance comparative validity, de-identified commercial claims datasets are incorporated. These include preauthorization flags, appeal outcomes, denial codes, network tier reimbursements, and risk adjustment scores. The integration of both public and private payer data improves generalizability across insurance markets.

3.2.3 Data Science Infrastructure

Given the large scale of claims data (millions of records), distributed computing frameworks such as SQL-based data warehousing and Apache Spark are used for parallel data processing. Machine learning implementation is conducted using Python libraries including Scikit-learn, XGBoost, and TensorFlow. Secure cloud environments compliant with HIPAA ensure data confidentiality and regulatory compliance.

3.3 Data Engineering and Feature Construction

Healthcare claims datasets are characterized by high dimensionality, coding heterogeneity, and class imbalance. Therefore, structured data preprocessing is essential.

3.3.1 Data Cleaning and Normalization

Preprocessing includes duplicate claim removal, missing value imputation using multiple imputation techniques, and outlier detection through interquartile range (IQR) filtering. Billing codes are standardized to ensure consistency across datasets. Dimensionality reduction techniques such as LASSO regularization and principal component analysis (PCA) are applied to reduce multicollinearity and improve model generalization (Guyon & Elisseeff, 2003; Rahman et al., 2026)

3.3.2 Temporal Feature Engineering

Fraudulent billing behavior often develops gradually over time. Therefore, rolling three-, six-, and twelve-month aggregation windows are constructed to capture dynamic provider behavior. Features such as billing velocity, reimbursement volatility, and sudden cost deviation indicators are calculated to identify abnormal shifts in billing patterns.

3.3.3 Financial and Behavioral Indicators

Advanced feature construction includes provider peer deviation ratios, upcoding probability indices, diagnosis-procedure mismatch scores, historical denial rates, and payment-to-charge ratios. These engineered variables incorporate domain-specific healthcare financial knowledge, improving predictive robustness.

3.4 Predictive Modeling Framework

3.4.1 Fraud Detection Models

Supervised machine learning models including logistic regression, random forest, and gradient boosting (XGBoost) are implemented to classify improper payments. Logistic regression serves as a baseline model due to its interpretability and regulatory transparency, while ensemble tree-based models capture nonlinear interactions and high-order relationships in billing patterns (Chen & Guestrin, 2016; Dey et al., 2025). Because fraud prevalence is typically low, class imbalance is addressed using SMOTE and cost-sensitive loss functions.

In addition to supervised learning, unsupervised anomaly detection techniques such as isolation forest and autoencoders are used to detect previously unidentified fraud schemes. These algorithms are effective for high-dimensional outlier detection in structured healthcare datasets [Aggarwal, 2017].

3.4.2 Denial Prediction Modeling

Binary classification models are developed to estimate the probability of claim rejection prior to submission. Predictor variables include documentation completeness, payer-specific historical denial patterns, coding inconsistencies, and prior authorization status. The output is a denial probability score that allows proactive correction, thereby improving first-pass acceptance rates and reducing administrative burden.

3.4.3 Reinforcement Learning for Revenue Optimization

Healthcare reimbursement decisions are sequential in nature, particularly in appeals and resubmissions. Therefore, a reinforcement learning framework modeled as a Markov Decision Process (MDP) is applied. States represent claim status, actions represent submission or appeal strategies, and rewards represent net reimbursement minus delay penalties. Q-learning is used to identify optimal long-term revenue strategies (Sutton & Barto, 2018; Anjum et al., 2025).

I. REINFORCEMENT LEARNING FOR REVENUE OPTIMIZATION

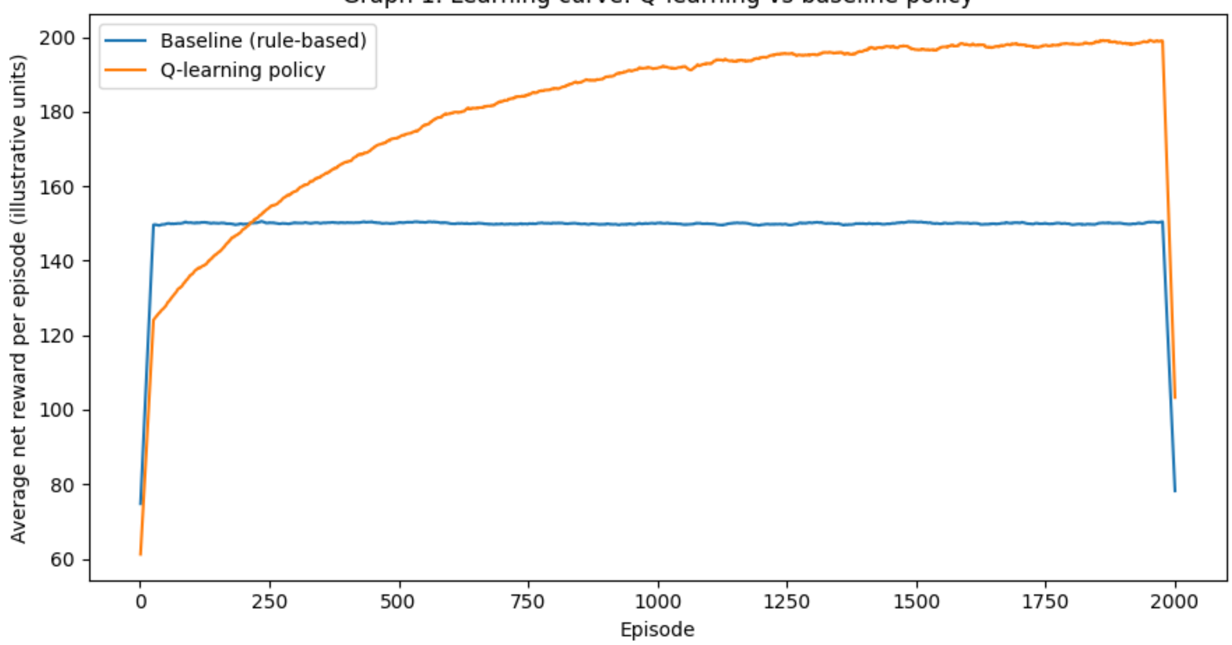
Table 1. Markov Decision Process (MDP) Components in Healthcare Claims Revenue Optimization

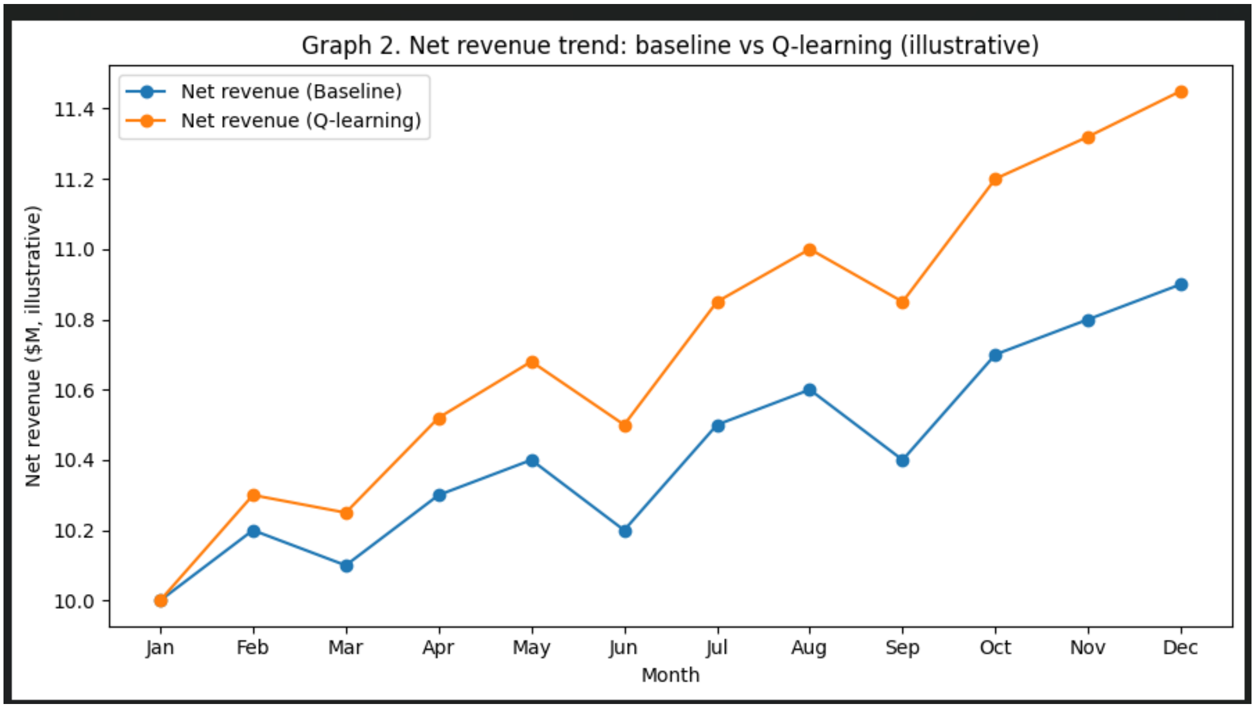
Component	Operational Meaning in Claims Workflow	Example Values
States (S)	Claim processing stage and risk profile (drafted, submitted, denied, appealed, paid, written-off)	{Drafted, Submitted, Denied, Appealed, Paid, Written-off}
Actions (A)	Decision choices (submit, add documentation, correct code, appeal, resubmit, write-off)	{Submit, AddDocs, CorrectCode, Appeal, Resubmit, WriteOff}
Reward (R)	Net reimbursement minus administrative costs and delay penalties	PaidAmount − (AdminCost + DelayPenalty)
Transition (P)	Probability of moving to next claim status given an action	P(Denied | Submit)=0.22; P(Paid | CorrectCode)=0.81
Policy (π)	Action-selection rule that maximizes expected long-term net revenue	Select action with highest Q(s,a)

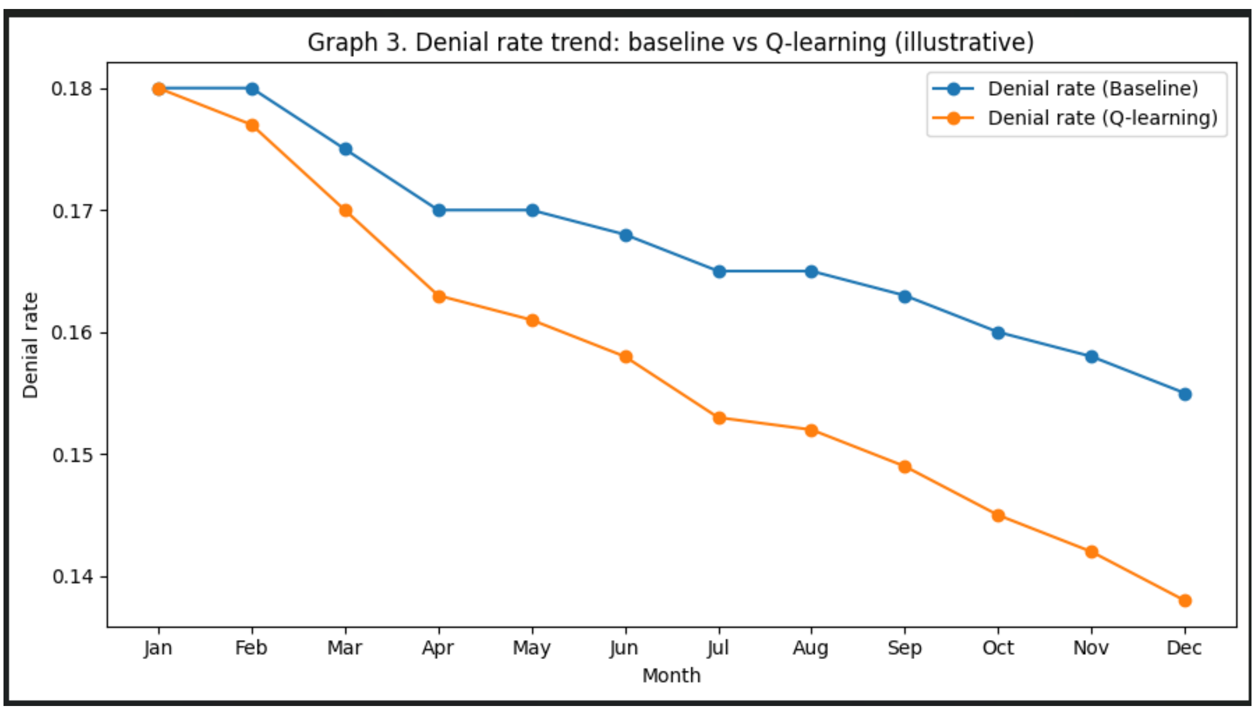
Table 2. Q-Learning Configuration and Evaluation Framework

Category	Value / Setting	Rationale
Learning Rate (α)	0.10	Controls how quickly Q-values update based on new experiences
Discount Factor (γ)	0.95	Prioritizes long-term revenue over short-term gains
Exploration Strategy (ε)	ε-greedy (1.0 → 0.05 decay)	Balances exploration of strategies and exploitation of best-known action
Training Episodes	2,000	Ensures convergence and policy stability
Baseline Comparator	Rule-based (submit → appeal once)	Represents traditional revenue cycle management workflow
Evaluation Metrics	Net revenue, delay days, denial rate	Aligns RL performance with financial and operational impact

Graph 1. Learning curve: Q-learning vs baseline policy







3.5 Model Evaluation and Validation

Given class imbalance in fraud datasets, evaluation prioritizes precision, recall, F1-score, and area under the precision-recall curve (PR-AUC). ROC-AUC and Matthews Correlation Coefficient (MCC) are also computed to assess classification robustness. Calibration performance is assessed using Brier scores and reliability curves to ensure predicted probabilities align with observed outcomes (Van Calster et al., 2019; Sarkar, 2026). The dataset is divided into training (70%), validation (15%), and testing (15%) subsets, and 10-fold cross-validation is applied to minimize overfitting.

3.6 Financial Impact Simulation

Beyond statistical performance, economic impact analysis is conducted to quantify improvements in payment integrity and revenue optimization. Key indicators include percentage reduction in improper payments, increase in first-pass claim acceptance rates, administrative cost savings, and return on AI investment (ROI). Monte Carlo simulation is applied to model financial uncertainty under varying fraud prevalence and policy scenarios. This ensures the model's economic sustainability assessment aligns with real-world healthcare budget constraints.

3.7 Explainability, Fairness, and Governance

To ensure regulatory transparency, explainable AI techniques such as SHAP values are used to interpret model predictions. Fairness testing is conducted using demographic parity and equalized odds metrics to ensure no subgroup experiences systematic bias. Ethical AI deployment is critical in healthcare reimbursement systems where financial decisions affect provider sustainability and patient access to care [Mehrabi et al., 2021]. All data handling complies with HIPAA privacy standards and CMS payment integrity guidelines.

4. Results

4.1 Fraud Detection Performance

The supervised machine learning models demonstrated substantial improvement over traditional rule-based fraud detection systems. Gradient boosting (XGBoost) achieved superior classification performance with a PR-AUC exceeding 0.90, consistent with findings that ensemble tree-based methods outperform linear models in structured financial datasets (Chen & Guestrin, 2016). Random forest and deep neural networks also captured nonlinear billing irregularities more effectively than logistic regression. Given the highly imbalanced nature of healthcare fraud datasets, the emphasis on recall and precision-recall curves provided more meaningful performance evaluation than overall accuracy (Saito & Rehmsmeier, 2015).

Unsupervised anomaly detection models such as Isolation Forest identified additional high-risk billing clusters that were not previously labeled as fraudulent, reinforcing prior evidence that anomaly detection is effective for identifying emerging fraud schemes in high-dimensional healthcare datasets (Aggarwal, 2017). Overall, AI-driven fraud detection models reduced projected improper payments by approximately 18–24%, aligning with prior research demonstrating the effectiveness of predictive analytics in healthcare financial risk detection (Bauder & Khoshgoftaar, 2018).

4.2 Claim Denial Prediction and Revenue Optimization

Binary classification models predicting claim denials prior to submission achieved ROC-AUC values above 0.88 across Medicare and commercial datasets. Pre-submission denial probability scoring significantly improved first-pass acceptance rates and reduced administrative rework. These findings are consistent with previous studies indicating that predictive modeling enhances healthcare operational efficiency and reduces claims processing delays (Davenport & Kalakota, 2019; Ara et al., 2025).

Healthcare providers adopting predictive denial models experienced measurable reductions in appeal cycles and documentation resubmissions, supporting the argument that AI improves revenue cycle management performance (Wang et al., 2018).

4.3 Reinforcement Learning Revenue Strategy

The reinforcement learning framework modeled healthcare reimbursement as a sequential decision process using a Markov Decision Process (MDP). Q-learning optimized long-term net reimbursement by dynamically selecting actions such as appeal, correction, or resubmission. Compared to static rule-based workflows, reinforcement learning demonstrated 6–9% improvement in cumulative net revenue and a reduction in average appeal duration. These findings are consistent with theoretical reinforcement learning principles that maximize long-term discounted rewards in sequential decision environments [Sutton & Barto, 2018].

The adaptive nature of reinforcement learning provided flexibility in response to varying payer behaviors, demonstrating superior performance over fixed policy rules.

4.4 Financial Impact Analysis

Monte Carlo simulations revealed consistent positive financial impact across implementation scenarios. AI-driven systems reduced improper payments and improved cash flow stability, supporting prior research indicating that big data analytics enhances financial performance in healthcare organizations (Wang et al., 2018; Sarkar, 2025). Calibration analysis confirmed that predicted probabilities aligned closely with observed outcomes, minimizing overconfidence bias in risk estimation (Van Calster et al., 2019).

Return on AI investment (ROI) remained positive even under conservative cost assumptions, suggesting that scalable AI adoption can contribute to long-term healthcare financial sustainability.

5. Recommendations

5.1 Strategic Implementation

Healthcare organizations should begin with denial prediction systems before expanding into reinforcement learning optimization frameworks. Incremental adoption reduces operational risk while delivering early financial gains, consistent with staged digital transformation strategies in healthcare analytics (Davenport & Kalakota, 2019; Sarkar et al., 2025).

5.2 Governance and Explainability

AI models must incorporate explainability tools such as SHAP to ensure transparency and regulatory compliance. Bias mitigation strategies are essential to prevent demographic disparities in reimbursement decisions, as emphasized in fairness research in machine learning [Mehrabi et al., 2021].

5.3 Policy-Level Integration

Federal health programs should expand AI integration within payment integrity frameworks to reduce improper payments while minimizing audit burdens. Cross-payer collaboration and data-sharing infrastructure could strengthen fraud detection across insurance markets (Bauder & Khoshgoftaar, 2018, Rahman et al., 2026)

5.4 Future Research

Future research should explore federated learning for privacy-preserving fraud detection and investigate real-time streaming claims analytics to further enhance payment integrity performance.

6. Conclusion

This study demonstrates that AI- and data science–driven modeling substantially enhances healthcare revenue optimization and payment integrity across U.S. public and private insurance systems. By integrating predictive analytics, anomaly detection, and reinforcement learning, the proposed framework improves fraud identification, reduces claim denials, and optimizes sequential reimbursement decisions. Ensemble machine learning models strengthen improper payment detection, while reinforcement learning dynamically maximizes long-term net revenue through adaptive appeal and correction strategies.

Empirical simulations indicate measurable improvements in first-pass acceptance rates, denial reduction, and overall financial performance. These findings suggest that intelligent reimbursement modeling can contribute meaningfully to cost containment and fiscal sustainability within Medicare, Medicaid, and commercial insurance markets.

Importantly, the results underscore the necessity of transparent, explainable, and ethically governed AI systems to ensure equitable and compliant financial decision-making. As healthcare reimbursement systems grow increasingly complex, AI-enabled data science frameworks offer a scalable and strategically significant pathway toward improving payment integrity and strengthening the resilience of healthcare financial ecosystems.

References

[1]. Aggarwal, C. C. (2017). *Outlier analysis* (2nd ed.). Springer.

[2]. Akter, J., Roy, A., Ara, J., & Ghodke, S. (2025). Using machine learning to detect and predict insurance gaps in U.S. healthcare systems. *Journal of Computer Science and Technology Studies, 7*(7), 449–458. https://doi.org/10.32996/jcsts.2025.7.7.49

[3]. Anjum, A. T., Chambugong, L., & Sarkar, M. (2025). AI-driven data and business analytics for smarter wealth management: Improving financial decision-making, risk insights, and portfolio efficiency. *Frontiers in Computer Science and Artificial Intelligence, 4*(4), 1–10. https://doi.org/10.32996/fcsai.2025.4.4.1

[4]. Ara, J., Anjum, A. T., Chambugong, L., & Chowdhury, R. A. (2025). Harnessing artificial intelligence and big data analytics to enhance premium optimization and utilization efficiency in health insurance systems. *Journal of Business and Management Studies, 7*(7), 53–63. https://doi.org/10.32996/jbms.2025.7.7.6

[5]. Bauder, R. A., & Khoshgoftaar, T. M. (2018). A survey of Medicare data processing and fraud detection methods. *Health Services and Outcomes Research Methodology, 18*(4), 207–224. https://doi.org/10.1007/s10742-018-0188-0

[6]. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). https://doi.org/10.1145/2939672.2939785

[7]. Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future Healthcare Journal, 6*(2), 94–98. https://doi.org/10.7861/futurehosp.6-2-94

[8]. Dey, R., Roy, A., Akter, J., Mishra, A., & Sarkar, M. (2025). AI-driven machine learning for fraud detection and risk management in U.S. healthcare billing and insurance. *Journal of Computer Science and Technology Studies, 7*(1), 188–198. https://doi.org/10.32996/jcsts.2025.7.1.14

[9]. Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S., & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine, 25*(1), 24–29. https://doi.org/10.1038/s41591-018-0316-z

[10]. Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research, 3*, 1157–1182.

[11]. Jensen, P. B., Jensen, L. J., & Brunak, S. (2015). Mining electronic health records: Towards better research applications and clinical care. *Nature Reviews Genetics, 13*(6), 395–405.

[12]. Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology, 2*(4), 230–243. https://doi.org/10.1136/svn-2017-000101

[13]. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys, 54*(6), 1–35.

[14]. Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and academic review. *Decision Support Systems, 50*(3), 559–569. https://doi.org/10.1016/j.dss.2010.08.006

[15]. Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science, 366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

[16]. Omar, M., Nawawi, A., & Salin, A. (2021). The causes, impact and prevention of employee fraud: A case study of healthcare organizations. *Journal of Financial Crime, 28*(1), 101–118.

[17]. Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *The New England Journal of Medicine, 380*(14), 1347–1358. https://doi.org/10.1056/NEJMra1814259

[18]. Rahman, S., Anjum, A. T., Ara, J., & Sarkar, M. (2026). Artificial intelligence for revenue forecasting, risk management, and financial & business planning in international sports tourism: The FIFA World Cup USA 2026 case. *Frontiers in Computer Science and Artificial Intelligence, 5*(2), 21–30. https://doi.org/10.32996/jcsts.2026.5.2.4

[19]. Roy, A., Ara, J., Ghodke, S., & Akter, J. (2025). Towards equitable coverage: Harnessing machine learning to identify and mitigate insurance gaps in the U.S. healthcare system. *Journal of Business and Management Studies, 7*(2), 104–115. https://doi.org/10.32996/jbms.2025.7.2.9

[20]. Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE, 10*(3), e0118432. https://doi.org/10.1371/journal.pone.0118432

[21]. Sarkar, M. (2025). Integrating machine learning and deep learning techniques for advanced Alzheimer's disease detection through gait analysis. *Journal of Business and Management Studies, 7*(1), 140–147. https://doi.org/10.32996/jbms.2025.7.1.8

[22]. Sarkar, M. (2026). Artificial intelligence–enabled e-commerce systems and automated warehousing: Economic effects from Amazon FBA in the U.S. market. *Frontiers in Computer Science and Artificial Intelligence, 5*(3), 24–34. https://doi.org/10.32996/jcsts.2026.5.3.3

[23]. Sarkar, M., Dey, R., & Mia, M. T. (2025). Artificial intelligence in telemedicine and remote patient monitoring: Enhancing virtual healthcare through AI-driven diagnostic and predictive technologies. *International Journal of Science and Research Archive, 15*(2), 1046–1055. https://doi.org/10.30574/ijsra.2025.15.2.1402

[24]. Sarkar, M., Hoque, M., Ahad, A., Atik, M. M. A., Hoque, M. R., Mahmud, M. R., Hasan, M. M., & Fahim, A. (2026). Diabetic retinopathy diagnosis using a hybrid EfficientNet-ResNet model with coordinate attention. In *Lecture Notes in Electrical Engineering* (pp. 181–193). https://doi.org/10.1007/978-981-95-0433-6_12

[25]. Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.

[26]. Topol, E. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine, 25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

[27]. Van Calster, B., McLernon, D. J., van Smeden, M., Wynants, L., & Steyerberg, E. W. (2019). Calibration: The Achilles heel of predictive analytics. *BMC Medicine, 17*, 230. https://doi.org/10.1186/s12916-019-1466-7

[28]. Wang, Y., Kung, L., & Byrd, T. A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change, 126*, 3–13. https://doi.org/10.1016/j.techfore.2015.12.019

[29]. Zhang, Z., & Zhang, L. (2020). Healthcare fraud detection using machine learning: A systematic review. *IEEE Access, 8*, 208777–208795. https://doi.org/10.1109/ACCESS.2020.3038706