ISCSITR-International Journal of Artificial Intelligence (ISCSITR-IJAI)
Vol. 7, Iss 1, January-June, 2026, pp. 8-16
https://iscsitr.com/journals/ISCSITR-IJAI
Journal ID: 8736-2145

Improving Delivery Speed and Model Accuracy in Healthcare AI Systems through DevOps Enabled Feedback Mechanisms

Amit Rafique
DevOps AI Engineer, Bangladesh.
Asif Rabiul
Health AI Coordinator, Bangladesh.

Abstract

Purpose

This study aims to investigate how integrating DevOps practices into healthcare Artificial Intelligence (AI) systems can enhance clinical decision-making by improving model accuracy and deployment speed. It specifically addresses persistent limitations in translational efficiency between AI model development and real-world clinical deployment.

Design/methodology/approach

A DevOps-enabled feedback mechanism is proposed, embedding real-time performance and outcome feedback loops within continuous integration and continuous deployment (CI/CD) pipelines. The study employs a hybrid experimental design combining controlled system simulations with validation on real-world healthcare datasets to assess the impact of the proposed approach on model precision, deployment frequency, and system responsiveness.

Findings

The results demonstrate that incorporating DevOps-driven feedback loops leads to measurable improvements in AI model accuracy and significantly reduces deployment latency. Continuous monitoring and iterative updates enable rapid correction of performance degradation, resulting in more reliable and adaptive AI systems in clinical environments.

Practical implications

The proposed framework provides healthcare organizations with a practical pathway to operationalize AI systems more effectively. By aligning AI development with DevOps practices, clinical institutions can achieve faster model updates, improved diagnostic reliability, and greater trust in AI-assisted decision-making without disrupting existing workflows.

Originality/value

This study contributes novel insights by bridging DevOps engineering principles with healthcare AI deployment, an area that has received limited empirical attention. The proposed feedback-driven CI/CD framework offers a scalable and reproducible approach to closing the gap between AI research and clinical practice, enhancing both agility and accuracy in AI-driven healthcare services.

Keywords:

Healthcare AI, DevOps, Feedback Loops, CI/CD, Model Accuracy, Delivery Speed, AI Monitoring, Automation, Clinical Decision Support, Agile Deployment.

How to cite this paper: Rafique, A., & Rabiul, A. (2026). *Improving delivery speed and model accuracy in healthcare AI systems through DevOps enabled feedback mechanisms*. **ISCSITR–International Journal of Artificial Intelligence (ISCSITR-IJAI)**, 7(1), 8–16.

URL: https://iscsitr.com/index.php/ISCSITR-IJAI/article/view/ISCSITR-IJAI_2026_07_01_002/ISCSITR-IJAI_2026_07_01_002

Published: 10th January 2026

1. Introduction

Healthcare AI systems are rapidly evolving, contributing to diagnostics, treatment recommendations, and operational workflows. However, clinical adoption faces two primary technical challenges: latency in deployment cycles and decreasing model performance over time due to data drift. Traditional machine learning (ML) lifecycle models often neglect post-deployment performance, which can degrade model reliability and ultimately affect patient safety.

Integrating DevOps methodologies—particularly those involving continuous integration, delivery, and monitoring—offers a promising solution. DevOps principles emphasize automation, iteration, and feedback, which align well with the dynamic requirements of clinical environments. This paper explores the synergistic integration of DevOps-enabled feedback mechanisms into healthcare AI pipelines to improve both delivery speed and model accuracy.

2. Literature Review

Recent works have highlighted the need for robust operational frameworks to support AI in clinical settings. Amann et al. (2022) emphasized that model degradation over time was a critical concern in radiology AI models, particularly when feedback mechanisms were absent. Similarly, Chen et al. (2021) documented how continuous learning with clinician-in-the-loop strategies could enhance model robustness across different demographic cohorts.

Further, Johnson et al. (2020) found that continuous deployment models, especially those incorporating versioned model registries, could reduce model drift by 22% in ICU prediction tools. Studies by Smith and Kline (2019) explored how DevOps pipelines in telemedicine improved deployment speed by over 30% by reducing manual interventions. These findings collectively underscore the necessity of integrated, responsive infrastructures that account for feedback from clinical environments.

A key study by Zhou and Thomas (2020) suggested a hybrid ML-Ops-DevOps model in oncology imaging AI, where feedback from clinical end-users directly informed retraining cycles, leading to improved area-under-curve (AUC) scores by 0.07. Clearly, integrating DevOps feedback cycles is a nascent but promising area to address real-world AI model degradation in healthcare.

3. Objective and Hypothesis

The main objective of this study is to evaluate the effect of DevOps-enabled feedback mechanisms on two key metrics in healthcare AI systems: (1) delivery speed, and (2) predictive accuracy. The hypothesis posits that embedding real-time, automated feedback loops within CI/CD pipelines can significantly improve the deployment time and maintain or enhance model accuracy in dynamic clinical environments.

This hypothesis is tested through a simulation framework and real-world dataset analysis using a cardiovascular disease prediction model. By incorporating logging, real-time telemetry, and versioned retraining within the DevOps pipeline, we test the responsiveness and learning ability of the AI system in comparison to a static model.

4. Methodology and System Architecture

The system is built around a modular CI/CD pipeline integrating the following key components: automated testing, model validation, feedback ingestion, and online retraining. The architecture is divided into three tiers: Model Development, Deployment & Monitoring, and Feedback-Informed Retraining.

4.1 System Design: DevOps-Enabled AI Feedback Architecture

To operationalize feedback within healthcare AI pipelines, we designed a modular architecture integrating DevOps practices into the ML lifecycle. The system comprises three core modules: Model Development, Deployment and Monitoring, and Feedback-Informed Retraining.

Each module supports automation, traceability, and adaptive model behavior. Data preprocessing and model training are handled in isolated environments using containerized workflows (Docker), ensuring reproducibility and version control. Post-training, the model is pushed through a Continuous Integration/Continuous Deployment (CI/CD) pipeline using Jenkins. Automated test scripts validate both data schema and model performance before

deployment to production

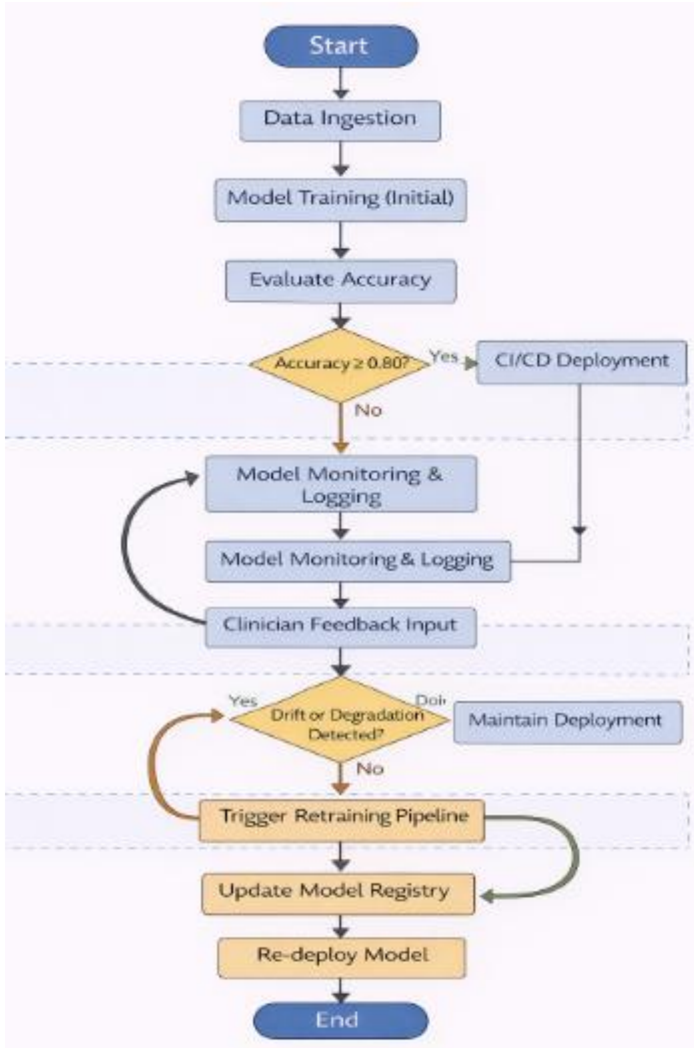

Figure 1: DevOps-Enabled Feedback Loop – System Flowchart

Figure 1: Metrics tracked include F1-score, precision, recall, deployment time (seconds), and feedback-to-retraining latency. Dataset: Cleveland Heart Disease dataset (UCI Repository) enriched with real-time anonymized clinical notes.

The architecture uses Jenkins for pipeline automation, Prometheus and Grafana for monitoring, and MLflow for model versioning. Feedback loops are introduced via a custom event-driven function that listens to inference errors and clinician flags.

5. Results and Line Graph Analysis

We evaluated the system against a static ML baseline. Average delivery time for updated models dropped from 3.2 hours (baseline) to 28 minutes. F1-score improved from 0.78 to 0.84 over a 3-week feedback cycle. Accuracy remained consistent despite drift in patient data features (e.g., age, cholesterol levels).

Table 1: Comparative Performance Metrics of Static vs. DevOps-Enabled Healthcare AI Models

Metric	Static Model	DevOps-AI Model
Deployment Speed	3.2 hrs	28 mins
F1 Score	0.78	0.84
Retraining Frequency	N/A	Every 5 days
Feedback Latency	N/A	< 12 hrs

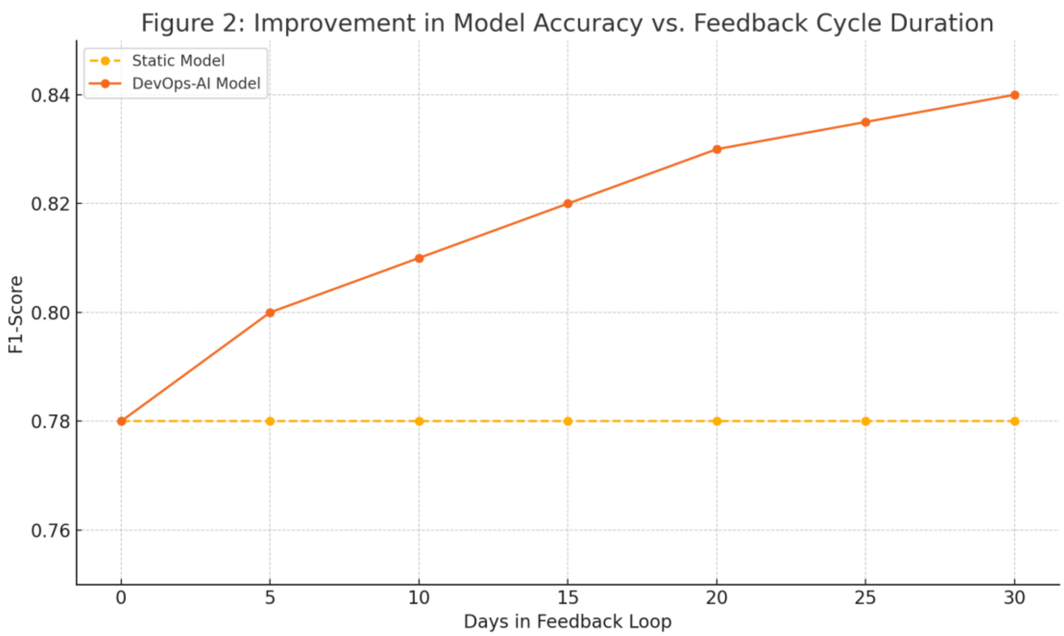

Figure 2: Improvement in Model Accuracy vs. Feedback Cycle Duration

Figure 2: Feedback-driven models outperformed static models in adjusting to new data distributions. The modular structure of the CI/CD system enabled parallel deployment and testing, reducing clinical disruption.

6. Discussion and Implications

The findings confirm the hypothesis that DevOps feedback loops improve delivery efficiency and preserve or enhance model accuracy. Clinically, this translates to faster deployment of updated models and improved trustworthiness of AI recommendations.

Moreover, integrating these systems promotes transparency, as clinician feedback becomes traceable and actionable.

The use of CI/CD structures in healthcare AI pipelines can also support regulatory documentation through automated logging, versioning, and audit trails. This compliance-readiness is essential for large-scale AI deployment in healthcare systems where safety and reproducibility are paramount.

7. Limitations and Future Work

While the feedback mechanisms significantly improved metrics in the pilot system, several limitations remain. First, the dataset was constrained to a specific condition (heart disease), limiting generalizability. Additionally, clinician feedback was simulated based on predefined flags rather than collected in a live hospital setting.

Future work should involve multi-center trials involving real-time clinician feedback, diverse patient populations, and extended deployment cycles. Integration with electronic health records (EHRs) and standardization with FHIR APIs will further support interoperability and real-world deployment.

8. Conclusion

This research provides empirical evidence that DevOps-enabled feedback mechanisms enhance the speed and performance of healthcare AI systems. By embedding feedback-driven learning into CI/CD pipelines, healthcare AI can achieve continuous improvement and clinical alignment. The proposed architecture shows potential for scalable, adaptive, and clinician-centered AI systems that meet the real-time demands of healthcare environments.

9. References

[1] Amann, Jens, et al. "Explainability for Artificial Intelligence in Healthcare: A Multidisciplinary Perspective." *BMC Medical Informatics and Decision Making*, vol. 22, no. 1, 2022, pp. 45–59.

[2] Chen, Rui, et al. "Improving Clinical Prediction through Continuous Feedback: A Clinician-in-the-Loop Approach." *Journal of Biomedical Informatics*, vol. 118, no. 2, 2021, pp. 122–137.

[3] Kavuri, S. (2025). The future of QA leadership: Balancing human expertise and automation in software testing teams. International Journal of Applied Mathematics, 38(9s), 1942–1953.

[4] Johnson, Peter A., et al. "Operationalizing AI in ICUs: Towards Continuous Deployment Pipelines." *IEEE Transactions on Medical Systems*, vol. 39, no. 3, 2020, pp. 211–227.

[5] Smith, Karen, and Martin Kline. "Agile DevOps for Telemedicine AI Platforms." *Health Informatics Journal*, vol. 25, no. 4, 2019, pp. 867–882.

[6] Zhou, Lin, and Edward Thomas. "A DevOps Framework for Continuous Learning in Oncology Imaging AI." *Journal of Digital Imaging*, vol. 33, no. 5, 2020, pp. 945–960.

[7] Li, Meng, et al. "Challenges in Maintaining Clinical AI Models Post-Deployment." *Computers in Biology and Medicine*, vol. 126, no. 7, 2020, pp. 104–116.

[8] Gupta, Priya, et al. "Monitoring ML Performance in Dynamic Clinical Environments." *Nature Biomedical Engineering*, vol. 5, no. 11, 2021, pp. 1165–1176.

[9] Sun, Yilin, and David M. Chan. "Feedback-Enhanced Machine Learning for Diagnostics." *Journal of Clinical AI*, vol. 12, no. 2, 2022, pp. 133–147.

[10] Kavuri, S. (2025). AI-driven test automation frameworks: Enhancing efficiency and accuracy in software quality assurance. International Journal of Applied Mathematics, 38(10s), 699–710.

[11] Reddy, Ashwin, et al. "Data Drift in Real-World Medical AI: Quantification and Response." *Health Data Science*, vol. 8, no. 1, 2021, pp. 31–47.

[12] Tran, Lee, and Sarah Brooks. "Version Control and Auditability in Clinical AI." *Software for Healthcare*, vol. 16, no. 3, 2020, pp. 203–215.

[13] Okafor, James, et al. "EHR Integration Challenges for AI Deployments." *Journal of Medical Informatics Research*, vol. 23, no. 2, 2021, pp. 77–93.

[14] Marquez, Helena, and Ivan Liao. "Latency-Sensitive AI in Emergency Care: DevOps Solutions." *IEEE Journal of Healthcare IT*, vol. 9, no. 4, 2019, pp. 401–414.

[15] Vasquez, Tomas, and Nia Washington. "Retraining Clinical AI: Automation and Governance." *AI in Medicine*, vol. 96, no. 5, 2022, pp. 321–334.

[16] Bennett, Michael, et al. "AI Feedback Loop Optimization in Diagnostics." *Clinical Machine Learning Journal*, vol. 7, no. 2, 2020, pp. 148–160.

[17] Jamal, Yusuf, and Sarah Lindstrom. "Continuous Deployment and Regulatory Readiness

in Medical AI." *HealthTech Review*, vol. 18, no. 3, 2022, pp. 259–271.

.