Challenge: Healthcare Predictive Analytics

Project: Planner: MUGABE GERARD

Title of the Project: Predicting Hospital Readmission with Automated Cloud Deployment

1) Project overview and challenge Addressed

Project overview

In Rwanda and across many parts of Africa, healthcare systems face serious resource constraints. Clinics and hospitals often operate under heavy patient loads, with not only limited access to diagnostic tools, staff, and real-time decision support systems but also are under immense pressure to manage limited healthcare resources efficiently while improving patient outcomes. As a result, patients at high risk of chronic illnesses like diabetes, hypertension, or cardiovascular diseases may go undiagnosed until complications arise.

Early risk prediction and timely intervention can significantly reduce both mortality and treatment costs. But achieving this at scale requires not just intelligent analytics , it also requires rapid deployment, testing, and reliability through DevOps best practices.

A significant challenge in hospital systems worldwide is unplanned patient readmission where patients return to the hospital shortly after being discharged, often due to complications that could have been prevented with timely care or monitoring.

Unplanned readmissions are:

> Costly: They increase healthcare spending and strain hospital resources.

> Risky: They expose patients to further complications or infections.

> Preventable: Many readmissions are avoidable with better data-driven decision support.

In resource-constrained settings, the impact of avoidable readmissions is even more severe, affecting national healthcare budgets and contributing to high morbidity rates. Rwanda's Ministry of Health and its partners continue to invest in digital health systems, presenting a perfect opportunity to apply predictive analytics and cloud-based automation.

Challenge addressed

Hospital readmissions are costly and often preventable. Furthermore, in Rwanda and globally, reducing avoidable readmissions is essential for improving patient outcomes and optimizing resource allocation.

This project directly addresses not only the "DevOps" challenge of the DTP Online Hackathon Rwanda 2025 but also addresses the Healthcare Predictive Analytics challenge of the DTP Online Hackathon by building a machine learning system that predicts whether a patient is likely to be readmitted within 30 days of discharge, based on clinical and demographic data.

It tackles the DevOps challenge by implementing a fully automated CI/CD pipeline for cloud deployment, ensuring the model can be updated, tested, and deployed seamlessly.

As a result, this project directly addresses two key challenges from the hackathon:

1. Healthcare predictive analytics

"Predict patient readmission risks or disease outbreaks from hospital or public health datasets."

By developing a machine learning model that predicts whether a patient is at high risk of being readmitted within 30 days after hospital discharge. The model is trained on structured health data, such as patient demographics, medical history, and hospitalization details.

2. DevOps for CI/CD

"Design a CI/CD pipeline for automated, cloud-based deployment."

By integrating the model into a FastAPI-based web service, which is containerized with Docker and automatically deployed using a GitHub Actions CI/CD pipeline to a cloud platform (Render or AWS EC2). This allows for fast, reliable updates and reproducible deployments.

2) Problem Statement

In Rwandan healthcare facilities, patient readmissions within 30 days of discharge often indicate gaps in post-discharge care, early detection of complications, or inadequate resource planning. Predicting these risks before they occur could help hospitals prioritize follow-up interventions, reduce healthcare costs, and improve patient outcomes.

Therefore, the key problem addressed in this project is:

'How can we use historical hospital data from Rwanda to predict the likelihood of patient readmission and support proactive healthcare interventions?'

3) Our goal is to:

Design and implement a CI/CD-enabled cloud deployment pipeline for a healthcare analytics solution, specifically a machine learning model that predicts disease risk based on patient data.

4) Dataset

For this project, data is derived or simulated to reflect common healthcare data points available in Rwanda's hospitals, such as data from King Faisal Hospital, CHUK, or district-level facilities.

The dataset contains records of patient admissions, demographics, diagnoses, treatments, and readmission status.

Key features include:

- Patient ID
- Age (years)
- Gender
- Primary diagnosis (e.g., diabetes, cardiovascular disease)
- Comorbidities (e.g., hypertension, HIV)
- Number of previous hospital visits in past 12 months

- ➢ Length of stay (days)
- ➢ Treatment plan (e.g., medication, surgery)
- ➢ Discharge condition (recovered, under care, referred)
- ➢ Readmission status (binary: 0 = No, 1 = Yes)

Example : Dataset (patient readmissions.csv)

Patient ID	Age	Gender	Primary Diagnosis	Comorbidities	Prev Visits	Treatment	Length of stay	Discharge Condition	Readmission Status
1	54	M	Diabetes	Hypertension	3	Medication	5	Recovered	1
2	67	F	Cardiovascular	HIV, Hypertension	1	Surgery	7	Under Care	0
3	40	M	Diabetes	None	0	Medication	4	Recovered	0

5) Technical Approach

Methodology:

Step 1: Data pre-processing:

- ➢ Handle missing values and outliers.
- ➢ Feature engineering (e.g., age groups, comorbidity counts).
- ➢ Data normalization or encoding categorical variables (e.g., one-hot encoding).

Step 2 : Model Selection:

- ➢ Start with baseline models: Logistic Regression, Random Forest.
- ➢ Compare performance with Gradient Boosting (XGBoost/LightGBM).
- ➢ Use cross-validation for model robustness.

Step 3 : Evaluation Metrics:

- ➢ Classification metrics: Accuracy, F1-score, ROC-AUC.
- ➢ Calibration: Ensure predictions are interpretable.

Step 4: Model Interpretability:

- ➢ Use SHAP or LIME to understand which features contribute most to readmission risk.

Step 5 : Deployment (DevOps):

- ➢ Containerize the app using **Docker**.

> ➢ CI/CD pipeline (e.g., **GitHub Actions or GitLab CI**) for automatic testing and cloud deployment.
> ➢ Host predictive API using **FastAPI/Flask + AWS/Azure/GCP**

Technological Stack

- ❖ Backend/ML: Python, Pandas, Scikit-learn, XGBoost, SHAP.
- ❖ Frontend (optional): Streamlit or React.js for a dashboard.
- ❖ DevOps: Docker, GitHub Actions, AWS/GCP.
- ❖ Visualization: Matplotlib, Seaborn, Plotly.

6) Use Case

A hospital discharge system integrates this tool to:

- ✓ Automatically flag high-risk patients before discharge
- ✓ Send alerts to doctors or care coordinators
- ✓ Recommend post-discharge monitoring or home visits

This can reduce readmissions, improve outcomes, and optimize patient flow, especially in public hospitals with high patient volumes.

7) Scalability & Social Good

The solution is designed with scalability in mind:

- Can be trained on other diseases or adapted to other populations
- Integrates easily into electronic health systems and dashboards
- Supports multilingual or mobile frontends for low-resource settings

This project aligns with Rwanda's Vision 2050 for a digital and health-secure future, offering a scalable model that can be adapted globally in similar healthcare ecosystems.

8) Expected Outcomes

Metric	Impact
Readmission Risk Prediction	Identify patients most likely to be readmitted
Model Accuracy (AUC)	85%+ with proper data tuning
Resource Efficiency	Optimize bed usage and reduce overcrowding
CI/CD Pipeline	Fully automated build, test, deploy workflow

9) DevOps Value Addition

With automated CI/CD:

- Model retraining and deployment take minutes, not days

- Error detection and rollback are easier via continuous integration
- Collaborative development is simplified via Git workflows

This ensures that as hospital data evolves, the predictive model can evolve with it automatically and reliably.

10) GitHub Repository Structure
See attached pdf file.

11) CI/CD DevOps Pipeline

CI: On every commit → run model tests, API tests

CD: Auto-deploy Docker container to Render or EC2

Monitoring: Basic logs + error handling

Infrastructure (Optional): Terraform for AWS provisioning

12) README.md

a. Project Overview

(above)

b. Setup / Usage Instructions

See attached pdf

c. Features & Technologies:

- List ML models,
- SHAP interpretability,
- Python,
- FastAPI
- Scikit-learn / XGBoost
- Docker
- GitHub Actions
- Render / EC2 deployment

13) Demo/Test Credentials

No login required; public API endpoint

14) Conclusion

This project delivers a meaningful solution to a real-world healthcare problem by blending machine learning with cloud-based DevOps automation. Through accurate prediction of hospital readmissions and a reliable, continuous deployment workflow, the system improves healthcare outcomes and resource efficiency.

SmartReadmit demonstrates that AI is not just about algorithms , it's about delivering actionable insights to the frontlines of healthcare, where they can save lives. Its modular

architecture and automated deployment pipeline make it suitable not only for Rwanda, but also for any health system facing similar challenges.

In short, this project is a scalable, impactful, and production-ready approach to modern, data-driven healthcare delivery.