Multidisease Patient Readmission Prediction using ML Models

Thota Sahithi, Reethu Gopavarapu, Sathwik Reddy Mareddy, Ganesh Srija Mandpati
Master's in Computer Science and Engineering
University of Missouri Kansas City

Introduction

Hospital readmissions are a major burden on healthcare systems.
This project focuses on predicting **30-day readmission risks** for **diabetic and heart failure patients** using machine learning.
The system integrates an **XGBoost model**, **random Forest, Logistic Regression** a **Flask backend**, and **HTML/CSS frontend** and **javscrip**t with **dual login portals** for doctors and patients.
It displays patient-specific risk scores, supports communication, and helps hospitals forecast bed and staff needs.
All patient and prediction data is stored securely in the cloud for scalability.

Problem Statement

- Hospital readmissions increase treatment cost and workload.
- Chronic diseases like diabetes and heart failure have high recurrence.
- Traditional follow-up systems lack predictive insights.
- Hospitals struggle with real-time resource planning.
- Patients often don't get timely follow-up or risk feedback.
- A data-driven, interactive platform is needed for prediction, planning, and communication.

Objectives

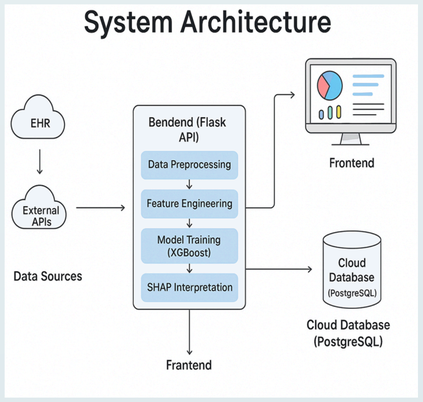
Predict 30-day hospital readmission risk.
Identify significant factors affecting readmission.
Generate interpretable risk scores (Low/Medium/High).
Simulate staffing and bed requirements.
Enable doctor-patient communication and report sharing.
Store and process all data securely on the cloud

Methodology

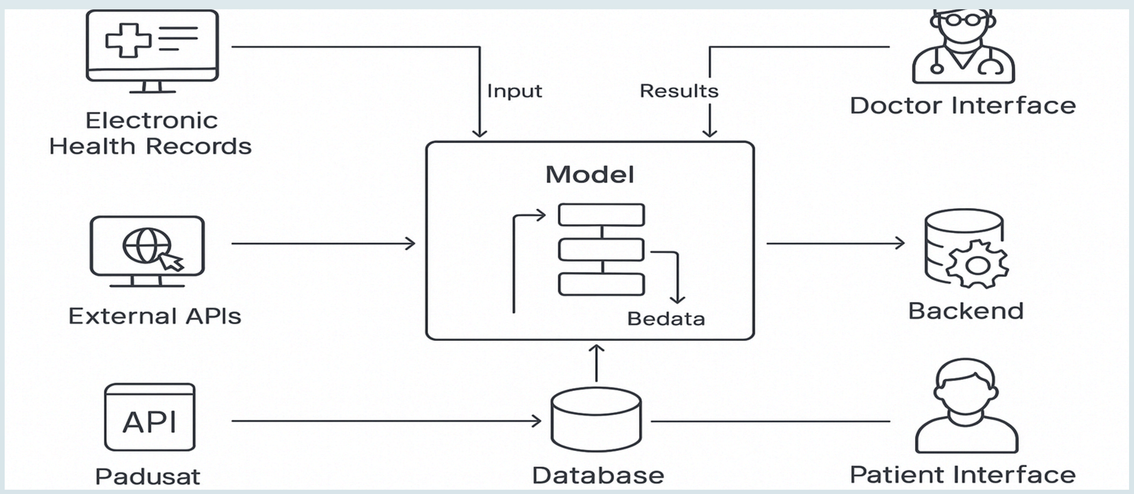
Ø**End-to-End Pipeline**: Starts from **data ingestion** (EHR + external APIs) → ends with **web-based prediction system**.
Ø**ML Model (Random Forest)**: Chosen for its high AUC and interpretability (with SHAP).
Ø**Dual-Login Architecture**: One system for doctors (dashboard, simulation) and one for patients (risk report, communication).
Ø**Cloud-Ready Design**: Deployed on Google Cloud or AWS with PostgreSQL for persistent, secure storage.
Ø**Interoperability**: APIs connect the ML model with the frontend; usable from multiple UIs (React/Streamlit).

System Architecture

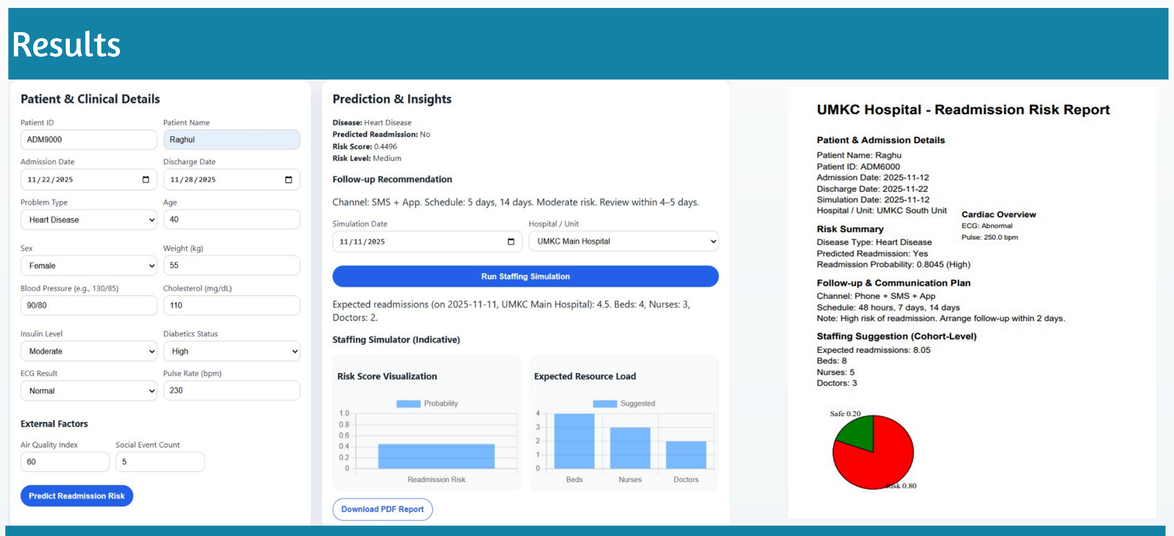
Frontend: HTML,CSS, JavaScript, plotly dashboards for visualization
Backend: Flask for API + business logic
ML Model: Trained scikit-learn/Random Forest models
Database: PostgreSQL for storing patient and prediction data
Cloud Integration: AWS for scalability
Doctor Login:
View patient readmission predictions, bed/staff forecasts
Patient Login:
View personal reports, receive notifications & doctor feedback



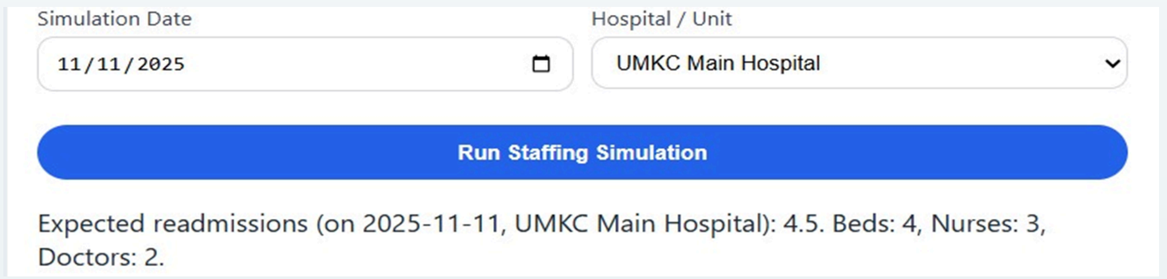
Data Flow Diagram

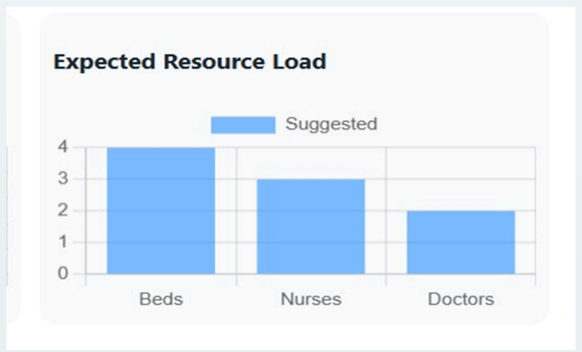


Results



Staff Simulation and visualization



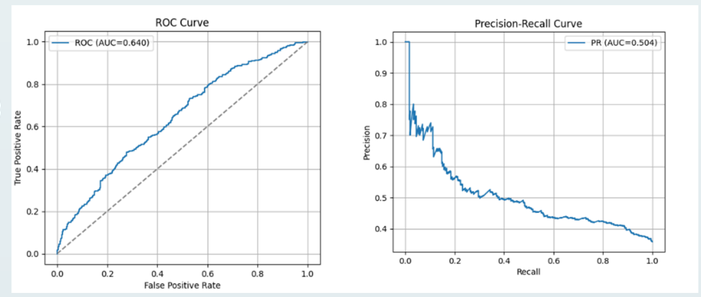


Date	Readmissions	Beds_Required	Nurses_Needed	Doctors_Needed	Support_Staff_Needed
########	3	3	1	1	1
########	6	6	2	1	1
########	3	3	1	1	1
########	3	3	1	1	1
########	2	2	1	1	1
########	2	2	1	1	1
########	4	4	1	1	1
########	1	1	1	1	1
########	2	2	1	1	1
########	5	5	1	1	1
########	3	3	1	1	1
########	5	5	1	1	1
########	4	4	1	1	1
########	3	3	1	1	1
########	4	4	1	1	1
########	8	8	2	1	1
########	6	6	2	1	1
########	5	5	1	1	1
########	3	3	1	1	1
########	1	1	1	1	1
########	5	5	1	1	1
########	0	0	0	0	0
########	4	4	1	1	1

ML Model and Risk Scoring

- Model: **Random Classifier**
- Input: Age, diagnosis, lab values, previous admissions
- Output: Readmission risk (probability 0–1)
- Risk Categories:
- 0.0–0.3: Low
- 0.3–0.7: Medium
- 0.7–1.0: High
- Performance:
  Diabetics
- Accuracy: 0.66
- AUC: 0.66

Heart failure
- Accuracy: 0.64
- AUC: 0.66



Feature	Coefficient
cat__Insulin_Moderate	0.5130139950115348
cat__Diabetics_Moderate	0.3815323793279493
cat__Insulin_High	0.2692355680653884
cat__Diabetics_High	0.2600239831686077
num__Age	0.18913087516768715
num__Cholesterol	0.14228978762886713
cat__Blood Group_AB+	0.12303286705644847
cat__weather_Windy	0.1094865570083328
cat__ECG Result_Abnormal	0.09480507378319951

Predicted_Risk_Score	Predicted_Readmis	RiskGroup
0.590928891	Yes	MODERATE
0.247362829	No	LOW
0.368831595	No	MODERATE
0.378492709	No	MODERATE
0.478002785	No	MODERATE
0.451816779	No	MODERATE
0.23532316	No	LOW
0.486479562	No	MODERATE
0.300389992	No	MODERATE

Conclusion

This project successfully developed a **machine learning–based healthcare readmission prediction system** that integrates clinical data, environmental factors, and cloud technology to improve patient care and hospital efficiency.
By accurately predicting readmission risks for **diabetes and heart failure patients**, the system enables doctors to take preventive actions and optimize staff and bed utilization.
The inclusion of **interactive dashboards and a dual login interface** enhances communication between doctors and patients, ensuring timely follow-ups and better outcomes.
Overall, the project demonstrates how **data-driven insights and automation** can significantly reduce preventable readmissions and support smarter healthcare management.

References

UMKC School of Computing and Engineering. (2025). Advanced Operating Systems and Data Science Research Projects. Kansas City, MO: University of Missouri – Kansas City.
Tahayna, B., et al. (2021). Predicting hospital readmission risk using data analytics: A machine learning approach. BMC Medical Informatics and Decision Making, 21(1), 1–12.

Acknowledgements

I extend my sincere gratitude to the University of Missouri, Kansas City, for providing the research opportunities that have greatly enriched my academic journey. I am especially thankful for Dr.Syed Jawad Husssian Shah for introducing me challenges of the problems and realtime implementation of Data Science through out the project.I am equally thankfull for the Dr. Rafida Zaman for her instrumental role for Guiding me about the cloud to implement in my ML project in advance. I am so proudfull for my research and dedication in th Data Science and the cloud.