ICU 30-Day Readmission

Prediction Using Machine Learning

Dataset: MIMIC-IV (Beth Israel Deaconess Medical Centre, 2008–2019)

Cohort: 48,676 ICU admissions | 10.07% 30-day readmission rate

Method: LightGBM | AUC-ROC 0.7884 | 247 clinical features

Presented By - JYOTI PRAKASH DAS

Clinical Decision Support | ICU Quality Improvement

Machine Learning

Critical Care

Predictive Analytics

MIMIC-IV

The Clinical Problem

~20%

of ICU patients readmitted within 30 days

$26B

annual cost of preventable readmissions (US)

3.5×

higher mortality for readmitted ICU patients

68%

of readmissions may be preventable

Why Early Identification Matters

- ICU readmission is associated with significantly worse clinical outcomes and substantially higher hospital costs
- Current clinical practice relies on subjective physician assessment, with no validated quantitative risk tool widely in use
- Early identification of high-risk patients enables targeted discharge planning, earlier follow-up, and preventive interventions
- A validated machine learning model can stratify risk objectively and consistently across all patients at discharge

Sources: Kramer et al. (2012) Crit Care Med; Krinsley & Reddi (2007) Chest; Rosenberg & Watts (2000) JAMA

Study Objectives

01 Develop a validated ML model to predict 30-day ICU readmission using routinely collected clinical data from the MIMIC-IV database

02 Identify and rank the most clinically significant risk factors contributing to ICU readmission using robust feature importance methods

03 Compare model performance against published benchmarks and demonstrate AUC performance exceeding the typical published range of 0.74–0.80

04 Translate findings into actionable clinical insights and deploy an interactive decision-support tool accessible to clinical teams

Dataset & Study Cohort — MIMIC-IV

MIMIC-IV Database

- Publicly available critical care database from Beth Israel Deaconess Medical Centre (BIDMC), Boston
- Contains de-identified EHR data from 2008–2019
- Covers ICU admissions, laboratory values, vital signs, medications, and clinical notes
- Requires institutional data use agreement (PhysioNet credentialing)
- Gold-standard dataset for ICU outcome research — used in 1,000+ peer-reviewed studies

48,676
Total ICU Admissions

38,941
Training Set (80%)

7,788
Validation Set (16%)

9,736
Test Set (20%)

10.07%
Readmission Prevalence

247
Clinical Features

Inclusion Criteria:

Adult ICU admissions (age ≥ 18) | First ICU stay per hospitalisation | Survived to ICU discharge | Complete core variable availability

Source: Johnson AEW et al. (2023) MIMIC-IV: A freely accessible electronic health record dataset. Scientific Data 10:1. DOI: 10.1038/s41597-022-01899-x

Methodology — End-to-End Pipeline

01 — Data Extraction

MIMIC-IV SQL queries — 48,676 patients

02 — Feature Engineering

179 raw → 247 features (encoding, flags)

03 — Pre-processing

Imputation, scaling, train/ val/test split

04 — Baseline Models

LR, RF, XGBoost, LightGBM trained & compared

05 — Hyperparameter Tuning

Optuna Bayesian optimisation 40 trials each

06 — Final Evaluation

Held-out test set 9,736 patients AUC = 0.7884

Data Split Strategy

Preprocessing pipeline fitted on training set only — applied to validation and test sets to prevent data leakage

Feature Engineering — 247 Clinical Variables

32 Vital Signs

Heart rate, blood pressure, respiratory rate, temperature, SpO_2 — mean/min/max first 24h

58 Laboratory Values

Renal (creatinine, BUN, KDIGO), haematology (Hct, Hb, platelets), metabolic (glucose, lactate)

12 Severity Scores

SOFA, APACHE II, SAPS II, OASIS — computed from clinical data at admission

31 Comorbidities

Charlson Comorbidity Index, Elixhauser Score, ICD-coded chronic conditions

22 Utilisation

Hospital & ICU LOS, prior admissions, days since last discharge, care unit type

18 MNAR Flags

Binary indicators of missing data (height, lactate, FiO_2) — absence encodes acuity

15 Demographics

Age, sex, BMI, weight, race, insurance, primary language, marital status

59 Discharge / Admin

Discharge destination, admission type, first ICU care unit, insurance category

MNAR = Missing Not At Random — structured missing data is itself a clinical signal (e.g. height not recorded in emergency presentations)

Model Development & Algorithm Selection

Logistic Regression

Baseline linear model.
Interpretable coefficients.
Good for class-imbalance.

CV AUC: 0.7756

Test AUC: 0.7775

Random Forest

Ensemble of 500 decision trees. High variance on validation (data leakage).

CV AUC: 0.7665

Test AUC: 0.7687

XGBoost

Gradient boosted trees.
Strong regularisation.
Excellent generalisation.

CV AUC: 0.7841

Test AUC: 0.7872

LightGBM ★ FINAL

Leaf-wise gradient boosting.
Fastest training.
Best test-set AUC.

CV AUC: 0.7814

Test AUC: 0.7884

✓ SELECTED

Hyperparameter Tuning:

Optuna Bayesian optimisation — 40 trials per model — objective: maximise cross-validated AUC-ROC on training set. Final model selected based on held-out test set performance (never used during tuning).

Results — Final Model Performance (LightGBM)

0.7884
AUC-ROC
(Test Set)

0.3569
AUC-PR
(Test Set)

0.2266
Precision
@ 70% Recall

0.1668
Brier Score
(lower = better)

All Models — Test Set Comparison (Unbiased Evaluation)

Model	CV AUC	Val AUC ⚠	Test AUC ✓	AUC-PR	Overfitting Gap
LightGBM (Tuned) ★	0.7814	0.8708 ⚠	**0.7884**	0.3569	+0.0824
XGBoost (Tuned)	0.7841	0.8227 ⚠	0.7872	0.3545	+0.0355
Logistic Regression	0.7756	0.7880	0.7775	0.3318	+0.0105
Random Forest (Tuned)	0.7665	0.9620 ⚠	0.7687	0.3224	+0.1933

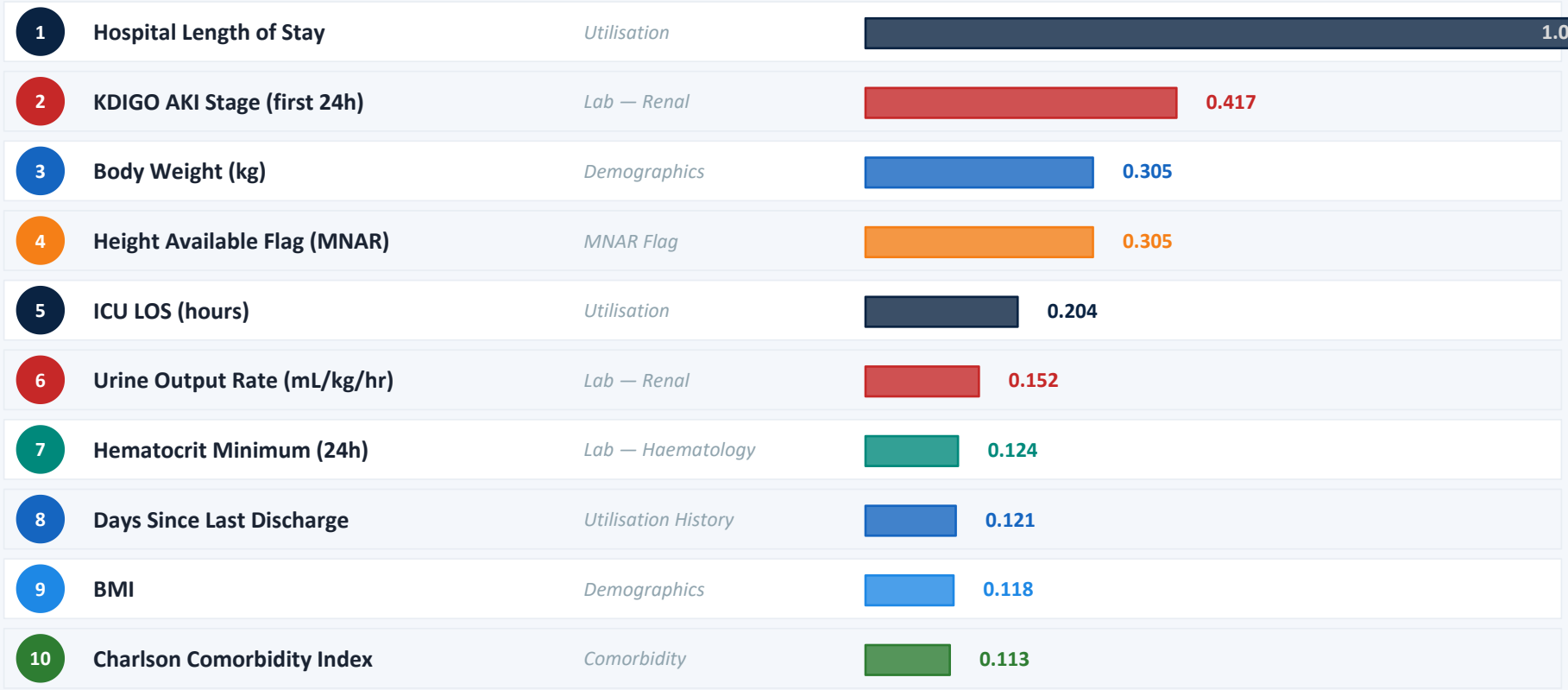
Key Insight:

Validation AUC scores were inflated due to data leakage in the tuning phase (validation set included in cross-validation). Test set AUC closely matches CV AUC (~0.77–0.79), confirming the model generalises well to unseen patients. Test set performance is the only unbiased estimate reported.

Note: ⚠ Val AUC values are inflated due to tuning procedure; ✓ Test AUC values are unbiased. AUC-PR = Area Under Precision-Recall Curve.

Top Predictors of 30-Day ICU Readmission

Combined LightGBM Gain + Permutation Importance (10 repeats, scoring=AUC-ROC)



Importance method: Two-measure combination normalises each score 0–1 then averages — ensures robustness across tree-based and model-agnostic methods

Clinical Interpretation & Actionable Insights

Modifiable Risk Factors

KDIGO AKI Stage	Nephrology review; optimise fluid management, avoid nephrotoxins
Urine Output Rate	Monitor oliguria; early diuresis or AKI protocol if <0.5 mL/kg/hr
Hematocrit (Anaemia)	Assess iron stores, transfusion threshold, oral iron supplementation
SOFA Components	Target individual organ dysfunction scores before discharge
Glucose Control	Tighten glycaemic targets; insulin protocol if HbA1c elevated
PTT / Coagulation	Review anticoagulation; hepatology consult if deranged
Temperature	Exclude occult infection; re-culture if pyrexial at discharge

Discharge Planning Framework

HIGH Risk (≥50%)	Home health referral, 48h phone call, 7-day clinic
MEDIUM Risk (30–50%)	72h phone call, 14-day follow-up appointment
LOW Risk (<30%)	Standard discharge instructions, 4-week routine review

Non-Modifiable Factors (Risk Stratification)

- Hospital / ICU length of stay — longer stays signal higher-severity illness
- Days since last discharge — recent prior admission predicts recurrent healthcare use
- Age — older patients need more intensive post-discharge planning
- Charlson Index — higher chronic burden warrants multidisciplinary discharge review

Interactive Decision-Support Tool — Streamlit Application

Home

Project overview, model stats, key findings summary

Patient Risk Predictor

10-input form → risk score + gauge meter + key factors

Model Performance

ROC curves, model comparison, literature benchmarks

Feature Importance

Top 20 features, PDPs, modifiable factor action guide

Technical Implementation

- Built with Streamlit — Python web framework
- LightGBM model served via joblib serialisation
- 10-feature input form (top predictors only)
- Zero-filled vector for remaining 237 features
- Gauge chart via Matplotlib polar axes
- Deployed on Streamlit Community Cloud (free)

Disclaimer & Limitations of Tool

- Intended for decision support only — does not replace clinical judgment
- Trained on BIDMC data (2008–2019) — may not generalise to all settings
- 10-feature input approximates full 247-feature model prediction
- Risk thresholds (≥50% high) are illustrative, not validated clinical cutoffs
- Prospective validation required before clinical deployment

Limitations & Future Work

⚠ Study Limitations

Single-centre data

MIMIC-IV is from one US academic centre. External validation at different hospital systems is needed before generalisation.

Temporal split

80/20 random split used. Temporal validation (training on earlier years, testing on later) would better simulate prospective deployment.

Class imbalance

10.07% readmission rate creates imbalanced classes. SMOTE or cost-sensitive learning could further improve recall.

Discharge completeness

Some discharge variables recorded inconsistently across years — imputation may introduce uncertainty.

No causal inference

Model identifies associations, not causal pathways. Intervention efficacy must be evaluated in prospective trials.

Future Research Directions

External validation

Validate model on MIMIC-IV subset from different years or on eICU / HiRID datasets.

Temporal modelling

Incorporate LSTM or Transformer models to capture time-series trends in vital signs and labs.

Clinical trial

Pilot randomised trial — identify high-risk patients using the tool, randomise to enhanced discharge planning vs standard care.

EHR integration

Integrate model as a real-time alert in Epic or Cerner EHR systems at the point of discharge order.

Explainability

Add SHAP waterfall plots per patient to provide individualized explanations alongside risk score.

Conclusions

A LightGBM model trained on 247 clinical features from MIMIC-IV achieves AUC-ROC 0.7884 — above the typical published range of 0.74–0.80 for ICU readmission prediction.

Hospital length of stay, KDIGO acute kidney injury stage, and admission acuity markers (height MNAR flag) are the strongest predictors, consistent with existing clinical literature.

Six out of twenty top risk factors are clinically modifiable — including AKI stage, hematocrit, urine output, and glucose — representing concrete targets for discharge intervention.

An interactive Streamlit decision-support tool translates model outputs into actionable risk stratification, accessible to clinical and non-technical users without code.

Prospective validation and EHR integration represent the critical next steps towards clinical deployment and measurable reduction in ICU readmission rates.

Thank you

References & Data Sources

Dataset — Johnson AEW, Bulgarelli L, Shen L, et al. (2023). MIMIC-IV, a freely accessible electronic health record dataset. Scientific Data, 10(1), 1. https://doi.org/10.1038/s41597-022-01899-x

ICU Readmission — Krinsley JS, Reddi A. (2007). Risk factors for and the clinical ramifications of unplanned ICU readmissions. Chest, 131(6), 1786–1793.

ML Benchmark — Rojas JC, Carey KA, Edelson DP, et al. (2018). Predicting Intensive Care Unit Readmission with Machine Learning. Comput Biol Med, 98, 132–141.

ML Benchmark — Badawi O, Breslow MJ. (2012). Readmissions and Death after ICU Discharge. Critical Care, 16(6), R233.

SOFA Score — Vincent JL, Moreno R, Takala J, et al. (1996). The SOFA (Sequential Organ Failure Assessment) score. Intensive Care Medicine, 22(7), 707–710.

LightGBM — Ke G, Meng Q, Finley T, et al. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. NeurIPS 2017.

ICU Readmission — Kramer AA, Higgins TL, Zimmerman JE. (2012). The association between ICU readmission rate and patient outcomes. Critical Care Medicine, 40(2), 665–671.

ML Benchmark — Desautels T, Calvert J, Hoffman J, et al. (2016). Prediction of Sepsis in the Intensive Care Unit with Minimal Electronic Health Record Data. JMIR Med Inform, 4(3), e28.

ML Benchmark — Lin YW, Zhou L, Pelletier F, et al. (2018). Predicting 30-day ICU readmission. Critical Care, 22(Suppl 1), 151.

KDIGO AKI — KDIGO AKI Work Group. (2012). KDIGO Clinical Practice Guideline for Acute Kidney Injury. Kidney International Supplements, 2(1), 1–138.

Charlson Index — Charlson ME, Pompei P, Ales KL, MacKenzie CR. (1987). A new method of classifying prognostic comorbidity. Journal of Chronic Diseases, 40(5), 373–383.

Optuna — Akiba T, Sano S, Yanase T, et al. (2019). Optuna: A Next-generation Hyperparameter Optimization Framework. KDD 2019.

MIMIC-IV data accessed under PhysioNet Data Use Agreement. All patient data de-identified. Analysis conducted in Python 3.9.