

ML Workflow Diagram

ML Workflow for Patient Outcome Prediction

1. Problem Definition

Predict whether a patient will recover based on clinical vitals and treatment.

2. Data Collection (EHRs/Synthetic Dataset)

Load synthetic ICU patient dataset (`patient_outcomes.csv`).

3. Data Preprocessing (cleaning, encoding, scaling)

Analyze outcome distribution, vital statistics, and feature correlations.

4. Exploratory Data Analysis (EDA)

Handle missing values, encode categorical variables, scale features.

5. Feature Engineering

80/20 split with stratified sampling to preserve class distribution.

6. Model Training (Logistic Regression, Random Forest)

Train Logistic Regression, Random Forest, and XGBoost models.

7. Model Evaluation (ROC-AUC, F1-score, Confusion Matrix)

Use Accuracy, F1-score, ROC-AUC, and confusion matrix.

8. Model Interpretation & Recommendation

Identify key predictors and suggest actionable insights.

9. Deployment (optional)