Classification of Hospital Ratings: Low vs High

Authors: Ananya C, Srinidhi N, Ananya A, Amine B, Aaditya N

Introduction

- Hospital performance ratings are crucial for evaluating the quality of care and guiding patient decisions.
- This project focuses on predicting whether a hospital belongs to the low (1-3 stars) or high (4-5 stars) rating category using various performance metrics.
- Accurate predictions can highlight key drivers of healthcare quality and help optimize improvement efforts.

Can we predict a hospital's rating based on its performance data?

Feature Selection & Visualization

Features: 21 columns (20 predictors + 1 target variable).
Target Variable: *Rating* (High or Low hospital rating).
Data Types
4 numerical features (costs).
17 categorical features (ratings, qualitative comparisons).
Categorical Features
Most categorical features (e.g., *Facility.Type*, *Rating.Safety*) have
3 unique values representing comparative ratings (e.g., Above, Same, Below).
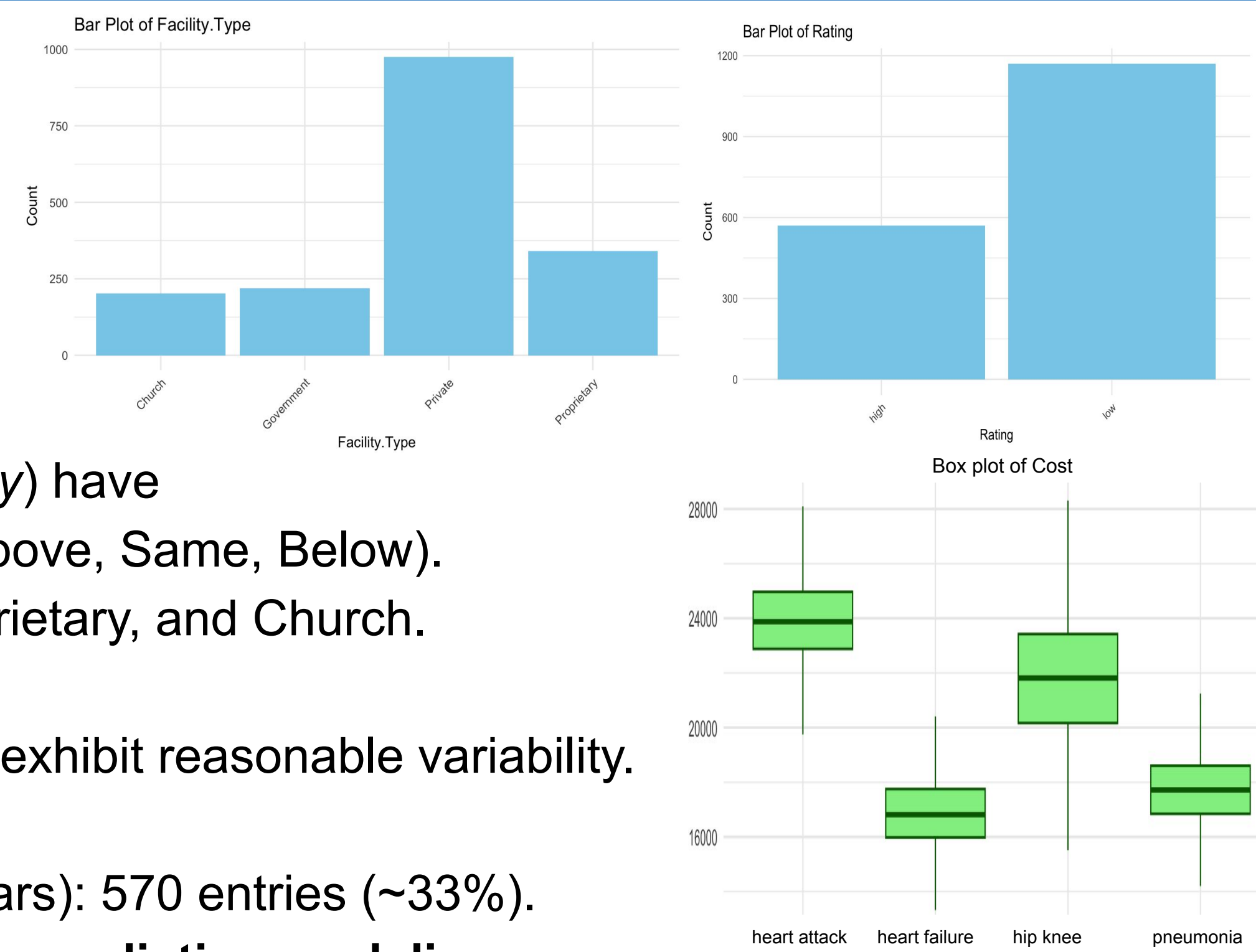*Facility.Type* has 4 categories: Government, Private, Proprietary, and Church.
Numerical Features
Costs (*Procedure.*.Cost*): Represented as integers. They exhibit reasonable variability.
Target Variable Distribution:
Rating: Low (1-3 stars): 1169 entries (~67%), High (4-5 stars): 570 entries (~33%).
Indicates an <u>imbalanced dataset</u>, requiring attention in predictive modeling.

Data

The data for this analysis was sourced from a dataset of 1,739 U.S. hospitals, featuring performance metrics and ratings.

| Predictors (20 variables)	Response
Facility Details: Organization type	Rating: Categorized as High (4-5 stars) or Low (1-3 stars) based on overall hospital performance.
Performance Ratings: Mortality, Safety, Readmission, Patient Experience, Effectiveness, Timeliness, and Imaging	
Cost Metrics: Average costs for Heart Attack, Heart Failure, Pneumonia, and Hip/Knee conditions	
Quality Metrics: Quality ratings for procedures like Heart Attack, Heart Failure, Pneumonia, and Hip/Knee conditions.	
Value Metrics: Cost effectiveness for the same procedures.	

Data Preprocessing

- Missing Data: Rows with missing values were removed to ensure data quality.
- Categorical Encoding: Features like Facility.Type and performance ratings were converted into factors for compatibility with modeling.
- Normalization: Numerical predictors (e.g., procedure costs) were scaled for consistency.
- Class Imbalance: Target variable (Rating) showed ~67% Low and ~33% High ratings; this was noted for modeling adjustments.
- Train-Test Split: Data was split 80%-20% to ensure robust evaluation.

Model Development

We implemented a machine learning workflow with 5-fold cross-validation for model selection and evaluation. The dataset was split 80/20 for training and testing, using ROC AUC as the metric. Logistic regression with ridge regularization and random forest were assessed using predefined hyperparameter grids. The best model, logistic regression with λ=0.001, achieved a cross-validation AUC of 0.94 and test AUC of 0.95, with 87.28% accuracy, indicating no overfitting. A confusion matrix and ROC curve illustrated performance. While logistic regression outperformed random forest, further tuning might improve the latter. The results suggest the dataset is well-suited for a linear model.
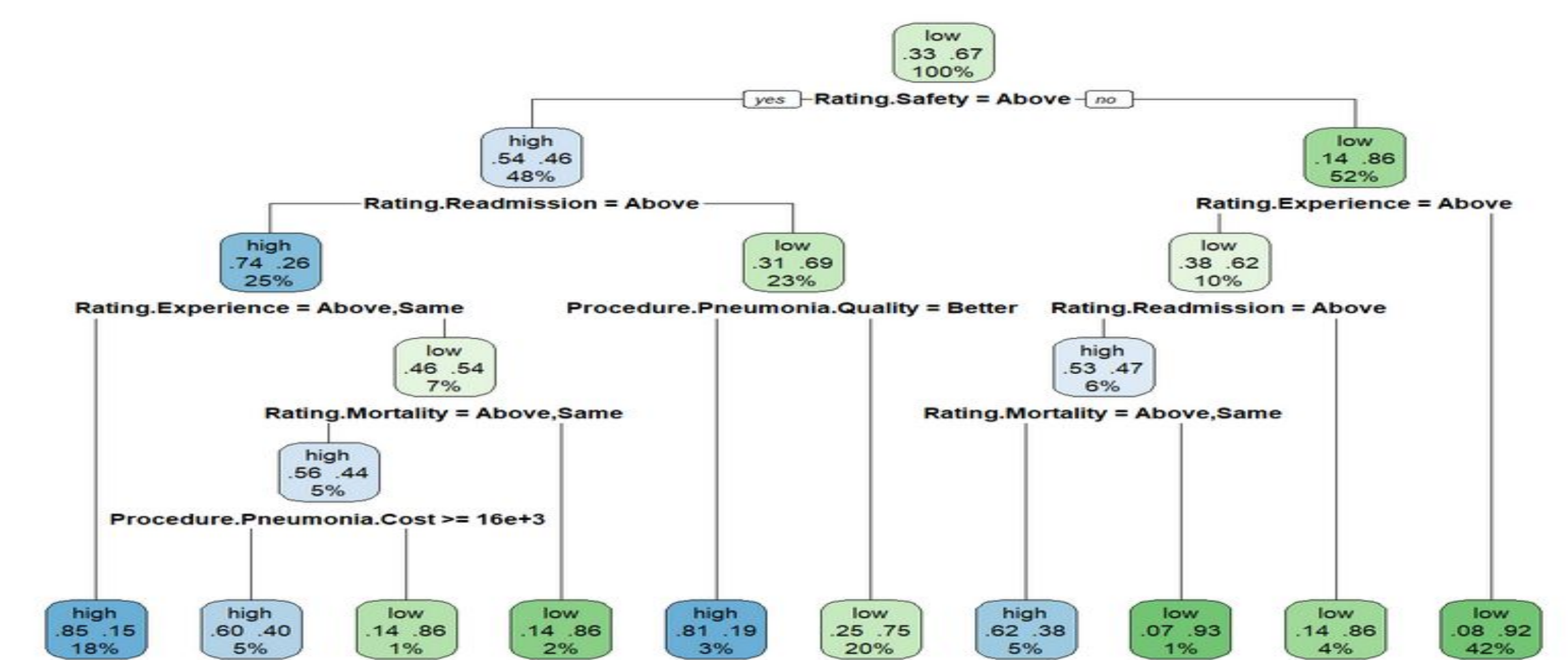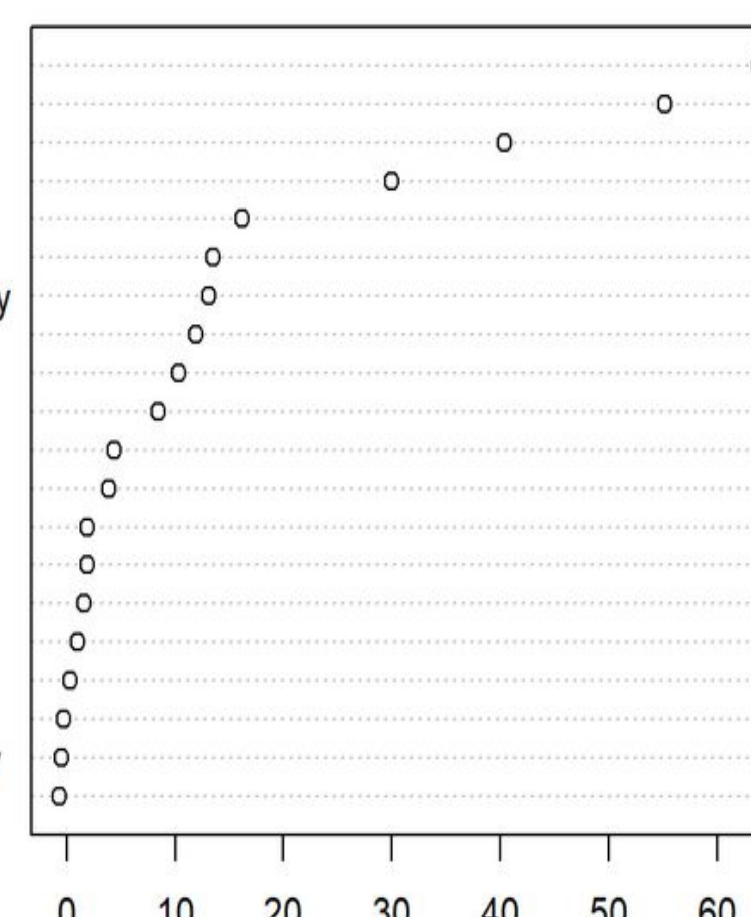
Analysis and Results
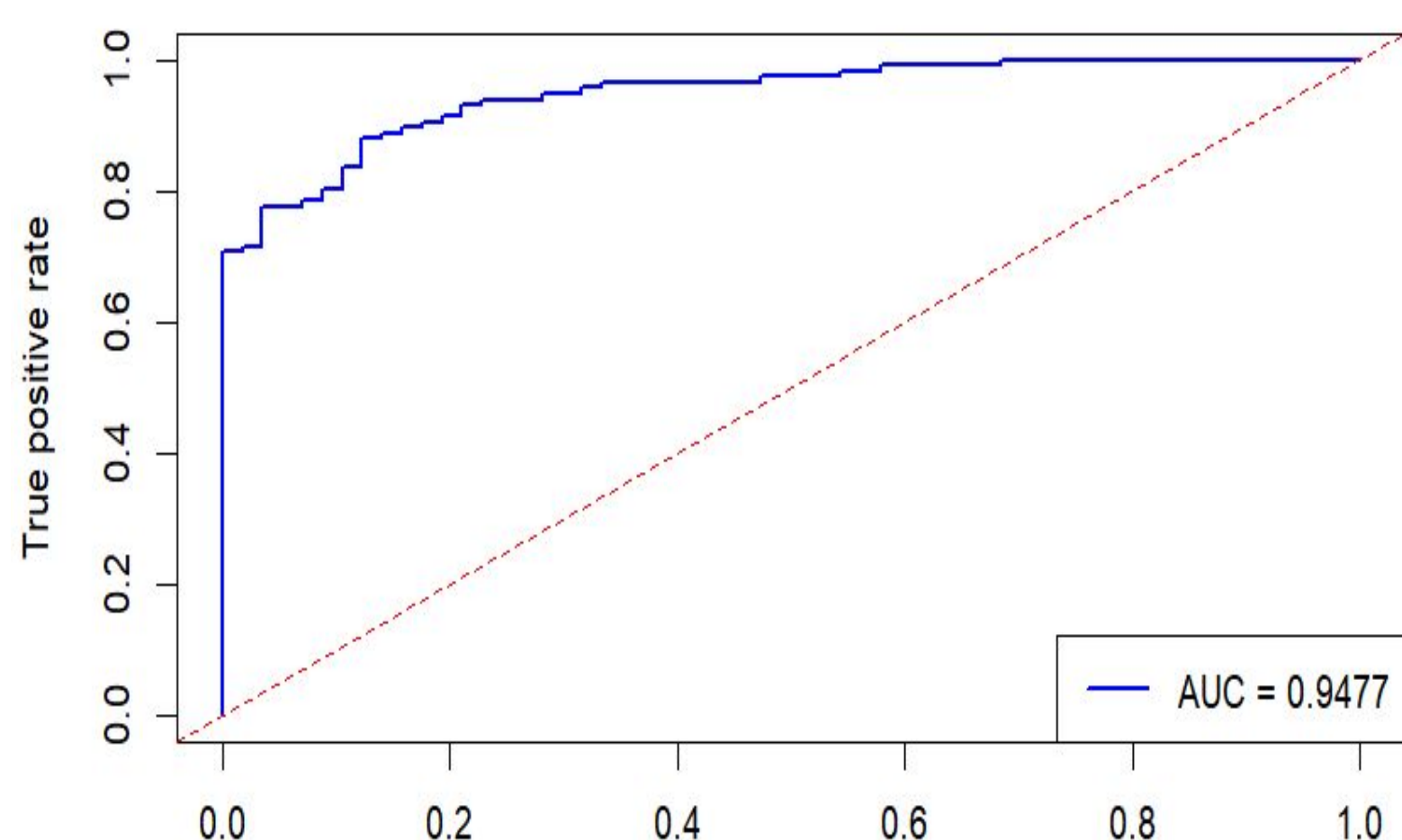
The ROC curve shows exceptional model performance with an AUC of 0.9477, with the curve's steep rise and high positioning indicating strong discriminative ability between classes.

The plot reveals Rating Safety, Rating Transmission, and Rating Experience as the top predictive features, while Facility Type has minimal impact on model performance.

Conclusion

We evaluated logistic regression with ridge regularization and random forest models using 5-fold cross-validation and ROC AUC. Logistic regression (λ=0.001) outperformed random forest with a cross-validated AUC of 0.94, a test AUC of 0.95, and an accuracy of 87.28%. The model effectively classified hospitals into low and high ratings, highlighting its potential for improving hospital performance assessments.