Machine Learning Models to Predict AKI Recovery

Thomas Tao-Min Huang MD.[1], Po-Chi Kuo PhD.[2], Venesia Ruan BA.[2], Fang-Sheng Mao MD.[1], Yi-Shiuan Tseng BA.[2], Cho-Hung Chiang MD.[1], Yu-Chang Yeh MD. PhD.[1], Vin-Cent Wu MD. PhD[1]

[1]National Taiwan University Hospital and College of Medicine, Taipei City, Taiwan

[2]Department of Computer Science, National Tsing Hua University (NTHU), Hsin-Chu City, Taiwan

Introduction

Acute kidney injury (AKI) is a severe complication in critically ill patients admitted to intensive care units (ICUs). AKI patients who recover their kidney function have better outcomes compared to those who do not. Identifying patients who are likely to recover within seven days of AKI onset can help healthcare providers plan for timely interventions and avoid unnecessary treatments. However, traditional predictive models have poor performance in identifying such patients. This study aims to develop machine learning models to predict seven-day kidney recovery using clinical parameters from the first days of AKI.

Methods and Materials

We conducted a retrospective cohort study using the Medical Information Mart for Intensive Care IV (MIMIC-IV) database to identify relevant features predicting seven-day AKI recovery. We included patients admitted to ICUs with AKI or incident AKI within seven days after ICU admission. Patients with a history of dialysis were excluded. We identified patients with complete kidney recovery (defined as returning to 110% of baseline serum creatinine) and used conventional logistic regression, random forest, XGBoost, and decision tree models to identify important features predicting recovery.

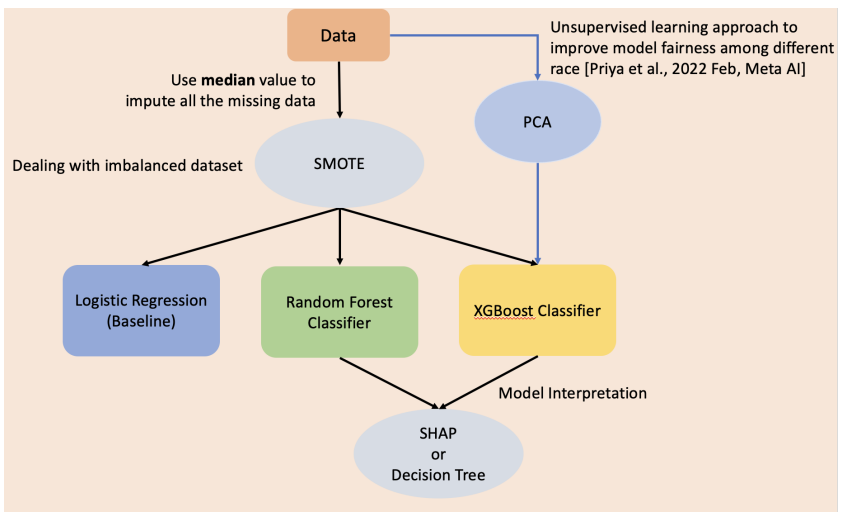

Figure 1. Process of Model Construction

Feature	Importance	Feature	Importance
Creatinine on the 1st Day	1.18	Age	0.28
Creatinine baseline	0.81	Gender	0.25
Ethnicity: White	0.55	Pulmonary disease	0.22
Kidney diseases	0.49	Blood Urea Nitrogen	0.2
AKI > stage 2	0.28	APACHE III	0.13

Table 1. Important Features and Their Weighting

Results

A total of 2449 patients were included in the study, of which 992 (40.5%) were female, and the mean age was 66.6±15.4 years. XGBoost had the highest area under the curve (AUC) of 0.80, followed by random forest (AUC: 0.75) and logistic regression (AUC: 0.61). Decision tree analysis found that creatinine on AKI day 1, blood urea nitrogen, underlying kidney diseases, age, and baseline creatinine were important features for identifying patients with kidney recovery.

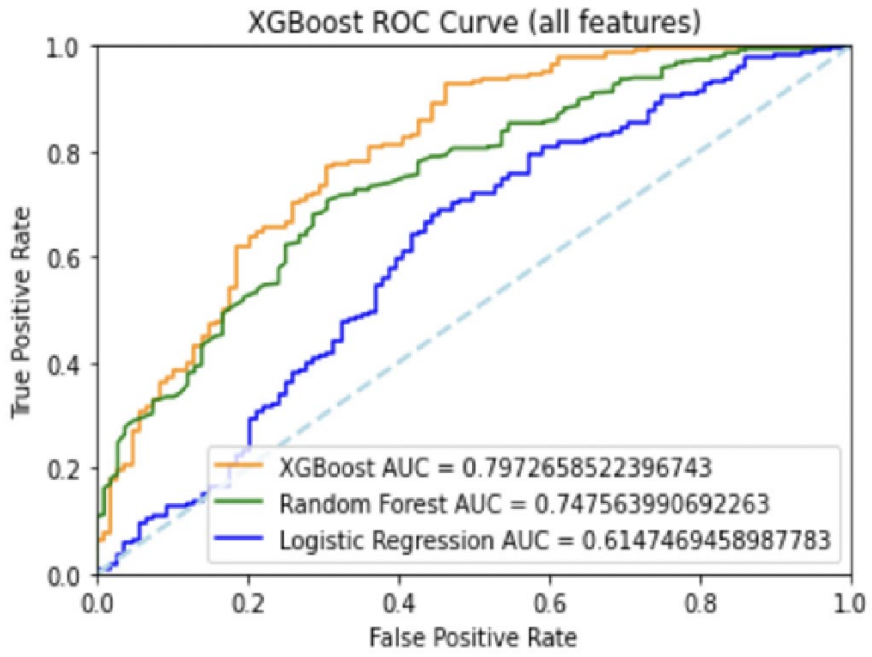

Figure 2. Performance of XGBoost, Random Forest and Logistic Regression to Predict 7 days AKI Recovery

Discussion

Our study found that machine learning models, especially XGBoost models, had better predictive performance than traditional models in identifying patients who are likely to recover from AKI within seven days. The use of machine learning models can aid in early risk stratification and timely intervention to improve patient outcomes. The important features identified by the decision tree analysis can also aid healthcare providers in identifying at-risk patients.

Conclusions

Our study showed that machine learning models can accurately predict seven-day kidney recovery in AKI patients. XGBoost models had the best predictive performance among the models evaluated. Identifying patients who are likely to recover within seven days of AKI onset can help healthcare providers plan for timely interventions and avoid unnecessary treatments. The use of machine learning models in clinical practice can aid in early risk stratification and timely intervention to improve patient outcomes in AKI patients.