Ensemble of Trees Approaches to Risk Adjustment For Evaluating a Hospital’s Performance

Yang Liu
Mikhail Traskin
Scott A. Lorch
Edward I. George
Dylan S. Small

Abstract: A commonly used method for evaluating a hospital’s performance on an outcome is to compare the hospital’s observed outcome rate to the hospital’s expected outcome rate given its patient (case) mix and service. The process of calculating the hospital’s expected outcome rate given its patient mix and service is called risk adjustment. Risk adjustment is critical for accurately evaluating and comparing hospitals’ performances since we would not want to unfairly penalize a hospital just because it treats sicker patients. The key to risk adjustment is accurately estimating the probability of an outcome given patient characteristics. For cases with binary outcomes, the method that is commonly used in risk adjustment is logistic regression. In this paper, we consider ensemble of trees methods as alternatives for risk adjustment, including random forests and Bayesian additive regression trees (BART). Both random forests and BART are modern machine learning methods that have been shown recently to have excellent performance for prediction of outcomes in many settings. We apply these methods to carry out risk adjustment for the performance of neonatal intensive care units (NICU). We show that these ensemble of trees methods outperform logistic regression in predicting mortality among babies treated in NICU, and provide a superior method of risk adjustment compared to logistic regression.