**ABSTRACT**

Predicting Patients health is a critical task in the Healthcare Industry. Healthcare datasets show a high degree of imbalance especially for rare diseases. The current work aims at predicting the post operative survival rate in thoracic surgery datasets. The dataset exhibits data imbalance with around 15% positive cases and remaining 85% negative cases. The commonly applicable machine learning techniques for prediction score poorly in predicting the positive cases in spite of high accuracy of the predictions for the negative cases. We use SMOTE (synthetic minority oversampling technique) to reduce the degree of imbalance and increase the positive samples proportion before the application of the following classifiers: Naive Bayes, Neural Networks, Random Forest, Boosting algorithms - Adaboost, Extreme Gradient boosting and Support Vector Machines and examine the results. The study shows that SVM and Naïve Bayes show significantly better performance on the imbalanced datasets than other models using synthetic datasets than under normal conditions.