

ABSTRACT

Heart disease is a prevalent and potentially fatal condition that affects millions of people worldwide. Early detection and accurate prediction of heart disease risk factors are crucial for effective prevention and treatment. In this study, we propose a machine learning approach using Random Forest Classification (RFC) to predict the risk of heart disease based on various demographic, clinical, and lifestyle factors. The dataset used in this study contains a comprehensive set of attributes such as age, sex, cholesterol levels, blood pressure, and smoking habits. We pre-process the dataset to handle missing values and normalize the features. We then train the RFC model on the pre-processed data and evaluate its performance using metrics such as accuracy, precision, recall, and F1-score. Our results show that the RFC model achieves high accuracy in predicting the risk of heart disease, demonstrating its potential as a valuable tool for early detection and risk stratification. Heart disease is a leading cause of mortality worldwide, and early detection of its risk factors is crucial for effective prevention and treatment. Machine learning techniques, such as Random Forest Classification (RFC), have shown promise in predicting heart disease risk. This study aims to develop and evaluate an RFC model for heart disease prediction using a comprehensive dataset of patient information, including demographic, clinical, and lifestyle factors. The dataset is pre-processed to handle missing values and normalize the features. The RFC model is trained and evaluated using performance metrics such as accuracy, precision, recall, and F1-score. Our results demonstrate that the RFC model achieves high accuracy in predicting heart disease risk, indicating its potential as a valuable tool for early detection and risk stratification.