

Heart Disease Prediction: using Machine Learning

Manas Tripathi, Kshitiz Jain, Jeraldin PJ, Ayush Gupta, Mohd. Armaan

Raj Kumar Goel Institute of Technology, Ghaziabad, UP, India

Abstract: Cardiovascular disease remains one of the foremost global health challenges, contributing substantially to morbidity and mortality rates worldwide. Early detection and accurate prediction are critical in mitigating the impact of heart-related conditions. This study explores the integration of Machine Learning (ML) algorithms with Explainable Artificial Intelligence (XAI) methodologies to enhance the prediction and interpretability of heart disease risk. Utilizing a publicly available dataset from Kaggle, various classification models—such as Decision Trees, Random Forest, Logistic Regression, Support Vector Machines, K-Nearest Neighbours, and Naive Bayes—were trained and evaluated. Among these, the Random Forest model demonstrated the highest predictive performance, closely followed by Logistic Regression. To provide interpretability, SHAP (Shapley Additive explanations) was applied to the Logistic Regression model, offering intuitive visual insights into the influence of individual features on model predictions. The use of SHAP plots further enhances transparency by highlighting the contribution of specific variables to the prediction outcomes. This approach not only facilitates accurate diagnosis but also aids clinicians and stakeholders in understanding the model's decision-making process.

Keywords: Cardiovascular Disease, Machine Learning, Explainable AI, SHAP, Predictive Modelling

