Optimizing Machine Learning Models for Heart Disease Prediction Using UCI Cleveland Dataset

Arjun Patil[1] and Vaishnavi S Gharat[2]

Assistant Professor and Head Department of IT[1]

Student P.G. Department of IT[2]

Veer Wajekar ASC College, Phunde, Uran

Abstract: *Heart disease remains one of the most critical health issues globally, with acute myocardial infarction (AMI) posing serious threats to human life. Early diagnosis can greatly enhance treatment outcomes and prevent fatalities. In this study, we explore the application of machine learning (ML) models to predict heart disease using the UCI Cleveland dataset. Four classifiers—Logistic Regression, Naïve Bayes, Support Vector Machines (SVM), and XGBoost—were evaluated. The work emphasizes preprocessing techniques to mitigate overfitting and ensure data quality. Among the evaluated models, XGBoost demonstrated superior performance with a 92% accuracy and an AUC score of 0.94. This research highlights the significance of ML in healthcare diagnostics and suggests directions for future improvements through feature enhancement and hybrid approaches.*

Keywords: *Heart disease*