Hybrid Feature Selection Approaches for Heart Disease Prediction Using Machine Learning Models

Dhanashri S. Karande*
Department of Electronics & Telecommunication Engineering, SVPM'S, COE, Malegaon (Bk)-Baramati, SPPU, Pune, Maharashtra, India

Dr. Shailendrakumar M. Mukane
Department of Electronics & Telecommunication Engineering, SVPM'S, COE, Malegaon (Bk)-Baramati, SPPU, Pune, Maharashtra, India

Abstract:

In healthcare analytics, precise and timely prediction is crucial because heart disease remains a major cause of death worldwide. This study presents a hybrid feature selection framework for heart disease prediction, combining filter (MI), wrapper (RFE), and embedded (Lasso) methods. Three hybrid approaches were developed. FSMR, which uses Mutual Information (MI) and Recursive Feature Elimination (RFE) for feature selection; FSML, which employs Mutual Information (MI) and Lasso for feature selection; and FSRL, which combines Recursive Feature Elimination (RFE) and Lasso for feature selection. Four benchmark datasets from the UCI repository, including Cleveland (D1), Switzerland (D2), Hungary (D3), and VA Long Beach (D4), were combined to enhance robustness and sample diversity. The selected features were tested using five Machine Learning (ML) models: Random Forest (RF), Extreme Gradient Boosting (XGBoost), Logistic Regression (LR), Support Vector Machine (SVM), and Naïve Bayes (NB). Performance was evaluated with accuracy, sensitivity, specificity, and precision. Among the methods tested, FSRL with RF achieved the highest accuracy, demonstrating the effectiveness of combining wrapper and embedded techniques. The results indicate that hybrid feature selection greatly improves diagnostic reliability compared to single-method approaches, providing a strong basis for clinical decision support systems in heart disease healthcare.

Keywords:

Heart disease prediction, Hybrid feature selection, Recursive Feature Elimination, Lasso, Mutual Information, Machine learning.