Exploring Machine Learning Applications in Predicting Heart Disease Risk: A Comparative Study of Logistic Regression, SVM, and Decision Trees

Asiana Holloway, Group 28

Abstract

Heart disease remains a leading cause of mortality worldwide, making early detection and risk assessment critical for improving patient outcomes. Artificial intelligence (AI) and machine learning (ML) have demonstrated significant potential in healthcare, particularly in developing predictive models for identifying individuals at high risk of heart disease. This research explores the application of ML algorithms to predict heart disease risk using the structured and widely recognized Heart Disease UCI dataset.

This study focuses on evaluating the effectiveness of commonly used ML models, such as Logistic Regression, Support Vector Machines (SVM), and Decision Trees, for cardiovascular risk prediction. The dataset includes clinical features such as age, cholesterol levels, blood pressure, and electrocardiogram results, providing comprehensive patient information for analysis. By examining the strengths and limitations of these models, the research highlights how AI can enhance traditional diagnostic processes, offering non-invasive, data-driven insights for preventative healthcare.

This project contributes to the growing integration of AI in healthcare by providing a deeper understanding of how ML algorithms can be leveraged to improve risk prediction systems, reduce the burden on healthcare professionals, and enable earlier interventions. Furthermore, it emphasizes the importance of transparent and interpretable AI models to ensure clinical adoption and trust in healthcare systems.

Citations:

Heart Disease UCI Dataset – Kaggle. Available at: https://www.kaggle.com/datasets/cherngs/heart-disease-cleveland-uci

Deo, R. C. (2015). Machine Learning in Medicine. Circulation, 132(20), 1920-1930. https://www.ahajournals.org/doi/10.1161/CIRCULATIONAHA.115.001593

Weng, S. F., Reps, J., Kai, J., et al. (2017). Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLoS ONE, 12(4), e0174944. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0174944