ScienceDirect

Expert Systems with Applications

Volume 263, 5 March 2025, 125678

Novel framework of significant risk factor identification and cardiovascular disease prediction

Soham Bandyopadhyay [a], Ananya Samanta [b], Monalisa Sarma [c], Debasis Samanta [d]

Show more

Share Cite

https://doi.org/10.1016/j.eswa.2024.125678

Get rights and content

Highlights

- A hybrid statistical approach to extract optimum key risk factors of heart disease.
- Substantiate the efficacy of optimum key risk factors with statistical measures.
- Implement Stacked Meta Neural Network classifier to predict heart disease.

Abstract

Cardiovascular disease (CVD) remains a major public health concern, characterized by high mortality rates and complex diagnostic challenges. Risk factor-based prediction models are commonly employed, but existing approaches often treat all factors equally important. This study proposes a novel method for predicting CVD using a minimal set of significant risk factors, without compromising predictive accuracy. The proposed framework consists of two main components: the Heart Disease Assessment and Review Technique (HEART) and the Stacked Meta Neural Network (SMNN). HEART has been introduced to identify the key risk factors of CVD. Different statistical techniques such as McFadden's pseudo-R-squared, mutual information, Friedman's and McNemar's tests, Shapley and Odds ratio analysis were used to signify and prioritize key risk factors over non-key risk factors. SMNN with 5-fold cross-validation uses these significant risk factors to predict CVD. The method is evaluated using multiple benchmark datasets, yielding an average accuracy of 90.5% ± 1.8% (IEEE Data-port), 88.5% ± 3.3% (Government College University, Faisalabad, Pakistan) and 80.3% ± 1.3% (South African Heart Disease dataset) underscoring its robustness. This study offers new insights into CVD risk factor assessment, enhancing diagnostic accuracy and informing more effective clinical interventions to improve patient outcomes.