Modelling Heart Attack Risk Based on Age, Hypertension, Gender, Psychology Factors, and Food Consumption Using Least Square Spline Semiparametric Binary Logistic Regression

Nur Chamidah
Faculty of Science and Technology, Universitas Airlangga, Indonesia

Hendri Susilo
Faculty of Medicine, Universitas Airlangga, Indonesia

Marisa Rifada
Faculty of Science and Technology, Universitas Airlangga, Indonesia

Dursun Aydin
Department of Statistics, Mugla Sitki Kocman University, Turkey

Naufal Ramadhan Al Akhwal Siregar
Faculty of Science and Technology, Universitas Airlangga, Indonesia

Muhammad Fikry Al Farizi
Faculty of Science and Technology, Universitas Airlangga, Indonesia

Abstract

Heart attacks are non-communicable diseases and remain one of the leading causes of death in Indonesia, affecting both young and elderly populations. Identifying the main contributing factors is crucial for prevention and early intervention. This study aims to predict the risk of heart attack based on key predictors including history of chest pain, gender, HbA1c levels, anxiety, and physical health. These factors represent both linear and non-linear effects on heart attack risk. This study aims to predict the risk of heart attacks through biostatistical modeling based on Least Square Spline Semiparametric Binary Logistic Regression (LSS-SBLR) method. The data used are secondary data obtained from Universitas Airlangga Hospital, Surabaya, Indonesia, in 2025. The results indicate that a history of chest pain, high HbA1c, and anxiety levels significantly increase the risk of heart attack, while good physical health reduce it. Moreover, female patients tend to have a lower risk of heart attack compared to males. The model achieved an in-sample accuracy of 87.5% with an AUC of 88.33%, and an out-sample accuracy and AUC of 100%, indicating excellent predictive performance. This study contributes to SDGs point 3, by providing a predictive framework for early detection and prevention of heart attacks in Indonesia.