

Comparison Analysis of Classification and Regression Tree (CART) and Random Forest performance in prediction of cardiovascular diseases

Nurul Afni Zohrah¹, Sri Suryani Prasetyowati², Yuliant Sibaroni³

^{1,2,3}Faculty of Informatics, Telkom University, Bandung

¹nurulafnizohrah@students.telkomuniversity.ac.id, ²srisuryani@telkomuniversity.ac.id,

³yuliant@telkomuniversity.ac.id.

Abstract

Cardiovascular or also known as heart disease is a condition where there is interference with the heart and blood vessels. The death rate caused by cardiovascular continues to increase with age, where 100 cases per 1000 people are over the age of 60. Lack of awareness of a healthy lifestyle and lack of information related to heart disease encourages the need for research or systems that can predict the causes of heart diseases with precise accuracy and high performance. In this final project, a comparative analysis was conducted on the performance of the CART and Random Forest algorithms in predicting factors causing heart diseases or cardiovascular. CART is a method for selecting a group of data in a space, which is then called a node, which is then divided into two child nodes and each child node can be split into the next two child nodes. Meanwhile, Random Forest is one of the ensemble methods used to improve the data classification method from a single unstable sorter through a combination of many sorters from the same method as the voting process to obtain final classification predictions. In this study, CART algorithm produces an optimal tree with an accuracy value is 76% with a precision value for class '0' is 81% and class '1' is 43%, recall value for class '0' is 84% and class '1' is 53%, F1-Score for class '0' is 84% and class '1' is 53%. While the classification tree produced by the Random Forest method has an accuracy is 80%, where the precision value is 83% for class '0' and 67% for class '1', the recall value is 91% for class '0' and 50% for class '1', and F1-Score, namely 87% for class '0' and 57% for class '1'

Keywords: Cardiovascular, CART, Random Forest