Abstract

This study proposes utilizing the machine learning technique CART to classify the spread of dengue hemorrhagic fever (DHF). To expand the features used, the CART classification model was developed based on data collected over the previous 2 to 4 years. The data sources included the Bandung City Health Office for the cases of DHF, the Bandung Meteorology, Climatology and Geophysics Agency for the climate data, the Bandung City Central Statistics Agency for population and educational history data. The top-performing CART classification model over the past 2, 3, and 4 years achieved accuracies of 93%, 93%, and 90%, respectively. The models that exhibited the highest accuracy values and optimal number of feature extensions were chosen as the best ones. CART is among several machine learning techniques that can effectively measure the most impactful features during the classification process. The meteorological parameters were found to be irrelevant in the classification process. This study reveals that the population size, male population proportion, and educational attainment levels are the most impactful features in the classification of DHF spread in Bandung City through feature expansion.

Keyword: CART, Classification, Dengue Hemorrhagic Fever, Feature Expansion, Machine Learning