ABSTRACT

Air is one of the crucial components for living beings. In 2021, DKI Jakarta ranked 12th as the most polluted capital city in the world. The impacts of polluted air are extremely dangerous. The air quality in Indonesia can be assessed using the Air Pollution Standard Index (ISPU). ISPU is calculated based on several parameters, including PM_{10}, NO_2, O_3, SO_2, and CO. ISPU categorizes air quality as good, moderate, unhealthy, very unhealthy, and hazardous, which will be shared as publicly accessible information. This aligns with the directive of the Menteri Lingkungan Hidup dan Kehutanan Republik Indonesia, as stated in Ministerial Regulation No. 14 of 2020. Therefore, classifying ISPU is necessary to assist the government in assessing air quality in DKI Jakarta. This study implements classification and regression trees (CART) and random forest using GridSearchCV for classifying ISPU in DKI Jakarta. CART is considered an easily interpretable, fast, and accurate classification method. Random forest is a classifier capable of reducing the risk of overfitting compared to CART. The predictor variables used in this study are PM_{10}, NO_2, O_3, SO_2, and CO. The results show that the CART model with GridSearchCV achieves an accuracy of 98,40%, and the random forest model with GridSearchCV achieves an accuracy of 98,52%. Based on this research, the application of GridSearchCV to CART and random forest effectively classifies ISPU in DKI Jakarta.

Keywords: Air Pollution Standard Index, Classification, Classification and Regression Tree, Random Forest, GridSearchCV