ABSTRACT

Indonesia is a region that is predominantly characterized by a tropical climate. As a predominantly tropical region, Indonesia experiences narrow temperature variations but diverse rainfall patterns. The rainfall in Indonesia exhibits significant levels of diversity. To mitigate the hazards associated with high rainfall, such as floods and landslides, the implementation of *data mining* classification is necessary. In addition to disaster preparedness, rainfall information is valuable in the fields of agriculture, transportation, and industry. By implementing *data mining* classification, it aids in the prediction of rainfall in Indonesia. This research utilizes daily climate data in Indonesia and adopts the random forest algorithm as the classification method. The selection of the random forest algorithm is based on its ability to produce accurate and stabel classification models without the need for complex parameter adjustments. Furthermore, the naïve bayes method is employed due to its ease of implementation and simple probability modeling, applicable to various types of classification data. Based on the research findings, it can be concluded that the random forest algorithm outperforms the naïve bayes algorithm in classifying climate datasets in Indonesia, achieving an accuracy rate of 86.55% compared to Naïve bayes accuracy rate of 36.61%. The results of this study are expected to serve as a reference for future research literature and contribute to monitoring daily rainfall in Indonesia for the prevention of natural disasters.

Keyword— ***classification, data mining, rainfall, naïve bayes, random forest***