ABSTRACT

Water is one of the main human needs. Even though the amount of water is very abundant in Indonesia, not all water is suitable for consumption. In this study, the Random Forest *and* Naive Bayes *algorithms were used to classify water quality in DKI Jakarta. This study was designed to determine the effectiveness of using Data Mining in classifying water quality. In the process, the* Random Forest *and* Naive Bayes *methods are used to determine the most accurate algorithm for classifying.* Random Forest *and* Naive Bayes *algorithms will be used in the water quality classification process. The classification process is carried out using the collected data with predetermined parameters. The testing process produces performance values from each category the* Random Forest *Algorithm has an accuracy of 85.23%, a precision of 88.11%, a recall of 79.89%, and an F1-score of 83.35. While the* Naïve Bayes *algorithm has an accuracy of 72.51%, 76.19% precision, 74.51% recall, and 73.08% F1-score. Thus it can be concluded that the* Random Forest *Algorithm has better performance than* Naïve Bayes.

Keywords: Water, Classification, Effectiveness, Accuracy