

ABSTRACT

Rivers are a source of life in the surrounding environment for household, agricultural and industrial purposes. Because rivers have such an important role, it is necessary to maintain water quality so that it does not have a negative impact on the surrounding environment. In fact, the water quality of the Citarum River was once the most polluted and dirty river in the world. River water quality can be determined by manually calculating the pollution index value. However, the manual calculation method is considered slow and ineffective. Therefore, an effective technique is needed to measure river water quality and publish the results to the public so that the water quality of the Citarum River can be maintained.

Machine learning was chosen as one of the proposed techniques to classify river water quality and visualize the results in a public-friendly information sistem. Three classification methods are applied in this final project: K-Nearest Neighbors, Support Vector Machines, and Random Forest. The visualization is done on a website that contains maps from upstream to downstream of the Citarum River.

In this research, we found that using random forest gives better results between the K-Nearest Neighbors method and the Support Vector Machines method, with an accuracy of 99.33%. The three methods will become the classification model of the Citarum River Mapping Sistem and will be implemented to show the water quality of the Citarum River at each monitoring point.

Keywords: River Water Quality, Machine learning, Classification, Mapping Sistem