

## ABSTRAK

Dalam aktivitas kehidupan sehari-hari manusia banyak memanfaatkan air terutama air bersih yang berasal dari berbagai macam sumber, sebagai contoh sungai. Namun pada saat ini banyak kerusakan lingkungan air baik yang ada di sungai merupakan ulah manusia. Banyak manusia yang membuang sampah sembarangan di sungai dan juga banyak sungai yang berdekatan dengan pabrik, kebanyakan saluran pembuangan dari tempat pabrik tersebut mengarah ke sungai. Hal ini tentunya menjadi penyebab air menjadi tercemar dan kualitas air menurun. Penelitian ini mengklasifikasikan kualitas air sungai di DKI Jakarta menjadi ‘cemar berat’ dan ‘cemar sedang’ menggunakan data mining dengan algoritma Support Vector Machine (SVM). Data dilatih dan diuji berdasarkan rasio data latih dengan jangkauan 0.5 sampai dengan 0.9 untuk melihat performa algoritma SVM dalam mengklasifikasikan kualitas air dengan menggunakan variasi parameter C (Cost). Hasil pengujian menunjukkan bahwa rasio data latih 0.9, dapat bekerja dengan sangat optimal dalam melakukan klasifikasi kualitas air sungai dengan performa akurasi sebesar 98.65%.

## ABSTRACT

In daily human activities, water is extensively utilized, particularly clean water from various sources, such as rivers. However, currently, there is significant environmental damage to water bodies, especially rivers, caused by human activities. Many people dispose of waste irresponsibly into rivers, and numerous rivers are located near factories, with most of their waste channels leading directly into the rivers. Consequently, this has led to water pollution and a decline in water quality. This research classifies the water quality of rivers in DKI Jakarta into two categories: 'heavy pollution' and 'moderate pollution' using data mining with the Support Vector Machine (SVM) algorithm. The data is trained and tested based on a training data ratio ranging from 0.5 to 0.9 to evaluate the SVM algorithm's performance in classifying water quality by varying the parameter C (Cost). The test results indicate that with a training data ratio of 0.9, the algorithm performs optimally, achieving a classification accuracy of 98.65% in determining the river water quality.