

ABSTRAK

Mengingat betapa pentingnya air bagi seluruh makhluk hidup, menjaga ekosistem, dan memiliki dampak positif pada kesehatan manusia. Dari banyaknya fungsi air bagi manusia, tidak semua air dapat dikonsumsi bagi tubuh manusia dikarenakan kandungan-kandungannya yang tidak cocok untuk dikonsumsi oleh manusia. Penelitian ini bertujuan untuk mengklasifikasi data air yang dapat diminum dengan algoritma SVM yang menghasilkan akurasi, presisi, F-1 Score, dan recall teroptimal dalam klasifikasi data kualitas air yang dapat diminum. Metode SVM telah diterapkan untuk memodelkan hubungan kompleks antara parameter kualitas air seperti *ph*, *Hardness*, *Solids*, *Chloramines*, *Sulfate*, *Conductivity*, *Organic carbon*, *Trihalomethanes*, *Turbidity*.

Penelitian ini melibatkan langkah-langkah analisis data, *pre-processing*, dan pelatihan model SVM menggunakan 3 parameter yaitu kernel, Cost, dan Gamma serta melakukan evaluasi model dengan metode *K-fold Cross Validation*. Hasil eksperimen menunjukkan bahwa SVM mampu memberikan klasifikasi dengan teroptimal pada kombinasi parameter kernel = Rbf, Cost=10, dan Gamma=1, dengan diterapkan K-fold cross validation yang paling optimal adalah K=9. Hasil tertinggi yang didapat adalah akurasi 76.58%, Presisi 77.48%, Recall 76.58%, dan F-1 Score 76.46%.

Kata Kunci : *Support Vector Machine*, Klasifikasi, Algoritma, Kualitas air.

ABSTRACT

Considering the crucial role of water for all living organisms in maintaining ecosystems and positively impacting human health, not all water is suitable for human consumption due to its varying contents. This research aims to classify drinkable water data using the SVM algorithm, achieving optimal accuracy, precision, F-1 Score, and recall in the classification of drinkable water quality data. The SVM method is applied to model complex relationships among water quality parameters such as pH, Hardness, Solids, Chloramines, Sulfate, Conductivity, Organic carbon, Trihalomethanes, and Turbidity.

This study involves data analysis, pre-processing steps, and SVM model training using three parameters: kernel, Cost, and Gamma. Model evaluation is conducted through K-fold Cross Validation. Experimental results indicate that SVM can provide optimal classification with the combination of parameters: kernel=Rbf, Cost=10, and Gamma=1. The most optimal K-fold cross-validation is achieved with K=9. The highest obtained results include an accuracy of 76.58%, precision of 77.48%, recall of 76.58%, and an F-1 Score of 76.46%.

Keywords : Support Vector Machine, Classification, Algorithm, Water Quality.