

PERANCANGAN SISTEM DAN PENERAPAN ALGORITMA K-NEAREST NEIGHBOR (KNN) UNTUK MEMPREDIKSI KUALITAS AIR YANG DAPAT DIKONSUMSI

Hardiana Said

ABSTRAK

Kualitas air yang aman untuk dikonsumsi sangatlah penting bagi kesehatan masyarakat luas di setiap daerah, namun kualitas air diberbagai daerah semakin menurun terutama untuk kebutuhan manusia dalam hal air minum, dampak dari kualitas air yang tidak aman untuk dikonsumsi dapat menyebabkan penyakit seperti kolera, diare, hepatitis A dan lainnya, hal ini dikarenakan air yang memiliki sanitasi yang buruk dan zat-zat yang melebihi kadar standar. Penelitian ini dilakukan untuk membuat sebuah sistem prediksi kualitas air dan untuk mendukung dilakukannya pemrosesan prediksi yaitu dengan menerapkan algoritma klasifikasi *data mining* yaitu adalah algoritma *K-Nearest Neighbor*. Algoritma ini diterapkan untuk menghitung probabilitas kualitas air yang aman atau tidak untuk dikonsumsi berdasarkan data rekaman yang diambil dari lingkungan sekitar terutama di daerah padat penduduk. Kumpulan data diperoleh dari website *kaggle* untuk percobaan pada sistem yang akan dibuat. Hasil pemodelan diukur menggunakan tabel *Confusion Matrix* untuk menghitung akurasi. Setelah diuji, model ini memiliki tingkat akurasi tertinggi 85,52% dengan nilai k (tetangga terdekat) = 3. Berdasarkan hasil permodelan kemudian telah dibuat aplikasi berbasis web untuk memprediksi kualitas air yang dapat dikonsumsi.

Kata kunci: Kualitas Air, Prediksi, Klasifikasi, Sistem, *K-Nearest Neighbor*.

**PERANCANGAN SISTEM DAN PENERAPAN ALGORITMA
K-NEAREST NEIGHBOR (KNN) UNTUK MEMPREDIKSI
KUALITAS AIR YANG DAPAT DIKONSUMSI**

Hardiana Said

ABSTRACT

The water quality that is safe for use is especially important for public health in every area, but water quality in various areas is decreasing for human needs in drinking water, the impact of water quality that is not safe for consumption can cause diseases such as cholera, diarrhea, hepatitis and others, this is due to water that has poor sanitation and substances that exceed standard levels. This research was conducted to create a prediction system about water and to support predictions by applying the application of data mining classification, namely the K-Nearest Neighbor algorithm. This algorithm is applied to calculate the probability that water quality is safe or not based on recorded data taken from the surrounding environment, especially in densely populated areas. A collection of data obtained from the kaggle website for experiments on the system to be created. The results of the modeling using the Confusion Matrix table to calculate accuracy. After being tested, this model has the highest accuracy rate of 85.52% with a value of k (nearest neighbor) = 3. Based on the modeling results, a web-based application is then made to predict the quality of water that can be consumed.

Keywords: Water Quality, Prediction, Classification, System, *K-Nearest Neighbor*.