

PREDIKSI KUALITAS UDARA MENGGUNAKAN ALGORITMA *K-NEAREST NEIGHBOR*

Adinda Amalia

Abstrak

Dalam kehidupan sehari-hari, udara digunakan untuk bernafas oleh makhluk hidup. Udara yang bersih mengandung banyak manfaat bagi kehidupan. Namun, udara yang terdapat di alam tidak selalu dalam keadaan bersih, sehingga dapat menyebabkan penurunan kualitas udara. Kualitas udara yang buruk dapat memberikan dampak terhadap kesehatan manusia dan lingkungan sekitarnya. Kualitas udara di DKI Jakarta dapat diketahui melalui Indeks Standar Pencemar Udara (ISPU). Penelitian ini bertujuan untuk melakukan prediksi kualitas udara yang ada di DKI Jakarta berdasarkan data ISPU. Prediksi dilakukan dengan menggunakan teknik *data mining* dengan metode klasifikasi. Algoritma yang digunakan untuk melakukan prediksi dalam penelitian ini yaitu algoritma *K-Nearest Neighbor* (KNN). Algoritma ini merupakan algoritma yang melakukan klasifikasi kelas objek baru berdasarkan tetangga terdekatnya. Jumlah data yang digunakan dalam penelitian ini sebanyak 445 data, kemudian membagi data tersebut menjadi 2 yaitu data latih dan data uji. Penelitian ini juga mencakup pengukuran performa yang meliputi nilai akurasi, presisi, *recall*, dan *f-measure* untuk setiap nilai K yang diujikan. Pengukuran ini dilakukan untuk mengetahui parameter yang optimal pada dataset yang digunakan. Adapun hasil yang diperoleh dari pengujian nilai K = 3 hingga K = 9, diperoleh bahwa nilai K = 7 memiliki performa terbaik dengan akurasi tertinggi sebesar 96%, presisi 92%, *recall* 95%, dan *f-measure* 93%.

Kata kunci: prediksi, kualitas udara, *data mining*, *k-nearest neighbor*

PREDIKSI KUALITAS UDARA MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR

Adinda Amalia

Abstract

In everyday life, air is used to breathe by living things. Clean air contains many benefits for life. However, the air found in nature is not always clean, so it can cause a decrease in air quality. Poor air quality can have an impact on human health and the surrounding environment. Air quality in DKI Jakarta can be known through the Indeks Standar Pencemar Udara (ISPU). This study aims to predict the air quality in DKI Jakarta based on ISPU data. Prediction is done using data mining techniques with classification methods. The algorithm used to make predictions in this study is the K-Nearest Neighbor (KNN) algorithm. This algorithm is an algorithm that classifies new object classes based on their closest neighbors. The amount of data used in this study was 445 data, then divided the data into 2, namely training data and test data. This study also includes performance measurements that include the values of accuracy, precision, recall, and f-measure for each tested K value. This measurement is carried out to determine the optimal parameters in the dataset used. The results obtained from testing the values of $K = 3$ to $K = 9$, it was found that the value of $K = 7$ had the best performance with the highest accuracy of 96%, precision of 92%, recall of 95%, and f-measurement of 93%.

Keyword: *classification, air quality, data mining, k-nearest neighbor*