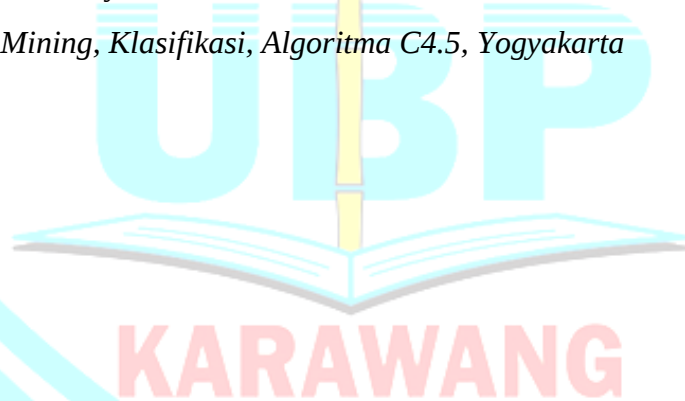


ABSTRAK

Kualitas udara merupakan faktor penting yang memengaruhi kesehatan dan kualitas hidup masyarakat. Yogyakarta, sebagai salah satu kota besar di Indonesia, menghadapi tantangan pencemaran udara yang memerlukan pemantauan dan pengelolaan yang efektif. Penelitian ini bertujuan untuk mengklasifikasikan kualitas udara di Yogyakarta menggunakan algoritma C4.5, salah satu metode pohon keputusan yang efektif dalam klasifikasi data. Data kualitas udara dikumpulkan dari beberapa stasiun pemantauan, mencakup parameter seperti PM2.5, PM10, dan gas polutan lainnya. Data yang diperoleh dianalisis menggunakan algoritma C4.5 untuk mengelompokkan tingkat kualitas udara ke dalam kategori seperti baik dan sedang. Hasil pengujian menunjukkan bahwa model klasifikasi ini memiliki akurasi sebesar 97,20%, recall 100%, precision 95,93%, dan f1-score 97,92% pada data pengujian. Sementara pada data pelatihan, diperoleh akurasi 96,47%, recall 100%, precision 94,77%, dan f1-score 97,31%. Model ini dapat digunakan sebagai alat bantu dalam pengambilan keputusan terkait pengelolaan kualitas udara dan perumusan kebijakan lingkungan di Yogyakarta. Temuan ini diharapkan berkontribusi pada peningkatan kualitas udara dan kesejahteraan masyarakat.

Kata Kunci: *Data Mining, Klasifikasi, Algoritma C4.5, Yogyakarta*



ABSTRACT

Air quality is an important factor affecting people's health and quality of life. Yogyakarta, as one of the big cities in Indonesia, is fighting air pollution that requires effective monitoring and management. This study aims to classify air quality in Yogyakarta using the C4.5 algorithm, one of the effective decision tree methods in data classification. Air quality data were collected from several monitoring stations, covering parameters such as PM2.5, PM10, and other pollutant gases. The data obtained were analyzed using the C4.5 algorithm to group air quality levels into categories such as good and moderate. The test results showed that this model classification had an accuracy of 97.20%, 100% recall, 95.93% precision, and 97.92% f1-score on the testing data. While on the training data, an accuracy of 96.47%, 100% recall, 94.77% precision, and 97.31% f1-score were obtained. This model can be used as a tool in decision making related to air quality management and environmental policies in Yogyakarta. These findings are expected to contribute to improving air quality and public welfare.

Keywords: Data Mining, Clasification, C4.5 Algorithm, Yogyakarta.

