

ABSTRAK

Pencemaran udara merupakan salah satu isu lingkungan yang berdampak penting terhadap masyarakat, ekosistem, serta kualitas hidup. Kota Yogyakarta sebagai wilayah dengan aktivitas kendaraan yang tinggi, pertumbuhan populasi, serta aktivitas pembakaran terbuka memiliki risiko pencemaran udara yang signifikan. Pemerintah telah menetapkan system Indeks Standar Pencemaran Udara (ISPU) melalui Kementerian Lingkungan Hidup dan Kehutanan (KLHK) untuk mengklasifikasikan kualitas udara ke dalam kategori tertentu seperti Baik, Sedang, Tidak Sehat, hingga Berbahaya. Namun, dalam praktiknya, data historis kualitas udara yang diperoleh dari Dinas Lingkungan Hidup (DLH) Kota Yogyakarta masih menghadapi keterbatasan cakupan wilayah dan distribusi kelas ISPU yang tidak seimbang. Keseimbangan ini menimbulkan tantangan dalam membangun model klasifikasi yang adil dan akurat, karena algoritma cenderung bias terhadap kelas mayoritas.

Penelitian ini bertujuan untuk membangun model klasifikasi ISPU berdasarkan data historis kualitas udara Kota Yogyakarta tahun 2023 dan 2024. Data yang digunakan berupa konsentrasi polutan utama, yaitu PM_{2.5}, PM₁₀, SO₂, CO, O₃, dan NO₂. Tahapan penelitian dimulai dengan studi literatur, pengumpulan data, dan pra-pemrosesan yang meliputi pembersihan data, serta penentuan kategori ISPU berdasarkan komponen kritis sesuai standar KLHK. Selanjutnya, untuk mengatasi masalah ketidakseimbangan kelas, digunakan metode *Synthetic Minority Oversampling Technique* (SMOTE). Model klasifikasi dibangun dengan algoritma *Support Vector Machine* (SVM) menggunakan kernel *confusi Function* (RBF). Evaluasi kinerja model dilakukan menggunakan metrik *Precision*, *Recall*, F1-Score, Macro-F1, serta *Confusion Matrix* guna memastikan performa klasifikasi dapat diukur secara adil di setiap kategori ISPU.

Hasil penelitian menunjukkan bahwa penerapan metode SMOTE mampu meningkatkan performa klasifikasi pada kategori minoritas dibandingkan tanpa penanganan ketidakseimbangan. Model SVM dengan kernel RBF berhasil memberikan hasil klasifikasi ISPU yang lebih seimbang, ditunjukkan dengan peningkatan nilai Macro-F1 dan *Recall* pada kelas minoritas. Dengan demikian, penelitian ini memberikan kontribusi dalam membangun model klasifikasi ISPU yang lebih adil, akurat, dan representatif terhadap kondisi pencemaran udara di Yogyakarta. Selain itu, hasil penelitian ini diharapkan dapat mendukung upaya pemantauan kualitas udara secara lebih efektif di tingkat daerah, serta menjadi acuan bagi penelitian lanjutan yang berfokus pada pengolahan data lingkungan dengan pendekatan pembelajaran mesin.

Kata Kunci: Indeks Standar Pencemaran Udara, *Synthetic Minority Oversampling Technique*, *Support Vector Machine*, Klasifikasi, Yogyakarta

ABSTRACT

Air pollution is one of the significant environmental issues impacting society, ecosystems, and quality of life. Yogyakarta City, as an area with high vehicle activity, population growth, and open burning practices, faces a significant risk of air pollution. The government has established the Air Pollution Standard Index (ISPU) system through the Ministry of Environment and Forestry (KLHK) to classify air quality into categories such as Good, Moderate, Unhealthy, and Hazardous. However, in practice, historical air quality data obtained from the Yogyakarta City Environmental Agency (DLH) still faces limitations in regional coverage and imbalanced distribution of ISPU classes. This imbalance poses challenges in building fair and accurate classification models, as algorithms tend to be biased toward majority classes.

This research aims to develop an ISPU classification model based on historical air quality data from Yogyakarta City in 2023 and 2024. The data used includes concentrations of major pollutants, namely PM_{2.5}, PM₁₀, SO₂, CO, O₃, and NO₂. The research stages begin with literature review, data collection, and preprocessing, including data cleaning and determination of ISPU categories based on critical components according to KLHK standards. Next, to address class imbalance issues, the Synthetic Minority Oversampling Technique (SMOTE) method is used. The classification model is built using the Support Vector Machine (SVM) algorithm with the Radial Basis Function (RBF) kernel. Model performance evaluation is conducted using metrics such as Precision, Recall, F1-Score, Macro-F1, and Confusion Matrix to ensure classification performance can be measured fairly across each ISPU category.

The research results show that the application of the SMOTE method can improve classification performance on minority categories compared to without imbalance handling. The SVM model with RBF kernel successfully provides more balanced ISPU classification results, indicated by increased Macro-F1 and Recall values for minority classes. Thus, this research contributes to building a fairer, more accurate, and representative ISPU classification model for air pollution conditions in Yogyakarta. Furthermore, the results of this research are expected to support more effective air quality monitoring efforts at the regional level and serve as a reference for further research focusing on environmental data processing using machine learning approaches.

Keywords: *Air Pollution Standard Index, Synthetic Minority Oversampling Technique, Support Vector Machine, Classification, Yogyakarta*