

Abstrak

Polusi udara adalah salah satu jenis masalah yang dipicu oleh zat-zat polutan udara. Dalam banyak kasus polusi udara dapat menyebabkan beberapa penyakit yang berdampak pada kesehatan jangka panjang. Beberapa tahun terakhir ini telah banyak diajukan metode untuk melakukan deteksi polusi udara. Pada umumnya metode deteksi ini menganalisis zat-zat penyebab polutan udara. Ada 3 tahapan dalam mendeteksi polusi udara dengan machine learning, yaitu *preprocessing*, ekstraksi ciri dan klasifikasi zat-zat yang dapat menyebabkan polusi udara. Nilai akurasi yang diperoleh dari keseluruhan proses deteksi sangat dipengaruhi oleh hasil pada tahap klasifikasi. Oleh karena itu pemilihan algoritma klasifikasi yang tepat menjadi penting. Dari sekian banyak literatur yang mengusulkan metode deteksi polusi udara, banyak diantaranya yang belum melakukan perbandingan algoritma machine learning untuk klasifikasi. Serta kebanyakan dari literatur menggunakan algoritma regresi. Selanjutnya pengembangan *prototype* deteksi polusi udara juga masih jarang yang menggunakan algoritma klasifikasi terbaik dikarnanya kurangnya perbandingan algoritma klasifikasi. Untuk menyelesaikan masalah-masalah di atas, tugas akhir ini mengusulkan implementasi algoritma klasifikasi yang mendukung peningkatan akurasi deteksi polusi udara. Di samping itu *prototype* berdasarkan algoritma yang diusulkan juga dikembangkan. Yang tidak kalah penting tugas akhir ini juga melakukan analisis terhadap akurasi deteksi *prototype* yang diusulkan. Metode yang digunakan dalam penelitian tugas akhir ini adalah 1. Studi literatur tentang deteksi polusi udara, 2. Digunakan algoritma *ensemble learning* untuk melakukan klasifikasi, 3. Pengembangan *prototype*, 4. Pengujian performansi dan analisis. Algoritma adaboost menjadi algoritma terburuk ketika dilakukan klasifikasi dengan akurasi sebesar 62% yang artinya algoritma tersebut kurang baik dalam melakukan klasifikasi polusi udara. Setelah dilakukan hyperparameter tuning algoritma adaboost memiliki akurasi sebesar 97% yang artinya algoritma adaboost mendapatkan peningkatan akurasi sebesar 35%. Algoritma terbaik dalam melakukan klasifikasi polusi udara adalah algoritma random forest dengan akurasi 98%.

Kata Kunci: Ensemble learning, klasifikasi, polusi udara.