

BAB V PENUTUP

5.1 Kesimpulan

1. Penggunaan metode imputasi Fillna dalam menangani missing value pada dataset kualitas udara memberikan peningkatan akurasi pada semua algoritma yang diuji. Dengan mengurangi jumlah data yang hilang, proses klasifikasi dapat berjalan lebih optimal, sehingga model memiliki informasi yang lebih lengkap untuk melakukan prediksi.
2. Metode resampling SMOTE yang diterapkan berhasil meningkatkan keseimbangan antara precision dan recall, khususnya pada kategori data dengan jumlah sampel lebih sedikit. Hasil evaluasi menunjukkan bahwa setelah penerapan SMOTE, akurasi model meningkat, dan model menjadi lebih efektif dalam mengklasifikasikan data dari kategori yang lebih kecil.
3. Dari hasil evaluasi model terhadap empat algoritma yang diuji, *Random Forest* menunjukkan performa terbaik dengan akurasi tertinggi sebesar 98,33% setelah diterapkannya SMOTE dan *tuning hyperparameter*. Selain itu, *Random Forest* juga memiliki nilai *precision*, *recall*, dan *F1-score* yang lebih tinggi dibandingkan algoritma lainnya.

5.2 Saran

1. Penelitian selanjutnya dapat mengeksplorasi metode machine learning yang lebih canggih serta menganalisis dampak kualitas udara terhadap kesehatan masyarakat.
2. Disarankan untuk membandingkan Regresi Logistik dengan algoritma lain seperti XGBoost atau LightGBM guna meningkatkan akurasi dan efisiensi model.

3. Hasil klasifikasi kualitas udara di DKI Jakarta dapat digunakan dalam sistem pemantauan kualitas udara berbasis AI serta mendukung kebijakan pengendalian polusi yang dikeluarkan oleh pemerintah dan instansi pengendali polusi, seperti penerapan zona emisi rendah dan pembatasan kendaraan di area dengan tingkat polusi tinggi.

