

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan implementasi *model machine learning* pada prediksi emisi CO₂ di Kanada, maka dapat diambil kesimpulan sebagai berikut:

1. Berdasarkan hasil evaluasi performa tujuh algoritma regresi sebelum dilakukan proses *tuning*, algoritma *Random Forest Regressor* menunjukkan kinerja prediktif paling optimal. Hal ini ditunjukkan oleh nilai R^2 sebesar 1.000, nilai *RMSE* sebesar 2.510, *MAE* sebesar 1.700, dan *MAPE* sebesar 0.01 (1%) pada data uji. Nilai-nilai tersebut menandakan bahwa model mampu memprediksi emisi dengan akurasi tinggi dan kesalahan minimal dibandingkan model lainnya.
2. Fitur yang memiliki pengaruh paling signifikan terhadap prediksi emisi adalah konsumsi bahan bakar gabungan (*Fuel Consumption Comb*) yang ditunjukkan melalui analisis *feature importance* dari model *Random Forest*. Hasil ini mengindikasikan bahwa efisiensi bahan bakar merupakan determinan utama dalam tingkat emisi karbon kendaraan.
3. Setelah dilakukan proses *hyperparameter tuning* menggunakan kombinasi metode *RandomizedSearchCV* dan *GridSearchCV*, model *Random Forest* mengalami peningkatan dari segi kestabilan dan kemampuan generalisasi. Hasil evaluasi model setelah *tuning* menunjukkan R^2 sebesar 0.9982, *RMSE* sebesar 2.4898, *MAE* sebesar 1.6928, dan *MAPE* sebesar 0.0068 (0.68%). Penurunan nilai R^2 yang sangat kecil diimbangi dengan penurunan metrik error lainnya, menandakan bahwa *tuning* berhasil mengurangi risiko *overfitting* tanpa menurunkan akurasi secara signifikan. Parameter terbaik yang diperoleh mencakup $n_estimators = 200$, $max_depth = 20$, dan $min_samples_leaf = 1$.
4. Model prediksi yang dikembangkan telah berhasil diimplementasikan dalam aplikasi web berbasis *Streamlit*, yang menyediakan antarmuka interaktif untuk prediksi emisi kendaraan. Aplikasi ini diharapkan dapat dimanfaatkan oleh masyarakat umum, produsen kendaraan, dan pembuat

kebijakan sebagai alat bantu dalam mendukung upaya pengurangan emisi karbon di sektor transportasi.

5.2 Saran

Penelitian ini masih memiliki beberapa keterbatasan, sehingga perlu dilakukan pengembangan lebih lanjut agar hasilnya dapat lebih optimal. Beberapa saran yang dapat diberikan oleh penulis antara lain sebagai berikut:

1. Perluasan *dataset* dapat dilakukan dengan menambahkan data kendaraan keluaran terbaru, termasuk jenis kendaraan hybrid dan listrik. Hal ini penting agar model dapat lebih relevan terhadap perkembangan teknologi otomotif saat ini.
2. Menambahkan variabel dinamis seperti pola mengemudi, kondisi lalu lintas, atau lingkungan jalan guna meningkatkan akurasi prediksi pada kondisi nyata di lapangan.
3. Dari sisi pemodelan, pendekatan *ensemble* atau kombinasi beberapa algoritma *machine learning* dapat dieksplorasi untuk memperoleh model yang lebih kuat dan memiliki generalisasi yang lebih baik.
4. Validasi model sebaiknya diperluas menggunakan data dari sumber lain atau data pengukuran emisi aktual, agar model tidak hanya bekerja baik pada data uji yang tersedia, tetapi juga pada data baru yang merepresentasikan kondisi nyata.
5. Penelitian lanjutan juga disarankan untuk menggunakan model ini sebagai alat bantu dalam simulasi kebijakan, misalnya untuk mengukur dampak penerapan insentif kendaraan rendah emisi atau regulasi terhadap kendaraan dengan emisi tinggi.