

BAB 5

SIMPULAN DAN SARAN

5.1 Simpulan

Penelitian ini bertujuan untuk membandingkan performa dua algoritma *ensemble learning* yaitu *random forest* dan *XGBoost* dalam memprediksi kategori penyebab kecelakaan lalu lintas yang dikategorikan menjadi *human error* dan *non-human error*. Proses pengembangan model dilakukan melalui dua tahap, yaitu pembangunan model awal dengan parameter default serta optimasi menggunakan metode *GridSearchCV*. Evaluasi performa dilakukan berdasarkan akurasi, metrik-metrik yang didapatkan dari *confusion matrix*, serta laporan klasifikasi untuk mengukur efektivitas dan kemampuan generalisasi dari masing-masing model terhadap data pengujian. Berdasarkan hasil evaluasi, diperoleh kesimpulan sebagai berikut:

1. Dengan *hyperparameter tuning* menggunakan *gridsearchcv*, model *random forest* menunjukkan akurasi sebesar 92,5%. Model *XGBoost* memperoleh akurasi sebesar 92,49%.
2. Berdasarkan hasil *confusion matrix*, model *random forest* berhasil mengklasifikasikan 174 *true positive*, 27.754 *true negative*, jumlah *false positive* sebesar 2.180, dan 85 *false negative*. Model *XGBoost* mencatat 185 *true positive*, 27.740 *true negative*, *false positive* dan 99 *false negative*.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, beberapa saran yang dapat diberikan untuk penelitian selanjutnya antara lain:

1. Menggabungkan pendekatan *ensemble* lain seperti *voting classifier* atau *stacking* untuk melihat apakah kombinasi beberapa model dapat memberikan hasil yang lebih baik dibandingkan model individual, seperti pada penelitian [20].
2. Memperluas ruang lingkup data dengan menambahkan informasi seperti lokasi kecelakaan untuk memberikan konteks tambahan dalam proses prediksi.