

BAB V

PENUTUP

5.1 SIMPULAN

Berdasarkan hasil penelitian dan pengujian model ada beberapa point penting yang dapat disimpulkan :

- a. Model berhasil mendapatkan akurasi sebesar 91.3% setelah percobaan beberapa parameter, dan untuk parameter grid terbaik ialah '*max_depth*': 50, '*max_features*': '*sqrt*', '*min_samples_leaf*': 4, '*min_samples_split*': 20 dengan '*n_estimator*' : 500. Dan untuk hasil klasifikasi mendapatkan nilai yang cukup tinggi dengan kelas FATAL INJURY mendapatkan precision 1.00 , recall 1.00 dan f1-score 1.00. NO APPARENT INJURY mendapatkan precision 0.95, recall 0.90 dan f1-score 0.92. POSSIBLE INJURY mendapatkan precision 0.82, recall 0.84 dan f1-score 0.83. SUSPECTED MINOR INJURY mendapatkan precision 0.86, recall 0.85 dan f1-score 0.85. SUSPECTED SERIOUS INJURY mendapatkan precision 0.96, recall 0.99 dan f1-score 0.98.
- b. Feature Importance yang diperoleh dari model Random forest menjelaskan bahwa fitur yang paling berpengaruh dalam model adalah ACRS Report Type diikuti Vehicle Damage Extent. Dua faktor ini memiliki bobot kepentingan yang lebih tinggi dibandingkan fitur lainnya.
- c. Faktor-faktor utama yang berpengaruh terhadap tingkat keparahan cedera akibat kecelakaan lalu lintas dengan menggunakan analisis korelasi non-parametrik (Spearman) untuk menentukannya ada 14 fitur yang paling berpengaruh. Dari 42 fitur yang ada 8 diantaranya dihapus atau disingkirkan. Dikarenakan fitur-fitur tersebut tidak memiliki hubungan langsung dengan tingkat keparahan cedera. "*Report Number*", "*Local Case Number*", "*Agency Name*": Merupakan informasi administratif yang bersifat unik untuk setiap laporan dan tidak berkontribusi pada pola

kecelakaan. "*Person ID*", "*Vehicle ID*": Hanya berfungsi sebagai identifikasi individu atau kendaraan dan tidak relevan dalam mempengaruhi variabel target. "*Latitude*", "*Longitude*", "*Location*": Meskipun terkait dengan lokasi kejadian, fitur ini tidak dapat digunakan secara langsung dalam model karena bersifat spesifik terhadap titik koordinat dan dapat menyebabkan overfitting.

5.2 SARAN

Penelitian ini masih memiliki beberapa keterbatasan yang perlu diperbaiki terutama dalam hal penanganan ketidakseimbangan data dan keakuratan model dalam memprediksi. Meskipun berbagai teknik diterapkan hasil yang diperoleh menunjukkan bahwa masih ada ruang untuk perbaikan baik dari segi metode, fitur yang digunakan maupun teknik analisis yang diterapkan. Penelitian selanjutnya disarankan untuk menambahkan jumlah data agar model dapat belajar dari lebih banyak variasi kasus dan meningkatkan generalisasi prediksi dan juga eksplorasi teknik penyeimbangan data yang lebih baik serta pemilihan fitur yang lebih relevan juga dapat dilakukan untuk meningkatkan akurasi dan keandalan model.