

INTISARI

PERBANDINGAN METODE *MACHINE LEARNING* DAN *ENSEMBLE LEARNING* DALAM KLASIFIKASI TINGKAT LUKA KORBAN KECELAKAAN

(Studi Kasus : Data Kecelakaan Lalu Lintas di Polres Kudus Tahun 2019 - 2021)

Freditasari Purwa Hidayat

Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Islam Indonesia

Kecelakaan lalu lintas merupakan kejadian yang tidak terduga dan tidak disengaja yang melibatkan kendaraan bermotor maupun pejalan kaki. Berdasarkan data pada Badan Pusat Statistik (BPS), kabupaten Kudus merupakan salah satu kabupaten di Jawa Tengah yang memiliki tingkat kasus kecelakaan yang tergolong tinggi. Kasus kecelakaan di kabupaten Kudus termasuk tinggi karena lebih besar dibandingkan kabupaten lainnya dengan kasus 964 pada tahun 2019, 831 kasus pada tahun 2020 dan 841 kasus pada tahun 2021. Kasus kecelakaan yang tinggi mengakibatkan korban mengalami luka ringan, luka berat, maupun meninggal dunia. Oleh karena itu, peneliti ingin mengetahui prediksi dari klasifikasi tingkat luka korban kecelakaan di kabupaten Kudus, dan mengidentifikasi faktor apa yang paling berpengaruh dalam penentuan tingkat luka korban. Menggunakan metode klasifikasi *machine* dan *ensemble learning* dengan menggabungkan metode klasifikasi seperti *decision tree*, dan *naïve bayes*. Hasil klasifikasi tunggal menggunakan data sebelum *oversampling*, didapatkan akurasi metode *decision tree* sebesar 95% dan *naïve bayes* sebesar 97%, sedangkan menggunakan data setelah *oversampling* didapatkan hasil akurasi metode *decision tree* sebesar 99% dan *naïve bayes* sebesar 93%. Sedangkan metode *ensemble learning* pada algoritma klasifikasi didapatkan akurasi sebesar 95% menggunakan data sebelum *oversampling* dan 93% setelah *oversampling*. Berdasarkan simulasi sebanyak 50 dan 100 kali dapat di diketahui bahwa pada penelitian ini, metode *ensemble* dengan *oversampling* dan tanpa *oversampling* tidak dapat menaikkan akurasi dalam klasifikasi. Dari seluruh variabel *independent*, variabel jenis kecelakaan merupakan variabel yang paling besar pengaruhnya dalam memprediksi tingkat luka korban kecelakaan lalu lintas, sedangkan variabel jenis kelamin merupakan variabel yang memiliki pengaruh paling kecil.

Kata Kunci : *Decision Tree*, *Ensemble*, Kecelakaan, Klasifikasi, *Naïve Bayes*