

BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan hasil penelitian terhadap data kecelakaan lalu lintas Polda Jatim tahun 2024, Faktor-faktor utama yang mempengaruhi tingkat kecelakaan lalu lintas adalah cuaca, kondisi cahaya, dan permukaan jalan. Berdasarkan analisis *feature importance* Decision Tree, cuaca memiliki kontribusi paling dominan sebesar 95% terhadap prediksi klasifikasi. Hasil korelasi menggunakan *heatmap* menampilkan korelasi positif Tingkat kecelakaan dengan Cuaca (+0.49) dan Kondisi Permukaan Jalan (+0.51). Dapat diartikan, semakin rusak atau licin kondisi permukaan dan buruknya cuaca pada saat berkendara dapat meningkatkan tingkat kecelakaan. Sedangkan korelasi negatif pada Kondisi cahaya (-0.39) menandakan semakin terang kondisi cahaya dapat mengurangi tingkat kecelakaan tinggi.

Analisis klasifikasi menggunakan model SVM, KNN, Decision Tree, dan Naïve Bayes, dengan hasil akurasi tertinggi pada model SVM dan Decision Tree yang mendapatkan akurasi 90.1%. Pada model KNN mendapatkan akurasi di 90% dan pada model Naïve Bayes mendapatkan akurasi di angka 87.9%. Dataset memiliki angka data kecelakaan mencapai 28.000 data, keseluruhan model menggunakan Teknik *sampling* SMOTE, ROS, dan RUS, namun performa yang didapatkan mengalami *overfitting* dan *underfitting* terhadap performa model. Berdasarkan akurasi yang didapatkan, model yang telah ditetapkan untuk melaksanakan klasifikasi tingkat kecelakaan pada dataset Polda Jatim tahun 2024 merupakan model SVM dan Decision Tree tanpa *sampling*.

Model dengan akurasi terbaik dikembangkan menjadi aplikasi *web* menggunakan Streamlit. Menampilkan hasil prediksi dengan UI sederhana dan interaktif dalam memprediksi tingkat kecelakaan berdasarkan faktor cuaca, kondisi cahaya, dan permukaan jalan. Hasil yang ditampilkan berupa prediksi

yang diuji menggunakan model SVM dan Decision Tree yang telah dikembangkan. Akurasi yang dituju terkait kasus ini, membagikan hasil prediksi menjadi 3 kategori, dengan akurasi diatas 70% dinyatakan Darurat, diatas 35% diharuskan untuk dilakukan pengecekan manual, dan dibawah 35% merupakan tingkat kecelakaan Ringan. Menggunakan Streamlit juga, dikembangkan aplikasi yang dapat melatih model, dan melakukan klasifikasi terhadap tingkat kecelakaan ringan secara berkala. Model ini akan di-*input* data untuk dilaksanakan pelatihan, lalu berkala untuk dilatih hingga dapat mengenal pola tingkat kecelakaan secara maksimal.

5.2 Saran

Beberapa saran dapat dipertimbangkan untuk mengangkat topik serupa pada penelitian selanjutnya, dengan harapan dapat memberikan hasil pengembangan model terhadap kasus kecelakaan yang lebih baik.

1. Eksplorasi dan menggunakan model yang lebih kompleks untuk membaca data kecelakaan lalu lintas, dapat juga mencoba menggunakan model hybrid seperti CNN-SVM.
2. Menggunakan dataset yang mencakup wilayah lain di Indonesia atau secara nasional, juga mencari data dengan fitur pendukung yang lebih banyak, seperti usia pengemudi, kendala kendaraan, dan faktor lainnya.
3. Eksplorasi Teknik augmentasi atau *sampling* data terhadap *imbalanced data*, perbandingan dapat mencoba teknik sampling Cluster Centroids dan ADASYN agar dapat menentukan teknik yang lebih adaptif terhadap dataset dengan ketidakseimbangan yang cukup signifikan.
4. Pengembangan aplikasi web dapat dijadikan referensi terhadap instansi untuk dijadikan proses laporan masyarakat terhadap kasus kecelakaan atau melaksanakan pendataan secara otomatis. Dapat juga mengembangkan aplikasi pada penelitian selanjutnya untuk kolaborasi dengan instansi agar teruji dan dapat menjadi kontribusi nyata dalam membantu pencegahan peningkatan angka kecelakaan.