

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam era teknologi informasi saat ini, data menjadi salah satu aset yang paling berharga bagi banyak organisasi. Pemanfaatan data untuk pengambilan keputusan telah menjadi suatu keharusan, terutama dalam industri otomotif (Nishfi et al., 2023). Dengan berkembangnya teknologi, klasifikasi otomatis terhadap data menjadi semakin penting, terutama dalam konteks industri otomotif untuk memahami tren pasar, preferensi konsumen, dan faktor-faktor lain yang mempengaruhi pengambilan keputusan terkait desain dan pemasaran mobil (Aldiyansyah et al., 2024). Algoritma klasifikasi, seperti Decision Tree dan Random Forest, telah terbukti efektif dalam menangani tugas-tugas klasifikasi semacam ini. Oleh karena itu, penelitian ini bertujuan untuk membandingkan kinerja kedua algoritma tersebut dalam mengklasifikasikan jenis transmisi mobil berdasarkan atribut-atribut tertentu, seperti model, fuel type, dan manufacturer, dengan menggunakan data mobil dari tahun 1970 hingga 2024.

Dalam konteks penelitian ini, dataset “Cars Dataset from 1970 to 2024” dipilih karena menyediakan informasi yang lengkap dan terstruktur tentang berbagai atribut mobil. Pemilihan atribut-atribut tertentu seperti model, fuel type, dan manufacturer didasarkan pada asumsi bahwa atribut-atribut ini dapat menjadi indikator penting dalam menentukan jenis transmisi mobil. (Nugraha Kasmara et al., n.d.) Selain memberikan wawasan tambahan tentang faktor-faktor yang memengaruhi pilihan transmisi konsumen, penelitian ini diharapkan dapat memberikan pemahaman yang lebih mendalam mengenai perbandingan kinerja antara algoritma Decision Tree dan Random Forest dalam klasifikasi jenis transmisi mobil. Temuan penelitian ini dapat bermanfaat bagi strategi pemasaran dan desain produk di industri otomotif.

1.2 Identifikasi Masalah

Berikut beberapa identifikasi masalah dari latar belakang diatas;

1. Kurangnya perbandingan akurasi algoritma Decision Tree dan Random Forest dalam klasifikasi transmisi mobil.
2. Metode yang digunakan sebelumnya belum menghasilkan hasil yang optimal.

Menjadi dasar permasalahan yang perlu diteliti dalam penelitian ini.

1.3 Tujuan Penelitian

Pada penelitian yang sedang dilakukan ini terdapat beberapa yang tujuan akan dicapai seperti berikut:

1. Menilai kinerja algoritma Decision Tree dan Random Forest berdasarkan metrix evaluasi seperti akurasi, recall, precision, dan F1-score.
2. Menentukan algoritma mana yang memiliki akurasi lebih baik dari kedua algoritma tersebut dalam klasifikasi transmisi mobil.

1.4 Batasan Masalah

Dari permasalahan yang telah dikemukakan, diperlukan penentuan batasan masalah agar dapat menghindari penyebaran topik yang terlalu luas, dan untuk itu batasan masalah diperlukan sebagai berikut:

1. Mengidentifikasi pengolahan data untuk mengklasifikasi transmisi mobil menggunakan algoritma Decision Tree dan Random Forest.
2. Penulisan code menggunakan software google colab, bahasa yang digunakan adalah *Python*.
3. Metode yang digunakan pada penelitian kali ini yaitu Decision Tree dan Random Forest.

1.5 Kontribusi

Penelitian ini diharapkan dapat memberikan kontribusi yang signifikan di sejumlah bidang. Pertama, perbandingan performa Decision Tree dan Random Forest dalam mengklasifikasikan tipe transmisi mobil dapat mengungkap kelebihan dan kekurangan masing-masing algoritma pada setting ini. Praktisi dan peneliti dapat menggunakan data ini untuk memilih algoritma terbaik untuk tugas klasifikasi serupa di industri otomotif dan bidang lainnya. Selain itu, analisis fitur berbasis informasi timbal balik dapat memberikan tambahan wawasan mengenai karakteristik yang memiliki dampak terbesar pada jenis transmisi mobil. Di masa depan, model prediksi yang lebih akurat dan efektif dapat dibangun berdasarkan hal ini. Faktor-faktor yang mempengaruhi keputusan klasifikasi dapat lebih dipahami melalui visualisasi struktur model Pohon Keputusan, yang dapat digunakan sebagai landasan untuk meningkatkan pengambilan keputusan otomotif. Oleh karena itu, temuan penelitian ini diharapkan dapat memajukan teknologi klasifikasi data secara signifikan dan penggunaannya dalam industri otomotif dan bidang terkait lainnya..

