

ABSTRAK

Industri otomotif Indonesia mengalami peningkatan signifikan dalam kehadiran mobil listrik buatan Cina seperti Wuling, BYD, Aion, dan Denza yang menawarkan harga kompetitif. Fenomena ini memicu tingginya aktivitas diskusi di media sosial, khususnya YouTube, di mana masyarakat memberikan beragam pendapat mengenai kualitas, performa, dan keandalan produk tersebut. Banyaknya komentar yang berjumlah besar dan tidak terstruktur menimbulkan tantangan bagi pelaku industri maupun peneliti untuk memahami kecenderungan opini publik secara manual. Oleh karena itu, penelitian ini dilakukan untuk mengevaluasi kinerja kombinasi metode vektorisasi dan algoritma machine learning dalam mengklasifikasikan sentimen masyarakat terhadap mobil listrik buatan Cina.

Penelitian ini menggunakan pendekatan text mining yang mencakup proses pengumpulan data dari kolom komentar YouTube, pelabelan data secara manual, serta tahapan pra-pemrosesan seperti pembersihan teks, normalisasi, tokenisasi, penanganan emoji, dan penghapusan stopword. Setelah data diproses, beberapa metode vektorisasi teks diterapkan untuk mengubah komentar menjadi representasi numerik. Metode-metode vektorisasi tersebut kemudian dikombinasikan dengan beberapa algoritma machine learning untuk membangun berbagai model klasifikasi. Setiap kombinasi model dievaluasi menggunakan pembagian data latih dan data uji untuk mengidentifikasi model yang paling optimal dalam menangani karakteristik teks komentar berbahasa Indonesia.

Hasil penelitian menunjukkan adanya variasi performa antara setiap kombinasi metode vektorisasi dan algoritma machine learning. Melalui pengujian menggunakan metrik akurasi, presisi, recall, dan f1-score, ditemukan bahwa efektivitas model sangat bergantung pada kesesuaian antara metode representasi teks dan algoritma klasifikasi yang digunakan. Secara umum, masyarakat menunjukkan sentimen yang terpolarisasi terhadap mobil listrik Cina, dengan aspek harga menjadi pendorong sentimen positif sementara isu infrastruktur pengisian daya memicu sentimen negatif. Penelitian ini memberikan kontribusi berupa analisis komprehensif mengenai kinerja berbagai kombinasi metode dalam analisis sentimen serta menyediakan dasar teknis yang dapat dimanfaatkan oleh industri, peneliti, dan pembuat kebijakan dalam memahami persepsi publik secara lebih akurat.

Kata Kunci: Analisis Sentimen, Mobil Listrik Cina, Vektorisasi Teks, Algoritma Machine Learning, Evaluasi Model