

BAB I

PENDAHULUAN

1.1 Latar Belakang

Di masa ini, ditandai dengan kemajuan teknologi yang cepat, masyarakat dapat merasakan dampak transformasi dalam banyak aspek kehidupan, termasuk sektor otomotif. Salah satu inovasi hebat di area ini adalah kendaraan listrik, kendaraan yang menggunakan motor listrik sebagai pengemudi utama dan baterai sebagai sumber energi [1]. Kendaraan listrik memiliki banyak keunggulan, termasuk efisiensi energi yang tinggi dan kontribusi positif bagi lingkungan. Sebagai moda transportasi lain, kendaraan listrik memainkan peran penting dalam mengurangi ketergantungan mereka pada energi fosil dan tidak menghasilkan emisi buang, membuat kebutuhan mobilitas masyarakat lebih ramah lingkungan.

Salah satu dampak menggunakan kendaraan listrik adalah mengurangi polusi udara yang disebabkan oleh kendaraan energi fosil. Melihat kondisi udara saat ini di Indonesia, kualitasnya pada tingkat yang mengkhawatirkan. Pada awal 2025, berbagai kota dan distrik di Indonesia mengalami variasi kualitas udara berdasarkan kategori sedang berdasarkan konsentrasi PM2.5 [2]. Konsentrasi PM2.5 yang tinggi sangat berbahaya karena partikel-partikel ini dapat dengan mudah memasuki sistem pernapasan dan dapat menyebabkan berbagai masalah kesehatan [2]. Dalam konteks ini, transisi ke kendaraan listrik dikaitkan dengan salah satu solusi potensial untuk mengurangi emisi buang dan meningkatkan kualitas udara di Indonesia.



Gambar 1.1 Indeks kualitas udara terburuk di dunia [3]

Pada awal tahun 2023 lalu, Indonesia sempat beberapa kali berada pada daftar 10 besar kota yang memiliki kualitas udara terburuk di dunia. Dari data di atas, menunjukkan bahwa pada Juni 2023, ibukota Indonesia pada tanggal 10 dengan jumlah AQI adalah 128, bahkan mencapai 155 indeks kualitas udara dan menjadi negara terburuk di timur [3]. Situasi ini dipengaruhi oleh sejumlah faktor, salah satunya adalah kepadatan lalu lintas, terutama di kota-kota besar di Indonesia. Kualitas udara yang buruk dan dampak negatif pada kesehatan manusia [3].

Untuk meminimalkan dampak negatif dari penggunaan kendaraan bertenaga bensin untuk lingkungan, beralih ke penggunaan pengguna listrik untuk beroperasi dapat menjadi salah satu solusi untuk masalah polusi. Menggunakan kendaraan listrik, mungkin merupakan langkah pertama pemerintah untuk mengurangi polusi udara, seperti yang ditentukan dalam Peraturan Presiden (Perpres) no 55 tahun 2019, diskusikan kemampuan untuk mempercepat mobil listrik berdasarkan baterai [4]. Kendaraan telah dipromosikan oleh pemerintah Indonesia sejak 2015 di National Electrical

Program (*National Electric Vehicle*) yang diluncurkan oleh Kementerian Industri Indonesia [4].

Selain itu, terdapat intensif atau keuntungan yang ditawarkan oleh pemerintah bagi masyarakat yang ingin melakukan pembelian kendaraan listrik, seperti memberikan kebebasan pajak kendaraan bermotor (PKB) untuk penggunaan lima tahun pertama dan bebas Pajak Penjualan atas Barang Mewah (PPnBM) selama sepuluh tahun pertama [5]. Dengan pembebasan pajak yang ditawarkan oleh Pemerintah, diharapkan mampu meningkatkan daya tarik masyarakat untuk beralih dari pengguna kendaraan konvensional menjadi pengguna kendaraan listrik.

Tabel 1.1 Mobil listrik terlaris sepanjang tahun 2024

No	Merek	Jenis Mobil	Unit Terjual
1	BYD	M6	6.124
2	Wuling	Binguo EV	5.156
3	BYD	Seal	4.828
4	Wuling	Air EV	4.440
5	Chery	Omoda E5	4.425
6	Wuling	Cloud EV	3.521
7	BYD	Atto 3	3.270
8	Morris Garages (MG)	MG4 EV	2.340
9	Hyundai	Ioniq 5	1.516
10	Hyundai	Kona Electric	1.196

Penjualan mobil listrik di Indonesia mengalami peningkatan pesat pada tahun 2024, meningkat 153% dibandingkan tahun 2023. BYD M6 menjadi mobil listrik terlaris di Indonesia pada tahun 2024 dengan penjualan sebanyak 6.124 unit terjual, diikuti oleh Wuling Binguo EV sebanyak 5.156 unit dan BYD Seal dengan 4.828 unit [6]. Merek-merek asal Tiongkok seperti BYD, Wuling, dan Chery mendominasi pasar berkat kombinasi desain modern, teknologi canggih, dan harga yang kompetitif. Minat masyarakat terhadap mobil listrik semakin tinggi, terlihat dari permintaan yang terus meningkat, terutama untuk model yang menawarkan efisiensi energi, jangkauan baterai

yang lebih baik, serta fitur-fitur inovatif yang sesuai dengan kebutuhan konsumen di Indonesia. Meskipun demikian, pangsa pasar mobil listrik masih relatif kecil, yaitu sekitar 4,98% dari total penjualan mobil nasional [6].

Kendaraan listrik menawarkan kenyamanan bagi penggunanya selain karena ramah lingkungan, kendaraan listrik juga memiliki tingkat kebisingan mesin yang rendah, dan para pengguna kendaraan listrik dapat lebih menghemat biaya karena tidak perlu mengisi bahan bakar bensin lagi melainkan memanfaatkan baterai sebagai daya penyimpanan listrik [6]. Namun kehadiran kendaraan listrik masih menuai pro dan kontra di tengah masyarakat, banyak masyarakat yang berpendapat bahwa untuk beralih menggunakan mobil listrik masih perlu persiapan yang matang, terutama dalam hal infrastruktur, banyak masyarakat yang memperdebatkan dari segi harga kendaraan listrik yang lebih mahal daripada kendaraan konvensional, tempat pengisian daya untuk kendaraan listrik yang masih minim di Indonesia, dan jarak tempuh serta kecepatan laju kendaraan yang lebih terbatas [7]. Hal ini yang menjadi tolak ukur untuk melakukan penelitian analisis sentimen ini.

Analisis sentimen merupakan suatu proses analisa yang dilakukan mengenai suatu objek atau kejadian menggunakan algoritma tertentu. Berdasarkan referensi penelitian sebelumnya, analisis sentimen menggunakan algoritma *Naïve Bayes* menjadi salah satu pilihan algoritma yang digunakan karena algoritma tersebut lebih efisien ketika dilakukan komputasi, sehingga lebih sesuai untuk diterapkan pada penelitian analisis sentimen yang memiliki data dalam jumlah besar [8]. Pada penelitian dengan topik Penerapan *Naive Bayes Classifier*, *K-Nearest Neighbor* (KNN) dan *Decision Tree* untuk Menganalisis Sentimen pada Interaksi Netizen dan Pemeritah, hasil pemodelan algoritma *Naïve Bayes* diperoleh akurasi terbesar yaitu 100%, KNN sebesar 98,25%, dan DT memberi hasil paling kecil yaitu 62,28% [9]. Penelitian lainnya terkait Komparasi Algoritma *Support Vector Machine* dengan *Naive Bayes* Untuk Analisis Sentimen Pada Aplikasi BRImo, didapatkan akurasi algoritma SVM sebesar 97,69% dan algoritma NB sebesar 96,53% [10]. Dari ketiga penelitian terdahulu, terdapat 3 algoritma dengan hasil memuaskan yaitu NB, SVM, dan

KNN. Ketiga algoritma tersebut yang akan digunakan pada penelitian ini, karena algoritma tersebut dapat membantu dalam melakukan prediksi pada peluang di masa yang akan datang berdasarkan kejadian atau pengalaman yang ada sebelumnya.

Penelitian akan dilakukan dengan melakukan eksplorasi informasi dari data yang diambil melalui media sosial Youtube, data yang diambil berupa tanggapan dari pengguna Youtube. Berdasarkan hasil survei yang telah dilakukan oleh Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) pada tahun 2022 kuartal, Youtube menjadi salah satu media sosial yang paling sering digunakan dengan persentase sebesar 63,02% [12]. Hal ini menjadikan YouTube sebagai sumber data yang sangat kaya dan representatif. *Scraping data* ini dilakukan untuk mendukung analisis sentimen yang dilakukan dengan menganalisa tanggapan dari masyarakat Indonesia terhadap video *review* mengenai kendaraan listrik yang dibahas oleh beberapa konten *creator* [11].

Dalam membantu mengolah dan memahami banyaknya komentar dari pengguna secara efektif, diperlukan teknologi yang mampu bekerja secara otomatis dan efisien. *Machine learning* memainkan peran penting dalam analisis review dengan kemampuan untuk memproses dan menganalisis data dalam skala besar secara cepat dan akurat. Algoritma *machine learning* seperti *Naive Bayes*, *Support Vector Machine (SVM)*, dan KNN dapat digunakan untuk mengklasifikasikan review menjadi sentimen positif, negatif, atau netral [13]. Hal ini membantu perusahaan memahami opini pelanggan dan membuat keputusan strategis berdasarkan umpan balik yang akurat. Selain itu, *machine learning* dapat mengidentifikasi pola dalam review yang mungkin tidak terlihat secara manual, seperti mengenali topik atau tema tertentu yang sering dibahas oleh pelanggan.

Algoritma *machine learning* dapat digunakan untuk melakukan analisis review secara otomatis dan *real-time*, mengurangi waktu dan biaya yang diperlukan untuk memproses data secara manual. Ini memungkinkan perusahaan untuk merespons isu-isu yang muncul dengan lebih cepat dan

efektif [14]. Lebih lanjut, *machine learning* dapat meningkatkan akurasi analisis review dengan mengurangi bias manusia dan memproses data dalam jumlah besar tanpa kelelahan, memastikan bahwa hasil analisis lebih objektif dan dapat diandalkan. Dalam konteks analisis review mobil listrik, *machine learning* dapat membantu memahami sentimen masyarakat terhadap fitur-fitur tertentu, seperti performa, efisiensi energi, dan desain, yang sangat berharga bagi produsen untuk meningkatkan kualitas produk dan strategi pemasaran mereka [15].

Penelitian ini dilakukan dengan tujuan untuk mengetahui tingkat akurasi yang didapatkan dan polaritas dari sentimen mana yang memiliki nilai tertinggi. Bahasa pemrograman yang akan digunakan adalah python untuk mengumpulkan data (*scraping data*) melalui video dari beberapa youtube channel yang membahas tentang kendaraan listrik menggunakan akses API (*Application Programming Interface*) [16]. Dari data yang telah terkumpul, berikutnya akan dilakukan pembersihan dan diklasifikasikan menjadi sentimen positif dan sentimen negatif, serta kesimpulan dari keunggulan dan kekurangan kendaraan listrik.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dari penelitian yang dilakukan adalah sebagai berikut:

1. Bagaimana proses pelabelan data sentimen dapat dilakukan secara otomatis menggunakan pendekatan *lexicon-based* dengan bantuan *library TextBlob* pada ulasan mobil listrik di YouTube menjadi 3 kelas sentimen positif, netral, dan negatif?
2. Bagaimana hasil evaluasi performa algoritma *Naïve Bayes*, *Support Vector Machine*, dan *K-Nearest Neighbours* dengan teknik *SMOTE* dan *grid search* terhadap peningkatan akurasi, presisi, *recall*, dan *f1-score* masing-masing model?

3. Bagaimana proses penerapan hasil *text preprocessing*, *text labeling*, dan visualisasi dapat dikembangkan ke dalam aplikasi berbasis web menggunakan *Streamlit* sebagai bentuk *deployment*?

1.3 Batasan Masalah

Berdasarkan pada penjelasan latar belakang sebelumnya, berikut ini merupakan rumusan masalah pada penelitian ini:

1. Sumber data yang digunakan berupa komentar pada video *review* mobil listrik di youtube.
2. Video ulasan yang digunakan sebagai sumber dataset dalam penelitian ini dibatasi pada konten yang diproduksi di Indonesia.
3. Pemilihan jenis mobil listrik yang di *review* sesuai dengan data penjualan mobil listrik tertinggi yaitu Wuling Binguo EV dan BYD M6.
4. Pengumpulan data pada video ulasan di *filter* berdasarkan urutan penonton terbanyak.
5. Data yang digunakan menggunakan bahasa Indonesia.
6. Klasifikasi akan dibagi menjadi tiga kelas yaitu sentimen positif, netral, dan negatif.
7. Menggunakan metode machine learning *Naïve Bayes*, *Support Vector Machine*, dan *K-Nearest Neighbours* untuk mengevaluasi akurasi dan menentukan metode yang memberikan hasil terbaik.
8. Penelitian ini dibatasi pada evaluasi performa algoritma *Naïve Bayes*, *Support Vector Machine*, dan *K-Nearest Neighbours* menggunakan *Confusion Matrix* dengan metrik akurasi, presisi, *recall*, dan *F1-score*.

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah ditetapkan, adapun tujuan dilakukannya penelitian adalah sebagai berikut:

1. Mengklasifikasikan sentimen pengguna Youtube mengenai video ulasan mobil listrik secara otomatis menggunakan

- pendekatan *lexicon-based* dengan bantuan *library TextBlob* menjadi sentimen positif, netral dan negatif.
2. Mengevaluasi dan membandingkan performa tiga algoritma *Naïve Bayes*, *Support Vector Machine (SVM)*, dan *K-Nearest Neighbors (KNN)* dalam menganalisis sentimen, serta pengaruh penggunaan teknik *SMOTE* dan *Grid Search* terhadap peningkatan kinerja model berdasarkan metrik akurasi, presisi, *recall*, dan *f1-score*.
 3. Mengembangkan aplikasi berbasis web menggunakan Streamlit yang mampu menerapkan hasil *text preprocessing*, pelabelan sentimen, dan visualisasi data, secara interaktif dan informatif.

1.4.2 Manfaat Penelitian

Manfaat yang didapatkan dari penelitian ini antara lain:

1. Memberikan pemahaman mengenai performa *Naïve bayes Classifier*, *Support Vector Machine*, dan *K-Nearest Neighbour* dalam menganalisis sentimen dari video ulasan mobil listrik di YouTube yang berasal dari reviewer Indonesia, sehingga dapat membantu menentukan metode klasifikasi opini publik.
2. Menghasilkan wawasan berbasis analisis sentimen mengenai persepsi masyarakat Indonesia terhadap mobil listrik, baik dari sisi keunggulan, kekurangan, maupun kekhawatiran yang diungkapkan dalam video-video ulasan. Informasi ini dapat digunakan oleh produsen mobil untuk meningkatkan kualitas produk dan pendekatan pemasaran yang lebih sesuai dengan kebutuhan konsumen Indonesia.
3. Menyediakan informasi yang dapat digunakan oleh pemerintah atau regulator terkait untuk memahami opini publik terhadap kebijakan dan infrastruktur mobil listrik di Indonesia, serta merancang regulasi yang lebih responsif terhadap kondisi pasar dan perlindungan konsumen.

4. Membantu masyarakat Indonesia dalam memahami tren sentimen publik terhadap berbagai merek dan jenis mobil listrik yang beredar, sehingga dapat dijadikan referensi dalam pengambilan keputusan pembelian yang lebih tepat dan sesuai dengan preferensi mereka.
5. Memperkaya wawasan akademik, khususnya bagi mahasiswa UMN, dalam memahami penerapan analisis sentimen pada mobil listrik serta bagaimana teknik ini dapat digunakan untuk mengkaji persepsi masyarakat terhadap perkembangan teknologi kendaraan listrik di Indonesia.

1.5 Sistematika Penulisan

Penulisan dalam pembuatan laporan skripsi ini disajikan dalam beberapa bab yang bertujuan untuk mempermudah dalam menemukan informasi yang dibutuhkan. Pembagian bab dijabarkan dengan detail sebagai berikut :

BAB I PENDAHULUAN

Pada Bab I ini, memiliki isi yang terdiri dari latar belakang penelitian ini dilakukan, rumusan masalah yang akan diselesaikan, batasan masalah selama penelitian dilakukan, tujuan dan manfaat yang ingin dicapai dari penelitian, dan sistematika penulisan laporan.

BAB II LANDASAN TEORI

Bagian ini menjelaskan teori tentang kendaraan listrik, algoritma naïve bayes, dan konsep dari pemanfaatan API dalam melakukan scraping data.

BAB III METODOLOGI PENELITIAN

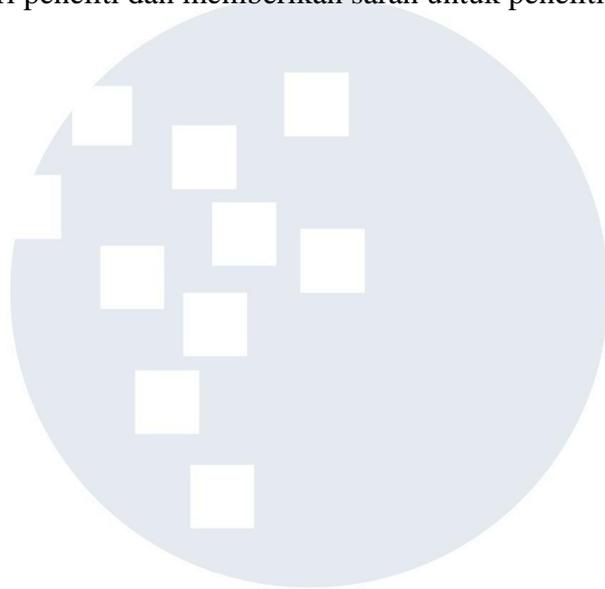
Bagian ini menjelaskan tentang metode dan tahap-tahap yang dilalui untuk menyelesaikan penelitian yang akan dilakukan.

BAB IV HASIL DAN DISKUSI

Bagian ini membahas mengenai hasil dari penelitian yang telah dilakukan dan menjelaskan perbedaan yang didapatkan.

BAB V KESIMPULAN DAN SARAN

Bagian ini menjelaskan hasil yang didapatkan secara keseluruhan berupa kesimpulan dari peneliti dan memberikan saran untuk peneliti selanjutnya.



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA