

# Pengukuran Sentimen Sosial Terhadap Teknologi Kendaraan Listrik: Bukti Empiris di Indonesia

Atha Fitrah Riyadi, Faiz Ramadhani Rahman, M. Aldiansyah Nofa Pratama, M. Khanif Khafidli\*, Harry Patria  
 Magister Manajemen Teknologi  
 Institut Teknologi Sepuluh Nopember Kampus ITS Tjokroaminoto  
 Surabaya, INDONESIA

riyadi.atha@gmail.com, faizramadhani09@gmail.com, nofaldiansyah@gmail.com, khanifkhafidli@gmail.com\*, hary.patria@sbm-itb.ac.id

**Abstract** – Currently, climate change due to global warming is a concern for many parties. The greenhouse gas emissions level is increasing day by day. The major contributors to Air pollution are the greenhouse effect. Transportation accounts for about 27% of air pollution, and governments in various countries use electric vehicles to reduce air pollution. However, the success of using electric vehicles is depending on perception, people's sentiment, and understanding. The main purpose of this research is to find out how public sentiment towards electric vehicles is through tweets and comments on the Twitter social media platform using sentiment analysis. The data obtained are 1084 tweets and comments. The data were classified using the Naïve Bayes method, K-Nearest Neighbor, and Decision Tree. The results showed that the Naïve Bayes Classification Method gave better results than K-Nearest Neighbor, and Decision Tree with an accuracy rate of 94% positive sentiment by 53%, negative sentiment by 38%, and neutral sentiment by 9%. So, it can be concluded that public sentiment towards electric vehicles is quite good based on conversations on the Twitter social media platform. In addition, the author also visualizes the results of the analysis in the form of graphs and word clouds so that they can help the electric vehicle industry players to understand public sentiment better and more accurately.

**Keywords:** Sentiment Analysis; Electric Vehicles; Climate Change; Social Media.

**Abstrak** – Saat ini, perubahan iklim akibat pemanasan global menjadi perhatian bagi banyak pihak. Tingkat emisi gas rumah kaca meningkat dari hari ke hari. Polusi udara adalah salah satu kontributor yang berperan signifikan terhadap efek rumah kaca. Transportasi menyumbang sekitar 27% dari polusi udara, dan pemerintah di berbagai negara mengambil langkah-langkah untuk mengurangi polusi udara dengan mendorong penggunaan kendaraan listrik. Namun keberhasilannya tergantung pada sentimen, persepsi dan pemahaman masyarakat terhadap kendaraan listrik. Tujuan utama penelitian ini adalah untuk mengetahui bagaimana sentimen publik terhadap kendaraan listrik melalui tweets dan komentar di platform sosial media twitter menggunakan analisis sentimen. Data yang diperoleh adalah 1084 tweets dan komentar. Data diklasifikasikan menggunakan metode Naïve Bayes, K-Nearest Neighbor, dan Decision Tree. Hasil penelitian menunjukkan bahwa bahwa Metode Klasifikasi Naïve Bayes memberikan hasil lebih baik dari K-Nearest Neighbor, dan Decision Tree dengan tingkat akurasi 94% sentimen positif sebesar 53%, sentimen negatif sebesar 38% dan sentimen netral sebesar 9%. Sehingga dapat disimpulkan bahwa sentimen masyarakat terhadap kendaraan listrik cukup baik berdasarkan percakapan di platform sosial media twitter. Selain itu, penelitian ini juga memvisualkan hasil analisa dalam bentuk grafik dan word cloud sehingga dapat membantu pelaku industri kendaraan listrik untuk memahami sentimen publik dengan lebih baik dan akurat.

**Kata Kunci:** Analisis Sentimen; Kendaraan Listrik; Perubahan Iklim; Media Sosial.

## 1. Pendahuluan

Perubahan iklim adalah isu yang banyak dibahas oleh berbagai negara secara serius pada masa kini. Kerusakan berbagai ekosistem, perubahan cuaca ekstrem, bencana alam, merupakan akibat terbesar dari perubahan iklim. Salah satu langkah terbaik untuk menekan perubahan iklim adalah dengan mengurangi emisi gas rumah kaca. Jika kenaikan suhu rata-rata di seluruh dunia terus meningkat dan potensi banyak bencana akan terjadi yang menimbulkan kerugian bagi seluruh makhluk hidup maka seluruh pihak dari berbagai negara harus tegas untuk mengambil keputusan untuk mendukung kegiatan yang ramah akan lingkungan.

Tanpa disadari, gas rumah kaca merupakan penyumbang terbesar dalam hal kenaikan suhu rata-rata global akibat kegiatan yang dilakukan manusia. Sebagai contoh penggunaan kendaraan pribadi untuk transportasi sehari-hari masih menggunakan bahan bakar fosil. Bahan bakar fosil merupakan bahan bakar yang turut andil dalam membentuk gas rumah kaca, dari lepasnya karbon monoksida, karbon dioksida hasil pembakaran ke atmosfer tersebut menjadi salah satu penyebab utama terbentuknya gas rumah kaca. Akibatnya lepasnya gas-gas tersebut menjadi penyebab yang kuat terjadinya perubahan iklim dan pemanasan global saat ini.

Dari laporan *Climate Transparency Report 2020* [1] dalam upaya mengurangi emisi gas di negara anggota G20 dengan target dari *Nationally Determined Contribution*



(NDC), bahwa transportasi adalah penyumbang 27 persen dari polusi udara. Kontributor utama yang signifikan terhadap polusi udara adalah transportasi dengan bahan bakar fosil, dan penggunaan alat berat.

Cara yang dapat diambil untuk mengurangi dan menekan emisi dari gas adalah dengan cara memulai penggunaan kendaraan listrik (*electric vehicle*) seperti yang dilakukan beberapa negara secara gencar akhir-akhir ini. Contohnya beberapa negara di Eropa perlahan mulai menggunakan kendaraan listrik. Sebuah penelitian di *European Environment Agency* [2] memperlihatkan kenaikan 3,5 persen penjualan mobil dari tahun sebelumnya merupakan kendaraan listrik dengan rasio 1:9 dari produksi mobil baru tipe hibrida yang rendah emisi dan berpenggerak listrik. Dari kenaikan penjualan tersebut ternyata berkontribusi menurunkan emisi CO<sub>2</sub> sebesar 12 persen dari tahun sebelumnya berdasarkan jumlah total mobil baru yang terjual, dan menjadi penurunan emisi paling besar sejak tahun 2010 dari mulai penetapan standar produksi mobil di Uni Eropa. 11 persen penjualan mobil baru dari total 11,6 juta yang terdaftar di Uni Eropa, adalah mobil listrik dan hibrida. Peningkatan tren pembelian mobil terus berlanjut meskipun kondisi saat ini masih dalam kondisi pandemi.

Pemerintah Indonesia berkomitmen untuk ikut andil dalam aksi nyata untuk mengurangi emisi dan perubahan iklim dunia. Penggunaan kendaraan listrik atau kendaraan dengan emisi rendah juga diharapkan menjadi tren yang mengikuti tren di Uni Eropa dan secara gencar diaplikasikan di Indonesia. Peraturan Presiden (Perpres) No 55 Tahun 2019 dari Pemerintah Indonesia adalah regulasi pemakaian kendaraan listrik mengenai transportasi jalan yang menggunakan kendaraan listrik dengan baterai, dan merupakan bukti nyata juga untuk meningkatkan upaya dalam menekan emisi gas dan perubahan iklim. Namun realisasinya, Indonesia masih tertinggal jika dibandingkan dengan Eropa. Dari catatan Gabungan Industri Kendaraan Bermotor Indonesia (Gaikindo) semester I/2020 menunjukkan penjualan kendaraan dengan emisi yang rendah adalah 1.900 unit terdiri dari 1.378 unit kendaraan *hybrid*, 34 unit kendaraan *plug-in hybrid electric vehicle* (PHEV) dan 488 unit kendaraan listrik dengan baterai (BEV) [3]. Hal ini juga dikuatkan oleh pernyataan Menteri Perhubungan Indonesia Budi Karya Sumadi bahwa polusi udara terbesar di Indonesia disumbang oleh kendaraan bermotor [4].

Salah satu hal yang berpengaruh terhadap tingkat adopsi kendaraan listrik di suatu negara adalah persepsi masyarakat terhadap kendaraan listrik. Persepsi positif terhadap kendaraan listrik dapat memainkan peran penting untuk mempromosikan kendaraan listrik ini di suatu negara. Tidak hanya terkait permasalahan teknis, namun persepsi sosial juga harus diperhatikan untuk meningkatkan distribusi komersial kendaraan listrik [5]. Sentimen dan perasaan konsumen juga ditemukan mempengaruhi tingkat adopsi kendaraan listrik [6]. Perasaan positif dalam mengendarai kendaraan listrik berkorelasi positif dengan niat konsumen untuk mengadopsi kendaraan listrik tersebut. Beberapa konsumen mengungkapkan sentimen mereka seperti “merasa baik” atau “kurang bersalah” setelah

mengendarai kendaraan listrik. Di sisi lain, konsumen menyatakan perasaan “malu” setelah mengendarai kendaraan listrik [7]. Temuan dari penelitian lainnya [8] menunjukkan bahwa persepsi positif pada gilirannya akan secara positif pula mempengaruhi niat untuk mengadopsi kendaraan listrik.

Selama beberapa tahun terakhir, terdapat banyak penelitian yang telah menganalisa bagaimana penerimaan pasar terhadap kendaraan hemat energi, banyak peneliti telah mempelajari dan mengidentifikasi faktor-faktor yang mempengaruhi penerimaan dan penolakan pasar terhadap kendaraan listrik. Menurut penelitian [9] disimpulkan bahwa harga, *driving range*, dan *charging rate* akan mempengaruhi kepemilikan kendaraan listrik. Sedangkan penelitian di negara tetangga yaitu Singapura mengungkapkan bahwa tingginya biaya *Certificate of Entitlement* (COE) dan nilai jual kembali menjadi perhatian utama konsumen untuk membeli kendaraan listrik [5]. Faktor penting lainnya untuk adopsi kendaraan listrik adalah kebijakan pemerintah [10]. Kebijakan pemerintah harus dijelaskan dengan baik kepada konsumen, karena jika tidak, kebijakan akan gagal meningkatkan peluang adopsi [11].

Dalam penelitian [12] disimpulkan bahwa jarak jangkauan yang terbatas sejauh 100 mil dianggap cukup untuk memiliki kendaraan listrik sebagai mobil kedua. 34 persen dari peserta penelitian menyatakan bahwa 150 mil akan membuat kendaraan listrik cocok sebagai mobil pertama. Jadi masih menjadi perdebatan bahwa jarak mengemudi kendaraan listrik yang terbatas merupakan penghalang atau tidak terhadap kepemilikan kendaraan listrik. Namun sebuah penelitian lain [13] terhadap 369 pengemudi di Denmark menemukan fakta bahwa jangkauan kendaraan listrik menjadi perhatian nyata karena kurang dari apa yang mereka inginkan dalam kendaraan listrik. Memahami persepsi sosial masyarakat untuk mengetahui hambatan penerimaan kendaraan listrik dapat membantu pelaku industri kendaraan listrik dalam menentukan segmen pasar dan memaksimalkan potensi masa depan kendaraan listrik, terutama di Indonesia.

Saat ini pemanfaatan teknologi informasi berkembang begitu pesatnya. Maka dari itu pemanfaatan teknologi informasi yang tepat dapat dijadikan sebagai salah satu senjata strategis dalam memberikan nilai tambah dalam persaingan bisnis [14], dalam hal ini memanfaatkan teknologi untuk mengetahui bagaimana persepsi sosial masyarakat Indonesia terhadap kendaraan listrik. Teknologi informasi yang dapat digunakan untuk penarikan informasi adalah melalui sosial media. Salah satu platform sosial media yang saat ini diminati oleh masyarakat Indonesia adalah *twitter* dengan menduduki peringkat ketujuh dari dua belas media sosial yang disediakan yaitu sebesar 27 persen [15]. *Tweets* maupun komentar publik terkait kendaraan listrik memiliki jumlah yang cukup banyak dengan aliran informasi yang sangat cepat. Salah satu cara untuk menarik kesimpulan terkait bagaimana persepsi masyarakat Indonesia terhadap kendaraan listrik adalah dengan menggunakan analisis sentimen.

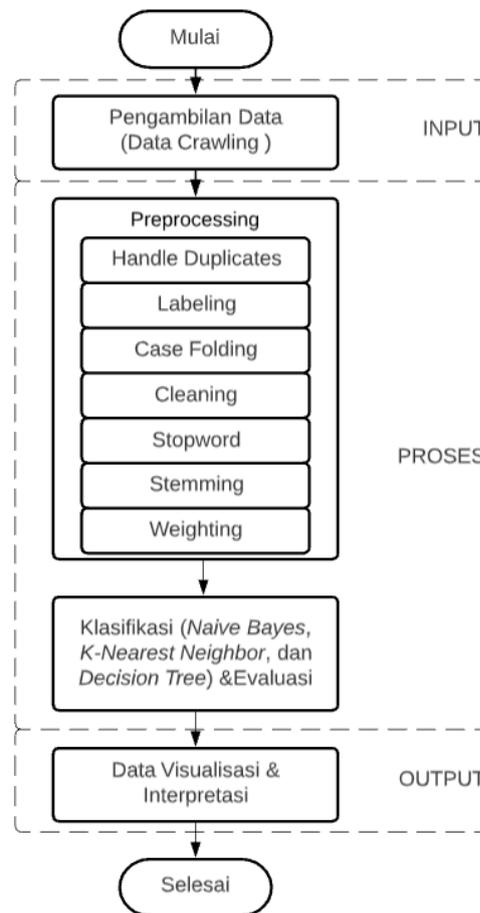
Pada penelitian ini, dilakukan perbandingan metode *K-Nearest Neighbor*, *Naïve Bayes*, dan *Decision Tree* untuk mengetahui sentimen masyarakat Indonesia terhadap kendaraan listrik. Dari penelitian ini diharapkan dapat menjadi bahan rekomendasi kepada pelaku industri kendaraan listrik dan pemerintah Indonesia untuk menentukan kebijakan dan strategi terbaik dalam meningkatkan tingkat adopsi masyarakat Indonesia terhadap kendaraan listrik.

### 2. Metodologi

Data yang digunakan dalam penelitian ini adalah data sekunder yang didapat dengan menggunakan *Twitter API (Application Programming Interface)* selama 1 Agustus 2021 hingga 30 September 2021. Data tersebut berupa *tweets*

maupun komentar dari pengguna *twitter* yang mengandung *keywords* Mobil Listrik, Tesla, *Electric Vehicle*, Gesits, Mobil SMK, Kendaraan Listrik, dan Motor Listrik.

Dalam penelitian ini menggunakan metode eksperimen. Metode eksperimen adalah metode sistematis yang menghubungkan antara sebab dan akibat. Penelitian ini tergolong penelitian kuantitatif yang memuat kondisi; mengontrol, memanipulasi dan mengamati [16]. Eksperimen dilakukan untuk mendapatkan model terbaik dari beberapa model yang digunakan, diantaranya adalah *Naïve Bayes*, *Decision Tree*, dan *K-Nearest Neighbor* sehingga didapatkan analisa yang akurat terkait sentimen masyarakat Indonesia terhadap kendaraan listrik.



Gambar 1. Metodologi Penelitian

Kegiatan penelitian ini dibagi menjadi 4 tahap: (1) *Data Crawling*, (2) *Data Preprocessing*, (3) *Sentiment Analysis* dan *Evaluation* (4) *Data Visualization* dan *Interpretation*, seperti pada Gambar 1.

#### A. Pengambilan Data (*Data Crawling*)

Tahap pertama yang dilakukan adalah *data crawling* yaitu teknik pengumpulan data pada sebuah *website* dengan menggunakan *Uniform Resource Locator (URL)*. URL ini adalah referensi untuk menemukan semua

*hyperlink* di situs *website*. Kemudian *indexing* dilakukan untuk menemukan *keywords* dalam dokumen pada setiap link yang ada. Akses terhadap data *tweets* di *twitter* memerlukan hak akses berupa *consumer key*, *consumer secret*, *access token* dan *access token secret*.

Dimana pada penelitian ini data diambil dari *twitter* menggunakan *Twitter API (Application Programming Interface)* dengan *keywords* Mobil Listrik, Tesla, *Electric Vehicle*, Gesits, Mobil SMK, Kendaraan Listrik, dan Motor Listrik.

### B. Data Preprocessing

Tahap kedua adalah *data preprocessing*. Data yang telah diambil dari *twitter* dianalisis dan diklasifikasikan. *Data preprocessing* dilakukan untuk mengolah data mentah menjadi sekumpulan data yang siap digunakan. Secara garis besar, ada dua tujuan *data preprocessing* yaitu untuk membersihkan data dari data yang jumlahnya ganda dan untuk pemilihan data yang akan diolah dalam dokumen [17].

Berikut *data preprocessing* yang dilakukan pada penelitian ini: (1) *Handle Duplicates*: Melakukan *filtering* terhadap *tweets* yang memiliki kemungkinan jumlah ganda dalam sekumpulan data yang diperoleh; (2) *Labeling*: Melakukan proses pelabelan dengan menggunakan dua *annotators*, *output* dari *annotators* pertama akan di *cross check* oleh *annotators* kedua; (3) *Case Folding*: Melakukan perubahan teks dalam dokumen menjadi bentuk baku yaitu berupa huruf kecil; (4) *Cleaning*: Menghilangkan komponen tertentu yang terdapat dalam *tweets* yaitu *Uniform Resource Locator* (URL), *username*, RT (*Retweet*),

karakter HTML, dan *hashtag*; (5) *Tokenizing*: Melakukan pemecahan teks menjadi kata-kata, dengan metode unigram, bigram, dan trigram; (6) *Stopwords Removal*: Menghapus kata-kata pada *tweets* yang terdapat pada daftar *stopwords* yang telah dibuat; (7) *Stemming*: Menghilangkan kata imbuhan untuk mendapatkan kata dasar; (8) *Weighting*: Melakukan proses pembobotan kata dengan TF-IDF.

### C. Sentiment Analysis dan Evaluation

Tahap ketiga adalah *sentiment analysis* dan *Evaluation* dilakukan untuk mengklasifikasikan persepsi sosial masyarakat terhadap kendaraan listrik [5].

Pada eksperimen ini dilakukan beberapa model yang digunakan, diantaranya adalah *Naïve Bayes*, *K-Nearest Neighbor*, dan *Decision Tree* menggunakan aplikasi *Rapid Miner* sebagai alat analisis. Sebanyak 1084 *tweets* yang di ambil pada periode bulan 1 Agustus sampai 30 September 2021. Dataset dibagi menjadi 70% untuk pelatihan dan 30% untuk pengujian seperti pada Tabel 1.

Tabel 1. Data Pelatihan dan Pengujian

Data Pelatihan	760
Data Pengujian	325
Total Data	1084

Hasil dari masing-masing akan dilakukan perbandingan dan dipilih model terbaik dengan melihat tingkat akurasi model berdasarkan *Confusion Matrix*: antara metode *Naïve Bayes*, *K-Nearest Neighbor*, dan *Decision Tree*.

### D. Naïve Bayes

*Naïve Bayes* adalah salah satu metode klasifikasi *supervised learning* yang paling umum digunakan untuk melakukan klasifikasi teks [18]. Untuk melakukan klasifikasi, kita perlu memilih fitur dari data terlebih dahulu. Untuk klasifikasi teks fitur *vector* juga disebut *term vector* yang merupakan struktur terpenting selama pelatihan dan proses klasifikasi dimana semua teks *tweet* akan akan diubah menjadi menjadi *term vector* untuk diproses oleh klasifikasi. *Term vector* dihasilkan berdasarkan kosakata unik yang diperoleh dari dataset pelatihan dan tidak ada duplikat dalam kosakata.

Dalam implementasi klasifikasi teks terdapat dua jenis *Naïve Bayes* yaitu *Naïve Bayes Bernoulli* dan *Naïve Bayes*

*Multinomial*, perbedaan antara keduanya adalah cara dimana fitur diekstraksi dari dokumen. Pada *Naïve Bayes Bernoulli*, *term vector* akan diinisialisasi dengan semua elemen menjadi nol, kemudian diperiksa pada setiap kata di kosakata untuk melihat apakah kata tersebut ada di *tweet*. Jika ada, maka tandai yang sesuai pada elemen dalam istilah vektor ke-1, jika tidak maka tandai yang sesuai pada elemen dalam istilah vektor ke-0, elemen dari *term vector* tidak hanya mewakili ada atau tidaknya kata namun juga dapat mewakili frekuensi kata.

Misalnya, jika kita memiliki dokumen D, dan kita memiliki kelas C, yang berisi beberapa kelas. Kemudian untuk mendapatkan kelas mana yang dimiliki oleh dokumen D, kita hanya perlu menghitung posterior probability P(C|D), dan memilih yang terbesar. P(C|D) bisa dihitung dengan Teorema Bayes (1), di mana prior probability dan likelihood sebelumnya dapat dihitung dari kumpulan data berlabel

$$P(C|D) = \frac{P(D|C)P(C)}{P(D)} \propto P(D|C)P(C) \quad (1)$$

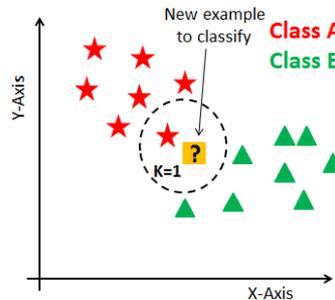
$$P(S_i|C_i) \propto P(T_j|C_i) = \prod_{j=1}^{|V|} [T_j P(W_j|C_i) + (1 - T_j)(1 - P(W_j|C_i))] \quad (2)$$

Berikut adalah cara kerja kedua algoritma ini. Setiap kelas memiliki kemungkinan yang sama, jadi kita bisa dengan mudah mendapatkan *prior probability* P(C<sub>i</sub>). Biarkan V menjadi kosakata, dan w<sub>j</sub> adalah kata ke-j di V, jadi dari data pelatihan, kita bisa mendapatkan probabilitas w<sub>j</sub> milik kelas C<sub>i</sub>, yaitu P(w<sub>j</sub> | C<sub>i</sub>).

Kemudian asumsikan bahwa S<sub>i</sub> adalah *tweet* ke-i dalam dataset pengujian, dan T<sub>j</sub> adalah *term vector* untuk *tweet* ini. Seperti yang telah kita ketahui, T<sub>j</sub> berisi 0 dan 1 itu berarti jika sesuai kata dalam kosakata ada di *tweet* S<sub>i</sub>. Maka probabilitas *tweet* S<sub>i</sub> milik kelas C<sub>i</sub> persamaan (2).

### E. K-Nearest Neighbor

K-Nearest Neighbor (KNN) adalah salah satu metode *machine learning* yang masuk dalam kategori *supervised learning* untuk melakukan klasifikasi teks dan data. Metode K-Nearest Neighbor melakukan klasifikasi terhadap objek berdasarkan data yang jaraknya paling dekat dengan objek tersebut atau biasa dapat disebut *neighborhood*.



Gambar 2. Model Klasifikasi KNN

Dari data sampel yang digunakan sebagai dimensi dalam penentuan keberadaan data sampel tersebut. Dimensi dibagi menjadi beberapa bagian berdasarkan *training sample*, dan akan ditandai dengan sebuah titik. Apabila titik tersebut berdekatan dengan  $k$  (*Neighbours*) maka akan masuk ke dalam kelas tersebut. *Euclidean Distance* merupakan Metode untuk mencari jarak terdekat.

Pada dasarnya tujuan algoritma ini adalah untuk klasifikasi objek dari atribut dan data latih. Jika algoritma tersebut diberikan *query* maka akan memunculkan data latih yang paling dekat dengan data baru yang dimasukkan dan dengan hasil voting paling banyak akan didapatkan klasifikasi diantara klasifikasi dari  $k$  obyek.

Berikut persamaan untuk menentukan jarak menggunakan *Euclidean Distance* [21] (3) yaitu:

- $A_i$  = kata ke- $i$  dokumen Testing atau Uji
- $B_i$  = kata ke- $i$  dokumen Training atau Latih
- $d(A, B)$  = Jarak dokumen  $A$  ke dokumen  $B$

$$d(A, B) = \sum_{i=1}^t \sqrt{(A_1 - B_1)^2 + \dots + (A_i - B_i)^2} \quad (3)$$

### F. Decision Tree

*Decision Tree* adalah *flowchart* sederhana yang memilih label untuk nilai input. *Flowchart* ini terdiri dari node keputusan, yang memeriksa nilai fitur, dan *leaf node*, yang memberikan label. Untuk memilih label nilai input, kita mulai dari *flowchart node* keputusan awal, yang dikenal sebagai *root node*. *Node* ini berisi kondisi yang memeriksa salah satu fitur nilai input, dan memilih cabang berdasarkan nilai fitur tersebut [13].

Mengikuti cabang yang menjelaskan nilai input, kemudian sampai di *node* keputusan baru dengan kondisi baru pada fitur nilai input. selanjutnya mengikuti cabang dipilih oleh kondisi masing-masing *node*, sampai di *leaf node* yang menyediakan label untuk nilai masukan. Untuk klasifikasi teks, *node* keputusan dapat berupa kata fitur yang telah kita pilih dan *leaf node* bisa menjadi kategori.

### G. Evaluation

Dalam melakukan evaluasi analisis sentiment terdapat beberapa cara bisa digunakan salah satunya adalah menggunakan acuan *confusion matrix* atau matrik yang dapat mempresentasikan prediksi dengan kondisi

sebenarnya dari data yang dihasilkan oleh algoritma *Machine Learning*, dengan acuan *confusion matrix* kita bisa meentukan nilai *accuracy*, *precision*, *recall* dan *specifity* [19][20]. *Accuracy* adalah rasio prediksi benar (positif dan negatif) dengan keseluruhan data. Akurasi menjawab pertanyaan “Berapa persen data benar yang relevan dan tidak relevan dari keseluruhan data uji” seperti pada persamaan (4).

*Precision* adalah rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif, seperti pada persamaan (5).

*Recall* adalah rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif, seperti pada persamaan (6).

*Specifity* adalah kebenaran memprediksi negatif dibandingkan dengan keseluruhan data negative, seperti pada persamaan (7) yaitu:

- False Negative* = FN
- False Positive* = FP
- True Negative* = TN
- True Positive* = TP

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \times 100\% \quad (4)$$

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad (5)$$

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\% \quad (6)$$

$$\text{Specifity} = \frac{TN}{TN+FP} \times 100\% \quad (7)$$

### H. Data Visualization dan Interpretation

Tahap keempat adalah *Data Visualization* dan *Interpretation*. Untuk memvisualisasikan hasil analisis pada tahapan sebelumnya, maka dibuatkan *data visualization* berupa grafik dan *word cloud* agar hasil analisa dapat lebih mudah dipahami dengan menggunakan berbagai visualisasi yang menarik. Sedangkan *interpretation* dilakukan untuk mendapatkan kesimpulan dari hasil visualisasi.

### 3. Hasil dan Pembahasan

Penelitian ini telah mengumpulkan semua *tweet* relevan yang berkaitan dengan Teknologi Kendaraan Listrik. Dibawah ini disajikan hasil untuk klasifikasi sentimen dan memvisualisasikan serta memberikan interpretasi terhadap hasil terkait.

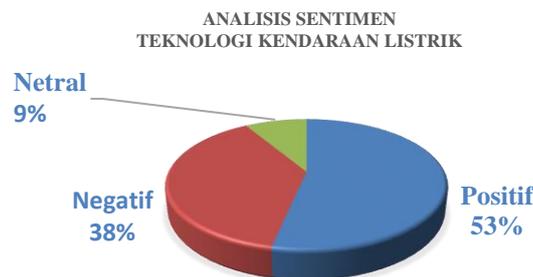
#### A. Sentiment Analysis

Eksperimen dilakukan dengan pengklasifikasi sentimen menggunakan algoritma *Decision Tree*, *K-NN*,

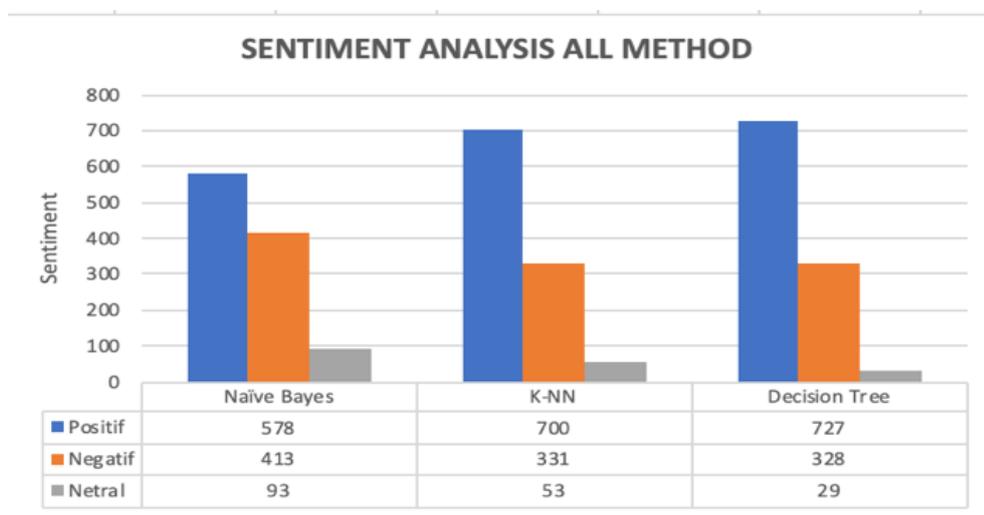
dan *Naïve Bayes*. Analisis sentimen dari penelitian ini menggunakan software *Rapid Miner* dengan tambahan *text processing extension*. Sebelum dilakukan klasifikasi, *tweet* berbahasa Indonesia diterjemahkan terlebih dahulu ke dalam inggris karena terbatasnya jumlah penelitian klasifikasi teks berbahasa Indonesia.

**Gambar 2** menunjukkan kondisi penerimaan masyarakat terhadap Teknologi Kendaraan Listrik di Indonesia pada bulan Agustus hingga September 2021. Dapat terlihat bahwa 53% masyarakat memberikan sentimen positif, 38% memberikan sentimen negatif, dan 9% netral.

**Gambar 3** merupakan hasil analisis sentimen menggunakan model dari metode *Naïve Bayes*. Penelitian ini juga membandingkan metode *Naïve Bayes* terhadap metode lain untuk mencari metode mana yang terbaik guna mendapatkan model yang yang terbaik. Dilakukan juga analisis sentimen menggunakan metode *K-Nearest Neighbor* dan *Decision Tree*. **Gambar 3** menunjukkan persebaran dan perbedaan yang tidak signifikan antara metode satu dengan yang lainnya



**Gambar 2.** Hasil Analisis Sentimen



**Gambar 3.** Perbandingan Hasil Analisis Sentimen Antar Model

### B. Evaluation

Selanjutnya dari hasil ketiga model yang telah dilakukan. Penelitian ini melakukan evaluasi model untuk membandingkan dan melakukan justifikasi model mana yang terbaik dari ketiga metode yang digunakan guna mencari hasil persebaran sentiment yang paling akurat dari data *tweet* yang diperoleh.

Gambar 4 menunjukkan bahwa diperoleh hasil akurasi sangat baik yaitu 94% dari model algoritma *Naïve*

*Bayes* dengan perolehan false positif yang relatif lebih besar dibandingkan lain.

Gambar 5 menunjukkan hasil analisa model menggunakan algoritma *K-Nearest Neighbor* dengan hasil akurasi baik yaitu 92.25%. dengan false negatif dan netral yang relatif lebih besar daripada yang lain.

Gambar 6 menunjukkan hasil analisa model menggunakan algoritma *K-Nearest Neighbor* dengan hasil akurasi cukup baik yaitu masih di angka 85.61% dengan false negatif dan netral yang relatif lebih besar dibandingkan yang lain.

accuracy: 94.00%

	true Positif	true Negatif	true Netral	true ?	class precision
pred. Positif	578	0	0	0	100.00%
pred. Negatif	46	367	0	0	88.86%
pred. Netral	12	7	74	0	79.57%
pred. ?	0	0	0	0	0.00%
class recall	90.88%	98.13%	100.00%	0.00%	

Gambar 4. Confusion Matrix Model *Naïve Bayes*

accuracy: 92.25%

	true Positif	true Negatif	true Netral	true ?	class precision
pred. Positif	634	61	5	0	90.57%
pred. Negatif	2	313	16	0	94.56%
pred. Netral	0	0	53	0	100.00%
pred. ?	0	0	0	0	0.00%
class recall	99.69%	83.69%	71.62%	0.00%	

Gambar 5. Confusion Matrix Model *K-Nearest Neighbor*

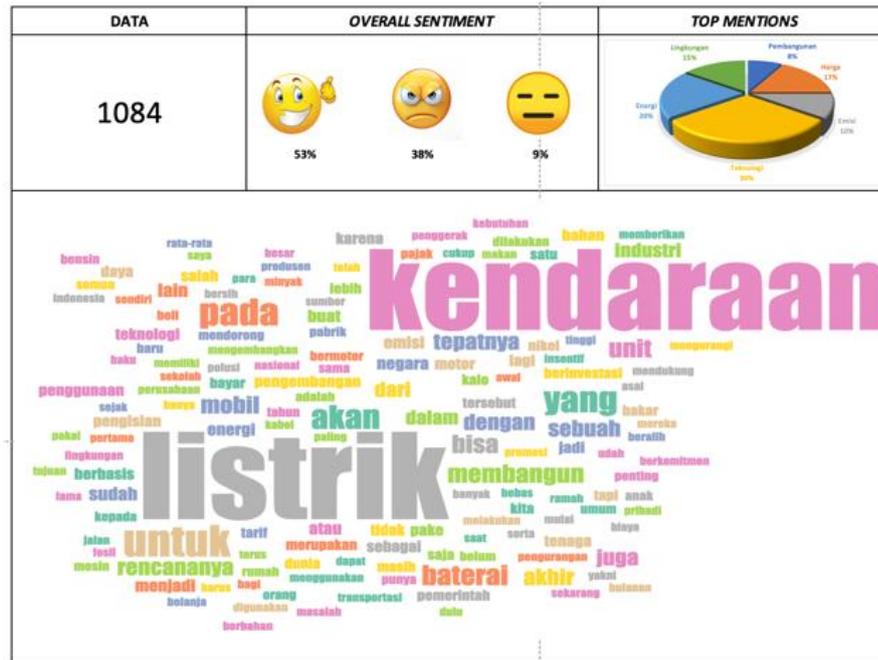
accuracy: 85.61%

	true Positif	true Negatif	true Netral	true ?	class precision
pred. Positif	610	78	39	0	83.91%
pred. Negatif	25	293	10	0	89.33%
pred. Netral	1	3	25	0	86.21%
pred. ?	0	0	0	0	0.00%
class recall	95.91%	78.34%	33.78%	0.00%	

Gambar 6. Confusion Matrix Model *Decision Tree*

Tabel 2. Perbandingan Pengukuran Evaluasi Antar Model

	<i>Naïve Bayes</i>	<i>K-NN</i>	<i>Decision Tree</i>
<i>Accuracy</i>	94,00%	92,25%	85,61%
<i>Weighted Mean Recall</i>	72,25%	63,75%	52,01%
<i>Weighted Mean Precision</i>	67,11%	71,28%	64,86%



Gambar 7. Visualisasi Data Sentimen

Model evaluasi juga dilakukan dengan melihat *weighted mean precision* dan *recall* dari setiap model yang dihasilkan. Tabel 2 memperlihatkan perbedaan hasil evaluasi model dari ketiga metode yang dilakukan.

Hasil evaluasi perbandingan metode dapat dilihat pada Tabel 2 yang menunjukkan bahwa Metode Naïve Bayes merupakan metode yang paling baik dibandingkan dengan Metode K-NN dan *Decision Tree* untuk mendapatkan model sentiment analisis pada kendaraan listrik di Indonesia, hal tersebut dapat dilihat dari hasil evaluasi menggunakan confusion matrix nilai *Accuracy* dan *Recall* pada metode Naïve Bayes mempunyai nilai paling tinggi dibandingkan dengan metode lain serta nilai *Precision* lebih rendah sedikit dibandingkan metode K-NN.

### C. Data Visualization dan Interpretation

Visualisasi data *tweet* juga dilakukan dengan menggunakan *word cloud* dan beberapa chart pendukung. Visualisasi ini dapat membantu melihat pembicaraan terbanyak dari masing-masing kata kunci yang terlibat dalam analisa kendaraan listrik ini. Gambar 7 menunjukkan bahwa topik dominan yang terbentuk di kalangan masyarakat pengguna *twitter* pada bulan Agustus hingga September 2021 terkait topik teknologi kendaraan listrik adalah Lingkungan, Pembangunan, Harga, Energi, Teknologi, dan Emisi.

Interpretasi dari hasil analisis sentimen dan juga visualisasi data terkait penerimaan teknologi kendaraan listrik di Indonesia adalah sudah cukup baik. Ini terlihat dari presentase sentiment positif pada bulan Agustus hingga September 2021 pada media sosial *twitter*. Hanya saja hal ini dirasa masih kurang optimal dimana masih hampir setengah dari sentimen masyarakat masih merasa teknologi kendaraan listrik ini negatif maupun *netizen* yang

memilih untuk netral dalam mengomentari perkembangan teknologi ini. Perlu adanya optimalisasi terhadap kondisi terkini penerimaan masyarakat dan diharapkan perusahaan teknologi kendaraan listrik dan pemerintah selaku pemangku kebijakan dapat saling mendukung dalam upaya meminimalisi penggunaan emisi karbon di Indonesia yang akan berdampak positif terhadap lingkungan.

### 4. Kesimpulan

Analisis pengukuran sentimen masyarakat dan visualisasi ini dimaksudkan untuk membantu para pelaku industri kendaraan listrik dalam memahami kondisi penerimaan masyarakat terhadap teknologi kendaraan listrik di Indonesia. Penelitian ini menggunakan data media sosial sebagai sumber untuk melihat peta sentimen terhadap industri kendaraan listrik. Berdasarkan hasil perbandingan metode Naïve Bayes, K-NN dan *Decision Tree* yang telah dilakukan untuk mendapatkan model sentiment analisis pada kendaraan listrik di Indonesia diperoleh bahwa metode klasifikasi yang terbaik adalah menggunakan Metode Naïve Bayes dengan tingkat *accuracy* 94% *Recall* 72,25% dan *Precision* 67,11% serta hasil analisis sentimen didapatkan sebesar 53% memberikan sentimen positif, 38% memberikan sentimen negatif dan 9% netral terhadap teknologi kendaraan listrik

Pengukuran sentimen analisis terhadap penerimaan masyarakat di Indonesia menggunakan Metode Naïve Bayes, K-NN dan *Decision Tree* semuanya memberikan hasil respon sentiment positif dominan atau lebih dari 50% pada masing-masing metode, ini memberikan gambaran bahwa penerimaan masyarakat di Indonesia masih bisa untuk dioptimalkan kedepannya. Dibutuhkan peran bersama oleh perusahaan teknologi kendaraan

listrik maupun pemerintah sebagai pemangku kebijakan untuk lebih intense dalam memberikan informasi yang menarik ke seluruh omnichannel media di tanah air terkait seberapa bermanfaatnya teknologi ini. Promosi dan sosialisasi diikuti dengan perubahan kebijakan yang tepat guna akan dapat membantu menyusul ketertinggalan terhadap negara-negara Eropa dalam upaya meminimalisir jumlah emisi gas rumah kaca di Indonesia.

Penelitian ini memberikan rekomendasi untuk pengembangan penelitian di masa depan. Penerapan analisis sentimen dalam *time frame* yang lebih besar dan dengan jumlah *tweet* yang lebih besar menjadi tantangan tersendiri untuk penelitian kedepannya termasuk penggunaan metode *lexicon sentiment analysis* untuk membantu mesin dalam melakukan klasifikasi daripada membiarkan mesin belajar sendiri. Pengembangan penelitian juga dapat dilakukan dengan mengekstraksi data dari platform media sosial lainnya seperti *Facebook*, *Instagram*, *Youtube* dan media-media online.

### 5. Daftar Pustaka

- [1] Climate Transparency, "The Climate Transparency Report," *Annual G20 Report*, 2020.
- [2] European Environment Agency, "Electric Vehicle in Europe," *EEA Report*, no. 20/2016, 2016.
- [3] Gaikindo "Harga dan Infrastruktur Jadi Tantangan Mobil Listrik Indonesia," 2021.
- [4] Tempo, "Menhub: Polusi Udara Disumbang Penggunaan Motor dan Mobil," *Bisnis Tempo*, 2020.
- [5] Min, X., Qiang, M., & Yisi, L., "Public's perception of adopting electric vehicles: A case study of Singapore," *Journal of the Eastern Asia Society for Transportation Studies*, 12, 2017.
- [6] Christidis, P., & Focas, C., "Factors affecting the uptake of hybrid and electric vehicles in the European Union," *Energies*, 12(18), 3414, 2019.
- [7] Graham-Rowe, E., Gardner, B., Abraham, C., Skippon, S., Dittmar, H., Hutchins, R., & Stannard, J., "Mainstream consumers driving plug-in battery-electric and plug-in hybrid electric cars: A qualitative analysis of responses and evaluations," *Transportation Research Part A*, 46, 140–153, 2012.
- [8] Schuitema, G., Anable, J., Skippon, S., & Kinnear, N., "The role of instrumental, hedonic and symbolic attributes in the intention to adopt electric vehicles," *Transportation Research Part A: Policy and Practice*, 48, 39-49, 2013.
- [9] Yang, S., Deng, C., Tang, T., & Qian, Y., "Electrical vehicle's energy consumption of car-following models," *Nonlinear Dynamics*, 71, 323-329, 2013.
- [10] Sun, X., & Xu, S., "The impact of government subsidies on consumer preferences for alternative fuel vehicles," *J. Dalian Univ. Technol. (Soc. Sci)*, 3, 8-16, 2018.
- [11] Lane, B., & Potter, S., "The adoption of cleaner vehicles in the UK: Exploring the consumer attitude-action gap," *Journal of Cleaner Production*, 15, 1085-1092, 2007.
- [12] Skippon, S., & Garwood, M., "Responses to battery electric vehicles: UK consumer attitudes and attributions of symbolic meaning following direct experience to reduce psychological distance," *Transportation Research Part D*, 16, 525-531, 2011.
- [13] Jensen, A. F., Cherchi, E., & Mabit, S. L., "On the stability of preferences and attitudes before and after experiencing an electric vehicle," *Transportation Research Part D. Transport and Environment*, 25, 24-32, 2013.
- [14] Wijaya, A., & Sensuse, D. I., "Perencanaan Strategis Sistem Informasi dan Teknologi Informasi Pada Perusahaan Otomotif dengan Menggunakan Metodologi Tozer," *Seminar Nasional Aplikasi Teknologi Informasi*, 2011.
- [15] Databooks, "Data Publish," *Online, KataData*, 2018.
- [16] Nazir, M., "Metode Penelitian," *Bogor: Ghalia Indonesia*, 2017.
- [17] Mujilawati, S., "Pre-processing Text Mining Pada Data Twitter," *Sentika*. pp.49-56, 2016.
- [18] Witten. H.I and Frank.E, "Data mining Practical Machine Learning Tools and Techniques Second Edition," *Elsevier, SanFrancisco*, 2005.
- [19] Yousef, Ahmed Hassan, Walaa Medhat, and Hoda Korashy Mohamed. "Sentiment Analysis Algorithms and Applications: A Survey." (2014).
- [20] Muhazir, Nurul Huzna., Omar, Faizal Mohd., Nawi, Mohd Nasrun Mohd. "Sentiment Analysis Visualization System for The Property Industry." (2018).
- [21] Alrajak, M.Suyudi., Ernawatu, Iin., Nurlaili, Ika. "Analisis Sentimen terhadap Pelayanan PT PLN di Jakarta pada Twitter dengan Algoritma K-Nearest Neighbor (K-NN)." (2020).

