

Analisis Sentimen Publik Indonesia Terhadap Motor Listrik pada Media Sosial Twitter

Imam Adi Nata¹, Damar Wicaksono², D. Jayus Nor Salim³

^{1,2,3} Jurusan Teknik Elektro Fakultas Teknik Universitas Tidar
Jl. Kapten S. Parman No. 39 Magelang 56116 INDONESIA

imamadinata@untidar.ac.id¹, damar@untidar.ac.id², djayus.nur@untidar.ac.id³

Intisari—Penggunaan kendaraan bermotor berbahan bakar minyak saat ini sedang ditekan di berbagai negara, termasuk di Indonesia. Pada tahun 2019, Presiden Joko Widodo meluncurkan program "Kendaraan Listrik Indonesia Goes to Campus" yang bertujuan untuk memperkenalkan kendaraan listrik kepada mahasiswa dan masyarakat luas di Indonesia. Program ini melibatkan beberapa perguruan tinggi dan perusahaan otomotif nasional dengan tujuan mempromosikan penggunaan kendaraan listrik yang ramah lingkungan. Riset mengenai analisis perbincangan publik tentang kendaraan listrik di media sosial menjadi perhatian para peneliti di seluruh dunia. Media sosial dapat digunakan sebagai sumber informasi tentang kampanye kendaraan listrik. Tujuan dari riset ini adalah untuk menganalisis tanggapan masyarakat terhadap kebijakan pemerintah mengenai kendaraan listrik. Dalam penelitian ini, akan dibandingkan lima metode machine learning untuk analisis sentimen publik dan hasil dari analisis sentimen menggunakan metode yang terpilih akan diungkapkan.

Keywords— Motor Listrik, Machine Learning, Analisis Sentimen.

I. PENDAHULUAN

Penggunaan kendaraan bermotor bahan bakar minyak saat ini mulai ditekan di beberapa negara salah satunya di Indonesia. Salah satu dampak dari kendaraan berbahan bakar minyak menghasilkan emisi gas buang seperti hidrokarbon (HC), Sulfur (Sox), karbon monoksida (CO), nitrogen oksida (NOx), Nitrogen dan partikel debu[1], [2]. Emisi ini dapat berdampak pada kualitas udara dan kesehatan manusia, terutama jika terjadi di daerah perkotaan dengan lalu lintas yang padat. Polusi udara yang dihasilkan oleh kendaraan dapat menyebabkan iritasi saluran pernapasan, sakit kepala, dan masalah kesehatan lainnya[2].

Selain itu, penggunaan bahan bakar minyak dari fosil, juga dapat berdampak pada perubahan iklim global. Penggunaan bahan bakar fosil mempercepat peningkatan suhu global karena emisi gas rumah kaca yang dihasilkan[3]. Hal ini dapat menyebabkan perubahan yang signifikan pada lingkungan, termasuk peningkatan tingkat laut, cuaca yang lebih ekstrem, dan kekeringan yang lebih sering terjadi.

Untuk mengurangi dampak negatif kendaraan bermotor yang menggunakan bahan bakar minyak terhadap lingkungan, sejumlah upaya dapat dilakukan. Misalnya, meningkatkan efisiensi bahan bakar kendaraan dengan teknologi yang lebih baik dan mengurangi jumlah kendaraan di jalan dengan cara menggunakan transportasi umum atau bersepeda. Selain itu, mengurangi penggunaan kendaraan pribadi dan mengganti kendaraan bermotor dengan kendaraan listrik atau kendaraan bertenaga alternatif dapat membantu mengurangi dampak negatif kendaraan bermotor terhadap lingkungan.

Pemerintah Indonesia mulai mempromosikan kendaraan listrik pada tahun 2015 melalui Program Mobil Listrik Nasional

(National Electric Car program) yang diluncurkan oleh Kementerian Perindustrian. Program ini bertujuan untuk mempercepat pengembangan kendaraan listrik di Indonesia melalui kerja sama antara pemerintah, industri, dan akademisi.

Selain itu, pada tahun 2019, Presiden Joko Widodo juga meluncurkan program KENDARAAN LISTRIK INDONESIA-GOES TO CAMPUS, yang bertujuan untuk memperkenalkan kendaraan listrik kepada mahasiswa dan masyarakat luas di Indonesia. Program ini melibatkan beberapa perguruan tinggi di Indonesia dan perusahaan otomotif nasional.

Selain itu, pemerintah Indonesia juga memberikan insentif untuk pembelian kendaraan listrik, seperti pembebasan pajak kendaraan bermotor (PKB) selama lima tahun dan pembebasan pajak penjualan atas barang mewah (PPnBM) selama 10 tahun untuk kendaraan listrik.

Meskipun demikian, kendaraan listrik masih belum banyak digunakan di Indonesia. Hal tersebut dapat dilihat kendaraan listrik di jalan Indonesia masih menjadi kendaraan minoritas. Kejadian ini memicu banyaknya penelitian tentang sikap masyarakat Indonesia tentang kendaraan listrik.

Penelitian ini akan membandingkan lima metode klasifikasi untuk mengetahui sentiment masyarakat terhadap kebijakan motor listrik di Indonesia. Lima metode yang akan dibandingkan yaitu Support Vector Machine(SVM), K-Nearest Neighbor(KNN), Decision Tree Classifier, Naïve Bayes Classifier, dan Random Forest. Selanjutnya metode yang paling optimal akan digunakan untuk mengetahui sentimen publik terhadap kebijakan motor listrik di Indonesia.

II. METODE

A. Pengambilan Data Tweet

Data yang digunakan dalam eksperimen ini diambil dari Twitter API antara bulan Januari sampai dengan April tahun 2023. Data berhasil dikumpulkan sebanyak 11731 tweet dengan beberapa atribut dari pengguna Twitter di Indonesia dengan kata kunci "Motor Listrik". Akan tetapi, dalam penelitian ini hanya akan menggunakan dua atribut yaitu atribut username dan tweet. Atribut yang akan digunakan dapat dilihat pada TABEL 1.

TABEL 1
ATRIBUT TWEET

| Atribut | Deskripsi |
|----------|---|
| Username | Nama user unik yang digunakan dalam twitter |
| Tweet | Tulisan dari pengguna di media social twitter |

B. Pre Processing

Langkah selanjutnya adalah data *preprocessing*. Secara umum data tweet terdapat banyak *noise*. Kata yang salah eja,

kata yang tidak mempunyai arti, juga kata-kata singkatan yang umum ditemukan pada kalimat. Oleh sebab itu diperlukan pra proses sebelum mengekstrasi fitur yang ada di dalam tweet[4]. Pra proses dilakukan dengan beberapa langkah, Langkah tersebut adalah:

- 1) *Tokenisasi* : Proses ini dilakukan untuk memisah kata-kata dalam kalimat menjadi lebih spesifik.
- 2) *Normalisasi* : Proses normalisasi dilakukan untuk mengubah kata kedalam bentuk standar. Proses ini meliputi *Case Folding* yaitu proses untuk mengubah huruf dari kapital menjadi huruf huruf kecil dan penghapusan tanda hubung dan karakter yang tidak dibutuhkan
- 3) *Pembersihan* : Proses ini meliputi penghapusan simbol-simbol yang umum ditemukan pada tweet seperti URL, @ dalam username, hastags (#).
- 4) *Penghapusan Stopword* : Menghapus kata-kata yang tidak ada arti misal kata penghubung. Kata tersebut akan dihapus dalam 2 langkah pra proses.
- 5) *Pemberian Label pada Tweet* : Proses ini dilakukan untuk membuat data latih pada *machine learning*. Tweet yang telah terkumpul akan diambil secara acak sejumlah 500 data dan akan diberi label positif atau negative.

C. Analisis Sentimen

Analisis sentimen merupakan suatu proses untuk memahami sebuah text secara otomatis dengan tujuan untuk mengambil informasi yang terkandung dalam suatu kalimat[5]. Informasi tersebut berisi tentang studi analitis terhadap sikap dan emosi seseorang terhadap suatu entitas yang diwakili oleh individu atau kelompok[6]. Penelitian ini akan membandingkan lima metode dalam machine learning untuk melakukan analisa sentimen pada twitter. Lima metode tersebut adalah sebagai berikut:

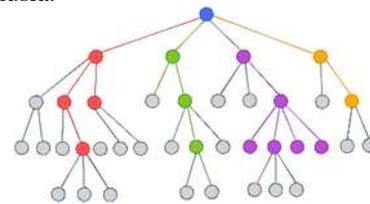
- 1) *Support Vector Machine (SVM)* : SVM merupakan Support Vector Machine. SVM adalah sebuah classifier linear biner non-probabilistik. Untuk sekumpulan data latihan yang terdiri dari titik-titik (x_i, y_i) , di mana x adalah vektor fitur, dan y adalah kelasnya, SVM bertujuan untuk menentukan hiperplane dengan margin maksimum yang memisahkan titik-titik dengan $x_i=1$ dan $x_i=-1$. Persamaan hiperplane tersebut adalah $w \cdot x + b = 0$. Untuk sebuah himpunan data yang terdiri dari himpunan fitur dan himpunan label, classifier SVM membangun sebuah model untuk memprediksi kelas-kelas pada contoh-contoh baru. SVM akan mengklasifikasikan kasus atau titik-titik data baru ke salah satu dari kategori-kategori yang ada[7].

Algoritma :

- Tentukan sebuah hiperplane optimal
 - Perluas langkah I untuk masalah-masalah yang tidak dapat dipisahkan secara linear
 - Pemetaan data ke dalam ruang berdimensi tinggi di mana mudah untuk mengklasifikasikan dengan menggunakan permukaan keputusan linear
- 2) *K-Nearest Neighbors (KNN)* : K-Nearest Neighbor (KNN) adalah metode sederhana pada algoritma machine learning. Algoritma ini bertujuan untuk mengklasifikasikan objek ke dalam salah satu kelas

yang telah ditetapkan dari kelompok sampel yang telah dibuat oleh machine learning. Algoritma KNN tidak membutuhkan pelatihan data untuk melakukan klasifikasi. Pada saat fase pengujian data pelatihan dapat digunakan [8].

- 3) *Decision Tree* : Decision tree adalah sebuah metode dalam machine learning yang digunakan untuk pengambilan keputusan berdasarkan serangkaian aturan dan kondisi. Pohon keputusan terdiri dari simpul-simpul yang mewakili keputusan atau kondisi, serta cabang-cabang yang menghubungkan simpul-simpul tersebut. Setiap simpul dalam pohon keputusan menunjukkan suatu fitur atau atribut yang digunakan untuk membagi data menjadi subset yang lebih kecil. Melalui serangkaian pertanyaan atau pengujian kondisi pada setiap simpul, pohon keputusan dapat mencapai keputusan akhir atau klasifikasi untuk suatu sampel data. Pohon keputusan sangat populer karena sifatnya yang mudah dipahami dan diinterpretasikan oleh manusia.

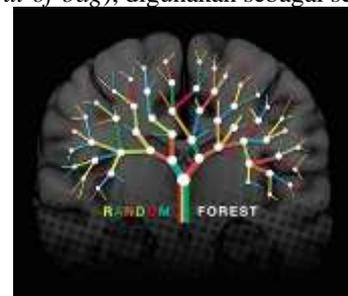


Gbr. 1 Decision Tree

- 4) *Multinomial Naive Bayes* : Teorema Naive Bayes adalah inti dari klasifikasi Naive Bayes, yang menggunakan kemungkinan untuk mengklasifikasikan set data. Perhitungan Naive Bayes Multinomial ditunjukkan dalam persamaan 1 [4].

$$P(X|Y) = \frac{P(Y|X) \cdot P(X)}{P(Y)} \quad (1)$$

- 5) *Random forest (RF)* : Random Forest adalah kumpulan dari Classification and Regression Trees (CART) yang dilatih pada dataset dengan ukuran yang sama dengan set pelatihan, yang disebut sebagai bootstrap, yang dibuat melalui pengambilan sampel acak pada set pelatihan itu sendiri. Setelah sebuah pohon dibangun, sekumpulan bootstraps yang tidak mencakup rekaman tertentu dari dataset asli sampel di luar tas (*out-of-bag*), digunakan sebagai set uji[9].



Gbr. 2 Random Forest[9]

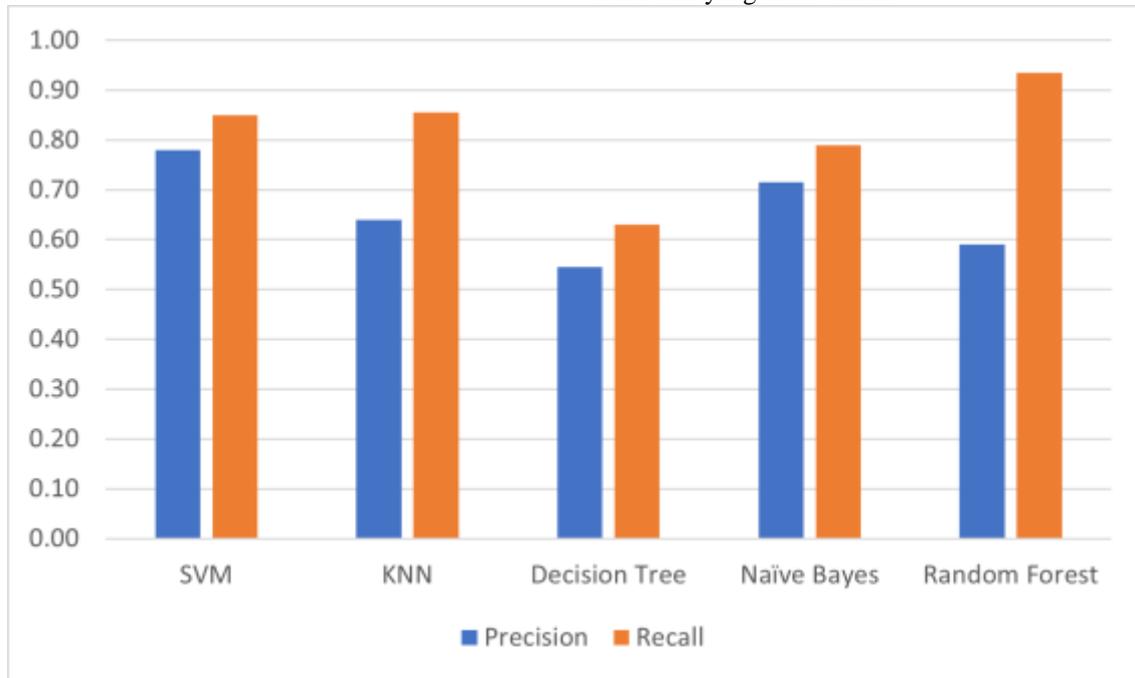
III. HASIL DAN PEMBAHASAN

Kami telah melakukan pembobotan klasifikasi dengan beberapa metode machine learning. Metode evaluasi yang digunakan untuk mengukur kinerja dari model klasifikasi adalah *precision*, *recall*, dan *F1-Score*[10]. Metode ini berguna

untuk melihat model klasifikasi mana yang memiliki kinerja paling optimal.

Metode pengukuran yang umum digunakan adalah presisi dan recall. Metode ini digunakan untuk mengekstrak fitur, opini, dan informasi [11].

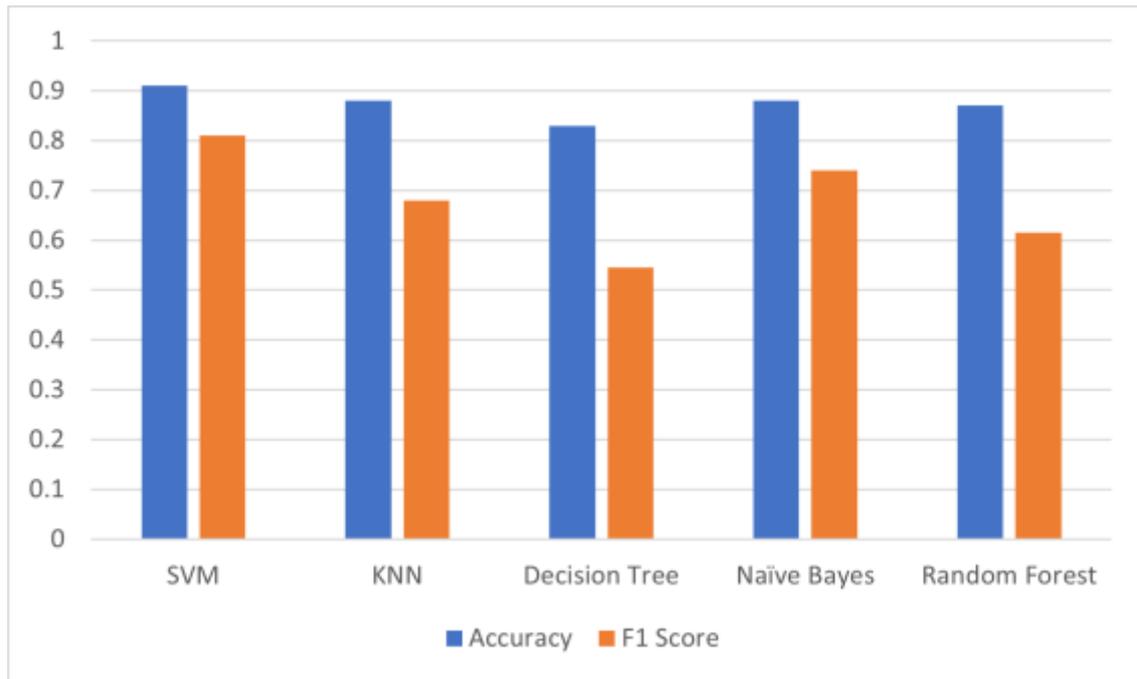
Presisi digunakan untuk mengukur tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem. Sedangkan recall adalah tingkat keberhasilan dalam menemukan kembali sebuah informasi sedangkan F1 Score merupakan perbandingan rata-rata presisi dan recall yang dibobotkan.



Gbr. 3 Perbandingan hasil perhitungan precision dan recall dari beberapa algoritma

Hasil perhitungan precision dan recall dapat dilihat pada Gambar 3. Gambar 3 menunjukkan bahwa metode *support vector machine* memiliki nilai precision tertinggi sedangkan pada metode *random forest* memiliki nilai recall tertinggi. Hal ini menunjukkan bahwa *support vector machine* dan *random forest* merupakan metode yang terbaik dibanding metode lain. Bagian ini menyajikan hasil perbandingan akurasi dan *F1-score* dari lima metode klasifikasi yang digunakan. Hasil perhitungan *F1-Score* dan *Accuracy* dari ke-lima metode klasifikasi ditunjukkan dalam Gambar 4. Kami menemukan bahwa hasil tertinggi dari perhitungan *F1-Score* dan *accuracy* terdapat pada algoritma *support vector machine* dengan nilai *F1-Score* sebesar 81% dan *accuracy* sebesar 91% dibanding dengan metode yang lain yang hanya mendapat nilai *F1-Score*

di bawah 75% dengan *accuracy* di bawah 89% sehingga *F1-Score* pada metode *support vector machine* mendapatkan nilai tertinggi. *F1-Score* dan *accuracy* yang tertinggi menunjukkan bahwa *support vector machine* menjadi metode terbaik untuk klasifikasi. Peringkat kedua diperoleh model klasifikasi *Multinomial Naive Bayes* dengan nilai *F1-Score* 74% dan *accuracy* 88%. Peringkat ketiga diperoleh model klasifikasi *K-Nearest Neighbors* (KNN) dengan *F1-Score* 68% dan *accuracy* 88%. Peringkat keempat diperoleh model klasifikasi *random forest* dengan *F1-Score* 62% dan *accuracy* 88% dan peringkat terakhir model *decision tree* dengan *F1-Score* 55% dan *akurasi* 83%. Di dalam hasil analisis kami *support vector machine* menyajikan akurasi yang lebih baik dari metode lain [12].

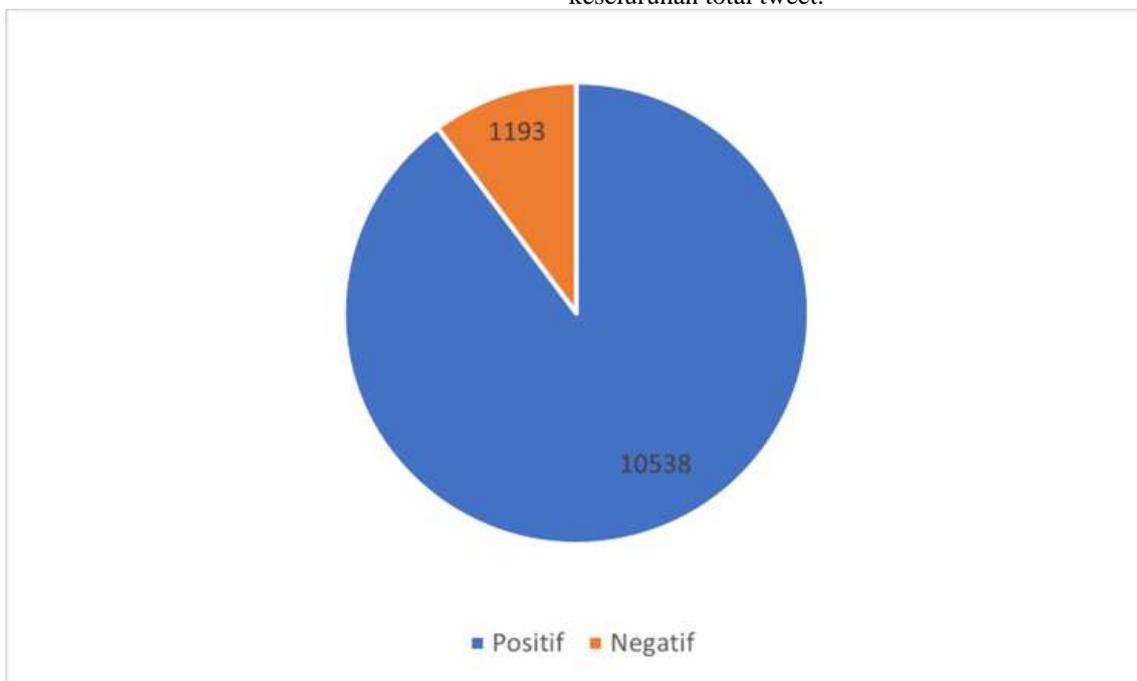


Gbr. 4 Perbandingan hasil perhitungan F1-Score dan accuracy dalam beberapa algoritma

Setelah mendapatkan metode terbaik, Langkah selanjutnya adalah klasifikasi sentimen. Penelitian ini bertujuan untuk menerapkan klasifikasi tweet terkait motor listrik di Indonesia. Klasifikasi bertujuan untuk menetapkan nilai positif dan negative dari suatu tweet. Model klasifikasi yang telah dibuat selama proses pelatihan digunakan untuk melakukan klasifikasi. Semua tweet tentang motor listrik di Indonesia akan

diberi label sentiment secara otomatis. Hasil klasifikasi dapat dilihat pada Gambar 5.

Hasil analisis sentimen menunjukkan bahwa pandangan publik Indonesia terkait motor listrik pada bulan Januari 2023 sampai dengan bulan April 2023 cukup baik dengan hasil positif sebanyak 10538 atau 90% sedangkan hasil analisis sentimen negatif sebanyak 1193 tweet atau 10% dari keseluruhan total tweet.



Gbr. 5 Hasil perhitungan sentimen positif dan negative

IV. PENUTUP

Dalam makalah ini kami membahas isu terkait motor listrik di Indonesia. Kami telah mengumpulkan data yang bersumber dari twitter mulai dari bulan Januari sampai dengan

bulan April tahun 2023. Kami melakukan analisis terhadap data tersebut dengan pendekatan *text mining* dan *machine learning*. Kami mendemonstrasikan lima metode *machine learning* untuk mencari metode yang paling baik dan dapat digunakan

untuk melakukan analisis sentimen publik di Indonesia terkait motor listrik. Hasil dari pembelajaran mesin dan pendekatan statistik tersebut cukup dapat dipahami. Dari hasil tersebut dapat disimpulkan bahwa masyarakat Indonesia khususnya di twitter memiliki kepedulian yang besar terhadap kebijakan motor listrik di Indonesia. Oleh karena itu pemerintah untuk melanjutkan kampanye penggunaan motor listrik di Indonesia untuk meningkatkan kesadaran masyarakat Indonesia dalam kebijakan motor listrik untuk membantu mengurangi kerusakan lingkungan akibat dari penggunaan bahan bakar fosil.

REFERENSI

- [1] K. Berbahan *et al.*, “BAHAYA EMISI GAS BUANG KENDARAAN BERBAHAN BAKAR BENSIN DAN MENUMBUHKAN LINGKUNGAN HIJAU DI PERKOTAAN,” *Prosiding Seminar Nasional Hukum, Bisnis, Sains dan Teknologi*, vol. 1, pp. 101–101, Sep. 2020, Accessed: May 24, 2023. [Online]. Available: <http://ojs.uadb.ac.id/index.php/HUBISINTEK/article/view/985>
- [2] A. Y. P. Wardoyo, “Emisi Partikulat Kendaraan Bermotor dan Dampak Kesehatan,” pp. xx–183, 2016.
- [3] S. Sulistyono, “PEMANASAN GLOBAL (GLOBAL WARMING) DAN HUBUNGANNYA DENGAN PENGGUNAAN BAHAN BAKAR FOSIL,” *Swara Patra : Majalah Ilmiah PPSDM Migas*, vol. 2, no. 2, Dec. 2012, Accessed: May 25, 2023. [Online]. Available: <http://ejournal.ppsdmmigas.esdm.go.id/sp/index.php/swarapatra/article/view/60>
- [4] M. Habibi, A. Priadana, M. Rifqi, M. ' Arif, and A. Yani Yogyakarta, “Sentiment Analysis and Topic Modeling of Indonesian Public Conversation about COVID-19 Epidemics on Twitter,” *IJID (International Journal on Informatics for Development)*, vol. 10, no. 1, pp. 23–30, Jun. 2021, doi: 10.14421/IJID.2021.2400.
- [5] G. A. Buntoro, “Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter,” *INTEGER: Journal of Information Technology*, vol. 2, no. 1, pp. 32–41, Mar. 2017, Accessed: May 25, 2023. [Online]. Available: <http://ejournal.itats.ac.id/integer/article/view/95>
- [6] R. Jena, “An empirical case study on Indian consumers’ sentiment towards electric vehicles: A big data analytics approach,” *Industrial Marketing Management*, vol. 90, pp. 605–616, Oct. 2020, doi: 10.1016/J.INDMARMAN.2019.12.012.
- [7] S. G. Kanakaraddi, A. K. Chikaraddi, K. C. Gull, and P. S. Hiremath, “Comparison Study of Sentiment Analysis of Tweets using Various Machine Learning Algorithms,” *Proceedings of the 5th International Conference on Inventive Computation Technologies, ICICT 2020*, pp. 287–292, Feb. 2020, doi: 10.1109/ICICT48043.2020.9112546.
- [8] M. Aziz, Y. Marcellino, I. A. Rizki, S. A. Ikhwanuddin, and J. W. Simatupang, “STUDI ANALISIS PERKEMBANGAN TEKNOLOGI DAN DUKUNGAN PEMERINTAH INDONESIA TERKAIT MOBIL LISTRIK,” *TESLA: Jurnal Teknik Elektro*, vol. 22, no. 1, pp. 45–55, Mar. 2020, doi: 10.24912/TESLA.V22I1.7898.
- [9] A. Sarica, A. Cerasa, and A. Quattrone, “Random forest algorithm for the classification of neuroimaging data in Alzheimer’s disease: A systematic review,” *Front Aging Neurosci*, vol. 9, no. OCT, p. 329, Oct. 2017, doi: 10.3389/FNAGI.2017.00329/BIBTEX.
- [10] “A Comparative Study of Classification Techniques in Data Mining Algorithms | Oriental Journal of Computer Science and Technology.” <http://www.computerscijournal.org/vol8no1/a-comparative-study-of-classification-techniques-in-data-mining-algorithms/> (accessed May 25, 2023).
- [11] A. A. Anees, H. Prakash Gupta, A. P. Dalvi, S. Gopinath, and B. R. Mohan, “Performance analysis of multiple classifiers using different term weighting schemes for sentiment analysis,” *2019 International Conference on Intelligent Computing and Control Systems, ICCS 2019*, pp. 637–641, May 2019, doi: 10.1109/ICCS45141.2019.9065895.
- [12] A. Rahman and M. S. Hossen, “Sentiment Analysis on Movie Review Data Using Machine Learning Approach,” *2019 International Conference on Bangla Speech and Language Processing, ICBSLP 2019*, Sep. 2019, doi: 10.1109/ICBSLP47725.2019.201470.