



Pendekatan *Machine Learning*: Analisis Sentimen Masyarakat Terhadap Kendaraan Listrik Pada Sosial Media X

Gathot Hanyokro Kusuma*, Inggih Permana, Febi Nur Salisah, M. Afdal, Muhammad Jazman, Arif Marsal

12050313348@students.uin-suska.ac.id*

*Penulis korespondensi

Universitas Islam Negeri Sultan Syarif Kasim Riau - Indonesia

Diterima: 27 Nov 2023 | Direvisi: 04 – 09 Des 2023
Disetujui: 26 Des 2023 | Dipublikasi: 30 Des 2023
Program Studi Sistem Informasi, Fakultas Sains dan Teknologi,
Universitas Islam Negeri Raden Fatah Palembang, Indonesia

ABSTRACT

Environmental issues and the depletion of fossil fuels continue to escalate as the number of fossil fuel-based vehicle users increases in Indonesia. Electric vehicles emerge as one of the potential alternative solutions to address current environmental challenges, given their eco-friendly nature and lack of pollution emissions. Sentiment analysis is conducted to understand public responses, both supportive and opposing, towards electric vehicles. This research aims to analyze the sentiment of X-social media users regarding electric vehicles using machine learning techniques. The research stages include data collection, data selection, preprocessing, and classification using Naïve Bayes Classifier (NBC), Support Vector Machine (SVM), and K-Nearest Neighbor (KNN) algorithms. The test results show that on a balanced dataset using ROS, SVM performs the best with accuracy = 68.7%, precision = 77.9%, and recall = 68.4%. Meanwhile, NBC yields an accuracy of 60.3%, precision of 61.3%, and recall of 60.3%, while KNN has an accuracy of 53.9%, precision of 54%, and recall of 53.9%.

Keywords: Sentiment Analysis, K-Nearest Neighbor, Naïve Bayes Classifier, Support Vector Machine

ABSTRAK

Masalah lingkungan dan penipisan bahan bakar minyak bumi terus meningkat karena jumlah pengguna kendaraan dengan bahan bakar minyak bumi terus meningkat di Indonesia. Kendaraan listrik menjadi salah satu solusi alternatif potensial untuk mengatasi tantangan lingkungan saat ini, mengingat sifatnya yang ramah lingkungan dan tidak menyebabkan polusi. Analisis sentimen dilakukan untuk memahami tanggapan masyarakat, baik yang mendukung maupun yang tidak mendukung terhadap kendaraan listrik. Penelitian ini bertujuan untuk menganalisis sentimen pengguna media sosial X mengenai kendaraan listrik dengan menggunakan teknik machine learning. Tahapan penelitian mencakup pengumpulan data, seleksi data, pra-pemrosesan, dan klasifikasi menggunakan algoritma Naïve Bayes Classifier (NBC), Support Vector Machine (SVM), dan K-Nearest Neighbor (KNN). Hasil pengujian menunjukkan bahwa pada dataset yang seimbang menggunakan ROS, SVM memberikan kinerja terbaik dengan akurasi = 68,7%, presisi = 77,9%, dan recall = 68,4%. Sementara itu, NBC menghasilkan akurasi 60,3%, presisi 61,3%, dan recall 60,3%, sedangkan KNN memiliki akurasi 53,9%, presisi 54%, dan recall 53,9%.

Kata Kunci: Analisis Sentimen, K-Nearest Neighbor, Naïve Bayes Classifier, Support Vector Machine

PENDAHULUAN

Jumlah kendaraan bermotor di Indonesia terus meningkat dari tahun ke tahun, meliputi sepeda motor, kendaraan mobil seperti mobil penumpang, mobil barang, dan mobil bis (Sidik & Ansawarman, 2022). Peningkatan jumlah kendaraan bermotor tersebut menunjukkan bahwa penggunaan Bahan Bakar Minyak (BBM) menjadi terus bertambah banyak sesuai dengan peningkatan tersebut. Menurut data Badan Pusat Statistik, jumlah kendaraan bermotor terus bertambah dari 126,5 juta unit pada tahun 2018 menjadi 148,3 juta unit pada tahun 2022 (Badan Pusat Statistik, 2022). Dengan hampir seluruh kendaraan menggunakan BBM, peningkatan jumlah pengguna kendaraan telah menyebabkan stok BBM yang semakin menipis dan kenaikan harga BBM (Erfini & Lestari, 2023). Selain itu, kendaraan yang menggunakan BBM juga menjadi penyumbang emisi dan pencemaran udara terbesar di Indonesia (Ghaniyyu & Husnita, 2021). Hal ini tentunya menimbulkan keprihatinan akan dampak lingkungan yang semakin serius.

Kendaraan listrik menjadi salah satu alternatif yang menjanjikan dalam mengatasi masalah lingkungan dan semakin meningkatnya harga (Efendi, 2020). Kelebihan kendaraan listrik terletak pada minimnya komponen dibandingkan kendaraan konvensional, sehingga perawatannya lebih mudah (Pratama et al., 2023). Diharapkan bahwa dengan keunggulan-keunggulannya, kendaraan listrik dapat secara signifikan menggantikan kendaraan konvensional yang bergantung pada BBM (Alfarizi & Fitriani, 2023; Mughtari et al., 2023). Peluang besar bagi Indonesia untuk mengadopsi kendaraan listrik menjadi nyata karena dukungan pemerintah terhadap kebijakan terkait kendaraan berbasis listrik dan memiliki sumber daya yang melimpah.

Penerbitan Peraturan Pemerintah (PP) Nomor 55 Tahun 2019 mengenai Percepatan Program Kendaraan Listrik Berbasis Baterai untuk angkutan jalan menjadi sorotan utama (I. Pambudi & Juwono, 2023; Subiantoro et al., 2023). Beragam tanggapan dari masyarakat melalui media sosial, terutama di platform seperti Twitter atau sekarang disebut sebagai media sosial X. Meskipun peraturan dan rencana pemerintah terkait transisi ke kendaraan listrik sudah diterbitkan, kehadiran kendaraan listrik di masyarakat masih menuai berbagai pandangan pro dan kontra (Fitrianto, 2023).

Menurut Liu (2012), bahwa analisis sentimen merujuk pada suatu proses sistematis dalam mengolah data tekstual untuk mengekstrak dan mengukur keadaan afektif serta informasi subjektif. Penilaian terkait dengan suatu produk sangat penting bagi calon konsumen (Wang & Wang, 2020). Analisis sentimen memiliki peran penting dalam memberikan wawasan kepada konsumen mengenai pengalaman orang lain. Lebih dari sekadar itu, analisis sentimen juga dapat membantu pemerintah dalam memantau persepsi masyarakat terhadap kendaraan listrik, baik dalam aspek keamanan, kenyamanan, maupun keberlanjutan. Evaluasi respons masyarakat terhadap kebijakan publik yang berkaitan dengan kendaraan listrik juga dapat dilakukan melalui analisis sentimen, yang kemudian dapat digunakan oleh pemerintah untuk merencanakan pengembangan infrastruktur yang mendukung kendaraan listrik (Christidis & Focas, 2019). Peran dari analisis sentimen meliputi pemahaman terhadap opini, penilaian, pandangan, atau perasaan masyarakat terhadap suatu produk, peristiwa, atau jasa (Handayani & Zufria, 2023).

Sosial media X, sebuah platform *microblogging*, memperoleh lebih dari 500 juta tweet setiap harinya (D. I. Pambudi & Sulastri, 2023). Platform ini memfasilitasi pengguna untuk dengan cepat berbagi pendapat dan pandangan mengenai berbagai topik dan produk (Agustian et al., 2022). Pengguna dapat membuat dan melihat tweet yang berisi teks, gambar, dan video tentang berbagai topik (Mustofa & Novita, 2022). Karena kualitas postingannya, sosial media X dipilih sebagai sumber data untuk menganalisis opini dan menemukan polaritas opini masyarakat.

Beberapa penelitian sebelumnya telah memanfaatkan algoritma *Naïve Bayes Classifier (NBC)*, *Support Vector Machine (SVM)*, atau *K-Nearest Neighbors (KNN)* untuk melakukan analisis sentimen. Beberapa rujukan seperti penelitian yang dilakukan oleh [Nurmalasari et al. \(2023\)](#) mengenai analisis menggunakan algoritma *SVM*, *KNN*, dan *NBC*. Hasil penelitian tersebut menunjukkan bahwa *SVM* memiliki akurasi terbaik, meskipun perbedaan performanya tidak signifikan dibandingkan dengan *NBC* dan *KNN*. Selanjutnya, di beberapa penelitian terkait algoritma *NBC* dan *SVM* menyebutkan hasil performa yang berbeda-beda. Beberapa menyebutkan *SVM* memiliki performa yang lebih baik ([Noviana & Rasal, 2023](#); [Rahayu et al., 2022](#); [Saddam et al., 2023](#)), akan tetapi [Wayan Ernawati et al. \(2023\)](#) menyebutkan bahwa *NBC* lebih baik.

Pada penelitian ini memanfaatkan teknik *Machine Learning (ML)* untuk melakukan analisis sentimen terhadap topik kendaraan listrik di sosial media X. Teknik *ML* yang digunakan meliputi *NBC*, *SVM*, dan *KNN*. Selain itu, penelitian ini juga menggunakan teknik *balancing data* guna meningkatkan performa model. Teknik *balancing* yang digunakan yaitu *Random Over Sampling (ROS)* dan *Random Under Sampling (RUS)* ([Fitriani et al., 2021](#)).

METODOLOGI PENELITIAN

Di dalam penelitian menggunakan langkah-langkah sistematis, yang terdiri dari empat tahapan penting seperti pada Gambar 1.



Gambar 1. Tahapan Penelitian

Pengumpulan Data

Dalam penelitian ini, data dikumpulkan menggunakan teknik *crawling data* yang diimplementasikan menggunakan bahasa pemrograman Python. Untuk melakukan *crawling*, digunakan *library* Python seperti *Tweet Harvesting*, *Node.js*, dan *Pandas*. *Keyword* yang digunakan adalah "kendaraan listrik". Data tweet yang terkait dengan topik kendaraan listrik dari bulan Januari 2023 hingga Juli 2023 diambil, kemudian disimpan dalam format *csv*.

Seleksi Data

Tahap seleksi data melibatkan penghapusan tweet yang tidak relevan dengan tujuan penelitian, seperti tweet yang tidak memuat pendapat tentang kendaraan listrik. Selain itu, semua tweet yang tidak menggunakan Bahasa Indonesia akan dieliminasi dari dataset.

Pra-proses Data

Tahap pra-proses ini dilakukan dengan menggunakan bahasa pemrograman Python. Beberapa proses pra-proses meliputi: **Pelabelan**, dimana tweet diberi label positif, negatif, atau netral oleh seorang annotator yang memiliki pengetahuan baik dalam Bahasa Indonesia; **Case Folding**, yang melibatkan perubahan semua huruf dalam tweet menjadi huruf kecil; **Cleaning**, yang menghapus karakter "@" (@), hashtag (#), emoticon, angka, dan tanda baca lainnya dari tweet; **Tokenizing**, yang memisahkan

tweet menjadi kata-kata berdasarkan spasi; **Normalisasi**, yang mengubah kata-kata singkatan atau tidak formal menjadi bahasa baku dalam Bahasa Indonesia; **Penghapusan Stopword**, yang menghilangkan kata-kata umum yang dianggap tidak penting seperti kata penghubung, kata ganti, dan kata depan; **Stemming**, yang mengubah kata-kata berimbuhan menjadi kata dasar; **TF-IDF**, yang memberi bobot pada setiap kata dalam tweet menggunakan *Term Frequency* dan *Inverse Document Frequency*; **Balancing Data**, yang memberikan prioritas yang sama pada setiap kelas dengan menggunakan metode *ROS* dan *RUS*. *ROS* menambahkan data pada kelas minoritas untuk menyamai jumlah data pada kelas mayoritas (Aryanti et al., 2023), sedangkan *RUS* mengurangi jumlah data pada kelas mayoritas (Purnajaya & Hanggara, 2021).

Pemodelan dan Evaluasi

Pemodelan dilakukan menggunakan algoritma *NBC*, *KNN*, dan *SVM* dengan menggunakan *Google Colabs* dan Bahasa Pemrograman Python sebagai *tool*-nya. Parameter-parameter yang digunakan untuk masing-masing algoritma tertera pada Tabel 1. Metode *tuning parameter* yang diterapkan adalah *Grid Search*, sementara metode evaluasi menggunakan *K-Fold Cross Validation* dengan nilai $k = 10$. Performa model dianalisis melalui *confusion matrix*, akurasi pada Persamaan (1), presisi pada Persamaan (2), dan *recall* pada Persamaan (3). Setelah itu, dilakukan perbandingan performa model klasifikasi sentimen yang dihasilkan.

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (1)$$

$$Presisi = \frac{TP}{TP+FP} \times 100\% \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (3)$$

Tabel 1. Parameter-Parameter Yang Digunakan

No.	Algoritma	Parameter	
1.	<i>KNN</i>	K	2; 3; 4; 5; 6; 7; 8; 9; 10
2.	<i>NBC</i>	-	-
3.	<i>SVM</i>	Kernel	Linear
		Kernel	Polynomial
		Gamma	Auto
		Cost (C)	1.00; 2.00; 3.00
		Degree (D)	1.00; 2.00; 3.00
		Kernel	Radial Basis Function (RBF)
		Gamma	Auto
		Kernel	Sigmoid
		Gamma	Scale
		Cost (C)	1.00; 2.00; 3.00

HASIL DAN PEMBAHASAN

Implementasi Pengumpulan Data, Seleksi, dan Pra-proses Data

Berdasarkan pengambilan data tweet dari sosial media X data periode januari 2023 sampai juli 2023, didapatkan sebanyak 2.341 tweet. Kemudian, setelah dilakukan seleksi didapatkan data sebanyak 1.527 tweet. *Library* python yang digunakan dapat dilihat pada Gambar 2. Hasil pra-proses data dapat dilihat pada Gambar 2. Hasil *TF-IDF* dapat dilihat pada Tabel 3 yang terdapat 3.452 kata. Statistik data setelah dilakukan *balancing* dapat dilihat pada Tabel 4.

```
[ ] pip install pandas

# Install Node.js (because tweet-harvest built using Node.js)
!sudo apt-get uourcepdate
!sudo apt-get install -y ca-certificates curl gnupg
!sudo mkdir -p /etc/apt/keyrings
!curl -fsSL https://deb.nodesource.com/gpgkey/nodesource-repo.gpg.key | sudo gpg --dearmor -o /etc/apt/keyrings/nodesource.gpg
!NODE_MAJOR=20 && echo "deb [signed-by=/etc/apt/keyrings/nodesource.gpg] https://deb.nodesource.com/node_${NODE_MAJOR}.x nodistro main" | sudo te
!sudo apt-get update
!sudo apt-get install nodejs -y

!node -v

Show hidden output

# Crawl Data
filename = 'apr2.csv'
search_keyword = 'kendaraan listrik until:2023-04-25 since:2023-04-01'
limit = 1000
```

Gambar 2. Pengumpulan Data

Pra-proses merupakan fase dimana data yang telah diperoleh diubah, dihapus, dan dipisahkan. Pra-proses data bertujuan untuk memudahkan pengolahan data pada tahap selanjutnya. Pelabelan data dilakukan oleh annotator yang memiliki latar belakang dalam Bahasa Indonesia. Hasil dari pra-proses data dapat dilihat pada Tabel 2.

Tabel 2. Hasil Pra-proses Data

No.	Tweet	Sentimen	Hasil Pra Proses
1	Beberapa program menteri BUMN @erickthohir seperti : Mendorong produksi minyak goreng merah, menurunkan harga BBM, program solar subsidi untuk nelayan, mengupayakan subsidi BBM tepat sasaran, mendorong energi baru terbarukan, serta penggunaan kendaraanlistrik. BangkitBersamaET https://t.co/MvIrAR9iC5a	positif	[program, menteri, badan, usaha, milik, negara, dorong, produksi, minyak, goreng, merah, turun, harga, bahan, bakar, minyak, program, solar, bantu, nelayan, upaya, bantu, bahan, bakar, minyak, sasar, dorong, energi, baru, guna]
2	@fotodakwah Nabi jg ga main Hp, ga pakai penerangan listrik, ga pakai kendaraan bermotor Lu bid'ah	negatif	[nabi, gak, main, handphone, gak, pakai, terang, listrik, gak, pakai, kendara, motor, bidah]
3	Kemarin kedua tempat tsb digunakan u/ argumen: 1) kondisi transport bagus pun mobil tetap dipakai, 2) transisi kendaraan listrik tetap dijalankan, 3) realistsnya kendaraan pribadi akan tetap ada. Agar jelas, yang dimaksud di sini mobil listrik pribadi. /2	positif	[kemarin, argumen, kondisi, transport, bagus, mobil, pakai, transisi, jalan, wujud, kendaraan, pribadi, mobil, listrik, pribadi]
...
1527	Mendadak mikir: Gimana ya nasibku yang ga bisa naik sepeda ini kalau ada invasi alien yang membuat semua kendaraan bermotor dan listrik mati?	negatif	[dadak, mikir, iya, nasib, gak, sepeda, invasi, alien, kendara, motor, listrik, mati]

Tabel 3. Hasil *TF-IDF*

No.	1 bahan	2 kendaraan	3 listrik	4 Motor	5 0	...	3452 zaman
1	0.227	0	0	0	0.142	...	0
2	0	0.102	0.093	0.142	0	...	0
..	0	0	0	0	0.166	...	0
1527	0	0.119	109	0.166	0	...	0

Tabel 4. Statistik Data

Label	Jumlah		
	Tanpa Balancing	ROS	RUS
Negatif	629	629	423
Netral	423	629	423
Positif	475	629	423
Total	1.527	1.887	1.269

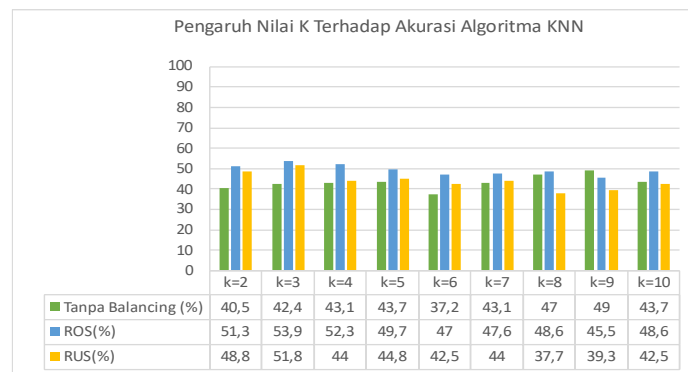
Hasil Pemodelan dan Evaluasi

Setelah data melalui tahap pra-proses, langkah selanjutnya adalah pengolahan data menggunakan algoritma *KNN*, *NBC*, dan *SVM* untuk mengevaluasi kinerja algoritma dalam penelitian ini. Pengolahan ini menggunakan metode *Grid Search* sebagai *tuning parameter*, dengan hasil eksperimen tercatat dalam Tabel 5, dimana A merepresentasikan Akurasi, P untuk Presisi, dan R untuk *Recall*.

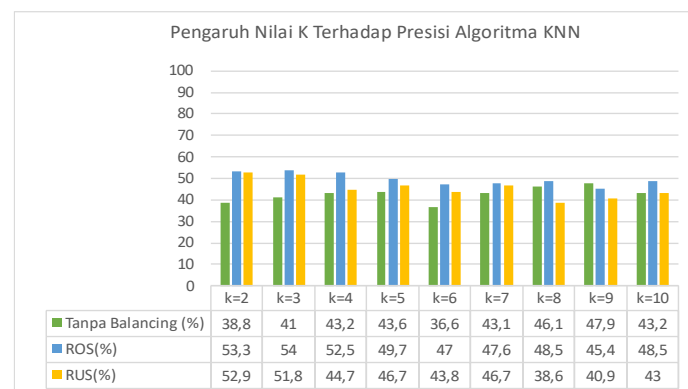
Tabel 5. Hasil Eksperimen

Algoritma	Parameter	Performa (%)								
		Tanpa Balancing			ROS			RUS		
		A	P	R	A	P	R	A	P	R
KNN	K=2	40.5	38.8	37.9	51.3	53.3	51.2	48.8	52.9	48.6
	K=3	42.4	41	41.4	53.9	54	53.9	51.8	51.8	51.6
	K=4	43.1	43.2	42.2	52.3	52.5	52.4	44	44.7	44.7
	K=5	43.7	43.6	43.3	49.7	49.7	49.8	44.8	46.7	45.3
	K=6	37.2	36.6	36.9	47	47	47.1	42.5	43.8	42.9
	K=7	43.1	43.1	43.1	47.6	47.6	47.7	44	46.7	44.7
	K=8	47	46.1	45.8	48.6	48.5	48.8	37.7	38.6	37.9
	K=9	49	47.9	47.9	45.5	45.4	45.5	39.3	40.9	39.7
	K=10	43.7	43.2	43.3	48.6	48.5	48.7	42.5	43	42.8
NBC	-	52.9	52.1	49.2	60.3	61.3	60.3	55.1	54.6	54.5
SVM	Kernel; Linear	57.5	55.8	54.7	62.9	62.9	62.5	55.1	55.2	55.2
	Kernel; Polynomial									
	C=1, D=1	57.5	55.8	54.7	62.9	62.9	62.5	55.1	55.2	55.2
	C=1, D=2	51.6	51.3	46.6	66.1	71.6	65.7	58.2	58.3	57.9
	C=1, D=3	53.5	61.8	47.9	67.1	76.5	66.7	55.1	55.1	53.8
	C=2, D=1	56.8	55.3	55.3	64.4	64	63.9	51.1	51.4	51.5
	C=2, D=2	54.9	53.9	50.8	66.1	69.3	65.9	58.2	58.3	58.1
	C=2, D=3	53.5	56.3	48.4	68.7	77.9	68.4	55.1	54.6	54.3
	C=3, D=1	57.5	56.3	56.3	63.4	64	63.4	50.3	50.8	50.6
	C=3, D=2	56.2	55.4	52.4	67.1	70.7	66.9	59	59	58.9
	C=3, D=3	52.9	55.4	47.9	68.7	77.9	68.4	55.1	54.6	54.3
	Kernels RBF	41.8	13.9	33.3	31.2	10	33	29.9	9.9	33.3
	Kernel Sigmoid:									
(c) = 1	54.3	54.6	54.5	54.3	54.6	54.5	54.3	54.6	54.5	
(c) = 2	48.8	48.9	48.9	48.8	48.9	48.9	48.8	48.9	48.9	
(c) = 3	45.6	45.7	45.7	45.6	45.7	45.7	45.6	45.7	45.7	

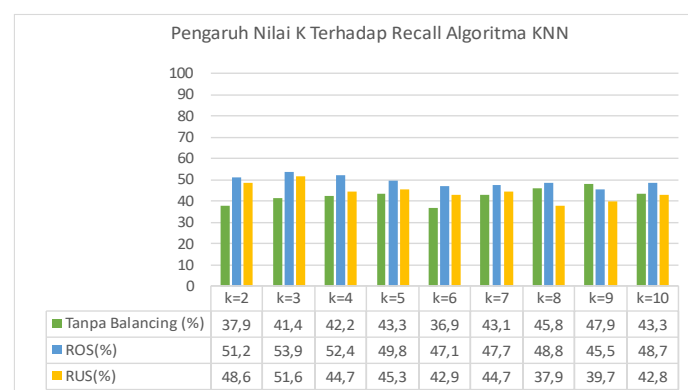
Selanjutnya, pengujian masing-masing algoritma yang telah ditentukan di dalam penelitian ini dilakukan. ilustrasi dalam Gambar 3, Gambar 4, dan Gambar 5 menunjukkan bagaimana nilai K mempengaruhi performa klasifikasi sentimen masyarakat terhadap kendaraan listrik dalam algoritma *KNN*. Dari gambar-gambar tersebut, dapat diamati bahwa performa terbaik terjadi ketika nilai $K = 3$. Performa optimal ini terlihat pada data yang sebelumnya diimbangi menggunakan *ROS*. Meskipun demikian, hasil yang optimal tersebut hanya mencapai akurasi sebesar 53,9%, presisi sebesar 54%, dan *recall* sebesar 53,9%.



Gambar 3. Pengaruh Nilai K Terhadap Akurasi Algoritma *KNN*



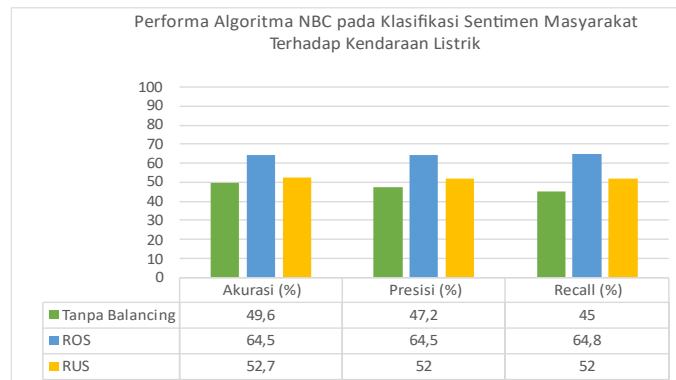
Gambar 4. Pengaruh Nilai K Terhadap Presisi Algoritma *KNN*



Gambar 5. Pengaruh Nilai K Terhadap Recall Algoritma *KNN*

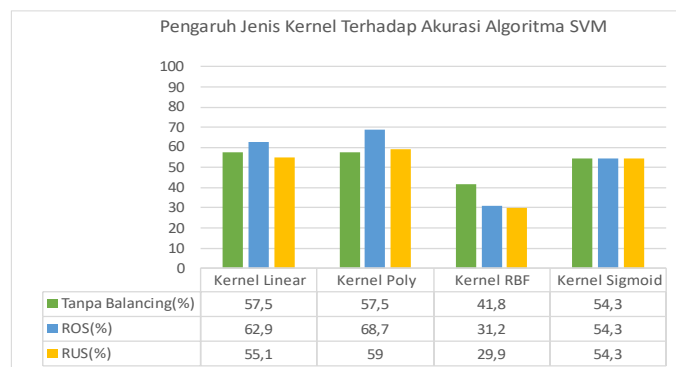
Kemudian, ilustrasi dalam Gambar 6 memperlihatkan perbandingan performa klasifikasi sentimen masyarakat terhadap kendaraan listrik menggunakan

Algoritma *NBC*. Dari gambar tersebut, dapat dilihat bahwa performa terbaik terjadi pada model *NBC* yang didukung oleh *ROS* dengan akurasi, presisi, dan *recall* secara berturut-turut mencapai 64,5%, 64,5%, dan 64,8%.

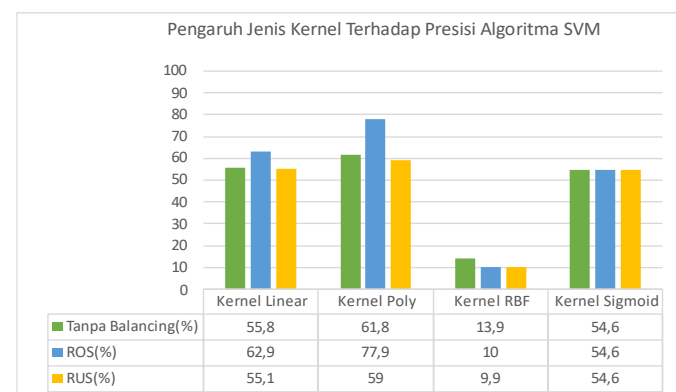


Gambar 6. Performa Algoritma NBC

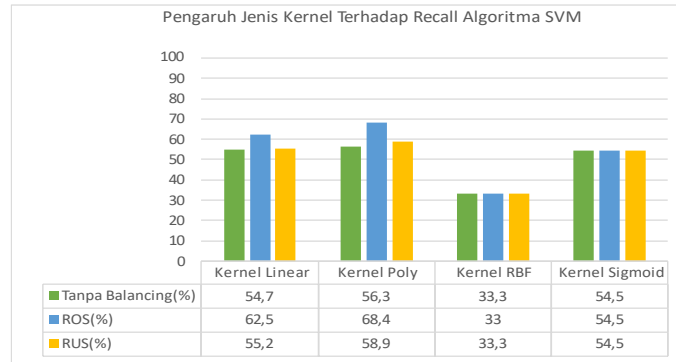
Ilustrasi dalam Gambar 7, Gambar 8, dan Gambar 9 menggambarkan pengaruh jenis kernel dan teknik *balancing data* pada algoritma *SVM* terhadap performa klasifikasi sentimen masyarakat terhadap kendaraan listrik. Dari gambar-gambar tersebut, dapat diamati bahwa performa terbaik terjadi ketika *SVM* menggunakan kernel polynomial (cost = 2.0; degree = 3.0), dengan teknik *balancing data* yang digunakan adalah *ROS*. Performa optimal ini dicapai dengan akurasi sebesar 68,7%, presisi sebesar 77,9%, dan recall sebesar 68,4%.



Gambar 7. Pengaruh Jenis Kernel Terhadap Akurasi Algoritma SVM

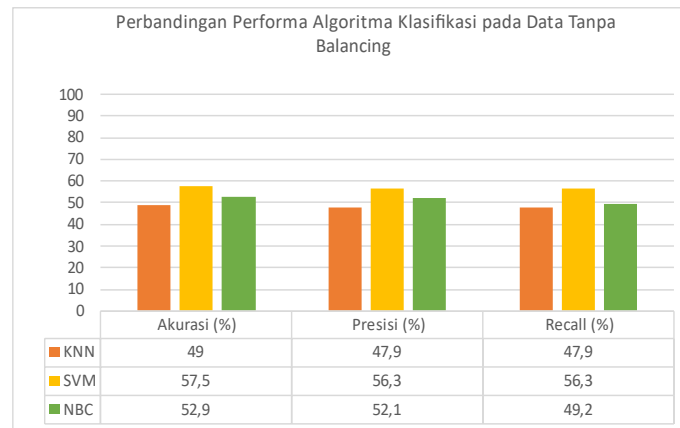


Gambar 8. Pengaruh Jenis Kernel Terhadap Presisi Algoritma SVM

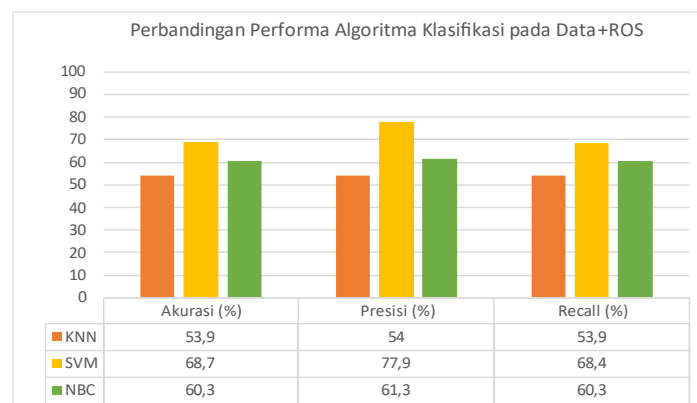


Gambar 9. Pengaruh Jenis Kernel Terhadap Recall Algoritma SVM

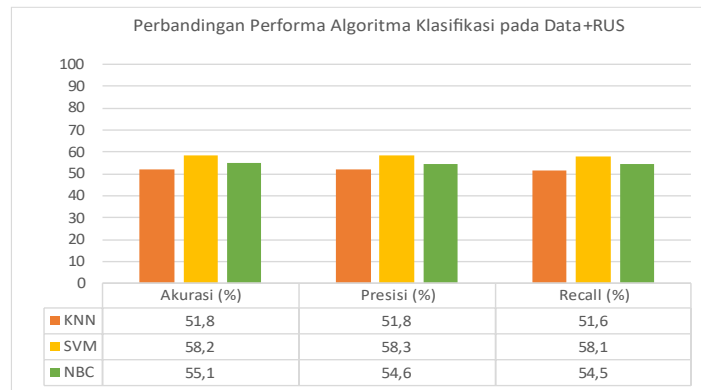
Selanjutnya, pada Gambar 10, Gambar 11, dan Gambar 12 dapat diketahui bahwa *SVM* selalu menjadi algoritma dengan performa terbaik, baik menggunakan teknik *balancing* maupun tanpa *balancing*. Pada gambar-gambar tersebut juga terlihat *SVM* memiliki performa terbaik ketika dilakukan *balancing data* menggunakan *ROS* dengan akurasi = 68,7%, presisi = 77,9%, dan *recall* = 68,4%.



Gambar 10. Perbandingan Performa Algoritma Klasifikasi Pada Data Tanpa *Balancing*



Gambar 11. Perbandingan Performa Algoritma Klasifikasi Pada Data+ROS



Gambar 12. Perbandingan Performa Algoritma Klasifikasi Pada Data+RUS

Di tahap akhir, *confusion matrix* dari hasil pemodelan dengan algoritma SVM Kernel Polynomial dengan $C=2$, $D=3$ disajikan pada Tabel 6.

Tabel 6. Confusion Matrix SVM Polynomial

		Aktual		
		Negatif	Netral	Positif
Prediksi	Negatif	63	1	1
	Netral	22	34	3
	Positif	29	3	33

Dari hasil pengujian menunjukkan bahwa algoritma SVM dengan kernel polynomial dan menggunakan teknik *balancing ROS* memberikan performa terbaik. Berdasarkan *confusion matrix* yang dihasilkan dari pengujian, total data yang dievaluasi sebanyak 189, dengan 63 data menunjukkan negatif, 34 data menunjukkan netral, dan 33 data menunjukkan positif. Dari evaluasi ini, diperoleh akurasi sebesar 68,7%, presisi 77,9%, dan *recall* 68,4%.

KESIMPULAN

Dari hasil pengujian, menunjukkan kecenderungan masyarakat Indonesia yang lebih cenderung memiliki pandangan negatif terhadap kendaraan listrik. Selanjutnya, pengujian tanpa *balancing* menunjukkan bahwa SVM memiliki akurasi 57,5%, presisi 56,3%, dan *recall* 56,3%. Pengujian algoritma NBC menghasilkan akurasi 52,9%, presisi 52,1%, dan *recall* 49,2%, sedangkan KNN menghasilkan akurasi 49%, presisi 47,9%, dan *recall* 47,9%. Namun, dengan penerapan teknik *balancing ROS*, algoritma SVM dapat mencapai performa terbaik dengan akurasi tertinggi sebesar 68,7%, presisi 77,9%, dan *recall* 68,4%. Sementara itu, NBC menghasilkan akurasi 60,3%, presisi 61,3%, dan *recall* 60,3%, sedangkan KNN memiliki akurasi 53,9%, presisi 54%, dan *recall* 53,9%.

Dari penelitian ini, diketahui bahwa algoritma SVM yang tidak menggunakan atau pun menggunakan teknik *balancing*, menunjukkan kinerja yang lebih baik daripada KNN dan NBC. Hal ini dapat dilihat dari nilai akurasi, presisi, dan *recall* yang dihasilkan.

DAFTAR RUJUKAN

Agustian, A., Tukino, T., & Nurapriani, F. (2022). Sentiment analysis, text mining application of sentiment analysis and naive bayes to electric vehicle usage opinions on twitter. *Jurnal Tika*, 7(3), 243–249. <https://doi.org/10.51179/TIKA.V7I3.1550>

- Alfarizi, S., & Fitriani, E. (2023). Analisis sentimen kendaraan listrik menggunakan algoritma naive bayes dengan seleksi fitur information gain dan particle swarm optimization. *Indonesian Journal on Software Engineering (IJSE)*, 9(1), 19–27. <https://ejournal.bsi.ac.id/ejurnal/index.php/ijse/article/view/15671>
- Aryanti, R., Misriati, T., & Hidayat, R. (2023). Klasifikasi risiko kesehatan ibu hamil menggunakan random oversampling untuk mengatasi ketidakseimbangan data. *Klik: Kajian Ilmiah Informatika Dan Komputer*, 3(5), 409–416. <https://djournals.com/klik/article/view/728>
- Badan Pusat Statistik. (2022). *Perkembangan jumlah kendaraan bermotor di indonesia*. Bps.Go.Id. <https://www.bps.go.id/id/statistics-table/2/NTcjMg==/perkembangan-jumlah-kendaraan-bermotor-menurut-jenis-unit.html>
- Christidis, P., & Focas, C. (2019). Factors affecting the uptake of hybrid and electric vehicles in the european union. *Energies*, 12(18), 3414. <https://doi.org/10.3390/EN12183414>
- Efendi, A. (2020). Rancang bangun mobil listrik sula politeknik negeri subang. *Jurnal Pendidikan Teknologi Dan Kejuruan*, 17(1), 75–84. <https://doi.org/10.23887/JPTK-UNDIKSHA.V17I1.23057>
- Erfina, A., & Lestari, R. A. (2023). Sentiment analysis of electric vehicles using the naïve bayes algorithm. *Sistemasi: Jurnal Sistem Informasi*, 12(1), 178–185. <https://doi.org/10.32520/STMSI.V12I1.2417>
- Fitriani, R. D., Yasin, H., & Tarno, T. (2021). Penanganan klasifikasi kelas data tidak seimbang dengan random oversampling pada naive bayes (studi kasus: status peserta kb iud di kabupaten kendal). *Jurnal Gaussian*, 10(1), 11–20. <https://ejournal3.undip.ac.id/index.php/gaussian/article/view/30243>
- Fitrianto, H. (2023). Analisis penggunaan kendaraan listrik sebagai upaya penurunan emisi lingkungan case study kendaraan listrik di provinsi sumatera utara. *Cakrawala Repositori IMWI*, 6(2), 1056–1067. <https://doi.org/10.52851/CAKRAWALA.V6I2.302>
- Ghaniyyu, F. F., & Husnita, N. (2021). Upaya pengendalian perubahan iklim melalui pembatasan kendaraan berbahan bakar minyak di indonesia berdasarkan paris agreement. *Morality: Jurnal Ilmu Hukum*, 7(1), 110–129. <https://doi.org/10.52947/MORALITY.V7I1.196>
- Handayani, A., & Zufria, I. (2023). Analisis sentimen terhadap bakal capres ri 2024 di twitter menggunakan algoritma svm. *Journal of Information System Research (JOSH)*, 5(1), 53–63. <https://doi.org/10.47065/JOSH.V5I1.4379>
- Liu, B. (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool Publishers. <https://www.cs.uic.edu/~liub/FBS/SentimentAnalysis-and-OpinionMining.pdf>
- Muchtari, F. A., Putra, A. M. N., & Bandri, S. (2023). Analisis pengaruh perubahan arus terhadap waktu dan temperatur pengisian baterai kendaraan listrik. *Ensiklopedia of Journal*, 5(3), 115–119. <https://jurnal.ensiklopediaku.org/ojs-2.4.8-3/index.php/ensiklopedia/article/view/1629>
- Mustofa, A., & Novita, R. (2022). Klasifikasi sentimen masyarakat terhadap pemberlakuan pembatasan kegiatan masyarakat menggunakan text mining pada twitter. *Building of Informatics, Technology and Science (BITS)*, 4(1), 200–208–200–208. <https://doi.org/10.47065/BITS.V4I1.1628>
- Noviana, R., & Rasal, I. (2023). Penerapan algoritma naive bayes dan svm untuk analisis sentimen boy band bts pada media sosial twitter. *Jurnal Teknik Dan Science*, 2(2), 51–60. <https://doi.org/10.56127/JTS.V2I2.791>
- Nurmalasari, D., Hermanto, T. I., & Nugroho, I. M. (2023). Perbandingan algoritma svm, knn dan nbc terhadap analisis sentimen aplikasi loan service. *Jurnal Media*

- Informatika Budidarma*, 7(3), 1521–1530. <https://ejurnal.stmik-budidarma.ac.id/index.php/mib/article/view/6427>
- Pambudi, D. I., & Sulastri, S. (2023). Perbandingan naïve bayes dan knn dalam klasifikasi tweet bbm subsidi. *Elkom: Jurnal Elektronika Dan Komputer*, 16(1), 35–44. <https://journal.stekom.ac.id/index.php/elkom/article/view/961>
- Pambudi, I., & Juwono, V. (2023). Electric vehicles in indonesia: public policy, impact, and challenges. *Asian Journal of Social and Humanities*, 2(2), 347–360. <https://doi.org/10.59888/AJOSH.V2I2.173>
- Pratama, Y., Murdiansyah, D. T., & Lhaksana, K. M. (2023). Analisis sentimen kendaraan listrik pada media sosial twitter menggunakan algoritma logistic regression dan principal component analysis. *Jurnal Media Informatika Budidarma*, 7(1), 529–535. <https://ejurnal.stmik-budidarma.ac.id/index.php/mib/article/view/5575>
- Purnajaya, A. R., & Hanggara, F. D. (2021). Perbandingan performa teknik sampling data untuk klasifikasi pasien terinfeksi covid-19 menggunakan rontgen dada. *Journal of Applied Informatics and Computing (JAIC)*, 5(1), 37–42. <https://pdfs.semanticscholar.org/6667/688f78652f9f0e9ff993d07f03cc195bfa16.pdf>
- Rahayu, A. S., Fauzi, A., & Rahmat, R. (2022). Komparasi algoritma naïve bayes dan support vector machine (svm) pada analisis sentimen spotify. *Jurnal Sistem Komputer Dan Informatika (JSON)*, 4(2), 349–354. <https://doi.org/10.30865/JSON.V4I2.5398>
- Saddam, M. A., Dewantara, E. K., & Solichin, A. (2023). Sentiment analysis of flood disaster management in jakarta on twitter using support vector machines. *Sinkron: Jurnal Dan Penelitian Teknik Informatika*, 7(1), 470–479. <https://doi.org/10.33395/SINKRON.V8I1.12063>
- Sidik, A. D., & Ansawarman, A. (2022). Prediksi jumlah kendaraan bermotor menggunakan machine learning. *Formosa Journal of Multidisciplinary Research*, 1(3), 559–568. <https://doi.org/10.55927/FJMR.V1I3.745>
- Subiantoro, H., Elok, A., & Maharani, P. (2023). Analysis of presidential regulation no. 55/2019 related to electric vehicle program in the context of realizing transportation environmentally friendly. *International Journal of Business, Economics and Law*, 30(2). <http://sdgs.bappenas.go.id/tujuan-7/>
- Wang, H., & Wang, Y. (2020). A review of online product reviews. *Journal of Service Science and Management*, 13(1), 88–96. <https://doi.org/10.4236/JSSM.2020.131006>
- Wayan Ernawati, N., Nyoman Satya Kumara, I., Setiawan, W., Raya Kampus Unud Jimbaran, J., Kuta Sel, K., & Badung, K. (2023). Perbandingan metode klasifikasi support vector machine dan naïve bayes pada analisis sentimen kendaraan listrik. *Jurnal SPEKTRUM*, 10(3), 106–114. <https://doi.org/10.24843/SPEKTRUM.2023.V10.I03.P12>