INTI NUSA MANDIRI
DOI: https://doi.org/10.33480/inti.v21i1.8588

VOL. 21. NO. 1 AGUSTUS 2026
P-ISSN: 0216-6933 | E-ISSN: 2685-807X
Diterbitkan Oleh: LPPM Universitas Nusa Mandiri

OPTIMASI HYPERPARAMETER INDOBERT UNTUK ANALISIS SENTIMEN KENDARAAN LISTRIK DENGAN KOMPARASI ALGORITMA MACHINE LEARNING

Elly Indrayuni[1*]; Acmad Nurhadi[2]

Sistem Informasi[1]
Informatika[2]
Universitas Bina Sarana Informatika, Indonesia[1,2]
https://www.bsi.ac.id/[1,2]
elly.eiy@bsi.ac.id[1*], achmad.ahh@bsi.ac.id.[2]

(*) Corresponding Author

Abstract— *The presence of electric vehicles has generated diverse public opinions on social media, creating the need for an automated approach to sentiment identification. Transformer-based models such as IndoBERT can capture semantic context more effectively than conventional machine learning methods that rely on feature representations such as TF-IDF, which are less effective in modeling word relationships, particularly in imbalanced datasets. This study aims to analyze public sentiment toward electric vehicles using an IndoBERT model optimized with Grid Search and compare its performance with Naive Bayes and Support Vector Machine (SVM). An experimental method was applied to a dataset of 1,517 Indonesian-language opinions. IndoBERT was fine-tuned using Grid Search by evaluating hyperparameter combinations of epochs (3, 4, and 5), learning rates (2e-5 and 3e-5), and a batch size of 16. Model performance was evaluated using accuracy, precision, recall, F1-score, and ROC-AUC. The best IndoBERT configuration was obtained with 5 epochs, a learning rate of 3e-5, and a batch size of 16, achieving 73% accuracy. Although its accuracy matched that of SVM, IndoBERT produced more balanced results, with a macro F1-score of 0.62 and the highest average AUC (0.780), outperforming SVM (0.758) and Naive Bayes (0.736). The novelty of this study lies in optimizing IndoBERT using Grid Search and comparing it with Naive Bayes and SVM based on ROC-AUC for Indonesian-language electric vehicle sentiment analysis. The findings demonstrate that Grid Search optimization enhances IndoBERT's contextual understanding, resulting in superior overall performance.*

Keywords*: Electric Vehicles, Hyperparameter Optimization, IndoBERT, Machine Learning, Sentiment Analysis.*

Abstrak— Kehadiran kendaraan listrik telah memunculkan beragam opini masyarakat di media sosial sehingga diperlukan pendekatan otomatis untuk mengidentifikasi sentimen. Model berbasis Transformer seperti IndoBERT mampu memahami konteks semantik lebih baik dibandingkan metode machine learning konvensional yang masih mengandalkan representasi fitur seperti TF-IDF, sehingga kurang optimal dalam memodelkan hubungan antarkata, terutama pada data dengan distribusi kelas yang tidak seimbang. Penelitian ini bertujuan menganalisis sentimen masyarakat terhadap kendaraan listrik menggunakan model IndoBERT yang dioptimalkan dengan Grid Search serta membandingkan kinerjanya dengan Naive Bayes dan Support Vector Machine (SVM). Penelitian menerapkan metode eksperimen menggunakan dataset sebanyak 1.517 opini berbahasa Indonesia. Proses fine-tuning IndoBERT dilakukan melalui Grid Search dengan menguji kombinasi hiperparameter berupa epoch (3, 4, dan 5), learning rate (2e-5 dan 3e-5), serta batch size 16. Evaluasi model dilakukan menggunakan metrik akurasi, presisi, recall, F1-score, dan ROC-AUC. Konfigurasi terbaik IndoBERT diperoleh pada epoch 5, learning rate 3e-5, dan batch size 16 dengan akurasi sebesar 73%. Meskipun nilai akurasinya setara dengan SVM, IndoBERT menghasilkan kinerja yang lebih seimbang dengan macro F1-score sebesar 0,62 dan rata-rata AUC tertinggi sebesar 0,780, melampaui SVM (0,758) dan Naive Bayes (0,736). Kebaruan penelitian ini terletak pada optimasi hiperparameter IndoBERT menggunakan Grid Search serta perbandingannya dengan Naive Bayes dan SVM berdasarkan metrik ROC-AUC pada analisis sentimen kendaraan listrik berbahasa Indonesia. Hasil

penelitian menunjukkan bahwa optimasi Grid Search meningkatkan kemampuan IndoBERT dalam memahami konteks teks sehingga menghasilkan kinerja keseluruhan yang lebih baik.

Kata kunci: *Kendaraan Listrik, Optimasi Hiperparameter, IndoBERT, Machine Learning, Analisis Sentimen.*

PENDAHULUAN

Perkembangan teknologi kendaraan listrik menjadi salah satu inovasi penting dalam sektor transportasi global. Hingga muncul ide berupa inovasi mesin mobil ramah lingkungan yang menjawab kekhawatiran masyarakat ditengah semakin kuatnya kesadaran akan pencemaran udara (Widagdo et al., 2023). Kehadiran kendaraan listrik memicu berbagai reaksi dari masyarakat baik pro dan kontra seperti yang ramai dibicarakan di media sosial *twitter* (Alin et al., 2023). Seiring meningkatnya perhatian masyarakat terhadap kendaraan listrik, berbagai pendapat dan ulasan mulai banyak muncul di platform digital dalam bentuk teks. Medsos atau media sosial adalah sebuah sarana yang sering digunakan untuk memperkenalkan diri, berbagi informasi, berinteraksi atau berkolaborasi satu sama lain (Husen et al., 2023). Media sosial menjadi salah satu bentuk utama dari transformasi ini, telah menjadi wadah terbuka bagi masyarakat untuk menyampaikan pendapat secara langsung dan spontan (Rizkia et al., 2025).

Kenaikan jumlah opini masyarakat membuat proses analisis secara manual semakin kurang efektif. Maka dari itu, dibutuhkan pendekatan otomatis yang mampu mengolah data teks dengan jumlah besar untuk mengetahui kecenderungan opini pengguna. Salah satu metode yang paling populer adalah *sentiment analysis*. Analisis sentimen merupakan metode untuk mendapatkan data dari berbagai platform yang ada di internet (Ningsih et al., 2024). Oleh karena itu, analisis sentimen merupakan salah satu cara untuk menyelesaikan masalah dalam mengelompokkan opini ke dalam kategori opini positif, negatif dan netral (Nurtikasari et al., 2022). Analisis sentimen ini memperbaiki pemahaman organisasi, bisnis, maupun peneliti tentang bagaimana masyarakat merasakan dan merespon berbagai hal yang sedang terjadi (Ernawati & Wati, 2024).

Dalam penelitian analisis sentimen, pendekatan berbasis *machine learning* seperti Naive Bayes dan Support Vector Machine (SVM) sering digunakan karena kemampuannya dalam mengklasifikasikan teks secara efektif. Pendekatan tersebut biasanya menggunakan metode ekstraksi fitur seperti Term *Frequency-Inverse Document Frequency (TF-IDF).* Metode TF-IDF adalah pendekatan yang digunakan untuk menentukan tingkat relevansi suatu kata *(term)* dalam sebuah dokumen (Sujadi et al., 2022). Teknik ini digunakan untuk mengidentifikasi fitur penting dalam teks (Azzahra & Mailoa, 2025).

Perkembangan *deep learning* berbasis arsitektur transformer memberikan pendekatan baru dalam bidang NLP. BERT merupakan model bahasa kontekstual yang memiliki kinerja lebih unggul dalam berbagai tugas NLP. Model BERT digunakan untuk memahami konteks serta hubungan antara kata dalam teks secara lebih dalam, yang telah terbukti sangat efektif dalam bidang pemrosesan bahasa alami atau dikenal dengan NLP (A. Bello & Leung, 2023). IndoBERT dilatih menggunakan teks dari berbagai sumber dalam bahasa Indonesia, termasuk Wikipedia dan teks dari situs web berita (Mustofa et al., 2025). IndoBERT memiliki kemampuan yang lebih baik dalam memahami konteks bahasa Indonesia dibandingkan model-model sebelumnya, sehingga diharapkan mampu memberikan hasil yang lebih akurat dan relevan (Simarmata & Sasongko, 2025).

Beberapa penelitian terdahulu telah menerapkan metode NLP untuk analisis sentimen maupun perbandingan model klasifikasi teks. Penelitian (Wijaya et al., 2024) tentang analisis sentimen penduduk Indonesia terhadap kendaraan listrik menggunakan metode *deep learning* berbasis FastText dan BERT. Temuan dari penelitian tersebut mengindikasikan bahwa pendekatan berbasis Transformer memiliki kinerja yang baik dalam menganalisis opini mengenai kendaraan listrik. Namun, penelitian tersebut belum menggunakan model bahasa Indonesia khusus seperti IndoBERT. Penelitian lain pernah dilakukan oleh (Salsabila et al., 2025) yaitu membandingkan kinerja algoritma *Naive Bayes, SVM,* dan *Bidirectional Encoder Representations* from Transformers (BERT) untuk klasifikasi sentimen pada aplikasi Yummy. Penelitian tersebut menggunakan 6.773 data ulasan dari Google Play Store dengan tahapan *preprocessing*, penerapan ekstraksi fitur TF-IDF pada algoritma Naive Bayes dan SVM, serta *fine-tuning Bidirectional Encoder Representations* from Transformers (BERT). Hasil penelitian menunjukkan bahwa SVM menghasilkan kinerja terbaik dengan memperoleh nilai akurasi sebesar 94%, nilai precision 95%, nilai recall 98%, dan F1-score sebesar 96% dibandingkan dengan algoritma Naive Bayes dan BERT yang memperoleh akurasi sebesar 90%. Hasil ini menunjukkan bahwa

VOL. 21. NO. 1 AGUSTUS 2026
P-ISSN: 0216-6933 | E-ISSN: 2685-807X
Diterbitkan Oleh: LPPM Universitas Nusa Mandiri

algoritma konvensional dapat memberikan performa tinggi pada dataset tertentu, namun model berbasis Transformer memiliki kelebihan dalam memahami konteks bahasa. Penelitian selanjutnya dilakukan (Nuryadi et al., 2025) yaitu penerapan metode *fine-tuning* IndoBERT untuk analisis sentimen ulasan pengguna aplikasi Tiket.com melalui Google Play Store menggunakan dataset SmSA IndoNLU. Hasil studi ini menunjukkan bahwa nilai akurasi model IndoBERT mampu mencapai 91%, dengan F1-score masing-masing pada kelas positif sebesar 93%, pada kelas negatif sebesar 94% dan pada kelas netral sebesar 77%. Namun nilai recall yang dihasilkan pada kelas netral masih relatif rendah, yaitu 65%. Hal ini menunjukkan bahwa klasifikasi sentimen netral masih menjadi tantangan dalam analisis sentimen berbahasa Indonesia.

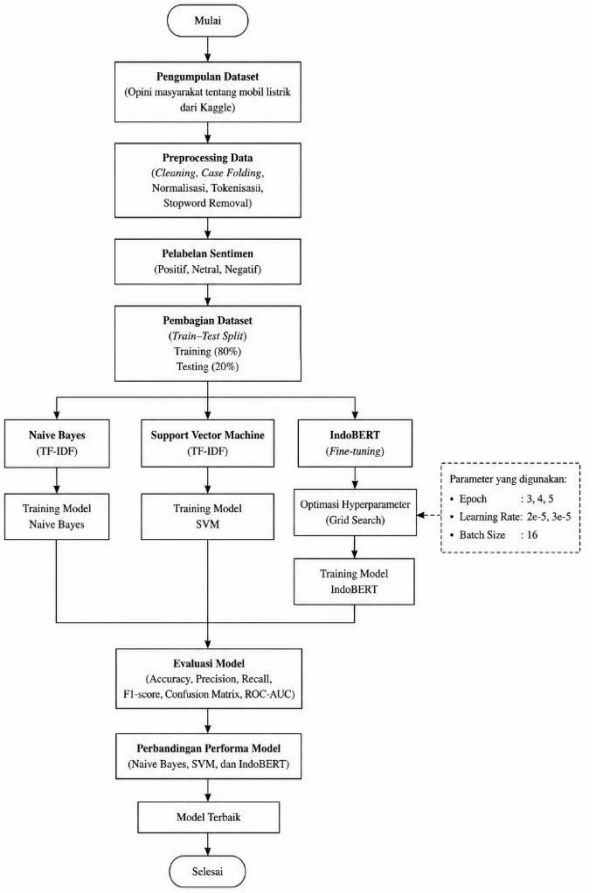
Berdasarkan penelitian terdahulu yang sudah dilakukan, IndoBERT terbukti memiliki kinerja yang lebih baik dalam analisis sentimen untuk teks berbahasa Indonesia. Namun banyak penelitian masih menggunakan konfigurasi *hyperparameter* yang telah ditentukan sebelumnya tanpa melakukan proses optimasi secara sistematis. Padahal, konfigurasi *hyperparameter* merupakan salah satu faktor yang dapat memengaruhi performa model *fine-tuning*. *Hyperparameter* diatur secara manual sesuai dengan kebutuhan peneliti untuk membantu memperkirakan parameter model (Haris et al., 2024). Oleh karena itu, penelitian ini mengusulkan optimasi *hyperparameter* melalui *Grid Search* untuk memperoleh konfigurasi IndoBERT yang lebih optimal serta membandingkan performanya dengan algoritma machine learning seperti Naive Bayes dan Support Vector Machine.

Kebaruan penelitian ini terletak pada penerapan optimasi *hyperparameter* menggunakan metode *Grid Search* pada model IndoBERT untuk analisis sentimen opini masyarakat mengenai kendaraan listrik berbahasa Indonesia. Berbeda dengan penelitian sebelumnya yang menggunakan konfigurasi hyperparameter tetap, penelitian ini mengevaluasi beberapa kombinasi nilai *epoch, learning rate,* dan *batch size* untuk mendapatkan konfigurasi model yang optimal. Selain itu, penelitian ini menganalisis performa IndoBERT dibandingkan dengan algoritma Naive Bayes dan Support Vector Machine (SVM) dengan menggunakan evaluasi yang lebih mendalam melalui metrik *accuracy, precision, recall, F1-score, confusion matrix*, dan *ROC-AUC.* Hal ini bertujuan untuk memberikan gambaran yang lebih menyeluruh terhadap kemampuan setiap model algoritma dalam mengklasifikasikan sentimen.

Berdasarkan uraian tersebut, tujuan dilakukannya penelitian ini untuk menganalisis pengaruh teknik optimasi hyperparameter menggunakan Grid Search terhadap performa model IndoBERT dalam analisis sentimen kendaraan listrik serta membandingkan hasilnya dengan algoritma Naive Bayes dan SVM. Hasil penelitian diharapkan dapat memberikan rekomendasi konfigurasi hyperparameter yang optimal dan menjadi sumber referensi dalam pengembangan model analisis sentimen berbahasa Indonesia.

BAHAN DAN METODE

Studi ini menerapkan teknik *text mining* berbasis *supervised machine learning* untuk menganalisis sentimen masyarakat terkait opini mengenai kendaraan listrik. Tahapan penelitian terdiri atas pengumpulan dataset, preprocessing data, pelabelan sentimen, pembagian dataset, penerapan model algoritma untuk klasifikasi sentimen menggunakan Naive Bayes, Support Vector Machine dan IndoBERT, optimasi hyperparameter pada model IndoBERT menggunakan Grid Search, evaluasi model, serta perbandingan performa ketiga algoritma. Tahapan penelitian dapat ditunjukkan pada Gambar 1.



Sumber: (Hasil Penelitian, 2026)

Gambar 1. Tahapan Penelitian

VOL. 21. NO. 1 AGUSTUS 2026
P-ISSN: 0216-6933 | E-ISSN: 2685-807X
Diterbitkan Oleh: LPPM Universitas Nusa Mandiri

INTI NUSA MANDIRI
DOI: https://doi.org/10.33480/inti.v21i1.8588

Langkah pertama yang dilakukan dalam penelitian ini adalah mengumpulkan dataset. Dataset yang digunakan berupa opini masyarakat mengenai kendaraan listrik yang diperoleh dari platform Kaggle. Dataset dipilih karena berisi kumpulan komentar berbahasa Indonesia yang merepresentasikan berbagai persepsi masyarakat terhadap penggunaan kendaraan listrik. Sebelum digunakan pada proses klasifikasi, data diperiksa untuk menghilangkan data duplikat maupun data yang tidak lengkap sehingga diperoleh dataset yang siap diproses pada tahap berikutnya.

Data teks yang telah terkumpul kemudian melalui tahap *preprocessing* guna meningkatkan kualitas data sebelum diklasifikasikan. Proses pertama diawali dengan *cleaning* untuk membersihkan simbol, angka, URL, mention, hashtag, emoji, dan tanda baca. Selanjutnya, *case folding* diterapkan untuk menyeragamkan seluruh karakter menjadi huruf kecil guna mencegah duplikasi representasi kata. Pada tahap *normalisasi*, kata-kata tidak baku maupun singkatan dikonversi ke dalam bentuk standar menggunakan kamus bahasa Indonesia (seperti mengolah kata 'gk' menjadi 'tidak' dan 'bgt' menjadi 'banget'). Setelah itu dilakukan *tokenisasi*, yaitu mengurai kalimat menjadi entitas kata *(token)*. Rangkaian ini diakhiri dengan *stopword removal*, yaitu mengeliminasi kata-kata yang bersifat umum yang memiliki kontribusi rendah terhadap proses klasifikasi sentimen seperti *dan*, *yang*, *di*, dan *ke.*

Tahap selanjutnya melibatkan pelabelan sentimen pada data yang telah diproses ke dalam tiga kelas utama: positif, netral, dan negatif. Label ini digunakan sebagai target pada proses pembelajaran terawasi (*supervised learning*). Dataset tersebut kemudian dibagi secara proporsional sebesar 80% sebagai data pelatihan dan 20% sebagai data pengujian melalui pendekatan *train-test split*. Skema pembagian ini diterapkan agar model dapat mempelajari pola data secara optimal sekaligus menguji kemampuan generalisasinya terhadap data yang belum pernah diuji sebelumnya.

Tahap klasifikasi dijalankan dengan membandingkan tiga algoritma, yaitu Naive Bayes, Support Vector Machine (SVM), dan IndoBERT. Pada algoritma Naive Bayes dan SVM data teks terlebih dahulu diubah menjadi representasi numerik memakai cara TF-IDF. TF-IDF menyajikan bobot terhadap setiap kata berdasarkan frekuensi kemunculannya pada suatu dokumen dan tingkat kemunculannya pada seluruh dokumen. Nilai Term Frequency (TF) dihitung menggunakan Persamaan (1).

$$TF(t,d) = \frac{f(t,d)}{\sum f(t,d)} \quad (1)$$

Dengan $f(t,d)$ mengungkapkan jumlah kemunculan kata t pada dokumen d. Selanjutnya nilai *Inverse Document Frequency (IDF)* dihitung memakai Persamaan (2).

$$IDF(t) = \log\left(\frac{N}{df(t)}\right) \quad (2)$$

Dimana N merupakan jumlah seluruh dokumen dan $df(t)$ menunjukkan jumlah dokumen yang mengandung kata t. Bobot akhir TF-IDF didapatkan menggunakan Persamaan (3).

$$TFIDF(t,d) = TF(t,d) \times IDF(t) \quad (3)$$

Model Naive Bayes menjalankan tahapan klasifikasi berdasarkan Teorema Bayes dengan mengasumsikan bahwa setiap fitur bersifat independen. Probabilitas suatu dokumen termasuk ke dalam kelas tertentu dapat dihitung menggunakan Persamaan (4).

$$P(C \mid X) = \frac{P(X|C)P(C)}{P(X)} \quad (4)$$

Karena nilai $P(X)$sama untuk setiap kelas, maka tahapan klasifikasi dijalankan dengan memilih kelas yang mempunyai probabilitas posterior terbesar sebagaimana ditunjukkan pada Persamaan (5).

$$\hat{C} = \arg\max P(C) \prod_{i=1}^{n} P(x_i \mid C) \quad (5)$$

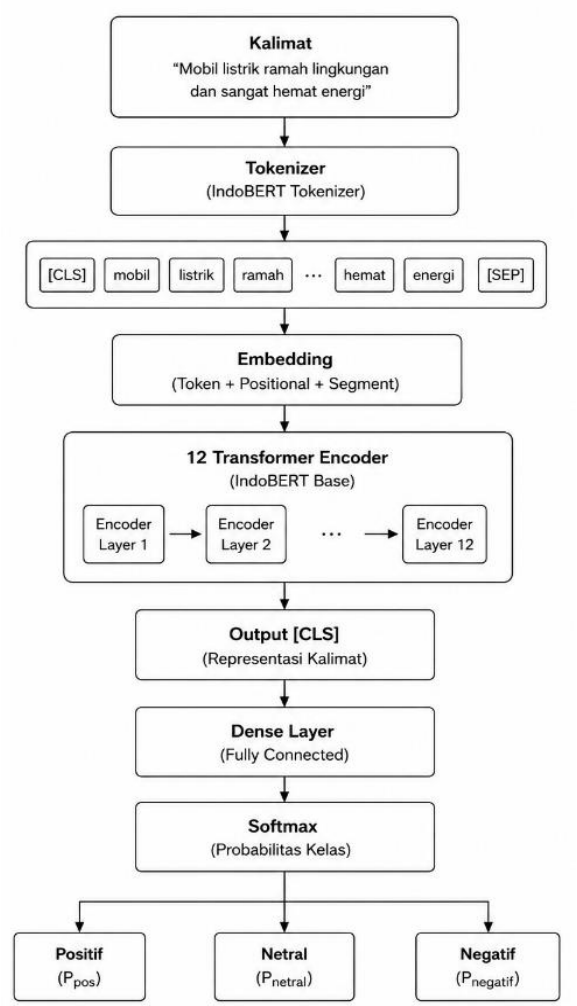
Sementara itu, *Support Vector Machine (SVM)* bekerja dengan mencari hyperplane terbaik yang mampu memisahkan kelas sentimen dengan margin terbesar. Fungsi hyperplane ditunjukkan pada Persamaan (6).

$$w \cdot x + b = 0 \quad (6)$$

Pada studi ini akan digunakan kernel linear karena sesuai untuk data teks berdimensi tinggi temuan ekstraksi TF-IDF. Berbeda dengan dua algoritma sebelumnya sedangkan IndoBERT dilakukan dengan pendekatan *fine-tuning*. Arsitektur Fine-tuning IndoBERT tampak jelas pada Gambar 2.

INTI NUSA MANDIRI
DOI: https://doi.org/10.33480/inti.v21i1.8588

VOL. 21. NO. 1 AGUSTUS 2026
P-ISSN: 0216-6933 | E-ISSN: 2685-807X
Diterbitkan Oleh: LPPM Universitas Nusa Mandiri

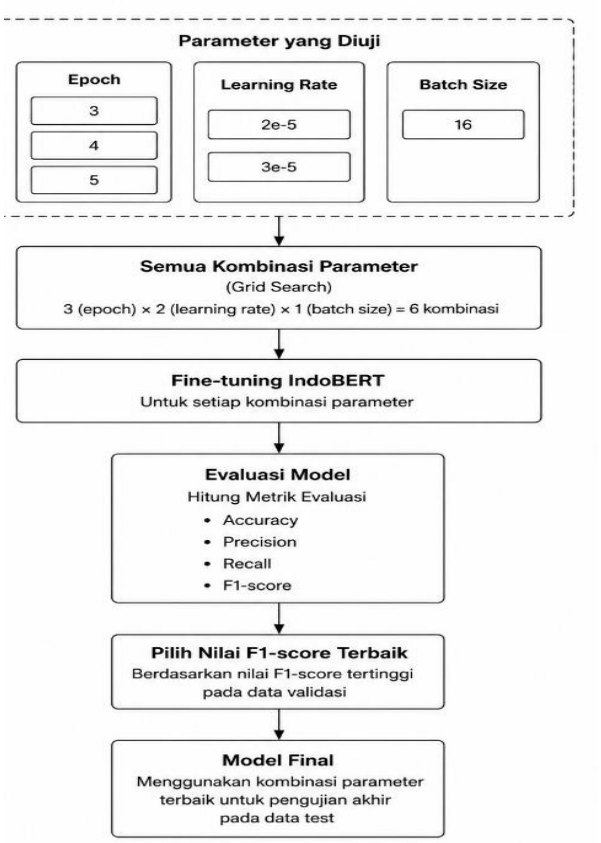


Sumber: (Hasil Penelitian, 2026)

Gambar 2. Arsitektur *Fine-tuning* IndoBERT

Pada tahapan *fine-tuning*, seluruh parameter model IndoBERT diperbarui menggunakan data pelatihan sehingga model dapat menyesuaikan representasi bahasa dengan tugas klasifikasi sentimen kendaraan listrik.

Pada model IndoBERT untuk memperoleh konfigurasi model terbaik, dilakukan optimasi hyperparameter seperti yang ditunjukkan pada Gambar 3. menggunakan *Grid Search* dengan parameter *epoch* (3, 4, 5), *learning rate* (2e-5, 3e-5), dan *batch size* 16. Dengan konfigurasi tersebut dihasilkan enam kombinasi hyperparameter yang masing-masing digunakan untuk melakukan proses fine-tuning secara penuh. Setiap kombinasi hyperparameter yang diterapkan kemudian dievaluasi menggunakan data validasi dengan menghitung nilai *Accuracy, Precision, Recall,* dan *F1-score*. Kombinasi yang menghasilkan nilai *F1-score* tertinggi dipilih sebagai konfigurasi terbaik dan selanjutnya digunakan pada proses pengujian menggunakan data testing. Pendekatan Grid Search dipilih karena mampu mengevaluasi seluruh kombinasi parameter secara sistematis sehingga menghasilkan konfigurasi model yang optimal.



Sumber: (Hasil Penelitian, 2026)

Gambar 3. Proses Optimasi Hyperparameter Menggunakan Grid Search

Evaluasi model dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, *confusion matrix*, dan ROC-AUC. Hasil evaluasi dari ketiga model dibandingkan untuk menentukan algoritma terbaik dalam klasifikasi sentimen opini masyarakat terhadap kendaraan listrik.

HASIL DAN PEMBAHASAN

Tahapan pertama dalam pelaksanaan penelitian adalah pengumpulan data dan analisis data yang bertujuan untuk mendapatkan informasi yang sesuai dengan kebutuhan penelitian. Pada penelitian ini, data yang digunakan didapatkan dari platform Kaggle melalui situs www.kaggle.com. Setelah dilakukan tahap preprocessing data, *data cleaning* yang dihasilkan sebanyak 1.517 yang kemudian dilakukan pelabelan ke dalam kategori ulasan positif, negatif atau netral. Hasil dari proses pemberian label secara manual menunjukkan bahwa data yang dikumpulkan terdiri dari 869

VOL. 21. NO. 1 AGUSTUS 2026
P-ISSN: 0216-6933 | E-ISSN: 2685-807X
Diterbitkan Oleh: LPPM Universitas Nusa Mandiri

INTI NUSA MANDIRI
DOI: https://doi.org/10.33480/inti.v21i1.8588

sentimen negatif, 504 sentimen positif dan 144 sentimen netral. Proses pelabelan sentimen atau opini masyarakat terhadap kendaraan listrik dapat ditunjukkan pada Tabel 1.

Tabel 1. Proses Pelabelan Sentimen

No	Review	Sentimen
0	problem subsidi kualitas diturunin harga dinaikin usaha gitu cari cuan subsidi sebab inflasi paling gede	Negatif
1	dukung terus doa negara makmur produksi sendirikan	Positif
2	iklan kudu kencang	Netral

Sumber: (Hasil Penelitian, 2026)

Visualisasi terhadap opini masyarakat dijalankan menggunakan *wordcloud* untuk memperlihatkan kata-kata yang paling sering muncul dalam setiap kategori. Untuk *wordcloud* sentimen positif dapat tampak jelas pada Gambar 4. Pada sentimen positif, kata yang memiliki frekuensi kemunculan tinggi meliputi *"harga"*, *"mau"*, *"mahal"*, *"lebih"*, *"laku"*, *"beli"*, *"subsidi"*, dan *"banyak"*.

Sumber: (Hasil Penelitian, 2026)
Gambar 4. *Wordcloud* Sentimen Positif

Pada sentimen netral kata yang paling sering muncul tampak jelas pada Gambar 5. yaitu *"subsidi"*, *"kendara"*, *"beli"*, *"orang"*, *"mau"*, *"laku"*, *"banyak"*, *"harga"*, serta *"motor"*.

Sumber: (Hasil Penelitian, 2026)
Gambar 5. *Wordcloud* Sentimen Netral

Sementara itu, *wordcloud* pada sentimen negatif terlihat pada Gambar 6. Berdasarkan Gambar 6. kata yang paling dominan adalah *"beli"*, *"kendara"*, *"harga"*, *"subsidi"*, *"banyak"*, *"rakyat"*, *"lebih"*, *"orang"*, dan *"mahal"*. Hasil visualisasi untuk setiap kategori sentimen memperlihatkan bahwa setiap kategori sentimen memiliki pola kata utama yang berbeda, sehingga dapat menyajikan gambaran mengenai topik atau isu yang paling banyak dibahas oleh masyarakat.

Sumber: (Hasil Penelitian, 2026)
Gambar 6. *Wordcloud* Sentimen Negatif

Setelah tahapan *preprocessing* dan pelabelan data selesai dijalankan, dataset kemudian dibagi menjadi data *training* dan data *testing* yang dipakai dalam proses pelatihan serta pengujian model *machine learning*. Pembagian data dilakukan dengan perbandingan 80% untuk data *training* dan 20% untuk data *testing*. Berdasarkan pembagian tersebut, sebanyak 1.213 data digunakan untuk data *training*, dan 304 data lainnya digunakan untuk data *testing.*

Untuk melakukan analisis sentimen pada penelitian ini akan digunakan dua model *machine learning* konvensional, yakni *Naïve Bayes* dan *Support Vector Machine (SVM)*, serta satu model berbasis Transformer, yaitu IndoBERT. Pada model algoritma *Naïve Bayes dan SVM*, pengaturan parameter awal menggunakan konfigurasi bawaan (*default*). Sementara itu, IndoBERT memiliki arsitektur yang lebih kompleks sehingga kinerjanya sangat dipengaruhi oleh perubahan *hyperparameter* yang diterapkan. Dalam proses pelatihan IndoBERT, digunakan beberapa parameter seperti *learning rate* untuk mengatur proses pembaruan bobot model, *batch size* untuk menentukan jumlah data yang diproses pada setiap pembaruan, serta *epoch* yang merupakan jumlah iterasi pelatihan model.

INTI NUSA MANDIRI
DOI: https://doi.org/10.33480/inti.v21i1.8588

VOL. 21. NO. 1 AGUSTUS 2026
P-ISSN: 0216-6933 | E-ISSN: 2685-807X
Diterbitkan Oleh: LPPM Universitas Nusa Mandiri

Tabel 2. *Hyperparameter* Model NB, SVM dan IndoBERT

Model	Hyperparameter	Nilai
Naive Bayes	Ngram_range	(1,2)
SVM	Kernel	linear
	probability	true
	random_state	42
IndoBERT	Epoch	3,4,5
	learning rate	2e-5, 3e-5
	batch size	16

Sumber: (Hasil Penelitian, 2026)

Pengujian model Naive Bayes dilakukan untuk mengukur performa generalisasi. Performa klasifikasi menggunakan model Naïve Bayes tertera pada Tabel 3.

Tabel 3. Performa Klasifikasi Naive Bayes

Sentimen	Precision	Recall	F1-score
Negatif	0.61	0.99	0.76
Netral	0.00	0.00	0.00
Positif	0.91	0.20	0.33
Accuracy			0.63

Sumber: (Hasil Penelitian, 2026)

Secara umum, pada Tabel 3. menunjukkan hasil klasifikasi Naive Bayes ini menunjukkan performa yang kurang prima. Model bisa mencapai akurasi 63% hanya karena ia sangat pintar menebak kelas negatif (yang jumlahnya mendominasi, yaitu 174 data), sementara model gagal total di kelas netral dan sangat lemah di kelas positif. Untuk pengujian yang telah dilakukan terhadap model *Support Vector Machine* (SVM) menggunakan total data uji sebanyak 304 sampel, didapatkan penilaian performa klasifikasi yang ditampilkan pada Tabel 4.

Tabel 4. Performa Klasifikasi *Support Vector Machine* (SVM)

Sentimen	Precision	Recall	F1-score
Negatif	0.72	0.91	0.81
Netral	0.67	0.07	0.12
Positif	0.77	0.61	0.68
Accuracy			0.73

Sumber: (Hasil Penelitian, 2026)

Tabel 4. menampilkan hasil performa model SVM yang cukup solid dengan nilai akurasi mencapai 73% (0,73). Angka ini mengindikasikan bahwa model SVM mampu mengklasifikasikan sebagian besar data uji ke dalam kategori sentimen yang tepat. Pendekatan pembobotan (*weighted average*) juga memperlihatkan konsistensi performa yang seimbang dengan nilai *precision*, *recall*, dan *f1-score* yang ada diantara 0,70 hingga 0,73. Untuk mengevaluasi kemampuan pendekatan berbasis *Deep Learning*, model transformer IndoBERT diuji menggunakan 304 sampel data yang sama. Pengujian model IndoBERT dilakukan menggunakan data pengujian dengan konfigurasi hyperparameter yang telah dirancang sebelumnya, yaitu Epoch sebanyak 3, 4, dan 5; Learning Rate sebesar 2e-5 dan 3e-5; serta Batch Size sebesar 16. Hasil pengujian hyperparameter dari IndoBERT ini ditunjukkan pada Tabel 5.

Tabel 5. Hasil Pengujian *Hyperparameter* IndoBERT

Epoch	Learning Rate	Batch Size	Accuracy	F1-score
3	2e-5	16	0.68	0.66
	3e-5		0.70	0.69
4	2e-5	16	0.68	0.66
	3e-5		0.72	0.71
5	2e-5	16	0.70	0.69
	3e-5		0.73	0.72

Sumber: (Hasil Penelitian, 2026)

Berdasarkan hasil pengujian hyperparameter pada Tabel 5. IndoBERT memperoleh akurasi rata-rata 0.70 dengan 0.73 untuk akurasi tertinggi dan 0.68 untuk akurasi terendah. Akurasi tertinggi yang diperoleh dengan konfigurasi *hyperparameter Epoch 5, Learning Rate* 3e-5, dan *Batch Size* 16. Sementara akurasi terendah diperoleh dengan konfigurasi *hyperparameter Epoch 3, Learning Rate* 2e-5, dan *Batch Size* 16. Berdasarkan Tabel 5, model IndoBERT mencatatkan nilai akurasi tertinggi sebesar 73% (0,73) dengan nilai *weighted average F1-score* mencapai 0,72. Untuk detail performa klasifikasi IndoBERT berdasarkan nilai akurasi tertinggi yang diperoleh dari hasil pengujian *Hyperparameter* IndoBERT dapat dilihat pada Tabel 6.

Tabel 6. Performa Klasifikasi IndoBERT

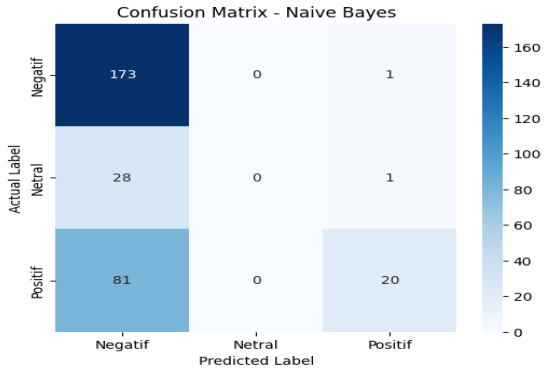
Sentimen	precision	Recall	F1-score
Negatif	0.74	0.90	0.81
Netral	0.73	0.28	0.40
Positif	0.72	0.58	0.64
Macro avg	*0.73*	*0.59*	*0.62*
Weighted average	*0.73*	*0.73*	*0.72*
Accuracy			0.73

Sumber: (Hasil Penelitian, 2026)

Pada Tabel 6, hal yang paling krusial untuk diperhatikan adalah peningkatan signifikan pada nilai *macro average F1-score* yang menyentuh angka 0,62. Kenaikan metrik makro ini mengindikasikan bahwa IndoBERT mempunyai kemampuan yang lebih seimbang dalam mengenali setiap kelas sentimen, meskipun bekerja di bawah kondisi data yang sangat tidak seimbang (*imbalanced data*). Secara keseluruhan, karakteristik distribusi nilai *precision* yang sangat merata dan stabil di angka 0,72–0,74 untuk seluruh kelas menunjukkan bahwa

VOL. 21. NO. 1 AGUSTUS 2026
P-ISSN: 0216-6933 | E-ISSN: 2685-807X
Diterbitkan Oleh: LPPM Universitas Nusa Mandiri
INTI NUSA MANDIRI
DOI: https://doi.org/10.33480/inti.v21i1.8588

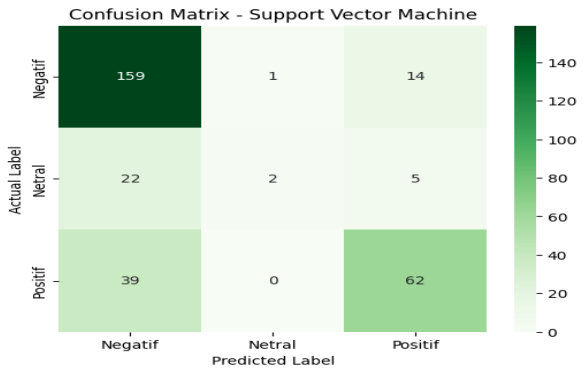
IndoBERT memiliki tingkat kepastian prediksi (*reliability*) yang jauh lebih konsisten dibandingkan model pendekatan tradisional. Setelah dijalankan pengujian model, penilaian evaluasi dilakukan dengan menggunakan confusion matrix. Untuk confusion matrix model algoritma Naive Bayes tampak jelas pada Gambar 7.



Sumber: (Hasil Penelitian, 2026)

Gambar 7. Confusion Matrix Naive Bayes

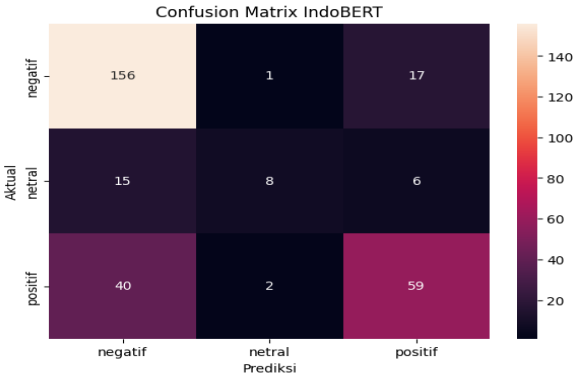
Berdasarkan Gambar 7. Confusion matrix Naive Bayes menunjukkan bahwa model mampu mengklasifikasikan kelas negatif dengan baik dengan 173 prediksi benar dari 174 data yang diberikan. Namun, hasil performa pada kelas netral dan positif masih kurang baik karena sebagian besar data positif dan netral salah diklasifikasikan sebagai negatif. Metode TF-IDF yang digunakan pada Naive Bayes masih kurang mampu memahami konteks dari opini, terutama pada kalimat yang memiliki makna yang tidak jelas. Evaluasi berikutnya untuk confusion matrix model algoritma *Support Vector Machine* (SVM) ditunjukkan pada Gambar 8.



Sumber: (Hasil Penelitian, 2026)

Gambar 8. Confusion Matrix *Support Vector Machine* (SVM)

Confusion matrix algoritma *Support Vector Machine* (SVM) yang ditunjukkan pada Gambar 8. menunjukkan bahwa model mampu mengklasifikasikan sentimen negatif dengan baik yaitu sebanyak 159 dari 174 data berhasil diprediksi dengan benar. Selain itu, SVM menunjukkan peningkatan kemampuan dalam mengenali sentimen positif dengan 62 prediksi benar dibandingkan Naive Bayes. Namun, kelas netral masih memiliki performa rendah karena sebagian besar data netral diklasifikasikan sebagai negatif. Dengan demikian meskipun SVM mampu bekerja baik pada data teks berbasis TF-IDF, metode ini masih memiliki keterbatasan dalam memahami konteks opini yang memiliki makna yang tidak jelas. Untuk evaluasi model yang terakhir yaitu confusion matrix dengan IndoBERT yang ditunjukkan pada Gambar 9.

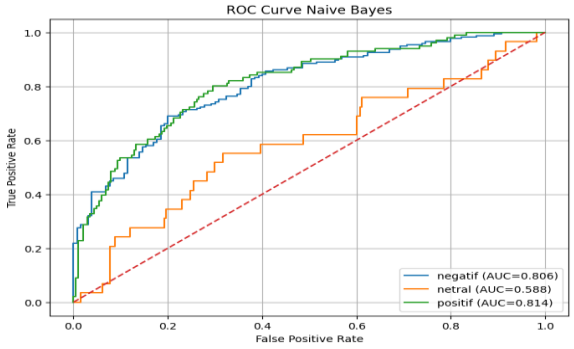


Sumber: (Hasil Penelitian, 2026)

Gambar 9. Confusion Matrix IndoBERT

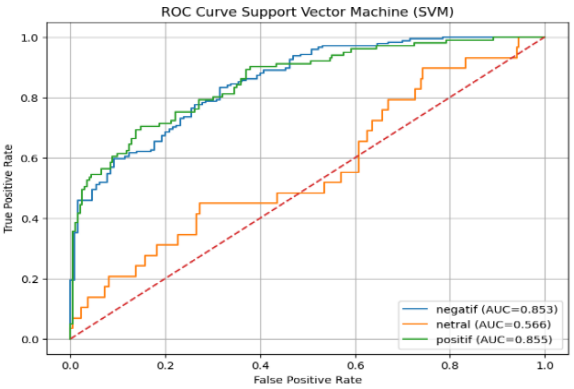
Confusion matrix model IndoBERT yang ditunjukkan pada Gambar 9. menunjukkan bahwa model IndoBERT mampu mengklasifikasikan sentimen negatif dengan baik yaitu sebanyak 156 data berhasil diprediksi dengan benar. Pada kelas positif, IndoBERT mampu mengenali 59 data dengan benar, sehingga IndoBERT memiliki performa yang lebih baik dibandingkan dengan Naive Bayes. Sementara itu, kelas netral masih memiliki tingkat kesalahan yang lebih besar dibandingkan kelas positif dan negatif karena karakteristik opini netral cenderung memiliki pola bahasa yang lebih samar. Secara umum, IndoBERT memberikan hasil klasifikasi yang lebih seimbang karena mampu memahami konteks bahasa dengan baik dibandingkan metode machine learning yang menggunakan TF-IDF. Selanjutnya dilakukan evaluasi menggunakan kurva *Receiver Operating Characteristic (ROC*) untuk tiga algoritma klasifikasi, yaitu *Naive Bayes, Support Vector Machine (SVM),* dan IndoBERT dalam menganalisis sentimen masyarakat terhadap opini mengenai kendaraan listrik.

INTI NUSA MANDIRI
DOI: https://doi.org/10.33480/inti.v21i1.8588
VOL. 21. NO. 1 AGUSTUS 2026
P-ISSN: 0216-6933 | E-ISSN: 2685-807X
Diterbitkan Oleh: LPPM Universitas Nusa Mandiri



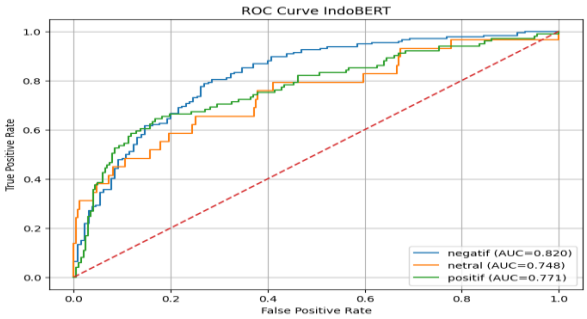
Sumber: (Hasil Penelitian, 2026)
Gambar 10. ROC Curve *Naive Bayes*

Berdasarkan kurva ROC pada Gambar 10, algoritma Naive Bayes menunjukkan performa klasifikasi yang optimal pada kelas positif dan negatif dengan capaian nilai AUC melampaui 0,80. Kendati demikian, efektivitas pemodelan pada kelas netral mengalami penurunan signifikan dengan nilai AUC sebesar 0,588. Keterbatasan ini berakar pada *conditional independence assumption* yang menjadi karakteristik dasar Naive Bayes, di mana hubungan kontekstual antar kata diabaikan. Pada analisis sentimen, kalimat netral umumnya tidak mengandung kata-kata bermuatan emosi (*sentiment-bearing words*) yang dominan dan sering memiliki kosakata yang tumpang tindih dengan kalimat positif maupun negatif. Kondisi tersebut menyebabkan distribusi probabilitas antar kelas menjadi saling berdekatan sehingga memperbesar margin kesalahan klasifikasi pada kelas netral. Selain itu, ketidakseimbangan jumlah data antar kelas (*imbalanced data*) juga dapat menyebabkan model lebih banyak mempelajari karakteristik kelas mayoritas dibandingkan kelas netral. Walaupun demikian, rata-rata AUC secara keseluruhan yang mencapai 0,736 mengindikasikan bahwa model masih berada pada kategori *fair-good classification* yang dapat diandalkan.



Sumber: (Hasil Penelitian, 2026)
Gambar 11. ROC Curve *Support Vector Machine (SVM)*

ROC Curve untuk model selanjutnya yaitu algoritma *Support Vector Machine (SVM)* dapat ditunjukkan pada Gambar 11. Pada Gambar 11. ROC Curve algoritma SVM menunjukkan nilai AUC terbaik untuk kelas positif sebesar 0.855 dan kelas negatif sebesar 0.853. Ini menunjukkan bahwa SVM sangat mampu dalam mengidentifikasi sentimen positif dan negatif. Namun untuk kelas netral masih menjadi tantangan dengan nilai AUC hanya sebesar 0.566. Rendahnya performa pada kelas netral bisa disebabkan oleh karakteristik teks netral yang memiliki kemiripan fitur dengan kelas positif dan negatif sehingga sulit dipisahkan secara tegas. Rata-rata AUC model algoritma SVM sebesar 0.758 menunjukkan bahwa SVM memiliki performa secara keseluruhan lebih baik dibandingkan Naive Bayes. Berbeda dengan Naive Bayes dan SVM, IndoBERT menunjukkan performa yang lebih seimbang pada ketiga kelas (positif, negatif, dan netral). ROC Curve untuk model IndoBERT dapat ditunjukkan pada Gambar 12.



Sumber: (Hasil Penelitian, 2026)
Gambar 12. ROC Curve IndoBERT

Pada Gambar 12. ROC Curve model IndoBERT menunjukkan bahwa representasi kontekstual yang dimiliki IndoBERT mampu menangkap makna kalimat dengan lebih baik sehingga dapat membedakan sentimen netral secara lebih akurat. Rata-rata nilai AUC model IndoBERT sebesar 0.780 merupakan nilai AUC tertinggi diantara ketiga algoritma yang diuji.

KESIMPULAN

Penelitian ini menyimpulkan bahwa model IndoBERT dengan konfigurasi *hyperparameter* optimal pada *Epoch* 5, *Learning Rate* 3e-5, dan *Batch Size* 16 menghasilkan performa terbaik dalam menganalisis sentimen masyarakat terhadap kendaraan listrik dengan nilai akurasi mencapai 73%. Meskipun nilai akurasi ini setara dengan algoritma SVM, IndoBERT terbukti lebih unggul dan seimbang dalam mengklasifikasikan data yang tidak seimbang (*imbalanced data*), yang ditunjukkan

dengan nilai *macro average F1-score* sebesar 0,62. Selain itu, berdasarkan evaluasi kurva ROC, IndoBERT memperoleh rata-rata nilai AUC tertinggi sebesar 0,780, yang mengindikasikan kemampuan model dalam membedakan kelas sentimen secara lebih menyeluruh dan memahami konteks bahasa ambigu terutama pada sentimen netral jauh lebih baik dibandingkan Naive Bayes dengan nilai AUC seebsar 0,736 dan SVM yang menghasilkan nilai AUC sebesar 0,758. Untuk pengembangan penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar dan melakukan penyeimbangan jumlah data antar kelas sentimen untuk mengatasi rendahnya deteksi pada kelas netral. Selain itu, dapat dieksplorasi penggabungan arsitektur IndoBERT dengan model *ensemble* atau penambahan lapisan *convolutional* yang merujuk pada CNN *(Convolutional Neural Network)* untuk meningkatkan ke dalaman ekstraksi fitur kontekstual.

REFERENSI

A. Bello, S. C. N., & Leung, M. F. (2023). A BERT Framework to Sentiment Analysis of Tweets. *Sensors*, *23*(1). https://doi.org/10.3390/s23010506

Alin, F. N., Totohendarto, M. H., & Muttaqin, M. R. (2023). Analisis Sentimen Terhadap Kendaraan Listrik Pada Platform Twitter Menggunakan Metode Naive Bayes. *INFORMATICS FOR EDUCATORS AND PROFESSIONAL : Journal of Informatics*, *8*(2), 96–107. https://doi.org/https://doi.org/10.51211/itbi.v8i2.2490

Azzahra, W. L., & Mailoa, E. (2025). Analisis Sentimen terhadap RSUD Salatiga Menggunakan SVM dan TF-IDF. *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, *6*(1), 478–489. https://doi.org/10.35870/jimik.v6i1.1208

Ernawati, S., & Wati, R. (2024). Evaluasi Performa Kernel SVM dalam Analisis Sentimen Review Aplikasi ChatGPT Menggunakan Hyperparameter dan VADER Lexicon. *Jurnal Buana Informatika*, *15*(01), 40–49. https://doi.org/10.24002/jbi.v15i1.7925

Haris, M., Suharso, A., & Nurkifli, E. H. (2024). Analisis Sentimen Pada Game Efootball Di Google Play Store Menggunakan Algoritma Indobert. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(6), 12108–12121. https://doi.org/10.36040/jati.v8i6.11810

Husen, R. A., Astuti, R., Marlia, L., Rahmaddeni, R., & Efrizoni, L. (2023). Analisis Sentimen Opini Publik pada Twitter Terhadap Bank BSI Menggunakan Algoritma Machine Learning. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *3*(2), 211–218. https://doi.org/10.57152/malcom.v3i2.901

Mustofa, R. L., Tarno, & Widodo, C. E. W. (2025). Analisis Sentimen Berbasis Aspek Pada Aplikasi Elektronik Survei Kepuasan Masyarakat (E-Skm) Jawa Tengah Menggunakan Indobert Aspect-Based Sentiment Analysis on the Central Java Public Satisfaction Survey Electronic Application (E-Skm) Using Indobert. *JTIIK(Jurnal Teknologi Informasi Dan Ilmu Komputer)*, *12*(4), 867–874. https://doi.org/https://doi.org/10.25126/jtiik.124

Ningsih, W., Alfianda, B., Rahmaddeni, R., & Wulandari, D. (2024). Perbandingan Algoritma SVM dan Naïve Bayes dalam Analisis Sentimen Twitter pada Penggunaan Mobil Listrik di Indonesia. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *4*(2), 556–562. https://doi.org/10.57152/malcom.v4i2.1253

Nurtikasari, Y., Syariful Alam, & Teguh Iman Hermanto. (2022). Analisis Sentimen Opini Masyarakat Terhadap Film Pada Platform Twitter Menggunakan Algoritma Naive Bayes. *INSOLOGI: Jurnal Sains Dan Teknologi*, *1*(4), 411–423. https://doi.org/10.55123/insologi.v1i4.770

Nuryadi, D., Metandi, F., Alam Hadiwijaya, N., Zainul Rohman, M., Hartanto, S., Syafrizal, A., & Yadie, E. (2025). Fine Tuning Indobert Untuk Analisis Sentimen Pada Ulasan Pengguna Aplikasi Tiket.Com Di Google Play Store. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *9*(2), 3577–3583. https://doi.org/10.36040/jati.v9i2.13204

Rizkia, A. S., Wufron, & Roji, F. F. (2025). Sentiment Analysis of Coretax: A Comparison of Manual, Transformers- Based, and Lexicon-Based Data Labeling on IndoBERT Performance. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *5*(3), 1037–1048. https://doi.org/https://doi.org/10.57152/malcom.v5i3.2151

Salsabila, S. A., Priyatna, B., Hananto, A., & Tukino. (2025). Komparasi Kinerja Model Naive Bayes, SVM, dan BERT dalam Klasifikasi Sentimen Ulasan Pada Aplikasi YUMMY. *STORAGE: Jurnal Ilmiah Teknik Dan Ilmu Komputer*, *4*(2), 42–47.

VOL. 21. NO. 1 AGUSTUS 2026
P-ISSN: 0216-6933 | E-ISSN: 2685-807X
Diterbitkan Oleh: LPPM Universitas Nusa Mandiri

https://doi.org/10.55123/storage.v4i2.5120

Simarmata, A. A. P., & Sasongko, T. B. (2025). Sentiment Analysis on BRImo Application Reviews Using IndoBERT. *Journal of Applied Informatics and Computing*, *9*(3), 851–862. https://doi.org/10.30871/jaic.v9i3.8162

Sujadi, H., Fajar, S., & Roni, C. (2022). Analisis Sentimen Pengguna Media Sosial Twitter Terhadap Wabah Covid-19 Dengan Metode Naive Bayes Classifier Dan Support Vector Machine. *INFOTECH Journal*, *8*(1), 22–27. https://doi.org/10.31949/infotech.v8i1.1883

Widagdo, A. S., Ardiansyah, Qodri, K. N., Saputra, F. E. N., & Putri, N. A. R. (2023). Analisis Sentimen Mobil Listrik di Indonesia Menggunakan Long-Short Term Memory (LSTM). *Jurnal Fasilkom*, *13*(3), 416–423. https://doi.org/https://doi.org/10.37859/jf.v13i3.6303

Wijaya, D. R., Sasmitha, G. M. A., & Vihikan, W. O. (2024). Sentiment Analysis of Indonesian Citizens on Electric Vehicle Using FastText and BERT Method. *Journal of Information Systems and Informatics*, *6*(3), 1360–1372. https://doi.org/10.51519/journalisi.v6i3.784