



Analisis Sentimen Masyarakat terhadap Kebijakan Penggunaan BBM Campuran Etanol di X (Twitter) Menggunakan Transformers (IndoBERT)

Arraffy Abbyu Arrasyid¹, Andreas Pradana²

raffy.rasyid@gmail.com¹, andreas.perdana@dharmawacana.ac.id²

¹Universitas Dharma Wacana, Indonesia

²Universitas Dharma Wacana, Indonesia

*Korespondensi: ✉ raffy.rasyid@gmail.com

Abstrak

The transition toward sustainable energy has become a strategic priority for Indonesia, particularly through the implementation of bioethanol-blended fuel policies. However, public perception toward this policy remains diverse and dynamic, especially on social media platforms such as X. This study aims to analyze public sentiment regarding the implementation of bioethanol-blended fuel policies and to compare the performance of Support Vector Machine (SVM) and IndoBERT in sentiment classification.

This research adopts a quantitative experimental approach using Natural Language Processing (NLP). A total of 2,501 tweets were collected through web crawling and processed using a dual pipeline preprocessing approach. Sentiment labeling was conducted using the VADER method with manual validation. SVM was used as the baseline model, while IndoBERT was used as the Transformer-based model. Additional experiments using class weighting and Synthetic Minority Oversampling Technique (SMOTE) were conducted to evaluate the effect of class imbalance on model performance. Model performance was evaluated using accuracy, precision, recall, F1-score, macro-F1, weighted-F1, and confusion matrix.

The results show that IndoBERT outperforms SVM, achieving an accuracy of 82.83% and a weighted F1-score of 82.72%, compared to SVM with an accuracy of 72.46% and a weighted F1-score of 72.02%. SMOTE improved SVM performance to 76.85% accuracy and 77.07% weighted F1-score, but remained lower than IndoBERT. The sentiment distribution indicates that public opinion is predominantly positive, although technical, economic, and infrastructure concerns remain.

Status Artikel:

Diterima: 15-05-2026

Direvisi: 29-05-2026

Diterima: 02-06-2026

Kata Kunci:

Natural language processing;

Support vector machine;

Grid search;

Dual pipeline;

VADER lexicon.



© 2026 Arraffy Abbyu Arrasyid, Andreas Pradana

This work is licensed under a

[Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

PENDAHULUAN

Ketahanan energi dan perubahan iklim merupakan isu strategis global yang menuntut transformasi sistem energi menuju sumber yang lebih berkelanjutan. Indonesia, sebagai negara

berkembang dengan tingkat konsumsi energi yang terus meningkat, masih menghadapi ketergantungan signifikan terhadap bahan bakar fosil impor (Herindrasti et al., 2024). Kondisi ini tidak hanya menekan neraca perdagangan, tetapi juga menjadi hambatan dalam pencapaian target Net Zero Emission (NZE) tahun 2060 (Zahira & Fadillah, 2022). Oleh karena itu, pemerintah mendorong transisi energi melalui pemanfaatan Energi Baru Terbarukan (EBT), salah satunya melalui kebijakan penggunaan bioetanol sebagai campuran bahan bakar minyak (BBM) sebagaimana diatur dalam Peraturan Presiden Nomor 40 Tahun 2023.

Implementasi kebijakan tersebut diwujudkan melalui peluncuran Pertamina Green 95 (E5) serta rencana peningkatan kadar campuran menjadi E10. Meskipun kebijakan ini diarahkan untuk mendukung energi bersih dan mengurangi ketergantungan impor BBM, respons masyarakat menunjukkan dinamika yang kompleks (Chaerul et al., 2025). Di satu sisi, kebijakan ini dipandang sebagai langkah progresif menuju keberlanjutan energi. Namun di sisi lain, muncul kekhawatiran terkait dampak teknis terhadap kendaraan, potensi peningkatan biaya operasional, serta isu food versus fuel akibat pemanfaatan bahan baku pangan untuk energi (Adit, 2025). Kompleksitas respons ini menunjukkan pentingnya analisis berbasis data untuk memahami persepsi publik secara objektif.

Media sosial, khususnya platform X (Twitter), telah menjadi ruang diskursif utama bagi masyarakat dalam menyampaikan opini terhadap kebijakan publik. Karakteristiknya yang real-time, terbuka, dan berbasis teks menjadikannya sumber data yang kaya untuk analisis sentimen (Laksono & Rahmadani, 2026). Namun demikian, teks pada media sosial memiliki sifat tidak terstruktur dan mengandung berbagai bentuk bahasa non-standar seperti slang, singkatan, serta ekspresi informal lainnya, sehingga menimbulkan tantangan signifikan dalam menangkap makna kontekstual secara akurat.

Penelitian sebelumnya menunjukkan bahwa analisis sentimen berbasis machine learning telah banyak digunakan untuk memahami opini publik pada media sosial. Studi oleh (Obie & M. Abu Jihad Plaza R., 2024) menunjukkan bahwa algoritma Support Vector Machine (SVM) mampu menghasilkan performa klasifikasi yang lebih baik dibandingkan Naive Bayes dengan akurasi mencapai 68,40%. Temuan ini menunjukkan bahwa pendekatan machine learning efektif dalam menangkap pola opini publik berbasis data teks, khususnya pada isu sosial yang berkembang di media digital. Selain itu, penelitian lain juga menunjukkan bahwa metode klasifikasi tradisional masih memiliki keterbatasan dalam menangkap hubungan semantik antar kata karena menggunakan pendekatan berbasis bag-of-words yang mengabaikan urutan kata (Darwis et al., 2020).

Seiring perkembangan Natural Language Processing (NLP), model berbasis Transformer seperti Bidirectional Encoder Representations from Transformers (BERT) yang diperkenalkan oleh (Devlin et al., 2019) menunjukkan kemampuan superior dalam memahami konteks bahasa melalui mekanisme self-attention. Pendekatan ini memungkinkan representasi teks yang bersifat kontekstual dan bidirectional, sehingga mampu menangkap makna kata berdasarkan konteks kalimat secara lebih akurat. Dalam konteks bahasa Indonesia, IndoBERT yang dikembangkan oleh (Koto et al., 2020; Santika et al., 2025) telah dilatih menggunakan korpus skala besar sehingga mampu merepresentasikan hubungan semantik secara lebih baik dibandingkan metode konvensional. Selain itu, pengembangan model lanjutan seperti IndoBERTweet (Koto et al., 2021) menunjukkan peningkatan performa dalam menangani teks media sosial yang bersifat informal. Namun demikian, penelitian ini menggunakan IndoBERT

sebagai model utama untuk menjaga konsistensi arsitektur model serta efisiensi komputasi dalam proses pelatihan dan evaluasi (Christian et al., 2025).

Perkembangan terbaru dalam Natural Language Processing mengindikasikan bahwa model berbasis Transformer dan Large Language Models (LLM) semakin mendominasi tugas analisis sentimen karena kemampuannya dalam memahami konteks bahasa yang kompleks dan representasi semantik yang lebih kaya dalam tugas analisis sentimen karena kemampuannya dalam memahami konteks bahasa yang kompleks dan representasi semantik yang lebih kaya (Tan et al., 2023; Aljabar et al., 2024; Amrullah, 2025). Studi-studi terkini juga menunjukkan bahwa pendekatan berbasis *pre-trained language models* mampu memberikan performa yang lebih stabil dan akurat dibandingkan metode tradisional, terutama pada data teks tidak terstruktur seperti media sosial (Sasikumar et al., 2026).

Berdasarkan uraian tersebut, terdapat kesenjangan penelitian (research gap) berupa keterbatasan studi yang secara spesifik mengkaji efektivitas model Transformer dalam menganalisis sentimen publik terhadap kebijakan energi berbasis bioetanol pada data media sosial berbahasa Indonesia. Sebagian besar penelitian terdahulu masih berfokus pada isu sosial umum atau hanya melakukan komparasi algoritma klasifikasi sentimen tanpa mempertimbangkan konteks kebijakan publik yang spesifik dan dinamis (A. D. Latief et al., 2023). Selain itu, pendekatan preprocessing yang digunakan pada penelitian sebelumnya umumnya bersifat seragam untuk seluruh model, sehingga belum mempertimbangkan kebutuhan representasi data yang berbeda antara metode machine learning tradisional dan model berbasis Transformer.

Berbeda dengan penelitian sebelumnya yang umumnya berfokus pada perbandingan performa algoritma klasifikasi sentimen secara generik, studi ini menekankan pendekatan *domain-specific policy analysis* dengan mengkaji persepsi publik terhadap kebijakan energi berbasis bioetanol di Indonesia. Pendekatan ini memungkinkan analisis yang lebih kontekstual karena mempertimbangkan dinamika opini masyarakat terhadap isu kebijakan publik yang spesifik dan aktual. Selain itu, penelitian ini secara khusus berfokus pada data media sosial berbahasa Indonesia dari platform X, yang memiliki karakteristik linguistik unik seperti penggunaan bahasa informal, slang, serta struktur kalimat tidak baku (*Indonesian social media context*), yang menjadi tantangan signifikan dalam analisis sentimen (Hidayatullah et al., 2024; Supriyadi & Makatita, 2025).

Lebih lanjut, penelitian ini mengadopsi pendekatan *dual pipeline preprocessing* sebagai strategi untuk menyesuaikan tahapan prapemrosesan dengan karakteristik masing-masing model tanpa mengubah arsitektur algoritma yang digunakan (Br Sembiring et al., 2026). Pada pendekatan ini, preprocessing untuk SVM dilakukan secara lebih agresif melalui stemming dan stopword removal untuk mengoptimalkan representasi fitur berbasis TF-IDF, sedangkan pada IndoBERT dilakukan preprocessing minimal guna mempertahankan konteks linguistik dan struktur semantik kalimat. Pendekatan adaptif ini masih relatif terbatas dalam literatur analisis sentimen berbahasa Indonesia, khususnya pada konteks kebijakan energi berbasis media sosial.

Dengan demikian, penelitian ini tidak hanya berkontribusi pada evaluasi performa model klasifikasi sentimen, tetapi juga menawarkan pendekatan analisis yang lebih kontekstual terhadap respons publik pada kebijakan energi nasional berbasis data media sosial. Selain itu, strategi dual pipeline preprocessing yang diterapkan dalam penelitian ini menunjukkan bahwa penyesuaian tahapan prapemrosesan terhadap karakteristik model dapat menjadi faktor penting

dalam meningkatkan kualitas representasi teks dan performa klasifikasi sentimen pada data tidak terstruktur berbahasa Indonesia.

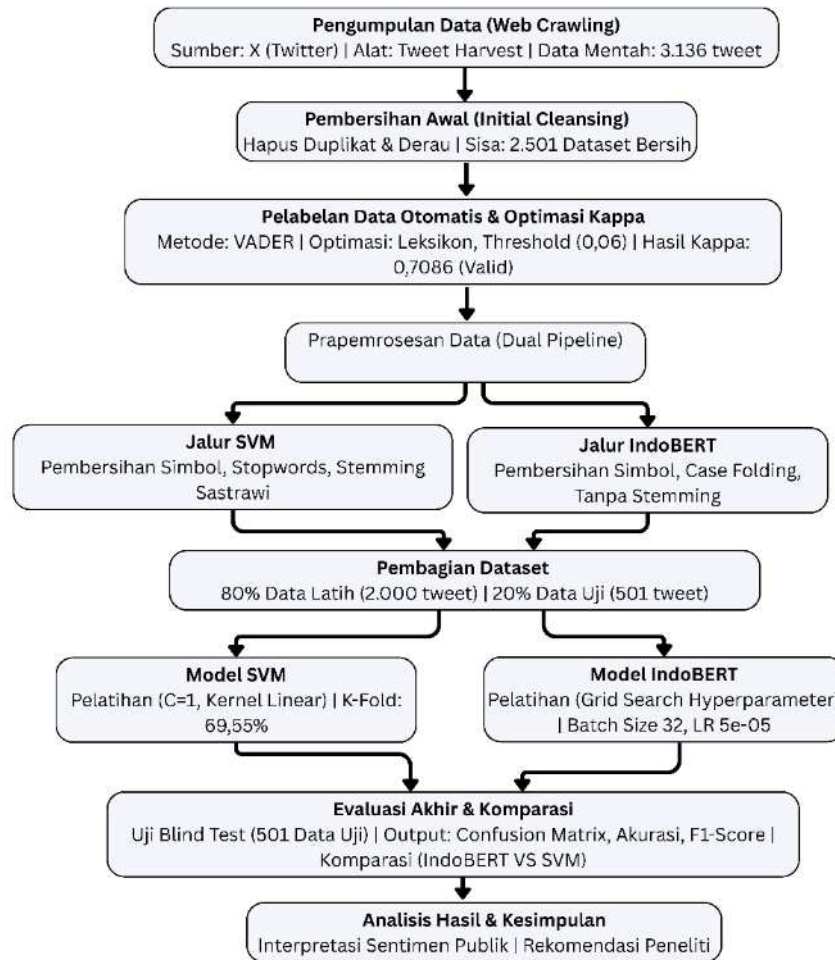
Penelitian ini menjadi relevan karena analisis sentimen berbasis media sosial dapat digunakan sebagai instrumen evaluasi awal terhadap penerimaan kebijakan publik secara real-time dan berbasis data.

Berdasarkan kesenjangan dan pendekatan yang diusulkan tersebut, penelitian ini bertujuan untuk: (1) menganalisis sentimen masyarakat terhadap kebijakan penggunaan BBM campuran etanol di platform X (Twitter), serta (2) membandingkan performa model IndoBERT dengan SVM dalam klasifikasi sentimen.

Kontribusi utama penelitian ini terletak pada dua aspek. Pertama, memberikan pemetaan empiris mengenai persepsi publik terhadap kebijakan energi berbasis bioetanol sebagai bagian dari transisi energi nasional. Kedua, mengevaluasi secara komparatif efektivitas model berbasis Transformer dibandingkan metode machine learning tradisional dalam konteks data teks tidak terstruktur berbahasa Indonesia. Dengan demikian, penelitian ini diharapkan tidak hanya memberikan kontribusi akademik dalam pengembangan metode analisis sentimen, tetapi juga menjadi referensi bagi pengambil kebijakan dalam memahami respons masyarakat secara lebih berbasis data.

METHODS

Penelitian ini merupakan penelitian kuantitatif eksperimental yang menerapkan pendekatan NLP untuk melakukan analisis sentimen terhadap data teks media sosial. Penelitian ini bertujuan untuk membandingkan performa model berbasis Transformer (IndoBERT) dengan metode machine learning tradisional (Support Vector Machine) dalam mengklasifikasikan sentimen publik terhadap kebijakan bahan bakar campuran etanol (E10) (Firdaus et al., n.d.; Sipayung et al., n.d.; Y. Santoso & S. Candra, 2025). Alur penelitian disusun secara sistematis yang meliputi tahapan pengumpulan data, filter Data, pelabelan data, pembagian dataset, pelatihan model, serta evaluasi performa untuk memastikan validitas eksperimen dan kemudahan replikasi (N. A. Haqimi et al., 2025). Alur tahapan penelitian secara keseluruhan diilustrasikan Pada Gambar 1.



Gambar 1 Diagram Alur Penelitian

Pengumpulan data dilakukan melalui teknik *web crawling* (Kahlon & Singh, 2026) pada platform X (Twitter) menggunakan tools *Tweet Harvest* yang terintegrasi dengan X API v2 (Ferdinan et al., n.d.). Proses *crawling* dilakukan dengan menggunakan sejumlah kata kunci yang relevan dengan kebijakan bahan bakar campuran etanol, yaitu “Pertamax Green”, “BBM etanol”, “E10 Indonesia”, “bioetanol”, dan “BBM campuran etanol”. Data dikumpulkan dalam rentang waktu 1 Januari 2024 hingga 25 Januari 2026 untuk menangkap dinamika opini publik pada fase awal implementasi kebijakan. Parameter data yang diambil meliputi teks tweet dan waktu publikasi (*timestamp*) tanpa menyertakan informasi identitas pribadi pengguna. Proses pengambilan data dibatasi pada bahasa Indonesia dan tweet publik, serta tidak mencakup retweet duplikatif. Data dikumpulkan sesuai dengan kebijakan platform X (Twitter) dan hanya menggunakan data publik tanpa menyertakan identitas personal pengguna. Seluruh proses pengambilan dan pengolahan data dilakukan dengan mempertimbangkan prinsip etika penelitian serta perlindungan privasi pengguna. Seluruh proses ini dilakukan dengan mempertimbangkan aspek etika penelitian dan kepatuhan terhadap kebijakan penggunaan data publik pada platform X (Twitter). Dari proses tersebut diperoleh sebanyak 3.136 tweet mentah, yang kemudian melalui tahap pembersihan data dari duplikasi dan noise sehingga diperoleh 2.501 tweet sebagai dataset utama.

Pelabelan sentimen dilakukan menggunakan pendekatan *weak labeling* dengan metode VADER (*Valence Aware Dictionary and Sentiment Reasoner*) (Satria & Husaini, n.d.). Mengingat VADER dikembangkan untuk bahasa Inggris, dilakukan proses *lexicon adaptation* untuk menyesuaikan dengan karakteristik bahasa Indonesia non-standar pada media social (A. Parashar et al., 2026). Adaptasi dilakukan dengan menambahkan sekitar ± 120 kosakata baru yang mencakup slang, singkatan, dan ekspresi informal seperti “gak”, “nggak”, “mantap”, “jelek”, “mahal”, dan “worth it”, yang masing-masing diberikan skor sentimen berdasarkan konteks penggunaannya. Selain itu, dilakukan penyesuaian terhadap beberapa skor pada leksikon asli untuk meningkatkan sensitivitas terhadap konteks lokal. Proses *threshold tuning* juga dilakukan untuk mengoptimalkan batas klasifikasi sentimen, di mana ambang batas ditetapkan menjadi $\geq 0,03$ untuk sentimen positif, $\leq -0,03$ untuk sentimen negatif, dan nilai di antaranya dikategorikan sebagai netral. Untuk memastikan kualitas pelabelan, proses anotasi dilakukan oleh peneliti sebagai annotator utama yang memiliki pemahaman terhadap analisis sentimen dan konteks bahasa Indonesia pada media sosial (A. D. Latief et al., 2023). Untuk meminimalkan subjektivitas, dilakukan validasi melalui anotasi pembandingan oleh seorang ahli (dosen) yang memiliki kompetensi di bidang terkait. Sebelum proses anotasi, disusun pedoman anotasi (*annotation guideline*) yang berisi definisi masing-masing kelas sentimen (positif, negatif, netral) beserta contoh kasus untuk menyamakan persepsi.

Tingkat kesesuaian antara hasil anotasi peneliti dan annotator pembandingan diukur menggunakan Cohen’s Kappa yang dirumuskan sebagai berikut:

$$k = \frac{P_0 - P_e}{1 - P_e}$$

Di mana k mewakili nilai *Cohen’s Kappa*, P_0 adalah proporsi kesepakatan aktual, dan P_e adalah probabilitas kesepakatan acak (A. D. Latief et al., 2023). Nilai $k > 0,70$ dikategorikan sebagai tingkat kesepakatan yang kuat (Ferdinan et al., n.d.). Hasil evaluasi awal menunjukkan nilai 0,5334, sehingga dilakukan proses optimasi melalui *lexicon adaptation*, *threshold tuning*, dan *ground truth refinement* hingga diperoleh nilai 0,7086 yang termasuk kategori kesepakatan kuat. Dengan demikian, dataset yang digunakan dalam penelitian ini dinilai memiliki tingkat reliabilitas yang memadai untuk digunakan dalam proses pelatihan model.

Pada tahap prapemrosesan, penelitian ini menerapkan pendekatan *Dual Pipeline Preprocessing* untuk mengakomodasi perbedaan karakteristik algoritma. Pada pipeline SVM, dilakukan pembersihan teks secara intensif meliputi case folding, penghapusan URL, mention, hashtag, simbol, tokenisasi, stopword removal, serta stemming menggunakan pustaka Sastrawi untuk menyederhanakan representasi fitur berbasis bag-of-words (N. Hasanati et al., 2022). Sementara itu, pada pipeline IndoBERT, prapemrosesan dilakukan secara minimal dengan hanya melakukan case folding dan penghapusan simbol non-alfabet guna mempertahankan struktur kalimat dan konteks semantik (Geni et al., 2023; Budianoor et al., 2026). Tahapan preprocessing memiliki peran penting dalam meningkatkan kualitas data dan performa model, karena data yang tidak terstruktur dapat menurunkan akurasi klasifikasi (Prasetya et al., 2025)

Dataset kemudian dibagi menggunakan metode *stratified random sampling* dengan proporsi 80% sebagai data latih dan 20% sebagai data uji untuk menjaga distribusi kelas tetap seimbang (R. N. Satrya et al., 2022; Wibowo et al., 2025). Untuk memastikan konsistensi hasil

eksperimen, seluruh proses yang melibatkan komponen acak dilakukan dengan menetapkan random seed sebesar 42. Pada tahap ekstraksi fitur, model SVM menggunakan metode TF-IDF dengan batas maksimum 5.000 fitur (Riansyahputra et al., n.d.; Zhao & Che, 2026), sedangkan model IndoBERT menggunakan *pre-trained tokenizer* dengan panjang maksimum input sebesar 128 token (Leelawat et al., 2022; Zahra et al., 2026).

Pipeline SVM menghasilkan representasi fitur berbentuk matriks TF-IDF dengan maksimum 5.000 fitur. Dengan jumlah data latih sebanyak 2.000 data, representasi data latih pada pipeline SVM berbentuk matriks sparse berukuran maksimum 2.000×5.000 . Proses pelatihan SVM relatif lebih ringan karena hanya melibatkan transformasi TF-IDF dan optimasi model berbasis machine learning tradisional.

Pipeline IndoBERT memiliki proses komputasi yang lebih kompleks karena setiap teks diubah menjadi token menggunakan tokenizer IndoBERT dengan panjang maksimum 128 token. Dengan 2.000 data latih, proses fine-tuning melibatkan pembaruan parameter model Transformer melalui beberapa epoch. Pada eksperimen IndoBERT dengan class weighting, pelatihan berhenti pada epoch ke-8 karena mekanisme early stopping, dengan waktu pelatihan sekitar 217 detik pada GPU Google Colab. Perbedaan ini menunjukkan bahwa pipeline IndoBERT membutuhkan sumber daya komputasi lebih besar dibandingkan pipeline SVM, tetapi memiliki kemampuan representasi konteks yang lebih kuat.

Pelatihan model utama dilakukan menggunakan dua pendekatan, yaitu SVM sebagai model baseline dan IndoBERT sebagai model berbasis Transformer. Pada pipeline SVM, teks direpresentasikan menggunakan TF-IDF dengan maksimum 5.000 fitur, kemudian diklasifikasikan menggunakan SVM. Pada pipeline IndoBERT, teks diproses menggunakan tokenizer IndoBERT dengan panjang maksimum 128 token dan dilatih melalui proses fine-tuning.

Selain eksperimen utama, penelitian ini menambahkan eksperimen tambahan untuk mengevaluasi pengaruh ketidakseimbangan kelas terhadap performa model. Pada eksperimen tambahan pipeline SVM, digunakan dua strategi penanganan imbalance, yaitu class weighting dan SMOTE. Class weighting diterapkan melalui parameter `class_weight='balanced'`, sedangkan SMOTE diterapkan pada representasi fitur TF-IDF dan hanya dilakukan pada data latih untuk menghindari data leakage. Pada pipeline IndoBERT, eksperimen tambahan dilakukan menggunakan class weighting pada fungsi loss berdasarkan distribusi kelas pada data latih.

Eksperimen tambahan ini digunakan sebagai analisis pendukung untuk menjawab potensi bias akibat distribusi kelas yang tidak seimbang, sehingga tidak menggantikan hasil utama penelitian.

Pada pipeline IndoBERT, penelitian ini menguji IndoBERT baseline dan IndoBERT dengan class weighting. Model IndoBERT dilatih melalui proses fine-tuning menggunakan optimizer AdamW, panjang maksimum input 128 token, batch size 32, learning rate $5e-05$, dan maksimum 50 epoch. Early stopping diterapkan dengan nilai patience sebesar 3 untuk mencegah overfitting. Pada skenario IndoBERT dengan class weighting, bobot kelas diterapkan pada fungsi Cross Entropy Loss berdasarkan distribusi kelas pada data latih.

Eksperimen dilakukan dengan random seed 42 untuk menjaga konsistensi hasil. Evaluasi dilakukan pada data uji yang sama untuk seluruh skenario agar perbandingan antar model tetap adil.

Evaluasi performa model dilakukan menggunakan confusion matrix yang terdiri dari True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN). Metrik evaluasi digunakan untuk mengukur kinerja model secara komprehensif dalam klasifikasi sentimen, yang dirumuskan sebagai berikut:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Mengingat karakteristik dataset yang cenderung tidak seimbang, metrik F1-score digunakan sebagai indikator utama dalam mengevaluasi performa model karena mampu memberikan keseimbangan antara precision dan recall.

Penanganan Ketidakseimbangan Data

Distribusi kelas pada dataset menunjukkan adanya ketidakseimbangan, dengan jumlah data positif sebanyak 1.319, netral sebanyak 654, dan negatif sebanyak 528. Setelah pembagian data menggunakan stratified random sampling, data latih terdiri atas 1.055 data positif, 523 data netral, dan 422 data negatif, sedangkan data uji terdiri atas 264 data positif, 131 data netral, dan 106 data negatif. Pembagian ini mempertahankan proporsi kelas pada data latih dan data uji.

Untuk mengevaluasi dampak ketidakseimbangan kelas terhadap performa model, penelitian ini menambahkan skenario penanganan imbalance. Pada pipeline SVM, digunakan dua pendekatan, yaitu class weighting dan Synthetic Minority Oversampling Technique (SMOTE). Class weighting diterapkan melalui parameter `class_weight='balanced'`, sedangkan SMOTE diterapkan pada representasi fitur TF-IDF dan hanya dilakukan pada data latih untuk menghindari data leakage.

Pada pipeline IndoBERT, penanganan imbalance dilakukan menggunakan class weighting pada fungsi loss. Bobot kelas dihitung berdasarkan distribusi data latih, yaitu 1,5798 untuk kelas negatif, 1,2747 untuk kelas netral, dan 0,6319 untuk kelas positif. Bobot yang lebih besar diberikan kepada kelas minoritas agar kesalahan prediksi pada kelas negatif dan netral memperoleh penalti yang lebih besar selama proses pelatihan. SMOTE tidak diterapkan secara langsung pada pipeline IndoBERT karena input IndoBERT berbentuk urutan token, sehingga oversampling berbasis vektor tidak secara langsung merepresentasikan struktur semantik teks.

Untuk memastikan model tidak bias terhadap kelas positif sebagai kelas mayoritas, evaluasi tidak hanya dilakukan menggunakan accuracy, tetapi juga macro-F1, weighted-F1, recall per kelas, dan confusion matrix.

RESULTS AND DISCUSSION

Hasil Pengumpulan dan Prapemrosesan Data

Tahap pengumpulan data menghasilkan 3.136 tweet mentah dari platform X (Twitter) yang diperoleh melalui proses crawling. Data tersebut mengandung berbagai bentuk noise seperti URL, mention, hashtag, dan emotikon. Setelah dilakukan pembersihan dan penghapusan duplikasi, diperoleh sebanyak 2.501 tweet yang digunakan sebagai dataset utama penelitian. Tabel 1 mendemonstrasikan perbandingan transformasi teks mentah menjadi format yang siap diekstraksi oleh SVM dan IndoBERT.

Table 1 Perbandingan Sampel Teks pada Dual Pipeline Preprocessing

Teks Mentah	Pipeline SVM	Pipeline IndoBERT
Pertamina melakukan penyesuaian harga pada BBM nonsubsidi kecuali Pertamina. #bisnisupdate #update #bisnis #onliner https://t.co/pBg5Sa9lL https://t.co/18jyoJUPIm	pertamina laku sesuai harga bbm nonsubsidi pertamax bisnisupdate update bisnis onliner	Pertamina melakukan penyesuaian harga pada bbm nonsubsidi kecuali pertamax bisnisupdate update bisnis onliner
@SBRwaras Hai Sobat. Saat ini BBM Pertamina Green 95 belum tersedia di wilayah Kota/Kabupaten Semarang. Nantikan di kota selanjutnya ya Sob. Terima kasih. - Firman	hai sobat ini bbm pertamax green sediaan wilayah kota kabupaten semarang nanti kota lanjut sob terima kasih firman	hai sobat saat ini bbm pertamax green belum tersedia di wilayah kota kabupaten semarang nantikan di kota selanjutnya ya sob terima kasih firman

Tabel 1 menunjukkan bahwa pipeline SVM yang menerapkan stopword removal dan stemming cenderung menghilangkan kata-kata penting, termasuk kata negasi seperti “tidak”, “belum”, dan “kecuali”, yang berperan dalam menentukan makna sentimen.

Sebaliknya, pipeline IndoBERT mempertahankan struktur kalimat dengan melakukan prapemrosesan minimal. Temuan ini menunjukkan bahwa pelestarian konteks linguistik menjadi faktor kunci dalam meningkatkan kualitas representasi teks, terutama pada data media sosial yang bersifat tidak baku. Hal ini sejalan dengan prinsip kerja model Transformer yang memanfaatkan hubungan kontekstual antar kata.

Hasil Validasi Pelabelan dan Distribusi Data

Evaluasi awal terhadap pelabelan menggunakan VADER menunjukkan nilai Cohen's Kappa sebesar 0,5334 yang termasuk kategori moderat. Setelah dilakukan optimasi melalui lexicon adaptation, threshold tuning, dan ground truth refinement, nilai tersebut meningkat menjadi 0,7086.

Peningkatan ini mengindikasikan bahwa penyesuaian leksikon dan parameter secara signifikan meningkatkan reliabilitas pelabelan otomatis, sehingga dataset layak digunakan untuk proses pelatihan model.

Distribusi sentimen menunjukkan dominasi kelas positif sebanyak 1.319 data, diikuti oleh netral (654) dan negatif (528). Dominasi sentimen positif mencerminkan kecenderungan penerimaan masyarakat terhadap kebijakan BBM campuran etanol, meskipun masih terdapat kritik yang berkaitan dengan aspek ekonomi dan teknis.

Dari perspektif kebijakan publik, hasil ini menegaskan bahwa analisis sentimen dapat digunakan sebagai indikator awal dalam mengevaluasi penerimaan kebijakan berbasis data sosial.

Table 2. Distribusi Train-Test

Kelas Sentimen	Total Data	Data Latih	Data Uji
Negatif	528	422	106
Netral	654	523	131
Positif	1.319	1.055	264
Total	2.501	2.000	501

Tabel 2 menunjukkan bahwa pembagian data menggunakan stratified random sampling mempertahankan proporsi kelas pada data latih dan data uji. Meskipun demikian, kelas positif tetap menjadi kelas mayoritas, sehingga eksperimen tambahan menggunakan class weighting dan SMOTE dilakukan untuk mengevaluasi potensi bias model terhadap kelas mayoritas.

Hasil Pelatihan Dan Evaluasi Model

Tahap pelatihan model dilakukan untuk mengevaluasi performa SVM sebagai model baseline dan IndoBERT sebagai model berbasis Transformer. Pada model SVM, teks terlebih dahulu direpresentasikan ke dalam bentuk vektor menggunakan TF-IDF. Selanjutnya, proses optimasi hyperparameter dilakukan menggunakan Grid Search dengan pendekatan 5-Fold Cross Validation. Grid Search dilakukan terhadap 12 kombinasi parameter, yang melibatkan variasi kernel, nilai C, dan gamma, sehingga menghasilkan total 60 proses fitting. Hasil tuning hyperparameter SVM ditunjukkan pada Tabel 3.

Table 3. Hasil Tuning Hyperparameter SVM

Parameter	Nilai
Metode evaluasi	5-Fold Cross Validation
Jumlah kandidat parameter	12 kombinasi
Total proses fitting	60 fits
Kernel terbaik	Linear
C terbaik	1

Gamma terbaik	Scale
Akurasi rata-rata validasi	69,55%
Standar deviasi	±1,65%

Berdasarkan Tabel 3, konfigurasi terbaik SVM diperoleh pada kernel linear dengan nilai $C = 1$ dan $\text{gamma} = \text{scale}$. Konfigurasi ini menghasilkan akurasi rata-rata validasi sebesar 69,55% dengan standar deviasi $\pm 1,65\%$. Nilai standar deviasi yang relatif kecil menunjukkan bahwa performa SVM cukup stabil pada proses validasi silang. Model terbaik tersebut kemudian digunakan sebagai baseline pada tahap evaluasi data uji.

Sementara itu, model IndoBERT dilatih melalui proses fine-tuning dengan beberapa kombinasi hyperparameter. Pengujian dilakukan terhadap kombinasi batch size dan learning rate untuk memperoleh konfigurasi terbaik. Rekapitulasi hasil pengujian hyperparameter IndoBERT ditunjukkan pada Tabel 4.

Table 4 Hasil Pengujian Hyperparameter IndoBERT (5-Fold)

Skenario	Batch Size	Learning Rate	Rata-rata Akurasi
1	16	2e-05	80,25%
2	16	3e-05	81,30%
3	16	5e-05	81,70%
4	32	2e-05	79,15%
5	32	3e-05	79,85%
6	32	5e-05	82,40%

Berdasarkan Tabel 4, konfigurasi terbaik IndoBERT diperoleh pada skenario 6, yaitu batch size 32 dan learning rate 5e-05, dengan rata-rata akurasi validasi sebesar 82,40%. Konfigurasi ini kemudian digunakan sebagai model IndoBERT terbaik untuk evaluasi pada data uji. Selama proses fine-tuning, early stopping diterapkan untuk mencegah overfitting dan mempertahankan model dengan performa validasi terbaik.

Komparasi Evaluasi Model pada Data Uji

Evaluasi pada data uji dilakukan untuk membandingkan performa model SVM dan IndoBERT terhadap data yang tidak digunakan dalam proses pelatihan. Hasil evaluasi ditunjukkan pada Tabel 5.

Table 5 Evaluasi Komparasi Kinerja Algoritma Klasifikasi

Algoritma Model	Accuracy	Precision	Recall	F1-score
SVM	72,46%	72,07%	72,46%	72,02%
IndoBERT	82,83%	82,81%	82,83%	82,72%

Berdasarkan Tabel 5, IndoBERT menunjukkan performa yang lebih baik dibandingkan SVM pada seluruh metrik evaluasi. IndoBERT memperoleh accuracy sebesar 82,83%, precision sebesar 82,81%, recall sebesar 82,83%, dan F1-score sebesar 82,72%. Sementara itu, SVM memperoleh accuracy sebesar 72,46%, precision sebesar 72,07%, recall sebesar 72,46%, dan F1-score sebesar 72,02%. Selisih akurasi sebesar 10,37 poin persentase menunjukkan

bahwa IndoBERT lebih efektif dalam menangkap konteks semantik pada teks media sosial dibandingkan pendekatan TF-IDF dan SVM.

Keunggulan IndoBERT dapat dijelaskan melalui kemampuan model Transformer dalam memahami konteks kata secara dua arah. Berbeda dengan SVM yang menggunakan representasi TF-IDF berbasis frekuensi kata, IndoBERT mampu mempertimbangkan hubungan antar kata dalam kalimat sehingga lebih baik dalam menangani teks tidak terstruktur, negasi, dan variasi bahasa informal pada platform X.

Analisis Performa Berdasarkan Kelas Sentimen

Selain mengevaluasi performa secara keseluruhan, analisis juga dilakukan pada masing-masing kelas sentimen untuk mengetahui kemampuan model dalam mengenali kelas negatif, netral, dan positif. Evaluasi per kelas dilakukan menggunakan metrik precision, recall, dan F1-score.

Hasil evaluasi per kelas pada model SVM ditunjukkan pada Tabel 6.

Table 6. Performa SVM Berdasarkan Kelas Sentimen

Kelas Sentimen	Precision	Recall	F1-score
Negatif	71,11%	60,38%	65,31%
Netral	65,55%	59,54%	62,40%
Positif	75,68%	83,71%	79,50%

Berdasarkan Tabel 6, model SVM menunjukkan performa terbaik pada kelas positif dengan F1-score sebesar 79,50%. Sebaliknya, performa pada kelas netral dan negatif relatif lebih rendah, dengan F1-score masing-masing sebesar 62,40% dan 65,31%. Perbedaan performa ini menunjukkan bahwa model SVM cenderung lebih mudah mengenali kelas positif yang merupakan kelas mayoritas dalam dataset. Rendahnya recall pada kelas netral dan negatif mengindikasikan bahwa sebagian data pada kedua kelas tersebut masih salah diklasifikasikan ke kelas lain.

Hasil evaluasi per kelas pada model IndoBERT ditunjukkan pada Tabel 7.

Table 7 Rincian Performa IndoBERT Berdasarkan Kelas Sentimen

Kelas Sentimen	Precision	Recall	F1-score
Negatif	76,64%	77,36%	77,00%
Netral	83,62%	74,05%	78,54%
Positif	84,89%	89,39%	87,08%
Rata-rata Weighted	82,81%	82,83%	82,72%

Tabel 7 menunjukkan bahwa IndoBERT menghasilkan performa yang lebih baik dibandingkan SVM pada seluruh kelas sentimen. Peningkatan paling signifikan terlihat pada kelas negatif dan netral yang sebelumnya memiliki performa terendah pada SVM. Pada kelas negatif, F1-score meningkat dari 65,31% menjadi 77,00%, sedangkan pada kelas netral meningkat dari 62,40% menjadi 78,54%.

Kelas positif tetap menjadi kelas dengan performa tertinggi pada kedua model. Namun demikian, peningkatan performa pada kelas negatif dan netral menunjukkan bahwa IndoBERT memiliki kemampuan yang lebih baik dalam memahami konteks kalimat dan hubungan antar kata dibandingkan pendekatan berbasis TF-IDF yang digunakan oleh SVM. Kemampuan representasi kontekstual ini memungkinkan IndoBERT mengurangi kesalahan klasifikasi pada kelas minoritas sehingga menghasilkan performa yang lebih seimbang antar kelas.

Meskipun performa kelas positif masih lebih tinggi dibandingkan kelas lainnya, hasil evaluasi menunjukkan bahwa IndoBERT mampu mempertahankan keseimbangan precision dan recall pada seluruh kelas sentimen. Oleh karena itu, dilakukan analisis tambahan terhadap ketidakseimbangan data menggunakan class weighting dan SMOTE untuk mengevaluasi potensi bias model terhadap kelas mayoritas.

Analisis Pengaruh Penanganan Ketidakseimbangan Data

Distribusi kelas pada dataset menunjukkan bahwa kelas positif memiliki jumlah data yang lebih besar dibandingkan kelas netral dan negatif. Kondisi ini berpotensi menyebabkan model lebih cenderung mempelajari pola dari kelas mayoritas dibandingkan kelas minoritas. Oleh karena itu, dilakukan eksperimen tambahan menggunakan strategi penanganan ketidakseimbangan data untuk mengevaluasi pengaruhnya terhadap performa model dan memastikan bahwa model tidak bias terhadap kelas positif.

Pada pipeline SVM, dua pendekatan digunakan, yaitu class weighting dan Synthetic Minority Oversampling Technique (SMOTE). Sementara itu, pada pipeline IndoBERT digunakan class weighting pada fungsi loss dengan bobot kelas sebesar 1,5798 untuk negatif, 1,2747 untuk netral, dan 0,6319 untuk positif. Eksperimen ini dilakukan sebagai analisis tambahan dan tidak menggantikan hasil utama penelitian.

Hasil evaluasi seluruh skenario ditunjukkan pada Tabel 8.

Table 8. Perbandingan Strategi Penanganan Ketidakseimbangan Data

Model	Accuracy	Macro-F1	Weighted-F1
SVM Baseline	72,46%	-	72,02%
SVM + Class Weight	75,85%	73,24%	76,34%
SVM + SMOTE	76,85%	73,85%	77,07%
IndoBERT Baseline	82,83%	-	82,72%
IndoBERT + Class Weight	73,05%	71,04%	73,49%

Berdasarkan Tabel 8, penerapan class weighting dan SMOTE berhasil meningkatkan performa SVM dibandingkan model baseline. SVM + Class Weight meningkatkan accuracy dari 72,46% menjadi 75,85%, sedangkan SVM + SMOTE menghasilkan performa terbaik pada kelompok model SVM dengan accuracy sebesar 76,85%, macro-F1 sebesar 73,85%, dan weighted-F1 sebesar 77,07%.

Meskipun demikian, performa terbaik SVM yang diperoleh melalui SMOTE masih berada di bawah IndoBERT baseline yang mencapai accuracy sebesar 82,83% dan weighted-F1 sebesar 82,72%. Hasil ini menunjukkan bahwa peningkatan kualitas representasi teks yang

diberikan oleh model Transformer masih memberikan keuntungan yang lebih besar dibandingkan strategi penanganan imbalance pada model machine learning tradisional.

Untuk mengevaluasi kemampuan model dalam mengenali kelas minoritas, analisis juga dilakukan berdasarkan nilai recall pada masing-masing kelas sentimen. Hasilnya ditunjukkan pada Tabel 9.

Table 9. Recall Per Kelas pada Strategi Penanganan Ketidakseimbangan Data

Model	Recall Negatif	Recall Netral	Recall Positif
SVM + Class Weight	72,64%	72,52%	78,79%
SVM + SMOTE	69,81%	70,99%	82,58%
IndoBERT + Class Weight	69,81%	74,05%	73,86%

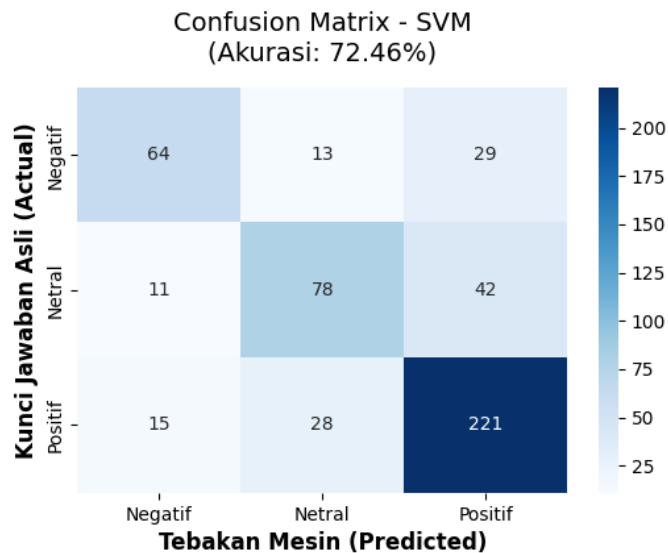
Tabel 9 menunjukkan bahwa class weighting pada SVM menghasilkan recall negatif dan netral yang lebih tinggi dibandingkan SMOTE. Temuan ini mengindikasikan bahwa class weighting lebih efektif dalam meningkatkan sensitivitas model terhadap kelas minoritas. Sebaliknya, SMOTE menghasilkan performa agregat yang lebih tinggi serta recall positif yang lebih besar.

Pada IndoBERT, penerapan class weighting menghasilkan peningkatan perhatian model terhadap kelas minoritas, tetapi menurunkan performa agregat dibandingkan IndoBERT baseline. Accuracy menurun dari 82,83% menjadi 73,05%, sedangkan weighted-F1 menurun dari 82,72% menjadi 73,49%. Hasil ini menunjukkan bahwa strategi penanganan ketidakseimbangan data tidak selalu memberikan dampak positif pada seluruh jenis model.

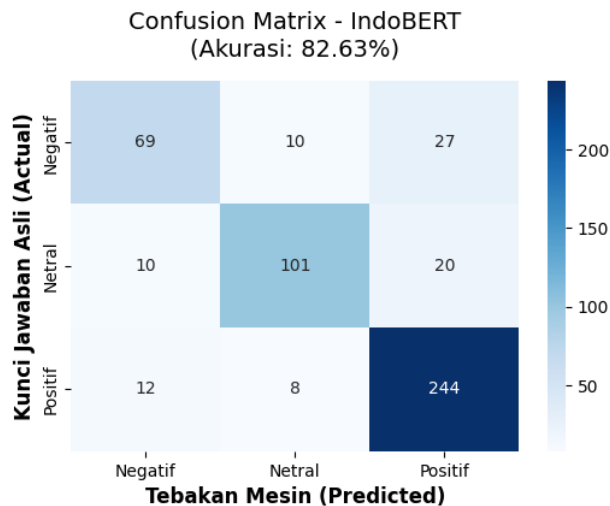
Secara keseluruhan, eksperimen tambahan ini menunjukkan bahwa penanganan ketidakseimbangan data dapat meningkatkan performa pada pipeline SVM, terutama melalui penggunaan SMOTE. Namun, IndoBERT baseline tetap menjadi model dengan performa terbaik pada penelitian ini. Temuan tersebut menunjukkan bahwa kemampuan representasi kontekstual yang dimiliki IndoBERT memiliki pengaruh yang lebih besar terhadap performa klasifikasi dibandingkan strategi penanganan imbalance yang diterapkan pada penelitian ini.

Analisis Confusion Matrix

Selain menggunakan metrik accuracy, precision, recall, dan F1-score, evaluasi model juga dilakukan menggunakan confusion matrix untuk melihat pola kesalahan klasifikasi pada masing-masing kelas sentimen. Analisis ini penting untuk mengetahui apakah model cenderung bias terhadap kelas tertentu, khususnya kelas positif yang merupakan kelas mayoritas dalam dataset.



Gambar 2 Confusion Matrix SVM



Gambar 3 Confusion Matrix IndoBERT

Berdasarkan Gambar 2, model SVM berhasil mengklasifikasikan 64 dari 106 data negatif, 78 dari 131 data netral, dan 221 dari 264 data positif secara benar. Kesalahan klasifikasi terbesar pada SVM terjadi pada kelas netral yang diprediksi sebagai positif sebanyak 42 data, serta kelas negatif yang diprediksi sebagai positif sebanyak 29 data. Pola ini menunjukkan bahwa SVM masih memiliki kecenderungan memprediksi sebagian data ke kelas positif sebagai kelas mayoritas.

Sebaliknya, berdasarkan Gambar 3, model IndoBERT menunjukkan pola prediksi yang lebih baik. IndoBERT berhasil mengklasifikasikan 69 dari 106 data negatif, 101 dari 131 data netral, dan 244 dari 264 data positif secara benar. Dibandingkan SVM, jumlah prediksi benar IndoBERT meningkat pada seluruh kelas, yaitu dari 64 menjadi 69 pada kelas negatif, dari 78 menjadi 101 pada kelas netral, dan dari 221 menjadi 244 pada kelas positif.

Peningkatan paling terlihat terjadi pada kelas netral. Pada model SVM, sebanyak 42 data netral masih salah diklasifikasikan sebagai positif, sedangkan pada IndoBERT jumlah

kesalahan tersebut menurun menjadi 20 data. Hal ini menunjukkan bahwa IndoBERT lebih mampu membedakan ekspresi netral dari ekspresi positif maupun negatif. Kemampuan ini berkaitan dengan representasi kontekstual pada model Transformer yang dapat menangkap hubungan antar kata secara lebih baik dibandingkan representasi TF-IDF pada SVM.

Secara keseluruhan, hasil confusion matrix memperkuat temuan evaluasi metrik sebelumnya bahwa IndoBERT memiliki performa klasifikasi yang lebih baik dibandingkan SVM. Distribusi prediksi IndoBERT yang lebih terkonsentrasi pada diagonal utama menunjukkan bahwa model ini lebih efektif dalam memetakan pola sentimen masyarakat terhadap kebijakan penggunaan BBM campuran etanol di platform X.

Implikasi Penelitian

Hasil penelitian ini memiliki implikasi metodologis dan praktis. Dari sisi metodologis, temuan penelitian menunjukkan bahwa model berbasis Transformer, khususnya IndoBERT, memiliki kemampuan yang lebih baik dibandingkan SVM dalam mengklasifikasikan sentimen pada teks media sosial berbahasa Indonesia. Keunggulan tersebut terlihat dari hasil evaluasi pada data uji, di mana IndoBERT memperoleh accuracy dan F1-score yang lebih tinggi dibandingkan SVM. Hal ini menunjukkan bahwa representasi kontekstual yang digunakan oleh IndoBERT lebih efektif dalam menangkap makna kalimat, termasuk pada teks yang mengandung bahasa informal, negasi, dan struktur kalimat tidak baku

Selain itu, hasil eksperimen tambahan terhadap ketidakseimbangan data menunjukkan bahwa strategi penanganan imbalance tidak memberikan dampak yang sama pada setiap model. Pada pipeline SVM, penerapan SMOTE dan class weighting mampu meningkatkan performa dibandingkan baseline. SMOTE menghasilkan performa agregat yang lebih tinggi, sedangkan class weighting lebih sensitif terhadap kelas minoritas. Sebaliknya, pada pipeline IndoBERT, penerapan class weighting tidak meningkatkan performa agregat dibandingkan IndoBERT baseline. Temuan ini mengindikasikan bahwa pemilihan strategi penanganan ketidakseimbangan data perlu disesuaikan dengan karakteristik model dan bentuk representasi data yang digunakan.

Dari sisi praktis, hasil analisis sentimen menunjukkan bahwa opini masyarakat terhadap kebijakan penggunaan BBM campuran etanol cenderung positif. Hal ini dapat menjadi indikasi awal adanya penerimaan publik terhadap kebijakan energi berbasis bioetanol. Namun demikian, masih terdapat sentimen netral dan negatif yang berkaitan dengan aspek teknis, ekonomi, kesiapan infrastruktur, serta kekhawatiran masyarakat terhadap dampak penggunaan bahan bakar campuran etanol pada kendaraan. Oleh karena itu, hasil penelitian ini dapat digunakan sebagai masukan bagi pemerintah dan pemangku kebijakan dalam merumuskan strategi komunikasi publik yang lebih tepat sasaran.

Secara lebih luas, penelitian ini menunjukkan bahwa analisis sentimen berbasis media sosial dapat digunakan sebagai instrumen pendukung dalam memahami respons masyarakat terhadap kebijakan publik. Meskipun tidak dapat menggantikan survei formal atau kajian kebijakan mendalam, pendekatan ini dapat memberikan gambaran awal yang bersifat real-time mengenai dinamika opini publik. Dengan demikian, integrasi analisis sentimen dalam proses evaluasi kebijakan dapat membantu pengambil keputusan untuk mengidentifikasi isu, kekhawatiran, dan persepsi masyarakat secara lebih cepat dan berbasis data..

SIMPULAN

Penelitian ini berhasil menganalisis sentimen masyarakat terhadap kebijakan penggunaan BBM campuran etanol berdasarkan data media sosial X menggunakan pendekatan NLP. Hasil analisis menunjukkan bahwa sentimen masyarakat didominasi oleh opini positif, yang mengindikasikan adanya kecenderungan penerimaan terhadap kebijakan energi berbasis bioetanol. Meskipun demikian, masih terdapat sentimen netral dan negatif yang menunjukkan adanya kekhawatiran masyarakat terkait aspek teknis kendaraan, biaya, kesiapan infrastruktur, serta implementasi kebijakan di lapangan.

Dari sisi metodologi, model berbasis Transformer, yaitu IndoBERT, menunjukkan performa yang lebih unggul dibandingkan SVM. IndoBERT memperoleh accuracy sebesar **82,83%** dan F1-score sebesar **82,72%**, sedangkan SVM memperoleh accuracy sebesar **72,46%** dan F1-score sebesar **72,02%**. Hasil ini menunjukkan bahwa IndoBERT lebih efektif dalam menangkap konteks semantik pada teks media sosial berbahasa Indonesia dibandingkan pendekatan berbasis TF-IDF dan SVM.

Eksperimen tambahan terhadap ketidakseimbangan data menunjukkan bahwa strategi penanganan imbalance memberikan dampak yang berbeda pada setiap model. Pada pipeline SVM, SMOTE menghasilkan performa agregat terbaik dengan accuracy sebesar **76,85%** dan weighted-F1 sebesar **77,07%**, sedangkan class weighting lebih efektif dalam meningkatkan sensitivitas terhadap kelas minoritas. Pada pipeline IndoBERT, penerapan class weighting tidak meningkatkan performa agregat dibandingkan IndoBERT baseline. Oleh karena itu, IndoBERT baseline tetap dipertahankan sebagai model utama penelitian.

Penelitian ini memberikan kontribusi metodologis dalam penerapan model Transformer untuk analisis sentimen bahasa Indonesia, khususnya pada konteks kebijakan energi berbasis bioetanol. Selain itu, penelitian ini juga memberikan kontribusi praktis sebagai dasar awal bagi pemangku kebijakan dalam memahami persepsi publik berbasis data media sosial. Hasil penelitian menunjukkan bahwa analisis sentimen dapat digunakan sebagai instrumen pendukung untuk mengidentifikasi respons, kekhawatiran, dan kecenderungan opini masyarakat terhadap kebijakan publik.

Meskipun demikian, penelitian ini memiliki beberapa keterbatasan. Dataset yang digunakan masih relatif terbatas, yaitu 2.501 tweet, sehingga berpotensi membatasi generalisasi model terhadap variasi opini masyarakat yang lebih luas. Selain itu, distribusi kelas yang tidak sepenuhnya seimbang masih menjadi tantangan dalam klasifikasi sentimen, terutama pada kelas minoritas. Penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar dan lebih beragam, mengeksplorasi model yang lebih sesuai untuk teks media sosial seperti IndoBERTweet, serta mempertimbangkan pendekatan lain seperti data augmentation, focal loss, weighted sampler, atau Large Language Model untuk meningkatkan performa klasifikasi pada kelas minoritas.

REFERENSI

- A. D. Latief, A. Jarin, M. T. Uliniansyah, E. Nurfadhilah, & D. I. N. Afra. (2023). A Proven Sentiment Annotation Guideline for Indonesian Twitter Data. *2023 International Conference on Computer, Control, Informatics and Its Applications (IC3INA)*, 31–36. <https://doi.org/10.1109/IC3INA60834.2023.10285807>
- A. Parashar, U. Ullah, J. Yadav, H. Rawat, Y. Singh, & B. Chari. (2026). Cultural and Sarcastic Context Aware Sentimental Analysis with Large Language Model and Rule Based Heuristics: An Empirical Study. *2026 IEEE 15th International Conference on Communication Systems and Network Technologies (CSNT)*, 255–260. <https://doi.org/10.1109/CSNT69054.2026.11502364>

- Adit, C. (2025). Dampak Penggunaan Bahan Bakar Minyak Dengan Kandungan Etanol 10% Pada Kendaraan. *Jurnal Exact: Kajian Kemahasiswaan*, 3(2), 231–234.
- Aljabar, A., Ali, I., & Karomah, B. M. (2024). Sentiment Analysis Using Transformer Method. *Journal of Informatics Information System Software Engineering and Applications (INISTA)*, 6(2), 90–97. <https://doi.org/10.20895/inista.v6i2.1383>
- Amrullah, A. (2025). Sentiment Analysis in the Age of Transformers and Large Language Models: A Comprehensive Review of Recent Advances and Future. *Intellithings Journal*, 2(1), 39–49.
- Br Sembiring, A. B. S., Robet, R., & Hoki, L. (2026). Comparison of IndoBERT and SVM Performance in Sentiment Analysis of Digital Education Platforms. *Sinkron*, 10(1), 64–74. <https://doi.org/10.33395/sinkron.v10i1.15472>
- Budianoor, R., Saputro, S. W., Abadi, F., Nugroho, R. A., & Farmadi, A. (2026). Quantifying the Impact of Text Preprocessing on IndoBERT Fine-Tuning for Indonesian Informal Culinary Sentiment Analysis. *Journal of Computing Theories and Applications*, 3(4), 564–581. <https://doi.org/10.62411/jcta.15980>
- Chaerul, M., Septiadi, S., & Triyono, G. (2025). Analisis Sentimen Kebijakan Pembatasan Subsidi Bahan Bakar Minyak di Indonesia Tahun 2024 Menggunakan Algoritma Klasifikasi. *Jurnal Pendidikan dan Teknologi Indonesia*, 5(5), 1471–1484. <https://doi.org/10.52436/1.jpti.825>
- Christian, W., Adamlu, D., Yu, A., & Suhartono, D. (2025). Leveraging IndoBERT and DistilBERT for Indonesian emotion classification in e-commerce reviews. *The 10th International Conference on Computer Science and Computational Intelligence 2025*, 269, 321–330. <https://doi.org/10.1016/j.procs.2025.08.284>
- Darwis, D., Pratiwi, E. S., & Pasaribu, A. F. O. (2020). PENERAPAN ALGORITMA SVM UNTUK ANALISIS SENTIMEN PADA DATA TWITTER KOMISI PEMBERANTASAN KORUPSI REPUBLIK INDONESIA. *Eduatic - Scientific Journal of Informatics Education*, 7(1). <https://doi.org/10.21107/edutic.v7i1.8779>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In J. Burstein, C. Doran, & T. Solorio (Eds.), *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (pp. 4171–4186). Association for Computational Linguistics. <https://doi.org/10.18653/v1/N19-1423>
- Ferdinan, R., Margareta, K., Christyan, S., & Achmad, S. (n.d.). *Utilizing Inter-Annotator Agreement For Effective Fake-Review Detection in E-Commerce*.
- Firdaus, M. A., Anugrah, M. F., Hidayati, S., Rahman, D., Sembiring, R. R., & Novelan, M. S. (n.d.). *PERBANDINGAN NLP DAN SVM DALAM ANALISIS SENTIMEN KOMENTAR INSTAGRAM TERKAIT STIGMA MASYARAKAT TERHADAP BANJIR SUMATERA 2025*.
- Geni, L., Yulianti, E., & Sensuse, D. I. (2023). Sentiment Analysis of Tweets Before the 2024 Elections in Indonesia Using Bert Language Models. *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika*, 9(3), 746–757. <https://doi.org/10.26555/jiteki.v9i3.26490>
- Herindrasti, S., Angelina, B., & Putriwinata, P. (2024). *Pengembangan Kebijakan Energi Terbarukan di Indonesia, Vietnam, dan Laos*. 10(2).
- Hidayatullah, A. F., Apong, R. A., Lai, D. T. C., & Qazi, A. (2024). Word Level Language Identification in Indonesian-Javanese-English Code-Mixed Text. *6th International Conference on AI in Computational Linguistics*, 244, 105–112. <https://doi.org/10.1016/j.procs.2024.10.183>

- Kahlon, N. K., & Singh, W. (2026). A systematic review of web scraping: Techniques, LLM-enhanced approaches, performance metrics, and legal–ethical issues. *Data & Knowledge Engineering*, 164, 102598. <https://doi.org/10.1016/j.datak.2026.102598>
- Koto, F., Lau, J. H., & Baldwin, T. (2021). IndoBERTweet: A Pretrained Language Model for Indonesian Twitter with Effective Domain-Specific Vocabulary Initialization. In M.-F. Moens, X. Huang, L. Specia, & S. W. Yih (Eds.), *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing* (pp. 10660–10668). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.emnlp-main.833>
- Koto, F., Rahimi, A., Lau, J. H., & Baldwin, T. (2020). IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP. *Proceedings of the 28th International Conference on Computational Linguistics*, 757–770. <https://doi.org/10.18653/v1/2020.coling-main.66>
- Laksono, A., & Rahmadani, F. A. (2026). *Analisis Sentimen Opini Publik di platform X (Twitter) terhadap Bencana Banjir Sumatra*. 5.
- Leelawat, N., Jariyapongpaiboon, S., Promjun, A., Boonyarak, S., Saengtabtim, K., Laosunthara, A., Yudha, A. K., & Tang, J. (2022). Twitter data sentiment analysis of tourism in Thailand during the COVID-19 pandemic using machine learning. *Heliyon*, 8(10). <https://doi.org/10.1016/j.heliyon.2022.e10894>
- N. A. Haqimi, D. L. Riatma, Masbahah, Y. F. Rachman, R. Khoirunisa, & A. F. Sani. (2025). Sentiment Analysis on Social Media X Related to Emission Mitigation Using IndoBert and Support Vector Machine Hyperparameters. *2025 5th International Conference of Science and Information Technology in Smart Administration (ICSINTESA)*, 305–310. <https://doi.org/10.1109/ICSINTESA68165.2025.11413636>
- N. Hasanati, Q. Aini, & A. Nuri. (2022). Implementation of Support Vector Machine with Lexicon Based for Sentiment Analysis on Twitter. *2022 10th International Conference on Cyber and IT Service Management (CITSM)*, 1–4. <https://doi.org/10.1109/CITSM56380.2022.9935887>
- Obie, M. & M. Abu Jihad Plaza R. (2024). Analisis Sentimen Terhadap Pengungsi Rohingya Di Aceh Menggunakan Algoritma Naive Bayes dan Support Vector Machine. *Sienna*, 5(2), 93–112. <https://doi.org/10.47637/sienna.v5i2.1430>
- Prasetya, F., Minarni, P. A., & Sonianto. (2025). Optimasi Klasifikasi Risiko Ibu Hamil Menggunakan Support Vector Machine (SVM) Berbasis Particle Swarm Optimization (PSO). *Sienna*, 6(2), 55–62. <https://doi.org/10.47637/sienna.v6i2.2192>
- R. N. Satrya, O. N. Pratiwi, R. Y. Fa'rifah, & J. Abawajy. (2022). Cryptocurrency Sentiment Analysis on the Twitter Platform Using Support Vector Machine (SVM) Algorithm. *2022 International Conference Advancement in Data Science, E-Learning and Information Systems (ICADEIS)*, 01–05. <https://doi.org/10.1109/ICADEIS56544.2022.10037413>
- Riansyahputra, M., Aziz, A., & Purwanto, A. (n.d.). *X Sentiment Analysis on Indonesia's New Capital (IKN)*.
- Santika, S. P., Saputra, M. A., Zahra, A., & Suhartono, D. (2025). Online Gambling Promotion Detection in Indonesian YouTube Comments Using Semi-Supervised IndoBERT Classification. *The 10th International Conference on Computer Science and Computational Intelligence 2025*, 269, 1269–1278. <https://doi.org/10.1016/j.procs.2025.09.068>
- Sasikumar, A., Nantish, A., Udayakumar, A., & Vairavasundaram, S. (2026). Parameter-efficient fine-tuning of Llama models for crisis-tweet categorization: A Pareto-

- informed trade-off analysis. *Results in Engineering*, 30, 110247.
<https://doi.org/10.1016/j.rineng.2026.110247>
- Satria, F., & Husaini, M. (n.d.). *PUBLIC SENTIMENT ANALYSIS OF THE INDONESIAN GOVERNMENT BUDGET EFFICIENCY NEWS USING VADER (VALENCE AWARE DICTIONARY AND SENTIMENT REASONER) ADAPTED TO THE INDONESIAN LANGUAGE LEXICON*.
- Sipayung, Y. R., Mujiyono, S., & Hawa, A. M. (n.d.). *The application of IndoBERT in tourist sentiment analysis: A comparative evaluation with SVM and LSTM*.
- Supriyadi, E., & Makatita, P. N. (2025). Sentiment Analysis of TikTok User Comments on QRIS Adoption in Indonesia Using IndoBERT. *The 10th International Conference on Computer Science and Computational Intelligence 2025*, 269, 121–130.
<https://doi.org/10.1016/j.procs.2025.08.265>
- Tan, K. L., Lee, C. P., & Lim, K. M. (2023). A Survey of Sentiment Analysis: Approaches, Datasets, and Future Research. *Applied Sciences*, 13(7), 4550.
<https://doi.org/10.3390/app13074550>
- Wibowo, R. S., Ramadhan, M. R. W., Nugroho, J. B., & Arifin, M. (2025). *Analisis Sentimen Menggunakan Support Vector Machine dan Naive Bayes Pengguna Game Roblox*. 8.
- Y. Santoso & S. Candra. (2025). Comparative Sentiment Analysis of App Reviews Using TF-IDF and IndoBERT as Feature Extraction with SVM. *2025 5th International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA)*, 13–18.
<https://doi.org/10.1109/ICICyTA68677.2025.11362384>
- Zahira, N. P., & Fadillah, D. P. (2022). *PEMERINTAH INDONESIA MENUJU TARGET NET ZERO EMISSION (NZE) TAHUN 2060 DENGAN VARIABLE RENEWABLE ENERGY (VRE) DI INDONESIA*. 2(2).
- Zahra, N. Z., Farhanatussaidah, S., Afifah, N. N., Muthoharoh, L., Satria, A., & Manullang, M. C. T. (2026). *Benchmarking Logistic Regression, SVM, Naive Bayes, and IndoBERT Fine-Tuning for Sentiment Analysis on Indonesian Product Reviews* (arXiv:2605.03439). arXiv. <https://doi.org/10.48550/arXiv.2605.03439>
- Zhao, Y., & Che, Y. (2026). TF-IDF joint SVM model in library automation bias risk assessment system. *Discover Artificial Intelligence*, 6(1), 249.
<https://doi.org/10.1007/s44163-026-00978-x>