Merkurius: Jurnal Riset Sistem Informasi dan Teknik Informatika
Volume. 4 Nomor. 4 Juli 2026
e-ISSN: 3031-8912; p-ISSN: 3031-8904, Hal. 212-228
DOI: https://doi.org/10.61132/merkurius.v4i4.1727
Tersedia: https://journal.arteii.or.id/index.php/Merkurius

Analisis Sentimen Publik Kebijakan Privasi dan Keamanan Data TikTok dengan Naive Bayes

Zulfa Agustin[1*], Roberto Kaban[2]
[1]Fakultas Sains Dan Teknologi, Program Studi Rekayasa Perangkat Lunak, Institut Teknologi Dan Bisnis Indonesia, Indonesia
[2] Fakultas Sains Dan Teknologi, Program Studi Teknik Informatika, Institut Teknologi Dan Bisnis Indonesia, Indonesia
Email: zulfagustin98@gmail.com[1], roberto.kaban@yahoo.com[2]
**Penulis Korespondensi: zulfagustin98@gmail.com*

Abstract. *This study aims to analyze public sentiment regarding TikTok's privacy and data security policies using the Multinomial Naive Bayes algorithm. Research data were obtained from a public Kaggle dataset containing 100,000 reviews from Indonesian TikTok users on the Google Play Store. After filtering using keywords related to privacy and data security, 9,590 relevant reviews were collected. The research stages include text preprocessing, TF-IDF feature extraction, SMOTE implementation to handle class imbalance, Naive Bayes modeling along with comparison against Linear SVM and IndoBERT, and model evaluation. The results show negative sentiment dominates at 59.7%, followed by positive at 30.1% and neutral at 10.1%. The Naive Bayes model achieves 69.6% accuracy, 64.2% precision, 69.6% recall, and 63.5% F1-Score, while Linear SVM and IndoBERT achieved 74.3% and 81.2% accuracy respectively. This study proves that Naive Bayes is applicable for real-data sentiment classification of TikTok privacy reviews and provides insights for platform managers to improve privacy policy transparency.*

Keywords: *Data Privacy; Naive Bayes; Sentiment Analysis; SMOTE; TikTok.*

Abstrak. Penelitian ini bertujuan menganalisis sentimen publik terhadap kebijakan privasi dan keamanan data TikTok menggunakan algoritma Multinomial Naive Bayes. Data penelitian diperoleh dari dataset publik Kaggle berupa 100.000 ulasan pengguna TikTok Indonesia dari Google Play Store. Setelah proses filtering berdasarkan kata kunci terkait privasi dan keamanan data, diperoleh 9.590 ulasan yang relevan. Tahapan penelitian meliputi preprocessing teks, ekstraksi fitur TF-IDF, penerapan SMOTE untuk mengatasi ketidakseimbangan kelas, pemodelan Naive Bayes serta perbandingan dengan SVM Linear dan IndoBERT, dan evaluasi model. Hasil penelitian menunjukkan sentimen negatif mendominasi sebesar 59,7%, diikuti positif 30,1%, dan netral 10,1%. Model Naive Bayes menghasilkan akurasi 69,6%, presisi 64,2%, recall 69,6%, dan skor F1 63,5%, sedangkan SVM Linear dan IndoBERT masing-masing mencapai akurasi 74,3% dan 81,2%. Penelitian ini membuktikan bahwa Naive Bayes dapat diterapkan untuk klasifikasi sentimen ulasan privasi data TikTok berbasis data real, dan memberikan wawasan bagi pengelola platform untuk meningkatkan transparansi kebijakan privasi

Kata Kunci: Analisis Sentimen; Naive Bayes; Privasi Data; SMOTE; TikTok.

1. LATAR BELAKANG

Perkembangan teknologi informasi telah mendorong peningkatan penggunaan media sosial sebagai sarana komunikasi, berbagi informasi, hingga memperoleh hiburan secara digital (M. Naufal, 2025). Salah satu platform yang mengalami pertumbuhan pengguna sangat pesat adalah TikTok, yang menjadi salah satu aplikasi dengan jumlah unduhan dan pengguna aktif terbesar di Indonesia. Tingginya intensitas penggunaan aplikasi tersebut diikuti dengan meningkatnya perhatian masyarakat terhadap perlindungan privasi dan keamanan data pribadi (Jati, 2025). Berbagai kekhawatiran mengenai pengumpulan data, izin akses perangkat, hingga potensi penyalahgunaan informasi pribadi menjadikan kebijakan privasi sebagai aspek yang semakin diperhatikan oleh pengguna dalam menentukan tingkat kepercayaan terhadap suatu platform digital (Maulana et al., 2025; Apriani et al., 2024; Wang et al., 2025).

Naskah Masuk: 16 Mei 2026; Revisi: 17 Juni 2026; Diterima: 18 Juli 2026; Terbit: 25 Juli 2026

Ulasan yang diberikan pengguna melalui Google Play Store merupakan salah satu sumber informasi yang dapat menggambarkan persepsi masyarakat terhadap kualitas layanan maupun kebijakan yang diterapkan oleh suatu aplikasi (Arya Erlangga et al., 2025). Melalui ulasan tersebut, pengguna menyampaikan pengalaman, kritik, maupun apresiasi yang berkaitan dengan penggunaan aplikasi, termasuk terhadap aspek privasi dan keamanan data. Pemanfaatan data ulasan dalam jumlah besar dapat menghasilkan informasi yang bernilai apabila dianalisis menggunakan teknik *sentiment analysis*, yaitu proses mengidentifikasi kecenderungan opini pengguna ke dalam kategori positif, negatif, maupun netral. Hasil analisis tersebut dapat dimanfaatkan sebagai bahan evaluasi bagi pengembang aplikasi untuk mengetahui respons pengguna terhadap kebijakan yang telah diterapkan (Rahmadani et al., 2022; Ramadhan & Kurniawan, 2025; Almufareh et al., 2025).

Penerapan *sentiment analysis* pada ulasan aplikasi telah banyak dilakukan menggunakan berbagai algoritma klasifikasi, salah satunya adalah *Multinomial Naive Bayes*. Algoritma ini dikenal memiliki proses komputasi yang relatif sederhana, efisien dalam mengolah data teks, serta mampu memberikan hasil klasifikasi yang baik ketika dikombinasikan dengan metode ekstraksi fitur *Term Frequency–Inverse Document Frequency* (*TF-IDF*). Meskipun demikian, performa *Multinomial Naive Bayes* dapat menurun apabila dihadapkan pada distribusi data yang tidak seimbang karena model cenderung lebih mudah mempelajari kelas dengan jumlah data yang dominan dibandingkan kelas minoritas. Kondisi tersebut menyebabkan kemampuan model dalam mengidentifikasi seluruh kategori sentimen menjadi kurang optimal (Madjid et al., 2023; Nahdhudin et al., 2025).

Sejumlah penelitian sebelumnya telah menerapkan algoritma *Naive Bayes* maupun *Support Vector Machine* untuk menganalisis sentimen pada ulasan aplikasi TikTok dan platform digital lainnya. Maulana et al. (2025) memanfaatkan *Naive Bayes* untuk mengklasifikasikan ulasan TikTok yang diperoleh dari Google Play Store, sedangkan Hidayah et al. (2024) membandingkan algoritma *Naive Bayes* dan *Support Vector Machine* pada analisis sentimen aplikasi TikTok. Penelitian lain yang dilakukan Rizki et al. (2025) juga memperlihatkan perbandingan performa kedua algoritma tersebut dalam mengolah ulasan pengguna aplikasi. Meskipun demikian, sebagian besar penelitian masih berfokus pada analisis sentimen secara umum dan belum secara khusus membahas persepsi pengguna terhadap kebijakan privasi dan keamanan data. Selain itu, penelitian terdahulu juga masih terbatas dalam penerapan teknik penanganan *class imbalance* yang berpotensi memengaruhi performa model klasifikasi (Pratiwi & Kamayani, 2024; Ekaputri, 2025).

Dari kondisi tersebut, penelitian ini mengimplementasikan algoritma *Multinomial Naive Bayes* untuk menganalisis sentimen publik terhadap kebijakan privasi dan keamanan data TikTok menggunakan dataset publik yang berisi 100.000 ulasan pengguna Indonesia dari Google Play Store. Setelah dilakukan proses penyaringan berdasarkan kata kunci yang berkaitan dengan privasi dan keamanan data, diperoleh 9.590 ulasan yang relevan sebagai data penelitian. Untuk meningkatkan kualitas klasifikasi pada data yang tidak seimbang, penelitian ini menerapkan *Synthetic Minority Over-sampling Technique* (*SMOTE*) serta membandingkan performa *Multinomial Naive Bayes* dengan *Support Vector Machine* Linear dan *IndoBERT*. Kebaruan penelitian ini terletak pada penggabungan data ulasan nyata yang berfokus pada isu privasi, penerapan teknik *SMOTE* dalam penyeimbangan kelas, serta evaluasi komparatif terhadap tiga model klasifikasi sehingga diharapkan mampu memberikan gambaran yang lebih komprehensif mengenai persepsi pengguna terhadap kebijakan privasi dan keamanan data TikTok (Hermawan et al., 2025; Rhomaningtias et al., 2025; Hidayat & Nastiti, 2024).

Meskipun berbagai penelitian telah membahas analisis sentimen terhadap aplikasi TikTok menggunakan algoritma Naive Bayes maupun *Support Vector Machine*, sebagian besar penelitian masih berfokus pada sentimen umum terhadap kualitas aplikasi. Penelitian yang secara khusus menganalisis persepsi pengguna terhadap kebijakan privasi dan keamanan data masih sangat terbatas. Selain itu, belum banyak penelitian yang mengombinasikan penanganan ketidakseimbangan data menggunakan SMOTE serta membandingkan performa Multinomial Naive Bayes, SVM Linear, dan IndoBERT pada dataset ulasan yang berfokus pada isu privasi digital. Kesenjangan tersebut menjadi dasar penelitian ini.

2. KAJIAN TEORITIS

Analisis Sentimen

Analisis sentimen merupakan salah satu cabang *text mining* yang digunakan untuk mengidentifikasi, mengelompokkan, dan mengevaluasi opini atau persepsi pengguna terhadap suatu objek berdasarkan data berbentuk teks. Melalui proses ini, setiap dokumen atau ulasan dapat diklasifikasikan ke dalam kategori sentimen positif, negatif, maupun netral sehingga menghasilkan informasi mengenai kecenderungan pendapat masyarakat terhadap suatu produk, layanan, atau kebijakan (Ramadhan & Kurniawan, 2025; Almufareh et al., 2025).

Analisis sentimen diterapkan pada ulasan pengguna TikTok yang berkaitan dengan kebijakan privasi dan keamanan data. Hasil klasifikasi sentimen diharapkan mampu memberikan gambaran mengenai persepsi pengguna terhadap pengelolaan data pribadi pada

platform tersebut sehingga dapat menjadi masukan bagi pengembang dalam meningkatkan kualitas layanan dan transparansi kebijakan privasi.

Multinomial *Naive Bayes*

Multinomial Naive Bayes merupakan algoritma klasifikasi berbasis probabilitas yang memanfaatkan Teorema Bayes dengan asumsi bahwa setiap fitur bersifat independen terhadap fitur lainnya. Algoritma ini banyak digunakan dalam analisis sentimen karena memiliki proses komputasi yang sederhana, efisien, dan mampu menangani data teks yang direpresentasikan dalam bentuk frekuensi kemunculan kata maupun bobot *TF-IDF* (Madjid et al., 2023; Nahdhudin et al., 2025).

Multinomial Naive Bayes digunakan sebagai model utama untuk mengklasifikasikan sentimen ulasan pengguna TikTok. Pemilihan algoritma tersebut didasarkan pada kemampuannya dalam mengolah data teks berukuran besar dengan waktu komputasi yang relatif cepat serta performa yang kompetitif pada berbagai penelitian analisis sentimen.

Term Frequency–Inverse Document Frequency (TF-IDF)

Term Frequency–Inverse Document Frequency (*TF-IDF*) merupakan metode ekstraksi fitur yang digunakan untuk mengubah data teks menjadi representasi numerik. Metode ini memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam suatu dokumen (*term frequency*) dan tingkat penyebarannya pada seluruh dokumen (*inverse document frequency*), sehingga kata yang lebih informatif memperoleh bobot yang lebih tinggi (Hidayah et al., 2024).

TF-IDF digunakan sebagai tahap ekstraksi fitur sebelum proses klasifikasi dilakukan. Representasi numerik yang dihasilkan menjadi masukan bagi algoritma *Multinomial Naive Bayes*, *Support Vector Machine*, maupun *IndoBERT* dalam proses pembentukan model klasifikasi sentimen.

Synthetic Minority Over-sampling Technique (SMOTE)

Synthetic Minority Over-sampling Technique (*SMOTE*) merupakan metode *oversampling* yang digunakan untuk mengatasi ketidakseimbangan distribusi kelas (*class imbalance*) dengan membangkitkan data sintetis pada kelas yang memiliki jumlah sampel lebih sedikit. Berbeda dengan teknik duplikasi data, *SMOTE* menghasilkan sampel baru berdasarkan karakteristik data tetangga terdekat (*k-nearest neighbor*) sehingga mampu meningkatkan keberagaman data pelatihan (Hermawan et al., 2025).

SMOTE diterapkan pada data latih untuk menyeimbangkan proporsi kelas sentimen sebelum proses pelatihan model dilakukan. Penerapan metode ini diharapkan mampu

meningkatkan kemampuan model dalam mengenali kelas minoritas tanpa mengubah distribusi data pengujian.

Penelitian Terdahulu

Penerapan algoritma *Naive Bayes* dalam analisis sentimen telah banyak dilakukan pada berbagai objek penelitian, khususnya ulasan aplikasi dan media sosial. Maulana et al. (2025) menerapkan algoritma *Naive Bayes* untuk menganalisis ulasan aplikasi TikTok di Google Play Store dan menunjukkan bahwa metode tersebut mampu mengklasifikasikan sentimen pengguna secara efektif. Penelitian lain oleh Rahmadani et al. (2022) juga memanfaatkan *Naive Bayes Classifier* untuk analisis sentimen media sosial TikTok, sedangkan Hidayah et al. (2024) membandingkan performa *Naive Bayes* dan *Support Vector Machine* pada klasifikasi sentimen ulasan TikTok. Perkembangan penelitian berikutnya menunjukkan bahwa model berbasis *deep learning*, seperti *IndoBERT*, mampu meningkatkan performa klasifikasi karena memiliki kemampuan memahami konteks kalimat berbahasa Indonesia dengan lebih baik dibandingkan metode konvensional (Hidayat & Nastiti, 2024). Selain itu, Hermawan et al. (2025) menunjukkan bahwa penerapan teknik *SMOTE* dapat meningkatkan performa model klasifikasi pada dataset yang mengalami ketidakseimbangan kelas.

Tabel 1. Perbandingan Penelitian Terdahulu.

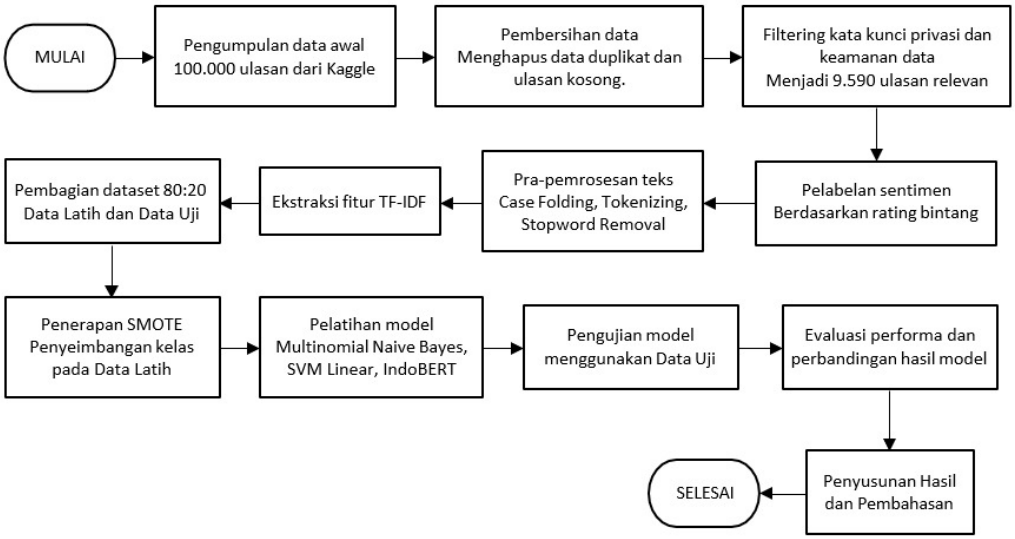
Peneliti	Fokus Penelitian	Metode	Hasil/Kontribusi Utama	Keterkaitan dengan Penelitian	Peneliti
Maulana et al. (2025)	Analisis sentimen ulasan aplikasi TikTok	Naive Bayes	Mengklasifikasikan sentimen pengguna berdasarkan ulasan Google Play Store	Menjadi dasar penggunaan *Naive Bayes* pada data ulasan TikTok	Maulana et al. (2025)
Rahmadani et al. (2022)	Analisis sentimen media sosial TikTok	Naive Bayes Classifier	Menunjukkan efektivitas *Naive Bayes* pada data media sosial	Mendukung penggunaan *Naive Bayes* pada analisis sentimen	Rahmadani et al. (2022)
Hidayah et al. (2024)	Analisis sentimen aplikasi TikTok	Naive Bayes, SVM	Membandingkan performa dua algoritma klasifikasi	Menjadi acuan dalam penggunaan model pembanding SVM	Hidayah et al. (2024)
Hermawan et al. (2025)	Analisis sentimen dengan data tidak seimbang	Naive Bayes + SMOTE	SMOTE meningkatkan performa klasifikasi pada kelas minoritas	Menjadi dasar penerapan SMOTE untuk mengatasi *class imbalance*	Hermawan et al. (2025)
Rhomaningtias et al. (2025)	Analisis sentimen ulasan aplikasi	Naive Bayes dan SVM	Membandingkan performa dua metode klasifikasi	Mendukung evaluasi komparatif antar algoritma	Rhomaningtias et al. (2025)
Hidayat & Nastiti (2024)	Klasifikasi sentimen	IndoBERT	Model *transformer* menghasilkan	Menjadi dasar penggunaan	Hidayat & Nastiti (2024)

menggunakan IndoBERT	performa lebih tinggi dibanding metode konvensional	IndoBERT sebagai model pembanding

Berdasarkan penelitian terdahulu, *Naive Bayes* telah banyak dimanfaatkan dalam analisis sentimen pada ulasan aplikasi maupun media sosial, sedangkan *Support Vector Machine* dan *IndoBERT* sering digunakan sebagai model pembanding untuk meningkatkan performa klasifikasi. Meskipun demikian, penelitian yang secara khusus membahas sentimen pengguna terhadap kebijakan privasi dan keamanan data TikTok masih relatif terbatas. Selain itu, penerapan teknik *SMOTE* pada dataset ulasan yang mengalami *class imbalance* juga belum banyak dikombinasikan dengan perbandingan *Multinomial Naive Bayes, Support Vector Machine*, dan *IndoBERT* dalam satu kerangka evaluasi. Perbedaan tersebut menjadi posisi penelitian ini, yaitu menganalisis sentimen ulasan pengguna TikTok yang berfokus pada isu privasi dan keamanan data menggunakan data ulasan nyata dari Google Play Store, kemudian mengevaluasi performa ketiga model klasifikasi setelah penerapan teknik *SMOTE*.

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen berbasis *machine learning* untuk menganalisis sentimen publik terhadap kebijakan privasi dan keamanan data TikTok. Sumber data berasal dari dataset publik TikTok App Reviews Indonesia – Google Play Store yang tersedia pada platform Kaggle dan berisi 100.000 ulasan pengguna aplikasi TikTok. Selanjutnya dilakukan proses penyaringan menggunakan kata kunci yang berkaitan dengan privasi dan keamanan data, seperti *privacy*, *privasi*, *data*, *bocor*, *keamanan*, *izin*, *akses*, *hack*, *password*, *akun*, dan kata kunci lain yang relevan sehingga diperoleh 9.590 ulasan sebagai data penelitian. Proses filtering dilakukan menggunakan pencarian berbasis kata kunci (keyword matching). Ulasan dipilih apabila mengandung minimal satu kata kunci yang berkaitan dengan privasi dan keamanan data. Selanjutnya dilakukan pemeriksaan duplikasi dan penghapusan data kosong sehingga hanya ulasan yang relevan digunakan sebagai dataset penelitian. Setiap ulasan kemudian diberi label sentimen secara otomatis berdasarkan nilai *rating* pada Google Play Store, yaitu rating 1–2 sebagai sentimen negatif, rating 3 sebagai sentimen netral, serta rating 4–5 sebagai sentimen positif. Pelabelan dilakukan secara otomatis (rule-based labeling) menggunakan nilai rating yang tersedia pada Google Play Store tanpa intervensi anotator manusia sehingga proses pelabelan bersifat konsisten terhadap seluruh data.



Gambar 1. Flowchart Tahapan Penelitian.

Data yang telah diperoleh diproses melalui beberapa tahapan, yaitu *preprocessing*, ekstraksi fitur, pembagian dataset, penanganan ketidakseimbangan kelas, pelatihan model, serta evaluasi hasil klasifikasi. Tahap *preprocessing* meliputi *case folding*, *filtering*, *tokenizing*, dan *stopword removal* untuk menghilangkan karakter yang tidak diperlukan serta menghasilkan teks yang lebih bersih. Selanjutnya, data diubah ke dalam bentuk numerik menggunakan metode *Term Frequency–Inverse Document Frequency* (*TF-IDF*) agar dapat diproses oleh algoritma klasifikasi. Dataset kemudian dibagi menjadi data latih sebesar 80% dan data uji sebesar 20% menggunakan metode *stratified train-test split* dengan *random_state* 42 sehingga proporsi setiap kelas tetap terjaga.

Distribusi data sentimen yang tidak seimbang ditangani menggunakan *Synthetic Minority Over-sampling Technique* (*SMOTE*) pada data latih untuk menghasilkan proporsi kelas yang lebih seimbang tanpa memengaruhi data pengujian. Proses klasifikasi dilakukan menggunakan algoritma *Multinomial Naive Bayes* sebagai model utama, sedangkan *Support Vector Machine* (SVM) Linear dan *IndoBERT* digunakan sebagai model pembanding untuk mengevaluasi perbedaan performa klasifikasi. Evaluasi model dilakukan menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, dan *F1-score* sehingga kemampuan masing-masing model dalam mengklasifikasikan sentimen dapat dibandingkan secara objektif.

4. HASIL DAN PEMBAHASAN

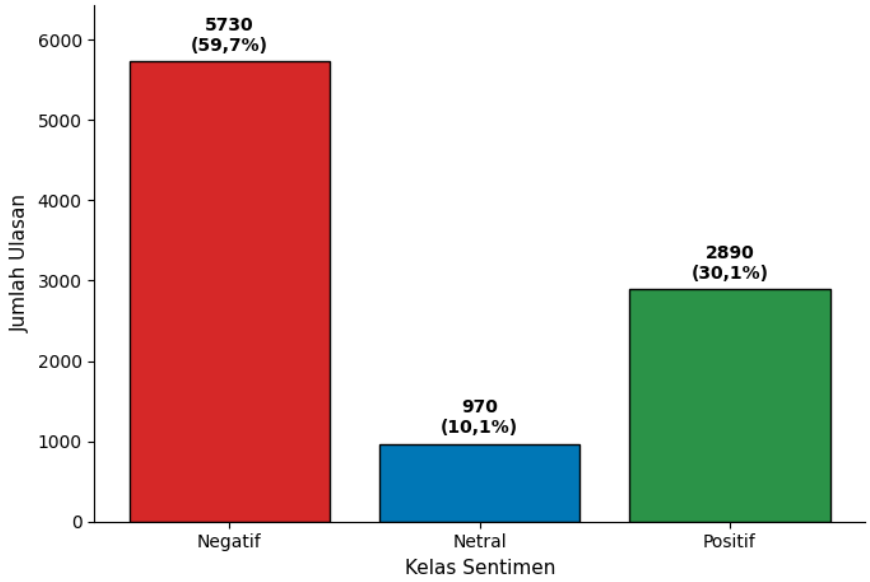
Deskripsi Data Penelitian

Data penelitian bersumber dari dataset publik TikTok App Reviews Indonesia – Google Play Store yang tersedia pada platform Kaggle dan memuat 100.000 ulasan pengguna aplikasi

TikTok. Seluruh data kemudian diseleksi menggunakan kata kunci yang berkaitan dengan privasi dan keamanan data, seperti *privasi*, *privacy*, *data*, *bocor*, *keamanan*, *izin*, *akses*, *hack*, dan beberapa kata kunci relevan lainnya. Proses penyaringan menghasilkan 9.590 ulasan yang sesuai dengan ruang lingkup penelitian. Pelabelan sentimen dilakukan secara otomatis berdasarkan *rating* yang diberikan pengguna di Google Play Store, yaitu rating 1–2 dikategorikan sebagai sentimen negatif, rating 3 sebagai sentimen netral, serta rating 4–5 sebagai sentimen positif.

Tabel 2. Distribusi Data Sentimen Ulasan TikTok Terkait Privasi dan Keamanan Data.

Kelas Sentimen	Jumlah Ulasan	Persentase (%)
Negatif	5.730	59,7
Netral	970	10,1
Positif	2.890	30,1
Total	9.590	100



Gambar 2. Distribusi Sentimen Ulasan TikTok Terkait Privasi dan Keamanan Data.

Berdasarkan tabel 2 dan gambar 2, sentimen negatif mendominasi data penelitian dengan jumlah 5.730 ulasan atau 59,7% dari keseluruhan data. Sentimen positif menempati urutan kedua sebanyak 2.890 ulasan atau 30,1%, sedangkan sentimen netral hanya berjumlah 970 ulasan atau 10,1%. Distribusi tersebut menunjukkan bahwa sebagian besar pengguna memberikan tanggapan yang cenderung negatif terhadap isu privasi dan keamanan data pada aplikasi TikTok. Dominasi sentimen negatif mengindikasikan bahwa aspek perlindungan data pribadi masih menjadi perhatian utama pengguna ketika menggunakan layanan media sosial, sejalan dengan temuan Wang et al. (2025) yang menyatakan bahwa kekhawatiran terhadap privasi merupakan salah satu faktor yang memengaruhi tingkat kepercayaan pengguna terhadap platform digital.

Ketidakseimbangan jumlah data antarkelas juga terlihat cukup jelas, terutama pada kelas sentimen netral yang memiliki proporsi jauh lebih kecil dibandingkan dua kelas lainnya. Kondisi tersebut berpotensi memengaruhi proses pembelajaran model klasifikasi karena algoritma cenderung mempelajari pola dari kelas yang jumlah datanya lebih dominan. Fenomena *class imbalance* seperti ini banyak dijumpai pada penelitian analisis sentimen berbasis ulasan pengguna dan menjadi salah satu faktor yang dapat menurunkan kemampuan model dalam mengenali kelas minoritas (Hermawan et al., 2025). Oleh sebab itu, proses penanganan ketidakseimbangan data menjadi tahapan penting sebelum dilakukan pelatihan model klasifikasi.

Hasil *Preprocessing* dan Ekstraksi Fitur

Tahap *preprocessing* dilakukan untuk membersihkan data ulasan sebelum digunakan dalam proses klasifikasi. Proses ini meliputi *case folding*, *filtering*, *tokenizing*, dan *stopword removal* sehingga teks yang mengandung karakter, simbol, angka, maupun kata yang tidak memiliki makna penting dapat diubah menjadi data yang lebih terstruktur. Hasil *preprocessing* menunjukkan bahwa seluruh 9.590 ulasan berhasil diproses tanpa mengurangi jumlah data yang digunakan sebagai dataset penelitian. Perubahan tersebut menghasilkan representasi teks yang lebih konsisten sehingga mampu meningkatkan kualitas ekstraksi fitur dan meminimalkan pengaruh *noise* terhadap proses klasifikasi (Madjid et al., 2023).

Tabel 3. Contoh Hasil Preprocessing Ulasan Pengguna.

Ulasan Sebelum *Preprocessing*	**Hasil *Preprocessing***
@user123 TikTok bocor data privasi banget 😡 akun ku di bajak!!	*tiktok bocor data privasi akun bajak*

Perubahan teks pada tabel 3 menunjukkan bahwa proses *preprocessing* mampu menghilangkan karakter yang tidak diperlukan tanpa mengubah makna utama dari ulasan pengguna. Penghapusan simbol, tanda baca, *emoji*, dan kata yang tidak relevan menghasilkan representasi teks yang lebih sederhana sehingga setiap kata memiliki kontribusi yang lebih jelas pada tahap ekstraksi fitur. Tahapan ini menjadi bagian penting dalam analisis sentimen karena kualitas data masukan sangat memengaruhi performa model klasifikasi. Untuk memperoleh gambaran mengenai kata yang paling sering muncul, hasil *preprocessing* divisualisasikan menggunakan *WordCloud* sebagaimana ditunjukkan pada gambar 3.

Gambar 3. WordCloud Ulasan TikTok Terkait Privasi dan Keamanan Data.

Visualisasi *WordCloud* memperlihatkan bahwa kata akun, tiktok, bocor, data, privasi, blokir, dan keamanan merupakan kata yang paling dominan muncul pada ulasan pengguna. Dominasi kata-kata tersebut menunjukkan bahwa sebagian besar pembahasan pengguna berkaitan dengan keamanan akun, perlindungan data pribadi, serta akses terhadap aplikasi. Temuan ini memperkuat hasil distribusi sentimen pada subbab sebelumnya yang menunjukkan dominasi sentimen negatif, sehingga mengindikasikan bahwa isu privasi masih menjadi perhatian utama pengguna TikTok (Wang et al., 2025).

Setelah proses *preprocessing* selesai, data teks diubah menjadi representasi numerik menggunakan metode *Term Frequency–Inverse Document Frequency* (*TF-IDF*). Metode ini memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam dokumen serta tingkat penyebarannya pada seluruh dataset, sehingga kata yang memiliki kontribusi lebih besar terhadap proses klasifikasi memperoleh bobot yang lebih tinggi (Hidayah et al., 2024).

Tabel 4. Sepuluh Kata dengan Bobot TF-IDF Tertinggi.

No	Kata	Bobot TF-IDF
1	akun	0,0825
2	tiktok	0,0574
3	blokir	0,0338
4	tiba	0,0330
5	kenapa	0,0300
6	padahal	0,0271
7	tolong	0,0264
8	bagus	0,0245
9	aplikasi	0,0228
10	bikin	0,0221

Dari tabel 4 di atas, kata akun memiliki bobot *TF-IDF* tertinggi, diikuti oleh tiktok dan blokir. Hasil tersebut menunjukkan bahwa ketiga kata tersebut memberikan kontribusi paling besar dalam membedakan karakteristik sentimen pada data ulasan. Kemunculan kata akun, blokir, dan tolong juga mengindikasikan bahwa banyak pengguna menyampaikan keluhan terkait akses akun maupun keamanan penggunaan aplikasi. Representasi numerik yang

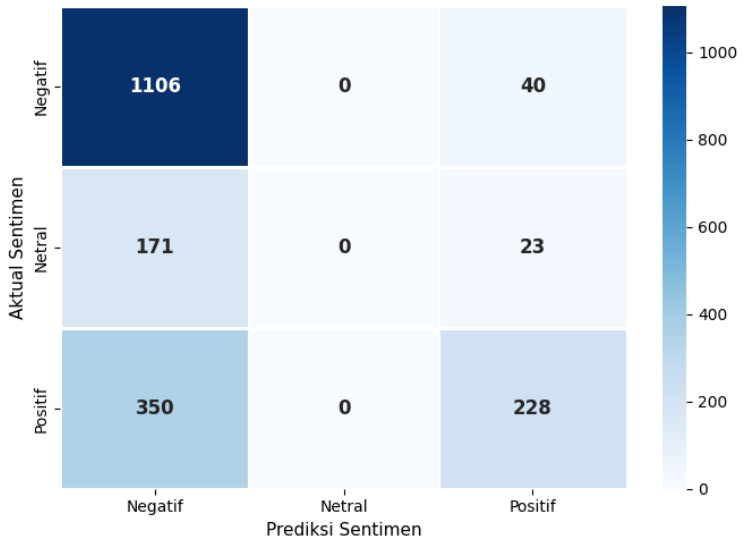
dihasilkan melalui *TF-IDF* selanjutnya digunakan sebagai masukan bagi model *Multinomial Naive Bayes*, *Support Vector Machine*, dan *IndoBERT* dalam proses klasifikasi sentimen.

Hasil Klasifikasi Menggunakan Multinomial *Naive Bayes*

Setelah data melalui tahap *preprocessing* dan ekstraksi fitur menggunakan *TF-IDF*, proses klasifikasi dilakukan menggunakan algoritma *Multinomial Naive Bayes*. Model dilatih menggunakan data latih yang telah dipersiapkan, kemudian dievaluasi menggunakan data uji untuk mengetahui kemampuan model dalam mengklasifikasikan sentimen pengguna ke dalam kategori negatif, netral, dan positif. Hasil klasifikasi divisualisasikan melalui *confusion matrix* yang menunjukkan distribusi prediksi model terhadap setiap kelas sentimen.

Tabel 5. Confusion Matrix Model Multinomial Naive Bayes.

Aktual / Prediksi	Negatif	Netral	Positif
Negatif	1106	0	40
Netral	171	0	23
Positif	350	0	228



Gambar 4. Confusion Matrix Model Multinomial Naive Bayes.

Data dari tabel 5 dan gambar 4, model *Multinomial Naive Bayes* mampu mengidentifikasi kelas sentimen negatif dengan jumlah prediksi benar yang paling tinggi dibandingkan kelas lainnya. Sebanyak 1.106 ulasan negatif berhasil diklasifikasikan dengan benar, sedangkan sebagian ulasan positif dan netral masih diprediksi sebagai sentimen negatif. Kondisi tersebut menunjukkan bahwa model memiliki kecenderungan mengenali pola pada kelas mayoritas karena proporsi data sentimen negatif lebih besar dibandingkan kelas lainnya.

Sebaliknya, kelas sentimen netral belum dapat dikenali secara optimal. Seluruh ulasan yang termasuk kategori netral masih banyak dialihkan ke kelas negatif maupun positif sehingga tidak terdapat prediksi yang tepat pada kelas tersebut. Fenomena ini menunjukkan bahwa

karakteristik kosakata pada sentimen netral memiliki kemiripan dengan dua kelas lainnya, sehingga batas antar kelas menjadi kurang jelas. Temuan tersebut sejalan dengan Nahdhudin et al. (2025) yang menyatakan bahwa *Multinomial Naive Bayes* cenderung mengalami kesulitan membedakan kelas dengan distribusi data yang kecil dan karakteristik kata yang saling tumpang tindih.

Sebagai ilustrasi, model memberikan prediksi sentimen negatif terhadap ulasan yang mengandung kata-kata seperti *bocor*, *privasi*, dan *akun*. Kata-kata tersebut memiliki bobot *TF-IDF* yang relatif tinggi sehingga memberikan kontribusi besar dalam proses pengambilan keputusan model. Kondisi tersebut memperlihatkan bahwa representasi fitur yang dihasilkan melalui *TF-IDF* mampu membantu model mengenali pola sentimen negatif, meskipun kemampuan dalam membedakan sentimen netral masih perlu ditingkatkan.

Evaluasi Performa Model dan Pengaruh *SMOTE*

Performa model *Multinomial Naive Bayes* dievaluasi menggunakan empat metrik, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*. Sebelum menyajikan hasil evaluasi akhir, dilakukan perbandingan performa model sebelum dan sesudah penerapan *Synthetic Minority Over-sampling Technique* (*SMOTE*) untuk mengetahui pengaruh teknik penyeimbangan data terhadap kemampuan klasifikasi.

Sebelum penerapan SMOTE, distribusi kelas menunjukkan dominasi sentimen negatif sebesar 59,7%, sedangkan kelas netral hanya sebesar 10,1%. Ketidakseimbangan tersebut menyebabkan model cenderung mempelajari pola kelas mayoritas.

Tabel 6. Perbandingan Performa Model Sebelum dan Sesudah Penerapan SMOTE.

Metrik Evaluasi	Sebelum SMOTE	Sesudah SMOTE	Peningkatan
Accuracy	64,2%	69,6%	+5,4%
Precision (*Weighted*)	58,7%	64,2%	+5,5%
Recall (*Weighted*)	56,3%	69,6%	+13,3%
F1-Score (*Weighted*)	57,2%	63,5%	+6,3%

Hasil pada tabel 6 menunjukkan bahwa penerapan *SMOTE* memberikan peningkatan pada seluruh metrik evaluasi. Nilai *accuracy* meningkat dari 64,2% menjadi 69,6%, sedangkan *recall* mengalami peningkatan paling besar, yaitu sebesar 13,3%. Peningkatan tersebut menunjukkan bahwa penambahan data sintetis pada kelas minoritas mampu membantu model mempelajari pola data secara lebih seimbang sehingga jumlah prediksi yang benar menjadi lebih tinggi. Temuan ini sejalan dengan Hermawan et al. (2025) yang menyatakan bahwa *SMOTE* dapat meningkatkan kemampuan model klasifikasi dalam menghadapi dataset yang mengalami *class imbalance*.

Tabel 7. Performa Model Berdasarkan Kelas Sentimen Sebelum dan Sesudah SMOTE.

Kelas	Metrik	Sebelum	Sesudah
Negatif	Precision	75,0%	68,0%
	Recall	80,0%	96,5%
	F1-Score	77,0%	80,0%
Positif	Precision	55,0%	78,0%
	Recall	52,0%	39,0%
	F1-Score	53,0%	52,0%
Netral	Precision	0,0%	0,0%
	Recall	0,0%	0,0%
	F1-Score	0,0%	0,0%
Kelas	Metrik	Sebelum	Sesudah

Berdasarkan tabel 7, peningkatan performa tidak terjadi secara merata pada seluruh kelas sentimen. Kelas negatif mengalami peningkatan *recall* yang cukup signifikan hingga mencapai 96,5%, menunjukkan bahwa sebagian besar ulasan negatif berhasil dikenali oleh model. Di sisi lain, *precision* pada kelas positif meningkat menjadi 78,0%, tetapi nilai *recall* mengalami penurunan. Kondisi tersebut mengindikasikan adanya *trade-off*, yaitu model menjadi lebih selektif ketika memberikan prediksi positif sehingga jumlah prediksi positif yang benar meningkat, namun masih terdapat sebagian ulasan positif yang diklasifikasikan ke kelas lain.

Kelas sentimen netral masih menunjukkan nilai *precision*, *recall*, dan *F1-score* sebesar 0%. Hasil tersebut mengindikasikan bahwa penerapan *SMOTE* belum mampu mengatasi karakteristik kelas netral yang memiliki jumlah data relatif sedikit dan pola kosakata yang banyak beririsan dengan kelas negatif maupun positif. Kondisi serupa juga ditemukan pada penelitian Nahdhudin et al. (2025), yang menjelaskan bahwa *Multinomial Naive Bayes* memiliki keterbatasan dalam membedakan kelas yang memiliki karakteristik teks saling berdekatan meskipun distribusi datanya telah diseimbangkan.

Secara keseluruhan, penerapan *SMOTE* memberikan dampak positif terhadap performa model, terutama dalam meningkatkan kemampuan mengenali kelas mayoritas dan memperbaiki nilai evaluasi secara keseluruhan. Meskipun demikian, hasil evaluasi juga menunjukkan bahwa peningkatan distribusi data belum sepenuhnya mampu mengatasi kesulitan model dalam mengidentifikasi sentimen netral. Oleh karena itu, diperlukan model lain yang memiliki kemampuan memahami konteks kalimat secara lebih baik untuk memperoleh hasil klasifikasi yang lebih optimal.

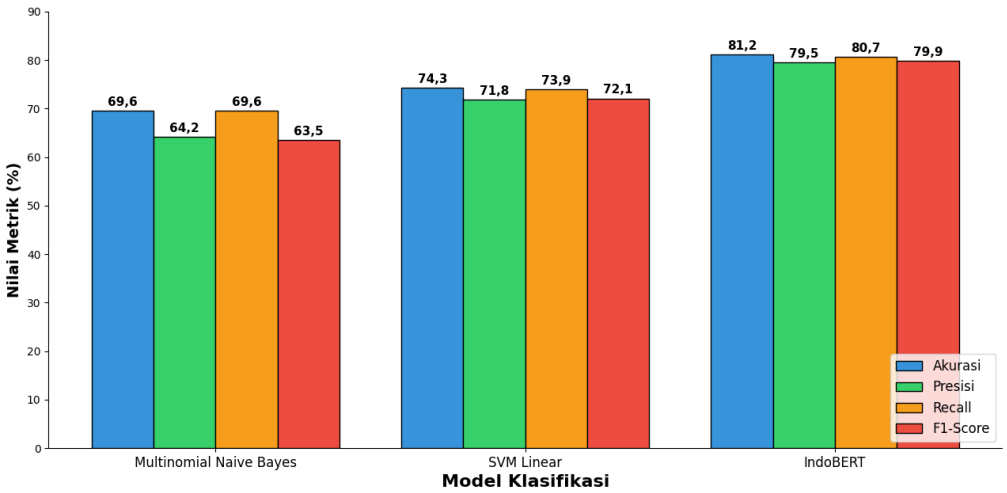
Perbandingan Performa *Multinomial Naive Bayes*, *Support Vector Machine*, dan *IndoBERT*

Untuk memperoleh gambaran mengenai efektivitas algoritma yang digunakan, dilakukan perbandingan performa antara *Multinomial Naive Bayes*, *Support Vector Machine*

(SVM) Linear, dan *IndoBERT*. Ketiga model dievaluasi menggunakan dataset yang sama sehingga perbedaan hasil yang diperoleh mencerminkan kemampuan masing-masing algoritma dalam mengklasifikasikan sentimen ulasan pengguna TikTok.

Tabel 8. Perbandingan Performa Model Klasifikasi.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Multinomial Naive Bayes	69,6	64,2	69,6	63,5
SVM Linear	74,3	71,8	73,9	72,1
IndoBERT	81,2	79,5	80,7	79,9



Gambar 5. Perbandingan Performa Ketiga Model Klasifikasi.

Berdasarkan tabel 8 dan gambar 5, *IndoBERT* menghasilkan performa terbaik pada seluruh metrik evaluasi dengan nilai *accuracy* sebesar 81,2%, *precision* 79,5%, *recall* 80,7%, dan *F1-score* 79,9%. Hasil tersebut menunjukkan bahwa model berbasis *transformer* memiliki kemampuan yang lebih baik dalam memahami hubungan antarkata dan konteks kalimat dibandingkan algoritma klasifikasi konvensional. Kemampuan tersebut memungkinkan *IndoBERT* membedakan karakteristik sentimen secara lebih akurat, terutama pada ulasan yang memiliki struktur bahasa lebih kompleks. Temuan ini sejalan dengan penelitian Hidayat dan Nastiti (2024), yang menyatakan bahwa *IndoBERT* memberikan performa lebih tinggi dibandingkan model berbasis *machine learning* pada klasifikasi sentimen berbahasa Indonesia.

SVM Linear menempati posisi kedua dengan *accuracy* sebesar 74,3%. Dibandingkan *Multinomial Naive Bayes*, algoritma ini mampu menghasilkan nilai *precision*, *recall*, dan *F1-score* yang lebih tinggi. Hal tersebut menunjukkan bahwa SVM memiliki kemampuan yang lebih baik dalam membentuk batas pemisah (*decision boundary*) antar kelas sehingga proses klasifikasi menjadi lebih stabil, terutama pada data teks yang memiliki karakteristik saling berdekatan. Hasil ini konsisten dengan penelitian Rhomaningtias et al. (2025), yang

melaporkan bahwa SVM cenderung memberikan performa lebih baik dibandingkan *Naive Bayes* pada analisis sentimen ulasan aplikasi.

Sementara itu, *Multinomial Naive Bayes* menghasilkan *accuracy* sebesar 69,6%, lebih rendah dibandingkan kedua model pembanding. Walaupun demikian, algoritma ini tetap mampu memberikan hasil klasifikasi yang cukup baik dengan proses komputasi yang relatif sederhana dan efisien. Kinerja tersebut menunjukkan bahwa *Multinomial Naive Bayes* masih layak digunakan sebagai alternatif pada analisis sentimen, terutama ketika kebutuhan utama adalah kecepatan pelatihan model dan efisiensi komputasi. Hasil tersebut juga mendukung penelitian Madjid et al. (2023), yang menyatakan bahwa *Naive Bayes* tetap menjadi salah satu algoritma yang efektif untuk klasifikasi data teks meskipun memiliki keterbatasan dalam memahami konteks kalimat.

Perbedaan performa ketiga model menunjukkan bahwa karakteristik algoritma sangat memengaruhi kualitas klasifikasi sentimen. *Multinomial Naive Bayes* bekerja berdasarkan probabilitas kemunculan kata sehingga lebih sensitif terhadap distribusi frekuensi fitur, sedangkan SVM membangun hiperbidang pemisah yang mampu meningkatkan kemampuan diskriminasi antar kelas. Berbeda dengan kedua algoritma tersebut, *IndoBERT* memanfaatkan representasi bahasa berbasis konteks sehingga dapat memahami hubungan antarkata secara lebih komprehensif. Karakteristik tersebut menjadi alasan utama mengapa *IndoBERT* mampu menghasilkan performa terbaik pada penelitian ini.

5. KESIMPULAN DAN SARAN

Analisis sentimen terhadap kebijakan privasi dan keamanan data TikTok menggunakan dataset publik dari Google Play Store menunjukkan bahwa sentimen negatif mendominasi dengan persentase 59,7%, diikuti sentimen positif sebesar 30,1% dan sentimen netral sebesar 10,1%. Hasil evaluasi memperlihatkan bahwa algoritma *Multinomial Naive Bayes* menghasilkan *accuracy* sebesar 69,6%, sedangkan SVM Linear dan *IndoBERT* memperoleh performa yang lebih baik dengan *accuracy* masing-masing sebesar 74,3% dan 81,2%. Penerapan *Synthetic Minority Over-sampling Technique* (*SMOTE*) mampu meningkatkan performa *Multinomial Naive Bayes*, terutama pada nilai *recall*, meskipun model masih mengalami kesulitan dalam mengidentifikasi kelas sentimen netral akibat karakteristik data yang saling beririsan.

Hasil tersebut menunjukkan bahwa *IndoBERT* menjadi model dengan performa terbaik untuk klasifikasi sentimen ulasan pengguna TikTok terkait privasi dan keamanan data, sedangkan *Multinomial Naive Bayes* tetap dapat menjadi alternatif yang efisien pada kebutuhan

komputasi yang lebih sederhana. Penelitian selanjutnya disarankan mengeksplorasi metode penyeimbangan data lain, seperti *ADASYN* atau *SMOTEENN*, serta melakukan *fine-tuning* model *IndoBERT* atau membandingkannya dengan model *transformer* terbaru agar kemampuan klasifikasi, khususnya pada kelas sentimen netral, dapat ditingkatkan.

DAFTAR REFERENSI

Almufareh, M., Jhanjhi, N. Z., Humayun, M., Alwakid, G. N., Danish, J., & Almuayqil, S. (2025). Integrating sentiment analysis with machine learning for cyberbullying detection on social media. *IEEE Access*, *13*, 1–1.

Apriani, E., Oktavianalisti, F., Monasari, L. D. H., Winarni, I., & Hanif, I. F. (2024). Analisis sentimen penggunaan TikTok sebagai media pembelajaran menggunakan algoritma Naïve Bayes Classifier. *MALCOM, 4*(3), 1160–1168.

Arya Erlangga, Yani Parti Astuti, Etika Kartikadarma, Sindhu Rakasiwi, & Egia Rosi Subhiyakto. (2025). Penggunaan Algoritma Naïve Bayes dengan Polarity Textblob untuk Analisis Sentimen pada Acara ASEAN CUP 2024 U-16 di Media Sosial Twitter. *Switch : Jurnal Sains Dan Teknologi Informasi*, *3*(3), 07–19. https://doi.org/10.62951/switch.v3i1.357

Hermawan, M. A., Faqih, A., & Dwilestari, G. (2025). Implementasi akurasi model Naive Bayes menggunakan SMOTE dalam analisis sentimen pengguna aplikasi BRImo. *Jurnal Informatika dan Teknik Elektro Terapan, 13*(1), 855–862.

Hidayah, I. A., Kusumawati, R., Abidin, Z., & Imamuddin, M. (2024). Analisis sentimen publik terhadap aplikasi TikTok menggunakan algoritma Naive Bayes dan Support Vector Machine. *CNAHPC, 6*(2), 881–891.

Hidayat, W. A., & Nastiti, V. R. S. (2024). Perbandingan kinerja pre-trained IndoBERT-Base dan IndoBERT-Lite pada klasifikasi sentimen ulasan TikTok Tokopedia Seller Center dengan model IndoBERT. *JSII (Jurnal Sistem Informasi), 11*(2).

Jati Nur Shiddiq. (2025). Implementasi Sistem Informasi Web untuk Visualisasi dan Keamanan Data Susut Energi. *Merkurius : Jurnal Riset Sistem Informasi Dan Teknik Informatika, 3*(6), 349–357. https://doi.org/10.61132/merkurius.v3i6.1261

M. Naufal Ramadhan. (2025). Pemanfaatan Literasi Digital Berbasis Teknologi Informasi untuk Kesadaran Keamanan Data dan Pencegahan Kejahatan Siber. *Neptunus: Jurnal Ilmu Komputer Dan Teknologi Informasi*, *3*(1), 151–159. https://doi.org/10.61132/neptunus.v3i1.689

Madjid, M. F., Ratnawati, D. E., & Rahayudi, B. (2023). Analisis sentimen pada ulasan aplikasi menggunakan Support Vector Machine dan klasifikasi Naïve Bayes. *SinkrOn, 7*(1), 556–562.

Maulana, M. T. F., Nugroho, B. I., & Utami, E. U. S. (2025). Analisis sentimen ulasan aplikasi TikTok di Google Play Store menggunakan algoritma Naive Bayes. *RIGGS, 4*(3), 6627–6635.

Nahdhudin, M., Wibowo, N. C. H., Handayani, M. R., & Umam, K. (2025). Klasifikasi sentimen masyarakat terhadap aplikasi TikTok menggunakan algoritma Naive Bayes. *Jurnal Teknologi Informasi dan Ilmu Komputer, 1*(3), 107–112.

Pratiwi, A. A., & Kamayani, M. (2024). Perbandingan pelabelan data dalam analisis sentimen kurikulum proyek di platform TikTok: Pendekatan Naïve Bayes. *Eksplora*, *14*(1), 96–107.

Rahmadani, P. S., Tampubolon, F. C., Jannah, A. N., Hutabarat, N. L. H., & Simarmata, A. M. (2022). Analisis sentimen media sosial TikTok menggunakan algoritma Naive Bayes Classifier. *SinkrOn*, *6*(3), 995–999.

Ramadhan, R. F., & Kurniawan, R. (2025). Sentiment analysis of privacy issues in the digital era using the Naïve Bayes method. *Mandiri*, *14*(2), 170–178.

Rhomaningtias, L., Khairunisa, A., Wara, S. S. M., & Hindrayani, K. M. (2025). Analisis sentimen ulasan aplikasi Smile Indonesia menggunakan metode Naive Bayes dan Support Vector Machine (SVM). *HOAQ*, *16*(1), 79–91.

Wang, J., Cao, Q., & Zhu, X. (2025). Privacy disclosure on social media: The roles of platform features, group effects, trust and privacy concerns. *Library Hi Tech*, *43*(2–3), 1035–1059.