TIN: Terapan Informatika Nusantara
Vol 7, No 1, June 2026, page 653-661
ISSN 2722-7987 (Media Online)
Website https://ejurnal.seminar-id.com/index.php/tin
DOI 10.47065/tin.v7i1.10243

Evaluasi Pseudo-Labeling IndoRoBERTa dan InSet Lexicon dengan SVM pada Komentar TikTok

Yehezkiel Juandro Metta*, Aswan Supriyadi Sunge, Asep Suprianto

Fakultas Teknik, Program Studi Teknik Informatika, Universitas Pelita Bangsa, Bekasi, Indonesia
Email: [1,*]yehezkiel.312210376@mhs.pelitabangsa.ac.id, [2]aswan.sunge@pelitabangsa.ac.id, [3]asep.supriyanto@pelitabangsa.ac.id
Email Penulis Korespondensi: yehezkiel.312210376@mhs.pelitabangsa.ac.id

Abstrak–TikTok menghasilkan komentar publik dalam jumlah besar yang dapat dimanfaatkan untuk mengidentifikasi kecenderungan opini masyarakat, termasuk pada kasus kecelakaan mobil program Makan Bergizi Gratis. Namun, bahasa media sosial yang informal dan distribusi kelas yang tidak seimbang dapat memengaruhi proses pelabelan serta klasifikasi sentimen. Penelitian ini bertujuan mengevaluasi pseudo-label sentimen yang dihasilkan IndoRoBERTa dan InSet Lexicon melalui klasifikasi Support Vector Machine dengan penerapan Synthetic Minority Over-sampling Technique. Penelitian menggunakan pendekatan kuantitatif eksperimental terhadap 10.309 komentar TikTok yang diperoleh melalui scraping menggunakan Apify. Karena penelitian tidak menggunakan anotasi manusia sebagai ground truth, label positif, negatif, dan netral yang dihasilkan kedua metode diperlakukan sebagai pseudo-label. Tahapan penelitian meliputi filtering, preprocessing, pembentukan pseudo-label, ekstraksi fitur Term Frequency-Inverse Document Frequency, pembagian data 80:20, penerapan SMOTE pada data latih, klasifikasi SVM, dan evaluasi menggunakan accuracy, precision, recall, serta F1-score. Hasil menunjukkan bahwa SVM paling mampu mereproduksi pseudo-label InSet Lexicon, dengan accuracy 0,865 tanpa SMOTE. Setelah SMOTE, precision meningkat menjadi 0,870 dengan F1-score tetap 0,865, serta recall kelas netral meningkat dari 0,767 menjadi 0,807. Temuan ini menunjukkan bahwa InSet Lexicon menghasilkan pseudo-label yang lebih konsisten untuk dipelajari SVM pada dataset penelitian, sementara SMOTE terutama membantu meningkatkan pengenalan kelas minoritas.

Kata Kunci: Analisis Sentimen; Pseudo-Labeling; IndoRoBERTa; InSet Lexicon; SVM; SMOTE

Abstract–TikTok generates large volumes of public comments that can be used to identify trends in public opinion, including responses to the accident involving a vehicle used in the Free Nutritious Meal program. However, informal social media language and imbalanced class distributions may affect sentiment labeling and classification performance. This study evaluates sentiment pseudo-labels generated by IndoRoBERTa and InSet Lexicon through Support Vector Machine classification with the application of the Synthetic Minority Over-sampling Technique. A quantitative experimental approach was applied to 10,309 TikTok comments collected through Apify scraping. Since no human annotations were used as ground truth, the positive, negative, and neutral labels produced by both methods were treated as pseudo-labels. The research stages included filtering, preprocessing, pseudo-label generation, Term Frequency-Inverse Document Frequency feature extraction, an 80:20 data split, SMOTE application to the training data, SVM classification, and evaluation using accuracy, precision, recall, and F1-score. The results show that SVM reproduced the InSet Lexicon pseudo-labels most effectively, achieving an accuracy of 0.865 without SMOTE. After SMOTE was applied, precision increased to 0.870 while the F1-score remained at 0.865, and neutral-class recall increased from 0.767 to 0.807. These findings indicate that InSet Lexicon produced pseudo-labels that were more consistently learned by SVM on this dataset, while SMOTE primarily improved minority-class recognition.

Keywords: Sentiment Analysis; Pseudo-Labeling; IndoRoBERTa; InSet Lexicon; SVM; SMOTE

1. PENDAHULUAN

Perkembangan teknologi informasi mendorong media sosial menjadi sumber data opini publik yang besar, cepat, dan dinamis. TikTok tidak hanya digunakan sebagai media hiburan, tetapi juga sebagai ruang komunikasi digital tempat pengguna menyampaikan komentar, kritik, dukungan, dan respons terhadap isu sosial maupun kebijakan publik. Fitur komentar pada TikTok membentuk kumpulan data teks yang dapat merepresentasikan kecenderungan opini masyarakat secara real-time. Dalam konteks penelitian berbasis komputasi, data tersebut dapat diolah melalui analisis sentimen untuk mengelompokkan komentar ke dalam polaritas positif, negatif, atau netral. Analisis sentimen menjadi relevan karena mampu mengubah teks tidak terstruktur menjadi informasi yang lebih terukur dan dapat digunakan untuk memahami persepsi publik terhadap suatu peristiwa (Indriyani et al., 2023a; Manalu et al., 2025a).

Salah satu isu yang banyak memperoleh tanggapan pengguna TikTok adalah kasus kecelakaan mobil program Makan Bergizi Gratis (MBG) yang menabrak siswa SDN Kalibaru 01. Peristiwa tersebut memunculkan komentar yang beragam, mulai dari kritik terhadap pelaksana program, simpati kepada korban, hingga dukungan terhadap keberlanjutan program. Keragaman respons tersebut menunjukkan bahwa komentar TikTok memiliki potensi besar sebagai sumber data opini publik, tetapi juga memiliki tantangan karena teks media sosial umumnya singkat, informal, mengandung slang, emoji, singkatan, variasi ejaan, dan ekspresi emosional. Karakteristik ini dapat menurunkan kualitas pemrosesan teks apabila tidak ditangani melalui tahapan preprocessing yang tepat (Istiqomah et al., 2025; Syafaah & Haryanto, 2024).

Penelitian analisis sentimen terhadap data TikTok telah banyak dilakukan dengan pendekatan machine learning dan deep learning. Indriyani et al. (2023) membandingkan Naive Bayes dan Support Vector Machine pada analisis sentimen aplikasi TikTok dan menunjukkan bahwa algoritma SVM memiliki performa yang kompetitif untuk klasifikasi teks. Sidupa & Dewi (2025) juga menemukan bahwa Support Vector Classification mampu memberikan hasil yang baik pada review aplikasi TikTok. Pada isu yang lebih spesifik, Azril & Crisnawati (2026) menggunakan SVM untuk mengklasifikasikan sentimen komentar pengguna TikTok mengenai program MBG. Sementara itu,

TIN: Terapan Informatika Nusantara
Vol 7, No 1, June 2026, page 653-661
ISSN 2722-7987 (Media Online)
Website https://ejurnal.seminar-id.com/index.php/tin
DOI 10.47065/tin.v7i1.10243

Pateman et al. (2025) mengombinasikan SVM dan SMOTE untuk menangani data sentimen dari platform TikTok dan X, sehingga menunjukkan pentingnya penanganan distribusi kelas yang tidak seimbang dalam analisis opini media sosial.

Selain pemilihan algoritma klasifikasi, metode pelabelan sentimen juga menjadi faktor penting karena label menjadi dasar pembelajaran model. Pendekatan berbasis transformer seperti IndoBERT, IndoBERTweet, dan XLM-RoBERTa mampu memahami konteks kalimat secara lebih mendalam karena menggunakan representasi bahasa berbasis contextual embedding. Model semacam ini terbukti unggul pada berbagai tugas klasifikasi teks berbahasa Indonesia, terutama saat data memuat konteks yang kompleks dan makna implisit (Iansyah et al., 2025; Ramadhan et al., 2025; Wijaya et al., 2025). Namun, model transformer membutuhkan sumber daya komputasi yang lebih besar dan performanya tetap dipengaruhi oleh kualitas data serta kesesuaian domain pelatihan.

Di sisi lain, pendekatan lexicon-based seperti InSet Lexicon masih banyak digunakan karena sederhana, efisien, dan tidak memerlukan data latih dalam jumlah besar. InSet Lexicon menghitung kecenderungan sentimen berdasarkan kamus kata positif dan negatif yang memiliki bobot tertentu. Metode ini efektif pada teks yang mengandung opini eksplisit, meskipun memiliki keterbatasan dalam memahami negasi, sarkasme, dan perubahan makna akibat konteks kalimat. Beberapa penelitian menunjukkan bahwa pendekatan lexicon dapat dikombinasikan dengan machine learning untuk menghasilkan klasifikasi yang lebih terstruktur (Firdaus & Herdiani, 2021; Musfiroh et al., 2021; Rufaida et al., 2023).

Support Vector Machine dipilih dalam penelitian ini karena dikenal stabil untuk klasifikasi teks berdimensi tinggi, terutama pada data yang direpresentasikan menggunakan Term Frequency-Inverse Document Frequency. SVM bekerja dengan membangun hyperplane optimal yang memisahkan kelas sentimen berdasarkan margin terbesar. Keunggulan ini membuat SVM banyak digunakan dalam text mining dan analisis sentimen, termasuk pada data media sosial. Namun, SVM dapat mengalami bias ketika distribusi kelas tidak seimbang, karena model cenderung lebih mudah mengenali kelas mayoritas dibandingkan kelas minoritas (Andrianto et al., 2024; Junaedi et al., 2024; Prasetya et al., 2024).

Ketidakseimbangan data menjadi persoalan penting pada komentar media sosial karena satu kelas sentimen sering kali mendominasi kelas lainnya. Pada kasus viral, komentar negatif dapat jauh lebih banyak dibandingkan komentar positif atau netral, sehingga model dapat memperoleh accuracy tinggi tetapi gagal mengenali kelas minoritas. Untuk mengatasi persoalan tersebut, Synthetic Minority Over-sampling Technique atau SMOTE digunakan untuk membentuk data sintetis pada kelas minoritas. SMOTE memperluas ruang fitur kelas minoritas melalui interpolasi antar sampel sehingga distribusi data latih menjadi lebih seimbang dan model memiliki peluang lebih besar untuk mempelajari pola setiap kelas (Hairani et al., 2024; Kamalov et al., 2025; Sulistiyono et al., 2021).

Dalam penelitian ini, IndoRoBERTa tidak digunakan sebagai klasifikator akhir, tetapi sebagai salah satu metode pembentuk pseudo-label sentimen. Penempatan tersebut dilakukan agar IndoRoBERTa dan InSet Lexicon dapat dibandingkan dalam fungsi yang setara, yaitu menghasilkan label positif, negatif, dan netral secara otomatis. SVM kemudian digunakan sebagai klasifikator yang sama pada seluruh skenario pengujian untuk mengevaluasi sejauh mana pseudo-label dari masing-masing metode dapat dipelajari dan direproduksi secara konsisten. Penggunaan satu algoritma klasifikasi yang sama juga bertujuan mengendalikan pengaruh perbedaan model, sehingga variasi performa yang muncul lebih dapat dikaitkan dengan karakteristik pseudo-label yang dihasilkan oleh IndoRoBERTa dan InSet Lexicon. Dengan demikian, penelitian ini tidak dimaksudkan untuk membandingkan IndoRoBERTa sebagai model klasifikasi langsung dengan SVM, melainkan untuk mengevaluasi kualitas operasional pseudo-label dari dua pendekatan yang berbeda melalui kerangka klasifikasi yang seragam.

Berdasarkan uraian tersebut, terdapat celah penelitian dalam evaluasi pseudo-label sentimen yang dihasilkan IndoRoBERTa dan InSet Lexicon melalui klasifikasi SVM dengan dan tanpa SMOTE pada komentar TikTok berbahasa Indonesia terkait kasus MBG. Sebagian besar penelitian sebelumnya berfokus pada penggunaan satu metode pelabelan, satu algoritma klasifikasi, atau belum secara eksplisit membandingkan konsistensi pseudo-label berbasis transformer dan lexicon dalam kerangka klasifikasi yang sama. Oleh karena itu, penelitian ini bertujuan mengevaluasi empat skenario, yaitu IndoRoBERTa + SVM, IndoRoBERTa + SVM + SMOTE, InSet Lexicon + SVM, dan InSet Lexicon + SVM + SMOTE. Kontribusi penelitian ini adalah memberikan bukti empiris mengenai metode pseudo-labeling yang lebih konsisten dipelajari oleh SVM pada data komentar TikTok serta menunjukkan pengaruh SMOTE terhadap kemampuan model dalam mengenali kelas sentimen minoritas.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan desain eksperimen komparatif. Pendekatan kuantitatif digunakan karena proses analisis dilakukan melalui pengolahan data numerik dan pengukuran performa model secara objektif menggunakan metrik evaluasi. Desain komparatif diterapkan untuk membandingkan dua metode pelabelan sentimen, yaitu IndoRoBERTa dan InSet Lexicon, serta membandingkan pengaruh penerapan SMOTE terhadap klasifikasi menggunakan SVM. Objek penelitian adalah komentar pengguna TikTok yang membahas kasus kecelakaan mobil program MBG. Komentar dipilih karena memuat opini publik yang beragam dan memiliki karakteristik bahasa tidak terstruktur yang relevan untuk analisis sentimen berbasis Natural Language Processing.


TIN: Terapan Informatika Nusantara
Vol 7, No 1, June 2026, page 653-661
ISSN 2722-7987 (Media Online)
Website https://ejurnal.seminar-id.com/index.php/tin
DOI 10.47065/tin.v7i1.10243

Pengumpulan data dilakukan menggunakan teknik scraping komentar TikTok dengan bantuan Apify. Proses scraping dilakukan terhadap dua belas video TikTok yang berkaitan dengan kasus kecelakaan mobil MBG dan memiliki interaksi komentar tinggi. Data yang diambil berupa teks komentar yang kemudian diekspor ke format CSV agar dapat diproses menggunakan bahasa pemrograman Python. Dari proses scraping diperoleh 14.490 komentar awal. Dataset mentah kemudian melalui filtering untuk menghapus komentar kosong, komentar duplikat, URL, simbol, emoji berlebih, dan komentar yang tidak relevan. Setelah filtering dan preprocessing awal, dataset akhir yang digunakan dalam penelitian berjumlah 10.309 komentar. Dataset tersebut dibagi menjadi data training dan testing dengan rasio 80:20, sehingga diperoleh 8.247 data training dan 2.062 data testing.

Tahap pengolahan teks dibedakan sesuai dengan karakteristik metode pembentukan pseudo-label. Dataset terlebih dahulu melalui filtering untuk menghapus komentar kosong dan komentar duplikat. Pada pelabelan menggunakan IndoRoBERTa, model menerima teks asli dari kolom text agar struktur kalimat, urutan kata, dan informasi kontekstual tetap terjaga. Teks tersebut tidak melalui stemming maupun stopword removal sebelum dimasukkan ke model. Tokenisasi dilakukan menggunakan tokenizer bawaan model w11wo/indonesian-roberta-base-sentiment-classifier, dengan panjang masukan dibatasi hingga 512 karakter. Model kemudian menghasilkan pseudo-label positif, negatif, atau netral beserta nilai confidence untuk setiap komentar (Iansyah et al., 2025; Wijaya et al., 2025).

Sementara itu, pelabelan menggunakan InSet Lexicon memanfaatkan teks dari kolom clean_text yang telah melalui cleaning, case folding, normalisasi, tokenizing, dan stopword removal. Cleaning digunakan untuk menghapus URL, mention, simbol, angka, emoji, tanda baca, karakter non-ASCII, dan elemen lain yang tidak relevan. Case folding mengubah seluruh teks menjadi huruf kecil, sedangkan normalisasi digunakan untuk menyeragamkan kata tidak baku, singkatan, dan slang. Tokenizing memecah komentar menjadi satuan kata, kemudian stopword removal menghapus kata umum yang dinilai tidak memberikan informasi sentimen yang signifikan (Rizki et al., 2023; Suhaeni et al., 2025; Wardhani et al., 2024).Pada proses pelabelan InSet Lexicon, setiap komentar dipecah menjadi token berdasarkan spasi. Setiap token kemudian dicocokkan dengan kamus kata positif dan negatif. Skor sentimen dihitung dengan menjumlahkan bobot kata yang ditemukan pada kedua kamus. Komentar dengan skor lebih besar dari nol diberi pseudo-label positif, skor lebih kecil dari nol diberi pseudo-label negatif, sedangkan skor sama dengan nol diberi pseudo-label netral. Skor nol dapat menunjukkan bahwa tidak terdapat kata bermuatan sentimen dalam kamus atau terdapat keseimbangan antara skor positif dan negatif (Firdaus & Herdiani, 2021; Musfiroh et al., 2021; Rufaida et al., 2023).

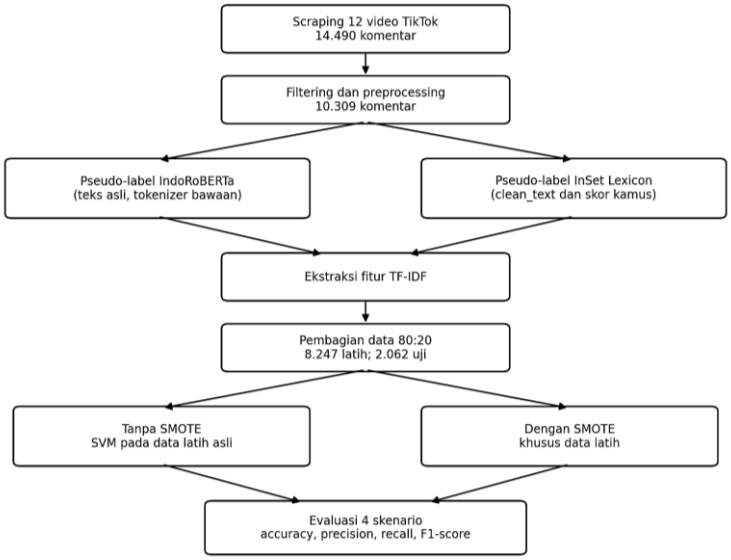
IndoRoBERTa dan InSet Lexicon digunakan secara terpisah sebagai pembentuk pseudo-label. Hasil pelabelan dari kedua metode tidak diperlakukan sebagai ground truth, tetapi sebagai target klasifikasi Support Vector Machine pada masing-masing skenario eksperimen. Oleh karena itu, metrik accuracy, precision, recall, dan F1-score digunakan untuk mengukur kemampuan SVM dalam mempelajari dan mereproduksi pseudo-label yang dihasilkan oleh setiap metode, bukan untuk menentukan kebenaran sentimen secara mutlak (Iansyah et al., 2025; Rufaida et al., 2023).

Penelitian ini tidak menggunakan anotasi manual oleh pakar atau anotator manusia sebagai standar pembanding. Ketiadaan ground truth manusia menjadi salah satu keterbatasan penelitian karena konsistensi pseudo-label secara komputasional belum secara langsung menunjukkan kesesuaian label dengan interpretasi manusia. Meskipun demikian, penggunaan algoritma SVM dan prosedur evaluasi yang sama pada seluruh skenario memungkinkan perbandingan dilakukan secara lebih terkontrol terhadap tingkat keterpelajaran pseudo-label IndoRoBERTa dan InSet Lexicon (Dinata et al., 2025; Helmiyah & Pramestiawan, 2025).

Data hasil pelabelan kemudian direpresentasikan dalam bentuk numerik menggunakan TF-IDF. Metode TF-IDF dipilih karena mampu memberikan bobot pada kata berdasarkan frekuensi kemunculan dalam dokumen dan tingkat kekhasan kata terhadap keseluruhan korpus. Representasi ini sesuai untuk algoritma SVM karena data teks umumnya berdimensi tinggi dan sparse. Setelah ekstraksi fitur, data dibagi menjadi training dan testing. Pada skenario tanpa SMOTE, data training langsung digunakan untuk melatih SVM. Pada skenario dengan SMOTE, proses oversampling hanya diterapkan pada data training untuk menghindari data leakage dan menjaga objektivitas evaluasi pada data testing.

Empat skenario eksperimen diterapkan dalam penelitian ini, yaitu SVM dengan pseudo-label IndoRoBERTa, SVM dengan pseudo-label IndoRoBERTa dan SMOTE, SVM dengan pseudo-label InSet Lexicon, serta SVM dengan pseudo-label InSet Lexicon dan SMOTE. Pada setiap skenario, teks hasil preprocessing untuk klasifikasi direpresentasikan menggunakan Term Frequency-Inverse Document Frequency dan dibagi menjadi data training serta testing dengan rasio 80:20. SMOTE hanya diterapkan pada data training setelah proses pembagian data untuk mencegah data leakage. Dengan menggunakan algoritma SVM dan prosedur evaluasi yang sama, perbedaan performa antar skenario dianalisis sebagai perbedaan tingkat konsistensi dan keterpelajaran pseudo-label, bukan sebagai bukti bahwa salah satu metode menghasilkan ground truth sentimen yang lebih benar (Hairani et al., 2024; Kamalov et al., 2025; Sulistiyono et al., 2021). Rangkaian tahapan penelitian yang digunakan, mulai dari pengumpulan data hingga evaluasi empat skenario model, ditunjukkan pada Gambar 1.

TIN: Terapan Informatika Nusantara
Vol 7, No 1, June 2026, page 653-661
ISSN 2722-7987 (Media Online)
Website https://ejurnal.seminar-id.com/index.php/tin
DOI 10.47065/tin.v7i1.10243



Gambar 1. Tahapan Penelitian

Komposisi dataset pada Tabel 1 menunjukkan bahwa proses filtering mengurangi data dari 14.490 komentar menjadi 10.309 komentar yang layak digunakan. Pengurangan ini dilakukan untuk memastikan bahwa data yang masuk ke tahap pelabelan dan klasifikasi tidak didominasi oleh noise, duplikasi, atau komentar yang tidak relevan terhadap kasus penelitian.

Tabel 1. Komposisi Dataset Penelitian

Tahap Dataset	Jumlah Data
Data hasil scraping awal	14.490
Dataset setelah filtering dan preprocessing	10.309
Data training	8.247
Data testing	2.062

Tabel 2. Distribusi Data Training Sebelum dan Sesudah SMOTE

Sentimen	Sebelum SMOTE	Sesudah SMOTE
Negatif	5.510	5.510
Positif	914	5.510
Netral	1.823	5.510

Tabel 2 memperlihatkan bahwa kelas negatif menjadi kelas mayoritas pada data training. Penerapan SMOTE dilakukan untuk membuat jumlah data positif dan netral setara dengan kelas negatif, sehingga model SVM tidak hanya mempelajari pola kelas mayoritas, tetapi juga memperoleh representasi yang cukup dari kelas minoritas.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Penelitian

Hasil penelitian diperoleh melalui rangkaian proses scraping, preprocessing, pelabelan sentimen, ekstraksi fitur TF-IDF, penerapan SMOTE, pelatihan SVM, dan evaluasi model. Dataset akhir berjumlah 10.309 komentar TikTok yang telah melalui filtering dan preprocessing. Komentar yang digunakan memuat beragam bentuk opini publik terhadap kasus kecelakaan mobil MBG. Sebagian komentar mengandung kritik langsung terhadap sopir dan pengelola program, sebagian berisi simpati terhadap korban, dan sebagian lain menanggapi program MBG secara umum. Karakteristik komentar yang informal, singkat, dan emosional memperlihatkan pentingnya preprocessing agar representasi fitur tidak terganggu oleh noise.

Pada tahap preprocessing, komentar yang semula memuat emoji, variasi ejaan, singkatan, dan kata tidak baku diubah menjadi bentuk yang lebih seragam. Misalnya komentar informal seperti “fix supirnya gak bisa nyupir” dinormalisasi menjadi bentuk yang lebih mudah dianalisis oleh model. Tahapan ini berpengaruh terhadap keberhasilan pelabelan dan klasifikasi karena kata-kata yang tidak distandarkan dapat mengurangi kemampuan sistem dalam mencocokkan kata pada lexicon maupun membentuk fitur TF-IDF yang konsisten. Preprocessing juga membantu mengurangi dimensi fitur yang tidak penting sehingga model dapat lebih fokus pada kata yang membawa informasi sentimen.

Hasil pelabelan menggunakan IndoRoBERTa dan InSet Lexicon menghasilkan label positif, negatif, dan netral yang kemudian dipakai sebagai target klasifikasi SVM. Perbedaan karakter kedua metode terlihat dari cara menentukan

TIN: Terapan Informatika Nusantara
Vol 7, No 1, June 2026, page 653-661
ISSN 2722-7987 (Media Online)
Website https://ejurnal.seminar-id.com/index.php/tin
DOI 10.47065/tin.v7i1.10243

polaritas. IndoRoBERTa menggunakan pendekatan kontekstual sehingga dapat mempertimbangkan susunan kata dan makna kalimat secara lebih luas. Sementara itu, InSet Lexicon mengandalkan skor polaritas kata. Pada komentar TikTok yang memuat opini eksplisit seperti kritik terhadap sopir, kata negatif yang kuat dapat dengan mudah dikenali oleh lexicon. Namun, pada komentar yang mengandung sarkasme, konteks tersembunyi, atau negasi kompleks, pendekatan lexicon dapat mengalami keterbatasan karena tidak selalu memahami relasi antar kata.

Rendahnya recall kelas netral pada skenario SVM dengan pseudo-label IndoRoBERTa perlu dilihat dalam konteks distribusi kelas dan karakteristik pseudo-label. Pada data testing, kelas negatif berjumlah 1.375 komentar, kelas positif 461 komentar, sedangkan kelas netral hanya 226 komentar. Ketimpangan tersebut menyebabkan model SVM lebih banyak mempelajari pola dari kelas negatif dan positif, sehingga batas keputusan cenderung berpihak pada kelas yang memiliki representasi lebih besar. Akibatnya, sebagian besar komentar yang memiliki pseudo-label netral diprediksi sebagai kelas lain dan recall kelas netral hanya mencapai 0,106. Selain faktor jumlah data, kelas netral juga memiliki karakteristik linguistik yang lebih sulit dipisahkan karena komentar netral dapat memuat kata yang secara leksikal tampak positif atau negatif, tetapi secara keseluruhan hanya menyampaikan informasi, pertanyaan, atau pernyataan tanpa orientasi sentimen yang kuat. Dalam representasi TF-IDF, kemiripan kosakata tersebut dapat menimbulkan tumpang tindih fitur antara kelas netral, positif, dan negatif.

Penerapan SMOTE pada pseudo-label IndoRoBERTa meningkatkan recall kelas netral dari 0,106 menjadi 0,544 dan F1-score dari 0,179 menjadi 0,445. Peningkatan tersebut menunjukkan bahwa penambahan sampel sintetis pada kelas minoritas membantu SVM membentuk batas keputusan yang lebih sensitif terhadap pola kelas netral. Namun, peningkatan kemampuan mengenali kelas minoritas diikuti penurunan accuracy dari 0,789 menjadi 0,751. Penurunan ini dapat terjadi karena model yang sebelumnya lebih dominan memprediksi kelas mayoritas menjadi lebih seimbang setelah menerima data sintetis. Perubahan batas keputusan tersebut meningkatkan jumlah prediksi netral, tetapi pada saat yang sama dapat menyebabkan sebagian data negatif atau positif beralih diprediksi sebagai netral. Dengan demikian, SMOTE tidak selalu meningkatkan accuracy keseluruhan, tetapi dapat memperbaiki keseimbangan deteksi antar kelas.

Pada skenario InSet Lexicon + SVM, model memperoleh accuracy 0,865, precision 0,865, recall 0,865, dan F1-score 0,865. Hasil ini lebih tinggi dibandingkan dua skenario berbasis IndoRoBERTa. Pada kelas negatif, model memperoleh precision 0,902, recall 0,914, dan F1-score 0,908. Pada kelas netral, precision mencapai 0,743, recall 0,767, dan F1-score 0,755. Pada kelas positif, precision 0,850, recall 0,809, dan F1-score 0,829. Nilai yang relatif seimbang pada tiga kelas menunjukkan bahwa label yang dihasilkan InSet Lexicon lebih konsisten untuk dataset komentar TikTok yang digunakan dalam penelitian ini, terutama karena banyak komentar berisi kata-kata opini eksplisit yang dapat dikenali dengan pendekatan kamus.

Skenario InSet Lexicon + SVM + SMOTE memperoleh accuracy 0,863, precision 0,870, recall 0,863, dan F1-score 0,865. Accuracy sedikit lebih rendah dibandingkan InSet Lexicon + SVM tanpa SMOTE, tetapi precision meningkat dari 0,865 menjadi 0,870. Pada kelas negatif, precision meningkat menjadi 0,933, sedangkan recall turun menjadi 0,887. Pada kelas netral, recall meningkat menjadi 0,807, meskipun precision turun menjadi 0,698. Pada kelas positif, recall meningkat menjadi 0,841 dan F1-score menjadi 0,830. Hasil ini menunjukkan bahwa SMOTE pada skenario InSet Lexicon membantu meningkatkan sensitivitas model terhadap kelas minoritas, khususnya netral dan positif, meskipun tidak selalu meningkatkan accuracy secara keseluruhan.

Rincian hasil evaluasi skenario IndoRoBERTa + SVM tanpa SMOTE disajikan pada Tabel 3 untuk menunjukkan kemampuan model dalam mereproduksi pseudo-label pada setiap kelas sentimen. Tabel tersebut memuat nilai precision, recall, F1-score, dan support untuk kelas negatif, netral, dan positif, serta nilai macro average dan weighted average. Secara umum, model memperoleh accuracy sebesar 0,789, tetapi performanya belum merata antarkelas karena kelas negatif memiliki recall tinggi sebesar 0,948, sedangkan kelas netral hanya mencapai recall 0,106. Perbedaan ini memperlihatkan bahwa ketidakseimbangan distribusi pseudo-label masih memengaruhi kemampuan SVM dalam mengenali kelas minoritas, sehingga hasil per kelas perlu dianalisis bersama dengan metrik agregat.

Tabel 3. Hasil Evaluasi IndoRoBERTa + SVM

Kelas Sentimen	Precision	Recall	F1-Score	Support
Negatif	0,799	0,948	0,867	1.375
Netral	0,571	0,106	0,179	226
Positif	0,769	0,651	0,705	461
Accuracy	0,789	0,789	0,789	2.062
Macro Average	0,713	0,568	0,584	2.062
Weighted Average	0,772	0,789	0,765	2.062

Perubahan performa setelah SMOTE diterapkan pada data latih dengan pseudo-label IndoRoBERTa ditampilkan pada Tabel 4. Penyajian hasil per kelas diperlukan untuk melihat apakah penyeimbangan data hanya mengubah accuracy keseluruhan atau juga memperbaiki sensitivitas model terhadap kelas minoritas. Setelah SMOTE, recall kelas netral meningkat dari 0,106 menjadi 0,544 dan F1-score kelas netral meningkat dari 0,179 menjadi 0,445, meskipun accuracy keseluruhan turun dari 0,789 menjadi 0,751. Tabel 4 dengan demikian memperlihatkan adanya trade-off antara

TIN: Terapan Informatika Nusantara
Vol 7, No 1, June 2026, page 653-661
ISSN 2722-7987 (Media Online)
Website https://ejurnal.seminar-id.com/index.php/tin
DOI 10.47065/tin.v7i1.10243

penurunan performa agregat dan peningkatan kemampuan model dalam mendeteksi pseudo-label netral serta positif secara lebih seimbang.

Tabel 4. Hasil Evaluasi IndoRoBERTa + SVM + SMOTE

Kelas Sentimen	Precision	Recall	F1-Score	Support
Negatif	0,870	0,796	0,831	1.375
Netral	0,376	0,544	0,445	226
Positif	0,695	0,718	0,707	461
Accuracy	0,751	0,751	0,751	2.062
Macro Average	0,647	0,686	0,661	2.062
Weighted Average	0,777	0,751	0,761	2.062

Hasil evaluasi SVM dengan pseudo-label InSet Lexicon tanpa SMOTE dirangkum pada Tabel 5. Tabel ini menyajikan performa tiap kelas untuk menilai konsistensi pseudo-label berbasis kamus ketika dipelajari oleh SVM melalui representasi TF-IDF. Model memperoleh accuracy, precision, recall, dan F1-score weighted average yang sama-sama mencapai 0,865, dengan F1-score kelas negatif sebesar 0,908, kelas netral sebesar 0,755, dan kelas positif sebesar 0,829. Sebaran nilai yang relatif lebih seimbang tersebut menunjukkan bahwa pada dataset penelitian ini, pseudo-label InSet Lexicon lebih mudah dipisahkan oleh SVM dibandingkan pseudo-label IndoRoBERTa tanpa SMOTE.

Tabel 5. Hasil Evaluasi InSet Lexicon + SVM

Kelas Sentimen	Precision	Recall	F1-Score	Support
Negatif	0,902	0,914	0,908	1.217
Netral	0,743	0,767	0,755	305
Positif	0,850	0,809	0,829	540
Accuracy	0,865	0,865	0,865	2.062
Macro Average	0,832	0,830	0,831	2.062
Weighted Average	0,865	0,865	0,865	2.062

Rincian performa SVM dengan pseudo-label InSet Lexicon setelah penerapan SMOTE ditunjukkan pada Tabel 6. Tabel ini digunakan untuk menilai perubahan kemampuan model terhadap kelas minoritas setelah jumlah data training positif dan netral diseimbangkan dengan kelas negatif. Walaupun accuracy turun tipis dari 0,865 menjadi 0,863, recall kelas netral meningkat dari 0,767 menjadi 0,807 dan recall kelas positif meningkat dari 0,809 menjadi 0,841, sedangkan precision weighted average naik menjadi 0,870. Hasil tersebut memperlihatkan bahwa SMOTE memberikan manfaat terutama pada peningkatan sensitivitas kelas minoritas, meskipun tidak menghasilkan peningkatan accuracy secara keseluruhan.

Tabel 6. Hasil Evaluasi InSet Lexicon + SVM + SMOTE

Kelas Sentimen	Precision	Recall	F1-Score	Support
Negatif	0,933	0,887	0,909	1.217
Netral	0,698	0,807	0,749	305
Positif	0,820	0,841	0,830	540
Accuracy	0,863	0,863	0,863	2.062
Macro Average	0,817	0,845	0,829	2.062
Weighted Average	0,870	0,863	0,865	2.062

Perbandingan metrik utama dari keempat skenario eksperimen secara keseluruhan disajikan pada Tabel 7. Tabel ini mempertemukan hasil IndoRoBERTa dan InSet Lexicon, masing-masing dengan dan tanpa SMOTE, sehingga pengaruh sumber pseudo-label dan teknik penyeimbangan data dapat diamati dalam kerangka klasifikasi SVM yang sama. Skenario InSet Lexicon + SVM tanpa SMOTE menghasilkan accuracy tertinggi sebesar 0,865, sedangkan InSet Lexicon + SVM + SMOTE menghasilkan precision tertinggi sebesar 0,870 dengan F1-score yang tetap sebesar 0,865. Oleh sebab itu, Tabel 7 menjadi dasar untuk menentukan bahwa tidak ada satu skenario yang unggul mutlak pada seluruh metrik; pemilihan skenario perlu disesuaikan dengan prioritas antara accuracy keseluruhan dan kemampuan mengenali kelas minoritas.

Tabel 7. Perbandingan Performa Seluruh Model

Model	Accuracy	Precision	Recall	F1-Score
IndoRoBERTa + SVM	0,789	0,772	0,789	0,765
IndoRoBERTa + SVM + SMOTE	0,751	0,777	0,751	0,761
InSet Lexicon + SVM	0,865	0,865	0,865	0,865
InSet Lexicon + SVM + SMOTE	0,863	0,870	0,863	0,865

TIN: Terapan Informatika Nusantara
Vol 7, No 1, June 2026, page 653-661
ISSN 2722-7987 (Media Online)
Website https://ejurnal.seminar-id.com/index.php/tin
DOI 10.47065/tin.v7i1.10243

Berdasarkan perbandingan seluruh skenario, SVM dengan pseudo-label InSet Lexicon tanpa SMOTE memperoleh accuracy tertinggi sebesar 0,865, sedangkan skenario dengan SMOTE memperoleh precision tertinggi sebesar 0,870 dan F1-score yang sama, yaitu 0,865. Oleh karena itu, tidak terdapat satu skenario yang unggul secara mutlak pada seluruh metrik. Skenario tanpa SMOTE lebih unggul dari sisi accuracy, sedangkan skenario dengan SMOTE menunjukkan peningkatan recall pada kelas netral dan positif. Temuan ini menegaskan bahwa pemilihan skenario terbaik perlu disesuaikan dengan tujuan evaluasi. Apabila prioritas utama adalah proporsi prediksi benar secara keseluruhan, skenario tanpa SMOTE lebih sesuai. Namun, apabila prioritas penelitian adalah meningkatkan pengenalan kelas minoritas, skenario dengan SMOTE memberikan hasil yang lebih seimbang.

3.2 Pembahasan

Hasil ini berbeda dari asumsi umum bahwa model berbasis transformer selalu lebih unggul dibandingkan pendekatan lexicon. Pada dataset ini, banyak komentar TikTok memuat kata-kata opini yang eksplisit, seperti kritik terhadap sopir, kecaman terhadap pelaksana program, atau pernyataan simpati yang relatif mudah dikenali berdasarkan skor kata. Kondisi tersebut membuat InSet Lexicon mampu menghasilkan label yang lebih konsisten untuk klasifikasi SVM. Di sisi lain, IndoRoBERTa tetap memiliki keunggulan konseptual dalam memahami konteks, tetapi performa akhirnya sangat bergantung pada kesesuaian model dengan domain komentar TikTok, kualitas preprocessing, dan distribusi label yang dihasilkan. Temuan ini sejalan dengan pandangan bahwa pemilihan metode analisis sentimen perlu disesuaikan dengan karakteristik dataset, bukan hanya berdasarkan kompleksitas model (Darusman & Gata, 2025; Rufaida et al., 2023).

Fenomena serupa terlihat pada skenario pseudo-label InSet Lexicon. Setelah SMOTE, accuracy turun tipis dari 0,865 menjadi 0,863, tetapi recall kelas netral meningkat dari 0,767 menjadi 0,807 dan recall kelas positif meningkat dari 0,809 menjadi 0,841. Hasil ini menunjukkan bahwa dampak utama SMOTE dalam penelitian ini bukan peningkatan accuracy, melainkan peningkatan sensitivitas model terhadap kelas yang memiliki jumlah lebih sedikit. Oleh karena itu, efektivitas SMOTE lebih tepat dievaluasi menggunakan recall per kelas, macro average, dan F1-score, bukan hanya accuracy. Pada pseudo-label IndoRoBERTa, macro recall meningkat dari 0,568 menjadi 0,686 dan macro F1-score meningkat dari 0,584 menjadi 0,661, meskipun weighted accuracy mengalami penurunan. Perubahan tersebut menunjukkan adanya pertukaran antara performa kelas mayoritas dan peningkatan pengenalan kelas minoritas.

Bila dibandingkan dengan penelitian terdahulu, performa SVM pada penelitian ini masih konsisten dengan beberapa studi analisis sentimen TikTok. Indriyani et al. (2023) menunjukkan bahwa SVM kompetitif untuk klasifikasi sentimen aplikasi TikTok. Sidupa & Dewi (2025) juga memperoleh hasil yang baik dengan Support Vector Classification. Pada penelitian ini, SVM mampu menghasilkan performa tinggi ketika label yang digunakan cukup konsisten dan fitur TF-IDF mampu merepresentasikan pola kata sentimen. Selain itu, hasil penelitian ini memperkuat temuan Pateman et al. (2025) bahwa SMOTE dapat membantu model dalam menghadapi data yang tidak seimbang, meskipun peningkatan tidak selalu tampak pada accuracy.

Jika dibandingkan dengan penelitian terdahulu, temuan penelitian ini menunjukkan persamaan sekaligus perbedaan. Kinerja SVM yang tinggi pada pseudo-label InSet Lexicon mendukung hasil Indriyani et al. (2023) dan Sidupa dan Dewi (2025), yang menunjukkan bahwa SVM kompetitif dalam klasifikasi sentimen TikTok. Namun, penelitian ini memperluas temuan tersebut dengan memperlihatkan bahwa performa SVM juga sangat dipengaruhi oleh sumber pseudo-label yang digunakan. Hasil accuracy 0,865 pada InSet Lexicon + SVM lebih tinggi daripada skenario berbasis IndoRoBERTa dalam eksperimen ini, tetapi temuan tersebut tidak bertentangan secara langsung dengan Iansyah et al. (2025) dan Wijaya et al. (2025), yang melaporkan keunggulan model transformer pada tugas klasifikasi kontekstual, karena IndoRoBERTa pada penelitian ini tidak di-fine-tuning sebagai klasifikator akhir dan hanya digunakan untuk membentuk pseudo-label. Selanjutnya, peningkatan recall kelas minoritas setelah SMOTE sejalan dengan Pateman et al. (2025), Hairani et al. (2024), dan Sulistiyono et al. (2021), meskipun accuracy keseluruhan tidak meningkat. Dengan demikian, kontribusi penelitian ini terletak pada bukti bahwa metode pelabelan, karakteristik bahasa komentar, dan tujuan evaluasi perlu dipertimbangkan secara bersamaan; pendekatan leksikon dapat lebih mudah dipelajari oleh SVM pada opini eksplisit, sedangkan pendekatan transformer tetap berpotensi lebih kuat untuk konteks implisit apabila dilakukan penyesuaian domain dan validasi terhadap ground truth manusia.

Secara praktis, hasil penelitian ini menunjukkan bahwa pemilihan metode pelabelan perlu mempertimbangkan karakteristik teks dan tujuan analisis. Untuk komentar yang banyak memuat opini eksplisit, InSet Lexicon dapat menjadi alternatif yang efisien dan tetap memberikan performa tinggi ketika dikombinasikan dengan SVM. Untuk data yang mengandung konteks lebih kompleks, sarkasme, atau makna implisit, model berbasis transformer seperti IndoRoBERTa tetap berpotensi dikembangkan melalui fine-tuning yang lebih sesuai dengan domain media sosial. Penerapan SMOTE disarankan ketika distribusi sentimen tidak seimbang dan tujuan penelitian tidak hanya mencari accuracy tertinggi, tetapi juga meningkatkan kemampuan model dalam mengenali kelas minoritas.

Kebaruan penelitian ini terletak pada perbandingan langsung antara pelabelan berbasis IndoRoBERTa dan InSet Lexicon dalam klasifikasi SVM dengan dan tanpa SMOTE pada komentar TikTok terkait kasus MBG. Penelitian ini tidak hanya menunjukkan model dengan nilai accuracy terbaik, tetapi juga menampilkan perubahan performa pada masing-masing kelas setelah data diseimbangkan. Hasil tersebut memberikan kontribusi terhadap pengembangan analisis sentimen bahasa Indonesia, terutama untuk data TikTok yang memiliki karakter informal dan distribusi sentimen tidak merata. Dari seluruh pengujian, InSet Lexicon + SVM + SMOTE dapat direkomendasikan sebagai skenario yang stabil karena menghasilkan precision tertinggi, F1-score yang setara dengan model terbaik, dan

TIN: Terapan Informatika Nusantara
Vol 7, No 1, June 2026, page 653-661
ISSN 2722-7987 (Media Online)
Website https://ejurnal.seminar-id.com/index.php/tin
DOI 10.47065/tin.v7i1.10243

peningkatan deteksi pada kelas minoritas, meskipun accuracy sedikit lebih rendah dibandingkan InSet Lexicon + SVM tanpa SMOTE.

4. KESIMPULAN

Penelitian ini mengevaluasi pseudo-label sentimen yang dihasilkan IndoRoBERTa dan InSet Lexicon melalui klasifikasi Support Vector Machine dengan dan tanpa penerapan Synthetic Minority Over-sampling Technique pada 10.309 komentar TikTok terkait kasus kecelakaan mobil program Makan Bergizi Gratis. Hasil pengujian menunjukkan bahwa SVM lebih konsisten mempelajari pseudo-label InSet Lexicon pada dataset penelitian, dengan accuracy tertinggi sebesar 0,865 pada skenario tanpa SMOTE. Skenario InSet Lexicon dengan SMOTE menghasilkan precision tertinggi sebesar 0,870 dan F1-score 0,865, serta meningkatkan recall kelas netral dari 0,767 menjadi 0,807 dan kelas positif dari 0,809 menjadi 0,841. Pada pseudo-label IndoRoBERTa, SMOTE meningkatkan recall kelas netral dari 0,106 menjadi 0,544, tetapi menurunkan accuracy dari 0,789 menjadi 0,751. Temuan tersebut menunjukkan bahwa SMOTE lebih berperan dalam meningkatkan sensitivitas terhadap kelas minoritas daripada menaikkan accuracy secara keseluruhan. Hasil penelitian ini tidak membuktikan bahwa InSet Lexicon menghasilkan label sentimen yang lebih benar daripada IndoRoBERTa, karena kedua metode digunakan sebagai pembentuk pseudo-label dan tidak divalidasi menggunakan anotasi manusia sebagai ground truth. Oleh sebab itu, kesimpulan penelitian dibatasi pada tingkat konsistensi pseudo-label ketika dipelajari oleh SVM dalam dataset dan skenario eksperimen yang digunakan. Penelitian selanjutnya perlu melibatkan anotator manusia, mengukur kesepakatan antar-anotator, melakukan fine-tuning IndoRoBERTa pada domain komentar TikTok, serta menguji metode pelabelan dan klasifikasi pada isu serta dataset yang lebih beragam.

REFERENCES

Andrianto, F., Fadlil, A., & Riadi, I. (2024). Linear kernel optimization of Support Vector Machine algorithm on online marketplace sentiment analysis. *Komputasi: Jurnal Ilmiah Ilmu Komputer Dan Matematika*, *21*(1), 68–82. https://doi.org/10.33751/komputasi.v21i1.9266

Azril, M., & Crisnawati, G. (2026). Klasifikasi sentimen komentar pengguna TikTok mengenai program Makan Bergizi Gratis (MBG) dengan Support Vector Machine (SVM). *RIGGS: Journal of Artificial Intelligence and Digital Business*, *5*(1), 8859–8866. https://doi.org/10.31004/riggs.v5i1.7237

Darusman, D., & Gata, W. (2025). Perbandingan kinerja machine learning dan deep learning untuk analisis sentimen Fufufafa. *Information System for Educators and Professionals: Journal of Information System*, *10*(1), 1–12. https://doi.org/10.51211/isbi.v10i1.3333

Dinata, R. M., Marhaeni, M., Atmadja, K., Rayhana, E., Hadi, V., & Al Kaf, U. (2025). Analisis komprehensif kinerja model klasifikasi sentimen: Evaluasi lintas metrik pada dataset tweet film bahasa Indonesia. *Jurnal Rekayasa Informasi*, *14*(1), 38–47. https://repository.istn.ac.id/13543/

Firdaus, R., & Herdiani, A. (2021). Lexicon-based sentiment analysis of Indonesian language student feedback evaluation. *Jurnal Computer*, *6*(1), 1–12. https://doi.org/10.34818/indojc.2021.6.1.408

Hairani, H., Widiyaningtyas, T., & Prasetya, D. D. (2024). Addressing class imbalance of health data: A systematic literature review on modified Synthetic Minority Oversampling Technique (SMOTE) strategies. *International Journal of Informatics and Visualization*, *8*(3), 1310–1318. https://doi.org/10.62527/joiv.8.3.2283

Helmiyah, S., & Pramestiawan, R. (2025). Analisis komparatif algoritma machine learning dengan metrik akurasi, presisi, recall, dan F1-score pada dataset kacang kering. *Jurnal Ilmu Komputer Dan Teknologi*, *6*(3), 152–159. https://doi.org/10.35960/ikomti.v6i3.2031

Iansyah, K., Nurlaili, A. L., & Al Haromainy, M. M. (2025). Comparative analysis of IndoBERT, IndoBERTweet, and XLM-RoBERTa for detecting online gambling comments on YouTube. *Bit-Tech*, *8*(2), 2379–2390. https://doi.org/10.32877/bt.v8i2.3257

Indriyani, F. A., Fauzi, A., & Faisal, S. (2023a). Analisis sentimen aplikasi TikTok menggunakan algoritma Naive Bayes dan Support Vector Machine. *TEKNOSAINS: Jurnal Sains, Teknologi Dan Informatika*, *10*(2), 176–184. https://doi.org/10.37373/tekno.v10i2.419

Indriyani, F. A., Fauzi, A., & Faisal, S. (2023b). Analisis sentimen aplikasi TikTok menggunakan algoritma Naive Bayes dan Support Vector Machine. *TEKNOSAINS: Jurnal Sains, Teknologi Dan Informatika*, *10*(2), 176–184. https://doi.org/10.37373/tekno.v10i2.419

Istiqomah, A., Safaah, T. N., & Haryati, G. (2025). Variasi bahasa dalam komentar TikTok: Kajian sosiolinguistik digital. *J-CEKI: Jurnal Cendekia Ilmiah*, *4*(6), 979–985. https://doi.org/10.56799/jceki.v4i6.10418

Junaedi, Gunawan, A. H., Kuswanto, V., & Jonathan. (2024). Tinjauan Support Vector Machine dalam text-mining untuk analisis sentimen di sektor pariwisata. *Bit-Tech*, *7*(2), 323–330. https://doi.org/10.32877/bt.v7i2.1810

Kamalov, F., Choutri, S. E., & Atiya, A. F. (2025). Analytical formulation of Synthetic Minority Oversampling Technique (SMOTE) for imbalanced learning. *Gulf Journal of Mathematics*, *19*(1), 400–415. https://doi.org/10.56947/gjom.v19i1.2639

Manalu, P. D., Simanjuntak, M., & Umri, C. (2025). Implementasi algoritma klasifikasi untuk analisis sentimen media sosial TikTok tahun 2025. *Jurnal Teknik Informatika Dan Teknologi Informasi*, *5*(1), 488–504. https://doi.org/10.55606/jutiti.v5i1.5644

TIN: Terapan Informatika Nusantara
Vol 7, No 1, June 2026, page 653-661
ISSN 2722-7987 (Media Online)
Website https://ejurnal.seminar-id.com/index.php/tin
DOI 10.47065/tin.v7i1.10243

Musfiroh, D., Khaira, U., Utomo, P. E. P., & Suratno, T. (2021). Analisis sentimen terhadap perkuliahan daring di Indonesia dari Twitter dataset menggunakan InSet Lexicon. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *1*(1), 24–33. https://doi.org/10.57152/malcom.v1i1.20

Pateman, D., Prasetyo, T. F., & Sujadi, H. (2025). Sentiment analysis of government on TikTok and X platforms with SVM and SMOTE approach. *JITK: Jurnal Ilmu Pengetahuan Dan Teknologi Komputer*, *10*(4), 900–908. https://doi.org/10.33480/jitk.v10i4.6645

Prasetya, H., Situmorang, Z., & Rosnelly, R. (2024). SVM optimization with kernel function for sentiment analysis on social media Twitter (X) in AFC U23 Asian Cup case study. *Proceedings of the International Conference of Science Technology UISU*, 227–233. https://doi.org/10.30743/wjxmmr59

Ramadhan, C., Atina, V., & Permatasari, H. (2025). Analisis perbandingan model CNN dan IndoBERT dalam sentimen berita politik Indonesia. *Prosiding Seminar Nasional Teknologi Informasi Dan Bisnis 2025*, 110–118. https://doi.org/10.47701/v1r9ka69

Rizki, A. S., Aristi, N. M., Ridha, N., Zulfahri, A. F., & Wibowo, D. A. (2023). Implementation of the Indonesian language stemming algorithm in Twitter data preprocessing: Case study Twitter Wargabanua and Instakalsel. *Fidelity: Jurnal Teknik Elektro*, *5*(3), 175–183. https://doi.org/10.52005/fidelity.v5i3.170

Rufaida, A. S. R., Permanasari, A. E., & Setiawan, N. A. (2023). Lexicon-based sentiment analysis using InSet Dictionary: A systematic literature review. *Proceedings of the 5th International Conference on Applied Engineering (ICAE 2022)*. https://doi.org/10.4108/eai.5-10-2022.2327474

Sidupa, B. C., & Dewi, C. (2025). Sentimen analisis terhadap aplikasi TikTok menggunakan Support Vector Classification. *Jurnal Mnemonic*, *8*(1), 1–8. https://doi.org/10.36040/mnemonic.v8i1.12635

Suhaeni, C., Kamila, S. A., Fahira, F., Yusran, M., & Dito, G. A. (2025). Exploring a large language model on the ChatGPT platform for Indonesian text preprocessing tasks. *Indonesian Journal of Statistics and Its Applications*, *9*(1), 100–116. https://doi.org/10.29244/ijsa.v9i1p100-116

Sulistiyono, M., Pristyanto, Y., Adi, S., & Gumelar, G. (2021). Implementasi algoritma Synthetic Minority Over-Sampling Technique untuk menangani ketidakseimbangan kelas pada dataset klasifikasi. *Sistemasi*, *10*(2), 445–457. https://doi.org/10.32520/stmsi.v10i2.1303

Syafaah, L. A., & Haryanto, S. (2024). Slang semantic analysis on TikTok social media Generation Z. *Proceeding ISETH: International Summit on Science, Technology, and Humanity*, 476–484. https://doi.org/10.23917/iseth.3898

Wardhani, D., Astuti, R., & Saputra, D. D. (2024). Optimasi feature selection text mining: Stemming dan stopword untuk sentimen analisis aplikasi SatuSehat. *Innovative: Journal of Social Science Research*, *4*(1), 7537–7548. https://doi.org/10.31004/innovative.v4i1.8759

Wijaya, I. N. S. W., Seputra, K. A., & Dewi, N. P. N. P. (2025). Fine tunning model IndoBERT untuk analisis sentimen berita pariwisata Indonesia. *Jurnal Pendidikan Teknologi Dan Kejuruan*, *22*(2), 195–204. https://doi.org/10.23887/jptk-undiksha.v22i2.104056