

INTEGRASI PELABELAN INSET LEXICON DAN SVM UNTUK KLASIFIKASI SENTIMEN DOOMSCROLLING DI TIKTOK

Robet Dwi Setiawan¹, Tri Listyorini², Endang Supriyati³

^{1,2,3}Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muria Kudus, Kudus, Indonesia

e-mail: ¹202351145@std.umk.ac.id, ²trilistyorini@umk.ac.id,
³endang.supriyati@umk.ac.id

ABSTRAK

Penelitian ini dilatarbelakangi oleh meningkatnya keluhan masyarakat tentang kebiasaan *doomscrolling* di sosial media, khususnya TikTok. Isu *doomscrolling* sering dikaitkan dengan dampak negatif pada kesehatan mental penggunanya. Opini publik tentang fenomena ini masih jarang dipetakan secara sistematis, khususnya pada data komentar TikTok yang bersifat informal dan tidak terstruktur. Penelitian ini bertujuan menganalisis sentimen fenomena *doomscrolling* melalui komentar TikTok dengan pendekatan berbasis *leksikon*, data dikumpulkan dengan mengambil video TikTok dengan kata pencarian “*doomscrolling*” dan “kecanduan TikTok” kemudian diolah menggunakan teknik *scraping*, lalu dilakukan pra-pemrosesan teks, pelabelan sentimen otomatis menggunakan *InSet Lexicon*, ekstraksi fitur dengan *Term Frequency Inverse Document Frequency* (TF-IDF), dan klasifikasi menggunakan algoritma *Support Vector Machine* (SVM). Hasil penelitian ini menunjukkan bahwa sentimen negatif mendominasi dengan persentase 70,72%, diikuti sentimen netral 16,57%, dan sentimen positif 12,71%. Model klasifikasi mencapai akurasi sebesar 75,76%, *precision* 76,66%, *recall* 75,76%, dan *F1-score* 76,07%. Hasil ini menggambarkan bahwa masyarakat cenderung memandang *doomscrolling* sebagai kebiasaan yang meresahkan, dan model yang dikembangkan cukup efektif dan dapat digunakan untuk memetakan opini publik terhadap fenomena *doomscrolling* di TikTok.

Kata kunci: doomscrolling; analisis sentimen; *inset lexicon*; support vector machine; tiktok.

ABSTRACT

This study was motivated by rising public complaints regarding the habit of doomscrolling on social media, particularly TikTok, a practice often linked to negative impact on users mental health. Public opinion on this phenomenon has yet to be systematically mapped, especially given the informal and unstructured nature of TikTok comments. This research aims to analyze sentiment regarding the doomscrolling phenomenon through TikTok comments using a lexicon based approach. Data were collected by retrieving TikTok videos using the search terms “doomscrolling” and “TikTok addiction” via scraping techniques, followed by text preprocessing, automated sentiment labeling using the InSet Lexicon, feature extraction via Term Frequency Inverse Document Frequency (TF-IDF), and classification using the Support Vector Machine (SVM) algorithm. The results indicate that negative sentiment dominates at 70.72%, followed by neutral sentiment 16.57%, and positive sentiment at 12.71%. The classification model achieved an accuracy of 75.76%, a precision of 76.66%, a recall of 75.76%, and an F1-score of 76.07%. These findings suggest that public generally perceives doomscrolling as a

troubling habit, and the developed model proves effective for mapping public opinion on the doomscrolling phenomenon on TikTok.

Keywords: *doomscrolling; sentiment analysis; inset lexicon; support vector machine; tiktok.*

1. PENDAHULUAN

Perkembangan media sosial mengubah pola konsumsi terhadap informasi secara signifikan. TikTok mengalami pertumbuhan pengguna sangat pesat dalam beberapa tahun terakhir, bukan hanya sarana hiburan, tetapi juga sebagai ruang berbagai isu, serta berkaitan dengan kesehatan mental dan kebiasaan digital [1]. Komentar pengguna dalam video TikTok telah dimanfaatkan sebagai sumber data penelitian analisis pada beberapa topik, penelitian yang menganalisis sentimen komentar TikTok salah satunya mengenai Ibu Kota Nusantara (IKN) [2]. Selain topik kebijakan publik, komentar pengguna juga mencerminkan pola perilaku digital penggunanya sendiri. Fenomena yang sekarang sedang dibicarakan pengguna sosial media adalah *doomscrolling*, yaitu dorongan individu untuk menggali linimasa (*scrolling*) media sosial dengan waktu dan pikiran yang tidak terkendali [3]. Kondisi ini sangat berdampak pada kesehatan mental pengguna, seperti meningkatnya kecemasan hingga gangguan tidur [4].

Beberapa penelitian sebelumnya telah menerapkan analisis sentimen pada komentar TikTok, meskipun belum ada yang spesifik untuk fenomena *doomscrolling*. Salah satu penelitian menganalisis sentimen terhadap kesehatan mental di TikTok menunjukkan bahwa pendekatan analisis sentimen dapat memetakan kecenderungan opini pengguna terhadap isu psikologis [5]. Pada sisi metode, algoritma *Naive Bayes* dan *Support Vector Machine* (SVM) pernah dibandingkan dalam mengklasifikasikan sentimen ulasan pengguna TikTok, dengan SVM lebih unggul dalam segi akurasi [6]. Kemampuan SVM dalam menangani topik sosial yang sensitif turut dibuktikan pada penelitian yang menerapkan algoritma untuk mendeteksi sentimen komentar TikTok terkait kekerasan seksual [7]. Selain itu, kombinasi *Term Frequency Inverse Document Frequency* (TF-IDF) dan SVM juga digunakan dalam menganalisis ulasan pengguna secara umum, serta dibandingkan dengan *Naive Bayes* pada komentar di ranah produk konsumen [8][9].

Meskipun SVM sering dipakai dalam menganalisis sentimen komentar TikTok, tetapi ketika masuk dalam ranah perilaku digital secara spesifik, cakupan penelitiannya jauh lebih sempit. Salah satu kasus memakai SVM dengan *kernel RBF* berdasarkan *screen time* dan interaksi sosial digital untuk mengelompokkan kondisi kesehatan mental pengguna ke dalam tiga kategori yaitu, sehat, stres, dan berisiko [10]. Persoalannya, penelitian ini memakai data sekunder dari Kaggle, bukan dari komentar publik yang ditulis langsung pengguna TikTok. Hal serupa terjadi pada studi lain yang menerapkan SVM untuk memetakan dampak psikologis remaja terhadap akibat paparan konten media sosial secara multi label. Pendekatannya cukup detail, namun bertumpu pada data *self report kuesioner*, dan penelitiannya sendiri mengakui bahwa ketidakseimbangan data menjadi kendala utama yang membatasi performa model [11]. Pada studi lain, evaluasi SVM bersama sepuluh algoritma *machine learning* lainnya untuk klasifikasi kanker payudara menunjukkan bahwa akurasi tinggi tetap bisa dipertahankan meski jumlah fitur dipangkas separuh lewat seleksi fitur, tetapi penelitian ini murni berfokus pada data medis terstruktur [12]. Di ranah aplikasi digital, penerapan SVM untuk mengklasifikasikan sentimen ulasan pengguna sebuah aplikasi perpustakaan digital justru menunjukkan

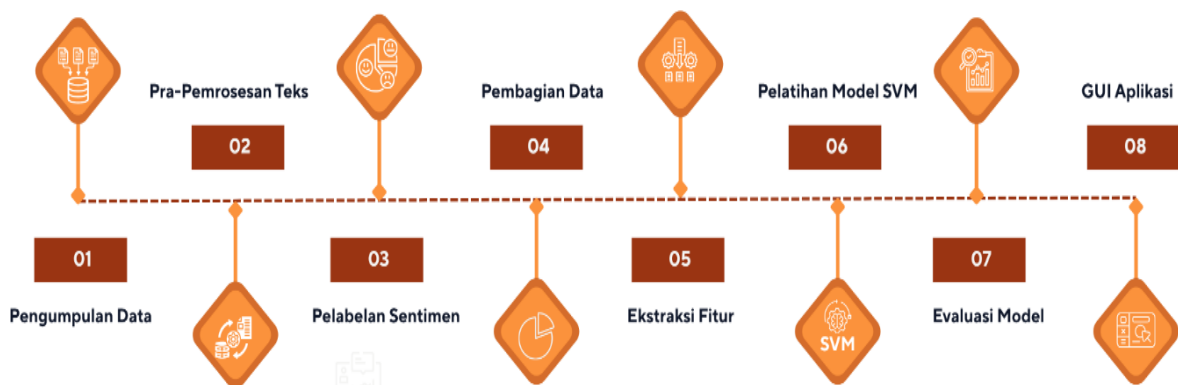
lompatan performa besar setelah ketidakseimbangan kelas ditangani lewat *SMOTE*, namun fokusnya sebatas evaluasi kualitas layanan aplikasi, bukan pola perilaku adiktif seperti *doomscrolling* [13]. SVM terbukti bekerja dengan baik pada berbagai jenis klasifikasi, tetapi belum ada yang menyatukan komentar TikTok, analisis sentimen berbasis SVM, dan fenomena *doomscrolling*.

Penelitian ini memiliki keunikan dibanding studi sebelumnya karena menerapkan algoritma SVM secara khusus pada dataset komentar TikTok yang membahas fenomena *doomscrolling*, sebuah topik yang belum banyak dipakai oleh penelitian analisis sentimen media sosial sebelumnya. Label sentimen yang diterapkan tidak dibuat manual melainkan melalui pendekatan berbasis *leksikon* menggunakan *InSet Lexicon*, sehingga pelabelan dapat digunakan dalam dataset besar. Selain itu, penelitian ini mengeksplorasi efektivitas kombinasi normalisasi kata tidak baku, *stopword*, *removal* yang diperkaya dengan daftar tambahan bahasa informal media sosial, serta *stemming* menggunakan pustaka Sastrawi untuk menghadapi karakter komentar TikTok yang cenderung singkat dan bercampur bahasa gaul. Pada akhirnya, setelah pengujian pada data uji, penelitian ini menghasilkan aplikasi antarmuka berbasis *streamlit* untuk memprediksi kelas dalam kalimat baru.

Berdasarkan uraian di atas, penelitian ini bertujuan untuk memberikan gambaran mengenai kecenderungan opini masyarakat terhadap kebiasaan *doomscrolling*, sekaligus menjadikan rujukan bagi pengembangan sentimen berbasis *machine learning* untuk isu kesehatan mental digital di TikTok maupun *platform* media sosial lainnya.

2. METODE PENELITIAN

Penelitian ini dilakukan melalui delapan tahapan utama yang saling berurutan sehingga dapat berjalan terstruktur dan konsisten [14]. Diawali dengan pengumpulan data kemudian proses scraping, dilanjutkan pra-pemrosesan data, pelabelan sentimen, kemudian proses pembagian data, dilanjutkan ekstraksi fitur dan evaluasi performa model, dan disempurnakan dengan aplikasi antarmuka. Untuk alur tahapan penelitian ditunjukkan Gambar 1.



Gambar 1. Tahapan Penelitian

Pengumpulan Data

Data penelitian didapat dari 30 video TikTok dengan kata pencarian “*doomscrolling*” dan “kecanduan TikTok”, kemudian diproses menggunakan teknik *scraping* dengan 2 atribut utama, isi komentar dan jumlah like [15]. Video yang diunggah pada tahun 2023-2026 dipilih secara manual dengan kriteria memiliki minimal 50

komentar. Pengambilan data dilakukan pada bulan Agustus-September 2026 menggunakan Google Colab dan API yang diambil dari apify.com. Selanjutnya komentar yang terlalu pendek serta duplikat dihapus agar dataset bersih dan siap ke tahap selanjutnya. Jumlah data setiap pembersihan ditunjukkan pada Tabel 1.

Tabel 1. Pembersihan Data

Keterangan	Jumlah
Jumlah Data awal	2899
Setelah Hapus Komentar Pendek	2176
Setelah Hapus Data Duplikat	2170

Pra-pemrosesan Teks

Teks mentah hasil Tabel 1, kemudian disempurnakan melalui tahapan berurutan sebagaimana lazim diterapkan pada sentimen media sosial [16], yaitu *cleaning*, *case folding*, normalisasi kata tidak baku, tokenisasi, *stopword removal*, dan *stemming*. Normalisasi kata tidak baku menggunakan repositori GitHub dengan link (<https://github.com/nasalsabila/kamus-alay>) sehingga bisa dikonversi ke bentuk formal secara konsisten. Proses *stemming* memakai pustaka Sastrawi, yang terbukti lebih baik dibanding Porter, Idris, Arifin, maupun Setiono [17]. Dataset akhir setelah tahap pra-pemrosesan dirangkum pada Tabel 2.

Tabel 2. Pembersihan Data Akhir

Keterangan	Jumlah
Data Awal Tabel 1	2170
Dataset akhir setelah Pra-pemrosesan	2124

Pelabelan Sentimen

Label sentimen penelitian ini dihasilkan melalui pendekatan berbasis *leksikon*, memanfaatkan *InSet Lexicon* dari GitHub ((<https://github.com/fajri91/InSet>), kamus sentimen yang dikembangkan untuk melabeli sentimen bahasa Indonesia secara otomatis [18].

Pembagian Data dan Ekstraksi Fitur

Dataset yang berlabel dibagi menjadi data latih dan data uji dengan rasio 80 banding 20 menggunakan *stratified sampling*, skema pembagian yang sering diterapkan pada klasifikasi sentimen berbasis *leksikon* untuk menjaga proporsi tiap kelas tetap seimbang [19].

Representasi teks diperoleh melalui TF-IDF yang menghitung seberapa sering kata muncul dalam satu komentar, lalu membandingkan dengan seluruh komentar, sehingga kata yang khas mendapat bobot yang besar [20]. Jumlah fitur pada penelitian ini dibatasi rentang tertentu agar model tidak terlalu kompleks akibat kosa kata yang jarang muncul.

$$TF - IDF(t, d) = TF(t, d) \times IDF(t) \quad (1)$$

Pemodelan dan Pelatihan

Proses klasifikasi dilakukan menggunakan algoritma SVM dengan implementasi *LinearSVC* dari pustaka *scikit-learn*, yang bekerja dalam mencari *hyperplane* pemisah terbaik kelas sentimen melalui *margin*, sebagaimana dijelaskan dalam kajian penerapan SVM pada klasifikasi teks berbahasa Indonesia [21].

$$\text{menimalkan } \frac{1}{2} \|w\|^2 \text{ dengan syarat } y_i(w \cdot x_i + b) \geq 1 \text{ untuk semua } i \quad (2)$$

Evaluasi Model

Performa model dievaluasi menggunakan metrik akurasi, presisi, *recall*, dan *F1-score* yang dihitung dari *confusion matrix* mengikuti rumus yang umum dipakai pada penelitian klasifikasi teks berbasis model pembelajaran mesin [22].

$$\text{Akurasi} = (TP + TN) \div (TP + TN + FP + FN) \quad (3)$$

$$\text{Presisi} = TP \div (TP + FP) \quad (4)$$

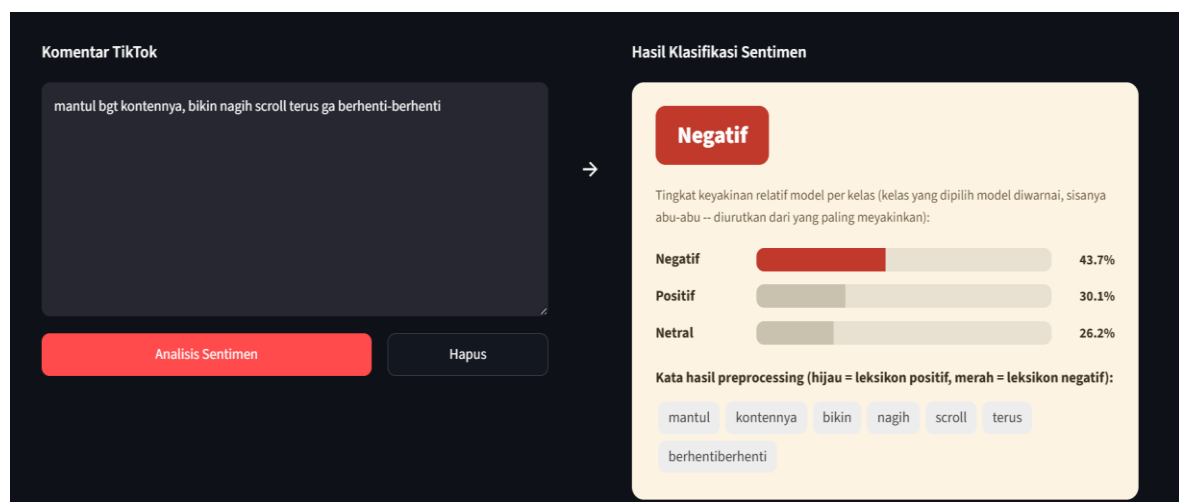
$$\text{Recall} = TP \div (TP + FN) \quad (5)$$

$$F1 \text{ score} = 2 \times (\text{presisi} \times \text{Recall}) \div (\text{presisi} + \text{Recall}) \quad (6)$$

Keterangan, TP diartikan jumlah prediksi benar pada kelas positif, TN merupakan hasil benar pada kelas negatif, FP adalah jumlah prediksi salah dianggap positif padahal sebenarnya negatif, dan FN ialah hasil dari prediksi salah yang dianggap negatif padahal sebenarnya positif.

Aplikasi Antarmuka

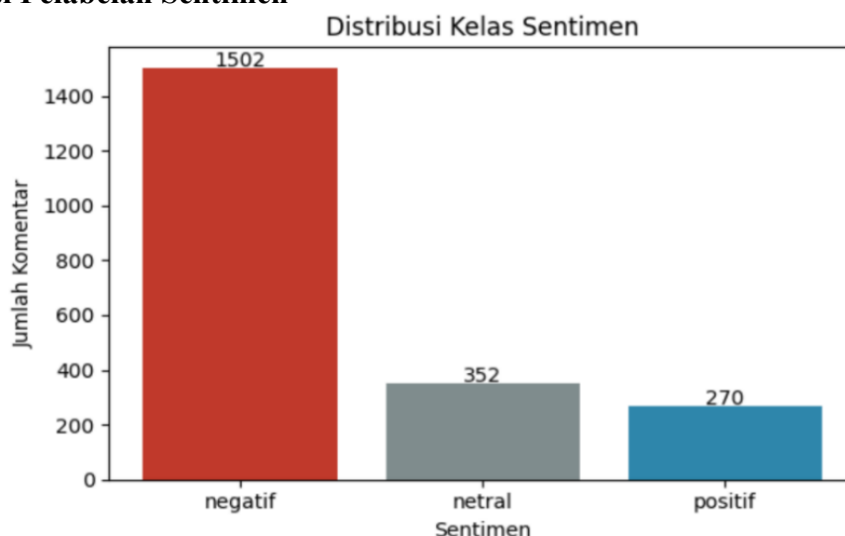
Model dan TF-IDF *vectorizer* yang telah dilatih kemudian diintegrasikan pada aplikasi antarmuka berbasis *streamlit*, sehingga prediksi sentimen terhadap kalimat baru dapat diuji langsung tanpa mengulang seluruh proses pelatihan dari awal.



Gambar 2 Aplikasi Antarmuka

3. HASIL DAN PEMBAHASAN

Visualisasi Pelabelan Sentimen

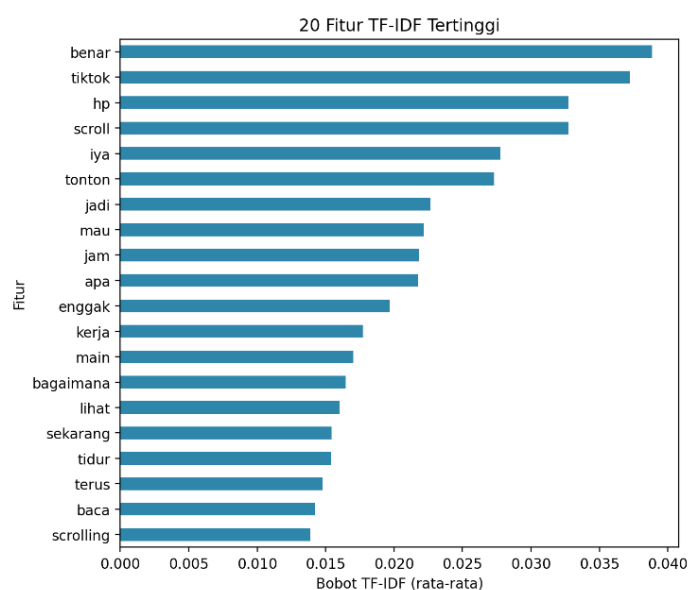


Gambar 3. Distribusi Kelas

Data dibagi menjadi 3 kelas dengan masing-masing kelas memiliki angka sentimen, yaitu kelas negatif mendominasi sentimen tertinggi dengan 1.502 komentar, diikuti kelas netral dengan 352 komentar, dan kelas positif yang paling rendah dengan 270 komentar.

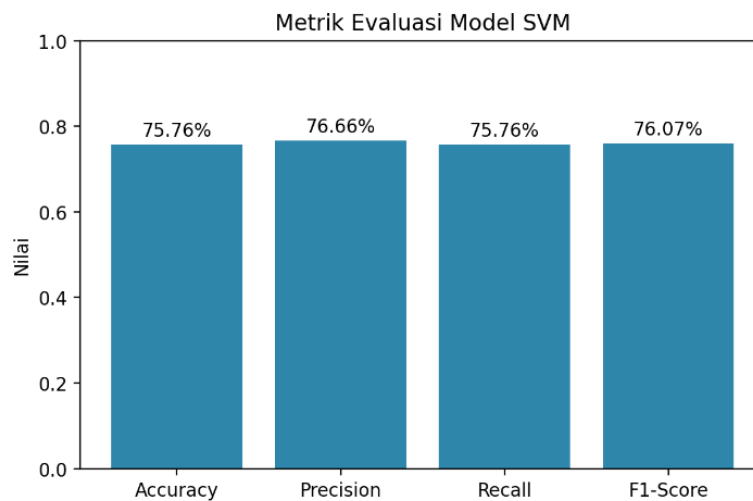
Ekstraksi Fitur

TF-IDF menghasilkan *matriks* berukuran 1.699 dari data latih dikali 220 *max features* yang dipilih dan 425 data uji dikali 220 *max features*, dengan dua puluh fitur berbobot tertinggi ditunjukkan pada Gambar 4.



Gambar 4. Fitur TF-IDF

Evaluasi Performa



Gambar 5. Evaluasi Model

Pada Gambar 5, SVM mencapai akurasi sebesar 75,76% sehingga model mampu memprediksi sentimen dengan benar sekitar tiga perempat dari total data uji. Nilai *precision* sebesar 76,66% dan *recall* sebesar 75,76% membuktikan bahwa keseimbangan model cukup baik diantara ketepatan prediksi dan kemampuan menangkap seluruh data pada tiap kelas, kemudian *F1-score* sebesar 76,07% menegaskan performa model tidak timpang ke salah satu sisi metrik.

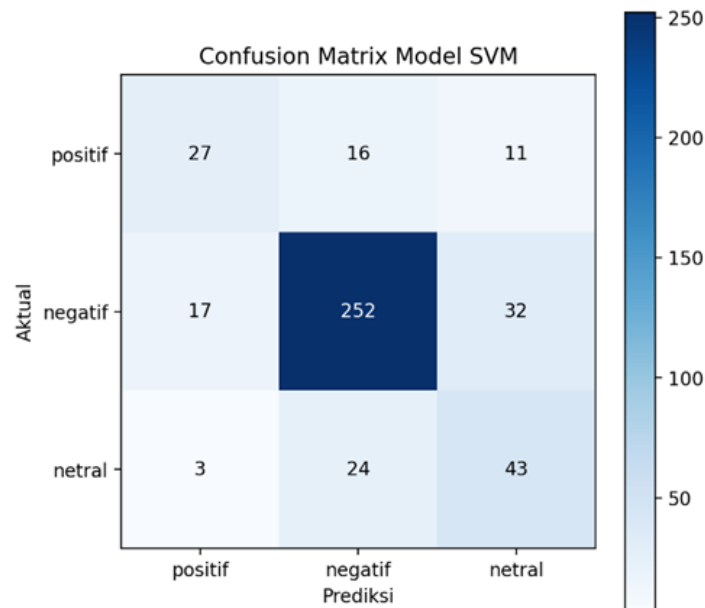
Nilai *precision* mendekati *accuracy* juga dapat diartikan bahwa model tidak terlalu bias terhadap kelas sentimen tertentu, namun performa ini tergolong sedang, ketimpangan jumlah data kelas negatif jauh mendominasi dari kelas netral dan positif menjadi faktor utama yang membatasi model mencapai akurasi yang lebih tinggi. Kondisi ini dapat dilihat jelas pada Tabel 3, kesenjangan performa antar kelas menjadi lebih nyata dibanding melihat angka akurasi keseluruhan.

Tabel 3. Classification Report

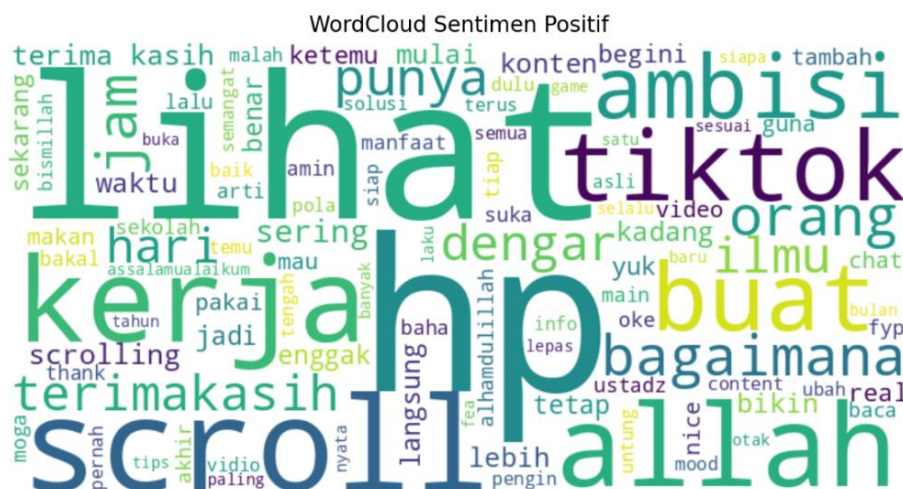
	Precision	Recall	F1-Score	Support
Negatif	0.8630	0.8372	0.8499	301
Netral	0.5000	0.6143	0.5513	70
Positif	0.5745	0.5000	0.5347	54
Accuracy			0.7576	425
Macro avg	0.6458	0.6505	0.6453	425
Weighted avg	0.7666	0.7576	0.7607	425

Berdasarkan Tabel 3, kelas negatif memiliki performa terbaik dengan *precision* 0,8630, *recall* 0,8372, dan *F1-score* 0,8499. Sebaliknya, kelas netral hanya mencapai *precision* 0,5000 dan *F1-score* 0,5513, sementara kelas positif mencapai *precision* 0,5745 dan *F1-score* 0,5347. Dengan hasil ini performa kelas positif dan kelas netral berada lebih rendah daripada kelas negatif, faktor ini dipengaruhi jumlah data yang jauh lebih sedikit dibanding kelas negatif.

Sebaran unt

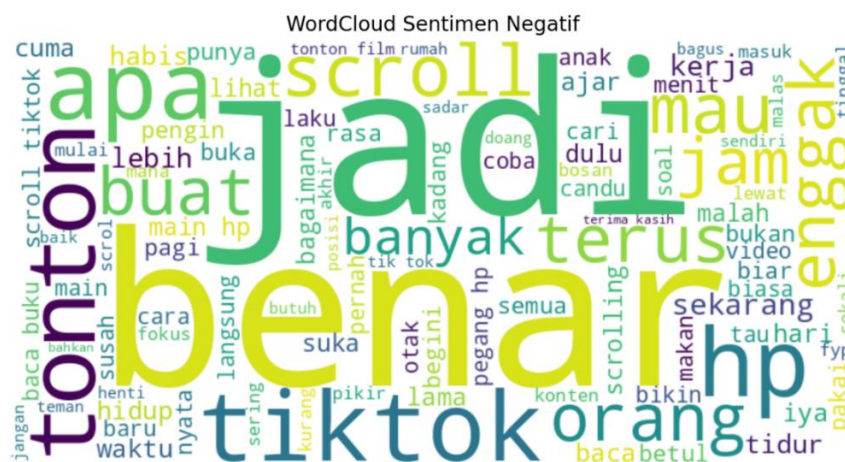


Visualisasi



[illegible]

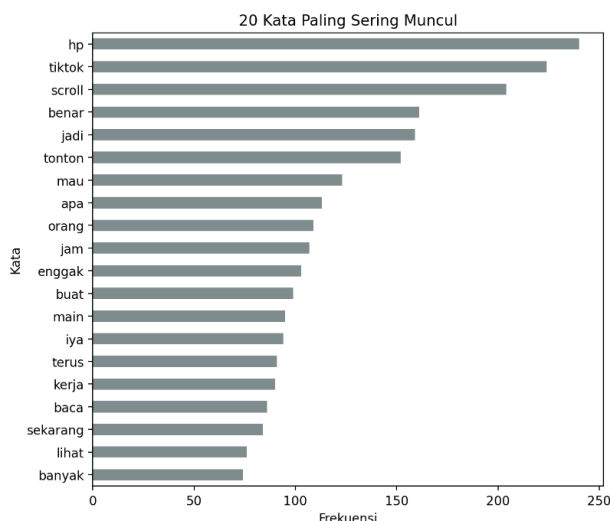
Berdasarkan Gambar 8, kata-kata bersentimen netral banyak memuat kata seperti “posisi”, “iya”, “jam”, “relate”, “scroll”, “tiap”. Sifat komentar pada kelas ini lebih banyak reaksi singkat, persetujuan sederhana, dan sekedar menyebut istilah *doomscrolling* tanpa menunjukkan sentimen yang tegas, sehingga wajar apabila model kesulitan membedakan dari kelas positif maupun negatif



Berdasarkan Gambar 9, kata-kata dominan pada komentar bersentimen negatif berpusat pada kata “benar”, “tonton”, “scroll” “toiktok” diikuti dengan kata keluhan

seperti “candu”, “bosan”, “malas”, “susah”, “tidur”. Kombinasi ini menggambarkan keresahan pengguna terhadap kebiasaan scrolling yang sulit dihentikan, sehingga berdampak pada waktu tidur, dan merasa malas beraktivitas lain, ini sesuai karakteristik *doomscrolling* sebagai kebiasaan yang dilakukan pengguna sendiri.

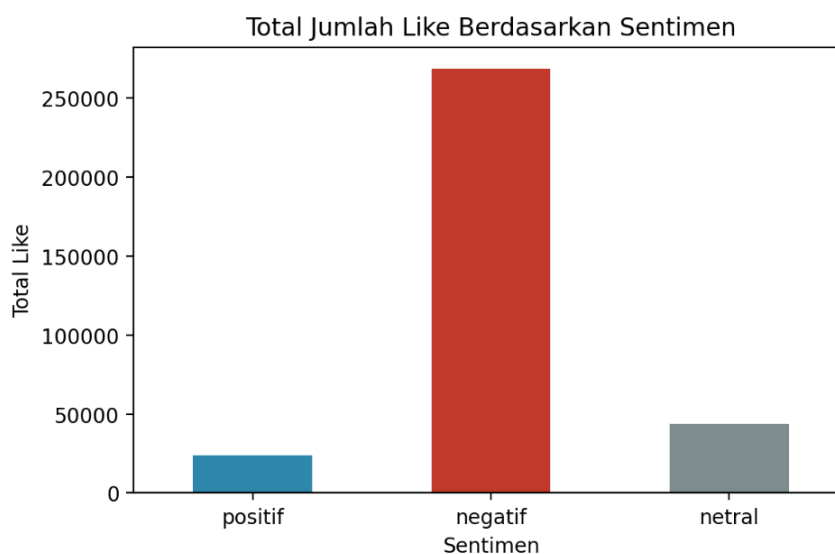
Sementara itu, dua puluh kata dengan frekuensi kemunculan tertinggi pada seluruh komentar dalam dataset penelitian ini ditunjukkan pada Gambar 10.



Gambar 10. Frekuensi Kata

Berdasarkan Gambar 10, kata “hp” menjadi yang paling sering muncul dengan frekuensi mendekati 240 kali, diikuti oleh “tiktok” dan “scroll” yang muncul lebih dari 200 kali. Kata lain seperti “benar”, “jadi” “tonton”, “jam” turut menempati posisi tertinggi. Dominasi kata-kata ini menegaskan bahwa pembahasan komentar memang terletak langsung pada aktivitas menggunakan ponsel dan kebiasaan *scrolling* di TikTok.

Adapun hasil analisis interaksi berdasarkan jumlah like pada setiap kelas sentimen digambarkan pada Gambar 11.



Gambar 11. Interaksi Like

Berdasarkan Gambar 11 komentar bersentimen negatif memperoleh total sekitar 268.779 *like*, sentimen netral sekitar 43.764 *like*, dan sentimen positif hanya sekitar 23.671 *like*. Tingginya interaksi pada sentimen negatif menunjukkan bahwa keluhan terkait kecanduan scrolling dan dampaknya banyak menarik perhatian dibanding komentar netral maupun positif.

Diskusi

Hasil penelitian ini menunjukkan kombinasi *InSet Lexicon* dan SVM dapat mengklasifikasikan sentimen komentar TikTok dengan isu *doomscrolling* yang mencapai akurasi 75,76%. Temuan ini sejalan dengan penelitian yang menerapkan pendekatan pada isu kebijakan publik lain, dimana *InSet Lexicon* dipadukan dengan SVM dan TF-IDF berhasil mendapatkan akurasi 71% pada klasifikasi sentimen kebijakan subsidi kendaraan listrik [18]. Lebih tingginya akurasi pada penelitian ini kemungkinan dipengaruhi oleh jumlah fitur TF-IDF yang dibatasi pada 220 kata.

Performa model jauh lebih baik pada kelas negatif dibanding kelas netral dan positif konsisten dengan tantangan umum penelitian klasifikasi teks berbasis *leksikon*, dimana ketidakseimbangan kelas membuat model cenderung mengenali kelas mayoritas [19]. Kondisi serupa juga ditemukan pada penelitian klasifikasi berbasis mesin lain, dimana performa model turun ketika kelas data uji tidak merata, meski proses pembagian data sudah dilakukan secara *stratified* [22].

Rendahnya precision kelas netral dan kelas positif mengindikasikan bahwa kedua kelas ini pada bahasa informal TikTok memang tidak selalu tegas. Persoalan ini umum terjadi pada analisis sentimen berbasis *leksikon*, pada proses pelabelan kata otomatis sangat bergantung pada skor setiap kata tanpa mempertimbangkan kalimat menyeluruh [18]. Dominasi sentimen negatif dari jumlah *like* turut memperkuat bahwa *doomscrolling* lebih banyak dikeluhkan sebagai kebiasaan yang meresahkan, sejalan dengan karakteristik komentar TikTok yang cenderung terbuka dan emosional [15].

4. KESIMPULAN

Integrasi *InSet Lexicon* dengan *Support Vector Machine* berbasis TF-IDF terbukti dapat memetakan opini publik terhadap *doomscrolling* di TikTok dengan akurasi 75,76%, dimana mayoritas warganet menilai kebiasaan ini secara kritis dan tingginya opini negatif, dan model penelitian dapat digunakan sebagai alat bantu pemetaan opini terhadap isu seperti ini. Namun, model tidak merata karena performa klasifikasi melemah pada kelas netral dan kelas positif karena timpangnya jumlah data latih, sehingga riset lanjutan disarankan untuk memperluas volume data pada kelas minoritas, menerapkan teknik penyeimbang, membuat penelitian lanjutan dengan *deep learning*, serta dapat memperluas data ke beberapa *platform* media sosial lainnya.

DAFTAR PUSTAKA

- [1] R. Tangke, D. Tineke Salaki, W. Widsli Kalengkongan, and E. Ketaren, "Analisis Sentimen Aplikasi Tiktok Menggunakan Algoritma Support Vector Machine (SVM) Dan Random Forest," *J. TIMES*, vol. 13, no. 2, pp. 53–62, 2024, doi: 10.51351/jtm.13.2.2024762.

-
- [2] Siti Rihastuti and Afnan Rosyidi, “Analisis Sentimen Pengguna Tiktok Tentang Progres Pembangunan Ikn Dengan Metode Random Forest,” *J. Comput. Sci. Technol.*, vol. 5, no. 1, pp. 19–23, 2025, doi: 10.54840/jcstech.v5i1.345.
 - [3] Sofiyyatul, T. Khoirunnisa, T. S. Wijayanti, Z. N. Atikoh, S. Nawawi, and F. Faruq, “EKSPLOKASI DOOMSCROLLING & SUBJECTIVE WELL-BEING PADA DEWASA AWAL PENGGUNA AKTIF MEDIA SOSIAL,” *J. PSIKOHUMANIKA*, vol. 17, no. 1, pp. 32–51, 2025, [Online]. Available: <http://ejurnal.setiabudi.ac.id/ojs/index.php/psikohumanika%0AEKSPLOKASI>
 - [4] N. Rohmah, N. Aminudin, A. Wantoro, D. Yana, and A. Andini, “Analisis Sentimen Kesehatan Mental di TikTok pada Generasi Milenial, Gen Z, dan Alpha Menggunakan SVM dan Random Forest,” *J. Rekayasa Perangkat Lunak*, vol. 04, no. 2, pp. 96–105, 2025, [Online]. Available: <https://journal.aisyahuniversity.ac.id/index.php/J-Rapa>
 - [5] V. A. Nugraha, “ANALISIS KONTEN KESEHATAN MENTAL PADA TIKTOK MENGGUNAKAN METODE SENTIMEN ANALISIS,” *PREPOTIF J. Kesehat. Masy.*, Vol. 9, Pp. 3003–3008, 2025.
 - [6] Friska Aditia Indriyani, Ahmad Fauzi, and Sutan Faisal, “Analisis sentimen aplikasi tiktok menggunakan algoritma naïve bayes dan support vector machine,” *TEKNOSAINS J. Sains, Teknol. dan Inform.*, vol. 10, no. 2, pp. 176–184, 2023, doi: 10.37373/tekno.v10i2.419.
 - [7] I. Wulandari, S. E. Prastya, M. Zulfadhilah, and R. Anshari, “Analisis Sentimen Terhadap Tindakan Kekerasan Seksual pada Media Sosial Tiktok Menggunakan Metode Support Vector Machine (SVM),” *J. Nas. Komputasi dan Teknol. Inf.*, vol. 8, no. 2, pp. 768–776, 2025.
 - [8] P. Husain and A. F. Syam, “Analisis Sentimen Ulasan Pengguna Tiktok pada Google Play Store Berbasis TF-IDF dan Support Vector Machine,” *J. Syst. Comput. Eng.*, vol. 5, no. 1, pp. 91–102, 2024.
 - [9] T. Setiawan, S. Liem, and M. Rizky, “Perbandingan Algoritma SVM dan Naïve Bayes dalam Analisis Sentimen Komentar Tiktok pada Produk Skincare,” *Appl. Inf. Technol. Comput. Sci.*, vol. 3, no. 2, pp. 28–32, 2024.
 - [10] Pendi and H. Sulistiani, “Klasifikasi Kesehatan Mental Menggunakan Support Vector Machine Berdasarkan Screen Time dan Interaksi Sosial Digital,” *Build. Informatics, Technol. Sci.*, vol. 7, no. 4, pp. 2312–2321, 2026, doi: 10.47065/bits.v7i4.9189.
 - [11] W. Febriani, Mambang, S. E. Prastya, B. Sabella, and F. D. Marleny, “Analysis of the Impact of Violent Content on Social Media on Adolescent Cyberpsychology Using Support Vector Machine and Random Forest,” *J. Appl. Informatics Comput.*, vol. 10, no. 1, pp. 707–711, 2026.
 - [12] S. Y. Putri, N. Ramadani, and N. A. Ginting, “EVALUASI RECURSIVE FEATURE ELIMINATION UNTUK KLASIFIKASI KANKER PAYUDARA MENGGUNAKAN,” *JIKSTRA*, vol. 8, no. 01, 2026.
 - [13] A. Mustafa, R. Kurniawan, B. Nurhakim, P. P. Marta, and K. Anam, “PENINGKATAN KINERJA SVM PADA KLASIFIKASI SENTIMEN ULASAN IPUSNAS BERBASIS IMBALANCE HANDLING,” *JIKSTRA*, vol. 8, no. 01, 2026.
 - [14] E. Supriyati, M. Iqbal, and T. Listyorini, “Social Networks in Forming Sentiment Analysis Climate Change Case Through Fuzzy Decision Tree

-
- Clustering,” *Int. Conf. Ser. Sci. Eng. Technol.*, vol. 3167, no. 1, 2024.
- [15] N. Apriliani *et al.*, “ANALISIS SENTIMEN REVIEW PENGGUNAAN TIKTOK MELALUI PENDEKATAN ALGORITMA NAÏVE BAYES,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 6, pp. 3725–3731, 2023.
- [16] M. Huda, E. Supriyati, and T. Listyorini, “ANALISIS SENTIMEN PENGGUNA YOUTUBE TERHADAP TAYANGAN #MATANAJWAMENANTITERAWAN DENGAN METODE NAÏVE BAYES CLASSIFIER,” *JiKO (Jurnal Inform. dan Komputer)*, vol. 5, no. 1, pp. 1–6, 2022, doi: 10.33387/jiko.
- [17] J. Pardede and D. Darmawan, “PERBANDINGAN ALGORITMA STEMMING PORTER , SASTRAWI , IDRIS , COMPARISON OF STEMMING ALGORITHMS PORTER , SASTRAWI , IDRIS , AND ARIFIN SETIONO PADA DOKUMEN TEKS BAHASA INDONESIA,” vol. 12, no. 1, 2025, doi: 10.25126/jtiik.2025128860.
- [18] D. Pratiwi, N. Khoerani, and S. Sari, “ANALISIS SENTIMEN TERHADAP KEBIJAKAN SUBSIDIPEMBELIAN KENDARAAN BERTENAGA LISTRIK DI INDONESIA MENGGUNAKAN PENDEKATAN INSET LEXICON DAN METODE SUPPORT VECTOR MACHINE,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 12, no. 6, pp. 1303–1314, 2025.
- [19] F. Olivian, S. Bariyah, G. C. Budiyanto, R. Annisa, L. A. Fitriana, and W. S. Dharmawan, “Performance Evaluation of Machine Learning Algorithms in Sentiment Analysis of Spotify Reviews,” *J. Artif. Intell. Eng. Appl.*, vol. 5, no. 3, 2026.
- [20] A. Saputra, M. L. Ferdilian, H. Irsyad, and A. Rahman, “ANALISIS PENCARIAN TEKS EMOSI DALAM TWEET BERBAHASA INDONESIA MENGGUNAKAN TF-IDF DAN COSINE SIMILARITY,” *J. Informatics Comput. Eng. Res.*, vol. 2, no. 2, pp. 38–42, 2025.
- [21] Honakan, Adiwijaya, and S. Al Faraby, “Analisis Dan Implementasi Support Vector Machine Dengan String Kernel Dalam Melakukan Klasifikasi Berita Berbahasa Indonesia Analysis and Implementation Support Vector Machine With String Kernel for Classification indonesian news,” *e-Proceeding Eng.*, vol. 5, no. 1, pp. 1701–1710, 2018.
- [22] A. N. Wahyuni, T. Listyorini, and E. Supriyati, “IMPLEMENTASI MODEL INDOBERT DAN MBERT UNTUK DETEKSI BERITA HOAKS BERBASIS WEB,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 10, no. 2, pp. 2736–2743, 2026.