



Analisis Sentimen Electronic *Word-of-mouth* TikTok pada Bisnis Jasa Photobox untuk Mendukung Strategi Pemasaran Digital

Kelvin Siahaan^{1*}, Muhamad Yusup², Tuti Nurhaeni³, Nuke Puji Lestari Santoso⁴

^{1,2,3,4}Program Studi Bisnis Digital, Universitas Raharja, Tangerang, Indonesia

*Penulis korespondensi: kelvin.siahaan@raharja.info

Abstract. *The self-photo (photobox) business in Indonesia has grown rapidly as a digital-era experience service, propelled largely by user-generated content (UGC) that acts as electronic word-of-mouth (eWOM) on TikTok. For small businesses, the public opinion embedded in these comments is a valuable yet underused marketing asset. This study analyses consumer sentiment toward photobox services and positions the analysis as a social-media-analytics tool to support digital marketing decisions. A machine-learning approach was applied to 1.086 labelled Indonesian TikTok comments grouped into positive, neutral, and negative classes. Comments were represented with TF-IDF features and classified using Naive Bayes, Logistic Regression, Random Forest, and Support Vector Machine, evaluated through five-fold stratified cross-validation, with SMOTE and SMOTE with Tomek Links applied to handle class imbalance. The results show that the eWOM is overwhelmingly positive (64,0%) and that most comments are extremely short. The best accuracy reached about 63,2%, but every model struggled with the neutral and negative classes; Logistic Regression with SMOTE detected negative opinion best. Read through the Elaboration Likelihood Model, the brevity of the comments reflects low-effort, peripheral-route reactions common on short-video platforms. Managerially, the strong positive buzz signals healthy brand momentum, yet the scarcity of substantive feedback means business owners must combine sentiment analytics with other listening channels to obtain actionable insight.*

Keywords: Digital Marketing; Electronic Word-Of-Mouth; Machine Learning; Sentiment Analysis; TikTok

Abstrak. Bisnis foto mandiri (photobox) di Indonesia tumbuh pesat sebagai jasa pengalaman di era digital dan banyak didorong oleh *user-generated content* (UGC) yang berperan sebagai *electronic word-of-mouth* (eWOM) di TikTok. Bagi usaha kecil, opini publik dalam komentar tersebut adalah aset pemasaran berharga yang belum banyak dimanfaatkan. Penelitian ini menganalisis sentimen konsumen terhadap jasa photobox dan memposisikan analisis tersebut sebagai alat social media analytics untuk mendukung keputusan pemasaran digital. Pendekatan *machine learning* diterapkan pada 1.086 komentar TikTok berbahasa Indonesia yang telah dilabeli ke dalam kelas positif, netral, dan negatif. Komentar direpresentasikan dengan fitur TF-IDF, lalu diklasifikasikan menggunakan *Naive Bayes*, *Logistic Regression*, *Random Forest*, dan *Support Vector Machine* yang dievaluasi melalui validasi silang berlapis lima, dengan SMOTE dan SMOTE + Tomek Links untuk menangani ketidakseimbangan kelas. Hasilnya, eWOM sangat didominasi sentimen positif (64,0%) dan mayoritas komentar teramat pendek. Akurasi terbaik hanya sekitar 63,2%, tetapi semua model kesulitan pada kelas netral dan negatif; *Logistic Regression* dengan SMOTE paling baik mendeteksi opini negatif. Ditinjau melalui Elaboration Likelihood Model, komentar yang pendek mencerminkan reaksi berupaya rendah pada jalur peripheral yang khas platform video pendek. Secara manajerial, kuatnya buzz positif menandakan momentum merek yang sehat, namun minimnya umpan balik substantif menuntut pelaku usaha memadukan analisis sentimen dengan kanal mendengarkan lain agar memperoleh wawasan yang dapat ditindaklanjuti.

Kata kunci: Analisis Sentimen; Electronic Word-Of-Mouth; Machine Learning; Pemasaran Digital; TikTok

1. LATAR BELAKANG

Ekonomi digital mengubah cara usaha kecil menjangkau konsumen. Platform video pendek seperti TikTok kini menjadi kanal pemasaran utama yang mempertemukan pelaku usaha dengan calon pelanggan tanpa biaya iklan besar. Di ruang ini, konten yang dibuat pengguna, atau *user-generated content* (UGC), menyebar cepat dan kerap menentukan apakah sebuah layanan ramai dikunjungi atau terlupakan (Hayat et al., 2024; Andriansyah et al., 2025).

UGC pada dasarnya adalah bentuk *electronic word-of-mouth* (eWOM), yaitu ulasan dan opini antar-konsumen yang tersebar melalui media daring. Karena dinilai lebih jujur daripada

iklan resmi, eWOM sangat berpengaruh pada keputusan pembelian, terutama untuk jasa yang belum pernah dicoba (Hennig-Thurau et al., 2004; Jain et al., 2021). Bagi pemasaran digital, komentar yang menumpuk di kolom TikTok karenanya bukan sekadar percakapan, melainkan cerminan persepsi pasar yang berharga.

Salah satu bisnis yang lahir dan tumbuh dari ekosistem ini adalah jasa photobox, yaitu bilik foto mandiri dengan properti dan pencahayaan yang tertata. Sebagai usaha berbasis pengalaman di era digital, nilai photobox lebih ditentukan oleh kesan, suasana, dan momen yang layak dibagikan daripada oleh spesifikasi teknis (Herlina et al., 2022; Rosanti, 2025). Model bisnis seperti ini sangat bergantung pada eWOM: semakin banyak konsumen membagikan pengalaman positif, semakin besar peluang usaha tersebut berkembang.

Persoalannya, volume komentar di TikTok terlalu besar untuk dibaca satu per satu, padahal di dalamnya tersimpan opini yang bisa memandu strategi pemasaran. Analisis sentimen berbasis *machine learning* menawarkan jalan keluar dengan memilah opini secara otomatis menjadi positif, netral, atau negatif, sehingga dapat berfungsi sebagai alat social media analytics yang terjangkau bagi pelaku usaha (Setiawan et al., 2024; Maharani, 2024; Rihastuti & Rosyidi, 2025). Pendekatan ini telah dipakai pada beragam topik bisnis dan publik (Nurrahman et al., 2025; Sofyan & Iqbal, 2025; Asmawi et al., 2025).

Namun komentar TikTok bukan bahan yang mudah diolah. Teksnya sangat pendek, penuh singkatan dan bahasa gaul, serta opini negatif jauh lebih sedikit dibanding pujian. Ketimpangan kelas seperti ini membuat model cenderung berat sebelah ke kelas mayoritas. Sejumlah studi menyiasatinya dengan teknik penyeimbangan seperti SMOTE dan kombinasi SMOTE dengan Tomek Links (Marzuki et al., 2023; Switrayana et al., 2023; Firmansyah et al., 2024; Widyaiswari et al., 2026), tetapi hasilnya tidak selalu konsisten antar-topik.

Untuk memahami mengapa eWOM photobox cenderung singkat dan reaktif, penelitian ini meminjam Elaboration Likelihood Model (ELM). Teori tersebut membedakan dua jalur pembentukan sikap: jalur central yang menimbang argumen secara mendalam, dan jalur peripheral yang bersandar pada isyarat cepat seperti daya tarik visual dan emosi (Petty & Cacioppo, 1986; Camilleri, 2022). Format hiburan yang singkat mendorong penonton menilai lewat jalur peripheral, sebagaimana ditemukan pada konteks video pendek pariwisata (Li & Rui, 2025).

Berpijak pada uraian itu, penelitian ini bertujuan menjawab empat hal: bagaimana distribusi sentimen eWOM terhadap jasa photobox di TikTok; seberapa baik kinerja empat algoritma *machine learning* dalam mengklasifikasikannya; sejauh mana SMOTE dan Tomek Links memperbaiki deteksi kelas minoritas; serta apa implikasi temuan tersebut

bagi strategi pemasaran digital usaha photobox. Kontribusi penelitian ini ada dua. Secara praktis, ia menunjukkan bagaimana analisis sentimen dapat menjadi alat bantu keputusan pemasaran bagi usaha kecil; secara metodologis, ia mendokumentasikan secara jujur batas kinerja pipeline sentimen klasik ketika berhadapan dengan komentar video pendek yang pendek dan timpang.

2. KAJIAN TEORITIS

Media Sosial dan Pemasaran di TikTok

TikTok menyajikan konten melalui algoritma laman For You yang menampilkan video sesuai pola keterlibatan pengguna, bukan sekadar jaringan pertemanan. Mekanisme ini membuat unggahan usaha kecil pun berpeluang menjangkau banyak orang, sehingga platform tersebut menjadi kanal pemasaran organik yang subur dan hemat biaya (Setiawan et al., 2024; Qolbidiasih et al., 2026).

User-generated content dan *Electronic Word-of-mouth*

UGC adalah konten yang dibuat dan disebarakan pengguna, bukan oleh merek, dan menjadi wujud paling nyata dari *electronic word-of-mouth* (eWOM) di media sosial. Tiga ciri membedakannya dari iklan: dipandang lebih autentik, mencerminkan pengalaman nyata sesama konsumen, dan menyebar nyaris tanpa biaya. Ketiga hal itu membuat eWOM efektif menurunkan persepsi risiko dan mendorong keputusan pembelian, terutama pada jasa yang baru akan dicoba (Hennig-Thurau et al., 2004; Jain et al., 2021; Hayat et al., 2024). Karena itu, memantau sentimen eWOM merupakan bagian penting dari pemasaran digital modern.

Analisis Sentimen sebagai Social Media Analytics

Analisis sentimen bertugas menentukan polaritas opini dalam sebuah teks, umumnya positif, netral, atau negatif, sehingga opini publik dalam jumlah besar dapat diringkas menjadi informasi yang berguna bagi pengambilan keputusan bisnis. Pada pendekatan supervised, teks lebih dulu diubah menjadi angka. Salah satu cara yang lazim adalah TF-IDF, yang memberi bobot lebih besar pada kata yang sering muncul di satu komentar tetapi jarang di komentar lain. Fitur tersebut kemudian diklasifikasikan dengan berbagai algoritma. *Naive Bayes* menghitung peluang kelas berdasarkan asumsi kebebasan antar-kata (Sofyan & Iqbal, 2025); *Logistic Regression* memodelkan peluang kelas secara linear dan mudah ditafsirkan; *Random Forest* menggabungkan banyak pohon keputusan untuk menekan overfitting (Wijanarko et al., 2024; Rihastuti & Rosyidi, 2025); sedangkan *Support Vector Machine* mencari batas pemisah paling lebar antar-kelas dan sering unggul pada data teks berdimensi tinggi (Prasetyo & Hilda, 2024; Nurrahman et al., 2025).

Ketidakseimbangan Kelas: SMOTE dan Tomek Links

Ketika satu kelas jauh lebih banyak daripada kelas lain, model cenderung mengabaikan kelas minoritas demi akurasi total. SMOTE mengatasinya dengan membuat contoh sintesis kelas minoritas di antara tetangga terdekatnya, bukan sekadar menggandakan data yang ada. Tomek Links melengkapinya dengan membuang pasangan data antar-kelas yang saling berdekatan sehingga batas antar-kelas menjadi lebih tegas. Kombinasi keduanya banyak dipakai pada analisis sentimen berbahasa Indonesia dengan hasil yang bervariasi (Marzuki et al., 2023; Switrayana et al., 2023; Firmansyah et al., 2024). Untuk menilai kinerja secara adil pada data timpang, evaluasi tidak cukup memakai akurasi, melainkan juga *precision*, *recall*, F1-score, dan *confusion matrix* (Rininda et al., 2024; Armaeni et al., 2024).

Elaboration Likelihood Model

ELM menjelaskan bahwa sikap terbentuk melalui dua jalur. Jalur central bekerja ketika seseorang termotivasi dan mampu menimbang argumen secara mendalam, sedangkan jalur peripheral bekerja ketika penilaian bersandar pada isyarat cepat seperti estetika, emosi, atau popularitas sumber. Saat keterlibatan kognitif rendah, misalnya ketika menonton hiburan singkat, jalur peripheral cenderung dominan (Petty & Cacioppo, 1986; Camilleri, 2022; Li & Rui, 2025). Dalam penelitian ini ELM dipakai sebagai lensa untuk menafsirkan karakteristik komentar dan perilaku konsumen, bukan sebagai hipotesis yang diuji secara statistik.

Bisnis Jasa Photobox dalam Ekonomi Digital

Photobox berkembang dari budaya foto instan Asia Timur sebelum masuk ke Indonesia dan menjelma menjadi bagian dari ekonomi kreatif digital. Sebagai jasa berbasis pengalaman, produknya tidak berwujud dan konsumsinya menyatu dengan produksinya. Karena calon pengunjung tidak dapat menilai hasil sebelum mencoba, mereka sangat bergantung pada eWOM berupa unggahan dan komentar konsumen sebelumnya. Kondisi ini menjadikan pemasaran photobox hampir sepenuhnya bertumpu pada konten media sosial (Herlina et al., 2022; Rosanti, 2025).

Kerangka Pemikiran

Kerangka penelitian ini menempatkan komentar TikTok sebagai eWOM yang membawa sentimen konsumen, lalu memanfaatkan *machine learning* untuk meringkasnya menjadi masukan bagi keputusan pemasaran digital. Dua sifat data, yaitu ketimpangan kelas dan komentar yang sangat pendek, diduga menekan kemampuan model mengenali kelas minoritas. Teknik penyeimbangan diharapkan memperbaiki deteksi opini negatif, tetapi keterbatasan isi teks diperkirakan tetap menyulitkan pengenalan kelas netral. ELM digunakan untuk

menjelaskan mengapa eWOM photobox cenderung singkat dan bermuatan emosi ringan, yang sekaligus membatasi sinyal yang tersedia bagi model.

3. METODE PENELITIAN

Desain Penelitian

Penelitian ini memakai pendekatan kuantitatif-komputasional dengan desain deskriptif-analitik. Pendekatan tersebut dipilih karena memungkinkan analisis opini konsumen berskala besar langsung dari perilaku organik pengguna, tanpa intervensi instrumen seperti pada survei. Alur kerjanya mengikuti tahapan baku pada Tabel 1, mulai dari penyiapan data hingga evaluasi model.

Tabel 1. Tahapan penelitian

Tahap	Kegiatan
1. Penyiapan data	Membaca korpus komentar berlabel dan membuang entri kosong
2. Praproses	Case folding, pembersihan, dan pembobotan fitur TF-IDF
3. Penyeimbangan	Penerapan SMOTE dan SMOTE + Tomek Links pada data latih
4. Pemodelan	Pelatihan <i>Naive Bayes</i> , <i>Logistic Regression</i> , <i>Random Forest</i> , SVM
5. Evaluasi	Validasi silang berlapis lima, metrik, dan <i>confusion matrix</i>

Sumber: diolah penulis (2026)

Sumber dan Karakteristik Data

Data berupa 1.087 komentar publik pada konten UGC photobox di TikTok yang dikumpulkan melalui teknik web scraping (Erama, 2025). Setiap komentar disertai kategori sentimen yang telah dilabeli ke dalam tiga kelas, yaitu positif, netral, dan negatif. Satu komentar tanpa isi teks dikeluarkan, sehingga 1.086 komentar dianalisis lebih lanjut. Ciri paling menonjol dari korpus ini adalah komentar yang teramat pendek, dengan panjang rata-rata 1,70 kata dan 86,8% di antaranya hanya terdiri atas satu kata, seperti seruan singkat atau penyebutan nama teman. Sifat ini penting dicatat sejak awal karena berpengaruh langsung pada kinerja model.

Pelabelan Sentimen

Kategori sentimen pada setiap komentar ditetapkan secara manual berdasarkan polaritas opininya. Komentar yang menyatakan kepuasan atau apresiasi diberi label positif, komentar bernada keluhan atau kekecewaan diberi label negatif, sedangkan pertanyaan dan pernyataan informatif yang tidak memihak diberi label netral. Korpus berlabel inilah yang menjadi acuan pada seluruh tahap klasifikasi. Perlu ditegaskan bahwa pelabelan komentar yang sangat pendek mengandung tingkat subjektivitas, dan hal ini kelak dibahas sebagai salah satu keterbatasan.

Praproses dan Ekstraksi Fitur

Komentar mentah dibersihkan melalui pelipatan huruf (case folding), penghapusan tautan dan simbol yang tidak bermakna, serta penyeragaman spasi. Teks yang telah bersih

kemudian diubah menjadi vektor angka dengan TF-IDF pada rentang unigram hingga bigram, yang membobot kata menurut kekhasannya pada tiap komentar. Representasi inilah yang menjadi masukan bagi seluruh model.

Pemodelan dan Penyeimbangan Kelas

Empat algoritma diuji, yakni *Naive Bayes*, *Logistic Regression*, *Random Forest*, dan *Support Vector Machine*. Untuk menghadapi ketimpangan kelas, penyeimbangan dilakukan hanya pada data latih di setiap lipatan agar tidak terjadi kebocoran informasi ke data uji. Dua skema dibandingkan, yaitu SMOTE saja serta SMOTE yang digabung dengan Tomek Links. Dengan begitu, pengaruh masing-masing teknik terhadap deteksi kelas minoritas dapat ditelusuri secara terpisah (Marzuki et al., 2023; Switrayana et al., 2023).

Evaluasi Model

Kinerja diukur menggunakan validasi silang berlapis 5 yang menjaga proporsi kelas pada tiap lipatan. Metrik yang dilaporkan mencakup akurasi, *precision*, *recall*, dan F1-score untuk setiap kelas, dilengkapi *confusion matrix* yang memperlihatkan pola kesalahan klasifikasi. Karena data sangat timpang, akurasi tidak dijadikan satu-satunya tolok ukur; perhatian utama diarahkan pada *recall* dan F1-score kelas minoritas (Rininda et al., 2024; Armaeni et al., 2024).

4. HASIL DAN PEMBAHASAN

Hasil

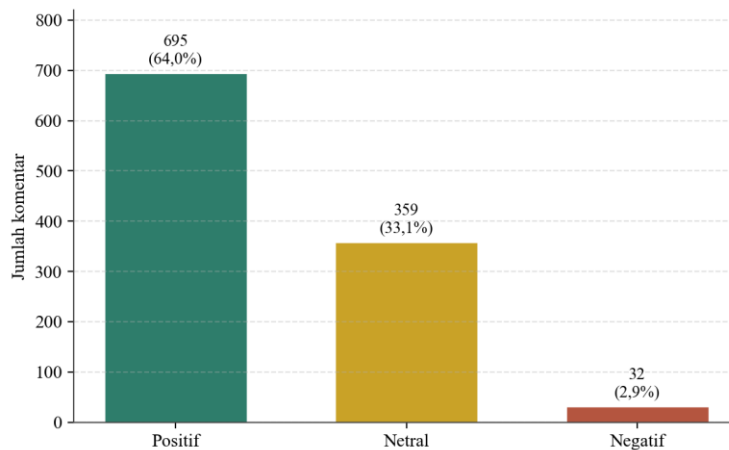
Karakteristik Data

Dari 1.086 komentar yang dianalisis, opini publik sangat condong ke arah positif. Sebanyak 64,0% komentar berlabel positif, 33,1% netral, dan hanya 2,9% negatif (Tabel 2 dan Gambar 1). Ketimpangan ini sangat tajam: jumlah komentar positif hampir dua puluh dua kali lipat komentar negatif. Dari sudut pandang pemasaran, dominasi eWOM positif seperti ini menandakan citra merek yang kuat di media sosial, meskipun bagi model klasifikasi justru menjadi tantangan (Maharani, 2024; Andriansyah et al., 2025).

Tabel 2. Distribusi kategori sentimen

Kategori	Jumlah	Persentase
Positif	695	64,0%
Netral	359	33,1%
Negatif	32	2,9%
Total	1.086	100,0%

Sumber: data penelitian (2026)

**Gambar 1.** Distribusi kategori sentimen komentar

Sumber: data penelitian (2026)

Selain timpang, komentar juga teramat pendek. Sebanyak 86,8% komentar hanya berisi satu kata dan 90,4% berisi tiga kata atau kurang. Contoh nyata pada Tabel 3 memperlihatkan bahwa banyak komentar berupa seruan singkat, pertanyaan sekilas, atau penyebutan nama teman. Teks sesingkat itu menyediakan sangat sedikit kata bagi TF-IDF, sehingga sinyal yang dapat dipelajari model pun terbatas.

Tabel 3. Contoh komentar pada tiap kategori sentimen

Kategori	Contoh komentar
Positif	“sumpah bru plg dri sna, bagus bgt woyy kameranya trs hasilnya jg bagus bgt”
Netral	“klo photobooth aja ga ke restonya bole gaa ka?”
Negatif	“padahal bagus tempatnya, masih ada aja yg ngerusak”

Sumber: data penelitian (2026)

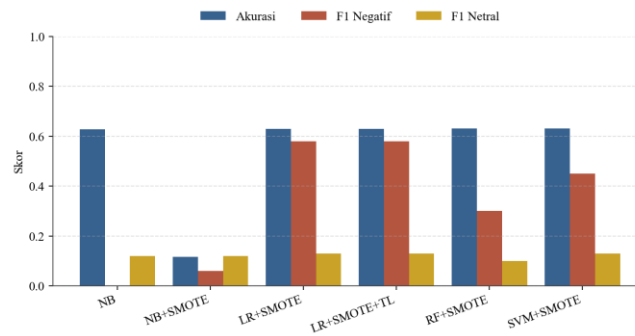
Perbandingan Kinerja Model

Kinerja keenam konfigurasi model dirangkum pada Tabel 4 dan Gambar 2. Empat dari enam konfigurasi mencapai akurasi yang berdekatan di kisaran 63%, dengan nilai tertinggi 63,2% pada *Random Forest* yang dipadukan SMOTE. Namun angka akurasi ini menipu: karena kelas positif mendominasi, model bisa terlihat cukup akurat hanya dengan cenderung menebak positif. Pengecualian mencolok terjadi pada *Naive Bayes* dengan SMOTE, yang akurasinya anjlok ke 11,6% karena contoh sintetis justru menambah derau bagi model tersebut.

Tabel 4. Perbandingan kinerja enam konfigurasi model

Konfigurasi	Akurasi	Recall Negatif	F1 Negatif	F1 Netral
<i>Naive Bayes</i> (tanpa penyeimbangan)	62,7%	0,00	0,00	0,12
<i>Naive Bayes</i> + SMOTE	11,6%	0,91	0,06	0,12
<i>Logistic Regression</i> + SMOTE	62,9%	0,66	0,58	0,13
<i>Logistic Regression</i> + SMOTE + Tomek	62,9%	0,66	0,58	0,13
<i>Random Forest</i> + SMOTE	63,2%	0,20	0,30	0,10
SVM + SMOTE	63,1%	0,41	0,45	0,13

Sumber: diolah penulis (2026)



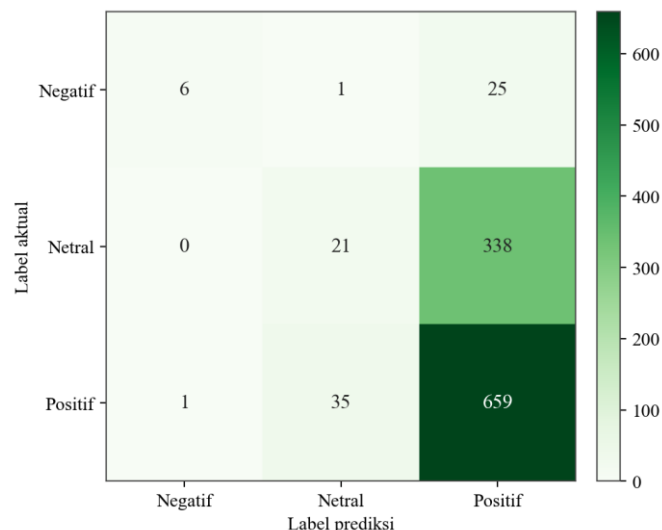
Gambar 2. Perbandingan akurasi, F1 negatif, dan F1 netral antar model

Sumber: diolah penulis (2026)

Perbedaan sesungguhnya baru terlihat pada kelas minoritas. *Logistic Regression* dengan SMOTE mendeteksi kelas negatif paling baik, dengan *recall* 0,66 dan F1-score 0,58, jauh di atas *Random Forest* yang *recall* negatifnya hanya 0,20. Penambahan Tomek Links pada *Logistic Regression* tidak mengubah hasil secara berarti, yang menandakan hanya sedikit pasangan data ambigu yang bisa dibuang pada korpus ini. Sementara itu, kelas netral tetap menjadi titik lemah semua model, dengan F1-score yang tidak pernah melampaui 0,13.

Analisis Kesalahan Klasifikasi

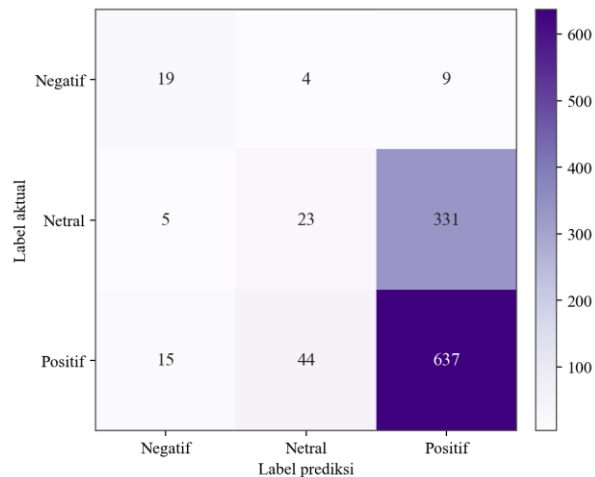
Confusion matrix pada Gambar 3 dan Gambar 4 menjelaskan sumber kesalahan. Pada *Random Forest* dengan SMOTE, hanya 21 dari 359 komentar netral yang dikenali benar, sedangkan sebagian besar sisanya diprediksi sebagai positif. Pola serupa terjadi pada kelas negatif yang hanya benar 6 dari 32. Sebaliknya, kelas positif nyaris selalu tepat (94,8% benar), yang menegaskan bahwa model condong ke kelas mayoritas.



Gambar 3. *Confusion matrix* Random Forest dengan SMOTE

Sumber: diolah penulis (2026)

Logistic Regression dengan SMOTE dan Tomek Links memperbaiki keadaan pada kelas negatif, dengan 19 dari 32 komentar dikenali benar. Meskipun begitu, kelas netral tetap banyak tertukar menjadi positif, sebagaimana tampak pada Gambar 4. Temuan ini konsisten dengan sifat data: komentar netral sering berupa pertanyaan pendek yang secara permukaan mirip komentar positif, sehingga sukar dibedakan tanpa pemahaman konteks.



Gambar 4. *Confusion matrix Logistic Regression dengan SMOTE dan Tomek Links*

Sumber: diolah penulis (2026)

Pembahasan

Rangkaian hasil di atas menunjuk pada satu simpulan yang konsisten: batas kinerja model lebih banyak ditentukan oleh sifat data daripada oleh pemilihan algoritma. Akurasi yang mentok di kisaran 63% bukan karena algoritmanya lemah, melainkan karena kelas positif mendominasi dan komentar terlalu pendek untuk memberi petunjuk yang cukup. SMOTE membantu menaikkan deteksi kelas negatif, terutama pada *Logistic Regression*, namun tidak mampu menolong kelas netral yang memang ambigu (Marzuki et al., 2023; Firmansyah et al., 2024).

Ditinjau melalui ELM, pola ini masuk akal. Menonton video photobox adalah aktivitas dengan keterlibatan kognitif rendah, sehingga penonton merespons lewat jalur peripheral berupa reaksi singkat dan emosional, bukan lewat argumen panjang di jalur central (Petty & Cacioppo, 1986; Li & Rui, 2025). Reaksi semacam itu menghasilkan eWOM satu-dua kata yang, meski jumlahnya melimpah, miskin informasi bagi model berbasis fitur permukaan seperti TF-IDF.

Bagi strategi pemasaran digital usaha photobox, temuan ini membawa beberapa implikasi. Pertama, dominasi eWOM positif menunjukkan momentum merek yang sehat dan modal sosial yang kuat untuk menarik pelanggan baru, sehingga konten yang menonjolkan pengalaman visual dan momen kebersamaan layak terus diproduksi. Kedua, karena sebagian

besar komentar bersifat reaktif dan minim substansi, buzz positif tidak otomatis berarti umpan balik yang dapat ditindaklanjuti; pelaku usaha perlu memadukan analisis sentimen dengan kanal lain seperti ulasan atau survei singkat. Ketiga, banyaknya komentar netral yang berupa pertanyaan tentang harga, lokasi, dan cara pemesanan justru menandakan minat pasar yang nyata dan peluang konversi bila usaha merespons dengan informasi yang jelas dan cepat.

Bagi kajian akademik, hasil ini menegaskan bahwa pipeline sentimen klasik memiliki langit-langit kinerja pada data video pendek, sehingga model peka konteks seperti berbasis transformer dan praproses yang lebih ketat menjadi arah yang menjanjikan untuk analitik pemasaran ke depan (Jain et al., 2021; Perwira et al., 2025).

5. KESIMPULAN DAN SARAN

Penelitian ini menyimpulkan bahwa eWOM terhadap jasa photobox di TikTok sangat didominasi sentimen positif (64,0%), diikuti netral (33,1%), dan sangat sedikit negatif (2,9%). Dari empat algoritma yang diuji, akurasi tertinggi hanya sekitar 63,2%, dan seluruh model kesulitan mengenali kelas netral. Teknik SMOTE, khususnya pada *Logistic Regression*, terbukti paling membantu meningkatkan deteksi kelas negatif, sedangkan penambahan Tomek Links tidak memberi perbedaan berarti. Batas kinerja ini terutama disebabkan oleh ketimpangan kelas dan komentar yang teramat pendek, bukan oleh kelemahan algoritma. Secara praktis, analisis sentimen tetap berguna sebagai alat social media analytics untuk memantau citra merek, asalkan pelaku usaha menyadari keterbatasannya dan melengkapinya dengan sumber umpan balik lain.

Penelitian ini memiliki sejumlah keterbatasan. Data hanya mencakup komentar publik yang telah dilabeli sebelumnya, sehingga pelabelan pada teks yang sangat pendek berpotensi subjektif. Analisis juga terbatas pada tingkat sentimen secara keseluruhan, belum menyentuh aspek yang dibicarakan konsumen. Untuk penelitian lanjutan, disarankan menerapkan praproses yang lebih ketat guna menyaring komentar tanpa muatan opini, mencoba model peka konteks seperti IndoBERT, memperluas data agar kelas negatif lebih terwakili, serta mengaitkan sentimen dengan indikator bisnis nyata seperti jumlah kunjungan atau penjualan agar manfaatnya bagi pemasaran digital semakin terukur (Perwira et al., 2025; Wijanarko et al., 2024).

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada dosen pembimbing dan seluruh pihak yang telah memberikan masukan pada penelitian ini. Artikel ini merupakan bagian dari tugas akhir penulis.

DAFTAR REFERENSI

- Andriansyah, Rahmawati, K. F., Dedu, M., & Agustina. (2025). Tiktok media in purchase decision: The effect of user generated content (UGC), brand equity, and digital marketing on skincare product. *Mabsya: Jurnal Manajemen Bisnis Syariah*, 7(2), 17-32. <https://doi.org/10.24090/mabsya.v7i2.13563>
- Armaeni, P. P., Wiguna, I. K. A. G., & Parwita, W. G. S. (2024). Sentiment analysis of YouTube comments on the closure of TikTok Shop using *Naive Bayes* and Decision Tree method comparison. *Jurnal Galaksi*, 1(2), 70-80. <https://doi.org/10.70103/galaksi.v1i2.15>
- Asmawi, M. A. H. A., Isawasan, P., Shanmugam, L., Salleh, K. A., & Savita, K. S. (2025). Exploring sentiment trends in TikTok comments using GPT for influencer content strategy. *e-Academia Journal*, 14(1).
- Camilleri, M. A. (2022). Walking the talk about corporate social responsibility communication: An elaboration likelihood model perspective. *Business Ethics, the Environment & Responsibility*, 31(3), 649-661. <https://doi.org/10.1111/beer.12427>
- Erama, R. (2025). Pemanfaatan platform cloud Google Colab untuk scraping komentar TikTok pada konten Gorontalo sebagai dasar analisis respons warganet. *Journal of Applied Engineering Science*, 1(2), 124-134. <https://doi.org/10.65177/jaes.v1i2.38>
- Firmansyah, A. S., Aziz, A., & Ahsan, M. (2024). Optimasi K-Nearest Neighbor menggunakan algoritma SMOTE untuk mengatasi imbalance class pada klasifikasi analisis sentimen. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(6), 3341-3347. <https://doi.org/10.36040/jati.v7i6.7257>
- Hayat, M. A., Hermawan, A., & Nuryadin, A. (2024). Pengaruh social media marketing dan user generated content (UGC) terhadap purchase decision produk Luxcrime pada generasi Z di TikTok. *Journal of Information Systems Management and Digital Business (JISMDB)*, 2(1).
- Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? *Journal of Interactive Marketing*, 18(1), 38-52. <https://doi.org/10.1002/dir.10073>
- Herlina, E., Sulaiman, E., Pratiwi, V. A., Permana, I. S., Pertiwi, W., & Saputra, A. (2022). Pelatihan mini studio photo box untuk meningkatkan kualitas foto produk dan daya jual produk UMKM. *Socircle: Journal of Social Community Services*, 1(1), 20-25. <https://doi.org/10.58468/socircle.v1i1.5>
- Jain, P. K., Saravanan, V., & Pamula, R. (2021). A hybrid CNN-LSTM: A deep learning approach for consumer sentiment analysis using qualitative user-generated contents. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 20(5). <https://doi.org/10.1145/3457206>

- Li, H., Yu, B. X. B., Li, G., & Gao, H. (2023). Restaurant survival prediction using customer-generated content: An aspect-based sentiment analysis of online reviews. *Tourism Management*, 96, 104707. <https://doi.org/10.1016/j.tourman.2022.104707>
- Li, K., & Rui, L. (2025). The impact of short-form videos on tourist travel intention toward cities: An integrated theory of planned behavior and elaboration likelihood model approach. *Humanities and Social Sciences Communications*, 12(1). <https://doi.org/10.1057/s41599-025-06300-x>
- Maharani, M. Z. (2024). Analisis sentimen positif terhadap Avoskin sebagai eco friendly brand di media sosial X dan TikTok. *Filosofi: Publikasi Ilmu Komunikasi, Desain, Seni Budaya*, 1(3), 125-140. <https://doi.org/10.62383/filosofi.v1i3.169>
- Marzuki, K., Putra, L. G. R., Hairani, H., Mardedi, L. Z. A., & Guterres, J. X. (2023). Performance improvement of the *Random Forest* method based on SMOTE-Tomek link on Lombok tourism sentiment analysis. *Jurnal Bumigora Information Technology (BITE)*, 5(2), 151-158. <https://doi.org/10.30812/bite.v5i1.3166>
- Nurrahman, A. A., Mauladi, M., & Rahman, A. (2025). Analisis sentimen masyarakat terhadap kenaikan harga bahan bakar minyak menggunakan *Support Vector Machine* dan SMOTE. *Sudo Jurnal Teknik Informatika*, 4(2), 50-56. <https://doi.org/10.56211/sudo.v4i2.908>
- Perwira, R. I., Permadi, V. A., Purnamasari, D. I., & Agusdin, R. P. (2025). Domain-specific fine-tuning of IndoBERT for aspect-based sentiment analysis in Indonesian travel user-generated content. *Journal of Information Systems Engineering and Business Intelligence*, 11(1), 30-40. <https://doi.org/10.20473/jisebi.11.1.30-40>
- Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, 19, 123-205. [https://doi.org/10.1016/S0065-2601\(08\)60214-2](https://doi.org/10.1016/S0065-2601(08)60214-2)
- Prasetyo, D. T., & Hilda, A. M. (2024). Metode *Support Vector Machine* untuk analisis sentimen aplikasi Threads di Google Play Store. *The Indonesian Journal of Computer Science*, 13(5), 76-83. <https://doi.org/10.33022/ijcs.v13i5.4446>
- Qolbidiasih, M. N., Putranto, T. D., Prasetyo, D., & Firmannandya, A. (2026). Interpretation of promotional messages in TikTok live streaming: Consumer reception of Glad2Glow. *Media Mahardhika*, 24(2).
- Rihastuti, S., & Rosyidi, A. (2025). Analisis sentimen pengguna TikTok tentang progres pembangunan IKN dengan metode *Random Forest*. *Journal of Computer Science and Technology (JCS-TECH)*, 5(1), 19-23.
- Rininda, G., Santi, I. H., & Kirom, S. (2024). Penerapan SVM dalam analisis sentimen pada Edlink menggunakan pengujian *confusion matrix*. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(5), 3335-3342. <https://doi.org/10.36040/jati.v7i5.7420>
- Rosanti, R. (2025). Peningkatan kreativitas mahasiswa melalui inovasi produk memory photo box sebagai hadiah personal berbasis nilai emosional. *Akademik: Jurnal Pengabdian Masyarakat*, 3(4), 12-18. <https://doi.org/10.61722/japm.v3i4.4901>
- Setiawan, T., Liem, S., & Pribadi, D. M. R. (2024). Perbandingan algoritma SVM dan *Naive Bayes* dalam analisis sentimen komentar TikTok pada produk skincare. *Applied Information Technology and Computer Science*, 3(2).

- Sofyan, S. N., & Iqbal, M. (2025). Analisis sentimen terhadap dampak inflasi menggunakan *Naive Bayes*. *Bulletin of Information Technology (BIT)*, 6(1), 1-8.
- Switrayana, I. N., Ashadi, D., Hairani, H., & Aminuddin, A. (2023). Sentiment analysis and topic modeling of Kitabisa applications using *Support Vector Machine (SVM)* and SMOTE-Tomek links methods. *International Journal of Engineering and Computer Science Applications (IJECSA)*, 2(2), 81-91. <https://doi.org/10.30812/ijecsa.v2i2.3406>
- Widyaiswari, R. P., Dzulkarnain, A., & Rausanfitra, A. (2026). Imbalanced data handling for optimizing *Random Forest* in sentiment analysis of East Java gubernatorial election. *Jurnal Sistem Informasi dan Informatika (Simika)*, 9(1).
- Wijanarko, R., Ratnawati, D. E., & Adikara, P. P. (2024). Analisis sentimen dampak perkembangan artificial intelligence (AI) pada media sosial X/Twitter menggunakan metode *Random Forest*. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 8(5).