

Identification of Buzzers in Skincare Reviews Using a Lexicon-Based Sentiment Analysis Method

Arfiana Diah Pramesti^{1*}, Khothibul Umam^{2*}, Maya Rini Handayani^{3*}

* Teknologi Informasi, Fakultas Sains dan Teknologi, UIN Walisongo Semarang

2208096072@student.walisongo.ac.id¹, khothibul_umam@walisongo.ac.id², maya@walisongo.ac.id³

Article Info

Article history:

Received 2025-09-02

Revised 2025-10-05

Accepted 2025-10-14

Keyword:

Buzzers,
Product Reviews,
Skincare,
Sentiment Analysis,
Lexicon,
Classification,
Social Media.

ABSTRACT

Along with the rapid development of digital technology, social media has become the main platform for consumers to share experiences about products, including skincare products. However, it is not uncommon for reviews provided by users to not reflect authentic experiences, but rather reviews created by certain parties, or buzzers, to manipulate public perception. The presence of buzzers in skincare reviews is important to consider, as they can affect consumer trust and influence purchasing decisions. This study aims to identify the presence of buzzers in skincare product reviews using a lexicon dictionary-based sentiment analysis. Of the 529 comments analyzed, 75 comments showed negative sentiment and 454 comments showed positive sentiment. The classification results revealed that 85.8% of the comments belonged to the non-buzzer category, while 14.2% were indicated as buzzers. Evaluation of the classification model showed high accuracy, reaching 93%, but performance in detecting buzzers was limited, with a recall metric of only 0.50. This shows that while the model managed to classify non-buzzer comments well, there are still difficulties in identifying buzzer comments, mostly due to data imbalance. This research emphasizes the importance of a proper analytical approach in detecting inauthentic reviews to ensure the information consumers receive remains accurate, transparent, and accountable.



This is an open access article under the [CC-BY-SA](#) license.

I. PENDAHULUAN

Di zaman serba digital saat ini, media sosial dan *platform belanja online* telah menjadi tempat utama bagi konsumen untuk saling berbagi pengalaman tentang produk, termasuk produk kecantikan dan perawatan kulit. Ulasan dari pengguna sangat berpengaruh karena dapat menjadi referensi bagi calon pembeli dalam mengambil keputusan. Namun, tidak semua ulasan mencerminkan pengalaman yang sebenarnya. Ada kalanya, ulasan-ulasan tersebut dibuat oleh individu atau kelompok yang dibayar untuk memberikan penilaian positif demi meningkatkan citra suatu produk [1]. Kelompok ini sering disebut sebagai *buzzers*, yang biasanya memberikan ulasan secara berlebihan dan terkesan tidak natural, sehingga bisa menyesatkan konsumen lain [2]. Dengan pesatnya perkembangan media sosial dan digital marketing, para *buzzers* ini menjadi aktor komunikasi yang mampu

menyampaikan pesan secara efektif, yang bisa mempengaruhi persepsi publik terhadap suatu produk [3].

Menurut Maulidatur dan Ernungtyas (2019), *buzzers* diartikan sebagai individu yang memanfaatkan *platform media sosial* untuk menyebarkan pesan tertentu, yang sering kali berhubungan dengan kampanye politik atau komunikasi pemerintah, dan dapat mempengaruhi audiens secara besar [4]. Dalam konteks ini, *buzzers* berfungsi sebagai alat yang efisien dalam hubungan masyarakat, memungkinkan pesan untuk mencapai audiens yang lebih luas dengan biaya yang relatif rendah.

Selain itu, Dewita (2020) mengemukakan bahwa *buzzers* juga dapat dimanfaatkan untuk promosi produk atau layanan, terutama dalam mengembangkan ekonomi kreatif di sektor pariwisata [5]. Penggunaan *buzzers* yang tepat dapat mempercepat penyebaran informasi dan meningkatkan minat masyarakat terhadap produk yang dipromosikan.

Di sisi lain, Arianto (2020) menambahkan bahwa influencer, yang memiliki pengaruh besar di media sosial, sering kali bekerja sama dengan merek untuk memperkuat ekosistem pemasaran digital [6]. *Influencer* tidak hanya menyampaikan pesan kepada pengikutnya, tetapi juga dapat membentuk keputusan pembelian, yang menjadikannya aset berharga dalam strategi pemasaran.

Keberadaan *buzzer* dalam ulasan *skincare* perlu mendapat perhatian karena dapat merusak kepercayaan konsumen sekaligus memengaruhi keputusan pembelian. Penelitian menunjukkan bahwa *review* konsumen dan ulasan daring, sebagai bentuk modern dari *word-of-mouth*, memiliki peran penting dalam membangun kepercayaan serta meningkatkan loyalitas terhadap merek *skincare* [7]. Konsumen umumnya lebih percaya pada rekomendasi yang dianggap autentik dan relevan, namun *buzzer* yang kerap memberikan opini berlebihan atau bersifat promosi dapat menciptakan persepsi yang menyesatkan. Selain itu, ulasan yang beragam dan proporsional dinilai lebih kredibel, sedangkan *review* yang terlalu positif atau seragam justru dapat menurunkan kredibilitas produk. Oleh karena itu, mendeteksi keberadaan *buzzer* dalam *review skincare* menjadi langkah penting untuk menjaga integritas informasi, melindungi reputasi merek, dan memastikan keputusan konsumen tetap berdasarkan fakta.

Untuk mendeteksi adanya campur tangan *buzzer* dalam ulasan produk, analisis sentimen merupakan pendekatan yang umum digunakan. Salah satu metode dasar adalah pendekatan lexicon-based, di mana kata-kata dalam ulasan dianalisis berdasarkan polaritasnya—positif, negatif, atau netral [8]. Metode ini mengandalkan kamus sentimen yang telah ditentukan sebelumnya untuk menilai dan mengklasifikasikan suatu teks [9]. Dengan demikian, metode berbasis lexicon berfungsi dengan cara menyusun kamus kata terlebih dahulu. Kata-kata yang terdapat dalam kamus tersebut digunakan untuk menganalisis makna suatu kalimat, kemudian mengklasifikasikan kalimat tersebut sesuai dengan kategori yang relevan [10].

Pendekatan dengan metode ini relatif mudah diimplementasikan karena tidak memerlukan data latih, serta efektif dalam mendeteksi sentimen eksplisit. Namun, metode ini memiliki keterbatasan, terutama dalam mengenali konteks kalimat yang ambigu, ironi, atau intensi tersembunyi yang kerap muncul dalam ulasan *buzzer*. Selain itu, lexicon-based cenderung menghasilkan klasifikasi yang kurang akurat apabila digunakan pada opini yang menggunakan gaya bahasa tidak langsung atau tersamar [11].

Untuk mengatasi keterbatasan tersebut, pendekatan berbasis pembelajaran mesin seperti Support Vector Machine (SVM) mulai banyak digunakan karena mampu menangkap pola dan ciri khas teks yang lebih kompleks [12]. SVM dikembangkan oleh Vapnik pada tahun 1992 sebagai metode klasifikasi yang efektif dalam menangani masalah nonlinier. Teknik ini kemudian dikenal sebagai salah satu inovasi terbaru dalam bidang pembelajaran mesin, menggantikan pendekatan sebelumnya yang menggunakan Neural Network (NN) [13]. SVM bekerja dengan memetakan data ke dalam

ruang berdimensi tinggi untuk menemukan hyperplane optimal yang memisahkan kelas sentimen secara efektif [14].

Penerapan metode lexicon-based telah digunakan dalam berbagai bidang, salah satunya dalam penelitian berjudul "Implementasi Lexicon Based Untuk Analisis Sentimen Dalam Mengetahui Trend Wisata Pantai Di DI Yogyakarta Berdasarkan Data Twitter". Penelitian tersebut memanfaatkan data dari media sosial Twitter untuk menganalisis opini publik terhadap destinasi wisata pantai di Yogyakarta. Melalui klasifikasi sentimen berdasarkan kamus lexicon, penelitian ini mampu mengidentifikasi tren kunjungan wisata dan persepsi masyarakat terhadap objek wisata yang dibahas. Hasilnya memberikan gambaran yang berguna bagi pengambil kebijakan dalam menyusun strategi promosi wisata berbasis data [15].

Adapun penelitian lain oleh Youllia Indrawaty Nurhasanah dan Muhammad Thoriq Naufal berjudul "Klasifikasi Sentimen Pada Buzzer Politik Jelang Pemilu 2024 menggunakan Metode Lexicon-based" membahas mengenai mengenai penerapan metode lexicon-based dalam mengidentifikasi dan mengklasifikasikan sentimen dari akun-akun yang diduga sebagai *buzzer* politik. Penelitian ini menyoroti bagaimana akun-akun tersebut menyebarkan opini menjelang Pemilu 2024, serta menganalisis kecenderungan sentimen yang dihasilkan, apakah bersifat positif, negatif, atau netral terhadap tokoh atau partai politik tertentu. Hasil penelitian menunjukkan bahwa metode lexicon-based mampu memberikan gambaran awal mengenai pola komunikasi *buzzer* di media sosial, meskipun masih memiliki keterbatasan dalam memahami konteks atau ironi dalam teks [11].

Pemilihan metode berbasis lexicon pada penelitian ini dilatarbelakangi oleh pertimbangan empiris dan metodologis. Sebelumnya, pendekatan ini telah banyak dimanfaatkan dalam penelitian deteksi *buzzer* pada ranah politik maupun isu publik, dan terbukti mampu mengidentifikasi opini yang dimanipulasi secara cukup efektif. Maka dari itu, penelitian ini mengadaptasi metode serupa untuk diterapkan pada domain berbeda, yakni ulasan produk *skincare*. Sementara itu, kajian yang secara khusus menyoroti keberadaan *buzzer* pada ulasan produk *skincare* masih sangat terbatas. Padahal, industri kecantikan memiliki basis konsumen yang luas serta dipengaruhi secara signifikan oleh opini yang berkembang di media sosial. Oleh karena itu, penelitian ini diarahkan untuk mengevaluasi efektivitas metode berbasis lexicon dalam mengidentifikasi *buzzer* pada *review skincare*, dengan mempertimbangkan permasalahan ketidakseimbangan data antara komentar *buzzer* dan *non-buzzer*. Pemilihan pendekatan *lexicon-based* didasarkan pada pertimbangan bahwa metode ini bersifat ringan, tidak memerlukan dataset berukuran besar maupun proses pelabelan yang kompleks, serta relevan untuk penelitian eksploratif awal yang bertujuan memberikan pemahaman dasar mengenai pola *buzzer* dalam konteks ulasan *skincare*.

Meskipun konteksnya berbeda, pendekatan yang sama dapat diterapkan untuk menganalisis ulasan konsumen

terhadap produk kecantikan. Namun, pada kasus deteksi *buzzer*, tantangan menjadi lebih kompleks karena opini sering kali disusun untuk tujuan manipulatif dan tidak eksplisit. Oleh karena itu, pengembangan analisis sentimen berbasis lexicon perlu disesuaikan, misalnya dengan menambahkan pembobotan terhadap kata-kata berlebihan atau tidak wajar, agar mampu mengenali pola-pola ulasan yang mencurigakan. Dengan pendekatan yang tepat, analisis ini dapat menjadi alat yang efektif untuk membedakan ulasan asli dan ulasan buatan dalam konteks pemasaran digital.

Untuk mengukur efektivitas pendekatan ini, hasil klasifikasi sentimen yang diperoleh dari metode lexicon berbobot akan dievaluasi menggunakan algoritma Support Vector Machine (SVM). Evaluasi ini bertujuan untuk mengetahui sejauh mana akurasi, presisi, dan recall dari sistem dalam membedakan komentar *buzzer* dan *non-buzzer*. Dengan demikian, kombinasi pendekatan lexicon dan metode pembelajaran mesin diharapkan mampu meningkatkan performa sistem dalam mendeteksi *buzzer* secara lebih akurat dan andal.

II. METODE

A. Data dan Metode

Data penelitian diperoleh dari komentar pada unggahan akun yang membahas pengalaman penggunaan produk *skincare* di platform TikTok, mengingat media sosial tersebut menjadi salah satu saluran utama dalam menyebarkan opini konsumen dan memengaruhi persepsi publik. Proses pengumpulan data dilakukan secara manual, dengan periode pengambilan data berlangsung pada bulan Maret 2025. Adapun variabel penelitian yang digunakan dapat dilihat pada Tabel 1.

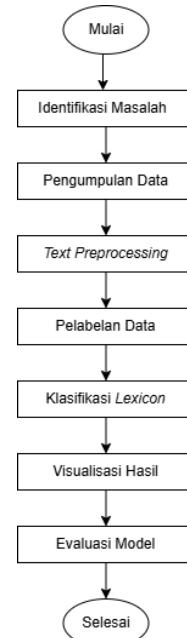
TABEL I
VARIABEL PENELITIAN

| Teks | Format (<i>Times New Roman</i>) |
|----------|---|
| Komentar | Komentar dari pengguna tiktok. |
| Skor | Nilai yang diberikan pada setiap teks berdasarkan kamus lexicon |
| Label | Teks yang sudah terlabel dan memiliki kategori sentimen positif dan negatif |

Metodologi yang diterapkan dalam penelitian ini adalah analisis sentimen dengan pendekatan berbasis lexicon. Tujuan penelitian ini adalah untuk mengklasifikasikan suatu komentar teridentifikasi sebagai *buzzer* atau tidak.

Dalam penelitian ini, proses anotasi dilakukan dengan pendekatan *hybrid*, yakni kombinasi antara anotasi manual oleh peneliti dan anotasi otomatis menggunakan sistem berbasis leksikon. Kamus sentimen yang digunakan bersifat manual (buatan sendiri) yang disesuaikan dengan konteks ulasan *skincare*, serta diperkaya dengan daftar kata positif dan negatif dari kamus NLTK lexicon. Proses pelabelan komentar ke dalam kategori *buzzer* maupun *non-buzzer* didasarkan pada pola sentimen tertentu, khususnya komentar dengan nuansa positif yang berlebihan, serta diperkuat dengan identifikasi

pola bahasa yang berulang-ulang. Dengan mekanisme ini, identifikasi *buzzer* tidak hanya bertumpu pada analisis sentimen semata, tetapi juga mempertimbangkan karakteristik linguistik yang khas dari komentar *buzzer*.

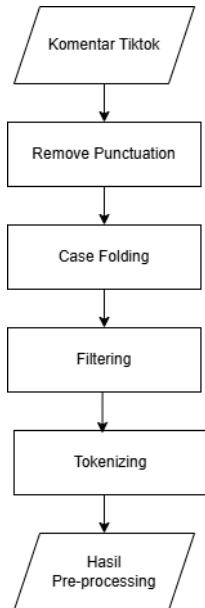


Gambar 1. Diagram Alur Penelitian

Penelitian ini terdiri dari beberapa tahapan, yaitu: (1) pengumpulan data yang dilakukan secara manual, (2) *text preprocessing*, mencakup pembersihan data dengan langkah-langkah seperti menghapus tanda baca, mengubah huruf menjadi huruf kecil (*case folding*), penyaringan kata (*filtering*), dan tokenisasi, (3) pelabelan data, yaitu memberikan skor pada teks komentar berdasarkan kamus kata positif dan negatif, (4) proses klasifikasi untuk menentukan kategori sentimen berdasarkan hasil penjumlahan skor, (5) visualisasi hasil pelabelan, yang dilakukan menggunakan diagram lingkaran untuk menampilkan frekuensi masing-masing sentimen, dan (6) Evaluasi Model, dilakukan dengan menggunakan data uji yang telah klasifikasi. Tujuan dari evaluasi ini adalah untuk mengukur tingkat akurasi pelabelan oleh metode Lexicon-Based [16].

B. Text Preprocessing

Komentar TikTok yang berhasil dikumpulkan mengandung banyak *noise*, sehingga diperlukan tahapan *text preprocessing*. Tahapan ini mencakup pembersihan data seperti menghapus tanda baca (*remove punctuation*), mengubah huruf menjadi bentuk kecil (*case folding*), normalisasi ejaan, penyaringan kata yang tidak relevan (*filtering*), serta pemecahan kalimat menjadi kata-kata terpisah (*tokenizing*) guna meningkatkan akurasi dalam proses analisis sentimen [17].



Gambar 2. Diagram Alur Text Preprocessing

C. Pelabelan Metode Lexicon

Setelah melalui proses *text preprocessing* untuk membersihkan data, tahap selanjutnya adalah pelabelan menggunakan metode lexicon. Pada tahap ini, setiap kalimat dianalisis berdasarkan keberadaan kata-kata yang terdapat dalam kamus lexicon, yang memuat daftar kata bermuatan positif dan negatif. Kata-kata yang dikenali dari kamus tersebut akan dihitung skornya berdasarkan jumlah kemunculannya dalam setiap kalimat atau teks.

$$S_{positive} \sum_{i \in t}^n positive score_i \quad (1)$$

$$S_{negative} \sum_{i \in t}^n negative score_i \quad (2)$$

Dalam konteks ini, ($S_{negative}$) merepresentasikan bobot negatif dalam sebuah kalimat yang dihitung dari total skor polaritas kata-kata opini negatif, dan digunakan sebagai indikator keberadaan *buzzer*. Sebaliknya, jika skor negatif tidak dominan, maka kalimat tersebut dikategorikan sebagai *non-buzzer*. Oleh karena itu, orientasi sentimen dalam penelitian ini dibagi menjadi dua kategori saja, yaitu negatif sebagai *buzzer* dan positif (termasuk 0) sebagai *non-buzzer*, yang ditentukan berdasarkan dominasi skor polaritas negatif atau positif dalam suatu teks [18].

$$\begin{aligned} Sentence_{sentiment} &\{ \text{positive if } S_{positive} > S_{negative} \} \\ &\text{negative if } S_{positive} < S_{negative} \end{aligned} \quad (3)$$

Apabila suatu teks memiliki jumlah kata negatif yang lebih dominan dibandingkan kata positif, maka teks tersebut akan diklasifikasikan sebagai sentimen negatif yang mengindikasikan keberadaan *buzzer*. Sebaliknya, jika kata

positif lebih banyak dari kata negatif, maka teks akan diberi label sentimen positif yang merepresentasikan *non-buzzer*. Dalam pendekatan ini, tidak digunakan kategori netral, sehingga orientasi sentimen dibatasi hanya pada dua klasifikasi utama, yaitu *buzzer* (negatif) dan *non-buzzer* (positif).

D. Visualisasi

Visualisasi dalam penelitian ini menggunakan diagram lingkaran untuk menggambarkan frekuensi hasil klasifikasi sentimen, yaitu jumlah komentar yang dikategorikan sebagai sentimen positif (*non-buzzer*) dan sentimen negatif (*buzzer*). Diagram lingkaran dipilih karena efektif dalam menunjukkan proporsi masing-masing kategori sentimen secara jelas dan mudah dipahami, sehingga pembaca dapat melihat perbandingan antara komentar yang bersentimen positif dan negatif.

E. Evaluasi Model

Tahapan ini bertujuan untuk menilai performa model *machine learning* yang telah dikonstruksi. Proses evaluasi dilakukan dengan menggunakan confusion matrix, yang merupakan metode umum dalam mengukur kinerja model klasifikasi. Confusion matrix menyajikan hasil klasifikasi dalam bentuk tabel yang menggambarkan jumlah prediksi benar maupun salah pada setiap kelas. Melalui penggunaan confusion matrix, peneliti dapat menilai tingkat kemampuan model dalam mengklasifikasikan data ke dalam kategori yang sesuai, baik pada kasus dengan dua kelas maupun lebih. Selain itu, metode ini juga memungkinkan perhitungan matrik evaluasi seperti akurasi, presisi, recall, dan F1-score [19].

Dalam confusion matrix, hasil prediksi dibagi menjadi empat komponen utama. Prediksi yang benar terhadap kelas positif disebut sebagai *True Positive* (TP), sedangkan jika model salah memprediksi kelas positif, maka disebut *False Positive* (FP). Sebaliknya, jika model berhasil mengidentifikasi kelas negatif dengan benar, hasilnya disebut *True Negative* (TN). Namun, jika model keliru memprediksi kelas negatif sebagai positif, maka hasil tersebut disebut *False Negative* (FN). Keempat komponen ini menjadi dasar dalam menilai kualitas dan keandalan model dalam tugas klasifikasi [20].

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Pengumpulan data dilakukan dengan cara manual dan disimpan ke dalam format .csv untuk memudahkan tahapan selanjutnya. *Scraping* manual dalam hal ini melibatkan pencarian dan pengambilan data secara langsung dari *platform* media sosial TikTok. Proses ini dilakukan dengan menyalin secara manual komentar-komentar yang terkait.

Meskipun membutuhkan waktu dan tenaga yang lebih banyak dibandingkan dengan teknik *scraping* otomatis, metode manual ini memungkinkan peneliti untuk memverifikasi dan memastikan kualitas data yang diperoleh

sebelum digunakan dalam analisis lebih lanjut. Contoh hasil *scraping* data komentar dapat dilihat pada Tabel 2.

TABEL 2
HASIL SCARPING KOMENTAR TIKTOK

| No | Komentar |
|----|---|
| 1. | Aku awal pakek moisnya bagus bngt, trs beli satu paket glowing abis jdinya |
| 2. | Kalo buat nyerahin bisa ka gak glowing aja soalnya kusem nnti KLO glowing aja kyak gimana gitu |
| 3. | Tergantung kecocokannya di kulit wajah klian.. gak semua skincare mahal itu cocok di kulit wajah klian yaa... |
| 4. | Cocok bangtt kk |
| 5. | Iya itu nyerahin iya glowing juga iya kalau mau lebih cerah kayaknya pakai yang paket vitamin c |

Data yang digunakan untuk analisis selanjutnya berasal dari komentar yang ditulis oleh pengguna media sosial TikTok. Karena data tersebut masih berbentuk teks yang tidak terstruktur, maka perlu dilakukan tahapan berikutnya, yaitu *text preprocessing*, guna mempersiapkan data agar siap untuk dianalisis lebih lanjut.

B. Text Preprocessing

Pada tahap ini, data teks yang semula tidak terstruktur akan diubah menjadi format yang terstruktur. Proses ini dilakukan dengan menerapkan metode text mining guna memperoleh informasi yang relevan dan memudahkan analisis lanjutan. Adapun tahapan *text preprocessing* yang digunakan dalam penelitian ini meliputi penghapusan tanda baca (*remove punctuation*), pengubahan seluruh huruf menjadi huruf kecil (*case folding*), penyaringan kata-kata tertentu (*filtering*), serta pemisahan teks menjadi kata-kata tunggal (*tokenizing*).

1) Remove Punctuation

Pada tahap ini, dilakukan penghapusan karakter, tanda baca, serta simbol-simbol yang tidak berkontribusi terhadap proses analisis sentimen. Proses ini mencakup penghilangan tanda titik (.), koma (,), tagar (#), mention (@), tautan URL, serta karakter lain yang bukan merupakan huruf atau angka. Namun, dalam penelitian ini, tanda tanya (?) tidak dihapus karena digunakan untuk mengidentifikasi keberadaan kata tanya secara otomatis, yang diasumsikan sebagai indikasi komentar informatif atau penasaran, dan dengan demikian dikategorikan sebagai *non-buzzer*. Hasil dari proses *remove punctuation* dapat dilihat melalui Tabel 3.

TABEL 3
PROSES REMOVE PUNCTUATION

| Sebelum | Sesudah |
|---|---|
| untuk kulit kusam & kering pake paket yang mana?? ka ? | untuk kulit kusam kering pake paket yang mana?? ka ? |

2) Case Folding

Pada tahap ini, seluruh huruf dalam teks diubah menjadi huruf kecil (*lowercase*) dengan tujuan untuk menyeragamkan format penulisan kata. Hal ini dilakukan agar kata-kata dengan makna sama namun berbeda dalam penulisan huruf

capital, seperti "TIDUR" dan "tidur" yang dapat dilihat dalam tabel dibawah, dikenali sebagai satu entitas yang sama dalam proses analisis sentimen. Hasil dari proses *case folding* dapat dilihat melalui Tabel 4.

TABEL 4
PROSES CASE FOLDING

| Sebelum | Sesudah |
|--|--|
| KAK DI BAWA TIDUR JUGA GAKPAPA KANN | kak di bawa tidur juga gakpapa kann |

3) Filtering

Pada tahap ini dilakukan eliminasi terhadap kata atau teks yang dinilai tidak memberikan pengaruh signifikan terhadap proses analisis sentimen. Dalam penelitian ini, daftar *stopword* yang digunakan diadopsi dari kamus yang tersedia pada *repository* GitHub. Hasil dari proses *filtering* dapat dilihat melalui Tabel 5.

TABEL 5
PROSES FILTERING

| Sebelum | Sesudah |
|---|---|
| kak serum yang pink itu bisa dipakai setiap malam atau 3x saja dlm seminggu ya maaf pemula jd kurang faham | kak serum pink dipakai malem 3x dlm seminggu maaf pemula jd kurang faham |

4) Tokenizing

Pada tahap ini, data teks akan dipisahkan menjadi potongan-potongan kata yang disebut sebagai token. Token-token ini nantinya berperan sebagai entitas dasar yang memiliki nilai dan akan digunakan dalam pembentukan matriks dokumen teks untuk keperluan analisis lebih lanjut. Hasil dari proses *tokenizing* dapat dilihat melalui Tabel 6.

TABEL 6
PROSES TOKENIZING

| Sebelum | Sesudah |
|---------------------------|--|
| dipake nya kapan aja kak? | 'dipake' 'nya' 'kapan' 'aja' 'kak' '?' |

C. Pelabelan Data

Kata-kata yang telah melewati proses *text preprocessing* selanjutnya perlu diberi label agar dapat dianalisis dan dihitung. Skor sentimen akan dihitung berdasarkan jumlah kemunculan kata-kata yang teridentifikasi dalam kamus lexicon. Perhitungan skor dilakukan dengan cara mengurangkan jumlah kata positif dari jumlah kata negatif yang terdapat dalam teks.

$$\text{Skor} = (\text{jumlah kata positif}) - (\text{jumlah kata negatif}) \quad (4)$$

Contoh hasil dari proses pelabelan data dapat dilihat pada Tabel 7, di mana kata-kata yang berhasil dikenali dari kamus lexicon akan diberikan skor, kemudian seluruh skor tersebut dijumlahkan berdasarkan rumus perhitungan yang telah ditentukan. Proses perhitungan skor dapat dilihat melalui Tabel 7.

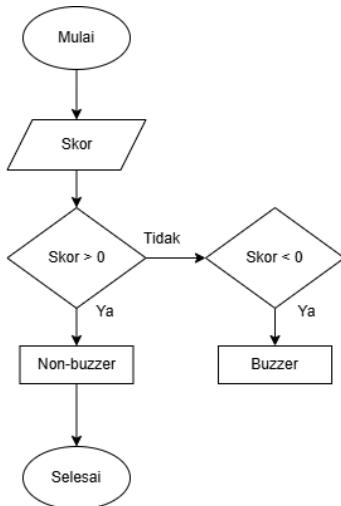
TABEL 7
PERHITUNGAN SKOR

| Teks | Kata Positif | Kata Negatif |
|---|---------------------------|--------------------------------------|
| cleansernya bagus banget pls langsung cerah | cleansernya pls | bagus banget langsung cerah |
| Jumlah | 2 | 4 |
| Perhitungan | Skor = 2 - 4 Skor = -2 | |

Setelah dilakukan perhitungan skor sentimen menggunakan kamus *lexicon*, tahap berikutnya adalah menetapkan label sentimen berdasarkan skor yang diperoleh guna mengidentifikasi kategori sentimen pada setiap komentar.

Pada penelitian ini, pelabelan sentimen dilakukan dalam dua kelas, yaitu sentimen negatif dan sentimen positif. Komentar yang akan diidentifikasi sebagai *buzzer* adalah komentar yang menunjukkan kecenderungan manipulatif atau bersifat hiperbolis untuk menggiring opini konsumen agar membeli suatu produk.

Sentimen dengan hasil skor = 0 (netral) akan dikategorikan Kelas positif diberikan untuk teks dengan skor > 0, sementara kelas negatif diberikan untuk teks dengan skor < 0. Nilai skor 0 dianggap sebagai *non-buzzer*, yang menunjukkan teks tersebut tidak teridentifikasi sebagai *buzzer* atau memiliki sentimen yang netral. Diagram alur klasifikasi sentimen dapat dilihat pada Gambar 3.



Gambar 3. Diagram Alur Klasifikasi Sentimen

Pada komentar yang memuat ungkapan seperti "bagus banget", "cerah seketika", atau "langsung glowing" dapat teridentifikasi sebagai *buzzer* karena mengandung ekspresi berlebihan yang tidak mencerminkan opini objektif. Pemilihan kata yang terlalu positif dan repetitif ini menjadi indikator utama dalam proses pembobotan kata, karena *buzzer* umumnya menggunakan gaya bahasa promosi yang khas dan tidak natural.

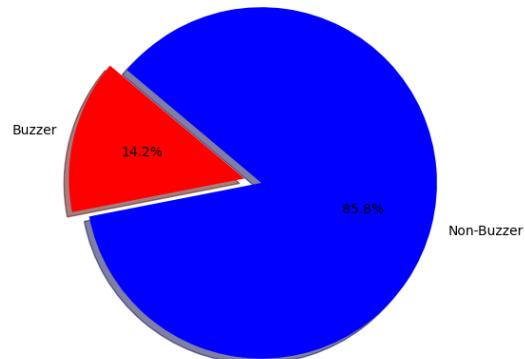
D. Visualisasi Hasil

Setelah seluruh tahapan analisis sentimen selesai dilakukan, hasil pelabelan kelas sentimen pada 529 komentar dalam postingan *review* penggunaan *skincare* ditunjukkan pada Tabel 8.

TABEL 8
HASIL PELABELAN SENTIMEN

| Sentimen | Frekuensi |
|-------------------------------|-----------|
| Negatif (<i>Buzzer</i>) | 75 |
| Positif (<i>Non-Buzzer</i>) | 454 |

Berdasarkan hasil analisis sentimen, dapat dilihat bahwa komentar-komentar dalam postingan *review* penggunaan *skincare* telah terbagi ke dalam beberapa kategori sentimen. Di antaranya, sentimen negatif terdeteksi pada sekitar 75 komentar, sementara sentimen positif ditemukan pada 454 komentar. Visualisasi hasil pelabelan kelas sentimen dalam konteks identifikasi *buzzer* dapat dilihat pada Gambar 4.



Gambar 4. Diagram Lingkaran sebagai Visualisasi Hasil Pelabelan Kelas Sentimen

E. Evaluasi Model

Setelah tahap klasifikasi selesai, langkah berikutnya adalah melakukan evaluasi terhadap model untuk menilai sejauh mana kinerjanya dalam melakukan prediksi. Untuk menampilkan confusion matrix, peneliti memanfaatkan pustaka scikit-learn dan memanggil fungsi matrik. Sementara itu, untuk menampilkan hasil perhitungan yang lebih mendalam, peneliti menggunakan fungsi classification report dari pustaka yang sama [21].

Evaluasi dilakukan dengan memanfaatkan berbagai metrik yang umumnya digunakan dalam klasifikasi teks, antara lain sebagai berikut:

1. Akurasi merupakan salah satu ukuran yang paling sering digunakan dalam mengevaluasi kinerja model. Nilai akurasi diperoleh dengan membandingkan jumlah prediksi yang benar terhadap keseluruhan jumlah data. Metrik ini memberikan gambaran umum mengenai tingkat kemampuan model dalam melakukan prediksi secara keseluruhan.

2. *Precision* (presisi) mengukur seberapa banyak dari seluruh prediksi positif yang benar-benar relevan. Dengan kata lain, dari semua prediksi yang dilakukan oleh model sebagai kelas positif, *precision* menghitung berapa persennya yang benar-benar positif.

3. *Recall* (sensitivitas atau daya ingat) merupakan metrik yang digunakan untuk mengukur kemampuan model dalam mengidentifikasi seluruh data positif yang tersedia. Metrik ini menunjukkan proporsi data positif yang berhasil dikenali oleh model dibandingkan dengan keseluruhan data positif yang sebenarnya terdapat dalam dataset.

4. *F1-score* adalah rata-rata gabungan antara *precision* dan *recall*. Metrik ini memberikan representasi mengenai keseimbangan kinerja keduanya, serta lebih relevan digunakan ketika terdapat ketidakseimbangan jumlah data antara kelas positif dan negatif dalam suatu dataset.

5. *Support* mengacu pada jumlah sampel aktual yang dimiliki oleh setiap kelas dalam dataset. Dengan kata lain, *support* menunjukkan frekuensi atau banyaknya data yang benar-benar termasuk ke dalam masing-masing kategori yang dianalisis [22].

Evaluasi menggunakan metrik-metrik ini penting dilakukan untuk memastikan bahwa model tidak hanya memiliki tingkat akurasi yang tinggi secara umum, tetapi juga efektif dalam mengenali kelas-kelas yang lebih spesifik dan relevan, seperti dalam hal deteksi *buzzer*, di mana keakuratan prediksi terhadap komentar positif dan negatif menjadi sangat penting.

Berdasarkan hasil evaluasi menggunakan fungsi classification report, model klasifikasi yang dikembangkan menghasilkan tingkat akurasi sebesar 93%. Persentase ini menunjukkan bahwa secara keseluruhan, model mampu melakukan prediksi dengan tepat terhadap 93% dari total data uji yang digunakan.

Pada kelas *non-buzzer* (label 0), model menunjukkan kinerja yang sangat baik. Nilai *precision* mencapai 0,94, yang menunjukkan bahwa sebagian besar prediksi model terhadap komentar *non-buzzer* sesuai dengan kenyataan. Nilai *recall* untuk kelas ini bahkan mencapai 0,99, yang berarti hampir seluruh data *non-buzzer* berhasil terdeteksi dengan benar oleh model. Selanjutnya, nilai *F1-score* tercatat sebesar 0,96, yang mengindikasikan keseimbangan prediksi model terhadap kelas ini. Adapun nilai *support* untuk kelas *non-buzzer* adalah 94, yang juga menunjukkan bahwa mayoritas data uji berasal dari kelas tersebut.

Sebaliknya, kinerja model terhadap kelas *buzzer* (label 1) menunjukkan hasil yang lebih rendah. Nilai *precision* sebesar 0,86 menunjukkan bahwa prediksi model terhadap komentar *buzzer* relatif akurat. Akan tetapi, nilai *recall* hanya mencapai 0,50, yang menunjukkan bahwa model hanya mampu mengidentifikasi setengah dari total komentar *buzzer* yang sebenarnya. Akibatnya, *F1-score* untuk kelas ini berada pada angka 0,63. Rendahnya *recall* dan *F1-score* ini menunjukkan bahwa model masih belum cukup sensitif dalam mengenali komentar dari *buzzer*, yang kemungkinan besar disebabkan oleh ketimpangan distribusi data, mengingat jumlah data pada

kelas *buzzer* jauh lebih sedikit (*support* = 12), sehingga memengaruhi efektivitas proses pembelajaran model.

Hasil evaluasi menunjukkan bahwa model memperoleh skor *macro average* sebesar 0,90 untuk *precision*, 0,74 untuk *recall*, dan 0,80 untuk *F1-score*. Nilai ini mencerminkan bahwa meskipun model mampu memberikan prediksi yang cukup tepat secara umum, kemampuannya dalam mengenali seluruh data *buzzer* masih terbatas. Sementara itu, nilai *weighted average* untuk ketiga metrik tersebut masing-masing berada pada angka 0,93. Angka ini dipengaruhi oleh dominasi data *non-buzzer* dalam dataset, sehingga performa keseluruhan tampak tinggi. Dengan demikian, meskipun akurasi model tergolong baik, efektivitasnya dalam mendeteksi komentar *buzzer* sebagai fokus utama penelitian masih perlu ditingkatkan.

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.94 | 0.99 | 0.96 | 94 |
| 1 | 0.86 | 0.50 | 0.63 | 12 |
| accuracy | | | 0.93 | 106 |
| macro avg | 0.90 | 0.74 | 0.80 | 106 |
| weighted avg | 0.93 | 0.93 | 0.93 | 106 |

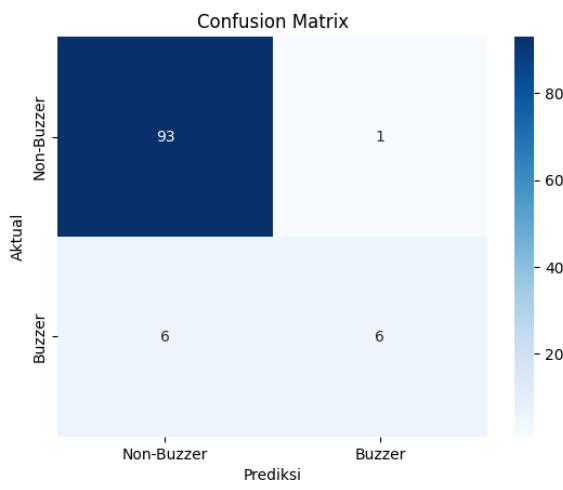
Gambar 5. Classification Report Evaluasi Model

Berdasarkan hasil evaluasi menggunakan fungsi confusion matrix, ditemukan bahwa model memiliki tingkat akurasi yang tinggi dalam mengenali kelas *non-buzzer*, namun memiliki kelemahan dalam mendeteksi kelas *buzzer*.

Untuk kelas *non-buzzer*, model berhasil dengan sangat baik memprediksi komentar yang bukan berasal dari *buzzer*. Dari keseluruhan data *non-buzzer*, model mengklasifikasikan 93 data dengan benar sebagai *non-buzzer* (*True Negative*). Namun, terdapat 1 data yang keliru diprediksi sebagai *buzzer* (*False Positive*), yang menunjukkan adanya sedikit kesalahan dalam klasifikasi kelas *non-buzzer*.

Di sisi lain, untuk kelas *buzzer*, meskipun model berhasil memprediksi 6 data dengan benar sebagai *buzzer* (*True Positive*), terdapat 6 data yang salah diprediksi sebagai *non-buzzer* (*False Negative*). Hal ini menunjukkan bahwa model hanya dapat mengenali sebagian kecil dari data *buzzer* yang sebenarnya ada dalam data uji.

Dari hasil ini, dapat disimpulkan bahwa meskipun model menunjukkan kinerja yang baik dalam mengenali kelas mayoritas (*non-buzzer*), ia mengalami kesulitan dalam mendeteksi kelas minoritas (*buzzer*). Salah satu faktor utama yang menyebabkan hal ini adalah ketidakseimbangan data antara kedua kelas, di mana jumlah data *non-buzzer* jauh lebih banyak dibandingkan dengan data *buzzer*. Ketidakseimbangan ini menyebabkan model lebih mudah untuk mengenali kelas yang lebih dominan, sementara kesulitan dalam mengidentifikasi kelas yang lebih jarang muncul, yaitu *buzzer*.



Gambar 7. Confusion Matrix Evaluasi Model

IV. KESIMPULAN

Berdasarkan hasil analisis terhadap komentar ulasan produk *skincare*, sebagian besar komentar yang ditemukan cenderung menunjukkan opini yang diberikan secara alami tanpa adanya indikasi keterlibatan *buzzer*. Dari total 529 komentar yang dianalisis, sebanyak 75 komentar diidentifikasi memiliki sentimen negatif, sementara 454 komentar lainnya menunjukkan sentimen positif, yang menggambarkan dominasi persepsi positif terhadap produk yang dibahas.

Hasil klasifikasi selanjutnya menunjukkan bahwa 85,8% komentar tergolong dalam kategori *non-buzzer*, yaitu komentar yang dianggap sebagai ulasan asli berdasarkan pengalaman pribadi pengguna. Sebaliknya, sekitar 14,2% komentar dikategorikan sebagai *buzzer*, yang mencerminkan adanya unsur promosi berlebihan, tampak tidak alami, dan cenderung bertujuan untuk mempengaruhi opini publik secara manipulatif.

Meskipun sebagian besar komentar bersifat autentik dan memberikan informasi yang berguna bagi konsumen, ada juga komentar yang berpotensi menyesatkan. Oleh karena itu, sangat penting untuk menggunakan pendekatan yang tepat dalam memverifikasi ulasan, guna memastikan bahwa informasi yang diterima konsumen tetap jujur, dapat dipertanggungjawabkan, dan memberikan gambaran yang akurat tentang produk tersebut.

Sementara itu, hasil evaluasi metode menggunakan SVM menunjukkan tingkat akurasi yang cukup tinggi, mencapai 93%, dalam mengidentifikasi komentar *buzzer*. Namun, meskipun model ini dapat mengklasifikasikan komentar *non-buzzer* dengan sangat baik, kemampuan model untuk mengenali komentar *buzzer* masih tergolong rendah. Metrik *recall* yang hanya mencapai 0,50 menunjukkan bahwa model belum sepenuhnya efektif dalam menangkap semua data *buzzer* yang ada.

Rendahnya performa model dalam mengidentifikasi *buzzer* lebih lanjut dapat dipengaruhi oleh ketidakseimbangan jumlah data, di mana data *buzzer* jauh lebih sedikit

dibandingkan dengan data *non-buzzer*. Dengan total dataset hanya berjumlah 529 komentar, model tidak cukup terlatih untuk mengenali *buzzer* secara optimal. Untuk itu, penelitian di masa depan disarankan untuk menggunakan dataset yang lebih besar dan seimbang antara kelas, agar dapat menghasilkan model klasifikasi yang lebih efektif dan dapat diandalkan.

Selain itu, penggunaan teknik penanganan data tidak seimbang seperti *oversampling* bisa menjadi solusi untuk meningkatkan performa model. Di samping itu, menggali metode klasifikasi lainnya, seperti Random Forest, Naïve Bayes, atau pendekatan *deep learning*, dapat memberikan hasil yang lebih baik dalam mengidentifikasi *buzzer* secara lebih akurat.

Meskipun pendekatan lexicon-based yang digunakan dalam penelitian ini tergolong sederhana dan belum dibandingkan dengan baseline lain seperti metode machine learning atau deep learning, kontribusi penelitian ini terletak pada penerapannya dalam konteks ulasan *skincare*, sebuah domain yang masih jarang dieksplorasi dalam studi deteksi *buzzer*. Penelitian ini juga menekankan evaluasi kritis terhadap keterbatasan lexicon-based pada kondisi ketidakseimbangan data (data imbalance), yang terbukti memengaruhi kemampuan model dalam mendeteksi *buzzer*. Dengan demikian, penelitian ini dapat menjadi pijakan awal yang penting untuk studi-studi selanjutnya yang berfokus pada pengembangan metode lebih lanjut, misalnya dengan pendekatan hybrid atau supervised learning, untuk meningkatkan akurasi dan sensitivitas dalam mengidentifikasi *buzzer* secara lebih komprehensif.

DAFTAR PUSTAKA

- [1] N. A. F. Firdausi, “Buzzer Skincare: Udah Bikin Rugi Konsumen, Jatuhin Brand Lagi,” Terminal Mojok. [Online]. Available: <https://mojok.co/terminal/buzzer-skincare-udah-bikin-rugi-konsumen-jatuhin-brand-lagi/>
- [2] C. Juditha, “Buzzer di Media Sosial pada Pilkada dan Pemilu Indonesia,” *Semin. Nas. Komun. dan Inform.*, vol. 2019, pp. 199–212, 2019.
- [3] U. Nisa, S. M. Amalia, R. Rahmawati, and W. Q. Khotimah, “Peran Buzzer Marketing dalam Melakukan Branding Usaha Mikro, Kecil, dan Menengah,” *Komuniti J. Komun. dan Teknol. Inf.*, vol. 17, no. 1, pp. 75–97, Mar. 2025, doi: 10.23917/komuniti.v17i1.7870.
- [4] N. F. E. Maulidatur Rohmah, “Branding Buzzer: Implementasi dan Implikasi pada Hubungan Masyarakat Pemerintahan,” *Inter Komunika J. Komun.*, vol. 4, no. 2, pp. 135–148, 2019, [Online]. Available: <https://elsam.or.id/pemrofilan-pemilih-dalam-penyebaran-iklan-politik-di-media-sosial-praktik-pada-pemilu-2019-dan-tantangannya-terhadap-privasi>
- [5] A. Dewita, “Pemanfaatan Buzzer Sebagai Media Promosi untuk Mengembangkan Ekonomi Kreatif di Dinas Pariwisata Provinsi Riau,” Universitas Negeri Sultan Syarif Kasim Riau, 2020.
- [6] B.-- Arianto, “Peran Buzzer Media Sosial Dalam Memperkuat Ekosistem Pemasaran Digital,” *Optim. J. Ekon. dan Pembang.*, vol. 10, no. 1, p. 52, Jun. 2020, doi: 10.12928/optimum.v10i1.14343.
- [7] N. A. Fitriyana, N. Suryadi, and R. D. V. Hapsari, “The Role of Social Media and Online Reviews in Building Brand Loyalty Through Trust,” *J. Ilm. Manaj. Kesatuan*, vol. 13, no. 5, pp. 3324–3336, Aug. 2025, doi: 10.37641/jimkes.v13i5.3602.
- [8] A. T. Siswanto, R. R. Muhima, and S. R. Wardhana, “Analisa

- [9] Sentimen Review Produk Handphone Pada Situs Amazon Menggunakan Pendekatan Lexicon Berdasarkan Sentiwordnet,” *KERNEL J. Ris. Inov. Bid. Inform. dan Pendidik. Inform.*, vol. 3, no. 1, pp. 14–19, Sep. 2022, doi: 10.31284/j.kernel.2022.v3i1.1928.
- [10] T. Hendrawati, N. L. W. S. R. Ginantra, and C. M. Saiman, “Analisis Sentimen Larangan Impor Pakaian Bekas Menggunakan Metode Support Vectore Machine dan Lexicon Based,” *TEMATIK*, vol. 11, no. 1, pp. 56–64, Jun. 2024, doi: 10.38204/tematik.v1i1.1890.
- [11] Y. Azhar, “Metode Lexicon-Learning Based Untuk Identifikasi Tweet Opini Berbahasa Indonesia,” *J. Nas. Pendidik. Tek. Inform.*, vol. 6, no. 3, p. 237, Jan. 2018, doi: 10.23887/janapati.v6i3.11739.
- [12] Y. I. Nurhasanah and M. T. Naufal, “Klasifikasi Sentimen Pada Buzzer Politik Jelang Pemilu 2024 menggunakan Metode Lexicon-based,” *MIND J.*, vol. 9, no. 2, pp. 166–178, Dec. 2024, doi: 10.26760/mindjournal.v9i2.166-178.
- [13] B. Pang and L. Lee, “Opinion Mining and Sentiment Analysis,” *Found. Trends® Inf. Retr.*, vol. 2, no. 1–2, pp. 1–135, 2008, doi: 10.1561/1500000011.
- [14] Z. Alhaq, A. Mustopa, S. Mulyatun, and J. D. Santoso, “Penerapan Metode Support Vector Machine Untuk Analisis Sentimen Pengguna Twitter,” *J. Inf. Syst. Manag.*, vol. 3, no. 2, pp. 44–49, Jul. 2021, doi: 10.24076/joism.2021v3i2.558.
- [15] A. M. Yolanda and R. T. Mulya, “Implementasi Metode Support Vector Machine untuk Analisis Sentimen pada Ulasan Aplikasi Sayurbox di Google Play Store,” *VARIANSI J. Stat. Its Appl. Teach. Res.*, vol. 6, no. 2, pp. 76–83, 2024, doi: 10.35580/variansiumn258.
- [16] A. R. Ismail and Raden Bagus Fajriya Hakim, “Implementasi Lexicon Based Untuk Analisis Sentimen Dalam Menentukan Rekomendasi Pantai Di DI Yogyakarta Berdasarkan Data Twitter,” *Emerg. Stat. Data Sci. J.*, vol. 1, no. 1, pp. 37–46, Jan. 2023, doi: 10.20885/esds.vol1.iss.1.art5.
- [17] R. Fauzi Akbar, M. Habibi, P. Winar Cahyo, and N. Alfi Sa'diya, “Metode Hybrid Menggunakan Pendekatan Lexicon Based dan Naive Bayes Classifier Untuk Analisis Sentimen Terkait Jaminan Hari Tua,” *Teknematika J. Inform. dan Komput.*, vol. 16, no. 2, pp. 73–79, Dec. 2023, doi: 10.30989/teknomatika.v16i2.1247.
- [18] S. Khairunnisa, A. Adiwijaya, and S. Al Faraby, “Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19),” *J. MEDIA Inform. BUDIDARMA*, vol. 5, no. 2, p. 406, Apr. 2021, doi: 10.30865/mib.v5i2.2835.
- [19] O. Manullang, C. Prianto, and N. H. Harani, “Analisis Sentimen Untuk Memprediksi Hasil Calon Pemilu Presiden Menggunakan Lexicon Based Dan Random Forest,” *J. Ilm. Inform.*, vol. 11, no. 02, pp. 159–169, 2023, doi: 10.33884/jif.v11i02.7987.
- [20] B. P. Pratiwi, A. S. Handayani, and S. Sarjana, “Pengukuran Kinerja Sistem Kualitas Udara Dengan Teknologi Wsn Menggunakan Confusion Matrix,” *J. Inform. Upgris*, vol. 6, no. 2, Jan. 2021, doi: 10.26877/jiu.v6i2.6552.
- [21] Arif Widiasan Subagio, Anggraini Puspita Sari, and Andreas Nugroho Sihananto, “Klasifikasi Lexicon-Based Sentiment Analysis Tragedi Kanjuruhan pada Twitter Menggunakan Algoritma Convolutional Neural Network,” *J. Ilm. Sist. Inf. dan Ilmu Komput.*, vol. 4, no. 1, pp. 166–177, Jan. 2024, doi: 10.55606/juisik.v4i1.759.
- [22] A. R. Hidayati, A. S. Fitriani, M. A. Rosid, F. Sains, and D. Teknologi, “Analisa Sentimen Pemilu 2019 Pada Judul Berita Online Menggunakan Metode Logistic Regression,” *Kesatria J. Penerapan Sist. Inf. (Komputer dan Manajemen)*, vol. 4, no. 2, pp. 298–305, 2023, [Online]. Available: <http://www.pkm.tunasbangsa.ac.id/index.php/kesatria/article/view/164>
- A. Cardova and A. Hermawan, “Implementasi Metode LSTM Untuk Mengklasifikasi Berita Palsu Pada PolitiFact,” *J. FASILKOM*, vol. 13, no. 3, pp. 471–479, Dec. 2023, doi: 10.37859/jf.v13i3.6175.