

Klasifikasi Opini Pengguna TikTok terhadap Keamanan dan Efektivitas Produk Skincare Lokal menggunakan Metode Naïve Bayes, Decision Tree, dan Random Fores

Sintia Ariyani, Styawati*

Fakultas Teknik dan Ilmu Komputer, Program Studi Sistem Informasi, Universitas Teknokrat Indonesia, Bandar Lampung, Indonesia

Email: ¹sintia_ariyani@teknokrat.ac.id, ^{2,*}styawati@teknokrat.ac.id

Email Penulis Korespondensi: styawati@teknokrat.ac.id

Submitted: 14/03/2026; Accepted: 31/03/2026; Published: 31/03/2026

Abstrak—Penelitian ini bertujuan untuk menganalisis dan membandingkan kinerja algoritma Naïve Bayes, Decision Tree, dan Random Forest dalam mengklasifikasikan opini pengguna TikTok terkait keamanan dan efektivitas produk skincare lokal. Berdasarkan hasil penelitian, algoritma Naïve Bayes, Decision Tree, dan Random Forest menunjukkan perbedaan kinerja yang signifikan dalam mengklasifikasikan opini pengguna TikTok mengenai keamanan dan efektivitas produk skincare lokal. Sebelum penerapan SMOTE, Random Forest memperoleh akurasi tertinggi sebesar 87%, diikuti Decision Tree sebesar 79% dan Naïve Bayes sebesar 65%, dengan kelemahan utama terdapat pada kelas minoritas seperti Aman dan Tidak Aman yang memiliki nilai recall rendah. Setelah penerapan SMOTE, ketiga algoritma mengalami peningkatan performa, terutama dalam mengenali kelas minoritas, sehingga nilai akurasi, presisi, recall, dan F1-score menjadi lebih seimbang pada seluruh kelas sentimen. Hasil analisis TF-IDF menunjukkan bahwa fitur yang terbentuk masih didominasi oleh kata-kata umum dan nama merek, sehingga belum sepenuhnya mampu merepresentasikan aspek keamanan dan efektivitas secara spesifik. Hal ini mengindikasikan bahwa proses preprocessing dan seleksi fitur masih dapat ditingkatkan agar menghasilkan representasi yang lebih relevan. Visualisasi hasil klasifikasi menunjukkan bahwa mayoritas komentar berada pada kategori Efektif dan Tidak Efektif, sementara kategori Netral memiliki jumlah yang lebih sedikit. Penerapan SMOTE memberikan peningkatan performa model, khususnya dalam menangani ketidakseimbangan data, namun penerapannya perlu dilakukan secara tepat pada data latih (training data) untuk menghindari bias evaluasi. Secara keseluruhan, Random Forest menunjukkan performa terbaik dibandingkan algoritma lainnya. Penelitian ini berkontribusi dalam pengembangan model analisis sentimen multi-kelas yang mampu membedakan aspek keamanan dan efektivitas produk skincare, serta menunjukkan bahwa penerapan SMOTE efektif dalam meningkatkan performa klasifikasi pada dataset yang tidak seimbang. Penelitian selanjutnya disarankan untuk meningkatkan metode pelabelan sentimen, memperbaiki kualitas fitur, serta mengeksplorasi algoritma yang lebih kompleks seperti deep learning agar hasil klasifikasi menjadi lebih akurat dan lebih robust.

Kata Kunci: TikTok; Skincare Lokal; Analisis Sentimen; Machine Learning; Random Forest; Naïve Bayes; Decision Tree; SMOTE; TF-IDF

Abstract—This study aims to analyze and compare the performance of Naïve Bayes, Decision Tree, and Random Forest algorithms in classifying TikTok users' opinions regarding the safety and effectiveness of local skincare products. The results show that these algorithms exhibit significant differences in performance for sentiment classification tasks. Before applying SMOTE, Random Forest achieved the highest accuracy of 87%, followed by Decision Tree at 79% and Naïve Bayes at 65%. The main weakness was observed in minority classes such as *Safe* and *Unsafe*, which had low recall values. After applying SMOTE, all models showed improved performance, particularly in recognizing minority classes, resulting in more balanced accuracy, precision, recall, and F1-score across all sentiment categories. The TF-IDF analysis revealed that the extracted features were still dominated by common words and brand names, indicating that they did not fully represent the specific aspects of safety and effectiveness. This suggests that the preprocessing and feature selection stages can be further improved to generate more relevant feature representations. The classification visualization showed that most comments were categorized as *Effective* and *Ineffective*, while the *Neutral* category contained fewer instances. The implementation of SMOTE improved model performance in handling imbalanced data; however, it must be applied carefully only to the training data to avoid evaluation bias. Overall, Random Forest demonstrated the best performance among the evaluated algorithms. This study contributes to the development of a multi-class sentiment analysis model capable of distinguishing between safety and effectiveness aspects of skincare products, and demonstrates that the application of SMOTE effectively improves classification performance on imbalanced datasets. Future research is recommended to enhance sentiment labeling methods, improve feature quality, and explore more advanced approaches such as deep learning to achieve more accurate and robust classification results.

Keywords: TikTok; Local Skincare; Sentiment Analysis; Machine Learning; Random Forest; Naïve Bayes; Decision Tree; SMOTE; TF-IDF

1. PENDAHULUAN

Industri skincare di Indonesia menunjukkan pertumbuhan signifikan, dengan nilai pasar mencapai Rp45,2 triliun pada tahun 2020 dan peningkatan jumlah perusahaan kosmetik sebesar 21,9% dalam satu tahun terakhir [1],[2]. Pertumbuhan ini diiringi transformasi digital, di mana TikTok, dengan 157,6 juta pengguna aktif di Indonesia, menjadi platform dominan dalam distribusi opini real-time mengenai keamanan dan efektivitas produk skincare lokal [3],[4]. Karakteristik konten video pendek memungkinkan pengguna memberikan testimoni spontan terkait efek samping, kandungan bahan, maupun klaim produk, sehingga menghasilkan data tidak terstruktur yang mengandung beragam sentimen. Aspek keamanan dan efektivitas menjadi sangat penting untuk diteliti karena produk skincare diaplikasikan langsung pada kulit, sehingga berpotensi menimbulkan reaksi seperti iritasi, alergi, atau ketidakcocokan pada jenis

kulit tertentu. Selain itu, efektivitas produk berkaitan dengan kemampuan produk memberikan hasil sesuai klaim, seperti mencerahkan, melembapkan, atau mengatasi permasalahan kulit. Dengan demikian, analisis sentimen yang memisahkan aspek keamanan dan efektivitas diperlukan untuk memberikan wawasan yang lebih mendalam dibandingkan klasifikasi sentimen umum (positif/negatif/netral)[3],[5].

Meskipun pertumbuhan industri skincare lokal dan penggunaan platform TikTok sebagai media berbagi opini semakin meningkat, terdapat permasalahan utama dalam memahami persepsi pengguna secara terstruktur. Banyaknya komentar yang tersebar dalam bentuk teks tidak terstruktur, penggunaan bahasa informal, singkatan, serta ekspresi sarkasme menyebabkan kesulitan dalam melakukan analisis secara manual. Selain itu, sebagian besar penelitian analisis sentimen masih berfokus pada klasifikasi umum seperti positif, negatif, dan netral, sehingga belum mampu menggambarkan secara spesifik persepsi pengguna terhadap aspek keamanan dan efektivitas produk skincare. Permasalahan lain yang muncul adalah ketidakseimbangan jumlah data pada setiap kategori sentimen yang dapat memengaruhi kinerja model klasifikasi. Oleh karena itu, diperlukan pendekatan analisis yang mampu mengklasifikasikan opini secara lebih spesifik dan mengatasi ketidakseimbangan data agar hasil analisis menjadi lebih akurat dan representatif

Namun demikian, volume komentar yang besar serta penggunaan bahasa informal, singkatan, dan sarkasme menimbulkan tantangan dalam pengolahan data secara manual [1],[6]. Penelitian terdahulu umumnya berfokus pada perbandingan algoritma klasifikasi seperti Penelitian terdahulu umumnya berfokus pada perbandingan algoritma klasifikasi. Misalnya, Liem et al. (2024) menunjukkan bahwa algoritma SVM dan Naïve Bayes berbasis SMOTE dapat meningkatkan kinerja klasifikasi pada data yang tidak seimbang[6]. Selain itu, penelitian oleh Rama Dani (2025) memanfaatkan algoritma Naïve Bayes, SVM, dan Random Forest untuk menganalisis sentimen komentar di YouTube, dan menunjukkan bahwa algoritma tersebut efektif dalam mengolah data teks tidak terstruktur[7]. Sebagian besar studi masih mengklasifikasikan sentimen secara umum (positif/negatif) dan belum memisahkan opini berdasarkan aspek keamanan dan efektivitas produk. Selain itu, meskipun Decision Tree memiliki interpretabilitas tinggi, penggunaannya dalam konteks analisis sentimen produk skincare di TikTok masih terbatas, sehingga evaluasi lebih lanjut diperlukan. Permasalahan ketidakseimbangan data (imbalanced class) juga menjadi faktor penting yang memengaruhi kinerja model klasifikasi [3],[7],[8].

Berdasarkan kondisi tersebut, penelitian ini bertujuan untuk mengembangkan dan membandingkan model klasifikasi multi-kelas opini pengguna TikTok terhadap keamanan dan efektivitas produk skincare lokal menggunakan algoritma Naïve Bayes, Decision Tree, dan Random Forest. Pemilihan ketiga algoritma ini didasarkan pada karakteristik masing-masing: Naïve Bayes memiliki performa baik pada data teks dengan proses komputasi sederhana, Decision Tree menawarkan interpretabilitas tinggi sehingga mudah dipahami, dan Random Forest mampu meningkatkan akurasi serta mengurangi overfitting melalui ensemble. Tahapan penelitian meliputi pengumpulan data komentar melalui web scraping, *preprocessing* teks (*case folding*, *tokenization*, *Stopword Removal*, dan *Stemming* menggunakan *Sastrawi*), ekstraksi fitur dengan TF-IDF, serta proses klasifikasi dengan pembagian data *training* dan *testing* sebesar 80:20. Evaluasi model dilakukan menggunakan *confusion matrix* serta metrik akurasi, presisi, *recall*, dan *F1-score*. Penelitian ini diharapkan dapat memberikan rekomendasi algoritma yang optimal untuk memantau opini publik secara otomatis, membantu industri skincare lokal dalam merumuskan strategi berbasis data, serta menyediakan sistem pendukung keputusan yang objektif dengan mempertimbangkan nuansa keamanan dan efektivitas produk berdasarkan ulasan pengguna di media sosial. Kontribusi utama penelitian ini adalah mengusulkan pendekatan klasifikasi sentimen multi-kelas yang mampu membedakan aspek keamanan dan efektivitas produk skincare lokal berbasis data TikTok, serta melakukan evaluasi komparatif terhadap tiga algoritma machine learning dengan penerapan SMOTE untuk meningkatkan performa klasifikasi pada data yang tidak seimbang.

2. METODOLOGI PENELITIAN

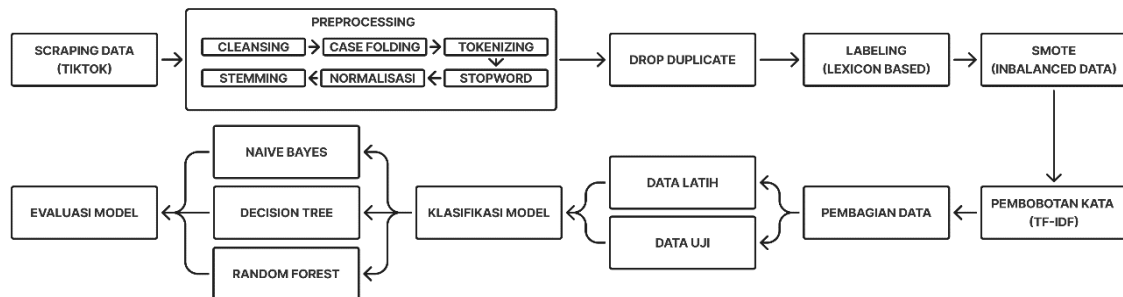
2.1. Tahapan Penelitian

Penelitian ini dilaksanakan melalui tahapan yang sistematis guna memastikan hasil yang diperoleh valid dan dapat dipertanggungjawabkan. Setiap tahapan dirancang untuk menghasilkan keluaran (output) yang jelas sehingga proses penelitian dapat ditelusuri secara ilmiah. Penelitian ini mengadopsi standar proses *Cross-Industry Standard Process for Data Mining* [9]. Alur tahapan penelitian yang digunakan dalam studi ini dapat dilihat pada Gambar 1, yang menggambarkan keterkaitan antar proses mulai dari pemahaman bisnis hingga evaluasi model.

Berdasarkan Gambar 1, alur penelitian ini dimulai dari pengambilan data komentar TikTok melalui teknik *web scraping* untuk memperoleh data opini pengguna secara real-time. Data yang telah diperoleh kemudian dilakukan *preprocessing* yang meliputi *cleansing*, *case folding*, *tokenizing*, *stopword removal*, normalisasi teks, dan *stemming* guna membersihkan data dari noise. Tahapan ini bertujuan untuk menghasilkan data yang bersih dan konsisten sebelum dilakukan analisis lebih lanjut [10]. Setelah proses pembersihan teks, dilakukan penghapusan data duplikat (*drop duplicate*) untuk memastikan keunikan setiap sampel data.

Selanjutnya, data diberi label sentimen menggunakan pendekatan *lexicon-based*. Untuk mengatasi ketidakseimbangan kelas (*imbalanced data*), dilakukan penyeimbangan data menggunakan metode *SMOTE*. Data yang telah seimbang kemudian dibobotkan menggunakan metode *Term Frequency-Inverse Document Frequency (TF-IDF)*, lalu dibagi menjadi data latih dan data uji. Proses klasifikasi dilakukan menggunakan tiga algoritma, yaitu Naïve

Bayes, Decision Tree, dan Random Forest. Hasil klasifikasi kemudian dievaluasi menggunakan metrik performa untuk mengetahui algoritma terbaik dalam mengklasifikasikan opini pengguna TikTok terhadap keamanan dan efektivitas produk skincare lokal.



Gambar 1. Tahapan Penelitian

2.2. Pengumpulan Data

Data diperoleh dari komentar pengguna TikTok melalui teknik *web scraping* dengan total sebanyak 8.082 baris komentar yang dikumpulkan dari berbagai video yang membahas produk *skincare* lokal Indonesia. Proses pengambilan data dilakukan menggunakan bahasa pemrograman Python untuk mengakses dan mengekstraksi data komentar secara terstruktur dari platform TikTok [11]. Informasi mengenai akun TikTok, topik video, tanggal unggah, jumlah tayangan, serta jumlah komentar dari masing-masing video disajikan pada Tabel 1. Seluruh komentar dikumpulkan sejak tanggal unggah masing-masing video sehingga mencakup interaksi pengguna pada periode September 2023 hingga Februari 2026.

Tabel 1. Akun TikTok Sumber Data

No.	Akun TikTok	Topik Video	Tanggal Unggah	Jumlah Tayangan	Jumlah Komentar
1	@it's syee	Rekomendasi Skincare Lokal Yang Jarang Diketahui	14 Feb 2023	28.800 x ditonton	41
2	@MaharajaSP	Rekomendasi Serum Pencerah Wajah	28 Des 2024	427.3000 x ditonton	253
3	@ Skincare by Jessica	Perbandingan Kandungan Niacinamide dan Vitamin C	11 Feb 2025	145.400 x ditonton	101
4	@Fernando Fen Lee	Perbandingan Skintific 377 Moisturizer dan Skintific Brightening Moisturizer	23 Maret 2025	15.400.000 x ditonton	5.032
5	@Abigel	Rekomendasi Skincare Jerawat Untuk Remaja	4 Apr 2025	99.000 x ditonton	94
6	@Claudia Chirstin	Rating Skincare Local	18 Jan 2026	246.400.000 x ditonton	269
7	@	Riview Paket Skincare	10 Feb 2026	337.00.000 x ditonton	601

Berdasarkan Tabel 1 yang menampilkan daftar video TikTok dari berbagai akun yang membahas produk *skincare* lokal Indonesia pada periode Februari 2023–Februari 2026. Setiap video memiliki tingkat perhatian publik yang berbeda, terlihat dari jumlah tayangan dan komentar yang cukup tinggi, misalnya video dari @Fernando Fen Lee yang mencapai lebih dari 15,4 juta tayangan dengan 5.032 komentar, serta video review paket *skincare* yang menembus 337 ribu tayangan dengan 601 komentar. Hal ini menunjukkan bahwa topik keamanan dan efektivitas *skincare* lokal memicu diskusi luas dan respon signifikan dari masyarakat, baik dalam bentuk penayangan maupun komentar di ruang publik digital.

2.3. Preprocessing

Tahap *preprocessing* merupakan proses awal dalam pengolahan data teks yang bertujuan untuk menyiapkan data agar lebih terstruktur dan mudah dianalisis oleh sistem. Data komentar yang diperoleh dari media sosial umumnya masih mengandung berbagai bentuk gangguan seperti simbol, singkatan, karakter khusus, maupun kata-kata yang tidak relevan. Oleh karena itu, diperlukan proses pembersihan dan normalisasi data agar kualitas data meningkat serta dapat menghasilkan performa model klasifikasi yang lebih baik[12]. Pada penelitian ini, proses *preprocessing* dilakukan melalui beberapa tahapan utama yaitu *cleansing*, *case folding*, *tokenizing*, *Stopword Removal*, normalisasi teks, dan *Stemming*. Setiap tahapan memiliki fungsi yang berbeda dalam mempersiapkan data teks sebelum masuk ke tahap ekstraksi fitur dan proses klasifikasi. Setiap tahapan dilakukan secara berurutan untuk memastikan data yang dihasilkan menjadi bersih, konsisten, dan siap digunakan pada tahap analisis selanjutnya[12].



a. *Cleansing*

Cleansing merupakan proses pembersihan data dengan menghilangkan berbagai karakter yang tidak memiliki makna dalam analisis teks. Tahap ini meliputi penghapusan tanda baca, angka, simbol, tautan (*URL*), *hashtag* (#), mention (@*username*), emoji, serta karakter lain yang tidak termasuk huruf alfabet. Proses ini dilakukan agar teks yang dianalisis menjadi lebih bersih dan fokus pada kata-kata yang memiliki makna penting dalam komentar pengguna[13].

b. *Case Folding*

Case Folding adalah proses penyamaan bentuk huruf dengan mengubah seluruh teks menjadi huruf kecil (*lowercase*). Langkah ini bertujuan untuk menghindari perbedaan representasi kata yang sebenarnya memiliki arti yang sama, misalnya kata “Bagus” dan “bagus”. Dengan melakukan *Case Folding*, sistem dapat mengenali kata-kata tersebut sebagai satu bentuk yang sama sehingga analisis menjadi lebih konsisten[14].

c. *Tokenizing*

Tokenizing merupakan proses memecah kalimat atau dokumen teks menjadi unit kata yang lebih kecil yang disebut token. Proses ini bertujuan untuk memisahkan setiap kata dalam teks sehingga dapat dianalisis secara individual oleh sistem. Hasil dari proses *Tokenizing* akan digunakan sebagai dasar dalam pembentukan fitur pada tahap selanjutnya[15].

d. *Stopword Removal*

Stopword Removal adalah proses penghapusan kata-kata umum yang sering muncul dalam teks tetapi tidak memiliki kontribusi signifikan terhadap makna atau sentimen suatu kalimat. Contoh kata yang termasuk *stopword* antara lain “dan”, “yang”, “di”, “ke”, serta kata-kata umum lainnya. Penghapusan kata tersebut dilakukan untuk mengurangi jumlah kata yang tidak penting sehingga proses analisis menjadi lebih efektif[16].

e. *Normalisasi Teks*

Normalisasi teks merupakan proses penyesuaian kata tidak baku atau kata yang sering disingkat pada media sosial menjadi bentuk kata yang lebih standar. Sebagai contoh, kata seperti “gk” diubah menjadi “tidak”, “bgt” menjadi “banget”, atau “bagusss” menjadi “bagus”. Proses ini bertujuan untuk meningkatkan konsistensi data sehingga mempermudah sistem dalam mengenali makna kata yang sebenarnya[17].

f. *Stemming*

Stemming adalah proses mengubah kata yang memiliki imbuhan menjadi bentuk kata dasar dengan cara menghilangkan awalan, akhiran, maupun kombinasi imbuhan lainnya. Dalam penelitian ini, proses *Stemming* dilakukan menggunakan library *Sastrawi* yang dirancang khusus untuk pemrosesan bahasa Indonesia sehingga mampu menghasilkan bentuk kata dasar secara lebih akurat[18].

Setelah seluruh tahapan preprocessing selesai dilakukan, data kemudian melalui proses penghapusan duplikat (*drop duplicate*) untuk memastikan tidak terdapat data yang berulang yang dapat memengaruhi hasil analisis. Proses ini penting untuk menghindari bias akibat pengulangan data yang sama dalam jumlah besar. Selanjutnya, data yang telah bersih dan unik diberi label sentimen menggunakan pendekatan *lexicon-based*, di mana setiap teks diklasifikasikan berdasarkan polaritas sentimen yang terkandung di dalamnya. Untuk mengatasi ketidakseimbangan kelas (*imbalanced data*), dilakukan penyeimbangan data menggunakan metode SMOTE. Tahap ini bertujuan untuk meningkatkan representasi kelas minoritas sehingga model klasifikasi tidak bias terhadap kelas mayoritas. Data yang telah seimbang kemudian siap untuk masuk ke tahap ekstraksi fitur menggunakan TF-IDF dan proses klasifikasi.

2.4. Drop Duplicate

Tahap pelabelan dilakukan untuk memberikan kategori sentimen pada setiap komentar yang telah melalui proses *preprocessing*. Pada penelitian ini, proses pelabelan menggunakan pendekatan *lexicon-based* dengan memanfaatkan kamus sentimen Bahasa Indonesia seperti *InSet (Indonesian Sentiment Lexicon)* yang berisi kumpulan kata dengan nilai sentimen tertentu. Setiap kata pada komentar akan dicocokkan dengan kamus *lexicon*, kemudian skor sentimen dihitung dengan menjumlahkan bobot dari setiap kata yang ditemukan. Nilai skor tersebut selanjutnya digunakan untuk menentukan kategori sentimen dari komentar pengguna[19]. Berbeda dengan klasifikasi sentimen pada umumnya yang hanya membagi sentimen menjadi positif, negatif, dan netral, penelitian ini mengelompokkan opini pengguna ke dalam lima kategori yang lebih spesifik, yaitu aman, tidak aman, efektif, tidak efektif, dan netral. Kategori ini digunakan untuk menggambarkan persepsi pengguna terhadap aspek keamanan dan efektivitas produk *skincare* lokal yang dibahas dalam video TikTok. Penentuan label sentimen dilakukan berdasarkan nilai skor sentimen yang dihasilkan dari proses perhitungan *lexicon*. Aturan klasifikasi sentimen tersebut ditunjukkan pada Persamaan (1)[19].

$$S_{sentimen} = \begin{cases} \text{Aman, skor} > 0 \text{ dan berkaitan dengan aspek keamanan} \\ \text{Tidak Aman, skor} < 0 \text{ dan berkaitan dengan aspek keamanan} \\ \text{Efektif, skor} > 0 \text{ dan berkaitan dengan efektivitas produk} \\ \text{Tidak/Efektif, skor} < 0 \text{ dan berkaitan dengan efektivitas produk} \\ \text{Netral, skor} = 0 \end{cases} \quad (1)$$

Persamaan (1) menunjukkan bahwa $S_{sentimen}$ merupakan label sentimen yang diperoleh berdasarkan hasil perhitungan skor *lexicon* pada setiap komentar. Skor tersebut dihitung dari penjumlahan bobot kata yang terdapat dalam komentar pengguna. Apabila skor menunjukkan kecenderungan positif terhadap aspek keamanan maka komentar dikategorikan sebagai aman, sedangkan jika menunjukkan kecenderungan negatif terhadap

keamanan maka dikategorikan sebagai tidak aman. Sementara itu, apabila komentar berkaitan dengan hasil atau kinerja produk maka akan dikategorikan sebagai efektif atau tidak efektif. Jika skor yang dihasilkan tidak menunjukkan kecenderungan tertentu maka komentar diklasifikasikan sebagai netral.

2.5. SMOTE (Synthetic Minority Over-sampling Technique)

SMOTE (*Synthetic Minority Over-sampling Technique*) merupakan metode yang digunakan untuk mengatasi permasalahan ketidakseimbangan kelas pada dataset. Ketidakseimbangan data terjadi ketika jumlah data pada suatu kelas jauh lebih sedikit dibandingkan kelas lainnya sehingga dapat mempengaruhi kinerja model klasifikasi. SMOTE bekerja dengan cara menghasilkan data sintesis pada kelas minoritas berdasarkan kemiripan dengan data yang sudah ada. Teknik ini membuat sampel baru dari kelas minoritas sehingga distribusi data menjadi lebih seimbang. Dengan menambahkan data sintesis tersebut, model pembelajaran mesin dapat mempelajari pola dari setiap kelas secara lebih baik dan mengurangi bias terhadap kelas mayoritas[20]. Dalam penelitian ini, dataset terlebih dahulu dibagi menjadi data latih dan data uji dengan perbandingan 80:20. Metode SMOTE hanya diterapkan pada data latih untuk menghindari terjadinya data leakage dan memastikan evaluasi model dilakukan secara objektif[20]. Dengan distribusi data latih yang lebih seimbang, model diharapkan dapat memberikan performa klasifikasi yang lebih akurat serta memiliki kemampuan generalisasi yang lebih baik terhadap data uji.

Penggunaan metode SMOTE telah terbukti efektif dalam meningkatkan performa model klasifikasi pada dataset yang tidak seimbang karena mampu menghasilkan distribusi data yang lebih proporsional selama proses pelatihan[20]. Oleh karena itu, SMOTE sering digunakan dalam tahap preprocessing pada penelitian *data mining* dan *machine learning* untuk meningkatkan kualitas hasil klasifikasi[20].

2.6. Pembobotan Kata (TF-IDF)

Setelah proses pelabelan dan penyeimbangan data selesai dilakukan, tahap selanjutnya adalah ekstraksi fitur menggunakan metode *TF-IDF* (*Term Frequency–Inverse Document Frequency*). Metode ini merupakan teknik pembobotan kata yang umum digunakan dalam pengolahan bahasa alami untuk mengubah data teks menjadi representasi numerik yang dapat diproses oleh algoritma machine learning. *TF-IDF* bekerja dengan menghitung dua komponen utama, yaitu *Term Frequency (TF)* dan *Inverse Document Frequency (IDF)*. *Term Frequency* mengukur seberapa sering suatu kata muncul dalam dokumen tertentu, sedangkan *Inverse Document Frequency* mengukur seberapa penting kata tersebut dengan melihat seberapa jarang kata itu muncul di seluruh koleksi dokumen. Kombinasi kedua nilai ini menghasilkan bobot yang mencerminkan tingkat relevansi suatu kata dalam dokumen[21]. Perhitungan *TF-IDF* dirumuskan dalam Persamaan (2), (3), dan (4) berikut:

$$TF(t, d) = f_{t,d} / \sum_k f_{k,d} \quad (2)$$

$$IDF(t, D) = \log \left(\frac{N}{|\{d \in D : t \in d\}|} \right) \quad (3)$$

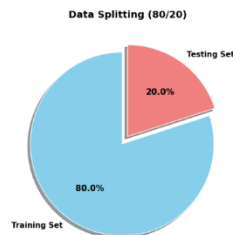
$$TF-IDF(t, d, D) = TF(t, d) \times IDF(t, D) \quad (4)$$

Keterangan: $f_{t,d}$ adalah frekuensi kemunculan kata t dalam dokumen d , $\sum_k f_{k,d}$ adalah jumlah seluruh kata dalam dokumen d , N adalah jumlah total dokumen dalam koleksi, dan $|\{d \in D : t \in d\}|$ adalah jumlah dokumen yang mengandung kata t . Penggunaan *TF-IDF* dalam penelitian ini bertujuan untuk memberikan bobot yang lebih tinggi pada kata-kata yang memiliki nilai informatif tinggi dan muncul secara spesifik dalam konteks tertentu, sementara kata-kata yang terlalu umum akan mendapatkan bobot lebih rendah. Hasil dari proses pembobotan ini berupa vektor fitur yang akan digunakan sebagai input untuk proses klasifikasi[21]. Selain itu, penerapan *TF-IDF* membantu mengubah data teks yang tidak terstruktur menjadi representasi numerik yang lebih sistematis. Metode ini mampu menyeleksi kata-kata penting dengan menekan pengaruh kata yang terlalu umum dalam dokumen. Proses pembobotan ini membuat setiap kata memiliki nilai yang merepresentasikan tingkat kepentingannya dalam suatu dokumen. Dengan demikian, kata yang sering muncul dalam satu dokumen namun jarang di dokumen lain akan memiliki bobot lebih tinggi. Sebaliknya, kata yang sering muncul di banyak dokumen akan memiliki bobot yang lebih rendah. Representasi vektor yang dihasilkan memungkinkan algoritma machine learning mengenali pola secara lebih efektif. Hal ini juga membantu dalam meningkatkan kemampuan model dalam membedakan antar kelas atau kategori. Selain itu, penggunaan *TF-IDF* dapat mengurangi noise dalam data teks yang digunakan. Dengan fitur yang lebih relevan, proses klasifikasi dapat berjalan dengan lebih akurat dan efisien. Oleh karena itu, *TF-IDF* menjadi metode yang tepat digunakan dalam tahap ekstraksi fitur pada penelitian ini.

2.7. Pembagian Data (Data Splitting)

Setelah proses pembobotan *TF-IDF* selesai dilakukan, tahap berikutnya dalam penelitian ini adalah membagi dataset menjadi dua bagian, yaitu data latih (*training data*) dan data uji (*testing data*). Pembagian ini bertujuan untuk mengevaluasi kinerja model klasifikasi yang akan dibangun secara objektif, menggunakan data yang belum pernah diproses oleh model sebelumnya[22]. Pada penelitian ini, pembagian data dilakukan dengan rasio 80:20, di mana 80% dari total dataset digunakan sebagai data latih untuk melatih dan menyesuaikan model, sedangkan 20% sisanya digunakan sebagai data uji untuk mengevaluasi kinerja model pada data baru[23]. Pemilihan rasio ini didasarkan pada pertimbangan bahwa jumlah data yang lebih besar untuk proses pelatihan akan membantu model dalam mempelajari

pola dan karakteristik data dengan lebih baik, sementara proporsi data uji yang memadai diperlukan untuk melakukan evaluasi yang objektif terhadap kemampuan generalisasi model. Proses pembagian data dilakukan secara acak (*random splitting*) untuk memastikan distribusi data pada setiap kelas sentimen terwakili secara proporsional pada kedua subset data. Hal ini penting untuk menghindari bias selama proses pelatihan dan pengujian model, sehingga hasil evaluasi dapat mencerminkan kemampuan model dalam memprediksi data yang belum pernah dilihat sebelumnya[24]. Distribusi pembagian data latih dan data uji ditampilkan pada Gambar 2. Dari total dataset, 80% dialokasikan sebagai data latih untuk melatih dan menyesuaikan model, sedangkan 20% sisanya digunakan sebagai data uji untuk mengukur kemampuan model dalam mengklasifikasikan data baru. Pembagian acak ini memastikan setiap kelas sentimen (misal: *Aman, Tidak Aman, Efektif, Tidak Efektif, Netral*) memiliki representasi yang seimbang pada kedua subset data, sehingga model tidak mempelajari pola yang bias dan hasil evaluasi menjadi lebih valid. Untuk memberikan gambaran visual mengenai proses pembagian dataset, distribusi data latih dan data uji ditampilkan pada Gambar 2.



Gambar 2. Pembagian Data Latih & Uji

Berdasarkan Gambar 2, terlihat bahwa dataset dibagi menjadi dua bagian, yaitu 80% sebagai data latih dan 20% sebagai data uji. Pembagian ini dilakukan secara acak sehingga setiap kelas sentimen, seperti Aman, Tidak Aman, Efektif, Tidak Efektif, dan Netral, memiliki distribusi yang seimbang pada kedua subset data. Hal ini bertujuan agar model tidak mengalami bias dalam proses pelatihan serta mampu menghasilkan evaluasi yang lebih akurat dan representatif. Data latih akan digunakan untuk melatih ketiga algoritma klasifikasi yang digunakan dalam penelitian ini, yaitu Naïve Bayes, Decision Tree, dan Random Forest. Sementara itu, data uji digunakan untuk menguji performa model yang telah dilatih menggunakan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score. Hasil evaluasi ini akan menunjukkan algoritma mana yang paling optimal dalam mengklasifikasikan opini pengguna TikTok terhadap keamanan dan efektivitas produk skincare lokal.

2.6. Klasifikasi Naïve Bayes

Naïve Bayes adalah algoritma klasifikasi yang bekerja berdasarkan Teorema Bayes, dengan asumsi bahwa fitur-fitur bersifat independen satu sama lain. Dalam konteks analisis sentimen, setiap kata dianggap memberikan kontribusi tersendiri terhadap prediksi kelas tanpa dipengaruhi oleh kata lain. Pendekatan ini sangat cocok untuk data teks berdimensi tinggi seperti hasil ekstraksi *TF-IDF* karena komputasinya efisien dan mampu bekerja dengan baik pada dataset besar [25]. Rumus dasar Teorema Bayes yang digunakan oleh algoritma ini adalah:

$$P(H | X) = \frac{P(X|H) \cdot P(H)}{P(X)} \tag{5}$$

Keterangan: $P(H|X)$ adalah probabilitas posterior hipotesis H diberikan fitur X , $P(X|H)$ adalah probabilitas likelihood bahwa fitur X muncul jika H benar, $P(H)$ adalah probabilitas prior dari kelas H , dan $P(X)$ adalah probabilitas total dari fitur X . Pada penelitian ini, hipotesis H merepresentasikan kelas sentimen (*Aman, Tidak Aman, Efektif, Tidak Efektif, dan Netral*), sedangkan X merupakan vektor fitur hasil pembobotan *TF-IDF*. Sistem akan memilih kelas dengan nilai $P(H|X)$ tertinggi sebagai hasil prediksi akhir.

2.7 Klasifikasi Decision Tree

Decision Tree adalah algoritma *supervised learning* yang membangun model prediksi dalam bentuk struktur pohon. Setiap cabang pohon merupakan keputusan berdasarkan atribut yang paling informatif, dan setiap daun merepresentasikan kelas keluaran. Decision Tree dipilih karena interpretabilitasnya tinggi pada akhirnya kamu bisa melihat aturan keputusan yang dibentuk oleh model secara jelas[26]. Proses pemilihan atribut terbaik pada *Decision Tree* biasanya menggunakan dua ukuran yakni Entropy dan Information Gain.

$$\text{Entropy}(S) = -\sum_{i=1}^n p_i \cdot \log_2(p_i) \tag{6}$$

Keterangan: S adalah himpunan data, p_i adalah proporsi data pada kelas ke- i , dan *Entropy*(S) digunakan untuk mengukur tingkat ketidakteraturan atau ketidakmurnian suatu dataset.

$$\text{InformationGain}(S, A) = \text{Entropy}(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|} \cdot \text{Entropy}(S_v) \tag{7}$$

Keterangan: S adalah himpunan data, A adalah atribut yang digunakan untuk pemisahan data, S_v adalah subset data berdasarkan nilai atribut A , dan *Entropy*(S) mengukur tingkat ketidakteraturan dataset. Atribut dengan nilai

Information Gain tertinggi akan dipilih sebagai node keputusan berikutnya. Dengan pendekatan ini, model dapat menangkap pola non-linear dalam data komentar TikTok tanpa memerlukan pra-normalisasi fitur yang kompleks.

2.7. Klasifikasi Random Forest

Random Forest adalah metode *ensemble learning* yang menggabungkan banyak *Decision Tree* untuk meningkatkan akurasi dan mengurangi risiko *overfitting*. Setiap pohon keputusan dilatih dengan subset data yang diambil secara *bootstrap* dan *subset* fitur yang dipilih secara acak pada setiap split. Dengan model *ensemble* seperti ini, variasi antar pohon meningkat sehingga kekuatan prediksi lebih stabil dibanding satu pohon saja[27]. Mekanisme *majority voting* digunakan untuk menentukan hasil akhir, di mana kelas dengan jumlah prediksi terbanyak dari seluruh pohon akan dipilih sebagai hasil klasifikasi. Selain itu, *Random Forest* memiliki keunggulan dalam menangani data yang mengandung *noise* dan *outlier*, serta mampu mengolah fitur berdimensi tinggi seperti hasil pembobotan *TF-IDF* dari 8.082 komentar TikTok secara lebih optimal

3. HASIL DAN PEMBAHASAN.

3.1. Data Collection

Data dikumpulkan melalui metode *crawling* komentar video TikTok yang membahas produk *skincare* lokal. Dataset yang digunakan berupa file *CSV* yang berisi komentar pengguna TikTok terkait *review* dan opini produk *skincare* dalam negeri. Data tersebut kemudian diunggah ke Google Drive untuk diakses dan diproses lebih lanjut menggunakan *Google Colab*. Proses ini berhasil memperoleh dataset komentar *skincare* yang selanjutnya disimpan dalam format *CSV* untuk tahap *preprocessing*. Dapat dilihat pada Tabel 3.

Tabel 3. Hasil Pengumpulan Data

No	Komentar Tiktok
1	whitening cream viva emang sebgas itu woyyy, omg ss cocok dimuka ak dan lumayan bagus whitestory serum niacin 10 persen jujur bagus poll, avoskin ga ada yang gagal skinkernya menurut gue, skinsena muka gue yg tadinya banyak bgt bruntusan jadi mulus
8.082	ss wardah varian serumnya bagus bagus pake banget, viva kok g masuk padahal murah dan bagus bgt, brightening cream wardah emg sebgas itu, wardah hydra rose fw emg sebgas itu ngelembabin, aku pke fw wardah hydrarose cuy awet mana murah lg.

Berdasarkan Tabel 3, hasil pengumpulan data menunjukkan jumlah komentar yang berhasil dikumpulkan dari platform TikTok terkait produk *skincare* lokal. Langkah selanjutnya adalah melakukan *preprocessing* agar data lebih siap dan dinamis untuk dianalisis.

3.2. Preprocessing

Pada periode Februari 2023 hingga Februari 2026, terkumpul ribuan komentar dari aplikasi TikTok yang kemudian diproses melalui tahapan *preprocessing*. Proses ini bertujuan untuk membersihkan data teks, menyamakan format penulisan, serta menyiapkan komentar agar lebih terstruktur sehingga dapat digunakan dalam analisis sentimen. Tahapan *preprocessing* yang dilakukan meliputi beberapa langkah, yaitu data cleaning, *Case Folding*, normalisasi, tokenisasi, *Stopword Removal*, dan *Stemming*. Tahap *data cleaning* dilakukan dengan menghapus karakter khusus, emotikon, simbol, tanda baca, serta *URL* yang tidak relevan. Selanjutnya, *case folding* digunakan untuk mengubah seluruh huruf menjadi huruf kecil agar penulisan menjadi konsisten. Pada tahap *normalisasi*, kata tidak baku atau bahasa slang diubah menjadi kata yang sesuai dengan bahasa Indonesia standar. Setelah itu dilakukan *tokenisasi* untuk memecah teks menjadi unit kata atau token. Proses berikutnya adalah *Stopword Removal*, yaitu menghapus kata-kata umum yang tidak memiliki makna penting dalam analisis, seperti “yang”, “dan”, dan “tidak”. Tahap terakhir adalah *Stemming*, yaitu mengubah kata menjadi bentuk dasar atau akar katanya.

Ringkasan hasil proses *preprocessing* ditampilkan pada Tabel 4, yang disusun secara vertikal agar setiap tahapan transformasi teks dapat terlihat dengan jelas dan memudahkan pembaca dalam memahami perubahan data dari bentuk mentah hingga siap digunakan dalam proses analisis. Adapun ringkasan hasil setiap tahapan *preprocessing* ditampilkan pada Tabel 4, yang menyajikan perubahan data teks secara bertahap dari bentuk mentah hingga siap digunakan dalam proses analisis sentimen.

Tabel 4. Hasil Preprocessing

Tanggal	Waktu	Tweet Asli	Cleaning	Case Folding	Normalisasi	Tokenize	Stopword Removal	Stemming
2025-12-04	13:23:45	apa cmn aku yg cari G2G?	apa cmn aku yg cari GG	apa cmn aku yg cari gg	apa cuman aku yang cari tidak	[apa, cuman, aku, yang, cari, tidak]	[cuman, cari]	cuman cari



Tanggal	Waktu	Tweet Asli	Cleaning	Case Folding	Normalisasi	Tokenize	Stopword Removal	Stemming
2025-12-05	14:03:26	IYA TAU GUA BASIC	IYA TAU GUA BASIC	iya tau gua basic	iya tau gua basic	[iya, tau, gua, basic]	[iya, tau, gua, basic]	iya tau gua basic
2026-02-03	12:40:32	yeyy aku standard yang azarine aku punya sunscreen anak	yeyy aku standard yang azarine aku punya sunsc...	yeyy aku standard yang azarine aku punya sunsc...	yeyy aku standard yang azarine aku punya sunsc...	[yeyy, aku, standard, yang, azarine, aku, puny...]	[yeyy, standard, azarine, sunscreen, anak, anak]	yeyy standard azarine sunscreen anak anak
2025-12-01	11:47:56	Tim finally found you 🌟	Tim finally found you	tim finally found you	tim finally found you	[tim, finally, found, you]	[tim, finally, found, you]	tim finally found you
2025-12-02	08:40:22	aku pigeon	aku pigeon	aku pigeon	aku pigeon	[aku, pigeon]	[pigeon]	pigeon

Berdasarkan Tabel 4, setiap komentar telah melalui semua tahapan preprocessing sehingga teks menjadi lebih bersih dari simbol dan tanda baca yang tidak relevan. Seluruh huruf diseragamkan ke format kecil melalui Case Folding, sementara kata-kata yang tidak baku atau slang diubah menjadi bentuk standar melalui proses normalisasi. Tokenisasi menghasilkan daftar kata yang siap dianalisis, dan Stopword Removal memastikan hanya kata-kata yang memiliki informasi penting yang tersisa. Selain itu, Stemming mengembalikan setiap kata ke bentuk dasarnya sehingga fitur TF IDF dapat diekstraksi secara optimal. Dengan seluruh tahapan ini, teks komentar TikTok menjadi lebih konsisten, terstruktur, dan siap digunakan dalam analisis sentimen maupun klasifikasi opini menggunakan algoritma machine learning.

3.3. Labeling

Setelah tahapan preprocessing selesai, langkah selanjutnya adalah memberikan label pada setiap komentar untuk keperluan analisis sentimen. Pelabelan ini bertujuan untuk mengkategorikan opini pengguna TikTok terhadap produk skincare ke dalam kelas tertentu agar dapat digunakan sebagai data latih bagi model klasifikasi. Kategori sentimen yang digunakan dalam penelitian ini terdiri dari lima kelas, yaitu Aman, Tidak Aman, Efektif, Tidak Efektif, dan Netral. Pelabelan sentimen dilakukan menggunakan pendekatan lexicon-based dengan memanfaatkan kamus sentimen Bahasa Indonesia (InSet) yang disesuaikan dengan konteks skincare. Dalam proses ini, kata-kata yang berkaitan dengan aspek keamanan seperti “aman” dan “tidak iritasi”, serta kata yang menunjukkan efektivitas seperti “cocok” dan “ampuh”, digunakan sebagai indikator dalam penentuan label. Penentuan kategori dilakukan berdasarkan dominasi kata dalam setiap komentar hasil stemming, sehingga setiap komentar dianalisis secara otomatis untuk diberikan label yang paling sesuai. Misalnya, komentar yang menyatakan bahwa produk “sangat bagus” diberi label Efektif karena kata “bagus” termasuk kata yang menandakan aspek efektivitas, sedangkan komentar yang menyebutkan produk “aman dipakai setiap hari” diberi label Aman karena kata “aman” menunjukkan aspek keamanan. Dengan pendekatan ini, setiap komentar dikategorikan secara konsisten dan otomatis, sehingga validitas metode pelabelan dapat lebih mudah dipahami.

Data yang digunakan dalam pelatihan model merupakan data yang telah melalui proses preprocessing dan pelabelan, kemudian dibagi menjadi data training dan testing untuk keperluan proses klasifikasi. Tabel 5 menampilkan beberapa contoh komentar dan label yang telah diberikan, sebagai ilustrasi metode pelabelan yang diterapkan pada seluruh dataset.

Tabel 5. Hasil Labeling

Komentar (Stemming)	Label
cocok di kulit dan bikin glowing	Efektif
bikin merah dan iritasi	Tidak Aman
tidak cocok di kulit sensitif	Tidak Efektif
aman dipakai setiap hari	Aman
mahal tapi hasil biasa saja	Tidak Efektif
aku pakai produk ini	Netral
skincare pigeon kumpul	Netral
iya tau gua basic	Netral

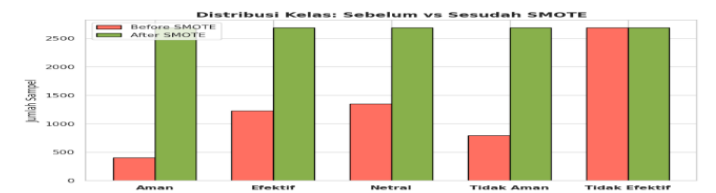
Dengan pelabelan ini, setiap komentar kini memiliki kategori sentimen yang jelas sehingga mempermudah proses pelatihan model klasifikasi untuk mengenali pola opini pengguna TikTok. Data yang telah dilabeli ini

kemudian digunakan untuk membangun model dengan algoritma seperti *Naïve Bayes*, *Decision Tree*, dan *Random Forest*.

3.4. Handle Imbalance (SMOTE)

Analisis awal terhadap data menunjukkan adanya ketidakseimbangan distribusi kelas pada dataset opini pengguna. Ketidakseimbangan ini terlihat dari perbedaan jumlah data pada setiap kategori label, yaitu Aman, Efektif, Netral, Tidak Aman, dan Tidak Efektif. Ketidakseimbangan kelas dapat menyebabkan model klasifikasi cenderung lebih fokus pada kelas dengan jumlah data yang lebih banyak sehingga berpotensi menurunkan kemampuan model dalam mengenali kelas dengan jumlah data yang lebih sedikit. Untuk mengatasi permasalahan tersebut, penelitian ini menerapkan Synthetic Minority Oversampling Technique (SMOTE). Metode SMOTE bekerja dengan cara membuat sampel sintetis baru pada kelas minoritas berdasarkan kedekatan antar data pada ruang fitur. Dengan demikian, jumlah data pada setiap kelas menjadi lebih seimbang tanpa harus menghapus data yang sudah ada.

Penerapan SMOTE dalam penelitian ini dilakukan setelah proses pembagian data menjadi data latih (training) dan data uji (testing). Teknik SMOTE hanya diterapkan pada data latih, sedangkan data uji tetap asli tanpa sampel sintetis, sehingga validitas evaluasi performa model tetap terjaga dan menghindari terjadinya data leakage. Penerapan metode ini diharapkan mampu meningkatkan kemampuan model dalam mengenali pola pada kelas minoritas secara lebih optimal. Selain itu, keseimbangan distribusi data yang lebih baik juga berkontribusi terhadap peningkatan stabilitas dan generalisasi model dalam proses klasifikasi. Dengan demikian, hasil klasifikasi yang diperoleh menjadi lebih representatif terhadap kondisi data yang sebenarnya. Distribusi kelas sebelum dan sesudah penerapan SMOTE dapat dilihat pada Gambar 3.



Gambar 3. Distribusi kelas sebelum dan sesudah penerapan SMOTE

Berdasarkan Gambar 3, terlihat perbedaan distribusi jumlah data pada setiap kelas sebelum dan sesudah penerapan SMOTE. Sebelum dilakukan SMOTE, distribusi data masih tidak seimbang, di mana kelas Tidak Efektif memiliki jumlah data paling dominan dibandingkan kelas lainnya, sedangkan kelas seperti Aman dan Tidak Aman memiliki jumlah data yang relatif lebih sedikit. Setelah penerapan SMOTE, jumlah data pada setiap kelas menjadi lebih seimbang, dengan jumlah yang relatif sama pada masing-masing kategori. Perlu diperhatikan bahwa distribusi setelah SMOTE yang ditampilkan merupakan hasil pada data latih, bukan keseluruhan dataset. Hal ini menunjukkan bahwa metode SMOTE berhasil mengatasi permasalahan ketidakseimbangan data dengan cara menambahkan sampel sintetis hanya pada kelas minoritas di data latih. Dengan distribusi data yang lebih seimbang, model klasifikasi dapat mempelajari pola dari setiap kelas secara lebih optimal, sehingga diharapkan dapat meningkatkan performa model dalam mengklasifikasikan data serta mengurangi bias terhadap kelas mayoritas.

3.5. Hasil Feature Extraction (TF-IDF)

TF-IDF (Term Frequency – Inverse Document Frequency) digunakan untuk mengukur tingkat kepentingan suatu kata dalam dokumen dengan membandingkan frekuensi kemunculannya pada dokumen tertentu dan keseluruhan dokumen. Metode ini banyak digunakan dalam analisis teks, pencarian informasi, serta pemrosesan bahasa alami (*Natural Language Processing / NLP*). Dengan menggunakan metode *TF-IDF*, kata yang sering muncul dalam suatu dokumen tetapi jarang muncul pada dokumen lainnya akan memiliki bobot yang lebih tinggi sehingga dianggap lebih penting dalam merepresentasikan isi dokumen. Hasil perhitungan *TF-IDF* pada data penelitian dirangkum pada Tabel 6.

Tabel 6. Hasil TF-IDF

Term	TF D1	TF D2	TF D3	IDF	TFIDF D1	TFIDF D2	TFIDF D3
anak	0	0	1	1,693	0	0	0,707
azarine	0	0	1	1,693	0	0	0,353
basic	0	1	0	1,693	0	0,5	0
cari	1	0	0	1,693	0,707	0	0
cuman	1	0	0	1,693	0,707	0	0
gua	0	1	0	1,693	0	0,5	0
iya	0	1	0	1,693	0	0,5	0
standard	0	0	1	1,693	0	0	0,353
sunscreen	0	0	1	1,693	0	0	0,353
tau	0	1	0	1,693	0	0,5	0
yeyy	0	0	1	1,693	0	0	0,353

Berdasarkan Tabel 6, ditampilkan hasil perhitungan TF-IDF dari tiga dokumen (D1, D2, D3). Nilai TF menunjukkan jumlah kemunculan kata pada setiap dokumen, sedangkan IDF menggambarkan tingkat kepentingan kata berdasarkan distribusinya di seluruh dokumen. Bobot TF-IDF, yang merupakan hasil perkalian antara TF dan IDF, menunjukkan tingkat relevansi suatu kata terhadap dokumen tertentu. Perlu dicatat bahwa contoh tiga dokumen ini hanya bersifat ilustratif dan tidak mewakili seluruh 8.082 komentar. Beberapa kata seperti “anak”, “cari”, dan “cuman” memiliki bobot TF-IDF tinggi karena muncul spesifik pada dokumen tertentu, namun kata-kata ini bersifat umum dan belum tentu merepresentasikan aspek keamanan atau efektivitas produk. Hal ini menunjukkan bahwa proses Stopword Removal perlu diperhatikan agar kata-kata umum tidak mendominasi bobot TF-IDF, dan seleksi fitur lebih kontekstual diperlukan untuk meningkatkan kualitas representasi teks dalam analisis sentimen. Hasil pembobotan TF-IDF ini kemudian digunakan sebagai fitur masukan dalam proses klasifikasi menggunakan algoritma machine learning untuk menganalisis opini pengguna terhadap keamanan dan efektivitas produk skincare lokal pada platform TikTok. Selain itu, penamaan algoritma yang sebelumnya tertulis “Model Skincare” telah diperbaiki menjadi “Naïve Bayes” agar lebih profesional dan jelas, menghindari kebingungan dengan algoritma lain seperti Decision Tree dan Random Forest. Hal ini penting agar interpretasi performa model menjadi valid dan mudah dipahami. Perlu ditekankan juga bahwa performa model dievaluasi hanya pada data uji yang asli, sehingga peningkatan akurasi setelah SMOTE mencerminkan performa model yang sah tanpa terpengaruh data leakage.

3.6. Hasil Klasifikasi Model

Penelitian ini menguji beberapa algoritma klasifikasi untuk menilai kemampuan model dalam mengklasifikasikan opini pengguna. Proses pengujian dilakukan dengan membandingkan hasil klasifikasi sebelum dan sesudah penerapan teknik SMOTE (Synthetic Minority Oversampling Technique). Teknik SMOTE digunakan untuk mengatasi ketidakseimbangan data pada dataset, di mana jumlah data pada setiap kelas tidak merata. Ketidakseimbangan ini dapat memengaruhi performa model karena model cenderung lebih mudah mengenali kelas yang memiliki jumlah data lebih banyak. Dalam penelitian ini digunakan tiga algoritma klasifikasi, yaitu Naïve Bayes, Decision Tree, dan Random Forest. Ketiga algoritma tersebut dipilih karena umum digunakan dalam penelitian text mining dan memiliki karakteristik berbeda dalam proses klasifikasi data. Perbandingan hasil ini bertujuan untuk mengetahui sejauh mana pengaruh SMOTE dalam meningkatkan keseimbangan data dan performa model. Selain itu, analisis dilakukan untuk melihat algoritma mana yang paling optimal dalam menangani data yang tidak seimbang. Hasil evaluasi juga memberikan gambaran mengenai efektivitas masing-masing metode dalam meningkatkan kualitas klasifikasi. Dengan demikian, dapat diperoleh pemahaman yang lebih mendalam terkait kontribusi setiap metode dalam proses klasifikasi data. Analisis ini juga membantu dalam mengidentifikasi kelebihan dan kekurangan dari masing-masing algoritma yang digunakan. Perbandingan kinerja antar model menjadi dasar dalam menentukan metode yang paling sesuai untuk data penelitian. Hal ini penting untuk memastikan bahwa model yang dipilih mampu memberikan hasil yang optimal dan konsisten. Evaluasi kinerja model dilakukan menggunakan beberapa metrik, yaitu accuracy, precision, recall, dan F1-score, agar dapat memberikan gambaran yang lebih lengkap mengenai kemampuan model dalam mengklasifikasikan setiap kelas. Hasil klasifikasi sebelum penerapan SMOTE disajikan pada Tabel 7, sedangkan hasil klasifikasi setelah penerapan SMOTE ditampilkan pada Tabel 8.

Tabel 7. Hasil Klasifikasi Before *SMOTE*

Model	Accuracy	Class	Precision	Recall	F1-Score
Naïve Bayes	65%	Aman	100%	10%	18%
		Efektif	93%	68%	78%
		Netral	99%	49%	65%
		Tidak Aman	75%	2%	3%
Decision Tree	79%	Tidak Efektif	55%	100%	71%
		Aman	72%	54%	62%
		Efektif	70%	60%	64%
		Netral	95%	100%	97%
Random Forest	87%	Tidak Aman	58%	84%	68%
		Tidak Efektif	83%	78%	81%
		Aman	84%	53%	65%
		Efektif	88%	72%	79%
Random Forest	87%	Netral	95%	100%	98%
		Tidak Aman	76%	93%	84%
		Tidak Efektif	86%	90%	88%

Berdasarkan hasil pada Tabel 7 dapat dilihat bahwa performa setiap model sebelum penerapan *SMOTE* menunjukkan hasil yang berbeda. Naïve Bayes memperoleh nilai akurasi sebesar 65%, yang merupakan nilai akurasi paling rendah dibandingkan dua model lainnya. Beberapa kelas juga menunjukkan nilai *recall* yang cukup rendah, misalnya pada kelas Aman yang hanya memiliki *recall* sebesar 10%, serta kelas Tidak Aman dengan *recall* sebesar 2%. Hal ini menunjukkan bahwa model *Skincare* masih mengalami kesulitan dalam mengenali beberapa kelas pada *dataset*.

Model *Decision Tree* memberikan hasil yang lebih baik dengan nilai akurasi sebesar 79%. Model ini mampu mengidentifikasi kelas Netral dengan sangat baik, terlihat dari nilai *precision* sebesar 95%, recall 100%, serta *f1-score* 97%. Hasil tersebut menunjukkan bahwa Naïve Bayes cukup mampu mempelajari pola data pada kelas tersebut. Sementara itu, model *Random Forest* menunjukkan performa terbaik sebelum penerapan *SMOTE* dengan nilai akurasi sebesar 87%. Model ini memiliki performa yang cukup stabil pada hampir seluruh kelas. Misalnya pada kelas Netral diperoleh *f1-score* sebesar 98%, sedangkan pada kelas Tidak Efektif diperoleh *f1-score* sebesar 88%. Hal ini menunjukkan bahwa *Random Forest* memiliki kemampuan yang lebih baik dalam memodelkan pola data dibandingkan dua algoritma lainnya. Hasil klasifikasi *After SMOTE* penerapan *SMOTE* ditampilkan pada Tabel 8.

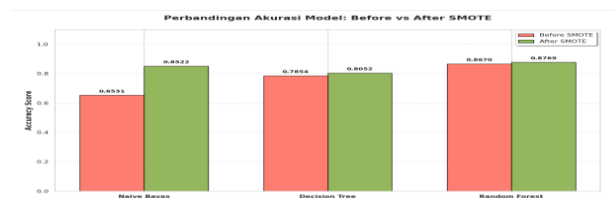
Tabel 8. Hasil Klasifikasi After *SMOTE*

Model	Accuracy	Class	Precision	Recall	F1-Score
Naïve Bayes	85%	Aman	58%	92%	71%
		Efektif	82%	90%	86%
		Netral	97%	80%	88%
		Tidak Aman	78%	87%	82%
		Tidak Efektif	91%	84%	87%
Decision Tree	81%	Aman	62%	60%	61%
		Efektif	74%	64%	69%
		Netral	94%	100%	97%
		Tidak Aman	61%	92%	73%
		Tidak Efektif	88%	78%	83%
Random Forest	88%	Aman	85%	69%	77%
		Efektif	85%	79%	82%
		Netral	95%	100%	98%
		Tidak Aman	71%	98%	82%
		Tidak Efektif	92%	85%	89%

Berdasarkan hasil pada Tabel 8 terlihat bahwa penerapan teknik *SMOTE* memberikan pengaruh terhadap performa model klasifikasi. Naïve Bayes mengalami peningkatan yang cukup signifikan, di mana nilai akurasi meningkat dari 65% menjadi 85%. Selain itu beberapa kelas juga menunjukkan peningkatan nilai recall, misalnya pada kelas Aman yang mencapai 92% serta kelas Efektif dengan *f1-score* sebesar 86%.

Pada model *Decision Tree*, nilai akurasi meningkat dari 79% menjadi 81% setelah penerapan *SMOTE*. Model ini tetap menunjukkan performa yang sangat baik pada kelas Netral, dengan nilai *f1-score* sebesar 97%. Meskipun peningkatan akurasi tidak terlalu besar, penerapan *SMOTE* tetap membantu model dalam mengenali data pada beberapa kelas yang sebelumnya kurang seimbang.

Sementara itu, model *Random Forest* tetap menunjukkan performa terbaik dibandingkan model lainnya. Setelah penerapan *SMOTE*, nilai akurasi meningkat dari 87% menjadi 88%. Model ini juga menunjukkan performa yang stabil pada hampir seluruh kelas, terutama pada kelas Netral dengan *f1-score* sebesar 98% dan kelas Tidak Efektif dengan *f1-score* sebesar 89%. Hal ini menunjukkan bahwa *Random Forest* memiliki kemampuan yang cukup baik dalam menangani dataset yang digunakan dalam penelitian ini. Untuk memberikan gambaran yang lebih jelas mengenai perbandingan akurasi model sebelum dan sesudah penerapan *SMOTE*, hasil tersebut ditampilkan pada Gambar 5.



Gambar 5. Perbandingan Akurasi Model Before dan After *SMOTE*

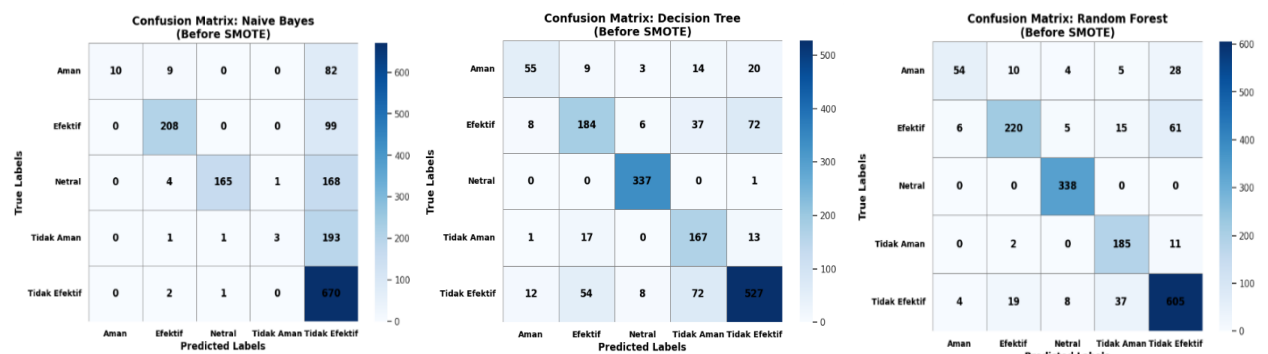
Berdasarkan Gambar 5 dapat dilihat perbandingan nilai akurasi dari ketiga model klasifikasi sebelum dan sesudah penerapan teknik *SMOTE*. Grafik tersebut menunjukkan bahwa penerapan *SMOTE* memberikan peningkatan performa pada seluruh model yang digunakan dalam penelitian ini.

Naïve Bayes mengalami peningkatan akurasi yang cukup signifikan, yaitu dari 0.6531 sebelum *SMOTE* menjadi 0.8522 setelah *SMOTE*. Hasil ini menunjukkan bahwa teknik *SMOTE* mampu membantu model dalam mengatasi ketidakseimbangan data sehingga model dapat mempelajari pola data dengan lebih baik. Pada model *Decision Tree*, nilai akurasi meningkat dari 0.7854 menjadi 0.8052 setelah penerapan *SMOTE*. Meskipun peningkatan yang terjadi tidak terlalu besar, hasil tersebut tetap menunjukkan bahwa teknik *oversampling* memberikan pengaruh positif terhadap performa model dalam proses klasifikasi. Sementara itu, model *Random Forest* menunjukkan performa terbaik dibandingkan dua model lainnya. Nilai akurasi meningkat dari 0.8670 sebelum *SMOTE* menjadi 0.8769 setelah *SMOTE*. Hal ini menunjukkan bahwa *Random Forest* memiliki kemampuan klasifikasi yang lebih

stabil dalam memproses dataset yang digunakan dalam penelitian ini. Secara keseluruhan, penerapan teknik *SMOTE* terbukti mampu meningkatkan performa model klasifikasi, khususnya pada Naïve Bayes yang mengalami peningkatan akurasi paling besar. Selain itu, *Random Forest* tetap menjadi model dengan performa terbaik baik sebelum maupun setelah penerapan *SMOTE*.

3.7. Evaluasi Model

Setelah proses klasifikasi menggunakan algoritma Naïve Bayes, Decision Tree, dan Random Forest, tahap selanjutnya adalah melakukan evaluasi performa model. Evaluasi dilakukan menggunakan confusion matrix untuk melihat sejauh mana model mampu mengklasifikasikan data ke kategori yang benar. Dari confusion matrix, dapat dihitung nilai True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) untuk setiap kelas. Nilai-nilai ini selanjutnya digunakan untuk menghitung metrik evaluasi seperti precision, recall, dan F1-score. Selain itu, confusion matrix memberikan gambaran visual distribusi prediksi model terhadap masing-masing kelas sentimen: Aman, Efektif, Netral, Tidak Aman, dan Tidak Efektif. Hasil evaluasi model sebelum dan sesudah penerapan teknik *SMOTE* ditunjukkan pada Gambar 6 dan Gambar 7

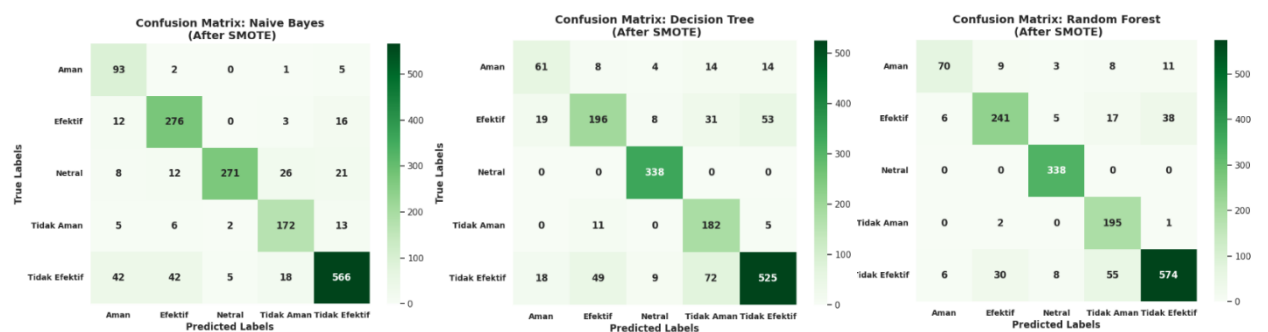


Gambar 6. Confusion Matrix Sebelum *SMOTE*

Berdasarkan Gambar 6, sebelum penyeimbangan data menggunakan *SMOTE*, Decision Tree menghasilkan total True Positive (TP) sebesar 1.443 (55+184+337+167+527). Masih terdapat sejumlah kesalahan klasifikasi, terutama pada kelas minoritas. Misalnya, kelas Aman sering diprediksi sebagai Tidak Efektif sebanyak 20 kasus, sedangkan kelas Efektif salah diklasifikasikan menjadi Tidak Efektif sebanyak 72 kasus. Pada kelas Tidak Aman, terdapat prediksi salah ke Efektif sebanyak 17 kasus dan ke Tidak Efektif sebanyak 13 kasus. Di sisi lain, kelas Netral menunjukkan akurasi cukup baik dengan 337 prediksi benar dan hanya 1 kesalahan.

Pada algoritma Naïve Bayes, model cenderung bias terhadap kelas mayoritas (Tidak Efektif) dengan 670 prediksi benar, sementara kelas minoritas sulit dikenali. Kelas Aman hanya benar 10 data, kelas Efektif 208 data, dan kelas Tidak Aman hanya 3 data yang benar, menunjukkan model hampir tidak mampu mengidentifikasi kelas minoritas.

Random Forest menunjukkan kinerja lebih stabil, dengan TP 1.602 (54+220+338+185+605). Model mampu mengenali kelas Netral dengan sempurna (338 prediksi benar) dan kelas Tidak Efektif 605 prediksi benar. Kesalahan klasifikasi masih terjadi pada kelas minoritas: 28 data Aman salah diprediksi Tidak Efektif, dan 61 data Efektif juga salah diprediksi Tidak Efektif. Secara umum, Random Forest lebih baik menangani distribusi data tidak seimbang dibanding Naïve Bayes dan Decision Tree.



Gambar 7. Confusion Matrix Setelah *SMOTE*

Setelah penerapan *SMOTE*, kemampuan model meningkat dalam mengenali kelas minoritas. Pada Decision Tree, TP mencapai 1.302 (61+196+338+182+525). Kelas Aman meningkat prediksi benar dari 55 menjadi 61, kelas Efektif dari 184 menjadi 196, dan kelas Tidak Aman dari 167 menjadi 182. Kesalahan prediksi ke kelas lain menurun, menunjukkan *SMOTE* membantu model mempelajari pola kelas minoritas.

Naïve Bayes juga meningkat signifikan: kelas Aman dari 10 menjadi 93 prediksi benar, Efektif dari 208 menjadi 276, Tidak Aman dari 3 menjadi 172, dengan total TP 1.378 (93+276+271+172+566). Random Forest tetap stabil, TP 1.418 (70+241+338+195+574), dengan peningkatan prediksi benar pada kelas minoritas dan kesalahan prediksi yang minimal.

Secara umum, SMOTE terbukti meningkatkan kemampuan model dalam mengidentifikasi kelas minoritas. Random Forest menunjukkan performa paling baik dengan keseimbangan optimal antara precision dan recall, Naïve Bayes meningkat signifikan, dan Decision Tree meningkat moderat namun tetap interpretatif.

3.8. Visualisasi WordCloud dan Frekuensi Kata

Visualisasi WordCloud dan grafik frekuensi kata digunakan untuk memahami kata-kata yang paling sering muncul pada komentar pengguna TikTok terkait produk skincare lokal Indonesia. Walaupun WordCloud dan grafik frekuensi memberikan gambaran umum kata dominan, perlu dicatat bahwa metode ini tidak secara langsung menjelaskan bagaimana kata-kata tersebut berkontribusi pada diferensiasi sentimen ‘Aman’ vs ‘Efektif’, yang menjadi fokus utama penelitian. Visualisasi WordCloud menampilkan kata-kata yang sering muncul, di mana ukuran kata menandakan frekuensi kemunculan: semakin besar, semakin sering kata muncul dalam dataset. Grafik frekuensi kata melengkapi informasi ini dengan menampilkan jumlah kemunculan setiap kata secara kuantitatif, sehingga mempermudah identifikasi kata yang paling dominan. Visualisasi dapat dilihat pada Gambar 8–12, sedangkan grafik frekuensi kata ditampilkan pada Gambar 13



Gambar 8. WordCloud Kategori Aman

Pada Gambar 8 terlihat WordCloud untuk kategori Aman. Kata-kata yang paling sering muncul adalah dokter, lihat, pigeon, viva, pakai, wardah, dan asli. Kemunculan kata dokter menunjukkan bahwa sebagian pengguna menilai keamanan produk berdasarkan rekomendasi tenaga medis, sementara kemunculan merek seperti pigeon, viva, dan wardah menandakan diskusi terkait keamanan produk tertentu. Kata-kata ini memberikan indikasi faktor-faktor yang memengaruhi persepsi keamanan produk.



Gambar 9. WordCloud Kategori Tidak Aman

Gambar 9 menunjukkan WordCloud kategori Tidak Aman. Kata dominan antara lain pigeon, china, merah, premium, dan lokal. Kata merah berkaitan dengan reaksi kulit seperti iritasi atau kemerahan, sedangkan kata china dan jepang menunjukkan perhatian pengguna terhadap asal bahan produk. Kata-kata ini membantu memahami persepsi risiko dan faktor yang memengaruhi penilaian keamanan.



Gambar 10. WordCloud Kategori Efektif

sentimen Aman dan Efektif. Untuk analisis lebih mendalam, diperlukan teknik NLP lanjutan untuk menilai kontribusi kata secara kuantitatif terhadap klasifikasi sentimen.

4. KESIMPULAN

Berdasarkan hasil penelitian, algoritma Naïve Bayes, Decision Tree, dan Random Forest menunjukkan perbedaan kinerja yang signifikan dalam mengklasifikasikan opini pengguna TikTok mengenai keamanan dan efektivitas produk skincare lokal. Sebelum penerapan SMOTE, Random Forest memperoleh akurasi tertinggi sebesar 87%, diikuti Decision Tree sebesar 79% dan Naïve Bayes sebesar 65%, dengan kelemahan utama terdapat pada kelas minoritas seperti Aman dan Tidak Aman yang memiliki nilai recall rendah. Setelah penerapan SMOTE, ketiga algoritma mengalami peningkatan performa, terutama dalam mengenali kelas minoritas, sehingga nilai akurasi, presisi, recall, dan F1-score menjadi lebih seimbang pada seluruh kelas sentimen. Hasil analisis TF-IDF menunjukkan bahwa fitur yang terbentuk masih didominasi oleh kata-kata umum dan nama merek, sehingga belum sepenuhnya mampu merepresentasikan aspek keamanan dan efektivitas secara spesifik. Hal ini mengindikasikan bahwa proses preprocessing dan seleksi fitur masih dapat ditingkatkan agar menghasilkan representasi yang lebih relevan. Visualisasi hasil klasifikasi menunjukkan bahwa mayoritas komentar berada pada kategori Efektif dan Tidak Efektif, sementara kategori Netral memiliki jumlah yang lebih sedikit. Penerapan SMOTE memberikan peningkatan performa model, khususnya dalam menangani ketidakseimbangan data, namun penerapannya perlu dilakukan secara tepat pada data latih (training data) untuk menghindari bias evaluasi. Secara keseluruhan, Random Forest menunjukkan performa terbaik dibandingkan algoritma lainnya. Penelitian selanjutnya disarankan untuk meningkatkan metode pelabelan sentimen, memperbaiki kualitas fitur, serta mengeksplorasi algoritma yang lebih kompleks seperti deep learning agar hasil klasifikasi menjadi lebih akurat dan lebih robust. Penelitian ini berkontribusi dalam pengembangan model analisis sentimen multi-kelas yang mampu membedakan aspek keamanan dan efektivitas produk skincare secara lebih spesifik, serta menunjukkan bahwa penerapan metode SMOTE efektif dalam meningkatkan performa klasifikasi pada dataset yang tidak seimbang. Hasil penelitian ini diharapkan dapat menjadi dasar dalam pengembangan sistem pemantauan opini publik berbasis media sosial yang lebih akurat dan informatif.

REFERENCES

- [1] 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.
- [2] A. N. Latif, M. Alianita, W. MUSAADAH, and R. Hidayat, "Mengatasi Tantangan dan Memanfaatkan Peluang Industri Halal Kosmetik dan Obat-Obatan Menuju Wajib Halal 2026: Analisis SWOT," *Indones. J. Halal*, vol. 7, no. 2, pp. 146–155, 2024, [Online]. Available: <https://doi.org/10.14710/halal.v7i2.23506>
- [3] Syifa Nurjanah and Yordan Hermawan Apidana, "Analisis Sentimen TikTok untuk Mengevaluasi Reputasi Merek Pasca Kasus *Overclaim*: Studi pada Daviena Skincare," *Technol. Informatics Insight J.*, vol. 4, no. 2, pp. 74–91, 2025, doi: 10.32639/h5x8te73.
- [4] A. Romadhony, S. Al Faraby, R. Rismala, U. N. Wisesti, and A. Arifianto, "Sentiment Analysis on a Large Indonesian Product Review Dataset," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 10, no. 1, pp. 167–178, 2024, doi: 10.20473/jisebi.10.1.167-178.
- [5] G. Khoerunnisa, S. S. Maesaroh, and M. R. Nugraha, "Analisis Sentimen Publik Terhadap Produk Rekomendasi Influencer Menggunakan Naive Bayes: Studi Kasus Tasya Farasya," *Data Sci. Indones.*, vol. 5, no. 1, pp. 84–95, 2025, doi: 10.47709/dsi.v5i1.6218.
- [6] S. Liem, T. Setiawan, and M. R. Pribadi, "Perbandingan Algoritma SVM dan Naïve Bayes Berbasis SMOTE dalam Analisis Sentimen Komentar Tiktok pada Produk Skincare," *Appl. Inf. Technol. Comput. Sci.*, vol. 3, no. 1, pp. 28–32, 2024, doi: 10.58466/aicoms.v3i1.1523.
- [7] J. Rama Dani, "Analisis Sentimen Komentar YouTube terhadap Kenaikan Tunjangan DPR RI menggunakan Naïve Bayes, SVM, dan Random Forest," *Technol. Sci.*, vol. 7, no. 3, pp. 1512–1524, 2025, doi: 10.47065/bits.v7i3.8513.
- [8] E. Permana, R. S. Eka Putri, P. D. Alfinda, and M. Mardhiyah, "Strategi Pemasaran Produk Skincare Somethinc Di Kalangan Generasi Z," *J. Pemasar. Kompetitif*, vol. 7, no. 2, pp. 119–135, 2024, doi: 10.32493/jpkpk.v7i2.29289.
- [9] D. Ruswanti, D. Susilo, and R. Riani, "Implementasi CRISP-DM pada Data Mining untuk Melakukan Prediksi Pendapatan dengan Algoritma C.45," *Go Infotech J. Ilm. STMIK AUB*, vol. 30, no. 1, pp. 111–121, 2024, doi: 10.36309/goi.v30i1.266.
- [10] Ridwan, E. H. Hermaliani, and M. Ernawati, "Penerapan Metode SMOTE Untuk Mengatasi Imbalanced Data Pada," *Comput. Sci.*, vol. 4, no. 1, pp. 80–88, 2024, [Online]. Available: <http://jurnal.bsi.ac.id/index.php/co-science>
- [11] C. Montag, H. Yang, and J. D. Elhai, "On the Psychology of TikTok Use: A First Glimpse From Empirical Findings," *Front. Public Heal.*, vol. 9, no. March, pp. 1–6, 2021, doi: 10.3389/fpubh.2021.641673.
- [12] D. Rifaldi and A. Fadlil, "DECODE: Jurnal Pendidikan Teknologi Informasi TEKNIK PREPROCESSING PADA TEXT MINING MENGGUNAKAN DATA TWEET 'MENTAL HEALTH,'" *Decod. J. Pendidik. Teknol. Inf.*, vol. 3, no. 2, pp. 161–171, 2023.
- [13] S. J. Angelina, A. Bijaksana, P. Negara, and H. Muhandi, "Analisis Pengaruh Penerapan Stopword Removal Pada Performa Klasifikasi Sentimen Tweet Bahasa Indonesia," *JUARA (Jurnal Apl. dan Ris. Inform.)*, vol. 02, no. 1, pp. 165–173, 2023, doi: 10.26418/juara.v2i1.69680.
- [14] P. Keterampilan and M. Berbasis, "RISET : Jurnal Ilmiah Multidisiplin Ilmu," vol. 1, no. 1, 2025.
- [15] A. Hermawan, I. Jowensen, J. Junaedi, and Edy, "Implementasi Text-Mining untuk Analisis Sentimen pada Twitter dengan Algoritma Support Vector Machine," *JST (Jurnal Sains dan Teknol.)*, vol. 12, no. 1, pp. 129–137, 2023, doi: 10.23887/jstundiksha.v12i1.52358.



- [16] A. T. J. H., “Preprocessing Text untuk Meminimalisir Kata yang Tidak Berarti dalam Proses Text Mining,” pp. 1–9.
- [17] L. Mu *et al.*, “DenseLoRA: Dense Low-Rank Adaptation of Large Language Models,” *Proc. Annu. Meet. Assoc. Comput. Linguist.*, vol. 1, pp. 10198–10211, 2025, doi: 10.18653/v1/2025.acl-long.503.
- [18] M. U. Albab, Y. K. P., and M. N. Fawaiq, “Optimization of the Stemming Technique on Text Preprocessing President 3 Periods Topic,” *J. Transform.*, vol. 20, no. 2, pp. 1–12, 2023, doi: 10.26623/transformatika.v20i2.5374.
- [19] S. Ghan, E. Diesen, C. Kunkel, K. Reuter, and H. Oberhofer, “Interpreting ultrafast electron transfer on surfaces with a converged first-principles Newns-Anderson chemisorption function,” *J. Chem. Phys.*, vol. 158, no. 23, pp. 1–13, 2023, doi: 10.1063/5.0151009.
- [20] Y. Elor and H. Averbuch-Elor, *To SMOTE, or not to SMOTE?*, vol. 1, no. 1. Association for Computing Machinery. [Online]. Available: <http://arxiv.org/abs/2201.08528>
- [21] Y. Lin and Y. Xie, “Semilinear heat equations and parabolic variational inequalities on graphs,” pp. 1–25, 2021, [Online]. Available: <http://arxiv.org/abs/2108.13007>
- [22] H. Bichri, A. Chergui, and M. Hain, “Investigating the Impact of Train / Test Split Ratio on the Performance of Pre-Trained Models with Custom Datasets,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 2, pp. 331–339, 2024, doi: 10.14569/IJACSA.2024.0150235.
- [23] R. Oktafiani, A. Hermawan, and D. Avianto, “Pengaruh Komposisi Split data Terhadap Performa Klasifikasi Penyakit Kanker Payudara Menggunakan Algoritma Machine Learning,” *J. Sains dan Inform.*, vol. 9, no. April, pp. 19–28, 2023, doi: 10.34128/jsi.v9i1.622.
- [24] H. Oktavianto, H. W. Sulistyono, G. Wijaya, and D. Irawan, “Analisis Komparasi Kinerja Metode Decision Tree dan Random Forest dalam Klasifikasi Teks Data Kesehatan,” vol. 11, no. 1, pp. 56–65, 2024.
- [25] A. Muzaki, V. Febriana, and W. N. Cholifah, “Analisis Sentimen Pada Ulasan Produk di E-Commerce dengan Metode Naive Bayes,” *J. Ris. dan Apl. Mhs. Inform.*, vol. 5, no. 4, pp. 758–765, 2024, doi: 10.30998/jrami.v5i4.9647.
- [26] S. A. Arnomo, A. A. Fajrin, Y. Siyamto, and S. F. N. Sadikin, “Evaluasi Model Decision Tree Pada Keputusan Kelayakan Kredit,” *J. Desain Dan Anal. Teknol.*, vol. 2, no. 2, pp. 200–206, 2023, doi: 10.58520/jddat.v2i2.39.
- [27] N. Fathirachman Mahing, A. Lazuardi Gunawan, A. Foresta Azhar Zen, F. Abdurrachman Bachtiar, and S. Agung Wicaksono, “Klasifikasi Tingkat Stress dari Data Berbentuk Teks dengan Menggunakan Algoritma Support Vector Machine (SVM) dan Random Forest,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 5, pp. 1067–1076, 2024, doi: 10.25126/jtiik.2024118010.