

# Analisis Sentimen Sunscreen Lokal Skintific, Somethinc, dan Avoskin dengan Naive Bayes dan SVM

## *Sentiment Analysis of Local Sunscreen Skintific, Somethinc, and Avoskin with Naive Bayes and SVM*

Windi Clarisha  
Prodi Teknik Komputer  
Universitas Negeri Makassar  
Makassar, Indonesia  
windiclarisha09@gmail.com

A. Astri Merilsa Fani  
Prodi Teknik Komputer  
Universitas Negeri Makassar  
Makassar, Indonesia  
astriandi39@gmail.com

Dewi Fatmarani Surianto  
Prodi Teknik Komputer  
Universitas Negeri Makassar  
Makassar, Indonesia  
dewifatmaranis@unm.ac.id

Nur Fadilah\*  
Prodi Pendidikan Teknologi  
Informasi  
Universitas Megarezky  
Makassar, Indonesia  
nurfadilahderman@unimerz.co.id\*

Diterima : Januari 2025  
Disetujui : Juli 2025  
Dipublikasi : Juli 2025

**Abstrak**— Industri kecantikan Indonesia, khususnya produk sunscreen lokal, menunjukkan pertumbuhan pesat seiring meningkatnya kesadaran masyarakat akan pentingnya perlindungan kulit dari sinar ultraviolet. Ulasan konsumen di platform digital menjadi sumber informasi penting untuk memahami persepsi dan preferensi pengguna. Penelitian ini bertujuan untuk menganalisis sentimen terhadap tiga merek sunscreen lokal—Skintific, Somethinc, dan Avoskin—dengan membandingkan dua metode klasifikasi teks, yaitu Naive Bayes dan Support Vector Machine (SVM). Untuk mengatasi ketidakseimbangan jumlah data antara sentimen positif dan negatif, digunakan teknik Synthetic Minority Over-sampling Technique (SMOTE). Hasil menunjukkan bahwa penerapan SMOTE pada Naive Bayes meningkatkan akurasi dari 81% menjadi 93%, serta memperbaiki precision, recall, dan F1-score secara signifikan. Sebaliknya, penerapan SMOTE pada SVM justru sedikit menurunkan akurasi dari 92% menjadi 91%, meskipun performa untuk kategori sentimen positif tetap stabil. Temuan ini menunjukkan bahwa kombinasi Naive Bayes dengan SMOTE lebih efektif dalam menangani data tidak seimbang untuk analisis sentimen produk kecantikan. Implikasi dari penelitian ini dapat digunakan oleh pelaku industri kecantikan sebagai dasar pengambilan keputusan dalam pengembangan dan pemasaran produk berbasis persepsi konsumen.

**Kata Kunci**—Analisis sentimen; Naive Bayes; Support Vector Machine (SVM); SMOTE; Sunscreen Lokal

**Abstract**— Indonesia's beauty industry, particularly local sunscreen products, has experienced rapid growth alongside increasing public awareness of the importance of skin protection against ultraviolet rays. Consumer reviews on digital platforms have become a vital source of information to understand user perceptions and preferences. This study aims to analyze sentiment toward three local sunscreen brands—Skintific, Somethinc, and Avoskin—by comparing two text classification methods: Naive Bayes and Support Vector Machine (SVM). To address the imbalance in the number of positive and negative sentiment data, the Synthetic Minority Over-sampling Technique (SMOTE) was

applied. The results show that applying SMOTE to Naive Bayes significantly improved the accuracy from 81% to 93%, along with notable enhancements in precision, recall, and F1-score. Conversely, applying SMOTE to SVM slightly reduced accuracy from 92% to 91%, although the performance for positive sentiment remained stable. These findings indicate that the combination of Naive Bayes and SMOTE is more effective in handling imbalanced data for sentiment analysis of beauty products. The implications of this study can serve as a basis for decision-making in product development and marketing strategies within the beauty industry, particularly in aligning with consumer sentiment.

**Keywords**—Sentiment analysis; Naive Bayes; Support Vector Machine (SVM); SMOTE; Local Sunscreen.

### I. PENDAHULUAN

Industri kecantikan, terutama di bidang kosmetik, mengalami pertumbuhan yang sangat cepat [1] Namun, konsumen kini menghadapi sejumlah masalah terkait produk perawatan kulit yang mereka pilih [2] Salah satu solusi yang populer dan semakin banyak digunakan oleh masyarakat adalah krim anti UV atau tabir surya, produk yang diformulasikan untuk memberikan perlindungan terhadap kulit dari paparan sinar ultraviolet (UV) [3] Dengan meningkatnya kesadaran masyarakat tentang pentingnya perlindungan kulit, terutama di negara tropis seperti Indonesia, permintaan untuk produk sunscreen terus bertambah, menciptakan variasi produk yang semakin beragam di pasar [4] Data dari Kompas Market Insight Dashboard menunjukkan bahwa penjualan sunscreen meningkat sebesar 99% pada kuartal pertama 2024 jika merujuk pada tahun sebelumnya dengan periode yang sama. Kenaikan tersebut menjadikan sunscreen sebagai salah satu produk dengan pertumbuhan tercepat di sektor kecantikan dan perawatan [5].

Era digital mendorong konsumen untuk mencari tinjauan dan rekomendasi produk sunscreen lokal lewat

platform e-commerce dan media sosial, agar sesuai dengan jenis kulit serta kebutuhan mereka [4] Akses informasi yang mudah melalui internet memungkinkan konsumen untuk menilai produk berdasar pengalaman pengguna lain [6] E-commerce menjadi pilihan utama karena kemudahan akses yang ditawarkannya [7] Perkembangan teknologi dan internet yang luas telah mengubah cara konsumen memilih produk yang tepat [8]

Seiring bertambahnya ulasan produk di *e-commerce* dan media sosial, analisis sentimen menjadi semakin signifikan. Ulasan atau kritik online dari pengguna dapat menjadi alat efektif untuk mendapatkan informasi mengenai produk atau layanan yang ada [9] Dengan menganalisis sentimen dalam ulasan tersebut, perusahaan dapat memperoleh informasi berharga tentang pandangan konsumen terhadap produk mereka [10] Selain itu, analisis sentimen juga membantu memberikan wawasan dan preferensi konsumen kepada publik, sehingga memperjelas mengenai kualitas suatu produk [11]

Penelitian mengenai analisis sentimen ulasan produk skincare telah maju dengan pesat. Misalnya, satu penelitian memakai Naïve Bayes Classifier untuk mengkategorikan sentimen ulasan produk Somethinc di Female Daily dengan akurasi 61%, namun menghadapi kesulitan dalam mengatasi kata-kata konteks [12] Penelitian lain menggunakan Support Vector Machine (SVM) untuk sentimen ulasan produk Skintific dengan akurasi 94%, tetapi memerlukan data dalam jumlah besar untuk mempertahankan akurasi yang tinggi [13]

Penelitian pengembangan Naïve Bayes Classifier untuk analisis sentimen merek skincare lokal seperti Avoskin, Azarine, dan Somethinc menunjukkan akurasi hingga 79% dengan K-fold Cross Validation, meskipun terbatas dalam menangani variasi pendapat di media sosial [14] Penelitian lain membandingkan kernel Polynomial dan RBF dalam algoritma SVM untuk analisis sentimen skincare di Indonesia menggunakan data Twitter, dengan hasil menjelaskan bahwa kernel RBF lebih unggul dengan akurasi 86,17% dibandingkan kernel Polynomial yang mencatat 84,56%, meskipun menghadapi masalah ketidakseimbangan data sentimen di media sosial [15]

Penelitian Naïve Bayes dengan TF-IDF dan Chi-Square pada ulasan kecantikan mencapai akurasi 80,18% dan f1-score 74,73%, meskipun terbatas untuk opini kompleks dan variasi Bahasa [16] Penelitian lain dengan Naïve Bayes untuk ulasan sunscreen di Female Daily menghasilkan akurasi 88% dan f1-score 89%, namun menghadapi tantangan pada opini kompleks dan dataset besar [17] Sementara itu, penerapan SVM pada ulasan e-commerce Amazon mencapai akurasi 80,06%, tetapi membutuhkan komputasi tinggi dan terkendala pada dataset besar dengan bahasa informal [18]

Penelitian sebelumnya berfokus pada analisis sentimen produk skincare secara terpisah dengan satu metode, seperti Naïve Bayes atau SVM. Beberapa studi hanya menganalisis ulasan produk di satu platform dan dengan metode berbeda, misalnya SVM pada produk Skintific dengan akurasi tinggi tetapi membutuhkan data besar untuk menjaga performa, serta Naïve Bayes pada produk Somethinc atau Avoskin yang efektif tetapi kurang mampu menangani kata-kata kontekstual dan variasi bahasa.

Belum ada penelitian yang secara komprehensif mengevaluasi kinerja secara komparatif Naïve Bayes dan SVM dalam mengklasifikasikan sentimen pada beberapa

merek sunscreen lokal dalam satu studi. Kesenjangan ini menjadi fokus penelitian ini, di mana Analisis Sentimen Komparatif terhadap Produk Sunscreen Lokal akan dilakukan dengan menguji kedua metode pada ulasan Skintific, Somethinc, dan Avoskin.

Pemilihan Skintific, Somethinc, dan Avoskin dalam penelitian ini difokuskan pada produk sunscreen mereka yang populer di Indonesia. Skintific menawarkan Produk Travel Size SKINTIFIC All Day Light Sunscreen Mist SPF50 PA++++ yang berhasil terjual sebanyak 13500 unit dalam dua minggu [19]. Somethinc meraih pangsa pasar 3,68% selama April-Juni 2022, dengan produk andalannya, Glowing Up Sunscreen Stick SPF 50+ PA++++, terjual sebanyak 1.000 unit [20]. Avoskin kerap mendapat respons positif berkat inovasi produknya dan konsistensinya dalam mengusung konsep green beauty. Produk terlarisnya adalah Avoskin The Great Shield SPF 50 PA++++, yang menjadi favorit di e-commerce [21]. Formulasi dan strategi ketiga merek ini relevan untuk dianalisis dalam konteks sentimen konsumen lokal.

Permasalahan yang sering dihadapi dalam analisis sentimen terhadap ulasan produk *sunscreen* lokal yaitu mengembangkan model yang mampu mengklasifikasikan opini dengan akurat di tengah informasi yang tidak lengkap, bias, dan variasi kata yang luas, terutama dalam bahasa informal dan campuran di media sosial. Model yang ada saat ini belum efektif menangani variasi bahasa dan ekspresi unik konsumen Indonesia, yang mempengaruhi akurasi klasifikasi sentiment [22]

Kendala dalam analisis sentimen tidak hanya terbatas pada permasalahan linguistik, tetapi juga mencakup ketidakseimbangan distribusi sentimen dalam dataset, di mana ulasan bernuansa positif cenderung lebih dominan dibandingkan ulasan negatif. Ketimpangan ini menjadi hambatan signifikan yang dapat mengurangi akurasi model klasifikasi. Untuk mengatasi permasalahan tersebut, penelitian ini menerapkan *Synthetic Minority Over-sampling Technique* (SMOTE), sebuah metode *resampling* yang terbukti efektif dalam meningkatkan kinerja model dengan menyeimbangkan jumlah data antar kelas melalui pembuatan sampel sintetis pada kelas minoritas. [23][24].

Tujuan dari penelitian ini untuk meningkatkan akurasi klasifikasi sentimen ulasan berbahasa Indonesia dengan menerapkan metode Naïve Bayes dan SVM yang dipadukan dengan teknik SMOTE, khususnya untuk menganalisis produk sunscreen lokal seperti Skintific, Somethinc, dan Avoskin. Diharapkan, pendekatan ini mampu menangani tantangan bahasa informal dan distribusi kelas yang tidak seimbang dalam data ulasan media sosial.

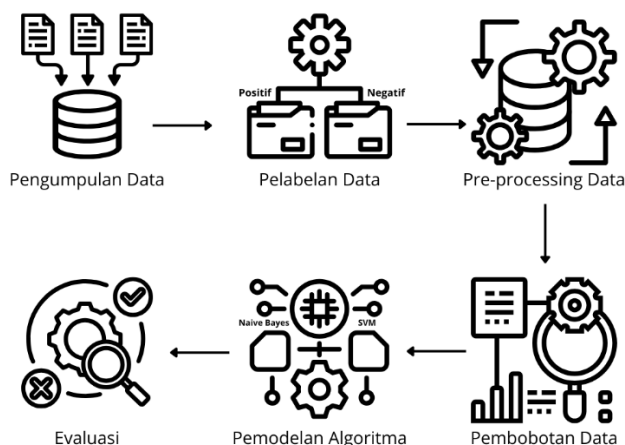
Penelitian ini memanfaatkan pendekatan klasifikasi menggunakan *Naïve Bayes* dan *Support Vector Machine* (SVM) guna menganalisis data berbasis teks. Naïve Bayes dipilih karena kemudahan dalam penerapan, kecepatan eksekusi, dan performa yang cukup baik untuk mengklasifikasikan data teks. Meski demikian, metode ini rentan terhadap pemilihan fitur dan berasumsi bahwa fitur bersifat independen, sehingga dapat mengurangi akurasi pada data yang kompleks [22]. Di sisi lain, Support Vector Machine (SVM) dirancang untuk bekerja secara efektif pada data berdimensi tinggi, sekaligus meningkatkan akurasi dalam proses klasifikasi antar kelas. Namun, SVM memerlukan waktu komputasi yang lebih lama, penggunaan

sumber daya yang lebih besar, serta sangat bergantung pada pemilihan kernel dan pengaturan parameter yang tepat [25]. Kedua metode ini digunakan untuk membandingkan hasil analisis sentimen sekaligus memahami preferensi konsumen terhadap produk *sunscreen* lokal.

Penelitian ini diharapkan dapat membantu membangun strategi pemasaran yang responsif dan terukur untuk memperkuat posisi produk *sunscreen* lokal, dengan memberikan wawasan tentang sentimen konsumen terhadap Skintific, Somethinc, dan Avoskin. Temuan ini juga diharapkan dapat membantu pelaku industri skincare lokal memahami preferensi pasar, meningkatkan kualitas produk, dan mendukung kebijakan yang memperkuat pertumbuhan industri kosmetik lokal, termasuk kesadaran akan pentingnya penggunaan *sunscreen*.

## II. METODE

Dalam sebuah studi, memiliki metode yang digambarkan dalam kerangka kerja penelitian adalah krusial. Langkah-langkah penelitian menjelaskan aktivitas yang akan dilakukan dari awal hingga selesai untuk mencapai tujuan yang diinginkan. Setiap langkah saling terhubung dan bersifat berkelanjutan. Proses dalam penelitian ini dijelaskan dengan lebih jelas melalui diagram alir yang terdapat pada Gambar 1.



Gambar 1. Tahapan Penelitian

### A. Pengumpulan Data

Penelitian ini menggunakan data kualitatif yang diambil dari teks ulasan yang menjadi objek kajian. Data tersebut dikumpulkan secara manual dari berbagai situs e-commerce. Data yang dikumpulkan berupa ulasan atau komentar terkait produk *sunscreen* dari Skintific, Somethinc, dan Avoskin dengan total 1500 data. Hasil pengumpulan data ini disimpan dalam format file .csv dan kemudian dicap untuk mengklasifikasikan pendapat atau pandangan dalam setiap ulasan.

### B. Pelabelan Data

Proses pelabelan data dilakukan untuk mengelompokkan pendapat atau pandangan dalam setiap ulasan. Pelabelan dilakukan secara manual dengan dua kategori, yaitu kelas positif dan negatif. Untuk menjamin konsistensi dan validitas hasil pelabelan, dilakukan evaluasi tingkat kesepakatan antar anotator melalui perhitungan nilai Cohen's Kappa. Dalam pelabelan sentimen, dua anotator independen memberikan label terhadap 1500 data teks. Hasil evaluasi menghasilkan nilai Cohen's Kappa sebesar 0,78,

yang mengindikasikan tingkat kesepakatan substansial. Data yang telah dilabeli dianggap valid untuk pelatihan model. Selanjutnya, dilakukan pra-pemrosesan terhadap dataset.

### C. Pre-processing Data

Pre-processing dilakukan untuk menyiapkan data agar lebih berkualitas dan siap untuk analisis. Proses ini penting karena data yang telah menjalani tahap pra-pemrosesan akan lebih akurat, konsisten, dan relevan, yang pada akhirnya mendukung pengambilan keputusan yang lebih tepat dan efektif [26]. Proses pra-pemrosesan mencakup beberapa langkah, seperti stemming, tokenisasi, penghapusan kata umum, dan casefolding.

### D. Pembobotan Data

Pasca tahapan pra-pemrosesan, data teks dikonversi menjadi representasi numerik melalui penerapan metode *Term Frequency-Inverse Document Frequency* (TF-IDF)[27] Metode TF-IDF memberikan pembobotan terhadap setiap kata berdasarkan tingkat frekuensinya dalam suatu dokumen. Adapun nilai *Term Frequency* (TF) dapat diperoleh melalui rumus sebagai berikut [28]:

$$TF = \frac{(\text{jumlah kemunculan kata dalam dokumen})}{(\text{jumlah kata dalam dokumen})} \quad (1)$$

Namun, nilai TF tidak menunjukkan pentingnya suatu kata dalam dokumen. Kata-kata umum yang sering muncul, seperti kata sambung, dapat memiliki nilai TF tinggi tetapi tidak signifikan. Oleh sebab itu, komponen *Inverse Document Frequency* (IDF) digunakan untuk memberikan bobot lebih besar pada kata-kata yang jarang muncul dalam korpus. Nilainya dihitung menggunakan rumus logaritmik berdasarkan rasio jumlah total dokumen terhadap jumlah dokumen yang memuat kata tersebut [28]:

$$IDF = \log \frac{N}{n} \quad (2)$$

Notasi  $N$  merepresentasikan total jumlah dokumen dalam korpus, sedangkan  $n$  menunjukkan jumlah dokumen yang mengandung kata tertentu. Nilai TF-IDF diperoleh dengan mengalikan nilai *Term Frequency* (TF) dan *Inverse Document Frequency* (IDF) yang telah dihitung sebelumnya. Kata dengan nilai TF-IDF tinggi dianggap penting dan lebih berkontribusi dalam menentukan topik dokumen.

### E. Pemodelan Algoritma Naïve Bayes Classifier dan Support Vector Machine (SVM)

Setelah proses pemberian nilai representatif terhadap kata dilakukan menggunakan metode seperti *Term Frequency-Inverse Document Frequency* (TF-IDF), data selanjutnya dianalisis menggunakan algoritma klasifikasi, yaitu *Naïve Bayes* dan *Support Vector Machine* (SVM). Dalam penelitian ini, digunakan pendekatan *Multinomial Naïve Bayes* sebagai salah satu metode klasifikasi, karena algoritma ini secara khusus dirancang untuk mengolah data teks yang telah dikonversi ke dalam bentuk frekuensi melalui teknik seperti TF-IDF maupun *Count Vectorizer*. *Multinomial Naïve Bayes* bekerja efektif pada fitur diskret dan dapat menangani klasifikasi dokumen pendek dengan kecepatan tinggi, meskipun memiliki kelemahan asumsi independensi antar fitur yang dapat memengaruhi akurasi pada data yang kompleks [12]. Proses klasifikasi ini untuk mengenali sentimen yang terdapat dalam ulasan yang telah diperoleh [29] *Naïve Bayes Classifier* merupakan algoritma klasifikasi berbasis probabilistik yang menerapkan Teorema

Bayes dengan asumsi independensi antar fitur. Metode ini banyak digunakan dalam *data mining* karena efisiensi komputasinya yang tinggi serta kemampuannya dalam menghasilkan performa klasifikasi yang kompetitif, khususnya pada data berdimensi tinggi dan domain teks. [30] Berikut adalah rumus untuk algoritma Naïve Bayes [31]:

$$P(X|H) = \frac{p(X|H)p(x)}{p(H)} \quad (3)$$

Rumus ini merupakan bentuk teorema Bayes yang digunakan untuk menghitung kemungkinan dari hipotesis  $H$  (seperti kelas) berdasarkan data atau fitur  $X$ . Dalam rumus ini,  $p(X|H)$  menunjukkan probabilitas bahwa data  $X$  terjadi jika hipotesis  $H$  benar, sedangkan  $p(X|H), p(x)$  menunjukkan probabilitas seluruh data  $X$ , dan  $p(H)$  adalah probabilitas awal dari hipotesis  $H$ .

Sementara itu, Support Vector Machine (SVM) adalah algoritma pembelajaran mesin yang paling sering digunakan untuk klasifikasi karena keefektifannya yang tinggi dalam mengklasifikasikan pola, dan tingkat keberhasilannya yang tinggi bila diterapkan di berbagai bidang [32] Dalam implementasinya, digunakan algoritma Support Vector Machine (SVM) dengan kernel linear dan parameter  $C=1.0$ , yang dipilih karena mampu bekerja secara optimal pada data berdimensi tinggi seperti data teks. Proses tuning parameter dilakukan menggunakan grid search sederhana pada nilai  $C$  (0.1, 1, 10) untuk mendapatkan performa terbaik [33]. Berikut adalah rumus untuk perhitungan SVM [34]:

$$(w \times xi) + b = 0 \quad (4)$$

Data  $xi$  masuk dalam kelas -1 yang dirumuskan dengan persamaan:

$$(w \times xi + b) \leq -1, yi = -1 \quad (5)$$

Sedangkan data  $xi$  yang termasuk dalam kelas +1 dapat dirumuskan dengan persamaan:

$$w \times xi + b \geq 1, yi = 1 \quad (6)$$

Sebelum pelatihan model dilakukan, diperhatikan bahwa terdapat ketidakseimbangan distribusi kelas dalam data, di mana kelas positif mendominasi sebanyak 1.196 sampel, sementara kelas negatif hanya 304 sampel. Ketidakseimbangan kelas dapat memicu bias terhadap kelas mayoritas. Untuk mengatasinya, digunakan metode SMOTE yang mensintesis data pada kelas minoritas guna menyeimbangkan distribusi dan diimplementasikan dengan pustaka imblearn. SMOTE bekerja dengan membuat sampel sintetis dari kelas minoritas (negatif) melalui interpolasi antar tetangga terdekat dalam ruang fitur. Teknik ini diterapkan dengan rasio 1:1, sehingga jumlah sampel pada kedua kelas menjadi seimbang, yakni 1.196 data positif dan 1.196 data negatif.

#### F. Evaluasi Model

Pada tahap ini, model akan dievaluasi untuk menilai seberapa akurat model dalam memprediksi sentimen. Proses evaluasi bertujuan untuk mengukur kinerja model dalam klasifikasi sentimen yang benar [35] Kemampuan model dalam memprediksi akan dinilai berdasarkan beberapa metrik, termasuk nilai dari Confusion Matrix, serta metrik lainnya seperti akurasi, presisi, recall, dan F-1 Score. Metrik-metrik ini memberikan pemahaman yang lebih menyeluruh mengenai kemampuan model dalam mengenali sentimen secara tepat dan seberapa baik model dalam menghindari

kesalahan klasifikasi. Berikut adalah rumus dari confusion matrix [36]:

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

Rumus yang disajikan digunakan untuk menghitung akurasi model klasifikasi, yakni rasio antara jumlah prediksi yang benar terhadap total prediksi. Nilai *True Positive* (TP) dan *True Negative* (TN) menunjukkan prediksi yang tepat, sedangkan *False Positive* (FP) dan *False Negative* (FN) merepresentasikan kesalahan klasifikasi

### III. HASIL DAN PEMBAHASAN

#### A. Pengumpulan Data

Pada penelitian ini, dataset yang dimanfaatkan adalah ulasan atau komentar produk sunscreen dari Skintific, Somethinc, dan Avoskin yang diambil dari berbagai platform e-commerce. Jumlah data yang digunakan mencapai 1500 ulasan atau komentar. Berikut adalah dataset yang dipakai yang terdapat dalam tabel 1

TABEL 1. CONTOH DATASET YANG DIGUNAKAN

No	Ulasan
1	Saya tidak terlalu menyukai sunscreen Skintific varian Aqua Light karena tidak cocok di wajah saya. Sunscreen ini justru membuat komedo putih bertambah dan menyebabkan lebih banyak jerawat kecil di wajah.
2	Teksturnya agak cair, tidak lengket di kulit, dan mudah meresap. Aman untuk kulit berjerawat. Kemasan produk ini juga bagus, sehingga cocok untuk dibawa saat bepergian.
3	Sejak dulu saya cocok dengan produk Skintific, terutama sunscreen varian Aqua Light Daily ini. Teksturnya sedikit kental, ringan, dan cepat meresap, sehingga sangat cocok digunakan saat terburu-buru berangkat kerja. Ukurannya yang kecil juga memudahkan untuk dibawa kemana saja. Setelah menggunakan sunscreen ini, wajah saya terlihat lebih cerah dan glowing. Saya sangat menyukai sunscreen Skintific yang satu ini. Pengirimannya juga cukup cepat dan aman dengan pembungkusan bubble wrap yang tebal. Terima kasih, Skintific.
...	...
1500	Saya agak kaget karena isinya tidak sampai penuh, tetapi produknya baik. Cepat meresap dan bebas white cast.

#### B. Pelabelan Data

Setelah data terkumpul, proses pelabelan data dilakukan secara manual oleh tim peneliti yang terdiri dari tiga orang. Hal ini dilakukan untuk memastikan bahwa pelabelan bersifat objektif. Pelabelan dilakukan dengan mengelompokkan ulasan menjadi sentimen positif apabila ulasan tersebut memberikan nilai tambah pada produk dan negatif apabila mengurangi nilai produk. Berikut merupakan hasil dari proses pelabelan manual pada dataset yang terdapat dalam tabel 2.

TABEL 2. HASIL PELABELAN DATA

No	Ulasan	Label
1	Saya tidak terlalu menyukai sunscreen Skintific varian Aqua Light karena tidak cocok di wajah saya. Sunscreen ini justru membuat komedo putih bertambah dan menyebabkan lebih banyak jerawat kecil di wajah.	Negatif
2	Teksturnya agak cair, tidak lengket di kulit, dan mudah meresap. Aman untuk kulit berjerawat. Kemasan produk ini juga	Positif

No	Ulasan	Label
	bagus, sehingga cocok untuk dibawa saat bepergian.	
3	Sejak dulu saya cocok dengan produk Skintific, terutama sunscreen varian Aqua Light Daily ini. Teksturnya sedikit kental, ringan, dan cepat meresap, sehingga sangat cocok digunakan saat terburu-buru berangkat kerja. Ukurannya yang kecil juga memudahkan untuk dibawa kemana saja. Setelah menggunakan sunscreen ini, wajah saya terlihat lebih cerah dan glowing. Saya sangat menyukai sunscreen Skintific yang satu ini. Pengirimannya juga cukup cepat dan aman dengan pembungkusan bubble wrap yang tebal. Terima kasih, Skintific.	Positif
...	...	...
1500	Saya agak kaget karena isinya tidak sampai penuh, tetapi produknya baik. Cepat meresap dan bebas white cast.	Positif

### C. Pre-processing Data

Setelah data dilabeli, data akan melewati proses preprocessing sebelum dianalisis. Metode yang digunakan dalam preprocessing adalah stemming, tokenisasi, penghapusan stopword, dan casefolding. Berikut adalah contoh hasil dari proses pra-proses data yang terdapat dalam tabel 3.

TABEL 3. CONTOH HASIL PRE-PROCESSING DATA

Sebelum Pre-processing	Sesudah Pre-processing
Saya tidak terlalu menyukai sunscreen Skintific varian Aqua Light karena tidak cocok di wajah saya. Sunscreen ini justru membuat komedo putih bertambah dan menyebabkan lebih banyak jerawat kecil di wajah.	suka sunscreen skintific varian aqua light cocok wajah sunscreen komedo putih tambah sebab jerawat wajah

### D. Pembobotan Data

Setelah tahap pra-pemrosesan diselesaikan, langkah selanjutnya yaitu menerapkan TF-IDF berfungsi mengukur tingkat kepentingan setiap kata dalam korpus melalui kombinasi frekuensi lokal dan global. Teknik ini dimanfaatkan untuk menentukan relevansi suatu kata dalam konteks dokumen tertentu terhadap keseluruhan korpus. TF-IDF terdiri dari dua komponen utama: *Term Frequency* (TF) yang mengukur frekuensi kemunculan kata dalam dokumen, dan *Inverse Document Frequency* (IDF) yang mengurangi bobot kata-kata umum berdasarkan sebarannya dalam korpus. Dengan demikian, TF-IDF tidak hanya mencerminkan frekuensi lokal suatu istilah, tetapi juga memperhitungkan kekhasan atau nilai diskriminatif kata tersebut di antara seluruh dokumen dalam korpus

### E. Pemodelan Algoritma Naïve Bayes Classifier dan Support Vector Machine (SVM)

Setelah pemberian nilai representatif pada kata menggunakan TF-IDF, tahap berikutnya adalah klasifikasi sentimen setiap kalimat. Studi ini menggunakan dua algoritma pembelajaran mesin, yaitu *Naive Bayes* dan *Support Vector Machine* (SVM), yang dipilih karena keduanya cocok untuk analisis data berbasis teks. Naïve Bayes merupakan metode klasifikasi berbasis probabilistik yang didasarkan pada Teorema Bayes dengan asumsi bahwa setiap fitur bersifat independen. Dalam implementasi ini, setiap kalimat dievaluasi berdasarkan probabilitas keanggotaannya dalam kelas positif maupun negatif; klasifikasi dilakukan dengan memilih kelas yang memiliki nilai probabilitas tertinggi. Sebaliknya, algoritma *Support Vector Machine* menerapkan pendekatan berbasis margin, di mana data teks yang telah direpresentasikan dalam ruang vektor dipisahkan oleh sebuah hiperbidang optimal. Untuk meningkatkan kemampuan diskriminatif antara dua kelas sentimen, SVM memanfaatkan fungsi kernel yang memungkinkan pemetaan data ke dalam ruang fitur berdimensi lebih tinggi. Transformasi ini bertujuan untuk menghasilkan batas pemisah yang lebih jelas dan akurat, terutama ketika data tidak dapat dipisahkan secara linear dalam ruang asli.

### F. Evaluasi Model

Setelah menyelesaikan proses pemodelan algoritma, tahap berikutnya adalah evaluasi model, di mana prediksi yang dihasilkan oleh model akan dibandingkan dengan label sebenarnya. Evaluasi dilakukan dengan menggunakan confusion matrix, yang bertujuan untuk memberikan gambaran rinci tentang performa model klasifikasi. Dengan metode ini, kita dapat menganalisis seberapa baik model dapat memprediksi kelas dengan benar serta kesalahan yang muncul dalam prediksi. Berikut adalah hasil dari model algoritma yang terdapat dalam tabel 4.

TABEL 4. MODEL SEBELUM SMOTE DAN SESUDAH SMOTE

Modeling		Sebelum SMOTE				Sesudah SMOTE			
		Acc ura cy	Pre cisi on	Rec all	F1- Sco re	Acc ura cy	Pre cisi on	Rec all	F1- Sco re
NB	Ne gati f	81 %	1.0 0	0.1 2	0.2 2	93 %	0.8 8	0.7 8	0.8 3
	Pos itif		0.8 1	1.0 0	0.8 9		0.9 4	0.9 7	0.9 6
SV M	Ne gati f	92 %	0.8 4	0.7 7	0.8 0	91 %	0.7 8	0.7 8	0.7 8
	Pos itif		0.9 4	0.9 6	0.9 5		0.9 4	0.9 4	0.9 4

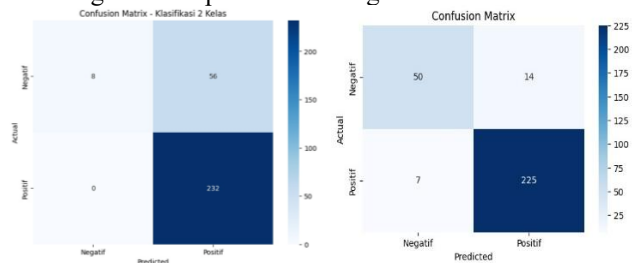
Hasil eksperimen menunjukkan bahwa pengoptimalan SMOTE meningkatkan kinerja model Naïve Bayes (NB) secara signifikan, tetapi mengurangi performa model Support Vector Machine (SVM). Sebelum penerapan SMOTE, Naïve Bayes menunjukkan tingkat akurasi 81%, yang mana pada sentimen negatif memiliki precision 1.00, recall 0.12, dan F1-Score 0.22; sementara pada sentimen positif, precision 0.81, recall 1.00, dan F1-Score 0.89. Setelah penerapan SMOTE, akurasi naik menjadi 93%, dengan sentimen negatif memiliki

precision 0.88, recall 0.78, dan F1-Score 0.83, dan untuk sentimen positif, precision 0.94, recall 0.97, serta F1-Score 0.96.

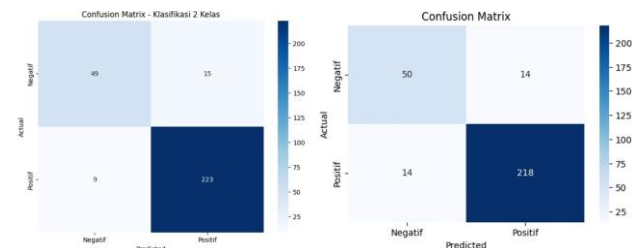
Sementara itu, sebelum SMOTE, SVM menunjukkan akurasi 92%, dengan sentimen negatif yang memiliki precision 0.84, recall 0.77, dan F1-Score 0.80; sedangkan pada sentimen positif, precision 0.94, recall 0.96, dan F1-Score 0.95. Namun, setelah SMOTE, terjadi penurunan akurasi menjadi 91%, dengan sentimen negatif menunjukkan precision 0.78, recall 0.78, dan F1-Score 0.78, sedangkan untuk sentimen positif, precision 0.94, recall 0.94, dan F1-Score 0.94.

Peningkatan kinerja Naïve Bayes setelah SMOTE terjadi karena algoritma ini sensitif terhadap ketidakseimbangan data, sehingga data sintetis membantu model mengenali pola pada kelas minoritas. Sebaliknya, kinerja SVM menurun karena data sintetis dari SMOTE dapat memperumit margin keputusan, membuat model lebih rentan terhadap noise dan kurang generalis.

Untuk menentukan model yang paling efektif, dilakukan perbandingan nilai confusion matrix dari kedua algoritma. Hasil confusion matrix untuk masing-masing model algoritma dapat dilihat sebagai berikut.



Gambar 2. Visualisasi Confusion Matrix Naïve Bayes



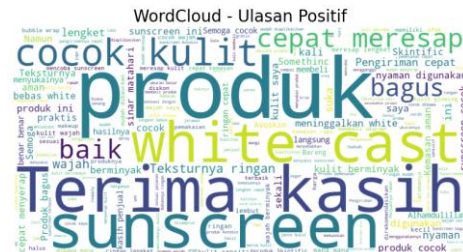
Gambar 3. Visualisasi Confusion Matrix SVM

Gambar tersebut menyajikan perbandingan hasil evaluasi model sebelum dan sesudah penerapan SMOTE. Pada visualisasi *Confusion Matrix* untuk algoritma *Naïve Bayes* sebelum penerapan SMOTE, tercatat 232 *True Positive* (TP), 8 *True Negative* (TN), 56 *False Positive* (FP), dan 0 *False Negative* (FN). Setelah diterapkan SMOTE, nilai evaluasi mengalami perubahan, yaitu 223 TP, 49 TN, 15 FP, dan 9 FN. Sementara itu, pada *Confusion Matrix* algoritma *Support Vector Machine* (SVM) sebelum SMOTE, diperoleh 223 TP, 49 TN, 15 FP, dan 9 FN. Setelah penerapan SMOTE, hasil evaluasi menunjukkan 218 TP, 50 TN, 14 FP, dan 14 FN. Perbandingan ini menunjukkan bahwa penggunaan SMOTE berdampak pada distribusi hasil klasifikasi, khususnya dalam meningkatkan representasi kelas minoritas.

### G. Visualisasi Data

Pada tahapan eksperimental dalam studi ini, digunakan pendekatan visualisasi *wordcloud* sebagai metode eksploratif untuk menganalisis karakteristik data teks. *Wordcloud* secara grafis merepresentasikan kata-kata yang terdapat dalam

korpus, dengan ukuran tipografi yang bervariasi sesuai dengan tingkat frekuensi kemunculan masing-masing kata, semakin tinggi frekuensinya, semakin besar pula ukuran visual yang ditampilkan. Teknik ini memungkinkan identifikasi cepat terhadap kata-kata dominan dalam teks. Hasil visualisasi *wordcloud* untuk ulasan bernuansa positif dan negatif masing-masing disajikan pada Gambar 4 dan Gambar 5.



Gambar 4. Word Cloud Ulasan Positif

Berdasarkan gambar tersebut dapat dilihat pada *wordcloud* ulasan positif, kata-kata seperti produk, terima kasih, sunscreen, teksturnya ringan, dan cocok untuk kulit mendominasi. Ini menunjukkan bahwa konsumen menghargai produk karena teksturnya ringan, cepat meresap, dan sesuai dengan jenis kulit mereka. Selain itu, banyak yang puas dengan pengiriman cepat dan pelayanan yang baik, ditunjukkan oleh kata-kata seperti pengiriman cepat dan terima kasih.



Gambar 5. Word Cloud Ulasan Negatif

Berdasarkan gambar tersebut dapat dilihat pada *wordcloud* ulasan negatif, kata-kata seperti produk, namun, isinya, teksturnya, dan wajah sering muncul. Kata-kata seperti kecewa, kecil, lengket, dan tidak sesuai menunjukkan bahwa ulasan negatif berfokus pada ketidakpuasan terhadap kualitas produk, ukuran kemasan, tekstur yang tidak nyaman, atau hasil yang tidak sesuai harapan. Selain itu, keluhan tentang harga juga muncul dalam beberapa ulasan.

## IV. KESIMPULAN

Berdasarkan analisis dan pengujian yang dilakukan mengenai sentimen produk sunscreen lokal dengan menggunakan metode Naïve Bayes dan Support Vector Machine (SVM), diperoleh kesimpulan bahwa penerapan teknik SMOTE secara signifikan dapat meningkatkan kinerja model Naïve Bayes. Setelah penerapan SMOTE, model Naïve Bayes menunjukkan peningkatan akurasi dari 81% menjadi 93%, dengan perbaikan nyata pada precision, recall, dan F1-Score untuk sentimen negatif maupun positif. Sebaliknya, penerapan SMOTE pada model SVM menyebabkan sedikit penurunan akurasi dari 92% menjadi 91%, meskipun metrik untuk sentimen positif tetap konsisten. Secara keseluruhan, model Naïve Bayes dengan SMOTE terbukti lebih efisien dalam memprediksi kategori sentimen

produk sunscreen lokal, dengan akurasi lebih tinggi dan keseimbangan yang lebih baik antara sentimen negatif dan positif dibandingkan SVM. Penelitian ini memberikan kontribusi teoritis dengan memberikan pemahaman mengenai dampak SMOTE terhadap dua algoritma berbeda pada data teks berbahasa Indonesia yang tidak seimbang. Namun, kontribusi praktis masih terbatas dan belum membahas skalabilitas model pada dataset besar maupun platform sosial media lain seperti TikTok dan Instagram yang memiliki karakteristik unik. Untuk pengembangan selanjutnya, direkomendasikan pemanfaatan model sebagai alat pemantauan sentimen real-time di e-commerce serta eksplorasi skalabilitas pada dataset heterogen dari berbagai platform sosial media agar model lebih adaptif dan aplikatif dalam strategi pemasaran sunscreen lokal. Langkah ini diharapkan dapat meningkatkan keandalan dan aplikasi praktis model dalam strategi pemasaran dan pengembangan produk sunscreen lokal secara lebih luas.

#### REFERENSI

- [1] J. Cardoso and E. Prasajo, "Pengaruh Kredibilitas Daya Tarik dan Kepercayaan terhadap Minat Pembelian Sunscreen Azarine di Kota Yogyakarta," *J. Econ. Business, Account. Manag.*, vol. 2, no. 1, pp. 124–140, 2024, doi: 10.61476/24zv8d23.
- [2] Via Aldina Auva and L. Nirawati, "Pengaruh Influencer Review dan Product Ingredients Terhadap Perilaku Konsumen pada Product Hydrasoothe Azarine Sunscreen Gel di Surabaya," *Al-Kharaj J. Ekon. Keuang. Bisnis Syariah*, vol. 6, no. 9, pp. 6771–6781, 2024, doi: 10.47467/alkharaj.v6i9.2820.
- [3] H. N. Sukmaningtyas, B. D. Maharani, and N. Kusuma, "Pengaruh Perceived Price dan Trust Terhadap Purchase Intention Melalui Perceived Value Sebagai Variabel Intervening Pada Produk Sunscreen Azarine," *Dyn. Manag. J.*, vol. 7, no. 2, pp. 179–188, 2023, [Online]. Available: <https://www.bing.com/ck/a?!&=&p=4e44e8cadbf7a3eJmltdHM9MTcwMDYxMTIwMCZpZ3VpZD0wMDNhM2E2Yi1kOTk0LTZkMmItMGEyYi0yODY2ZDhjMjZjNWQmaW5zaWQ9NTI4MA&pptn=3&ver=2&hsh=3&fclid=003a3a6b-d994-6d2b-0a2b-2866d8c26c5d&psq=http%3A%2F%2Fdx.doi.org%2F10.31000%2Fdmj.v7i2&u=>
- [4] A. N. S. Rahayu, T. I. Hermanto, and I. M. Nugroho, "Sentiment Analysis Using K-Nearest Neighbor Based on Particle Swarm Optimization According To Sunscreen'S Reviews," *J. Tek. Inform.*, vol. 3, no. 6, pp. 1639–1646, 2022, doi: 10.20884/1.jutif.2022.3.6.425.
- [5] N. S. Andini, "Compas Market Insight Dashboard: Alami Penurunan Tajam di Desember 2023, Azarine dan Facetology Berhasil Tingkatkan Nilai Penjualan Kategori Sunscreen Sepanjang Periode Q1 2024," [compas.co.id](https://compas.co.id). Accessed: May 26, 2025. [Online]. Available: [https://compas.co.id/article/nilai-penjualan-kategori-sunscreen-q1-2024/?utm\\_source=chatgpt.com](https://compas.co.id/article/nilai-penjualan-kategori-sunscreen-q1-2024/?utm_source=chatgpt.com)
- [6] N. Luh and P. Septyani, "Pengaruh Harga Terhadap Keputusan Pembelian Produk Skincare di Kota Denpasar pada E-Commerce Shopee Dengan Kepuasan Konsumen Sebagai Variabel Mediasi," *J. Manajemen, Kewirausahaan dan Pariwisata Juni*, vol. 4, no. 6, pp. 1153–1163, 2021.
- [7] R. N. Anwar and F. A. Wardani, "Pengaruh Kualitas Produk Dan Kualitas Pelayanan Terhadap Minat Beli Ulang Produk Scarlett Di E-Commerce Shopee," *Nusant. Ilmu Pengetah. Sos.*, vol. 8, no. 5, pp. 1370–1379, 2021, [Online]. Available: <http://jurnal.um-tapsel.ac.id/index.php/nusantara/index>
- [8] M. R. Putri and K. B. Sienarta, "Pemediasian Antara Pengaruh E-Wom Terhadap Repurchase Intention Dengan Mediasi Brand Image Pada Produk Skincare Merek Lokal Indonesia Yang Ada Di E-Commerce," *Parsimonia - J. Ekon. dan Bisnis*, vol. 10, no. 2, pp. 100–109, 2023, doi: 10.33479/parsimonia.v10i2.779.
- [9] D. Meisa Azzahra, M. Hafid Totohendarto, and S. Alam, "Analisis Sentimen Ulasan Produk Serum Wajah Pada Beauty Brand Somethinc Menggunakan Metode Naïve Bayes Classifier," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 3, pp. 1604–1611, 2023, doi: 10.36040/jati.v7i3.6929.
- [10] N. Tf-id, "Analisis Sentimen Ulasan Herborist Sistem Pengambilan Keputusan Menggunakan Klasifikasi," *J. SWABUMI*, vol. 12, no. 2, pp. 176–181, 2024.
- [11] M. Hamka, N. Alfatari, and D. Ratna Sari, "Analisis Sentimen Produk Kecantikan Jenis Serum Menggunakan Algoritma Naïve Bayes Classifier," *J. Sist. Komput. dan Inform.*, vol. 4, no. 1, p. 64, 2022, doi: 10.30865/json.v4i1.4740.
- [12] D. A. WP, J. D. Firizqi, and Z. A. Amalia, "Analisis Sentimen Produk Skincare Somethinc Niacinamide di Female Daily dengan Naïve Bayes Classifier," *J. Media Inform. Budidarma*, vol. 8, no. 2, p. 946, 2024, doi: 10.30865/mib.v8i2.7571.
- [13] H. Harnelia, "Analisis Sentimen Review Skincare Skintific Dengan Algoritma Support Vector Machine (Svm)," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 2, 2024, doi: 10.23960/jitet.v12i2.4095.
- [14] K. S. Putri, I. R. Setiawan, and A. Pambudi, "Analisis Sentimen Terhadap Brand Skincare Lokal Menggunakan Naïve Bayes Classifier," *Technol. J. Ilm.*, vol. 14, no. 3, p. 227, 2023, doi: 10.31602/tji.v14i3.11259.
- [15] V. D. Yunanda and N. Hendrastuty, "Perbandingan Kernel Polynomial dan RBF Pada Algoritma SVM Untuk Analisis Sentimen Skincare di Indonesia," *J. Media Inform. Budidarma*, vol. 8, no. 2, p. 726, 2024, doi: 10.30865/mib.v8i2.7425.
- [16] F. Novitasari and M. D. Purbolaksono, "Analysis Sentiment Aspect Level on Beauty Product Reviews Using Chi-Square and Naïve Bayes," *J. Data ...*, no. January, pp. 18–30, 2021, [Online]. Available: <http://commdis.telkomuniversity.ac.id/jdsa/index.php/jdsa/article/view/72%0Ahttps://scholar.archive.org/work/kefndqsmtz4foilip5bh41vbe/access/wayback/https://commdis.telkomuniversity.ac.id/jdsa/index.php/jdsa/article/download/72/30/>
- [17] Melisa Nur Aini, Rita Yulfani, and Nurul Jariah, "Penerapan Metode Naïve Bayes Untuk Analisis

- Sentimen Ulasan Produk Sunscreen Berdasarkan Female Daily Review,” *J. Multimed. dan Teknol. Inf.*, vol. 6, no. 01, pp. 24–34, 2024, doi: 10.54209/jatilima.v6i01.421.
- [18] Muhammad Rio Pratama, Faza Abdillah Gunawan Soerawinata, Rafdi Reyhan Zhafari, Rendy, and Helena Nurramdhani Imanda, “Sentiment Analysis of Beauty Product E-Commerce Using Support Vector Machine Method,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 2, pp. 269–274, 2022, doi: 10.29207/resti.v6i2.3876.
- [19] Cemara Nimas, “6 Top Produk Sunscreen Paling Laris di E-commerce,” *Compas.co.id*. Accessed: Jun. 30, 2025. [Online]. Available: <https://compas.co.id/article/produk-sunscreen-paling-laris/>
- [20] R. P. Wiwaha, “Data Penjualan Sunscreen di Tokopedia dan Blibli, Ini 7 Brand Terlarisnya,” *compas.co.id*. Accessed: Jun. 30, 2025. [Online]. Available: <https://compas.co.id/article/sunscreen-kulit-kering/>
- [21] I. D. Rukmana, “10 Top Brand Sunscreen Lokal Terlaris 2022, Tidak Takut Matahari Lagi!,” *compas.co.id*. Accessed: Jun. 30, 2025. [Online]. Available: <https://compas.co.id/article/10-top-brand-sunscreen-lokal-terlaris-2022/>
- [22] T. Astuti and Y. Astuti, “Analisis Sentimen Review Produk Skincare Dengan Naïve Bayes Classifier Berbasis Particle Swarm Optimization (PSO),” *J. Media Inform. Budidarma*, vol. 6, no. 4, p. 1806, 2022, doi: 10.30865/mib.v6i4.4119.
- [23] Jasmarizal, Junadhi, Rahmaddeni, and M. Khairul Anam, “Penerapan Metode Support Vector Machine Untuk Analisis Sentimen Terhadap Produk Skincare,” *Indones. J. Comput. Sci.*, vol. 13, no. 1, pp. 1438–1450, 2024, doi: 10.33022/ijcs.v13i1.3654.
- [24] P. M. Berlianti and E. Y. Hidayat, “Implementasi Naïve Bayes Classifier untuk Sentimen Produk Kecantikan Berdasarkan Ulasan Female Daily,” *Indones. J. Comput. Sci.*, vol. 13, no. 6, pp. 10248–10256, 2024.
- [25] R. W. Pratiwi, S. F. H. D. Dairoh, D. I. Af’idah, Q. R. A., and A. G. F., “Analisis Sentimen Pada Review Skincare Female Daily Menggunakan Metode Support Vector Machine (SVM),” *J. Informatics, Inf. Syst. Softw. Eng. Appl.*, vol. 4, no. 1, pp. 40–46, 2021, doi: 10.20895/inista.v4i1.387.
- [26] I. N. T. I. Nabawi, Rudiman, and F. Yulianto, “Analisis Sentimen Pada Ulasan Aplikasi Google Maps Terhadap Pelayanan Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan Samarinda Menggunakan Metode K-Nearest Neighbor Dengan Fitur Ekstraksi TF-IDF,” *J. Teknol. Inf.*, vol. Vol 18, no. 2, p. 17, 2024.
- [27] O. I. Gifari, M. Adha, F. Freddy, and F. F. S. Durrand, “Film Review Sentiment Analysis Using TF-IDF and Support Vector Machine,” *J. Inf. Technol.*, vol. 2, no. 1, pp. 36–40, 2022.
- [28] R. Wati, S. Ernawati, and H. Rachmi, “Pembobotan TF-IDF Menggunakan Naïve Bayes pada Sentimen Masyarakat Mengenai Isu Kenaikan BIPIH,” *J. Manaj. Inform.*, vol. 13, no. 1, pp. 84–93, 2023, doi: 10.34010/jamika.v13i1.9424.
- [29] D. Atmajaya, A. Febrianti, and H. Darwis, “Metode SVM dan Naive Bayes untuk Analisis Sentimen ChatGPT di Twitter,” *Indones. J. Comput. Sci.*, vol. 12, no. 4, pp. 2173–2181, 2023, doi: 10.33022/ijcs.v12i4.3341.
- [30] D. F. Zhafira, B. Rahayudi, and I. Indriati, “Analisis Sentimen Kebijakan Kampus Merdeka Menggunakan Naive Bayes dan Pembobotan TF-IDF Berdasarkan Komentar pada Youtube,” *J. Sist. Informasi, Teknol. Informasi, dan Edukasi Sist. Inf.*, vol. 2, no. 1, pp. 55–63, 2021, doi: 10.25126/justsi.v2i1.24.
- [31] K. Anwar, “Analisa sentimen Pengguna Instagram Di Indonesia Pada Review Smartphone Menggunakan Naive Bayes,” *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 2, no. 4, pp. 148–155, 2022, doi: 10.30865/klik.v2i4.315.
- [32] E. Suryati, Styawati, and A. A. Aldino, “Analisis Sentimen Transportasi Online Menggunakan Ekstraksi Fitur Model Word2vec Text Embedding Dan Algoritma Support Vector Machine (SVM),” *J. Teknol. Dan Sist. Inf.*, vol. 4, no. 1, pp. 96–106, 2023, [Online]. Available: <https://doi.org/10.33365/jtsi.v4i1.2445>
- [33] J. Anggraini and D. Alita, “Implementasi Metode SVM Pada Sentimen Analisis Terhadap Pemilihan Presiden (Pilpres) 2024 Di Twitter,” *J. Inform. J. Pengemb. IT*, vol. 9, no. 2, pp. 102–111, 2024, doi: 10.30591/jpit.v9i2.6560.
- [34] 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.
- [35] B. Ramadhani and R. R. Suryono, “Komparasi Algoritma Naïve Bayes dan Logistic Regression Untuk Analisis Sentimen Metaverse,” *J. Media Inform. Budidarma*, vol. 8, no. 2, p. 714, 2024, doi: 10.30865/mib.v8i2.7458.
- [36] M. R. Humaidi and A. Maulani, “Klasifikasi Naïve Bayes Dan Confusion Matrix Pada Pengguna Aplikasi E-Commerce Di Play Store,” *J. Ilm. Inform.*, vol. 8, no. 2, pp. 132–139, 2023.