

Analisis Sentimen Netizen YouTube Terhadap Keracunan Program MBG Perbandingan Metode SVM dan Naïve Bayes

Djati Lumayung*¹, Khairullah²

^{1,2}Prodi Studi Teknik Informatika, Universitas Muhammadiyah Bengkulu

E-mail: ¹djatilumayung931@gmail.com, ²khairullah@umb.ac.id

Abstrak

Kasus Keracunan Masal Pada Program Makan Bergizi Gratis (MBG) yang terjadi di berbagai daerah Indonesia sepanjang tahun 2025 memicu gelombang besar opini publik di kolom komentar YouTube. Penelitian ini bertujuan membandingkan performa algoritma Support Vector Machine (SVM) dan Naïve Bayes (NB) dalam menganalisis sentimen komentar netizen YouTube terkait kasus keracunan MBG. Sebanyak 2.039 komentar berbahasa Indonesia dikumpulkan menggunakan YouTube Data API v3, kemudian dilabeli ke dalam tiga kelas sentimen: Positif, Negatif, dan Netral. Tahapan preprocessing meliputi case folding, cleansing, normalisasi kata tidak baku, tokenisasi, stopword removal, dan stemming menggunakan library Sastrawi. Ekstraksi fitur dilakukan dengan TF-IDF dan data dibagi dengan rasio 80% training dan 20% testing. Hasil distribusi menunjukkan sentimen Netral mendominasi (46,2%), diikuti Negatif (37,8%) dan Positif (16,0%). Evaluasi menggunakan akurasi, presisi, recall, dan F1-Score menunjukkan SVM memperoleh akurasi 72,55% dengan F1-Score 71,62%, lebih unggul dibandingkan Naïve Bayes yang memperoleh akurasi 64,22% dan F1-Score 59,84%. SVM terbukti lebih optimal untuk klasifikasi sentimen komentar YouTube berbahasa Indonesia pada konteks kebijakan publik.

Kata kunci - Analisis Sentimen, Support Vector Machine, Naïve Bayes, YouTube, Makan Bergizi Gratis

1. PENDAHULUAN

Salah satu upaya pemerintah dalam meningkatkan kualitas gizi anak di Indonesia adalah melalui Program Makan Bergizi Gratis (MBG) yang diperkenalkan oleh Presiden Prabowo Subianto [1]. Program ini dirancang untuk menyediakan makanan bergizi secara gratis dengan harapan dapat membantu mengatasi masalah kekurangan gizi dan mendukung tumbuh kembang anak secara optimal. Perkembangan media sosial menyebabkan berbagai informasi mengenai pelaksanaan program tersebut cepat menyebar dan memicu beragam respons dari masyarakat. YouTube menjadi salah satu media yang aktif digunakan pengguna untuk menyampaikan pandangan mereka melalui kolom komentar. Kumpulan komentar tersebut dapat dimanfaatkan sebagai sumber informasi untuk mengidentifikasi opini publik terhadap MBG, terutama pada saat muncul peristiwa seperti kasus keracunan yang menimbulkan perhatian luas [2].

Pada bulan Oktober 2025, sejumlah insiden keracunan pangan yang berkaitan dengan program MBG mulai dilaporkan dari berbagai daerah di Indonesia [3]. Berdasarkan laporan Kementerian Kesehatan RI tertanggal 5 Oktober 2025, tercatat sebanyak 119 kejadian keracunan pangan tersebar di 25 provinsi (88 kabupaten/kota), dengan total korban mencapai 11.660 jiwa. Peristiwa ini tentu saja memicu reaksi luas di media sosial. Para *netizen* ramai-ramai memberikan komentar di berbagai platform, termasuk YouTube, sehingga dalam waktu singkat terkumpul ribuan opini dari berbagai sudut pandang. Sayangnya, data opini publik yang begitu besar ini belum pernah dianalisis secara sistematis khususnya untuk mengukur polaritas sentimen *netizen*

YouTube terhadap kasus keracunan MBG secara terukur dan objektif. Inilah yang menjadi titik tolak penelitian ini.

Untuk menjawab kebutuhan tersebut, penelitian ini menggunakan pendekatan analisis sentimen. Secara sederhana, analisis sentimen adalah teknik komputasi yang digunakan untuk menentukan apakah sebuah teks mengandung opini positif, negatif, atau netral. Sebagai bagian dari text mining, metode ini memanfaatkan algoritma machine learning untuk memproses ribuan komentar secara otomatis, sehingga bisa menghasilkan gambaran yang terukur tentang bagaimana publik bersikap terhadap suatu isu[4]. Dalam konteks isu sensitif seperti kasus keracunan Program MBG, pendekatan ini sangat berguna untuk membantu pemerintah dan pemangku kepentingan memahami persepsi publik secara cepat. Namun demikian, hasil analisis sentimen sangat bergantung pada pemilihan algoritma klasifikasi yang tepat, sehingga perlu dilakukan perbandingan untuk menemukan metode yang paling sesuai dengan karakteristik data.

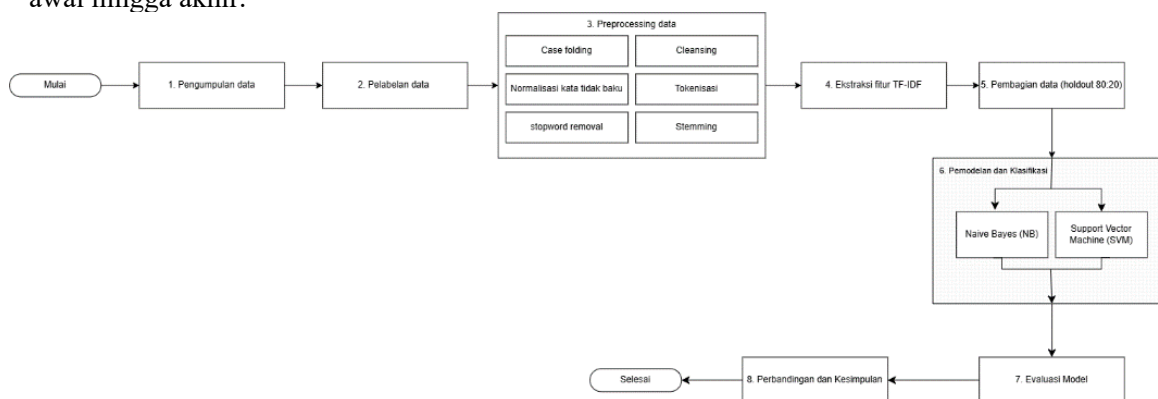
Dua algoritma yang cukup banyak digunakan dalam penelitian analisis sentimen adalah Support Vector Machine (SVM) dan Naïve Bayes (NB). Dari beberapa studi yang telah dilakukan, hasilnya tidak selalu konsisten. Pada kasus analisis sentimen cyberbullying di aplikasi X misalnya, Naive Bayes sedikit lebih unggul dengan akurasi 87% dibanding SVM yang memperoleh 86%[5]. Namun penelitian Sarimole terhadap ulasan aplikasi Satu Sehat di Twitter memberikan gambaran yang berbeda SVM justru lebih unggul dengan akurasi 87,95% dibanding Naive Bayes yang hanya mencapai 81,65% [6]. Perbedaan temuan ini mengisyaratkan bahwa tidak ada satu algoritma yang selalu unggul di semua kondisi performa keduanya sangat bergantung pada jenis data, konteks bahasa, dan topik yang dianalisis.

Atas dasar itu, penelitian ini mencoba mengisi celah tersebut dengan membandingkan algoritma Support Vector Machine dan Naïve Bayes dalam mengklasifikasikan opini *netizen*. Dengan membandingkan kedua algoritma pada dataset yang sama, penelitian ini diharapkan dapat memberikan rekomendasi yang lebih konkret mengenai metode analisis sentimen yang paling efektif untuk konteks komentar berbahasa Indonesia di media sosial, sekaligus memberi gambaran nyata tentang persepsi publik terhadap program strategis pemerintah.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif berbasis komputasi dengan menerapkan metode analisis sentimen menggunakan algoritma Support Vector Machine (SVM) dan Naïve Bayes (NB) secara komparatif. Alur penelitian secara keseluruhan terdiri dari enam tahapan utama yang saling berurutan, yaitu: (1) pengumpulan data, (2) pelabelan data, (3) preprocessing data, (4) ekstraksi fitur, (5) pembagian data, (6) pemodelan dan klasifikasi, (7) evaluasi model, serta (8) perbandingan dan kesimpulan, sebagaimana ditunjukkan pada Gambar 1 berikut.

Keenam tahapan tersebut disajikan secara visual dalam bentuk diagram alur pada Gambar 1 berikut, guna memberikan gambaran menyeluruh mengenai rangkaian proses penelitian dari awal hingga akhir.



Gambar 1 Tahapan Penelitian

Sebagaimana ditunjukkan pada Gambar 1, proses penelitian dimulai dari tahap pengumpulan data komentar YouTube, dilanjutkan dengan pelabelan sentimen berbasis leksikon, preprocessing teks, ekstraksi fitur menggunakan TF-IDF, pembagian data, pemodelan dan klasifikasi menggunakan algoritma SVM dan Naïve Bayes, evaluasi model, hingga tahap akhir berupa perbandingan dan penarikan kesimpulan. Alur ini disusun secara sekuensial, di mana keluaran dari satu tahap menjadi masukan bagi tahap berikutnya.

2.1 Pengumpulan Data

Data dalam penelitian ini diambil dari komentar publik pada video-video YouTube yang membahas kasus keracunan Program MBG. Pemilihan video dilakukan secara purposif, yaitu memprioritaskan video-video viral dari kanal berita besar dengan jumlah komentar tinggi, seperti Kompas TV, CNN Indonesia, dan kanal berita nasional lainnya yang secara spesifik mengangkat isu keracunan MBG. Pengambilan data dilakukan menggunakan YouTube Data API v3 berbasis Python yang dijalankan melalui Google Colaboratory, sehingga komentar bisa diekstraksi secara otomatis dan terstruktur. Dari proses ini terkumpul 2.157 komentar berbahasa Indonesia yang relevan dengan kasus keracunan MBG, diambil sejak isu ini mulai ramai dibahas di YouTube hingga jumlah data dianggap cukup. Komentar yang dimasukkan ke dalam dataset harus memenuhi tiga kriteria: berbahasa Indonesia, mengandung opini atau ekspresi sentimen yang jelas, dan berkaitan langsung dengan kasus keracunan MBG. Sementara itu, komentar yang berisi tautan, spam, atau hanya emoji tanpa teks tidak dimasukkan ke dalam dataset, sehingga jumlah data final yang digunakan untuk pelatihan dan pengujian model menjadi 2.039 komentar.

2.2 Pelabelan Data

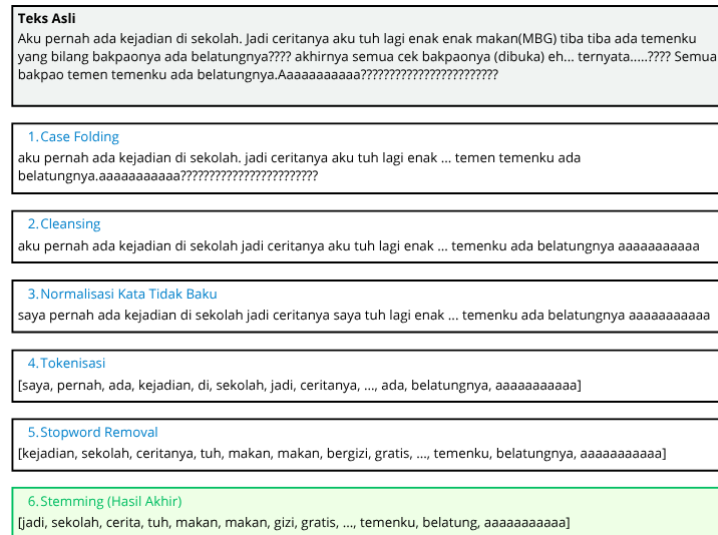
Untuk memberi label sentimen pada setiap komentar, penelitian ini menggunakan pendekatan berbasis leksikon (*lexicon-based*) yang dijalankan melalui Google Colaboratory [7]. Leksikon yang digunakan disusun secara mandiri oleh peneliti dan disesuaikan dengan konteks isu keracunan Program MBG, terdiri atas kamus kata positif dan kamus kata negatif berbahasa Indonesia. Kamus kata positif mencakup tiga kategori: (1) ungkapan dukungan umum, misalnya "bagus", "mendukung", "apresiasi", "berhasil", dan "bermanfaat"; (2) istilah terkait konteks program, misalnya "bergizi", "solusi", "inovasi", dan "diperbaiki"; serta (3) ungkapan apresiasi terhadap kebijakan pemerintah, misalnya "kerja bagus", "peduli", dan "progresif". Kamus kata negatif mencakup lima kategori: (1) istilah terkait dampak kesehatan, misalnya "keracunan", "muntah", "diare", dan "rawat inap"; (2) kritik terhadap pelaksanaan program, misalnya "gagal", "amburadul", "tidak higienis", dan "kadaluarsa"; (3) ekspresi kekecewaan, misalnya "kecewa", "mengecewakan", dan "memprihatinkan"; (4) dugaan penyimpangan pengelolaan, misalnya "korupsi", "manipulasi", dan "mark up"; serta (5) ungkapan penolakan, misalnya "tolak", "hentikan", dan "batalkan". Skor sentimen masing-masing komentar dihitung berdasarkan selisih jumlah kata yang cocok dengan kamus kata positif dan kamus kata negatif tersebut, dengan aturan: skor positif menghasilkan label Positif, skor negatif menghasilkan label Negatif, dan skor nol (tidak ada kata kunci yang cocok atau jumlah kata positif dan negatif seimbang) menghasilkan label Netral.

Untuk memvalidasi keakuratan label yang dihasilkan secara otomatis oleh pendekatan *lexicon-based*, penelitian ini melakukan validasi manual pada sampel acak sebanyak 150 komentar (7,4% dari total 2.039 data), yang diambil secara stratified mengikuti proporsi distribusi label pada data lengkap. Sampel tersebut diberi label ulang secara independen oleh peneliti tanpa melihat label hasil *lexicon-based* terlebih dahulu, kemudian dibandingkan untuk mengukur tingkat kesesuaian. Hasil validasi menunjukkan tingkat kesesuaian (*agreement rate*) sebesar 74,00% dengan nilai Cohen's Kappa sebesar 0,5456, yang menurut kategorisasi Landis dan Koch (1977) termasuk dalam kategori kesesuaian moderat (*moderate agreement*). Nilai ini menunjukkan bahwa label yang dihasilkan oleh pendekatan *lexicon-based* cukup dapat diandalkan untuk digunakan sebagai *ground truth* dalam melatih dan mengevaluasi model klasifikasi SVM dan Naïve Bayes pada penelitian ini, meskipun demikian tingkat kesesuaian yang moderat juga mengindikasikan adanya ruang perbaikan pada cakupan kamus leksikon, khususnya untuk menangkap ekspresi sentimen implisit yang tidak menggunakan kata kunci eksplisit dalam

kamus. Pendekatan *lexicon-based* dipilih dalam penelitian ini karena lebih efisien untuk memproses data dalam jumlah besar secara konsisten dan objektif, dibandingkan jika dilakukan pelabelan manual[8].

2.3 Preprocessing Data

Untuk memperjelas proses preprocessing data, Gambar 2 menampilkan ilustrasi transformasi teks pada salah satu komentar dataset secara bertahap, mulai dari teks asli hingga hasil akhir setelah melalui keenam tahap preprocessing [9].



Gambar 2 diagram bertahap menunjukkan transformasi 1 komentar asli melalui 6 tahap preprocessing Terdapat enam tahap yang dilakukan. Tahap pertama adalah *case folding*, yaitu mengubah seluruh huruf menjadi huruf kecil [10]. Tahap berikutnya adalah *cleansing*, yang membersihkan teks dari karakter-karakter yang tidak relevan seperti URL, mention (@), tagar (#), angka, tanda baca, dan karakter non-alfabet lainnya. Tahap ketiga adalah normalisasi kata tidak baku, yaitu mengubah kata-kata slang atau singkatan yang umum dipakai di media sosial menjadi bentuk baku (contohnya “gk” menjadi “tidak”, “bgt” menjadi “sangat”) menggunakan kamus normalisasi bahasa Indonesia [11]. Tahap keempat adalah *tokenization*, yang memecah teks menjadi satuan kata atau token. Tahap kelima adalah *stopword removal*, yaitu membuang kata-kata yang tidak memiliki makna penting menggunakan daftar *stopword* bahasa Indonesia. Tahap terakhir adalah *stemming*, yang mengubah kata-kata berimbuhan kembali ke bentuk kata dasarnya menggunakan library Sastrawi yang memang dirancang khusus untuk bahasa Indonesia [12].

2.4 Ekstraksi Fitur dengan TF-IDF

Setelah melalui tahap preprocessing, teks komentar perlu diubah menjadi representasi numerik agar bisa diproses oleh algoritma machine learning. Untuk itu, penelitian ini menggunakan metode TF-IDF (*Term Frequency-Inverse Document Frequency*). Secara sederhana, TF-IDF menggabungkan dua ukuran: TF (Term Frequency) yang menghitung seberapa sering sebuah kata muncul dalam satu dokumen, dan IDF (Inverse Document Frequency) yang mengukur seberapa jarang kata tersebut muncul di seluruh kumpulan dokumen. Nilai TF dihitung sebagaimana ditunjukkan pada persamaan (1) berikut

$$TF(t, d) = \frac{f(t, d)}{\sum f(t', d)} \quad (1)$$

di mana $f(t, d)$ menyatakan jumlah kemunculan kata t pada dokumen d .

Sementara itu, IDF dihitung sebagaimana ditunjukkan pada persamaan (2) berikut

$$IDF(t) = \log \left(\frac{N}{df(t)} \right) \quad (2)$$

di mana N adalah jumlah total dokumen dan $df(t)$ adalah jumlah dokumen yang memuat kata t . Skor TF-IDF akhir untuk setiap kata diperoleh sebagaimana ditunjukkan pada persamaan (3).

$$TF(t, d) \times IDF(t) \quad (3)$$

Dalam penelitian ini, perhitungan TF-IDF diimplementasikan menggunakan *TfidfVectorizer* dari library Scikit-learn, dengan konfigurasi kombinasi *unigram* dan *bigram* (*ngram_range*=(1,2)) sehingga model dapat menangkap konteks dari frasa dua kata, bukan hanya kata tunggal [13][14]. Ringkasan hasil proses ekstraksi fitur TF-IDF ditunjukkan pada Tabel 1 berikut.

Tabel 1 Ekstraksi fitur TF-IDF

Jumlah dokumen	2039
Jumlah fitur	2698
Label kelas	'Negatif', 'Netral', 'Positif'

Berdasarkan Tabel 1, proses ekstraksi fitur menghasilkan 2.698 fitur (kombinasi unigram dan bigram) dari 2.039 dokumen komentar yang telah melalui tahap preprocessing, dengan tiga label kelas sentimen yaitu Negatif, Netral, dan Positif. Jumlah fitur yang cukup besar ini mencerminkan keragaman kosakata yang muncul dalam komentar-komentar terkait isu keracunan Program MBG.

Dua puluh fitur dengan bobot TF-IDF rata-rata tertinggi ditampilkan pada Tabel 2 berikut, untuk memberikan gambaran mengenai kata atau frasa yang paling dominan dalam dataset.

Tabel 2 20 fitur TF-IDF Tertinggi

makan	0.0761
gratis	0.0668
gizi	0.0585
makan gizi	0.0568
gizi gratis	0.0561
racun	0.0356
gratis makan	0.0256
anak	0.0237
makan racun	0.0217
sekolah	0.0207
racun gratis	0.0189
program	0.0160
basi	0.0153
kasih	0.0125
masak	0.0124
korupsi	0.0123
uang	0.0113
perintah	0.0110
henti	0.0107
indonesia	0.0099

Sebagaimana terlihat pada Tabel 2, kata-kata dengan bobot TF-IDF tertinggi didominasi oleh istilah yang berkaitan langsung dengan topik penelitian, seperti “makan”, “gratis”, dan “gizi”, yang merupakan bagian dari frasa “makan bergizi gratis” (MBG). Tingginya bobot kata-kata tersebut menunjukkan bahwa TF-IDF berhasil menangkap istilah-istilah kunci yang relevan dengan konteks isu yang dibahas dalam komentar.

2.5 Pembagian Data dan Pemodelan

Dari total 2.039 komentar yang telah melalui tahap preprocessing, dataset kemudian dibagi menjadi dua bagian menggunakan metode *holdout*. Pembagian dilakukan dengan rasio 80% untuk data latih (*training set*), yang setara dengan 1.631 komentar, dan 20% untuk data uji (*testing set*) sebanyak 408 komentar. Proses pembagian ini dilakukan secara acak dengan fungsi *train_test_split* dari Scikit-learn, dengan nilai parameter *random_state* yang dikunci agar hasil pembagian data tetap konsisten setiap kali dijalankan (*reproducibility*). Setelah pembagian data selesai, dua model klasifikasi dibangun dan dilatih secara terpisah pada data latih. Model pertama adalah SVM dengan kernel Linear, dibangun menggunakan kelas *LinearSVC* dari Scikit-learn.

Model kedua adalah Naïve Bayes, tepatnya varian *Multinomial Naïve Bayes*, yang dibangun menggunakan kelas *MultinomialNB* dari Scikit-learn. Penting untuk dicatat bahwa kedua model dilatih menggunakan representasi vektor TF-IDF yang sama persis, sehingga perbandingan performa antar keduanya bisa dilakukan secara adil.

Tabel 3 pembagian data

Keterangan	Jumlah	Persentase
Data Training	1.631	(80%)
Data Testing	408	(20%)

Berdasarkan Tabel 3, dari total 2.039 komentar yang telah melalui tahap preprocessing, sebanyak 1.631 komentar (80%) dialokasikan sebagai data latih dan 408 komentar (20%) sebagai data uji. Pembagian ini menggunakan metode holdout dengan stratifikasi, sehingga proporsi masing-masing kelas sentimen tetap terjaga baik pada data latih maupun data uji. Distribusi label sentimen pada data latih ditunjukkan pada Tabel 4 berikut.

Tabel 4 Distribusi label di training

Label	Jumlah
Netral	758
Negatif	613
Positif	260

Tabel 4 menunjukkan bahwa pada data latih, kelas Netral mendominasi dengan 758 komentar, diikuti oleh kelas Negatif sebanyak 613 komentar, dan kelas Positif sebanyak 260 komentar. Proporsi ini konsisten dengan distribusi label pada keseluruhan dataset, yang mengindikasikan bahwa proses pembagian data secara stratified berhasil mempertahankan keseimbangan proporsi antar kelas.

Sebagai pembandingan, distribusi label sentimen pada data uji disajikan pada Tabel 5 berikut.

Tabel 5 Distribusi label di testing

Label	Jumlah
Netral	190
Negatif	153
Positif	65

Sebagaimana ditunjukkan pada Tabel 5, proporsi ketiga kelas pada data uji (Netral 190, Negatif 153, dan Positif 65 komentar) sebanding dengan proporsi pada data latih di Tabel 4, yang menegaskan bahwa metode pembagian data stratified telah diterapkan secara konsisten pada kedua subset data.

2.6 Evaluasi Model

Untuk menilai performa kedua model, penelitian ini menggunakan empat metrik evaluasi yang umum dipakai dalam penelitian klasifikasi teks: Accuracy, Precision, Recall, dan F1-Score [15]. Nilai Accuracy, Precision, recall, dan F1-Score dihitung menggunakan komponen True Positive (TP), True Negative (TN), False Positive (FP), serta False Negative (FN) yang dihasilkan dari confusion matrix sebagai dasar evaluasi model [16]. Rumus masing-masing metrik ditunjukkan pada persamaan (4) hingga (7) berikut, yaitu Accuracy, Precision, Recall, dan F1-Score secara berurutan.

$$\frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (4)$$

$$\frac{TP}{(TP + FP)} \quad (5)$$

$$\frac{TP}{(TP + FN)} \quad (6)$$

$$2 \times \frac{(Precision \times recall)}{(Precision + Recall)} \quad (7)$$

Karena penelitian ini melibatkan tiga kelas sentimen (Positif, Negatif, Netral), nilai Precision, Recall, dan F1-Score dihitung dengan pendekatan weighted average, sehingga

ketidakseimbangan jumlah data antar kelas tetap diperhitungkan. Di samping keempat metrik tersebut, confusion matrix juga divisualisasikan untuk menggambarkan secara lebih jelas bagaimana distribusi hasil prediksi masing-masing model [17]. Berdasarkan keempat metrik inilah perbandingan performa antara SVM dan Naïve Bayes dilakukan, untuk melihat algoritma mana yang bekerja lebih baik [4] dalam mengklasifikasikan sentimen komentar YouTube terhadap kasus keracunan Program MBG.

3. HASIL DAN PEMBAHASAN

Bagian ini memaparkan hasil penelitian secara sistematis mengikuti alur tahapan penelitian, mulai dari pengumpulan dan pelabelan data, proses preprocessing, distribusi sentimen, hingga evaluasi dan perbandingan performa kedua model klasifikasi.

3.1 Pengumpulan dan Pelabelan Data

Proses pengumpulan data dilakukan menggunakan YouTube Data API v3 berbasis Python melalui Google Colaboratory. Data dikumpulkan dari kolom komentar video-video viral pada kanal berita nasional besar BBC News Indonesia, Kompas.com, dan kanal sejenis yang secara spesifik membahas kasus keracunan Program Makan Bergizi Gratis (MBG). Total komentar yang berhasil dikumpulkan adalah sebanyak 2.157 komentar berbahasa Indonesia. Komentar yang mengandung spam, tautan, atau hanya berisi emoji tanpa teks dieksklusi dari dataset, sehingga jumlah data yang digunakan dalam penelitian ini menjadi 2.039 komentar. Informasi umum mengenai data mentah hasil pengumpulan ditunjukkan pada Tabel 6 berikut.

Tabel 6 Info Data Mentah

Jumlah baris	2157
Jumlah kolom	5
Nama kolom	komentar, username, likes, tanggal, video_id

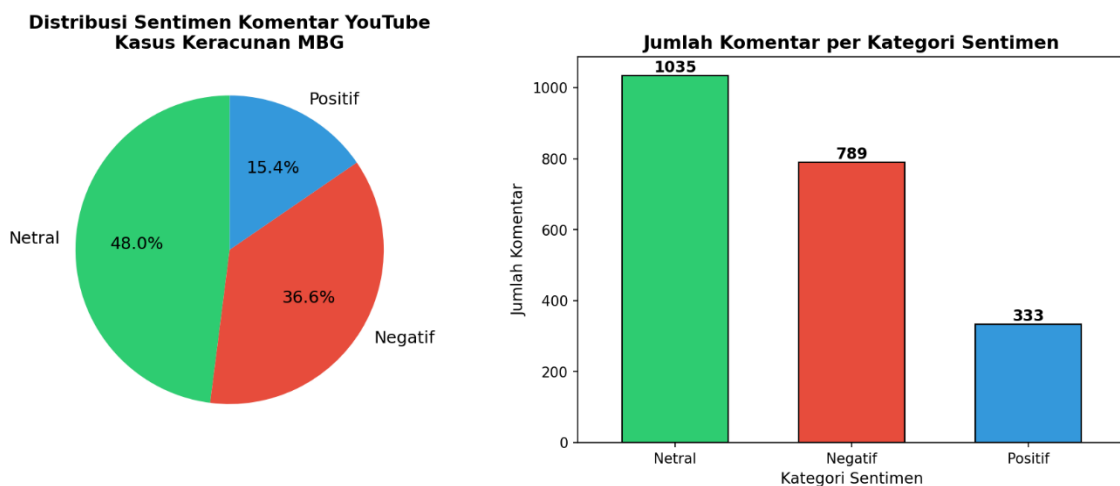
Sebagai gambaran awal, lima baris pertama data mentah ditampilkan pada Tabel 7 berikut. Berdasarkan Tabel 6, data mentah yang berhasil dikumpulkan terdiri atas 2.157 baris dengan 5 kolom, yaitu komentar, username, likes, tanggal, dan video_id. Kolom komentar menjadi fokus utama analisis, sedangkan kolom lainnya digunakan sebagai metadata pendukung.

Tabel 7 Contoh Data (5 baris pertama)

No	komentar	username	likes	tanggal	Video_id
0	Jadii gimana menu MBG kalian? Atau malah belu...	@KokBisa	395	2025-10-16	k7lZDu6_GnU
1	1 korban sudahh terlalu banyakk 🤔\nMiriss mel...	@ferryakob489	275	2025-10-16	k7lZDu6_GnU
2	"Korban MBG NUNGGU ADA YG MATI DULU, baru diev...	@farifairis	37	2025-10-18	k7lZDu6_GnU
3	Video edukasi tema keracunan kesukaan anak say...	@kakashibaiq123	56	2025-10-24	k7lZDu6_GnU
4	Fun Fact:\n\nTulisan thumbnails vidio ini dar...	@BlackyBlox_30	461	2025-10-17	k7lZDu6_GnU

Tabel 7 memperlihatkan bahwa data komentar yang dikumpulkan masih dalam bentuk mentah, mengandung berbagai elemen seperti emoji, tanda baca berulang, serta gaya penulisan informal khas media sosial, yang selanjutnya perlu dibersihkan melalui tahap preprocessing sebelum dapat digunakan dalam proses pemodelan.

Pelabelan data dilakukan menggunakan pendekatan berbasis leksikon (lexicon-based) yang dijalankan melalui Google Colaboratory, dengan menghitung skor kemunculan kata pada kamus kata positif dan kamus kata negatif yang disusun sesuai konteks isu keracunan MBG, kemudian mengklasifikasikannya ke dalam tiga kategori: Positif (skor kata positif lebih besar), Negatif (skor kata negatif lebih besar), dan Netral (skor seimbang atau tidak ditemukan kata kunci sentimen). Proses pelabelan ini dilakukan terhadap 2.157 komentar mentah yang berhasil dikumpulkan, sebelum data melalui tahap preprocessing. Dari total tersebut, diperoleh distribusi sebagai berikut: sentimen Netral sebanyak 1.035 komentar (48,0%), Negatif sebanyak 789 komentar (36,6%), dan Positif sebanyak 333 komentar (15,4%), sebagaimana ditampilkan pada Gambar 3 berikut.



Gambar 3 Distribusi sentiment dan jumlah per kategori Sentimen Komentar Youtube Kasus Keracunan MBG

Perlu dicatat bahwa setelah data melalui tahap preprocessing yang akan dijelaskan pada subbagian berikutnya, sejumlah komentar tidak valid (spam, tautan, atau emoji tanpa teks) turut dieksklusi dari dataset, sehingga distribusi label pada 2.039 komentar bersih sedikit berbeda dari distribusi awal: sentimen Netral menjadi 941 komentar (46,1%), Negatif 771 komentar (37,8%), dan Positif 327 komentar (16,0%). Distribusi pada data bersih inilah yang selanjutnya digunakan sebagai acuan dalam pembagian data latih dan data uji. Ketidakseimbangan distribusi kelas ini merupakan kondisi yang wajar dan mencerminkan pola opini publik yang sesungguhnya terjadi di media sosial.

3.2 Preprocessing Data

Tahap preprocessing bertujuan membersihkan dan menyiapkan teks komentar agar dapat diproses secara optimal oleh algoritma klasifikasi [18]. Enam tahapan preprocessing diterapkan secara berurutan menggunakan library Python NLTK, Sastrawi, dan PySastrawi, yaitu: (1) case folding, (2) cleansing, (3) normalisasi kata tidak baku, (4) tokenisasi, (5) stopword removal, dan (6) stemming. Sebagai contoh penerapan preprocessing, komentar “Korban MBG NUNGGU ADA YG MATI DULU, baru dievaluasi..” diubah menjadi “korban makan gizi gratis nunggu mati evaluasi” setelah melalui seluruh tahapan. Proses normalisasi kata berhasil menstandarkan kata-kata tidak baku khas media sosial seperti “gk” menjadi “tidak”, “bgt” menjadi “sangat”, dan sebagainya, sehingga representasi teks menjadi lebih seragam dan informatif bagi model. Sebagaimana ditampilkan pada tabel 8 berikut.

Lima contoh hasil transformasi teks melalui seluruh tahapan preprocessing ditunjukkan pada Tabel 8 berikut.

Tabel 8 contoh hasil processing

No	Sebelum	Sesudah
1	Jadiiii gimana menu MBG kalian? Atau malah belum dapet? Cer....	jadiiii menu makan gizi gratis dapet ce....
2	1 korban sudahh terlalu banyakk 🤔....	Miriss melihatt program korban sudahh terlalu banyakk miriss....
3	"Korban MBG NUNGGU ADA YG MATI DULU, baru dievaluasi.."	ki korban makan gizi gratis nunggu mati e....
4	Video edukasi tema keracunan kesukaan anak saya yg baru us....	video edukasi tema racun suka anak usi....
5	Fun Fact: Tulisan thumbnails vidio ini dari kata MBG ata....	fun fact tulis thumbnails vidio makan

Berdasarkan kelima contoh pada Tabel 8, terlihat bahwa proses preprocessing secara konsisten berhasil mengubah teks komentar yang awalnya tidak terstruktur dan mengandung

elemen non-formal menjadi representasi teks yang lebih ringkas, baku, dan siap diproses oleh tahap ekstraksi fitur.

3.3 Distribusi Sentimen

Sebagaimana telah disinggung pada bagian 3.1, distribusi sentimen pada data final pasca-preprocessing (2.039 komentar) menunjukkan bahwa sentimen Netral mendominasi dengan 941 komentar (46,1%), diikuti sentimen Negatif sebanyak 771 komentar (37,8%), dan sentimen Positif sebanyak 327 komentar (16,0%). Dominasi sentimen Netral mengindikasikan bahwa sebagian besar masyarakat merespons kasus keracunan MBG dengan cara informatif dan deskriptif, yakni dengan menyebarkan informasi, bertanya, atau menceritakan pengalaman tanpa memberikan penilaian emosional yang tegas. Sentimen Negatif yang cukup besar (37,8%) mencerminkan adanya ketidakpuasan dan kekhawatiran signifikan dari sebagian masyarakat terhadap kualitas pengelolaan program MBG, terutama terkait standar keamanan pangan dan respons pemerintah dalam menangani insiden keracunan. Sementara itu, sentimen Positif yang relatif kecil (16,0%) mengindikasikan bahwa meskipun terdapat kasus keracunan, masih ada segmen masyarakat yang tetap mendukung dan optimis terhadap keberlangsungan program MBG.

Untuk memperdalam pemahaman terhadap topik yang dibicarakan pada masing-masing kategori sentimen, dilakukan visualisasi word cloud sebagaimana ditampilkan pada Gambar 4.



Gambar 4 World Cloud Analisis Sentimen Komentar Youtube Kasus Keracunan Program Makan Bergizi Gratis (MBG)

Gambar 4 menampilkan word cloud untuk masing-masing kategori sentimen. Pada sentimen Negatif (771 komentar), kata-kata yang paling dominan adalah “makan”, “gratis”, “racun”, “anak”, “korupsi”, dan “basi”, yang menggambarkan kecaman publik terhadap kualitas makanan dan dugaan korupsi dalam program. Pada sentimen Netral (941 komentar), kata-kata utama meliputi “makan”, “gratis”, “sekolah”, “indonesia”, “program”, dan “masak”, mencerminkan diskusi yang bersifat faktual dan deskriptif. Sedangkan pada sentimen Positif (327 komentar), kata-kata yang menonjol antara lain “makan”, “gizi”, “sehat”, “anak”, “sekolah”, dan “prabowo”, yang menunjukkan dukungan terhadap tujuan program untuk meningkatkan gizi anak sekolah. Kata “makan” dan “gratis” yang muncul di semua kategori merupakan kata kunci topik utama, sementara kata “racun”, “korupsi”, dan “basi” menjadi pembeda kuat yang mendefinisikan sentimen Negatif.

3.4 Evaluasi dan Perbandingan Model

Kedua model klasifikasi, SVM dan Naïve Bayes, dilatih menggunakan 80% data (data latih) dan diuji pada 20% data sisanya (data uji). Hasil evaluasi kuantitatif menggunakan empat metrik utama disajikan pada Tabel 9 berikut.

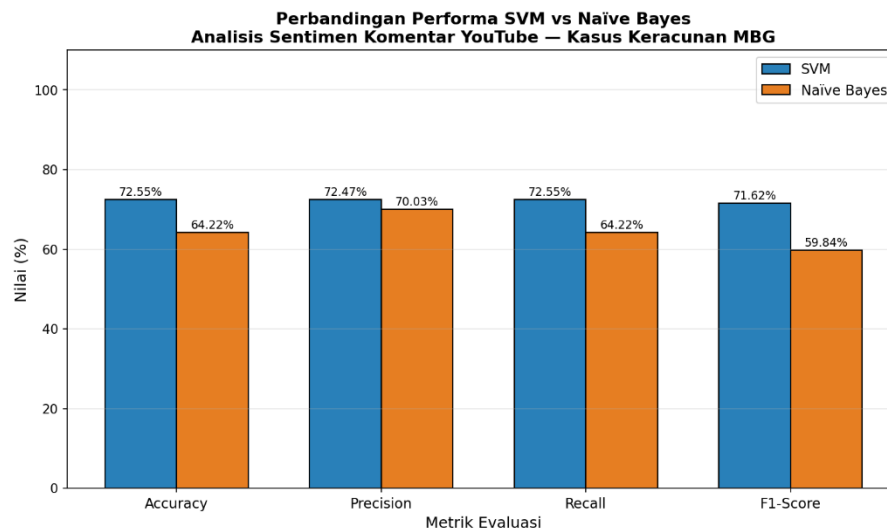
Tabel 9 Perbandingan Metrik Evaluasi SVM dan Naïve Bayes

Metrik	SVM	Naïve Bayes
Accuracy	72,55	64,22
Precision	72,47	70,03
Recall	72,55	64,22
F1-Score	71,62	59,84

Berdasarkan Tabel 9, model SVM secara konsisten mengungguli Naïve Bayes pada seluruh metrik evaluasi, dengan selisih akurasi mencapai 8,33 poin persentase. Selisih performa

yang cukup signifikan ini, khususnya pada metrik Recall dan F1-Score, mengindikasikan bahwa SVM lebih mampu menangani kompleksitas dan ketidakseimbangan data pada kasus klasifikasi sentimen tiga kelas ini dibandingkan Naïve Bayes.

SVM secara konsisten mengungguli Naïve Bayes pada seluruh metrik evaluasi. SVM memperoleh akurasi 72,55%, presisi 72,47%, recall 72,55%, dan F1-Score 71,62%. Sementara itu, Naïve Bayes memperoleh akurasi 64,22%, presisi 70,03%, recall 64,22%, dan F1-Score 59,84%. Selisih akurasi antara kedua model mencapai 8,33 poin persentase, dengan SVM berada di posisi yang lebih unggul. Visualisasi perbandingan keempat metrik evaluasi secara berdampingan disajikan pada Gambar 5.

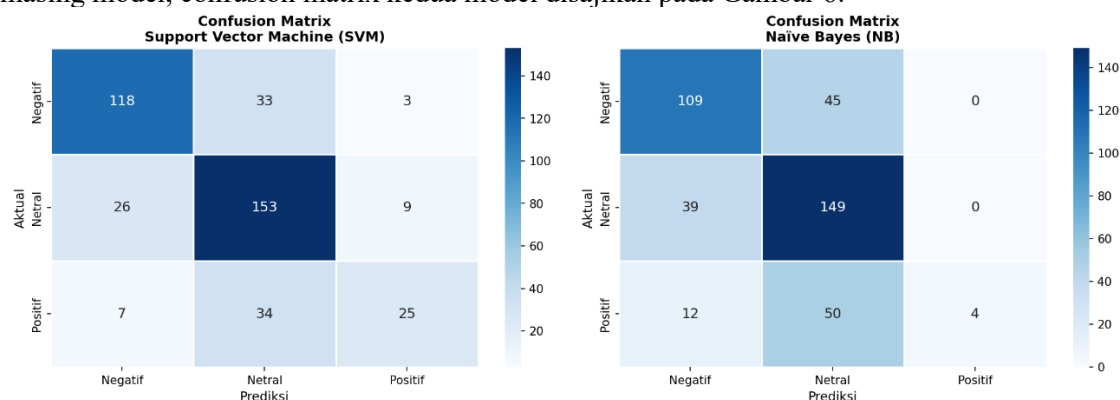


Gambar 5 Perbandingan Performa SVM vs Naïve Bayes

Gambar 5 mempertegas keunggulan SVM di semua metrik evaluasi. Selain metrik keseluruhan, perlu diperhatikan bahwa nilai Precision Naïve Bayes (70,03%) hanya sedikit lebih rendah dari SVM (72,47%), namun nilai Recall dan F1-Score Naïve Bayes jauh lebih rendah, masing-masing 64,22% dan 59,84%. Perbedaan besar antara Precision dan Recall pada Naïve Bayes mengindikasikan bahwa model ini cenderung melakukan misklasifikasi pada sejumlah kelas, terutama kelas dengan jumlah sampel yang lebih sedikit (kelas Positif).

3.5 Analisis Confusion Matrix

Untuk memperoleh gambaran yang lebih mendalam mengenai pola prediksi masing-masing model, confusion matrix kedua model disajikan pada Gambar 6.



Gambar 6 Confusion Matrix SVM dan Naïve Bayes

Berdasarkan confusion matrix pada Gambar 6, model SVM mampu mengklasifikasikan kelas Negatif dengan benar sebanyak 118 data, kelas Netral sebanyak 153 data, dan kelas Positif sebanyak 25 data. Kesalahan klasifikasi yang paling banyak terjadi pada kelas Positif, di mana 34 data Positif salah diklasifikasikan sebagai Netral dan 7 data diklasifikasikan sebagai Negatif. Hal

ini dapat dipahami karena komentar bernada positif dalam konteks kasus keracunan MBG cenderung menggunakan kosakata yang netral dan tumpang tindih dengan kelas Netral.

Pada model Naïve Bayes, kelas Negatif berhasil diklasifikasikan dengan benar sebanyak 109 data dan kelas Netral sebanyak 149 data. Namun, model Naïve Bayes sama sekali gagal mengklasifikasikan kelas Negatif dan Netral ke kelas Positif (prediksi Positif masing-masing 0), dan hanya mampu memprediksi 4 data Positif dengan benar. Hal ini mengindikasikan bahwa model Naïve Bayes sangat bias terhadap kelas mayoritas (Netral dan Negatif) dan kesulitan mengenali fitur-fitur pembeda kelas Positif, yang merupakan kelas minoritas dalam dataset. Kelemahan ini sejalan dengan karakteristik Naïve Bayes yang sensitif terhadap distribusi kelas yang tidak seimbang [19].

Keunggulan SVM dalam penelitian ini sejalan dengan penelitian Sarimole (2024) yang menemukan SVM lebih unggul dari Naïve Bayes dalam analisis sentimen ulasan aplikasi Satu Sehat di Twitter dengan selisih akurasi 6,30% [6], serta penelitian Jakak dan Rahman (2025) yang membuktikan efektivitas SVM untuk analisis sentimen MBG di Instagram dengan akurasi 86,7% [2]. Pada penelitian ini, algoritma Support Vector Machine (SVM) digunakan untuk menentukan hyperplane optimal yang berfungsi sebagai batas pemisah antar kelas sentimen. Hyperplane tersebut dipilih dengan memaksimalkan margin antar kelas pada ruang fitur berdimensi tinggi yang dihasilkan dari representasi teks menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF), sehingga model lebih tangguh terhadap noise dan lebih mampu menangani ambiguitas semantik pada komentar media sosial berbahasa Indonesia yang tidak formal. Selain itu, kernel Linear yang digunakan pada SVM terbukti cocok untuk data teks TF-IDF yang pada umumnya sudah bersifat linearly separable di ruang berdimensi tinggi [20].

4. KESIMPULAN

Penelitian ini berhasil melakukan analisis sentimen terhadap 2.039 komentar netizen YouTube terkait kasus keracunan Program Makan Bergizi Gratis (MBG) menggunakan algoritma Support Vector Machine (SVM) dan Naïve Bayes (NB) dengan ekstraksi fitur TF-IDF dan pembagian data 80% training dan 20% testing. Hasil distribusi sentimen menunjukkan bahwa sentimen Netral mendominasi dengan 941 komentar (46,2%), diikuti sentimen Negatif sebanyak 771 komentar (37,8%), dan sentimen Positif sebanyak 327 komentar (16,0%), yang mengindikasikan bahwa sebagian besar masyarakat menyikapi kasus keracunan MBG secara informatif dan objektif, sementara sepertiga lainnya mengekspresikan ketidakpuasan dan kekhawatiran terhadap pengelolaan program tersebut. Hasil perbandingan kedua algoritma menunjukkan bahwa Support Vector Machine (SVM) secara konsisten memperoleh nilai yang lebih tinggi pada seluruh metrik evaluasi, yakni akurasi 72,55%, presisi 72,47%, recall 72,55%, dan F1-Score 71,62%, dibandingkan Naïve Bayes yang memperoleh akurasi 64,22%, presisi 70,03%, recall 64,22%, dan F1-Score 59,84%. Dengan selisih akurasi sebesar 8,33 poin persentase, SVM terbukti lebih optimal untuk mengklasifikasikan sentimen komentar YouTube berbahasa Indonesia pada konteks isu kebijakan publik yang sensitif. Untuk penelitian selanjutnya, disarankan untuk memperluas jumlah dataset, menerapkan teknik penyeimbangan data seperti SMOTE guna mengatasi ketidakseimbangan kelas sentimen, serta mengeksplorasi model berbasis transformer seperti IndoBERT untuk meningkatkan akurasi klasifikasi sentimen pada teks bahasa Indonesia yang lebih kompleks dan tidak formal.

5. SARAN

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat dijadikan acuan untuk penelitian lanjut. Pertama, penelitian selanjutnya disarankan untuk menerapkan teknik penyeimbangan data seperti SMOTE (Synthetic Minority Over-sampling Technique) atau undersampling guna mengatasi ketidakseimbangan kelas sentimen yang terjadi, khususnya pada kelas Positif yang jumlahnya jauh lebih sedikit dibandingkan kelas Netral dan

Negatif, sehingga performa klasifikasi pada kelas minoritas dapat ditingkatkan secara signifikan. Kedua, penelitian selanjutnya dapat mengeksplorasi penggunaan model berbasis transformer seperti IndoBERT atau RoBERTa-base yang dirancang khusus untuk bahasa Indonesia, mengingat model-model tersebut mampu memahami konteks semantik yang lebih dalam pada teks informal khas media sosial dibandingkan pendekatan TF-IDF berbasis kata. Ketiga, perluasan sumber data dengan menambahkan komentar dari platform media sosial lain seperti X (Twitter) atau TikTok dapat memberikan gambaran yang lebih komprehensif dan representatif terhadap opini publik mengenai program MBG. Keempat, pelabelan data pada penelitian ini menggunakan pendekatan lexicon-based yang memiliki keterbatasan dalam menangkap konteks, sarkasme, atau ambiguitas makna yang kompleks pada komentar berbahasa Indonesia yang bersifat informal. Penelitian selanjutnya disarankan untuk menggunakan pelabelan manual oleh multiple annotator dengan pengujian inter-annotator agreement, atau memanfaatkan model bahasa besar (LLM) untuk pelabelan yang lebih kontekstual, guna meningkatkan validitas label sentimen yang dihasilkan.

DAFTAR PUSTAKA

- [1] A. A. Merlinda and Y. Yusuf, "Analisis Program Makan Gratis Prabowo Subianto Terhadap Strategi Peningkatan Motivasi Belajar Siswa di Sekolah Tinjauan dari Perspektif Sosiologi Pendidikan," *Ranah Res. J. Multidiscip. Res. Dev.*, vol. 7, no. 2, pp. 1364–1373, 2025, doi: <https://doi.org/10.38035/rj.v7i2>.
- [2] Wardianto, P. M. Jakak, and M. Rahman, "Analisis Sentimen Publik Terhadap Program Makan Bergizi Gratis di Instagram Menggunakan Algoritma Support Vector Machine," *SMARTICS J.*, vol. 11, no. 1, pp. 14–20, 2025, doi: <https://doi.org/10.21067/smartics.v11i1.11852>.
- [3] A. Leader *et al.*, "Kasus Keracunan Makanan Program Makan Bergizi Gratis (MBG) pada Siswa : Tinjauan Nilai Kemanusiaan Sila Kedua Pancasila dalam Perspektif Keamanan Pangan dan Tanggung Jawab Negara Faculty of Languages and Arts (FBS), Surabaya State University , Indones," *J. Ilm. Nusant. (JINU)*, vol. 3, no. 2, pp. 922–935, 2026, doi: <https://doi.org/10.61722/jinu.v3i2.9336>.
- [4] L. Rhomeaningtias, A. Khairunisa, S. Shella, M. Wara, and K. M. Hindrayani, "Analisis Sentimen Ulasan Aplikasi Smile Indonesia Menggunakan Metode Naive Bayes Dan Support Vector Machine (SVM)," *HOAQ J. Teknol. Inf.*, vol. 16, no. 1, pp. 79–91, 2025, doi: <https://doi.org/10.52972/hoaq.vol16no1>.
- [5] A. M. N. Sari, M. Jazman, T. Ahsyar, Syaifullah, "Penerapan Algoritma Klasifikasi Naive Bayes dan Support Vector Machine untuk Analisis Sentimen Cyberbullying Bilingual di Aplikasi X," *Sist. J. Sist. INFORMSI*, vol. 14, no. 1, 2025.
- [6] F. M. Sarimole and Kudrat, "Analisis Sentimen Terhadap Aplikasi Satu Sehat Pada Twitter Menggunakan Algoritma Naive Bayes Dan Support Vector Machine," *J. Sains dan Teknologi*, vol. 5, no. 3, pp. 783–790, 2024, doi: <https://doi.org/10.55338/saintek.v5i3.2702>.
- [7] N. Habibah, F. Sains, D. A. N. Teknologi, U. Islam, N. Sultan, and S. Kasim, "Analisis Sentimen Mengenai Penggunaan E-Wallet Pada Google Play Menggunakan Lexicon Based Dan K-Nearest Neighbor," *JURIKOM (Jurnal Ris. Komputer)*, vol. 10, no. 1, pp. 192–200, 2023, doi: [10.30865/jurikom.v10i1.5429](https://doi.org/10.30865/jurikom.v10i1.5429).
- [8] I. Najiyah and M. Rizal, "Analisis Sentimen Media Sosial X Terhadap Kebijakan Presiden Republik Indonesia Prabowo Subianto," *J. FASILKOM*, vol. 15, no. 3, pp. 443–455, 2025, doi: <https://doi.org/10.37859/jf.v15i3.10385>.
- [9] D. H. Saifulloh, B. Perdian, and A. I. Panca, "Analisis Sentimen Ulasan Produk E-Commerce Berbasis Teks Dengan Pendekatan Probabilistik Naive Bayes," *J. SIKOM (Sistem Inf. Komputer)*, vol. 2, no. 1, pp. 61–74, 2025.
- [10] F. Abdusyukur, "View Of Penerapan Algoritma Support Vector Machine (Svm) Untuk Klasifikasi Pencemaran Nama Baik Di Media Sosial Twitter," *Komputa J. Ilm. Komput. dan Inform.*, vol. 12, no. 1, pp. 2715–7849, 2023, doi: <https://doi.org/10.34010/komputa.v12i1.9418>.
- [11] H. S. Wafa *et al.*, "Prediksi Penyakit Diabetes Menggunakan Algoritma Support Vector Machine (SVM)," *INFORMATICS Digit. Expert*, vol. 4, no. 1, pp. 40–45, 2022.
- [12] A. D. Pratama, "Analisa sentimen masyarakat terhadap penggunaan chatgpt menggunakan metode support vector machine (svm)," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 9, no. 1,

-
- pp. 327–338, 2024.
- [13] H. Barus, I. N. Fajri, and Y. Pristyanto, “Sentiment Classification Analysis of Tokopedia Reviews Using TF-IDF , SMOTE , and Traditional Machine Learning Models,” *J. Appl. Informatics Comput.*, vol. 9, no. 5, pp. 2552–2561, 2025.
- [14] O. I. Gifari, M. Adha, I. R. Hendrawan, F. Freddy, S. Durrand, and A. S. Literature, “Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine,” *JIFOTECH (JOURNAL Inf. Technol.*, vol. 2, no. 1, pp. 36–40, 2022.
- [15] D. Atmajaya, A. Febrianti, and H. Darwis, “Metode SVM dan Naive Bayes untuk Analisis Sentimen ChatGPT di Twitter Dedy,” *Indones. J. Comput. Sci.*, vol. 12, no. 1, pp. 2173–2181, 2023.
- [16] A. Azmi, Febri and A. Voutama, “View of Prediksi Churn Nasabah Bank Menggunakan Klasifikasi Random Forest Dan Decision Tree Dengan Evaluasi Confusion Matrix,” *KOMPUTA J. Ilm. Komput. dan Inform.*, vol. 13, no. 1, pp. 2089–9033, 2024.
- [17] B. Ramadhani and R. R. Suryono, “Komparasi Algoritma Naïve Bayes dan Logistic Regression Untuk Analisis Sentimen Metaverse,” *J. MEDIA Inform. BUDIDARMA*, vol. 8, no. April, pp. 714–725, 2024, doi: 10.30865/mib.v8i2.7458.
- [18] D. Safryda and T. Ridwan, “Analisis Sentimen Ulasan Aplikasi Pospay dengan Algoritma Support Vector Machine,” *J. Ilm. Inform.*, vol. 11, no. 1, 2023.
- [19] I. M. Arya, A. Dwija, I. M. Gede, and I. G. Aris, “JTIM : Jurnal Teknologi Informasi dan Multimedia Perbandingan Algoritma Naive Bayes Berbasis Feature Selection Gain Ratio dengan Naive Bayes Kovensional dalam Prediksi Komplikasi Hipertensi,” vol. 6, no. 1, pp. 37–49, 2024.
- [20] Sukirman, Sajidah, N. P. Husain, A. F. Syam, and R. Mustikosari, “Analisis Sentimen Ulasan Pengguna Tiktok pada Google Play Store Berbasis TF-IDF dan Support Vector Machine,” *J. Syst. Comput. Eng.*, vol. 5, no. 1, pp. 91–102, 2024.
-