



Evaluasi algoritma KNN dan Naive Bayes untuk analisis sentimen kebijakan program makan bergizi gratis

Nursakina, Farid Wajidi, Muh. Rafli Rasyid

Program Studi Informatika, Universitas Sulawesi Barat, Indonesia

Riwayat Artikel:

Diterima 12 September 2025

Direvisi 8 Oktober 2025

Disetujui 13 Oktober 2025

Kata Kunci:

Analisis Sentimen

Media Sosial

Makan Bergizi Gratis

Naive Bayes

K-Nearest Neighbor

ABSTRACT. Social media has become a primary platform for the public to express opinions on government policies, including Indonesia's Free Nutritious Meal (MBG) program. This study analyzes public sentiment toward MBG by comparing the K-Nearest Neighbor (KNN) and Naive Bayes algorithms. A total of 9,723 tweets were collected, and after removing the neutral category, 6,322 tweets remained (3,955 positive and 2,367 negative), indicating a dominance of positive opinions. The experimental results show that KNN performed best at $k=9$ with a 70:30 split, achieving an accuracy of 72.27% and an F1-score of 65.92%, with an average cross-validation accuracy of 73.30%. Naive Bayes with parameter $\alpha=0.5$ consistently outperformed KNN, achieving an average accuracy of 79.61% and an F1-score of 77.33%, along with better precision-recall balance. The main contribution of this research is providing empirical evidence that Naive Bayes is more effective than KNN for sentiment analysis of Indonesian-language text with a large dataset, as well as offering a methodological framework applicable to the evaluation of other public policies.

ABSTRAK. Media sosial menjadi ruang utama masyarakat untuk mengekspresikan opini terhadap kebijakan publik, termasuk Program Makan Bergizi Gratis (MBG) di Indonesia. Penelitian ini menganalisis sentimen publik terhadap MBG dengan membandingkan algoritma K-Nearest Neighbor (KNN) dan Naive Bayes. Data diperoleh dari 9.723 tweet, kemudian setelah penghapusan kategori netral tersisa 6.322 tweet (3.955 positif dan 2.367 negatif), yang menunjukkan dominasi opini positif. Hasil eksperimen menunjukkan bahwa KNN terbaik pada $k=9$ dengan rasio 70:30 menghasilkan akurasi 72,27% dan F1-score 65,92%, dengan akurasi rata-rata *cross-validation* 73,30%. Naive Bayes dengan parameter $\alpha=0,5$ unggul dengan akurasi rata-rata 79,61% dan F1-score 77,33%, serta keseimbangan presisi dan *recall* yang lebih baik. Kontribusi penelitian ini adalah memberikan bukti empiris bahwa Naive Bayes lebih efektif dibandingkan KNN dalam analisis sentimen teks berbahasa Indonesia dengan dataset besar, serta menawarkan kerangka metodologis yang dapat diterapkan untuk evaluasi kebijakan publik lainnya.

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



Penulis Korespondensi:

Nursakina

Fakultas Teknik Universitas Sulawesi Barat

Jl. Prof. Dr. Baharuddin Lopa, Majene

Email: nsakina212@gmail.com

PENDAHULUAN

Media sosial telah berkembang menjadi salah satu ruang utama bagi masyarakat Indonesia dalam menyampaikan opini terhadap isu sosial maupun kebijakan pemerintah. Dengan jumlah pengguna aktif yang melebihi 167 juta orang dan rata-rata durasi penggunaan harian sekita 7 jam, *platform* seperti X (sebelumnya dikenal sebagai *Twitter*) tidak hanya berfungsi sebagai sarana komunikasi, tetapi tetapi juga sebagai ruang ekspresi dan pertukaran opini masyarakat yang mencerminkan

respons publik terhadap kebijakan (Rizqika dkk., 2025). Kondisi ini menjadikan media sosial sebagai sumber data yang potensial untuk menangkap respons masyarakat secara luas dan aktual.

Salah satu kebijakan strategis yang saat ini tengah dijalankan oleh pemerintah adalah Program Makan Bergizi Gratis (MBG), yang bertujuan untuk meningkatkan asupan gizi peserta didik serta mendukung kualitas pendidikan nasional. Program ini menyasar balita, siswa PAUD hingga SMA, serta ibu hamil. Namun, alokasi anggaran yang besar mencapai 71 triliun rupiah menimbulkan beragam respons publik (Trisno Aji, 2025). Sebagian masyarakat menilai program ini sebagai langkah positif dalam pembangunan sumber daya manusia yang unggul, sementara sebagian lainnya menyuarakan kekhawatiran terhadap beban fiskal dan efektivitas implementasinya. Perbedaan persepsi ini menunjukkan pentingnya dilakukan analisis sentimen berbasis media sosial dengan pendekatan *machine learning* untuk memperoleh gambaran objektif mengenai penerimaan masyarakat terhadap kebijakan MBG.

Sejumlah studi terdahulu telah menerapkan analisis sentimen pada berbagai topik. Syifa dkk. (2023) membandingkan algoritma *K-Nearest Neighbor* (KNN) dan *Naive Bayes* dalam analisis sentimen terkait Covid-19, dengan akurasi tertinggi sebesar 72,37%. Sitanggang dkk. (2024) menganalisis sentimen terhadap Program Makan Siang Gratis menggunakan algoritma *Naive Bayes* pada 2.211 *tweet*, menghasilkan akurasi sebesar 72,2%. Sementara itu, Iman & Ujianto (2024) menerapkan KNN untuk isu pemindahan ibu kota dengan akurasi mencapai 77%. Penelitian serupa dilakukan oleh Abdillah dkk. (2024) yang membandingkan algoritma *Naive Bayes* dan *K-Nearest Neighbor* pada analisis sentimen pengguna aplikasi Zenius. Hasil penelitian mereka menunjukkan bahwa performa kedua algoritma dipengaruhi oleh distribusi data yang tidak seimbang, di mana *Naive Bayes* cenderung lebih stabil dalam menghadapi ketimpangan kelas dibandingkan KNN. Meskipun relevan, studi-studi tersebut masih terbatas pada ukuran dataset yang kecil, hanya mengevaluasi satu algoritma, sehingga belum memberikan gambaran komparatif yang representatif terhadap opini publik.

Berdasarkan keterbatasan penelitian terdahulu, studi ini berkontribusi dengan memanfaatkan dataset yang lebih besar, yaitu 9.723 *tweet* mengenai Program MBG, membandingkan dua algoritma klasifikasi populer dalam analisis teks, yakni *K-Nearest Neighbor* (KNN) dan *Naive Bayes*, serta menyajikan analisis sentimen publik sebagai bahan evaluasi bagi pemerintah dalam merumuskan strategi implementasi kebijakan gizi nasional. Dengan demikian, penelitian ini tidak hanya memperkaya kajian akademik mengenai analisis sentimen berbasis media sosial, tetapi juga menghadirkan bukti empiris baru terkait persepsi publik terhadap kebijakan MBG di Indonesia.

METODE

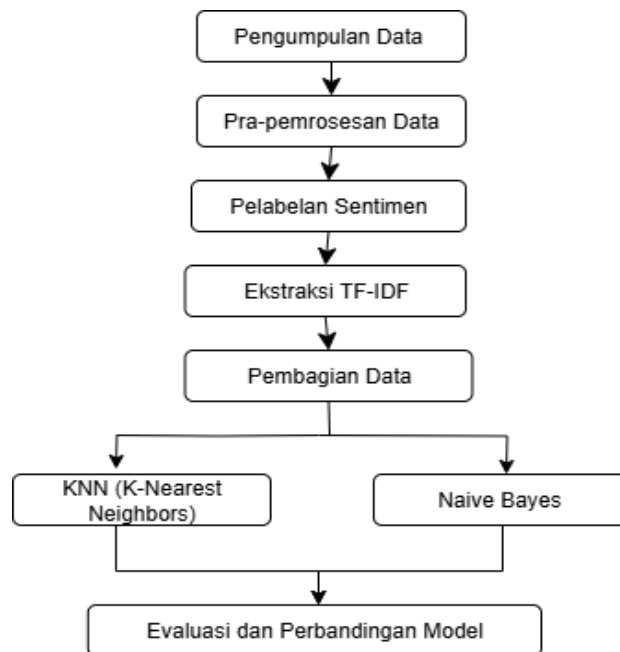
Jenis penelitian

Penelitian ini merupakan studi kuantitatif dengan pendekatan komputasional. Jenis penelitian yang digunakan adalah eksperimen, karena melibatkan implementasi dan pengujian dua algoritma klasifikasi berbasis *supervised learning*, yaitu *K-Nearest Neighbor* (KNN) dan *Naive Bayes*.

Tahapan penelitian

Penelitian ini dilaksanakan melalui beberapa tahapan, dimulai dengan pengumpulan data dari media sosial *Platform X* menggunakan kata kunci terkait Program Makan Bergizi Gratis. Data yang diperoleh kemudian melalui pra-pemrosesan teks untuk memastikan kebersihan dan konsistensi, dilanjutkan dengan pelabelan sentimen ke dalam tiga kategori: positif, negatif, dan netral. Tahap berikutnya adalah ekstraksi fitur menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF) untuk mengubah teks menjadi representasi numerik, kemudian data dibagi menjadi data latih dan data uji. Selanjutnya dilakukan klasifikasi menggunakan algoritma *K-Nearest Neighbor* (KNN)

dan *Naive Bayes*, dan tahap akhir adalah evaluasi serta perbandingan kinerja model untuk menentukan algoritma dengan hasil terbaik (Gambar 1).



Gambar 1. Alur penelitian

Pengumpulan Data

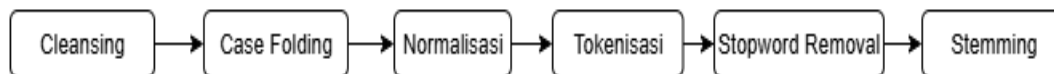
Pengumpulan data dilakukan menggunakan *tweet-harvest* versi 2.6.1 yang dijalankan melalui perintah *npx* pada lingkungan *Node.js* versi 20.18.3. Tool ini memerlukan autentikasi token untuk memastikan akses yang lebih stabil ke data *Twitter*. Dalam tahap awal penelusuran, kata kunci “makan siang gratis” dipilih karena muncul sebagai trending topik dengan tingkat keterlibatan dan frekuensi kemunculan tertinggi dibandingkan istilah lain seperti “makan bergizi gratis”. Oleh sebab itu, istilah tersebut dianggap paling representatif untuk menjangkau opini publik terkait Program Makan Bergizi Gratis (MBG). Pengambilan data difokuskan pada *tweet* berbahasa Indonesia selama periode 6 Januari hingga 1 Juli 2025, dengan total 9.723 *tweet* mentah. Seluruh data disimpan dalam format *Comma Separated Values (CSV)* untuk kemudian diproses pada tahap pra-pemrosesan sebelum analisis lebih lanjut. Meski demikian, penggunaan satu trending topik sebagai kata kunci utama diakui sebagai keterbatasan karena berpotensi belum mencakup seluruh variasi istilah publik terkait program MBG.

Pra-pemrosesan Data

Tahapan *pra-pemrosesan* dilakukan untuk memastikan data teks yang diperoleh lebih bersih, konsisten dan siap digunakan dalam analisis sentimen. Proses ini meliputi beberapa langkah utama (Gambar 2), yaitu :

1. *Cleansing*: Menghapus elemen-elemen yang tidak relevan seperti tautan, mention pengguna, tanda baca, angka, simbol, atau karakter khusus.
2. *Case folding*: Mengubah seluruh huruf menjadi bentuk konsisten, yaitu huruf kecil. (*lowercase*).
3. *Normalisasi*: Mengonversi kata tidak baku, singkatan, atau bahasa gaul menjadi bentuk formal. Contoh: “gak” menjadi “tidak”. Tujuannya agar sesuai dengan kamus referensi.

4. *Tokenisasi*: memecah kalimat menjadi unit-unit (*token*).
5. *Stopword removal*: Menghilangkan kata-kata yang sering muncul namun tidak berkontribusi terhadap identifikasi makna, seperti “dan”, “atau”, “di”
6. *Stemming*: Mengubah kata berimbuhan menjadi bentuk dasar agar proses analisis lebih konsisten. Misalnya, “berlari” menjadi “lari”.



Gambar 2. Tahapan alur pra-pemrosesan

Pelabelan Sentimen

Tahap pelabelan dilakukan untuk mengklasifikasikan setiap *tweet* ke dalam tiga kategori, yaitu positif, negatif, dan netral. Proses ini menggunakan model *Bidirectional Encoder Representations from Transformers (BERT)* berbahasa Indonesia dengan nama *taufiqdp/indonesian-sentiment*, yang telah *di-fine-tune* khusus untuk klasifikasi sentimen pada teks berbahasa Indonesia. Model ini dipilih karena telah dilatih menggunakan data berbahasa Indonesia dengan domain opini umum, sehingga relevan untuk analisis sentimen publik. Untuk memastikan kesesuaian dan keandalan model dalam konteks kebijakan publik, dilakukan evaluasi manual terhadap sampel hasil pelabelan guna menilai konsistensi klasifikasi polaritas. Model ini memanfaatkan pembelajaran dua arah (*bidirectional*) sehingga mampu memahami konteks kata secara akurat, dengan keluaran berupa label sentimen dan tingkat kepercayaan (*confidence score*) yang digunakan dalam analisis (Berliana & Yusuf, 2025). Meskipun demikian, pelabelan otomatis berbasis BERT memiliki keterbatasan, terutama pada dataset yang tidak seimbang sehingga kelas minoritas seperti netral cenderung kurang akurat (Habbat dkk., 2023)(Bau et al., 2023). Oleh karena itu, hasil pelabelan dievaluasi kembali secara manual untuk memastikan fokus analisis hanya pada opini publik yang jelas bersifat positif atau negatif.

Ekstraksi Fitur

Term Frequency-Inverse Document Frequency (TF-IDF) digunakan untuk memberi bobot pada kata berdasarkan frekuensinya dalam suatu dokumen dibandingkan dengan seluruh korpus. Semakin sering kata muncul pada dokumen tertentu, bobot TF meningkat, sedangkan semakin umum kata muncul di banyak dokumen, nilai IDF menurun (Aziz & Fauziah, 2022). Perhitungan TF-IDF dilakukan melalui dua persamaan, yaitu IDF dan bobot kata $w(t,d)$ sebagai hasil perkalian TF dan IDF sebagaimana dijelaskan oleh (Khoerunnisa dkk., 2025).

$$IDF(w) = \log\left(\frac{N}{DF(W)}\right) \quad (1)$$

Keterangan :

$IDF(w)$: *Inverse Document Frequency* dari istilah w
 N : Jumlah total dokumen dalam korpus
 $DF(W)$: Jumlah dokumen yang mengandung istilah w

$$w(t, d) = t(f, d) \times idf(t) \quad (2)$$

Keterangan:

$w(t, d)$: Bobot istilah t pada dokumen
 $t(f, d)$: Jumlah kemunculan istilah t dalam dokumen d
 $idf(t)$: Nilai *Inverse Document Frequency* dari istilah t

Pembagian Data

Dataset dibagi menjadi data latih dan data uji dengan tiga skenario proporsi 90:10, 80:20, dan 70:30. Model KNN diuji dengan variasi jumlah tetangga ($k = 3, 5, 7, 9$), sementara *Naive Bayes* menggunakan variasi parameter *smoothing* ($\alpha = 0,1; 0,5; 1,0; 2,0$). Hasil terbaik dari tiap skenario kemudian divalidasi menggunakan metode *5-fold cross-validation*, yang membagi data menjadi lima lipatan untuk memperoleh evaluasi yang lebih objektif. Metode ini telah digunakan secara relevan dalam studi sebelumnya (Widodo dkk., 2022).

Klasifikasi algoritma

Pemilihan algoritma *K-Nearest Neighbor (KNN)* dan *Naive Bayes* didasarkan pada karakteristik yang saling melengkapi. KNN merupakan algoritma berbasis kedekatan spasial yang tidak mengasumsikan distribusi data, sehingga cocok untuk data teks yang bersifat tidak linier. Sementara itu, *Naive Bayes* menggunakan pendekatan probabilistik yang efisien dan telah terbukti efektif dalam klasifikasi teks Bahasa Indonesia. Kedua algoritma dipilih untuk memberikan perbandingan performa yang representatif terhadap data *tweet* yang telah dilabeli secara otomatis (Puspita & Widodo, 2021).

• KNN

Algoritma *K-Nearest Neighbor (KNN)* merupakan metode klasifikasi berbasis *supervised learning* yang melakukan prediksi terhadap data uji dengan mengidentifikasi sejumlah tetangga terdekat sebanyak k dari data latih, berdasarkan ukuran jarak tertentu (Taufiqurrahman dkk., 2023). Salah satu ukuran jarak yang sering digunakan adalah *Euclidean Distance*, yang dirumuskan sebagai berikut (Habib Kusuma & Cahyono, 2023) :

$$d = \sqrt{\sum_{i=1}^p (x_1 - x_2)^2} \quad (3)$$

Keterangan :

d	: Jarak <i>Euclidean</i> antara dua vektor
i	: Fitur data
x_1	: data latih
x_2	: data uji
p	: Jumlah filter

• Naive Bayes

Algoritma *Naive Bayes* adalah salah satu algoritma dalam *machine learning* yang digunakan pada tugas klasifikasi, termasuk untuk analisis sentimen. Algoritma ini berlandaskan *Teorema Bayes* dengan asumsi bahwa setiap fitur atau atribut bersifat independen (Sanjaya dkk., 2023). Dalam konteks analisis sentimen, *Naive Bayes* berfungsi untuk memprediksi kategori sentimen dari suatu teks berdasarkan frekuensi atau keberadaan kata-kata dalam teks. Secara matematis, *Teorema Bayes* dirumuskan sebagai berikut (Punusingon dkk., 2025) :

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \quad (4)$$

Keterangan :

$P(A B)$: Probabilitas terjadinya A jika B terjadi
$P(B A)$: Probabilitas terjadinya B jika A terjadi
$P(A)$: Probabilitas A terjadi
$P(B)$: Probabilitas B terjadi

Evaluasi dan Perbandingan Model

Evaluasi performa model dilakukan untuk menilai efektivitas algoritma *K-Nearest Neighbor (KNN)* dan *Naive Bayes* dalam mengklasifikasi sentimen publik terhadap kebijakan Program Makan Bergizi Gratis. Kinerja model diukur menggunakan metrik akurasi, presisi, *recall*, dan *F1-score*, sedangkan *confusion matrix* digunakan untuk melihat distribusi prediksi pada tiap kelas sentimen. Untuk memperoleh hasil yang lebih stabil, pengujian dilakukan dengan *5-fold cross-validation* sehingga setiap subset data berkesempatan menjadi data latih maupun uji. Seluruh metrik evaluasi dihitung berdasarkan rumus evaluasi klasifikasi standar (Saragih & Kurniawan, 2025) :

Akurasi: mengukur seberapa sering model membuat prediksi yang benar.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (5)$$

Presisi: mengukur seberapa baik model menangkap semua kasus positif

$$precision = \frac{TP}{TP + FP} \times 100 \quad (6)$$

Recall: mengukur proporsi prediksi positif yang benar

$$recall = \frac{TP}{TP + FN} \times 100 \quad (7)$$

F1-Score: digunakan untuk menilai keseimbangan antara Precision dan Recall

$$F1 = \frac{2 \times (precision \times recall)}{(precision + recall)} \times 100 \quad (8)$$

HASIL DAN DISKUSI

Pengumpulan Data

Proses pengumpulan data dilakukan menggunakan *tweet-harvest* versi 2.6.1, dengan kata kunci “makan siang gratis” sebagai acuan pencarian. Selama periode 6 Januari hingga 1 Juli 2025, berhasil dikumpulkan sebanyak 9.723 *tweet* mentah. Data tersebut disimpan dalam format *Comma-Separated Values (CSV)* dan selanjutnya digunakan sebagai dasar untuk tahap pra-pemrosesan.

Pra-pemrosesan Data

Dataset awal yang dikumpulkan berjumlah 9.723 *tweet*. Untuk memastikan kualitas dan konsistensi data, dilakukan tahap pra-pemrosesan yang mencakup *cleansing*, *case folding*, *normalisasi*, *tokenisasi*, *stopword removal*, dan *stemming*. Selain itu, dilakukan penghapusan data duplikat serta eliminasi kata-kata dengan panjang tiga karakter atau kurang, yang dinilai tidak memberikan kontribusi signifikan. Setelah seluruh proses pra-pemrosesan diterapkan, jumlah teks yang tersisa menjadi 8.471. Hasil pra-pemrosesan ditampilkan secara rinci pada Tabel 1, yang memperlihatkan contoh transformasi teks sebelum dan sesudah setiap tahap.

Tabel 1. Pra-pemrosesan data

Tahapan	Sebelum	Sesudah
<i>Cleansing</i>	@BandaniDani @ilhampid Makan siang gratis dan sekolah gratis keduanya sama,² penting dan dibutuhkan dgn adanya program MBG harapannya anak,² bisa mendapatkan asupan gizi yg seimbang sehingga dpt meningkatkan	Makan siang gratis dan sekolah gratis keduanya sama penting dan dibutuhkan dgn adanya program MBG harapannya anak bisa mendapatkan asupan gizi seimbang sehingga dpt meningkatkan konsentrasi dan kecerdasan anak Lanjutkan MBG

Tahapan	Sebelum	Sesudah
<i>Case folding</i>	konsentrasi dan kecerdasan anak. Lanjutkan MBG Makan siang gratis dan sekolah gratis keduanya sama penting dan dibutuhkan dgn adanya program MBG harapannya anak bisa mendapatkan asupan gizi seimbang sehingga dpt meningkatkan konsentrasi dan kecerdasan anak Lanjutkan MBG	makan siang gratis dan sekolah gratis keduanya sama penting dan dibutuhkan dgn adanya program mbg harapannya anak bisa mendapatkan asupan gizi seimbang sehingga dpt meningkatkan konsentrasi dan kecerdasan anak lanjutkan mbg
<i>Normalisasi</i>	makan siang gratis dan sekolah gratis keduanya sama penting dan dibutuhkan dgn adanya program mbg harapannya anak bisa mendapatkan asupan gizi seimbang sehingga dpt meningkatkan konsentrasi dan kecerdasan anak lanjutkan mbg	makan siang gratis dan sekolah gratis keduanya sama penting dan dibutuhkan dengan adanya program makan bergizi gratis harapannya anak bisa mendapatkan asupan gizi seimbang sehingga dapat meningkatkan konsentrasi dan kecerdasan anak lanjutkan makan bergizi gratis
<i>Tokenisasi</i>	makan siang gratis dan sekolah gratis keduanya sama penting dan dibutuhkan dengan adanya program makan bergizi gratis harapannya anak bisa mendapatkan asupan gizi seimbang sehingga dapat meningkatkan konsentrasi dan kecerdasan anak lanjutkan makan bergizi gratis	['makan', 'siang', 'gratis', 'dan', 'sekolah', 'gratis', 'keduanya', 'sama', 'penting', 'dan', 'dibutuhkan', 'dengan', 'adanya', 'program', 'makan', 'bergizi', 'gratis', 'harapannya', 'anak', 'bisa', 'mendapatkan', 'asupan', 'gizi', 'seimbang', 'sehingga', 'dapat', 'meningkatkan', 'konsentrasi', 'dan', 'kecerdasan', 'anak', 'lanjutkan', 'makan', 'bergizi', 'gratis']
<i>Stopword</i>	['makan', 'siang', 'gratis', 'dan', 'sekolah', 'gratis', 'keduanya', 'sama', 'penting', 'dan', 'dibutuhkan', 'dengan', 'adanya', 'program', 'makan', 'bergizi', 'gratis', 'harapannya', 'anak', 'bisa', 'mendapatkan', 'asupan', 'gizi', 'seimbang', 'sehingga', 'dapat', 'meningkatkan', 'konsentrasi', 'dan', 'kecerdasan', 'anak', 'lanjutkan', 'makan', 'bergizi', 'gratis']	['makan', 'siang', 'gratis', 'sekolah', 'gratis', 'dibutuhkan', 'program', 'makan', 'bergizi', 'gratis', 'harapannya', 'anak', 'asupan', 'gizi', 'seimbang', 'meningkatkan', 'konsentrasi', 'kecerdasan', 'anak', 'lanjutkan', 'makan', 'bergizi', 'gratis']
<i>Stemming</i>	['makan', 'siang', 'gratis', 'sekolah', 'gratis', 'dibutuhkan', 'program', 'makan', 'bergizi', 'gratis', 'harapannya', 'anak', 'asupan', 'gizi', 'seimbang', 'meningkatkan', 'konsentrasi', 'kecerdasan', 'anak', 'lanjutkan', 'makan', 'bergizi', 'gratis']	makan siang gratis sekolah gratis butuh program makan gizi gratis harap anak asupan gizi imbang tingkat konsentrasi cerdas anak lanjut makan gizi gratis

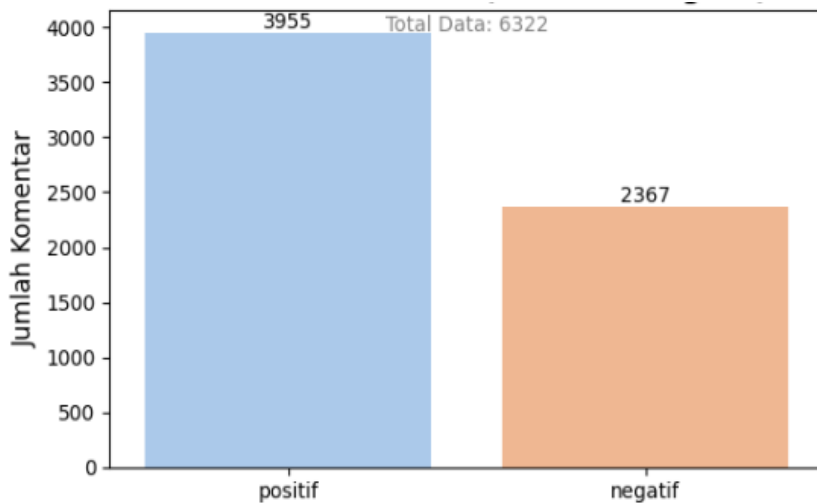
Pelabelan Data

Setelah tahap pra-pemrosesan, sebanyak 8.471 *tweet* digunakan untuk pelabelan otomatis menggunakan model *taufiqdp/indonesian-sentiment*, yang mengklasifikasikan setiap *tweet* ke dalam tiga kategori sentimen: positif, negatif, dan netral. Contoh hasil pelabelan ditunjukkan pada Tabel 2.

Tabel 2. Hasil pelabelan sentimen

<i>Stemming</i>	Bert sentimen	<i>Confidence</i>
senang makan siang gratis	Positif	0,997
mending bubar makan siang gratis	Negatif	0,961
potret anak sekolah dasar program makan siang gratis	Netral	0,860

Hasil pelabelan awal memperlihatkan bahwa terdapat 3.955 *tweet* dengan sentimen positif, 2.367 *tweet* dengan sentimen negatif, dan 2.149 *tweet* dengan sentimen netral. Kategori netral kemudian dihapus dari hasil akhir. Keputusan ini didasarkan pada evaluasi manual terhadap sampel data netral, yang menunjukkan bahwa sebagian besar teks dalam kategori tersebut tidak mengandung opini atau sikap yang dapat dianalisis secara emosional (Khadapi & Pakpahan, 2024). Sebagian besar hanya bersifat informatif atau deskriptif, seperti pada contoh teks “potret anak sekolah dasar program makan siang gratis” yang diklasifikasikan sebagai netral dengan *confidence* sebesar 0,860. Meskipun *confidence* tinggi, isi teks tidak menunjukkan dukungan atau penolakan terhadap isu yang diteliti.



Gambar 3. Distribusi kelas sentimen

Setelah kategori netral dihapus, sebanyak 6.322 *tweet* digunakan dalam analisis lanjutan, terdiri dari 3.955 *tweet* berlabel positif dan 2.367 berlabel negatif. Distribusi kelas sentimen ini divisualisasikan pada Gambar 3. Rata-rata skor *confidence* di atas 0,90 menunjukkan bahwa tingkat keyakinan model yang tinggi dalam klasifikasi polaritas teks.



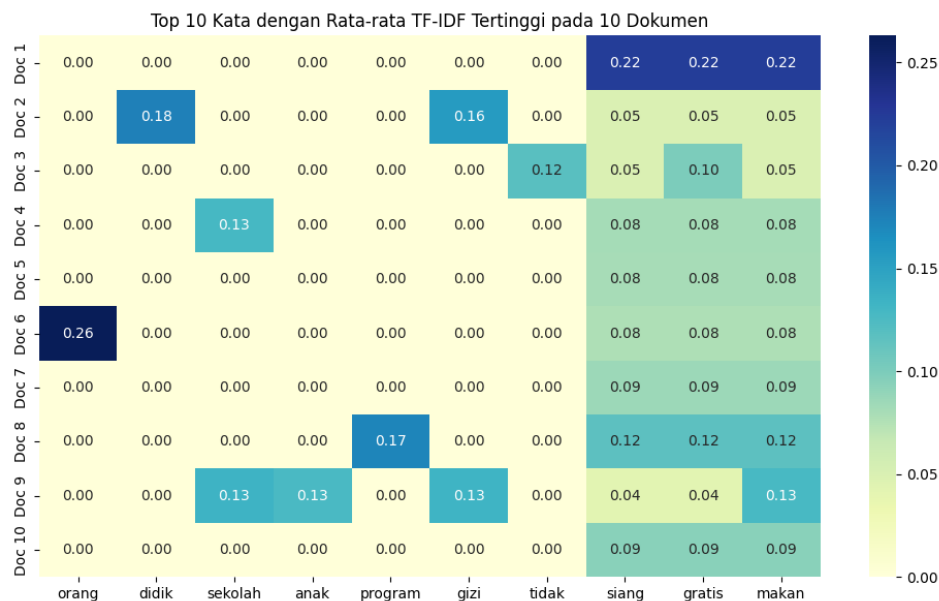
Gambar 4. Word cloud sentimen negatif dan positif

Gambar 4 menunjukkan visualisasi *word cloud* yang memuat kata-kata dominan dari *tweet* berlabel positif dan negatif. Kata-kata tersebut berasal dari hasil pelabelan sentimen dan mencerminkan ekspresi masyarakat terhadap isu yang dibahas. Pada bagian negatif, kata-kata seperti “bodoh”, “miskin”, dan “bohong” mencerminkan kritik, ketidakpuasan, atau penolakan terhadap pelaksanaan program. Sebagai contoh, salah satu komentar publik menyatakan, “program makan siang gratis bodoh, pilih apa tempuh cari dana halal haram korupsi laksana lapang ini.” Kutipan ini menunjukkan bahwa kemunculan kata-kata negatif bukan sekadar ekspresi emosional, melainkan bentuk kritik

terhadap sumber pendanaan, integritas pelaksanaan, dan transparansi kebijakan. Sementara itu, bagian positif menampilkan kata-kata seperti “gizi”, “bagus”, dan “bijak”, yang menunjukkan dukungan, apresiasi, serta harapan publik terhadap manfaat yang diharapkan dari program Makan Bergizi Gratis.

Pembobotan TF-IDF

Pada tahap ini digunakan *Term Frequency–Inverse Document Frequency (TF-IDF)* untuk mengubah teks menjadi data numerik dengan menekankan kata penting dan mengurangi dominasi kata umum.



Gambar 5. Top 10 kata TF-IDF

Gambar 5 memperlihatkan distribusi bobot dari sepuluh kata dominan pada sepuluh dokumen hasil ekstraksi fitur menggunakan *Term Frequency–Inverse Document Frequency (TF-IDF)*. Kata “makan”, “siang”, dan “gratis” secara konsisten memiliki bobot tinggi, yang menegaskan relevansinya dengan isu Program Makan Bergizi Gratis. Kata-kata lain seperti “gizi”, “anak”, dan “program” juga muncul menonjol pada beberapa dokumen, menunjukkan keterkaitan tematik yang kuat. Sebaliknya, kata “orang” hanya dominan pada Dokumen 6 dengan bobot sebesar 0,26, yang menunjukkan bahwa kontribusi kata tersebut bersifat kontekstual dan tidak merata di seluruh korpus. Visualisasi ini menggambarkan variasi kontribusi kata kunci dalam dokumen serta menegaskan pentingnya kata-kata tersebut sebagai fitur dalam proses klasifikasi sentimen.

Klasifikasi Algoritma

KNN

Pengujian algoritma *K-Nearest Neighbor (KNN)* dilakukan dengan variasi jumlah tetangga ($k = 3, 5, 7, 9$) dan proporsi data latih-uji (90:10, 80:20, 70:30) untuk mengevaluasi pengaruh parameter tersebut terhadap akurasi klasifikasi sentimen.

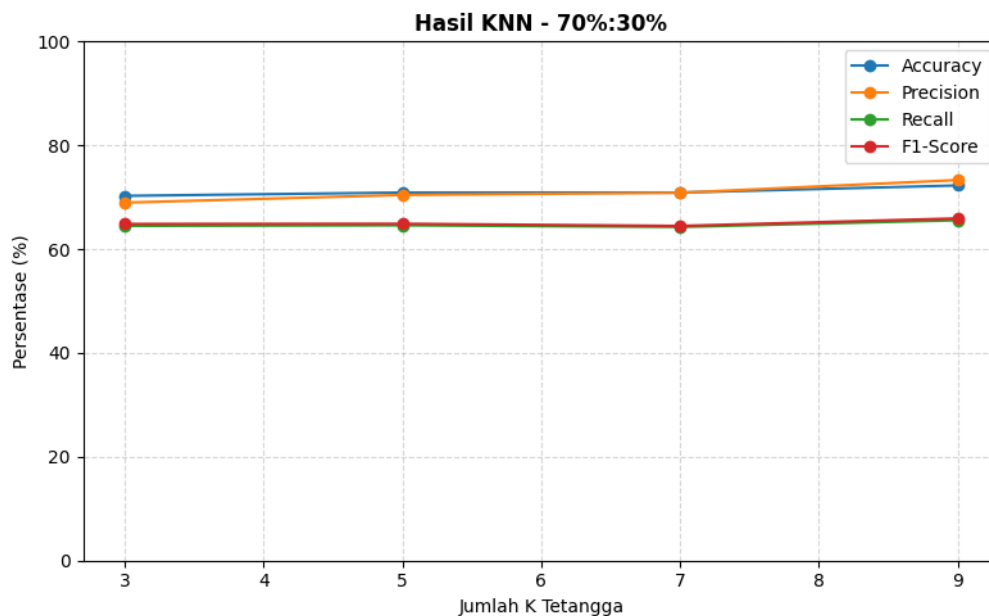
Hasil pengujian KNN ditunjukkan pada Tabel 3. Performa model cenderung meningkat seiring dengan bertambahnya nilai k . Kondisi ini dapat dijelaskan karena jumlah tetangga yang lebih besar membuat keputusan klasifikasi lebih stabil dan tidak mudah dipengaruhi oleh data menyimpang (*outlier*). Konfigurasi optimal diperoleh pada rasio 70:30 dengan $k = 9$, menghasilkan akurasi 72,27%

dan *F1-score* 65,92%. Tren performa KNN pada rasio data ini dapat dilihat pada Gambar 6. Rasio ini menunjukkan keseimbangan antara ketersediaan data latih untuk membangun model dan data uji untuk evaluasi yang representatif.

Tabel 3. Hasil pengujian KNN berdasarkan variasi nilai K dan proporsi data

Latih% : Uji%	K Tetangga	Akurasi%	Presisi%	Recall%	<i>F1-Score</i> %
90:10	3	66,67	63,95	60,57	60,51
	5	68,56	66,89	62,17	62,18
	7	69,98	69,23	63,47	63,61
	9	70,93	71,03	64,23	64,41
80:20	3	69,09	67,27	63,36	63,63
	5	70,43	69,35	64,52	64,88
	7	70,36	69,74	63,95	64,16
	9	71,62	72,11	65,01	65,30
70:30	3	70,27	68,96	64,47	64,85
	5	70,90	70,44	64,58	64,89
	7	70,90	70,86	64,27	64,47
	9	72,27	73,31	65,56	65,92

Namun, meskipun terjadi peningkatan, performa KNN tetap terbatas. Akurasi hanya berkisar pada 70%–72%, menandakan bahwa metode ini tidak cukup mampu menangkap keragaman kosakata serta kompleksitas semantik dalam opini publik. Selain itu, kebutuhan perhitungan jarak terhadap seluruh data latih menjadikan KNN kurang efisien pada dataset berskala besar. Sensitivitas terhadap distribusi data dan nilai k juga membuat prediksi tidak konsisten pada skenario yang berbeda. Dengan demikian, meskipun KNN mampu memberikan hasil yang cukup baik, algoritma ini tidak menunjukkan keunggulan yang signifikan pada kasus analisis sentimen dengan data besar.



Gambar 6. Performa KNN pada rasio data 70:30

Naive Bayes

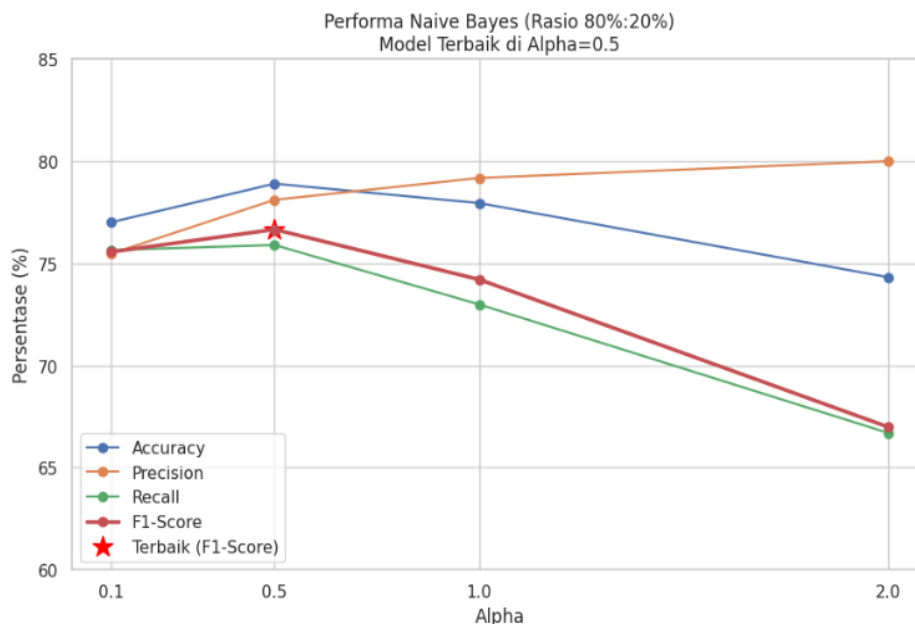
Pengujian *Naive Bayes* dilakukan dengan variasi proporsi data latih-uji (90:10, 80:20, 70:30) dan nilai *smoothing* α (0,1, 0,5, 1,0, 2,0) untuk menilai stabilitas dan kinerja klasifikasi sentimen terhadap distribusi data dan tingkat regularisasi.

Tabel 4. Hasil pengujian *Naive Bayes* berdasarkan proporsi data latih dan uji

	Latih% : Uji %	Alpha terbaik	Akurasi%	Presisi%	Recall%	F1-Score%
0	90:10	0,1	76,30	74,74	74,20	74,44
1		0,5	78,36	77,77	74,91	75,79
2		1,0	76,62	77,46	71,49	72,58
3		2,0	74,41	78,65	67,18	67,67
4	80:20	0,1	77,00	75,46	75,64	75,55
5		0,5	78,89	78,10	75,89	76,65
6		1,0	77,94	79,17	72,98	74,20
7	70:30	2,0	74,31	79,99	66,69	66,98
8		0,1	76,70	75,13	75,27	75,20
9		0,5	78,44	78,05	74,79	75,75
10		1,0	77,12	79,36	71,39	72,57
11		2,0	73,75	80,91	65,61	65,51

Hasil pengujian *Naive Bayes* pada Tabel 4 menunjukkan bahwa performa terbaik diperoleh pada rasio data latih-uji 80:20 dengan parameter *smoothing* $\alpha = 0,5$. Konfigurasi ini menghasilkan akurasi sebesar 78,89%, presisi 78,10%, *recall* 75,89%, dan *F1-score* 76,65%. Visualisasi tren performa *Naive Bayes* terhadap variasi Alpha ditunjukkan pada Gambar 7.

Keseimbangan antara presisi dan *recall* mengindikasikan bahwa model cukup andal dalam mengidentifikasi sentimen positif maupun negatif secara proporsional. Kinerja tersebut dipengaruhi oleh peran α sebagai faktor *smoothing*. Nilai α yang terlalu kecil cenderung mengabaikan kata-kata jarang muncul, sehingga *recall* menurun. Sebaliknya, nilai α yang terlalu besar menyebabkan distribusi probabilitas menjadi terlalu merata, yang berdampak pada penurunan presisi. Nilai $\alpha = 0,5$ terbukti optimal karena mampu menjaga keseimbangan antara generalisasi dan sensitivitas model terhadap variasi data.

Gambar 7. Tren performa *Naive Bayes* terhadap variasi nilai Alpha (80:20)

Hasil ini menunjukkan bahwa kesederhanaan asumsi independensi antar fitur pada *Naive Bayes* menjadi keunggulan saat diterapkan pada opini publik berskala besar. Namun, asumsi ini juga dapat menjadi keterbatasan ketika dihadapkan pada teks dengan struktur bahasa yang lebih kompleks atau konteks yang lebih spesifik.

Evaluasi dan pebandingan model

Evaluasi performa dilakukan menggunakan skema *5-fold cross-validation* untuk kedua algoritma, yaitu *K-Nearest Neighbor (KNN)* dan *Naive Bayes*. Pendekatan ini digunakan untuk memastikan bahwa hasil klasifikasi tidak bergantung pada satu pembagian data tertentu, melainkan mencerminkan stabilitas model secara umum terhadap variasi data latih dan uji.

Rasio terbaik dari masing-masing algoritma dipertahankan dalam evaluasi *5-fold cross-validation* untuk mencerminkan konfigurasi optimal yang diperoleh selama eksperimen awal. Tujuannya agar setiap algoritma diuji dalam kondisi performa terbaiknya. Hasil *cross-validation* untuk KNN disajikan pada Tabel 5, sedangkan hasil *Naive Bayes* disajikan pada Tabel 6.

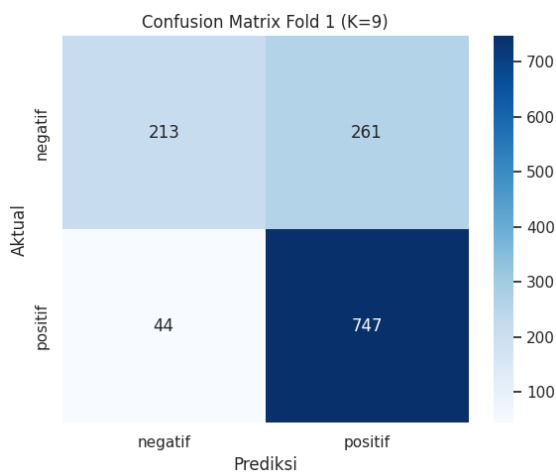
Tabel 5. Hasil evaluasi KNN berdasarkan *5-fold cross-validation* ($K = 9$)

Fold	Akurasi%	Presisi%	Recall%	F1-Score%
1	75,89	78,49	69,69	70,66
2	73,12	74,77	64,46	66,94
3	73,02	72,98	67,27	67,95
4	72,39	74,57	65,19	65,37
5	72,07	72,65	65,53	65,92
Rata -rata	73,30	74,69	66,83	67,37

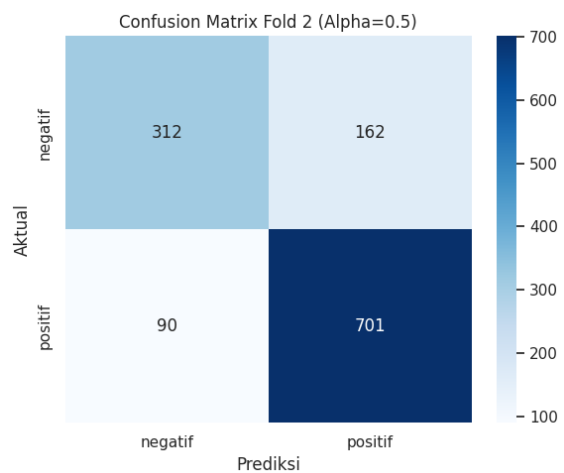
Tabel 6. Hasil Evaluasi *Naive Bayes* Berdasarkan *5-Fold Cross-Validation* ($\alpha = 0.5$)

Fold	Akurasi%	Presisi%	Recall%	F1-Score%
1	79,92	79,45	76,80	77,68
2	80,08	79,42	67,22	78,00
3	78,64	77,69	75,75	76,44
4	79,98	79,82	76,53	77,54
5	79,43	79,05	76,04	76,99
Rata -rata	79,61	79,09	76,47	77,33

Tabel 5 dan Tabel 6 menyajikan hasil evaluasi algoritma KNN ($k=9$) dan *Naive Bayes* ($\alpha=0,5$) dengan skema *5-fold cross-validation*. Secara keseluruhan, *Naive Bayes* menunjukkan performa lebih unggul dengan akurasi rata-rata 79,61% dan *F1-score* 77,33%, sedangkan KNN hanya mencapai akurasi 73,30% dan *F1-score* 67,37%. Hasil ini menunjukkan bahwa pendekatan probabilistik lebih efektif dalam menangani distribusi kata pada teks opini publik dibandingkan metode berbasis jarak yang sensitif terhadap variasi data latih.

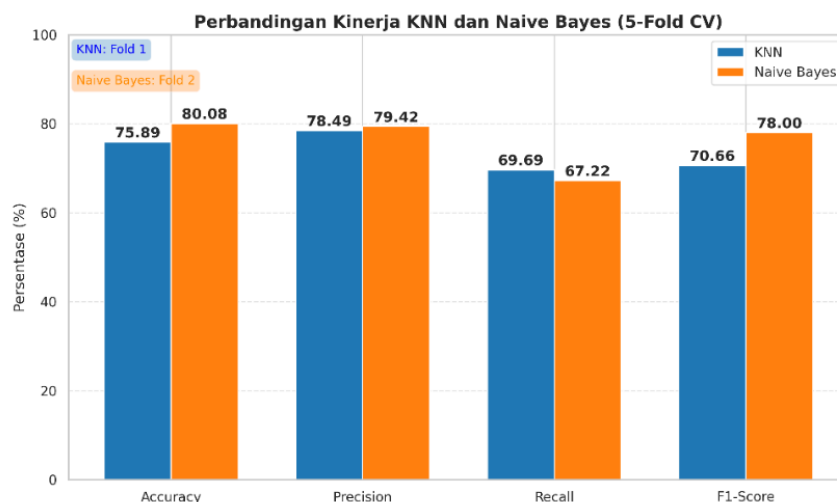


Gambar 8. Confusion matrix KNN



Gambar 9. Confusion matrix Naive Bayes

Perbedaan kinerja semakin jelas terlihat pada *confusion matrix* (Gambar 8 dan Gambar 9). Model KNN menghasilkan banyak kesalahan klasifikasi pada kelas negatif, dengan 261 data salah diprediksi sebagai positif, sehingga *recall* rata-rata hanya mencapai 66,83%. Hal ini menunjukkan bahwa KNN kurang mampu menangkap pola distribusi teks dengan representasi vektor berdimensi tinggi. Sebaliknya, *Naive Bayes* mampu menghasilkan distribusi prediksi yang lebih seimbang, dengan 312 data negatif dan 701 data positif terklasifikasi benar, yang berkontribusi pada nilai *F1-score* lebih tinggi (77,33%). Keunggulan ini konsisten dengan sifat *Naive Bayes* sebagai algoritma berbasis probabilitas yang efektif dalam menangani fitur teks berskala besar (Ma'aly dkk., 2024). Temuan ini memperkuat bahwa *Naive Bayes* lebih andal dibanding KNN dalam membedakan sentimen positif dan negatif pada teks media sosial, sehingga dapat memberikan gambaran yang lebih akurat terkait persepsi publik terhadap kebijakan pemerintah.



Gambar 10. Perbandingan Kinerja KNN dan Naive Bayes

Gambar 10 mempertegas keunggulan *Naive Bayes* pada semua metrik evaluasi, khususnya akurasi dan *F1-score*. Temuan ini konsisten dengan hasil penelitian Syifa dkk. (2023) yang juga melaporkan kinerja lebih baik dari *Naive Bayes* dibanding KNN dalam analisis sentimen Covid-19, serta lebih tinggi dari capaian Sitanggang dkk. (2024) yang hanya memperoleh akurasi 72,2% pada analisis sentimen Program Makan Siang Gratis. Dengan dataset yang lebih besar dan pengaturan parameter yang tepat, penelitian ini mampu meningkatkan akurasi hingga 79,61%. Dari sisi kebijakan, dominasi

sentimen positif menunjukkan bahwa Program Makan Bergizi Gratis memperoleh penerimaan yang baik dari masyarakat. Temuan ini dapat menjadi indikator bagi pemerintah untuk memperkuat strategi komunikasi publik, menjaga transparansi kebijakan, serta merespons kritik secara konstruktif untuk mempertahankan tingkat dukungan masyarakat terhadap program yang dijalankan.

KESIMPULAN

Penelitian ini membandingkan algoritma *K-Nearest Neighbor (KNN)* dan *Naive Bayes* untuk klasifikasi sentimen publik terhadap Program Makan Bergizi Gratis (MBG). Hasil evaluasi menunjukkan bahwa *Naive Bayes* memiliki kinerja lebih baik dengan akurasi rata-rata 79,61% dan *F1-score* 77,33%, lebih tinggi dibandingkan KNN yang hanya mencapai akurasi 73,30% dan *F1-score* 67,37%. Hal ini membuktikan bahwa pendekatan probabilistik lebih efektif dibandingkan metode berbasis jarak dalam analisis sentimen teks berbahasa Indonesia pada dataset berskala besar. Secara ilmiah, penelitian ini memperkaya kajian analisis sentimen berbasis *machine learning* dengan menyajikan bukti empiris bahwa *Naive Bayes* unggul dibanding KNN pada konteks kebijakan sosial. Secara praktis, dominasi sentimen positif menunjukkan bahwa Program MBG memperoleh penerimaan yang baik di masyarakat, sehingga pemerintah dapat menjadikannya sebagai indikator dukungan publik sekaligus memperkuat strategi komunikasi kebijakan.

Untuk penelitian selanjutnya, disarankan penggunaan algoritma lain seperti *Support Vector Machine (SVM)*, *Random Forest*, maupun pendekatan *deep learning* (misalnya *Long Short-Term Memory (LSTM)* atau *BERT fine-tuning*) agar diperoleh hasil yang lebih akurat. Selain itu, dataset yang lebih luas, termasuk data dari berbagai platform media sosial, dapat meningkatkan generalisasi model dan memberikan gambaran lebih komprehensif mengenai persepsi publik.

REFERENSI

- Abdillah, T., Khaira, U., & Hutabarat, B. F. (2024). Komparasi Metode Naive Bayes dan K-Nearest Neighbors Terhadap Analisis Sentimen Pengguna Aplikasi Zenius. *Processor: Jurnal Ilmiah Sistem Informasi, Teknologi Informasi dan Sistem Komputer*, 19(1). <https://doi.org/10.33998/processor.2024.19.1.1596>
- Aziz, A., & Fauziah. (2022). Analisis Sentimen Identifikasi Opini Terhadap Produk, Layanan dan Kebijakan Perusahaan Menggunakan Algoritma TF-IDF dan SentiStrength. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 6(1), 115–125. <https://tunasbangsa.ac.id/ejurnal/index.php/jsakti/article/view/430>
- Bau, R. T. R. L., Hermila, A., Farman, I., Hidayat, L. M., & Salim, S. (2023). AI Perspectives in Education: A BERT-based Exploration of Informatics Students' Attitudes to ChatGPT. *2023 9th International Conference on Education and Technology (ICET)*, 36–41. <https://doi.org/10.1109/ICET59790.2023.10435040>
- Berliana, H., & Yusuf, R. (2025). Analisis sentimen terhadap penggunaan donasi korban penyiraman air keras pada media sosial X.com menggunakan metode BERT. *Journal of Science and Social Research*, 8(2), 1134–1142. <https://www.jurnal.goretanpena.com/index.php/JSSR/article/view/3078>
- Habbat, N., Nouri, H., Anoun, H., & Hassouni, L. (2023). Sentiment analysis of imbalanced datasets using BERT and ensemble stacking for deep learning. *Engineering Applications of Artificial Intelligence*, 126, 106999. <https://doi.org/10.1016/j.engappai.2023.106999>
- Habib Kusuma, I., & Cahyono, N. (2023). Analisis Sentimen Masyarakat Terhadap Penggunaan E-Commerce Menggunakan Algoritma K-Nearest Neighbor. *Jurnal Informatika: Jurnal pengembangan IT (JPIT)*, 8(3), 302–307. <https://doi.org/10.30591/jpit.v8i3.5734>
- Iman, A. K., & Ujianto, E. I. H. (2024). Analisis Sentimen Pemindahan Ibu Kota Indonesia Menggunakan K-Nearest Neighbor. *Jurnal Pendidikan dan Teknologi Indonesia (JPTI)*, 4(12), 759–768. <https://doi.org/10.52436/1.jpti.546>
- Khadapi, M., & Pakpahan, V. M. (2024). Analisis Sentimen Berbasis Jaringan LSTM dan BERT terhadap Diskusi Twitter tentang Pemilu 2024. *JUKI: Jurnal Komputer dan Informatika*, 6(2), 130–137. <https://ioinformatic.org/index.php/JUKI/article/view/681>

- Khoerunnisa, S., Shiddieq, D. F., & Nurhayati, D. (2025). Penerapan Algoritma Naive Bayes dengan Teknik TF-IDF dan Cross Validation untuk Analisis Sentimen Terhadap Starlink. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 5(2), 566–577. <https://doi.org/10.57152/malcom.v5i2.1852>
- Ma'aly, A. N., Pramesti, D., Fathurahman, A. D., & Fakhurroja, H. (2024). Exploring Sentiment Analysis for the Indonesian Presidential Election Through Online Reviews Using Multi-Label Classification with a Deep Learning Algorithm. *Information (Switzerland)*, 15(11). <https://doi.org/10.3390/info15110705>
- Punusingon, P. V., Yusupa, A., & Tarigan, V. (2025). Analisis Sentimen Pemilihan Remaja Teladan Wilayah Rataan Menggunakan Algoritma Naive Bayes. *Journal Computer and Technology*, 3(1), 65–73. <https://ojs.ninetyjournal.com/index.php/COMTECHNO/article/view/337>
- Puspita, R., & Widodo, A. (2021). Perbandingan Metode KNN, Decision Tree, dan Naive Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS. *Jurnal Informatika Universitas Pamulang*, 5(4), 646. <https://doi.org/10.32493/informatika.v5i4.7622>
- Rizqika, V., Handoko, P., & Rochmania, A. (2025). Self disclosure generasi Z melalui media sosial X (Twitter). *Jurnal Ilmu Komunikasi UHO*, 10(2), 397–411. <https://jurnalilmukomunikasi.uho.ac.id/index.php/journal/article/view/348/205>
- Sanjaya, T. P. R., Fauzi, A., & Masruriyah, A. F. N. (2023). Analisis Sentimen Ulasan Pada E-Commerce Shopee Menggunakan Algoritma Naive Bayes dan Support Vector Machine. *INFOTECH: Jurnal Informatika & Teknologi*, 4(1), 16–26. <https://doi.org/10.37373/infotech.v4i1.422>
- Saragih, N. N., & Kurniawan, R. (2025). Analisis Sentimen Menggunakan Metode Naive Bayes Tentang Program Mudik Gratis Pemerintah Kota Medan 2024. *Journal of Computer Engineering, System and Science*, 10(1), 299–311. <https://jurnal.unimed.ac.id/2012/index.php/cess>
- Sitanggang, Umaidah, Y., & Adam, R. I. (2024). Analisis sentimen masyarakat terhadap Program Makan Siang Gratis pada media sosial X menggunakan Algoritma Naive Bayes. *Jurnal Informatika dan Teknik Elektro Terapan*, 12(3), 2755–2762. <https://doi.org/10.23960/jitet.v12i3.4902>
- Syifa, H. A. N., Nugroho, A., & Firliana, R. (2023). Perbandingan Algoritma Naive Bayes Classifier Dan K-Nearest Neighbors Untuk Analisis Sentimen Covid-19 Di Twitter. *Jurnal Ilmiah Informatika (JIF)*, 11(01), 54–62. <https://doi.org/10.33884/jif.v11i01.7069>
- Taufiqurrahman, H., Tri Anggraeny, F., & Al Haromainy, M. M. (2023). Perbandingan Algoritma Naive Bayes Dan K-Nearest Neighbor Pada Analisis Sentimen Ulasan Aplikasi Mypertamina. *Jurnal Mahasiswa Teknik Informatika*, 7(6), 3934–3939.
- Trisno Aji, W. (2025). Makan Bergizi Gratis di Era Prabowo-Gibran: Solusi untuk Rakyat atau Beban Baru? *NAAFI: Jurnal Ilmiah Mahasiswa*, 2(2). <https://jurnal.stkip-majenang.ac.id/index.php/naafi/article/view/134>
- Widodo, S., Brawijaya, H., & Samudi, S. (2022). Stratified K-fold cross validation optimization on machine learning for prediction. *Sinkron*, 7(4), 2407–2414. <https://doi.org/10.33395/sinkron.v7i4.11792>