

Evaluasi Komparatif Algoritma Machine Learning dalam Analisis Sentimen Program Makan Bergizi Gratis di Media Sosial X

Angelina Putri Ariani*, Norhikmah, Yoga Pristyanto

Fakultas Ilmu Komputer, Sistem Informasi, Universitas Amikom Yogyakarta, Yogyakarta, Indonesia
Email: ¹angelinaputri5@students.amikom.ac.id, ²hikmah@amikom.ac.id, ³yoga.pristyanto@amikom.ac.id
Email Penulis Korespondensi: angelinaputri5@students.amikom.ac.id
Submitted 28-04-2026; Accepted 09-06-2026; Published 30-06-2026

Abstrak

Pengkajian opini masyarakat di platform media sosial melalui analisis sentimen memiliki peranan yang krusial dalam memahami bagaimana publik merespons kebijakan yang ditetapkan oleh pemerintah, termasuk Program Makan Bergizi Gratis (MBG). Kendati demikian, ketidakseimbangan karakteristik data media sosial serta penggunaan bahasa informal seperti sarkasme menjadi hambatan pada tahap pengklasifikasian sentimen. Oleh karena itu, penelitian ini bertujuan untuk mengkaji pandangan masyarakat mengenai program MBG di platform X sekaligus mengevaluasi efektivitas sejumlah algoritma machine learning dalam mengategorikan sentimen. Dataset yang digunakan dalam penelitian ini mencakup 2.000 komentar yang dikumpulkan pada periode Februari hingga April 2026, dilabeli menggunakan pendekatan berbasis leksikon, serta melewati tahap preprocessing dan ekstraksi fitur melalui TF-IDF. Pengklasifikasian data dilakukan dengan menerapkan enam algoritma, yakni Naïve Bayes, K-Nearest Neighbor (K-NN), Support Vector Machine (SVM), Decision Tree, Random Forest, dan Logistic Regression. Berdasarkan hasil penelitian, dapat diketahui bahwa Random Forest merupakan model yang mendapatkan tingkat akurasi terbaik, yakni mencapai 92%, serta didukung oleh hasil cross validation sebesar 89% yang menunjukkan stabilitas model. Berdasarkan hasil klasifikasi, distribusi sentimen publik dominan sentimen netral sejumlah 66,3% disusul oleh sentimen negatif sejumlah 22,6% dan sentimen positif sejumlah 11,1%. Temuan tersebut mengungkap bahwa pandangan masyarakat terhadap program MBG cenderung bersifat netral dengan kecenderungan kritik yang lebih dominan, serta membuktikan bahwa ketepatan dalam pemilihan algoritma memegang peranan penting dalam meningkatkan akurasi analisis sentimen pada data teks yang rumit dan tidak seimbang.

Kata Kunci: Machine Learning; Analisis Sentimen; Random Forest; TF-IDF; Media Sosial X; Program Makan Bergizi Gratis

Abstract

The analysis of public opinion on social media platforms through sentiment analysis plays a crucial role in understanding how the public responds to government policies, including the Free Nutritious Meal Program (MBG). However, the imbalanced nature of social media data and the use of informal language such as sarcasm pose challenges in the sentiment classification process. Therefore, this study aims to examine public perceptions of the MBG program on platform X while also evaluating the effectiveness of several machine learning algorithms in categorizing sentiment. The dataset used in this study consists of 2,000 comments collected between February and April 2026. The data were labeled using a lexicon-based approach and processed through preprocessing and feature extraction using TF-IDF. The classification process was carried out using six algorithms: Naïve Bayes, K-Nearest Neighbor (K-NN), Support Vector Machine (SVM), Decision Tree, Random Forest, and Logistic Regression. The results show that Random Forest achieved the highest accuracy, reaching 92%, supported by a cross-validation score of 89%, indicating strong model stability. Based on the classification results, public sentiment is predominantly neutral at 66.3%, followed by negative sentiment at 22.6% and positive sentiment at 11.1%. These findings suggest that public opinion toward the MBG program tends to be neutral, with a stronger inclination toward criticism than support. Furthermore, the results highlight the importance of selecting appropriate algorithms to improve the accuracy of sentiment analysis on complex and imbalanced textual data.

Keywords: Machine Learning; Sentiment Analysis; Random Forest; TF-IDF; Social Media X; Free Nutritious Meal Program

1. PENDAHULUAN

Program Makan Bergizi Gratis (MBG) adalah program unggulan yang dilaksanakan oleh pemerintah Presiden Prabowo Subianto yang diluncurkan Januari 2025 untuk menangani stunting dan kurang gizi sebagai upaya mewujudkan visi Indonesia Emas 2024 [1]. Berdasarkan data 2023, angka stunting nasional masih berada di 21,5% sehingga program ini menargetkan 82,9 juta penerima manfaat hingga 2029, mencakup ibu hamil, balita, dan siswa di seluruh jenjang pendidikan [2][3]. Dengan anggaran Rp 71 triliun pada 2025, program ini diharapkan turut menggerakkan UMKM dan pemanfaatan produk pertanian lokal [4]. Sejalan dengan pelaksanaannya, program ini turut menjadi topik yang mendapatkan sorotan luas di ranah publik digital, termasuk di media sosial X.

Platform X kini menjelma sebagai salah satu wadah utama yang dimanfaatkan masyarakat untuk mengutarakan pendapat terhadap kebijakan publik secara cepat dan luas [5]. Interaksi pengguna di platform ini menghasilkan beragam sentimen, baik positif, negatif, maupun netral, yang dapat dimanfaatkan sebagai analisis untuk memahami opini masyarakat. Oleh karena itu, informasi yang diperoleh dari platform ini memiliki potensi besar untuk dapat dimanfaatkan sebagai rujukan dalam proses penentuan arah kebijakan.

Analisis sentimen termasuk dalam salah satu bagian dari cabang ilmu *Natural Language Processing*/NLP yang berfungsi untuk mengenali, mengekstraksi, serta mengklasifikasi emosi, sikap, dan opini subjektif dalam teks secara otomatis [5]. Pada konteks Program MBG, analisis ini memiliki peran yang signifikan dalam mendukung para peneliti untuk memahami respon masyarakat secara terstruktur [6]. Namun demikian, pemilihan algoritma *machine learning* yang tepat menjadi tantangan tersendiri, mengingat teks media sosial cenderung tidak terstruktur, informal, serta mengandung singkatan dan *slang* yang dapat mempengaruhi kinerja model.

Sejumlah algoritma *machine learning* telah banyak diterapkan pada studi analisis sentimen, beberapa di antaranya *Naïve Bayes* yang efisien dalam mengolah data berdimensi tinggi [5], SVM memiliki kemampuan dalam membentuk *hyperplane* yang optimal untuk memisahkan kelas secara akurat [7][8], K-NN yang bekerja berdasarkan kedekatan jarak antara data [9], *Decision Tree* dan *Random Forest* menawarkan tingkat interpretabilitas yang baik serta kemampuan dalam mengurangi *overfitting* [10][11], serta *Logistic Regression* memberikan interpretasi berbasis probabilistik dalam proses klasifikasi [12].

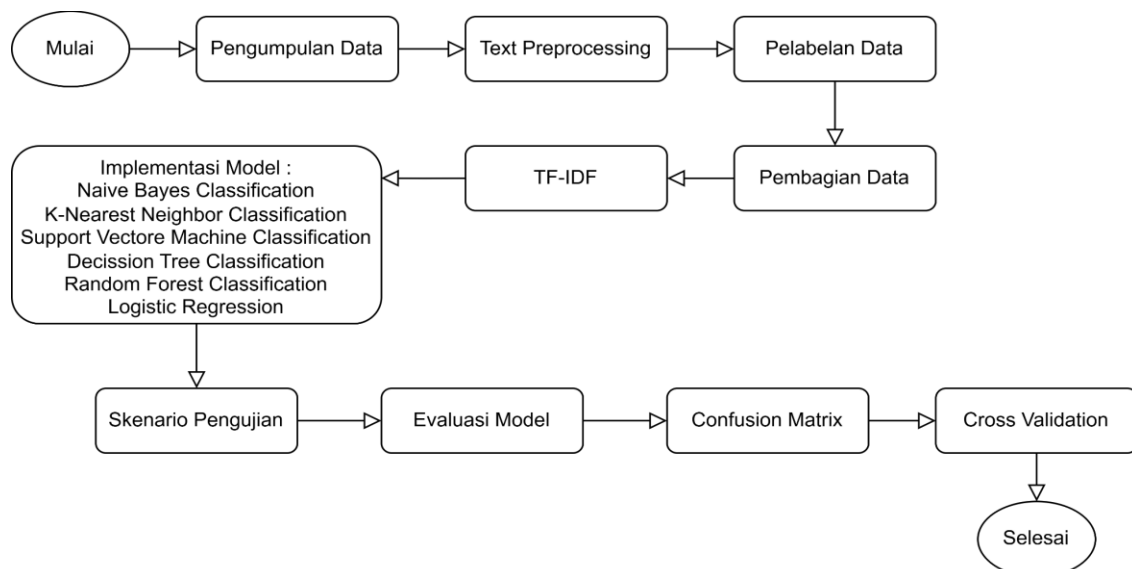
Beberapa penelitian sebelumnya telah mengevaluasi algoritma *machine learning* untuk memahami respons publik pada suatu topik. Nursingah dkk. (2024) menggunakan *Naïve Bayes* pada topik MBG di platform dan mencapai hasil akurasi 80.31% dengan dominasi opini positif sebesar 53.04% [13], sementara Krisnawan dkk. (2025) yang menerapkan *Bidirectional Gated Recurrent Unit* (BiGRU) pada topik yang sama menemukan sentimen publik justru cenderung negatif sebesar 57% dengan akurasi 80% [1].

Dalam studi perbandingan performa algoritma, Aditya dkk. (2025) menemukan *Naïve Bayes* paling optimal (79.84%) dibandingkan SVM dan K-NN pada klasifikasi sentiment aplikasi Info BMKG [14], sementara Ainunnisa & Sulastris (2023) menemukan bahwa algoritma *Logistic Regression* unggul dengan akurasi 84% dibandingkan dengan SVM dan K-NN pada ulasan TikTok [12]. Kurniawan dkk. (2025) merekomendasikan algoritma *Random Forest* untuk analisis sentimen ulasan investasi dengan *recall* mencapai 95%, meski SVM mencatat akurasi tertinggi sebesar 91% [11], sedangkan Octa.N dkk (2023) menunjukkan *Decision Tree* mencapai tingkat akurasi 70.27% pada sentimen *marketplace* Indonesia [10].

Penelitian sebelumnya umumnya hanya berfokus pada satu atau beberapa algoritma pada objek tertentu, sehingga belum memberikan gambaran menyeluruh mengenai performa berbagai algoritma dalam konteks dataset dan permasalahan yang sama. Hal ini menyebabkan kesulitan dalam menentukan algoritma yang paling *robust*, baik pendekatan berbasis probabilitas (*Naïve Bayes*), berbasis jarak (K-NN), berbasis fungsi pemisah (*hyperplane*) seperti SVM, berbasis pohon keputusan (*Decision Tree* dan *Random Forest*), maupun berbasis regresi (*Logistic Regression*) dalam menangani variasi ekspresi opini publik di platform X terkait Program Makan Bergizi Gratis. Dengan demikian, penelitian ini diharapkan mampu berkontribusi melalui perbandingan kinerja beberapa algoritma *machine learning*, meliputi *Naïve Bayes*, K-NN, SVM, *Decision Tree*, *Random Forest*, dan *Logistic Regression*, dan memberikan rekomendasi model yang memiliki kinerja terbaik dan *robust* dalam analisis sentimen opini publik pada platform X terkait Program Makan Bergizi Gratis.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif guna mengevaluasi dan melakukan perbandingan kinerja beberapa algoritma *machine learning* dalam proses klasifikasi sentimen. Proses penelitian mengikuti tahapan data mining dan *Natural Language Processing*, dimulai dari pengumpulan data melalui *web scraping* pada media sosial X, dilanjutkan dengan tahap *text preprocessing* (*cleaning, case folding, tokenization, stopword removal, dan stemming*), pelabelan data, pembagian dataset menjadi data latih dan data uji, kemudian ekstraksi fitur melalui TF-IDF guna menghasilkan representasi numerik [1]. Selanjutnya dilakukan implementasi model dan skenario pengujian untuk membandingkan enam algoritma, yaitu *Naïve Bayes*, K-NN, SVM, *Decision Tree*, *Random Forest*, dan *Logistic Regression*. Evaluasi model dinilai menggunakan matrik evaluasi dari *confusion matrix* (akurasi, presisi, *recall*, dan *F1-Score*) guna menentukan model paling optimal dalam mengelompokkan sentimen masyarakat pada Program MBG, serta dilakukan pengujian tambahan *cross validation* untuk menguatkan hasil evaluasi model, sebagaimana ditunjukkan oleh Gambar 1.



Gambar 1. Alur Penelitian

Berdasarkan Gambar 1, alur penelitian diawali dengan tahapan pengumpulan data hingga tahap *cross validation* yang menghasilkan model klasifikasi sentimen terbaik.

2.1 Pengumpulan Data

Tahap awal penelitian dimulai dengan mengumpulkan opini masyarakat dari platform media sosial X melalui teknik *web scraping* secara otomatis. Proses ini dijalankan dengan memanfaatkan ekstensi *Chrome* bernama *Web Scraper* untuk mengakses data secara *real-time* [15]. Data yang dikumpulkan berfokus pada cuitan yang memuat kata kunci spesifik terkait topik penelitian, yaitu "mbg", "program mbg", dan "program makan bergizi gratis" guna memperoleh opini yang paling relevan dengan kebijakan nasional tersebut [16].

2.2 Teks Preprocessing

Data mentah yang diperoleh umumnya bersifat tidak beraturan dan mengandung banyak *noise*, sehingga perlu tahapan *preprocessing* teks untuk memastikan kualitas dan konsistensi [17]. Tahapan *preprocessing* meliputi *cleaning* untuk menghilangkan elemen yang tidak relevan seperti URL, *mention*, angka, tanda baca, dan karakter khusus; *case folding* untuk mengonversi seluruh teks menjadi huruf kecil; *tokenization* untuk memecah kalimat menjadi token individual [18][19]. Dilakukan penghapusan kata-kata umum atau *stopword removal* dengan *library* Sastrawi, serta *stemming* yang juga diterapkan dengan memanfaatkan *library* Sastrawi guna mengonversi kata berimbuhan ke bentuk dasar [20].

2.3 Pelabelan Data

Proses pelabelan bertujuan untuk memberikan identitas pada setiap data komentar agar mesin dapat memahami kategori sentimen yang ada [17]. Data dikelompokkan ke dalam tiga label sentimen yaitu netral, positif, dan negatif [16]. Penelitian ini menggunakan pelabelan data berbasis leksikon (*InSet Lexicon*) yang mengklasifikasikan teks dengan mencocokkan kata terhadap kamus sentimen berskor polaritas, serta validasi manual untuk memastikan akurasi pelabelan [21].

2.4 Pembagian Data

Dataset pada penelitian ini dibagi menjadi dua bagian utama (*baseline*), yakni 80% digunakan sebagai data latih (*training data*) dan 20% sisanya sebagai data uji (*testing data*). Skema ini dinilai ideal karena mampu menyeimbangkan antara kecukupan data bagi algoritma dalam mempelajari pola serta ketersediaan data uji yang representatif untuk keperluan validasi model [15].

2.5 TF-IDF

TF-IDF diterapkan untuk mengubah teks menjadi bentuk numerik dengan cara menghitung bobot masing-masing kata, dimana TF (*Term Frequency-Inverse*) menghitung jumlah kemunculan suatu kata dalam suatu dokumen, sementara IDF (*Inverse Document Frequency*) berperan dalam pemberian bobot yang lebih tinggi pada kata-kata yang memiliki frekuensi kemunculan rendah di seluruh koleksi dokumen [11][17]. Teknik ini membantu sistem untuk mengenali kata-kata yang paling diskriminatif dan relevan dalam menentukan pola sentimen [20]. Perhitungan TF-IDF ditunjukkan pada persamaan (1) berikut :

$$W(t, d) = TF(t, d) \times IDF(t) \quad (1)$$

Nilai IDF dihitung untuk mengukur tingkat keunikan suatu kata di seluruh dataset guna mengurangi pengaruh kata-kata umum dengan tingkat kemunculan tinggi tetapi tidak informatif, dengan rumus (2) sebagai berikut [15]:

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

Dalam studi ini, TF-IDF dikonfigurasi menggunakan *n-gram* dengan rentang (1,2) yang memungkinkan pengambilan fitur dalam bentuk kata tunggal (*unigram*) serta pasangan kata (*bigram*) secara bersamaan, untuk menangkap konteks leksikal yang lebih rumit dalam teks.

2.6 Implementasi Model

Penelitian ini menguji enam algoritma *machine learning* untuk dibandingkan kinerjanya dalam mengklasifikasikan sentimen publik. Seluruh model dibangun menggunakan *library Scikit-learn* pada bahasa pemrograman *Python*, dengan fitur yang diperoleh melalui ekstraksi menggunakan TF-IDF.

a. Naïve Bayes

Algoritma yang berbasis probabilitas dan Teorema Bayes yang sangat efisien dalam mengelola data teks berdimensi tinggi [5][14][22]. Dalam penelitian ini digunakan varian Multinomial *Naïve Bayes* yang ideal untuk data teks dengan representasi frekuensi kata seperti TF-IDF. Persamaan (3) algoritma *Naïve Bayes* sebagai berikut [11]:

$$P(y) = \frac{P(x)P(x)}{P(y)} \quad (3)$$

b. K-NN

Algoritma klasifikasi berbasis jarak (*instance-based learning*) yang menentukan label data baru berdasarkan mayoritas label dari tetangga terdekatnya. Metode ini diterapkan dengan jumlah tetangga terdekat (k) sebesar 3. Persamaan (4) algoritma K-NN ditunjukkan sebagai berikut [9]:

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^n ((a_r(x_i) - a_r(x_j)))^2} \quad (4)$$

c. SVM

Algoritma *machine learning* berbasis *supervised learning* yang mengidentifikasi batas pemisah optimal antarkelas data pada ruang fitur berdimensi tinggi [23]. Algoritma ini dikonfigurasi menggunakan kernel *Radial Basis Function* (RBF) yang dapat memberikan hasil prediksi paling akurat dibandingkan kernel Linear maupun Polynomial [24], dengan parameter $C=1.0$ dan $\gamma='scale'$. Selain itu, parameter $class_weight='balanced'$ diterapkan guna menangani ketidakseimbangan distribusi data pada setiap kelas sentimen.

d. *Decision Tree*

Model prediksi berbasis struktur pohon hirarki yang memecah pengambilan keputusan menjadi aturan-aturan sederhana yang mudah dipahami [10]. *Decision Tree* digunakan dengan kriteria *entropy* serta pembatasan parameter seperti $min_samples_split=5$, $min_samples_leaf=2$, dan $max_depth=10$ untuk mengurangi risiko *overfitting*. Algoritma *Decision Tree* dengan kriteria *entropy* memiliki persamaan (5) sebagai berikut [25]:

$$Entropy(S) = -\sum_{i=1}^n p_i \log \log 2p_i \quad (5)$$

e. *Random Forest*

Algoritma *ensemble learning* yang menyusun kumpulan pohon keputusan guna memperkuat stabilitas dan akurasi klasifikasi [19]. Pada penelitian ini digunakan $n_estimators=200$, kriteria *entropy*, dan $class_weight='balanced'$. Persamaan (6) algoritma *Random Forest* sebagai berikut [11]:

$$f(x) = Average(f_1(x), f_2(x), \dots, f_n(x)) \quad (6)$$

f. *Logistic Regression*

Teknik statistik yang memodelkan hubungan antara variabel independen dengan probabilitas klasifikasi biner [12]. *Logistic Regression* diterapkan dengan $class_weight='balanced'$ dan pengaturan jumlah maksimum iterasi ($max_iter=1000$) guna memastikan proses konvergensi berjalan secara optimal. Algoritma *Logistic Regression* memiliki persamaan (7), (8) sebagai berikut [26]:

$$g(x) = sigmoid(a + \beta X) \quad (7)$$

$$sigmoid(x) = \frac{1}{1+exp^{-x}} \quad (8)$$

2.7 Skenario Pengujian

Tahap pengujian dirancang untuk mengevaluasi kinerja dan reliabilitas model klasifikasi sentimen melalui beberapa skenario, guna memastikan model mampu mengenali pola pada data yang belum pernah ditemui sebelumnya [11][27]. Skenario pertama (*baseline*) menggunakan pembagian data 80:20 dengan $n\text{-gram}$ (1,2) sebagai acuan utama evaluasi. Skenario kedua menerapkan pembagian 70:30 dengan $n\text{-gram}$ (1,2) untuk menguji konsistensi performa terhadap perubahan komposisi data, dengan konfigurasi fitur yang sama, sehingga perbedaan hasil dipengaruhi oleh variasi pembagian data [1][7]. Skenario ketiga menguji pengaruh representasi fitur dengan menggunakan TF-IDF *unigram* (1,1) pada pembagian data 80:20, guna melihat kontribusi konteks kata terhadap peningkatan performa model [15].

2.8 Evaluasi Model

Tahap evaluasi model merupakan proses krusial untuk menilai sejauh mana performa algoritma dalam menyelesaikan tugas klasifikasi sentimen. Efektivitas model diukur menggunakan metrik standar sebagai berikut [28]:

a. Akurasi (*Accuracy*): Mengukur rasio prediksi secara tepat terhadap keseluruhan data pada rumus (9):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

b. Presisi (*Precision*): Mengukur seberapa tepat prediksi kelas positif pada rumus (10):

$$Presisi = \frac{TP}{TP+FP} \quad (10)$$

c. *Recall*: Mengukur keberhasilan model dalam mengenali seluruh data positif aktual pada rumus (11):

$$Recall = \frac{TP}{TP+FN} \quad (11)$$

d. *F1-Score*: Merupakan rata-rata harmonic presisi dan *recall* pada rumus (12) :

$$F1-Score = 2 \times \frac{(Presisi \times recall)}{(Presisi + recall)} \quad (12)$$

2.9 Confusion Matrix

Confusion Matrix adalah salah satu metode yang digunakan dalam evaluasi kinerja yang diterapkan dalam penelitian ini guna mengukur efektivitas model *machine learning* dalam mengklasifikasikan data komentar ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif [5]. Matriks ini menyajikan hasil klasifikasi dalam bentuk tabel yang membandingkan label aktual (sebenarnya) dengan hasil prediksi yang dikeluarkan oleh model [17]. Selain itu, matriks ini sering kali divisualisasikan dalam ke bentuk *heatmap* untuk mempermudah interpretasi terhadap distribusi kesalahan serta kekuatan dan kelemahan model di tingkat kelas individual [5].

2.10 Cross Validation

Cross Validation adalah teknik validasi yang diterapkan untuk menilai kemampuan generalisasi dan kestabilan model. Dalam penelitian ini digunakan *Stratified K-Fold* dengan $k=10$, dimana dataset dibagi menjadi 10 *fold* dengan distribusi kelas yang seimbang, kemudian dilakukan pelatihan dan pengujian secara bergantian pada setiap *fold*. Hasil evaluasi dinyatakan dalam bentuk nilai rata-rata (*mean accuracy*) dan deviasi standar (*standard deviation*), dimana nilai rata-rata menunjukkan performa keseluruhan model, sedangkan deviasi standar menunjukkan tingkat konsistensi model terhadap variasi pembagian data [27].

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Dataset pada penelitian ini berupa 2.000 komentar dari media sosial X terkait Program Makan Bergizi Gratis (MBG) yang dikumpulkan selama Februari–April 2026. Gambar 2 menampilkan contoh data yang diambil sebagai sampel dalam studi ini.

	komentar
0	gausah diefisiensi pak, dimusnahkan saja anggarannya. output programnya cuman jadi tai. udah gitu udah terbukti ada ribuan kasus keracunan. bikin bergizi engga membunuh iya
1	Keknya pemerintah pun gaakan peduli juga mau anggaran defisit inflasi tinggi dan kejadian krisis moneter 98 atau maybe lebih parah penting MBG
2	Bubarkan MBG !!!
3	Pilihan: Stop MBG → BBM terkendali → fiskal selamat: Gak akan ada rakyat yang turun ke jalan → Rezim fokus kendalikan Fiskal dan Penegakan hukum MBG terus → BBM naik → Fiskal Jebol: Rakyat turun → Rezim OUT!
4	Harus sih, dampak perang berkepanjangan itu teramat seram loh. Bisa2 stagflation lagi ekonomi dunia seperti tahun 70an.

Gambar 2. Contoh Data Komentar

Berdasarkan Gambar 2 tersebut, dataset menampilkan kolom komentar yang berisi opini masyarakat terkait Program MBG. Data ini selanjutnya digunakan sebagai input dalam proses analisis sentimen.

3.2 Teks Preprocessing

Dataset opini publik dikumpulkan dari media sosial X diproses melalui serangkaian tahapan *preprocessing* guna meminimalkan gangguan (*noise*) serta meningkatkan konsistensi teks. Tahap ini menghasilkan data yang lebih bersih dan terorganisir, sehingga kata-kata yang digunakan menjadi lebih merepresentasikan kebutuhan pada analisis sentimen. Hasil teks *preprocessing* ditunjukkan pada Tabel 1 berikut.

Tabel 1. Hasil Teks *Preprocessing*

Komentar	Komen <i>clean</i>
gausah diefisiensi pak, dimusnahkan saja angga...	gausah efisiensi musnah anggar output program ...
Keknya pemerintah pun gaakan peduli juga mau a...	kek perintah gaakan peduli mau anggar defisit ...
Bubarkan MBG !!!	bubar mbg
Pilihan: Stop MBG → BBM terkendali → fiskal ...	pilih stop mbg bbm kendali fiskal selamat gak ...
Harus sih, dampak perang berkepanjangan itu te...	dampak perang panjang amat seram loh stagflati...

Berdasarkan Tabel 1, hasil tahapan *preprocessing* menunjukkan keberhasilan dalam membersihkan dan menyederhanakan data teks, dimana berbagai variasi penulisan kata informal seperti karakter khusus, tanda baca, dan *stopword* berhasil diubah ke dalam bentuk baku, serta kata berimbuhan telah dikembalikan ke bentuk dasarnya melalui proses *stemming*. Sehingga, kualitas data semakin meningkat dan siap untuk dimanfaatkan pada tahap ekstraksi fitur.

3.3 Pelabelan Data

Proses pelabelan data untuk penelitian ini menggunakan pendekatan berbasis leksikon (*lexicon-based*) dengan memanfaatkan kamus sentimen yang mengelompokkan kata ke dalam polaritas positif dan negatif. Setiap komentar dianalisis berdasarkan kemunculan kata dalam kamus tersebut, kemudian diberikan skor sentimen untuk menentukan kategori positif, netral, atau negatif. Distribusi hasil pelabelan ditunjukkan pada Tabel 2.

Tabel 2. Distribusi Kelas Sentimen

<i>Sentiment</i>	<i>count</i>
<i>Neutral</i>	1313
<i>Negative</i>	481
<i>Positive</i>	206

Berdasarkan Tabel 2, sentimen netral mendominasi dataset sebanyak 1313 data, diikuti sentimen negatif 481 data dan positif 206 data. Kondisi ini menunjukkan adanya ketidakseimbangan distribusi kelas yang dipengaruhi karakteristik media sosial, di mana banyak komentar bersifat informatif tanpa mengandung sentimen yang eksplisit. Selain itu, penggunaan bahasa nonformal, singkatan, dan sarkasme menjadi tantangan bagi pendekatan leksikon dalam mengidentifikasi sentimen secara tepat.

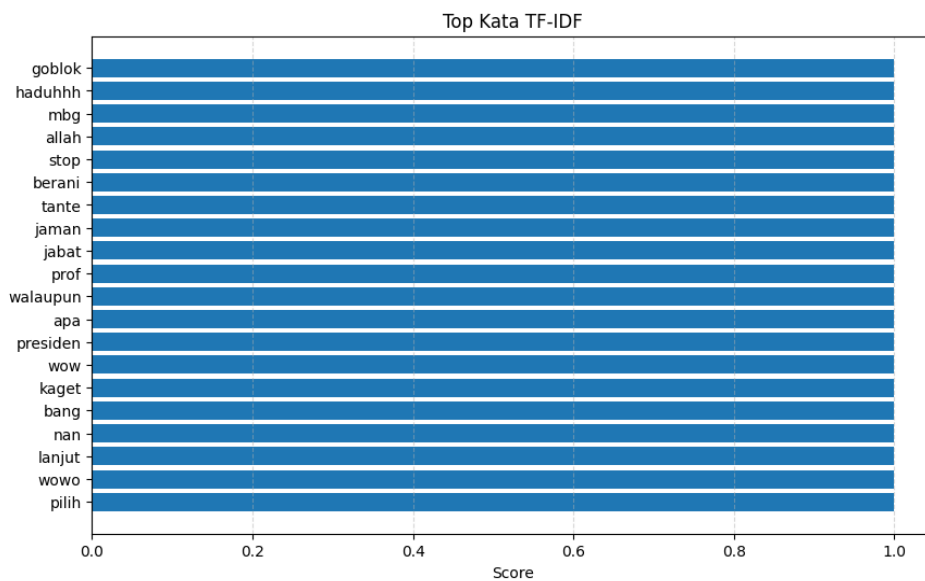
Untuk mengevaluasi kualitas pelabelan otomatis, dilakukan validasi terhadap 100 sampel data yang dipilih secara acak. Hasil pengujian menunjukkan tingkat kesesuaian sebesar 70%, yang mengindikasikan bahwa metode leksikon cukup efektif dalam mengidentifikasi polaritas sentimen secara umum, meskipun masih memiliki keterbatasan dalam memahami konteks bahasa yang kompleks.

3.4 Pembagian Data

Tahapan analisis dilanjutkan dengan membagi dataset ke dalam kategori data latih (*training set*) dan data uji (*testing set*). Pada skenario utama yaitu pembagian 80:20, terdapat sebanyak 1.600 data yang diterapkan sebagai data latih serta 400 data sebagai dimanfaatkan sebagai data uji, proporsi tersebut dinilai mampu memberikan distribusi data yang proporsional antara tahap pelatihan dan evaluasi model, sehingga model mampu menangkap pola data secara maksimal tanpa mengurangi kemampuan model dalam menghadapi data yang belum pernah dikenali sebelumnya.

3.5 TF-IDF

Setelah melalui tahapan pembagian data, data teks ditransformasikan menjadi representasi numerik menggunakan TF-IDF dengan konfigurasi *n-gram* (1,2). Gambar 3 menunjukkan hasil ekstraksi fitur melalui TF-IDF yang divisualisasikan dalam bentuk 20 kata dengan bobot terbesar.



Gambar 3. Bentuk 20 Kata TF-IDF dengan Bobot Tertinggi

Berdasarkan Gambar 3, proses ini menghasilkan matriks fitur berdimensi tinggi yang mencerminkan keragaman kata dalam komentar, di mana kata-kata seperti “goblok”, “haduhhh”, “mbg”, “stop” mendominasi bobot tertinggi sebagai fitur paling diskriminatif dalam menentukan pola sentiment, sehingga model *machine learning* dapat menangkap pola kata yang relevan dalam menentukan sentimen.

Untuk memperjelas penerapan TF-IDF, berikut disajikan contoh perhitungan manual menggunakan empat sampel data dengan kata sampel “mbg”, “program”, dan “efisien” (N=4). Perhitungan dilakukan menggunakan persamaan (1) dan (2), dengan hasil ditunjukkan pada Tabel 3 berikut.

Tabel 3. Hasil Perhitungan TF-IDF

Kata	TF(D1)	TF(D2)	TF(D3)	TF(D4)	IDF	TF-IDF(D1)	TF-IDF(D2)	TF-IDF(D3)	TF-IDF(D4)
mbg	1	0	0	1	0,301	0,301	0	0	0,301
program	1	1	0	0	0,301	0,301	0,301	0	0
efisien	1	0	0	0	0,602	0,602	0	0	0

Berdasarkan Tabel 3, kata “efisiensi” memperoleh bobot TF-IDF tertinggi pada D1 sebesar 0,602 karena hanya muncul pada satu dokumen sehingga memiliki nilai IDF tertinggi dan dianggap paling diskriminatif. sementara itu, kata “mbg” dan “program” memperoleh bobot yang sama sebesar 0.301 karena keduanya muncul pada dua dokumen. Hasil ini menunjukkan bahwa TF-IDF secara optimal mengalokasikan nilai yang lebih tinggi pada kata-kata yang bersifat unik dan informatif, sehingga representasi fitur yang dihasilkan mampu mendukung proses klasifikasi sentimen secara optimal.

3.6 Implementasi Model

Pada tahap ini, implementasi model diterapkan dengan membandingkan enam algoritma *machine learning* yang memanfaatkan *Python* dan *library Scikit-learn*. Data hasil ekstraksi fitur TF-IDF dibagi menjadi 80% data latih dan 20% data uji. Untuk mengatasi ketidakseimbangan kelas, model SVM, *Decision Tree*, *Random Forest*, dan *Logistic Regression* dikonfigurasi dengan parameter *class_weight='balanced'*.

Setiap algoritma menerapkan pendekatan yang berbeda dalam proses klasifikasi sentimen. *Naïve Bayes* melakukan klasifikasi berdasarkan peluang kemunculan tiap kata pada setiap kelas. K-NN menentukan label sentimen berdasarkan kedekatan data terhadap tetangga terdekat ($k=3$). SVM memisahkan kelas menggunakan *hyperplane* optimal dengan kernel RBF. *Decision Tree* membentuk aturan keputusan secara bertahap berdasarkan kriteria *entropy*, sedangkan *Random Forest* menggabungkan hasil dari 200 pohon keputusan guna menghasilkan prediksi yang lebih stabil. *Logistic Regression* memprediksi probabilitas setiap kelas sentimen melalui pendekatan regresi yang sesuai untuk data teks berdimensi tinggi.

3.7 Evaluasi Model

Kinerja model dinilai dengan menggunakan *Confusion Matrix* yang diturunkan menjadi *accuracy*, *precision*, *recall*, dan *F1-score*. Pengujian dilakukan pada data uji (*testing data*) sebanyak 400 data guna menilai kemampuan model dalam mengelompokkan sentimen ke dalam tiga kategori, yaitu positif, netral, dan negatif. Tabel 4 menyajikan hasil evaluasi performa model yang diimplementasikan.

Tabel 4. Hasil Evaluasi Model

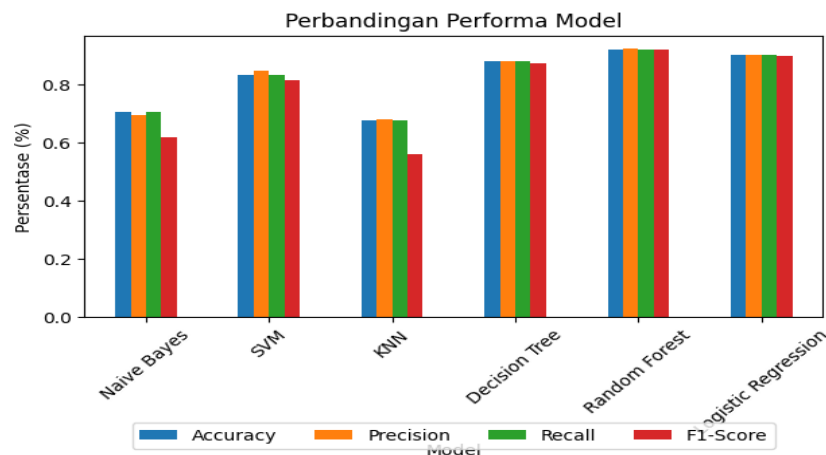
Model	Kelas	Precision	Recall	F1-score	Support
<i>Naïve Bayes</i>	<i>Negative</i>	1.00	0.19	0.32	95
	<i>Neutral</i>	0.69	1.00	0.82	264
	<i>Positive</i>	0.00	0.00	0.00	41
	<i>accuracy</i>			0.70	400
	<i>weighted avg</i>	0.69	0.70	0.62	400
<i>K-NN</i>	<i>Negative</i>	1.00	0.06	0.12	95
	<i>Neutral</i>	0.67	1.00	0.80	264
	<i>Positive</i>	0.00	0.00	0.00	41
	<i>accuracy</i>			0.68	400
	<i>weighted avg</i>	0.68	0.68	0.56	400
<i>SVM</i>	<i>Negative</i>	0.94	0.62	0.75	95
	<i>Neutral</i>	0.81	0.99	0.89	264
	<i>Positive</i>	0.87	0.32	0.46	41
	<i>accuracy</i>			0.83	400
	<i>weighted avg</i>	0.85	0.83	0.81	400
<i>Decision Tree</i>	<i>Negative</i>	0.90	0.63	0.74	95
	<i>Neutral</i>	0.87	0.98	0.93	264
	<i>Positive</i>	0.89	0.76	0.82	41
	<i>accuracy</i>			0.88	400
	<i>weighted avg</i>	0.88	0.88	0.87	400
<i>Random Forest</i>	<i>Negative</i>	0.96	0.80	0.87	95
	<i>Neutral</i>	0.92	0.98	0.95	264
	<i>Positive</i>	0.82	0.80	0.81	41
	<i>accuracy</i>			0.92	400

Model	Kelas	Precision	Recall	F1-score	Support
Logistic Regression	<i>weighted avg</i>	0.92	0.92	0.92	400
	<i>Negative</i>	0.85	0.80	0.83	95
	<i>Neutral</i>	0.92	0.96	0.94	264
	<i>Positive</i>	0.89	0.76	0.82	41
	<i>accuracy</i>			0.90	400
	<i>weighted avg</i>	0.90	0.90	0.90	400

Berdasarkan Tabel 4, model *Naïve Bayes* menghasilkan akurasi sebesar 70% dengan *weighted avg F1-Score* sebesar 0.62. Meskipun akurasi untuk kelas negatif sudah mencapai angka yang ideal, model ini sama sekali tidak berhasil mengidentifikasi kelas positif. Hal ini menunjukkan bahwa anggapan tentang *independensi* antar fitur dalam *Naïve Bayes* kurang efektif dalam mengenali pola kata yang membedakan antara sentimen positif dan netral pada komentar di media sosial yang cenderung ambigu. Hasil evaluasi model *K-Nearest Neighbor* (K-NN), yang mendapatkan nilai akurasi 68% dengan *weighted avg F1-Score* terendah sebesar 0.56, menunjukkan bahwa metode yang berfokus pada kemiripan jarak kurang berhasil dalam ruang fitur TF-IDF yang memiliki dimensi tinggi karena jarak antara titik data biasanya seragam, sehingga model mengalami kesulitan dalam membedakan kelas minoritas.

Selanjutnya, model *Support Vector Machine* (SVM) dengan akurasi 83% dan *weighted avg F1-Score* 0.81, menunjukkan peningkatan yang signifikan dibanding dua model sebelumnya. Namun *recall* kelas positif yang hanya 0.32 mengindikasikan bahwa meskipun kernel RBF mampu membentuk batas keputusan non-linear, ketidakseimbangan distribusi kelas masih mempengaruhi kemampuan model dalam mengenali sentimen positif secara optimal. Model *Decision Tree* dengan akurasi 88% dan *weighted avg F1-Score* 0.87 berhasil mengatasi kelemahan tersebut, di mana penggunaan *criterion entropy* terbukti efektif dalam memilih fitur kata yang paling diskriminatif untuk memisahkan ketiga kelas sentiment termasuk kelas minoritas.

Model *Random Forest* mencapai performa paling baik dengan akurasi 92% dan *weighted avg F1-Score* 0.92, menunjukkan bahwa mekanisme *ensemble* dari 200 pohon keputusan berhasil menurunkan variasi dan meningkatkan konsistensi dalam prediksi di semua kelas. Model *Logistic Regression* dengan akurasi 90% dan *weighted avg F1-Score* 0.90, menempati posisi kedua, membuktikan bahwa pendekatan probabilistik linear masih bisa bersaing dalam merepresentasikan fitur TF-IDF walaupun tidak dapat menangkap hubungan non-linear yang rumit seperti yang dilakukan oleh *Random Forest*. Hasil evaluasi pada skenario utama ditampilkan dalam grafik batang pada Gambar 4 guna mempermudah perbandingan performa seluruh model secara visual.



Gambar 4. Visualisasi Hasil Evaluasi Model

Mengacu pada Gambar 4, perbandingan visual keenam model menegaskan hasil pada Tabel 4 bahwa *Random Forest* mendominasi seluruh matrik evaluasi secara konsisten, disusul *Logistic Regression* dan *Decision Tree* dengan performa yang cukup kompetitif. Di sisi lain, *Naïve Bayes* dan K-NN tertinggal terutama pada *recall* dan *F1-score*, menunjukkan kurang optimalnya kemampuan kedua model dalam mengenali kelas minoritas.

Berdasarkan hasil evaluasi terhadap skenario utama tersebut, selanjutnya dilakukan perbandingan performa model pada beberapa skenario pengujian untuk melihat konsistensi kinerja model. Perbandingan hasil evaluasi seluruh model pada ketiga skenario pengujian ditampilkan oleh Tabel 5.

Tabel 5. Perbandingan Evaluasi Model Berdasarkan Skenario

Model	Skenario 1 (<i>baseline</i>)		Skenario 2		Skenario 3	
	<i>Accuracy</i>	<i>F1-Score</i>	<i>Accuracy</i>	<i>F1-Score</i>	<i>Accuracy</i>	<i>F1-Score</i>
<i>Naïve Bayes</i>	70.50%	61.51%	69.67%	60.20%	69.00%	58.96%
<i>SVM</i>	83.25%	81.29%	81.50%	79.42%	82.50%	80.45%
<i>K-NN</i>	67.50%	55.78%	66.83%	54.65%	73.00%	69.23%
<i>Decision Tree</i>	87.75%	87.02%	88.17%	87.39%	87.75%	87.03%

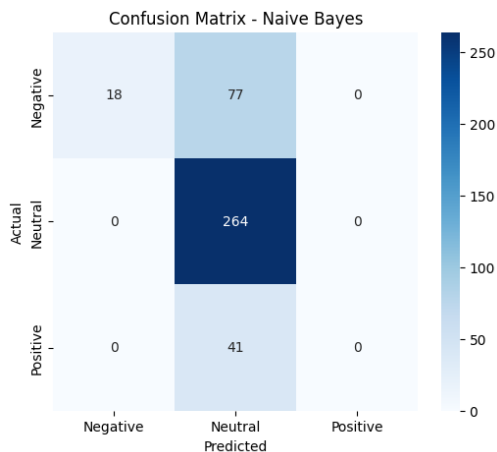
Model	Skenario 1 (<i>baseline</i>)		Skenario 2		Skenario 3	
	<i>Accuracy</i>	<i>F1-Score</i>	<i>Accuracy</i>	<i>F1-Score</i>	<i>Accuracy</i>	<i>F1-Score</i>
<i>Random Forest</i>	92.00%	91.83%	90.17%	89.76%	91.50%	91.28%
<i>Logistic Regression</i>	90.00%	89.83%	86.00%	85.76%	89.75%	89.52%

Berdasarkan hasil pada Tabel 5, *Random Forest* secara konsisten memperoleh performa terbaik pada seluruh skenario pengujian. Pada skenario utama (80:20 dengan TF-IDF *n-gram* (1,2)), model ini mencapai akurasi 92.00% dan *F1-Score* 91.83%, yang membuktikan bahwa model berbasis *ensemble* mampu menghasilkan model *robust* terhadap variasi pembagian data maupun representasi fitur. Pada skenario pembagian data 70:30 dengan penggunaan *n-gram* (1,2), sebagian besar model mengalami penurunan performa, namun *Random Forest* tetap menjadi model terbaik dengan akurasi 90.17% dan *F1-Score* 89.76% yang menunjukkan berkurangnya data latih mempengaruhi kemampuan model dalam mempelajari pola data.

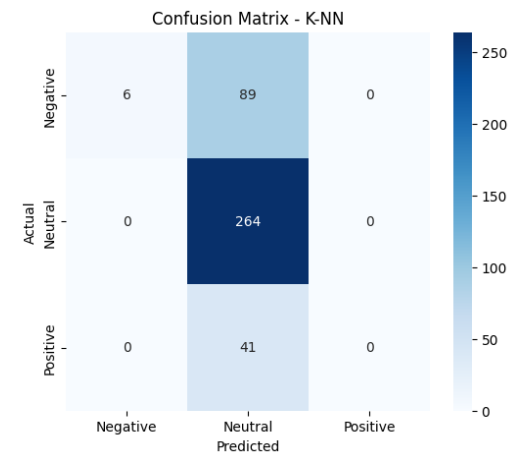
Sementara itu, pada skenario penggunaan TF-IDF *unigram* (1,1) dan pembagian data 80:20, performa model secara umum tidak mengalami perubahan yang signifikan dibandingkan *baseline*. *Random Forest* kembali memperoleh hasil tertinggi dengan akurasi 91.50% dan *F1-Score* 91.28%, sedangkan *Decision Tree* menunjukkan performa yang relatif stabil pada seluruh skenario. Secara keseluruhan, *Random Forest*, *Logistic Regression*, dan *Decision Tree* merupakan tiga model dengan performa paling kompetitif, sedangkan *Naïve Bayes* dan K-NN cenderung menghasilkan nilai akurasi dan *F1-Score* yang lebih rendah di seluruh skenario membuktikan bahwa kedua algoritma ini kurang mampu menangani karakteristik data yang tidak seimbang dan berdimensi tinggi.

3.8 Confusion Matrix

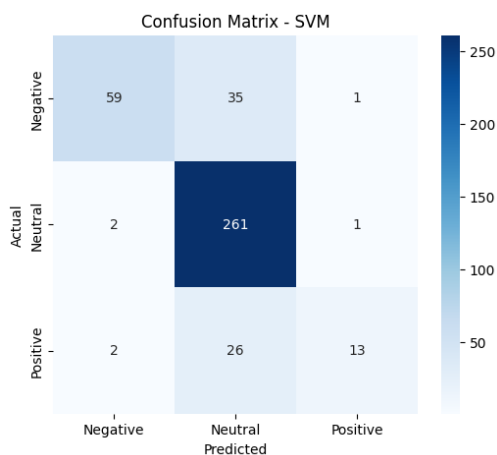
Dalam penelitian ini, *Confusion matrix* digunakan untuk memvisualisasikan kinerja serta kesalahan klasifikasi dari keenam algoritma yang diuji berdasarkan skenario utama. Hasil evaluasi tersebut disajikan pada Gambar 5 hingga Gambar 10, yang masing-masing merepresentasikan performa algoritma *Naïve Bayes*, K-NN, SVM, *Decision Tree*, *Random Forest*, dan *Logistic Regression*.



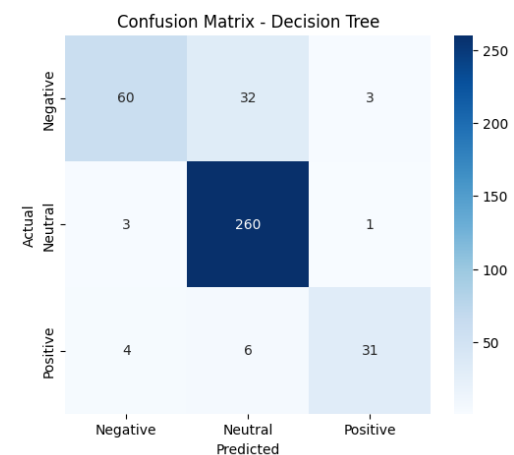
Gambar 5. Confusion Matrix Naïve Bayes



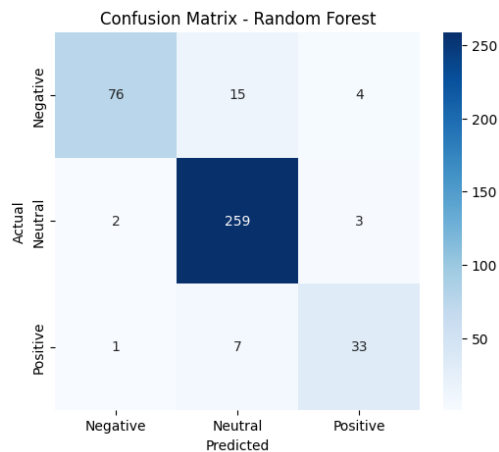
Gambar 6. Confusion Matrix K-NN



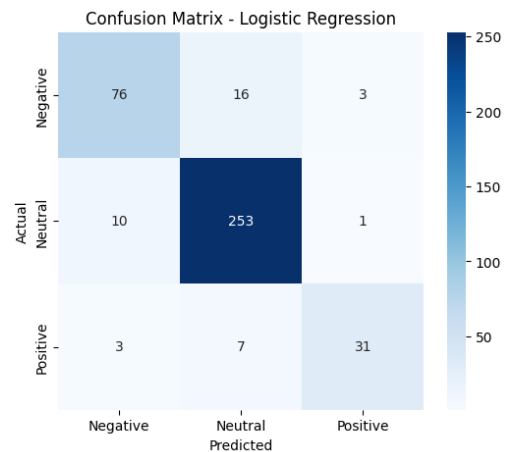
Gambar 7. Confusion Matrix SVM



Gambar 8. Confusion Matrix Decision Tree



Gambar 9. *Confusion Matrix Random Forest*



Gambar 10. *Confusion Matrix Logistic Regression*

Berdasarkan *Confusion Matrix* pada Gambar 5 hingga 10, model *Naïve Bayes* menunjukkan bias yang kuat terhadap kelas netral, dengan 77 dari 95 data negatif dan seluruh 41 data positif salah diklasifikasikan sebagai netral, hal ini menunjukkan ketidakmampuan model dalam mengenali kelas minoritas. Model *K-NN* hampir seluruh data negatif (89 dari 95) dan seluruh data positif (41) diprediksi sebagai netral, mempertegas pendekatan berbasis jarak kurang efektif pada fitur berdimensi tinggi. Model *SVM* menunjukkan peningkatan kinerja dengan mampu mengklasifikasikan secara tepat sebanyak 59 data negatif dan 13 data positif, meskipun masih terdapat sejumlah besar data yang keliru diprediksi ke dalam kelas netral (35 negatif dan 26 positif). Model ini tetap menunjukkan kecenderungan bias terhadap kelas netral (261 data benar).

Model *Decision Tree* menunjukkan performa lebih seimbang dengan 260 data netral diklasifikasikan dengan benar, meskipun misklasifikasi kelas ke kelas netral masih ditemukan, mengindikasikan bahwa pemisahan berbasis *entropy* belum sepenuhnya optimal pada kelas minoritas. *Random Forest* mencapai performa terbaik (76 data negatif, 259 data netral, 33 data positif), yang membuktikan bahwa metode *ensemble* mampu meminimalkan kecenderungan dominasi kelas mayoritas dan menghasilkan klasifikasi yang proporsional. *Logistic Regression* juga menunjukkan kinerja lebih stabil dengan kemampuan yang lebih baik dalam mengelompokkan ketiga kelas (76 negatif, 253 netral, dan 31 positif), meskipun masih terdapat kecenderungan kesalahan ke kelas netral.

Secara keseluruhan, *Confusion Matrix* memperlihatkan bahwa sebagian besar model memiliki kecenderungan bias terhadap kelas mayoritas (netral), namun model berbasis *ensemble* dan linear seperti *Random Forest* dan *Logistic Regression* dapat mengurangi bias tersebut dan memberikan hasil yang lebih seimbang.

3.9 Cross Validation

Untuk memperkuat hasil analisis perbandingan model, evaluasi tambahan juga dilakukan dengan menerapkan teknik *Stratified K-Fold Cross Validation*. Pada Tabel 6 di bawah menampilkan hasil *mean Cross Validation* dengan *fold* = 10 yang dibandingkan dengan akurasi model sebelum dilakukan *Cross Validation*.

Tabel 6. Hasil *Cross Validation*

Model	Akurasi	CV Mean	CV Std
<i>Naïve Bayes</i>	70.50%	70.88%	1.37
<i>K-NN</i>	67.50%	67.12%	0.80
<i>SVM</i>	83.25%	78.69%	1.90
<i>Decision Tree</i>	87.75%	86.75%	1.76
<i>Random Forest</i>	92.00%	89.00%	1.66
<i>Logistic Regression</i>	90.00%	85.19%	2.56

Hasil pengujian pada Tabel 6 ini menunjukkan bahwa model *Random Forest* tetap memperoleh rata-rata akurasi tertinggi sebesar 89.00%, diikuti oleh *Decision Tree* (86.75%) dan *Logistic Regression* (85.19%). Sementara itu, *SVM* memperoleh akurasi sebesar 78.69%, sedangkan *Naïve Bayes* dan *K-NN* menunjukkan performa yang lebih rendah, masing-masing sebesar 70.88% dan 67.12%. Nilai deviasi standar pada seluruh model relatif rendah (0.80-2.56), yang mengindikasikan bahwa performa model cukup stabil pada berbagai pembagian data. Meskipun *K-NN* memiliki deviasi terendah, performanya tetap kurang optimal, sedangkan *Logistic Regression* menunjukkan variasi yang sedikit lebih tinggi. Secara keseluruhan, hasil *Cross Validation* konsisten dengan hasil evaluasi sebelumnya, di mana *Random Forest* tetap menjadi model terbaik. Hal ini membuktikan bahwa model menunjukkan kinerja yang stabil dengan tingkat *overfitting* yang tidak signifikan.

3.10 Evaluasi Perbandingan dengan Penelitian Terdahulu

Perbandingan hasil penelitian ini dengan penelitian sejenis disajikan pada Tabel 7 berikut.

Tabel 7. Perbandingan dengan Penelitian Terdahulu

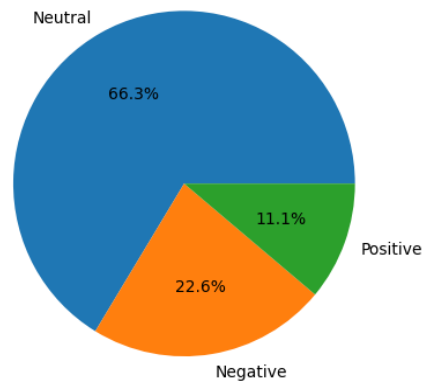
Peneliti	Metode Terbaik	Akurasi	Presisi	Recall	F1-Score
Nursinggah dkk. (2024)	<i>Naïve Bayes</i>	80.31%	-	-	-
Krisnawan dkk. (2025)	<i>BiGRU</i>	80.00%	80.00%	80.00%	80.00%
Aditya dkk. (2025)	<i>Naïve Bayes</i>	79.84%	60.21%	58.07%	57.31%
Ainunnisa & Sulastri (2023)	<i>Logistic Regression</i>	84.00%	83.00%	82.00%	83.00%
Kurniawan dkk. (2025)	<i>Random Forest</i>	85.00%	87.00%	95.00%	91.00%
Octa,N dkk. (2025)	<i>Decision Tree</i>	70.27%	-	-	-
Penelitian ini	<i>Random Forest</i>	92.00%	92.00%	92.00%	91.83%

Berdasarkan Tabel 7, penelitian ini memperoleh hasil paling baik dengan akurasi 92%, *precision* 92%, *recall* 92%, dan *F1-Score* 91,83% menggunakan algoritma *Random Forest*. Jika dibandingkan dengan penelitian Nursinggah dkk. (2024) dan Krisnawan dkk. (2025) yang sama-sama membahas sentimen Program MBG dengan akurasi masing-masing 80,31% dan 80%, hasil tersebut menunjukkan bahwa penggunaan TF-IDF *n-gram* (1,2) serta penerapan *class_weight* berkontribusi dalam meningkatkan performa model. Hasil ini juga sejalan dengan penelitian Kurniawan dkk. (2025) yang merekomendasikan *Random Forest* dengan akurasi 85%, sehingga memperkuat efektivitas metode *ensemble* dalam klasifikasi sentimen. Selain itu, akurasi *Decision Tree* pada penelitian ini mencapai 88%, lebih tinggi dibandingkan penelitian Octa. N dkk. (2023) sebesar 70,27%, yang mengindikasikan bahwa pengaturan parameter seperti *entropy* dan *class_weight* berpengaruh terhadap kemampuan model dalam menangani data yang tidak seimbang.

3.11 Analisis Sentimen Publik

Berdasarkan hasil evaluasi, *Random Forest* dipilih sebagai model terbaik dan digunakan untuk menganalisis distribusi sentimen publik terhadap Program MBG. Hasil distribusi kelas sentimen ditampilkan pada Gambar 11.

Distribusi Sentimen Masyarakat



Gambar 11. Diagram Distribusi Sentimen *Random Forest*

Hasil analisis menunjukkan sentimen netral sebesar 66,3%, disusul oleh sentimen negatif sebesar 22,6%, dan sentimen positif sebesar 11,1%. Dominasi sentimen netral mengindikasikan bahwa mayoritas komentar memiliki sifat informatif tanpa ekspresi emosi yang jelas, sementara proporsi sentimen negatif menunjukkan adanya kekhawatiran publik terkait implementasi program, seperti pengelolaan, transparansi, dan efektivitas. Sementara itu, sentimen positif mencerminkan dukungan terhadap manfaat program. Tingginya sentimen netral juga dapat dipengaruhi keterbatasan model dalam memahami bahasa non-formal seperti sarkasme, sehingga secara keseluruhan opini publik cenderung netral dengan kecenderungan kritik yang lebih terlihat dibandingkan dukungan.

4. KESIMPULAN

Berdasarkan hasil penelitian, dapat ditarik kesimpulan bahwa dari enam algoritma yang diuji, *Random Forest* menunjukkan performa terbaik dengan tingkat akurasi sebesar 92.00% yang didukung oleh hasil *cross validation* sebesar 89.00%. *Logistic Regression* dan *Decision Tree* juga menunjukkan kinerja yang stabil dan cukup seimbang dengan akurasi masing-masing sebesar 90.00% dan 87.75% dalam mengklasifikasikan ketiga kelas sentimen, sedangkan SVM memperoleh akurasi sebesar 83.25% yang menunjukkan model SVM masih mengalami keterbatasan dalam mendeteksi kelas positif sebagai kelas minoritas. Di sisi lain, *Naïve Bayes* dan *K-Nearest Neighbor* (K-NN) menunjukkan performa yang kurang optimal dengan akurasi sebesar masing-masing sebesar 70.50% dan 67.50% karena cenderung bias terhadap kelas mayoritas, yaitu netral. Berdasarkan model terbaik, distribusi sentimen publik menunjukkan dominasi sentimen netral sebesar 66.3% yang kemudian disusul sentimen negatif sebesar 22.6% dan sentimen positif sebesar 11.1%.

Ketidakseimbangan data dan kompleksitas bahasa nonformal seperti sarkasme terbukti memengaruhi kinerja model meskipun telah diterapkan *class_weight*. Penelitian selanjutnya disarankan mengkasji pendekatan *deep learning* dan peningkatan kualitas pelabelan data untuk hasil klasifikasi yang lebih optimal.

REFERENCES

- [1] Krisnawan, Z. A. Rabbani, Trimono, and M. Idhom, "Analisis sentimen program makan bergizi gratis menggunakan bidirectional gated recurrent unit," *It-Explore*, vol. 4, no. 3, pp. 282–294, 2025, doi: 10.24246/itexplore.v4i3.2025.pp282-294.
- [2] Marlina dan Y. Yusuf, "ANALISIS EFEKTIVITAS PROGRAM MAKAN BERGIZI GRATIS DI SEKOLAH DASAR INDONESIA TAHUN 2025.pdf," vol. 2, no. 2, pp. 2855–2866, 2025.
- [3] I. Febryanti, I. Indiati, Pane, and P. Astuti, "Implementasi Kebijakan Makan Bergizi Gratis (MBG) Studi Kasus pada SDN 3 Kepanjen Kabupaten Malang," *Dialogue J. Ilmu Adm. Publik*, vol. 7, no. 1, pp. 067–079, 2025, [Online]. Available: <https://ejournal2.undip.ac.id/index.php/dialogue/article/view/26628>
- [4] S. D. Waluyo, "Kebijakan Makanan Bergizi Gratis: Tinjauan Ekonomi Politik Dalam Kesejahteraan Dan Ketahanan Pangan," *J. Ilm. Ilmu Adm. Negara*, vol. 12, no. 1, pp. 2614–2945, 2025.
- [5] A. S. Muliana, D. Lestari, and S. P. Rafflesia, "Analysis of Public Sentiment on Election Results using Naïve Bayes in Social Media X," *Sistemasi*, vol. 13, no. 6, p. 2467, 2024, doi: 10.32520/stmsi.v13i6.4592.
- [6] H. H. Limbong and N. N. Norhikmah, "Optimization of Sentiment Analysis for Amikom One Application Reviews Using SMOTE with Artificial Neural Network Algorithm," *Sistemasi*, vol. 13, no. 5, p. 2048, 2024, doi: 10.32520/stmsi.v13i5.4437.
- [7] A. F. Rizkyllah, A. Meiriza, and D. Y. Hardiyanti, "Comparative Study of KNN and SVM Methods for Analyzing College Major Consistency Based on High School Background," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 14, no. 4, pp. 429–436, 2025, doi: 10.32736/sisfokom.v14i4.2521.
- [8] I. N. Amalina, N. Norhikmah, and W. M. Ashari, "Analysis of the Performance Comparison between Random Forest and SVM RBF in Detecting Cyberbullying on Imbalanced Data with the SMOTE Approach," *Sistemasi*, vol. 14, no. 6, p. 2768, 2025, doi: 10.32520/stmsi.v14i6.5574.
- [9] D. Munandar, M. Afdal, Z. Zarnelly, and R. Novita, "Analisis Sentimen Ulasan Pengguna Aplikasi Mobile Banking Menggunakan Algoritma K-Nearest Neighbor," *J. Teknol. Sist. Inf. dan Apl.*, vol. 7, no. 3, pp. 1309–1318, 2024, doi: 10.32493/jtsi.v7i3.41409.
- [10] A. T. Octa Nuryawan, M. Hasbullah, M. Rizal, M. F. Rajab, and N. Agustina, "Algoritma Decision Tree Untuk Analisis Sentimen Public Terhadap Marketplace Diindonesia," *Naratif J. Nas. Riset, Apl. dan Tek. Inform.*, vol. 5, no. 1, pp. 18–25, 2023, doi: 10.53580/naratif.v5i1.186.
- [11] N. H. Nanda Dwi Kurniawan, Praditya Rendi Ferdian, "Analisis Sentimen Algoritma Naïve Bayes, Support Vector Machine, dan Random Forest Pada Ulasan Aplikasi Ajaib," *J. Nas. Tenknologi dan Sist. Inf.*, vol. 11, no. 1, pp. 087–097, 2024, doi: 10.25077/TEKNOSI.v11i1.2025.87-97.
- [12] I. R. Ainunnisa and S. Sulastri, "Analisis Sentimen Aplikasi Tiktok dengan Metode Support Vector Machine (SVM), Logistic Regression dan Naïve Bayes," *J. Teknol. Sist. Inf. dan Apl.*, vol. 6, no. 3, pp. 423–430, 2023, doi: 10.32493/jtsi.v6i3.31076.
- [13] L. Nursingah, R. Ruuhwan, and T. Mufizar, "Analisis Sentimen Pengguna Aplikasi X Terhadap Program Makan Siang Gratis Dengan Metode Naïve Bayes Classifier," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 3, 2024, doi: 10.23960/jitet.v12i3.4336.
- [14] N. Aditiya, P. Setiaji, and Supriyono, "Analisis Sentimen Kepuasan Masyarakat terhadap Aplikasi 'INFO BMKG' menggunakan Naive Bayes, SVM, dan KNN," *Sist. J. Sist. Inf.*, vol. 14, no. 3, pp. 2540–9719, 2025, [Online]. Available: <http://sistemasi.ftik.unisi.ac.id>
- [15] U. Farid, A. Faqih, A. E. Pajri, and M. J. Vikri, "Analisis Perbandingan Algoritma SVM , Logistic Regression , Naive Bayes , dan XGBoost Untuk Deteksi Fake News," vol. 13, no. 1, 2026, doi: 10.30865/jurikom.v13i1.9492.
- [16] J. Friadi and D. E. Kurniawan, "Analisis Sentimen Ulasan Wisatawan Terhadap Alun-Alun Kota Batam: Perbandingan Kinerja Metode Naive Bayes dan Support Vector Machine," *J. Sist. Inf. Bisnis*, vol. 14, no. 4, pp. 403–407, 2024, doi: 10.21456/vol14iss4pp403-407.
- [17] F. Fahraini, "Analisis Sentimen Masyarakat Indonesia terhadap Keterlibatan Bill Gates dalam Program Vaksin TBC di Media Sosial X Menggunakan SVM," vol. 13, no. 1, pp. 29–47, 2026, doi: 10.30865/jurikom.v13i1.9431.
- [18] S. Nur, Z. Wati, and H. Darwis, "Naïve Bayes Classifier dan K-Nearest Neighbor pada Analisis Sentimen Perkuliahan Daring di Universitas Muslim Indonesia Naïve Bayes Classifier dan K-Nearest Neighbor on Sentiment Analysis of Online Lectures at Universitas Muslim Indonesia," vol. 5, no. 1, pp. 47–54, 2025.
- [19] F. A. Larasati, D. E. Ratnawati, and B. T. Hanggara, "Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 9, pp. 4305–4313, 2022, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [20] D. Kurniawan, H. D. Purnomo, and A. Iriani, "Analisis Sentimen Komentar Konsumen Industri Jamu di Media Sosial menggunakan Artificial Neural Network dan K-Nearest Neighbor," *J. Sist. Inf. Bisnis*, vol. 14, no. 3, pp. 210–223, 2024, doi: 10.21456/vol14iss3pp210-223.
- [21] A. S. Rizkia, Wufron, and F. F. Roji, "Sentiment Analysis of Coretax: A Comparison of Manual, Transformers- Based, and Lexicon-Based Data Labeling on IndoBERT Performance," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 5, no. 3, pp. 1037–1048, 2025.
- [22] Norhikmah, W. Nurastuti, A. Aminuddin, A. Sidauruk, and P. H. Gunawan, "Optimization and Collaboration of Fuzzy C-Mean, K-Mean, and Naïve Bayes Algorithms Using the Elbow Method for Micro, Small, and Medium Enterprises," *Int. J. Informatics Vis.*, vol. 9, no. 5, pp. 2062–2071, 2025, doi: 10.62527/joyv.9.5.3292.
- [23] ISK Idris, YA Mustofa, and IA Salihi, "Analisis Sentimen Terhadap Penggunaan AplikasiShopee Menggunakan Algoritma Support VectorMachine (SVM)," *Jambura J. Electr. Electron. Eng.*, vol. 5, no. 1, pp. 32–35, 2023.
- [24] A. W. Nugroho and Norhikmah, "Analisis Sentimen menggunakan Algoritma Support Vector Machine pada Covid_19 Sentiment Analysis using the Support Vector Machine Algorithm on Covid_19," *J. Sist. Inf.*, vol. 13, no. 4, pp. 2540–9719, 2024, [Online]. Available: <http://sistemasi.ftik.unisi.ac.id>
- [25] F. Qothrunnada et al., "Tone Detection On Teranika Musical Instrument Using Discrete Wavelet Transform And Decision Tree Classification" vol. 4, no. 2, pp. 373–380, 2023.

- [26] I. Rahmawati, T. Rika Fitriani, A. No'eman, and A. Y. P. Yusuf, "Analisis Sentimen Menggunakan Algoritma Logistic Regression Pada Penerbangan Lion Air berdasarkan Ulasan Platform Online," *J. Ris. Inform. dan Teknol. Inf.*, vol. 1, no. 1, pp. 11–16, 2023, doi: 10.58776/jriti.v1i1.60.
- [27] A. I. Kamil, O. N. Pratiwi, and D. Witarsyah, "Analisis Sentimen dan Pemodelan Topik terhadap Aplikasi Pembelajaran Online pada Platform Google Play," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 10, no. 2, pp. 836–849, 2025, doi: 10.29100/jupi.v10i2.6023.
- [28] A. Sagita *et al.*, "Penerapan Metode Random Forest Dalam Menganalisis," vol. 7, no. 6, pp. 3307–3313, 2023.