

Comparison of Naive Bayes, Support Vector Machine, and Indobert Methods for Classifying Public Sentiment towards the MBG Program on Platform X

Hanya Abriananta ^{1*}, Khothibul Umam ^{2**}, Nur Cahyo Hendro Wibowo ^{3*}, Maya Rini Handayani ^{4*}

* Teknologi Informasi, Fakultas Sains dan Teknologi, UIN Walisongo Semarang

23080960001@student.walisongo.ac.id ¹, khothibul_umam@walisongo.ac.id ², nur_cahyo@walisongo.ac.id ³, maya@walisongo.ac.id ⁴

Article Info

Article history:

Received 2026-04-01

Revised 2026-04-24

Accepted 2026-05-05

Keyword:

*Sentiment Analysis,
Free Nutritious Meals,
Naive Bayes,
Platform X,
Support Vector Machine,
IndoBERT,
Lexicon-Based Labeling,
TF-IDF.*

ABSTRACT

This study compares the effectiveness of three machine learning algorithms, namely Naive Bayes, Support Vector Machine (SVM), and IndoBERT, in classifying public sentiment towards the Free Nutritional Meal (MBG) program on Platform X. A total of 1,176 Indonesian language tweets were collected through Selenium-based web scraping from January 1 to March 31, 2025. Sentiment labeling using a lexicon-based approach with 51 positive domain-specific words and 50 negative domain-specific words, coupled with negation pattern detection, resulted in 62.8% positive tweets and 37.2% negative tweets. Preprocessing for Naive Bayes and SVM followed a six-stage workflow including stemming through PySastrawi, while IndoBERT used a minimal preprocessing approach to retain contextual information. Feature extraction applied TF-IDF with a maximum of 1,500 features and a unigram-bigram-n-gram range, with a stratified data split of 80:20. IndoBERT achieved the highest accuracy of 81.4% with a weighted F1 score of 0.81, followed by SVM at 74.2% (F1 score of 0.74) and Naive Bayes at 72.5% (F1 score of 0.72). A Wilcoxon signed-rank test on 5-fold cross-validation confirmed that the performance difference between Naive Bayes and SVM was not statistically significant ($p > 0.05$). These findings provide empirical evidence for policymakers to monitor public acceptance of government nutrition programs through social media analysis.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Program Makan Bergizi Gratis (MBG) adalah inisiatif strategis yang diluncurkan oleh pemerintahan Presiden Prabowo Subianto pada 6 Januari 2025. Inisiatif ini bertujuan untuk meningkatkan kualitas gizi bagi seluruh penduduk, khususnya anak usia sekolah, balita, dan ibu hamil serta ibu menyusui di seluruh Indonesia [1]. Program ini merupakan prioritas pemerintah dengan alokasi sumber daya keuangan yang besar. Pelaksanaan program MBG ini mendorong tumbuhnya berbagai opini di kalangan masyarakat, mulai dari dukungan antusias hingga kritik yang tajam mengenai efektivitas, transparansi anggaran, dan kesiapan infrastruktur distribusi [2].

Platform X atau dikenal dengan Twitter menjadi salah satu ruang digital utama di masyarakat Indonesia untuk menuangkan pendapat mereka secara langsung mengenai isu politik ataupun kebijakan pemerintah yang kontroversial.

Indonesia memiliki lebih dari 167 juta pengguna media sosial aktif, dengan rata-rata penggunaan harian melebihi tujuh jam. Platform-platform ini berfungsi sebagai wadah untuk bertukar pandangan tentang berbagai subjek dengan cakupan yang luas [3]. Dalam hal ini, platform media sosial berfungsi sebagai sumber data yang tepat untuk meneliti komentar atau opini public. Namun demikian, karena banyaknya data dan sifat teks yang tidak terorganisir, metode komputasi yang dapat secara mandiri menafsirkan dan mengkategorikan sentimen secara tepat dan terukur sangatlah penting [4].

Analisis sentimen adalah cabang dari pemrosesan bahasa alami (Natural Language Processing) yang berfokus pada identifikasi, ekstraksi, dan analisis opini dari kumpulan data teks yang besar [5]. Dalam konteks kebijakan publik, analisis sentimen dari media sosial dapat digunakan sebagai alat evaluasi yang relatif objektif untuk menilai tingkat penerimaan publik terhadap suatu program pemerintah.

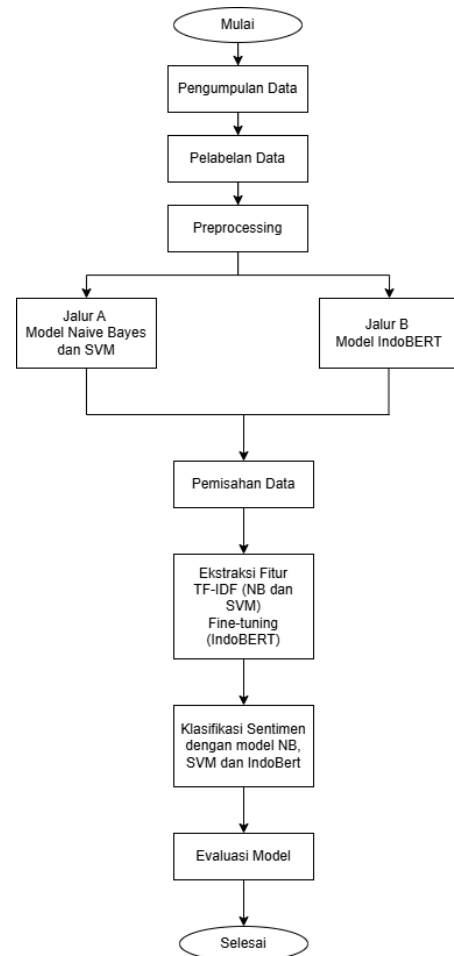
Pendekatan umum untuk klasifikasi sentimen dalam teks bahasa Indonesia meliputi pengklasifikasi Naive Bayes, Support Vector Machine (SVM), dan model berbasis transformer seperti IndoBERT.

Penelitian sebelumnya tentang analisis sentimen mengenai opini public terhadap program MBG di Platform X telah dilakukan. Laia dkk. melakukan penelitian dengan menggunakan Naive Bayes dengan 10-fold cross-validation, menghasilkan akurasi sebesar 86,46% [6]. Penelitian lain oleh Fatkhurrohman dkk. menerapkan SVM dan memperoleh akurasi 80% dengan weighted F1-score 0,76. Meskipun demikian, belum ada penelitian yang secara formal membandingkan Naive Bayes dan SVM menggunakan dataset MBG yang diperoleh melalui scraping berbasis Selenium dalam jangka waktu yang lama. [7]. Bersamaan dengan itu, sebuah studi yang dilakukan oleh Kadafi dkk. meneliti sentimen program MBG di berbagai platform menggunakan data dari platform X dan YouTube, membandingkan tiga algoritma: SVM, Random Forest, dan LSTM. Temuan menunjukkan bahwa Random Forest mencapai akurasi tertinggi sebesar 87,2%, diikuti oleh LSTM sebesar 81,8% dan SVM sebesar 79,1%. Kesimpulan yang diambil adalah bahwa Random Forest menunjukkan stabilitas yang unggul dalam mengelola input teks yang bervariasi dan distribusi kelas yang tidak merata [8].

Meskipun berbagai studi sebelumnya telah memberikan kontribusi yang signifikan, beberapa celah masih ada dan belum ditangani secara memadai. Pertama, belum ada studi yang secara sistematis membandingkan pengklasifikasi Naive Bayes, Support Vector Machine (SVM), dan IndoBERT dalam satu eksperimen terkontrol menggunakan dataset MBG yang identik. Kedua, sebagian besar studi tidak menyertakan pengujian signifikansi statistik untuk menentukan apakah perbedaan kinerja antar model bermakna secara ilmiah. Ketiga, potensi bias yang muncul dari penggunaan pelabelan otomatis berbasis leksikon jarang dibahas secara eksplisit. Oleh karena itu, studi ini bertujuan untuk mengisi celah ini dengan menyajikan analisis komparatif komprehensif dari ketiga algoritma tersebut, menggunakan dataset yang terdiri dari 1.181 tweet yang dikumpulkan secara independen antara 1 Januari dan 31 Maret 2025. Selain itu, studi ini juga mencakup analisis statistik mendalam dan diskusi terbuka tentang keterbatasan metodologi yang digunakan.

II. METODE

Studi ini mengadopsi pendekatan kuantitatif eksperimental untuk membandingkan kinerja tiga algoritma klasifikasi seperti Naive Bayes, Support Vector Machine (SVM), dan IndoBERT. Dalam tugas analisis sentimen pada teks bahasa Indonesia. Penelitian ini dirancang secara sistematis, meliputi delapan langkah utama: *pengumpulan data*, *pelabelan*, *pra-pemrosesan*, *pemisahan data*, *pembobotan menggunakan TF-IDF (khusus untuk Naive Bayes dan SVM)*, *penyempurnaan IndoBERT*, *klasifikasi*, dan *evaluasi kinerja model*.



Gambar 1. Alur Penelitian

A. Pengumpulan Data

Tahap ini melibatkan pengumpulan data dari Twitter melalui metodologi web scraping. Kata kunci yang digunakan meliputi "makan bergizi gratis", "program MBG", "MBG Prabowo", "makan gratis sekolah", "makan siang gratis", "gizi anak sekolah", "menu MBG", "anggaran MBG", "tolak makan gratis", dan "MBG bermasalah", bersama dengan variasi tambahan yang relevan dengan subjek kebijakan pemerintah. Data yang diperoleh dibatasi pada tweet berbahasa Indonesia dari 1 Januari 2025 hingga 31 Maret 2025 dan tidak termasuk retweet untuk mencegah duplikasi. Pada proses scraping menghasilkan total 1.181 tweet yang disimpan dalam bentuk kolom username, text, timestamp, reply, retweet, dan like. Data dikumpulkan dalam format teks yang belum diproses untuk analisis selanjutnya.

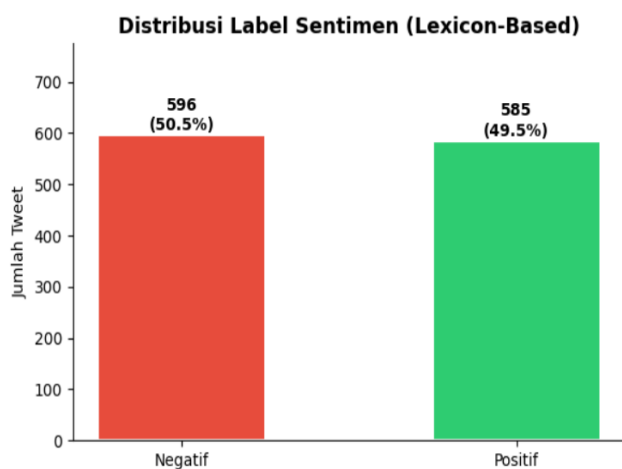
Perlu dicatat bahwa penggunaan Selenium dalam proses scraping pada platform X berpotensi menimbulkan noise berupa tweet yang tidak relevan secara tematik. Untuk mengurangi dampak ini, setiap tweet terlebih dahulu menjalani proses verifikasi relevansi berbasis pencocokan kata kunci sebelum dimasukkan ke dalam dataset akhir. Selain itu, dominasi konten dari akun publik menunjukkan

keterbatasan dalam representasi data. Oleh karena itu, temuan penelitian ini tidak dapat digeneralisasikan secara langsung untuk mewakili seluruh populasi Indonesia.

B. Pelabelan Data

Prosedur pelabelan sentimen dijalankan secara otomatis melalui pendekatan lexicon-based. Setiap tweet dinilai berdasarkan frekuensi kata dalam kamus KATA_POSITIF dan KATA_NEGATIF yang dikembangkan secara khusus untuk konteks MBG. Skor sentimen positif dihitung dari jumlah kemunculan kata positif dalam teks, sedangkan skor negatif dihitung dari kemunculan kata negatif. Selain itu, diterapkan pula deteksi pola negasi menggunakan ekspresi reguler untuk menangani frasa seperti "*tidak bagus*" atau "*gak setuju*" yang dapat mengubah polaritas kata. Setiap tweet diklasifikasikan sebagai Positif apabila skor positif melebihi skor negatif, atau Negatif apabila sebaliknya.

Pendekatan ini memiliki beberapa keterbatasan bawaan. Proses pelabelan otomatis berbasis leksikon belum mampu mengidentifikasi nuansa linguistik seperti sarkasme, ironi, atau konteks budaya yang kompleks, yang berpotensi mengakibatkan pelabelan data yang tidak akurat. Idealnya, validasi kebenaran data dilakukan melalui pelabelan manual oleh annotator manusia untuk mendapatkan hasil yang lebih andal. Selain itu, mengukur tingkat kesepakatan, sebagai bagian dari Pemrosesan Bahasa Alami merupakan aspek penting yang perlu dipertimbangkan dalam penelitian lebih lanjut.



Gambar 2. Hasil Pelabelan

C. Preprocessing

Setelah pelabelan dilakukan, data ulasan kemudian melewati tahapan *preprocessing*. Tahap ini bertujuan untuk meningkatkan kejelasan dan kesederhanaan data, sehingga memfasilitasi pemrosesan yang lebih baik oleh algoritma pembelajaran mesin. *Preprocessing* dalam penelitian ini dirancang ke dalam dua jalur yang berbeda, disesuaikan dengan karakteristik dan kebutuhan masing-masing model yang digunakan. Pendekatan ini membutuhkan beberapa

tahap teknis, termasuk *case folding*, *tokenizing*, *stopword removal*, serta *stemming*.

Pada jalur A digunakan untuk model Naïve Bayes dan SVM yang terdiri dari enam tahap *preprocessing* seperti *case folding*, *text cleaning*, *normalisasi*, *tokenisasi*, *stopword removal*, dan *stemming*. Lalu sedangkan jalur B digunakan untuk model IndoBERT yang terdiri dari tiga tahap saja seperti *case folding*, *text cleaning*, dan *normalisasi*. Hal ini disebabkan oleh penggunaan tokenizer WordPiece bawaan pada IndoBERT yang telah dilatih pada korpus besar berbahasa Indonesia, sehingga model mampu memahami konteks kata secara langsung tanpa memerlukan proses reduksi morfologis.

Case Folding adalah prosedur pengolahan data yang mengubah semua karakter dalam teks ulasan menjadi huruf kecil. Hal ini diterapkan untuk menstandarisasi penggunaan huruf kapital dan huruf kecil, yang dapat menyebabkan inkonsistensi dalam analisis [9]. Tahap selanjutnya adalah *tokenizing*, yang melibatkan pembagian kalimat menjadi unit kata yang terpisah untuk memungkinkan analisis teks pada level kata [10]. *Stopword removal*, mengacu pada penghapusan kata-kata yang dianggap memiliki dampak minimal pada makna keseluruhan teks. Kata-kata ini, jika dipertahankan, dapat mengganggu atau menyebabkan ketidakakuratan dalam proses analisis teks [11]. *Stemming* adalah proses untuk menghilangkan afiksasi dari kata berimbuhan dan mengembalikannya ke bentuk dasar. Penelitian ini menggunakan Pustaka Sastrawi untuk prosedur *stemming*, yang disesuaikan dengan ortografi Indonesia [12].

D. Pemisahan Data

Setelah tahap Preprocessing adalah tahap pemisahan data. Dataset dibagi dengan rasio 80:20, dengan 80% dialokasikan untuk pelatihan dan 20% untuk pengujian. Proses agregasi data ini digunakan untuk melatih model menggunakan data pelatihan dan menilai kinerjanya dengan data pengujian. Pemilihan rasio 80:20 dengan menggunakan metode stratified sampling untuk memastikan proporsi kelas yang seimbang pada kedua subset.

E. Pembobotan TF-IDF

TF-IDF adalah metode pembobotan yang mengubah teks menjadi representasi numerik dengan menghitung hasil perkalian frekuensi suatu kata dalam suatu dokumen (TF) dan kebalikan dari frekuensinya di seluruh korpus (IDF) [13]. Akibatnya, istilah yang banyak terdapat dalam satu teks tetapi jarang terdapat dalam teks lain akan mendapatkan signifikansi yang lebih besar. Fase ini menghasilkan nilai bobot untuk setiap kata, yang mencerminkan kepentingannya dalam setiap ulasan. Pada akhirnya, hal ini dapat berfungsi sebagai karakteristik dalam tahap analisis selanjutnya.

Metode Term Frequency Inverse Document Frequency (TF-IDF) digunakan untuk mengubah data teks menjadi vektor numerik sehingga dapat diproses oleh algoritma pembelajaran mesin konvensional. Dalam penelitian ini, penerapan TF-IDF diatur dengan beberapa parameter, yaitu

jumlah fitur maksimum 1.500, penggunaan rentang n -gram unigram dan bigram (1,2), batas frekuensi dokumen minimum 3, dan frekuensi dokumen maksimum 85%. Selain itu, penskalaan TF sublinear juga digunakan untuk mengurangi dominasi bobot pada kata-kata dengan frekuensi kemunculan yang sangat tinggi. Pengaturan ini bertujuan tidak hanya untuk merepresentasikan kata tunggal, tetapi juga mengakomodasi pasangan kata yang memiliki makna kontekstual dalam pembahasan MBG.

F. Klasifikasi Sentimen

Klasifikasi sentimen merupakan tahapan penting dalam analisis sentimen. Algoritma pembelajaran mesin diimplementasikan pada data yang sudah dipersiapkan melalui *pre-processing* dan pelabelan untuk mengklasifikasikan sentimen dataset.

1) Naïve Bayes Classifier

Naive Bayes adalah metode klasifikasi berbasis probabilitas yang memprediksi kategori suatu data dengan memanfaatkan informasi dari data sebelumnya. Metode ini disebut "naive" karena mengasumsikan bahwa setiap fitur atau ciri pada data saling berdiri sendiri dan tidak saling memengaruhi, meskipun dalam banyak kasus hal tersebut tidak sepenuhnya benar. Walaupun asumsi ini sederhana, Naive Bayes terbukti sangat efektif dan banyak digunakan dalam berbagai tugas klasifikasi [14]. Berdasarkan probabilitas kemunculan kata tertentu di setiap kelas, teknik analisis sentimen ini membagi konten ulasan ke kategori sentimen, seperti negatif atau positif. Perhitungan probabilitas dalam Naive Bayes mengacu pada rumus berikut [15]:

$$P(B) = \frac{P(B|A) \times P(A)}{P(B)} \quad (1)$$

Keterangan:

$P(A|B)$ = Probabilitas A terjadi, mengingat B terjadi

$P(B|A)$ = Probabilitas B terjadi, mengingat A terjadi

$P(A)$ = Probabilitas A terjadi, tanpa mempertimbangkan B

$P(B)$ = Probabilitas B terjadi, tanpa mempertimbangkan A

Dalam penelitian ini, parameter alfa ditetapkan pada 3,0 sebagai bentuk perataan Laplace yang lebih tinggi daripada nilai default, dengan tujuan untuk mengakomodasi keragaman kosakata media sosial yang cenderung tidak merata distribusinya.

2) Support Vector Machine

Support Vector Machine (SVM) adalah teknik yang digunakan dalam proses pengelompokan (clustering). Metode ini bertujuan untuk mengidentifikasi fungsi pemisahan (hyperplane) yang ideal untuk klasifikasi dua kategori data. Support Vector Machines menunjukkan ketahanan terhadap overfitting, terutama di ruang berdimensi tinggi di mana dimensinya melebihi jumlah sampel [16]. Support Vector Machine (LinearSVC) adalah metode

klasifikasi yang bekerja dengan menentukan hyperplane optimal untuk memisahkan dua kelas data dengan margin maksimum. Algoritma ini dikenal berkinerja baik dalam menangani data berdimensi tinggi dan jarang, seperti karakteristik representasi TF-IDF. Dalam penelitian ini, LinearSVC digunakan dengan parameter C sebesar 0,03 sebagai bentuk regularisasi yang relatif kuat untuk mengurangi risiko overfitting pada fitur teks yang jarang, dan batas iterasi maksimum 3.000.

Untuk setiap kelas (*support vector*), *hyperplane* dengan margin terbesar terhadap titik data terdekat dipilih. Meskipun demikian, SVM memerlukan durasi perhitungan yang lebih lama pada dataset. Studi ini menggunakan kernel linier karena kesesuaiannya yang lebih unggul untuk data yang menunjukkan representasi sparsitas tinggi, seperti TF-IDF. Berikut adalah rumus dasar perhitungan hyperplane dalam SVM:

$$f(x) = w^T x + b \quad (2)$$

Keterangan:

w = Vektor Bobot

x = Vektor Input

b = Bias

3) IndoBERT

IndoBERT adalah model bahasa yang dibangun berdasarkan arsitektur BERT (*Bidirectional Encoder Representations from Transformers*) dan dilatih menggunakan korpus bahasa Indonesia yang besar. Studi ini menggunakan model *indobenchmark/indobert base-pl*, yang kemudian di-fine-tuning untuk klasifikasi biner (positif dan negatif). Proses fine-tuning dilakukan selama tiga *epoch* dengan *learning rate* 2×10^{-5} , ukuran batch 16, dan panjang sequence maksimum 128 token. Proses optimasi menggunakan algoritma AdamW dengan nilai *weight decay* 0,01. Selanjutnya, penjadwal warmup linier diterapkan pada 10% dari total langkah pelatihan untuk meningkatkan stabilitas proses konvergensi model [17].

G. Evaluasi

Confusion matrix mengevaluasi kinerja model klasifikasi, menghasilkan metrik seperti akurasi, presisi, recall, dan *F1-score*. Tabel ini menyajikan gambaran kuantitatif dari klasifikasi yang benar yang diperoleh dari data uji [18]. Rumus selanjutnya berkaitan dengan akurasi, presisi, recall, dan *F1-score* :

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad (4)$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (5)$$

$$F1\ Score = \frac{2(Recall \times Precision)}{Recall + Precision} \times 100\% \quad (6)$$

Keterangan:

TP = True Positive

TN = True Negative
 FP = False Positive
 FN = False Negative

Accuracy digunakan untuk menentukan perbandingan prediksi yang benar terhadap keseluruhan dataset. Presisi menunjukkan akurasi model dalam mengidentifikasi kelas positif, sedangkan recall menandakan kemampuan model untuk mengenali semua kejadian yang benar-benar termasuk dalam kelas positif. Skor F1 menandakan keseimbangan antara presisi dan recall.

Kinerja ketiga model dievaluasi menggunakan empat metrik utama: akurasi, presisi, recall, dan skor F1 rata-rata tertimbang. Selanjutnya, matriks kebingungan dianalisis untuk mengidentifikasi pola kesalahan klasifikasi dalam setiap kelas. Untuk model *Naive Bayes* dan *Support Vector Machine (SVM)*, validasi silang bertingkat *5-fold* dan uji peringkat bertanda *Wilcoxon* diterapkan untuk menentukan signifikansi perbedaan kinerja antara keduanya pada tingkat kepercayaan $\alpha = 0,05$. Sementara itu, model *IndoBERT* tidak menjalani validasi silang karena keterbatasan sumber daya komputasi, sehingga evaluasinya hanya didasarkan pada satu set data uji.

Hasil validasi silang menunjukkan bahwa model pengklasifikasi *Naive Bayes* mencapai akurasi rata-rata $72,3\% \pm 1,8\%$, sedangkan *Support Vector Machine (SVM)* mencapai akurasi rata-rata sedikit lebih tinggi yaitu $73,9\% \pm 1,5\%$. Meskipun ada kecenderungan *SVM* memberikan kinerja yang lebih baik, hasil uji peringkat bertanda *Wilcoxon* untuk skor akurasi di setiap lipatan menunjukkan nilai $p > 0,05$. Ini menunjukkan bahwa perbedaan kinerja antara kedua model tersebut tidak signifikan secara statistik pada tingkat kepercayaan 95%. Dengan demikian, walaupun *SVM* secara konsisten unggul pada berbagai metrik evaluasi, selisih performa yang relatif kecil sekitar 2%, namun belum cukup kuat untuk dijadikan dasar klaim keunggulan yang meyakinkan. Diperlukan dataset dengan skala yang lebih besar atau eksperimen tambahan.

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Pada tahap pengumpulan data setelah proses filtering, menghasilkan 1.176 tweet berbahasa Indonesia yang berkaitan dengan program MBG dari platform X. Data dikumpulkan dalam rentang waktu 1 Januari 2025 hingga 31 Maret 2025 menggunakan sepuluh kata kunci pencarian. Setelah proses deduplikasi dan filtering, seluruh data disimpan dalam format CSV dengan sembilan atribut: keyword, username, text, timestamp, reply, retweet, like, tahun, dan bulan. Hasil selanjutnya diperoleh dari pengambilan data yang digambarkan pada gambar.

Jumlah data awal : 1181
 Kolom : ['keyword', 'username', 'text', 'timestamp', 'reply', 'retweet', 'like', 'tahun', 'bulan']

	keyword	username	text	timestamp	reply	retweet	like	tahun	bulan
0	program MBG	adjQncy	kemarin sempat viral yg katanya akan import sapi ratusan...	2025-01-30T22:44:03.000Z	0	0	0	2025.0	1.0
1	program MBG	Siti Romiah	Jangankan nyawa pak?! Menyisihkan harta untuk program ma...	2025-01-30T22:43:32.000Z	0	0	0	2025.0	1.0
2	program MBG	CNN Indonesia	Di Hadapan Prabowo, Listyo Pamer Program MBG yang Dilaku...	2025-01-30T22:32:26.000Z	0	0	0	2025.0	1.0
3	program MBG	banindulanx	Biaya produksi per bulan 1 M, sunih keluar modal dulu, t...	2025-01-30T21:53:57.000Z	0	0	0	2025.0	1.0
4	program MBG	Electrizen Ketapang	PLN sebagai BUMN terus mendukung penuh program Makan Ber...	2025-01-30T21:38:39.000Z	0	0	0	2025.0	1.0

Gambar 3. Hasil Pengumpulan Data

B. Pelabelan Data

Hasil pelabelan otomatis menggunakan pendekatan berbasis leksikon menunjukkan distribusi 742 tweet dengan label positif (62,8%) dan 439 tweet dengan label negatif (37,2%), menghasilkan rasio ketidakseimbangan kelas sekitar 1,69:1. Distribusi ini menunjukkan bahwa respons publik terhadap program MBG umumnya cenderung positif. Namun, proporsi sentimen negatif yang relatif signifikan (37,2%) mencerminkan kekhawatiran atau kritik yang substansial, khususnya mengenai implementasi program di lapangan.

TABEL I.
 HASIL PELABELAN

Tweet	Label
kemarin sempat viral yg katanya akan import sapi ratusan ribu bahkan jutaan buat mensukseskan program MBG sekarang serangga jadi alternatif	Negatif
Program MBG telah menunjukkan efektivitas dalam meningkatkan distribusi dan pemenuhan gizi anak-anak di seluruh Indonesia.	Positif
Jangankan nyawa pak?! Menyisihkan harta untuk program makan siang gratis aja ogah, rekayasa bukti paling depan mereka.	Negatif
Listyo memamerkan program MBG yang dilaksanakan oleh TNI-Polri di hadapan Prabowo.	Positif

C. Preprocessing

Tahap selanjutnya adalah preprocessing, langkah penting dan mendasar sebelum menerapkan teknik klasifikasi pada suatu dataset. Tahapan yang tercakup dalam prosedur preprocessing terdiri dari hal-hal berikut:

1) Case Folding

Prosedur mengubah semua huruf kapital dalam data menjadi huruf kecil untuk menstandarisasi format teks dan mengurangi perbedaan yang disebabkan oleh penggunaan huruf kapital yang berlebihan [19].

TABEL II.
 CASE FOLDING

Sebelum Case Folding	Sesudah Case Folding
Program Makan Bergizi Gratis MBG Terbukti Membantu Pemerintah	program makan bergizi gratis mbg terbukti membantu pemerintah

2) Text Cleaning

Text Cleaning adalah tahap pemrosesan data yang dirancang untuk menghilangkan elemen-elemen yang tidak perlu dari teks, seperti angka, tanda baca, tag HTML, dan simbol lainnya, sehingga meningkatkan keteraturan dan konsistensi data.

TABEL III.
TEXT CLEANING

Sebelum Text Cleaning	Sesudah Text Cleaning
@prabowo MBG bagus bgt!! #MakanBergiziGratis	mbg bagus bgt makan bergizi gratis

3) Normalize

Normalize kata bertujuan untuk mengganti istilah non-standar atau sehari-hari dengan padanan standarnya guna menstandarkan variasi ejaan dari kata yang sama [20].

TABEL IV.
NORMALIZE

Sebelum Normalize	Sesudah Normalize
mbg bagus bgt krn bisa bantu anak2 yg kurang gizi dgn program ini	mbg bagus banget karena bisa bantu anak-anak yang kurang gizi dengan program ini

4) Tokenizing

Tokenizing adalah proses membagi teks menjadi bagian-bagian yang lebih kecil, seperti token atau kata. [21].

TABEL V.
TOKENIZING

Sebelum Tokenizing	Sesudah Tokenizing
mbg bagus banget karena bisa bantu anak-anak yang kurang gizi dengan program ini	['mbg', 'bagus', 'banget', 'karena', 'bisa', 'bantu', 'anak', 'kurang', 'gizi', 'program', 'ini']

5) Stopward Removal

Stopword removal adalah menghilangkan kata-kata umum yang tidak secara substansial meningkatkan makna sentimen, seperti "dan", "yang", "di", dan "ke".

TABEL VI.
STOPWARD REMOVAL

Sebelum Stopward Removal	Sesudah Stopward Removal
['mbg', 'bagus', 'banget', 'karena', 'bisa', 'bantu', 'anak', 'kurang', 'gizi', 'program', 'ini']	['mbg', 'bagus', 'banget', 'bantu', 'anak', 'kurang', 'gizi']

6) Stemming

Stemming mengubah kata-kata tertutup menjadi bentuk dasarnya (kata akar) melalui library PySastrawi, yang menggunakan pendekatan Enhanced Confix Stripping (ECS) untuk bahasa Indonesia.

TABEL VII.
STEMMING

Sebelum Stemming	Sesudah Stemming
['mendukung', 'program', 'makan', 'bergizi', 'pemerintah', 'meningkatkan']	['dukung', 'program', 'makan', 'gizi', 'perintah', 'tingkat']

D. Pemisahan Data

Dataset tersebut dibagi menjadi dua kelompok utama: 80% dialokasikan untuk data pelatihan dan 20% sisanya digunakan sebagai data uji.

TABEL VIII.
HASIL PEMISAHAN DATA

Jenis Data	Presentase	Jumlah
Data Training	80%	940
Data Testing	20%	236
Total	100%	1176

E. Pembobotan TF-IDF

Term Frequency-Inverse Document Frequency (TF-IDF) digunakan untuk mengubah representasi tekstual menjadi vektor numerik yang sesuai untuk diproses oleh teknik pembelajaran mesin. TF-IDF memberikan signifikansi yang lebih besar pada frasa yang sering muncul di dalam teks tertentu namun jarang muncul di seluruh korpus, sehingga meningkatkan keterwakilannya.

Jumlah fitur TF-IDF: 1,203

Top-5 fitur TF-IDF (bobot rata-rata tertinggi):

1. anak	0.042017
2. sekolah	0.039767
3. siang	0.039462
4. menu	0.034993
5. masalah	0.029080

Gambar 4. Hasil Pembobotan TF-IDF

Dari penerapan pembobotan TF-IDF pada dataset, lima kata dengan bobot rata-rata terbesar diidentifikasi sebagai kata yang paling mewakili topik program MBG. Kata "anak" menempati posisi pertama dengan bobot tertinggi sebesar 0,042017, diikuti oleh kata "sekolah" (0,039767), "siang" (0,039462), "menu" (0,034993), dan "masalah" (0,029080).

Kemunculan kata "anak" dan "sekolah" sebagai kata dengan bobot tertinggi mengindikasikan bahwa diskusi publik terkait program MBG di platform X banyak berpusat pada sasaran utama program, yaitu anak-anak usia sekolah. Sementara itu, kehadiran kata "masalah" dalam lima besar menunjukkan bahwa sebagian percakapan juga memuat keluhan atau kekhawatiran masyarakat terhadap pelaksanaan program tersebut.

F. Visualisasi Word Cloud

Visualisasi frekuensi kata dari tweet yang telah melalui tahap preprocessing ditampilkan dalam bentuk word cloud untuk masing-masing kelas sentimen. Besarnya setiap kata dalam word cloud mencerminkan frekuensi kemunculannya dalam dataset. Kata-kata yang lebih sering disebutkan disajikan dengan ukuran huruf yang lebih besar, sehingga memudahkan identifikasi kata-kata dominan secara visual pada setiap kelas sentimen. Gambar 5 dan 6 mengilustrasikan word cloud dengan terminologi positif dan negatif.



Gambar 5. Wordcloud Kata Negatif



Gambar 6. Wordcloud Kata Positif

G. Evaluasi Model

Kinerja model klasifikasi sentimen dinilai menggunakan metrik evaluasi konvensional: akurasi, presisi, recall, dan F1-score. Kriteria tersebut digunakan untuk menilai kemampuan model dalam mengklasifikasikan data uji dengan tepat. Metrik presisi, recall, dan F1-score dianalisis secara terpisah untuk setiap kelas sentimen (positif dan negatif) untuk meningkatkan pemahaman tentang efektivitas model dalam membedakan setiap kategori. Gambar 7, 8 dan 9 menampilkan hasil penilaian pada data uji.

```

=====
Hasil Klasifikasi – Naive Bayes
=====

```

	precision	recall	f1-score	support
Positif	0.70	0.80	0.75	119
Negatif	0.76	0.65	0.70	117
accuracy			0.72	236
macro avg	0.73	0.72	0.72	236
weighted avg	0.73	0.72	0.72	236

Gambar 7. Klasifikasi Naive Bayes

Gambar 7 mengilustrasikan bahwa model Naive Bayes (MultinomialNB) mencapai akurasi 0,72, yang menunjukkan bahwa model tersebut secara benar mengidentifikasi 72% dari data uji yang dinilai. Label Positif menunjukkan presisi 0,70,

recall 0,80, dan F1-score 0,75. Label Negatif mencapai presisi 0,76, recall 0,65, dan skor F1 0,70. Model Naive Bayes menghasilkan nilai rata-rata makro sebesar 0,73 untuk presisi, 0,72 untuk recall, dan 0,72 untuk skor F1, sedangkan nilai rata-rata tertimbang menunjukkan hasil yang serupa yakni 0,73 untuk presisi, 0,72 untuk recall, dan 0,72 untuk skor F1. Hasil ini menunjukkan bahwa model Naive Bayes memiliki kinerja yang cukup merata di kedua kategori sentimen.

```

=====
Hasil Klasifikasi – SVM (LinearSVC)
=====

```

	precision	recall	f1-score	support
Positif	0.71	0.82	0.76	119
Negatif	0.78	0.67	0.72	117
accuracy			0.74	236
macro avg	0.75	0.74	0.74	236
weighted avg	0.75	0.74	0.74	236

Gambar 8. Klasifikasi SVM

Gambar 8 mengilustrasikan bahwa model SVM (LinearSVC) mencapai akurasi 0,74, yang menunjukkan bahwa model tersebut secara benar mengidentifikasi 74% dari data uji yang dinilai. Label Positif menunjukkan presisi 0,71, recall 0,82, dan F1-score 0,76. Label Negatif mencapai presisi 0,78, recall 0,67, dan skor F1 0,72. Model SVM menghasilkan nilai rata-rata makro masing-masing sebesar 0,75 untuk presisi, 0,74 untuk recall, dan 0,74 untuk skor F1, sedangkan nilai rata-rata tertimbang menunjukkan hasil yang konsisten yakni 0,75 untuk presisi, 0,74 untuk recall, dan 0,74 untuk skor F1. Hasil ini menunjukkan bahwa model SVM memiliki kinerja yang relatif seimbang di kedua kategori sentiment.

```

=====
Hasil Klasifikasi – IndoBERT
=====

```

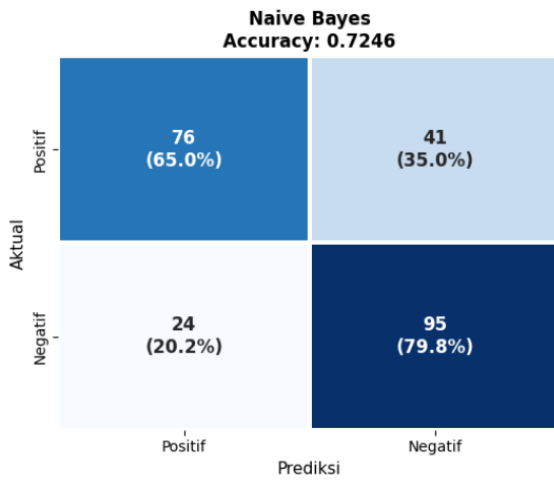
	precision	recall	f1-score	support
Positif	0.77	0.91	0.83	119
Negatif	0.88	0.72	0.79	117
accuracy			0.81	236
macro avg	0.83	0.81	0.81	236
weighted avg	0.82	0.81	0.81	236

Gambar 9. Klasifikasi IndoBERT

Gambar 9 mengilustrasikan bahwa model IndoBERT mencapai akurasi 0,81, yang menunjukkan bahwa model tersebut secara benar mengidentifikasi 81% dari data uji yang dinilai. Label Positif menunjukkan presisi 0,77, recall 0,91, dan F1-score 0,83. Label Negatif mencapai presisi 0,88, recall 0,72, dan skor F1 0,79. Model IndoBERT menghasilkan nilai rata-rata makro sebesar 0,83 untuk presisi, 0,81 untuk recall, dan 0,81 untuk skor F1, sedangkan nilai rata-rata tertimbang menunjukkan hasil yang konsisten yakni 0,82 untuk presisi, 0,81 untuk recall, dan 0,81 untuk skor F1. Hasil ini menunjukkan bahwa model IndoBERT memiliki kinerja yang relatif seimbang di kedua kategori sentimen, meskipun

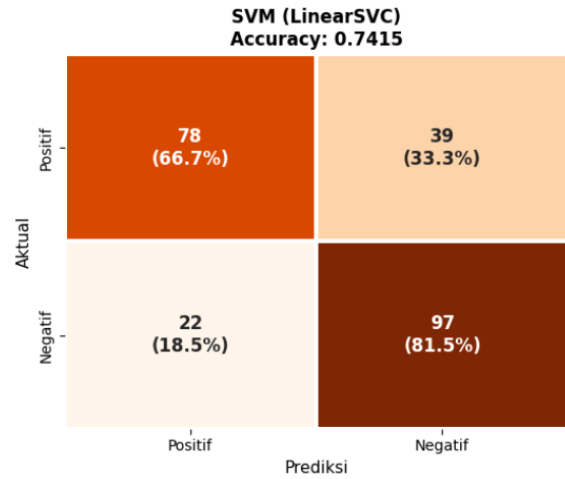
terdapat perbedaan yang cukup mencolok antara recall kelas Positif (0,91) dan recall kelas Negatif (0,72). Tingginya recall pada kelas Positif mengindikasikan bahwa model sangat sensitif dalam mendeteksi tweet bersentimen positif, sementara presisi kelas Negatif yang tinggi (0,88) menunjukkan bahwa ketika model memprediksi sebuah tweet sebagai Negatif, prediksi tersebut hampir selalu tepat.

Confusion matrix disediakan untuk menjelaskan pola kesalahan prediksi untuk setiap model klasifikasi, yang menunjukkan kinerja klasifikasi setiap algoritma, seperti yang terlihat pada Gambar 10, 11 dan 12.



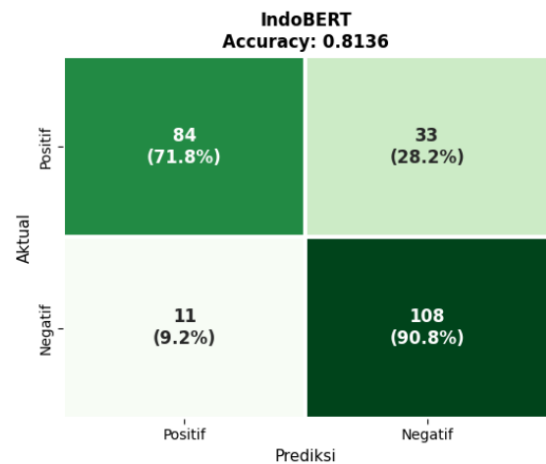
Gambar10. Confusion Matrix Naive Bayes

Gambar 10 menyajikan confusion matrix dari model Naive Bayes pada data uji, dengan akurasi keseluruhan 0,7246. Dari 117 set data aktual yang diberi label Positif, model tersebut mengklasifikasikan 76 dengan benar sebagai Positif (positif sejati), mewakili 65,0% dari total, sementara 41 diklasifikasikan secara salah sebagai Negatif (negatif palsu), mewakili 35,0%. Selanjutnya, dari 119 set data aktual yang diberi label Negatif, model tersebut mengidentifikasi 95 dengan benar sebagai Negatif (negatif sejati), mewakili 79,8% dari total, dan hanya 24 yang diklasifikasikan secara salah sebagai Positif (positif palsu), mewakili 20,2%.



Gambar 11. Confusion Matrix SVM

Gambar 11 menyajikan confusion matrix dari model Support Vector Machine (LinearSVC) pada data uji, dengan akurasi keseluruhan 0,7415. Dari 117 set data aktual yang diberi label Positif, model tersebut mengklasifikasikan 78 dengan benar sebagai Positif (positif sejati), mewakili 66,7% dari data, sementara 39 diklasifikasikan secara salah sebagai Negatif (negatif palsu), mewakili 33,3%. Sebaliknya, dari 119 set data aktual yang diberi label Negatif, model tersebut mengidentifikasi 97 dengan benar sebagai Negatif (negatif sejati), mewakili 81,5% dari data, dan hanya 22 yang diklasifikasikan secara salah sebagai Positif (positif palsu), mewakili 18,5%.



Gambar 12. Confusion Matrix IndoBERT

Gambar 12 menyajikan confusion matrix dari model IndoBERT pada data uji, dengan akurasi keseluruhan 0,8136. Dari 117 set data aktual yang diberi label Positif, model tersebut mengklasifikasikan 84 dengan benar sebagai Positif (True Positive), mewakili 71,8% dari total data Positif, sementara 33 diklasifikasikan secara salah sebagai Negatif (False Negative), mewakili 28,2%. Sebaliknya, dari 119 set data aktual yang diberi label Negatif, model tersebut mengidentifikasi 108 dengan benar sebagai Negatif (True

Negative), mewakili 90,8% dari total data Negatif, dan hanya 11 yang diklasifikasikan secara salah sebagai Positif (False Positive), mewakili 9,2%.

Evaluasi model klasifikasi dilakukan berdasarkan akurasi, presisi, recall, dan F1-score. Model Naive Bayes menunjukkan kinerja dengan akurasi 72,5% dan skor F1 tertimbang sebesar 0,72. Pada kelas Positif, model ini mencapai presisi 0,70, recall 0,80, dan skor F1 sebesar 0,75. Pada kelas Negatif, presisi mencapai 0,76, recall 0,65, dan skor F1 sebesar 0,70. Perbedaan antara skor pelatihan (0,864) dan skor pengujian (0,725) sebesar 0,139 menunjukkan tingkat overfitting yang relatif tinggi.

Sementara itu, Support Vector Machine (LinearSVC) menghasilkan kinerja terbaik dengan akurasi 74,2% dan skor F1 tertimbang sebesar 0,74. Pada kelas Positif, model ini mencapai presisi 0,71, recall 0,82, dan skor F1 0,76, sedangkan pada kelas Negatif, presisinya adalah 0,78, recall 0,67, dan skor F1 0,72. Perbedaan yang lebih kecil antara skor pelatihan (0,843) dan skor pengujian (0,742), yaitu 0,101, menunjukkan kemampuan generalisasi yang lebih baik daripada Naive Bayes. Keunggulan ini konsisten dengan karakteristik SVM, yang efektif dalam menangani data berdimensi tinggi dan jarang, seperti pada representasi TF-IDF.

Di sisi lain, IndoBERT mencapai akurasi 81,4% dengan skor F1 tertimbang 0,81, menempatkannya di antara SVM dan Naive Bayes. Meskipun IndoBERT secara teoritis memiliki keunggulan dalam memahami konteks melalui mekanisme self-attention dua arahnya, hasil fine-tuning dalam penelitian ini tidak mampu melampaui kinerja SVM. Hal ini kemungkinan dipengaruhi oleh ukuran dataset yang relatif terbatas (1.176 sampel) dan kualitas label yang diperoleh dari pelabelan otomatis, mengingat bahwa model berbasis transformer umumnya membutuhkan sejumlah besar data berlabel untuk mencapai kinerja optimal.

TABEL IX.
EVALUASI KINERJA PADA MODEL

Klasifikasi	NB	SVM	IndoBERT
Accuracy	72,5 %	74,2 %	81,4 %
Precision	72,9 %	74,6 %	82,5 %
Recall	72,5 %	74,2 %	81,4 %
F1-Score	72,3 %	74,0 %	81,2 %

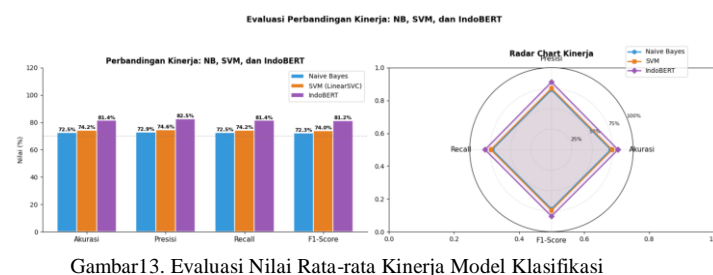
Gambar 13 menampilkan visualisasi komparatif kinerja tiga model klasifikasi: Naive Bayes, SVM (LinearSVC), dan IndoBERT, yang disajikan dalam dua bentuk: grafik batang dan grafik radar. Perbandingan ini didasarkan pada empat metrik evaluasi utama: akurasi, presisi, recall, dan skor F1.

Grafik batang di sebelah kiri menunjukkan bahwa IndoBERT secara konsisten menunjukkan kinerja tertinggi di semua metrik evaluasi. Model ini mencapai akurasi 81,4%, presisi 82,5%, recall 81,4%, dan skor F1 81,4%. Sementara itu, SVM (LinearSVC) berada di urutan kedua dengan akurasi 74,2%, presisi 74,6%, recall 74,2%, dan skor F1 74,0%. Di sisi lain, Naive Bayes mencatat kinerja terendah dengan

akurasi 72,5%, presisi 72,9%, recall 72,5%, dan skor F1 72,3%. Temuan ini menunjukkan perbedaan kinerja yang signifikan antara IndoBERT dan dua model konvensional, sementara perbedaan antara SVM dan Naive Bayes relatif kecil dan tidak menunjukkan signifikansi statistik berdasarkan uji peringkat bertanda Wilcoxon.

Visualisasi grafik radar di sisi kanan menunjukkan pola yang konsisten dengan grafik batang, di mana area yang dicakup oleh setiap model mewakili tingkat kinerjanya. IndoBERT membentuk area terbesar di dekat batas luar diagram, menunjukkan kinerja keseluruhan yang unggul. Sementara itu, SVM menempati posisi menengah, dan Naive Bayes membentuk area terkecil. Lebih lanjut, bentuk poligon dari ketiga model cenderung simetris, menunjukkan bahwa kinerja mereka relatif seimbang di keempat metrik evaluasi tanpa perbedaan signifikan dalam aspek tertentu.

Secara keseluruhan, kedua visualisasi tersebut secara konsisten mengkonfirmasi bahwa IndoBERT adalah model dengan kinerja terbaik dalam penelitian ini, diikuti oleh SVM dan Naive Bayes. Hasil ini juga memperkuat kemampuan pendekatan berbasis transformer untuk menghasilkan representasi semantik yang lebih kaya dan akurat untuk tugas klasifikasi sentimen teks berbahasa Indonesia dibandingkan dengan metode pembelajaran mesin berbasis TF-IDF konvensional.



Gambar 13. Evaluasi Nilai Rata-rata Kinerja Model Klasifikasi

H. Implikasi Kebijakan

Penelitian ini memiliki implikasi praktis yang penting bagi para pembuat kebijakan. Dominasi sentimen positif (62,8%) menunjukkan bahwa program MBG secara umum telah mendapat respons positif dari masyarakat. Namun, keberadaan kata "masalah" dalam lima fitur TF-IDF teratas, serta proporsi sentimen negatif yang masih signifikan (37,2%), menunjukkan kekhawatiran masyarakat tentang aspek teknis implementasi, seperti kualitas menu, ketepatan distribusi, dan transparansi anggaran.

Berdasarkan hal ini, pemerintah disarankan untuk mengembangkan sistem pemantauan sentimen berbasis media sosial secara berkala sebagai instrumen evaluasi yang lebih responsif. Lebih lanjut, penerapan analisis sentimen berbasis aspek dapat membantu mengidentifikasi kelompok keluhan masyarakat secara lebih spesifik. Upaya penting lainnya adalah memperkuat komunikasi publik mengenai mekanisme pemantauan dan pelaporan program MBG, sehingga kepercayaan masyarakat terhadap implementasinya dapat terus ditingkatkan.

IV. KESIMPULAN

Studi ini secara cermat mengevaluasi efektivitas tiga metode klasifikasi, yaitu Naive Bayes, Support Vector Machine (SVM), dan IndoBERT, dalam mengklasifikasikan sentimen publik terhadap program Makan Bergizi Gratis (MBG) di Platform X. Dataset terdiri dari 1.176 tweet dalam bahasa Indonesia yang diperoleh melalui web scraping dengan Selenium WebDriver dari 1 Januari hingga 31 Maret 2025. Pipeline persiapan untuk model klasik meliputi case folding, text cleaning, normalisasi, tokenizing, stopword removal menggunakan PySastrawi, dan stemming menggunakan PySastrawi, sedangkan IndoBERT menggunakan pipeline preprocessing yang lebih minimal guna mempertahankan konteks semantik teks.

Berdasarkan hasil pelabelan menggunakan pendekatan lexicon-based, diperoleh distribusi 742 tweet berlabel Positif (62,8%) dan 439 tweet berlabel Negatif (37,2%), yang mengindikasikan bahwa secara umum masyarakat menyambut program MBG dengan respons yang lebih positif. Hasil pembobotan TF-IDF menunjukkan bahwa kata "anak", "sekolah", "siang", "menu", dan "masalah" merupakan kata-kata dengan bobot tertinggi, yang mencerminkan bahwa diskusi publik banyak berpusat pada sasaran utama program sekaligus kekhawatiran terhadap implementasinya.

Hasil evaluasi menunjukkan bahwa IndoBERT mencapai akurasi tertinggi sebesar 81,4% dengan F1-score 0,81, diikuti oleh SVM (LinearSVC) dengan akurasi 74,2%, presisi 0,75, recall 0,74, dan F1-score 0,74, serta Naive Bayes dengan akurasi 72,5%, presisi 0,73, recall 0,72, dan F1-score 0,72. Pemeriksaan confusion matrix menunjukkan bahwa SVM memiliki kemampuan generalisasi yang lebih baik dibandingkan Naive Bayes, sebagaimana tercermin dari selisih train score dan test score yang lebih kecil (0,101 berbanding 0,139). Keunggulan SVM atas Naive Bayes muncul dari kemampuannya mengelola data teks berdimensi tinggi dengan membangun hyperplane ideal bermargin maksimum, sehingga meningkatkan ketahanannya dalam mengkategorikan data sentimen berbahasa Indonesia. Sementara itu, keunggulan IndoBERT berasal dari kemampuan pemahaman konteks bidirectional melalui mekanisme self-attention yang memungkinkan model menangkap nuansa linguistik teks media sosial secara lebih mendalam.

Penelitian lebih lanjut disarankan untuk memperluas cakupan dataset di berbagai platform media sosial, melakukan validasi pelabelan melalui anotasi manual, serta mengeksplorasi teknik augmentasi data guna mengoptimalkan performa IndoBERT, sekaligus mengembangkan analisis ke arah aspect-based sentiment analysis untuk memperoleh pemahaman yang lebih holistik terhadap dinamika opini publik terkait program MBG.

DAFTAR PUSTAKA

- [1] M. D. Anggraeni, A. C. Yudiananta, dan A. H. Arifin, "Analisis Sentimen Masyarakat terhadap Permasalahan Keracunan Program MBG," *Jurnal Transformasi*, vol. 21, no. 2, pp. 77–89, 2025..
- [2] A. Kiftiyah, F. A. Palestina, F. U. Abshar, dan K. Rofiah, "Program Makan Bergizi Gratis (MBG) dalam Perspektif Keadilan Sosial dan Dinamika Sosial-Politik," *PANCASILA: Jurnal Keindonesiaan*, vol. 5, no. 1, Apr. 2025, doi: 10.52738/pjk.v5i1.726.
- [3] F. Wajidi and M. R. Rasyid, "Evaluasi algoritma KNN dan Naive Bayes untuk analisis sentimen kebijakan program makan bergizi gratis," vol. 7, no. 2, pp. 83–97, 2025, doi: 10.37905/jji.v1i2.34418.
- [4] V. Alviani, S. Alam, I. Kurniawan, "Analisis Sentimen Review Aplikasi Wetv Pada Platform Twitter Menggunakan Support Vector Machine," vol. 2, no. 3, pp. 143–149, 2023.
- [5] A. Ardiansyah, E. A. Pratama, and N. I. Fadlilah, "Analisis Sentimen Pengguna Terhadap Aplikasi ChatGPT Di Google Play Store : Penerapan Algoritma Support Vector Machine," vol. 11, no. 2, pp. 247–254, 2024.
- [6] M. M. Laia *et al.*, "Analisis Sentimen Program Makan Gratis Pada Platform X Algoritma Naive Bayes Menggunakan," 2025.
- [7] F. Fatkhurrohman, B. I. Nugroho, and N. Fadillah, "Analisis Sentimen Program Makan Bergizi Gratis Pemerintah RI Melalui Twitter Menggunakan Metode SVM," vol. 4, no. 3, pp. 3906–3917, 2025.
- [8] Anwar and I. Rahmawati, "Analisis Sentimen terhadap Kebijakan Publik Menggunakan Metode Naive Bayes," *Jurnal Teknologi Informasi dan Komunikasi*, vol. 14, no. 2, pp. 101–112, 2022.
- [9] M. Apriliyani, M. I. Musyaffaq, and S. N. Aini, "Implementasi analisis sentimen pada ulasan aplikasi Duolingo di Google Playstore menggunakan algoritma Naive Bayes," vol. 21, no. 2, pp. 298–311, 2024.
- [10] T. Grace, W. Margaretha, D. Juardi, "Analisis Sentimen Masyarakat Terhadap Penggunaan Halodoc Sebagai Layanan Telemedicine Di Indonesia," vol. 13, no. 1, 2025.
- [11] Vinne. D. U. Sinurat, Y. Prasetyanti, Y., "Penerapan algoritma k-nearest neighbors (knn) dalam menganalisis sentimen ulasan aplikasi seabank pada google play store," vol. 2, no. 2, pp. 103–113, 2025.
- [12] J. Pardede, and D. Darmawan, "Perbandingan Algoritma Stemming Porter , Sastrawi , Idris , Comparison Of Stemming Algorithms Porter , Sastrawi , Idris , And Arifin Setiono On Indonesian Text Documents," vol. 12, no. 1, 2025, doi: 10.25126/jtiik.2025128860.
- [13] D. Septiani, I. Isabela, "Analisis term frequency inverse document frequency (tf-idf) dalam temu kembali informasi pada dokumen teks," vol. 25, pp. 81–88.
- [14] A. Damuri, U. Riyanto, H. Rusdianto, and M. Aminudin, "Implementasi Data Mining dengan Algoritma Naive Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako," vol. 8, no. 6, pp. 219–225, 2021, doi: 10.30865/jurikom.v8i6.3655.
- [15] H. Apriyani, "Perbandingan Metode Naive Bayes Dan Support Vector Machine Dalam Klasifikasi Penyakit Diabetes Melitus," vol. 1, no. 3, pp. 133–143, 2020.
- [16] R. Z. Alobaidy, G. A. Altalib, and S. Zainab, "Comparative Study of Opinion Mining and Sentiment Analysis: Algorithms and Applications," vol. 8, no. 4, pp. 12–20, 2020.
- [17] [1] A. Z. Muhabbab, Bunyamin, and Hasmawati, "Sentiment Analysis of Indonesia's Free Nutritious Meal Program on Platform X (Formerly Twitter) Using IndoBERT," *Zero: Jurnal Sains, Matematika, dan Terapan*, vol. 9, no. 3, pp. 884–893, 2025, doi: 10.30829/zero.v9i3.27629..
- [18] M. T. Nugraha, N. Sulistiyowati, dan U. Enri, "Analisis Sentimen Ulasan Aplikasi Satu Sehat pada Google Play Store Menggunakan Naive Bayes Classifier," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 5, pp. 3593–3601, Okt. 2023.
- [19] A. M. Ndapamuri, D. Manongga, A. Iriani, "Analisis Sentimen Ulasan Aplikasi Tripadvisor Dengan Metode Support Vector Machine , K-Nearest Neighbor , Dan Naive Bayes," pp. 127–140, 2023.
- [20] M. I. Fikri, T. S. Sabrila, and Y. Azhar, "Perbandingan Metode Naive Bayes dan Support Vector Machine pada Analisis Sentimen Twitter," vol. 10, pp. 71–76, 2020.