

# Sentiment Analysis of Free Nutritious Meal Program on Platform X: Comparing Naive Bayes, SVM, Random Forest, and IndoBERT

Alrijal Nur Ilham <sup>1\*</sup>, Etika Kartikadarma <sup>2\*</sup>

\* Teknik Informatika, Universitas Dian Nuswantoro

[111202214113@mhs.dinus.ac.id](mailto:111202214113@mhs.dinus.ac.id)<sup>1</sup>, [etika.kartikadarma@dsn.dinus.ac.id](mailto:etika.kartikadarma@dsn.dinus.ac.id)<sup>2</sup>

## Article Info

### Article history:

Received 2026-05-08

Revised 2026-05-19

Accepted 2026-05-25

### Keyword:

*Sentiment Analysis,  
TF-IDF,  
MBG Program,  
Social Media.*

## ABSTRACT

The Free Nutritious Meal Program (MBG), launched by the Indonesian government in January 2025, generated various public responses on social media, particularly on platform X. This study aims to analyze public sentiment toward the MBG Program and compare the performance of four sentiment classification methods: Naive Bayes, Support Vector Machine (SVM), Random Forest, and IndoBERT. The dataset was collected through tweet crawling using the keywords “MBG” and “Makan Bergizi Gratis” during the period of July–December 2025, resulting in 1,906 Indonesian-language tweets. The preprocessing stage included cleaning, case folding, normalization, tokenization, stopword removal, and stemming. Sentiment labeling was performed using the InSet Lexicon and produced 1,113 negative tweets and 793 positive tweets. Manual validation on part of the dataset was conducted by two independent annotators and achieved a Cohen’s Kappa score of 0.78, indicating substantial agreement. For classical machine learning models, feature extraction was carried out using TF-IDF, while IndoBERT used contextual text representations without stemming. Class imbalance in classical models was handled using SMOTE, whereas IndoBERT applied class weighting. The experimental results show that IndoBERT achieved the best performance with an accuracy of 92.93%. Among the classical models, SVM produced the highest performance with an accuracy of 92.15%, followed by Naive Bayes and Random Forest. Word frequency analysis also revealed that positive sentiment was mainly associated with support for the program and nutrition-related topics, while negative sentiment was dominated by concerns about food safety, budget management, and criticism of the program. Based on the findings, IndoBERT is more effective in understanding the context of Indonesian-language tweets. However, TF-IDF-based classical models, especially SVM, still provide competitive performance with lower computational requirements, making them suitable for sentiment analysis in public policy studies.



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

## I. PENDAHULUAN

Perkembangan media sosial yang sangat pesat telah mengubah cara masyarakat dalam menyampaikan opini, kritik, maupun dukungan terhadap berbagai isu public [1]. Platform media sosial seperti X menjadi ruang digital yang aktif digunakan masyarakat untuk berdiskusi mengenai kebijakan pemerintah, kondisi sosial, hingga berbagai peristiwa yang sedang terjadi [2]. Di Indonesia, tingginya jumlah pengguna media sosial menjadikan platform tersebut

sebagai sumber data yang relevan untuk memahami persepsi dan sentimen publik terhadap suatu kebijakan [3]. Dengan memanfaatkan teknik Natural Language Processing (NLP) dan analisis sentimen, opini masyarakat yang tersebar dalam jumlah besar dapat dianalisis secara otomatis untuk mengetahui kecenderungan respons publik terhadap suatu topik tertentu [2], [3].

Salah satu kebijakan pemerintah yang banyak menjadi perhatian publik pada tahun 2025 adalah Program Makan Bergizi Gratis (MBG). Program ini diluncurkan sebagai salah

satu upaya pemerintah dalam meningkatkan kualitas gizi masyarakat serta mendukung pembangunan sumber daya manusia menuju visi Indonesia Emas 2045 [4]. Program MBG ditujukan kepada siswa sekolah, ibu hamil, dan ibu menyusui melalui penyediaan makanan bergizi yang didanai oleh anggaran negara. Besarnya anggaran yang digunakan serta luasnya cakupan penerima manfaat menyebabkan program ini menjadi topik yang sering dibahas di media sosial, khususnya pada platform X [5].

Respons masyarakat terhadap Program MBG menunjukkan beragam sentimen. Sebagian masyarakat memberikan dukungan karena program ini dinilai dapat membantu pemenuhan gizi anak dan meningkatkan kualitas kesehatan masyarakat [6]. Namun, tidak sedikit pula pengguna media sosial yang menyampaikan kritik terkait pengelolaan anggaran, efektivitas pelaksanaan program, hingga isu keamanan pangan yang muncul selama implementasi program berlangsung. Perbedaan opini tersebut menjadikan analisis sentimen penting dilakukan untuk mengetahui kecenderungan pandangan masyarakat terhadap Program MBG secara lebih sistematis [7].

Analisis sentimen merupakan salah satu cabang dari text mining yang digunakan untuk mengidentifikasi opini atau emosi dari suatu teks ke dalam kategori tertentu, seperti positif atau negatif. Pada data media sosial, proses analisis sentimen memiliki tantangan tersendiri karena tweet umumnya menggunakan bahasa informal, singkatan, bahasa gaul, serta terkadang mengandung sarkasme dan campuran bahasa. Kondisi tersebut dapat mempengaruhi performa model klasifikasi sentimen sehingga diperlukan metode yang mampu memahami konteks bahasa dengan baik [8], [9].

Berbagai penelitian sebelumnya telah menggunakan metode machine learning untuk melakukan analisis sentimen pada data media sosial berbahasa Indonesia. Algoritma klasik seperti Naive Bayes, Support Vector Machine (SVM), dan Random Forest masih banyak digunakan karena relatif sederhana, efisien, dan memiliki performa yang cukup baik pada klasifikasi teks berbasis TF-IDF [10]. Di sisi lain, perkembangan model deep learning berbasis transformer seperti IndoBERT menunjukkan kemampuan yang lebih baik dalam memahami konteks kalimat karena menggunakan representasi bahasa secara kontekstual [11]. Meskipun demikian, model transformer umumnya membutuhkan sumber daya komputasi yang lebih besar dibandingkan metode klasik [12].

Beberapa penelitian terkait Program MBG sebelumnya telah dilakukan dengan menggunakan model klasifikasi tertentu, baik model machine learning klasik maupun model transformer. Namun, sebagian besar penelitian masih berfokus pada penggunaan satu atau dua model saja, sehingga perbandingan performa antar metode pada dataset yang sama masih terbatas. Selain itu, beberapa penelitian belum menjelaskan secara rinci proses preprocessing, penanganan ketidakseimbangan data, maupun analisis kesalahan klasifikasi yang terjadi pada model [13]. Oleh karena itu, diperlukan penelitian yang dapat membandingkan beberapa

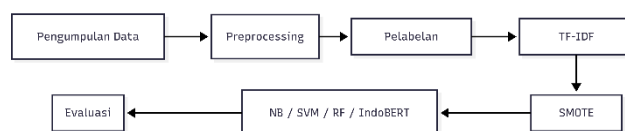
metode klasifikasi secara langsung menggunakan dataset yang sama agar hasil evaluasi menjadi lebih objektif dan terukur [14].

Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap Program Makan Bergizi Gratis (MBG) pada platform X serta membandingkan performa empat metode klasifikasi, yaitu Naive Bayes, Support Vector Machine (SVM), Random Forest, dan IndoBERT. Model Naive Bayes, SVM, dan Random Forest menggunakan fitur TF-IDF, sedangkan IndoBERT menggunakan representasi kontekstual berbasis transformer [15], [16]. Selain membandingkan performa model, penelitian ini juga menganalisis distribusi sentimen serta kata-kata yang dominan pada masing-masing kelas sentimen untuk memperoleh gambaran mengenai opini publik terhadap Program MBG.

Hasil penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan analisis sentimen berbahasa Indonesia, khususnya pada topik kebijakan publik. Selain itu, penelitian ini juga diharapkan dapat menjadi referensi dalam menentukan metode klasifikasi sentimen yang efektif dengan mempertimbangkan performa model dan kebutuhan sumber daya komputasi.

## II. METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis sentimen berbasis machine learning terhadap data tweet berbahasa Indonesia yang berkaitan dengan Program Makan Bergizi Gratis (MBG). Tahapan penelitian terdiri dari pengumpulan data, preprocessing teks, pelabelan sentimen, ekstraksi fitur, penanganan ketidakseimbangan data, proses klasifikasi, dan evaluasi model. Alur penelitian dilakukan secara bertahap agar proses analisis dapat berjalan secara sistematis dan menghasilkan evaluasi yang objektif terhadap setiap model klasifikasi yang digunakan.



Gambar 1. Tahapan Penelitian

Gambar 1 menunjukkan tahapan penelitian yang dilakukan dalam analisis sentimen Program Makan Bergizi Gratis (MBG). Penelitian diawali dengan proses pengumpulan data tweet dari platform X menggunakan metode crawling berdasarkan kata kunci “MBG” dan “Makan Bergizi Gratis”. Data yang diperoleh kemudian melalui tahap preprocessing yang meliputi cleaning, case folding, normalisasi kata, tokenisasi, stopword removal, dan stemming.

Setelah preprocessing, dilakukan pelabelan sentimen menggunakan InSet Lexicon untuk mengelompokkan tweet ke dalam kelas positif dan negatif. Dataset selanjutnya dibagi menjadi data latih dan data uji menggunakan metode stratified split. Pada model machine learning klasik, teks diubah menjadi representasi numerik menggunakan TF-IDF dan dilakukan penanganan ketidakseimbangan data menggunakan SMOTE pada data latih. Data kemudian digunakan untuk

proses klasifikasi menggunakan algoritma Naive Bayes, Support Vector Machine (SVM), dan Random Forest [17].

Sementara itu, pada model IndoBERT, data teks yang telah dipreprocessing digunakan langsung pada proses fine-tuning tanpa tahap stemming dan TF-IDF karena IndoBERT menggunakan representasi bahasa secara kontekstual [18]. Seluruh model kemudian dievaluasi menggunakan confusion matrix, accuracy, precision, recall, dan F1-score. Tahap akhir penelitian dilakukan melalui analisis hasil klasifikasi dan visualisasi wordcloud untuk mengetahui pola kata dominan pada masing-masing sentiment [13].

#### A. Pengumpulan Data

Data penelitian diperoleh melalui proses crawling dari platform X menggunakan tools *tweet-harvest* berbasis Node.js dan Python. Pengambilan data dilakukan menggunakan kata kunci “MBG” dan “Makan Bergizi Gratis” selama periode Juli hingga Desember 2025. Pemilihan periode tersebut dilakukan untuk menangkap dinamika opini publik selama masa implementasi program dan berbagai isu yang berkembang di media sosial.

Hasil crawling menghasilkan sebanyak 1.906 tweet berbahasa Indonesia yang kemudian disimpan dalam format *.xlsx*. Dataset yang diperoleh memuat beberapa atribut utama seperti isi tweet (*full\_text*), waktu publikasi (*created\_at*), dan label sentimen. Sebelum digunakan pada tahap analisis, data terlebih dahulu diperiksa untuk mengurangi tweet duplikat dan tweet yang tidak relevan dengan topik penelitian.

#### B. Preprocessing Teks

Tahap preprocessing dilakukan untuk membersihkan dan menormalisasi data teks agar lebih mudah diproses oleh model klasifikasi. Preprocessing menjadi tahap penting karena data media sosial umumnya mengandung banyak noise seperti simbol, singkatan, bahasa gaul, dan karakter yang tidak diperlukan.

Tahapan preprocessing yang digunakan dalam penelitian ini meliputi:

TABEL I  
TAHAPAN PREPROCESSING TEKS

Tahap	Keterangan
Cleaning	Menghapus URL, mention, hashtag, angka, tanda baca, emoji, dan karakter non-alfanumerik yang tidak diperlukan dalam proses klasifikasi.
Case Folding	Mengubah seluruh huruf menjadi huruf kecil untuk menjaga konsistensi data teks.
Normalisasi	Mengubah kata tidak baku, singkatan, dan bahasa gaul menjadi bentuk kata baku menggunakan kamus normalisasi Bahasa Indonesia.
Tokenizing	Memecah kalimat menjadi token-token kata individual agar dapat diproses lebih lanjut.
Stopword Removal	Menghapus kata-kata umum yang tidak memiliki pengaruh signifikan terhadap sentimen, seperti “dan”, “yang”, atau “di”.

Stemming	Mengubah kata menjadi bentuk dasar menggunakan pustaka Sastrawi. Tahap stemming diterapkan pada model machine learning klasik, sedangkan IndoBERT menggunakan teks asli tanpa stemming karena model telah dilatih menggunakan struktur bahasa alami Bahasa Indonesia.
----------	---

Tabel 1 Tahapan preprocessing dilakukan untuk membersihkan dan menstandarkan data teks sebelum proses klasifikasi dilakukan. Proses ini bertujuan mengurangi noise pada data media sosial sehingga model dapat mempelajari pola sentimen dengan lebih optimal. Pada penelitian ini, preprocessing meliputi cleaning, case folding, normalisasi, tokenizing, stopword removal, dan stemming. Namun, tahap stemming hanya diterapkan pada model machine learning klasik berbasis TF-IDF, sedangkan IndoBERT menggunakan teks asli tanpa stemming untuk mempertahankan konteks bahasa alami [15].

#### C. Pelabelan Sentimen

Pelabelan sentimen dilakukan menggunakan pendekatan lexicon-based dengan memanfaatkan InSet Lexicon. Setiap kata dalam tweet dicocokkan dengan skor polaritas yang terdapat pada kamus sentimen Bahasa Indonesia. Tweet dengan total skor positif diberi label positif, sedangkan tweet dengan total skor negatif diberi label negatif [19].

Hasil pelabelan menghasilkan 1.113 tweet negatif dan 793 tweet positif. Untuk memastikan kualitas label, dilakukan validasi manual terhadap sebagian data menggunakan dua anotator independen. Tingkat kesepakatan anotator diukur menggunakan Cohen’s Kappa dan menghasilkan nilai sebesar 0,78 yang menunjukkan tingkat kesepakatan yang baik.

#### D. Ekstraksi Fitur

Pada model machine learning klasik, representasi numerik teks dilakukan menggunakan metode TF-IDF (*Term Frequency–Inverse Document Frequency*). TF-IDF digunakan untuk mengubah data teks menjadi vektor numerik berdasarkan frekuensi kemunculan kata dan tingkat kepentingannya dalam dokumen [12].

Parameter TF-IDF yang digunakan dalam penelitian ini meliputi:

- `max_features = 5000`
- `ngram_range = (1,2)`
- `min_df = 2`

Proses fitting TF-IDF hanya dilakukan pada data latih untuk menghindari terjadinya data leakage.

#### E. Penanganan Ketidakseimbangan Data

Distribusi data sentimen pada dataset penelitian tidak seimbang karena jumlah tweet negatif lebih banyak dibandingkan tweet positif. Untuk mengatasi permasalahan tersebut, digunakan metode SMOTE (Synthetic Minority Over-sampling Technique) pada model machine learning klasik. SMOTE bekerja dengan membuat sampel sintesis pada

kelas minoritas sehingga distribusi data menjadi lebih seimbang [20].

Penerapan SMOTE hanya dilakukan pada data latih agar data uji tetap merepresentasikan kondisi data asli. Sementara itu, pada model IndoBERT penanganan ketidakseimbangan data dilakukan menggunakan class weighting untuk memberikan bobot lebih besar pada kelas minoritas selama proses pelatihan model.

#### F. Model Klasifikasi

Penelitian ini membandingkan empat metode klasifikasi sentimen, yaitu Naive Bayes, Support Vector Machine (SVM), Random Forest, dan IndoBERT.

1) *Naive Bayes*: Naive Bayes merupakan algoritma probabilistik yang sering digunakan pada klasifikasi teks karena sederhana dan efisien dalam memproses data berdimensi tinggi. Dengan rumus sebagai berikut:

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)} \quad (2)$$

di mana  $P(C|X)$  adalah probabilitas kelas C diberikan fitur X,  $P(X|C)$  adalah likelihood fitur,  $P(C)$  adalah probabilitas prior kelas, dan  $P(X)$  adalah probabilitas marginal fitur.

2) *SVM*: SVM digunakan karena memiliki performa yang baik pada klasifikasi teks berbasis TF-IDF dan mampu memisahkan data menggunakan hyperplane optimal. Dengan rumus sebagai berikut:

$$f(x) = w^T x + b \quad (3)$$

di mana  $w$  adalah vektor bobot,  $x$  adalah vektor fitur TF-IDF, dan  $b$  adalah bias. Tanda dari  $f(x)$  menentukan kelas prediksi.

3) *Random Forest*: Random Forest merupakan metode ensemble berbasis decision tree yang bekerja menggunakan mekanisme voting untuk menentukan hasil klasifikasi. Dengan rumus sebagai berikut:

$$\{y\} = \text{mode}(h_1(x), h_2(x), \dots, h_n(x)) \quad (4)$$

di mana  $h_i(x)$  adalah prediksi decision tree ke- $i$ , dan  $\hat{y}$  adalah hasil voting mayoritas dari seluruh pohon.

4) *IndoBERT*: IndoBERT merupakan model transformer yang telah dilatih menggunakan korpus Bahasa Indonesia sehingga mampu memahami konteks kalimat secara lebih baik dibandingkan metode klasik. Pada penelitian ini, IndoBERT digunakan melalui proses fine-tuning terhadap dataset sentimen MBG.

#### G. Evaluasi Model

Seluruh model dievaluasi menggunakan data uji yang sama agar hasil perbandingan lebih objektif. Evaluasi dilakukan menggunakan confusion matrix dan beberapa metrik evaluasi, yaitu accuracy, precision, recall, dan F1-score.

Accuracy digunakan untuk mengukur tingkat ketepatan prediksi model secara keseluruhan. Precision menunjukkan

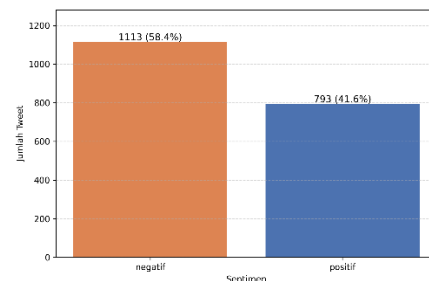
kemampuan model dalam menghasilkan prediksi positif yang benar, sedangkan recall mengukur kemampuan model dalam mendeteksi seluruh data positif. F1-score digunakan sebagai metrik utama karena mampu menyeimbangkan nilai precision dan recall, terutama pada kondisi data yang tidak seimbang [21].

### III. HASIL DAN PEMBAHASAN

#### A. Distribusi Dataset

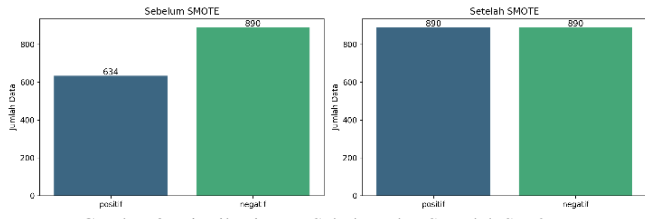
Dataset pada penelitian ini diperoleh melalui proses crawling dari platform X menggunakan kata kunci “MBG” dan “Makan Bergizi Gratis” selama periode Juli hingga Desember 2025. Setelah dilakukan proses preprocessing dan pelabelan sentimen menggunakan InSet Lexicon, diperoleh total 1.906 tweet berbahasa Indonesia yang terdiri dari 1.113 tweet negatif dan 793 tweet positif.

Distribusi dataset menunjukkan bahwa sentimen negatif lebih mendominasi dibandingkan sentimen positif. Dominasi sentimen negatif menunjukkan bahwa pembahasan terkait Program Makan Bergizi Gratis (MBG) di media sosial selama periode penelitian lebih banyak dipengaruhi oleh kritik, kekhawatiran, dan isu kontroversial dibandingkan dukungan terhadap program. Kondisi tersebut dapat dipengaruhi oleh karakteristik media sosial yang cenderung lebih aktif dalam menyebarkan opini kritis dan isu viral dibandingkan informasi positif atau netral. Selain itu, beberapa isu yang berkaitan dengan keamanan pangan, pengelolaan anggaran, dan implementasi program turut mempengaruhi tingginya jumlah sentimen negatif pada dataset penelitian.



Gambar 2. Distribusi Sentimen Dataset MBG

Berdasarkan Gambar 2, distribusi sentimen pada dataset MBG menunjukkan kondisi yang tidak seimbang, di mana jumlah sentimen negatif lebih dominan dibandingkan sentimen positif. Ketidakseimbangan distribusi kelas tersebut berpotensi menyebabkan model klasifikasi menjadi bias terhadap kelas mayoritas. Pada kondisi tersebut, model cenderung lebih sering memprediksi kelas negatif dibandingkan kelas positif. Oleh karena itu, penelitian ini menerapkan metode SMOTE pada data latih untuk meningkatkan keseimbangan distribusi kelas pada model machine learning klasik.



Gambar 3. Distribusi Data Sebelum dan Sesudah SMOTE

Berdasarkan Gambar 3, penerapan SMOTE berhasil meningkatkan keseimbangan distribusi kelas pada data latih. Setelah proses oversampling dilakukan, jumlah data pada masing-masing kelas menjadi lebih seimbang sehingga model dapat mempelajari pola sentimen positif dan negatif secara lebih optimal.

Meskipun SMOTE membantu meningkatkan keseimbangan data, penggunaan metode ini pada data teks berbasis TF-IDF tetap memiliki keterbatasan. Sampel sintesis yang dihasilkan tidak sepenuhnya merepresentasikan struktur bahasa alami karena dibentuk berdasarkan interpolasi vektor fitur. Oleh karena itu, evaluasi model tetap dilakukan menggunakan data uji asli untuk menjaga validitas hasil eksperimen dan mengurangi risiko overfitting terhadap data sintesis [20].

**B. Analisis Word Frequency dan Wordcloud**

Analisis word frequency dilakukan untuk mengetahui kata-kata yang paling dominan pada masing-masing kelas sentimen. Selain itu, visualisasi wordcloud digunakan untuk memberikan gambaran umum mengenai topik yang paling sering dibahas pengguna media sosial terkait Program MBG. Analisis ini membantu memahami kecenderungan opini publik terhadap program secara lebih mendalam [22].



Gambar 4. Wordcloud Sentimen Positif

Pada Gambar 4 sentimen positif, kata-kata yang paling dominan antara lain “gizi”, “makan”, “gratis”, “sehat”, “dukung”, dan “generasi”. Kemunculan kata-kata tersebut menunjukkan bahwa sebagian masyarakat menilai Program MBG sebagai kebijakan yang bermanfaat dalam meningkatkan kualitas gizi masyarakat, khususnya bagi anak-anak sekolah. Kata “generasi” dan “masa depan” juga menunjukkan adanya harapan publik terhadap dampak jangka panjang program terhadap pembangunan sumber daya manusia Indonesia.

Selain itu, banyak tweet positif yang mengaitkan Program MBG dengan peningkatan kesejahteraan masyarakat dan pemerataan akses makanan bergizi. Hal ini menunjukkan bahwa sebagian pengguna media sosial melihat program tersebut sebagai bentuk perhatian pemerintah terhadap kesehatan masyarakat dan pendidikan generasi muda.

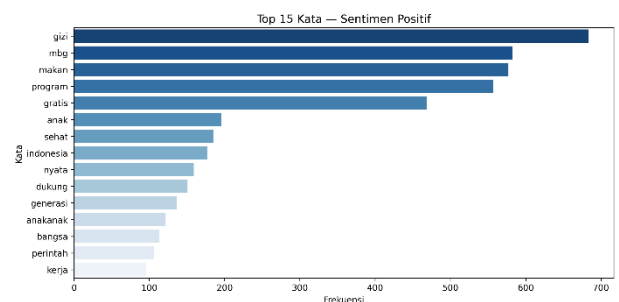


Gambar 5. Wordcloud Sentimen Negatif

Untuk Gambar 5 wordcloud sentimen negatif didominasi oleh kata-kata seperti “racun”, “korban”, “korupsi”, “dana”, “proyek”, “anggaran”, dan “henti”. Dominasi kata-kata tersebut menunjukkan bahwa opini negatif publik lebih banyak berkaitan dengan isu keamanan pangan dan pengelolaan program. Kemunculan kata “racun” dan “korban” dipengaruhi oleh adanya pembahasan terkait kasus makanan bermasalah yang dikaitkan dengan program MBG di media sosial.

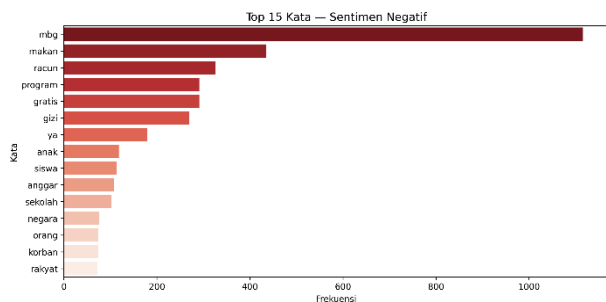
Di sisi lain, kata-kata seperti “korupsi”, “anggaran”, dan “dana” menunjukkan adanya kekhawatiran masyarakat terhadap transparansi pengelolaan anggaran program. Hal ini menunjukkan bahwa opini publik terhadap kebijakan pemerintah tidak hanya dipengaruhi oleh manfaat program, tetapi juga oleh tingkat kepercayaan masyarakat terhadap pengelolaan kebijakan tersebut.

Untuk memperjelas pola kata dominan pada masing-masing sentimen, dilakukan analisis top 15 kata dengan frekuensi tertinggi pada setiap kelas sentimen.



Gambar 6. Top 15 Kata Dominan pada Sentimen Positif

Hasil analisis menunjukkan bahwa kata-kata dominan pada Gambar 6 sentimen positif cenderung berkaitan dengan dukungan terhadap program dan isu kesehatan Masyarakat.



Gambar 7. Top 15 Kata Dominan pada Sentimen Negatif

Gambar 7 sentimen negatif lebih banyak berkaitan dengan kritik terhadap implementasi program. Perbedaan pola kata tersebut menunjukkan bahwa opini publik terhadap Program MBG dipengaruhi oleh berbagai faktor sosial dan isu yang berkembang selama periode pengambilan data.

### C. Hasil Klasifikasi Model

Penelitian ini membandingkan performa empat model klasifikasi sentimen, yaitu Naive Bayes, Support Vector Machine (SVM), Random Forest, dan IndoBERT. Seluruh model dievaluasi menggunakan data uji yang sama untuk menjaga objektivitas hasil eksperimen. Evaluasi dilakukan menggunakan metrik accuracy, precision, recall, dan F1-score.

TABEL II  
PERBANDINGAN PERFORMA MODEL KLASIFIKASI

Metrik	Naive Bayes	SVM	Random Forest	IndoBERT
Accuracy	92.15%	92.15%	90.31%	92.93%
Precision Negatif	0.95	0.93	0.90	0.96
Recall Negatif	0.91	0.93	0.93	0.92
F1-score Negatif	0.93	0.93	0.92	0.94
Precision Positif	0.88	0.91	0.90	0.89
Recall Positif	0.94	0.91	0.86	0.94
F1-score Positif	0.91	0.91	0.88	0.92

Berdasarkan Tabel 2 hasil pengujian, IndoBERT memperoleh performa terbaik dengan akurasi sebesar 92,93%. Hasil tersebut menunjukkan bahwa model transformer memiliki kemampuan yang lebih baik dalam memahami konteks bahasa pada tweet berbahasa Indonesia. Kemampuan IndoBERT dalam mempelajari hubungan kontekstual antar kata membantu model mengenali pola sentimen pada teks informal yang umum ditemukan di media sosial.

Meskipun demikian, perbedaan performa antara IndoBERT dan SVM tidak terlalu besar. Hal ini menunjukkan bahwa model machine learning klasik berbasis TF-IDF masih

mampu memberikan hasil yang kompetitif untuk tugas analisis sentimen berbahasa Indonesia.

1) *Analisis Model Naive Bayes:* Naive Bayes menghasilkan akurasi sebesar 92,15% dengan performa yang cukup baik pada kedua kelas sentimen. Tingginya precision pada kelas negatif menunjukkan bahwa model cukup efektif dalam mengidentifikasi tweet negatif dengan tingkat kesalahan prediksi yang rendah. Performa Naive Bayes dipengaruhi oleh karakteristik algoritma yang bekerja berdasarkan probabilitas kemunculan kata pada masing-masing kelas sentimen. Pada data teks berbasis TF-IDF, Naive Bayes mampu bekerja secara efisien karena proses klasifikasi dilakukan berdasarkan distribusi token yang dominan pada setiap kelas. Namun, model ini masih memiliki keterbatasan dalam memahami hubungan kontekstual antar kata karena menggunakan asumsi independensi fitur. Akibatnya, beberapa tweet yang mengandung opini campuran atau struktur kalimat kompleks masih sulit diklasifikasikan dengan tepat oleh model.

2) *Analisis Model Support Vector Machine (SVM):* SVM memperoleh akurasi sebesar 92,15% dan menjadi model klasik dengan performa paling kompetitif pada penelitian ini. Nilai precision dan recall yang relatif seimbang menunjukkan bahwa model mampu mengklasifikasikan kedua kelas sentimen secara stabil tanpa terlalu bias terhadap salah satu kelas. Performa SVM yang tinggi dipengaruhi oleh kemampuannya dalam menangani data berdimensi tinggi dan sparse seperti representasi TF-IDF. Algoritma SVM bekerja dengan mencari hyperplane optimal yang mampu memisahkan kelas positif dan negatif secara maksimal. Oleh karena itu, SVM sering menghasilkan performa yang stabil pada tugas klasifikasi teks. Selain memiliki performa yang tinggi, SVM juga membutuhkan sumber daya komputasi yang relatif lebih ringan dibandingkan model transformer. Hal ini menjadikan SVM sebagai alternatif yang efisien untuk penelitian analisis sentimen dengan keterbatasan perangkat keras.

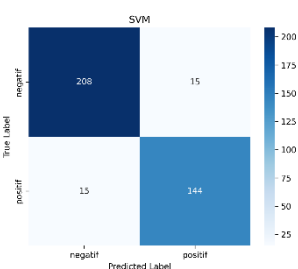
3) *Analisis Model Random Forest:* Random Forest memperoleh akurasi sebesar 90,31%, lebih rendah dibandingkan model lainnya. Meskipun demikian, model masih mampu menghasilkan performa yang cukup baik pada kedua kelas sentimen. Penurunan performa Random Forest dipengaruhi oleh karakteristik data TF-IDF yang memiliki dimensi sangat tinggi dan sparse. Pendekatan ensemble tree pada Random Forest kurang optimal dalam menangani representasi fitur teks dibandingkan model linear seperti SVM. Selain itu, recall pada kelas positif yang lebih rendah menunjukkan bahwa model masih mengalami kesulitan dalam mendeteksi seluruh tweet positif secara optimal. Walaupun memiliki performa lebih rendah dibandingkan model lain, Random Forest tetap menunjukkan kemampuan klasifikasi yang cukup stabil dan mampu mengurangi risiko overfitting melalui mekanisme ensemble learning.

4) *Analisis Model IndoBERT*: IndoBERT memperoleh performa terbaik di antara seluruh model yang diuji. Tingginya nilai F1-score menunjukkan bahwa model mampu menghasilkan keseimbangan yang baik antara precision dan recall pada kedua kelas sentimen. Keunggulan IndoBERT dipengaruhi oleh arsitektur transformer yang mampu memahami hubungan kontekstual antar kata dalam suatu kalimat. Berbeda dengan model berbasis TF-IDF yang hanya bergantung pada frekuensi kata, IndoBERT mampu memahami makna kata berdasarkan konteks penggunaannya. Hal tersebut membantu model dalam menangani bahasa informal, singkatan, dan variasi struktur kalimat yang umum ditemukan pada media sosial. Meskipun demikian, IndoBERT masih mengalami kesalahan klasifikasi pada beberapa tweet yang mengandung sarkasme, opini ambigu, dan campuran bahasa. Selain itu, proses pelatihan IndoBERT membutuhkan waktu dan sumber daya komputasi yang lebih besar dibandingkan model klasik.

*D. Analisis Confusion Matrix dan Error Analysis*

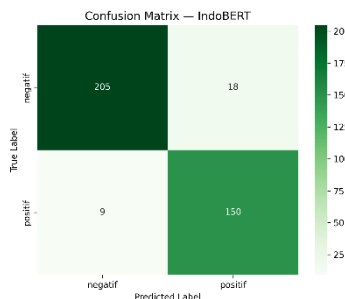
Confusion matrix digunakan untuk melihat distribusi hasil prediksi model terhadap masing-masing kelas sentimen. Pada penelitian ini, confusion matrix ditampilkan untuk model SVM sebagai model klasik terbaik dan IndoBERT sebagai model dengan performa terbaik secara keseluruhan.

Confusion Matrix — Model Klasik (TF-IDF)



Gambar 8. Confusion Matrix Model SVM

Hasil confusion matrix pada Gambar 8 menunjukkan bahwa model SVM mampu mengklasifikasikan sebagian besar tweet negatif dan positif dengan benar. Namun, masih terdapat beberapa kesalahan klasifikasi pada tweet yang mengandung opini campuran dan penggunaan bahasa informal.



Gambar 9. Confusion Matrix Model IndoBERT

Dari Confusion Matrix Gambar 9 IndoBERT menunjukkan jumlah kesalahan klasifikasi yang lebih sedikit dibandingkan

model klasik. Hal ini menunjukkan bahwa model transformer lebih mampu memahami konteks bahasa pada tweet berbahasa Indonesia.

TABEL III  
CONTOH KESALAHAN KLASIFIKASI MODEL

Tweet	Label Asli	Prediksi	Analisis
“program bagus tapi makanannya basi”	Negatif	Positif	Mengandung opini campuran
“gratis sih bagus, tapi kualitasnya dipertanyakan”	Negatif	Positif	Kata positif lebih dominan
“makan gratis malah bikin masalah baru”	Negatif	Positif	Konteks negatif tidak dominan secara token
“anak sekolah jadi lebih semangat makan”	Positif	Positif	Sentimen jelas dan langsung

Berdasarkan tabel 3 hasil error analysis, sebagian besar kesalahan klasifikasi terjadi pada tweet yang mengandung opini campuran, ironi, dan struktur kalimat ambigu. Model berbasis TF-IDF cenderung kesulitan memahami konteks kalimat karena proses klasifikasi lebih bergantung pada frekuensi token dibandingkan hubungan makna antar kata.

Selain itu, penggunaan bahasa informal dan singkatan pada media sosial juga mempengaruhi performa model. Tweet dengan struktur bahasa yang tidak baku sering kali menyebabkan model kesulitan memahami maksud sebenarnya dari pengguna. Kondisi tersebut menunjukkan bahwa analisis sentimen pada media sosial masih menjadi tantangan karena tingginya variasi bahasa dan konteks pada data teks.

*E. Perbandingan dan Analisis Model*

Berdasarkan hasil pengujian, IndoBERT menjadi model dengan performa terbaik pada penelitian ini dengan akurasi sebesar 92,93% dan F1-score tertinggi pada kedua kelas sentimen. Hasil tersebut menunjukkan bahwa model transformer memiliki kemampuan yang lebih baik dalam memahami konteks bahasa pada tweet berbahasa Indonesia dibandingkan model machine learning klasik berbasis TF-IDF.

Kemampuan IndoBERT dalam mempelajari hubungan kontekstual antar kata membantu model dalam memahami variasi bahasa informal yang umum digunakan pada media sosial. Hal ini menjadi keunggulan utama dibandingkan model klasik yang lebih bergantung pada frekuensi kata. Pada beberapa kasus, IndoBERT mampu mengenali konteks negatif atau positif meskipun kata-kata yang digunakan tidak secara langsung menunjukkan polaritas sentimen tertentu.

Namun demikian, perbedaan performa antara IndoBERT dan SVM tidak terlalu besar. SVM tetap mampu menghasilkan akurasi sebesar 92,15% dengan nilai precision dan recall yang relatif seimbang pada kedua kelas sentimen. Hal ini menunjukkan bahwa metode machine learning klasik berbasis TF-IDF masih relevan digunakan untuk analisis

sentimen berbahasa Indonesia, khususnya ketika sumber daya komputasi terbatas.

SVM menunjukkan performa yang stabil karena algoritma ini sangat cocok digunakan pada data teks berdimensi tinggi dan sparse seperti representasi TF-IDF. Selain itu, proses pelatihan SVM relatif lebih cepat dan ringan dibandingkan IndoBERT. Oleh karena itu, SVM dapat menjadi alternatif yang efisien untuk implementasi sistem analisis sentimen pada lingkungan dengan keterbatasan perangkat keras.

Naive Bayes juga menunjukkan performa yang kompetitif dengan akurasi yang sama seperti SVM. Model ini memiliki keunggulan dari sisi kesederhanaan algoritma dan efisiensi komputasi. Meskipun demikian, asumsi independensi antar fitur pada Naive Bayes menyebabkan model kurang optimal dalam memahami hubungan antar kata pada kalimat yang kompleks. Hal tersebut menyebabkan model masih mengalami kesalahan klasifikasi pada tweet yang mengandung opini campuran atau konteks ambigu.

Sementara itu, Random Forest memperoleh performa paling rendah dibandingkan model lainnya. Hal ini menunjukkan bahwa pendekatan ensemble tree kurang optimal untuk menangani representasi fitur teks berbasis TF-IDF yang memiliki dimensi tinggi dan sparse. Meskipun demikian, Random Forest tetap mampu menghasilkan performa klasifikasi yang cukup baik dan stabil pada kedua kelas sentimen.

Hasil penelitian ini menunjukkan bahwa pemilihan model klasifikasi tidak hanya bergantung pada tingkat akurasi, tetapi juga perlu mempertimbangkan kebutuhan komputasi dan kompleksitas implementasi. IndoBERT memberikan performa terbaik dalam memahami konteks bahasa, namun membutuhkan sumber daya komputasi yang lebih besar. Sebaliknya, model klasik seperti SVM mampu memberikan performa yang kompetitif dengan proses pelatihan yang lebih sederhana dan efisien.

Selain itu, hasil penelitian ini juga menunjukkan bahwa kombinasi preprocessing yang tepat, penggunaan TF-IDF, serta penanganan imbalance menggunakan SMOTE mampu meningkatkan performa model machine learning klasik secara signifikan. Hal ini membuktikan bahwa model klasik masih memiliki potensi yang baik dalam analisis sentimen media sosial berbahasa Indonesia apabila didukung oleh pipeline preprocessing yang optimal.

Secara umum, seluruh model pada penelitian ini mampu menghasilkan performa klasifikasi yang cukup tinggi dengan akurasi di atas 90%. Hal tersebut menunjukkan bahwa pendekatan machine learning maupun transformer dapat digunakan secara efektif untuk menganalisis sentimen publik terhadap Program Makan Bergizi Gratis (MBG) pada media sosial.

Meskipun demikian, penelitian ini masih memiliki beberapa keterbatasan. Data penelitian hanya diperoleh dari platform X sehingga belum sepenuhnya merepresentasikan opini seluruh masyarakat Indonesia. Selain itu, data media sosial sering mengandung spam, bot, bahasa informal, serta opini ekstrem yang dapat mempengaruhi hasil analisis

sentimen. Penggunaan pelabelan berbasis lexicon juga masih memiliki keterbatasan dalam memahami sarkasme dan konteks implisit pada tweet.

Untuk penelitian selanjutnya, analisis sentimen dapat dikembangkan dengan menggunakan dataset yang lebih besar dan berasal dari berbagai platform media sosial lainnya. Selain itu, penelitian berikutnya juga dapat menerapkan pendekatan aspect-based sentiment analysis atau topic modeling untuk memperoleh pemahaman yang lebih mendalam mengenai isu-isu spesifik yang dibahas masyarakat terkait Program MBG.

#### IV. KESIMPULAN

Penelitian ini berhasil menganalisis sentimen publik terhadap Program Makan Bergizi Gratis (MBG) pada platform X menggunakan metode Naive Bayes, Support Vector Machine (SVM), Random Forest, dan IndoBERT. Hasil penelitian menunjukkan bahwa sentimen negatif lebih dominan dibandingkan sentimen positif, yang dipengaruhi oleh isu keamanan pangan, pengelolaan anggaran, dan implementasi program.

Berdasarkan hasil pengujian, IndoBERT memperoleh performa terbaik dengan accuracy sebesar 92,93%, sedangkan SVM menjadi model klasik dengan performa paling kompetitif dengan accuracy sebesar 92,15%. Hasil tersebut menunjukkan bahwa model transformer lebih unggul dalam memahami konteks bahasa, namun model klasik berbasis TF-IDF masih mampu memberikan performa yang kompetitif dengan kebutuhan komputasi yang lebih ringan.

Selain itu, analisis wordcloud dan frekuensi kata menunjukkan bahwa sentimen positif lebih banyak berkaitan dengan manfaat program terhadap gizi dan kesehatan masyarakat, sedangkan sentimen negatif didominasi oleh kritik terhadap implementasi dan pengelolaan program. Penelitian selanjutnya dapat menggunakan dataset yang lebih besar dan pendekatan analisis yang lebih mendalam untuk memperoleh hasil yang lebih komprehensif.

#### DAFTAR PUSTAKA

- [1] E. M. Sipayung, "Sentiment on Public Trust Using the NLP Rule Based Method," *Jurnal Sistem dan Teknologi Informasi (JustIN)*, vol. 12, no. 1, p. 175, 2024, doi: 10.26418/justin.v12i1.72426.
- [2] A. M. Rizqiyah and I. K. D. Nuryana, "Analisis Sentimen Masyarakat terhadap Kebijakan Iuran Tabungan Perumahan Rakyat (Tapera) pada Platform X Menggunakan Algoritma Naive Bayes Classifier dan Support Vector Machine," *Journal of Emerging Information Systems and Business Intelligence (JEISBI)*, vol. 5, no. 3, pp. 298–306, 2024, [Online]. Available: <https://ejournal.unesa.ac.id/index.php/JEISBI/article/view/64074>
- [3] F. Fathoni, A. Ibrahim, F. Rizka Mumtaz, M. Azmi Zaky, M. Jodi Pratama, and I. Akbar Kurniawan, "Analisis Sentimen Public Twitter Terhadap Kebijakan Pemerintah Menggunakan Metode Svm (Studi Kasus : Ruu Tni )," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 4, pp. 6322–6329, 2025, doi: 10.36040/jati.v9i4.14036.
- [4] L. S. Memory *et al.*, "SMATIKA: STIKI Informatika Jurnal Sentiment analysis on public opinion trends on #

- MakanBergiziGratis programs on,” vol. 16, no. 79, pp. 175–189, 2026.
- [5] L. Nursingghah, R. Ruuhwan, and T. Mufizar, “Analisis Sentimen Pengguna Aplikasi X Terhadap Program Makan Siang Gratis Dengan Metode Naïve Bayes Classifier,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, 2024, doi: 10.23960/jitet.v12i3.4336.
- [6] N. N. Aprianti, N. M. M. R. Desmayani, L. G. B. Libraeni, I. G. A. Indrawan, and M. L. Radhitya, “Public Sentiment Analysis of the Free Nutritious Meals Program (MBG) on Social Media X Using the Naïve Bayes Method,” *Journal of Applied Informatics and Computing*, vol. 9, no. 6, pp. 3929–3936, 2025, doi: 10.30871/jaic.v9i6.11420.
- [7] M. F. Saleh and R. Imanda, “Public Sentiment Analysis of the Free Meal Program: A Comparison of Naïve Bayes and Support Vector Machine Methods on the Twitter (X) Social Media Platform,” *Journal of Applied Informatics and Computing*, vol. 9, no. 1, pp. 131–139, 2025, doi: 10.30871/jaic.v9i1.8895.
- [8] A. Sitanggang, Y. Umaidah, Y. Umaidah, R. I. Adam, and R. I. Adam, “Analisis Sentimen Masyarakat Terhadap Program Makan Siang Gratis Pada Media Sosial X Menggunakan Algoritma Naïve Bayes,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, 2024, doi: 10.23960/jitet.v12i3.4902.
- [9] Z. Purwanti and Sugiyono, “Pemodelan Text Mining untuk Analisis Sentimen Terhadap Program Makan Siang Gratis di Media Sosial X Menggunakan Algoritma Support Vector Machine (SVM),” *Jurnal Indonesia : Manajemen Informatika dan Komunikasi*, vol. 5, no. 3, pp. 3065–3079, 2024, doi: 10.35870/jimik.v5i3.1001.
- [10] N. Nur, A. Aryanti, and O. Suria, “Analisis Sentimen Terhadap Pemutusan Hubungan Kerja Di Indonesia : Komparasi Indobert Dengan Svm , Random Forest , Dan Decision Tree Dengan Optimasi Tf - Idf Pendahuluan Pemutusan Hubungan Kerja (PHK) merupakan salah satu fenomena sosial dan ekonomi yan,” vol. 10, no. 2, pp. 1158–1176, 2025.
- [11] R. Rahmadani, A. Rahim, and R. Rudiman, “Analisis Sentimen Ulasan ‘Ojol the Game’ Di Google Play Store Menggunakan Algoritma Naive Bayes Dan Model Ekstraksi Fitur Tf-Idf Untuk Meningkatkan Kualitas Game,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, 2024, doi: 10.23960/jitet.v12i3.4988.
- [12] A. Winston, N. D. Rafael, and H. Lucky, “Analyzing Public Sentiment Toward the Makan Bergizi Gratis ( MBG ) Program on TikTok Using SVM and IndoBERT,” vol. 3, no. 1, pp. 33–39, 2026, doi: 10.21512/ijcshaijournal.v3i1.15184.
- [13] J. Homepage, U. Arfan, and E. Badii, “MALCOM: Indonesian Journal of Machine Learning and Computer Science Sentiment Analysis of Papuan Public Perceptions Toward the Free Nutritional Meal Program in Indonesia Analisis Sentimen Masyarakat Papua Terhadap Program Makan Gizi Gratis di Indonesia,” vol. 5, no. October, pp. 1538–1546, 2025.
- [14] V. B. Lestari and C. A. Hutagalung, “Evaluation of TF-IDF Extraction Techniques in Sentiment Analysis of Indonesian-Language Marketplaces Using SVM, Logistic Regression, and Naïve Bayes,” *Journal of Computer Science and Applications*, vol. 8, no. 1, pp. 22–2025, 2025, [Online]. Available: <https://doi.org/10.21009/j->
- [15] N. Husin, “Komparasi Algoritma Random Forest, Naïve Bayes, dan Bert Untuk Multi-Class Classification Pada Artikel Cable News Network (CNN),” *Jurnal Esensi Infokom : Jurnal Esensi Sistem Informasi dan Sistem Komputer*, vol. 7, no. 1, pp. 75–84, 2023, doi: 10.55886/infokom.v7i1.608.
- [16] Heti Aprilianti, Khothibul Umam, and Maya Rini Handayani, “Comparative Study of SVM, KNN, and Naïve Bayes for Sentiment Analysis of Religious Application Reviews,” *Journal of Applied Informatics and Computing*, vol. 9, no. 3, pp. 920–927, 2025, doi: 10.30871/jaic.v9i3.9482.
- [17] H. Jayadianti, W. Kaswidjanti, A. T. Utomo, S. Saifullah, F. A. Dwiyanto, and R. Drezewski, “Sentiment analysis of Indonesian reviews using fine-tuning IndoBERT and R-CNN,” *ILKOM Jurnal Ilmiah*, vol. 14, no. 3, pp. 348–354, 2022, doi: 10.33096/ilkom.v14i3.1505.348-354.
- [18] M. Samantri and Afiyati, “Perbandingan Algoritma Support Vector Machine dan Random Forest untuk Analisis Sentimen Terhadap Kebijakan Pemerintah Indonesia Terkait Kenaikan Harga BBM Tahun 2022,” *Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi)*, vol. 8, no. 1, pp. 1–9, 2024, doi: 10.35870/jtik.v8i1.1202.
- [19] N. S. Sediarmoko, Y. Nataliani, and I. Suryady, “Sentiment Analysis of Customer Review Using Classification Algorithms and SMOTE for Handling Imbalanced Class,” *Indonesian Journal of Information Systems*, vol. 7, no. 1, pp. 38–52, 2024, doi: 10.24002/ijis.v7i1.8879.
- [20] M. F. As Shidiq and D. Alita, “Analisis Sentimen Masyarakat Terhadap Kasus Judi Online Menggunakan Data Dari Media Sosial X Pendekatan Naive Bayes Dan Svm,” *Jurnal Sistem Informasi dan Informatika (Simika)*, vol. 8, no. 1, pp. 24–35, 2025, doi: 10.47080/simika.v8i1.3624.
- [21] B. F. S. Supriyanto and S. Rosalin, “Analisis Sentimen Program Merdeka Belajar dengan Text Analysis Wordcloud & Word Frequency,” *Jurnal Minfo Polgan*, vol. 12, no. 1, pp. 25–32, 2023, doi: 10.33395/jmp.v12i1.12312.