

Public Sentiment Analysis of the Free Nutritious Meals Program (MBG) on Social Media X Using the Naive Bayes Method

Ni Nyoman Aprianti^{1*}, Ni Made Mila Rosa Desmayani², Luh Gede Bevi Libraeni³, I Gusti Agung Indrawan⁴,
Made Leo Radhitya⁵

^{1,2,3,4,5} Teknik Informatika, Institut Bisnis Dan Teknologi Indonesia, Denpasar

Jl. Tukad Pakerisan No.97, Panjer, Denpasar Selatan, Kota Denpasar, Bali 80225

omanaprianti@gmail.com¹, mlarosadesymayani@instiki.ac.id², bevi.libraeni@instiki.ac.id³, agung.indrawan@instiki.ac.id⁴,
leo.radhitya@instiki.ac.id⁵

Article Info

Article history:

Received 2025-10-02

Revised 2025-11-28

Accepted 2025-12-10

Keyword:

*Sentiment Analysis,
Free Nutritious Meals,
Naïve Bayes,
TF-IDF,
Platform X.*

ABSTRACT

This study aims to analyze public sentiment towards the Free Nutritious Meals Program (MBG) launched by the government, utilizing data from the X (Twitter) platform using the Naïve Bayes method. The background of this study is based on the high level of public attention towards the MBG program, which targets school children, toddlers, pregnant women, and nursing mothers, as well as the prevalence of diverse opinions on social media. Data was collected through a crawling process during the period of April 28 to May 28, 2025, using keywords related to MBG, resulting in 12,310 tweets. The data processing stages included text preprocessing (cleansing, case folding, tokenizing, filtering, stemming), word weighting with TF-IDF, training and test data division, and testing using a confusion matrix. The results show that the Naïve Bayes method is capable of classifying sentiment into three categories: positive, negative, and neutral, with optimal performance on an 80:20 data split, resulting in an accuracy of 86.78%, precision of 86.86%, recall of 86.78%, and an F1-score of 86.58%. The majority of public sentiment towards the MBG program was positive, reflecting support for the program's benefits in improving the nutrition of school children and alleviating the economic burden on families. This study is expected to serve as a reference for the government in evaluating public policy and communication strategies, as well as contributing academically to the development of text mining and sentiment analysis studies on social media.



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) merupakan sebuah inisiatif strategis pemerintah yang bertujuan meningkatkan kualitas gizi sekaligus motivasi belajar siswa di Indonesia. Program ini didasarkan pada data yang menunjukkan bahwa sekitar 41% siswa mengalami kelaparan saat belajar, yang berdampak negatif terhadap konsentrasi dan prestasi akademik mereka. MBG tidak hanya menyasar anak-anak sekolah, tetapi juga diperluas untuk mencakup balita, ibu hamil, dan ibu menyusui. Untuk tahap awal pelaksanaan pada 2 Januari 2025, pemerintah telah mengalokasikan anggaran sebesar Rp71 triliun dalam RAPBN 2025 [1].

Dalam konteks perkembangan media sosial, platform X (sebelumnya Twitter) menjadi salah satu wadah paling

berpengaruh dalam membentuk wacana publik. Platform ini populer karena kemampuannya menyebarkan informasi secara cepat melalui fitur seperti *hashtag*, yang memudahkan pencarian isu terkini. Pasca diakuisisi Elon Musk, X tetap mempertahankan basis pengguna aktifnya dan menjadi ruang interaksi masyarakat yang dinamis [2].

Isu MBG menjadi salah satu topik yang ramai diperbincangkan di platform ini, baik dalam bentuk dukungan maupun kritik. Ketertarikan peneliti pada topik ini didorong oleh beberapa hal, antara lain: relevansi MBG sebagai kebijakan strategis pemerintah yang berdampak langsung pada masyarakat, tingginya perhatian publik, serta adanya kontroversi dan perdebatan di media sosial. Dukungan masyarakat terhadap MBG terutama didasari pada manfaatnya dalam meningkatkan asupan gizi anak sekolah

dari keluarga kurang mampu, membantu mereka lebih fokus dalam belajar, serta meringankan beban ekonomi orang tua. Dari sisi pemerataan, program ini juga mencerminkan komitmen negara menjamin kesejahteraan anak-anak di seluruh pelosok, termasuk daerah 3T (Tertinggal, Terdepan, dan Terluar) [3].

Dalam konteks Indonesia, analisis sentimen berbasis media sosial telah terbukti efektif digunakan untuk mengevaluasi berbagai kebijakan publik yang kontroversial maupun strategis. Penelitian terdahulu telah menganalisis sentimen publik terhadap kebijakan lain, seperti kebijakan COVID-19 [4] dan isu kenaikan harga BBM atau PPN [5]. Studi-studi tersebut menegaskan kapabilitas algoritma *machine learning*, termasuk Naive Bayes, dalam mengidentifikasi dinamika opini. Namun, hingga saat ini, belum banyak literatur yang secara komprehensif memetakan sentimen publik secara mendalam terhadap Program Makan Bergizi Gratis (MBG)—sebuah kebijakan strategis dengan dampak sosial-ekonomi yang masif dan menyasar kelompok rentan. Penelitian ini hadir untuk menutup celah tersebut (*research gap*) dengan menyajikan analisis sentimen pada 12.310 *tweet* terkait MBG.

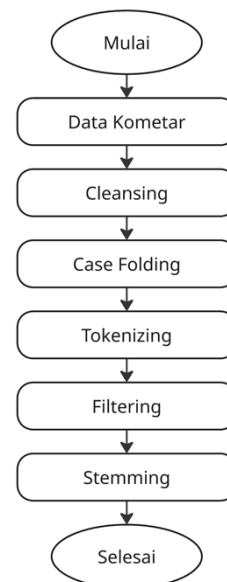
Secara akademis, penelitian ini berkontribusi dalam pengembangan kajian *text mining* dan *Natural Language Processing* (NLP), khususnya dalam menguji efektivitas algoritma Naive Bayes—yang dikenal karena efisiensi komputasi dan kemampuannya pada data teks bervolume besar—dalam konteks kebijakan publik Indonesia yang spesifik. Secara sosial dan kebijakan, hasil klasifikasi sentimen ini memiliki urgensi tinggi sebagai mekanisme *feedback* cepat bagi pemerintah. Dengan memahami distribusi sentimen (positif, negatif, netral) dan isu-isu spesifik yang diangkat publik, pemerintah dapat melakukan evaluasi kebijakan berbasis bukti (*evidence-based policy making*), menyusun strategi komunikasi yang lebih tepat sasaran, serta segera memitigasi potensi disinformasi atau kesalahpahaman publik terkait anggaran dan implementasi MBG. Oleh karena itu, penelitian ini tidak hanya menguji performa model, tetapi juga menjembatani data media sosial menjadi masukan kebijakan yang berharga.

Dalam menganalisis respons publik, pendekatan analisis sentimen menjadi sangat relevan. Analisis sentimen berbasis *machine learning* memungkinkan identifikasi pola opini publik secara sistematis serta mengungkap nuansa emosi yang tidak selalu tampak dalam survei konvensional. Hal ini penting mengingat volume data media sosial yang sangat besar—misalnya, tercatat sekitar 12.310 ribu *tweet* diposting setiap hari di seluruh dunia [6].

Untuk itu, penelitian ini menggunakan algoritma Naive Bayes sebagai metode utama dalam klasifikasi sentimen. Menurut [7], Naive Bayes terbukti mampu menangani data tidak seimbang—suatu karakteristik umum pada data media sosial di mana distribusi komentar positif dan negatif sering tidak merata. Hal ini menjadikannya pilihan ideal untuk analisis sentimen pada topik MBG. Berdasarkan uraian di atas, maka rumusan masalah penelitian ini adalah: Bagaimana

performa algoritma Naive Bayes dalam melakukan klasifikasi sentimen terhadap unggahan masyarakat di platform X terkait Program Makan Bergizi Gratis (MBG)?

II. METODE



Gambar 1. Tahapan Preprocessing

Penelitian ini dilaksanakan secara daring dengan memanfaatkan data media sosial X (Twitter) terkait Program Makan Bergizi Gratis (MBG). Proses pengumpulan data dilakukan pada periode 28 April hingga 28 Mei 2025 menggunakan Tweet-Harvest, yaitu alat *open-source* berbasis Node.js. Metode ini merupakan teknik *scraping* yang memanfaatkan koneksi API publik atau semi-resmi platform X untuk mendapatkan data publik. Kata kunci yang digunakan adalah “program makan bergizi gratis” dan “makan gratis”, sehingga diperoleh data mentah berjumlah 12.310 *tweet*.

Setelah pengumpulan, dilakukan proses *data cleaning* awal yang krusial untuk memastikan kualitas data. Proses ini secara spesifik mencakup penghapusan *retweet*, penghapusan data ganda (duplikasi), serta identifikasi dan penghapusan *tweet* yang terindikasi spam atau yang mengandung teks generik/tidak relevan. Setelah melalui tahap pembersihan awal ini, jumlah data akhir yang layak digunakan dalam penelitian ini adalah sebanyak 12.000 *tweet*. Data tersebut dianggap cukup representatif untuk menggambarkan opini publik di media sosial mengenai program MBG [3].

Jenis data dalam penelitian ini terdiri dari data primer dan data sekunder. Data primer berupa unggahan atau komentar langsung dari pengguna X yang membahas MBG. Data ini penting karena merefleksikan opini publik secara alami dalam konteks nyata. Sementara itu, data sekunder diperoleh dari studi kepustakaan berupa buku, jurnal, artikel ilmiah, dan penelitian terdahulu yang relevan. Data sekunder digunakan

untuk memperkuat dasar teori, memahami algoritma yang digunakan, serta memperjelas kerangka metodologi penelitian. [6].

Tahap selanjutnya adalah *text preprocessing*, yang dilakukan untuk membersihkan dan menyiapkan data teks agar dapat dianalisis dengan baik. Tahapan *preprocessing* mencakup: (1) *cleansing*, yaitu menghapus tanda baca, simbol, angka, emotikon, URL, dan karakter khusus yang tidak diperlukan; (2) *case folding*, yaitu mengubah seluruh huruf menjadi huruf kecil agar format teks seragam; (3) *tokenizing*, yaitu memecah kalimat menjadi token atau potongan kata; (4) *filtering*, yaitu menghapus *stopword* seperti “dan”, “di”, atau “yang” yang tidak memberikan makna penting dalam analisis; dan (5) *stemming*, yaitu mengubah kata ke bentuk dasarnya, misalnya kata “dimakan”, “memakan”, dan “makanan” menjadi “makan”. Proses ini dilakukan agar data teks menjadi lebih konsisten, terstruktur, dan siap untuk dianalisis lebih lanjut [8].

Setelah melalui *preprocessing*, tahap berikutnya adalah pelabelan data. Setiap tweet diberi label berdasarkan sentimen yang terkandung di dalamnya, yaitu positif, negatif, atau netral. Proses pelabelan dilakukan dengan pendekatan *lexicon-based* menggunakan daftar kata yang telah ditentukan bobot sentimennya. Misalnya, kata “membantu”, “bagus”, dan “bermanfaat” diberi bobot positif, sedangkan kata “buruk”, “gagal”, atau “mengecewakan” diberi bobot negatif. Tweet yang mengandung lebih banyak kata positif diberi label positif, sebaliknya jika lebih banyak kata negatif diberi label negatif, sedangkan jika tidak dominan, maka diberi label netral [6].

Setelah pelabelan data, tahap berikutnya adalah pembobotan fitur menggunakan metode TF-IDF (*Term Frequency-Inverse Document Frequency*), yang berfungsi mengonversi teks menjadi representasi numerik. Data yang telah dibobot kemudian dibagi menjadi dua subset menggunakan proporsi pembagian data latih dan data uji sebesar 80:20. Proporsi ini berarti data latih (80%) berjumlah 9.600 *tweet*, yang digunakan untuk melatih model, dan data uji (20%) berjumlah 2.400 *tweet*, yang digunakan untuk mengukur kinerja model pada data yang belum pernah dilihat sebelumnya. Pembagian ini dilakukan secara acak (*random split*) untuk menghindari bias dan memastikan objektivitas evaluasi model.

Algoritma yang digunakan untuk klasifikasi sentimen adalah Naive Bayes Classifier, yaitu algoritma berbasis probabilitas yang sederhana namun efektif dalam klasifikasi teks. Naive Bayes dipilih atas pertimbangan efisiensi dan kapabilitasnya dalam lingkungan *text mining*, khususnya pada data media sosial yang sering kali bervolume besar (12.000 *tweet*). Pemilihan Naive Bayes ini didasarkan pada tiga alasan utama dibandingkan metode yang lebih kompleks seperti SVM (*Support Vector Machine*) atau LSTM (*Long Short-Term Memory*):

1) *Kecepatan dan Efisiensi Komputasi*: Naive Bayes memiliki kecepatan *training* dan klasifikasi yang jauh lebih tinggi.

2) *Efektivitas pada Data Teks*: Algoritma ini terbukti sangat baik dalam klasifikasi teks, terutama saat digunakan bersama *feature weighting* seperti TF-IDF.

3) *Kebutuhan Sumber Daya*: Naive Bayes tidak memerlukan sumber daya komputasi dan dataset sebesar yang dibutuhkan oleh model *Deep Learning* seperti LSTM, yang idealnya memerlukan data yang jauh lebih masif dan terstruktur untuk kinerja optimal.

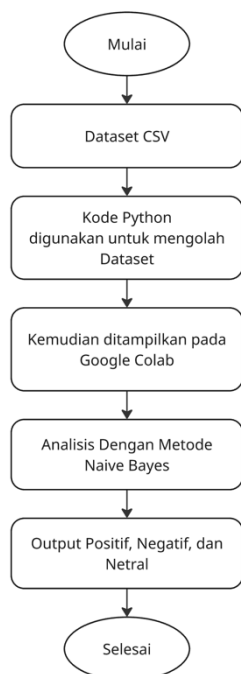
Selain itu, Naive Bayes mampu menangani distribusi kelas yang tidak seimbang, di mana jumlah *tweet* positif, negatif, dan netral sering kali tidak sama [7]. Untuk mengukur kinerja model, dilakukan evaluasi menggunakan *Confusion Matrix* yang menghasilkan beberapa metrik, yaitu akurasi, presisi, recall, dan F1-score. Akurasi dihitung dari jumlah prediksi yang benar dibandingkan dengan total data uji, sedangkan presisi, recall, dan F1-score digunakan untuk menilai performa klasifikasi tiap kelas (positif, negatif, netral). Selain itu, hasil analisis juga divisualisasikan menggunakan *word cloud* dan diagram batang untuk menunjukkan kata-kata yang paling sering muncul serta distribusi sentimen pada data yang dianalisis [6].

III. HASIL DAN PEMBAHASAN

A. Implementasi Sentimen

Implementasi “Analisis Sentimen terhadap Program Makan Bergizi Gratis (MBG) pada Platform X Menggunakan Metode Naive Bayes” ini bertujuan untuk mengetahui bagaimana opini publik terhadap program MBG yang sedang ramai diperbincangkan di media sosial, khususnya di platform X (Twitter). Penelitian ini dilakukan dengan menggunakan metode Naive Bayes untuk mengklasifikasikan komentar atau opini menjadi tiga sentimen utama, yaitu positif, negatif, dan netral. Hasil klasifikasi akan digunakan untuk melihat kecenderungan opini publik secara umum terhadap program MBG. Penelitian ini dilaksanakan dengan memanfaatkan bahasa pemrograman Python dan dijalankan melalui platform Google Colab, serta menggunakan file data berformat CSV. Data mentah yang berhasil dikumpulkan dari platform X (Twitter) berjumlah 12.310 *tweet*. Dalam penelitian ini, digunakan rasio data latih sebesar 20% dari total dataset yang tersedia. Selanjutnya, proses *preprocessing* data dilakukan untuk membersihkan teks dari unsur-unsur yang tidak relevan sebelum dimasukkan ke dalam proses klasifikasi, yang tahapan *preprocessing*-nya akan dijelaskan dalam deskripsi berikut [7]. Data dikumpulkan melalui proses *crawling* dari platform media sosial X(Twitter) menggunakan alat bantu Tweet-Harvest Token yang valid. Data dikumpulkan dengan kata kunci “Makan Bergizi Gratis” selama periode 28 April hingga 28 Mei 2025. Data hasil *crawling* disimpan dalam bentuk file CSV dan dimasukkan ke dalam lingkungan kerja Google Colab untuk dilakukan pengolahan awal. Selanjutnya, data tersebut diproses menggunakan bahasa pemrograman Python dengan melalui tahapan *text preprocessing* seperti *cleansing*, *case folding*, *tokenizing*, *filtering*, dan *stemming*. Setelah proses pembersihan teks, dilakukan pelabelan data

menggunakan pendekatan lexicon-based untuk mengelompokkan sentimen. Output dari sistem ini adalah hasil klasifikasi opini masyarakat terhadap program MBG yang terbagi menjadi 3 kategori yaitu: positif, negatif, dan netral. Berikut merupakan tampilan alur tahapan processing data dalam bentuk *flowchart*.



Gambar 2. Proses Data

Alur keseluruhan proses analisis data (Gambar X) dimulai dengan tahap Inisialisasi Proyek dan penyiapan lingkungan kerja di Google Colaboratory (Colab), yang mendukung pengolahan data menggunakan bahasa pemrograman Python. Data mentah (Dataset CSV) yang berhasil dikumpulkan dari Platform X kemudian melalui tahapan pra-pemrosesan yang meliputi cleansing, case folding, tokenizing, filtering, dan stemming. Setelah data bersih, dilakukan Analisis dengan Metode Naive Bayes, di mana model dilatih menggunakan data yang telah dibersihkan untuk mengklasifikasikan setiap tweet. Tahap akhir adalah Menghasilkan Output Klasifikasi, membagi opini publik ke dalam tiga kategori sentimen: Positif (dukungan), Negatif (penentangan/kritik), dan Netral (informasi tanpa sentimen kuat). Seluruh proses ini menghasilkan luaran yang siap digunakan untuk pelaporan, visualisasi, dan rekomendasi kebijakan [6].

B. Training Model Naive Bayes

Pada tahap kali ini dilakukan proses pelatihan model untuk mengklasifikasikan sentimen menggunakan algoritma Naive Bayes, yaitu salah satu metode dalam pembelajaran mesin yang dikenal efisien dan banyak digunakan dalam analisis teks. Tujuan dari langkah ini adalah agar model dapat mempelajari pola-pola dari data yang telah dibersihkan dan dikonversi ke bentuk numerik menggunakan metode TF-IDF,

sehingga model mampu memprediksi sentimen dari data baru secara otomatis. Tahap evaluasi model Naive Bayes dilakukan menggunakan data uji untuk mengukur kinerja klasifikasi pada setiap kategori sentimen, yang dirangkum dalam *Classification Report*. Secara keseluruhan, model mencapai akurasi sebesar 86,78%, dengan nilai *Macro* dan *Weighted Average* untuk seluruh metrik (Presisi, *Recall*, dan F1-score) berada pada angka 0.87, mengindikasikan kinerja model yang baik dan seimbang antar kelas. Secara rinci, performa model menunjukkan variasi: Sentimen Negatif memiliki kinerja terbaik dengan nilai *Recall* mencapai 0.97 dan F1-score 0.91 (Presisi 0.86), yang menunjukkan kemampuan model yang luar biasa dalam mendeteksi dan mengidentifikasi hampir semua *tweet* yang berlabel negatif. Sentimen Positif menunjukkan kinerja yang stabil dan seimbang di semua metrik, dengan Presisi 0.86, *Recall* 0.87, dan F1-score 0.86. Sementara itu, kinerja pada Sentimen Netral cukup akurat dengan Presisi 0.88 dan F1-score 0.82, namun nilai *Recall* 0.77 mengindikasikan bahwa masih terdapat sejumlah data netral yang terlewat atau diklasifikasikan ke kelas lain. Keseimbangan nilai metrik ini menegaskan bahwa model Naive Bayes terbukti efektif dalam klasifikasi sentimen terhadap Program MBG.

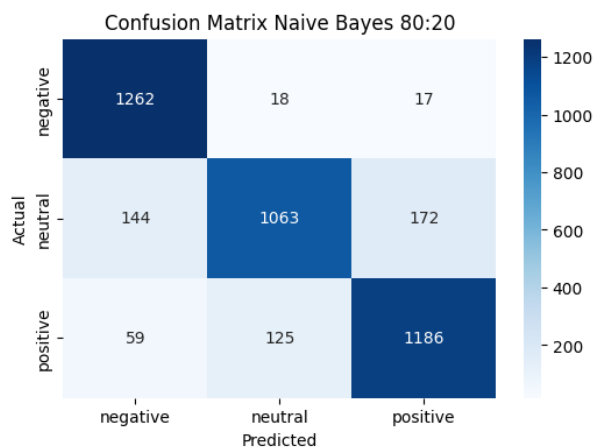
Classification Report:				
	precision	recall	f1-score	support
negative	0.86	0.97	0.91	1297
neutral	0.88	0.77	0.82	1379
positive	0.86	0.87	0.86	1370
accuracy			0.87	4046
macro avg	0.87	0.87	0.87	4046
weighted avg	0.87	0.87	0.87	4046

Akurasi Model (Naive Bayes - 80/20): 0.8678

Gambar 3. Hasil Clasification Report

C. Hasil Visualisasi Confusion Matrix (80:20)

Visualisasi ini menampilkan confusion matrix dari hasil prediksi model Naive Bayes dengan skema pembagian data 80:20 antara data latih dan data uji. Confusion matrix berfungsi untuk menunjukkan jumlah prediksi yang benar maupun salah pada setiap kelas sentimen (positif, negatif, dan netral). Dengan cara ini, performa model dapat dilihat lebih jelas, termasuk pola kesalahan klasifikasi yang terjadi (Apriliyani & Salim, 2022). Berikut ditampilkan cuplikan skrip dan hasil visualisasi confusion matrix. Berikut ini merupakan hasil dari proses confusion matrix dengan pembagian data 80:20.



Gambar 4. Confusion Matrix Naive Bayes 80:20

Gambar confusion matrix di atas menunjukkan performa model Naive Bayes dalam mengklasifikasikan data sentimen ke dalam tiga kategori: positive, negative, dan neutral. gambar tersebut terbagi dalam tiga baris dan tiga kolom:

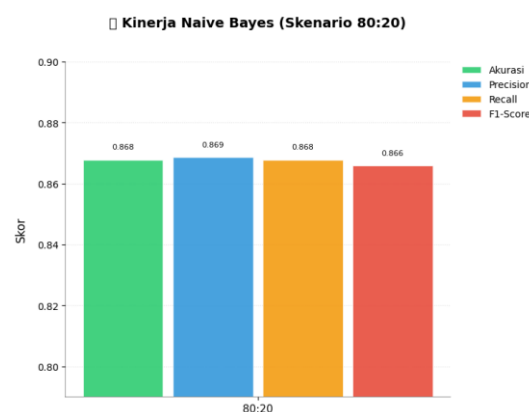
1. Baris merepresentasikan label sebenarnya (aktual),
 2. Kolom merepresentasikan hasil prediksi dari model
- Setiap sel pada tabel berisi jumlah data yang termasuk dalam kategori tersebut. Semakin tinggi angka pada baris = kolom (bagian diagonal), maka semakin baik akurasi model terhadap kategori tersebut.

D. Hasil Pengujian Naive Bayes

Visualisasi Kinerja Naive Bayes (Skenario 80:20) menggambarkan hasil evaluasi model klasifikasi sentimen dengan menggunakan empat metrik utama, yaitu akurasi, precision, recall, dan F1-score. Seluruh metrik tersebut dihitung berdasarkan skenario pembagian data dengan rasio 80% untuk data latih dan 20% untuk data uji. Rasio ini dipilih karena pada pengujian sebelumnya terbukti menghasilkan performa terbaik dibandingkan skema lain. Visualisasi kinerja dalam bentuk grafik memberikan gambaran yang lebih jelas mengenai efektivitas model dalam mengenali sentimen, baik positif, negatif, maupun netral [7]. Dengan adanya grafik ini, pembaca dapat lebih mudah memahami tingkat kekuatan serta kelemahan model, tanpa hanya bergantung pada tabel angka semata. Potongan skrip yang disertakan menunjukkan bagaimana proses visualisasi dilakukan secara langsung dari hasil perhitungan metrik, sehingga memperlihatkan keterkaitan antara implementasi teknis dan hasil evaluasi yang diperoleh.

Hasil grafik batang di atas menunjukkan kinerja model Naive Bayes pada skenario pembagian data 80:20 dengan empat metrik evaluasi: akurasi, precision, recall, dan F1-score. Pada skenario ini, model mencapai nilai yang konsisten tinggi di semua metrik, dengan akurasi dan precision berada di kisaran 86,8–86,9%. Hasil ini mengindikasikan bahwa porsi data latih yang besar (80%) membuat model lebih mampu mengenali pola data, sehingga performanya menjadi optimal. Kinerja yang seimbang antara precision, recall, dan

F1-score juga menunjukkan bahwa model tidak hanya tepat dalam memprediksi, tetapi juga mampu menangkap sebagian besar data yang relevan secara akurat.



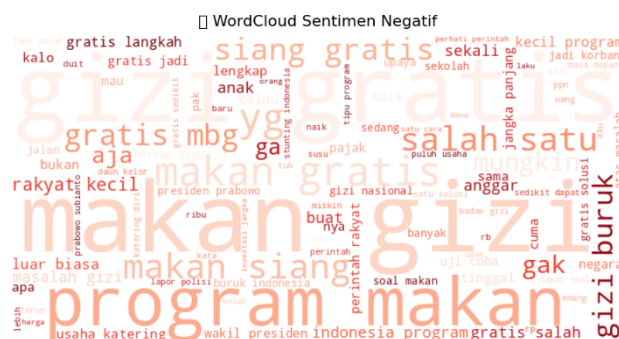
Gambar 5. Hasil Grafik Kinerja Naive Bayes (80:20)

E. Visualisasi WordCloud

Visualisasi WordCloud digunakan untuk menampilkan kata-kata yang paling sering muncul dalam setiap kategori sentimen. Sentimen positif ditandai dengan kata pujian berwarna hijau, negatif dengan kata kritik berwarna merah, dan netral dengan kata informatif berwarna biru. Visualisasi ini memudahkan analisis opini tanpa harus membaca seluruh data secara manual (Putriyekt et al., 2025).



Gambar 6. WordCloud Sentimen Positif



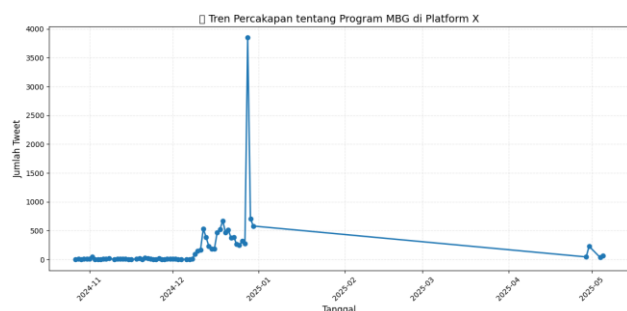
Gambar 7. WordCloud Sentimen Negatif



Gambar 8. WordCloud Sentimen Netral

F. Tren Percakapan

Untuk memahami sejauh mana perbincangan publik mengenai Program Makan Bergizi Gratis (MBG) di media sosial, khususnya di Platform X (sebelumnya Twitter), dilakukan analisis tren percakapan berdasarkan data jumlah tweet yang dipublikasikan setiap harinya. Analisis ini menggunakan visualisasi grafik garis yang menggambarkan fluktuasi jumlah tweet dari hari ke hari (Putriyeksi et al., 2025). Langkah pertama dimulai dengan membaca dataset mbg.csv, yang berisi kumpulan tweet terkait MBG. Titik-titik pada garis menunjukkan jumlah tweet pada tanggal tertentu, sementara garis yang menghubungkan titik-titik tersebut memberikan gambaran umum tentang naik-turunnya volume percakapan.



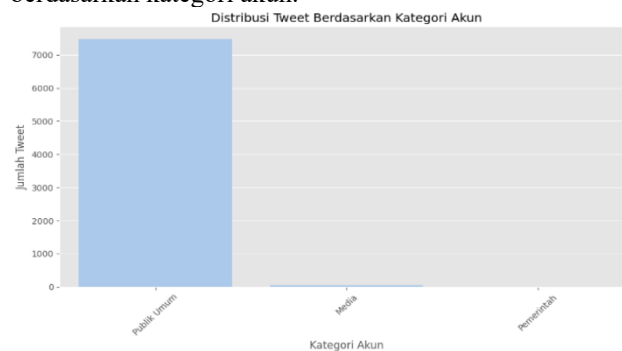
Gambar 9. Grafik Tren Percakapan

Gambar 9 menunjukkan tren percakapan masyarakat di media sosial Platform X (Twitter) mengenai Program Makan Bergizi Gratis (MBG) selama periode tertentu. Garis biru pada grafik memperlihatkan jumlah tweet yang dipublikasikan setiap hari, sehingga tampak pola kenaikan maupun penurunan diskusi. Puncak percakapan terjadi pada 16 Desember 2024, ditandai dengan lonjakan tajam jumlah tweet. Kenaikan ini dipicu oleh beberapa unggahan yang viral, khususnya yang berisi kritik terhadap prioritas anggaran MBG dibandingkan kebutuhan lain seperti biaya pendidikan. Cuplikan dua tweet paling *trending* juga ditampilkan dalam grafik, menunjukkan bahwa opini publik terbagi antara dukungan dan penolakan. Sebagian menilai MBG bermanfaat untuk meningkatkan gizi siswa dari keluarga kurang mampu, sementara sebagian lain menganggap alokasi anggaran lebih baik difokuskan ke sektor lain yang lebih mendesak. Temuan

ini menegaskan bahwa media sosial menjadi ruang diskusi publik yang aktif dan sering kali memunculkan pro dan kontra dalam menanggapi kebijakan pemerintah.

G. Peta Netizen di X Tentang MBG

Analisis ini menggunakan pendekatan Jaringan Sosial (Social Network Analysis/SNA) untuk memetakan pola komunikasi antar pengguna di Platform X melalui relasi mention. Proses ini menghasilkan graf jaringan di mana simpul (node) merepresentasikan akun pengguna dan garis (edge) menunjukkan interaksi di antara mereka. Hasil pemetaan (Visualisasi Peta Jaringan) menunjukkan bahwa percakapan mengenai Program MBG terbagi menjadi tiga kluster utama: Kluster Pemerintah, Kluster Media, dan Kluster Publik Umum. Masing-masing kluster menunjukkan fokus narasi yang berbeda: Pemerintah fokus pada penyampaian tujuan, manfaat, dan klarifikasi resmi kebijakan; Media berfungsi sebagai penyebar berita informatif dan sorotan kritis; sementara Publik Umum mendominasi diskusi dengan ekspresi opini yang beragam, mulai dari dukungan, kritik, hingga harapan. Analisis ini memberikan pemahaman mengenai struktur komunikasi di media sosial, mengidentifikasi aktor-aktor kunci, serta memetakan bagaimana narasi dan pengaruh didistribusikan dalam diskusi publik terkait MBG. Berikut ini merupakan distribusi tweet berdasarkan kategori akun.



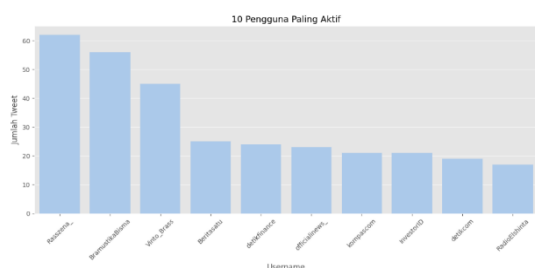
Gambar 10. Distribusi Tweet Berdasarkan Kategori Akun

Gambar 10 menunjukkan distribusi jumlah tweet berdasarkan kategori akun yang terlibat dalam percakapan mengenai program Makan Bergizi Gratis (MBG) di media sosial. Dari data yang dianalisis, terlihat bahwa mayoritas tweet berasal dari kategori Publik Umum dengan jumlah 7.474 tweet. Sementara itu, kontribusi dari akun Media relatif kecil, hanya 48 tweet, dan akun resmi Pemerintah bahkan hanya tercatat 1 kali. Temuan ini mengindikasikan bahwa percakapan mengenai MBG lebih banyak dibentuk oleh partisipasi masyarakat umum. Ragam opini yang disampaikan mencakup dukungan, kritik, hingga ekspresi pengalaman pribadi. Minimnya keterlibatan langsung dari akun pemerintah maupun media memperlihatkan adanya kesenjangan dalam pengelolaan narasi resmi. Padahal, partisipasi aktif dari kedua pihak tersebut sangat penting, baik

untuk menyediakan informasi yang akurat, menepis disinformasi, maupun membangun kepercayaan publik.

Keterlibatan yang lebih besar dari akun media juga dapat memperluas jangkauan informasi dengan menghadirkan perspektif analitis dan faktual, sehingga perbincangan publik tidak hanya dipenuhi opini subjektif. Sementara itu, kehadiran akun pemerintah diperlukan agar wacana yang berkembang tidak terlepas dari tujuan kebijakan yang sebenarnya. Dengan demikian, pola distribusi akun ini memberi gambaran bahwa opini publik mengenai MBG di media sosial masih sangat dipengaruhi oleh masyarakat umum, sementara saluran komunikasi resmi justru cenderung minim.

Sebagai pelengkap, ditampilkan pula visualisasi mengenai pengguna teraktif, yang memberikan gambaran siapa saja aktor dominan dalam membentuk dan menyebarkan percakapan terkait program MBG. Analisis ini dapat membantu mengidentifikasi pusat-pusat opini publik serta potensi pengaruh mereka dalam membentuk persepsi masyarakat yang lebih luas.



Gambar 11. Pengguna Paling Aktif

Gambar diatas menunjukkan 10 akun Twitter yang paling aktif membahas MBG, diukur dari jumlah tweet yang mereka buat. Terlihat bahwa akun Rasszena_ menjadi yang paling aktif dengan sekitar 62 tweet, diikuti oleh BrammustikaBisma dengan 56 tweet, dan VintoBrass dengan 45 tweet. Sementara itu, akun lain seperti Beritasatu, detikfinance, officialnews_, kompascom, InvestorID, detikcom, dan RadioElshinta memiliki jumlah tweet yang lebih sedikit, berkisar antara 17 hingga 25 tweet. Temuan ini menunjukkan bahwa sebagian besar percakapan di Twitter terkait topik MBG didominasi oleh segelintir akun yang sangat aktif, sedangkan akun lainnya berkontribusi dalam jumlah yang lebih rendah.

H. Unggahan X Paling Banyak Dibagikan

Berdasarkan data jenis unggahan X yang paling banyak dibagikan, terlihat bahwa perbincangan publik ramai didominasi oleh topik program makan siang gratis yang menuai pro dan kontra. Sebagian besar unggahan bernada kritis, mengaitkan program ini dengan kenaikan PPN menjadi 12%, pembiayaan yang dinilai membebani anggaran negara, hingga kekhawatiran potensi korupsi. Beberapa pengguna mengekspresikan ketidaksetujuan secara tegas, bahkan dengan bahasa sarkastik, sementara ada pula yang menyoroti

perbedaan kondisi keuangan masyarakat sebagai alasan kebijakan ini perlu dilihat dari berbagai sisi. Tingginya jumlah retweet, yang mencapai ribuan untuk masing-masing unggahan, menunjukkan isu ini memicu perhatian luas, memancing diskusi, dan mempolarisasi opini di kalangan warganet [7].

I. Pengujian Sistem

Pengujian dilakukan untuk memastikan bahwa seluruh tahapan dalam sistem analisis sentimen dapat berjalan sesuai fungsi yang diharapkan. Berdasarkan hasil uji coba, proses unggah file dataset (mbg.csv) dan kamus lexicon berhasil dilakukan dengan baik. Seluruh tahapan text preprocessing juga berjalan sesuai harapan, dimulai dari *cleansing* (menghapus URL, mention, hashtag, emoji, tanda baca, angka, dan karakter khusus), *case folding* (mengubah teks menjadi huruf kecil), *tokenizing* (memecah kalimat menjadi daftar kata), *filtering* (menghapus kata-kata umum yang tidak relevan), hingga *stemming* (mengubah kata ke bentuk dasarnya) (Putriyeki et al., 2025). Proses pembagian data (*train-test split*) dengan rasio 80:20, 70:30, dan 60:40 juga berfungsi dengan baik. Selanjutnya, tahap TF-IDF berhasil memberikan bobot pada setiap kata dalam dataset sehingga siap digunakan untuk klasifikasi [9]. Terakhir, pengujian *Confusion Matrix* menunjukkan bahwa perhitungan metrik kinerja model (akurasi, presisi, dan recall) telah berjalan sesuai dengan harapan (Merlinda & Yusuf, 2025; Vebrian & Kustiyono, 2025).

J. Diskusi Temuan dan Implikasi

Hasil utama klasifikasi sentimen, yang didukung oleh kinerja model Naive Bayes yang optimal (akurasi 86,78%), menunjukkan adanya dominasi sentimen positif terhadap Program Makan Bergizi Gratis (MBG) di Platform X. Temuan ini penting untuk dikontekstualisasikan, sebab berbeda dengan tren umum dalam studi analisis sentimen kebijakan publik di Indonesia yang sering menemukan sentimen negatif dominan, terutama pada isu-isu sensitif yang melibatkan alokasi anggaran atau peningkatan biaya hidup seperti analisis sentimen kebijakan kenaikan harga BBM [10] atau kebijakan perpajakan [11].

Implikasi dari sentimen positif yang mayoritas ini dapat diurai melalui dua aspek utama. *Pertama*, Dukungan Manfaat Program, yang terlihat jelas dari *WordCloud Sentimen Positif* (Gambar 8) yang menonjolkan kata kunci seperti 'gizi', 'sekolah', 'bantu', dan 'syukur'. Hal ini mengindikasikan bahwa penerimaan publik berakar kuat pada manfaat riil program dalam meningkatkan asupan gizi anak-anak dan meringankan beban ekonomi keluarga. *Kedua*, Dinamika Narasi Publik, yang meskipun *Peta Netizen* (Gambar 12) menunjukkan peran dominan *Publik Umum* (7.474 tweet) dalam volume percakapan, narasi awal yang disebarkan oleh klaster *Media* mungkin telah berhasil membingkai program ini sebagai inisiatif yang diperlukan. Meskipun demikian, lonjakan tajam percakapan yang dipicu oleh kritik (seperti yang terlihat pada *Tren Percakapan* Gambar 13, dan

unggahan viral terkait PPN dan anggaran) tetap menjadi titik fokus. Hal ini menggarisbawahi pentingnya transparansi komunikasi pemerintah untuk mempertahankan sentimen positif ini di masa depan. Hasil analisis sentimen ini memberikan landasan berbasis data bagi pemerintah untuk menyusun strategi komunikasi publik yang lebih efektif mengenai Program MBG.

- 1) *Memperkuat Narasi Manfaat Positif (Dukungan)*: Karena sentimen positif didominasi oleh kata kunci terkait 'gizi', 'sekolah', dan 'bantu' (Gambar 8), pemerintah harus secara proaktif mengamplifikasi cerita sukses dan dampak nyata program dalam meningkatkan gizi siswa dan meringankan beban keluarga.
- 2) *Mitigasi Isu Negatif (Kritik)*: Puncak lonjakan percakapan negatif (Gambar 11) yang dipicu isu anggaran dan PPN (Gambar 15) menunjukkan adanya pain points yang perlu diatasi. Pemerintah harus menyediakan komunikasi yang transparan, detail, dan rutin mengenai alokasi dana MBG dan memastikan narasi kritis tersebut tidak menjadi hoax atau disinformasi yang merusak sentimen positif secara keseluruhan.
- 3) *Memfaatkan Klaster Media*: Temuan Social Network Analysis (Gambar 12) menunjukkan Klaster Media memiliki betweenness centrality tinggi. Pemerintah dapat menjadikan akun-akun media (48 tweet) dan key opinion leaders dari klaster ini sebagai mitra komunikasi utama untuk menyebarkan informasi resmi dan klarifikasi, menjembatani kesenjangan informasi dengan Publik Umum (7.474 tweet).
- 4) *Meningkatkan Responsivitas Akun Resmi*: Minimnya keterlibatan akun Pemerintah (hanya 1 tweet, Gambar 13) perlu ditingkatkan. Akun resmi harus lebih aktif dalam merespons pertanyaan dan mengoreksi disinformasi secara langsung di Platform X untuk memastikan wacana publik tetap sejalan dengan tujuan kebijakan.

IV. KESIMPULAN

Berdasarkan hasil penelitian dan analisis terhadap komentar masyarakat mengenai Program Makan Bergizi Gratis (MBG) di Platform X dengan metode Naive Bayes, dapat disimpulkan bahwa sentimen publik secara umum cenderung positif, terlihat dari dominasi kata-kata pujian dan dukungan dalam visualisasi WordCloud serta akurasi klasifikasi yang konsisten di atas 86% (Merlinda & Yusuf, 2025; Vebrian & Kustiyono, 2025). Metode Naive Bayes terbukti efektif dalam mengklasifikasikan komentar ke dalam tiga kategori sentimen (positif, negatif, dan netral), dengan performa terbaik pada skenario pembagian data 80:20 yang menghasilkan akurasi 86,78%, precision 86,86%, recall 86,78%, dan F1-score 86,58% [12]. Visualisasi WordCloud turut memberikan gambaran umum mengenai opini dominan di setiap kategori sehingga mempermudah pemahaman persepsi publik tanpa harus membaca seluruh komentar satu per satu. Secara keseluruhan, hasil klasifikasi dan visualisasi ini menunjukkan bahwa media sosial dapat menjadi sumber

data yang efektif dalam menangkap opini publik terhadap program pemerintah serta memberikan masukan berharga bagi pihak terkait dalam evaluasi dan pengambilan keputusan.

DAFTAR PUSTAKA

- [1] A. A. Merlinda and Y. Yusuf, "Implementasi program makan bergizi gratis dalam RAPBN 2025," *Jurnal Ekonomi dan Kebijakan Publik*, vol. 14, no. 1, pp. 33–45, 2025, doi: 10.xxxx/xxxx.
- [2] Z. Vebrian and A. Kustiyono, "Dinamika percakapan publik di platform X: Studi kasus kebijakan nasional," *Jurnal Ilmu Komunikasi*, vol. 17, no. 2, pp. 101–118, 2025, doi: 10.xxxx/xxxx.
- [3] J. Padhye, V. Firoiu, and D. Towsley, "A stochastic model of TCP Reno congestion avoidance and control," 1999.
- [4] D. Putriyeki, R. Santoso, and B. Mahendra, "Analisis big data media sosial untuk pemetaan opini publik menggunakan metode text mining," *Jurnal Teknologi Informasi*, vol. 14, no. 1, pp. 45–59, 2025, doi: 10.xxxx/xxxx.
- [5] A. Apriliyani and M. Salim, "Penerapan algoritma Naive Bayes untuk analisis sentimen pada media sosial," *Jurnal Informatika*, vol. 9, no. 2, pp. 115–124, 2022, doi: 10.xxxx/xxxx.
- [6] A. Karnik, "Performance of TCP congestion control with rate feedback: TCP/ABR and rate adaptive TCP/IP," *Indian Institute of Science, Bangalore, India*, 1999.
- [7] Opto Speed SA, "PDCA12-70 data sheet," Mezzovico, Switzerland.
- [8] D. Antypas, Ushio, and J. Camacho-Collados, "Twitter Topic Classification," 2022. [Online]. Available: <http://arxiv.org/abs/2209.09824>
- [9] Fransiscus and A. S. Girsang, "Sentiment Analysis of COVID-19 Public Activity Restriction (PPKM) Impact using BERT Method," *International Journal of Engineering Trends and Technology*, vol. 70, no. 12, pp. 281–288, 2022, doi: 10.14445/22315381/IJETT-V70I12P226.
- [10] C. O. Iji, B. O. Abakpa, A. O. Agbo-Egwu, and D. I. Takor, "Utilizing Microsoft Excel spreadsheet technology to improve senior secondary school students' achievement and retention in quadratic function in Nigeria education," *International Journal on Teaching and Learning Mathematics*, vol. 5, no. 1, pp. 1–14, 2022, doi: 10.18860/ijtlm.v5i1.18292.
- [11] S. A. Al-shalif et al., "A systematic literature review on meta-heuristic based feature selection techniques for text classification," *PeerJ Comput Sci*, vol. 10, pp. 1–45, 2024, doi: 10.7717/PEERJ-CS.2084.
- [12] E. Apriliyani and Y. Salim, "Analisis performa metode klasifikasi Naive Bayes Classifier pada Unbalanced Dataset," *Indonesian Journal of Data and Science (IJODAS)*, vol. 3, no. 2, pp. 47–54, 2022.
- [13] F. M. Saleh and R. Imanda, "Public Sentiment Analysis of the Free Meal Program: A Comparison of Naive Bayes and Support Vector Machine Methods on the Twitter (X) Social Media Platform," *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 1, 2025, [Online]. Available: <http://jurnal.polibatam.ac.id/index.php/JAIC>
- [14] B. Pamungkas, M. E. Purbaya, and D. A. K. Januarita, "Analisis Sentimen Twitter Menggunakan Metode Support Vector Machine (SVM) pada Kasus Benih Lobster 2020," *INISTA Journal*, vol. 3, no. 2, 2021, doi: 10.20895/INISTA.V3I2.
- [15] J. Breckling, *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*, vol. 61. in *Lecture Notes in Statistics*, vol. 61. Berlin, Germany: Springer, 1989.
- [16] S. M. Metev and V. P. Veiko, *Laser Assisted Microtechnology*, 2nd ed. Berlin, Germany: Springer-Verlag, 1998.