Perbandingan Kinerja Metode Naive Bayes dan Support Vector Machine (SVM) dalam Analisis Sentimen Publik terhadap Program Makan Bergizi Gratis (MBG)

[Performance Comparison of Naive Bayes and Support Vector Machine (SVM) Methods in Public Sentiment Analysis of the Free Nutritious Meal Program (MBG)]

Mohamad Yachob Oktavian[1)], Yulian Findawati[2*)], Suhendro Busono[3)], Uce Indahyanti[4)]

[1)] *Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia*
[2)] *Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia*
[3)] *Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia*
[4)] *Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia*

*Email Penulis Korespondensi: yulianfindawati@umsida.ac.id

Abstract. *The Free Nutritious Meal Program (MBG) launched by the Indonesian government has triggered diverse public discourse on social media. This study compares Naive Bayes and Support Vector Machine (SVM) in classifying public sentiment toward the MBG program using 1,050 valid Indonesian-language tweets collected via web scraping during January–December 2025. Text preprocessing included case folding, cleaning, tokenizing, stopword removal, and stemming using Sastrawi, followed by TF-IDF feature extraction, yielding 2,920 features. The dataset was split into 840 training and 210 testing samples. Results show Naive Bayes achieved slightly higher scores than SVM across all metrics, with 70.48% accuracy, 71.14% precision, 70.48% recall, and 69.73% F1-score, versus SVM's 68.57% accuracy. Naive Bayes was also far more computationally efficient, training about 52 times faster than SVM. Sentiment distribution was relatively balanced (Neutral 34.2%, Positive 33.0%, Negative 32.8%), indicating the public remained in a critical evaluation phase toward the program.*

Keywords - *sentiment analysis; Naive Bayes; Support Vector Machine; Makan Bergizi Gratis; X*

Abstrak. *Program Makan Gratis Bergizi (MBG) yang diluncurkan oleh pemerintah Indonesia telah memicu beragam wacana publik di media sosial. Studi ini membandingkan Naive Bayes dan Support Vector Machine (SVM) dalam mengklasifikasikan sentimen publik terhadap program MBG menggunakan 1.050 tweet berbahasa Indonesia yang valid yang dikumpulkan melalui web scraping selama Januari–Desember 2025. Praproses teks meliputi case folding, pembersihan, tokenisasi, penghapusan stopword, dan stemming menggunakan Sastrawi, diikuti oleh ekstraksi fitur TF-IDF, menghasilkan 2.920 fitur. Dataset dibagi menjadi 840 sampel pelatihan dan 210 sampel pengujian. Hasil menunjukkan Naive Bayes mencapai skor sedikit lebih tinggi daripada SVM di semua metrik, dengan akurasi 70,48%, presisi 71,14%, recall 70,48%, dan skor F1 69,73%, dibandingkan dengan akurasi SVM sebesar 68,57%. Naive Bayes juga jauh lebih efisien secara komputasi, pelatihan sekitar 52 kali lebih cepat daripada SVM. Distribusi sentimen relatif seimbang (Netral 34,2%, Positif 33,0%, Negatif 32,8%), menunjukkan bahwa publik masih berada dalam fase evaluasi kritis terhadap program tersebut.*

Kata Kunci - *analisis sentimen; Naive Bayes; Support Vector Machine; Makan Bergizi Gratis; X*

I. PENDAHULUAN

Program Makan Bergizi Gratis (MBG) diinisiasi oleh pemerintah Indonesia sebagai langkah strategis dalam menanggulangi problematika malnutrisi di kalangan populasi rentan, termasuk pelajar, ibu hamil, dan ibu menyusui. Pasca peluncuran, program ini memperoleh ragam tanggapan dari masyarakat, terutama di platform media sosial X, mulai dari apresiasi atas upaya pemerintah mengentaskan stunting, kerisauan teknis pelaksanaan, hingga kritik terhadap sejumlah kejadian di tingkat implementasi. Mengidentifikasi sudut pandang publik terhadap Program MBG menjadi esensial bagi evaluasi kebijakan, namun volume cuitan yang mencapai ribuan hingga jutaan entri membuat analisis manual menjadi tidak praktis sehingga diperlukan pendekatan machine learning untuk memproses data secara otomatis dan efektif.

Dua algoritma yang banyak digunakan dalam analisis sentimen adalah Naive Bayes, yang bekerja melalui pendekatan probabilistik, dan Support Vector Machine (SVM), yang menggunakan pendekatan geometris untuk memaksimalkan pemisahan antar kelas. Studi-studi terdahulu pada berbagai domain kebijakan publik menunjukkan hasil perbandingan yang bervariasi: pada sentimen kebijakan Ibu Kota Nusantara, SVM mencapai akurasi 94% berbanding Naive Bayes 91% [1]; pada RUU Kesehatan, SVM mencapai 87% berbanding Naive Bayes 82% [2];

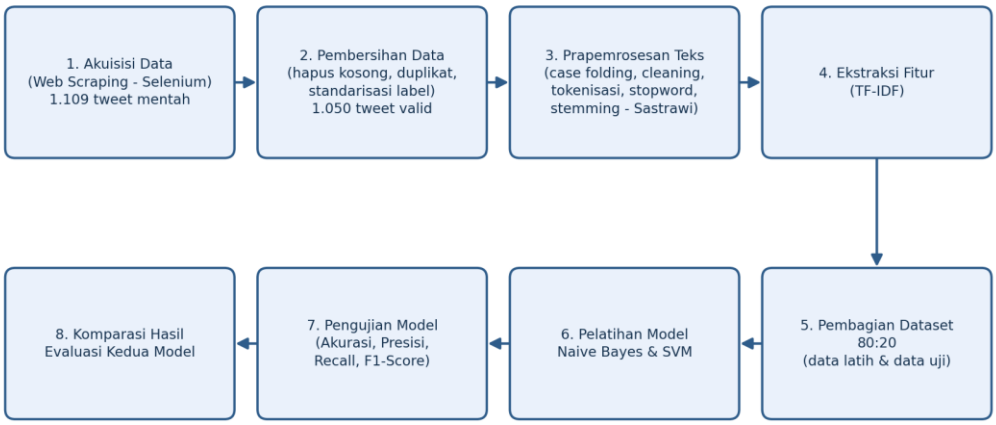
sementara pada pelayanan BPJS, selisihnya jauh lebih besar yaitu SVM 79,32% berbanding Naive Bayes 61,63% [3]. Pada topik dengan variasi bahasa tinggi seperti mobil listrik, SVM juga unggul dengan akurasi 70,82% berbanding 63,02% [4], dan pada isu COVID-19, SVM mencapai 81,6% berbanding 78,3% [5].

Beberapa penelitian sebelumnya telah secara spesifik mengkaji respons publik terhadap Program Makan Siang/Bergizi Gratis. Sitanggang dkk. [8] menerapkan algoritma Naive Bayes secara tunggal tanpa pembanding metode lain pada data media sosial X, sedangkan Arif dan Kustiyono [10] menggunakan pendekatan lexicon-based (TextBlob) pada data periode Januari–Februari 2025, yaitu masa awal implementasi program. Kedua penelitian tersebut belum melakukan perbandingan kinerja antar algoritma machine learning secara sistematis, dan cakupan datanya terbatas pada periode awal peluncuran program sehingga belum menangkap dinamika sentimen publik yang lebih matang setelah program berjalan lebih lama. Penelitian ini berupaya mengisi kesenjangan tersebut melalui tiga penekanan, yaitu: (1) perbandingan langsung dua algoritma dengan pendekatan berbeda—Naive Bayes yang bersifat probabilistik dan SVM yang bersifat geometris—pada objek kajian yang sama; (2) cakupan data selama satu tahun penuh (Januari–Desember 2025) yang mencakup fase awal hingga fase evaluasi kritis publik terhadap program, berbeda dengan penelitian [10] yang hanya mencakup dua bulan awal implementasi; dan (3) evaluasi kinerja model berbasis metrik machine learning standar (akurasi, presisi, recall, F1-Score, serta efisiensi komputasi), berbeda dengan pendekatan lexicon-based yang digunakan pada penelitian sebelumnya. Penelitian ini bertujuan melakukan perbandingan kinerja antara metode Naive Bayes dan SVM dalam mengklasifikasikan sentimen masyarakat terhadap Program MBG berdasarkan data X periode Januari–Desember 2025, dengan fokus pada metrik akurasi, presisi, recall, F1-Score, dan efisiensi komputasi, guna memberikan arahan praktis bagi pemantauan sentimen program pemerintah serupa di masa depan.

II. METODE

A. Tahapan Penelitian

Penelitian ini disusun secara bertahap, dimulai dari akuisisi data hingga komparasi performa model, sebagaimana diilustrasikan pada diagram alir Gambar 1.



Gambar 1. Diagram Alir Tahapan Penelitian

B. Data Penelitian

Data dikumpulkan dari platform X melalui teknik web scraping menggunakan library Selenium, mencakup periode Januari hingga Desember 2025. Total 1.109 tweet berhasil dikumpulkan menggunakan empat kata kunci pencarian, sebagaimana ditunjukkan pada Tabel 1. Setelah proses data cleaning yang menghilangkan 59 tweet (5,3%) berupa data dengan label kosong, tidak valid, duplikat, atau terlalu pendek, diperoleh 1.050 tweet yang valid untuk digunakan dalam pelatihan dan pengujian model.

Tabel 1. Kata Kunci Pencarian Data

No	Kata Kunci	Jumlah	Keterangan
1	“mbg”	404 (38,5%)	Singkatan program

No	Kata Kunci	Jumlah	Keterangan
2	“program makan gratis”	268 (25,5%)	Nama populer program
3	“dampak mbg”	193 (18,4%)	Diskusi dampak program
4	“kritik mbg”	185 (17,6%)	Kritik implementasi

Pelabelan sentimen dilakukan dengan tiga kategori: Positif (dukungan, apresiasi, ekspektasi positif), Negatif (kritik, sarkasme, kekecewaan), dan Netral (informatif, pertanyaan, tanpa opini eksplisit).

Proses anotasi/pelabelan dilakukan secara manual oleh satu orang anotator (peneliti sendiri) berdasarkan kriteria kategori sentimen yang telah ditetapkan sebelumnya, tanpa melibatkan multiple annotator agreement (IAA). Penggunaan anotator tunggal ini menjadi salah satu keterbatasan penelitian karena berpotensi menimbulkan bias subjektivitas dalam proses pelabelan data.

Untuk menjaga konsistensi proses pelabelan, disusun pedoman anotasi (labeling guideline) berupa definisi operasional untuk setiap kategori sentimen sebelum proses pelabelan dimulai, sebagaimana telah dijabarkan di atas: Positif untuk cuitan yang menunjukkan dukungan, apresiasi, atau ekspektasi positif; Negatif untuk cuitan yang memuat kritik, sarkasme, keluh kesah, atau kekecewaan; dan Netral untuk cuitan yang bersifat informatif, berupa pertanyaan, atau tidak mengandung opini eksplisit. Pedoman ini diterapkan secara konsisten oleh anotator yang sama pada keseluruhan 1.050 tweet untuk meminimalkan variasi interpretasi antarwaktu, meskipun potensi subjektivitas akibat penggunaan anotator tunggal tetap menjadi keterbatasan penelitian ini.

C. Model Klasifikasi

Model Multinomial Naive Bayes bekerja berdasarkan teorema Bayes dengan asumsi independensi antar fitur, sehingga sangat efisien secara komputasi dan cocok untuk klasifikasi teks berskala besar [12]. Probabilitas suatu dokumen D yang memuat fitur x_1, x_2, …, x_n untuk masuk ke dalam kelas C ditentukan berdasarkan probabilitas prior dan likelihood, dengan rumus dasar sebagai berikut:

$$P(C|X) = [P(X|C) \times P(C)] / P(X) \quad (1)$$

dengan P(C|X) adalah probabilitas kelas C diberikan fitur X (posterior), P(X|C) adalah probabilitas kemunculan fitur X pada kelas C (likelihood), P(C) adalah probabilitas prior kelas C, dan P(X) adalah probabilitas fitur X. Kelas dengan nilai probabilitas posterior tertinggi ditetapkan sebagai hasil klasifikasi sentimen dari tweet yang dianalisis.

Model Support Vector Machine (SVM) yang digunakan dalam penelitian ini menggunakan kernel Radial Basis Function (RBF) untuk mencari hyperplane optimal yang memisahkan kelas-kelas sentimen pada ruang fitur berdimensi tinggi [13]. Fungsi keputusan SVM dinyatakan sebagai:

$$f(x) = sign(w \cdot x + b) \quad (2)$$

dengan w adalah vektor bobot (weight vector), x adalah vektor fitur input berupa representasi TF-IDF dari tweet, b adalah bias, dan sign adalah fungsi tanda yang menentukan kelas prediksi. Hyperplane optimal dicari dengan memaksimalkan margin, yang dirumuskan sebagai problem optimasi berikut:

$$Minimize \tfrac{1}{2}\|w\|^2, subject\ to\ y_i(w \cdot x_i + b) \geq 1, untuk\ semua\ i \quad (3)$$

Dalam kasus data yang tidak terpisah secara linier, kernel trick digunakan untuk memindahkan data ke ruang baru yang memungkinkan pemisahan linier, dan pada penelitian ini digunakan kernel RBF yang populer untuk klasifikasi teks [13]. Kedua model dilatih menggunakan representasi vektor TF-IDF dari 840 data latih (80%) dan dievaluasi pada 210 data uji (20%) menggunakan metrik akurasi, presisi, recall, F1-Score, serta waktu komputasi (training dan testing).

III. HASIL DAN PEMBAHASAN

A. Gambaran Umum Dataset

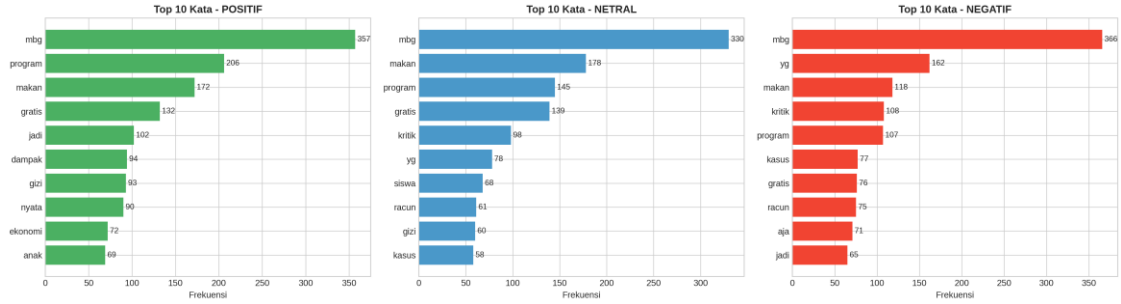
Dataset penelitian terdiri atas 1.050 tweet berbahasa Indonesia yang membahas Program MBG, dengan distribusi label yang relatif seimbang sebagaimana ditunjukkan pada Tabel 2. Keseimbangan distribusi kelas ini merupakan kondisi ideal dalam membangun model klasifikasi karena mencegah bias model terhadap kelas mayoritas.

Tabel 2. Distribusi Label Sentimen Dataset

No	Label Sentimen	Jumlah Tweet	Persentase
1	Netral	359	34,2%

No	Label Sentimen	Jumlah Tweet	Persentase
2	Positif	347	33,0%
3	Negatif	344	32,8%
Total		1.050	100%

Analisis kata dominan menunjukkan bahwa pada sentimen Positif, kata seperti “program”, “makan”, “gratis”, “gizi”, dan “dampak” mencerminkan apresiasi terhadap manfaat konkret program. Pada sentimen Negatif, kata “kritik”, “kasus”, dan “racun” mendominasi, mengindikasikan kekhawatiran masyarakat terkait insiden keamanan pangan dalam implementasi program.



Gambar 2. Sepuluh Kata dengan Frekuensi Tertinggi per Kelas Sentimen

B. Hasil Ekstraksi Fitur dan Pelatihan Model

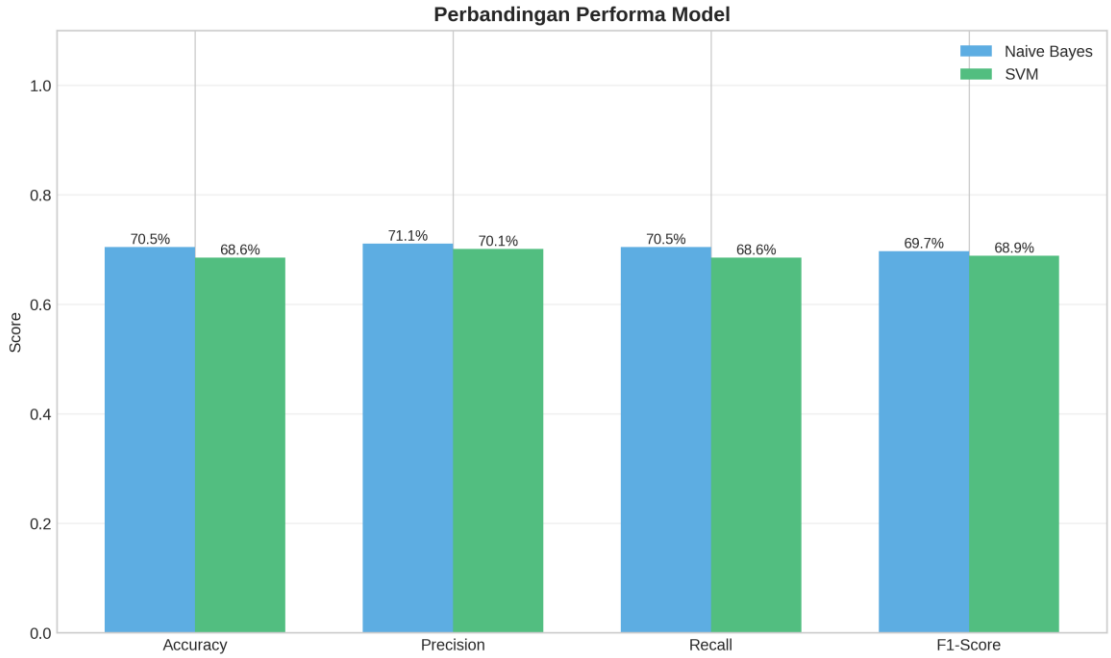
Proses ekstraksi fitur TF-IDF menghasilkan 2.920 fitur unik dari keseluruhan korpus tweet, yang kemudian dibagi menjadi 840 data latih dan 210 data uji. Model Naive Bayes dilatih dengan waktu training 0,0047 detik dan testing 0,0006 detik, jauh lebih cepat dibanding SVM yang membutuhkan waktu training 0,2447 detik dan testing 0,0371 detik (sekitar 52 kali lebih lambat). Hasil evaluasi kedua model pada data uji disajikan pada Tabel 3.

Tabel 3. Perbandingan Metrik Evaluasi Naive Bayes dan SVM

Metrik	Naive Bayes	SVM	Selisih (NB−SVM)
Akurasi	70,48%	68,57%	+1,91%
Presisi	71,14%	70,10%	+1,04%
Recall	70,48%	68,57%	+1,91%
F1-Score	69,73%	68,87%	+0,86%
Waktu Training	0,0047 dtk	0,2447 dtk	NB 52× lebih cepat
Waktu Testing	0,0006 dtk	0,0371 dtk	NB 62× lebih cepat

Berdasarkan Tabel 3, Naive Bayes memperoleh nilai evaluasi sedikit lebih tinggi dibandingkan SVM pada seluruh metrik yang diuji, dengan selisih akurasi sebesar 1,91 poin persentase. Dari sisi efisiensi komputasi, Naive Bayes jauh lebih unggul dengan waktu training 52 kali lebih cepat dan testing 62 kali lebih cepat dibandingkan SVM.

Perlu dicermati bahwa selisih akurasi sebesar 1,91 poin persentase tersebut tergolong kecil dan diperoleh dari satu kali pembagian data uji (210 tweet), sehingga belum tentu mencerminkan keunggulan yang signifikan secara statistik. Selisih sekecil ini berpotensi dipengaruhi oleh variasi acak pada pembagian data (data-split randomness) atau karakteristik spesifik sampel uji yang digunakan, dan bukan semata-mata karena superioritas algoritmik Naive Bayes atas SVM. Penelitian ini tidak melakukan pengujian signifikansi statistik (misalnya paired t-test atau McNemar’s test) maupun validasi silang berulang untuk mengonfirmasi konsistensi selisih tersebut pada pembagian data yang berbeda, sehingga kesimpulan keunggulan Naive Bayes pada penelitian ini sebaiknya dimaknai sebagai keunggulan praktis pada skenario pengujian spesifik ini. Keunggulan yang lebih meyakinkan justru terlihat pada aspek efisiensi komputasi, di mana Naive Bayes secara konsisten 52–62 kali lebih cepat dibandingkan SVM, sehingga faktor ini menjadi pertimbangan yang lebih kuat dibandingkan selisih akurasi semata dalam memilih Naive Bayes untuk kebutuhan pemantauan sentimen program pemerintah secara praktis.



Gambar 3. Perbandingan Skor Akurasi, Presisi, Recall, dan F1-Score

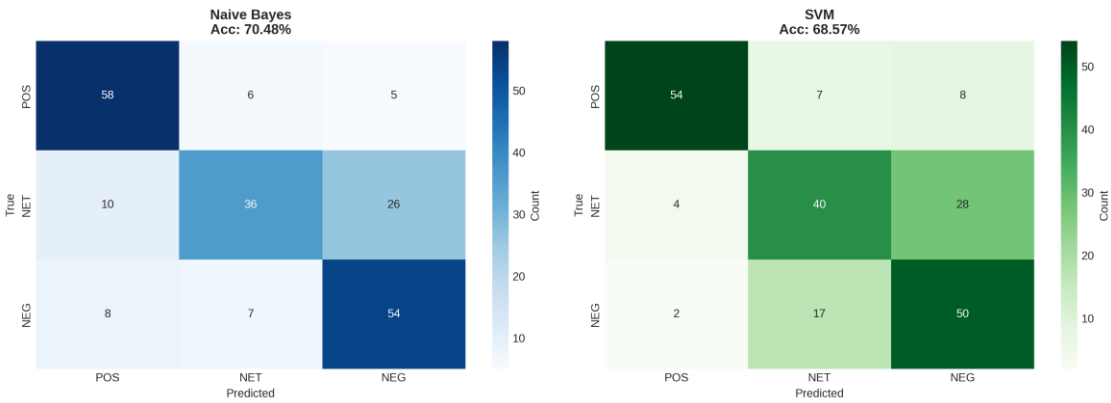
C. Analisis per Kelas Sentimen dan Confusion Matrix

Evaluasi lebih mendalam pada Tabel 4 menunjukkan pola yang menarik: pada kelas Positif, SVM memiliki presisi lebih tinggi (90,00% berbanding 76,32%), namun Naive Bayes unggul pada recall (84,06% berbanding 78,26%). Pada kelas Netral, Naive Bayes lebih baik pada seluruh metrik, sementara kelas ini menjadi kelas dengan performa terendah pada kedua model karena karakteristiknya yang tidak memiliki polaritas emosional jelas. Pada kelas Negatif, Naive Bayes unggul pada presisi dan F1-Score.

Tabel 4. Hasil Evaluasi per Kelas Sentimen

Kelas	NB Precision	NB Recall	NB F1	SVM Precision	SVM Recall	SVM F1
Positif	76,32%	84,06%	80,00%	90,00%	78,26%	83,72%
Netral	73,47%	50,00%	59,50%	62,50%	55,56%	58,82%
Negatif	63,53%	78,26%	70,13%	58,14%	72,46%	64,52%

Analisis confusion matrix menunjukkan pola kesalahan paling menonjol pada kedua model adalah misklasifikasi kelas Netral menjadi Negatif: Naive Bayes salah mengklasifikasikan 26 dari 72 tweet Netral sebagai Negatif, sedangkan SVM mencatat 28 kesalahan serupa. Hal ini mengindikasikan tumpang tindih fitur linguistik antara kelas Netral dan Negatif, terutama karena tweet netral sering membahas isu negatif (seperti kasus keracunan) secara deskriptif tanpa ekspresi emosi eksplisit. Kedua model menunjukkan performa terbaik pada kelas Positif, dengan Naive Bayes mengklasifikasikan 58 dari 69 tweet Positif dengan benar (84,06%) dan SVM 54 dari 69 (78,26%).

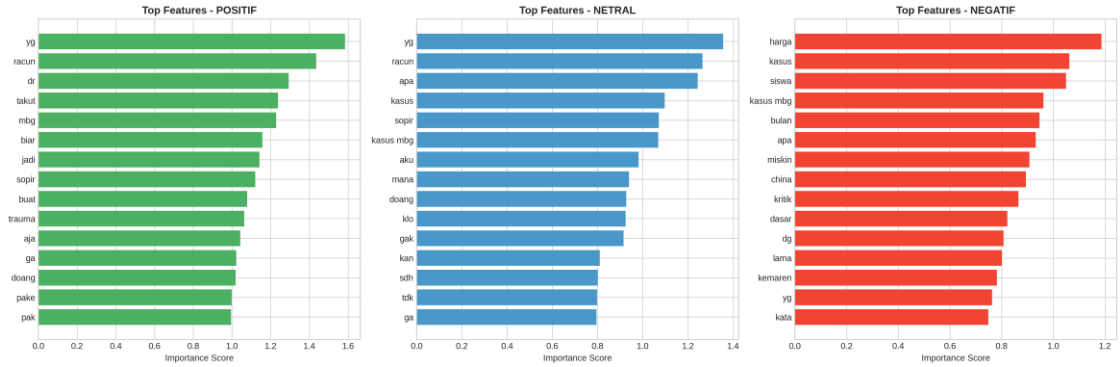


Gambar 4. Confusion Matrix Naive Bayes dan SVM pada Data Uji

D. Pembahasan

Distribusi sentimen yang relatif seimbang (Netral 34,2%, Positif 33,0%, Negatif 32,8%) mengindikasikan bahwa diskusi publik terkait Program MBG selama tahun 2025 bersifat heterogen, dengan masyarakat masih berada pada tahap evaluasi kritis terhadap program yang baru berjalan. Temuan ini berbeda dengan penelitian [10] yang menemukan dominasi sentimen negatif sebesar 71,8% pada data Januari–Februari 2025, kemungkinan karena periode awal implementasi cenderung menuai lebih banyak kritik dibanding periode yang lebih panjang.

Temuan utama penelitian ini menunjukkan Naive Bayes memperoleh nilai evaluasi sedikit lebih tinggi dibandingkan SVM pada seluruh metrik yang diuji, berbeda dengan mayoritas penelitian terdahulu yang menempatkan SVM sebagai algoritma unggul [2], [3], [6]. Beberapa faktor yang dapat menjelaskan hal ini meliputi: (1) distribusi kelas yang sangat seimbang (33–34% per kelas) menguntungkan Naive Bayes yang bekerja berdasarkan prior probability; (2) ukuran dataset yang moderat (840 data latih) merupakan kondisi di mana Naive Bayes umumnya bersaing ketat dengan SVM; dan (3) karakteristik bahasa Indonesia informal di media sosial yang bervariasi tinggi dapat memengaruhi konsistensi SVM dalam menemukan hyperplane optimal, sementara pendekatan probabilistik Naive Bayes lebih toleran terhadap variasi tersebut.



Gambar 5. Importance Score Fitur SVM Teratas per Kelas Sentimen

Hasil analisis feature importance pada model SVM (Gambar 5) menunjukkan bahwa kata seperti “racun” dan “kasus mbg” menjadi penanda kuat pada kelas Negatif dan Netral, sementara kata “takut” dan “trauma” justru berkontribusi pada kelas Positif, mengindikasikan adanya konteks pemulihan/relief setelah insiden yang sempat dikhawatirkan. Pola ini memperkuat temuan pada Gambar 2 bahwa isu keamanan pangan (kasus keracunan) menjadi tema sentral yang melintasi ketiga kelas sentimen, sehingga model perlu menangkap konteks kalimat secara lebih dalam agar tidak salah mengelompokkan tweet deskriptif sebagai sentimen negatif.

Beberapa keterbatasan penelitian ini meliputi ukuran dataset yang relatif terbatas (1.050 tweet) dibandingkan penelitian sejenis [3] yang menggunakan 3.503 tweet, metode pelabelan manual yang rentan subjektivitas, penggunaan parameter default SVM tanpa hyperparameter tuning ekstensif, serta tidak digunakannya validasi silang k-fold yang dapat memberikan estimasi performa lebih robust. Skema pembagian data tunggal (single hold-out split) 80:20 dipilih dalam penelitian ini karena penelitian ini diposisikan sebagai studi komparasi awal (preliminary comparative study) antara Naive Bayes dan SVM pada kasus Program MBG, mengacu pada pendekatan yang lazim digunakan pada sebagian besar penelitian sejenis [1]–[6]. Konsekuensinya, sebagaimana didiskusikan pada bagian sebelumnya, hasil evaluasi pada skema ini lebih rentan terhadap variasi akibat karakteristik satu partisi data tertentu, sehingga validasi silang k-fold menjadi saran pengembangan yang relevan untuk penelitian selanjutnya guna memperoleh estimasi performa yang lebih stabil dan dapat diandalkan.

IV. KESIMPULAN

Penelitian ini berhasil mengimplementasikan algoritma Naive Bayes dan SVM untuk mengklasifikasikan sentimen publik terhadap Program MBG ke dalam tiga kategori menggunakan representasi fitur TF-IDF dari 2.920 kosakata yang diekstraksi dari 1.050 tweet. Perbandingan kinerja menunjukkan bahwa Naive Bayes memperoleh nilai evaluasi sedikit lebih tinggi dibandingkan SVM pada seluruh metrik yang diuji, dengan akurasi 70,48% berbanding 68,57%, serta jauh lebih efisien secara komputasi (52 kali lebih cepat dalam waktu training). Distribusi sentimen publik terhadap Program MBG relatif seimbang antara Netral, Positif, dan Negatif, dengan kelas Netral menjadi kategori yang paling sulit diklasifikasikan oleh kedua model akibat tumpang tindih fitur linguistik dengan kelas Negatif. Hasil penelitian ini memberikan arahan praktis bahwa Naive Bayes dapat menjadi pilihan yang efektif dan efisien untuk pemantauan sentimen program kebijakan pemerintah serupa di masa depan, khususnya pada kondisi distribusi kelas

yang seimbang dan ukuran dataset moderat. Penelitian selanjutnya disarankan memperluas ukuran dataset, menerapkan validasi silang k-fold untuk memperoleh estimasi performa yang lebih robust, melakukan optimasi hyperparameter SVM melalui Grid Search, serta mengeksplorasi pendekatan deep learning seperti IndoBERT untuk meningkatkan akurasi klasifikasi sentimen berbahasa Indonesia. Selain itu, proses anotasi pada penelitian selanjutnya disarankan melibatkan minimal tiga orang anotator untuk menghitung tingkat kesepakatan antaranotator (inter-annotator agreement) sehingga bias subjektivitas dalam pelabelan data dapat diminimalkan, serta memperluas variasi kata kunci pencarian data agar cakupan tweet yang diperoleh lebih banyak dan lebih bervariatif.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Program Studi Informatika, Fakultas Sains dan Teknologi, Universitas Muhammadiyah Sidoarjo atas dukungan yang diberikan selama proses penelitian dan penulisan artikel ini.

REFERENSI

[1] A. Supian, B. Tri Revaldo, N. Marhadi, L. Efrizoni, and R. Rahmaddeni, “Perbandingan Kinerja Naive Bayes Dan SVM Pada Analisis Sentimen Twitter Ibukota Nusantara,” J. Ilm. Inform., vol. 12, no. 01, pp. 15–21, 2024.

[2] T. Widyanto, I. Ristiana, and A. Wibowo, “Komparasi Naive Bayes dan SVM Analisis Sentimen RUU Kesehatan di Twitter,” SINTECH, vol. 6, no. 3, pp. 147–161, 2023.

[3] F. Amandasari and Damayanti, “Perbandingan Kinerja Support Vector Machine dan Naive Bayes dalam Klasifikasi Sentimen Twitter Terhadap Pelayanan BPJS,” J. Pendidik. dan Teknol. Indones., vol. 5, no. 3, pp. 645–653, 2025.

[4] W. Ningsih, B. Alfianda, and D. Wulandari, “Comparison of Naive Bayes and SVM Algorithms in Twitter Sentiment Analysis on Electric Car Use in Indonesia,” MALCOM, vol. 4, no. 2, pp. 556–562, 2024.

[5] H. Sujadi, S. Fajar, and C. Roni, “Analisis Sentimen Pengguna Media Sosial Twitter terhadap Wabah Covid-19 dengan Metode Naive Bayes Classifier dan Support Vector Machine,” Infotech, vol. 8, no. 1, pp. 22–27, 2022.

[6] F. Matheos Sarimole, “Analisis Sentimen Terhadap Aplikasi Satu Sehat Pada Twitter Menggunakan Algoritma Naive Bayes Dan Support Vector Machine,” J. Sains dan Teknol., vol. 5, no. 3, pp. 1–8, 2024.

[7] P. Syahputra and R. Kurniawan, “Analisis Sentimen Terhadap Kesehatan Mental Remaja Menggunakan Metode Naive Bayes,” J. Inf. Syst. Res., vol. 5, no. 4, pp. 1216–1224, 2024.

[8] A. Sitanggang, Y. Umaidah, and R. I. Adam, “Analisis Sentimen Masyarakat Terhadap Program Makan Siang Gratis Pada Media Sosial X Menggunakan Algoritma Naive Bayes,” J. Inform. dan Tek. Elektro Terap., vol. 12, no. 3, 2024.

[9] N. P. G. Naraswati, D. C. Rosmilda, D. Desinta, F. Khairi, R. Damaiyanti, and R. Nooraeni, “Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification,” J. Sist. Inf., vol. 10, no. 1, pp. 228–238, 2021.

[10] M. W. Arif and K. Kustiyono, “Analisis Sentimen Kebijakan Makan Bergizi Gratis di Media Sosial Menggunakan Natural Language Processing Berbasis Python TextBlob di Indonesia,” J. Pendidik. dan Teknol. Indones., vol. 5, no. 9, pp. 2463–2471, 2025.

[11] N. A. Maulana and D. Darwis, “Perbandingan Metode SVM dan Naive Bayes untuk Analisis Sentimen pada Twitter tentang Obesitas di Kalangan Gen Z,” J. Pendidik. dan Teknol. Indones., vol. 5, no. 3, pp. 655–666, 2025.

[12] Y. Findawati, I. R. I. Astutik, A. S. Fitroni, I. Indrawati, and N. Yuniasih, “Comparative analysis of Naïve Bayes, K Nearest Neighbor and C.45 method in weather forecast,” J. Phys. Conf. Ser., vol. 1402, no. 6, 2019, doi: 10.1088/1742-6596/1402/6/066046.

[13] Ratih Puspitasari, Y. Findawati, and M. A. Rosid, “Sentiment Analysis of Post-COVID-19 Inflation Based on Twitter Using the K-Nearest Neighbor and Support Vector Machine Classification Methods,” J. Tek. Inform., vol. 4, no. 4, pp. 669–679, 2023, doi: 10.52436/1.jutif.2023.4.4.801.

Conflict of Interest Statement:
The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.