

Sistem Monitoring Sentimen Publik Real-Time Program MBG Berbasis Machine Learning: Studi Komparatif SVM dan Naïve Bayes di Platform X

Azzahra Azka Tsaqofa^{1*}, Sri Bakti Handayani², Leni Nurhayati³

^{1 2 3} Sistem Informasi, Fakultas Teknologi Informasi, Universitas Sebelas April, Indonesia

*Corresponding email: tsakofa21@gmail.com

Abstrak: Program Makan Bergizi Gratis (MBG) menimbulkan berbagai respons masyarakat pada media sosial X. Penelitian ini bertujuan membandingkan kinerja algoritma Support Vector Machine (SVM) dan Naïve Bayes dalam melakukan klasifikasi sentimen publik terhadap Program MBG. Metode penelitian menggunakan CRISP-DM yang meliputi business understanding, data understanding, data preparation, modeling, evaluation, dan deployment. Data diperoleh melalui web scraping menggunakan Tweet Harvest dengan kata kunci terkait MBG sebanyak 4.754 tweet pada periode 20 Mei–25 Juni 2026. Data diproses melalui tahapan text mining, preprocessing teks, dan ekstraksi fitur TF-IDF. Hasil analisis menunjukkan distribusi sentimen yaitu netral sebesar 41,9%, negatif sebesar 40,5%, dan positif sebesar 17,5%. Evaluasi model menghasilkan akurasi SVM sebesar 85,04% dan Naïve Bayes sebesar 83,66%. Berdasarkan hasil evaluasi, SVM dipilih sebagai model terbaik karena menghasilkan performa klasifikasi yang lebih seimbang. Penelitian ini memberikan kontribusi dalam pemanfaatan machine learning untuk mendukung pemantauan opini publik terhadap program pemerintah berbasis data.

Kata kunci: Analisis Sentimen, Program MBG, Support Vector Machine, Naïve Bayes, Text Mining.

Abstract: The implementation of the Free Nutritious Meal Program (MBG) has generated various public responses on social media X. This study aims to compare the performance of Support Vector Machine (SVM) and Naive Bayes algorithms for classifying public sentiment toward the MBG program. The research applies the CRISP-DM methodology consisting of business understanding, data understanding, data preparation, modeling, evaluation, and deployment. Data were collected using Tweet Harvest web scraping from platform X with MBG-related keywords, producing 4,754 tweet data collected during May 20–June 25, 2026. The data were processed through text mining stages, preprocessing, and TF-IDF feature extraction. The classification results show sentiment distribution of 41.9% neutral, 40.5% negative, and 17.5% positive. Model evaluation shows that SVM achieved an accuracy of 85.04%, while Naive Bayes achieved 83.66%. Based on the evaluation results, SVM was selected as the better model because it provides more balanced classification performance. This research demonstrates that machine learning-based sentiment analysis can support monitoring of public perception toward government programs.

Keywords: Sentiment Analysis, MBG Program, Support Vector Machine, Naive Bayes, Text Mining, Social Media X

PENDAHULUAN

Program Makan Bergizi Gratis (MBG) merupakan program prioritas nasional yang diluncurkan pada 6 Januari 2025 oleh Pemerintahan Presiden Prabowo Subianto sebagai bagian dari upaya mewujudkan visi Indonesia Emas 2045 (Kiftiyah, 2025). Program yang diselenggarakan oleh Badan Gizi Nasional (BGN) ini bertujuan untuk mengatasi masalah kekurangan gizi, menurunkan prevalensi stunting, meningkatkan kualitas sumber daya manusia, serta meringankan beban ekonomi masyarakat (Perpres No. 83, 2024). Dengan pagu anggaran sebesar Rp71 triliun pada tahun 2025 dan target penerima manfaat mencapai 82,9 juta orang, program ini merupakan salah satu inisiatif pemerintah dengan skala terbesar dalam sejarah Indonesia. Dalam proses implementasinya, program ini memperoleh beragam tanggapan masyarakat yang disampaikan melalui media sosial X.

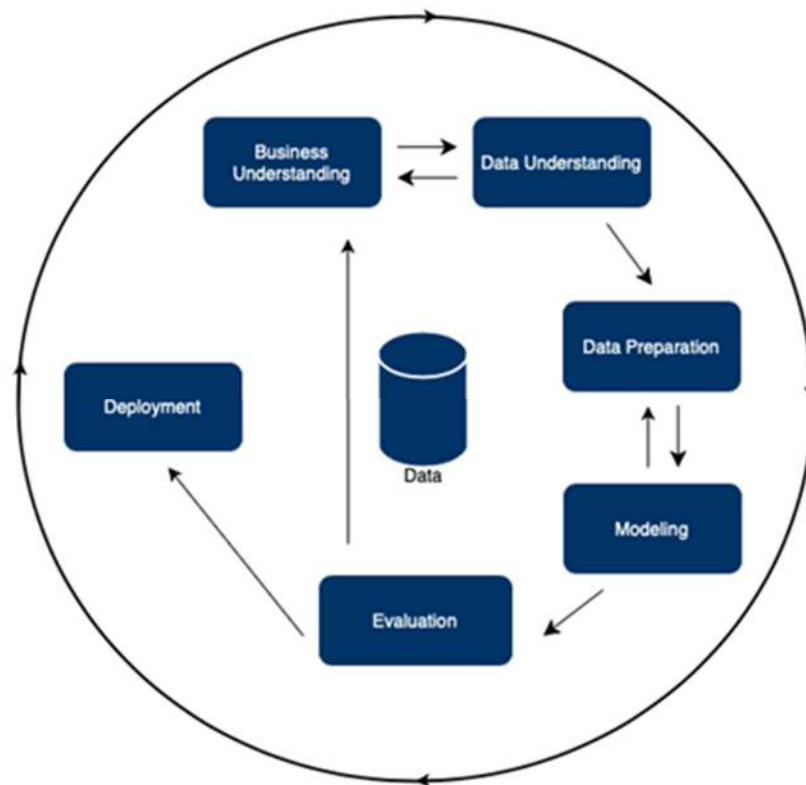
Di era digital saat ini, media sosial khususnya platform X (sebelumnya Twitter) telah menjadi ruang publik strategis untuk memahami opini publik terhadap program pemerintah (Harwenda et al., 2025). Platform X memiliki 586 juta pengguna aktif bulanan secara global, dengan 25,2 juta pengguna di Indonesia (Singh, 2025). Karakteristik platform ini, seperti format pesan singkat (maksimal 280 karakter), mekanisme retweet untuk penyebaran informasi cepat, serta penggunaan hashtag untuk mengelompokkan diskusi, menjadikannya

sumber data yang kaya dan real-time untuk memahami sentimen publik (Mulyono et al., 2023; Schröer et al., 2021; Wirth & Hipp, 2000).

Karakteristik media sosial X yang menyediakan percakapan secara cepat dan terbuka menjadikannya sumber data penting untuk memahami persepsi publik secara real-time. Analisis sentimen berbasis *machine learning* dapat digunakan untuk mengolah opini publik yang bersifat tidak terstruktur menjadi informasi yang lebih mudah dianalisis (Khan & Srivastava, 2024; Wnkhade et al., 2022). Penelitian ini berfokus pada perbandingan dua algoritma klasifikasi teks, yaitu *Support Vector Machine* (SVM) dan Naïve Bayes, untuk mengetahui model yang memiliki performa terbaik dalam mengklasifikasikan sentimen terhadap Program MBG (Wahid et al., 2025). Kontribusi utama penelitian ini adalah: (1) penerapan CRISP-DM untuk analisis sentimen kebijakan publik spesifik, (2) perbandingan komprehensif SVM dan Naïve Bayes pada data bahasa Indonesia, dan (3) pengembangan prototipe sistem pemantauan sentimen real-time.

METODE

Penelitian ini menggunakan metodologi CRISP-DM (*Cross-Industry Standard Process for Data Mining*) yang terdiri dari enam fase: *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment*. Metodologi ini terdiri dari enam fase yang saling terkait (Wirth & Hipp, 2000).



Gambar 1. Proses CRISP-DM

Pengumpulan Data

Tahap data understanding dilakukan melalui pengumpulan data tweet dari platform X menggunakan web scraping dengan tools Tweet Harvest berbasis Node.js dan metode cookie-based authentication (Rahman, 2023). Kueri pencarian yang digunakan adalah: "MBG OR mbg OR makanbergizigratis since:2026-05-20 until:2026-06-25 lang=id". Proses

pengumpulan dilakukan dalam 10 kali eksekusi dengan batas maksimal 500 tweet per eksekusi untuk menghindari rate limit, menghasilkan total 4.754 tweet.

Pra-Pemrosesan Data

Berikut merupakan proses pra-pemrosesan data yang dilakukan dalam penelitian ini (Joshi & Patel, 2021):

- **Data Cleaning:** Menghapus kolom yang tidak relevan, kata duplikat, dan menangani outlier.
- **Text Processing:** Melakukan pembersihan teks pada kolom `full_text` dengan menghapus URL, mention (@username), hashtag (#), karakter khusus, dan angka, kemudian mengubah semua teks menjadi huruf kecil (lowercase).
- **Tokenization:** Memecah teks menjadi unit kata (token).
- **Stopword Removal:** Menghapus kata umum yang tidak informatif menggunakan daftar stopwords bahasa Indonesia dan Sastrawi.
- **Stemming:** Mengubah kata berimbuhan menjadi kata dasar menggunakan library Sastrawi.
- **Feature Extraction:** Mengekstraksi fitur teks menggunakan metode TF-IDF (Term Frequency-Inverse Document Frequency) dengan mempertahankan 5.000 fitur teratas (term) yang memiliki nilai variansi tertinggi, mencakup unigram dan bigram.

Pelabelan Sentimen

Pelabelan sentimen dilakukan secara manual berbasis lexicon dengan tiga kategori:

- **Kata Positif:** ['mendukung', 'setuju', 'bagus', 'baik', 'gratis', 'manfaat', 'senang', 'gizi', 'sehat', 'pro', 'mantap', 'keren', 'hebat', 'luar biasa', 'terima kasih', 'sukses', 'berhasil', 'berguna', 'bermanfaat', 'tepat', 'puas', 'membantu', 'layak', 'patut', 'teladan', 'maju', 'berkualitas', 'transparan', 'tepat sasaran', 'relevan', 'inovatif', 'efektif', 'efisien', 'progresif', 'unggul', 'cerdas', 'solusi', 'harapan', 'optimis', 'membaik', 'perbaikan', 'peningkatan', 'berkembang', 'solutif', 'visioner', 'produktif', 'percaya', 'yakin', 'diharapkan', 'menyenangkan', 'bahagia', 'cemerlang', 'brilian', 'andal', 'terpercaya', 'akurat', 'sederhana', 'mudah', 'lancar', 'cepat', 'hemat', 'murah', 'peduli', 'bijaksana', 'inspiratif', 'edukatif', 'modern', 'terdepan', 'pionir', 'gemilang', 'makmur', 'sejahtera', 'dukung', 'lanjutkan', 'bela', 'pro-rakyat', 'pro mbg', 'terima kasih prabowo', 'bergizi', 'kenyang', 'anak sehat', 'ibu bahagia', 'semangat', 'terbantu', 'program presiden', 'cetak generasi emas', 'bangga', 'bersyukur', 'alhamdulillah', 'terbaik', 'program bagus']
- **Kata Negatif:** ['stop', 'tolak', 'hentikan', 'bubar', 'gagal', 'berhenti', 'distop', 'diboongin', 'dibohongi', 'dikerahkan', 'korupsi', 'koruptor', 'bocor', 'mark up', 'markup', 'rampok', 'uang rakyat', 'pajak', 'fiktif', 'siluman', 'blusukan', 'bancakan', 'ladang korupsi', 'penjarahan', 'maling', 'skala prioritas', 'anggaran pendidikan', 'anggaran kesehatan', 'guru honorer', 'keracunan', 'racun', 'phk', 'dipecat', 'parah', 'biadab', 'bangsat', 'gila', 'kocak', 'tai', 'laknat', 'rezim', 'bodoh', 'tuli', 'buta', 'tolol', 'goblog', 'dungu', 'anjing', 'anjir', 'ampun', 'capek', 'muak', 'miris', 'nyesek', 'bayaran', 'demonstan', 'polisi', 'aparatus', 'demo', 'mahasiswa', 'emak-emak', 'guru', 'kerugian', 'utang', 'rupiah', 'listrik', 'pemadaman', 'mati lampu', 'pln', 'boongin', 'adu domba', 'buruk', 'menolak', 'lawan', 'jelek', 'sial', 'benci', 'kecewa', 'mengecewakan', 'curang', 'bohong', 'tidak jelas', 'amburadul', 'acak-acakan', 'busuk', 'basi', 'tidak layak', 'tidak sehat', 'merugikan', 'boros', 'sia-sia', 'sampah', 'kacau', 'bermasalah', 'darurat', 'protes', 'demonstrasi', 'tidak setuju', 'rugi', 'susah', 'sulit', 'lambat', 'mahal', 'menipu', 'tipuan', 'fitnah', 'manipulasi', 'diskriminasi', 'ketidakadilan', 'pelanggaran', 'penyalahgunaan', 'ancaman', 'bahaya', 'resiko', 'krisis', 'frustasi', 'marah', 'geram', 'bingung', 'ragu', 'takut', 'cemas', 'stres', 'penderitaan', 'kemiskinan', 'kelaparan', 'keluhan', 'keluh kesah', 'tidak efektif', 'tidak efisien', 'tidak terorganisir', 'lemah', 'payah', 'rendah', 'kotor', 'rusak', 'hancur', 'embel-embel', 'ompreng', 'omprengan', 'kutukan', 'gizi buruk', 'stunting', 'gagal', 'dibayar', 'upah', 'imbalan', 'hadiah', 'doorprize', 'wajan', 'aib', 'terbongkar', 'sandiwara', 'adegan', 'bodohin', 'boneka', 'kelabakan', 'ambruk', 'jebol']
- **Kata Netral:** ['tapi', 'bukannya', 'seharusnya', 'tolong', 'diperhatikan', 'semoga', 'dipantau']

Pemodelan

Pada tahap pemodelan, sistem klasifikasi sentimen dibangun menggunakan dua algoritma Machine Learning yang dibandingkan, yaitu *Support Vector Machine* (SVM) dengan parameter $C=0.1$ dan kernel RBF, serta Naïve Bayes varian multinomial dengan parameter smoothing $\alpha=0.1$. Pembagian data dilakukan dengan proporsi 80% data latih

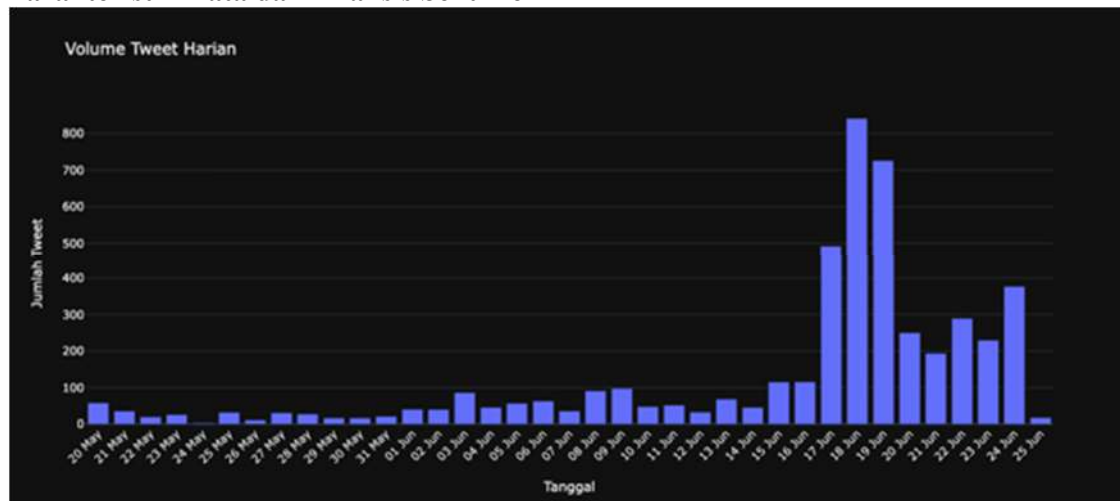
(*training set*) dan 20% data uji (*testing set*). Penanganan ketidakseimbangan kelas menggunakan metode SMOTE (*Synthetic Minority Over-sampling Technique*) pada data latih.

Evaluasi Model dan Implementasi Sistem

Model dievaluasi menggunakan metrik accuracy, precision, recall, dan F1-Score melalui classification report, confusion matrix, serta cross-validation 5-fold untuk mengukur stabilitas model. Selanjutnya, sistem diimplementasikan sebagai dashboard berbasis Flask dengan komponen: form upload, visualisasi distribusi sentimen (*pie chart*), grafik tren temporal (*line chart*), word cloud per kategori, tabel sampel tweet, dan sistem alert untuk deteksi lonjakan sentimen negatif >30% dalam 24 jam.

HASIL DAN PEMBAHASAN

Karakteristik Data dan Analisis Sentimen



Gambar 2. Volume Harian Tweet

Berdasarkan grafik Gambar 2, sumbu horizontal menunjukkan rentang waktu dari 20 Mei hingga 25 Juni 2026, sedangkan sumbu vertikal merepresentasikan jumlah tweet per hari yang berkisar antara 0 hingga lebih dari 800 tweet. Terlihat bahwa volume tweet mengalami fluktuasi yang signifikan, dengan beberapa lonjakan tajam yang mengindikasikan adanya peristiwa atau momen tertentu yang memicu tingginya diskusi publik.

Tabel 1. Sample Data Distribusi Sentimen Awal Harian (Periode Mei - Juni 2026)

Tanggal	Negatif	Netral	Positif	Total	% Negatif
2026-05-20	15	40	3	58	25,86%
2026-05-21	15	15	6	36	41,67%
2026-05-27	19	7	5	31	61,29%
2026-06-08	40	48	3	91	43,96%
2026-06-12	28	0	5	33	84,85%
2026-06-17	123	306	63	492	25,00%
2026-06-18	303	355	183	841	36,03%
2026-06-19	199	381	146	726	27,41%
2026-06-25	13	4	1	18	72,22%

Dataset penelitian terdiri dari 4.754 tweet terkait Program MBG yang dikumpulkan selama periode 20 Mei – 25 Juni 2026. Setelah melalui tahap preprocessing dan pembersihan data, diperoleh 2.538 data unik yang siap dianalisis. Distribusi sentimen setelah pembersihan disajikan pada Tabel 2.

Tabel 2. Distribusi Sentimen Akhir

Kategori Sentimen	Jumlah	Persentase
Positif	444	17,5%
Negatif	1.028	40,5%
Netral	1.066	41,9%
Total	2.538	100%

Berdasarkan Tabel 2, terlihat bahwa sentimen Netral mendominasi dengan persentase 41,9%, diikuti oleh sentimen Negatif sebesar 40,5%, sementara sentimen Positif hanya mencapai 17,5%. Hal ini mengindikasikan bahwa perbincangan publik di media sosial X lebih banyak bersifat informatif atau kritis dibandingkan dengan dukungan penuh terhadap Program MBG.

Analisis temporal menunjukkan fluktuasi sentimen yang signifikan selama periode pengamatan. Lonjakan sentimen negatif tertinggi terjadi pada 12 Juni 2026 dengan proporsi 84,8%, dan pada 25 Juni 2026 sebesar 72,2%. Sebaliknya, sentimen positif sempat mencapai titik terendah 0% pada 12 Juni 2026. Volume tweet tertinggi terjadi pada 18 Juni 2026 dengan 841 tweet, diikuti 17 Juni (492) dan 19 Juni (726). Fluktuasi ini mengindikasikan adanya peristiwa atau pemberitaan tertentu yang memicu perubahan opini publik secara drastis.

Perbandingan Model Support Vector Machine (SVM) dan Naïve Bayes

Setelah proses preprocessing dan ekstraksi fitur TF-IDF, data digunakan untuk melatih dan menguji kedua model klasifikasi. Tabel 3 menyajikan perbandingan metrik kinerja keseluruhan antara model SVM dan Naïve Bayes.

Tabel 3. Perbandingan Metrik Kinerja Model SVM dan Naïve Bayes

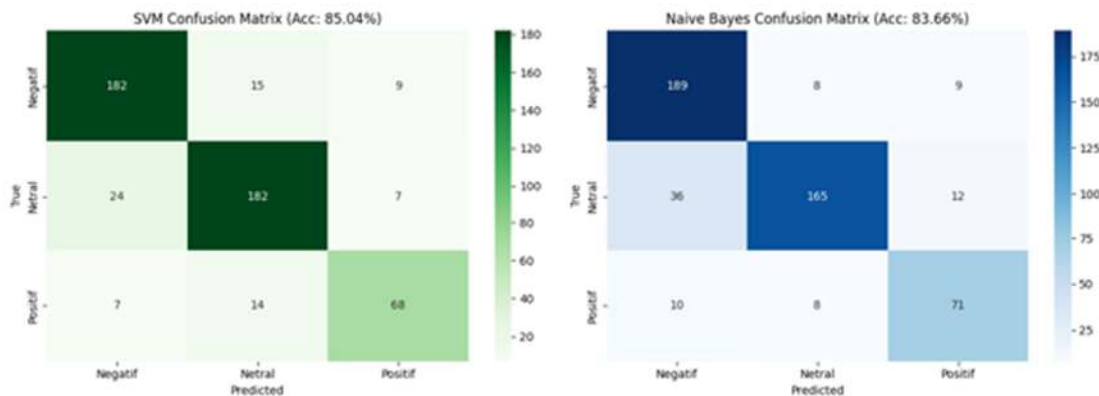
Metrik	SVM	Naïve Bayes
Akurasi Data Uji	85,04%	83,66%
Akurasi Data Latih	94,73%	96,45%
Selisih Akurasi	9,69%	12,79%
Cross-Validation 5-Fold	0,8888 ($\pm 0,0234$)	0,8825 ($\pm 0,0178$)
F1-Macro	83,78%	82,64%
F1-Weighted	85,00%	83,00%

Berdasarkan Tabel 3, model SVM menunjukkan akurasi data uji 85,04%, lebih tinggi dibandingkan Naïve Bayes (83,66%). Selisih antara akurasi latih dan uji pada SVM (9,69%) lebih kecil dibandingkan Naïve Bayes (12,79%), mengindikasikan bahwa SVM lebih stabil dan tidak mengalami overfitting yang signifikan. Hasil cross-validation 5-fold mengkonfirmasi stabilitas kedua model, dengan SVM pada 0,8888 ($\pm 0,0234$) dan Naïve Bayes pada 0,8825 ($\pm 0,0178$). Analisis per kelas (Tabel 4) memberikan gambaran lebih mendalam tentang kekuatan masing-masing model.

Tabel 4. Perbandingan Performa Per Kelas Sentimen

Kelas	Metrik	SVM	Naïve Bayes
Negatif	Precision	85,45%	80,43%
Negatif	Recall	88,35%	91,75%
Negatif	F1-Score	86,88%	85,75%
Netral	Precision	86,26%	91,16%
Netral	Recall	85,45%	77,46%
Netral	F1-Score	85,85%	83,77%
Positif	Precision	80,95%	77,17%
Positif	Recall	76,40%	79,78%
Positif	F1-Score	78,61%	78,45%

Naïve Bayes unggul dalam recall untuk kelas Negatif (91,75%) dan Positif (79,78%), yang berarti model ini lebih sensitif dalam menangkap sentimen negatif dan positif. Namun, SVM menunjukkan precision yang lebih tinggi untuk kelas Negatif (85,45%) dan Positif (80,95%), serta keseimbangan yang lebih baik pada kelas Netral (precision 86,26%, recall 85,45%). Keunggulan SVM dalam menjaga keseimbangan antara precision dan recall di semua kelas menjadikannya model yang lebih andal untuk sistem pemantauan opini publik.



Gambar 3. Confusion Matrix Hasil Pelatihan dan Pengujian SVM dan Naïve Bayes

Confusion matrix dari kedua model (Gambar 3) menunjukkan pola kesalahan yang berbeda. SVM lebih baik dalam mengklasifikasikan tweet netral dengan benar, sementara Naïve Bayes cenderung lebih "agresif" dalam mengkategorikan tweet ambigu ke arah negatif. Hal ini penting dalam konteks pemantauan kebijakan karena kesalahan klasifikasi netral ke negatif dapat memicu alarm palsu yang berlebihan.

Dari model terbaik yang sudah disimpan, sistem pemantauan berhasil diimplementasikan sebagai dashboard berbasis Flask dengan fitur-fitur utama: (1) Form Upload untuk prediksi sentimen instan, (2) Statistik Deskriptif total tweet dan interaksi, (3) Pie Chart Distribusi Sentimen, (4) Line Chart Tren Temporal, (5) Word Cloud per Kategori, dan (6) Alert System untuk deteksi lonjakan sentimen negatif >30% dalam 24 jam. Sistem dideploy melalui ngrok dan dapat diakses secara publik.

Diskusi

Hasil penelitian menunjukkan bahwa kedua algoritma memiliki performa yang baik dengan akurasi di atas 83%. Temuan ini sejalan dengan penelitian Rahman (2023) yang menyimpulkan bahwa SVM dan Naïve Bayes efektif untuk analisis sentimen bahasa Indonesia. Namun, penelitian ini memberikan kontribusi tambahan dengan menerapkan kedua algoritma dalam konteks pemantauan kebijakan publik real-time.

Keunggulan SVM dalam menjaga keseimbangan antar kelas sentimen sangat penting dalam konteks pemantauan kebijakan. Seperti yang diungkapkan oleh Alba et al. (2025), analisis sentimen kebijakan membutuhkan model yang tidak bias terhadap satu kelas tertentu. Dalam penelitian ini, SVM menunjukkan performa yang lebih seimbang pada kelas Netral yang memiliki proporsi data terbesar (41,9%). Sementara itu, Naïve Bayes memiliki recall terendah pada kelas Netral (77,46%), yang berarti banyak tweet netral yang salah diklasifikasikan ke kelas lain.

Keunggulan Naïve Bayes dalam mendeteksi sentimen negatif (recall 91,75%) menunjukkan bahwa algoritma ini lebih sensitif terhadap kritik publik. Namun, presisi negatifnya yang lebih rendah (80,43%) dibandingkan SVM (85,45%) berarti Naïve Bayes cenderung menghasilkan lebih banyak false positive, yang dapat menyebabkan overreaksi dalam manajemen krisis. SVM lebih unggul dalam menangani kompleksitas fitur teks

berdimensi tinggi hasil ekstraksi TF-IDF, sehingga lebih sesuai digunakan sebagai model utama dalam sistem pemantauan sentimen Program MBG.

Penelitian ini memiliki beberapa keterbatasan: pertama, pengumpulan data terbatas pada satu platform (X) dan periode 30 hari. Kedua, pelabelan sentimen dilakukan berbasis lexicon yang mungkin tidak seakurat pelabelan manual. Penelitian selanjutnya disarankan untuk memperluas cakupan data dan menerapkan model deep learning seperti IndoBERT.

SIMPULAN

Penelitian ini berhasil membandingkan algoritma SVM dan Naïve Bayes untuk analisis sentimen publik terhadap Program Makan Bergizi Gratis pada media sosial X. Hasil evaluasi menunjukkan SVM memiliki akurasi 85,04%, lebih tinggi dibandingkan Naïve Bayes sebesar 83,66%. Dengan demikian, SVM dipilih sebagai model terbaik untuk mendukung sistem pemantauan sentimen publik berbasis machine learning karena presisi yang lebih tinggi dan stabilitas yang lebih baik. Kemudian, sistem dashboard berbasis Flask berhasil diimplementasikan untuk mendukung pemantauan opini publik secara real-time.

DAFTAR PUSTAKA

- Alba, C., Warner, B., Saxena, A., Huang, J., & An, R. (2025). Towards Robust Sentiment Analysis of Temporally-Sensitive Policy-Related Online Text. Association for Computational Linguistics (ACL). <https://doi.org/10.18653/v1/2025.acl-srw.70>
- Harwenda, R., Angelo, M., Budi, I., Santoso, A., & Putra, P. (2025). Sentiment Analysis on Government Public Policies: A Systematic Literature Review. *Dinasti International Journal of Education Management And Social Science*, 6(5), 4192–4211. <https://doi.org/10.38035/dijemss.v6i5.4699>
- James, T. R., et al. (2022). Application of Machine Learning Methods on Climate Data and Commercial Microwave Link Attenuations for Estimating Meteorological Visibility in Dusty Condition. *Engineering*, 14(2).
- Joshi, D.-A., & Patel, B. (2021). Data Preprocessing: The Techniques for Preparing Clean and Quality Data for Data Analytics Process. *Oriental Journal of Computer Science and Technology*, 13, 78–81. <https://doi.org/10.13005/ojest13.0203.03>
- Khan, M., & Srivastava, A. (2024). Sentiment Analysis of Twitter Data Using Machine Learning Techniques. *International Journal of Engineering and Management Research*, 14, 196–203. <https://doi.org/10.5281/zenodo.10791485>
- Kiftiyah, K. R. A. (2025). Program Makan Bergizi Gratis (MBG) dalam Perspektif Keadilan Sosial dan Dinamika Sosial – Politik. *Keindonesiaan*, 5(1), 101–112.
- Mulyono, B., Affandi, I., Suryadi, K., & Darmawan, C. (2023). Online civic engagement through social media: An analysis of twitter big data. *Jurnal Cakrawala Pendidikan*, 42(1). <https://doi.org/10.21831/cp.v42i1.54201>
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes classifier dan confusion matrix pada analisis sentimen berbasis teks pada Twitter. *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, 5(2), 697–711.
- Peraturan Presiden Republik Indonesia. (2024). Peraturan Presiden Republik Indonesia Nomor 83 Tahun 2024 tentang Badan Gizi Nasional. Jakarta: Sekretariat Negara.
- Rahman, D. (2023). Analisis Chatgpt Tweet Menggunakan EDA Dan Sentiment Analysis: Data Pengguna Twitter Di Indonesia. Skripsi/Laporan Penelitian.
- Schröer, C., Kruse, F., & Marx Gómez, J. (2021). A Systematic Literature Review on Applying CRISP-DM Process Model. *Procedia Computer Science*, 181, 526–534. <https://doi.org/10.1016/j.procs.2021.01.199>
- Singh, S. (2025). Twitter/X statistics 2025: Users, revenue, and growth. DemandSage.
- Wahid, Y., Sanatang, & Andayani, D. (2025). Performance Comparison of SVM and Naïve Bayes For Indonesian-Language Sentiment Analysis On Free Fire Reviews Using TF-IDF And SMOTE. *Journal of Embedded Systems, Security and Intelligent Systems*, 6(4), 674–689. <https://doi.org/10.59562/jessi.v6i4.10818>

- Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining.
- Wnkhade, M., Rao, A., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55, 5731–5780.
<https://doi.org/10.1007/s10462-022-10144-1>