Jurnal Pustaka Cendekia Hukum dan Ilmu Sosial
Volume 4, No. 2, June-September 2026, pp. 3314-3324
E-ISSN: 2987-4475
https://pchukumsosial.org/index.php/pchs

Klasifikasi Sentimen Social Media Tentang Program Makan Bergizi Gratis dengan Metode *Support Vector Machine* (SVM)

Fikri Fadilah[1], Imam Budiawan[2*], Yumi Novita Dewi[3]

[1,2,3]Informatika, Universitas Bina Sarana Informatika, Jl. Kramat Raya No.98, Kec. Senen, Kota Jakarta Pusat, DKI Jakarta
imam.imb@bsi.ac.id

Abstract

The Free Nutritional Meal Program (MBG) is a government policy aimed at improving the nutritional quality of the community, particularly students. The implementation of this program has generated diverse responses from the public on social media, necessitating sentiment analysis to objectively understand public perception. This study aims to analyze public sentiment towards the Free Nutritional Meal Program on platform X, implement the Support Vector Machine (SVM) method in the sentiment classification process, and evaluate model performance using accuracy, precision, recall, and F1-score metrics. The study used a quantitative approach based on text mining, with a dataset of 848 tweets that underwent preprocessing stages, including cleaning, case folding, tokenizing, normalization, stopword removal, and stemming. Next, feature extraction was performed using Term Frequency–Inverse Document Frequency (TF-IDF), splitting the training data to the test data with an 80:20 ratio, and classification using the linear kernel SVM algorithm. The results showed that public sentiment was dominated by neutral sentiment at 51.7%, followed by positive sentiment at 33% and negative sentiment at 15.3%. The SVM model with a linear kernel produced an accuracy of 80%, a weighted F1-score of 0.80, and a macro F1-score of 0.77. These results indicate that the combination of TF-IDF and SVM is able to effectively classify social media sentiment and can be used to support the evaluation of the Free Nutritional Meal Program implementation.

Keywords: Sentiment Analysis, Support Vector Machine, TF-IDF, Text Mining, Free Nutritional Meal Program.

Abstrak

Program Makan Gratis Bergizi (MBG) adalah kebijakan pemerintah yang telah menarik perhatian luas dan menimbulkan berbagai reaksi dari masyarakat, khususnya di platform media sosial seperti X. Beragam opini ini perlu diteliti secara terstruktur untuk memahami bagaimana masyarakat berpikir dan sebagai referensi untuk penyiaran implementasi program tersebut. Algoritma digunakan untuk analisis. Proses tersebut meliputi pembersihan dan pengolahan teks, serta pembobotan kata berdasarkan metode TF-IDF. Penelitian ini menggunakan teknik penambangan teks. Dataset ini berisi 848 tweet berbahasa Indonesia yang telah melalui beberapa proses, termasuk pembersihan, case-folding, tokenisasi, normalisasi, penghapusan stopword, stemming, ekstraksi fitur menggunakan TF-IDF, dan penentuan sentimen, yang dibagi menjadi tiga kelas: positif, negatif, dan netral.Data kemudian dibagi 80:20 untuk pelatihan dan pengujian menggunakan algoritma SVM dengan kernel linier. Penelitian ini menyelidiki bahwa jumlah tweet sama dengan total tweet. Model SVM yang menggunakan kernel linier memberikan hasil terbaik dengan akurasi 80%, skor F1 tertimbang 0,80, dan skor F1 makro 0,77. Temuan ini menunjukkan bahwa kombinasi pra-pemrosesan teks, TF-IDF, dan teknik SVM berhasil dalam mengenali perasaan masyarakat terhadap Program Makanan Bergizi Gratis yang dibahas di media sosial.

Kata Kunci: Analisis Sentimen, Support Vector Machine, TF-IDF, Penambangan Teks, Program Makanan Bergizi Gratis.

✉ Corresponding author: Fikri Fadilah
Email Address: fadilahfikri159@gmail.com (Jl. Kramat Raya No.98, Kota Jakarta Pusat, DKI Jakarta)
Received 01 August 2026, Accepted 13 August 2026, Published 26 August 2026

PENDAHULUAN

Pada masa kini, di era kemajuan digital yang pesat, publik semakin gampang menyampaikan pendapat mengenai kebijakan pemerintah. Akhir-akhir ini, program pemberian makanan bergizi gratis menjadi pusat perhatian—di mana saja, khususnya di platform social

media X, banyak orang membicarakan program ini secara serentak. Beberapa orang melihat sisi positif: anak-anak itu memperoleh asupan gizi yang cukup, perhatian mereka saat pelajaran menjadi lebih baik, serta terdapat peluang usaha bagi UMKM setempat. Namun bukan semua orang sependapat. Banyak orang yang merasa cemas, "Sumber dana pendidikan katanya berkurang, sumber pendanaannya pun tidak transparan—barangkali ini cuma program populis tanpa rencana jangka panjang?" Pendapat yang cukup menusuk hati, "Siswa sekolah benar-benar mendapatkan makan siang gratis, tetapi kesempatan untuk melanjutkan sekolah terbatas. Kenyang memang, tapi harapan masa depannya masih gelap.

Fokus utama penelitian adalah mengidentifikasi program persepsi masyarakat dengan menggunakan data media sosial secara sistematis dan obyektif. Pilih sebagai sumber data karena merupakan platform yang sangat aktif digunakan masyarakat Indonesia dalam menyampaikan opini. Kecepatan penyebaran informasi dan tingginya interaksi membuat platform ini menjadi pusat percakapan dan tren nasional.

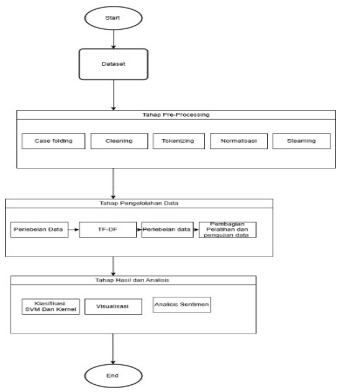
Media sosial juga menjadi ruang bagi masyarakat untuk Perluas pesannya. Berbagai bentuk media sosial mencapai 139 juta orang pada Januari 2024, mencerminkan tingginya aktivitas dan keterlibatan digital masyarakat. Langkah selanjutnya adalah pra-pemrosesan teks, yang juga dikenal sebagai tindakan mengubah teks menjadi format standar yang dapat dianalisis lebih mendalam, dengan tujuan utama untuk mengurangi noise. Berdasarkan Survei Status Gizi Indonesia (SSGI) tahun 2022, prevalensi stunting di Indonesia tercatat sebesar 21,6%, menurun dari 24,4% pada tahun sebelumnya. Penurunan ini menunjukkan adanya kemajuan, namun masih diperlukan upaya strategis untuk mencapai target nasional.

Mulai tanggal 30 September 2025, Badan Gizi Nasional (BGN) menginformasikan kepada masyarakat bahwa 56 Satuan Pelayanan Pemenuhan Gizi (SPPG) akan mengikuti program Makan Bergizi Gratis (MBG), dengan Jawa Barat menjadi provinsi dengan MBG dan keracunan korban tertinggi, yaitu 4.125 korban. Dilansir dari Merdeka.com. Fokus penelitian ini adalah menganalisis persepsi publik telusuri fitur media sosial dari program ini.Algoritma ini terbukti efektif dalam mengolah data teks dan mampu memberikan performa klasifikasi yang tinggi, dengan akurasi mencapai 80%.

METODE

sentimen media sosial mengenai program "Makan Bergizi Gratis" dengan menggunakan metode Support Vector Machine (SVM). Metode ini dirancang untuk memproses data tekstual yang tidak terstruktur dari platform digital menjadi informasi sentimen yang terstruktur, yaitu positif, negatif, atau netral. Penelitian ini merupakan penelitian kuantitatif dengan pendekatan text mining yang berfokus pada analisis sentimen publik terhadap Program Makan Bergizi Gratis pada media sosial X. Pendekatan.

Proses Langkah Penelitian



Sumber : Hasil Penelitian 2026
Gambar III 1 Proses Langkah Penelitian

Pengumpulan Data

Dataset dimanfaatkan cuitan dari sumber terbuka kaggle (https://www.kaggle.com/datasets/brianchang248/makan-bergizi-gratis-twitter-february-2025-dataset) mengenai yang di dapatkan meliki total data 2.500 dataset kaggle setelah dilakukan proses *di google colab* pembersihan data dan penghapusan data berisi berisi 890 komentar publik terkait Program Makan Bergizi Gratis. Data diunduh pada 13 Febuari 2025 dan hanya menggunakan kolom full_text sebagai objek analisis

Proscesing

Tahap pre-processing bertujuan mengubah data teks tidak terstruktur menjadi format yang siap diproses oleh algoritma klasifikasi. Tahapan pre- processing yang diterapkan meliputi:

1. *Cleaning:* Menghapus elemen tidak relevan seperti URL, mention (@username), hashtag, retweet (RT), emoji, dan karakter khusus lainnya menggunakan ekspresi reguler (regular expression).
2. *Case folding*: Mengonversi seluruh karakter teks menjadi huruf kecil (lowercase) untuk menghindari duplikasi kata akibat perbedaan kapitalisasi.
3. *Tokenizing:* Memecah teks menjadi unit kata (token) menggunakan library NLTK.
4. *Normalisasi*: Mengubah kata tidak baku, singkatan, atau typo menjadi bentuk baku berdasarkan kamus normalisasi bahasa Indonesia (contoh: "gk" "tidak","gua" → "aku").
5. *Stopword* removal: Menghapus kata-kata umumyang tidak membawa makna spesifik seperti "yang", "di", "ke", "dan" menggunakan daftar stopword bahasa Indonesia dari *library NLTK*.
6. *Stemming*: Mengubah kata berimbuhan menjadi bentuk dasarnya menggunakan Sastrawi *Stemmer* (contoh: "makanan" → "makan", "bergizi"

Labeling

Tahap pelabelan data diimplementasikan setelah proses preprocessing selesai, dengan menggunakan model pre-trained RoBERTa varian Bahasa Indonesia. Model berbasis transformer ini bertugas mengklasifikasikan dataset ke dalam beberapa parameter sentimen. Dalam penelitian

ini, klasifikasi dilakukan secara menyeluruh dengan membagi data ke dalam kategori sentimen positif, negatif, dan netral Confusion.

Yaitu:

1. Sentimen Positif: Opini yang mendukung, memuji, atau menekankan manfaat program gizi seimbang, peningkatan konsentrasi, dampak ekonomi bagi UMKM
2. Sentimen Negatif: Opini yang berisi kritik, kekhawatiran anggaran, ketidakjelasan sumber dana, atau anggapan bahwa program ini hanya bersifat populis.
3. Sentimen Netral: Opini berupa teks berita factual atau pernyataan yang tidak menunjukkan keberpihakan emosional secara jelas.

Pembobotan Kata (Feature Extraction menggunakan TF-IDF)

Teks yang sudah bersih diekstraksi fiturnya menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). Proses ini mengubah representasi kata yang bersifat tekstual menjadi bentuk vektor numerik berdasarkan tingkat kepentingan kata tersebut di dalam dataset, sehingga dapat dipahami secara komputasional oleh algoritma SVM.

Pembagian Dataset

(Data Splitting)Datavektor numerik akan dibagi secara acak (random) ke dalam dua bagian utama menggunakan rasio tertentu.

Kinerja evaluasi model

Evaluasi kinerja model di evaluasikan mengguanakan confusion matrix Tediri dari empat komponen True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) . Berdasarkan komponen tersebut, dilakukan : perhitungan yang terhadap empat & metrik evaluasi, yaitu : accuracy, precision, recall, dan F1-score.

1. Akurasi (Accuracy)

   Mengukur proporsi prediksi yang benar terhadap seluruh data uji. Rumus:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

2. Presisi (Precision)

   Mengukur ketepatan model dalam memprediksi kelas positif. Rumus:

$$Precision = \frac{TP}{TP + FP}$$

3. Recall

   Mengukur kemampuan model untuk menemukan seluruh data positif yang sebenarnya. Rumus:

$$Recall = \frac{TP}{TP + FN}$$

4. F1-Score

   Merupakan rata-rata harmonik antara presisi dan recall. Rumus:

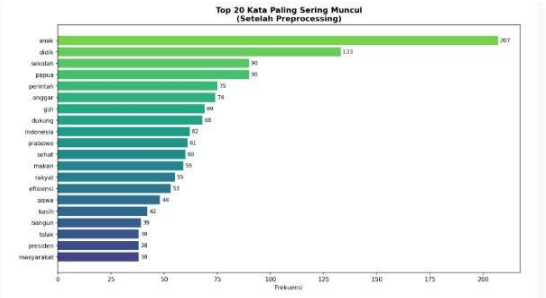
$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Evaluasi performa model dilakukan menggunakan Confusion Matrix guna menggambarkan visualisasi distribusi prediksi terhadap kelas sentimen positif, negatif, dan netral. Instrumen ini menyajikan rincian perbandingan antara nilai aktual dengan hasil prediksi model, sehingga pola klasifikasi pada seluruh dataset dapat terpetakan dengan jelas. berdimensi 3×3 menampilkan jumlah prediksi benar dan salah dari setiap kelas, sehingga membantu dalam mengidentifikasi pola kesalahan model. Seluruh tahapan pengolahan data dilakukan secara terstruktur menggunakan berbagai pustaka Python seperti Scikit-learn untuk implementasi SVM dan evaluasi model, NLTK dan Sastrawi untuk pra-pemrosesan teks bahasa Indonesia, serta Pandas dan NumPy untuk pengelolaan dan manipulasi data. Dengan pendekatan yang sistematis ini, diharapkan penelitian mampu menghasilkan model klasifikasi sentimen yang akurat, andal, dan dapat dipertanggungjawabkan secara ilmiah dalam menganalisis persepsi publik terhadap Program Makan Bergizi Gratis di media sosial.

HASIL DAN DISKUSI

Penelitian ini memanfaatkan dataset yang terdiri dari serangkaian cuitan opini.berbahasa Indonesia di platform media sosial X yang berhubungan dengan ProgramMakan Sehat Gratis (MBG). Data kumpulkan melalui 3 metode crawling dari Kaggle dan telah menjalani proses pelabelan manual ke dalam tiga kategori sentimen: Positif (n=280), Negatif (n=130), dan Netral (n=438). Jumlah keseluruhan data yang dipakai dalam penelitian Ini mewakili 848 pesan Twitter. Temuan akhir penelitian diuraikan di bawah ini.secara komprehensif, mulai dari analisis data usai preprocessing, pemilihan kernel SVM yang paling optimal, hingga penilaian kinerja model akhir.

Kata Paling Sering Muncul Setelah Preprocessing



Gambar 1. Top 20 Kata Paling Sering Muncul Setelah Preprocessing
Sumber Dari google colab Hasil Penelitian

Memperlihatkan bahwa kata “anak” merupakan kata yang paling dominan dengan frekuensi kemunculan sebesar 207 kali, diikuti oleh kata “didik” (133 kali), “sekolah” dan “papua” (masing-masing 90 kali), serta “perintah” (75 kali). Kata-kata lain yang juga sering muncul meliputi “anggar”, “gizi”, “dukung”, “indonesia”, dan “prabowo”. Kemunculan kata-kata ini

relevan dengan konteks Program Makan Bergizi Gratis yang erat kaitannya dengan isu pendidikan anak, kebijakan pemerintah, anggaran negara, serta tokoh yang memprakarsainya. Dominasi kata "anak" dan "didik" mengindikasikan bahwa masyarakat banyak membahas dampak program ini terhadap dunia pendidikan dan tumbuh kembang anak sekolah.Analisa dalam processing keyword kata dalam dominasi dari tertinggi hingga terendah dan di aritkan fokus dalam beropini .Analisa dalam dominasi keyword kata dominasi tinggi (anak, didik, sekolah, papua, perintah): Fokus utama masyarakat pada penerima manfaat, pendidikan, pemerataan wilayah, dan peran pemerintah. Dominasi menengah (anggar, gizi, dukung, indonesia, prabowo): Dimensi finansial, kesehatan, dukungan sosial, identitas nasional, serta tokoh politik.Dominasi rendah (sehat, rakyat, makan, efisiensi, siswa, kasih, bangun, tolak, presiden, masyarakat): Aspek kesehatan, sosial, moral, pembangunan, kritik, dan legitimasi publik.

visualisasi word cloud

analisis tekstual, dilakukan juga per kelas sentimen guna memperlihatkan kata-kata yang paling menonjol dalam kelas

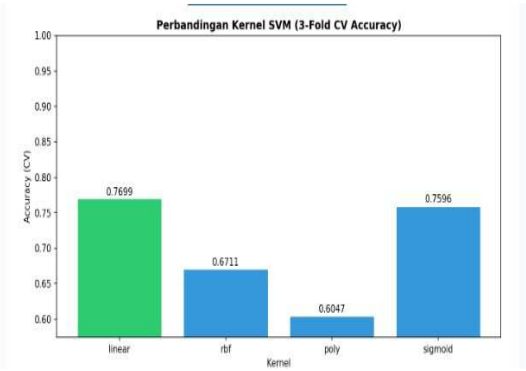
Gambar 2. visualisasi word cloud
Sumber: Dari google colab Hasil Penelitian

Sentimen positif (n=280).Program makan bergizi mendapat dukungan luas, terutama bagi anak-anak di Papua dan sekolah, karena dianggap mampu meningkatkan kesehatan dan kualitas gizi.Kata-kata dominan seperti gizi, anak, dukung, sehat, Papua, dan sekolah menunjukkan bahwa masyarakat menilai program ini bermanfaat bagi perkembangan anak dan mendukung pendidikan serta kesehatan.Sentimen negatif (n=130).Sebagian masyarakat menilai program makan bergizi tidak efisien dalam penggunaan anggaran, bahkan menimbulkan kritik terkait proses didik dan aksi demo.Kata-kata seperti didik, anak, efisiensi, anggar, ajar, dan demo menandakan adanya persepsi negatif, terutama terkait efektivitas, alokasi dana, dan protes terhadap pelaksanaan program.Sentimen netral (n=438). Program makan bergizi dipandang sebagai bagian dari kegiatan sekolah dan pendidikan anak, dengan fokus pada makan dan pengelolaan anggaran. Kata-kata seperti didik, anak, sekolah, makan, dan anggar menunjukkan pandangan netral, lebih bersifat deskriptif tanpa penilaian emosional, hanya menggambarkan keterkaitan program dengan aktivitas pendidikan.

Pemilihan Kernel Svm Terbaik

Sebelum melatih model final, dilakukan perbandingan terhadap empat jenis kernel SVM yang umum digunakan, yaitu linear, RBF (Radial Basis Function), polynomial (poly), dan

sigmoid. Perbandingan dilakukan menggunakan metode 3-Fold Cross Validation (CV) untuk memperoleh estimasi akurasi yang lebih stabil dan tidak bergantung pada pembagian data tertentu. Hasil perbandingan

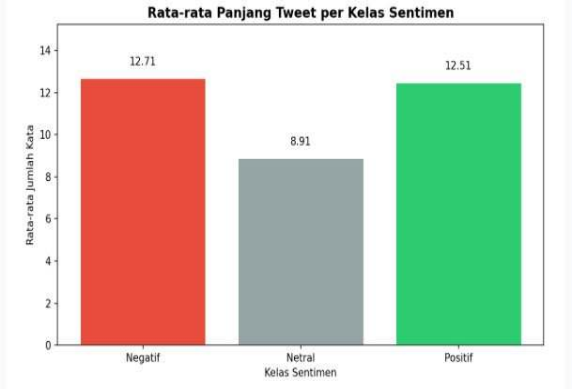


Gambar 3. Perbandingan Kernel SVM (3-Fold CV Accuracy)
Sumber: Dari google colab Hasil Penelitian, 2026

hasil perbandingan akurasi 3-Fold Cross Validation untuk keempat kernel SVM adalah sebagai berikut: kernel linear menghasilkan akurasi CV tertinggi sebesar 0,7699 (76,99%), kernel sigmoid berada di posisi kedua dengan akurasi 0,7596 (75,96%), kernel RBF berada di posisi ketiga dengan akurasi 0,6711 (67,11%), sedangkan kernel polynomial (poly) menghasilkan akurasi terendah sebesar 0,6047 (60,47%). Berdasarkan hasil perbandingan ini, kernel linear terpilih sebagai kernel terbaik dan digunakan dalam pelatihan model final. Pemilihan kernel linear sejalan dengan karakteristik data teks berdimensi tinggi yang dihasilkan dari proses ekstraksi fitur TF-IDF, di mana data teks pada umumnya lebih mudah dipisahkan secara linear dalam ruang fitur berdimensi tinggi.

Perbandingan Rata-rata Panjang Tweet per Kelas Sentimen

Analisis berikutnya mengkaji rata-rata panjang tweet per kelas sentimen untuk mengetahui apakah terdapat perbedaan karakteristik panjang teks antara kelas Positif, Negatif, dan Netral.



Gambar 4. Perbandingan Rata-rata Panjang Tweet Per Kelas Sentimen
Sumber: Dari google colab Hasil Penelitian, 2026

Menunjukkan bahwa tweet bersentimen Negatif memiliki rata-rata panjang tertinggi sebesar 12,71 kata, diikuti sentimen Positif sebesar 12,51 kata, sedangkan sentimen Netral memiliki

rata-rata panjang terendah sebesar 8,91 kata. Perbedaan panjang ini mengindikasikan bahwa pengguna yang mengekspresikan sentimen negatif maupun positif cenderung menggunakan lebih banyak kata untuk menguraikan pendapat dan argumen mereka secara lebih elaboratif. Sebaliknya, tweet netral yang umumnya berupa pernyataan faktual atau penyebaran informasi singkat cenderung lebih pendek dan padat. Temuan ini konsisten dengan karakteristik umum opini emosional di media sosial yang membutuhkan lebih banyak kata untuk mengekspresikan keberpihakan.

Pembagian Dataset (Data Splitting)

Tabel .1 Output Pembagian Dataset (Data Splitting)

Data Latih	678	Total
Data Uji	170	848
Presentas Data Latih	80%	
Presente Data Uji	20%	

Sumber Dari: Ouput google colab Hasil Penelitian, 2026

Berdasarkan output program menampilkan informasi pembagian dataset sebagai berikut: total data yang digunakan adalah 848 data, dengan rincian data latih sebanyak 678 data (80%) dan data uji sebanyak 170 data (20%). Rasio 80:20 ini dipilih karena dinilai optimal dalam menyeimbangkan antara ketersediaan data yang cukup untuk proses pembelajaran model serta keterwakilan data yang memadai untuk proses evaluasi yang objektif dan representatif.

Ekstraksi Fitur dengan TF-IDF

Gambar 5. Output Matriks TF-IDF dan Informasi Fitur
Sumber Dari: Ouput google colab Hasil Penelitian, 2026

Berdasarkan hasil ekstraksi fitur menggunakan TF-IDF menghasilkan matriks berukuran (848, 2801), yang berarti terdapat 848 dokumen (tweet) dan 2.801 fitur unik (kata-kata unik) yang berhasil diekstraksi dari seluruh corpus data setelah preprocessing. Kelas label yang dihasilkan terdiri dari tiga kategori, yaitu ['Negatif', 'Netral', 'Positif']. Jumlah fitur sebanyak 2.801 ini mencerminkan dimensionalitas tinggi yang merupakan karakteristik umum data teks,

dan hal ini justru menjadi salah satu keunggulan SVM dengan kernel linear yang mampu bekerja secara efisien pada ruang fitur berdimensi tinggi.

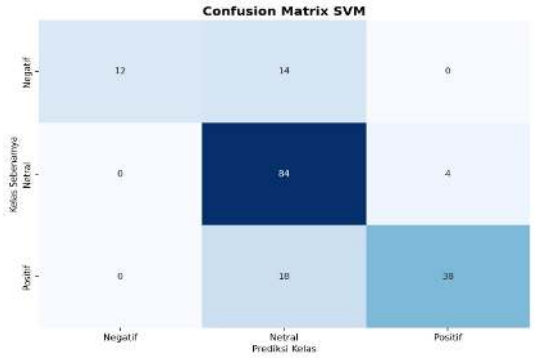
Pelatihan Model dan Laporan Klasifikasi

```
...  • Melatih model SVM...
     ✓ Model berhasil dilatih.

     Laporan Klasifikasi:
                   precision    recall  f1-score   support

          Negatif       0.71      0.65      0.68        26
           Netral       0.78      0.88      0.82        88
          Positif       0.89      0.75      0.82        56

         accuracy                           0.80       170
        macro avg       0.79      0.76      0.77       170
     weighted avg       0.81      0.80      0.80       170
```

Gambar 6. Laporan Klasifikasi Final Model SVM (Kernel Linear)
Sumber: Dari google colab .Hasil Penelitian, 2026

Model SVM dilatih menggunakan pipeline yang mengintegrasikan TfidfVectorizer dan Support Vector Classifier (SC) dengan kernel linear. Pemilihan kernel linear didasarkan pada hasil perbandingan kernel yang dilakukan sebelumnya menggunakan 3-Fold Cross Validation, di mana kernel linear menghasilkan akurasi CV tertinggi sebesar 76,99% dibandingkan kernel lainnya. Parameter random_state=42 ditetapkan untuk memastikan reprodusibilitas hasil eksperimen.Proses pelatihan dilakukan menggunakan data latih (X_train, y_train), kemudian model digunakan untuk memprediksi label sentimen pada data uji (X_test). Kode implementasi pipeline TF-IDF dan SVM serta hasil laporan klasifikasi yang diperoleh ditampilkan.

Confusion Matrix SVM



Gambar 7. Confusion Matrix Model SVM (Kernel Linear)
Sumber: Dari google colab .Hasil Penelitian, 2026

Confusion matrix model svm menampilkan informasi distribusi prediksi untuk ketiga kelas sentimen secara lengkap. Analisis terperinci dari setiap sel pada confusion matrix dijabarkan sebagai berikut.Untuk kelas Negatif, dari total 26 data yang sebenarnya berlabel Negatif, model berhasil memprediksi dengan benar (True Negative) sebanyak 12 data. Namun, sebanyak 14 data Negatif salah diklasifikasikan sebagai Netral, dan tidak ada satu pun data Negatif yang

salah diprediksi sebagai Positif. Tingginya misklasifikasi Negatif ke Netral (14 dari 26 data, atau 53,8%) mengindikasikan bahwa batas antara sentimen negatif dan netral dalam data media sosial berbahasa Indonesia cukup tipis, terutama pada opini yang disampaikan dengan nada kritis namun tidak eksplisit. Untuk kelas Netral, dari total 88 data berlabel Netral, sebanyak 84 data berhasil diprediksi dengan tepat (True Netral), sedangkan hanya 4 data yang salah diklasifikasikan sebagai Positif, dan tidak ada data Netral yang salah diprediksi sebagai Negatif. Performa yang sangat baik pada kelas Netral ini konsisten dengan tingginya nilai recall (0,88) yang diperoleh pada laporan klasifikasi. Model terbukti mampu membedakan teks informatif atau faktual (netral) dari teks opini dengan sangat baik.Untuk kelas Positif, dari total 56 data berlabel Positif, sebanyak 38 data berhasil diprediksi secara tepat (True Positif). Namun, sebanyak 18 data Positif salah diklasifikasikan sebagai Netral, dan tidak ada satu pun data Positif yang salah diprediksi sebagai Negatif. Misklasifikasi Positif ke Netral (18 dari 56 data, atau 32,1%) menunjukkan bahwa sebagian tweet yang mengandung dukungan terhadap program MBG ditulis dalam gaya bahasa yang cukup deskriptif sehingga model sulit membedakannya dari pernyataan netral.

KESIMPULAN

Sentimen masyarakat terhadap Program Makan Bergizi Gratis di platform X didominasi oleh sentimen netral sebesar 51,7%, diikuti oleh positif 33% dan negatif 15,3%. Hal ini menunjukkan bahwa mayoritas pengguna bersikap informatif dan objektif, meskipun tetap terdapat dukungan serta kritik yang cukup menonjol.Tahapan text pre-processing seperti case folding, tokenizing, stopword removal, stemming, dan normalisasi slang bekerja efektif dalam menangani karakteristik data media sosial yang tidak terstruktur..

REFERENSI

Fatkhurrohman, F., Nugroho, B. I., & Fadillah, N. (2025). Analisis Sentimen Program Makan Bergizi Gratis Pemerintah RI Melalui Twitter Menggunakan Metode SVM. 4(3), 3906–3917.

Ilmiah, J., Informasi, S., Vallen, J., & Azizah, A. H. (2025). Analisis Sentimen Media Sosial X Program Makanan Sehat Gratis dengan Support Vector Machine dan Naive Bayes tentang berbagai topik . Pendapat dapat mengambil berbagai bentuk , mulai dari kritik , pujian , adalah Twitter (X), merupakan sosial media yang menjadi pusat trending di Indonesia dan memahami pola respons masyarakat terhadap program makanan bergizi gratis . Analisis. 5(November).

Ilmiah, J., Informasi, S., Vallen, J., & Azizah, A. H. (2025). Analisis Sentimen Media Sosial X Program Makanan Sehat Gratis dengan Support Vector Machine dan Naive Bayes tentang berbagai topik . Pendapat dapat mengambil berbagai bentuk , mulai dari kritik , pujian , adalah Twitter (X), merupakan sosial media yang menjadi pusat trending di

Indonesia dan memahami pola respons masyarakat terhadap program makanan bergizi gratis . Analisis. 5(November).

Akbar, Z., Riadi, I., & Umar, R. (2026). Analisis Sentimen Program Makan Bergizi Gratis Menggunakan Lexicon-Based dan Support Vector Machine. 16(1), 150–156.

Amelia, F., Khoniaro, A. F., Nurasih, R. T., Rusyida, W. Y., Data, S., Islam, U., Abdurrahman, N. K. H., & Pekalongan, W. (2025). Unveiling public sentiment: support vector machine. 2(2), 466–479.

Go, R. Y., Sarah, H., Gaspersz, D., & Kusumawardhana, R. H. (2026). Analisis Sentimen Publik Program Makan Bergizi Gratis (MBG) di Youtube : Perbandingan Kinerja Algoritma Support Vector Machine (SVM) dan Random Forest. 7(1), 24–31.

Amelia, F., Khoniaro, A. F., Nurasih, R. T., Rusyida, W. Y., Data, S., Islam, U., Abdurrahman, N. K. H., & Pekalongan, W. (2025). Unveiling public sentiment: support vector machine. 2(2), 466–479.