Indonesian Journal of Multidisciplinary on Social and Technology
Homepage: https://journal.ilmudata.co.id/index.php/ijmst
Vol. 4 No. 3 (2026) pp: 1044-1055
P-ISSN: 2986-6790, e-ISSN: 2986-6782

Analisis Sentimen MBG Di Media Sosial X Menggunakan Support Vector Machine Dan Naive Bayes

Akbar Supia Dirja[1], Lana Aulia Salsabila[2], M. Akhdan Putra Hermawan[3], Cika Jelita[4], Hanif Zaidan Sinaga[5]
[1,2,3,4,5] Program Studi Bisnis Digital, Fakultas Ekonomi dan Bisnis, Universitas Ibn Khaldun Bogor,
[1]akbarsupiad20@gmail.com, [2]lanaaulia627@gmail.com, [3]akhdanptr78@gmail.com, [4]cika0312@gmail.com,
[5*]hanif.zaidan@gmail.com

Abstrak
Analisis sentimen merupakan salah satu penerapan Natural Language Processing yang digunakan untuk mengidentifikasi kecenderungan opini masyarakat ke dalam kategori positif, negatif, dan netral. Penelitian ini bertujuan menganalisis sentimen masyarakat terhadap Program Makan Bergizi Gratis pada media sosial X serta membandingkan kinerja algoritma Support Vector Machine dan Naive Bayes. Data penelitian menggunakan dataset publik yang diperoleh melalui platform Kaggle dan berasal dari unggahan pengguna media sosial X mengenai Program Makan Bergizi Gratis. Setelah proses pembersihan, penelitian menggunakan 1.335 data yang terdiri atas 612 sentimen positif, 377 sentimen negatif, dan 346 sentimen netral. Tahapan pengolahan data meliputi cleaning, case folding, tokenizing, stopword removal, stemming, dan pembobotan fitur menggunakan Term Frequency-Inverse Document Frequency. Hasil pengujian menunjukkan bahwa Support Vector Machine dengan kernel linear memperoleh akurasi sebesar 77,9% dan Macro F1-Score sebesar 0,78. Sementara itu, Multinomial Naive Bayes memperoleh akurasi sebesar 75,7% dan Macro F1-Score sebesar 0,75. Berdasarkan hasil tersebut, Support Vector Machine memiliki performa yang lebih baik dalam mengklasifikasikan sentimen masyarakat, terutama pada kelas netral. Temuan penelitian menunjukkan bahwa sentimen positif berhubungan dengan manfaat program dan peningkatan gizi siswa, sedangkan sentimen negatif didominasi oleh kekhawatiran terhadap anggaran, efisiensi, dan pelaksanaan program. Hasil penelitian ini dapat menjadi gambaran awal mengenai respons masyarakat terhadap kebijakan Program Makan Bergizi Gratis di media sosial.

Kata kunci: Makan Bergizi Gratis, Analisis Sentimen, Media Sosial X, Support Vector Machine, Naive Bayes

1. Pendahuluan

Suhendra dan Pratiwi (2024) menjelaskan bahwa perkembangan teknologi informasi dan komunikasi telah mengubah cara masyarakat memperoleh informasi, berinteraksi, serta menyampaikan pendapat mengenai berbagai isu publik. Media sosial saat ini tidak hanya digunakan sebagai sarana komunikasi, tetapi juga menjadi wadah bagi masyarakat untuk mengekspresikan opini, kritik, dan dukungan terhadap berbagai kebijakan pemerintah. Melalui platform media sosial, opini publik dapat tersebar secara cepat dan menjangkau audiens yang luas sehingga mampu membentuk persepsi masyarakat terhadap suatu isu tertentu.

Suhendra dan Pratiwi (2024) juga menjelaskan bahwa media sosial telah berkembang menjadi ruang publik digital yang memungkinkan masyarakat untuk menyampaikan opini dan berpartisipasi dalam diskusi mengenai berbagai isu sosial maupun kebijakan pemerintah. Salah satu platform yang banyak digunakan untuk menyampaikan pendapat adalah media sosial X. Karakteristik X yang memungkinkan pengguna menyampaikan opini secara langsung menjadikan platform ini sebagai sumber data yang relevan dalam analisis sentimen. Melalui unggahan dan komentar yang dibagikan pengguna, berbagai persepsi masyarakat terhadap suatu kebijakan dapat dianalisis secara lebih cepat dan luas.

Purwanti dan Sugiyono (2024) menjelaskan bahwa Program Makan Bergizi Gratis (MBG) merupakan salah satu program pemerintah yang bertujuan meningkatkan kualitas gizi masyarakat, khususnya bagi anak-anak usia sekolah, serta mendukung upaya penurunan angka stunting. Sejak diperkenalkan, program ini menjadi salah satu topik yang banyak diperbincangkan oleh masyarakat di media sosial. Berbagai tanggapan muncul, mulai dari dukungan terhadap manfaat program hingga kritik mengenai implementasi dan efektivitasnya. Beragamnya opini yang berkembang menjadikan Program Makan Bergizi Gratis (MBG) sebagai objek yang menarik untuk dianalisis menggunakan pendekatan analisis sentimen.

Menurut Sipayung dan Hakim (2024), analisis sentimen merupakan salah satu cabang dari Natural Language Processing (NLP) yang digunakan untuk mengidentifikasi, mengekstraksi, dan mengklasifikasikan opini atau emosi yang terkandung dalam suatu teks ke dalam kategori tertentu, seperti positif, negatif, dan netral. Analisis sentimen banyak digunakan untuk memahami persepsi masyarakat terhadap suatu produk, layanan, kebijakan, maupun isu sosial yang berkembang di media sosial. Dengan memanfaatkan teknik analisis sentimen, data teks dalam jumlah besar dapat diolah menjadi informasi yang lebih mudah dipahami dan digunakan sebagai dasar pengambilan keputusan.

Maulana et al. (2024) menjelaskan bahwa algoritma Support Vector Machine (SVM) merupakan salah satu metode machine learning yang memiliki performa baik dalam tugas klasifikasi teks. Algoritma ini mampu menangani data berdimensi tinggi dan menghasilkan tingkat akurasi yang relatif tinggi dibandingkan beberapa metode klasifikasi lainnya. Selain itu, Ma'rufudin & Yudhistira (2025) menunjukkan bahwa SVM efektif digunakan dalam analisis sentimen pada data media sosial karena mampu mengklasifikasikan sentimen positif, negatif, dan netral dengan performa yang baik. Oleh karena itu, algoritma SVM dipilih dalam penelitian ini untuk melakukan klasifikasi sentimen masyarakat terhadap Program Makan Bergizi Gratis (MBG).

Selain SVM, algoritma Naive Bayes juga banyak digunakan dalam analisis sentimen karena kesederhanaan dan efisiensinya dalam menangani data teks berskala besar. Saputra dkk. (2025) dalam penelitiannya yang membandingkan performa algoritma Naive Bayes dan SVM pada teks media sosial menemukan bahwa SVM mencapai akurasi 88,85% dan F1-score 88,86%, lebih unggul dibandingkan Naive Bayes yang hanya mencapai akurasi 72,64% dan F1-score 72,26%, terutama dalam memprediksi sentimen netral. Berdasarkan temuan tersebut, penelitian ini membandingkan performa algoritma SVM dan Naive Bayes dalam mengklasifikasikan sentimen masyarakat terhadap Program Makan Bergizi Gratis (MBG), guna mengetahui metode yang lebih efektif diterapkan pada data komentar media sosial.

Magdalena dan Hakim (2024) , Saputra dkk. (2025), serta Ma'rufudin & Yudhistira (2025) menunjukkan bahwa analisis sentimen telah digunakan untuk mengkaji opini masyarakat terhadap kebijakan publik dan membandingkan performa beberapa algoritma klasifikasi pada data media sosial. Namun, masih diperlukan kajian yang secara khusus membandingkan kemampuan Support Vector Machine dan Naive Bayes dalam mengklasifikasikan sentimen masyarakat terhadap Program Makan Bergizi Gratis pada media sosial X. Perbandingan kedua algoritma tersebut penting karena komentar masyarakat di media sosial memiliki karakteristik bahasa yang beragam, seperti penggunaan bahasa tidak baku, singkatan, serta kemiripan kosakata antara sentimen positif, negatif, dan netral. Penelitian ini tidak hanya berfokus pada nilai akurasi model, tetapi juga menganalisis precision, recall, dan F1-score pada setiap kelas sentimen untuk mengetahui kemampuan masing-masing algoritma dalam mengenali opini masyarakat secara lebih menyeluruh.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap Program Makan Bergizi Gratis pada media sosial X serta membandingkan performa algoritma Support Vector Machine dan Naive Bayes dalam mengklasifikasikan sentimen positif, negatif, dan netral. Perbandingan dilakukan menggunakan metrik accuracy, precision, recall, dan F1-score untuk menentukan algoritma yang memberikan performa terbaik pada dataset yang digunakan. Selain membandingkan performa model, penelitian ini juga bertujuan mengidentifikasi kecenderungan opini masyarakat, seperti dukungan terhadap manfaat program, kekhawatiran terhadap pengelolaan anggaran dan efisiensi, serta kebutuhan informasi mengenai pelaksanaan program. Hasil penelitian diharapkan dapat memberikan gambaran yang lebih terukur mengenai penerimaan dan kekhawatiran masyarakat terhadap Program Makan Bergizi Gratis.

2. Metode Penelitian

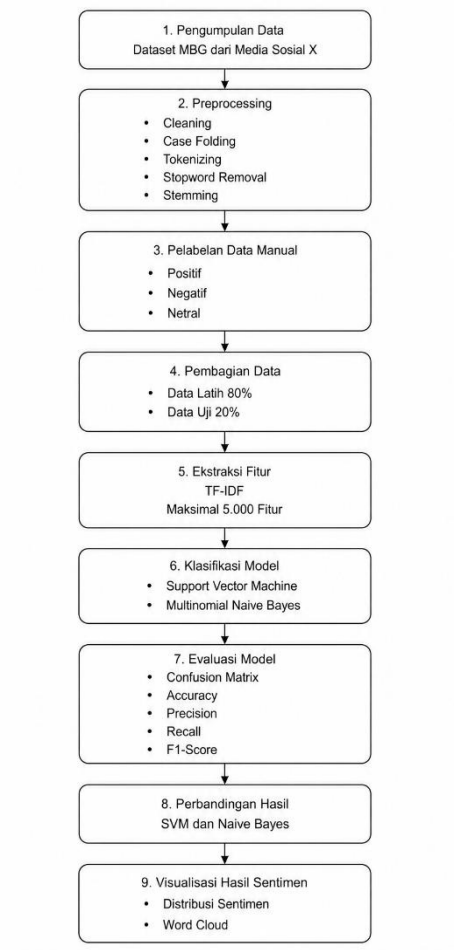
Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis sentimen berbasis Natural Language Processing (NLP). Penelitian ini bertujuan untuk mengidentifikasi dan mengklasifikasikan sentimen masyarakat terhadap Program Makan Bergizi Gratis (MBG) pada media sosial X ke dalam kategori positif, negatif, dan netral. Proses klasifikasi dilakukan menggunakan algoritma Support Vector Machine (SVM) dan Naive Bayes untuk membandingkan performa kedua algoritma berdasarkan nilai accuracy, precision, recall, dan F1-score

Gambar 1 menunjukkan tahapan penelitian analisis sentimen masyarakat terhadap Program Makan Bergizi Gratis (MBG) pada media sosial X. Penelitian dimulai dengan pengumpulan data berupa teks unggahan atau komentar masyarakat mengenai program MBG yang diperoleh dari dataset pada platform Kaggle.

Data yang telah dikumpulkan kemudian melalui tahap preprocessing untuk membersihkan dan menyeragamkan teks. Tahapan preprocessing meliputi cleaning, case folding, tokenizing, stopword removal, dan stemming. Setelah preprocessing, dilakukan pelabelan data secara manual ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral, berdasarkan isi dan kecenderungan makna dari setiap komentar. Dataset yang telah melalui preprocessing dan pelabelan selanjutnya dibagi menjadi data latih sebesar 80% dan data uji sebesar 20%. Pembagian data dilakukan menggunakan fungsi train_test_split dengan parameter random_state=42 dan stratify berdasarkan label sentimen agar proporsi setiap kelas tetap terjaga.

DOI: https://doi.org/10.69693/ijmst.v4i2.11800

Gambar 1. Alur Penelitian Analisis Sentimen Menggunakan Algoritma SVM dan Naive Bayes



Setelah proses pembagian data, dilakukan ekstraksi fitur menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). Pembentukan kosakata dan pembobotan TF-IDF dilakukan pada data latih menggunakan fit_transform, sedangkan data uji ditransformasi menggunakan kosakata yang telah dibentuk dari data latih. Jumlah fitur dibatasi maksimal sebanyak 5.000 fitur. Hasil transformasi TF-IDF digunakan sebagai data masukan dalam proses klasifikasi menggunakan Support Vector Machine dan Multinomial Naive Bayes. Model SVM menggunakan kernel linear dengan parameter (C = 1,0), sedangkan Multinomial Naive Bayes menggunakan parameter smoothing alpha = 1,0.

Kinerja kedua model dievaluasi menggunakan confusion matrix serta metrik accuracy, precision, recall, dan F1-score. Hasil evaluasi kemudian dibandingkan untuk menentukan model yang memberikan performa lebih baik. Tahap terakhir adalah visualisasi hasil sentimen dalam bentuk distribusi kelas sentimen dan word cloud.

2.1. Tahapan Penelitian

Tahapan penelitian terdiri atas pengumpulan data, preprocessing teks, pelabelan data secara manual, pembagian data latih dan data uji, ekstraksi fitur menggunakan Term Frequency-Inverse Document Frequency (TF-IDF), klasifikasi menggunakan Support Vector Machine dan Multinomial Naive Bayes, evaluasi kinerja model, perbandingan hasil kedua algoritma, serta visualisasi hasil sentimen.

Seluruh tahapan dilakukan secara berurutan untuk menghasilkan model yang mampu mengklasifikasikan komentar masyarakat ke dalam kategori sentimen positif, negatif, dan netral. Urutan tahapan penelitian disesuaikan dengan proses pengolahan data pada kode penelitian, yaitu pembagian dataset dilakukan terlebih dahulu sebelum pembentukan fitur TF-IDF.

2.2 Pengumpulan Data

Data penelitian diperoleh dari dataset publik berjudul *Sentimen Publik Terhadap Makan Bergizi Gratis* yang diunggah oleh pengguna Hanief_Nief melalui platform Kaggle. Dataset tersebut diakses pada 23 Juni 2026 melalui URL:

DOI: https://doi.org/10.69693/ijmst.v4i2.11800

https://www.kaggle.com/datasets/hanif281103/sentimen-publik-terhadap-makan-bergizi-gratis. Data dalam dataset berasal dari komentar pengguna Twitter atau media sosial X mengenai Program Makan Bergizi Gratis (MBG).

Dataset awal berjumlah 3.462 komentar. Kolom yang digunakan dalam penelitian ini adalah kolom komentar karena memuat teks opini masyarakat mengenai program MBG. Selanjutnya, dilakukan proses pemeriksaan, pembersihan, dan seleksi data dengan menghapus data kosong, data duplikat, serta komentar yang tidak relevan dengan topik penelitian. Setelah proses tersebut, diperoleh sebanyak 1.335 komentar yang digunakan dalam tahap analisis.

Dataset dari Kaggle hanya digunakan sebagai sumber data teks. Data yang telah melalui proses pembersihan kemudian diberi label sentimen secara manual dengan bantuan AI ke dalam tiga kategori, yaitu positif, negatif, dan netral. Penentuan akhir label dilakukan oleh peneliti berdasarkan isi dan kecenderungan makna dari setiap komentar.

2.3. Preprocessing Data

Tahap preprocessing dilakukan untuk membersihkan dan menyeragamkan data teks sebelum digunakan dalam proses ekstraksi fitur dan klasifikasi. Proses ini bertujuan mengurangi unsur yang tidak relevan serta meningkatkan kualitas representasi teks. Tahapan preprocessing yang digunakan dalam penelitian ini meliputi cleaning, case folding, tokenizing, stopword removal, dan stemming.

- **Cleaning**
  Tahap cleaning dilakukan dengan menghapus unsur yang tidak diperlukan dalam teks, seperti URL, mention, hashtag, emoji, angka, tanda baca, karakter khusus, serta spasi berlebih.
- **Case Folding**
  Case folding dilakukan dengan mengubah seluruh huruf pada komentar menjadi huruf kecil agar perbedaan penggunaan huruf kapital tidak dianggap sebagai kata yang berbeda.
- **Tokenizing**
  Tokenizing merupakan proses memecah kalimat menjadi unit-unit kata atau token sehingga setiap kata dapat diproses secara terpisah.
- **Stopword Removal**
  Stopword removal dilakukan untuk menghapus kata-kata umum yang memiliki frekuensi tinggi tetapi tidak memberikan pengaruh besar terhadap penentuan sentimen, seperti kata “yang”, “dan”, “di”, atau “ke”.
- **Stemming**
  Stemming dilakukan untuk mengubah kata berimbuhan menjadi bentuk kata dasar. Proses ini bertujuan menyatukan kata-kata yang memiliki makna dasar yang sama sehingga tidak dianggap sebagai fitur yang berbeda.

2.4. Pelabelan Data

Pelabelan sentimen dilakukan secara manual oleh peneliti dengan bantuan AI sebagai alat bantu dalam mengidentifikasi kecenderungan awal sentimen pada setiap komentar. AI tidak digunakan sebagai penentu akhir label. Peneliti tetap membaca, memeriksa, dan menetapkan label akhir berdasarkan isi serta konteks komentar terhadap Program Makan Bergizi Gratis (MBG).

Setiap komentar dikategorikan ke dalam tiga kelas sentimen, yaitu positif, negatif, dan netral. Komentar diberi label positif apabila mengandung dukungan, apresiasi, harapan baik, atau penilaian yang menunjukkan manfaat program MBG. Komentar diberi label negatif apabila mengandung kritik, ketidaksetujuan, kekecewaan, kekhawatiran, atau penilaian buruk terhadap program. Sementara itu, komentar diberi label netral apabila hanya menyampaikan informasi, fakta, pertanyaan, atau pernyataan yang tidak menunjukkan kecenderungan sentimen positif maupun negatif.

Untuk menjaga konsistensi pelabelan, kriteria sentimen tersebut diterapkan secara seragam pada seluruh komentar, yaitu sentimen positif untuk komentar yang mengandung dukungan atau manfaat, sentimen negatif untuk komentar yang mengandung kritik atau kekhawatiran, serta sentimen netral untuk komentar yang bersifat informatif atau berupa pertanyaan tanpa kecenderungan opini tertentu.

Berdasarkan hasil pelabelan terhadap 1.335 komentar, diperoleh 612 komentar dengan sentimen positif, 377 komentar dengan sentimen negatif, dan 346 komentar dengan sentimen netral.

2.5. Pembagian Data Latih dan Data Uji

Dataset yang telah melalui tahap preprocessing selanjutnya dibagi menjadi data latih dan data uji menggunakan fungsi train_test_split. Pembagian dilakukan dengan proporsi 80% sebagai data latih dan 20% sebagai data uji. Dari total 1.335 data, sebanyak 1.068 data digunakan sebagai data latih dan 267 data digunakan sebagai data uji.

Parameter random_state=42 digunakan agar pembagian data menghasilkan komposisi yang konsisten ketika proses pengujian dijalankan kembali. Selain itu, parameter stratify diterapkan berdasarkan label sentimen untuk

DOI: https://doi.org/10.69693/ijmst.v4i2.11800

menjaga agar proporsi kelas positif, negatif, dan netral pada data latih dan data uji tetap menyerupai distribusi kelas pada keseluruhan dataset.
Data latih digunakan untuk membentuk fitur TF-IDF dan melatih model Support Vector Machine serta Multinomial Naive Bayes. Sementara itu, data uji digunakan untuk mengevaluasi kemampuan kedua model dalam mengklasifikasikan komentar yang tidak digunakan selama proses pelatihan.
Skema evaluasi yang digunakan dalam penelitian ini adalah stratified hold-out validation, yaitu membagi dataset satu kali menjadi data latih dan data uji dengan tetap mempertahankan proporsi setiap kelas sentimen.

2.6. Ekstraksi Fitur Menggunakan TF-IDF

Data teks yang telah melalui tahap preprocessing selanjutnya diubah menjadi representasi numerik menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). Metode ini digunakan untuk memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam suatu komentar dan tingkat kepentingannya terhadap keseluruhan dokumen dalam dataset.
Term Frequency menunjukkan seberapa sering suatu kata muncul dalam sebuah dokumen, sedangkan Inverse Document Frequency digunakan untuk mengurangi bobot kata yang terlalu sering muncul pada banyak dokumen. Dengan demikian, kata yang lebih spesifik dan memiliki nilai informasi lebih tinggi akan memperoleh bobot yang lebih besar.
Proses ekstraksi fitur dilakukan menggunakan TfidfVectorizer dengan jumlah fitur maksimal sebanyak 5.000 melalui parameter max_features=5000. Parameter ngram_range menggunakan nilai bawaan (1,1), sehingga fitur yang digunakan berupa kata tunggal atau unigram. Proses fit_transform diterapkan pada data latih untuk membentuk kosakata dan matriks TF-IDF, sedangkan proses transform diterapkan pada data uji menggunakan kosakata yang telah dibentuk dari data latih. Hasil pembobotan TF-IDF kemudian digunakan sebagai data masukan dalam proses klasifikasi menggunakan Support Vector Machine dan Multinomial Naive Bayes.

2.7. Perangkat dan Library Penelitian

Proses pengolahan dan analisis data dilakukan menggunakan bahasa pemrograman Python melalui Google Colaboratory. Library pandas digunakan untuk membaca dan mengelola dataset. Library scikit-learn digunakan untuk membagi data, membentuk fitur TF-IDF, membangun model Support Vector Machine dan Multinomial Naive Bayes, serta mengevaluasi performa model. Visualisasi hasil dilakukan menggunakan Matplotlib dan Seaborn, sedangkan visualisasi kata menggunakan library WordCloud.

2.8. Klasifikasi Menggunakan Support Vector Machine (SVM)

Support Vector Machine (SVM) digunakan sebagai salah satu algoritma untuk mengklasifikasikan komentar masyarakat ke dalam tiga kelas sentimen, yaitu positif, negatif, dan netral. Model SVM menerima data masukan berupa matriks fitur numerik yang dihasilkan melalui proses pembobotan TF-IDF.
Pada penelitian ini, model Support Vector Machine dibangun menggunakan kernel linear dengan nilai parameter (C = 1,0). Untuk mengurangi pengaruh ketidakseimbangan jumlah data pada setiap kelas sentimen, digunakan parameter class_weight='balanced'. Parameter tersebut memberikan bobot yang lebih besar pada kelas dengan jumlah data lebih sedikit sehingga proses pelatihan tidak terlalu didominasi oleh kelas mayoritas. Selain itu, parameter probability=True digunakan agar model dapat menghasilkan estimasi probabilitas untuk setiap kelas sentimen.
Setelah proses pelatihan menggunakan data latih selesai, model digunakan untuk memprediksi kelas sentimen pada data uji. Hasil prediksi kemudian dievaluasi menggunakan accuracy, precision, recall, F1-score, dan confusion matrix.

2.8.1. Klasifikasi Menggunakan Multinomial Naive Bayes

Multinomial Naive Bayes digunakan sebagai algoritma pembanding untuk mengklasifikasikan komentar masyarakat ke dalam tiga kelas sentimen, yaitu positif, negatif, dan netral. Algoritma ini dipilih karena sesuai digunakan pada data teks yang telah direpresentasikan dalam bentuk numerik melalui pembobotan TF-IDF.
Model Multinomial Naive Bayes bekerja dengan menghitung probabilitas suatu komentar termasuk ke dalam setiap kelas sentimen berdasarkan bobot fitur kata yang terdapat dalam komentar tersebut. Kelas dengan nilai probabilitas tertinggi kemudian ditetapkan sebagai hasil prediksi.
Pada penelitian ini, model Multinomial Naive Bayes menggunakan parameter smoothing alpha=1,0. Model dilatih menggunakan data latih hasil transformasi TF-IDF, kemudian digunakan untuk memprediksi label sentimen pada data uji. Evaluasi model dilakukan menggunakan accuracy, precision, recall, F1-score, dan confusion matrix, kemudian hasilnya dibandingkan dengan performa model Support Vector Machine.

2.8.2. Evaluasi Model

Evaluasi model dilakukan untuk mengukur dan membandingkan kinerja algoritma Support Vector Machine dan Multinomial Naive Bayes dalam mengklasifikasikan sentimen positif, negatif, dan netral. Pengujian dilakukan menggunakan data uji yang berjumlah 20% dari keseluruhan dataset.
Kinerja kedua model dievaluasi menggunakan confusion matrix serta beberapa metrik, yaitu accuracy, precision, recall, dan F1-score. Confusion matrix digunakan untuk menunjukkan jumlah prediksi yang benar dan salah pada setiap kelas sentimen. Accuracy digunakan untuk mengukur persentase keseluruhan prediksi yang benar,

DOI: https://doi.org/10.69693/ijmst.v4i2.11800

sedangkan precision menunjukkan ketepatan model dalam memprediksi suatu kelas sentimen. Recall digunakan untuk mengukur kemampuan model dalam menemukan seluruh data yang sebenarnya termasuk dalam suatu kelas, sementara F1-score merupakan nilai rata-rata harmonis antara precision dan recall.

Karena penelitian ini menggunakan tiga kelas sentimen, nilai precision, recall, dan F1-score dihitung pada setiap kelas, yaitu positif, negatif, dan netral. Selain itu, digunakan nilai macro average dan weighted average untuk memberikan gambaran performa model secara keseluruhan. Hasil evaluasi kedua algoritma kemudian dibandingkan untuk menentukan model yang memberikan performa terbaik pada dataset penelitian.

2.8.3. Perbandingan Hasil Model

Hasil evaluasi model Support Vector Machine dan Multinomial Naive Bayes dibandingkan berdasarkan nilai accuracy, precision, recall, dan F1-score. Perbandingan dilakukan untuk mengetahui algoritma yang memberikan performa lebih baik dalam mengklasifikasikan sentimen masyarakat terhadap Program Makan Bergizi Gratis.

Selain membandingkan nilai akurasi secara keseluruhan, analisis juga dilakukan pada performa setiap kelas sentimen, yaitu positif, negatif, dan netral. Hal ini dilakukan karena nilai akurasi saja belum cukup untuk menggambarkan kemampuan model dalam mengenali seluruh kelas sentimen, terutama ketika jumlah data pada setiap kelas tidak sama.

Model dengan nilai accuracy dan Macro F1-score yang lebih tinggi dianggap memiliki performa yang lebih baik secara keseluruhan. Hasil perbandingan kedua algoritma selanjutnya digunakan sebagai dasar dalam menentukan model yang paling sesuai untuk klasifikasi sentimen pada dataset penelitian.

2.8.4. Alur Penelitian

Alur penelitian dimulai dari pengumpulan data komentar masyarakat mengenai Program Makan Bergizi Gratis (MBG) pada media sosial X. Dataset awal yang diperoleh melalui platform Kaggle berjumlah 3.462 data. Setelah dilakukan pemeriksaan dan seleksi data, diperoleh sebanyak 1.335 komentar yang digunakan dalam penelitian.

Data yang telah dipilih kemudian melalui tahap preprocessing yang meliputi cleaning, case folding, tokenizing, stopword removal, dan stemming. Setelah preprocessing, setiap komentar diberi label secara manual oleh peneliti ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral.

Dataset selanjutnya dibagi menjadi data latih sebesar 80% atau 1.068 data dan data uji sebesar 20% atau 267 data menggunakan fungsi train_test_split. Parameter random_state=42 digunakan agar hasil pembagian data tetap konsisten ketika proses dijalankan kembali, sedangkan parameter stratify digunakan untuk mempertahankan proporsi setiap kelas sentimen pada data latih dan data uji.

Setelah pembagian data, dilakukan ekstraksi fitur menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). Proses ekstraksi fitur menggunakan TfidfVectorizer dengan jumlah fitur maksimal sebanyak 5.000. Pembentukan kosakata dan pembobotan TF-IDF dilakukan hanya pada data latih menggunakan fit_transform, sedangkan data uji ditransformasi menggunakan kosakata dari data latih melalui fungsi transform.

Hasil transformasi TF-IDF digunakan untuk melatih model Support Vector Machine dan Multinomial Naive Bayes. Model SVM menggunakan kernel linear dengan parameter (C = 1,0), *class_weight='balanced'*, dan *probability=True*, sedangkan Multinomial Naive Bayes menggunakan parameter smoothing *alpha = 1,0*.

Kinerja kedua model dievaluasi menggunakan confusion matrix serta metrik accuracy, precision, recall, dan F1-score. Hasil evaluasi kedua algoritma kemudian dibandingkan untuk menentukan model yang memberikan performa lebih baik dalam mengklasifikasikan sentimen masyarakat terhadap Program Makan Bergizi Gratis. Tahap terakhir adalah visualisasi distribusi sentimen dan word cloud untuk menggambarkan kata-kata yang dominan pada setiap kategori sentimen.

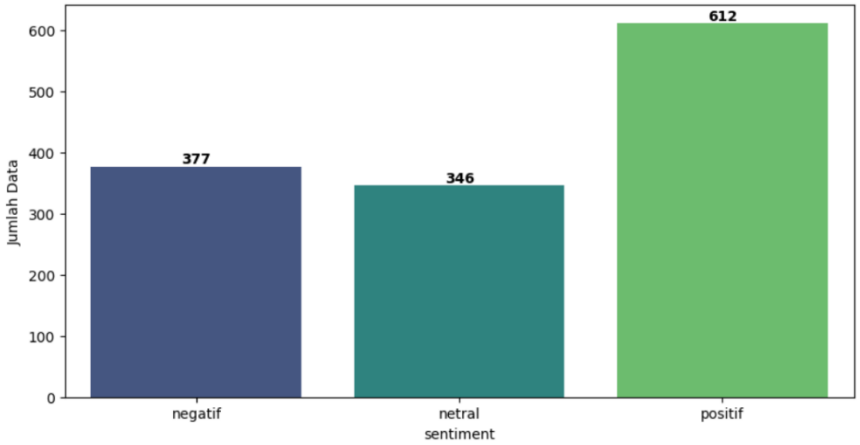
3. Hasil dan Pembahasan

3.1. Pengumpulan Data

Dataset yang digunakan dalam penelitian ini diperoleh dari platform Kaggle yang berisi komentar masyarakat mengenai Program Makan Bergizi Gratis (MBG). Setelah proses seleksi data, sebanyak 1.335 komentar diberi label secara manual oleh peneliti ke dalam tiga kelas sentimen, yaitu positif, negatif, dan netral. Jumlah data yang digunakan dalam penelitian ini sebanyak 1.335 komentar.

Sebelum dilakukan proses klasifikasi, data terlebih dahulu diperiksa untuk memastikan tidak terdapat data kosong (missing value) maupun duplikasi yang dapat memengaruhi proses analisis. Berdasarkan hasil pemeriksaan, seluruh data dapat digunakan pada tahapan penelitian selanjutnya.

DOI: https://doi.org/10.69693/ijmst.v4i2.11800



Gambar 2. Distribusi Kelas Sentimen

Berdasarkan Gambar 2, penelitian ini menggunakan sebanyak 1.335 data yang telah melalui tahap pembersihan dan pelabelan. Distribusi sentimen terdiri atas 612 data positif, 377 data negatif, dan 346 data netral. Distribusi tersebut menunjukkan bahwa kelas positif memiliki jumlah data paling banyak, sedangkan kelas netral memiliki jumlah data paling sedikit. Untuk mengurangi pengaruh perbedaan jumlah data antarkelas, model Support Vector Machine menggunakan pembobotan kelas melalui parameter *class_weight='balanced'*. Oleh karena itu, evaluasi model tidak hanya menggunakan accuracy, tetapi juga precision, recall, dan F1-score pada setiap kelas untuk memperoleh gambaran performa yang lebih menyeluruh.

3.2. Hasil Preprocessing Data

Tahap preprocessing dilakukan untuk meningkatkan kualitas data teks sebelum digunakan dalam proses klasifikasi menggunakan algoritma Support Vector Machine dan Multinomial Naive Bayes. Tahapan preprocessing yang dilakukan meliputi cleaning, case folding, tokenizing, stopword removal, dan stemming.

Pada tahap *cleaning*, karakter khusus, tanda baca, angka, tautan (*URL*), serta simbol yang tidak diperlukan dihapus dari komentar. Selanjutnya dilakukan *case folding* untuk mengubah seluruh huruf menjadi huruf kecil (*lowercase*). Setelah itu dilakukan *tokenizing* untuk memisahkan kalimat menjadi kumpulan kata (*token*). Kata-kata umum yang tidak memiliki makna penting dihilangkan melalui proses *stopword removal*, kemudian setiap kata dikembalikan ke bentuk dasarnya menggunakan proses *stemming*.

Contoh hasil preprocessing ditunjukkan pada Tabel berikut.

Tabel 1. Contoh Hasil Preprocessing Data Komentar

Tahapan	Hasil
Data Asli	Program MBG Sangat Membantu Siswa!!!
Cleaning	Program MBG Sangat Membantu Siswa
Case Folding	program mbg sangat membantu siswa
Tokenizing	program, mbg, sangat, membantu, siswa
Stopword Removal	program, mbg, membantu, siswa
Stemming	program, mbg, bantu, siswa

3.3. Hasil Ekstraksi Fitur Menggunakan TF-IDF

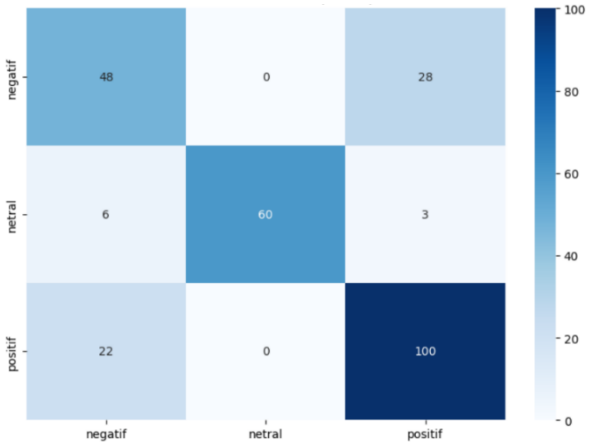
Setelah proses preprocessing selesai dilakukan, setiap komentar diubah menjadi representasi numerik menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF). Metode ini digunakan untuk memberikan bobot pada setiap kata berdasarkan tingkat kemunculannya pada keseluruhan dokumen.

Hasil ekstraksi fitur menghasilkan matriks TF-IDF yang kemudian digunakan sebagai data masukan pada proses pelatihan model Support Vector Machine dan Multinomial Naive Bayes. Melalui TF-IDF, setiap komentar direpresentasikan dalam bentuk vektor numerik sehingga dapat diproses oleh kedua algoritma klasifikasi.

DOI: https://doi.org/10.69693/ijmst.v4i2.11800

3.4. Hasil Pelatihan Model Support Vector Machine (SVM)

Berdasarkan hasil pengujian menggunakan kernel linear dengan parameter (C = 1,0), *class_weight='balanced'*, dan *probability=True*, model SVM memperoleh akurasi sebesar 77,9%. Sebagaimana divisualisasikan dalam confusion matrix pada Gambar 3, model ini menunjukkan performa yang baik dalam mengenali kelas netral, dengan keberhasilan memprediksi 60 data dari total 69 sampel uji secara akurat. Hasil tersebut menunjukkan bahwa SVM mampu membentuk hyperplane pemisah yang optimal meskipun terdapat kemiripan kata kunci antara sentimen netral dan positif pada data teks media sosial X.



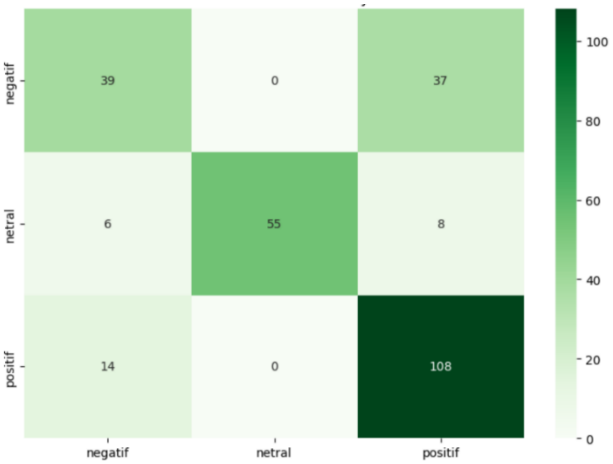
Gambar 3. Matriks Evaluasi Model SVM

3.5. Hasil Pelatihan Model Multinomial Naive Bayes

Berdasarkan hasil pengujian, model Multinomial Naive Bayes memperoleh akurasi sebesar 75,7%. Sebagaimana ditunjukkan pada confusion matrix pada Gambar 4, model berhasil memprediksi 39 data negatif, 55 data netral, dan 108 data positif dengan benar.

Pada kelas negatif, sebanyak 37 data salah diprediksi sebagai sentimen positif. Pada kelas netral, sebanyak 6 data salah diprediksi sebagai sentimen negatif dan 8 data salah diprediksi sebagai sentimen positif. Sementara itu, pada kelas positif, sebanyak 14 data salah diprediksi sebagai sentimen negatif.

Hasil tersebut menunjukkan bahwa Multinomial Naive Bayes memiliki kemampuan yang baik dalam mengenali sentimen positif, tetapi masih mengalami kesulitan dalam membedakan beberapa komentar negatif dari komentar positif.



Gambar 4. Matriks Evaluasi Model Multinomial Naive Bayes

3.6. Evaluasi Model

Evaluasi model dilakukan untuk mengukur dan membandingkan kinerja Support Vector Machine dan Multinomial Naive Bayes dalam mengklasifikasikan sentimen positif, negatif, dan netral. Pengujian dilakukan terhadap 267 data uji atau 20% dari keseluruhan dataset.

DOI: https://doi.org/10.69693/ijmst.v4i2.11800

Kinerja kedua model diukur menggunakan accuracy, precision, recall, dan F1-score. Accuracy digunakan untuk menunjukkan proporsi keseluruhan prediksi yang benar. Precision menunjukkan ketepatan model dalam memprediksi suatu kelas sentimen, sedangkan recall mengukur kemampuan model dalam mengenali seluruh data yang sebenarnya termasuk dalam kelas tersebut. F1-score digunakan untuk menggambarkan keseimbangan antara precision dan recall.

Selain itu, confusion matrix digunakan untuk menunjukkan jumlah prediksi yang benar dan kesalahan klasifikasi pada setiap kelas sentimen. Hasil evaluasi juga disajikan dalam classification report yang memuat nilai precision, recall, F1-score, dan support untuk setiap kelas, serta nilai macro average dan weighted average sebagai gambaran performa model secara keseluruhan.

Tabel **2.** Hasil Classification Report Model SVM

	Precision	Recall	F1-score	Support
Negatif	0.63	0.63	0.63	76
Netral	1.00	0.87	0.93	69
Positif	0.76	0.82	0.79	122
Accuracy			0.78	1
Macro avg	0.80	0.77	0.78	267
Weighted avg	0.79	0.78	0.78	267

3.7. Hasil Support Vector Machine (SVM)

Model Support Vector Machine memperoleh accuracy sebesar 0,78 atau 78%. Pada kelas negatif, model menghasilkan precision 0,63, recall 0,63, dan F1-score 0,63. Nilai tersebut menunjukkan bahwa model mampu mengenali 63% dari seluruh data negatif dan 63% dari prediksi negatif yang dihasilkan merupakan prediksi yang benar.

Pada kelas netral, model memperoleh precision 1,00, recall 0,87, dan F1-score 0,93. Hasil ini menunjukkan bahwa seluruh data yang diprediksi sebagai netral merupakan prediksi yang benar, meskipun masih terdapat sebagian data netral yang belum berhasil dikenali oleh model.

Pada kelas positif, model memperoleh precision 0,76, recall 0,82, dan F1-score 0,79. Hal ini menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam mengenali sentimen positif.

Secara keseluruhan, model SVM memperoleh Macro F1-Score sebesar 0,78 dan Weighted F1-Score sebesar 0,78. Nilai tersebut menunjukkan bahwa performa model relatif konsisten dalam mengklasifikasikan ketiga kelas sentimen.

Tabel 3. Hasil Classification Report Model Multinomial Naive Bayes

	Precision	Recall	F1-score	support
Negatif	0.66	0.51	0.58	76
Netral	1.00	0.80	0.89	69
Positif	0.71	0.89	0.79	122
Accuracy			0.76	1
Macro avg	0.79	0.73	0.75	267
Weighted avg	0.77	0.76	0.75	267

3.8. Hasil Multinomial Naive Bayes

Model Multinomial Naive Bayes memperoleh accuracy sebesar 0,76 atau 76%. Pada kelas negatif, model menghasilkan precision 0,66, recall 0,51, dan F1-score 0,58. Nilai recall yang lebih rendah menunjukkan bahwa model masih mengalami kesulitan dalam mengenali seluruh data negatif.

Pada kelas netral, model memperoleh precision 1,00, recall 0,80, dan F1-score 0,89. Hasil ini menunjukkan bahwa seluruh data yang diprediksi sebagai netral merupakan prediksi yang benar, meskipun masih terdapat sebagian data netral yang belum berhasil dikenali oleh model.

Pada kelas positif, model memperoleh precision 0,71, recall 0,89, dan F1-score 0,79. Nilai recall yang tinggi menunjukkan bahwa model mampu mengenali sebagian besar data positif, tetapi nilai precision yang lebih rendah menunjukkan bahwa beberapa data dari kelas lain masih salah diprediksi sebagai positif.

Secara keseluruhan, model Multinomial Naive Bayes memperoleh Macro F1-Score sebesar 0,75 dan Weighted F1-Score sebesar 0,75. Hasil tersebut menunjukkan bahwa performa Multinomial Naive Bayes sedikit lebih rendah dibandingkan Support Vector Machine.

3.9. Pengujian Model

Setelah proses pelatihan dan evaluasi selesai, model Support Vector Machine diuji menggunakan beberapa komentar baru yang tidak termasuk dalam data latih maupun data uji. Pengujian tambahan ini dilakukan untuk melihat penerapan model dalam mengklasifikasikan komentar baru ke dalam kategori sentimen positif, negatif, dan netral.

Contoh hasil prediksi ditunjukkan pada Tabel 4. Komentar “Program MBG sangat membantu masyarakat” diprediksi sebagai sentimen positif, komentar “Program ini hanya menghabiskan anggaran” diprediksi sebagai

DOI: https://doi.org/10.69693/ijmst.v4i2.11800

sentimen negatif, sedangkan komentar "Program MBG mulai dilaksanakan bulan depan" diprediksi sebagai sentimen netral.

Tabel 4. Hasil Prediksi Sentimen Menggunakan Model SVM

Komentar	Hasil Prediksi
Program MBG sangat membantu masyarakat	Positif
Program ini hanya menghabiskan anggaran	Negatif
Program MBG mulai dilaksanakan bulan depan	Netral

Hasil pengujian menunjukkan bahwa model dapat memberikan prediksi yang sesuai dengan kecenderungan makna pada ketiga contoh komentar tersebut. Pengujian ini digunakan sebagai ilustrasi penerapan model, sedangkan penilaian kinerja utama tetap didasarkan pada hasil evaluasi terhadap 267 data uji.

3.10. Pembahasan

Hasil penelitian menunjukkan bahwa Support Vector Machine dan Multinomial Naive Bayes dapat digunakan untuk mengklasifikasikan sentimen masyarakat terhadap Program Makan Bergizi Gratis (MBG) pada media sosial X ke dalam kelas positif, negatif, dan netral. Tahap preprocessing berperan dalam membersihkan dan menyeragamkan data teks sehingga unsur yang tidak relevan dapat dikurangi sebelum proses klasifikasi. Selanjutnya, TF-IDF digunakan untuk mengubah data teks menjadi representasi numerik yang dapat diproses oleh kedua algoritma.

Berdasarkan hasil pengujian, Support Vector Machine memperoleh akurasi sebesar 77,9% dengan Macro F1-Score sebesar 0,78, sedangkan Multinomial Naive Bayes memperoleh akurasi sebesar 75,7% dengan Macro F1-Score sebesar 0,75. Hasil tersebut menunjukkan bahwa SVM memberikan performa keseluruhan yang sedikit lebih tinggi dibandingkan Multinomial Naive Bayes pada dataset penelitian.

Perbedaan performa kedua model terlihat pada masing-masing kelas sentimen. SVM menghasilkan F1-score sebesar 0,63 pada kelas negatif, 0,93 pada kelas netral, dan 0,79 pada kelas positif. Sementara itu, Multinomial Naive Bayes memperoleh F1-score sebesar 0,58 pada kelas negatif, 0,89 pada kelas netral, dan 0,79 pada kelas positif. Dengan demikian, kedua model menghasilkan performa yang sama pada kelas positif, tetapi SVM memiliki performa lebih baik dalam mengenali kelas negatif dan netral.

Multinomial Naive Bayes memiliki recall yang tinggi pada kelas positif, yaitu 0,89, yang menunjukkan bahwa model mampu mengenali sebagian besar komentar positif. Namun, precision pada kelas positif hanya sebesar 0,71 karena beberapa komentar negatif dan netral masih salah diprediksi sebagai positif. Pada kelas negatif, recall Multinomial Naive Bayes hanya sebesar 0,51, sehingga model masih mengalami kesulitan dalam mengenali seluruh komentar negatif. Sebaliknya, SVM memberikan hasil yang lebih seimbang pada kelas negatif dengan precision dan recall masing-masing sebesar 0,63.

Hasil tersebut menunjukkan bahwa pendekatan SVM lebih sesuai untuk dataset penelitian ini karena mampu membentuk batas pemisah antarkelas pada data teks berdimensi tinggi. Meskipun demikian, selisih akurasi kedua model hanya sekitar 2,2 poin persentase. Oleh karena itu, hasil penelitian ini menunjukkan bahwa SVM memiliki performa lebih tinggi pada dataset yang digunakan, bukan berarti selalu lebih unggul untuk seluruh dataset analisis sentimen.

Gambar **5**. Visualisasi Kata Kunci (Word Cloud) pada Sentimen Positif, Negatif, dan Netral

Visualisasi word cloud digunakan untuk menggambarkan kata-kata yang sering muncul pada setiap kelas sentimen. Ukuran kata yang lebih besar menunjukkan frekuensi kemunculan yang lebih tinggi, tetapi tidak secara langsung menunjukkan tingkat kepentingan atau makna sentimen dari kata tersebut.

DOI: https://doi.org/10.69693/ijmst.v4i2.11800

- Sentimen Positif (Gambar 5 bagian kiri)

Word cloud sentimen positif didominasi oleh kata-kata yang berkaitan langsung dengan program, seperti "makan", "bergizi", "gratis", "program", dan "siang". Kemunculan kata-kata tersebut menunjukkan bahwa komentar positif banyak membahas program makan bergizi dan manfaat yang diharapkan dari pelaksanaannya. Interpretasi sentimen positif tidak hanya didasarkan pada frekuensi kata, tetapi juga pada konteks komentar yang mengandung dukungan, apresiasi, atau harapan baik terhadap program MBG.

- Sentimen Negatif (Gambar 5 bagian tengah)

Pada sentimen negatif, kata-kata yang sering muncul antara lain "makan", "siang", "bergizi", "gratis", dan "program". Kata-kata tersebut masih berkaitan dengan topik utama MBG, tetapi muncul dalam komentar yang mengandung kritik, ketidaksetujuan, atau kekhawatiran. Hasil ini menunjukkan bahwa sentimen negatif tidak selalu dapat dikenali hanya dari satu kata tertentu, melainkan perlu dilihat berdasarkan konteks keseluruhan kalimat, seperti kritik terhadap pelaksanaan, penggunaan anggaran, kualitas makanan, atau efektivitas program.

- Sentimen Netral (Gambar 5 bagian kanan)

Word cloud sentimen netral memperlihatkan kata-kata seperti "di sekolah", "hari ini", "makan siang", "bergizi gratis", dan "sesuai jadwal". Kosakata tersebut menunjukkan bahwa komentar netral cenderung berisi informasi, laporan kegiatan, jadwal pelaksanaan, atau keterangan mengenai proses pembagian makanan tanpa menunjukkan dukungan maupun penolakan secara langsung.

3.11. Evaluasi Akhir

Hasil **Tabel 5.** Perbandingan Performa Algoritma

	Metode	Akurasi (%)	F1-Score (Macro)
0	Support Vector Machine (SVM)	77.9%	0.78
1	Naive Bayes	75.7%	0.75

Tabel 5 menyajikan perbandingan performa Support Vector Machine dan Multinomial Naive Bayes berdasarkan nilai accuracy dan Macro F1-Score. Model Support Vector Machine memperoleh accuracy sebesar 77,9% dan Macro F1-Score sebesar 0,78. Sementara itu, Multinomial Naive Bayes memperoleh accuracy sebesar 75,7% dan Macro F1-Score sebesar 0,75.

Hasil tersebut menunjukkan bahwa Support Vector Machine memiliki performa keseluruhan yang sedikit lebih tinggi dibandingkan Multinomial Naive Bayes pada dataset penelitian. Selisih accuracy kedua model sebesar 2,2 poin persentase, sedangkan selisih Macro F1-Score sebesar 0,03. SVM juga menghasilkan performa yang lebih baik pada kelas negatif dan netral, sedangkan kedua model memperoleh F1-score yang sama pada kelas positif, yaitu 0,79.

Berdasarkan hasil pengujian tersebut, Support Vector Machine dipilih sebagai model dengan performa terbaik pada dataset yang digunakan. Namun, hasil ini hanya berlaku pada karakteristik data dan parameter penelitian ini sehingga tidak dapat digunakan untuk menyimpulkan bahwa SVM selalu lebih unggul pada seluruh kasus analisis sentimen.

4. Kesimpulan

Berdasarkan hasil penelitian, Support Vector Machine dan Multinomial Naive Bayes dapat digunakan untuk mengklasifikasikan sentimen masyarakat terhadap Program Makan Bergizi Gratis pada media sosial X ke dalam kategori positif, negatif, dan netral. Penelitian menggunakan 1.335 komentar yang terdiri atas 612 sentimen positif, 377 sentimen negatif, dan 346 sentimen netral.

Hasil pengujian menunjukkan bahwa Support Vector Machine dengan kernel linear dan parameter (C = 1,0) memperoleh accuracy sebesar 77,9% dan Macro F1-Score sebesar 0,78. Sementara itu, Multinomial Naive Bayes dengan parameter smoothing alpha = 1,0 memperoleh accuracy sebesar 75,7% dan Macro F1-Score sebesar 0,75. Berdasarkan hasil tersebut, Support Vector Machine memberikan performa keseluruhan yang sedikit lebih tinggi dibandingkan Multinomial Naive Bayes pada dataset penelitian.

Perbedaan performa terlihat terutama pada kelas negatif dan netral. SVM memperoleh F1-score sebesar 0,63 pada kelas negatif dan 0,93 pada kelas netral, sedangkan Multinomial Naive Bayes memperoleh F1-score sebesar 0,58 pada kelas negatif dan 0,89 pada kelas netral. Pada kelas positif, kedua model menghasilkan F1-score yang sama, yaitu 0,79.

Visualisasi word cloud menunjukkan bahwa pembahasan masyarakat berpusat pada istilah yang berkaitan dengan program makan bergizi, pelaksanaan makan siang, sekolah, dan jadwal program. Sentimen positif mencerminkan

DOI: https://doi.org/10.69693/ijmst.v4i2.11800

dukungan dan harapan terhadap manfaat program, sentimen negatif menunjukkan kritik dan kekhawatiran terhadap pelaksanaan program, sedangkan sentimen netral lebih banyak berisi informasi mengenai kegiatan dan jadwal pelaksanaan.

Penelitian ini masih memiliki keterbatasan karena menggunakan dataset sekunder dari satu sumber dan belum secara khusus menangani sarkasme, ironi, serta variasi bahasa tidak baku pada media sosial. Penelitian selanjutnya dapat menggunakan dataset yang lebih besar, melibatkan lebih dari satu pelabel, serta membandingkan algoritma lain untuk memperoleh hasil klasifikasi yang lebih kuat.

Reference

Almas Hediawan, Reihan, Ahmad Faisol, And Suryo Adi Wibowo. 2026. "Analisis Sentimen Terhadap Penggunaan Sistem Paylater Pada Media Sosial X Menggunakan Metode Support Vector Machine." *JATI (Jurnal Mahasiswa Teknik Informatika)* 9(6): 10914–21. Doi:10.36040/Jati.V9i6.15439.

Ditami, Gientry Rachma, Eva Faja Ripanti, And Herry Sujaini. 2022. "Implementasi Support Vector Machine Untuk Analisis Sentimen Terhadap Pengaruh Program Promosi Event Belanja Pada Marketplace." *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)* 8(3): 508. Doi:10.26418/Jp.V8i3.56478.

Fauzi, Ahmad, And Heri Agus Yuniak. 2024. "JEPIN (Jurnal Edukasi Dan Penelitian Informatika) Analisis Sentimen US Airline Pada Media Sosial Twitter/X Menggunakan Perbandingan Algoritma Data Mining." *JEPIN (Jurnal Edukasi Dan Penelitian Informatika)* 10(2): 277–86. Https://Www.Kaggle.Com/Datasets/Crowdflower/Twitt.

Haerani, Febri. 2026. "Analisis Sentimen Program Makan Bergizi Gratis Berdasarkan Komentar Di Media Sosial X Menggunakan Gru." *Jurnal Informatika Dan Teknik Elektro Terapan* 14(1): 749–59. Doi:10.23960/Jitet.V14i1.8754.

Ma'rufudin, Ma'rufudin, And Aditia Yudhistira. 2025. "Analisis Sentimen Petani Milenial Pada Media Sosial X Menggunakan Algortitma Support Vector Machine (SVM)." *Jurnal Pendidikan Dan Teknologi Indonesia* 5(3): 845–57. Doi:10.52436/1.Jpti.717.

Magdalena, Evasaria, And Bhustomy Hakim. 2024. "Sentimen Kebijakan Pemerintah Terhadap Kepercayaan Masyarakat Menggunakan Metode NLP Rule Based Sentiment On Public Trust Using The NLP Rule Based Method." 12(1): 175–82. Doi:10.26418/Justin.V12i1.72426.

Maulana, Bagas Akbar, Muhammad Jazilul Fahmi, Ari Muhamad Imran, And Nutriana Hidayati. 2024. "Analisis Sentimen Terhadap Aplikasi Pluang Menggunakan Algoritma Naive Bayes Dan Support Vector Machine (SVM)." *MALCOM: Indonesian Journal Of Machine Learning And Computer Science* 4(2): 375–84. Doi:10.57152/Malcom.V4i2.1206.

Prasetyoningrum, Devi, And Pulung Nurtantio Andono. 2024. "Perbandingan Model Machine Learning Dalam Analisis Sentimen Pada Kasus Monkeypox Di Media Sosial X." *Building Of Informatics, Technology And Science (BITS)* 6(3): 2005–14. Doi:10.47065/Bits.V6i3.6447.

Purba, Uliano Wilyam, Andre Hadiman Rotua Parhusip, Sahid Maulana, Luluk Muthoharoh, Ardika Satria, And Martin C. T. Manullang. 2026. "Sentiment Analysis Of Indonesian Spotify Reviews Using Machine Learning And Bilstm." (1): 1–8. Http://Arxiv.Org/Abs/2605.03443.

Saputra, Juliandri, Lily Maryani, Rahmaddeni, Denok Wulandari, And Wisnu Eka. 2025. "Analisis Performa Naive Bayes Dan Svm Terhadap Sentimen Teks Media Sosial Dengan Word2Vec Dan Smote." *Jurnal INSTEK (Informatika Sains Dan Teknologi)* 10(1): 143–55. Doi:10.24252/Instek.V10i1.54889.

Sipayung, Evasaria Magdalena. 2024. "Sentiment On Public Trust Using The NLP Rule Based Method." *Jurnal Sistem Dan Teknologi Informasi (Justin)* 12(1): 175. Doi:10.26418/Justin.V12i1.72426.

Srg, M Rio Pratama. 2025. "Global Research And Innovation Journal (GREAT) ANALISIS SENTIMEN PADA MEDIA SOSIAL TWITTER TERHADAP PENEGAKAN HUKUM DI INDONESIA TAHUN 2024 MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM) Analisis Sentimen , Twitter , Penegakan Hukum , Support Vector ." 1: 43–51.

Suhendra, Suhendra, And Feny Selly Pratiwi. 2024. "Peran Komunikasi Digital Dalam Pembentukan Opini Publik: Studi Kasus Media Sosial." *Iapa Proceedings Conference*: 293. Doi:10.30589/Proceedings.2024.1059.

Syahira, Melani Alka, And Rakhmat Kurniawan. 2024. "Analisis Sentimen Cyberbullying Pada Media Sosial X Menggunakan Metode Support Vector Machine." *Jurnal Media Informatika Budidarma* 8(3): 1724. Doi:10.30865/Mib.V8i3.7926.

Wirawan, Alwan, Hasan Dwi Cahyono, And Winarno. 2023. "Easy Data Augmentation In Sentiment Analysis Of Cyberbullying." *2023 6th International Conference On Information And Communications Technology, ICOIACT 2023*: 443–47. Doi:10.1109/ICOIACT59844.2023.10455817.

DOI: https://doi.org/10.69693/ijmst.v4i2.11800