

Public Sentiment Analysis of the Free Meal Program: A Comparison of Naive Bayes and Support Vector Machine Methods on the Twitter (X) Social Media Platform

Muhammad Farhan Saleh ^{1*}, Rahmi Imanda ^{2*}

* Teknik Informatika, Universitas Muhammadiyah Prof. Dr. Hamka
farhansaleh135@gmail.com ¹, rahmi.imanda@uhamka.ac.id ²

Article Info

Article history:

Received 2024-11-14
Revised 2024-12-23
Accepted 2025-01-16

Keyword:

*Sentiment Analysis,
Naive Bayes,
SVM,
Free Meal Program,
Twitter.*

ABSTRACT

The problems of nutrition, including stunting, remain a challenge in Indonesia. Therefore, Prabowo and Gibran launched the 2024 Free Meal Program, which provides free lunch to every school child as well as pregnant mothers. This research analyzes public sentiment towards this program using data from X with Naive Bayes and Support Vector Machine (SVM) methods. The data was analyzed through crawling, preprocessing, labeling, and feature extraction using TF-IDF. The results showed a predominance of positive sentiment towards the program, with SVM performing better in sentiment classification, achieving 86.42% accuracy compared to Naive Bayes with 67.9%. The findings can guide policymakers in improving the communication strategy and implementation of the Free Meal Program to make it more impactful for Indonesians.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Pendahuluan Masalah gizi di Indonesia, khususnya di kalangan anak-anak, masih menjadi tantangan besar. Kondisi ini meliputi kekurangan gizi hingga gizi buruk yang berdampak serius pada pertumbuhan dan perkembangan anak. Beberapa faktor penyebabnya termasuk kurangnya asupan makanan bergizi, minimnya edukasi tentang nutrisi yang tepat, dan sulitnya akses ke layanan kesehatan serta gizi yang memadai di beberapa wilayah [1]. Gizi buruk dan kurang gizi adalah penyebab utama masalah kesehatan, termasuk stunting. Stunting, yaitu terhambatnya pertumbuhan tubuh akibat kekurangan gizi jangka panjang, menyebabkan anak-anak memiliki tinggi badan di bawah standar usia dan sering mengalami keterlambatan perkembangan kognitif [2].

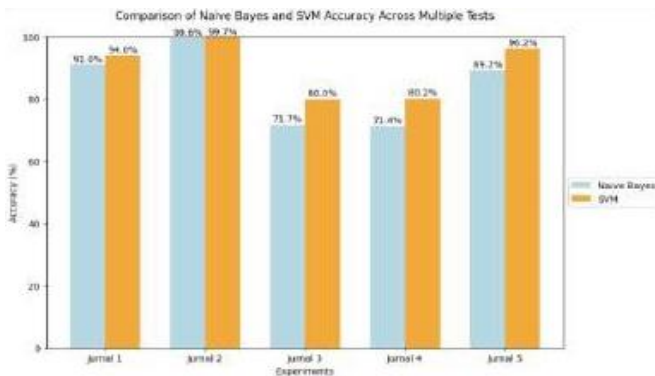
Stunting dapat meningkatnya risiko penyakit kronis di usia dewasa, termasuk diabetes dan penyakit jantung. Prevalensi stunting di Indonesia masih tinggi, disebabkan oleh kemiskinan, kurangnya akses ke makanan bergizi, dan rendahnya pengetahuan tentang pola makan sehat [3]. Dalam upaya mengatasi masalah ini, pasangan calon presiden (CAPRES) dan wakil presiden (WAPRES) Prabowo Subianto dan Gibran Rakabuming mengumumkan sebuah program kebijakan baru dalam kampanye mereka pada tahun 2024,

yaitu Program Makan Gratis [4]. Program ini bertujuan untuk menyediakan makan siang gratis kepada anak-anak sekolah dan ibu hamil di seluruh Indonesia [5]. Diharapkan, kebijakan ini tidak hanya membantu mengurangi angka stunting tetapi juga meningkatkan status gizi dan kesehatan masyarakat [6]. Program ini direncanakan untuk dilaksanakan secara bertahap dan melibatkan berbagai pihak, termasuk Usaha Mikro, Kecil, dan Menengah (UMKM) berfungsi untuk menyuplai makanan [7].

Untuk memahami persepsi publik terhadap program ini, analisis sentimen digunakan sebagai metode untuk memvisualisasikan reaksi masyarakat. Analisis sentimen adalah suatu proses untuk mengidentifikasi maupun mengategorisasi pandangan yang diekspresikan dalam suatu kelompok teks, khususnya untuk memastikan respon atau sentimen masyarakat terhadap suatu program, apakah positif atau negatif [8]. Metode proses ini biasanya diterapkan diberbagai aspek seperti pelayanan pelanggan dan penelitian sosial untuk dapat memahami pandangan dan respons khalayak mengenai produk atau layanan [9].

Terdapat beberapa metode yang banyak digunakan untuk menganalisis sentimen tersebut. Naive Bayes adalah metode yang sering digunakan dalam pengolahan bahasa alami

(Natural Language Processing/NLP) karena kesederhanaan dan efektivitasnya dalam mengklasifikasikan data teks [10]. Metode lain, Support Vector Machine (SVM) adalah sebuah proses klasifikasi yang menemukan hyperplane optimal guna membedakan data ke dalam kelas-kelas yang berbeda, dikenal karena kemampuannya menangani data berdimensi tinggi dan menghasilkan klasifikasi yang akurat [11]. Kedua metode ini banyak digunakan oleh para peneliti untuk menganalisa sentimen masyarakat. Namun dari berbagai jurnal yang sudah ditelusuri, kedua metode ini memberikan hasil yang berbeda.



Gambar 1. Perbandingan Naive Bayes dan SVM Referensi Jurnal

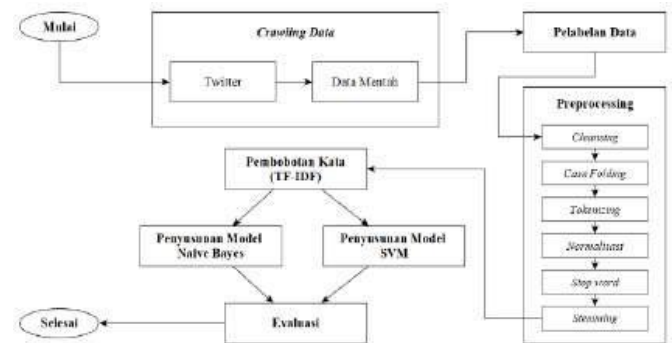
Perbedaan yang tergambar pada gambar 1 menunjukkan, metode SVM dapat memberikan keakuratan yang lebih akurat dibandingkan metode Naive Bayes dengan selisih paling tinggi mencapai 8,8%. Namun juga terdapat perbandingan yang memiliki selisih paling kecil yaitu 0,1%. Sehingga pada penelitian, dilakukan perbandingan untuk analisis sentimen. Tujuannya untuk mengetahui metode mana yang memiliki akurasi lebih tinggi dan berapa selisihnya. Platform Twitter dipilih sebagai sumber data utama untuk penelitian ini karena popularitasnya sebagai media sosial yang digunakan oleh berbagai kalangan masyarakat untuk mengekspresikan opini. Twitter memungkinkan pengguna untuk berbagi pandangan secara real-time dengan format pesan singkat yang memudahkan analisis. Dengan lebih dari jutaan pengguna di Indonesia, Twitter menjadi salah satu platform utama yang merefleksikan dinamika opini publik terhadap isu-isu terkini [12].

Penelitian ini berfokus untuk mengidentifikasi dan melakukan analisis terhadap sentimen masyarakat Indonesia terhadap kebijakan makan gratis yang diusulkan melalui perbandingan metode Naive Bayes dan SVM, diharapkan hasil analisis ini memberikan gambaran komprehensif tentang penerimaan publik, serta wawasan bagi penyusun keputusan dalam menyusun rancangan dan menyesuaikan kebijakan agar lebih efektif. Temuan tersebut juga dapat digunakan untuk merumuskan strategi komunikasi yang lebih baik dan rencana implementasi yang terperinci, memastikan program ini mencapai sasaran dan memberikan manfaat maksimal bagi masyarakat. Dengan pemahaman mendalam mengenai sentimen publik, kebijakan diharapkan dapat dioptimalkan

untuk mengatasi tantangan dan memaksimalkan dampak positif program tersebut.

II. METODE

Tahapan yang dilakukan dalam penelitian dapat dilihat pada Gambar 1. Gambar 1 menjelaskan alur atau serangkaian dari penelitian ini.



Gambar 2. Metode Penelitian

A. Crawling Data

Pada tahap ini, data mentah akan dikumpulkan dari platform Twitter. Data ini akan berisi postingan dan tweet yang relevan dengan topik penelitian. Data mentah ini akan menjadi bahan utama untuk analisis sentimen [13].

B. Pelabelan Data

Setelah mendapatkan data mentah, langkah berikutnya adalah memberikan label sentimen pada setiap tweet. Sentimen umumnya dapat berupa positif dan negatif berdasarkan konten teksnya. Proses pelabelan ini akan membantu dalam pemodelan dan analisis sentimen selanjutnya [14].

C. Preprocessing Data

Pada tahap ini, Data yang telah dilabeli akan menjalani tahap preprocessing. Proses ini melibatkan berbagai langkah [15], seperti:

- 1) *Cleansing*: Menghapus karakter atau elemen data yang tidak relevan atau mengganggu.
- 2) *Case Folding*: Merubah seluruh teks ke dalam huruf kecil atau huruf besar untuk memperoleh konsistensi.
- 3) *Tokenizing*: Memecah teks menjadi kata-kata atau token.
- 4) *Normalisasi*: Mengganti kata-kata dengan sinonim atau bentuk dasar mereka.
- 5) *Stopword*: Menghilangkan penggunaan kosakata umum tersebut yang bukan merupakan nilai sentimen, semisal “dan”, “atau”, “pada”, dsb.
- 6) *Stemming*: Mengubah kosakata ke dalam bentuk aslinya atau kata-kata akar untuk memperlancar analisa. data mentah akan dikumpulkan dari platform Twitter. Data ini akan berisi postingan dan tweet yang relevan dengan topik penelitian. Data mentah ini akan menjadi bahan utama untuk analisis sentimen.

D. Pembobotan Data

Pada proses selanjutnya, dilakukan pemberian bobot kata dengan cara Term Frequency-Inverse Document Frequency (TF-IDF) pada data teks. Metode ini akan menentukan nilai bobot untuk masing-masing kata pada teks berdasarkan intensitas kemunculan kata yang ada pada dokumen dan relevansi kata dalam konteks analisis sentimen [16].

E. Pemodelan Data

Setelah data sudah diproses dan dibobotkan, langkah berikutnya adalah membangun model klasifikasi sentimen. Peneliti memilih algoritma klasifikasi Naive Bayes. Model ini memanfaatkan bobot kata dari langkah awal untuk mengelompokkan kalimat sentimen dari cuitan tersebut tergolong positif atau negatif. Tidak hanya memanfaatkan Naive Bayes [17], penelitian ini akan memanfaatkan Support Vector Machine (SVM) untuk membandingkan hasil pengelompokkan sentimen.

SVM ialah salah satu metode pembelajaran mesin yang baik dalam memisahkan kelas yang berbeda dengan menemukan hyperplane optimal yang memaksimalkan margin antara kelas-kelas tersebut [18]. Pada penelitian ini, SVM akan digunakan untuk pengklasifikasian data tweet ke dalam kelompok sentimen positif atau negatif dan hasilnya akan dibandingkan dengan model Naive Bayes.

F. Evaluasi

Setelah model Naive Bayes dan SVM diterapkan, hasil uji sentimen akan diperoleh dari kedua model tersebut. Hasil ini akan memberikan gambaran tentang sentimen yang ada dalam data, termasuk sejauh mana sentimen positif atau negatif yang terdapat dalam data tersebut. Perbandingan antara hasil klasifikasi Naive Bayes dan SVM juga akan dilakukan guna meninjau capaian kinerja kedua model tersebut. Model Naive Bayes akan memberikan output berupa tingkat peluang terjadinya sentimen positif atau negatif berdasarkan bobot kata yang telah dibentuk dengan metode TF-IDF. Sementara itu, SVM akan menghasilkan klasifikasi sentimen berdasarkan hyperplane yang memisahkan kelas-kelas sentimen. Hasil uji ini akan dianalisis untuk menentukan model mana yang lebih akurat dalam mengklasifikasikan sentimen tweet terkait program makanan gratis. Hasil uji sentimen ini akan menjadi fokus analisis dan interpretasi untuk menjawab pertanyaan penelitian serta melihat bagaimana kedua model tersebut bekerja dalam konteks data yang digunakan.

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Dalam penelitian yang dilakukan menggunakan 1 dataset dengan 6 kata kunci pencarian berbeda yang berasal dari media sosial X. Setelah mengumpulkan 6 file dengan kata kunci pencarian yang berbeda lalu menjadikannya 1 file dataset untuk di olah. Periode waktu pencarian ini mencakup

rentang dari 1 Februari 2024 hingga 1 September 2024. Crawling data ini memanfaatkan library tweet harvest. Menjalankan perintah kode program pada tweet harvest untuk pengumpulan data dari media sosial X. Menjalankan command “search_keyword = 'kebijakan makan prabowo lang:id'” untuk mengambil data pencarian pertama lalu, “search_keyword = 'kebijakan makan gibran lang:id'” untuk pencarian kedua, “search_keyword = 'program makan prabowo lang:id'” untuk pencarian ketiga, “search_keyword = 'program makan gibran lang:id'” untuk pencarian keempat, “search_keyword = 'subsidi makan prabowo lang:id'” untuk pencarian kelima, dan pencarian keenam “search_keyword = 'subsidi makan gibran lang:id'”. Hasil pengumpulan data setelah menjadi satu sebanyak 401 data.

TABEL I
PENJELASAN COMMAND KEBIJAKAN MAKAN PRABOWO

No	Kata	Keterangan
1	Kebijakan makan prabowo	Tweet harus mengandung kata 'kebijakan makan prabowo'
2	Lang:id	Tweet harus berbahasa indonesia

TABEL II
PENJELASAN COMMAND KEBIJAKAN MAKAN GIBRAN

No	Kata	Keterangan
1	Kebijakan makan gibran	Tweet harus mengandung kata 'kebijakan makan gibran'
2	Lang:id	Tweet harus berbahasa Indonesia

TABEL III
PENJELASAN COMMAND PROGRAM MAKAN PRABOWO

No	Kata	Keterangan
1	Kebijakan makan gibran	Tweet harus mengandung kata 'program makan prabowo'
2	Lang:id	Tweet harus berbahasa indonesia

TABEL IV
PENJELASAN COMMAND PROGRAM MAKAN PRABOWO

No	Kata	Keterangan
1	Kebijakan makan gibran	Tweet harus mengandung kata 'program makan gibran'
2	Lang:id	Tweet harus berbahasa indonesia

TABEL V
PENJELASAN COMMAND SUBSIDI MAKAN PRABOWO

No	Kata	Keterangan
1	Kebijakan makan gibran	Tweet harus mengandung kata 'subsidi makan prabowo'
2	Lang:id	Tweet harus berbahasa indonesia

TABEL VI
PENJELASAN COMMAND SUBSIDI MAKAN PRABOWO

No	Kata	Keterangan
1	Kebijakan makan gibran	Tweet harus mengandung kata 'subsidi makan gibran'
2	Lang:id	Tweet harus berbahasa indonesia

B. Pelabelan Data

Labelin sentimen disini menggunakan Library Vader Lexicon dari NLTK. Library berguna untuk memberikan skor sentimen pada setiap kata, lalu mengakumulasiannya untuk tiap data, dan menentukan apakah hasil skornya menunjukkan sentimen positif atau negatif [19]. Perlu dicatat bahwa library ini hanya mendeteksi bahasa Inggris, sehingga sebelum proses pelabelan, data harus terlebih dahulu diterjemahkan ke dalam bahasa Inggris menggunakan library googletrans [20].

TABEL VII
PREPROCESSING

Text Indonesia	Text English	Sentimen
Menko Perekonomian Airlangga Hartarto simulasi makan siang gratis di SMPN 2 Curug Tangerang turun gunung memastikan kebijakan berjalan baik ketika dilaksanakan. Ini Hasil Dipilih Rakyat Indonesia kemenangan prabowo-gibran SIAPA SANGKA PENGEN BANGET	Coordinating Minister for the Airlangga Economy Hartarto free lunch simulation at SMPN 2 Curug Tangerang down the mountain ensures that the policy goes well when implemented. This is the result of the Indonesian people's victory Prabowo-gibran who thinks I really want	Positive
Yang lebih menyebalkan dari pendukung fanatik salah satu capres adalah orang ignorant yg bilang siapapun presidennya kira tetep cari makan sendiri dikira kebijakan presiden ga ngaruh ke kehidupan kamu	What is more annoying than the fanatical supporters of one of the presidential candidates are ignorant people who say whoever the president is, you still have to find your own food, thinking that the president's policies do not affect your life.	Negative

Sentimen positif yang mendominasi dalam analisis ini dapat dijelaskan melalui beberapa faktor. Secara teknis, pelabelan sentimen menggunakan library NLTK/VADER berbasis lexicon, di mana kata-kata yang sering muncul dalam dataset seperti "bagus," "membantu," atau "sukses" cenderung menghasilkan skor sentimen positif. Namun, dominasi sentimen positif juga dapat dikaitkan dengan beberapa faktor eksternal. Secara budaya, masyarakat Indonesia cenderung mendukung program pemerintah yang dirasa memberikan manfaat langsung, terutama jika berkaitan dengan pendidikan atau kesejahteraan anak.

Selain itu, keberhasilan program makan gratis ini dalam menyampaikan narasi yang positif melalui media sosial, seperti cerita sukses dan testimoni dari penerima manfaat, turut memperkuat persepsi positif publik. Kepercayaan terhadap tokoh atau institusi yang menginisiasi program ini juga menjadi faktor lain, terutama jika tokoh tersebut memiliki reputasi baik dalam melaksanakan kebijakan sosial.

Meskipun demikian, keterbatasan alat pelabelan dalam menangkap konteks penuh dari komentar sosial media harus diakui, sehingga analisis ini perlu dipandang sebagai representasi umum dari persepsi publik.

C. Preprocessing

Preprocessing merupakan tahapan penting dalam analisis data teks karena data mentah yang diperoleh dari platform media sosial umumnya masih mengandung banyak noise atau informasi yang tidak relevan [21]. Proses ini melakukan perubahan data mentah ke data yang lebih bersih sehingga digunakan untuk analisis lebih lanjut [22]. Langkah-langkah preprocessing dapat dilakukan sebagai berikut:

1) Cleansing

Pada tahap ini, teks tweet membersihkan berbagai elemen karakter yang tidak dibutuhkan seperti URL, simbol, angka, serta tanda baca koma titik yang tidak diperlukan [23]. Hal ini dilakukan untuk mengurangi gangguan pada proses analisis sentimen.

TABEL VIII
CLEANSING

Sebelum	Sesudah
Mengenai gagasan Paslon nomor 2 Prabowo - Gibran makan siang dan susu gratis pengamat sosial dan kebijakan publik https://t.co/aZM51W8151	Mengenai gagasan Paslon nomor Prabowo Gibran makan siang dan susu gratis pengamat sosial dan kebijakan publik

2) Case Folding

Setelah proses pembersihan, langkah berikutnya adalah case folding. Seluruh kalimat diubah menjadi huruf lowercase untuk memastikan bahwa kata yang sama dengan perbedaan kapitalisasi dapat dikenali sebagai kata yang sama [24].

TABEL IX
CASE FOLDING

Sebelum	Sesudah
Mengenai gagasan Paslon nomor Prabowo Gibran makan siang dan susu gratis pengamat sosial dan kebijakan publik	mengenai gagasan paslon nomor prabowo gibran makan siang dan susu gratis pengamat sosial dan kebijakan publik

3) Tokenizing

Proses tokenizing melakukan pemecahan kalimat menjadi kata yang terpisah atau token. Setiap kata dalam tweet dianggap sebagai satu token [25], yang nantinya akan diproses lebih lanjut.

TABEL X
TOKENIZING

Sebelum	Sesudah
mengenai gagasan paslon nomor prabowo gibran makan siang dan susu gratis pengamat sosial dan kebijakan publik	[mengenai, gagasan, paslon, nomor, prabowo, gibran, makan, siang, dan, susu, gratis, pengamat, sosial, dan, kebijakan, publik]

4) Normalisasi

Pada tahap ini, dilakukan perubahan terhadap kata-kata tidak baku atau singkatan yang sering digunakan dalam bahasa informal media sosial menjadi bentuk baku [26]. Kata-kata seperti “gk” menjadi “tidak”, “dgn” menjadi “dengan”, dan “sm” menjadi “sama” adalah beberapa contoh normalisasi yang diterapkan. Hal ini penting untuk meningkatkan akurasi analisis karena kata-kata baku dan tidak baku dianggap berbeda oleh model.

TABEL XI
NORMALISASI

Sebelum	Sesudah
mengenai gagasan paslon nomor prabowo gibran makan siang dan susu gratis pengamat sosial dan kebijakan publik	[mengenai, gagasan, paslon, nomor, prabowo, gibran, makan, siang, dan, susu, gratis, pengamat, sosial, dan, kebijakan, publik]

5) Stopword Removal

Stop words memiliki kosakata umum seperti “yang”, “di”, “ke”, dan “dan” yang biasanya tidak memiliki kontribusi yang penting terhadap analisis sentimen. Pada tahap ini, kata-kata tersebut dihapus dari teks untuk mengurangi kompleksitas dan meningkatkan relevansi fitur yang digunakan oleh model [27].

TABEL XII
STOPWORD REMOVAL

Sebelum	Sesudah
[mengenai, gagasan, paslon, nomor, prabowo, gibran, makan, siang, dan, susu, gratis, pengamat, sosial, dan, kebijakan, publik]	[gagasan, paslon, nomor, prabowo, gibran, makan, siang, susu, gratis, pengamat, sosial, kebijakan, publik]

6) Stemming

Stemming merupakan suatu proses pengubahan kata ke dalam bentuk asli atau dasarnya [28]. Misalnya, kata “makanan” diubah menjadi “makan”, “pembelajaran” diubah menjadi “ajar”, dan “menyukai” diubah menjadi “suka”. Proses ini penting untuk menyederhanakan variasi kata dan mengurangi redundansi dalam analisis.

TABEL XIII
STEMMING

Sebelum	Sesudah
[gagasan, paslon, nomor, prabowo, gibran, makan, siang, susu, gratis, pengamat, sosial, kebijakan, publik]	[gagas, paslon, nomor, prabowo, gibran, makan, siang, susu, gratis, amat, sosial, bijak, publik]

D. Pembobotan Kata

Tahap ini ditujukan dengan melihat nilai bobot dari tiap kata dengan permodelan TF-IDF. Pembentukan fitur dilakukan menggunakan TfidfVectorizer dari sklearn. Pada penelitian ini, digunakan parameter `max_features=5000`

untuk memastikan bahwa model hanya mempertimbangkan 5000 kata-kata terpenting. Selain itu, parameter `ngram_range = (1, 2)` digunakan untuk mempertimbangkan tidak hanya kata tunggal (unigram) tetapi juga kombinasi kata dua kata (bigram). Model TfidfVectorizer diinisialisasi dengan parameter yang sesuai. Data teks yang sudah dibersihkan kemudian dipasangkan dengan TfidfVectorizer untuk mengubah teks menjadi fitur numerik. Pada tahap ini, `X_tfidf` adalah matriks sparse yang berisi nilai TF-IDF untuk setiap dokumen (tweet) dan fitur (kata atau bigram) yang terpilih. Setelah pembentukan fitur TF-IDF selesai, matriks `X_tfidf` dipisahkan menjadi fitur (X) dan label (y). Label adalah hasil pelabelan sentimen pada tahapan sebelumnya.

TABEL XIV
TF-IDF

No	Ada	Airlangga	Wilayah
0	0.000000	0.212282	0.000000
1	0.000000	0.000000	0.000000
2	0.000000	0.000000	0.000000
3	0.000000	0.000000	0.000000

E. Pemodelan Naïve Bayes (NB) dan Support Vector Machine (SVM)

1) Naive Bayes (NB)

Setelah proses tersebut, dataset dibagi untuk keperluan pelatihan dan pengujian model. Pembagian dataset dilakukan dengan 80% Train dan 20% Test: Pada pembagian ini, 80% data digunakan untuk melatih model dan 20% sisanya untuk menguji performa model. Ini adalah rasio pembagian yang umum digunakan untuk memastikan bahwa model memiliki cukup data untuk belajar dan cukup data untuk diuji.

TABEL XV
PEMODELAN NAÏVE BAYES

	Precision	Recall	F1-score	Support
Negative	0.43	0.75	0.55	4
Neutral	0.17	0.67	0.28	6
Positive	0.94	0.68	0.79	71
Accuracy			0.68	81
Macro avg	0.51	0.70	0.54	81
Weighted avg	0.86	0.68	0.74	81
Accuracy: 67.90%				

Hasil evaluasi model Naïve Bayes menunjukkan kinerja yang bervariasi untuk masing-masing kelas sentimen (negative, neutral, dan positive). Pada sentimen negative, model berhasil mencapai precision sebesar 0.43, recall 0.75, dan f1-score 0.55, dengan jumlah dukungan (support) sebanyak 4 data. Untuk sentimen neutral, precision tercatat lebih rendah di angka 0.17, tetapi recall lebih tinggi yaitu 0.67, menghasilkan f1-score sebesar 0.28, dengan dukungan 6 data. Sentimen positive menunjukkan kinerja terbaik, dengan precision 0.94, recall 0.68, dan f1-score 0.79, dari total 71 data.

Secara keseluruhan, akurasi model berada di angka 67.90%, menunjukkan bahwa model mampu memprediksi dengan benar sekitar dua per tiga dari total data uji. Rata-rata makro (macro average) yang menghitung metrik secara seimbang untuk semua kelas menunjukkan precision 0.51, recall 0.70, dan f1-score 0.54. Rata-rata berbobot (weighted average), yang mempertimbangkan distribusi data di setiap kelas, menunjukkan precision 0.86, recall 0.68, dan f1-score 0.74.

Hasil ini menunjukkan bahwa model Naïve Bayes lebih baik dalam menangani data sentimen positive, tetapi kurang optimal dalam memprediksi sentimen neutral dan negative, terutama karena jumlah data pada kedua kelas tersebut lebih sedikit dibandingkan dengan kelas positive. Hal ini mengindikasikan bahwa distribusi data yang tidak seimbang (imbalanced dataset) dapat memengaruhi performa model, sehingga perlu dipertimbangkan metode tambahan seperti SMOTE atau penyesuaian threshold untuk meningkatkan performa pada kelas dengan data yang lebih sedikit.

2) Support Vector Machine

Setelah proses tersebut, dataset dibagi untuk keperluan pelatihan dan pengujian model. Pembagian dataset dilakukan dengan 80% Train dan 20% Test: Pada pembagian ini, 80% data digunakan untuk melatih model dan 20% sisanya untuk menguji performa model. Ini adalah rasio pembagian yang umum digunakan untuk memastikan bahwa model memiliki cukup data untuk belajar dan cukup data untuk diuji.

TABEL XVI
PEMOODELAN SVM

	Precision	Recall	F1-score	Support
Negative	1.00	0.25	0.40	4
Neutral	0.33	0.33	0.33	6
Positive	0.91	0.94	0.92	71
Accuracy			0.86	81
Macro avg	0.75	0.51	0.55	81
Weighted avg	0.87	0.86	0.85	81
Accuracy: 86,42%				

Hasil evaluasi model Support Vector Machine (SVM) menunjukkan performa yang kuat secara keseluruhan, terutama pada kelas sentimen positive, namun terdapat variasi dalam menangani kelas lain. Untuk sentimen negative, model mencapai precision 1.00, namun recall hanya sebesar 0.25, menghasilkan f1-score 0.40 dari total 4 data yang tersedia.

Pada sentimen neutral, precision dan recall sama-sama berada di angka 0.33, menghasilkan f1-score 0.33, dengan jumlah dukungan sebanyak 6 data. Sentimen positive menjadi kelas dengan kinerja terbaik, di mana precision tercatat 0.91, recall 0.94, dan f1-score 0.92, dari 71 data.

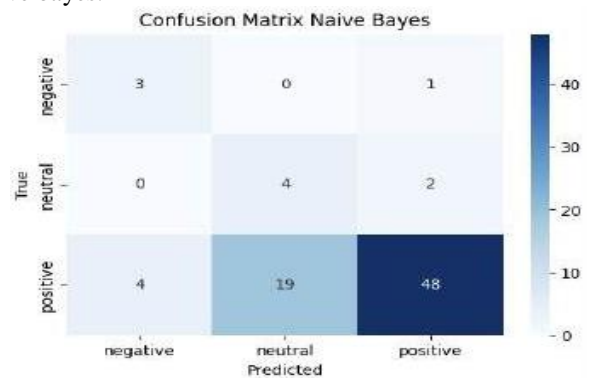
Secara keseluruhan, akurasi model mencapai 86.42%, menunjukkan bahwa model dapat memprediksi dengan benar sebagian besar data uji. Rata-rata makro (macro average), yang menghitung performa secara merata untuk semua kelas, mencatat precision 0.75, recall 0.51, dan f1-score 0.55.

Sementara itu, rata-rata berbobot (weighted average), yang mempertimbangkan distribusi data di tiap kelas, mencatat precision 0.87, recall 0.86, dan f1-score 0.85.

Hasil ini menunjukkan bahwa model SVM sangat efektif dalam memprediksi sentimen positive, tetapi cenderung mengalami kesulitan pada kelas dengan data lebih sedikit, seperti neutral dan negative. Ketidakseimbangan jumlah data pada masing-masing kelas dapat menjadi salah satu faktor yang memengaruhi performa model, terutama pada metrik recall untuk kelas negative. Namun, secara keseluruhan, SVM menunjukkan performa yang lebih konsisten dan akurat dibandingkan Naïve Bayes, dengan akurasi lebih tinggi dan f1-score yang lebih baik untuk sebagian besar kelas.

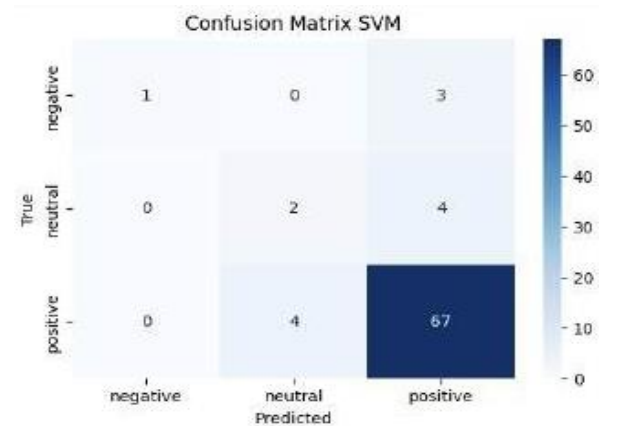
F. Evaluasi

Proses ini didapatkan dari confusion matrix pada algoritma naive bayes.



Gambar 3. Confusion Matrix Naïve Bayes

Naive Bayes tampaknya cukup kuat dalam mengidentifikasi sentimen positif, dengan 48 prediksi positif yang benar. Namun, model ini menunjukkan kesulitan dalam membedakan antara sentimen netral dan positif, karena 19 instance sentimen positif diklasifikasikan sebagai netral. Naive Bayes juga memiliki beberapa kesalahan klasifikasi untuk sentimen negatif, tetapi cenderung relatif baik dalam menjaga keseimbangan antara kelas. Sedangkan untuk confusion Support Vector Machine, menghasilkan beberapa perbedaan, berikut gambar 4.



Gambar 4. Confusion Matrix SVM

SVM mempunyai tingkat performa lebih baik untuk mengklasifikasikan sentimen positif dibandingkan Naive Bayes, dengan 67 prediksi positif yang benar dan sedikit kesalahan dalam klasifikasi sentimen positif ke netral atau negatif. Namun, SVM menunjukkan kelemahan dalam mengenali sentimen negatif dan netral, dengan beberapa instance negatif dan netral diklasifikasikan sebagai positif.

Identifikasi dan Analisis Persepsi Publik. Penelitian ini mengidentifikasi dan menganalisis persepsi masyarakat terhadap Program Makan Gratis yakni diusulkan Prabowo dan Gibran melalui Twitter. Mayoritas sentimen publik positif, menyambut baik inisiatif untuk kesejahteraan masyarakat, terutama di sektor pendidikan. Namun, ada juga sentimen negatif terkait skeptisisme dan kekhawatiran implementasi program. Sentimen netral muncul dari pengguna yang ragu akan kelayakan program. Walaupun dukungan luas, perhatian perlu diberikan pada kekhawatiran masyarakat terkait implementasi jangka panjang.

Perbandingan Performa Algoritma Naive Bayes dan SVM. Penelitian ini bertujuan untuk membandingkan dua algoritma dalam mengklasifikasikan sentimen publik terhadap program makan gratis, yaitu Naive Bayes dan Support Vector Machine (SVM). Berdasarkan hasil pengujian yang dilakukan, terdapat perbedaan performa yang signifikan antara kedua algoritma tersebut.

- 1) *Akurasi*: Algoritma SVM menunjukkan performa yang lebih unggul dibandingkan Naive Bayes dengan akurasi sebesar 86,42%. SVM mampu mendeteksi sentimen positif dengan sangat baik, yang terlihat pada nilai f1-score untuk sentimen positif sebesar 0.92. Meskipun demikian, SVM memiliki kelemahan dalam mendeteksi sentimen negatif, dengan nilai precision dan recall yang rendah pada kategori ini.
- 2) *Sentimen Negatif*: Naive Bayes lebih responsif terhadap sentimen negatif dengan recall sebesar 0.75, meskipun precision-nya hanya 0.43. Hal ini menunjukkan bahwa Naive Bayes cenderung lebih baik dalam menangkap sentimen negatif meskipun seringkali salah klasifikasi sentimen lainnya sebagai negatif.
- 3) *Sentimen Positif*: Pada kategori sentimen positif, SVM memiliki kinerja lebih baik, dengan f1-score yang tinggi (0.92) dibandingkan Naive Bayes yang memiliki f1-score 0.79. Hal ini menunjukkan bahwa SVM lebih efektif dalam mendeteksi sentimen positif di dalam dataset.
- 4) *Sentimen Netral*: Untuk sentimen netral, kedua algoritma menunjukkan performa yang kurang optimal, dengan precision dan recall yang relatif rendah. Meskipun Naive Bayes sedikit lebih baik dalam hal recall, SVM tetap menunjukkan kinerja lebih baik dalam hal precision.
- 5) *Pemilihan Algoritma*: Secara keseluruhan, pemilihan algoritma bergantung pada konteks analisis sentimen yang diinginkan. Jika fokus utama adalah pada akurasi secara umum dan kinerja pada sentimen positif, maka SVM lebih disarankan. Namun, jika analisis perlu lebih sensitif terhadap sentimen negatif, Naive Bayes dapat

menjadi pilihan yang lebih baik meskipun akurasinya lebih rendah.

Berikut adalah tabel yang menunjukkan perbandingan kinerja antara kedua algoritma berdasarkan beberapa metrik evaluasi.

TABEL XVII
PERBANDINGAN KINERJA PERFORMA

Metrik	SVM	Naive Bayes
Akurasi	86.42%	67.90%
Precision Negatif	1.00	0.43
Recall Negatif	0.25	0.75
Precision Netral	0.33	0.17
Recall Netral	0.33	0.67
Precision Positif	0.91	0.94
Recall Positif	0.94	0.68

Berikut adalah perbandingan pada algoritma Naive Bayes dan Support Vector Machine.



Gambar 5. Perbandingan Akurasi Naive Bayes

Hasil evaluasi memberikan perbedaan bahwa model Naive Bayes menghasilkan akurasi sebesar 67,90%, sedangkan model Support Vector Machine (SVM) memiliki tingkat akurasi yang lebih tinggi, yaitu 86,42%. Dari perbandingan ini, terlihat bahwa akurasi SVM lebih unggul dengan selisih 18,52% dibandingkan Naive Bayes. Perbedaan ini menunjukkan metode pada SVM mampu menghasilkan kinerja yang lebih unggul dalam mengklasifikasikan data pada dataset yang digunakan.

Berdasarkan analisis sentimen publik, mayoritas masyarakat memberikan respon positif terhadap program makan gratis, terutama dalam hal manfaat langsung yang dirasakan oleh siswa dan orang tua. Strategi komunikasi dapat difokuskan untuk memperkuat citra positif ini dengan membagikan lebih banyak cerita sukses dari sekolah yang sudah menjalankan program. Selain itu, penggunaan media sosial seperti Twitter dan Instagram dapat dioptimalkan untuk membagikan testimoni siswa dan orang tua yang menerima manfaat, disertai visualisasi seperti foto atau video kegiatan program.

Penelitian ini memiliki beberapa keterbatasan yang perlu diperhatikan. Pertama, data sosial media dari Twitter sering kali mengandung bias, baik karena profil pengguna yang tidak mewakili populasi umum maupun keberadaan noise seperti spam dan retweet tanpa konteks. Kedua, pelabelan sentimen otomatis menggunakan library NLTK dan VADER terbatas dalam menangkap konteks lokal dan bahasa informal, sehingga dapat memengaruhi akurasi pelabelan sentimen. Ketiga, distribusi data yang tidak seimbang menyebabkan model lebih optimal dalam memprediksi sentimen dominan (positive) tetapi kurang akurat untuk kelas negative dan neutral. Penelitian selanjutnya dapat mempertimbangkan pengumpulan data yang lebih representatif, penggunaan teknik manual atau semi-otomatis untuk pelabelan sentimen, serta penerapan metode penanganan data tidak seimbang seperti SMOTE.

V. KESIMPULAN

Penelitian ini mengidentifikasi persepsi publik terhadap Program Makan Gratis yang diusulkan oleh Prabowo dan Gibran melalui Twitter, dengan hasil mayoritas sentimen positif, meskipun terdapat kekhawatiran terkait implementasi program. Dalam analisis performa algoritma, Support Vector Machine (SVM) menunjukkan akurasi lebih tinggi (86,42%) dibandingkan Naïve Bayes (67,90%), dengan kinerja yang lebih baik dalam mendeteksi sentimen positif. Sebaliknya, Naïve Bayes lebih responsif terhadap sentimen negatif, meski memiliki tingkat kesalahan klasifikasi yang lebih tinggi. Sentimen netral sulit diklasifikasikan secara optimal oleh kedua algoritma. Penelitian ini menyarankan strategi komunikasi yang memanfaatkan media sosial untuk memperkuat citra positif program dengan berbagi testimoni dan visualisasi dari penerima manfaat. Namun, terdapat keterbatasan, termasuk bias data media sosial, keterbatasan akurasi pelabelan otomatis, dan distribusi data yang tidak seimbang. Untuk penelitian selanjutnya, disarankan pengumpulan data yang lebih representatif, metode pelabelan yang lebih akurat, dan penanganan data tidak seimbang seperti SMOTE guna meningkatkan kualitas prediksi.

DAFTAR PUSTAKA

- [1] O. Martony, "Stunting di Indonesia: Tantangan dan solusi di era modern," *Journal of Telenursing (JOTING)*, vol. 5, no. 2, pp. 1734–1745, 2023.
- [2] N. Nasriyah and S. Ediyono, "Dampak Kurangnya Nutrisi Pada Ibu Hamil Terhadap Risiko Stunting Pada Bayi Yang Dilahirkan," *Jurnal Ilmu Keperawatan dan Kebidanan*, vol. 14, no. 1, pp. 161–170, 2023, doi: 10.26751/jikk.v14i1.1627.
- [3] R. Hizriyani, "Pemberian asi eksklusif sebagai pencegahan stunting," *Jurnal Jendela Bunda Program Studi PG-PAUD Universitas Muhammadiyah Cirebon*, vol. 8, no. 2, pp. 55–62, 2021.
- [4] N. A. Rahmawati, S. A. Prasetyo, and M. W. Ramadhani, "Memetakan Visi Prabowo Gibran Pada Masa Kampanye Dalam Prespektif Pembangunan: (Analisis Wacana Kritis Visi Dan Misi Prabowo Gibran Dalam Prespektif Modernisasi)," *WISSEN: Jurnal Ilmu Sosial dan Humaniora*, vol. 2, no. 3, pp. 97–120, 2024.
- [5] R. Saputra and F. N. Hasan, "Analisis Sentimen Terhadap Program Makan Siang & Susu Gratis Menggunakan Algoritma Naive Bayes," *Jurnal Teknologi Dan Sistem Informasi Bisnis*, vol. 6, no. 3, pp. 411–419, 2024.
- [6] F. I. Septiani, N. Rosiana, and A. Azzahra, "Dampak Makan Siang Gratis Pada Kondisi Keuangan Negara Dan Peningkatan Mutu Pendidikan," *JUPENSAL: Jurnal Pendidikan Universal*, vol. 1, no. 2, pp. 191–196, 2024.
- [7] P. A. Maharani, A. R. Namira, and T. V. Chairunnisa, "Peran Makan Siang Gratis Dalam Janji Kampanye Prabowo Gibran Dan Realisasinya," *Journal Of Law And Social Society*, vol. 1, no. 1, pp. 1–10, 2024.
- [8] M. R. Alwi, S. Subandi, and M. R. Hariyadi, "Analisis Sentimen Data Twitter Terhadap Pelaksanaan Pembelajaran Online Di Indonesia Pada Masa Pandemi Covid-19 Menggunakan Metode Natural Language Processing," *Jurnal Riset Sistem Informasi dan Teknologi Informasi (JURSISTEKNI)*, vol. 5, no. 3, pp. 319–331, 2023.
- [9] F. Panjaitan, "Perbandingan Penggunaan Tf-idfvectorizer, Countvectorizer, Dan Hashingvectorizer Dengan Optimalisasi Parameter Pada Machine Learning Untuk Analisis Sentimen Pemilu 2024," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 4, pp. 7413–7419, 2024.
- [10] N. Nurwanda, N. Suarna, and W. Prihartono, "Penerapan Nlp (Natural Language Processing) Dalam Analisis Sentimen Pengguna Telegram Di Playstore," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 2, pp. 1841–1846, 2024.
- [11] T. Wuriyanto, H. B. Setiawan, and A. B. Tjandrarini, "Penerapan Model CRISP-DM pada Prediksi Nasabah Kredit yang Berisiko Menggunakan Algoritma Support Vector Machine," *Jurnal Ilmiah Scroll (Jendela Teknologi Informasi)*, vol. 10, no. 1, pp. 1–6, 2022.
- [12] N. B. Tsaniyah, N. A. R. Putri, H. N. Kafidani, and D. Amalia, "Interaksi Sosial Akun Twitter PT. Angkasa Pura II (Studi Analisis Isi)," *Al-DYAS*, vol. 3, no. 1, pp. 87–102, 2024.
- [13] A. Gaizka, A. R. Dzikrillah, and E. Sinduningrum, "Analisis Sentimen Masyarakat Sebelum Dan Sesudah Terpilihnya Gibran Sebagai Cawapres Prabowo Menggunakan Naïve Bayes," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 6, pp. 2830–2841, 2024.
- [14] M. F. Latif, "Visualisasi Analisis Sentimen Ulasan Pengunjung Mengenai Objek Wisata Di Sulawesi Selatan (Studi Kasus Kabupaten Gowa)," 2023, Politeknik Negeri ujung Pandang.
- [15] I. Budianto and S. N. Anwar, "Analisis Sentiment Pengguna Twitter Mengenai Program Vaksinasi Covid-19 Menggunakan Algoritma Naïve Bayes," *(JurTI) Jurnal Teknologi Informasi*, vol. 6, no. 1, pp. 37–43, 2022.
- [16] R. Al Rasyid and D. H. U. Ningsih, "Penerapan Algoritma TF-IDF dan Cosine Similarity untuk Query Pencarian Pada Dataset Destinasi Wisata," *Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi)*, vol. 8, no. 1, pp. 170–178, 2024.
- [17] M. Rinestu and B. Marsanto, "Klasifikasi Keputusan Investasi Di Masa Pandemi Covid-19 Dengan Menggunakan Naive Bayes," *Management Studies and Entrepreneurship Journal (MSEJ)*, vol. 3, no. 3, pp. 1784–1796, 2022.
- [18] A. S. Rahayu, A. Fauzi, and R. Rahmat, "Komparasi Algoritma Naïve Bayes Dan Support Vector Machine (SVM) Pada Analisis Sentimen Spotify," *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 2, pp. 349–354, 2022.
- [19] S. Ernawati and R. Wati, "Evaluasi Performa Kernel SVM dalam Analisis Sentimen Review Aplikasi ChatGPT Menggunakan Hyperparameter dan VADER Lexicon," *Jurnal Buana Informatika*, vol. 15, no. 01, pp. 40–49, 2024.
- [20] S. N. Qofifa, R. Ardiansyah, Y. Y. Joefrie, and N. T. Lapatta, "Artikel Analisis Sentimen terhadap Resolusi Genjatan Senjata PBB 2023: Studi pada 10 Negara Penolak Resolusi Konflik Israel-Palestina," *The Indonesian Journal of Computer Science*, vol. 13, no. 5, 2024.
- [21] L. A. Fudholi, N. Rahaningsih, and R. D. Dana, "Sentimen Analisis Perilaku Penggemar Coldplay Di Media Sosial Twitter Menggunakan Metode Naïve Bayes," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 3, pp. 4150–4159, 2024.
- [22] R. Rahmadani, A. Rahim, and R. Rudiman, "Analisis Sentimen Ulasan 'Ojol The Game' Di Google Play Store Menggunakan Algoritma Naive Bayes Dan Model Ekstraksi Fitur Tf-Idf Untuk Meningkatkan

- Kualitas Game,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, 2024.
- [23] S. Djusar and S. F. Abdillah, “Perbandingan Analisis Sentimen Masyarakat Terhadap Calon Presiden Negara Indonesia 2024 Berdasarkan Opini di Twitter Menggunakan Metode Rule Based,” in *SEMASTER: Seminar Nasional Teknologi Informasi & Ilmu Komputer*, 2024, pp. 364–375.
- [24] M. S. Mahendrasah and T. Hariguna, “Analisis Sentimen Pengguna Aplikasi Bukalapak di Platform Playstore Menggunakan Metode Naïve Bayes,” *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 2, pp. 733–745, 2024.
- [25] W. Wartumi, R. Kurniawan, and Y. A. Wijaya, “Analisis Data Sentimen Ulasan Pengguna Aplikasi Shopee di Google Play Store dengan Klasifikasi Algoritma Naïve Bayes,” *Jurnal Informatika dan Rekayasa Perangkat Lunak*, vol. 6, no. 1, pp. 164–170, 2024.
- [26] S. Sahilla, F. Amalia, and K. Mariskhana, “Klasifikasi Sentimen Pengguna Terhadap Akun Twitter Official Dana Dengan Menggunakan Algoritma Naïve Bayes Classifier,” *Jurnal Riset Sistem Informasi dan Teknologi Informasi (JURSISTEKNI)*, vol. 6, no. 3, pp. 580–591, 2024.
- [27] F. Baehaqi and N. Cahyono, “Analisis Sentimen Terhadap Cyberbullying Pada Komentar Di Instagram Menggunakan Algoritma Naïve Bayes,” *Indonesian Journal of Computer Science*, vol. 13, no. 1, 2024.
- [28] V. A. Sulistiani and M. Hamka, “Analisis Sentimen Pengguna Media Sosial Terhadap Identitas Kependudukan Digital Menggunakan Metode Support Vector Machine (SVM),” *Journal of Information System Research (JOSH)*, vol. 5, no. 4, pp. 1323–1332, 2024.
- [29] D. Vonega, A. Fadila, and D. Kurniawan, “Analisis Sentimen Twitter Terhadap Opini Publik Atas Isu Pencalonan Puan Maharani dalam PILPRES 2024,” *JAIC*, vol. 6, no. 2, pp. 129–135, Nov. 2022.