

Analisis Sentimen Tanggapan Publik di Twitter Terkait Program Kerja Makan Siang Gratis Prabowo–Gibran Menggunakan Algoritma Naïve Bayes Classifier dan Support Vector Machine

Annisa Ramadhani*, Inggih Permana, M Afdal, Mona Fronita

Fakultas Sains Dan Teknologi, Sistem Informasi, Universitas Islam Negeri Sultan Syarif Kasim, Riau, Indonesia

Email: ^{1,*}11950324534@student.uin-suska.ac.id, ²inggihermana@uin-suska.ac.id, ³m.afdal@uin-suska.ac.id,

⁴monafronita@uin-suska.ac.id.

Email Penulis Korespondensi: 11950324534@student.uin-suska.ac.id

Submitted: 02/11/2024; Accepted: 01/12/2024; Published: 03/12/2024

Abstrak—Indonesia menghadapi tantangan serius terkait stunting, dengan angka mencapai 21% pada tahun 2024, meskipun terjadi penurunan dari 24% pada tahun 2021. Sebagai langkah strategis, pemerintah meluncurkan program-program untuk mengatasi masalah ini, termasuk edukasi gizi, pemeriksaan kesehatan ibu hamil, dan pemberian makanan tambahan. Di tengah upaya tersebut, program makan siang gratis yang diusulkan diharapkan dapat meningkatkan kualitas gizi bagi anak-anak dan ibu hamil. Namun, program ini menuai kontroversi terkait anggaran yang dibutuhkan, diperkirakan mencapai Rp450 triliun, yang dapat mempengaruhi keseimbangan anggaran negara dan menimbulkan inflasi. Penelitian ini menganalisis sentimen masyarakat terhadap program makan siang gratis menggunakan Algoritma Naïve Bayes Classifier (NBC) dan Support Vector Machine (SVM). Dari analisis terhadap 1.028 tweet, ditemukan bahwa sentimen negatif mendominasi dengan 44,84%, diikuti oleh sentimen positif (32,39%) dan netral (22,76%). SVM menunjukkan performa lebih baik dengan akurasi 75,39% dibandingkan NBC yang hanya 68,97%. Hasil penelitian ini memberikan wawasan penting tentang persepsi masyarakat terhadap program tersebut dan menyoroti kebutuhan akan penelitian lebih lanjut untuk memperbaiki metodologi analisis sentimen.

Kata Kunci: Stunting; Makan Siang Gratis; Analisis Sentimen; Naïve Bayes Classifier; Support Vector Machine

Abstract—Indonesia faces a serious challenge related to stunting, with rates reaching 21% in 2024, although this represents a decrease from 24% in 2021. In response, the government has launched various programs to address this issue, including nutrition education, health check-ups for pregnant women, and supplementary food provisions. Amid these efforts, the proposed free lunch program aims to improve nutritional quality for children and pregnant women. However, this program has sparked controversy over the required budget, estimated at IDR 450 trillion, which could impact the national budget balance and lead to inflation. This study analyzes public sentiment toward the free lunch program using the Naïve Bayes Classifier (NBC) and Support Vector Machine (SVM) algorithms. An analysis of 1,028 tweets revealed that negative sentiment predominates at 44.84%, followed by positive sentiment (32.39%) and neutral sentiment (22.76%). SVM outperformed NBC with an accuracy of 75.39%, compared to NBC's 68.97%. The findings provide important insights into public perceptions of the program and highlight the need for further research to improve sentiment analysis methodologies.

Keywords: Stunting; Free Lunch Program; Sentiment Analysis; Naïve Bayes Classifier; Support Vector Machine

1. PENDAHULUAN

Indonesia memiliki tingkat stunting yang tinggi, khususnya di kalangan anak balita. Stunting menjadi masalah serius karena dapat memengaruhi pertumbuhan fisik dan perkembangan kognitif anak. Upaya penanganan stunting perlu melibatkan berbagai pihak, termasuk pemerintah, masyarakat, dan sektor kesehatan, untuk memastikan setiap anak mendapatkan nutrisi yang cukup dan berkualitas. Berdasarkan data Kementerian Kesehatan Republik Indonesia tahun 2024, angka stunting pada tahun 2021 sekitar 24%, namun mengalami penurunan sebesar 3% menjadi 21% pada tahun 2024 [1]. Meskipun terjadi penurunan, Indonesia tetap menjadi negara dengan tingkat stunting tertinggi di dunia, terutama pada anak-anak usia balita. WHO menetapkan standar target minimal stunting di bawah 20%, sehingga pemerintah merencanakan program untuk mengurangi stunting pada anak-anak balita berdasarkan permasalahan ini [2].

Sebelum diberlakukannya program makan siang gratis, pemerintah Indonesia telah meluncurkan tiga program utama untuk mengatasi stunting pada balita dan anak-anak, menurut Kementerian Kesehatan. Program-program tersebut meliputi: Generasi Sehat Tanpa Anemia dan Stunting adalah program yang bertujuan mendidik remaja mengenai pentingnya konsumsi gizi seimbang untuk mencegah anemia dan stunting. Melalui program ini, diharapkan masyarakat lebih memahami pentingnya asupan nutrisi yang seimbang dalam mencegah stunting dan anemia. Pemeriksaan USG di Semua Puskesmas untuk Ibu Hamil: Program ini menyediakan pemeriksaan ultrasonografi (USG) kepada ibu hamil di seluruh puskesmas. Pemeriksaan ini tidak hanya berfungsi untuk memantau perkembangan janin, tetapi juga untuk mengukur status gizi dan kadar zat besi ibu hamil. Asupan gizi yang memadai dan cukup zat besi sangat penting untuk mencegah stunting pada bayi yang sedang dikandung. Pemberian Makanan Tambahan (PMT) bertujuan untuk meningkatkan status gizi anak-anak dan masyarakat secara keseluruhan. Program ini melibatkan penyediaan makanan tambahan bergizi untuk balita dan anak-anak yang membutuhkan, guna memastikan mereka mendapatkan asupan gizi yang cukup untuk pertumbuhan dan perkembangan yang optimal [3].

Program makan siang gratis yang diusulkan oleh presiden dan wakil presiden bertujuan untuk menyediakan makanan dan susu gratis di sekolah serta pesantren, sekaligus memberikan bantuan gizi bagi balita dan ibu hamil.



Inisiatif ini mencakup pemberian makan siang setiap hari kepada siswa dari berbagai jenjang pendidikan, dengan tujuan meningkatkan asupan gizi, partisipasi sekolah, dan pengembangan karakter siswa. Selain itu, program ini juga memberikan dukungan ekonomi kepada keluarga yang kurang mampu. Meskipun ada pendapat bahwa usaha untuk mengatasi stunting melalui program ini dianggap terlambat, pemberian susu gratis kepada ibu hamil dianggap lebih penting. Pengalaman dari berbagai negara menunjukkan bahwa program makan siang gratis dapat meningkatkan gizi anak, kehadiran di sekolah, dan memperkuat karakter siswa [4].

Namun, program ini telah menimbulkan banyak kontroversi, terutama terkait dengan anggaran yang dibutuhkan. Biaya yang diperlukan untuk menjalankan program makan siang gratis ini cukup besar, diperkirakan sekitar Rp450 triliun, seperti yang disampaikan dalam sebuah wawancara [5]. Oleh karena itu, ada sejumlah pihak yang tidak setuju dengan keberadaan program ini. Menurut penelitian, program ini akan menghabiskan sebagian besar anggaran, bahkan mungkin mencapai setengah dari anggaran bantuan sosial (bansos) di era pemerintahan Jokowi pada tahun 2024. Hal yang perlu diperhatikan adalah jika dana tambahan untuk program makan siang gratis tidak tersedia, hal ini dapat menyebabkan defisit anggaran. Selain itu, ada kekhawatiran tentang potensi inflasi dan peningkatan utang negara akibat alokasi anggaran yang besar untuk program ini [6].

Meskipun demikian, tujuan utama dari program ini adalah untuk mengurangi ketimpangan ekonomi dan mendorong pemerataan ekonomi. Program ini berfokus pada peningkatan kualitas gizi anak-anak dan mendukung pertumbuhan perekonomian nasional. Diharapkan program ini dapat menurunkan angka gizi buruk dan stunting pada balita dan anak-anak di Indonesia, sehingga berkontribusi pada peningkatan kualitas sumber daya manusia menuju generasi emas pada tahun 2045 [7].

Dengan demikian, tujuan utama dari penelitian ini adalah untuk memberikan informasi yang berguna bagi masyarakat, pemangku kepentingan, serta pembuat kebijakan terkait pemahaman tentang sikap masyarakat terhadap kebijakan Makan Siang Gratis. Hasil penelitian ini diharapkan dapat menjadi rujukan untuk merencanakan dan mengevaluasi kebijakan lebih lanjut, serta untuk mengidentifikasi tantangan dan peluang dalam pelaksanaan program tersebut, dengan harapan program ini dapat memberikan dampak sosial dan ekonomi yang positif di Indonesia [8].

Penelitian ini memberikan kontribusi yang signifikan bagi para pembuat kebijakan, masyarakat, dan pemangku kepentingan dengan menyediakan informasi yang berguna dalam mengevaluasi sentimen publik terhadap program Makan Siang Gratis. Selain itu, penelitian ini dapat menjadi dasar yang kuat dalam merancang kebijakan yang lebih tepat dan responsif terhadap kebutuhan masyarakat, serta memberikan wawasan tentang tantangan atau potensi keberhasilan yang mungkin terjadi. Dengan demikian, hasil dari penelitian ini dapat berperan penting dalam membantu pemerintah untuk mengimplementasikan program yang efektif dan berdampak positif, serta memberikan pemahaman yang lebih baik kepada masyarakat mengenai dampak sosial dan ekonomi dari kebijakan tersebut [9].

X (dulu dikenal sebagai Twitter) telah menjadi platform media sosial yang sangat populer untuk mengungkapkan pandangan tentang topik-topik yang sedang hangat dibicarakan, baik berupa opini positif maupun negatif melalui komentar dan kutipan. X memungkinkan penggunaannya Untuk menjalin interaksi dengan orang lain dengan mudah dan cepat. Berbeda dari platform media sosial lainnya, pengguna X tetap dapat terhubung dan berinteraksi secara bebas tanpa harus berteman [10]. Dengan banyaknya orang yang berbagi opini di X, media sosial ini dapat dimanfaatkan untuk melakukan analisis sentimen terhadap isu-isu yang sedang tren, sehingga menjadi sumber informasi mengenai sentimen masyarakat tentang isu tertentu [11].

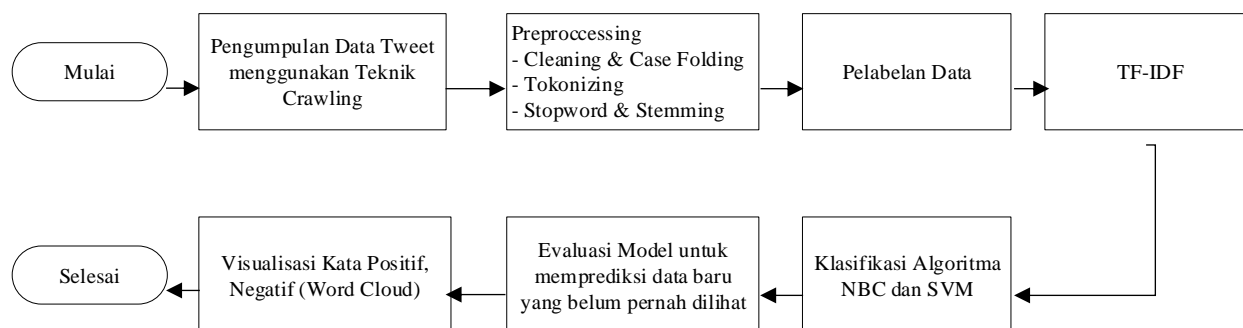
Penelitian ini menerapkan Algoritma Naïve Bayes Classifier (NBC) dan Support Vector Machine (SVM) untuk analisis sentimen. NBC merupakan salah satu algoritma klasifikasi yang umum diterapkan dalam machine learning untuk mengidentifikasi probabilitas kelas dari data tertentu. [12] NBC dikenal karena kesederhanaan dan kemudahan implementasinya, serta telah terbukti efektif dalam berbagai aplikasi, termasuk analisis sentimen dan pengenalan teks [13]. Sementara itu, SVM bekerja dengan membangun model yang Mengelompokkan dua kelas data dengan mencari hyperplane terbaik yang memiliki jarak terdekat dengan data dari masing-masing kelas [14]. SVM unggul dalam menciptakan batas keputusan yang optimal, sehingga mampu mengklasifikasikan data yang belum dikenal dengan akurasi tinggi [15].

Beberapa penelitian sebelumnya telah melakukan analisis sentimen menggunakan Algoritma NBC dan SVM. Perbandingan algoritma SVM dan NBC dalam Analisa sentimen pilkada pada twitter hasil perbandingan tersebut algoritma NBC memiliki akurasi 81,7%, recall 81,7% dan presisi 80%. Sedangkan algoritma SVM memiliki akurasi 80,7%, recall 80,7% dan presisi 84%. Maka dapat disimpulkan bahwa algoritma NBC lebih unggul dalam akurasi dan recall sedangkan dalam presisi yang lebih unggul adalah algoritma SVM [16]. Analisis sentimen pulik terhadap pengungsi rohingnya di Indonesia dengan metode *Support Vector Machine* dan *Naïve Bayes Classifier* hasil penelitian ini menunjukkan bahwa algoritma SVM memiliki akurasi sebesar 76%, sementara algoritma NBC memiliki akurasi sebesar 70%. Hal ini menunjukkan bahwa algoritma SVM mampu memprediksi sentimen dengan lebih baik dan memiliki angka kesalahan yang lebih rendah dibandingkan algoritma NBC [17]. Analisis sentiment pada ulasan aplikasi kredivo dengan algoritma SVM dan NBC hasil pengujian menggunakan algoritma SVM menghasilkan akurasi 83,3% dengan nilai presisi kelas positif 77% dan kelas negatif 87% sedangkan nilai recall untuk kelas positif sebesar 89% dan 73% untuk kelas negatif. Kemudian untuk algoritma Naive Bayes Classifier sendiri menghasilkan nilai akurasi sebesar 80,8% dengan nilai presisi untuk kelas positif sebesar 81% dan untuk kelas negatif sebesar 87%, sedangkan nilai recall untuk kelas positif sebesar 88% dan untuk kelas negatif sebesar 79%. Jadi untuk tingkat keseluruhan dapat dilihat dari nilai akursi dengan algoritma SVM lebih tinggi dibanding Naive Bayes Classifier [18]. Analisa sentimen aplikasi peduli lindungi dengan metode NBC dan SVM hasil penujian menggunakan algoritma NBC

menghasilkan akurasi 76%, recall 82%, presisi 76% dan F1-Score 79%. Sedangkan menggunakan algoritma SVM menghasilkan akurasi 80%, presisi 83%, recall 80% dan F1-Score 81% [19].

2. METODOLOGI PENELITIAN

Penelitian ini melakukan analisis sentimen pada program kerja makan siang gratis Prabowo – Gibran dengan menggunakan algoritma Naïve Bayes Classifier dan Support Vector Machine. Tahapan penelitian ini disajikan dalam bentuk diagram alir agar dapat berjalan sesuai tujuan yang akan dicapai. Tahapan penelitian dapat dilihat pada gambar 1.



Gambar 1. Tahapan Penelitian

Dalam penelitian ini dilakukan analisis sentiment pada program kerja makan siang gratis Prabowo – Gibran pada media sosial X. Adapun data yang diambil berupa data tweet dengan Teknik crawling data untuk mendapatkan data tweet selama 3 bulan terakhir (Januari – Maret 2024). Kemudian data yang telah dikumpulkan tersebut selanjutnya masuk ke tahap pengolahan data yakni *text processing* mulai dari *case folding* untuk mengubah teks menjadi bentuk standard yakni menjadi huruf kecil, *tokenizing* untuk pemotongan kalimat menjadi sebuah kata, *stopword* untuk menghilangkan kata-kata umum yang sering muncul seperti “yang”, “dan”, “di”, lalu *stemming* yakni untuk menyelaraskan kata untuk mereduksi daftar kata dalam data latih. Kemudian dilakukan ekstraksi fitur menggunakan TF-IDF dan setelahnya dilakukan pelabelan data yakni komentar positif, komentar negatif dan komentar netral.

Tahap selanjutnya yaitu klasifikasi menggunakan algoritma Naïve Bayes Classifier dan Support Vector Machine. Kemudian akan diperoleh akurasi masing-masing algoritma dan selanjutnya masuk ke tahap evaluasi model untuk memprediksi data baru yang belum pernah dilihat. Kemudian masuk ke tahap visualisasi kata menggunakan word cloud.

2.1 Analisis Sentimen

Analisis sentiment adalah proses mengklasifikasi dokumen teks yang berupa opini berdasarkan sentimen untuk menentukan apakah suatu tanggapan bersifat positif, negatif, atau netral [20]. Tujuan dari analisis sentimen untuk menentukan opini dari seseorang dengan memperhatikan topik tertentu dan menentukan ekstraksi atribut pada suatu dokumen atau teks yang berisi komentar untuk mengetahui ekspresi yang ada di dalamnya sehingga komentar tersebut dapat dikategorikan positif atau negatif [21]. Langkah yang dilakukan dalam analisis sentimen adalah mendefinisikan domain dataset, preprocessing, featured selection, pelabelan, klasifikasi dan evaluasi [22].

2.2 Algoritma Naïve Bayes Classifier (NBC)

Naïve Bayes Classifier merupakan machine learning yang menggunakan perhitungan probabilitas yang menggunakan konsep pendekatan Bayes. Penggunaan teorema Bayes pada algoritma Naïve Bayes adalah dengan menggabungkan prior probability dan conditional probability dalam suatu rumus yang dapat digunakan untuk menghitung probabilitas dari setiap kemungkinan klasifikasi [23]. Rumus dari teorema Bayes adalah :

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \tag{1}$$

Dalam teori probabilitas, khususnya dalam konteks pengklasifikasian data, kita sering kali berurusan dengan konsep-konsep yang terkait dengan Teorema Bayes. Misalnya, X merujuk pada data yang belum diketahui kelasnya atau kategori apa yang dimilikinya. Kita ingin mengetahui kelas atau kategori dari data tersebut, yang dalam hal ini diwakili oleh H, yang merupakan hipotesis mengenai kelas yang mungkin dimiliki oleh data X. Teorema Bayes menghubungkan probabilitas dari hipotesis (kelas) dengan data yang diamati. Probabilitas bahwa data X berasal dari kelas H, setelah mengamati X, disebut sebagai posteriori probabilitas dan dilambangkan sebagai $P(H|X)$. Ini adalah probabilitas yang ingin kita hitung dalam pengklasifikasian. Untuk menghitungnya, kita membutuhkan beberapa informasi lain, seperti prior probabilitas $P(H)$, yang menggambarkan seberapa besar keyakinan kita terhadap kelas H sebelum kita melihat data X, dan likelihood $P(X|H)$, yaitu probabilitas bahwa data X akan teramati jika kelas H benar.

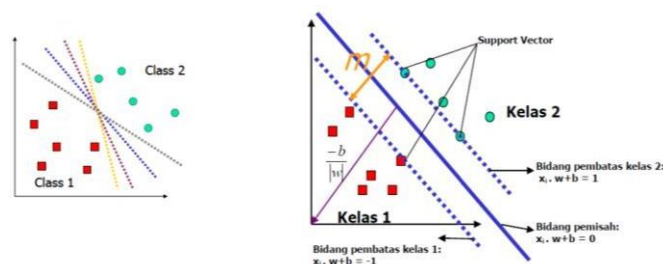
Selain itu, terdapat probabilitas total $P(X)$, yang merupakan probabilitas umum terjadinya data X , yang berfungsi sebagai faktor normalisasi untuk memastikan bahwa total probabilitas posteriori adalah 1.

Dengan menggunakan rumus Teorema Bayes, kita dapat menghitung probabilitas posteriori $P(H/X)$ sebagai rasio antara $P(X/H) \cdot P(H)$ dan $P(X)$. Proses ini memungkinkan kita untuk memilih kelas yang paling mungkin dimiliki oleh data X dengan mengidentifikasi kelas H yang memiliki probabilitas posterior tertinggi. Dalam praktiknya, pendekatan ini banyak digunakan dalam algoritma klasifikasi seperti *Naive Bayes Classifier*, yang mengasumsikan bahwa fitur-fitur dalam data bersifat independen dan menghitung probabilitas untuk setiap kelas berdasarkan informasi yang diberikan. Teorema Bayes, dengan demikian, memberikan dasar yang kuat untuk pendekatan probabilistik dalam pengklasifikasian dan membuat keputusan berdasarkan data yang terbatas.

2.3 Algoritma Support Vector Machine (SVM)

Untuk masalah *klasifikasi dan prediksi*, *Support Vector Machine* adalah pembelajaran mesin supervised yang sangat populer. SVM (algoritma pembelajaran mesin yang diawasi) menggunakan pendekatan berbasis kernel atau penerjemah untuk menemukan hyperplane atau bidang pemisah yang optimal anatar berbagai kelas dalam dataset. Tujuan utama SVM adalah untuk memaksimalkan jarak atau jarak antara hyperplane dengan data terdekat dari setiap kelas, sehingga diperoleh pemisahan yang jelas dan baik antara kelas-kelas tersebut (Isnain et al., 2021) [24]. Untuk menghubungkan data dari ruang input awal ke ruang fitur yang lebih besar, kernel digunakan. Metode optimasi digunakan untuk melatih algoritma SVM untuk menemukan hyperplane ideal yang memaksimalkan margin pemisah antara kelas-kelas. Pemisahan kelas dan generalisasi model pada data baru meningkat seiring dengan margin yang diperoleh (Rozaqi et al., 2021) [25].

Dalam proses pelatihan SVM, algoritma menggunakan data pelatihan yang sudah diberi label kelas untuk membangun model klasifikasi. Proses pelatihan ini digunakan untuk menemukan hyperplane yang optimal dan mempelajari pola dan karakteristik yang membedakan kelas-kelas dalam dataset. Setelah model terbentuk, model tersebut dapat digunakan untuk memprediksi kelas dari data pengujian, yang merupakan data baru yang belum pernah dilihat sebelumnya. Keunggulan utama SVM adalah kemampuan mereka untuk menangani data yang kompleks dan besar. SVM dapat mentransformasi data ke dalam ruang fitur yang lebih besar, di mana pemisahan kelas menjadi lebih mudah dengan menggunakan kernel yang sesuai. Ini termasuk masalah klasifikasi dan prediksi pada berbagai jenis data, termasuk data dengan dimensi yang tinggi, seperti data teks, gambar, atau sinyal (Suryati et al., 2023) [26]. Konsep kerja dari algoritma Support Vector Machine (SVM) dapat dilihat pada Gambar 2



Gambar 2. Hyperplane Support Vector Machine

Untuk memperoleh garis hyperplane yang paling optimal dalam memisahkan data ke dua buah kelas tersebut, maka digunakan perhitungan margin hyperplane dan menemukan titik maksimal. Dalam memperoleh hyperplane pada SVM. Berikut adalah beberapa rumus dari SVM.

$$\text{Fungsi keputusan SVM} \quad F(x) = w^T x + B \quad (2)$$

$$\text{Margin yang ingin dimaksimalkan} \quad \text{Margin} = \frac{1}{\|w\|} \quad (3)$$

$$\text{Kendala untuk data yang terklasifikasi dengan benar} \quad y_i(w^T x_i + b) \geq 1, \quad (4)$$

$$\text{Fungsi tujuan untuk optimasi} \quad \min_{w,b} \frac{1}{2} \|w\|^2 \quad (5)$$

Dalam proses klasifikasi dengan Support Vector Machine (SVM) biasanya ditemui kondisi dimana kernel linear bekerja tidak optimal yang mengakibatkan hasil klasifikasi terhadap data menjadi buruk. Hal tersebut dapat diatasi terhadap data menjadi buruk. Hal tersebut dapat diatasi dengan menggunakan kernel non-linear dengan memanfaatkan kernel trick. Dengan memanfaatkan kernel trick, akan dilakukan mapping data input ke feature space yang dimensinya lebih tinggi sehingga membuat data input yang dihasilkan akan terpisah secara linear dan membentuk hyperplane yang optimal. Persamaan dari setiap kernel SVM dapat dilihat pada Tabel 1

Tabel 1. Persamaan setiap Kernel SVM

Kernel	Persamaan
Linear	$K(x_i, x) = x^T x$

Polynomial	$K(x_i, x) = (\gamma \cdot x_i^T x + r)^p, \gamma > 0$
RBF	$K(x_i, x) = \exp(-\gamma x_i - x ^2), \gamma > 0$
Sigmoid	$K(x_i, x) = \tanh(\gamma x_i^T x + r)$

3. HASIL DAN PEMBAHASAN

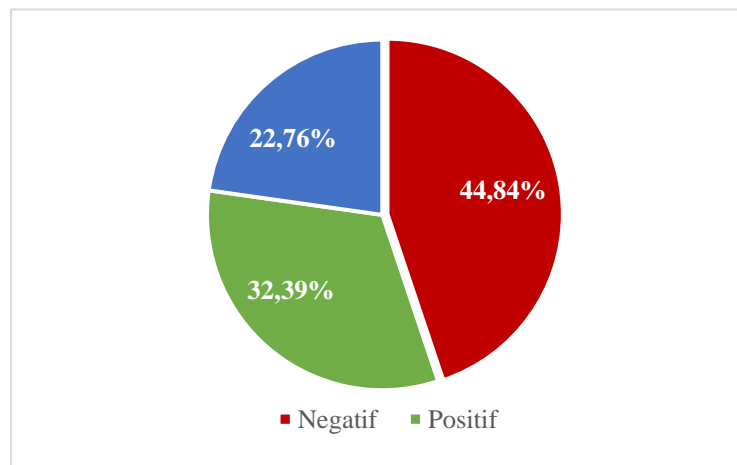
Pada data yang digunakan dalam penelitian ini diambil dari tweet dengan kata kunci “Makan siang gratis”. Pengambil data menggunakan teknik crawling data pada media sosial X. proses crawling menghasilkan 1028 data yang merupakan tweet dari kata kunci makan siang gratis selama 3 bulan terakhir (Januari – Maret 2024). Seluruh data kemudian dilakukan preprocessing yang terdiri dari beberapa proses yaitu cleaning, case folding, tokenizing, stopword & stemming. Pada proses cleaning semua emotikon, mention, hastag, pada data tweet akan dihapus. selanjutnya pada proses case folding semua teks akan diubah menjadi huruf kecil. Kemudian pada proses tokenizing kalimat akan dipecah menjadi kata per kata. Selanjutnya pada proses stopword akan dihapus kata-kata yang bukan merupakan kata kunci pada komentar dan terakhir pada tahapan stemming akan diselaraskan kalimat-kalimat yang ada untuk mengurangi daftar kata pada data training. Hasil dari tahapan preprocessing dapat dilihat pada Tabel 1

Tabel 1. Hasil Preprocessing Data

Proses	Hasil
Data Awal	@OppungTp makan siang gratis siap nya di tahun berapa ?? terus apakah nanti pemerintah yakin makan siang gratis bakal mencapai sampai ke pelosok ? dimana program ini bnyk sekali celah untuk korupsi nya
Cleaning	makan siang gratis siap di tahun berapa terus apakah nanti pemerintah yakin makan siang gratis bakal mencapai sampai ke pelosok dimana program ini bnyk sekali celah untuk korupsi
Tokenizing	['makan', 'siang', 'gratis', 'siap', 'di', 'tahun', 'berapa', 'terus', 'apakah', 'nanti', 'pemerintah', 'yakin', 'makan', 'siang', 'gratis', 'bakal', 'mencapai', 'sampai', 'ke', 'pelosok', 'dimana', 'program', 'ini', 'bnyk', 'sekali', 'celah', 'untuk', 'korupsi']
Stopword Removal	['makan', 'siang', 'gratis', 'pemerintah', 'makan', 'siang', 'gratis', 'mencapai', 'pelosok', 'dimana', 'program', 'bnyk', 'celah', 'korupsi']
Stemming	['makan', 'siang', 'gratis', 'perintah', 'makan', 'siang', 'gratis', 'capai', 'pelosok', 'diamana', 'program', 'bnyk', 'celah', 'korupsi']

3.1 Pelabelan Data

Data yang dimiliki saat ini masih belum memiliki label, sehingga masih perlu melakukan proses pelabelan terlebih dahulu. Pada penelitian ini proses pelabelan dilakukan secara otomatis menggunakan kamus sentimen bahasa Indonesia (InSet). Jika hasil polaritas memiliki nilai diatas satu (> 1) maka teks akan diberi label positif. Kemudian jika hasil polaritas memiliki nilai dibawah satu (< 1) maka teks akan diberi label negatif. Dan jika hasil polaritas sama dengan nol (0) maka teks akan diberi label netral. Hasil dari proses pelabelan data dapat dilihat pada gambar 3.



Gambar 3. Hasil Pelabelan Data

Melalui hasil pelabelan data yang dilakukan secara otomatis menggunakan kamus sentimen bahasa Indonesia, menghasilkan sentimen positif sebanyak 32.39% (333 sentimen), negatif 44.84% (461 sentimen) dan netral 22.76% (234 sentimen) dari total sebanyak 1028 data komentar. Berdasarkan hasil tersebut menunjukkan bahwa mayoritas

masyarakat saat ini masih mempertanyakan dan mengkritik program makan gratis, sehingga sentimen yang diberikan cenderung negatif.

3.2 Hasil Implementasi TF-IDF

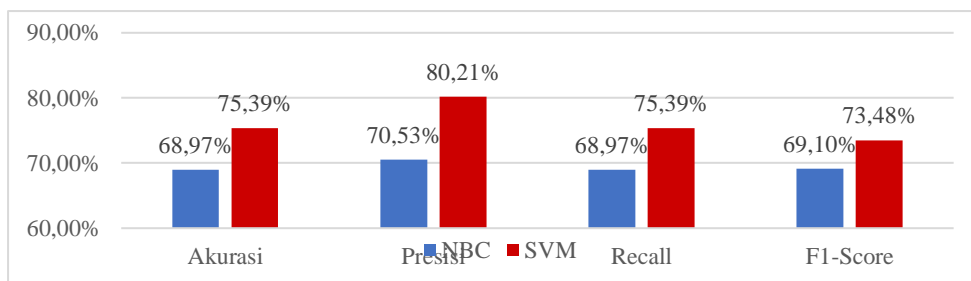
Data yang telah diberi label akan dilakukan pembobotan menggunakan TF-IDF. Ini merupakan tahapan terpenting sebelum melakukan pemodelan. Semakin sering suatu kata muncul dalam satu kalimat, maka bobot yang dihasilkan akan semakin tinggi. Hasil implementasi pembobotan dengan TF-IDF dapat dilihat pada tabel 2.

Tabel 2. Hasil Implementasi Teknik TF-IDF

No	Anak	demo	Duit	Gratis	Makan	...	yakin
1	0,0000	0,0000	0,5036	0,0932	0,0930	...	0,0000
2	0,0000	0,6685	0,0000	0,0986	0,1969	...	0,0000
...
1027	0,0000	0,0000	0,0000	0,0457	0,0456	...	0,0000
1028	0,0000	0,0000	0,0000	0,0386	0,0771	...	0,0000

3.3 Hasil Evaluasi Model NBC dan SVM

Tahapan pemodelan dilakukan menggunakan dua algoritma yaitu Naïve Bayes (NBC) dan Support Vector Machine (SVM). Algoritma NBC yang diimplementasikan adalah tipe Gaussian, sedangkan pada SVM adalah kernel Radial Basis Function (RBF). Model yang telah dilatih kemudian dievaluasi menggunakan beberapa metrik yaitu akurasi, presisi, recall, dan F1-score untuk mengetahui kemampuan model melakukan klasifikasi. Hasil evaluasi masing-masing algoritma akan dibandingkan untuk mendapatkan algoritma terbaik pada kasus analisis sentimen program makan gratis. Hasil evaluasi model NBC dan SVM dapat dilihat pada gambar 4.



Gambar 4. Hasil Evaluasi Model NBC dan SVM

Berdasarkan hasil evaluasi model tersebut, dapat diketahui bahwa algoritma SVM memiliki kemampuan yang lebih baik dibandingkan dengan NBC dengan nilai akurasi sebesar 75.39%, presisi 80.21%, recall 75.39% dan F1-score 73.48%. Sementara itu NBC hanya memiliki akurasi sebesar 68.97%, presisi 70.53%, recall 68.97% dan F1-score 69.10%. Meskipun SVM memiliki akurasi yang lebih baik, namun SVM memerlukan waktu komputasi yang sedikit lebih lama dibandingkan NBC.

3.4 Hasil Evaluasi Data

Tahapan berikutnya adalah melakukan visualisasi kemunculan kata menggunakan word cloud. Proses ini dilakukan untuk menampilkan kata yang sering muncul pada setiap kategori sentimen. Pada sentimen positif, 5 kata yang sering muncul yaitu ‘ajar’, ‘india’, ‘china’, ‘didik’, ‘sekolah’. Kata tersebut menunjukkan bahwa mayoritas komentar pada sentimen ini membahas tentang pemerintah yang mulai memperelajari program makan siang gratis, bahkan melakukan kunjungan ke sekolah di negara lain yang telah menjalankan program makan gratis sebelumnya, yaitu china dan india. Hasil visualisasi sentimen positif dapat dilihat pada gambar 5.



Gambar 5. Visualisasi Sentimen Positif



- J. Indones. Manaj. Inform. dan Komun.*, vol. 5, no. 3, 2024, [Online]. Available: <https://journal.stmiki.ac.id/index.php/jimik/article/view/978/770>
- [7] Z. Purwanti, “Pemodelan Text Mining untuk Analisis Sentimen Terhadap Program Makan Siang Gratis di Media Sosial X Menggunakan Algoritma Support Vector Machine (SVM),” vol. 5, no. 3, pp. 3065–3079, 2024.
- [8] A. Sitanggang, R. I. Umaidah, Yuyun, and Adam, “Analisis Sentimen Masyarakat Terhadap Program Makan Siang Gratis Pada Media Sosial X Menggunakan Algoritma Naïve Bayes,” *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 3, 2024, doi: 10.23960/jitet.v12i3.4902.
- [9] M. A. S. Putra, I. Permana, M. Mustakim, and M. Afdal, “Analisis Sentimen Masyarakat Mengenai Gerakan Childfree di Media Sosial X Menggunakan Algoritma NBC dan SVM,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 4, pp. 1189–1198, 2024, doi: 10.57152/malcom.v4i4.1356.
- [10] R. Afandi, M. Afdal, and R. Novita, “Analisis Sentimen Masyarakat Terhadap Pinjaman Online di Twitter Menggunakan Algoritma Naïve Bayes Classifier dan K-Nearest Neighbor,” vol. 6, no. 2, pp. 596–605, 2024, doi: 10.47065/bits.v6i2.5300.
- [11] R. N. Yudistira Arya Wibisono, M. Afdal, Mustakim, “Implementasi Algoritma Support Vector Machine Untuk Analisa Sentimen Data Ulasan Aplikasi Pinjaman Online di Google Play Store,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 4, pp. 1244–1252, 2024, doi: 10.57152/malcom.v4i4.1435.
- [12] Y. A. Putri, S. Defit, and G. W. Nurcahyo, “Analisis Sentimen Publik Terhadap Program Penurunan Angka Prevalensi Stunting Indonesia Menggunakan Data Twitter Dengan Metode Naïve Bayes,” vol. 4, pp. 1978–1989, 2024.
- [13] D. A. Warraihan, I. Permana, M. Mustakim, and R. Novita, “Analisis Sentimen Pengguna Transportasi Online Maxim Pada Instagram Menggunakan Naïve Bayes Classifier dan K-Nearest Neighbor,” *J. Media Inform. Budidarma*, vol. 7, no. 3, p. 1134, 2023, doi: 10.30865/mib.v7i3.6336.
- [14] I. P. Rahayu, A. Fauzi, and J. Indra, “Analisis Sentimen Terhadap Program Kampus Merdeka Menggunakan Naive Bayes Dan Support Vector Machine,” *J. Sist. Komput. dan Inform.*, vol. 4, no. 2, p. 296, 2022, doi: 10.30865/json.v4i2.5381.
- [15] G. Ginabila and A. Fauzi, “Analisis Sentimen Terhadap Pemutar Musik Online Spotify Dengan Algoritma Naive Bayes dan Support Vector Machine,” *J. Ilm. Ilk. - Ilmu Komput. Inform.*, vol. 6, no. 2, pp. 111–122, 2023, doi: 10.47324/ilkominfo.v6i2.180.
- [16] E. S. Romaito, M. K. Anam, Rahmadden, Ulfah, and A. Noviciate, “Perbandingan Algoritma SVM Dan NBC Dalam Analisa Sentimen Pilkada Pada Twitter,” pp. 169–179, 2021, [Online]. Available: <https://doi.org/10.22303/csrid.13.3.2021.169-179>
- [17] D. Ananda and R. R. Suryono, “Analisis Sentimen Publik Terhadap Pengungsi Rohingya di Indonesia dengan Metode Support Vector Machine dan Naïve Bayes,” *J. Media Inform. Budidarma*, vol. 8, no. 2, p. 748, 2024, doi: 10.30865/mib.v8i2.7517.
- [18] A. Muhammadin and I. A. Sobari, “Analisis Sentimen Pada Ulasan Aplikasi Kredivo Dengan Algoritma Svm Dan Nbc,” *Reputasi J. Rekayasa Perangkat Lunak*, vol. 2, no. 2, pp. 85–91, 2021, doi: 10.31294/reputasi.v2i2.785.
- [19] G. K. Locarso, “Analisis Sentimen Review Aplikasi Pedulilindungi Pada Google Play Store Menggunakan Nbc,” *JTIK (Jurnal Tek. Inform. Kaputama)*, vol. 6, no. 2, pp. 353–361, 2022, doi: 10.59697/jtik.v6i2.207.
- [20] S. Kurniawan, W. Gata, D. A. Puspitawati, M. Tabrani, and K. Novel, “Perbandingan Metode Klasifikasi Analisis Sentimen Tokoh Politik Pada Komentar Media Berita Online,” vol. 1, no. 10, pp. 2–8, 2021.
- [21] M. H. Saragih and A. S. Girsang, “Sentiment analysis of customer engagement on social media in transport online,” *Proc. - 2017 Int. Conf. Sustain. Inf. Eng. Technol. SIET 2017*, vol. 2018-Janua, pp. 24–29, 2017, doi: 10.1109/SIET.2017.8304103.
- [22] P. S. M. Suryani, L. Linawati, and K. O. Saputra, “Penggunaan Metode Naïve Bayes Classifier pada Analisis Sentimen Facebook Berbahasa Indonesia,” *Maj. Ilm. Teknol. Elektro*, vol. 18, no. 1, p. 145, 2019, doi: 10.24843/mite.2019.v18i01.p22.
- [23] B. agus Sholekha Inez, Faqih Ahmad, “Sentiment Analysis Of Publik Opinion Covid Vaccine Using Naive Bayes And Random Forest Methods,” *J. Tek. Inform. Atmaluhur*, vol. 6, no. 1, p. 4, 2022.
- [24] A. R. Isnain, A. I. Sakti, D. Alita, and N. S. Marga, “Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm,” *J. Data Min. dan Sist. Inf.*, vol. 2, no. 1, p. 31, 2021, doi: 10.33365/jdmsi.v2i1.1021.
- [25] A. J. Rozaqi, A. Sunyoto, and M. rudyanto Arief, “Deteksi Penyakit Pada Daun Kentang Menggunakan Pengolahan Citra dengan Metode Convolutional Neural Network,” *Creat. Inf. Technol. J.*, vol. 8, no. 1, p. 22, 2021, doi: 10.24076/citec.2021v8i1.263.
- [26] A. A. A. Suryati emi, Styawati Styawati, “Analisis Sentimen Transportasi Online Menggunakan Ekstraksi Fitur Model Word2vec Text Embedding Dan Algoritma Support Vector Machine,” *J. Teknol. Dan Sist. Inf.*, vol. 4, no. 1, pp. 96–106, 2023.