



Department of Digital Business

Journal of Artificial Intelligence and Digital Business (RIGGS)

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 4 No. 3 (2025) pp: 3906-3917

P-ISSN: 2963-9298, e-ISSN: 2963-914X

Analisis Sentimen Program Makan Bergizi Gratis Pemerintah RI Melalui Twitter Menggunakan Metode SVM

Fatkhurrohman Fatkhurrohman¹, Bangkit Indarmawan Nugroho², Nurul Fadillah³

^{1,2}Sistem Informasi, STMIK YMI Tegal

³Teknik Informatika, STMIK YMI Tegal

fatkhurrohman150602@gmail.com*, bangkit_in@stmik-tegal.ac.id, nurul@stmik-tegal.ac.id

Abstrak

Program Pemberian Makanan Sehat secara Gratis yang diluncurkan oleh pemerintah menerima beragam reaksi dari masyarakat, khususnya melalui media sosial seperti Twitter. Studi ini dilakukan untuk mengevaluasi perasaan publik mengenai program itu dengan tujuan memberikan gambaran opini publik yang lebih terstruktur dan terukur. Data diperoleh dari Twitter dengan teknik crawling, kemudian melewati proses preprocessing, normalisasi, dan konversi menjadi representasi numerik menggunakan teknik TF-IDF. Selanjutnya, sentimen dikategorikan melalui penerapan algoritma Support Vector Machine (SVM) ke dalam tiga kelas: positif, negatif, dan netral. Hasil penelitian menunjukkan bahwa model tersebut berhasil dengan akurasi mencapai 80% dan weighted F1-score di angka 0,76. Sentimen positif menjadi kategori yang paling dominan dan terdeteksi dengan akurasi tinggi, ditunjukkan oleh nilai recall sebesar 0,98 dan F1-score sebesar 0,88. Sebaliknya, pendapat netral dan negatif masih memperlihatkan kelemahan dalam klasifikasi, terutama dari sisi recall dan F1-score yang sangat rendah. Temuan ini mengindikasikan bahwa model SVM bisa digunakan secara efisien untuk menganalisis opini publik dalam bahasa Indonesia, khususnya pada isu kebijakan sosial yang banyak dibicarakan di media sosial.

Kata kunci: Analisis Sentimen, Klasifikasi Teks, Makan Bergizi, SVM, Twitter

1. Latar Belakang

Kebijakan Program Makan Bergizi Gratis yang diusung pemerintah menuai berbagai reaksi dari masyarakat [1]. Sebagian masyarakat menyambutnya sebagai langkah positif dalam pemenuhan gizi anak-anak sekolah dan kelompok rentan, namun tak sedikit pula yang meragukan efektivitas dan kesiapan implementasi program ini, terutama dari sisi anggaran dan distribusi [2]. Di zaman digital, platform seperti Twitter menjadi saluran utama bagi orang-orang untuk menyuarakan pendapat secara real-time, sehingga platform ini menyimpan data opini publik yang sangat bernilai untuk dianalisis [3]. Permasalahan utama yang disoroti dalam penelitian ini adalah bagaimana mengidentifikasi persepsi masyarakat terhadap program tersebut secara terstruktur dan objektif melalui data media sosial [4].

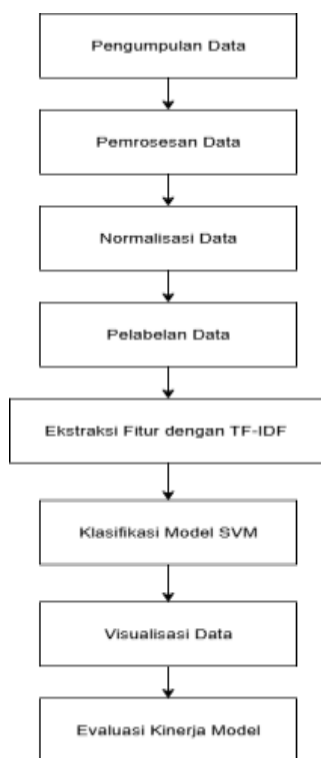
Penelitian terdahulu telah banyak memanfaatkan media sosial untuk menganalisis respons publik terhadap kebijakan, seperti pada isu PSBB, pandemi COVID-19, dan kenaikan harga bahan pokok [5]. Metode yang umum diterapkan untuk klasifikasi sentimen mencakup algoritma kecerdasan buatan seperti Naive Bayes dan Support Vector Machine (SVM), yang terbukti efektif dalam mengklasifikasikan pendapat ke dalam kategori positif, negatif, dan netral [6]. Namun, belum ada studi yang khusus menerapkan metode SVM untuk menilai pandangan publik mengenai program makan gratis lewat Twitter dalam konteks kebijakan baru [7]. Hal ini menunjukkan adanya celah riset yang bisa dilengkapi oleh studi ini, serta menekankan kontribusi orisinal dalam konteks analisis sentimen berbasis teks berbahasa Indonesia.

Penelitian ini memiliki tujuan untuk menilai pandangan masyarakat mengenai program makan bergizi gratis dengan memanfaatkan informasi yang diperoleh dari Twitter. Untuk itu, metode Support Vector Machine digunakan guna mengklasifikasikan tweet ke dalam tiga jenis sentimen, yaitu positif, negatif, dan netral [8]. Selain itu, penelitian ini juga mengevaluasi kinerja metode SVM dalam tugas klasifikasi teks berbahasa Indonesia melalui pengukuran akurasi, presisi, recall, dan F1-score. Dengan pendekatan ini, penelitian diharapkan dapat memberikan wawasan berbasis data bagi pengambil kebijakan dalam menilai efektivitas program serta memperbaiki strategi komunikasi publik. Dengan demikian, penelitian ini difokuskan pada analisis sentimen masyarakat di media sosial Twitter sebagai objek penelitian, dengan permasalahan utama yaitu bagaimana mengklasifikasikan opini publik terhadap Program Makan Bergizi Gratis menggunakan metode SVM.

2. Metode Penelitian

2.1. Desain Penelitian

Penelitian ini menerapkan metode kuantitatif karena informasi yang dianalisis berbentuk data numerik yang didapat dari pengelompokan sentimen terkait tweet masyarakat mengenai Program Makan Bergizi Gratis. Adapun metode deskriptif bertujuan untuk menggambarkan hasil analisis berdasarkan pola-pola yang ditemukan dalam data. Proses penelitian terdiri dari beberapa tahapan yang dilakukan secara sistematis, mulai dari pengumpulan data hingga evaluasi model klasifikasi sentimen. Tahapan penelitian ini dapat dijelaskan melalui alur berikut.



Gambar 1. Alur Penelitian

Prosedur penilaian opini dilaksanakan melalui sejumlah langkah yang berhubungan, diawali dengan mengumpulkan dan membersihkan informasi berupa teks, dilanjutkan dengan normalisasi untuk memastikan konsistensi bahasa. Data yang telah diproses kemudian dilabeli sesuai kategori sentimen yang ditentukan, sebelum dipresentasikan dalam bentuk angka melalui teknik TF-IDF. Representasi ini memungkinkan data diterapkan dalam pelatihan model klasifikasi dengan memanfaatkan algoritma Support Vector Machine (SVM). Setelah model dilatih, dilakukan evaluasi untuk mengukur performa klasifikasi, serta visualisasi data untuk membantu interpretasi hasil secara lebih intuitif. Tahapan-tahapan ini dirancang agar sistem mampu melakukan analisis sentimen secara efektif dan menyajikan hasil yang mudah dipahami. Berikut adalah penjelasan setiap tahapannya:

2.2. Pengumpulan Data

Tahap ini dilaksanakan dengan mengumpulkan informasi melalui Twitter menggunakan teknik web scraping. Kata kunci yang digunakan antara lain “makan bergizi gratis”, “program makan gratis”, dan variasi lain yang relevan dengan topik kebijakan pemerintah. Data yang diambil difilter agar hanya mencakup tweet berbahasa Indonesia, dalam rentang waktu tertentu, dan tidak mengandung retweet untuk menghindari duplikasi. Data dikumpulkan dalam bentuk teks mentah untuk kemudian diproses lebih lanjut.

2.3. Pemrosesan Data

Tweet yang telah dikumpulkan kemudian melalui proses preprocessing agar siap untuk dianalisis. Tahap ini mencakup penghapusan karakter yang tidak biasa dan alamat web dari data (cleansing), mengubah semua karakter menjadi huruf kecil, pembagian kata, penghilangan istilah umum yang tidak mengandung arti signifikan (stopword removal), dan pengurangan kata menjadi bentuk aslinya. Tujuan preprocessing adalah untuk meningkatkan akurasi klasifikasi dengan mengurangi noise dalam data.

2.4. Normalisasi Data

Setelah preprocessing, dilakukan normalisasi teks. Langkah ini digunakan untuk mengganti kata-kata tidak formal atau istilah slang menjadi bentuk resmi yang sesuai dengan norma bahasa Indonesia. Proses ini penting karena banyak tweet yang menggunakan gaya bahasa informal, sehingga tanpa normalisasi, makna kata dapat terdistorsi.

2.5. Pemberian Label pada Dataset

Data yang telah dinormalisasi kemudian diberi label sentimen secara otomatis. Pelabelan dilakukan dengan pendekatan berbasis leksikon, yaitu menggunakan daftar kata yang sudah dikelompokkan menjadi emosi positif, negatif, atau netral. Setiap tweet akan dianalisis menurut istilah yang terdapat di dalamnya untuk menentukan kategori sentimen yang paling sesuai.

2.6. Ekstraksi Fitur

Setelah data diberi label, langkah selanjutnya adalah mengubah teks menjadi bentuk angka dengan memanfaatkan metode TF-IDF. TF-IDF menghitung seberapa penting suatu kata dalam sebuah dokumen relatif terhadap seluruh kumpulan dokumen. Hasil dari tahapan ini ialah vektor numerik yang bisa digunakan sebagai input dalam algoritma machine learning.

2.7. Klasifikasi SVM.

Data yang telah dikonversi menjadi vektor fitur selanjutnya dimanfaatkan dalam tahapan pelatihan model klasifikasi dengan penerapan algoritma SVM. Algoritma ini dipilih karena dapat menangani data dengan banyak dimensi seperti teks dengan efektif dan menunjukkan performa cukup baik dalam klasifikasi dua atau lebih kelas. Model akan mempelajari pola dari data latih untuk kemudian digunakan dalam mengklasifikasikan data uji.

2.8. Visualisasi Data

Visualisasi data dilakukan untuk mendukung proses analisis serta memahami pola opini masyarakat mengenai inisiatif Makan Bergizi Gratis. Visualisasi yang diterapkan meliputi awan kata untuk memperlihatkan istilah yang paling sering diucapkan, grafik distribusi tweet berdasarkan waktu untuk mengamati tren percakapan, diagram *pie chart* untuk menunjukkan proporsi sentimen, dan *heatmap* untuk menggambarkan intensitas waktu penggunaan dalam membuat tweet.

2.9. Evaluasi Kinerja Model

Evaluasi terhadap kinerja model dilakukan menggunakan Matrik Evaluasi dan *Confusion Matrik* seperti akurasi, presisi, *recall*, *F1-score*, *macro avg fl*, dan *micro avg fl*. Selain itu, confusion matrix juga digunakan untuk melihat jumlah prediksi akurat dan tidak akurat dalam setiap kategori sentimen. Evaluasi penting dalam mengetahui sejauh mana sistem mampu mengklasifikasikan tweet dengan baik [9].

2.9.1. Akurasi

Nilai akurasi dihitung dengan membagi jumlah prediksi yang tepat (baik kategori positif maupun negatif) dengan keseluruhan data yang ada. Nilai akurasi yang semakin tinggi menunjukkan bahwa model memiliki performa yang lebih baik dalam mendeteksi pola data [10]. Namun demikian, dalam kasus klasifikasi dengan data tidak seimbang, akurasi saja tidak cukup untuk merepresentasikan performa model secara menyeluruh. Rumus Akurasi dapat ditulis pada persamaan (1)

$$\text{Akurasi} = \frac{\text{Jumlah prediksi data}}{\text{Total data}} \quad (1)$$

2.9.2. Presisi

Presisi mengukur seberapa akurat model dalam meramalkan kategori positif. Tingginya nilai presisi menunjukkan bahwa mayoritas dari prediksi positif adalah benar, yaitu memiliki *True Positive (TP)* yang tinggi dan *False Positive (FP)* yang rendah [11]. Presisi menjadi penting terutama pada kasus di mana kesalahan klasifikasi positif membawa konsekuensi besar. Dengan demikian, semakin tinggi nilai presisi, semakin kecil kemungkinan sistem memberikan hasil positif yang salah. Dalam penelitian ini, presisi digunakan sebagai salah satu metrik utama untuk menilai kinerja metode SVM dalam klasifikasi sentimen. Rumus Presisi dapat ditulis pada persamaan (2)

$$\text{Presisi} = \frac{\text{True Positif (TP)}}{\text{TP+False Positive (FP)}} \quad (2)$$

2.9.3. Recall

Recall mengukur seberapa besar data positif yang berhasil ditemukan oleh model dari seluruh data yang memang positif. *False Negative (FN)* merujuk pada informasi yang seharusnya positif tetapi gagal dikenali oleh model. *Recall* yang tinggi menunjukkan bahwa sistem memiliki kepekaan yang baik terhadap data positif. *Recall* sangat penting pada kasus di mana kehilangan data positif dapat menimbulkan kerugian besar, misalnya dalam deteksi fraud atau analisis opini positif masyarakat [12]. Semakin tinggi nilai recall, semakin kecil kemungkinan data positif terlewat oleh sistem. Dalam penelitian ini, recall digunakan untuk mengukur sejauh mana metode SVM mampu menangkap sentimen positif secara menyeluruh. Rumus *Recall* dapat ditulis pada persamaan (3)

$$Recall = \frac{True\ Positif\ (TP)}{TP + False\ Negative\ (FN)} \quad (3)$$

2.9.4. F1-Score

F1-Score mengombinasikan precision dan recall menjadi satu ukuran yang seimbang. Nilai F1 akan rendah jika salah satu dari presisi atau recall rendah. Metrik ini cocok digunakan saat distribusi kelas tidak seimbang karena mencerminkan keseimbangan kinerja model [13]. Selain itu, F1-Score mampu memberikan evaluasi yang lebih adil dibandingkan hanya menggunakan precision atau recall secara terpisah. Dengan demikian, metrik ini sering dijadikan acuan utama dalam penelitian analisis sentimen maupun klasifikasi teks. Rumus *F1-Score* dapat ditulis pada persamaan (4)

$$F1\text{-Score} = 2 \times \frac{Presisi \times Recall}{Presisi + Recall} \quad (4)$$

2.9.5. Macro Average F1

Dalam pendekatan ini, setiap kelas dihitung F1-nya terlebih dahulu, kemudian dirata-ratakan. Hal ini bermanfaat dalam situasi data yang tidak seimbang, karena metrik ini tidak berat sebelah terhadap kelas yang jumlah datanya lebih banyak. Nilai *Macro F1* yang tinggi menandakan bahwa model bekerja dengan baik pada semua kelas secara merata [14]. Rumus *Macro Average F1* dapat ditulis pada persamaan (5)

$$Macro\ F1 = \frac{F1_{positif} + F1_{negatif}}{3} \quad (5)$$

2.9.6. Weighted Avg F1

Weighted Avg F1 merupakan rata-rata F1-Score dari tiap kelas yang disesuaikan dengan proporsi jumlah data pada masing-masing kelas. Metrik ini memperhitungkan ketidakseimbangan data, sehingga memberikan penilaian performa model yang lebih adil. *Weighted Avg F1* sangat berguna ketika distribusi kelas tidak seimbang, karena metrik ini mencegah hasil evaluasi didominasi oleh kelas mayoritas [15]. Semakin tinggi nilai *Weighted Avg F1*, semakin baik model dalam menjaga keseimbangan performa di semua kelas. Dalam penelitian ini, *Weighted Avg F1* dipakai untuk menilai konsistensi kinerja SVM dalam mengklasifikasikan sentimen positif, negatif, dan netral. Rumus *Weighted Avg F1* dapat ditulis pada persamaan (6)

$$Weighted\ F1 = \frac{n_1 \cdot F1_1 + n_2 \cdot F1_2 + n_3 \cdot F1_3}{n} \quad (6)$$

3. Hasil dan Diskusi

3.1 Pengumpulan Data

Data sebanyak 562 tweet mentah dikumpulkan selama bulan Mei 2025, yang seluruhnya berkaitan langsung dengan topik penelitian. Tweet yang diperoleh kemudian disaring agar hanya mencakup tweet berbahasa Indonesia dan tidak mengandung retweet, untuk memastikan data bersih dari duplikasi dan sesuai konteks penelitian. Pemilihan kata kunci dilakukan secara spesifik agar data yang dikumpulkan benar-benar mencerminkan opini publik terhadap kebijakan tersebut. Proses crawling dilakukan secara otomatis dalam jangka waktu tertentu untuk memperoleh data yang aktual dan relevan. Data hasil crawling disimpan dalam format terstruktur sehingga siap digunakan pada tahap pra-proses. Langkah ini juga memastikan konsistensi serta validitas data penelitian. Selanjutnya dilakukan pengecekan ulang terhadap dataset untuk menghindari adanya noise. Dataset akhir kemudian digunakan sebagai dasar dalam proses analisis sentimen.

Tabel 1. Pengumpulan Data

No.	Nomor Tweet	Waktu	Nomor Penulis	Tweet
1.	19243466375 66279932	2025-05- 19T06:09:25.000Z	1307860511317 352450	Dengerin nih guess kata Neneng Jebred yg sering nyinyir dan minta progrm makan bergizi gratis dihentikn, pdahal jlas program MBG itu memberikn gizi untuk mencerdaskan generasi. Mari bersama dukung Lanjutkan MBG #MBGMembangunSDMUnggul Amanda Zahra Villains Parkinson Trump is 6 https://t.co/vE4Xi0u58l
2.	19243412427 56010006	2025-05- 19T05:47:59.000Z	1725815701674 590208	Program Makan Bergizi Gratis terbukti membantu pemerintah dalam meningkatkan pemerataan dan pemenuhan gizi anak-anak di seluruh wilayah Indonesia. #MakanBergiziGratis #GiziAnakMerata #IndonesiaSehat https://t.co/9T3s0uGyfj
...
561.	19292138635 53057167	2025-06- 01T16:30:02.000Z	1859990476381 995008	Makan Bergizi Gratis mampu meningkatkan gizi masyarakat, mencegah stunting, dan menumbuhkan ekonomi kerakyatan. Dukung terus Program MBG demi Indonesia maju. #FaktaIndonesia #MakanBergiziGratis #ManfaatMakanBergiziGratis https://t.co/nKhVtyE96g
562.	19292110274 44359519	2025-06- 01T16:18:46.000Z	1725413776508 497920	@afrizalpradito @idextratime Program Makan Bergizi Gratis (MBG), selain untuk anak sekolah, juga dibagikan ke ibu hamil/menyusui & kelompok bayi dibawah usia 5 th Hal ini bertujuan mencukupi kebutuhan gizi anak, yg dibutuhkan sejak periode golden age, yakni 1.000 hari pertama sejak lahir Lanjutkan MBG ❤️

Tabel 1 merupakan hasil dari proses pengambilan data menggunakan API Twitter. Tabel ini memuat beberapa informasi penting dari setiap tweet yang berhasil dikumpulkan, seperti Tweet ID, Author ID, Created At, dan Text. Setiap baris pada tabel mewakili satu tweet yang relevan dengan topik penelitian, yaitu program makan bergizi gratis dari pemerintah. Proses pengumpulan ini dilakukan dengan memanfaatkan kata kunci tertentu dan alat bantu seperti Google Colab serta pustaka requests dan pandas untuk melakukan scraping.

3.2 Pemrosesan Data

Tahap ini membantu mengurangi noise dalam data dan menekankan pada penggunaan kata-kata yang memiliki makna. Dengan data yang telah bersih dan terstruktur, proses klasifikasi sentimen pada tahap selanjutnya dapat berjalan lebih optimal. Proses ini juga berperan penting dalam mengurangi beban komputasi saat ekstraksi fitur dan pelatihan model dilakukan.

Tabel 2. Hasil Pemrosesan

No.	Tweet	Hasil pemrosesan
1.	Dengerin nih guess kata Neneng Jebred yg sering nyinyir dan minta progrm makan bergizi gratis dihentikn, pdahal jlas program MBG itu memberikn gizi untuk mencerdaskan generasi. Mari bersama dukung Lanjutkan MBG #MBGMembangunSDMUnggul Amanda Zahra Villains Parkinson Trump is 6 https://t.co/vE4Xi0u58l	dengerin nih guess kata neneng jebred yg sering nyinyir dan minta progrm makan bergizi gratis dihentikn pdahal jlas program mbg itu memberikn gizi untuk mencerdaskan generasi mari bersama dukung lanjutkan mbg mbgmembangunsdmunggul amanda zahra villains parkinson trump is 6
2.	Program Makan Bergizi Gratis terbukti membantu pemerintah dalam meningkatkan pemerataan dan pemenuhan gizi anak-anak di seluruh wilayah Indonesia. #MakanBergiziGratis #GiziAnakMerata #IndonesiaSehat https://t.co/9T3s0uGyfj	program makan bergizi gratis terbukti membantu pemerintah dalam meningkatkan pemerataan dan pemenuhan gizi anakanak di seluruh wilayah indonesia makanbergizigratis gizianakmerata indonesiasehat
...
561.	Makan Bergizi Gratis mampu meningkatkan gizi masyarakat, mencegah stunting, dan menumbuhkan ekonomi kerakyatan. Dukung terus Program MBG demi Indonesia maju. #FaktaIndonesia #MakanBergiziGratis #ManfaatMakanBergiziGratis https://t.co/nKhVtyE96g	makan bergizi gratis mampu meningkatkan gizi masyarakat mencegah stunting dan menumbuhkan ekonomi kerakyatan dukung terus program mbg demi indonesia maju faktaindonesia makanbergizigratis manfaatmakanbergizigratis
562.	@afrizalpradito @idextratime Program Makan Bergizi Gratis (MBG), selain untuk anak sekolah, juga dibagikan ke ibu hamil/menyusui & kelompok bayi dibawah usia 5 th Hal ini bertujuan mencukupi kebutuhan gizi anak, yg dibutuhkan sejak periode golden age, yakni 1.000 hari pertama sejak lahir Lanjutkan MBG ❤️	program makan bergizi gratis mbg selain untuk anak sekolah juga dibagikan ke ibu hamilmenyusui amp kelompok bayi dibawah usia 5 th hal ini bertujuan mencukupi kebutuhan gizi anak yg dibutuhkan sejak periode golden age yakni 1000 hari pertama sejak lahir lanjutkan mbg

Tabel 2 disusun dalam dua bagian utama, salah satunya adalah 'Tweet' yang memuat teks mentah berupa cuitan asli pengguna, dan kolom "Hasil Pemrosesan" yang berisi hasil teks setelah dilakukan tahapan preprocessing.

3.3 Normalisasi Data

Proses normalisasi dibantu oleh kamus konversi yang diambil dari repositori GitHub, yang memuat pasangan kata tidak baku dan bentuk bakunya. Hasil dari proses ini digunakan untuk memastikan setiap kata dalam tweet dapat dikenali dengan benar oleh sistem dan diklasifikasikan ke dalam polaritas sentimen yang sesuai.

Tabel 3. Hasil Normalisasi

No.	Hasil pemrosesan	Hasil normalisasi
1.	dengerin nih guess kata neneng jebred yg sering nyinyir dan minta progrm makan bergizi gratis dihentikn pdahal jlas program mbg itu memberikn gizi untuk mencerdaskan generasi mari bersama dukung lanjutkan mbg mbgmembangunsdmunggul amanda zahra villains parkinson trump is 6	dengerin nih guess kata neneng jebred yang sering nyinyir dan minta progrm makan bergizi gratis dihentikn pdahal jlas program mbg itu memberikan gizi untuk mencerdaskan generasi mari bersama dukung lanjutkan mbg mbgmembangunsdmunggul amanda zahra villains parkinson trump is 6
2.	program makan bergizi gratis terbukti membantu pemerintah dalam meningkatkan pemerataan dan pemenuhan gizi anakanak di seluruh wilayah indonesia makanbergizigratis gizianakmerata indonesia sehat	program makan bergizi gratis terbukti membantu pemerintah dalam meningkatkan pemerataan dan pemenuhan gizi anakanak di seluruh wilayah indonesia makanbergizigratis gizianakmerata indonesia sehat
...
561.	makan bergizi gratis mampu meningkatkan gizi masyarakat mencegah stunting dan menumbuhkan ekonomi kerakyatan dukung terus program mbg demi indonesia maju faktaindonesia makanbergizigratis manfaatmakanbergizigratis	makan bergizi gratis mampu meningkatkan gizi masyarakat mencegah stunting dan menumbuhkan ekonomi kerakyatan dukung terus program mbg demi indonesia maju faktaindonesia makanbergizigratis manfaatmakanbergizigratis
562.	program makan bergizi gratis mbg selain untuk anak sekolah juga dibagikan ke ibu hamil menyusui amp kelompok bayi dibawah usia 5 th hal ini bertujuan mencukupi kebutuhan gizi anak yg dibutuhkan sejak periode golden age yakni 1000 hari pertama sejak lahir lanjutkan mbg	program makan bergizi gratis mbg selain untuk anak sekolah juga dibagikan ke ibu hamil menyusui amp kelompok bayi dibawah usia 5 tahun hal ini bertujuan mencukupi kebutuhan gizi anak yang dibutuhkan sejak periode golden age yakni 1000 hari pertama sejak lahir lanjutkan mbg

Tabel 3 menunjukkan hasil dari proses normalisasi. Kolom Tweet berisi teks mentah hasil praproses awal, namun masih mengandung berbagai kata tidak baku, singkatan, dan istilah informal yang umum digunakan oleh pengguna media sosial. Sedangkan kolom Tweet_Normalisasi, yaitu perubahan bentuk baku sesuai dengan kamus normalisasi yang digunakan. Misalnya, pada entri pertama, kata "yg" diubah menjadi "yang" dan "pdahal" menjadi "padahal". Proses ini dilakukan secara sistematis terhadap seluruh data untuk menghasilkan representasi teks yang lebih bersih, seragam, dan sesuai dengan kaidah Bahasa Indonesia.

3.4 Pemberian Label pada Dataset

Pelabelan dilaksanakan otomatis melalui pendekatan berbasis lexicon, yaitu dengan membandingkan setiap kata dalam tweet terhadap kamus kata berkonotasi positif dan negatif yang diperoleh dari sumber daring. Skor sentimen ditentukan berdasarkan jumlah kemunculan istilah bernuansa positif maupun negatif dalam setiap tweet. Apabila frekuensi kosakata bernuansa positif melebihi kosakata bernuansa negatif, tweet dikategorikan sebagai sentimen positif. Sebaliknya, jika kata negatif lebih dominan, tweet diberi label "negatif"; dan jika keduanya seimbang atau tidak ditemukan kata dari kedua kamus tersebut, maka tweet diberi label "netral".

Tabel 4. Hasil Pelabelan

No.	Tweet	Label
1.	dengerin nih guess kata neneng jebred yg sering nyinyir dan minta progrm makan bergizi gratis dihentikn pdahal jlas program mbg itu memberikn gizi untuk mencerdaskan generasi mari bersama dukung lanjutkan mbg mbgmembangunsdmunggul amanda zahra villains parkinson trump is 6	Positif
2.	program makan bergizi gratis terbukti membantu pemerintah dalam meningkatkan pemerataan dan pemenuhan gizi anakanak di seluruh wilayah indonesia makanbergizigratis gizianakmerata indonesia sehat	Positif
...
562.	makan bergizi gratis mampu meningkatkan gizi masyarakat mencegah stunting dan menumbuhkan ekonomi kerakyatan dukung terus program mbg demi indonesia maju faktaindonesia makanbergizigratis manfaatmakanbergizigratis	Positif
563.	program makan bergizi gratis mbg selain untuk anak sekolah juga dibagikan ke ibu hamil menyusui amp kelompok bayi dibawah usia 5 th hal ini bertujuan mencukupi kebutuhan gizi anak yg dibutuhkan sejak periode golden age yakni 1000 hari pertama sejak lahir lanjutkan mbg	Positif

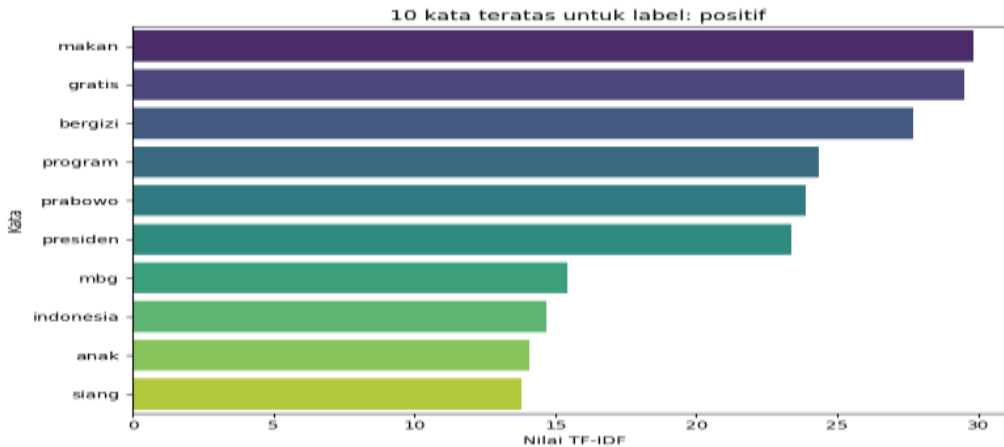
Tabel 4 menampilkan hasil pelabelan terhadap sejumlah tweet yang dianalisis. Setiap tweet diberikan label sentimen berdasarkan dominasi kata-kata bernuansa positif atau negatif yang terkandung di dalamnya. Sebagian

besar tweet dalam tabel tersebut menunjukkan dukungan terhadap program Makan Bergizi Gratis, sehingga memperoleh label Positif.

3.5 Ekstraksi Fitur

Ekstraksi fitur dilakukan menggunakan pendekatan berbasis frekuensi kata, yang menekankan pada kosakata dengan tingkat kemunculan tinggi pada tweet di tiap kelas sentimen. Ekstraksi dilakukan secara terpisah untuk setiap label sentimen.

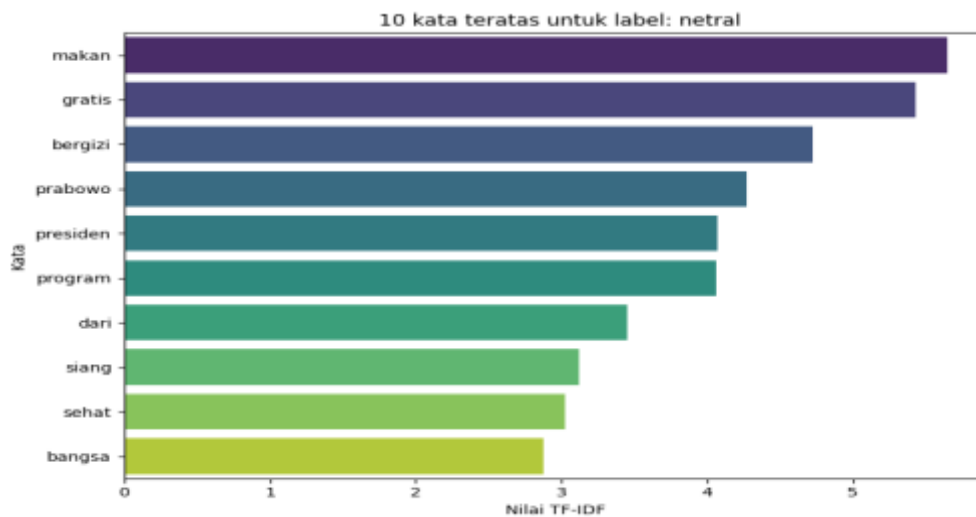
3.5.1. Sentimen Positif



Gambar 2. TF-Idf Sentimen Positif

Berdasarkan Gambar 2, terlihat bahwa kata-kata seperti "makan", "gratis", "bergizi", dan "program" memiliki skor TF-IDF tertinggi untuk label positif. Temuan tersebut mengindikasikan bahwa akun Twitter dengan kecenderungan sentimen positif mengenai program makan bergizi umumnya menyoroti elemen-elemen tersebut secara eksplisit. Kata "makanbergizigratis" dan "dukung" juga muncul sebagai kata kunci yang relevan, menandakan adanya dukungan nyata terhadap program yang digagas pemerintah.

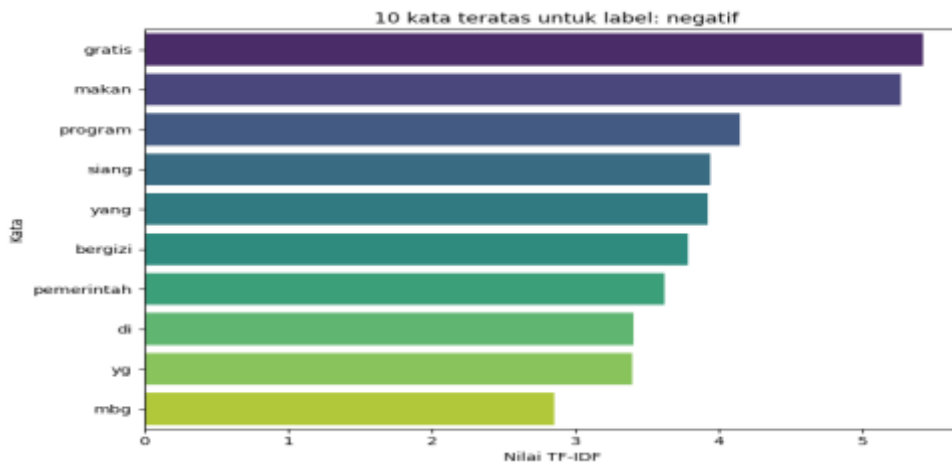
3.5.2. Sentimen Netral



Gambar 3. TF-idf Sentimen Netral

Pada gambar 3 kategori netral, kata-kata seperti "makan", "siang", "gratis", "program", dan "bergizi" tetap muncul dengan skor TF-IDF yang tinggi. Namun, tidak ditemukan kata-kata bernuansa emosional atau penilaian subjektif yang kuat. Kemunculan kata seperti "yg", "pendidikan", dan "generasi" mengindikasikan bahwa tweet-tweet tersebut lebih banyak bersifat informatif atau netral dalam menyampaikan pendapat.

3.5.3. Sentimen Negatif

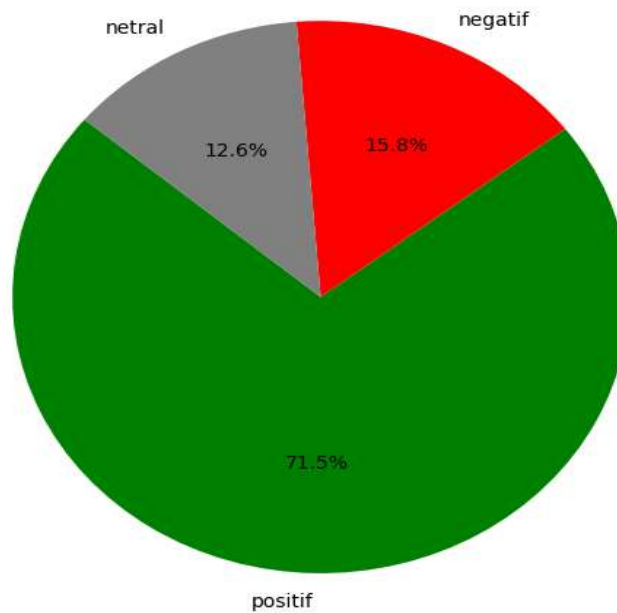


Gambar 4. Tf-idf Sentimen Negatif

Untuk gambar 4 dengan kategori negatif, meskipun kata "gratis", "makan", dan "siang" masih muncul sebagai kata dominan, konteks penggunaannya cenderung mengandung kritik atau keraguan terhadap program. Kata-kata seperti "pilih", "ya", dan "orang" juga muncul, yang mungkin terkait dengan kritik terhadap motif politis di balik pelaksanaan program tersebut.

3.6 Visualisasi Data

3.6.1. Pie Chart

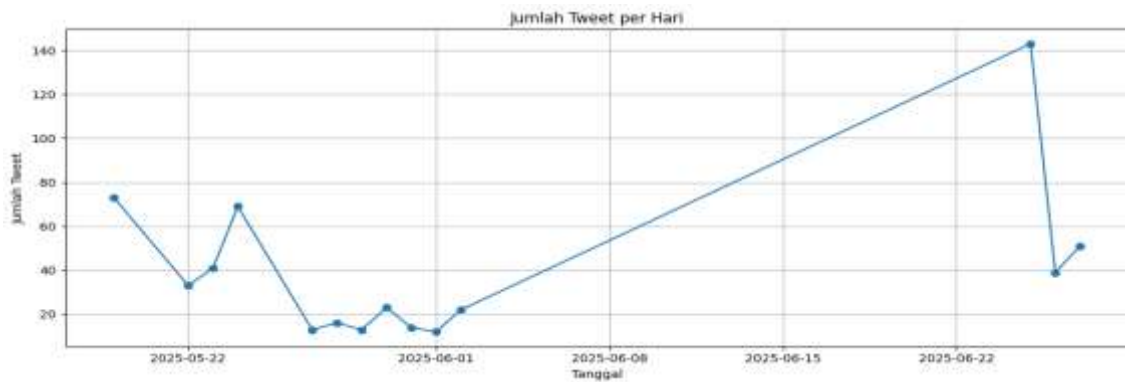


Gambar 5. Pie Chart

Berdasarkan gambar 5, mayoritas tweet masyarakat menunjukkan sentimen positif terhadap program makan bergizi gratis, yaitu sebesar 71.5%. Sementara itu, sentimen negatif tercatat sebesar 15.8%, dan sentimen netral sebesar 12.6%. Hal ini menunjukkan bahwa secara umum, respons publik di Twitter cenderung mendukung program tersebut, meskipun masih terdapat sebagian kecil pengguna yang menyampaikan kritik atau bersikap netral.

konteks penggunaannya dalam sentimen negatif cenderung mengandung kritik, ketidakpercayaan, atau sindiran terhadap program. Kemunculan kata seperti “yg”, “ga”, dan “cuma” mengindikasikan adanya nada skeptis atau ketidakpuasan dari sebagian pengguna Twitter.

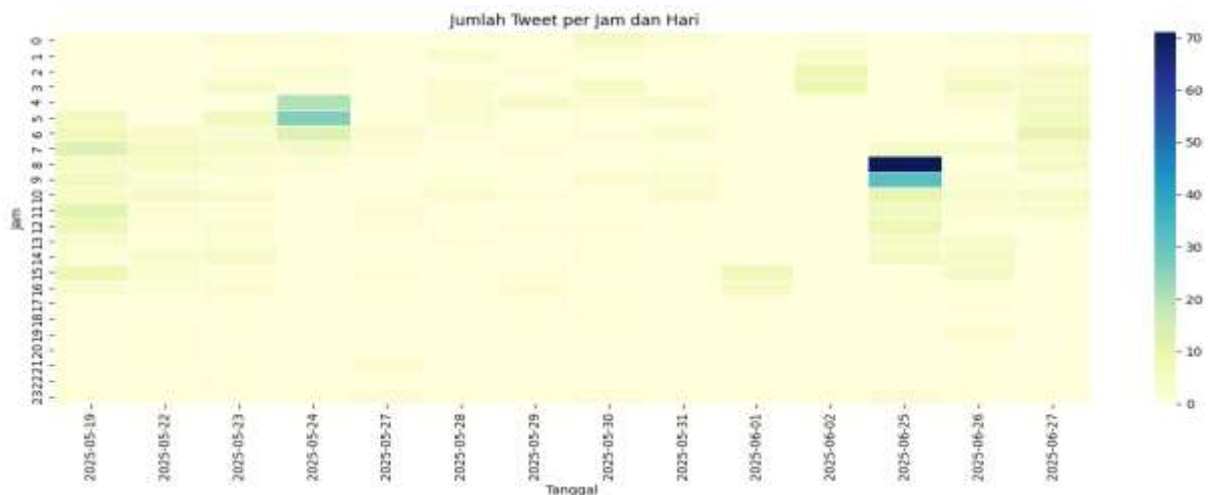
3.6.5. Grafik Waktu (per hari)



Gambar 9. Grafik Waktu per Hari

Grafik jumlah tweet per hari menunjukkan percakapan publik mengalami lonjakan pada tanggal-tanggal tertentu, seperti 19 Mei dan 23 Mei 2025. Hal ini kemungkinan dipicu oleh peristiwa atau pengumuman pemerintah yang relevan.

3.6.6. Grafik Waktu (per jam dan hari)



Gambar 10. Grafik per Jam dan Hari

heatmap distribusi tweet berdasarkan jam dan tanggal menunjukkan bahwa tweet paling sering diposting pada rentang pukul 6 hingga 10 pagi, terutama pada tanggal-tanggal dengan volume tinggi. Pola ini memberikan wawasan penting mengenai waktu-waktu puncak diskusi publik di media sosial yang dapat dimanfaatkan dalam strategi komunikasi kebijakan.

3.5.2 Evaluasi Kinerja Model SVM

Proses evaluasi diterapkan guna menilai kinerja algoritma klasifikasi Support Vector Machine dalam mengidentifikasi opini publik mengenai kebijakan makan bergizi tanpa biaya. Output evaluasi ditampilkan melalui classification report serta confusion matrix guna memberikan gambaran akurasi, presisi, recall, serta distribusi klasifikasi antar kelas sentimen.

Tabel 5. Klasifikasi

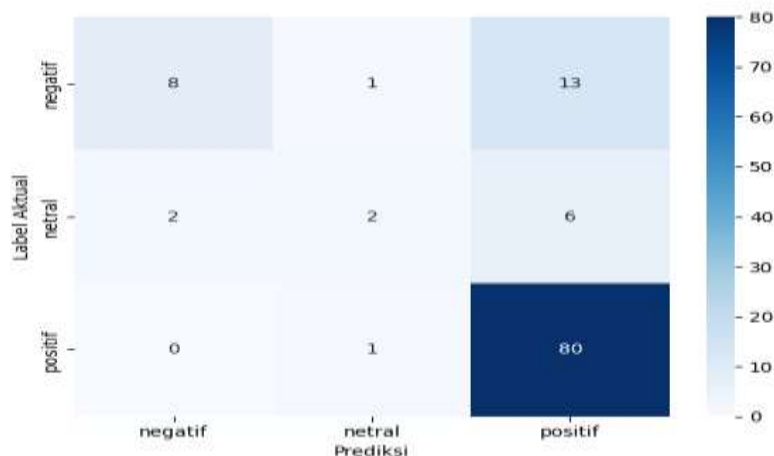
Label sentimen	Presisi	Recall	F1-Score	Support
Negatif	0.80	0.36	0.50	22
Netral	0.50	0.20	0.28	10
Positif	0.80	0.98	0.88	81

Tabel 5 menggambarkan performa model SVM dalam mengklasifikasikan polaritas negatif, netral, dan positif pada tweet hasil perolehan. Dalam kategori sentimen negatif, algoritma memberikan nilai presisi sebesar 0.80, recall sebesar 0.36, serta F1-score sebesar 0.50. Artinya, model cukup baik dalam mengenali tweet bernuansa negatif, meskipun masih terdapat kesalahan dalam mendeteksi semua tweet negatif secara menyeluruh, terlihat dari nilai recall yang rendah. Untuk kategori netral, performa model terlihat masih kurang optimal dengan presisi 0.50, recall 0.20, dan F1-score 0.28. Temuan ini memperlihatkan model masih kesulitan dalam membedakan tweet netral dari kategori lainnya. Sementara itu, performa tertinggi ditunjukkan pada kategori sentimen positif dengan presisi 0.80, recall 0.98, dan F1-score 0.88. Hasil tersebut mengindikasikan kemampuan algoritme dalam mendeteksi serta mengelompokkan tweet positif dengan sangat baik, meskipun masih ada potensi kesalahan kecil dalam mengidentifikasi keseluruhan data. Secara keseluruhan, performa model dapat dikatakan baik, terutama dalam mengenali sentimen positif, namun masih perlu peningkatan dalam menangani kategori netral.

Tabel 6. Matrik Evaluasi

Matrik Evaluasi	Nilai
Akurasi	0.80
Macro Avg F1	0.56
Weighted Avg F1	0.76

Berdasarkan Tabel 6, dapat dilihat bahwa sistem SVM mampu bekerja dengan baik pada tugas klasifikasi sentimen tweet terhadap tweet. Nilai akurasi sebesar 0,80 yang mengindikasikan bahwa 80% prediksi yang dilakukan oleh model sesuai dengan label aslinya. Selain itu, nilai *Macro Average F1* dan *Weighted Average F1* masing-masing sebesar 0,56 dan 0,76 juga memperkuat performa model secara keseluruhan. Nilai *Macro Average F1* mencerminkan rata-rata performa model dalam mengenali semua kelas secara seimbang tanpa mempertimbangkan jumlah data di masing-masing kelas, sementara *Weighted Average F1* mempertimbangkan proporsi jumlah data pada setiap kelas, sehingga memberikan gambaran performa yang lebih realistis perihal distribusi data dengan ketidakseimbangan antar kelas. Konsistensi dari ketiga metrik ini menandakan bahwa sistem tidak hanya mencatat akurasi tinggi, tetapi juga mampu mendeteksi opini positif, netral, dan negatif secara proporsional dan adil. Dengan demikian, model SVM ini dapat diandalkan untuk analisis sentimen terhadap data Twitter secara efektif.



Gambar 11. Matrik Konfusi

Berdasarkan gambar 10, menunjukkan performa yang cukup baik dalam mengenali dan memisahkan setiap kategori sentimen. Hal ini terlihat dari banyaknya prediksi yang benar yang berada di sepanjang garis diagonal matriks, yaitu sebanyak 8 data dengan label negatif, 2 data netral, dan 80 data positif yang berhasil diklasifikasikan dengan tepat sesuai label aslinya. Kesalahan klasifikasi relatif lebih banyak terjadi pada kelas negatif dan netral, seperti 13 data negatif yang keliru dikategorikan ke positif serta 6 data netral yang juga salah terdeteksi pada label positif. Sementara itu, kelas positif dapat dikenali dengan sangat baik karena hanya terdapat 1 data positif yang

salah diprediksi sebagai netral. Hasil ini mengindikasikan bahwa model mampu membedakan karakteristik dari kelas positif dengan sangat baik, namun masih mengalami kesulitan dalam mengenali kelas netral dan sebagian negatif. Confusion matrix ini memperkuat kesimpulan bahwa model lebih konsisten dalam mengenali pola data berlabel positif, tetapi masih perlu perbaikan agar kinerja pada kelas negatif dan netral dapat meningkat.

4. Kesimpulan

Berdasarkan hasil penelitian, model *Support Vector Machine (SVM)* berhasil diterapkan secara efektif untuk mengklasifikasikan opini publik mengenai Program Makan Bergizi Gratis yang dianalisis lewat data Twitter. Model menunjukkan performa yang baik dalam mengenali sentimen positif dengan nilai F1-score tertinggi, sedangkan klasifikasi untuk sentimen netral dan negatif masih menunjukkan kelemahan, khususnya dalam hal recall dan ketepatan identifikasi. Proses analisis yang dilakukan meliputi pengumpulan data, preprocessing, dan evaluasi model yang menunjukkan kemampuan pendekatan yang digunakan mampu menangkap persepsi publik secara umum. Namun, distribusi data yang tidak seimbang antar kategori sentimen menjadi salah satu tantangan utama yang memengaruhi hasil klasifikasi. Untuk mengatasi keterbatasan tersebut, disarankan agar penelitian selanjutnya melakukan pelabelan data secara manual guna meningkatkan kualitas data latih serta mempertimbangkan penggunaan algoritma lain atau pendekatan berbasis *deep learning* agar klasifikasi sentimen, terutama kategori netral dan negatif, dapat lebih optimal dan akurat.

Referensi

- [1] A. Kiftiyah, F. A. Palestina, F. U. Abshar, and K. Rofiah, "Program Makan Bergizi Gratis (MBG) dalam Perspektif Keadilan Sosial dan Dinamika Sosial – Politik," *Pancasila J. Keindonesiaan*, vol. 5, no. 1, pp. 101–112, 2025, doi: 10.52738/pjk.v5i1.726.
- [2] I. Irayana and I. Yarliani, "Intervensi gizi, sanitasi, dan kesehatan bagi anak jalanan melalui pelibatan partisipasi masyarakat," *J. Warn. Pendidik. dan Pembelajaran Anak Usia Dini*, vol. 7, no. 2, pp. 108–129, 2022, doi: 10.24903/jw.v7i2.1619.
- [3] F. S. Pratiwi, "Peran Komunikasi Digital dalam Pembentukan Opini Publik : Studi Kasus Media Sosial," pp. 293–315, 2024, doi: 10.30589/proceedings.2024.1059.
- [4] E. Yudianto, I. Suchayo, and U. P. Marga, "Persepsi Masyarakat Mengenai Program Bantuan Sembako di Kecamatan Balongpanggang Kabupaten Gresik," vol. 5, pp. 3819–3827, 2025, doi: 10.31004/innovative.v5i2.18805.
- [5] M. D. Febrian, T. R. Munjazi, and M. I. Al-mubarak, "Partisipasi Aktor Politik Dalam Wacana Makan Bergizi Gratis: Tinjauan Literatur Atas Kepentingan Politik," *J. Media Akad.*, vol. 3, no. 6, 2025, doi: 10.62281/v3i6.2338.
- [6] B. A. Maulana, M. J. Fahmi, A. M. Imran, and N. Hidayati, "Analisis Sentimen Terhadap Aplikasi Pluang Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM)," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 2, pp. 375–384, 2024, doi: 10.57152/malcom.v4i2.1206.
- [7] A. Sitanggang, Y. Umaidah, Y. Umaidah, R. I. Adam, and R. I. 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.
- [8] A. Zidna, M. Ali, M. Sri, and S. Joko Dwi, "Penerapan Metode Support Vector Machine Untuk Analisis Sentimen Pengguna Twitter," *J. Inf. Syst. Manag.*, vol. 3, no. 1, pp. 16–21, 2021, doi: 10.24076/joism.2021v3i2.558.
- [9] S. Rabbani, D. Safitri, N. Rahmadhani, A. A. F. Sani, and M. K. Anam, "Perbandingan Evaluasi Kernel SVM untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 3, no. 2, pp. 153–160, 2023, doi: 10.57152/malcom.v3i2.897.
- [10] S. M. Natzir, "Perbandingan Kinerja Model Pembelajaran Mesin Dalam Prediksi Banjir Menggunakan Knn, Naive Bayes, Dan Random Forest," *HOAQ (High Educ. Organ. Arch. Qual. J. Teknol. Inf.*, vol. 14, no. 2, pp. 59–64, 2023, doi: 10.52972/hoaq.vol14no2.p59-64.
- [11] A. Cahyana and E. R. Susanto, "Penerapan Algoritma XGBoost untuk Prediksi Diabetes: Analisis Confusion Matrix dan ROC Curve," *Fountain Informatics J.*, vol. 10, no. 1, pp. 2548–5113, 2025, [Online]. Available: <http://dx.doi.org/10.21111/fij.v10i1.14311>
- [12] T. Puspandoyo *et al.*, "Prediksi Kualitas Laporan Keuangan Kementerian Negara / Lembaga Menggunakan," vol. 3, pp. 15–36, 2022, doi: 10.33105/jmp.v3i1.417.
- [13] A. Widyanto, K. Kusriani, and K. Kusnawi, "Pengaruh Keseimbangan Data terhadap Akurasi Model Support Vector Machine pada Data Set Donor Darah," *J. Teknol. Terpadu*, vol. 9, no. 2, pp. 79–88, 2023, doi: 10.54914/jtt.v9i2.771.
- [14] I. N. Husada and H. Toba, "Pengaruh Metode Penyeimbangan Kelas Terhadap Tingkat Akurasi Analisis Sentimen pada Tweets Berbahasa Indonesia," *J. Tek. Inform. dan Sist. Inf.*, vol. 6, no. 2, pp. 400–413, 2020, doi: 10.28932/jutisi.v6i2.2743.
- [15] V. No, A. Maizaliyanti, K. Umam, W. D. Yuniarti, and M. R. Handayani, "Edumatic : Jurnal Pendidikan Informatika Implementasi Algoritma Random Forest dalam Klasifikasi Ulasan Pengunjung Mall Semarang untuk Pengambilan Keputusan Layanan," vol. 9, no. 2, pp. 452–461, 2025, doi: 10.29408/edumatic.v9i2.30379.