Jurnal Ilmiah Sistem Informasi dan Ilmu Komputer
Volume 5, Nomor 3, November 2025
p-ISSN: 2827-8135; e-ISSN: 2827-7953, Hal. 308-318
DOI: https://doi.org/10.55606/juisik.v5i3.1665
Tersedia: https://journal.sinov.id/index.php/juisik

Analisis Sentimen Media Sosial X Program Makanan Sehat Gratis dengan Support Vector Machine dan Naive Bayes

Jack Vallen[1*], Anik Hanifatul Azizah [2]
[1,2] Universitas Esa Unggul
**Penulis Korespondensi: Jckvallen@student.esaunggul.ac.id [1]*

Abstract*. The free nutritious meal program, has sparked various reactions from the public, particularly on social media. This study aims to determine how netizens feel about the program through data analysis from social media platform X (formerly Twitter). Data was collected using web scraping techniques with the keyword "free nutritious meals," then processed through text cleaning and automatic labeling stages using the Indonesian language version of the RoBERTa model. With this approach, each tweet was efficiently and accurately classified into positive or negative sentiment. The labeled data was then analyzed using two classification algorithms: Support Vector Machine (SVM) and Naïve Bayes. Test results showed that SVM performed better, with an average accuracy of 0.8367 and a deviation of 0.0117. Meanwhile, Naïve Bayes recorded an accuracy of 0.7716 with a deviation of 0.0101. Visualization through WordCloud also shows the dominant words in each sentiment. Words such as "support," "healthy," and 'grow' appear frequently in positive sentiments, while words such as "budget," "poison," and "cost" dominate negative sentiments. These findings illustrate significant public support for the program, but also concerns regarding its implementation and funding. The results of this analysis are expected to provide input for policymakers in understanding public opinion more objectively.*

Keywords*: Free Nutritious Meals; Naïve Bayes; Sentiment Analysis; Social Media; SVM.*

Abstrak. Program makanan bergizi gratis telah memicu berbagai reaksi dari masyarakat, terutama di media sosial. Studi ini lakukan untuk menentukan bagaimana netizen merasa tentang program tersebut melalui analisis data dari platform media sosial X (Twitter). Data dikumpulkan menggunakan teknik web scraping dengan kata kunci "makanan bergizi gratis", kemudian diproses melalui tahap pembersihan teks dan penandaan otomatis menggunakan versi bahasa Indonesia dari model RoBERTa. Dengan pendekatan ini, setiap tweet diklasifikasikan secara efisien dan akurat ke dalam sentimen positif atau negatif. Data yang telah ditandai kemudian dianalisis menggunakan dua algoritma klasifikasi: Support Vector Machine (SVM) dan Naïve Bayes. Hasil uji menunjukkan bahwa SVM berkinerja lebih baik, dengan akurasi rata-rata 0,8367 dan deviasi 0,0117. Sementara itu, Naïve Bayes mencatat akurasi 0,7716 dengan deviasi 0,0101. Visualisasi melalui WordCloud juga menunjukkan kata-kata dominan dalam masing-masing sentimen. Kata-kata seperti "dukungan," "sehat," dan "tumbuh" sering muncul dalam sentimen positif, sementara kata-kata seperti "anggaran," "racun," dan "biaya" mendominasi sentimen negatif. Temuan ini menggambarkan dukungan publik yang signifikan terhadap program tersebut, namun juga kekhawatiran terkait implementasi dan pendanaannya. Hasil analisis ini diharapkan dapat memberikan masukan bagi pembuat kebijakan dalam memahami opini publik secara lebih objektif.

Kata kunci: Analisis Sentimen; Makanan Bergizi Gratis; Media Sosial; Naïve Bayes; SVM.

1. LATAR BELAKANG

Di era digital ini, penyebaran informasi telah menjadi lebih mudah diakses di berbagai tempat. Informasi dapat diakses di mana saja, di rumah, saat bepergian, dan di mana pun. Salah satu sumber informasi paling banyak saat ini adalah media sosial, di mana orang dapat dengan mudah mengakses dan berbagi informasi dengan pengguna lain dalam bentuk video, audio, teks, atau gambar. Salah satu contoh penggunaan media sosial adalah ketika orang berbagi pendapat mereka tentang suatu produk, berita, atau individu lain.

Naskah Masuk: 30 Agustus 2025; Revisi: 14 September 2025; Diterima: 18 Oktober 2025; Tersedia: 24 Oktober 2025

Media sosial secara luas digunakan sebagai platform bagi orang untuk berbagi pendapat tentang berbagai topik. Pendapat dapat mengambil berbagai bentuk, mulai dari kritik, pujian, berita palsu, hingga ujaran kebencian, yang dapat memicu perdebatan di antara pengguna media sosial, Di Indonesia sendiri media sosial dapat berkembang dengan sangat cepat Dimana pengguna aktif sosial media di Indonesia sudah mencapai 139 juta pengguna pada Januari 2024 (Simon Kemp, 2024).

Salah satu platform media sosial yang banyak digunakan untuk mengekspresikan pendapat adalah Twitter (X), merupakan sosial media yang menjadi pusat trending di Indonesia dikarenakan cepatnya penyebaran informasi dan respon dari berbagai sudut pandang (Rizki, Auliasari, & Prasetya, 2021).

Salah satu topik pembicaraan yang hangat adalah program makanan bergizi gratis yang diusulkan oleh pemerintahan Prabowo-Gibran, yang merupakan salah satu tujuan utama Presiden Prabowo dan Gibran untuk mencapai Indonesia Emas 2025 (Sugiarto, 2025). Program makanan bergizi gratis ini merupakan contoh konkret dari upaya pemerintah untuk meningkatkan kualitas gizi masyarakat Indonesia, Program ini ditujukan untuk bayi atau anak di bawah lima tahun, wanita hamil, ibu menyusui, dan siswa di berbagai tingkatan pendidikan (Saptati, 2025).

Pada tanggal 6 Januari 2025, Program Makanan Bergizi Gratis secara resmi diluncurkan, dimulai secara bertahap dari setiap tingkatan pendidikan, mulai dari taman kanak-kanak hingga sekolah menengah atas (Muallifa, 2025). Meskipun program ini memiliki tujuan mulia, program ini masih menimbulkan reaksi yang beragam dari masyarakat. Beberapa mendukung program ini karena mereka percaya program ini akan meningkatkan kesejahteraan masyarakat dan membantu pertumbuhan gizi anak-anak, namun banyak juga yang mengkritik program ini karena khawatir program ini akan manfaatkan oleh pihak yang tidak bertanggung jawab.

Fenomena ini menunjukkan betapa bermanfaatnya analisis sentimen dalam mencari tahu dan memahami pola respons masyarakat terhadap program makanan bergizi gratis. Analisis sentimen merupakan proses mengidentifikasi dan mengelompokkan pendapat yang diungkapkan dalam teks serta mengklasifikasikannya. Selain itu, pemangku kepentingan dapat menggunakan analisis sentimen untuk memahami persepsi masyarakat terhadap isu-isu tertentu, yang dapat dimanfaatkan untuk membantu mengembangkan kebijakan atau keputusan yang lebih selaras dengan kebutuhan masyarakat (Natasuwarna, 2020).

Dalam studi ini, data akan dikumpulkan menggunakan metode Web Scraping, yang dipilih karena web scraping memungkinkan pengumpulan data tanpa perlu memperoleh API dari situs web (Sembiring & Sari, 2019).

Untuk pelabelan, versi Indonesia dari model RoBERTa, yang merupakan model transformer yang telah dilatih sebelumnya, akan digunakan. Model ini dipilih untuk melabeli dataset (Atmajaya, Febrianti, & Darwis, 2023). Analisis sentimen akan dilakukan menggunakan algoritma SVM (Support Vector Machine) dan Naïve Bayes untuk menentukan akurasi masing-masing algoritma.

2. KAJIAN TEORITIS

Penelitian ini menggunakan data yang diperoleh melalui web scraping berupa tweet publik dalam bahasa Indonesia yang diambil dari Twitter (X) dengan kata kunci "makanan bergizi gratis". Data yang digunakan dikumpulkan dari tanggal 6 Januari hingga 30 Juni 2025, dan akan diklasifikasikan menggunakan dua jenis algoritma analisis sentimen, yaitu Naïve Bayes dan SVM (Support Vector Machine).

Analisis Sentimen

Analisis sentimen merupakan salah satu teknik dalam Pemrosesan Bahasa Alami (NLP), yang merupakan bidang studi yang mengeksplorasi cara komputer dapat bekerja dan berpikir seperti manusia (Salsabila, 2022). Salah satu fungsi analisis sentimen adalah untuk menentukan apakah isi suatu pendapat memiliki makna negatif atau positif.

Analisis sentimen dilakukan dengan menganalisis pendapat, sikap, atau evaluasi publik terhadap topik, produk, atau layanan tertentu (Rusdiaman & Rosiyadi, 2019).

SVM (Support Vector Machine)

Support Vector Machine atau SVM adalah salah satu metode pembelajaran mesin terawasi yang sering digunakan dalam klasifikasi dan regresi. Metode ini digunakan dalam klasifikasi, yang merupakan bagian dari pembelajaran terawasi dan memiliki konsep yang lebih jelas. Dibandingkan dengan metode lain, metode SVM (Support Vector Machine) memiliki tingkat keakuratan yang lebih tinggi (Pratiwi, Febbi, Af'idah, & Rifa, 2021).

Naïve Bayes

Naive Bayes merupakan algoritma yang digunakan dalam *Teks Mining*. Algoritma Naive Bayes menggunakan teori probabilitas untuk menemukan probabilitas klasifikasi tertinggi dengan memeriksa frekuensi klasifikasi dalam data pelatihan. Keuntungan dari algoritma ini adalah Naive Bayes hanya memerlukan jumlah data yang sedikit sebagai data pelatihan (Trisna, 2023).

Word Cloud

Penelitian ini akan menggunakan word clouds, yang merupakan metode yang dapat digunakan untuk memberikan representasi visual dari kumpulan kata yang digunakan atau untuk memvisualisasikan kata-kata dominan yang muncul dalam data yang digunakan (Ibrahim, Bakar, Harun, & Abdulateef, 2021).

Web Scraping

Pada tahap pengumpulan data, Perpustakaan Selenium akan digunakan. Selenium adalah perpustakaan Python yang dapat melakukan fungsi pengujian otomatis situs web. Salah satu fungsi ini adalah web scraping, yang dapat digunakan untuk mengambil data dari situs web seperti situs media sosial (Yondra, Triyanto, & Bahri, 2022), Metode ini dilakukan dengan cara mengekstrak data dari website dan menyimpannya dalam bentuk CSV atau *spreadsheet* (Tarale, 2025).

Text - Preprocessing

Tahap ini melibatkan pembersihan data untuk memastikan bahwa data yang akan digunakan sesuai dengan persyaratan model sehingga tidak menimbulkan kebingungan selama tahap pelabelan. Prasunting teks meliputi case folding, tokenisasi, normalisasi, dan stemming. Tahap ini dilakukan untuk mengetahui bahwa data yang digunakan sesuai dengan kebutuhan model dan untuk menghilangkan semua jenis teks yang tidak sesuai dengan persyaratan tahap prasanting teks, termasuk case folding, pembersihan, stopwords, stemming, dan tokenisasi (Isnain, Sakti, Alita, & Marga, 2021).

Labelling

Pada tahap ini, pelabelan akan dilakukan pada data yang telah melalui tahap Pra-pemrosesan Teks. Pada tahap ini, data akan dilabeli menggunakan versi Indonesia dari model RoBERTa, yang merupakan model transformer yang telah dilatih sebelumnya. Model ini dipilih untuk melakukan pelabelan pada dataset (Atmajaya, Febrianti, & Darwis, 2023) dan dibagi menjadi beberapa kelas atau label. Dalam studi ini, sentimen positif dan negatif akan digunakan, sementara sentimen netral akan dihilangkan.

Confusion Matrix

Setelah klasifikasi menggunakan algoritma Naïve Bayes dan SVM, data akan diperiksa akurasinya untuk menentukan seberapa akurat masing-masing model dalam mengklasifikasikan data yang diberikan. Metode yang digunakan untuk memeriksa akurasi adalah Matriks Kebingungan. Dimana Matriks Kebingungan mengukur performa dari model dengan membandingkan nilai yang sebenarnya dengan nilai prediksi, Dimana *output* yang dihasilkan akan berupa matriks yang menunjukan jumlah prediksi yang benar dan salah. dalam

confusion matrix terdapat beberapa matriks yang digunakan untuk mengambarkan hasil dari klasifikasi yaitu: (Zendrato A et al., 2024).

a. *True* positif: data positif yang diprediksi Negatif
b. *True* negatif: data negatif yang diprediksi Positif
c. *False* positif: data negatif tapi hasil prediksi positif
d. *False* negatif: data positif tapi hasil prediksi negatif

Dengan metode ini, akurasi data yang diklasifikasikan oleh algoritma sebelumnya akan diperiksa. Metrik yang diukur dalam metode ini meliputi akurasi, presisi, recall, dan F1-score (Nugroho, 2025).

K-Fold Cross Validation

Dalam studi ini, metode Validasi Silang K-Fold 10 digunakan untuk menguji validitas model yang digunakan. Teknik ini dipilih karena memberikan evaluasi yang lebih komprehensif dan akurat dibandingkan dengan metode pembagian data sederhana. Uji validasi juga dilakukan untuk memastikan bahwa model yang digunakan tidak mengalami overfitting (Rangga & Hayaty, 2019). Selain memastikan bahwa model tidak mengalami overfitting, 10 K-Fold Cross Validation juga dilakukan untuk memastikan bahwa model dapat melakukan klasifikasi secara konsisten.

Roberta

RoBERTa merupakan model *pre trained* yang berbasis arsitektur transformer yang mampu memahami konteks bahasa alami secara mendalam. RoBERTa merupakan suatu peningkatan dari model BERT yang di kembangkan oleh Facebook AI *Research* dan dapat membantu dalam pemahaman bahasa (Atmajaya et al., 2023).

3. METODE PENELITIAN

Pada penelitian ini akan melakukan pengambilan data dari media sosial Twitter (X) menggunakan teknik *web scraping*, dengan kata kunci “makan bergizi gratis” dan filter bahasa Indonesia. Setiap permintaan dibatasi maksimal 2000 tweet, namun jumlah tweet per bulan bervariasi tergantung aktivitas pengguna. Data yang diperoleh kemudian diproses melalui tahapan *preprocessing* seperti *cleaning, case folding*, normalisasi, *stopword*, *stemming*, dan tokenisasi. Setelah itu, data dilabeli secara otomatis menggunakan model RoBERTa ke dalam dua jenis sentimen, yaitu positif dan negatif. Proses klasifikasi dilakukan menggunakan algoritma *Support Vector Machine* (SVM) dan *Naive bayes*, lalu dievaluasi menggunakan metrik akurasi, *Precision*, *recall, F1-score*, serta divalidasi dengan 10 *K-Fold Cross Validation*

4. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan data yang diperoleh melalui web scraping, yaitu data yang diambil atau dikumpulkan dari platform media sosial Twitter (X) berbahasa Indonesia yang mengandung kata kunci "makanan bergizi gratis", menghasilkan 12.767 titik data. Setelah data berhasil dikumpulkan, langkah selanjutnya adalah melakukan prapemrosesan teks pada data yang ada untuk memastikan format data sesuai dan memudahkan proses pelabelan.

Table 1. Text Pre-processing.

Tweet	Text Preprocessing Result
Program Makan Bergizi Gratis bantu anak Indonesia siap hadapi masa depan! Gizi seimbang untuk prestasi cemerlang sejak dini #MakanBergiziGratis	['program', 'gizi', 'bantu', 'anak', 'indonesia', 'hadap', 'imbang', 'prestasi', 'cemerlang']

Selanjutnya, proses labelling data dilakukan menggunakan model RoBERTa dengan tujuan mempercepat dan menyederhanakan proses klasifikasi sentimen. Model ini dipilih karena telah dilatih sebelumnya menggunakan data bahasa Indonesia, sehingga mampu memahami makna dan konteks kalimat dengan lebih mendalam. Dengan bantuan RoBERTa, proses pelabelan dapat dilakukan secara otomatis tanpa perlu melalui langkah-langkah manual yang biasanya memakan waktu lama dan berisiko menimbulkan bias subjektif dari para pelabel. Model ini akan mengkategorikan setiap tweet ke dalam sentimen positif, atau negatif berdasarkan pola bahasa yang telah dipelajarinya. Penggunaan RoBERTa diharapkan dapat menghasilkan pelabelan yang lebih akurat dan konsisten, serta menjadi landasan yang kuat untuk tahap pemodelan berikutnya.

Table 2 Labelling

Tweet	*labeling*
['program', 'gizi', 'bantu', 'anak', 'indonesia', 'hadap', 'imbang', 'prestasi', 'cemerlang']	positive
['budget', 'gizi', 'anak', 'harga', 'air', 'mineral', 'jamu']	negative

Setelah proses penandaan selesai, langkah berikutnya adalah memvisualisasikan keberadaan kata-kata dominan dalam setiap sentimen. Hal ini dilakukan untuk mengidentifikasi kata-kata yang paling menonjol dalam sentimen tersebut. Dalam melakukan visualisasi, perpustakaan Word Cloud akan digunakan untuk memvisualisasikan kata-kata dominan yang muncul dalam data yang digunakan (Ibrahim et al., 2021).

Hasil visualisasi menggunakan Wordcloud untuk menampilkan kata-kata dominan yang paling banyak muncul dari setiap sentimen setelah pelabelan menggunakan RoBERTa ditampilkan pada gambar di bawah ini untuk sentimen positif dan negatif.

Gambar 1 WordCloud Negative

Gambar 2 WordCloud Positive

Gambar 3.1 dan 3.2 menampilkan visualisasi WordCloud yang memvisualisasikan kata-kata paling dominan yang muncul dalam setiap jenis sentimen, yaitu negatif dan positif. Pada Gambar 3.1 (WordCloud Sentimen Negatif), kata-kata yang paling menonjol meliputi “racun,” “anggaran,” “korupsi,” “hentikan,” “laporan,” dan beberapa kata lain yang mencerminkan kekhawatiran, kritik, atau penolakan terhadap program makanan bergizi gratis. Kata-kata ini menunjukkan persepsi publik yang mengaitkan program tersebut dengan masalah negatif seperti penyalahgunaan anggaran atau dugaan korupsi.

Sebaliknya, pada Gambar 3.2 (WordCloud Sentimen Positif), kata-kata seperti “dukungan,” “sehat,” “manfaat,” “tumbuh,” dan “generasi” muncul, menunjukkan bahwa banyak orang juga mendukung program ini dan memandangnya sebagai upaya positif untuk membangun generasi yang sehat dan berkualitas tinggi. Visualisasi ini membantu peneliti memahami konteks opini publik secara lebih intuitif dan berfungsi sebagai pelengkap penting dalam menganalisis konten sentimen yang tidak hanya didasarkan pada angka tetapi juga mempertimbangkan nuansa kata-kata yang digunakan oleh netizen.

Setelah tahap penandaan data selesai, klasifikasi akan dilakukan menggunakan algoritma SVM dan Naïve Bayes, masing-masing menggunakan data yang sama dan akan dievaluasi menggunakan matriks kebingungan untuk menentukan kinerja masing-masing algoritma dalam mengklasifikasikan data.

Table 3. Confusion Matrix.

Model	Precision (Positive)	Precision (Negative)	Recall (Positive)	Recall (Negative)	F1 Score (Positive)	F1 Score (Negative)	Accuracy
SVM	87%	84%	82%	89%	84%	86%	85%
Naïve Bayes	77%	78%	76%	79%	77%	79%	78%

Selanjutnya, akan dilakukan validasi silang K-fold. Metode ini digunakan untuk memastikan bahwa model secara keseluruhan berfungsi dengan baik pada data yang belum pernah dilihat sebelumnya. Metode ini juga digunakan untuk memastikan bahwa model tidak mengalami overfitting dan memberikan pengukuran kinerja yang lebih baik.

Table 4. Hasil K-Fold Cross Validation.

Accuracy of each fold (entire data):			
	Naïve Bayes		SVM (Support Vector Machine)
FOLD	Accuracy of each Fold	FOLD	Accuracy of each Fold
FOLD 1	0.7943	FOLD 1	0.8344
FOLD 2	0.7656	FOLD 2	0.85
FOLD 3	0.7691	FOLD 3	0.8439
FOLD 4	0.7795	FOLD 4	0.8431
FOLD 5	0.77	FOLD 5	0.816
FOLD 6	0.7578	FOLD 6	0.8265
FOLD 7	0.7689	FOLD 7	0.8309
FOLD 8	0.7602	FOLD 8	0.857
FOLD 9	0.7706	FOLD 9	0.8265
FOLD 10	0.7802	FOLD 10	0.8387
Average accuracy	0.7716	Average accuracy	0.8367
Standard deviation	0.0101	Standard deviation	0.0117

Berdasarkan hasil uji validitas menggunakan metode 10 K-Fold Cross Validation, dua hasil rata rata akurasi diperoleh dari dua model yang diuji, yaitu Naïve Bayes dan Support Vector Machine (SVM). Untuk model Naïve Bayes, akurasi rata-rata yang diperoleh dari sepuluh lipatan adalah 0,7716 dengan nilai simpangan baku 0,0101. Hasil ini menunjukkan bahwa model Naïve Bayes memiliki kinerja yang cukup baik dan stabil dalam mengklasifikasikan sentimen, meskipun masih ada ruang untuk perbaikan. Sementara itu, model SVM menunjukkan kinerja yang lebih unggul dengan akurasi rata-rata 0,8367 dan simpangan baku 0,0117, artinya variasi kinerja antar fold sangat kecil. Kedua hasil menunjukkan bahwa baik Naïve Bayes maupun SVM memiliki kemampuan klasifikasi yang konsisten, tetapi SVM

mendapatkan akurasi yang lebih baik dan dapat diandalkan dalam menganalisis sentimen netizen terhadap program makanan bergizi gratis. Teknik validasi silang ini membuktikan bahwa model yang dibangun tidak mengalami overfitting dan dapat digeneralisasikan dengan baik ke data serupa.

5. KESIMPULAN DAN SARAN

Dari hasil klasifikasi, ditemukan bahwa algoritma SVM berkinerja lebih baik daripada Naïve Bayes. Model SVM mencatat akurasi rata-rata 0,8367 dengan standar deviasi 0,0117, sementara model Naïve Bayes memiliki akurasi rata-rata 0,7716 dengan deviasi 0,0101. Perbedaan ini menunjukkan bahwa SVM mampu mengklasifikasikan sentimen dengan lebih akurat dan konsisten. Visualisasi data melalui word cloud juga mengungkapkan perbedaan kata-kata dominan antara sentimen positif dan negatif, mencerminkan opini publik tentang program ini secara lebih mendalam.

Berdasarkan hasil analisis visualisasi WordCloud, gambaran umum persepsi publik terhadap program makan bergizi gratis, khususnya dalam sentimen positif, kata-kata dominan yang muncul adalah “dukung”, “sehat”, “cerdas”, “kuat”, dan “tumbuh”. ” Kemunculan kata-kata ini menunjukkan dukungan dan harapan publik terhadap implementasi program ini, terutama dalam membantu meningkatkan kesehatan dan pertumbuhan anak-anak sebagai sasaran utamanya. Masyarakat tampaknya memandang program ini sebagai langkah positif yang memiliki dampak baik bagi generasi muda. Dan dalam sentimen negatif, kata-kata yang paling dominan yang muncul adalah “anggaran”, “racun”, “biaya”, “tolak”, dan “korupsi”. Dominasi kata-kata ini mencerminkan kekhawatiran dan keraguan masyarakat terkait implementasi program, terutama mengenai masalah anggaran, efektivitas implementasi, dan risiko yang mungkin timbul.

Hal ini menunjukkan bahwa meskipun sebagian anggota masyarakat memberikan dukungan, sebagian lainnya tetap skeptis terhadap kelayakan dan kesiapan program tersebut dari segi teknis maupun pendanaan.

DAFTAR REFERENSI

Atmajaya, D., Febrianti, A., & Darwis, H. (2023). Metode SVM dan Naive Bayes untuk analisis sentimen ChatGPT di Twitter. *Indonesian Journal of Computer Science Attribution, 12*(4), 2173–2181. https://doi.org/10.33022/ijcs.v12i4.3341

Ibrahim, V., Bakar, J. A., Harun, N. H., & Abdulateef, A. F. (2021). A word cloud model based on hate speech in an online social media environment. *Baghdad Science Journal, 18*(2), 937–946. https://doi.org/10.21123/bsj.2021.18.2(Suppl.).0937

Isnain, A. R., Sakti, A. I., Alita, D., & Marga, N. S. (2021). Sentimen analisis publik terhadap kebijakan lockdown pemerintah Jakarta menggunakan algoritma SVM. *Jurnal Data Mining dan Sistem Informasi, 2*(1), 31–37. https://doi.org/10.33365/jdmsi.v2i1.1021

Muallifa, R. N. L. (2025, January 6). Program makan bergizi gratis mulai kapan? Cek mekanismenya. *Liputan6.com*. https://www.liputan6.com/hot/read/5864880/program-makan-bergizi-gratis-mulai-kapan-cek-mekanismenya?page=2

Muallifa, R. N. L., Reni, S. D. I., dan Sugiarto, E. C. sudah termasuk sumber web, sehingga dapat dicatat dengan tanggal akses jika perlu untuk dokumen akademik.

Natasuwarna, A. P. (2020). Seleksi fitur support vector machine pada analisis sentimen keberlanjutan pembelajaran daring. *Techno.Com, 19*(4), 437–448. https://doi.org/10.33633/tc.v19i4.4044

Nugroho, K. S. (2019, November 13). Confusion matrix untuk evaluasi model pada supervised learning. https://ksnugroho.medium.com/confusion-matrix-untuk-evaluasi-model-pada-unsupervised-machine-learning-bc4b1ae9ae3f

Pratiwi, R. W., H, S. F., Af'idah, D. I., A, Q. R., & F, A. G. (2021). Analisis sentimen pada review skincare Female Daily menggunakan metode support vector machine (SVM). *Journal of Informatics, Information System, Software Engineering and Applications, 4*(1), 40–46. https://doi.org/10.20895/inista.v4i1.387

Rangga, N., & Hayaty, M. (2019). Perbandingan akurasi dan waktu proses algoritma K-NN dan SVM dalam analisis sentimen Twitter. *Jurnal Informatika, 6*(2), 212–218. https://doi.org/10.31311/ji.v6i2.5129

Reni Saptati, D. I. (2025). Pemerintah salurkan makan bergizi gratis (MBG), ini sasaran utama penerimanya. *Media Keuangan*. https://mediakeuangan.kemenkeu.go.id/article/show/pemerintah-salurkan-makan-bergizi-gratis-mbg-ini-sasaran-utama-penerimanya

Rusdiaman, D., & Rosiyadi, D. (2019). Analisa sentimen terhadap tokoh publik menggunakan metode Naïve Bayes Classifier dan Support Vector Machine. *CESS (Journal of Computer Engineering System and Science), 4*(2), 230–235. https://doi.org/10.24114/cess.v4i2.13796

Salsabila, N. A. (2022). Skripsi analisis sentimen pada media sosial Twitter terhadap tokoh Gus Dur menggunakan metode Naïve Bayes dan support vector machine (SVM).

Sembiring, F., & Sari, D. P. (2019). Design process data storage and organize data scraping. *Integrated (Information Technology and Vocational Education), 1*(1), 19–22. https://doi.org/10.17509/integrated.v1i1.16837

Simon Kemp. (2024, February 21). Digital 2024: Indonesia. *Datareportal - Global Digital Insights*. https://datareportal.com/reports/digital-2024-indonesia

Sugiarto, E. C. (2024, December 18). Makan bergizi gratis dan SDM unggul. *Sekretariat Negara*. https://www.setneg.go.id/baca/index/makan_bergizi_gratis_dan_sdm_unggul?utm_source=chatgpt.com

Tarale, H. (2025). Web scraping using Python. *The Academic International Journal of Multidisciplinary Research*, *3*(1), 1023–1033. https://doi.org/10.5281/zenodo.14852767

Trisna, K. W. (2023). Model penerimaan pinjaman nasabah menggunakan algoritma Naïve Bayes dalam dataset bank. *JBASE - Journal of Business and Audit Information Systems*, *6*(1), 1–13. https://doi.org/10.30813/jbase.v6i1.4309

Yondra, A. S., Triyanto, D., & Bahri, S. (2022). Implementasi web scraping untuk mengumpulkan informasi produk dari situs e-commerce dan marketplace dengan teknik pemrosesan paralel. *Jurnal Komputer dan Aplikasi*, *10*(5), 93–102. https://doi.org/10.26418/coding.v10i01.52722

Zendrato, A. D., Berutu, S. S., Sumihar, Y. P., & Budiati, H. (2024). Pengembangan model klasifikasi sentimen dengan pendekatan VADER dan algoritma Naive Bayes terhadap ulasan aplikasi Indodax. *Journal of Information System Research (JOSH)*, *5*(3), 755–764. https://doi.org/10.47065/josh.v5i3.5050