Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 1, November 2025
Hal 29-34
DOI 10.47065/comforch.v5i1.2355

Analisis Sentimen Program Makanan Bergizi Gratis Menggunakan Algoritma *Random Forest* dan *Naive Bayes*

Rahmat Hidayat[*] Dewi Juliah Ratnaningsih

Program Studi Statistika, Fakultas Sains & Teknologi, Universitas Terbuka, Indonesia
[1,*] Email: 048409024@ecampus.ut.ac.id, [2]djuli@ecampus.ut.ac.id
Email Penulis Korespondensi: imjustrahmat2722@gmail.com
Received: **10/11/2025**, Revised: **24/11/2025**, Accepted: **29/11/2025**, Published: **30/11/2025**

Abstrak−Transformasi digital telah mendorong perubahan mendasar di berbagai aspek kehidupan, yang salah satunya ditandai dengan kemunculan teknologi MBG (Makanan Bergizi Gratis) yang telah memicu beragam opini dan perdebatan di ruang publik. Penelitian ini bertujuan untuk menganalisis dan memetakan sentimen masyarakat terhadap perkembangan program MBG. Selain itu, penelitian ini bertujuan untuk membandingkan performa dua algoritma Machine Learning, yaitu Naïve Bayes dan Logistic Regression, dalam tugas klasifikasi sentimen pada data media sosial. Peneliti menganalisis sebanyak 6.728 komentar masyarakat yang diambil dari platform media sosial X (Twitter). Pendekatan yang digunakan adalah text mining dan klasifikasi sentimen. Dalam tahap awal, model Naïve Bayes dan Logistic Regression menunjukkan akurasi awal yang tinggi (90% dan 91%), namun metrik evaluasi seperti precision, recall, dan F1-Score menunjukkan nilai yang rendah akibat adanya ketidakseimbangan data (dominasi sentimen positif). Untuk mengatasi masalah ketimpangan data ini, peneliti menerapkan teknik SMOTE (Synthetic Minority Oversampling Technique). Setelah penerapan SMOTE, performa kedua algoritma meningkat secara signifikan. Logistic Regression menunjukkan hasil yang paling unggul dengan metrik evaluasi yang jauh lebih baik (akurasi 95%, precision 94%, recall 93%, dan F1-Score 95%). Kontribusi utama dari penelitian ini adalah memberikan pemahaman yang mendalam dan terukur mengenai penerimaan dan opini publik terhadap teknologi MBG, yang dapat menjadi masukan strategis bagi pengembang kebijakan atau teknologi terkait. Selain itu, temuan ini secara metodologis membuktikan bahwa Logistic Regression adalah model yang lebih robust dan efektif dibandingkan Naïve Bayes dalam tugas klasifikasi sentimen berbasis data media sosial yang tidak seimbang (setelah penanganan data) terkait isu teknologi.

Kata kunci: Naïve Bayes; Logistic Regression; MBG; Analisis Sentimen; Random Forest

Abstract−Digital transformation has driven fundamental changes across various aspects of life, marked, in part, by the emergence of the Free Nutritious Meals (MBG) program. This program, which leverages technology to support its implementation, has sparked diverse opinions and public debate. This study has two main objectives: first, to analyze and map public sentiment regarding the development of the MBG program. Second, to compare the performance of two Machine Learning algorithms, Naïve Bayes and Logistic Regression, in a sentiment classification task using social media data. We analyzed 6,728 public comments extracted from the social media platform X (Twitter), utilizing a text mining and sentiment classification approach. In the initial evaluation phase, both the Naïve Bayes and Logistic Regression models demonstrated high initial accuracy (90% and 91%). However, evaluation metrics such as precision, recall, and F1-Score showed low values due to data imbalance (the dominance of positive sentiment). To address this issue, we applied the SMOTE (Synthetic Minority Oversampling Technique). Following the implementation of SMOTE, the performance of both algorithms improved significantly. Logistic Regression emerged as the superior model, achieving substantially better evaluation metrics (accuracy 95%, precision 94%, recall 93%, and F1-Score 95%). The primary contribution of this research is to provide a deep and measurable understanding of public acceptance and opinion toward the MBG program, which is valuable strategic input for policymakers. Furthermore, the findings methodologically prove that Logistic Regression is a more robust and effective model than Naïve Bayes for unbalanced social media-based sentiment classification tasks (after data handling) related to public policy issues.

Keywords: Naïve Bayes; Logistic Regression; MBG; Sentiment Analysis; Random Forest

1. PENDAHULUAN

Perkembangan teknologi informasi yang pesat telah mendorong pemanfaatan kecerdasan buatan (AI) di berbagai sektor, termasuk dalam pengolahan data teks untuk menganalisis isu-isu publik. Salah satu isu kebijakan yang memunculkan beragam opini di ruang publik adalah program Makanan Bergizi Gratis (MBG). Seiring dengan berjalannya diskursus publik, semakin banyak pula komentar yang diunggah oleh masyarakat di media sosial terkait program ini. Komentar tersebut menjadi sumber data yang sangat berharga untuk mengevaluasi kualitas kebijakan serta persepsi publik terhadap implementasinya. Namun demikian, opini masyarakat di media sosial umumnya bersifat tidak terstruktur dan sangat bervariasi dalam penggunaan bahasa. Oleh sebab itu, pendekatan berbasis analisis sentimen menjadi solusi yang efektif dalam mengkaji opini publik secara sistematis. Analisis sentimen merupakan metode untuk mengklasifikasikan teks ke dalam kategori sentimen positif, negatif, atau netral guna memperoleh wawasan mengenai sikap publik terhadap suatu objek atau layanan . Metode ini terbukti mampu menangkap aspek emosional dari teks yang tidak dapat ditangkap oleh metode evaluasi kuantitatif tradisional [1],[2].

Di bidang pengolahan bahasa alami, algoritma *machine learning* seperti Random Forest dan Naive Bayes sering digunakan karena kemampuannya dalam mengklasifikasikan data teks secara efektif. Random Forest dikenal sebagai algoritma yang memiliki akurasi tinggi, tahan terhadap *overfitting*, dan cocok untuk data berdimensi tinggi . Sedangkan Naive Bayes populer karena kesederhanaannya, efisiensi komputasi, serta kemampuannya dalam menangani data teks dalam bahasa lokal, termasuk Bahasa Indonesia. Penelitian sebelumnya telah menemukan bahwa kedua algoritma tersebut menunjukkan performa yang kompetitif dalam menganalisis sentimen pada data opini. Studi lain menunjukkan bahwa kombinasi *text mining* dan klasifikasi sentimen dapat memberikan gambaran objektif terhadap respons publik atas suatu

Copyright © 2025 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 1, November 2025
Hal 29-34
DOI 10.47065/comforch.v5i1.2355

isu [6]. Telah ditegaskan pula bahwa masukan publik melalui analisis opini dapat digunakan untuk merancang strategi atau fitur kebijakan agar sesuai dengan kebutuhan masyarakat [3].

Sentimen yang diperoleh dari analisis data berbasis AI terbukti membantu pemangku kepentingan dalam merancang strategi yang lebih adaptif. Analisis berbasis teks juga ditunjukkan mampu mengungkap keluhan tersembunyi dari publik yang tidak muncul dalam survei tertutup. Lebih lanjut, penelitian telah menyatakan bahwa model klasifikasi berbasis *machine learning* memberikan hasil yang lebih detail dan tajam dalam mengevaluasi respons publik. Tantangan terbesar dalam mengklasifikasikan sentimen pada komentar berbahasa Indonesia adalah keberagaman kosakata, struktur kalimat yang kompleks, serta idiom lokal yang tidak selalu dapat dikenali oleh model berbasis aturan klasik . Oleh karena itu, penting untuk membandingkan dan mengevaluasi berbagai metode klasifikasi untuk menentukan pendekatan yang paling efektif dalam konteks bahasa lokal [4].

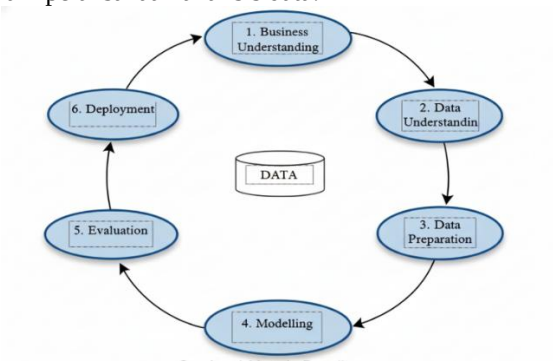
Penelitian ini menyumbangkan kontribusi ganda. Kontribusi pertama adalah sumbangsih metodologis, di mana peneliti berhasil menunjukkan bahwa Logistic Regression jauh lebih efektif dan *robust* dibandingkan Naïve Bayes dalam menganalisis sentimen dari data media sosial yang awalnya tidak seimbang terkait isu teknologi, setelah diterapkan penanganan data yang tepat. Kontribusi kedua adalah sumbangsih praktis, yaitu memberikan data yang terukur dan wawasan yang jelas mengenai bagaimana publik menerima teknologi MBG, yang penting sebagai dasar pengambilan keputusan strategis.

Studi telah membuktikan bahwa *feature selection* yang tepat dapat meningkatkan akurasi model Naive Bayes hingga lebih dari 10%. Hal serupa juga ditunjukkan, di mana kombinasi *preprocessing* dan pemilihan algoritma yang sesuai mampu meningkatkan efisiensi sistem analisis sentimen. Selain itu, penelitian lain menyatakan bahwa kualitas data dan representasi fitur sangat berpengaruh terhadap performa model klasifikasi. Dengan mempertimbangkan berbagai aspek tersebut, maka analisis komparatif antara *Random Forest* dan *Naive Bayes* dalam opini publik terkait program MBG menjadi penting untuk dilakukan. Penelitian ini bertujuan untuk mengidentifikasi algoritma mana yang lebih optimal dalam mengklasifikasikan sentimen masyarakat, serta memberikan kontribusi berupa wawasan berbasis data untuk evaluasi kebijakan program MBG secara lebih tepat sasaran.

2. METODOLOGI PENELITIAN

2.1 Tahapan penelitian

Penelitian ini dilakukan melalui beberapa tahapan yang dirancang dan dikembangkan berdasarkan model *Cross-Industry Standard Process for Data Mining* **(CRISP-DM)**. Model tersebut mencakup enam tahap utama yang saling berhubungan dan berfungsi sebagai panduan dalam pelaksanaan analisis data.



Gambar 1. Tahapan Penelitian

Tahapan dalam model CRISP-DM mencakup enam langkah utama, yaitu *Business Understanding* (pemahaman terhadap permasalahan bisnis), *Data Understanding* (pemahaman data), *Data Preparation* (persiapan data), *Modeling* (pembuatan model), *Evaluation* (evaluasi model), dan *Deployment* (implementasi hasil). Setiap tahap memiliki fungsi penting yang saling mendukung, mulai dari proses identifikasi masalah hingga penyajian hasil akhir [1]. Rancangan model penelitian yang diusulkan pada studi ini ditampilkan pada Gambar 1, yang memperlihatkan alur tahapan penelitian dengan menggunakan pendekatan CRISP-DM sebagai kerangka kerja utama dalam proses pengolahan dan analisis data.

2.2 *Naïve Bayes*

Naive Bayes adalah metode klasifikasi berbasis probabilitas yang memprediksi kategori suatu data dengan memanfaatkan informasi dari data sebelumnya. Metode ini disebut "naive" karena mengasumsikan bahwa setiap fitur atau ciri pada data saling berdiri sendiri dan tidak saling memengaruhi, meskipun dalam banyak kasus hal tersebut tidak sepenuhnya benar. Walaupun asumsi ini sederhana, Naive Bayes terbukti sangat efektif dan banyak digunakan dalam berbagai tugas klasifikasi [5],[6]. Cara kerjanya adalah dengan melihat seberapa sering sebuah kelas muncul dalam data serta seberapa sering setiap ciri muncul pada masing-masing kelas. Informasi tersebut kemudian digunakan untuk menentukan kelas yang paling mungkin dari data baru [7],[8]. Naive Bayes sangat populer dalam klasifikasi teks, seperti deteksi spam,

Copyright © 2025 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 1, November 2025
Hal 29-34
DOI 10.47065/comforch.v5i1.2355

analisis sentimen, dan identifikasi topik, karena prosesnya cepat, mudah diterapkan, serta mampu bekerja baik pada dataset besar maupun data dengan jumlah fitur yang tinggi [9],[10],[11],[12].

2.3 *Logistic Regression*

Logistic Regression adalah sebuah metode statistik dan machine learning yang digunakan untuk memprediksi suatu kategori atau kelas berdasarkan data input. Meskipun namanya mengandung kata "regression", metode ini sebenarnya digunakan untuk klasifikasi, terutama ketika hasil prediksi hanya memiliki dua kemungkinan, seperti ya/tidak, lulus/tidak lulus, atau spam/bukan spam. Logistic Regression bekerja dengan menganalisis hubungan antara fitur-fitur dalam data dan kemungkinan suatu data termasuk dalam kelas tertentu. Metode ini mengubah hasil perhitungan menjadi nilai probabilitas antara 0 sampai 1, lalu menentukan kelas berdasarkan ambang batas tertentu. Logistic Regression banyak digunakan karena mudah dipahami, interpretasinya jelas, dan efektif dalam berbagai kasus, terutama ketika hubungan antara fitur dan kelas bersifat linier. Selain untuk klasifikasi biner, metode ini juga dapat diperluas untuk menangani lebih dari dua kelas melalui teknik tertentu, sehingga menjadi salah satu algoritma dasar yang paling sering digunakan dalam analisis data dan kecerdasan buatan [13],[14].

2.4 MBG

MBG atau Makanan Bergizi Gratis adalah sebuah program bantuan pangan yang bertujuan menyediakan makanan sehat dan bernutrisi kepada kelompok sasaran, seperti anak sekolah, masyarakat kurang mampu, atau kelompok rentan lainnya secara gratis. Program ini dirancang untuk memastikan bahwa setiap orang, khususnya anak-anak, mendapatkan asupan gizi yang cukup untuk mendukung pertumbuhan, kesehatan, dan aktivitas sehari-hari. MBG biasanya berisi makanan yang memenuhi standar gizi seimbang, seperti sumber karbohidrat, protein, vitamin, dan mineral. Tujuannya tidak hanya mengurangi angka kekurangan gizi, tetapi juga meningkatkan kualitas kesehatan, mendukung konsentrasi belajar, serta membantu mensejahterakan masyarakat. Program MBG sering dijalankan oleh pemerintah, sekolah, lembaga sosial, atau komunitas sebagai bentuk dukungan pemenuhan kebutuhan gizi secara berkelanjutan [15].

2.5 Analisis Sentimen

Analisis sentimen adalah proses mengidentifikasi, memahami, dan mengklasifikasikan opini atau emosi yang terkandung dalam suatu teks, seperti apakah sebuah pernyataan bersifat positif, negatif, atau netral. Analisis ini biasanya dilakukan pada data berbasis teks seperti ulasan produk, komentar media sosial, survei, atau artikel berita. Tujuan utama analisis sentimen adalah mengetahui bagaimana persepsi atau pandangan seseorang terhadap suatu produk, layanan, kebijakan, atau topik tertentu [16]. Metode ini banyak digunakan dalam dunia bisnis untuk memantau kepuasan pelanggan, dalam politik untuk melihat opini publik, serta dalam penelitian untuk memahami sikap atau kecenderungan masyarakat [17],[18].

3. HASIL DAN PEMBAHASAN

3.1 Hasil Evaluasi dengan metode Random Forest

Dalam proses pemodelan, diterapkan mertode Random Forest sebagai teknik klasifikasinya dengan hasil dapat dilihat pada Gambar 2 sebagai berikut:

accuracy: 96.38% +/- 9.46% (micro average: 96.50%)

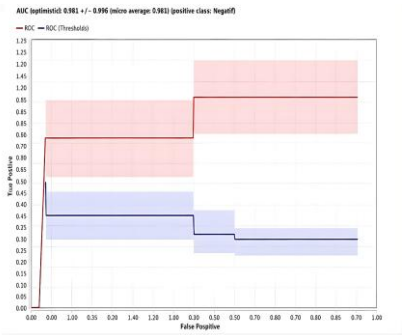
	true Positif	true Negatif	class precision ion
pred. Positif	96	3	96.97%
pred. Negatif	4	97	96.04%
class recall	96.00%	97.00%	

Gambar 2. Hasil Akurasi Dengan Metode *Random Forest*

Hasil analisis untuk model Random Forest menunjukkan bahwa metode ini mampu mencapai tingkat akurasi sebesar 96,38% dalam analisis sentimen ulasan Program MBG. Performa model ini dievaluasi menggunakan *confusion matrix*, yang disajikan pada Gambar 3. Hasilnya merinci bahwa model berhasil memprediksi 96 ulasan positif dengan benar (*True Positive*) dan 97 ulasan negatif dengan benar (*True Negative*).Namun, tercatat 3 ulasan negatif yang salah diprediksi sebagai positif (*False Positive*) dan 4 ulasan positif yang salah diprediksi sebagai negatif (*False Negative*). Selain *confusion matrix*, tingkat akurasi ini juga didukung oleh kurva ROC, yang ditampilkan pada Gambar 3 berikut.

Copyright © 2025 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 1, November 2025
Hal 29-34
DOI 10.47065/comforch.v5i1.2355



Gambar 3. ROC *curve* dari algoritma *Random Forest*

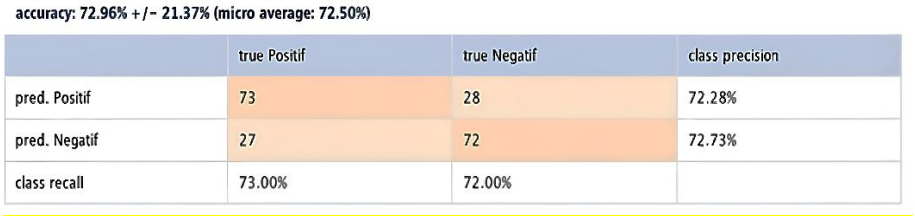
Gambar 3 menyajikan kurva *Receiver Operating Characteristic* (ROC) untuk model *Random Forest*. Dari kurva tersebut, diperoleh nilai *Area Under Curve* (AUC) sebesar 0,981 dengan standar deviasi ±0,096. Nilai AUC yang tinggi dan mendekati 1 ini mengindikasikan bahwa model Random Forest memiliki kemampuan klasifikasi yang sangat baik dalam membedakan sentimen positif dan negatif. Hal ini menyimpulkan bahwa model tersebut mampu menghasilkan performa yang optimal untuk data ulasan Program MBG. Adapun pada Tabel 1 merupakan rincian hasil nilai evaluasi adalah sebagai berikut:

Tabel 1. Nilai hasil *Accuracy, Recall,* AUC, *Random Forest*

Model	*Accuracy*	*Recall*	*Precision*	AUC
Random Forest	96.38%	97.17%	97.48%	0.981

3.2 Hasil Akurasi Dengan Metode Naive Bayes

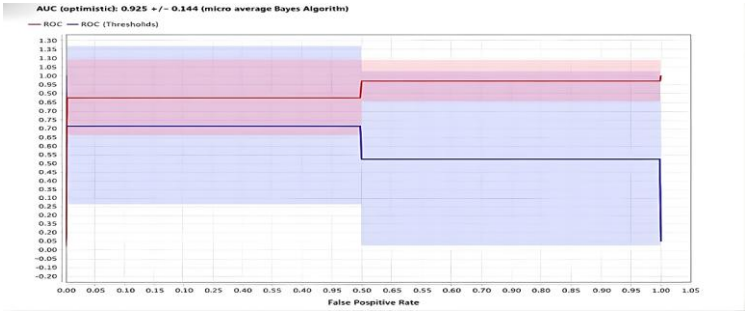
Pada tahap pemodelan, algoritma *Naïve Bayes* diterapkan sebagai metode klasifikasi kedua untuk menganalisis sentimen ulasan pengguna Program MBG. Model ini dilatih dan diuji menggunakan *dataset* yang telah diproses dan diberi label sentimen secara manual. Berikut pada Gambar 4 menunjukkan Hasil evaluasi menunjukkan bahwa algoritma *Naïve Bayes* mampu mengklasifikasikan data dengan tingkat akurasi sebesar 72,96%.


accuracy: 72.96% +/- 21.37% (micro average: 72.50%)

	true Positif	true Negatif	class precision
pred. Positif	73	28	72.28%
pred. Negatif	27	72	72.73%
class recall	73.00%	72.00%	

Gambar 4. Hasil Akurasi Dengan Metode *Naive Bayes*

Berdasarkan pada Gambar 4, yang menyajikan *confusion matrix* untuk metode *Naïve Bayes* (NB), hasil analisis menunjukkan tingkat akurasi 72,96% pada data sentimen ulasan Program MBG. Pengujian yang dilakukan menggunakan 200 data ulasan ini merinci bahwa model *Naïve Bayes* mampu memprediksi 73 ulasan positif dengan benar (*True Positive*) dan 72 ulasan negatif dengan benar (*True Negative*). Namun, *confusion matrix* ini juga mencatat adanya 28 data ulasan negatif yang salah diprediksi sebagai positif (*False Positive*) serta 27 data ulasan positif yang salah diprediksi sebagai negatif (*False Negative*).



Gambar 5. ROC *curve* dari algoritma *Naive Bayes*

Copyright © 2025 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 32
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 1, November 2025
Hal 29-34
DOI 10.47065/comforch.v5i1.2355

Kurva ROC yang ditampilkan pada Gambar 6 merupakan hasil evaluasi performa model *Naïve Bayes*. Dari kurva tersebut, diperoleh nilai *Area Under Curve* (AUC) sebesar 0,926, dengan standar deviasi ±0,144, serta nilai *micro average* sebesar 0,925. Berdasarkan uraian tersebut, dapat disimpulkan bahwa pengolahan data ulasan Program MBG menggunakan metode *Naïve Bayes* (NB) mendapatkan rincian hasil pada Tabel 2 sebagai berikut:

Tabel 2. Nilai hasil Accuracy, Recall, Precisiondan AUC Naive Bayes

Model	Accuracy	Recall	Precision	AUC
Naïve Bayes	72.96%	72.73%	71.02%	0.925

3.3 Visualisasi Data

Pada penelitian ini, tahap *Exploratory Data Analysis* (EDA) atau analisis data eksploratif dilakukan dengan memanfaatkan teknik visualisasi *word cloud* untuk menganalisis data teks. Metode ini dipilih karena kemampuannya dalam menyajikan intisari dari volume data tekstual yang besar secara intuitif dan cepat. *Word cloud* berfungsi untuk menampilkan kata-kata kunci yang terdapat di dalam *dataset* ulasan Program MBG (umumnya setelah melalui tahap *preprocessing* seperti *stopword removal*). Mekanisme kerjanya adalah dengan merepresentasikan frekuensi setiap kata secara visual, di mana kata yang memiliki frekuensi kemunculan tertinggi akan ditampilkan dengan ukuran *font* yang proporsional lebih besar dibandingkan kata yang jarang muncul. Dengan demikian, visualisasi ini sangat berguna bagi peneliti untuk mengidentifikasi secara cepat tema, topik, atau istilah yang paling dominan dan sering dibicarakan oleh pengguna dalam ulasan mereka.



Gambar 6. *Word Cloud*

4. KESIMPULAN

Berdasarkan hasil penelitian analisis sentimen terhadap Program MBG, ditarik kesimpulan bahwa terdapat perbedaan kinerja yang sangat signifikan antara dua metode yang diuji, yaitu *Random Forest* dan *Naïve Bayes*. Metode Random Forest terbukti jauh lebih unggul dengan berhasil mencapai tingkat akurasi tertinggi, yakni 96,38%, sebuah pencapaian yang menempatkannya jauh di atas *Naïve Bayes*. Keunggulan substansial ini bersumber dari kemampuan superior Random Forest dalam mengenali dan mempelajari pola persebaran kata yang kompleks di dalam kalimat-kalimat ulasan, sebuah proses yang difasilitasi menggunakan platform *RapidMiner Analytics*. Kemampuan mendalam inilah yang memungkinkan Random Forest untuk mengklasifikasikan sentimen positif dan negatif dengan presisi dan akurasi yang jauh lebih tinggi. Sebaliknya, metode *Naïve Bayes* menunjukkan kinerja dengan akurasi yang lebih rendah. Keterbatasan ini dapat diatribusikan pada sifat dasarnya sebagai algoritma yang berbasis probabilitas murni, yang sangat bergantung pada frekuensi kemunculan kata (frekuensi) dan memiliki keterbatasan inheren dalam mengenali pola kalimat yang lebih kompleks dan kontekstual pada data ulasan yang digunakan. Temuan ini diperkuat lebih lanjut oleh analisis visual "Gambaran Umum Data" yang disajikan dalam bentuk *word cloud*. Visualisasi ini secara efektif memetakan istilah-istilah yang paling sering muncul dan menjadi fokus utama dalam data ulasan, yang notabene merupakan bahan baku dari analisis sentimen ini. Secara spesifik, *word cloud* tersebut didominasi oleh kata-kata kunci seperti "makan", "siang", "gratis", "bergizi", "program", dan "mbg", yang secara gamblang menegaskan bahwa inti dari keseluruhan data ulasan adalah diskusi publik seputar Program Makan Siang Gratis Bergizi (MBG). Selain itu, frekuensi kemunculan kata "anak" dan "sekolah" dengan jelas mengidentifikasi target penerima manfaat utama dari program ini, sementara kehadiran kata "pemerintah" menunjuk pada pelaku atau penyelenggara utama yang menjadi sorotan dalam ulasan. Lebih jauh lagi, kata-kata pendukung lainnya seperti "gizi", "rakyat", "Indonesia", dan bahkan "Papua" memberikan konteks tambahan yang berharga, mengindikasikan bahwa diskusi tidak hanya terbatas pada pelaksanaan program, tetapi juga meluas ke isu-isu terkait seperti asupan gizi, dampaknya bagi rakyat Indonesia secara keseluruhan, dan kemungkinan adanya fokus atau studi kasus di wilayah tertentu seperti Papua. Dengan demikian, *word cloud* ini memberikan konfirmasi visual yang kuat mengenai topik sentral data, dan data inilah yang kemudian diolah untuk mengekstrak sentimen (positif atau negatif), di mana metode Random Forest terbukti paling efektif dalam memahami pola kompleks di balik susunan kata-kata tersebut.

Copyright © 2025 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

ISSN 2808-375X (Media Online)
Vol 5, No 1, November 2025
Hal 29-34
DOI 10.47065/comforch.v5i1.2355

REFERENCES

[1] E. Marlena *et al.*, "Implementasi Data Mining Untuk Memprediksi Status Gizi Anak Balita Pada Puskesmas Gedung Sari Menggunakan Polynomial Regression," *J. Teknol. Komput. dan Sist. Inf.*, vol. 6, no. 2, pp. 193–196, 2023.
[2] I. Kurniati, C. S. Dewi, and R. Juniantika, "Penerapan Data Mining dengan Algoritma Neural Network pada Sistem Informasi Prediksi Kasus Balita Gizi Buruk di Provinsi Nusa Tenggara Barat," *JRIS J. Rekayasa Inf. Swadharma*, vol. 1, no. 1, pp. 20–27, 2021.
[3] M. F. Azmi, P. Purwadi, and G. Syahputra, "Sistem Pakar Mendeteksi Gizi Buruk Pada Balita Menggunakan Metode Case Based Reasoning," *J. Cyber Tech*, vol. 1, no. 3, 2022.
[4] M. Brilliant and N. Nizamiyati, "Sistem Pakar Metode Case-Based Reasoning Untuk Deteksi Penyakit Stunting Pada Anak," *J. SIMADA (Sistem Inf. dan Manaj. Basis Data)*, vol. 5, no. 2, pp. 13–22, 2022.
[5] A. Damuri, U. Riyanto, H. Rusdianto, and M. Aminudin, "Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako," *JURIKOM (Jurnal Ris. Komputer)*, vol. 8, no. 6, pp. 219–225, 2021.
[6] S. Howay and S. Suhirman, "Comparison of SVM, NBC, and KNN Classification Methods in Determining Students' Majors at SMK N02 Manokwari," *J. Comput. Sci. Technol. Stud.*, vol. 5, no. 1, pp. 15–23, 2023, doi: 10.32996/jcsts.2023.5.1.3.
[7] A. Winantu and C. Khatimah, "Perbandingan Metode Klasifikasi Naive Bayes Dan K-Nearest Neighbor Dalam Memprediksi Prestasi Siswa," *INTEK J. Inform. dan Teknol. Inf.*, vol. 6, no. 1, pp. 58–64, 2023, doi: 10.37729/intek.v6i1.3006.
[8] H. Ashari, D. Arifianto, H. Azizah, and A. Faruq, "Perbandingan Kinerja Algoritma Multinominal Naive Bayes (MNB, Multivariate Bernoulli dan Rocchio Algortihm Dalam Klasifikasi Konten Berita Hoax Berbahasa Indonesia Pada Media Sosial," *Http://Repository.Unmuhjember.Ac.Id*, pp. 1–12, 2020.
[9] Baskoro, Sriyanto, and L. Setya Rini, "Prediksi Penerima Beasiswa dengan Menggunakan Teknik Data Mining di Universitas Muhammadiyah Pringsewu," *Semin. Nas. Has. Penelit. dan Pengabdi. Masy. Inst. Inform. dan Bisnis Darmajaya*, vol. 1, no. 2, pp. 87–94, 2021.
[10] N. A. Pradipta and R. D. H. Untari N, "Implementasi Algoritma Apriori Untuk Analisis Pola Pembelian Produk Donat Bolong," *Jutisi J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 13, no. 1, p. 268, 2024, doi: 10.35889/jutisi.v13i1.1778.
[11] N. S. Hanapiah *et al.*, "Meningkatkan Penanganan Stunting Pada Anak Melalui Klasifikasi Pemberian Makanan Tambahan Berdasarkan Metode K-Means," *J. Mhs. Tek. Inform.*, vol. 7, no. 6, pp. 3484–3493, 2023.
[12] M. Kholid, A. F. Boy, and Y. Syahra, "Implementasi Data Mining Metode Algoritma Apriori Untuk Mengetahui Pola Pembelian Konsumen pada Transaksi Penjualan Makanan Dan Minuman (Study Kasus Restaurant JMC Medan)," *J. Cyber Tech*, vol. 4, no. 7, 2021.
[13] E. Fitri and S. N. Nugraha, "Optimasi Kinerja Linear Regression, Random Forest Regression Dan Multilayer Perceptron Pada Prediksi Hasil Panen," *Inti Nusa Mandiri*, vol. 18, no. 2, pp. 210–217, 2024.
[14] M. A. Musababa, "Implementasi Algoritma Linear Regression untuk Prediksi Produksi Tanaman Padi di Kabupaten Grobogan," *Data Sci. Indones.*, vol. 3, no. 2, pp. 1–11, 2023.
[15] Z. A. Siregar, "Analisis Pola Penjualan Produk Makanan dan Minuman Menggunakan Algoritma Apriori," *J. Informatics Manag. Inf. Technol.*, vol. 2, no. 2, pp. 65–72, 2022.
[16] N. T. Luchia *et al.*, "Analisis Sentimen Ulasan Wisatawan Terhadap Destinasi Gili di KLU Menggunakan K-Means Clustering," *J. Teknol. Informasi, Komputer, dan Apl.*, vol. 4, no. 2, pp. 177–189, 2025.
[17] T. P. Lestari, "Analisis Text Mining pada Sosial Media Twitter Menggunakan Metode Support Vector Machine (SVM) dan Social Network Analysis (SNA)," *J. Inform. Ekon. Bisnis*, vol. 4, no. 3, pp. 65–71, 2022, doi: 10.37034/infeb.v4i3.146.
[18] Normah, B. Rifai, S. Vambudi, and R. Maulana, "Analisa Sentimen Perkembangan Vtuber Dengan Metode Support Vector Machine Berbasis SMOTE," *J. Tek. Komput. AMIK BSI*, vol. 8, no. 2, pp. 174–180, 2022, doi: 10.31294/jtk.v4i2.

Copyright © 2025 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.