Journal of Information System, Applied, Management, Accounting and Research.
http://journal.stmikjayakarta.ac.id/index.php/jisamar ,
jisamar@stmikjayakarta.ac.id , jisamar2017@gmail.com
e-ISSN: 2598-8719 (Online), p-ISSN: 2598-8700 (Printed) , Vol. 10 No.3 (August 2026)

ANALISIS KOMPARATIF KINERJA *LOGISTIC REGRESSION* DAN *RANDOM FOREST* BERBASIS TF-IDF UNTUK KLASIFIKASI SENTIMEN KOMENTAR TIKTOK TERHADAP PROGRAM MAKAN BERGIZI GRATIS

Febriza Evan Nugraha[1], Ibni Faiq Athallah[2*], Marrylinteri Istoningtyas[3]

Program Studi Teknik Informatik[1], Program Studi Teknik Informatik [2], Program Studi Teknik Informatik [3]
Fakultas Ilmu Komputer[1], Fakultas Ilmu Komputer [2], Fakultas Ilmu Komputer [3]
Universitas Dinamika Bangsa[1], Universitas Dinamika Bangsa[2], Universitas Dinamika Bangsa[3]

*Correspondent Author: ibnifaiqathallah28@gmail.com

Authors Email: febrizaevan23@gmail.com[1], ibnifaiqathallah28@gmail.com[2], marrylinteri.kuliah@gmail.com[3]

Received: May 20, 2026**. Revised:** July 08, 2026**. Accepted:** July 15, 2026**. Issue Period:** Vol.10 No.3 (2026), Pp. 787-798

Abstrak: Analisis sentimen terhadap kebijakan publik di media sosial menjadi semakin penting seiring meningkatnya partisipasi masyarakat dalam ruang digital. Penelitian ini membandingkan kinerja algoritma Logistic Regression dan Random Forest berbasis TF-IDF untuk mengklasifikasikan sentimen komentar TikTok terhadap Program Makan Bergizi Gratis ke dalam tiga kelas, yaitu positif, netral, dan negatif. Data yang digunakan berjumlah 4.811 komentar publik dari enam video TikTok yang dikumpulkan pada rentang Januari hingga Juni 2026. Setelah tahap preprocessing dan pelabelan manual, diperoleh 3.884 komentar valid dengan distribusi sentimen negatif 39,73%, positif 37,97%, dan netral 22,30%. Ketidakseimbangan kelas ditangani menggunakan SMOTE pada data latih, dan pembagian data dilakukan dengan stratified split 80:20. Hasil evaluasi menunjukkan bahwa Logistic Regression mengungguli Random Forest pada seluruh metrik, dengan akurasi 0,76 dan macro F1-score 0,74 berbanding akurasi 0,73 dan macro F1-score 0,71 milik Random Forest. Pada kedua model, kelas netral secara konsisten menjadi kelas dengan performa terendah, mengindikasikan ambiguitas semantik yang tidak dapat ditangkap secara optimal oleh representasi berbasis frekuensi kata. Penelitian ini memberikan bukti empiris bahwa Logistic Regression lebih unggul untuk klasifikasi sentimen teks media sosial berbahasa Indonesia dengan representasi TF-IDF, dan merekomendasikan eksplorasi model berbasis konteks seperti IndoBERT untuk penelitian selanjutnya.

DOI: 10.52362/jisamar.v10i3.2515

Journal of Information System, Applied, Management, Accounting and Research.
http://journal.stmikjayakarta.ac.id/index.php/jisamar ,
jisamar@stmikjayakarta.ac.id , jisamar2017@gmail.com
e-ISSN: 2598-8719 (Online), p-ISSN: 2598-8700 (Printed) , Vol. 10 No.3 (August 2026)

Kata kunci: Analisis Sentimen; Logistic Regression; Random Forest; TF-IDF; Program Makan Bergizi Gratis; TikTok

Abstract: *Sentiment analysis of public policy on social media has become increasingly important as public participation in digital spaces continues to grow. This study compares the performance of Logistic Regression and Random Forest algorithms based on TF-IDF for classifying the sentiment of TikTok comments regarding the Free Nutritious Meal Program into three classes: positive, neutral, and negative. The dataset consists of 4,811 public comments collected from six TikTok videos between January and June 2026. After preprocessing and manual labeling, 3,884 valid comments were obtained with a sentiment distribution of 39.73% negative, 37.97% positive, and 22.30% neutral. Class imbalance was addressed using SMOTE on the training data, and the dataset was split using an 80:20 stratified split. Evaluation results show that Logistic Regression outperformed Random Forest across all metrics, achieving an accuracy of 0.76 and a macro F1-score of 0.74 compared to Random Forest's accuracy of 0.73 and macro F1-score of 0.71. In both models, the neutral class consistently showed the lowest performance, indicating semantic ambiguity that cannot be optimally captured by frequency-based feature representations. This study provides empirical evidence that Logistic Regression is more suitable for Indonesian social media text sentiment classification with TF-IDF representation, and recommends exploring context-based models such as IndoBERT for future research.*

Keywords: *Sentiment Analysis; Logistic Regression; Random Forest; TF-IDF; Free Nutritious Meal Program; TikTok*

I. PENDAHULUAN

Era digital telah mengubah media sosial menjadi forum publik tempat masyarakat secara bebas mengekspresikan opini terhadap berbagai isu [1]. TikTok, yang menempati posisi kelima sebagai platform paling banyak digunakan di Indonesia berdasarkan laporan *Digital* 2026 [2], kini berfungsi tidak hanya sebagai sarana hiburan, tetapi juga sebagai ruang publik digital dengan tingkat partisipasi tinggi melalui fitur komentar [3]. Namun, besarnya volume komentar menimbulkan tantangan analisis yang tidak dapat diselesaikan secara manual [4].

Salah satu isu yang ramai diperbincangkan adalah Program Makan Bergizi Gratis (MBG), program pemerintah yang menyasar balita, siswa PAUD hingga SMA, serta ibu hamil [5]. Dengan alokasi anggaran mencapai 335 triliun rupiah dalam APBN 2026 [6], program ini memicu beragam respons publik di media sosial, mulai dari dukungan hingga kritik [7]. Memahami sentimen tersebut penting agar pemangku kepentingan dapat mengevaluasi penerimaan masyarakat dan mengidentifikasi potensi masalah implementasi [8].

Analisis sentimen berbasis *machine learning* telah terbukti efektif untuk tujuan tersebut [9], namun sebagian besar kajian terdahulu hanya membandingkan satu atau dua algoritma sederhana [10], dan kajian yang khusus menganalisis komentar TikTok terkait MBG masih sangat terbatas [11]. Penelitian ini membandingkan *Logistic Regression* (LR) dan *Random Forest* (RF) dengan ekstraksi fitur TF-IDF, karena keduanya merepresentasikan pendekatan yang berbeda—linear-probabilistik dan *ensemble* berbasis pohon keputusan—serta telah digunakan dalam berbagai studi klasifikasi teks bahasa Indonesia [12], [13], [14]. Hasil studi terdahulu menunjukkan pola yang tidak konsisten: LR mengungguli RF pada beberapa domain [4], [12], [13], sementara RF lebih unggul pada domain lain [10], [16], mengonfirmasi bahwa performa keduanya sangat bergantung pada karakteristik data. Oleh karena itu, penelitian ini bertujuan memberikan gambaran empiris perbandingan kinerja LR dan RF dalam mengklasifikasikan sentimen komentar TikTok terhadap Program MBG.

II. METODE DAN MATERI

DOI: 10.52362/jisamar.v10i3.2515

Journal of Information System, Applied, Management, Accounting and Research.
http://journal.stmikjayakarta.ac.id/index.php/jisamar ,
jisamar@stmikjayakarta.ac.id , jisamar2017@gmail.com
e-ISSN: 2598-8719 (Online), p-ISSN: 2598-8700 (Printed) , Vol. 10 No.3 (August 2026)

2.1. Metode yang Digunakan

Penelitian ini menggunakan pendekatan *text mining* dengan metode klasifikasi untuk mengkategorikan komentar TikTok ke dalam tiga kelas sentimen, yaitu positif, netral, dan negatif [19], [20]. Dua algoritma yang dibandingkan adalah Logistic Regression dan Random Forest, keduanya dilatih menggunakan representasi fitur berbasis TF-IDF. Pemilihan kedua algoritma ini didasarkan pada perbedaan mendasar pendekatannya—Logistic Regression bekerja secara linear-probabilistik sementara Random Forest mengandalkan mekanisme *ensemble* berbasis pohon keputusan [29], [30]—sekaligus karena keduanya telah digunakan secara berdampingan dalam berbagai studi klasifikasi teks berbahasa Indonesia [12], [13].

Untuk menangani ketidakseimbangan distribusi kelas yang lazim ditemukan pada data sentimen kebijakan publik [17], penelitian ini menerapkan *Synthetic Minority Over-sampling Technique* (SMOTE) pada data latih sebelum proses pelatihan model [27]. Evaluasi dilakukan menggunakan confusion matrix, akurasi, presisi, recall, dan F1-score per kelas, dengan macro F1 sebagai metrik penentu dalam perbandingan kinerja antaralgoritma [18].

2.2. Penelitian Terdahulu

Penelitian mengenai analisis sentimen di media sosial telah berkembang cukup pesat, khususnya yang memanfaatkan pendekatan *machine learning* untuk memahami opini publik terhadap berbagai isu. Sejumlah studi telah menunjukkan bahwa data teks dari platform media sosial dapat diolah secara otomatis untuk mengklasifikasikan sentimen pengguna [8], [9].

Studi perbandingan berskala besar yang dilakukan Couronné et al. [15] terhadap ratusan dataset nyata menemukan bahwa Random Forest cenderung mengungguli Logistic Regression pada sekitar 69% kasus, meskipun selisihnya bervariasi tergantung karakteristik data. Namun temuan tersebut tidak selalu berlaku untuk teks berbahasa Indonesia. Simanjuntak et al. [12] justru menemukan sebaliknya saat mengklasifikasikan emosi pada tweet, di mana Logistic Regression menghasilkan akurasi lebih tinggi dibanding Random Forest. Pola serupa juga ditemukan oleh Junianto et al. [13] pada ulasan aplikasi Disdukcapil dan oleh Sulistyawati dan Prabowo [4] pada komentar TikTok secara umum. Di sisi lain, Saepudin et al. [16] dan Butsianto dan Rifa'i [10] menunjukkan bahwa Random Forest justru lebih unggul pada domain ulasan e-commerce dan aplikasi berbahasa Indonesia. Inkonsistensi ini menunjukkan bahwa performa kedua algoritma sangat bergantung pada domain dan karakteristik data yang digunakan [15].

Kajian terkait sentimen pada Program Makan Bergizi Gratis (MBG) juga telah mulai muncul. Prianto et al. [17] menemukan bahwa distribusi kelas pada data sentimen kebijakan publik cenderung tidak seimbang, dengan kelas netral sebagai minoritas, kondisi yang berpotensi membuat model bias terhadap kelas mayoritas. Meskipun demikian, kajian yang secara khusus membandingkan kinerja *Logistic Regression* dan *Random Forest* pada komentar TikTok terkait kebijakan MBG masih belum ditemukan [4], [11], sehingga penelitian ini mengisi celah tersebut.

2.3. Data Penelitian

Data yang digunakan dalam penelitian ini adalah komentar publik dari video TikTok yang membahas Program Makan Bergizi Gratis, dikumpulkan dalam rentang Januari hingga Juni 2026 menggunakan layanan ekspor komentar pihak ketiga bernama ExportTok. Pencarian video sumber dilakukan menggunakan variasi kata kunci seperti 'makan bergizi gratis', 'program MBG', dan 'makanan gratis pemerintah'. Pengumpulan hanya dilakukan pada komentar yang bersifat publik dan semata-mata untuk keperluan penelitian akademik non-komersial [3], [7].

Total dataset yang berhasil dikumpulkan berjumlah 4.811 komentar yang bersumber dari enam video TikTok. Dari sisi waktu, sebagian besar komentar berasal dari Maret 2026, yaitu sebanyak 85,0 persen dari total. Setiap baris dataset merepresentasikan satu komentar dengan 15 kolom metadata, namun hanya kolom teks komentar yang digunakan sebagai masukan pada *pipeline* klasifikasi. Pelabelan manual dilakukan oleh kedua peneliti secara kolektif berdasarkan pedoman tertulis, dengan setiap komentar didiskusikan hingga dicapai kesepakatan label [20].

2.4. Preprocessing Data

DOI: 10.52362/jisamar.v10i3.2515

Journal of Information System, Applied, Management, Accounting and Research.
http://journal.stmikjayakarta.ac.id/index.php/jisamar ,
jisamar@stmikjayakarta.ac.id , jisamar2017@gmail.com
e-ISSN: 2598-8719 (Online), p-ISSN: 2598-8700 (Printed) , Vol. 10 No.3 (August 2026)

Komentar TikTok berbahasa Indonesia memiliki karakteristik yang cukup khas: banyak mengandung emoji, singkatan tidak baku, pencampuran kode bahasa, serta elemen non-tekstual lain yang tidak relevan untuk analisis sentimen. Sebelum dapat digunakan untuk melatih model, data teks mentah harus melewati serangkaian tahapan preprocessing agar representasinya menjadi bersih, konsisten, dan informatif [21], [22].

Tahap preprocessing dimulai dari *data cleaning*, yaitu penghapusan URL, mention, hashtag, emoji, karakter Unicode non-ASCII, angka, dan tanda baca dari teks komentar [23]. Setelah itu dilakukan *case folding* untuk menyeragamkan seluruh huruf menjadi huruf kecil [22]. Selanjutnya dilakukan normalisasi kata tidak baku menggunakan kamus gabungan dari kamus colloquial bahasa Indonesia yang tersedia secara terbuka dan kamus tambahan yang disusun secara manual untuk menangani slang spesifik domain MBG [22]. Tahap berikutnya adalah tokenisasi untuk memecah teks menjadi token individual [21], dilanjutkan dengan *stopword removal* menggunakan daftar stopword Sastrawi yang dimodifikasi—kata negasi, *intensifier*, dan kata kontras dikecualikan dari penghapusan agar tidak merusak muatan sentimen teks [20], [22]. Tahap terakhir adalah *stemming* menggunakan algoritma Nazief dan Adriani melalui library Sastrawi [24].

2.5. Ekstrasi Fitur TF-IDF

Data hasil preprocessing dikonversi menjadi representasi numerik menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Metode ini memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam dokumen dan tingkat keunikannya dalam keseluruhan korpus [23], [25]. Nilai TF dihitung dengan membagi frekuensi kemunculan suatu kata dengan jumlah total kata dalam dokumen, seperti pada persamaan 1.

$$tf_{t,d} = \frac{n_{t,d}}{\sum_k n_{k,d}} \quad \textbf{(1)}$$

di mana $tf_{t,d}$ adalah nilai Term Frequency kata t dalam dokumen d, $n_{t,d}$adalah frekuensi kemunculan kata t dalam dokumen d, dan $\sum_k n_{k,d}$ adalah jumlah total seluruh kata dalam dokumen d.

Nilai TF saja belum cukup untuk menentukan tingkat kepentingan suatu kata karena kata yang sering muncul di banyak dokumen juga dapat memiliki nilai TF tinggi. Oleh karena itu digunakan Inverse Document Frequency (IDF) untuk memberikan bobot lebih besar pada kata yang jarang muncul dan bobot lebih kecil pada kata yang umum [25], [26]. Nilai IDF dihitung menggunakan Persamaan 2.

$$idf_t = \log \frac{N}{n_t} \quad \textbf{(2)}$$

di mana N adalah jumlah total dokumen dalam korpus dan n_t adalah jumlah dokumen yang mengandung kata t [26], [25]. Bobot TF-IDF akhir diperoleh dari perkalian kedua komponen tersebut, seperti pada persamaan 3.

$$tfidf_{t,d} = tf_{t,d} \times idf_t \quad \textbf{(2)}$$

Kata dengan nilai TF-IDF tinggi dianggap lebih informatif karena sering muncul pada dokumen tertentu tetapi jarang ditemukan pada dokumen lain. Sebaliknya, kata yang muncul secara umum pada banyak dokumen akan memiliki bobot yang lebih rendah. Bobot TF-IDF yang dihasilkan kemudian digunakan sebagai vektor fitur untuk proses klasifikasi [23], [25].

2.4. Penyeimbangan Data dengan SMOTE

Distribusi data sentimen menunjukkan bahwa jumlah data pada masing-masing kelas tidak sepenuhnya seimbang, terutama pada kelas netral yang memiliki jumlah data lebih sedikit dibandingkan kelas lainnya. Untuk mengurangi potensi bias model terhadap kelas mayoritas digunakan metode *Synthetic Minority Oversampling Technique* (SMOTE). Metode ini menghasilkan sampel sintetis pada kelas minoritas sehingga distribusi kelas menjadi lebih seimbang sebelum proses pelatihan model dilakukan [27]. Secara matematis, pembangkitan sampel sintetis dirumuskan seperti pada persamaan 4.

$$x_{(baru)} = x_i + \delta \times (x_{(nn)} - x_i) \quad \textbf{(3)}$$

DOI: 10.52362/jisamar.v10i3.2515

Journal of Information System, Applied, Management, Accounting and Research.
http://journal.stmikjayakarta.ac.id/index.php/jisamar ,
jisamar@stmikjayakarta.ac.id , jisamar2017@gmail.com
e-ISSN: 2598-8719 (Online), p-ISSN: 2598-8700 (Printed) , Vol. 10 No.3 (August 2026)

dengan x_i adalah sampel asli dari kelas minoritas, x_{nn} salah satu tetangga terdekatnya, dan δ merupakan bilangan acak antara 0 dan 1 [27].

2.5. *Logistic Regression*

Logistic Regression merupakan algoritma klasifikasi yang digunakan untuk memprediksi probabilitas suatu data termasuk ke dalam kelas tertentu berdasarkan fitur yang dimiliki [29]. Algoritma ini banyak digunakan pada klasifikasi teks karena mampu bekerja dengan baik pada data berdimensi tinggi serta memiliki kompleksitas komputasi yang relatif rendah [30], [31]. Dalam penelitian ini, *Logistic Regression* digunakan untuk mengklasifikasikan sentimen komentar TikTok berdasarkan fitur TF-IDF yang telah dihasilkan pada tahap sebelumnya. Fungsi *sigmoid* yang digunakan untuk menghasilkan probabilitas prediksi dirumuskan sebagai berikut [31]:

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad \textbf{(4)}$$

dengan (z) merupakan kombinasi linear dari fitur-fitur masukan dan (e) adalah konstanta Euler [31].

2.6. *Random Forest*

Random Forest merupakan algoritma machine learning berbasis *ensemble* yang membangun sejumlah pohon keputusan (*decision tree*) dan menggabungkan hasil prediksi dari setiap pohon untuk menghasilkan klasifikasi yang lebih akurat dan stabil [32]. Algoritma ini menggunakan teknik *bootstrap sampling* dan pemilihan fitur secara acak pada setiap pohon untuk meningkatkan kemampuan generalisasi model serta mengurangi risiko *overfitting* [32], [33]. Pada tugas klasifikasi, hasil akhir ditentukan menggunakan mekanisme *majority voting*, yaitu kelas yang memperoleh suara terbanyak dari seluruh pohon akan menjadi hasil prediksi [32].

2.7. Pembagian Data dan Evaluasi Model

Dataset hasil pelabelan dibagi menjadi data latih dan data uji dengan proporsi 80:20 menggunakan *stratified split* sebelum ekstraksi fitur dilakukan. Strategi ini memastikan proporsi setiap kelas sentimen tetap terjaga pada kedua subset [18]. Nilai *random state* ditetapkan secara konsisten untuk menjamin reprodusibilitas eksperimen.

Evaluasi kinerja model dilakukan pada data uji menggunakan *confusion matrix*, akurasi, presisi, *recall*, dan F1-*score* yang dihitung per kelas sentimen [18]. Macro F1 digunakan sebagai metrik utama dalam perbandingan kinerja antara Logistic Regression dan Random Forest karena memberikan bobot setara pada setiap kelas tanpa terpengaruh dominasi kelas mayoritas. Seluruh implementasi dilakukan menggunakan bahasa pemrograman Python dengan pustaka *scikit-learn* dan Sastrawi, dijalankan di lingkungan Visual Studio Code dan Google Colaboratory.

III. PEMBAHASA DAN HASIL

3.1. Hasil Preprocessing Data

Tahap preprocessing berhasil mengurangi jumlah data dari 4.811 komentar menjadi 4.012 komentar yang siap digunakan pada proses klasifikasi. Pengurangan terbesar terjadi pada tahap penghapusan komentar duplikat, yaitu sebanyak 539 data yang dihapus karena merupakan komentar identik dari pengguna yang sama, kemungkinan besar akibat artefak proses ekspor data. Selain itu, ditemukan 246 komentar yang menjadi kosong setelah melewati seluruh tahapan preprocessing — komentar-komentar ini umumnya hanya berisi emoji, tautan, atau karakter non-alfabet yang seluruhnya terhapus pada tahap cleaning — sehingga tidak dapat digunakan pada tahap selanjutnya. Tahap awal pembuangan komentar kosong sebelum preprocessing juga menyingkirkan 14 entri yang sejak awal tidak memiliki konten. Secara keseluruhan, *pipeline* preprocessing berjalan sesuai yang diharapkan dan menghasilkan dataset yang lebih bersih dan konsisten untuk proses klasifikasi.

Tabel 1. Ringkasan Hasil Preprocessing

Tahap	*Sebelum*	*Sesudah*	*Selisih*

DOI: 10.52362/jisamar.v10i3.2515

Journal of Information System, Applied, Management, Accounting and Research.
http://journal.stmikjayakarta.ac.id/index.php/jisamar ,
jisamar@stmikjayakarta.ac.id , jisamar2017@gmail.com
e-ISSN: 2598-8719 (Online), p-ISSN: 2598-8700 (Printed) , Vol. 10 No.3 (August 2026)

Pembuangan komentar kosong	4.811	4.797	14
Penghapusan komentar duplikat	4.797	4.258	539
Pembuangan komentar kosong pasca-preprocessing	4.258	4.012	246

3.2. Distribusi Data Sentimen

Hasil pelabelan manual terhadap 3.884 komentar menunjukkan bahwa sentimen negatif mendominasi dataset dengan jumlah 1.543 komentar atau 39,73% dari total data, diikuti oleh sentimen positif sebanyak 1.475 komentar (37,97%), dan sentimen netral sebanyak 866 komentar (22,30%). Dominasi sentimen negatif pada data komentar TikTok terkait Program MBG ini mengindikasikan bahwa masyarakat lebih banyak mengekspresikan kritik, kekhawatiran, atau ketidaksetujuan terhadap program tersebut dibandingkan dukungan. Distribusi ini juga memperlihatkan adanya ketidakseimbangan kelas yang cukup signifikan, terutama pada kelas netral yang jumlahnya hampir setengah dari dua kelas lainnya. Kondisi tersebut sejalan dengan temuan Prianto et al. [17] yang menyebutkan bahwa kelas netral cenderung menjadi minoritas pada data sentimen kebijakan publik, sehingga penanganan ketidakseimbangan kelas menjadi langkah yang diperlukan sebelum melatih model.

Tabel 2. Distribusi Data per-Kelas Sentimen

Kelas	***Jumlah***	***Persentase***
4.811	1.475	37,97%
4.797	4. 866	22,30%
4.258	1.543	39,73%
Total	**3.884**	**100%**

3.3. Hasil Penyeimbangan Data Menggunakan SMOTE

Sebelum diterapkan SMOTE, data training menunjukkan distribusi kelas yang tidak seimbang dengan kelas netral hanya berjumlah 693 data, jauh lebih sedikit dibandingkan kelas positif (1.180 data) dan kelas negatif (1.234 data). Ketidakseimbangan ini berpotensi membuat model cenderung mengabaikan kelas netral dalam proses klasifikasi karena model lebih sering "melihat" kelas mayoritas selama pelatihan [27], [28].

Setelah diterapkan SMOTE pada data training, jumlah data pada setiap kelas menjadi seimbang, yaitu masing-masing 1.234 data, sehingga total data training meningkat dari 3.107 menjadi 3.702 data. SMOTE mencapai penyeimbangan ini dengan membangkitkan sampel sintetis pada kelas netral dan positif melalui interpolasi terhadap tetangga terdekat di ruang fitur, bukan sekadar menduplikasi data yang sudah ada [27]. Perlu dicatat bahwa SMOTE hanya diterapkan pada data training, sedangkan data testing dipertahankan dalam kondisi distribusi aslinya agar hasil evaluasi mencerminkan kondisi nyata.

Tabel 3. Distribusi Data Training per Kelas Sebelum dan Sesudah SMOTE

Kelas	***Sebelum***	***Sesudah***
Positif	1.180	1.234
Netral	693	1.234
Negatif	1.234	1.234

DOI: 10.52362/jisamar.v10i3.2515

Journal of Information System, Applied, Management, Accounting and Research.
http://journal.stmikjayakarta.ac.id/index.php/jisamar ,
jisamar@stmikjayakarta.ac.id , jisamar2017@gmail.com
e-ISSN: 2598-8719 (Online), p-ISSN: 2598-8700 (Printed) , Vol. 10 No.3 (August 2026)

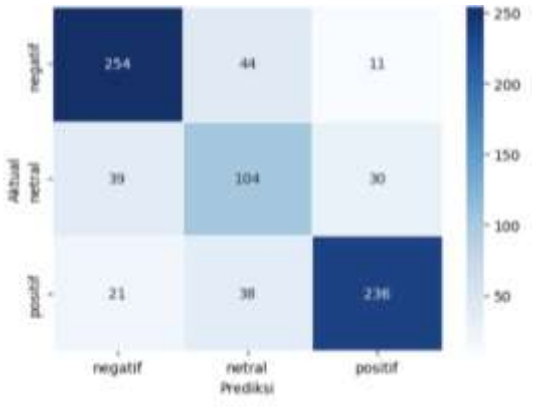
Total	**3.107**	**3.702**

3.4. Hasil Klasifikasi *Logistic Regression*

Model *Logistic Regression* menghasilkan akurasi keseluruhan sebesar 0,76 pada data uji. Berdasarkan *confusion matrix* pada Gambar 1, dari 309 komentar negatif yang sebenarnya, model berhasil mengklasifikasikan dengan benar sebanyak 254 data, sementara 44 data salah diprediksi sebagai netral dan 11 data sebagai positif. Pada kelas netral, dari 173 data aktual, model hanya berhasil mengklasifikasikan 104 data dengan benar, sedangkan 39 data salah diprediksi sebagai negatif dan 30 data sebagai positif. Pada kelas positif, dari 295 data aktual, sebanyak 236 berhasil diklasifikasikan dengan benar, sementara 21 data salah diprediksi sebagai negatif dan 38 sebagai netral.

Tabel 4. Hasil Klasifikasi Logistic Regression

Kelas	***Precision***	***Recall***	***F1-Score***
Positif	0,85	0,80	0,83
Netral	0,56	0,60	0,58
Negatif	0,81	0,82	0,82
Accuracy			**0,76**
Macro avg	**0,74**	**0,74**	**0,74**



Gambar 1. Confusion Matrix Logistic Regression

Dilihat per kelas, kelas positif memiliki performa terbaik dengan *precision* 0,85, *recall* 0,80, dan *F1-score* 0,83. Kelas negatif juga menunjukkan performa yang baik dengan *precision* 0,81, *recall* 0,82, dan *F1-score* 0,82. Sementara itu, kelas netral menjadi kelas dengan performa paling rendah, yaitu *precision* 0,56, *recall* 0,60, dan *F1-score* 0,58. Pola ini wajar mengingat kelas netral merupakan kelas dengan jumlah data terkecil dan secara semantik memiliki batas yang lebih kabur dengan dua kelas lainnya, sehingga lebih rentan mengalami misklasifikasi. Secara keseluruhan, model *Logistic Regression* mencatatkan *macro average precision*, *recall*, dan *F1-score* masing-masing sebesar 0,74.

3.5. Hasil Klasifikasi *Random Forest*

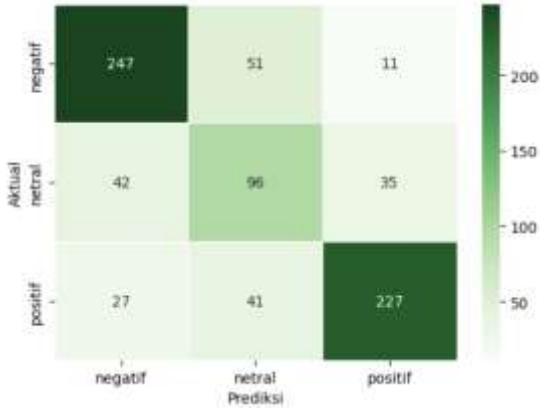
DOI: 10.52362/jisamar.v10i3.2515

Journal of Information System, Applied, Management, Accounting and Research.
http://journal.stmikjayakarta.ac.id/index.php/jisamar ,
jisamar@stmikjayakarta.ac.id , jisamar2017@gmail.com
e-ISSN: 2598-8719 (Online), p-ISSN: 2598-8700 (Printed) , Vol. 10 No.3 (August 2026)

Model *Random Forest* menghasilkan akurasi keseluruhan sebesar 0,73 pada data uji. Berdasarkan *confusion matrix* pada Gambar 2, dari 309 komentar negatif aktual, model mengklasifikasikan 247 data dengan benar, sementara 51 data salah diprediksi sebagai netral dan 11 sebagai positif. Pada kelas netral, dari 173 data aktual, hanya 96 yang berhasil diklasifikasikan dengan benar, sementara 42 data salah diprediksi sebagai negatif dan 35 sebagai positif. Pada kelas positif, dari 295 data aktual, sebanyak 227 berhasil diklasifikasikan dengan benar, sedangkan 27 salah diprediksi sebagai negatif dan 41 sebagai netral.

Tabel 5. Hasil Klasifikasi Random Forest

Kelas	*Precision*	*Recall*	*F1-Score*
Positif	0,83	0,77	0,80
Netral	0,51	0,55	0,53
Negatif	0,78	0,80	0,79
Accuracy			**0,73**
Macro avg	**0,71**	**0,71**	**0,71**



Gambar 2. Confusion Matrix Random Forest

Serupa dengan *Logistic Regression*, kelas netral juga menjadi kelas dengan performa paling lemah pada *Random Forest*, dengan *precision* 0,51, *recall* 0,55, dan *F1-score* 0,53. Kelas positif mencatatkan *precision* 0,83, *recall* 0,77, dan *F1-score* 0,80, sedangkan kelas negatif menghasilkan *precision* 0,78, *recall* 0,80, dan *F1-score* 0,79. Secara keseluruhan, *Random Forest* mencatatkan *macro average precision, recall,* dan *F1-score* masing-masing sebesar 0,71, lebih rendah dibandingkan *Logistic Regression* pada semua metrik.

3.6. Perbandingan Kinerja Model

Berdasarkan hasil evaluasi, *Logistic Regression* secara konsisten mengungguli *Random Forest* pada seluruh metrik evaluasi. *Logistic Regression* mencatatkan akurasi 0,76 berbanding 0,73 milik *Random Forest*, dengan selisih *macro F1-score* sebesar 0,03 (0,74 vs 0,71). Keunggulan ini konsisten pada setiap kelas sentimen, di mana *Logistic Regression* menghasilkan *F1-score* yang lebih tinggi untuk kelas positif (0,83 vs 0,80), kelas netral (0,58 vs 0,53), maupun kelas negatif (0,82 vs 0,79).

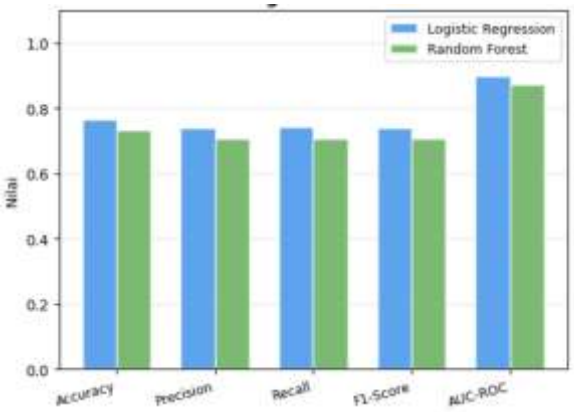
Tabel 6. Perbandingan Kinerja Keseluruhan Model

Metrik	*Logistic Regression*	*Random Forest*

DOI: 10.52362/jisamar.v10i3.2515

Journal of Information System, Applied, Management, Accounting and Research.
http://journal.stmikjayakarta.ac.id/index.php/jisamar ,
jisamar@stmikjayakarta.ac.id , jisamar2017@gmail.com
e-ISSN: 2598-8719 (Online), p-ISSN: 2598-8700 (Printed) , Vol. 10 No.3 (August 2026)

Accuracy	0,76	0,73
Precision	0,74	0,71
Recall	0,74	0,71
F1-Score	0,74	0,71



Gambar 3. Perbandingan Matrik Evaluasi Setiap Model

Tabel 7. Perbandingan Kinerja Keseluruhan Model per Kelas

Metrik	***Logistic Regression***			***Random Forest***		
	Prec.	**Recall**	**F1**	**Prec.**	**Recall**	**F1**
Precision	0,85	0,80	0,83	0,83	0,77	0,80
Recall	0,56	0,60	0,58	0,51	0,55	0,53
F1-Score	0,81	0,82	0,82	0,78	0,80	0,79

Keunggulan *Logistic Regression* pada penelitian ini sejalan dengan temuan beberapa studi terdahulu pada domain teks berbahasa Indonesia, seperti yang dilaporkan oleh Simanjuntak et al. [12], Junianto et al. [13], dan Sulistyawati dan Prabowo [4] yang juga menemukan *Logistic Regression* lebih unggul dari *Random Forest* pada data teks media sosial berbahasa Indonesia. Hal ini mengindikasikan bahwa untuk klasifikasi teks dengan representasi TF-IDF yang menghasilkan fitur berdimensi tinggi namun sparse, pendekatan linear seperti *Logistic Regression* cenderung lebih efektif dibandingkan model berbasis pohon seperti *Random Forest*. Sementara itu, pola kelas netral yang konsisten menjadi yang paling sulit diklasifikasikan oleh kedua model menunjukkan

DOI: 10.52362/jisamar.v10i3.2515

bahwa batas semantik antara opini netral dengan positif maupun negatif pada komentar TikTok memang inherently ambigu, dan ini merupakan tantangan yang perlu diatasi pada penelitian selanjutnya, misalnya melalui penggunaan representasi fitur berbasis konteks seperti BERT bahasa Indonesia.

IV. KESIMPULAN

Penelitian ini membandingkan kinerja Logistic Regression dan Random Forest berbasis TF-IDF untuk mengklasifikasikan sentimen komentar TikTok terhadap Program Makan Bergizi Gratis ke dalam tiga kelas, yaitu positif, netral, dan negatif. Dari 4.811 komentar yang dikumpulkan, sebanyak 3.884 komentar berhasil dilabeli setelah melewati tahap preprocessing dan pelabelan manual, dengan distribusi sentimen negatif sebagai yang terbanyak (39,73%), diikuti positif (37,97%), dan netral (22,30%).

Hasil evaluasi menunjukkan bahwa *Logistic Regression* secara konsisten mengungguli *Random Forest* pada seluruh metrik. *Logistic Regression* mencatatkan akurasi 0,76 dan *macro F1-score* 0,74, sementara *Random Forest* menghasilkan akurasi 0,73 dan *macro F1-score* 0,71. Keunggulan *Logistic Regression* berlaku pada setiap kelas sentimen, dan pola ini sejalan dengan sejumlah studi terdahulu pada domain klasifikasi teks berbahasa Indonesia yang menemukan hasil serupa [4], [12], [13]. Temuan ini mengindikasikan bahwa pendekatan linear seperti *Logistic Regression* lebih sesuai untuk data teks berrepresentasi TF-IDF yang bersifat berdimensi tinggi dan sparse dibandingkan model berbasis ensemble pohon keputusan.

Pada kedua model, kelas netral secara konsisten menjadi kelas dengan performa paling rendah, yang mengindikasikan bahwa batas semantik antara opini netral dengan positif maupun negatif pada komentar TikTok bersifat ambigu dan sulit ditangkap oleh representasi fitur berbasis frekuensi kata. Untuk penelitian selanjutnya, disarankan untuk mengeksplorasi penggunaan representasi fitur berbasis konteks seperti IndoBERT yang lebih mampu menangkap makna semantik kalimat secara keseluruhan, menambahkan fitur leksikal berbasis kamus sentimen bahasa Indonesia, serta mempertimbangkan pendekatan pelabelan dengan lebih dari dua anotator untuk meningkatkan kualitas ground truth, terutama pada komentar-komentar yang bersifat ambigu.

REFERENASI

[1] D. A. Sulistyo and E. Setiadi, “Analisis Sentimen Kebijakan Makan Bergizi Gratis Menggunakan IndoBERT dan Machine Learning,” *Jurnal FASILKOM (Teknologi Informasi dan Ilmu Komputer)*, vol. 15, no. 3, pp. 507–513, Dec. 2025, doi: 10.37859/jf.v15i3.10546.

[2] “Digital 2026: Top digital and social media trends in Indonesia,” We Are Social. [Online]. Available: https://wearesocial.com/id/blog/2025/11/digital-2026-top-digital-and-social-media-trends-in-indonesia/

[3] S. Yusra and R. A. Putri, “Klasifikasi Stance Opini Publik Komentar TikTok Kasus Guru Menampar Murid memakai TF-IDF dan SVM,” *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 7, no. 3, pp. 910–921, 2026, doi: 10.30865/json.v7i3.9544.

[4] A. Sulistyawati and D. W. Prabowo, “Analisis Sentimen Komentar Pengguna Tiktok terhadap Konten Fashion Brand Clotiva dengan Random Forest, Logistic Regression, dan Support Vector Machine,” *Journal of Community Service and Research (JURAGAN)*, vol. 3, no. 1, pp. 2860–2872, Mar. 2026, doi: 10.62710/v6exn887.

[5] “Pemerintah Salurkan Makan Bergizi Gratis (MBG), Ini Sasaran Utama Penerimanya,” Media Keuangan. [Online]. Available: https://mediakeuangan.kemenkeu.go.id/article/show/pemerintah-salurkan-makan-bergizi-gratis-mbg-ini-sasaran-utama-penerimanya

[6] “Anggaran Makan Bergizi Gratis Rp 335 Triliun, untuk Intervensi Gizi hingga Digitalisasi,” Tempo. [Online]. Available: https://www.tempo.co/politik/anggaran-makan-bergizi-gratis-rp-335-triliun-untuk-intervensi-gizi-hingga-digitalisasi-2060952

[7] R. A. Munir, “Analisis Sentimen Cuitan di Media Sosial X tentang Program Makan Bergizi Gratis dengan Metode NLP,” *Jurnal Informatika Dan Teknik Elektro Terapan*, vol. 13, no. 3, pp. 589–595, Jul. 2025, doi: 10.23960/jitet.v13i3.6912.

[8] M. A. Hasan and N. P. Bimby, “Analisis Sentimen Publik Terhadap Kenaikan Pajak PPN di Indonesia Tahun 2024 Menggunakan Algoritma Machine Learning,” *Jurnal FASILKOM (Teknologi Informasi dan Ilmu Komputer)*, vol. 15, no. 1, pp. 179–184, Apr. 2025, doi: 10.37859/jf.v15i1.8556.

DOI: 10.52362/jisamar.v10i3.2515

[9] M. Samantri and Afiyati, "Perbandingan Algoritma Support Vector Machine dan Random Forest untuk Analisis Sentimen Terhadap Kebijakan Pemerintah Indonesia Terkait Kenaikan Harga BBM Tahun 2022," *Jurnal Teknologi Informasi Dan Komunikasi (JTIK)*, vol. 8, no. 1, pp. 1–9, Jan. 2024, doi: 10.35870/jtik.v8i1.1202.
[10] S. Butsianto and A. M. Rifa'i, "Analisis Sentimen Ulasan Aplikasi Jamsostek dengan SVM, Random Forest, dan Logistic Regression," *Jurnal Informatika Ekonomi Bisnis*, vol. 7, no. 3, pp. 700–706, Sep. 2025, doi: 10.37034/infeb.v7i3.1266.
[11] L. N. Afifah, S. Rahayu, and Purwadi, "Sentiment Analysis of TikTok User Comments on The Free Nutritious Meal Program Using Support Vector Machine," *Journal of Artificial Intelligence and Engineering Applications (JAIEA)*, vol. 5, no. 2, pp. 2366–2371, Feb. 2026, doi: 10.59934/jaiea.v5i2.1879.
[12] W. O. Simanjuntak, A. B. P. Negara, and R. Septriana, "Perbandingan Algoritma Logistic Regression dan Random Forest (Studi Kasus : Klasifikasi Emosi Tweet)," *Jurnal Aplikasi dan Riset Informatika (JUARA)*, vol. 2, no. 1, pp. 160–164, Aug. 2023, doi: 10.26418/juara.v2i1.69682.
[13] H. Junianto, R. E. Saputro, B. A. Kusuma, and D. I. S. Saputra, "Comparison of Logistic Regression and Random Forest in Sentiment Analysis of Disdukcapil Application Reviews," *Jurnal Teknik Informatika (JUTIF)*, vol. 5, no. 6, pp. 1539–1547, Feb. 2024, doi: 10.52436/1.jutif.2024.5.6.1802.
[14] A. M. Wahid, Turino, K. A. Nugroho, T. Safitri, and F. S. Utomo, "Optimasi Logistic Regression dan Random Forest untuk Deteksi Berita Hoax Berbasis TF-IDF," *Jurnal Pendidikan dan Teknologi Indonesia (JPTI)*, vol. 4, no. 8, pp. 381–392, Aug. 2024, doi: 10.52436/1.jpti.602.
[15] R. Couronné, P. Probst, and A.-L. Boulesteix, "Random forest versus logistic regression: a large-scale benchmark experiment," *BMC Bioinformatics*, vol. 19, no. 1, p. 270, Jul. 2018, doi: 10.1186/s12859-018-2264-5.
[16] A. Saepudin, A. Faqih, and G. Dwilestari, "Perbandingan Algoritma Klasifikasi Support Vector Machine, Random Forest dan Logistic Regression Pada Ulasan Shopee," *Jurnal TEKNO KOMPAK*, vol. 18, no. 1, pp. 178–192, Feb. 2024, doi: 10.33365/jtk.v18i1.3764.
[17] S. E. Prianto, Berlilana, and R. E. Saputro, "Analisis Sentimen Program Makan Bergizi Gratis Menggunakan Random Forest dan Support Vector Machine dengan Penyeimbangan Data Berbasis SMOTE," *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 6, no. 2, pp. 353–365, Mar. 2026.
[18] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Morgan Kaufmann, 2012.
[19] B. Liu, *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers, 2012.
[20] P. A. Permatasari, L. Linawati, and L. Jasa, "Survei Tentang Analisis Sentimen Pada Media Sosial," *Majalah Ilmiah Teknologi Elektro*, vol. 20, no. 2, pp. 177–186, Dec. 2021, doi: 10.24843/mite.2021.v20i02.p01.
[21] M. Siino, I. Tinnirello, and M. La Cascia, "Is text preprocessing still worth the time? A comparative survey on the influence of popular preprocessing methods on Transformers and traditional classifiers," *Inf. Syst.*, vol. 121, Mar. 2024, doi: 10.1016/j.is.2023.102342.
[22] V. W. D. Thomas and F. Rumaisa, "Analisis Sentimen Ulasan Hotel Bahasa Indonesia Menggunakan Support Vector Machine dan TF-IDF," *Jurnal Media Informatika Budidarma*, vol. 6, no. 3, pp. 1767–1774, Jul. 2022, doi: 10.30865/mib.v6i3.4218.
[23] B. Ramadhani and Suryono, "Komparasi Algoritma Naïve Bayes dan Logistic Regression Untuk Analisis Sentimen Metaverse," *Jurnal Media Informatika Budidarma*, vol. 8, no. 2, pp. 714–725, Apr. 2024, doi: 10.30865/mib.v8i2.7458.
[24] M. U. Albab, Y. K. P., and M. N. Fawaiq, "Optimization of the Stemming Technique on Text Preprocessing President 3 Periods Topic," *Jurnal Transformatika*, vol. 20, no. 2, pp. 1–12, Jan. 2023, doi: 10.26623/transformatika.v20i2.5374.
[25] R. Wati and H. Rachmi, "Pembobotan TF-IDF Menggunakan Naïve Bayes Pada Sentimen Masyarakat Mengenai Isu Kenaikan BIPIH," *Jurnal Manajemen Informatika (JAMIKA)*, vol. 13, no. 1, pp. 84–93, Apr. 2023, doi: 10.34010/jamika.v13i1.9424.
[26] M. H. Mahendra, D. T. Murdiansyah, and K. M. Lhaksamana, "Analisis Sentimen Tweet COVID-19 Menggunakan Metode K-Nearest Neighbors dengan Ekstraksi Fitur TF-IDF dan CountVectorizer," *Jurnal Ilmu Multidisiplin*, vol. 1, no. 2, pp. 37–43, Aug. 2023, doi: 10.69688/dike.v1i2.35.
[27] N. V Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.

DOI: 10.52362/jisamar.v10i3.2515

[28] J. L. Leevy, T. M. Khoshgoftaar, R. A. Bauder, and N. Seliya, "A survey on addressing high-class imbalance in big data," *J. Big Data*, vol. 5, no. 1, Dec. 2018, doi: 10.1186/s40537-018-0151-6.
[29] Ash Shiddicky and S. Agustian, "Analisis Sentimen Masyarakat Terhadap Kebijakan Vaksinasi Covid-19 pada Media Sosial Twitter menggunakan Metode Logistic Regression," *Jurnal Computer Science and Information Technology (CoSciTech)*, vol. 3, no. 2, pp. 99–106, Aug. 2022, doi: 10.37859/coscitech.v3i2.3836.
[30] A. A. Putri, S. Agustian, and R. Abdillah, "Penerapan Metode Logistic Regression untuk Klasifikasi Sentimen pada Dataset Twitter Terbatas," *Jurnal Sistem Informasi (ZONAsi)*, vol. 7, no. 1, pp. 95–107, Jan. 2025, doi: 10.31849/zn.v7i1.24804.
[31] O. Shobayo, S. Adeyemi-Longe, O. Popoola, and B. Ogunleye, "Innovative Sentiment Analysis and Prediction of Stock Price Using FinBERT, GPT-4 and Logistic Regression: A Data-Driven Approach," *Big Data and Cognitive Computing*, vol. 8, no. 11, Nov. 2024, doi: 10.3390/bdcc8110143.
[32] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer, 2006.
[33] L. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

DOI: 10.52362/jisamar.v10i3.2515