Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9458

Analisis Sentimen Publik Terhadap Makan Bergizi Gratis Menggunakan Bi-LSTM dan IndoBERT

Almira Putri Wibowo*, Dedi Gunawan
Teknik Informatika, Fakultas Komunikasi dan Informatika, Universitas Muhammadiyah Surakarta, Surakarta, Indonesia
Email: 1,*l200220068@student.ums.ac.id, 2dedi.gunawan@ums.ac.id
Email Penulis Korespondensi: l200220068@student.ums.ac.id*
Submitted: **07/01/2026**; Accepted: **22/01/2026**; Published: **31/03/2026**

Abstrak– Program Makan Bergizi Gratis (MBG) yang direncanakan pemerintah pada tahun 2024 menjadi perbincangan luas di media sosial, khususnya Twitter dan TikTok, dengan beragam opini yang mencerminkan persepsi publik terhadap implementasinya. Penelitian ini bertujuan menganalisis sentimen masyarakat terhadap Program Makan Bergizi Gratis menggunakan pendekatan deep learning. Data dikumpulkan melalui metode crawling dan scraping dari Twitter dan TikTok, menghasilkan 7.341 data setelah proses data cleaning. Pelabelan dilakukan menggunakan dua pendekatan, yaitu pelabelan manual dan pelabelan otomatis berbasis IndoBERTweet. Penelitian ini membandingkan kinerja dua model, yaitu Bidirectional Long Short-Term Memory (Bi-LSTM) dan IndoBERT, dengan variasi learning rate 3e-5 dan 5e-5. Hasil eksperimen menunjukkan bahwa IndoBERT secara konsisten menghasilkan performa yang lebih baik dibandingkan Bi-LSTM pada kedua metode pelabelan. Konfigurasi optimal diperoleh pada pelabelan otomatis dengan learning rate 3e-5 yang mencapai akurasi 82% dan F1-Score 79%. Hasil analisis juga menunjukkan bahwa sentimen publik terhadap Program Makan Bergizi Gratis didominasi oleh sentimen negatif. Selain itu, dikembangkan aplikasi web berbasis Flask untuk prediksi sentimen secara real-time. Penelitian ini diharapkan dapat menjadi acuan dalam pengembangan analisis sentimen bahasa Indonesia yang lebih efektif dan akurat.
Kata Kunci: Analisis Sentimen; Makan Bergizi Gratis; Bi-LSTM; IndoBERT; Deep Learning; Flask

Abstract– The Free Nutritious Meals Program (MBG) planned by the government in 2024 has become a hot topic on social media, particularly Twitter and TikTok, with a variety of opinions reflecting public perceptions of its implementation. This study aims to analyze public sentiment towards the Free Nutritious Meal Program using a deep learning approach. Data was collected through crawling and scraping methods from Twitter and TikTok, resulting in 7,341 data points after data cleaning. Labeling was done using two approaches, namely manual and automatic based on IndoBERTweet. This study compares the performance of two models, namely Bidirectional Long Short-Term Memory (Bi-LSTM) and IndoBERT, with learning rate variations of 3e-5 and 5e-5. The results of the experiment show that IndoBERT consistently produces better performance than Bi-LSTM in both labeling methods. The optimal configuration was obtained in automatic labeling with a learning rate of 3e-5, which achieved an accuracy of 82% and an F1-Score of 79%. The analysis also shows that public sentiment towards the Free Nutritious Meals Program is dominated by negative sentiment. In addition, a Flask-based web application was developed for real-time sentiment prediction. This research is expected to serve as a reference for the development of more effective and accurate Indonesian language sentiment analysis.

Keywords: Sentiment Analysis; Free Nutritious Meals; Bi-LSTM; IndoBERT; Deep Learning; Flask

1. PENDAHULUAN

Indonesia saat ini masih menghadapi permasalahan serius dalam bidang Kesehatan masyarakat, khususnya terkait tingginya angka stunting pada balita dan anak-anak. Stunting merupakan kondisi gagal tumbuh akibat kekurangan gizi kronis yang terjadi pada waktu lama. Berdasarkan data Survei Status Gizi Indonesia (SSGI) 2024, prevalensi stunting turun mencapai 19,8% berada sedikit di bawah standar WHO yaitu tidak lebih dari 20%. Menurut Kementerian Kesehatan Republik Indonesia (Kemenkes RI) pemerintah menargetkan angka stunting di Indonesia 14,2% pada tahun 2029. Stunting yang tidak terdeteksi atau bahkan sampai dibiarkan tanpa diatasi akan menimbulkan efek dalam jangka pendek maupun efek jangka panjang yang bisa meningkatkan kesakitan, kematian, bahkan menurunnya kemampuan kognitif dan motorik anak [1].

Sebagai strategi dalam pemenuhan gizi pada anak-anak untuk mencapai target tersebut dan upaya dalam meningkatkan kualitas sumber daya manusia di masa depan, Presiden Prabowo Subianto dan Gibran Rakabuming Raka meluncurkan Program Makanan Bergizi Gratis (MBG) [2]. Program ini menyasar anak-anak sekolah, pesantren, dan ibu hamil untuk memastikan pemenuhan gizi merata, sekaligus memberikan perlindungan sosial bagi masyarakat rentan terhadap kemiskinan dan ketidaksetaraan [3].

Namun, pelaksanaan Program MBG menimbulkan beragam opini dari masyarakat, terutama pada media sosial. Sosial media merupakan wadah yang dimanfaatkan oleh masyarakat untuk mengungkapkan segala sesuatu [4]. Twitter dan TikTok menjadi sarana utama masyarakat menyampaikan opini secara terbuka dan cepat. Beberapa masyarakat mendukung program karena memperbaiki gizi anak, sementara yang lain mengkritik anggaran dan kasus keracunan, sehingga diperlukan analisis lebih lanjut untuk memahami opini publik secara mendalam [5].

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9458

Analisis sentimen digunakan untuk memahami bagaimana masyarakat merespons Program MBG, apakah mereka menilai program tersebut bermanfaat atau tidak. [6]. Proses ini secara otomatis mengklasifikasikan teks menjadi sentimen positif, negatif, atau netral, sehingga efisien untuk menangani data dalam jumlah besar [7]. Tujuannya adalah memahami opini dan sikap masyarakat terhadap Program Makan Bergizi Gratis, karena pandangan mereka dapat mempengaruhi tingkat dukungan terhadap kebijakan pemerintah.

Berbagai penelitian sebelumnya menunjukkan bahwa pendekatan *deep learning* lebih efektif daripada algoritma *machine learning* tradisional dalam analisis sentimen berbahasa Indonesia. Misalnya, SVM berbasis TF-IDF pada analisis sentimen Program MBG hanya mencapai akurasi 67%, menunjukkan keterbatasan metode klasik dalam menangkap konteks bahasa secara mendalam, terutama akibat ketidakseimbangan data yang menyebabkan bias terhadap kelas mayoritas [8]. Penelitian lain yang membandingkan KNN dan Naive Bayes menunjukkan bahwa Naive Bayes memiliki performa lebih baik dengan akurasi 79,61% dan F1-score 77,33%, namun masih terbatas dalam hal generalisasi dan pemahaman konteks semantik, sehingga direkomendasikan penggunaan pendekatan *deep learning* [9].

Deep learning menjadi alternatif unggulan dibanding *machine learning* tradisional karena dapat mempelajari fitur secara otomatis, fleksibel, dan skalabel, sehingga efektif untuk menangani data besar dan masalah kompleks seperti analisis sentimen serta pengolahan bahasa alami [10]. Berdasarkan kajian tersebut, Bi-LSTM dan IndoBERT dipilih sebagai arsitektur yang relevan untuk dianalisis lebih lanjut. Bi-LSTM banyak digunakan dalam analisis sentimen karena kemampuannya memproses urutan teks dari dua arah, yang memungkinkan pemahaman konteks kata secara lebih komprehensif. Hal ini dibuktikan pada penelitian analisis sentimen ulasan pengguna aplikasi My Pertamina, di mana Bi-LSTM mencapai akurasi hingga 90%, melampaui LSTM konvensional [11]. Adapun, IndoBERT, sebagai model berbasis transformer yang dilatih khusus pada korpus bahasa Indonesia, juga menunjukkan performa yang sangat baik dalam analisis sentimen. Penelitian mengenai sentimen publik terhadap kebijakan kenaikan PPN 12% menunjukkan bahwa IndoBERT mampu mencapai akurasi 84,94%, yang menegaskan keunggulannya dalam menangkap konteks bahasa Indonesia yang kompleks dan beragam [12].

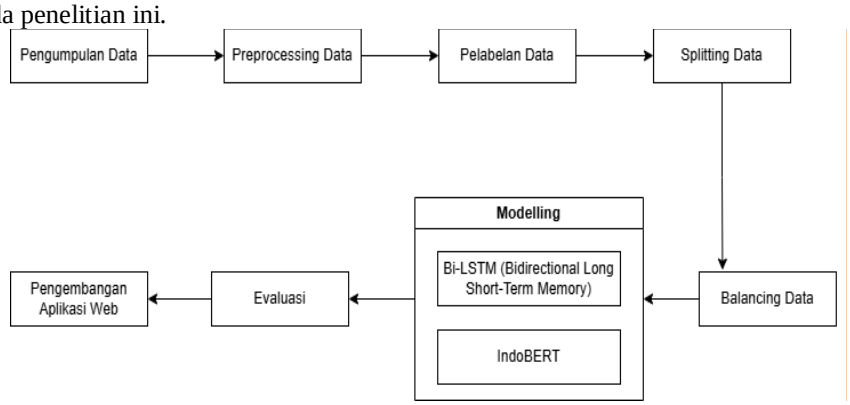
Selain itu, penelitian yang membandingkan Naive Bayes dan Bi-LSTM pada ulasan BSI Mobile memperlihatkan bahwa Bi-LSTM lebih unggul dengan akurasi 94,90% dibandingkan Naive Bayes 94%, dan penerapan *oversampling* pada kelas minoritas meningkatkan performa model secara signifikan [13]. Studi lain pada sentimen publik terkait Pemutusan Hubungan Kerja (PHK) di Indonesia membandingkan IndoBERT, SVM, Random Forest, dan Decision Tree, menunjukkan IndoBERT unggul dengan akurasi 89,6% [14].

Walaupun demikian, sebagian besar studi masih terbatas pada dataset tertentu dan jarang membandingkan langsung Bi-LSTM dan IndoBERT dalam analisis sentimen kebijakan publik, sehingga penelitian ini mengevaluasi kinerja kedua model pada data multi platform (Twitter dan TikTok) dengan penerapan dua metode *labeling*, yaitu pelabelan manual dan otomatis, serta teknik *Random Oversampling* untuk meningkatkan akurasi kelas minoritas dan memberikan wawasan baru serta lebih komprehensif mengenai opini masyarakat terhadap Program Makan Bergizi Gratis melalui media sosial.

Fokus penelitian ini adalah menganalisis persepsi publik terhadap Program MBG serta mengevaluasi kinerja Bi-LSTM dan IndoBERT dalam mengklasifikasikan sentimen. Tujuannya adalah membandingkan efektivitas kedua arsitektur dalam analisis sentimen kebijakan publik. Hasil penelitian ini berkontribusi pada pengembangan kajian *Natural Language Processing* (NLP) berbahasa Indonesia, khususnya analisis sentimen berbasis *deep learning*, melalui gambaran empiris mengenai efektivitas model dalam memahami karakteristik bahasa Indonesia yang bersifat informal.

2. METODOLOGI PENELITIAN

Pada penelitian ini dilaksanakan melalui serangkaian tahapan yang ditampilkan pada Gambar 1. Setiap tahapan dijelaskan dengan sistematis agar memberikan pandangan yang jelas mengenai proses analisis sentimen yang dilakukan pada penelitian ini.



Gambar 1. Alur Analisis Sentimen

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9458

2.1 Pengumpulan Data

Data yang digunakan pada penelitian ini berasal dari platform Twitter dan TikTok, khususnya dari komentar masyarakat mengenai Program Makan Bergizi Gratis. Kedua platform tersebut dipilih karena sering digunakan masyarakat dalam menyampaikan pendapat mengenai isu-isu tertentu. Pengumpulan data dilakukan menggunakan *Python* untuk proses *crawling* data Twitter dan *scraping* data TikTok.

2.2 Preprocessing Data

Preprocessing adalah proses memperbaiki data mentah agar menjadi data bersih yang siap digunakan untuk analisis atau pelatihan model. Pada *preprocessing* ini memiliki beberapa tahapan dalam memperbaiki data, seperti berikut ini.

2.2.1 Cleaning

Data *cleaning* merupakan proses meningkatkan kualitas data dengan menghapus atau memperbaiki data yang tidak relevan, tidak lengkap, atau tidak akurat. Tahapan yang dilakukan meliputi penghapusan tanda baca, angka, URL, hashtag, emoji, serta data kosong [15]. Selain itu, seluruh teks diubah menjadi huruf kecil *(case folding)* dan data duplikat dihapus untuk memastikan kualitas data yang digunakan dalam proses pemodelan

2.2.2 Normalisasi

Normalisasi adalah proses *preprocessing* untuk mengubah kata-kata tidak baku atau slang menjadi bentuk sesuai kaidah bahasa Indonesia [16].

2.2.3 Tokenisasi

Tokenisasi adalah proses memisahkan teks menjadi potongan kecil berupa token seperti kata, kalimat atau huruf untuk mengubah teks mentah menjadi kumpulan token sehingga dapat dianalisis lebih lanjut [17].

2.2.4 Stemming

Stemming adalah proses penghapusan imbuhan seperti awalan, akhiran atau sisipan agar kata menjadi bentuk dasar [18].

2.3 Pelabelan

Pelabelan adalah proses pemberian label pada data dengan tujuan untuk mengklasifikasikan setiap sentimen ke dalam kategori positif, negatif, atau netral [19]. Penelitian ini menggunakan dua jenis pelabelan yaitu pelabelan manual dan pelabelan otomatis. Pelabelan manual dilakukan untuk menghasilkan dataset berlabel manual, sedangkan pelabelan otomatis dilakukan menggunakan *IndoBERTweet* untuk menghasilkan dataset berlabel otomatis. Kedua dataset tersebut digunakan secara terpisah dalam proses pelatihan model, sehingga penelitian ini memiliki dua skenario eksperimen berdasarkan metode pelabelan yang digunakan.

2.4 Splitting Data

Pada tahap ini, memisahkan data menjadi dua bagian, yaitu data *training* dan data *testing*. Tujuannya supaya model bisa belajar dari training set yang kemudian diuji pada data lain (*testing*) guna melihat seberapa baik model dalam melakukan prediksi data baru. Pembagian komposisi data nya yaitu, 80% untuk data training dan 20% untuk data testing.

2.5 Balancing Data

Balancing data adalah proses penyeimbangan jumlah data antar kelas agar tidak bias terhadap kelas tertentu. Ketidakseimbangan kelas dapat mempengaruhi performa model, karena tanpa penyeimbangan, model akan cenderung bias terhadap kelas mayoritas, sehingga *recall* pada kelas yang minoritas menjadi rendah [20].

Pada penelitian ini, dataset memiliki distribusi kelas yang tidak seimbang dan didominasi oleh sentimen negatif, sehingga berpotensi menyebabkan bias terhadap kelas minoritas. Oleh karena itu, *Random Oversampling* dipilih untuk meningkatkan jumlah data pada kelas minoritas tanpa menghilangkan data asli, sehingga distribusi data antar kelas menjadi lebih seimbang. Pemilihan metode ini karena pada pada penelitian [21] menunjukkan bahwa penerapan *Random Oversampling* pada model IndoBERT mampu meningkatkan kinerja analisis sentimen bahasa Indonesia pada dataset yang tidak seimbang dengan memperbaiki representasi kelas minoritas dalam konteks deep learning NLP.

2.6 Modelling

Pada penelitian ini menggunakan dua model deep learning, yaitu Bi-LSTM (*Bidirectional Long Short-Term Memory*) dan IndoBERT, untuk melakukan analisis sentimen. Pemilihan kedua model ini didasarkan pada kemampuannya dalam menangkap konteks bahasa secara mendalam, namun dengan pendekatan arsitektur yang berbeda.

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9458

2.7.1 Bi-LSTM (Bidirectional Long Short-Term Memory)

Bi-LSTM merupakan salah satu model *deep learning* berasal dari pengembangan LSTM yang memungkinkan untuk menangkap konteks dari dua arah [22]. Berbeda dengan LSTM konvensional yang hanya memproses urutan satu arah, Bi-LSTM memungkinkan pemanfaatan informasi dari masa lalu dan masa depan secara bersamaan sehingga konteks yang diperoleh menjadi lebih lengkap. Proses pelatihan Bi-LSTM pada dasarnya serupa dengan LSTM, namun diperlukan langkah tambahan pada proses *backpropagation through time* karena pembaruan bobot input dan output tidak dapat dilakukan secara bersamaan [23].

Arsitektur Bi-LSTM yang digunakan pada penelitian ini terdiri atas *embedding layer*, *Bi-LSTM layer*, *dropout layer*, dan *dense layer*. *Embedding layer* berfungsi merepresentasikan teks ke dalam bentuk vektor numerik berdimensi tetap. Selanjutnya, representasi tersebut diproses oleh *Bi-LSTM layer* secara dua arah. *Dropout layer* diterapkan untuk mengurangi risiko *overfitting*, sedangkan *dense layer* digunakan untuk menghasilkan keluaran akhir berupa klasifikasi sentimen.

2.7.2 IndoBERT

IndoBERT merupakan model berbasis berbahasa tunggal yang dikembangkan berdasarkan arsitektur *Bidirectional Encoder Representation from Transformers* (BERT) dan dilatih dalam kumpulan teks bahasa Indonesia [24]. Cara kerja IndoBERT mirip dengan BERT yaitu dengan menggunakan *Transformer Bidirectional* untuk memahami konteks. IndoBERT terbukti memiliki pemrosesan data yang lebih baik dibandingkan BERT dalam memahami konteks Bahasa Indonesia [25].

Pada tahap pemodelan ini, digunakan model "IndoBERT-base-p1" yang telah melalui proses *pre-training*. Model tersebut kemudian dilakukan proses *fine-tuning* untuk tugas klasifikasi sentimen dengan menyesuaikan lapisan klasifikasi pada bagian akhir model. Proses *fine-tuning* memungkinkan IndoBERT memanfaatkan pengetahuan linguistik yang telah dipelajari sebelumnya sehingga mampu meningkatkan performa dalam mengklasifikasikan sentimen pada teks berbahasa Indonesia.

2.7.3 Konfigurasi Pelatihan Model

Pada penelitian ini, kedua model deep learning, yaitu Bi-LSTM dan IndoBERT, dilatih dengan pengaturan hyperparameter yang seragam untuk memastikan perbandingan performa yang objektif. *Batch size*, jumlah *epoch* maksimum, *learning rate*, dan mekanisme *early stopping* disamakan pada kedua model. Perbedaan hanya terdapat pada karakteristik arsitektur masing-masing, termasuk *optimizer*, fungsi *loss*, dan panjang maksimum sekuens, yang disesuaikan dengan kebutuhan Bi-LSTM dan IndoBERT. Konfigurasi detail *hyperparameter* untuk kedua model disajikan pada Tabel 1.

Tabel 1. Konfigurasi pelatihan

Parameter	Bi-LSTM	IndoBERT
Optimizer	Adam	AdamW
Loss function	Sparse Categorical Cross-Entropy	Cross-Entropy
Batch size	32	32
Learning rate	3e-5 dan 5e-5	3e-5 dan 5e-5
Epoch max	10	10
Early stopping patience	3	3
Token max length	50	128
LSTM units / Hidden size	128	-
Dense layer units	64	-
Dropout	0,3	-

2.7 Evaluasi

Tahap evaluasi dilakukan guna menilai kinerja model dalam mengklasifikasikan sentimen publik Program MBG serta menentukan model mana yang paling akurat dan efisien dalam memproses teks informal bahasa indonesia dengan menggunakan empat matrik evaluasi, yaitu *accuracy*, *precision*, *recall*, dan *f1-score* [26]. *Accuracy* digunakan untuk mengukur persentase prediksi yang benar dari keseluruhan data yang dilatih. *Precision* berguna untuk menentukan seberapa akurat model dalam prediksi kelas positif. *Recall* adalah penilaian seberapa banyak data positif dapat ditemukan oleh model. *F1-score* untuk mengukur keseimbangan antara *precision* dan *recall* terutama pada kondisi distribusi data yang tidak seimbang.

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9458

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$f1-score = \frac{(recall\ x\ precision)}{(precision+recall)} \quad (4)$$

Keterangan
TP : Jumlah data positif yang diprediksi positif oleh model
TN : Jumlah data negatif yang diprediksi negatif oleh model
FP : Jumlah data negatif tetapi diprediksi positif oleh model
FN : Jumlah data positif tetapi diprediksi negatif oleh model

2.8 Pengembangan Website Prediksi

Website dikembangkan sebagai bentuk penerapan dari model klasifikasi sentimen yang sebelumnya telah training dan dievaluasi. Hasil pemodelan akan disimpan dalam formal pkl, kemudian dimuat ulang dalam website untuk melakukan prediksi secara langsung terhadap input pengguna. Melalui website ini, pengguna dapat memasukkan teks dan memperoleh hasil prediksi klasifikasi sentimen berdasarkan model yang telah dilatih sebelumnya.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Penelitian ini menggunakan dua sumber data, yaitu dari Twitter dan TikTok. Pengambilan data dilakukan dengan metode *Crawling* pada platform Twitter dan *Scraping* pada platform TikTok. Data Twitter dikumpulkan berdasarkan enam kata kunci dengan rentang waktu antara 6 Januari 2024 hingga 1 November 2025, sedangkan data TikTok diperoleh dari video berita terbaru yang membahas MBG. Proses ini menghasilkan 4.633 data Tweet dan 5.425 TikTok. Selanjutnya, kolom 'full_text' dari kedua dataset digabungkan menjadi satu dataset untuk digunakan pada tahap data *cleaning*.

3.2 Preprocessing

Tahap *preprocessing* dilakukan untuk membersihkan dan menyiapkan dataset sebelum digunakan pada proses analisis sentimen. Proses ini meliputi tahapan data *cleaning*, normalisasi, tokenisasi, dan *stemming*.

3.2.1 Data Cleaning

Hasil proses data cleaning menunjukkan bahwa dari 10.058 data mentah, sebanyak 7.341 data dinyatakan layak untuk diproses. Tahapan ini meliputi penghapusan tanda baca, angka, URL, hashtag, emoji, baris kosong, serta data duplikat, dan normalisasi teks ke huruf kecil (*case folding*) untuk menjaga konsistensi data. Proses ini bertujuan mengurangi *noise* dan potensi bias yang dapat mempengaruhi performa model analisis sentimen. Hasil data cleaning ditampilkan pada Tabel 2 yang memperlihatkan perbandingan data sebelum dan sesudah pembersihan, di mana elemen non informatif berhasil dihilangkan sehingga teks menjadi lebih bersih dan mendukung penangkapan konteks semantik pada tahap pemodelan selanjutnya.

Tabel 2. Contoh Data Cleaning

Sebelum	Sesudah
Ya ampun ini mah yang buat mbg protein di tubuhnya dah terdenaturasi semua ajig ALIAS GA BISA MIKIR LU SEMUA TOLOL	ya ampun ini mah yang buat mbg protein di tubuhnya dah terdenaturasi semua ajig alias ga bisa mikir lu semua tolol

3.2.2 Normalisasi

Pada tahap normalisasi, digunakan kamus normalisasi yang berisi pasangan kata tidak baku dan kata baku sesuai karakteristik dataset. Kamus ini diterapkan pada seluruh teks untuk menyeragamkan penggunaan kata. Hasil normalisasi disajikan pada Tabel 3 yang menunjukkan perubahan kata tidak baku dan singkatan menjadi bentuk baku, sehingga meningkatkan konsistensi data dan kualitas representasi teks untuk analisis sentimen.

Tabel 3. Contoh normalisasi

Sebelum	Normalisasi
ya ampun ini mah yang buat mbg protein di tubuhnya dah terdenaturasi semua ajig alias ga bisa mikir lu semua tolol	ya ampun ini memang yang untuk makan bergizi gratis protein di tubuhnya sudah

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9458

	terdenaturasi semua astaga alias tidak dapat berpikir kamu semua bodoh

3.2.3 Tokenisasi

Data hasil normalisasi diproses melalui tahap tokenisasi untuk memecah teks menjadi unit-unit kata (token) sebagai masukan pemodelan. Hasil tokenisasi disajikan pada Tabel 4 yang menunjukkan perubahan teks menjadi deretan token individual dengan urutan yang tetap, sehingga memudahkan ekstraksi fitur dan mempertahankan konteks semantik.

Tabel 4. Contoh tokenisasi

Normalisasi	Tokenisasi
ya ampun ini memang yang untuk makan bergizi gratis protein di tubuhnya sudah terdenaturasi semua astaga alias tidak dapat berpikir kamu semua bodoh	['ya', 'ampun', 'ini', 'memang', 'yang', 'untuk', 'makan', 'bergizi', 'gratis', 'protein', 'di', 'tubuhnya', 'sudah', 'terdenaturasi', 'semua', 'astaga', 'alias', 'tidak', 'dapat', 'berpikir', 'kamu', 'semua', 'bodoh']

3.2.4 Stemming

Tahap *stemming* dilakukan untuk menghilangkan imbuhan pada token kata sehingga diperoleh bentuk kata dasar yang lebih seragam menggunakan algoritma *stemming* Bahasa Indonesia. Hasil proses ini disajikan pada Tabel 5 yang menunjukkan perubahan token berimbuhan menjadi kata dasar, sehingga meningkatkan konsistensi representasi teks dan kualitas input pada tahap pemodelan sentimen.

Tabel 5. Contoh stemming

Tokenisasi	Stemming
['ya', 'ampun', 'ini', 'memang', 'yang', 'untuk', 'makan', 'bergizi', 'gratis', 'protein', 'di', 'tubuhnya', 'sudah', 'terdenaturasi', 'semua', 'astaga', 'alias', 'tidak', 'dapat', 'berpikir', 'kamu', 'semua', 'bodoh']	ya ampun ini memang yang untuk makan gizi gratis protein di tubuh sudah terdenaturasi semua astaga alias tidak dapat pikir kamu semua bodoh

3.3 Pelabelan

Tahap pelabelan dataset dilakukan menggunakan dua metode, yaitu pelabelan manual oleh peneliti dan pelabelan otomatis menggunakan *IndoBERTweet*. Pelabelan manual dilakukan dengan membaca teks dan menentukan label sentimen berdasarkan konteks kalimat, sedangkan pelabelan otomatis memanfaatkan model *transformer* yang telah dilatih pada data Twitter berbahasa Indonesia untuk mengenali pola sentimen secara kontekstual dan konsisten. Hasil pelabelan ditampilkan pada Tabel 6 yang menunjukkan distribusi sentimen setelah penerapan kedua metode, di mana kelas negatif memiliki jumlah lebih banyak dibandingkan kelas positif dan netral. Distribusi ini perlu diperhatikan karena dapat mempengaruhi kinerja model, sehingga diperlukan penerapan teknik *balancing* agar proses klasifikasi tetap proporsional dan akurat pada ketiga kategori sentimen.

Tabel 6. Distribusi label

Metode	Positif	Negatif	Netral	Total
Manual	2.548	3.218	1.575	7.341
Otomatis	1.883	3.775	1.683	7.341

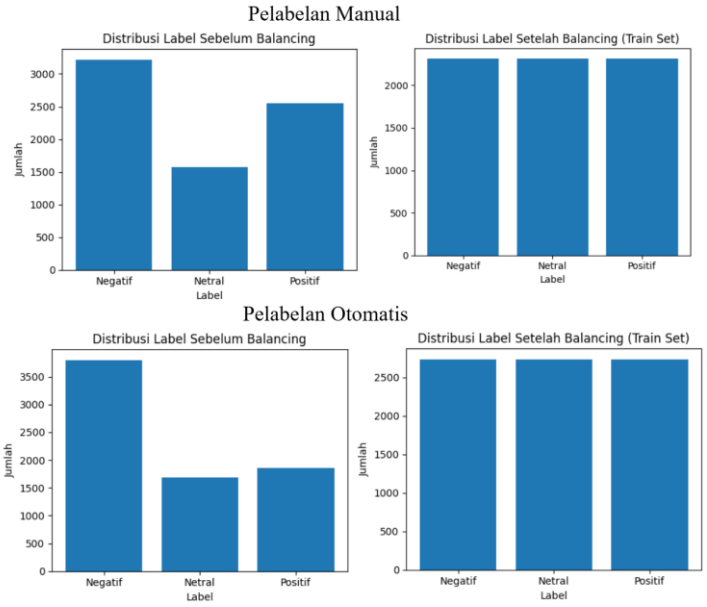
3.4 Splitting Data

Dataset yang telah bersih kemudian dibagi dengan metode *stratified* agar proporsi setiap label tetap seimbang pada data *training*, *validation* dan *testing*. Pembagian awal dilakukan dengan memisahkan 20% data untuk data *testing* dan 80% data untuk data *training*. Selanjutnya, sebanyak 10% dari data *training* digunakan untuk data *validation*. Sehingga komposisi akhir data menjadi 72% untuk data *training*, 8% untuk data *validation*, dan 20% untuk data *testing*.

3.5 Balancing Data

Hasil pelabelan dataset menunjukkan bahwa distribusi kelas cenderung tidak seimbang, dimana pada pelabelan manual maupun otomatis jumlah data berlabel negatif lebih dominan. Ketidakseimbangan ini berpotensi mempengaruhi kinerja model dan menyebabkan bias terhadap kelas mayoritas. Oleh karena itu, dilakukan penyeimbangan data pada data latih dengan menerapkan metode *Random Oversampling*, yaitu dengan menambah jumlah sampel pada kelas minoritas melalui penggandaan data yang ada, sehingga mengurangi bias pada kelas mayoritas seperti gambar 2.

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9458



Gambar 2. Distribusi Label Sebelum dan Sesudah Balancing

3.6 Modelling

Pelatihan kedua model deep learning, yaitu Bi-LSTM dan IndoBERT, dilakukan sesuai dengan konfigurasi *hyperparameter* yang telah ditetapkan sebelumnya, termasuk *batch size*, jumlah *epoch* maksimum, *learning rate*, serta mekanisme *early stopping* dengan *patience* 3. Pelatihan dilakukan dengan variasi metode pelabelan, yaitu manual dan otomatis, serta dua nilai *learning rate* berbeda (3e-5 dan 5e-5) untuk menilai pengaruh kombinasi parameter tersebut terhadap proses *training*. Tabel 7 menyajikan hasil *training* secara lengkap, meliputi akurasi dan *loss* pada data *training* dan validasi, jumlah *epoch* terbaik, waktu pelatihan, penggunaan memori, serta penerapan *early stopping*.

Tabel 7. Hasil training

Model	Pelabelan	Learning Rate	Best Epoch	Train Acc	Val Acc	Train Loss	Val Loss	Time	Memory	Early Stop
Bi-LSTM	Manual	5e-5	10	0.7144	0.6548	0.6882	0.7977	805.55	**21.8**	10
	Otomatis	5e-5	9	0.7936	0.6963	0.5255	0.6692	1015.61	21.93	10
Bi-LSTM	Manual	3e-5	10	0.6015	0.5861	0.9124	0.9202	714.68	22.1	10
	Otomatis	3e-5	10	0.7237	0.6861	0.6925	0.7065	828.68	25.07	10
Indo BERT	Manual	5e-5	2	**0.8768**	0.7585	**0.3372**	0.6292	142.43	2668.35	5
	Otomatis	5e-5	1	0.8040	0.8146	0.4848	0.4797	168.11	2949.60	4
Indo BERT	Manual	3e-5	1	0.6973	0.7687	0.6851	0.5542	**140.16**	2310.54	4
	Otomatis	3e-5	1	0.8083	**0.8333**	0.4745	**0.4496**	160.22	2993.47	4

Tabel 6 menunjukkan hasil pelatihan model Bi-LSTM dan IndoBERT dengan variasi *learning rate* (3e-5 dan 5e-5) serta metode pelabelan (manual dan otomatis). Secara umum, IndoBERT unggul dalam akurasi dan efisiensi waktu, sedangkan Bi-LSTM lebih efisien dari sisi penggunaan memori.

Model Bi-LSTM mencapai performa terbaik pada *learning rate* 5e-5 dengan pelabelan otomatis, menghasilkan *validation accuracy* 0,6963 dan *train accuracy* 0,7936, dengan *train loss* 0,5255 dan *validation loss* 0,6692. Sementara itu, pada *learning rate* 3e-5, akurasi Bi-LSTM menurun saat menggunakan pelabelan manual tetapi meningkat kembali dengan pelabelan otomatis, menunjukkan sensitivitas model terhadap kombinasi *hyperparameter* dan kualitas pelabelan. Meskipun waktu pelatihan relatif panjang, penggunaan memorinya kecil berkisar 21 hingga 25 MB, sehingga model ini lebih efisien untuk sistem dengan keterbatasan sumber daya.

Di sisi lain, IndoBERT menunjukkan performa yang lebih stabil dan unggul pada seluruh konfigurasi, konsisten menunjukkan akurasi tinggi baik pada pelabelan manual maupun otomatis. Pelabelan otomatis terbukti meningkatkan *validation accuracy*, dengan konfigurasi terbaik tercapai pada *learning rate* 3e-5 dan pelabelan otomatis, menghasilkan *validation accuracy* 0,8333, *train accuracy* 0,8083, serta *train* dan *validation loss* rendah

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9458

yaitu 0,4745 dan 0,4496. Waktu pelatihan singkat berkisar 140 hingga 168 detik menegaskan efektivitas model dalam mempelajari representasi bahasa, meskipun penggunaan memorinya lebih besar (2310-2993 MB).

Dengan demikian, IndoBERT unggul dalam hal efektivitas performa, sementara Bi-LSTM lebih efisien dari sisi memori tetapi sensitif terhadap kombinasi learning rate dan metode pelabelan. Pelabelan otomatis terbukti meningkatkan kinerja kedua model, namun efeknya lebih menonjol pada Bi-LSTM karena kapasitas representasinya terbatas. IndoBERT mampu memanfaatkan data pelabelan otomatis untuk belajar representasi bahasa yang lebih baik, sehingga menghasilkan *validation accuracy* tertinggi dengan waktu pelatihan singkat, meskipun membutuhkan memori besar.

Secara keseluruhan, IndoBERT dengan pelabelan otomatis merupakan pilihan optimal untuk performa maksimal, sedangkan Bi-LSTM lebih cocok untuk sistem dengan keterbatasan sumber daya. Hasil ini menegaskan pentingnya pemilihan model, pengaturan *hyperparameter*, dan kualitas pelabelan dalam optimasi model *deep learning* berbasis NLP, sekaligus memberikan gambaran bagaimana kedua arsitektur merespons data teks yang besar dan kompleks pada analisis sentimen.

3.7 Evaluasi

Tahap evaluasi kinerja model dilakukan untuk menilai kinerja model dalam mengklasifikasikan sentimen publik terhadap Program MBG serta menentukan model yang paling akurat dan efisien dalam memproses teks informal berbahasa Indonesia. Evaluasi dilakukan menggunakan 4 matrik, yaitu *accuracy, Precision, Recall,* dan *F1-Score*. Model Bi-LSTM dan IndoBERT dievaluasi menggunakan dua konfigurasi *learning rate* untuk mengetahui pengaruh parameter pembelajaran terhadap performa model. Seluruh pengujian dilakukan pada dua skema pelabelan data, yaitu pelabelan manual dan pelabelan otomatis.

Tabel 8. Hasil evaluasi

Model	Pelabelan	Learning Rate	Accuracy	Precision	Recall	F1-Score
Bi-LSTM	Manual	5e-5	0.64	0.63	0.63	0.62
	Otomatis	5e-5	0.70	0.67	0.69	0.68
Bi-LSTM	Manual	3e-5	0.59	0.57	0.59	0.58
	Otomatis	3e-5	0.69	0.66	0.68	0.67
IndoBERT	Manual	5e-5	0.76	0.74	0.70	0.71
	Otomatis	5e-5	0.80	0.79	0.75	0.77
IndoBERT	Manual	3e-5	0.76	0.73	0.71	0.72
	Otomatis	3e-5	**0.82**	**0.80**	**0.78**	**0.79**

Berdasarkan Tabel 8, IndoBERT menunjukkan performa terbaik dalam klasifikasi sentimen publik terhadap Program MBG. Konfigurasi optimal diperoleh pada *learning rate* 3e-5 dengan pelabelan otomatis, menghasilkan *accuracy* 82%, *F1-Score* 79%, *precision* 80%, dan *recall* 78%, yang menunjukkan kemampuan model dalam mengklasifikasikan sentimen positif, negatif, dan netral secara proporsional. Sebaliknya, Bi-LSTM mencapai performa terbaik pada *learning rate* 5e-5 dengan pelabelan otomatis, yang menghasilkan *accuracy* 70% dan *F1-Score* 68%, menunjukkan bahwa meskipun model ini efektif dalam mengenali pola sentimen eksplisit, kemampuannya masih terbatas dalam memahami konteks semantik yang kompleks. Hasil ini menegaskan bahwa IndoBERT lebih unggul secara numerik dalam memproses teks informal berbahasa Indonesia, sementara Bi-LSTM cukup efektif pada pola sentimen sederhana.

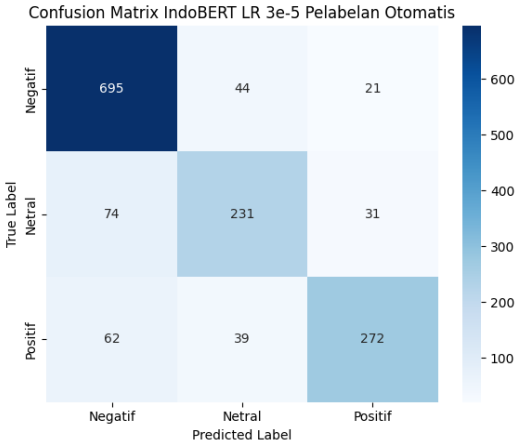
Perbedaan performa kedua model dapat dijelaskan oleh arsitekturnya. Bi-LSTM menggunakan pendekatan RNN sekuensial, sehingga pemahaman konteks terbatas pada urutan kata, kurang efektif untuk teks panjang atau informal. IndoBERT, berbasis transformer dengan mekanisme *attention bidirectional*, yang mampu menangkap hubungan antar kata dari berbagai arah, sehingga lebih akurat dalam memproses bahasa informal, slang, atau frasa ambigu yang umum pada opini publik terkait Program MBG.

Learning rate 3e-5 terbukti optimal untuk IndoBERT karena model belajar lebih stabil, mengurangi risiko *overfitting*, dan mampu menangkap pola sentimen dengan presisi tinggi. Selain itu, pelabelan otomatis secara konsisten meningkatkan performa karena label lebih konsisten dan minim subjektivitas, meskipun pelabelan manual tetap unggul dalam menangkap nuansa semantik yang kompleks.

Dari sisi interpretasi sentimen publik, performa tinggi IndoBERT menunjukkan bahwa opini masyarakat terhadap Program MBG cukup jelas dalam menyampaikan sentimen positif, negatif, maupun netral. Model mampu menangkap nuansa bahasa informal, variasi ekspresi, serta kontras opini yang muncul di media sosial. Hal ini menandakan bahwa mayoritas teks mengandung informasi sentimen yang relatif eksplisit, dan IndoBERT mampu mengenali konteks serta pola semantik yang kompleks untuk menyajikan klasifikasi yang proporsional.

Secara keseluruhan, IndoBERT dengan *learning rate* 3e-5 dan pelabelan otomatis merupakan kombinasi optimal untuk analisis sentimen Program MBG pada teks informal berbahasa Indonesia. Model ini unggul secara numerik dan relevan secara praktis untuk menangkap opini publik yang kompleks, sehingga dapat dijadikan acuan bagi penelitian analisis sentimen pada program pemerintah atau topik sosial lainnya.

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9458



Gambar 3. *Confusion Matrix*

Berdasarkan gambar 3 yang merupakan *confusion matrix* model IndoBERT dengan *learning rate* 3e-5 pelabelan otomatis, terlihat bahwa model mampu mengklasifikasikan sentimen dengan cukup baik pada seluruh kelas. Pada kelas negatif, sebanyak 695 data berhasil diprediksi dengan benar, sementara sebagian kecil masih salah diklasifikasikan sebagai netral (44 data) dan positif (21 data).

Pada kelas netral, model mengklasifikasikan 231 data dengan benar, namun masih terjadi kesalahan prediksi ke kelas negatif (74 data) dan positif (31 data). Hal ini menunjukkan bahwa kelas netral memiliki karakteristik yang lebih ambigu dan cenderung tumpeng tindik dengan sentimen lainnya. Sementara itu, sebanyak 272 data berhasil diklasifikasikan dengan benar, meskipun masih terdapat kesalahan ke kelas negatif (62 data) dan netral (39 data).

Secara keseluruhan, hasil *confusion matrix* menunjukkan bahwa model IndoBERT dengan *learning rate* 3e-5 memiliki kemampuan klasifikasi yang stabil dan proporsional pada ketiga kelas sentimen. Kesalahan yang muncul umumnya disebabkan oleh kemiripan makna antar kelas, terutama antara sentimen netral dengan sentimen negatif atau positif.

3.8 Perbandingan dengan Penelitian Sebelumnya

Hasil penelitian menunjukkan bahwa IndoBERT dengan *learning rate* 3e-5 dan pelabelan otomatis mencapai performa terbaik dalam klasifikasi sentimen publik terhadap Program MBG, dengan *accuracy* 82% dan *F1-score* 79%, mampu menangkap sentimen positif, negatif, maupun netral secara proporsional. Sebaliknya, Bi-LSTM mencapai performa maksimal pada *learning rate* 5e-5 dengan pelabelan otomatis, menghasilkan *accuracy* 70% dan *F1-score* 68%. Meskipun Bi-LSTM lebih efisien dari sisi penggunaan memori, model ini sensitif terhadap kombinasi hyperparameter dan metode pelabelan, serta kurang efektif dalam memahami konteks bahasa informal dan kompleks dibanding IndoBERT.

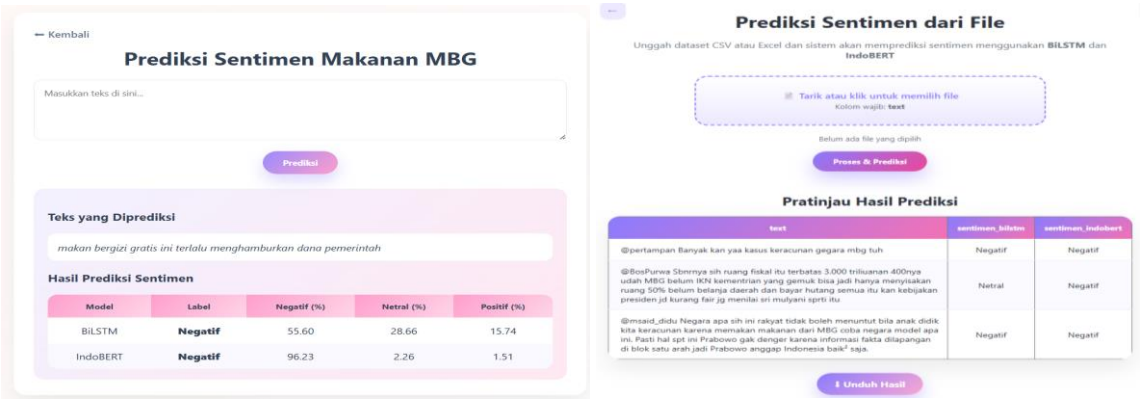
Sebagai perbandingan dengan penelitian terdahulu, model SVM dengan TF-IDF menunjukkan nilai *precision*, *recall*, dan *F1-score* tinggi (masing-masing 98%, 99%, dan 98%) namun *accuracy* keseluruhan hanya 67% [8], akibat ketidakseimbangan data dan keterbatasan jumlah dataset. Penelitian lain membandingkan KNN dan Naive Bayes untuk analisis sentimen Program MBG, di mana Naive Bayes unggul dengan *accuracy* 79,61% dan *F1-score* 77,33%, lebih tinggi dibanding KNN (*accuracy* 73,30% dan *F1-score* 67,37%) [9].

Dalam penelitian ini, ketidakseimbangan kelas dengan jumlah data negatif lebih banyak berhasil diatasi melalui teknik balancing, sehingga klasifikasi tetap proporsional. Dibandingkan kedua pendekatan tersebut, IndoBERT menunjukkan peningkatan signifikan, baik dari sisi akurasi maupun kemampuannya dalam menangkap konteks bahasa informal, slang, dan frasa ambigu, sekaligus mempertahankan distribusi klasifikasi yang seimbang pada ketiga kelas sentimen. Hasil ini memperkuat temuan bahwa model *deep learning* berbasis *transformer* tidak hanya unggul secara numerik, tetapi juga lebih efektif secara praktis dalam menangkap opini publik yang kompleks, sehingga relevan sebagai acuan dalam penelitian analisis sentimen untuk kebijakan sosial maupun program pemerintah.

3.9 Pengembangan Website Prediksi

Aplikasi web dikembangkan menggunakan *framework Flask* yang berfungsi sebagai media untuk melakukan prediksi sentimen teks mengenai topik Program Makan Bergizi Gratis. Pengguna dapat memasukkan teks secara langsung pakai kolom *input text* atau mengunggah file CSV maupun Excel untuk dianalisis. Sistem kemudian mengolah data tersebut menggunakan model Bi-LSTM dan IndoBERT yang telah dilatih sebelumnya. Output prediksi ini berupa tabel yang berisi teks dan prediksi label positif, netral atau negatif hasil klasifikasi sentimen seperti pada gambar 4.

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9458



Gambar 4. Tampilan Website Prediksi

4. KESIMPULAN

Penelitian ini menganalisis sentimen publik terhadap Program Makan Bergizi Gratis (MBG) menggunakan model Bi-LSTM dan IndoBERT berdasarkan data Twitter dan TikTok, dengan hasil yang menunjukkan bahwa IndoBERT secara konsisten memiliki performa lebih unggul dibandingkan Bi-LSTM. Konfigurasi terbaik diperoleh pada IndoBERT dengan learning rate 3e-5 dan pelabelan otomatis, yang mencapai *accuracy* 82% dan *F1-score* 79%. Penerapan pelabelan otomatis dan teknik *Random Oversampling* terbukti efektif dalam meningkatkan konsistensi label serta mengurangi dampak ketidakseimbangan data, sementara Bi-LSTM lebih unggul dari sisi efisiensi memori namun sensitif terhadap variasi hyperparameter. Penelitian ini memberikan kontribusi empiris dalam perbandingan arsitektur *deep learning* untuk analisis sentimen kebijakan publik serta menghasilkan aplikasi prediksi sentimen berbasis web, dengan saran pengembangan selanjutnya berupa perluasan sumber data dan eksplorasi model atau teknik penanganan ketidakseimbangan data yang lebih lanjut. Sebagai saran, penelitian selanjutnya dapat memperluas sumber data dengan melibatkan lebih banyak platform media sosial agar gambaran opini publik yang diperoleh menjadi lebih beragam dan representatif. Selain itu, teknik penanganan ketidakseimbangan data seperti SMOTE atau pendekatan *cost-sensitive learning* dapat digunakan sebagai pembanding untuk meningkatkan kinerja klasifikasi. Penelitian berikutnya juga dapat mempertimbangkan penggunaan model *transformer* yang lebih mutakhir atau pendekatan *aspect-based sentiment analysis* agar analisis sentimen tidak hanya bersifat umum, tetapi mampu menggambarkan aspek-aspek tertentu yang menjadi fokus perhatian masyarakat secara lebih mendalam.

REFERENCES

[1] R. Imeldawati, “Dampak Terjadinya Stunting terhadap Perkembangan Kognitif Anak : Literature Review,” *J. Med. Nusant.*, vol. 3, no. 1, pp. 101–107, 2025, doi: 10.59680/medika.v3i1.1632.

[2] A. Kiftiyah, F. A. Palestina, F. U. Abshar, and K. Rofiah, “Program Makan Bergizi Gratis (MBG) dalam Perspektif Keadilan Sosial dan Dinamika Sosial – Politik,” *Pancasila J. Keindonesiaan*, vol. 5, no. 1, pp. 101–112, 2025, doi: 10.52738/pjk.v5i1.726.

[3] R. Amelia Br Siregar, A. Ramadhan Manik, A. Syahri, M. Budi Akbar, A. Arnita, and F. Ramadhani, “Penerapan Naïve Bayes Untuk Analisis Sentimen Netizen Terhadap Program Makan Siang Gratis Pada Media Social X,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 4, pp. 5621–5628, 2025, doi: 10.36040/jati.v9i4.13870.

[4] A. Z. Ahmad, E. Asril, M. Sadar, Walhidayat, Syahtriatna, and Y. Turnandes, “Analisis Sentimen Opini Terhadap Vaksin Covid-19 Pada Media Sosial Twitter Menggunakan Naïve Bayes Dan Decision Tree,” *Zo. J. Sist. Inf.*, vol. 5, no. 1, pp. 100–110, 2023, doi: 10.31849/zn.v5i1.5553.

[5] F. Fathoni, A. P. Maretta, A. N. Kusuma, R. M. Sasmita, A. F. Rizkyllah, and A. Ibrahim, “Perbandingan Metode Naïve Bayes Dan K-Nearest Neighbor Terhadap Analisis Sentimen Ulasan Program Makan Siang Gratis Di Indonesia,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 4, pp. 6385–6390, 2025, doi: 10.36040/jati.v9i4.14084.

[6] B. Rahmatullah, S. A. Saputra, P. Budiono, and D. P. Wigandi, “Sentimen Analisis Makan Bergizi Gratis Menggunakan Algoritma Naive Bayes,” *J. Inf. Technol.*, vol. 5, no. 1, 2025, doi: 10.46229/jifotech.v5i1.978.

[7] S. A. S. Mola, S. N. R. Djawa, and A. Y. Mauko, *Text Mining: Analisis Sentimen dengan Naïve Bayes*. Bandung: Kaizen Media Publishing, 2025. Accessed: Jan. 17, 2026. [Online]. Available: https://books.google.co.id/books?id=qrxNEQAAQBAJ&printsec=frontcover&hl=id&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

[8] M. S. Dewi and A. H. Hasugian, “PENERAPAN SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN PADA TANGGAPAN MASYARAKAT DI MEDIA SOSIAL TERHADAP PROGRAM MAKAN SIANG GRATIS,” vol. 10, no. 2, pp. 911–922, 2025, doi: /10.36341/rabit.v10i2.6425.

[9] N. Sakina, F. Wajidi, and M. R. Rasyid, “Evaluasi Algoritma KNN dan Naive Bayes untuk Analisis Sentimen Kebijakan Program Makan Bergizi Gratis,” *Jambura J. Informatics*, vol. 1, no. 2, pp. 83–97, 2025, doi: 10.37905/jji.v1i2.34418.

[10] S. Nurmaini, A. Darmawahyuni, A. I. Sapitri, M. N. Rachmatullah, Firdaus, and B. Tutuko, *PENGENALAN DEEP*

LEARNING DAN IMPLEMENTASINYA. Palembang: UPT Penerbit dan Percetakan Universitas Sriwijaya, 2021.

[11] A. Saputra, R. C. Sigitta Hariyono, and N. M. Saraswati, "Analisis Sentimen Pengguna Aplikasi MyPertamina Menggunakan Algoritma Bidirectional Long Short Term Memory," *J. Eksplora Inform.*, vol. 13, no. 2, pp. 156–163, 2024, doi: 10.30864/eksplora.v13i2.973.

[12] M. R. Manoppo, I. C. Kolang, D. N. Fiat, R. Michelly, and C. Mawara, "Analisis Sentimen Publik Di Media Sosial Terhadap Kenaikan Ppn 12 % Di Indonesia Menggunakan Indobert Analysis of Public Sentiment on Social Media Towards the 12 % Ppn Increase in Indonesia Using Indobert," vol. 4, no. 2, pp. 152–163, 2025, doi: 10.69916/jkbti.v4i2.322.

[13] H. Ma'we, A. Y. Husodo, and B. Irmawati, "Performance Comparison of Naive Bayes and Bidirectional Lstm Algorithms in Bsi Mobile Review Sentiment Analysis," *J. Tek. Inform.*, vol. 6, no. 1, pp. 159–172, 2024, doi: 10.52436/1.jutif.2024.5.6.4178.

[14] N. N. A. Aryanti and O. Suria, "ANALISIS SENTIMEN TERHADAP PEMUTUSAN HUBUNGAN KERJA DI INDONESIA : KOMPARASI INDOBERT DENGAN SVM, RANDOM FOREST, DAN DECISION TREE DENGAN OPTIMASI TF - IDF 1," *Rabit J. Teknol. dan Sist. Inf. Univrab*, vol. 10, no. 2, pp. 1158–1176, 2025, doi: 10.36341/rabit.v10i2.6364.

[15] V. Chang, L. Liu, Q. Xu, T. Li, and C. H. Hsu, "An improved model for sentiment analysis on luxury hotel review," *Expert Syst.*, vol. 40, no. 2, 2023, doi: 10.1111/exsy.12580.

[16] L. Jannah, "Analisis Sentimen Publik Terhadap Kenaikan Pajak Pertambahan Nilai (PPN) Sebesar 12 % Menggunakan Naïve Bayes Classifier," vol. 7, pp. 601–612, 2025, doi: 10.30865/json.v7i2.9151.

[17] M. D. H. Fahri and D. Gunawan, "Analisis Sentimen Pengguna X terhadap Perempuan di Lingkungan Kerja Menggunakan Algoritma Machine Learning," *J. Technol. Informatics*, vol. 7, no. 2, pp. 134–146, 2025, doi: 10.37802/joti.v7i2.1087.

[18] R. Hidayat and W. Gata, "Pemanfaatan IndoBERT Untuk Analisis Sentimen Ulasan Aplikasi Depok Single Window," *Inf. Syst. Educ. Prof. J. Inf. Syst.*, vol. 10, no. 1, p. 13, 2025, doi: 10.51211/isbi.v10i1.3334.

[19] R. R. Suryono, "Sentiment Classification of Indonesian-Language Roblox Reviews Using IndoBERT with SMOTE Optimization," *J. Appl. Informatics Comput.*, vol. 9, no. 4, pp. 1868–1877, 2025, doi: 10.30871/jaic.v9i4.10155.

[20] H. A. Nabila and E. W. Pamungkas, "PERBANDINGAN ALGORITMA MACHINE LEARNING: SVM, RANDOM FOREST, DAN XGBOOST UNTUK PREDIKSI STROKE 1)," *Rabit J. Teknol. Dan Sist. Inf. Univrab*, vol. 10, no. 2, pp. 1098–1110, 2025, doi: 10.36341/rabit.v10i2.6444.

[21] D. R. Alfinsyah and B. P. Hartato, "Evaluating the Impact of Random Over Sampling on IndoBERT Performance for Indonesian Sentiment Analysis," vol. 9, no. 6, pp. 3270–3282, 2025.

[22] X. Liu, Z. Ma, H. Guo, Y. Xu, and Y. Cao, "Short-Term Power Load Forecasting Based on DE-IHHO Optimized BiLSTM," *IEEE Access*, vol. 12, pp. 145341–145349, 2024, doi: 10.1109/ACCESS.2024.3437247.

[23] U. B. Mahadevaswamy and P. Swathi, "Sentiment Analysis using Bidirectional LSTM Network," *Procedia Comput. Sci.*, vol. 218, pp. 45–56, 2022, doi: 10.1016/j.procs.2022.12.400.

[24] G. A. Pradnyana, W. Anggraeni, E. M. Yuniarno, and M. H. Purnomo, "Fine-Tuning IndoBERT Model for Big Five Personality Prediction from Indonesian Social Media," in *2023 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, IEEE, 2023, pp. 93–98. doi: 10.1109/ISITIA59021.2023.10221074.

[25] H. Imaduddin, F. Y. A'la, and Y. S. Nugroho, "Sentiment Analysis in Indonesian Healthcare Applications using IndoBERT Approach," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 8, pp. 113–117, 2023, doi: 10.14569/IJACSA.2023.0140813.

[26] M. I. Abidin and E. W. Pamungkas, "Analisis Sentimen Terhadap Timnas Indonesia Di Piala Asia 2023 Dengan Model Transformer Berbahasa Indonesia," *Rabit J. Teknol. dan Sist. Inf. Univrab*, vol. 10, no. 2, pp. 482–496, 2025, doi: 10.36341/rabit.v10i2.6142.