JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9460
Hal 265-276
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Analisis Sentimen Multi-Platform Media Sosial pada Program Makan Bergizi Gratis Menggunakan Ensemble IndoBERT-SVM

Muhammad Bisri Mustofa*, Ifnu Wisma Dwi Prastya, Sahri

Fakultas Sains dan Teknologi, Teknik Informatika, Universitas Nahdlatul Ulama Sunan Giri, Bojonegoro, Indonesia
Email: [1*]bisri3522@gmail.com, [2]ifnuprastya@unugiri.ac.id, [3]sahriunugiri@gmail.com
Email Penulis Korespondensi: bisri3522@gmail.com
Submitted **02-01-2026;** Accepted **15-02-2026;** Published **28-02-2026**

Abstrak

Analisis sentimen kebijakan publik di media sosial menghadapi tantangan signifikan akibat heterogenitas linguistik lintas platform dan keterbatasan model tunggal dalam menangkap keragaman ekspresi opini. Penelitian sebelumnya cenderung menggunakan pendekatan single-platform dan single-model yang berpotensi menghasilkan bias representasi dan penurunan akurasi hingga 12–15% ketika diterapkan pada konteks platform berbeda. Penelitian ini bertujuan mengembangkan model ensemble berbasis soft voting yang mengintegrasikan Support Vector Machine (SVM) dan IndoBERT untuk menganalisis sentimen publik terhadap Program Makan Bergizi Gratis (MBG) secara multi-platform, serta mengevaluasi efektivitas pendekatan ensemble dibandingkan model tunggal dalam mengatasi variasi karakteristik linguistik platform digital. Dataset penelitian terdiri dari 7.500 komentar dari X, TikTok, dan YouTube yang dikumpulkan periode 6 Januari–28 September 2025, diproses melalui preprocessing bahasa Indonesia informal, pelabelan berbasis leksikon, dan pembagian stratified split. Hasil menunjukkan bahwa SVM optimal pada TikTok (akurasi 98,1%, F1 makro 98,0%) namun lemah pada kelas netral di X (F1 51,0%), sementara IndoBERT unggul menangani ambiguitas pragmatik di X (F1 netral 74,0%) meski sedikit menurun di TikTok (F1 makro 93,0%). Model ensemble menghasilkan performa paling seimbang dengan akurasi 92,53% (YouTube), 95,73% (TikTok), 92,53% (X), dan F1 makro 85,07%, 94,33%, 84,92%. Kontribusi penelitian ini meliputi pengembangan pendekatan multi-platform sentiment analysis yang mengatasi bias single-platform, peningkatan kemampuan generalisasi klasifikasi lintas ekosistem digital heterogen, serta penyediaan instrumen evidence-based evaluation untuk perbaikan implementasi dan komunikasi kebijakan pemerintah.

Kata Kunci: Analisis Sentimen; IndoBERT; Multi-Platform; Soft Voting Ensemble; Support Vector Machine

Abstract

Sentiment analysis of public policy on social media faces significant challenges due to linguistic heterogeneity across platforms and limitations of single models in capturing the diversity of opinion expressions. Previous studies tend to employ single-platform and single-model approaches that potentially generate representational bias and accuracy degradation of up to 12–15% when applied to different platform contexts. This study aims to develop a soft voting-based ensemble model that integrates Support Vector Machine (SVM) and IndoBERT to analyze public sentiment toward the Free Nutritious Meal (MBG) Program across multiple platforms, and to evaluate the effectiveness of the ensemble approach compared to single models in addressing variations in linguistic characteristics of digital platforms. The research dataset consists of 7,500 comments from X, TikTok, and YouTube collected from January 6 to September 28, 2025, processed through informal Indonesian language preprocessing, lexicon-based labeling, and stratified split division. Results demonstrate that SVM performs optimally on TikTok (accuracy 98.1%, macro F1 98.0%) but weakly on the neutral class in X (F1 51.0%), while IndoBERT excels in handling pragmatic ambiguity in X (neutral F1 74.0%) despite slightly declining on TikTok (macro F1 93.0%). The ensemble model produces the most balanced performance with accuracies of 92.53% (YouTube), 95.73% (TikTok), 92.53% (X), and macro F1 scores of 85.07%, 94.33%, 84.92% respectively. The contributions of this research include the development of a multi-platform sentiment analysis approach that addresses single-platform bias, improved classification generalization capability across heterogeneous digital ecosystems, and provision of evidence-based evaluation instruments for improving government policy implementation and communication.

Keywords: Sentiment Analysis; IndoBERT; Multi-Platform; Soft Voting Ensemble; Support Vector Machine

1. PENDAHULUAN

Program Makan Bergizi Gratis (MBG) mulai diimplementasikan secara bertahap pada Januari 2025 dengan dukungan anggaran Rp71 triliun dan target jutaan penerima manfaat. Selain meningkatkan pemenuhan gizi, MBG terbukti berkontribusi pada peningkatan minat belajar siswa sekolah dasar dan menjadi fokus perhatian publik terkait tata kelola, mekanisme distribusi, serta keberlanjutan pelaksanaannya [1]. Permasalahan utama dalam evaluasi kebijakan publik seperti MBG adalah sulitnya memperoleh gambaran opini masyarakat secara objektif dan representatif, mengingat opini publik kini tersebar di berbagai platform media sosial dengan karakteristik komunikasi yang heterogen mulai dari diskusi reflektif di YouTube, ekspresi spontan di TikTok, hingga debat argumentatif di X. Perbedaan ekologi komunikasi ini berpotensi menghasilkan bias interpretasi jika analisis hanya mengandalkan satu platform atau satu model klasifikasi. Media sosial kini berfungsi tidak hanya sebagai sarana komunikasi, tetapi juga sebagai ruang diskusi publik yang mencerminkan persepsi masyarakat terhadap kebijakan pemerintah, sehingga analisis sentimen berbasis data media sosial menjadi penting untuk memahami opini publik secara objektif dan komprehensif [2].

Model berbasis *Bidirectional Encoder Representations from Transformers* (BERT), khususnya *IndoBERT*, terbukti unggul dalam analisis sentimen berbahasa Indonesia [3]. *IndoBERT* telah berhasil diterapkan pada berbagai domain, dari ulasan *e-commerce* seperti Tiket.com (akurasi 91%) [4] dan Shopee (93%) [5], hingga isu sosial-politik seperti film *Dirty Vote* [6] dan investasi publik Danantara [7]. Studi komparatif menunjukkan BERT mengungguli model

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9460
Hal 265-276
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

unidirectional seperti GPT dalam klasifikasi sentimen berbahasa Indonesia, dengan akurasi 79,36% [6] dan 81% dalam deteksi perundungan siber [8], menjadikan *IndoBERT* kandidat ideal untuk analisis opini kebijakan strategis nasional.

Beberapa penelitian terdahulu telah mengeksplorasi topik serupa namun dengan keterbatasan. Triningsih dkk. [2] menganalisis sentimen Program MBG menggunakan *machine learning* di X, namun terbatas pada satu platform. Agustina dan Ihsan [9] mengembangkan *ensemble* untuk sentimen COVID-19 menggunakan *soft voting*, tetapi fokus pada satu platform tanpa mempertimbangkan variasi komunikasi antarplatform. Cahyadi dan Rochadiani [10] mengimplementasikan *ensemble deep learning* untuk sentimen *game mobile* (akurasi 90%), namun tidak mengeksplorasi konteks *multi-platform*. Sinapoy dkk. [11] membandingkan LSTM dan *IndoBERT* untuk deteksi hoaks di Twitter, menemukan keunggulan *IndoBERT*, namun juga terbatas satu platform.

Telaah terhadap penelitian *IndoBERT* periode 2022–2024 menunjukkan sebagian besar studi memusatkan analisis pada satu platform, terutama X. Kesenjangan penelitian (*research gap*) yang diidentifikasi mencakup: (1) Bias *single-platform* yang mengabaikan heterogenitas ekosistem digital. Perbedaan konteks platform dapat memengaruhi persepsi publik dengan variasi sentimen hingga 23%, di mana X cenderung lebih kritis dibanding YouTube [9]. Penelitian ini mengintegrasikan tiga platform utama: X, YouTube, dan TikTok. (2) Keterbatasan generalisasi *single model* lintas platform yang berpotensi menurunkan akurasi 12–15% akibat perbedaan karakteristik linguistik [10]. Model berbasis fitur statistik seperti *Support Vector Machine* (SVM) efektif pada teks singkat eksplisit namun lemah pada ambiguitas pragmatik, sementara *IndoBERT* unggul dalam nuansa semantik tetapi dapat menurun pada teks sangat singkat. (3) Belum adanya integrasi *IndoBERT*–SVM dalam kerangka *ensemble* untuk analisis sentimen kebijakan publik *multi-platform*, meskipun pendekatan *ensemble* terbukti meningkatkan performa [10].

Untuk mengatasi keterbatasan tersebut, penelitian ini menggunakan *ensemble learning* yang menggabungkan *IndoBERT* dengan SVM melalui *soft voting*. SVM dipilih karena kemampuannya menangani *high-dimensional feature space* dari *embedding* BERT sekaligus memberikan interpretabilitas relevan untuk evaluasi kebijakan publik. Pendekatan *ensemble deep learning* mampu meningkatkan akurasi hingga 90% dengan standar deviasi kesalahan lebih rendah [10]. Kombinasi *IndoBERT*–SVM terbukti meningkatkan akurasi 7–9% serta menurunkan bias klasifikasi, mendukung *evidence-based policy* [12], [11].

Penelitian ini menganalisis 7.500 komentar publik (2.500 per platform: X, TikTok, YouTube) periode 6 Januari–28 September 2025, mencakup fase implementasi awal hingga evaluasi keberlanjutan MBG. Kontribusi teoretis: (1) Pengembangan pendekatan *multi-platform sentiment analysis* berbasis *ensemble* untuk mengatasi bias *single-platform* dan meningkatkan generalisasi lintas ekosistem digital heterogen; (2) Validasi empiris efektivitas *soft voting ensemble* dibanding model tunggal dalam analisis sentimen kebijakan publik berbahasa Indonesia; (3) Identifikasi pola hubungan antara karakteristik komunikasi platform dengan performa klasifikasi sentimen. Kontribusi praktis: (1) Peta komprehensif dinamika sentimen publik terhadap Program MBG di tiga ekosistem digital untuk *evidence-based evaluation* kebijakan; (2) Bukti empiris bahwa *ensemble IndoBERT*–SVM mengurangi bias klasifikasi *single model*; (3) Dasar empiris bahwa efektivitas analisis sentimen kebijakan publik ditentukan kemampuan adaptasi model terhadap karakteristik linguistik platform, mendukung strategi komunikasi tersegmentasi.

Berdasarkan kesenjangan penelitian tersebut, studi ini bertujuan mengembangkan dan mengevaluasi model *ensemble IndoBERT*–SVM dalam analisis sentimen *multi-platform* terhadap Program MBG sebagai instrumen pemetaan opini publik untuk evaluasi kebijakan. Secara spesifik: (1) Menganalisis distribusi dan karakteristik sentimen publik terhadap Program MBG di tiga platform (X, YouTube, TikTok) untuk memahami perbedaan pola ekspresi opini yang dipengaruhi ekologi komunikasi platform; (2) Mengevaluasi efektivitas *ensemble IndoBERT*–SVM dibanding model tunggal dalam klasifikasi sentimen *multi-platform*, dengan fokus pada akurasi, *F1-score* per kelas, dan stabilitas performa lintas karakteristik linguistik platform digital.

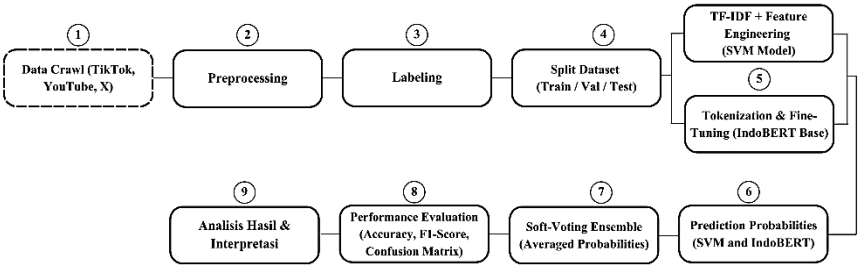
2. METODOLOGI PENELITIAN

Penelitian ini terdiri atas sembilan tahapan utama yang mencakup proses pengumpulan data dari tiga platform media sosial hingga evaluasi model *ensemble* IndoBERT SVM. Metode ini menggabungkan pendekatan *machine learning* dan *deep learning* untuk memperoleh klasifikasi sentimen publik yang akurat terhadap Program Makan Bergizi Gratis (MBG).

2.1 Alur Penelitian

Alur penelitian mencakup proses *crawling* data dari tiga platform media sosial (YouTube, TikTok, dan X), tahap preprocessing, pelabelan data, pembagian dataset menjadi data latih, validasi, dan uji, pembentukan model berbasis SVM dengan fitur TF-IDF serta *fine-tuning* model *IndoBERT*, penggabungan hasil prediksi probabilitas menggunakan metode soft voting ensemble, serta evaluasi performa model melalui metrik akurasi, F1-*score*, dan *confusion matrix*. Diagram alur penelitian tersebut disajikan pada Gambar 1.

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9460
Hal 265-276
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom



Gambar 1. *Flowchart* Diagram Alur Penelitian

2.2 Crawling Data

Tahap *crawling* dilakukan pada tiga *platform* media sosial YouTube, TikTok, dan X yang dipilih karena masing-masing mencerminkan pola komunikasi publik yang berbeda opini reflektif di YouTube, ekspresi spontan pada TikTok, dan percakapan argumentatif *real-time* di X [13], [14]. Pendekatan *multi-platform* ini memastikan analisis menangkap keragaman persepsi publik terhadap Program Makan Bergizi Gratis (MBG).

Data dikumpulkan pada periode 2025-01-06 hingga 2025-09-28 dengan kuota 2.500 komentar per *platform* (total 7.500 komentar), ditetapkan untuk menjaga representativitas sentimen, meminimalkan duplikasi dan spam, serta mempertahankan keseimbangan antar-platform [15], [16]. Proses *crawling* dijalankan menggunakan *Python* 3.10 pada *Google Colab* (*Ubuntu* 22.04) dengan *pustaka pandas, requests, googleapiclient*, serta *integrasi Node.js*.

Data YouTube diperoleh melalui YouTube Data API v3 (*googleapiclient.discovery*), mengekstraksi atribut seperti *publishedAt, authorDisplayName, textDisplay, dan likeCount*, sesuai prinsip *reproducible research* dan etika akses konten publik [17]. Untuk TikTok, data dikumpulkan melalui web *scraping* pada konten publik yang dapat diakses tanpa autentikasi, menggunakan *requests* dan *cursor-based pagination*, disertai pembersihan HTML entities, emoji, dan karakter non-alfabet [18]. Data X dikumpulkan menggunakan *tweet-harvest* v2.6.1 (*Node.js*) dengan kueri: ("Makan Bergizi Gratis" OR "MBG" OR "Makan Siang Gratis") *since*:2025-01-06 *until*:2025-09-28 lang:id.

Seluruh data dikompilasi dan disimpan dalam format CSV berpengkodean UTF-8 yang mencakup atribut utama berupa username, teks komentar, platform asal, serta penanda waktu (*timestamp*). Pengelolaan data dilakukan dengan mengacu pada prinsip *ethical digital research*, dengan membatasi pengambilan data hanya pada konten yang bersifat publik, menerapkan anonimisasi identitas pengguna sebagai bentuk perlindungan privasi, serta melakukan penyaringan awal terhadap ujaran kebencian dan konten bermasalah. Pendekatan ini selaras dengan prinsip perlindungan hak privasi dan *right to be forgotten* dalam konteks penggunaan data media sosial di Indonesia [19].

2.3 Preprocessing Data

Tahap *preprocessing* bertujuan menstandarkan teks agar bersih dan seragam sebelum proses klasifikasi sentimen [20]. Mengingat komentar media sosial bersifat tidak terstruktur, tahap ini krusial untuk memastikan konsistensi pembacaan pola bahasa oleh model. Proses *preprocessing* mencakup enam langkah utama. *Cleansing* dilakukan untuk menghapus elemen tidak relevan seperti URL, mention, hashtag, angka, emoji, dan karakter non-ASCII [21]. Selanjutnya, *tokenization* berbasis spasi dan tanda baca diterapkan untuk memecah teks menjadi unit kata [22]. *Case folding* digunakan untuk menyatukan huruf ke bentuk kecil [23], diikuti *stopword removal* menggunakan pustaka *Sastrawi* guna menghilangkan kata umum yang tidak berkontribusi signifikan terhadap polaritas sentimen [24]. Normalisasi bahasa tidak baku (*slang normalization*) dilakukan dengan kamus *alay* berbasis dataset, misalnya “gk” menjadi “tidak” dan “bgt” menjadi “banget” [25]. Terakhir, *stemming* dengan algoritma *Nazief–Adriani* mengembalikan kata ke bentuk dasarnya untuk mengurangi ambiguitas morfologis [26].Seluruh tahapan diimplementasikan menggunakan Python 3.10 pada Google Colab (Ubuntu 22.04) dengan pustaka *re*, *pandas*, dan *Sastrawi*. Dataset hasil *preprocessing* disimpan sebagai *dataset_clean.xlsx* dengan kolom *text* dan *clean_text*. Sebagai ilustrasi, kalimat “Makan siang gratisnya keren banget!!!” diproses menjadi “makan siang gratis keren”.

2.4 Labeling Data

Tahap *labeling* dilakukan secara otomatis menggunakan metode *lexicon-based sentiment analysis* dengan memberikan label positif, negatif, atau netral pada setiap teks. Proses ini membandingkan kata dalam teks dengan kamus polaritas yang berisi daftar kata positif dan negatif untuk menentukan kecenderungan sentimen. Aturan linguistik tambahan diterapkan untuk menangani negasi seperti “tidak + kata positif” yang dikategorikan negatif dan “tidak + kata negatif” sebagai positif. Pendekatan ini dipilih karena bersifat sederhana, cepat, serta efisien untuk data besar tanpa memerlukan pelatihan model, meskipun memiliki keterbatasan dalam mendeteksi sarkasme, ironi, atau sentimen implisit [27].

2.5 Split Dataset

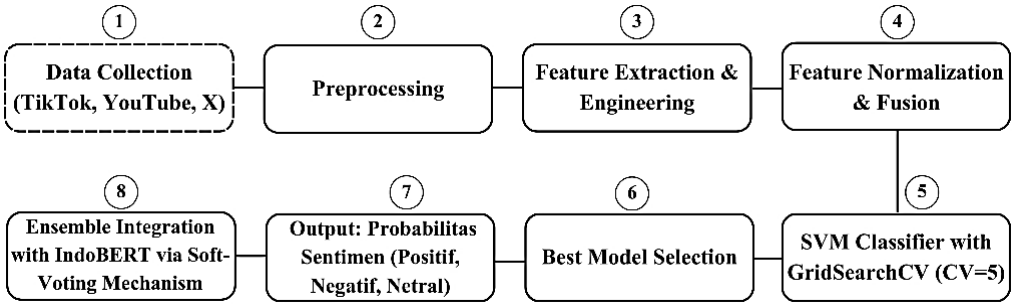
Tahap ini bertujuan membagi data hasil pelabelan menjadi tiga subset menggunakan metode *stratified split* agar distribusi kelas sentimen tetap seimbang pada setiap bagian dataset. Komposisi pembagian ditetapkan sebesar 70% untuk *training*, 15% untuk *validation*, dan 15% untuk *testing*. Metode ini dilakukan dengan pustaka *scikit-learn* di Python, sebagaimana diterapkan dalam penelitian serupa untuk menjaga objektivitas dan menghindari *overfitting* [28]. Setiap subset disimpan

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9460
Hal 265-276
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

secara terpisah dalam format *.xlsx* menggunakan *pandas*, dan dilakukan validasi akhir untuk memastikan kolom *clean_text* serta *label* bebas dari nilai kosong (*NaN*) sebelum tahap pelatihan mode.

2.6 SVM Model

Model SVM dibangun berbasis representasi TF-IDF dengan n-gram 1–2 (maks. 30.000 fitur), diperkaya fitur linguistik tambahan: skor sentimen (TextBlob, VADER), panjang teks, rasio kategori kata, jumlah negasi, pola tanda baca/ karakter berulang, proporsi huruf kapital, serta fitur khas media sosial (emoji, hashtag, mention, URL, dan rasio POS) [29], [30]. Seluruh fitur numerik dinormalisasi menggunakan StandardScaler dan digabungkan dengan matriks *TF-IDF* melalui teknik *sparse concatenation* untuk menjaga efisiensi komputasi. Arsitektur lengkap model SVM ditunjukkan pada Gambar 2.



Gambar 2. Arsitektur Model SVM

Gambar 2 memperlihatkan tiga komponen utama: (1) ekstraksi TF-IDF dari teks terpreproses, (2) integrasi fitur linguistik tambahan, dan (3) klasifikasi SVM kernel linear yang menghasilkan distribusi probabilitas sentimen. Secara matematis, fungsi keputusan SVM dinyatakan sebagai:

$$f(\mathbf{x}) = \text{sign}\left(\sum_{i=1}^{n} \alpha_i \, y_i K(\mathbf{x}_i, \mathbf{x}) + b\right) \tag{1}$$

dengan $\alpha_i$sebagai koefisien vektor pendukung, $y_i$label kelas, $K(\cdot,\cdot)$fungsi kernel, dan $b$bias. Untuk memperoleh probabilitas kelas, keluaran fungsi keputusan dikonversi menggunakan *Platt scaling*:

$$P(y = c \mid \mathbf{x}) = \frac{1}{1 + \exp\left(A f_c(\mathbf{x}) + B\right)} \tag{2}$$

Sebagai ilustrasi mekanisme kerja model, diberikan contoh perhitungan manual pada satu sampel data. Untuk komentar *"mbg enak banget"* setelah tahap *preprocessing*, vektor *TF-IDF* diperoleh, misalnya bobot kata "mbg" = 0.40, "enak" = 0.85, dan "banget" = 0.60, sehingga vektor fitur dinyatakan sebagai:

$$\mathbf{x} = [0.40,\ 0.85,\ 0.60,\ \ldots\ ] \tag{3}$$

Dengan menggunakan SVM kernel linear, skor keputusan untuk masing-masing kelas sentimen diperoleh sebagai:

$$f_{\text{neg}}(\mathbf{x}) = -1.8, f_{\text{net}}(\mathbf{x}) = -0.3, f_{\text{pos}}(\mathbf{x}) = 2.1 \tag{4}$$

Dengan parameter *Platt scaling* $A = 0.5$dan $B = -1.0$, probabilitas kelas positif dihitung sebagai:

$$P(\text{positif} \mid \mathbf{x}) = \frac{1}{1 + \exp\left(0.5 \cdot 2.1 - 1.0\right)} = 0.85 \tag{5}$$

Setelah normalisasi terhadap seluruh kelas, distribusi probabilitas akhir model SVM dinyatakan sebagai:

$$P_{\text{SVM}} = [0.05,\ 0.10,\ 0.85] \tag{6}$$

yang masing-masing merepresentasikan kelas negatif, netral, dan positif. Contoh perhitungan ini disajikan sebagai ilustrasi mekanisme kerja model dan tidak dimaksudkan merepresentasikan seluruh proses inferensi.

Pelatihan model dilakukan menggunakan kernel linear, *Radial Basis Function* (RBF), dan polynomial melalui *Grid Search* dengan lima lipatan validasi silang berdasarkan skor *F1-weighted* [31]. Kernel linear terpilih karena menunjukkan stabilitas dan kinerja terbaik pada data teks Bahasa Indonesia. Model diaktifkan dengan *probability = True* agar keluaran probabilitas dapat diintegrasikan dalam skema *soft voting ensemble* bersama *IndoBERT*.

2.7 IndoBERT Fine-tuning

Model *IndoBERT* (*indolem/indobert-base-uncased*) diadaptasi untuk klasifikasi sentimen Bahasa Indonesia berbasis konteks dengan memanfaatkan arsitektur *Transformer encoder* [32]. Teks hasil *preprocessing* ditokenisasi menggunakan *AutoTokenizer* dengan panjang maksimum 128 token. Representasi token khusus $[CLS]$digunakan sebagai ringkasan

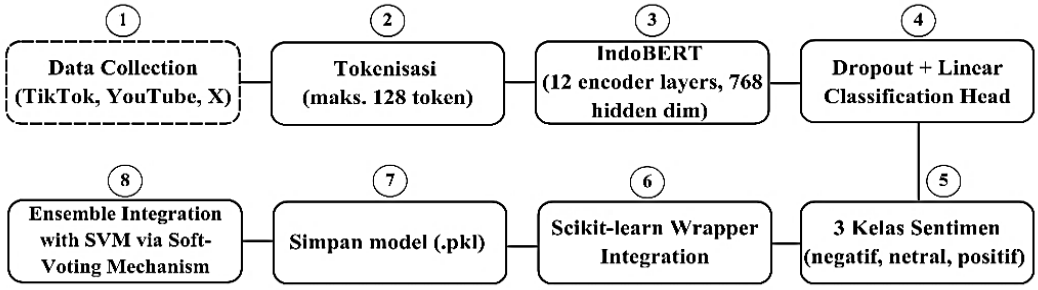
JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9460
Hal 265-276
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

semantik seluruh urutan teks dan dipetakan ke tiga kelas sentimen (positif, netral, negatif) melalui *classification head* berupa lapisan linear dengan *dropout* sebesar 0,3.

Secara matematis, probabilitas kelas sentimen diperoleh menggunakan fungsi *softmax* sebagai berikut:

$$P(y = c \mid \mathbf{x}) = \text{softmax}(W\mathbf{h}_{[\text{CLS}]} + b)_c \quad (7)$$

dengan W dan b masing-masing merepresentasikan bobot dan bias lapisan klasifikasi, serta c menunjukkan kelas sentimen. Skema arsitektur dan alur pemrosesan lengkap model *IndoBERT* disajikan pada Gambar 3.



Gambar 3. Arsitektur Model *IndoBERT*

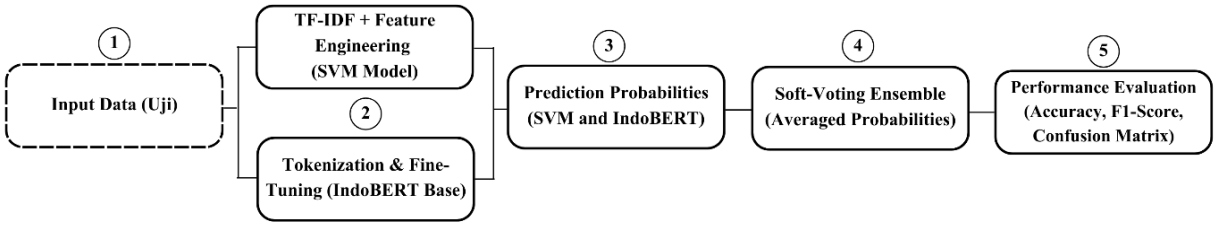
Gambar 3 memperlihatkan arsitektur *IndoBERT* yang terdiri dari lapisan tokenisasi, *Transformer encoder* untuk ekstraksi fitur kontekstual, dan *classification head* untuk prediksi sentimen. Pelatihan model dilakukan selama 4 epoch menggunakan AdamW *optimizer* dengan *learning rate* sebesar 2×10^{-5} dan batch size 16. Fungsi *CrossEntropyLoss* digunakan untuk mengukur kesalahan prediksi, sementara *gradient clipping* dengan batas maksimum 1,0 diterapkan untuk menjaga stabilitas konvergensi. Model kemudian dikemas dalam *Scikit-learn Wrapper* agar kompatibel dengan antarmuka *BaseEstimator*, sehingga memungkinkan integrasi langsung ke dalam skema *soft voting ensemble* bersama model SVM. Seperti halnya SVM, *IndoBERT* menghasilkan distribusi probabilitas kelas yang digunakan sebagai masukan pada tahap *ensemble*. Pemilihan *IndoBERT* didasarkan pada kemampuannya dalam menangkap konteks semantik dan ragam bahasa informal media sosial Indonesia secara lebih unggul dibandingkan model multibahasa [33], [34].

2.8 Soft Voting Ensemble

Model *ensemble* dikembangkan menggunakan pendekatan *weighted soft voting* yang mengintegrasikan dua pendekatan komplementer: *Support Vector Machine* (SVM) berbasis fitur statistik dan *IndoBERT* sebagai model representasi kontekstual Bahasa Indonesia. Masing-masing model menghasilkan distribusi probabilitas kelas sentimen melalui fungsi *predict_proba*(), yang selanjutnya digabungkan menggunakan mekanisme *soft voting* berbasis rata-rata tertimbang (*weighted averaging*). Secara matematis, probabilitas akhir untuk kelas sentimen $c$dirumuskan sebagai:

$$P_{\text{ensemble}}(c) = w_1 \cdot P_{\text{SVM}}(c) + w_2 \cdot P_{\text{IndoBERT}}(c) \quad (8)$$

dengan $w_1 = w_2 = 0{,}5$, yang merepresentasikan bobot setara karena tidak ditemukan dominasi performa yang konsisten dari salah satu model pada seluruh platform media sosial. Prediksi akhir ditentukan berdasarkan kelas dengan nilai P_*ensemble* tertinggi. Model SVM berkontribusi melalui representasi TF-IDF dan fitur tambahan, termasuk skor sentimen (*TextBlob* dan *VADER*), karakteristik linguistik, serta indikator khas media sosial. Sebaliknya, *IndoBERT* menangkap konteks semantik dan variasi bahasa informal secara mendalam melalui arsitektur *Transformer* yang telah dipratrain pada korpus Bahasa Indonesia. Arsitektur lengkap ensemble secara keseluruhan disajikan pada Gambar 4.



Gambar 4. Arsitektur *Ensemble* Model

Gambar 4 memperlihatkan arsitektur *soft voting ensemble* yang mengintegrasikan keluaran probabilitas dari SVM dan *IndoBERT* melalui mekanisme *weighted averaging*. Kombinasi kedua model ini dirancang untuk memanfaatkan keunggulan masing-masing pendekatan dalam menghadapi keragaman ekspresi linguistik pada data media sosial. Pendekatan *ensemble* ini sejalan dengan temuan penelitian sebelumnya yang menunjukkan bahwa integrasi model statistik dan *pre-trained language* model mampu meningkatkan ketahanan dan stabilitas sistem klasifikasi sentimen [35], [36].

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9460
Hal 265-276
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

2.9 Evaluasi Kinerja

Evaluasi model menggunakan akurasi, presisi, *recall*, *F1-score* (metrik utama untuk distribusi kelas tidak merata), dan matriks konfusi untuk identifikasi pola kesalahan (*false positive/negative*) serta bias antarkelas. Distribusi *confidence score* diperiksa untuk menilai konsistensi terhadap variasi bahasa media sosial (slang, singkatan, emoji, kalimat tidak baku). Tinjauan kualitatif prediksi keliru mengidentifikasi situasi ambigu (ironi, pernyataan berlapis konteks), memberikan gambaran menyeluruh stabilitas model pada data dunia nyata.

3. HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil dan evaluasi komprehensif model *Ensemble IndoBERT–SVM* dalam menganalisis opini publik terhadap Program Makan Bergizi Gratis (MBG) di ruang digital Indonesia berdasarkan 7.500 komentar dari *X*, *TikTok*, dan *YouTube* (periode 6 Januari–28 September 2025). Penerapan model mengikuti alur metodologis Bab II: *preprocessing* → pelabelan otomatis berbasis *lexicon-based* → *stratified split* → pelatihan terpisah *SVM* dan *IndoBERT* → integrasi prediksi melalui *soft voting ensemble* → evaluasi pada *test set* proporsional untuk memastikan generalisasi lintas platform dan objektivitas metrik. Pembahasan difokuskan pada (1) distribusi sentimen lintas platform yang mencerminkan keragaman ekspresi publik, dan (2) perbandingan kinerja model tunggal dan *ensemble*.

3.1 Karakteristik Dataset Hasil Crawling

Dataset terdiri dari 7.500 komentar publik dari X, TikTok, dan YouTube (masing-masing 2.500), dikumpulkan periode 6 Januari–28 September 2025 dalam format *.csv* (Tabel 1). Pemilihan data seimbang bertujuan menjaga representasi lintas platform.

Tabel 1. Distribusi Data Hasil *Crawling* per Platform

Platform	Jumlah Data	Periode Pengambilan	Format Penyimpanan
YouTube	2.500	6 Jan – 28 Sep 2025	.csv
TikTok	2.500	6 Jan – 28 Sep 2025	.csv
X	2.500	6 Jan – 28 Sep 2025	.csv
Total	7.500		

Ketiga platform menunjukkan karakteristik linguistik berbeda: komentar di X cenderung argumentatif dan ringkas, TikTok lebih emosional dengan bahasa gaul dan emoji, sedangkan YouTube didominasi komentar deskriptif dan reflektif. Variasi ini menjadi dasar perlunya pendekatan multi-model untuk menangkap kompleksitas opini publik terhadap kebijakan MBG.

Gambar 5. Contoh Komentar per Platform

3.2 Hasil Preprocessing Data

Preprocessing dilakukan untuk membersihkan dan menstandarkan komentar publik agar siap dianalisis oleh model pembelajaran mesin, menggunakan pustaka Sastrawi untuk *stemming* dan *stopword removal*. Proses ini mencakup normalisasi slang ("gk" → "tidak"), pembersihan simbol/URL/emoji, konversi ke huruf kecil, serta penghapusan dokumen kosong pasca-pembersihan. Hasil akhir disimpan dalam dataset_*clean*.xlsx dengan kolom *text* (mentah) dan *clean_text* (hasil *preprocessing*). Alur transformasi lengkap ditunjukkan pada Tabel 2.

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9460
Hal 265-276
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Tabel 2. Hasil Preprocessing Data

No	Text Mentah	Lowercase & Hapus Simbol/Emoji	Normalisasi Slang	Hapus Stopword	Stemming (Clean Text)
1	tadi mbg disekola ak ayam katsu 🤪😋	tadi mbg disekola ak ayam katsu	tadi mbg di sekolah dapat ayam katsu	tadi mbg sekolah dapat ayam katsu	tadi mbg sekolah dapat ayam katsu
2	Program MBG bikin anak-anak nggak cuma hadir, tapi juga berkembang.	program mbg bikin anak anak nggak cuma hadir tapi juga berkembang	program mbg membuat anak anak tidak cuma hadir tetapi juga berkembang	program mbg membuat anak anak cuma hadir berkembang	program mbg buat anak anak cuma hadir kembang

Sebagai ilustrasi alur preprocessing, Tabel 2 menampilkan transformasi teks dari data mentah hingga bentuk akhir yang siap dianalisis. Misalnya, pada contoh komentar *"Makan siang gratisnya keren banget!!! 🤣"* mengalami serangkaian transformasi: pertama, dikonversi ke huruf kecil dan dibersihkan dari emoji serta tanda seru; kedua, tidak memerlukan normalisasi slang karena tidak mengandung bentuk alay; ketiga, kata "gratisnya" dan "banget" difilter, di mana "banget" dihapus sebagai *stopword*; terakhir, dilakukan stemming yang mempertahankan bentuk dasar kata, menghasilkan *"makan siang gratis keren"* sebagai representasi akhir. Proses serupa diterapkan secara konsisten pada seluruh dataset, memastikan konsistensi linguistik dan kesiapan data untuk pemodelan.

3.3. Hasil Pelabelan dan Distribusi Sentimen Awal

Pelabelan sentimen dilakukan secara otomatis menggunakan pendekatan *lexicon-based* (Bagian 2.4) secara terpisah per platform untuk mempertahankan karakteristik linguistik masing-masing. Data kemudian dibagi menggunakan *stratified split* (70% latih, 15% validasi, 15% uji) guna menjaga keseimbangan kelas.

Sebelum analisis lebih lanjut, ditegaskan definisi operasional sentimen yang digunakan dalam penelitian ini. Komentar diklasifikasikan sebagai positif jika mengandung ekspresi apresiatif (bagus, enak, membantu) atau dukungan eksplisit terhadap program. Sebaliknya, komentar dikategorikan sebagai negatif apabila memuat kritik eksplisit (buruk, gagal, kecewa), keluhan, atau pertanyaan retoris bernada skeptis seperti "Gratis? Katanya sih gratis…". Sementara itu, komentar bersifat netral mencakup pernyataan informatif atau faktual tanpa muatan emosional, misalnya "Program MBG mulai Januari 2025". Aturan penanganan negasi diterapkan (tidak bagus → negatif), sedangkan ekspresi ironi atau sarkasme yang tidak dapat ditangkap secara memadai oleh pendekatan leksikal cenderung diklasifikasikan sebagai netral. Distribusi sentimen hasil pelabelan ditunjukkan pada Tabel 3.

Tabel 3. Distribusi Sentimen Lintas Platform (Setelah Labeling dan *Stratified Split*)

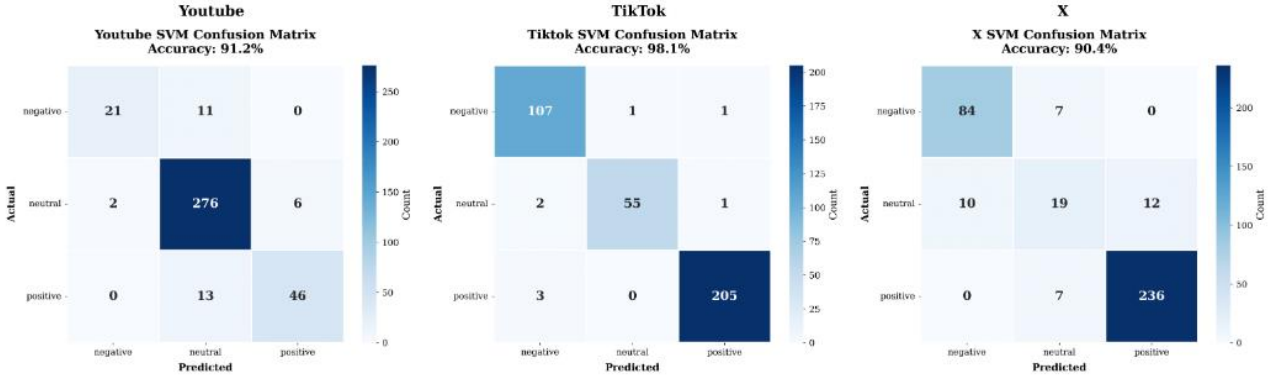
Platform	Kelas Sentimen	Train	Validasi	Test	Total	Persentase
YouTube	Netral	1.325	284	284	1.893	75,72%
	Positif	274	59	59	392	15,68%
	Negatif	151	32	32	215	8,60%
	Subtotal	1.750	375	375	2.500	100%
TikTok	Positif	972	208	208	1.388	55,52%
	Negatif	506	108	109	723	28,92%
	Netral	272	59	58	389	15,56%
	Subtotal	1.750	375	375	2.500	100%
X	Positif	1.134	243	243	1.620	64,82%
	Negatif	425	91	91	607	24,29%
	Netral	190	41	41	272	10,89%
	Subtotal	1.749	375	375	2.499	100%
Total Keseluruhan		5.249	1.125	1.125	7.499	100%

Tabel 3 menunjukkan heterogenitas distribusi sentimen lintas platform. YouTube didominasi sentimen netral (75,72%), mencerminkan komunikasi informatif- deskriptif yang berbeda signifikan dengan TikTok dan X. Kedua platform tersebut didominasi sentimen positif (TikTok 55,52%; X 64,82%), menunjukkan respons pengguna yang lebih ekspresif-emosional terhadap konten audiovisual dan evaluatif- langsung terhadap wacana publik. Sentimen negatif juga lebih tinggi di TikTok (28,92%) dan X (24,29%) dibanding YouTube (8,60%), mengisyaratkan dinamika interaksi lebih intens seperti viralitas atau polarisasi pada platform berbasis audiovisual atau *real-time*. Data X berjumlah 2.499 akibat penghapusan duplikat pada tahap praproses. Secara agregat, sentimen positif paling dominan (45,34%), diikuti netral (34,06%) dan negatif (20,60%). Temuan ini menegaskan bahwa karakter komunikasi digital sangat dipengaruhi oleh arsitektur dan norma tiap platform, sehingga analisis sentimen lintas platform harus mempertimbangkan konteks sosioteknisnya secara spesifik.

3.4. Analisis Performa Model SVM Antar-Platform

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9460
Hal 265-276
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Model *Support Vector Machine* (SVM) dievaluasi pada *stratified test set* setiap platform menggunakan representasi TF-IDF (*n-gram* 1–2, *max_features* = 30.000) dan *kernel linear* terbaik hasil *Grid Search* lima lipatan. Kinerja diukur melalui akurasi, presisi–recall per kelas, dan analisis kesalahan klasifikasi. Hasil evaluasi lengkap disajikan dalam Tabel 7, sedangkan visualisasi pola prediksi ditunjukkan melalui panel *confusion matrix* gabungan pada Gambar 6.



Gambar 6. Panel *Confusion Matrix* SVM Gabungan

Gambar 6 memperlihatkan confusion matrix model SVM pada tiga platform, menunjukkan pola kesalahan klasifikasi yang berbeda sesuai karakteristik linguistik masing-masing. Secara umum, TikTok menunjukkan performa paling tinggi (akurasi 98,1%), disusul YouTube (91,2%) dan X (90,4%). Variasi ini mencerminkan perbedaan karakter linguistik tiap platform.

Pada YouTube, dominasi kelas netral (75,72%) menghasilkan kecenderungan model mengarah ke label mayoritas. Walaupun kelas netral mencapai F1 = 94,0%, kategori negatif menunjukkan kinerja lebih rendah (F1 = 79,0%). Panel *confusion matrix* menampilkan bahwa sebagian kritik bersifat implisit misalnya "Programnya bagus, tapi porsinya kurang" sering terpetakan sebagai netral. Hal ini menegaskan keterbatasan TF-IDF dalam menangkap makna tersirat pada komentar panjang yang bersifat argumentatif.

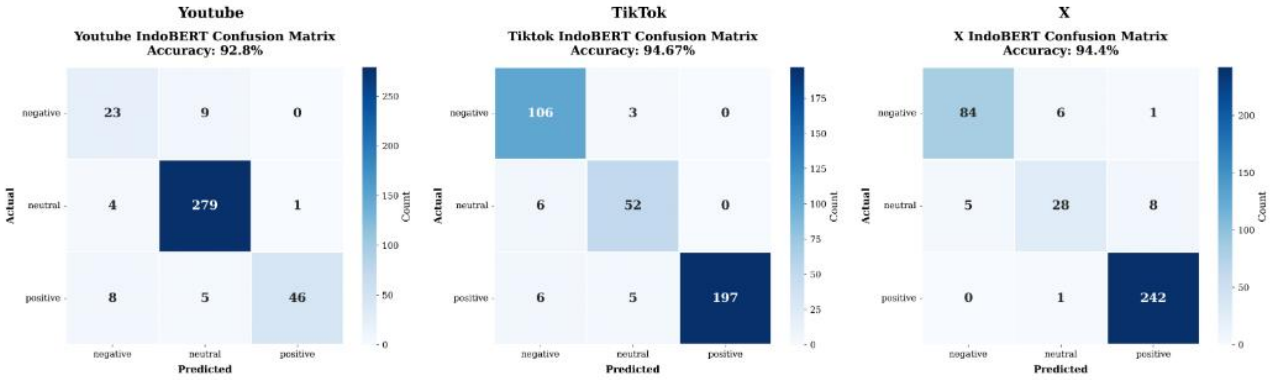
TikTok mencatat performa hampir sempurna untuk seluruh kelas (F1 ≥ 97,0%). Hal ini tampak jelas pada confusion matrix yang memperlihatkan minimnya *misclassification* antar label. Teks TikTok yang singkat, lugas, dan emosional membuat representasi berbasis kata kunci bekerja sangat efektif tanpa kebutuhan pemodelan konteks mendalam.

Pada X, akurasi keseluruhan mencapai 90,4%, namun kelas netral berada pada performa terendah (F1 = 51,0%). Gambar 6 panel *confusion matrix* menunjukkan banyak pernyataan netral dalam konteks debat publik misalnya "Mereka bilang gratis, tapi kok bayar?" salah diklasifikasi sebagai negatif. Model menangkap kata bermuatan emosional tetapi gagal menafsirkan ironi dan pertanyaan retoris, karakteristik yang umum pada diskursus X.

Secara keseluruhan, temuan ini menegaskan bahwa efektivitas SVM sangat bergantung pada ekologi komunikasi platform. Pendekatan berbasis fitur statistik bekerja optimal pada teks yang eksplisit dan sederhana (TikTok), namun kurang tangguh menghadapi ketidakseimbangan kelas, implikatur, dan gaya bahasa argumentatif (YouTube dan X). Keterbatasan tersebut menjadi dasar empiris perlunya model berbasis konteks, yaitu *IndoBERT*, pada tahap analisis berikutnya.

3.5. Analisis Performa Model IndoBERT Antar-Platform

Model *IndoBERT* (*indolem/indobert-base-uncased*) dievaluasi pada stratified test set tiap platform setelah *fine-tuning* selama empat *epoch* (learning rate = 2×10^{-5}, batch size = 16). Evaluasi mencakup akurasi dan *F1-score* per kelas, dengan hasil evaluasi lengkap disajikan dalam Tabel 7**,** pola kesalahan divisualisasikan melalui *confusion matrix* pada Gambar 7.



Gambar 7. Panel *Confusion Matrix Indobert* Gabungan

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9460
Hal 265-276
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Gambar 7 memperlihatkan confusion matrix IndoBERT pada tiga platform. Rata-rata akurasi mencapai 93,96%, dengan variasi performa yang mengikuti karakter wacana tiap platform. Kinerja kelas negatif (F1 = 69,0%) masih di bawah SVM (79,0%) karena komentar YouTube sering mengandung kritik implisit dalam struktur kontras seperti *"tapi"* atau *"meskipun"*. Sesuai Gambar 7, sebanyak 23 komentar negatif diklasifikasi dengan benar dan 9 salah sebagai netral, menunjukkan tantangan model dalam membaca kritik yang terselubung dalam paragraf panjang.

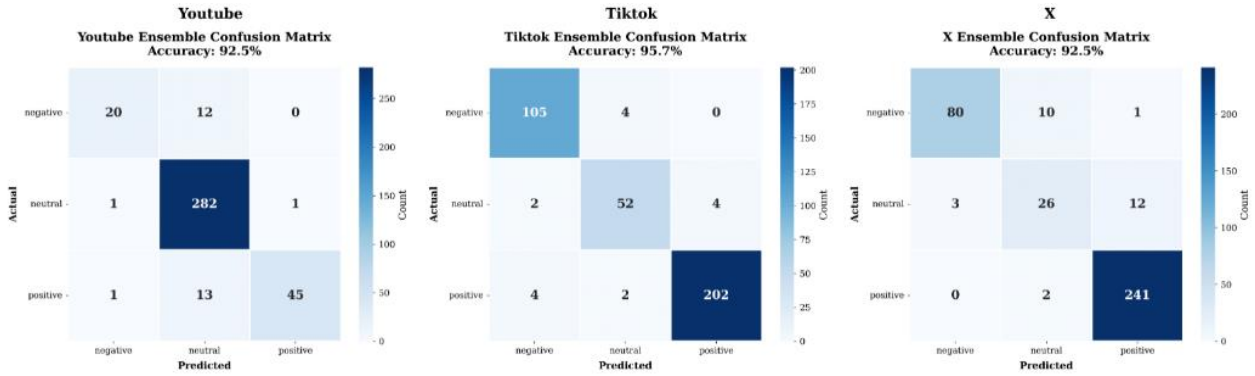
TikTok menjadi platform paling stabil, dengan F1 seluruh kelas di atas 88%. Teks yang pendek dan langsung memudahkan pemanfaatan *embedding IndoBERT*. Gambar 7 memperlihatkan 106 komentar negatif diklasifikasi dengan benar dan hanya 3 keliru sebagai netral, menegaskan akurasi model pada ekspresi eksplisit.

Pada platform X, peningkatan terbesar terjadi pada kelas netral, dengan *F1-score* naik dari 51% pada SVM menjadi 74%. *Confusion matrix* (Gambar 7) menunjukkan bahwa *IndoBERT* mengklasifikasikan 28 komentar netral secara benar, sementara 5 keliru sebagai negatif dan 8 sebagai positif. Hasil ini menegaskan kemampuan model dalam mengenali ambiguitas pragmatik seperti pertanyaan retoris, meskipun masih kesulitan membedakan komentar netral yang memiliki nada evaluatif samar.

Temuan ini mengungkap komplementaritas fungsional antara *IndoBERT* dan SVM: yang pertama unggul dalam wacana kompleks, yang kedua dalam teks eksplisit. *IndoBERT* lebih unggul pada platform dengan wacana kompleks (X), namun sedikit menurun pada platform yang ekspresinya sangat langsung (TikTok dan sebagian YouTube). Hal ini menunjukkan bahwa representasi kontekstual lebih efektif untuk komentar ambigu, sedangkan SVM tetap kompetitif pada teks yang lugas. Perbedaan ini menjadi dasar penggunaan pendekatan *ensemble* pada subbab berikutnya.

3.6. Analisis Model Ensemble IndoBERT–SVM

Model *ensemble* dikembangkan menggunakan pendekatan *weighted soft voting* yang mengintegrasikan probabilitas keluaran SVM dan *IndoBERT*. Strategi ini dirancang untuk memanfaatkan keunggulan komplementer kedua model: SVM yang efektif pada teks eksplisit berbasis *n-gram*, dan *IndoBERT* yang unggul dalam memahami konteks semantik serta ekspresi implisit. Evaluasi dilakukan pada *stratified test set* tiap platform, dengan ringkasan performa ditampilkan pada Tabel 7 dan pola kesalahan divisualisasikan melalui *confusion matrix* pada Gambar 8.



Gambar 8. Panel *Confusion Matrix Ensemble* Gabungan

Gambar 8 memperlihatkan *confusion matrix* model *ensemble* pada tiga platform, menunjukkan peningkatan stabilitas klasifikasi dibanding model tunggal. YouTube. Akurasi mencapai 92,53% dengan *F1 makro* 85,07%. *Confusion matrix* (Gambar 8a) menunjukkan bahwa dari 32 komentar negatif, hanya 20 yang benar diklasifikasi, sementara 12 salah sebagai netral mengonfirmasi tantangan dalam menangani kritik implisit pada teks panjang. Namun, stabilitas kelas netral tetap terjaga (282 benar dari 284), menjadikan F1 netral tetap tinggi (95,43%).

TikTok. Performa paling optimal akurasi 95,73%, *F1 makro* 94,33%. *Confusion matrix* (Gambar 8b) memperlihatkan minimnya *misclassification*: 105/109 negatif benar, 202/208 positif benar, dan 52/58 netral benar. Hasil ini menegaskan efektivitas integrasi SVM–*IndoBERT* pada teks singkat, lugas, dan emosional.

X. Akurasi 92,53%, namun kelas netral tetap menjadi tantangan (F1 = 65,82%). *Confusion matrix* (Gambar 8c) menunjukkan bahwa dari 41 sampel netral, hanya 26 benar diklasifikasi; sedangkan 12 salah sebagai positif dan 3 sebagai negatif. Meskipun tidak sepenuhnya mengatasi ambiguitas pragmatik khas X, *ensemble* mampu mereduksi kecenderungan SVM untuk melakukan prediksi ekstrem pada kelas negatif.

Secara keseluruhan, integrasi *IndoBERT* dan SVM menghasilkan peningkatan stabil pada seluruh platform baik dari sisi akurasi rata-rata maupun penurunan bias antarkelas. Hasil ini menunjukkan bahwa pendekatan hybrid berbasis fitur statistik dan representasi kontekstual lebih adaptif terhadap keragaman gaya bahasa lintas platform, sehingga lebih andal untuk pemetaan sentimen publik dalam konteks kebijakan pemerintah.

Tabel 7. Perbandingan Performa Model Lintas Platform

Model	Platform	Akurasi	F1 Makro	F1 Negatif	F1 Netral	F1 Positif
SVM	YouTube	91,2%	91,0%	79,0%	94,0%	82,0%
	TikTok	98,1%	98,0%	97,0%	97,0%	99,0%

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9460
Hal 265-276
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

	X	90,4%	90,0%	91,0%	51,0%	96,0%
	YouTube	92,8%	84,0%	69,0%	97,0%	87,0%
IndoBERT	TikTok	94,7%	93,0%	93,0%	88,0%	97,0%
	X	94,4%	88,0%	93,0%	74,0%	98,0%
	YouTube	92,53%	85,07%	74,07%	95,43%	85,71%
Ensemble	TikTok	95,73%	94,33%	95,00%	90,00%	98,00%
	X	92,53%	84,92%	91,95%	65,82%	96,98%

Tabel 7 merangkum performa komparatif ketiga model (SVM, IndoBERT, dan Ensemble) di seluruh platform berdasarkan metrik akurasi dan F1-score per kelas. Secara keseluruhan, tidak ada model tunggal yang secara konsisten unggul di semua platform. SVM mencatat performa tertinggi pada TikTok (akurasi 98,1%), IndoBERT unggul pada platform X (akurasi 94,4%), sedangkan Ensemble memberikan keseimbangan terbaik dengan stabilitas lintas kelas—terutama terlihat pada peningkatan F1 negatif di YouTube (dari 69,0% menjadi 74,07%) dan F1 netral di X (dari 51,0% menjadi 65,82%). Temuan ini menegaskan bahwa pendekatan ensemble lebih andal untuk analisis sentimen multi-platform yang memerlukan representasi seimbang lintas kategori sentimen.

3.7. Analisis Komparatif Lintas-Platform

Evaluasi menunjukkan tidak ada model universal yang optimal di semua platform kinerja sangat bergantung pada karakter linguistik tiap ekosistem. Di YouTube, SVM unggul (*F1-makro* 91,0%) berkat stabilitas pada kelas netral, namun lemah pada kritik implisit (F1 negatif 79,0%). *IndoBERT* justru turun (*F1-makro* 84,0%) akibat kesulitan menangkap kritik terselubung, meski unggul pada kelas netral (97,0%). Ensemble menyeimbangkan keduanya (F1 negatif 74,07%, F1 netral 95,43%).

Di TikTok, SVM mencatat performa hampir sempurna (*F1-makro* 98,0%) karena efektivitas pendekatan berbasis *n-gram* pada teks singkat dan emosional. *IndoBERT* sedikit menurun (93,0%), sementara *ensemble* memberikan keseimbangan prediksi antarkelas.

Pada X, perbedaan paling mencolok terjadi pada kelas netral: SVM gagal (F1 = 51,0%), *IndoBERT* unggul (74,0%) berkat kemampuan menangani ambiguitas pragmatik, dan *ensemble* mencapai stabilitas menengah (65,82%) dengan menghindari prediksi ekstrem.

Temuan ini menegaskan bahwa efektivitas analisis sentimen ditentukan oleh konteks platform: YouTube (reflektif), TikTok (emosional), dan X (argumentatif). Pendekatan *ensemble IndoBERT*–SVM menghasilkan representasi opini publik yang lebih seimbang dan adaptif.

Secara komparatif, performa ensemble kami (F1-makro 84,92%–94,33%) melampaui studi sebelumnya yang menggunakan model tunggal pada platform tunggal [2], [9]. Kontribusi utama terletak pada integrasi pendekatan multi-platform dan fitur linguistik khas media sosial, yang memungkinkan adaptasi terhadap variasi ekspresi lintas ekosistem digital sebuah ekstensi signifikan dari temuan Cahyadi & Rochadiani [10] ke konteks kebijakan publik nasional.[2]

4. KESIMPULAN

Penelitian ini menunjukkan bahwa opini publik terhadap Program Makan Bergizi Gratis (MBG) di ruang digital Indonesia bersifat heterogen dan tidak dapat dianalisis menggunakan pendekatan yang seragam. Dari 7.500 komentar yang dikumpulkan dari X, TikTok, dan YouTube selama periode 6 Januari hingga 28 September 2025, terlihat bahwa setiap platform memiliki ekologi komunikasi yang berbeda: YouTube didominasi sentimen netral (75,72%) yang mencerminkan pola diskursif dan reflektif; TikTok menampilkan respons spontan yang cenderung positif (55,52%); sedangkan X menjadi arena debat publik dengan dukungan mayoritas (64,82%) meskipun sarat ambiguitas pragmatik. Dalam konteks keberagaman ini, model *ensemble* berbasis *soft voting* yang menggabungkan SVM dan *IndoBERT* terbukti lebih adaptif daripada model tunggal, mencapai akurasi rata-rata 93,60% dengan F1 makro berkisar antara 84,92% hingga 94,33% tergantung platform. Keunggulan ini bukan semata karena memberikan akurasi tertinggi, tetapi karena mampu mengintegrasikan kekuatan komplementer representasi leksikal dan pemahaman kontekstual sehingga menghasilkan pemetaan sentimen yang lebih stabil dan representatif, terutama dalam menangani kasus ambiguitas seperti kritik implisit pada YouTube (F1 negatif meningkat dari 69,0% menjadi 74,07%) dan ironi pada X (F1 netral 65,82% dibandingkan SVM 51,0%). Temuan ini menegaskan bahwa keberhasilan analisis sentimen dalam evaluasi kebijakan publik tidak hanya ditentukan oleh aspek teknis model, tetapi juga oleh kemampuannya menangkap variasi linguistik, pragmatik, dan budaya komunikasi antarplatform digital. Dengan demikian, pendekatan yang digunakan tidak hanya memperkaya metodologi pemrosesan bahasa alami untuk bahasa Indonesia informal, tetapi juga menyediakan dasar empiris bagi pemerintah dalam merancang strategi komunikasi kebijakan yang lebih responsif, tersegmentasi, dan berbasis bukti di era media sosial.

REFERENCES

[1] R. Puspa, I. Sarifah, and M. Yunus, "Dampak Makan Bergizi Gratis terhadap Minat Belajar Siswa Kelas 5 SD," Pendas: Jurnal Ilmiah Pendidikan Dasar, vol. 10, no. 2, pp. 381–392, 2025, doi: 10.23969/jp.v10i02.25526.

[2] E. Triningsih, M. Afdal, I. Permana, and N. E. Rozanda, "Analisis Sentimen Terhadap Program Makan Bergizi Gratis Menggunakan Algoritma Machine Learning Pada Sosial Media X," Building of Informatics, Technology and Science (BITS), vol. 6, no. 4, pp. 2240–2250, 2025, doi: 10.47065/bits.v6i4.5123.

[3] H. Jayadianti, W. Kaswidjanti, A. T. Utomo, and S. Saifullah, "Sentiment Analysis of Indonesian Reviews Using Fine-Tuning IndoBERT and R-CNN," ILKOM Jurnal Ilmiah, vol. 14, no. 3, pp. 348–354, 2022, doi: 10.33096/ilkom.v14i3.1223.348-354.
[4] D. Nuryadi et al., "Fine Tuning IndoBERT untuk Analisis Sentimen pada Ulasan Pengguna Aplikasi Tiket.com di Google Play Store," JATI (Jurnal Mahasiswa Teknik Informatika), vol. 9, no. 2, pp. 3577–3583, 2025.
[5] S. Aras, M. Yusuf, R. Ruimassa, E. A. B. Wambrauw, and E. B. Palalangan, "Sentiment Analysis on Shopee Product Reviews Using IndoBERT," Journal of Information Systems and Informatics, vol. 6, no. 3, pp. 1616–1627, 2024, doi: 10.51519/journalisi.v6i3.814.
[6] P. Ashari, N. Mutiah, and D. Prawira, "Perbandingan Kinerja Model Deep Learning BERT dan GPT dalam Analisis Sentimen Komentar Video Youtube (Studi Kasus: Film Dirty Vote)," JEPIN (Jurnal Edukasi dan Penelitian Informatika), vol. 11, no. 1, pp. 66–75, 2025.
[7] A. Y. Pratama, G. A. Sanjaya, N. K. Lubis, and M. R. Aditya, "Analisis Sentimen Publik Terkait Danantara Menggunakan Algoritma IndoBERT pada Platform Media Sosial," METIK Jurnal, vol. 9, no. 1, pp. 92–99, 2025.
[8] N. P. V. D. Saraswati, N. Yudistira, and P. P. Adikara, "Analisis Sentimen terhadap Perundungan Siber pada Twitter menggunakan Algoritma Bidirectional Encoder Representations from Transformer (BERT)," Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, vol. 7, no. 2, pp. 909–916, 2023.
[9] N. Agustina and C. N. Ihsan, "Pendekatan Ensemble untuk Analisis Sentimen Covid-19 Menggunakan Pengklasifikasi Soft Voting," Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK), vol. 10, no. 2, pp. 263–270, Apr. 2023, doi: 10.25126/jtiik.2023106215.
[10] M. F. Cahyadi and T. H. Rochadiani, "Implementasi Ensemble Deep Learning untuk Analisis Sentimen terhadap Genre Game Mobile," Jurnal Media Informatika Budidarma, vol. 8, no. 3, pp. 1512–1523, 2024, doi: 10.30865/mib.v8i3.7234.
[11] M. I. K. Sinapoy, Y. Sibaroni, and S. S. Prasetyowati, "Comparison of LSTM and IndoBERT Method in Identifying Hoax on Twitter," Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), vol. 7, no. 3, pp. 657–662, 2023, doi: 10.29207/resti.v7i3.4912.
[12] N. R. Setiawan and E. R. Kaburuan, "Sentimen Analisis Review Aplikasi Digital Korlantas Pada Google Play Store Menggunakan Metode SVM," Jurnal SISFOKOM, vol. 12, no. 1, pp. 105–116, 2023, doi: 10.32736/sisfokom.v12i1.1489.
[13] N. K. A. Krisnasari, I. M. A. D. Suarjaya, and I. M. S. Raharja, "Classification of Public Figures Sentiment on Twitter Using Big Data Technology," JITE (Journal of Informatics and Telecommunication Engineering), vol. 6, no. 1, pp. 157–169, 2022.
[14] S. Uyun, R. P. Rosalin, L. V Sari, and H. H. Sucinta, "A Hybrid Classification Model Based on BERT for Multi-Class Sentiment Analysis on Twitter," JITEKI: Jurnal Ilmiah Teknik Elektro Komputer dan Informatika, vol. 11, no. 2, pp. 194–205, 2025.
[15] D. Napitupulu, R. Fauzi, and M. Lubis, "Sentiment Analysis on Indonesian Tweets about the 2024 Election," Sinkron: Jurnal dan Penelitian Teknik Informatika, vol. 9, no. 1, pp. 44–53, 2025.
[16] R. Agustina, A. U. El Majid, and R. Nuari, "Performance Comparison of BERT Metrics and Classical Machine Learning Models (SVM, Naïve Bayes) for Sentiment Analysis," INOVTEK Polbeng – Seri Informatika, vol. 10, no. 2, pp. 1–10, 2025.
[17] A. N. Syafia, M. F. Hidayattullah, and W. Suteddy, "Studi Komparasi Algoritma SVM dan Random Forest pada Analisis Sentimen Komentar YouTube," JPIT, 2024.
[18] B. Winarko and others, "The Halodoc's Social Media Data Analysis Amidst Covid-19 Pandemic: An Evidence from Indonesia," Jurnal Sosial Humaniora (JSH), 2021.
[19] S. K. Syari, S. P. Simarmata, and K. Widyanti, "From Privacy to Oblivion: The Right to Be Forgotten in Indonesia's Social Media Platforms," Law Journal, vol. 6, no. 1, 2025, doi: 10.46576/lj.v6i1.6895.
[20] A. M. U. Albab, "Optimization of the Stemming Technique on Text Preprocessing," Transformatika: Jurnal Bahasa, Sastra, dan Pembelajaran, vol. 6, no. 1, 2023.
[21] J. T. I. dan Ilmu Komputer, "Perbandingan Algoritma Stemming Porter, Sastrawi, Idris, dan Arifin–Setiono untuk Bahasa Indonesia," JTIIK, vol. 12, no. 2, 2025.
[22] A. B. Mutiara, "Improving the Accuracy of Text Classification using Stemming," J. Big Data, vol. 8, no. 4, 2021.
[23] T. Salsabilla and D. Alita, "Analisis Sentimen Inses di Social Media menggunakan Algoritma Naïve Bayes," JPIT, vol. 9, no. 3, 2024, doi: 10.30591/jpit.v9i3.6611.
[24] R. R. Gunawan and D. Y. Rahayu, "Evaluation Analysis of the Necessity of Stemming and Stopword Removal in Indonesian Text Classification," Jurnal Matrik, vol. 19, no. 1, 2023.
[25] N. Adila, "Implementation of Web Scraping for Journal Data Collection on the SINTA Website," Sinkron: Jurnal dan Penelitian Teknik Informatika, vol. 7, no. 4, 2022.
[26] R. Anggoro, "Exploring Stemming Techniques in Ambon Malay Languages," Jurnal Teknologi Informasi UNG, vol. 9, no. 2, 2024.
[27] R. L. Mustofa and B. Prasetiyo, "Sentiment analysis using lexicon-based method with Naive Bayes classifier algorithm on #newnormal hashtag in Twitter," in Journal of Physics: Conference Series, 2021, p. 42155. doi: 10.1088/1742-6596/1918/4/042155.
[28] R. Oktafiani, A. Hermawan, and D. Avianto, "Pengaruh Komposisi Split Data terhadap Performa Klasifikasi Penyakit Kanker Payudara Menggunakan Algoritma Machine Learning," Jurnal Sains dan Informatika (JSI), vol. 9, no. 1, pp. 45–54, 2023, doi: 10.34128/jsi.v9i1.622.
[29] N. Hafidz and D. Y. Liliana, "Klasifikasi Sentimen pada Twitter Terhadap WHO Terkait Covid-19 Menggunakan SVM, N-Gram, PSO," Jurnal RESTI, vol. 5, no. 2, pp. 213–219, 2021, doi: 10.29207/resti.v5i2.2960.
[30] F. Illia, M. P. Eugenia, and S. A. Rutba, "Sentiment Analysis on PeduliLindungi Application Using TextBlob and VADER Library," in International Conference on Data Science and Official Statistics (ICDSOS), 2021, pp. 278–288. doi: 10.34123/icdsos.v2021i1.236.
[31] M. Desiawan and A. Solichin, "SVM Optimization with Grid Search Cross Validation for Improving Accuracy of Schizophrenia Classification Based on EEG Signal," Jurnal Teknik Informatika, vol. 17, no. 1, pp. 10–20, 2024, doi: 10.15408/jti.v17i1.37422.
[32] L. Geni, E. Yulianti, and D. S. Indra, "Sentiment Analysis of Tweets Before the 2024 Elections in Indonesia Using IndoBERT Language Models," JITEKI, vol. 9, no. 3, pp. 746–757, 2023, doi: 10.26555/jiteki.v9i3.26490.
[33] C. J. L. Tobing, I. G. W. Wijayakusuma, and L. P. I. Harini, "Perbandingan Kinerja IndoBERT dan mBERT untuk Deteksi Berita Hoaks Politik dalam Bahasa Indonesia," Jurnal Sains dan Teknologi, vol. 14, no. 1, pp. 114–123, 2025.

[34] A. A. P. Simarmata and T. B. Sasongko, "Sentiment Analysis on BRImo Application Reviews Using IndoBERT," Journal, 2023.
[35] A. N. Nurrahman, R. Ramadhani, and A. Rahman, "Sentiment Analysis of Indonesian Datasets Based on a Hybrid Deep-Learning Strategy," Indonesian Journal of Electrical Engineering and Computer Science, vol. 29, no. 2, pp. 950–958, 2023.
[36] D. H. Handayani and T. K. Nur, "The Hybrid of BERT and Deep Learning Models for Indonesian Sentiment Analysis," Bulletin of Electrical Engineering and Informatics, vol. 13, no. 1, pp. 221–230, 2024.