Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9507

Analisis Sentimen Komentar iPhone 17 pada Platform YouTube Menggunakan IndoBERT dan Support Vector Machine

Siti Mar'atus Sholihah*, Afril Efan Pajri, Ita Aristia Sa'ida
Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Nahdlatul Ulama Sunan Giri, Bojonegoro, Indonesia
Email: [1,*]sitimaratuss893@gmail.com, [2]afril@unugiri.ac.id, [3]itaaristia@unugiri.ac.id
Email Penulis Korespondensi: sitimaratuss893@gmail.com*
Submitted: **24/01/2026**; Accepted: **20/02/2026**; Published: **31/03/2026**

Abstrak– Penelitian ini bertujuan untuk menganalisis sentimen komentar *YouTube* berbahasa Indonesia terkait iPhone 17 dengan membandingkan metode *Support Vector Machine* berbasis *Term Frequency–Inverse Document Frequency* dan model *IndoBERT*. Data diperoleh melalui proses *crawling* komentar pada kanal *YouTube* GadgetIn, kemudian diproses melalui tahapan *pre-processing* untuk mengurangi *noise* dan menormalkan teks. Pelabelan sentimen dilakukan secara otomatis menggunakan *InSetLexicon* dengan dua kelas, yaitu positif dan negatif. *Dataset* selanjutnya dibagi menggunakan teknik *stratified split* menjadi data latih, validasi, dan uji. Selain dua model utama, pendekatan *ensemble IndoBERT–SVM* diuji sebagai metode tambahan untuk menilai stabilitas performa klasifikasi. Evaluasi dilakukan menggunakan *confusion matrix* serta metrik *Accuracy*, *Precision*, *Recall*, dan *F1-score*. Hasil pengujian menunjukkan bahwa *IndoBERT* memperoleh performa terbaik dengan nilai *Accuracy* sebesar 92, 29%, diikuti oleh model *ensemble* sebesar 91,63%, dan *Support Vector Machine* sebesar 88,99%. Temuan ini mengindikasikan bahwa model berbasis *transformer* lebih efektif dalam memahami konteks bahasa informal pada komentar *YouTube* dibandingkan metode berbasis fitur tradisional. Dengan demikian, penelitian ini memberikan bukti empiris mengenai efektivitas pendekatan *machine learning* dan *transformer* dalam analisis sentimen *media sosial* berbahasa Indonesia.

Kata Kunci: Analisis Sentimen; *YouTube*; iPhone 17; *IndoBERT*; SVM

Abstract– This study aims to analyze the sentiment of Indonesian-language *YouTube* comments related to the iPhone 17 by comparing a *Support Vector Machine* model based on *Term Frequency–Inverse Document Frequency* with the *IndoBERT* model. The data were collected through a comment *crawling* process from the GadgetIn *YouTube channel* and subsequently processed through a series of *pre-processing* steps to reduce *noise* and normalize the text. Sentiment *labeling* was performed automatically using *InSetLexicon* with two classes, namely positive and negative. The *dataset* was then divided using a *stratified split* technique into training, validation, and *testing* sets. In addition to the two primary models, an *IndoBERT–SVM ensemble* approach was evaluated as an additional method to assess classification *performance* stability. Model evaluation was conducted using a *confusion matrix* and the metrics *Accuracy*, *Precision*, *Recall*, and *F1-score*. The experimental results indicate that *IndoBERT* achieved the best *performance* with an *Accuracy* of 92.29%, followed by the *ensemble* model at 91.63% and the *Support Vector Machine* at 88.99%. These findings suggest that *transformer*-based models are more effective in capturing the *context* of *informal language* in *YouTube* comments compared to traditional *feature*-based methods. Therefore, this study provides *empirical* evidence of the effectiveness of *machine learning* and *transformer*-based approaches for sentiment analysis of Indonesian *social media* text.

Keywords: Sentiment Analysis; *YouTube*; iPhone 17; *IndoBERT*; SVM

1. PENDAHULUAN

Perkembangan industri *smartphone* menunjukkan pertumbuhan yang sangat cepat, dengan *Apple Inc.* sebagai salah satu perusahaan yang secara konsisten menghadirkan inovasi teknologi. Kehadiran iPhone 17 menimbulkan berbagai respons dari masyarakat, khususnya di platform *YouTube*. Kanal *GadgetIn* sebagai salah satu *reviewer* teknologi dengan jumlah audiens besar menjadi wadah diskusi yang aktif, sehingga kolom komentarnya dipenuhi beragam opini, kritik, ekspektasi, serta pertimbangan pembelian. Opini publik terhadap produk baru memiliki peran penting dalam mendukung strategi pemasaran [1] karena mencerminkan persepsi dan sikap konsumen sebelum produk resmi dirilis. Oleh sebab itu, analisis sentimen diperlukan untuk mengidentifikasi kecenderungan opini secara sistematis dan terukur [2], [3]. Informasi yang dihasilkan tidak hanya bermanfaat bagi perusahaan dalam merumuskan strategi yang lebih tepat, tetapi juga membantu calon konsumen dalam mengambil keputusan berdasarkan pandangan pengguna lain.

Meskipun analisis sentimen telah banyak dikembangkan, penerapannya pada komentar *YouTube* berbahasa Indonesia masih menghadapi berbagai kendala. Karakteristik bahasa yang digunakan cenderung tidak baku, mengandung singkatan, campuran bahasa, sarkasme, hingga ekspresi emosional yang kompleks [4]. Permasalahan utama pada studi sebelumnya terletak pada penggunaan pendekatan berbasis fitur statistik yang kurang mampu memahami konteks kalimat secara menyeluruh [5]. Selain itu, ketidakseimbangan distribusi kelas sentimen berpotensi menyebabkan bias terhadap kelas tertentu [6]. Keterbatasan jumlah *dataset* berlabel pada produk teknologi pra-rilis juga turut memengaruhi kualitas model yang dihasilkan . Kondisi tersebut menunjukkan perlunya pendekatan yang lebih adaptif dalam menangani teks informal di *media sosial*.

Pendekatan berbasis fitur statistik sering kali hanya merepresentasikan frekuensi kemunculan kata tanpa mempertimbangkan hubungan semantik dan konteks kalimat. Akibatnya, model tradisional cenderung kesulitan

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9507

memahami struktur bahasa yang kompleks dan nuansa implisit pada teks informal. Pada konteks produk pra-rilis seperti iPhone 17, opini publik dapat bersifat spekulatif dan ambigu, sehingga dibutuhkan pendekatan yang mampu menangkap representasi kontekstual secara lebih mendalam .

Berbagai penelitian terdahulu telah mencoba mengatasi kendala tersebut dengan menerapkan metode yang berbeda. Penggunaan *Support Vector Machine* dengan fitur *Term Frequency–Inverse Document Frequency* dilaporkan mampu mencapai akurasi 85,3% pada analisis ulasan produk elektronik [7]. Namun demikian, metode tersebut masih memiliki keterbatasan dalam menangkap hubungan semantik pada kalimat yang kompleks. Studi lain menunjukkan bahwa model berbasis *Bidirectional Encoder Representations from Transformers* mampu meningkatkan performa klasifikasi dibandingkan metode tradisional [8], meskipun tantangan terkait ketidakseimbangan data di *media sosial* belum sepenuhnya teratasi [9], [10]. Pendekatan *ensemble* juga dikembangkan untuk meningkatkan stabilitas prediksi dengan menggabungkan beberapa model [11], sementara teknik seperti *Synthetic Minority Over-sampling Technique* digunakan untuk menangani distribusi kelas yang tidak merata [12]. Hasil-hasil tersebut memperlihatkan bahwa pemilihan metode dan strategi pengolahan data sangat menentukan kualitas analisis sentimen [13].

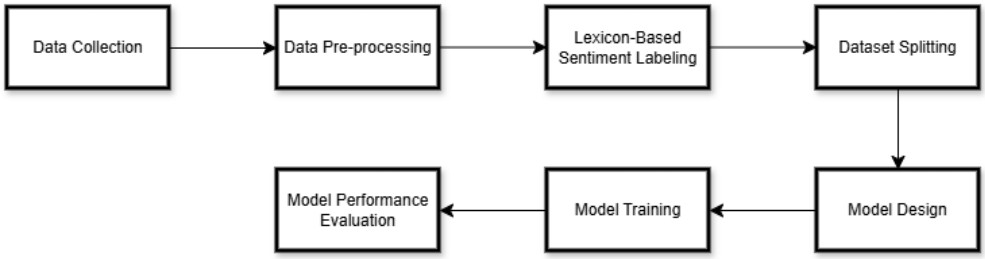
Berdasarkan kajian tersebut, masih terdapat celah penelitian yang belum banyak dieksplorasi, terutama pada analisis sentimen komentar *YouTube* berbahasa Indonesia mengenai produk teknologi yang belum dirilis secara resmi. Perbandingan langsung antara metode klasik seperti *Support Vector Machine* dan pendekatan berbasis *transformer* seperti *IndoBERT* pada konteks ini masih terbatas [14], [15]. Selain itu, penelitian yang secara khusus membahas opini publik terhadap iPhone 17 pada fase pra-rilis belum banyak ditemukan. Kebaruan penelitian ini terletak pada penerapan evaluasi komparatif antara pendekatan berbasis fitur statistik dan model bahasa kontekstual pada *dataset* yang sama, serta pengujian model *ensemble IndoBERT–SVM* untuk memperoleh pendekatan yang lebih stabil dan representatif [16].

Penelitian ini bertujuan untuk mengembangkan model klasifikasi sentimen berbasis *Support Vector Machine* dengan fitur *Term Frequency–Inverse Document Frequency*, melatih serta mengevaluasi model *IndoBERT*, dan menguji performa model *ensemble IndoBERT–SVM* pada *dataset* komentar *YouTube* yang sama. Evaluasi dilakukan menggunakan *confusion matrix* serta metrik *Accuracy*, *Precision*, *Recall*, dan *F1-score* guna membandingkan kinerja setiap model secara objektif. Kontribusi penelitian ini meliputi penyediaan analisis komparatif berbasis metrik kuantitatif pada teks informal berbahasa Indonesia, serta pemberian gambaran empiris mengenai kecenderungan opini publik terhadap iPhone 17 pada fase pra-rilis. Temuan ini diharapkan dapat memperkaya literatur analisis sentimen sekaligus menjadi rujukan dalam pengembangan pendekatan berbasis *transformer* dan *ensemble* untuk data *media sosial*.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Pada penelitian ini, penulis mengembangkan sistem analisis sentimen untuk mengidentifikasi sentimen yang diungkapkan oleh pengguna platform *YouTube* terhadap produk iPhone 17. Sentimen yang dianalisis diklasifikasikan ke dalam dua kategori, yaitu sentimen positif dan negatif, dengan menerapkan metode *Support Vector Machine* dan *IndoBERT*. Tujuan dari penelitian ini adalah untuk memperoleh gambaran tingkat klasifikasi sentimen pengguna *YouTube* serta mengevaluasi kinerja kedua metode tersebut dalam konteks analisis sentimen komentar berbahasa Indonesia. Seluruh eksperimen dilakukan menggunakan skema pembagian data yang sama, yaitu data latih, data validasi, dan data uji, guna memastikan perbandingan kinerja antar model berlangsung secara adil dan objektif. Penelitian ini dilakukan melalui beberapa tahapan, yaitu pengumpulan data, prapemrosesan teks, pelabelan sentimen, pembagian *dataset*, ekstraksi fitur, perancangan dan pelatihan model, serta evaluasi kinerja. Seluruh tahapan disusun secara sistematis sebagaimana ditunjukkan pada Gambar 1 untuk memastikan proses eksperimen berjalan terstruktur dan dapat direplikasi [17].



Gambar 1. Rancangan penelitian

2.2 Data Collection

Data Collection merupakan tahap awal penelitian yang bertujuan untuk memperoleh data utama berupa komentar pengguna *YouTube* berbahasa Indonesia yang berkaitan dengan iPhone 17. Tahap *Data Collection* dilakukan melalui proses *crawling* menggunakan *YouTube Application Programming Interface*. Data dikumpulkan dari *video* kanal *GadgetIn* yang membahas iPhone 17 pada periode Oktober–November 2024 [18]. Sebanyak 3.053

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9507

komentar berhasil dikumpulkan. Setelah dilakukan penyaringan komentar kosong dan duplikat, diperoleh 3.022 komentar valid. Data tersebut merepresentasikan opini, pertanyaan, serta tanggapan pengguna terhadap iPhone 17.

Tabel 1. Ringkasan *Dataset*

Keterangan	Jumlah
Total komentar awal	3.053
Setelah pembersihan	3.022
Label Positif	1.527
Label Negatif	1.495

Distribusi yang relatif seimbang menunjukkan tidak diperlukan teknik *oversampling* maupun *undersampling*.

2.3 Data *Pre-processing*

Data *Pre-processing* dilakukan untuk meningkatkan kualitas data teks sebelum digunakan dalam proses analisis sentimen. Pada tahap ini, komentar dibersihkan dari karakter noninformatif seperti tanda baca, angka, simbol, dan tautan. Selanjutnya dilakukan penyamaan bentuk huruf untuk menghindari perbedaan makna akibat variasi penulisan. Proses tokenisasi diterapkan untuk memisahkan teks menjadi unit kata, diikuti dengan normalisasi kata tidak baku ke bentuk baku. Kata-kata umum yang tidak memiliki kontribusi terhadap makna sentimen dihapus, kemudian dilakukan *stemming* untuk mengubah kata ke bentuk dasar. Selain itu, komentar kosong dan komentar duplikat dihapus guna mengurangi *noise* serta mencegah potensi kebocoran data pada tahap pemodelan [19].

2.4 Lexicon-Based Sentiment *Labeling*

Lexicon-Based Sentiment *Labeling* komentar yang telah melalui tahap prapemrosesan diberi label sentimen secara otomatis menggunakan pendekatan berbasis leksikon. Proses pelabelan dilakukan dengan menghitung skor polaritas kata-kata yang terdapat dalam setiap komentar berdasarkan kamus sentimen yang digunakan. Berdasarkan skor tersebut, komentar diklasifikasikan ke dalam dua kategori sentimen, yaitu positif dan negatif. Hasil pelabelan ini berperan sebagai data berlabel yang menjadi dasar dalam proses pelatihan dan evaluasi model klasifikasi sentimen [20].

$$Score(d) = \sum_{i=1}^{n} polarity(word_i) \quad (1)$$

Aturan klasifikasi:

a. Skor > 0 → Positif
b. Skor < 0 → Negatif
c. Skor = 0 → ditinjau konteks

Contoh Perhitungan Manual

Komentar: “kamera bagus tapi mahal”

Tabel 2. Perhitungan Manual

Kata	Skor
bagus	+1
mahal	-1

Total skor = 0 → ditentukan berdasarkan dominasi konteks → Negatif.

2.5 *Dataset* Splitting

Dataset Splitting bertujuan untuk membagi data berlabel ke dalam beberapa subset agar proses pelatihan dan evaluasi model dapat dilakukan secara objektif. Data dibagi menjadi data latih, data validasi, dan data uji dengan menggunakan metode pembagian berstrata. Pendekatan ini dipilih untuk menjaga proporsi kelas sentimen positif dan negatif agar tetap seimbang pada setiap subset data. Pembagian *dataset* dilakukan satu kali dan digunakan secara konsisten pada seluruh eksperimen guna memastikan perbandingan kinerja antar model bersifat adil [21].
Dataset dibagi menggunakan metode *Stratified split* dengan rasio:

a. 70% *Training*
b. 15% *Validation*
c. 15% *Testing*

Dengan *random_state* = 42.
Jika jumlah keseluruhan data dinyatakan sebagai N, maka pembagian *dataset* dapat dirumuskan sebagai berikut:

$$N_{train} = \alpha \times N \quad (2)$$
$$N_{validation} = \beta \times N \quad (3)$$
$$\boldsymbol{N_{test} = \gamma \times N} \quad (4)$$

dengan ketentuan $\alpha + \beta + \gamma = 1$.
Selain itu, untuk menjaga keseimbangan kelas, proporsi kelas pada setiap subset memenuhi persamaan:

$$\frac{N_{train}^{(c)}}{N_{train}} = \frac{N_{validation}^{(c)}}{N_{validation}} = \frac{N_{test}^{(c)}}{N_{test}} = \frac{N^{(c)}}{N} \quad (5)$$

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9507

dengan c menyatakan kelas sentimen.
Tabel 3 menyajikan distribusi *dataset* yang digunakan dalam penelitian ini.

Tabel 3. Distribusi *Dataset*

Subset	Jumlah Data	Positif	Negatif
Data Latih	2.115	1.069	1.046
Data Validasi	453	229	224
Data Uji	454	229	225

2.6 Model Design

Model Design, penelitian ini membandingkan dua pendekatan klasifikasi sentimen, yaitu *Support Vector Machine* dan *IndoBERT*. *Support Vector Machine* digunakan sebagai metode klasifikasi berbasis fitur, di mana komentar terlebih dahulu direpresentasikan dalam bentuk vektor numerik menggunakan bobot *Term Frequency–Inverse Document Frequency*. Model *Support Vector Machine* dilatih menggunakan kernel *Radial Basis Function* dengan parameter hasil pengujian, yaitu nilai parameter regularisasi sebesar 100 dan nilai gamma sebesar 0,01, untuk menghasilkan pemisahan kelas yang optimal pada ruang fitur [22].

2.6.1. *Term Frequency–Inverse Document Frequency*

Secara matematis, bobot *Term Frequency–Inverse Document Frequency* dihitung sebagai berikut:

$$TF - IDF(d,t) = TF(d,t) \times IDF(t) \quad (6)$$

$$TF(d,t) = \frac{ft,d}{\sum t' \in d ft',d} \quad (7)$$

$$IDF(t) = \log\left(\frac{N}{dft}\right) \quad (8)$$

dengan $f_{t,d}$menyatakan frekuensi kemunculan term $t$pada dokumen d, Nmerupakan jumlah seluruh dokumen, dan df_tadalah jumlah dokumen yang mengandung term t [23].

Contoh Perhitungan Manual

Misal:

a. N = 3022
b. df(mahal) = 300

$$IDF(mahal) = \log\frac{3022}{300} = 1.00 \quad (9)$$

ika kata muncul 2 kali dari 10 kata:

$$TF = 0.2$$

$$TF - IDF = 0.2 \times 1.00 = 0.2 \quad (10)$$

2.6.2. *Support Vector Machine*

Klasifikasi menggunakan *Support Vector Machine* dilakukan dengan fungsi keputusan:

$$f(x) = w \cdot x + b \quad (11)$$

2.6.3. *IndoBERT Fine-tuning*

Untuk menangani data yang bersifat nonlinier, digunakan fungsi kernel *Radial Basis Function* yang dirumuskan sebagai berikut:

$$K(xi,xj) = exp(-\gamma \parallel xi - xj \parallel^2) \quad (12)$$

Sebagai pembanding, *IndoBERT* digunakan melalui pendekatan *fine-tuning* untuk tugas klasifikasi sentimen biner. *IndoBERT* membangun representasi kontekstual dua arah sehingga mampu menangkap makna kata berdasarkan konteks kalimat secara menyeluruh. Token khusus [CLS] dimanfaatkan sebagai representasi ringkas dari keseluruhan komentar, yang selanjutnya dipetakan ke kelas sentimen positif dan negatif melalui lapisan klasifikasi. Proses pelatihan dilakukan menggunakan optimizer AdamW, sedangkan pengujian dilakukan pada data validasi dan data [24], [25].

2.7 Model Training

Model Training dilakukan dengan melatih masing-masing model menggunakan data latih yang telah disiapkan. Pada *Support Vector Machine*, proses pelatihan bertujuan untuk menentukan parameter model yang menghasilkan pemisahan kelas sentimen secara optimal berdasarkan representasi fitur *Term Frequency–Inverse Document Frequency*. Sementara itu, pada *IndoBERT*, proses pelatihan dilakukan melalui penyesuaian parameter model menggunakan pendekatan *fine-tuning* agar model mampu mempelajari pola sentimen dalam komentar *YouTube* berbahasa Indonesia [26].

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9507

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (13)$$

$$Precision = \frac{TP}{TP+FP} \quad (14)$$

$$Recall = \frac{TP}{TP+FN} \quad (15)$$

$$F1-score = \frac{2\times Precision\times Recall}{Precision\times Recall} \quad (16)$$

Penggunaan *Accuracy* bertujuan untuk mengukur ketepatan prediksi secara keseluruhan, sedangkan *Precision* dan *Recall* digunakan untuk menilai ketepatan serta kelengkapan prediksi pada masing-masing kelas sentimen. *F1-score* digunakan sebagai ukuran gabungan untuk menyeimbangkan nilai *Precision* dan *Recall*, sehingga evaluasi kinerja model menjadi lebih representatif.

2.8 Model *Performance* Evaluation

Model *Performance* Evaluation bertujuan untuk menilai kinerja model dalam mengklasifikasikan sentimen komentar. Evaluasi dilakukan menggunakan *confusion matrix* untuk mengetahui jumlah prediksi benar dan salah pada setiap kelas sentimen. Berdasarkan *confusion matrix* tersebut, kinerja model dianalisis menggunakan metrik *Accuracy*, *Precision*, *Recall*, dan *F1-score* yang dirumuskan sebagai berikut [27]:

Tabel 4. *Confusion matrix*

	Predicted Positive	**Predicted Negative**
Actual Positive	TP	FN
Actual Negative	FP	TN

Evaluasi dilakukan pada data uji dengan skema pembagian data yang sama untuk seluruh model sehingga perbandingan kinerja antar model dapat dilakukan secara objektif.

2.9 Tools dan Platform Pengembangan

1. ***IndoBERT*** ialah model pra-latih berlandaskan *Transformer* yang dirancang untuk bahasa Indonesia dan dilatih melalui berbagai sumber teks lokal. Model ini mampu memahami konteks secara dua arah sehingga efektif menangkap makna komentar yang bersifat informal, termasuk singkatan dan variasi bahasa percakapan pada *YouTube* [28].
2. ***Support Vector Machine* (SVM)** ialah algoritma klasifikasi yang menentukan *hyperplane* optimal dalam membagi kelas positif dan negatif pada ruang fitur. SVM dikenal stabil untuk data berdimensi tinggi dan pada penelitian ini digunakan bersama fitur *TF-IDF*, kemudian dievaluasi pada data uji yang sama dengan *IndoBERT* agar perbandingan kinerja tetap objektif [29].
3. **Kondisi Distribusi Data Sentimen** pada *dataset* penelitian ini relatif seimbang antara kelas positif dan negatif. Karena itu, tidak diperlukan teknik penyeimbangan tambahan seperti *oversampling* atau *undersampling*. Penggunaan data asli dipilih untuk menjaga karakteristik alami data dan menghindari distorsi maupun potensi *data leakage* saat evaluasi [30].

2.10 Experimental Setup

Eksperimen pada penelitian ini dilakukan secara sistematis untuk menjaga reproduksibilitas dan mencegah *data leakage*. *Dataset* dibagi satu kali menggunakan metode *stratified split* menjadi data latih, validasi, dan uji dengan *seed* tetap (*random_state* = 42), sehingga perbandingan kinerja model berlangsung secara adil. Karena distribusi kelas sentimen relatif seimbang, penelitian ini tidak menerapkan teknik penyeimbangan data tambahan. Proses ekstraksi fitur *TF-IDF* dilakukan dengan prinsip *fit on training data* dan *transform on validation/testing data* agar evaluasi mencerminkan kemampuan generalisasi model terhadap data baru. Selanjutnya, tahapan penelitian dilanjutkan pada penyajian hasil implementasi yang meliputi data komentar, hasil *pre-processing*, distribusi label sentimen, serta evaluasi kinerja model *Support Vector Machine*, *IndoBERT*, dan model *ensemble* [31].

3. HASIL DAN PEMBAHASAN

3.1 Paparan Data dan *Pre-processing*

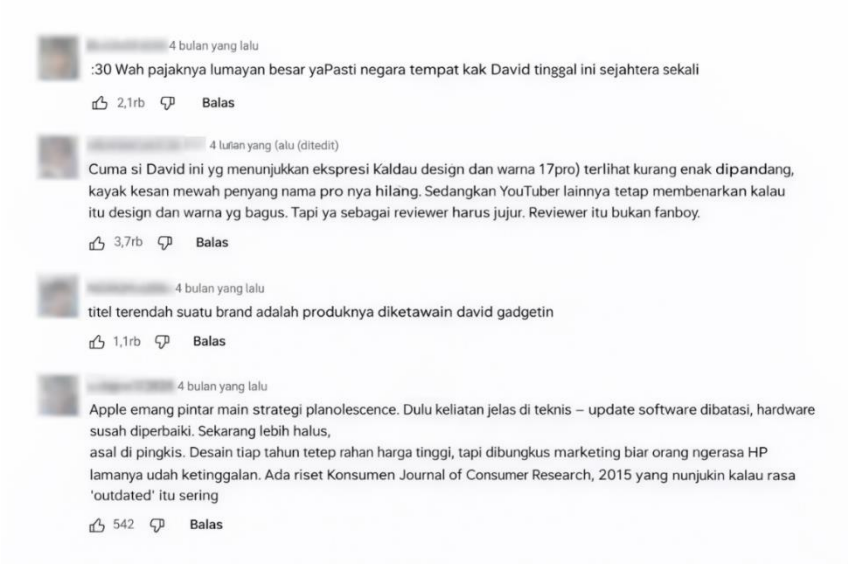
3.1.1 Sumber dan Karakteristik Data

Penelitian ini menggunakan data primer berupa komentar pengguna *YouTube* untuk merepresentasikan persepsi publik terhadap produk iPhone 17. Data dikumpulkan melalui proses *crawling* menggunakan *YouTube* Application Programming Interface dengan sumber utama berasal dari *video* kanal GadgetIn berjudul *"iPhone 17 First Look! Inilah Spesifikasi dan Harganya"* yang diunggah pada 25 Oktober 2024. Pengumpulan komentar dilakukan pada periode Oktober hingga November 2024 dengan fokus pada komentar awal setelah *video* dipublikasikan, sehingga opini yang diperoleh mencerminkan respons spontan pengguna.

Untuk memberikan gambaran awal mengenai karakteristik data yang digunakan, contoh sampel komentar pengguna ditampilkan dalam bentuk visual pada **Gambar 2**. Sampel ini menunjukkan variasi opini pengguna, mulai dari kritik terhadap harga, penilaian desain produk, hingga pertanyaan terkait keputusan pembelian. Identitas

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9507

pengguna pada gambar disamarkan untuk menjaga privasi, sedangkan isi komentar tetap dipertahankan guna merepresentasikan data asli yang dianalisis.



Gambar 2. Sampel Komentar *YouTube* tentang iPhone 17

Selain ditampilkan dalam bentuk gambar, beberapa contoh komentar juga disajikan dalam bentuk teks pada Tabel 1 untuk memudahkan pembacaan isi data. Tabel tersebut menunjukkan bahwa komentar pengguna bersifat beragam, mencakup sentimen positif maupun negatif. Setelah melalui tahap prapemrosesan, jumlah data berkurang dari 3.053 komentar awal menjadi 3.022 komentar valid yang selanjutnya digunakan pada tahap pelabelan dan pemodelan.

Tabel 5. Sampel Data Komentar

No	Komentar
1	*Timestamp*: 00:00 iPhone 17 SERIES! + pajak. Wah kelihatannya makin mahal ya dibanding seri sebelumnya, tapi fiturnya juga makin canggih.
2	Aslinya iPhone itu bisa banget lo bikin kamera jadi kayak profesional, hasil *video*nya stabil dan tajam banget, cocok buat konten kreator.
3	Tolong sarannya 🙏 iPhone saya 13 Pro, lebih bagus ganti ke iPhone 17 atau nunggu seri berikutnya ya? Soalnya masih bingung dari segi kamera dan baterai.

Tabel 1 menunjukkan beragam respons pengguna terhadap iPhone 17, mulai dari komentar apresiatif hingga pertanyaan terkait keputusan pembelian. Setelah prapemrosesan, jumlah data berkurang dari **3.053** menjadi **3.022** komentar valid yang selanjutnya digunakan pada tahap pelabelan dan pemodelan.

3.1.2 Tahapan *Pre-processing* Teks

Sebelum dilakukan pelabelan dan pemodelan, data komentar yang telah dikumpulkan diproses melalui tahapan *pre-processing* untuk meningkatkan kualitas data teks. Proses ini bertujuan untuk mengurangi *noise* seperti URL, emoji, simbol, dan karakter tidak relevan lainnya yang umum ditemukan pada komentar *YouTube*. Tahapan *pre-processing* meliputi pembersihan teks (*cleaning*), penyamaan huruf (*case folding*), tokenisasi, normalisasi kata tidak baku, penghapusan *stopword*, *stemming*, serta penyusunan kembali token menjadi teks akhir (*detokenized*).

Selain itu, komentar kosong maupun komentar yang menjadi kosong setelah pembersihan dihapus dari *dataset*. Komentar duplikat juga dieliminasi untuk meminimalkan kemiripan antar data yang berpotensi memengaruhi evaluasi model atau menyebabkan kebocoran data. Hasil dari tahap *pre-processing* menghasilkan satu kolom teks utama, yaitu *clean_text*, yang selanjutnya digunakan pada tahap pelabelan sentimen dan pelatihan model. Contoh hasil pembersihan teks ditampilkan pada Tabel 6.

Tabel 6. Tahap *Cleaning*

No	*Full_text*	*Cleaning*
1	*Timestamp*:\n00:00 iPhone 17 SERIES! + pajak ini kalau masuk Indonesia harganya pasti jadi mahal banget dibanding generasi sebelumnya.	*timestamp* iphone series pajak unboxing iphone masuk indonesia harga mahal
2	Aslinya iphone itu bisa banget lo bikin kamera langsung bagus tanpa edit tapi tetap keliatan natural.	asli iphone banget lo bikin kamera langsung bagus natural
3	Tolong sarannya 🙏 iphone sy 13 pro. Bagus ganti ke iphone 17 atau nunggu seri berikutnya ya?	saran iphone sy pro bagus ganti iphone pro nunggu seri berikut

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9507

Pada tahap pembersihan, komentar kosong dan komentar yang menjadi kosong setelah pembersihan dihapus. Selain itu, komentar duplikat juga dihapus sebelum *dataset* digunakan lebih lanjut untuk mengurangi potensi kemiripan antar data yang dapat memengaruhi evaluasi.

3.2 Pelabelan Sentimen dan Analisis Distribusi Data

3.2.1 Hasil Pelabelan Sentimen

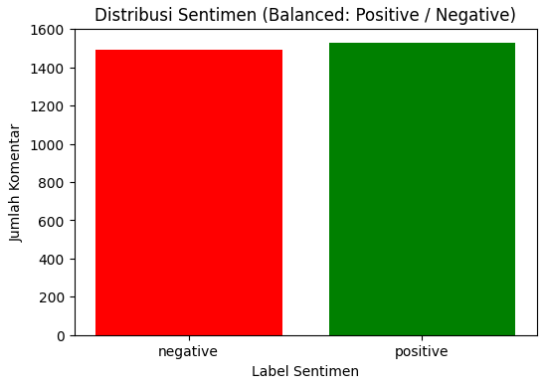
Setelah tahap *pre-processing* selesai, pelabelan sentimen dilakukan secara otomatis menggunakan *InSetLexicon* dengan dua kategori sentimen, yaitu positif dan negatif. Dari total 3.053 komentar yang diperoleh pada tahap awal, proses pembersihan menghasilkan 3.022 komentar valid yang kemudian digunakan pada tahap analisis. Hasil pelabelan menunjukkan bahwa terdapat 1.527 komentar berlabel positif dan 1.495 komentar berlabel negatif. Perbedaan jumlah antar kelas relatif kecil, sehingga distribusi data dapat dikatakan seimbang. Kondisi ini mengindikasikan bahwa *dataset* dapat digunakan langsung pada tahap pelatihan model tanpa memerlukan teknik penyeimbangan tambahan. Rincian distribusi label sentimen disajikan pada Tabel 7, sedangkan visualisasi proporsi kelas ditampilkan pada Gambar 3 untuk memperjelas sebaran sentimen dalam *dataset*.

Tabel 7. Distribusi Label Sentimen Setelah *Cleaning*

Label	Jumlah Data
Positive	1.527
Negative	1.495
Total	3.022

Gambar 3 menyajikan distribusi sentimen komentar *YouTube* terkait iPhone 17 yang menunjukkan keseimbangan relatif antara sentimen positif dan negatif.

Gambar 3. Visualisasi Distribusi Data Sentimen iPhone 17



3.2.2 Analisis Distribusi Data Sentimen

Berdasarkan distribusi label sentimen yang ditunjukkan pada Tabel 3 dan Gambar 3, dapat diamati bahwa jumlah komentar dengan sentimen positif dan negatif berada pada proporsi yang hampir seimbang. Kondisi ini menunjukkan bahwa opini publik terhadap iPhone 17 bersifat beragam dan tidak didominasi oleh satu kecenderungan sentimen tertentu. Keseimbangan distribusi kelas ini memberikan keuntungan dalam proses pelatihan model klasifikasi, karena dapat mengurangi potensi bias prediksi terhadap salah satu kelas. Selain itu, distribusi data yang relatif seimbang memungkinkan evaluasi kinerja model dilakukan secara lebih objektif tanpa memerlukan penerapan teknik penyeimbangan tambahan seperti *oversampling* atau *undersampling*. Dengan demikian, karakteristik alami data tetap terjaga dan hasil pengujian model dapat merepresentasikan kondisi data sebenarnya.

3.3 Pembagian Data (Train, Validation, Test)

Dataset berlabel selanjutnya dibagi menjadi data latih, data validasi, dan data uji menggunakan teknik *stratified split*. Pendekatan ini dipilih untuk memastikan proporsi kelas positif dan negatif tetap konsisten pada setiap subset data. Hasil pembagian menunjukkan bahwa data latih terdiri dari 2.115 komentar, data validasi sebanyak 453 komentar, dan data uji sebanyak 454 komentar. Pembagian *dataset* dilakukan satu kali menggunakan seed tetap agar hasil eksperimen dapat direplikasi. Selain itu, penggunaan subset data yang sama pada seluruh model bertujuan untuk memastikan bahwa perbandingan performa antar metode dilakukan secara objektif dan adil. Rincian pembagian data beserta distribusi label pada masing-masing subset disajikan pada Tabel 8.

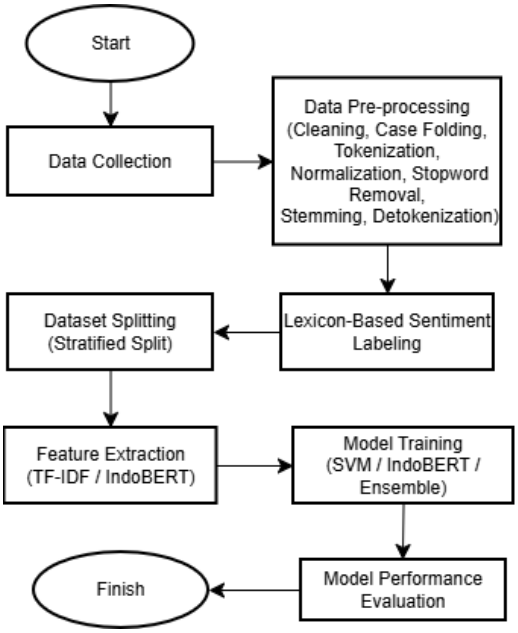
Tabel 8. Rekap Pembagian Data dan Distribusi Label

Subset	Total Data	Positif	Negatif
Train	2.115	1.069	1.046
Validation	453	229	224
Test	454	229	225
Total	**3.022**	**1.527**	**1.495**

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9507

3.4 Tahapan Penerapan Metode Klasifikasi Sentimen

Penerapan metode dilakukan melalui beberapa tahapan sistematis yang mengikuti rancangan penelitian. Proses dimulai dari pengumpulan data komentar, dilanjutkan dengan prapemrosesan teks, pelabelan sentimen, serta pembagian *dataset*. Pada tahap pemodelan, fitur teks diekstraksi menggunakan *Term Frequency–Inverse Document Frequency* untuk model *Support Vector Machine*, sedangkan model *IndoBERT* memanfaatkan representasi kontekstual melalui proses *fine-tuning*. Model dilatih menggunakan data latih dan divalidasi menggunakan data validasi. Evaluasi akhir dilakukan menggunakan *confusion matrix* dan metrik evaluasi.



Gambar 4. *Pipeline* Analisis Sentimen

3.5 Pengujian Skenario

Pengujian skenario dilakukan untuk mengevaluasi pengaruh representasi fitur terhadap performa model. Pengujian dilakukan dengan membandingkan fitur TF-IDF menggunakan konfigurasi *unigram* dan *unigram+bigram* pada model *Support Vector Machine*.

Tabel 9. Hasil Pengujian Skenario Fitur

Fitur TF-IDF	*Accuracy* (%)	*Precision* (%)	*Recall* (%)	*F1-score* (%)
Unigram	86,75	88,12	84,30	86,15
Unigram+Bigram	88,99	90,87	86,90	88,80

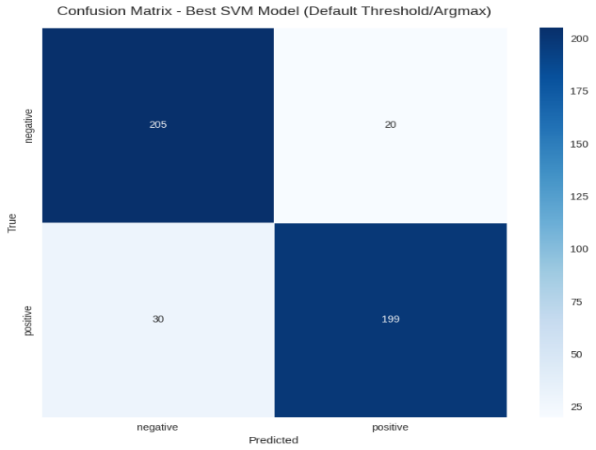
Hasil menunjukkan penggunaan *unigram+bigram* menghasilkan performa lebih baik karena mampu menangkap konteks frasa.

3.6 Hasil Klasifikasi dan Evaluasi Model

3.4.1 Hasil Model SVM

Model *Support Vector Machine* (SVM) dilatih menggunakan fitur *TF-IDF* dan dievaluasi pada data uji melalui *confusion matrix*. Hasil evaluasi menunjukkan bahwa model mampu mengklasifikasikan dengan benar 205 komentar negatif sebagai negatif (*True Negative*/TN) dan 199 komentar positif sebagai positif (*True Positive*/TP). Namun, masih terdapat kesalahan prediksi, yaitu 20 komentar negatif yang diprediksi sebagai positif (*False Positive*/FP) serta 30 komentar positif yang diprediksi sebagai negatif (*False Negative*/FN). Berdasarkan hasil tersebut, performa model dihitung menggunakan metrik evaluasi utama, yaitu *Accuracy*, *Precision*, *Recall*, dan *F1-score*, dengan nilai *Accuracy* sebesar 88,99%, *Precision* 90,87%, *Recall* 86,90%, serta *F1-score* 88,80%. Temuan ini menunjukkan bahwa SVM cukup efektif dalam mengklasifikasikan sentimen, meskipun masih rentan mengalami kesalahan pada komentar yang bersifat ambigu atau memiliki makna tidak tersurat. *Confusion matrix* digunakan untuk membandingkan label aktual dan prediksi pada data uji, dan visualisasinya disajikan pada Gambar 5 berikut.

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9507



Gambar 5. Visualisasi Hasil *Confusion matrix* Model SVM pada Data iPhone 17

Untuk memperjelas distribusi hasil klasifikasi, ringkasan *confusion matrix* disajikan pada Tabel 5 berikut.

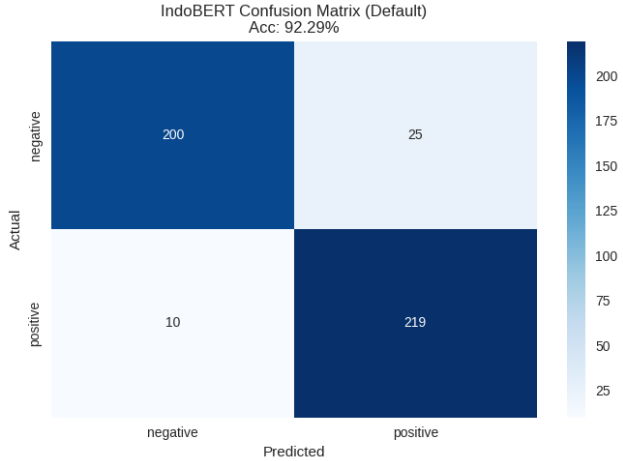
Tabel 10. Perhitungan *Confusion matrix* Model SVM

Aktual / Prediksi	Positif	Negatif	Total
Positif	TP = 199	*FN* = 30	229
Negatif	FP = 20	TN = 205	225
Total	219	235	454

Setelah nilai pada *confusion matrix* diperoleh, kinerja model SVM dievaluasi menggunakan metrik *Accuracy*, *Precision*, *Recall*, dan *F1-score*. Metrik tersebut dipilih karena mampu menggambarkan performa model secara menyeluruh, baik dari sisi ketepatan prediksi maupun kemampuan identifikasi kelas positif.

3.4.2 Hasil Model *IndoBERT*

Model *IndoBERT* dievaluasi melalui skema *fine-tuning* untuk tugas klasifikasi sentimen biner. Berdasarkan hasil evaluasi menggunakan *confusion matrix*, model ini mampu mengidentifikasi secara tepat 200 komentar negatif sebagai kelas negatif (*True Negative*/TN) dan 219 komentar positif sebagai kelas positif (*True Positive*/TP). Meskipun demikian, masih ditemukan sejumlah kesalahan prediksi, yaitu 25 komentar negatif yang salah diklasifikasikan sebagai positif (*False Positive*/FP) serta 10 komentar positif yang terklasifikasi sebagai negatif (*False Negative*/FN). Pengukuran kinerja menunjukkan bahwa *IndoBERT* memperoleh nilai *Accuracy* sebesar 92, 29%, *Precision* 89,75%, *Recall* 95,63%, dan *F1-score* 92,60%. Nilai *Recall* yang tinggi mengindikasikan kemampuan model yang sangat baik dalam mendeteksi sentimen positif, sehingga menjadikan *IndoBERT* sebagai pendekatan dengan performa paling unggul dibandingkan metode lain dalam penelitian ini. Visualisasi perbandingan antara label aktual dan hasil prediksi disajikan melalui *confusion matrix* pada Gambar 6.



Gambar 6. Visualisasi Hasil *Confusion matrix* Model *IndoBERT* pada Data iPhone 17

Untuk memperjelas distribusi hasil klasifikasi, ringkasan *confusion matrix* disajikan pada Tabel 6 berikut.

Tabel 11. Perhitungan *Confusion matrix* Model *IndoBERT*

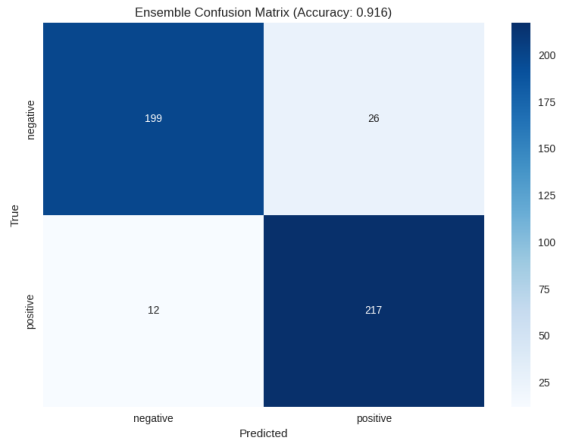
Aktual / Prediksi	Negatif	Positif	Total
Negatif	**TN = 200**	**FP = 25**	225
Positif	***FN* = 10**	**TP = 219**	229
Total	210	244	**454**

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9507

Berdasarkan perhitungan metrik evaluasi, model *IndoBERT* menunjukkan performa yang sangat baik pada klasifikasi sentimen biner. Nilai *Recall* yang tinggi menandakan model efektif dalam mendeteksi komentar positif, sehingga menjadi yang paling unggul dibandingkan model lainnya.

3.4.3 Hasil Model *Ensemble IndoBERT*–SVM

Selain model tunggal, penelitian ini turut menguji model *ensemble IndoBERT*–SVM dengan menerapkan mekanisme *soft voting*. Berdasarkan evaluasi menggunakan *confusion matrix*, model *ensemble* mampu mengklasifikasikan secara benar 217 komentar positif sebagai positif (*True Positive*/TP) dan 199 komentar negatif sebagai negatif (*True Negative*/TN). Adapun kesalahan prediksi masih ditemukan, yaitu 26 komentar negatif yang salah diprediksi sebagai positif (*False Positive*/FP) serta 12 komentar positif yang terklasifikasi sebagai negatif (*False Negative*/FN). Hasil pengukuran metrik menunjukkan nilai *Accuracy* 91,63%, *Precision* 89,30%, *Recall* 94,76%, dan *F1-score* 91,90%. Temuan ini mengindikasikan bahwa model *ensemble* memiliki performa yang relatif stabil dan mendekati capaian *IndoBERT*, meskipun belum mampu melampaui kinerja model *transformer* tersebut. Untuk memperlihatkan perbandingan antara label aktual dan hasil prediksi pada data uji, visualisasi *confusion matrix* disajikan pada Gambar 7.



Gambar 7. *Confusion matrix* Model *Ensemble IndoBERT* - SVM pada Data Uji

Untuk memperjelas distribusi hasil klasifikasi, ringkasan *confusion matrix* disajikan pada Tabel 7 berikut.

Tabel 12. Rekapitulasi *Confusion matrix* Model *Ensemble IndoBERT*

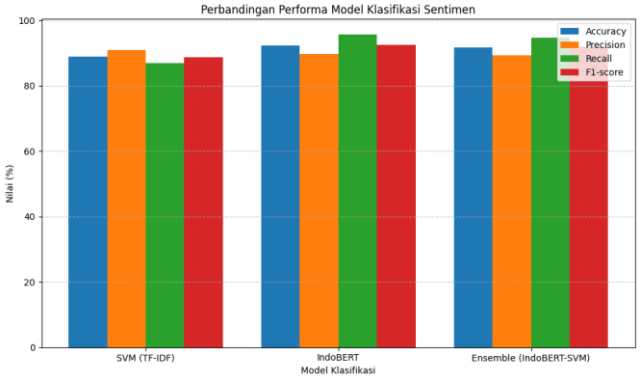
Aktual / Prediksi	Negatif	Positif	Total
Negatif	**TN = 199**	**FP = 26**	225
Positif	***FN* = 12**	**TP = 217**	229
Total	211	243	454

Setelah nilai *TP*, *TN*, *FP*, dan *FN* diperoleh dari *confusion matrix*, performa model *ensemble* kemudian dihitung melalui empat metrik evaluasi utama, yakni *Accuracy*, *Precision*, *Recall*, dan *F1-score*. Pemilihan metrik ini didasarkan pada kapabilitasnya dalam memberi gambaran komprehensif mengenai performa model, baik dari sisi ketepatan prediksi secara umum maupun kemampuan model dalam mendeteksi kelas tertentu.

3.7 Perbandingan Hasil Pengujian Model

Perbandingan hasil pengujian menunjukkan bahwa model *IndoBERT* memiliki performa paling unggul dalam klasifikasi sentimen komentar *YouTube* terkait iPhone 17. Keunggulan ini didukung oleh kemampuan model berbasis *transformer* dalam memahami konteks kalimat secara menyeluruh dan menangkap makna implisit pada teks berbahasa informal. Model *Support Vector Machine* berbasis *Term Frequency–Inverse Document Frequency* menunjukkan kinerja yang kompetitif, khususnya dari sisi ketepatan prediksi, namun kurang optimal dalam menangani komentar yang bersifat ambigu. Sementara itu, model *ensemble IndoBERT–SVM* menghasilkan performa yang relatif stabil dan seimbang, sehingga dapat dijadikan alternatif ketika diperlukan kompromi antara ketepatan dan kelengkapan deteksi sentimen. Secara keseluruhan, hasil pada tabel dan visualisasi grafik menegaskan efektivitas pendekatan berbasis *transformer* dalam analisis sentimen komentar *YouTube* berbahasa Indonesia.

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9507



Gambar 8. Perbandingan Performa Model Klasifikasi Sentimen

Gambar tersebut memberikan gambaran visual perbandingan performa model klasifikasi sentimen berdasarkan beberapa metrik evaluasi utama. Nilai kuantitatif lengkap dari masing-masing model selanjutnya disajikan pada tabel rekapitulasi hasil pengujian.

Tabel 13. Rekapitulasi Performa Model Klasifikasi Sentimen

Model	*Accuracy* (%)	*Precision* (%)	*Recall* (%)	*F1-score* (%)
SVM (TF-IDF)	88,99	90,87	86,90	88,80
IndoBERT	92, 29	89,75	95,63	92,60
Ensemble (*IndoBERT* - SVM)	91,63	89,30	94,76	91,90

3.8 Analisis Kesalahan Klasifikasi

Berdasarkan hasil evaluasi menggunakan *confusion matrix*, kesalahan klasifikasi umumnya muncul pada komentar yang memiliki struktur bahasa ambigu, penggunaan bahasa informal, serta ekspresi opini yang tidak disampaikan secara eksplisit. Komentar dengan makna implisit, ironi, atau sarkasme cenderung sulit diklasifikasikan secara tepat, terutama oleh model berbasis fitur statistik. Model *Support Vector Machine* yang menggunakan representasi *Term Frequency–Inverse Document Frequency* masih bergantung pada frekuensi kemunculan kata, sehingga kurang optimal dalam menangkap konteks kalimat secara menyeluruh. Sebaliknya, *IndoBERT* sebagai model berbasis *transformer* mampu memahami representasi kontekstual dua arah, sehingga dapat mengurangi sebagian kesalahan klasifikasi pada komentar yang memiliki makna kontekstual. Meskipun demikian, kesalahan prediksi masih ditemukan pada komentar dengan pola bahasa yang sangat tidak baku atau mengandung makna ganda.

3.9 Pengaruh Pelabelan Otomatis

Pelabelan sentimen pada penelitian ini dilakukan secara otomatis menggunakan *InSetLexicon* untuk mempercepat proses anotasi data dalam skala besar. Pendekatan ini efektif dalam mengurangi beban anotasi manual dan memungkinkan pembentukan *dataset* berlabel dalam waktu relatif singkat. Namun demikian, pelabelan berbasis leksikon memiliki keterbatasan dalam memahami konteks kalimat yang kompleks, seperti negasi, sarkasme, atau perubahan makna akibat susunan kata. Akibatnya, terdapat potensi kesalahan label pada sebagian komentar, yang dapat memengaruhi hasil pelatihan model. Kondisi ini menunjukkan bahwa kualitas label memiliki peran penting terhadap performa model klasifikasi, terutama pada data media sosial yang bersifat informal dan dinamis.

3.10 Perbandingan dengan Penelitian Terdahulu

Hasil penelitian ini sejalan dengan berbagai penelitian terdahulu yang melaporkan bahwa model berbasis *transformer* menunjukkan performa yang lebih baik dibandingkan metode *machine learning* tradisional dalam analisis sentimen teks berbahasa Indonesia. Studi sebelumnya menunjukkan bahwa pendekatan berbasis *BERT* mampu menangkap konteks semantik dengan lebih efektif dibandingkan metode berbasis fitur statistik. Temuan pada penelitian ini memperkuat hasil tersebut, khususnya pada komentar YouTube yang memiliki karakteristik bahasa tidak baku. Sementara itu, model *Support Vector Machine* tetap menunjukkan performa kompetitif, terutama dari sisi ketepatan prediksi, namun masih memiliki keterbatasan dalam memahami makna implisit.

3.11 Implikasi Praktis

Secara praktis, hasil penelitian ini dapat dimanfaatkan oleh perusahaan atau pelaku industri teknologi untuk memahami persepsi konsumen terhadap produk sebelum peluncuran resmi. Analisis sentimen komentar YouTube dapat memberikan gambaran awal mengenai respons pasar, potensi penerimaan produk, serta aspek yang menjadi perhatian konsumen. Selain itu, penerapan model berbasis *transformer* dapat mendukung pengambilan keputusan strategis berbasis data, khususnya dalam penyusunan strategi pemasaran dan komunikasi produk di media sosial.

3.12 Keterbatasan Penelitian

Penelitian ini memiliki beberapa keterbatasan yang perlu diperhatikan. Pertama, pelabelan sentimen dilakukan secara otomatis menggunakan *InSetLexicon*, sehingga masih berpotensi menghasilkan label yang kurang akurat pada komentar tertentu. Kedua, sumber data terbatas pada satu kanal YouTube, sehingga hasil penelitian

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9507

belum sepenuhnya merepresentasikan opini publik secara luas. Ketiga, klasifikasi sentimen hanya menggunakan dua kelas, yaitu positif dan negatif, sehingga variasi sentimen lain belum terakomodasi.

3.13 Saran Penelitian Lanjutan

Berdasarkan keterbatasan tersebut, penelitian selanjutnya disarankan untuk menggunakan anotasi manual atau pendekatan *semi-supervised* guna meningkatkan kualitas label data. Selain itu, perluasan sumber dan jumlah *dataset* dari berbagai kanal atau platform media sosial dapat meningkatkan generalisasi model. Penelitian lanjutan juga dapat mengembangkan klasifikasi *multiclass* untuk menangkap variasi sentimen yang lebih beragam serta mengeksplorasi penerapan model *ensemble* yang lebih kompleks.

4. KESIMPULAN

Berdasarkan hasil akhir penelitian, dapat disimpulkan bahwa model *IndoBERT* menunjukkan kinerja paling optimal dalam mengklasifikasikan sentimen komentar *YouTube* berbahasa Indonesia terkait iPhone 17. Temuan ini menegaskan bahwa pendekatan berbasis *transformer* lebih efektif dalam memahami konteks kalimat, variasi bahasa informal, serta makna implisit yang umum muncul pada komentar *media sosial*. Model *Support Vector Machine* berbasis *Term Frequency–Inverse Document Frequency* tetap memperlihatkan performa yang kompetitif, khususnya dalam ketepatan prediksi, namun memiliki keterbatasan ketika menghadapi komentar ambigu dan kontekstual. Sementara itu, model *ensemble IndoBERT–SVM* memberikan performa yang relatif stabil dan seimbang, sehingga dapat dipertimbangkan sebagai alternatif ketika diperlukan kompromi antara ketepatan dan kelengkapan klasifikasi sentimen. Secara keseluruhan, hasil penelitian ini menunjukkan bahwa pemilihan model bahasa yang tepat berpengaruh signifikan terhadap kualitas analisis sentimen pada data *YouTube* yang tidak terstruktur. Kesimpulan ini sepenuhnya didasarkan pada hasil akhir pengujian model menggunakan metrik evaluasi yang konsisten dan objektif. Dengan demikian, penelitian ini memberikan kontribusi empiris dalam pengembangan analisis sentimen berbahasa Indonesia serta dapat menjadi rujukan bagi penelitian lanjutan dan penerapan praktis di bidang analisis opini digital. Temuan tersebut juga memperkuat relevansi penggunaan model modern dalam mendukung pengambilan keputusan berbasis data, evaluasi persepsi publik, dan pengembangan strategi komunikasi digital yang lebih adaptif berkelanjutan pada konteks *media sosial*.

REFERENCES

[1] R. Obiedat *et al.*, “Sentiment Analysis of Customers’ Reviews Using a Hybrid Evolutionary SVM-Based Approach in an Imbalanced Data Distribution,” *IEEE Access*, vol. 10, pp. 22260–22273, 2022, doi: 10.1109/ACCESS.2022.3149482.

[2] T. Walasary, “Survey Paper tentang Analisis Sentimen,” *KONSTELASI Konvergensi Teknol. dan Sist. Inf.*, vol. 2, no. 1, pp. 201–206, 2022, doi: 10.24002/konstelasi.v2i1.5378.

[3] Keerthana Prabhu B and Dr. Vinay K S, “Voice of the Customer: a Comparative Study of Sentiment Across Countries for Samsung Galaxy Buds Pro,” *EPRA Int. J. Econ. Bus. Manag. Stud.*, no. June, pp. 139–145, 2025, doi: 10.36713/epra22664.

[4] S. S. Almalki, “Sentiment Analysis and Emotion Detection Using Transformer Models in Multilingual Social Media Data,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 16, no. 3, pp. 324–333, 2025, doi: 10.14569/IJACSA.2025.0160332.

[5] I. O. Wibowo *et al.*, “Analisis Sentimen Komentar Youtube Pidato Kemenangan Prabowo Subianto Menggunakan Naïve Bayes Optimasi Particle Swarm Optimization,” vol. 12, no. 2, pp. 3344–3349, 2025, [Online]. Available: https://www.youtube.com/watch?v=5DbCvqfg-9I.

[6] F. M. Hidayat and H. Sanjaya, “Analisis Sentimen Publik Terhadap Penjualan Iphone 16 Dan Kebijakan Tkdn Di Indonesia,” *INFOTECH J.*, vol. 11, no. 1, pp. 74–80, 2025, doi: 10.31949/infotech.v11i1.13159.

[7] N. Fitriyah, B. Warsito, D. Asih, and I. Maruddani, “Sentiment Analysis of GOJEK on Twitter Social Media Using Support Vector Machine (SVM) Classification,” *J. Gaussian*, vol. 9, no. 3, pp. 376–390, 2020, [Online]. Available: https://ejournal3.undip.ac.id/index.php/gaussian/

[8] H. Jayadianti, W. Kaswidjanti, A. T. Utomo, S. Saifullah, F. A. Dwiyanto, and R. Drezewski, “Sentiment analysis of Indonesian reviews using fine-tuning IndoBERT and R-CNN,” *Ilk. J. Ilm.*, vol. 14, no. 3, pp. 348–354, 2022, doi: 10.33096/ilkom.v14i3.1505.348-354.

[9] M. M. Rahman, A. I. Shiplu, Y. Watanobe, and M. A. Alam, “RoBERTa-BiLSTM: A Context-Aware Hybrid Model for Sentiment Analysis,” *IEEE Trans. Emerg. Top. Comput. Intell.*, vol. 9, no. 6, pp. 3788–3805, 2025, doi: 10.1109/TETCI.2025.3572150.

[10] S. Astuti, “Analisis Sentimen Pandangan Publik Terhadap Kenaikan Pajak 12% Dari Twitter Menggunakan Indonesian Roberta Base Classifier,” *Smart Comp Jurnalnya Orang Pint. Komput.*, vol. 14, no. 3, pp. 698–706, 2025, doi: 10.30591/smartcomp.v14i3.8354.

[11] R. Ardiyansyah, I. Makmur, and S. P. Phai, “Sentimen Komentar Youtube Dengan Sentiment Intensity Analyzer Dari NLTK,” *Semin. Nas. Corisindo*, pp. 31–36, 2024.

[12] Risca Lusiana Pratiwi, Zulia Imami Alfianti, Ahmad Fauzi, and Ginabila Ginabila, “Penerapan Algoritma Naive Bayes dan SVM untuk Analisis Sentimen terhadap Penggunaan True Wireless Stereo (TWS),” *SKANIKA Sist. Komput. dan Tek. Inform.*, vol. 8, no. 2, pp. 257–268, 2025, doi: 10.36080/skanika.v8i2.3535.

[13] Dhea Ferdiana Merpatika and Albert Yakobus Chandra, “Analisis Sentimen Publik Di Media Sosial Terhadap Kasus Dugaan Korupsi Impor Minyak Pertamina Menggunakan Xgboost,” *Rabit J. Teknol. dan Sist. Inf. Univrab*, vol. 10, no. 2, pp. 576–593, 2025, doi: 10.36341/rabit.v10i2.6307.

[14] 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.

[15] V. D. Setiawan, D. U. Iswavigra, and E. Anggiratih, "Implementation of IndoBERT for Sentiment Analysis of the Constitutional Court's Decision Regarding the Minimum Age of Vice Presidential Candidates," *Sci. J. Informatics*, vol. 12, no. 3, pp. 397–406, 2025, doi: 10.15294/sji.v12i3.26320.

[16] K. S. Nugroho, A. Y. Sukmadewa, H. Wuswilahaken Dw, F. A. Bachtiar, and N. Yudistira, "BERT Fine-Tuning for Sentiment Analysis on Indonesian Mobile Apps Reviews," *ACM Int. Conf. Proceeding Ser.*, pp. 258–264, 2021, doi: 10.1145/3479645.3479679.

[17] M. Kumar, L. Khan, and H. T. Chang, "Evolving techniques in sentiment analysis: a comprehensive review," *PeerJ Comput. Sci.*, vol. 11, pp. 1–61, 2025, doi: 10.7717/PEERJ-CS.2592.

[18] st Yuyun Khanafiyah and D. Kartika Sari, "Analisis Sentimen Komentar Youtube Kanal Dirty Vote Menggunakan Metode Naive Bayes Classifier," *Agustus*, vol. 12, no. 4, p. 6723, 2025.

[19] N. Amalia Putri, A. Srirahayu, and N. Arif Sudibyo, "Sentiment Analysis Towards the KitaLulus Application Using the Naive Bayes Method from Google Play Store Reviews," *J. Indones. Sos. Teknol.*, vol. 5, no. 10, pp. 4593–4603, 2024, doi: 10.59141/jist.v5i10.1244.

[20] F. F. Mailoa, "Analisis sentimen data twitter menggunakan metode text mining tentang masalah obesitas di indonesia," *J. Inf. Syst. Public Heal.*, vol. 6, no. 1, p. 44, 2021, doi: 10.22146/jisph.44455.

[21] A. D. Hananto, A. M. Erfiana, B. L. P. Putri, P. D. Putri, and F. Kurniawan, "Algoritma Machine Learning Naïve Bayes pada Analisis Sentimen Kesepakatan Polri dan GNPF-MUI pada Aksi Bela Islam III '212,'" *SINTA J. (Science, Technol. Agric.*, vol. 4, no. 2, pp. 151–160, 2023, doi: 10.37638/sinta.4.2.151-160.

[22] S. K. Al Sarayrah, N. M. Alkudah, and H. Y. Al-kharabsheh, "Twitter Sentiment Analysis Using Machine Learning and Deep Learning Techniques," *J. Comput. Sci.*, vol. 21, no. 8, pp. 1785–1794, 2025, doi: 10.3844/jcssp.2025.1785.1794.

[23] Wildan Amru Hidayat and V. R. S. Nastiti, "Perbandingan Kinerja Pre-Trained Indobert-Base Dan Indobert-Lite Pada Klasifikasi Sentimen Ulasan Tiktok Tokopedia Seller Center Dengan Model Indobert," *JSiI (Jurnal Sist. Informasi)*, vol. 11, no. 2, pp. 13–20, 2024, doi: 10.30656/jsii.v11i2.9168.

[24] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP," *COLING 2020 - 28th Int. Conf. Comput. Linguist. Proc. Conf.*, pp. 757–770, 2020, doi: 10.18653/v1/2020.coling-main.66.

[25] M. Ahmad, S. Aftab, M. S. Bashir, and N. Hameed, "Sentiment analysis using SVM: A systematic literature review," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 2, pp. 182–188, 2018, doi: 10.14569/IJACSA.2018.090226.

[26] M. A. Bert, "Analisis Sentimen Komentar YouTube pada Video Clash of Champions di Channel Ruangguru Analisis Sentimen Komentar YouTube pada Video Clash of Champions di," 2024.

[27] A. N. A. Saputra, R. E. Saputro, and D. I. S. Saputra, "Enhancing Sentiment Analysis Accuracy Using SVM and Slang Word Normalization on YouTube Comments," *Sinkron*, vol. 9, no. 2, pp. 687–699, 2025, doi: 10.33395/sinkron.v9i2.14613.

[28] M. Dhingra, P. Saini, and M. Scholar, "Sentiment Analysis on US Elections Using Deep Learning and Transformer Models," *Int. J. Sci. Dev. Res.*, vol. 10, no. 7, pp. 143–153, 2025, [Online]. Available: www.ijsdr.org

[29] Gishella Septania Al-Husna, Dian Asmarajati, Iman Ahmad Ihsannuddin, and Rina Mahmudati, "Perbandingan Metode Naïve Bayes Dan Support Vector Machine Untuk Analisis Sentimen Pada Ulasan Pengguna Aplikasi Linkedin," *STORAGE J. Ilm. Tek. dan Ilmu Komput.*, vol. 3, no. 2, pp. 139–144, 2024, doi: 10.55123/storage.v3i2.3602.

[30] S. George and V. Srividhya, "Performance Evaluation of Sentiment Analysis on Balanced and Imbalanced Dataset Using Ensemble Approach," *Indian J. Sci. Technol.*, vol. 15, no. 17, pp. 790–797, 2022, doi: 10.17485/ijst/v15i17.2339.

[31] M. Iqbal, M. Afdal, and R. Novita, "Implementasi Algoritma Support Vector Machine Untuk Analisa Sentimen Data Ulasan Aplikasi Pinjaman Online di Google Play Store," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 4, pp. 1244–1252, 2024, doi: 10.57152/malcom.v4i4.1435.