

Analisis Sentimen Komentar Tiktok Pasca Gempa Mindanao Dengan Pemodelan SVM-TFIDF dan IndoBERT-BiLSTM

Reza Okta Pratama, Arfany Dhimas Muftareza*, Nardi, Sugiarto

Program Studi Instrumentasi-MKG, Sekolah Tinggi Meteorologi Klimatologi dan Geofisika
Jalan Meteorologi Nomor 5, Kelurahan Tanah Tinggi, Kecamatan Tangerang, Kota Tangerang, Banten, Indonesia 15119
E-mail: arfanydhimuf@gmail.com

Naskah Masuk: 25 Juli 2026; Diterima: 25 September 2026 ; Terbit: 30 September 2026

ABSTRAK

Abstrak - Penelitian ini menganalisis sentimen komentar TikTok berbahasa Indonesia pascagempa Mw 7,8 di Sarangani, Filipina, pada 8 Juni 2026 menggunakan dua pendekatan, yaitu *Support Vector Machine* (SVM) berbasis *Term Frequency-Inverse Document Frequency* (TF-IDF) dan *IndoBERT-Bidirectional Long Short-Term Memory* (BiLSTM). Data diperoleh melalui *scraping* komentar dari enam video TikTok menggunakan *Apify TikTok Comment Scraper* dan menghasilkan 2.294 komentar, dengan 2.023 komentar valid setelah *preprocessing*. Pelabelan tiga kelas sentimen (negatif, netral, dan positif) dilakukan menggunakan *Indonesian RoBERTa-base* sebagai pseudo-label, sedangkan data uji divalidasi secara manual oleh dua anotator. SVM menggunakan *preprocessing* lengkap, TF-IDF dengan maksimum 5.000 fitur dan *n-gram* 1-3, serta kernel RBF. IndoBERT-BiLSTM menggunakan *preprocessing* ringan, *tokenizer* dengan panjang maksimum 128 token, dan *embedding* IndoBERT yang dibekukan. Dengan pembagian *stratified* 80:20, SVM memperoleh akurasi 67,16% dan *macro F1-score* 0,58, sedangkan IndoBERT-BiLSTM memperoleh akurasi 66,42% dan *macro F1-score* 0,64. IndoBERT-BiLSTM lebih baik pada kelas positif, tetapi membutuhkan waktu komputasi jauh lebih besar. Hasil menunjukkan perbedaan kinerja bergantung pada metrik evaluasi dan karakteristik kelas sentimen.

Kata kunci: Analisis Sentimen, Gempa Mindanao, Tiktok, SVM-TFIDF, IndoBERT-BiLSTM

ABSTRACT

Abstract - This study analyzes the sentiment of Indonesian-language TikTok comments following the Mw 7.8 earthquake in Sarangani, Philippines, on June 8, 2026, using two approaches: *Support Vector Machine* (SVM) based on *Term Frequency-Inverse Document Frequency* (TF-IDF) and *IndoBERT-Bidirectional Long Short-Term Memory* (BiLSTM). Data were obtained by scraping comments from six TikTok videos using the *Apify TikTok Comment Scraper*, yielding 2,294 comments, with 2,023 valid comments remaining after preprocessing. Three-class sentiment labeling (negative, neutral, and positive) was performed using *Indonesian RoBERTa-base* for pseudo-labeling, while the test data were manually validated by two annotators. The SVM model employed comprehensive preprocessing, TF-IDF with a maximum of 5,000 features and 1–3 *n-grams*, and an RBF kernel. The IndoBERT-BiLSTM model utilized light preprocessing, a tokenizer with a maximum length of 128 tokens, and frozen IndoBERT embeddings. Using an 80:20 stratified split, the SVM achieved an accuracy of 67.16% and a *macro F1-score* of 0.58, whereas the IndoBERT-BiLSTM achieved an accuracy of 66.42% and a *macro F1-score* of 0.64. IndoBERT-BiLSTM performed better on the positive class but required significantly more computational time. The results indicate that performance differences depend on the evaluation metrics and the characteristics of the sentiment classes.

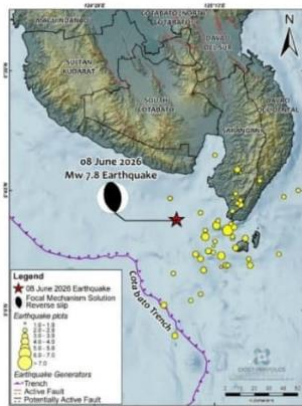
Keywords: Sentiment Analysis, Mindanao Earthquake, Tiktok, SVM-TFIDF, IndoBERT-BiLSTM

Copyright © 2026 Jurnal Teknik Elektro dan Komputasi (ELKOM)

1. PENDAHULUAN

Pada 8 Juni 2026 pukul 07.37 Waktu Standar Filipina (PST), gempa bumi berkekuatan magnitudo (Mw) 7,8 mengguncang wilayah Sarangani, Filipina Selatan. Episentrum gempa terletak di laut pada koordinat 1,25° LU; 126,27° BT, sekitar 32 km barat laut Maasim, Sarangani, pada kedalaman 33 km. Gambar 1 merupakan plot gempa bumi lepas pantai Sarangani berkekuatan 7,8 magnitudo pada tanggal 8 Juni 2026 dan gempa susulan hingga pukul 12:00 siang yang diterbitkan oleh *Philippine Institute of Volcanology and Seismology* [1]. Guncangan terasa sangat kuat, dengan intensitas tertinggi PEIS VII (*Destruktif*) di Kota

General Santos yang berpenduduk lebih dari 720.000 jiwa. Dalam waktu empat jam pascagempa, PHIVOLCS mencatat tidak kurang dari 138 gempa susulan dengan magnitudo tertinggi mencapai M6,7. USGS melaporkan mekanisme sumber berupa *thrust faulting* akibat subduksi Lempeng Filipina ke bawah Lempeng Sunda di zona deformasi Palung Cotabato [2].



Gambar 1. Plot gempa bumi mindanao, filipina [1]

BMKG mencatat magnitudo setara 7,7 dalam sistem pemantauan nasional Indonesia [3]. Selain itu, gempa ini membangkitkan tsunami lokal dengan gelombang leih dari 1 meter yang terpantau di Kiamba, Sarangani, sementara gelombang lebih kecil turut terdeteksi hingga wilayah Indonesia, Palau, dan Jepang. Berdasarkan dokumen *Situational Report No. 12*, gempa bumi magnitudo 7,8 di Maasim, Sarangani pada 8 Juni 2026 menimbulkan dampak besar berupa korban jiwa, ratusan luka-luka, puluhan ribu pengungsi, serta kerusakan luas pada infrastruktur dan permukiman [4].

Perkembangan teknologi informasi dan komunikasi telah mengubah cara masyarakat merespons bencana, dengan media sosial menjadi saluran utama untuk berbagi informasi, mengekspresikan emosi, dan membentuk opini publik secara cepat [5]. Dalam konteks ini, TikTok memainkan peran penting sebagai platform komunikasi krisis. Studi tentang kebakaran Maui 2023 menunjukkan efektivitasnya dalam menyebarkan informasi *real-time*, memobilisasi bantuan, dan melawan disinformasi [6]. TikTok memungkinkan publik mengalami bencana secara *vicarious* melalui digital *storytelling*, sekaligus memfasilitasi koordinasi selama fase krisis [7]. Pada kasus gempa bumi dan tsunami, analisis konten media sosial mengungkap pola keterlibatan publik yang mencerminkan kepedulian, kecemasan, dan solidaritas masyarakat. Penelitian terhadap lebih dari 150.000 unggahan terkait gempa Türkiye–Suriah 2023 bahkan menunjukkan adanya disparitas antara unggahan informatif dan komentar emosional yang berfokus pada ekspresi duka serta doa [5].

Analisis sentimen sebagai cabang NLP bertujuan mengekstrak opini, sikap, dan polaritas emosi dari teks [8], dan dalam konteks bencana terbukti memberikan *insight* cepat, layak, serta hemat biaya mengenai perilaku masyarakat [9]. Dua pendekatan utama adalah SVM dengan TF-IDF, yang menunjukkan akurasi tinggi hingga 93,65% dan konsisten mengungguli algoritma klasik lainnya [10], serta tetap kompetitif dengan akurasi 82,14% dibandingkan model *Transformer* [11], dan pendekatan berbasis IndoBERT yang menawarkan representasi kontekstual lebih kaya; arsitektur hibrida IndoBERT–BiLSTM mencapai F1-score 0,92 [12], sementara IndoBERT yang dilatih pada korpus Indo4B terbukti efektif untuk berbagai tugas NLP [13]. Implementasi pada data media sosial juga menjanjikan, seperti kombinasi BiLSTM–IndoBERT dan IndoBERTweet–LSTM yang berhasil mengklasifikasikan ulasan TikTok dengan hasil kompetitif [14][15], serta penggunaan IndoBERT large-p1 dalam analisis sentimen tweet Pemilu 2024 yang meraih akurasi 83,5%, melampaui SVM sebesar 3,5% [16].

Meskipun analisis sentimen media sosial dalam konteks bencana telah banyak diteliti, kajian khusus mengenai respons masyarakat Indonesia di TikTok terhadap gempa bumi di negara Asia Tenggara masih terbatas. Penerapan metode *machine learning* komprehensif untuk memahami keterlibatan publik lintas negara pascagempa [17]. Media sosial menawarkan perspektif unik yang mencerminkan ekspresi spontan publik dibandingkan saluran tradisional [5]. Penelitian ini hadir untuk mengisi celah tersebut dengan menganalisis sentimen komentar TikTok berbahasa Indonesia terkait gempa M7,7 Mindanao, Filipina, menggunakan pendekatan SVM + TF-IDF dan IndoBERT + BiLSTM. Kontribusinya meliputi penyediaan dataset komentar lintas negara, perbandingan empiris antara metode klasik dan Transformer, serta pemahaman mendalam mengenai pola sentimen dan karakteristik linguistik respons publik Indonesia terhadap bencana internasional, termasuk potensi ancaman tsunami yang turut dirasakan masyarakat pesisir Indonesia.

2. KAJIAN PUSTAKA

Gempa bumi berkekuatan magnitudo 7,8 yang terjadi pada 8 Juni 2026 di lepas pantai Sarangani, Filipina Selatan, dengan episentrum pada kedalaman 33 km, menimbulkan guncangan destruktif hingga intensitas PEIS VII di General Santos dan memicu lebih dari 138 gempa susulan dengan magnitudo tertinggi M6,7 dalam empat jam pertama. Mekanisme sumber berupa *thrust faulting* akibat subduksi Lempeng Filipina ke bawah Lempeng Sunda di zona Palung Cotabato, sementara BMKG mencatat magnitudo setara 7,7. Peristiwa ini juga membangkitkan tsunami lokal dengan ketinggian gelombang lebih dari 1 meter di Kiamba serta terdeteksi hingga Indonesia, Palau, dan Jepang, serta menurut *Situational Report No. 12* berdampak besar berupa korban jiwa, ratusan luka-luka, puluhan ribu pengungsi, serta kerusakan luas pada infrastruktur [1][2][3][4].

2.1. Analisis Sentimen

Analisis sentimen merupakan cabang dari pemrosesan bahasa alami (NLP) yang berfokus pada pengidentifikasian dan pengklasifikasian opini, sikap, serta polaritas emosi yang terkandung dalam teks. Tujuannya adalah untuk mengekstrak informasi subjektif dari data linguistik sehingga dapat menggambarkan persepsi, reaksi, dan respons masyarakat terhadap suatu peristiwa atau isu. Dalam konteks kebencanaan, analisis sentimen memiliki nilai strategis karena mampu menyediakan gambaran cepat mengenai perilaku publik, tingkat kepanikan, maupun dukungan sosial yang muncul di media sosial [5][6][7][8][9].

2.2. SVM + TF-IDF

Teks direpresentasikan sebagai vektor numerik menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). TF-IDF mengukur tingkat kepentingan suatu kata dalam sebuah dokumen relatif terhadap seluruh korpus. Kata yang sering muncul dalam satu dokumen namun jarang di dokumen lain memperoleh bobot tinggi [18]. Bobot TF-IDF dihitung menggunakan persamaan 1:

$$TF-IDF(t, d) = \frac{f(t, d)}{\sum_d t_d} \times \log \frac{N}{df(t)} \quad (1)$$

TF (*Term Frequency*) digunakan untuk menghitung seberapa sering suatu kata t muncul dalam dokumen d , sedangkan IDF (*Inverse Document Frequency*) digunakan untuk mengukur seberapa penting suatu kata dengan melihat jumlah dokumen yang mengandung kata tersebut [18].

Support Vector Machine (SVM) adalah algoritma pembelajaran terawasi yang bekerja dengan mencari *hyperplane* optimal berdimensi tinggi yang memisahkan kelas-kelas data dengan margin maksimum [19]. Pada kasus data yang tidak terpisah secara linear, SVM menggunakan fungsi kernel untuk memetakan data ke ruang dimensi lebih tinggi sehingga pemisahan linear menjadi dimungkinkan [20]. Fungsi keputusan SVM diformulasikan pada persamaan 2:

$$f(x) = \text{sign}(\sum \alpha_i y_i K(x_i, x) + b) \quad (2)$$

Di mana α_i adalah koefisien Lagrange, y_i adalah label kelas, b adalah bias, dan $K(x_i, x)$ adalah fungsi kernel. Pada penelitian ini digunakan kernel *Radial Basis Function* (RBF) diformulasikan pada persamaan 3:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (3)$$

Di mana γ adalah parameter skala yang menentukan jangkauan pengaruh setiap titik data. Kernel RBF dipilih karena kemampuannya menangani data tidak linear secara efektif dan merupakan konfigurasi *default* yang terbukti kompetitif pada berbagai tugas klasifikasi teks [21].

2.3. IndoBERT + BiLSTM

IndoBERT adalah model bahasa berbasis arsitektur BERT (*Bidirectional Encoder Representations from Transformers*) yang dilatih khusus pada korpus Indo4B, korpus teks berbahasa Indonesia berukuran besar yang terdiri dari lebih dari 4 miliar kata dari berbagai domain [13]. Arsitektur IndoBERT terdiri dari lapisan *Transformer encoder* yang masing-masing menerapkan mekanisme *Multi-Head Self-Attention* (MHSA). Mekanisme *self-attention* dihitung pada persamaan 4:

$$\text{Attention}(Q, K, V) = \text{softmax}(QK^T / \sqrt{d_k}) V \quad (4)$$

Di mana Q (*Query*), K (*Key*), dan V (*Value*) adalah proyeksi linear dari representasi input, dan d_k adalah dimensi vektor kunci yang berfungsi sebagai faktor normalisasi untuk menstabilkan gradien. Output dari lapisan terakhir IndoBERT berupa *hidden state* berdimensi 768 untuk setiap token, yang selanjutnya digunakan sebagai input lapisan BiLSTM. *Bidirectional Long Short-Term Memory* (BiLSTM) merupakan pengembangan dari arsitektur LSTM standar yang memproses sekuens input secara dua arah, maju (*forward*) dan mundur (*backward*), sehingga setiap *hidden state* mengandung informasi konteks dari kedua sisi sekuens [22]. Setiap sel LSTM mengatur aliran informasi melalui tiga mekanisme *gating* yang diformulasikan sebagai berikut [23]:

Forget gate, menentukan informasi yang dibuang dari *cell state* diformulasikan pada persamaan 5:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (5)$$

Input gate, menentukan informasi baru yang disimpan ke *cell state* diformulasikan pada persamaan 6:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i); c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (6)$$

Output gate, menentukan output berdasarkan *cell state* yang diperbarui diformulasikan pada persamaan 7:

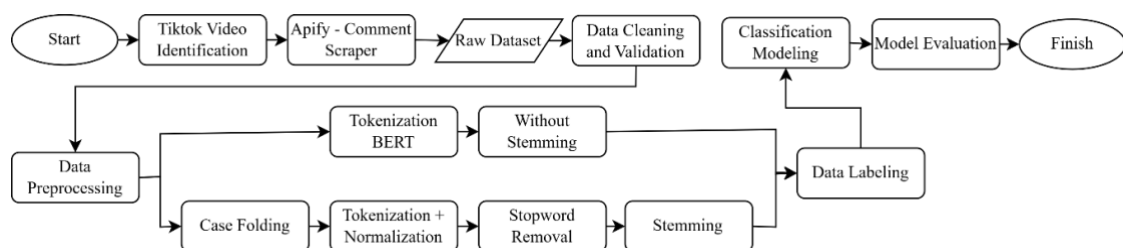
$$h_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \odot \tanh(c_t) \quad (7)$$

Di mana σ adalah fungsi sigmoid, $\odot$ adalah operasi perkalian elemen per elemen (*Hadamard product*), x_t adalah input pada waktu t , h_t adalah *hidden state*, dan c_t adalah *cell state*. Representasi gabungan ini selanjutnya dilewatkan ke lapisan *fully connected* dengan aktivasi *softmax* untuk menghasilkan distribusi probabilitas atas tiga kelas sentimen yang diformulasikan pada persamaan 8:

$$\hat{y} = \text{softmax}(W \cdot h_t + b) \quad (8)$$

3. METODE PENELITIAN

Metode penelitian yang digunakan mengikuti alur sistematis yang disajikan dalam gambar 2 yang mencakup pengumpulan data, preprocessing data, pelabelan data, pemodelan klasifikasi, dan evaluasi model.



Gambar 2. Diagram alir penelitian

Penelitian dimulai dari identifikasi video TikTok sebagai sumber data, dilanjutkan dengan proses scraping komentar untuk memperoleh *raw dataset*. Data mentah kemudian melalui tahap pembersihan dan validasi, termasuk preprocessing sesuai kebutuhan. Setelah itu dilakukan pelabelan data untuk menentukan kategori sentimen. Tahap berikutnya adalah pemodelan klasifikasi menggunakan pendekatan SVM-TF-IDF dan IndoBERT-BiLSTM, yang kemudian dievaluasi melalui metrik performa untuk menilai akurasi dan efektivitas. Seluruh rangkaian ini membentuk kerangka kerja analisis sentimen yang terstruktur dari awal hingga akhir penelitian.

Pengumpulan data dilakukan melalui teknik *scraping* terhadap komentar pengguna pada enam video TikTok yang berkaitan dengan peristiwa gempa bumi M7.7 Mindanao, Filipina, menggunakan platform Apify dengan aktor TikTok *Comment Scraper*. Video dipilih berdasarkan relevansi konten dengan topik kejadian gempa dan tsunami serta tingkat keterlibatan pengguna yang signifikan, mencakup video dari akun media berita seperti detikcom, okezone, Antara News, mistar.id, dan Bulelengterkini. Informasi keenam video tercatat ketika proses scraping komentar dilakukan pada 21 Juni 2026 pukul 21.15 WIB dan disajikan pada Tabel 1.

Tabel 1. Identifikasi video tiktok

No	Akun	Topik	Like	Komen	Simpan	Share
1	detikcom	Gempa Mindanao M 7,7	63.000	480	1.776	3.434
2	officialokezone	Sirene peringatan dini tsunami	20.200	430	1.127	1.212
3	Antara News	Gelombang tsunami hantam Gorontalo Utara	14.300	140	625	1.325
4	mistar.id	BMKG Deteksi Gelombang Tsunami di Sulut	41.500	732	3.558	17.400
5	Bulelengterkini	Rekaman video detik-detik tsunami	140.300	346	3.408	7.139
6	Antara News	Tsunami setinggi 18 cm terpantau BMKG	108.700	1.999	4.627	9.453

Pemilihan video sumber data dilakukan melalui proses pencarian sistematis pada platform TikTok menggunakan kata kunci yang relevan dengan peristiwa gempa bumi dan tsunami Mindanao, mencakup variasi istilah dalam Bahasa Indonesia yang umum digunakan publik maupun media untuk merujuk peristiwa tersebut. Video kandidat yang muncul dari pencarian ini kemudian disaring melalui dua kriteria utama. Pertama, keterkaitan konten, yaitu video harus secara langsung membahas atau melaporkan peristiwa gempa dan tsunami Mindanao, bukan sekadar menyinggung topik tersebut secara tangensial. Kedua, tingkat keterlibatan (*engagement*) yang memadai, yaitu volume komentar yang cukup untuk mendukung analisis sentimen secara bermakna, mengingat video dengan jumlah komentar yang sangat sedikit tidak akan memberikan sampel yang representatif untuk proses pelabelan maupun pemodelan. Video-video yang dipilih juga mencakup sumber dari akun media berita arus utama, sehingga komentar yang terkumpul merepresentasikan reaksi publik terhadap pemberitaan peristiwa tersebut, bukan opini yang muncul dari konten hiburan atau opini pribadi kreator video.

Diterapkan dua jalur *preprocessing* yang berbeda sesuai kebutuhan masing-masing model. Jalur pertama untuk SVM + TF-IDF, sedangkan jalur kedua untuk IndoBERT + BiLSTM. Jalur SVM mencakup enam tahapan berurutan, yaitu Case Folding, mengubah seluruh huruf dalam teks menjadi huruf kecil [24]; Cleaning Data, membersihkan teks dari noise seperti tanda baca, angka, URL, emoji, dan karakter khusus [25]; Tokenization, memecah kalimat atau dokumen menjadi unit kata (token) [26] menggunakan fungsi `nltk.word_tokenize()`; Normalisasi, mengubah kata tidak baku atau bahasa gaul menjadi bentuk baku [27] menggunakan kamus slang kustom (`slang_dict.json`) yang dikembangkan secara khusus untuk konteks bahasa Indonesia informal; Stopword Removal, menghapus kata-kata umum yang tidak memiliki pengaruh besar terhadap makna sentimen, seperti “dan”, “yang”, atau “di” [24] menggunakan gabungan stopwords NLTK bahasa Indonesia dan stopwords Tala; Stemming, mengubah kata berimbuhan menjadi bentuk kata dasar [28] menggunakan PySastrawi StemmerFactory yang dirancang khusus untuk morfologi bahasa Indonesia. Jalur IndoBERT + BiLSTM hanya meliputi pembersihan ringan, mengingat model berbasis Transformer telah memiliki kemampuan pemahaman kontekstual secara mandiri dengan hanya menerapkan tokenisasi berbasis *subword* menggunakan tokenizer bawaan IndoBERT dengan panjang sekuen maksimum 128 token [13].

Pelabelan sentimen dilakukan secara otomatis menggunakan model pra-latih Indonesian RoBERTa-base yang diimplementasikan melalui pustaka Hugging Face Transformers [29]. Pemanfaatan model pra-latih dipilih karena mampu memahami konteks semantik dan sintaksis bahasa Indonesia secara lebih baik dibandingkan pendekatan berbasis kamus maupun metode machine learning konvensional. Selain itu, pendekatan pelabelan otomatis berbasis model transformer terbukti efisien dan andal dalam melakukan anotasi data teks berukuran besar tanpa memerlukan proses pelabelan manual yang memakan waktu dan sumber daya yang signifikan [30]. Pada proses ini, setiap komentar hasil scraping dikonversi menjadi representasi token menggunakan tokenizer yang sesuai dengan model Indonesian RoBERTa-base. Selanjutnya, model melakukan inferensi untuk menghasilkan probabilitas pada masing-masing kelas sentimen. Kelas dengan nilai probabilitas tertinggi dipilih sebagai label akhir komentar. Sentimen yang digunakan dalam penelitian ini terdiri atas tiga kategori, yaitu positif, negatif, dan netral.

Mengingat label dari Indonesian RoBERTa-base merupakan keluaran model dan bukan anotasi yang divalidasi oleh manusia, pelabelan otomatis pada penelitian ini diperlakukan sebagai *pseudo-labeling* (*weak supervision*) untuk keperluan pelatihan model. Untuk mengatasi keterbatasan ini sekaligus menyediakan tolok ukur evaluasi yang independen dari model pelabel itu sendiri, data uji dianotasi secara manual oleh dua annotator yang bekerja secara independen satu sama lain. Proses anotasi telah disepakati mengenai definisi dan kriteria operasional untuk masing-masing kelas sentimen (positif, negatif, dan netral) secara khusus dalam konteks komunikasi bencana, termasuk arahan untuk menangani kasus yang berpotensi ambigu seperti komentar yang memuat unsur doa, ungkapan duka, maupun sarkasme. Perbedaan hasil anotasi antara kedua annotator diselesaikan melalui diskusi hingga tercapai konsensus.

Komentar yang dikumpulkan pada penelitian ini bersifat publik dan dapat diakses tanpa autentikasi pada platform TikTok. Meskipun demikian, penelitian ini tetap memperhatikan aspek privasi pengguna dalam proses pengumpulan dan pengolahan data. Data mentah hasil scraping yang memuat metadata identitas

pengguna dihapus dari dataset sebelum tahap pelabelan, pemodelan, dan analisis lebih lanjut, sehingga dataset yang digunakan pada penelitian ini hanya memuat teks komentar, tautan video sumber, dan label sentimen, tanpa informasi yang dapat mengidentifikasi penulis komentar secara langsung. Data disimpan dalam bentuk teranonimisasi pada penyimpanan internal tim peneliti dengan akses yang dibatasi hanya untuk keperluan penelitian ini. Contoh komentar yang ditampilkan pada Tabel 4 dipilih dengan mempertimbangkan aspek anonimitas, sehingga tidak memuat informasi tambahan yang memungkinkan penelusuran kembali ke akun pengguna tertentu. Dataset lengkap tidak dipublikasikan secara terbuka mengingat pertimbangan privasi pengguna, namun tersedia atas permintaan yang wajar kepada penulis korespondensi untuk keperluan replikasi atau penelitian lanjutan.

Pemodelan klasifikasi dalam penelitian ini dilakukan dengan dua pendekatan utama, yaitu SVM TF-IDF dan IndoBERT BiLSTM. Pada metode SVM TF-IDF, teks direpresentasikan ke dalam bentuk vektor numerik menggunakan bobot TF-IDF yang menekankan kata-kata penting dalam dokumen relatif terhadap korpus, kemudian diklasifikasikan menggunakan algoritma SVM dengan kernel *Radial Basis Function* (RBF) untuk memisahkan kelas sentimen secara optimal. Sementara itu, pendekatan IndoBERT BiLSTM memanfaatkan representasi kontekstual dari model bahasa IndoBERT yang dilatih pada korpus besar berbahasa Indonesia, kemudian diproses lebih lanjut oleh arsitektur BiLSTM yang membaca sekuens data dua arah untuk menangkap konteks secara lebih komprehensif. Hasil keluaran BiLSTM kemudian diteruskan ke lapisan *fully connected* dengan aktivasi *softmax* untuk menghasilkan prediksi kelas sentimen. Dengan kombinasi ini, kedua metode diharapkan mampu memberikan performa yang kompetitif dalam klasifikasi sentimen berbasis teks media sosial.

Seluruh eksperimen dirancang agar dapat diterapkan konsisten pada pembagian data, pelatihan SVM, learning curve, dan BiLSTM. Pelabelan otomatis pada data latih menggunakan `w11wo/indonesian-roberta-base-sentiment-classifier`, sedangkan ekstraksi embedding memakai `indobenchmark/indobert-base-p1` (MAX_LEN=128). Eksperimen dijalankan pada CPU dengan versi pustaka utama: `scikit-learn 1.6.1`, `tensorflow 2.20.0`, `torch 2.11.0+cpu`, `transformers 5.16.1`, dan `PySastrawi 1.2.1`. Konfigurasi pelatihan tiap model dirangkum pada Tabel 2.

Tabel 2. Konfigurasi Pelatihan Model

Komponen	Nilai
Seed global	42 (<code>train_test_split</code> , SVM, <code>learning_curve</code> , <code>tf.random.set_seed</code>)
Model pelabelan otomatis	<code>w11wo/indonesian-roberta-base-sentiment-classifier</code>
Model embedding	<code>indobenchmark/indobert-base-p1</code> , MAX_LEN=128
Versi pustaka utama	<code>scikit-learn 1.6.1</code> ; <code>tensorflow 2.20.0</code> ; <code>torch 2.11.0+cpu</code> ; <code>transformers 5.16.1</code> ; <code>PySastrawi 1.2.1</code>
TF-IDF	<code>max_features=5000</code> ; <code>ngram_range=(1,3)</code> ; <code>min_df=2</code> ; <code>sublinear_tf=True</code>
SVM	<code>kernel=rbf</code> ; <code>C=1.0</code> ; <code>gamma=scale</code> ; <code>random_state=42</code>
IndoBERT (BiLSTM)	Frozen, <code>mode eval()</code> + <code>torch.no_grad()</code>
Arsitektur BiLSTM	BiLSTM 128→64 unit (<code>dropout 0.3</code> , <code>recurrent_dropout 0.2</code>); Dense 64 (ReLU); <code>dropout akhir 0.4</code>
Optimizer & Loss	Adam, <code>lr=2e-4</code> ; <code>categorical_crossentropy</code>
Pelatihan BiLSTM	<code>batch_size=32</code> ; <code>max epoch=30</code> ; <code>validation_split=0.1</code> (dari data latih); <code>early stopping</code> (<code>patience=5</code>); <code>ReduceLR (factor=0.5, patience=3)</code>
Pembagian data	Stratified 80:20, <i>cross validation</i> 5-fold, <code>random_state=42</code>

Evaluasi model menggunakan metrik akurasi, presisi, *recall*, *F1-Score* [31], [32], [33], dan *cross validation* 5-fold [34] pada skema pembagian data latih dan data uji sebesar 80:20. Skema 5-fold cross-validation pada penelitian ini tidak digunakan sebagai protokol evaluasi utama, melainkan secara khusus sebagai bagian dari analisis learning curve untuk mengukur konsistensi dan kestabilan generalisasi model terhadap data yang tidak digunakan dalam proses pelatihan. Secara matematis evaluasi model dapat dituliskan sebagai berikut:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

$$Precision = \frac{TP}{TP + FP} \quad (10)$$

$$Recall = \frac{TP}{TP + FN} \quad (11)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (12)$$

4. HASIL DAN PEMBAHASAN

4.1. Pengumpulan dan Distribusi Data

Pengumpulan data dilakukan melalui teknik *scraping* terhadap komentar pengguna pada enam video TikTok yang berkaitan dengan peristiwa gempa bumi Filipina. Dari proses tersebut, dihimpun sebanyak 2.294 komentar mentah dengan distribusi yang tercantum pada Tabel 3.

Tabel 3. Distribusi komentar per video tiktok

Video	Jumlah Komentar	Presentase
1	262	11,42%
2	329	14,34%
3	84	3,66%
4	390	17,00%
5	129	5,62%
6	1.100	47,95%
Total	2.294	100,00%

Berdasarkan Tabel 2, Video 6 mendominasi jumlah komentar dengan 1.100 data (47,95% dari keseluruhan korpus), sementara Video 3 memiliki jumlah paling sedikit dengan 84 komentar (3,66%). Ketimpangan distribusi ini mencerminkan perbedaan tingkat popularitas dan jangkauan masing-masing video di platform TikTok.

4.2. Preprocessing Data

Diterapkan dua jalur preprocessing yang berbeda sesuai dengan kebutuhan masing-masing algoritma. Jalur pertama untuk model SVM + TF-IDF mencakup preprocessing lengkap, sedangkan jalur kedua untuk model IndoBERT + BiLSTM hanya meliputi pembersihan ringan tanpa stemming dan penghapusan stopword, mengingat model berbasis BERT telah memiliki kemampuan pemahaman kontekstual secara mandiri. Contoh hasil preprocessing bertahap disajikan pada Tabel 4.

Tabel 4. Contoh hasil preprocessing bertahap pada komentar sampel

Tahapan	Output
Input Asli	Semoga korban gempa di Filipina cepet dapet bantuan ya Allah #PrayForPhilippines
Cleaning & Case Folding	semoga korban gempa di filipina cepet dapet bantuan ya allah prayforphilippines
Tokenisasi	['semoga', 'korban', 'gempa', 'di', 'filipina', 'cepet', 'dapet', 'bantuan', 'ya', 'allah', 'prayforphilippines']
Normalisasi Slang	Tidak ada perubahan (11 token)
Stopword Removal	['semoga', 'korban', 'gempa', 'filipina', 'cepet', 'dapet', 'bantuan', 'allah', 'prayforphilippines'] (9 token)
Stemming	['moga', 'korban', 'gempa', 'filipina', 'cepet', 'dapet', 'bantu', 'allah', 'prayforphilippines']
Output Akhir	moga korban gempa filipina cepet dapet bantu allah prayforphilippines

Dari total 2.294 komentar yang dikumpulkan, sebanyak 2.023 komentar dinyatakan valid setelah preprocessing. Sebanyak 271 komentar tidak dapat diproses lebih lanjut karena berupa stiker, emoji tunggal, atau teks yang tidak mengandung konten linguistik bermakna.

4.3. Pelabelan Sentimen

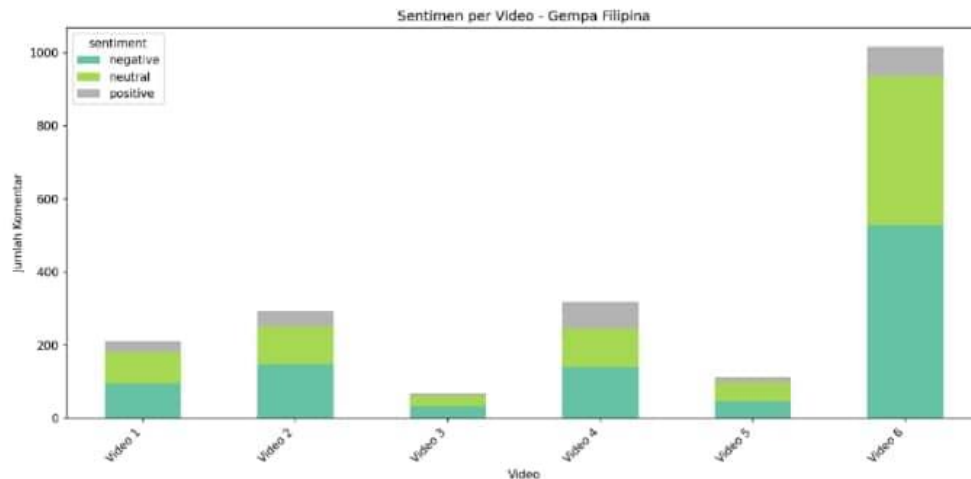
Pelabelan sentimen dilakukan secara otomatis menggunakan model pra-latih Indonesian RoBERTa-base melalui pustaka Hugging Face Transformers. Setiap komentar diklasifikasikan ke dalam tiga kelas sentimen: positif, negatif, dan netral. Hasil distribusi pelabelan disajikan pada Tabel 5.

Tabel 5. Distribusi kelas sentimen pada data berlabel

Kelas Sentimen	Jumlah Komentar	Presentase
Negatif	995	49,23%
Netral	784	38,79%
Positif	244	12,08%

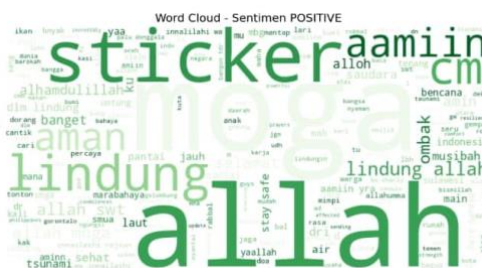
Distribusi kelas sentimen bersifat tidak seimbang (*imbalanced*). Kelas negatif mendominasi dataset dengan 995 komentar (49,23%), diikuti kelas netral sebanyak 784 komentar (38,79%), dan kelas positif sebagai kelas minoritas dengan 244 komentar (12,08%). Ketidakseimbangan ini

merupakan karakteristik umum analisis sentimen pada konteks bencana, mengingat respons emosional masyarakat cenderung mengekspresikan keprihatinan dan kecemasan. Seluruh data berlabel dibagi dengan rasio 80:20 menggunakan stratified split, menghasilkan 1.618 data latih dan 405 data uji. Gambar 3 menyajikan Distribusi Sentimen Keseluruhan pada Komentar TikTok terkait Gempa Filipina.



Gambar 3. Distribusi sentimen per video tiktok terkait gempa filipina

Analisis per video menunjukkan bahwa sentimen negatif menjadi proporsi terbesar pada sebagian video, terutama Video 2, Video 4, dan Video 6, sementara pada video lain seperti Video 1, Video 3, dan Video 5, proporsi sentimen negatif dan netral relatif berimbang. Gambar 4 menyajikan *Word Cloud* Komentar Sentimen Positif terkait Gempa Filipina.



Gambar 4. Word cloud komentar sentimen positif terkait gempa filipina

Pada kelas positif, kata-kata yang paling dominan adalah "allah", "aamiin", "lindung", "aman", dan "saudara". Pola ini mencerminkan ekspresi doa dan harapan dari pengguna yang memanjatkan perlindungan bagi korban gempa. Gambar 5 menyajikan *Word Cloud* Komentar Sentimen Negatif terkait Gempa Filipina.



Gambar 5. Word cloud komentar sentimen negatif terkait gempa filipina

Pada kelas negatif, kata dominan meliputi "tsunami", "panik", "lari", "ombak", "bahaya", dan "takut". Dominasi kosakata yang berkaitan dengan ancaman fisik bencana mencerminkan ekspresi kecemasan dan ketakutan masyarakat terhadap dampak gempa. Gambar 6 menyajikan *Word Cloud* Komentar Sentimen Netral terkait Gempa Filipina.



Gambar 6. Word cloud komentar sentimen netral terkait gempa filipina

Pada kelas netral, kata menonjol antara lain "tsunami", "gempa", "waspada", "berita", dan "BMKG", yang mengindikasikan bahwa komentar netral lebih bersifat informatif berupa penyebaran informasi faktual dan imbauan kewaspadaan.

Word cloud pada Gambar 4, 5, dan 6 menampilkan kata-kata dengan frekuensi ternormalisasi tertinggi pada masing-masing kelas sentimen. Pada kelas positif, kata-kata yang menonjol didominasi kosakata religius dan solidaritas, konsisten dengan pola umum komunikasi bencana di Indonesia yang sering menyertakan doa dan harapan. Pada kelas negatif, kata-kata yang menonjol umumnya berkaitan dengan ekspresi kekhawatiran atau ketakutan terhadap penanganan peristiwa, sementara pada kelas netral kata-kata yang muncul cenderung bersifat informatif dan deskriptif, sejalan dengan karakteristik komentar yang menyampaikan informasi tanpa muatan afektif yang kuat. Meski demikian, kemunculan sebuah kata dengan frekuensi tinggi tidak dengan sendirinya merepresentasikan keadaan psikologis penulis komentar secara individual, karena kata yang sama dapat digunakan dalam konteks dan dengan maksud yang berbeda-beda.

4.4. Hasil Klasifikasi Model

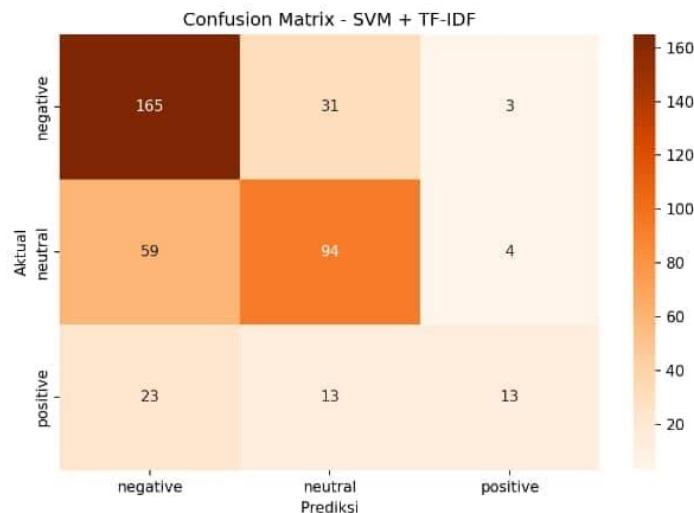
4.4.1. SVM + TF-IDF

Model pertama mengombinasikan *Support Vector Machine* (SVM) dengan pembobotan fitur Term Frequency-Inverse Document Frequency (TF-IDF). Kernel yang digunakan adalah *Radial Basis Function* (RBF) sebagai konfigurasi default scikit-learn. Selain evaluasi utama menggunakan skema pembagian data 80:20, penelitian ini juga mengkombinasikan dengan skema 5-fold *cross-validation* (cv=5) pada model SVM dengan kernel RBF. Pada skema ini, data latih dibagi secara acak ke dalam lima subset dengan ukuran yang relatif seimbang. Pada setiap iterasi, empat subset digunakan sebagai data latih dan satu subset sisanya digunakan sebagai data validasi, sehingga proses ini diulang sebanyak lima kali hingga setiap subset pernah berperan sebagai data validasi. Skor akurasi validasi kemudian dirata-ratakan dari kelima iterasi tersebut untuk memperoleh estimasi performa model yang lebih stabil dan tidak bias terhadap satu pembagian data tertentu.

Pendekatan ini dipilih dengan pertimbangan mengurangi bias evaluasi, karena performa yang diamati tidak hanya bergantung pada satu kali pembagian data latih-uji, melainkan merupakan rata-rata dari lima skema pembagian yang berbeda. Skema ini memaksimalkan efisiensi pemanfaatan data, yang relevan mengingat jumlah data pada penelitian ini relatif terbatas, sehingga seluruh data tetap dimanfaatkan baik sebagai data latih maupun validasi pada tiap iterasi. Waktu pelatihan model tercatat 0,21 detik dengan total waktu pipeline sebesar 0,31 detik. Hasil evaluasi pada data uji disajikan pada Tabel 6 dan confusion matrix pada gambar 7.

Tabel 6. Laporan klasifikasi model SVM + TF-IDF

Kelas Sentimen	Precision	Recall	F1-Score	Support
Negatif	0,67	0,83	0,74	199
Netral	0,68	0,60	0,64	157
Positif	0,65	0,27	0,38	49
Macro Avg	0,67	0,56	0,58	405
Weighted Avg	0,67	0,67	0,66	405

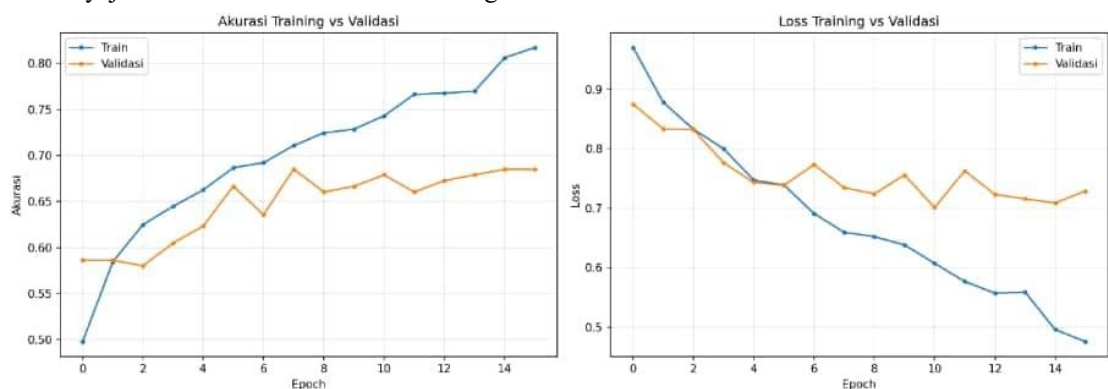


Gambar 7. Confusion matrix model SVM + TF-IDF

Model SVM + TF-IDF mencapai akurasi keseluruhan sebesar 67,16%. Kelas negatif memperoleh F1-score tertinggi sebesar 0,74 dengan recall 0,83, didukung oleh ketersediaan data latih yang paling banyak pada kelas tersebut. Sebaliknya, kelas positif menunjukkan kinerja terendah dengan F1-score 0,38 dan recall 0,27, di mana hanya 13 dari 49 sampel positif pada data uji yang berhasil diklasifikasikan dengan benar.

4.4.2. IndoBERT + BiLSTM

Model kedua merupakan arsitektur hibrida yang menggabungkan IndoBERT sebagai lapisan ekstraksi fitur kontekstual dengan Bidirectional Long Short-Term Memory (BiLSTM) sebagai lapisan klasifikasi sekuensial. Setiap komentar ditokenisasi menggunakan tokenizer IndoBERT dengan panjang sekuens maksimum 128 token. Embeddings kontekstual dari lapisan terakhir IndoBERT kemudian digunakan sebagai input lapisan BiLSTM. Ekstraksi embeddings memerlukan waktu 768,44 detik karena dijalankan pada CPU tanpa akselerasi GPU, sementara pelatihan BiLSTM berlangsung selama 16 epoch dengan early stopping, dengan total waktu latih 1.469,86 detik. Gambar 8 menyajikan kurva akurasi dan loss training.

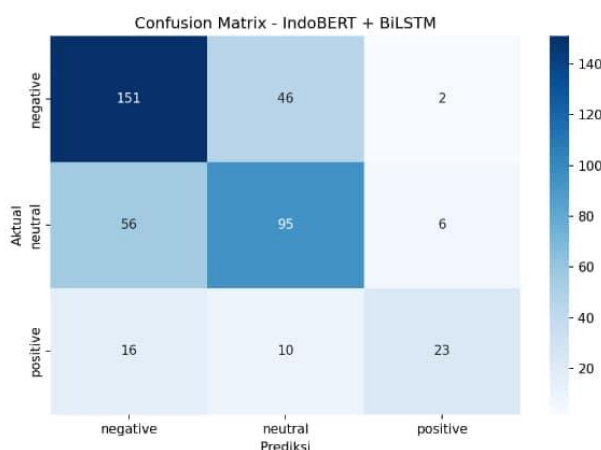


Gambar 8. Kurva akurasi dan loss training vs validasi model IndoBERT + BiLSTM

Kurva pelatihan menunjukkan bahwa akurasi latih meningkat konsisten dari 49,79% pada epoch pertama hingga 81,73% pada epoch ke-16, sementara akurasi validasi bergerak fluktuatif dengan nilai tertinggi 68,52%. Hasil evaluasi pada data uji disajikan pada Tabel 7 dan confusion matrix pada gambar 9.

Tabel 7. Laporan Klasifikasi model IndoBERT + BiLSTM

Kelas Sentimen	Precision	Recall	F1-Score	Support
Negatif	0,68	0,76	0,72	199
Netral	0,63	0,61	0,62	157
Positif	0,74	0,47	0,57	49
Macro Avg	0,68	0,61	0,64	405
Weighted Avg	0,67	0,66	0,66	405



Gambar 9. Confusion matrix model IndoBERT + BiLSTM

Model IndoBERT + BiLSTM memperoleh akurasi keseluruhan sebesar 66,42%. Dibandingkan SVM + TF-IDF, model ini menunjukkan kinerja yang lebih baik pada kelas positif dengan F1-score 0,57 dan precision 0,74, serta berhasil mengklasifikasikan 23 dari 49 sampel positif dengan benar. Namun, model ini mencatat recall yang lebih rendah pada kelas negatif (0,76 berbanding 0,83 pada SVM).

4.5. Perbandingan Kinerja Model

Komparasi kinerja model antara SVM + TF-IDF dan IndoBERT + BiLSTM disajikan pada tabel 8.

Tabel 8. Komparasi kinerja model SVM+TF-IDF dan IndoBERT+BiLSTM

Model	Accuracy	Precision (Macro)	Recall (Macro)	F1 (Macro)	F1 (Weighted)	Waktu Total (dtk)
SVM + TF-IDF	0,67	0,67	0,56	0,58	0,66	0,31
IndoBERT + BiLSTM	0,66	0,68	0,61	0,64	0,66	2.249,36

Berdasarkan Tabel 7, SVM + TF-IDF sedikit mengungguli IndoBERT + BiLSTM dalam hal akurasi keseluruhan (67,16% berbanding 66,42%). Namun, jika dievaluasi menggunakan F1-score macro average sebagai metrik yang lebih representatif pada kondisi kelas tidak seimbang, IndoBERT + BiLSTM (0,64) mengungguli SVM + TF-IDF (0,58) dengan selisih 0,06 poin. F1-score weighted average kedua model hampir setara, yaitu 0,66 untuk masing-masing model. Perbedaan paling signifikan antara kedua model terletak pada efisiensi komputasi. Komponen training BiLSTM saja memerlukan 1.469,86 detik, namun tahap ekstraksi embedding dari IndoBERT turut menyumbang porsi waktu yang besar, yakni 768,44 detik, sehingga total waktu kedua tahap ini jauh melampaui waktu tokenisasi (0,37 detik) dan waktu prediksi (10,69 detik). Keunggulan IndoBERT + BiLSTM lebih terlihat pada kelas minoritas, dengan peningkatan recall dari 0,27 menjadi 0,47 dan precision dari 0,65 menjadi 0,74 untuk kelas positif.

5. KESIMPULAN

Penelitian ini berhasil menganalisis sentimen 2.294 komentar TikTok terkait gempa bumi dan tsunami Mindanao, dengan 2.023 komentar dinyatakan valid setelah preprocessing. Pelabelan menggunakan Indonesian RoBERTa-base menghasilkan distribusi sentimen negatif sebesar 49,23%, netral 38,79%, dan positif 12,08%, yang menunjukkan ketidakseimbangan kelas dengan sentimen positif sebagai kelas minoritas. Hasil analisis kata dominan memperlihatkan bahwa komentar positif banyak memuat ekspresi doa, harapan, dan solidaritas, komentar negatif didominasi kosakata yang berkaitan dengan ketakutan serta ancaman tsunami, sedangkan komentar netral cenderung berisi informasi dan imbauan kewaspadaan. Pada evaluasi data uji dengan pembagian stratified 80:20, SVM-TF-IDF memperoleh akurasi 67,16%, macro precision 0,67, macro recall 0,56, dan macro F1-score 0,58. Sementara itu, IndoBERT-BiLSTM memperoleh akurasi 66,42%, macro precision 0,68, macro recall 0,61, dan macro F1-score 0,64. SVM menunjukkan recall lebih tinggi pada kelas negatif, sedangkan IndoBERT-BiLSTM memberikan peningkatan kinerja pada kelas positif, dengan precision 0,74, recall 0,47, dan F1-score 0,57 dibandingkan SVM yang memperoleh precision 0,65, recall 0,27, dan F1-score 0,38.

Secara keseluruhan, kedua pendekatan memiliki karakteristik kinerja yang berbeda. SVM-TF-IDF memberikan akurasi sedikit lebih tinggi serta efisiensi komputasi yang sangat besar, dengan waktu total

pipeline hanya 0,31 detik. Sebaliknya, IndoBERT-BiLSTM menghasilkan macro F1-score lebih tinggi sebesar 0,64 dibandingkan 0,58 pada SVM-TF-IDF, yang menunjukkan performa yang lebih seimbang terhadap kelas sentimen, khususnya kelas positif yang jumlahnya lebih sedikit. Namun, pendekatan IndoBERT-BiLSTM memerlukan waktu komputasi jauh lebih besar, yaitu total 2.249,36 detik, yang terdiri atas 768,44 detik untuk ekstraksi embedding dan 1.469,86 detik untuk pelatihan BiLSTM. Dengan demikian, hasil penelitian menunjukkan bahwa evaluasi model pada dataset sentimen yang tidak seimbang perlu mempertimbangkan lebih dari sekadar akurasi, dengan macro F1-score memberikan perspektif tambahan mengenai kemampuan model dalam menangani setiap kelas sentimen.

REFERENSI

- [1] P. I. of V. and Seismology, "Primer on the 08 June 2026 magnitude (Mw) 7.8 offshore Sarangani earthquake," 2026. [Online]. Available: <https://www.phivolcs.dost.gov.ph/primer-on-the-08-june-2026-magnitude-mw-7-8-offshore-sarangani-earthquake>
- [2] U. S. G. Survey, "M 7.8 – 26 km SW of Kablalan, Philippines," 2026. [Online]. Available: <https://earthquake.usgs.gov/earthquakes/eventpage>
- [3] K. Badan Meteorologi dan Geofisika, "BMKG akhiri peringatan dini tsunami gempa M7,7 Filipina," 2026. [Online]. Available: <https://www.bmkg.go.id/siaran-pers/bmkg-akhiri-peringatan-dini-tsunami-gempa-m77-filipina-kenaikan-tinggi-muka-air-laut-terpantau-di-sejumlah-wilayah>
- [4] N. D. R. R. and M. Council, "Situational Report No. 12 for the Effects of Magnitude 7.8 Earthquake in Maasim, Sarangani," NDRRMC, Quezon City, Philippines, 2026.
- [5] Y. Liu, "Earthquake resilience and public engagement: A social media perspective," *Risk Anal.*, 2025, doi: 10.1111/risa.70041.
- [6] J. P. Stimpson, A. Srivastava, K. Tamirisa, J. K. Kaholokula, and A. N. Ortega, "Crisis communication about the Maui wildfires on TikTok," *JMIR Form. Res.*, vol. 9, p. e67515, 2025, doi: 10.2196/67515.
- [7] W. Liu, X. Zhao, M. Zhan, and S. Hernandez, "Streaming disasters on TikTok: Examining social mediated crisis communication, public engagement, and emotional responses during the 2023 Maui wildfire," *Public Relat. Rev.*, vol. 50, no. 3, 2024, doi: 10.1016/j.pubrev.2024.102479.
- [8] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Found. Trends Inf. Retr.*, vol. 2, no. 1–2, pp. 1–135, 2008, doi: 10.1561/1500000011.
- [9] S. Mendon, P. Dutta, A. Behl, and S. Lessmann, "A hybrid approach of machine learning and lexicons to sentiment analysis," *Inf. Syst. Front.*, vol. 23, no. 5, pp. 1145–1168, 2021, doi: 10.1007/s10796-021-10107-x.
- [10] H. Murfi, S. T. Gowandi, G. Ardaneswari, and S. Nurrohmah, "BERT-based combination of convolutional and recurrent neural network for Indonesian sentiment analysis," *Appl. Soft Comput.*, vol. 151, p. 111112, 2024, doi: 10.1016/j.asoc.2023.111112.
- [11] R. Sutoyo, "Sentiment analysis of AI adoption in Indonesian higher education," 2025. [Online]. Available: <https://arxiv.org/abs/2604.27439>
- [12] M. Ahmadian, "Hybrid models for emotion classification and sentiment analysis in Indonesian language," *Appl. Comput. Intell. Soft Comput.*, vol. 2024, p. Article 2826773, 2024, doi: 10.1155/2024/2826773.
- [13] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A benchmark dataset and pre-trained language model for Indonesian NLP," in *Proceedings of the 28th International Conference on Computational Linguistics (COLING 2020)*, 2020, pp. 757–770. doi: 10.18653/v1/2020.coling-main.66.
- [14] M. A. Farizi, "Implementation of BiLSTM and IndoBERT for sentiment analysis of TikTok reviews," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 10, no. 1, 2025, doi: 10.29100/jipi.v10i1.5815.
- [15] J. C. Setiawan, K. M. Lhaksana, and B. Bunyamin, "Sentiment analysis of Indonesian TikTok review using LSTM and IndoBERTweet algorithm," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 8, no. 3, pp. 774–780, 2023, doi: 10.29100/jipi.v8i3.3911.
- [16] L. Geni, "Sentiment analysis of tweets before the 2024 elections in Indonesia using IndoBERT language model," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 9, no. 3, pp. 746–757, 2023,.
- [17] R. Chaudhari, M. Biradar, and T. Mondal, "Understanding post-disaster queries using social media network analysis and a machine learning model," in *Proceedings of the International Conference on Intelligent Systems*, Springer, 2026. doi: 10.1007/978-3-032-18480-1_36.
- [18] M. Karim, M. M. S. Missen, M. Umer, A. Fida, A. A. Eshmawi, A. Mohamed, and I. Ashraf, "Comprehension of polarity of articles by citation sentiment analysis using TF-IDF and ML classifiers," *PeerJ Comput. Sci.*, vol. 8, Art. no. e1107, 2022, doi: 10.7717/peerj-cs.1107.
- [19] C. Cortes and V. Vapnik, "Support-Vector Networks," *Kluwer Acad. Publ.*, vol. 20, no. 2, pp. 273–297, 1995, doi: 10.1111/j.1747-0285.2009.00840.x.
- [20] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in

- Proc. 5th Annu. Workshop Comput. Learn. Theory (COLT), 1992, pp. 144–152, doi: 10.1145/130385.130401.
- [21] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, and B. Thirion, “Scikit-learn: Machine Learning in Python,” *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011, doi: 10.1289/EHP4713.
- [22] M. Schuster and K. K. Paliwal, “Bidirectional recurrent neural networks,” *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2673–2681, 1997, doi: 10.1109/78.650093.
- [23] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
- [24] A. F. Hidayatullah and M. R. Ma’Arif, “Pre-processing tasks in Indonesian Twitter messages,” in *Journal of Physics: Conference Series*, 2017, doi: 10.1088/1742-6596/801/1/012072.
- [25] R. B. Hadiprakoso, H. Setiawan, and R. N. Yasa, “Text preprocessing for optimal accuracy in Indonesian sentiment analysis using a deep learning model with word embedding,” in *AIP Conference Proceedings*, 2023, doi: 10.1063/5.0126116.
- [26] A. W. Pradana and M. Hayaty, “The effect of stemming and removal of stopwords on the accuracy of sentiment analysis on Indonesian-language texts,” *Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control*, vol. 4, no. 4, 2019, doi: 10.22219/kinetik.v4i4.912.
- [27] E. Nugraheni, F. I. Haekal, and A. Arisal, “Optimizing Indonesian tweet preprocessing on halal domain,” in *IEEE International Conference Proceedings*, 2024. [Online]. Available: <https://ieeexplore.ieee.org/document/10732128/>
- [28] R. Rianto, A. B. Mutiara, E. P. Wibowo, and P. I. Santosa, “Improving the accuracy of text classification using stemming method, a case of non-formal Indonesian conversation,” *J. Big Data*, vol. 8, no. 1, 2021, doi: 10.1186/s40537-021-00413-1.
- [29] T. Wolf *et al.*, “Transformers: State-of-the-art natural language processing,” in *Proceedings of EMNLP 2020*, 2020, pp. 38–45. doi: 10.18653/v1/2020.emnlp-demos.6.
- [30] J. Hartmann, M. Heitmann, C. Siebert, and C. Schamp, “More than a feeling: Accuracy and application of sentiment analysis,” *Int. J. Res. Mark.*, vol. 40, no. 1, pp. 75–87, 2023, doi: 10.1016/j.ijresmar.2022.05.005.
- [31] H. Nguyen, A. Veluchamy, and M. Diop, “Comparative study of sentiment analysis with product reviews using machine learning and lexicon-based approaches,” *SMU Data Sci. Rev.*, vol. 1, no. 4, 2018, [Online]. Available: <https://scholar.smu.edu/datasciencereview/vol1/iss4/7/>
- [32] K. Gulati, S. S. Kumar, R. S. K. Boddu, and K. Sarvakar, “Comparative analysis of machine learning-based classification models using sentiment classification of tweets related to COVID-19 pandemic,” *Mater. Today Proc.*, vol. 56, pp. 3505–3514, 2022, doi: 10.1016/j.matpr.2021.03.584.
- [33] A. Madasu and S. Elango, “Efficient feature selection techniques for sentiment analysis,” *Multimed. Tools Appl.*, vol. 79, pp. 6313–6335, 2020, doi: 10.1007/s11042-019-08409-z.
- [34] M. F. Fakhrezi and A. F. Rochim, “Comparison of sentiment analysis methods based on accuracy value case study: Twitter mentions of academic article,” *J. RESTI*, 2023, doi: <https://doi.org/10.29207/resti.v7i1.4767>.