VOL. 21. NO. 1 AGUSTUS 2026
P-ISSN: 0216-6933 | E-ISSN: 2685-807X
Diterbitkan Oleh: LPPM Universitas Nusa Mandiri

INTI NUSA MANDIRI
DOI: https://doi.org/10.33480/inti.v21i1.8550

ANALISIS SENTIMEN INFORMASI GEMPA BUMI BMKG PADA APLIKASI X MENGGUNAKAN SVM DAN RANDOM FOREST

Arfany Dhimas Muftareza[1*]; Reza Okta Pratama[2]; I Dewa Gede Loka Maheswara[3]; Giarno[4]; Agustina Rachmawardani[5]

Instrumentasi[1,2,5]
Meteorologi[3]
Klimatologi[4]
Sekolah Tinggi Meteorologi Klimatologi dan Geofisika, Kota Tangerang, Indonesia[1,2,3,4,5]
www.stmkg.ac.id[1,2,3,4,5]
arfanydhimuf@gmail.com[1*], rezaoktapp@gmail.com[2], maheswaradewaloka@gmail.com[3], giarnostmkg@gmail.com[4], agustina.rahmawardani@stmkg.ac.id[5]

(*) Corresponding Author

Abstract— *This study analyzes public sentiment towards the dissemination of earthquake information by BMKG through the X application with modeling using the Support Vector Machine (SVM) and Random Forest (RF) algorithms. Data were collected from 2,711 tweets mentioning @infoBMKG using tweet-harvest, then processed through the stages of case folding, cleaning, tokenization, slang normalization, stopword removal, and stemming. Automatic sentiment labeling was performed using a hybrid approach of VADER and InSet Lexicon. Feature representation used TF-IDF (Term Frequency–Inverse Document Frequency) with 1,000 features and data distribution 80% train and 20% validation. The results show that RF achieved an accuracy of 81.92% and SVM 81.17%, with almost identical Macro F1 (RF: 0.7445; SVM: 0.7443). Neutral sentiment indicates informative tweets without emotional content (61.75%), negative sentiment represents the public's emotional response that is not solely intended as a form of negative assessment of BMKG (24.71%), and positive sentiment is an expression of appreciation, gratitude, and hope for the delivery of information (13.54%). SVM excels in cross-validation stability (std ±0.0574) and negative sentiment recall (0.71), making it more suitable for real-time disaster communication monitoring.*

Keywords*: BMKG, Random Forest, Sentiment Analysis, Support Vector Machine, X Application.*

Abstrak—Penelitian ini menganalisis sentimen publik terhadap diseminasi informasi gempa bumi oleh BMKG melalui aplikasi X dengan pemodelan menggunakan algoritma *Support Vector Machine* (SVM) dan *Random Forest* (RF). Data dikumpulkan dari 2.711 *tweet* yang menyebut @infoBMKG menggunakan *tweet-harvest*, kemudian diproses melalui tahapan *case folding*, *cleaning*, tokenisasi, normalisasi slang, *stopword removal*, dan *stemming*. Pelabelan sentimen otomatis dilakukan menggunakan pendekatan *hybrid* VADER dan InSet Lexicon. Representasi fitur menggunakan TF-IDF (*Term Frequency–Inverse Document Frequency*) dengan 1.000 fitur dan pembagian data 80% *train* dan 20% *validation*. Hasil menunjukkan RF mencapai akurasi 81,92% dan SVM 81,17%, dengan Macro F1 yang hampir identik (RF: 0,7445; SVM: 0,7443). Sentimen netral mengindikasikan tweet bersifat informatif tanpa muatan emosi (61,75%), sentimen negatif merepresentasikan respons emosional masyarakat yang bukan semata-mata ditujukan sebagai bentuk penilaian negatif terhadap BMKG (24,71%), dan sentimen positif berupa ungkapan apresiasi, rasa syukur, dan harapan terhadap penyampaian informasi (13,54%). SVM unggul dalam stabilitas *cross-validation* (std ±0,0574) dan *recall* sentimen negatif (0,71), menjadikannya lebih sesuai untuk pemantauan komunikasi kebencanaan *real-time*.

Kata kunci: *BMKG, Analisis Sentimen, Aplikasi X, Random Forest, Support Vector Machine.*

INTI NUSA MANDIRI
DOI: https://doi.org/10.33480/inti.v21i1.8550

VOL. 21. NO. 1 AGUSTUS 2026
P-ISSN: 0216-6933 | E-ISSN: 2685-807X
Diterbitkan Oleh: LPPM Universitas Nusa Mandiri

PENDAHULUAN

Indonesia merupakan negara kepulauan yang terletak pada pertemuan tiga lempeng tektonik utama, yaitu Lempeng Eurasia, Lempeng Indo-Australia, dan Lempeng Pasifik, sehingga menjadikannya sebagai salah satu wilayah dengan tingkat aktivitas seismik tertinggi di dunia (Handoyo, 2024; Zulfakriza, 2024). Kecepatan dan ketepatan penyebarluasan informasi gempa bumi menjadi faktor krusial dalam upaya melindungi keselamatan masyarakat dan meminimalisir dampak dari bencana.

Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) berwenang memanfaatkan *platform* Twitter yang saat ini menjadi aplikasi X melalui akun resmi @infoBMKG untuk menyampaikan informasi gempa bumi secara *real-time* kepada jutaan pengguna (Hayati et al., 2023a). Setiap unggahan memunculkan beragam respons masyarakat, mulai dari apresiasi terhadap kecepatan informasi hingga kritik terkait format penyampaian, yang mencerminkan dinamika sentimen publik penting untuk dikaji (Aldamen & Hacimic, 2023; Rudyawan, 2023).

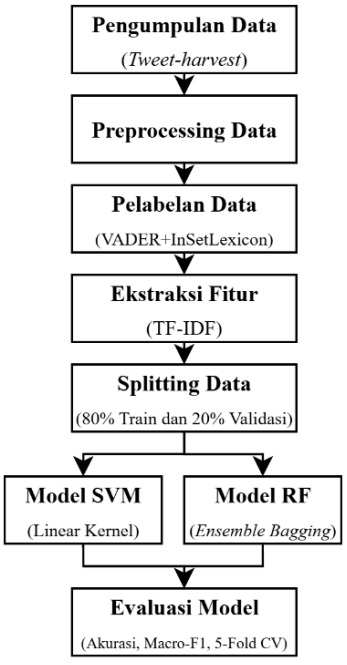
Analisis sentimen merupakan cabang *Natural Language Processing* (NLP) yang bertujuan mengklasifikasikan opini teks ke dalam kategori positif, negatif, atau netral (Tan et al., 2023). Tantangan utama pada data Aplikasi X berbahasa Indonesia adalah keberagaman bahasa, penggunaan bahasa gaul, dan banyaknya singkatan, sehingga diperlukan *preprocessing* dan pelabelan data yang cermat (Nugraheni et al., 2024). Penggunaan leksikon sentimen InSet Lexicon penting untuk meningkatkan kualitas pelabelan data bahasa Indonesia secara otomatis (Tiwari et al., 2026).

Beberapa penelitian sebelumnya memanfaatkan SVM dan Random Forest untuk analisis sentimen pada Aplikasi X berbahasa Indonesia. Alita (2024) menggunakan SVM untuk sentimen Pilpres tahun 2024, sementara Khoirunnisa & Ningrum (2026) membandingkan Naive Bayes, Random Forest, dan SVM pada data bencana banjir. Kedua penelitian tersebut belum mengeksplorasi kombinasi lebih dari satu pendekatan lexicon untuk pelabelan otomatis dan hanya berfokus pada isu politik dan bencana hidrometeorologi, belum menyentuh konteks diseminasi informasi gempa bumi oleh lembaga resmi seperti BMKG.

Penelitian Hayati (2023) menunjukkan bahwa Aplikasi X efektif digunakan BMKG sebagai media komunikasi digital untuk menyampaikan informasi gempa bumi secara cepat dan meningkatkan keterlibatan masyarakat melalui *hashtag* dan *citizen journalism*. Sementara itu, Rudyawan (2023) menjelaskan bahwa media sosial berperan penting dalam penyebaran informasi dan edukasi kebencanaan pascagempa Cianjur, namun juga memunculkan misinformasi dan dampak sosial selama bencana berlangsung.

Dengan demikian, terdapat gap penelitian pada aspek objek kajian spesifik berupa akun resmi @infoBMKG sebagai sumber informasi gempa bumi melalui analisis sentimen publik terhadap diseminasi informasi gempa bumi oleh BMKG dan metode pelabelan sentimen otomatis yang menggabungkan dua leksikon berbeda (VADER dan InSet Lexicon) untuk meningkatkan akurasi pelabelan data berbahasa Indonesia. Novelty penelitian ini terletak pada penerapan pendekatan hybrid VADER–InSet Lexicon untuk pelabelan sentimen otomatis yang dipadukan dengan perbandingan performa SVM dan *Random Forest* secara khusus pada domain diseminasi informasi gempa bumi BMKG, yang belum pernah dilakukan pada penelitian-penelitian sebelumnya. Hasil penelitian diharapkan memberikan masukan bagi BMKG dalam memahami persepsi masyarakat dan merumuskan strategi komunikasi kebencanaan yang lebih efektif.

BAHAN DAN METODE



Sumber: (Hasil Penelitian, 2026)
Gambar 1. Diagram Alir Penelitian

Penelitian ini menggunakan metode kuantitatif dengan pendekatan eksperimen (*experimental research*) berbasis *machine learning*

dan *text mining* untuk klasifikasi sentimen yang menggunakan data numerik dan pengujian eksperimental untuk menganalisis dan menemukan pola informasi dari data teks menggunakan algoritma komputasi (NURSUKUWAGUS et al., 2026). Alur penelitian ini dilakukan dengan pengumpulan data, *preprocessing* data, pelabelan sentimen, splitting data, ekstraksi fitur TF-IDF, seleksi fitur menggunakan Chi-Square, pelatihan model klasifikasi SVM dan Random Forest, serta evaluasi model menggunakan akurasi, Macro F1, dan validasi silang 5-fold.

Pengumpulan Data

Data dikumpulkan dari Aplikasi X menggunakan *tweet-harvest* untuk menangkap respons masyarakat berbahasa Indonesia terhadap informasi gempa bumi yang disebarkan BMKG, namun secara eksplisit mengecualikan tweet langsung dari akun resmi tersebut dan tautan berita. Pengumpulan data dilakukan pada 18 April 2026 secara berkala yang mencakup data komentar pada aplikasi X mulai dari tanggal 1 Februari 2026 hingga 11 April 2026. Query yang digunakan dalam penelitian ini:

(@infoBMKG OR "BMKG") AND ("gempa") -from:infoBMKG -from:BMKG -filter:links lang:id

Preprocessing Data

Preprocessing dilakukan secara berurutan melalui enam tahap untuk memastikan kualitas teks yang optimal untuk analisis sentimen, diantaranya:

1. *Case Folding*, yaitu mengubah seluruh huruf menjadi huruf kecil akibat variasi kapitalisasi (Ramadhani et al., 2026).
2. *Cleaning*, yaitu menghapus noise seperti URL, emoji, dan tanda baca agar data lebih relevan (Hadiprakoso et al., 2023).
3. *Tokenization*, yaitu memecah kalimat menjadi unit kata (Ramadhani et al., 2026) menggunakan nltk.word_tokenize().
4. Normalisasi Slang, yaitu mengubah kata tidak baku menjadi bentuk baku (Nugraheni et al., 2024) menggunakan kamus slang_dict.json.
5. *Stopword Removal*, yaitu menghapus kata umum (Ramadhani et al., 2026) menggunakan NLTK dan stopwords Tala.
6. *Stemming*, yaitu mengubah kata berimbuhan ke bentuk dasar (Rianto et al., 2021) menggunakan PySastrawi.

Pelabelan Data

Pelabelan sentimen dilakukan secara otomatis menggunakan pendekatan hybrid VADER dan InSet Lexicon (Fauziah et al., 2021; Khairani & Setiawan, 2024). Penentuan skor untuk setiap tweet dilakukan dengan ketentuan, jika token *tweet* cocok dengan InSet Lexicon, skor akhir merupakan gabungan berbobot 70% skor leksikon Indonesia dan 30% skor *compound* VADER. Jika tidak ada kecocokan, 100% skor VADER digunakan. Kelas sentimen ditentukan berdasarkan ambang batas (*Threshold)*, yaitu positif (≥ 0,05), negatif (≤ -0,05), dan netral (-0,05 < skor < 0,05).

Pendekatan pelabelan otomatis dipilih dengan pertimbangan efisiensi dan konsistensi dalam menangani volume data yang besar (2.711 tweet), yang apabila dilakukan secara manual oleh anotator manusia akan memerlukan waktu yang signifikan dan berisiko menimbulkan inkonsistensi antar-anotator (*inter-annotator disagreement*), khususnya untuk kebutuhan pemantauan sentimen publik yang bersifat mendekati waktu nyata (*near real-time*) pada konteks bencana. Pendekatan hybrid VADER-InSet Lexicon dipilih karena kombinasi keduanya telah terbukti meningkatkan akurasi pelabelan otomatis pada teks berbahasa Indonesia dibandingkan penggunaan satu leksikon tunggal (Fauziah et al., 2021; Khairani & Setiawan, 2024).

TF-IDF *Vectorization*

TF-IDF (*Term Frequency–Inverse Document Frequency*) adalah metode pembobotan kata dalam analisis sentimen yang digunakan untuk mengukur seberapa penting suatu kata dalam dokumen dibandingkan dengan keseluruhan dokumen pada kumpulan data (Karim et al., 2022).

$$TF(t,d) = \frac{f(t,d)}{\sum t_d} \quad (1)$$

TF (*Term Frequency*) digunakan untuk menghitung seberapa sering suatu kata t muncul dalam dokumen d (Karim et al., 2022).

$$IDF(t) = \log \frac{N}{df(t)} \quad (2)$$

IDF (*Inverse Document Frequency*) digunakan untuk mengukur seberapa penting suatu kata dengan melihat jumlah dokumen yang mengandung kata tersebut (Karim et al., 2022).

$$TF\text{-}IDF(t,d) = TF(t,d) \times IDF(t) \quad (3)$$

INTI NUSA MANDIRI
DOI: https://doi.org/10.33480/inti.v21i1.8550

VOL. 21. NO. 1 AGUSTUS 2026
P-ISSN: 0216-6933 | E-ISSN: 2685-807X
Diterbitkan Oleh: LPPM Universitas Nusa Mandiri

TF-IDF merupakan hasil perkalian TF dan IDF untuk menentukan bobot akhir suatu kata dalam dokumen (Karim et al., 2022). Konfigurasi yang digunakan adalah max_features = 1.000, ngram_range (1,2), min_df = 2, max_df = 0,8. Setelah itu dataset dibagi dengan rasio 80:20 menggunakan *GroupShuffleSplit* untuk mencegah *data leakage. GroupShuffleSplit* mengelompokkan berdasarkan kolom *cleaned_text*. Setiap nilai unik pada kolom diperlakukan sebagai satu grup, sehingga seluruh data yang memiliki isi teks yang sama akan ditempatkan secara utuh pada salah satu subset data. Dengan demikian, tidak ada teks yang identik muncul pada kedua subset secara bersamaan untuk mencegah terjadinya data leakage, yaitu kondisi ketika informasi yang sama tersedia pada data pelatihan dan data pengujian sehingga dapat menghasilkan estimasi performa model yang terlalu optimistis dan terkesan "menghafal". Apabila terdapat beberapa data dengan nilai *cleaned_text* yang identik, seluruh data tersebut akan tetap berada dalam satu grup dan tidak akan dipisahkan selama proses pembagian data.

Setelah proses ekstraksi fitur TF-IDF, dilakukan seleksi fitur menggunakan uji Chi-Square (χ^2) untuk mengukur tingkat ketergantungan statistik antara setiap fitur (kata/n-gram) dengan label kelas sentimen, sehingga hanya fitur yang paling diskriminatif terhadap kelas yang dipertahankan (Sami'Un et al., 2024). Nilai Chi-Square untuk setiap fitur dihitung dengan persamaan:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \quad (4)$$

Dengan O_i adalah frekuensi observasi fitur pada suatu kelas dan E_i adalah frekuensi harapan apabila fitur tersebut tidak memiliki hubungan dengan kelas. Semakin tinggi nilai χ^2, semakin kuat hubungan fitur tersebut dengan kelas sentimen. Seleksi fitur dilakukan menggunakan SelectKBest dari pustaka scikit-learn dengan parameter k = 500, sehingga dari 1.000 fitur hasil TF-IDF, dipertahankan 500 fitur (50%) dengan skor Chi-Square tertinggi. Proses fitting Chi-Square dilakukan hanya pada 2.180 data latih untuk menghindari kebocoran informasi ke 531 data uji, kemudian fitur terpilih yang sama diterapkan pada transformasi data uji.

Model Klasifikasi

SVM dengan kernel linear bekerja dengan mencari *hyperplane* terbaik sebagai pemisah antar kelas dengan memaksimalkan lebar margin (Subrata et al., 2026). Random Forest merupakan algoritma *ensemble learning (bagging)* yang menggabungkan banyak *decision tree* untuk meningkatkan akurasi dan mengurangi *overfitting* dengan mengambil rata-rata seluruh *decision tree* (Cherliana et al., 2026). Kedua model menggunakan parameter *class_weight balanced* untuk menangani ketidakseimbangan kelas. Secara matematis SVM dan *Random Forest* masing-masing dapat dituliskan sebagai persamaan (5) dan (6).

$$\min_{\mathbf{w},b,\xi} \frac{1}{2} \| \mathbf{w} \|^2 + \sum_{i=1}^{n} C_i \ \xi_i \quad (5)$$

$$\hat{f} = \frac{1}{B} \sum_{b=1}^{B} f_b(x) \quad (6)$$

Evaluasi model menggunakan metrik akurasi, presisi, *recall*, *F1-Score* (Gulati et al., 2022), dan *cross validation* 5-fold (Fakhrezi & Rochim, 2023). Secara matematis evaluasi model dapat dituliskan sebagai berikut.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

$$Precision = \frac{TP}{TP+FP} \quad (8)$$

$$Recall = \frac{TP}{TP+FN} \quad (9)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (10)$$

HASIL DAN PEMBAHASAN

Distribusi Data dan Pelabelan Sentimen

Proses scraping awal diperoleh 2.724 tweet. Setelah dilakukan penyaringan data untuk menghapus tweet yang berasal dari akun resmi BMKG, maka jumlah data yang digunakan dalam penelitian ini adalah 2.711 tweet (13 baris dihapus). Sebelum data dianalisis, dilakukan beberapa tahapan dalam preprocessing data secara berurutan. Tabel 1 berikut ini adalah contoh hasil dari tahapan preprocessing data:

Tabel 1. Hasil Preprocessing Data

Tahapan	Output
Data Asli	@infoBMKG update gempa min tadi wilayah bandung subang
Case Folding	@infobmkg update gempa min tadi wilayah bandung subang

VOL. 21. NO. 1 AGUSTUS 2026
P-ISSN: 0216-6933 | E-ISSN: 2685-807X
Diterbitkan Oleh: LPPM Universitas Nusa Mandiri

INTI NUSA MANDIRI
DOI: https://doi.org/10.33480/inti.v21i1.8550

Tahapan	Output
Cleaning Data	update gempa min tadi wilayah bandung subang
Tokenizing	[update, gempa, min, tadi, wilayah, bandung, subang]
Normalisasi Slang	[update, gempa, admin, tadi, wilayah, bandung, subang]
Stopword Removal	[update, gempa, admin, wilayah, bandung, subang]
Stemming	update gempa admin wilayah bandung subang

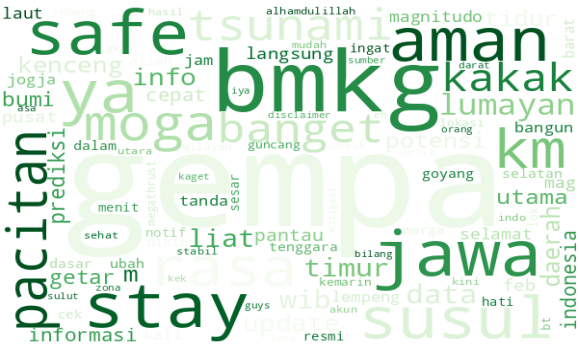
Sumber: (Hasil Penelitian, 2026)

Distribusi label sentimen hasil pelabelan otomatis disajikan pada Tabel 2. Kelas Netral mendominasi dengan 61,75%, kelas Negatif (24,71%), dan kelas Positif (13,54%). Rincian sentimen tiap kelas ditunjukkan pada *Wordcloud*. *Wordcloud* positif disajikan pada Gambar 2, *Wordcloud* negatif disajikan pada Gambar 3, dan *Wordcloud* netral disajikan pada Gambar 4.

Tabel 2. Distribusi Kelas Sentimen Hasil Pelabelan

Kelas	Jumlah	Persentase
Netral	1.674	61,75%
Negatif	670	24,71%
Positif	367	13,54%
Total	2.711	100%

Sumber: (Hasil Penelitian, 2026)



Sumber: (Hasil Penelitian, 2026)

Gambar 1. *Wordcloud* Kelas Positif

Sumber: (Hasil Penelitian, 2026)

Gambar 2. *Wordcloud* Kelas Negatif

Sumber: (Hasil Penelitian, 2026)

Gambar 3. Wordcloud Kelas Netral

Karakteristik leksikal masing-masing kelas sentimen dapat divisualisasikan melalui *word cloud* yang disajikan pada Gambar 2, 3, dan 4. Pada kelas Positif (Gambar 2), kata-kata dengan frekuensi tinggi seperti "aman", "safe", "stay", dan "bmkg" berdampingan dengan ekspresi rasa syukur dan doa seperti "alhamdulillah" dan "moga", mengindikasikan bahwa sentimen positif banyak muncul sebagai bentuk apresiasi atas kecepatan informasi sekaligus doa keselamatan pascagempa. Kelas Negatif (Gambar 3) didominasi kata-kata yang merepresentasikan respons emosional spontan seperti "takut", "panik", "trauma", "pusing", dan "susul", serta kata "tsunami" dan "waspada" yang mengindikasikan kekhawatiran terhadap potensi bencana susulan. Kemunculan kata "bmkg" pada kelas ini menunjukkan bahwa sentimen negatif lebih banyak diarahkan pada kondisi gempa itu sendiri, bukan terhadap institusi. Sementara itu, kelas Netral (Gambar 4) diwarnai istilah operasional dan faktual seperti "admin", "update", "barusan", "kenceng", dan "cek", yang mencerminkan tweet bersifat informatif dan berfokus pada verifikasi kejadian gempa tanpa muatan emosi yang dominan.

Penerapan TF-IDF menangkap informasi kontekstual dua kata yang sering membentuk ekspresi sentimen khas, memastikan hanya fitur yang muncul minimal dua kali yang disertakan, menyaring kata-kata yang terlalu umum. Hasil ekstraksi menghasilkan matriks fitur berukuran (2.180 × 1.000) untuk data latih dan (531 × 1.000) untuk data uji. Encoding label dilakukan dengan pemetaan: negatif→0, netral→1, dan positif→2. Kemudian dataset dibagi menjadi data latih sebesar 80% (2180) dan data uji sebesar 20% (531). Hasil seleksi fitur Chi-Square terhadap 1.000 fitur TF-IDF menghasilkan 500 fitur dengan tingkat asosiasi tertinggi terhadap kelas sentiment, Fitur "safe", "stay safe", "stay", dan "aman" menempati peringkat teratas dengan skor Chi-Square jauh di atas fitur lainnya, mengonfirmasi secara statistik bahwa istilah-istilah tersebut memang paling diskriminatif dalam membedakan kelas Positif, konsisten dengan

INTI NUSA MANDIRI
DOI: https://doi.org/10.33480/inti.v21i1.8550
VOL. 21. NO. 1 AGUSTUS 2026
P-ISSN: 0216-6933 | E-ISSN: 2685-807X
Diterbitkan Oleh: LPPM Universitas Nusa Mandiri

temuan kualitatif pada wordcloud (Gambar 2). Sebaliknya, fitur "takut" dan "pusing" menegaskan perannya sebagai penciri utama kelas Negatif. Menariknya, fitur lokasi seperti "jawa", "malang", dan "jawa timur" turut memiliki skor Chi-Square tinggi, mengindikasikan bahwa sentimen publik pada periode pengumpulan data kemungkinan dipengaruhi oleh peristiwa gempa spesifik di wilayah Malang/Jawa Timur, sehingga penyebutan lokasi tersebut berasosiasi kuat dengan kelas sentimen tertentu.

Perbandingan Performa Model

Hasil evaluasi menggunakan *cross validation* 5-fold kedua model pada 531 data uji disajikan pada Tabel 3.

Tabel 3. Perbandingan Performa SVM dan Random Forest

Model	Accuracy	Macro F1	CV F1	CV Std
SVM	0,8117	0,7443	0,7287	±0,0574
Random Forest	0,8192	0,7445	0,7228	±0,0670

Sumber: (Hasil Penelitian, 2026)

Random Forest sedikit lebih unggul dalam akurasi (81,92%) dibandingkan SVM (81,17%), namun kedua model hampir identik pada nilai Macro F1 (0,7445 vs 0,7443). Perbedaan yang sangat kecil ini menunjukkan kemampuan generalisasi yang setara (Indriani & Wiranata, 2024; Suasnawa, 2025). Untuk menguji signifikansi statistik dari selisih akurasi tersebut, dilakukan McNemar Test terhadap prediksi kedua model pada 531 data uji yang sama. Dari seluruh data uji, hanya ditemukan 5 kasus yang diklasifikasikan berbeda oleh kedua model (SVM benar namun Random Forest salah pada 2 kasus, dan sebaliknya pada 3 kasus). Uji McNemar (exact binomial, karena b+c < 25) menghasilkan statistik sebesar 2,00 dengan p-value sebesar 1,0000, jauh di atas taraf signifikansi $\alpha = 0,05$. Hasil ini menunjukkan bahwa selisih akurasi 0,75% antara Random Forest dan SVM tidak signifikan secara statistik, sehingga kedua model dapat dianggap memiliki performa klasifikasi yang setara pada data uji ini. Temuan ini memperkuat argumen bahwa pemilihan model akhir sebaiknya tidak semata-mata didasarkan pada akurasi, melainkan juga mempertimbangkan stabilitas *cross-validation*. Dari sisi stabilitas, SVM menunjukkan keunggulan lebih nyata dengan standar deviasi *cross-validation* (±0,0574) lebih kecil dibandingkan Random Forest (±0,0670). Nilai CV F1 *mean* yang mendekati F1 data uji pada kedua model mengindikasikan tidak adanya *overfitting* yang signifikan.

Evaluasi Per Kelas Sentimen

Tabel 4 dan 5 menyajikan *classification report* masing-masing model. Kelas Netral secara konsisten mencapai F1-Score tertinggi (0,88) pada kedua model, sejalan dengan proporsi sampel yang mendominasi dataset. Kelas Positif mencatat F1-Score terendah (0,62) akibat jumlah sampel yang paling sedikit (13,54%).

Tabel 4. *Classification Report* SVM

Kelas	Precision	Recall	F1
Negatif	0,75	0,71	0,73
Netral	0,88	0,89	0,88
Positif	0,6	0,64	0,62
Macro avg	0,75	0,74	0,74

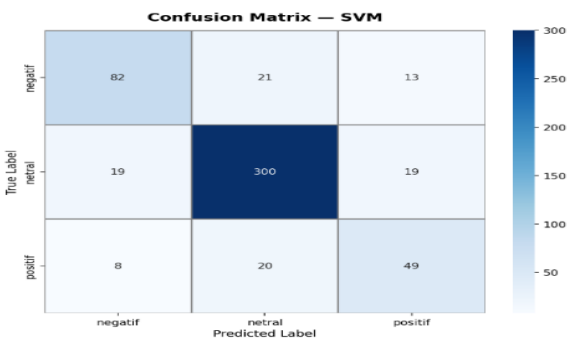
Sumber: (Hasil Penelitian, 2026)

Tabel 5. *Classification Report* Random Forest

Kelas	Precision	Recall	F1
Negatif	0,83	0,66	0,73
Netral	0,84	0,93	0,88
Positif	0,69	0,56	0,62
Macro avg	0,79	0,72	0,74

Sumber: (Hasil Penelitian, 2026)

Perbedaan pola yang mencolok terlihat pada kelas Negatif. Random Forest menunjukkan *precision* lebih tinggi (0,83 vs 0,75), namun *recall* lebih rendah (0,66 vs 0,71). Artinya, SVM memiliki sensitivitas yang lebih tinggi terhadap kelas negatif yang ditunjukkan oleh nilai *recall* sebesar 0,71 dibandingkan *Random Forest* sebesar 0,66 yang lebih diinginkan dalam konteks kebencanaan di mana deteksi kepanikan memiliki konsekuensi praktis lebih besar. Pola kesalahan utama pada *confusion matrix* kedua model adalah misklasifikasi *tweet* Positif dan Negatif ke kelas Netral, akibat ekspresi emosional ringan yang mendapat skor mendekati nol. Visualisasi *confusion matrix* dari model SVM ditunjukkan pada Gambar 5, sedangkan model Random Forest ditujukkan pada Gambar 6.

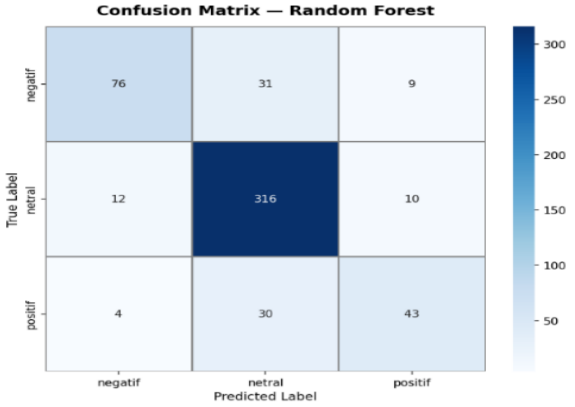


Sumber: (Hasil Penelitian, 2026)

Gambar 4. *Confusion Matrix* SVM

VOL. 21. NO. 1 AGUSTUS 2026
P-ISSN: 0216-6933 | E-ISSN: 2685-807X
Diterbitkan Oleh: LPPM Universitas Nusa Mandiri

INTI NUSA MANDIRI
DOI: https://doi.org/10.33480/inti.v21i1.8550



Sumber: (Hasil Penelitian, 2026)

Gambar 5. *Confusion Matrix* Random Forest

Pola kesalahan yang paling sering terjadi pada kedua model adalah misklasifikasi tweet Positif dan Negatif ke dalam kelas Netral. Dibandingkan SVM, diagonal netral pada Random Forest menunjukkan nilai yang lebih tinggi (316 vs 300), mengindikasikan kemampuan RF yang sedikit lebih baik dalam mengklasifikasikan kelas mayoritas. *Random Forest* lebih jarang salah mengklasifikasikan Positif sebagai Negatif (hanya 4 kasus) dibandingkan SVM (8 kasus), namun lebih banyak kehilangan tweet Positif ke dalam kelas Netral. Pola ini konsisten dengan nilai recall Positif yang lebih rendah pada Random Forest (0,56 vs 0,64 pada SVM).

Pembahasan

Dominasi sentimen Netral (61,75%) mencerminkan karakteristik komunikasi kebencanaan di Aplikasi X, yaitu masyarakat cenderung meneruskan atau mengonfirmasi data faktual tanpa disertai muatan emosional eksplisit, terutama bagi pengguna yang tidak berada pada wilayah terdampak. Temuan ini mengindikasikan bahwa format komunikasi berbasis data teknis BMKG telah diterima dengan baik oleh publik.

Proporsi sentimen Negatif yang cukup signifikan (24,71%) hanya merepresentasikan respons emosional masyarakat berupa rasa takut, kekhawatiran, dan kepanikan yang muncul akibat peristiwa gempa bumi, bukan semata-mata ditujukan sebagai bentuk penilaian negatif terhadap BMKG. Rendahnya sentimen Positif (13,54%) mengindikasikan bahwa ekspresi apresiasi bersifat implisit dan menggunakan bahasa informal yang sulit ditangkap oleh kamus InSet Lexicon. Pengembangan kamus sentimen berbahasa Indonesia untuk domain kebencanaan perlu untuk dikembangkan dan ditindaklanjuti.

Pertimbangan dalam memilih model akhir perlu mempertimbangkan konteks penggunaan. Apabila prioritas adalah stabilitas dan konsistensi prediksi, SVM lebih direkomendasikan berdasarkan nilai standar deviasi *cross-validation* yang lebih kecil (±0,0574). Untuk kepentingan monitoring *real-time* dalam konteks kebencanaan, SVM lebih direkomendasikan berdasarkan stabilitas *cross-validation* dan kemampuan *recall* sentimen negatif yang lebih baik (Rochmawati et al., 2025). Penelitian ini juga sejalan dengan upaya peningkatan literasi kebencanaan berbasis MKGI dalam memahami persepsi masyarakat terhadap penyebaran informasi kebencanaan melalui media digital sebagai bagian dari penguatan kesiapsiagaan dan komunikasi risiko bencana (Munawar et al., 2026).

KESIMPULAN

Penelitian ini menunjukkan bahwa SVM dan *Random Forest* mampu mengklasifikasikan sentimen *tweet* berbahasa Indonesia tentang informasi gempa bumi BMKG dengan akurasi di atas 81%. Sentimen Netral mendominasi (61,75%), diikuti Negatif (24,71%) dan Positif (13,54%), mencerminkan bahwa komunikasi berbasis data teknis BMKG diterima dengan baik namun masih memerlukan penguatan pesan ketenangan. SVM lebih direkomendasikan untuk pemantauan *real-time* berkat stabilitas *cross-validation* yang lebih baik (±0,0574) dan *recall* sentimen negatif yang lebih tinggi (0,71). Pengembangan kamus sentimen bahasa Indonesia yang lebih komprehensif untuk domain kebencanaan menjadi rekomendasi untuk penelitian selanjutnya.

Penelitian ini memiliki beberapa keterbatasan. Pertama, pelabelan sentimen dilakukan secara otomatis menggunakan pendekatan hybrid VADER dan InSet Lexicon tanpa validasi manual oleh anotator manusia. Hal ini merupakan batasan cakupan studi (*scope limitation*) yang disengaja demi efisiensi pemrosesan data berskala besar, namun berimplikasi pada keterbatasan model berbasis leksikon dalam menangkap konteks bahasa yang ambigu, sarkastik, atau bercampur bahasa daerah, sehingga dapat ditingkatkan melalui validasi manual atau pendekatan berbasis pembelajaran kontekstual pada penelitian mendatang. Kedua, data penelitian hanya bersumber dari satu akun resmi (@infoBMKG) dalam periode pengumpulan tertentu, sehingga hasil belum tentu dapat digeneralisasi untuk platform media sosial lain atau periode waktu yang berbeda. Ketiga, distribusi kelas data yang tidak seimbang (netral

mendominasi 61,75%) berpotensi memengaruhi kemampuan model dalam mengenali kelas minoritas seperti sentimen positif, meskipun telah ditangani dengan parameter *class_weight balanced*.

Pengembangan kamus sentimen bahasa Indonesia yang lebih komprehensif untuk domain kebencanaan, disertai validasi label secara manual dan perluasan sumber data, menjadi rekomendasi untuk penelitian selanjutnya. Selain itu, penelitian mendatang disarankan untuk mengeksplorasi model berbasis transformer seperti IndoBERT atau varian pralatih bahasa Indonesia lainnya, yang berpotensi menangkap konteks semantik, sarkasme, dan hubungan antarkata secara lebih baik dibandingkan pendekatan berbasis leksikon dan fitur TF-IDF yang digunakan pada penelitian ini, sehingga dapat mengatasi keterbatasan yang telah diuraikan sebelumnya.

REFERENSI

Aldamen, Y., & Hacimic, E. (2023). Positive determinism of Twitter usage development in crisis communication: Rescue and relief efforts after the 6 February 2023 earthquake in Türkiye. *Social Sciences*, *12*(8), 436. https://doi.org/10.3390/socsci12080436

Alita, D. (2024). Implementasi metode SVM pada sentimen analisis terhadap pemilihan presiden (Pilpres) 2024 di Twitter. *Jurnal Informatika*, *9*(2). https://doi.org/https://doi.org/10.30591/jpit.v9i2.6560

Cherliana, Sanjaya, M. R., Indah, D. R., & Kurniawan, D. (2026). Penerapan Random Forest dan XGBoost untuk Analisis Sentimen pada Ulasan Aplikasi M-Pajak. *Jurnal Algoritma*, *23*(1 SE-Artikel), 850–861. https://doi.org/10.33364/algoritma/v.23-1.3370

Fakhrezi, M. F., & Rochim, A. F. (2023). Comparison of sentiment analysis methods based on accuracy value case study: Twitter mentions of academic article. *Jurnal RESTI*. https://doi.org/https://doi.org/10.29207/resti.v7i1.4767

Fauziah, Y., Yuwono, B., & Aribowo, A. S. (2021). Lexicon based sentiment analysis in Indonesia languages: A systematic review. *Semantic Scholar*. https://pdfs.semanticscholar.org/32f3/8cdf16d2695748ca785730a232cb6575b7a7.pdf

Gulati, K., Kumar, S. S., Boddu, R. S. K., & Sarvakar, K. (2022). Comparative analysis of machine learning-based classification models using sentiment classification of tweets related to COVID-19 pandemic. *Materials Today: Proceedings*, *56*, 3505–3514. https://doi.org/10.1016/j.matpr.2021.03.584

Hadiprakoso, R. B., Setiawan, H., & Yasa, R. N. (2023). Text preprocessing for optimal accuracy in Indonesian sentiment analysis using a deep learning model with word embedding. *AIP Conference Proceedings*, *2680*(1). https://doi.org/10.1063/5.0126116

Handoyo, F. W. (2024). Enhancing disaster resilience: Insights from the Cianjur earthquake to improve Indonesia's risk financing strategies. *SAGE Open*, *14*(2). https://doi.org/10.1177/21582440241256777

Hayati, K., Wardani, A., & Misbah, N. A. (2023a). Digital communication of BMKG: Studies on conveyance of earthquake disaster information via social media Twitter @infoBMKG. *Jurnal Desentralisasi Dan Kebijakan Publik*, *4*(2), 1–15.

Hayati, K., Wardani, A., & Misbah, N. A. (2023b). Digital Communication of BMKG Non-Departmental Government Institutions: Studies on Conveyance of Earthquake Disaster Information via Social Media Twitter@ infoBMKG. *Jurnal Desentralisasi Dan Kebijakan Publik (JDKP)*, *4*(2). https://doi.org/10.30656/jdkp.v4i2.6928

Karim, M., Missen, M. M. S., Umer, M., Fida, A., & Nawaz, M. (2022). Comprehension of polarity of articles by citation sentiment analysis using TF-IDF and ML classifiers. *PeerJ Computer Science*, *8*, e1107. https://doi.org/10.7717/peerj-cs.1107

Khairani, D., & Setiawan, A. (2024). Enhancing understanding of public sentiment on Twitter using SVM and lexicon methods. *2024 3rd International Conference Proceedings*. https://ieeexplore.ieee.org/document/10701213/

Khoirunnisa, A., & Ningrum, N. K. (2026). Comparison of Naïve Bayes, Random Forest, and SVM algorithm performance in analyzing sentiment regarding the Aceh floods on platform X. *Journal of Applied Informatics and Computing*, *10*(2). https://doi.org/https://doi.org/10.30871/jaic.v10i2.12250

Munawar, Haryanto, Y. D., Abigael, F. D., Muftareza, A. D., Muthahhari, I., Mardiyansyah, A., & Sinambela, M. (2026). Peningkatan Literasi Kebencanaan untuk Optimalisasi Informasi Peringatan Dini Bencana Geo-Hidrometeorologi. *JPM: Jurnal Pengabdian Masyarakat*, *6*(3), 521–530.

https://doi.org/10.47065/jpm.v6i3.2629

Nugraheni, E., Haekal, F. I., & Arisal, A. (2024). Optimizing Indonesian tweet preprocessing on halal domain. *IEEE International Conference Proceedings*. https://ieeexplore.ieee.org/document/10732128/

NURSUKUWAGUS, A., DEWI, D. A., ABDILLAH, H. F., NAJMI, R. A. N., & ZULKARNAEN, R. R. (2026). SENTIMENT ANALYSIS WITH CLASSIFICATION LEARNING FOR COMMENTS ON THE X (TWITTER) APPLICATION. *Journal of Engineering Science and Technology*, *21*(2), 649–661.

Ramadhani, A., Aryanti, R., & Agustiani, S. (2026). Optimasi Model Machine Learning Menggunakan Teknik SMOTE pada Analisis Sentimen Pengguna RedBus. *Journal of Artificial Intelligence and Technology Information (JAITI)*, *4*(1 SE-Articles), 1–12. https://doi.org/10.58602/jaiti.v4i1.182

Rianto, R., Mutiara, A. B., Wibowo, E. P., & Santosa, P. I. (2021). Improving the accuracy of text classification using stemming method, a case of non-formal Indonesian conversation. *Journal of Big Data*, *8*(1). https://doi.org/10.1186/s40537-021-00413-1

Rochmawati, N., Zyen, A. K., & Widiastuti, N. A. (2025). Comparison of support vector machine (SVM) and random forest algorithms in the analysis of social media X user sentiment towards the TNI Bill. *Journal of Applied Informatics and Computing*, *9*(5), 2854–2860. https://doi.org/https://doi.org/10.30871/jaic.v9i5.10883

Rudyawan, A. (2023). Cianjur earthquake — roles of social media and the distribution of sciences. *IOP Conference Series: Earth and Environmental Science*, *1245*(1), 12016. https://doi.org/10.1088/1755-1315/1245/1/012016

Sami'Un, D. C., Sugiharto, A., & Jie, F. (2024). CHI SQUARE FEATURE SELECTION FOR IMPROVING SENTIMENT ANALYSIS OF NEWS DATA PRIVACY TREATS . In *Journal of Theoretical and Applied Information Technology* (Vol. 102, Issue 18, pp. 6601–6610). https://www.scopus.com/pages/publications/85206448265

Subrata, B. K., Jumaryadi, Y., Sulistyo, F. P., & Rahayu, S. (2026). Analisis Sentimen Masyarakat terhadap Profesionalisme Generasi Z di Dunia Kerja Menggunakan Support Vector Machine (SVM). *Journal of Artificial Intelligence and Technology Information (JAITI)*, *4*(2 SE-Articles), 314–329. https://doi.org/10.58602/jaiti.v4i2.279

Tan, K. L., Lee, C. P., & Lim, K. M. (2023). A survey of sentiment analysis: Approaches, datasets, and future research. *Applied Sciences*, *13*(7), 4550. https://doi.org/10.3390/app13074550

Tiwari, D., Bhati, B. S., Garg, P., & Babbar, N. (2026). A novel lexicon dictionary and CNN-LSTM employed hybrid approach for sentiment detection of COVID-19 vaccines. *Scientific Reports*, *16*(1), 22615. https://doi.org/10.1038/s41598-026-51418-w

Zulfakriza, Z. (2024). Seismic source analysis of the destructive earthquake November 21, 2022, Mw 5.6 Cianjur (Indonesia) from relocated aftershock. *Scientific Reports*, *14*(1), 12142. https://doi.org/10.1038/s41598-024-60408-9