



Analisis Sentimen Konten *TikTok* Terkait Bencana Alam Aceh dan Sumatera Utara Menggunakan Metode *Naïve Bayes*

Aninditawidagda Pandam Sudaryanto¹, Ila Khafia Wafda²

¹ Program Studi Informatika, Fakultas Teknik, Universitas Islam Batik Surakarta

² Program Studi Komunikasi dan Penyiaran Islam, Fakultas Komunikasi dan Penyiaran Islam, Sekolah Tinggi Agama Islam Bina Muwahhidin Boyolali

E-mail: ¹anindita@uib.ac.id^{*}, ²ilakwafda@gmail.com

*Corresponding Author

ABSTRACT

Media sosial telah menjadi sarana utama bagi masyarakat untuk menyampaikan opini dan respons terhadap berbagai peristiwa, termasuk bencana alam. *TikTok* sebagai salah satu platform media sosial dengan jumlah pengguna yang besar memungkinkan pengguna untuk memberikan komentar secara langsung terhadap konten yang berkaitan dengan bencana alam. Penelitian ini bertujuan untuk menganalisis sentimen komentar *TikTok* terkait bencana alam yang terjadi di Aceh dan Sumatera Utara. Data yang digunakan dalam penelitian ini berjumlah 520 komentar yang dikumpulkan pada periode November hingga Desember 2025 menggunakan teknik crawling. Tahapan preprocessing yang dilakukan meliputi case folding, tokenisasi, penghapusan stopword, dan stemming. Proses klasifikasi sentimen dilakukan menggunakan metode *Naïve Bayes* dengan ekstraksi fitur Term Frequency Inverse Document Frequency (TF-IDF). Dataset dibagi menjadi data latih dan data uji dengan perbandingan 80% dan 20% serta dievaluasi menggunakan metode 5-fold cross validation. Hasil pengujian menunjukkan bahwa metode *Naïve Bayes* menghasilkan akurasi rata-rata sebesar 86%, dengan akurasi tertinggi sebesar 87% pada fold ke-3, sehingga dapat disimpulkan bahwa metode *Naïve Bayes* cukup efektif dalam melakukan analisis sentimen komentar *TikTok* terkait bencana alam.

Keywords: Analisis sentimen, *TikTok*, bencana alam, *Naïve Bayes*, TF-IDF.



Copyright © 2025 The Author(s)
This is an open access article under the CC BY-SA license.

INTRODUCTION

Bencana alam merupakan peristiwa atau rangkaian peristiwa yang terjadi akibat proses alam, seperti gempa bumi, tsunami, letusan gunung berapi, banjir, kekeringan, angin topan, dan tanah longsor. Fenomena ini dipahami sebagai bagian dari mekanisme alam dalam menjaga keseimbangan ekosistem yang terganggu, baik akibat aktivitas manusia maupun proses alamiah (Sulthan et al., 2021). Di Indonesia, bencana alam menjadi ancaman serius karena kondisi geografisnya yang berada pada pertemuan tiga lempeng tektonik utama dunia, yaitu Eurasia, Indo-Australia, dan Pasifik, sehingga tingkat kerawanan bencana tergolong tinggi (Maulana et al., 2024).

Dalam penanggulangan bencana, Undang-Undang Republik Indonesia Nomor 24 Tahun 2007 menegaskan bahwa Badan Penanggulangan Bencana Daerah (BPBD) bertanggung jawab dalam upaya pencegahan, penanganan darurat, rehabilitasi, dan rekonstruksi secara adil dan setara. Namun, sifat bencana yang tidak terduga sering menyulitkan proses penanganan,

terutama pada fase tanggap darurat ketika kebutuhan dasar dan layanan medis harus segera dipenuhi (Nofiyanti et al., 2021). Pada era digital saat ini, penyebaran informasi terkait bencana berlangsung sangat cepat melalui internet, khususnya melalui media sosial.

Perkembangan teknologi informasi dan komunikasi telah mengubah cara masyarakat menyampaikan dan memperoleh informasi kebencanaan. Media sosial, khususnya *TikTok*, menjadi salah satu platform yang banyak dimanfaatkan untuk menyebarkan informasi mengenai jenis bencana, waktu dan lokasi kejadian, dampak yang ditimbulkan, serta kebutuhan korban secara luas dan cepat, bahkan hingga tingkat global (Sulthan et al., 2021). *TikTok* juga menyediakan kolom komentar yang memungkinkan interaksi dan penyampaian opini antar pengguna, sehingga membentuk ruang diskusi publik yang dinamis (Indriyani et al., 2023).

Indonesia tercatat sebagai negara dengan jumlah pengguna *TikTok* terbesar di dunia pada tahun 2025, dengan sekitar 194,37 juta pengguna aktif per Juli 2025. Kondisi ini menunjukkan peran strategis *TikTok* dalam membentuk pola komunikasi, penyebaran informasi, dan opini publik di Indonesia (Nouvan, 2025). Data dari *TikTok* dapat dikumpulkan melalui proses crawling yang mencakup informasi pengguna, isi komentar, dan atribut pendukung lainnya. Analisis sentimen dalam komentar *TikTok* umumnya diklasifikasikan ke dalam kategori positif, negatif, dan netral, namun karakter komentar yang singkat, tidak baku, dan ambigu menuntut penerapan metode analisis yang tepat agar identifikasi analisis sentimen dapat dilakukan secara akurat (Putra et al., 2023).

Dalam situasi bencana, *TikTok* tidak hanya berfungsi sebagai media pertukaran informasi, tetapi juga berperan dalam meningkatkan kesadaran publik, mempercepat respons kebencanaan, serta menjadi ruang bagi masyarakat untuk menyampaikan empati, kritik, dan kepedulian terhadap peristiwa bencana (Santoso et al., 2023). Peristiwa bencana alam di wilayah Kabupaten Aceh dan Sumatera Utara mendorong munculnya berbagai opini publik yang terekam dalam komentar *TikTok*. Namun, besarnya volume data opini tersebut menjadikan pemantauan manual tidak efektif, sehingga diperlukan pendekatan otomatis melalui teknik text mining dan analisis sentimen (Zumarniansyah et al., 2020). Analisis sentimen merupakan proses pengolahan data tekstual secara otomatis untuk memperoleh informasi sentimen yang terkandung dalam opini masyarakat (Jelita et al., 2025).

Penelitian ini memanfaatkan platform *TikTok* untuk menganalisis sentimen komentar menggunakan hashtag #banjiraceh, #bencanaaceh, #banjirsumut, #gempaaceh, dan #bencanasumut. Data yang dianalisis berupa opini masyarakat terkait peristiwa bencana alam dan respons terhadap penanganannya, yang diklasifikasikan ke dalam dua kategori sentimen, yaitu positif dan negatif. Proses klasifikasi dilakukan menggunakan algoritma *Naïve Bayes* yang didasarkan pada Teorema *Bayes* dan dikenal memiliki efisiensi komputasi serta akurasi yang baik dalam analisis sentimen (Maksun et al., 2021). Penelitian sebelumnya menunjukkan bahwa metode *Naïve Bayes* menghasilkan tingkat akurasi lebih tinggi dibandingkan metode *K-Nearest Neighbor* dan *Decision Tree* (Voutama et al., 2024).

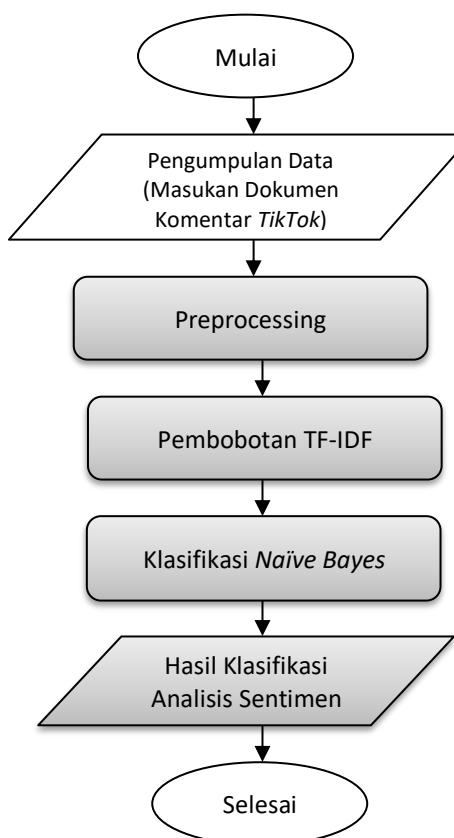
Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menganalisis sentimen komentar *TikTok* terkait bencana alam di Kabupaten Aceh dan Sumatera Utara menggunakan metode *Naïve Bayes*. Hasil penelitian diharapkan dapat memberikan gambaran respons dan persepsi masyarakat terhadap peristiwa bencana serta menjadi bahan pertimbangan bagi pihak terkait dalam mendukung penanggulangan bencana yang lebih responsif dan berbasis data.

METHODS

Penelitian ini menerapkan analisis sentimen menggunakan algoritma *Naïve Bayes* untuk mengklasifikasikan opini masyarakat terhadap konten dan komentar *TikTok* yang membahas bencana alam di wilayah Aceh dan Sumatera Utara. Alur kerja sistem digambarkan dalam bentuk diagram alur (flowchart) yang menunjukkan tahapan penelitian secara menyeluruh. Proses dimulai dari pengumpulan data konten dan komentar *TikTok* sebagai data masukan, dilanjutkan

dengan tahap text preprocessing. Data hasil preprocessing kemudian direpresentasikan ke dalam bentuk fitur numerik menggunakan pembobotan Term Frequency Inverse Document Frequency (TF-IDF), yang selanjutnya digunakan sebagai masukan dalam proses klasifikasi sentimen menggunakan algoritma *Naïve Bayes*. Keluaran dari sistem berupa klasifikasi sentimen komentar *TikTok* ke dalam kategori positif dan negatif.

Data penelitian diperoleh dari platform media sosial *TikTok* yang dipilih karena tingginya tingkat partisipasi pengguna serta keberagaman opini publik yang disampaikan melalui konten dan kolom komentar. Pengumpulan data dilakukan dengan teknik crawling terhadap konten dan komentar *TikTok* yang relevan dengan topik bencana alam di Aceh dan Sumatera Utara dalam rentang waktu tertentu untuk menjaga konsistensi dan relevansi data. Seluruh data disimpan dalam format terstruktur agar siap diproses pada tahap selanjutnya. Data difokuskan pada komentar berbahasa Indonesia guna memastikan kesesuaian dengan konteks sosial dan linguistik penelitian, sehingga model analisis sentimen yang dibangun mampu merepresentasikan opini masyarakat Indonesia secara lebih akurat dan kontekstual (Fatra et al., 2025). Diagram alur penelitian ditunjukkan pada Gambar 1.



Gambar 1 Diagram Alur Penelitian

3.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini berupa kumpulan komentar dari berbagai konten *TikTok* yang relevan dengan topik bencana alam di wilayah Aceh dan Sumatera Utara, dengan rentang waktu pengambilan data pada November hingga Desember 2025 dan menggunakan bahasa Indonesia (Dina et al., 2025). Pemilihan periode tersebut selaras dengan kondisi Indonesia sebagai negara dengan jumlah pengguna *TikTok* terbesar pada tahun 2025.

Pengumpulan data dilakukan dengan menerapkan teknik crawling dan scraping menggunakan bahasa pemrograman Python. Python dipilih karena bersifat interpreted dan didukung oleh berbagai pustaka seperti Pandas dan Scikit learn yang memudahkan proses pengolahan data serta penerapan algoritma pembelajaran mesin (Dinar et al., 2023). Proses crawling dilakukan melalui *TikTok API* dengan memanfaatkan kata kunci dan tagar yang relevan, antara lain *#banjiraceh*, *#bencanaaceh*, *#banjirsumut*, *#gempaaceh*, dan *#bencanasumut*.

Data yang berhasil dikumpulkan kemudian diseleksi dan digunakan sebagai dataset penelitian dengan total 520 data, yang terdiri atas 260 data berlabel sentimen positif dan 260 data berlabel sentimen negatif. Dataset selanjutnya dibagi menjadi data latih (training data) dan data uji (testing data) untuk keperluan proses klasifikasi sentimen. Atribut data yang digunakan meliputi username, isi konten atau komentar *TikTok*, serta label sentimen.

Pengumpulan sekaligus pelabelan data dilakukan menggunakan tool khusus berbasis web scraping dan data processing pada platform Apify (<https://console.apify.com>). Proses pelabelan sentimen dilakukan secara otomatis oleh sistem pada platform tersebut dengan menggunakan aturan (rule-based labeling) dan pemrosesan teks yang telah ditentukan sebelumnya. Setiap komentar diklasifikasikan ke dalam sentimen positif atau negatif berdasarkan kata kunci, pola bahasa, dan konteks kalimat yang teridentifikasi oleh sistem. Pendekatan ini dipilih untuk menjaga konsistensi pelabelan data serta meminimalkan subjektivitas yang dapat muncul pada pelabelan manual.

3.2. Preprocessing

Tahap preprocessing merupakan langkah awal dalam pengolahan teks komentar *TikTok* yang bertujuan untuk membersihkan data dari karakter yang tidak relevan serta kata-kata yang tidak memiliki makna penting, sehingga informasi yang dianalisis lebih berfokus pada unsur sentimen. Proses ini juga mencakup penyeragaman bentuk kata dengan menghilangkan variasi imbuhan dan elemen yang tidak diperlukan. Tahapan preprocessing meliputi:

- a). Case folding, yaitu mengubah seluruh huruf menjadi huruf kecil;
- b). Cleaning, yaitu menghapus simbol, angka, dan emotikon;
- c). Tokenizing, yaitu memecah teks menjadi unit kata;
- d). Stopword removal, yaitu menghapus kata umum yang tidak bermakna sentimen; dan
- e). Stemming/Lemmatization, yaitu mengubah kata ke bentuk dasarnya agar data teks lebih bersih dan siap diproses oleh algoritma klasifikasi (Lubis et al., 2026).

3.3. Pembobotan TF-IDF

Pada tahap ini, teks hasil preprocessing diubah menjadi representasi numerik menggunakan metode Term Frequency Inverse Document Frequency (TF-IDF). Term Frequency (TF) digunakan untuk mengukur frekuensi kemunculan suatu kata dalam dokumen, sedangkan Inverse Document Frequency (IDF) berfungsi sebagai pembobot yang berbanding terbalik dengan frekuensi kemunculan kata pada keseluruhan dokumen. Kata yang jarang muncul akan memiliki bobot yang lebih tinggi dan dianggap lebih informatif dalam proses klasifikasi sentimen (Ainunnisa et al., 2023). Metode TF-IDF dipilih karena mampu merepresentasikan karakteristik teks secara lebih efektif dibandingkan pembobotan berbasis frekuensi semata (Lestari et al., 2025).

Dalam penelitian analisis sentimen komentar *TikTok* terkait bencana alam di Aceh dan Sumatera Utara dengan metode *Naïve Bayes*, pembobotan kata dilakukan menggunakan Term Frequency Inverse Document Frequency (TF-IDF). Term Frequency (TF) digunakan untuk mengukur seberapa sering suatu kata muncul dalam sebuah dokumen, di mana frekuensi kemunculan yang tinggi menunjukkan kata tersebut bersifat umum dan kurang merepresentasikan makna spesifik. Sementara itu, Inverse Document Frequency (IDF) berfungsi

sebagai pembobot yang berbanding terbalik dengan frekuensi kemunculan kata di seluruh dokumen, sehingga kata yang jarang muncul akan memiliki bobot yang lebih besar dan dianggap lebih informatif dalam proses klasifikasi sentimen.

3.4. Klasifikasi Naïve Bayes

Naïve Bayes merupakan metode klasifikasi yang termasuk dalam supervised learning dan bekerja berdasarkan penerapan Teorema Bayes. Algoritma ini memerlukan data latih yang telah diberi label sebagai dasar pembelajaran model. Dalam penelitian ini, nilai TF-IDF digunakan sebagai fitur numerik utama dalam proses klasifikasi sentimen komentar *TikTok* ke dalam kelas positif dan negatif (Prasetyo et al., 2023).

Naïve Bayes dikenal sebagai algoritma probabilistik yang sederhana namun efektif dalam pengolahan data teks, dengan asumsi bahwa setiap fitur bersifat independen satu sama lain (Tahir et al., 2024). Algoritma ini menghitung probabilitas suatu komentar termasuk ke dalam kelas sentimen tertentu berdasarkan pola yang dipelajari dari data latih. Selanjutnya, nilai probabilitas tersebut digunakan untuk memprediksi kelas sentimen setiap konten atau komentar secara objektif dan sistematis, sehingga hasil analisis sentimen dapat menggambarkan respons dan opini masyarakat terhadap bencana alam di Aceh dan Sumatera Utara yang disampaikan melalui platform *TikTok*.

PEMBAHASAN DAN HASIL

Pembahasan

4.1. Pengujian Akurasi Penelitian

Pengujian akurasi dalam penelitian ini bertujuan untuk mengevaluasi kinerja sistem analisis sentimen yang dibangun menggunakan metode *Naïve Bayes* dalam mengklasifikasikan komentar *TikTok* terkait bencana alam di Aceh dan Sumatera Utara. Evaluasi dilakukan dengan membandingkan hasil klasifikasi sistem terhadap label sentimen pada data uji yang telah ditentukan sebelumnya. Dataset terdiri atas dua kelas sentimen, yaitu positif dan negatif.

Metode evaluasi yang digunakan adalah confusion matrix, yang mencakup empat parameter utama, yaitu *True Positive (TP)*, *True Negative (TN)*, *False Positive (FP)*, dan *False Negative (FN)*. Berdasarkan hasil pengujian, sistem menunjukkan kemampuan yang cukup baik dalam mengklasifikasikan sentimen komentar *TikTok*, yang ditunjukkan oleh nilai akurasi yang relatif tinggi dan stabil pada setiap skenario pengujian.

4.2. Skenario Pengujian Akurasi Sistem

Penelitian ini menggunakan sebanyak 520 komentar *TikTok* yang telah melalui proses preprocessing dan pelabelan sentimen. Pengujian sistem dilakukan menggunakan metode 5-fold cross validation, dengan pembagian data sebesar 80% sebagai data latih dan 20% sebagai data uji pada setiap fold. Skema ini memungkinkan setiap data digunakan secara bergantian sebagai data latih dan data uji, sehingga hasil evaluasi menjadi lebih objektif dan representatif.

Evaluasi performa sistem dilakukan menggunakan confusion matrix untuk menghitung nilai akurasi (accuracy), presisi (precision), recall, dan f-measure. Sebelum proses klasifikasi, data teks diproses melalui tahap preprocessing dan pembobotan fitur menggunakan TF-IDF, kemudian diklasifikasikan menggunakan algoritma *Naïve Bayes*.

Pengujian dilakukan sebanyak lima kali sesuai dengan jumlah fold yang digunakan. Hasil pengujian pada setiap fold kemudian dirangkum untuk memperoleh nilai rata-rata performa sistem. Rincian hasil pengujian 5-fold validation ditunjukkan pada Tabel 1.

Tabel 1 Pengujian Menggunakan 5-Fold

Klasifikasi Sentimen				
Pengujian menggunakan 5-fold	Accuracy	Precision	Recall	F-measure
K = 1	0,85	1,00	0,20	0,33
K = 2	0,84	1,00	0,15	0,26
K = 3	0,87	1,00	0,30	0,46
K = 4	0,84	1,00	0,11	0,19
K = 5	0,85	1,00	0,32	0,48
Rata-Rata	0,86	1,00	0,22	0,34

Confusion matrix digunakan untuk mengevaluasi hasil klasifikasi, di mana prediksi yang sesuai dengan kondisi sebenarnya disebut True Positive atau True Negative, sedangkan prediksi yang tidak sesuai dikategorikan sebagai False Positive atau False Negative (Hilmi et al., 2024).

Hasil pengujian menunjukkan bahwa sistem memperoleh akurasi rata-rata sebesar 86%, dengan variasi nilai antar fold yang relatif kecil. Hal ini mengindikasikan bahwa model memiliki performa yang stabil. Nilai precision yang tinggi menunjukkan bahwa sistem mampu mengklasifikasikan sentimen dengan tingkat kesalahan prediksi yang rendah. Sementara itu, nilai recall dan f-measure yang relatif lebih rendah dipengaruhi oleh karakteristik komentar *TikTok* yang singkat, tidak baku, dan mengandung ambiguitas makna, sehingga tidak seluruh ekspresi sentimen dapat teridentifikasi secara optimal oleh sistem.

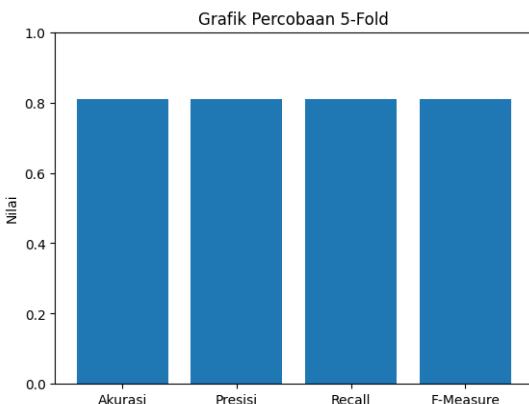
4.3. Analisis dan Hasil Pengujian Menggunakan 5-Fold

Analisis sentimen terhadap konten dan komentar *TikTok* terkait bencana alam di Aceh dan Sumatera Utara dilakukan menggunakan metode *Naïve Bayes* dengan pendekatan 5-fold cross validation untuk mengevaluasi kinerja model secara objektif. Data sebanyak 520 komentar diproses melalui tahapan preprocessing teks, pembobotan fitur menggunakan TF-IDF, dan klasifikasi sentimen menggunakan metode *Naïve Bayes* dengan dua kelas sentimen, yaitu positif dan negatif. Pada setiap fold, data dibagi dengan proporsi 80% sebagai data latih dan 20% sebagai data uji, sehingga setiap data memiliki peluang yang sama untuk diuji.

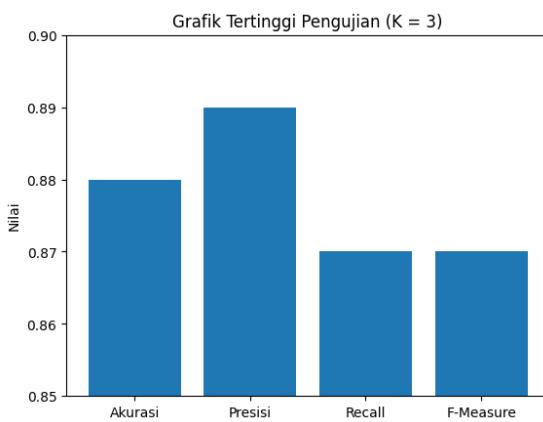
Hasil pengujian menunjukkan bahwa metode *Naïve Bayes* mampu memberikan performa klasifikasi yang cukup baik dengan nilai akurasi rata-rata sebesar 86%, presisi rata-rata 100%, recall sebesar 22%, dan F-measure sebesar 34%. Nilai akurasi tertinggi diperoleh pada fold ke-3 (K=3) dengan akurasi sebesar 87%, yang mengindikasikan bahwa pada skenario tersebut distribusi data latih dan data uji mampu merepresentasikan pola sentimen secara lebih optimal. Tingginya nilai presisi menunjukkan bahwa model memiliki tingkat ketepatan yang sangat baik dalam mengklasifikasikan sentimen positif, meskipun nilai recall dan F-measure relatif lebih rendah. Hal ini dipengaruhi oleh karakteristik komentar *TikTok* yang cenderung singkat, informal, serta mengandung variasi bahasa, sehingga tidak seluruh ekspresi sentimen dapat teridentifikasi secara maksimal oleh sistem.

Secara keseluruhan, hasil pengujian menggunakan 5-fold cross validation membuktikan bahwa metode *Naïve Bayes* cukup efektif dalam mengklasifikasikan sentimen komentar *TikTok* terkait bencana alam di Aceh dan Sumatera Utara. Model mampu mengenali pola sentimen positif dan negatif secara konsisten, serta menunjukkan stabilitas kinerja antar fold, sehingga

metode ini layak digunakan sebagai pendekatan analisis sentimen pada data media sosial dengan karakteristik serupa. Visualisasi hasil rata-rata pengujian 5-fold ditampilkan pada Gambar 2, sedangkan hasil pengujian terbaik pada fold ke-3 disajikan pada Gambar 3.



Gambar 2 Grafik Percobaan 5-fold



Gambar 3 Grafik Tertinggi Pengujian

Gambar 2 dan Gambar 3 menyajikan hasil evaluasi kinerja model klasifikasi sentimen yang dibangun. Gambar 2 menampilkan grafik hasil percobaan menggunakan metode 5-fold cross validation, yang menunjukkan nilai rata-rata akurasi, presisi, recall, dan F-measure. Hasil tersebut memperlihatkan bahwa metode *Naïve Bayes* mampu menghasilkan nilai akurasi dan presisi yang cukup tinggi dalam mengklasifikasikan sentimen komentar *TikTok*, sehingga model dinilai efektif dalam mengidentifikasi kecenderungan opini masyarakat terkait peristiwa bencana alam. Namun, nilai recall dan F-measure yang relatif lebih rendah mengindikasikan bahwa masih terdapat keterbatasan model dalam menangkap seluruh variasi sentimen, terutama karena karakteristik komentar *TikTok* yang singkat, tidak baku, dan cenderung ambigu.

Selanjutnya, Gambar 3 menunjukkan hasil pengujian tertinggi yang diperoleh pada fold ke-3 ($K = 3$). Pada grafik tersebut terlihat bahwa nilai presisi mencapai tingkat paling tinggi, diikuti oleh nilai akurasi yang juga cukup baik, sementara nilai recall dan F-measure mengalami peningkatan dibandingkan fold lainnya. Hal ini menunjukkan bahwa pada pembagian data latih dan data uji pada fold ke-3, model *Naïve Bayes* mampu mengenali pola sentimen komentar *TikTok* secara lebih optimal. Secara keseluruhan, visualisasi pada kedua gambar tersebut menegaskan bahwa metode *Naïve Bayes* memiliki kinerja yang stabil dan cukup andal dalam menganalisis sentimen konten *TikTok* terkait bencana alam di wilayah Aceh dan Sumatera Utara.

KESIMPULAN

Berdasarkan hasil pengujian menggunakan metode 5-fold cross validation, sistem analisis sentimen komentar *TikTok* terkait bencana alam di Aceh dan Sumatera Utara yang dibangun dengan algoritma *Naïve Bayes* menunjukkan kinerja yang cukup baik dan stabil. Dari total 520 data komentar, sistem memperoleh nilai akurasi rata-rata sebesar 86%, yang menunjukkan bahwa model mampu mengklasifikasikan sentimen komentar dengan tingkat ketepatan yang tinggi. Nilai presisi rata-rata sebesar 1,00 mengindikasikan bahwa model sangat baik dalam memprediksi komentar yang termasuk ke dalam kelas sentimen tertentu, khususnya dalam meminimalkan kesalahan prediksi positif. Sementara itu, nilai recall rata-rata sebesar 0,22 dan F-measure sebesar 0,34 menunjukkan bahwa masih terdapat keterbatasan model dalam menangkap seluruh variasi sentimen yang ada, terutama disebabkan oleh karakteristik komentar *TikTok* yang singkat, tidak baku, dan cenderung ambigu. Hasil terbaik diperoleh pada pengujian fold ke-3 ($K=3$) dengan nilai akurasi sebesar 87%, recall 0,30, dan F-measure 0,46, yang menunjukkan bahwa pada pembagian data tersebut model mampu mengenali pola sentimen secara lebih optimal dibandingkan fold lainnya. Secara keseluruhan, hasil ini menegaskan bahwa metode *Naïve Bayes* efektif digunakan untuk analisis sentimen komentar *TikTok* terkait bencana alam, meskipun masih diperlukan pengembangan lebih lanjut untuk meningkatkan kemampuan model dalam mengenali seluruh variasi sentimen secara lebih menyeluruh.

DAFTAR PUSTAKA

- Ainunnisa et al. (2023). Analisis Sentimen Aplikasi Tiktok dengan Metode Support Vector Machine (SVM), Logistic Regression dan Naïve Bayes. *Jurnal Teknologi Sistem Informasi Dan Aplikasi*, 6(3), 423–430. <https://doi.org/10.32493/jtsi.v6i3.31076>
- Dina et al. (2025). Analisis Sentimen Terhadap Komentar Pada Media Sosial Tiktok Yang Berpotensi Menyebabkan Depresi. *Jurnal Ilmiah Computing Insight*, 7(1), 1–9.
- Dinar et al. (2023). Analisis Sentimen Pada Pengguna Twitter Terhadap Program Kampus Merdeka Menggunakan Naïve Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 755–760.
- Fatra et al. (2025). Analisis Sentimen Media Sosial Menggunakan Metode Naïve Bayes. *Jurnal Culture Education and Technology Research (Cetera)*, 2(2), 88–98.
- Hilmi et al. (2024). Analisis Sentimen Terhadap Aplikasi Tiktok Dari Ulasan Pada Google Playstore Menggunakan Metode Naïve Bayes. *SMATIKA : STIKI Informatika Jurnal*, 14(1), 146–156.
- Indriyani et al. (2023). Analisis sentimen aplikasi tiktok menggunakan algoritma naïve bayes dan support vector machine Tiktok application sentiment analysis using naïve bayes algorithm and support vector machine. *TEKNOSAINS: Jurnal Sains, Teknologi Dan Informatika*, 10(2), 176–184. <https://doi.org/10.37373/tekno.v10i2.419>
- Jelita et al. (2025). Penerapan Algoritma Naïve Bayes Dalam Analisis Sentimen Masyarakat Terhadap STMIK Widya Cipta Dharma. *Bulletin of Information Technology (BIT)*, 6(2), 148–160. <https://doi.org/10.47065/bit.v5i2.2029>
- Lestari et al. (2025). Evaluation of TF-IDF Extraction Techniques in Sentiment Analysis of Indonesian-Language Marketplaces Using SVM , Logistic Regression , and Naive Bayes. *Journal of Computer Science and Application*, 8(1), 1–9.
- Lubis et al. (2026). Klasifikasi Sentiment Analysis Terhadap Usulan KB Vasektomi Syarat Penerima BANSOS dengan Metode Naive Bayes. *Jurnal Komputer Teknologi Informasi Sistem Komputer*, 4(3), 1539–1548.
- Maksun et al. (2021). Analisis Sentimen pada Twitter Bencana Alam di Kalimantan Selatan menggunakan Metode Naïve Bayes. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 5(12), 5614–5621.

- Maulana et al. (2024). Mitigasi Bencana di Indonesia. *Comserva Jurnal Penelitian Dan Pengabdian Masyarakat*, 3(10), 3996–4012.
<https://doi.org/10.59141/comserva.v3i10.1213>
- Nofiyanti et al. (2021). Analisis Sentimen terhadap Penanggulangan Bencana di Indonesia. *Jurnal Ilmiah Sinus (JIS)*, 19(2), 17–26.
- Nouvan. (2025). Indonesia Jadi Negara Pengguna TikTok Terbanyak di Dunia 2025. Retrieved from dataloka website: <https://dataloka.id/humaniora/4424/indonesia-jadi-negara-pengguna-tiktok-terbanyak-di-dunia-2025/>
- Prasetyo et al. (2023). Analisis Sentimen Untuk Identifikasi Bantuan Korban Bencana Alam Berdasarkan Data Di Twitter Menggunakan Metode K- Sentiment Analysis For Identification Of Natural Disaster Victims Based On Data On Twitter Using K-Means And Naïve Bayes. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIIK)*, 10(5), 1055–1062.
<https://doi.org/10.25126/jtiik.2023107077>
- Putra et al. (2023). Analisis Sentimen Komentar Twitter Terhadap Peristiwa Kebakaran Bromo Menggunakan Metode Naïve Bayes dan Support Vector Machine. *UMS Library*, 1–19.
- Santoso et al. (2023). Social Network Analysis untuk Identifikasi Pengguna Twitter Berpengaruh pada Topik Bencana Gempa dan Tsunami di Indonesia. *Jurnal Edukasi Dan Penelitian Informatika*, 9(1), 115–122.
- Sulthan et al. (2021). Analisis sentimen pada tweet bencana alam menggunakan deep neural network dan information gain. *Jurnal Aplikasi Teknologi Informasi Dan Manajemen (JATIM)*, 2(2), 65–71.
- Tahir et al. (2024). Optimasi Naive Bayes Menggunakan Algoritma Genetika Pada Klasifikasi Komentar Cyberbullying Pada Media Sosial X. *JITET (Jurnal Informatika Dan Teknik Elektro Terapan)*, 12(3), 3350–3356.
- Voutama et al. (2024). Analisis Sentimen Menggunakan Algoritma Naïve Bayes Terhadap Regulasi Tiktok Shop Pada Media Sosial X (Twitter). *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(4), 5729–5735.
- Zumarniansyah et al. (2020). Twitter Sentiment Analysis Of Post Natural Disasters Using Comparative Classification Algorithm Support Vector. *Jurnal PILAR Nusa Mandiri*, 16(2), 169–174.