Djtechno : Jurnal Teknologi Informasi
Vol. 7, No. 2, Agustus 2026
E-ISSN: 2745-3758, P-ISSN : 2776-8546 DOI: 10.46576/djtechno

ANALISIS SENTIMEN MASYARAKAT TERHADAP PEMERINTAH DI ERA KABINET PRABOWO SUBIANTO BERDASARKAN SOSIAL MEDIA X MENGGUNAKAN NAÏVE BAYES CLASSIFIER

Adetia Irvanda[1], Wahyu Fuadi[2], Maryana[3]
1,2,3) Prodi Teknik Informatika, Fakultas Teknik, Universitas Malikussaleh, Indonesia

Article Info

Article history:

Received: 12 Juli 2026
Revised: 14 Juli 2026
Accepted: 16 Juli 2026

ABSTRACT

Abstrak

Kebijakan awal pemerintahan Kabinet Prabowo Subianto memicu beragam opini masyarakat di media sosial X, namun opini tersebut belum terklasifikasi ke dalam polaritas sentimen sehingga sulit dijadikan bahan evaluasi. Penelitian ini bertujuan mengimplementasikan dan mengevaluasi algoritma Naïve Bayes Classifier (NBC) dalam mengklasifikasikan sentimen masyarakat terhadap pemerintahan Kabinet Prabowo Subianto ke dalam kelas positif, negatif, dan netral. Sebanyak 587 tweet dikumpulkan melalui teknik scraping sejak 20 Oktober 2024, kemudian diproses melalui enam tahap text preprocessing, dilabeli otomatis menggunakan InSet Lexicon, dan dibobotkan dengan TF-IDF yang menghasilkan 317 fitur. Data dibagi 70:30 secara stratified menjadi 410 data latih dan 177 data uji, lalu diklasifikasikan menggunakan Multinomial Naïve Bayes dengan Laplace smoothing α=1,0. Hasil pelabelan menunjukkan dominasi sentimen negatif sebesar 55,7% (327 tweet), diikuti netral 30,7% (180 tweet) dan positif 13,6% (80 tweet). Model memperoleh akurasi 61,58% dengan presisi weighted 60,39%, recall weighted 61,58%, dan F1-score weighted 55,60%. Ketidakseimbangan kelas menyebabkan model bias terhadap kelas negatif, sehingga penyeimbangan data disarankan pada penelitian berikutnya.

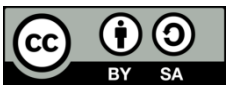
Kata Kunci: Analisis Sentimen, Naïve Bayes Classifier, TF-IDF, InSet Lexicon, Media Sosial X.

Abstract

The early policies of the Prabowo Subianto Cabinet administration triggered diverse public opinions on social media X, yet these opinions remain unclassified by sentiment polarity, making them difficult to use for evaluation. This study aims to implement and evaluate the Naïve Bayes Classifier (NBC) algorithm in classifying public sentiment toward the Prabowo Subianto Cabinet administration into positive, negative, and neutral classes. A total of 587 tweets were collected through scraping starting October 20, 2024, processed through six text preprocessing stages, automatically labeled using the InSet Lexicon, and weighted with TF-IDF, producing 317 features. The data were split 70:30 with stratification into 410 training and 177 testing samples, then classified using Multinomial Naïve Bayes with Laplace smoothing α=1.0. Labeling results show negative sentiment dominance at 55.7% (327 tweets), followed by neutral at 30.7% (180 tweets) and positive at 13.6% (80 tweets). The model achieved 61.58% accuracy with 60.39% weighted precision, 61.58% weighted recall, and 55.60% weighted F1-score. Class imbalance caused the model to be biased toward the negative class, so data balancing is recommended for future research.

Keywords: Sentiment Analysis, Naïve Bayes Classifier, TF-IDF, InSet Lexicon, Social Media X

Djtechno : Jurnal Teknologi Informasi
Vol. 7, No. 2, Agustus 2026
E-ISSN: 2745-3758, P-ISSN : 2776-8546 DOI: 10.46576/djtechno



Corresponding Author:
E-mail: adetia.210170166@mhs.unimal.ac.id

1. PENDAHULUAN

Kebijakan publik merupakan instrumen utama yang digunakan pemerintah untuk mengatur berbagai aspek kehidupan masyarakat, mulai dari ekonomi, pendidikan, kesehatan, hingga lingkungan. Di tengah perubahan cepat dalam masyarakat dan tantangan yang kompleks, kebijakan pemerintah menjadi fokus utama dalam upaya mencapai tujuan pembangunan yang berkelanjutan. Masyarakat menuangkan opininya terkait kebijakan baru pemerintah melalui berbagai cara, mulai dari aksi unjuk rasa hingga penyampaian pendapat pada platform media sosial, salah satunya X. Opini yang disampaikan melalui platform tersebut belum terbedakan ke dalam sentimen positif, negatif, atau netral, sehingga pemerintah kesulitan melihat penilaian publik terhadap kebijakan yang dikeluarkan. Media sosial telah berkembang menjadi alat bagi masyarakat Indonesia untuk menyuarakan pendapat mereka tentang kebijakan pemerintah dan masalah sosial lainnya [1].

Pada tahun 2023, pengguna media sosial di Indonesia mencapai 167 juta, dengan jumlah pengguna X pada awal tahun 2023 mencapai 24 juta pengguna. Laporan *We Are Social* per Oktober 2023 menyatakan Indonesia memiliki sekitar 27,5 juta pengguna X dan menempati peringkat keempat negara dengan pengguna X terbanyak [2]. Platform X memiliki karakteristik pesan singkat yang membuat pengguna dari semua kalangan mudah memberikan tanggapan positif maupun negatif terhadap topik yang sedang diperbincangkan, dan API yang disediakan X memudahkan pengambilan data sebagai sumber informasi penelitian [3]. Dengan begitu, media sosial X dapat dijadikan acuan untuk mengetahui kecenderungan opini masyarakat terhadap suatu kejadian, termasuk isu politik dan pemerintahan, karena analisis data X dapat memberikan gambaran yang bermanfaat tentang opini publik terhadap masalah politik [1].

Pemerintahan Prabowo Subianto dimulai sejak pelantikannya pada 20 Oktober 2024 dalam Sidang Paripurna MPR RI di Kompleks Parlemen, Jakarta. Pada hari yang

Djtechno : Jurnal Teknologi Informasi
Vol. 7, No. 2, Agustus 2026
E-ISSN: 2745-3758, P-ISSN : 2776-8546 DOI: 10.46576/djtechno

sama diumumkan Kabinet Merah Putih untuk periode 2024-2029 yang terdiri dari 48 menteri serta 5 pejabat setingkat menteri, struktur yang lebih besar dibandingkan kabinet-kabinet sebelumnya [4], [5]. Di awal pemerintahannya, Presiden Prabowo Subianto mengeluarkan beberapa kebijakan yang dinilai kontroversial oleh sebagian masyarakat, sehingga persepsi publik terhadap kebijakan tersebut sangat bervariasi antara dukungan dan kritik. Oleh karena itu, diperlukan analisis sentimen untuk mengetahui bagaimana sentimen di media sosial mengenai kepuasan masyarakat semasa pemerintahan Kabinet Prabowo Subianto.

Analisis sentimen adalah teknik untuk mengekstrak data opini, memahami dan memproses data teks secara otomatis, serta mengidentifikasi emosi yang terkandung dalam opini [6], sehingga informasi yang sebelumnya tidak terstruktur dapat diubah menjadi data yang lebih terorganisir. Proses analisis sentimen membutuhkan *text mining*, yaitu proses yang dilakukan komputer untuk mengekstrak informasi yang bermanfaat dari sejumlah dokumen dengan format tidak terstruktur yang bersumber dari teks [7].

Salah satu algoritma yang sering digunakan untuk analisis sentimen adalah *Naïve Bayes Classifier* (NBC). NBC adalah teknik klasifikasi teks yang menghitung data dengan cepat dan akurat sehingga populer digunakan pada data X [8]. Penelitian terdahulu menunjukkan NBC memiliki akurasi 80% dan presisi 75%, lebih tinggi dibandingkan metode *Holistic Lexicon Based* yang hanya mencapai 70% [9]. Penelitian pada data tweet era Kabinet Joko Widodo memperoleh akurasi Naïve Bayes sebesar 93,02% [2], penerapan NBC pada analisis sentimen dugaan pelanggaran Pemilu 2024 menunjukkan hasil klasifikasi yang baik [10], demikian pula pada analisis sentimen program Kartu Indonesia Pintar Kuliah [11] dan opini vaksin Covid-19 [12]. Di sisi lain, penerapan NBC pada data X BMKG Nasional hanya mencapai akurasi 69,97% [13], yang menunjukkan performa NBC dipengaruhi karakteristik data yang digunakan.

Berdasarkan uraian tersebut, penelitian ini mengimplementasikan dan mengevaluasi algoritma NBC dalam mengklasifikasikan sentimen masyarakat terhadap pemerintahan Kabinet Prabowo Subianto berdasarkan data media sosial X ke dalam tiga kelas, yaitu positif, negatif, dan netral. Penelitian ini juga mengukur kinerja model melalui akurasi, presisi, *recall*, dan F1-*score* dari *confusion matrix*. Hasil penelitian diharapkan memberikan informasi mengenai sentimen masyarakat

Djtechno : Jurnal Teknologi Informasi
Vol. 7, No. 2, Agustus 2026
E-ISSN: 2745-3758, P-ISSN : 2776-8546 DOI: 10.46576/djtechno

terhadap pemerintah di era Kabinet Prabowo Subianto sekaligus menjadi referensi bagi pemerintah untuk merumuskan kebijakan yang lebih responsif terhadap kebutuhan masyarakat.

2. METODE PENELITIAN

Penelitian ini terdiri dari beberapa tahapan yang berurutan, yaitu pengumpulan data, *text preprocessing*, pelabelan sentimen menggunakan InSet Lexicon, pembobotan TF-IDF, pembagian data latih dan data uji, klasifikasi dengan *Naïve Bayes Classifier*, dan evaluasi menggunakan *confusion matrix*. Sistem analisis sentimen dibangun berbasis web menggunakan bahasa pemrograman Python dan *framework* Flask, dengan perangkat keras berupa laptop berprosesor Intel Core i5 dan RAM 8 GB serta sistem operasi Windows 11 64 bit. Studi kepustakaan terhadap jurnal dan literatur yang relevan dilakukan sebagai landasan teori penelitian.

A. Pengumpulan Data

Pengumpulan data dilakukan dengan teknik *scraping*, yaitu metode ekstraksi data dari sebuah situs web yang kemudian disimpan dalam format tertentu [2]. Proses *scraping* memanfaatkan *Application Programming Interface* (API) yang disediakan platform X [3]. Data diambil mulai 20 Oktober 2024, bertepatan dengan pelantikan Presiden Prabowo Subianto, dengan kata kunci antara lain “Prabowo Subianto”, “Kabinet Prabowo”, “Menteri Prabowo”, “pemerintahan Prabowo”, “kebijakan Prabowo”, “program Prabowo”, “Presiden Prabowo”, “kinerja pemerintah Prabowo”, dan “Prabowo dan Gibran”. Hanya tweet berbahasa Indonesia yang digunakan, dengan total 587 tweet relevan yang disimpan sebagai dataset berformat .csv beratribut *username*, tweet, dan label.

B. Text Preprocessing

Text *preprocessing* merupakan proses mengubah bentuk data yang tidak terstruktur menjadi data terstruktur untuk mengecilkan dimensi data sehingga proses komputasi menjadi lebih efisien dan presisi [14]. Enam tahap dilakukan secara berurutan, yaitu *cleansing* (penghapusan hashtag, URL, mention, dan simbol), *case folding* (penyeragaman seluruh huruf menjadi huruf kecil), *tokenizing* (pemecahan kalimat menjadi token berdasarkan tanda baca dan spasi), *stopword removal* (pembuangan kata tidak deskriptif seperti “yang”, “dan”, “di”, “dari” menggunakan *stoplist* Tala), normalisasi kata tidak baku menggunakan kamus normalisasi sebanyak

Djtechno : Jurnal Teknologi Informasi
Vol. 7, No. 2, Agustus 2026
E-ISSN: 2745-3758, P-ISSN : 2776-8546 DOI: 10.46576/djtechno

4.330 entri [15], dan *stemming* ke bentuk kata dasar menggunakan algoritma *Enhanced Confix Stripping* (ECS) melalui pustaka PySastrawi v1.0.1, yang merupakan algoritma *stemming* bahasa Indonesia dengan tingkat kesalahan paling sedikit.

C. Pelabelan Sentimen dengan InSet Lexicon

Pelabelan sentimen dilakukan otomatis menggunakan InSet Lexicon yang memuat 3.609 kata positif berbobot +1 sampai +5 dan 6.609 kata negatif berbobot −1 sampai −5. Pendekatan pelabelan berbasis leksikon InSet telah digunakan pada berbagai penelitian analisis sentimen berbahasa Indonesia [16], [17]. Pelabelan diterapkan pada teks hasil *case folding* sebelum *stopword removal* agar kata negasi seperti "tidak" dan "bukan" tetap terjaga. Skor sentimen S dihitung sebagai penjumlahan bobot seluruh kata dalam dokumen sesuai Persamaan (1), dengan $w(t_i)$ adalah bobot kata ke-i dalam InSet yang bernilai 0 apabila kata tidak terdapat dalam leksikon dan n adalah jumlah token dalam dokumen.

$$S = \sum_{i=1}^{n} w(t_i) \quad (1)$$

Label kemudian ditetapkan berdasarkan ambang batas $\theta^+ = +0{,}5$ dan $\theta^- = -0{,}5$ sesuai Persamaan (2). Dokumen berlabel positif jika $S > \theta^+$, negatif jika $S < \theta^-$, dan netral jika $\theta^- \leq S \leq \theta^+$.

$$label(d) = \{positif, S > 0{,}5;\ negatif, S < -0{,}5;\ netral, -0{,}5 \leq S \leq 0{,}5\} \quad (2)$$

Sebagai contoh perhitungan manual, digunakan korpus baku tiga dokumen, yaitu D1 = "kabinet bagus maju", D2 = "gagal korupsi buruk", dan D3 = "kabinet biasa saja", dengan bobot InSet w(bagus)=+3, w(maju)=+2, w(gagal)=−3, w(korupsi)=−5, w(buruk)=−3, serta w(kabinet)=w(biasa)=w(saja)=0. Skor D1 dihitung S(D1) = 0 + 3 + 2 = 5, dan karena 5 > 0,5 maka D1 berlabel positif. Skor D2 dihitung S(D2) = (−3) + (−5) + (−3) = −11 sehingga berlabel negatif, sedangkan skor D3 bernilai 0 pada rentang −0,5 ≤ S ≤ 0,5 sehingga berlabel netral.

D. Pembobotan TF-IDF

Term Frequency-*Inverse Document Frequency* (TF-IDF) adalah metode pemberian bobot pada setiap kata dalam suatu dokumen, di mana *term* yang dianggap penting mendapatkan bobot tinggi dan *term* yang kurang relevan mendapatkan bobot rendah [18]. *Term Frequency* (TF) mengukur frekuensi kemunculan kata dalam sebuah dokumen sesuai Persamaan (3), sedangkan *Inverse Document Frequency* (IDF)

Djtechno : Jurnal Teknologi Informasi
Vol. 7, No. 2, Agustus 2026
E-ISSN: 2745-3758, P-ISSN : 2776-8546 DOI: 10.46576/djtechno

mengukur sebaran kata dalam seluruh koleksi dokumen sesuai Persamaan (4). Bobot akhir TF-IDF merupakan perkalian keduanya sesuai Persamaan (5).

$$TF(t,d) = \frac{f(t,d)}{\Sigma f(t\prime,d)} \quad (3)$$

$$IDF(t,D) = log\frac{|D|}{1+df(t,D)} + 1 \quad (4)$$

$$TFIDF(t,d,D) = TF(t,d) \times IDF(t,D) \quad (5)$$

Pada persamaan tersebut, f(t,d) adalah frekuensi kemunculan *term* t dalam dokumen d, Σf(t′,d) adalah jumlah seluruh *term* dalam dokumen d, |D| adalah jumlah total dokumen dalam korpus, dan df(t,D) adalah jumlah dokumen yang mengandung *term* t. Sebagai contoh perhitungan manual untuk term "gagal" pada dokumen D2 dengan korpus baku |D| = 3, diperoleh TF(gagal, D2) = 1/3 = 0,333, IDF(gagal) = log(3/(1+1)) + 1 = 1,176, sehingga TFIDF(gagal, D2) = 0,333 × 1,176 = 0,392. Pada implementasi sistem, pembobotan dilakukan menggunakan *TfidfVectorizer* Scikit-learn dengan parameter *max features*=1.000, *min df*=0,01, dan *max df*=0,95.

E. Pembagian Data

Dataset dibagi menjadi data latih dan data uji dengan rasio 70:30 menggunakan *train test split* Scikit-learn dengan parameter *stratify*=y dan *random state*=42 agar proporsi kelas terjaga pada kedua himpunan dan pembagian bersifat *reproducible*. Komposisi split 70:30 merupakan salah satu komposisi yang umum digunakan dan memberikan performa yang baik pada analisis sentimen dengan algoritma Naïve Bayes [19]. Dari 587 tweet, jumlah data latih dihitung n_{train} = ⌊587 × 0,70⌋ = 410 dan data uji n_{test} = 587 − 410 = 177. Alokasi per kelas mengikuti proporsi stratifikasi, yaitu kelas negatif 327 × 0,70 ≈ 228 latih dan 99 uji, kelas netral 180 × 0,70 = 126 latih dan 54 uji, serta kelas positif 80 × 0,70 = 56 latih dan 24 uji.

F. Naïve Bayes Classifier

Naïve Bayes Classifier adalah metode klasifikasi berbasis probabilitas sederhana yang mengaplikasikan teorema Bayes dengan asumsi independensi antar-atribut yang tinggi, sesuai untuk banyak dataset dengan performa cepat dan akurasi tinggi [20]. Probabilitas posterior kelas C_i terhadap data X dihitung sesuai Persamaan (6), dengan $P(C_i)$ sebagai *prior probability*, $P(X|C_i)$ sebagai *likelihood*, dan P(X) sebagai probabilitas data X [21].

$$P(C_i|X) = \frac{P(X|C_i) \cdot P(C_i)}{P(X)} \quad (6)$$

Djtechno : Jurnal Teknologi Informasi
Vol. 7, No. 2, Agustus 2026
E-ISSN: 2745-3758, P-ISSN : 2776-8546 DOI: 10.46576/djtechno

Model yang digunakan adalah *Multinomial Naïve Bayes* dengan *Laplace smoothing* α = 1,0 untuk menghindari probabilitas nol pada kata yang tidak muncul di suatu kelas. *Prior probability* setiap kelas dihitung sesuai Persamaan (7), dengan N_k adalah jumlah dokumen kelas C_k pada data latih, N total dokumen latih, dan K jumlah kelas (K = 3). *Likelihood* setiap term dihitung sesuai Persamaan (8), dengan $n(t,C_k)$ frekuensi term t pada seluruh dokumen kelas C_k, $n(C_k)$ total term pada kelas tersebut, dan |V| ukuran *vocabulary*. Kelas prediksi ditentukan melalui *maximum a posteriori* (MAP) pada Persamaan (9).

$$P(C_k) = \frac{N_k + \alpha}{N + \alpha \cdot K} \quad (7)$$

$$P(t|C_k) = \frac{n(t,C_k) + \alpha}{n(C_k) + \alpha \cdot |V|} \quad (8)$$

$$\hat{C} = argmax\,[\,log\,P(C_k) + \Sigma\,log\,P(t_i|C_k)\,] \quad (9)$$

Sebagai contoh perhitungan manual, digunakan korpus baku (N = 3, K = 3, α = 1, |V| = 8) dan dokumen uji "kabinet gagal". *Prior probability* ketiga kelas identik, yaitu P(positif) = P(negatif) = P(netral) = (1+1)/(3+1×3) = 2/6 = 0,333. *Likelihood* pada kelas positif diperoleh P(kabinet|positif) = (1+1)/(3+1×8) = 2/11 = 0,182 dan P(gagal|positif) = (0+1)/11 = 0,091, sedangkan pada kelas negatif berlaku sebaliknya. Skor MAP kelas positif dihitung log(0,333) + log(0,182) + log(0,091) = −5,202, yang pada contoh sederhana ini bernilai sama untuk ketiga kelas karena simetri data baku. Pada data nyata dengan 317 fitur TF-IDF, akumulasi perbedaan *likelihood* antar-kelas menghasilkan skor yang berbeda nyata sehingga prediksi dapat ditentukan.

G. Evaluasi Model

Kinerja model dievaluasi menggunakan *confusion matrix*, yaitu tabel yang menunjukkan jumlah data uji yang diklasifikasikan dengan benar dan salah [22]. Karena penelitian ini menggunakan tiga kelas sentimen, digunakan *confusion matrix* 3×3. Dari matriks tersebut dihitung akurasi sesuai Persamaan (10), presisi per kelas sesuai Persamaan (11), *recall* per kelas sesuai Persamaan (12), dan F1-*score* per kelas sesuai Persamaan (13). Karena distribusi kelas tidak seimbang, dihitung pula rata-rata berbobot (*weighted average*) sesuai Persamaan (14), dengan n_k adalah jumlah data uji berlabel kelas k.

$$Akurasi = \frac{TP_{total}}{N_{test}} \times 100\% \quad (10)$$

$$Presisi_k = \frac{TP_k}{TP_k + FP_k} \quad (11)$$

Djtechno : Jurnal Teknologi Informasi
Vol. 7, No. 2, Agustus 2026
E-ISSN: 2745-3758, P-ISSN : 2776-8546 DOI: 10.46576/djtechno

$$Recall_k = \frac{TP_k}{TP_k + FN_k} \quad (12)$$

$$F1_k = \frac{2 \times Presisi_k \times Recall_k}{Presisi_k + Recall_k} \quad (13)$$

$$Metrik_w = \Sigma \frac{n_k}{N_{test}} \times Metrik_k \quad (14)$$

3. HASIL DAN PEMBAHASAN

A. Hasil Text Preprocessing

Sebanyak 587 tweet mentah diproses melalui enam tahap *text preprocessing* secara berurutan. Proses ini berhasil menghapus 235 URL, 927 mention, dan 169 hashtag, disertai penghapusan 1.996 *stopword*, normalisasi 1.038 kata tidak baku, dan reduksi 2.118 kata berimbuhan menjadi kata dasar. Rata-rata jumlah kata per tweet berkurang dari 24,5 menjadi 18,4 kata atau tereduksi 24,8%. Penelusuran satu tweet melalui keenam tahap disajikan pada Tabel 1.

Tabel 1. Penelusuran Satu Tweet Melalui Enam Tahap Preprocessing

Tahap	Proses	Hasil
Input	-	@bahlillahadalia copot aja ni org @bahlillahadalia dari kabinet @prabowo @gibran_tweet
1. Cleansing	Hapus 4 mention	copot aja ni org dari kabinet
2. Case Folding	Ubah ke huruf kecil	copot aja ni org dari kabinet (tidak berubah)
3. Tokenizing	Pisah menjadi token	[“copot”,“aja”,“ni”,“org”,“dari”,“kabinet”] → 6 token
4. Stopword Removal	Hapus “dari”	[“copot”,“aja”,“ni”,“org”,“kabinet”] → 5 token
5. Normalisasi	aja→saja, ni→ini, org→orang	[“copot”,“saja”,“ini”,“orang”,“kabinet”]
6. Stemming	Tidak ada perubahan	copot saja ini orang kabinet
Output	-	copot saja ini orang kabinet (5 kata)

Sumber: Hasil Penelitian (2026)

Tabel 1 memperlihatkan tahap *cleansing* memberikan kontribusi terbesar dalam pengurangan token pada tweet tersebut dengan menghapus empat mention sekaligus, sehingga jumlah token berkurang dari 10 menjadi 6. Tahap normalisasi mengubah tiga kata tidak baku (“aja”, “ni”, “org”) menjadi bentuk baku yang dapat dikenali oleh InSet Lexicon maupun pembobot TF-IDF. Tahap *stemming* tidak menghasilkan perubahan karena seluruh kata yang tersisa sudah berbentuk kata dasar, sehingga hasil akhir berupa lima kata informatif yang mencerminkan sentimen negatif tweet tersebut. Contoh hasil *preprocessing* pada beberapa tweet lain disajikan pada Tabel 2.

Djtechno : Jurnal Teknologi Informasi
Vol. 7, No. 2, Agustus 2026
E-ISSN: 2745-3758, P-ISSN : 2776-8546 DOI: 10.46576/djtechno

Tabel 2. Contoh Hasil Tweet Tahap Preprocessing

No	Teks Asli	Teks Hasil Preprocessing	Kata Awal	Kata Akhir
1	Presiden Prabowo akan kesulitan benahi masalah bangsa yang carut-marut bilamana orang-orang Jokowi m...	presiden prabowo sulit benah masalah bangsa carutmarut bilamana orangorang jokowi masih ad	33	27
2	@ferrykoto Dan sekarang tetap dipelihara oleh @prabowo jadilah kabinet omon2 .	sekarang tetap pelihara jadi kabinet omon	11	6
3	@124hmahHastuti1 DIPECAT DARI KABINET PRABOWO !	pecat kabinet prabowo	6	3
4	@NgkongRoses @ch_chotimah2 ga ada yang bisa dibanggakan dari kabinet prabowo sekarang.. to*ol semua.	enggak ada bisa bangga kabinet prabowo sekarang tool semua	13	9

Sumber: Hasil Penelitian (2026)

B. Hasil Pelabelan Sentimen

Seluruh 587 tweet diberi label sentimen otomatis menggunakan InSet Lexicon dengan ambang batas $\theta^+ = +0{,}5$ dan $\theta^- = -0{,}5$ sesuai Persamaan (1) dan (2). Distribusi label hasil pelabelan disajikan pada Tabel 3. Kelas negatif mendominasi dengan 327 tweet atau 55,7% yang mencerminkan kecenderungan masyarakat mengungkapkan kritik terhadap pemerintahan Kabinet Prabowo Subianto melalui platform X pada periode pengumpulan data.

Tabel 3. Distribusi Label Sentimen Hasil Pelabelan InSet Lexicon

Kelas Sentimen	Jumlah Tweet	Persentase (%)
Negatif	327	55,7
Netral	180	30,7
Positif	80	13,6
Total	587	100,0

Sumber: Hasil Penelitian (2026)

Tabel 3 menunjukkan ketidakseimbangan distribusi kelas yang signifikan. Kelas netral sebagian besar berasal dari tweet pelaporan perkembangan pemerintahan tanpa muatan evaluatif eksplisit, sedangkan kelas positif sebagai kelas minoritas menandakan ekspresi dukungan lebih jarang ditemukan dibandingkan kritik. Ketidakseimbangan distribusi ini diantisipasi melalui parameter *stratify* pada tahap pembagian data agar proporsi kelas terjaga secara proporsional pada data latih maupun data uji.

C. Hasil Pembobotan TF-IDF

Pembobotan TF-IDF terhadap 587 tweet hasil *preprocessing* menghasilkan 317 fitur sehingga terbentuk matriks berukuran 587×317. Kebenaran perhitungan sistem diverifikasi melalui perhitungan manual pada korpus baku sesuai Persamaan (3), (4),

Djtechno : Jurnal Teknologi Informasi
Vol. 7, No. 2, Agustus 2026
E-ISSN: 2745-3758, P-ISSN : 2776-8546 DOI: 10.46576/djtechno

dan (5), di mana kata "kabinet" yang muncul pada dua dokumen memperoleh bobot lebih rendah (0,333) dibandingkan kata yang hanya muncul pada satu dokumen seperti "gagal" (0,392) karena nilai IDF-nya lebih kecil (1,000 berbanding 1,176). Pola serupa terjadi pada data nyata, sebagaimana disajikan pada Tabel 4 yang memuat sampel nilai TF-IDF aktual dari sistem pada salah satu tweet data uji.

Tabel 4. Sampel Nilai TF-IDF Aktual dari Sistem

No	Term	Nilai TF-IDF	Keterangan
1	nkri	0,6303	Bobot tertinggi; term sangat spesifik pada tweet ini
2	tri	0,4856	Term langka; jarang muncul di dokumen lain dalam korpus
3	titip	0,4692	Term kontekstual negatif; sering muncul dalam konteks kritik
4	jokowi	0,3423	Term umum; muncul di banyak tweet dalam dataset
5	prabowo	0,1719	Bobot terendah; muncul hampir di seluruh 587 tweet

Sumber: Hasil Penelitian (2026)

Tabel 4 menunjukkan term "nkri" mendapatkan nilai TF-IDF tertinggi (0,6303) karena kemunculannya yang langka dalam korpus menghasilkan nilai IDF tinggi, sedangkan term "prabowo" mendapatkan nilai terendah (0,1719) karena muncul hampir merata di seluruh 587 tweet. Fenomena ini membuktikan TF-IDF efektif menekan bobot kata yang terlalu umum dan mengangkat bobot kata yang spesifik pada dokumen tertentu sebagai dasar keputusan klasifikasi NBC.

D. Hasil Pembagian Data

Dataset 587 tweet dibagi menjadi 410 data latih (70%) dan 177 data uji (30%) secara *stratified* seperti disajikan pada Tabel 5. Verifikasi aritmetika (228 + 126 + 56 = 410 dan 99 + 54 + 24 = 177) memastikan tidak ada data yang hilang atau terduplikasi dalam proses pembagian.

Tabel 5. Distribusi Label pada Data Latih dan Data Uji

Kelas Sentimen	Total	Data Latih (70%)	Data Uji (30%)
Negatif	327	228	99
Netral	180	126	54
Positif	80	56	24
Total	587	410	177

Sumber: Hasil Penelitian (2026)

Mekanisme *stratified split* berhasil mempertahankan proporsi kelas pada kedua himpunan, yaitu kelas negatif 55,6% pada data latih berbanding 55,9% pada data uji, kelas netral 30,7% berbanding 30,5%, dan kelas positif 13,7% berbanding 13,6%, sehingga selisih proporsi tidak melebihi 0,3 poin persentase pada semua kelas. Konsistensi ini memastikan model dievaluasi pada distribusi kelas yang representatif,

Djtechno : Jurnal Teknologi Informasi
Vol. 7, No. 2, Agustus 2026
E-ISSN: 2745-3758, P-ISSN : 2776-8546 DOI: 10.46576/djtechno

sementara penggunaan *random state*=42 menjamin hasil penelitian dapat direproduksi secara independen.

E. Hasil Pelatihan Model NBC

Model *Multinomial Naïve Bayes* dilatih pada 410 data latih dengan 317 fitur TF-IDF dan parameter α = 1,0. Proses pelatihan menghasilkan estimasi *prior probability* setiap kelas melalui Persamaan (7) yang disajikan pada Tabel 6.

Tabel 6. Nilai Prior Probability Hasil Pelatihan Model

Kelas	Nk (Latih)	Perhitungan	P(Ck)	log P(Ck)
Negatif	228	(228+1)/(410+3) = 229/413	0,5545	−0,5897
Netral	126	(126+1)/(410+3) = 127/413	0,3075	−1,1793
Positif	56	(56+1)/(410+3) = 57/413	0,1380	−1,9804
Total	410	229+127+57 = 413	1,0000	-

Sumber: Hasil Penelitian (2026)

Tabel 6 menunjukkan *prior probability* kelas negatif (0,5545) jauh lebih tinggi dibandingkan kelas netral (0,3075) dan positif (0,1380), mencerminkan dominasi kelas negatif pada data latih. Ketimpangan ini berkontribusi pada bias prediksi model ke arah kelas negatif, karena skor log-posterior kelas negatif mendapat keunggulan awal sebesar log(0,5545) = −0,5897 dibandingkan log(0,1380) = −1,9804 untuk kelas positif. Verifikasi total $P(C_k)$ = 1,0000 mengonfirmasi ketiga probabilitas membentuk distribusi yang valid. Nilai *likelihood* beberapa fitur teratas per kelas hasil pelatihan disajikan pada Tabel 7.

Tabel 7. Nilai Likelihood Fitur Teratas per Kelas Sentimen

Term	TF-IDF	P(t|Neg)	P(t|Net)	P(t|Pos)
nkri	0,6303	0,001796	0,002699	0,002039
ganggu	0,5875	0,002065	0,002062	0,002039
tri	0,4856	0,003708	0,003655	0,004928
titip	0,4692	0,005359	0,002894	0,002552
harus	0,4231	0,005306	0,002156	0,002356

Sumber: Hasil Penelitian (2026)

Tabel 7 memperlihatkan term “titip” memiliki nilai P(t|Neg) lebih tinggi (0,005359) dibandingkan P(t|Net) (0,002894) dan P(t|Pos) (0,002552), yang mengindikasikan term ini lebih sering muncul dalam konteks tweet negatif pada data latih. Term “nkri” memiliki nilai TF-IDF tertinggi namun *likelihood* yang sangat kecil karena kemunculannya langka dalam keseluruhan data latih. Perbedaan nilai *likelihood* antar-kelas untuk setiap term inilah yang secara akumulatif membedakan skor MAP ketiga kelas melalui Persamaan (9) sehingga prediksi dapat dilakukan pada data nyata.

Djtechno : Jurnal Teknologi Informasi
Vol. 7, No. 2, Agustus 2026
E-ISSN: 2745-3758, P-ISSN : 2776-8546 DOI: 10.46576/djtechno

F. Hasil Pengujian dan Evaluasi Model

Model diuji terhadap 177 data uji dan menghasilkan *confusion matrix* 3×3 pada Tabel 8. Sampel hasil prediksi model terhadap data uji beserta skor probabilitas posteriornya disajikan pada Tabel 9.

Tabel 8. Confusion Matrix Hasil Pengujian Model NBC

Aktual \ Prediksi	Negatif	Netral	Positif
Negatif (99)	92	7	0
Netral (54)	39	14	1
Positif (24)	12	9	3

Sumber: Hasil Penelitian (2026)

Tabel 9. Sampel Hasil Prediksi Model NBC terhadap Data Uji

No	Tweet (Sesudah Preprocessing)	Aktual	Prediksi	P(Neg)	P(Net)	P(Pos)
1	aceh deh tenteram damai kenapa harus ganggu ganggu lagi pak prabowo...	negatif	negatif	0,7557	0,2062	0,0381
2	panjang titip orang jokowi masih kabinet apa prabowo pasti jalan agenda yan	negatif	negatif	0,6916	0,2227	0,0857
3	tri titip jokowi kabinet prabowo khianat bangsatperusak nkri	negatif	negatif	0,6163	0,2974	0,0863
4	presiden prabowo pimpin rapat batas bentuk tanggul laut pulau jawa	netral	negatif	0,4996	0,2872	0,2132

Sumber: Hasil Penelitian (2026)

Berdasarkan Tabel 8, akurasi model dihitung sesuai Persamaan (10) sebagai berikut.

$$Akurasi = \frac{92 + 14 + 3}{177} \times 100\% = \frac{109}{177} \times 100\% = 61{,}58\%$$

Presisi per kelas dihitung sesuai Persamaan (11), yaitu presisi kelas negatif 92/(92+51) = 92/143 = 64,34%, presisi kelas netral 14/(14+16) = 14/30 = 46,67%, dan presisi kelas positif 3/(3+1) = 3/4 = 75,00%. *Recall* per kelas dihitung sesuai Persamaan (12), yaitu *recall* kelas negatif 92/(92+7) = 92/99 = 92,93%, *recall* kelas netral 14/(14+40) = 14/54 = 25,93%, dan *recall* kelas positif 3/(3+21) = 3/24 = 12,50%. F1-*score* per kelas sesuai Persamaan (13) menghasilkan 76,03% (negatif), 33,33% (netral), dan 21,43% (positif). Rata-rata berbobot dihitung sesuai Persamaan (14) sebagai berikut.

$$Presisi_w = \frac{99}{177}(0{,}6434) + \frac{54}{177}(0{,}4667) + \frac{24}{177}(0{,}7500) = 60{,}39\%$$

$$Recall_w = 0{,}5197 + 0{,}0792 + 0{,}0169 = 61{,}58\%$$

$$F1_w = 0{,}4252 + 0{,}1017 + 0{,}0291 = 55{,}60\%$$

Rekapitulasi seluruh metrik evaluasi per kelas disajikan pada Tabel 10.

Tabel 10. Rekapitulasi Metrik Evaluasi Klasifikasi per Kelas

Djtechno : Jurnal Teknologi Informasi
Vol. 7, No. 2, Agustus 2026
E-ISSN: 2745-3758, P-ISSN : 2776-8546 DOI: 10.46576/djtechno

Kelas	nk	TP	FP	FN	Presisi (%)	Recall (%)	F1-Score (%)
Negatif	99	92	51	7	64,34	92,93	76,03
Netral	54	14	16	40	46,67	25,93	33,33
Positif	24	3	1	21	75,00	12,50	21,43
Weighted Avg	177	109	-	-	60,39	61,58	55,60

Sumber: Hasil Penelitian (2026)

G. Pembahasan

Tabel 8 mengungkapkan pola kesalahan klasifikasi yang dominan, yaitu 39 tweet netral dan 12 tweet positif salah diprediksi sebagai negatif. Total prediksi negatif mencapai 143 tweet, jauh melampaui jumlah aktual kelas negatif pada data uji sebanyak 99 tweet, yang konsisten dengan bias model terhadap kelas mayoritas akibat ketimpangan *prior probability* pada Tabel 6. Kelas negatif menunjukkan performa terbaik dengan hanya 7 kesalahan dari 99 sampel dan tidak ada satu pun tweet negatif yang salah diprediksi sebagai positif. Sebaliknya, model hampir tidak pernah memprediksi kelas positif dengan hanya 4 prediksi dari 177 data uji.

Akurasi model diperoleh sebesar 61,58% yang berasal dari 109 prediksi benar, terdiri dari 92 prediksi benar kelas negatif, 14 kelas netral, dan 3 kelas positif. Nilai akurasi ini perlu dibaca secara hati-hati karena data uji tidak seimbang, di mana kelas negatif mencakup 55,93% dari total data uji sehingga sebagian besar kontribusi akurasi berasal dari keberhasilan pada kelas mayoritas. Kelas negatif memiliki *recall* tertinggi (92,93%) namun presisi rendah (64,34%), yang berarti model mampu mendeteksi hampir semua tweet negatif tetapi sering salah mengklasifikasikan tweet kelas lain sebagai negatif. Kelas positif menunjukkan pola berlawanan dengan presisi tinggi (75,00%) namun *recall* sangat rendah (12,50%), yang berarti prediksi positif model cenderung benar namun sebagian besar tweet positif terlewat. Nilai F1-*score weighted* sebesar 55,60% mencerminkan ketidakseimbangan performa antar-kelas yang membutuhkan penanganan lebih lanjut.

Hasil akurasi 61,58% pada penelitian ini lebih rendah dibandingkan penelitian sejenis pada data era Kabinet Joko Widodo yang mencapai 93,02% [2] maupun penerapan NBC pada topik lain yang umumnya melampaui 90% [8], [10], [11], [12], meskipun terdapat pula penerapan NBC dengan akurasi 69,97% pada data X BMKG Nasional [13]. Perbedaan ini dapat dijelaskan oleh tiga faktor. Pertama, ketidakseimbangan distribusi kelas yang tajam (55,7% berbanding 13,6%) membuat model bias terhadap kelas mayoritas. Kedua, penggunaan tiga kelas sentimen lebih sulit dibandingkan klasifikasi biner karena batas antara kelas netral dengan kedua

Djtechno : Jurnal Teknologi Informasi
Vol. 7, No. 2, Agustus 2026
E-ISSN: 2745-3758, P-ISSN : 2776-8546 DOI: 10.46576/djtechno

kelas lainnya tidak selalu tegas. Ketiga, pelabelan otomatis berbasis leksikon bergantung pada cakupan kosakata InSet; temuan serupa dilaporkan pada pelabelan berbasis leksikon yang menghasilkan akurasi 55,6% [23], sehingga kualitas pelabelan menjadi faktor penting bagi performa klasifikasi.

Dari sisi fungsionalitas, pengujian sistem terhadap 15 skenario yang meliputi *login*, *dashboard*, *scraping*, *preprocessing*, pelabelan, konversi TF-IDF, pembagian data, pelatihan model, pengujian model, dan ekspor CSV seluruhnya memberikan hasil valid sesuai harapan, dengan antarmuka web Flask yang mengintegrasikan seluruh alur analisis dalam satu sistem. Untuk meningkatkan performa klasifikasi, penanganan ketidakseimbangan kelas seperti *oversampling*, penyesuaian ambang klasifikasi, kombinasi pelabelan manual, atau perbandingan dengan algoritma lain seperti SVM, Random Forest, dan Logistic Regression layak dipertimbangkan pada penelitian berikutnya [24], [25], [26].

4. SIMPULAN

Penelitian ini berhasil mengimplementasikan algoritma *Naïve Bayes Classifier* untuk mengklasifikasikan sentimen masyarakat terhadap pemerintahan Kabinet Prabowo Subianto berdasarkan 587 tweet dari media sosial X ke dalam kelas positif, negatif, dan netral. Hasil pelabelan InSet Lexicon menunjukkan sentimen masyarakat didominasi kelas negatif sebesar 55,7% (327 tweet), diikuti netral 30,7% (180 tweet) dan positif 13,6% (80 tweet), yang mengindikasikan kritik lebih banyak diekspresikan dibandingkan dukungan pada periode awal pemerintahan. Model *Multinomial Naïve Bayes* dengan pembobotan TF-IDF 317 fitur memperoleh akurasi 61,58%, presisi *weighted* 60,39%, *recall weighted* 61,58%, dan F1-*score weighted* 55,60%, dengan performa terbaik pada kelas negatif dan terendah pada kelas positif akibat ketidakseimbangan distribusi kelas. Pengembangan selanjutnya disarankan menggunakan dataset yang lebih besar dan seimbang, menerapkan algoritma pembanding seperti SVM atau Random Forest, serta mengembangkan analisis sentimen secara *real-time* dari data media sosial.

REFERENCES

[1] N. I. Purnamasari and A. Wijanarko, "Peran buzzer politik Ganjar-Mahfud dalam pembentukan opini publik di media sosial X," Warta Ikatan Sarjana Komunikasi Indonesia, vol. 7, no. 2, 2024, doi: 10.25008/wartaiski.v7i2.290.

Djtechno : Jurnal Teknologi Informasi
Vol. 7, No. 2, Agustus 2026
E-ISSN: 2745-3758, P-ISSN : 2776-8546 DOI: 10.46576/djtechno

[2] F. Naufal Rahman and S. Lestari, "Analisis sentimen masyarakat terhadap pemerintah di era Kabinet Joko Widodo berdasarkan sosial media X menggunakan Naïve Bayes dan K-Nearest Neighbor (KNN)," J. Inf. Technol. Comput. Sci. (INTECOMS), vol. 7, no. 5, 2024.

[3] Y. Afrillia, L. Rosnita, and D. Siska, "Analisis sentimen pengguna Twitter terhadap isu kesetaraan gender dalam penerapan Permendikbudristek Nomor 30 Tahun 2021 menggunakan TextBlob," J. Informatics Comput. Sci., vol. 8, no. 2, 2022.

[4] Humas Sekretariat Kabinet, "Presiden Prabowo Subianto umumkan susunan Kabinet Merah Putih di Istana Merdeka," 2024. [Online]. Available: https://setkab.go.id/presiden-prabowo-subianto-umumkan-susunan-kabinet-merah-putih-di-istana-merdeka-jakarta/

[5] Sekretariat Presiden, "Presiden Prabowo Subianto umumkan Kabinet Merah Putih periode 2024-2029," 2024. [Online]. Available: https://www.presidenri.go.id/siaran-pers/presiden-prabowo-subianto-umumkan-kabinet-merah-putih-periode-2024-2029/

[6] F. V. Sari and A. Wibowo, "Analisis sentimen pelanggan toko online JD.id menggunakan metode Naïve Bayes Classifier berbasis konversi ikon emosi," J. SIMETRIS, vol. 10, no. 2, 2020.

[7] A. Firdaus and W. I. Firdaus, "Text mining dan pola algoritma dalam penyelesaian masalah informasi: (sebuah ulasan)," J. JUPITER, vol. 13, no. 1, 2021.

[8] Y. A. Singgalen, "Analisis sentimen wisatawan melalui data ulasan Candi Borobudur di Tripadvisor menggunakan algoritma Naïve Bayes Classifier," Building Informatics, Technol. Sci. (BITS), vol. 4, no. 3, 2022, doi: 10.47065/bits.v4i3.2486.

[9] E. E. Amelia and I. Yustiana, "Analisis sentimen pada ulasan produk UNIQLO dengan algoritma Naive Bayes," J. Sains Komput. Inform. (J-SAKTI), vol. 8, no. 1, 2024.

[10] D. S. Nugroho, I. F. Hanif, M. A. Hasbi, F. Fredianto, A. M. Saputra, and R. Zildjian, "Analisis sentimen dugaan pelanggaran Pemilu 2024 berdasarkan tweet menggunakan algoritma Naïve Bayes Classifier," MALCOM: Indones. J. Mach. Learn. Comput. Sci., vol. 4, no. 3, pp. 1169-1176, 2024, doi: 10.57152/malcom.v4i3.1496.

[11] D. Pramudita, Y. Akbar, and T. Wahyudi, "Analisis sentimen terhadap program Kartu Indonesia Pintar Kuliah pada media sosial X menggunakan algoritma Naive Bayes," MALCOM: Indones. J. Mach. Learn. Comput. Sci., vol. 4, no. 4, pp. 1420-1430, 2024, doi: 10.57152/malcom.v4i4.1565.

[12] F. Fathonah and A. Herliana, "Penerapan text mining analisis sentimen mengenai vaksin Covid-19 menggunakan metode Naïve Bayes," J. Sains dan Inform., vol. 7, no. 2, pp. 155-164, 2021, doi: 10.34128/jsi.v7i2.331.

[13] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan algoritma Naive Bayes untuk analisis sentimen review data Twitter BMKG Nasional," J. Teknoinfo, vol. 15, no. 1, 2021.

[14] Nurdin, M. Suhendri, Y. Afrilia, and Rizal, "Klasifikasi karya ilmiah (tugas akhir) mahasiswa menggunakan metode Naïve Bayes Classifier (NBC)," SISTEMASI: J. Sist. Inf., 2021.

[15] M. I. Raif, N. N. Hidayati, and T. Matulatan, "Otomatisasi pendeteksi kata baku dan tidak baku pada data Twitter berbasis KBBI," J. Teknol. Inf. Ilmu Komput., vol. 11, no. 2, pp. 337-348, 2024, doi: 10.25126/jtiik.20241127404.

[16] M. Fernanda and N. Fathoni, "Perbandingan performa labeling lexicon InSet dan VADER pada analisa sentimen Rohingya di aplikasi X dengan SVM," J. Informatika dan Sains Teknol., vol. 1, no. 3, pp. 62-76, 2024, doi: 10.62951/modem.v1i3.112.

[17] D. Pratiwi, N. Khoerani, and S. Sari, "Analisis sentimen terhadap kebijakan subsidi pembelian kendaraan bertenaga listrik di Indonesia menggunakan pendekatan InSet Lexicon dan metode Support Vector Machine," J. Teknol. Inf. Ilmu Komput., vol. 12, no. 6, pp. 1303-1314, 2025.

[18] S. Khairunnisa, A. Adiwijaya, and S. Al Faraby, "Pengaruh text preprocessing terhadap analisis sentimen komentar masyarakat pada media sosial Twitter (studi kasus pandemi COVID-19)," J. Media Inform. Budidarma, vol. 5, no. 2, p. 406, 2021, doi: 10.30865/mib.v5i2.2835.

[19] Y. A. Prasetyo, E. Utami, and A. Yaqin, "Pengaruh komposisi split data terhadap performa akurasi analisis sentimen algoritma Naïve Bayes dan SVM," J. Electr. Eng. Comput. (JEECOM), vol. 6, no. 2, 2024, doi: 10.33650/jeecom.v4i2.

Djtechno : Jurnal Teknologi Informasi
Vol. 7, No. 2, Agustus 2026
E-ISSN: 2745-3758, P-ISSN : 2776-8546 DOI: 10.46576/djtechno
[20] A. Safira and F. N. Hasan, "Analisis sentimen masyarakat terhadap paylater menggunakan metode Naive Bayes Classifier," J. Sist. Inf., vol. 5, no. 1, 2023.

[21] H. F. Putro, R. T. Vulandari, and W. L. Y. Saptomo, "Penerapan metode Naive Bayes untuk klasifikasi pelanggan," J. Teknol. Inf. Komun. (TIKomSiN), vol. 8, no. 2, 2020, doi: 10.30646/tikomsin.v8i2.500.

[22] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier dan confusion matrix pada analisis sentimen berbasis teks pada Twitter," J. Sains Komput. Inform. (J-SAKTI), vol. 5, no. 2, 2021.

[23] A. Undap, A. Sambul, and S. Pakpahan, "Analisis sentimen situs pembajak artikel penelitian menggunakan metode Lexicon-Based," J. Tek. Inform., 2021.

[24] B. A. Santoso, I. Nugroho, and D. U. Asyfiya, "Perbandingan algoritma Naïve Bayes, Support Vector Machine, dan Random Forest untuk analisis sentimen komentar politik YouTube," TIN: Terapan Inform. Nusantara, vol. 6, no. 4, 2025, doi: 10.47065/tin.v6i4.8326.

[25] A. Atikah Putri, S. Agustian, and R. Abdillah, "Penerapan metode Logistic Regression untuk klasifikasi sentimen pada dataset Twitter terbatas," J. Sist. Inf., vol. 7, no. 1, 2025.

[26] V. A. Sulistiani and M. Hamka, "Analisis sentimen pengguna media sosial terhadap Identitas Kependudukan Digital menggunakan metode Support Vector Machine (SVM)," J. Inf. Syst. Res. (JOSH), vol. 5, no. 4, pp. 1312-1322, 2024, doi: 10.47065/josh.v5i4.5211.