

ANALISIS SENTIMEN PUBLIK TERHADAP TREN #KABURAJA DULU DI MEDIA SOSIAL X MENGGUNAKAN ALGORITMA *NAÏVE BAYES*

Indira Fildah Fakhrana^{1*}, Hannie², Azhari Ali Ridha³

^{1,2,3}Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang; Jl. HS Ronggo Waluyo, Puseurjaya, Telukjambe Timur, Karawang, Jawa Barat 41361

Keywords:

Analisis Sentimen, *Naïve Bayes*, KDD, Media Sosial X, #KaburAjaDulu.

Correspondent Email:

raraindira79@gmail.com

Abstrak. Penelitian ini menganalisis sentimen publik terhadap tren #KaburAjaDulu di platform X menggunakan algoritma *Naïve Bayes* dengan pendekatan *Knowledge Discovery in Databases* (KDD). Data dikumpulkan melalui proses *crawling* pada periode 1 Desember 2024–1 Desember 2025 sebanyak 3.102 tweet berbahasa Indonesia. Pelabelan sentimen dilakukan melalui dua pendekatan, yaitu pelabelan manual oleh ahli bahasa dan pelabelan berbasis leksikon. Representasi fitur menggunakan TF-IDF, sedangkan klasifikasi dilakukan dengan *Naïve Bayes*. Evaluasi model menggunakan skenario pembagian data latih dan uji 90:10, 80:20, dan 70:30, serta penanganan ketidakseimbangan kelas dengan SMOTE. Hasil menunjukkan dominasi sentimen negatif sebesar 83% dan sentimen positif 16%. Performa terbaik diperoleh pada skenario 70:30 dengan akurasi 74%. Temuan ini menunjukkan efektivitas *Naïve Bayes* dalam klasifikasi sentimen teks serta pentingnya kualitas pelabelan data.

Abstract. This study analyzes public sentiment toward the #KaburAjaDulu trend on the X platform using the *Naïve Bayes* algorithm with a *Knowledge Discovery in Databases* (KDD) approach. Data was collected through a *crawling* process during the period of December 1, 2024–December 1, 2025, amounting to 3,102 tweets in Indonesian. Sentiment labeling was carried out using two approaches, namely manual labeling by linguists and lexicon-based labeling. Feature representation used TF-IDF, while classification was performed using *Naïve Bayes*. Model evaluation used training and testing data division scenarios of 90:10, 80:20, and 70:30, as well as class imbalance handling with SMOTE. The results show a dominance of negative sentiment at 83% and positive sentiment at 16%. The best performance was obtained in the 70:30 scenario with an accuracy of 74%. These findings demonstrate the effectiveness of *Naïve Bayes* in text sentiment classification and the importance of data labeling quality.



Copyright © JITET (Jurnal Informatika dan Teknik Elektro Terapan). This article is an open access article distributed under terms and conditions of the Creative Commons Attribution (CC BY NC)

1. PENDAHULUAN

Kemajuan teknologi digital telah memposisikan media sosial sebagai platform utama interaksi sosial dan penyampaian opini publik. Di Indonesia, jumlah pengguna aktif media sosial mencapai 167 juta orang, yang mengindikasikan kuatnya pengaruh platform ini dalam membentuk wacana publik [1]. Media

sosial X (sebelumnya Twitter), dengan fitur trending *topic*-nya, menjadi salah satu ruang utama bagi publik untuk mendiskusikan isu-isu sosial secara *real-time* [2].

Salah satu fenomena digital yang viral di platform X adalah tagar #KaburAjaDulu. Tagar ini tidak hanya menjadi hiburan semata, melainkan juga cerminan dari kecemasan

kollektif, ekspresi keresahan, dan kritik terhadap kondisi sosial, ekonomi, serta politik di Indonesia [3], [4]. Tagar ini merepresentasikan aspirasi sebagian masyarakat akan transformasi sosial dan kesempatan yang lebih baik, sebagai respons terhadap gejolak ekonomi dan kesenjangan sosial [5]. Maraknya diskusi publik dengan tagar tersebut, seperti tercatat 2.704 cuitan dalam periode singkat di awal tahun 2025 [6], membuat analisis opini secara manual menjadi tidak efisien dan rawan subjektivitas. Oleh karena itu, diperlukan pendekatan komputasional yang sistematis dan terukur untuk memahami persepsi publik terhadap fenomena ini.

Analisis sentimen, atau opinion mining, adalah pendekatan komputasional yang digunakan untuk mengidentifikasi dan mengkategorikan opini dalam teks ke dalam kelas sentimen seperti positif atau negatif [7], [8]. Salah satu algoritma yang populer dan efektif untuk tugas klasifikasi teks adalah *Naïve Bayes Classifier* (NBC). Algoritma ini bekerja berdasarkan *Teorema Bayes* dengan asumsi independensi antar fitur. Meskipun sederhana, NBC terbukti memiliki akurasi yang baik, cepat dalam pemrosesan, dan efektif untuk data berdimensi tinggi seperti data dari media sosial [9]. Keberhasilan algoritma ini juga ditunjukkan dalam beberapa penelitian terdahulu, seperti analisis sentimen terhadap Pilkada 2024 yang mencapai akurasi 77% [10], serta analisis sentimen pada fenomena #KaburAjaDulu yang menghasilkan akurasi 75% [11].

Penelitian ini bertujuan untuk Menganalisis sentimen publik di media sosial X terhadap tren #KaburAjaDulu dengan melakukan klasifikasi komentar menjadi kategori positif dan negatif menggunakan Algoritma *Naïve Bayes*. Mengevaluasi akurasi dan performa model klasifikasi yang dihasilkan dari penerapan metode KDD dalam pengolahan persepsi publik tersebut.

2. TINJAUAN PUSTAKA

2.1 Text Mining

Text mining merupakan proses pengolahan data teks untuk menemukan informasi atau pola yang berguna dari kumpulan dokumen yang tidak terstruktur. Teknik ini melibatkan beberapa tahapan seperti *preprocessing*, transformasi teks, serta penerapan algoritma

klasifikasi untuk memperoleh informasi yang relevan dari data teks [12].

2.2 Analisis Sentimen

Analisis sentimen adalah teknik dalam *text mining* yang digunakan untuk mengidentifikasi dan mengklasifikasikan opini atau perasaan yang terdapat dalam suatu teks ke dalam kategori tertentu seperti positif, negatif, atau netral [13]. Analisis sentimen banyak digunakan untuk mengetahui persepsi publik terhadap suatu produk, layanan, kebijakan, atau fenomena sosial.

2.3 Algoritma Naïve Bayes

Naïve Bayes merupakan algoritma klasifikasi berbasis probabilitas yang menggunakan prinsip Teorema Bayes dengan asumsi independensi antar fitur [14]. Algoritma ini sering digunakan dalam klasifikasi teks karena memiliki kemampuan yang baik dalam memproses data dengan dimensi tinggi serta membutuhkan waktu komputasi yang relatif cepat.

Rumus dasar *Naïve Bayes* dapat dituliskan sebagai berikut:

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)} \quad (1)$$

Dimana:

- C = kelas sentimen
- X = fitur atau kata dalam dokumen
- P(C|X) = probabilitas kelas C dengan kondisi X
- P(X|C) = probabilitas X dengan kondisi C
- P(C) = probabilitas awal kelas C
- P(X) = probabilitas awal X

2.4 Penelitian Terdahulu

Penelitian ini merujuk pada sejumlah studi terdahulu yang relevan dengan penerapan metode *Naïve Bayes* berdasarkan kata kunci yang digunakan untuk dianalisa sentimennya, biasanya sentimen yang akan dianalisis yaitu sentimen positif dan sentimen negatif. Berikut beberapa referensi jurnal sebagai tolak ukur penelitian serta sebagai perbandingan dapat dilihat pada Tabel 2.1.

Tabel 2. 1 Penelitian Terdahulu

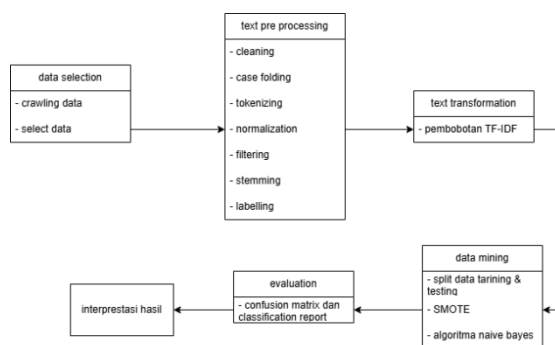
No	Penulis (Tahun)	Judul	Metode	Hasil
1	Ni Made Ayu Juli Astari,	Analisis Sentimen Dokumen	K-NN, Naïve Bayes,	NB memperoleh akurasi

	Dewa Gede Hendra Divanaya, Gede Indrawan (2020)	Twitter Mengenai Dampak Virus Corona Menggunakan Metode Naive Bayes Classifier	Decision Tree	tertinggi (84,08%), mengungguli K-NN (83,38%) dan DT (81,09%)
2	Franly Salmon Pattiiha, Hendery (2022)	Analisis Sentimen Pelanggan Tokopedia Menggunakan Metode Naive Bayes Classifier	NBC	Akurasi mencapai 95,10% pada klasifikasi sentimen positif dan negatif dari ulasan pengguna
3	Sifa Melina Salsabila, Aang Alim Murtopo, Nurul Fadhilah (2022)	Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification	NBC	Akurasi 87,34% dengan dominasi sentimen negatif terhadap kebijakan pemerintah

3. METODE PENELITIAN

Penelitian ini menggunakan metodologi *Knowledge Discovery in Databases (KDD)* sebagai kerangka kerja utama. KDD merupakan proses sistematis untuk menemukan pengetahuan yang sah, baru, dan bermanfaat dari kumpulan data berukuran besar [16]. Tahapan KDD yang diterapkan dalam penelitian ini meliputi seleksi data, *pre-processing*, transformasi data, data mining, dan evaluasi.

Berdasarkan prosedur KDD yang telah dijelaskan, dapat disusun langkah-langkah penelitian dengan tahapan sebagai berikut:



Gambar 3. 1 Rancangan Alur Penelitian

3.1 Pengumpulan dan Seleksi Data

Objek penelitian ini adalah tweet berbahasa Indonesia yang membahas tren #KaburAjaDulu di media sosial X. Data dikumpulkan melalui teknik *crawling* menggunakan *library tweet-harvest* pada lingkungan *Google Colaboratory* dengan bahasa pemrograman *Python*. Kata kunci yang digunakan adalah "#KaburAjaDulu" dan "#kaburajadulu" dalam rentang waktu 1 Desember 2024 hingga 1 Desember 2025. Dari proses ini, terkumpul 3.102 data mentah. Selanjutnya, dilakukan seleksi data dengan menghapus tweet duplikat dan atribut yang tidak relevan (seperti *ID*, *URL*, *username*), sehingga menyisakan 2.135 data unik yang siap diproses lebih lanjut.

3.2. Pelabelan Data

Proses pelabelan sentimen dilakukan dengan dua pendekatan untuk tujuan komparasi:

- 1. Pelabelan Manual:** Data diberikan label secara manual oleh seorang ahli bahasa Indonesia ke dalam dua kategori, yaitu positif dan negatif.
- 2. Pelabelan *Lexicon-Based*:** Data dilabeli secara otomatis menggunakan kamus *InSet Lexicon* berbahasa Indonesia. Label ditentukan berdasarkan nilai *polarity score*, di mana skor ≥ 1 dikategorikan sebagai positif dan skor ≤ -1 sebagai negatif.

3.3. Text Pre-processing

Tahap ini bertujuan membersihkan dan menyiapkan data teks mentah menjadi data terstruktur agar siap untuk proses pemodelan. Tahapan *pre-processing* meliputi:

- 1. Cleaning:** Menghapus elemen non-teks seperti tautan (URL), angka, simbol, tanda baca, emoji, dan spasi berlebih.
- 2. Case Folding:** Mengubah seluruh huruf dalam teks menjadi huruf kecil (*lowercase*).
- 3. Normalization:** Memperbaiki kata tidak baku, singkatan, atau typo menjadi kata baku sesuai kamus yang telah disusun.
- 4. Tokenizing:** Memecah kalimat menjadi kata-kata penyusunnya (*token*).
- 5. Filtering (Stopword Removal):** Menghapus kata-kata yang tidak memiliki makna signifikan menggunakan *library NLTK* yang dilengkapi dengan daftar *stopword* khusus.

6. **Stemming**: Mengubah kata berimbuhan menjadi kata dasar menggunakan library Sastrawi untuk bahasa Indonesia.

3.4. Transformasi Data (Pembobotan TF-IDF)

Setelah data bersih, dilakukan transformasi data teks menjadi representasi numerik menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Metode ini menghitung bobot setiap kata berdasarkan seberapa sering kata tersebut muncul dalam sebuah dokumen (*Term Frequency*) dan seberapa jarang kata tersebut muncul di seluruh dokumen (*Inverse Document Frequency*).

3.5. Pemodelan dengan Algoritma Naïve Bayes dan Penanganan Ketidakseimbangan Kelas

Tahap ini adalah inti dari proses KDD, di mana data yang telah ditransformasi digunakan untuk membangun model klasifikasi.

1. **Split Data**: Data dibagi menjadi data latih (*training*) dan data uji (*testing*) dengan tiga skenario rasio pembagian yang berbeda, yaitu 90:10, 80:20, dan 70:30.
2. **Handling Imbalanced Data dengan SMOTE**: Untuk mengatasi dominasi salah satu kelas sentimen, diterapkan metode *Synthetic Minority Over-sampling Technique* (SMOTE) yang bekerja dengan membuat sampel sintetis baru pada kelas minoritas.
3. **Klasifikasi dengan Naïve Bayes**: Klasifikasi sentimen dilakukan menggunakan algoritma Multinomial *Naïve Bayes* yang diimplementasikan dengan library *scikit-learn* di *Python*.

3.6. Evaluasi Model

Kinerja model klasifikasi dievaluasi menggunakan confusion matrix. Metrik evaluasi yang dihitung meliputi:

- **Akurasi**: Proporsi prediksi benar dari keseluruhan data.
- **Presisi**: Proporsi prediksi positif benar dari keseluruhan data yang diprediksi positif.
- **Recall** (Sensitivitas): Proporsi data positif yang berhasil diprediksi dengan benar.
- **F1-Score**: Rata-rata harmonik antara presisi dan recall.

4. HASIL DAN PEMBAHASAN

4.1 Hasil Penelitian

Penelitian ini menganalisis sentimen publik terhadap tren #KaburAjaDulu di media sosial X dengan menerapkan metodologi *Knowledge Discovery in Databases* (KDD) dan algoritma *Naïve Bayes*. Dari proses *crawling* data menggunakan kata kunci "#KaburAjaDulu" dan "#kaburajadulu" pada periode 1 Desember 2024 hingga 1 Desember 2025, diperoleh 3.102 data mentah. Setelah melalui tahap seleksi dan pembersihan data duplikat, diperoleh 2.135 data unik yang siap diproses.

4.2. Hasil Pre-processing dan Pelabelan Data

Tahap *pre-processing* yang terdiri dari *cleaning*, *case folding*, *normalisasi*, *tokenizing*, *filtering*, dan *stemming* berhasil membersihkan data dari elemen-elemen yang tidak relevan seperti tautan, angka, simbol, dan tanda baca. Data yang telah bersih kemudian dilabeli menggunakan dua pendekatan. Hasil distribusi sentimen dari kedua pendekatan pelabelan disajikan pada Tabel 1.

Tabel 1. Distribusi Sentimen Data #KaburAjaDulu

No	Pelabelan Data	Sentimen Positif	Sentimen Negatif
1	Manual	340	1.790
2	<i>Lexicon-Based</i>	358	1.777

Berdasarkan Tabel 1, sentimen negatif mendominasi pembahasan tren #KaburAjaDulu di kedua pendekatan pelabelan. Pada pelabelan manual, sentimen negatif mencapai 1.790 data (84%), sementara pada pelabelan *lexicon-based* sebanyak 1.777 data (83%). Hal ini mengindikasikan bahwa publik lebih banyak mengekspresikan kekecewaan, kritik, dan keresahan terhadap isu tersebut dibandingkan dukungan atau harapan.

4.3. Hasil Pemodelan dan Evaluasi

Data hasil *pre-processing* ditransformasi menggunakan metode TF-IDF untuk memperoleh bobot kata. Selanjutnya, dilakukan

pemodelan dengan algoritma *Naïve Bayes* pada enam skenario uji coba. Skenario tersebut mengombinasikan pendekatan pelabelan (manual dan *lexicon-based*), rasio pembagian data latih-uji (90:10, 80:20, 70:30), serta penerapan metode SMOTE untuk menangani ketidakseimbangan kelas. Hasil evaluasi model menggunakan *confusion matrix* dirangkum dalam Tabel 2 dan Tabel 3.

Tabel 2. Performa Model *Naïve Bayes* dengan Pelabelan Manual dan SMOTE

No	Metrik	Skenario 1 (90:10)	Skenario 2 (80:20)	Skenario 3 (70:30)
1	Accuracy	71%	74%	74%
2	Precision	32%	35%	35%
3	Recall	59%	58%	61%
4	F1-Score	41%	44%	45%

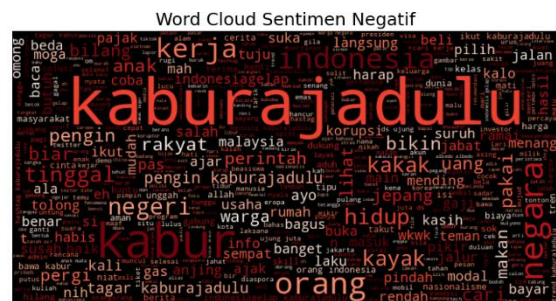
Tabel 3. Performa Model *Naïve Bayes* dengan Pelabelan *Lexicon-Based* dan SMOTE

No	Metrik	Skenario 4 (90:10)	Skenario 5 (80:20)	Skenario 6 (70:30)
1	Accuracy	71%	73%	74%
2	Precision	35%	33%	34%
3	Recall	61%	55%	60%
4	*F1-Score*	44%	41%	43%

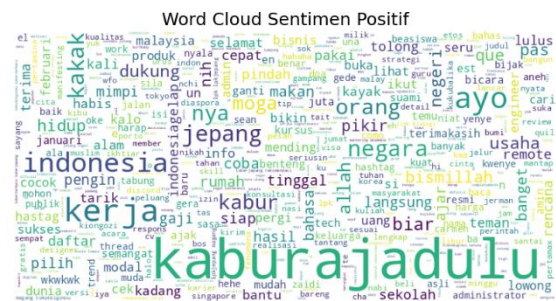
Hasil menunjukkan bahwa performa terbaik secara keseluruhan dicapai pada Skenario 3 (pelabelan manual dengan rasio data 70:30) dengan akurasi sebesar 74%, *recall* 61%, dan *F1-score* 45%. Pada pendekatan *lexicon-based*, performa tertinggi juga diperoleh pada rasio 70:30 (Skenario 6) dengan akurasi 74%. Nilai akurasi yang relatif stabil di berbagai skenario menunjukkan bahwa algoritma *Naïve Bayes* cukup konsisten dalam mengklasifikasikan sentimen.

4.4. Visualisasi Data

Visualisasi *word cloud* digunakan untuk melihat kata-kata yang paling dominan muncul dalam setiap kategori sentimen. Pada sentimen positif, kata-kata seperti "kerja", "ayo", "jepang", dan "peluang" mendominasi, yang mengindikasikan adanya motivasi dan harapan untuk mencari kehidupan yang lebih baik di luar negeri. Sebaliknya, pada sentimen negatif, kata-kata seperti "negara", "orang", "pajak", "salah", dan "takut" lebih dominan, mencerminkan kekecewaan, kritik terhadap kondisi dalam negeri, serta kekhawatiran akan dampak sosial dan ekonomi dari fenomena ini.



Gambar 3. 2 *Word Cloud* Sentimen Negatif



Gambar 2. *Word Cloud* Sentimen Positif

4.5. Pembahasan

Penelitian ini berhasil mengidentifikasi bahwa sentimen negatif mendominasi percakapan publik seputar tren #KaburAjaDulu di media sosial X. Temuan ini sejalan dengan penelitian sebelumnya oleh Aftoni et al. (2025) yang juga menemukan dominasi sentimen negatif pada tagar yang sama [11]. Hal ini mengonfirmasi bahwa fenomena #KaburAjaDulu lebih banyak digunakan sebagai wadah ekspresi ketidakpuasan dan keresahan terhadap situasi sosial-ekonomi-politik di Indonesia, dibandingkan sebagai ajakan positif semata. Kata-kata seperti "negara", "pajak", dan "salah" yang muncul dalam sentimen negatif memperkuat indikasi

bahwa kritik publik tertuju pada tata kelola dan kebijakan dalam negeri.

Dari sisi performa model, algoritma *Naïve Bayes* menunjukkan kinerja yang cukup efektif dengan nilai akurasi mencapai 74%. Hasil ini konsisten dengan penelitian-penelitian terdahulu yang menggunakan algoritma serupa untuk analisis sentimen, seperti pada sentimen Pilkada 2024 [10] dan isu-isu sosial lainnya. Stabilitas nilai akurasi di berbagai rasio pembagian data mengindikasikan bahwa model yang dibangun cukup *robust*.

Meskipun demikian, nilai presisi yang relatif rendah (di kisaran 32-35%) menunjukkan bahwa model masih sering salah dalam mengklasifikasikan sentimen positif. Hal ini dapat disebabkan oleh dominasi data negatif yang sangat besar (ketidakseimbangan kelas) dan kompleksitas bahasa di media sosial seperti sarkasme, ironi, atau konteks ganda yang sulit ditangkap oleh model, meskipun SMOTE telah diterapkan.

Perbandingan antara dua metode pelabelan menunjukkan bahwa pelabelan manual oleh ahli bahasa (Skenario 3) menghasilkan performa yang sedikit lebih unggul dan stabil dibandingkan pelabelan otomatis berbasis leksikon (Skenario 6). Hal ini menegaskan bahwa kualitas label data merupakan faktor krusial yang memengaruhi kinerja model klasifikasi sentimen. Pelabelan manual lebih mampu menangkap nuansa bahasa dan konteks, sementara pendekatan *lexicon-based* cenderung menghasilkan label yang lebih *noisy* karena keterbatasannya dalam memahami konteks kalimat secara utuh.

5. KESIMPULAN

5.1 Kesimpulan

Berdasarkan hasil analisis sentimen publik terhadap tren #KaburAjaDulu di media sosial X menggunakan algoritma *Naïve Bayes*, dapat ditarik kesimpulan sebagai berikut:

- Penelitian ini berhasil melakukan analisis sentimen publik di media sosial X terhadap tren #KaburAjaDulu dengan mengklasifikasikan komentar ke dalam kategori positif dan negatif menggunakan Algoritma *Naïve Bayes*. Hasil klasifikasi menunjukkan bahwa sentimen negatif mendominasi sebesar 83-84% (1.790 data manual dan 1.777 data leksikon) dan sentimen positif sebesar 16-17% (340 data

manual dan 358 data leksikon). Hal ini mengindikasikan bahwa isu #KaburAjaDulu lebih banyak mendapatkan sentimen negatif yang merefleksikan keresahan dan kritik terhadap kondisi sosial-ekonomi-politik di Indonesia.

- Evaluasi performa model melalui penerapan metodologi KDD menunjukkan bahwa algoritma *Naïve Bayes* cukup efektif dengan akurasi terbaik 74% pada skenario 70:30. Pelabelan manual menghasilkan performa lebih unggul (*F1-score* 45%) dibandingkan pelabelan *lexicon-based* (*F1-score* 43%), yang membuktikan bahwa kualitas label data merupakan faktor penting dalam menentukan kinerja model klasifikasi sentimen.

5.2 SARAN

Beberapa saran yang dapat diterapkan untuk perbaikan dan pengembangan penelitian adalah sebagai berikut:

- Penelitian selanjutnya dapat mempertimbangkan perluasan objek kajian pada isu-isu lanjutan terkait migrasi tenaga kerja serta mengkolaborasikan algoritma *Naïve Bayes* dengan metode lain seperti *Support Vector Machine*, *Random Forest*, *K-Nearest Neighbor*, atau *Decision Tree* untuk perbandingan performa.
- Penelitian mendatang dapat menambahkan label sentimen netral, positif kuat, dan negatif kuat untuk mendapatkan pemahaman yang lebih mendalam mengenai spektrum opini publik.
- Implementasi metode seleksi fitur seperti *N-Gram*, *Particle Swarm Optimization*, *Chi Square*, atau *Stable Mutation Jump Strategy* dapat diterapkan untuk meningkatkan kualitas model klasifikasi.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Ibu Hannie, S.Kom, MMSI., CEC dan Bapak Azhari Ali Ridha, S.Kom., MMSI. selaku dosen pembimbing, serta seluruh pihak yang telah memberikan dukungan dalam penelitian ini.

DAFTAR PUSTAKA

- [1] A. D. Riyanto, "Data Digital Indonesia 2025," 2025. [Online]. Tersedia: <https://andi.link/hootsuite-we-are-social-data-digital-indonesia-2025/>
- [2] F. Ramadhani, I. D. Arianto, "Gema Suara Rakyat dalam Media Sosial (Analisis Jaringan Komunikasi pada #KaburAjaDulu di Twitter (X))," vol. 8, no. X, 2025.
- [3] N. Abelia, K. J. Farah, D. Wibisono, and I. Mahmud, "Dampak Framing Tagar #Kaburajadulu Terhadap Opini Publik dan Kebijakan Sosial di Indonesia," 2025.
- [4] P. Sari, M. Julianty, D. Mirza, N. Nafilah, and S. Zulfachrinal, "Menghadapi Ancaman Nasionalisme Disintegrasi Bangsa di Tengah Trend Kabur Aja Dulu," vol. 3, 2025.
- [5] Z. P. Ajie, "Analisis Sentimen Negatif Publik pada Tagar #KaburAjaDulu di Media Sosial 'X'," vol. 6, no. 3, 2025.
- [6] Muda Bicara, "Kabur Aja Dulu Trending di Medsos: Gambaran Frustrasi Generasi Muda terhadap Kondisi Indonesia," 2025. [Online]. Tersedia: <https://www.mudabicara.id/kajian/kabur-aja-dulu-trending-di-medsos-gambaran-frustrasi-generasi-muda-terhadap-kondisi-indonesia/>
- [7] C. C. Aggarwal, *Machine Learning for Text*, 2nd ed. Springer Nature Switzerland AG, 2022. <https://doi.org/10.1007/978-3-030-96623-2>
- [8] Samsir, Ambiyar, U. Verawardina, F. Edi, and R. Watianthos, "Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes," *Jurnal Media Informatika Budidarma*, vol. 5, pp. 157–163, 2021.
- [9] L. N. Idris, F. N. I. S. Kumala, L. N. Mustofa, F. N. Y. Aril, L. N. Salihi, and F. N. I. Abraham, "Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (SVM)," vol. 5, pp. 32–35, 2023.
- [10] M. J. Siddiq, S. Jayasri, A. Suhendi, T. Hidayat, R. Rizky, U. Islam, S. Yusuf, J. Raya, and L. Cikaliung, "Masyarakat Terhadap Pilkada 2024 Di Media Sosial Twitter Menggunakan Algoritma Naïve," vol. 13, no. 2, pp. 609–622, 2025.
- [11] Aftoni et al., "Analisis sentimen terhadap tagar #kaburajadulu di media sosial x menggunakan metode naïve bayes," 2025.
- [12] E. Buulolo, *Data Mining untuk Perguruan Tinggi*. Deepublish Publisher, 2020.
- [13] Rohmatun and A. Baita, "Analisis Sentimen Masyarakat Terhadap Vaksin Covid-19 Pada Media Sosial Twitter Menggunakan Metode Naïve Bayes," *Jurnal Teknologi Sistem Informasi dan Aplikasi*, vol. 4, no. 2, 2025.
- [14] Mola et al., "Penerapan Algoritma Naïve Bayes untuk Analisis Sentimen Ulasan Aplikasi MyPertamina," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 3, no. 1, 2025.
- [15] D. Duei Putri, G. F. Nama, and W. E. Sulistiono, "Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier," *JITET*, vol. 10, no. 1, Jan. 2022