INOMATEC: Jurnal Inovasi dan Kajian Multidisipliner Kontemporer
Vol. 01, No. 04, Tahun 2026, Hal. 533-542, ISSN: 3123-5573 (Online)

Penerapan Algoritma Multinomial Naïve Bayes untuk Klasifikasi Sentimen Pada Hashtag #Kaburajadulu di Media Sosial X

Muhammad Ni'mad Rahmatullah[1], Moh. Dasuki[2], Habibatul Azizah Al Faruq[3]
[1,2,3]Program Studi Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember
E-mail: m.nikmad.r@gmail.com[1], moh.dasuki22@unmuhjember.ac.id[2], habibatulazizah@unmuhjember.ac.id[3]

Article Info

Article history:

Received February 04, 2026
Revised February 06, 2026
Accepted February 12, 2026

Keywords:

#Kaburajadulu, Sentiment Analysis, Multinomial Naïve Bayes, Confusion Matrix

ABSTRACT

This study aims to analyze the sentiment of social media X users toward issues discussed through the hashtag #KaburAjaDulu by applying the Multinomial Naïve Bayes algorithm. The data were collected through a crawling process from January 10 to March 3, 2025, resulting in 1,164 comments consisting of 724 positive and 440 negative sentiments. The dataset was divided into training and testing data with an 80:20 ratio and processed through several preprocessing stages, including text cleaning, tokenization, stopword removal, stemming, and term weighting using the TF–IDF method. Hyperparameter optimization was performed using Random Search, while model performance was evaluated using K-Fold Cross Validation. The experimental results show that the model achieved an accuracy of 72%, with precision, recall, and F1-score values of 76%, 83%, and 79% for the positive sentiment class, and 63%, 53%, and 57% for the negative sentiment class. These findings indicate that the proposed approach is able to classify user sentiment effectively and can be used to understand public opinion trends on social media.

This is an open access article under the CC BY-SA license.

Article Info

Article history:

Received February 04, 2026
Revised February 06, 2026
Accepted February 12, 2026

Kata Kunci:

#Kaburajadulu, Analisis Sentiment, Multinomial Naïve Bayes, Confusion Matrix

ABSTRAK

Penelitian ini bertujuan untuk mengkaji sentimen pengguna media sosial *X* terhadap isu yang ramai diperbincangkan melalui hashtag *#KaburAjaDulu* dengan menerapkan algoritma *Multinomial Naïve Bayes*. Data diperoleh melalui proses *crawling* pada periode 10 Januari hingga 3 Maret 2025 dan menghasilkan sebanyak 1.164 komentar yang terdiri atas 724 sentimen positif dan 440 sentimen negatif. Data selanjutnya dibagi menjadi data latih dan data uji dengan perbandingan 80:20 serta diproses melalui tahapan pembersihan teks, tokenisasi, penghapusan *stopword*, *stemming*, dan pembobotan kata menggunakan metode *TF–IDF*. Optimasi parameter dilakukan menggunakan metode *Random Search*, sedangkan evaluasi kinerja model menggunakan *K-Fold Cross Validation*. Hasil pengujian menunjukkan bahwa model mencapai tingkat akurasi sebesar 72%, dengan nilai presisi, *recall*, dan *F1-score* pada sentimen positif masing-masing sebesar 76%, 83%, dan 79%, serta pada sentimen negatif sebesar 63%, 53%, dan 57%. Temuan ini menunjukkan bahwa

pendekatan yang digunakan mampu mengklasifikasikan sentimen pengguna secara cukup efektif dan dapat dimanfaatkan untuk memahami kecenderungan opini publik di media sosial.

This is an open access article under the CC BY-SA license.

Corresponding Author:
Muhammad Ni'mad Rahmatullah
Universitas Muhammadiyah Jember
Email: m.nikmad.r@gmail.com

PENDAHULUAN

Natural Language Processing (NLP) merupakan salah satu cabang dari *Artificial Intelligence* (AI) yang berfokus pada pemrosesan dan pemahaman bahasa alami yang digunakan manusia dalam komunikasi sehari-hari (Rahmadani & Tasrif 2023). Penerapan NLP menjadi semakin penting seiring dengan meningkatnya interaksi digital di media sosial, di mana bahasa yang digunakan cenderung singkat, tidak terstruktur, dan bersifat informal. Dengan kemampuannya dalam mengenali pola bahasa, konteks, serta emosi, NLP memungkinkan sistem komputer untuk menganalisis opini dan sentimen pengguna secara lebih akurat, khususnya dalam percakapan publik di ruang digital (Rivaldi et al. 2024).

Media sosial *X* merupakan salah satu platform yang berperan besar dalam membentuk dan mencerminkan dinamika sosial masyarakat. Platform ini menjadi ruang terbuka bagi pengguna untuk menyampaikan opini, kritik, dan aspirasi terhadap berbagai isu yang sedang berkembang. Salah satu fenomena yang menarik perhatian publik adalah viralnya tagar *#KaburAjaDulu*, yang digunakan untuk mengekspresikan keresahan terkait kondisi ketenagakerjaan, tekanan sosial, serta keinginan mencari peluang yang lebih baik di luar negeri. Munculnya tagar ini mencerminkan adanya kecenderungan opini kolektif yang perlu dikaji secara sistematis sebagai bagian dari pemahaman terhadap dinamika sosial (Winahyu & Suharjo 2021).

Melimpahnya data opini yang tersebar di media sosial menjadikan platform tersebut sebagai sumber data yang potensial untuk analisis sentimen. Analisis sentimen dapat memberikan gambaran mengenai persepsi dan sikap masyarakat terhadap suatu isu secara cepat dan aktual. Untuk mengolah data teks dalam jumlah besar, khususnya teks pendek seperti cuitan, diperlukan metode klasifikasi yang efektif. Algoritma *Multinomial Naïve Bayes* dipilih karena mampu mengolah data berbasis frekuensi kata dan telah terbukti efektif dalam berbagai penelitian analisis sentimen (Hasri & Alita 2022).

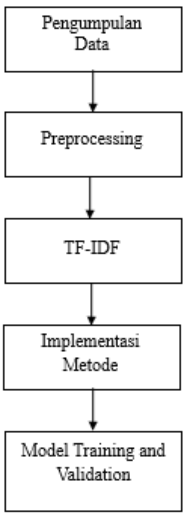
Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap tagar *#KaburAjaDulu* di media sosial *X* menggunakan algoritma *Multinomial Naïve Bayes*. Analisis dilakukan terhadap komentar berbahasa Indonesia yang dikumpulkan dalam periode tertentu dan diklasifikasikan ke dalam dua kategori sentimen, yaitu positif dan negatif. Kinerja model dievaluasi menggunakan metrik *accuracy, precision, recall,*

dan *F1-score* dengan validasi *K-Fold Cross Validation*. Hasil penelitian diharapkan dapat memberikan gambaran mengenai kecenderungan opini publik serta menjadi referensi bagi pengembangan kajian analisis sentimen dan perumusan kebijakan yang lebih responsif terhadap aspirasi masyarakat.

.

METODE PENELITIAN

Penelitian ini bertujuan untuk menganalisis dan mengklasifikasikan sentimen dari teks yang memuat hashtag *#KaburjaDulu* di media sosial *X*. Pendekatan kuantitatif digunakan



Gambar 2. 1 alur penelitian

karena fokus penelitian adalah pengolahan data numerik dan pengujian hipotesis melalui model statistik serta algoritma *machine learning*. Alur penelitian dilakukan sebagai berikut:

1.1. Pengumpulan Data

Proses pengumpulan data dilakukan menggunakan library TWScrape 0.15.0 di Python pada media sosial X, dengan mengambil retweet yang menggunakan hashtag #KaburAjaDulu sebagai indikator. Tahap ini dijalankan di platform JupyterLab. Data dikumpulkan dari 10 Januari hingga 3 Maret 2025, periode saat hashtag tersebut menjadi trending topic di media sosial X.

1.2. Preproccesing Data

Preprocessing teks merupakan tahap awal dalam pengolahan data teks yang bertujuan untuk membersihkan dan merapikan data mentah (Agung et al. 2023). Pada tahap ini, data yang awalnya tidak terstruktur diubah menjadi bentuk yang lebih teratur dan mudah diproses oleh sistem. Proses ini meliputi pemilihan elemen penting serta transformasi konten agar lebih relevan dan siap dianalisis, sehingga meningkatkan kualitas data dan efisiensi analisis (Aisah, Irawan & Suprapti 2023).

1. Labeling adalah tahap penting dalam preprocessing data teks, terutama pada metode supervised learning seperti klasifikasi sentimen. Tahap ini bertujuan memberikan label pada setiap data sehingga dapat dikategorikan, misalnya ke dalam kelas positif atau negatif (Virgiansyah, Stephanie & Rizky Pribadi 2024).

2. *Cleaning* data bertujuan membersihkan data untuk meningkatkan kualitasnya. Proses ini meliputi penghapusan elemen yang tidak relevan, seperti hashtag, simbol, dan emoji yang tidak memiliki makna penting, sehingga data menjadi lebih siap untuk dianalisis (Khairunnisa, Adiwijaya & Faraby 2021).
3. *Case folding* adalah tahap *preprocessing* teks yang mengubah seluruh huruf menjadi huruf kecil. Tujuannya untuk memastikan konsistensi penulisan sehingga perbedaan antara huruf kapital dan kecil tidak memengaruhi analisis, sehingga proses pengolahan data menjadi lebih efektif (Albab, P. & Fawaiq 2023).
4. *Stopword removal* adalah tahap *preprocessing* teks yang bertujuan menghapus kata-kata umum yang tidak memiliki nilai signifikan dalam analisis, seperti kata penghubung, kata depan, dan kata bantu ("dan", "yang", "di"). Proses ini menyederhanakan data sehingga algoritma klasifikasi dapat lebih fokus pada kata-kata yang bermakna (Ajis, Utomo & Barkah 2025).
5. *Stemming* adalah proses mengubah kata berimbuhan menjadi bentuk dasarnya dengan menghapus awalan, akhiran, sisipan, atau kombinasi imbuhan lainnya (Saraswati et al. 2025).
6. *Tokenization* adalah proses memecah teks menjadi unit-unit kecil yang disebut token atau kata (Rifaldi, Abdul Fadlil & Herman 2023). Tahap ini juga menghapus tanda baca, angka, *mention*, hashtag, dan URL, sehingga memudahkan perhitungan frekuensi kata dan mendukung analisis data secara lebih efektif (Krawczyk, Probierz & Kozak 2024).

1.3 TF-IDF

TF–IDF digunakan untuk memberikan bobot pada kata, sehingga dapat menilai seberapa penting suatu kata dalam sebuah dokumen atau dataset. *Term Frequency (TF)* menghitung frekuensi kemunculan kata dengan asumsi bahwa kata yang muncul lebih sering memiliki kontribusi lebih besar terhadap makna dokumen (Andrian et al. 2023).

$$TF - IDF(t,d) = tf(t,d) \times log(N \,/\, df(t)) \quad (1)$$

1.4 *Multinomial Naïve Bayes*

Multinomial Naïve Bayes adalah tipe *Naïve Bayes* yang cocok untuk analisis sentimen komentar, karena setiap dokumen dianggap sebagai *vektor* kata dan setiap nilai kata merepresentasikan frekuensi kemunculannya (Agustina et al. 2022). Probabilitas setiap kelas dihitung berdasarkan frekuensi kata dalam dokumen (Ionendri, Feri Candra & Afdi Rizal 2025).

$$c_{map} = P(c) \prod_{k=1}^{n_d} P(t_k|c) \quad (2)$$

1.5 Laplace Smooting

Laplace Smoothing adalah teknik sederhana namun penting dalam klasifikasi teks menggunakan Naïve Bayes. Metode ini mengatasi masalah probabilitas nol yang muncul ketika suatu kata tidak ditemukan dalam data latih untuk suatu kelas. Dengan menambahkan nilai 1 pada setiap frekuensi kata, Laplace Smoothing memastikan semua kata memiliki probabilitas lebih dari nol, sehingga model dapat mempertimbangkan semua kelas secara adil (Rizky, Aziz & Harianto 2024).

$$P(t_k|c) = \frac{N_{ct} + 1}{\sum_t N_{ct} + |V|} \quad (3)$$

1.6 Random Search

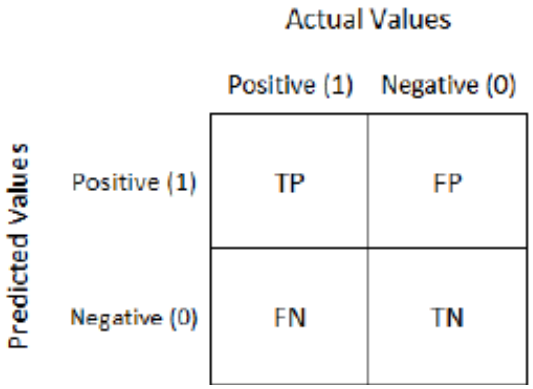
Random Search adalah teknik optimasi *hyperparameter* dalam *machine learning* yang memilih kombinasi parameter secara acak dari ruang parameter yang telah ditentukan sebelumnya (Rom, Jamil & Ibrahim 2024).

1.7 K-Fold Cross Validation

K-Fold Cross Validation adalah metode validasi yang membagi data menjadi beberapa *subset* untuk meminimalkan bias selama proses pelatihan dan pengujian (Lusiyanti et al. 2025).

1.8 Confusion Matrix

Pada tahap evaluasi, digunakan *Confusion Matrix*, yaitu tabel yang membandingkan hasil prediksi model dengan label asli. Tabel ini mencakup *true positive* (TP), *true negative* (TN), *false positive* (FP), dan *false negative* (FN). Dari informasi ini, metrik evaluasi seperti *accuracy*, *precision*, *recall*, dan *F1-score* dapat dihitung untuk menilai kinerja model secara menyeluruh, terutama pada data yang tidak seimbang. *Confusion Matrix* juga membantu



Gambar 2. 2 evaluasi matrix

mengidentifikasi jenis kesalahan model dan menjadi dasar perbaikan model selanjutnya(Sathyanarayanan 2024).

2.1 Spesifikasi Penulisan

Berikut spesifikai atau aturan penulisan jurnal JUSTINDO:

2.1.2 Penulisan Judul

Judul yang harus ditulis secara ringkas dan menggambarkan isi naskah (maksimal 12 kata), dalam judul hindari penulisan sub judul atau studi kasus. Judul artikel ditulis dalam bahasan Indonesia dan Inggris dengan format rata tengah dengan menggunakan huruf *Capitalize Each Word*, font arial ukuran 12 pt di cetak dengan huruf tebal dan spasi tunggal *(bold)* dengan spasi 1,0.

2.1.3 Penulisan Identitas Penulis

Nama penulis (tanpa gelar akademik); afiliasi penulis (prodi, fakultas, universitas); alamat email. Penulisan identitas penulis ditulis di bawah judul dengan rata tengah menggunakan font arial, 10 pt, spasi tunggal dan di cetak huruf miring *(italic)* dan terdapat penulis koresponden yang ditandai dengan symbol asterisk.

2.1.4 Penulisan Abstrak

Abstrak (150-250 kata) ditulis dalam bahasa Indonesia dan *Inggris*; kata kunci (minimal tiga buah dan maksimal 5 buah); abstrak ditulis menggunakan huruf *arial* dengan ukuran10pt spasi tunggal. Abstrak bahasa inggris di cetak huruf miring *(italic)*.

2.1.5 Penulisan Isi Artikel

Artikel utama ditulis dengan menggunakan huruf arial, ukuran 11 poin, spasi 1,0 dan paragraf rata kiri-kanan (justify). Artikel secara garis besar terdiri dari judul Pendahuluan, Metode Penelitian, Hasil dan Pembahasan, dan Kesimpulan. Penulisan judul ditulis dengan huruf kapital dengan didahului penomoran menggunakan angka. Gunakan *style heading* dalam *template* ini secara langsung. *Style* sudah diformat sedemikian rupa sehingga memberikan jarak *heading* yang sesuai dan level penomoran judul dan subjul.

HASIL DAN PEMBAHASAN

Setelah *preprocessing* yang sudah di jelaskan di metode penelitian, dilakukan visualisasi menggunakan *word cloud*. Visualisasi ini menampilkan sebaran dan frekuensi kata yang paling dominan, sehingga memberikan gambaran awal mengenai pola atau tema umum dalam data setelah tahapan *preprocessing* selesai.

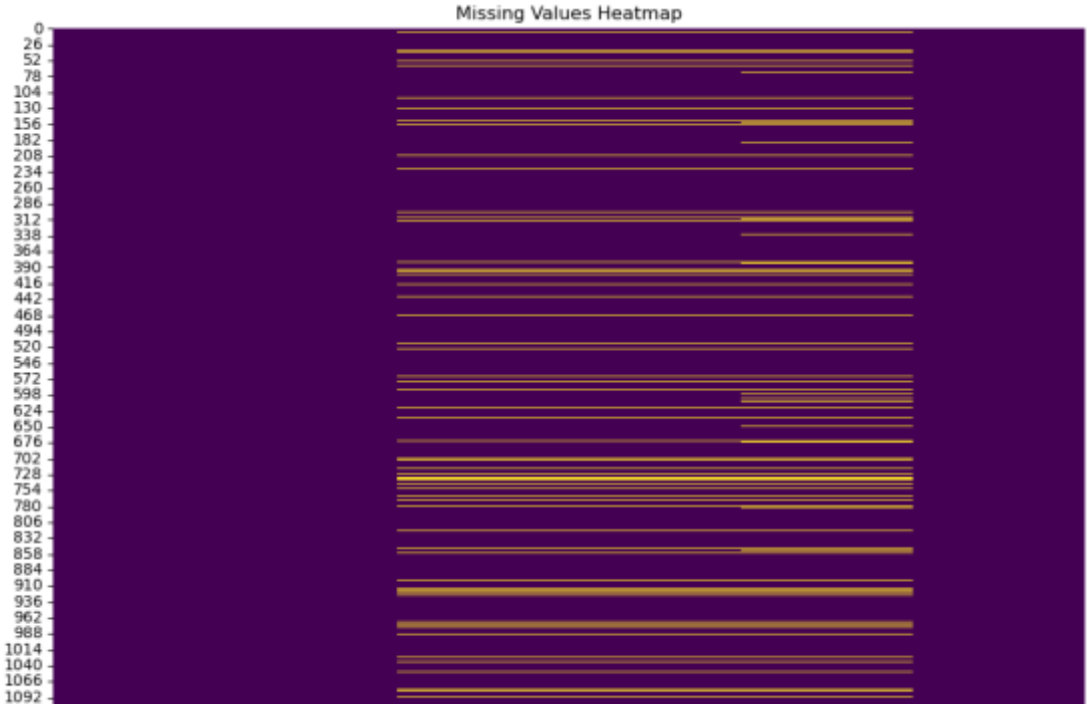
Gambar 3. 1 hasil Word Cloud

Dari hasil *word cloud* dapat dilihat terdapat beberapa kata yang tidak relevan atau tidak memiliki makna berarti sehingga perlu menambahkan cleaning secara manual.

```
sw_indo = stopwords.words("indonesian") + list(punctuation) + ["''"]
sw_indo.extend(["yg", "gw", "gue", "gak", "gk", "ga", "dg", "rt", "dgn", "ny", "d", 'klo',
                'kalo', 'amp', 'krn', 'nya', 'ni', 'nih', 'sih', 'bang', 'deh', 'dr', 'an',
                'si', 'tau', 'tdk', 'tuh', 'utk', 'ya', 'org', 'st', 'that', 'lg', 'so', 'gt', 'wa', 'ln', 'in',
                'jd', 'jg', 'jgn', 'sdh', 'aja', 'n', 's', 'f', 'of', 'tp', 'de', 'b', 'are', 'w',
                'nyg', 'hehe', 'pen', 'u', 'nan', 'loh', 'rt', 'is', 'isi', 'if', 'hi', 'bgt', 'lo',
                '&amp', 'yah' , 'the' , 'but' , 'they' , 'where' , 'to' , 'and' , 'this' , 'for', 'were' , 'ken'])
```

Gambar 3. 2 Code Cleaning

Setelah dilakukan pembersihan kata kemudian bisa dilihat beberapa dokumen yang kosong dan dilakukan penghapusan dokument kosong.



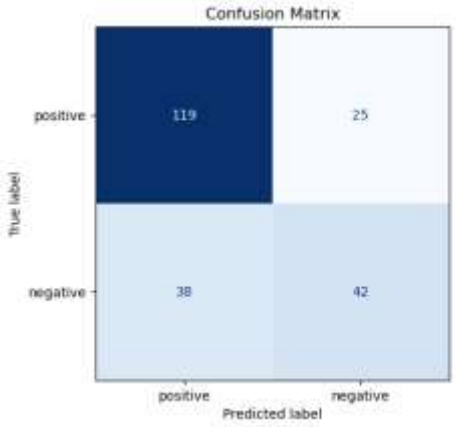
Gambar 3. 3 Visualisasi Dokument Kosong

Tahap selanjutnya adalah perhitungan metrik dengan pembagian data latih dan uji 80:20, menggunakan rumus dan metode sebelumnya, untuk menentukan model paling optimal melalui *K-Fold Cross Validation*.

Gambar 3. 4 Hasil K-fold Crossvalidation

Pada Gambar 3.4, kombinasi *hyperparameter* terbaik hasil *Random Search* adalah 0,1 dengan *fit_prior* = True. Nilai *test score* rata-rata yang diperoleh dari *K-Fold Cross Validation* 5 fold adalah 0,700967. Berikut ditunjukkan perhitungan untuk memperoleh nilai rata-rata pada proses *K-Fold Cross Validation*.

Tahap berikutnya adalah perhitungan menggunakan data uji untuk memperoleh hasil evaluasi model sebagai berikut:



	precision	recall	f1-score	support
0	0.76	0.83	0.79	144
1	0.63	0.53	0.57	80
accuracy			0.72	224
macro avg	0.69	0.68	0.68	224
weighted avg	0.71	0.72	0.71	224

Gambar 3. 5 Hasil Evaluasi Model

Berdasarkan evaluasi, model *Multinomial Naïve Bayes* menghasilkan 119 data sebagai *True Positive* (TP), 42 data sebagai *True Negative* (TN), 25 data sebagai *False Positive* (FP), dan 38 data sebagai *False Negative* (FN). Dari distribusi ini, diperoleh akurasi sebesar 72%, menunjukkan sekitar 72% prediksi diklasifikasikan dengan benar. Faktor yang memengaruhi hasil antara lain pembersihan data yang belum optimal, penggunaan bahasa *slang* di komentar platform *X*, serta banyaknya kata singkatan yang memengaruhi bobot kata. Untuk kelas positif, *precision*, *recall*, dan *F1-score* masing-masing 76%, 83%, dan 79%, menunjukkan kemampuan model yang baik dalam mengidentifikasi kelas positif. Sedangkan untuk kelas negatif, *precision*, *recall*, dan *F1-score* masing-masing 63%, 53%, dan 57%.

KESIMPULAN DAN SARAN

Dari 1.117 entri yang berhasil diolah, 718 berlabel positif dan 399 berlabel negatif, menunjukkan respons masyarakat terhadap tagar *#KaburAjaDulu* cenderung positif, meskipun distribusi komentar relatif seimbang. Model *Multinomial Naïve Bayes* mencapai akurasi 72%, dengan kinerja lebih baik pada kelas positif (*precision* 76%, *recall* 83%, *F1-score* 79%), sementara pada kelas negatif masih rendah (*precision* 63%, *recall* 53%, *F1-score* 57%). Hasil ini dipengaruhi oleh pembersihan data yang belum optimal, penggunaan bahasa *slang*, dan banyaknya kata singkatan, sehingga model perlu perbaikan untuk menghasilkan prediksi yang lebih seimbang.

Nomor dan judul tabel ditulis diposisi tengah. Tabel dinomori dengan angka arab sesuai dengan urutannya. Judul tabel ditulis dibagian atas tabel dengan cara *title case*, kecuali untuk

kata sambung dan kata depan. Ukuran huruf untuk judul tabel dan isi tabel adalah 8 pt (delapan). Sisi paling luar tabel tidak boleh melebihi batas margin halaman. Jarak baris yang digunakan antara tabel dengan kalimat diatasnya dan dibawahnya adalah 1 (satu) baris kosong. Tabel wajib menggunakan layout sesuai dengan Tabel 1 tanpa menggunakan garis lurus/vertikal. Setiap tabel harus diacu dalam tulisan dengan disertai nomor tabel dan diawali dengan huruf besar, misalnya Tabel 1.

DAFTAR PUSTAKA

Agung, A., Daniswara, A., Kadek, I. & Nuryana, D., 2023, 'Data Preprocessing Pola Pada Penilaian Mahasiswa Program Profesi Guru', *Journal of Informatics and Computer Science*, 05.

Agustina, N., Citra, D.H., Purnama, W., Nisa, C. & Kurnia, A.R., 2022, 'Implementasi Algoritma Naive Bayes untuk Analisis Sentimen Ulasan Shopee pada Google Play Store', *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 2(1), 47–54.

Aisah, I.S., Irawan, B. & Suprapti, T., 2023, 'Algoritma Support Vector Machine (SVM) Untuk Analisis Sentimen Ulasan Aplikasi Al Qur'an Digital', *Jurnal Mahasiswa Teknik Informatika*, 7(6).

Ajis, A.S., Utomo, F.S. & Barkah, A.S., 2025, 'Evaluasi Pengaruh Varian Daftar Stopword terhadap Kinerja Klasifikasi Teks Al-Qur'an dengan Support Vector Machine dan Backpropagation Neural Network', *Jurnal Pendidikan dan Teknologi Indonesia*, 5(7), 1867–1880.

Albab, M.U., P., Y.K. & Fawaiq, M.N., 2023, 'Optimization of the Stemming Technique on Text Preprocessing President 3 Periods Topic', *Jurnal Transformatika*, 20(2), 1–12.

Andrian, B.W., Tobing, F.A.T., Pane, I.Z. & Kusnadi, A., 2023, 'Implementation of Naïve Bayes Algorithm in Sentiment Analysis of Twitter Social Media Users Regarding Their Interest to Pay the Tax', *International Journal of Science, Technology & Management*, 4(6), 1733–1742.

Hasri, C.F. & Alita, D., 2022, 'Penerapan Metode Naive Bayes Classifier dan Support Vector Machine pada Aanalisis Sentimen Terhadap Dampak Virus Corona di Twitter', *Jurnal Informatika dan Rekayasa Perangkat Lunak*, 3(2), 145–160.

Ionendri, N.A., Feri Candra & Afdi Rizal, 2025, 'News Classification using Natural Language Processing with TF-IDF and Multinomial Naïve Bayes', *Journal of Applied Computer Science and Technology*, 6(1), 37–45.

Khairunnisa, S., Adiwijaya, A. & Faraby, S. Al, 2021, 'Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19)', *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(2), 406.

Krawczyk, N., Probierz, B. & Kozak, J., 2024, 'Towards AI-Generated Essay Classification Using Numerical Text Representation', *Applied Sciences (Switzerland)*, 14(21).

Lusiyanti, D., Musdalifah, S., Sahari, A. & Fajri, I. Al, 2025, 'Evaluasi Kinerja Algoritma Machine learning pada Dataset Skala Besar', *MathVision : Jurnal Matematika*, 7(1), 84–92.

Rahmadani, S. & Tasrif, E., 2023, 'Implementasi Natural Language Processing (NLP) pada Layanan Pengelolaan Surat Kantor Camat Bukit Barisan', *Jurnal Teknik Komputer dan Informatika*, 4(1), 19.

Rifaldi, D., Abdul Fadlil & Herman, 2023, 'Teknik Preprocessing Pada Text Mining Menggunakan Data Tweet "Mental Health"', *Decode: Jurnal Pendidikan Teknologi Informasi*, 3(2), 161–171.

Rivaldi, R.C., Wismarini, T.D., Lomba, J.T. & Semarang, J., 2024, 'Analisis Sentimen Pada Ulasan Produk Dengan Metode Natural Language Processing (NLP) (Studi Kasus Zalika Store 88 Shopee)', 17(1), 120–128.

Rizky, Y.A., Aziz, A. & Harianto, W., 2024, 'Implementasi Naive Bayes Dengan Menggunakan Metode Laplace Smoothing', *RAINSTEK: Jurnal Terapan Sains dan Teknologi*, 6(3), 164–172.

Rom, A.R.M., Jamil, N. & Ibrahim, S., 2024, 'Multi objective hyperparameter tuning via random search on deep learning models', *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 22(4), 956.

Saraswati, N.W.S., Yanti, C.P., Muku, I.D.M.K. & Dewi, D.A.P.R., 2025, 'Evaluation Analysis of the Necessity of Stemming and Lemmatization in Text Classification', *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, 24(2), 321–332.

Sathyanarayanan, S., 2024, 'Confusion Matrix-Based Performance Evaluation Metrics', *African Journal of Biomedical Research*, 4023–4031.

Virgiansyah, M.R., Stephanie, S. & Rizky Pribadi, M., 2024, 'Analisis Sentimen terhadap Jalan Rusak di Palembang Pada Media Sosial Menggunakan Algoritma Naïve Bayes', *TIN: Terapan Informatika Nusantara*, 5(1), 1–10.

Winahyu, J. & Suharjo, I., 2021, 'Aplikasi Web Analisis Sentimen Dengan Algoritma Multinomial Naïve Bayes', *Kumpulan Artikel Mahasiswa Pendidikan Teknik Informatika (KARMAPATI)*, 10(2), 206.