JEMSI: Jurnal Ekonomi Manajemen Sistem Informasi
E-ISSN: 2686-5238
P-ISSN: 2686-4916
https://dinastirev.org/JEMSI dinasti.info@gmail.com +62 811 7404 455
DOI: https://doi.org/10.38035/jemsi.v7i2
https://creativecommons.org/licenses/by/4.0/
Received: 29 November 2025 | **Revised:** 10 Desember 2025 | **Accepted:** 20 Desember 2025

Perbandingan Alogaritma antara SVM dan Naïve Bayes guna sebagai Sentimen Mahasiswa Chat Ai

Armanta Tarigan[1], Dedi Trisnawarman[2]
[1]Universitas Tarumanagara, Jakarta Barat, Indonesia, armantatarigan@gamil.com
[2]Universitas Tarumanagara, Jakarta Barat, Indonesia, dedit@fti.untar.ac.id

Corresponding Author: armantatarigan@gmail.com[1]

Abstract: *This study analyzes Indonesian students' sentiment toward the use of ChatAI in education using the Naive Bayes algorithm and Support Vector Machine (SVM) using data from Twitter and questionnaires. Through the application of Natural Language Processing (NLP) techniques and TF-IDF feature extraction, the evaluation results demonstrated that SVM was the best model, with an F1-score of 0.802 on Twitter data and 0.896 on questionnaire data, significantly outperforming Naive Bayes. Based on visualizations in the Streamlit-based system, student sentiment was predominantly neutral, but there was a higher positive trend in the questionnaire data compared to social media, indicating that ChatAI is accepted as an effective academic tool.*

Keywords: *ChatAI, Sentiment Analysis, Support Vector Machine (SVM), Naive Bayes, NLP, TF-IDF, Higher Education.*

Abstrak: Penelitian ini menganalisis sentimen mahasiswa Indonesia terhadap penggunaan ChatAI dalam dunia pendidikan menggunakan algoritma *Naive Bayes* dan *Support Vector Machine* (SVM) dengan sumber data dari Twitter dan kuesioner. Melalui penerapan teknik *Natural Language Processing* (NLP) dan ekstraksi fitur TF-IDF, hasil evaluasi membuktikan bahwa SVM merupakan model terbaik dengan F1-Score sebesar 0.802 pada data Twitter dan 0.896 pada data kuesioner, mengungguli performa *Naive Bayes* secara signifikan. Berdasarkan visualisasi pada sistem berbasis *Streamlit*, ditemukan bahwa sentimen mahasiswa didominasi oleh respon netral, namun terdapat kecenderungan positif yang lebih tinggi pada data kuesioner dibandingkan media sosial, yang mengindikasikan bahwa ChatAI diterima sebagai alat bantu akademik yang efektif.

Kata Kunci: *ChatAI, Analisis Sentimen, Support Vector Machine (SVM), Naive Bayes, NLP, TF-IDF, Pendidikan Tinggi.*

PENDAHULUAN

Era digital kontemporer ditandai oleh akselerasi inovasi teknologi kecerdasan buatan (Artificial Intelligence/AI) yang semakin merambah ke berbagai aspek kehidupan, termasuk

dalam sektor pendidikan tinggi (Santoso, 2025). Salah satu terobosan paling signifikan dalam beberapa tahun terakhir adalah kemunculan *Generative AI*, yakni teknologi yang tidak hanya mampu menganalisis data, tetapi juga menghasilkan konten baru secara otonom. Dalam konteks ini, *Large Language Models* (LLM) seperti GPT dari OpenAI, Gemini dari Google, dan Claude dari Anthropic telah menjadi sorotan utama karena kemampuannya meniru pola bahasa manusia dan menyusun teks yang kompleks (Rolando & Mulyono, 2025).

Perkembangan teknologi kecerdasan buatan (*Artificial Intelligence* / AI) telah membawa perubahan besar dalam berbagai bidang kehidupan, termasuk dunia pendidikan tinggi (Rifky, 2024). Salah satu penerapan AI yang kini banyak digunakan mahasiswa adalah *ChatAI* seperti ChatGPT, Gemini, atau Copilot. Platform tersebut berfungsi sebagai asisten digital yang mampu membantu mahasiswa dalam menyusun ide, menjawab pertanyaan akademik, hingga menyusun laporan atau tugas kuliah (Manuaba *et al*., 2025).

Fenomena meningkatnya penggunaan ChatAI di Indonesia tidak terlepas dari mudahnya akses dan kemampuannya menghasilkan jawaban secara cepat (Sulaiman *et al*., 2024). Mahasiswa kerap menggunakannya sebagai *first step* untuk memahami topik yang sulit ataupun sebagai sarana mencari inspirasi saat mengalami *mental block* dalam penyelesaian tugas (Hartati, 2025). Namun demikian, pemanfaatan ini juga memunculkan perdebatan baru terkait etika akademik, ketergantungan alat, hingga akurasi informasi yang diberikan oleh sistem AI. Di satu sisi, ChatAI dapat mempercepat proses belajar; namun di sisi lain, muncul kekhawatiran bahwa penggunaannya dapat mengurangi kemampuan berpikir kritis dan kreativitas mahasiswa apabila tidak digunakan secara bijak (Oktafia *et al*., 2025).

Kondisi tersebut menimbulkan pertanyaan penting tentang bagaimana persepsi dan sentimen mahasiswa terhadap ChatAI (Oktafia *et al*., 2025). Tidak semua mahasiswa menerima teknologi ini secara positif. Sebagian menilai ChatAI sangat membantu, tetapi sebagian lainnya merasa bahwa hasil yang diberikan kurang akurat, terlalu "robotic", atau tidak mampu memahami konteks lokal. Selain itu, banyak mahasiswa yang masih meragukan keandalan ChatAI sebagai alat bantu akademik jangka panjang (Tontowi, 2025).

Untuk memahami fenomena ini secara objektif, diperlukan pendekatan berbasis data. Salah satu metode yang banyak digunakan untuk menganalisis opini publik adalah analisis sentimen, yaitu proses mengklasifikasikan teks menjadi kategori positif, negatif, atau netral (Tontowi, 2025). Dengan menerapkan teknik Natural Language Processing (NLP) dan algoritma machine learning, analisis sentimen dapat memberikan gambaran yang komprehensif mengenai bagaimana mahasiswa memandang peran ChatAI dalam kegiatan akademik (Aplikasi *et al*., 2025).

Penelitian ini menggunakan dua sumber data utama, yaitu data Twitter sebagai representasi opini spontan dan data kuesioner sebagai representasi opini terstruktur (Safitri *et al*., 2025). Penggunaan dua jenis data ini penting untuk memberikan perspektif yang lebih seimbang. Data Twitter mencerminkan ekspresi alami yang muncul tanpa batasan format, sedangkan data kuesioner memberikan jawaban lebih formal dan mendalam sesuai pertanyaan yang diajukan.

Untuk mengolah data tersebut, digunakan dua model machine learning populer, yaitu Naive Bayes dan Support Vector Machine (SVM). Kedua model ini dipilih karena telah terbukti efektif dalam berbagai penelitian analisis sentimen dan mampu menangani teks dalam jumlah besar. Proses analisis meliputi *preprocessing* data, pembentukan fitur menggunakan TF-IDF, pelatihan model, evaluasi performa, hingga visualisasi hasil melalui aplikasi Streamlit yang dikembangkan dalam penelitian ini (Prasetyo *et al*., 2024).

Dengan adanya analisis yang sistematis dan dukungan visualisasi interaktif, penelitian ini diharapkan dapat memberikan pemahaman mendalam mengenai bagaimana mahasiswa Indonesia menilai penggunaan ChatAI sebagai alat bantu akademik, serta seberapa efektif model machine learning dalam memprediksi sentimen tersebut. Selain itu, hasil penelitian

diharapkan dapat menjadi pertimbangan bagi institusi pendidikan dalam merumuskan kebijakan terkait penggunaan teknologi AI di lingkungan akademik (Agustiani *et al*., 2025).

Meskipun menawarkan kemudahan, penggunaan ChatAI dalam konteks akademik menimbulkan berbagai persepsi di kalangan mahasiswa. Sebagian menganggap ChatAI sebagai alat bantu yang efektif dan efisien, sementara sebagian lainnya menilai penggunaannya dapat menurunkan kemampuan berpikir kritis dan orisinalitas. Oleh karena itu, diperlukan analisis yang lebih mendalam untuk memahami bagaimana persepsi dan sentimen mahasiswa terhadap penggunaan ChatAI sebagai alat bantu akademik.

Analisis sentimen merupakan salah satu teknik *text mining* yang digunakan untuk mengidentifikasi opini publik terhadap suatu topik berdasarkan data teks. Melalui pendekatan ini, peneliti dapat mengelompokkan opini menjadi tiga kategori utama: positif, negatif, dan netral (Xue *et al*., 2009). Dalam penelitian ini, data yang dianalisis berasal dari dua sumber utama: (1) unggahan mahasiswa di platform X (Twitter) dan (2) hasil kuesioner yang diisi oleh mahasiswa dari berbagai universitas di Indonesia. Pendekatan ini memungkinkan perbandingan antara persepsi mahasiswa secara daring (Twitter) dengan hasil survei langsung.

Untuk mengelompokkan sentimen secara otomatis, digunakan dua algoritma *machine learning* yang umum dalam analisis teks, yaitu Naive Bayes dan Support Vector Machine (SVM) (Rahat *et al*., 2019). Kedua algoritma ini akan dibandingkan berdasarkan kinerja *F1-Score* untuk menentukan model yang paling akurat dalam mengklasifikasikan sentimen mahasiswa terhadap ChatAI.

Dengan demikian, penelitian ini diharapkan dapat memberikan gambaran mengenai bagaimana mahasiswa Indonesia memandang ChatAI sebagai alat bantu akademik, serta memberikan kontribusi bagi pihak universitas dalam mengatur penggunaan teknologi AI di lingkungan pendidikan.

METODE

Penelitian ini dilaksanakan mulai bukan juli 2025 hingga selesai di lingkungan Universitas Tarumanagara. Penelitian dilakukan menggunakan pendekatan kuantitatif dengan metode analisis komparatif untuk membandingkan performa dua algoritma klasifikasi teks Naive Bayes dan Support Vector Machine (SVM) dalam melakukan analisis sentimen terhadap opini mahasiswa mengenai penggunaan ChatAI dalam konteks akademik. Metodologi disusun melalui tiga tahapan utama: perancangan, pengumpulan serta pengolahan data, dan pembangunan serta evaluasi model.

Pada tahap perancangan penelitian, peneliti melakukan identifikasi masalah terkait fenomena meningkatnya penggunaan ChatAI oleh mahasiswa. Tahap ini kemudian dilanjutkan dengan perumusan tujuan penelitian serta kajian literatur untuk memahami teori dasar NLP, teknik preprocessing teks, dan karakteristik algoritma Naive Bayes dan SVM yang menjadi dasar analisis.

Tahap berikutnya adalah pengumpulan dan pengolahan data, yang melibatkan dua sumber data yaitu, data primer dari kuesioner mahasiswa Indonesia dan data sekunder dari media sosial X (Twitter). Seluruh data teks kemudian diproses melalui rangkaian *text preprocessing*, meliputi *case folding*, pembersihan karakter dan simbol non-relevan, penghapusan URL dan mention, tokenisasi, penghapusan stopwords, serta stemming menggunakan algoritma bahasa Indonesia. Tahap ini bertujuan memastikan kualitas data yang optimal sebelum digunakan dalam pelatihan model.

Pada tahap pembangunan dan evaluasi model, data yang telah dibersihkan diubah menjadi representasi numerik menggunakan metode TF-IDF, yang memberikan bobot berbeda pada setiap kata berdasarkan tingkat kepentingannya dalam dokumen. Kedua model, Naive Bayes dan SVM, kemudian dilatih menggunakan dataset tersebut dan diuji performanya melalui metrik *accuracy, precision, recall,* dan F1-*score*. Evaluasi ini memungkinkan peneliti mengidentifikasi model yang paling efektif dalam mengklasifikasikan sentimen positif, negatif,

atau netral. Selain itu, hasil klasifikasi digunakan untuk menganalisis kecenderungan umum sentimen mahasiswa terhadap penggunaan ChatAI dalam kegiatan akademik.

Secara keseluruhan, metode ini dirancang untuk menghasilkan analisis yang komprehensif melalui tahapan terstruktur: mulai dari perancangan penelitian, pengumpulan data, preprocessing, pembangunan model klasifikasi, hingga evaluasi performa algoritma. Pendekatan ini memastikan validitas metodologis sekaligus memberikan dasar ilmiah yang kuat dalam menarik kesimpulan terkait persepsi mahasiswa terhadap ChatAI.

HASIL DAN PEMBAHASAN

Evaluasi Hasil Prepocessing

Tahap preprocessing berhasil merapikan teks pada kedua dataset X (twitter) dan kuesioner dengan menghilangkan elemen-elemen tidak relevan seperti URL, mention, hashtag, emoji, serta tanda baca. Pada dataset Twitter yang cenderung informal, proses ini mengubah teks mentah yang penuh singkatan dan simbol menjadi bentuk ringkas dan bermakna tanpa menghilangkan konteks sentimen. Contoh pada tabel menunjukkan bahwa tweet yang awalnya panjang dan tidak baku dapat disederhanakan menjadi kumpulan kata inti yang siap digunakan oleh model.

Pada dataset kuesioner, yang memiliki karakteristik lebih panjang dan terstruktur, preprocessing membantu mereduksi kalimat deskriptif menjadi kata-kata kunci yang tetap menggambarkan opini utama responden. Transformasi seperti case folding, penghapusan stopwords, dan stemming berperan besar dalam menyatukan variasi kata dan menjaga konsistensi fitur. Secara keseluruhan, preprocessing terbukti menghasilkan data yang lebih bersih, padat, dan informatif sehingga mendukung kinerja model secara optimal.

Evaluasi Model

Berdasarkan hasil evaluasi pada dataset Twitter, terlihat perbedaan performa yang cukup jelas antara model Naive Bayes dan Support Vector Machine (SVM). Nilai F1-Score menunjukkan bahwa SVM mencapai skor 0.802, lebih tinggi dibandingkan Naive Bayes yang hanya 0.702. Keunggulan ini menunjukkan bahwa SVM lebih mampu mengenali pola pada teks Twitter yang sangat bervariasi, informal, serta sering mengandung singkatan, emoji, dan struktur kalimat tidak baku. Kemampuan SVM dalam membentuk margin pemisah yang optimal membuatnya lebih efektif dalam menangani keragaman tersebut.

Classification Report juga memperlihatkan perbedaan kinerja yang signifikan antar model. Naive Bayes gagal memprediksi kelas negatif dan positif (precision, recall, dan F1-Score = 0.00), karena cenderung bias pada kelas netral yang memang mendominasi pada dataset. Akurasi total NB memang mencapai 0.79, tetapi akurasi ini tidak mencerminkan performa sebenarnya (accuracy paradox) karena hampir seluruh prediksi diarahkan ke kelas mayoritas.

Sebaliknya, SVM lebih stabil dalam mengklasifikasikan ketiga kelas. Meskipun masih belum mampu mengenali sentimen negatif, SVM dapat memprediksi kelas positif dengan F1-Score 0.46, jauh lebih baik dibanding NB. Untuk kelas netral, SVM memiliki precision 0.84 dan recall sempurna 1.00, menunjukkan efektivitasnya dalam mendeteksi pola bahasa yang umum dalam percakapan informatif di Twitter.

Jika dilihat dari macro average dan weighted average, SVM tetap unggul. Macro F1-Score SVM mencapai 0.46, lebih tinggi dari NB yang hanya 0.29. Weighted F1-Score SVM sebesar 0.80 juga memperlihatkan konsistensi model meskipun kelas tidak seimbang.

Secara keseluruhan, hasil ini menunjukkan bahwa SVM merupakan model yang paling sesuai untuk dataset Twitter dalam penelitian ini. Variasi bahasa yang tinggi dan ketidakseimbangan kelas membuat Naive Bayes kurang efektif, sementara SVM dapat menghasilkan prediksi yang lebih akurat dan stabil. Tabel 1 kemudian merangkum F1-Score kedua model sebagai pembanding utama.

Tabel 1. Ringkasan F1-Score Model untuk Dataset Twitter

Model	F1-Score
Naive Bayes	0.702
SVM	0.802

Pembahasan:

1. Model SVM memiliki performa lebih baik dengan F1-Score sebesar 0.802, dibandingkan Naive Bayes yang hanya 0.702.
2. Hal ini menunjukkan bahwa SVM mampu menangkap pola teks secara lebih baik dalam kasus data Twitter yang cenderung beragam dan informal.

Selain ringkasan tersebut, berikut adalah *Classification Report* masing-masing model untuk data Twitter Pada Tabel 2 dan Tabel 3 berikut.

Tabel 2. Classification Report Model Naive Bayes – Dataset Twitter

Sentimen	Precision	Recall	F1-Score	Support
Negatif	0.00	0.00	0.00	2
Netral	0.79	1.00	0.88	46
Positif	0.00	0.00	0.00	10
Accuracy	0.79			58
Macro Avg	0.26	0.33	0.29	
Weighted Avg	0.63	0.79	0.70	

Tabel 3. Classification Report Model SVM – Dataset Twitter

Sentimen	Precision	Recall	F1-Score	Support
Negatif	0.00	0.00	0.00	2
Netral	0.84	1.00	0.91	46
Positif	1.00	0.30	0.46	10
Accuracy	0.84			58
Macro Avg	0.61	0.43	0.46	
Weighted Avg	0.84	0.84	0.80	

Pembahasan per metrik:

1. *Precision & Recall:* Model SVM lebih baik secara keseluruhan dalam memprediksi sentimen netral dibanding NB.
2. *F1-Score:* Naive Bayes gagal memprediksi positif dan negatif dengan baik (F1-Score 0.00), sedangkan SVM mampu menangani positif dengan lebih baik meski recall rendah.

Kesimpulan: SVM menunjukkan kinerja lebih stabil dan akurat pada dataset Twitter.

Evaluasi Model pada Dataset Twitter

Evaluasi model pada dataset kuesioner menunjukkan hasil yang lebih baik dibandingkan dataset Twitter karena teks kuesioner lebih terstruktur dan berfokus pada konteks akademik. Tabel 18 menunjukkan bahwa SVM kembali menjadi model terbaik dengan F1-Score 0.896, jauh lebih tinggi daripada Naive Bayes (NB) yang hanya 0.595. Hal ini menunjukkan bahwa SVM lebih mampu menangkap pola bahasa dalam jawaban yang panjang, formal, dan konsisten.

Classification Report mengungkapkan bahwa NB kembali gagal memprediksi kelas negatif dan positif, dengan precision, recall, dan F1-Score 0.00 di kedua kelas. Model ini tetap bias pada kelas netral sehingga akurasi 0.71 tidak mencerminkan performa sebenarnya (accuracy paradox). Macro average F1-Score NB yang hanya 0.28 menegaskan bahwa NB tidak mampu membedakan pola antar kelas pada data ini.

Sebaliknya, SVM menunjukkan performa sangat stabil di semua kelas. Model ini mencapai precision, recall, dan F1-Score 1.00 pada kelas negatif; F1-Score 0.94 pada kelas netral; dan F1-Score 0.75 pada kelas positif. Nilai macro average dan weighted average masing-

masing sekitar 0.90, menandakan kemampuan SVM dalam menangani kelas mayoritas dan minoritas secara seimbang.

Secara keseluruhan, hasil ini menunjukkan bahwa struktur teks yang formal dan jelas membuat SVM jauh lebih optimal dibanding NB. SVM mampu memahami konteks dan korelasi antar kata yang kompleks, sedangkan NB kesulitan menangani variasi makna dan kalimat panjang. Oleh karena itu, SVM merupakan model yang paling efektif untuk data kuesioner dalam penelitian ini.

Analisis Visualisasi Hasil

Visualisasi hasil menjadi elemen penting dalam analisis karena memberikan gambaran yang lebih jelas dan intuitif mengenai performa model serta pola sentimen dari data. Dalam penelitian ini, seluruh visualisasi disajikan melalui antarmuka Streamlit, yang memungkinkan pengguna melihat hasil secara interaktif dan real-time. Tampilan ini membantu pengguna memahami bagaimana data diproses pada setiap tahap dan bagaimana model bekerja.

Salah satu visualisasi utama adalah pie chart distribusi sentimen pada setiap dataset. Diagram ini memudahkan pengguna melihat proporsi sentimen negatif, netral, dan positif secara cepat. Pada Twitter, sentimen didominasi oleh kelas netral, sedangkan pada kuesioner distribusi lebih terarah karena responden memberikan opini yang lebih jelas terkait penggunaan ChatAI dalam akademik. Perbedaan pola ini jauh lebih mudah dipahami ketika divisualisasikan daripada hanya melalui tabel.

Selain itu, grafik perbandingan performa model—biasanya dalam bentuk bar chart—menunjukkan perbedaan F1-Score antara Naive Bayes dan SVM. Visualisasi ini memperjelas keunggulan SVM, terutama pada dataset kuesioner yang lebih kompleks. Penempatan grafik secara berdampingan juga membantu pengguna memahami bagaimana konteks dan karakteristik platform memengaruhi hasil sentimen.

Sistem juga menyediakan tabel interaktif berisi teks asli, teks hasil preprocessing, dan label sentimen prediksi. Fitur ini mendukung verifikasi manual sehingga pengguna dapat memastikan bahwa hasil klasifikasi sesuai dengan makna teks sebenarnya.

Secara keseluruhan, visualisasi melalui Streamlit memberikan kontribusi penting dalam memahami pola data, mengevaluasi performa model, dan memvalidasi hasil prediksi. Visualisasi menjadi penghubung antara analisis teknis dan interpretasi hasil, sehingga mendukung penyusunan kesimpulan penelitian secara lebih kuat dan berbasis data.

1. Distribusi Sentimen pada Dataset Twitter dan Kuesioner

Distribusi sentimen menunjukkan bahwa:

a. Twitter didominasi sentimen netral (>70%). Hal ini menunjukkan bahwa pengguna Twitter cenderung menulis opini informatif atau tanpa ekspresi emosional kuat.
b. Kuesioner memiliki proporsi sentimen positif yang lebih besar (23.8%), mencerminkan opini lebih reflektif dan terarah terhadap manfaat ChatAI dalam konteks akademik.
c. Sentimen negatif pada kedua dataset relatif kecil (3–5%), tetapi tetap signifikan karena mencerminkan kritik terhadap ChatAI.

Perbedaan ini menegaskan bahwa platform mempengaruhi cara opini disampaikan, sehingga memengaruhi hasil klasifikasi sentimen.

2. Perbandingan Performa Model: Naive Bayes vs SVM

Perbandingan performa menunjukkan bahwa SVM secara konsisten lebih unggul pada kedua dataset:

a. Twitter: SVM (0.802) vs NB (0.702)
b. Kuesioner: SVM (0.896) vs NB (0.595)

NB memiliki kelemahan dalam menangani ketidakseimbangan kelas dan konteks kalimat, sehingga cenderung memprediksi sebagian besar data sebagai netral. Sebaliknya, SVM mampu

memisahkan kelas lebih akurat karena kemampuan menentukan margin optimal pada ruang fitur TF-IDF.

3. Perbandingan Distribusi Sentimen: Twitter vs Kuesioner

Tabel 4. Perbandingan Persentase Sentimen Twitter dan Kuesioner

Sentimen	Twitter (%)	Kuesioner (%)
Positif	17.2	23.8
Netral	79.3	71.4
Negatif	3.5	4.8

Berdasarkan persentase pada Tabel 4, sentimen netral mendominasi kedua dataset. Pada Twitter, 79.3% unggahan tergolong netral, sedangkan pada kuesioner 71.4%. Dominasi ini menunjukkan bahwa banyak pengguna Twitter hanya memberi komentar informatif atau menyebutkan ChatAI tanpa ekspresi emosional yang jelas. Karakter komunikasi Twitter yang singkat, cepat, dan spontan turut membuat sentimen eksplisit jarang muncul.

Pada kuesioner, meskipun netral tetap dominan, persentase sentimen positif lebih tinggi dibanding Twitter 23.8% dibanding 17.2%. Hal ini karena kuesioner mendorong responden untuk merefleksikan pengalaman akademik mereka secara langsung, sehingga opini yang muncul lebih terarah dan mencerminkan manfaat nyata penggunaan ChatAI, misalnya dalam memahami materi atau menyelesaikan tugas.

Sentimen negatif pada kedua dataset relatif kecil (3.5% di Twitter dan 4.8% di kuesioner). Namun, jenis kritik berbeda: di Twitter bersifat spontan seperti keluhan atas jawaban yang salah, sementara pada kuesioner lebih terstruktur, menyangkut akurasi, kedalaman analisis, atau konteks jawaban.

Secara keseluruhan, perbandingan ini menunjukkan bahwa pola sentimen sangat dipengaruhi konteks platform. Twitter menghasilkan opini yang lebih spontan dan tidak eksplisit sehingga didominasi sentimen netral, sedangkan kuesioner memicu jawaban lebih reflektif sehingga meningkatkan proporsi sentimen positif. Temuan ini menegaskan bahwa sumber dan konteks pengumpulan data berperan besar dalam pembentukan sentimen.

Diskusi Hasil

Perbandingan Performa Model

Hasil evaluasi menunjukkan bahwa Support Vector Machine (SVM) secara konsisten unggul dibandingkan Naive Bayes (NB) pada kedua dataset. Pada Twitter, SVM memperoleh F1-Score 0.802 sementara NB hanya 0.702. Pada kuesioner, perbedaannya semakin besar: SVM mencapai 0.896 dan NB hanya 0.595. Keunggulan SVM disebabkan kemampuannya membentuk hyperplane optimal pada ruang fitur TF-IDF, sehingga lebih efektif dalam membedakan sentimen meskipun teks bervariasi dan kelas tidak seimbang.

SVM juga mampu mengenali sentimen positif dan negatif lebih baik, bahkan ketika jumlah datanya sedikit seperti pada Twitter. Sebaliknya, Naive Bayes menunjukkan performa sangat rendah pada kelas minoritas bahkan F1-Score 0.00 pada kelas positif dan negatif untuk kedua dataset. Keterbatasan ini dipengaruhi oleh asumsi independensi fitur NB serta ketergantungannya pada frekuensi kata, yang membuatnya sulit memahami konteks kalimat yang kompleks atau frasa yang maknanya tidak dapat dipisahkan per kata.

NB juga sangat sensitif terhadap ketidakseimbangan kelas dan cenderung memprediksi kelas mayoritas, menyebabkan akurasi terlihat tinggi tetapi kemampuan klasifikasinya buruk untuk kelas minoritas. Sebaliknya, SVM lebih tangguh dalam kondisi distribusi data tidak merata.

Kedua model tampil lebih baik pada dataset kuesioner karena teks yang lebih panjang, jelas, dan terstruktur memberikan representasi fitur yang lebih stabil. Namun, SVM jauh lebih mampu memanfaatkan kekayaan fitur tersebut dibanding NB. Temuan ini menegaskan bahwa SVM adalah model yang paling sesuai untuk analisis sentimen terkait ChatAI dalam konteks

akademik, sementara NB hanya cocok sebagai baseline. Hasil ini konsisten dengan penelitian sebelumnya yang menunjukkan keandalan SVM untuk klasifikasi teks dengan variasi linguistik tinggi.

Perbedaan Distribusi dan Karakteristik Data Twitter vs Kuesioner

Perbedaan karakteristik dataset Twitter dan dataset kuesioner berpengaruh kuat terhadap performa model klasifikasi. Dataset Twitter berisi opini spontan yang tidak terstruktur, singkat, dan menggunakan bahasa informal seperti singkatan, emoji, slang, hingga frasa kontekstual. Karakteristik ini menyebabkan mayoritas tweet dikategorikan netral karena banyak unggahan hanya berupa informasi, pertanyaan, atau komentar tanpa ekspresi emosional jelas. Kondisi tersebut menyulitkan model terutama Naive Bayes, untuk membedakan sentimen positif dan negatif karena indikator linguistik yang eksplisit sangat minim. Ambiguitas, ironi, dan sarkasme dalam tweet juga memperbesar potensi kesalahan prediksi.

Sebaliknya, dataset kuesioner jauh lebih terstruktur, panjang, dan formal. Jawaban diberikan dalam konteks yang jelas, yaitu pengalaman mahasiswa dalam menggunakan ChatAI untuk kegiatan akademik. Hal ini menghasilkan teks yang lebih deskriptif, memiliki makna yang lebih eksplisit, serta menawarkan variasi kosakata yang stabil, sehingga sentimen lebih mudah dikenali oleh model. Distribusi sentimennya juga lebih seimbang, memungkinkan model belajar dari pola yang lebih representatif. Inilah sebabnya SVM mencapai F1-Score tinggi (0.896) pada dataset kuesioner, jauh melampaui hasil pada Twitter.

Pada sisi lain, keterbatasan Twitter yaitu, teks pendek, minim konteks, dan penggunaan bahasa yang variative membuat representasi fitur sulit dibentuk secara stabil meskipun preprocessing telah dilakukan. Banyak tweet tidak memberikan cukup informasi bagi model, sehingga NB gagal total pada kelas minoritas dan SVM pun hanya bekerja optimal pada kelas tertentu.

Secara keseluruhan, analisis ini menegaskan bahwa struktur dan kualitas data sangat menentukan efektivitas model klasifikasi. Dataset kuesioner yang lebih konsisten dan kaya makna memungkinkan model belajar lebih baik, sementara dataset Twitter yang lebih kompleks membutuhkan metode lanjutan seperti context-aware embeddings atau model deep learning untuk menghasilkan prediksi yang lebih akurat.

Efektivitas Preprocessing dan Pengaruhnya terhadap Kinerja Model

Proses preprocessing dalam penelitian ini memberikan kontribusi besar terhadap peningkatan kualitas data sebelum pemodelan. Tahapan seperti case folding, penghapusan URL, mention, emoji, tanda baca, serta stopword removal berhasil menghilangkan berbagai noise sehingga teks mentah menjadi lebih terstruktur dan ringkas. Perbandingan antara Teks Mentah dan Teks Bersih menunjukkan bahwa preprocessing mampu menyederhanakan kalimat tanpa menghilangkan makna inti, sehingga fitur TF-IDF dapat bekerja lebih optimal.

Efektivitas preprocessing sangat terlihat pada dataset kuesioner. Karena respon umumnya panjang dan deskriptif, tahapan ini membantu mereduksi kalimat menjadi kata kunci yang relevan, sekaligus meningkatkan konsistensi bahasa melalui stemming. Hasilnya, model khususnya SVM yang lebih mudah memetakan pola sentimen karena fitur yang dihasilkan lebih stabil dan representatif.

Pada dataset Twitter, preprocessing berhasil menghapus elemen-elemen yang tidak relevan seperti hashtag atau URL. Namun, sifat teks Twitter yang singkat, informal, dan penuh variasi gaya bahasa tetap menjadi tantangan. Meskipun noise berkurang, variasi konteks membuat Naive Bayes sulit menangkap hubungan kata, sehingga F1-Score untuk kelas positif dan negatif sangat rendah. Sebaliknya, SVM tetap menunjukkan kinerja lebih baik karena mampu memanfaatkan representasi fitur berdimensi tinggi untuk membentuk batas keputusan yang lebih akurat.

Secara keseluruhan, preprocessing terbukti menjadi tahap fundamental yang meningkatkan kualitas teks, meminimalkan ambiguitas, dan memperjelas representasi fitur. Proses ini berperan besar dalam membantu model terutama SVM dalam memahami pola sentimen dengan lebih baik, sehingga menghasilkan prediksi yang lebih akurat pada kedua dataset.

KESIMPULAN

Kesimpulan

Penelitian ini berhasil membangun sistem analisis sentimen berbasis NLP dan machine learning untuk memahami persepsi mahasiswa Indonesia terhadap penggunaan ChatAI sebagai alat bantu akademik, menggunakan data Twitter dan kuesioner. Sistem mampu menjalankan seluruh proses mulai dari pengumpulan data, preprocessing, ekstraksi fitur TF-IDF, hingga pelatihan model Naive Bayes dan SVM yang diintegrasikan dalam aplikasi Streamlit. Hasil evaluasi menunjukkan bahwa SVM secara konsisten unggul dengan F1-Score tertinggi pada kedua dataset (0.802 pada Twitter dan 0.896 pada kuesioner), sementara Naive Bayes kesulitan memprediksi kelas minoritas dan menghasilkan F1-Score 0.00 pada beberapa kelas. Distribusi sentimen pada kedua dataset didominasi sentimen netral, namun kuesioner menunjukkan peningkatan sentimen positif karena responden memberikan opini reflektif dan terarah. Proses preprocessing terbukti sangat penting karena meningkatkan kualitas teks, menjaga makna inti, dan menghasilkan representasi fitur yang stabil, yang berdampak signifikan terhadap performa model khususnya SVM. Selain itu, implementasi berbasis Streamlit memudahkan pengguna dalam melihat alur analisis secara lengkap melalui tabel preprocessing, visualisasi, dan hasil evaluasi, sehingga sistem ini tidak hanya bersifat akademik tetapi juga aplikatif dan dapat digunakan kembali untuk berbagai kebutuhan analisis sentimen.

Saran

Penelitian selanjutnya disarankan menggunakan model berbasis deep learning seperti LSTM, GRU, atau BERT karena memiliki kemampuan lebih baik dalam menangkap konteks dan makna implisit, terutama untuk kalimat panjang dan kompleks. Jumlah dan sumber dataset juga perlu diperluas misalnya dari forum akademik, YouTube comments, atau platform diskusi edukasi agar model lebih representatif serta mampu menangani variasi bahasa yang lebih luas. Upaya balancing data seperti SMOTE, oversampling, atau cost-sensitive learning penting dilakukan untuk meningkatkan performa pada kelas minoritas yang masih rendah pada penelitian ini. Selain analisis sentimen, dapat ditambahkan analisis emosi untuk menghasilkan pemahaman yang lebih kaya mengenai persepsi mahasiswa, misalnya mengidentifikasi emosi spesifik seperti puas, bingung, atau frustrasi. Pengembangan sistem juga dapat ditingkatkan melalui integrasi API, fitur dashboard interaktif, laporan otomatis, serta tampilan antarmuka yang lebih modern. Terakhir, modul preprocessing perlu dibuat lebih adaptif terhadap Bahasa Indonesia dengan mencakup normalisasi slang, pemetaan singkatan, dan teknik lemmatization agar kualitas representasi fitur semakin baik. Secara keseluruhan, pengembangan ini diharapkan dapat menghasilkan sistem analisis sentimen yang lebih akurat, skalabel, dan relevan untuk berbagai kebutuhan akademik maupun industri.

REFERENSI

A. M. Rahat, A. Kahir, and A. K. M. Masum, “Comparison of Naive Bayes and SVM Algorithm based on Sentiment Analysis Using Review Dataset,” *Proc. 2019 8th Int. Conf. Syst. Model. Adv. Res. Trends, SMART 2019*, pp. 266–270, Feb. 2020, doi: 10.1109/SMART46866.2019.9117512.

“Analisis Sentimen: Metode Alternatif Penelitian Big Data” - Reza Safitri, Nisa Alfira, Dian Tamitiadini, Wayan Weda Asmara Dewi, Nufian Febriani - Google Books.” Accessed: Nov. 19, 2025.

B. Rolando and H. Mulyono, "E-Commerce as a Catalyst for Digital Economy Development: A Study of Marketing Strategies and Their Impact," *J. Distrib. Sci.*, vol. 23, no. 4, pp. 61–79, 2025, doi: 10.15722/JDS.23.04.202504.61.

H. Xue, Q. Yang, and S. Chen, "SVM: Support Vector Machines," *Top Ten Algorithms Data Min.*, pp. 37–59, Jan. 2009, doi: 10.1201/9781420089653-10/SVM-SUPPORT-VECTOR-MACHINES-HUI-XUE-QIANG-YANG-SONGCAN-CHEN.

J. Khatib Sulaiman, M. Abid Nadzif, and U. Stikubank Semarang, "Penggunaan Teknologi Natural Language Processing dalam Sistem Chatbot untuk Peningkatan Layanan Informasi Administrasi Publik," *Indones. J. Comput. Sci.*, vol. 13, no. 1, pp. 2024–1227, Feb. 2024, doi: 10.33022/IJCS.V13I1.3645.

J. T. Santoso, "KECERDASAN BUATAN (Artificial Intelligence)," *Penerbit Yayasan Prima Agus Tek.*, pp. 1–227, Aug. 2023, Accessed: Nov. 12, 2025. [Online]. Available: https://penerbit.stekom.ac.id/index.php/yayasanpat/article/view/437

K. Korovkinas, P. Danėnas, and G. Garšva, "SVM and naive bayes classification ensemble method for sentiment analysis," *Balt. J. Mod. Comput.*, vol. 5, no. 4, pp. 398–409, Dec. 2017, doi: 10.22364/BJMC.2017.5.4.06.

"KECERDASAN BUATAN BERBASIS PENGETAHUAN - Sri Hartati - Google Books." Accessed: Nov. 19, 2025. [Online]. Available: https://books.google.co.id/books?hl=en&lr=&id=cnlREAAAQBAJ&oi=fnd&pg=PA1&dq=kecerdasan+buatan&ots=mFFeW4Rzdz&sig=fOn8yaQWJYta9xDU_i9zpewpWXw&redir_esc=y#v=onepage&q=kecerdasan buatan&f=false

N. Oktafia *et al.*, "Mahasiswa dan AI: Transformasi Cara Berpikir Kritis dan Penyelesaian Masalah di Era Digital," *Pros. Konseling Kearifan Nusant.*, vol. 4, pp. 10–33, Jan. 2025, doi: 10.29407/VPNFQ046.

"Pembelajaran Matematika dengan Artifical Intelligence: Menutup Celah ... - Riza Agustiani, Atika Zahra, Agustiany Dumeva Putri, Muhammad Hadi, Aji Marwan, Asna Markhotun Solehah, Anggun Meylani Saputri, Ahmad Sopani - Google Books." Accessed: Nov. 19, 2025.

S. Rifky, "Dampak Penggunaan Artificial Intelligence Bagi Pendidikan Tinggi," *Indones. J. Multidiscip. Soc. Technol.*, vol. 2, no. 1, pp. 37–42, Feb. 2024, doi: 10.31004/ijmst.v2i1.287.

T. Aplikasi *et al.*, "Analisis Sentimen Multi-Kelas untuk Menilai Kepuasan Mahasiswa terhadap Aplikasi Manajemen Akademik Berbasis Web di Lingkungan Perguruan Tinggi," *J. Komput. Teknol. Inf. Sist. Inf.*, vol. 4, no. 1, pp. 194–197, Jun. 2025, doi: 10.62712/JUKTISI.V4I1.363.

"TEKNOLOGI ChatGPT : Pengetahuan Dasar dan Pemanfaatan kombinasi keahlian ... - Ida Bagus Kerthyayana Manuaba, Danang Erwanto, Loso Judijanto, Budi Harto, H. Sa'dianoor, I Kadek Dwi Gandika Supartha, Farid Wahyudi, Mahdianta Pandia, Kelvin Kelvin - Google…." Accessed: Nov. 12, 2025. [Online]. Available: https://books.google.co.id/books?hl=en&lr=&id=O5ATEQAAQBAJ&oi=fnd&pg=PA1&dq=adalah+ChatAI+seperti+ChatGPT,+Gemini,+atau+Copilot.+Platform+tersebut+berfungsi+sebagai+asisten+digital+yang+mampu+membantu+mahasiswa+dalam+menyusun+ide,+menjawab+pertanyaan+akademik,+hingga+menyusun+laporan+atau+tugas+kuliah.&ots=jvvWIAT6h2&sig=CwZMpBtX_dX9EZwMMc6dC1_lzFc&redir_esc=y#v=onepage&q&f=false

"Transformasi Digital, Harmonisasi Inovasi, dan Kemanusiaan - Alva E Tontowi - Google Books." Accessed: Nov. 19, 2025. [Online]. Available: https://books.google.co.id/books?hl=en&lr=&id=AYOFEQAAQBAJ&oi=fnd&pg=PA1&dq=Tidak+semua+mahasiswa+menerima+teknologi+ini+secara+positif.+Sebagian+menilai+ChatAI+sangat+membantu,+tetapi+sebagian+lainnya+merasa+bahwa+hasil+yang+diberikan+kurang+akurat,+terlalu+"robotic",+atau+tidak+mampu+memahami+ko

nteks+lokal.+Selain+itu,+banyak+mahasiswa+yang+masih+meragukan+keandalan+ChatAI+sebagai+alat+bantu+akademik+jangka+panjang.&ots=yejOCftEAi&sig=eF5uu9SpN1gcYVhyoDTgPYReeBo&redir_esc=y#v=onepage&q&f=false

Y. A. Prasetyo, E. Utami, and A. Yaqin, "Pengaruh Komposisi Split Data Terhadap Performa Akurasi Analisis Sentimen Algoritma Naïve Bayes dan SVM," *J. homepage J. Electr. Eng. Comput.*, vol. 6, no. 2, 2024, doi: 10.33650/jeecom.v4i2.