

# Evaluasi Sentimen Pengguna ChatGPT Menggunakan Naive Bayes: Tinjauan dari *Confusion Matrix* dan *Classification Report*

## *Evaluation ChatGPT User Sentiment using Naive Bayes: A Review of Confusion Matrix and Classification Report*

Dianda Rifaldi<sup>1\*</sup>, Tri Stiyo Famuji<sup>2</sup>, Bella Okta Sari Miranda<sup>3</sup>, Fauzan Purma Ramadhan<sup>4</sup>, Iriene Putri Mulyadi<sup>5</sup>, Vanji Saputra<sup>6</sup>, Galih Pramuja Inngam Fanani<sup>7</sup>

<sup>1,4,5</sup>Informatika, Fakultas Teknologi Informasi, Universitas Riau Indonesia, Rengat, Indonesia.

<sup>2</sup>Informatika, Fakultas Farmasi, Sains, dan Teknologi, Universitas Al-Irsyad, Cilacap, Indonesia.

<sup>3</sup>Informatika, Fakultas Teknologi Industri, Universitas Ahmad Dahlan, Yogyakarta, Indonesia.

<sup>6</sup>Sistem Informati, Fakultas Teknologi Informasi, Universitas Riau Indonesia, Rengat, Indonesia.

<sup>7</sup>Sistem dan Teknologi Informasi, Fakultas Sains dan Teknologi, Universitas 'Aisyiyah Surakarta, Indonesia.

\*E-mail: [diandarifaldi@gmail.com](mailto:diandarifaldi@gmail.com)

[tristiyofamuji@gmail.com](mailto:tristiyofamuji@gmail.com); [mirandabella1110@gmail.com](mailto:mirandabella1110@gmail.com); [fauzanpurmar@gmail.com](mailto:fauzanpurmar@gmail.com);

[irieneputrimulyadi5@gmail.com](mailto:irieneputrimulyadi5@gmail.com); [vanjisaputra34@gmail.com](mailto:vanjisaputra34@gmail.com); [galihfanani58@gmail.com](mailto:galihfanani58@gmail.com)

### Article History

Submitted : Juni 24, 2025

Revised : Juli 08, 2025

Accepted : Juli 29, 2025

Available Online : Juli 31, 2025

Published Regularly : Juli 31, 2025

**Kata Kunci:** AI, Chatgpt, Kaggle, NLP, Naïve Bayes,

**Keywords:** AI, Chatgpt, Kaggle, NLP, Naïve Bayes

### Contact



Author

[diandarifaldi@gmail.com](mailto:diandarifaldi@gmail.com)

### ABSTRAK

Perkembangan teknologi kecerdasan buatan (AI), khususnya dalam pemrosesan bahasa alami (NLP), telah menghasilkan berbagai inovasi, salah satunya ChatGPT. Popularitasnya yang tinggi mendorong perlunya analisis terhadap sentimen pengguna terhadap ChatGPT menggunakan algoritma Naive Bayes. Data diambil dari platform Kaggle berupa 500 tweet berbahasa Inggris yang telah dilabeli positif, netral, dan negatif. Proses meliputi preprocessing, ekstraksi fitur TF-IDF, pembagian data (80% latih, 20% uji), dan pelatihan model. Hasil menunjukkan akurasi sebesar 56%, dengan f1-score tertinggi pada kelas negatif (0.67) dan terendah pada kelas netral (0.38). Model menunjukkan ketimpangan klasifikasi, dengan precision tinggi tetapi recall rendah pada kelas netral, serta recall tinggi namun precision rendah pada kelas positif. Confusion matrix menegaskan banyaknya kesalahan klasifikasi antar kelas. Hal ini mencerminkan keterbatasan Naive Bayes dalam menangani konteks kata dalam data teks. Perbaikan dapat dilakukan melalui penyeimbangan data, penguatan fitur NLP, serta penerapan algoritma klasifikasi yang lebih kompleks..

## ABSTRACT

The development of artificial intelligence (AI) technology, particularly in natural language processing (NLP), has led to various innovations, including ChatGPT. Its growing popularity highlights the need for user sentiment analysis. This study evaluates user sentiment toward ChatGPT using the Naive Bayes algorithm. The dataset, obtained from Kaggle, consists of 500 labeled English tweets categorized as positive, neutral, or negative. The process involved text preprocessing, TF-IDF feature extraction, data splitting (80% training, 20% testing), and model training. The results show an accuracy of 56%, with the highest f1-score in the negative class (0.67) and the lowest in the neutral class (0.38). The model exhibits classification imbalance, with high precision but low recall in the neutral class, and high recall but low precision in the positive class. The confusion matrix further confirms frequent misclassifications between classes. These findings reflect the limitations of Naive Bayes in handling contextual relationships in text data. Improvements can be achieved through data balancing, enhanced NLP-based feature representation, and the application of more complex classification algorithms.

## 1. Pendahuluan

Kemajuan dalam teknologi kecerdasan buatan (*Artificial Intelligence/AI*) telah melahirkan berbagai inovasi, salah satunya adalah ChatGPT, sebuah model bahasa generatif berbasis transformer yang dikembangkan oleh OpenAI [1]. Sejak diluncurkan, ChatGPT telah digunakan secara luas di berbagai sektor, mulai dari pendidikan, layanan pelanggan, hingga pembuatan konten digital. Namun, popularitasnya yang besar juga menimbulkan berbagai macam opini dan sentimen dari para pengguna, baik yang positif, netral, maupun negatif.

Analisis sentimen terhadap tanggapan pengguna menjadi hal penting untuk memahami persepsi masyarakat terhadap ChatGPT, terutama terkait pemakaian dan tingkat kepercayaan terhadap teknologi AI ini [2]. Salah satu metode yang sering dipakai untuk klasifikasi sentimen adalah algoritma *Naive Bayes*, karena kehandalannya dalam mengolah data teks serta efisiensi dalam komputasi [3][4][5][6].

Dalam beberapa tahun terakhir, perkembangan teknologi kecerdasan buatan terutama di bidang pemrosesan bahasa alami (*Natural Language Processing/NLP*) semakin pesat [7][8]. ChatGPT sebagai salah satu inovasi terdepan memiliki potensi besar dalam mempermudah berbagai aktivitas manusia, mulai dari komunikasi hingga otomatisasi layanan. Namun, seperti halnya teknologi baru lainnya, penerimaan dan persepsi masyarakat terhadap ChatGPT sangat beragam dan dipengaruhi oleh berbagai faktor seperti pengalaman penggunaan, tingkat kepercayaan, serta pemahaman terhadap fungsi teknologi ini.

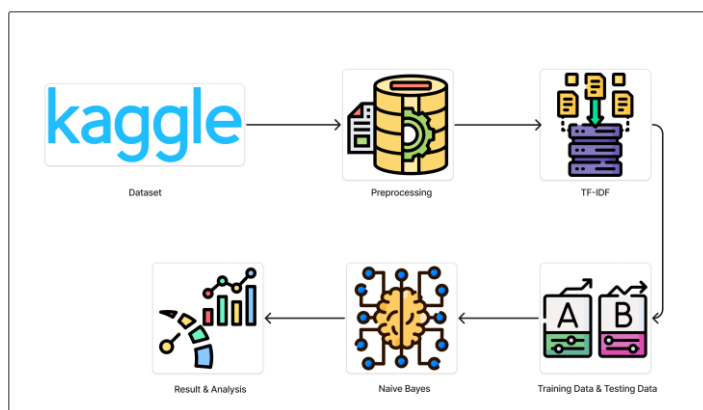
Oleh karena itu, analisis sentimen menjadi alat penting untuk menggali pola persepsi pengguna secara lebih mendalam [9][10]. Algoritma *Naive Bayes* dipilih dalam penelitian ini karena kemampuannya dalam mengklasifikasikan teks secara efisien dan akurat, meskipun dengan data yang relatif sederhana. Selain itu, untuk mengevaluasi performa model klasifikasi, penelitian ini menggunakan metrik-metrik yang umum dipakai seperti confusion matrix dan *classification report* yang terdiri dari *precision*, *recall*, dan *f1-score*. Pendekatan ini memungkinkan penilaian tidak hanya terhadap akurasi keseluruhan, tetapi juga pemahaman terhadap kekuatan dan kelemahan model dalam mengklasifikasikan setiap kategori sentimen.

Melalui evaluasi ini, diharapkan hasil yang diperoleh dapat memberikan wawasan lebih mendalam mengenai bagaimana model *Naive Bayes* bekerja dalam konteks analisis sentimen terhadap ChatGPT, serta memberikan dasar yang kuat untuk pengembangan metode analisis sentimen yang lebih baik di masa depan. Penelitian ini bertujuan untuk mengkaji komentar

pengguna mengenai ChatGPT dengan menggunakan algoritma *Naive Bayes*, serta menilai kinerja model klasifikasi tersebut melalui metrik evaluasi seperti *confusion matrix*, *precision*, *recall*, dan *f1-score*. Dengan pendekatan ini, diharapkan penelitian dapat memberikan gambaran tentang pola sentimen pengguna dan tingkat akurasi model dalam mengelompokkan opini publik terhadap teknologi ChatGPT.

## 2. Metode Penelitian

Metode penelitian ini dirancang untuk menganalisis sentimen pengguna terhadap ChatGPT menggunakan pendekatan pembelajaran mesin. Penelitian ini terdiri dari beberapa tahapan utama, yaitu pengumpulan data, pra-pemrosesan data, ekstraksi fitur, pemodelan menggunakan algoritma *Naive Bayes*, serta evaluasi performa model. Data yang digunakan merupakan kumpulan teks ulasan atau komentar dari pengguna, yang kemudian diproses dan diklasifikasikan ke dalam tiga kategori sentimen: positif, negatif, dan netral. Seluruh tahapan dirancang secara sistematis untuk memastikan bahwa proses klasifikasi dilakukan secara objektif dan dapat diandalkan. Pemilihan metode ini bertujuan untuk mengevaluasi efektivitas model *Naive Bayes* dalam mengenali pola sentimen dalam data teks. Detail metode penelitian dapat dilihat pada gambar 1 berikut



**Gambar 1.** Metode Penelitian

Berikut ini keterangan setiap tahapan dari Gambar 1 metode penelitian.

1. Dataset (Kaggle)  
Data diambil dari Kaggle sebagai sumber utama. Dataset ini berisi informasi yang relevan untuk proses klasifikasi, misalnya teks ulasan, label sentimen, atau data kategori lainnya.
2. Preprocessing  
Data dibersihkan dan disiapkan, seperti menghapus data kosong, duplikasi, simbol, serta melakukan tokenisasi, stopword removal, dan stemming untuk data teks.
3. TF-IDF  
Data teks dikonversi menjadi vektor angka menggunakan TF-IDF, yang mengukur pentingnya kata dalam dokumen relatif terhadap seluruh korpus, untuk keperluan pemodelan.
4. 4. Training & Testing Data  
Data dibagi menjadi data latih dan data uji, biasanya dengan rasio 80:20. Data latih digunakan untuk membangun model, sementara data uji untuk mengevaluasi performanya.

## 5. Naive Bayes

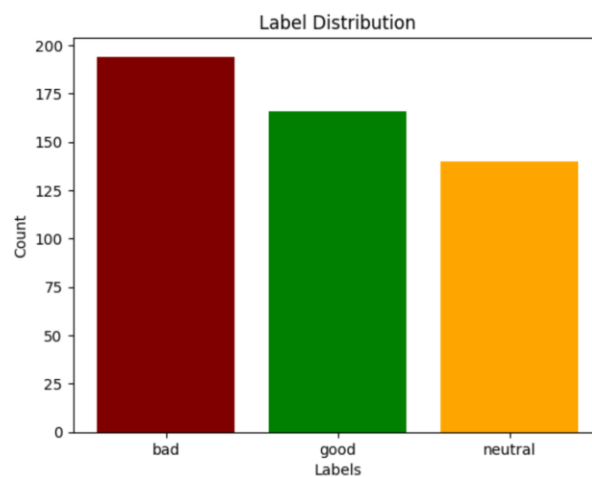
Model klasifikasi Naive Bayes dilatih menggunakan data latih. Algoritma ini menghitung probabilitas kelas berdasarkan fitur, dengan asumsi independensi antar fitur.

## 6. Result & Analysis

Hasil prediksi dievaluasi menggunakan metrik seperti akurasi, presisi, recall, F1-score, dan confusion matrix untuk menilai kinerja model.

### 2.1 Dataset

Dataset diperoleh dari data *open source* Kaggle. Data berisi 500 baris yang memiliki 3 kolom terdiri dari nomor, tweet, dan label dari tweet dengan berbahasa Inggris. Dataset yang diperoleh sudah memiliki label yang langsung diberikan oleh admin unggahan yang diambil selama 1 bulan pada tahun 2023 lalu sehingga termasuk kedalam *supervised learning*.



**Gambar 2.** Distribusi Labeling

Gambar 2. merupakan hasil dari labeling dataset yang diperoleh yang dibagi menjadi 3 class yaitu Bad = Jelek, Good = Bagus, Neutral = Netral. Label ini diperlukan agar dapat melihat performa yang dihasilkan oleh model seberapa akurat nya, karena dari dataset didapatkan sudah mendapatkan label nya maka tidak perlu memberikan labeling pada saat pemrosesan nya sebelum implementasi model tetapi jika label belum didapatkan maka perlu dilakukan perlabelan dengan menggunakan library yang disediakan oleh scikit-learn sesuai dengan kebutuhan.

### 2.2 Preprocessing

Setelah data diperoleh, data melewati tahap *preprocessing* dengan menggunakan tahap *case folding*, *tokenizing* berdasarkan mengikuti tahap-tahap yang dilakukan oleh peneliti terdahulu dan menyesuaikan kondisi dataset yang dimiliki [11][12][13][14]. *Case folding* merupakan tahap untuk memproses dataset yang memiliki campuran karakter CAPSLOCK dan *lowercase*, oleh karena itu sebelum masuk ke tahap analisis diperlukan *case folding* untuk membuat seluruh text yang akan di proses menjadi *lowercase*. *Tokenizing* merupakan tahap untuk membuat kalimat menjadi kata sehingga menghasilkan token-token kata menggunakan library NLTK. Lalu tahap akhir yaitu normalisasi dengan mengubah kata menjadi kata dasar berdasarkan corpus yang digunakan dengan menggunakan library.

### 2.3 TF-IDF

Sebelum digunakan dalam proses pelatihan model, data teks perlu dikonversi ke dalam format yang dapat diproses oleh algoritma *machine learning*. Salah satu pendekatan yang diterapkan dalam penelitian ini adalah **Term Frequency-Inverse Document Frequency (TF-IDF)**. TF-IDF merupakan teknik representasi teks yang mengubah kumpulan kata menjadi nilai numerik berdasarkan frekuensi kemunculan kata dalam satu dokumen (*term frequency*) dan kelangkaan kata tersebut di seluruh dokumen (*inverse document frequency*) [15][16][17]. Dengan metode ini, kata-kata yang memiliki makna penting dalam konteks tertentu akan diberikan bobot lebih besar, sementara kata-kata umum yang sering muncul di banyak dokumen—seperti "the", "is", atau "and"—akan diberikan bobot lebih kecil.

Penggunaan TF-IDF dalam penelitian ini bertujuan untuk mengekstraksi informasi penting dari teks tweet agar model *Naive Bayes* dapat mengidentifikasi pola sentimen secara lebih akurat. Teknik ini dianggap tepat untuk analisis sentimen karena mampu menyoroti kata-kata yang berperan dalam membentuk opini atau ekspresi emosi pengguna. Hasil transformasi dari TF-IDF kemudian digunakan sebagai input untuk melatih dan menguji model klasifikasi berbasis *Naive Bayes*. Persamaan 1,2 dan 3 merupakan detail yang digunakan untuk fitur ekstraksi pada penelitian ini.

$$TF(t, d) = \frac{N(t, d)}{T} \quad (1)$$

Persamaan (1) di dalam konteks ini, TF(t,d) menunjukkan seberapa sering istilah t muncul dalam dokumen d, N(t,d) adalah banyaknya kemunculan istilah t dalam dokumen d, sedangkan T merupakan total keseluruhan istilah yang ada di dalam dokumen tersebut. Dengan demikian, setiap kombinasi dokumen dan kata akan memiliki nilai TF(t,d) yang berbeda-beda [18].

$$IDF(t) = \frac{\log N}{N(t)} \quad (2)$$

Persamaan (2) menggambarkan cara menghitung IDF(t), yaitu frekuensi dokumen invers dari istilah t. Di sini, N merupakan jumlah total dokumen, sedangkan N(t) adalah jumlah dokumen yang mengandung istilah t [18]

$$TF - IDF = TF * IDF \quad (3)$$

Persamaan (3) rumus perhitungan dari TF-IDF setelah didapatkan nilai frekuensi dari TF dan IDF [18].

## 2.4 Split Data

Pada penelitian ini, dataset dibagi menjadi dua bagian utama, yaitu data untuk pelatihan (*training*) dan data untuk pengujian (*testing*), dengan perbandingan 80:20. Sebanyak 80% dari keseluruhan data dimanfaatkan untuk melatih model agar dapat mempelajari pola-pola yang terkandung dalam data, sementara 20% sisanya digunakan untuk mengevaluasi kemampuan model dalam mengklasifikasikan data yang belum dikenalnya. Strategi pembagian ini dilakukan untuk memastikan bahwa model tidak hanya mampu bekerja dengan baik pada data pelatihan, tetapi juga dapat melakukan prediksi yang akurat terhadap data baru. Teknik *train-test split* ini merupakan metode yang lazim digunakan dalam proses pembelajaran mesin, karena dapat membantu mencegah terjadinya *overfitting* serta memberikan gambaran yang lebih objektif mengenai performa model.

## 2.5 Implementasi Model

Setelah melewati proses pra-pemrosesan dan ekstraksi fitur menggunakan metode TF-IDF, tahap berikutnya dalam penelitian ini adalah menerapkan algoritma *Naive Bayes* sebagai model untuk klasifikasi sentimen. Algoritma ini dipilih karena kemampuannya yang tinggi dalam mengelola data berbasis teks secara cepat dan akurat, khususnya untuk permasalahan klasifikasi seperti analisis opini. *Naive Bayes* bekerja dengan pendekatan probabilistik dan

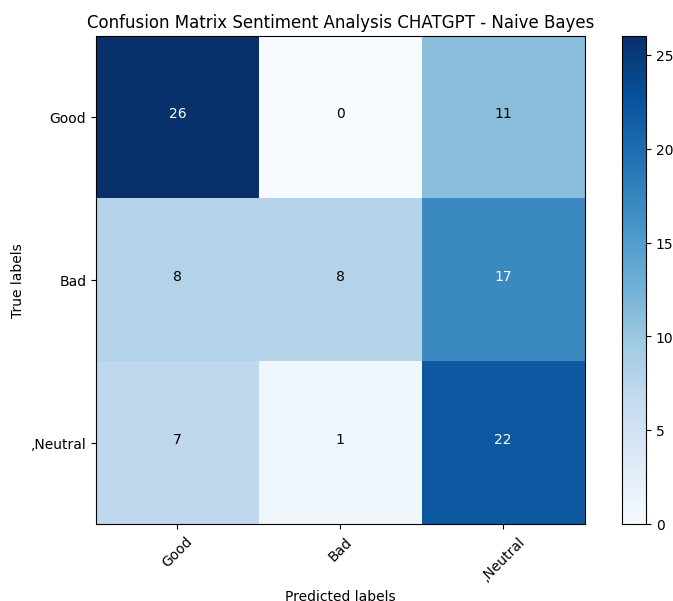
mengasumsikan bahwa setiap fitur bersifat independen terhadap kelas target, meskipun asumsi tersebut cukup sederhana, namun terbukti efektif dalam banyak kasus. Dalam pelaksanaannya, model dilatih menggunakan data pelatihan yang telah dibagi sebelumnya, kemudian diuji menggunakan data pengujian untuk menilai kinerjanya. Model akan menghasilkan prediksi kategori sentimen—positif, netral, atau negatif—yang selanjutnya dianalisis menggunakan metrik evaluasi seperti *confusion matrix*, *precision*, *recall*, dan *f1-score*. Melalui pendekatan ini, diharapkan model mampu menghasilkan klasifikasi yang andal dan memberikan wawasan mengenai kecenderungan opini pengguna terhadap ChatGPT.

### 3. Hasil dan Pembahasan

Berdasarkan hasil evaluasi model *Naive Bayes*, diperoleh akurasi sebesar 56% yang menunjukkan bahwa model masih memiliki keterbatasan dalam mengklasifikasikan sentimen pengguna terhadap ChatGPT. Dari laporan klasifikasi, terlihat bahwa kelas negatif (-1) memiliki *f1-score* tertinggi sebesar 0.67, sedangkan kelas netral (0) justru memiliki *f1-score* terendah yaitu 0.38, meskipun *precision*-nya tergolong tinggi (0.89). Hal ini mengindikasikan bahwa model cukup yakin saat memprediksi kelas netral, namun tidak mampu mengenali sebagian besar data netral secara akurat (*recall* hanya 0.24). Di sisi lain, kelas positif (1) memiliki *recall* tinggi (0.73) namun *precision*-nya rendah (0.44), menunjukkan bahwa model cenderung sering memprediksi kelas positif meskipun tidak selalu benar. Hasil *confusion matrix* memperkuat temuan ini, di mana banyak data dari kelas negatif dan netral justru salah diklasifikasikan ke dalam kelas positif. Ketidakseimbangan ini menunjukkan adanya bias model terhadap kelas tertentu, khususnya kelas positif. Secara keseluruhan, performa model masih belum optimal dan dapat ditingkatkan melalui metode seperti *balancing data*, *optimasi parameter*, atau penggunaan algoritma klasifikasi lain yang lebih kompleks. Untuk detail lebih lanjut pada tabel 1 dan gambar 3

**Tabel 1.** Classification Report

Label	Precision	Recall	F1-Score	Support
-1	0.63	0.70	0.67	37
0	0.89	0.24	0.38	33
1	0.44	0.73	0.55	30
Accuracy			0.56	100
Macro avg	0.65	0.56	0.53	100
Weighted	0.66	0.56	0.54	100



**Gambar 3.** Confusion Matrix Sentiment Analysis Chatgpt

Table 1. Report hasil kinerja algoritma *Naive Bayes* dalam menganalisis sentimen terhadap ChatGPT masih tergolong kurang optimal, dengan tingkat akurasi keseluruhan hanya mencapai 56%. Terdapat ketidakseimbangan performa antar kelas yang cukup signifikan, terutama pada kelas netral. Meskipun precision untuk kelas ini cukup tinggi (0.89), nilai *recall*-nya sangat rendah (0.24), menunjukkan bahwa model sering kali tidak mampu mengenali data netral dengan benar. Hal ini mengindikasikan bahwa model cenderung mengklasifikasikan data sebagai netral tanpa mampu mengenali karakteristik khas dari sentimen tersebut. Pola yang serupa juga terlihat pada kelas positif (1), di mana *recall* tinggi (0.73) tidak diimbangi dengan precision yang memadai (0.44), mengakibatkan banyak data dari kelas lain seperti negatif atau netral salah terklasifikasi sebagai positif. Kondisi ini dapat disebabkan oleh distribusi data yang tidak seimbang atau kurangnya efektivitas fitur yang digunakan selama proses pelatihan. *Confusion matrix* semakin menegaskan kecenderungan model dalam melakukan kesalahan klasifikasi, khususnya data dari kelas ‘Bad’ dan ‘Neutral’ yang banyak diklasifikasikan secara keliru ke dalam kelas ‘Neutral’ atau ‘Good’. Masalah ini kemungkinan besar juga berkaitan dengan asumsi independensi antar fitur dalam *Naive Bayes* yang kurang sesuai untuk data teks, karena kata-kata dalam kalimat umumnya saling bergantung secara kontekstual. Akibatnya, model tidak mampu menangkap keterkaitan antar kata secara menyeluruh sehingga menyebabkan fitur representasi yang dihasilkan tidak cukup diskriminatif. Kecenderungan *overgeneralization* terhadap kata-kata yang diasosiasikan dengan sentimen tertentu turut memperparah akurasi klasifikasi. Oleh karena itu, untuk meningkatkan kinerja model, beberapa strategi dapat diterapkan oleh peneliti berikutnya. Pertama, analisis distribusi data secara menyeluruh sangat diperlukan. Jika ditemukan ketimpangan jumlah antar kelas, metode seperti SMOTE (*Synthetic Minority Over-sampling Technique*) atau teknik undersampling bisa digunakan untuk menciptakan dataset yang lebih seimbang. Kedua, kualitas fitur dapat ditingkatkan melalui teknik feature engineering, seperti penggunaan n-gram, TF-IDF, atau representasi kata yang lebih canggih seperti Word2Vec dan BERT untuk menangkap konteks makna kata secara lebih baik. Ketiga, Naive Bayes dapat digantikan atau dikombinasikan dengan algoritma lain seperti *Random Forest*, SVM, atau model berbasis deep learning seperti LSTM, yang lebih mampu menangani kompleksitas data teks. Selain itu, pengaturan parameter model (hyperparameter tuning) menggunakan grid search atau *cross-validation* juga penting dilakukan untuk mencapai performa optimal. Evaluasi hasil klasifikasi pun sebaiknya tidak hanya mengandalkan akurasi, namun juga mempertimbangkan metrik seperti *macro-averaged F1-score* agar evaluasi antar kelas lebih adil, khususnya jika data tidak seimbang. Temuan ini menunjukkan bahwa dalam analisis sentimen, kualitas hasil klasifikasi sangat dipengaruhi oleh pemilihan teknik pra-proses, representasi data teks, dan algoritma yang digunakan.

#### 4. Kesimpulan

Berdasarkan Berdasarkan hasil evaluasi terhadap komentar pengguna ChatGPT menggunakan algoritma Naive Bayes, ditemukan bahwa model belum mampu memberikan performa klasifikasi yang optimal. Akurasi yang dicapai hanya sebesar 56%, dengan *f1-score* tertinggi pada kelas negatif (0.67) dan terendah pada kelas netral (0.38). Ketidakseimbangan performa terlihat jelas, terutama pada kelas netral yang memiliki precision tinggi namun recall rendah, serta kelas positif yang menunjukkan *recall* tinggi tetapi precision rendah. *Confusion matrix* turut memperkuat temuan ini dengan memperlihatkan tingginya tingkat kesalahan klasifikasi antar kelas. Keterbatasan ini berkaitan dengan asumsi independensi fitur dalam algoritma Naive Bayes, yang kurang sesuai untuk data teks yang bersifat kontekstual. Oleh karena itu, untuk meningkatkan kinerja model, disarankan penerapan strategi seperti penyeimbangan distribusi data, penguatan representasi fitur melalui teknik NLP yang lebih canggih, serta eksplorasi algoritma klasifikasi lain yang lebih kompleks dan kontekstual.

#### Daftar Pustaka

- [1] S. Aggrawal dan A. J. Magana, “Teamwork Conflict Management Training and Conflict

- Resolution Practice via Large Language Models,” *Futur. Internet*, vol. 16, no. 5, hal. 1–25, 2024, doi: 10.3390/fi16050177.
- [2] M. Nanda Fahriza dan N. Riza, “Analisis Sentimen Pada Ulasan Aplikasi Chat Generative Pre-Trained Transformer Gpt Menggunakan Metode Klasifikasi K-Nearest Neighbor(Knn),” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 2, hal. 1351–1358, 2023, doi: 10.36040/jati.v7i2.6767.
- [3] M. S. Adhi, M. Z. Nafan, dan E. Usada, “Pengaruh Semantic Expansion pada Naïve Bayes Classifier untuk Analisis Sentimen Tokoh Masyarakat,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 3, no. 2, hal. 141–147, 2019, doi: 10.29207/resti.v3i2.901.
- [4] Merinda Lestandy, Abdurrahim Abdurrahim, dan Lailis Syafa’ah, “Analisis Sentimen Tweet Vaksin COVID-19 Menggunakan Recurrent Neural Network dan Naïve Bayes,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, hal. 802–808, 2021, doi: 10.29207/resti.v5i4.3308.
- [5] K. Aulia dan L. Amelia, “Analisis Sentimen Twitter Pada Isu Mental Health Dengan Algoritma Klasifikasi Naive Bayes,” *Siliwangi J. (Seri Sains Teknol.*, vol. 6, no. 2, hal. 60–65, 2020.
- [6] S. O. S. Dinauni, Herlawati, dan Rasim, “Analisis Sentimen Mengenai Gangguan Bipolar Pada Twitter Menggunakan Algoritma Naïve Bayes,” vol. 6, no. 2, hal. 63–73, 2022.
- [7] K. Li, C. Zhou, X. (Robert) Luo, J. Benitez, dan Q. Liao, “Impact of information timeliness and richness on public engagement on social media during COVID-19 pandemic: An empirical investigation based on NLP and machine learning,” *Decis. Support Syst.*, vol. 162, no. February, hal. 113752, 2022, doi: 10.1016/j.dss.2022.113752.
- [8] J. Frei, L. Frei-Stuber, dan F. Kramer, “GERNERMED++: Semantic annotation in German medical NLP through transfer-learning, translation and word alignment,” *J. Biomed. Inform.*, vol. 147, hal. 104513, 2023, doi: 10.1016/j.jbi.2023.104513.
- [9] A. F. Rahma, A. R. W. Wibowo, N. N. Nidya, dan Agussalim, “Analisis Sentimen Hashtag ‘Dirumahaja’ Saat Pandemi Covid-19 Di Indonesia Menggunakan Nlp,” *J. Inform. dan Sist. Inf.*, vol. 1, hal. 343–353, 2020, [Daring]. Tersedia pada: <http://jifosi.upnjatim.ac.id/index.php/jifosi/article/view/239>
- [10] N. L. P. M. Putu, Ahmad Zuli Amrullah, dan Ismarmiaty, “Analisis Sentimen dan Pemodelan Topik Pariwisata Lombok Menggunakan Algoritma Naive Bayes dan Latent Dirichlet Allocation,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 1, hal. 123–131, 2021, doi: 10.29207/resti.v5i1.2587.
- [11] A. I. Kadhim, Y. N. Cheah, dan N. H. Ahamed, “Text Document Preprocessing and Dimension Reduction Techniques for Text Document Clustering,” *Proc. - 2014 4th Int. Conf. Artif. Intell. with Appl. Eng. Technol. ICAIET 2014*, hal. 69–73, 2015, doi: 10.1109/ICAIET.2014.21.
- [12] H. T. Duong dan T. A. Nguyen-Thi, “A review: preprocessing techniques and data augmentation for sentiment analysis,” *Comput. Soc. Networks*, vol. 8, no. 1, hal. 1–16, 2021, doi: 10.1186/s40649-020-00080-x.
- [13] M. K. Sohrabi dan F. Hemmatian, “An efficient preprocessing method for supervised sentiment analysis by converting sentences to numerical vectors: a twitter case study,” *Multimed. Tools Appl.*, vol. 78, no. 17, hal. 24863–24882, 2019, doi: 10.1007/s11042-019-7586-4.
- [14] D. Rifaldi, A. Fadlil, dan Herman, “Teknik Preprocessing Pada Text Mining Menggunakan Data Tweet Mental Health,” *J. Pendidik. Teknol. Inf.*, vol. 3, no. 2, hal. 161–171, 2023.
- [15] R. Ahuja, A. Chug, S. Kohli, S. Gupta, dan P. Ahuja, “ScienceDirect ScienceDirect The Impact of Features Extraction on the Sentiment Analysis,” *Procedia Comput. Sci.*, vol. 152, hal. 341–348, 2019, doi: 10.1016/j.procs.2019.05.008.
- [16] D. Sunitha, R. K. Patra, N. V. Babu, A. Suresh, dan S. C. Gupta, “Twitter sentiment

- analysis using ensemble based deep learning model towards COVID-19 in India and European countries,” *Pattern Recognit. Lett.*, vol. 158, hal. 164–170, 2022, doi: 10.1016/j.patrec.2022.04.027.
- [17] D. Rifaldi, A. Fadlil, dan Herman, “Implementation of Word Trends Using a Machine Learning Approach with TF-IDF and Latent Dirichlet Allocation,” *Int. J. Informatics Vis.*, vol. 8, no. 4, hal. 2297–2304, 2024, doi: 10.62527/joiv.8.4.2452.
- [18] A. S. Neogi, K. A. Garg, R. K. Mishra, dan Y. K. Dwivedi, “Sentiment analysis and classification of Indian farmers’ protest using twitter data,” *Int. J. Inf. Manag. Data Insights*, vol. 1, no. 2, hal. 100019, 2021, doi: 10.1016/j.jjime.2021.100019.