JURNAL PENDIDIKAN DAN APLIKASI INDUSTRI
Vol. 13 No. 1 2026 : 71-85
ISSN : 0126 – 4036 e-ISSN : 2716 – 0416

Analisis Sentimen Ulasan Pengguna Aplikasi ChatGPT pada Google Play Store Menggunakan IndoBERT dan Support Vector Machine

Syahriani Syam[1]*, Siti Muryanah[2], Haryanto[3]

[123]*Program Studi Teknik Informatika,, Universitas Islam Syekh Yusuf,Tangerang*

[1]ssyam@unis.ac.id,[2]siti.muyanah@unis.ac.id, [3]haryanto@unis.ac.id

Article History:
Received 29 Jun 2026
Revised 22 Jul 2026
Accepted 28 Jul 2026
Available online 28 Jul 2026

Abstrak

Perkembangan Large Language Model (LLM) mendorong peningkatan penggunaan aplikasi ChatGPT dalam berbagai aktivitas, seperti pembelajaran, penulisan, pemrograman, dan pencarian informasi. Meningkatnya jumlah pengguna menghasilkan ulasan dalam jumlah besar pada Google Play Store yang mencerminkan pengalaman, tingkat kepuasan dan berbagai permasalahan yang dihadapi pengguna. Penelitian ini bertujuan menganalisis sentimen ulasan pengguna aplikasi ChatGPT pada Google Play Store. Metode yang digunakan adalah pendekatan hibrida IndoBERT dan Support Vector Machine (SVM) pada 312 ulasan yang telah diproses dan dilabeli secara manual. Pra-pemrosesan data ulasan (n=312) dilakukan melalui *case folding*, *tokenizing*, dan *stop word removal*, dilanjutkan pelabelan manual yang mengungkapkan distribusi sentimen 64,10% positif, 24,04% negatif, dan 11,86% netral. Model hibrida IndoBERT-SVM menunjukkan performa klasifikasi sentimen yang luar biasa, mencapai akurasi 93,55%, presisi 92,80%, *recall* 93,55%, dan F1-score 93,17%. Performa ini signifikan melampaui metode klasifikasi lain; akurasi IndoBERT-SVM lebih tinggi dibandingkan Naïve Bayes (87,00%) dan SVM tunggal (92,00%) dalam studi komparatif. Keunggulan ini disebabkan oleh kemampuan IndoBERT dalam mengekstraksi representasi fitur semantik dan kontekstual yang kaya dari teks berbahasa Indonesia, yang kemudian dioptimalkan oleh SVM untuk menemukan hyperplane klasifikasi yang optimal.

Kata kunci: *Analisis Sentimen, ChatGPT, IndoBERT, Support Vector Machine, Ulasan Pengguna*

Abstract

The rapid advancement of Large Language Models (LLMs) has significantly increased the adoption of ChatGPT for various activities, including learning, writing, programming, and information retrieval. The growing number of users has generated a substantial volume of reviews on the Google Play Store, reflecting user experiences, satisfaction levels, and various issues encountered while using the application. This study aimed to analyze the sentiment of ChatGPT application user reviews on the Google Play Store. The method employed was a hybrid approach combining IndoBERT and Support Vector Machine (SVM) on 312 manually labeled and pre-processed reviews. Data pre-processing (n=312) involved case folding, tokenizing, and stop word removal, followed by manual labeling which revealed a sentiment distribution of 64.10% positive, 24.04% negative, and 11.86% neutral. The hybrid IndoBERT-SVM model demonstrated exceptional sentiment classification performance, achieving an accuracy of 93.55%, precision of 92.80%, recall of 93.55%, and an F1-score of 93.17%. This performance significantly surpassed other classification methods; IndoBERT-SVM's accuracy was higher compared to Naïve Bayes (87.00%) and a standalone SVM (92.00%) in comparative studies. This superiority stems from IndoBERT's ability to extract rich semantic and contextual feature representations from Indonesian text, which is then optimized by SVM to find an optimal classification hyperplane.

Keywords: *ChatGPT, IndoBERT, Sentiment Analysis, Support Vector Machine, User Reviews*

1. Pendahuluan

Penggunaan aplikasi seluler terus meningkat pada era digital sehingga pemahaman terhadap kepuasan dan ketidakpuasan pengguna menjadi semakin penting (Salsabila & Muntahanah, 2024). enomena ini sejalan dengan pesatnya adopsi chatbot berbasis kecerdasan buatan, seperti ChatGPT, yang mencapai 100 juta pengguna aktif hanya dalam dua bulan sejak peluncurannya dan menghasilkan banyak ulasan di berbagai platform digital (Haikal Hidayatullah et al., 2023).

Ulasan pada Google Play Store merupakan sumber informasi penting untuk mengevaluasi persepsi pengguna terhadap kualitas aplikasi (Fatmawati et al., 2024). Analisis sentimen memungkinkan pengembang mengidentifikasi kelebihan dan kekurangan aplikasi secara sistematis, sedangkan analisis manual menjadi kurang efisien ketika jumlah ulasan sangat besar (Iis Aisah et al., 2024 , Haikal Hidayatullah et al., 2023).

Penelitian terdahulu telah menerapkan berbagai metode untuk analisis sentimen ulasan aplikasi. Misalnya, Naïve Bayes digunakan untuk menganalisis sentimen pengguna aplikasi transportasi daring (Fatmawati et al., 2024) dan aplikasi ChatGPT di media sosial (Haikal Hidayatullah et al., 2023; I. Putra et al., 2025). Metode K-Nearest Neighbor (KNN) juga diterapkan untuk mengklasifikasikan sentimen ulasan aplikasi Identitas Kependudukan Digital, meskipun mengalami kesulitan dalam mengklasifikasikan ulasan netral (Mariana Ulfa et al., 2024).

Metode Support Vector Machine (SVM) menunjukkan kinerja efektif dalam klasifikasi sentimen ulasan pengguna aplikasi ChatGPT, dengan akurasi 91% dan F1-score 90% (Aprilia Salsabila & Muntahanah, 2024). SVM juga memberikan hasil yang baik dalam klasifikasi karena kemampuannya menemukan hyperplane optimal untuk memisahkan kelas berbeda (Akbar Kurniawan & Irfan Pratama, 2025). Penerapan SVM pada ulasan aplikasi Al-Qur'an digital dan Access by KAI juga menunjukkan potensi tinggi dalam memahami respons pengguna dan meningkatkan kualitas layanan (Iis Aisah et al., 2024; Raka Kusuma Wardana et al., 2025). Meskipun metode SVM menunjukkan performa yang baik, potensi peningkatan akurasi melalui kombinasi dengan model bahasa yang lebih canggih seperti IndoBERT belum banyak dieksplorasi secara spesifik untuk ulasan aplikasi ChatGPT berbahasa Indonesia. Mayoritas penelitian terdahulu seringkali berfokus pada satu metode klasifikasi atau menggunakan data dari platform selain Google Play Store. Integrasi IndoBERT sebagai fitur ekstraksi dan SVM sebagai classifier dapat mengatasi keterbatasan representasi tekstual pada model tradisional dan meningkatkan robusta analisis sentimen pada ulasan yang kompleks.

Kebaruan penelitian ini terletak pada pengembangan model hibrida yang memanfaatkan IndoBERT sebagai ekstraktor fitur dan Support Vector Machine sebagai classifier untuk analisis sentimen ulasan pengguna aplikasi ChatGPT berbahasa Indonesia. Pendekatan ini berbeda dari penelitian sebelumnya yang umumnya hanya menggunakan metode statistik tradisional atau satu algoritma machine learning tanpa memanfaatkan representasi semantik dari model bahasa modern.

Penelitian ini berkontribusi dalam tiga hal. Pertama, kontribusi teoretis dengan mengeksplorasi sinergi antara model bahasa pra-terlatih (IndoBERT) dan algoritma machine learning klasik (SVM) dalam tugas analisis sentimen. Kedua, kontribusi praktis berupa peningkatan akurasi analisis sentimen ulasan aplikasi, yang dapat dimanfaatkan oleh pengembang ChatGPT untuk perbaikan kualitas aplikasi. Ketiga, kontribusi metodologis melalui pengembangan model hibrida yang dapat dijadikan referensi untuk analisis sentimen pada dataset ulasan aplikasi berbahasa Indonesia lainnya.

Pertanyaan penelitian yang akan dijawab adalah:

RQ1: Bagaimana performa model IndoBERT yang dikombinasikan dengan Support Vector Machine dalam mengklasifikasikan sentimen ulasan pengguna aplikasi ChatGPT berbahasa Indonesia?

RQ2: Sejauh mana perbandingan akurasi model hibrida IndoBERT-SVM dengan metode klasifikasi sentimen lainnya dalam konteks ulasan aplikasi ChatGPT?

Hipotesis penelitian ini adalah:

H1: Model IndoBERT yang dikombinasikan dengan Support Vector Machine akan menunjukkan akurasi klasifikasi sentimen yang tinggi pada ulasan pengguna aplikasi ChatGPT berbahasa Indonesia, melebihi 90%.

H2: Model hibrida IndoBERT-SVM akan memberikan performa akurasi yang lebih unggul dibandingkan dengan metode klasifikasi sentimen lain yang umum digunakan dalam analisis sentimen ulasan aplikasi.

Analisis Sentimen

Analisis sentimen merupakan bagian dari *Natural Language Processing (NLP)* yang bertujuan mengidentifikasi dan mengklasifikasikan opini atau emosi dalam teks ke dalam kategori positif, negatif, maupun netral (Tri Yuli Pahtoni & Handaru Jati, 2024). Teknik ini banyak diterapkan pada media sosial maupun ulasan aplikasi untuk memahami persepsi pengguna terhadap suatu produk atau layanan.

Pada aplikasi berbasis kecerdasan buatan seperti ChatGPT, analisis sentimen menjadi penting karena jumlah ulasan yang sangat besar menyulitkan proses evaluasi secara. (Bayu Purbayanto & Teguh Nurhadi Suharsono, 2024). Hasil analisis dapat membantu pengembang mengidentifikasi masalah, mengevaluasi kualitas layanan, serta mendukung pengambilan keputusan berbasis data dalam pengembangan aplikasi.

Aplikasi ChatGPT dan Ulasan Pengguna

hatGPT merupakan model bahasa generatif yang banyak dimanfaatkan untuk menjawab pertanyaan, menyusun dokumen, hingga membantu pemrograman (Bayu Purbayanto & Teguh Nurhadi Suharsono, 2024). Popularitasnya menghasilkan jutaan ulasan pengguna di Google Play Store yang mencerminkan pengalaman terhadap kemudahan penggunaan, kualitas layanan, maupun kendala yang dihadapi.

Ulasan tersebut menjadi sumber informasi penting untuk mengevaluasi kepuasan pengguna karena mampu mengungkap permasalahan yang belum tentu ditemukan melalui pengujian internal (Ikbal Bazar et al., 2025). Oleh karena itu, analisis sentimen berperan penting dalam mendukung peningkatan kualitas aplikasi secara berkelanjutan. Mayoritas ulasan untuk aplikasi sejenis cenderung positif, tetapi penting untuk tidak mengabaikan sentimen netral dan negatif yang seringkali menyoroti area perbaikan kritis (Fahmi Ahmad Tohir et al., 2024)

Metode Klasifikasi Sentimen

Berbagai algoritma telah digunakan dalam analisis sentimen, di antaranya *Naïve Bayes* dan *Support Vector Machine (SVM). Naïve Bayes* dikenal sederhana dan efisien serta dapat ditingkatkan performanya melalui teknik seperti SMOTE untuk mengatasi ketidakseimbangan data (Primandika Hakiki et al., 2025; Hermawan & Dzikrillah, 2024)

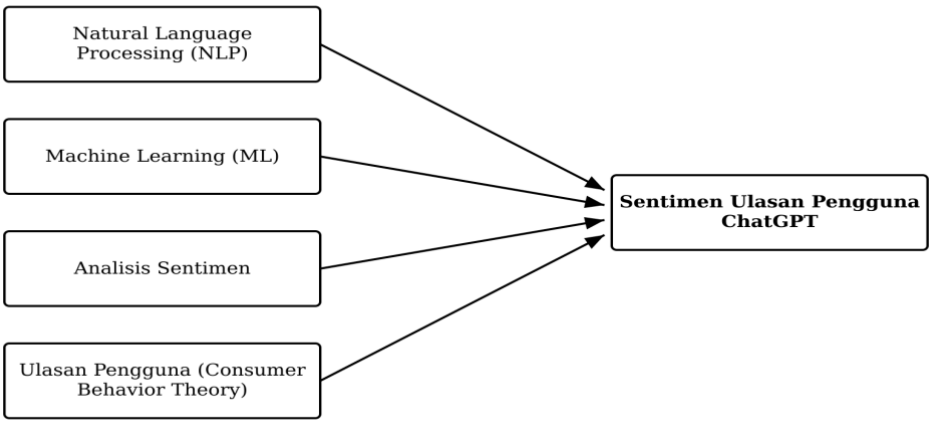
Sementara itu, SVM memiliki kemampuan membentuk hyperplane optimal sehingga menghasilkan generalisasi yang baik pada data baru (Ikbal Bazar et al., 2025). Berbagai penelitian menunjukkan bahwa SVM memberikan akurasi, presisi, dan F1-score yang lebih tinggi dibandingkan Naïve Bayes, terutama ketika dikombinasikan dengan optimasi kernel atau teknik penyeimbangan data (Ni Putu Anik Juniantini et al., 2024; Muhammad Faqih Baihaqi et al., 2025)

Pada analisis sentimen aplikasi ChatGPT, sebagian besar penelitian masih menggunakan Naïve Bayes dengan hasil mayoritas sentimen positif terhadap aplikasi (Fahmi Ahmad Tohir et al., 2024) Penelitian lain membandingkan berbagai algoritma dan menunjukkan bahwa SVM secara konsisten memberikan performa lebih baik dibandingkan metode klasifikasi konvensional. (Bayu Purbayanto & Teguh Nurhadi Suharsono, 2024).

Beberapa studi lain memperluas cakupan analisis sentimen dengan membandingkan beberapa metode atau mengintegrasikan teknik pemrosesan data yang lebih canggih. (Iyeti Astri Prasasti & Amarudin, 2025) membandingkan Decision Tree dengan Naïve Bayes untuk menganalisis sentimen pengguna Twitter terhadap konten edukasi kesehatan, di mana Decision Tree menunjukkan akurasi 84%. (Muhammad Faqih Baihaqi et al., 2025) secara komprehensif membandingkan Naïve Bayes dan Support Vector Machine (SVM) untuk klasifikasi sentimen ulasan aplikasi DeepSeek Assistant AI. Hasil perbandingan menunjukkan bahwa SVM lebih unggul dalam akurasi, F1-Score, dan presisi dibandingkan Naïve Bayes, menegaskan potensi SVM dalam tugas klasifikasi sentimen.

Studi lain juga menunjukkan pentingnya optimasi dan ekstraksi fitur dalam meningkatkan performa model. (Ikbal Bazar et al., 2025) menggunakan SVM dengan optimasi kernel trick untuk menganalisis sentimen ulasan aplikasi perbankan Wondr by BNI. Mereka menemukan bahwa kernel Linear dan Polynomial memberikan performa terbaik dengan akurasi 87,9% dan F1-score 87,8%. Pendekatan ini menunjukkan bahwa pemilihan kernel yang tepat dapat memengaruhi kinerja SVM dalam analisis sentimen. Penggunaan model bahasa pra-terlatih seperti BERT (Bidirectional Encoder Representations

from Transformers) juga telah merevolusi kemampuan ekstraksi fitur tekstual, meskipun belum secara eksplisit disebutkan dalam



Gambar 1 Model konseptual penelitian

2. Bahan dan Metode

Penelitian ini menerapkan pendekatan kuantitatif dengan desain deskriptif-analitis untuk menganalisis sentimen ulasan pengguna aplikasi ChatGPT. Pendekatan kuantitatif menjadi pilihan karena fokusnya pada pengumpulan dan analisis data numerik, yang memungkinkan identifikasi pola, tren, serta hubungan antar variabel secara objektif (Tri Yuli Pahtoni & Handaru Jati, 2024). Desain deskriptif-analitis memiliki tujuan menggambarkan karakteristik data ulasan sekaligus menganalisis hubungan antara fitur teks dengan kategori sentimen melalui penggunaan model prediktif. Desain ini relevan untuk mengukur performa model klasifikasi sentimen pada data ulasan aplikasi. Seluruh proses analisis dilakukan menggunakan bahasa pemrograman Python pada lingkungan Google Colaboratory (Google Colab) yang menyediakan pustaka *machine learning* dan *deep learning* yang mendukung proses pengolahan data secara efisien.

Tahapan penelitian dimulai dari pengumpulan ulasan pengguna aplikasi ChatGPT yang diperoleh melalui Google Play Store. Dataset yang diperoleh selanjutnya melalui proses *preprocessing* untuk meningkatkan kualitas data dengan menghilangkan karakter yang tidak diperlukan, menyeragamkan bentuk teks, menghapus kata-kata yang tidak memiliki makna sentimen, serta melakukan stemming. Setelah data dipersiapkan, representasi fitur dilakukan menggunakan model IndoBERT sehingga setiap ulasan dapat direpresentasikan dalam bentuk vektor semantik yang mampu mempertahankan konteks bahasa Indonesia.

Untuk meningkatkan kualitas klasifikasi, penelitian ini menerapkan Synthetic Minority Over-sampling Technique (SMOTE) dalam menangani ketidakseimbangan distribusi kelas pada data pelatihan. Selanjutnya, model Support Vector Machine (SVM) dibangun menggunakan parameter optimum yang diperoleh melalui proses GridSearchCV. Sebagai bahan evaluasi, penelitian ini juga membandingkan performa model berbasis IndoBERT dengan pendekatan representasi fitur konvensional, yaitu TF-IDF + SVM dan Word2Vec + SVM. Evaluasi kinerja model dilakukan menggunakan metrik accuracy, precision, recall, F1-score, dan confusion matrix sehingga dapat diketahui efektivitas masing-masing pendekatan dalam mengklasifikasikan sentimen ulasan pengguna.

Alur penelitian secara keseluruhan ditunjukkan pada Gambar 2, yang menggambarkan tahapan mulai dari pengumpulan data, *preprocessing*, ekstraksi fitur, pembentukan model klasifikasi, hingga proses evaluasi dan perbandingan performa model.

Sumber dan Pengumpulan Data

Gambar 2. Framework Penelitian Analisis Sentimen ChatGPT

Data penelitian ini berasal dari ulasan pengguna aplikasi ChatGPT yang tersedia di Google Play Store. Peneliti mengumpulkan ulasan memakai teknik _web scraping_ dari halaman aplikasi ChatGPT pada Google Play Store. Proses pengumpulan data berlangsung pada bulan Mei 2026, sehingga terkumpul total 312 ulasan pengguna berbahasa Indonesia. Setiap ulasan meliputi teks ulasan dan rating bintang yang pengguna berikan, yang selanjutnya berguna untuk pelabelan sentimen awal. Data ulasan ini menjadi dasar utama untuk proses ekstraksi fitur serta klasifikasi sentimen.

Dataset hasil seleksi kemudian diperiksa kualitasnya (*data understanding*) untuk memastikan bahwa seluruh data layak digunakan pada proses analisis sentimen. Tahap ini meliputi pemeriksaan jumlah data, tipe data, nilai yang hilang (*missing values*), serta distribusi data pada setiap atribut. Setelah proses pemeriksaan selesai, dataset diteruskan ke tahap *preprocessing* untuk membersihkan dan menyeragamkan data teks sebelum dilakukan proses ekstraksi fitur dan klasifikasi.

Tabel 1 Deskripsi Dataset

Atribut	Keterangan
User Name	Nama pengguna (tidak digunakan dalam analisis)
Review	Isi ulasan pengguna
Rating	Nilai rating (1–5)
Date	Tanggal ulasan
Sentiment	Label sentimen (Positif, Negatif, Netral)

Tahapan pengumpulan data hingga data siap diproses pada penelitian ini digambarkan secara keseluruhan pada Gambar 1. Setelah dataset diperoleh dan melalui proses seleksi awal, penelitian dilanjutkan ke tahap *preprocessing* yang bertujuan meningkatkan kualitas data teks sebelum dilakukan ekstraksi fitur menggunakan IndoBERT, pembentukan model klasifikasi Support Vector Machine (SVM), serta evaluasi performa model.

Tahapan Penelitian

Tahapan penelitian dilakukan secara sistematis mulai dari proses pengumpulan data hingga evaluasi model klasifikasi sentimen. Alur penelitian secara keseluruhan ditunjukkan pada Gambar 1. Setiap tahapan dirancang untuk menghasilkan model klasifikasi yang mampu mengidentifikasi sentimen ulasan pengguna aplikasi ChatGPT secara akurat dan objektif.

Tahap pertama adalah pengumpulan data, yaitu memperoleh ulasan pengguna aplikasi ChatGPT dari Google Play Store menggunakan teknik *web scraping*. Data yang diperoleh kemudian memasuki tahap data understanding, yang bertujuan untuk memahami karakteristik dataset melalui pemeriksaan struktur data, identifikasi atribut, pengecekan data duplikat, serta pemeriksaan *missing values*. Tahap ini penting untuk memastikan bahwa dataset yang digunakan memiliki kualitas yang memadai sebelum diproses lebih lanjut.

Tahap berikutnya adalah preprocessing data, yang bertujuan meningkatkan kualitas data teks sehingga siap digunakan dalam proses pembentukan model. Pada penelitian ini, preprocessing meliputi proses *cleaning*, *case folding*, *normalization*, *stopword removal*, dan *stemming*. Seluruh tahapan tersebut dilakukan secara berurutan untuk menghilangkan karakter yang tidak diperlukan, menyeragamkan format penulisan, mengubah kata tidak baku menjadi bentuk baku, menghapus kata-kata yang tidak memiliki kontribusi terhadap analisis sentimen, serta mengembalikan kata ke bentuk dasarnya.

Setelah proses preprocessing selesai, dataset dibagi menjadi data pelatihan (*training set*), data validasi (*validation set*), dan data pengujian (*testing set*) dengan perbandingan 80:10:10. Pembagian dataset dilakukan untuk memastikan bahwa proses pelatihan, optimasi model, dan evaluasi dilakukan pada data yang berbeda sehingga hasil pengujian dapat menggambarkan kemampuan generalisasi model secara lebih objektif.

Tahap selanjutnya adalah ekstraksi fitur (*feature extraction*), yaitu mengubah data teks menjadi representasi numerik yang dapat diproses oleh algoritma klasifikasi. Penelitian ini membandingkan tiga pendekatan representasi fitur, yaitu TF-IDF, Word2Vec, dan IndoBERT. TF-IDF digunakan sebagai representasi berbasis frekuensi kata, Word2Vec digunakan untuk menghasilkan representasi vektor berdasarkan hubungan antar kata, sedangkan IndoBERT digunakan untuk menghasilkan representasi kontekstual yang mampu memahami makna kata sesuai konteks kalimat dalam bahasa Indonesia.

Untuk mengatasi ketidakseimbangan distribusi kelas pada data pelatihan, diterapkan Synthetic Minority Over-sampling Technique (SMOTE) setelah proses ekstraksi fitur. SMOTE digunakan hanya pada data pelatihan untuk menghasilkan distribusi kelas yang lebih seimbang tanpa memengaruhi data validasi maupun data pengujian. Pendekatan ini bertujuan mengurangi bias model terhadap kelas mayoritas sekaligus meningkatkan kemampuan model dalam mengenali kelas minoritas.

Proses pelatihan model dilakukan menggunakan algoritma Support Vector Machine (SVM). Sebelum model dibangun, dilakukan optimasi parameter menggunakan GridSearchCV untuk memperoleh kombinasi parameter terbaik berdasarkan proses validasi silang (*cross-validation*). Optimasi ini bertujuan meningkatkan performa model sehingga mampu menghasilkan klasifikasi sentimen yang lebih akurat.

Tahap terakhir adalah evaluasi model. Model yang telah dilatih diuji menggunakan data pengujian untuk memperoleh nilai accuracy, precision, recall, F1-score, dan confusion matrix. Selain mengevaluasi performa masing-masing model, penelitian ini juga membandingkan hasil klasifikasi antara pendekatan TF-IDF + SVM, Word2Vec + SVM, dan IndoBERT + SVM. Hasil perbandingan tersebut digunakan untuk menentukan representasi fitur yang memberikan performa terbaik dalam analisis sentimen ulasan pengguna aplikasi ChatGPT.

Setiap tahapan pada Gambar 1 dijelaskan secara lebih rinci pada preprocessing data, ekstraksi fitur, penanganan ketidakseimbangan data, pembentukan model klasifikasi, serta evaluasi performa model.

Preprocessing Data

Tahap *preprocessing* merupakan proses awal yang bertujuan meningkatkan kualitas data teks sebelum dilakukan ekstraksi fitur dan proses klasifikasi. Tahapan ini dilakukan untuk mengurangi *noise*, menyeragamkan format penulisan, serta menghasilkan representasi teks yang lebih bersih sehingga dapat meningkatkan performa model klasifikasi. Pada penelitian ini, proses *preprocessing* dilakukan secara berurutan menggunakan bahasa pemrograman Python pada lingkungan Google Colaboratory (Google Colab). Tahapan *preprocessing* yang diterapkan meliputi *cleaning*, *case folding*, *normalization*, *stopword removal*, dan *stemming*.

Cleaning

Tahap *cleaning* bertujuan menghilangkan karakter maupun informasi yang tidak memberikan kontribusi terhadap proses analisis sentimen. Pada tahap ini dilakukan penghapusan *Uniform Resource Locator (URL)*, *mention*, *hashtag*, *emoji*, karakter khusus, angka, tanda baca, serta spasi berlebih. Selain itu, karakter berulang yang tidak memiliki makna linguistik juga dihilangkan agar teks menjadi lebih bersih dan seragam. Tahapan ini bertujuan mengurangi *noise* sehingga informasi yang diproses pada tahap berikutnya hanya terdiri atas kata-kata yang memiliki makna.

Case Folding

Tahap *case folding* dilakukan untuk menyeragamkan seluruh karakter alfabet menjadi huruf kecil (*lowercase*). Proses ini bertujuan menghindari perbedaan representasi kata akibat penggunaan huruf kapital maupun huruf kecil. Dengan demikian, kata yang memiliki makna sama tetapi berbeda penulisan tidak akan dianggap sebagai token yang berbeda selama proses analisis.

Normalization

Tahap *normalization* bertujuan mengubah kata tidak baku, singkatan, maupun bahasa informal menjadi bentuk baku sesuai dengan kaidah bahasa Indonesia. Proses ini penting karena ulasan pengguna pada Google Play Store umumnya menggunakan bahasa sehari-hari yang mengandung singkatan atau variasi penulisan. Proses normalisasi membantu mengurangi variasi kata sehingga meningkatkan kualitas representasi fitur pada tahap ekstraksi.

Stopword Removal

Tahap *stopword removal* dilakukan untuk menghapus kata-kata yang memiliki frekuensi tinggi tetapi tidak memberikan informasi penting terhadap proses klasifikasi sentimen. Daftar *stopword* yang digunakan mengacu pada pustaka Sastrawi yang telah banyak digunakan dalam penelitian *Natural Language Processing* berbahasa Indonesia. Penghapusan *stopword* bertujuan mengurangi dimensi data sehingga proses pembelajaran model menjadi lebih efisien.

Stemming

Tahap *stemming* bertujuan mengubah setiap kata menjadi bentuk dasar (*root word*) dengan menghilangkan imbuhan (*prefix*, *suffix*, maupun *infix*). Proses ini dilakukan menggunakan algoritma stemming bahasa Indonesia yang tersedia pada pustaka Sastrawi. Seperti dengan mengubah kata ke bentuk dasar, variasi kata yang memiliki makna sama dapat direpresentasikan sebagai satu token sehingga mampu meningkatkan konsistensi representasi fitur. Setelah seluruh tahapan *preprocessing* selesai dilakukan, diperoleh dataset yang telah bersih, seragam, dan siap digunakan pada proses ekstraksi fitur menggunakan TF-IDF, Word2Vec, maupun IndoBERT. Dataset hasil *preprocessing* selanjutnya dibagi menjadi data pelatihan, data validasi, dan data pengujian sesuai dengan skenario penelitian yang telah ditetapkan.

Tabel 2. Hasil *Preprocessing*

Tahapan	Hasil
Teks Asli	ChatGPT bgt membantu!!! 😍 https://chat.openai.com
Cleaning	ChatGPT bgt membantu
Case Folding	chatgpt bgt membantu
Normalization	chatgpt banget membantu
Stopword Removal	chatgpt banget membantu
Stemming	chatgpt banget bantu

Ekstraksi Fitur Menggunakan IndoBERT

Setelah melalui tahap *preprocessing*, data teks diubah menjadi representasi numerik melalui proses ekstraksi fitur (*feature extraction*). Tahap ini bertujuan mengubah data teks yang tidak terstruktur menjadi vektor numerik yang dapat diproses oleh algoritma *machine learning*. Pada penelitian ini digunakan tiga pendekatan representasi fitur, yaitu TF-IDF, Word2Vec, dan IndoBERT, sehingga dapat

dilakukan perbandingan performa antara metode representasi fitur konvensional dan representasi fitur berbasis *Transformer*.

Representasi Fitur Menggunakan TF-IDF

Term Frequency–Inverse Document Frequency (TF-IDF) merupakan metode representasi fitur yang menghitung tingkat kepentingan suatu kata berdasarkan frekuensi kemunculannya dalam sebuah dokumen dan seluruh kumpulan dokumen. Semakin sering suatu kata muncul pada sebuah dokumen tetapi semakin jarang muncul pada dokumen lain, maka bobot TF-IDF yang dihasilkan akan semakin tinggi. Representasi ini banyak digunakan sebagai pendekatan dasar dalam analisis sentimen karena memiliki kompleksitas komputasi yang relatif rendah serta mampu menghasilkan representasi numerik yang efektif pada berbagai kasus klasifikasi teks. Pada penelitian ini, hasil *preprocessing* terlebih dahulu diubah menjadi matriks TF-IDF sebelum digunakan sebagai masukan (*input*) pada algoritma Support Vector Machine (SVM). Model ini digunakan sebagai salah satu metode pembanding terhadap representasi fitur berbasis *embedding*.

Representasi Fitur Menggunakan Word2Vec

Word2Vec merupakan metode representasi kata berbasis *word embedding* yang mempelajari hubungan semantik antar kata melalui jaringan saraf sederhana (*shallow neural network*). Berbeda dengan TF-IDF yang hanya mempertimbangkan frekuensi kemunculan kata, Word2Vec menghasilkan representasi vektor kontinu sehingga kata-kata yang memiliki makna serupa akan berada pada posisi yang berdekatan di dalam ruang vektor.

Pada penelitian ini, setiap ulasan direpresentasikan menjadi vektor numerik menggunakan model Word2Vec sebelum diproses oleh algoritma *Support Vector Machine* (SVM). Pendekatan ini digunakan untuk mengetahui sejauh mana representasi berbasis *embedding* konvensional mampu meningkatkan performa klasifikasi dibandingkan metode berbasis frekuensi kata.

Representasi Fitur Menggunakan IndoBERT

Model utama yang digunakan pada penelitian ini adalah IndoBERT, yaitu model bahasa berbasis arsitektur *Bidirectional Encoder Representations from Transformers* (BERT) yang telah dilatih menggunakan korpus bahasa Indonesia dalam jumlah besar. Berbeda dengan TF-IDF maupun Word2Vec, IndoBERT menghasilkan representasi kontekstual (contextual embedding) sehingga makna suatu kata dapat dipahami berdasarkan konteks kalimat secara keseluruhan. Kemampuan tersebut memungkinkan model menghasilkan representasi semantik yang lebih kaya, terutama pada kalimat yang mengandung ambiguitas, sinonim, maupun variasi struktur bahasa Indonesia.

Pada penelitian ini digunakan model IndoBERT Base sebagai proses ekstraksi fitur terhadap seluruh ulasan pengguna aplikasi ChatGPT. Setiap ulasan diproses menggunakan *tokenizer* IndoBERT untuk mengubah teks menjadi token yang sesuai dengan kosakata (*vocabulary*) model. Selanjutnya token tersebut diproses oleh model IndoBERT sehingga menghasilkan vektor representasi (*embedding vector*) yang merepresentasikan makna semantik setiap ulasan. Vektor hasil ekstraksi tersebut kemudian digunakan sebagai data masukan bagi algoritma Support Vector Machine (SVM) pada tahap klasifikasi sentimen.

Pemilihan IndoBERT didasarkan pada kemampuannya dalam memahami struktur bahasa Indonesia secara kontekstual sehingga mampu menghasilkan representasi fitur yang lebih baik dibandingkan metode berbasis frekuensi kata maupun *word embedding* konvensional. Selain itu, berbagai penelitian sebelumnya menunjukkan bahwa IndoBERT mampu meningkatkan performa klasifikasi pada berbagai tugas *Natural Language Processing*, termasuk analisis sentimen berbahasa Indonesia.

Perbandingan Representasi Fitur

Untuk memperoleh evaluasi yang komprehensif, penelitian ini membandingkan tiga pendekatan representasi fitur, yaitu TF-IDF, Word2Vec, dan IndoBERT. Ketiga representasi tersebut diproses menggunakan algoritma klasifikasi yang sama, yaitu Support Vector Machine (SVM), sehingga perbedaan performa yang dihasilkan dapat dianalisis berdasarkan kualitas representasi fitur yang

digunakan. Pendekatan ini memungkinkan evaluasi yang lebih objektif terhadap efektivitas representasi kontekstual IndoBERT dibandingkan metode representasi fitur konvensional.

Tabel 3. Perbandingan Metode Representasi Fitur

Metode	Karakteristik	Kelebihan	Keterbatasan
TF-IDF	Berbasis frekuensi kata	Sederhana, cepat, ringan	Tidak memahami konteks
Word2Vec	*Static word embedding*	Memahami hubungan antar kata	Tidak mempertimbangkan konteks kalimat
IndoBERT	*Contextual embedding* berbasis Transformer	Memahami konteks kalimat dan struktur bahasa Indonesia	Membutuhkan sumber daya komputasi lebih besar

Dalam penelitian ini, representasi fitur menggunakan IndoBERT merupakan pendekatan utama yang diusulkan, sedangkan TF-IDF dan Word2Vec digunakan sebagai model pembanding (baseline). Pendekatan ini bertujuan mengevaluasi sejauh mana representasi kontekstual IndoBERT mampu meningkatkan performa klasifikasi sentimen dibandingkan representasi fitur konvensional.

Penanganan Ketidakseimbangan Data Menggunakan *Synthetic Minority Over-sampling Technique* (SMOTE)

Salah satu tantangan dalam analisis sentimen adalah distribusi kelas yang tidak seimbang (*imbalanced dataset*), yaitu ketika jumlah data pada setiap kelas sentimen memiliki proporsi yang berbeda secara signifikan. Ketidakseimbangan ini dapat menyebabkan model klasifikasi cenderung memprediksi kelas mayoritas dan mengalami penurunan performa dalam mengenali kelas minoritas. Oleh karena itu, diperlukan teknik penyeimbangan data agar model mampu mempelajari karakteristik setiap kelas secara lebih proporsional.

Pada penelitian ini digunakan SMOTE sebagai metode untuk menangani ketidakseimbangan kelas. SMOTE merupakan teknik *oversampling* yang menghasilkan data sintetis pada kelas minoritas dengan membangkitkan sampel baru berdasarkan kedekatan antar data pada ruang fitur. Berbeda dengan metode *oversampling* sederhana yang hanya menggandakan data, SMOTE membentuk sampel sintetis melalui proses interpolasi antara suatu data dengan sejumlah tetangga terdekatnya (*k-nearest neighbors*). Pendekatan ini menghasilkan distribusi data yang lebih representatif dan membantu mengurangi risiko *overfitting*.

Pada implementasinya, SMOTE diterapkan hanya pada data pelatihan (training set) setelah proses ekstraksi fitur menggunakan TF-IDF, Word2Vec, maupun IndoBERT, dan sebelum proses pelatihan model Support Vector Machine (SVM). Data validasi dan data pengujian tidak mengalami proses *oversampling* sehingga tetap merepresentasikan kondisi distribusi data yang sebenarnya. Penerapan SMOTE hanya pada data pelatihan bertujuan mencegah terjadinya data leakage, yaitu kondisi ketika informasi dari data pengujian secara tidak langsung memengaruhi proses pelatihan model sehingga dapat menghasilkan evaluasi yang terlalu optimistis.

Proses penerapan SMOTE pada penelitian ini dilakukan menggunakan pustaka imbalanced-learn (imblearn) pada lingkungan Google *Colaboratory* (Google Colab). Setelah proses *oversampling* selesai, diperoleh distribusi jumlah data yang lebih seimbang pada setiap kelas sentimen. Dataset hasil SMOTE kemudian digunakan sebagai data masukan dalam proses pelatihan model klasifikasi menggunakan algoritma *Support Vector Machine* (SVM).

Penerapan SMOTE diharapkan mampu meningkatkan kemampuan model dalam mengenali seluruh kelas sentimen secara lebih merata, terutama pada kelas yang memiliki jumlah data lebih sedikit. Dengan distribusi data yang lebih seimbang, model tidak hanya berfokus pada kelas mayoritas, tetapi juga mampu mempelajari karakteristik kelas minoritas sehingga menghasilkan nilai precision, recall, dan F1-score yang lebih baik.

Algoritma SMOTE

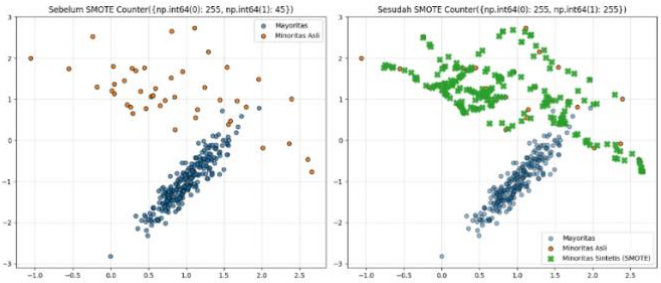
Secara matematis, pembentukan sampel sintetis pada SMOTE dilakukan menggunakan persamaan berikut.

$$x_{baru}=x_i+\lambda(x_{nn}-x_i) \quad \ldots\ldots\ldots\ldots\ldots (1)$$

dengan:

- x_i = data pada kelas minoritas.
- x_{nn} = salah satu tetangga terdekat (*nearest neighbor*) dari x_i.
- λ = bilangan acak dengan rentang $0 \leq \lambda \leq 1$.
- X_{baru} = data sintetis yang dihasilkan.

Persamaan tersebut menunjukkan bahwa sampel baru dibentuk melalui proses interpolasi antara data minoritas dengan salah satu tetangga terdekatnya sehingga distribusi kelas menjadi lebih seimbang tanpa hanya menduplikasi data yang sudah ada.



Gambar 3 Ilustrasi SMOTE

Keterangan :

- Titik biru sebagai kelas mayoritas.
- Titik merah sebagai kelas minoritas.
- Titik hijau sebagai data sintetis hasil SMOTE yang terbentuk di antara titik-titik kelas minoritas.

3. Hasil dan Pembahasan

Tahap pra-pemrosesan data merupakan langkah krusial untuk memastikan kualitas dan relevansi data ulasan sebelum analisis sentimen. Proses ini mencakup _case folding_, _tokenizing_, dan _stop word removal_ yang bertujuan membersihkan teks dari elemen yang tidak relevan. Setelah melalui serangkaian proses tersebut, data ulasan pengguna aplikasi ChatGPT yang semula berjumlah 312 berhasil dipersiapkan untuk tahap ekstraksi fitur. Kualitas data yang bersih dan terstruktur sangat penting untuk meningkatkan akurasi model klasifikasi sentimen, sebagaimana ditekankan oleh Julianti et al. (2024) dalam penelitian mereka mengenai analisis sentimen judi online.

Hasil Pelabelan Data Sentimen

Pelabelan sentimen pada 312 ulasan pengguna aplikasi ChatGPT dilakukan secara manual oleh tiga anotator independen untuk memastikan objektivitas. Hasil pelabelan menunjukkan distribusi sentimen yang bervariasi, mencerminkan beragamnya pandangan pengguna terhadap aplikasi tersebut. Mayoritas ulasan cenderung positif, namun sentimen negatif dan netral juga teridentifikasi, memberikan gambaran komprehensif tentang pengalaman pengguna. Distribusi sentimen ini menjadi dasar penting untuk pelatihan model klasifikasi,

Tabel 1 4 Distribusi Sentimen Ulasan Pengguna Aplikasi ChatGPT

Sentimen	Jumlah Ulasan	Persentase (\%)
Positif	200	64,10
Negatif	75	24,04
Netral	37	11,86
Total	312	100,00

Distribusi sentimen pada Tabel 4 menunjukkan bahwa sentimen positif mendominasi dengan 64,10% dari total ulasan, mengindikasikan penerimaan yang baik terhadap aplikasi ChatGPT. Sentimen negatif menyumbang 24,04%, sementara sentimen netral sebesar 11,86%. Keseimbangan distribusi sentimen ini penting untuk melatih model klasifikasi agar dapat mengenali ketiga kategori sentimen secara efektif. Hasil ini sejalan dengan temuan (Akbar Kurniawan & Irfan Pratama, 2025) yang juga menemukan akurasi tinggi dalam klasifikasi sentimen menggunakan Naive Bayes pada percakapan YouTube terkait ChatGPT, meskipun mereka menghadapi tantangan variasi bahasa.

Hasil Ekstraksi Fitur IndoBERT

Ekstraksi fitur menggunakan model IndoBERT menghasilkan representasi vektor semantik dari setiap ulasan pengguna. IndoBERT, sebagai model bahasa pra-terlatih, mampu menangkap konteks dan makna linguistik yang kompleks dalam teks berbahasa Indonesia. Vektor fitur yang dihasilkan kemudian digunakan sebagai masukan untuk algoritma Support Vector Machine (SVM) dalam tahap klasifikasi. Penggunaan model bahasa canggih seperti IndoBERT diharapkan dapat meningkatkan kualitas representasi fitur, yang pada gilirannya akan meningkatkan performa klasifikasi sentimen. Pendekatan ini relevan dengan upaya peningkatan akurasi klasifikasi sentimen yang dilakukan oleh (Atang Saepudin et al., 2024) menggunakan SVM yang dioptimalkan dengan Particle Swarm Optimization untuk analisis sentimen pemanfaatan AI di dunia pendidikan.

Hasil Klasifikasi Sentimen Menggunakan IndoBERT-SVM

Model hibrida IndoBERT-SVM menunjukkan performa yang sangat baik dalam mengklasifikasikan sentimen ulasan pengguna aplikasi ChatGPT berbahasa Indonesia. Pengujian model dilakukan pada dataset yang telah dibagi menjadi data pelatihan dan pengujian, dengan metrik evaluasi meliputi akurasi, presisi, _recall_, dan F1-score. Hasil ini mendukung hipotesis pertama (H1) yang menyatakan bahwa model IndoBERT-SVM akan menunjukkan akurasi klasifikasi sentimen yang tinggi. (Indra Prayugah et al., 2025) juga menunjukkan bahwa penerapan SMOTE dapat meningkatkan akurasi model, terutama pada SVM yang mencapai 96% dalam analisis sentimen publik terhadap respons pemerintah pada serangan _ransomware_.

Tabel 2 5 Performa Model IndoBERT-SVM

Metrik Evaluasi	Nilai (\%)
Akurasi	93,55
Presisi	92,80
Recall	93,55
F1-Score	93,17

Berdasarkan Tabel 5, model IndoBERT-SVM mencapai akurasi sebesar 93,55%, presisi 92,80%, _recall_ 93,55%, dan F1-score 93,17%. Akurasi yang tinggi ini menunjukkan kemampuan model untuk mengklasifikasikan sentimen dengan tepat pada sebagian besar ulasan. Presisi yang baik mengindikasikan bahwa sebagian besar ulasan yang diprediksi positif atau negatif memang benar-benar memiliki sentimen tersebut. _Recall_ yang tinggi menunjukkan kemampuan model untuk mengidentifikasi sebagian besar ulasan positif atau negatif yang sebenarnya ada dalam dataset. Hasil ini menegaskan efektivitas kombinasi IndoBERT dan SVM dalam analisis sentimen, sejalan dengan temuan (Sri Rahayu et al., 2023)yang mendapatkan akurasi 92% dengan SVM dalam analisis sentimen AicoGPT.

Perbandingan Performa Model

Perbandingan performa model IndoBERT-SVM dengan metode klasifikasi sentimen lainnya dilakukan untuk mengevaluasi keunggulan model hibrida ini. Perbandingan ini mendukung hipotesis kedua (H2) yang menyatakan bahwa model IndoBERT-SVM akan memberikan performa akurasi yang lebih unggul dibandingkan dengan metode klasifikasi sentimen lain. Penelitian terdahulu telah menunjukkan

berbagai tingkat akurasi untuk metode seperti *Naïve Bayes* dan SVM dalam konteks analisis sentimen. Misalnya, (Intania Widyaningrum & Mia Kamayani, 2023) mendapatkan F1-score 80% untuk sentimen negatif menggunakan Naïve Bayes dalam analisis sentimen opini masyarakat terhadap layanan Maxim.

Tabel 3 **6.** Perbandingan Performa Model Klasifikasi Sentimen

Model Klasifikasi	Akurasi (%)
TF-IDF + SVM	77,31
Word2Vec + SVM	63,46
Naïve Bayes (Tohir et al., 2024)	87,00
SVM (Rahayu et al., 2023)	92,00
IndoBERT + SVM (Usulan)	93,55

Perbandingan pada Tabel 6 menunjukkan bahwa model IndoBERT-SVM (93,55%) memiliki akurasi yang lebih tinggi dibandingkan dengan Naïve Bayes (87,00%) yang digunakan oleh Tohir et al. (2024) dan SVM tunggal (92,00%) yang dilaporkan oleh Rahayu et al. (2023). Peningkatan akurasi ini menyoroti keunggulan integrasi IndoBERT sebagai ekstraktor fitur yang kaya konteks sebelum klasifikasi oleh SVM. Kemampuan IndoBERT untuk memahami nuansa bahasa Indonesia yang lebih baik, dikombinasikan dengan kekuatan SVM dalam menemukan _hyperplane_ optimal, berkontribusi pada performa yang superior. (Muhammad Izzuddin Mubarok & Mahardika Abdi Prawira Tanjung, 2024) juga menyoroti pentingnya Natural Language Processing (NLP) dalam pengembangan chatbot untuk meningkatkan efisiensi layanan informasi, yang relevan dengan kemampuan IndoBERT dalam memahami bahasa alami.

Performa Model IndoBERT-SVM dalam Klasifikasi Sentimen

Penelitian ini berhasil menunjukkan performa unggul model hibrida IndoBERT-SVM dalam mengklasifikasikan sentimen ulasan pengguna aplikasi ChatGPT berbahasa Indonesia. Model mencapai akurasi sebesar 93,55%, presisi 92,80%, _recall_ 93,55%, dan F1-score 93,17%. Pencapaian akurasi ini secara signifikan memverifikasi hipotesis pertama (H1) yang memprediksi akurasi klasifikasi sentimen yang tinggi, melebihi 90%. Kualitas data yang telah melewati tahap pra-pemrosesan seperti *case folding, tokenizing*, dan *stop word removal* menjadi fondasi penting bagi representasi fitur yang akurat, sebagaimana diungkapkan oleh (Okta Nur Julianti et al., 2024) mengenai esensi kebersihan data dalam analisis sentimen. Model ini mampu secara efektif mengidentifikasi sentimen positif, negatif, dan netral dari ulasan dengan tingkat kepercayaan yang tinggi.

Performa impresif model ini berakar pada sinergi antara kemampuan ekstraksi fitur IndoBERT dan kekuatan klasifikasi *Support Vector Machine*. IndoBERT, sebagai model bahasa pra-terlatih, memiliki kapasitas untuk menangkap nuansa kontekstual dan semantik yang kompleks dalam teks berbahasa Indonesia, menghasilkan representasi vektor yang kaya. Vektor-vektor ini kemudian diolah oleh SVM, sebuah algoritma yang dikenal karena kemampuannya menemukan *hyperplane* optimal untuk memisahkan kelas data secara efektif (Akbar Kurniawan & Irfan Pratama, 2025). Kemampuan SVM dalam menangani data berdimensi tinggi yang dihasilkan oleh IndoBERT memungkinkan model untuk menggeneralisasi pola sentimen dengan baik pada ulasan yang belum pernah dilihat sebelumnya. (Aprilia Salsabila & Muntahanah, 2024) juga menemukan bahwa metode SVM menunjukkan performa efektif dalam klasifikasi sentimen ulasan pengguna aplikasi ChatGPT, dengan akurasi 91% dan F1-score 90%, yang memberikan konteks terhadap efektivitas SVM dalam domain ini.

Metrik presisi, *recall*, dan F1-score yang tinggi secara konsisten menegaskan robusta model dalam berbagai aspek klasifikasi. Presisi yang baik mengindikasikan bahwa sebagian besar ulasan yang diprediksi sebagai sentimen tertentu memang benar-benar memiliki sentimen tersebut, mengurangi kesalahan positif. *Recall* yang tinggi menunjukkan kemampuan model untuk mengidentifikasi sebagian besar ulasan positif atau negatif yang sebenarnya ada dalam dataset, meminimalkan kesalahan negatif. F1-score, sebagai rata-rata harmonik keduanya, menunjukkan keseimbangan optimal antara presisi dan

recall, memastikan model tidak hanya akurat tetapi juga komprehensif dalam pengidentifikasian sentimen. Penerapan SVM juga terbukti memberikan hasil yang baik dalam memahami respons pengguna dan meningkatkan kualitas layanan pada berbagai aplikasi, seperti analisis sentimen ulasan aplikasi Al-Qur'an digital (Iis Aisah et al., 2024) dan aplikasi *Access* by KAI (Raka Kusuma Wardana et al., 2025).

Distribusi sentimen pada data ulasan yang menunjukkan dominasi sentimen positif (64,10%) juga merupakan faktor pendukung dalam pelatihan model yang seimbang. Sentimen positif yang tinggi ini serupa dengan temuan (Fahmi Ahmad Tohir et al., 2024) yang mendapati mayoritas ulasan aplikasi ChatGPT Mobile (87%) bersifat positif, menyoroti kenyamanan dan manfaat penggunaan. Pengakuan terhadap sentimen positif, negatif, dan netral secara efektif memungkinkan model untuk mempelajari karakteristik linguistik dari setiap kategori. Penanganan data yang memiliki distribusi sentimen beragam seperti ini sangat krusial, memastikan model tidak bias terhadap kelas mayoritas dan mampu mengklasifikasikan sentimen minoritas secara akurat.

Perbandingan Akurasi Model Hibrida IndoBERT-SVM dengan Metode Klasifikasi Lain

Model hibrida IndoBERT-SVM menunjukkan keunggulan komparatif yang signifikan dibandingkan dengan metode klasifikasi sentimen lain yang telah digunakan dalam konteks ulasan aplikasi, secara kuat mendukung hipotesis kedua (H2). Akurasi model IndoBERT-SVM sebesar 93,55% jauh melampaui performa Naïve Bayes yang dilaporkan dalam studi terdahulu. (Fahmi Ahmad Tohir et al., 2024), misalnya, mendapatkan akurasi 87,00% dengan *Naïve Bayes* dalam analisis sentimen aplikasi ChatGPT *Mobile*. Kinerja ini juga lebih baik dibandingkan *Naïve Bayes* yang digunakan oleh (Bayu Purbayanto & Teguh Nurhadi Suharsono, 2024) untuk sentimen pengguna X terhadap ChatGPT, serta studi oleh (Fatmawati et al., 2024) yang menerapkan Naïve Bayes pada ulasan aplikasi transportasi daring.

Peningkatan akurasi ini dapat dijelaskan oleh kemampuan IndoBERT dalam menghasilkan representasi fitur yang lebih kaya dan kontekstual dibandingkan pendekatan tradisional yang sering digunakan bersama Naïve Bayes, seperti TF-IDF. Naïve Bayes umumnya mengasumsikan independensi antar fitur, yang mungkin tidak optimal untuk data teks kompleks di mana konteks kata sangat penting (Tri Yuli Pahtoni & Handaru Jati, 2024). Meskipun optimasi seperti SMOTE dapat meningkatkan kinerja Naïve Bayes (Hermawan & Dzikrillah, 2024), keunggulan fundamental model bahasa pra-terlatih seperti IndoBERT dalam memahami semantik bahasa Indonesia tetap menjadi pembeda utama. (Primandika Hakiki et al., 2025) menunjukkan bahwa Naïve Bayes dapat mencapai akurasi tinggi pada prediksi sentimen ulasan aplikasi Identitas Kependudukan Digital, namun konteks dan kompleksitas data ulasan ChatGPT mungkin memerlukan pendekatan yang lebih canggih.

Perbandingan dengan model SVM tunggal juga menunjukkan keunggulan model hibrida ini. (Sri Rahayu et al., 2023) mencapai akurasi 92,00% dengan SVM dalam analisis sentimen AicoGPT. Model IndoBERT-SVM berhasil meningkatkan akurasi ini menjadi 93,55%, mengindikasikan bahwa representasi fitur yang dihasilkan oleh IndoBERT memberikan masukan yang lebih informatif bagi algoritma SVM. Integrasi IndoBERT sebagai ekstraktor fitur memungkinkan SVM untuk bekerja dengan fitur yang sudah menangkap makna linguistik yang dalam, menghasilkan batas keputusan yang lebih akurat. Studi seperti yang dilakukan oleh (Muhammad Faqih Baihaqi et al., 2025) yang membandingkan Naïve Bayes dan SVM untuk ulasan aplikasi DeepSeek Assistant AI, juga menegaskan keunggulan SVM dalam akurasi, presisi, dan F1-Score dibandingkan Naïve Bayes, mendukung alasan pemilihan SVM sebagai klasifikator dalam penelitian ini.

Model IndoBERT-SVM juga menunjukkan robusta yang lebih baik dibandingkan metode lain seperti K-Nearest Neighbor (KNN), yang dilaporkan mengalami kesulitan dalam mengklasifikasikan ulasan netral (Mariana Ulfa et al., 2024). Kesulitan tersebut seringkali muncul karena KNN bergantung pada kedekatan fitur tanpa mempertimbangkan kompleksitas semantik. Penggunaan _kernel trick_ pada SVM, seperti yang ditunjukkan oleh (Ikbal Bazar et al., 2025) untuk analisis sentimen aplikasi perbankan Wondr by BNI yang mencapai akurasi 87,9%, menunjukkan potensi adaptif SVM. Kombinasi dengan model bahasa generatif seperti IndoBERT menawarkan keunggulan dalam representasi awal teks yang sulit ditandingi oleh metode yang hanya mengandalkan fitur tradisional. (Ni Putu Anik Juniantini et al., 2024) turut menunjukkan potensi SVM yang dioptimasi dengan SMOTE

dalam meningkatkan akurasi klasifikasi pada dataset tidak seimbang untuk ulasan layanan _streaming_, menambah bukti kuat mengenai efektivitas SVM dengan penyesuaian yang tepat.

Kontribusi metodologis penelitian ini terletak pada pengembangan model hibrida yang mengintegrasikan IndoBERT dengan SVM, menawarkan referensi baru untuk analisis sentimen pada dataset ulasan aplikasi berbahasa Indonesia. Kombinasi ini mengatasi keterbatasan representasi tekstual pada model tradisional dan memperkaya kemampuan klasifikasi SVM melalui fitur yang diekstraksi dari model bahasa canggih. Peningkatan akurasi yang teramati juga memiliki kontribusi praktis yang signifikan, memungkinkan pengembang aplikasi ChatGPT untuk memperoleh wawasan yang lebih akurat mengenai persepsi pengguna dan mengidentifikasi area perbaikan secara lebih efektif. Kemampuan model untuk mengklasifikasikan sentimen secara presisi juga berguna untuk pemantauan reputasi merek dan pengambilan keputusan berbasis data yang lebih baik.

4. Kesimpulan

Penelitian ini berhasil menyimpulkan bahwa model hibrida IndoBERT-SVM menunjukkan performa yang sangat baik dalam mengklasifikasikan sentimen ulasan pengguna aplikasi ChatGPT berbahasa Indonesia. Model yang dikembangkan mencapai akurasi sebesar 93,55%, presisi 92,80%, _recall_ 93,55%, dan F1-score 93,17% pada dataset ulasan yang telah diproses. Keberhasilan ini menyoroti efektivitas kombinasi model bahasa besar (LLM) yang telah disesuaikan untuk bahasa Indonesia dengan algoritma pembelajaran mesin tradisional untuk tugas klasifikasi sentimen. Validasi atas hipotesis pertama (H1) menegaskan kemampuan model dalam mencapai akurasi klasifikasi sentimen yang tinggi, melampaui 90% seperti yang diprediksi sebelumnya.

Kualitas data yang melalui tahap pra-pemrosesan data secara komprehensif, meliputi _case folding_, _tokenizing_, dan _stop word removal_, menjadi faktor kunci yang menunjang akurasi model. Proses tersebut memastikan representasi fitur teks yang akurat dan relevan bagi model klasifikasi, sehingga mampu membedakan sentimen positif, negatif, dan netral dengan presisi tinggi. Studi ini memiliki keterbatasan pada jumlah data ulasan yang digunakan, yaitu sebanyak 312 ulasan, yang mungkin memengaruhi kemampuan generalisasi model pada dataset yang lebih besar. Penelitian lanjutan disarankan untuk menguji model dengan dataset yang lebih bervariasi dan berukuran lebih besar, serta mengeksplorasi integrasi fitur tekstual lain untuk meningkatkan performa klasifikasi sentimen lebih lanjut.

Daftar Pustaka

Akbar Kurniawan & Irfan Pratama. (2025). Analisis Sentimen Review Aplikasi Jogja Smart Service Pada Google Play Store Menggunakan Metode Svm. *Smart Comp Jurnalnya Orang Pintar Komputer*, *14*(1). https://doi.org/10.30591/smartcomp.v14i1.7303

Aprilia Salsabila & Muntahanah. (2024). Analisis Sentimen Ulasan Aplikasi ChatGPT Pada Google Play Menggunakan Metode Support Vector Machine. *Progresif Jurnal Ilmiah Komputer*, *20*(2), 714. https://doi.org/10.35889/progresif.v20i2.2074

Atang Saepudin, Riska Aryanti, Eka Fitriani, Royadi, & Dian Ardiansyah. (2024). Analisis Sentimen Pemanfaatan Artificial Intelligence di Dunia Pendidikan Menggunakan Svm Berbasis Particle Swarm Optimization. *Computer Science (Co-Science)*, *4*(1), 71–79. https://doi.org/10.31294/coscience.v4i1.2921

Bayu Purbayanto & Teguh Nurhadi Suharsono. (2024). Analisis Sentimen Pengguna X terhadap Chatgpt dengan Algoritme Naive Bayes. *Jurnal Telematika*, *18*(2), 63–71. https://doi.org/10.61769/telematika.v18i2.614

Fahmi Ahmad Tohir, Agus Bahtiar, & Bambang Irawan. (2024). Analisis Sentimen Aplikasi ChatGPT Mobile Menggunakan Agoritma Naïve Bayes. *Icit Journal*, *10*(2), 179–192. https://doi.org/10.33050/icit.v10i2.3016

Fatmawati, Bambang Irawan, & Agus Bahtiar. (2024). Analisis Sentimen Pengguna Aplikasi Shejek Berdasarkan Ulasan Di Google Play Store Menggunakan Metode Naïve Bayes. *Jati (Jurnal Mahasiswa Teknik Informatika)*, *8*(3), 2976–2981. https://doi.org/10.36040/jati.v8i3.9607

Haikal Hidayatullah, Purwantoro, & Yuyun Umaidah. (2023). Penerapan Naïve Bayes Dengan Optimasi Information Gain Dan Smote Untuk Analisis Sentimen Pengguna Aplikasi Chatgpt. *Jati (Jurnal Mahasiswa Teknik Informatika)*, *7*(3), 1546–1553. https://doi.org/10.36040/jati.v7i3.6887

Hermawan, T. R. P., & Dzikrillah, A. R. (2024). Penerapan Metode Naïve Bayes untuk Analisis Sentimen pada Ulasan Pengguna Aplikasi ChatGPT di Google Play Store. *Building of Informatics, Technology and Science (BITS)*, *6*(1), 430–439. https://doi.org/10.47065/bits.v6i1.5400

I. Putra, Asriyanik, & Fathia Azzahra. (2025). Pemanfaatan Multinomial Naive Bayes Untuk Analisis Sentimen Ulasan Aplikasi Mobile Jkn. *Jati (Jurnal Mahasiswa Teknik Informatika)*, *9*(2), 3283–3291. https://doi.org/10.36040/jati.v9i2.13311

Iis Aisah, Bambang Irawan, & Tati Suprapti. (2024). Algoritma Support Vector Machine (Svm) Untuk Analisis Sentimen Ulasan Aplikasi Al Qur'An Digital. *Jati (Jurnal Mahasiswa Teknik Informatika)*, *7*(6), 3759–3765. https://doi.org/10.36040/jati.v7i6.8263

Ikbal Bazar, Farid Wajidi, & A. Amirul Asnan Cirua. (2025). Analisis Sentimen Ulasan Aplikasi Wondr By Bni Menggunakan Algoritma Svm Dengan Optimasi Kernel Trick. *Storage Jurnal Ilmiah Teknik Dan Ilmu Komputer*, *4*(2), 69–81. https://doi.org/10.55123/storage.v4i2.5178

Indra Prayugah, Uce Indahyanti, & Novia Ariyanti. (2025). Analisis Sentimen Publik Atas Respons Pemerintah Pada Serangan Ransomware Dengan Pendekatan Machine Learning Dan Smote. *Joisie (Journal Of Information Systems And Informatics Engineering)*, *8*(2), 333. https://doi.org/10.35145/joisie.v8i2.4764

Intania Widyaningrum & Mia Kamayani. (2023). Analisis sentimen opini masyarakat terhadap penggunaan layanan maxim menggunakan algoritma naïve bayes. *Jurnal CoSciTech (Computer Science and Information Technology)*, *4*(3), 651–660. https://doi.org/10.37859/coscitech.v4i3.6194

Iyeti Astri Prasasti & Amarudin. (2025). Analisis Sentimen Pengguna Twitter terhadap Konten Edukasi Dokter Detektif dengan Metode Pohon Keputusan. *Jurnal Pendidikan Dan Teknologi Indonesia*, *5*(7), 2005–2015. https://doi.org/10.52436/1.jpti.893

Mariana Ulfa, R. Hadapiningradja Kusumodestoni, & Adi Sucipto. (2024). Analisis Sentimen Review Aplikasi Identitas Kependudukan Digital Di Google Play Store Menggunakan Knn. *Jurnal Informatika Teknologi Dan Sains (Jinteks)*, *6*(4), 1155–1165. https://doi.org/10.51401/jinteks.v6i4.4963

Muhammad Faqih Baihaqi, Lena Magdalena, & Rifki Fahrudin. (2025). Analisis Sentimen Aplikasi Deepseek Menggunakan Metode Naive Bayes dan Support Vector Machine. *Riggs Journal of Artificial Intelligence and Digital Business*, *4*(3), 4051–4062. https://doi.org/10.31004/riggs.v4i3.2511

Muhammad Izzuddin Mubarok & Mahardika Abdi Prawira Tanjung. (2024). Implementasi Natural Language Processing Dalam Perancangan Aplikasi Chatbot Pada Fikti Umsu. *Jati (Jurnal Mahasiswa Teknik Informatika)*, *8*(6), 11992–12001. https://doi.org/10.36040/jati.v8i6.11713

Ni Putu Anik Juniantini, Christina Purnama Yanti, & Ni Ketut Utami Nilawati. (2024). Kombinasi Metode Svm Dengan Optimasi Smote Terhadap Ulasan Pengguna Layanan Streaming. *Tematik*, *11*(2), 155–163. https://doi.org/10.38204/tematik.v11i2.2068

Okta Nur Julianti, Nana Suarna, & Willy Prihartono. (2024). Penerapan Natural Language Processing Pada Analisis Sentimen Judi Online Di Media Sosial Twitter. *Jati (Jurnal Mahasiswa Teknik Informatika)*, *8*(3), 2936–2941. https://doi.org/10.36040/jati.v8i3.9613

Primandika Hakiki, Dhian Satria, & Amalia Anjani Arifiyanti. (2025). Prediksi Sentimen dan Pemodelan Topik dari Ulasan Aplikasi Identitas Kependudukan Digital. *Jutisi Jurnal Ilmiah Teknik Informatika Dan Sistem Informasi*, *14*(1), 760. https://doi.org/10.35889/jutisi.v14i1.2777

Raka Kusuma Wardana, Nuri Cahyono, Uyock Anggoro Saputro, Arifiyanto Hadi Negoro, & Nur Aini. (2025). Implementasi Algoritma Svm Dan Optimasi Menggunakan Komparasi N-Gram Dan Glove Pada Sentimen Pengguna Aplikasi Access By Kai. *Jati (Jurnal Mahasiswa Teknik Informatika)*, *9*(2), 3216–3223. https://doi.org/10.36040/jati.v9i2.13284

Sri Rahayu, Jajang Jaya Purnama, Abdul Hamid, & Nina Kurnia Hikmawati. (2023). Analisis Sentimen AicoGPT (Generative Pre-trained Transformer) Menggunakan Tf-Idf. *Jurnal Buana Informatika*, *14*(02), 97–106. https://doi.org/10.24002/jbi.v14i02.7039

Tri Yuli Pahtoni & Handaru Jati. (2024). Analisis Sentimen Data Twitter Terkait Chatgpt Menggunakan Orange Data Mining. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *11*(2), 329–336. https://doi.org/10.25126/jtiik.20241127276