ANALISIS SENTIMEN PENGGUNAAN APLIKASI *CHATGPT* PADA LINGKUNGAN AKADEMIK MENGGUNAKAN ALGORITMA *NAÏVE BAYES*

Izmi Panji Gumilang[1], Nurul Mega Saraswati[2], Fathulloh[3]
[1,2,3]Program Studi Informatika, Fakultas Sains dan Teknologi, Universitas Peradaban
gilanxjr13@gmail.com, nurul.mega.s@gmail.com, fathul.peradaban@gmail.com
Jl.Raya Pagojengan Km 3 Paguyangan, Brebes, Jawa Tengah, 52276

Keywords:

ChatGPT, Naïve Bayes, sentiment analysis, text mining.

Abstract

Generative Pre-trained Transformer (ChatGPT) is an artificial intelligence technology developed by OpenAI, designed to understand natural language and interact with humans contextually. Its high popularity has brought diverse experiences and opinions from users, particularly regarding its use in academic environments. Sentiment analysis research is necessary to understand user perceptions of the ChatGPT application. User review data were collected from the Google Play Store through web scraping, totaling 5,000 reviews from July 2023 to July 2025. The Naïve Bayes algorithm was used to classify the review texts using a Fine-Grained Sentiment Analysis approach into three sentiment categories: positive (72.3%), negative (18.5%), and neutral (9.2%). Testing using the Confusion Matrix with an 80% training and 20% testing data split showed that the model achieved an average accuracy of 81%, precision of 82%, recall of 81%, and F1-score of 81%. The results of this study are expected to provide insights into the extent to which ChatGPT is considered helpful or effective, especially in contexts related to academic environments.

Kata Kunci:

ChatGPT, Naïve Bayes, analisis sentimen*, Text Mining.*

Abstrak

Generative Pre-trained Transformer (ChatGPT) adalah teknologi kecerdasan buatan yang dikembangkan oleh *OpenAI* yang dirancang untuk memahami bahasa alami dan mampu berinteraksi dengan manusia secara kontekstual. *ChatGPT* memiliki popularitas yang tinggi membawa beragam pengalaman serta opini dari pengguna yang ada pada ulasan aplikasi *ChatGPT*, khusunya penggunaanya pada lingkungan akademik. Penelitian analisis sentimen diperlukan untuk mengetahui persepsi pengguna terhadap aplikasi *ChatGPT*. Data ulasan yang diambil dari *Google Play Store* melalui proses *Scraping* dengan jumlah total 5.000 data ulasan dari bulan Juli 2023 sampai Juli 2025. Algoritma *Naïve Bayes* digunakan untuk mengklasifikasikan data teks ulasan dengan pendekatan analisis *Fine-Grained Sentiment Analysis* ke dalam tiga kategori sentimen positif 72.3%, negatif 18.5% dan netral 9.2%. Hasil uji menggunakan *Confusion Matrix* dengan perbandingan 80% data latih dan 20% data uji menunjukan bahwa model *Confusion Matrix* menghasilkan rata-rata nilai *Accuracy* sebesar 81%, *Precission* 82%, *Recall* 81% dan *F1-Score* 81%. Hasil penelitian ini diharapkan memberikan gambaran sejauh mana *ChatGPT* dianggap membantu atau efektif digunakan terutama pada konteks yang berkaitan dengan lingkungan akademik.

I. Pendahuluan

Perkembangan teknologi informasi yang pesat memberikan dampak signifikan pada berbagai aspek kehidupan, termasuk pendidikan sebagai sektor strategis dalam pembangunan sumber daya

manusia [1]. Di era digital, akses informasi semakin luas dan didukung oleh inovasi berbasis kecerdasan buatan (*Artificial Intelligence*/AI), salah satunya *Generative Pre-trained Transformer* (ChatGPT) [2]. ChatGPT dikembangkan oleh OpenAI sebagai teknologi AI yang mampu memahami bahasa alami dan berinteraksi melalui teks, sehingga berpotensi besar mendukung proses pembelajaran seperti memahami materi, menerjemahkan teks, dan menyusun ringkasan [3]. Namun, dalam lingkungan akademik, pemanfaatannya juga menimbulkan kekhawatiran terkait keakuratan informasi, keterbatasan konteks, potensi penurunan kemampuan berpikir kritis, serta isu plagiarisme yang dapat memengaruhi persepsi dan efektivitas penggunaannya [4].

Analisis sentimen merupakan metode untuk mengidentifikasi opini publik terhadap suatu topik melalui data teks tidak terstruktur dengan tujuan mengklasifikasikan sentimen ke dalam kategori positif, negatif, dan netral menggunakan *Fine-Grained Sentiment Analysis* [5]. Melalui analisis sentimen terhadap ulasan pengguna, khususnya di *Google Play Store*, dapat diperoleh gambaran objektif mengenai persepsi pengguna terhadap aplikasi ChatGPT dalam lingkungan akademik. Ulasan negatif berpotensi menurunkan tingkat kepercayaan dan menghambat pemanfaatan ChatGPT, sedangkan ulasan positif dapat meningkatkan citra serta kepercayaan terhadap manfaat dan keandalannya dalam mendukung kegiatan akademik.

Perkembangan *Machine Learning* dan *Natural Language Processing* (NLP) membuka peluang besar dalam pengolahan data teks, termasuk analisis sentimen, dengan salah satu algoritma yang umum digunakan adalah Naïve Bayes [6]. Algoritma ini dipilih karena sederhana, efisien, dan memiliki kinerja yang baik dalam mengklasifikasikan data teks tidak terstruktur. Penelitian sebelumnya menunjukkan bahwa berbagai algoritma seperti KNN, Random Forest, dan Decision Tree menghasilkan tingkat akurasi yang bervariasi [7][8][9]. Berdasarkan keunggulan tersebut, Naïve Bayes dinilai relevan untuk menganalisis sentimen opini publik terhadap aplikasi ChatGPT di lingkungan akademik. Hasil analisis diharapkan mampu memberikan gambaran dominasi sentimen pengguna serta menjadi dasar pemanfaatan teknologi AI secara adaptif dan bertanggung jawab dalam dunia pendidikan..

II. Landasan Teori

1. *Text Mining*

Text mining adalah proses pengolahan data yang bertujuan untuk mengekstraksi informasi bermakna dari data berbentuk teks, umumnya bersumber dari dokumen tidak terstruktur seperti artikel, komentar atau laporan. Proses ini mencakup identifikasi kata atau frasa yang mewakili isi teks serta menemukan pola atau hubungan antar dokumen, untuk analisis lebih lanjut. Data teks diubah ke dalam bentuk numerik melalui tahapan awal yaitu preprocessing data. Selain itu, *text mining* juga dapat digunakan dalam analisis sentimen untuk mengidentifikasi opini atau emosi yang terkandung dalam teks[10].

2. *Natural Language Processing*

Natural Language Processing (NLP) merupakan cabang ilmu komputer yang berfokus pada pemerosesan bahasa manusia menggunakan teknologi mesin. Tujuannya adalah agar mesin mampu memahami dan merespon bahasa alami manusia. NLP banyak di implementasikan dalam berbagai bidang, seperti analisis sentimen, analisis media sosial, *chatbot,* penerjemah bahasa dan ekstraksi informasi. NLP dan pembelajaran mesin *(Machine Learning)* sangat berkaitan, Output NLP, seperti *preprocessing* teks, sering digunakan sebagai input algoritma pembelajaran mesin *(Machine Learning)*. Sebaliknya, metode pembelajaran mesin yang di awasi *(Supervised Learning)* dapat dimanfaatkan dalam membangun sumber daya NLP, misalnya pembuatan kamus untuk mengenali entitas. Aplikasi NLP bergantung pada penggunaan korpus, yaitu kumpulan data besar bahasa alami baik dalam bentuk tertulis maupun lisan, yang disimpan dalam sistem komputer untuk dianalisis. Korpus memiliki peranan penting dalam analisis statistik bahasa, seperti penghitungan frekuensi kata, serta dalam validasi aturan linguistik, untuk menemukan kesalahan tata bahasa[11].

3. *Preprocessing*

Preprocessing adalah langkah awal dalam pemrosesan data teks untuk mengubah data mentah yang tidak terstruktur menjadi format yang lebih teratur, sehingga analisis ini dapat dilakukan secara lebih akurat[12].

Berikut adalah proses tahapan penting teks *preprocessing* yaitu:

a. *Case folding* adalah proses mengubah semua huruf kapital menjadi huruf kecil *(Lowercase).*
b. *Cleaning* adalah proses untuk menghapus elemen-elemen yang tidak dibutuhkan dalam teks seperti emoji, mention dan data yang mengandung *noise.*
c. *Tokenizing* adalah proses memecah teks menjadi kata-kata atau bagian kecil.
d. *Normalization* adalah proses mengubah kata tidak baku menjadi kata baku.
e. *Stopword removal* adalah proses menghapus kata-kata yang tidak memiliki makna atau nilai dalam analisis teks.
f. *Stemming* adalah proses mengubah kata ke bentuk dasar atau akar-akarnya. Proses ini menggunakan *library Sastrawi*, hasil *stemming* akan sesuai aturan bahasa Indonesia seusai *library* tersebut.

4. TF IDF

Term Frequency-Inverse Document Frequency (TF-IDF) merupakan salah satu metode pembobotan kata yang digunakan untuk menilai seberapa penting sebuah kata *(term)* terhadap suatu dokumen dalam kumpulan dokumen *(corpus)*. Bobot yang dihasilkan oleh metode ini digunakan dalam tahap representasi teks ke dalam bentuk numerik agar dapat diproses[13]. TF-IDF banyak digunakan dalam analisis data teks seperti klasifikasi dokumen dan analisis sentimen karena mampu menyoroti kata-kat kunci yang relevan dan mengurangi bobot kata-kata umum yang kurang informatif.

1. *Term Frequency*

   Term Frequency (TF) menunjukan seberapa sering sebuah kata muncul dalam sebuah dokumen. Semakin sering kata tersebut muncul, semakin tinggi bobot yang diberikan. Nilai TF dihitung dengan membagi jumlah kemunculan kata dalam dokumen dengan total jumlah kata dalam dokumen tersebut. Berikut perhitungan TF pada rumus 1.

$$TF(t,d) = \frac{ft,d}{\sum kfk,d} \quad (1)$$

   Keterangan rumus TF:

$TF(t,d)$	=	Menunjukkan seberapa sering kata *t* muncul di dokumen *d*
ft,d	=	Jumlah kemunculan kata *t* dalam dokumen *d*
$\sum kfk,d$	=	Total jumlah kata dalam dokumen *d*

2. *Inverse Document Frequency*

   Inverse Document Frequency (IDF) digunakan untuk mengukur sejauh mana kata tertentu jarang muncul seluruh dokumen. Kata yang sering muncul di banyak dokumen akan mendapatkan bobot yang lebih rendah karena dianggap tidak memiliki ciri khas atau keunikan. Berikut perhitungan IDF pada rumus 2.

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

   Keterangan rumus IDF

$IDF(t)$	=	Mengukur seberapa unik kata *t* di seluruh dokumen
N	=	Total jumlah dokumen pada korpus
$DF(t)$	=	Jumlah dokumen yang mengandung kata *t*

3. TF-IDF

Nilai akhir pembobotan TF-IDF didapatkan dengan mengalikan hasil perhitungan TF dan IDF. Perhitungan ini memeberikan bobot yang tinggi pada kata yang sering muncul dalam suatu dokumen tetapi jarang ditemukan dalam dokumen lain. Berikut perhitungan TF-IDF pada rumus 3.

$$TF\ IDF(t,d) = TF\ (t,d) \times IDF(t) \quad (3)$$

Keterangan rumus TF-IDF:

$TF\ IDF(t,d)$	=	Nilai bobot akhir kata *t* dalam dokumen *d*, menggambarkan pentingnya kata tersebut
$TF\ (t,d)$	=	Menunjukkan seberapa sering kata *t* muncul di dokumen *d*
$IDF(t)$	=	Mengukur seberapa unik kata *t* di seluruh dokumen

5. *Naïve Bayes*

Naïve Bayes merupakan salah satu algoritma klasifikasi yang berbasis probabilistik yang menggunakan *Teorema Bayes*. Algoritma ini digunakan secara luas dalam berbagai bidang, terutama dalam analisis sentimen karena kemampuannya yang baik dalam menangani data teks. *Naïve Bayes* bekerja dengan mengasumsikan bahwa setiap fitur (atribut) bersifat independen terhadap fitur lainnya, meskipun dalam praktiknya asumsi ini jarang sepenuhnya terpenuhi[14]. Secara matematis, algoritma *Naïve Bayes* menggunakan *Teorema Bayes*, yang menggunakan rumus 4 sebagai berikut:

$$P(X|H) = \frac{P(X|H)P(H)}{P(X)} \quad (4)$$

Keterangan rumus *Teorema Bayes*:

X	=	Data dengan kelas yang belum diketahui.
H	=	Hipotesis data X termasuk dalam kelas tertentu.
P(H | X)	=	Probabilitas hipotesis H berdasarkan data x (posterior).
P(H)	=	Probabilitas awal hipotesis H (prior).
P(X | H)	=	Probabilitas data X muncul jika hipotesis H benar.
P(X)	=	Probabilitas keseluruhan dari data X.

Algoritma *Naïve Bayes* memiliki beberapa varian jenis dan bentuk data yang di gunakan. Masing-masing memiliki karakteristik dan penerapan tersendiri, antara lain:

1. *Multinomiinal Naïve Bayes* Merupakan varian yang paling umum digunakan klasifikasi teks seperti analisis sentimen. Model menghitung probabilitas berdasarkan jumlah kemunculan fitur (kata) dalam dokumen.
2. *Bernoulli Naïve Bayes* merupakan varian yang digunakan untuk data biner, yaitu mencatat apakah suatu fitur muncul atau tidak dalam dokumen, biasanya untuk representasi data dengan nilai (1 atau 0).
3. *Gaussian Naïve Bayes* merupakan varian yang cocok untuk data numerik yang di asumsikan mengikuti apakah mengikuti distribusi normal. Varian ini tidak digunakan dalam analisis teks, tetapi lebih banyak diterapkan pada data yang bersifat kontinu.

Analisis sentimen, *Naïve Bayes* digunakan untuk mengklasifikasi data teks ke dalam kategori sentimen tertentu, misalnya positif dan negatif. Langkah-langkah utama penggunaan algoritma ini sebagai berikut:

- Menentukan kategori atau kelas (positif dan negatif).
- Membuat tabel frekuensi kata berdasarkan kategori.
- Menghitung probabilitas kata dalam setiap kategori.
- Menghitung probabilitas keseluruhan untuk setiap kategori berdasarkan data input.
- Menentukan kategori dengan probabilitas tertinggi sebagai hasil klasifikasi.

6. *Confusion Matrix*

Confusion Matrix merupakan alat dalam data mining yang digunakan untuk mengevaluasi kinerja model klasifikasi dalam *Machine Learning*. Tabel ini membandingkan nilai aktual (kelas sebenarnya) dengan hasil prediksi model, di mana baris mewakili kelas aktual dan kolom mewakili prediksi. Dari tabel ini dapat dihitung berbagai matrix evaluasi, seperti *Accuracy, Precision, Recall* dan *F1-Score*. *Confusion Matrix* terdiri dari empat komponen utama, yaitu *True Positive* (TP), *False Positive* (FP), *False Negative* (FN), dan *True Negative* (TN) yang memberikan gambaran tentang keberhasilan dan kesalahan klasifikasi model[15]. *Confusion Matrix* bisa di lihat pada Tabel 1.

Tabel 1. *Confusion Matrix*

		Nilai Aktual	
		Positif	Negatif
NILAI PREDIKSI	Positif	*True Positive* (TP)	*False Positive* (FP)
	Negatif	*False Negative* (FN)	*True Negative* (TN)

Keterangan Tabel 1 *Confusion Matrix:*

1. *True Positive* (TP) adalah data yang memang positif dan berhasil dikenali dengan benar sebagai positif.
2. *False negative* (FN) adalah data yang sebenarnya positif, tapi salah dikenali sebagai negatif.
3. *False Positive* (FP) adalah data yang sebenarnya negatif, tapi salah dikenali sebagai positif.
4. *True Negative* (TN) adalah data yang memang negatif dan berhasil di kenal dengan benar sebagai negatif.

Cara menggunakan *Confusion Matrix*

Pengujian menggunakan *Confusion Matrix* memberikan empat jenis keluaran, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP) dan *False Negative* (FN). Keempat nilai ini digunakan untuk menghitung nilai *accuracy, precision, recall,* dan *F1-score*, yang merupakan metode evaluasi umum dalam menilai kinerja model klasifikasi pada *Machine Learning*. Evaluasi ini membantu mengukur seberapa baik algoritma dalam memprediksi data secara tepat.

a) *Accuracy*

Accuracy adalah metrik yang digunakan untuk mengukur ketepatan model dalam mengklasifikasikan data. Nilai ini dihitung berdasarkan perbandingan antara jumlah prediksi yang benar dengan total data yang diuji. Berikut perhitungan *accuracy* pada rumus 5.

$$Accuracy = \frac{TP + TN}{(TP + TN + FP + FN)} \quad (5)$$

b) *Precision*

Precision adalah metrik yang digunakan untuk mengukur seberapa banyak hasil prediksi positif yang benar dibandingkan dengan seluruh data yang diprediksi sebagai positif. Metrik ini penting untuk meminimalkan kesalahan dalam mengklasifikasikan data negatif sebagai positif *(false positive).* Berikut perhitungan *Precision* pada rumus 6.

$$Precision = \frac{TP}{(TP) + (FP)} \quad (6)$$

c) *Recall*

Recall adalah metrik yang digunakan untuk mengukur seberapa banyak data positif yang berhasil dikenali dengan benar oleh model. Metrik ini penting untuk menghindari kesalahan dalam mengabaikan data yang seharusnya termasuk ke dalam kelas positif *(false negative)*. Berikut perhitungan Recall pada rumus 7.

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

d) *F1 Score*

F1 Score adalah rata-rata dari perhitungan *Precision* dan *Recall* yang digunakan untuk mengevaluasi performa model, terutama saat diperlukan keseimbangan antara ketepatan prediksi dan kemampuan dalam mendeteksi data positif secara menyeluruh. Berikut perhitungan *F1 Score* pada rumus 8.

$$F1\ Score = 2\frac{(Precision\ x\ Recall)}{Precision + Recall} \quad (8)$$

7. Analisis Sentimen

Analisis Sentimen merupakan bidang ilmu komputasi yang berfokus pada pengkajian opini, pendapat, emosi dan perasaan individu yang diungkapkan dalam bentuk teks. Tujuan utama dari analisis sentimen adalah mengidentifikasi atribut dan komponen yang berkaitan dengan entitas dalam sebuah dokumen, serta menentukan apakah opini tersebut bersifat positif, negatif atau netral. Entitas tersebut bisa berupa produk, layanan, peristiwa, atau topik tertentu. Proses ini memanfaatkan *text mining*, yang merupakan bagian dari d*ata mining,* untuk mengekstrak informasi penting dari kumpulan teks besar dan tidak terstruktur. Dokumen-dokumen tersebut kemudian dianalisis dan diklasifikasikan berdasarkan sentimen yang terkandung didalamnya, sehingga membantu memahami persepsi masyarakat terhadap suatu topik[15].

8. *ChatGPT*

Generative Pre-trained Transformer (ChatGPT) merupakan teknologi kecerdasan buatan yang dikembangkan oleh *Open AI* yang dirancang untuk memahami bahasa alami dan berinteraksi dengan pengguna melalui percakapan berbasiskan teks. *ChatGPT* memiliki berbagai fungsi, seperti menerjemahkan bahasa, memberikan rekomendasi, meningkatkan produktivitas serta mendukung kegiatan di bidang pendidikan. Penggunaanya cukup sederhana dengan memasukan pertanyaan, kemudian sistem akan menghasilkan respon yang relevan, serta mampu memperbaiki jawaban apabila ditemukan ketidakakuratan. *ChatGPT* memanfaatkan informasi dari berbagai sumber, seperti artikel, dan media daring lainya yang tersedia di internet, kemudian mengolahnya menjadi jawaban yang ringkas dan terstruktur. *ChatGPT* mendukung personalisasi, dan aksesibilitas pembelajaran, menyediakan sumber belajar interaktif serta membantu pemecahan masalah. Namun demikian, teknologi AI ini juga memiliki keterbatasan, seperti pemahaman konteks yang terbatas, ketidakmampuan membedakan fakta dan opini, serta ketergantung pada koneksi internet[4].

9. *Web Scraping*

Web Scraping merupakan metode untuk mengumpulkan data informasi dari internet yang bersifat semi-terstruktur, seperti halaman web berbasis *HTML* atau *XHTML*. Teknik ini mempermudah proses pengolahan data dengan memanfaatkan pustaka pemrograman, salah satunya *Python. Web scraping* banyak diterapkan untuk keperluan riset pasar, analisis kompetitor, pengumpulan prospek, serta mengekstrak informasi spesifik dari berbagai situs *web,* terutama situs yang tidak

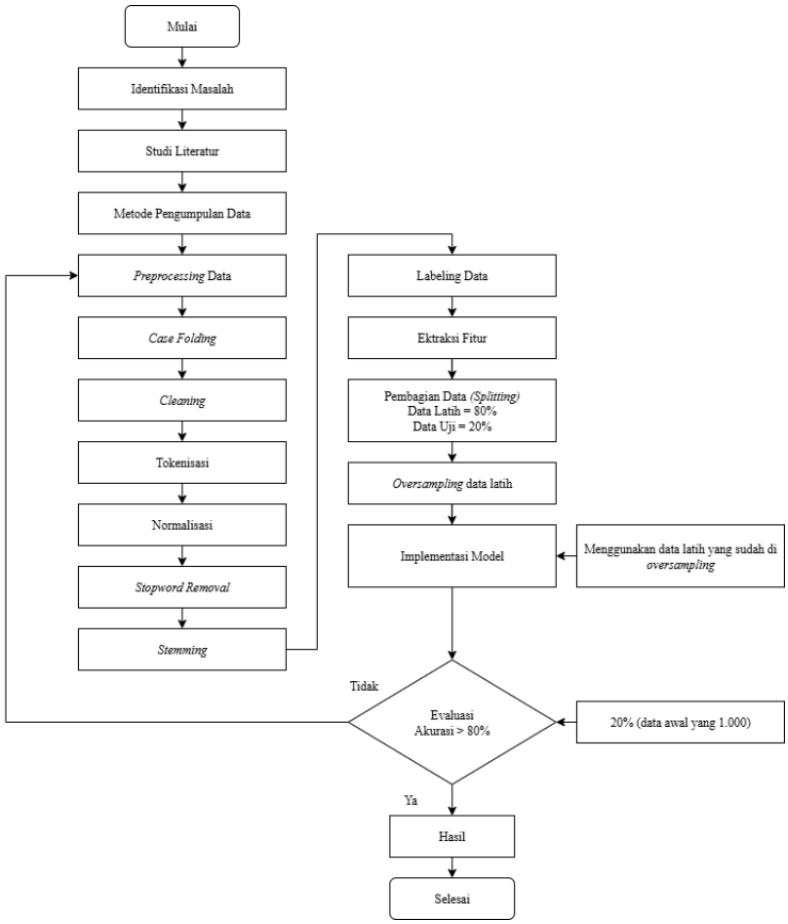
menyediakan cara mudah mengambil data, seperti API atau layanan ekspor data. Proses *scraping* dilakukan dengan mengunjungi halaman *web*, menganalisis konten dan mengekstrak elemen-elemen tertentu seperti teks, gambar, tautan atau data lainnya. Data yang dikumpulkan kemudian di konversi ke dalam format yang lebih struktur dan mudah digunakan seperti *speeadsheet* atau basis data, untuk analisis lebih lanjut atau aplikasi lainnya[16].

10. *Python*

Python adalah bahasa pemrograman interpretatif yang dikenal karena keterbacaan kode dan sintaks yang jelas. Bahasa pemrograman ini mendukung berbagai paradigma pemrograman, seperti pemrograman berorientasi objek, imperatif, fungsional dan prosedural. *Python* dikenal dengan pustaka komperhensif dan standar fungsional yang baik. Salah satu karakteristik utama *Python* yaitu sistem tipe dinamis dan manjemen memori otomatis, yang memudahkan pengelolaan kode. Meskipun sering digunakan untuk *scripting*, *Python* juga banyak dimanfaatkan dalam pengembangan perangkat lunak dan dapat di jalankan di berbagai platform sistem operasi, seperti *Windows, Linux, Unix, Mac OS X, dan Java Virtual Machine*[16].

III. Metodologi

Tahapan penelitian ini terdiri dari pengumpulan data, metode yang diusulkan serta pengujian dan kesimpulan. Dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

1. Identifikasi Masalah

Langkah Awal dalam penelitian ini adalah mengidentifikasi permasalahan yang akan diteliti. Permasalahan utama yang diangkat dalam penelitian ini adalah mengetahui persepsi atau opini pengguna aplikasi *ChatGPT* dari ulasan *Google Play Store*, semakin banyaknya penggunaan alat

bantu berbasis AI yang digunakan untuk pembelajaran dan pencarian informasi. Penggunaanya cukup populer, namun banyak ditemukan ulasan yang beragam dari para penggunanya, Tujuan analisis ini dilakukan untuk mengukur respon pengguna ke dalam tiga kategori positif, negatif dan netral.

2. Studi Literatur

Studi literatur dilakukan dengan mempelajari artikel jurnal yang berkaitan dengan metode analisis sentimen yang digunakan pada aplikasi berbasis teks. Manfaat studi literatur ini adalah untuk memperluas pemahaman mengenai konsep dan teori yang mendasari penelitian, khususnya *Naïve Bayes* terkait algoritma yang digunakan dalam klasifikasi sentimen mengenai penggunaan aplikasi *ChatGPT* pada lingkungan akademik.

3. Metode Pengumpulan Data

Informasi dataset yang digunakan dalam penelitian ini adalah data yang diperoleh dari ulasan pengguna aplikasi *ChatGPT* di *Google Play Store,* berisi teks ulasan melalui proses *web scraping* dengan menggunakan bahasa pemrograman *python*. Data yang di ambil 5000 ulasan dari periode juli 2023 sampai juli 2025 tujuan utama dari pengambilan data ini yaitu untuk menganalisis sentimen ulasan *Google Play Store* tersebut mengenai penggunaan aplikasi *ChatGPT* pada lingkungan akademik.

4. *Preprocessing* Data

Tahapan ini dilakukan untuk membersihkan dan menyiapkan data teks agar siap digunakan dalam proses klasifikasi. Beberapa tahapan penting seperti *Casefolding* merubah huruf kapital menjadi huruf kecil, *cleaning* menghapus elemen yang tidak dibutuhkan, *tokenizing, normalization, stopword* dan *stemming*. Tahapan ini bertujuan untuk mengubah data mentah menjadi data yang lebih terstrukur, menyederhanakan format teks agar lebih sesuai serta meningkatkan kualitasnya sehingga dapat diproses secara lebih optimal dalam analisis sentimen menggunakan algoritma *Naïve Bayes*. Data penelitian ini menggunakan ulasan aplikasi *ChatGPT* di *Google Play Store* melalui proses *scarping,* ulasan tersebut mengenai penggunaan aplikasi *ChatGPT* yang berkaitan dengan lingkungan akademik. Tahapan *preprocessing* ini berperan penting untuk menjamin kualitas data serta meningkatkan akurasi hasil klasifikasi menggunakan algoritma *Naïve Bayes.*

Berikut Langkah-langkah *preprocessing* data:

a. *Case folding* mengubah semua huruf kapital dalam teks menjadi huruf kecil *(lowercase)* bertujuan untuk representasi kata, sehingga kata seperti ("Bagus") dan ("bagus") dianggap sama, jika proses ini tidak dilakukan sistem akan menanggap sebagai dua kata berbeda padahal maknanya sama.
b. *Cleaning* menghapus karakter-karakter yang tidak dibutuhkan dalam teks seperti tanda baca, emoji, mention dan karakter khusus lainnya yang tidak memiliki kontribusi signifikan terhadap analisis.
c. *Tokenizing* memecah teks menjadi kata-kata atau bagian kecil. Misalnya (“aplikasi ini sangat membantu”) menjadi token (“aplikasi”, “ini”, “sangat”, “membantu”). Fungsinya setiap kata dapat dianalisis secara individu dalam proses klasifikasi atau pembobotan kata.
d. *Normalization* mengubah kata tidak baku atau singkatan menjadi kata baku. Misalnya ("gk", "nggak", "tdk") menjadi bentuk baku ("tidak") menggunakan kamus normalisasi
e. *Stopword removal* menghapus kata-kata umum misalnya (dan, yang, adalah, ke, di), kata tersebut tidak memiliki makna penting dalam konteks analisis, sedangkan kata yang dianggap penting adalah kata-kata yang membawa informasi utama seperti kata yang menunjukan sentimen (bagus, buruk, puas, kecewa) karena kata tersebut mempengaruhi hasil analisis.
f. *Stemming* mengubah kata berimbuhan ke bentuk dasar atau kata akar. Misalnya, kata ("bermain", "memainkan", dan "bermainlah") akan diubah menjadi ("main"). Proses ini menggunakan *library stemmer* Bahasa Indonesia *Sastrawi.*

5. Labelling Data

Labelling data dilakukan setelah proses *preprocessing* selesai, yaitu memberikan kategori sentimen positif, negatif dan netral pada setiap data ulasan. Pelabelan dengan menggunakan jenis analisis *sentimen Fine-Grained Sentiment Analysis*, Proses pelabelan ini data ulasan skor atau bintang tidak dijadikan acuan yang ada di *Google Play Store*, tapi menggunakan *Wilwo Sentimen Classifier*, yaitu

sebuah model *Pretrained* sehingga tidak memerlukan proses pelatihan dari awal dan langsung dapat digunakan untuk mengkategorikan data secara otomatis.

6. **Ekstraksi Fitur**
   Teks yang telah diproses dikonversi menjadi representasi numerik menggunakan metode *Term Frequency - Inverse Document Frequency* (TF-IDF). Data yang telah berbentuk numerik kemudian fitur di seleksi menggunakan metode *SelectKBest* dengan acuan nilai *Chi2,* Hal ini dilakukan agar data dapat diproses oleh algoritma pembelajaran mesin seperti *Naïve Bayes*.
7. **Pembagian Data**
   a. Penentuan Jumlah Dataset: total jumlah ulasan yang berhasil di kumpulkan, misalnya ada 5000 data ulasan.
   b. Pembagian Data: Dataset dibagi menjadi dua bagian data latih (training data) dan data uji (testing data) dengan rasio 80:20, contohnya jika menggunakan rasio 80:20 dan memiliki data 5000 ulasan, maka:
   - Data Latih: 4000 komentar (80%)
   - Data Uji: 1000 komentar (20%)

8. ***Oversampling* Data Latih**
   Proses setelah data dibagi menjadi data latih 80% dan data uji 20%, dilakukan teknik *oversampling* pada data latih untuk mengatasi ketidakseimbangan jumlah kelas sentimen agar memastikan jumlah data pada masing-masing kelas (positif, negatif dan netral) menjadi seimbang, sehingga model tidak bias terhadap kelas mayoritas
9. **Implementasi Model**
   Implementasi model dilakukan setelah data siap digunakan hasil (*preprocessing*, pelabelan dan pembagian data). Implementasi ini bertujuan untuk membangun melatih dan mengevaluasi model klasifikasi sentimen. Data latih yang telah melalui proses ektraksi fitur dan *oversampling* digunakan untuk membangun model klasifikaksi algoritma *Multinominal Naïve Bayes* dengan melakukan pencarian *tuning hyperparameter* nilai *alpha* terbaik menggunakan *GridSearchCV* kemudian parameter terbaik digunakan untuk melakukan prediksi pada data uji. Hasil prediksi dibandingkan dengan label sebenarnya untuk dilakukan evaluasi model, evaluasi dilakukan menggunakan *Confusion matrix*.
10. **Evaluasi**
    Evaluasi ini dilakukan kinerja model menggunakan *confusion matrix,* untuk melihat seberapa baik model dapat mengklasifikasikan data uji ke dalam label yang sesuai untuk mendapatkan nilai *accuracy, precision, recall,* dan *F1-Score* dari klasifikasi. Jika hasil evaluasi belum berhasil (Tidak) maka di lakukan kembali pada pengolahan data karena ada hal yang perlu diperbaiki seperti error ataupun akurasi yang belum tercapai > 80% dengan melakukan perbaikan *preprocessing, balancing* data, ataupun *tuning hyperparameter*, jika evaluasi berhasil (Ya) maka hasil evaluasi model sudah sesuai yang diinginkan maka ke tahap terakhir hasil.
11. **Hasil**
    Tahapan ini merupakan langkah akhir setelah seluruh proses penelitian dilaksanakan dan menghasilkan output yang sesuai dengan yang di harapkan, hasil tersebut kemudian diimplementasikan ke dalam sebuah program *Graphical User Interface* (GUI). Seluruh rangkaian tahapan penelitian dirangkum dan disimpulkan pada tahap akhir ini.

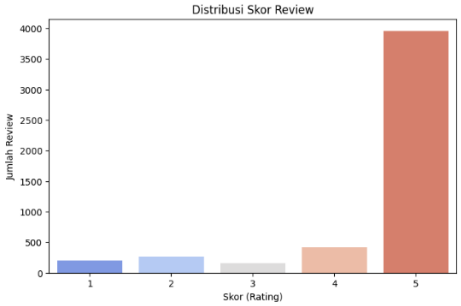
IV. Hasil dan Pembahasan

Pembahasan ini menyampaikan hasil identifikasi sentimen pengguna aplikasi *ChatGPT* dengan 5000 data ulasan dari *Google Play Store.* Data ulasan di analisis dengan pendekatan *Fine-Grained Sentiment Analysis* yang mengelompokan ulasan ke dalam tiga kategori positif, negatif dan netral. Proses analisis mencakup pengumpulan data, tahapan *preprocessing* teks, pelabelan data, hingga klasifikasi menggunakan algoritma *Multinomial Naïve Bayes*. Hasil klasifikasi kemudian dievaluasi menggunakan

Confusion Matrix untuk mengukur performa model berdasarkan nilai *Accuracy, Precision, Recall,* dan *F1-Score.*

1. Pengumpulan Data

Data yang di peroleh merupakan hasil *scraping* dari yang diambil langsung dari ulasan komentar aplikasi *ChatGPT* pada *Google Play Store* melalui tautan 'com.openai.chatgpt' Dataset yang di kumpulkan sebesar 5000 data, dengan rentang waktu Juli 2023 hingga Juli 2025. Proses *scraping* menggunakan bahasa pemrograman *python* dengan *library google_play_scraper* dengan parameter *MOST_RELEVANT*. Data ulasan hasil *scraping* mengambil sejumlah informasi seperti, *username, score, at,* dan *content,* setiap pengguna memberikan score yang berbeda-beda dari jumlah skor tertinggi 5 dan skor terendah 1 dapat dilihat pada gambar 2.



Gambar 2. Diagram Pengumpulan Data

Data ulasan sebanyak 5.000 yang sudah dikumpulkan dari ulasan pengguna aplikasi *ChatGPT* di *Google Play Store* skor penilaian dalam ulasan bervariasi dari skor 1 hingga skor 5 dengan distribusi jumlah data yang berbeda pada setiap tingkat skor. Skor 1 memiliki 200 ulasan, skor 2 sebanyak 269 ulasan, skor 3 sebanyak 162 ulasan, skor 4 sebanyak 413 ulasan, dan skor 5 mendominasi dengan jumlah data 3956 ulasan. Distribusi jumlah review berdasarkan skor yang diberikan dapat di lihat pada Tabel 2.

Tabel 2. Distribusi jumlah review berdasarkan skor

No	Skor (Bintang)	Jumlah Data Ulasan
1.	1	200
2.	2	269
3.	3	162
4.	4	413
5.	5	3956
6.	**Total**	5000

Data yang di ambil dari ulasan memiliki skor yang diberikan pengguna, tetapi skor tersebut tidak selalu mencerminkan isi ulasan secara tepat sebagai contoh terdapat pengguna memberikan skor bintang 5 namun dalam ulasanya menyebutkan "aplikasi *ChatGPT* membuat malas belajar" dan masih banyak contoh ketidaksesuaian antara skor dan isi ulasan. Hasil *scraping* data ulasan berupa username, skor dan tanggal ulasan dapat dilihat pada Tabel 3.

Tabel 3. Ulasan Hasil *Scraping*

No	*username*	*score*	*date*	*content*
1.	Pengguna Google	5	31/1/2025	aplikasi ya bagus buat ngerjain tugas

No	*username*	*score*	*date*	*content*
2.	Pengguna Google	5	31/1/2025	ChatGPT ini bagus untuk di tanyakan pertanyaan penting atau untuk material sekolah
3.	Pengguna Google	3	22/1/2025	kurang karena mengirim soal menggunakan foto di batasi karena premium
4.	Pengguna Google	4	23/1/2025	sangat membantu sekali bagi para mahasiswa untuk mencari referensi, dan juga lengkap
5.	Pengguna Google	1	19/5/2025	banyak soal yg dijawab salah, bikin kesel
6.	Pengguna Google	2	19/5/2025	Aplikasinya kurang efektif kalo digunakan untuk mengerjakan tugas tapi bisa slebew in orang

2. *Preprocessing Data*

Proses *web scraping* di gunakan untuk mengambil data ulasan aplikasi *ChatGPT* dari *Google Play Store* yang di simpan dalam format CSV. Data ulasan yang di ambil untuk analisis 5000 data, semua data yang digunakan harus melalui beberapa tahapan *Preprocessing*, untuk mengubah data mentah menjadi data yang lebih terstrukur, menyederhanakan format teks agar lebih sesuai serta meningkatkan kualitasnya sehingga dapat diproses secara lebih optimal dalam analisis. Beberapa tahapan *Preprocessing* yaitu *Casefolding*, *cleaning, tokenizing, normalization, stopword* dan *stemming*.

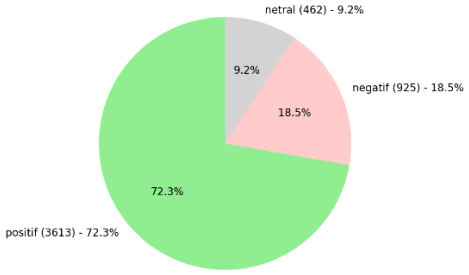
Tabel 4. Hasil Tahapan *Preprocessing*

No	*Content*	*Casefolding*	*Cleaning*	*Tokenizing*	*Nomalization*	*Stopword Removal*	*Stemming*
1.	aplikasi eror udh ngisi pertanyaan jawaban ga muncul	aplikasi eror udh ngisi pertanyaan jawaban ga muncul	aplikasi eror udh ngisi pertanyaan jawaban ga muncul	[“aplikasi”, “eror”, “udh”, “ngisi”, “pertanyaan”, “jawaban”, “ga”, “muncul”]	[“aplikasi”, “eror”, “sudah”, “ngisi”, “pertanyaan”, “jawaban”, “tidak”, “muncul”]	[“aplikasi”, “eror”, “ngisi”, “pertanyaan”, “jawaban”, “tidak”, “muncul”]	[“aplikasi”, “eror”, “ngisi”, “tanya”, “jawab”, “tidak”, “muncul”]
2.	sangat membantu tugas harian saya 🙏 👍	sangat membantu tugas harian saya 🙏 👍	sangat membantu tugas harian saya	[“sangat”, “membantu”, “tugas”, “harian”, “saya”]	[“sangat”, “membantu”, “tugas”, “harian”, “saya”]	[“sangat”, “membantu”, “tugas”, “harian”]	[“sangat”, “bantu”, “tugas”, “hari”]

3. Labelling Data

Proses pemberian label pada data dilakukan dengan Pelabelan dengan menggunakan jenis analisis sentimen *Fine-Grained Sentiment,* proses pelabelan 5000 data ulasan skor atau bintang tidak

dijadikan acuan yang ada di *Google Play Store*, tapi menggunakan *Wilwo Sentimen Classifier*, yaitu sebuah model *Pretrained* sehingga tidak memerlukan proses pelatihan dari awal dan langsung dapat digunakan untuk mengkategorikan data secara otomatis. Penelitian ini menggunakan *Fine-Grained Sentiment Analysis* dengan tiga label kategori positif, negatif dan netral, label tersebut diberikan pada sebuah data untuk menunjukkan apakah data tersebut masuk ke dalam kategori positif, negatif maupun netral dengan suatu opini atau persepsi tertentu. Misalnya jika ulasan kalimatnya itu bersifat membantu, memuji, biasa saja, ataupun mengkritik makan label memuji akan masuk kategori positif sebaliknya jika isinya kritikan maka label negatif dan kalimat biasa saja bisa masuk ke netral dapat dilihat pada Gambar 4.2, dimana label positif lebih dominan mencapai 72.3%, data berlabel negatif mencapai 18.5% dan netral 9.2%.



Gambar 3. Labelling Data Ulasan

4. Ektraksi Fitur

Proses ektraksi fitur merupakan tahapan penting analisis sentimen yaitu memberikan nilai atau bobot pada setiap kata dalam teks untuk menunjukkan kontribusi pada kata tersebut terhadap seluruh sentimen positif, negatif dan netral dari teks tersebut. Tujuannya untuk mengidentifikasi kata-kata yang lebih signifikan dalam menentukan sentimen yang jelas. Data penelitian ini menggunakan 5000 data ulasan pengguna aplikasi *ChatGPT* di *Google Play Store*, namun karena data yang sangat besar maka ditampilkan beberapa hasil pembobotan TF-IDF menggunakan contoh tiga kalimat ulasan yang sudah melalui tahapan *preprocessing* dan labelling.

1. Aplikasi sangat bagus bisa kerja tugas baik. D1
2 Sangat bantu kerja tugas yang sulit. D2
3. Banyak soal jawab salah bikin kesel. D3

Proses untuk mengekstraksi fitur dari kalimat tersebut digunakan metode pembobotan teks yaitu

1) *Term Frequency* (TF) menunjukan seberapa sering sebuah kata muncul dalam sebuah dokumen. TF = jumlah kemunculan kata / total kata dalam dokumen.

Tabel 4. *Term Frequency* Jumlah Kemunculan Kata dalam Dokumen

No	Kata	Dokumen d1	Dokumen d2	Dokumen d3
1.	Aplikasi	1	0	0
2.	Sangat	1	1	0
3.	Bagus	1	0	0
4.	Bisa	1	0	0
5.	Kerja	1	1	0
6.	Tugas	1	1	0
7.	Baik	1	0	0
8.	Bantu	0	1	0

No	Kata	Dokumen d1	Dokumen d2	Dokumen d3
9.	Sulit	0	1	0
10.	banyak	0	0	1
11.	Soal	0	0	1
12.	Jawab	0	0	1
13	Salah	0	0	1
14.	Bikin	0	0	1
15.	Kesel	0	0	1
16.	**Total kata**	**7**	**5**	**6**

Tabel 5. *Term Frequency* Proporsi Frekuensi Kata dalam Dokumen

No	Kata	TF-IDF d1	TF-IDF d2	TF-IDF d3
1.	Aplikasi	0.1429	0	0
2.	Sangat	0.1429	0.2000	0
3.	Bagus	0.1429	0	0
4.	Bisa	0.1429	0	0
5.	Kerja	0.1429	0.2000	0
6.	Tugas	0.1429	0.2000	0
7.	Baik	0.1429	0	0
8.	Bantu	0	0.2000	0
9.	Sulit	0	0.2000	0
10.	Banyak	0	0	0.1667
11.	Soal	0	0	0.1667
12.	Jawab	0	0	0.1667
13.	Salah	0	0	0.1667
14.	Bikin	0	0	0.1667
15.	Kesel	0	0	0.1667

2) *Inverse Document Frequency (IDF)* mengukur sejauh mana kata tertentu jarang muncul seluruh dokumen. Nilai IDF dihitung berdasarkan jumlah dokumen yang mengandung kata tersebut, dengan perhitungan IDF = LOG10 (jumlah dokumen / jumlah dokumen yang mengandung kata).

Tabel 6. *Inverse Document Frequency* (IDF)

No	Kata	Df	Idf
1.	Aplikasi	1	0.4771
2.	Sangat	2	0.1761
3.	Bagus	1	0.4771
4.	Bisa	1	0.4771

No	Kata	Df	Idf
5.	Kerja	2	0.1761
6.	Tugas	2	0.1761
7.	Baik	1	0.4771
8.	Bantu	1	0.4771
9.	Sulit	1	0.4771
10.	Banyak	1	0.4771
11.	Soal	1	0.4771
12.	Jawab	1	0.4771
13.	Salah	1	0.4771
14.	Bikin	1	0.4771
15.	Kesel	1	0.4771

3) *Term Frequency - Inverse Document Frequency* (TF-IDF) mengalikan nilai TF dan IDF untuk menentukan bobot akhir dari setiap kata dalam setiap dokumen.

Tabel 7. TF-IDF Bobot Akhir Tiap Kata dalam Setiap Dokumen

No	Kata	TF-IDF d1	TF-IDF d2	TF-IDF d3
1.	Aplikasi	0.0682	0	0
2.	Sangat	0.0252	0.0352	0
3.	Bagus	0.0682	0	0
4.	Bisa	0.0682	0	0
5.	Kerja	0.0252	0.0352	0
6.	Tugas	0.0252	0.0352	0
7.	Baik	0.0682	0	0
8.	Bantu	0	0.0954	0
9.	Sulit	0	0.0954	0
10.	Banyak	0	0	0.0795
11.	Soal	0	0	0.0795
12.	Jawab	0	0	0.0795
13.	Salah	0	0	0.0795
14.	Bikin	0	0	0.0795
15.	Kesel	0	0	0.0795

5. Pembagian Data

Proses pembagian data atau mengelompokan data menjadi data latih dan data uji digunakan *library sklearn* dengan mengambil fungsi *train_test _split* dari *sklearn.model_selection*. Fungsi ini digunakan untuk membagi dataset menjadi dua bagian, yaitu data latih 80% *(training set)* yang digunakan dalam pelatihan model, dan data uji 20% *(testing set)* yang di pakai untuk mengukur performa model setelah pelatihan selesai. Persentase pembagian data biasanya disesuaikan dengan

kebutuhan penelitian agar performa model memperoleh cukup data untuk melatih sekaligus dapat dievaluasi secara objektif. Pembagian data dilakukan secara acak *(random sampling)* sehingga distribusi label pada kedua subset tetap proporsional dan representatif terhadap keseluruhan dataset, jumlah pembagian dataset dapat dilihat pada Tabel 8.

Tabel 8. Jumlah Pembagian Dataset

No	Dataset	Jumlah	Data Latih	Data Uji
1	Positif	3613	2890	722
2	Negatif	925	740	185
3	Netral	462	369	92
Total		5000	4000	1000

6. *Oversampling*

Proses *oversampling* ini hanya digunakan pada data latih setelah proses pembagian data untuk menghindari kebocoran data *(data leakage)* pada data uji. Evaluasi model tetap menggunakan data uji yang mempersentasikan distribusi asli. Dapat dilihat pada Tabel 9 Jumlah Data Latih *Oversampling.*

Tabel 9. Jumlah Data Latih *Oversampling*

No	Label sentimen	Jumlah	Data Latih	Data Latih *Oversampling*
1.	Positif	3613	2890	2890
2.	Negatif	925	740	2890
3.	Netral	462	369	2890
4.	**Total**	5000	4000	8670

7. Klasifikasi Algortima *Naïve Bayes*

Proses klasifikasi data hasil ektraksi fitur dilakukan menggunakan algoritma *Multinominal Naïve Bayes* untuk mendapatkan atau meningkatkan performa model yang optimal, dilakukan pencarian *hyperparameter* untuk menemukan nilai terbaik dari *parameter alpha (smoothing)* menggunakan *GridSearchCV*.

- Algoritma *Multinominal Naïve Bayes* digunakan untuk klasifikasi teks dengan menghitung probabilitas frekuensi kata.
- *Parameter alpha (smoothing)* untuk menghindari nilai probabilitas nol pada kata-kata yang tidak muncul di data latih.

Pencarian nilai *hyperparameter alpha* yang optimal dilakukan secara sistematis menggunakan metode *GridSearchCV*.

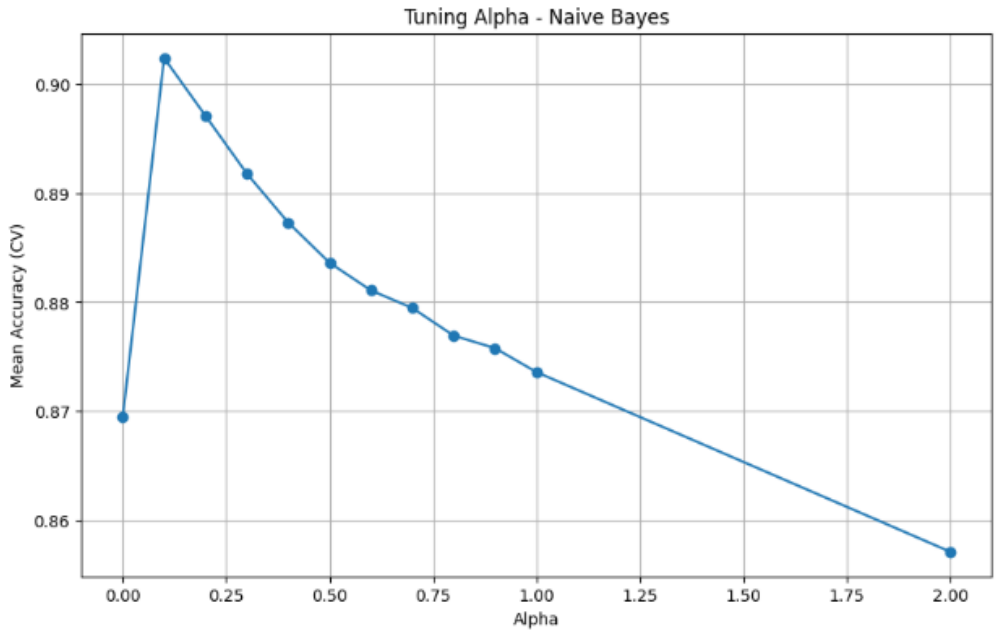
- Nilai *alpha* yang diuji: [0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 2.0]
- Teknik Validasi Silang: 10-*fold cross-validation* (CV=10)

Metode ini akan menguji setiap nilai *alpha* secara sistematis menggunakan validasi silang yang dilakukan sebanyak 10 kali pembagian data untuk menemukan nilai yang memberikan kinerja model terbaik dapat dilihat pada Tabel 10.

Tabel 10. Hasil *Tuning Hyperparameter Alpha*

No	*Alpha*	*Mean Accuracy (CV=10)*
1	0.0	0.869435
2	0.1	0.902422
3	0.2	0.897116
4	0.3	0.891811
5	0.4	0.887313
6	0.5	0.881804
7	0.6	0.881084
8	0.7	0.879049
9	0.8	0.876932
10	0.9	0.875779
11	1.0	0.873527
12	2.0	0.857093

Hasil *tuning hyperparameter alpha* pada algoritma *Multinominal Naïve Bayes* menggunakan *GridSearchCV* dengan 10-*fold cross-validation* diperoleh nilai *alpha* optimal terbaik sebesar 0.1 dengan rata-rata akurasi 90.24%.



Gambar 4. Grafik *Tuning Alpha*

8. Evaluasi

Proses untuk mendapatkan nilai akurasi yang optimal dan gambaran yang jelas tentang kinerja model, digunakan *Confusion Matrix, yaitu Accuracy, Precision, Recall, F1-Score,* untuk mendapatkan nilai dari model proses evaluasi model dilakukan dengan metode perhitungan *Confusion Matrix* dapat dilihat pada Tabel 11.

Tabel 11. Hasil *Confusion Matrix*

		Nilai Prediksi		
		Prediksi Negatif	Prediksi Netral	Prediksi Positif
Nilai Aktual	Negatif	142	6	37
	Netral	21	40	31
	Positif	43	45	635

Hasil *Confusion Matrix* untuk mendapatkan nilai *accuracy, precission, recall* dan *F1-Score* dari model menggunakan pengujian atau perhitungan *Confusion Matrix.* Hasil Evaluasi *Confusion Matrix* dapat dilihat pada Tabel 11.

Tabel 12. Hasil Evaluasi *Confusion Matrix*

	Precission	***Recall***	***F1-Score***	***Support***
Negatif	0.6893	0.7676	0.7263	185
Netral	0.4396	0.4348	0.4372	92
Positif	0.9033	0.8783	0.8906	723
Macro Avg	**0.6774**	**0.6935**	**0.6847**	**1000**
Weighted AVg	**0.821**	**0.817**	**0.8185**	**1000**
Accuracy				**0.817**

Hasil dari Evaluasi model *Confusion Matrix* pada tabel 4.13, model memperoleh Akurasi sebesar 81%, dengan nilai *Precission* 82%, nilai *Recall* 81%, nilai *F1-Score* 81%. Hasil Evaluasi model *Confusion Matrix* dapat di simpulkan bahwa model sudah bekerja dengan baik untuk analisis sentimen penggunaan aplikasi *ChatGPT.*

V. Kesimpulan dan Saran

1. Kesimpulan

Berdasarkan hasil penelitian Analisis Sentimen penggunaan Aplikasi *ChatGPT* dengan menggunakan algoritma *Naïve Bayes* berhasil dilakukan dan memperoleh akurasi sebesar 81%. Penelitian ini menggunakan 5.000 data yang di ambil dari ulasan pengguna aplikasi *ChatGPT* di *Google Play Store* dari bulan Juli 2023 hingga Juli 2025. Hasil labelling data menunjukan distibusi sentimen positif sebesar 72.3%, negatif 18.5% dan netral 9.2%. Hasil evaluasi model menggunakan *Confusion Matrix* menghasilkan nilai rata-rata *Accuracy* sebesar 81%, *Precission* 82%, *Recall* 81% dan *F1-Score* 81%. Hasil kesimpulan algoritma *Naïve Bayes* mampu mengklasifikasikan data teks ulasan dengan akurasi sebesar 81%, menunjukan algoritma *Naïve Bayes* cukup baik untuk analisis sentimen penggunaan aplikasi *ChatGPT*.

2. Saran

Hasil analisis pembahasan dan kesimpulan penelitan yang telah dilakukan, penulis menyadari penelitian ini masih jauh dari kata sempurna, maka dari itu untuk mengembangkan penelitian yang lebih baik penulis memberikan saran sebagai berikut:

a. Pengujian data dapat dilakukan menggunakan algoritma klasifikasi teks lainnya agar dapat diketahui perbandingan kinerja algoritma seperti SVM, BERT, KNN sehingga dapat diperoleh model terbaik dengan akurasi tertinggi.
b. Penelitian selanjutnya dapat mengambil data dari platform lain seperti X, *Instagram* dan forum diskusi ataupun mengisi *kuesioner Google Froms* sehingga hasil analisis lebih beragam sehingga dapat merepresenatikan sudut pandang secara lebih luas.

c. Pra-pemerosesan data dapat ditingkatkan dengan memperhatikan normalisasi kata, penghapusan *stopword y*ang lebih spesifik atau teknik *stemming* dan *lemmatization* yang lebih baik agar kualitas data menjadi lebih optimal.
d. Penelitian ini jika hasil evaluasi belum mencapai 80% maka perlu dilakukan perbaikan dari tahap *preprocessing* data, penyesuaian *parameter* TF-IDF, *balancing* data, serta *tuning hyperparameter* agar kualitas data meningkat dan performa model meningkat, kemudian model dilatih dan dievaluasi hingga performa mencapai target lebih dari 80%.

Referensi

[1] T. Saputra and S. Serdianus, "Peran A*rtificial Intelligence ChatGPT* dalam Perencanaan Pembelajaran di era Revolusi Industri 4.0," *J. Ilmu Sos. dan Pendidik.*, vol. 3, no. 1, pp. 1–18, 2023.

[2] H. N. Cahyanto, P. Pamungkas, and O. Zulkarnain, "Pengaruh Penggunaan *ChatGPT* Terhadap Kemandirian Mahasiswa dalam Menyelesaikan Tugas Akademik," *PREPOTIF J. Kesehat. Masy.*, vol. 8, no. 1, pp. 930–935, 2024.

[3] M. Ramli, "Mengeksplorasi Tantangan Etika Dalam Penggunaan *Chat GPT* Sebagai Alat Bantu Penulisan Ilmiah: Pendekatan Terhadap Integritas Akademik," *TA'DIBAN J. Islam. Educ.*, vol. 4, no. 1, pp. 1–10, 2023, doi: 10.61456/tjie.v4i1.129.

[4] W. Suharmawan, "Pemanfaatan Chat GPT dalam Dunia pendidikan," *Educ. J. J. Educ. Res. Dev.*, vol. 7, no. 2, pp. 13–20, 2023.

[5] Aluisius Dwiki Adhi and Safitri Juanita, "Analisis Sentimen pada Ulasan Pengguna Aplikasi Bibit dan Bareksa dengan Algoritma KNN," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 2, pp. 636–646, 2021.

[6] N. Agustina, Adrian, and M. Hermawati, "Implementasi Algoritma *Naïve Bayes Classifier* untuk Mendeteksi Berita Palsu pada Sosial Media," *Fakt. Exacta*, vol. 14, no. 4, p. 206, 2021, doi: 10.30998/faktorexacta.v14i4.11259.

[7] I. H. Kusuma and N. Cahyono, "Analisis Sentimen Masyarakat Terhadap Penggunaan E-Commerce Menggunakan Algoritma *K-Nearest Neighbor*," *J. Inform. J. Pengemb. IT*, vol. 8, no. 3, pp. 302–307, 2023, doi: 10.30591/jpit.v8i3.5734.

[8] M. F. Y. Herjanto and C. Carudin, "Analisis Sentimen Ulasan Pengguna Aplikasi Sirekap Pada *Play Store* Menggunakan Algoritma *Random Forest Classifer*," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 2, pp. 1204–1210, 2024, doi: 10.23960/jitet.v12i2.4192.

[9] A. T. Octa Nuryawan, M. Hasbullah, M. Rizal, M. F. Rajab, and N. Agustina, "Algoritma *Decision Tree* Untuk Analisis Sentimen Public Terhadap Marketplace Di indonesia," *Naratif J. Nas. Riset, Apl. dan Tek. Inform.*, vol. 5, no. 1, pp. 18–25, 2023, doi: 10.53580/naratif.v5i1.186.

[10] H. P. Doloksaribu and Yusran Timur Samuel, "Komparasi Algoritma Data Mining Untuk Analisis Sentimen Aplikasi Pedulilindungi," *J. Teknol. Inf. J. Keilmuan dan Apl. Bid. Tek. Inform.*, vol. 16, no. 1, pp. 1–11, 2022, doi: 10.47111/jti.v16i1.3747.

[11] Ernianti Hasibuan and Elmo Allistair Heriyanto, "Analisis Sentimen Pada Ulasan Aplikasi Amazon Shopping Di *Google Play Store* Menggunakan *Naive Bayes Classifier*," *J. Tek. dan Sci.*, vol. 1, no. 3, pp. 13–24, 2022, doi: 10.56127/jts.v1i3.434.

[12] K. S. Putri, I. R. Setiawan, and A. Pambudi, "Analisis Sentimen Terhadap Brand Skincare Lokal Menggunakan *Naïve Bayes Classifier*," *Technol. J. Ilm.*, vol. 14, no. 3, p. 227, 2023, doi: 10.31602/tji.v14i3.11259.

[13] G. Rininda, I. Hartami Santi, and S. Kirom, "Penerapan Svm Dalam Analisis Sentimen Pada Edlink Menggunakan Pengujian *Confusion Matrix*," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 5, pp. 3335–3342, 2023, doi: 10.36040/jati.v7i5.7420.

[14] A. H. Hasugian, "Analisis Sentimen Publik Pada Aplikasi X Terhadap Kenaikan UKT Di

Indonesia Menggunakan Algoritma *Naive Bayes*," *J. Pendidik. dan Teknol. Indones. Vol.*, vol. 5, no. 2, pp. 381–392, 2025.

[15] N. Fibriyanti Arminda, N. Sulistiyowati, and T. Nur Padilah, "Implementasi Algoritma *Multinomial Naive Bayes* Pada Analisis Sentimen Terhadap Ulasan Pengguna Aplikasi Brimo," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 3, pp. 1817–1822, 2023, doi: 10.36040/jati.v7i3.7012.

[16] A. Suryadi, W. A. Syb'an, N. Alfa'inna, and E. H. Hermaliani, "Implementasi *Web Scraping* dan *Sentiment Analysis* Terhadap Berita Menggunakan *Machine Learning*," *Swabumi*, vol. 11, no. 1, pp. 28–34, 2023, doi: 10.31294/swabumi.v11i1.15145.