

**OPTIMALISASI ALGORITMA MACHINE LEARNING
UNTUK ANALISIS SENTIMEN ULASAN PENGGUNA
CHATGPT DI GOOGLE PLAY STORE**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh
RAJULAIN JUWANTO
21.11.3926

Kepada
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025

**OPTIMALISASI ALGORITMA MACHINE LEARNING
UNTUK ANALISIS SENTIMEN ULASAN PENGGUNA
CHATGPT DI GOOGLE PLAY STORE**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

RAJULAIN JUWANTO

21.11.2926

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**OPTIMALISASI ALGORITMA MACHINE LEARNING UNTUK
ANALISIS SENTIMEN ULASAN PENGGUNA CHATGPT DI
GOOGLE PLAY STORE**

yang disusun dan diajukan oleh

Rajulain Juwanto

21.11.3926

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 11 Agustus 2024

Dosen Pembimbing,



Ferlan Fauzi Abdulloh, S.Kom., M.Kom

NIK. 190302276

HALAMAN PENGESAHAN
SKRIPSI
OPTIMALISASI ALGORITMA MACHINE LEARNING UNTUK
ANALISIS SENTIMEN ULASAN PENGGUNA CHATGPT DI
GOOGLE PLAY STORE

yang disusun dan diajukan oleh

Rajulain Juwanto

21.11.3926

Telah diperintahkan di depan Dewan Penguji
pada tanggal 11 Agustus 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Windha Mega Pradnya Dhuhrta, S.Kom., M.Kom.
NIK. 190302185



Theophilus Bayu Sasongko, S.Kom., M.Eng.
NIK. 190302375



Ferian Fauzi Abdullah, S.Kom., M.Kom.
NIK. 190302276



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 11 Agustus 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Rajulain Juwanto
NIM : 21.11.3926

Menyatakan bahwa Skripsi dengan judul berikut:

Optimalisasi Algoritma Machine Learning Untuk Analisis Sentimen Ulasan Pengguna Chatgpt Di Google Play Store

Dosen Pembimbing : Ferjan Fauzi Abdulloh, S.Kom., M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 11 Agustus 2025

Yang Menyatakan,



Rajulain Juwanto

HALAMAN PERSEMBAHAN

Puji syukur kehadiran Allah SWT atas segala penyertaan dan rahmat-Nya yang tak terhingga, sehingga skripsi ini dapat diselesaikan dengan sebaik-baiknya.

Dengan rasa syukur yang mendalam, karya ini dipersembahkan kepada Allah SWT, yang telah melimpahkan nikmat dan kemampuan sehingga penyusunan skripsi ini dapat terselesaikan.

Selanjutnya, skripsi ini dipersembahkan kepada keluarga, terkhusus kepada Bapak, Ibu, dan Kakak, atas doa, pengertian, serta dukungan yang tulus, yang menjadi pilar utama dalam menyelesaikan studi dan meraih gelar sarjana.

Tidak lupa, terima kasih disampaikan kepada semua pihak yang telah memberikan bantuan dalam penyusunan skripsi ini, yang tidak dapat disebutkan satu per satu. Semoga Allah SWT senantiasa melimpahkan perlindungan, kesejahteraan, dan kedamaian.

KATA PENGANTAR

Puji syukur kehadirat Allah SWT, atas rahmat dan karunia-Nya, sehingga skripsi yang berjudul "OPTIMALISASI ALGORITMA MACHINE LEARNING UNTUK ANALISIS SENTIMEN ULASAN PENGGUNA CHATGPT DI GOOGLE PLAY STORE" ini dapat terselesaikan dengan baik. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas AMIKOM Yogyakarta.

Penyusunan skripsi ini tidak lepas dari bimbingan, dukungan, dan bantuan dari berbagai pihak. Oleh karena itu, ucapan terima kasih yang tulus disampaikan kepada:

1. Allah SWT, atas segala petunjuk dan kemudahan yang dilimpahkan.
2. Orang tua tercinta, atas doa, kasih sayang, dan pengorbanan yang tak terhingga.
3. Bapak Ferian Fauzi Abdulloh, M.Kom, selaku Dosen Pembimbing Skripsi, atas waktu, arahan, bimbingan, dan masukan yang sangat berharga selama proses penyusunan skripsi ini.
4. Ibu Prof. Dr. Kusri, M.Kom., selaku Dekan Fakultas Ilmu Komputer, Universitas AMIKOM Yogyakarta, atas dukungan dan fasilitas yang diberikan.
5. Ibu Windha Mega Pradnya Dhuhi M. Kom., Theopilus Bayu Sasongko, S.Kom., M.Eng dan Bapak Ferian Fauzi Abdulloh M. Kom., selaku Dewan Penguji Skripsi, atas waktu, saran, dan kritik membangun yang diberikan demi kesempurnaan skripsi ini.
6. Segenap jajaran dosen dan staf karyawan Fakultas Ilmu Komputer, Universitas AMIKOM Yogyakarta, atas ilmu dan fasilitas yang telah diberikan selama masa perkuliahan.
7. Teman-teman seperjuangan Program Studi Informatika, atas kebersamaan, motivasi, dan dukungan selama masa perkuliahan.
8. Semua pihak yang tidak dapat disebutkan satu per satu, atas bantuan dan dukungan yang diberikan.

Penyusun menyadari bahwa skripsi ini masih jauh dari kata sempurna. Oleh karena itu, segala kritik dan saran yang membangun akan diterima dengan senang hati demi perbaikan di masa mendatang. Besar harapan semoga skripsi ini dapat memberikan manfaat bagi perkembangan ilmu pengetahuan, khususnya di bidang analisis sentimen dan machine learning.

Yogyakarta, 11 Agustus 2025

Penulis

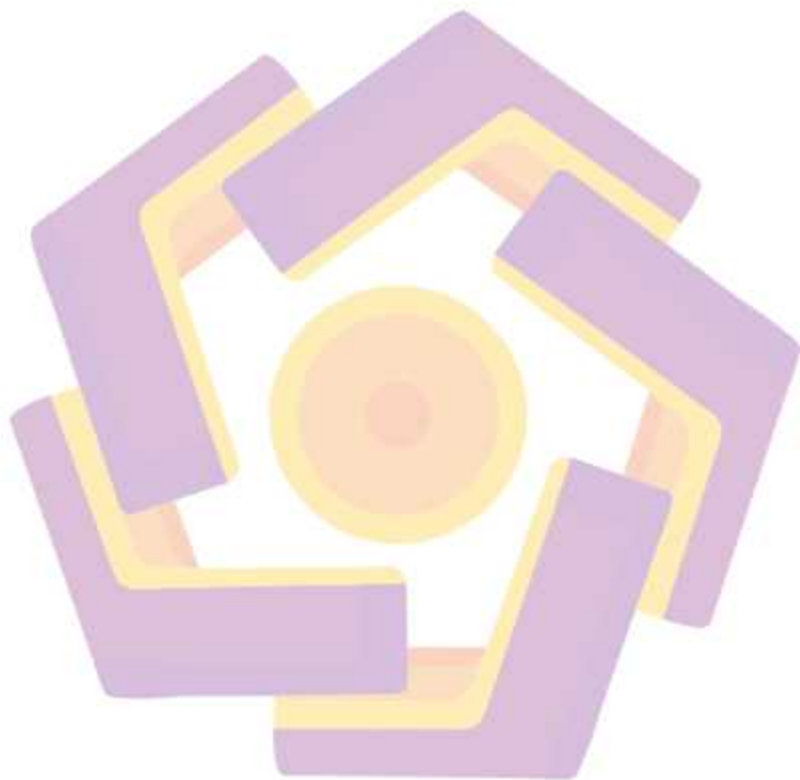


DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH	xiv
INTISARI	xvii
<i>ABSTRACT</i>	xviii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6
2.2 Dasar Teori	11

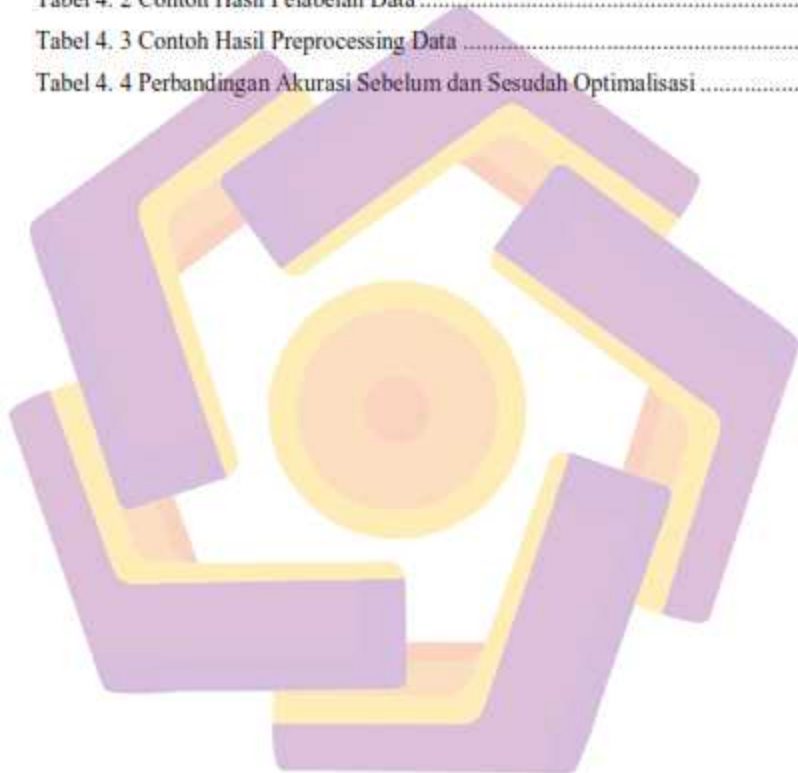
2.2.1 Analisis Sentimen.....	11
2.2.2 Teks Preprocessing.....	11
2.2.3 Representasi Fitur Teks (Vectorization).....	14
2.2.4 Machine Learning untuk Klasifikasi Teks.....	14
2.2.5 Algoritma Klasifikasi	14
2.2.6 Optimalisasi Hyperparameter (GridSearchCV)	17
2.2.7 Metrik Evaluasi Model.....	18
2.2.8 ChatGPT sebagai Aplikasi Chatbot AI.....	19
BAB III METODE PENELITIAN	21
3.1 Objek Penelitian.....	21
3.2 Alur Penelitian	21
3.3 Alat dan Bahan.....	23
3.3.1 Data Penelitian	23
3.3.2 Alat/Instrumen Penelitian	24
BAB IV HASIL DAN PEMBAHASAN	27
4.1 Deskripsi Data Penelitian	27
4.2 Hasil Preprocessing Data.....	28
4.3 Analisis Performa Model dan Pembahasan.....	30
4.3.1 Performa Support Vector Machine (SVM).....	30
4.3.2 Performa K-Nearest Neighbors (K-NN).....	32
4.3.3 Performa Naïve Bayes.....	34
4.3.4 Perbandingan Kinerja Antar Algoritma dan Pembahasan	36
4.4 Visualisasi Sentimen.....	37
BAB V PENUTUP	40
5.1 Kesimpulan.....	40

5.2	Saran.....	40
	REFERENSI.....	42
	LAMPIRAN.....	44



DAFTAR TABEL

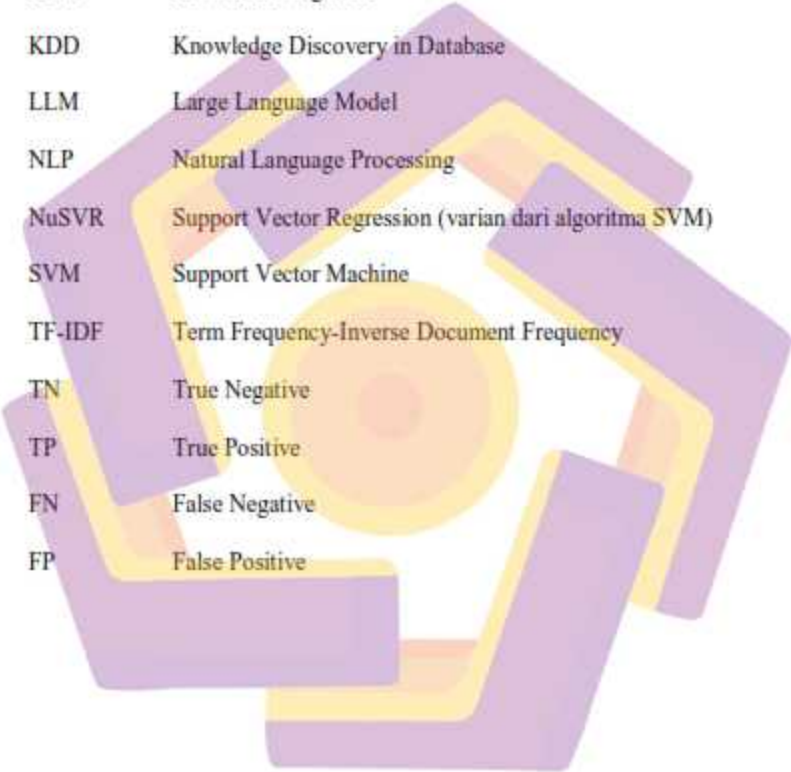
Tabel 2. 1 Keaslian Penelitian	9
Tabel 2. 2 Contoh Hasil Preprocessing Data	13
Tabel 4. 1 Contoh Hasil Scraping Data	27
Tabel 4. 2 Contoh Hasil Pelabelan Data	28
Tabel 4. 3 Contoh Hasil Preprocessing Data	29
Tabel 4. 4 Perbandingan Akurasi Sebelum dan Sesudah Optimalisasi	37



DAFTAR GAMBAR

Gambar 2. 1 Alur Proses Preprocessing Teks	11
Gambar 2. 2 Diagram Kerja Support Vector Machine (SVM)	15
Gambar 2. 3 Diagram Kerja K-Nearest Neighbors (K-NN)	16
Gambar 2. 4 Diagram Kerja Naïve Bayes.....	16
Gambar 2. 5 Diagram GridSearchCV	18
Gambar 2. 6 Contoh Confusion Matrix	19
Gambar 3. 1 Alur Penelitian.....	22
Gambar 4. 1 GridSearch CV SVM.....	31
Gambar 4. 2 Akurasi SVM Sebelum vs Sesudah Optimasi	31
Gambar 4. 3 Confusion Matrix SVM	32
Gambar 4. 4 GridSearch CV K-NN.....	33
Gambar 4. 5 Akurasi K-NN Sebelum vs Sesudah Optimasi.....	33
Gambar 4. 6 Confusion Matrix K-NN.....	34
Gambar 4. 7 GridSearch CV Naïve Bayes.....	35
Gambar 4. 8 Akurasi Naïve Bayes Sebelum vs Sesudah Optimasi.....	35
Gambar 4. 9 Confusin Matrix Naïve Bayes	36
Gambar 4. 10 Wordcloud Sentimen Negatif.....	38
Gambar 4. 11 Wordcloud Sentimen Positif	39

DAFTAR LAMBANG DAN SINGKATAN



AI	Artificial Intelligence (kecerdasan buatan)
CRISP-DM	Cross-Industry Standard Process for Data Mining
K-NN	K-Nearest Neighbors
KDD	Knowledge Discovery in Database
LLM	Large Language Model
NLP	Natural Language Processing
NuSVR	Support Vector Regression (varian dari algoritma SVM)
SVM	Support Vector Machine
TF-IDF	Term Frequency-Inverse Document Frequency
TN	True Negative
TP	True Positive
FN	False Negative
FP	False Positive

DAFTAR ISTILAH



Akurasi	Metrik evaluasi yang mengukur proporsi total prediksi yang benar dari semua prediksi yang dilakukan model.
Analisis Sentimen	Bidang dalam pemrosesan bahasa alami (NLP) yang berfokus pada penentuan sikap, opini, atau emosi dari sebuah teks.
Case Folding	Tahapan <i>preprocessing</i> teks yang mengubah seluruh huruf menjadi huruf kecil.
ChatGPT	Aplikasi <i>chatbot</i> kecerdasan buatan generatif berbasis <i>Large Language Model</i> (LLM) dari OpenAI.
Chatbot	Program komputer yang dirancang untuk mensimulasikan percakapan manusia, baik melalui teks maupun suara.
Confusion Matrix	Tabel yang menggambarkan kinerja model klasifikasi, menampilkan jumlah True Positives, False Positives, False Negatives, dan True Negatives.
F1-Score	<i>Harmonic mean</i> dari <i>precision</i> dan <i>recall</i> , berguna untuk mengevaluasi model dengan kelas yang tidak seimbang.
GridSearchCV	Metode untuk melakukan optimalisasi <i>hyperparameter</i> dengan menguji setiap kombinasi parameter yang telah ditentukan.
K-Nearest Neighbors	Algoritma klasifikasi yang bekerja dengan mencari 'K' sampel terdekat dari data pelatihan untuk menentukan kelas titik data baru.



Large Language Model	Model bahasa berbasis <i>deep learning</i> yang sangat besar, mampu memahami dan menghasilkan teks yang menyerupai manusia.
Machine Learning	Bidang dalam kecerdasan buatan yang memungkinkan sistem belajar dari data dan membuat prediksi tanpa diprogram secara eksplisit.
Naïve Bayes	Algoritma klasifikasi probabilistik yang didasarkan pada Teorema Bayes dengan asumsi independensi fitur yang kuat.
Normalisasi	Proses penyeragaman format teks atau perbaikan kesalahan penulisan/singkatan.
Optimalisasi Hyperparameter	Proses menemukan kombinasi parameter terbaik untuk model <i>machine learning</i> guna memaksimalkan performanya.
Penrosesan Bahasa Alami	Bidang ilmu yang menggabungkan linguistik, ilmu komputer, dan kecerdasan buatan untuk memungkinkan komputer memahami dan memproses bahasa manusia.
Precision	Metrik evaluasi yang mengukur proporsi prediksi positif yang sebenarnya positif.
Preprocessing Data	Serangkaian tahapan untuk mengubah data mentah menjadi format yang siap dianalisis oleh algoritma.
Recall	Metrik evaluasi yang mengukur proporsi <i>instance</i> positif yang berhasil diprediksi benar oleh model.
Scraping	Proses otomatis untuk mengekstrak data dari situs web atau platform <i>online</i> .
Stemming	Proses mereduksi kata-kata berlebihan menjadi bentuk kata dasarnya (kata akar).
Stopword Removal	Proses menghilangkan kata-kata umum yang tidak memiliki makna sentimen signifikan dari teks.

Support Vector Machine

Algoritma klasifikasi yang bekerja dengan mencari *hyperplane* terbaik yang memisahkan kelas-kelas data dalam ruang berdimensi tinggi.

TF-IDF

Teknik representasi fitur teks yang menghitung bobot suatu kata dalam dokumen berdasarkan frekuensi kemunculan dan jarang tidaknya kata tersebut.

Tokenisasi

Proses memecah kalimat atau paragraf menjadi unit-unit kata atau *token* individual.



INTISARI

Seiring pesatnya perkembangan teknologi kecerdasan buatan, aplikasi chatbot seperti ChatGPT telah menjadi bagian integral dari interaksi digital. Penelitian ini berfokus pada evaluasi respons pengguna terhadap aplikasi ChatGPT dengan menerapkan algoritma machine learning untuk analisis sentimen. Kami menguji tiga algoritma klasifikasi sentimen yang umum digunakan, yaitu Support Vector Machine (SVM), K-Nearest Neighbors (K-NN), dan Naïve Bayes, untuk mengategorikan ribuan ulasan pengguna dari Google Play Store ke dalam kelas sentimen positif dan negatif.

Proses diawali dengan pengumpulan data ulasan dan dilanjutkan dengan tahapan prapemrosesan data, termasuk normalisasi teks dan stemming, untuk memastikan kualitas data. Selanjutnya, optimasi hyperparameter dilakukan menggunakan GridSearchCV guna meningkatkan kinerja model klasifikasi. Hasil studi menunjukkan bahwa algoritma SVM, dengan konfigurasi hyperparameter yang telah dioptimalkan (misalnya, penggunaan kernel rbf dan nilai parameter C yang disesuaikan), mampu mencapai akurasi klasifikasi sentimen yang paling tinggi di antara ketiga algoritma yang diuji.

Temuan ini diharapkan dapat memberikan pemahaman mendalam mengenai persepsi dan pengalaman pengguna aplikasi ChatGPT, yang akan menjadi masukan berharga bagi pengembang dan penyedia layanan dalam meningkatkan fitur, kinerja, dan kepuasan pengguna di masa mendatang. Selain itu, penelitian ini membuka jalan bagi eksplorasi lebih lanjut dalam bidang analisis sentimen, khususnya untuk aplikasi berbasis kecerdasan buatan, guna menciptakan pengalaman pengguna yang lebih responsif dan relevan.

Kata kunci: Aplikasi ChatGPT, Analisis Sentimen, *Support Vector Machine*, *K-Nearest Neighbors*, Naïve Bayes.

ABSTRACT

In the rapidly evolving digital era, chatbot applications like ChatGPT have become an integral part of digital interaction. This research focuses on evaluating user responses to the ChatGPT application by applying machine learning algorithms for sentiment analysis. We tested three commonly used sentiment classification algorithms: Support Vector Machine (SVM), K-Nearest Neighbors (K-NN), and Naïve Bayes, to categorize thousands of user reviews from the Google Play Store into positive and negative sentiment classes.

The process began with collecting review data, followed by data preprocessing steps, including text normalization and stemming, to ensure data quality. Subsequently, hyperparameter optimization was conducted using GridSearchCV to enhance the performance of the classification models. The study results indicate that the SVM algorithm, with optimized hyperparameter configurations (e.g., specific rbf kernel usage and adjusted C parameter values), achieved the highest sentiment classification accuracy among the three algorithms tested.

These findings are expected to provide deep insights into the perceptions and experiences of ChatGPT application users, offering valuable input for developers and service providers to improve features, performance, and user satisfaction in the future. Furthermore, this research opens avenues for further exploration in sentiment analysis, particularly for AI-based applications, to create a more responsive and relevant user experience.

Keywords: ChatGPT Application, Sentiment Analysis, Support Vector Machine, K-Nearest Neighbors, Naïve Bayes.