



**ANALISIS SENTIMEN PENGGUNA CHATGPT BERDASARKAN
ULASAN PADA GOOGLE PLAY STORE MENGGUNAKAN
ALGORITMA *NAÏVE BAYES***

SKRIPSI

MUHAMMAD ZUL FIKAR

2010512091

PROGRAM STUDI S1 SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA

2024



**ANALISIS SENTIMEN PENGGUNA CHATGPT BERDASARKAN
ULASAN PADA GOOGLE PLAY STORE MENGGUNAKAN
ALGORITMA *NAÏVE BAYES***

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana
Komputer**

MUHAMMAD ZUL FIKAR

2010512091

PROGRAM STUDI S1 SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA

2024

PERNYATAAN ORISINALITAS

Tugas Akhir ini hasil karya sendiri dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Muhammad Zul Fikar

NIM : 2010512091

Program Studi : SI – Sistem Informasi

Judul : Analisis Sentimen Pengguna ChatGPT Berdasarkan Ulasan Pada Google

Play Store Menggunakan Algoritma *Naïve Bayes*

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan sayaini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 29 Juli 2024



(Muhammad Zul Fikar)

PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Pembangunan Nasional “Veteran” Jakarta, saya yang ber tanda tangan di bawah ini:

Nama : Muhammad Zul Fikar
NIM : 2010512091
Fakultas : Fakultas Ilmu Komputer
Program Studi : SI Sistem Informasi

Demi pembangunan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional “Veteran” Jakarta Hak Royalti Eksklusif (*non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

ANALISIS SENTIMEN PENGGUNA CHATGPT BERDASARKAN ULASAN PADA GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA *NAÏVE BAYES*

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional “Veteran” Jakarta berhak menyimpan, mengalih data/formatkan, Mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian Pernyataan ini saya buat dengan sesungguhnya.

Dibuat di : Jakarta
Pada Tanggal : 29/07/2024

Yang Menyatakan,



Muhammad Zul Fikar

LEMBAR PENGESAHAN

Skripsi ini diajukan oleh:

Nama : Muhammad Zul Fikar
NIM : 2010512091
Program Studi : S1 Sistem Informasi
Judul Tugas Akhir : Analisis Sentimen Pengguna ChatGPT Berdasarkan Ulasan Pada Google Play Store Menggunakan Algoritma Naïve Bayes

Telah Berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Ilmu Komputer pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jakarta.



Nurhafifah Matondang, S.Kom., M.M., M.T.I.

Dosen Pembimbing I



Sarika, S.Kom., M.Kom.

Dosen Pembimbing II



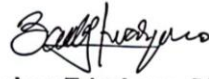
Iin Ernawati, S.Kom., M.Si.

Dosen Penguji I



Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM

Dekan Fakultas Ilmu Komputer UPN "Veteran"
Jakarta



Bambang Triwahyono, S.Kom., M.Si.

Dosen Penguji II



Anita Muliawati, S.Kom., M.T.I.

Ketua Program Studi S1 Sistem Informasi

Ditetapkan di : Jakarta
Tanggal Ujian : 27 Juni 2024

ANALISIS SENTIMEN PENGGUNA CHATGPT BERDASARKAN ULASAN PADA GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA *NAÏVE BAYES*

Muhammad Zul Fikar

ABSTRAK

Kecerdasan Buatan (*Artificial Intelligence* atau AI) telah mencapai kemajuan luar biasa dalam beberapa tahun terakhir. ChatGPT merupakan kecerdasan buatan yang dikembangkan oleh OpenAI yang dapat menghasilkan teks, menerjemahkan bahasa, menulis berbagai jenis konten kreatif, dan menjawab pertanyaan dengan cara yang informatif. Seiring dengan popularitas yang meningkat, muncul kebutuhan untuk memahami bagaimana pengguna merasakan dan mengalami aplikasi ini. Evaluasi kualitasnya dapat dilakukan melalui analisis sentimen yaitu sebuah teknik untuk mengidentifikasi dan memahami pendapat atau sentimen dari pengguna berdasarkan ulasan yang diberikan. Penelitian ini menggunakan metode *Naïve Bayes* untuk menganalisis sentimen pengguna terhadap aplikasi ChatGPT berdasarkan ulasan di Google Play Store. Data ulasan yang dikumpulkan dimulai dari Juli 2023 hingga Februari 2024 kemudian dilabeli secara manual oleh 3 anotator. Data akan melalui berbagai tahapan sebelum dilakukannya klasifikasi seperti tahap *preprocessing*, pembobotan kata dengan metode *Term Frequency – Inverse Document Frequency* (TF-IDF), dan pembagian data. Hasil penelitian menunjukkan bahwa model klasifikasi *Naïve Bayes* menghasilkan akurasi yang baik, dengan nilai akurasi mencapai 86%, presisi sebesar 82%, *recall* sebesar 98%, dan *f1-score* sebesar 89% dengan menggunakan rasio data latih dan uji 80:20. Selain hasil klasifikasi, penelitian juga membuat visualisasi data berupa *word cloud* untuk mengidentifikasi kata-kata kunci yang sering muncul dalam ulasan positif dan negatif. Luaran lainnya yang diperoleh sebuah sistem sederhana untuk memprediksi label berdasarkan data ulasan yang diberikan.

Kata Kunci : Analisis Sentimen, ChatGPT, *Naïve Bayes*, Kecerdasan Buatan

SENTIMENT ANALYSIS OF CHATGPT USERS BASED ON REVIEWS ON GOOGLE PLAY STORE USING THE NAÏVE BAYES ALGORITHM

Muhammad Zul Fikar

ABSTRACT

Artificial Intelligence (AI) has achieved remarkable advancements in recent years. ChatGPT, developed by OpenAI, is an AI that can generate text, translate languages, write various types of creative content, and answer questions in an informative manner. With its growing popularity, there is a need to understand how users perceive and experience this application. The quality evaluation can be conducted through sentiment analysis, a technique for identifying and understanding users' opinions or sentiments based on their reviews. This study employs the Naïve Bayes method to analyze user sentiment towards the ChatGPT application based on reviews from the Google Play Store. The review data collected ranges from July 2023 to February 2024 and is manually labeled by three annotators. The data undergoes various stages before classification, including preprocessing, word weighting using the Term Frequency – Inverse Document Frequency (TF-IDF) method, and data splitting. The results of the study show that the Naïve Bayes classification model produces good accuracy, with an accuracy rate of 86%, precision of 82%, recall of 98%, and an F1-score of 89% using an 80:20 training and test data ratio. In addition to the classification results, the study also creates data visualizations in the form of word clouds to identify key words frequently appearing in positive and negative reviews. Another outcome of the study is a simple system to predict labels based on the given review data.

Keywords : Sentiment Analysis, ChatGPT, Naïve Bayes, Artificial Intelligence

KATA PENGANTAR

Puji syukur dipanjatkan kehadirat Allah SWT karena atas berkat dan rahmat karunia-Nya penulis mampu menyelesaikan Skripsi yang berjudul “Analisis Sentimen Pengguna ChatGPT Berdasarkan Ulasan Pada Google Play Store Menggunakan Algoritma *Naïve Bayes*”. Dalam penyelesaian Skripsi ini tidak lepas dari bantuan berbagai pihak yang telah memberikan bantuan dan dukungan. Untuk itu peneliti mengucapkan banyak terima kasih kepada:

1. Allah SWT atas segala rahmat, ridha, dan karunia-Nya.
2. Kedua orang tua yang selalu memberikan dukungan baik secara moral maupun materiil.
3. Bapak Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM selaku Dekan Fakultas Ilmu Komputer.
4. Ibu Anita Muliawati, S.Kom., M.T.I. selaku Ketua Program Studi Sarjana Sistem Informasi.
5. Ibu Nurhafifah Matondang, S.Kom., M.M., M.T.I., selaku Dosen Pembimbing 1 yang sudah memberikan bimbingan, saran, dan motivasi dalam penyusunan penelitian.
6. Ibu Sarika, S.Kom., M.Kom., selaku Dosen Pembimbing 2 yang sudah memberikan bimbingan, saran, dan motivasi dalam penyusunan penelitian.
7. Seluruh kawan, saudara, dan teman-teman seperjuangan yang telah memberikan bantuan dan dukungan dengan sepenuh hati.
8. Serta seluruh pihak yang tidak bisa penulis sebutkan satu persatu tanpa mengurangi rasa hormat.

Peneliti menyadari masih banyaknya kekurangan dari Skripsi ini, baik dari materi maupun teknik penulisan, mengingat kurangnya pengetahuan dan pengalaman peneliti. Oleh karena itu, kritik dan saran yang membangun akan sangat berharga bagi peneliti.

Jakarta, 20 Juni 2024

Peneliti

DAFTAR ISI

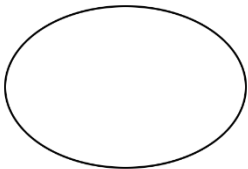


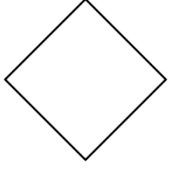


PERNYATAAN ORISINALITAS.....	iii
LEMBAR PERSETUJUAN.....	iv
LEMBAR PENGESAHAN.....	v
ABSTRAK	vi
<i>ABSTRACT</i>	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	ix
DAFTAR SIMBOL	xiii
DAFTAR GAMBAR	xiv
DAFTAR TABEL.....	xv
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Tujuan Penelitian.....	3
1.4. Manfaat Penelitian.....	4
1.5. Ruang Lingkup	4
1.6. Luaran yang Diharapkan	5
1.7. Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA	7
2.1. Analisis Sentimen	7
2.2. <i>Text Mining</i>	7
2.3. ChatGPT	8
2.4. Pelabelan Data	8
2.5. <i>Preprocessing Data</i>	9
2.5.1. <i>Case Folding</i>	9
2.5.2. <i>Cleaning</i>	9

2.5.3. <i>Normalization</i>	10
2.5.4. <i>Tokenizing</i>	10
2.5.5. <i>Stopword Removal</i>	10
2.5.6 <i>Stemming</i>	10
2.6. Pembobotan Kata	11
2.7. <i>Naïve Bayes</i>	12
2.8. Python.....	14
2.9. <i>Confusion Matrix</i>	14
2.10. Penelitian Terdahulu	16
BAB III METODOLOGI PENELITIAN	21
3.1. Tahapan Penelitian.....	21
3.1.1. Identifikasi Masalah	21
3.1.2. Studi Literatur	22
3.1.3. Pengumpulan Data	22
3.1.4. Pelabelan Data.....	23
3.1.5. <i>Preprocessing Data</i>	23
3.1.6. Pembobotan Kata	27
3.1.7. Klasifikasi <i>Naïve Bayes</i>	28
3.1.8. Hasil Evaluasi.....	28
3.1.9. Pengembangan Sistem	28
3.1.10. Pengujian Sistem.....	29
3.2. Perangkat Penelitian	29
3.2.1. Perangkat Keras	29
3.2.2. Perangkat Lunak.....	30
3.3. Jadwal Penelitian	30
BAB IV HASIL DAN PEMBAHASAN	32

4.1.	Pengumpulan Data	32
4.2.	Pelabelan Data	32
4.3.	<i>Preprocessing Data</i>	33
4.3.1.	<i>Case Folding</i>	34
4.3.2.	<i>Cleaning</i>	35
4.3.3.	<i>Normalization</i>	36
4.3.4.	<i>Tokenizing</i>	38
4.3.5.	<i>Stopwords Removal</i>	39
4.3.6.	<i>Stemming</i>	41
4.4.	Pembobotan Kata TF-IDF	41
4.5.	Klasifikasi <i>Naïve Bayes</i>	42
4.6.	Hasil Evaluasi	43
4.7.	Analisis Hasil Klasifikasi dan Visualisasi	44
4.7.1.	Distribusi Sentimen	45
4.7.2.	Visualisasi <i>Word Cloud</i> Sentimen Positif	45
4.7.3.	Visualisasi <i>Word Cloud</i> Sentimen Negatif	46
4.7.4.	Frekuensi Kata Terbanyak	47
4.8.	Implementasi Pengembangan Sistem	49
4.8.1.	Tampilan Halaman <i>Landing Page</i>	49
4.8.2.	Tampilan Halaman <i>Dashboard</i>	50
4.8.3.	Tampilan Halaman <i>Sentimen Prediction</i>	51
4.9.	Pengujian Sistem	51
BAB V PENUTUP		55
5.1.	Kesimpulan	55
5.2.	Saran	56
DAFTAR PUSTAKA		57

LAMPIRAN.....	62
---------------	----

DAFTAR SIMBOL

No.	Simbol <i>Flowchart</i>	Nama	Deskripsi
1.		<i>Terminator</i>	Titik mulai atau titik awal dari alur proses yang akan dijelaskan.
2.		<i>Process</i>	Merepresentasikan tahapan atau aktivitas yang dilakukan dalam alur proses.
3.		<i>Parallelogram</i>	Menyatakan proses <i>input/ouput</i> tanpa tergantung jenis peralatan.
4.		<i>Decision</i>	Simbol yang mewakili pengambilan keputusan untuk langkah selanjutnya pada alur program.
5.		<i>Flow Line</i>	Penghubung antara dua segmen diagram yang terpisah secara fisik, namun tetap terhubung dalam alur proses atau informasi yang sedang dijelaskan.
6.		<i>Dotted Line</i>	Menunjukkan bahwa hubungan antara dua elemen tersebut bersifat pilihan (opsional), tidak langsung, atau bergantung pada situasi khusus.

DAFTAR GAMBAR

Gambar 3.1 Tahapan Penelitian	21
Gambar 4.1 Hasil <i>Scraping Data</i>	32
Gambar 4.2 Proses Klasifikasi <i>Naive Bayes</i>	43
Gambar 4.3 Diagram Pie Distribusi Sentimen.....	45
Gambar 4.4 <i>Word Cloud</i> Sentimen Positif.....	46
Gambar 4.5 <i>Word Cloud</i> Sentimen Negatif.....	47
Gambar 4.6 Frekuensi Kata Terbanyak.....	48
Gambar 4.7 Halaman <i>Landing Page</i>	49
Gambar 4.8 Halaman <i>Dashboard</i> Bagian 1	50
Gambar 4.9 Halaman <i>Dashboard</i> Bagian 2	50
Gambar 4.10 Halaman <i>Dashboard</i> Bagian 3	50
Gambar 4.11 Halaman Prediksi Sentimen	51

DAFTAR TABEL

Tabel 2.1 Tabel <i>Confusion Matrix</i>	15
Tabel 2.2 Penelitian Terdahulu.....	16
Tabel 3.1 Contoh Proses <i>Case Folding</i>	24
Tabel 3.2 Contoh Proses <i>Cleaning</i>	24
Tabel 3.3 Contoh Proses <i>Normalization</i>	25
Tabel 3.4 Contoh Proses <i>Tokenizing</i>	26
Tabel 3.5 Contoh Proses <i>Stopword Removal</i>	26
Tabel 3.6 Contoh Proses <i>Stemming</i>	27
Tabel 3.7 Jadwal Penelitian.....	30
Tabel 4.1 Tahap Pelabelan Data oleh Anotator	33
Tabel 4.2 Proses <i>Case Folding</i>	34
Tabel 4.3 Proses <i>Cleaning</i>	35
Tabel 4.4 Proses <i>Normalization</i>	37
Tabel 4.5 Proses <i>Tokenizing</i>	38
Tabel 4.6 Proses <i>Stopword Removal</i>	39
Tabel 4.7 Proses <i>Stemming</i>	41
Tabel 4.8 Contoh 10 Kata Teratas Hasil Pembobotan TF-IDF	42
Tabel 4.9 Hasil <i>Confusion Matrix</i>	43
Tabel 4.10 Hasil Pengujian <i>Black box Testing</i>	52