Analisis Sentimen Pengguna Aplikasi ChatGPT Berdasarkan Ulasan Google Play Store Menggunakan Metode TF-IDF dan Support Vector Machine (SVM)

Rangga Yudha Syahbana[1*], Wijiyanto[2], Agustina Srirahayu[3]

[1]Teknik Informatika/Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta
[1*]20011203@mhs.udb.ac.id

[2] Teknik Informatika/Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta
[2]wijiyanto@udb.ac.id

[3] Teknik Informatika/Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta
[3]agustina@udb.ac.id

Abstrak*— Perkembangan teknologi kecerdasan buatan generatif seperti ChatGPT telah meningkatkan penggunaan aplikasi berbasis *Artificial Intelligence* (AI) dalam berbagai bidang, seperti pendidikan, pekerjaan, dan pencarian informasi. Meningkatnya jumlah pengguna menyebabkan munculnya berbagai ulasan yang berisi opini, pengalaman, serta tingkat kepuasan pengguna terhadap aplikasi tersebut. Oleh karena itu, diperlukan analisis sentimen untuk mengetahui kecenderungan opini pengguna berdasarkan ulasan yang diberikan. Penelitian ini bertujuan untuk membangun sistem analisis sentimen terhadap ulasan pengguna aplikasi ChatGPT menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) sebagai metode ekstraksi fitur dan *Support Vector Machine* (SVM) sebagai metode klasifikasi. Dataset yang digunakan berupa 5.000 ulasan pengguna aplikasi ChatGPT yang diperoleh dari Google Play Store. Tahapan penelitian meliputi pengumpulan dataset, pelabelan sentimen berdasarkan rating pengguna, *preprocessing* teks, pembobotan fitur menggunakan TF-IDF, klasifikasi menggunakan algoritma SVM, serta evaluasi model menggunakan *Accuracy*, *Precision*, *Recall*, dan F1-*Score*. Hasil pengujian menunjukkan bahwa kombinasi metode TF-IDF dan SVM menghasilkan nilai *Accuracy* sebesar 90,10%, *Precision* sebesar 87,19%, *Recall* sebesar 90,10%, dan *F1-Score* sebesar 88,38%. Sistem yang dikembangkan berhasil mengklasifikasikan ulasan pengguna ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral, sehingga dapat membantu memahami persepsi pengguna terhadap aplikasi ChatGPT berdasarkan ulasan yang diberikan.
Kata kunci— Analisis Sentimen, ChatGPT, TF-IDF, *Support Vector Machine*, *Google Play Store*.

Abstract*— *The development of generative artificial intelligence technology such as ChatGPT has increased the adoption of Artificial Intelligence (AI)-based applications in various fields, including education, work, and information retrieval. The increasing number of users has generated various reviews containing opinions, experiences, and satisfaction levels toward the application. Therefore, sentiment analysis is needed to identify user opinions based on provided reviews. This study aims to develop a sentiment analysis system for ChatGPT user reviews using Term Frequency-Inverse Document Frequency (TF-IDF) as a feature extraction method and Support Vector Machine (SVM) as a classification algorithm. The dataset used consists of 5,000 ChatGPT user reviews collected from the Google Play Store. The research stages include dataset collection, sentiment labeling based on user ratings, text preprocessing, feature extraction using TF-IDF, sentiment classification using SVM, and model evaluation using Accuracy, Precision, Recall, and F1-Score. The experimental results show that the combination of TF-IDF and SVM achieved an Accuracy of 90.10%, Precision of 87.19%, Recall of 90.10%, and an F1-Score of 88.38%. The developed system successfully classifies user reviews into three sentiment categories: positive, negative, and neutral, providing useful insights into user perceptions of the ChatGPT application based on user reviews.
Keywords— *Sentiment Analysis, ChatGPT, TF-IDF, Support Vector Machine, Google Play Store*

I. PENDAHULUAN

Perkembangan teknologi kecerdasan buatan (*Artificial Intelligence*/AI), khususnya *generative artificial intelligence*, telah mengalami perkembangan yang pesat dalam beberapa tahun terakhir. Salah satu implementasi teknologi tersebut adalah ChatGPT yang dikembangkan menggunakan model bahasa besar (*Large Language Model*/LLM) dan mampu memahami bahasa alami serta menghasilkan respons berbasis teks sesuai konteks pengguna. Kemampuan tersebut menyebabkan ChatGPT banyak digunakan dalam berbagai bidang seperti pendidikan, pekerjaan, layanan informasi, dan pembuatan konten digital. Wang et al. menjelaskan bahwa model bahasa besar seperti ChatGPT memiliki kemampuan dalam memahami konteks bahasa dan melakukan berbagai tugas berbasis teks, termasuk analisis sentimen, meskipun masih memiliki keterbatasan tertentu dalam memahami opini yang kompleks [1]. Penelitian terbaru mengenai ulasan pengguna ChatGPT menunjukkan bahwa pendekatan *machine learning* dapat digunakan untuk

menganalisis opini pengguna terhadap layanan berbasis kecerdasan buatan, sehingga dapat memberikan gambaran mengenai tingkat kepuasan dan pengalaman pengguna terhadap aplikasi tersebut [13].

Perkembangan ChatGPT juga menunjukkan bahwa teknologi AI generatif telah memberikan peluang baru dalam berbagai bidang, namun masih terdapat tantangan seperti akurasi informasi, interpretasi pengguna, dan penerimaan terhadap teknologi tersebut. Aljanabi et al. menjelaskan bahwa penggunaan ChatGPT membawa peluang sekaligus tantangan yang perlu dievaluasi berdasarkan pengalaman dan persepsi pengguna [14].

Meningkatnya jumlah pengguna aplikasi ChatGPT menyebabkan bertambahnya jumlah ulasan yang diberikan pengguna melalui platform distribusi aplikasi seperti *Google Play Store*. Ulasan pengguna mengandung informasi mengenai pengalaman penggunaan, tingkat kepuasan, serta permasalahan yang ditemukan selama menggunakan aplikasi. Data ulasan tersebut dapat dimanfaatkan sebagai sumber informasi untuk mengetahui persepsi pengguna terhadap suatu layanan digital. Hu dan Liu menjelaskan bahwa ulasan pelanggan (*customer review*) dapat dianalisis untuk memperoleh informasi opini pengguna melalui pendekatan pengolahan teks secara otomatis [8].

Namun, jumlah ulasan yang besar menyebabkan proses analisis manual menjadi kurang efektif dan membutuhkan waktu yang lama. Oleh karena itu, diperlukan metode otomatis untuk mengidentifikasi kecenderungan opini pengguna melalui analisis sentimen. Analisis sentimen merupakan bidang dalam *Natural Language Processing* (NLP) yang bertujuan untuk mengklasifikasikan teks berdasarkan polaritas opini menjadi kategori tertentu seperti positif, negatif, dan netral. Pang et al. menunjukkan bahwa metode *machine learning* dapat digunakan secara efektif untuk melakukan klasifikasi sentimen berbasis teks [2]. Selain itu, Pak dan Paroubek menunjukkan bahwa data dari platform digital dapat digunakan sebagai sumber analisis sentimen untuk mengetahui kecenderungan opini pengguna [7].

Pada analisis sentimen, data teks perlu melalui tahap preprocessing untuk menghasilkan representasi data yang lebih bersih sehingga memudahkan algoritma klasifikasi dalam mengenali pola. Tahapan seperti *case folding*, tokenisasi, penghapusan kata tidak penting (*stopword removal*), dan stemming digunakan untuk mengurangi variasi data yang tidak diperlukan. Manning et al. menjelaskan bahwa *preprocessing* merupakan bagian penting dalam pengolahan dokumen digital karena berpengaruh terhadap kualitas representasi teks yang dihasilkan [4].

Setelah proses *preprocessing* dilakukan, data teks perlu dikonversi menjadi bentuk numerik. Salah satu metode ekstraksi fitur yang banyak digunakan adalah *Term Frequency-Inverse Document Frequency* (TF-IDF). Metode TF-IDF memberikan bobot terhadap kata berdasarkan tingkat kemunculan kata dalam dokumen dan tingkat kepentingannya terhadap keseluruhan dokumen. Salton dan Buckley menjelaskan bahwa pendekatan pembobotan berbasis TF-IDF efektif digunakan dalam sistem pengolahan informasi dan klasifikasi dokumen [3].

Pada tahap klasifikasi, penelitian ini menggunakan algoritma *Support Vector Machine* (SVM). SVM merupakan metode *machine learning* yang mampu menangani data berdimensi tinggi dan banyak diterapkan pada klasifikasi teks. Joachims menunjukkan bahwa SVM memiliki kemampuan yang baik dalam melakukan klasifikasi dokumen dengan jumlah fitur yang besar [5]. Selain itu, penelitian mengenai analisis sentimen berbasis data digital menunjukkan bahwa kombinasi teknik ekstraksi fitur dengan algoritma machine learning masih banyak digunakan karena mampu memberikan performa yang baik dalam klasifikasi opini pengguna [15]. Sebastiani juga menjelaskan bahwa metode *machine learning* telah banyak digunakan dalam otomatisasi klasifikasi teks karena mampu melakukan pembelajaran berdasarkan pola pada data dokumen [6].

Selain pemilihan metode klasifikasi, evaluasi performa model menjadi bagian penting untuk mengetahui tingkat keberhasilan sistem. Fawcett menjelaskan bahwa evaluasi model klasifikasi

dapat dilakukan menggunakan berbagai metrik pengukuran untuk mengetahui kemampuan model dalam membedakan kelas data [9]. Sokolova dan Lapalme menyatakan bahwa penggunaan metrik seperti *precision*, *recall*, dan *F1-score* diperlukan untuk memberikan gambaran performa model secara lebih menyeluruh terutama pada permasalahan klasifikasi multi-kelas [10].

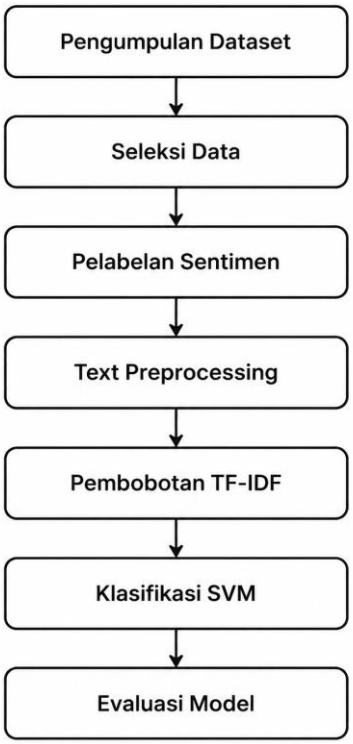
Dalam penelitian analisis sentimen, permasalahan ketidakseimbangan jumlah data pada setiap kelas (*imbalanced dataset*) sering ditemukan. Kondisi tersebut dapat menyebabkan model lebih dominan mengenali kelas mayoritas dibandingkan kelas minoritas. Chawla et al. memperkenalkan metode SMOTE (*Synthetic Minority Over-sampling Technique*) sebagai salah satu pendekatan untuk menangani permasalahan ketidakseimbangan data pada proses klasifikasi [11].

Selain pendekatan berbasis *machine learning*, penelitian mengenai klasifikasi teks juga berkembang menggunakan metode *deep learning*. Kim menunjukkan bahwa pendekatan *Convolutional Neural Network* (CNN) mampu digunakan untuk melakukan klasifikasi teks dengan memanfaatkan representasi fitur otomatis dari data kalimat [12].

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk membangun sistem analisis sentimen ulasan pengguna aplikasi ChatGPT menggunakan metode TF-IDF dan Support *Vector Machine*. Sistem yang dikembangkan menggunakan dataset ulasan pengguna aplikasi ChatGPT dari *Google Play Store* dan melakukan klasifikasi sentimen menjadi tiga kategori, yaitu positif, negatif, dan netral. Hasil penelitian diharapkan dapat memberikan informasi mengenai persepsi pengguna terhadap aplikasi ChatGPT berdasarkan ulasan yang tersedia.

II. METODOLOGI PENELITIAN

Metodologi penelitian ini dilakukan melalui beberapa tahapan, yaitu pengumpulan dataset, preprocessing teks, pelabelan sentimen, ekstraksi fitur menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF), klasifikasi menggunakan *Support Vector Machine* (SVM), serta evaluasi performa model. Alur penelitian dapat dilihat pada Gambar 1.



Gambar 1. Alur Metodologi Penelitian

A. Pengumpulan Dataset

Dataset yang digunakan pada penelitian ini berupa ulasan pengguna aplikasi ChatGPT yang diperoleh dari *platform Google Play Store*. Dataset disimpan dalam format CSV dengan nama dataset-gpt.csv yang terdiri dari 11 atribut, seperti reviewId, userName, content, score, dan atribut pendukung lainnya.

Pada penelitian ini, atribut yang digunakan adalah kolom *content* sebagai data teks ulasan pengguna dan kolom *score* sebagai nilai rating pengguna. Dataset yang digunakan berjumlah 5.000 data ulasan**.** Pengambilan data dilakukan melalui teknik scraping pada Google Play Store menggunakan library Google Play Scraper pada bulan Juli 2026. Dataset yang diperoleh terdiri atas ulasan berbahasa Indonesia dan bahasa Inggris. Sebelum digunakan pada tahap analisis, data diperiksa untuk menghapus ulasan duplikat serta entri yang kosong sehingga diperoleh sebanyak 5.000 ulasan yang layak digunakan dalam penelitian. Data tersebut digunakan untuk

mengetahui kecenderungan sentimen pengguna terhadap aplikasi ChatGPT berdasarkan pengalaman penggunaan yang diberikan melalui ulasan.

B. Pelabelan Sentimen

Tahap pelabelan dilakukan untuk menentukan kategori sentimen pada setiap ulasan pengguna. Proses pelabelan dilakukan berdasarkan nilai rating (*score*) yang diberikan pengguna pada *Google Play Store*.

Aturan pelabelan yang digunakan adalah:

Tabel 1 Pelabelan Sentimen

No	Rating	Kategori Sentimen
1	4-5	Positif
2	3	Netral
3	1-2	Negatif

Pendekatan berbasis rating banyak digunakan dalam analisis sentimen ulasan aplikasi karena rating pengguna dapat digunakan sebagai representasi awal dari opini pengguna terhadap suatu produk atau layanan.

C. *Text Preprocessing*

Tahap *preprocessing* dilakukan untuk membersihkan dan mempersiapkan data teks sebelum dilakukan proses klasifikasi. Tahapan ini bertujuan untuk mengurangi *noise* pada data sehingga fitur yang dihasilkan dapat merepresentasikan informasi penting dari suatu dokumen.

Tahapan *preprocessing* yang digunakan meliputi:

1. *Cleaning*

Proses *cleaning* dilakukan dengan menghapus karakter yang tidak diperlukan seperti URL, simbol, angka, tanda baca, dan karakter khusus.

2. *Case Folding*

Case folding digunakan untuk mengubah seluruh teks menjadi huruf kecil (*lowercase*) sehingga kata dengan bentuk huruf berbeda dianggap sama.

3. *Tokenizing*

Tokenizing merupakan proses memecah teks menjadi bagian-bagian kata (*token*) yang akan digunakan pada proses berikutnya.

4. *Stopword Removal*

Tahap ini bertujuan menghapus kata umum yang tidak memiliki pengaruh besar terhadap proses klasifikasi.

5. *Stemming*

Stemming dilakukan untuk mengubah kata menjadi bentuk dasar sehingga variasi kata yang memiliki makna sama dapat direpresentasikan sebagai satu bentuk kata.

Tahapan *preprocessing* merupakan bagian penting dalam pengolahan teks karena kualitas data yang digunakan dapat mempengaruhi performa model klasifikasi [4].

D. Ekstraksi Fitur Menggunakan TF-IDF

Setelah proses *preprocessing* selesai dilakukan, data teks dikonversi menjadi representasi numerik menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF).

TF-IDF merupakan metode pembobotan kata yang memberikan nilai berdasarkan frekuensi kemunculan suatu kata dalam dokumen serta tingkat kepentingan kata tersebut pada keseluruhan dokumen. Metode ini digunakan karena mampu mengurangi pengaruh kata yang sering muncul namun memiliki informasi rendah.

Perhitungan TF-IDF dilakukan berdasarkan persamaan berikut:

$$\mathrm{TFIDF}(t,d) = \mathrm{TF}(t,d) \times \mathrm{IDF}(t)$$

E. Klasifikasi Menggunakan *Support Vector Machine* (SVM)

Tahap klasifikasi dilakukan menggunakan algoritma *Support Vector Machine* (SVM). SVM merupakan algoritma *machine learning* yang bekerja dengan mencari batas pemisah (*hyperplane*) terbaik untuk membedakan kelas data.

Pada penelitian ini, SVM digunakan untuk mengklasifikasikan ulasan pengguna ke dalam tiga kelas sentimen, yaitu positif, negatif, dan netral.

F. Evaluasi Model

Evaluasi dilakukan untuk mengetahui performa model SVM dalam melakukan klasifikasi sentimen. Pengujian dilakukan menggunakan *Confusion Matrix* dengan parameter evaluasi:

1. *Accuracy*

Mengukur persentase keseluruhan prediksi yang benar dibandingkan dengan jumlah seluruh data pengujian.

2. *Precision*

Mengukur tingkat ketepatan model dalam memprediksi suatu kelas sentimen.

3. *Recall*

Mengukur kemampuan model dalam menemukan data yang benar pada suatu kelas.

4. F1-*Score*

Merupakan nilai gabungan antara precision dan recall yang digunakan untuk mengetahui keseimbangan performa model.

III. HASIL DAN PEMBAHASAN

A. Hasil Pengumpulan Dataset

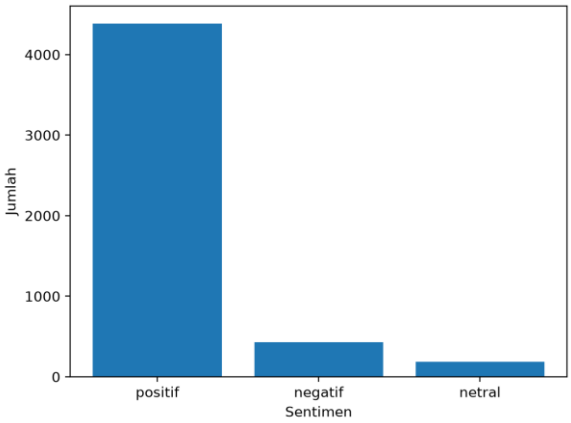
Dataset yang digunakan pada penelitian ini merupakan ulasan pengguna aplikasi ChatGPT yang diperoleh dari *Google Play Store*. Dataset disimpan dalam format CSV dengan nama dataset-gpt.csv. Dataset tersebut terdiri dari 11 atribut, dengan atribut utama yang digunakan dalam penelitian ini yaitu *content* sebagai teks ulasan pengguna dan *score* sebagai nilai rating pengguna.

Berdasarkan hasil pengolahan dataset, diperoleh sebanyak 5.000 data ulasan yang digunakan dalam proses analisis sentimen. Data tersebut kemudian melalui proses seleksi dan pembersihan sebelum dilakukan tahap klasifikasi. Dataset dibagi menjadi dua bagian menggunakan metode *train-test split*, yaitu sebanyak 4.000 data (80%) digunakan sebagai data latih dan 1.000 data (20%) digunakan sebagai data uji.

B. Distribusi Sentimen Dataset

Tahap awal analisis dilakukan dengan melakukan pelabelan sentimen berdasarkan nilai rating (*score*) yang diberikan pengguna. Rating 4-5 dikategorikan sebagai sentimen positif, rating 3 sebagai sentimen netral, dan rating 1-2 sebagai sentimen negatif.

Berdasarkan hasil pelabelan, distribusi kelas sentimen menunjukkan bahwa data didominasi oleh sentimen positif. Hal tersebut menunjukkan bahwa sebagian besar pengguna memberikan pengalaman penggunaan yang baik terhadap aplikasi *ChatGPT*.



Gambar 2. Distribusi Sentimen Dataset

Berdasarkan grafik distribusi sentimen, jumlah ulasan positif memiliki jumlah paling tinggi dibandingkan dengan kategori lainnya. Sementara itu, jumlah ulasan negatif dan netral memiliki jumlah yang lebih sedikit. Kondisi tersebut menunjukkan bahwa dataset mengalami ketidakseimbangan kelas (*imbalanced dataset*), dimana jumlah data pada setiap kategori sentimen tidak memiliki distribusi yang sama.

Ketidakseimbangan data dapat mempengaruhi performa model karena algoritma klasifikasi cenderung lebih mudah mengenali kelas dengan jumlah data yang lebih besar. Hal tersebut terlihat pada hasil evaluasi model dimana kategori positif memiliki performa lebih baik dibandingkan kategori negatif dan netral.

C. Hasil *Text Preprocessing*

Sebelum dilakukan proses klasifikasi, data ulasan terlebih dahulu dilakukan *preprocessing* untuk menghilangkan *noise* dan meningkatkan kualitas representasi teks. Tahapan preprocessing yang digunakan meliputi *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*.

Contoh hasil preprocessing dapat dilihat pada tabel berikut:

Tabel 2 Hasil *Text Preprocessing*

No	Data Awal	Hasil *Preprocessing*
1	DI UPDATE BUKANNYA BAGUS MALAH MAKIN PARAH. EMANG DEVELOPER TUTUP MATA DENGAN KRITIKAN	di update bukan bagus makin parah emang developer tutup mata kritik
2	bagus dan bermanfaat 👍👍	bagus manfaat
3	very helpful and easy to use	*helpful easy use*

Berdasarkan hasil *preprocessing*, data teks menjadi lebih sederhana dengan menghilangkan karakter yang tidak diperlukan. Proses ini membantu mengurangi variasi kata yang dapat mengganggu proses pembelajaran model sehingga fitur yang dihasilkan lebih relevan untuk klasifikasi.

D. Hasil Ekstraksi Fitur Menggunakan TF-IDF

Setelah proses *preprocessing* selesai dilakukan, data teks dikonversi menjadi bentuk numerik menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF).

Pada penelitian ini, metode TF-IDF menghasilkan sebanyak 5.000 fitur yang digunakan sebagai representasi data teks sebelum masuk ke tahap klasifikasi menggunakan *Support Vector Machine.*

Jumlah fitur tersebut menunjukkan bahwa proses ekstraksi berhasil mengubah data teks menjadi bentuk vektor numerik berdasarkan tingkat kepentingan setiap kata pada dokumen. Kata yang memiliki frekuensi tinggi tetapi muncul pada banyak dokumen akan memiliki bobot lebih rendah, sedangkan kata yang lebih spesifik akan memiliki bobot lebih tinggi.

E. Hasil Klasifikasi Menggunakan *Support Vector Machine*

Setelah proses pembobotan TF-IDF dilakukan, data kemudian digunakan sebagai input pada algoritma *Support Vector Machine* (SVM). Model dilatih menggunakan 4.000 data latih dan diuji menggunakan 1.000 data uji.

Hasil evaluasi model dapat dilihat pada tabel berikut:

Tabel 3 Hasil Evaluasi SVM

No	Parameter Evaluasi	Nilai
1	Accuracy	90,10%
2	Precision	87,19%
3	Recall	90,10%
4	F1-Score	88,38%

Berdasarkan hasil evaluasi tersebut, model SVM dengan ekstraksi fitur TF-IDF mampu menghasilkan tingkat akurasi sebesar 90,10%**.** Nilai tersebut menunjukkan bahwa model mampu melakukan klasifikasi sentimen dengan baik pada data ulasan pengguna aplikasi *ChatGPT*.

F. Analisis *Classification Report*

Hasil klasifikasi berdasarkan masing-masing kelas sentimen ditampilkan pada tabel berikut:

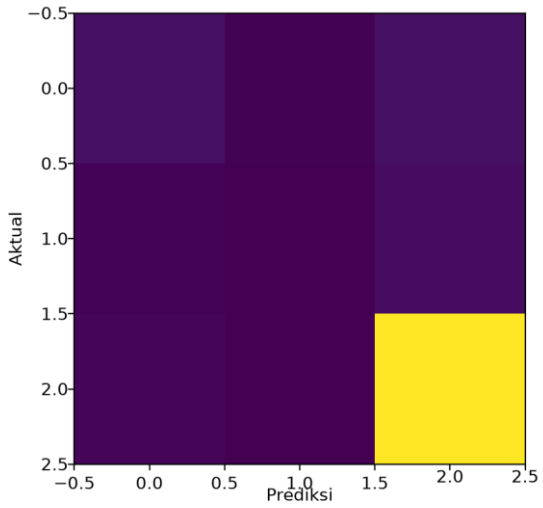
Tabel 4 Hasil Evaluasi SVM

No	Sentimen	*Precision*	*Recall*	F1-*Score*	*Support*
1	Negatif	0,64	0,48	0,55	86
2	Netral	0,11	0,03	0,04	37
3	Positif	0,93	0,98	0,95	877

Berdasarkan hasil tersebut, kelas positif memperoleh performa terbaik dengan nilai F1-*Score* sebesar 0,95. Hal ini disebabkan oleh jumlah data positif yang jauh lebih banyak dibandingkan kelas lainnya sehingga model lebih banyak mempelajari pola dari kelas tersebut.

Sebaliknya, kelas netral memperoleh performa paling rendah dengan nilai F1-Score sebesar 0,04. Rendahnya performa tersebut dipengaruhi oleh jumlah data netral yang jauh lebih sedikit dibandingkan kelas positif sehingga model kesulitan mempelajari karakteristik sentimen netral. Selain itu, ulasan dengan sentimen netral sering kali memiliki karakteristik bahasa yang berada di antara sentimen positif dan negatif sehingga batas antar kelas menjadi kurang jelas. Pada penelitian selanjutnya, performa kelas netral dapat ditingkatkan melalui penerapan metode penyeimbangan data seperti SMOTE, penambahan jumlah data netral, serta penggunaan metode klasifikasi berbasis transformer seperti IndoBERT atau BERT yang memiliki kemampuan lebih baik dalam memahami konteks kalimat [16].

G. *Confusion Matrix*



Gambar 3. *Confusion Matrix* Model SVM

Confusion Matrix digunakan untuk melihat distribusi hasil prediksi model terhadap kelas sebenarnya. Berdasarkan hasil *confusion matrix*, model mampu mengenali sebagian besar data positif dengan baik. Namun, masih terdapat kesalahan klasifikasi terutama pada kelas negatif dan netral.

Kesalahan klasifikasi tersebut dipengaruhi oleh ketidakseimbangan jumlah data pada setiap kelas. Untuk penelitian lanjutan, peningkatan performa pada kelas minoritas dapat dilakukan dengan menerapkan metode penyeimbangan data seperti *Synthetic Minority Oversampling Technique* (SMOTE) atau menggunakan pendekatan *deep learning* berbasis transformer

H. Pembahasan

Berdasarkan hasil pengujian, kombinasi metode TF-IDF dan Support Vector Machine mampu memberikan performa yang baik untuk melakukan analisis sentimen ulasan pengguna aplikasi *ChatGPT*. Nilai akurasi sebesar 90,10% menunjukkan bahwa metode tersebut efektif dalam melakukan klasifikasi teks berbasis ulasan pengguna.

Namun, hasil penelitian juga menunjukkan bahwa distribusi data memiliki pengaruh besar terhadap performa model. Model memiliki kemampuan lebih baik dalam mengenali sentimen positif karena jumlah datanya dominan, sedangkan kelas negatif dan netral masih memiliki performa yang lebih rendah.

Dengan demikian, sistem yang dikembangkan telah mampu melakukan proses analisis sentimen secara otomatis mulai dari pengolahan dataset, *preprocessing*, ekstraksi fitur menggunakan TF-IDF, klasifikasi menggunakan SVM, hingga evaluasi performa model.

IV. KESIMPULAN

Penelitian ini berhasil membangun sistem analisis sentimen terhadap ulasan pengguna aplikasi ChatGPT menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) sebagai ekstraksi fitur dan *Support Vector Machine* (SVM) sebagai algoritma klasifikasi. Sistem yang dikembangkan berhasil mengklasifikasikan ulasan ke dalam kategori positif, negatif, dan netral dengan performa klasifikasi yang baik.

Penelitian ini memberikan kontribusi berupa penerapan kombinasi metode TF-IDF dan SVM sebagai pendekatan untuk menganalisis sentimen ulasan pengguna aplikasi ChatGPT sehingga dapat memberikan gambaran mengenai persepsi pengguna berdasarkan ulasan di Google Play Store. Namun, penelitian ini masih memiliki keterbatasan berupa ketidakseimbangan jumlah data pada setiap kelas sentimen sehingga performa klasifikasi pada kelas netral dan negatif belum optimal. Oleh karena itu, penelitian selanjutnya dapat menerapkan teknik penyeimbangan data seperti SMOTE atau menggunakan pendekatan berbasis *transformer* untuk meningkatkan performa klasifikasi, terutama pada kelas minoritas.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan dan bantuan dalam proses pelaksanaan penelitian ini. Terima kasih disampaikan kepada dosen pembimbing yang telah memberikan arahan, masukan, serta bimbingan selama proses penelitian dan penyusunan artikel ini. Penulis juga mengucapkan terima kasih kepada pihak yang telah mendukung penyediaan data, serta kepada keluarga dan rekan-rekan yang telah memberikan motivasi dan dukungan selama

proses penelitian berlangsung. Semoga penelitian ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan, khususnya dalam bidang analisis sentimen dan penerapan metode *machine learning* pada pengolahan data teks.

REFERENSI

[1] Z. Wang, Q. Xie, Z. Xu, Y. Zhang, Y. Yu, and R. Fu, "Is ChatGPT a Good Sentiment Analyzer? A Preliminary Study," *arXiv preprint arXiv:2304.04339*, 2023.

[2] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment Classification using Machine Learning Techniques," in *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing*, pp. 79–86, 2002.

[3] G. Salton and C. Buckley, "Term-weighting Approaches in Automatic Text Retrieval," *Information Processing & Management*, vol. 24, no. 5, pp. 513–523, 1988.

[4] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge, United Kingdom: Cambridge University Press, 2008.

[5] T. Joachims, "Text Categorization with Support Vector Machines: Learning with Many Relevant Features," in *Proceedings of the 10th European Conference on Machine Learning (ECML)*, pp. 137–142, 1998.

[6] F. Sebastiani, "Machine Learning in Automated Text Categorization," *ACM Computing Surveys*, vol. 34, no. 1, pp. 1–47, Mar. 2002.

[7] A. Pak and P. Paroubek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining," in *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC)*, pp. 1320–1326, 2010.

[8] M. Hu and B. Liu, "Mining and Summarizing Customer Reviews," in *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 168–177, 2004.

[9] T. Fawcett, "An Introduction to ROC Analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, Jun. 2006.

[10] M. Sokolova and G. Lapalme, "A Systematic Analysis of Performance Measures for Classification Tasks," *Information Processing & Management*, vol. 45, no. 4, pp. 427–437, Jul. 2009.

[11] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.

[12] Y. Kim, "Convolutional Neural Networks for Sentence Classification," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1746–1751, 2014.

[13] K. K. Chaturvedi, A. K. Gupta, and P. Prasad, Sentiment Analysis of ChatGPT User Reviews Using Machine Learning Approaches,"`International Journal of Information Management Data Insights, vol. 4, 2024.

[14] M. A. Aljanabi, M. Ghazi, A. H. Ali, and S. A. Abed, "ChatGPT: Open Possibilities, Challenges, and Implications in Education," Journal of Artificial Intelligence and Applications, vol. 3, no. 1, pp. 1–10, 2023.

[15] A. A. Alsaeedi and M. K. Khan,"A Study on Sentiment Analysis Techniques of Social Media Data," International Journal of Advanced Computer Science and Applications, vol. 12, no. 2, pp. 1–10, 2021.

[16] R. K. Kaliyar, A. Goswami, and P. Narang, "FakeBERT: Fake News Detection in Social Media with a BERT Based Deep Learning Approach," Multimedia Tools and Applications, vol. 80, pp. 11765–11788, 2021.