

BAB I

PENDAHULUAN

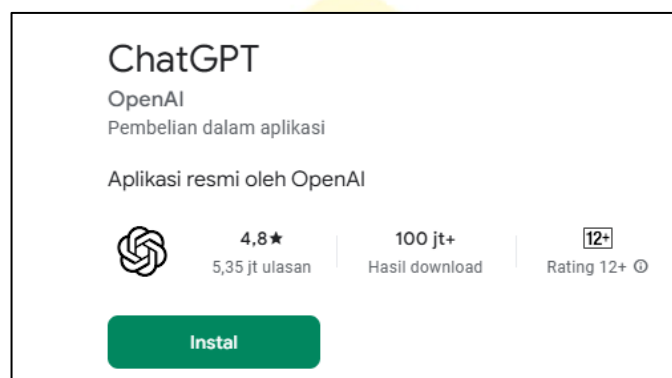
A. Latar Belakang

Dalam realitas digital kontemporer, *Artificial Intelligence* (AI) atau kecerdasan buatan telah menjadi salah satu inovasi teknologi paling revolusioner di era digital, dengan kemampuan untuk mensimulasikan kecerdasan manusia dalam berbagai bidang (Oktavian et al., 2024). Teknologi ini digunakan untuk memproses data, menganalisis pola, dan memberikan solusi otomatis yang mendukung berbagai aktivitas manusia. Penerapan AI dapat ditemukan di banyak sektor, seperti kesehatan, pendidikan, bisnis, hingga hiburan, yang semuanya bertujuan untuk meningkatkan efisiensi dan produktivitas (Faris et al., 2024). Dengan kemampuannya yang terus berkembang, AI tidak hanya menjadi alat pendukung, tetapi juga berperan sebagai inovator dalam menciptakan peluang dan solusi baru yang relevan dengan kebutuhan masyarakat *modern* (Ramadhana & Nasution, 2024).

Salah satu implementasi AI yang menonjol dalam kehidupan sehari-hari adalah aplikasi *ChatGPT* yang dikembangkan oleh OpenAI (Nita et al., 2023). *ChatGPT* memanfaatkan kecerdasan buatan untuk memahami dan merespons teks dengan cara yang menyerupai komunikasi manusia (Leippold, 2023). Aplikasi ini telah digunakan secara luas untuk berbagai keperluan seperti, memberikan saran, menjawab pertanyaan, hingga mendukung pembelajaran (Hassani & Silva, 2023). Kemampuannya untuk memahami konteks percakapan membuat *ChatGPT* semakin populer di kalangan individu maupun profesional (Niyu et al., 2024). Sebagai salah satu aplikasi AI, *ChatGPT* telah menarik perhatian banyak pengguna, yang kemudian memberikan ulasan dan umpan balik melalui *platform* distribusi aplikasi seperti *Google Play Store* (Sagala & Samuel, 2024).

Sebagai *platform* distribusi aplikasi Android, *Google Play Store* menyediakan ruang bagi pengguna untuk memberikan ulasan dan

penilaian terhadap aplikasi yang mereka gunakan (Eskiyaturrofikoh & Suryono, 2024), termasuk *ChatGPT*. Melalui fitur ini, pengguna dapat menyampaikan pengalaman mereka, baik yang positif maupun negatif, sehingga memberikan gambaran mengenai kualitas aplikasi secara keseluruhan (Viriya et al., 2024). Hingga November 2024, *ChatGPT* telah menerima 5,35 juta ulasan dengan *rating* rata-rata 4.8 di *Google Play Store* yang bisa dilihat pada gambar 1.1. Namun, tingginya volume ulasan tersebut menghadirkan tantangan bagi pengembang aplikasi untuk memahami opini pengguna secara menyeluruh (Sagala & Samuel, 2024). Dalam konteks inilah analisis sentimen menjadi relevan sebagai solusi untuk mengekstraksi informasi penting dari data ulasan yang sangat besar.



Gambar 1. 1 *Rating* aplikasi *ChatGPT* pada *Google Play Store*

Untuk mengatasi tantangan tersebut, analisis sentimen menjadi pendekatan yang tepat dalam memahami pola opini pengguna dari ulasan yang sangat besar (Sagala & Samuel, 2024). Analisis sentimen adalah proses yang bertujuan untuk mengelompokkan ulasan atau opini pengguna ke dalam kategori tertentu, seperti sentimen positif, negatif, atau netral (Cindo et al., 2019). Pendekatan ini memungkinkan pengembang untuk menganalisis persepsi pengguna secara efisien tanpa harus membaca setiap ulasan secara manual (Pribadi, 2019). Dalam konteks aplikasi *ChatGPT*, analisis sentimen dapat membantu mengidentifikasi kekuatan dan kelemahan aplikasi berdasarkan umpan balik pengguna, sehingga memberikan landasan yang kuat untuk pengembangan fitur dan

peningkatan kualitas layanan di masa mendatang (Salsabila & Muntahanah, n.d.).

Penelitian oleh Gilbert Jeffson Sagala dan Yusran Timur Samuel (2024) menganalisis sentimen ulasan pengguna aplikasi *ChatGPT* di *Google Play Store* menggunakan algoritma *Random forest*, SVM, dan Naïve Bayes. Sebanyak 2.652 ulasan dikumpulkan melalui *web scraping* (28 Juli 2023–28 Januari 2024). Proses prapemrosesan meliputi pembersihan teks, tokenisasi, *stemming*, dan pembobotan TF-IDF. Ketidakseimbangan data diatasi dengan SMOTE. Hasil menunjukkan bahwa *Random forest* dan SVM mencapai akurasi 91% dan *F1-score* 90%. Naïve Bayes menghasilkan akurasi 90% dan *F1-score* 87%. Penelitian ini menunjukkan dominasi sentimen positif dan menegaskan efektivitas *Random forest* dan SVM dalam klasifikasi sentimen. (Sagala & Samuel, 2024).

Penelitian oleh Ghanim Kanugrahan, Vito Hafizh Cahaya Putra, dan Yudi Ramdhani (2024) menganalisis sentimen pengguna terhadap aplikasi Gojek berdasarkan 65.584 ulasan dari Kaggle. Sentimen diklasifikasikan menggunakan algoritma SVM, *Random forest*, dan *Decision tree*, dengan pelabelan berdasarkan skor bintang. Proses prapemrosesan mencakup normalisasi teks, penghapusan tanda baca dan duplikat, serta *filtering* teks pendek. Fitur diekstraksi menggunakan TF-IDF dan *word count*. Ketidakseimbangan data ditangani dengan SMOTE. Hasil evaluasi menunjukkan bahwa model SVM dengan kernel RBF memiliki performa terbaik (akurasi 87,72% dan *F1-score* 87,69%), disusul SVM Linear (87,62%). *Random forest* menunjukkan performa yang stabil dengan akurasi sekitar 84,5%, sedangkan *Decision tree* memiliki performa terendah (74–78%). Penelitian ini menyimpulkan bahwa SVM merupakan algoritma paling unggul, sementara *Random forest* tetap kompetitif dalam klasifikasi sentimen pengguna aplikasi Gojek. (Kanugrahan et al., 2024).

Penelitian ini menggunakan empat algoritma klasifikasi, yaitu *Naive bayes*, *Support vector machine* (SVM), *Decision tree*, dan *Random*

forest. *Naive bayes* merupakan algoritma berbasis probabilitas yang mengasumsikan independensi antar fitur, *SVM* memisahkan data menggunakan *hyperplane* optimal, *Decision tree* membagi data berdasarkan fitur-fitur tertentu dalam struktur pohon keputusan, sementara *Random forest* merupakan metode *ensemble* yang menggabungkan sejumlah pohon keputusan untuk meningkatkan akurasi dan mengurangi risiko *overfitting*. Keempat algoritma ini dipilih karena mewakili pendekatan klasifikasi yang berbeda seperti probabilistik, *margin-based*, *rule-based*, dan *ensemble*, serta telah banyak digunakan dalam penelitian analisis sentimen. Pemilihan ini memungkinkan perbandingan performa model secara lebih menyeluruh sesuai dengan karakteristik data yang digunakan.

Data yang digunakan dalam penelitian ini berasal dari ulasan pengguna aplikasi *ChatGPT* di *Google Play Store*. Data diambil dalam rentang waktu tiga bulan, yaitu mulai bulan Januari sampai bulan Maret 2025 untuk memastikan kebaruan data yang dianalisis, dengan jumlah data yang didapatkan sebanyak 2289 sampel ulasan. Pengambilan data dilakukan menggunakan teknik *web scraping*, yang memungkinkan pengambilan data secara otomatis dari situs web. Penelitian ini tidak hanya berfokus pada penerapan keempat algoritma klasifikasi tersebut, tetapi juga pada evaluasi performa masing-masing algoritma. Pengukuran performa dilakukan menggunakan *Confusion matrix* dengan metrik seperti akurasi, presisi, *recall*, dan *F1-score* (Sari & Hayuningtyas, 2019).

Penulis juga melakukan analisis terhadap ulasan pengguna untuk mengetahui tanggapan yang diberikan berdasarkan data hasil pelabelan awal, yang kemudian divisualisasikan dalam bentuk *wordcloud* sehingga kata-kata kunci yang muncul dapat merepresentasikan sentimen pengguna secara lebih akurat. Analisis ini bertujuan untuk membantu memahami bagaimana pengguna menilai fitur, kinerja dan manfaat *ChatGPT* di kehidupan sehari-hari dan membandingkan performa antar algoritma klasifikasi yang digunakan, guna mengetahui algoritma mana yang paling

efektif dalam mengklasifikasikan sentimen ulasan pengguna, sehingga dapat memberikan rekomendasi mengenai metode klasifikasi yang paling efektif.

Berdasarkan uraian di atas, peneliti mengusulkan penelitian mengenai “Analisis Sentimen Aplikasi *ChatGPT* pada *Play Store* Menggunakan Pendekatan Data Mining Dengan Evaluasi Kinerja Model”.

B. Rumusan Masalah

1. Bagaimana melakukan analisis sentimen ulasan masyarakat Indonesia terhadap *ChatGPT*?
2. Bagaimana melakukan pengukuran dan perbandingan evaluasi performa antar algoritma *Naive bayes*, *Support vector machine (SVM)*, *Decision tree*, *Random forest* dalam mengklasifikasikan sentimen?

C. Tujuan Penelitian

1. Untuk mengetahui analisis sentimen ulasan masyarakat Indonesia terhadap *ChatGPT*.
2. Untuk mengetahui pengukuran dan perbandingan evaluasi performa antar algoritma *Naive bayes*, *Support vector machine (SVM)*, *Decision tree*, *Random forest* dalam mengklasifikasikan sentimen.

D. Batasan Masalah

1. Sumber data berasal dari *Google Play Store* berupa komentar dan *rating*.
2. Algoritma yang digunakan untuk mengklasifikasikan sentimen pada penelitian ini adalah *Naive bayes Classifier (NBC)*, *Support vector machine (SVM)*, *Decision tree*, *Random forest*.
3. Data yang digunakan terbatas, hanya ulasan dengan bahasa Indonesia dan dengan lokasi ID negara Indonesia.
4. Pelabelan sentimen dilakukan secara otomatis berdasarkan nilai *rating* pengguna (> 3 bernilai positif, $= 3$ bernilai netral, < 3 bernilai negatif).
5. Hasil keluaran pada penelitian ini berupa sentimen ulasan aplikasi yang bernilai positif, negatif, dan netral.

6. *Tools* yang digunakan mulai dari pengumpulan data, teks *preprocessing* sampai evaluasi algoritma adalah Python.
7. Data diambil dalam rentang waktu 3 bulan dari bulan Januari – Maret 2025 dengan jumlah data 2289.
8. Perbandingan yang dicari adalah tingkat akurasi (*accuracy*), Presisi (*precision*), sensitivitas (*recall*) dan *F1-score*.

E. Manfaat Penelitian

Manfaat yang dapat diperoleh pada penelitian ini adalah:

1. Manfaat praktis

Memberikan rekomendasi berbasis data bagi pengembang *ChatGPT* dalam meningkatkan kualitas layanan. Analisis terhadap ulasan pengguna yang divisualisasikan melalui *wordcloud* dapat membantu memahami tanggapan pengguna terkait fitur, kinerja, dan manfaat *ChatGPT* dalam kehidupan sehari-hari. Hasil penelitian ini juga dapat menjadi acuan bagi pengembang aplikasi lain dalam memahami kebutuhan dan ekspektasi pengguna terhadap layanan berbasis kecerdasan buatan.

2. Manfaat Teoritis

Penelitian ini menjadi contoh penerapan text *mining* dalam analisis sentimen pada data publik. Selain itu, penelitian ini memperkaya kajian mengenai perbandingan kinerja algoritma klasifikasi (Naive Bayes dan SVM) untuk menentukan metode yang lebih efektif dalam mengklasifikasikan sentimen ulasan pengguna. Hasilnya diharapkan dapat menjadi referensi bagi penelitian selanjutnya yang berfokus pada evaluasi dan pengembangan metode klasifikasi dalam pemrosesan data teks.