

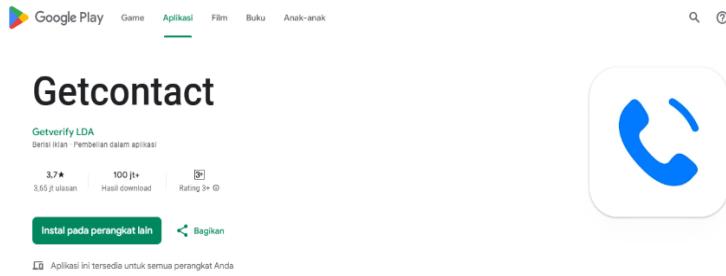
BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi yang pesat telah membawa perubahan besar dalam cara masyarakat berinteraksi dan menyampaikan opini. Salah satu interaksi pengguna dengan berbagai aplikasi, salah satunya adalah melalui ulasan pengguna (*user review*) di platform distribusi aplikasi seperti *Google Play Store*. Ulasan pengguna menjadi salah satu sumber informasi penting yang dapat digunakan oleh pengembang untuk memahami kepuasan, keluhan, serta saran dari pengguna aplikasi mereka. Ulasan-ulasan ini dapat berisi tanggapan positif maupun negatif terhadap suatu aplikasi, yang sangat bermanfaat bagi pengembang untuk meningkatkan kualitas aplikasinya, serta membantu calon pengguna dalam pengambilan keputusan.

Salah satu aplikasi populer yang banyak dibahas oleh pengguna adalah *GetContact*. Aplikasi *Getcontact* adalah alat populer yang membantu pengguna mengidentifikasi panggilan tak dikenal dan memblokir panggilan spam. Salah satu fitur unik *GetContact* adalah kemampuannya menampilkan "tag" atau label yang diberikan oleh pengguna lain terhadap nomor tertentu. Fitur ini memungkinkan pengguna untuk melihat bagaimana nomor telepon mereka atau nomor lain diberi nama dalam kontak pengguna lain. Hal ini membuka ruang untuk analisis sentimen publik terhadap nomor tertentu berdasarkan tag yang muncul (Rita & Pambudi, 2023). Namun, seperti aplikasi lainnya, *GetContact* menerima berbagai ulasan dari penggunanya, mulai dari pujian atas fungsionalitasnya hingga keluhan terkait privasi dan performa aplikasi. Volume ulasan yang sangat besar membuat proses analisis manual menjadi tidak efisien.

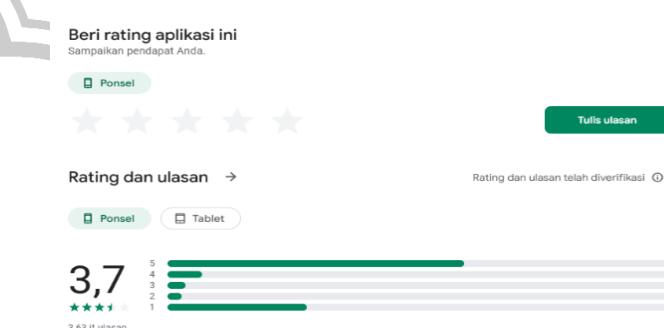


Gambar 1. 1 Rating Aplikasi *GetContact*

Sumber: *Google Play store*, 2024

Pada gambar 1.1 dijelaskan bahwa di *Google Play Store*, aplikasi ini telah diunduh lebih dari 100 juta kali dan menerima lebih dari 3,6 juta ulasan dengan rating rata-rata 3,7 bintang. Dengan lebih dari 100 juta unduhan, *Getcontact* memiliki basis pengguna yang luas, menghasilkan volume ulasan yang signifikan untuk dianalisis. Namun, pada kenyataan nya dalam ulasan aplikasi *Getcontact* ini juga banyak sekali ulasan pengguna yang bersifat negatif maupun positif.

Sentimen terhadap aplikasi ini cukup beragam. Beberapa orang merasa fitur ini membantu mereka mengenali nomor yang tidak dikenal, seperti panggilan spam atau potensi penipuan. Namun, ada juga kekhawatiran terkait privasi, karena *GetContact* mengakses daftar kontak pengguna untuk mengelola tag dan identifikasi nomor. Banyak pihak mengkritik aplikasi ini karena dianggap terlalu invasif terhadap data pribadi. Dengan banyaknya data berupa tag yang dihasilkan oleh pengguna, arah pengembangan aplikasi ini bisa melibatkan teknologi kecerdasan buatan untuk menganalisis sentimen publik terhadap nomor tertentu. Misalnya, *GetContact* dapat memberikan peringkat sentimen positif, atau negatif, terhadap nomor berdasarkan pola tag yang diberikan.



Gambar 1. 2 Rating Ulasan *Getcontact*

Sumber: *Google Play store*, 2024

Dalam gambar 1.2 Analisis sentimen ini terdapat ulasan pengguna yang dapat memberikan wawasan berharga mengenai persepsi publik terhadap aplikasi ini. Algoritma *Naïve Bayes*, yang sering digunakan dalam klasifikasi teks, cocok untuk tugas akhir ini karena kemampuannya dalam menangani data dengan fitur yang banyak dan kompleksitas yang rendah. *Natural Language Processing* (NLP) adalah penerapan ilmu komputer, khususnya linguistik komputasional (*computational linguistics*), untuk mengkaji interaksi antara komputer dengan bahasa (alami) manusia. Salah satu metode yang sering digunakan dalam pengolahan bahasa alami (NLP) untuk klasifikasi teks adalah algoritma *Naïve Bayes*, yang memiliki keunggulan dalam hal kesederhanaan, efisiensi, dan kemampuan untuk memberikan hasil yang cukup baik dalam berbagai kasus klasifikasi teks (Ratnawati, 2018).

Algoritma *Naïve Bayes* telah terbukti efektif dalam menganalisis sentimen melalui berbagai penelitian. Berikut adalah beberapa contoh studi yang mendukung penggunaan algoritma ini:

1. Analisis Sentimen Menggunakan Algoritma *Naïve Bayes* Terhadap Review Restoran: Studi ini mengaplikasikan *Naïve Bayes* untuk menganalisis sentimen dalam ulasan restoran, dengan hasil yang menunjukkan efektivitas algoritma dalam mengklasifikasikan sentimen positif dan negatif (Permadi, 2020).
2. Penerapan Algoritma *Naïve Bayes* dalam Menganalisis Sentimen pada Review Pengguna *E-Commerce*: Studi ini mengaplikasikan *naïve bayes* untuk menganalisis sentimen pada review pengguna *e-commerce* dengan nilai sebesar *accuracy* sebesar 99,5%, *precision* sebesar 99,49%, *recall* sebesar 100%. Dengan demikian dapat disimpulkan metode klasifikasi algoritma *Naïve Bayes* cukup relevan meskipun akurasi nya belum 100% (Laia & Yamin, 2023).

Berdasarkan bukti-bukti tersebut, algoritma *Naïve Bayes* terbukti efektif dalam analisis sentimen, terutama dalam mengklasifikasikan opini atau ulasan menjadi kategori positif, negatif, atau netral.

Dengan menggunakan algoritma ini, analisis sentimen pada ulasan *GetContact* dapat dilakukan secara otomatis dan lebih efisien. Metode ini dikenal

karena kesederhanaan dan efektivitasnya dalam pengklasifikasian data teks ke dalam kategori tertentu seperti sentimen positif atau negatif. Dengan memanfaatkan algoritma ini dapat melakukan analisis untuk mengidentifikasi faktor utama yang mempengaruhi kepuasan atau ketidakpuasaan terhadap aplikasi ini.

Fokus penelitian ini adalah analisis sentimen ulasan aplikasi GetContact di Google Play Store yang menggunakan algoritma Naive Bayes. Tujuan utama dari penelitian ini adalah untuk menentukan apakah ulasan pengguna terhadap aplikasi ini cenderung bersifat positif atau negatif. Diharapkan bahwa hasil penelitian ini akan memberikan pemahaman yang lebih mendalam tentang bagaimana persepsi pengguna memengaruhi reputasi aplikasi. Hasil ini juga akan menjadi dasar bagi peneliti selanjutnya untuk meningkatkan kualitas penelitian dan kualitasnya. Sebagai tanggapan terhadap masalah yang sudah dijelaskan sebelumnya, tugas akhir adalah "Klasifikasi Sentimen Positif dan Negatif pada Ulasan GetContact dengan Algoritma Naive Bayes."

Dengan latar belakang ini, penelitian ini menjadi penting untuk dilakukan mengingat peran sentimen pengguna dalam membentuk opini publik terhadap suatu aplikasi. Selain itu, implementasi metode *naïve bayes* diharapkan dapat menjadi solusi yang efektif dalam mengolah data ulasan yang berjumlah besar secara efisien dan akurat.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang di atas, dapat dirumuskan permasalahan sebagai berikut, yaitu :

1. Bagaimana penerapan algoritma *naïve bayes* dalam mengklasifikasi sentimen positif dan negatif pada ulasan aplikasi *GetContact*
2. Berapa nilai akurasi, presisi, recall dan *F-1 Score* dalam mengklasifikasi sentimen positif dan negatif pada ulasan pengguna di aplikasi *GetContact* ?

1.3 Tujuan Penelitian

Dengan mempertimbangkan permasalahan yang dihadapi, tujuan dari penelitian ini adalah sebagai berikut :

1. Untuk mengevaluasi kemampuan algoritma *naïve bayes* dalam mengklasifikasi sentimen positif dan negatif pada ulasan aplikasi *GetContact*.

1.4 Manfaat Penelitian

Adapun manfaat dari tercapainya penelitian ini yaitu :

1. Bagi peneliti selanjutnya, tugas akhir ini dapat menjadi referensi atau acuan yang berkaitan dengan analisis sentimen pada ulasan aplikasi atau untuk mengembangkan dan membandingkan algoritma klasifikasi sentimen.
2. Sebagai referensi pembelajaran dan pengembangan keterampilan di bidang machine learning, analisis sentimen, dan pemrosesan data teks untuk mendukung Pendidikan berbasis teknologi.

1.5 Batasan Masalah

Adapun beberapa Batasan masalah yang diangkat dalam melakukan penelitian ini, antara lain:

1. Data yang digunakan berasal dari platform *Kaggle*.
2. Panjang teks minimal 5 kata.
3. Metode yang digunakan adalah *naïve bayes* tanpa perbandingan dengan metode lain.
4. Klasifikasi sentimen pengguna yang dilakukan berdasarkan sentimen positif dan negatif.
5. Data ulasan yang digunakan pada tahun 2024.
6. Jumlah ulasan yang ada di google playstore yaitu sebanyak 3,71 juta ulasan.
7. Dataset yang diambil tidak melibatkan gambar, video, atau tautan.
8. Tidak mengklasifikasi emoticon.
9. Perangkat lunak yang digunakan sebagai analisis adalah *Google Collab* dengan Bahasa pemrograman *phyton*.
10. Sentimen yang digunakan yaitu sentimen umum yang berfokus pada persepsi keseluruhan pengguna aplikasi.
11. Menggunakan matrik dasar seperti akurasi, *presisi*, *recall* dan *f1-score*.
12. Dataset yang di gunakan adalah Bahasa indonesia