

BAB I

PENDAHULUAN

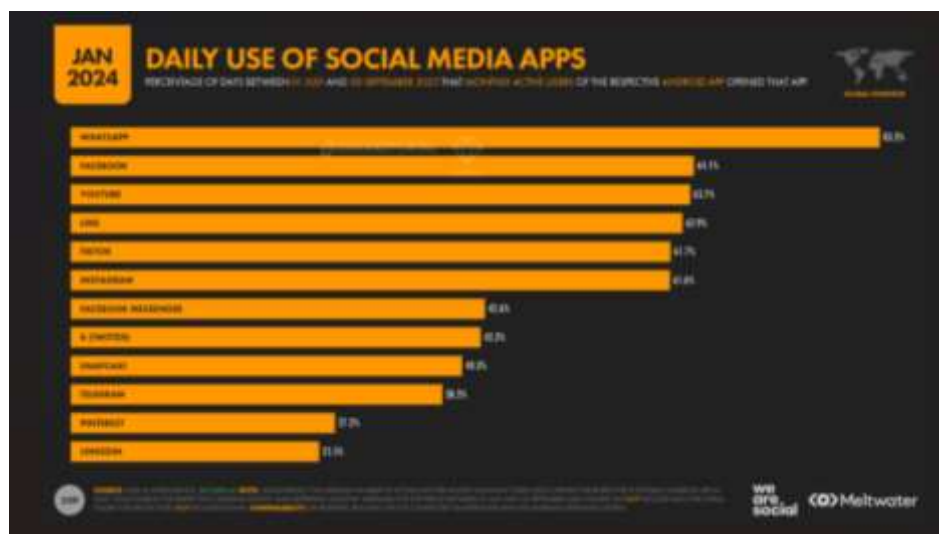
1.1 LATAR BELAKANG MASALAH

Opini atau sentimen publik semakin berkembang dan bebas disampaikan melalui berbagai media, sentimen tersebut menjadi potensi besar bagi perusahaan yang ingin memahami umpan balik dari masyarakat terhadap merek dagang mereka.[1] Analisis Sentimen merupakan teknik ekstraksi data teks yang digunakan untuk memperoleh informasi sentimen, baik yang bernilai positif maupun negatif.[2] Hal ini merupakan metode otomatis yang digunakan untuk menganalisis ulasan produk yang bertujuan memperoleh informasi terkait sentimen, yang merupakan bagian dari isi ulasan tersebut. Dengan demikian, sistem ini mampu mengetahui opini atau perasaan pengguna.[3]

Sentimen atau ulasan pengguna juga tersedia di *Google Play Store*, hal ini mencerminkan pengalaman mereka, baik yang positif maupun negatif, terhadap aplikasi yang digunakan. *Play Store*, sebagai salah satu layanan utama Google, menawarkan berbagai kategori, seperti game, penelusuran, penawaran, buku, dan aplikasi, termasuk media sosial serta e-commerce. Dengan beragam pilihan aplikasi ini, *Play Store* menjadi platform penting bagi pengguna untuk menyampaikan opini mereka melalui ulasan. *Google Play Store* dapat diakses melalui Android, situs web dan Android TV. Berdasarkan Data *Mobile Operating System Market Share in Indonesia* mulai dari September 2023 sampai September 2024, terdapat 87,56% pengguna Android dan 12,33% pengguna ios, yang berarti pengguna *Play Store* lebih banyak dibanding dengan *App Store*, yang mana pada

kedua aplikasi tersebut terdapat kolom penilaian dan ulasan para pengguna tentang aplikasi yang sudah di unduh.[4]

Salah satu aplikasi populer yang di gunakan sehari-hari untuk media komunikasi adalah WhatsApp, Menurut laporan *We Are Social* pada Januari 2024, 83,2% pengguna internet yang berusia mulai dari 16 tahun menggunakan WhatsApp dengan alasan untuk berhubungan dengan teman dan keluarga.



Gambar 1. 1 penggunaan media social [We Are Sosial]

WhatsApp juga tersedia dalam versi berbasis web, memungkinkan pengguna untuk mengakses layanan tanpa perlu mengunduh aplikasi, pengguna cukup membuka situs web melalui *Chrome*, *Microsoft Edge*, dan lainnya. Ulasan pengguna tentang aplikasi WhatsApp di *Google Play Stores* saat ini mencapai rating 4,2, ini akan menjadi sumber informasi kritis untuk memahami persepsi dan pengalaman pengguna WhatsApp dan penting bagi pengembang dan manajemen WhatsApp untuk meningkatkan layanan, mengelola reputasi, dan mempertahankan keunggulan kompetitif.[5]

Terdapat beberapa penelitian sebelumnya terkait analisis sentiment review whatsapp menggunakan Algoritma *Naïve Bayes*, *Support Vector Machine*, dan salah satunya oleh Muhammad Riski[6], yaitu klasifikasi sentiment ulasan aplikasi whatsapp menggunakan Algoritma *KNN(K-Nearest Neighbor)* melalui 3 skenario dengan pembagian data latih dan data uji yang berbeda. Skenario pertama membagi data dengan rasio 70% data latih dan 30% data uji, skenario kedua dengan rasio 80% data latih dan 20% data uji, dan skenario ketiga dengan rasio 90% data latih dan 10% data uji. Pada penelitian tersebut, dihasilkan akurasi tertinggi adalah ketika pembagian data 70:30 diperoleh dengan nilai $K=14$ dan $\text{threshold} = 20$ yang menghasilkan *accuracy* mencapai 84%, nilai presisi sebesar 87,65%, *recall* sebesar 92,21%, dan *f1-score* sebesar 89,87% untuk kelas positif. Sementara pada kelas negatif, nilai *precision* mencapai 68,42%, *recall* mencapai 56,52%, dan *f1-score* mencapai 61,90%. *Confusion matrix* menunjukkan model mengklasifikasikan positif sebanyak 81 dan negatif sebanyak 19.[7]

Dalam tahapan analisis sentimen terdapat metode klasifikasi, yang umum digunakan adalah *Logistic Regression*, metode *Logistic Regression* merupakan suatu metode klasifikasi yang dipakai untuk membentuk model hubungan antar variabel independen dan variabel dependen, Pada algoritma klasifikasi ini berfungsi untuk mencari hubungan antara input dengan probabilitas hasil output dimana input bisa berupa variabel diskrit atau kontinu sedangkan output dari *Logistic Regression* yang memiliki karakteristik biner (dua kelas) seperti positif atau negatif, **spam** atau **bukan spam**, (1) atau (0).[8]

Salah satu penelitian yang membahas penggunaan *Logistic Regression* dalam analisis sentimen adalah penelitian berjudul "Analisis Sentimen Nasabah pada Layanan Perbankan Menggunakan Metode Regresi Logistik Biner, Naïve Bayes Classifier (NBC), dan Support Vector Machine (SVM)". Penelitian ini membandingkan beberapa metode klasifikasi untuk mengelompokkan sentimen nasabah terhadap layanan perbankan. Hasilnya, meskipun metode SMOTE-SVM kernel RBF memiliki kinerja terbaik, *Logistic Regression* tetap relevan untuk analisis sentimen. Algoritma ini mampu mengolah data ulasan pengguna dan mengklasifikasikan sentimen secara akurat. Oleh karena itu, *Logistic Regression* sering digunakan dalam analisis sentimen berbasis teks[9].

Penelitian yang dilakukan ini adalah tentang sentimen analisis untuk menerapkan algoritma *Logistic Regression*. Penelitian ini termasuk kedalam sebuah penelitian yang bernama *fined grained sentiment analysis* yaitu analisis pada suatu kalimat komentar.[10]Ulasan yang diberikan pengguna dalam bahasa Indonesia di aplikasi *Google Play Store* menjadi sumber informasi yang digunakan untuk proses pengambilan keputusan. Ulasan ini memberikan gambaran tentang kesesuaian aplikasi dengan kebutuhan pengguna, kemudahan penggunaan aplikasi, serta efektivitas dan efisiensi desain aplikasi dalam memungkinkan interaksi pengguna dengan sistem. Semua aspek ini dapat diketahui melalui ulasan yang diberikan oleh pengguna.[11]Data dari ulasan *google play store* tersebut akan diproses menggunakan *machine learning*, kemudian melakukan pencarian, perhitungan komentar positif dan negatif, serta menghasilkan klasifikasi menggunakan algoritma *Logistic Regression* , Penelitian

ini dilakukan fokus kepada kritik dan saran dalam penggunaan aplikasi Whatsapp yang berasal dari akun resmi *google play store*. Proses pengolahan data dan klasifikasi sentimen analisis menggunakan bahasa pemrograman *Python*.

1.2 RUMUSAN MASALAH

Berdasarkan latar belakang permasalahan yang telah diuraikan diatas maka dapat diambil rumusan yang akan menjadi pembahasan penelitian ini, yaitu :
“Bagaimana cara menerapkan algoritma *Logistic Regression* untuk menganalisis sentimen pada ulasan aplikasi WhatsApp di *Google Play Store*”.

1.3 BATASAN MASALAH

Agar penelitian lebih fokus dan terarah, terdapat beberapa batasan masalah yang ditetapkan dalam penelitian ini, yaitu:

1. Jenis dataset yang digunakan adalah dataset *Whatsappreview* yang diambil dari *Kaggle Datasets* dan berasal dari *Google Play Store* dalam rentang waktu tertentu dan menggunakan bahasa Indonesia sebanyak 49.502 ulasan, dengan fokus pada dataset yang sudah ada.
2. Penelitian hanya menggunakan fitur yang tersedia dalam dataset yaitu *score* dan *content*.
3. Penelitian ini terbatas pada penggunaan algoritma *Logistic Regression* untuk analisis sentimen. Algoritma lain tidak dibahas dalam penelitian ini.
4. Evaluasi model hanya akan dilakukan dengan metrik seperti *accuracy*, *precision*, *recall*, dan *F1-score*. Metrik evaluasi lain yang lebih kompleks tidak akan dibahas. Selain itu, pembobotan fitur dalam analisis sentiment

hanya akan menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF), tanpa membahas teknik pembobotan lainnya.

1.4 TUJUAN DAN MANFAAT PENELITIAN

1.4.1 Tujuan Penelitian

Adapun tujuan yang ingin dicapai pada penelitian dan penyusunan tugas akhir ini adalah :

1. Melihat sejauh mana tingkat akurasi algoritma *Logistic regression* pada tahapan analisis sentimen.
2. Mengklasifikasi nilai sentimen pada review aplikasi WhatsApp yaitu positif atau negative.

1.4.2 Manfaat Penelitian

Adapun manfaat yang didapat dari penelitian dan penyusunan tugas akhir ini adalah :

1. Mengetahui tingkat akurasi algoritma *Logistic Regression* pada tahapan analisis sentiment.
2. Mengetahui analisis sentimen pada ulasan aplikasi whatsapp di *google play store* menggunakan metode *Logistic Regression*.

1.5 SISTEMATIKA PENULISAN

Adapun sistematika penulisan yang digunakan dalam menyusun penelitian ini adalah sebagai berikut :

BAB I : PENDAHULUAN

Pada bab ini penulis membahas tentang latar belakang masalah, perumusan masalah, pembatasan masalah, tujuan dan manfaat penelitian, dan sistematika penulisan.

BAB II : LANDASAN TEORI

Pada bab landasan teori ini membahas tentang teori-teori dan pendapat para ahli yang berhubungan langsung dengan permasalahan yang dianalisis, antara lain :Pengertian *data mining*, *text mining*, analisis sentimen, *Logistic Regression*, WhataApp.

BAB III : METODOLOGI PENELITIAN

Pada bab ini menjelaskan tentang kerangka kerja penelitian, metode pengumpulan data, metode pengklasifikasian, dan alat bantu yang digunakan selama mengerjakan penelitian.

BAB IV: ANALISIS

Pada bab ini akan dilakukan tahap analisis menggunakan algoritma *Logistic Regression classifier* dengan memanfaatkan data *review* aplikasi WhatsApp yang telah dikumpulkan.

BAB V :: PENUTUP

Pada bab ini membahas tentang kesimpulan dan saran selama melaksanakan proses peneliti.