

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi digital telah mendorong transformasi besar dalam berbagai aspek kehidupan, termasuk dalam aktivitas jual beli secara daring (*online*)[1]. *E-commerce* menjadi solusi yang semakin diminati oleh masyarakat karena kemudahan, kecepatan, dan kenyamanan yang ditawarkannya. Salah satu platform *e-commerce* terbesar di Indonesia saat ini adalah Shopee[2]. Berdasarkan laporan DataReportal dan GoodStats, Shopee termasuk platform *e-commerce* dengan jumlah kunjungan dan pengguna aktif tertinggi di Indonesia dalam beberapa tahun terakhir[3][4]. Dengan meningkatnya interaksi pengguna melalui ulasan, muncul kebutuhan untuk memahami opini publik secara sistematis sebagai dasar pengambilan keputusan bisnis, analisis terhadap ulasan pengguna menjadi penting karena dapat menggambarkan persepsi pelanggan secara langsung terhadap kualitas layanan, kemudahan penggunaan, serta tingkat kepuasan mereka terhadap platform, studi mengenai sentimen ulasan shopee menjadi krusial untuk mengevaluasi popularitas platform dan mengidentifikasi area peningkatan berdasarkan umpan balik[5].

Sebagai sebuah platform, Shopee tidak hanya menjadi perantara antara penjual dan pembeli dalam melakukan transaksi, tetapi juga menyediakan berbagai fitur seperti pencarian produk, sistem pembayaran, pengiriman barang, program loyalitas, serta layanan pelanggan[2]. Berbagai fitur tersebut dirancang untuk memberikan kemudahan, kenyamanan, dan keamanan bagi pengguna saat menggunakan aplikasi. Pengalaman pengguna terhadap fitur dan layanan tersebut sering kali dituangkan dalam bentuk ulasan atau komentar pada kolom ulasan aplikasi maupun media sosial. Ulasan yang diberikan pengguna dapat berupa kritik, saran, maupun penilaian terhadap kualitas layanan yang tersedia pada aplikasi Shopee[6].

Ulasan-ulasan tersebut mengandung informasi penting mengenai kepuasan atau ketidakpuasan pengguna terhadap performa aplikasi Shopee. Namun, jumlah ulasan yang sangat banyak dan bersifat tidak terstruktur membuat proses analisis secara manual menjadi tidak efisien. Oleh karena itu, dibutuhkan pendekatan berbasis teknologi yang mampu menganalisis opini pengguna secara otomatis dan akurat[7]. Terlebih lagi, dalam konteks *e-commerce*, ulasan pelanggan secara signifikan memengaruhi calon pembeli lainnya dalam memilih produk atau layanan sehingga analisis sentimen yang akurat menjadi semakin relevan[8].

Salah satu pendekatan yang dapat digunakan untuk menganalisis ulasan pengguna adalah *Natural Language Processing (NLP)*. NLP merupakan cabang kecerdasan buatan yang memungkinkan komputer memahami, mengolah, dan menganalisis bahasa manusia dalam bentuk teks. Melalui NLP, data ulasan yang sebelumnya tidak terstruktur dapat diproses menjadi informasi yang lebih mudah dianalisis. Dengan pendekatan tersebut, analisis sentimen terhadap ulasan pengguna dapat dilakukan secara otomatis dan lebih efisien. Penelitian ini perlu dilakukan karena banyaknya ulasan pengguna pada aplikasi Shopee sulit dianalisis secara manual, sehingga diperlukan metode yang mampu mengolah dan mengklasifikasikan sentimen secara cepat dan akurat. Hasil penelitian diharapkan dapat membantu memahami opini pengguna terhadap layanan aplikasi Shopee serta menjadi bahan evaluasi bagi pihak terkait dalam meningkatkan kualitas layanan yang diberikan[9] [10].

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana cara melakukan analisis sentimen terhadap ulasan pengguna aplikasi Shopee menggunakan pendekatan *Natural Language Processing (NLP)*?
2. Bagaimana representasi teks ulasan dengan metode TF-IDF dapat membantu dalam proses klasifikasi sentimen?

3. Seberapa baik performa algoritma *Logistic Regression* dalam mengklasifikasikan sentimen ulasan pengguna Shopee ke dalam kategori positif, negatif dan netral?

1.3 Batasan Masalah

Agar penelitian ini lebih terarah dan tidak menyimpang dari tujuan utama, maka batasan masalah yang digunakan adalah sebagai berikut:

1. Data yang digunakan dalam penelitian ini berupa ulasan pengguna aplikasi Shopee yang diambil dari Google Play Store. Penggunaan data dari Google Play Store dipilih karena platform tersebut menyediakan ulasan langsung dari pengguna aplikasi Shopee sehingga relevan untuk menggambarkan opini dan pengalaman pengguna terhadap sistem serta fitur aplikasi.
2. Data ulasan yang dianalisis dalam penelitian ini berjumlah 5.000 data. Jumlah tersebut dipilih karena dinilai telah mampu mewakili berbagai opini pengguna, baik sentimen positif, netral, maupun negatif, sehingga cukup untuk proses pelatihan dan pengujian model analisis sentimen. Data ulasan diambil dalam rentang waktu 12 Januari 2026 sampai 09 April 2026, yaitu ketika aplikasi Shopee berada pada versi v3.73.36.7.5.6, sehingga hasil penelitian mencerminkan persepsi pengguna terhadap fitur dan layanan aplikasi pada periode tersebut.
3. Penelitian ini hanya berfokus pada analisis sentimen teks ulasan dalam bahasa Indonesia. Pembatasan ini dilakukan agar proses *preprocessing*, normalisasi kata, *stemming*, dan analisis sentimen dapat dilakukan secara lebih optimal sesuai dengan karakteristik bahasa yang digunakan. Selain itu, penggunaan satu bahasa membantu meningkatkan konsistensi data dan mengurangi kesalahan interpretasi makna yang dapat memengaruhi performa model klasifikasi. Perbedaan bahasa dapat berpengaruh terhadap performa model karena setiap bahasa memiliki struktur kalimat, kosakata, singkatan, serta penggunaan slang yang berbeda, sehingga memerlukan pendekatan pemrosesan teks yang sesuai dengan bahasa tersebut.

4. Metode representasi teks yang digunakan adalah TF-IDF (*Term Frequency–Inverse Document Frequency*) untuk mengubah teks menjadi bentuk numerik.
5. Algoritma yang digunakan untuk klasifikasi sentimen adalah *Logistic Regression*, tanpa melakukan perbandingan dengan algoritma lain.
6. Kategori sentimen yang dianalisis menjadi tiga kelas, yaitu positif, negatif, dan netral.
7. Evaluasi performa model dilakukan menggunakan metrik akurasi, presisi, *recall*, dan *F1-score* untuk menilai tingkat keberhasilan model dalam mengklasifikasikan sentimen.

1.4 Tujuan Penelitian

Penelitian ini memiliki beberapa tujuan yang saling berkaitan dalam proses analisis sentimen terhadap ulasan pengguna aplikasi Shopee. Setiap tujuan dirancang untuk menjawab pertanyaan penelitian secara sistematis agar hasil yang diperoleh relevan dengan fokus penelitian. Adapun tujuan penelitian ini adalah sebagai berikut:

1. Untuk menjelaskan proses analisis sentimen terhadap ulasan pengguna aplikasi Shopee menggunakan pendekatan *Natural Language Processing* (NLP).
2. Untuk menerapkan metode TF-IDF dalam merepresentasikan teks ulasan menjadi bentuk numerik yang dapat diolah oleh algoritma pembelajaran mesin.
3. Untuk mengevaluasi performa algoritma *Logistic Regression* dalam mengklasifikasikan sentimen ulasan pengguna Shopee ke dalam kategori positif, negatif, dan netral menggunakan metrik akurasi, presisi, *recall*, dan *F1-score*.

1.5 Manfaat Penelitian

1. Manfaat Teoritis

Secara teoritis, penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan ilmu pengetahuan di bidang *Natural Language*

Processing (NLP) dan analisis sentimen. Melalui penerapan metode TF-IDF dan *Logistic Regression*, penelitian ini menunjukkan bagaimana pendekatan berbasis pembelajaran mesin dapat dimanfaatkan untuk menganalisis opini atau sentimen pengguna terhadap suatu aplikasi *e-commerce*. Hasil penelitian ini dapat menjadi rujukan bagi penelitian selanjutnya yang ingin mengembangkan model analisis sentimen dengan tingkat akurasi yang lebih tinggi, atau yang ingin membandingkan performa berbagai algoritma klasifikasi dalam konteks data teks berbahasa Indonesia

2. Manfaat Praktis

Secara praktis, penelitian ini diharapkan dapat memberikan manfaat bagi pihak Shopee sebagai objek penelitian, khususnya dalam memahami persepsi dan tingkat kepuasan pengguna terhadap performa aplikasi. Dengan adanya sistem analisis sentimen yang dibangun menggunakan metode NLP, TF-IDF, dan *Logistic Regression*, Shopee dapat memperoleh wawasan yang lebih mendalam terkait umpan balik pengguna secara otomatis dan dalam skala besar. Hasil analisis tersebut dapat dijadikan dasar untuk evaluasi, perbaikan layanan, peningkatan kualitas fitur, serta pengambilan keputusan strategis dalam pengembangan aplikasi. Selain itu, penelitian ini juga dapat menjadi acuan bagi perusahaan *e-commerce* lain dalam menerapkan teknologi analisis sentimen sebagai bagian dari sistem manajemen pengalaman pelanggan (*customer experience management*).

Selain itu, hasil penelitian ini juga dapat dijadikan referensi bagi pengembang sistem maupun peneliti lain yang tertarik dalam bidang analisis teks dan kecerdasan buatan.

1.6 Sistematika Penulisan

Sistematika penulisan dalam penelitian yang berjudul “Penerapan Metode *Logistic Regression* dan TF-IDF dalam Analisis Sentimen Terhadap Aplikasi Shopee” disusun sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi latar belakang penelitian mengenai pentingnya analisis sentimen pada

ulasan produk di *e-commerce* Shopee, rumusan masalah, tujuan penelitian, manfaat penelitian, serta sistematika penulisan skripsi.

BAB II TINJAUAN PUSTAKA

Bab ini membahas teori-teori yang mendukung penelitian, seperti konsep analisis sentimen, *Natural Language Processing* (NLP), *text mining*, *preprocessing* teks, TF-IDF, algoritma *Logistic Regression*, evaluasi model klasifikasi, serta penelitian terdahulu yang relevan.

BAB III METODE PENELITIAN

Bab ini menjelaskan metode yang digunakan dalam penelitian, meliputi objek penelitian, sumber dan jenis data, teknik pengumpulan data ulasan Shopee, tahapan *preprocessing* data, pelabelan sentimen, pembagian data latih dan data uji, proses ekstraksi fitur menggunakan TF-IDF, penerapan algoritma *Logistic Regression*, serta alat dan bahan yang digunakan dalam penelitian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi hasil implementasi penelitian, mulai dari proses pengolahan data, *preprocessing*, pelabelan sentimen, pelatihan model *Logistic Regression*, pengujian model, hingga evaluasi performa menggunakan *accuracy*, *precision*, *recall*, dan *F1-score*. Pada bab ini juga dilakukan pembahasan terhadap hasil analisis sentimen yang diperoleh.

BAB V PENUTUP

Bab ini berisi kesimpulan dari hasil penelitian yang telah dilakukan serta saran yang dapat digunakan untuk pengembangan penelitian selanjutnya.