

BAB I

PENDAHULUAN

A. Latar Belakang

Penyebaran ulasan pengguna pada platform e-commerce, seperti Tokopedia, memberikan dampak signifikan terhadap pengalaman dan keputusan yang diambil oleh pengguna. Sebagai salah satu platform e-commerce terbesar di Indonesia, Tokopedia tidak hanya berfungsi sebagai tempat transaksi, tetapi juga sebagai sumber informasi penting bagi penggunanya melalui ulasan yang mereka berikan. Namun, jumlah ulasan yang sangat beragam sering kali menyulitkan pengguna dalam menyaring informasi yang relevan dan akurat. Selain itu, Tokopedia dan platform e-commerce lainnya dihadapkan pada tantangan dalam memahami pola perilaku pengguna, yang berpotensi meningkatkan kualitas layanan, personalisasi pengalaman pengguna, dan efektivitas strategi pemasaran.

Berbagai penelitian telah dilakukan untuk menganalisis ulasan pengguna menggunakan pendekatan yang beragam. Pada penelitian (Kursiyanto dkk, 2024), diusulkan model *Bidirectional Encoder Representations from Transformers* (BERT) yang mampu menangkap konteks bahasa secara mendalam untuk analisis sentimen. Pada penelitian (Kurnia, 2023) mengaplikasikan algoritma *Stochastic Gradient Descent* (SGD) untuk analisis sentimen pada ulasan *e-commerce*. Pada penelitian (Adniana dkk,2024) mengadopsi algoritma X-Means, yang merupakan pengembangan dari K-Means, untuk menentukan jumlah klaster secara otomatis. Pada penelitian (Puspitasari dkk,2024) menggunakan algoritma Naïve Bayes untuk klasifikasi sentimen pada data media sosial seperti Twitter.

Metode-metode yang gunakan dalam analisis sentimen dan segmentasi ulasan pengguna pada platform e-commerce memiliki kelebihan dan kekurangannya masing-masing. Metode BERT unggul dalam memahami konteks kata secara mendalam, tetapi membutuhkan sumber daya komputasi besar dan waktu pelatihan yang lama (Kurnia, 2023). Metode SGD mampu memberikan akurasi yang tinggi pada ulasan Tokopedia, namun kurang optimal untuk data tidak seimbang dan sensitif terhadap pengaturan hyperparameter (Kursiyanto dkk, 2024). algoritma x-means mampu menentukan jumlah klaster secara otomatis dengan hasil klasterisasi lebih optimal, tetapi memerlukan waktu komputasi lebih lama pada dataset besar (Adniana dkk,2024). Algoritma Naïve Bayes unggul dalam menangani dataset kecil dan memberikan hasil cepat, tetapi sering kali kurang akurat dalam memahami konteks kalimat yang kompleks (Puspitasari dkk,2024). Dengan optimalisasi yang tepat, metode-metode ini dapat disesuaikan untuk kebutuhan klasifikasi yang lebih spesifik dan efisien.

Penelitian ini mengusulkan penerapan model BERT untuk analisis sentimen dan algoritma x-means untuk segmentasi ulasan pengguna pada *platform* e-commerce. Kombinasi kedua metode ini bertujuan untuk memberikan hasil analisis yang lebih akurat dan segmentasi yang lebih relevan. Model BERT akan digunakan untuk memahami sentimen dalam ulasan pengguna berdasarkan konteks yang mendalam, sementara algoritma x-means akan digunakan untuk mengelompokkan pengguna berdasarkan pola ulasan mereka secara otomatis.

Tujuan penelitian ini adalah mengimplementasikan model BERT dan algoritma x-means untuk analisis sentimen dan segmentasi ulasan pengguna pada

platform e-commerce. Implementasi ini diharapkan mampu memberikan wawasan yang lebih baik mengenai pola perilaku pengguna dan meningkatkan pengalaman pengguna secara keseluruhan. Dengan hasil analisis yang lebih mendalam dan segmentasi yang lebih presisi, *platform e-commerce* dapat memanfaatkan informasi ini untuk meningkatkan layanan, strategi pemasaran, dan personalisasi.

B. Rumusan Masalah

Berdasarkan uraian latar belakang yang telah dipaparkan, penelitian ini merumuskan beberapa rumusan masalah sebagai berikut:

1. Apakah model BERT dapat diimplementasikan untuk melakukan analisis sentimen terhadap ulasan pengguna pada platform e-commerce?
2. Apakah algoritma X-Means dapat diaplikasikan untuk melakukan segmentasi ulasan pengguna secara otomatis berdasarkan pola ulasan mereka?

C. Tujuan Penelitian

Penelitian ini bertujuan untuk mencapai beberapa tujuan utama, yaitu:

1. Mengimplementasikan model BERT untuk analisis sentimen ulasan pengguna pada *platform e-commerce*.
2. Menerapkan algoritma X-Means untuk segmentasi ulasan pengguna secara otomatis berdasarkan pola ulasan mereka.

D. Batasan Masalah

Penelitian ini memiliki ruang lingkup yang terbatas untuk memastikan fokus pada tujuan dan hasil yang optimal. Adapun batasan-batasan yang diterapkan dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini berfokus pada *platform e-commerce* Tokopedia dan data ulasan pengguna.
2. Penelitian ini menggunakan model BERT *Multilingual* (mBERT) untuk mengklasifikasikan ulasan pengguna ke dalam tiga kategori sentimen: positif, netral, dan negatif.
3. Segmentasi pengguna dilakukan menggunakan algoritma X-Means untuk menentukan jumlah klaster secara otomatis berdasarkan pola ulasan.

E. Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Manfaat bagi Peneliti

Penelitian ini memberikan kesempatan bagi peneliti untuk mendalami penerapan model BERT dan algoritma x-means, serta meningkatkan pemahaman tentang analisis sentimen dan segmentasi pengguna pada data ulasan platform e-commerce.

2. Manfaat bagi Pengguna

Hasil penelitian ini diharapkan dapat membantu *platform e-commerce*, khususnya pengguna tokopedia, dalam memahami pola perilaku dan kebutuhan konsumen melalui analisis sentimen dan segmentasi yang lebih akurat.