

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi dalam dunia mulai makin meluas dan mudah dipelajari [1]. Teknologi membawa perubahan yang signifikan dalam sektor kehidupan, termasuk dalam bidang perdagangan atau biasa disebut *e-commerce*, kemajuan ini dapat memungkinkan masyarakat untuk melakukan transaksi jual beli secara daring tanpa perlu mengeluarkan tenaga yang lebih dan melakukan dengan lebih mudah [2]. Yes setDi Indonesia, pertumbuhan *e-commerce* mengalami peningkatan yang pesat seiring berjalannya waktu dan seiring meningkatnya pengguna jumlah internet. Salah satu platform *e-commerce* yang berdiri di Indonesia dan cukup populer yaitu Tokopedia dan Shopee[3]. Platform ini menyediakan berbagai layanan transaksi secara daring yang memungkinkan berbagai pengguna untuk membeli produk bahkan pengguna memeberikan ulasan dan penilaian pada suatu produk [4]. Pertumbuhan ini juga didukung oleh data Bank Indonesia yang menunjukkan bahwa pada kuartal III tahun 2025 jumlah transaksi e-commerce meningkat sebesar 20,5% secara tahunan dengan total frekuensi transaksi mencapai 1,44 miliar kali dan nilai nominal sebesar Rp134,67 triliun. Selain itu, transaksi e-commerce di Indonesia diproyeksikan dapat mencapai Rp503 triliun pada tahun 2025, setelah sebelumnya mencapai Rp487 triliun pada tahun 2024[5].

Seiring meningkatnya jumlah pengguna e-commerce, jumlah ulasan yang dibuat oleh pengguna juga meningkat, dan jenis ulasan yang diberikan juga semakin beragam[6]. Pilihan Shopee dan Tokopedia sebagai subjek penelitian didasarkan pada dominasi keduanya di ekosistem pasar Indonesia. Shopee menempati posisi pertama dari platform e-commerce dengan 133,1 juta kunjungan bulanan, sedangkan Tokopedia menempati posisi kedua dengan 71,68 juta kunjungan bulanan.

Selain itu, Shopee menguasai 46% pangsa pasar nasional, sedangkan Tokopedia memegang 23% GMV. Ulasan ini memberikan pendapat, kritik, dan rekomendasi tentang fitur aplikasi, kualitas layanan, dan pengalaman pengguna saat menggunakan platform [7]. Namun, jumlah data yang sangat besar membuat analisis manual tidak efisien dan memerlukan waktu yang lama. Oleh karena itu, alat otomatis diperlukan untuk menilai opini pengguna dengan efektif [8]. Analisis sentimen, yang dilakukan menggunakan teknik pemrosesan bahasa natural (NLP), bertujuan untuk membagi pendapat pengguna menjadi kategori positif, negatif, atau netral. Ini adalah metode yang banyak digunakan. Analisis sentimen membantu organisasi memahami secara lebih sistematis dan berbasis data persepsi pengguna terhadap layanan mereka [9].

Analisis sentiment secara umum mempunyai keterbatasan karena hanya bisa memberikan opini secara keseluruhan terhadap objek yang tertera. Pendekatan ini belum mampu melakukan identifikasi aspek spesifik yang menjadi kepuasan atau ketidakpuasan bagi para pengguna [10]. Secara akademis, Shopee dan Tokopedia dapat menjadi subjek penelitian karena data ulasan pengguna dapat digunakan untuk mempelajari penggunaan pemrosesan bahasa alami, khususnya analisis sentimen berbasis aspek, dalam konteks e-commerce berbahasa Indonesia. Selain itu, ulasan pengguna dari kedua platform mencakup ulasan umum dan pendapat tentang hal-hal tertentu seperti produk, harga, pengiriman, dan fitur aplikasi. Dengan menggunakan metode ABSA, pengguna dapat melakukan analisa lebih rinci berdasarkan aspek[11].

Dalam penerapan analisis sentiment berbasis aspek, pemilihan metode klasifikasi merupakan factor penting yang dapat mempengaruhi hasil analisis. Machine learning dan Deep Learning menjadi salah satu cara metode yang digunakan dalam analisis sentiment untuk melihat dan melatih performa model menjadi baik dan sempurna [12]. Model *machine learning* yang dipakai adalah SVM yang bekerja mencari hyperlane optimal yang mampu memisahkan data ke dalam kelas dengan margin

maksimal [13]. Model berbasis machine learning tradisional masih memiliki keterbatasan. Hal ini disebabkan karena model tersebut bergantung pada representasi fitur seperti TF-IDF atau Bag of Words yang tidak mampu menangkap hubungan kontekstual antar kata secara mendalam [14].

Metode baru untuk memahami konteks bahasa yang lebih baik dihasilkan dari kemajuan dalam teknologi pembelajaran mendalam dalam bidang pemrosesan bahasa alami. IndoBERT, model bahasa berbasis transformer yang diadaptasi dari BERT yang dikembangkan oleh Google, adalah salah satu model yang paling banyak digunakan dalam pengolahan bahasa Indonesia [15]. Model ini telah dilatih untuk menangani dataset besar bahasa Indonesia. Dengan menggunakan mekanisme perhatian, IndoBERT dapat memahami hubungan kontekstual antar kata dalam sebuah kalimat. Ini memungkinkan model untuk memahami makna teks lebih dalam daripada pendekatan pembelajaran mesin konvensional. Akibatnya, IndoBERT dinilai dapat melakukan analisis sentimen berbasis teks dengan kinerja yang lebih baik [16].

Dalam konteks Aspect Based Sentiment Analysis pada ulasan pengguna e-commerce di Indonesia, penelitian yang secara khusus membandingkan kinerja model tradisional seperti SVM dengan model deep learning seperti IndoBERT masih relatif terbatas. Meningkatnya jumlah ulasan yang diberikan oleh pengguna aplikasi e-commerce yang tidak terstruktur dan memiliki berbagai macam pendapat menunjukkan kebutuhan akan analisis sentimen dalam penelitian ini. Ulasan pengguna tidak hanya berisi komentar umum positif atau negatif, tetapi juga keluhan, kritik, dan pengalaman khusus tentang hal-hal tertentu seperti harga, produk, pengiriman, dan sistem aplikasi. Analisis sentimen diperlukan untuk membantu mengolah opini pengguna secara sistematis, otomatis, dan berbasis data karena analisis ulasan secara manual akan membutuhkan waktu yang lama, berpotensi subjektif, dan sulit dilakukan secara konsisten pada jumlah data yang besar.

1.2 Rumusan Masalah

Pada penelitian ini, terdapat 3 Rumusan masalah yang menjadi pokok pada penelitian, yaitu :

1. Bagaimana proses penerapan Aspect-Based Sentiment Analysis pada ulasan pengguna aplikasi Tokopedia dan Shopee berdasarkan aspek harga, produk, pengiriman, dan sistem aplikasi?
2. Bagaimana hasil deployment model IndoBERT dan SVM dalam menampilkan hasil prediksi aspect based sentiment analisis pada ulasan pengguna aplikasi Tokopedia dan Shopee?
3. Bagaimana perbandingan performa model Support Vector Machine dan IndoBERT berdasarkan metrik evaluasi accuracy, precision, recall, dan F1-score?

1.3 Batasan Masalah

Pada penelitian terdapat Batasan masalah yang menjadi pokok pada penelitian, yaitu:

1. Data yang digunakan didapatkan dari *scrapping* aplikasi *play store* untuk lakukan analisis sentiment
2. Penelitian ini hanya berfokus pada perbandingan dua model machine learning dan deep learning yaitu SVM dan IndoBert dalam melakukan analisis sentimen, tanpa mempertimbangkan model machine learning atau deep learning lainnya sebagai pembanding.
3. Analisis sentimen yang dilakukan hanya mencakup tiga klasifikasi sentimen yaitu positif, negatif, dan netral, serta proses analisis menggunakan metode CRISP-DM serta dilakukan analisa aspek sentiment analyst (ABSA)

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

1. Untuk Penelitian :

- a. Mampu menerapkan metode Aspect Based Sentimen Analisis pada ulasan penggunaan aplikasi *e-commerce*.
- b. Dapat mengembangkan hasil deployment dari model IndoBERT dan SVM sebagai bentuk hasil sentiment analisis berdasarkan aspek untuk tujuan penelitian.
- c. Membandingkan model dalam melakukan sentiment analisis berbasis aspek berdasarkan metrik evaluasi.

1.4.2 Manfaat Penelitian

1. Manfaat bagi Pendidikan :

- a. Penelitian dijadikan sebagai refrensi bagi Pendidikan dalam mempelajari penerapan sentimen analisis berbasis aspek pada ulasan aplikasi *e-commerce* yaitu Tokopedia dan Shopee.
- b. Memberikan pemahaman mengenai model SVM dan IndoBERT dalam membantu memahami pendekatan model machine learning dan deep learning berbasis transformer dalam melakukan sentiment analisis berbasis aspek.
- c. Memberikan gambaran proses data preparation pada model yang berbeda bahwa proses data preparation pada SVM dan IndoBERT memiliki perbedaan, karena SVM membutuhkan representasi fitur seperti TF-IDF, sedangkan IndoBERT membutuhkan tokenisasi khusus berbasis transformer.

1.5 Sistematika Penulisan

Dokumentasi penelitian ini mengikuti susunan tertentu untuk memberikan struktur yang terorganisir dan mempermudah pemahaman. Struktur penulisan terbagi menjadi lima bab dengan fokus berbeda, sebagai berikut:

1. BAB I PENDAHULUAN

Pendahuluan mencakup latar belakang, perumusan masalah, batasan masalah, tujuan dan manfaat penelitian. Bab ini bertujuan untuk memberikan konteks dan alasan pentingnya penelitian ini dilakukan.

2. BAB II LANDASAN TEORI

Bagian ini menjelaskan teori yang mendukung penelitian termasuk framework dan algoritma yang digunakan. Teori ini bersumber dari jurnal atau artikel yang membahas metode atau penelitian terkait.

3. BAB III METODOLOGI PENELITIAN

Metodologi penelitian membahas objek penelitian, langkah-langkah yang digunakan dalam penelitian, teknik pengambilan data, dan penjelasan mengenai variabel penelitian.

4. BAB IV ANALISIS DAN HASIL PENELITIAN

Bab ini mencakup implementasi metodologi pada objek penelitian untuk mencapai tujuan penelitian. Di dalamnya terdapat penjelasan tentang proses pengolahan data, analisis data, dan hasil akhir yang diperoleh.

5. BAB V SIMPULAN DAN SARAN

Bagian ini berisi kesimpulan dari hasil penelitian beserta saran berdasarkan kendala atau analisis yang ditemukan. Saran untuk penelitian selanjutnya juga akan disampaikan untuk memberikan arah bagi penelitian di masa depan.

