

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Indonesia mengalami pertumbuhan signifikan dalam *e-commerce*, dengan proyeksi tingkat pertumbuhan sebesar 30,5% pada tahun 2024, hampir tiga kali lipat dari rata-rata global sebesar 10,4% [1]. Pertumbuhan ini disebabkan oleh tingginya jumlah pengguna internet dan pergeseran gaya hidup masyarakat, yang mengindikasikan transformasi dalam kebiasaan berbelanja yang difasilitasi oleh platform digital [2]. Platform *e-commerce* terkemuka seperti Shopee, Tokopedia, dan Bukalapak telah menjadi bagian penting dari pengalaman belanja online di Indonesia. Shopee, dengan 93 juta pengguna aktif, menunjukkan dominasi di pasar dan menekankan pentingnya memahami faktor-faktor yang mempengaruhi kepuasan pengguna [3]. Pertumbuhan pesat ini mencerminkan keberhasilan Shopee dan menyoroti kebutuhan untuk memahami sentimen pengguna terhadap aplikasi tersebut.

Analisis sentimen merupakan teknik yang sangat penting untuk memahami opini dan umpan balik pengguna, terutama dalam konteks ulasan aplikasi platform seperti *Google Play Store* [4]. Dalam era digital, opini dan sentimen pengguna tidak hanya mencerminkan pengalaman mereka tetapi juga merupakan aset berharga bagi perusahaan. Ulasan di platform seperti *Google Play Store* memberikan umpan balik langsung dari pengguna, yang dapat memberikan wawasan berharga tentang kekuatan dan kelemahan aplikasi. Khususnya untuk *e-commerce* seperti Shopee, memahami sentimen pengguna melalui ulasan ini sangat penting untuk meningkatkan layanan dan strategi pemasaran. Ulasan positif dapat memperkuat loyalitas pelanggan, sedangkan umpan balik negatif dapat memberikan informasi berharga untuk perbaikan aplikasi [5]. Dengan menggunakan metode analisis sentimen, perusahaan dapat memperoleh wawasan mengenai persepsi pengguna, yang dapat membantu dalam pengambilan keputusan terkait pengembangan produk dan strategi pemasaran.

Berbagai teknik telah diterapkan dalam analisis sentimen ulasan aplikasi. Algoritma klasik seperti *Naive Bayes* dan *Support Vector Machine (SVM)* yang sering digunakan, dengan hasil yang bervariasi tergantung pada dataset dan konteks aplikasi [5], [6]. Penelitian yang dilakukan oleh Ilmawan menunjukkan bahwa SVM sering kali memberikan akurasi yang lebih tinggi dibandingkan dengan *Naive Bayes*, dengan model SVM mencapai akurasi hingga 89,49%, sedangkan *Naive Bayes* mencapai 83,87% [5]. Penelitian selanjutnya yang dilakukan oleh Jannah juga menunjukkan bahwa SVM memiliki kinerja yang lebih baik dengan rata-rata akurasi sebesar 86,14%, dibandingkan dengan *Naive Bayes*. Hal ini mengindikasikan bahwa SVM lebih efektif dalam mengklasifikasikan sentimen *review* produk perawatan kecantikan di Shopee [6]. Meskipun metode ini dapat memberikan hasil yang baik, keterbatasan dalam menangkap konteks kalimat seringkali mengurangi akurasi, terutama dalam domain e-commerce. Sebagai solusi, metode seperti Bidirectional Long Short-Term Memory (BiLSTM) telah diperkenalkan untuk menangkap konteks data sekuensial dengan lebih efektif.

Metode BiLSTM adalah arsitektur jaringan saraf canggih yang memperluas kemampuan algoritma LSTM. Tidak seperti LSTM, yang memproses informasi dalam satu arah, BiLSTM dapat menganalisis data dalam arah maju dan mundur, sehingga memungkinkannya untuk menangkap informasi kontekstual yang lebih kaya dari data sekuensial [10]. Pemrosesan dua arah ini memungkinkan BiLSTM untuk unggul dalam tugas-tugas yang melibatkan *data sequential* atau deret waktu dengan meningkatkan kemampuannya untuk memahami dependensi dan pola dalam data. Dengan kemajuan dalam *Deep Learning*, metode BiLSTM telah menunjukkan potensi besar dalam meningkatkan akurasi analisis sentimen yang menawarkan keunggulan dalam menangkap konteks yang lebih kompleks dibandingkan dengan metode sebelumnya [7], [8], [9]. Namun, penelitian sebelumnya tentang BiLSTM umumnya dilakukan pada domain yang tidak langsung berkaitan dengan e-commerce, seperti analisis sentimen publik figur atau prediksi harga saham [7], [8].

Penelitian yang dilakukan oleh Komang, untuk menganalisis sentimen publik figur di Indonesia menggunakan data twitter berjumlah 1.034.329 tweet, dengan metode BiLSTM, dan menemukan bahwa sentimen negatif 45,98%, sentimen positif 28,04%, dan sentimen netral 25,98%, menunjukkan bahwa sentimen positif meningkat dengan konten relevan dan negatif dengan konten bertentangan. [7]. Penelitian lain yang dilakukan oleh Afrianto, penelitian ini membangun model prediksi harga saham menggunakan data historis dan sentimen publik dengan metode BiLSTM, menghasilkan MSE 0,094 dan RMSE 0,306, menunjukkan performa model yang efektif dalam memprediksi harga saham [8]. Penelitian ini menunjukkan keunggulan BiLSTM, tetapi penerapannya untuk menganalisis ulasan aplikasi Shopee masih jarang dilakukan, sehingga memunculkan kebutuhan untuk mengeksplorasi metode ini dalam domain e-commerce.

Berdasarkan permasalahan yang telah diuraikan dan temuan dari penelitian sebelumnya, penelitian ini mengusulkan model BiLSTM untuk analisis sentimen ulasan aplikasi Shopee. Penelitian ini akan menggunakan data komentar pengguna dari *Google Play Store* dan menerapkan pembobotan kata serta ekstraksi fitur melalui *Word2Vec*. Dengan pendekatan ini, diharapkan dapat diperoleh wawasan yang lebih mendalam mengenai sentimen pengguna terhadap aplikasi Shopee, serta memberikan kontribusi pada pengembangan aplikasi *e-commerce* di Indonesia.

1.2 Rumusan Masalah

Rumusan masalah dari penelitian ini adalah Bagaimana implementasi model BiLSTM dengan pembobotan kata dan ekstraksi fitur menggunakan *Word2Vec* untuk analisis sentimen pada ulasan aplikasi Shopee?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah Mengimplementasikan model BiLSTM dengan pembobotan kata dan ekstraksi fitur menggunakan *Word2Vec* untuk analisis sentimen pada ulasan aplikasi Shopee.

1.4 Batasan Masalah

Batasan masalah dari penelitian ini adalah sebagai berikut:

- a. Data yang digunakan diambil dari ulasan aplikasi Shopee pada *Google Play Store* dengan teknik *scraping* menggunakan Bahasa Indonesia, dari 01 Juli hingga 08 Agustus 2024, dengan jumlah total 5000 data.
- b. Data yang diambil terdiri dari ulasan yang dikategorikan sebagai positif, netral, dan negatif.
- c. Metode analisis yang digunakan dalam penelitian ini adalah BiLSTM.
- d. Pengolahan analisis sentimen dilakukan menggunakan *Google Colaboratory* dengan Bahasa Pemrograman *Python*.

