

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pariwisata merupakan sebuah ekosistem yang melibatkan pergerakan individu ke destinasi di luar lingkungan domestik mereka guna mengejar tujuan rekreasi dan relaksasi. Fenomena ini menggabungkan niat perjalanan personal dengan ketergantungan pada sektor layanan komersial, yang bertujuan untuk memberikan pengalaman menyenangkan bagi para pelaku perjalanan tersebut [1]. Berdasarkan data dari Badan Pusat Statistik (BPS) Provinsi Sumatera Utara, jumlah wisatawan mancanegara (wisman) yang berkunjung pada periode Januari hingga Desember 2024 mencapai 250.413 kunjungan, meningkat sebesar 26,32% dibandingkan tahun sebelumnya yang berjumlah 198.240 kunjungan [2]. Peningkatan jumlah wisatawan yang berkunjung ke Sumatera Utara menunjukkan bahwa sektor pariwisata di daerah ini terus berkembang dan menarik minat wisatawan mancanegara. Namun, kenaikan kunjungan tidak selalu mencerminkan tingkat kepuasan dan loyalitas wisatawan terhadap destinasi yang mereka datangi.

Di era digital saat ini, ulasan daring telah menjadi sumber informasi yang sangat berpengaruh bagi wisatawan dalam menentukan destinasi yang akan mereka kunjungi [3]. Hal ini karena wisatawan cenderung mempercayai pengalaman dan pendapat pengunjung lain yang dibagikan melalui *platform* seperti *Google Maps* untuk memperoleh gambaran yang lebih jelas sebelum memutuskan tempat wisata tujuan mereka [4]. Banyaknya ulasan yang tersedia di *Google Maps* memiliki isi yang beragam dan sentimen yang berbeda terhadap berbagai aspek dalam satu ulasan yang sama [5]. Kondisi ini membuat wisatawan kesulitan memperoleh informasi yang detail mengenai kelebihan atau kekurangan pada aspek tertentu dan pengelola wisata tidak mendapatkan umpan balik yang terperinci, padahal ulasan wisatawan juga berperan penting sebagai sumber informasi yang dapat membantu pengelola merancang program pengembangan pariwisata [6]. Oleh karena itu, penerapan analisis sentimen berbasis aspek menjadi solusi yang tepat untuk mengolah ulasan destinasi wisata, karena setiap ulasan sering kali mengandung berbagai aspek dengan sentimen yang berbeda-beda, sehingga metode ini mampu memberikan pemahaman yang lebih spesifik dan mendalam terhadap persepsi wisatawan [7].

Beberapa penelitian terdahulu mengenai analisis sentimen telah menunjukkan keunggulan signifikan dalam penggunaan algoritma *Deep Learning*, khususnya arsitektur

Long Short-Term Memory (LSTM). Penelitian [8] menunjukkan bahwa model LSTM sangat optimal dalam mengevaluasi persepsi wisatawan pada *TripAdvisor* dengan capaian akurasi 96,71% dan *loss* 14,22% melalui klasifikasi opini yang dibagi menjadi dua kelas yaitu positif dan negatif. Temuan ini didukung oleh penelitian Abdurahmonov dkk. terkait analisis sentimen cuaca ekstrem di Twitter, di mana model LSTM menghasilkan akurasi sebesar 98,94% pada klasifikasi aspek dan 97,53% pada klasifikasi sentimen [9]. Sejalan dengan hal tersebut, penelitian pada media sosial berbahasa Arab menggunakan dataset ASAD (55.000 tweet) juga mengonfirmasi efektivitas LSTM sebagai sinyal *time series*, di mana model *triple-layer* LSTM dengan *word embedding FastText* berhasil mencapai akurasi tertinggi sebesar 90,9% [10]. Secara keseluruhan, berbagai literatur tersebut mengonfirmasi bahwa penambahan lapisan pada arsitektur LSTM serta penggunaan teknik *word embedding FastText* mampu meningkatkan akurasi klasifikasi opini secara signifikan dalam menangani data teks yang kompleks.

Penelitian ini melakukan pendekatan Analisis Sentimen Berbasis Aspek dengan memanfaatkan kombinasi *FastText* untuk ekstraksi fitur teks dan LSTM untuk klasifikasi sentimen. Karakteristik ulasan wisatawan Indonesia ditandai dengan ragam bahasa informal, seperti penggunaan kata slang, singkatan, ejaan tidak baku, dan kesalahan ketik seperti "gk enak dn tmptny jorok", "hotelnya mantul", "pelayanannya jelek bgt". Pemilihan *FastText* didasarkan pada kemampuannya dalam mengenali kata langka, kata yang tidak terdaftar (*out-of-vocabulary*), variasi penulisan, serta bentuk morfologis dengan lebih baik, sehingga memberikan pemahaman teks yang lebih mendalam [11]. Selain itu, dalam analisis sentimen berbasis aspek, satu ulasan sering kali memuat beberapa aspek dengan sentimen yang berbeda, seperti "Pemandangannya mantul tapi makanannya gak enak". Situasi ini memerlukan model yang dapat memahami konteks baik antar kalimat maupun antar aspek yang ada. LSTM memiliki kemampuan untuk mempertahankan dan memperbarui informasi kontekstual jangka panjang, sehingga efektif dalam menangkap perubahan sentimen antar aspek [12].

Berdasarkan tingginya jumlah ulasan wisata yang bersifat tekstual dan beragam aspek penilaian di dalamnya, diperlukan pendekatan analisis sentimen berbasis aspek untuk mengidentifikasi opini wisatawan secara lebih akurat. Oleh karena itu, penelitian ini mengangkat judul **“Analisis Sentimen Berbasis Aspek terhadap Ulasan Tempat Wisata pada Provinsi Sumatera Utara Menggunakan FastText dan LSTM.”**

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah penelitian ini adalah bahwa ulasan destinasi wisata memuat berbagai aspek dan sentimen berbeda dalam satu ulasan yang sama.

1.3 Tujuan

Tujuan dari penelitian ini adalah mengembangkan dan mengimplementasikan aplikasi berbasis *website* untuk mengidentifikasi aspek dan mengklasifikasikan sentimen pada ulasan destinasi wisata di *Google Maps* menggunakan *FastText* dan LSTM.

1.4 Manfaat

Adapun manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Penelitian ini memberikan informasi berupa aspek dan sentimen yang dapat menjadi dasar evaluasi bagi wisatawan dan pengelola wisata dalam mengambil tindakan, seperti menentukan destinasi dan melakukan pengembangan destinasi wisata.
2. Penelitian ini diharapkan dapat memperkaya literatur dalam bidang analisis sentimen berbasis aspek, khususnya yang menggunakan metode *FastText* dan LSTM.

1.5 Ruang Lingkup

Agar pembahasan dalam tugas akhir ini lebih terarah dan mencegah adanya perluasan masalah serta pembahasan yang terlalu kompleks, maka ruang lingkup terhadap tugas akhir ini ditetapkan sebagai berikut:

1. Dataset penelitian ini terdiri dari 8.506 ulasan yang dikumpulkan melalui proses *crawling* data dari *Google Maps Reviews*. Proses *crawling* data dilakukan dengan bantuan *tool Apify*.
2. Seluruh proses pelatihan dan pengujian model dibatasi pada penggunaan dataset berbahasa Indonesia, sehingga penelitian ini tidak mencakup analisis sentimen pada dataset berbahasa lain.
3. Aspek-aspek pariwisata yang menjadi target klasifikasi dalam ABSA dibatasi pada lima kategori utama: pelayanan, fasilitas, harga, daya tarik, kebersihan dan aksesibilitas.
4. Hasil klasifikasi yang dihasilkan hanya dikategorikan berdasarkan aspek dengan sentimen yang bersifat positif dan negatif.

5. Kinerja model klasifikasi pada penelitian ini diukur menggunakan *confusion matrix* beserta metrik turunannya, yaitu *accuracy*, *precision*, *recall*, dan *f1-score*.

