

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sektor pariwisata menempati posisi yang cukup strategis dalam perekonomian Indonesia. Selain menjadi salah satu sumber devisa negara, sektor ini juga turut membuka lapangan kerja dan menggerakkan berbagai bidang usaha yang menyertainya. Merujuk data World Travel & Tourism Council (WTTC), kontribusi pariwisata terhadap Produk Domestik Bruto (PDB) Indonesia pada tahun 2025 mencapai sekitar 5,5%, atau setara dengan Rp1.269,8 triliun, dengan daya serap tenaga kerja mencapai lebih dari 14 juta orang atau sekitar 9,3% dari total angkatan kerja nasional [1]. Pada periode yang sama, wisatawan mancanegara membelanjakan sekitar Rp344 triliun, sementara wisatawan domestik mencatatkan pengeluaran sebesar Rp381,4 triliun [1]. Selain menjadi penyumbang devisa, angka-angka tersebut juga memperlihatkan bagaimana pariwisata ikut menopang sektor-sektor lain seperti transportasi, akomodasi, kuliner, dan ekonomi kreatif. Kebun Binatang Surabaya menjadi salah satu contoh destinasi yang merepresentasikan potensi pariwisata tersebut pada skala daerah.

Sebagai salah satu destinasi wisata edukasi dan konservasi tertua di Jawa Timur, Kebun Binatang Surabaya (KBS) tidak hanya berperan sebagai tempat rekreasi, melainkan juga sebagai sarana edukasi dan pelestarian satwa. KBS telah berdiri sejak tahun 1916 dan terus dikembangkan melalui penambahan berbagai fasilitas guna menunjang pengalaman berkunjung wisatawan [2]. Beberapa kajian sebelumnya memperlihatkan bahwa kualitas layanan, kelengkapan sarana prasarana, serta keragaman koleksi satwa memiliki pengaruh yang cukup besar terhadap tingkat kepuasan pengunjung, dan menjadi salah satu indikator keberlanjutan pengelolaan destinasi wisata [3]. Di sisi lain, penelitian mengenai komunikasi dan strategi promosi KBS turut menunjukkan bahwa pemanfaatan media sosial secara tepat berkontribusi dalam mempertahankan citra KBS sebagai destinasi warisan budaya sekaligus daya

tarik wisata yang tetap relevan pada era digital [4]. Kondisi ini menjadikan pemahaman terhadap opini dan persepsi pengunjung sebagai hal yang cukup penting untuk terus dikaji dalam upaya pengembangan kualitas destinasi.

Salah satu pendekatan yang dapat dimanfaatkan untuk memahami opini tersebut adalah analisis sentimen, yaitu cabang text mining yang berupaya mengenali polaritas suatu teks apakah cenderung bermuatan positif, negatif, atau netral [5]. Pada konteks pariwisata, pendekatan ini menjadi relevan karena ulasan wisatawan yang tersebar di berbagai platform digital, seperti Google Maps, TripAdvisor, maupun media sosial, sesungguhnya memuat pengalaman nyata pengunjung terhadap suatu destinasi. Beberapa studi memperlihatkan bahwa analisis sentimen atas ulasan wisata dapat membantu mengungkap tingkat kepuasan wisatawan sekaligus memberikan gambaran mengenai berbagai aspek destinasi yang dinilai baik maupun yang masih perlu diperbaiki [6]. Dengan bantuan teknik text mining dan machine learning, data ulasan dalam jumlah besar dapat diolah secara lebih efisien dan objektif dibandingkan jika dilakukan secara manual, sehingga hasilnya dapat menjadi masukan yang berguna bagi pengelola destinasi dalam mengambil keputusan [7].

Agar proses analisis sentimen tersebut dapat berjalan optimal, dibutuhkan cara merepresentasikan teks ke dalam bentuk numerik yang tetap mempertahankan makna kata secara memadai. Salah satu pendekatan yang banyak digunakan untuk keperluan ini adalah Word2Vec, sebuah metode word embedding yang memetakan kata ke dalam vektor numerik berdasarkan konteks kemunculannya, sehingga relasi semantik antar kata dapat lebih tertangkap dibandingkan metode berbasis frekuensi seperti TF-IDF [8]. Studi yang membandingkan sejumlah metode machine learning konvensional seperti Logistic Regression, Naïve Bayes, dan SVM dengan pendekatan deep learning berbasis Word2Vec pada ulasan hotel menunjukkan bahwa kombinasi Word2Vec dan deep learning cenderung menghasilkan akurasi yang lebih baik [8]. Word2Vec sendiri memiliki dua varian arsitektur utama, yaitu Continuous Bag of Words (CBOW) dan Skip-gram, yang bekerja dengan cara berbeda dalam mempelajari konteks kata,

sehingga menjadi menarik untuk dibandingkan performanya pada tugas analisis sentimen.

Selain pemilihan metode representasi teks, model klasifikasi yang digunakan juga sangat menentukan performa analisis sentimen. Bidirectional Long Short-Term Memory (BiLSTM) dipilih dalam penelitian ini karena kemampuannya memahami konteks kalimat secara dua arah, yaitu dari awal ke akhir dan dari akhir ke awal, sehingga mampu menangkap hubungan antar kata yang lebih kompleks dibandingkan model sekuensial satu arah seperti LSTM biasa. Kemampuan ini membuat representasi makna teks menjadi lebih lengkap, terutama pada data ulasan yang memiliki struktur bahasa alami dan panjang kalimat yang bervariasi. Penelitian sebelumnya menunjukkan bahwa BiLSTM mampu memberikan performa yang lebih baik dibandingkan LSTM karena dapat menangkap konteks temporal secara lebih efektif serta meningkatkan akurasi pada tugas analisis sentimen [9]. Pada studi analisis sentimen berbasis deep learning lainnya, BiLSTM juga terbukti mampu merepresentasikan informasi kontekstual secara lebih optimal dan meningkatkan kinerja klasifikasi dibandingkan pendekatan sekuensial satu arah [10]. Oleh karena itu, penggunaan BiLSTM dinilai relevan untuk penelitian ini yang berfokus pada analisis sentimen ulasan wisata, karena model ini dapat memahami pola bahasa dan hubungan antar kata secara lebih mendalam.

Permasalahan lain yang umum dijumpai pada data ulasan daring adalah ketidakseimbangan jumlah data antar kelas sentimen (class imbalance), di mana jumlah ulasan positif jauh lebih banyak dibandingkan ulasan negatif. Kondisi ini dapat menyebabkan model klasifikasi menjadi bias terhadap kelas mayoritas sehingga performanya kurang optimal dalam mengenali kelas minoritas. Untuk mengatasi permasalahan tersebut, penelitian ini menerapkan teknik Synthetic Minority Over-sampling Technique (SMOTE) pada tahap pembagian data guna menyeimbangkan jumlah data antar kelas sebelum proses pelatihan model dilakukan.

Meskipun berbagai penelitian telah menerapkan Word2Vec dan model deep learning untuk analisis sentimen, kajian yang secara khusus membandingkan arsitektur CBOW dan Skip-gram menggunakan BiLSTM pada data ulasan wisata lokal berbasis Google Maps masih terbatas. Sebagian besar penelitian terdahulu lebih berfokus pada satu konfigurasi embedding atau hanya membandingkan algoritma klasifikasi, sehingga perbandingan langsung antara kedua arsitektur Word2Vec pada objek wisata edukasi seperti Kebun Binatang Surabaya belum banyak dikaji. Oleh karena itu, penelitian ini dilakukan untuk mengisi celah tersebut dengan menganalisis sentimen ulasan Kebun Binatang Surabaya pada Google Maps serta membandingkan performa Word2Vec CBOW dan Skip-gram yang diintegrasikan dengan model BiLSTM, sehingga dapat diketahui kombinasi metode yang paling optimal dalam mengklasifikasikan sentimen pengunjung dan memahami persepsi mereka terhadap destinasi wisata tersebut.

1.2 Perumusan Masalah

1.2.1 Permasalahan

Ulasan wisatawan terhadap Kebun Binatang Surabaya yang tersebar di platform daring seperti Google Maps mengandung banyak informasi penting mengenai pengalaman berkunjung, namun belum dimanfaatkan secara optimal oleh pengelola untuk peningkatan kualitas layanan. Selain itu, belum banyak penelitian yang secara khusus mengevaluasi arsitektur word embedding mana, antara CBOW atau Skip-gram, yang paling optimal ketika dipadukan dengan model BiLSTM untuk mengklasifikasikan sentimen pada data ulasan wisata lokal berbahasa Indonesia yang cenderung tidak baku dan tidak seimbang antar kelas.

1.2.2 Metode Usulan

Metode yang diusulkan adalah penerapan analisis sentimen terhadap ulasan pengunjung Kebun Binatang Surabaya menggunakan representasi teks Word2Vec dengan dua arsitektur, yaitu CBOW dan Skip-gram, yang selanjutnya diklasifikasikan menggunakan model Bidirectional Long Short-Term Memory

(BiLSTM). Data ulasan dari Google Maps diproses melalui tahapan preprocessing teks, pelabelan sentimen berbasis leksikon, penanganan ketidakseimbangan data menggunakan SMOTE, pelatihan model, serta evaluasi menggunakan confusion matrix dengan metrik accuracy, precision, recall, F1-score, dan hamming loss guna menghasilkan perbandingan performa yang objektif antara CBOW dan Skip-gram.

1.2.3 Pertanyaan Penelitian

1. Bagaimana hasil analisis sentimen terhadap ulasan Kebun Binatang Surabaya pada Google Maps menggunakan Word2Vec dan Bidirectional Long Short-Term Memory (BiLSTM)?
2. Bagaimana perbandingan performa Word2Vec Continuous Bag of Words (CBOW) dengan Word2Vec Skip-gram dalam mengklasifikasikan sentimen menggunakan model BiLSTM?

1.3 Tujuan dan Manfaat

1.3.1 Tujuan

1. Menganalisis sentimen ulasan pengunjung Kebun Binatang Surabaya pada Google Maps menggunakan Word2Vec dan Bidirectional Long Short-Term Memory (BiLSTM).
2. Membandingkan performa ekstraksi fitur Word2Vec Continuous Bag of Words (CBOW) dan Word2Vec Skip-gram dalam klasifikasi sentimen menggunakan BiLSTM berdasarkan metrik accuracy, precision, recall, dan F1-score.

1.3.2 Manfaat

1. Menjadi referensi akademik terkait penerapan Word2Vec dan Bidirectional Long Short-Term Memory (BiLSTM) dalam analisis sentimen di bidang pariwisata, khususnya terkait perbandingan arsitektur CBOW dan Skip-gram.
2. Memberikan wawasan kepada pengelola wisata mengenai opini dan tingkat kepuasan wisatawan terhadap Kebun Binatang Surabaya yang dapat dijadikan

dasar pengambilan keputusan strategis dalam meningkatkan kualitas layanan dan pengelolaan destinasi wisata.

1.4 Batasan Masalah

Penelitian ini memiliki sejumlah batasan untuk memperjelas ruang lingkup dan fokus pembahasan, sehingga hasil yang diperoleh lebih terarah dan sesuai tujuan penelitian. Batasan-batasan tersebut adalah sebagai berikut:

1. Data ulasan diambil dari platform Google Maps terkait Kebun Binatang Surabaya melalui teknik web scraping.
2. Data mentah hasil scraping berjumlah 5.000 ulasan, yang setelah proses cleaning dan seleksi menghasilkan 2.583 data ulasan valid yang digunakan dalam penelitian.
3. Sentimen yang dianalisis hanya terdiri dari dua kelas, yaitu positif dan negatif, pada level dokumen (setiap ulasan diberi satu label sentimen, bukan berbasis aspek).
4. Pelabelan sentimen dilakukan secara otomatis menggunakan pendekatan berbasis leksikon (lexicon-based).
5. Ekstraksi fitur teks menggunakan Word2Vec dengan dua skenario arsitektur, yaitu Continuous Bag of Words (CBOW) dan Skip-gram.
6. Penanganan ketidakseimbangan data (class imbalance) dilakukan menggunakan Synthetic Minority Over-sampling Technique (SMOTE) pada data latih.
7. Model klasifikasi yang digunakan adalah Bidirectional Long Short-Term Memory (BiLSTM).

1.5 Metodologi Penulisan Skripsi

Penulisan skripsi ini disusun melalui beberapa tahapan, yaitu studi literatur untuk memperoleh landasan teori yang relevan, pengumpulan dan pengolahan data ulasan Kebun Binatang Surabaya, perancangan dan implementasi model analisis sentimen menggunakan Word2Vec dan BiLSTM, pengujian serta evaluasi performa model, hingga penarikan kesimpulan dan penyusunan laporan secara sistematis sesuai dengan

Pedoman Penyusunan Laporan Skripsi Program Studi Sistem Informasi Universitas Trunodjoyo Madura.

1.6 Sistematika Penulisan

BAB I. PENDAHULUAN

Bab ini membahas latar belakang penelitian mengenai analisis sentimen terhadap ulasan wisatawan Kebun Binatang Surabaya, perumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, metodologi penulisan, serta sistematika penulisan yang digunakan dalam penelitian ini.

BAB II. TINJAUAN PUSTAKA

Bab ini mencakup penjelasan tentang landasan teori yang meliputi konsep analisis sentimen, pemrosesan bahasa alami, Word2Vec (CBOW dan Skip-gram), pelabelan sentimen berbasis leksikon, SMOTE, Long Short-Term Memory (LSTM) dan Bidirectional Long Short-Term Memory (BiLSTM), metrik evaluasi model, serta tinjauan penelitian terdahulu yang relevan.

BAB III. METODE PENELITIAN

Bab ini membahas tentang gambaran umum Kebun Binatang Surabaya, sumber dan karakteristik data, serta tahapan penelitian yang meliputi pengumpulan data, pelabelan sentimen berbasis leksikon, preprocessing teks, ekstraksi fitur dengan Word2Vec (CBOW dan Skip-gram), penanganan ketidakseimbangan data dengan SMOTE, pelatihan model BiLSTM, dan evaluasi kinerja menggunakan confusion matrix.

BAB IV. HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil pengolahan data dan pembahasan yang mencakup hasil preprocessing, hasil pelabelan sentimen, hasil klasifikasi menggunakan BiLSTM dengan Word2Vec CBOW dan Skip-gram, perbandingan performa kedua skenario,

serta interpretasi temuan yang dapat digunakan untuk pengembangan destinasi wisata Kebun Binatang Surabaya.

BAB V. KESIMPULAN DAN SARAN

Bab ini memuat kesimpulan yang menjawab tujuan penelitian serta saran yang dapat diberikan kepada pihak terkait maupun penelitian selanjutnya berdasarkan hasil analisis sentimen yang telah dilakukan.

DAFTAR PUSTAKA

Memuat daftar jurnal, buku, dan sumber referensi lain yang digunakan sebagai dasar dalam penyusunan penelitian ini.