

Penerapan *Natural Language Processing (NLP)* Dalam Sistem Rekomendasi Wisata Alam Indonesia Berbasis Sentimen Pengunjung

Aulia Hafiz^{1*}, Martiano

¹Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara, Medan, Indonesia

Article Info

Article history:

Received Oktober 6, 2026

Revised Oktober 6, 2026

Accepted Oktober 6, 2026

Kata Kunci:

Natural Language Processing,
Analisis Sentimen,
TF-IDF,
Support Vector Machine,
Sistem Rekomendasi

Keywords:

*Natural Language Processing,
Sentiment Analysis,
TF-IDF,
Support Vector Machine,
Recommendation System*

ABSTRAK

Perkembangan teknologi digital telah mengubah cara wisatawan memperoleh informasi sebelum melakukan perjalanan. Kondisi tersebut dapat menyulitkan calon wisatawan dalam memperoleh informasi yang relevan karena harus membaca dan memahami sejumlah besar ulasan secara manual. Penelitian ini bertujuan menerapkan Natural Language Processing (NLP) dalam sistem rekomendasi wisata alam Indonesia berbasis sentimen pengunjung, mengklasifikasikan ulasan ke dalam sentimen positif, negatif, dan netral, serta memanfaatkan hasil klasifikasi sebagai dasar pemeringkatan destinasi wisata. Data penelitian berupa 70 ulasan pengunjung dari 11 destinasi wisata yang diperoleh dari sumber digital. Tahapan pengolahan data meliputi seleksi dan pembersihan data, pelabelan sentimen, case folding, cleaning dan normalisasi, tokenizing, stopword removal, stemming, serta pembobotan Term Frequency-Inverse Document Frequency (TF-IDF). Klasifikasi sentimen dilakukan menggunakan algoritma Support Vector Machine (SVM) dengan kernel linear. Hasil pengujian menggunakan pembagian data 80:20 menunjukkan bahwa model memperoleh accuracy sebesar 85,71%. Sementara itu, hasil validasi silang stratified 4-fold menghasilkan rata-rata accuracy sebesar 91,42% dengan standar deviasi 2,87%. Hasil penelitian menunjukkan bahwa penerapan NLP dengan TF-IDF dan SVM dapat digunakan untuk mengolah ulasan pengunjung dan menghasilkan rekomendasi destinasi berdasarkan kecenderungan sentimen.

ABSTRACT

The development of digital technology has resulted in an increasing number of visitor reviews being available on various tourism platforms. This condition can make it difficult for potential tourists to obtain relevant information because they have to read and understand a large number of reviews manually. This study aims to apply Natural Language Processing (NLP) in a visitor sentiment-based recommendation system for natural tourism destinations in Indonesia, classify reviews into positive, negative, and neutral sentiments, and utilize the classification results as the basis for ranking tourism destinations. The research data consisted of 70 visitor reviews from 11 tourism destinations obtained from digital sources. The data processing stages included data selection and cleaning, sentiment labeling, case folding, cleaning and normalization, tokenizing, stopword removal, stemming, and Term Frequency-Inverse Document Frequency (TF-IDF) weighting. Sentiment classification was performed using the Support Vector Machine (SVM) algorithm with a linear kernel. The results of the 80:20 data split showed that the model achieved an accuracy of 85.71%. Meanwhile, the stratified 4-fold cross-validation achieved an average accuracy of 91.42% with a standard deviation of 2.87%. The results indicate that the application of NLP using TF-IDF and SVM can be used to process visitor reviews and generate destination recommendations based on sentiment tendencies.



Corresponding Author:

Aulia Hafiz
Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara
Medan, Indonesia
Email: hafizaulia260@gmail.com

1. PENDAHULUAN

Sektor pariwisata merupakan salah satu sektor yang memiliki peranan penting dalam mendukung aktivitas ekonomi Indonesia. Indonesia memiliki kekayaan alam dan keragaman geografis yang menghasilkan berbagai jenis objek daya tarik wisata, termasuk wisata alam, wisata tirta, dan wisata buatan. Keberadaan objek daya tarik tersebut tersebar di berbagai wilayah Indonesia dan menjadi salah satu potensi utama dalam pengembangan industri pariwisata nasional. Badan Pusat Statistik menunjukkan bahwa objek wisata alam, tirta, dan buatan merupakan jenis objek daya tarik wisata yang dominan di Indonesia. Pada tahun 2024, jumlah objek daya tarik wisata juga mengalami pertumbuhan, sementara pendapatan dari objek daya tarik wisata mencapai Rp11,63 triliun. Kondisi tersebut menunjukkan bahwa pengembangan dan pengelolaan destinasi wisata, termasuk wisata alam, memiliki peranan penting dalam mendukung keberlanjutan sektor pariwisata Indonesia [1].

Besarnya potensi pariwisata Indonesia juga tercermin dari tingginya aktivitas perjalanan wisata. Data Badan Pusat Statistik menunjukkan bahwa perjalanan wisatawan nusantara dan kunjungan wisatawan mancanegara merupakan aktivitas yang terus dipantau sebagai bagian penting dari perkembangan pariwisata nasional. Statistik Wisatawan Nusantara 2024 menunjukkan bahwa perjalanan wisatawan nusantara memiliki karakteristik yang beragam berdasarkan tujuan perjalanan, lama perjalanan, akomodasi, serta pengeluaran wisatawan. Sementara itu, Statistik Kunjungan Wisatawan Mancanegara 2024 mencatat karakteristik kunjungan berdasarkan pintu masuk, kebangsaan, dan pola kunjungan. Hal tersebut menunjukkan bahwa Indonesia memiliki pasar wisata yang luas sehingga diperlukan penyediaan informasi destinasi yang dapat membantu wisatawan menentukan pilihan perjalanan secara lebih efektif [2], [3].

Perkembangan teknologi digital telah mengubah cara wisatawan memperoleh informasi sebelum melakukan perjalanan. Informasi mengenai destinasi yang sebelumnya diperoleh melalui sumber konvensional kini semakin banyak tersedia melalui internet, media sosial, platform perjalanan, dan layanan digital lainnya. Wisatawan dapat membandingkan berbagai destinasi berdasarkan foto, penilaian, komentar, lokasi, fasilitas, serta pengalaman wisatawan sebelumnya. Dalam konteks pariwisata, ulasan daring menjadi salah satu sumber informasi yang digunakan wisatawan dalam merencanakan perjalanan dan menentukan pilihan destinasi. Penelitian menunjukkan bahwa ulasan perjalanan daring digunakan wisatawan untuk merencanakan dan memesan perjalanan, sehingga keberadaan informasi tersebut memiliki peran dalam proses pengambilan keputusan wisatawan [4].

Ulasan pengunjung memiliki karakteristik yang berbeda dengan informasi promosi yang disediakan oleh pengelola destinasi. Ulasan merupakan bentuk user-generated content yang dapat berisi pengalaman, penilaian, pendapat, emosi, maupun rekomendasi berdasarkan pengalaman langsung pengunjung. Informasi tersebut dapat memberikan gambaran yang lebih rinci mengenai pengalaman wisatawan terhadap suatu destinasi. Penelitian Bigne et al. [5] menunjukkan bahwa ulasan daring dapat menggambarkan pengalaman wisatawan pada berbagai aspek destinasi dan mengandung informasi yang tidak selalu tercermin hanya melalui nilai rating bintang. Bahkan, isi tekstual sebuah ulasan dapat

mengandung aspek positif dan negatif secara bersamaan sehingga memberikan informasi yang lebih kaya mengenai pengalaman pengunjung.

Dari sisi pengelola destinasi wisata, ulasan pengunjung juga merupakan sumber informasi yang penting untuk mengevaluasi kualitas layanan dan pengalaman wisatawan. Ulasan dapat digunakan untuk mengidentifikasi aspek yang dianggap memuaskan maupun aspek yang masih memerlukan perbaikan. Informasi mengenai fasilitas, aksesibilitas, kebersihan, harga, pelayanan, serta kondisi lingkungan dapat muncul secara langsung dalam teks ulasan. Penelitian mengenai analisis ulasan pariwisata menunjukkan bahwa pengolahan ulasan secara otomatis dapat membantu mengidentifikasi topik, aspek, serta sentimen yang berkaitan dengan kekuatan dan kelemahan suatu destinasi [6]. Oleh karena itu, pengolahan ulasan tidak hanya bermanfaat bagi wisatawan, tetapi juga dapat digunakan sebagai sumber informasi evaluasi bagi pengelola destinasi.

Jumlah ulasan yang besar dan bentuk data yang tidak terstruktur menyebabkan analisis secara manual menjadi kurang efektif. Ulasan yang tersedia dalam bentuk teks bebas perlu diproses terlebih dahulu agar dapat dikelompokkan berdasarkan informasi yang terkandung di dalamnya. Salah satu teknologi yang dapat digunakan untuk menangani permasalahan tersebut adalah Natural Language Processing (NLP). NLP memungkinkan komputer melakukan pemrosesan terhadap bahasa manusia dalam bentuk teks sehingga informasi yang sebelumnya tidak terstruktur dapat dianalisis secara sistematis. Dalam penelitian pariwisata terkini, NLP telah digunakan untuk melakukan analisis sentimen dan ekstraksi kata kunci dari ulasan wisatawan untuk memperoleh informasi mengenai preferensi, kepuasan, serta persepsi pengunjung terhadap suatu layanan atau destinasi [7].

Salah satu penerapan NLP yang relevan dengan penelitian ini adalah analisis sentimen. Analisis sentimen digunakan untuk mengidentifikasi kecenderungan opini atau polaritas yang terkandung dalam teks, seperti positif, negatif, atau netral. Dalam konteks pariwisata, analisis sentimen dapat digunakan untuk mengolah ulasan pengunjung sehingga opini yang awalnya berupa teks bebas dapat diubah menjadi informasi terstruktur. Bigne et al. [5] menerapkan NLP dan machine learning terhadap lebih dari 20.000 ulasan TripAdvisor untuk mengidentifikasi komponen positif dan negatif dalam pengalaman wisatawan. Penelitian tersebut menunjukkan bahwa analisis sentimen dapat digunakan untuk memahami hubungan antara isi ulasan, pengalaman wisatawan, dan rating yang diberikan terhadap suatu destinasi [13], [14].

Dalam penelitian analisis sentimen, salah satu metode yang dapat digunakan untuk melakukan klasifikasi teks adalah Support Vector Machine (SVM). SVM merupakan metode machine learning yang dapat digunakan untuk membedakan data ke dalam beberapa kelas berdasarkan karakteristik fitur yang dimiliki. Penggunaan SVM dalam analisis sentimen pada data ulasan perjalanan juga telah diteliti. Nugroho et al. [8], misalnya, membandingkan SVM dan Random Forest dalam melakukan klasifikasi sentimen pada ulasan aplikasi Tiket.com dengan pembobotan Delta TF-IDF. Hasil penelitian menunjukkan bahwa SVM menghasilkan kinerja klasifikasi yang baik pada data ulasan tersebut, dengan akurasi sebesar 94,33%. Hasil tersebut menunjukkan bahwa SVM memiliki potensi untuk digunakan dalam klasifikasi sentimen pada data tekstual yang berkaitan dengan layanan perjalanan dan pariwisata [8].

Hasil analisis sentimen selanjutnya dapat dikembangkan menjadi dasar dalam sistem rekomendasi wisata. Sistem rekomendasi pada dasarnya bertujuan membantu pengguna menemukan pilihan yang sesuai dengan preferensi tertentu dari sejumlah alternatif yang tersedia. Dalam konteks pariwisata, penggunaan ulasan sebagai sumber informasi dapat membantu sistem memahami preferensi wisatawan secara lebih mendalam dibandingkan hanya menggunakan informasi umum mengenai destinasi. Abbasi-Moud et al. [9] mengembangkan sistem rekomendasi wisata yang memanfaatkan ulasan pengguna untuk mengekstraksi preferensi, melakukan pengelompokan semantik, dan menganalisis sentimen sebelum menghasilkan rekomendasi destinasi. Penelitian tersebut menunjukkan

bahwa kombinasi antara pemrosesan ulasan dan analisis sentimen dapat digunakan untuk menghasilkan rekomendasi wisata yang lebih sesuai dengan karakteristik dan preferensi pengguna.

Selain membantu wisatawan, sistem rekomendasi berbasis sentimen juga dapat memberikan manfaat bagi pengelola destinasi. Hasil klasifikasi dapat digunakan untuk mengetahui proporsi ulasan positif, negatif, dan netral pada suatu destinasi. Selanjutnya, informasi tersebut dapat digunakan untuk mengidentifikasi aspek yang paling sering memperoleh tanggapan positif maupun keluhan dari pengunjung. Penelitian mengenai analisis ulasan wisata menunjukkan bahwa pengolahan sentimen dan topik dapat membantu praktisi pariwisata memahami kekuatan dan kelemahan destinasi berdasarkan pengalaman pengguna yang tercermin dalam ulasan daring [4], [6]. Dengan demikian, sistem yang menggabungkan analisis sentimen dan rekomendasi dapat memiliki fungsi ganda, yaitu membantu wisatawan memilih destinasi sekaligus menyediakan informasi evaluatif bagi pengelola wisata.

Meskipun penelitian mengenai analisis sentimen ulasan wisata telah berkembang, masih terdapat peluang untuk mengembangkan hasil analisis tersebut menjadi sistem rekomendasi yang secara khusus berfokus pada wisata alam Indonesia. Sebagian penelitian berfokus pada klasifikasi sentimen, identifikasi aspek, atau analisis pengalaman wisatawan, sedangkan penelitian lain mengembangkan sistem rekomendasi berdasarkan preferensi pengguna. Abbasi-Moud et al. [9] menunjukkan bahwa ulasan dan sentimen dapat digunakan untuk menghasilkan rekomendasi wisata yang lebih sesuai dengan preferensi pengguna. Namun, penerapan pendekatan tersebut dalam konteks wisata alam Indonesia dengan menggabungkan klasifikasi sentimen pengunjung dan pemeringkatan beberapa destinasi masih menjadi ruang pengembangan yang relevan.

Berdasarkan permasalahan tersebut, diperlukan suatu sistem yang mampu mengolah data ulasan pengunjung secara otomatis, mengklasifikasikan sentimen yang terkandung di dalamnya, serta menggunakan hasil klasifikasi tersebut sebagai salah satu dasar dalam menentukan rekomendasi destinasi wisata alam. Penerapan NLP dapat digunakan untuk melakukan tahapan pengolahan teks, pembentukan fitur, dan analisis sentimen, sedangkan hasil klasifikasi sentimen dapat digunakan untuk menghasilkan nilai atau skor yang selanjutnya digunakan dalam proses pemeringkatan destinasi. Pendekatan semacam ini sejalan dengan perkembangan penelitian pariwisata yang memanfaatkan user-generated content, NLP, analisis sentimen, dan sistem rekomendasi untuk mendukung pengambilan keputusan wisatawan [9], [7].

Berdasarkan uraian tersebut, penelitian ini diarahkan pada penerapan Natural Language Processing (NLP) dalam sistem rekomendasi wisata alam Indonesia berbasis sentimen pengunjung. Sistem yang dikembangkan diharapkan mampu mengolah ulasan pengunjung menjadi informasi sentimen yang terstruktur, mengidentifikasi kecenderungan penilaian terhadap destinasi, dan menghasilkan urutan rekomendasi berdasarkan hasil analisis tersebut. Dengan demikian, penelitian ini tidak hanya berfokus pada kemampuan model dalam melakukan klasifikasi teks, tetapi juga pada pemanfaatan hasil klasifikasi sebagai informasi yang dapat digunakan dalam proses rekomendasi. Pendekatan tersebut diharapkan dapat membantu calon wisatawan memperoleh rekomendasi yang lebih informatif dan membantu pengelola memahami persepsi pengunjung terhadap destinasi wisata alam Indonesia [7], [9].

Sistem rekomendasi (recommender system) merupakan sistem berbasis komputer yang dirancang untuk membantu pengguna menemukan atau memilih item yang paling relevan dari sejumlah besar alternatif yang tersedia. Sistem ini bekerja dengan melakukan proses penyaringan informasi berdasarkan karakteristik item, preferensi pengguna, riwayat interaksi, maupun informasi lain yang berkaitan dengan pengguna. Keberadaan sistem rekomendasi menjadi semakin penting seiring dengan meningkatnya jumlah informasi digital karena pengguna dapat mengalami kesulitan dalam menentukan pilihan ketika dihadapkan pada terlalu banyak alternatif [10].

Sistem rekomendasi berbasis sentimen merupakan pengembangan sistem rekomendasi yang memanfaatkan informasi mengenai sentimen atau opini yang terdapat dalam data pengguna, terutama

data berbentuk teks seperti ulasan dan komentar. Pendekatan ini muncul karena data rating saja terkadang belum mampu menggambarkan preferensi pengguna secara menyeluruh. Ulasan tekstual dapat mengandung informasi yang lebih spesifik mengenai alasan pengguna menyukai atau tidak menyukai suatu item. Patel dan Chhinkaniwala [11] menjelaskan bahwa analisis sentimen terhadap user-generated content dapat memberikan informasi preferensi pengguna yang lebih spesifik dibandingkan hanya menggunakan data rating.

Natural Language Processing (NLP) merupakan salah satu bidang dalam kecerdasan buatan dan ilmu komputer yang berfokus pada bagaimana komputer dapat memproses, menganalisis, memahami, dan menghasilkan bahasa yang digunakan manusia. NLP menggabungkan konsep dari ilmu komputer, linguistik, dan pembelajaran mesin untuk mengubah bahasa manusia yang bersifat kompleks dan tidak terstruktur menjadi bentuk yang dapat diproses oleh komputer. Dalam perkembangannya, NLP telah digunakan pada berbagai aplikasi, seperti penerjemahan mesin, ekstraksi informasi, klasifikasi teks, analisis sentimen, sistem tanya jawab, dan peringkasan teks [12], [15].

2. METODE

2.1 Pendekatan dan Data Penelitian

Penelitian ini menggunakan pendekatan penelitian terapan untuk merancang sistem rekomendasi wisata alam berbasis analisis sentimen ulasan pengunjung menggunakan Natural Language Processing. Penelitian terapan dipilih karena tujuan penelitian bukan menemukan teori baru, melainkan menerapkan metode yang telah ada untuk menyelesaikan permasalahan nyata, yaitu kesulitan calon wisatawan dalam menyimpulkan ulasan yang berjumlah banyak.

Data yang diolah bersifat kuantitatif dalam bentuk hasil perhitungan sentimen dan skor rekomendasi, meskipun sumber datanya berupa teks. Perubahan dari teks menjadi angka inilah yang menjadi inti pengolahan pada penelitian ini.

Data penelitian merupakan bahan utama yang digunakan dalam proses perancangan sistem. Pada penelitian ini, data yang digunakan berfokus pada ulasan atau komentar pengunjung terhadap sejumlah destinasi wisata unggulan di Indonesia. Dalam penelitian ini, data yang digunakan hanya berupa teks ulasan pengunjung. Data seperti gambar, video, suara, harga tiket, jarak lokasi, dan fasilitas secara rinci tidak dibahas karena penelitian ini dibatasi pada pengolahan teks menggunakan Natural Language Processing. Data penelitian terdiri dari data primer dan data sekunder. Data primer adalah data utama yang digunakan dalam penelitian. Data ini diperoleh dari ulasan atau komentar pengunjung terhadap destinasi wisata melalui platform online seperti Google Review, media sosial, atau sumber digital lainnya. Data ulasan yang telah dikumpulkan akan diseleksi terlebih dahulu. Ulasan yang kosong, duplikat, atau tidak berkaitan dengan objek wisata tidak digunakan dalam penelitian. Setelah itu, data akan diproses melalui tahapan preprocessing agar teks menjadi lebih bersih dan siap dianalisis.

2.2 Tahapan Pengolahan Sentimen

Alur pengolahan sentimen merupakan tahapan yang dilakukan untuk mengubah ulasan pengunjung menjadi informasi yang dapat digunakan dalam sistem rekomendasi wisata alam. Pada penelitian ini, ulasan pengunjung akan diproses menggunakan tahapan Natural Language Processing (NLP) agar teks ulasan dapat dianalisis oleh sistem.

Proses pengolahan sentimen dimulai dari data ulasan yang telah dikumpulkan. Ulasan tersebut kemudian melalui tahap preprocessing, yaitu membersihkan teks agar lebih rapi dan terstruktur. Tahapan preprocessing meliputi case folding, cleaning, tokenizing, stopword removal, dan stemming. Setelah proses preprocessing selesai, data ulasan akan diubah menjadi bentuk numerik menggunakan metode TF-IDF. Metode ini digunakan untuk memberikan bobot pada kata berdasarkan tingkat kepentingannya dalam kumpulan ulasan. Kata yang memiliki nilai penting terhadap isi ulasan akan memiliki bobot lebih besar sehingga dapat membantu proses klasifikasi sentimen.

Tahap berikutnya adalah klasifikasi sentimen. Pada tahap ini, ulasan akan dikelompokkan ke dalam tiga kategori, yaitu sentimen positif, negatif, dan netral. Hasil klasifikasi ini kemudian digunakan sebagai dasar dalam menghitung skor rekomendasi wisata.

Pengumpulan ulasan beberapa destinasi wisata. Tahap pertama adalah mengumpulkan data ulasan pengunjung dari beberapa destinasi wisata di Indonesia melalui platform digital seperti Google Review. Seleksi dan pembersihan data. Data diseleksi untuk menghapus ulasan duplikat, ulasan kosong, spam, atau komentar yang tidak berkaitan dengan objek wisata. Pelabelan sentimen. Setiap ulasan diberi label positif, negatif, atau netral sebagai kelas pada proses pembelajaran mesin menggunakan algoritma Support Vector Machine. Case folding. Seluruh huruf pada teks diubah menjadi huruf kecil agar tidak terjadi perbedaan makna akibat penggunaan huruf kapital. Cleaning dan normalisasi. Cleaning menghapus karakter yang tidak diperlukan seperti angka, URL, emoji, dan tanda baca. Normalisasi mengubah kata tidak baku menjadi kata baku. Tokenizing. Kalimat dipecah menjadi kumpulan kata (token) sehingga setiap kata dapat diproses secara individual. Stopword removal. Menghapus kata-kata yang tidak memiliki pengaruh besar terhadap proses klasifikasi. Stemming. Mengubah kata berimbuhan menjadi kata dasar. Pembobotan TF-IDF. Setiap dokumen diubah menjadi representasi numerik. TF-IDF memberikan bobot tinggi pada kata yang sering muncul pada suatu dokumen tetapi jarang muncul pada keseluruhan dokumen.

Pembagian data training dan testing. Dataset dibagi menjadi data training sebesar 80% dan data testing sebesar 20%. Klasifikasi menggunakan Support Vector Machine. Model SVM dilatih menggunakan data training, kemudian digunakan untuk memprediksi kategori sentimen pada data testing menjadi tiga kelas, yaitu positif, negatif, dan netral. Evaluasi model. Kinerja model diukur menggunakan confusion matrix dan metrik Accuracy, Precision, Recall, serta F1-Score. Pengelompokan hasil per destinasi. Hasil sentimen dikelompokkan berdasarkan masing-masing destinasi wisata dengan menghitung jumlah ulasan positif, negatif, dan netral pada setiap objek wisata. Perhitungan skor rekomendasi. Skor dihitung berdasarkan proporsi sentimen pada setiap destinasi. Destinasi dengan persentase sentimen positif lebih tinggi memperoleh skor lebih besar. Peningkatan wisata alam. Daftar rekomendasi disusun berdasarkan skor yang telah dihitung, dengan destinasi berskor tertinggi ditempatkan pada urutan pertama.

2.3 Rancangan Pengujian Model dan Sistem

Pengujian sistem dilakukan untuk mengetahui apakah sistem rekomendasi wisata alam berbasis sentimen pengunjung dapat berjalan sesuai dengan fungsi yang telah dirancang. Metode pengujian yang digunakan dalam penelitian ini adalah black box testing. Metode ini digunakan karena pengujian berfokus pada fungsi sistem berdasarkan input dan output yang dihasilkan, tanpa melihat kode program secara langsung.

Skema hold-out digunakan sebagai pengujian utama dengan pembagian data latih sebesar 80% dan data uji sebesar 20%. Pembagian dilakukan menggunakan teknik stratified sampling agar proporsi setiap kelas sentimen pada kedua bagian data tetap sebanding dengan proporsi pada dataset asli. Skema validasi silang digunakan sebagai pengujian tambahan. Skema ini diperlukan karena jumlah data pada penelitian ini relatif sedikit, sehingga hasil pengujian hold-out sangat bergantung pada data mana yang kebetulan terpilih sebagai data uji. Dengan validasi silang, seluruh data pernah berperan sebagai data uji sehingga ketergantungan tersebut berkurang.

Metrik evaluasi yang digunakan pada kedua skema meliputi Accuracy, Precision, Recall, dan F1-Score. Nilai Precision, Recall, dan F1-Score dihitung untuk setiap kelas secara terpisah, kemudian diringkas menggunakan perhitungan macro average dan weighted average. Kedua cara peringkasan tersebut dilaporkan bersama-sama karena perbedaan di antara keduanya memberikan petunjuk mengenai ada atau tidaknya permasalahan ketidakseimbangan kelas.

2.4 Perhitungan Skor Rekomendasi

Hasil klasifikasi sentimen tidak langsung ditampilkan kepada pengguna, melainkan diolah terlebih dahulu menjadi skor rekomendasi. Skor ini berfungsi mengubah sebaran sentimen yang berupa hitungan jumlah menjadi satu nilai tunggal yang dapat diurutkan.

Perhitungan dilakukan dengan memberikan bobot berbeda pada setiap kategori sentimen. Sentimen positif diberi bobot penuh karena mencerminkan kepuasan, sentimen netral diberi bobot separuh karena tidak menunjukkan kecenderungan tertentu, sedangkan sentimen negatif diberi bobot nol karena mencerminkan ketidakpuasan. Rumus perhitungan skor dinyatakan sebagai berikut.

$$\text{Skor} = ((1 \times \text{Positif}) + (0,5 \times \text{Netral}) + (0 \times \text{Negatif})) / \text{Total Ulasan} \times 100.$$

Pembagian terhadap total ulasan dilakukan agar skor bersifat proporsional dan tidak dipengaruhi oleh jumlah ulasan. Tanpa pembagian ini, destinasi dengan jumlah ulasan banyak akan selalu memperoleh nilai lebih besar meskipun proporsi keluhannya tinggi. Dengan pembagian tersebut, nilai skor selalu berada pada rentang 0 sampai 100 sehingga dapat dibandingkan antardestinas.

Destinasi kemudian diurutkan berdasarkan skor dari nilai tertinggi ke nilai terendah. Apabila terdapat dua destinasi dengan skor sama, urutan ditentukan berdasarkan jumlah ulasan yang lebih banyak, dengan pertimbangan bahwa skor yang didukung lebih banyak ulasan memiliki tingkat keyakinan yang lebih tinggi.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan, Seleksi, dan Pelabelan Data

Data penelitian dikumpulkan dari ulasan pengunjung pada platform digital Google Review terhadap sejumlah destinasi wisata di Indonesia. Pengumpulan data dilakukan terhadap beberapa destinasi agar sistem dapat melakukan perbandingan dan pemeringkatan antardestinas. Jumlah data ulasan yang berhasil dikumpulkan adalah sebanyak 70 ulasan yang tersebar pada 11 destinasi wisata. Setiap data ulasan memiliki sembilan atribut, yaitu `id_ulasan`, `nama_wisata`, `kategori`, `lokasi`, `provinsi`, `deskripsi`, `rating`, `ulasan`, dan `tanggal`. Atribut yang digunakan pada proses analisis sentimen adalah atribut ulasan sebagai data teks dan atribut rating sebagai dasar pelabelan sentimen.

Data ulasan yang telah dikumpulkan diseleksi terlebih dahulu untuk menghilangkan data yang tidak layak digunakan. Kriteria seleksi meliputi penghapusan ulasan yang bernilai kosong, penghapusan ulasan duplikat pada destinasi yang sama, serta penghapusan ulasan yang tidak berkaitan dengan objek wisata. Tidak terdapat data yang dibuang pada tahap seleksi. Hal ini disebabkan seluruh data ulasan telah melalui penyaringan awal pada saat proses pengumpulan data, sehingga jumlah data yang digunakan pada tahap selanjutnya tetap berjumlah 70 ulasan.

Pelabelan sentimen dilakukan. Label sentimen diperlukan sebagai kelas target pada proses pembelajaran mesin menggunakan algoritma Support Vector Machine. Pada penelitian ini, pelabelan dilakukan berdasarkan nilai rating yang diberikan pengunjung dengan ketentuan rating bernilai 4 atau 5 dilabeli Positif, rating bernilai 3 dilabeli Netral, dan rating bernilai 1 atau 2 dilabeli Negatif. Hasil pelabelan ditampilkan pada Tabel 1.

Berdasarkan Tabel 1 terlihat bahwa distribusi kelas sentimen pada dataset sangat tidak seimbang. Kelas Positif mendominasi dengan proporsi 82,86%, sedangkan kelas Netral hanya berjumlah 4 ulasan atau 5,71% dari keseluruhan data. Kondisi ketidakseimbangan kelas ini menjadi catatan penting yang akan dibahas lebih lanjut pada bagian evaluasi model dan keterbatasan penelitian.

Tabel 1. Hasil Pelabelan Sentimen

No	Kelas Sentimen	Rentang Rating	Jumlah	Persentase
1	Positif	Rating 4 - 5	58	82,86%
2	Negatif	Rating 1 - 2	8	11,43%
3	Netral	Rating 3	4	5,71%
	Total		70	100,00%

3.2 Preprocessing dan Pembobotan TF-IDF

Preprocessing teks dilakukan melalui tahapan yang meliputi case folding, cleaning dan normalisasi, tokenizing, stopwords removal, serta stemming. Proses ini bertujuan mengurangi gangguan (noise) pada data teks sehingga algoritma dapat bekerja lebih akurat. Contoh hasil pengolahan pada setiap tahapan ditampilkan pada Tabel 2.

Tabel 2. Contoh Hasil Tiap Tahapan Preprocessing

No	Tahapan	Hasil
0	Teks Asli	Pemandangan Danau Toba sangat indah dan udaranya sejuk. Cocok untuk liburan bersama keluarga.
1	Case Folding	pemandangan danau toba sangat indah dan udaranya sejuk. cocok untuk liburan bersama keluarga.
2	Cleaning	pemandangan danau toba sangat indah dan udaranya sejuk cocok untuk liburan bersama keluarga
3	Tokenizing	['pemandangan', 'danau', 'toba', 'sangat', 'indah', 'dan', 'udaranya', 'sejuk', 'cocok', 'untuk', 'liburan', 'bersama', 'keluarga']
4	Normalisasi	['pemandangan', 'danau', 'toba', 'sangat', 'indah', 'dan', 'udaranya', 'sejuk', 'cocok', 'untuk', 'liburan', 'bersama', 'keluarga']
5	Stopword Removal	['pemandangan', 'danau', 'toba', 'sangat', 'indah', 'udaranya', 'sejuk', 'cocok', 'liburan', 'bersama', 'keluarga']
6	Stemming	pandang danau toba sangat indah udara sejuk cocok libur sama keluarga

Berdasarkan Tabel 2 dapat dilihat bahwa tahap cleaning berhasil menghilangkan tanda baca berupa titik pada akhir kalimat. Tahap stopwords removal menghilangkan kata sambung "dan" dan "untuk" yang tidak memiliki pengaruh terhadap penentuan sentimen. Selanjutnya tahap stemming mengubah kata berimbuhan menjadi kata dasar, seperti kata "pemandangan" menjadi "pandang", kata "udaranya" menjadi "udara", dan kata "liburan" menjadi "libur".

Perlu dijelaskan bahwa pada tahap stopwords removal dilakukan penyesuaian terhadap daftar stopwords bawaan pustaka Sastrawi. Kata negasi dan kata penghubung kontras seperti "tidak", "bukan", "kurang", "tetapi", dan "namun" dikecualikan dari proses penghapusan. Penyesuaian ini diperlukan karena penghapusan kata negasi dapat mengubah makna sentimen sebuah kalimat. Sebagai contoh, kalimat "fasilitasnya tidak bagus" akan berubah menjadi "fasilitas bagus" apabila kata "tidak" dihapus, sehingga sentimen negatif akan salah dikenali sebagai sentimen positif.

Setelah preprocessing selesai, data teks diubah menjadi representasi numerik menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). Hasil pembobotan menghasilkan matriks berukuran 70 baris dokumen dikali 224 kolom fitur kata.

Hasil pembobotan menunjukkan bahwa kata "sangat" memperoleh bobot tertinggi sebesar 0,1004. Hal ini menunjukkan bahwa kata penguat intensitas banyak digunakan pengunjung dalam menyampaikan penilaian. Kata "pandang" yang merupakan bentuk dasar dari kata "pemandangan" menempati posisi kedua, yang menandakan bahwa aspek pemandangan merupakan hal yang paling

sering diulas pengunjung. Kata "indah" dan "cocok" yang bermuatan positif juga muncul pada peringkat atas, sejalan dengan dominasi kelas Positif pada dataset.

3.3 Klasifikasi dan Evaluasi Model

Dataset dibagi menjadi data training sebesar 80% dan data testing sebesar 20%. Pembagian dilakukan menggunakan teknik stratified sampling agar proporsi setiap kelas sentimen pada data training dan data testing tetap sebanding dengan proporsi pada dataset asli. Data training berjumlah 56 ulasan dan data testing berjumlah 14 ulasan. Perlu diperhatikan bahwa jumlah data testing pada kelas Netral hanya sebanyak 1 ulasan dan kelas Negatif hanya sebanyak 2 ulasan. Jumlah yang sangat sedikit ini menyebabkan hasil evaluasi pada kedua kelas tersebut menjadi kurang stabil dan sangat sensitif terhadap kesalahan prediksi.

Model klasifikasi dibangun menggunakan algoritma Support Vector Machine dengan kernel linear. Kernel linear dipilih karena data hasil pembobotan TF-IDF memiliki dimensi fitur yang tinggi namun jumlah sampel relatif sedikit, sehingga kernel linear cenderung memberikan hasil yang lebih stabil dan tidak mudah mengalami overfitting.

Parameter `class_weight` diatur pada nilai `balanced` sebagai upaya menangani ketidakseimbangan kelas. Dengan pengaturan tersebut, algoritma memberikan bobot penalti yang lebih besar terhadap kesalahan klasifikasi pada kelas minoritas, yaitu kelas Netral dan kelas Negatif. Evaluasi model dilakukan menggunakan confusion matrix serta metrik Accuracy, Precision, Recall, dan F1-Score.

Berdasarkan confusion matrix, dari 14 data testing terdapat 12 data yang diprediksi dengan benar dan 2 data yang salah diprediksi. Seluruh data kelas Positif sebanyak 11 ulasan berhasil diprediksi dengan benar. Pada kelas Negatif, dari 2 ulasan hanya 1 ulasan yang berhasil dikenali, sedangkan 1 ulasan lainnya salah diprediksi sebagai Positif. Pada kelas Netral, 1 ulasan yang tersedia juga salah diprediksi sebagai Positif. Berdasarkan confusion matrix tersebut, dihitung nilai metrik evaluasi untuk setiap kelas sebagaimana ditampilkan pada Tabel 3.

Tabel 3. Hasil Evaluasi Model per Kelas Sentimen

Kelas Sentimen	Precision	Recall	F1-Score	Support
Positif	0,85	1,00	0,92	11
Negatif	1,00	0,50	0,67	2
Netral	0,00	0,00	0,00	1
Macro Average	0,62	0,50	0,53	14
Weighted Average	0,81	0,86	0,82	14

Nilai Accuracy yang diperoleh model adalah sebesar 85,71%. Nilai Precision, Recall, dan F1-Score dengan perhitungan weighted average berturut-turut adalah 80,77%, 85,71%, dan 81,55%. Namun demikian, apabila dilihat menggunakan perhitungan macro average, nilai Precision, Recall, dan F1-Score menurun tajam menjadi 61,54%, 50,00%, dan 52,78%.

Perbedaan yang besar antara nilai weighted average dan macro average tersebut menunjukkan bahwa nilai Accuracy sebesar 85,71% tidak dapat dijadikan satu-satunya acuan keberhasilan model. Nilai Accuracy yang tinggi tersebut lebih banyak disumbang oleh keberhasilan model mengenali kelas Positif yang jumlahnya mendominasi dataset. Model bahkan gagal sepenuhnya mengenali kelas Netral dengan nilai F1-Score sebesar 0,00.

Untuk memperoleh gambaran kinerja yang lebih stabil, dilakukan pengujian tambahan menggunakan validasi silang (cross validation) dengan skema stratified 4-fold. Hasil pengujian tersebut ditampilkan pada Tabel 4.

Tabel 4. Hasil Pengujian Validasi Silang 4-Fold

Fold	Jumlah Data Uji	Akurasi	Keterangan
Fold 1	18	88,89%	Baik
Fold 2	18	94,44%	Sangat baik
Fold 3	17	88,24%	Baik
Fold 4	17	94,12%	Sangat baik
Rata-rata	70	91,42%	Standar deviasi 2,87%

Berdasarkan Tabel 4, rata-rata akurasi hasil validasi silang adalah sebesar 91,42% dengan standar deviasi sebesar 2,87%. Nilai standar deviasi yang kecil menunjukkan bahwa kinerja model relatif konsisten pada seluruh pembagian data. Meskipun demikian, nilai akurasi ini tetap harus dibaca bersama dengan nilai macro F1-Score, mengingat kondisi dataset yang sangat tidak seimbang.

3.4 Pemeringkatan Rekomendasi Wisata

Setelah seluruh ulasan berhasil diklasifikasikan, hasil sentimen dikelompokkan berdasarkan masing-masing destinasi wisata. Pada tahap ini dihitung jumlah ulasan positif, negatif, dan netral pada setiap objek wisata.

Perhitungan skor rekomendasi dan pemeringkatan wisata alam. Skor rekomendasi dihitung berdasarkan proporsi sentimen pada setiap destinasi dengan memberikan bobot berbeda pada setiap kategori sentimen. Rumus perhitungan skor rekomendasi yang digunakan adalah sebagai berikut.

$$\text{Skor} = ((1 \times \text{Positif}) + (0,5 \times \text{Netral}) + (0 \times \text{Negatif})) / \text{Total Ulasan} \times 100$$

Bobot 1 diberikan pada sentimen positif karena mencerminkan kepuasan pengunjung, bobot 0,5 diberikan pada sentimen netral karena bersifat umum dan tidak menunjukkan kecenderungan tertentu, sedangkan bobot 0 diberikan pada sentimen negatif karena mencerminkan ketidakpuasan pengunjung. Nilai skor berada pada rentang 0 sampai 100.

Sebagai contoh perhitungan, destinasi Pantai Kuta memiliki 7 ulasan positif, 1 ulasan netral, dan 2 ulasan negatif dari total 10 ulasan, sehingga skornya adalah $((1 \times 7) + (0,5 \times 1) + (0 \times 2)) / 10 \times 100 = 75,00$. Hasil perhitungan skor dan pemeringkatan seluruh destinasi ditampilkan pada Tabel 5.

Tabel 5. Hasil Pemeringkatan Rekomendasi Wisata Alam

Ranking	Nama Destinasi	Total Ulasan	% Positif	Skor
1	Lompat Batu Bawomataluo	5	100,00%	100,00
2	Banda Neira	6	83,33%	91,67
3	Bukit Lawang	6	83,33%	91,67
4	Kawah Ijen	6	83,33%	91,67
5	Danau Toba	7	85,71%	85,71
6	Candi Borobudur	6	83,33%	83,33
7	Gunung Bromo	6	83,33%	83,33
8	Jam Gadang	6	83,33%	83,33
9	Ngarai Sianok	6	83,33%	83,33
10	Raja Ampat	6	83,33%	83,33
11	Pantai Kuta	10	70,00%	75,00

Berdasarkan Tabel 5, destinasi Lompat Batu Bawomataluo menempati peringkat pertama dengan skor sempurna 100,00 karena seluruh ulasannya bersentimen positif. Sebaliknya, Pantai Kuta

menempati peringkat terakhir dengan skor 75,00 meskipun memiliki jumlah ulasan positif terbanyak yaitu 7 ulasan. Hal ini terjadi karena Pantai Kuta juga memiliki jumlah ulasan negatif terbanyak, yaitu 2 ulasan yang umumnya mengeluhkan kebersihan area wisata dan kepadatan pengunjung.

Temuan ini menunjukkan bahwa sistem rekomendasi berbasis sentimen memberikan hasil yang berbeda dengan sistem rekomendasi berbasis popularitas. Destinasi dengan jumlah ulasan terbanyak tidak secara otomatis menempati peringkat teratas, karena yang dinilai adalah proporsi kepuasan pengunjung, bukan jumlah kunjungan atau jumlah ulasan.

Perlu dicatat bahwa destinasi dengan jumlah ulasan sedikit berpotensi memperoleh skor tinggi secara tidak proporsional. Lompat Batu Bawomataluo memperoleh skor sempurna hanya berdasarkan 5 ulasan, sehingga tingkat keyakinan terhadap skor tersebut lebih rendah dibandingkan destinasi dengan jumlah ulasan lebih banyak.

3.5 Pengujian Sistem, Pembahasan, dan Keterbatasan

Pengujian sistem dilakukan menggunakan metode black box testing. Pengujian berfokus pada fungsionalitas sistem berdasarkan input dan output yang dihasilkan tanpa melihat struktur kode program. Seluruh fitur yang diuji sebanyak 10 skenario pengujian memberikan hasil sesuai dengan yang diharapkan. Dengan demikian dapat dinyatakan bahwa sistem telah berjalan sesuai dengan fungsi yang dirancang. Berdasarkan hasil penelitian yang telah diuraikan, penerapan Natural Language Processing dalam sistem rekomendasi wisata alam berbasis sentimen pengunjung dapat dilakukan melalui rangkaian tahapan yang saling berkesinambungan. Tahapan preprocessing berhasil mengubah ulasan pengunjung yang berbentuk teks tidak terstruktur menjadi teks yang lebih bersih dan seragam. Pembobotan TF-IDF kemudian berhasil mengubah teks tersebut menjadi representasi numerik berdimensi 224 fitur yang dapat diproses oleh algoritma pembelajaran mesin.

Pada sisi rekomendasi, penerapan pembobotan sentimen berhasil menghasilkan urutan destinasi yang berbeda dari urutan berbasis popularitas. Hal ini membuktikan bahwa hasil analisis sentimen benar-benar memberikan pengaruh terhadap keluaran sistem rekomendasi, sesuai dengan tujuan penelitian. Perbedaan hasil antara skema hold-out dan validasi silang juga perlu dicermati. Nilai akurasi pada validasi silang sebesar 91,42% lebih tinggi dibandingkan hold-out sebesar 85,71%. Perbedaan ini bukan berarti model bekerja lebih baik, melainkan menunjukkan bahwa hasil pada skema hold-out sangat dipengaruhi oleh data mana yang kebetulan terpilih sebagai data uji. Pada penelitian dengan jumlah data terbatas, satu kesalahan prediksi saja sudah menurunkan akurasi sekitar tujuh persen, sehingga hasil pengujian tunggal menjadi kurang dapat diandalkan.

Simpangan baku sebesar 2,87% pada validasi silang menunjukkan bahwa kinerja model relatif konsisten pada berbagai pembagian data. Konsistensi ini merupakan pertanda baik, namun tetap harus dibaca bersama nilai macro F1-Score yang rendah. Model konsisten dalam mengenali kelas positif, dan sekaligus konsisten pula dalam gagal mengenali kelas netral. Penelitian ini memiliki sejumlah keterbatasan yang perlu disampaikan secara terbuka sebagai dasar pengembangan penelitian selanjutnya. Jumlah data ulasan yang digunakan relatif sedikit, yaitu 70 ulasan pada 11 destinasi. Jumlah tersebut belum cukup untuk membangun model klasifikasi yang benar-benar andal, terutama untuk kelas minoritas.

Distribusi kelas sangat tidak seimbang dengan dominasi kelas Positif sebesar 82,86%. Akibatnya nilai Accuracy menjadi kurang representatif dan model gagal mengenali kelas Netral. Jumlah data testing pada skema pembagian 80:20 hanya sebanyak 14 ulasan, dengan kelas Netral hanya terwakili 1 ulasan. Hasil evaluasi pada kelas tersebut menjadi sangat sensitif terhadap satu kesalahan prediksi. Pelabelan sentimen dilakukan secara otomatis berdasarkan nilai rating, bukan berdasarkan pembacaan isi teks oleh pelabel manusia. Pendekatan ini berpotensi menimbulkan ketidaksesuaian apabila terdapat ulasan yang isi teksnya tidak sejalan dengan rating yang diberikan.

Objek penelitian mencakup destinasi berkategori sejarah dan budaya seperti Candi Borobudur, Jam Gadang, Banda Neira, dan Lompat Batu Bawomataluo, sehingga belum sepenuhnya terbatas pada kategori wisata alam.

Sistem belum memiliki mekanisme pembobotan berdasarkan jumlah ulasan, sehingga destinasi dengan jumlah ulasan sedikit dapat memperoleh skor tinggi secara tidak proporsional. Terdapat dua perbedaan mendasar antara penelitian ini dengan penelitian terdahulu. Perbedaan pertama terletak pada jumlah objek yang diteliti. Penelitian terdahulu umumnya meneliti satu destinasi sehingga keluaran yang dihasilkan hanya berupa sebaran sentimen pada destinasi tersebut. Penelitian ini meneliti sebelas destinasi sekaligus sehingga memungkinkan dilakukannya perbandingan antardestinasi.

Perbedaan kedua terletak pada keluaran akhir. Penelitian terdahulu berhenti pada tahap pelaporan hasil klasifikasi, sedangkan penelitian ini melanjutkannya menjadi skor rekomendasi dan pemeringkatan destinasi. Dengan demikian, hasil analisis sentimen tidak hanya menjadi informasi deskriptif, tetapi juga menjadi dasar pengambilan keputusan bagi calon wisatawan. Meskipun demikian, perlu diakui bahwa jumlah data pada penelitian ini jauh lebih sedikit dibandingkan penelitian terdahulu yang umumnya menggunakan ratusan hingga ribuan ulasan. Keterbatasan ini menyebabkan nilai evaluasi yang diperoleh belum dapat dibandingkan secara setara, terutama pada kelas minoritas.

4. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan mengenai penerapan Natural Language Processing dalam sistem rekomendasi wisata alam Indonesia berbasis sentimen pengunjung, dapat ditarik kesimpulan sebagai berikut. Metode Natural Language Processing dapat diterapkan untuk menganalisis sentimen ulasan pengunjung wisata alam melalui rangkaian tahapan preprocessing yang meliputi case folding, cleaning dan normalisasi, tokenizing, stopword removal, serta stemming menggunakan pustaka Sastrawi. Penyesuaian daftar stopword dengan mempertahankan kata negasi terbukti diperlukan agar makna sentimen kalimat tidak berubah pada saat proses penghapusan kata. Pengolahan data teks agar dapat diklasifikasikan menjadi sentimen positif, negatif, atau netral dilakukan dengan mengubah teks hasil preprocessing menjadi representasi numerik menggunakan pembobotan Term Frequency-Inverse Document Frequency, yang menghasilkan matriks berukuran 70 dokumen dikali 224 fitur kata. Representasi numerik tersebut selanjutnya diklasifikasikan menggunakan algoritma Support Vector Machine dengan kernel linear dan pembagian data 80% training serta 20% testing. Sistem rekomendasi wisata alam berhasil dirancang dan dibangun dengan memanfaatkan hasil analisis sentimen. Skor rekomendasi dihitung menggunakan pembobotan proporsi sentimen setiap destinasi, di mana sentimen positif diberi bobot 1, sentimen netral diberi bobot 0,5, dan sentimen negatif diberi bobot 0. Hasil pemeringkatan menempatkan Lompat Batu Bawomataluo pada peringkat pertama dengan skor 100,00 dan Pantai Kuta pada peringkat terakhir dengan skor 75,00. Tingkat akurasi model dalam mengklasifikasikan sentimen adalah sebesar 85,71% pada pengujian hold-out 80:20 dan sebesar 91,42% dengan standar deviasi 2,87% pada pengujian validasi silang 4-fold. Namun demikian, nilai macro F1-Score hanya sebesar 52,78% karena model gagal mengenali kelas Netral yang jumlah datanya sangat sedikit. Hal ini menunjukkan bahwa nilai akurasi yang tinggi belum menjamin kinerja model yang merata pada seluruh kelas sentimen. Sistem yang dibangun dapat membantu calon wisatawan dalam memilih destinasi wisata berdasarkan pengalaman pengunjung sebelumnya. Urutan rekomendasi yang dihasilkan berbeda dengan urutan berbasis popularitas, karena destinasi dengan jumlah ulasan terbanyak tidak secara otomatis memperoleh peringkat tertinggi apabila proporsi ulasan negatifnya juga tinggi. Hasil pengujian sistem menggunakan metode black box testing terhadap 10 skenario pengujian menunjukkan bahwa seluruh fitur sistem berjalan sesuai dengan fungsi yang dirancang. Meskipun demikian, pencapaian tersebut perlu dimaknai secara berimbang. Keberhasilan sistem dalam menghasilkan peringkat rekomendasi tidak serta-merta berarti bahwa peringkat tersebut sepenuhnya dapat diandalkan, mengingat sebagian destinasi hanya didukung oleh lima sampai tujuh ulasan. Nilai

akurasi yang tinggi juga tidak dapat dilepaskan dari dominasi kelas positif pada dataset. Oleh karena itu, hasil penelitian ini sebaiknya diposisikan sebagai bukti bahwa pendekatan yang digunakan dapat berjalan, bukan sebagai bukti bahwa sistem telah siap digunakan secara luas.

REFERENSI

- [1] Badan Pusat Statistik, Statistik Objek Daya Tarik Wisata 2024. Jakarta, Indonesia: Badan Pusat Statistik, 2025.
- [2] Badan Pusat Statistik, Statistik Wisatawan Nusantara 2024. Jakarta, Indonesia: Badan Pusat Statistik, 2025.
- [3] Badan Pusat Statistik, Statistik Kunjungan Wisatawan Mancanegara 2024. Jakarta, Indonesia: Badan Pusat Statistik, 2025.
- [4] S. Alam et al., "Use and impact of online travel reviews for planning free and easy holidays," *International Journal of Information Systems in the Service Sector*, vol. 14, no. 1, 2022, doi: 10.4018/IJISSS.287583.
- [5] E. Bigne, C. Ruiz, C. Perez-Cabañero, and A. Cuenca, "Are customer star ratings and sentiments aligned? A deep learning study of the customer service experience in tourism destinations," *Service Business*, vol. 17, pp. 281–314, 2023, doi: 10.1007/s11628-023-00524-0.
- [6] N. Saraswathi and T. Sasirooba, "Improving the accuracy of sentiment analysis using a linguistic rule-based feature selection method in tourism reviews," *Measurement: Sensors*, vol. 29, Art. no. 100888, 2023, doi: 10.1016/j.measen.2023.100888.
- [7] D. Guidotti, L. Pandolfo, and L. Pulina, "Discovering sentiment insights: Streamlining tourism review analysis with large language models," *Information Technology & Tourism*, vol. 27, pp. 227–261, 2025, doi: 10.1007/s40558-024-00309-9.
- [8] R. A. Nugroho, B. Warsito, and A. R. Hakim, "Performance comparison of Support Vector Machine and Random Forest using Delta TF-IDF on sentiment classification of Tiket.com reviews," in *2023 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, 2023, pp. 632–637, doi: 10.1109/ISITIA59021.2023.10221142.
- [9] Z. Abbasi-Moud, H. Vahdat-Nejad, and J. Sadri, "Tourism recommendation system based on semantic clustering and sentiment analysis," *Expert Systems with Applications*, vol. 167, Art. no. 114324, 2021, doi: 10.1016/j.eswa.2020.114324.
- [10] D. Roy and M. Dutta, "A systematic review and research perspective on recommender systems," *Journal of Big Data*, vol. 9, Art. no. 59, 2022, doi: 10.1186/s40537-022-00592-5.
- [11] J. Patel and H. Chhinkaniwala, "The role of sentiment analysis in a recommender system: A systematic survey," *International Journal of Web Engineering and Technology*, vol. 17, no. 1, pp. 29–62, 2022, doi: 10.1504/IJWET.2022.125086.
- [12] D. Khurana, A. Koli, K. Khatter, and S. Singh, "Natural language processing: State of the art, current trends and challenges," *Multimedia Tools and Applications*, vol. 82, pp. 3713–3744, 2023, doi: 10.1007/s11042-022-13428-4.
- [13] M. Birjali, M. Kasri, and A. Beni-Hssane, "A comprehensive survey on sentiment analysis: Approaches, challenges and trends," *Knowledge-Based Systems*, vol. 226, Art. no. 107134, 2021, doi: 10.1016/j.knosys.2021.107134.
- [14] A. Ligthart, C. Catal, and B. Tekinerdogan, "Systematic reviews in sentiment analysis: A tertiary study," *Artificial Intelligence Review*, vol. 54, pp. 4997–5053, 2021, doi: 10.1007/s10462-021-09973-3.
- [15] T.-X. Sun, X.-Y. Liu, X.-P. Qiu, and X.-J. Huang, "Paradigm shift in natural language processing," *Machine Intelligence Research*, vol. 19, pp. 169–183, 2022, doi: 10.1007/s11633-022-1331-6.