

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pariwisata merupakan salah satu sektor yang memiliki kontribusi signifikan terhadap perekonomian Indonesia [1]. Taman Safari Indonesia, sebagai salah satu atraksi wisata terkemuka di negeri ini, menarik banyak pengunjung dari berbagai kalangan. Setiap tahun, taman ini berfungsi sebagai destinasi konservasi sekaligus menampilkan keberagaman satwa dari seluruh dunia dalam lingkungan yang menyerupai habitat alami mereka [2]. Kesuksesan Taman Safari Indonesia dalam menarik pengunjung menjadikan pentingnya pemahaman mendalam tentang persepsi masyarakat terhadap layanan dan fasilitas yang disediakan [3].

Dalam upaya memahami pengalaman pengunjung, ulasan yang ditinggalkan di *platform* digital seperti *Google Maps* menjadi sumber data yang berharga [4]. Ulasan-ulasan ini tidak hanya mengandung penilaian numerik, tetapi juga menyimpan informasi penting tentang kepuasan, harapan, dan pengalaman unik dari setiap pengunjung. Mengingat meningkatnya penggunaan *platform* digital, terdapat kebutuhan untuk menganalisis data ulasan ini secara sistematis. Analisis sentimen, yang merupakan bagian dari pengolahan bahasa alami NLP (*Natural Language Processing*), menjadi alat utama dalam mengevaluasi dan mengategorikan emosi yang terkandung dalam ulasan [5].

Melalui analisis sentimen, setiap ulasan dapat diklasifikasikan menjadi beberapa kategori positif, negatif, atau netral [6]. Dengan memanfaatkan algoritma *machine learning* seperti *Logistic Regression*, *Naïve Bayes*, *Support Vector Machine (SVM)*, *Random Forest*, dan *K-Nearest Neighbors*. penelitian bertujuan untuk mengidentifikasi pola perilaku dari pengunjung, serta meningkatkan pemahaman manajemen tentang faktor-faktor yang memengaruhi kepuasan pengunjung. Hasil analisis ini akan memberikan

wawasan yang berguna bagi pengelola Taman Safari untuk merespon umpan balik dan mengoptimalkan kualitas layanan [7].

Penelitian ini memilih lima algoritma yang digunakan untuk mendukung penelian, yaitu *Logistic Regression*, *Naïve Bayes*, *Support Vector Machine (SVM)*, *Random Forest*, dan *K-Nearest Neighbors*, karena masing-masing memiliki keunggulan dalam analisis sentimen. *Logistic Regression* dipilih karena kesederhanaan dan efisiensinya dalam menghasilkan model yang terinterpretasikan dengan baik, *Naïve Bayes* dikenal cepat dan efektif dalam menangani data yang tidak seimbang. *Support Vector Machine* unggul dalam menangani data dengan dimensi tinggi, yang sering ditemukan dalam analisis teks. *Random Forest*, dapat memberikan akurasi yang tinggi dengan mengurangi risiko *overfitting* melalui pengambilan keputusan berbasis banyak pohon keputusan. Terakhir, *K-Nearest Neighbors* dipilih karena kemudahan implementasinya dan kemampuan untuk beradaptasi dengan pola data yang kompleks. Dengan memanfaatkan kombinasi dari algoritma-algoritma ini, penelitian diperkirakan dapat memberikan pemahaman yang lebih mendalam mengenai perilaku pengunjung serta meningkatkan kualitas analisis sentimen yang dilakukan.

Pemilihan algoritma ini didasarkan pada dukungan dari berbagai jurnal yang relevan, di mana penelitian yang menunjukkan efektivitas SVM mendukung penggunaannya dalam menganalisis sentimen pengunjung Taman Safari Indonesia. Selain itu, penelitian yang menguatkan pilihan *Naïve Bayes* dan *Logistic Regression* yang juga diterapkan dalam analisis ini [8]. Penelitian lain yang juga menunjukkan akurasi tinggi untuk SVM yaitu 92% dan *Naïve Bayes* sebesar 90%, memberikan validasi tambahan terkait algoritma tersebut [9]. Secara keseluruhan, hubungan antar penelitian tersebut memperkuat argumen untuk pemilihan algoritma yang dilakukan di sini dan memberikan konteks praktis untuk meningkatkan kualitas analisis dan rekomendasi yang dihasilkan.

Selain itu, tantangan yang dihadapi oleh pengunjung adalah kesulitan dalam menentukan pilihan destinasi yang sesuai di tengah banyaknya informasi dan

pilihan yang tersedia. Dengan demikian, rekomendasi berbasis ulasan menjadi semakin relevan [10]. Melalui penerapan *Term Frequency*, pengunjung dapat diarahkan pada atraksi atau pengalaman yang sesuai dengan preferensi mereka berdasarkan ulasan yang ada. Sistem ini akan menganalisis konten ulasan untuk menawarkan rekomendasi yang relevan, sehingga membantu pengunjung menemukan atraksi yang lebih memuaskan [11].

Rekomendasi berbasis konten ini tidak hanya menghadirkan keuntungan bagi pengunjung tetapi juga memberikan perspektif baru bagi pengelola Taman Safari Indonesia. Dengan memanfaatkan hasil analisis sentimen, pengelola dapat mengidentifikasi atraksi mana yang paling banyak disukai atau dikritik, serta memperbaiki aspek-aspek layanan yang perlu ditingkatkan [12]. Melalui rekomendasi yang lebih dipersonalisasi, Taman Safari Indonesia dapat meningkatkan pengalaman berkunjung dan menciptakan loyalitas pengunjung yang lebih tinggi.

Penelitian bertujuan untuk menganalisis sentimen pengunjung dan memberikan rekomendasi yang relevan untuk meningkatkan pengalaman di Taman Safari Indonesia. Dengan pemahaman yang lebih baik mengenai persepsi pengunjung, diperkirakan manajemen dapat mengambil keputusan strategis yang berbasis pada data untuk inovasi layanan [13]. Penelitian difokuskan pada pengolahan data ulasan yang diambil dari *Google Maps*, yang akan dianalisis untuk mengidentifikasi tema utama dan pola sentimen yang muncul.

Penggunaan *Google Maps* sebagai sumber data utama dalam penelitian didasarkan pada beberapa alasan yang relevan. *Google Maps* adalah *platform* yang sangat populer dan banyak digunakan oleh wisatawan untuk membagikan pengalaman mereka, sehingga menyediakan volume ulasan yang besar dan variatif. Ulasan yang terdapat di *Google Maps* sering kali mencerminkan pengalaman langsung pengunjung, yang membuat data tersebut lebih autentik dan berharga untuk analisis sentimen. Menganalisis ulasan dari satu *platform* dapat memfokuskan metode analisis secara lebih mendalam dan konsisten,

mengurangi variabilitas yang mungkin muncul jika menggunakan banyak sumber data. Alasan-alasan tersebut membuat *Google Maps* dipilih sebagai sumber data yang optimal untuk mendalami analisis sentimen dan rekomendasi dalam penelitian.

Penelitian juga bertujuan untuk menghadapi tantangan dalam pengambilan keputusan saat merencanakan kunjungan. Banyak pengunjung yang mungkin merasa terjebak dalam lautan informasi yang membingungkan mengenai berbagai pilihan atraksi yang ada [14]. Oleh karena itu, melalui rekomendasi yang efisien, penelitian diperkirakan dapat memudahkan pengunjung dalam menemukan pengalaman yang sesuai dengan minat mereka, membuat mereka merasa lebih puas dan terlayani dengan baik.

Perkembangan riset di bidang analisis sentimen dan rekomendasi berbasis konten, hasil penelitian diperkirakan memberikan kontribusi signifikan terhadap pengembangan teknologi informasi dalam pariwisata [15]. Dengan menganalisis sentimen yang terkandung dalam ulasan dan mengimplementasikan rekomendasi, penelitian bertujuan untuk meningkatkan kualitas layanan dan interaksi antara pengelola taman safari dan pengunjung [16]. Diperkirakan bahwa hasil penelitian dapat menjadi acuan bagi penelitian lebih lanjut di bidang yang sama, sehingga mendorong adopsi teknologi baru dalam pengelolaan pariwisata.

Secara keseluruhan, penelitian akan memberikan wawasan yang lebih dalam tentang pengalaman pengunjung Taman Safari Indonesia dan menyoroti pentingnya analisis sentimen serta rekomendasi dalam meningkatkan kualitas layanan [17]. Dengan mengandalkan data ulasan sebagai dasar analisis, pengelola Taman Safari dapat berinovasi dalam memenuhi harapan pengunjung serta menciptakan pengalaman yang lebih memuaskan, yang pada akhirnya dapat berdampak positif terhadap pertumbuhan sektor pariwisata di Indonesia secara keseluruhan.

Urgensi penelitian terletak pada kemampuannya untuk membantu wisatawan dengan memberikan rekomendasi tempat atau hewan yang sering

dikunjungi. Dengan menyediakan informasi yang akurat dan relevan, penelitian bertujuan untuk memudahkan pengunjung dalam menentukan pilihan yang paling sesuai dengan minat dan harapan mereka. Rekomendasi yang efisien tidak hanya akan meningkatkan kepuasan pengunjung, tetapi juga mendorong mereka untuk mengeksplorasi lebih banyak atraksi yang ditawarkan oleh Taman Safari Indonesia. Penelitian diperkirakan dapat berkontribusi pada pengalaman berkunjung yang lebih memuaskan dan optimal, serta meningkatkan daya tarik Taman Safari sebagai destinasi wisata.

1.2 Rumusan Masalah

1. Bagaimana sentimen pengunjung Taman Safari Indonesia yang diungkapkan melalui ulasan di *Google Maps*?
2. Apa saja tema utama yang muncul dari analisis sentimen terhadap ulasan pengunjung?
3. Bagaimana kinerja algoritma *machine learning* dalam mengklasifikasikan sentimen ulasan pengunjung?
4. Bagaimana cara menerapkan *Term Frequency* untuk merekomendasikan atraksi di Taman Safari Indonesia berdasarkan sentimen dan preferensi pengunjung?

1.3 Batasan Masalah

1. Penelitian hanya akan menggunakan ulasan pengunjung Taman Safari Indonesia yang tersedia di *platform Google Maps* dengan rentang waktu 10 tahun terakhir.
2. Ulasan pengunjung hanya akan diklasifikasikan menjadi tiga kategori sentimen: positif, negatif, dan netral.
3. Penelitian akan membatasi analisis kinerja pada lima algoritma *machine learning* tertentu, yaitu *Logistic Regression*, *Naive Bayes*, dan *Support Vector Machine (SVM)*, *K-Nearest Neighbor (KNN)* dan *Random Forest*
4. Rekomendasi atraksi yang dihasilkan hanya akan berdasarkan analisis sentimen dari ulasan yang sudah diklasifikasikan dan tidak akan mempertimbangkan faktor eksternal lain di luar data ulasan yang dianalisis.

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

1. Mengetahui dan menganalisis sentimen yang diungkapkan oleh pengunjung Taman Safari Indonesia melalui ulasan di *Google Maps* untuk memahami persepsi umum terhadap pengalaman mereka.
2. Mengidentifikasi tema-tema utama yang muncul dari analisis sentimen terhadap ulasan pengunjung, guna memberikan insight terkait aspek-aspek yang berpengaruh terhadap kepuasan pengunjung.
3. Mengevaluasi kinerja algoritma machine learning dalam mengklasifikasikan sentimen ulasan pengunjung untuk menentukan algoritma mana yang paling efektif dalam pengolahan data dengan kompleksitas tinggi
4. Menerapkan *Term Frequency* untuk merekomendasikan atraksi di Taman Safari Indonesia berdasarkan sentimen dan preferensi pengunjung, sehingga membantu pengunjung dalam memilih atraksi yang sesuai dengan minat mereka.

1.4.2 Manfaat Penelitian

1. Penelitian ini dapat memberikan informasi bagi pengelola Taman Safari Indonesia mengenai sentimen pengunjung, yang dapat digunakan untuk meningkatkan kualitas layanan dan fasilitas. Dengan memahami pengalaman pengunjung, pengelola dapat melakukan perbaikan yang lebih terarah
2. Dengan mengidentifikasi tema utama dalam ulasan, penelitian ini memberikan wawasan tentang aspek-aspek spesifik yang paling dihargai dan yang perlu diperbaiki, memungkinkan pengelola untuk menyusun strategi yang lebih baik dalam mengelola atraksi dan pelayanan
3. Penelitian ini menunjukkan potensi penggunaan algoritma machine learning dalam menganalisis sentimen, yang dapat menjadi acuan

bagi peneliti dan praktisi di bidang data *science* untuk menerapkan metode serupa dalam analisis data ulasan di berbagai sektor.

4. Dengan menerapkan *Term Frequency*, penelitian ini akan membantu peneliti lain yang ingin mengembangkan aplikasi menggunakan penerapan *Term Frequency* untuk menemukan atraksi yang sesuai dengan minat dan preferensi pengunjung.

1.5 Sistematika Penulisan

BAB I: PENDAHULUAN

Bab ini memperkenalkan penelitian mengenai analisis sentimen dan rekomendasi berbasis ulasan *Google Maps* untuk Taman Safari Indonesia. Dengan meningkatnya dampak ulasan *online* terhadap pengambilan keputusan, analisis terhadap sentimen pengunjung yang dibagikan di *platform* tersebut menjadi krusial. Penelitian ini bertujuan untuk mengklasifikasikan ulasan pengunjung menjadi kategori sentimen positif, negatif, atau netral menggunakan tiga algoritma: *Support Vector Machine (SVM)*, *Logistic Regression*, *Naive Bayes*, *Random Forest*, dan *K-Nearest Neighbors*. Selain itu, penelitian juga mengembangkan rekomendasi berbasis *Term Frequency* yang membantu pengunjung dalam memilih atraksi sesuai dengan preferensi mereka. Diperkirakan, hasil dari penelitian dapat memberikan wawasan untuk meningkatkan kualitas layanan Taman Safari Indonesia sembari membantu pengunjung dalam membuat keputusan yang lebih informatif. Batasan penelitian mencakup analisis terhadap ulasan yang diambil dari *Google Maps*, dengan fokus pada sentimen dan rekomendasi yang berbasis teks.

BAB II: LANDASAN TEORI

Bab ini menjelaskan konsep-konsep utama yang diintegrasikan dalam penelitian, seperti analisis sentimen yang bertujuan untuk mengklasifikasikan teks ulasan ke dalam kategori positif, negatif, dan netral. Algoritma yang digunakan termasuk *Support Vector Machine (SVM)*, *Logistic Regression*, *Naive Bayes*, *K-Nearest Neighbor (KNN)*, dan *Random Forest*, di mana masing-masing algoritma memiliki keunggulan dalam menangani data teks. Selain itu, bab ini juga membahas

Term Frequency sebagai metode rekomendasi yang menganalisis ulasan untuk memberikan saran atraksi yang relevan berdasarkan kesamaan konten yang terdapat dalam ulasan pengunjung. Landasan teori ini mendasari pemilihan metode dan pendekatan yang digunakan dalam penelitian untuk analisis yang efektif terhadap sentimen dan rekomendasi.

BAB III: METODOLOGI PENELITIAN

Bab ini menguraikan pendekatan penelitian yang diambil dalam proyek ini, mencakup metode pengumpulan data dari ulasan yang diambil dari *Google Maps*. Teknik analisis terdiri dari penerapan algoritma *K-Nearest Neighbor (KNN)*, *Random Forest*, *Support Vector Machine*, *Logistic Regression*, dan *Naive Bayes* untuk tujuan analisis sentimen. Data yang diperoleh dari ulasan akan diproses secara sistematis untuk mengklasifikasikan sentimen serta menghasilkan rekomendasi berbasis *Term Frequency*. Model yang dibangun akan dievaluasi menggunakan metrik akurasi untuk memastikan efektivitas dalam memprediksi sentimen serta memberikan rekomendasi yang relevan bagi pengunjung Taman Safari Indonesia, sehingga memberikan kontribusi yang signifikan terhadap pengambilan keputusan yang lebih baik.

BAB IV: ANALISIS DAN HASIL PENELITIAN

Dalam bab ini, analisis dan hasil penelitian dipaparkan berdasarkan langkah-langkah yang dilakukan untuk mengelola dan memproses *dataset* yang diperoleh dari aplikasi *Google Maps Extractor*. Penggunaan alat *CSV Editor* dijelaskan, baik untuk memodifikasi *dataset* dengan menghapus kolom yang tidak diperlukan maupun untuk fokus pada kolom-kolom yang relevan seperti nama pengulas, *rating*, tanggal publikasi, dan teks ulasan. Proses *preprocessing* meliputi penanganan nilai yang hilang, memeriksa duplikat, serta tokenisasi, penghilangan *stopwords*, dan *stemming*. Setelah itu, analisis sentimen dilakukan menggunakan teknik yang telah disebutkan, seperti SVM dan *Logistic Regression*. Hasil menunjukkan bahwa sebagian besar ulasan bersifat positif, dengan analisis mendalam memperlihatkan kualitas atraksi, layanan, dan pengalaman interaksi dengan hewan sebagai tema utama. Bab ini menyoroti kinerja berbagai algoritma

dalam kategorisasi data sentimen serta penemuan pola yang berguna untuk analisis dan pengembangan rekomendasi.

BAB V: KESIMPULAN DAN SARAN

Bab ini memberikan kesimpulan dari penelitian yang menunjukkan bahwa pengalaman pengunjung Taman Safari Indonesia umumnya positif, berdasarkan analisis sentimen yang dilakukan. Model analisis berhasil mengklasifikasikan ulasan menjadi tiga kategori sentimen dan mengidentifikasi tema-tema penting dalam pengalaman pengunjung, seperti kepuasan terhadap layanan dan kualitas atraksi. Selain itu, bab ini menyoroti pentingnya pemilihan algoritma dalam analisis teks, dengan SVM dan *Logistic Regression* menunjukkan performa yang lebih baik dibandingkan dengan *Naive Bayes*. Saran untuk penelitian selanjutnya mencakup eksplorasi penggunaan algoritma lain seperti *Random Forest* atau KNN untuk meningkatkan akurasi klasifikasi, serta mendorong penelitian di industri atau lokasi yang berbeda untuk memperluas pemahaman tentang pengalaman pengunjung di taman hiburan.

