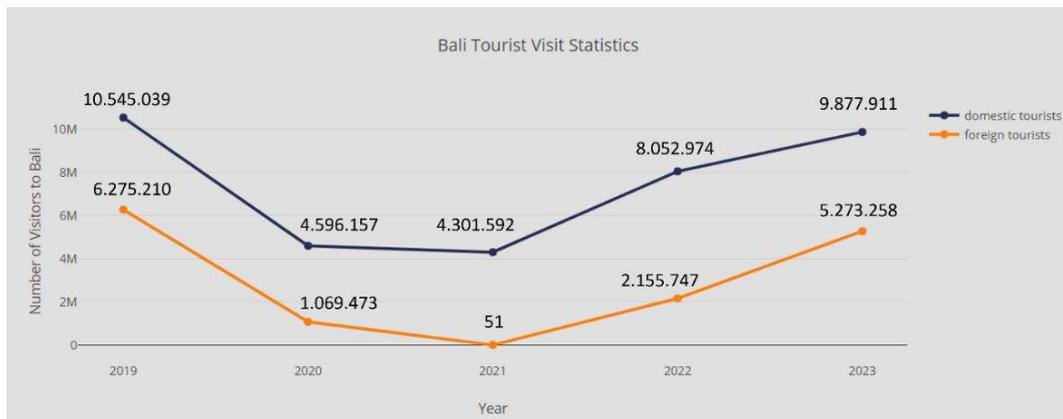


BAB I

PENDAHULUAN

1.1 Latar Belakang

Pariwisata telah menjadi sektor prioritas dalam pembangunan perekonomian Indonesia (Antara & Sumarniasih, 2017). Indonesia memiliki kota-kota yang kaya akan destinasi wisata salah satunya adalah Bali. Bali merupakan salah satu destinasi wisata paling populer di Indonesia karena memiliki beragam destinasi wisata (Utami & Erfina, 2022). Gambar 1.1 menampilkan sebuah grafik data kunjungan wisatawan ke Bali dari Januari 2018 – Desember 2023 yang bersumber dari situs resmi Badan Pusat Statistika (BPS) Provinsi Bali pada tanggal 30 Mei 2024. Dari data tersebut, tercatat sebanyak 9.877.911 wisatawan nusantara dan 5.273.258 wisatawan mancanegara melakukan perjalanan atau kunjungan ke Bali pada Tahun 2023. Dapat dilihat bahwa kunjungan wisatawan ke Bali mengalami kenaikan dalam 2 tahun belakangan ini.



Gambar 1.1 Statistik Kunjungan Wisatawan Bali

Menurut *World Tourism Organization* (UNWTO) (2007) sebagai organisasi pariwisata dunia, terdapat enam aspek yang menjadi hal dasar yang dapat menarik wisatawan pada sebuah destinasi pariwisata. Enam aspek tersebut diantaranya adalah *attractions*, *amenity*, *accessibility*, *image*, *price*, dan *human resource*. Semua aspek yang telah disebutkan perlu diperhatikan agar tercipta suatu destinasi pariwisata yang ramah wisatawan (Arianto & Budi, 2023). UNWTO juga menjelaskan bahwa terdapat sebuah konsep yang menggambarkan sebuah perjalanan wisatawan (dimulai dari tahap perencanaan hingga tahap mengenang

kembali pengalaman berwisata) yang akan mempengaruhi sebuah pengalaman wisata wisatawan tersebut hal ini disebut *tourist journey*. Terdapat tahapan *planning* dalam sebuah *tourist journey*, sebuah tahapan saat wisatawan akan mencari informasi lebih lanjut dan spesifik terkait destinasi wisata yang akan dituju. Sebagai sumber informasi, wisatawan cenderung menggunakan sumber informasi yang dianggap berkredibilitas tinggi, seperti informasi mulut ke mulut atau *word-of-mouth (WOM)* (Becken & Wilson, 2007). Ulasan pada situs pariwisata merupakan sebuah contoh dari electronic WOM (eWOM) (Tan & Chen, 2012).

Ulasan *online* yang terdapat pada situs wisata seperti situs TripAdvisor memiliki peran yang signifikan dalam memengaruhi keputusan wisatawan. Berdasarkan survei Ipsos MORI yang dilakukan bersama TripAdvisor, 72% responden selalu atau sering membaca ulasan sebelum membuat keputusan berwisata. TripAdvisor merupakan salah satu situs wisata populer dimana pengguna dapat berbagi pengalaman mereka. Situs ini memberikan informasi mengenai hotel, restoran, dan kegiatan berdasarkan ulasan dari wisatawan yang telah berada di sana sebelumnya. TripAdvisor menyediakan lebih dari satu miliar ulasan dan opini dari para wisatawan, serta membantu para pengguna dalam merencanakan, memesan, dan menjalani perjalanan mereka. Opini dan pengalaman positif atau negatif dalam ulasan online dapat mempengaruhi persepsi calon wisatawan tentang destinasi wisata tertentu. Volume besar data ulasan di situs TripAdvisor membuatnya sulit untuk secara manual mengevaluasi sentimen dari setiap ulasan. Untuk dapat melakukan sebuah evaluasi terhadap sentimen dari setiap ulasan maka diperlukan adanya implementasi dari sebuah teknik yang disebut analisis sentimen.

Analisis sentimen adalah suatu teknik mengekstrak data teks untuk mendapatkan informasi tentang sentimen bernilai positif, netral maupun negatif. (Sari & Hayuningtyas, 2019). Informasi ini dapat memberikan wawasan berharga yang dapat digunakan oleh berbagai pihak diantaranya pemerintah, industri pariwisata, dan pengelola tempat wisata untuk meningkatkan kualitas layanan, menanggapi keluhan, dan meningkatkan pengalaman pengunjung. *Natural Language Processing (NLP)* digunakan sebagai metode dalam analisis sentimen untuk mengolah data teks. NLP atau pemrosesan bahasa alami adalah teknik komputasi untuk melakukan analisis dan merepresentasikan bahasa alami dengan

tingkatan analisis linguistik yang memiliki tujuan untuk memproses bahasa seperti manusia dengan pendekatan teoritis (Liddy, 2001).

Penelitian terkait analisis sentimen sudah banyak dilakukan terhadap berbagai topik seperti produk, layanan, pariwisata, politik, dan topik lainnya menggunakan sumber data yang beragam. Diantaranya penelitian analisis sentimen yang dilakukan oleh Yang et al. (2020) terhadap data ulasan produk *e-commerce* dalam bahasa Cina menggunakan model baru berdasarkan leksikon sentimen yang dikombinasikan dengan *Convolutional Neural Network (CNN)*. Terdapat juga studi analisis sentimen yang berfokus pada pariwisata. Penelitian yang melakukan analisis sentimen pariwisata selama pandemi Covid-19 menggunakan data dari Twitter (Flores-Ruiz et al., 2021; Leelawat et al., 2022) dan menggunakan data dari Instagram (Priadana & Rizal, 2021). Penelitian yang berfokus pada pariwisata di Indonesia juga dilakukan oleh Oktaviana et al. (2021) yang mengevaluasi kinerja *Long Short-Term Memory (LSTM)* dan pembobotan *word2vec* dalam analisis sentimen ulasan *Google Maps* mengenai pantai-pantai di Yogyakarta. Matrutty et al. (2023) melakukan penelitian *review* hotel di Manado dari TripAdvisor menggunakan Naïve Bayes, Christanto & Singgalen (2022) menganalisis data *review* restoran di Labuan Bajo menggunakan beberapa algoritma *machine learning*. Penelitian terhadap Tanah Lot Bali yang memiliki *title Traveller's Choice 2023* (*title* yang didapatkan karna objek wisata tersebut mendapatkan ulasan luar biasa dari wisatawan dan termasuk dalam 1% properti terbaik di dunia) pada situs TripAdvisor belum pernah dilakukan.

Oleh karena itu, penelitian ini bertujuan untuk melakukan perancangan dan implementasi model komputasi serta menganalisis hasil eksperimen sistem analisis sentimen terhadap data ulasan Tanah Lot Bali di Tripadvisor dengan menggunakan metode *random forest* (Breiman, 2001). Algoritma *random forest* merupakan algoritma *machine learning* yang digunakan untuk melakukan prediksi sentimen terhadap data teks ulasan (El Naqa & Murphy, 2015). *Random forest* dipilih karena memiliki performa yang baik pada beberapa penelitian seperti penelitian di bidang edukasi (Utari et al., 2020; Xu & Yin, 2021), bidang kesehatan (Iwendi et al., 2020), dan bidang *finance* (Abraham et al., 2022; Yin et al., 2023). Hasil yang diperoleh dari penelitian ini diharapkan menjadi informasi yang dapat memberikan wawasan

yang berharga dan mendukung pengambilan keputusan yang lebih baik dalam upaya meningkatkan pengalaman wisata wisatawan.

1.2 Rumusan Masalah

Sesuai latar belakang masalah yang ada, maka muncul rumusan masalah sebagai berikut:

1. Bagaimana perancangan model komputasi sistem analisis sentimen terhadap data ulasan Tanah Lot Bali di Tripadvisor dengan metode *Random Forest*?
2. Bagaimana implementasi model komputasi sistem analisis sentimen menggunakan algoritma *Random Forest* sebagai algoritma klasifikasi sentimen data ulasan Tanah Lot Bali di Tripadvisor?
3. Bagaimana analisis terhadap hasil eksperimen sistem analisis sentimen data ulasan pengunjung Tanah Lot Bali di Tripadvisor dengan metode *random forest*?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah tersebut, maka tujuan dari penelitian ini adalah sebagai berikut:

1. Merancang model komputasi sistem analisis sentimen terhadap data ulasan Tanah Lot Bali di Tripadvisor dengan metode *random forest*.
2. Mengimplementasikan model komputasi analisis sentimen serta *Random Forest* sebagai algoritma klasifikasi sentimen dengan penerapan pemodelan perangkat lunak.
3. Menganalisa hasil eksperimen sistem analisis sentimen terhadap data ulasan Tanah Lot Bali di Tripadvisor dengan metode *random forest*.

1.4 Manfaat Penelitian

Manfaat yang diperoleh dari penelitian ini antara lain sebagai berikut:

1. Diharapkan dapat menambah pengetahuan tentang analisis sentimen, *natural language processing*, dan *random forest* serta penerapannya

dalam analisis sentimen pada data ulasan pengunjung Tanah Lot Bali di TripAdvisor.

2. Dapat memperoleh informasi yang diharapkan dapat memberikan wawasan yang berharga dan mendukung pengambilan keputusan yang lebih baik dalam upaya meningkatkan pengalaman wisata wisatawan.
3. Dapat menjadi salah satu referensi dalam penelitian mendatang terkait dengan analisis sentimen menggunakan bahasa pemrograman Python.

1.5 Batasan Masalah

Batasan masalah dalam penelitian ini antara lain sebagai berikut:

1. Penelitian ini menggunakan data ulasan pada situs Tripadvisor.co.id berbahasa Indonesia dari November 2010 hingga September 2023.
2. Ulasan dilabeli dan diklasifikasi menjadi 1 sentimen (positif, netral, dan negatif) secara keseluruhan.
3. Aspek pada *aspect filtering* adalah *attractions*, *amenities*, *accessibility*, *image*, *price*, dan *human resource*.
4. Model komputasi dibangun dengan menggunakan bahasa pemrograman *Python*.

1.6 Sistematika Penulisan

Berikut ini merupakan sistematika penulisan dan rincian penjelasan tiap bab.

BAB I PENDAHULUAN

Bab ini berisi latar belakang, rumusan masalah, batasan masalah, tujuan penelitian yang akan dilakukan, serta metodologi penelitian dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini berisi penjelasan tentang teori-teori dan konsep algoritma yang digunakan dalam penelitian.

BAB III METODOLOGI PENELITIAN

Bab ini berisi penjelasan tentang desain penelitian serta perangkat-perangkat pendukung penelitian.

BAB IV HASIL PENELITIAN DAN PEMBAHASAN

Bab ini berisi tentang pembahasan mengenai perancangan model komputasi, implementasi model komputasi, skenario eksperimen, serta pembahasan hasil dan analisis dari eksperimen yang dilakukan.

BAB V KESIMPULAN DAN SARAN

Bab ini menjelaskan kesimpulan yang merupakan jawaban dari masalah dalam penelitian, serta berisi saran yang dapat menjadi rujukan untuk penelitian selanjutnya.