

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Perkembangan teknologi informasi telah mendorong transformasi digital pada berbagai sektor, termasuk sektor pariwisata. Pariwisata merupakan salah satu sektor yang memiliki peran penting dalam meningkatkan perekonomian daerah melalui peningkatan kunjungan wisatawan dan pengembangan potensi destinasi wisata (Khumaidi et al., 2025). Provinsi Jawa Timur, khususnya wilayah Madura, memiliki berbagai destinasi wisata pantai yang menarik seperti Pantai Lombang, Pantai Slopeng, Pantai Siring Kemuning, dan Pantai Lon Malang yang menawarkan keindahan alam, fasilitas wisata, serta aktivitas rekreasi yang beragam. Banyaknya destinasi wisata pantai di Madura menyebabkan wisatawan sering mengalami kesulitan dalam menentukan lokasi wisata yang sesuai dengan kebutuhan dan preferensi pengguna, seperti pantai dengan fasilitas lengkap, kebersihan yang baik, maupun suasana yang tenang (Andriani et al., 2025). Wisata pantai di Madura merupakan salah satu destinasi yang banyak dicari wisatawan melalui Google Maps sebelum menentukan tujuan perjalanan. Sistem pada Google Maps bekerja dengan mengumpulkan informasi dari pengguna dan menampilkannya dalam bentuk data lokasi, rating, serta ulasan yang dapat diakses oleh calon pengunjung. Informasi tersebut membantu wisatawan memahami karakteristik suatu pantai sebelum melakukan kunjungan (Arianto et al., 2025).

Meskipun informasi wisata pantai di Google Maps cukup banyak tersedia, masih terdapat beberapa permasalahan dalam proses pencarian destinasi wisata. Pengguna sering mengalami kesulitan dalam menemukan pantai yang sesuai dengan kebutuhan dan preferensi pengguna karena banyaknya jumlah destinasi wisata yang tersedia (Cholil et al., 2023). Selain itu, informasi yang ditampilkan pada Google Maps umumnya hanya berupa daftar lokasi, rating, dan ulasan tanpa memberikan rekomendasi yang lebih personal sesuai kebutuhan pengguna (Febrian & Nasir, 2021). Permasalahan lain yang sering muncul adalah pengguna harus membaca banyak ulasan secara manual untuk mengetahui karakteristik suatu pantai, seperti apakah pantai tersebut memiliki fasilitas lengkap, cocok untuk

keluarga, memiliki akses jalan yang mudah, atau memiliki suasana yang tenang(Maulidya Prastita Syah et al., 2025). Proses tersebut membutuhkan waktu cukup lama dan kurang efisien. Tidak semua pengguna juga mampu menyimpulkan informasi penting dari banyaknya ulasan yang tersedia. Banyaknya informasi pada platform wisata digital justru dapat menyebabkan information *overload* sehingga pengguna mengalami kesulitan dalam menentukan pilihan destinasi yang paling sesuai(Kopi, 2024).

Berdasarkan berbagai permasalahan tersebut, penelitian ini memfokuskan pada masalah sulitnya pengguna memperoleh rekomendasi destinasi pantai yang sesuai dengan preferensi pengguna berdasarkan informasi ulasan pengunjung. Sebagai contoh, seorang pengguna ingin mencari pantai dengan suasana bersih, ombak tenang, akses mudah, dan cocok untuk keluarga. Namun, Google Maps belum mampu secara spesifik memberikan rekomendasi berdasarkan karakteristik tersebut. Pengguna harus membaca satu per satu ulasan dari berbagai pantai untuk memperoleh informasi yang diinginkan. Kondisi tersebut menyebabkan proses pencarian destinasi menjadi kurang efektif dan memerlukan waktu yang lebih lama(Putri & Apsiswanto, 2024). Oleh karena itu, diperlukan suatu sistem rekomendasi yang mampu membantu pengguna memperoleh destinasi pantai yang relevan berdasarkan isi ulasan pengunjung.

Berbagai penelitian sebelumnya telah mengembangkan solusi sistem rekomendasi menggunakan beberapa metode yang berbeda. Salah satu metode yang sering digunakan adalah *Collaborative Filtering* yang memberikan rekomendasi berdasarkan kesamaan perilaku antar pengguna. Metode ini cukup efektif apabila tersedia data interaksi pengguna dalam jumlah besar. Selain itu, terdapat metode *Hybrid Recommendation* yang menggabungkan beberapa pendekatan untuk meningkatkan kualitas rekomendasi. Metode *Content Based Filtering* juga banyak digunakan dalam sistem rekomendasi karena mampu memberikan rekomendasi berdasarkan karakteristik atau konten dari suatu item.

Meskipun berbagai metode sistem rekomendasi telah banyak diterapkan, masing-masing metode masih memiliki keterbatasan. *Collaborative Filtering* memiliki kelemahan berupa *cold start problem* dan ketergantungan yang tinggi terhadap data interaksi pengguna. Penelitian mengenai implementasi *Collaborative*

*Filtering* menggunakan algoritma *Cosine Similarity* dan *Jaccard Similarity* pada sistem *e-commerce* menunjukkan bahwa kualitas rekomendasi dapat menurun ketika data pengguna bersifat *sparse* atau jumlah rating yang tersedia terbatas. Kondisi tersebut menyebabkan sistem kesulitan dalam menemukan pola kesamaan antar pengguna sehingga rekomendasi yang dihasilkan menjadi kurang optimal (Waskito et al., 2024). Selain itu, metode berbasis *deep learning* seperti BERT juga telah banyak digunakan dalam pengolahan teks karena mampu memahami hubungan semantik dan konteks kalimat dengan lebih baik dibandingkan metode representasi teks tradisional. Penelitian Perbandingan Representasi Teks Tf-Idf Dan Bert Terhadap Akurasi *Cosine Similarity* Dalam Penilaian Otomatis Jawaban Berbasis Teks menunjukkan bahwa BERT menghasilkan korelasi yang lebih tinggi dalam pengukuran kemiripan teks karena kemampuannya memahami makna semantik. Namun, metode ini memiliki tingkat kompleksitas yang lebih tinggi dibandingkan TF-IDF sehingga implementasinya membutuhkan proses komputasi yang lebih besar (Maulidya Prastita Syah et al., 2025). Sementara itu, *Content-Based Filtering* memiliki keunggulan karena tidak bergantung pada data pengguna lain dan dapat memberikan rekomendasi berdasarkan karakteristik item yang tersedia. Namun, kualitas rekomendasi yang dihasilkan sangat bergantung pada kualitas fitur dan informasi yang dimiliki setiap item (Rekomendasi et al., 2025). Oleh karena itu, diperlukan metode yang mampu memanfaatkan informasi ulasan pengunjung secara efektif, memiliki kompleksitas yang relatif rendah, serta sesuai dengan karakteristik data teks yang diperoleh dari Google Maps.

Berdasarkan kelebihan dan kekurangan dari berbagai metode tersebut, penelitian ini memilih menggunakan metode *Content Based Filtering* dengan pendekatan TF-IDF dan *Cosine Similarity*. Metode TF-IDF digunakan untuk menghitung bobot kata pada ulasan pengunjung sehingga sistem dapat mengetahui kata-kata penting yang merepresentasikan karakteristik suatu pantai. Setelah proses pembobotan dilakukan, metode *Cosine Similarity* digunakan untuk menghitung tingkat kemiripan antara input preferensi pengguna dengan data ulasan pantai yang tersedia. Pemilihan metode ini dilakukan karena TF-IDF dan *Cosine Similarity* memiliki proses yang relatif sederhana, efisien, serta cocok digunakan pada data

teks berupa ulasan pengunjung. Selain itu, metode ini mampu menghasilkan rekomendasi berdasarkan isi informasi secara langsung tanpa memerlukan riwayat pengguna dalam jumlah besar. Penelitian Pengembangan Sistem Rekomendasi Pembimbing Tugas Akhir Menggunakan Teknik *Content Based Filtering* menunjukkan bahwa kombinasi *Content-Based Filtering*, TF-IDF, dan *Cosine Similarity* mampu menghasilkan rekomendasi yang akurat serta meningkatkan efisiensi proses pemilihan berdasarkan tingkat kemiripan dokumen (Astuti & Andriyani, 2025).

Penelitian ini berhasil membangun sistem rekomendasi destinasi wisata pantai di Madura yang mampu membantu wisatawan memperoleh rekomendasi destinasi secara lebih cepat, efektif, dan sesuai dengan preferensi pengguna. Sistem menggunakan metode *Content-Based Filtering* dengan pembobotan TF-IDF dan perhitungan *Cosine Similarity* untuk mengukur tingkat kemiripan antara preferensi pengguna dan informasi yang terdapat pada ulasan destinasi wisata. Sistem juga menyediakan dua skenario pencarian, yaitu berdasarkan query dan pemilihan fitur wisata, serta dilengkapi dengan filter berdasarkan kabupaten. Berdasarkan hasil evaluasi, sistem mampu memberikan rekomendasi yang relevan, sehingga dapat membantu wisatawan dalam menentukan destinasi wisata pantai yang sesuai dengan kebutuhan dan preferensi wisatawan.

## **1.2 Rumusan Masalah**

### **1.2.1 Permasalahan**

Wisatawan mengalami kesulitan dalam menemukan destinasi pantai di Madura yang sesuai dengan preferensi pengguna karena banyaknya pilihan destinasi yang tersedia. Informasi yang terdapat pada Google Maps umumnya berupa rating dan ulasan pengunjung dalam jumlah besar sehingga pengguna harus membaca ulasan secara manual untuk memperoleh informasi yang dibutuhkan. Proses tersebut membutuhkan waktu yang relatif lama dan kurang efisien dalam membantu pengambilan keputusan destinasi wisata.

### **1.2.2 Metode Usulan**

Untuk mengatasi permasalahan tersebut, penelitian ini mengusulkan pembangunan sistem rekomendasi destinasi pantai di Madura menggunakan metode *Content-Based Filtering*. Ulasan pengunjung yang diperoleh dari

Google Maps akan diolah menggunakan TF-IDF untuk menentukan bobot kata dan *Cosine Similarity* untuk menghitung tingkat kemiripan antara preferensi pengguna dengan karakteristik destinasi pantai yang terdapat dalam ulasan.

### **1.2.3 Pertanyaan Penelitian**

Bagaimana kinerja sistem rekomendasi destinasi wisata pantai di Madura menggunakan metode *Content-Based Filtering* dengan pembobotan TF-IDF dan perhitungan *Cosine Similarity* berdasarkan nilai *Precision@10* dan *Recall@10*?

## **1.3 Tujuan dan Manfaat**

### **1.3.1 Tujuan Penelitian**

Tujuan penelitian ini adalah mengimplementasikan metode *Content-Based Filtering* dengan pembobotan TF-IDF dan perhitungan *Cosine Similarity* untuk membangun sistem rekomendasi destinasi wisata pantai di Madura berdasarkan ulasan pengunjung Google Maps serta mengevaluasi kinerja sistem menggunakan metrik *Precision@10* dan *Recall@10*.

### **1.3.2 Manfaat Penelitian**

Penelitian ini diharapkan dapat membantu wisatawan dalam memperoleh rekomendasi destinasi pantai di Madura yang sesuai dengan preferensi dan kebutuhan berdasarkan ulasan pengunjung melalui penerapan metode *Content-Based Filtering*. Selain itu, penelitian ini diharapkan dapat menjadi acuan dalam pengembangan sistem rekomendasi berbasis teks serta menjadi referensi bagi penelitian selanjutnya yang berkaitan dengan pemanfaatan ulasan pengguna dengan memanfaatkan metode *Content-Based Filtering*, TF-IDF, dan *Cosine Similarity*.

## **1.4 Batasan Masalah**

Batasan masalah dalam penelitian ini adalah objek penelitian difokuskan pada destinasi wisata pantai yang berada di wilayah Pulau Madura, Jawa Timur. Data yang digunakan berupa ulasan pengunjung yang diperoleh dari Google Maps melalui proses *scraping*. Sistem rekomendasi dibangun menggunakan metode *Content-Based Filtering* dengan pembobotan TF-IDF dan perhitungan tingkat

kemiripan menggunakan *Cosine Similarity*. Ulasan pengunjung digunakan sebagai representasi dokumen dalam proses pembobotan TF-IDF dan perhitungan kemiripan, serta dilakukan ekstraksi fitur untuk memperoleh karakteristik destinasi wisata yang meliputi aspek kebersihan, fasilitas, aksesibilitas, suasana, dan aktivitas wisata.

Selain itu, sistem hanya menampilkan 10 rekomendasi pantai teratas (Top-10) berdasarkan nilai kemiripan tertinggi. Evaluasi sistem dilakukan menggunakan metrik *Precision* dan *Recall*. Sistem rekomendasi yang dibangun menggunakan informasi yang diperoleh ulasan pengunjung, namun belum mempertimbangkan faktor lain seperti lokasi pengguna secara *real-time*, harga tiket masuk, jarak tempuh, kondisi cuaca, maupun riwayat preferensi pengguna dalam memberikan rekomendasi.

## **1.5 Sistematikan Penulisan**

Sistematika penulisan penelitian ini sebagai berikut:

### **BAB I PENDAHULUAN**

Bab ini mencakup latar belakang masalah, perumusan masalah, tujuan dan manfaat, penggambaran ruang lingkup masalah, serta struktur sistematis penulisan.

### **BAB II KAJIAN PUSTAKA**

Bab ini membahas teori-teori yang menjadi dasar dalam penelitian, meliputi konsep pariwisata dan wisata pantai, sistem rekomendasi, ulasan pengunjung, *preprocessing text*, ekstraksi fitur wisata, metode *Content-Based Filtering*, pembobotan *Term Frequency–Inverse Document Frequency* (TF-IDF), perhitungan *Cosine Similarity*, evaluasi sistem menggunakan *Precision@10* dan *Recall@10*, serta penelitian terdahulu yang berkaitan dengan sistem rekomendasi dan pengolahan teks.

### **BAB III METODOLOGI PENELITIAN**

Bab ini menjelaskan tahapan penelitian yang dilakukan mulai dari pengumpulan data ulasan wisata pantai melalui proses scraping Google Maps menggunakan Apify, *preprocessing text*, ekstraksi fitur wisata, pembobotan TF-IDF, perhitungan *Cosine Similarity*, proses ranking rekomendasi, hingga evaluasi model menggunakan *Precision@10* dan *Recall@10*. Selain itu, bab ini juga

menjelaskan skenario pengujian yang digunakan untuk menguji fungsionalitas sistem rekomendasi berbasis Streamlit yang dikembangkan.

#### BAB IV HASIL DAN PEMBAHASAN

Bab ini membahas hasil penelitian yang meliputi gambaran umum data, tahapan preprocessing, ekstraksi fitur, penggabungan dokumen, pembobotan TF-IDF, perhitungan *Cosine Similarity*, implementasi sistem rekomendasi, serta hasil pengujian menggunakan metrik *Precision@10* dan *Recall@10* yang disertai pembahasan hasil penelitian

#### BAB V PENUTUP

Bab ini berisi kesimpulan yang diperoleh berdasarkan hasil penelitian serta saran yang diharapkan dapat menjadi masukan untuk pengembangan sistem rekomendasi dan penelitian selanjutnya.