

Nama : Aditya Ulil Albab

NIM : A11.2023.15093

Mata Kuliah : Sistem Temu Kembali Informasi (STKI)

Dosen Pengampu : Abu Salam, M.Kom

## 1. Pendahuluan

### 1.1 Tujuan

Tujuan utama proyek ini adalah mengembangkan sebuah search engine berbasis web untuk membantu pengguna mencari informasi pantai di Karimunjawa berdasarkan rating, sentimen, dan kata kunci. Sistem ini juga menyediakan rekomendasi pantai berdasarkan sentimen dan ranking otomatis berdasarkan rating, jumlah review, dan persentase sentimen positif.

### 1.2 Ruang Lingkup

Proyek ini mencakup pengembangan aplikasi web interaktif berbasis Streamlit yang dilengkapi dengan proses preprocessing data ulasan pantai Karimunjawa, implementasi sistem pencarian multi-filter (rating, sentimen, dan kata kunci), visualisasi data berupa ranking pantai dan rekomendasi berdasarkan sentimen, serta penyajian statistik ulasan secara real-time.

### 1.3 Kontribusi Proyek

Proyek ini berkontribusi dalam menyediakan alat pencarian pantai Karimunjawa yang mudah digunakan, memberikan insight berbasis sentimen dan rating ulasan pengunjung, serta mendorong pengambilan keputusan wisata yang lebih informasional bagi calon wisatawan.

## 2. Data & Preprocessing

### 2.1 Ringkas

Data yang digunakan berasal dari file CSV ('label\_data.csv') yang berisi 720 ulasan dari 15 pantai di Karimunjawa. Data terdiri dari kolom:

- 'title': Nama pantai
- 'stars': Rating 1–5
- 'text': Ulasan teks
- 'sentimen': Label sentimen (positive, neutral, negative)

## 2.2 Contoh Before/After

Before

"pantai yang menakjubkan dengan pasirpasir putih..."

After (setelah preprocessing):

"pantai menakjubkan pasirpasir putih"

## 2.3 Preprocessing yang dilakukan mencakup:

1. Normalisasi teks: Mengganti singkatan dan slang (misal: `yg` → `yang`, `bgtttt` → `banget`).
2. Stopword removal: Menghapus kata tidak penting menggunakan Sastrawi.
3. Tokenisasi: Memisahkan teks menjadi kata-kata.
4. Stemming: Mengubah kata ke bentuk dasar menggunakan Sastrawi.
5. Labelling: Melabel hasil sentiment dibagi menjadi positive, negative, neutral

## 3. Metode

Metode yang digunakan dalam sistem ini meliputi:

- Pencarian berbasis teks dengan preprocessing lowercase dan penghapusan karakter khusus.
- Filter multi-kriteria (rating, sentimen, kata kunci).
- Algoritma ranking berdasarkan:
  - Rata-rata rating
  - Persentase sentimen positif
  - Logaritma jumlah review untuk menghindari bias populasi.
- Visualisasi interaktif menggunakan Streamlit.

## 4. Arsitektur Search Engine

### 4.1 Diagram Alir Singkat

Input Pengguna → Filter Rating → Filter Sentimen → Pencarian Kata Kunci →

Preprocessing Teks → Pencocokan → Hasil Filtered →

Perhitungan Statistik & Ranking → Tampilan Visual & Tabel

## 5. Eksperimen & Evaluasi

### 5.1 Skenario

Skenario pengujian meliputi:

1. Pencarian dengan filter rating 4–5.
2. Filter sentimen positif.
3. Pencarian kata kunci seperti "sunset", "bersih", "indah".

### 5.2 Metrik

- Jumlah hasil pencarian
- Rata-rata rating
- Jumlah review positif/negatif
- Akurasi rekomendasi berdasarkan sentimen

### 5.3 Hasil

Hasil pengujian menunjukkan bahwa sistem berhasil menampilkan pantai terbaik berdasarkan perhitungan skor ranking, rekomendasi pantai berdasarkan sentimen berjalan sesuai ekspektasi, dan pencarian kata kunci mampu menemukan ulasan yang relevan dengan kata kunci yang dimasukkan pengguna.

### 5.4 Analisi

Analisis terhadap hasil menunjukkan bahwa ranking pantai sangat dipengaruhi oleh jumlah review dan persentase sentimen positif. Sistem juga terbukti responsif terhadap perubahan filter yang diterapkan, serta visualisasi ranking dengan badge emas, perak, dan perunggu berhasil meningkatkan engagement pengguna dalam berinteraksi dengan aplikasi.

## 6. Diskusi

### 6.1 Kelebihan

- Antarmuka intuitif dengan sidebar filter.
- Visualisasi ranking yang menarik.
- Kemampuan melihat detail review per pantai.
- Responsif terhadap interaksi pengguna.

## 6.2 Keterbatasan

- Data terbatas hanya 720 ulasan.
- Preprocessing masih sederhana, belum menggunakan embedding atau TF-IDF.
- Belum ada sistem login atau personalisasi.

## 6.3 Saran Pengembangan

- Integrasi dengan model NLP untuk analisis sentimen otomatis.
- Penambahan fitur “pantai terdekat” berdasarkan lokasi pengguna.
- Implementasi sistem rekomendasi berbasis collaborative filtering.
- Penambahan visualisasi peta interaktif.

## 7. Kesimpulan

Proyek Search Engine Pantai Karimunjawa telah berhasil dikembangkan dengan kemampuan pencarian multi-kriteria, ranking otomatis, dan rekomendasi berdasarkan sentimen. Sistem ini dapat membantu wisatawan dalam memilih pantai yang sesuai dengan preferensi mereka. Dengan pengembangan lebih lanjut, sistem ini dapat menjadi platform yang lebih komprehensif untuk pariwisata Karimunjawa.