

ABSTRAK

Sistem rekomendasi berperan penting dalam membantu pengguna menemukan informasi yang relevan di tengah banyaknya pilihan yang tersedia pada berbagai platform digital. Dalam konteks pariwisata, sistem rekomendasi dimanfaatkan untuk membantu pengguna menemukan destinasi wisata yang sesuai dengan preferensi pengguna. Permasalahan utama dalam sistem rekomendasi berbasis konten adalah bagaimana merepresentasikan informasi destinasi wisata agar tingkat kemiripan antar tempat dapat diukur secara akurat. Oleh karena itu, penelitian ini melakukan studi komparatif antara metode representasi teks berbasis TF-IDF dan IndoBERT, dengan mempertimbangkan IndoBERT *pre-trained* dan IndoBERT yang telah melalui proses *Fine-Tuning*, dalam penerapan *Content-Based Filtering* untuk rekomendasi destinasi wisata. TF-IDF digunakan sebagai pendekatan berbasis frekuensi yang sederhana dan efisien, sedangkan IndoBERT merupakan model bahasa Indonesia berbasis Transformer yang mampu menghasilkan representasi kontekstual. Representasi teks dari kedua metode digunakan untuk membentuk vektor fitur destinasi wisata, yang selanjutnya dihitung tingkat kemiripannya menggunakan *Cosine Similarity*. Evaluasi kinerja sistem rekomendasi dilakukan menggunakan metrik *Normalized Discounted Cumulative Gain* (NDCG) dengan skema *graded relevance* dan *binary relevance*. Hasil eksperimen menunjukkan bahwa IndoBERT dengan proses *Fine-Tuning* menggunakan *learning rate* 5×10^{-5} menghasilkan kinerja terbaik, dengan nilai NDCG *Graded* sebesar 0,727768 pada *cutoff* 5 dan 0,739136 pada *cutoff* 10, serta nilai NDCG *binary* sebesar 0,626778 pada *cutoff* 5 dan 0,625513 pada *cutoff* 10. Di sisi lain, TF-IDF menunjukkan efisiensi waktu komputasi tertinggi, dengan waktu maksimum tercepat sebesar 0,001468 detik pada konfigurasi unigram. Secara keseluruhan, penelitian ini menunjukkan bahwa representasi teks berbasis *embedding* kontekstual seperti IndoBERT lebih efektif dalam meningkatkan kualitas rekomendasi destinasi wisata, sementara TF-IDF tetap kompetitif sebagai metode *baseline* yang efisien.

Kata Kunci : Sistem Rekomendasi, *Content-Based Filtering*, *Normalized Discounted Cumulative Gain*, *Cosine Similarity*, TF-IDF, IndoBERT.