

Analisis Sentimen Terhadap Daya Tarik Wisata Kabupaten Karangasem Menggunakan Model IndoBERT

Oleh
Elma Regina Nababan, NIM 2115091040
Jurusan Teknik Informatika
Program Studi Sistem Informasi

ABSTRAK

Sektor pariwisata merupakan penopang utama perekonomian Bali, termasuk di Kabupaten Karangasem yang memiliki potensi besar melalui 28 destinasi wisata unggulan, namun tercatat adanya penurunan jumlah kunjungan wisatawan dari sekitar 1,3 juta pada tahun 2023 menjadi 1,2 juta pada tahun 2024 yang mengindikasikan adanya tantangan dalam pengelolaan destinasi. Penelitian ini bertujuan untuk menganalisis sentimen wisatawan secara otomatis guna memberikan gambaran objektif mengenai pengalaman pengunjung di platform Google Maps dan TripAdvisor melalui pengumpulan 22.835 data ulasan dengan model IndoBERT yang menggunakan teknik *fine-tuning* dan evaluasi *10-Fold Cross Validation*. Untuk mengatasi ketidakseimbangan data, penelitian ini membandingkan empat skenario penyeimbangan data yang menghasilkan performa sebagai berikut: skenario *Raw Data* memperoleh akurasi 0,84 dan *F1-Score* 0,73; skenario *Downsampling* memperoleh akurasi 0,75 dan *F1-Score* 0,75; skenario *Easy Data Augmentation* (EDA) memperoleh akurasi 0,86 dan *F1-Score* 0,79; serta skenario kombinasi *Downsampling* dan EDA yang menghasilkan akurasi 0,83 dengan *F1-Score* tertinggi sebesar 0,81. Meskipun skenario EDA unggul dalam akurasi, skenario kombinasi dianggap paling optimal dalam menangani klasifikasi di seluruh kelas sentimen, di mana analisis menunjukkan bahwa meskipun persepsi didominasi ulasan positif, aspek manajemen layanan dan transparansi biaya menjadi keluhan utama pada sentimen negatif yang perlu diperbaiki untuk meningkatkan kualitas pariwisata di Kabupaten Karangasem.

Kata Kunci: Analisis Sentimen, Karangasem, IndoBERT, Google Maps, TripAdvisor, *Easy Data Augmentation*, *Downsampling*

***Sentiment Analysis Of The Tourist Attractiveness Of Karangasem Regency
Using The IndoBERT Model***

By

Elma Regina Nababan, NIM 2115091040

Informatics Engineering

Information Systems Study Program

ABSTRACT

The tourism sector is the primary backbone of Bali's economy, including Karangasem Regency, which possesses significant potential through 28 leading tourist destinations; however, a decline in tourist visits from approximately 1.3 million in 2023 to 1.2 million in 2024 has been recorded, indicating challenges in destination management. This study aims to automatically analyze tourist sentiment to provide an objective overview of visitor experiences on Google Maps and TripAdvisor platforms by collecting 22,835 review data using the IndoBERT model with fine-tuning and 10-fold cross-validation evaluation. To address data imbalance, this study compares four data balancing scenarios, yielding the following performance results: the Raw Data scenario achieved an accuracy of 0.84 and an F1-score of 0.73; the Downsampling scenario achieved an accuracy of 0.75 and an F1-score of 0.75; the Easy Data Augmentation (EDA) scenario achieved an accuracy of 0.86 and an F1-score of 0.79; and the combined Downsampling and EDA scenario produced an accuracy of 0.83 with the highest F1-score of 0.81. Although the EDA scenario excelled in accuracy, the combined scenario is considered the most optimal for handling classification across all sentiment classes, where analysis reveals that while perceptions are dominated by positive reviews, service management and fee transparency remain the primary complaints in negative sentiments that need to be addressed to enhance tourism quality in Karangasem Regency.

Keywords: *Sentiment Analysis, Karangasem, IndoBERT, Google Maps, TripAdvisor, Easy Data Augmentation, Downsampling*