

ABSTRAK

Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna terhadap enam objek wisata di Provinsi Nusa Tenggara Timur dengan memanfaatkan data ulasan dalam jumlah besar yang diperoleh dari *Google Maps*. Data dikumpulkan melalui proses *scraping* menggunakan *SerpAPI*, kemudian dilakukan tahapan *Cleaning* dan pra-pemrosesan teks untuk meningkatkan kualitas data. Pelabelan sentimen dilakukan secara otomatis menggunakan model *IndoBERT* sehingga diperoleh tiga kelas sentimen, yaitu positif, negatif, dan netral. Representasi fitur teks dilakukan menggunakan metode *Term Frequency–Inverse Document Frequency (TF-IDF)*, kemudian diklasifikasikan menggunakan model *Support Vector Machine (SVM) baseline* dan model *SVM* teroptimasi dengan *GridSearchCV*. Hasil evaluasi menunjukkan bahwa model *SVM baseline* menghasilkan akurasi sebesar 83,87%, namun menunjukkan ketidakseimbangan performa antar kelas dengan nilai *Macro F1-score* sebesar 0,4287. Setelah dilakukan optimasi parameter menggunakan *GridSearchCV*, model *SVM* teroptimasi menghasilkan akurasi sebesar 78,27% dengan peningkatan nilai *Macro F1-score* menjadi 0,4818. Peningkatan tersebut menunjukkan perbaikan dalam kemampuan model mengenali kelas sentimen minoritas meskipun terjadi penurunan akurasi keseluruhan. Secara keseluruhan, model *SVM* teroptimasi memberikan hasil klasifikasi yang lebih seimbang dan representatif dalam menggambarkan persepsi wisatawan berdasarkan ulasan daring, sehingga dapat digunakan sebagai dasar analisis sentimen pada sektor pariwisata.

Kata Kunci: Analisis Sentimen, *Support Vector Machine (SVM)*, *TF-IDF*, *Google Maps*, Ulasan Wisata

ABSTRACT

This study aims to analyze user sentiment toward six tourist destinations in East Nusa Tenggara Province by utilizing a large volume of online review data obtained from Google Maps. The data were collected through a scraping process using SerpAPI, followed by data Cleaning and text preprocessing to improve data quality. Sentiment labeling was performed automatically using the IndoBERT model, resulting in three sentiment classes: positive, negative, and neutral. Text features were represented using the Term Frequency–Inverse Document Frequency (TF-IDF) method and classified using a baseline Support Vector Machine (SVM) model as well as an optimized SVM model with GridSearchCV. The evaluation results indicate that the baseline SVM model achieved an accuracy of 83.87%, but exhibited imbalanced performance across sentiment classes with a Macro F1-score of 0.4287. After parameter optimization using GridSearchCV, the optimized SVM model achieved an accuracy of 78.27% with an improved Macro F1-score of 0.4818. This improvement demonstrates enhanced capability in recognizing minority sentiment classes, despite a decrease in overall accuracy. Overall, the optimized SVM model provides a more balanced and representative classification of tourist perceptions based on online reviews, and can serve as a basis for sentiment analysis in the tourism sector.

Keywords: Sentiment Analysis, Support Vector Machine (SVM), TF-IDF, Google Maps, Tourism Reviews