ANALISIS SENTIMEN TERHADAP WISATA DI PROVINSI BALI MENGGUNAKAN METODE *SUPPORT VECTOR MACHINE (SVM)*

(Studi Kasus : Tanah Lot)

Maulana Indra Wicaksono Harjanto[1], Henny Wahyu Sulistyo[2], Dudi Irawan[3]

Program Studi Teknik Informatika, Fakultas Teknik Universitas Muhammadiyah Jember

maulanaindra2507@gmail.com

ABSTRAK

Perkembangan teknologi digital mendorong meningkatnya penggunaan platform daring sebagai sarana penyampaian opini dan ulasan wisata oleh masyarakat. Ulasan tersebut mengandung informasi penting yang dapat dimanfaatkan untuk mengetahui persepsi dan tingkat kepuasan pengunjung terhadap suatu destinasi wisata. Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap ulasan wisata Tanah Lot menggunakan metode *Support Vector Machine* (SVM) dengan ekstraksi fitur *Term Frequency–Inverse Document Frequency* (TF-IDF) serta membandingkan performa model berdasarkan pelabelan pakar dan pelabelan menggunakan *Artificial Intelligence* (AI). Data penelitian diperoleh melalui teknik web scraping terhadap situs TripAdvisor, dengan jumlah data sebanyak 500 ulasan wisata. Data yang diperoleh kemudian melalui tahapan *preprocessing* yang meliputi *cleaning*, *tokenizing*, *stopword removal*, dan *stemming*. Selanjutnya, data diekstraksi menggunakan metode TF-IDF dan diklasifikasikan ke dalam tiga kelas sentimen, yaitu negatif, netral, dan positif. Proses evaluasi model dilakukan menggunakan metode 10-*Fold Cross Validation* dengan parameter penilaian berupa akurasi, *precision*, *recall*, dan *f1-score*. Hasil penelitian menunjukkan bahwa model SVM dengan label pakar menghasilkan akurasi terbesar 0,76, sedangkan model SVM dengan label AI menghasilkan akurasi terkecil 0,58. Model cenderung memiliki performa terbaik pada kelas sentimen netral, sementara kelas negatif dan positif masih menunjukkan tingkat kesalahan klasifikasi yang cukup tinggi. Perbedaan hasil tersebut menunjukkan bahwa kualitas pelabelan data berpengaruh signifikan terhadap performa model klasifikasi sentimen. Berdasarkan hasil yang diperoleh, dapat disimpulkan bahwa metode *Support Vector Machine* (SVM) dengan ekstraksi fitur TF-IDF mampu digunakan untuk analisis sentimen ulasan wisata Tanah Lot. Namun, peningkatan performa model masih dapat dilakukan melalui penyeimbangan data, optimalisasi parameter, serta perbaikan proses pelabelan data.

Kata kunci: *Support Vector Machine, Web Scrapper, Confusion Matrix*

Sentiment Analysis of Tourism in Bali Province Using the Support Vector Machine (SVM) Method (Case Study: Tanah Lot)

Maulana Indra Wicaksono Harjanto[1], Henny Wahyu Sulistyo[2], Dudi Irawan[3]

Informatic Enginnering Study Program, Faculty of Enginnering

University of Muhammadiyah Jember

maulanaindra2507@gmail.com

ABSTRACT

The rapid development of digital technology has led to an increasing use of online platforms as a medium for sharing opinions and reviews related to tourism. These reviews contain valuable information that can be utilized to understand visitors' perceptions and satisfaction levels toward tourist destinations. This study aims to conduct sentiment analysis on tourism reviews of Tanah Lot using the Support Vector Machine (SVM) method with Term Frequency–Inverse Document Frequency (TF-IDF) feature extraction, as well as to compare model performance based on expert labeling and Artificial Intelligence (AI) labeling. The research data were obtained through a web scraping technique from the TripAdvisor website, resulting in a total of 500 tourism reviews. The collected data underwent preprocessing stages including cleaning, tokenization, stopword removal, and stemming. The data were then transformed using the TF-IDF method and classified into three sentiment classes: negative, neutral, and positive. Model evaluation was carried out using 10-Fold Cross Validation, with performance metrics including accuracy, precision, recall, and f1-score. The results show that the SVM model using expert-labeled data an average accuracy of 0.76, while the SVM model using AI-labeled data achieved an small accuracy of 0.58. The model demonstrated better performance in classifying neutral sentiment, whereas the negative and positive sentiment classes still exhibited relatively high misclassification rates. These findings indicate that data labeling quality has a significant impact on sentiment classification performance. Based on the results, it can be concluded that the Support Vector Machine (SVM) method combined with TF-IDF feature extraction is capable of performing sentiment analysis on Tanah Lot tourism reviews. However, further improvements can be achieved through data balancing techniques, parameter optimization, and enhancements in the data labeling process.

Keywords: *Support Vector Machine, Web Scraper, Confusion Matrix*