Web-Based Sentiment Analysis Application of Hotel Reviews in Indonesia

Rizki Ramadhan[1], Putu Harry Gunawan[2]

[1,2]Fakultas Informatika, Universitas Telkom, Bandung
[1]rizkyy@student.telkomuniversity.ac.id, [2]phgunawan@telkomuniversity.ac.id

Abstract

For many nations, including Indonesia, tourism is crucial. Following the oil and gas sector, tourism is Indonesia's second-largest source of foreign currency earnings. A comfortable operation of tourism activities depends on the presence of amenities like hotels and tourist attractions. To ensure that tourism operates smoothly, it is also crucial to understand the meaning behind customer feedback to maintain good facilities and services and improve inadequate or poor facilities.

The data used in this work is based on reviews of 150 hotels in the "Best Value Hotels" category in Indonesia from the Tripadvisor website. The reviews contained in the data are reviews in the form of positive and negative. The data is then processed for machine learning model training to classify sentiments as positive or negative reviews. We then applied feature extraction methods such as TF-IDF and SMOTE sampling method and compared the three models' performance with and without using them. This work also implements the trained model into web-based application, so the user can perform sentiment analysis on hotel reviews independently. The result is the overall performance of the three models in performing sentiment analysis on hotel reviews are great. SMOTE and TF-IDF also play a vital role in improving the three models' with an overall F1 Score above 80%.

Keywords: sentiment analysis, machine learning, hotel reviews