

ABSTRACT

The hotel is one of the most important tourism products, usually tourists look for hotels by considering both in terms of services, facilities, and the distance of the tour to be visited. Because currently technology is increasingly developing, tourists can book hotels through the website. Therefore many hotel reviews are found on the internet. With the availability of hotel reviews on the internet with a large number, so tourists can not understand all the reviews that are read whether they contain positive or negative opinions. Sentiment analysis is needed to overcome this problem. This study provides a solution by classifying positive and negative reviews using the support vector machine method by comparing the chi-square and TF-IDF feature selection, and comparing the use of stemming and without stemming. The best experimental results obtained are using the TF-IDF feature selection with 10-fold cross validation producing an accuracy value of 91.43%. The process of using without stemming on Chi-Square with 10-fold cross validation produces an accuracy value of 89.38%, while the use without stemming on TF-IDF with 10-fold cross validation produces an accuracy value of 88.5%.

Keywords: Hotel, Review, Sentiment Analysis, Support Vector Machine, Chi-Square, TF-IDF