Abstract

Film is an art that combines literature, acting, music, and visual arts. In general, films can be judged in various aspects, which will make the film liked. For example in the presentation of stories, visual effects, and so on. Assessment in these aspects can be described in a review or opinion. Currently, public opinion is an important thing in making a decision, for example in movie reviews. Sentiment analysis is a process that functions to classify the content of the dataset which makes the text positive and negative. This research will use the TF-IDF technique to perform feature extraction in the review, the Chi Square technique as a feature selection method, and SVM as a classification method. In this study, there were 5529 reviews on the http://movfreak.blogspot.com/ site consisting of 2884 positive reviews and 2645 negative reviews. The results of the test of the scenario without using the chi square feature selection resulted in an accuracy value of 85.71%, while the use of the chi square feature selection resulted in an accuracy value of 86.13%. The results using chi square feature selection get a higher accuracy value with a difference of 0.42%, because it can choose features optimally. So it can be concluded that this study can get better accuracy results by using the Support Vector Machine scenario by applying the Chi-square feature selection.

Keywords: Classification, Sentiment Analysis, Movie Review, TF-IDF, Support Vector Machine, Chi Square .