

ABSTRACT

A social media named Twitter which is currently widely used to explore data that aims to research, from people who talked about film in twitter can make a trend for prediction. Through a prediction is expected to assist a production house strategize so that the film aired sold in the market. In previous studies have been conducted a study using hollywood movie data, this time using the data of Indonesian films for research. These features are used that are from two sides, namely in terms of users and tweets, from the user side to use features number of followers official, follower production house, follower producer, follower director and follower 3 artist protagonist, of the tweet-use features number of retweets, like, the number of tweets containing the URL and not, as well as the number of tweets with sentiment scores of positive and negative. The purpose of this study is to apply a pre-processing techniques to make an appropriate input data for Support Vector Machine method, and with method of Support Vector Machine obtained an accuracy of 76%. Also conducted the comparison method between the Support Vector Machine to Simple Logistic Regression, and Logistic method, which obtained an accuracy value by 54% and 44%.

Keywords: *Twitter, Crawling Data, Pre-processing, Support Vector Machine*