ABSTRACT

Quality of local movies in Indonesia is improving years by years, this encourage people to watch the movies. There is also many contributing factors behind the success of a movie. Quality of the director and the ability of the actors to play a character can be an attraction for a movie on the market. As another important thing to note is how often a film discussed in social media. More people discuss it and it will make people want to know more about the movie. One of the social media that can be used as a benchmark if a movie is being discussed on social media is Twitter. By using Twitter a film can be investigated whether a film will be sold in the market or not. In this study writer use dataset which is taken from Twitter by analyzing of the user and the tweet. On the user, writer analyzing features follower official, follower production house, follower producer, follower director, and follower of the three main actor of the movie, on the tweet features writer analyzing the tweet, retweet, likes, the number of tweets which contains the URL, and the number of tweets contain positive and negative sentiment. The purpose of this study is to apply the pre-processing techniques to form a better input data into the Random Forest algorithm. The test results using the Random Forest algorithm obtained an accuracy of 88%.

Keywords : *Twitter, Pre-processing, Random Forest*