Abstract

Social media has grown exponentially since 2004 as a forum for discussion and exchanging opinions. This development received good enthusiasm by Indonesian citizens. Where out of 170 million people in Indonesia are active users of social media. However, the purpose of making social media is often misused to spread toxic comments, such as spreading hate, pornography, radicalism, SARA, and many more. So, lately, sentiment analysis of toxic comments on social media is being carried out. Therefore, in this study, sentiment analysis will be carried out on toxic comments on social media to sort out toxic or non-toxic comments based on labels or *multilabels*. In this research, Word2Vec method is used as feature extraction. Word2Vec as feature extraction has been widely applied in *Natural Language Processing* (NLP) research and shows a high potential impact on sentiment analysis performance. Then the classification is carried out using the method *Support Vector Machine* (SVM)to find the best results. The optimum results obtained from several tests carried out showed the highest F1-Score value of 73.69% of the data classified correctly using Word2Vec as feature extraction and without using *stemming* at the *preprocessing*.

Keywords: **Toxic Comment, Sentiment Analysis, *Multilabel*, Word2Vec, *Support Vector Machine* (SVM).**