

Abstract

In recent years, social media users have increased significantly. In January 2022 social media users in Indonesia reached 191 million people, an increase of 12.35% from the previous year. With this significant increase, more and more people tend to seek information through social media. Despite the many advantages provided by social media, the quality of information on social media is lower than traditional news media because a lot of hoax information is spread. With the many disadvantages felt by hoax information, it has led to many studies to detect hoax information on social media, especially information that is widely spread on Twitter. The purpose of this study is to compare the LSTM and IndoBERT methods in detecting hoaxes using datasets taken from Twitter. In this study, experiments were conducted to find the best parameters of the LSTM method and experiments using the IndoBERT method. The IndoBERT method shows good performance with an average accuracy value of 91.62%, and the LSTM model provides an average accuracy value of 87.61% using a combination of Word2Vec dimension value parameters of 200 and LSTM units of 128. In this study IndoBERT is shown to provide better accuracy than LSTM in hoax detection.

Keywords: hoax detection, social media, twitter, IndoBERT, LSTM