ABSTRACT

Accessing information through social media in the digital era has made it easier for people to obtain and exchange information. However, this has also led to an increase in the spread of fake news (hoaxes), which can mislead or manipulate public opinion. Handling fake news through media literacy and education has been implemented, but this requires substantial time and resources, while manual detection by experts is costly and time-consuming. Therefore, the application of machine learning is chosen to solve this problem quickly and efficiently. Previous studies have used classical machine learning methods, but these methods have limitations in word representation and context understanding. Deep learning models such as Convolutional Neural Network (CNN) and Long Short Term Network (LSTM) have shown better results, especially when using BERT as the text representation. This research aims to solve the problem of classifying Indonesian fake news using the BERT-CNN model. The dataset used consists of 14,794 scraped data from detik.com and turnbackhoax.id websites. The BERT-CNN model was trained and tested with four scenarios to determine the best data processing methods and parameters. The model achieved optimal performance using oversampling dataset, a 70/30 split for training and testing data, and trained with a learning rate of 0.00001 and dropout rate of 0.5. Model achieved an accuracy of 0.9798, precision of 1, recall of 0.9798, and an f1-score of 0.9897 for predicting fake news and an accuracy of 1, precision of 0.8746, recall of 1, and an f1-score of 0.9331 for predicting real news.

Keywords : Indonesian fake news classification, text representation, Bidirectional Encoder Representations from Transformers, Convolutional Neural Network.