ABSTRACT

Cyberbullying is often defined as any behavior carried out by an individual or group via social networks that includes negative, intimidating, or demeaning messages, causing discomfort to others. Cyberbullying has a harmful impact on the psychological well-being of its victims. The context-sensitive nature of tweets presents challenges in interpreting message content, particularly in languages such as Indonesian, which often exhibit substantial vocabulary variations. To address these challenges, this research employs feature expansion with FastText, utilizing a corpus generated from the IndoNews dataset containing 127,580 entries to enhance vocabulary comprehension in Indonesian-language tweets. Additionally, the study employs a Hybrid Deep Learning methodology for text classification, combining Convolutional Neural Networks (CNN) with Bidirectional Gated Recurrent Units (BiGRU). The CNN-BiGRU hybrid combines CNN's pattern recognition with BiGRU's context understanding for improved text classification. This research utilizes a dataset of 30,084 tweets retrieved from the X platform, limited to the scope of the Indonesian language. An analysis conducted on 30,085 datasets indicates that the implementation of FastText-enhanced feature expansion within the optimized hybrid deep learning model BiGRU-CNN achieved the highest accuracy of 81.72%, representing an improvement of +2.19% over the baseline CNN model and +2.09% over the baseline BiGRU model. As a result, this study has successfully detected cyberbullying on Twitter, supporting the creation of a safer and more constructive social media environment for users.

Keywords: BiGRU, CNN, *cyberbullying,* FastText, *hybrid deep learning, optimization*