

ABSTRACT

Cyberbullying on social media platforms has become widespread in society. Cyberbullying can take many forms, including hate speech, trolling, adult content, racism, harassment, or rants. One social media platform that has many cyberbullies is Twitter, which has been renamed 'X'. The anonymous nature of this 'X' platform allows users from all over the world to commit cyberbullying as they can freely share their thoughts and expressions without having to account for their identity. The purpose of this study is to see how the semantic influence of IndoBERT affects the hybrid deep learning model used to detect cyberbullying on platform 'X'. This study has 30,084 tweets with a hybrid deep learning approach that combines CNN and LSTM. In the IndoBERT scenario, IndoBERT features were first combined with TF-IDF, then expanded using FastText before being applied to the hybrid deep learning model. The test results produced the accuracy rate by: CNN (79.41%), LSTM (78.77%), CNN-LSTM (81.18%), CNN-LSTM-IndoBERT (82.05%). The results of this study show that the hybrid-IndoBERT model improves accuracy by 2.64% over the initial CNN baseline and 3.28% over the initial LSTM baseline. This research has successfully identified cyberbullying on platform 'X', which is an important step in the effort to build a safer environment on platform 'X'. The main contribution of this research is to explore the influence of IndoBERT semantic features on hybrid deep learning models for cyberbullying detection and integrate TF-IDF feature extraction and FastText feature expansion to improve text classification performance.

Keywords: cyberbullying; hybrid deep learning ; indoBERT; TF-IDF; fasttext