Ekspansi Fitur dengan *GloVe* untuk Deteksi Ujaran Kebencian Menggunakan Metode *Convolutional Neural Network (CNN)* dan *Long Short-term Memory (LSTM)* di Twitter

Muh. Fachrul Hidayat[1], Erwin Budi Setiawan[2]

[1,2]Fakultas Informatika, Universitas Telkom, Bandung
[1]fahrulhidayat@students.telkomuniversity.ac.id, [2]erwinbudisetiawan@telkomuniversity.ac.id

Abstract

Hate speech is unwanted behavior that aims to attack individuals or groups. The spread of hate speech can be found on social media, one of which is on the Twitter platform. This problem can be solved by classifying hate speech. This research proposes to build a hybrid model of hate speech detection using three deep learning models CNN (Convolutional Neural Network), LSTM (Long-Short Term Memory), and a combination of both models. The dataset used comes from Twitter in the form of Indonesian tweets. TF-IDF (Term Frequency - Inverse Document Frequency) and GloVe (Global Vector) are used as feature extraction and feature expansion to improve the accuracy of each model. Several scenarios were tested to find the best features. The best results in this study used a dataset with a ratio of 90% train and 10% test after pre-processing, TF-IDF with Unigram + Bigram + Trigram weighting, and 10,000 feature vectors. The method with Top-10 feature expansion in the tweet corpus achieved the highest accuracy of 90.83% for the CNN-LSTM hybrid model and improved by 0.79% against the previously determined baseline followed by the LSTM-CNN model with a result of 91.72% which improved by 1.88% against the baseline.

Keywords: hate speech, CNN, LSTM, GloVe