Abstract

Hate speech is a crime that is used as a means of provocation to a group, where a group is divided into several groups such as race, skin color, gender, disability, and nationality. Hate speech that occurs is usually in the form of sentences and images and is disseminated via the internet. Twitter social media is a social network that accommodates a lot of public opinion about anything that can be disseminated quickly received by other Twitter users. In previous research, the accuracy obtained was 79% using the Convolutional Neural Network method. Based on this research, in this study a system was developed to identify hate speech and classified using the Convolutional Neural Network method by comparing several tuning hyperparameters and producing the best tuning hyperparameters for the CNN model, namely dropout 0.3 and learning rate 0.001 which resulted in a CNN model accuracy value of 81%.

Keywords: Hate Speech, Convolutional Neural Network (CNN), Hyperparameter tuning