

INTRODUCTION

Social media is a medium that allows users to represent themselves and interact in social media, collaborate, share, and communicate with other users, connecting social virtually (Septian, Fachrudin and Nugroho, 2019). One of the popular social media platforms in Indonesia is Twitter. Twitter itself is one of the most widely used platforms, with around 19.5 million users out of 245 million users in Indonesia (Kominfo, 2022). Twitter is a social media platform that has the concept of disseminating information in a concise, direct, and real-time manner, with messages limited to less than 280 characters. Twitter serves as a medium for readers around the world and can be used as a means of spreading information to everyone (Fadli and Hidayatullah, 2021). However, many netizens misuse to use this platform to spread hate speech against certain ethnic groups or communities. According to Law No. 11 of 2008 in Indonesia, there are certain restrictions on how to communicate through social media (Eka Sembodo, Budi Setiawan and Abdurahman Baizal, 2016). Therefore, research about hate speech in advanced is needed to detect hate speech problem on Twitter.

Research on Twitter in Indonesian language has been conducted, in studies (Eka Sembodo, Budi Setiawan and Abdurahman Baizal, 2016; Isnain, Sihabuddin and Suyanto, 2020) data crawling was used to obtain some datasets of tweets in the Indonesian language. Based on previous research, researchers using data crawling to obtain datasets from Twitter in Indonesian language. Hybrid Deep Learning was introducing as a method to detection hate speech in general. In this study (Dewi and Ciptayani, 2022), Hybrid Deep Learning was performed by combining two or more methods to achieve better accuracy for detection hate speech. In another study (Salur and Aydin, 2020) used Hybrid Deep Learning with a CNN+BiLSTM approach, achieving the best accuracy of 82.14% compared to using standard Deep Learning methods, where Convolutional Neural Networks (CNN) had an accuracy of only 78.06%, and Bidirectional Long Short Term Memory (Bi-LSTM) had an accuracy of 80.44%. This study proves that using Hybrid Deep Learning can enhance the accuracy of a method.

The Glove method was using for feature expansion technique at reducing problem about discrepancies of words on Twitter in the Indonesian language. This is necessary because tweets on Twitter especially in Indonesian are often difficult to understand without context. Therefore, a feature is needed to facilitate this understanding. In study (Lim, Setiawan and Santoso, 2019), Glove outperformed Word2Vec, achieving an average F1 accuracy of 0.476 compared to Word2Vec's F1 score of 0.470. this researcher chose to use Glove for feature expansion due to its limited use in combination with Hybrid Deep Learning.

Study (Carracedo and Mondéjar, 2021), inform hate speech detection using Bidirectional Long Short Term Memory achieved a better accuracy of 79% compared to using SVM, which only reached 75%. However, that research only focused on the CNN method. Therefore, the researcher intends to utilize both CNN and Bi-LSTM methods, as CNN shown the best accuracy among all methods. Meanwhile, the Bi-LSTM method has not been previously applied in the Indonesian language. This research is expected to improve the performance of feature expansion in detecting hate speech.

From this method (Jamjoom *et al.*, 2024), it can be observed that GloVe has better accuracy than Word2Vec when using Bi-LSTM. This is evident as Word2Vec only achieved an accuracy of 82%, while GloVe saw an increase in accuracy by about 4%, reaching 86%. It can also be seen that the CNN method, when combined with GloVe, has better accuracy compared to using CNN alone. CNN with Word2Vec achieved 77% accuracy, whereas with GloVe, it increased to 79%. This also illustrates the potential of GloVe to outperform Word2Vec, suggesting that using both methods together could yield better accuracy than using just one method alone. This also provides initial evidence that GloVe has better accuracy than the existing Word2Vec, that information also reinforcing the theory of using GloVe as the preferred method.