Abstract

This research evaluates the performance of Gated Recurrent Unit-Support Vector Machine (GRU-SVM) model in sentiment analysis on Indonesian tweets. Generally, sentiment analysis research only performs binary classification based on the polarity value generated. However, binary sentiment analysis has its own limitations, such as in classifying sentences that contain both negative and positive expressions. Finegrained sentiment analysis provides more precise results by associating values to more than two target classifications. The GRU-SVM model is proposed as a deep learning method that can overcome the limitations of traditional machine learning models in terms of efficiency and performance. GRU is used because it can adaptively control dependencies and is more efficient in memory usage, while SVM is used as it is a state-of-the-art method in sentiment analysis on social media. The results show that adding feature extraction techniques can improve the accuracy of the model. N-gram extraction technique improves accuracy by 1.95%, GloVe improves accuracy by 2.10%. Dataset augmentation can improve model performance by 25% from the initial accuracy. The appropriate selection of training and testing data ratio can affect model performance, the ratio of 90:10 provides the highest accuracy score of 94.97%. The highest accuracy is achieved at 96.02% which surpasses several processes in sentiment analysis. Overall, the results of this research show that technique of vector representation, feature extraction, used dataset, and training and testing data ratio are essential in model testing.

Keywords: sentiment analysis, social media, gru, svm, granularity.