Abstract

The dissemination of information through social media has been rampant, especially on the Twitter platform. This information eventually invites various opinions from users as their points of view on a topic being discussed. These opinions can be collected and processed using sentiment analysis to assess public tendencies to obtain a fundamental source of decision-making. However, the procedure is not optimal enough due to its inability to recognize the word meaning of the opinion sentences. By using sentence-level granularity-oriented sentiment analysis, the system can explore the "sense of the word" in each sentence by giving it a granularity weight as the system's consideration in recognizing word meaning. To construct the procedure, this research utilizes LSTM as the classification model combined with TF-IDF and IndoBERTweet as feature extraction. Not only that, but this research also conducts the Word2Vec feature expansion method which was built using Twitter and IndoNews corpus to produce word similarity corpus and find effective word semantics. To be fully compliant with the granularity requirements, manual labeling, and system labeling were performed by considering weight granularity as a model performance comparison. This research succeeded in getting 88.97% accuracy for manual labeling data and 97.80% for system labeling data after combining these methods. The experimental results show that the granularity-oriented sentiment analysis model can outperform the conventional sentiment analysis system which can be seen based on the high performance of the resulting system.

Keywords: Sentiment Analysis sentence-level granularity oriented, LSTM, TF-IDF, IndoBERTweet, Word2Vec