

Abstract

After being hit by COVID-19 for a long time around the world which resulted in the paralysis of all countries, especially the economic aspects of all countries that dropped dramatically, the world was again shocked by the conflict between Russia and Ukraine which resulted in an increase in world oil prices including in Indonesia, many people complained and opposed the government's policy of increasing fuel prices because fuel affects various aspects, including economic aspects. Based on these problems, researchers use sentiment analysis methods that aim to find out people's opinions on issues that are being discussed throughout Indonesia and this research focuses on comparing the SVM algorithm with TF-IDF feature extraction then using K-Fold Cross Validation after that it is compared with the Lexicon Inset dictionary, in this case the model with Lexicon Inset which contains weighting on each word. In this study, it was found that the dataset model using the SVM algorithm with TF-IDF feature extraction and then using K-Fold Cross Validation obtained an average accuracy of 0.85 using the SVM algorithm. While the model using the automatic labeling dataset using the Indonesian sentiment Lexicon (Lexicon Inset) obtained an average accuracy of 0.68. Classification using SVM with TF-IDF feature extraction is superior to using Lexicon Inset.

Keywords : Sentiment, K-Fold, Lexicon, SVM, TF-IDF