

Analisis Sentimen Twitter: Penanganan Pandemi COVID-19 Menggunakan Metode Hybrid Naïve Bayes, Decision Tree, dan Support Vector Machine

Muyassar Akmal Iftikar ¹, Dr. Yuliant Sibaroni, S.Si., M.T ²

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹ akmaliftikar@student.telkomuniversity.ac.id, ² yuliant@telkomuniversity.ac.id.

Abstract

Indonesia is ranked sixth as the country with the most Twitter social media users. In December 2019, a virus called COVID-19 emerged and spread to almost all over world. A big number of Twitter users in Indonesia use Twitter to serve as a forum for sharing information as well as a means to express opinions about the Indonesian government's efforts in dealing with COVID-19. The author intends to do sentiment analysis of Twitter users regarding the handling of COVID-19 in Indonesia and compare the accuracy results of the hybrid model (Stacking Ensemble) and other individual models. Machine learning classification models have advantages and disadvantages of their own, and each model has different characteristics in carrying out the classification process. Therefore, one single machine learning model cannot be used as a definitive model in completing the classification process. The author uses the Stacking Ensemble as a hybrid classification model. Stacking Ensemble works by combining prediction results from other classification models. Then these results will be combined with a meta-classifier (Logistic Regression) with the aim of getting a final prediction result that is more accurate than the results of a single model classification. From this study, it was found that the general reaction of Twitter users in Indonesia to the handling of COVID-19 in Indonesia was generally positive, with a positive sentiment percentage of 75.3% and 60.39%, respectively. So, it can be concluded that Indonesian citizens' response to the handling of COVID-19 in Indonesia is positive. In addition, it was found that the hybrid Stacking Ensemble method can increase the accuracy values produced by other individual classifiers, with a difference of 0.63% and 1.02%.

Keywords: sentiment analysis, COVID-19, hybrid, Stacking Ensemble
