Abstract

The role of the internet and the growth of information that is reported on social media makes the development and dissemination of news easier, as well as in accessing it. During the current Covid-19 pandemic, a lot of news is spreading so that the wider community is looking for or getting information about this virus. This news entitled Covid-19 contains a lot of unimportant information and even reports hoax information. This makes the international community, especially Indonesia nervous about the news circulating during the Covid-19 period. Therefore, the author makes a system model to classify news according to what happens in the field. The information spread on social media is so varied that many unimportant news even contain hoax information. News classification will be done with K-Nearest Neighbor (KNN). The existing news is divided into several classes based on the category, then the news is classified as text using the K-Nearest Neighbor (KNN) method and k-fold cross validation as validation of the model created. The classification process is carried out using a scheme using 80% train data and 20% test data and changing the parameter value of k on K-Nearest Neighbor with k = 3, k = 5, k = 7, k = 9, and on k-fold cross validation as much as k = 5 and k = 10. Confusion matrix is used for evaluation. Finally, from each model that was carried out by changing the value of k on the K-Nearest Neighbor, the best accuracy results were obtained with an F1-Score of 48% from the value of k = 5, the validation results of the k-fold cross validation k = 10 by 45%.

Keywords: Covid-19, K-Nearest Neighbor, Hoax, internet, classification.