ABSTRACT

Natural disasters are natural and unpredictable. The impact of natural disasters depends on their intensity, which can be in the form of landslides, floods, earthquakes, and even fatalities. Many things can be minimized before the disaster spreads. For this reason, it is important to have information about the occurrence of disasters.

Social media is a place that can connect one information to another, so that it can be spread widely. The development of social media today is very fast, especially with the news related to disasters around. One of the most widely used social media is Twitter. By using Twitter, the public can quickly disseminate disaster information through tweets so that the impact of the response can be accelerated. Therefore, it is important to know real-time information about the number of natural disasters so that anticipation can be done early on.

So from the explanation mentioned above, a system is needed that can sort out natural disaster data by itself on Tweets. The test results are made to display the mapping of disasters that occur in the territory of Indonesia which are classified based on the area that has tweet data in the form of a map visualization of which disaster area has the higher frequency and type of disaster. Regional classification is done using the K-Nearest Neighbor algorithm method. In this study. In this study, the results of the Confusion matrix test have the best accuracy using the Jaccard metric of 86% and for data sharing performance using k-fold cross validation 10 Fold, the best accuracy result is the Jaccard metric, which is 83% at fold 8.

Keywords: *twitter, natural disaster, classification, confusion matrix, k-fold cross validation*