ABSTRACT

Traffic congestion in various cities has become an endless problem and seems to have become a daily routine of people in the Capital Jakarta. The causes are very diverse, ranging from traffic accidents, road users attitude, limited public transport, and increased vehicle ownership per year increases, and the level of online riders who are ready to look for any orderan.

Information on traffic conditions is certainly needed by the riders in avoiding congestion. This information can be obtained easily through social networks such as twitter. However, the information shared on twitter is still text that has not been categorized. In addition, the information obtained has not been visualized into the form of maps. Twitter accounts that provide traffic information are Twitter @TMCPoldaMetro, @lewatmana, @radioelshinta, @ sonorafm92. The information provided by the account is not all posts about traffic conditions. This is where the pre-processing process is needed to generate the required traffic condition data.

In this study, a congestion classification system was established in DKI Jakarta with one of datamining technique, which is classification using classification method of decesion tree that is C4.5. This C4.5 method transforms a very large fact into a decision tree presenting the rules. Locations obtained will be plotted by geocoding and the classification process will be tested using a data partition with a confusion matrix. The results in this study show average accuracy rate 99.08%, precision 99.46%, and Recall 97.99%.

Keywords: Congestion, twitter, Algorithm C4.5