## *ABSTRACT*

*The system developed by the author is a system that can determine the best path if one of the roads experience vehicle density or loss. The roads specified in this system are very limited, just the road* Bundaran HI Jakarta.

*In this Final Project research, congestion data was taken from Twitter social media because of the many tweets on Twitter which stated the traffic situation in Jakarta. Before being classified, the data will go through a pre-processing process consisting of Case Folding, Cleaning, Tokenization, and Data Transformation. This system uses the K-Nearest Neighbor (KNN) algorithm to classify Twitter data. The amount of data used in this study is 600 data with data tested three times. In the first test the data is divided into 50% training data and 50% testing data, while in the second test, the data is divided by 67% training data and 33% testing data. Finally, the data is divided into 80% training data and 20% testing data. From these tests obtained the highest precision value in training data as much as 80% and testing data as much as 20% with the value of Accuracy = 84,16%, Precision = 96,00%, and Recall 84,00% with the number of neighbors (K) is 27.*



***Keywords:*** *Density, Data Mining, Pre-processing, Classification*