

ABSTRACT

SENTIMENT ANALYSIS ON TWITTER ABOUT MANDALIKA CIRCUIT USING NAIVE BAYES ALGORITHM

By

Muhammad Alvi Awliya

18102239

Technological developments make it easy for everyone to give their opinion on something or a trending discussion on social media. Twitter is one of the social media that is the destination for opinion escape. Discussion of the Mandalika circuit is an interesting topic discussed by the public on Twitter social media. The great attention pinned by the public and all Indonesian people is not without cause because Mandalika itself is an international circuit that is becoming a discourse in Indonesia. In this study, sentiment analysis was carried out by taking data from Twitter regarding the Mandalika circuit and classifying it using the Naive Bayes algorithm, and comparing labeling using three sentiments and two sentiments. The data taken was in the form of Indonesian language data on social media Twitter with a total of 1148 data. Data processing is done through a preprocessing process, namely cleaning, case folding, removing duplicates, tokenizing, stopword removing, and stemming. Weighing words using the TF-IDF method and creating models using the Naive Bayes algorithm. Validation uses 10-fold Cross-validation and evaluation with Confusion Matrix. The test results of sentiment analysis by labeling the three sentiment categories show an accuracy of 59%, 50% precision, 50% recall, and 50% f1-score. In addition, the results for the two sentiments get 83% accuracy, 56% precision, 58% recall, and 56% f1-score. Based on the results of the comparison between the use of three labeling sentiments and two sentiments, it can be seen through the accuracy results that the use of two sentiments is better than three sentiments.

Keywords: *Mandalika Circuit, Naive Bayes, Sentiment Analysis, Twitter*