

ABSTRACT

Social media Twitter is one of the popular social media and contains a lot of information. Every tweet or tweet written can provide information about a person's personality. The problem is how to classify existing text on Twitter social media into classes that will be created with good performance values. In this final assignment research, the author built a system to classify one's personality on social media twitter using the classification method Adaptive Boosting (AdaBoost) using the TF-IDF and WIDF weighting methods. Another thing that distinguishes this research is the weighting of each word using the weighting method TF-IDF and WIDF by adding new features to the approach based on the user's social behavior such as the number of characters in the tweet, the average character of each tweet, the average word on the tweet, media URL that reads how many users upload photos or videos, punctuation marks that are counting the number of question marks (?) and exclamation marks (!) on every tweet, capitalizing letters, and emojis on tweets that can be detected as many as 2,552 different characters, and the last is to try to integrate a personality approach based on social behavior with a linguistic approach. From the results of the experiment on the training data ratio of 70% and the test data 30% (70:30), the results of accuracy were 62.79% based on social behavior and based on the linguistic approach using TF-IDF.

Keywords: Personality, Twitter, Big 5 Personality, AdaBoost, TF-IDF and WIDF.