Abstract

A person's personality type can generally be known by writing content on social media. A web-based application that performs the classification of a person's personality type based on writing on Twitter. It is hoped that it can be another way to find out a person's personality type. In this study, the author performs a person's personality classification system using the Naive Bayes–Support Vector Machine classification method. The Personality Model used is the Big Five Personality. Testing the data is based on the content of Twitter users' tweets. The system was built by collecting data from Twitter and labeling them, namely Openness, Consciousness, Extraversion, Agreeableness, Neuroticsm. The dataset used is obtained from the Twitter API. Experiments were carried out with scenarios to get the data ratio of the accuracy of social behavior as a baseline, adding TF-IDF and LIWC data and using the SMOTE method which was tested using a hyperparameter tuning technique using Grid Search with social behavior as the baseline. Accuracy results obtained by adding TF-IDF and LIWC data and applying the SMOTE method can increase the accuracy value by 82.41% on the Support Vector Machine algorithm. Meanwhile, Naive Bayes has increased accuracy by 28.32% with an accuracy of 71.42%. And the Naive Bayes-Support Vector Machine algorithm has an accuracy increase of 17.64% with an accuracy value of 53.86%.

Keywords: Personality, NBSVM, Big Five Personality, Twitter