Abstract

Personality is a psychological construction that leads to variations in individual thought patterns, distinguishing human characteristics based on traits, feelings, and behavior. Hence, human beings have personalities and tend to be social creatures. Social media was created to facilitate people for remote communication. Twitter is one of the largest social media that is often used by the world community. Because the activities of social media users are quite numerous and varied, we are interested to predict the behavior of Twitter users based on their personality. These personalities are classified into five categories called Big Five Personality traits. By using Artificial Neural Network (ANN), it is one of the classification methods used as a reference in personality prediction, this method has good performance in classification. To overcome the overfitting while training the data, we used Synthetic Minority Oversampling Technique (SMOTE) technique to handle the imbalance data. Other features such as Linguistic Inquiry Word Count (LIWC) and Robustly Optimized BERT pre-training Approach (RoBERTa) with hyperparameter tuning can increase the performance of the system that we have built. The focus of this research is to present an analysis of Twitter user behavior that it can be used for the Big Five Personality Prediction from Twitter users using the ANN method. The use of this method has aspect to be considered, that is the accuracy of the classification of Openness, Conscientiousness, Extra-version, Agreeableness, and Neuroticism (OCEAN) traits. The experimental results have shown that the ANN method gets 87.09% of accuracy.

Keywords: Artificial Neural Network, OCEAN, Big Five Personality, LIWC, RoBERTa, Twitter