

**DETEKSI KEPRIBADIAN BIG FIVE PENGGUNA TWITTER
DENGAN METODE C4.5****Shantika Valerin Therik¹, Erwin Budi Setiawan²**^{1,2}Fakultas Informatika, Universitas Telkom, Bandung¹shantikavt@students.telkomuniversity.ac.id, ²erwinbudisetiawan@telkomuniversity.ac.id

Abstract

Twitter is a social media that provides a feature for its users to read and write messages known as "tweets". To be able to understand a person's personality, posts on social media, namely Twitter, can be used as a source of information. The Big Five in psychology is one method for interpreting a person's personality. Appropriate personality has a direct impact on performance at work. Personality fit will see how a person's personality fits the job and company culture. In this study, the C4.5 method was used to create a personality classification model for Twitter users which consisted of five classes, namely Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. Dataset obtained using the Twitter API. An experiment was conducted with the scenario of getting the data ratio from the accuracy of social behavior as a baseline, adding TF-IDF and LIWC data and applying the SMOTE method which was tested using a hyperparameter tuning technique using Grid Search with social behavior as the baseline. The accuracy results obtained by adding TF-IDF and LIWC data are 62.06% with an increase in accuracy of 17.24% from the baseline and using the SMOTE method can increase the accuracy value to 76.92% with an increase in the accuracy value of 32.1% from the baseline. From the features with the best accuracy values in the experiment, the Big Five personality detection decision tree model was generated.

Keywords : *Big Five, C4.5 Classification, TF-IDF, LIWC*