

Comparative Analysis of Personality Detection using Random Forest and Multinomial Naive Bayes

Azka Zainur Azifa^{1*}, Warih Maharani², Prati Hutari Gani³

^{1,2,3} School of Computing, Informatics Study Program, Telkom University, Bandung, Indonesia

Email: ^{1*}azkazainur@students.telkomuniversity.ac.id, ²wmaharani@telkomuniversity.ac.id,

³pratihutarigani@telkomuniversity.ac.id

Correspondence Author Email: azkazainur@students.telkomuniversity.ac.id

Abstract— Personality is a difference that is owned by each individual in thinking, feeling, and behaving. Personality is an individual characteristic that is formed based on biological parents and environmental influences. Personality type is one of the determinants of the type of work performed. The Big Five personality is a method used to detect personality. This theory divides characteristics into five dimensions, namely Openness, Conscientiousness, Extraversion, Neuroticism, and Agreeableness. Several studies have shown that personality identification can be done through social media, one of which is by using Twitter. Much research related to personality detection has been carried out using machine learning, but only focuses on one machine learning model. In the case of text detection, multinomial naive bayes have a more stable performance than random forest, while random forest has better accuracy than multinomial naive bayes. therefore this study focuses on conducting a comparative analysis using random forest and multinomial naive Bayes. the best accuracy is produced by a system with a random forest model of 60.71% and a precision value of 62% for openness personality and 57% for agreeableness personality.

Keywords: Twitter; Personality Detection; Big Five Personality; Random Forest; Multinomial Naive Bayes