

ABSTRAK

Pada tahun 2020, virus Covid-19 menyebabkan pandemi yang membuat sebagian besar orang lebih aktif di media sosial, seperti Twitter. Twitter memiliki fitur tweet yang memungkinkan penggunanya mengirimkan pesan singkat tentang apa yang mereka rasakan dan pikirkan saat itu. Berdasarkan tweet seseorang, kita mengetahui pola pikirnya, dan memungkinkan kita mengetahui kepribadian orang tersebut. Salah satu model kepribadian adalah kepribadian Lima Besar. Lima Besar membagi kepribadian menjadi lima kelas: keterbukaan, kesadaran, ekstraversi, keramahan, dan neurotisme. Ada beberapa cara yang bisa dilakukan untuk mengetahui kepribadian, salah satunya yaitu mengikuti tes psikologi. Namun, hal itu bisa memakan waktu lama dan konsentrasi total. Oleh karena itu, penelitian ini melakukan deteksi kepribadian Big Five pada pengguna Twitter dengan menggunakan metode Gradient Boosted Decision Tree (GBDT). Penelitian ini bertujuan untuk memperoleh nilai akurasi yang tinggi dengan melakukan pembobotan melalui metode TF-IDF dan menggunakan fitur sentimen dan emosi. Penelitian ini menggunakan dataset Indonesia yang dikumpulkan melalui Twitter API. Penelitian ini terdiri dari dua pengujian skenario, pengujian skenario pertama dilakukan dengan dataset tidak seimbang dan pengujian skenario kedua dilakukan dengan menerapkan teknik oversampling dengan metode SMOTE untuk menangani dataset tidak seimbang. Dengan menerapkan metode SMOTE, penelitian ini memperoleh akurasi yang tinggi dengan nilai 60,36%.

Kata Kunci: Twitter; Tweet; Personalitas; Big Five; Gradient Boosted Decision Tree;