

Perkembangan dunia modern yang pesat telah meningkatkan penggunaan media sosial. Salah satu platform media sosial yang paling populer di Indonesia adalah X, di mana pengguna dapat berbagi kicauan untuk mengekspresikan diri. Aktivitas yang terjadi di X dapat mengungkapkan aspek kepribadian pengguna. Penelitian ini berfokus pada penggunaan algoritma Regresi Logistik, yang menggabungkan berbagai skenario preprocessing data, teknik penyeimbangan data, rasio pemisahan data, dan metode ekstraksi fitur, untuk mengklasifikasikan kepribadian pengguna X berdasarkan tweet mereka. Data untuk penelitian ini dikumpulkan dengan menggunakan kuesioner Big Five Personality yang didistribusikan secara luas dan diperkaya dengan merayapi tweet dari 257 responden pengguna X di Indonesia. Data diklasifikasikan ke dalam lima sifat kepribadian, yaitu openness, conscientiousness, extraversion, agreeableness, dan neuroticism. Teknik prapemrosesan data diterapkan dalam dua skenario untuk membuat dataset lebih cocok untuk membangun model. Ketidakseimbangan dataset diatasi dengan menerapkan dua teknik penyeimbangan data, yaitu Random Removal Augmentation dan Synthetic Minority Oversampling Technique (SMOTE). Hyperparameter tuning dilakukan dengan menggunakan grid search untuk mengidentifikasi parameter model yang optimal. Eksperimen dilakukan dengan beberapa kombinasi ekstraksi fitur seperti TF-IDF, TF-IDF dengan BOW, dan TF-IDf dengan N-gram. Akurasi tertinggi yang dicapai adalah 71%.

Kata kunci-Kepribadian Big Five, Regresi Logistik, Klasifikasi Kepribadian