

DETEKSI KEPRIBADIAN BIG FIVE PENGGUNA TWITTER DENGAN METODE C4.5

Shantika Valerin Therik¹, Erwin Budi Setiawan²

^{1,2}Fakultas Informatika, Universitas Telkom, Bandung

¹shantikavt@students.telkomuniversity.ac.id, ²erwinbudisetiawan@telkomuniversity.ac.id

Abstrak

Twitter merupakan media sosial yang memberikan fitur bagi penggunaannya untuk membaca dan menulis pesan yang disebut “*tweet*”. Untuk dapat memahami kepribadian seseorang, postingan di media sosial yaitu Twitter dapat digunakan sebagai sumber informasi. Big Five dalam dunia psikologi merupakan salah satu metode untuk menginterpretasi kepribadian seseorang. Kepribadian yang sesuai memiliki dampak secara langsung pada kinerja di tempat kerja. Personality fit akan melihat bagaimana kepribadian seseorang sesuai dengan pekerjaan dan budaya perusahaan. Pada penelitian ini, metode C4.5 digunakan untuk membuat model klasifikasi kepribadian pengguna Twitter yang terdiri dari lima kelas yaitu *Openness*, *Consciousness*, *Extraversion*, *Agreeableness*, *Neuroticism*. Dataset yang diperoleh menggunakan API Twitter. Dilakukan percobaan dengan skenario mendapat rasio data dari akurasi perilaku sosial sebagai baseline, penambahan data TF-IDF dan LIWC dan menerapkan metode SMOTE yang diujikan menerapkan teknik *hyperparameter tuning* menggunakan *Grid Search* dengan perilaku sosial sebagai *baseline*. Hasil akurasi yang diperoleh dengan penambahan data TF-IDF dan LIWC sebesar 62.06% dengan kenaikan akurasi sebesar 17.24% dari *baseline* dan menggunakan metode SMOTE dapat meningkatkan nilai akurasi menjadi 76.92% dengan kenaikan nilai akurasi sebesar 32.1% dari *baseline*. Dari fitur dengan nilai akurasi terbaik dalam percobaan, dihasilkan model pohon keputusan deteksi kepribadian Big Five.

Kata kunci : Big Five, Klasifikasi C4.5, TF-IDF, LIWC