

Prediksi Kepribadian Big Five Media Sosial Twitter Dengan Metode Naive Bayes-Support Vector Machine (NBSVM)

Alvinda Amalia Rizkita¹, Erwin Budi Setiawan²

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

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

Abstrak

Secara umum tipe kepribadian dapat diketahui dengan penulisan konten pada sosial media. Sebuah aplikasi berbasis web yang melakukan klasifikasi tipe kepribadian seseorang berdasarkan penulisan pada Twitter. Diharapkan dapat menjadi salah satu cara lain untuk mengetahui tipe kepribadian seseorang. Di penelitian ini, penulis melakukan sistem klasifikasi kepribadian seseorang dengan menggunakan metode klasifikasi *Naive Bayes-Support Vector Machine*. Model Kepribadian yang digunakan yaitu *Big Five Personality*. Pengujian data dilakukan berdasarkan konten *tweet* pengguna Twitter. Sistem dibangun dengan mengumpulkan data dari Twitter dan dilakukan pelabelan yaitu *Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism*. Dataset yang digunakan diperoleh dari API Twitter. Dilakukan percobaan dengan skenario untuk mendapatkan rasio data dari akurasi perilaku sosial sebagai *baseline*, penambahan data TF-IDF dan LIWC serta dilakukan 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 serta menerapkan metode SMOTE dapat menaikkan nilai akurasi sebesar 82.41% pada algoritma *Support Vector Machine*. Sedangkan *Naive Bayes* mengalami kenaikan akurasi sebesar 28.32% dengan akurasi sebesar 71.42%. Dan algoritma *Naive Bayes-Support Vector Machine* mengalami kenaikan akurasi 17.64% dengan nilai akurasi yaitu 53.86%.

Kata kunci : Kepribadian, NBSVM, Big Five Personality, Twitter