

Prediksi kepribadian *Big Five* dengan *Term-Frequency Inverse Document Frequency* Menggunakan Metode *k-Nearest Neighbor* pada *Twitter*

Roji Ellandi¹, Erwin Budi Setiawan S.Si., M.T², Dr. Fida Nirmala Nugraha, s.Psi.,M.Psi.³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹rojiellandi@students.telkomuniversity.ac.id, ²erwinbudisetiawan@telkomuniversity.ac.id,
³fidanurmalanugraha@telkomuniversity.ac.id

Abstrak

Twitter merupakan media sosial yang sampai saat ini sangat digemari dan menjadi penyebaran informasi yang sangat cepat. Informasi yang beredar juga sangat banyak, mulai dari berita, opini, komentar, dan kritik semuanya ada yang bersifat positif, negatif, dan netral. Menurut data yang dilansir secara statistik dan berdasarkan penelitian *PeerReach*, Indonesia termasuk pengguna *Twitter* yang paling aktif ke 3 di dunia dibawah Amerika Serikat dan Jepang. Dari kumpulan data tersebut kita dapat melakukan analisis kepribadian terhadap suatu keadaan untuk melihat respon masyarakat, media ataupun pemerintahan terhadap suatu objek tersebut dan proses klasifikasi itu sendiri. Metode yang digunakan dalam penelitian prediksi kepribadian ini dilakukan untuk mengklasifikasi sebuah *tweet* ke dalam bentuk 5 kepribadian. Metode kepribadian yang digunakan peneliti adalah *Big Five Personality* dan metode perhitungan klasifikasi dengan *k-NN (k-Nearest Neighbor)*. Hasil dari penelitian ini adalah dapat memperoleh nilai akurasi 60,97% dengan pembobotan melalui tahap *TF-IDF (Term-Frequency Invert Document Frequency)*.

Kata kunci : Klasifikasi, *Big Five Personality*, *Twitter*, *TF-IDF*, *k-NN*
