

Analisis Perbandingan Deteksi Kepribadian menggunakan Random Forest dan Multinomial Naïve Bayes

Azka Zainur Azifa¹, Warih Maharani², Prati Hutari Gani³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung
¹azkazainur@students.telkomuniversity.ac.id, ²wmaharani@telkomuniversity.ac.id,
³pratihutarigani@telkomuniversity.ac.id

Abstrak

Kepribadian merupakan sebuah perbedaan yang dimiliki oleh setiap individu dalam berpikir, berperasaan dan berperilaku. Kepribadian adalah karakteristik individu yang terbentuk berdasarkan biologis orang tua dan pengaruh lingkungan. Jenis kepribadian menjadi salah satu penentu jenis pekerjaan yang dilakukan. Kepribadian Big Five merupakan model yang digunakan dalam mendeteksi kepribadian. Teori ini membagi karakteristik menjadi lima dimensi berupa Openness, Conscientiousness, Extraversion, Neuroticism dan Agreeableness. Beberapa penelitian menunjukkan bahwa identifikasi kepribadian dapat dilakukan melalui media sosial, salah satunya dengan menggunakan Twitter. Penelitian terkait deteksi kepribadian sudah banyak dilakukan dengan menggunakan machine learning, namun hanya berfokus pada satu model machine learning. Pada kasus deteksi text, multinomial naive bayes memiliki performa yang lebih stabil dibandingkan random forest, sedang random forest memiliki akurasi yang lebih baik dibandingkan multinomial naive bayes. maka dari itu penelitian ini berfokus untuk melakukan analisis perbandingan dengan menggunakan random forest dan multinomial naive bayes. akurasi dan presisi terbaik dihasilkan oleh sistem dengan model random forest sebesar 60.71% dan nilai presisi sebesar 62% untuk kepribadian openness dan 57% untuk kepribadian agreeableness.

Kata kunci : Twitter, Big Five, Random Forest, Multinomial Naive Bayes
