

Abstrak

Twitter merupakan salah satu media sosial yang sangat populer digunakan yang dapat memfasilitasi pengguna untuk menuliskan pemikiran mereka kedalam sebuah *tweet*. Dari *tweet* yang bersifat personal tersebut dapat diambil informasi seperti jenis emosi, sentimen, dan kepribadian. Kepribadian penting untuk diketahui karena berguna dalam memahami potensi pengguna, meningkatkan akurasi untuk sistem rekomendasi, dan membantu diagnosis kesehatan mental. Saat ini sudah banyak penelitian sebelumnya yang memprediksi kepribadian seseorang menggunakan data berbahasa Inggris dengan menggunakan banyak fitur, namun untuk bahasa Indonesia belum diteliti dengan baik. Maka dalam penelitian ini dibangun sistem prediksi kepribadian yang diklasifikasikan kedalam model *Big Five*, karena dianggap model yang paling sesuai untuk menentukan kepribadian pengguna di media sosial. Fitur yang digunakan adalah fitur sosial, deteksi sentimen, dan deteksi emosi dengan menghitung bobot kata setiap *tweet* menggunakan *Term Frequency Inverse Document Frequency* (TF-IDF). Metode klasifikasi akan menggunakan *Naive Bayes* karena cepat dalam perhitungan data yang besar, dan implementasi yang cenderung lebih mudah untuk dibangun. Hasil eksperimen menunjukkan akurasi sebesar 45,00% untuk gabungan semua fitur. Rata-rata fitur emosi sendiri menghasilkan akurasi 38,12%, fitur sentimen menghasilkan akurasi 38,95%, dan fitur sosial dengan akurasi 23,15%.

Kata kunci : big five, deteksi emosi, analisis sentimen, tf-idf, naive bayes.