

Abstrak-Berkembangnya platform media sosial telah menyediakan data tekstual yang sangat besar yang dapat digunakan untuk prediksi kepribadian dengan menggunakan teknik pemrosesan bahasa alami (NLP). Penelitian ini menyajikan analisis perbandingan kinerja dan efisiensi komputasi antara dua model berbasis transformator, BERT dan XLNet, dalam memprediksi ciri-ciri kepribadian Big Five dari tweet berbahasa Indonesia yang dikumpulkan dari Twitter (sekarang dikenal sebagai X). Dataset yang terdiri dari 9.708 tweet yang dilabeli oleh psikolog, telah melalui proses preprocessing dan dievaluasi dengan menggunakan klasifikasi multi-label. Berbagai konfigurasi diuji, termasuk pengoptimal yang berbeda, laju pembelajaran, ukuran batch, epoch, dan tingkat putus sekolah. Hasilnya menunjukkan bahwa BERT mengungguli XLNet dalam hal skor F1 secara keseluruhan (43,362% vs 42,822%) dan waktu inferensi (215,72 ms vs 7714,83 ms), yang menunjukkan efisiensi dan stabilitas yang lebih tinggi. Meskipun XLNet menunjukkan hasil yang sedikit lebih baik pada sifat-sifat tertentu seperti Extraversion dan Openness, kinerjanya lebih sensitif terhadap hiperparameter dan dibatasi oleh waktu komputasi yang lebih lama. Penelitian ini menyoroti kesesuaian BERT untuk tugas klasifikasi teks bahasa Indonesia secara real-time dan menyarankan eksplorasi lebih lanjut mengenai potensi XLNet menggunakan dataset yang lebih besar dan konfigurasi yang dioptimalkan.

Kata kunci-kepribadian big five, prediksi kepribadian, BERT, XLNet, pemrosesan bahasa alami.