

ABSTRAK

Maraknya penggunaan Artificial Intelligence (AI) generatif dalam lingkup perguruan tinggi membawa ancaman atas penggunaan tidak bertanggung jawab yang mengancam integritas akademik serta kualitas luaran sumber daya manusia perguruan tinggi. Dalam penelitian ini ditujukan untuk mengembangkan sebuah model pendeteksi teks hasil AI sebagai upaya penanganan penggunaan AI generatif yang tak bertanggung jawab. Digunakan model IndoBERT yang dilatih melalui metode fine-tuning untuk mengklasifikasikan teks hasil AI. Digunakan data berupa teks yang dikumpulkan dari paper-paper jurnal lokal berbahasa Indonesia, hasil parafrase teks paper menggunakan model AI generatif, serta data yang digunakan dalam pengembangan-pengembangan terdahulu. Dihasilkan model pendeteksi teks hasil AI dengan akurasi di atas 97% yang dilatih menggunakan data yang dikumpulkan secara mandiri, serta model dengan akurasi di atas 99% dengan tambahan data dari pengembangan terdahulu. Model yang dihasilkan dapat digunakan sebagai alat kontrol karya-karya tulis ilmiah yang dihasilkan oleh perguruan tinggi maupun submisi lain berbentuk teks.

Kata kunci: *IndoBERT, BERT, pendeteksi, artificial intelligence*