

ABSTRAK

Pesatnya jangkauan internet di Indonesia telah mengubah pola komunikasi masyarakat, namun di sisi lain memicu peningkatan Kekerasan Berbasis Gender Online (KBGO), khususnya pada platform Instagram. IndoBERT digunakan sebagai metode untuk menangkap fitur semantik dan konteks dua arah (*bidirectional*) yang mendalam, sementara SVM diterapkan sebagai klasifikator pada tahap akhir untuk mengoptimalkan hasil *embedding* berdimensi tinggi. Tahapan penelitian meliputi akuisisi data komentar Instagram, *preprocessing*, ekstraksi fitur menggunakan IndoBERT, serta klasifikasi kelas KBGO dan Non-KBGO menggunakan SVM. Hasil evaluasi menunjukkan model mencapai nilai akurasi sebesar 0,79, *precision* 0,50, *recall* 0,67 dan *F-1 Score* 0,57. Berdasarkan *confusion matrix* terhadap 14 data uji, model berhasil mengklasifikasikan 11 data dengan benar. Meskipun performa model berada pada kategori cukup baik efektivitas deteksi masih dipengaruhi oleh ketidakseimbangan dataset antara kelas mayoritas (Non-KBGO) dan minoritas (KBGO). Penelitian ini membuktikan bahwa pendekatan *Deep Learning* berbasis konteks memiliki potensi besar dalam mengenali pola kekerasan diskriminatif di ruang digital.

Kata kunci: KBGO, Instagram, IndoBERT, Support Vector Machine, Deep Learning.