

ABSTRAK

Perdagangan manusia atau dikenal dengan *human trafficking* mengalami peningkatan signifikan pada Januari hingga Maret tahun 2025 dibandingkan periode sebelumnya. Perkembangan modus pendekatan baru ke ranah digital melalui iklan lowongan kerja palsu di media sosial semakin sulit dideteksi secara manual akibat volume postingan yang besar dan pola penipuan yang terus berkembang. Penelitian ini bertujuan mengembangkan sistem deteksi otomatis berbasis *fusion model* yang mengintegrasikan IndoBERT dan *meta-feature* terstruktur untuk mengklasifikasikan postingan lowongan kerja berisiko perdagangan manusia di Facebook. Sistem dibangun menggunakan pipeline ekstraksi teks berbasis EasyOCR, ekstraksi *meta-feature domain-specific*, dan model Fusion IndoBERT + *Meta-Feature*. Eksperimen dilakukan pada 324 sampel postingan dengan 27 konfigurasi yang mencakup variasi arsitektur, strategi *fine-tuning*, dan rasio pembagian data. Model terpilih menghasilkan akurasi 0,8980, *precision* 0,8333, *recall* 0,9524, *F1-score* 0,8889, dan AUC 0,9439, dengan penambahan *meta-feature* terbukti meningkatkan *recall* sebesar 4,76 poin persentase dengan performa yang lebih stabil di seluruh metrik dibandingkan model *baseline* IndoBERT. Evaluasi *cross-validation* mengonfirmasi kemampuan generalisasi sistem dengan *recall* rata-rata $0,8155 \pm 0,0641$ dan AUC $0,8826 \pm 0,0432$.

Kata kunci: *human trafficking*, deteksi risiko, IndoBERT, *meta-feature*, *fusion model*, klasifikasi teks

ABSTRACT

Human trafficking saw a significant increase from January to March 2025 compared to the previous period. The emergence of new methods in the digital realm, such as fake job postings on social media, has made manual detection increasingly difficult due to the large volume of posts and constantly evolving fraud patterns. This study aims to develop an automated detection system based on a fusion model that integrates IndoBERT and structured meta-features to classify job postings on Facebook that pose a risk of human trafficking. The system was built using a text extraction pipeline based on EasyOCR, domain-specific meta-feature extraction, and the IndoBERT + Meta-Feature fusion model. Experiments were conducted on 324 sample posts using 27 configurations that included variations in architecture, fine-tuning strategies, and data split ratios. The selected model achieved an accuracy of 0.8980, precision of 0.8333, recall of 0.9524, F1-score of 0.8889, and AUC of 0.9439; the addition of meta-features was shown to increase recall by 4.76 percentage points, with more stable performance across all metrics compared to the IndoBERT baseline model. Cross-validation evaluation confirmed the system's generalization ability with an average recall of 0.8155 ± 0.0641 and an AUC of 0.8826 ± 0.0432 .

Keywords: *human trafficking, risk detection, IndoBERT, meta-feature, fusion model, text classification*