

RINGKASAN

Rancang Bangun Model IndoBERT untuk Deteksi Ujaran Kebencian Studi Kasus Diskursus RKUHAP di Media Sosial X dan Threads, Firman Agung Alamsyah, NIM E41253341. Tahun 2026, Teknologi Informasi, Politeknik Negeri Jember, Ery Setiyawan Jullev Atmadji, S.Kom, M.Cs (Dosen Pembimbing).

Perkembangan media sosial seperti X dan Threads telah menjadi ruang utama pertukaran wacana publik di Indonesia. Perdebatan mengenai Rancangan Kitab Undang-Undang Hukum Acara Pidana (RKUHAP) memicu tingginya polaritas opini yang kerap mengandung ujaran kebencian terselubung dalam sarkasme dan kritik tajam. Pemisahan kritik yang sah dengan ujaran kebencian sulit diatasi metode konvensional. Oleh karena itu, penelitian ini merancang model *deep learning* IndoBERT dengan arsitektur *Multi-Head Classification* bertingkat guna mendeteksi ujaran kebencian sekaligus memetakan sub-kategorinya secara kontekstual.

Penelitian ini menggunakan 6.450 data teks dari X dan Threads yang terbagi secara bertingkat, yaitu Level 1 untuk deteksi biner serta Level 2 untuk lima sub-kategori melalui pelabelan terverifikasi. Pembagian dataset menggunakan rasio 80% data latih, 10% data validasi, dan 10% data uji. Model terbaik dengan formulasi *Hierarchical Masked Loss* berhasil memperoleh performa mumpuni, yaitu dengan tingkat akurasi 93,95% serta nilai Macro F1 sebesar 78,26% pada Level 1. Selain itu pada Level 2 diperoleh akurasi 93,33% dengan nilai Macro F1 64,98%, serta capaian *Composite Macro F1* sebesar 71,62%. Model inferensi ini kemudian diintegrasikan ke dalam sistem web berbasis Laravel dan FastAPI agar pengguna dapat memantau serta menguji deteksi ujaran kebencian secara praktis.