

ABSTRAK

Media sosial seperti Twitter menjadi sumber data penting untuk memahami emosi publik, namun tantangan utama dalam klasifikasi emosi berbahasa Indonesia adalah distribusi kelas yang tidak seimbang (*imbalanced data*) sehingga model cenderung bias terhadap kelas mayoritas. Penelitian ini bertujuan membangun model klasifikasi emosi menggunakan *IndoBERTweet*, *transformer* yang dilatih khusus pada data Twitter Indonesia, dengan penerapan *Focal Loss* sebagai fungsi kerugian alternatif untuk mengatasi ketidakseimbangan tersebut. Dataset yang digunakan adalah *Emotion Dataset from Indonesian Public Opinion* yang berisi 7.080 tweet dengan enam kategori emosi (*anger, fear, joy, love, sadness, dan neutral*). Penelitian ini menguji empat skenario loss function, yaitu *Cross-Entropy*, *Weighted Cross-Entropy*, *Focal Loss Gamma*, dan *Focal Loss Balanced*, dengan evaluasi menggunakan metrik *accuracy, precision, recall, dan F1-score*. Hasil penelitian menunjukkan bahwa *Focal Loss balanced* memberikan performa terbaik dengan *F1-macro* 79,7%, melampaui metode lain sekaligus mengungguli penelitian sebelumnya (*CNN, LSTM, IndoBERT*), sehingga penelitian ini berhasil mencapai *state-of-the-art* baru pada dataset *benchmark* klasifikasi emosi berbahasa Indonesia.

Kata Kunci: klasifikasi emosi, *IndoBERTweet*, *Focal Loss*, *imbalanced data*, Twitter