

ABSTRAK

Keluhan kesehatan ibu hamil dan bayi yang disampaikan melalui platform konsultasi kesehatan daring mengandung informasi penting yang dapat dimanfaatkan untuk mendukung deteksi dini tingkat risiko kesehatan. Penelitian ini bertujuan membangun model klasifikasi tingkat risiko keluhan ibu hamil dan bayi berbasis teks menggunakan IndoBERT yang dikombinasikan dengan *Convolutional Neural Network* (CNN) dan *Bidirectional Long Short-Term Memory* (BiLSTM). *Dataset* penelitian diperoleh dari forum kesehatan Alodokter dan diklasifikasikan ke dalam tiga kategori risiko, yaitu rendah, sedang, dan tinggi yang mengacu pada kategori risiko pada *dataset* UCI *Maternal Health Risk*. Tahapan penelitian meliputi pra pemrosesan data, augmentasi data, tokenisasi, optimasi *hyperparameter*, dan pelatihan serta evaluasi model. Evaluasi performa model dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil evaluasi menunjukkan bahwa model IndoBERT-CNN memperoleh akurasi sebesar 80,15% dan *weighted F1-score* sebesar 80,03%, sedangkan model IndoBERT-BiLSTM memperoleh akurasi sebesar 83,90% dan *weighted F1-score* sebesar 83,84%. Model IndoBERT-BiLSTM mengungguli IndoBERT-CNN pada seluruh metrik evaluasi, yang menunjukkan bahwa kemampuan BiLSTM dalam memproses urutan teks secara *bidirectional* lebih efektif dalam menangkap konteks teks keluhan kesehatan dibandingkan ekstraksi pola lokal oleh CNN.

Kata kunci: IndoBERT, CNN, BiLSTM, klasifikasi risiko, kesehatan ibu dan bayi, pemrosesan bahasa alami