

ABSTRAK

Kesehatan ibu dan anak (KIA) merupakan aspek penting dalam pembangunan kesehatan nasional Indonesia yang memerlukan pengelolaan informasi medis secara efektif. Peningkatan pemanfaatan platform *telemedicine* seperti Alodokter menghasilkan data teks konsultasi KIA dalam jumlah besar. Namun, karakteristik data yang tidak terstruktur dan tidak baku menyulitkan proses ekstraksi informasi klinis secara otomatis. Untuk mengatasi permasalahan tersebut, penelitian ini mengembangkan pendekatan *Named Entity Recognition* (NER) berbasis arsitektur *Indonesian Bidirectional Encoder Representations from Transformers* (IndoBERT) yang dikombinasikan dengan dua varian *Recurrent Neural Network* (RNN), yaitu *Bidirectional Gated Recurrent Unit* (BiGRU) dan *Bidirectional Long Short-Term Memory* (BiLSTM), serta *Conditional Random Field* (CRF). Penelitian ini menggunakan 1.000 data pertanyaan medis terkait KIA dari Alodokter yang dianotasi ke dalam sembilan kategori entitas dan dikonversi ke skema *Beginning, Inside, Outside* (BIO). Pengembangan model dilakukan melalui beberapa skenario eksperimen. Skenario 1 berfokus pada optimasi *hyperparameter* model dasar IndoBERT-BiGRU-CRF dan IndoBERT-BiLSTM-CRF, sedangkan Skenario 2 menguji tiga strategi modifikasi arsitektur, yaitu peningkatan dimensi *hidden* BiGRU (Skenario 2a), implementasi *stacked layer* BiGRU dan *dropout* (Skenario 2b), serta optimasi lapisan klasifikasi menggunakan *Gaussian Error Linear Unit* (GELU) (Skenario 2c). Hasil penelitian menunjukkan bahwa model terbaik diperoleh pada skenario 2b, yaitu arsitektur IndoBERT-BiGRU-CRF dengan *stacked layer* BiGRU yang menghasilkan *test Micro F1-score* sebesar 0,8455. Hasil tersebut menunjukkan bahwa optimasi arsitektur mampu meningkatkan performa model dalam mengenali entitas medis pada teks KIA berbahasa Indonesia.

Kata kunci: *Named Entity Recognition*, IndoBERT, BiGRU, BiLSTM, CRF