

## ABSTRACT

Maternal and child health (MCH) is a key aspect of Indonesia's national health development that requires effective management of medical information. The growing use of telemedicine platforms such as Alodokter has generated a large volume of MCH-related consultation texts. However, its highly unstructured nature and frequent use of informal language pose significant challenges for automated clinical information extraction. To address this challenge, this study proposes a Named Entity Recognition (NER) approach based on IndoBERT combined with two Recurrent Neural Network (RNN) variants, namely Bidirectional Gated Recurrent Unit (BiGRU) and Bidirectional Long Short-Term Memory (BiLSTM), as well as Conditional Random Field (CRF). The study utilizes 1,000 MCH-related medical questions collected from Alodokter, which were manually annotated into nine entity categories and converted into the Beginning, Inside, Outside (BIO) tagging scheme. Model development was conducted through several experimental scenarios. Scenario 1 focused on hyperparameter optimization of the baseline IndoBERT-BiGRU-CRF and IndoBERT-BiLSTM-CRF architectures, while Scenario 2 evaluated three architectural modification strategies: increasing the BiGRU hidden dimension (Scenario 2a), incorporating stacked BiGRU layers with dropout (Scenario 2b), and optimizing the classification layer using the Gaussian Error Linear Unit (GELU) activation function (Scenario 2c). The results show that the best-performing model was obtained in Scenario 2b, where the IndoBERT-BiGRU-CRF architecture with two stacked BiGRU layers achieved a test Micro F1-score of 0.8455. These findings demonstrate that architectural optimization can effectively improve the recognition of medical entities in Indonesian MCH texts.

**Keywords:** Named Entity Recognition, IndoBERT, BiGRU, BiLSTM, CRF