ABSTRACT

The analysis of court decision documents presents a significant challenge when the number of documents is vast, and every part of the ruling must be carefully examined to determine the length of the sentence in a criminal case. This research proposes a regression approach as a solution. Although regression for predicting sentence length in court documents has not been extensively studied, previous research predicted categories and sentence lengths based on archived court records using Word2Vec and LSTM methods. However, this research faced several limitations, one of which was the mismatch of Word2Vec word embeddings with texts that share many similarities. This issue arose from the repetitive use of phrases and formats in Indonesian court documents, rendering LSTM ineffective when combined with Word2Vec, as Word2Vec uses the same word embedding for similar words in different contexts. Based on these shortcomings, this study proposes a regression approach using the BERT model combined with LSTM. In the BERT model, due to the lengthy trial history text, this research utilizes Hierarchical BERT. The *dataset* includes 22.630 court decision documents from general and specific criminal cases, containing sections on legal matters such as charges, facts, legal facts, and legal considerations. This study employs the hyperparameters IndoBERT-base-uncased with a batch size of 32, a learning rate of 0,0001, a hidden size of 128, 1 layer, and 20 epochs, as well as Google-BERT/bert-base-multilingual-uncased with a batch size of 32, a learning rate of 0,0001, a hidden size of 256, 2 layers, and 20 epochs. The results show that the BERT model combined with LSTM without an Attention mechanism yielded the best performance with an R2 Score of 0,5006. Meanwhile, the combination of BERT and LSTM with the Attention mechanism provided an R2 Score of 0,4927. Although slightly lower, this research contributes to the development of legal document regression, particularly in predicting sentence lengths in Supreme Court decision documents.

Keywords : Legal Documents, Sentence Length Prediction, Hyperparameters, BERT (Bidirectional Encoder Representations from Transformers), LSTM (Long Short-Term Memory)