

Abstract

The demand for text summarization is rising due to the increasing amount of online content, particularly in the creation of brief news from news documents. However, text summarization in Indonesian is still not widely used. Indonesian language processing itself poses several difficulties, particularly when it comes to text summarization. This study uses the pre-trained model to combine extractive and abstractive methodologies to provide a hybrid text summarizing method designed for the Indonesian language. The IndoSUM and Liputan6 Indonesian language corpus datasets are used to train the proposed model. The ROUGE score evaluation metric and human evaluation are used to assess the summary findings from the suggested model. It is evident from the evaluation results that the suggested model is comparable to earlier studies. For further study, a deeper investigation was conducted regarding the impact of extractive summarization on the outcomes of abstractive summarization. The experimental results showed that the length of the selected extractive summaries has a slight effect on the outcomes of abstractive summaries. The greater the percentage of sentences' length selected in each dataset, the higher the ROUGE score.

Keywords: summary, extractive, abstractive.