Abstract

Reading news articles is a routine habit for the majority of people, which is usually used as reading material or reference to get the latest information that is currently circulating. One form or way of getting information quickly can be made easier by the existence of news article summaries. News article summarization can also minimize information overload. The aim of this Final Project is to produce better quality summaries on several news topics whose performance is low due to limited data. The summary approach used is an abstract method by utilizing pre-trained models based on transformers. The pre-trained models used are T5-small and BART-BASE and to improve performance, a data augmentation process is applied. Tests in this Final Project were carried out on the XLSum Multi News dataset, which contains categories of government & politics, health, economics, infrastructure development, crime and technology industry with crime and economic having the lowest summary quality. Evaluation of the summary results is done automatically with the ROUGE metric and subjective assessment of the respondents. The evaluation results showed that the ROUGE score did not increase performance, but respondents concluded that in terms of relevance, coherence and suitability, the summary produced after data augmentation was of better quality.

Keywords: Articles,News , Summarization, Abstract, Data Augmentation, Transformers