

Performance Analysis of ChatGPT for Indonesian Abstractive Text Summarization

Rifqi Aulia Rahman¹, Suyanto²

^{1,2}Faculty of Informatics, Telkom University, Bandung

leucopsar@students.telkomuniversity.ac.id, suyanto@telkomuniversity.ac.id

Abstract

In the age of big data, the ability to quickly comprehend and analyze large volumes of information has made effective text summarization increasingly critical. This research focuses on generating concise summaries with human-like coherence by employing prompt engineering and few-shot learning techniques, applied to two Indonesian datasets: Liputan6 and Indosum. The approach demonstrates that advanced summarization algorithms can create brief yet clear summaries. The effectiveness of these summaries is assessed through both human evaluation and automated metrics, particularly emphasizing the ROUGE score. Notably, the study identifies discrepancies between ROUGE scores and human judgments, indicating that traditional automated evaluation methods for text summarization may not fully align with human evaluative standards. This gap suggests a need for enhanced evaluation techniques that better capture the nuances of human judgment. The findings propose a path forward for enhancing the effectiveness and evaluative accuracy of automated summarization systems across various applications.

Keywords: GPT, IndoBART, IndoBERT, ROUGE, text summarization
