

ABSTRACT

The rapid growth of short-form video content across social media platforms has created a demand for efficient automated analysis methods to understand content, characteristics, and dissemination performance, Instagram Reels has emerged as one of the most widely used platforms, while manual content analysis requires substantial time and computational resources, therefore this study focuses on developing a Reels content analysis model using the large language model Qwen2.5-7B-Instruct, which is selected for its strong instruction-following capability and multilingual text processing performance, and fine-tuned using the Quantized Low-Rank Adaptation (QLoRA) method, which is chosen to reduce computational requirements without significantly degrading model performance, the primary objective of this research is to develop a model capable of accurately and efficiently performing video content summarization and content classification, including advertisement identification, the research dataset consists of 655 Instagram Reels posts that are automatically transcribed and subsequently processed through a data cleaning pipeline, the Generate–Annotate–Learn (GAL) approach is applied to enrich the diversity of the training data, the methodological pipeline includes transcript extraction and preprocessing, data augmentation, fine-tuning using QLoRA with 4-bit quantization and LoRA adapters, and performance evaluation using ROUGE-L F1 for summarization and accuracy and F1-score for classification, the evaluation results show that the model achieves an accuracy and F1-score of 0.86 for advertisement detection and approximately 0.61 for content category classification, for the summarization task the ROUGE-L F1 score reaches 0.36 in transcript-based evaluation and 0.33 in generative evaluation, these findings demonstrate that QLoRA is an efficient and resource-efficient approach for adapting Large Language Models (LLM) to Indonesian short-form video content analysis, while also opening opportunities for the development of LLM-based analytical systems that are easier to deploy in industrial settings and future research contexts

Keywords: Short-form Video Content, Instagram Reels, Qwen2.5-7B-Instruct, QLoRA, Generate–Annotate–Learn, Summarization, Content Classification, Advertisement Detection