UNIVERSITAS
GADJAH MADA
FINE-TUNING DAN KOMPARASI MODEL BAHASA PRE-TRAINED UNTUK SUMARISASI TEKS BAHASA INDONESIA
Fariz Muhammad, Dr. Mardhani Riasetiawan, SE Ak, M.T.
Universitas Gadjah Mada, 2025 | Diunduh dari http://etd.repository.ugm.ac.id/

ABSTRACT

Fine-Tuning and Comparing Pre-Trained Language Models for Indonesian Text Summarization

by

Fariz Muhammad

21/475103/PA/20528

The rapid increase in digital textual content has created a pressing need for effective text summarization tools, especially for Indonesian texts where tailored research remains limited. Despite significant advancements in natural language processing, few studies have focused on fine-tuning pre-trained models specifically for Indonesian language tasks. This gap underscores the necessity for specialized summarization systems that can efficiently distill large volumes of text while preserving essential information.

This study fine-tunes and compares three pre-trained language models—IndoBART, IndoT5, and IdT5—for Indonesian text summarization using the IndoSum dataset. The research process involves merging pre-split dataset folds, standardizing and tokenizing the text, and then fine-tuning the models using the Low-Rank Adaptation (LoRa) technique to enhance computational efficiency. Model performance is rigorously evaluated using multiple metrics: ROUGE-1, ROUGE-2, ROUGE-L, METEOR, and BERTScore, which collectively assess content accuracy, semantic similarity, and fluency.

Evaluation reveals that IndoBART achieves a ROUGE-1 of 38.30, ROUGE-2 of 33.91, and ROUGE-L of 37.36, alongside a METEOR score of 23.37 and a BERTScore F1 of 89.53. Idt5 shows moderate improvements with a ROUGE-1 of 40.0, ROUGE-2 of 35.02, ROUGE-L of 39.36, METEOR score of 25.55, but falls slightly with BERTScore F1 of 89.03. Notably, IndoT5 outperforms both models by achieving the highest scores—ROUGE-1 of 41.66, ROUGE-2 of 37.06, and ROUGE-L of 40.72; a METEOR score of 26.43; and a BERTScore F1 of 90.06. These results demonstrate IndoT5's superior ability to capture essential vocabulary, complex phrase structures, and nuanced semantic relationships in Indonesian text summarization.

Keywords: Text Summarization, Pre-trained Language Models, Natural Language Processing (NLP), Low-Rank Adaptation (LoRa), Fine-tuning, Evaluation Metrics.

UNIVERSITAS
GADJAH MADA
FINE-TUNING DAN KOMPARASI MODEL BAHASA PRE-TRAINED UNTUK SUMARISASI TEKS BAHASA INDONESIA
Fariz Muhammad, Dr. Mardhani Riasetiawan, SE Ak, M.T.
Universitas Gadjah Mada, 2025 | Diunduh dari http://etd.repository.ugm.ac.id/

ABSTRAK

Fine-Tuning dan Komparasi Model Bahasa Pre-Trained untuk Sumarisasi Teks Bahasa Indonesia

oleh

Fariz Muhammad

21/475103/PA/20528

Pesatnya pertumbuhan konten tekstual digital mendorong kebutuhan akan alat peringkasan teks yang efektif, terutama untuk bahasa Indonesia yang masih kurang mendapat perhatian dalam penelitian. Meski *natural language processing (NLP)* telah berkembang pesat, studi yang secara khusus mengeksplorasi *fine-tuning* model *pre-trained* untuk bahasa Indonesia masih terbatas. Hal ini menegaskan perlunya sistem peringkasan yang mampu merangkum teks dalam jumlah besar secara efisien tanpa kehilangan informasi penting.

Penelitian ini melakukan *fine-tuning* dan evaluasi terhadap tiga model bahasa *pre-trained*—IndoBART, IndoT5, dan IdT5—dalam tugas peringkasan teks bahasa Indonesia menggunakan *dataset* IndoSum. Prosesnya mencakup penggabungan lipatan *dataset*, standarisasi, tokenisasi, dan pelatihan ulang dengan teknik *Low-Rank Adaptation (LoRa)* untuk efisiensi komputasi. Kinerja model diukur menggunakan metrik ROUGE-1, ROUGE-2, ROUGE-L, METEOR, dan BERTScore yang menilai akurasi isi, kesamaan semantik, dan kefasihan.

Hasil evaluasi menunjukkan bahwa IndoBART memperoleh ROUGE-1 sebesar 38,8, ROUGE-2 sebesar 34,56, dan ROUGE-L sebesar 37,95, METEOR sebesar 23,75, serta BERTScore F1 sebesar 87,95. IdT5 memberikan peningkatan sedang dengan ROUGE-1 sebesar 40,83, ROUGE-2 sebesar 35,92, ROUGE-L sebesar 40,17, METEOR 26,19, dan BERTScore F1 sebesar 88,14. IndoT5 mencatat performa terbaik dengan ROUGE-1 sebesar 42,01, ROUGE-2 sebesar 37,56, ROUGE-L sebesar 41,13, METEOR 26,72, dan BERTScore F1 sebesar 88,66. Ini menunjukkan kemampuan IndoT5 dalam menangkap kosakata penting, struktur kalimat kompleks, dan hubungan semantik yang halus pada tugas peringkasan teks bahasa Indonesia.

Keywords: Sumarisasi Teks, Pre-trained Language Models, Natural Language Processing (NLP), Low-Rank Adaptation (LoRa), Fine-tuning, Metrik Evaluasi.