

**PERBANDINGAN PERFORMA MODEL INDOBERT DAN PEGASUS
UNTUK PERINGKASAN TEKS ABSTRAKTIF PADA DATASET
ARTIKEL BERITA *ONLINE* BERBAHASA INDONESIA**

Oleh

Ketut Tuter Diatmika, NIM 2015101011

Jurusan Teknik Informatika

ABSTRAK

Penelitian ini bertujuan untuk membandingkan performa dua model berbasis Transformer, yaitu IndoBERT+GPT-2 decoder dan PEGASUS-Large, dalam tugas peringkasan teks abstraktif pada artikel berita berbahasa Indonesia. Dataset yang digunakan adalah *Indonesian News Dataset* dari Kaggle dengan jumlah awal 32.148 data, yang setelah melalui tahap *pre-processing* menjadi 29.985 pasangan teks dan ringkasan. Data dibagi menjadi tiga bagian dengan proporsi 70% untuk pelatihan, 15% untuk validasi, dan 15% untuk pengujian. Metodologi penelitian meliputi tahapan *pre-processing*, implementasi model, pelatihan, inferensi, serta evaluasi menggunakan metrik ROUGE (ROUGE-1, ROUGE-2, ROUGE-L), BERTScore-F1, dan METEOR. IndoBERT diimplementasikan dalam arsitektur *encoder-decoder* dengan GPT-2 sebagai *decoder*, sedangkan PEGASUS-Large dilatih menggunakan pendekatan LoRA untuk mengatasi keterbatasan sumber daya komputasi. Hasil penelitian menunjukkan bahwa IndoBERT+GPT-2 decoder unggul pada seluruh metrik evaluasi dibandingkan PEGASUS-Large. Skor ROUGE-2 yang lebih tinggi menegaskan kemampuan model dalam menangkap struktur frasa dua kata berurutan, sementara skor BERTScore dan METEOR yang lebih baik menunjukkan kesesuaian semantik dan kualitas parafrase yang lebih tinggi. Analisis kualitatif berbasis *case-by-case* juga memperlihatkan bahwa IndoBERT lebih konsisten menghasilkan ringkasan kategori *Top*, sedangkan PEGASUS cenderung menghasilkan lebih banyak ringkasan kategori *Divergent* dan *Bottom*. Temuan ini menegaskan bahwa model monolingual yang dilatih khusus untuk bahasa Indonesia memiliki keunggulan dalam tugas peringkasan teks dibandingkan model generik yang diadaptasi.

Kata Kunci: IndoBERT, PEGASUS, peringkasan teks, NLP, METEOR.

***PERFORMANCE COMPARISON OF INDOBERT AND PEGASUS MODELS
FOR ABSTRACTIVE TEXT SUMMARIZATION ON INDONESIAN ONLINE
NEWS ARTICLE DATASETS***

By

Ketut Tuter Diatmika, NIM 2015101011

Informatic Engineering Major

ABSTRACT

This study aims to compare the performance of two Transformer-based models, namely IndoBERT+GPT-2 decoder and PEGASUS-Large, in generating abstractive summaries of Indonesian online news articles. The dataset used was the Indonesian News Dataset from Kaggle, comprising 32,148 raw entries, which after pre-processing resulted in 29,985 text-summary pairs. The dataset was split into training (70%), validation (15%), and testing (15%) sets. The research methodology consisted of data pre-processing, model implementation, training, inference, and evaluation using ROUGE metrics (ROUGE-1, ROUGE-2, ROUGE-L), BERTScore-F1, and METEOR. IndoBERT was implemented in an encoder-decoder architecture with GPT-2 as the decoder, while PEGASUS-Large was fine-tuned using the LoRA approach to address computational resource limitations. The results indicate that IndoBERT+GPT-2 decoder outperformed PEGASUS-Large across all evaluation metrics. Higher ROUGE-2 scores demonstrate its ability to capture bigram structures more effectively, while superior BERTScore and METEOR values highlight better semantic alignment and paraphrasing quality. The case-by-case qualitative analysis further shows that IndoBERT consistently produced more *Top* summaries, whereas PEGASUS tended to generate more *Divergent* and *Bottom* cases. These findings confirm that monolingual models trained specifically for Indonesian are more effective for abstractive summarization tasks than generic models adapted from other languages.

Keywords: IndoBERT, PEGASUS, text summarization, NLP, METEOR.