

INTISARI

Mesin Penerjemah Jawa ke Indonesia dengan Arsitektur Berbasis Transformer

Oleh

Vincent Yeozeziel
21/481885/PA/20999

Penurunan penggunaan bahasa daerah, khususnya Bahasa Jawa, di Indonesia menimbulkan kekhawatiran terkait pelestarian budaya lokal di era modern. Upaya pelestarian memerlukan pendekatan yang relevan dengan perkembangan teknologi saat ini. Teknologi mesin penerjemah berbasis *neural* telah menunjukkan potensi untuk menjembatani bahasa-bahasa daerah dengan bahasa nasional, serta berkontribusi pada pelestarian bahasa melalui aksesibilitas digital.

Mesin penerjemah berbasis *neural* dengan arsitektur Transformer menjadi model yang mutakhir karena algoritmanya dapat diparalelkan dan dapat menangkap dependensi jarak jauh dari kata ke kata sehingga mudah untuk diskalakan menjadi lebih besar. Meskipun demikian, penerjemah bahasa di Indonesia masih sedikit yang mengeksplorasi metode ini. Metode yang saat ini lebih banyak dilakukan adalah dengan melakukan *fine-tuning* terhadap model bahasa yang telah ada yang belum dirancang untuk melakukan tugas penerjemahan. Selain itu, *fine-tuning* memerlukan sumber daya yang besar untuk melakukannya.

Studi ini bertujuan untuk mengusulkan pendekatan pembangunan model mesin penerjemahan dengan Transformer untuk tugas penerjemahan bahasa Jawa ke bahasa Indonesia. Studi ini mencakup pembangunan dan percobaan hiperparameter arsitektur Transformer untuk mencapai metrik evaluasi yang baik. Evaluasi BLEU skor yang didapat dalam penelitian ini adalah 39.33.

Kata-kata kunci : *Deep Learning*, Pemrosesan Bahasa Alami, Komputasi Linguistik, BLEU, Mekanisme *Attention*

ABSTRACT

Javanese-to-Indonesian Machine Translation with Transformer-Based Architecture

By

Vincent Yezekiel
21/481885/PA/20999

The decline in the use of regional languages, particularly Javanese, in Indonesia raises concerns regarding the preservation of local culture in the modern era. Preservation efforts require approaches that are relevant to current technological advancements. Neural machine translation technology has shown potential in bridging regional languages with the national language, as well as contributing to language preservation through digital accessibility.

Neural machine translation with a Transformer architecture has become a state-of-the-art model because its algorithm is parallelizable and capable of capturing long-range dependencies between words, making it scalable to larger sizes. Nevertheless, language translation research in Indonesia has rarely explored this method. Current approaches more commonly involve fine-tuning existing language models that were not initially designed for translation tasks. Moreover, fine-tuning requires substantial computational resources.

This study aims to propose a model development approach for machine translation using the Transformer architecture for the task of translating Javanese to Indonesian. The study involves the construction and experimentation of Transformer hyperparameters to achieve good evaluation metrics. The BLEU score achieved in this research is 39.33.

Keywords : Deep Learning, Natural Language Processing, Computational Linguistic, BLEU, Attention Mechanism