

**IMPLEMENTASI TRANSFORMER MODEL
UNTUK SISTEM PENERJEMAHAN
LOW RESOURCE LANGUAGE
BAHASA KAWI KE BAHASA INDONESIA**

Oleh

I Putu Arsana Putra, NIM 1815091016

Jurusan Teknik Informatika

Program Studi Sistem Informasi

ABSTRAK

Bahasa Kawi adalah salah satu bahasa yang berkembang pada abad 9 sampai ke 15 pada peradaban Hindu Jawa. Bahasa Kawi saat ini sering dipakai sebagai bahasa pengantar pada upacara Agama Hindu. Adanya potensi punahnya bahasa Kawi diakibatkan beberapa hal seperti media dari bahasa Kawi itu yang mudah rusak, penutur atau penggiat bahasa Kawi yang terus berkurang. Upaya yang telah dilakukan oleh pada linguistik adalah preservasi terhadap manuskrip seperti pemeliharaan terhadap manuskrip ataupun penerjemahan bahasa yang terkandung dalam manuskrip dalam hal ini Bahasa Kawi ke dalam bahasa Indonesia. Walaupun begitu hal ini terbatas pada jumlah manuskrip yang mampu di terjemahkan dikarenakan waktu yang diperlukan dalam melakukan penerjemahan atau anggaran yang perlu di alokasikan. Penelitian ini merupakan kelanjutan dari penelitian sebelumnya dalam bidang pemanfaatan *neural network* dalam penerjemahan bahasa Kawi dengan pengembangannya menggunakan Transformer. Transformer adalah salah satu arsitektur berbasis *neural network* yang dimanfaatkan dalam penyelesaian masalah penerjemahan bahasa dan sampai karya tulis ini dibuat merupakan *state of the art* dalam *neural machine translation* (NMT). Metodologi penelitian dari karya tulis ini di antaranya akuisisi data yang dari penelitian (Budaya, et. all), *pre processing* yang terdiri dari pembersihan data serta tokenisasi data yang terdiri dari *Word Level Tokenizer* dan *Subword Level Tokenizer*, pembangunan model Transformer, pelatihan model Transformer, proses decoding dengan Greedy dan Beam dan evaluasi model menggunakan skor BLEU. Hasil penelitian menunjukkan model Transformer terbaik memiliki skor BLEU masing-masing gram sebesar 21.13, 13.54, 7.36, 2.90 dengan Mean skor BLEU adalah 11.23 pada *Word Level*. Sedangkan pada *Subword Level* memiliki skor BLEU terbaik sebesar 15.67, 9.52, 5.64, 2.77 dengan Mean skor BLEU adalah 8.40.

Kata Kunci : Bahasa Kawi, Bahasa Jawa Kuno, *Neural Machine Translation*, Tokenizer, Transformer.

**IMPLEMENTASI TRANSFORMER MODEL
UNTUK SISTEM PENERJEMAHAN
LOW RESOURCE LANGUAGE
BAHASA KAWI KE BAHASA INDONESIA**

Oleh

I Putu Arsana Putra, NIM 1815091016

Jurusan Teknik Informatika

Program Studi Sistem Informasi

ABSTRAK

Kawi language is one of the languages that developed between the 9th and 15th centuries during the Hindu-Javanese civilization. Today, it is often used as a ceremonial language in Hindu religious rituals. However, the Kawi language is at risk of extinction due to several factors, such as the fragile nature of its traditional medium and the declining number of speakers or enthusiasts. Linguistic efforts to preserve the language have focused on manuscript conservation and the translation of Kawi texts into Indonesian. Nevertheless, these efforts are limited by the small number of manuscripts that can be translated, mainly due to the time and budget required for such work. This research is a continuation of previous studies on the use of neural networks for translating the Kawi language, with further development using the Transformer architecture. Transformer is a neural network-based architecture widely used in machine translation tasks and, at the time of this writing, remains the state-of-the-art model in Neural Machine Translation (NMT). The methodology of this study includes data acquisition (from Budaya et al.), preprocessing consisting of data cleaning and tokenization using both word-level and subword-level tokenizers, Transformer model construction, model training, decoding using Greedy and Beam search methods, and model evaluation using BLEU scores. The results show that the best-performing Transformer model at the word level achieved BLEU scores of 21.13, 13.54, 7.36, and 2.90 for 1-gram to 4-gram, with a mean BLEU score of 11.23. At the subword level, the best BLEU scores were 15.67, 9.52, 5.64, and 2.77, with a mean BLEU score of 8.40.

Key Word: *Kawi Language, Old Javanese Language, Neural Machine Translation, Tokenizer, Transformer.*