

Mesin Penerjemah Hasil Transliterasi Lontar Bali dalam Bahasa Kawi ke dalam Bahasa Indonesia dengan *Deep Learning*

Oleh

I Gede Bintang Arya Budaya, NIM 2029101004

Program Studi Ilmu Komputer

ABSTRAK

Mesin penerjemah adalah pilihan utama bagi orang-orang, terutama untuk mempelajari pengetahuan yang tidak menggunakan bahasa asli mereka. Bahasa Kawi sebagai bahasa lokal, meskipun jarang digunakan secara aktif, merupakan bahasa yang banyak digunakan dalam berbagai literatur lama seperti lontar, khususnya di Bali yang masih dipelajari hingga saat ini. Banyak mesin penerjemah dengan bahasa pengguna aktif tinggi, seperti bahasa Indonesia dan Inggris yang terus dikembangkan tetapi tidak untuk bahasa Kawi. Pengembangan model mesin penerjemah berbasis *neural* dan implementasinya dengan *microframework Flask* merupakan langkah awal pengembangan mesin penerjemah dari bahasa Kawi ke Bahasa Indonesia. Pengembangan dimulai dengan membuat korpus paralel Kawi ke Indonesia dan dilanjutkan dengan membuat arsitektur mesin penerjemah dengan model berbasis *Recurrent Neural Network (RNN)* yaitu *Simple RNN*, *Bidirectional RNN*, *Gated Recurrent Unit (GRU)*, *Bidirectional GRU*, *Long Short-Term Memory (LSTM)*, *Bidirectional LSTM*, LSTM dengan mekanisme *Attention*, dan terakhir *Bidirectional LSTM* dengan mekanisme *Attention*. Model mesin penerjemah tersebut lalu dibandingkan berdasarkan performa melalui parameter *BLEU scores*. Berdasarkan hasil pengujian, dengan dataset yang masih terbatas model LSTM dengan mekanisme *Attention* mampu mencapai skor rata-rata tertinggi BLEU sebesar 20.86. Meskipun menghasilkan performa terbaik dalam percobaan, namun model mesin penerjemah LSTM dengan mekanisme *Attention* masih belum mampu mencapai rata – rata BLEU yang disebutkan dalam level memahami konteks yaitu di nilai *BLEU Scores* 60 - 70. Penambahan dataset korpus paralel Kawi ke Indonesia dapat menjadi salah satu solusi pengembangan mesin penerjemah kedepannya.

Kata-kata kunci: mesin penerjemah, bahasa Kawi, bahasa Indonesia, *deep learning*, Nilai BLEU.

*Translation Machine for Balinese Palm Leaf Manuscript Transliteration from
Kawi into Indonesian with Deep Learning*

By

I Gede Bintang Arya Budaya, NIM 2029101004

Program Studi Ilmu Komputer

ABSTRACT

Machine translation is the primary choice for people, especially for learning knowledge beyond their native language. Kawi language as a local language, although rarely used actively, is a language used in various old literature such as palm-leaf manuscripts, especially in Bali, and still studied today. Many machine translations with high active user languages, such as Indonesian and English are developed continuously but not for Kawi. Building the neural-based machine translation model and its implementation with the microframework Flask is the initial steps for a robust machine translation from Kawi to Indonesia. The first step was to make a parallel corpus from Kawi to Indonesia. The machine translation architecture builds based on Recurrent Neural networks (RNN), there are Simple RNN, Bidirectional RNN, Gated Recurrent Unit (GRU), Bidirectional GRU, Long Short-Term Memory (LSTM), Bidirectional LSTM, LSTM with Attention mechanism, and finally Bidirectional LSTM with an Attention mechanism. The comparison for each machine translation model performance through the BLEU scores parameter. Based on the test results, with a dataset that is still limited, the LSTM model with the Attention mechanism can achieve the highest BLEU average score of 20.86. However, although producing the best performance in the experiment, the LSTM with the Attention mechanism model is unable to achieve the average BLEU stated in the level of understanding the context, with the BLEU Scores score of 60 - 70. Increasing the dataset size of the Kawi to Indonesia can be one solution for the future development of machine translation.

Keywords: machine translation, Kawi language, Indonesian language, deep learning, BLEU Scores