

ABSTRAK

PENERJEMAHAN BAHASA LAMPUNG-BAHASA INDONESIA MENGUNAKAN *LONG SHORT-TERM MEMORY*

Oleh

MOHAMMED RAIHAN AKBAR

Penelitian ini bertujuan untuk mengembangkan model penerjemahan otomatis dari bahasa Lampung Api ke bahasa Indonesia menggunakan *long short-term memory* (LSTM). Mengingat bahasa Lampung merupakan bahasa daerah yang terancam punah, penelitian ini fokus pada penerapan model LSTM untuk mengatasi tantangan dalam menerjemahkan bahasa tersebut. Model dilatih dengan data yang telah dibersihkan dan dilakukan *tokenisasi*, lalu dievaluasi menggunakan analisis kuantitatif berupa plot *loss* dan metrik BLEU dan analisis kualitatif berupa contoh hasil penerjemahan. Hasil evaluasi pada set latih menunjukkan skor BLEU 1-gram sebesar 0,7900, namun menurun pada n-gram lebih tinggi: 2-gram sebesar 0,7025, 3-gram sebesar 0,6205, dan 4-gram sebesar 0,5290. Pada set uji, skor BLEU 1-gram turun menjadi 0,5118, dengan penurunan pada n-gram berikutnya. Model menunjukkan kesulitan dalam menangani konteks kompleks, terutama pada kalimat panjang, di mana skor 4-gram turun menjadi 0,0532 di set uji pada kalimat 10 kata. Temuan ini mengindikasikan keterbatasan LSTM dalam menangani *long-range dependencies*. Penambahan mekanisme *attention* dapat dipertimbangkan untuk perbaikan. Penelitian ini diharapkan berkontribusi pada pengembangan teknologi penerjemahan bahasa daerah di Indonesia dan mendukung pelestarian Bahasa Lampung.

Kata kunci: penerjemahan mesin, bahasa Lampung, bahasa Indonesia, *long short-term memory*, pemrosesan bahasa alami, pelestarian bahasa

ABSTRACT

LONG SHORT-TERM MEMORY FOR LAMPUNG-INDONESIA TRANSLATION

BY

MOHAMMED RAIHAN AKBAR

This research aims to develop an automatic translation model from the Lampung Api language to Indonesian using long short-term memory (LSTM). Considering that the Lampung language is an endangered regional language, this study focuses on applying the LSTM model to address the challenges in translating the language. The model was trained on cleaned and tokenized data and evaluated using quantitative analysis in the form of loss plots and BLEU metrics, as well as qualitative analysis through translation examples. Evaluation results on the training set showed a BLEU 1-gram score of 0.7900, but it decreased with higher n-grams: 2-gram at 0.7025, 3-gram at 0.6205, and 4-gram at 0.5290. On the test set, the BLEU 1-gram score dropped to 0.5118, with further declines in the subsequent n-grams. The model demonstrated difficulties in handling complex contexts, especially in long sentences, where the 4-gram score dropped to 0.0532 in the test set for sentences with 10 words. These findings indicate the limitations of LSTM in handling long-range dependencies. The addition of an attention mechanism may be considered for improvement. This research is expected to contribute to the development of regional language translation technology in Indonesia and support the preservation of the Lampung language.

Keywords: machine translation, Lampung language, Indonesian language, long short-term memory, natural language processing, language preservation