

ABSTRAK

Ronggo Haikal, 2024. **Implementasi Arsitektur Transformers penerjemahan Otomatis Bahasa Indonesia ke Bahasa Papua Kokas**. Tugas Akhir, Program Studi Informatika (S1), Sekolah Tinggi Informatika & Komputer Indonesia, Pembimbing: Nira Radita, Co. Pembimbing: Mukhlis Amien

Kata kunci: Arsitektur Transformers, MarianMT, BLEU, Bahasa Papua Kokas, Penerjemahan otomatis

Penelitian ini bertujuan mengimplementasikan arsitektur Transformers untuk penerjemahan otomatis dari Bahasa Indonesia ke Bahasa Papua Kokas menggunakan framework MarianMT. Tantangan yang dihadapi adalah kurangnya sumber daya bahasa Papua Kokas yang mengancam kelestariannya. Dataset yang digunakan dalam penelitian ini terdiri dari 2908 pasangan kalimat. Model dilatih dengan teknik tokenisasi dan diuji menggunakan metrik BLEU. Evaluasi yang dilakukan menunjukkan bahwa model dapat menerjemahkan teks dengan akurasi yang memadai, di mana peningkatan skor BLEU terlihat pada setiap epoch. Model kemudian diimplementasikan dalam aplikasi berbasis web yang memungkinkan penerjemahan secara real-time. Hasil penelitian menunjukkan bahwa pendekatan arsitektur Transformers layak diterapkan untuk penerjemahan Bahasa Indonesia ke Bahasa Papua Kokas. Saran untuk pengembangan lebih lanjut termasuk memperluas dataset, mengeksplorasi teknik optimasi lain seperti TER dan METEOR, serta mengadaptasi model untuk bahasa daerah lain.

ABSTRACT

Ronggo Haikal, 2024. Implementation of Transformers Architecture for Automatic Translation of Indonesian to Papuan Language Kokas. Final Project, Study Program Informatika S1, Sekolah Tinggi Informatika & Komputer Indonesia, Advisor 1: Nira Radita, Advisor 2: Mukhlis Amien

Keyword: Transformers Architecture, MarianMT, BLEU, Papuan Kokas Language, Automatic translation

This study aims to implement the Transformers architecture for automatic translation from Indonesian to Papuan Kokas language using the MarianMT framework. The challenge faced is the lack of Papuan Kokas language resources that threaten its sustainability. The dataset used in this study consists of 2908 sentence pairs. The model was trained using tokenization techniques and tested using the BLEU metric. The evaluation conducted showed that the model can translate text with adequate accuracy, where an increase in the BLEU score is seen in each epoch. The model was then implemented in a web-based application that allows real-time translation. The results of the study indicate that the Transformers architecture approach is feasible to be applied for translation from Indonesian to Papuan Kokas language. Suggestions for further development include expanding the dataset, exploring other optimization techniques such as TER and METEOR, and adapting the model to other regional languages.