Statistical Machine Translation Dayak Language – Indonesia Language

Muhammad Fiqri Khaikal [1], Arie Ardiyanti Suryani [2]
Informatika, Fakultas Informatika, Universitas Telkom
Jl.Telekomunikasi Terusan Buah Batu Sukapura Kec. Dayeuhkolot, Kota Bandung, 40257
E-Mail : fiqrikhaikal@student.telkomuniversity.ac.id [1]; ardiyanti@telkomuniversity.ac.id [2]

ABSTRACT

This Paper aims to discuss how to create the local language machine translation of Indonesia Language where the reason of local language selection was carried out as considering the using of machine translator for local language   are still infrequently found mainly for Dayak Language machine translator. Machine Translation on this research had used statistical approach where the resource data that was taken originated from articles on dayaknews.com pages with total parallel corpus was approximately 1000 Dayak Language – Indonesia Language furthermore as this research contains the corpus with total 1000 sentences accordingly divided into three sections in order to comprehend the certain analysis from a pattern that was created. The monolingual corpus was collected approximately 1000 sentences of Indonesia Language. The testing was carried out using Bilingual Evaluation Understudy (BLEU) tool and had result the highest accuracy value amounting to 49.15% which increase from some the others machine translator amounting to approximately 3%.

Key words - Machine Translator, Statistical, Parallel Corpus, Monolingual Corpus, BLEU