

Abstract

This paper will discuss the translation machine with a statistical approach where this machine can do Sundanese-Javanese translation based on phrases. The data used is in the form of 1233 pairs of sentences or parallel corpus lines of Sundanese and Javanese, while monolingual corpus consists of 1273 sentences or lines of Javanese. Testing is done by creating two testing schemes, namely by adding quantities of parallel and monolingual corpus. Test results are evaluated automatically using Bilingual Evaluation Understudy (BLEU). In the first scheme, several configurations of the translation machine using 1273 sentences or lines of a monolingual corpus in Javanese. From some tests, the best BLEU score was 42.13%. Based on the test results showed the quantity of parallel corpus and monolingual corpus can increase the evaluation value of Sundanese -Javanese statistical translation machine.

Keywords: statistical machine translator, Sundanese-Javanese language, corpus, BLEU.
