Abstract

Evaluation of machine translation is an important problem in machine translation development. Evaluation is a process that purpose to assess a machine translation based on its output.

This final assignment implements BLEU (BiLingual Evaluation Understudy) method for evaluating machine translation (MT) system that based on modified n-gram precision computation. The basic idea in calculation of this metric is to compare n-gram(s) of the candidate with the n-gram(s) of the reference translation and count the number of matches. The more matches, the better candidate translation is. Moreover, this final assignmet also implements the modification of BLEU method with Information Extraction concept that used weighting model.

Evaluating is done by looking up the correlation or the association of score that produce by human (human judgement). The result of this experiment shows that the computation of BLEU weighting method has more correlation with human judgement rather than the baseline of BLEU.

Keywords *: evaluation, BLEU method, N-gram(s) Co-occurance, modified n-gram(s) precision, weighting, correlation*