

Abstrak

Mengevaluasi terjemahan merupakan permasalahan yang penting dalam perkembangan sistem mesin penterjemah atau *Machine Translation*. Evaluasi ini merupakan suatu proses yang bertujuan untuk menilai suatu sistem mesin penterjemah yang dilihat dari hasil atau keluaran suatu sistem mesin penterjemah teks tersebut.

Pada Tugas Akhir ini diimplementasikan metoda BLEU (BiLingual Evaluation Understudy) yang merupakan metoda untuk mengevaluasi sistem mesin penterjemah (MT) yang berbasis presisi yang dimodifikasi (*modified n-gram(s) precision*). Ide dari metoda BLEU ini adalah untuk membandingkan *n-gram(s)* terjemahan keluaran dari suatu MT (kandidat) dengan *n-gram(s)* terjemahan lain yang dibuat oleh manusia dan menghitung jumlah *n-gram(s)* yang cocok. Makin banyak yang cocok, maka terjemahan kandidat semakin bagus dan *score* yang dihasilkan dari evaluasi dengan metoda BLEU ini semakin tinggi. Selain itu, diterapkan juga pengembangan metoda BLEU ini dengan konsep *Information Extraction* yaitu dengan dengan teknik pembobotan.

Pengujian dilakukan dengan melihat korelasi atau keterhubungan *score* yang dihasilkan dengan penilaian manusia (*human judgement*). Hasil pengujian menunjukkan bahwa perhitungan *score* BLEU dengan pembobotan memiliki korelasi yang lebih besar dengan penilaian manusia daripada metoda BLEU dasar.

Kata kunci : evaluasi, metoda BLEU, kecocokan *n-gram(s)*, *modified n-gram precision*, pembobotan, korelasi