

Abstrak

Mengevaluasi ringkasan merupakan salah satu permasalahan yang penting dalam perkembangan *Automatic Text Summarization* (ATS). Evaluasi ini merupakan suatu proses yang bertujuan untuk menilai suatu sistem peringkasan teks otomatis yang dilihat dari hasil atau keluaran suatu sistem peringkasan teks otomatis tersebut.

Pada Tugas Akhir ini diimplementasikan pengembangan variasi metoda ROUGE (*Recall-Oriented Understudy for Gisting Evaluation*) yang merupakan metoda untuk mengevaluasi sistem peringkasan teks otomatis (ATS), yaitu ROUGE-NW yang merupakan penggabungan dari metode ROUGE-1(*n*-gram 1) dengan metode ROUGE-W. ROUGE merupakan suatu teknik evaluasi yang digunakan untuk menentukan kualitas dari sebuah ringkasan secara otomatis dengan cara membandingkan ringkasan yang dihasilkan oleh ATS dengan ringkasan (ideal) lain yang dibuat oleh manusia [5]. Selain itu, sebagai pembanding diterapkan juga metode asli ROUGE yaitu ROUGE-1, ROUGE-2, dan ROUGE-W.

Pengujian dilakukan dengan melihat korelasi atau keterhubungan *score* yang dihasilkan oleh masing-masing metode ROUGE dengan penilaian yang dihasilkan oleh manusia (*human judgement*). Hasil pengujian menunjukkan bahwa perhitungan *score* ROUGE-2 dan ROUGE-NW cenderung memiliki korelasi yang lebih tinggi dengan penilaian manusia dibandingkan dengan metode ROUGE-1 dan ROUGE-W.

Kata kunci : evaluasi, metode ROUGE, *n*-gram(*s*), korelasi