Abstrak

Text Summarization adalah sebuah proses untuk menghasilkan ringkasan *(summary)* dari suatu artikel tapi tetap memiliki gambaran yang akurat dari isi suatu artikel. Tujuannya adalah mengambil sumber informasi dengan mengutip sebagian besar isi yang penting dan menampilkan kepada pembaca dalam bentuk yang ringkas dan sesuai dengan kebutuhan pembaca.

Sistem ini mengimplementasikan metode semantik Lexical Chain [1] yaitu salah satu metode pada *text summarization* yang membentuk rantai leksikal berdasarkan hubungan semantik antar kata dalam teks. Sebagai pembanding digunakan metode peringkasan teks dengan TF-IDF *(Term Frequency*Inverse Document Frequency)* dengan menambahkan konsep ekstraksi frase utama *(Keyphrase Extraction)* [10] dari suatu teks. Hasil ringkasan yang dihasilkan tetap memiliki/ mengandung bagian-bagian yang penting dari artikel asli sehingga diharapkan dapat membantu pembaca untuk menyerap informasi yang ada dalam artikel melalui ringkasan *(summary)*.

Hasil pengujian dari 40 *single document* menunjukkan bahwa pengunaan metode Lexical Chain memberikan performansi yang baik pada kisaran *summary length* 50%-20% dengan panjang dokumen < 29 kalimat (rata-rata panjang kalimat dari 40 dokumen uji). Metode ini memberikan nilai precision >= 0.502 dan F-Measure >= 0.312, sehingga lebih baik daripada TF-IDF dengan *Keyphrase Extraction*.

Pengujian *multidocuments* dari 6 folder dokumen uji menunjukkan bahwa pengunaan metode TF-IDF dengan *Keyphrase Extraction* memberikan performansi yang baik pada kisaran *summary length* 20%-10%. Metode ini memberikan nilai precision >= 0.524 dan F-Measure >= 0.191, sehingga lebih baik bila dibandingkan dengan hasil metode Lexical Chain dan hasil evaluasi TIPSTER SUMMAC.

Kata kunci: *Text Summarization*, Lexical Chain, TF-IDF, Frase Utama.