

Abstrak

Peringkasan teks otomatis (*automatic text summarization*) adalah proses mengambil informasi yang paling penting dari sebuah teks atau beberapa teks untuk membuat sebuah versi ringkas dari teks dengan menggunakan aplikasi berbasis komputer.

Pada Tugas Akhir ini diimplementasikan suatu teknik *Automatic Text Summarization* berdasarkan pendekatan Graf untuk multi dokumen berita. Proses ini menghasilkan keluaran berupa ringkasan ekstraktif yang terdiri dari kalimat-kalimat. Metode perankingan berbasis graf yang diterapkan adalah *LexRank*[6] yang akan meranking kalimat dari multi dokumen yang digunakan, dengan menghitung nilai *centrality* suatu kalimat, ranking kalimat tertinggi kemudian akan di ekstrak menjadi ringkasan. Adapun metode *centrality* yang digunakan pada tugas akhir ini adalah metode *degree centrality* yang merupakan adaptasi langsung dari *LexRank* dan metode *modified LexRank centrality*, yang merupakan modifikasi *LexRank* dengan ide dasar *PageRank*. Perhitungan Kedua metode tersebut ditentukan oleh perhitungan *similarity* antar kalimat, metode *similarity* yang digunakan pada tugas akhir ini, yaitu metode *idf modified cosine similarity*, yang merupakan bawaan langsung dari *LexRank*, dan metode *Long Common Subsequences similarity* yang merupakan modifikasi untuk melihat *similarity* antar kalimat berdasarkan makna. Selanjutnya hasil ringkasan yang didapat belum terurut untuk itu di perlukan proses *ordering*. *Ordering* yang akan digunakan adalah metode *chronological ordering*. Permasalahan suatu peringkasan multi dokumen adalah terdapatnya *redundancy data* atau informasi yang berulang untuk menghilangkannya di lakukan proses *Reranker*.

Pengevaluasian dilakukan menggunakan *ROUGE evaluation toolkit*. “Hasil pengujian menunjukkan metode *degree centrality* dengan *idf modified cosine similarity* memiliki akurasi lebih baik dibandingkan ke tiga metode lainnya yang diimplementasikan pada Tugas Akhir ini” .

Kata kunci: peringkasan teks, *LexRank*, *Degree centrality*, *Modified LexRank centrality*, *idf modified cosine similarity*, *Long common subsequences similarity*, *Reranker*, *chronological ordering*.