Abstrak

Peringkasan teks otomatis adalah aplikasi berbasis komputer yang mampu meringkas sebuah atau beberapa teks dengan menyaring informasi paling penting didalamnya. Kemampuan meringkas ini menghasilkan ringkasan yang langsung dapat dipergunakan tanpa proses editing kembali. Terdapat 2 tipe peringkasan teks yaitu abstraksi dan ekstraksi. Tipe abstraksi, meringkas teks dengan melibatkan parafrase dari teks. Sedangkan tipe ekstraksi, hanya menyalin informasi yang dianggap penting dari teks.

Pada Tugas Akhir ini, diimplementasikan sebuah metode ekstraksi yaitu *Maximal Marginal Importance* (MMI), dimana metode ini melakukan pembobotan nilai terhadap fitur kalimat pada teks, yang sebelumnya telah mengalami tahap *preprocessing*. Setelah tiap kalimat mempunyai bobot nilai, kalimat–kalimat tersebut dikelompokkan menggunakan algoritma *k-means clustering* berdasarkan nilai kemiripan (*similarity*) antar 2 kalimat. Dari tiap kluster kalimat yang didapatkan, dibentuklah sebuah /lebih pohon biner kalimat. Pembentukan pohon biner kalimat ini ditujukan untuk mengetahui pada level manakah kalimat–kalimat dari tiap kluster berada. Dimana level kalimat ini dilibatkan dalam perhitungan nilai MMI kalimat. Kalimat dengan nilai MMI tertinggi pada tiap kluster terpilih sebagai hasil ringkasan. Pada tahap akhir, dilakukan proses *ordering* terhadap kalimat ringkasan jika indeksnya belum terurut.

Evaluasi dilakukan menggunakan *ROUGE Evaluation Toolkit*. Dengan ROUGE, pengukuran akurasi hasil ringkasan (ringkasan kandidat) terhadap ringkasan referensi dilihat berdasarkan pada parameter *recall*, *precision*, dan *f-measure*. “Hasil evaluasi menunjukkan isi ringkasan kandidat mendekati isi ringkasan referensi berdasarkan parameter ROUGE”.

Kata kunci : peringkasan teks otomatis, ekstraksi, MMI, *similarity*, *k-means clustering*, ROUGE