

Abstrak

Berita elektronik merupakan media informasi yang paling populer dan interaktif saat ini. Begitu interaktifnya, hingga perkembangannya cukup pesat. Terbukti bertambah banyaknya situs perusahaan maupun situs personal, yang berarti semakin meningkatkan jumlah informasi dan data. Peningkatan yang pesat ini juga dipacu oleh penggunaan internet yang semakin berkembang dibandingkan era sebelumnya. Sebagai akibatnya, jumlah informasi meningkat secara eksponensial.

Banyaknya data yang ada, semestinya dapat memberikan manfaat yang banyak pula. *Summarization* merupakan salah satu proses menyaring informasi paling penting dari sebuah sumber atau beberapa sumber dokumen untuk membuat sebuah versi ringkas dari teks. Saat ini, kebanyakan metode *Summarization* seperti *TextRank*, *PageRank* dan lain sebagainya hanya mengandalkan perhitungan kesamaan berdasarkan kata dan tidak memperhatikan aspek lain, misalnya kesamaan frasa. Pada tugas akhir ini berusaha meringkas suatu dokumen dengan metode *Document Index Graph* yang dikombinasikan dengan *PageRank* yang merupakan metode untuk menentukan seberapa penting sebuah *vertex* dalam graf, berdasarkan informasi yang diambil dari struktur graf keseluruhan.

Metode ini diuji coba dengan menggunakan sampel berita berbahasa Indonesia dari media massa berbasis web. Pemilihan *threshold* yang tepat akan meningkatkan kualitas ringkasan. Pengujian berdasarkan *fmeasure* membuktikan bahwa ringkasan yang dihasilkan cukup bagus.

Kata kunci: *summarization, threshold, PageRank, vertex, graf*