

ABSTRAK

Membaca merupakan aktivitas tiap individu yang bertujuan untuk menangkap intisari dari suatu sumber bacaan. Semakin ringkas sumber bacaan maka proses perpindahan informasi dari penulis ke pembaca akan semakin cepat. Oleh karena itu dibutuhkan peringkasan teks otomatis sehingga dapat meminimalkan masalah biaya data dan waktu baca.

Penelitian ini bertujuan untuk mengembangkan metode peringkasan teks berita online berbahasa Indonesia. Dengan mengintegrasikan *genetic algorithm* (GA) untuk menangkap konten utama serta *Vector Space Model* (VSM) yang diterapkan untuk menangkap hubungan semantik antar kalimat yang didahului dengan tahap *text preprocessing* yang meliputi *sentence splitting*, *case folding*, *tokenizing*, *stopword removal* dan *stemming*. Sehingga dapat meningkatkan akurasi dari peringkasan teks tersebut.

Hasil penelitian menunjukkan bahwa dengan integrasi antara GA dan VSM dapat meningkatkan akurasi peringkasan teks dalam bahasa Indonesia. Dengan melakukan pembobotan GA pada metode fitur teks, terjadi peningkatan akurasi sebesar 4,8%. Peningkatan akurasi sebesar 2,9% setelah ditambahkan fitur hubungan semantik antar kalimat.

Kata kunci:

Peringkasan teks, fitur teks, *genetic algorithm*, *vector space model*, semantik