

Abstrak

Peringkasan teks (*text summarization*) adalah proses pengambilan informasi yang paling penting dari satu atau beberapa dokumen teks sehingga dihasilkan sebuah versi yang lebih ringkas dengan menggunakan aplikasi berbasis komputer.

Pada Tugas Akhir ini diimplementasikan metode *clustering* dan *nonnegative matrix factorization* (NMF) untuk meringkas multi dokumen berita. Ringkasan yang dihasilkan berupa ringkasan ekstraktif yang berisi kalimat-kalimat penting dari dokumen berita. NMF akan menguraikan kalimat menjadi kombinasi *semantic feature*. Hal ini dapat meningkatkan kualitas hasil ringkasan, karena dengan melihat kesamaan antara *semantic feature* kalimat dengan topik berita, akan diketahui kalimat yang penting dan benar-benar relevan dengan topik berita. Dalam metode peringkasan ini juga akan dilakukan proses *clustering* yang bertujuan untuk menghilangkan kalimat-kalimat *noise* dan meminimalkan *redundancy* pada ringkasan.

Pengevaluasian dilakukan menggunakan *ROUGE evaluation toolkit*. Hasil pengujian menunjukkan bahwa metode NMF dengan *clustering* menghasilkan ringkasan yang mempunyai nilai *precision* lebih tinggi dan *redundancy* lebih sedikit dibandingkan dengan ringkasan yang dihasilkan metode NMF tanpa *clustering*.

Kata kunci : peringkasan teks, *nonnegative matrix factorization*, *clustering*.