

Abstrak

Penelitian ini mencoba mengimplementasi algoritma Nazief dan Andriani, yaitu algoritma yang dikembangkan untuk stemming dokumen Bahasa Indonesia. Yaitu suatu proses pemotongan imbuhan-imbuhan suatu kata menjadi kata dasarnya yang merupakan suatu tahap penting dalam *Information Retrieval*.

Setelah diimplementasi, kata-kata yang gagal di-*stem* kemudian diteliti untuk penyusunan skema perbaikan algoritma tersebut.

Setelah dilakukan pengujian pada skema perbaikannya serta membandingkan dengan skema yang lama, terbukti bahwa skema tersebut dapat menghasilkan nilai akurasi yang lebih tinggi serta nilai indeks MWC dan ICF yang lebih baik. Namun, muncul permasalahan lain yaitu munculnya kasus *unstemmed word* yang tidak muncul sebelumnya, walau jumlahnya tidak signifikan jumlah kasus yang berhasil di tangani.

Kata kunci: Stemming, Information Retrieval, Bahasa Indonesia