

Abstrak

Sistem *Information Retrieval* adalah sistem yang digunakan untuk menemukan kembali informasi-informasi yang relevan terhadap kebutuhan pengguna dari suatu kumpulan informasi berdasarkan *query* dari pengguna. Suatu sistem *information retrieval* dikatakan ideal jika sistem tersebut dapat menemukan seluruh dokumen yang relevan saja. Akan tetapi, *term-term* yang terdapat di dokumen dan di *query* sering memiliki banyak varian morfologik, sehingga pasangan *term* seperti "compute", "computing" dan "computation" tidak akan dianggap ekuivalen sistem.

Teknik *stemming* diterapkan pada sistem *information retrieval* untuk mengubah varian morfologik dari *term-term* menjadi sebuah bentuk kata sehingga dianggap ekuivalen oleh sistem. Selain itu *stemming* juga mampu mengurangi ukuran *index file*. Pada tugas akhir ini dibuat sebuah sistem *information retrieval* yang mengimplementasikan teknik *stemming* menggunakan algoritma STANS.

Algoritma STANS merupakan salah satu teknik *affix removal stemmer* yang terbentuk dari modifikasi algoritma *porter's stemming*. Pada tugas ini dilakukan pula analisis pengaruh penerapan *stemming* menggunakan algoritma STANS. Hasil penelitian menunjukkan bahwa dengan *stemming* sistem mampu mereduksi *term* yang dihasilkan sehingga mampu mengurangi ukuran *index*. Dari sudut pandang performansi sistem, bisa dikatakan penerapan algoritma STANS lebih baik dibanding algoritma Porter dimana nilai *precision*, *iap* dan *non-iap* yang dihasilkan menunjukkan peningkatan tanpa mengurangi nilai *recall* yang dihasilkan.

Kata kunci: Sistem *information retrieval*, *Stemming*, Algoritma STANS