

Abstrak

Stemming dalam sistem *Information Retrieval* digunakan untuk membatasi varian bentuk kata yang berbeda menjadi bentuk dasarnya, sehingga nantinya dapat meningkatkan kemampuan sistem dalam menemukan dokumen relevan sesuai *query* yang ada. Sistem *Information retrieval* dikatakan ideal apabila dapat menemukan informasi yang relevan saja. *Term* yang ada pada *query* ataupun kumpulan informasi memiliki banyak varian morfologi sehingga *term-term* yang memiliki bentuk beda tidak akan dianggap ekivalen oleh sistem. Dalam tugas akhir ini, dibuat sebuah sistem *Information retrieval* yang mengimplementasikan teknik *stemming* dengan menggunakan algoritma Nazief & Adriani dan algoritma Ahmad Yusoff Sembok.

Algoritma Nazief & Adriani dan algoritma Ahmad Yusoff Sembok merupakan algoritma *stemming* untuk teks berbahasa Indonesia. Perbedaan kedua algoritma ini ada pada proses penghilangan imbuhan (*affixes*). Untuk Algoritma Nazief & Adriani, proses penghilangan imbuhan dilakukan pada akhiran (*suffixes*) terlebih dahulu. Sedangkan pada algoritma Ahmad Yusoff Sembok penghilangan imbuhan terlebih dahulu dilakukan pada awalan (*prefixes*).

Pada tugas akhir ini dilakukan analisis pengaruh penerapan *stemming* menggunakan algoritma Nazief&Adriani dan algoritma Ahmad Yusoff Sembok. 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 Nazief&Adriani lebih baik dibanding algoritma Ahmad Yusoff Sembok dimana nilai *recall* selalu lebih besar daripada nilai *precision*, dan nilai *recall* pada algoritma Nazief&Adriani lebih besar dari nilai *recall* pada algoritma Ahmad Yusoff Sembok.

Kata kunci : *Information Retrieval*, *Stemming*, Algoritma Nazief & Adriani, Algoritma Ahmad Yusoff Sembok