

Abstrak

Stemming merupakan proses pemotongan imbuhan seperti awalan, akhiran, sisipan, dan kombinasi yang dilakukan dengan menggunakan algoritma tertentu untuk mengembalikan suatu kata ke bentuk dasarnya. Stemming merupakan bagian dari *information retrieval* dan digunakan pada mesin pencari. Dengan menerapkan stemming pada mesin pencari, maka dapat diperoleh informasi yang lebih spesifik dari banyaknya informasi yang ada. Stemming merupakan inti dari teknik pemrosesan *natural language* untuk mendapatkan *information retrieval* yang efektif dan efisien.

Pada Bahasa Indonesia, suatu kata dapat digabungkan dengan awalan, akhiran, sisipan, ataupun kombinasi. Hal tersebut dapat menyebabkan sulitnya pencocokan kata yang berhubungan. Beberapa penggabungan dapat menyebabkan perubahan pada akar kata. Dalam pengolahan teks berbahasa Indonesia, dapat digunakan teknik stemming. *Stemming* dalam *information retrieval system* digunakan untuk membatasi varian bentuk kata yang berbeda menjadi bentuk dasarnya, sehingga nantinya akan dapat meningkatkan kemampuan sistem dalam menemukan dokumen relevan sesuai *query* yang ada. Dalam tugas akhir ini, akan mengimplementasikan teknik *stemming* dengan menggunakan algoritma *Confix-Stripping* pada *information retrieval system*.

Algoritma *Confix-Stripping* (CS) berdasarkan pada aturan morfologi bahasa Indonesia yang dikelompokkan menjadi satu dan di-enkapsulasi pada imbuhan, termasuk awalan, akhiran, sisipan, dan kombinasi. Algoritma *Confix-Stripping* (CS) menggunakan tiga komponen: kumpulan imbuhan, rule-rule, dan kamus. Pada tugas akhir ini akan dilakukan analisis penerapan algoritma *Confix-Stripping* (CS) tersebut pada *Information Retrieval system*.

Kata kunci : sistem temu kembali informasi, *stemming*, *Confix-Stripping*