ABSTRAK

Stemming adalah suatu teknik pencarian bentuk dasar dari suatu *term* pada *information retrieval*. Yang dimaksud dengan *term* itu sendiri adalah tiap kata yang berada pada suatu dokumen teks. *Stemming* dilakukan pada saat pembuatan indeks dari suatu dokumen. Teknik *stemming* diperlukan selain untuk memperkecil jumlah indeks yang berbeda dari suatu dokumen, juga untuk melakukan pengelompokan kata-kata lain yang memiliki kata dasar dan arti yang serupa namun memiliki bentuk atau *form* yang berbeda karena mendapatkan imbuhan yang berbeda. Teknik *stemming* terdiri dari berbagai macam metode. Metode pertama yakni *stemming* dengan acuan tabel pemenggalan imbuhan. Proses *stemming* suatu *term* dengan metode ini dilakukan dengan cara menghilangkan imbuhan dari *term* tersebut sesuai dengan table acuan pemenggalan imbuhan yang digunakan. Metode kedua merupakan pengembangan dari metode pertama. Metode kedua ini selain menggunakan tabel acuan pemenggalan imbuhan, juga menggunakan suatu kamus kata dasar. Kamus kata dasar ini digunakan sebagai acuan hasil *stemming* saat proses pemenggalan imbuhan selesai dilakukan. Hasil dari proses *stemming* dengan metode ini harus ada pada kamus kata dasar, jika tidak maka *term* yang diinputkan dianggap sebagai bentuk dasar.Salah satu teknik stemming adalah algoritma Idris. Algoritma Idris menggunakan teknik berbeda dalam menentukan hasil *stemming*, yaitu menggunakan 2 kamus : kamus umum dan kamus khusus (sesuai isi dokumen). Algoritma ini akan mengoreksi ulang hasil *stemming*, lalu mencocokkannya dengan kata-kata dalam kamus.

Kata kunci : stemming, algoritma Idris, term, information retrieval