

Abstrak

Stemming merupakan suatu proses untuk menemukan kata dasar dari sebuah kata dengan menghilangkan semua imbuhan. Berdasarkan asumsi bahwa *term* yang memiliki akar kata yang sama akan selalu memiliki makna yang sama, *stemming* digunakan pada *information retrieval* untuk meningkatkan keakuratan *retrieval* (perolehan informasi). Sehingga membuat penulis melakukan implementasi *stemming* teks berbahasa Indonesia pada *information retrieval*.

Porter stemmer merupakan salah satu algoritma *stemming* yang telah lama digunakan. *Porter Stemmer* pertama kali dipublikasikan pada tahun 1980. Selain itu, *Porter stemmer* merupakan *stemmer* ini tidak menggunakan kamus, karena berdasarkan kenyataannya resource seperti kamus digital yang besar sangatlah mahal

Pada tugas akhir ini penulis mengimplementasikan *Porter stemmer* dan melakukan modifikasi pada *Porter stemmer* bila ditemukan stem yang tidak terdapat pada kamus yang digunakan. Setelah modifikasi, dilakukan perbandingan nilai keakuratan dari stem yang dihasilkan oleh *Porter stemmer* dan gabungan porter stemmer(stemmer hasil modifikasi Porter stemmer). Hasil dari pengujian menunjukkan bahwa *stemming* yang menggunakan gabungan porter stemmer menghasilkan keakuratan stem yang lebih baik dibandingkan dengan menggunakan Porter stemmer. Sehingga dapat disimpulkan bahwa modifikasi yang dilakukan dapat meningkatkan keakuratan stem.

Pada tugas akhir ini dilakukan pula analisis pengaruh proses *stemming* terhadap relevansi dokumen yang diretrieve pada *Information Retrieval*. *Stemming* dilakukan dengan *Porter stemmer* dan gabungan porter stemmer. Hasil penelitian menunjukkan bahwa peningkatan keakuratan stem pada banyak kasus menurunkan nilai precision akan tetapi tidak akan berpengaruh terhadap nilai recall.

Kata Kunci: *Stemming, Information Retrieval, Precision, Recall*