

Abstrak

Semakin berkembangannya pengaksesan berita melalui internet membuat kemudahan pengaksesan tersebut menjadi sebuah hal yang sangat penting bagi masyarakat. Guna memudahkan proses pengaksesannya, maka perlu adanya suatu proses pengolahan berita agar menjadi lebih teratur dan efisien, sehingga pengaksesan untuk berita-berita lampau pun menjadi relatif lebih mudah. Salah satu proses pengolahan tersebut adalah klasifikasi. Sebelum melakukan klasifikasi, terlebih dahulu perlu adanya *preprocessing* data dimana salah satu tahap *preprocessing* adalah *stemming*.

Stemming adalah proses untuk memecah setiap varian-varian suatu kata menjadi kata dasar. Salah satu metode *stemming* yang ada adalah *successor variety*. *Successor variety* adalah metode *stemming* yang berdasarkan pada batas-batas morfem, yang menggunakan pengetahuan dari *structural linguistic*. Ada beberapa metode *successor variety*, beberapa diantaranya adalah *cutoff* dan *entropy method*. Kedua metode *stemming* tersebut diterapkan pada sebuah sistem yang kemudian dilihat bagaimana performansinya.

Hasil dari implementasi kedua metode tersebut adalah perbandingan performansi kedua metode yang diterapkan berdasarkan data uji yang digunakan, metode *entropy* lebih baik dari pada metode *cutoff* dari segi akurasi, sedangkan dari segi ICF, ternyata *cutoff* yang lebih baik dari *entropy*. Sedangkan dari sisi klasifikasi, kedua metode tersebut memiliki performansi yang setara.

Kata kunci: klasifikasi, *successor variety stemming*, *cutoff*, *entropy*