

BAB V PENUTUP

5.1. Kesimpulan

Berdasarkan penelitian implementasi nazief & adriani menggunakan n-gram dan *sorensen dice coefficient* dalam typography kata bahasa indonesia, maka dapat diambil kesimpulan sebagai berikut, yaitu:

- a. Metode *N-Gram* dan *SDC* dapat memperbaiki kesalahan *stemming* dan *typo* kata pada Nazief-Adriani.
- b. Berdasarkan hasil uji sebanyak 3 kali pengujian didapatkan relevansi,
 1. Pengujian pertama dilakukan pada 10 kata didapatkan relevansi,
 - . *N3-Gram* 80%
 - . *N4-Gram* 70%
 - . *N5-Gram* 30%
 2. Pengujian kedua dilakukan pada 20 kata didapatkan relevansi,
 - . *N3-Gram* 75%
 - . *N4-Gram* 50%
 - . *N5-Gram* 15%
 3. Pengujian ketiga dilakukan pada 30 kata didapatkan relevansi,
 - . *N3-Gram* 76%
 - . *N4-Gram* 56%
 - . *N5-Gram* 26%
- c. Kesalahan pada *N-gram* dan *sorensen dice similarity* dipengaruhi oleh urutan huruf pada setiap kata dan juga kata pada urutan pertama bergantung pada urutan huruf abjad.
- d. Penelitian ini dipengaruhi oleh korpus kata dasar apabila kata yang dibandingkan tidak ada kesamaan maka kata pertama yang ditampilkan adalah kata yang mempunyai nilai relevansi terhadap kata yang di bandingkan.

5.2. Saran

Berdasarkan kesimpulan yang didapatkan dari hasil analisis, maka penulis mencoba memberikan saran sebagai berikut :

- a. Pada penelitian selanjutnya dapat ditambahkan algoritma yang dapat menstemming kata berulang.
- b. Metode yang digunakan adalah *N-Gram* dan *Sorensen Dice Coefficient* untuk memperbaiki pemenggalan kata pada Nazief & Adriani *Stemmer*, untuk penelitian selanjutnya mungkin bisa digunakan metode-metode yang lain untuk melihat dan membandingkan dengan hasil penelitian sebelumnya.

