

BAB V

SIMPULAN DAN REKOMENDASI

5.2 Simpulan

Berdasarkan pada penelitian “Penerapan Deteksi Kata Majemuk pada *Information Retrieval* dalam *Knowledge Management System*” dan mengacu pada rumusan masalah yang telah disampaikan pada BAB1, dapat disimpulkan bahwa:

1. Hasil retrieval menunjukkan bahwa sistem dengan deteksi kata majemuk dapat melakukan penelusuran lebih baik. Terbukti dengan ditemukannya hampir semua dokumen yang relevan pada titik retrieval pertama pada pencarian menggunakan kata majemuk. Hal ini dipengaruhi oleh deteksi kata majemuk yang diterapkan dapat mengoptimalkan *term/token* sehingga pencarian pada dokumen dalam matriks *termdocument* dapat menghasilkan hasil yang lebih sesuai dengan *query*.
2. kualitas index yang dibangun sangat mempengaruhi performansi penelusuran. Dengan digunakannya deteksi kata majemuk index akan bertambah namun kualitasnya menjadi lebih baik. Index yang baik akan memberikan hasil penelusuran yang baik.
3. Berdasarkan hasil evaluasi nilai mean average precision mempengaruhi frekuensi kemunculan dokumen relevan yang muncul untuk setiap retrieval. Semakin tinggi nilai MAP, semakin baik dan cepat sistem menampilkan dokumen yang relevan. Nilai MAP tertinggi diperoleh pada scenario nilai r yang sama yaitu 60, dengan nilai sebesar 0.689411 atau 68.94% pada sistem dengan deteksi kata majemuk dan 0.530522 atau 53.22% pada sistem tanpa deteksi kata majemuk.

5.2 Rekomendasi

Berikut merupakan saran-saran pada penelitian ini untuk pengembangan lebih lanjut :

1. Untuk meningkatkan kualitas *index* yang digunakan, dapat dilakukan penelitian lanjutan terkait pemrosesan teks dan pembagunan *index*, diantaranya :
 - a. Dapat meningkatkan kinerja deteksi kata majemuk dengan menambah kemampuan untuk mendeteksi kata majemuk senyawa atau kata majemuk yang sudah melebur menjadi satu.
 - b. Pada proses *Part-Of-Speech Tagging* menggunakan POST untuk Bahasa Indonesia dapat meningkatkan
 - c. Perlu dilakukan pengembangan aplikasi dengan menggunakan Bahasa Inggris mengingat banyaknya istilah Bahasa Inggris yang digunakan dalam ilmu komputer diantaranya menambah daftar *Stopwords*, *stemming* dan aturan pada deteksi kata majemuk
 - d. Perlu dilakukan pengujian dengan menggunakan token tanpa melalui proses *stemming*, karena struktur bahasa Indonesia berbeda dengan bahasa asing lainnya, sehingga ada kemungkinan proses pembuatan *index* dengan menggunakan token yang tidak di-*stemming* menghasilkan presisi yang lebih baik.
 - e. Perlu dilakukan pengembangan menentukan *threshold* dari nilai *similarity* pada dokumen sesuai dengan *query* yang dimasukkan untuk menghindari munculnya dokumen acak jika *query* tidak ditemukan.