

BAB 5

KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari penelitian yang dilakukan pada tugas akhir ini. Selain itu, bab ini juga berisi saran-saran untuk pengembangan penelitian lebih lanjut.

5.1 Kesimpulan

Pada tugas akhir ini telah dilakukan penelitian mengenai pencarian opini pada dokumen blog berbahasa Indonesia. Eksperimen dilakukan dengan mencoba berbagai teknik pemrosesan kueri, perbaikan kueri, serta pengurutan kembali dokumen berdasarkan opini. Evaluasi dengan menggunakan nilai MAP kemudian dilakukan untuk membandingkan akurasi dari masing-masing eksperimen yang dilakukan.

Dari keseluruhan eksperimen yang dilakukan, ada beberapa hal yang dapat disimpulkan dari penelitian pada tugas akhir ini, yaitu:

1. Dari segi pemrosesan kueri, kueri yang memberikan hasil perolehan yang terbaik adalah kueri dengan menggunakan pencarian berdasarkan frase dan pembobotan *field body* pada dokumen yang dijalankan pada pengindeksan yang menggunakan pemotong kata berimbunan.
2. Dari segi perbaikan kueri, jumlah dokumen dan kata yang digunakan untuk *pseudo-relevance feedback* sebaiknya tidak terlalu banyak dan tidak terlalu sedikit. Pada penelitian perolehan dokumen blog ini, jumlah dokumen yang tepat digunakan adalah 50 dokumen dan jumlah kata yang tepat digunakan adalah 10 kata.
3. Dari segi pengurutan kembali dokumen berdasarkan opini, teknik pengurutan dengan menggunakan mesin pencari Indri dan pengurutan kembali dari 50 dokumen teratas yang terbaik ternyata memberikan hasil yang hampir sama. Hasil ini ternyata tidak begitu berpengaruh terhadap nilai akurasi perolehan dokumen blog. Untuk eksperimen pengurutan kembali dengan menambahkan

nilai *similarity* dokumen berdasarkan kata-kata opini, hasil dengan normalisasi jauh lebih baik daripada hasil tanpa normalisasi.

5.2 Saran

Penulis menyadari masih ada banyak kekurangan dalam pengerjaan tugas akhir ini. Berikut ini adalah saran-saran dari penulis untuk penelitian selanjutnya:

1. Penambahan jumlah koleksi blog berbahasa Indonesia untuk pengujian efektivitas dan efisiensi dari metode perolehan dokumen blog.
2. Pengelompokkan korpus blog berupa *feed* dan *permalinks* untuk dibandingkan hasil akurasi perolehannya.
3. Penggunaan teknik *machine learning*, khususnya dalam hal pendeteksian dokumen blog yang berupa opini agar bisa dibandingkan hasilnya dengan teknik dengan pendekatan statistik.
4. Pendeteksian blog berupa *spam*, untuk dapat menghindari perolehan dokumen blog yang berupa *spam*.