

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Adapun kesimpulan akhir dari penelitian pengelompokan dokumen elektronik berbahasa Indonesia dengan menggunakan metode *Bisecting K-Means* yaitu:

1. Sistem pengelompokan dokumen dengan menggunakan metode *Bisecting K-Means* dapat mengelompokkan dokumen-dokumen pada suatu *root directory* secara otomatis.
2. Berdasarkan referensi dari penelitian *document clustering* menggunakan metode *Bisecting K-Means* yang dilakukan Steinbach *et al* (2000: 14), didapatkan nilai *entropy* sebesar 0,2940 – 1,1872. Penulis menyimpulkan bahwa nilai *entropy* untuk *dataset IlmuKomputer.org* sebesar 0,7268 cukup baik, dan sangat baik untuk *dataset file.upi.edu* dengan nilai *entropy* 0,0719.
3. Berdasarkan referensi dari penelitian *document clustering* menggunakan metode *Bisecting K-Means* yang dilakukan Steinbach *et al* (2000: 15), didapatkan nilai *overall similarity* sebesar 0,2228 – 0,5502. Penulis menyimpulkan bahwa nilai *overall similarity* untuk *dataset IlmuKomputer.org* sebesar 0,2594 dan *dataset file.upi.edu* sebesar 0,3546 cukup baik.

4. Dari nilai *purity* yang didapatkan, pengelompokan *dataset IlmuKomputer.org* hanya 65% yang sesuai dengan pengelompokan manual, dan 97% untuk *dataset file.upi.edu*.
5. *Bisecting K-Means* kurang baik dalam mengelompokkan dokumen dengan dokumen-dokumen yang memiliki keterkaitan dengan kelompok-kelompok lainnya.
6. Kompleksitas waktu dari metode *Bisecting K-Means* berbanding lurus dengan banyaknya ragam kata pada dokumen-dokumen suatu *dataset*. Semakin banyak ragam kata pada suatu *dataset* maka semakin lama pula waktu yang diperlukan untuk mengelompokkan dokumen.

5.2 Saran

Untuk pengembangan lebih lanjut, saran-saran yang diberikan pada penelitian ini yaitu sebagai berikut:

1. Perlu dilakukan penelitian lebih lanjut dengan menggunakan jumlah dokumen elektronik yang lebih banyak dan lebih beragam untuk mengetahui kemampuan metode *Bisecting K-Means* dalam melakukan pengelompokan dokumen.
2. Perlu dikembangkan kembali performansi sistem pengelompokan dokumen dengan *Bisecting K-Means* untuk memperbaiki kompleksitas waktu yang diperlukan dalam melakukan proses pengelompokan dokumen.
3. Perlu dikembangkan kembali sistem yang dapat melakukan ekstraksi teks pada dokumen yang terenkripsi.