BAB VI

PENUTUP

6.1 Kesimpulan

Kesimpulan yang dapat diambil dari penelitian tugas akhir ini adalah sebagai berikut :

a. *Fingerprint biword winnowing* dapat digunakan sebagai *feature*/ciri untuk mengelompokkan dokumen teks.
b. *K-Means* adalah metode pengelompokkan yang dapat mengelompokan dokumen dengan cepat, akan tetapi sangat sensitif terhadap nilai *centroid* awal dan tidak dapat menentukan jumlah kelompok yang tepat bagi suatu koleksi dokumen.
c. Pembentukan nilai *centroid* awal dengan cara memilih secara random dari dokumen yang ada memberikan hasil pengelompokkan yang lebih baik, yakni nilai *precision* yang mencapai 88.33% untuk masukan jumlah kelompok 7 dan jumlah dokumen 20. Sedangkan pembentukan nilai *centroid* awal secara random berdasarkan range nilai dimensi dokumen hanya menghasilkan nilai precision mencapai 76.67% untuk masukan jumlah kelompok 7 dan jumlah dokumen 20.

6.2 Saran

Saran yang dapat penulis berikan untuk pengembangan apllikasi *clustering* dokumen teks yang lebih baik di masa yang akan datang adalah :

a. Mengkombinasikan metode *K-means* dengan metode *clustering* lainnya, karena *k-means* memiliki kelebihan yaitu kemudahan dalam pengimplementasian dan dapat bekerja dengan cepat. Akan tetapi tidak bisa menentukan jumlah kelompok yang tepat bagi suatu koleksi dokumen. Dengan cara mengkombinasikan metode *k-means* dengan metode *clustering* lain, diharapkan dapat menghasilkan pengelompokkan dokumen yang lebih sempurna.

b. Pada kasus *clustering*, akan lebih baik jika pembentukan nilai *centroid* awal dilakukan dengan cara memilih secara acak dari dokumen yang telah ada.