

## BAB V

### PENUTUP

#### 5.1 Kesimpulan

Kesimpulan dari penelitian yang sudah dilakukan dengan judul “Implementasi Clustering Untuk Pengelompokan Karya Ilmiah Bertema COVID-19 Berdasarkan Kemiripan Abstrak Menggunakan K-Means Clustering” adalah sebagai berikut:

1. Hasil *clustering* terbaik dihasilkan dari proses *clustering* menggunakan fitur yang dihasilkan dari metode pengurangan dimensi LDA dengan metode pengukuran jarak *cosine distance* dengan menggunakan *cluster* sejumlah 15 dengan nilai *silhouette coefficient* sebesar 0.5336, dimana hasil *cluster* ini sudah terbentuk dengan layak.
2. Hasil analisis menggunakan *silhouette plot* menyimpulkan bahwa proses *clustering* yang dihasilkan sudah cukup baik, namun analisis dokumen berdasarkan *WordCloud* menyimpulkan bahwa hasil *clustering* yang dilakukan masih tidak sempurna, dikarenakan terdapat abstrak karya ilmiah yang tidak tergolongkan dengan baik.

#### 5.2 Saran

Berdasarkan penelitian yang dilakukan, terdapat beberapa saran yang dapat meningkatkan kecepatan dan hasil dari proses *clustering*, antara lain:

1. Diperlukan metode *pre-processing* data yang lebih baik sehingga abstrak karya ilmiah yang terdapat campuran Bahasa Indonesia dan Bahasa Inggris dapat dihilangkan.
2. Diperlukan dokumen yang lebih terformat dengan baik sehingga tidak ada kata-kata yang tergabung tanpa spasi.

