

## BAB V

### PENUTUP

#### 5.1 Hasil dan Kesimpulan

Berdasarkan hasil yang diperoleh dan dijelaskan pada bab sebelumnya, maka dapat disimpulkan,

1. Algoritma K-Means mampu mengelompokkan data berdasarkan kemiripan kata dalam judul secara otomatis tanpa label (*unsupervised learning*), sementara proses *text mining* berperan penting dalam meningkatkan kualitas struktur data sehingga pola kemiripan antar judul dapat dikenali dengan lebih baik. Secara keseluruhan, *model* yang dihasilkan dapat membantu pengorganisasian koleksi perpustakaan secara lebih sistematis dan mendukung pengelompokan berbasis topik.
2. Cara kerja *text mining* dalam proses pengenalan teks judul buku pada Perpustakaan UPN Veteran Jakarta dilakukan melalui tahapan *preprocessing text mining* yang meliputi *Case Folding*, *Tokenizing*, *Stopword Removal*, *Stemming* dan pembobotan TF-IDF untuk mengubah data teks tidak terstruktur menjadi representasi numerik. Penentuan jumlah *cluster* optimal dilakukan menggunakan metode *Elbow* serta evaluasi dari hasil *David-Bouldin Index*.

Berdasarkan hasil penelitian yang telah dilakukan, hasil penelitian menunjukkan bahwa  $k$  optimal berada di  $k = 5$  dengan nilai DBI sebesar 2.103. Pada tahap akhir pratinjau penamaan cluster, meskipun *Cluster 1* dan *Cluster 3* sama-sama berada dalam domain hukum, proses *clustering* dilakukan berdasarkan representasi numerik TF-IDF dan pengukuran kemiripan antar dokumen dimana pada TF-IDF, setiap dokumen direpresentasikan sebagai vektor bobot kata.

#### 5.2 Saran

Untuk meningkatkan kualitas pengelompokan pada penelitian selanjutnya, disarankan untuk melakukan penyempurnaan pada tahap klastering dengan

menerapkan teknik *word embedding* seperti Word2Vec atau BERT guna menangkap makna semantik yang lebih mendalam dibandingkan TF-IDF. Selain itu, penggunaan metode evaluasi tambahan seperti *Silhouette Coefficient* dan pengujian algoritma alternatif seperti DBSCAN perlu dipertimbangkan untuk memvalidasi stabilitas jumlah klaster yang dihasilkan serta meminimalisir tumpang tindih fitur antar kelompok.