

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan pembahasan pada bab-bab sebelumnya maka pada bab terakhir ini penulis dapat mengambil kesimpulan diantaranya:

1. Aplikasi ini dapat mengetahui jenis bisnis yang sedang populer dibahas pada portal berita dengan mengimplementasikan text preprocessing, normalisasi pada hasil TF.IDF, *Word Cloud*, algoritma K-Means serta visualisasi teks dengan *Word Cloud* dan di uji dengan silh *Silhouette Coefficient*.
2. Proses text preprocessing berpengaruh tinggi pada hasil penelitian karena data yang diproses pada rumus tidak boleh berupa kata melainkan angka dari TF.IDF untuk mendapatkan bobot pada setiap data berita.
3. Normalisasi berguna untuk mengantisipasi jarak data terlalu tinggi terhadap data lain. Sehingga meningkatkan kualitas perhitungan jarak data terhadap *cluster*.
4. *Word Cloud* dapat mempresentasikan kata yang sering muncul pada setiap *cluster* sehingga dapat diketahui jenis kata terpopuler yang dibahas pada analisis setiap bulannya.
5. Nilai *Silhouette Coefficient* dihitung dari jarak dokumen terhadap dokumen satu *cluster* dan jarak dokumen terhadap dokumen *cluster* lain,

sehingga didapatkan hasil interpretasi kekuatan dokumen terhadap *cluster* tersebut. *Cluster* yang mempunyai anggota banyak belum tentu hasil pengujiannya lebih bagus daripada *cluster* yang memiliki anggota sedikit.

6. Semakin banyak bobot data yang relevan meningkatkan kualitas *cluster* dan semakin banyak kata yang sama akan memperjelas *Word Cloud* pada analisis yang dilakukan. Semakin banyak bobot data yang di analisa dapat mempengaruhi kualitas *cluster* itu sendiri.
7. Aplikasi ini dapat menghitung presentase interpretasi setiap anggota *cluster* melalui pengujian *Silhouette Coefficient* yang di kategorikan menurut hasil perhitungannya.

5.2 Saran

1. Aplikasi ini dapat dikembangkan lebih lanjut tidak hanya membuat *cluster* pada data berita yang sama melainkan dapat mengetahui klasifikasi dokumen sehingga mengetahui positif dan negatifnya.
2. Disarankan untuk mencoba algoritma *clustering* lain guna perbandingan.