

# Spesifikasi Tugas Besar

IF5172 – Data Warehouse & Data Mining

# Text Mining

1. Dengan menggunakan dataset yang telah diberikan, lakukan pengelompokan dokumen (document clustering) dengan menggunakan
  - a) Dengan bag of word & dengan MFS sbg fitur dokumen
  - b) Untuk masing2 fitur teks, implementasikan dengan 3 teknik klustering
2. Dari hasil (1), lakukan perbandingan tiap pendekatan (teknik), baik dari sisi waktu proses maupun dari sisi akurasi. Catatan : jelaskan metode akurasi yang dipilih.
3. Dari hasil (2), jelaskan mengapa waktu proses dan akurasi tiap teknik berbeda. Identifikasi faktor utama dari kinerja tersebut.
4. Tuliskan kesimpulan dari proses (1), (2), dan (3)

# Graph Mining

1. Dengan menggunakan dataset yang sudah diberikan, lakukan **graph clustering** untuk mendeteksi 'komunitas' dengan menggunakan 3 teknik yg berbeda,
2. Dari hasil (1), jelaskan perbedaan 'komunitas' (jika ada) hasil 3 teknik yg diimplementasikan. Maknai perbedaan tersebut dari sudut pandang domainnya.

- Tugas dikerjakan secara berkelompok sebanyak 2-3 orang,
- Laporan Tugas (dalam format pdf) dikumpulkan secara kolektif melalui ketua kelas dan dikirim ke semua dosen kelas pada tanggal 30 Desember 2016, maksimal pukul 22.00

Catatan: format laporan akan disusulkan.