

Abstrak

Berita *online* saat ini merupakan sesuatu yang sangat umum di kalangan masyarakat Indonesia bahkan dunia. Saat ini data berita *online* yang telah tersimpan pada suatu penyimpanan mencapai ratusan miliar berita. Untuk memudahkan proses pencarian, manipulasi atau pengolahan data diperlukan suatu permodelan data dan model yang cocok untuk data berita adalah model *graph*. Untuk memudahkan pembaca maka berita *online* tersebut perlu dikelompokkan berdasarkan keterkaitan isi beritanya. Salah satu metode yang bisa dimanfaatkan untuk mengelompokkan berita adalah dengan *graph clustering*.

Sebelum melakukan *graph clustering*, data berita *online* terlebih dahulu harus diubah menjadi model *graph*. Langkah pertama yang harus dilakukan untuk mengubah data berita ke bentuk *graph* adalah melakukan *preprocessing* terhadap data berita, lalu dihitung bobot keterkaitan beritanya dengan memanfaatkan *cosine similarity*, setelah itu bobot hasil penghitungan *cosine similarity* dinormalisasi untuk dijadikan *edge* yang menghubungkan dokumen berita. Setelah berbentuk *graph*, barulah dilakukan proses *graph clustering*. Dalam tugas akhir ini algoritma *graph clustering* yang digunakan adalah *Chinese Whispers*, karena algoritma *Chinese Whispers* mampu membentuk *cluster* dari data *graph* yang besar dengan waktu yang relatif cepat, sehingga sangat cocok digunakan untuk kasus *clustering* berita *online*.

Pada penelitian ini telah diuji performansi algoritma *Chinese Whispers* dari segi kualitas serta tingkat akurasi *cluster* yang dihasilkan. Setelah dilakukan pengujian diperoleh bahwa kualitas hasil *cluster Chinese Whisper* cukup baik karena hampir 95% *node* hasil *cluster* sudah memiliki nilai *intra-cluster* yang lebih tinggi dari pada *inter-cluster*-nya, sedangkan rata-rata akurasi dari proses *clustering* menggunakan algoritma *Chinese Whispers* adalah 80.0 %.

Kata kunci : *Graph, Graph Database, Clustering, Graph Clustering, Chinese Whispers*