

BAB I

PENDAHULUAN

1.1 Latar Belakang

Ada tiga karakteristik penting dalam ilmu sosial komputasi yaitu keterhubungan (*connectivity*) yang membentuk keterhubungan antar pengguna dalam suatu *cluster*, kerjasama (*collaboration*) yang memodelkan bagaimana mereka berinteraksi dan komunitas (*community*) yaitu pengelompokan atau klasterisasi berdasarkan kemiripan dan hubungan kedekatan antar pengguna.[1] Klasterisasi pada jejaring sosial merupakan mengelompokkan *vertex* pada *graph* jejaring sosial yang memiliki hubungan kedekatan antar *vertex*. Pada data berbasis *graph* jejaring sosial *twitter*, *vertex* didefinisikan sebagai *username* sedangkan *edge* didefinisikan sebagai hubungan yang terjadi antar *username*. Hubungan tersebut bisa berupa interaksi *mention* maupun *reply*. *Vertex* yang saling berdekatan (bertetangga) biasanya merupakan *vertex* yang sering berinteraksi satu sama lain sehingga nantinya akan membentuk suatu *sub-graph* pada *graph* jejaring sosial dan didefinisikan sebagai suatu *cluster*

Data yang digunakan yaitu data *twitter* yang terdiri atas *username* (*vertex*) dan hubungan antar *username* (*edge*). Suatu *vertex* dinyatakan terhubung dengan *vertex* lainnya apabila ada interaksi satu sama lain baik berupa *mention* atau *reply*. Nantinya data *twitter* yang berbentuk *graph* dikonversi ke bentuk *adjacency* matriks terlebih dahulu kemudian diubah ke bentuk matriks *geodesic distance* agar dimensi matriksnya bisa direduksi menggunakan metode *Multidimensional Scaling* (MDS). Kemudian hasil dari reduksi dimensi MDS akan menghasilkan sebuah matriks yang kemudian akan di *cluster* menggunakan metode *Clustering K – Medoids*. Pada metode *K – Medoids* kualitas sebuah *cluster* dapat diukur dengan *purity*. *Purity* digunakan sebagai ukuran kemurnian *cluster* dari algoritma klasterisasi. Tiap data yang ada di *cluster* akan dibandingkan dengan *cluster* yang asli apakah data tersebut masuk dalam *cluster* yang sesuai. Kualitas *cluster* dikategorikan baik jika nilai *purity* mendekati 1 dan dikategorikan buruk jika nilai *purity* mendekati 0.

1.2 Perumusan Masalah

Berdasarkan latar belakang masalah yang telah diungkapkan sebelumnya, maka permasalahan pokok yang dibahas dalam tugas akhir ini adalah :

1. Bagaimana proses klasterisasi pada *graph* jejaring sosial dengan menerapkan metode *K-Medoids*?
2. Berapa persen akurasi klasterisasi *graph* jejaring sosial menggunakan metode *K-Medoids*?

1.3 Tujuan dan Manfaat

1.3.1 Tujuan

Adapun tujuan yang ingin dicapai dari penelitian ini antara lain :

1. Mengklasterisasi *graph* jejaring sosial menggunakan metode *K-Medoids*
2. Mengetahui akurasi penggunaan metode *K-Medoids* dalam klasterisasi *graph* jejaring sosial

1.3.2 Manfaat

Hasil dari penelitian ini dapat digunakan sebagai referensi penelitian-penelitian selanjutnya dengan metode-metode yang berbeda serta data *graph* yang berbeda, tidak hanya menggunakan *graph* dalam bentuk data jejaring sosial saja.

1.4 Batasan Masalah

Adapun batasan masalah dari penelitian ini adalah sebagai berikut :

1. Metode yang digunakan dalam klasterisasi menggunakan *Multidimensional Scaling (MDS)* dan *K – Medoids*
2. Metode yang digunakan untuk pengujian kualitas klasterisasi adalah *purity*
3. Data yang digunakan adalah data twitter
4. Pengambilan data menggunakan *software NodeXL*
5. Maksimal 500 data interaksi mention dan reply antar username.
6. Data yang diposting di twitter antara Februari 2016 – Juni 2016
7. Bahasa pemrograman yang digunakan adalah *Python 2.7* dan *Tkinter*

1.5 Metodologi Pengerjaan Tugas Akhir

Pembuatan tugas akhir ini dilakukan dengan mengikuti metodologi sebagai berikut :

1.5.1 Studi Pustaka

Mencari dan mempelajari berbagai macam literatur dari *paper*, *website*, buku, atau artikel yang terkait mengenai deteksi *cluster*, *K – Medoids*, dan *purity*

1.5.2 Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data yang akan digunakan, data tersebut diperoleh dari jejaring sosial *Twitter* menggunakan *software NodeXL*

1.5.3 Perancangan Aplikasi

Pada tahap ini akan dilakukan perancangan aplikasi yang akan dibangun, meliputi perancangan arsitektur sistem, perancangan data serta antarmuka.

1.5.4 Pembuatan Aplikasi

Dalam tahapan ini dilakukan pembuatan aplikasi menggunakan bahasa pemrograman *Python 2.7* dan *Tkinter*

1.5.5 Penyusunan Laporan

Tahap ini merupakan tahap terakhir yaitu penyusunan laporan dari hasil implementasi aplikasi yang telah dibuat dan melakukan penarikan kesimpulan.

1.6 Sistematika Penulisan

Sistematika penulisan tugas akhir terdiri dari beberapa bab yang akan di jelaskan pada setiap babnya sebagai berikut :

BAB I PENDAHULUAN

Bab Pendahuluan terdiri dari berbagai subbab yaitu latar belakang, perumusan masalah, tujuan dan manfaat, batasan masalah, metodologi pengerjaan tugas akhir dan sistematika penulisan.

BAB II STUDI PUSTAKA

Bab studi pustaka berisi mengenai teori-teori yang berhubungan dengan pokok bahasan Tugas Akhir. Terdiri atas tinjauan pustaka yang menjelaskan penelitian sebelumnya dan landasan teori yakni penjelasan mengenai teori yang sudah ada.

BAB III PERANCANGAN SISTEM

Berisi perancangan sistem atau alur sistem yang akan dibuat berupa *flowchart* ataupun dalam bentuk diagram yang lain.

BAB IV IMPLEMENTASI SISTEM

Pada bab ini berisi implementasi dari pembahasan-pembahasan bab-bab sebelumnya, menerapkan rancangan sistem yang telah dibuat.

BAB V PENUTUP

Kesimpulan dari hasil ujicoba sistem yang telah dibuat dan saran untuk pengembangan selanjutnya.