

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi telah mengakibatkan meningkatnya data dalam jumlah besar. Data berukuran besar yang sudah disimpan jarang digunakan secara optimal karena manusia seringkali tidak memiliki waktu dan kemampuan yang cukup untuk mengelolanya. Data bervolume besar seperti data teks, jauh melampaui kapasitas pengolahan manusia yang sangat terbatas [1].

Kasus yang disoroti adalah data tugas akhir mahasiswa UTM khususnya jurusan teknik informatika. Setiap tahun fakultas dan perpustakaan menerima dan menyimpan data penelitian tugas akhir mahasiswa. Namun data tersebut dibiarkan begitu saja tanpa adanya pengolahan atau tindak lanjut. Padahal data tersebut dapat dijadikan pembelajaran bagi pihak lainnya khususnya mahasiswa yang melakukan penelitian terkait.

Data mining sangat sesuai untuk diterapkan pada data berukuran besar. Penerapan data mining pada data tugas akhir mahasiswa UTM diharapkan bisa menambang ilmu pengetahuan dan informasi yang penting dan berguna untuk pengambilan keputusan di masa depan. Metode data mining yang akan diterapkan dalam penelitian ini adalah clustering atau pengelompokan data [2].

Pengelompokan artikel atau teks berbahasa Indonesia merupakan salah satu solusi yang dapat digunakan untuk mempermudah mencerna informasi penting yang ada di dalamnya. Clustering dapat digunakan untuk membantu menganalisis berita dengan mengelompokkan secara otomatis berita yang memiliki kesamaan. Pada *text clustering* terdapat suatu permasalahan yaitu adanya fitur-fitur yang berdimensi tinggi sehingga menyebabkan waktu komputasi menjadi besar. Oleh karena itu dalam tugas akhir ini digunakan metode *Self-Organizing Maps (SOM)* dengan menerapkan pembatasan wilayah pencarian dan reduksi dimensi sehingga diharapkan dapat mengurangi waktu komputasi dalam proses *clustering* [3].

1.2 Perumusan Masalah

Dari latar belakang yang sudah dijelaskan di atas, maka terdapat rumusan masalah sebagai berikut:

- a. Bagaimana menerapkan metode *Self-Organizing Maps (SOM)* dalam pengelompokan data teks berskala besar?
- b. Bagaimana mengurangi waktu komputasi SOM ketika menangani data teks berskala besar?

1.3 Tujuan

Pada tugas akhir ini akan dilakukan penelitian yang bertujuan untuk:

- a. Menerapkan metode *Self-Organizing Maps* untuk mengelompokan data teks berskala besar.
- b. Mengimplementasikan metode *Efficient Self-Organizing Maps (SOM)* dengan menerapkan reduksi dimensi sehingga dapat mengurangi waktu komputasi.

1.4 Batasan Masalah

Batasan Masalah dalam pengerjaan tugas akhir ini adalah:

- a. Data yang digunakan adalah tugas akhir mahasiswa UTM jurusan teknik informatika dari berbagai bidang minat tahun 2005-2012.
- b. Data inputan adalah ekstraksi dari abstrak dan laporan tugas akhir.
- c. Jumlah data yang digunakan adalah 140 abstrak dan 7 laporan tugas akhir mahasiswa.

1.5 Metodologi Penelitian

Pembuatan tugas akhir ini terbagi menjadi beberapa tahap pengerjaan yang tertera sebagai berikut:

1. Studi Literatur dan Identifikasi Permasalahan
Tahap ini dilakukan dengan mempelajari konsep-konsep yang berkaitan dengan tugas akhir ini untuk memahami metode analisa kebutuhan, perancangan sistem dan pendukung lainnya melalui

literature-literatur seperti buku, paper, sumber ilmiah lain seperti *internet* ataupun artikel dokumen teks lainnya yang berhubungan.

2. Analisa dan Perancangan sistem

Dalam tahap ini merupakan tahapan pertama dari pengembangan sistem. Pada tahap ini mendefinisikan permasalahan yang ada pada penelitian dan menentukan tujuan perancangan sebagai acuan untuk menentukan langkah-langkah selanjutnya. Perancangan sistem dan deskripsi sistem berdasarkan data hasil observasi menggunakan *flowchart*, diagram dan lain-lain.

3. Pembuatan Program

Tahap ini merupakan tahap pembuatan dan pengembangan program sesuai dengan perancangan sistem yang ditetapkan pada tahap sebelumnya.

4. Uji Coba dan Evaluasi

Melakukan ujicoba terhadap sistem yang sudah dibuat dengan scenario-skenario yang sudah ditentukan. Dari percobaan tersebut dapat diketahui scenario mana yang menjadi scenario terbaik untuk menentukan optimum cluster.

5. Dokumentasi

Pada Tahap ini dilakukan dokumentasi mengenai apa yang diperoleh dan dikerjakan mulai dari kegiatan studi literatur dan identifikasi permasalahan hingga menghasilkan program baru.

1.6 Sistematika Penulisan

Adapun sistematika pembahasan tugas akhir ini adalah sebagai berikut:

BAB I: PENDAHULUAN

Pada bab ini terdiri dari: latar belakang, permusan masalah, batasan masalah, tujuan penelitian dan sistematika penulisan.

BAB II: STUDI PUSTAKA

Pada bab ini dijelaskan mengenai hal-hal yang melandasi atau yang akan digunakan sebagai pedoman untuk menyelesaikan permasalahan yang terapat dalam tugas akhir ini.

BAB III: PERANCANGAN SISTEM

Pada bab ini akan dijelaskan analisis permasalahan, solusi dalam penanganan masalah, analisis kebutuhan dan perancangan sistem.

BAB IV: IMPLEMENTASI SISTEM

Pada bab ini akan dijelaskan implementasi SOM untuk meng-*cluster* tugas akhir mahasiswa. Selain mengimplementasikan som, reduksi dimensi menggunakan *sparse matrix* juga dilakukan untuk mengurangi waktu komputasi SOM dan meningkatkan akurasi hasil.

BAB V: PENUTUP

Pada bab ini berisikan kesimpulan dari hasil ujicoba serta saran yang dapat digunakan untuk pengembangan selanjutnya.