

BAB 1

PENDAHULUAN

1.1. LATAR BELAKANG

Memasuki era globalisasi, diperlukan informasi yang cepat, tepat, akurat untuk kemajuan di segala bidang. Informasi yang diperlukan berasal dari sejumlah data yang sangat besar. Perkembangan teknologi yang pesat memungkinkan data jumlah besar terakumulasi sehingga menciptakan kondisi ledakan data tetapi informasi yang dihasilkan sedikit.

Dari ledakan data yang terjadi, kesulitan yang dihadapi adalah bagaimana mendapatkan nilai tambah dari basis data skala besar yang makin banyak terakumulasi sejalan dengan perkembangan teknologi informasi sehingga data-data tersebut dapat berfungsi serta menghasilkan suatu informasi yang berguna.

Dengan perkembangan teknologi saat ini, kesulitan tersebut dapat dipecahkan dengan menggunakan penambangan data atau lebih dikenal dengan data mining. Data mining adalah serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data. Dengan demikian dari kumpulan data tersebut dapat menghasilkan informasi.

Untuk mendapatkan nilai tambah berupa karakteristik data, salah satu teknik data mining yang dapat digunakan adalah klasterisasi. Pada dunia bisnis nilai tambah tersebut sangat penting untuk pendukung pembuatan suatu keputusan.

Data mining dengan kluster adalah mengelompokkan objek-objek berdasarkan kesamaan karakteristik di antara objek-objek tersebut dan ketaksamaan pada objek-objek dalam kluster lainnya. Fuzzy klustering adalah salah satu teknik untuk menentukan kluster optimal dalam suatu ruang vektor yang didasarkan pada bentuk normal Euclidian untuk jarak antar vektor di mana keberadaan tiap-tiap titik data dalam suatu kluster ditentukan oleh derajat keanggotaan.

Berdasarkan ulasan di atas, penulis akan membuat suatu aplikasi untuk melakukan data mining dari data yang tersimpan pada basis data relasional SQL Server 2000 dengan teknik klusterisasi menggunakan algoritma *Fuzzy C-Means*.

1.2. RUMUSAN MASALAH

1. Bagaimana membangun perangkat lunak untuk membentuk kluster-kluster dengan *Fuzzy C-Means* (FCM).
2. Bagaimana membangun perangkat lunak untuk membentuk kluster-kluster dengan *Fuzzy C-Means* (FCM) dengan jumlah kluster valid berdasarkan metode *Total Within Cluster Variation* (TWCV)

1.3. BATASAN MASALAH

1. Metode yang digunakan pada proses data mining adalah klustering.
2. Algoritma yang digunakan adalah *Fuzzy C-Means*.
3. Data yang digunakan adalah data yang sudah bersih siap untuk dimining yang disimpan pada suatu tabel dalam SQL Server 2000.

4. Pengklasteran data hanya dapat dilakukan pada data yang tersimpan di SQL Server.
5. Data yang digunakan hanya dari satu tabel, kolom input bertipe numerik dengan maksimal jumlah kolom = 20

1.4. TUJUAN

1. Membangun perangkat lunak untuk membentuk kluster-kluster dengan *Fuzzy C-Means* (FCM)
2. Membangun perangkat lunak untuk membentuk kluster-kluster dengan *Fuzzy C-Means* (FCM) dengan jumlah kluster valid berdasarkan metode *Total Within Cluster Variation* (TWCV)

1.5. METODE YANG DIGUNAKAN

Metode yang digunakan adalah:

1. Studi Pustaka
Dilakukan dengan mencari literatur yang terkait baik melalui perpustakaan maupun internet.
2. Membangun Perangkat Lunak
 - a. Analisis Kebutuhan Perangkat Lunak
Menganalisis data-data yang dibutuhkan dan informasi yang didapatkan untuk merancang Perangkat Lunak
 - b. Desain
Merancang Perangkat Lunak sebagai acuan implementasi Perangkat Lunak tersebut
 - c. Implementasi
Mengimplementasikan perancangan Perangkat Lunak ke dalam suatu program

d. Pengujian

Menguji Perangkat Lunak yang telah dibuat

1.6. SISTEMATIKA PENULISAN

Bab 1 : PENDAHULUAN

Bab ini berisi latar belakang masalah, rumusan masalah, batasan masalah, tujuan, sistematika masalah.

Bab 2 : LANDASAN TEORI

Bab ini berisi uraian sistematis dari teori yang ada pada literatur maupun penjabaran tinjauan pustakan yang mendasari pemecahan masalah.

Bab 3 : ANALISIS DAN PERANCANGAN SISTEM

Bab ini berisi analisis dan perancangan dari sistem yang akan dikembangkan.

Bab 4 : IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini berisi tentang implementasi dan pengujian sistem dari sistem yang dikembangkan.

Bab 5 : KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dan saran dari sistem yang dikembangkan.

Daftar Pustaka:

Halaman ini berisi sumber referensi yang digunakan pada sistem yang dikembangkan.