

BAB I PENDAHULUAN

1.1 Latar Belakang

Klasterisasi merupakan suatu teknik dalam data mining yang cukup penting dan banyak dipakai. Tujuan utama dari metode *cluster* adalah pengelompokan sejumlah data/objek ke dalam *cluster* (group) sehingga dalam setiap *cluster* akan berisi data yang semirip mungkin tetapi sangat tidak mirip dengan objek/data dari *cluster* yang lain^[3].

Salah satu teknik klastering yang cukup terkenal adalah *Fuzzy C-Means* (FCM). FCM adalah teknik pengklasteran fuzzy dimana keberadaan tiap titik ditentukan oleh derajat keanggotaan. Dengan menggunakan pendekatan fuzzy setiap objek dapat menjadi anggota dari semua kluster dengan derajat keanggotaan antara 0 sampai 1. Untuk menghasilkan formulasi yang presisi dalam menentukan kriteria klastering dapat ditempuh dengan metode fungsi objektif, yaitu dengan mengukur kemampuan untuk dilibatkan dalam *cluster* sebagai fungsi dari cacah *cluster* dengan fungsi objektif tertentu. Pemilihan fungsi objektif ini tergantung pada sebaran data objek. Struktur *cluster* yang baik adalah dengan meminimalkan fungsi objektif tersebut. Sering kali dalam meminimasi fungsi objektif tersebut dapat menghantarkan pada penyelesaian yang merupakan optimum lokal dikarenakan dalam mengaplikasikan metode ini harus diketahui secara pasti terlebih dahulu jumlah *cluster* yang akan dibuat dari sebaran data yang diberikan serta inisialisasi awal pusat *cluster*. Sehingga performansi dari FCM sangat tergantung pada pemilihan nilai pusat *cluster* awal^[1].

Evolutionary Programming (EP) merupakan salah satu teknik atau metode untuk mengatasi masalah optimasi. Walaupun pada awal diperkenalkannya metode ini ditujukan untuk menghasilkan suatu bentuk kecerdasan (*intelligence*) yang dipandang sebagai suatu tingkah laku yang adaptif (*adaptif behaviour*). Akan tetapi dalam perkembangannya EP justru mengalami pergeseran menuju bentuk yang mirip ES (*Evolutionary Strategies*), digunakan untuk mengatasi masalah-masalah optimasi numerik.

Dengan pertimbangan tersebut, tugas akhir ini difokuskan pada penggabungan FCM dengan EP. Hal ini dimaksudkan untuk menghilangkan masalah optimum lokal pada FCM karena EP berfungsi untuk membangkitkan pilihan solusi yang lebih bervariasi sehingga dapat meminimalisir terjadi kejenuhan fungsi objektif dan diharapkan bisa menghasilkan pengelompokan data yang lebih homogen, penyebaran pusat kluster yang optimum global, serta menghasilkan suatu *cluster* yang memiliki tingkat akurasi yang tinggi.

1.2 Identifikasi Masalah

Berdasarkan latar belakang yang telah diuraikan di atas, maka dapat dirumuskan beberapa masalah antara lain :

1. Bagaimana EP dapat digunakan untuk mengoptimasi FCM
2. Bagaimana menerapkan EP pada FCM dalam melakukan klasterisasi data.

3. Se jauh mana peningkatan kinerja FCM dengan pendekatan EP serta parameter-parameter FCM apa yang kemudian harus dipertimbangkan dalam penerapan EP.

Dalam implementasi tugas akhir ini dibatasi oleh beberapa hal, sebagai berikut :

1. Dataset yang digunakan adalah data yang sudah dilakukan *pre-processing* dan siap di-*mining*.
2. Jenis data pada dataset adalah data numeric saja.
3. Tidak ada fitur perubahan data uji tapi jumlah data dapat di tambah
4. Perubahan data uji akan mengakibatkan perubahan pada *source code* dan penyesuaian pada *dataset*.

1.3 Tujuan

Tujuan yang ingin dicapai dalam tugas akhir ini adalah

1. Menerapkan metoda EP untuk menghindari solusi yang optimum lokal pada penyelesaian pusat *cluster* dengan FCM biasa.
2. Menganalisis sejauh mana algoritma *Evolutionary programming* dapat meningkatkan kinerja *fuzzy clustering* dalam hal *cluster validity*, kompleksitas komputasi, dan *classification rate*.

1.4 Metode Penelitian

Metode yang digunakan penulis dalam mengumpulkan data-data untuk pembuatan aplikasi dan penyusunan Tugas Akhir dilakukan dengan cara, yaitu :

1. Studi literatur
Pencarian referensi dan sumber-sumber yang berhubungan dengan *Clustering*, *Fuzzy c-Means*, *Evolutionary Programming*, dan referensi yang dapat membantu dalam menyelesaikan tugas akhir ini
2. Pengumpulan data
Mengumpulkan beberapa dataset yang diperlukan untuk menguji sistem yang dibuat.
3. Analisis dan perancangan sistem
Melakukan analisis dan perancangan terhadap sistem yang dibangun, menganalisis metode yang akan digunakan untuk menyelesaikan permasalahan, termasuk menentukan bahasa pemrograman yang digunakan, arsitektur, fungsionalitas, dan antarmuka sistem. Input sistem berupa data set uji dan outputnya dari sistem adalah nilai validitas klastering.
4. Implementasi dan pembangunan sistem
 - a. Membangun sistem klastering dengan menggunakan *fuzzy c-means* secara konvensional.
 - b. Membangun sistem klastering dengan menggunakan *fuzzy c-means* dengan menambahkan metode *Evolutionary programming* dalam penentuan titik pusat klaster.

5. Pengujian dan analisis
Pengujian dan analisis dilakukan oleh penulis tanpa melibatkan pihak lain, meliputi:
 - a. Pengujian untuk mengetahui tingkat validitas klastering dan *classification rate* dari system klastering FCM konvensional.
 - b. Pengujian untuk mengetahui tingkat validitas klastering dan *classification rate* dari system klastering FCM dengan penambahan *Evolutionary Programming*.
 - c. Analisis perbandingan performansi antara FCM konvensional dengan penambahan EP dari segi tingkat validitas klaster dan nilai *classification rate*-nya.
6. Pengambilan kesimpulan dan penyusunan laporan Tugas Akhir.