

1. Pendahuluan

1.1 Latar belakang

Kebutuhan akan analisis data semakin berkembang. Salah satu cara dalam menganalisis data tersebut dengan cara membentuk suatu klaster. Suatu klaster ditentukan berdasarkan persamaan karakteristik dari setiap data. Dari klaster yang dihasilkan, dapat diketahui informasi mengenai kemiripan data dalam klaster tersebut. Salah satu keuntungan dalam membentuk klaster untuk proses analisis adalah dapat dengan mudah membaca karakteristik dari tiap klaster data sehingga memudahkan dalam menganalisa data. Analisis klaster merupakan cara yang bagus untuk mempercepat dalam meninjau data khususnya jika data diklaster ke dalam beberapa kelompok[5]. Oleh karena itu dibutuhkan suatu mekanisme tertentu untuk membentuk klaster yang tepat. Pada proses pengklasteran, pembentukan klaster dilakukan sedemikian rupa sehingga setiap data berada tepat dalam satu klaster. Namun, terkadang kita tidak dapat menempatkan suatu data dengan tepat pada satu klaster, karena terletak diantara dua atau lebih klaster. Kadangkala perbedaan data dapat memberikan hasil klaster yang kurang akurat, sehingga diperlukan suatu cara atau metode untuk menghasilkan klaster yang lebih akurat. Banyak cara atau metode yang dapat digunakan untuk menghasilkan klaster yang lebih akurat. Akan tetapi metode tersebut terkadang juga dapat memberikan hasil yang kurang optimal, sehingga diperlukan lagi cara untuk mengoptimalkan hasil klaster tersebut.

Salah satu metode pengklasteran yang dirasa dapat memberikan hasil klaster yang lebih akurat adalah *Fuzzy Subtractive Clustering (FSC)*. FSC merupakan salah satu algoritma klasterisasi dalam mencari klaster yang tepat. FSC memiliki kemampuan yang baik dalam menemukan klaster karena mempertimbangkan seluruh dimensi atau variabel yang dimiliki suatu data. FSC merupakan algoritma yang tidak terawasi sehingga pengguna tidak perlu menentukan jumlah klaster terlebih dahulu. Pengguna akan mendapatkan informasi jumlah klaster yang tepat hasil algoritma FSC. Hasil dari algoritma FSC ini berupa pusat klaster, standar deviasi, dan derajat keanggotaan tiap data terhadap klaster yang dihasilkan. Derajat keanggotaan tersebut mencerminkan kedekatan tiap data terhadap pusat klaster sehingga dapat dilihat anggota dari masing-masing klaster.

Untuk mengetahui lebih lanjut tingkat keakuratan dari klaster yang dihasilkan oleh FSC, maka dapat dibangun suatu *Fuzzy Inference System (FIS)*. FIS merupakan sistem yang dapat melakukan penalaran dengan prinsip serupa seperti manusia melakukan penalaran dengan nalurinya dengan input berupa bilangan tertentu dan output yang dihasilkan juga harus berupa bilangan tertentu. Pada FIS dibagi atas tiga proses yaitu fuzzifikasi, *rule evaluation*, dan defuzzifikasi. Pada proses fuzzifikasi dihasilkan suatu fungsi keanggotaan. Fungsi keanggotaan hasil fuzzifikasi menggunakan standar deviasi hasil FSC yang kemudian digunakan dalam defuzzifikasi. Perbedaan hasil defuzzifikasi FIS dengan data target (variabel output) akan digunakan untuk mengukur tingkat akurasi.

Pada tugas akhir ini dilakukan pengujian terhadap data Iris yang dikenalkan oleh Sir Ronald Aylmer Fisher (1936) sebagai contoh untuk analisis perbedaan[6].

Iris merupakan *genus* (marga) dari suatu spesies tanaman bunga. Kumpulan data Iris tersebut terdiri dari lima puluh contoh dari masing-masing subgenus Iris (*Iris setosa*, *Iris virginica*, dan *Iris versicolor*). Empat fitur yang diukur dari masing-masing contoh yaitu panjang dan lebar dari sepal (kelopak) dan petal (mahkota) Iris dalam ukuran sentimeter. Berdasarkan kombinasi dari keempat fitur tersebut, Fisher mengembangkan suatu model untuk membedakan spesies dari masing-masing data. Data Fisher tersebut biasa digunakan sebagai uji tipikal untuk teknik-teknik klasifikasi. Selain data Iris, tugas akhir ini juga melakukan percobaan untuk data yang lain yaitu data nilai mahasiswa yang diambil dari referensi 2. Data ini terdiri dari lima variabel yaitu nilai UTS (Ujian Tengah Semester), nilai tugas I, nilai UAS (Ujian Akhir Semester), nilai tugas II, dan nilai akhir. Selain kedua data yang telah disebutkan sebelumnya, tugas akhir ini juga menggunakan data *dummy*. Penelitian menggunakan data ini dimaksudkan untuk mengetahui kemungkinan suatu data terletak tidak tepat pada satu kluster, karena terletak diantara dua atau lebih kluster.

Pada permasalahan di atas, pemilihan model *fuzzy* dikarenakan berkemampuan untuk menjelaskan secara linguistik suatu sistem. Aturan-aturan dalam model *fuzzy* pada umumnya dibangun berdasarkan keahlian manusia dan pengetahuan heuristik dari sistem yang dimodelkan. Model *fuzzy* sangat sederhana dan mudah untuk dimengerti serta mengaplikasikan pengalaman tanpa proses pelatihan. Selain itu model *fuzzy* lebih fleksibel dan memiliki toleransi terhadap data-data yang tidak tepat dan didasarkan pada bahasa alami. Teknik ini selanjutnya dikembangkan menjadi teknik yang dapat mengidentifikasi aturan-aturan dari suatu basis data yang telah dikelompokkan berdasarkan persamaan strukturnya. Dalam hal ini metode pengelompokan *fuzzy* berfungsi untuk mencari kluster-kluster data. Informasi yang dihasilkan dari metode pengelompokan ini, yaitu informasi tentang pusat kelompok, digunakan untuk membentuk aturan-aturan dalam sistem penalaran *fuzzy*.

1.2 Perumusan masalah

Permasalahan yang akan dibahas pada tugas akhir ini adalah

1. bagaimana performansi sistem ditinjau dari tingkat akurasi hasil klusterisasi FSC
2. faktor-faktor apa saja pada FSC yang mempengaruhi ketepatan hasil kluster yang dijadikan sebagai dasar dalam penentuan jumlah aturan FIS

Agar tujuan akhir tercapai maka diperlukan adanya ruang lingkup yang menjadi batasan masalah. Pada tugas akhir ini yang menjadi batasan masalah diantaranya

1. metode *fuzzy inference system* yang digunakan adalah metode Sugeno orde satu.
2. fungsi keanggotaan yang digunakan pada proses fuzzifikasi adalah fungsi Gaussian karena punya tingkat ketelitian yang baik
3. tipe data berupa data numerik (angka) baik bilangan bulat maupun desimal dan tidak *null* dan salah satu variabelnya dapat dijadikan sebagai variabel output yang akan digunakan pada metode Sugeno orde satu.
4. faktor-faktor pengaruh ketepatan penentuan jumlah aturan pada *fuzzy inference system* yang akan diteliti merupakan variabel pada algoritma

fuzzy subtractive clustering yang terdiri dari jari-jari, accept ratio, reject ratio, dan squash factor

1.3 Tujuan

Melihat permasalahan di atas, tujuan yang dapat ditetapkan diantaranya

1. mengetahui performansi dari sistem yang ditinjau dari tingkat akurasi hasil klasterisasi algoritma FSC
2. membangun klaster yang optimal dengan tujuan demi performansi yang maksimal
3. menguji faktor-faktor yang berpengaruh dalam ketepatan hasil klaster.

1.4 Metodologi penyelesaian masalah

Metodologi yang akan digunakan untuk menyelesaikan tugas akhir ini adalah :

1. Studi literatur

Untuk menyelesaikan masalah tugas akhir ini, pertama kali dilakukan pembelajaran terhadap logika fuzzy dengan fungsi keanggotaan berbentuk Kurva *Gauss*. Kemudian pada tahap selanjutnya mempelajari proses pengklasteran data yang menggunakan algoritma *Fuzzy Subtractive Clustering*. Sedangkan tahap terakhir mempelajari literatur mengenai *Fuzzy Inference System* menggunakan metode Sugeno orde-1 serta mempelajari proses penentuan tingkat akurasi untuk mengetahui performansi sistem.

2. Perancangan sistem, yaitu:

Pada proses perancangan sistem, data akan dibagi menjadi data latih dan data uji. Selanjutnya data latih diolah dan dilakukan pengklasteran data menggunakan algoritma *Fuzzy Subtractive Clustering* hingga menghasilkan *sigma* (standart deviasi) dan pusat klaster. Tahap berikutnya variabel-variabel data latih akan dibedakan menjadi variabel input dan variabel output, selanjutnya menghitung derajat keanggotaan tiap data dengan hanya melibatkan variabel-variabel input saja terhadap pusat klaster yang telah dihasilkan. Berikutnya membangun *Fuzzy Inference System* menggunakan metode Sugeno orde-1 berdasar derajat keanggotaan, nilai *sigma*, dan pusat klaster hasil klasterisasi algoritma *Fuzzy Subtractive Clustering*. Hasil defuzzifikasi pada proses *Fuzzy Inference System* selanjutnya akan dilakukan perhitungan performansi sistem yang dilihat dari tingkat akurasi antara variabel output data dengan hasil defuzzifikasi. Tahap terakhir dilakukan pengujian dari data uji terhadap data latih.

3. Implementasi

Setelah mendapatkan rancangan yang jelas, barulah sistem diterapkan dalam bentuk suatu program simulasi dengan mengimplementasikan *fuzzy inference system* pada hasil klasterisasi algoritma *fuzzy subtractive clustering* untuk selanjutnya akan dilakukan analisa.

4. Evaluasi dan analisis hasil

Pada tahap ini hasil sistem akan dievaluasi sehingga dapat diketahui apakah hasil sistem yang diimplementasikan sesuai dengan tujuan yang dilakukan untuk melihat performansi sistem berdasar tingkat akurasi, membentuk klaster yang optimal, serta untuk mengetahui faktor-faktor yang

mempengaruhi ketepatan hasil klaster. Kemudian dilakukan analisis terhadap sistem berdasar data-data yang diperoleh pada tahap implementasi

5. Pembuatan Dokumentasi

Tahap akhir dari penelitian ini adalah pembuatan dokumentasi disertai menarik kesimpulan dari seluruh langkah-langkah yang dilakukan dalam penelitian serta saran-saran yang perlu diperhatikan untuk memperbesar manfaat penelitian ini terutama dalam hal perbaikan hasil penelitian.

1.5 Sistematika Penulisan

BAB I Pendahuluan

Bab ini menguraikan tugas akhir secara umum, meliputi latar belakang, perumusan masalah, tujuan, dan metodologi penyelesaian masalah yang digunakan.

BAB 2 Dasar Teori

Bab ini membahas uraian teori yang berhubungan dengan sistem fuzzy, algoritma *fuzzy subtractive clustering*, *fuzzy inference system* menggunakan metode Sugeno orde-1, serta prosentase menentukan tingkat akurasi.

BAB 3 Analisis Kebutuhan dan Pemodelan Sistem

Bab ini berisi gambaran umum sistem, analisa kebutuhan sistem, pengolahan data, serta pemodelan sistem

BAB 4 Percobaan dan Analisis Hasil Percobaan

Bab ini membahas mengenai hasil percobaan yang telah dilakukan yaitu pengklasteran data latih dan selanjutnya akan dianalisa faktor-faktor yang berpengaruh terhadap hasil klaster tersebut serta performansi sistem yang dilihat dari aspek akurasi

BAB 5 Kesimpulan dan Saran

Berisi kesimpulan dari pembuatan tugas akhir ini dan saran-saran yang diperlukan untuk pengembangan lebih lanjut.