

Penerapan Metode *Fuzzy C-Means* Untuk Pengelompokan Data Kredit

Ni Luh Gede Pivin Suwirmayanti

Program Studi Sistem Komputer, STIKOM Bali
Jl. Raya Puputan No. 86 Renon, Denpasar, Bali, Telp. (0361)244445
e-mail: pivin@stikom-bali.ac.id

Abstrak

Segmentasi data merupakan suatu proses pengelompokan data yang semula berperilaku berbeda-beda menjadi beberapa kelompok yang sekarang berperilaku lebih seragam. Metode clustering cukup banyak digunakan untuk mengatasi permasalahan yang terkait dengan segmentasi data. Clustering ini merupakan metode pengelompokan berdasarkan ukuran kedekatan. Salah satu metode clustering yang dapat digunakan untuk mengelompokkan data adalah *Fuzzy C-Means*, yang merupakan suatu metode pengelompokan data yang ditentukan oleh derajat keanggotaan. Adapun tujuan penelitian ini adalah mengkaji penerapan fuzzy clustering, utamanya *Fuzzy C-Means*, dalam penentuan pusat cluster dalam permasalahan klasifikasi untuk penilaian kelayakan pemberian kredit. Data Uji yang digunakan adalah German Credit Dataset terdiri dari 1000 record dan 20 variabel dengan tipe data campuran yaitu tipe data numerik dan kategorikal. Studi kasus yang diangkat dalam penelitian ini adalah mengenai segmentasi pasar, dengan menerapkan metode *Fuzzy C-Means*. Pada tahapan desain sistem menggunakan use case diagram dan pada tahapan implementasi sistem dilakukan inputan data uji, menginputkan atribut, menginputkan jumlah cluster yang dihasilkan, sehingga memperoleh pusat cluster yang terbentuk yang nantinya akan dapat ditindaklanjuti sebagai bahan pertimbangan untuk pengambilan keputusan dan juga memungkinkan untuk melakukan pendekatan marketing yang sesuai dengan ciri-ciri dominan dari kelompok-kelompok yang terbentuk. Terakhir dilakukan proses pengujian dengan dengan memperoleh hasil sesuai dengan skenario pengujian.

Kata kunci: Clustering, *Fuzzy C-Means*, Data Kredit, Segmentasi Data.

1. Pendahuluan

Segmentasi data merupakan pengelompokan data yang nantinya dapat dipilih sebagai target tujuan pemasaran suatu produk. Segmentasi data juga merupakan proses pengelompokan data yang semula berperilaku berbeda-

beda menjadi beberapa kelompok pasar yang sekarang berperilaku lebih seragam. Pengelompokan data pada perusahaan seharusnya dapat dianalisis dan dimanfaatkan lebih lanjut dengan menggunakan metode Clustering. Analisis ini memberikan banyak output yang dapat dimanfaatkan perusahaan. Semua output tersebut dapat digunakan perusahaan untuk memenangkan persaingan ataupun meningkatkan income dan omzet perusahaan itu tersebut. Seperti halnya pengelompokan data penilaian kelayakan pemberian kredit yang terdiri dari 1000 record dan 20 variabel ditambah dengan sebuah variabel target atau variabel response [1]. Pengelompokan data ini diperlukan untuk mendapatkan nilai-nilai dominan yang akan menjadi ciri-ciri dari masing-masing kelompok atau segment.

Metode clustering cukup banyak digunakan untuk mengatasi permasalahan yang terkait dengan segmentasi data. Clustering ini merupakan metode pengelompokan berdasarkan ukuran kedekatan (kemiripan) dimana cluster tidak harus sama persis akan tetapi merupakan pengelompokan berdasarkan pada kedekatan dari suatu karakteristik sample data yang ada, salah satunya dengan menggunakan rumus jarak euclidean [2]. Penerapan metode clustering dapat digunakan untuk mengelompokkan pelanggan yang memiliki kesamaan dalam perilaku pelanggan dalam berbelanja. Semakin akurat cluster yang terbentuk maka akan semakin jelas kesamaan pola perilaku dari pelanggan. Sehingga, perusahaan dapat menentukan strategi pemasaran dengan lebih tepat, berdasarkan pada pola perilaku pelanggan yang telah didapatkan dari proses cluster yang sudah terbentuk.

Salah satu metode clustering yang dapat digunakan untuk mengelompokkan data adalah *Fuzzy C-Means* (FCM), yang merupakan suatu teknik pengelompokan data dimana keberadaan tiap titik data dalam suatu kelompok (cluster) ditentukan oleh derajat keanggotaan. Dengan cara memperbaiki pusat cluster dan nilai keanggotaan tiap-tiap data secara berulang, maka akan didapat bahwa pusat cluster menuju lokasi yang tepat [2]. Nilai awal titik pusat cluster *Fuzzy C-Means* dibangkitkan secara acak sehingga terjadi optimum

lokal, dimana proses selanjutnya bergantung pada nilai awal yang dibangkitkan secara acak.

Berdasarkan pembahasan diatas maka penelitian ini mengusulkan Penerapan Metode *Fuzzy C-Means* untuk Pengelompokan Data Kredit. Metode *Fuzzy C-Means* ini dipilih karena dapat menentukan jumlah cluster yang akan dibentuk. Adapun tujuan penelitian ini adalah mengkaji penerapan fuzzy clustering, utamanya *Fuzzy C-Means*, serta mengkaji sejauh mana unjuk kerja *Fuzzy C-Means* dalam pembentukan cluster dalam permasalahan klasifikasi untuk penilaian kelayakan pemberian kredit. Sehingga informasi ini nantinya akan dapat ditindaklanjuti sebagai bahan pertimbangan untuk pengambilan keputusan dan juga memungkinkan untuk melakukan pendekatan marketing yang sesuai dengan ciri-ciri dominan dari kelompok-kelompok yang terbentuk.

2. Pembahasan

Fuzzy C-Means Clustering (FCM)

Fuzzy C-Means adalah salah satu teknik pengelompokan data yang mana keberadaan tiap titik data dalam suatu kelompok (cluster) ditentukan oleh derajat keanggotaan. Berbeda dengan K-Means clustering, dimana suatu objek hanya akan menjadi anggota satu cluster, dalam *Fuzzy C-Means* setiap objek dapat menjadi anggota dari beberapa cluster. Batas-batas dalam K-Means adalah tegas (hard) sedangkan dalam *Fuzzy C-Means* adalah soft. *Fuzzy C-Means* bersifat sederhana, mudah diimplementasikan, memiliki kemampuan untuk melakukan pengelompokan data yang besar, lebih kokoh terhadap data outlier. Algoritma *Fuzzy C-Means* adalah sebagai berikut [3]:

1. Pemasukan data yang akan di cluster X, berupa matriks berukuran n x m
dengan:
n = jumlah sample data
m = atribut setiap data
X_{ij} = data sample ke-i (i = 1,2,...,n), atribut ke-j (j = 1,2,...,m)

2. Menentukan beberapa hal diantaranya:

- Jumlah cluster = c;
- Pangkat = w;
- Maksimum iterasi = MaxIter;
- Error terkecil yang diharapkan = ε;
- Fungsi obyektif awal = P₀ = 0;
- Iterasi awal = t=1;

3. Membangkitkan nilai acak μ_{ik}, i=1,2,...,n; k=1,2,...,c; sebagai elemen-elemen matriks partisi awal u. μ_{ik} adalah derajat keanggotaan yang merujuk pada seberapa besar kemungkinan suatu data bisa menjadi anggota ke dalam

suatu cluster. Posisi dan nilai matriks dibangun secara random. Dimana nilai keanggotaan terletak pada interval 0 sampai dengan 1. Pada posisi awal matriks partisi U masih belum akurat begitu juga pusat clusternya. Sehingga kecendrungan data untuk masuk suatu cluster juga belum akurat.

Selanjutnya menghitung jumlah setiap kolom (atribut)

$$Q_j = \sum_{k=1}^c \mu_{ik} \dots \dots \dots (1)$$

Q_j adalah jumlah nilai derajat keanggotaan perkolom = 1 dengan j=1,2,...,m

Selanjutnya dilakukan perhitungan sebagai berikut:

$$\mu_{ik} = \frac{\mu_{ik}}{Q_j} \dots \dots \dots (2)$$

4. Menghitung pusat Cluster ke-k: V_{kj}, dengan k=1,2,...,c; dan j=1,2,...,m.

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w * X_{ij})}{\sum_{i=1}^n (\mu_{ik})^w} \dots \dots \dots (3)$$

5. Menghitung fungsi obyektif pada iterasi ke-t, P_t. Fungsi obyektif digunakan sebagai syarat perulangan untuk mendapatkan pusat cluster yang tepat. Sehingga diperoleh kecendrungan data untuk masuk ke cluster mana pada langkah akhir.

$$P_t = \sum_{i=1}^n \sum_{k=1}^c ([\sum_{j=1}^m (X_{ij} - V_{kj})^2] (\mu_{ik})^w) \dots \dots (4)$$

6. Menghitung perubahan matriks partisi:

$$\mu_{ik} = \frac{[\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{-\frac{1}{w-1}}}{\sum_{k=1}^c [\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{-\frac{1}{w-1}}} \dots \dots \dots (5)$$

dengan: i=1,2,...,n; dan k=1,2,...,c.

7. Mengecek kondisi berhenti:

Jika:
(|P_t - P_{t-1}| < ζ)(6)

atau
(t > maxIter)(7)

maka berhenti;
Jika tidak: t=t+1, dilakukan pengulangan langkah ke-4.

Validity Cluster

Metode cluster validity yang digunakan yaitu [4]:

- Korelasi
Menggunakan 2 jenis matriks, yaitu proximity matrix dan incidence matrix
 - Proximity matrix adalah matriks yang berisi jarak antar obyek
 - Incidence matrix adalah matriks biner yang mengindikasikan keanggotaan klaster
 - 0 jika anggota dari klaster yang berbeda
 - 1 jika anggota dari klaster yang sama

2. Cohession dan Separation

Cohession merupakan cluster validity yang menghitung varian intra kluster (WSS).

$$WSS = \sum_{i=1}^k \sum_{x \in C_i} (x - m_i)^2 \dots\dots\dots (8)$$

Sedangkan Separation merupakan cluster validity yang menghitung varian inter kluster(BSS)

$$BSS = \sum_{i=1}^k |C_i|(m - m_i)^2 \dots\dots\dots(9)$$

3. Silhouette Coefficient

Menggunakan gabungan kedua prinsip dasar dari cohesion dan separation

$$s = \begin{cases} 1 - \frac{a}{b}, & \text{jika } a < b \\ \frac{b}{a} - 1, & \text{jika } a \geq b \end{cases} \dots\dots\dots (3)$$

dimana:

a = rata-rata jarak ke obyek satu kluster

b = minimal rata-rata jarak ke obyek berbeda kluster

4. Dunn Index

Memiliki dasar pemikiran bahwa kluster yang baik adalah yang memiliki diameter yang kecil dan jarak yang besar terhadap kluster lainnya.

$$D = \min_{i=1 \dots nc} \left(\min_{j=i+1 \dots nc} \left(\frac{d(c_i, c_j)}{\max_{k=1 \dots nc} (diam(c_k))} \right) \right) \dots (4)$$

dimana:

a. d(c_i, c_j) adalah jarak terhadap kluster lain

b. diam(c_k) adalah jarak terhadap sesama kluster

5. Davies-Bouldin Index

Memiliki prinsip dasar dalam menghitung similarity antar kluster

$$DB = \left(\frac{1}{nc} \right) \sum_{i=1}^{nc} \left(\max_{\substack{j=1 \dots nc \\ i \neq j}} (R_{ij}) \right) \dots\dots\dots (5)$$

dimana nilai

$$R_{ij} = \frac{s_i + s_j}{d_{ij}} \dots\dots\dots(6)$$

- a. s_i adalah rata-rata jarak obyek seluruh kluster i dengan pusatnya
- b. d_{ij} adalah jarak pusat kluster i dan pusat kluster j

Fitur Selection

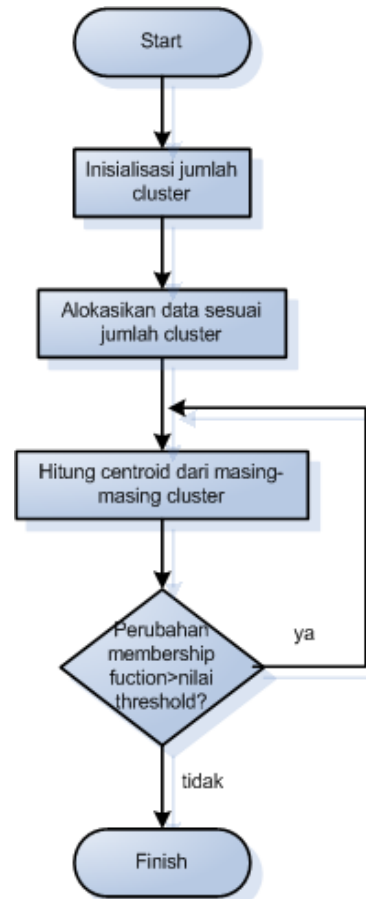
Feature selection atau Feature reduction adalah suatu kegiatan yang umumnya bisa dilakukan secara preprocessing dan bertujuan untuk memilih fitur yang berpengaruh dan mengesampingkan fitur yang tidak berpengaruh dalam suatu kegiatan pemodelan atau penganalisaan data. Ada banyak alternatif yang bisa digunakan dan harus dicoba-coba untuk mencari yang cocok. Secara garis besar ada dua kelompok besar dalam pelaksanaan feature selection, yaitu ranking selection dan subset selection [4].

Pemilihan fitur merupakan kegiatan yang termasuk ke dalam preprocessing yang bertujuan untuk memilih fitur yang berpengaruh dan mengesampingkan fitur yang

tidak berpengaruh dalam suatu kegiatan pemodelan atau penganalisaan data. Terdapat banyak alternatif yang dapat digunakan dan juga dilakukan proses mencoba-coba untuk mencari fitur yang cocok [4].

Analisis Cluster

Secara umum sistem yang akan dibuat dalam penelitian ini adalah sistem untuk pengelompokan data kredit menggunakan perhitungan Fuzzy clustering Berikut adalah diagram alur kerja penelitian yang dilakukan:



Gambar 1. Alur Kerja Fuzz C-Means

Langkah-langkah kegiatan yang akan dilakukan untuk permasalahan ini adalah sebagai berikut [5] :

1. Inisialisasi Jumlah Cluster, pertama kali adalah menentukan pusat cluster yang akan menandai lokasi rata-rata untuk tiap cluster.
2. Alokasi data sesuai jumlah cluster, dan pada kondisi awal jumlah cluster masih belum akurat, tiap-tiap titik data memiliki derajat keanggotaan untuk tiap cluster.
3. Menghitung Centroid dari masing-masing cluster, dengan cara memperbaiki pusat cluster dan derajat keanggotaan tiap titik data secara berulang maka akan dilihat bahwa pusat cluster akan bergerak menuju lokasi yang tepat.

- Melihat perubahan membership function jika lebih besar dari nilai threshold, dengan perulangan didasarkan pada minimisasi fungsi obyek yang menggambarkan jarak dari titik data yang diberikan ke pusat cluster yang terbobot oleh derajat keanggotaan titik data tersebut.

Data Uji

Penelitian ini nantinya akan menggunakan data uji yang banyak dipakai dalam permasalahan klasifikasi untuk penilaian kelayakan pemberian kredit yang disebut German Credit Dataset, data set ini didonasikan oleh Prof. Hofman dari Hamburg University, Jerman. Dataset ini terdiri dari 1000 record dan 20 variabel ditambah dengan sebuah variabel target atau variabel response, dimana 13 variabel diantaranya bertipe kategori dan sisanya sebanyak 7 variabel bertipe numerik. German Credit Dataset ini dapat diunduh di UCI Machine Learning Repository.

Studi kasus yang diangkat dalam penelitian ini adalah mengenai segmentasi pasar. Pendefinisian pemasaran memberikan definisi yang lebih terinci dengan mengatakan bahwa segmentasi pasar adalah proses membagi-bagi pasar yang semula berperilaku heterogen menjadi beberapa kelompok pasar yang sekarang berperilaku lebih seragam [1].

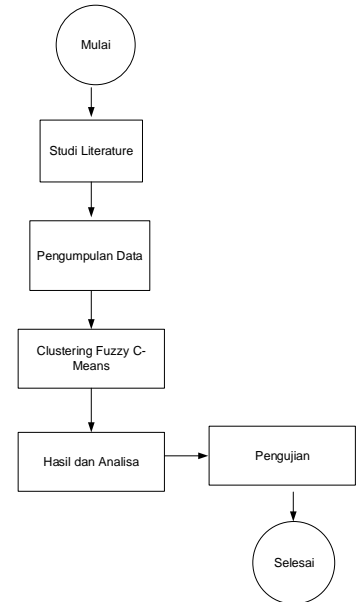
Penentuan segmentasi pasar berdasarkan pada beberapa kriteria diantaranya: Segmentasi demografis didasarkan pada karakteristik populasi yang dapat diukur dari umur, jenis kelamin, pendapatan, pendidikan, dan pekerjaan. Segmentasi Psikografis adalah proses pengelompokkan orang dalam arti sikap, nilai-nilai yang dianut, dan gaya hidup. Segmentasi tingkah laku memfokuskan pada apakah orang akan membeli dan menggunakan suatu produk atau tidak, disamping seberapa sering dan berapa banyak yang dipakainya. Konsumen dapat dikategorikan menurut tingkat penggunaan, misalnya, pengguna berat, sedang, ringan, dan bukan pengguna. Segmentasi manfaat (Benefit Segmentation), memfokuskan pada nilai persamaan (value equation) [3].

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U |
|----|-----|----|-----|-----|-------|-----|-----|---|-----|------|---|------|----|------|------|---|------|---|------|------|---|
| 1 | A11 | 6 | A34 | A43 | 1169 | A65 | A75 | 4 | A93 | A101 | 4 | A121 | 67 | A143 | A152 | 2 | A173 | 1 | A192 | A201 | 1 |
| 2 | A12 | 48 | A32 | A43 | 5951 | A61 | A73 | 2 | A92 | A101 | 2 | A121 | 22 | A143 | A152 | 1 | A173 | 1 | A191 | A201 | 2 |
| 3 | A14 | 12 | A34 | A46 | 2096 | A61 | A74 | 2 | A93 | A101 | 3 | A121 | 49 | A143 | A152 | 1 | A172 | 2 | A191 | A201 | 1 |
| 4 | A11 | 42 | A32 | A42 | 7882 | A61 | A74 | 2 | A93 | A103 | 4 | A122 | 45 | A143 | A153 | 1 | A173 | 2 | A191 | A201 | 1 |
| 5 | A11 | 24 | A33 | A40 | 4870 | A61 | A73 | 3 | A93 | A101 | 4 | A124 | 53 | A143 | A153 | 2 | A173 | 2 | A191 | A201 | 2 |
| 6 | A14 | 36 | A32 | A46 | 9055 | A65 | A73 | 2 | A93 | A101 | 4 | A124 | 35 | A143 | A153 | 1 | A172 | 2 | A192 | A201 | 1 |
| 7 | A14 | 24 | A32 | A42 | 2835 | A63 | A75 | 3 | A93 | A101 | 4 | A122 | 53 | A143 | A152 | 1 | A173 | 1 | A191 | A201 | 1 |
| 8 | A12 | 36 | A32 | A41 | 6948 | A61 | A73 | 2 | A93 | A101 | 2 | A123 | 35 | A143 | A151 | 1 | A174 | 1 | A192 | A201 | 1 |
| 9 | A14 | 12 | A32 | A43 | 3059 | A64 | A74 | 2 | A91 | A101 | 4 | A121 | 61 | A143 | A152 | 1 | A172 | 1 | A191 | A201 | 1 |
| 10 | A12 | 30 | A34 | A40 | 5234 | A61 | A71 | 4 | A94 | A101 | 2 | A123 | 28 | A143 | A152 | 2 | A174 | 1 | A191 | A201 | 2 |
| 11 | A12 | 12 | A32 | A40 | 1295 | A61 | A72 | 3 | A92 | A101 | 1 | A123 | 25 | A143 | A151 | 1 | A173 | 1 | A191 | A201 | 2 |
| 12 | A11 | 48 | A32 | A49 | 4308 | A61 | A72 | 3 | A92 | A101 | 4 | A122 | 24 | A143 | A151 | 1 | A173 | 1 | A191 | A201 | 2 |
| 13 | A12 | 12 | A32 | A43 | 1567 | A61 | A73 | 1 | A92 | A101 | 1 | A123 | 22 | A143 | A152 | 1 | A173 | 1 | A192 | A201 | 1 |
| 14 | A11 | 24 | A34 | A40 | 1199 | A61 | A75 | 4 | A93 | A101 | 4 | A123 | 60 | A143 | A152 | 2 | A172 | 1 | A191 | A201 | 2 |
| 15 | A11 | 15 | A32 | A40 | 1403 | A61 | A73 | 2 | A92 | A101 | 4 | A123 | 28 | A143 | A151 | 1 | A173 | 1 | A191 | A201 | 1 |
| 16 | A11 | 24 | A32 | A43 | 1282 | A62 | A73 | 4 | A92 | A101 | 2 | A123 | 32 | A143 | A152 | 1 | A172 | 1 | A191 | A201 | 2 |
| 17 | A14 | 24 | A34 | A43 | 2424 | A65 | A75 | 4 | A93 | A101 | 4 | A122 | 53 | A143 | A152 | 2 | A173 | 1 | A191 | A201 | 1 |
| 18 | A11 | 30 | A30 | A49 | 8072 | A65 | A72 | 2 | A93 | A101 | 3 | A123 | 25 | A141 | A152 | 3 | A173 | 1 | A191 | A201 | 1 |
| 19 | A12 | 24 | A32 | A41 | 12579 | A61 | A75 | 4 | A92 | A101 | 2 | A124 | 44 | A143 | A153 | 1 | A174 | 1 | A192 | A201 | 2 |
| 20 | A14 | 24 | A32 | A43 | 3430 | A63 | A75 | 3 | A93 | A101 | 2 | A123 | 31 | A143 | A152 | 1 | A173 | 2 | A192 | A201 | 1 |
| 21 | A14 | 9 | A34 | A40 | 2134 | A61 | A73 | 4 | A93 | A101 | 4 | A123 | 48 | A143 | A152 | 3 | A173 | 1 | A192 | A201 | 1 |
| 22 | A11 | 6 | A32 | A43 | 2647 | A63 | A73 | 2 | A93 | A101 | 3 | A121 | 44 | A143 | A151 | 1 | A173 | 2 | A191 | A201 | 1 |
| 23 | A11 | 10 | A34 | A40 | 2241 | A61 | A72 | 1 | A93 | A101 | 3 | A121 | 48 | A143 | A151 | 2 | A172 | 2 | A191 | A202 | 1 |

Gambar 2. Data Uji

Alur Analisis

Berikut merupakan alur analisis dari sistem penerapan metode Fuzzy C-Means untuk pengelompokan data kredit.



Gambar 3. Alur Analisis

Tahapan-tahapan dari alur analisis penelitian yang akan dilakukan, diantaranya sebagai berikut :

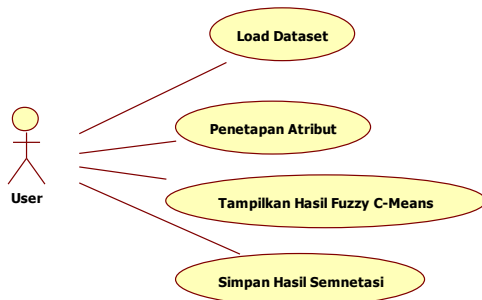
- Manfaat dari studi literatur dapat memberikan gambaran menyeluruh mengenai sejauh mana perkembangan penelitian-penelitian yang terkait dengan penelitian yang akan diambil. Dalam penelitian ini referensi diperoleh dari jurnal, artikel laporan penelitian, dan buku yang berhubungan mengenai penelitian yang *Fuzzy C-Means* untuk pengelompokan data.
- Pengumpulan Data. Berdasarkan bentuk dan sifatnya, data penelitian yang digunakan bersifat campuran yaitu data kualitatif (yang berbentuk kata-kata/kalimat) dan data kuantitatif (yang berbentuk angka). Data kuantitatif dapat dikelompokkan berdasarkan cara mendapatkannya yaitu data diskrit dan data kontinum. Berdasarkan sifatnya, data kuantitatif terdiri atas data nominal, data ordinal, data interval dan data rasio.
- Metode yang diusulkan dalam penelitian ini adalah metode *Fuzzy C-Means* yang. Dimana dalam penelitian ini akan mencoba melakukan pembentukan cluster pada data bertipe campuran yaitu penentuan pemberian kredit.
- Tahapan Hasil dan Analisis merupakan tahapan evaluasi yang melakukan pengukuran terhadap hasil yang diperoleh. Hasil akhir yang diharapkan dari penelitian mengenai penerapan metode *Fuzzy*

C-Means dalam pengelompokan data adalah memperoleh Cluster data kredit.

- Pengujian, Tahapan pengujian dialkuakn setelah tahapan hasil dan analisis, dimana pengujian menggunakan metode pengujian Black Box yang memfokuskan pada fungsional dari software.

Desain Sistem

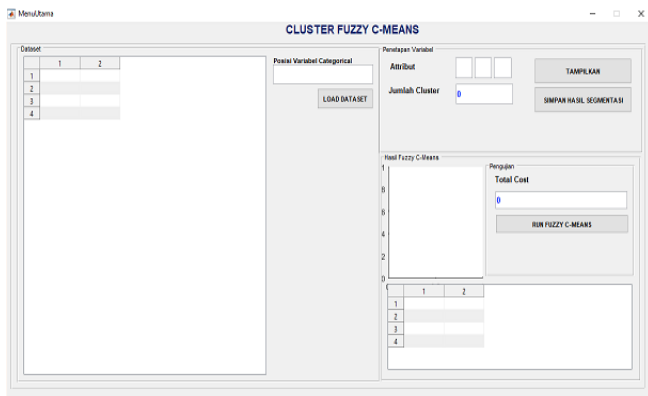
Desain Sistem menggunakan Use Case Diagram untuk mejabarkan fungsi sistem yang ada pada sistem pengelompokan data kredit dengan menerapkan Metode Fuzzy C-Means dapat dilihat pada gambar berikut:



Gambar 4. Tampilan Menu Utama

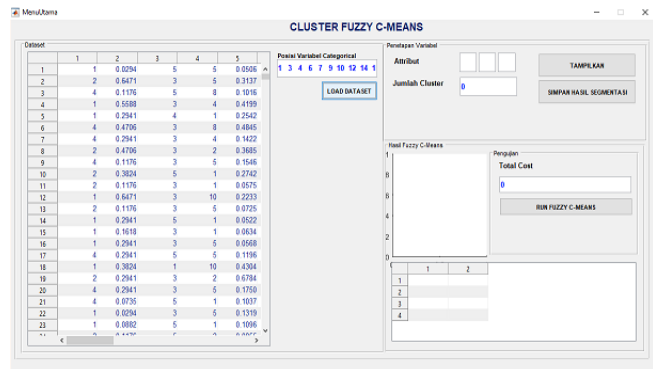
Implementasi Sistem

Implementasi sistem dilakukan setelah tahapan perancangan sistem. Adapun pengelompokan data kredit menggunakan metode Fuzzy C-Means dapat dilihat pada gambar 5.



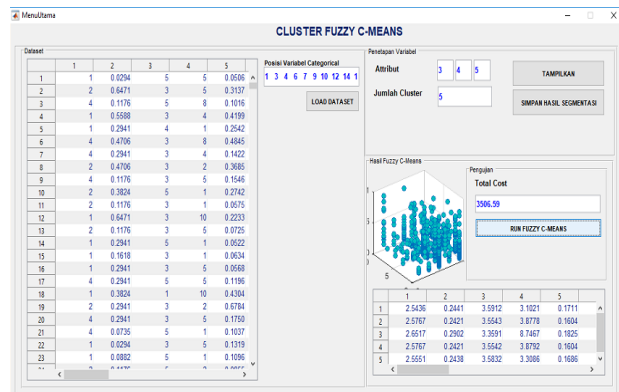
Gambar 5. Tampilan Menu Utama

Gambar 6 menunjukkan inputan data kredit, dimana data yang diinputkan bertipe data campuran yaitu numerik dan kategorikal. Akan muncul posisi data kategorikal untuk menunjukkan posisi data tersebut.



Gambar 6. Inputan Data

Gambar 7 merupakan tampilan dari penetapan variabel dengan menginputkan 3 atribut, dimana atribut merupakan posisi kolom data yang dipilih dari 1 sampai 20 kolom. Jumlah cluster yang ditentukan yaitu 5 cluster. Maka akan diperoleh hasil grafik tiga dimensi untuk menggambarkan cluster yang terbentuk. Dari Grafik juga akan muncul tabel pusat cluster yang terbentuk dari 5 cluster yang sudah dihasilkan.



Gambar 7. Hasil pembentukan pusat cluster dengan Fuzzy C-Means

Pengujian

Adapun hasil pengujian sistem dapat dilihat sebagai berikut:

Tabel 1. Pengujian dengan Black Box

| No | Komponen | Skenario | Hasil |
|----|---------------------------|--|----------|
| 1. | Load Dataset | Untuk Pengambilan data yang akan dioleh. | Berhasil |
| 2. | Posisi Data Kategorikal | Untuk memunculkan posisi variabel kategorikal pada kolom data. | Berhasil |
| 3. | Penentuan Variabel | Untuk memilih inputan atribut | Berhasil |
| 4. | Menampilkan Hasil Cluster | Untuk menentukan jumlah cluster yang | Berhasil |

| No | Komponen | Skenario | Hasil |
|----|------------------------|--|----------|
| | | diinginkan dan akan muncul dalam grafik 3D | |
| 5. | Total Cost | Total cost muncul untuk pengukuran hasil clustering, untuk mengukur total jarak yang dihasilkan. | Berhasil |
| 6. | Tampilan Pusat Cluster | Untuk menampilkan pusat cluster yang dibentuk sesuai jumlah cluster yang ditentukan, yang dapat dilihat pada Grid. | Berhasil |
| 7. | Simpan Hasil Sementara | Untuk Menyimpan Hasil Cluster dan Total Cost | Berhasil |

3. Kesimpulan

Kesimpulan yang dapat ditarik dari penelitian ini sebagai berikut :

1. Perancangan sistem ini menggunakan Use Case Diagram.
2. Telah berhasil diterapkan metode *Fuzzy C-Means* untuk Pengelompokan data kredit, dimana data uji ini banyak dipakai dalam permasalahan klasifikasi untuk penilaian kelayakan pemberian kredit yang disebut German Credit Dataset.
3. Hasil pengujian menunjukkan sistem berhasil dibuat sesuai dengan yang diharapkan.

Daftar Pustaka

- [1]. M. V. Jagannatha Reddy and B. Kavitha. "Clustering the Mixed Numerical and Categorical Dataset using Similarity Weight and Filter Method", *International Journal of Database Theory and Application* Vol. 5, 2012.
- [2]. D.R. Ningrat, D. Asih. I Maruddani, T. Wuryandari. "Analisis Cluster Dengan Algoritma K-Means Dan Fuzzy C-Means Clustering Untuk Pengelompokan Data Obligasi Korporasi", *Jurnal Gaussian*, Vol.5, No. 4, 2016.
- [3]. C. Zhao, H. Zhao, W. Yao. "Fuzzy C-Means Clustering Based on Improved Marked Watershed Transformation", *Jurnal Telekomika*, Vol.14, No.3, 2016.
- [4]. P. Suwirmayanti, I.K.G. Darma Putra, I.N. Satya Kumara." Optimasi Pusat Cluster K-Prototype dengan Algoritma Genetika", *Jurnal Teknologi Elektro*, Vol. 13, No. 2, 2014.
- [5]. P.I. Ciptayani, K.C. Dewi, I.W. Budi Sentana. Student Grouping Using Adaptive Genetic Algorithm. *International Electronics Symposium (IES)*, 2016.