

APLIKASI METODE *FUZZY C-MEANS* UNTUK PENGLASTERAN KELAYAKAN RUMAH DI DESA WAYAME, AMBON

R. P. A. Sormin¹, F. Y. Rumlawang², L. J. Sinay³

^{1, 2, 3} Jurusan Matematika FMIPA Universitas Pattimura
Jl. Ir. M. Putuhena, Kampus Unpatti, Poka-Ambon, Indonesia
e-mail: ³lj.sinay@staff.unpatti.ac.id

Abstrak

Pengklasteran adalah proses pengelompokan data ke dalam kluster berdasarkan parameter tertentu sehingga obyek-obyek dalam sebuah kluster memiliki tingkat kemiripan yang tinggi satu sama lain dan sangat tidak mirip dengan obyek yang lain pada kluster yang berbeda. Algoritma *Fuzzy C-Means* termasuk salah satu teknik pengklasteran data yang mana keberadaan pada setiap titik data dalam suatu kluster ditentukan oleh derajat keanggotaan. Pada penelitian ini Algoritma *Fuzzy C-Means* digunakan untuk menentukan kelayakan rumah. Hasil yang diperoleh menunjukkan bahwa masih banyak rumah yang tidak layak di Desa wayame yang harus lebih diperhatikan.

Kata Kunci: *C-Means*, *fuzzy*, kelayakan, pengklasteran.

APPLICATION OF FUZZY C-MEANS ALGORITHM FOR CLUSTERING HOUSE FEASIBILITY IN WAYAME VILLAGE, AMBON

Abstract

Clustering is a process of organizing objects into groups whose members are similar in a cluster but different with members of other cluster. Fuzzy C-Means Algorithm is a data clustering technique in which a dataset is grouped into clusters with every data point in the data sets belonging to every cluster to a certain degree. In this research Fuzzy C-Means Algorithm is used to determine house feasibility. The result shows that there are many houses which are not feasible and need to get more attention.

Keywords: C-means, clustering, fuzzy, feasibility.

1. Pendahuluan

Secara garis besar kelayakan merupakan tahap dimana pantas atau tidaknya sesuatu berada pada tempat tertentu. Penentuan kelayakan merupakan hal yang sangat penting dilakukan dalam mempertimbangkan dan mengambil sebuah keputusan. Masalah penentuan kelayakan seringkali menjadi masalah yang sangat rumit dan kompleks, sehingga membutuhkan solusi yang sangat tepat dan sesuai. Sehingga output yang dihasilkan sesuai dengan yang diharapkan.

Penentuan kelayakan bisa terjadi dalam berbagai kasus, salah satunya adalah penentuan kelayakan kelayakan rumah hunian. Menurut Peraturan Menteri Negara Perumahan Rakyat Republik Indonesia nomor: 22/Permen/M/2008 tentang standar pelayanan minimal bidang perumahan rakyat daerah provinsi dan daerah kabupaten/kota menyatakan bahwa rumah layak huni adalah rumah yang memenuhi persyaratan keselamatan bangunan dan kecukupan minimum luas bangunan serta keselamatan penghuninya.

Beberapa kriteria sudah ditentukan untuk melihat rumah yang layak huni, sehingga diperlukan pengujian untuk mengetahui yang layak dan mana yang tidak layak untuk dihuni oleh masyarakat. Metode pengklasteran dengan algoritma *Fuzzy C-Means* akan digunakan untuk menentukan kelayakan rumah tersebut.

Pada penelitian ini, kelayakan rumah masyarakat di Desa Wayame ditentukan menggunakan Metode pengklasteran dengan algoritma *Fuzzy C-Means* berdasarkan kriteria-kriteria yang telah ditentukan. Dengan demikian, tujuan dari penelitian ini adalah untuk menentukan kelayakan rumah masyarakat di Desa Wayame menggunakan Metode Pengklasteran dengan algoritma *Fuzzy C-Means* berdasarkan kriteria-kriteria yang telah ditentukan.

2. Tinjauan Pustaka

Fuzzy C-Means (FCM) adalah suatu teknik pengklasteran data yang mana keberadaan tiap-tiap titik data dalam suatu klaster ditentukan oleh derajat keanggotaan. Teknik ini pertama kali diperkenalkan oleh Jim Bezdek pada tahun 1981. *Fuzzy C-Means* memungkinkan pengelompokan dimana kelompok data tidak terdistribusi secara jelas [1].

Pengelompokan data dengan metode ini sebelumnya sudah dibuat dan digunakan, namun dengan aplikasi yang berbeda bahkan menggunakan program yang dapat mempermudah. Beberapa yang pernah dibuat adalah:

Metode Klastering dengan Algoritma Fuzzy C-Means untuk Rekomendasi pemilihan bidang keahlian pada program studi teknik informatika [2]. Dalam penelitian ini Algoritma Fuzzy C-Means untuk rekomendasi penjurusan dapat diterapkan dengan manfaat penentu mahasiswa masuk kelompok mana dan setiap kelompok diidentifikasi sebagai jurusan apa.

Kemudian Analisis Penyandang Masalah Kesejahteraan Sosial di Indonesia Menggunakan Metode Fuzzy C-Means Klastering dan Biplot [3]. Dalam penelitian ini Fuzzy C-Means pengklasteran digunakan untuk mengelompokkan data ke dalam klaster tertentu sedangkan biplot untuk memberikan penerangan ke dalam plot.

2.1. Definisi Pengklasteran

Pengklasteran adalah suatu metode pengelompokan berdasarkan ukuran kedekatan (kemiripan). Pengklasteran membagi data menjadi kelompok-kelompok atau klusters berdasarkan suatu kemiripan atribut-atribut diantara data tersebut. Karakteristik tiap klaster tidak ditentukan sebelumnya, melainkan tercermin dari kemiripan data yang terkelompok di dalamnya. Beberapa teknik pengklasteran dalam data mining meliputi: skalabilitas, kemampuan untuk menangani tipe atribut yang berbeda, menangani data yang mengandung *noise*, mampu menangani dimensionalitas yang tinggi, dan dapat diterjemahkan dengan mudah. Pengklasteran beda dengan grup, kalau grup berarti kelompok yang sama, kondisinya kalau tidak ya pasti bukan kelompoknya. Tetapi kalau klaster tidak harus sama akan tetapi pengelompokannya berdasarkan pada kedekatan dari suatu karakteristik sampel yang ada.

2.2. Fuzzy C-Means

Ada beberapa algoritma klastering data, salah satu diantaranya adalah Fuzzy C-Means (FCM). Fuzzy C-Means (FCM) adalah suatu teknik pengklasteran data yang mana keberadaan tiap-tiap titik data dalam suatu klaster ditentukan oleh derajat keanggotaan. Teknik ini pertama kali diperkenalkan oleh Jim Bezdek pada tahun 1981. Fuzzy C-Means memungkinkan pengelompokan dimana kelompok data tidak terdistribusi secara jelas.

Fuzzy berarti sesuatu yang bersifat kabur dan tidak jelas, dimana logika fuzzy itu sendiri berarti mengalokasikan nilai kebenaran dari sebuah variabel ke dalam rentang nilai 0 dan 1 agar berbeda dengan logika tradisional yang mengalokasikan nilai kebenaran tepat pada nilai 0 atau 1. C disini berarti banyaknya konstanta klaster yang diinginkan, oleh sebab itu pada algoritma ini banyaknya jumlah klaster yang ingin digunakan harus ditentukan di awal. Means diartikan sebagai nilai rerata dalam suatu kelompok data, dalam hal ini didefinisikan sebagai klaster. Secara harafiah jika ketiganya digabungkan maka dapat diartikan bahwa algoritma ini menghitung nilai rerata.

Konsep dasar Fuzzy C-Means (FCM), pertama kali adalah menentukan pusat klaster, yang akan menandai lokasi rata-rata untuk tiap-tiap klaster. Pada lokasi awal, pusat klaster ini masih belum akurat. Tiap-tiap titik data memiliki derajat keanggotaan untuk tiap-tiap klaster. Dengan cara memperbaiki pusat klaster dan derajat keanggotaan tiap-tiap titik data secara berulang, maka akan dapat dilihat bahwa pusat klaster akan

bergerak menuju lokasi yang tepat. Perulangan ini didasarkan pada minimisasi fungsi obyektif yang menggambarkan jarak dari titik data yang diberikan ke pusat kluster yang terbobot oleh derajat keanggotaan titik data tersebut.

2.3. Algoritma Fuzzy C-Means

Algoritma Fuzzy C-Means (FCM) adalah sebagai berikut.

- a. Input data yang akan dikluster X , berupa matriks berukuran $n \times m$ (n = jumlah sampel data, m = atribut setiap data). X_{ij} = data sampel ke- i , atribut ke- j .

$$\begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \dots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix}$$

- b. Tentukan:

- 1) Jumlah kluster = c
- 2) Pangkat = w
- 3) Maksimum iterasi = MaxIter
- 4) Error terkecil yang diharapkan = ξ
- 5) Fungsi obyektif awal = $P_0 = 0$
- 6) Iterasi awal = $t = 1$.

- c. Bangkitkan bilangan random μ_{ik} , $i = 1, 2, \dots, n$; $k = 1, 2, \dots, c$; sebagai elemen-elemen matriks partisi awal U . Dengan nilai μ_{ik} diantara 0 dan 1 dan jumlah setiap barisnya sama dengan 1.

Hitung jumlah setiap kolom:

$$Q_i = \sum_{k=1}^c \mu_{ik},$$

dengan $j = 1, 2, \dots, n$.

Hitung : $\mu_{ik} = \frac{\mu_{ik}}{Q_i}$

- d. Hitung pusat kluster ke- k : V_{kj} , dengan $k = 1, 2, \dots, c$; dan $j = 1, 2, \dots, m$.

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

- e. Hitung fungsi obyektif pada iterasi ke- t , P_t .

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

- f. Hitung perubahan matriks partisi:

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

dengan $i = 1, 2, \dots, n$ dan $k = 1, 2, \dots, c$.

- g. Cek kondisi berhenti:

- 1) Jika $(|P_t - (P_t - 1)| < \xi)$ atau $(t > \text{MaxIter})$ maka berhenti;
- 2) Jika tidak : $t = t + 1$, ulangi langkah ke-4 (langkah d).

2.4. Kelayakan Rumah

Rumah adalah tempat untuk tumbuh dan berkembang, baik secara jasmani, rohani dan sosial. Definisi ini membawa banyak konsekuensi yakni bahwa selain kualitas rumah yang harus baik, diperlukan pula segala fasilitas yang dibutuhkan untuk tumbuh dan berkembang. Fasilitas itu misalnya fasilitas pendidikan, pasar/toko, tempat kerja, fasilitas air bersih dan sanitasi.

Berdasarkan Undang-Undang No. 4 Tahun 1992 tentang Perumahan dan Permukiman [4], disebutkan bahwa permukiman adalah bagian dari lingkungan hidup diluar kawasan hutan lindung, baik yang berupa kawasan perkotaan maupun pedesaan yang berfungsi sebagai lingkungan tempat tinggal atau lingkungan hunian dan tempat kegiatan yang mendukung perikehidupan dan penghidupan. Dan berdasarkan Peraturan Pemerintah No. 80 tahun 1999 tentang kawasan siap bangun dan lingkungan siap bangun berdiri sendiri [5], rumah layak huni adalah rumah yang memenuhi persyaratan kesehatan, keselamatan dan kenyamanan. Menurut Johan Silas, rumah disebut layak bila memenuhi aspek sehat, aman, terjamin, dapat dicapai dan mampu dibayar, termasuk kebutuhan dasar, bebas diskriminasi dan kepastian kepemilikannya. Rumah akan menjadi tempat tinggal yang aman dan nyaman, bila memiliki kualitas bangunan yang baik, lengkap dengan fasilitasnya, serta berada dalam lingkungan yang bersih dan sehat. Semakin baik kondisi dan kualitas rumah yang ditempati menunjukkan semakin baik keadaan sosial ekonomi rumah tangga.

Secara umum rumah dapat dikatakan layak huni apabila memiliki jenis lantai, dinding dan atap yang memenuhi syarat atau memiliki kualitas yang baik, serta mempunyai luas lantai yang mencukupi. Selain itu, rumah layak huni juga ditentukan oleh fasilitas penerangan, air minum, dan tempat pembuangan akhir kotoran/tinja.

Belum ada kriteria khusus untuk menentukan rumah dikatakan layak, namun menurut Kementerian Pekerjaan Umum dan Perumahan Rakyat, definisi rumah tidak layak huni adalah sebagai berikut:

1. Bahan lantai berupa tanah atau kayu kelas IV
2. Bahan dinding berupa bilik bambu/kayu/rotan atau kayu kelas IV
3. Bahan atap berupa daun atau genteng plentong yang sudah rapuh
4. Rusak berat, dan/atau
5. Rusak sedang dan luas bangunan tidak mencukupi standar minimal luas per anggota keluarga yaitu 9m²

Pada penelitian ini, yang akan menjadi kriteria penilaian rumah hunian yang layak yaitu ukuran rumah, jenis lantai, jenis dinding, jenis atap, kepemilikan rumah, sumber air minum, sumber penerangan dan keberadaan jamban.

Berdasarkan [6], untuk menentukan kelayakan digunakan persamaan:

$$\text{Kelayakan} = \left(\frac{\left(\frac{\sum \text{nilai yang mendekati layak}}{\sum \text{nilai per cluster}} \right) \times 100\%}{\sum \text{kriteria}} \right).$$

2.5. Validitas Fuzzy C-Means

Bezdek [1] mengusulkan validitas dengan menghitung koefisien partisi atau *partition coefficient* (PC) sebagai evaluasi nilai keanggotaan data pada setiap kluster. Nilai PC Index (PCI) hanya mengevaluasi nilai derajat keanggotaan, tanpa memandang nilai vektor (data) yang biasanya mengandung informasi geometrik (sebaran data). Nilainya dalam rentang [0,1], nilai yang semakin besar (mendekati 1) mempunyai arti bahwa kualitas kluster yang didapat semakin baik. Berikut formula untuk menghitung PC Index:

$$PCI = \frac{1}{N} \left(\sum_{i=1}^N \sum_{j=1}^K u_{ij}^2 \right),$$

dimana N merupakan jumlah data dalam set data, K merupakan jumlah kluster, sedangkan u_{ij} menyatakan nilai keanggotaan data ke- i pada kluster ke- j .

Bezdek [7] sebelumnya telah mengusulkan validitas dengan menghitung entropi partisi atau *partition entropy* (PE). Nilai PE Index (PEI) mengevaluasi keteracakan data dalam kluster. Nilainya dalam rentang [0,1],

nilai yang semakin kecil (mendekati 0) mempunyai arti bahwa kualitas kluster yang didapat semakin baik. Berikut formula untuk menghitung PE Index:

$$PEI = -\frac{1}{N} \left(\sum_{i=1}^N \sum_{j=1}^K u_{ij} \times \log_2 u_{ij} \right).$$

Kedua metrik PCI dan PEI memiliki kecenderungan monotonik terhadap K. Modifikasi nilai PCI (MPCI) dilakukan oleh Dave (1996) terhadap kecenderungan monotonik tersebut. Formula yang digunakan seperti berikut:

$$MPCI = 1 - \frac{K}{K-1} (1 - PCI).$$

Nilai MPCI yang didapat adalah $0 \leq MPCI \leq 1$. Nilai MPCI ekuivalen dengan Non-Fuzziness Index (NFI).

Fukuyama dan Sugeno [8] mengusulkan validitas dengan formula seperti pada persamaan berikut:

$$FSI = \sum_{j=1}^K \sum_{i=1}^N u_{ij}^m \times d(x_i, c_j)^2 - \sum_{j=1}^K \sum_{i=1}^N u_{ij}^m \times d(c_j, \bar{x})^2 = J_m(u, c) - K_m(u, c);$$

dimana m merupakan bobot pangkat (*weighting exponent*), nilainya $m > 1$. $d(x_i, c_j)$ merupakan jarak antara data ke- i terhadap centroid kluster ke- j . c_j adalah centroid kluster ke- j . $d(c_j, \bar{x})$ merupakan jarak antara centroid hasil klustering terhadap rata-rata semua data. $J_m(u, c)$ adalah nilai fungsi obyektif yang mengukur kohesi, sedangkan $K_m(u, c)$ adalah nilai obyektif yang mengukur nilai separasi. Secara umum nilai Fukuyama Sugeno Index (FSI) yang semakin kecil mempunyai arti bahwa kualitas kluster yang didapat semakin baik [9].

Validitas untuk mengevaluasi kluster yang dimodifikasi oleh Pal dan Bezdek [10] diberikan oleh formula berikut:

$$XBI = \frac{\sum_{j=1}^K \sum_{i=1}^N u_{ij}^m \times d(x_i, c_j)^2}{N \times \min_{i,j} (d(c_i, c_j)^2)} = \frac{J_m(u, c)}{Sep(c)}$$

$J_m(u, c)$ adalah ukuran kohesi, sedangkan $Sep(c)$ adalah ukuran separasi.

Secara umum, nilai yang terbaik untuk Xie Beni Index (XBI) adalah nilai index yang semakin kecil. Nilai XBI yang semakin kecil mempunyai arti kualitas hasil pengelompokkan yang semakin baik [9].

Wu dan Yang [9] mengusulkan cara mengevaluasi kluster dengan menghitung *Partition Coefficient and Exponential Separation* (PCAES) Index. PCAES Index (PCAESI) untuk kluster ke- i didefinisikan sebagai gabungan antara kohesi dan separasi kluster tersebut. Ukuran kohesi kluster ke- j relatif terhadap kohesi keseluruhan kluster diukur terhadap u_M , seperti pada persamaan berikut:

$$Koh_j = \sum_{i=1}^N \frac{u_{ij}^2}{u_M}$$

Sementara u_M didapatkan dari persamaan berikut:

$$u_M = \min_{1 \leq j \leq K} \left\{ \sum_{i=1}^N u_{ij}^2 \right\}$$

Separasi kluster ke- j terhadap kluster lain yang terdekat relatif terhadap separasi semua kluster diukur terhadap β_T , seperti pada persamaan berikut:

$$Sep_j = \exp \left(\frac{\min_{k \neq j} \{d(c_j, c_k)^2\}}{\beta_T} \right), \text{ untuk } k = 1, \dots, K$$

Untuk β_T dinyatakan oleh persamaan berikut:

$$\beta_T = \frac{1}{K} \sum_{j=1}^K d(c_j, \bar{x})^2;$$

Untuk menghitung PCAESI pada kluster ke- j dirumuskan sebagai berikut:

$$PCAESI_j = Koh_j - Sep_j.$$

Nilai $PCAESI_j$ yang besar berarti kluster ke- j bersifat kohesif (kompak) didalam dan terpisah dari (K-1) kluster yang lain. Nilai yang kecil atau negatif menunjukkan bahwa kluster ke- j dikenali sebagai kluster yang kurang baik [9].

Nilai kohesi total semua kluster didapatkan dengan menjumlahkan semua nilai kohesi dari setiap kluster, seperti pada persamaan berikut:

$$Koh = \sum_{j=1}^K Koh_j;$$

Sementara separasi total semua kluster didapatkan dengan menjumlahkan semua nilai separasi dari setiap kluster seperti pada persamaan berikut:

$$Sep = \sum_{j=1}^K Sep_j.$$

Validitas total dalam PCAES Index didefinisikan oleh persamaan berikut:

$$PCAESI = \sum_{j=1}^K PCAESI_j = Koh - Sep = \sum_{j=1}^K Koh_j - \sum_{j=1}^K Sep_j$$

Nilai PCAESI yang besar berarti K kluster bersifat kohesif (kompak) dan terpisah satu sama lain. Nilai PCAESI yang kecil berarti ada beberapa K kluster yang tidak kompak atau terpisah dari yang lain. Nilai maksimal PCAESI, yaitu K, dapat digunakan untuk mendeteksi struktur kluster data dengan partisi kompak dan terpisah dengan baik [9].

3. Metode Penelitian

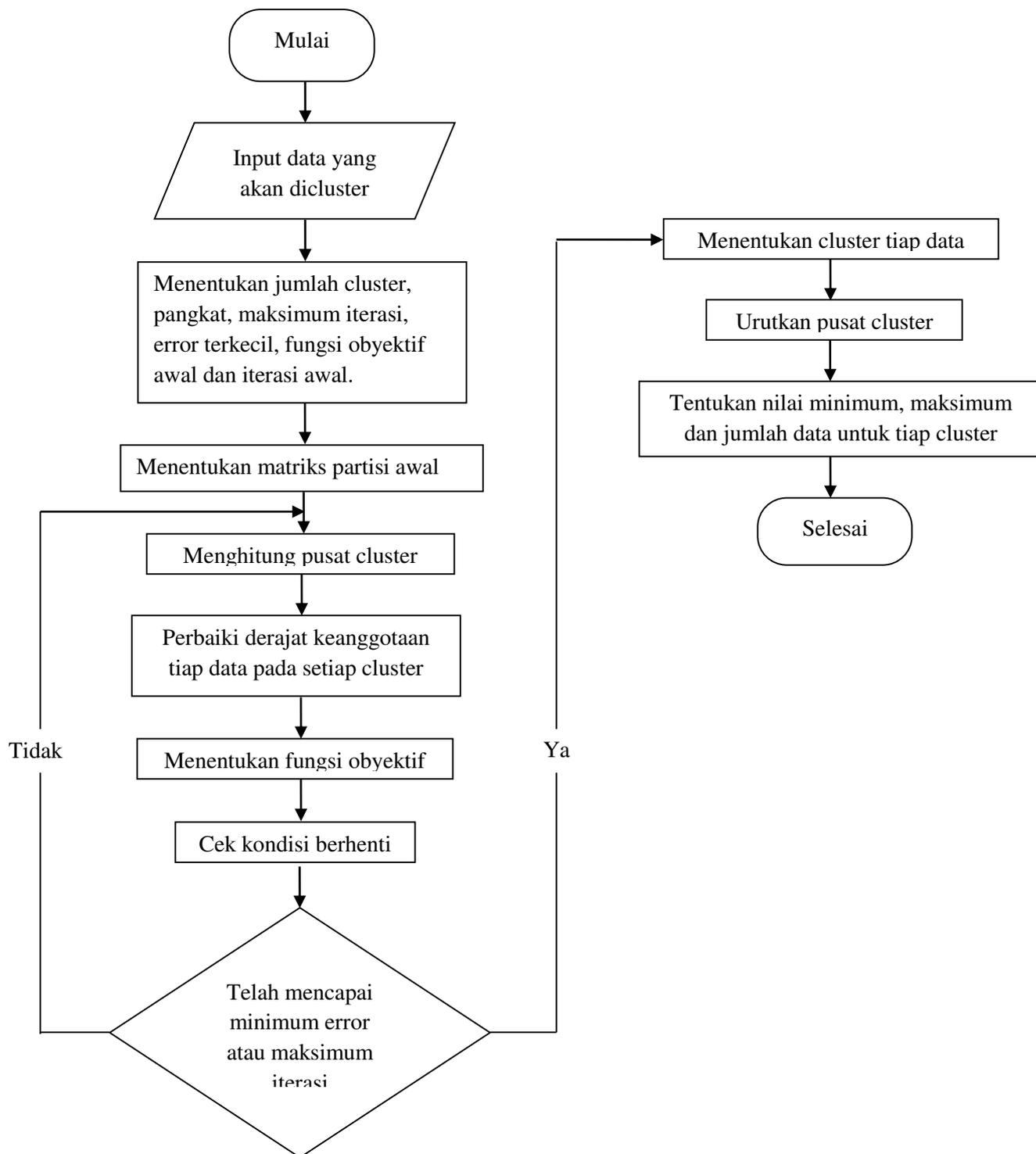
Penelitian ini menggunakan studi kasus yaitu dengan menerapkan metode pengklasteran dengan Algoritma Fuzzy C-Means berdasarkan kriteria-kriteria yang mempengaruhi kelayakan rumah hunian. Pada penelitian ini yang dijadikan unit observasi adalah Desa Wayame, Kota Ambon.

Data yang digunakan dalam penelitian ini adalah data primer yang diambil secara random dari rumah masyarakat setempat. Dalam penelitian ini variabel yang digunakan untuk penentuan kelayakan rumah hunian menggunakan algoritma Fuzzy C-Means adalah sebagai berikut:

- a. X_1 : Atap rumah
- b. X_2 : Dinding rumah
- c. X_3 : Lantai rumah
- d. X_4 : Keberadaan jamban
- e. X_5 : Kepemilikan rumah
- f. X_6 : Ukuran rumah
- g. X_7 : Sumber air minum
- h. X_8 : Sumber penerangan

4. Prosedur Penelitian

Prosedur yang dilakukan dalam penelitian ini berkaitan dengan tujuan penelitian adalah menentukan kelayakan rumah hunian berdasarkan kriteria-kriteria yang telah ditetapkan dengan langkah-langkah sebagai berikut:



Gambar 1. Diagram Alur Algoritma Fuzzy C-Means

5. Hasil dan Pembahasan

Bagian ini memuat tentang analisis data dan hasil pengklasteran menggunakan metode Fuzzy C-Means dengan menggunakan software MATLAB. Data yang digunakan dalam penelitian ini, diambil secara random pada masyarakat di Desa Wayame (dua dusun yaitu dusun Keranjang dan Dusun Waringin Cap). Untuk menentukan kelayakan rumah, prosesnya dimulai dari pengumpulan data rumah yang didasarkan atas kriteria-kriteria penilaian rumah. Kriteria-kriteria yang digunakan untuk melakukan penilaian tersebut adalah:

a. Jenis Atap

Jenis atap terluas yang digunakan pada rumah tersebut dibagi atas tiga kategori yaitu: Daun/rumbia (1), Seng/abses (2), dan genteng (3).

b. Jenis dinding

Jenis dinding dibagi atas tiga kategori yaitu: Tidak permanen (1), semi permanen (2) dan permanen (3).

c. Jenis Lantai

Jenis lantai dibagi atas tiga kategori yaitu: Lantai masih tanah (1), lantai menggunakan semen atau papan (2) dan lantai menggunakan ubin (3).

d. Keberadaan Jamban

Keberadaan pada tiap rumah huni dibagi atas tiga kategori yaitu: jamban umum (1), jamban bersama (2) dan jamban sendiri (3).

e. Status Kepemilikan rumah

Status kepemilikan rumah dibagi atas dua kategori yaitu: bukan pemilik (1) dan pemilik (2).

f. Luas bangunan rumah

Luas bangunan rumah (m^2) tidak dikategorikan.

g. Sumber air minum

Sumber air minum dibagi atas 3 kategori yaitu air hujan/air sungai (1), sumur/pompa (2), dan ledeng/kemasan (3).

h. Sumber penerangan

Sumber penerangan dibagi atas 3 kategori yaitu lampu minyak (1), genset/disel (2), dan listrik (3).

Berdasarkan hasil pengumpulan data dari 8 kriteria di atas, ternyata hanya 5 kriteria yang berpengaruh dalam proses pengklasteran. Kelima kriteria tersebut yaitu jenis atap, jenis dinding, jenis lantai, keberadaan jamban dan luas bangunan rumah. Dengan demikian, kriteria sumber penerangan, sumber air minum dan kepemilikan rumah tidak berpengaruh karena hasil yang diperoleh saat menggunakan 8 kriteria di atas sama dengan hasil yang diperoleh menggunakan 5 kriteria (jenis atap, jenis dinding, jenis lantai, keberadaan jamban dan luas bangunan rumah). Hal ini dapat terjadi karena semua observasi memiliki kesamaan dalam 3 kriteria tersebut (sumber penerangan, sumber air minum dan kepemilikan rumah).

Langkah awal dari proses pengklasteran ini adalah menentukan parameter awal yang akan digunakan untuk menyelesaikan masalah dengan algoritma Fuzzy C-means. Parameter-parameter tersebut adalah

- | | |
|---|-------------|
| a. Jumlah klaster (c) | = 2 |
| b. Pangkat (w) | = 2 |
| c. Maksimum iterasi (MaxIter) | = 100 |
| d. Error terkecil yang diharapkan (ξ) | = 10^{-5} |
| e. Fungsi objektif awal (P_0) | = 0 |

Beberapa hal penting yang diperoleh dalam proses pengklasteran ini adalah nilai fungsi obyektif selama iterasi, pusat klaster serta derajat keanggotaan rumah setiap klaster pada iterasi terakhir.

Tabel 1. Nilai obyektif selama 14 iterasi

| Iterasi ke- | Nilai Obyektif |
|-------------|----------------|
| 1 | 7440,00566 |
| 2 | 6148,68789 |
| 3 | 4827,08567 |
| 4 | 2986,68849 |
| 5 | 2631,28815 |
| 6 | 2614,98795 |
| 7 | 2614,48678 |
| 8 | 2614,45738 |
| 9 | 2614,45252 |
| 10 | 2614,45142 |
| 11 | 2614,45116 |
| 12 | 2614,45110 |
| 13 | 2614,45109 |
| 14 | 2614,45108 |

Nilai fungsi obyektif yang diperoleh dalam penelitian ini didasarkan atas proses iterasi. Pada Tabel 1, dapat dilihat bahwa pada iterasi pertama diperoleh nilai fungsi obyektif 7440,00566. Proses iterasi ini terus berjalan hingga iterasi ke-14 sehingga diperoleh nilai fungsi obyektif adalah 2614,45108. Proses iterasi berhenti pada iterasi ke-14 dimana nilai $|P_t - P_{t-1}| < \xi$.

Setelah diperoleh nilai fungsi obyektif, maka dapat diperoleh nilai pusat kluster atau *centriod* yang akan digunakan. Pada iterasi terakhir (iterasi ke-14), diperoleh nilai-nilai pusat kluster yang diberikan dalam bentuk matriks V_{kj} . Untuk $k = 1, 2$ dan $j = 1, 2, 3, 4, 5$, nilai pusat kluster adalah

$$V_{kj} = \begin{pmatrix} 1,9989 & 2,9812 & 2,5257 & 2,9378 & 68,2309 \\ 1,8341 & 2,3606 & 2,0676 & 2,3933 & 36,5991 \end{pmatrix}$$

Berdasarkan hasil pengolahan data maka dapat diberikan derajat keanggotaan tiap observasi untuk masing-masing kluster pada iterasi terakhir (iterasi ke-14).

Tabel 2. Derajat Keanggotaan pada Iterasi 14

| Observasi ke-i | μ_{i1} | μ_{i2} |
|----------------|------------|------------|
| 1 | 0,0291 | 0,9709 |
| 2 | 0,2414 | 0,7586 |
| 3 | 0,0025 | 0,9975 |
| 4 | 0,5982 | 0,4018 |
| 5 | 0,0293 | 0,9707 |
| 6 | 0,9970 | 0,0030 |
| 7 | 0,9620 | 0,0380 |
| \vdots | \vdots | \vdots |
| 42 | 0,9970 | 0,0030 |

Dari derajat keanggotaan pada iterasi terakhir tersebut dapat diperoleh informasi mengenai kecenderungan untuk setiap observasi masuk ke kluster mana. Derajat keanggotaan terbesar menunjukkan bahwa kecenderungan tertinggi observasi untuk masuk mejadi anggota kluster tertentu.

Pada observasi pertama, nilai derajat keanggotaan untuk kluster pertama 0,0291 sedangkan nilai derajat keanggotaan untuk kluster kedua 0,9709. Dari nilai tersebut observasi pertama masuk dalam kluster kedua. Hal itu dikarenakan observasi pertama mempunyai derajat keanggotaan tertinggi di kluster kedua daripada kluster pertama.

Selanjutnya, pada observasi kedua nilai derajat keanggotaan untuk kluster pertama 0,2414 sedangkan nilai derajat keanggotaan untuk kluster kedua 0,7586. Dari nilai tersebut observasi kedua masuk dalam kluster kedua.

Penentuan berlanjut hingga observasi ke-42, dengan nilai derajat keanggotaan untuk kluster pertama 0,9970 sedangkan nilai derajat keanggotaan untuk kluster kedua 0,0030. Dari nilai tersebut observasi ke-42 masuk dalam kluster pertama. Kecenderungan kluster secara keseluruhan dapat dilihat dalam Tabel 3.

Tabel 3. Kecenderungan Kluster

| Observasi ke- | Kluster 1 | Kluster 2 | Observasi ke- | Kluster 1 | Kluster 2 |
|---------------|-----------|-----------|---------------|-----------|-----------|
| 1 | | * | 22 | * | |
| 2 | | * | 23 | * | |
| 3 | | * | 24 | | * |
| 4 | * | | 25 | | * |
| 5 | | * | 26 | | * |
| 6 | * | | 27 | * | |
| 7 | * | | 28 | * | |
| 8 | * | | 29 | | * |
| 9 | * | | 30 | * | |
| 10 | * | | 31 | | * |
| 11 | | * | 32 | * | |
| 12 | * | | 33 | * | |
| 13 | | * | 34 | | * |
| 14 | * | | 35 | | * |
| 15 | * | | 36 | * | |
| 16 | * | | 37 | | * |
| 17 | * | | 38 | * | |
| 18 | * | | 39 | * | |
| 19 | | * | 40 | * | |
| 20 | | * | 41 | | * |
| 21 | | * | 42 | * | |

Setelah didapatkan kluster, maka dilanjutkan dengan memasukkan bobot yang akan menentukan persentase kelayakan, berikut langkah-langkahnya:

- 1) Jenis atap = 2-3
- 2) Jenis dinding = 3
- 3) Jenis lantai = 3
- 4) Keberadaan jamban = 3
- 5) Luas rumah = 45m²

$$\text{Kelayakan} = \left(\frac{\left(\frac{\sum \text{nilai yang mendekati layak}}{\sum \text{nilai per cluster}} \right) \times 100\%}{\sum \text{kriteria}} \right)$$

Pada kluster pertama, perhitungan persentase untuk menentukan kelayakan rumah untuk setiap kriteria yaitu: berdasarkan jenis atap 100%, berdasarkan jenis dinding 95,833%, berdasarkan jenis lantai 54,167%, berdasarkan keberadaan jamban 91,667%, berdasarkan luas rumah 100%. Rata-rata persentase pada kluster pertama adalah 90,278%. Sehingga persentase kelayakan rumah hunian yang termasuk pada kluster pertama yaitu 90,278%.

Sedangkan pada klaster kedua, perhitungan persentase untuk menentukan kelayakan rumah untuk setiap kriteria yaitu: berdasarkan jenis atap 83,333%, berdasarkan jenis dinding 44,444%, berdasarkan jenis lantai 5,556%, berdasarkan keberadaan jamban 50%, berdasarkan luas rumah 16,667%. Rata-rata persentase pada klaster kedua adalah 50%. Sehingga persentase kelayakan rumah hunian yang termasuk pada klaster kedua yaitu 50%.

Dari hitungan di atas, diambil nilai persentase yang terbesar untuk mendapatkan kelayakan rumah hunian. Sehingga yang termasuk ke dalam kategori layak adalah klaster 1 dengan kelayakan 90%. Dengan hasil lengkap dari pengklasteran rumah layak dan tidak layak dapat dilihat pada Tabel 4.

Tabel 4. Hasil Pengklasteran

| Klaster | Observasi ke- |
|-------------------|--|
| 1 (kelayakan 90%) | 4,6,7,8,9,10,12,14,15,16,17,18,22,23,27,28,30,32,33,36,38,39,40,42 |
| 2 (kelayakan 50%) | 1,2,3,5,11,13,19,20,21,24,25,26,29,31,34,35,37,41 |

Dari tabel di atas dapat dilihat bahwa masih terdapat 18 rumah dengan kelayakan 50% dari 42 rumah yang menjadi observasi. Dengan persentase rumah yang masuk klaster 1 adalah 57,14% dan yang masuk klaster 2 adalah 42,86%. Sehingga masih diperlukannya perhatian khusus untuk melihat kondisi perumahan di daerah tersebut.

Validitas Pengklasteran Menggunakan Fuzzy

Metode pengelompokan yang menggunakan konsep fuzzy, sebuah data bisa menjadi anggota di semua klaster dengan nilai derajat keanggotaan yang dimilikinya. Semakin tinggi nilai derajat keanggotaan pada sebuah klaster maka semakin besar kecenderungan menjadi anggota klaster tersebut.

Dari pengolahan data menggunakan Matlab diperoleh hasil validitas Fuzzy seperti pada Tabel 5.

Tabel 5. Validitas dengan Jarak *Euclidean* dan *City Block*

| Jarak | <i>Euclidean</i> | <i>City Block</i> |
|--------------|-------------------|-------------------|
| PCI | 0,8614 | 0,8614 |
| PEI | 0,3408 | 0,3408 |
| MPCI | 0,7229 | 0,7229 |
| FSI | -6,3955e+003 | -6,8651e+003 |
| XBI | 0,0622 | 0,0679 |
| PCAESI | 1,1452 dan 0,9810 | 1,1452 dan 0,9810 |
| PCAESI total | 2,1262 | 2,1262 |

Dari hasil validasi, dapat dilihat bahwa nilai yang didapat dengan kedua jarak yang digunakan hampir sama. Dengan penjelasan:

- Nilai PCI (*partition coefficient index*) yaitu 0,8614 (mendekati 1) sama pada kedua jarak mempunyai arti bahwa kualitas klaster yang didapat baik.
- Nilai PEI (*partition entropy index*) yaitu 0,3408 (mendekati 0) sama pada kedua jarak mempunyai arti bahwa kualitas klaster yang didapat baik.
- Nilai MPCI yaitu 0,7229 (mendekati 1) sama pada kedua jarak yang digunakan mempunyai arti bahwa kualitas klaster yang didapat baik`
- Nilai FSI hampir sama pada kedua jarak yaitu -6,3955e+003 untuk *Euclidean* dan -6,8651e+003 untuk *City Block*, mempunyai arti bahwa kualitas klaster yang didapat baik.
- Nilai XBI yaitu 0,0622 untuk *Euclidean* dan 0,0679 untuk *City Block*, mempunyai arti bahwa kualitas hasil pengelompokan baik.

- Nilai PCAESI untuk klaster ke 1 dan klaster ke 2 yaitu 1,1452 dan 0,9810, karena nilainya cukup besar dan tidak negatif maka berarti masing-masing klaster bersifat kohesif (kompak) di dalam dan terpisah dari klaster yang lain.
- Nilai PCAESI total yaitu 2,1262 untuk kedua jarak yang digunakan, mempunyai arti bahwa klaster bersifat kohesif (kompak) dan terpisah dari klaster yang lain.

Sehingga dapat disimpulkan bahwa 2 klaster baik untuk digunakan pada pengelompokan data untuk menentukan kelayakan rumah.

6. Kesimpulan

Algoritma *Fuzzy C-Means* merupakan salah satu teknik pengklasteran data yang mana keberadaan tiap-tiap titik data dalam suatu klaster ditentukan oleh derajat keanggotaannya. Dalam penelitian ini dilakukan pengklasteran terhadap 42 sampel rumah dari dua dusun di desa Wayame. Berdasarkan pengklasteran yang dilakukan diperoleh 2 klaster. Dari kedua klaster tersebut diketahui bahwa klaster pertama merupakan klaster yang beranggotakan rumah yang tidak layak yaitu 18 rumah terdiri atas 9 rumah pada Dusun Keranjang dan 9 rumah pada Dusun Waringin Cap, sedangkan klaster kedua merupakan klaster yang beranggotakan rumah yang layak yaitu 24 rumah terdiri atas 13 rumah pada dusun Keranjang dan 11 rumah pada Dusun Waringin Cap.

Daftar Pustaka

- [1] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, New York: Plenum, 1981.
- [2] M. Mirza and Faisal, "Metode Klastering dengan Algoritma Fuzzy C-Means untuk Rekomendasi Pemilihan Bidang Keahlian Pada Program Studi Teknik Informatika," 2013.
- [3] U. Fadilah, "Analisis Penyandang Masalah Kesejahteraan Sosial di Indonesia Menggunakan Metode Fuzzy C-Means Klastering dan Biplot," 2011.
- [4] Undang-Undang Nomor 4 Tahun 1992 Tentang Perumahan dan Permukiman.
- [5] Peraturan Pemerintah No. 80 tahun 1999 Tentang Kawasan Siap Bangun dan Lingkungan Siap Bangun Berdiri Sendiri.
- [6] R. Putra, N. Nasution and Yummastian, "Aplikasi E-Zakat dan Penyaluran Menggunakan Fuzzy C-Means (Studi Kasus: LAZISMU Pekanbaru)," 2015.
- [7] J. C. Bezdek, "Cluster Validity with Fuzzy Set," *J. Cybernet*, vol. 3, no. 3, pp. 58-72, 1974.
- [8] Y. Fukuyama and M. Sugeno, "A New Method Measurements in Taxonomic Problems," *Annals of Eugenics*, vol. 7, no. 11, pp. 179-188, 1989.
- [9] K. L. Wu and M. S. Yang, "A Cluster Validity Index for Fuzzy Clustering," *Pattern Recognition Letters*, vol. 26, pp. 1275-1291, 2005.
- [10] N. R. Pal and J. C. Bezdek, "On Cluster Validity for Fuzzy C-Means Model," *IEEE Transactions of Fuzzy Systems*, vol. 3, no. 3, pp. 370-379, 1995.