

Pembentukan Aturan Fuzzy Untuk Pemberian Rekomendasi Penerima Bantuan Keluarga Berumah Tidak Layak Huni Menggunakan K-means Clustering

Aidil¹, Judi Prajetno², Esther Irawati Setiawan³, and Adi Surya Putra⁴

¹Departemen Informatika, STMIK PPKIA Tarakanita Rahmawati, Tarakan, Indonesia

²Departemen Elektro, Fakultas Sains dan Teknologi, Institut Sains dan Teknologi Terpadu Surabaya, Surabaya, Indonesia

³Departemen Teknologi Informasi, Fakultas Sains dan Teknologi, Institut Sains dan Teknologi Terpadu Surabaya, Surabaya, Indonesia

⁴Departemen Sistem Informasi, Fakultas Sains dan Teknologi, Institut Sains dan Teknologi Terpadu Surabaya, Surabaya, Indonesia

Corresponding author: Esther Irawati Setiawan (e-mail: esther@istts.ac.id).

ABSTRACT Assistance for families whose houses are not livable is one of the social benefits given to families who are experiencing financial difficulties and/or have houses that are not livable. The variables considered when determining beneficiaries often make decisions difficult to make. Therefore, we need a fuzzy reasoning system that automatically generates rules as the expected decision makers. To form fuzzy rules, an expert is needed. An expert is a person who is experienced and able to explain a rule related to a field. In this research, a rule is formed automatically which does not depend on an expert. The fuzzy rules generated can be obtained from several techniques such as the clustering process. The method used in generating this fuzzy rule is the k-means clustering method. K-means clustering is used to group data and generate rules in the case of recommendations for recipients of uninhabitable housing. The results of the generation of fuzzy rules are used for the fuzzy inference process using the Sugeno Fuzzy Inference System method. The Sugeno method produces an output (consequently) in the form of a constant or a linear equation. In this study, 1000 training data were used and a process of testing 300 test data was carried out to obtain recommendations for recipients of uninhabitable housing assistance. The test results are used to assess the accuracy of the rules developed. The study's findings reveal that the results of k-means clustering may automatically produce rules for the production of Sugeno Fuzzy Inference System rules in more than 75% of cases.

KEYWORDS Clustering, Fuzzy, Fuzzy Inference System, Fuzzy Rules, K-Means, Sugeno, Uninhabitable Houses

ABSTRAK Bantuan bagi keluarga yang rumah tidak layak huni merupakan salah satu manfaat sosial yang diberikan kepada keluarga yang mengalami kesulitan keuangan dan/atau memiliki rumah tidak layak huni. Variabel yang dipertimbangkan saat menentukan penerima manfaat sering kali membuat keputusan sulit diambil. Oleh karena itu, diperlukan sistem penalaran fuzzy yang secara otomatis menghasilkan aturan-aturan sebagai pembuat keputusan yang diharapkan. Untuk membentuk aturan fuzzy diperlukan seorang pakar. Pakar adalah seorang ahli yang berpengalaman dalam suatu bidang yang mampu menjelaskan suatu aturan yang terkait dengan suatu bidang. Dalam penelitian ini dibentuk suatu rule secara otomatis yang tidak tergantung dengan seorang pakar. Aturan fuzzy dibangkitkan bisa diperoleh dari beberapa teknik seperti proses clustering. Metode yang digunakan dalam membangkitkan aturan fuzzy ini yaitu metode k-means clustering. Dalam hal rekomendasi penerima bantuan rumah tidak layak huni, K-means clustering digunakan untuk mengelompokkan data dan mengembangkan aturan. Hasil dari pembangkitan aturan fuzzy digunakan

untuk proses inferensi fuzzy menggunakan metode Fuzzy Inference System Sugeno. Metode sugeno menghasilkan output (konsekuen) berupa konstanta atau persamaan linier. Dalam penelitian ini digunakan 1000 data training dan dilakukan proses pengujian 300 data uji untuk mendapatkan rekomendasi penerima bantuan rumah tidak layak huni. Hasil pengujian digunakan untuk mengetahui akurasi aturan yang terbentuk. Dari hasil penelitian menunjukkan bahwa hasil k-means clustering dapat membentuk rule secara otomatis untuk pembangkitan aturan Fuzzy Inference System Sugeno dapat dilihat dari hasil akurasi perhitungan pengujian data uji skenario global sama-sama menghasilkan akurasi minimal di atas 75%.

KATA KUNCI Aturan Fuzzy, Clustering, Fuzzy, Fuzzy Inference System, K-Means, Rumah Tidak Layak Huni, Sugeno

I. PENDAHULUAN

Bantuan keluarga berumah tidak layak huni adalah salah satu bantuan sosial yang dihibahkan kepada keluarga yang menghadapi kesulitan ekonomi dan atau mempunyai rumah yang tidak layak huni [1]. Terdapat berbagai variabel-variabel yang perlu diperhatikan dalam penentuan penerima bantuan menyebabkan kesulitan untuk mengambil keputusan. Untuk itulah diperlukan suatu sistem otomatis inferensi Fuzzy yang membangkitkan aturan sebagai pengambil keputusan.

Dari tantangan ini, muncul ide pemikiran untuk mengembangkan suatu penelitian berbasis fuzzy yang menghasilkan aturan-aturan fuzzy secara otomatis. Untuk membentuk aturan fuzzy diperlukan seorang pakar. Pakar adalah seorang ahli yang berpengalaman dalam suatu bidang yang mampu menjelaskan suatu aturan yang terkait dengan suatu bidang.

Dalam penelitian ini dibentuk suatu rule secara otomatis yang tidak tergantung dengan seorang pakar. Aturan fuzzy dibangkitkan bisa diperoleh dari beberapa teknik seperti proses clustering. Metode yang digunakan dalam membangkitkan aturan fuzzy ini yaitu metode k-means clustering. K-means clustering digunakan untuk mengelompokkan data dan membangkitkan aturan pada kasus rekomendasi penerima bantuan keluarga berumah tidak layak huni. Hasil dari pembangkitan aturan fuzzy digunakan untuk memproses inferensi fuzzy.

Langkah yang diambil adalah membuat aplikasi pendukung keputusan yang diharapkan dapat mempermudah untuk melakukan seleksi rekomendasi penerima bantuan keluarga berumah tidak layak huni di dinas sosial dan tenaga kerja kota Tarakan. Dalam makalah ini, penentuan penerimaan bantuan keluarga berumah berdasarkan beberapa kriteria diantaranya umur, penghasilan, jumlah tanggungan (orang) dan kondisi rumah (lantai, dinding, atap dan WC).

Selain itu, penelitian yang diusulkan hanya menentukan seleksi rekomendasi penerima bantuan berumah tidak layak huni tanpa menentukan nilai nominal bantuannya. Sedangkan data yang digunakan pada proses pelatihan adalah data keluarga berumah tidak layak huni tahun 2012 sebanyak 500 data dan 2013 sebanyak 500 data. Data yang digunakan untuk pengujian sistem adalah data keluarga berumah tidak layak huni 2014 sebanyak 300 data. Data yang di ambil sebagai data uji dari kota Tarakan.

Pada makalah ini, clustering data diproses menggunakan K-means dengan analisis varian untuk menentukan cluster terbaik dan pengujian datanya menggunakan sistem inferensi fuzzy model Sugeno. Makalah ini bertujuan untuk mengimplementasikan sistem yang dapat memberikan rekomendasi penerima bantuan keluarga berumah tidak layak huni. Selain itu, diharapkan penelitian ini dapat mempermudah dan mempercepat dalam proses pencarian keluarga yang direkomendasikan mendapat bantuan keluarga berumah tidak layak huni berdasarkan kebutuhan kuota penerima.

Target lainnya yang ingin dicapai adalah meminimalisir berkurangnya kuota penerima bantuan berumah tidak layak huni di tahun berikutnya karena kesalahan dalam seleksi pemilihan. Hasil penelitian ini akan dibandingkan dengan hasil seleksi pemberian bantuan keluarga berumah tidak layak huni yang dilaksanakan oleh Departemen Sosial dan Tenaga Kerja

Kontribusi penelitian ini adalah :

1. Membantu para pegawai dinas sosial dan tenaga kerja untuk menentukan pemberian bantuan keluarga berumah tidak layak huni sesuai dengan kriteria yang di inginkan.
2. Mengetahui sejauh mana kegunaan Fuzzy dalam kehidupan sehari – hari.
3. Sampai dengan saat ini, masih belum ada program aplikasi yang dapat membantu dalam penunjang pengambilan keputusan untuk memberikan bantuan keluarga berumah tidak layak huni.

II. TINJAUAN PUSTAKA

Beberapa hal yang menjadi dasar teori dari penelitian ini yaitu teori program bedah rumah (Rumah Tidak Layak Huni), data mining, k-means clustering, analisa varian, logika fuzzy, fuzzy inference system sugeno.

A. PROGRAM BEDAH RUMAH (RUMAH TIDAK LAYAK HUNI)

Rumah itu adalah kebutuhan manusia mendasar yang berfungsi sebagai lokasi untuk hidup atau dihuni serta metode membesarkan keluarga. Pada dasarnya, setiap individu membutuhkan rumah yang layak, tetapi bagi sebagian orang, memenuhi permintaan untuk perumahan yang layak huni adalah kesulitan.

Selain itu, Rumah adalah salah satu indikator kemakmuran. Kondisi rumah yang layak dihuni mencerminkan kesejahteraan yang tinggal di rumah, sementara kehadiran banyak rumah komunitas yang tidak dapat dihuni menunjukkan bahwa orang-orang di wilayah tersebut tidak kaya. Pemerintah Kota Tarakan secara resmi mempresentasikan hasil program kepada pemilik rumah pada 2012, setelah menyelesaikan program bedah rumah atau rehabilitasi rumah tidak layak dihuni (RTLH).

Layanan Sosial Kota Tarakan, bekerja sama dengan pihak ketiga, meluncurkan program renovasi DPR pada 30 Oktober 2012, dengan target 100 RTLH yang tersebar di empat distrik, yaitu Distrik Tarakan Barat, Tarakan Utara, Tarakan Tengah, dan Tarakan Timur, yang memiliki sejak diperpanjang selama beberapa tahun [2].

Salah satu tanda kemiskinan adalah keadaan rumah yang tidak layak huni. Dengan mengkaji masalah kesejahteraan masyarakat, Pemerintah Kota Tarakan menginisiasi kegiatan pemulihan tempat tinggal yang tidak layak huni bagi masyarakat kurang mampu melalui program New Singapore. Pada tahun 2012 dilakukan perbaikan 1000 unit rumah tidak layak huni.

Dalam APBN 2013, dana yang disiapkan untuk penyediaan rumah layak huni bagi masyarakat kurang mampu mencapai Rp 78 miliar. Program Rehabilitasi Rumah Tidak Layak Huni (RTLH), yang juga dikenal dengan renovasi rumah, merupakan salah satu komponen komitmen pemerintah dalam pengentasan kemiskinan melalui Dinas Sosial. Program renovasi rumah 2013 diluncurkan pada akhir Mei, dengan peletakan batu pertama.

Several features of livable aid have been researched. In the publication [3,] research data was gathered by directly visiting the Mesuji District Office and the poor who got aid to obtain data that corresponded to reality in the field through observations, interviews, and recording. Data is examined using an analytical descriptive technique after collection. Based on the findings of this research, the government-funded livable housing aid program in Mesuji Regency has been carried out correctly and in compliance with the activity's operational procedures. Meanwhile, according to Islamic economics, the root cause of poverty is a failure to meet the requirements of the community.

Ada beberapa metode untuk menentukan penerima manfaat rumah layak huni. Studi ini menggunakan pendekatan VIKOR, dan temuannya diharapkan dapat membantu pemerintah dalam mengelola Dana Bantuan RUTLAHU dan memilih pelamar yang memenuhi persyaratan. Berdasarkan temuan penelitian ini, pendekatan VIKOR dapat mengidentifikasi individu yang layak mendapatkan pembiayaan RUTLAHU dan melakukan pemeringkatan yang efektif [4].

Teknik weighted product (WP) telah diteliti untuk mengatasi tantangan terkait penyediaan bantuan rumah layak huni [5]. Data untuk penelitian ini dikumpulkan dari masyarakat Desa Pacinan per kepala keluarga. Selain itu,

penulis mewawancarai seorang anggota staf kantor desa. Pada bulan Maret 2014, data dikumpulkan.. Dari hasil penelitian dapat didapatkan bahwa metode Weighted Product dapat diimplementasikan ke dalam sistem dan telah dibuktikan pada saat pengujian penelitian.

Metodologi TOPSIS, metode Sistem Pendukung Keputusan yang secara optimal/praktis dapat mendukung proses pengambilan keputusan dengan menggunakan ide-ide sederhana/mudah dipahami, digunakan dalam penelitian [6]. Data dari Desa Sumbaga digunakan dalam proses pemilihan bantuan rumah tidak layak huni, meliputi ciri-ciri dinding, atap, lantai, pekerjaan, pendapatan, jumlah tanggungan, luas rumah dan fasilitas MCK. Keberadaan SPK di Desa Sumbaga, menurut temuan studi, dapat membantu masyarakat menilai kelayakan mereka untuk mendapatkan bantuan RTLH dengan mempercepat dan mempersingkat berbagai proses.

Penelitian ini menggunakan teknik Fuzzy [1]. Dalam penelitian ini Fuzzy Multiple-Attribute Decision-Making (FMADM) digunakan untuk menemukan hasil dari pemilihan dan penghitungan setiap alternatif dengan Simple Additive Weighting (SAW). Pendekatan deskriptif atau survei digunakan dalam penelitian ini untuk mengumpulkan daftar nama yang diusulkan dari desa jamban sebagai data calon penerima bantuan rumah layak huni, serta berbagai kriteria yang akan dijadikan acuan dalam proses pengambilan keputusan. Berdasarkan temuan penelitian, tingkat akurasi metode SAW sebesar 95,44% dan tingkat akurasi metode FMADM sebesar 94,24%.

B. DATA MINING

Data Mining merupakan istilah untuk menggambarkan proses penemuan pengetahuan tersembunyi di dalam basis data. Data mining adalah suatu proses yang semi-otomatis untuk melakukan ekstraksi dan identifikasi informasi pengetahuan yang berpotensi dan berguna untuk disimpan dalam kumpulan data besar menggunakan pendekatan statistik, matematika, kecerdasan buatan, dan pembelajaran mesin [7]. Data mining, menurut Grup Gartner, adalah tindakan menemukan tautan, pola, dan tren yang signifikan dengan menganalisis sejumlah data dengan alat deteksi pola menggunakan perhitungan statistik dan matematis [8].

C. K-MEANS CLUSTERING

Salah satu algoritma clustering adalah K-Means Clustering. Clustering adalah teknik penambangan data yang membagi data menjadi banyak pengelompokan (grup, kluster, atau segmen), yang masing-masing dapat berisi beberapa anggota. Setiap objek ditugaskan ke grup yang paling mirip. Ini analog dengan mengelompokkan hewan dan tumbuhan ke dalam keluarga dengan anggota yang sebanding.

J.B. MacQueen mengembangkan algoritme K-Means pada tahun 1976, yang merupakan salah satu algoritme pengelompokan yang paling banyak digunakan untuk mengelompokkan data berdasarkan properti yang dapat

dibandingkan atau dibagi. Grup data ini dinamakan sebagai cluster. Data di dalam suatu cluster mempunyai ciri-ciri serupa dan tidak serupa dengan data cluster lain. Kompleksitas $O(nKt)$ algoritma K-Means menunjukkan cukup efisien, dengan asumsi n adalah jumlah objek data, k adalah jumlah cluster yang dihasilkan, dan t adalah jumlah iterasi [9]. nilai k dan t biasanya jauh lebih kecil dari nilai n . Selanjutnya, algoritma ini akan berhenti pada situasi optimal selama iterasinya [10].

1) ALGORITMA K-MEANS CLUSTERING

Di bawah menggambarkan bagaimana algoritma K-Means membagi dataset menjadi cluster. [11]:

1. Tentukan k yang merupakan jumlah cluster yang dihasilkan. Untuk menghitung jumlah k cluster, perlu dilakukan pertimbangan teoretis dan kontekstual.
2. Secara acak menghasilkan k Centroid pertama (lokasi pusat cluster). Centroid awal dipilih secara acak dari antara item yang tersedia sebanyak k cluster, dan rumus berikut digunakan untuk menghasilkan centroid cluster ke- i berikutnya:

$$v = \frac{\sum_{i=1}^n x_i}{n}; i = 1, 2, 3, \dots, n \quad (1)$$

Dimana ; v = centroid cluster

x_i = objek yang ke- i

n = banyak objek yang menjadi anggota cluster

3. Tentukan jarak antara setiap objek dan setiap cluster centroid. Penulis menggunakan Euclidean Distance untuk menghitung jarak antara objek dan centroid.

$$d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}; i = 1, 2, 3, \dots, n \quad (2)$$

Dimana ; x_i = objek x yang ke- i

y_i = daya y yang ke- i

n = jumlah objek

4. Setiap objek harus di alokasikan ke centroid terdekat. Lakukan iterasi, lalu temukan posisi centroid baru, untuk menetapkan objek ke setiap cluster selama iterasi.
5. Jika posisi centroid baru tidak sama dengan yang lama, ulangi langkah 3. Konvergensi diperiksa dengan membandingkan penentuan grup matriks pada iterasi sebelumnya dengan penentuan grup matriks pada iterasi saat ini. Jika hasilnya sama, maka algoritma k-means cluster analysis sudah konvergen; jika berbeda, proses belum konvergen, dan iterasi berikutnya diperlukan.

Sebelum proses analisa cluster, terlebih dahulu dilakukan perhitung mean dan standar deviasi untuk data yang sudah dikelompokkan dengan menggunakan persamaan 3 dan persamaan 4.

$$\mu = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} \quad (3)$$

$$\sigma_{ki} = \sqrt{\frac{\sum_{j=1}^n (x_j - \mu_{ki})^2}{n-1}} \quad (4)$$

Dimana ; k = jumlah cluster 1... k

i = atribut 1 ... i

n = banyaknya objek

x_j = data ke- j

\bar{y} = rata-rata cluster

μ = rata-rata cluster

Kepadatan klaster dapat digunakan untuk melakukan analisis klaster (cluster density). *Variance Within Cluster* (V_w) dan *Variance Between Cluster* (V_b) dapat digunakan untuk menghitung kepadatan cluster. Persamaan 5 digunakan untuk menghitung varian dari setiap pembentukan cluster.

$$V_c^2 = \frac{1}{n_c - 1} \sum_{i=1}^n (y_i - \bar{y}_c)^2 \quad (5)$$

Dimana ; V_c = variance cluster c

$c = 1..k$, dengan k adalah banyak cluster

n_c = banyak data cluster c

y_i = data ke- i suatu cluster

\bar{y}_c = rata-rata data suatu cluster

Dengan menggunakan rumus tersebut, kita dapat menentukan *variance within cluster* (V_w) dari varians di atas:

$$V_w = \frac{1}{N - c} \sum_{i=1}^c (n_i - 1) v_i^2 \quad (6)$$

Dimana ; N = total data

$c = 1..k$, dengan k mewakili jumlah cluster

n_i = total data dalam suatu cluster

v_i = varian dalam suatu cluster

Lalu *Variance Between Cluster* (V_b) menggunakan persamaan:

$$V_b = \frac{1}{c-1} \sum_{i=1}^c n_i (\bar{y}_i - \bar{y})^2 \quad (7)$$

Dimana ; \bar{y} = rata-rata dari y_i

Nilai variance digunakan untuk mengidentifikasi cluster ideal melalui perhitungan kerapatan cluster sebagai *variance within clusters* (V_w) dan *variance between clusters* (V_b) menggunakan persamaan 8.

$$V = \frac{V_w}{V_b} \quad (8)$$

D. LOGIKA FUZZY

Secara linguistik, *fuzzy* diartikan sebagai kabur atau. Nilai bisa sangat besar sementara juga salah. Dalam logika fuzzy, derajat memiliki rentang nilai dari 0 (nol) sampai 1 (satu). [12].

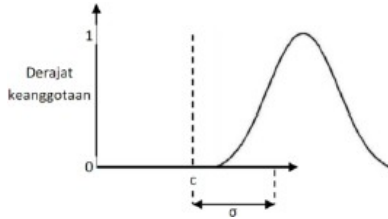
Logika fuzzy adalah penalaran dengan nilai antara benar dan salah. Akan tetapi, keberadaan dan kesalahan suatu ditentukan oleh besarnya keanggotaannya. Logika fuzzy memiliki derajat keanggotaan mulai dari 0 sampai

1) FUNGSI KEANGGOTAAN

ke dalam nilai penurunan (juga dikenal sebagai tingkat yang keanggotaan), yang memiliki rentang 0 hingga 1. Teknik fungsi adalah salah satu metode untuk menentukan nilai penarikan.

2) KURVA GAUSS

Fungsi keanggotaan Gaussian ditentukan dengan 2 parameter $\{c, \theta\}$ dengan mengikuti persamaan :



GAMBAR 1. Alur Penelitian.

Fungsi Keanggotaan:

$$G(x; k, \gamma) = e^{-k(\gamma-x)^2} \quad (9)$$

3) FUZZY INFERENCE SYSTEM (FIS) SUGENO

Metode fuzzy Sugeno adalah metode inferensi fuzzy untuk aturan yang didefinisikan sebagai *IF - THEN*, dimana output sistem (konsekuensi) adalah dalam bentuk konstanta atau persamaan linier daripada himpunan fuzzy. Takagi-Sugeno Kang memelopori pendekatan ini pada tahun 1985 [13]. Model Sugeno menggunakan fungsi fitur Singleton, yang memiliki derajat kehancuran satu untuk satu nilai crisp dan derajat kehancuran nol untuk nilai crisp lainnya. Formula Orde 0:

$$\begin{aligned} & \text{IF } (x_1 \text{ is } a_1)^\circ (x_2 \text{ is } A_2)^\circ \dots (x_n \text{ is } A_n) \\ & \text{THEN } z = k, \end{aligned} \quad (10)$$

Ai adalah himpunan fuzzy ke i sebagai antaseden (alasan). Sedangkan k adalah konstanta *assertive* sebagai konsekuen (kesimpulan), dan $^\circ$ adalah operator fuzzy (AND atau OR). Sementara rumus Orde 1 adalah:

$$\begin{aligned} & \text{IF } (x_1 \text{ is } a_1)^\circ (x_2 \text{ is } A_2)^\circ \dots (x_n \text{ is } A_n) \\ & \text{THEN } z = p_1 * x_1 + \dots + p_n * x_n + q, \end{aligned} \quad (11)$$

dimana Ai adalah himpunan fuzzy ke i sebagai antaseden, $^\circ$ adalah operator fuzzy (AND atau OR), q merupakan konstanta dalam konsekuen, dan pi adalah konstanta ke I.

4) EKSTRAKSI ATURAN FUZZY

Aturan Fuzzy kadang-kadang dapat diperoleh dari ahli manusia. Akuisisi pengetahuan, bagaimanapun juga adalah tugas yang rumit, dan beberapa bagian sistem tidak diketahui, sistem pakar belum tersedia [14]. Data mining dapat didefinisikan sebagai proses otomatis mencari ukuran pola-pola data yang besar. Data mining menggunakan proses pencarian melalui ukuran data yang besar menggunakan teknik clustering (K-means, Fuzzy K-means, Subtractive) [15] untuk memperoleh data yang relevan dan signifikan dalam pengenalan pola, dan logika fuzzy dari fuzzy inferensi sistem (Mamdani dan TSK) merupakan teknik berbasis untuk mengesktraksi pembuatan aturan-aturan (IF- THEN) [16].

Proses pembangkitan aturan fuzzy dilakukan dengan proses pencarian nilai output melalui langkah-langkah berikut [17]:

- Menentukan derajat keanggotaan setiap titik data i dalam setiap cluster k dengan menggunakan fungsi gauss berdasarkan persamaan 12.

$$\mu_{ki} = e^{-\sum_{j=1}^m \frac{(x_{ij}-C_{kj})^2}{2\sigma_j^2}} \quad (12)$$

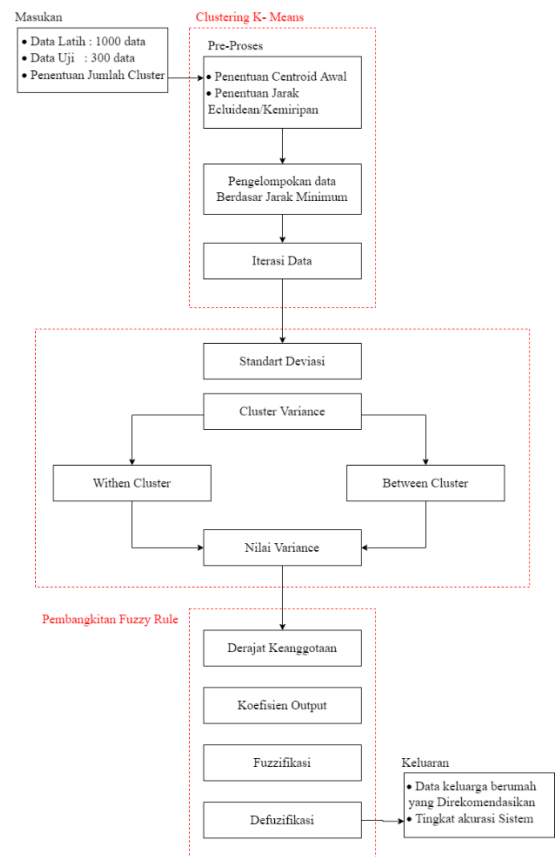
disusun menjadi satu vektor k:

$$k = [k_{11} \dots k_{1m} k_{10} k_{21} \dots k_{2m} k_{20} \dots k_{c1} \dots k_{cm} k_{10}]^T \quad (13)$$

Dari proses di atas terbentuklah koefisien output yang digunakan sebagai nilai output (konsekuensi) dari inferensi Sugeno.

III. DESAIN ARSITEKTURAL

Arsitektur sistem digunakan untuk menggambarkan kerja sistem yang digunakan dalam proses analisis dan implementasi, serta arsitektur sistem secara keseluruhan yang terlihat pada Gambar 2.



GAMBAR 2. Arsitektur Sistem

1) MASUKAN

Masukan atau input merupakan data-data yang diperlukan selama proses analisa sampai implementasi sehingga bisa menghasilkan output yang sesuai dengan harapan. User memasukkan data keluarga untuk mendapatkan rekomendasi penerima bantuan rumah tidak layak huni dengan variabel seperti umur, penghasilan, jumlah tanggungan, kondisi lantai, kondisi dinding, kondisi atap dan kondisi WC, kemudian sistem melakukan proses clustering untuk mengelompokkan

data keluarga, untuk menghasilkan cluster yang ideal dilakukan perhitungan varian, within cluster dan between cluster.

2) CLUSTERING K-MEANS

Sebelum melakukan proses clustering, penulis menentukan terlebih dahulu centroid awal yang menjadi dasar clustering. Untuk melakukan pemilihan centroid awal caranya dengan menekan tombol cari centroid pada form iterasi, lalu tentukan centroid awal secara random sesuai dengan jumlah cluster. Setelah melakukan pemilihan centroid awal, lalu lakukan perhitungan iterasi dengan menekan tombol "1" yang terdapat di group Hitung Iterasi.

3) ANALISA CLUSTER

Fitur analisis cluster digunakan untuk menghitung standar deviasi, varians, di dalam dan lintas cluster. Nilai varian cluster terkecil digunakan untuk menguji cluster ideal dengan menentukan nilai varian di dalam dan antar cluster.

4) PEMBANGKITAN FUZZY RULE

Fitur fuzzy digunakan untuk mengekstrak aturan fuzzy dengan menggunakan fuzzy inference system sugeno. Hasil dari fuzzy inference system sugeno akan menghasilkan output rekomendasi penerima bantuan keluarga berumah tidak layak huni. Nilai jumlah cluster merupakan jumlah cluster yang digunakan untuk pengembangan sistem inferensi fuzzy aturan Sugeno, aturan yang dihasilkan jumlahnya sama dengan banyaknya cluster. Pada makalah ini, penulis menggunakan 3 (tiga) cluster, oleh karena itu aturan fuzzy akan dibentuk dengan 3 cluster (tiga) Jumlah data pelatihan yang digunakan, nilai varians, adalah nilai varians terkecil dari beberapa percobaan yang dilakukan penulis untuk membentuk cluster optimal. Fungsi tombol proses digunakan untuk melakukan proses mencari derajat keanggotaan kemudian langkah selanjutnya melakukan defuzzifikasi.

5) KELUARAN

Output pada penelitian ini berupa rekomendasi penerima bantuan rumah tidak layak huni setelah dilakukan proses perhitungan dimana algoritma akan menghasilkan pengelompokkan data dengan kondisi yang optimum. Dalam kasus saran untuk penerima bantuan keluarga dengan perumahan tidak layak huni, K-means clustering digunakan untuk mengelompokkan data dan mengembangkan pedoman. Hasil dari pembangkitan aturan fuzzy digunakan untuk memproses inferensi fuzzy.

Sistem ini diharapkan memberikan sebuah hasil yang lebih relevan berupa rekomendasi penerima bantuan rumah tidak layak huni, sehingga bisa dibandingkan dengan perhitungan yang dilakukan dinas social dan tenaga kerja di kota Tarakan.

IV. UJI COBA

Dalam proses uji coba yang dilakukan menggunakan data sampel 1000 data dan menggunakan 300 data uji. Dalam proses pengujian dilakukan dalam 20 kali uji sebanyak 300 data setiap skenarionya.

A. JUMLAH DATA UNTUK REKOMENDASI PEMBERIAN BANTUAN

Lokasi penelitian ini dilakukan pada Dinas Sosial dan Tenaga Kerja (Dinsosnaker). Alamat Dinas Sosial dan Tenaga kerja Jl. Teuku Umar no.36 Kecamatan Tarakan Tengah Kota Tarakan, Telp (0551) 21329-34499 Fax. 34499.

Sedangkan kurun waktu penelitian dilakukan sejak September tahun 2014 hingga Januari 2015. Pengambilan data input yang diperlukan dalam penelitian ini, diambil dari Dinsosnaker dengan mengambil data sejak tahun 2012, 2013 dan 2014.

B. JUMLAH DATA UNTUK REKOMENDASI PEMBERIAN BANTUAN

Pada penelitian ini dibutuhkan suatu sampel data yang nantinya diolah menjadi sebuah informasi. Sampel data ini merupakan sebagian dari populasi yang menjadi objek pengamatan dari penelitian ini.

Sampel data diambil dari bagian Penyandang Masalah Kesejahteraan Sosial (PMKS) rumah tidak layak huni yang ada di kota Tarakan. Untuk besaran bantuan yang diberikan mulai dari 13,5 juta dan 27 juta sesuai dengan kondisi rumah yang akan diberikan bantuan. Sampel data latih yang digunakan dalam proses k-means clustering sebanyak 1000 data yang diambil dari 4 Kecamatan yang ada di kota Tarakan yaitu Kecamatan Tarakan Barat yang terdiri antara delapan kecamatan, Desa Karang Anyar, Karang Anyar Pantai, Karang Harapan, Karang Balik, dan Karang Rejo. Kecamatan Tarakan Tengah dibagi lagi menjadi lima kecamatan: Desa Kampung Satu, Pamusian, Selumit, Selumit Pantai, dan Sebengkok. Kabupaten Tarakan Timur dibagi lagi menjadi tujuh kecamatan: Lingkas Ujung, Mambirdan, Kampung Empat, Enam Desa, Pantai Amal, Mamdindingan Timur, dan Gunung Lingkas. Kecamatan Tarakan Utara terbagi menjadi tiga kecamatan yaitu Juata Kerikil, Juata Laut, dan Juata Permai. Untuk data uji data yang digunakan sebanyak 300 data yang diambil dari 3 Kecamatan yang ada di kota Tarakan yaitu terdiri dari delapan kelurahan, yaitu Desa Karang Anyar, Karang Anyar Pantai, Karang Harapan, Karang Balik, dan Karang Rejo, disusul Kecamatan Tarakan Tengah yang terdiri dari lima desa, yaitu Kampung Satu, Pamusian, Selumit, Selumit Pantai, dan Desa Sebengkok, dan terakhir Kecamatan Tarakan Timur yang terdiri dari tiga kecamatan yaitu Kelurahan Gunung Lingkas, Mamburungan Timur dan Pantai Amal.

C. JUMLAH DATA UNTUK REKOMENDASI PEMBERIAN BANTUAN

Pada penelitian ini, data training yang digunakan sebanyak 1000 data dari data keluarga yang mengajukan permohonan untuk mendapatkan bantuan rumah tidak layak huni dari tahun 2012 dan tahun 2013 Data ini yang nantinya diproses untuk menghasilkan sebuah cluster untuk membentuk aturan fuzzy secara otomatis dengan menggunakan fuzzy inference system sugeno. Data uji yang digunakan adalah data keluarga

berumah tidak layak huni sebanyak 300 data dari keluarga yang mengajukan permohonan untuk mendapatkan bantuan rumah tidak layak huni tahun 2014.

D. PERHITUNGAN AKURASI

Hasil yang diamati pada penelitian ini adalah seberapa tingkat akurasi metode K-means clustering dalam membangkitkan aturan fuzzy dalam pengujian data bantuan keluarga berumah tidak layak huni. Dalam penelitian ini akurasi dihitung dengan cara membagi jumlah data uji dukungan keluarga terhadap rumah tidak layak huni dengan jumlah data. Akurasi merujuk pada seberapa dekat suatu hasil pengukuran dengan nilai sebenarnya (true value atau nilai referensi). Besarnya presisi dapat dihitung dengan menggunakan persamaan 11.

$$Akurasi = \frac{\sum Data\ Uji\ Benar}{\sum Total\ Data\ Uji} * 100 \quad (14)$$

E. HASIL PENGUJIAN

Dalam proses pengujian dilakukan pengujian sebanyak 300 data uji. Sedangkan untuk proses pembangkitan aturan fuzzy digunakan sebanyak 1000 data latih yang diambil dari data keluarga berumah tidak layak huni. Dan dilakukan perhitungan nilai varian untuk setiap jumlah cluster yang diuji. Nilai akurasi dilakukan dengan mencari rata-rata akurasi global untuk seluruh nilai akurasi skenario pada masing-masing jumlah cluster uji. Dari hasil proses pelatihan nantinya dicatat hasil perhitungan varian dan dibandingkan dengan nilai akurasi global untuk seluruh skenario di masing-masing jumlah cluster apakah nilai varian terkecil dari jumlah cluster memiliki nilai akurasi tertinggi.

TABEL I
HASIL PERHITUNGAN AKURASI

NO	ID	VARIANCE	JUMLAH CLUSTER	AKURASI
1	10,59,647	0,00111	3	82,33 %
2	16,59,622	0,00110	3	81 %
3	10,59,907	0,00126	3	81 %
4	382,300,948	0,00126	3	80,33 %
5	289,301,774	0,00126	3	79,33%
6	620,347,149	0,00126	3	79,33%
7	326,633,207	0,00126	3	79,33%
8	695,980,243	0,00127	3	78,33%
9	275,985,802	0,00127	3	78,33%
10	272,755,946	0,00127	3	78,33%
11	331,852,540	0,00127	3	78,33%
12	331,121,540	0,00127	3	78,33%
13	284,45,295	0,00127	3	78,33%
14	272,755,622	0,00127	3	78,33%
15	480,254,340	0,00128	3	77%
16	357,149,704	0,00129	3	76,33%
17	289,301,744	0,00129	3	76,33%
18	543,814,540	0,00131	3	74,33%
19	829,824,589	0,00131	3	74,33%
20	949,364,524	0,00131	3	74,33 %

Tabel 1 menampilkan hasil perhitungan varian yang digunakan untuk menentukan cluster ideal dengan

menggunakan varian terkecil. Bisa dilihat untuk rata-rata akurasi global di setiap percobaan, menunjukkan bahwa pengujian akurasi paling maksimal dilakukan dengan menggunakan keluarga 10, keluarga 59 dan keluarga 647 dengan akurasi sitem 82 %.

V. KESIMPULAN

Dari penelitian berbasis K-Means clustering untuk pembangkitan aturan fuzzy dalam rangka rekomendasi pemberian bantuan rumah tidak layak huni yang telah dilakukan, dapat diambil kesimpulan:

1. K-Means clustering untuk pembangkitan aturan fuzzy merupakan metode yang tepat. Terbukti dari hasil perhitungan akurasi pengujian data uji setiap skenario dan akurasi pengujian data uji skenario global sama-sama menghasilkan akurasi minimal di atas 75%.
2. Dari hasil perhitungan nilai varian dan hasil perhitungan akurasi. Bisa diambil kesimpulan, proses penentuan jumlah cluster paling ideal masih belum optimal dengan menggunakan nilai varian terkecil. Dengan melihat hasil penelitian diperlukan adanya optimalisasi dalam menentukan cluster paling ideal.
3. Dalam prosedur ekstraksi aturan sugeno FIS, pendekatan pengelompokan K-Means dapat digunakan untuk membangun aturan fuzzy dengan penggunaan pusat cluster dan sigma.

Untuk pengembangan penelitian selanjutnya diperlukan adanya optimalisasi dalam menentukan cluster paling ideal. Selain itu, dapat pula dilakukan penyempurnaan proses pengelompokan yang bukan hanya pada satu periode saja, serta menambah kriteria-kriteria lain sehingga hasil yang diperoleh semakin sempurna.

DAFTAR PUSTAKA

- [1] B. Satria and L. Tambunan, "Sistem Pendukung Keputusan Penerima Bantuan Rumah Layak Huni Menggunakan FMADM dan SAW," *JOINTECS (Journal of Information Technology and Computer Science)*, vol. 5, no. 3, pp. 167–176, 2020.
- [2] "Wali Kota Tarakan Tinjau Program Bedah Rumah Agar Layak Huni - KANTOR BERITA KALIMANTAN." <https://kbbk.news/wali-kota-tarakan-tinjau-program-bedah-rumah-agar-layak-huni/> (accessed Nov. 17, 2022).
- [3] K. Khotimah, "Analisis Program Bantuan Rumah Layak Huni Terhadap Pengentasan Kemiskinan di Kecamatan Mesuji Dalam Perspektif Ekonomi Islam," UIN Raden Intan Lampung, 2018.
- [4] H. Tumanggor, M. Haloho, P. Ramadhani, and S. D. Nasution, "Penerapan Metode VIKOR Dalam Penentuan Penerima Dana Bantuan Rumah Tidak Layak Huni," *JURIKOM (Jurnal Riset Komputer)*, vol. 5, no. 1, pp. 71–78, 2018.
- [5] D. Kusumawardani, "Sistem Pendukung Keputusan Penerima Bantuan Rumah Layak Huni Dengan Menggunakan Metode Weighted Product (WP)," *FASILKOM UDINUS*, 2014.
- [6] H. Nalattissifa and Y. Ramdhani, "Sistem Penunjang Keputusan Menggunakan Metode Topsis Untuk Menentukan Kelayakan Bantuan Rumah Tidak Layak Huni (RTLH)," *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 19, no. 2, pp. 246–256, 2020.
- [7] Efraim. Turban, J. E. Aronson, and T.-P. Liang, "Decision support systems and intelligent systems," p. 936, 2005.
- [8] "Straubhaar, J. and LaRose, R. (2006) Communications Media in the Information Society. Wadsworth Publishing Company,

- Belmont, CA. - References - Scientific Research Publishing.”
<https://www.scirp.org/%28S%28351jmbntvnsjt1aadkozje%29%29/reference/referencespapers.aspx?referenceid=2515740>
(accessed Nov. 18, 2022).
- [9] K. A. A. Nazeer and M. P. Sebastian, “Improving the Accuracy and Efficiency of the k-means Clustering Algorithm,” 2009.
- [10] R. Arapoglou, K. Kolomvatsos, and S. Hadjiefthymiades, “Buyer agent decision process based on automatic fuzzy rules generation methods,” *International Conference on Fuzzy Systems*, pp. 1–8, 2010.
- [11] A. Bakri and M. F. M. Adini, “PENGELOMPOKAN DATA KAJI CUACA MENGGUNAKAN K-MEANS BAGI PERAMALAN TABURAN HUJAN.”.
- [12] K. Sri, “Aplikasi logika fuzzy untuk pendukung keputusan / Sri Kusumadewi, Hari Purnomo,” 2010.
- [13] J. Moreno, O. Castillo, J. Castro, L. Martinez, and P. Melin, “Data Mining for extraction of fuzzy IF-THEN rules using Mamdani and Takagi-Sugeno-Kang FIS,” *Engineering Letters*, vol. 15, 2007.
- [14] P. C. Chang and C.-H. Liu, “A TSK type fuzzy rule based system for stock price prediction,” *Expert Syst. Appl.*, vol. 34, pp. 135–144, 2008.
- [15] A. PRIYONO, M. Ridwan, A. ALIAS, R. Rahmat, A. Hassan, and M. Mohd Ali, “Generation of Fuzzy Rules with Subtractive Clustering,” *Jurnal Teknologi*, vol. 43, p. 143, 2005, doi: 10.11113/jt.v43.782.
- [16] E. Turban, E. McLean, and J. Wetherbe, “Information Technology for Management : Making Connections for Strategic Advantage / E. Turban, E. McLean, J. Wetherbe.,” 2001.
- [17] L. T. 1 Kóczy, “Fuzzy Rule Base Model Identification Techniques,” 2005.