

Contents list available at www.jurnal.unimed.ac.id

CESS
(Journal of Computing Engineering, System and Science)

journal homepage: <https://jurnal.unimed.ac.id/2012/index.php/cess>



**Analisis Data Penyandang Masalah Kesejahteraan Sosial Menggunakan
Algoritma Fuzzy C-Means**

**Data Analysis of People with Social Welfare Problems using Fuzzy C-Means
Algorithm**

Triswanto^{1*}, Rudi Kurniawan², Yudhistira Arie Wijaya³

^{1,2,3}STMIK IKMI Cirebon

Jl. Perjuangan No.10B, Karyamulya, Kesambi, Kota Cirebon, Jawa Barat

email: ¹trisw123@gmail.com, ²rudi226@gmail.com, ³yudhistira010471@gmail.com

ABSTRAK

Kesejahteraan sosial menjadi perhatian utama dalam berbagai negara, dan identifikasi dan pemahaman yang lebih baik tentang masalah ini menjadi penting untuk merancang kebijakan dan program yang lebih efektif. Masalah kesejahteraan sosial seperti kemiskinan, disabilitas, dan ketimpangan sosial menjadi isu yang mendapat perhatian luas. Tujuan penelitian ini adalah untuk menganalisis data penyandang masalah kesejahteraan sosial dengan mengoptimalkan nilai K berdasarkan *Davies Bouldin Index*. Penelitian ini menggunakan metode Knowledge Discovery in Database dengan algoritma *Fuzzy C-Means* untuk mengelompokkan data penyandang masalah kesejahteraan sosial dengan lebih baik. Algoritma ini digunakan karena pengelompokan datanya berdasarkan dengan derajat keanggotaan sehingga pusat cluster yang dihasilkan dalam mencapai fungsi sasaran mencari solusi terbaik. Dalam penelitian ini, data yang digunakan adalah data penyandang masalah kesejahteraan sosial di Kabupaten Karawang. Hasil dari penelitian ini menunjukkan bahwa metode *Fuzzy C-Means Clustering* dapat memberikan pengelompokan data penyandang masalah kesejahteraan sosial yang lebih akurat dan representatif. Dalam penelitian ini dapat ditemukan nilai *Davies Bouldin Index* yaitu 0,564 dan nilai K optimal yaitu 4. Dengan menggunakan metode ini, kelompok-kelompok yang memiliki karakteristik dan kebutuhan yang serupa dapat diidentifikasi dengan lebih baik.

Kata Kunci: *Fuzzy C-Means, Masalah Kesejahteraan Sosial, Clustering, Davies Bouldin Index, Nilai K Optimal*

ABSTRACT

Social welfare is a major concern in many countries, and better identification and understanding of these issues is important for designing more effective policies and programs.

*Penulis Korespondensi:
email: trisw123@gmail.com

Social welfare problems such as poverty, disability, and social inequality are issues that have received widespread attention. The purpose of this research is to analyze data on people with social welfare problems by optimizing the K value based on the *Davies Bouldin Index*. This research uses the Knowledge Discovery in Database method with the *Fuzzy C-Means* algorithm to better categorize data on people with social welfare problems. This algorithm is used because the data grouping is based on the degree of membership so that the resulting cluster center achieves the target function of finding the best solution. In this study, the data used is data on people with social welfare problems in Karawang Regency. The results and discussion of this research show that the *Fuzzy C-Means Clustering* method can provide a more accurate and representative grouping of data on people with social welfare problems. In this study, it can be found that the *Davies Bouldin Index* value is 0.564 and the optimal K value is 4. By using this method, groups that have similar characteristics and needs can be better identified.

Keywords: *Fuzzy C-Means, Social welfare, Clustering, Davies Bouldin Index, Optimal K Value*

1. PENDAHULUAN

Penyandang Masalah Kesejahteraan Sosial (PMKS) adalah seseorang, keluarga, bahkan sekelompok masyarakat yang tidak dapat melaksanakan kegiatan sosialnya yang disebabkan oleh suatu hambatan, kesulitan, atau juga gangguan [1]. Kesejahteraan sosial menjadi fokus utama yang terkait dengan berbagai tantangan di wilayah Indonesia, termasuk di Kabupaten Karawang. Permasalahan ini muncul akibat perubahan gaya hidup masyarakat dalam era globalisasi. Proses globalisasi meningkatkan tuntutan hidup, intensifikasi persaingan dalam kehidupan sehari-hari, serta ketidaksesuaian dan pembatasan yang diterapkan oleh masyarakat [2]. Masalah kesejahteraan sosial tidak mampu sepenuhnya diselesaikan, hal tersebut dapat dikurangi dan dihadapi secara serius selama evolusi kehidupan. Informasi mengenai situasi yang melibatkan bantuan sosial menjadi penting, dan analisis diperlukan untuk mengidentifikasi lingkup masalah yang ada saat ini serta menyediakan layanan sosial [3]. Untuk mengelompokkan data penyandang masalah kesejahteraan sosial di Kabupaten Karawang khususnya di tahun 2020 dan 2021 dapat dilakukan analisis dengan menggunakan salah satu metode *Fuzzy C-Means Clustering* yang dapat membagi data menjadi beberapa cluster [4]. Algoritma dalam metode *Fuzzy C-Means* memiliki tingkat ketelitian cukup tinggi terhadap ukuran suatu objek dibandingkan algoritma *K-Means* [5]. Sehingga algoritma dalam *Fuzzy C-Means* dapat membantu mengklasifikasi data penyandang masalah kesejahteraan sosial ini [6]. Berdasarkan informasi yang tersedia mengenai PMKS di Kabupaten Karawang, setiap kecamatan memiliki jenis PMKS yang berbeda.

Alasan mengapa penelitian ini menggunakan metode *Fuzzy C-Means* karena didasarkan pada jurnal penelitian yang ditulis oleh Ghina Nabila Saputro Putri, Dwi Ispriyanti, Tatik Widiharih dengan judul "Implementasi Algoritma *Fuzzy C-Means* Dan *Fuzzy Possibilistics C-Means* Untuk Klasterisasi Data Tweets Pada Akun Twitter Tokopedia", dimana pada penelitian ini menggunakan metode *Fuzzy C-Means* (FCM) dan *Fuzzy Possibilistics C-Means* (FPCM) untuk mengelompokkan data tweets dari akun Twitter @tokopedia. Hasilnya menunjukkan bahwa FCM lebih baik daripada FPCM dalam mengelompokkan data tweets. Dari hasil *clustering*, empat jenis konten utama diidentifikasi. Cluster 1, yang berisi tweet tentang promo barang otomotif, alat olahraga, dan merchandise artis ternama, memiliki rata-rata like dan

retweet tertinggi. Sementara itu, cluster 2, yang berisi tweet mengenai penawaran tiket pesawat, alat elektronik, dan gadget, memiliki rata-rata like dan retweet terendah [4].

Tujuan utama dari penelitian ini adalah untuk meningkatkan pemahaman dan penerapan metode *clustering Fuzzy C-Means* dalam konteks kesejahteraan sosial. Tujuan ini mencakup pengidentifikasian nilai K terhadap *Davies Bouldin Index*, yang selanjutnya dapat digunakan untuk mengkategorikan penyandang masalah kesejahteraan sosial secara lebih efisien [7]. Selain itu, penelitian ini bertujuan untuk mengetahui cluster dari data penyandang masalah kesejahteraan sosial. Pendekatan statistik dan matematika digunakan untuk mengaplikasikan metode *clustering Fuzzy C-Means* [8]. Langkah ini melibatkan pemilihan jumlah cluster yang optimal dan perhitungan jarak antara titik data [9]. Selanjutnya, pendekatan visualisasi data juga diterapkan untuk mempresentasikan hasil *clustering* dengan jelas. Penelitian ini memanfaatkan perangkat lunak analisis data yaitu Rapidminer versi 10.1 [10].

Penelitian ini mengimplementasikan metode *clustering Fuzzy C-Means* untuk pengelompokan data penyandang masalah kesejahteraan sosial di Kabupaten Karawang. Hasil penelitian ini memberikan pandangan yang lebih jelas dan terperinci tentang beragam jenis penyandang masalah kesejahteraan sosial [11][12]. Dengan kata lain, hasil penelitian ini dapat membantu merancang program intervensi yang lebih efektif dan berkelanjutan. Metode *Fuzzy C-Means Clustering* yang digunakan dalam penelitian ini menunjukkan pengelompokan data penyandang masalah kesejahteraan sosial yang lebih akurat dan representatif. Berbeda dengan penelitian sebelumnya yang menggunakan metode *Fuzzy C-Means Clustering* dengan berdasarkan nilai *silhouette* dan *Matriks Partisi* [1][13], penelitian ini menggunakan metode *Fuzzy C-Means Clustering* berdasarkan nilai DBI. Dengan metode ini, kelompok-kelompok yang memiliki karakteristik dan kebutuhan yang serupa dapat diidentifikasi dengan lebih baik. Hasilnya memberikan gambaran dan identifikasi kelompok yang dibagi menjadi beberapa klaster berdasarkan hasil dari analisis data dengan mencari nilai K optimal berdasarkan DBI [14].

2. DASAR/TINJAUAN TEORI

2.1. Penelitian Terdahulu

Sumber acuan penelitian ini adalah studi literatur yang dilakukan terhadap jurnal-jurnal yang memiliki tema yang sama dengan penelitian yang akan dilakukan. Dalam hal ini, peneliti menggunakan studi literatur sebagai sumber acuan untuk penelitian ini. Penelitian sebelumnya pernah dilakukan oleh Nafa Nur Adifia Nanda, Nurissaidah Ulinnuha, Hani Khaulasari, dengan judul "Pengelompokan Kecamatan di Wilayah Kabupaten Bojonegoro Berdasarkan Jenis Penyandang Masalah Kesejahteraan Sosial Menggunakan *Fuzzy C-Means Clustering*", Dalam penelitian ini, *Fuzzy C-Means (FCM) Clustering* digunakan untuk mengelompokkan kecamatan di Kabupaten Bojonegoro berdasarkan karakteristik Penyandang Masalah Kesejahteraan Sosial (PMKS). Hasil dari penggunaan FCM dalam penelitian ini adalah pengelompokkan kecamatan berdasarkan jenis PMKS menjadi 2 cluster (tinggi dan rendah) dengan nilai *silhouette* 0,7275 [1].

Penelitian lain yang dilakukan oleh Rahmawati dkk, dengan judul "Pengelompokkan Data Penerima Bantuan Program Keluarga Harapan (PKH) di Kecamatan Tuah Madani Pekanbaru Menggunakan *Fuzzy C-Means*", Dalam penelitian ini, data PKH Kecamatan Tuah Madani dipecah menjadi dua cluster, yaitu tingkat tinggi dan rendah, berdasarkan jumlah

penerima bantuan PKH dan jumlah penduduk di setiap kelurahan. Hasilnya menunjukkan bahwa Kelurahan Sialang Munggu, Sidomulyo Barat, dan Tuah Karya termasuk dalam cluster tinggi penerima PKH, sementara Kelurahan Air Putih dan Tuah Madani termasuk dalam cluster rendah penerima PKH [13].

Berdasarkan studi literatur di atas, penelitian ini berfokus dalam penggunaan algoritma *Fuzzy C-Means* dalam pengelompokan data penyandang masalah kesejahteraan sosial (PMKS). Dimana data PMKS diperoleh dari web resmi Open Data Jabar. *Fuzzy C-Means* ialah teknik pengelompokan data yang keberadaan setiap data dalam cluster dipastikan oleh nilai keanggotaan.

2.2. Dasar Teori

2.2.1. Data Mining

Teknik ini bukanlah hal yang baru dalam mengolah data menjadi pengetahuan. Namun, salah satu kesulitan dalam menggambarkan *Data mining* adalah bahwa teknik ini menurun banyak aspek dari bidang-bidang ilmu yang sudah ada terlebih dahulu. *Data mining* berasal dari beberapa disiplin ilmu dan bertujuan untuk mengubah pandangan data tradisional sehingga bisa menangani jumlah data yang sangat besar, dimensi data yang tinggi, dan data yang heterogen dan berbeda sifat. *Data mining* memiliki banyak manfaat dalam mengelola data mentah menjadi pengetahuan atau kumpulan informasi yang dapat digunakan sebagai penunjang keputusan secara efektif [15].

2.2.2. Fuzzy C-Means Clustering

Metode *Fuzzy C-Means* ditemukan pertama kali oleh Dunn pada tahun 1973, kemudian dikembangkan lagi oleh Bezdek pada tahun 1981. Ide dasar dari metode ini mirip dengan metode *K-Means*. *Fuzzy C-Means* didasarkan pada logika fuzzy. Setiap titik data dimasukkan ke suatu kelompok berdasarkan nilai keanggotaannya pada kelompok tersebut. Metode FCM merupakan suatu teknik pengelompokan data yang keberadaan tiap-tiap data dalam suatu *cluster* ditentukan oleh nilai keanggotaannya dan nilai derajat keanggotaan yang jangkauan nilainya 0 sampai 1 [16]. Algoritma FCM sebagai berikut:

1. Menentukan banyak kelompok (c), fuzzifier (m), maksimum iterasi (*MaxIter*) perubahan nilai fungsi objektif terkecil yang diharapkan (c), fungsi objektif awal ($P_0=0$), dan iterasi awal ($t = 1$).
2. Membangkitkan bilangan random U_{ik} dengan i merupakan banyak data dan k merupakan banyak kelompok sebagai elemen-elemen awal matriks keanggotaan awal U .
3. Menghitung pusat kelompok ke- i dengan persamaan:

$$P_i = \frac{\sum_k^n = 1(U_{ik})^m X_k}{\sum_k^n = 1(U_{ik})^m}$$

dengan U_{ik} adalah nilai keanggotaan objek ke- k dengan pusat kelompok ke- i , X_k adalah objek data ke- k , N adalah banyaknya objek penelitian, dan m adalah fuzzifier.

4. Menghitung fungsi objektif pada iterasi ke- t dengan persamaan:

$$J(P, U, X, c, m) = \sum_{i=1}^c \sum_{k=1}^N (U_{ik})^m D_{ik}^2 (X_k, P_i)$$

dengan c adalah banyak kelompok yang diinginkan, N adalah banyak objek penelitian, U_{ik} adalah nilai keanggotaan objek ke- k pada kelompok ke- i yang merupakan bagian dari matriks U , m adalah fuzzifier, dan $d_{ik}^2(X_k, P_i)$ adalah jarak antara vektor pengamatan ke- k dengan pusat kelompok ke- i .

5. Menghitung perubahan matriks keanggotaan dengan persamaan:

$$U_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{d_{ik}^2}{d_{jk}^2}\right)^{\frac{1}{m-1}}}$$

dengan U_{ik} adalah nilai keanggotaan objek ke- k dengan pusat kelompok ke- i , d_{ik}^2 adalah jarak antara objek ke- k dengan pusat kelompok ke- i , dan d_{jk}^2 adalah jarak antara objek ke- k dengan pusat kelompok ke- j , dan m adalah fuzzifier.

6. Cek kondisi berhenti:

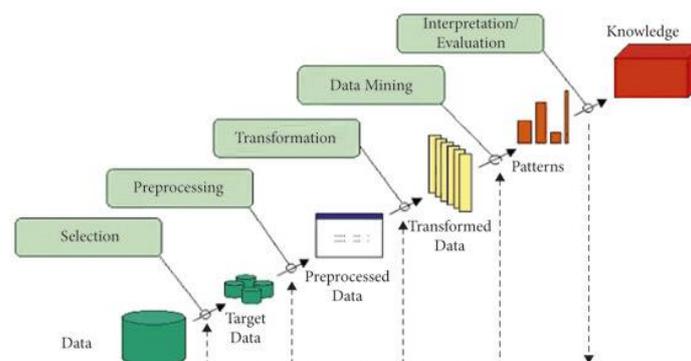
- Jika $|J_t - J_{t-1}| < \varepsilon$ atau $t > MaxIter$ maka berhenti;
- Jika tidak: $t = t + 1$, ulangi langkah ke-3.

2.2.3. RapidMiner

RapidMiner adalah sebuah perangkat lunak open-source yang digunakan untuk menganalisis data processing. Pada RapidMiner, digunakan berbagai teknik deskriptif dan prediksi. RapidMiner ini menggunakan bahasa Java untuk pengoperasiannya. Dengan adanya kecanggihan teknologi algoritma komputasi dan analisis data berbasis komputer, datamining dapat diolah menggunakan software Rapidminer [10]. Algoritma *Fuzzy C-Means* dapat dihitung dengan menggunakan RapidMiner. Penelitian ini diharapkan dapat membantu pemerintah dalam mengenal tingkat penyandang masalah kesejahteraan sosial di wilayah Kecamatan Kabupaten Karawang pada Provinsi Jawa Barat.

3. METODE

Metode yang digunakan dalam penelitian ini adalah Knowledge Discovery in Database (KDD). KDD adalah proses yang digunakan untuk mengidentifikasi validitas data, potensi, guna, dan pada akhirnya menghasilkan pola data yang dapat dimengerti. KDD berhubungan dengan teknik integrasi dan penemuan ilmiah, interpretasi, dan visualisasi dari pola-pola sejumlah kumpulan data. *Data mining* sendiri adalah bagian dari tahapan proses KDD. Tahapan proses KDD meliputi seleksi data, pra proses data, transformasi data, *data mining*, dan yang terakhir interpretasi dan evaluasi. Berikut adalah ilustrasi mengenai proses KDD yang dapat dilihat pada Gambar 1 berikut ini.



Gambar 1. Tahapan Knowledge Discovery in Database

Berdasarkan gambar di atas dapat dijelaskan bahwa tahapan KDD adalah:

1. Data Selection

Data selection (pemilihan data) merupakan proses penyiapan data dengan memilih sekumpulan data yang menjadi target/atribut/indikator dari dalam data yang besar.

2. Preprocessing/Cleaning

Proses cleaning (pembersihan data) merupakan proses memperbaiki kesalahan (data error), membuang data ganda, dan memeriksa data yang inkonsisten seperti kesalahan cetak (tipografi).

3. Data Transformation

Proses Data Transformation (transformasi data) merupakan proses perubahan data ataupun pengkodean data. Data transformasi dilakukan untuk mempermudah pengolahan data dan merupakan data yang telah siap untuk dilakukan proses *data mining*.

4. *Data mining*

Data mining merupakan proses pencarian pola atau informasi yang menarik/bermanfaat dalam data yang besar dengan menggunakan teknik atau metode tertentu.

5. Evaluation

Evaluation merupakan proses menampilkan pola/informasi/pengetahuan dari hasil proses *data mining* ke dalam bentuk yang mudah dimengerti.

4. HASIL DAN PEMBAHASAN

Implementasi algoritma *Fuzzy C-Means Clustering* ini dirancang dengan menggunakan tools RapidMiner versi 10.1. Tools ini dipilih karena memiliki fitur yang lengkap dan mudah digunakan untuk mengimplementasikan algoritma *clustering*. Pada penelitian ini, digunakan tahapan Knowledge Discovery in Database.

4.1. Hasil

4.1.1. Data

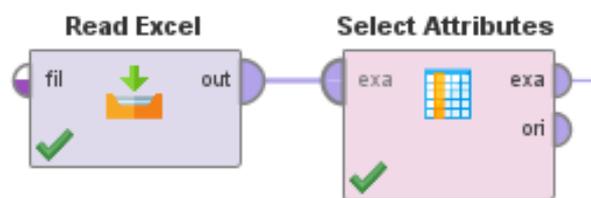
Penelitian ini menggunakan data sekunder, yaitu data yang diperoleh dari instansi pemerintah atau swasta. Data sekunder yang digunakan dalam penelitian ini berasal dari sumber resmi pemerintah, yaitu situs web <https://opendata.jabarprov.go.id>. Dataset tersebut dapat diakses dan diunduh pada halaman yang berjudul "Dataset penyandang masalah kesejahteraan sosial per kecamatan di Kabupaten Karawang". Data ini dikumpulkan dari 30 sampel yang mewakili 30 kecamatan di Kabupaten Karawang tahun 2020. Hasil pengumpulan data disimpan dalam format Excel dan kemudian disimpan ke dalam aplikasi RapidMiner. Dataset berikut hanya disajikan sebanyak 10 record, dari data 5 data teratas dan 5 data terbawah. Seperti terlampir pada Gambar 2 berikut ini.

KODE_KABU PATEN	NAMA_KAB UPATEN	KODE_KECA MATAN	NAMA_KECA MATAN	ANAK BALITA TERLANTAR	ANAK TERLANTAR	WANITA RAWAN SOSIAL EKONOMI	LANSIA TERLANTAR	PENYANDANG DISABILITAS	KELUARGA FAKIR MISKIN	KORBAN BENCANA ALAM	PEKERJA MIGRAN	JUML AH	SATU AN	TAH UN
3215	KARAWANG	3215010	PANGKALAN	410	3643	390	2946	174	5585	548	1	1369	7	ORA NG 0
3215	KARAWANG	3215011	TEGALWARU	239	2986	398	2147	113	4266	29	1	1017	9	ORA NG 0
3215	KARAWANG	3215012	CIAMPEL	399	3677	176	2527	181	5225	143	0	1232	8	ORA NG 0
3215	KARAWANG	3215013	TELUK JAMBE TIMUR	273	4155	274	3631	203	6872	10	0	1541	8	ORA NG 0
3215	KARAWANG	3215014	TELUK JAMBE BARAT	284	3678	435	3612	186	6397	2939	0	1753	1	ORA NG 0
3215	KARAWANG	3215152	PEDES	863	7924	822	5423	354	12457	7114	2	3495	9	ORA NG 0
3215	KARAWANG	3215161	CILEBAR	289	4911	354	3728	195	7668	1604	2	1875	1	ORA NG 0
3215	KARAWANG	3215170	CIBUAYA	595	6264	128	4334	303	9711	4253	2	2559	0	ORA NG 0
3215	KARAWANG	3215180	TIRTAJAYA	452	6660	775	4708	291	10405	16	6	2331	3	ORA NG 0
3215	KARAWANG	3215190	BATUJAJA	622	9013	297	5118	296	13213	3	1	2856	3	ORA NG 0
3215	KARAWANG	3215200	PAKISIJAYA	387	5029	48	2390	136	6698	463	0	1515	1	ORA NG 0

Gambar 2. Dataset Penyandang Masalah Kesejahteraan Sosial

4.1.2. Data Selection

Untuk dapat menggunakan dataset tersebut di RapidMiner versi 10.1, data tersebut perlu diimpor menggunakan operator *Read Excel* yang tersedia di RapidMiner. Operator ini berfungsi untuk membaca file Excel dan penting untuk mengimpor data dari file Excel yang tersimpan di komputer pengguna ke dalam RapidMiner. Operator *Select Attributes* pada aplikasi RapidMiner digunakan untuk memfilter atribut data yang akan digunakan dalam pemrosesan dan analisis data selanjutnya. Dataset penyandang masalah kesejahteraan sosial memiliki 16 atribut. Melalui operator *Select Attributes*, data difilter menjadi hanya 11 atribut. Tampilan operator *Read Excel* dan *Select Attributes* dapat dilihat pada Gambar 3 berikut ini.



Gambar 3. Tahap Data Selection

Dari hasil pembacaan operator *Select Attributes* yang telah dilakukan didapat informasi pada Tabel 1 sebagai berikut:

Tabel 1. Statistik Dataset

Uraian	Keterangan
<i>Record</i>	30
<i>Special Attributes</i>	0
<i>Regular Attributes</i>	11
<i>Attributes:</i>	
No	Numerical, missing 0
Nama Kabupaten	Nominal, missing 0
Nama Kecamatan	Nominal, missing 0
Anak Balita Terlantar	Numerical, missing 0
Anak Terlantar	Numerical, missing 0
Wanita Rawan Sosial Ekonomi	Numerical, missing 0
Lansia Terlantar	Numerical, missing 0
Penyandang Disabilitas	Numerical, missing 0

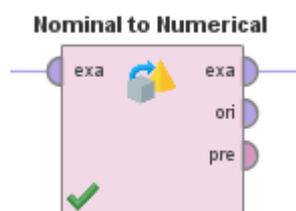
Keluarga Fakir Miskin	Numerical, missing 0
Korban Bencana Alam	Numerical, missing 0
Pekerja Migran	Numerical, missing 0

4.1.3. Preprocessing

Tahap *preprocessing* dalam *Knowledge Discovery in Database* (KDD) bertujuan untuk menghilangkan *Missing Value* pada data yang akan digunakan agar proses selanjutnya tidak mengalami kesalahan. *Missing Value* perlu ditangani karena dapat mengganggu proses *modelling data mining* selanjutnya. Pada dataset Penyandang Masalah Kesejahteraan Sosial, tidak ditemukan *Missing Value* atau data yang tidak memiliki nilai.

4.1.4. Transformation

Proses transformasi data diperlukan dalam *data mining* untuk mengubah tipe data atribut agar sesuai dengan analisis yang dibutuhkan. Dalam penelitian ini, transformasi data dilakukan karena terdapat beberapa atribut dengan tipe data nonnumerik yang perlu diubah menjadi bentuk numerik agar dapat diproses lebih lanjut. Atribut nonnumerik yang perlu diubah adalah Nama_Kecamatan dan Nama_Kabupaten. Kedua atribut ini kemudian ditransformasikan ke dalam bentuk numerik menggunakan operator *Nominal to Numerical*. Tampilan operator *Nominal to Numerical* dapat dilihat pada Gambar 4 berikut ini.



Gambar 4. Operator *Nominal to Numerical*

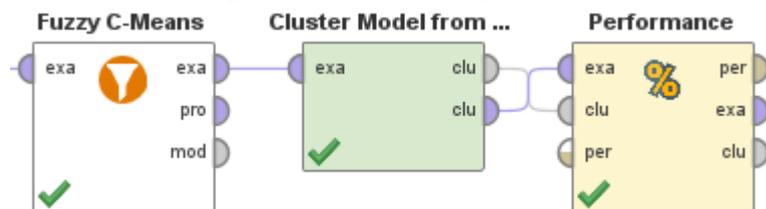
Dari hasil pembacaan operator *Nominal to Numerical* yang telah dilakukan didapat informasi pada Tabel 4.2 sebagai berikut:

Tabel 2. Statistik Dataset

Uraian	Keterangan
<i>Record</i>	30
<i>Special Attributes</i>	0
<i>Regular Attributes</i>	11
<i>Attributes:</i>	
No	Numerical, missing 0
Nama Kabupaten	Numerical, missing 0
Nama Kecamatan	Numerical, missing 0
Anak Balita Terlantar	Numerical, missing 0
Anak Terlantar	Numerical, missing 0
Wanita Rawan Sosial Ekonomi	Numerical, missing 0
Lansia Terlantar	Numerical, missing 0
Penyandang Disabilitas	Numerical, missing 0
Keluarga Fakir Miskin	Numerical, missing 0
Korban Bencana Alam	Numerical, missing 0
Pekerja Migran	Numerical, missing 0

4.1.5. Data mining

Pada tahap ini, teknik yang digunakan adalah pengklasteran dengan mengimplementasikan algoritma *Fuzzy C-Means*. Operator ini merupakan operator utama pemodelan ini agar dapat diketahui hasil pengklasteran dataset ini. Selanjutnya ditambahkan operator *cluster model from data* untuk menghubungkan operator *Fuzzy C-Means* dengan Operator *Cluster Distance Performance* dengan metode *Davies Bouldin Index (DBI)*. pemodelan pada tahap *Data mining* di RapidMiner, dapat dilihat pada Gambar 5 berikut ini.



Gambar 5. Pemodelan *Data mining*

4.1.6. Evaluation

Pada tahap evaluasi, dilakukan evaluasi *cluster* untuk menentukan jumlah *cluster* yang menghasilkan nilai *Davies Bouldin Index (DBI)* terkecil. Nilai DBI terkecil diperoleh dari perbandingan DBI dengan algoritma *Fuzzy C-Means*, dengan percobaan pengujian dari K=2 hingga K=10 dengan iterasi sebanyak 10. Hasilnya, *cluster* dengan nilai DBI terkecil adalah K=4 sebesar 0.564. Informasi hasil DBI dapat dilihat pada Tabel 3 berikut ini.

Tabel 3. Evaluasi Tabel nilai K Optimal

<i>Clusters</i>	<i>Iteration</i>	<i>Davies Bouldin Index</i>
2	10	1,215
3	10	0,890
4	10	0,564
5	10	0,883
6	10	1,317
7	10	0,607
8	10	1,039
9	10	0,993
10	10	0,649

4.2. Pembahasan

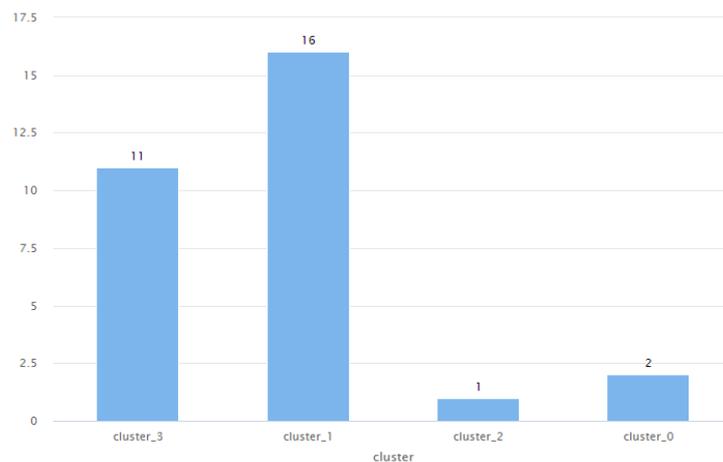
Berdasarkan penelitian sebelumnya yang dilakukan oleh Nafa Nur Adifia Nanda, Nurissaidah Ulinuha, dan Hani Khaulasari dengan judul “Pengelompokan Kecamatan di Wilayah Kabupaten Bojonegoro Berdasarkan Jenis Penyandang Masalah Kesejahteraan Sosial Menggunakan Fuzzy C-Means Clustering”, ditemukan bahwa hasil pengelompokan kabupaten di Kabupaten Bojonegoro berdasarkan jenis penyandang masalah kesejahteraan sosial dibagi menjadi 2 cluster (tinggi dan rendah) dengan nilai silhouette sebesar 0,7275. Nilai tersebut menunjukkan kekuatan cluster mempunyai struktur kuat. Terdapat 7 kecamatan dengan penyandang masalah kesejahteraan sosial yang tinggi dan tergolong dalam cluster 1. Di cluster 2 terdapat 21 kecamatan dengan penyandang masalah kesejahteraan sosial yang rendah dibandingkan dengan cluster 1 [1]. Selain itu, hasil yang dilakukan oleh peneliti tentang Analisis Data Penyandang Masalah Kesejahteraan Sosial di kabupaten karawang menunjukkan

bahwa nilai DBI yang paling rendah menunjukkan cluster yang lebih baik dibandingkan dengan cluster dengan nilai DBI yang lebih tinggi. Berdasarkan hasil pengujian pengelompokan data dengan jumlah cluster dari K=2 hingga K=20, masing-masing dengan nilai iterasi dari 1 hingga 10, diperoleh nilai K optimal pada K=5 dengan nilai iterasi 10 dan nilai Davies Bouldin Index sebesar 0,564. Hasil pengujian dapat dilihat pada Tabel 4 sebagai berikut.

Tabel 4. Nilai K Optimal

<i>Clusters</i>	<i>Iteration</i>	<i>Davies Bouldin Index</i>
4	10	0,564

Berdasarkan Hasil *clustering* dengan menggunakan algoritma *Fuzzy C-Means* menghasilkan 4 kelompok terbaik, dengan nilai DBI 0.564. Jumlah anggota *cluster* 0 adalah 2 item, *cluster* 1 adalah 16 item, *cluster* 2 adalah 1 item dan *cluster* 3 adalah 11 item. Grafik bar *cluster* terlampir pada Gambar 6 berikut ini.



Gambar 6. Jumlah Cluster

Anggota masing-masing *cluster* berdasarkan nama kecamatan di Kabupaten Karawang dapat dilihat pada tabel 5 berikut ini.

Tabel 5. Data Cluster per Kecamatan

<i>Cluster 0</i>	<i>Cluster 1</i>	<i>Cluster 2</i>	<i>Cluster 3</i>
KARAWANG BARAT RENGASDENGKLOK	KLARI TIRTAMULYA JATISARI BANYUSARI KOTABARU CILAMAYA KULON LEMAHABANG TELAGASARI KARAWANG TIMUR	CIKAMAYA WETAN	PANGKALAN TEGALWARU CIAMPEL TELUK JAMBE TIMUR TELUK JAMBE BARAT CIKAMPEK PURWASARI MAJALAYA RAWAMERTA

Cluster 0	Cluster 1	Cluster 2	Cluster 3
	TEMPURAN		CILEBAR
	KUTAWALUYA		PAKISJAYA
	JAYAKERTA		
	PEDES		
	CIBUAYA		
	TIRTAJAYA		
	BATUJAYA		

5. KESIMPULAN

Setelah dilakukan penelitian *data mining* dengan metode *Fuzzy C-Means Clustering* terkait data penyandang masalah kesejahteraan sosial (PMKS) di Kabupaten Karawang, maka dapat diambil kesimpulan bahwa nilai K optimal untuk pengelompokan data PMKS adalah 4 dengan nilai *Davies Bouldin Index* 0,564. Dengan *cluster* 0 berjumlah 2 di Kecamatan Karawang Barat dan Rengasdengklok, *cluster* 1 berjumlah 16 di Kecamatan Klari, Tirtamulya, Jatisari, Banyusari, Kotabaru, Cilamaya Kulon, Lemahabang, Telagasari, Karawang Timur, Tempuran, Kutawaluya, Jayakarta, Pedes, Cibuyaya, Tirtajaya dan Batujaya, *cluster* 2 berjumlah 1 di Kecamatan Cikamaya Wetan dan *cluster* 3 berjumlah 11 di Kecamatan Pangkalan, Tegalwaru, Ciampel, Teluk Jambe Timur, Teluk Jambe Barat, Cikampek, Purwasari, Majalaya, Rawamerta, Cilebar Dan Pakisjaya. Berdasarkan analisis menggunakan metode *Fuzzy C-Means*, data penyandang masalah kesejahteraan sosial ini menjadi hal baru untuk diteliti terutama dalam penggunaan algoritma *Fuzzy C-Means* dan aplikasi Rapidminer versi 10.1. Saran dari penelitian ini diharapkan dapat dilakukan penelitian kembali dengan algoritma yang berbeda.

REFERENSI

- [1] N. N. A. Nanda, N. Ulinuha, and H. Khaulasari, "Pengelompokan Kecamatan di Wilayah Kabupaten Bojonegoro Berdasarkan Jenis Penyandang Masalah Kesejahteraan Sosial Menggunakan Fuzzy C-Means Clustering," *Techno.COM*, vol. 22, no. 2, pp. 499–507, 2023.
- [2] R. A. Prasetyo *et al.*, "Tekanan Kemiskinan Penyandang Masalah Kesejahteraan Sosial (PMKS) pada Masa Pandemi di Kabupaten Gresik Program Studi Ilmu Politik , Universitas Wijaya Kusuma Surabaya Poverty Pressure for People with Social Welfare Problems (PMKS) During the Pandemic," *J. Sociol. Pendidik. Humanis*, vol. 7, no. 2, pp. 12–26, 2022.
- [3] E. Syafaqoh, N. Ulinuha, and L. Hakim, "Klusterisasi Penyandang Masalah Kesejahteraan Sosial (PMKS) Di Kabupaten Bojonegoro Menggunakan Algoritma K-Medoids," *KUBIK J. Publ. Ilm. Mat.*, vol. 7, no. 2, pp. 78–87, 2023, doi: 10.15575/kubik.v7i2.21653.
- [4] G. N. S. Putri, D. Ispriyanti, and T. Widiharih, "Implementasi Algoritma Fuzzy C-Means Dan Fuzzy Possibilistics C-Means Untuk Klasterisasi Data Tweets Pada Akun Twitter Tokopedia," *J. Gaussian*, vol. 11, no. 1, pp. 86–98, 2022, doi: 10.14710/j.gauss.v11i1.33996.
- [5] R. Rahmati and A. W. Wijayanto, "Analisis Cluster Dengan Algoritma K-Means, Fuzzy C-Means Dan Hierarchical Clustering," *JIKO (Jurnal Inform. dan Komputer)*, vol. 5, no. 2,

- pp. 73–80, 2021.
- [6] K. E. Setiawan, A. Kurniawan, A. Chowanda, and D. Suhartono, “Model pengelompokan rumah sakit di Jakarta menggunakan fuzzy,” *Procedia Ilmu Komput.*, vol. 216, pp. 356–363, 2023.
- [7] I. F. Ashari, E. Dwi Nugroho, R. Baraku, I. Novri Yanda, and R. Liwardana, “Analysis of Elbow, Silhouette, Davies-Bouldin, Calinski-Harabasz, and Rand-Index Evaluation on K-Means Algorithm for Classifying Flood-Affected Areas in Jakarta,” *J. Appl. Informatics Comput.*, vol. 7, no. 1, pp. 89–97, 2023, doi: 10.30871/jaic.v7i1.4947.
- [8] I. Nanda Nur Rahmawati, M. Ivan Ariful Fathoni, “Penentuan Penerima KIP Kuliah Mahasiswa S1 Unugiri Menggunakan Fuzzy C-Means Clustering,” *Transform. J. Pendidik. Mat. dan Mat.*, vol. 6, no. 2, pp. 121–130, 2022, doi: 10.36526/tr.v.
- [9] V. S. Thalapala and K. Guravaiah, “FCMCP : Fuzzy C-Means untuk Penempatan Kontroler dalam Jaringan yang Ditentukan Perangkat Lunak,” *Procedia Ilmu Komput.*, vol. 201, pp. 109–116, 2022.
- [10] M. F. Julianto, S. W. Hadi, S. Setiaji, W. Gata, and R. Pebrianto, “Clustering Pencapaian Target Penjualan Rumah Para Karyawan Marketing Menggunakan Rapid Miner Dan Algoritma K-Means,” *Bianglala Inform.*, vol. 8, no. 2, pp. 79–85, 2020, doi: 10.31294/bi.v8i2.8189.
- [11] I. A. Darmawan, M. F. Randy, I. Yuniarto, M. M. Mutoffar, and M. T. P. Salis, “Penerapan Data Mining Menggunakan Algoritma Apriori Untuk Menentukan Pola Golongan Penyandang Masalah Kesejahteraan Sosial,” *Sebatik*, vol. 26, no. 1, pp. 223–230, 2022, doi: 10.46984/sebatik.v26i1.1622.
- [12] F. I. Putri, R. Damayanti, and Kismiantini, “Penerapan Algoritma K-Means Untuk Mengelompokan Kecamatan Di Kabupaten Gunungkidul Berdasarkan Program Keluarga Harapan,” *Pros. Semin. Nas. Mat. Stat. dan Apl.*, vol. 2, pp. 408–418, 2022.
- [13] T. Fitri, A. N. Rahma, and C. C. Marzuki, “Pengelompokan Data Penerima Bantuan Program Keluarga Harapan (Pkh) di Kecamatan Tuah Madani Pekanbaru Menggunakan Fuzzy C-Means,” *JSMS (Jurnal Sains Mat. dan Stat.)*, vol. 9, no. 1, pp. 1–10.
- [14] S. Khairunnisa and M. I. Jambak, “Pengelompokan Cuaca Kota Palembang Menggunakan Algoritma K-Means Clustering Untuk Mengetahui Pola Karakteristik Cuaca,” *J. Media Inform. Budidarma*, vol. 6, no. 4, p. 2352, 2022, doi: 10.30865/mib.v6i4.4810.
- [15] Y. R. Sari, A. Sudewa, D. A. Lestari, and T. I. Jaya, “Penerapan Algoritma K-Means Untuk Clustering Data Kemiskinan Provinsi Banten Menggunakan Rapidminer,” *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 5, no. 2, p. 192, 2020, doi: 10.24114/cess.v5i2.18519.
- [16] F. N. Hayati, M. Silfiani, and D. Nurlaily, “Perbandingan Pengelompokan Pusat Kesehatan Masyarakat Di Kota Balikpapan Menggunakan Metode K-Means Dan Fuzzy C-Means, Comparison of Groupings of Community Health Centers in Balikpapan City Using K-Means and Fuzzy C-Means Methods,” *Var. J. Stat. Its Appl.*, vol. 5, no. April, pp. 55–66, 2023.