

PERANCANGAN APLIKASI *DATA MINING* UNTUK KLUSTERISASI PENYEBARAN PENYAKIT DI KABUPATEN XYZ DENGAN ALGORITMA K-MEANS DAN FUZZY C-MEANS

Anugrayani Bustamin, Angela Hervina Gosal

Departemen Teknik Informatika, Fakultas Teknik, Universitas Hasanuddin
Jl. Poros Malino Km.6, Bontomarannu, Gowa, Sulawesi Selatan, Indonesia
anugrayani@gmail.com, angelahervina@gmail.com

ABSTRAK

Perkembangan data yang semakin meningkat saat ini menjadi tantangan tersendiri untuk memberikan kontribusi agar data tersebut dapat diolah menjadi informasi yang lebih bermanfaat. Salah satu pemanfaatan data yang dapat diterapkan yaitu pengelompokan data penyebaran penyakit sebagai informasi acuan bagi Pemerintah Daerah (Pemda) untuk membuat program intervensi yang lebih tepat sasaran baik dari sisi alokasi dana maupun penanganan cepat tanggap. Teknik yang dapat digunakan untuk menyelesaikan masalah tersebut dengan data mining. Penelitian ini akan membandingkan algoritma data mining K-Means dan *Fuzzy C-Means* yang berfokus pada pengelompokan atau klustering data. Hasil yang diperoleh pada penelitian ini, baik Algoritma K-Means maupun *Fuzzy C-Means* sama-sama berhasil melakukan pengelompokan data penyebaran penyakit dengan 3 kluster pada iterasi 5. Fungsi objektif yang diperoleh K-Means sebesar 0,0575 sedangkan *Fuzzy C-Means* sebesar 0,0783.

Kata kunci : *Data Mining, Algoritma K-Means, Algoritma Fuzzy C-Means*

1. Pendahuluan

1.1. Latar Belakang

Kementerian Kesehatan melalui Dinas Kesehatan di setiap daerah secara rutin mengumpulkan data kesehatan penduduk untuk diolah sehingga dihasilkan ranking Provinsi dan Kabupaten/Kota Sehat^[1]. Selain dalam rangka melakukan perankingan, data tersebut sebenarnya dapat diolah untuk membuat program intervensi yang lebih tepat sasaran baik dari sisi alokasi dana maupun penanganan cepat tanggap terhadap penyakit tertentu. Dengan pemanfaatan seperti ini, tidak ada lagi istilah kita berada pada kondisi “kebanjiran” data tetapi “miskin” informasi.

Data mining pada dasarnya merujuk pada proses mengekstrak atau “menambang” informasi (*knowledge*) dari data yang sangat banyak. *Data mining* biasanya disebut juga dengan penambangan informasi, ekstraksi informasi, analisis pola/data, atau penemuan informasi dalam database (*Knowledge Discovery in Databases*)^[2]. Sehingga dapat disimpulkan bahwa *data mining* adalah proses menemukan informasi yang penting dari data yang sangat banyak melalui proses ekstraksi pola atau fitur, yang tersimpan dalam database, gudang data (*data warehouse*), atau tempat penyimpanan data lainnya.

Proses *data mining* dapat dilakukan dengan menggunakan metode-metode *data mining*. Masing-masing metode tersebut memiliki beberapa algoritma yang dapat diterapkan untuk menyelesaikan berbagai permasalahan *data mining*. Permasalahan *data mining* dapat berupa proses pengelompokan data, pembentukan pola, serta pencarian hubungan sebab-akibat antarobjek untuk data dengan skala yang besar. Dalam menemukan informasi yang bermanfaat pada kasus data dengan skala besar, salah satu teknik alternatif yang dapat diimplementasikan yaitu klusterisasi. Teknik klusterisasi sangat bermanfaat dalam prediksi, pemetaan dan sebagainya. Hasil klustering yang baik akan menghasilkan tingkat kesamaan yang tinggi dalam satu kelas dan tingkat kesamaan yang rendah antar kelas. Kesamaan yang dimaksud merupakan pengukuran secara numerik terhadap dua buah objek.

Penelitian yang dilakukan Ridlo dkk tahun 2017 melakukan pemetaan produktivitas tanaman padi di Kabupaten Karawang dengan menggunakan algoritma K-Means. Pemetaan dilakukan dengan melibatkan atribut luas tanaman dan produksi padi. Hasil dari pemetaan divisualisasikan pada sebuah peta di web sehingga dapat diketahui kecamatan yang memiliki produktivitas kurang, sesuai dan lebih dari target ^[3]

Pada tahun 2018, Cheng melakukan penelitian mengenai *Importance-Performance Analysis* (IPA) dengan menggunakan algoritma *Fuzzy C-Means* untuk melakukan evaluasi terhadap destinasi wisata yang ramah terhadap turis. Hasil dari penelitian ini menyatakan bahwa IPAFCM merupakan metode yang lebih simpel, efektif, dan lebih mudah dimengerti. Metode ini juga dapat mengatasi permasalahan klasifikasi atribut yang tidak dapat diklasifikasikan pada IPA tradisional^[4].

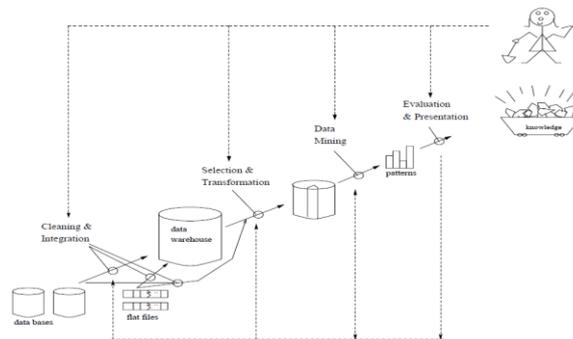
Berdasarkan beberapa penelitian terkait tersebut, penelitian ini akan membandingkan kinerja algoritma pengelompokan data yaitu Algoritma K-Means dan Fuzzy C-Means dengan studi kasus data penyebaran penyakit diare di Kabupaten XYZ untuk mendapatkan informasi yang dapat dimanfaatkan kedepannya.

2. Studi Literatur

2.1. Data Mining

Data yang sangat melimpah, bersama dengan kebutuhan analisis data, seringkali dijelaskan sebagai situasi “kaya data tetapi miskin informasi”. Jumlah data yang berkembang pesat, yang dikumpulkan dan disimpan dalam *database* yang besar, telah jauh melampaui kemampuan pemahaman manusia tanpa adanya bantuan alat (*tools*) yang hebat. Sehingga akhirnya, data yang terkumpul dalam *database* akan menjadi “kuburan data” yakni arsip data yang jarang dikunjungi. Akibatnya, keputusan penting sering diambil bukan berdasarkan informasi yang didapat dari kekayaan data yang tersimpan dalam *database* tetapi lebih kepada insting seorang pengambil keputusan. Dengan adanya *tools data mining* dapat dilakukan analisis data berupa penemuan pola data yang penting, memberikan kontribusi besar terhadap strategi bisnis, dasar

pengetahuan, serta penelitian ilmiah dan medis. Setelah proses analisis data dilakukan, data bertransformasi menjadi informasi yang berguna dan bermanfaat. Jarak besar antara data dan informasi memerlukan pengembangan *tools data mining* secara sistematis yang akan mengubah “data kuburan” menjadi “bongkahan emas” dari pengetahuan [2].



Gambar 2.1. Tahapan *data mining*

Secara sederhana, *data mining* merujuk kepada proses menambang atau mengekstrak pengetahuan dari data dengan jumlah yang sangat besar. Terdapat pula berbagai istilah yang memiliki makna yang sama dengan *data mining*, seperti ekstraksi informasi, analisis data/pola, pengerukan data, dan penemuan pengetahuan dalam *database* (*Knowledge Discovery in Databases* atau KDD). Gambar 2.1 merupakan tahapan proses *data mining* yang terbagi menjadi beberapa langkah berurut sebagai berikut [2] :

- a. *data cleaning* atau pembersihan data yang bertujuan untuk menghilangkan bising (*noise*) dan data-data yang tidak relevant
- b. *data integration* atau integrasi data yang bertujuan untuk menggabungkan beberapa sumber data
- c. *data selection* atau seleksi data di mana data yang relevan dengan tugas analisis diambil kembali dari *database*
- d. *data transformation* atau transformasi data di mana data diubah ke bentuk yang sesuai untuk proses pengolahan data

- e. *data mining* atau penambangan data yang merupakan proses inti di mana metode cerdas diterapkan agar dapat mengekstrak pola data
- f. *pattern evaluation* atau evaluasi pola untuk mengidentifikasi pola penting yang sesungguhnya yang dapat menginterpretasikan suatu pengetahuan berdasarkan tingkat kepentingannya
- g. *knowledge presentation* atau presentasi informasi/pengetahuan di mana teknik visualisasi dan representasi pengetahuan digunakan untuk menunjukkan informasi yang didapatkan pada pengguna.

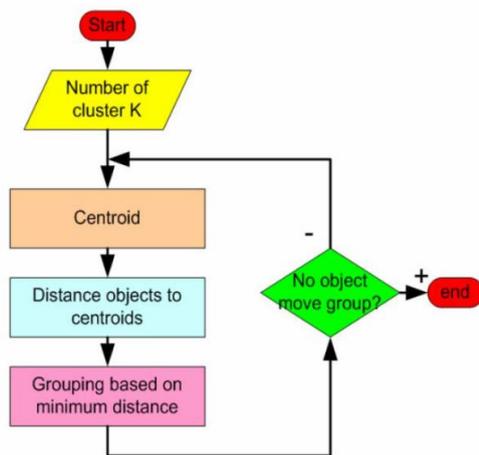
2.3. Metode Clustering

Clustering merupakan salah satu metode *data mining* yang mengelompokkan data berdasarkan informasi yang menjelaskan mengenai objek tersebut dan hubungan antarobjeknya. Tujuan yang ingin dicapai adalah bahwa objek dalam satu kelompok bersifat mirip atau berelasi antara yang satu terhadap yang lainnya dan berbeda atau tidak berelasi dengan objek dalam kelompok lainnya. Kemiripan objek tersebut dapat ditentukan dari besar kecilnya jarak antarobjek pengelompokannya. Semakin besar jarak antarobjek dalam kelompok dan semakin kecil jarak objek antarkelompok, maka semakin baik dan semakin jelas kelompok data yang terbentuk [2].

2.3.1. Algoritma K-Means

Salah satu algoritma yang digunakan dari metode *clustering* adalah algoritma *K-Means*. Algoritma *K-Means* merupakan algoritma *clustering* yang melakukan pengelompokan data sejumlah n sebanyak k cluster, di mana n dan k merupakan bilangan bulat positif. Pengelompokan ini dilakukan dengan meminimalisir total jarak antara data dan titik pusat (*centroid*) dari cluster yang bersesuaian [6].

Secara sederhana, langkah-langkah dalam menyelesaikan algoritma *K-Means* digambarkan dalam Gambar 2.2 berikut [7] :



Gambar 2.2. Flowchart algoritma K-Means

Flowchart di Gambar 2.2 melakukan proses sebagai berikut [2] :

- Menentukan jumlah *cluster* (K) yang ingin dibentuk. Jika perlu, tetapkan ambang batas (T) perubahan posisi *centroid*.
- Menentukan titik *centroid* dari setiap *cluster* yang akan dibentuk. Pada awal iterasi, penentuan *centroid* dilakukan secara acak. Sedangkan untuk iterasi selanjutnya, penentuan *centroid* dilakukan dengan menghitung rata-rata nilai dari setiap data dari masing-masing *cluster*.
- Menghitung jarak antara setiap data terhadap setiap titik *centroid*. Perhitungan jarak ini biasanya dilakukan dengan menggunakan persamaan *Euclidean Distance* yakni:

$$D(x_i, c_j) = \sqrt{\sum (x_i - c_j)^2}$$

di mana:

- $D(x_i, c_j)$ adalah jarak antara data x_i terhadap c_j
 - x_i adalah data ke- i di mana i adalah jumlah data
 - c_j adalah *centroid* cluster j di mana j adalah jumlah *cluster*
- Mengelompokkan objek. Penentuan anggota *cluster* dilakukan dengan membandingkan jarak *centroid* dari masing-masing *cluster*. Data tersebut akan dikelompokkan dalam *cluster* yang memiliki jarak *centroid* paling kecil dari semua *cluster*.

- e. Jika anggota setiap *cluster* masih berpindah, maka dilakukan perulangan dari langkah ke-2 hingga akhir dan berhenti saat semua data anggota *cluster* tidak berpindah lagi ke *cluster* lain.

2.3.2. Algoritma *Fuzzy C-Means*

Metode *clustering* dengan algoritma *Fuzzy C-Means* (FCM) didasarkan pada teori logika *fuzzy*. Teori ini pertama kali diperkenalkan oleh Lotfi Zadeh pada tahun 1965 dengan nama himpunan *fuzzy* (*fuzzy set*). Dalam teori *fuzzy*, keanggotaan sebuah data tidak diberikan nilai secara tegas dengan nilai 1 (menjadi anggota) dan 0 (tidak menjadi anggota), melainkan dengan suatu nilai derajat keanggotaan yang jangkauan nilainya 0 sampai 1. Nilai keanggotaan suatu data dalam sebuah himpunan menjadi 0 ketika sama sekali tidak menjadi anggota, dan menjadi 1 ketika menjadi anggota secara penuh dalam suatu himpunan. Umumnya, nilai keanggotaannya antara 0 dan 1. Semakin tinggi nilai keanggotaannya maka semakin tinggi derajat keanggotaannya, dan semakin kecil nilainya maka semakin rendah derajat keanggotaannya^[8].

FCM sebenarnya merupakan versi *fuzzy* dari algoritma *K-Means*. Contoh sederhananya adalah berupa suhu, yang umumnya terbagi menjadi dua yakni panas atau dingin. Suhu 10° disebut dingin sedangkan suhu 40° disebut panas. Bagaimana dengan suhu 20° ? Jika pada himpunan tegas menyatakan batas suhu dingin adalah 20° maka suhu 21° dinyatakan panas. Dalam teori *fuzzy*, untuk menentukan suhu panas atau dingin tersebut ditentukan dengan derajat keanggotaan, misalnya suhu 20° disebut 70% dingin dan 30% panas, dan sebagainya. Dengan demikian, penentuan status sebuah data pada setiap himpunan didasarkan pada nilai derajat keanggotaan pada setiap himpunan.

Asumsikan terdapat sejumlah data dalam set data X yang berisi n data dan berdimensi m . Data tersebut akan dikelompokkan dengan menghitung jaraknya terhadap setiap *centroid* ($c = c_1, c_2, \dots, c_k$) yang terbentuk, di mana k adalah jumlah *cluster*. Setiap data memiliki nilai derajat keanggotaan pada setiap *cluster*, yang dinyatakan dengan u_{ij} yang bernilai

antara 0 sampai 1, i menyatakan data x_i dan j menyatakan cluster c_j . Algoritma FCM terbagi menjadi beberapa tahap sebagai berikut ^[9]:

1. Melakukan inisialisasi dengan menentukan jumlah cluster yang akan dibentuk ($k \geq 2$), menentukan bobot pangkat ($w > 1$), menentukan ambang batas perubahan nilai fungsi objektif (T). Parameter c_j adalah cluster ke- j sedangkan w adalah bobot pangkat (*weighting exponent*) yang bernilai $w > 1$ dan umumnya diberi nilai 2.
2. Menetapkan nilai derajat keanggotaan awal secara acak pada matriks partisi fuzzy (U) dengan syarat jumlah setiap data selalu sama dengan 1. Nilai keanggotaan tersebut disimpan dalam matriks partisi fuzzy berukuran $N \times k$, di mana baris merupakan data sedangkan kolom merupakan nilai keanggotaan pada setiap cluster. Bentuknya adalah sebagai berikut:

$$U = \begin{bmatrix} u_{11}(x_1) & u_{12}(x_1) & \dots & u_{1k}(x_1) \\ \vdots & \ddots & & \vdots \\ u_{n1}(x_n) & \dots & & u_{nk}(x_n) \end{bmatrix} \dots\dots\dots(2)$$

3. Menghitung titik *centroid* dari masing-masing cluster menggunakan persamaan berikut:

$$c_{ij} = \frac{\sum_{i=1}^N (U_{ij})^w \cdot x_{ij}}{\sum_{i=1}^N (U_{ij})^w} \dots\dots\dots(3)$$

4. Membentuk matriks partisi yang baru dengan menghitung derajat keanggotaan setiap data pada setiap cluster menggunakan persamaan berikut:

$$u_{ij} = \frac{D(x_i, c_j)^{\frac{-2}{w-1}}}{\sum_{l=1}^k D(x_i, c_l)^{\frac{-2}{w-1}}} \dots\dots\dots(4)$$

5. Menghitung nilai fungsi objektif dengan menggunakan persamaan berikut:

$$J = \sum_{i=1}^N \sum_{l=1}^k (u_{ij})^w D(x_i, c_l)^2 \dots\dots\dots(5)$$

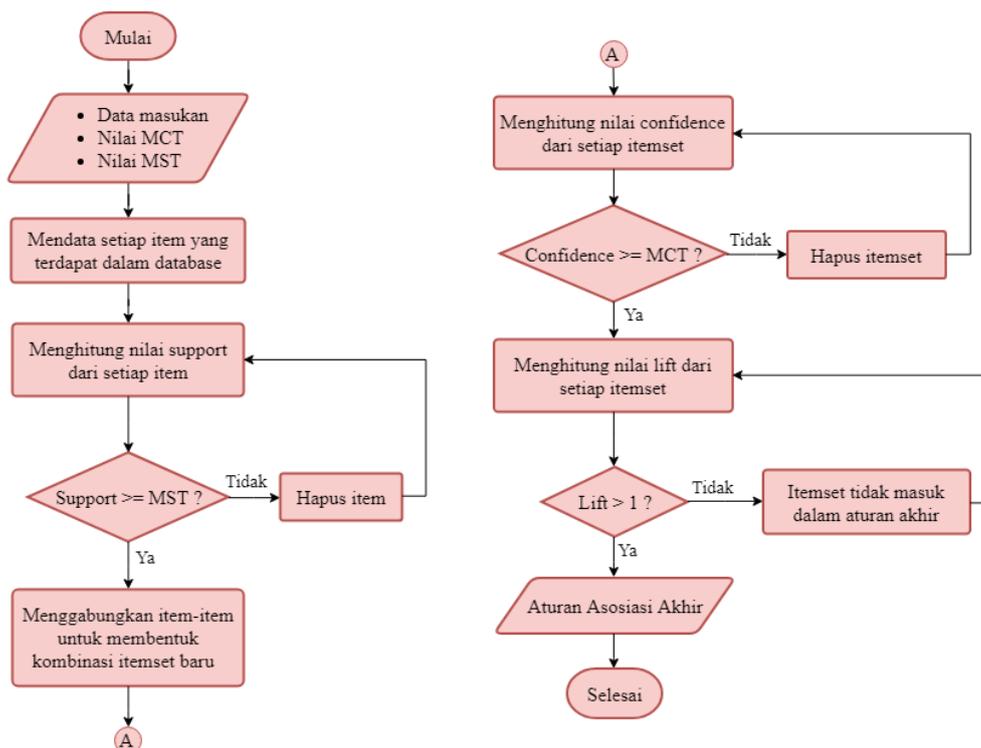
6. Mengulangi langkah 3 sampai 5 sampai nilai perubahan fungsi objektif yang didapatkan lebih rendah dari nilai ambang batas yang diterapkan.

3. Metodologi

Penelitian ini dilakukan dengan beberapa tahapan yaitu pengumpulan data, perancangan algoritma, implementasi data ke algoritma dan terakhir analisa hasil perbandingan algoritma. Data yang digunakan merupakan data penyebaran penyakit di Kabupaten XYZ. Setiap kecamatan pada Kabupaten XYZ berisi data terkait jumlah desa dan jumlah penderita diare setiap bulannya. Data tersebut akan dikelompokkan berdasarkan zona tingkat penderita. Untuk penelitian ini klusterisasi akan dikategorikan dalam 3 kelompok kluster.

3.1. Perancangan Algoritma *K-Means*

Program algoritma *K-Means* dirancang berdasarkan alur proses kerja algoritma sebagaimana yang ditunjukkan oleh Gambar 3.2 berikut.



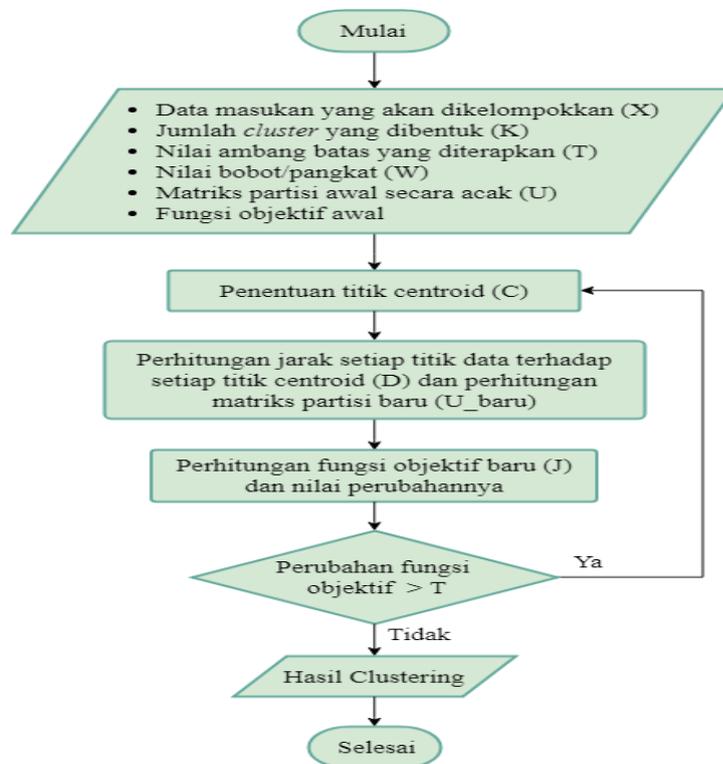
Gambar 3.2. Proses algoritma *K-Means*

Flowchart pada Gambar 3.2 menjelaskan proses kerja algoritma *K-Means* yang diuraikan sebagai berikut:

- a. Memasukkan parameter yang digunakan yakni data yang ingin dikelompokkan, jumlah *cluster* yang ingin dibentuk (K), dan nilai ambang batas (T) perubahan posisi *centroid*.
- b. Menentukan titik *centroid* awal dari setiap *cluster* yang akan dibentuk. Penentuan titik *centroid* awal dapat dilakukan secara acak. Namun, agar hasil pengelompokan yang didapatkan tidak berubah-ubah, maka penentuan titik *centroid* awal dalam aplikasi modul praktikum ini dilakukan dengan mengambil nilai dari titik data itu sendiri. Jadi, semisal dari 10 data ingin diambil 3 titik *centroid*, maka titik *centroid* pertama diambil dari titik data pertama, titik *centroid* kedua diambil dari titik data kedua, dan titik *centroid* ketiga diambil dari nilai rata-rata data lainnya yakni data ketiga sampai kesepuluh.
- c. Memulai iterasi dengan menghitung jarak setiap titik data terhadap setiap titik *centroid* dengan menggunakan persamaan *Euclidean Distance*. Setelah itu, membandingkan jarak data dari *centroid* yang satu terhadap *centroid* yang lainnya. Data tersebut akan dikelompokkan ke dalam *cluster* yang memiliki jarak *centroid* paling kecil dari semuanya.
- d. Mengamati apakah terdapat data yang berpindah cluster dengan menghitung nilai fungsi objektif. Nilai fungsi objektif didapatkan dengan menjumlahkan jarak setiap data dalam masing-masing *cluster* yang ditempatinya. Setelah itu, membandingkan nilai fungsi objektif lama dan baru untuk mendapatkan nilai perubahan fungsi objektif. Nilai perubahan fungsi objektif yang melebihi nilai ambang batas menandakan bahwa perubahan posisi *centroid* belum memenuhi nilai ambang batas yang ditetapkan sehingga proses iterasi dilanjutkan.
- e. Menghitung titik *centroid* yang baru dan melanjutkan proses iterasi.

3.3. Perancangan Algoritma *Fuzzy C-Means*

Program algoritma FCM dirancang berdasarkan alur proses kerja algoritma sebagaimana yang ditunjukkan oleh Gambar 3.3 berikut.



Gambar 3.3. Proses Algoritma FCM

Flowchart pada Gambar 3.3 menjelaskan proses kerja algoritma FCM yang diuraikan sebagai berikut:

1. Memasukkan parameter yang digunakan yakni X , K , T , W , U . Berikut ini merupakan potongan kode program untuk memuat parameter masukan.
2. Menentukan titik *centroid* dari setiap *cluster* yang akan dibentuk. Berikut ini merupakan potongan kode program untuk menghitung titik *centroid*.
3. Menghitung jarak dari setiap data terhadap masing-masing titik *centroid* dengan menggunakan persamaan *Euclidean Distance*. Bersamaan dengan itu, dilakukan pula perhitungan nilai derajat

keanggotaan baru untuk setiap elemen data sehingga membentuk matriks partisi yang baru (U_{baru}). Berikut ini merupakan potongan kode program untuk menghitung jarak dan matriks partisi baru.

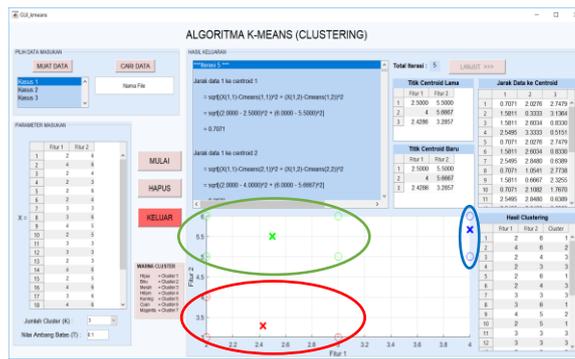
4. Menghitung nilai fungsi objektif baru (J). Berikut ini merupakan potongan program untuk menghitung nilai J .
5. Mengulangi langkah 2 sampai 4 jika nilai $J > T$ dan berhenti jika nilai $J < T$.

4. Hasil dan Pembahasan

Penelitian ini membandingkan kinerja dari algoritma K-Means dan *Fuzzy C-Means* untuk mengelompokkan data penyebaran penyakit diare di Kabupaten XYZ.

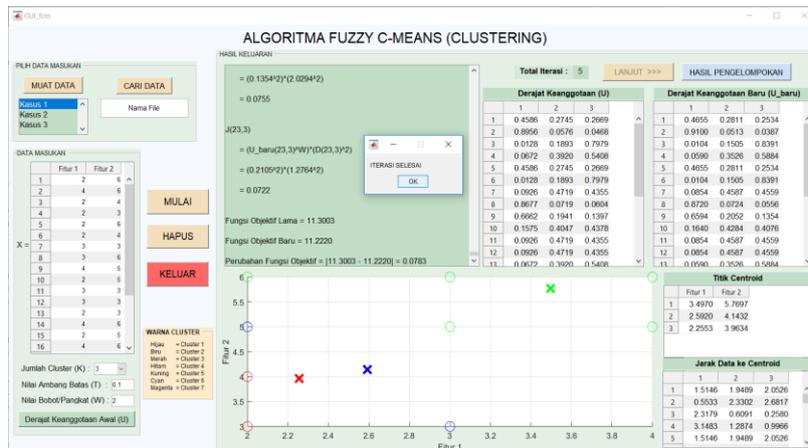
Algoritma *K-Means* diawali dengan inisialisasi 3 titik *centroid* awal. Titik *centroid* pertama diambil dari dataset 1, titik *centroid* kedua diambil dari dataset 2, dan titik *centroid* ketiga diambil dari rata-rata dataset lainnya. Posisi dari setiap dataset dan setiap *centroid* ditampilkan melalui grafik yang tersedia. Kemudian dilakukan perhitungan jarak dari setiap titik dataset terhadap setiap titik *centroid* dan dilanjutkan dengan melakukan pengelompokan berdasarkan jarak *centroid* terkecil yang didapatkan. Pada iterasi kedua, titik *centroid* berpindah posisi sesuai dengan titik *centroid* baru yang didapatkan pada iterasi pertama. Dari hasil pengelompokan pada iterasi kedua, terjadi perpindahan *cluster*.

Proses algoritma yang dilalui pada iterasi selanjutnya sama dengan proses algoritma pada iterasi kedua. Pada penelitian ini, hasil terbaik di peroleh pada iterasi kelima karena nilai perubahan fungsi objektif yang sudah di bawah nilai ambang batas (0.1) yakni yakni 0.0575 sebagaimana yang ditunjukkan pada Gambar 4.1.



Gambar 4.1 Hasil klusterisasi algoritma *K-Means*

Tidak jauh berbeda dengan Algoritma K-Means, Algoritma FCM juga diawali dengan pembentukan 3 titik *centroid* awal sebagai iterasi pertama. Titik *centroid* tersebut didapatkan dari perhitungan derajat keanggotaan yang dipangkatkan dengan bobot W . Setelah titik *centroid* dari setiap *cluster* didapatkan maka langkah selanjutnya adalah menghitung jarak setiap data terhadap setiap titik *centroid* dengan menggunakan persamaan *Euclidean distance*. Hasil perhitungan jarak ditampilkan pada tabel jarak data ke *centroid*. Bersamaan dengan itu, dilakukan pula perhitungan nilai derajat keanggotaan yang baru (U_{baru}). Karena iterasi ini merupakan iterasi yang pertama maka hasil pengelompokan yang didapatkan masih belum tetap dan perlu dilanjutkan ke iterasi selanjutnya. Pada iterasi kedua didapatkan nilai perubahan fungsi objektif sebesar 1.2464. Nilai tersebut masih di atas nilai ambang batas yang ditentukan sehingga proses algoritma dilanjutkan ke iterasi selanjutnya dengan tahapan proses yang sama. Proses iterasi selesai pada iterasi kelima dengan nilai perubahan fungsi objektif sebesar 0.0783. Karena nilai tersebut lebih kecil dari nilai ambang batas maka proses iterasi selesai.



Gambar 4.2 Tampilan ketika iterasi pada kasus 3 algoritma FCM selesai.

Tabel 4.1 akan menampilkan data hasil kluster yang telah diimplementasikan pada 2 algoritma klusterisasi yaitu *K-Means* dan *Fuzzy C-Means*.

Tabel 4.1 Hasil Klustering Data dengan Algoritma *K-Means* dan *Fuzzy C-Means*

Fitur 1	Fitur 2	<i>K-Means</i>	<i>FCM</i>
2	6	1	1
4	6	2	1
2	4	3	3
2	3	3	3
2	6	1	1
2	4	3	3
3	3	3	2
3	6	1	1
4	5	2	1
2	5	1	2
3	3	3	2
3	3	3	2
2	3	3	3
4	6	2	1

Fitur 1	Fitur 2	<i>K-Means</i>	<i>FCM</i>
2	5	1	2
4	6	2	1
3	6	1	1
4	6	2	1
3	5	3	1
2	6	1	1
3	5	3	1
4	5	2	1
3	5	3	1

Berdasarkan hasil yang diperoleh, kinerja kedua algoritma dapat dilihat dari jumlah iterasi yang digunakan dalam menghasilkan fungsi objektif terbaik. Algoritma K-Means dengan fungsi objektif sebesar 0,0575 memberikan hasil yang baik dibandingkan dengan Fuzzy C-Means sebesar 0,0783 pada kasus data pengelompokan penyebaran penyakit diare di Kabupaten XYZ.

5. Kesimpulan & Saran

Dari hasil analisis dan pengujian yang telah dilakukan penelitian ini, Algoritma K-Means dan Fuzzy C-Means sama-sama berhasil melakukan pengelompokan pada iterasi kelima dengan perubahan fungsi objektif yang sudah di bawah nilai ambang batas (0.1) . Fungsi objektif yang diperoleh K-Means sebesar 0,0575 sedangkan *Fuzzy C-Means* sebesar 0,0783 pada masing-masing penentuan 3 kluster data. Penelitian selanjutnya dapat dikembangkan dengan algoritma klusterisasi lainnya dengan jumlah kluster yang berbeda.

DAFTAR PUSTAKA

- [1] D. Kesehatan, "24 Indikator Kesehatan Dalam IPKM," 1 April 2012. [Online]. Available: <http://www.depkes.go.id/index.php?vw=2&id=1337>. [Accessed 18 November 2018]
- [2] M. Kamber, M. Kaufmann, and P. All, "Data Mining: Concepts and Techniques," 2000.
- [3] Ridlo. M. Rosyd, Defiyanti. Sofi, Primajaya. Aji, "Implementasi Algoritma K-Means untuk Pemetaan Produktivitas Panen Padi di Kabupaten Karawang, CITEE 2017, pp. 426-433
- [4] Cheng. Shou-Cheng, "Applying Importance-Performance Analysis by Fuzzy C-Means to Evaluate Tourist Friendly Destination", International Conference on Machine Learning and Cybernetics, 2018.
- [5] P.-N. Tan, M. Steinbach, and V. Kumar, "Chap 8 : Cluster Analysis: Basic Concepts and Algorithms," *Introd. to Data Min.*, p. Chapter 8, 2005.
- [6] B. K. Teknomo, "K-Means Clustering Tutorial," pp. 1–12, 2007.
- [7] N. Wakhidah, "Clustering Menggunakan K-Means Algorithm (K-Means Algorithm Clustering)," *Fak. Teknol. Inf.*, vol. 21, no. 1, pp. 70–80, 2014.
- [8] Zadeh, L.a, (1956). Fuzzy Sets. *Information and Control*, 8(3), 338-353.
- [9] Bezdek, J.C. Ehrlich, R & Full. W, (1984), *The Fuzzy C-Means Clustering Algorithm*. *Computers and Geosience*, 10(2-3), 191-203