

OPTIMASI FUZZY C-MEANS DAN K-MEANS MENGGUNAKAN ALGORITMA GENETIKA UNTUK PENGLASTERAN DATASET *DIABETIC RETINOPATHY*

Muhammad Ezar Al Rivan*¹, Steven², William Tanzil³

^{1,2,3}Teknik Informatika STMIK Global Informatika MDP

Email: ¹meedzhar@mdp.ac.id, ²stevenkaigeru@gmail.com, ³williamtanzil@gmail.com

*Penulis Korespondensi

(Naskah masuk: 12 Maret 2019, diterima untuk diterbitkan: 07 Oktober 2020)

Abstrak

Diabetic Retinopathy adalah komplikasi dari diabetes yang mengakibatkan gangguan pada retina mata. Gangguan tersebut dapat diketahui dengan deteksi awal melalui data yang diekstraksi dari citra mata. Deteksi awal dapat dilakukan dengan menggunakan metode *clustering*. Metode yang digunakan yaitu *Fuzzy C-Means* dan *K-Means*. *Fuzzy C-Means* dan *K-Means* memiliki kelemahan dari jumlah iterasi yang besar. Jumlah iterasi pada *Fuzzy C-Means* dan *K-Means* dapat dioptimasi dengan menggunakan Algoritma Genetika. Optimasi dilakukan dengan cara mengganti bagian pada *Fuzzy C-Means* dan *K-Means* pada saat menentukan pusat *cluster*. Dataset yang digunakan pada penelitian adalah dataset *Diabetic Retinopathy*. Hasil optimasi sebelum dan sesudah *hybrid* Algoritma Genetika pada *Fuzzy C-Means* terlihat pada nilai rata-rata iterasi dari 17,1 menjadi 6,65 terjadi penurunan sebesar 61,11% dan pada *K-Means* terlihat pada nilai rata-rata iterasi dari 10,85 menjadi 7,35 terjadi penurunan sebesar 32,25%. Berdasarkan hasil perbandingan nilai rata-rata iterasi Algoritma Genetika-*Fuzzy C-Means* dan Algoritma Genetika-*K-Means* maka dapat disimpulkan bahwa Algoritma Genetika-*Fuzzy C-Means* memiliki jumlah iterasi yang lebih baik dibanding Algoritma Genetika-*K-Means*. Algoritma Genetika-*Fuzzy C-Means* juga memiliki *inter cluster distance* yang paling kecil dan *intra cluster distance* yang paling besar.

Kata kunci: *Dataset diabetic retinopathy, GA, GA K-Means, GA FCM, Klaster*

FUZZY C-MEANS AND K-MEANS OPTIMIZATION USING GENETIC ALGORITHM FOR DIABETIC RETINOPATHY DATASET CLUSTERING

Abstract

Diabetic Retinopathy is a diabetic complication that causes retina disorder. Retina disorder can be known from data extracted from the eye image. Early detection conduct using clustering. These methods are *Fuzzy C-Means* and *K-Means*. These methods have a large number of iteration as a weakness. The number of iteration can be optimized using a genetic algorithm. Optimization conducted by replacing a part from *Fuzzy C-Means* dan *K-Means* that use to generate early centroid. The dataset used in the study is a dataset of diabetic retinopathy. The optimization results before and after the hybrid Genetic Algorithm on *Fuzzy C-Means* are the average iteration values decreased from 17.1 to 6.65, decreasing 61,11% and in *K-Means* are the average iteration values decreased from 10.85 to 7.35 decreasing 32,25%. Based on the comparison of Genetic Algorithm *Fuzzy C-Means* and Genetic Algorithm *K-Means* iterations, it can be concluded that the Genetic Algorithm *Fuzzy C-Means* has a better number of iteration than Genetic Algorithm *K-Means*. Genetic Algorithm-*Fuzzy-C-Means* has the smallest inter-cluster distance and biggest intracuster distance.

Keywords: *Diabetic retinopathy dataset, GA, GA K-Means, GA FCM, cluster*

1. PENDAHULUAN

Diabetic retinopathy adalah salah satu bentuk komplikasi diabetes melitus, di mana kadar gula yang tinggi pada akhirnya mengakibatkan kerusakan pada pembuluh darah retina mata, terutama di jaringan-jaringan yang sensitif terhadap cahaya. Kondisi ini dapat diderita oleh siapapun yang

menderita diabetes tipe 1 maupun 2, terutama mereka yang gula darahnya tidak terkontrol dan telah menderita diabetes dalam jangka waktu yang lama. Hal ini dikarenakan penderita diabetes memiliki resiko yang tinggi untuk memunculkan berbagai masalah kesehatan (Tjin, 2017).

Deteksi dini *Diabetic Retinopathy* dilakukan pada penelitian yang dilakukan (Heryawan, 2017) dengan morfologi matematika dengan teknik pengolahan citra untuk mengolah dataset gambar yaitu DIARETDB1 sebanyak 89 citra fundus mata dan e-optha sebanyak 200 fundus mata. Pengujian menggunakan dataset DIARETDB1 tersebut diperoleh hasil akurasi 86,51%, sensitivitas 90% dan spesifitas 55%, sementara pengujian menggunakan dataset e-optha diperoleh hasil akurasi 70.5%, sensitivitas 80% dan spesifisitas sebesar 60%.

Diabetic retinopathy dapat didiagnosis dengan melihat ciri-ciri pada mata. Dataset *Diabetic Retinopathy Debrecen* yang terdiri dari 1151 model mata dalam bentuk data numerik yang merupakan hasil ekstraksi dari citra mata. Dataset ini juga digunakan pada penelitian yang dilakukan oleh (Bahrami and Sajedi, 2018) untuk menguji performa algoritma *Random Forest*. Pada penelitian lain dataset ini juga digunakan oleh (Prayitno, 2018) untuk menguji performa algoritma *Kernel K-Means*.

Dataset *Diabetic Retinopathy* dapat digunakan untuk *clustering* seperti yang dilakukan oleh (Bahrami and Sajedi, 2018) dan (Prayitno, 2018). Namun dalam penelitian ini digunakan metode yang berbeda. Pada penelitian ini dataset *Diabetic Retinopathy* digunakan untuk menguji performa *hybrid K-Means* dan *hybrid Fuzzy C-Means*.

Beberapa metode *clustering* yang digunakan adalah *K-Means* dan *Fuzzy C-Means*. Kedua metode tersebut memiliki kelebihan dan kekurangan masing-masing. Penelitian mengenai perbandingan algoritma *K-Means* dan *Fuzzy C-Means* telah dilakukan (Yohannes, 2016). Berdasarkan penelitian tersebut, diperoleh jumlah iterasi yang dibutuhkan algoritma *K-Means* lebih sedikit dibandingkan dengan algoritma *Fuzzy C-Means*.

Algoritma *K-Means* dan *Fuzzy C-Means* dapat dioptimasi untuk mengurangi jumlah iterasi. Algoritma Genetika digunakan untuk optimasi *K-Means* dan *Fuzzy C-Means*. Pada penelitian yang dilakukan oleh (Singh *et al.*, 2013) *hybrid* Algoritma Genetika dilakukan pada *Fuzzy C-Means*. Penelitian tersebut dilakukan untuk mendeteksi perubahan pada dua citra yang berbeda. *Hybrid* Algoritma Genetika terhadap *Fuzzy C-Means* juga dilakukan pada penelitian (Liu and Hou, 2009). Penelitian yang dilakukan oleh (Hadi, Gede Darma Putra and Satya Kumara, 2017) juga menggunakan *hybrid* Algoritma Genetika dan *Fuzzy C-Means*. Pada penelitian lain *hybrid* Algoritma Genetika pada *K-Means* dilakukan oleh (Yong-Guo Liu, Ke-Fei Chen and Xue-Ming Li, 2005). Selain itu terdapat penelitian lain yaitu yang dilakukan oleh (Malki, Rizk, Shorbagy, Moussa, 2016) terkait dengan penggunaan Algoritma Genetika pada *K-Means* untuk menguji permasalahan *clustering*.

Pada penelitian ini Algoritma Genetika akan digunakan sebagai optimasi untuk *K-Means* dan *Fuzzy C-Means*. Pada algoritma *K-Means*,

Algoritma Genetika digunakan untuk menghasilkan kandidat pusat cluster. Kandidat pusat cluster kemudian akan dihitung jaraknya dengan data. Kandidat pusat cluster akan mengalami proses seleksi, *crossover* dan mutasi. Proses ini akan terus berlangsung sampai konvergensi tercapai.

Pada algoritma *Fuzzy C-Means*, Algoritma Genetika juga digunakan untuk menghasilkan kandidat pusat *cluster*. Kemudian antara data dan pusat *cluster* dihitung matriks *fuzzy* yang menunjukkan derajat keanggotaan *cluster*. Kandidat pusat *cluster* kemudian melalui proses seleksi, *crossover* dan mutasi. Proses ini akan berlangsung terus menerus sampai tercapai konvergensi.

Konvergensi pada *K-Means* dan *Fuzzy C-Means* tercapai ketika tidak terjadi perubahan fungsi objektif. Perubahan fungsi objektif dihitung dari selisih fungsi objektif pada iterasi sebelumnya dengan fungsi objektif saat ini. Optimasi dilakukan untuk mempercepat konvergensi atau dengan kata lain mengurangi jumlah iterasi.

2. METODE PENELITIAN

Dalam penelitian ini, metode yang diusulkan yaitu dengan melakukan optimasi pada algoritma *Fuzzy C-Means* dan *K-Means*. Optimasi dilakukan dengan menggunakan Algoritma Genetika. Algoritma Genetika akan disisipkan pada *Fuzzy C-Means* dan *K-Means*. Proses penyisipan yang dilakukan pada algoritma *Fuzzy C-Means* dan *K-Means* yaitu pada bagian inisialiasi pusat *cluster* dan *update* pusat *cluster*.

Pada algoritma *Fuzzy C-Means*, inisialisasi awal pusat *cluster* ditentukan oleh bilangan acak. Pada bagian ini, inisialisasi pusat *cluster* diganti oleh Algoritma Genetika. Algoritma Genetika akan menghasilkan kandidat pusat *cluster*. Selanjutnya bagian dari *Fuzzy C-Means* yang diganti oleh Algoritma Genetika yaitu pada bagian *update* pusat kluster. Kandidat pusat kluster yang dihasilkan oleh Algoritma Genetika kemudian mengalami proses seleksi, *crossover* dan mutasi.

Pada algoritma *K-Means*, inisialisasi awal pusat *cluster* dilakukan dengan menggunakan bilangan acak. Algoritma Genetika digunakan pada bagian ini untuk menghasilkan kandidat pusat *cluster*. Selanjutnya kandidat pusat *cluster* akan mengalami proses seleksi, *crossover* dan mutasi. Bagian ini digunakan untuk mengganti peran *update* pusat *cluster* pada algoritma *K-Means*.

Representasi kromosom Algoritma Genetika pada penelitian ini menggunakan representasi biner. Representasi biner dipilih karena representasi biner merupakan representasi kromosom yang sederhana. Proses seleksi menggunakan *Roulette Wheel Selection*. Untuk proses *crossover* menggunakan *singlepoint crossover*. Proses mutasi menggunakan *flip bit* menyesuaikan dengan representasi biner. Untuk menghitung *fitness* dilakukan dengan cara yang sama pada *Fuzzy C-Means* dan *K-Means*.

Fuzzy C-Means dan *K-Means* memiliki fungsi objektif sehingga untuk menghitung *fitness* maka fungsi objektif digunakan sebagai *fitness*.

2.1. Hybrid Algoritma Genetika - Fuzzy C-Means

Adapun langkah-langkah *hybrid* Algoritma Genetika dan *Fuzzy C-Means* sebagai berikut :

1. Menentukan parameter Algoritma Genetika
2. Menentukan bilangan acak sebanyak populasi sebagai kandidat pusat *cluster*
3. Menghitung derajat keanggotaan dan *fitness* masing-masing pusat kluster.
4. Melakukan proses seleksi, *crossover* dan mutasi.
5. Ulangi langkah 3 dan 4 sampai stabil

2.2. Hybrid Algoritma Genetika - K-Means

Adapun langkah-langkah *hybrid* Algoritma Genetika dan *K-Means* sebagai berikut:

1. Menentukan parameter Algoritma Genetika
2. Menentukan bilangan acak sebanyak populasi sebagai kandidat pusat *cluster*
3. Menghitung jarak minimum dan *fitness* masing-masing pusat *cluster*
4. Melakukan proses seleksi, *crossover* dan mutasi
5. Ulangi langkah 3 dan 4 sampai stabil

2.3. Diabetic Retinopathy Dataset

Dataset yang digunakan adalah dataset berupa data numerik mengenai ciri-ciri mata yang mengalami *diabetic retinopathy* pada laman resmi *UCI Machine Learning Repository* (Balint & Andras, 2014). Jumlah data sebanyak 1151 baris dengan atribut sebanyak 19. Adapun atribut tersebut yaitu nilai biner kualitas penilaian, nilai biner *pre-screening*, hasil dari bakteri MA yang terdeteksi dengan tingkat keyakinan alfa, hasil dari *exudates* (sesuatu yang keluar dari luka berupa cairan di mata) berupa *pixel* dari luka yang dinormalisasi dengan membagi luka dengan ROI, jarak *euclidean* dari titik tengah *macula* dan titik tengah *optic disc*, diameter *optic disc* dan nilai biner untuk hasil *AM/FM based*.

3. LANDASAN TEORI

Adapun tahapan pada algoritma *Fuzzy C-Means* (Dembéle and Kastner, 2003) dalam (Bezdek, 1981) yaitu :

1. Inisialisasi. Pada tahap ini yang dilakukan adalah menetapkan banyak kluster *K*, derajat *fuzziness* *m* dan matriks acak *u* yang menampung derajat keanggotaan. Pilih secara acak *K* sampel sebagai *centroid*. Lalu hitung matriks derajat keanggotaan antara data dengan *centroid*.
2. *Update centroid*. Pada tahap ini, dicari *centroid* baru.
3. Menghitung derajat keanggotaan baru

4. Ulangi langkah 1 sampai 3 hingga stabil yang dicirikan dengan tidak ada perubahan derajat keanggotaan.

Tahapan pada algoritma *K-Means* (Jain, 2008) dalam (Jain and Dubes, 1988) yaitu :

1. Memilih partisi awal sebanyak *K cluster*. Ulangi langkah 2 sampai 3 hingga stabil yang dicirikan tidak ada perubahan pusat *cluster*
2. Menentukan partisi baru berdasarkan jarak terdekat data dengan pusat *cluster*
3. Menghitung pusat *cluster* baru

4. HASIL DAN PEMBAHASAN

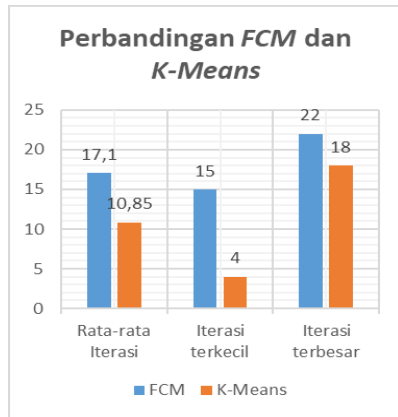
Dalam penelitian ini, ada 4 skenario eksperimen yang dilakukan yaitu melakukan *clustering* dengan menggunakan algoritma *Fuzzy C-Means*, *K-Means*, *Hybrid Algoritma Genetika - Fuzzy C-Means* dan *Hybrid Algoritma Genetika - K-Means*. Setiap skenario dilakukan sebanyak 20 kali. Setiap skenario akan dicatat iterasi yang diperlukan sampai konvergen. Selain itu, evaluasi dilakukan dengan menggunakan *intra cluster distance* dan *inter cluster distance* (Soler *et al.*, 2013). Hasil *clustering* yang baik adalah yang memiliki *intra cluster distance* yang besar dan *intercluster* kecil.

Pada Tabel 1. terdapat perbandingan iterasi antara *Fuzzy C-Means* dan *K-Means*. Dari Tabel 1 algoritma yang memiliki iterasi terkecil minimum yaitu *K-Means* dengan nilai 4. Begitu juga iterasi terbesar minimum yaitu *K-Means* dengan nilai 18. Secara rata-rata algoritma yang memiliki iterasi minimum adalah *K-Means* dengan nilai 10,85. Dapat diketahui bahwa *K-Means* memiliki iterasi minimum dibandingkan dengan *Fuzzy C-Means*. Gambar 1 menunjukkan perbandingan *Fuzzy C-Means* dan *K-Means* dalam bentuk Histogram.

Tabel 1. Perbandingan Iterasi FCM dan *K-Means*

	FCM	<i>K-Means</i>
Iterasi terkecil	15	4
Iterasi terbesar	22	18
Rata-rata iterasi	17,1	10,85

Pada Tabel 2 terdapat perbandingan *inter cluster distance* antara *Fuzzy C-Means* dan *K-Means*. Dari Tabel 2 jika dilihat dari rata-rata *inter cluster* maka *Fuzzy C-Means* lebih baik dibandingkan *K-Means*. Pada Tabel 3 menunjukkan perbandingan *intra cluster distance* antara *Fuzzy C-Means* dan *K-Means*. Dari Tabel 3 jika dilihat dari rata-rata *intra cluster* maka *Fuzzy C-Means* lebih baik dibandingkan dengan *K-Means*.



Gambar 1. Perbandingan FCM dan K-Means

Tabel 2. Evaluasi *inter cluster* FCM dan K-Means

	FCM	K-Means
<i>Inter cluster</i> terkecil	1,0045	1,0497
<i>Inter cluster</i> terbesar	1,0045	1,0497
Rata-rata <i>inter cluster</i>	1,0045	1,0497

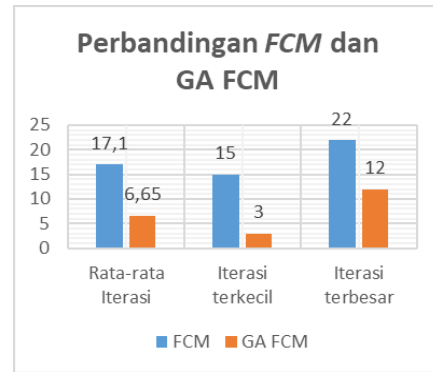
Tabel 3. Evaluasi *intra cluster* FCM dan K-Means

	FCM		K-Means	
	1	2	1	2
<i>Intra cluster</i> terkecil	2,0807	2,1073	0,2518	0,4718
<i>Intra cluster</i> terbesar	2,1605	2,1605	1,2368	1,085
Rata-rata <i>intra cluster</i>	2,0807	2,1073	0,7361	0,7640

Pada Tabel 4. terdapat perbandingan iterasi antara *Fuzzy C-Means* dan Algoritma Genetika - *Fuzzy C-Means*. Dari Tabel 4 algoritma yang memiliki iterasi terkecil minimum yaitu Algoritma Genetika - *Fuzzy C-Means* dengan nilai 3. Begitu juga, algoritma yang memiliki iterasi terbesar minimum yaitu Algoritma Genetika - *Fuzzy C-Means* dengan nilai 12. Secara rata-rata algoritma yang memiliki iterasi minimum yaitu Algoritma Genetika-*Fuzzy C-Means* dengan nilai 6,65. Dapat diketahui bahwa algoritma yang memiliki iterasi minimum yaitu Algoritma Genetika-*Fuzzy C-Means*. Gambar 2 menunjukkan perbandingan *Fuzzy C-Means* dan Algoritma Genetika-*Fuzzy C-Means* dalam bentuk Histogram.

Tabel 4. Perbandingan Iterasi FCM dan GA-FCM

	FCM	GA-FCM
Iterasi terkecil	15	3
Iterasi terbesar	22	12
Rata-rata iterasi	17,1	6,65



Gambar 2. Perbandingan FCM dan GA-FCM

Pada Tabel 5 terdapat perbandingan *inter cluster distance* antara *Fuzzy C-Means* dan Algoritma Genetika - *Fuzzy C-Means*. Dari Tabel 5 jika dilihat dari rata-rata *inter cluster* maka Algoritma Genetika-*Fuzzy C-Means* lebih baik dibandingkan *Fuzzy C-Means*. Pada Tabel 6 menunjukkan perbandingan *intra cluster distance* antara *Fuzzy C-Means* dan Algoritma Genetika - *Fuzzy C-Means*. Dari Tabel 6 jika dilihat dari rata-rata *intra cluster* maka Algoritma Genetika - *Fuzzy C-Means* lebih baik dibandingkan dengan *Fuzzy C-Means*.

Tabel 5. Evaluasi *inter cluster* FCM dan GA-FCM

	FCM	GA-FCM
<i>Inter cluster</i> terkecil	1,0045	0,1618
<i>Inter cluster</i> terbesar	1,0045	0,3021
Rata-rata <i>inter cluster</i>	1,0045	0,2369

Tabel 6. Evaluasi *intra cluster* FCM dan GA-FCM

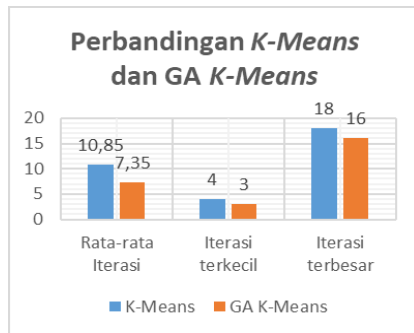
	FCM		GA-FCM	
	1	2	1	2
<i>Intra cluster</i> terkecil	2,0807	2,1073	1,8964	1,9338
<i>Intra cluster</i> terbesar	2,1605	2,1605	2,2697	2,3676
Rata-rata <i>intra cluster</i>	2,0807	2,1073	2,1389	2,1753

Pada Tabel 7. terdapat perbandingan iterasi antara *K-Means* dan Algoritma Genetika - *K-Means*. Dari Tabel 7 algoritma yang memiliki iterasi terkecil minimum yaitu Algoritma Genetika - *K-Means* dengan nilai 3. Begitu juga algoritma yang memiliki iterasi terbesar minimum yaitu Algoritma Genetika - *K-Means* dengan nilai 16. Secara rata-rata algoritma yang memiliki iterasi minimum yaitu Algoritma Genetika - *K-Means* dengan nilai 7,35. Dapat diketahui bahwa algoritma yang memiliki iterasi minimum yaitu Algoritma Genetika - *K-Means*. Gambar 3 menunjukkan perbandingan *K-Means* dan

Algoritma Genetika - *K-Means* dalam bentuk Histogram.

Tabel 7. Perbandingan Iterasi *K-Means* dan *GA-K-Means*

	<i>K-Means</i>	<i>GA K-Means</i>
Iterasi terkecil	4	3
Iterasi terbesar	18	16
Rata-rata iterasi	10,85	7,35



Gambar 3. Perbandingan *K-Means* dan *GA-K-Means*

Pada Tabel 8 terdapat perbandingan *inter cluster distance* antara *K-Means* dan Algoritma Genetika - *K-Means*. Dari Tabel 8 jika dilihat dari rata-rata *inter cluster* maka Algoritma Genetika - *K-Means* sama baiknya dibandingkan *K-Means*. Pada Tabel 9 menunjukkan perbandingan *intra cluster distance* antara *K-Means* dan Algoritma Genetika - *K-Means*. Dari Tabel 9 jika dilihat dari rata-rata *intra cluster* maka Algoritma Genetika - *K-Means* lebih baik dibandingkan dengan *K-Means*.

Tabel 8. Evaluasi *inter cluster K-Means* dan *GA-K-Means*

	<i>K-Means</i>	<i>GA-K-Means</i>
<i>Inter cluster</i> terkecil	1,0497	1,0497
<i>Inter cluster</i> terbesar	1,0497	1,0497
Rata-rata <i>inter cluster</i>	1,0497	1,0497

Tabel 9. Evaluasi *intra cluster K-Means* dan *GA-K-Means*

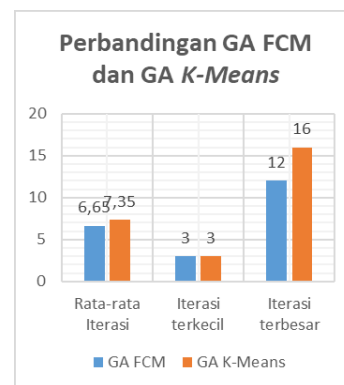
	<i>K-Means</i>		<i>GA-K-Means</i>	
	1	2	1	2
<i>Intra cluster</i> terkecil	0,2518	0,4718	1,7868	1,714
<i>Intra cluster</i> terbesar	1,2368	1,085	2,3681	2,0415
Rata-rata <i>intra cluster</i>	0,7361	0,7640	1,9993	1,9053

Pada Tabel 10. terdapat perbandingan iterasi antara Algoritma Genetika - *Fuzzy C-Means* dan Algoritma Genetika - *K-Means*. Dari Tabel 10 Algoritma Genetika - *Fuzzy C-Means* dan Algoritma Genetika - *K-Means* memiliki iterasi terkecil minimum yang sama yaitu 3. Algoritma Genetika-*Fuzzy C-Means* memiliki iterasi terbesar minimum

dengan nilai 12. Secara rata-rata iterasi minimum dimiliki oleh Algoritma Genetika - *Fuzzy C-Means* yaitu 6,65. Dapat diketahui bahwa antara Algoritma Genetika - *Fuzzy C-Means* lebih baik dibandingkan Algoritma Genetika - *K-Means*. Gambar 4 menunjukkan perbandingan Algoritma Genetika - *Fuzzy C-Means* dan Algoritma Genetika - *K-Means* dalam bentuk Histogram.

Tabel 10. Perbandingan Iterasi *GA-FCM* dan *GA-K-Means*

	<i>GA-FCM</i>	<i>GA-K-Means</i>
Iterasi terkecil	3	3
Iterasi terbesar	12	16
Rata-rata iterasi	6,65	7,35



Gambar 4. Perbandingan *GA-FCM* dan *GA K-Means*

Pada Tabel 11 terdapat perbandingan *inter cluster distance* antara Algoritma Genetika - *Fuzzy C-Means* dan Algoritma Genetika - *K-Means*. Dari Tabel 11 jika dilihat dari rata-rata *inter cluster* maka Algoritma Genetika - *Fuzzy C-Means* lebih baik dibandingkan Algoritma Genetika - *K-Means*. Pada Tabel 12 menunjukkan perbandingan *intra cluster distance* antara Algoritma Genetika - *Fuzzy C-Means* dan Algoritma Genetika - *K-Means*. Dari Tabel 12 jika dilihat dari rata-rata *intra cluster* maka Algoritma Genetika - *Fuzzy C-Means* lebih baik dibandingkan dengan Algoritma Genetika - *K-Means*.

Tabel 11. Evaluasi *inter cluster GA-FCM* dan *GA-K-Means*

	<i>GA-FCM</i>	<i>GA-K-Means</i>
<i>Inter cluster</i> terkecil	0,1618	1,0497
<i>Inter cluster</i> terbesar	0,3021	1,0497
Rata-rata <i>inter cluster</i>	0,2369	1,0497

Tabel 12. Evaluasi *intra cluster GA-FCM* dan *GA-K-Means*

	<i>GA-FCM</i>		<i>GA-K-Means</i>	
	1	2	1	2
<i>Intra cluster</i> terkecil	1,8964	1,9338	1,7868	1,714
<i>Intra cluster</i> terbesar	2,2697	2,3676	2,3681	2,0415
Rata-rata <i>intra cluster</i>	2,1389	2,1753	1,9993	1,9053

Untuk menentukan algoritma yang memiliki keunggulan berupa iterasi minimum dapat dilihat pada Tabel 13.

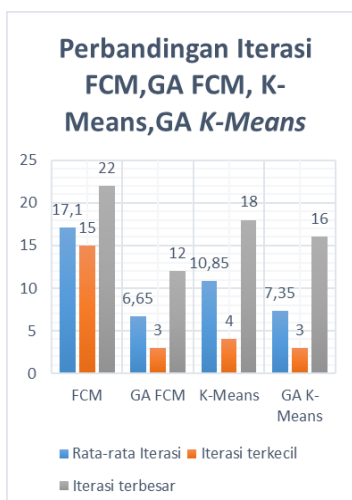
Tabel 13. Perbandingan Iterasi FCM, GA-FCM, K-Means dan GA-K-Means

	FCM	GA-FCM	K-Means	GA-K-Means
Iterasi terkecil	15	3	4	3
Iterasi terbesar	22	12	18	16
Rata-rata iterasi	17,1	6,65	10,85	7,35

Tabel 14. Evaluasi *inter cluster* FCM, GA-FCM, K-Means dan GA-K-Means

	FCM	GA-FCM	K-Means	GA-K-Means
<i>Inter cluster</i> terkecil	1,0045	0,1618	1,0497	1,0497
<i>Inter cluster</i> terbesar	1,0045	0,3021	1,0497	1,0497
Rata-rata <i>inter cluster</i>	1,0045	0,2369	1,0497	1,0497

Untuk menentukan algoritma yang memiliki keunggulan berupa iterasi minimum dapat dilihat pada Tabel 13. Dari Tabel 13 dapat diketahui bahwa secara keseluruhan algoritma yang memiliki iterasi yang terbaik yaitu Algoritma Genetika - *Fuzzy C-Means*. Namun jika dibandingkan antara *Fuzzy C-Means* dan *K-Means* yang memiliki iterasi terbaik adalah *K-Means*. Algoritma Genetika ketika digunakan untuk optimasi memberikan pengaruh yang besar pada Algoritma Genetika - *Fuzzy C-Means* dibandingkan dengan Algoritma Genetika - *K-Means*.



Gambar 5. Perbandingan Secara Keseluruhan

Pengaruh tersebut terlihat pada jumlah iterasi rata-rata yang menurun dari 17,1 menjadi 6,65 pada Algoritma Genetika - *Fuzzy C-Means* dibandingkan dengan Algoritma Genetika - *K-Means* dari 10,85

menjadi 7,35. Penurunan iterasi Gambar 5 menunjukan perbandingan iterasi secara keseluruhan.

Berdasarkan perbandingan iterasi, Algoritma Genetika - *Fuzzy C-Means* merupakan algoritma yang paling baik dengan jumlah rata-rata iterasi 6,65. Berdasarkan Tabel 14, dengan melakukan evaluasi terhadap hasil *clustering* menggunakan *inter cluster distance* terlihat bahwa Algoritma Genetika - *Fuzzy C-Means* memiliki *inter cluster distance* yang kecil dibandingkan algoritma yang lain. Selain *inter cluster distance*, evaluasi dilakukan menggunakan *intra cluster distance*. Perbandingan *intra cluster distance* dapat dilihat pada Tabel 12. Algoritma Genetika - *Fuzzy C-Means* memiliki *intra cluster distance* yang lebih besar dibandingkan dengan algoritma yang lain.

5. KESIMPULAN

Algoritma Genetika dapat diterapkan untuk mengoptimasi algoritma *clustering Fuzzy C-Means* dan *K-Means*. Algoritma genetika berperan sebagai pengganti peran untuk mencari pusat *cluster*. Penggunaan Algoritma Genetika memberikan pengaruh yang signifikan pada *Fuzzy C-Means*. Iterasi rata-rata *Fuzzy C-Means* sebelum optimasi adalah 17,1 kemudian setelah dioptimasi maka menjadi 6,65. Penurunan iterasi sebanyak 61,11%. Iterasi rata-rata *K-Means* sebelum optimasi adalah 10,85, kemudian setelah dioptimasi menjadi 7,35. Penurunan iterasi sebanyak 32,25%. Optimasi pada *Fuzzy C-Means* lebih besar dibandingkan dengan optimasi pada *K-Means*. Dengan melakukan evaluasi terhadap hasil *clustering*, *inter cluster distance* pada Algoritma Genetika-*Fuzzy C-Means* memiliki angka yang paling kecil sedangkan *intra cluster distance* merupakan yang paling besar.

DAFTAR PUSTAKA

BAHRAMI, M. & SAJEDI, H. 2018. Prediction of diabetic retinopathy based on a committee of random forests. *International Journal of Intelligent Machines and Robotics*, 1(2), pp. 133–139. doi: 10.1504/IJIMR.2018.094910.

BALINT, A. & ANDRAS, H. 2014. *Diabetic Retinopathy Debrecen Data Set Data Set*. Available at: <https://archive.ics.uci.edu/ml/datasets/Diabetic+Retinopathy+Debrecen+Data+Set> (Accessed: 15 October 2018).

BEZDEK, J. C. 1981. *Pattern Recognition with Fuzzy Objective Function Algorithm*. New York: Plenum Press.

DEMBELE, D. & KASTNER, P. 2003. Fuzzy C-means method for clustering microarray data', *Bioinformatics*, 19(8), pp. 973–980. doi: 10.1093/bioinformatics/btg119.

HADI, R., GEDE DARMA PUTRA, I. K. and

- SATYA KUMARA, I. N. 2017. Penentuan Kompetensi Mahasiswa dengan Algoritma Genetik dan Metode Fuzzy C-Means. *Majalah Ilmiah Teknologi Elektro*, 15(2), pp. 101–106. doi: 10.24843/mite.1502.15.
- HERYAWAN, L. 2017. Deteksi Dini Retinopati Diabetik dengan Pengolahan Citra Berbasis Morfologi Matematika. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 11(2), p. 209. doi: 10.22146/ijccs.24761.
- JAIN, A. K. 2008. Data Clustering: 50 Years Beyond K-Means. *19th International Conference on Pattern Recognition (ICPR)*, pp. 2–33.
- JAIN, A. K. & DUBES, R. C. 1988. Algorithms for Clustering Data’.
- LIU, S. H. & HOU, H. F. 2009. A combination of mixture genetic algorithm and fuzzy C-means clustering algorithm. *ITME2009 - Proceedings 2009 IEEE International Symposium on IT in Medicine and Education*, pp. 254–258. doi: 10.1109/ITIME.2009.5236422.
- MALKI, A. AL, dkk. 2016. Hybrid Genetic Algorithm with K-Means for Clustering Problems. *Open Journal of Optimization*, 05(02), pp. 71–83. doi: 10.4236/ojop.2016.52009.
- PRAYITNO, E. 2018. Identifikasi Penyakit Retinopati Diabetika Menggunakan Algoritma Kernel K-Means. *MATHunesa Jurnal Ilmiah Matematika*, 6(2), pp. 60–65.
- SINGH, K. K. dkk. 2013. Unsupervised change detection from remote sensing images using hybrid genetic FCM. *2013 Students Conference on Engineering and Systems (SCES)*, pp. 1–5. doi: 10.1109/SCES.2013.6547529.
- SOLER, J. dkk. 2013. Data Clustering and Similarity. *Proceedings of the Twenty-Sixth International Florida Artificial Intelligence Research Society Conference*, pp. 492–495. Available at: <https://hal.archives-ouvertes.fr/hal-00828152>.
- TJIN, W. 2017. *Retinopati Diabetik*. Available at: <https://www.alodokter.com/retinopati-diabetik> (Accessed: 10 October 2018).
- YOHANNES, Y. 2016. Analisis Perbandingan Algoritma Fuzzy C-Means dan K-Means. *Annual Research Seminar 2016*, 2(1), pp. 151–155.
- YONG-GUO LIU, KE-FEI CHEN & XUE-MING LI. 2005. A hybrid genetic based clustering algorithm. *Proceedings of the Third International Conference on Machine Learning and Cybernetics*, (August), pp. 1677–1682. doi:

10.1109/icmlc.2004.1382045.

Halaman ini sengaja dikosongkan