

Perbandingan Metode K-Means dan Fuzzy C-Means Untuk Pengelompokan Pegawai Berdasarkan Nilai Kinerja dan Tingkat Kedisiplinan Pegawai

Wikarno, Rheo Malani, Bedi Suprpty

Teknologi Informasi
Politeknik Negeri Samarinda
Samarinda, Indonesia

arnoseven93@gmail.com, rheo@polnes.ac.id, bedirheody@gmail.com

Abstract—Bank Indonesia (BI) adalah bank sentral Republik Indonesia. Sebagai bank sentral, BI membutuhkan pegawai-pegawai dengan kompetensi terbaik dalam bidangnya masing-masing serta perilaku yang baik. Untuk mengetahui kualifikasi setiap pegawai dibutuhkan sebuah pengelompokan/clustering. Melalui proses penilaian kinerja serta tingkat kedisiplinan pegawai dilakukan proses pengelompokan agar dapat diketahui cluster-cluster yang terdapat didalamnya. Pembagian cluster pada penelitian ini mencirikan kualifikasi pegawai, dimana cluster 1 berisi pegawai dengan kualifikasi terbaik, cluster 2 berisi pegawai dengan kualifikasi baik, cluster 3 berisi pegawai dengan kualifikasi cukup, cluster 4 berisi pegawai dengan kualifikasi kurang dan cluster 5 berisi pegawai dengan kualifikasi Buruk. Rata-rata Persentase MAPE untuk keseluruhan cluster pada metode K-means lebih kecil dibandingkan dengan metode Fuzzy C-means. Persentase MAPE pada metode K-means sebesar 18,21% sementara pada metode Fuzzy C-means sebesar 21,07%. Rata-rata varian jarak antar anggota pada masing-masing cluster untuk keseluruhan cluster pada metode K-means lebih kecil dibandingkan dengan metode Fuzzy C-means. Rata-rata varian jarak pada metode K-means sebesar 0,350 sementara pada metode Fuzzy C-means sebesar 0,863. Rata-rata jarak antar cluster pada keseluruhan jarak antar tiap-tiap cluster pada metode K-means lebih besar dibandingkan dengan metode Fuzzy C-means. Rata-rata jarak antar cluster pada metode K-means sebesar 9,227 sementara pada metode Fuzzy C-means sebesar 6,465.

Keywords—Penilaian Kinerja, MAPE, Clustering, K-means, Fuzzy C-means

I. PENDAHULUAN

Bank Indonesia (BI) adalah bank sentral Republik Indonesia. Sebagai bank sentral, BI mempunyai satu tujuan tunggal, yaitu mencapai dan memelihara kestabilan nilai rupiah. Dalam hal mencapai tujuan tersebut, BI membutuhkan pegawai-pegawai dengan kompetensi terbaik dalam bidangnya masing-masing serta perilaku yang baik. BI memiliki dua komponen dalam penilaian kinerja pegawai, yaitu Prestasi Kerja dan Perilaku Kerja. Untuk mengetahui kualifikasi setiap pegawai dibutuhkan sebuah pengelompokan/clustering, dengan pengelompokan tersebut

akan dihasilkan pegawai dengan kualifikasi terbaik hingga buruk.

Clustering adalah proses membagi data ke dalam kelompok-kelompok yang mempunyai obyek yang karakteristiknya sama. *Clustering* adalah mengelompokkan item data ke dalam sejumlah kecil grup sedemikian sehingga masing-masing grup mempunyai sesuatu persamaan yang esensial [1]. Ada berbagai macam metode *clustering*, antara lain adalah K-means dan Fuzzy C-means.

K-means *clustering* merupakan salah satu metode data *clustering* non-hirarki yang mengelompokkan data dalam bentuk satu atau lebih *cluster*/kelompok. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu *cluster*/kelompok dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan *cluster*/kelompok yang lain sehingga data yang berada dalam satu *cluster*/kelompok memiliki tingkat variasi yang kecil [2,3,4].

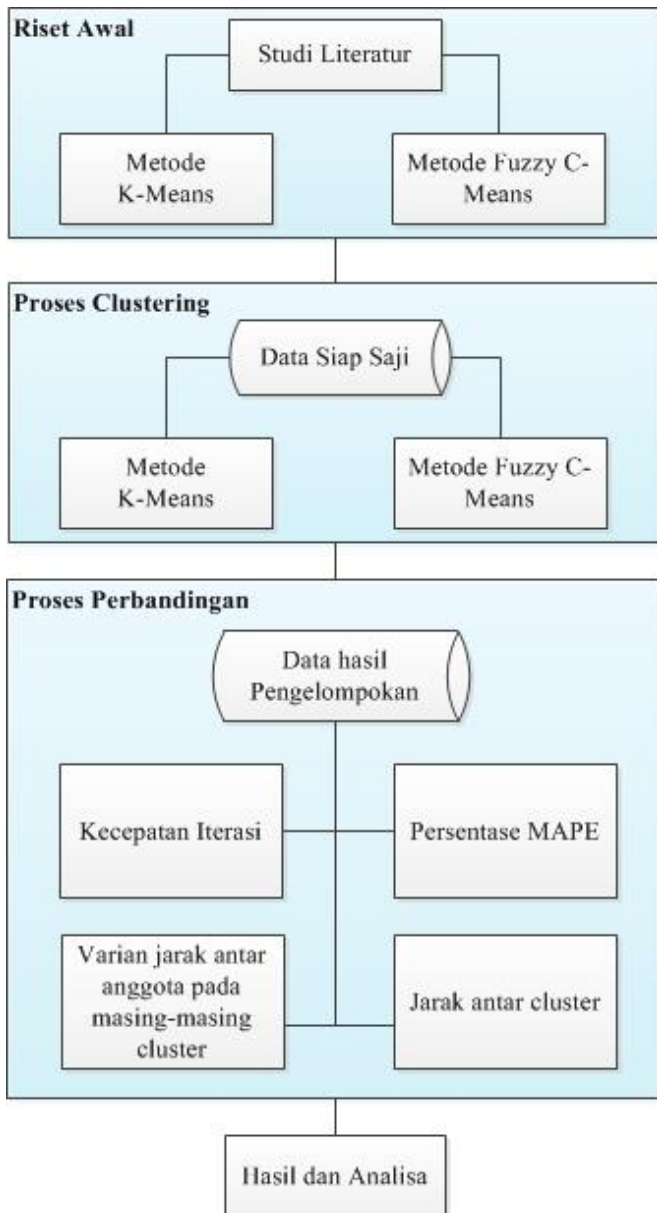
Fuzzy C-means (FCM) adalah salah satu teknik pengelompokan data yang mana keberadaan tiap-tiap titik data dalam suatu *cluster* ditentukan oleh derajat keanggotaan. Teknik ini pertama kali diperkenalkan oleh Jim Bezdek pada tahun 1981 [5]. Konsep dasar FCM pertama kali adalah menentukan pusat *cluster* yang akan menandai lokasi rata-rata untuk tiap-tiap *cluster*. Pada kondisi awal, pusat *cluster* ini masih belum akurat. Tiap-tiap data memiliki derajat keanggotaan untuk tiap-tiap *cluster*. Dengan cara memperbaiki pusat *cluster* dan nilai keanggotaan tiap-tiap data secara berulang, maka dapat dilihat bahwa pusat *cluster* akan menuju lokasi yang tepat. Perulangan ini didasarkan pada minimasi fungsi obyektif yang menggambarkan jarak dari titik data yang diberikan ke pusat *cluster* yang terbobot [6].

Pada penelitian ini pengelompokan pegawai akan dilakukan menggunakan metode K-means dan Fuzzy C-means. Dimana Pembagian *cluster* pada penelitian ini mencirikan kualifikasi pegawai, dimana *cluster* 1 berisi pegawai dengan kualifikasi terbaik, *cluster* 2 berisi pegawai dengan kualifikasi baik, *cluster* 3 berisi pegawai dengan kualifikasi cukup, *cluster* 4 berisi pegawai dengan kualifikasi kurang dan *cluster* 5 berisi pegawai dengan kualifikasi Buruk.

Setelah hasil pengelompokan didapatkan, selanjutnya dilakukan proses uji perbandingan antara hasil *clustering* menggunakan metode K-means dan hasil *clustering* menggunakan metode Fuzzy C-means. Beberapa hal yang akan dibandingkan antara lain adalah, kecepatan iterasi, nilai MAPE, varian jarak antar anggota pada setiap *cluster* dan jarak antar *cluster*.

II. METODOLOGI

A. Metode Penelitian



Gambar. 1. Metodologi Penelitian

Terdapat beberapa tahapan dalam metode penelitian yang perlu dilakukan diantaranya.

1) Riset Awal

Riset awal diperlukan sebagai dasar penelitian dalam hal mempelajari segala hal yang terkait dengan topik penelitian. Pada tahapan riset awal peneliti perlu melakukan studi literatur berbagai teori yang sudah dikumpulkan dari beberapa sumber seperti buku, jurnal dan internet yang diperlukan untuk menunjang penelitian ini.

2) Proses Clustering

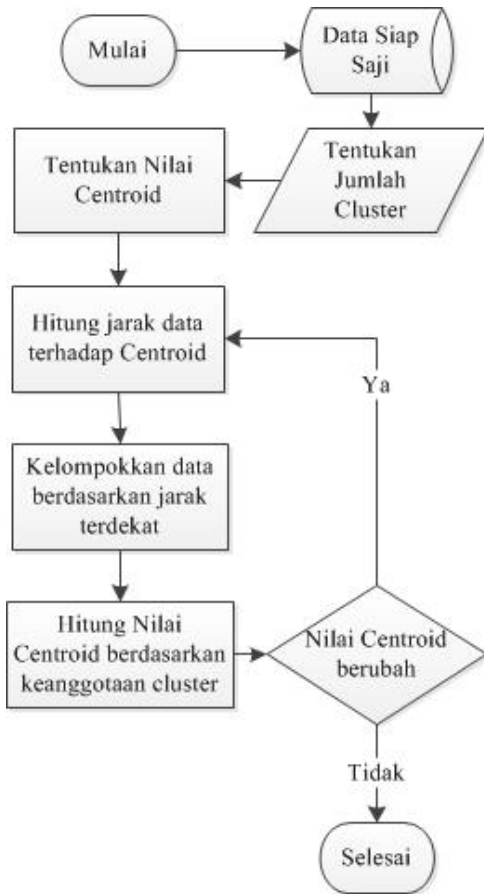
Pada tahapan ini data siap saji dikelompokkan menggunakan metode K-means dan Fuzzy C-means. Setelah data dikelompokkan, data hasil peengelompokkan akan diproses pada tahapan selanjutnya.

3) Proses Perbandingan

Pada tahapan ini data hasil pengelompokan akan dibandingkan. Perbandingan yang akan dilakukan adalah membandingkan kecepatan iterasi, persentase MAPE, varian jarak antar anggota pada masing-masing *cluster* dan jarak antar *cluster*.

B. Metode K-means

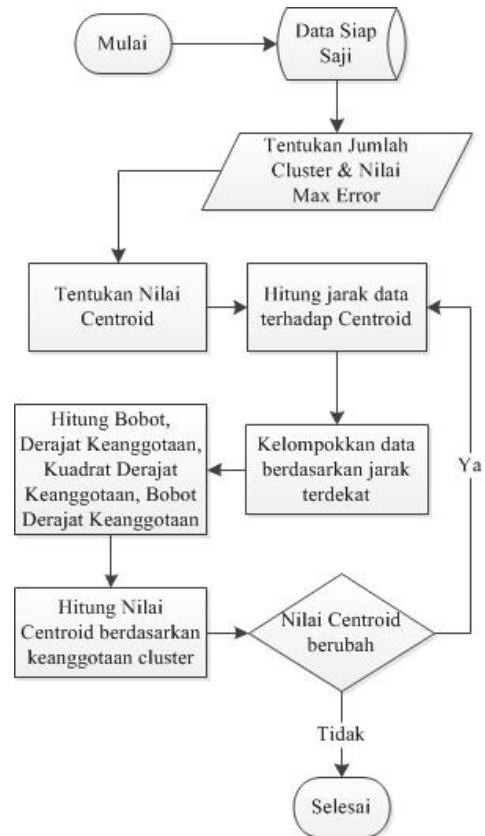
Metode K-means merupakan metode yang membutuhkan parameter *input* sebanyak k dan membagi sekumpulan n objek kedalam k *cluster* sehingga tingkat kemiripan antar anggota dalam satu *cluster* tinggi sedangkan tingkat kemiripan dengan anggota pada *cluster* lain sangat rendah. Kemiripan anggota terhadap *cluster* diukur dengan kedekatan objek terhadap nilai *mean* K-means hanya melihat jarak data ke masing-masing *centroid* pada setiap *cluster*. Berikut adalah alur proses *clustering* menggunakan metode K-means:



Gambar. 2. Diagram Alir Metode K-means

C. Metode Fuzzy C-Means

Metode Fuzzy C-means adalah suatu metode pengklasteran data yang ditentukan oleh derajat keanggotaan. Berikut adalah alur proses *clustering* menggunakan metode Fuzzy C-means:



Gambar. 3. Diagram Alir Metode Fuzzy C-Means

D. Euclidean Distance

Euclidean Distance merupakan salah satu teknik pengukuran kemiripan atau jarak. Ukuran ini sering dipergunakan dalam *clustering* karena sederhana walaupun sangat sensitif terhadap penciliran. Ukuran ini memiliki masalah jika skala nilai atribut yang satu sangat besar dibandingkan nilai atribut lainnya. Rumus perhitungan *Euclidean distance* dapat dilihat pada persamaan 1.

$$d^2_{rs} = (\chi_r - \chi_s)(\chi_r - \chi_s) \tag{1}$$

E. Mean Absolute Percentage Error (MAPE)

MAPE merupakan rata-rata dari keseluruhan persentase kesalahan (selisih) antara data aktual dengan data hasil peramalan. Ukuran akurasi dicocokkan dengan data *time series*, dan ditunjukkan dalam persentase. Proses perhitungan MAPE ditunjukkan pada persamaan 2.

$$MAPE = \frac{\sum \frac{X_t - F_t}{X_t} \times 100}{n} \tag{2}$$

Xt = Data aktual
 Ft = Hasil peramalan
 n = banyaknya data

III. HASIL & PEMBAHASAN

A. Data Pegawai

Data pegawai dapat dilihat pada tabel 1.

TABLE I. TABEL DATA PEGAWAI

Nama	Atribut 1	Atribut 2	Atribut 3	Atribut 4
A1	3,000	3,500	21,000	1,000
A2	3,500	3,500	17,000	1,000
A3	3,750	4,000	23,000	2,000
A4	3,250	3,250	22,000	1,000
A5	2,750	3,000	23,000	3,000
A6	3,750	4,000	25,000	2,000
A7	3,250	3,500	25,000	1,000
B1	3,500	4,000	24,000	1,000
B2	3,000	3,250	22,000	4,000
B3	2,750	3,000	10,000	1,000
B4	3,250	3,250	20,000	2,000
B5	3,750	3,750	23,000	2,000
B6	3,000	3,000	17,000	4,000
B7	3,250	3,250	19,000	1,000
C1	3,750	3,750	25,000	2,000
C2	3,500	3,500	21,000	3,000
C3	3,000	3,000	20,000	1,000
C4	3,000	3,000	9,000	4,000
C5	3,000	3,000	19,000	2,000
C6	3,000	3,000	10,000	5,000
C7	3,250	3,250	25,000	3,000
D1	2,000	2,250	8,000	6,000
D2	2,000	2,500	5,000	5,000
D3	4,000	4,000	23,000	1,000
D4	2,750	3,000	21,000	2,000
D5	3,000	3,000	15,000	1,000
D6	3,250	3,500	25,000	1,000
D7	3,000	3,000	18,000	4,000
E1	3,000	3,000	21,000	5,000
E2	3,250	3,250	23,000	1,000
E3	3,000	3,500	21,000	1,000
E4	3,250	3,250	17,000	2,000
E5	3,500	3,500	22,000	4,000
E6	3,000	3,250	16,000	3,000
E7	3,000	3,500	18,000	2,000

Keterangan :

Atribut 1 : Nilai Kinerja

Atribut 2 : Predikat Kerja

Atribut 3 : Persentase Jumlah Kelebihan Jam Kerja

Atribut 4 : Jumlah Lupa Absen

Nilai kinerja dan predikat kerja pegawai memiliki rentang/batasan nilai dari 2,000 hingga 5,000. Sementara untuk jumlah kelebihan jam kerja dan jumlah lupa absen tidak memiliki batasan nilai. Pada atribut 1, 2, dan 3, semakin besar nilainya semakin baik. Sementara pada atribut 4, semakin besar nilainya semakin buruk.

B. Penentuan Nilai Centroid Awal

Penentuan nilai centroid awal dalam penelitian ini menggunakan teknik pengambilan acak secara sistematis. Dalam penelitian ini banyaknya jumlah cluster adalah 5

cluster. Tahapan pertama adalah menentukan rentang nilai antar kelas dengan rumus sebagai berikut:

$$\text{Rentang Nilai} = \text{Nilai Max} - \text{Nilai Min} \quad (3)$$

Tahapan selanjutnya adalah menghitung panjang interval kelas dengan rumus sebagai berikut:

$$\text{Panjang Interval Kelas} = \frac{\text{Rentang Nilai}}{\text{Banyaknya Kelas}} \quad (4)$$

Contoh perhitungan pada atribut 1 :

Diketahui :

Nilai Max = 4,000

Nilai Min = 2,000

Banyaknya kelas = 5

Maka :

Rentang Nilai = 4,000 - 2,000 = 2,000

Panjang Interval kelas = 2,000 / 5 = 0,400

Hasil perhitungan untuk setiap atribut dapat dilihat pada tabel 2.

TABLE II. TABEL PANJANG INTERVAL SETIAP KELAS

Atribut 1	Rentang Nilai Antar Kelas	2,000
	Banyak Kelas	5,000
	Panjang Interval Kelas	0,400
Atribut 2	Rentang Nilai Antar Kelas	1,750
	Banyak Kelas	5,000
	Panjang Interval Kelas	0,350
Atribut 3	Rentang Nilai Antar Kelas	20,000
	Banyak Kelas	5,000
	Panjang Interval Kelas	4,000
Atribut 4	Rentang Nilai Antar Kelas	5,000
	Banyak Kelas	5,000
	Panjang Interval Kelas	1,000

Setelah nilai panjang interval kelas didapat, selanjutnya diambil nilai tengah setiap kelas disetiap atribut untuk dijadikan nilai *centroid* awal.

Contoh pengambilan nilai tengah pada atribut 1:

Diketahui :

Nilai Max = 4,000

Nilai Min = 2,000

Panjang Interval Kelas = 0,400

Maka :

Kelas 1 = 2,000 s/d 2,400

Kelas 2 = 2,400 s/d 2,800

Kelas 3 = 2,800 s/d 3,200

Kelas 4 = 3,200 s/d 3,600

Kelas 5 = 3,600 s/d 4,000

Sehingga :

Nilai tengah pada kelas 1 adalah 2,200 dan nilai tersebut menjadi nilai *centroid* untuk cluster Buruk pada atribut 1. Keseluruhan nilai *centroid* awal dapat dilihat pada tabel 3.

TABLE III. NILAI CENTROID AWAL

Cluster	Atribut 1	Atribut 2	Atribut 3	Atribut 4
---------	-----------	-----------	-----------	-----------

Terbaik	3,800	3,825	23,000	1,500
Baik	3,400	3,475	19,000	2,500
Cukup	3,000	3,125	15,000	3,500
Kurang	2,600	2,775	11,000	4,500
Buruk	2,200	2,425	7,000	5,500

C. Clustering Dengan Metode K-Means

1) Menentukan Jumlah Cluster dan Nilai Centroid Awal

Jumlah cluster dan nilai centroid awal yang digunakan dapat dilihat pada tabel 3.

2) Menghitung Jarak Data dengan Centroid

Untuk menghitung jarak antara data dengan centroid, pada penelitian ini menggunakan rumus *euclidean distance*. Hasil perhitungan jarak tersebut ditunjukkan pada tabel 4.

TABLE IV. HASIL PERHITUNGAN JARAK ITERASI KE-1

Nama	ED Cluster Terbaik	ED Cluster Baik	ED Cluster Cukup	ED Cluster Kurang	ED Cluster Buruk
A1	2,235	2,532	6,511	10,627	14,766
A2	6,037	2,502	3,262	7,042	11,095
A3	0,532	4,080	8,221	12,372	16,527
A4	1,372	3,365	7,438	11,571	15,717
A5	2,008	4,111	8,020	12,096	16,214
:	:	:	:	:	:
E7	5,099	1,188	3,375	7,479	11,621

Pada iterasi ke-2 data A1 memiliki jarak terdekat ke cluster Baik, sehingga A1 pindah menjadi anggota cluster Baik. Data A2 memiliki jarak terdekat ke cluster Cukup, sehingga A2 pindah menjadi anggota cluster Cukup. Data E3 memiliki jarak terdekat ke cluster Baik, sehingga E3 pindah menjadi anggota cluster Baik. Data E4 memiliki jarak terdekat ke cluster Cukup, sehingga E4 pindah menjadi anggota cluster Cukup.

Pada iterasi ke-3 data D7 memiliki jarak terdekat ke cluster Cukup, sehingga D7 pindah menjadi anggota cluster Cukup. Data E7 memiliki jarak terdekat ke cluster Cukup, sehingga E7 pindah menjadi anggota cluster Cukup.

Pada iterasi ke-4, iterasi dihentikan dikarenakan semua data sudah memiliki jarak terdekat ke cluster yang sama dengan iterasi ke-3 sehingga tidak ada lagi data yang berpindah cluster.

3) Pengelompokan Data

Setelah mendapat hasil perhitungan jarak masing-masing data terhadap ke 5 cluster. Langkah selanjutnya adalah mengelompokkan data dengan cara memilih jarak terdekat terhadap cluster.

Pada penelitian ini, clustering menggunakan metode K-means berhenti pada iterasi ke-4 dikarenakan nilai centroid tidak mengalami perubahan. Hasil pengelompokan data dapat dilihat pada tabel 5.

TABLE V. HASIL PENGELOMPOKAN DATA ITERASI KE-4

Cluster	Nilai Centroid	Anggota	Jumlah
Terbaik	3,429	A3, A4, A5, A6, A7, B1, B2, B5, C1, C7, D3, D6, E2, E5	14
	3,571		
	23,571		
	2,000		
Baik	3,083	A1, B4, B7, C2, C3, C5, D4, E1, E3	9
	3,222		
	20,333		
	2,000		
Cukup	3,107	A2, B6, D5, D7, E4, E6, E7	7
	3,214		
	16,857		
	2,429		
Kurang	2,917	B3, C4, C6	3
	3,000		
	9,667		
	3,333		
Buruk	2,000	D1, D2	2
	2,375		
	6,500		
	5,500		

D. Clustering Dengan Fuzzy C-Means

Pada metode Fuzzy C-means, nilai centroid awal yang digunakan sama dengan nilai centroid awal pada metode K-means yang terdapat pada tabel 3. Proses clustering dengan metode Fuzzy C-means hampir serupa dengan metode K-means. Untuk metode Fuzzy C-means tingkat keberadaan data dalam suatu kelas atau cluster ditentukan oleh derajat keanggotaannya. Sehingga data dapat menjadi anggota dari semua kelas atau cluster terbentuk dengan derajat atau tingkat keanggotaan yang berbeda antara 0 hingga 1. Penghentian iterasi dilakukan ketika nilai absolute antara centroid sebelumnya dengan centroid baru lebih kecil dari 10^{-4} .

1) Menghitung Bobot ED

Pembobotan hasil perhitungan jarak menggunakan *euclidean distance* (ED), yaitu dengan cara 1 dibagi dengan hasil perhitungan jarak. Hasil pembobotan ED ditunjukkan pada tabel 6.

TABLE VI. HASIL PEMBOBOTAN EUCLIDEAN DISTANCE

Bobot Euclidean Distance					
V1	0,447	0,166	1,879	0,729	0,498
V2	0,395	0,400	0,245	0,297	0,243
V3	0,154	0,307	0,122	0,134	0,125
V4	0,094	0,142	0,081	0,086	0,083
V5	0,068	0,090	0,061	0,064	0,062
Jml	1,158	1,104	2,387	1,310	1,010
Jumlah seluruh Bobot ED = 44,976					

2) Menghitung Derajat Keanggotaan

Derajat keanggotaan merupakan bobot fuzzy. Nilai derajat keanggotaan dihitung dengan cara membagi bobot *euclidean distance* dengan total bobot untuk masing-masing data. Hasil perhitungan derajat keanggotaan ditunjukkan pada tabel 7.

TABLE VII. HASIL DERAJAT KEANGGOTAAN

Derajat Keanggotaan					
V1	0,386	0,150	0,787	0,556	0,493
V2	0,341	0,362	0,103	0,227	0,241
V3	0,133	0,278	0,051	0,103	0,123
V4	0,081	0,129	0,034	0,066	0,082
V5	0,058	0,082	0,025	0,049	0,061

3) Menghitung Kuadrat Derajat Keanggotaan

Setelah mendapatkan hasil derajat keanggotaan kemudian nilai setiap derajat keanggotaan dikuadratkan. Hasil Kuadrat derajat keanggotaan ditunjukkan pada tabel 8.

TABLE VIII. HASIL KUADRAT DERAJAT KEANGGOTAAN

Kuadrat Derajat Keanggotaan						Jml	
V1	0,149	0,023	0,620	:	:	0,016	5,956
V2	0,116	0,131	0,011	:	:	0,293	3,369
V3	0,018	0,077	0,003	:	:	0,036	1,302
V4	0,007	0,017	0,001	:	:	0,007	0,897
V5	0,003	0,007	0,001	:	:	0,003	0,978

4) Menghitung Bobot Derajat Keanggotaan

Pembentukan *centroid* baru dihitung dengan menggunakan bobot derajat keanggotaan. Vektor 1 pada bobot derajat keanggotaan merupakan *centroid* pertama. Menghitung bobot derajat keanggotaan dengan cara Vektor 1 (V1) kuadrat derajat keanggotaan dikali dengan *feature vector*. Begitu juga dengan V2 sampai dengan V5. Hasil perhitungan bobot derajat keanggotaan dapat dilihat pada tabel 9.

TABLE IX. HASIL BOBOT DERAJAT KEANGGOTAAN VEKTOR 1

Vektor 1					
C1	0,448	0,079	2,324	1,005	0,668
C2	0,523	0,079	2,479	1,005	0,729
C3	3,136	0,383	14,252	6,804	5,588
C4	0,149	0,023	1,239	0,309	0,729

Berdasarkan hasil bobot derajat keanggotaan masing-masing vektor yang terlihat pada tabel 9 selanjutnya dihitung jumlah dari masing-masing *cluster*. Total dari bobot derajat keanggotaan ditunjukkan pada tabel 10.

TABLE X. TOTAL BOBOT DERAJAT KEANGGOTAAN

V1	V2	V3	V4	V5
20,571	10,549	4,001	2,626	2,270
21,463	10,947	4,158	2,708	2,541
137,095	66,366	22,490	10,935	8,586
10,512	7,326	3,308	3,238	4,646

5) Pembangkitan Centroid Baru

Dari hasil jumlah bobot derajat keanggotaan maka dapat ditentukan *centroid* baru yaitu dengan cara membagi jumlah masing-masing vektor pada kuadrat derajat keanggotaan dengan total kuadrat derajat keanggotaan untuk setiap vektor. Sehingga terbentuklah *centroid* baru. *Centroid* baru dapat dilihat pada tabel 11.

TABLE XI. PEMBENTUKAN CENTROID BARU

Cen 1	Cen 2	Cen 3	Cen 4	Cen 5
3,454	3,132	3,073	2,929	2,321
3,604	3,250	3,194	3,021	2,598
23,019	19,701	17,275	12,195	8,779
1,765	2,175	2,541	3,611	4,750

6) Pengelompokan Data

Pada kasus ini, proses *clustering* dihentikan pada iterasi ke 33 dikarenakan nilai *absolute* antara *centroid* pada iterasi ke-32 dengan *centroid* pada iterasi ke-33 telah lebih kecil dari 10^{-4} .

TABLE XII. NILAI CENTROID ITERASI KE-32 DAN KE-33

Iterasi 32					
	K1	K2	K3	K4	K5
Atribut 1	3,478	3,171	3,157	3,118	2,711
Atribut 2	3,624	3,334	3,252	3,239	2,833
Atribut 3	23,104	21,274	20,054	17,305	9,228
Atribut 4	1,852	2,27	2,045	2,316	4,336
Iterasi 33					
Atribut 1	3,478	3,171	3,157	3,118	2,711
Atribut 2	3,624	3,334	3,252	3,239	2,833
Atribut 3	23,104	21,274	20,055	17,305	9,228
Atribut 4	1,852	2,27	2,045	2,316	4,336

Selanjutnya data dikelompokkan sesuai dengan jarak terdekat data pada iterasi ke-33. Hasil pengelompokan data dapat dilihat pada tabel 13.

TABLE XIII. HASIL PENGELOMPOKAN DATA ITERASI KE-33

Cluster	Nilai Centroid	Anggota	Jumlah
Terbaik	3,478	A3, A4, A5, A6, A7, B1, B5, C1, C7, D3, D6, E2	12
	3,624		
	23,104		
	1,852		
Baik	3,171	A1, B2, C2, D4, E1, E3, E5	7
	3,334		
	21,274		
	2,270		
Cukup	3,157	B4, B7, C3, C5	4
	3,252		
	20,055		
	2,045		
Kurang	3,118	A2, B6, D5, D7, E4, E6, E7	7
	3,239		
	17,305		
	2,316		
Buruk	2,711	B3, C4, C6, D1, D2	5
	2,833		
	9,228		
	4,336		

E. Perbandingan Hasil

1) Kecepatan Iterasi

Dengan menggunakan metode K-means proses *clustering* berhenti pada iterasi ke-4 sementara proses *clustering* menggunakan metode Fuzzy C-means berhenti pada iterasi ke-33. Hasil *clustering* dapat dilihat pada tabel 14.

TABLE XIV. NILAI CENTROID DIAKHIR ITERASI

Cluster	K-means		Fuzzy C-means	
	Nilai Centroid	Jumlah Anggota	Nilai Centroid	Jumlah Anggota
Terbaik	3,429	14	3,478	12
	3,571			
	23,571			
	2,000			
Baik	3,083	9	3,171	7
	3,222			
	20,333			
	2,000			
Cukup	3,107	7	3,157	4
	3,214			
	16,857			
	2,429			

Kurang	2,917	3	3,118	7
	3,000		3,239	
	9,667		17,305	
	3,333		2,316	
Buruk	2,000	2	2,711	5
	2,375		2,833	
	6,500		9,228	
	5,500		4,336	

2) *Persentase MAPE*

Perhitungan persentase *error* menggunakan metode MAPE (*Mean Absolute Percentage Error*). Hasil perhitungan MAPE dapat dilihat pada tabel 15.

TABLE XV. TABEL HASIL PERHITUNGAN MAPE

Means Absolute Percentage Error (MAPE)		
Cluster	K-means	Fuzzy C-means
Terbaik	19,05%	18,27%
Baik	17,50%	19,53%
Cukup	19,02%	16,14%
Kurang	25,75%	18,30%
Buruk	9,70%	33,09%
Rata-rata MAPE	18,21%	21,07%

3) *Varian Jarak antar anggota pada masing-masing cluster*

Varian jarak antar anggota pada masing-masing *cluster* didapatkan dengan menghitung nilai varian dari nilai jarak/nilai ED setiap anggota dengan nilai *centroid* masing-masing *cluster*. Hasil perhitungan varian jarak antar anggota pada masing-masing *cluster* dapat dilihat pada tabel 16.

TABLE XVI. VARIAN JARAK ANTAR ANGGOTA PADA MASING-MASING CLUSTER

Varian Jarak antar anggota pada masing – masing cluster		
Cluster	K-means	Fuzzy C-means
Terbaik	0,308	0,409
Baik	0,561	0,516
Cukup	0,377	0,341
Kurang	0,502	0,515
Buruk	0,000	2,534
Rata-rata Varian Jarak	0,350	0,863

Semakin kecil nilai varian jarak antar anggota pada masing-masing *cluster* menandakan bahwa setiap anggota memiliki jarak yang dekat dengan nilai *centroid cluster*.

4) *Jarak antar cluster*

Jarak antar *cluster* didapatkan dengan cara menghitung nilai jarak/nilai ED antara nilai *centroid* antar *cluster*. Hasil perhitungan jarak antar *cluster* dapat dilihat pada tabel 17.

TABLE XVII. JARAK ANTAR CLUSTER

Jarak Antar Cluster				
Antar Cluster			K-means	Fuzzy C-means
Terbaik	ke	Baik	3,275	1,924
Terbaik	ke	Cukup	6,745	3,095
Terbaik	ke	Kurang	13,990	5,841
Terbaik	ke	Buruk	17,526	14,139
Baik	ke	Cukup	3,503	1,243
Baik	ke	Kurang	10,753	3,971
Baik	ke	Buruk	14,335	12,241
Cukup	ke	Kurang	7,253	2,763

Cukup	ke	Buruk	10,892	11,083
Kurang	ke	Buruk	3,994	8,346
Rata-rata Jarak			9,227	6,465

Nilai besar pada jarak antar *cluster* menandakan bahwa jarak antar nilai *centroid* di tiap *cluster* berjauhan. Sehingga kemungkinan adanya anggota *cluster* yang berdekatan dengan anggota pada *cluster* lainnya semakin kecil.

F. *Analisa*

Berdasarkan hasil perbandingan diatas, Proses *clustering* menggunakan metode K-means berhenti pada iterasi ke-4. Dimana *Cluster 1* berisi pegawai dengan kualifikasi terbaik dan jumlah anggota sebanyak 14 orang. *Cluster 2* berisi pegawai dengan kualifikasi baik sebanyak 9 orang. *Cluster 3* berisi pegawai dengan kualifikasi cukup sebanyak 7 orang. *Cluster 4* berisi pegawai dengan kualifikasi kurang sebanyak 3 orang. *Cluster 5* berisi pegawai dengan kualifikasi buruk sebanyak 2 orang. Rata-rata persentase MAPE untuk keseluruhan *cluster* sebesar 18,21%. Rata-rata varian jarak antar anggota pada masing-masing *cluster* untuk keseluruhan *cluster* sebesar 0,350. Rata-rata jarak antar *cluster* pada keseluruhan jarak antar tiap-tiap *cluster* sebesar 9,227.

Sementara proses *clustering* menggunakan metode Fuzzy C-means berhenti pada iterasi ke-33. Dimana *Cluster 1* berisi pegawai dengan kualifikasi terbaik dan jumlah anggota sebanyak 12 orang. *Cluster 2* berisi pegawai dengan kualifikasi baik sebanyak 7 orang. *Cluster 3* berisi pegawai dengan kualifikasi cukup sebanyak 4 orang. *Cluster 4* berisi pegawai dengan kualifikasi kurang sebanyak 7 orang. *Cluster 5* berisi pegawai dengan kualifikasi buruk sebanyak 5 orang. Rata-rata nilai MAPE untuk keseluruhan *cluster* sebesar 21,07%. Rata-rata varian jarak antar anggota pada masing-masing *cluster* untuk keseluruhan *cluster* sebesar 0,863. Rata-rata jarak antar *cluster* pada keseluruhan jarak antar tiap-tiap *cluster* sebesar 6,465.

IV. KESIMPULAN

A. *Kesimpulan*

Berdasarkan hasil penelitian dan pembahasan diatas, dapat diperoleh beberapa kesimpulan antara lain :

- Kecepatan iterasi metode K-means lebih cepat dibandingkan kecepatan iterasi metode Fuzzy C-means.
- Rata-rata Persentase MAPE untuk keseluruhan *cluster* pada metode K-means lebih kecil dibandingkan dengan metode Fuzzy C-means. Persentase MAPE pada metode K-means sebesar 18,21% sementara pada metode Fuzzy C-means sebesar 21,07%.
- Rata-rata varian jarak antar anggota pada masing-masing *cluster* untuk keseluruhan *cluster* pada metode K-means lebih kecil dibandingkan dengan metode Fuzzy C-means. Rata-rata varian jarak pada metode K-means sebesar 0,350 sementara pada metode Fuzzy C-means sebesar 0,863.

- Rata-rata jarak antar *cluster* pada keseluruhan jarak antar tiap-tiap *cluster* pada metode K-means lebih besar dibandingkan dengan metode Fuzzy C-means. Rata-rata jarak antar *cluster* pada metode K-means sebesar 9,227 sementara pada metode Fuzzy C-means sebesar 6,465.

DAFTAR PUSTAKA

- [1] Khotimah, T., "Pengelompokan Surat Dalam Al Qur'an Menggunakan Algoritma K-Means", Jurnal SIMETRIS, ISSN: 2252-4983, 2014.
- [2] Malani, R., Suprpty, B., "Implementasi Sistem Seleksi Kinerja Pegawai Negeri Khususnya Tenaga Pengajar Terbaik Menggunakan Metode K-MEANS Clustering pada Politeknik Negeri Samarinda", Seminar Nasional Teknologi Terapan 2016 Sekolah Vokasi Universitas Gajah Mada, ISBN: 978-602-1159-18-7, 2016.
- [3] Ramadhan, A., Efendi, Z., Mustakim., "Perbandingan K-Means dan Fuzzy C-Means untuk Pengelompokan Data User Knowledge Modeling", Seminar Nasional Teknologi Informasi, Komunikasi dan Industri (SNTIKI) 9, ISSN: 2579-7271, 2017.
- [4] Battiti, R., Brunato, M., "Machine Learning plus Intelligent Optimization", Los Angeles: Lionsolver Inc, 2013.
- [5] Suprpty, B., Malani, R., Nurhayati, O.D., "Design of Information System for Acceptance Selection of Prospective Employees Online Using Tahani Fuzzy Logic Method and Simple Additive Weighting (SAW)", International Journal of Computing and Informatics, Vol. 1, No.1, ISSN: 2502-2334, 2016.
- [6] Nastiti, P.R., Putra, A.B.W., "Perbandingan Algoritma K-Means dan Fuzzy C-Means Clustering Untuk Kualifikasi Data Kinerja Dosen", Seminar Nasional Serba Informatika, ISSN: 977-2579-7247-01, 2017.