

Penerapan Algoritma *K-Means* dan *Fuzzy C-Means* untuk Pengelompokan Kabupaten Kota Berdasarkan Produksi Padi di Provinsi Jawa Barat

Abdul Aziz
Universitas Buana Perjuangan
Karawang, Indonesia
if16.abdulaziz@mhs.ubpkarawang.ac.id

Amril Mutoi Siregar
Universitas Buana Perjuangan
Karawang, Indonesia
amrilmutoi@ubpkarawang.ac.id

Candra Zonyfar
Universitas Buana Perjuangan
Karawang, Indonesia
candra@ubpkarawang.ac.id

Abstract— Tanaman padi menjadi sumber bahan pangan utama hampir dari setengah penduduk dunia. Tidak terkecuali Indonesia memenuhi kebutuhan bahan pangannya dari tanaman padi. Dalam dunia komputer, *data mining* dikenal luas sebagai teknik penggalian data untuk mencari sebuah pola tersembunyi demi menghasilkan sebuah pengetahuan baru di dalam sekumpulan data. Secara khusus data mining memiliki metode tersendiri berdasarkan tujuan dari pemanfaatan himpunan data yaitu estimasi, prediksi, klasifikasi, klusterisasi dan asosiasi. Salah satu teknik yang bisa digunakan untuk tujuan pemetaan sebuah data adalah teknik klusterisasi. Klusterisasi merupakan sebuah teknik dalam data mining yang berfungsi untuk mengelompokkan data berdasarkan kemiripannya ke dalam kluster. Oleh karena itu penelitian ini menggunakan metode *K-Means* dan *Fuzzy C-Means*.

Kata kunci — *Algoritma, Fuzzy C-Means, K-Means, Kluster, Padi*

I. PENDAHULUAN

Tanaman padi menjadi sumber bahan pangan utama hampir dari setengah penduduk dunia. Tidak terkecuali Indonesia, hampir seluruh penduduk Indonesia memenuhi kebutuhan bahan pangannya dari tanaman padi. Dengan demikian tanaman padi merupakan tanaman yang mempunyai nilai spiritual, budaya, ekonomi, dan politik yang penting bagi bangsa Indonesia karena memengaruhi hajat hidup orang banyak [1]. Produksi padi di Jawa Barat dari Januari hingga September 2018 sebesar 8.108 ton Gabah Kering Giling (GKG). Produksi tertinggi terjadi pada bulan Maret yaitu sebesar 302 ton. Sementara itu, potensi produksi padi pada bulan Oktober, November, dan Desember masing-masing sebesar 521 ton, 446 ton, dan 445 ton. Dengan demikian, perkiraan total produksi padi 2018 adalah 9.539 ton. Tiga kabupaten dengan produksi padi (GKG) tertinggi terjadi di kabupaten Indramayu, Karawang, dan Subang dengan produksi masing-masing sebesar 1.392 ton, 1.124 ton, dan 991 ton [2].

Dalam dunia komputer, *data mining* dikenal luas sebagai teknik penggalian data untuk mencari sebuah pola tersembunyi demi menghasilkan sebuah pengetahuan baru di dalam sekumpulan data. *Data mining* secara khusus memiliki metode tersendiri berdasarkan tujuan dari pemanfaatan himpunan data yaitu estimasi, prediksi, klasifikasi, klusterisasi dan asosiasi. Salah satu teknik yang bisa digunakan untuk tujuan pemetaan sebuah data adalah teknik klusterisasi (*clustering*). *Clustering* merupakan sebuah teknik dalam *data mining* yang berfungsi untuk mengelompokkan data (*grouping*) berdasarkan kemiripannya ke dalam kluster. Ada beberapa algoritma yang dapat digunakan, salah satu yang populer digunakan untuk klusterisasi suatu *dataset* adalah *k-means* [3].

Data mining mulai dikenal sejak tahun 1990, ketika pekerjaan pemanfaatan data menjadi sesuatu yang penting dalam berbagai bidang, mulai dari bidang akademik, bisnis hingga medis. Munculnya *data mining* didasarkan pada jumlah data yang tersimpan dalam basis data semakin besar. Dalam berbagai literatur, teori-teori pada *data mining* sudah ada sejak lama seperti antara lain *Naïve-Bayes* dan *Nearest Neighbour*, pohon keputusan, aturan asosiasi, *k-means clustering* dan *text mining*. Dalam *data mining* disebut juga dengan *knowledge-discovery in database* (KDD) atau disebut penemuan pengetahuan data karena tujuan utama *data mining* adalah untuk memanfaatkan data dalam basis data dengan mengolahnya sehingga menghasilkan informasi baru yang berguna [4].

Clustering atau analisis *cluster* ialah suatu pengelompokan suatu set benda-benda fisik dan yang ilustrasi kedalam kelas kesamaan objek. Tujuannya ialah untuk mendapatkan kelompok objek yang sama atau mirip satu sama lain dalam kelompoknya. Semakin tinggi tingkat kesamaan objek dalam satu kelompok dan semakin tinggi tingkat beda pada tiap kelompok baik analisa kelompok semakin baik *performance*. *Clustering* dapat mengenali *cluster* alami dari dalam data yang berdasarkan kemiripan atribut yang disebut dengan segmentasi. *Unsupervised learning method* adalah proses ekstraksi tanpa dapat diprediksi hasilnya dan tidak satupun atribut yang digunakan untuk dapat membantu proses belajar yang tidak berlabel yaitu seperti metode, *fuzzy c-means*, *k-means*, *k-medoids*, hierarki dan lain-lain [5]. Menurut Daniel t. Larose *Clustering* berbeda dengan *classification* yang tidak ada variabel target untuk *clustering*. Sebaiknya, algoritma klusterisasi mencari ke segmen data seluruh set menjadi sub kelompok yang relative homogeny atau kelompok, di mana kesamaan catatan dalam *cluster* dimaksimalkan, dan kesamaan catatan luar kluster ini diminimalkan [6].

II. DATA DAN METODE

A. Algoritma *K-Means*

Algoritma *k-means* ditemukan oleh MacQueen pada tahun 1967. *K-Means* adalah salah satu teknik pengelompokan yang paling banyak digunakan karena kesederhanaan dan kecepatan. Algoritma ini mempartisi data ke dalam *k cluster* dengan menetapkan setiap objek *cluster centroid* terdekatnya, nilai rata-rata dari variabel untuk semua objek dalam *cluster* tertentu berdasarkan ukuran jarak yang digunakan. Terdapat beberapa teknik perhitungan jarak pada setiap data dalam *cosine*, dan *correlatin* [7].

Menurut Agusta *k-means* merupakan salah satu algoritma *clustering* (pengelompokan). *K-means clustering* merupakan metode *clustering* non-hierarki yang mengelompokkan data dalam bentuk satu atau lebih *cluster/kelompok*. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu *cluster/kelompok* yang lain, sehingga data yang berada dalam satu *cluster/kelompok* memiliki tingkat variasi kecil. Sedangkan menurut Wu dan Kumar *k-means* adalah algoritma *clustering* yang mempartisi himpunan D menjadi *k cluster* data. Algoritma *k-means* mengklaster semua titik data pada D demikian sehingga titik data X_i menjadi satu-satunya *k* partisi. Dengan kata lain, satu titik data hanya masuk ke dalam satu *cluster*.

Menurut Tan, Steinbech, dan Kumar adapun langkah-langkah dari algoritma *k-means*

1. Tentukan *k* data sebagai *centroid*, *k* adalah jumlah *cluster* yang digunakan (ditentukan oleh peneliti).
2. Tiap titik (data) kemudian dicari *centroid* terdekatnya.
3. Setiap himpunan titik (data) yang menjadi *centroid* disebut *cluster*.
4. Hitung kembali *centroid* dari setiap *cluster*.
5. Ulangi langkah 1-4 sampai *centroid* tidak berubah.

Menurut Wu dan Kumar metode *clustering* menggunakan algoritma *k-means*, ukuran kedekatan data dihitung menggunakan jarak *euclidean*. Algoritma *k-means* bertujuan untuk meminimumkan jarak total *euclidean* diantara setiap titik X_i , dan *cluster* terdekat yakni C_j . Sedangkan menurut Agusta jarak *euclidean* ditentukan dengan menggunakan persamaan berikut [8].

$$d_{ij} = \sqrt{\sum_{k=1}^n \{x_{ik} - x_{jk}\}^2} \quad (1)$$

Keterangan :

d_{ij} = jarak antara data ke-i dan ke-j

n = dimensi data

X_{ik} = koordinat data ke-i pada dimensi k

X_{jk} = koordinat data ke-j pada dimensi k

B. Algoritma *Fuzzy C-Means*

Algoritma *fuzzy c-means* (FCM) adalah suatu teknik pengklasteran data yang keberadaan tiap-tiap titik data suatu *cluster* ditentukan oleh nilai keanggotaan. Nilai keanggotaan tersebut akan mencakup bilangan *real* pada *interval* 0-1. FCM adalah salah satu metode *optimizing partitioned cluster*. Kelebihan metode FCM adalah penempatan pusat *cluster* yang lebih tepat dibandingkan dengan metode *cluster* lain. Caranya adalah dengan memperbaiki pusat *cluster* secara berulang, maka akan dapat dilihat bahwa pusat *cluster* akan bergerak menuju lokasi yang tepat.

Menurut Yan Algoritma dari *fuzzy c-means* adalah sebagai berikut :

1. Input data yang akan dikelompokkan, yaitu X berupa matriks berukuran $n \times m$ (n = jumlah sampel data, m = atribut setiap data). X_{ij} data sampel ke-I ($i=1,2,\dots,n$), atribut ke-j ($j=1,2,\dots,m$).
2. Tentukan *cluster* (c), pangkat untuk matriks partisi (w), maksimum iterasi (*MaxIter*), *error* terkecil yang diharapkan (ξ), fungsi objektif awal ($P_0=0$), dan iterasi awal ($t=1$).
3. Bangkitkan bilangan random $\mu_{ik}, i=1,2,\dots,n; k=1,2,\dots,c$ sebagai elemen matriks partisi awal U.
4. Hitung pusat *cluster* ke-k: V_{kj} dengan $k=1,2,\dots,c$; dan $j=1,2,\dots,m$, menggunakan persamaan berikut :

$$V_{kj} = \frac{\sum_{i=1}^n \left((\mu_{ik})^w \cdot X_{ij} \right)}{\sum_{i=1}^n (\mu_{ik})^w} \quad (1)$$

dengan :

V_{kj} = pusat *cluster* ke-k untuk atribut ke-j

μ_{ik} = derajat keanggotaan untuk data sampel ke-I pada *cluster* ke-k

X_{ij} = data ke-i, atribut ke-j

1. Hitung fungsi objektif pada iterasi ke-t menggunakan persamaan berikut :

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right) \tag{2}$$

dengan :

V_{kj} = pusat *cluster* ke-k untuk atribut ke-j

μ_{ik} = derajat keanggotaan untuk data sampel ke-i pada *cluster* ke-k

X_{ij} = data ke-i, atribut ke-j

P_t = antara data ke-i dan data

2. Hitung perubahan matriks partisi menggunakan persamaan berikut :

$$\mu_{ik} = \frac{\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{1}{w-1}}}{\sum_{k=1}^c \left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{1}{w-1}}} \tag{3}$$

dimana :

V_{kj} = pusat *cluster* ke-k untuk atribut ke-j

X_{ij} = data ke-i, atribut ke-j

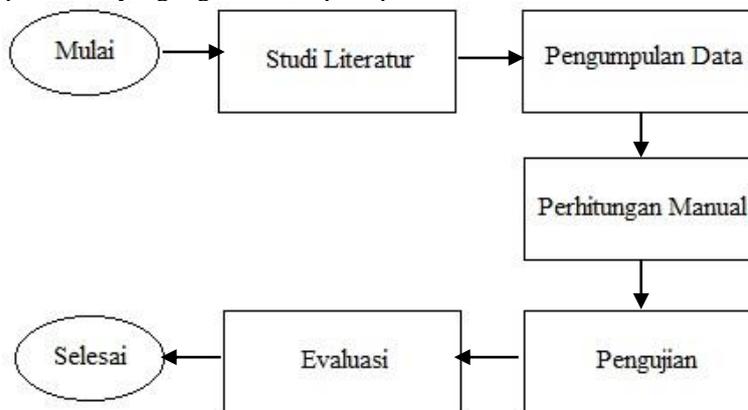
μ_{ik} = derajat keanggotaan untuk data sampel ke-i pada *cluster* ke-k

3. Cek kondisi berhenti :

Jika : $(|P_t - P_{t-1}| < \xi)$ atau $(t > MaxIter)$ maka berhenti. Jika tidak : $t = t + 1$, ulangi langkah ke-4 [9].

C. Gambaran Umum Penelitian

Adapun gambaran umum penelitian yang digunakan, seperti pada Gambar 1.



Gambar 1 Gambaran Umum Penelitian

D. Persiapan Data

Tahapan awal yang dilakukan pada penelitian ini adalah dengan mempersiapkan *dataset* yang diperoleh dari situs resmi dari situs resmi Kementerian Pertanian Republik Indonesia. Data yang diambil yaitu mengenai hasil produksi padi setiap provinsi di Jawa Barat. Data yang akan diolah yaitu data dari tahun 2010-2015 dengan menggunakan 27 data provinsi Jawa Barat.

E. Seleksi Data

Dari empat atribut hanya satu atribut yang digunakan pada proses perhitungan data. Atribut yang dipakai yaitu hasil produksi padi yang terdiri dari tahun 2010-2015. Atribut tersebut dipilih karena cocok digunakan dalam perhitungan, dan nilai dari atribut tersebut dapat digunakan sebagai dasar untuk menentukan daerah mana yang tingkat produksi padinya cukup baik, sangat baik dan baik.

III. HASIL DAN PEMBAHASAN

A. Hasil Algoritma *K-Means*

1. Menentukan jumlah kluster. Pada penelitian ini akan dibagi menjadi tiga kluster.
2. Tentukan titik *centroid* atau pusat kluster. Titik *centroid* di tentukan secara acak.

Tabel 1 merupakan nilai titik *centroid* awal yang akan digunakan dalam proses perhitungan algoritma *K-Means* dimana c1 merupakan data ke-23, c2 data ke-19 dan c3 adalah data ke-13.

Tabel 1. Titik *Centroid* Awal

C1	4300	3842	1820	2384	2271	2427
C2	0	0	0	0	157035	166428
C3	1358441	1415050	1376604	1435938	1361374	1294158

3. Hitung jarak data terhadap masing-masing *centroid* dengan rumus *Euclidean*.

Hasil iterasi pertama dengan menggunakan data sebanyak 27 data menghasilkan anggota kluster c1 sebanyak 9 anggota c2 sebanyak 11 anggota dan c3 sebanyak 7 anggota yang ditunjukkan pada Tabel 2.

Tabel 2. Hasil Iterasi 1

Data ke-	C1	C2	C3	DC1	DC2	DC3	Kluster
1.	1260222,542	1154029,669	2101092,772		ok		2
2.	1983239,11	1857555,781	1397076,416			ok	3
3.	2070914,386	1955871,603	1295270,04			ok	3
.
.
.
27.	93364,07492	196229,8312	3266424,705	ok			1
Total				9	11	7	

4. Perbarui nilai titik *centroid*.

Mencari nilai *centroid* baru adalah dengan membagi hasil total jumlah atribut dengan total jumlah anggota setiap masing-masing kluster. Hasil dalam perhitungan mencari nilai *centroid* baru ditandai warna biru yang selanjutnya hasil tersebut adalah angka yang akan digunakan dalam perhitungan selanjutnya, dengan kata lain hasil tersebut adalah nilai *centroid* baru. Tabel 3 menunjukkan nilai *centroid* baru.

Tabel 3. Mencari Nilai *Centroid* Baru

Tahun	2010	2011	2012	2013	2014	2015
Total Jumlah Atribut Total	193131	175985	163958	170095	147335	152167
Hasil Bagi Total	21459	19533,88889	18217,5556	18899,44444	16370,55556	16907,4444
Total Jumlah Atribut Total	4738276	4626300	4330556	4894484	4466817	4252808

Jumlah	11	11	11	11	11	11
Hasil Bagi	430752,3636	420572,7273	393686,9091	444953,0909	406074,2727	386618,9091
Total						
Jumlah Atribut	6805665	6841586	6777347	7018581	7030747	6968169
Total						
Jumlah Anggota	7	7	7	7	7	7
Klaster 1						
Hasil Bagi	972237,8571	977369,4286	968192,4286	1002654,429	1004392,429	995452,7143

5. Ulangi langkah 3 dan 4 dengan memakai pusat klaster yang baru. Karena pada iterasi ke-4 dan ke-5 anggota klaster tidak ada yang berubah maka proses dihentikan karena sudah konvergen dengan hasil klaster 1 sebanyak 12 anggota, klaster 2 sebanyak 8 anggota dan klaster 3 sebanyak 7 anggota.

B. Hasil Algoritma *Fuzzy C-Means*

1. Secara acak pilih bilangan dengan total dari nilai ketiga klaster adalah 1.
2. Membentuk matriks partisi awal U (derajat keanggotaan dalam *cluster*) dengan ukuran 27x3.

Tabel 4. Matriks Partisi Awal U

Data ke-	μ_1	μ_2	μ_3
1.	0,2	0,4	0,3
2.	0,4	0,5	0,1
3.	0,3	0,2	0,1
.	.	.	.
.	.	.	.
.	.	.	.
Data ke-	μ_1	μ_2	μ_3
27.	0,3	0,4	0,1

Selanjutnya yaitu menghitung *centroid* klaster sebagai berikut.

• Data 1 *Cluster 1*

1.	Kab. Bogor	538804	497711	494815	559367	517442	492207
----	------------	--------	--------	--------	--------	--------	--------

μ_1^w . A = 21552,16

μ_1^w . B = 19908,44

μ_1^w . C = 19792,6

μ_1^w . D = 22374,68

μ_1^w . E = 20697,68

μ_1^w . F = 19688,28

• Data 1 *Cluster 2*

μ_1^w . A = 86208,64

μ_1^w . B = 79633,76

μ_1^w . C = 79170,4

μ_1^w . D = 89498,72

μ_1^w . E = 82790,72

μ_1^w . F = 78753,12

• Data 1 *Cluster 3*

μ_1^w . A = 48492,36

μ_1^w . B = 44793,99

μ_1^w . C = 44533,35

μ_1^w . D = 50343,03

μ_1^w . E = 46569,78

μ_1^w . F = 44298,63

Tabel 5. Centroid Cluster 1

Centroid C1							
Data ke-	$\mu 1^w$	$\mu 1^{w*} A$	$\mu 1^{w*} B$	$\mu 1^{w*} C$	$\mu 1^{w*} D$	$\mu 1^{w*} E$	$\mu 1^{w*} F$
1.	0,04	21552,16	19908,44	19792,6	22374,68	20697,68	19688,28
2.	0,16	128947,84	115844	132126,08	122826,88	143597,6	134824,64
3.	0,09	77600,52	71174,16	78168,42	79439,58	74749,05	76648,41
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
27.	0,09	4309,92	3816,99	3564,99	3701,16	3196,71	3457,17
Σ	2,89	1391869,1	1395257,	1339097,2	1413875,0	1359934,9	1318388,23
		1	11	8	6		
V1		481615,60	482787,9	463355,46	489230,12	470565,70	456189,699
			273	02	46	93	

Sehingga untuk iterasi pertama, didapat titik centroid pada Table 6.

Tabel 6. Centroid Perulangan I

	A	B	C	D	E	F
V1	481615,60	482787,92	463355,46	489230,12	470565,70	456189,6
	9	73	02	46	93	99
V2	368404,32	355610,51	350845,95	375000,52	368415,01	355880,5
	41	78	65	57	58	81
V3	452615,36	447775,28	428508,14	467829,25	450191,40	429140,9
	11	24	81		28	444

Jumlahkan seluruh hasil perhitungan dari semua data, maka akan didapatkan nilai fungsi objektif dari iterasi pertama. Tabel 7 adalah hasil perhitungan fungsi objektif.

Tabel 7. Nilai Fungsi Objektif Perulangan I

Data ke-	L1	L2	L3	LT
1.	8.759.808.989	38.695.063.482	18.878.316.870	66.333.189.341
2.	125.159.722.469	286.454.293.449	8.543.791.158	420.157.807.076
3.	80.105.425.386	51.681.925.128	9.743.749.370	141.533.099.883
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
27.	105.147.420.629	104.456.809.216	10.255.477.329	219.859.707.175
			Σ	7.160.205.192.13
				7

Fungsi objektif didapatkan pada iterasi pertama sebesar 7.160.205.192.137. Nilai tersebut akan digunakan untuk menentukan apakah iterasi akan berhenti atau berlanjut. Setelah dapat fungsi objektif pada iterasi pertama. Matriks partisi baru yang sudah di hitung seperti pada Tabel 8.

Tabel 8. Matriks Partisi Baru

Data ke-	$\mu 1$	$\mu 2$	$\mu 3$
1.	0,339034274	0,307003035	0,353962691
2.	0,384870858	0,262750959	0,352378184
3.	0,384296717	0,264732587	0,350970696
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮
27.	0,254535041	0,455498467	0,289966492

Matriks partisi baru tersebut akan digunakan untuk perhitungan iterasi selanjutnya jika memenuhi untuk lanjut ke iterasi selanjutnya, jika tidak matriks partisi tersebut adalah matriks partisi akhir yang akan digunakan untuk mencari cluster dari setiap data.

Jika lebih kecil dari batas kesalahan atau sudah mencapai t_{max} , maka iterasi berhenti. Pada iterasi pertama fungsi objektif sebesar 7160205192137, maka masih berlanjut ke iterasi selanjutnya. Pada iterasi selanjutnya dilakukan perhitungan yang sama hanya saja menggunakan matriks partisi yang sudah diperbarui. Berikut hasil akhir clustering pada tabel 9.

Tabel 9. Hasil *Clustering Fuzzy C-Means*

No.	Lokasi	μ_1	μ_2	μ_3	C
1.	Kab. Bogor	0,266146815	0467621942	0,266231243	C2
2.	Kab. Sukabumi	0,431924962	0,136044877	0,432030161	C3
3.	Kab. Cianjur	0,438162761	0,123581431	0,438255808	C3
.
.
.
27.	Kota Banjar	0,012945906	0,974105646	0,012948448	C2

Proses *clustering* sudah selesai dilakukan dengan hasil seperti yang terlihat pada Tabel 9. Hasil dari nilai *centroid* akhir *Fuzzy C-Means* pada Tabel 10.

Tabel 10. Nilai *Centroid* Akhir *Fuzzy C-Means*

C	A	B	C	D	E	F
C1	855980,5247	858883,8637	842413,9067	886554,7915	863581,8689	850559,5988
C2	131658,2388	125536,1923	117808,0353	130302,5367	130509,6132	125877,477
C3	855906,3554	858797,3031	842332,4523	886473,5792	863502,2317	850475,4402

C. Implementasi *RapidMiner*

Tahap selanjutnya yaitu melakukan pengujian menggunakan *Software RapidMiner Studio*. Cara mengolah dataset pada *RapidMiner* sebagai berikut :

1. Pilih *Import Data*.
2. Pilih data yang akan diolah.
3. Simpan pada *Local Repository* dan beri nama *File* kemudian klik *Finish*.
4. Pilih *Design* kemudian drag *File* yang sudah disimpan tadi dan pilih algoritma yang akan digunakan pada kolom *Operators*.
5. Hubungkan data dengan algoritma yang sudah dipilih, lalu klik algoritma yang dipilih sehingga muncul *Box Parameters*.
6. Ketik berapa kluster yang akan dipakai dengan berapa iterasi yang dijalankan pada proses tersebut, kemudian pada *Measure Types* pilih *Numerical Measures* sehingga pada kolom *Numerical Measure* akan otomatis berubah menjadi *Euclidean Distance*.
7. Klik tombol *Play* sehingga akan muncul hasil, kemudian pilih *Cluster Model (Clustering)* untuk menampilkan hasil kluster.

Hasil dari perhitungan menggunakan *RapidMiner* yang didapatkan pada Tabel 11.

Tabel 11. Hasil Perhitungan *K-Means* Menggunakan *RapidMiner*

Jumlah Anggota Kluster
3 Anggota
4 Anggota
20 Anggota
27 Anggota

IV. KESIMPULAN DAN SARAN

Berdasarkan hasil perhitungan algoritma *K-Means* dan *Fuzzy C-Means* dengan *dataset* yang digunakan sebanyak 27 data produksi padi di provinsi Jawa Barat yang digunakan, Kedua algoritma tersebut mengelompokkan *dataset* sebanyak 27 data kabupaten kota kedalam tiga kluster yaitu kluster cukup baik, kluster sangat baik dan kluster baik. Algoritma *K-Means* menghasilkan jumlah anggota kluster cukup baik sebanyak 12, kluster sangat baik sebanyak 8 dan kluster baik sebanyak 7 kabupaten kota.

PENAKUAN

Naskah ilmiah ini adalah sebagian dari penelitian Tugas Akhir milik Abdul Aziz dengan judul Penerapan Algoritma *K-Means* dan *Fuzzy C-Means* Untuk Pengelompokan Kabupaten Kota Berdasarkan Produksi Padi di Provinsi Jawa Barat, yang dibimbing oleh Amril Mutoi Siregar dan Candra Zonyfar.

DAFTAR PUSTAKA

- [1] Isaq, M. Rumiati, A.T. dan Permatasari, E.O., 2017. Analisis Faktor-Faktor yang Mempengaruhi Produksi Padi di Provinsi Jawa Timur Menggunakan Regresi Semiparametrik *Spline*. *Jurnal Sains dan Seni ITAS* : 101-107.
- [2] Metodologi, Perbaikan, Perhitungan Data, dan Produksi Beras., 2018. Luas Panen Dan Produksi Padi di Jawa Barat : 1-10.

- [3] Ridlo, M.R. S. dan Primajaya, A., 2017. Implementasi Algoritma *K-Means* Untuk Pemetaan Produktivitas Panen Padi Di Kabupaten Karawang. *Departemen Teknik Elektro dan Teknologi Informasi* : 426-433.
- [4] Nur, F. Zarlis, M. dan Nasution, B.B., 2017. Penerapan Algoritma *K-Means* Pada Siswa Baru Sekolah Menengah Kejuruan Untuk *Clustering* Jurusan. *Jurnal Nasional Informatika dan Teknologi Jaringan* : 100-105.
- [5] Siregar, A.M., 2019. Pengelompokan Bidang Laju Pertumbuhan Ekonomi Indonesia Menggunakan Algoritma *K-Means*. *Jurnal Accounting Information System* : 140-151.
- [6] Taslim dan Fajrizal., 2016. Penerapan Algoritma *K-Means* Untuk *Clustering* Data Obat pada Puskesmas Rumbai. *Jurnal Teknologi Information & Komunikasi Digital Zone* : 108-114.
- [7] Yohannes., 2016. Analisis Perbandingan Algoritma *Fuzzy C-Means* dan *K-Means* : 151-155.
- [8] Abdurrahman, G., 2016. *Clustering* Data Ujian Tengah Semester (UTS) *Data Mining* Menggunakan Algoritma *K-Means*. *JUSTINDO, Jurnal Sistem & Teknologi Informasi Indonesia, Vol.1, No.2* : 71-79.
- [9] Wijaya, A.K, 2014. Implementasi *Data Mining* Dengan ALgoritma *Fuzzy C-Means* Studi Kasus Penjualan di UD Subur Baru : 1-8.