

IMPLEMENTASI ALGORITMA K-MEANS UNTUK KLASTERISASI POTENSI DESA PADA SEKTOR PRODUKSI PERTANIAN DI KABUPATEN BOJONEGORO

Teguh Pribadi¹⁾, Rahmad Irsyada²⁾, Hastie Audytra³⁾ Doni Abdul Fatah⁴⁾

^{1,2,3}Program Studi Teknik Informatika, Fakultas Teknik, Universitas Sunan Giri
Bojonegoro (UNUGIRI)

Jl Jenderal Ahmad Yani no. 10 Bojonegoro

⁴Jurusan Teknik Informatika, Fakultas Teknik, Universitas Trunojoyo Madura

Jl. Raya Telang, PO BOX 2 Kamal, Bangkalan

Email : ¹pribadi.teguh90@gmail.com, ²irsyada.rahmat@gmail.com,

³hastie.audytra@gmail.com, ⁴doni.fatah@trunojoyo.ac.id

ABSTRAK

Hasil pertanian (panen) di lingkup desa merupakan salah satu deteksi kondisi desa tersebut dalam bidang pertanian. Terkait kondisi hasil pertanian, pemetaan klusterisasi hasil pertanian desa juga perlu dilakukan dan diketahui agar penetapan program desa khususnya dalam bidang pertanian dapat menentukan prioritas sesuai dengan kondisinya. Beberapa data menjadi fokus utama pemetaan hasil pertanian di Bojonegoro adalah hasil panen Padi, Jagung, Kacang Tanah, dan Kedelai. Pengelompokan potensi hasil pertanian desa sesuai laporan BPS Bojonegoro tahun 2018 dimana jumlah klaster yang diusulkan 2 klaster yaitu klaster baik dan klaster tidak baik. Metode klusterisasi menggunakan algoritma K-Means dengan perhitungan jarak menggunakan Euclidean distance sedangkan perhitungan evaluasi jarak hasil klaster menggunakan Davies-Bouldin Index (DBI). Hasil klusterisasi adalah klaster 1 (baik) sebanyak 295 desa dan klaster 2 (tidak baik) 106 dengan DBI sebesar 0.70.

Kata kunci: *Davies-Bouldin Index (DBI), Hasil Pertanian Desa, K-Means, Potensi*

ABSTRACT

Agriculture (harvest) in the village area is one of the detection conditions of the village in agriculture. Regarding the condition of agricultural products, mapping of clustering of village agricultural products also needs to be carried out and it is known that the establishment of village programs, especially in agriculture, can determine priorities according to their conditions. Some data are the main focus of mapping agricultural products in Bojonegoro are rice, corn, peanuts and soybeans. The grouping of rural agricultural yields is in accordance with the BPS Bojonegoro report in 2018 where the number of proposed 2 cluster clusters is good and bad clusters. The clustering method uses the K-Means algorithm with distance calculation using Euclidean distance while the calculation of the distance evaluation results using cluster Davies-Bouldin Index (DBI). The results of clustering are cluster 1 (good) as many as 295 villages and cluster 2 (not good) 106 with DBI of 0.70.

Keywords : *Davies-Bouldin Index (DBI), Village Agricultural Products, K-Means, Potential*

PENDAHULUAN

Berdasarkan UU nomor 19 tahun 2013 Bab I pasal 1 ayat 4, Pertanian [1] adalah kegiatan mengelola sumber daya alam hayati dengan bantuan teknologi, modal, tenaga kerja, dan manajemen untuk menghasilkan Komoditas Pertanian yang mencakup tanaman pangan, hortikultura, perkebunan, dan/atau peternakan dalam suatu *agroekosistem*. Dalam UU tersebut juga mengatur tentang perlindungan dan pemberdayaan petani sehingga terwujud kedaulatan dan kemandirian petani dalam rangka meningkatkan taraf kesejahteraan, kualitas, dan kehidupan petani yang lebih baik. Di kabupaten Bojonegoro terdapat 430 desa yang tersebar di 28 kecamatan yang mana memiliki total produksi padi 963.137,12 ton, non-padi 324.241,493ton, ternak besar 510.080 ekor, dan ternak kecil 2.358.436 ekor [2]. Persebaran masing-masing 4 sektor tersebut tersebut masih relatif distributif dan diperlukan klasterisasi untuk menggolongkan desa yang unggul, berkembang, dan perlu penanganan.

Clustering [3] adalah metode yang digunakan dalam data mining yang cara kerjanya mencari dan menglompokkan data yang mempunyai kemiripan karakteristik antara data satu dengan data lainnya yang telah diperoleh. Ciri khas dari teknik data mining ini adalah mempunyai sifat tanpa arahan (*unsupervised*), yang dimaksud adalah teknik ini diterapkan tanpa perlunya data training dan tanpa ada *teacher* serta tidak memerlukan target *output*. Metode *clustering* yang mempunyai sifat efisien dan cepat yang dapat digunakan salah satunya adalah metode *k-means* dimana metode ini bertujuan untuk membuat *cluster* objek berdasarkan atribut menjadi *k* partisi. cara kerja metode ini adalah mula – mula ditentukan *cluster* yang akan dibentuk, pada elemen pertama dalam tiap *cluster* dapat dipilih untuk dijadikan sebagai titik tengah (*centroid*), selanjutnya akan dilakukan pengulangan langkah-langkah hingga tidak ada objek yang dapat dipindahkan lagi [4].

Berdasarkan hal-hal tersebut diatas, penelitian ini berfokus pada masalah terkait pengklasterisasian/ pengelompokan potensi desa-desa pada sektor produksi pertanian di kabupaten Bojonegoro menggunakan algoritma *K-Means* sehingga mampu memberikan informasi/ pengetahuan kepada kepala daerah dalam mengambil kebijakan terkait penanganan pertanian.

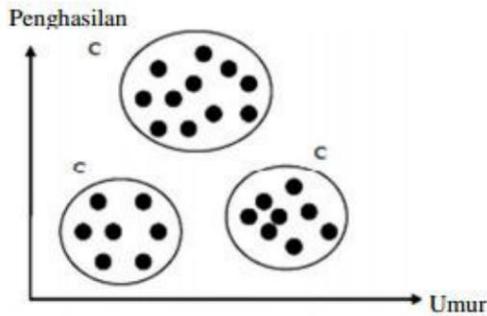
Data mining merupakan metode yang digunakan dalam pengolahan data berskala besar oleh karena itu data mining memiliki peranan yang sangat penting dalam beberapa bidang kehidupan diantaranya yaitu bidang industri, bidang keuangan, cuaca, ilmu dan teknologi [3]. Dalam data mining juga terdapat metode – metode yang dapat digunakan seperti klasifikasi, *clustering*, *regresi*, seleksi variabel, dan market basket analisis.

Clustering atau pengklasteran adalah suatu teknik data mining yang digunakan untuk menganalisis data untuk memecahkan permasalahan dalam mengelompokkan data atau lebih tepatnya mempartisi dari dataset ke dalam subset. Pada teknik *clustering* targetnya adalah untuk kasus pendistribusian (objek, orang, peristiwa dan lainnya) ke dalam suatu kelompok, hingga derajat tingkat keterhubungan antar anggota cluster yang sama adalah kuat dan lemah antara anggota cluster yang berbeda [5]. Pada proses analisis cluster metode yang digunakan untuk membagi data menjadi subset data berdasarkan kesamaan atau kemiripan yang telah ditentukan sebelumnya. Jadi analisis *cluster* secara umum dapat dikatakan bahwa [5] :

- a) Data yang terdapat dalam satu *cluster* memiliki tingkat kesamaan yang tinggi, dan
- b) Data yang terdapat dalam suatu *cluster* yang berbeda memiliki tingkat kesamaan yang rendah.

Sebagai contoh dapat dilihat pada gambar 1. Pada gambar 1 dapat dilihat misalkan data tersebut merupakan data konsumen sederhana yang terdapat dua

atribut didalamnya, yaitu umur dan penghasilan. Pada data yang berdasarkan dua atribut tersebut kemudian dibagi menjadi tiga cluster yaitu *cluster C1* yang terdiri dari konsumen usia muda dan berpenghasilan rendah, *cluster C2* terdiri dari konsumen usia muda dan tua berpenghasilan tinggi, dan *cluster C3* terdiri dari konsumen usia tua dan berpenghasilan relatif rendah.



Gambar 1. Grafik Clustering

Salah satu algoritma *clustering* adalah *K-Means* dimana *K-Means* merupakan suatu algoritma yang digunakan dalam pengelompokan secara pertisi yang memisahkan data ke dalam kelompok yang berbeda – beda. Algoritma ini mampu meminimalkan jarak antara data ke clusternya. Pada dasarnya penggunaan algoritma ini dalam proses clustering tergantung pada data yang didapatkan dan konklusi yang ingin dicapai di akhir proses [5]. Salah satu penghitung jarak adalah manhattan distance yang mana prosesnya adalah:

- a. Tentukan jumlah cluster
- b. Alokasikan data ke dalam *cluster* secara *random*
- c. Hitung *centroid*/ rata-rata dari data yang ada di masing-masing cluster dengan persamaan

$$D_{L1}(x_1, x_2) = |x_2 - x_1| = \sqrt{\sum_{j=1}^p (x_{2j} - x_{1j})^2} \quad (1)$$

- d. Alokasikan masing-masing data ke *centroid*/rata-rata terdekat
- e. Kembali ke Step c, apabila masih ada data yang berpindah *cluster* atau apabila perubahan nilai *centroid*, ada yang di atas nilai *threshold* yang

ditentukan atau apabila perubahan nilai pada *objective function* yang digunakan di atas nilai *threshold* yang ditentukan.

METODE

Metodologi penelitian yang dilakukan dengan langkah-langkah sebagai berikut:

- a. Studi Literatur
Pencarian literatur dalam rangka melengkapi pengetahuan dasar dan teori-teori melalui buku-buku, jurnal, maupun media internet.
- b. Pengumpulan Data
Pengumpulan data sekunder melalui situs <https://bojonegorokab.bps.go.id/publication.html> yang mana didapatkan data atau *dataset* sebanyak 430 data sesuai jumlah desa yang tersebar pada 28 kecamatan di Bojonegoro yang ditunjukkan pada tabel 1.

Tabel 1. Data Produksi Hasil Pertanian Bojonegoro Tahun 2017

No	Desa	Kecamatan	Hasil Pertanian			
			P	J	K T	K d
1	Sidobandung	Balen	Balen	3 1 4 6	27 7.2 3	74 .3 3
2	Mayangkawis	Balen	Balen	3 2 8 4	45 1.6 8	75 .2 9
3	Kenep	Balen	Balen	3 0 9 8	41 3.4 4	79 .3 4
4	Pohbogo	Balen	Balen	1 6 5 8	40 4.5 9	25 .1 7
5	Penganten	Balen	Balen	7 0 3 3	31 9.6 8	84 .8 5
...

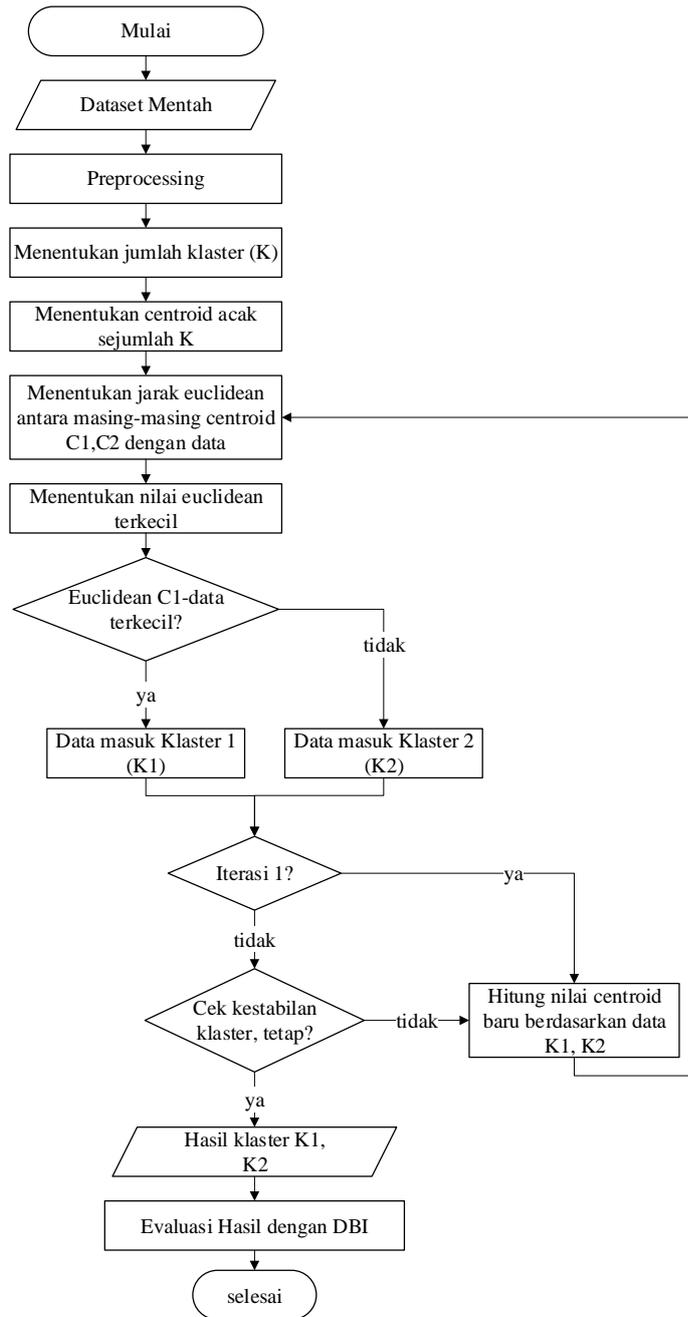
No	Desa	Kecamatan	Hasil Pertanian			
			P	J	KT	Kd
430	Banjar sari	Trucuk	704	270	4238	

c. Identifikasi Masalah

Klasterisasi yang diusulkan peneliti sebanyak 2 kluster yaitu baik dan tidak baik.

d. Klasterisasi - Evaluasi

Flowchart dari *preprocessing*, klasterisasi *K-Means*, sampai hasil kluster serta evaluasinya dapat dilihat pada gambar 2.



Gambar 2. Metode yang diusulkan

HASIL DAN PEMBAHASAN

Proses pengujian dan pembahasan sesuai flowchat metodologi dalam bab iii penelitian ini adalah sebagai berikut:

a. Mempersiapkan data hasil pertanian

Dataset hasil pertanian diambil dari data publikasi Laporan tahun 2018 Badan Pusat Statistik Bojonegoro (<https://bojonegorokab.bps.go.id/publication.html>) salah satunya tentang laporan kecamatan atas desa-desa termasuk data hasil pertanian. Data dari hasil pertanian dapat dilihat pada tabel 1.

b. Preprocessing

Metode *preprocessing* yang digunakan adalah *missing value*, *inconsistent data*, dan *noisy data* dengan *data cleaning* yaitu data yang masuk pada 3 kategori tersebut akan dihapus. Dari 430 data hasil pertanian, ada 29 data yang di *cleaning* sehingga tersisa sebanyak 401. Sebanyak 401 data tersebut siap untuk diproses data metode atau algoritma. Data lengkap setelah di *preprocessing* dapat dilihat pada [5] dan sebagiannya dapat dilihat pada tabel 2.

Tabel 2. Dataset hasil *processing*

No	P	J	KT	Kd
1	3146	277.23	74.33	243.1
2	3284	451.68	75.29	211.5
3	3098	413.44	79.34	343.4

4	1658	404.59	25.17	151.2
..
401	704	270	42.38	18.21

c. Menentukan jumlah kluster

Salah satu jumlah kluster yang diusulkan adalah 2 kluster yaitu kluster baik dan tidak baik.

d. Menentukan centroid awal

Centroid ditentukan secara acak dengan rincian seperti tabel 3. Dimana P merupakan Padi, J merupakan Jagung, KT merupakan Kacang Tanah, Kd merupakan Kedelai, dan satuan nilai dalam bilangan tersebut adalah ton.

Tabel 3. Daftar 2 *centroid* Awal

centroid ke-	P	J	KT	Kd
centroid 1 (c1)	779	68.8	92.31	44.64
centroid 2 (c2)	2584	18.5	73.26	1.43

e. Menghitung euclidean (iterasi)

1) Iterasi 1

Menghitung jarak *dataset* dengan pusat kluster dengan menggunakan jarak *euclidean* (persamaan 1). Berikut adalah perhitungan jarak *Euclidean* (d) antara centroid 1 dan *dataset* hasil *preprocessing* (lampiran 3) data ke-1.

$$d(c1, data1) = \sqrt{\sum_i (c1 - data1)^2} \quad (2)$$

$$d(c1, data1)$$

$$= \sqrt{(779 - 3146)^2 + (68.8 - 277.23)^2 + (92.31 - 74.33)^2 + (44.64 - 243.1)^2}$$

$$d(c1, data1) = \sqrt{(-2367)^2 + (-208.43)^2 + (-17.98)^2 + (-198.46)^2}$$

$$d(c1, data1) = \sqrt{5602689 + 43443.0649 + 323.2804 + 39386.3716}$$

$$d(c1, data1) = \sqrt{5685841.717}$$

$$d(c1, data1) = 2384.500308 \text{ (dibulatkan 2384.50)}$$

Sedangkan berikut adalah *centroid 2* dan *dataset* hasil perhitungan jarak *Euclidean* (d) antara *preprocessing* (lampiran 3) data ke-1.

$$d(c1, data1) = \sqrt{\sum_i (c2 - data1)^2} \quad (3)$$

$$d(c1, data1)$$

$$= \sqrt{(2584 - 3146)^2 + (18.5 - 277.23)^2 + (73.26 - 74.33)^2 + (1.43 - 243.1)^2}$$

$$d(c1, data1) = \sqrt{(-562)^2 + (-258.73)^2 + (-1.07)^2 + (-241.67)^2}$$

$$d(c1, data1) = \sqrt{315844 + 66941.2129 + 1.1449 + 58404.3889}$$

$$d(c1, data1) = \sqrt{441190.7467}$$

$$d(c1, data1) = 664.2219107 \text{ (dibulatkan 664.22)}$$

Setelah perhitungan jarak selesai, jarak tersebut dipakai untuk penentuan kluster dari data 1 yaitu didasarkan pada jarak terpendek/terkecil. Berdasarkan jarak antara c1-data1 (2384.50) dengan c2-data1 (664.50), maka data 1 termasuk kluster 2. Lakukan perhitungan antara data selanjutnya dengan c1 dan c2, sehingga menghasilkan seperti tabel 3.

Tabel 4. Perhitungan *Euclidean dataset* dengan c1, c2 iterasi 1

No	P	J	KT	Kd	Jarak data - C1	Jarak data - C2	Kluster
1	3	27	74	24	238	664	K2
	1	7.2	.3	3.	4.5	.22	
	4	3	3	1	0		
2	3	45	75	21	253	849	K2
	2	1.6	.2	1.	9.6	.58	
	8	8	9	5	4		
3	3	41	79	34	236	732	K2
	0	3.4	.3	3.	3.4	.91	
	9	4	4	4	6		
4	1	40	25	15	949	101	K1
	6	4.5	.1	1.	.35	5.5	
	5	9	7	2		2	
...
...
4	7	27	42	18	222	189	K1
0	0	.3	.2	.03	7.0		
1	4	8	1			7	

Setelah dilakukan seluruh perhitungan, banyaknya data pada kluster 1 adalah 224 dan kluster 2 adalah 177.

2) Iterasi 2

Dalam K-Means, iterasi kedua pasti akan dilakukan karena digunakan untuk mengecek kestabilan kluster. Pada iterasi kedua, penentuan pusat kluster/centroid menggunakan hasil kluster tabel 4. Pada iterasi 1, Kluster 1 memiliki banyak data 224 dengan total *value* atribut P=200089.3, J=103089.3.1, KT=10948.95, Kd=11612.38 sedangkan kluster 2 memiliki banyak data 177 dengan total *value* atribut P=557608.3, J=82267.84, KT=12961.56,

Kd=13129.05. *Centroid* iterasi kedua dihasilkan dari total atribut dibagi banyaknya data pada kluster tersebut.

atribut P pada c1

$$= \sum \frac{ai}{n} = \frac{200089.3}{224} = 893.26$$

atribut P pada c2

$$= \sum \frac{ai}{n} = \frac{557608.3}{177} = 3150.33$$

Lakukan seterusnya untuk seluruh atribut dengan masing-masing centroid sehingga menghasilkan seperti tabel 5.

Tabel 5. Centroid iterasi 2

Iterasi 2	P	J	KT	Kd
centroid 1	893.26	460.22	48.88	51.84
centroid 2	3150.33	464.79	73.23	74.18

Lakukan perhitungan jarak *Euclidean dataset* dengan *centroid* iterasi 2 seperti langkah diiterasi 1, hasilnya dapat dilihat pada tabel 6.

Tabel 6. Perhitungan *Euclidean dataset* dengan c1, c2 iterasi 2

No	P	J	KT	Kd	Jarak data - C1	Jarak data - C2	Kluster
1	3	64	11	24	229	439	K2
	1	.3	.1	3.	5.5	.08	
	4	1	4	1	6		
2	3	73	11	21	242	439	K2
	2	.9	.0	1.	7.3	.72	
	8	5	3	5	0		
3	3	58	19	34	226	493	K2
	0	.5	.6	3.	0.1	.12	
	9	8	3	4	2		
4	1	73	6.	15	863	154	K1
	6		72	1.	.96	6.2	
	5			2		5	
...
...

No	P	J	K T	K d	Jarak data - C1	Jarak data - C2	Klaster
401	4166	140	1284	1339	3288.80	1069.77	K2

Pada iterasi kedua dilakukan cek kestabilan terkait hasil kluster iterasi 1 dan iterasi 2. Iterasi akan dilanjutkan lagi jika ada data yang berpindah kluster dari hasil iterasi 1 ke iterasi 2 atau dikatakan tidak stabil serta iterasi akan dihentikan jika hasil kluster sama atau stabil. Pada iterasi 2, data yang tetap klusternya sebanyak 302 data dan data yang berpindah kluster sebanyak 99 data. Tabel 7 memperlihatkan hasil pengecekan kestabilan antara iterasi 1 dan iterasi 2.

Tabel 7. Cek kestabilan kluster iterasi 1 dan 2

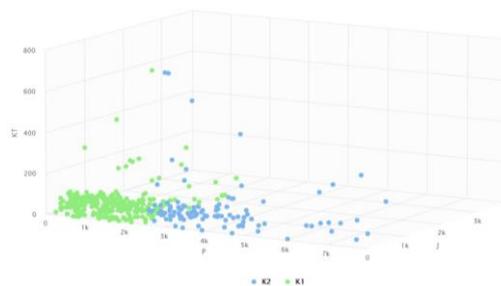
No	Hasil Kluster Iterasi 1	Hasil Kluster Iterasi 2	0: tetap 1: pindah
1	K2	K2	0
2	K2	K2	0
3	K2	K2	0
4	K1	K1	0
....
399	K1	K2	1
400	K1	K1	0
401	K1	K2	1

Dari hasil pengecekan kestabilan hasil kluster iterasi 1 dengan iterasi 2 yang diperlihatkan tabel 7, masih terdapat data yang berpindah kluster maka masih harus dilakukan iterasi lagi.

3) Ulangi iterasi

Lakukan iterasi sampai semua kluster tetap/0. Berdasarkan penelitian iterasi stabil sampai iterasi 6 dengan hasil akhir kluster 1 ada sebanyak 295 data (data ke 4, 9, 10, 13, 17, 18, 19, 21, 22, 24, 27, 28, 29, 30, 31, 32, 33, 34, 36, 37, 38, 41, 44, 49, 51, 52, 53, 54, 55, 56, 58, 60, 61, 62, 63, 64, 72, 80, 81, 85, 88, 89, 90, 91, 92, 94, 95, 96, 97, 98, 99, 104, 108, 114, 115, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 164, 165, 169,

173, 188, 190, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 263, 264, 265, 266, 267, 268, 269, 270, 272, 273, 274, 275, 276, 277, 278, 279, 280, 282, 283, 284, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 332, 336, 337, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 362, 363, 364, 365, 366, 367, 369, 370, 371, 372, 373, 374, 375, 377, 378, 379, 380, 381, 382, 383, 385, 387, 388, 389, 391, 393, 395, 397, 398, 399, 400, 401) dan kluster 2 sebanyak 106 data (data ke 1, 2, 3, 5, 6, 7, 8, 11, 12, 14, 15, 16, 20, 23, 25, 26, 35, 39, 40, 42, 43, 45, 46, 47, 48, 50, 57, 59, 65, 66, 67, 68, 69, 70, 71, 73, 74, 75, 76, 77, 78, 79, 82, 83, 84, 86, 87, 93, 100, 101, 102, 103, 105, 106, 107, 109, 110, 111, 112, 113, 116, 148, 163, 166, 167, 168, 170, 171, 172, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 189, 191, 192, 213, 262, 271, 281, 285, 330, 331, 333, 334, 335, 338, 361, 368, 376, 384, 386, 390, 392, 394, 396) serta sebaran dengan *scatter plot* 3D dapat dilihat pada gambar 3.



Gambar 3. Sebaran K-Means dengan *Scatter Plot* 3D

f. Evaluasi DBI

Hasil nilai evaluasi menggunakan *Davies-Bouldin-Index* (DBI) sebagai berikut:

1) Perhitungan *Sum of Square Within* (SSW) Cluster

Berdasarkan centroid terakhir pada iterasi 6 yaitu tabel 7 dan banyak kluster 1 adalah 295 data, kluster 2 adalah 106 data, maka diproses selisih jarak antara masing-masing data dengan centroidnya. Jumlah jarak untuk K1 213736.31 dan K2 124337.91, sehingga sesuai persamaan 2 SSW nya adalah SSW

Klaster 1=(1/295)*213736.3=724.52 dan
SSW Klaster
2=(1/106)*124337.91=1172.99.

2) Perhitungan *Sum of Square Between* (SSB) *Cluster*

Perhitungan SSB dengan menghitung jarak (*euclidean*) antar centroid sebagai berikut.

$$\begin{aligned} \text{SSB}_{1,2} &= d(C_1, C_2) \\ &= \text{SQRT}((1172.44- \\ &3885.15)^2+(477.59- \\ &419.49)^2+(50.78-84.23)^2+(57.4- \\ &73.63)^2) \\ &= 2713.59 \end{aligned}$$

3) Perhitungan Nilai Rasio

Nilai rasio didapatkan dari jumlah SSW pada rasio yang dihitung tersebut dibagi dengan SSBnya:

$$\begin{aligned} R_{1,2} &= (\text{SSW}_1 + \text{SSW}_2) / \text{SSB}_{1,2} \\ &= 724.52/1172.99 = 0.70 \end{aligned}$$

4) *Davies-Bouldin-Index* (DBI)

Nilai DBI didapatkan dari rata-rata seluruh nilai R Max, perhitungan DBInya sebagai berikut:

$$\begin{aligned} \text{DBI} &= (R_{1\text{Max}} + R_{2\text{Max}}) = (0.70 + \\ &0.70) = 0.70 \end{aligned}$$

SIMPULAN

Hasil analisa dan komputasi dari metode *K-Means* untuk hasil pertanian desa di Kabupaten Bojonegoro tahun 2017 dengan 2 klaster dapat disimpulkan bahwa klaster baik sebanyak 295 data dan klaster tidak baik sebanyak 106 data. Untuk nilai evaluasi DBI dari hasil klasterisasi tersebut adalah 0.70 dimana nilai DBI mendekati 1 adalah indikator kualitas klaster semakin baik. Dalam penelitian ini, perhitungan jarak antara data dan *centroid* hanya diuji menggunakan *Euclidean* sehingga diperlukan uji metode jarak yang lain sebagai pembandingan seperti manhattan dan minkowski.

DAFTAR PUSTAKA

- [1] Undang-Undang No.19 Tahun 2013 tentang Perlindungan & Pemberdayaan Petani.

- [2] BPS Kabupaten Bojonegoro. "Bojonegoro dalam angka 2018". Bojonegoro : Badan Pusat Statistik. 2018.

- [3] J. O. Ong, "Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing President University," *Jurnal Ilmiah Teknik Industri*, vol. 12, no. 1, pp. 10 - 20, 2013.

- [4] Ardilla, Y. H. Tjandrasa dan I. Arieshanti, "Deteksi Penyakit Epilepsi dengan Menggunakan Entropi Permutasi, K-Means Clustering, dan Multilayer Perceptron," *Jurnal Teknik POMITS* , vol. 3, no. 1, pp. A70 - A74, 2014.

- [5] Agustina, S. D. Yhudo, H. Santoso, N. Marnasusanto, A. Tirtana dan F. Khusnu, "Clustering Kualitas Beras Berdasarkan Ciri Fisik Menggunakan Metode K-Means," Universitas Brawijaya Malang, Malang, 2012.

- [6] Dataset : https://drive.google.com/open?id=1FuUKlCfo9P_KyRRw2R_J60oUHZigxS6r