

Klasterisasi Dana Bantuan Pada Program Keluarga Harapan (PKH) Menggunakan Metode K-Means

Said Abdul Azis^{1✉}, Sarjon Defit², Yuhandri Yunus³

^{1,2,3}Universitas Putra Indonesia YPTK Padang

said.azis@gmail.com

Abstract

The Family of Hope Program (PKH) is a program that aims to reduce poverty and improve the quality of human resources. Optimizing the provision of assistance in accordance with the expectations of those in need. Data on the poor or integrated social welfare data is needed as a reference for grouping. This study aims to make it easier for the selection team to provide assistance in accordance with the predetermined criteria whether or not they deserve to receive the assistance. The data used in the study is data from 2019. The data processing in this study uses the K-Means Clustering method with 3 clusters, namely Cluster 1 (C1) Nearly Poor Households (RTHM), Cluster 2 (C2) Poor Households (RTM), Cluster 3 (C3) Very Poor Households (RTSM). The results of the clustering process with 2 iterations state that for Cluster 1 the amount of data is, for Cluster 2 the amount of data, and for Cluster 3 the amount of data. So this research is very helpful in relocating targeted assistance according to the family hope cluster.

Keywords: Data Mining, K-Means, Clustering, Hope Family Program, Poverty.

Abstrak

Program Keluarga Harapan (PKH) merupakan program yang bertujuan untuk mengurangi angka kemiskinan dan meningkatkan kualitas sumber daya manusia. Mengoptimalkan pemberian bantuan sesuai dengan harapan masyarakat yang membutuhkan. Data-data masyarakat miskin atau data terpadu kesejahteraan sosial diperlukan sebagai acuan untuk melakukan pengelompokan. Penelitian ini bertujuan untuk mempermudah tim penyeleksi memberikan bantuan sesuai dengan kriteria yang telah ditentukan apakah layak atau tidak layak menerima bantuan tersebut. Data yang digunakan dalam penelitian adalah data tahun 2019. Pengolahan data dalam penelitian ini menggunakan metode K-Means Clustering dengan jumlah 3 Cluster yaitu Cluster 1 (C1) Rumah Tangga Hampir Miskin (RTHM), Cluster 2 (C2) Rumah Tangga Miskin (RTM), Cluster 3 (C3) Rumah Tangga Sangat Miskin (RTSM). Hasil dari proses clustering dengan 2 kali iterasi menyatakan bahwa untuk Cluster 1 berjumlah data, untuk Cluster 2 berjumlah data, dan untuk Cluster 3 berjumlah data. Maka penelitian ini sangat membantu dalam merelokasikan bantuan yang tepat sasaran sesuai klaster keluarga harapan.

Kata kunci: Data Mining, K-Means, Clustering, Program Keluarga Harapan, Kemiskinan.

© 2021 INFEB

1. Pendahuluan

Kabupaten Indragiri Hilir merupakan salah satu daerah hasil pemekaran dengan pelaksanaannya terhitung pada tanggal 20 November 1965. Kabupaten Indragiri Hilir berada di Pulau Sumatera yang memiliki luas daratan sekitar 11.605 km² dan memiliki luas perairan sekitar 7.207 km², dengan data luas daratan yang begitu besar, Kabupaten Indragiri Hilir berpenduduk sekitar 703.734 jiwa pada tahun 2017 yang tercatat pada Badan Pusat Statistik (BPS) Kabupaten Indragiri Hilir.

Kemiskinan bisa diartikan dengan kondisi seseorang yang tidak dapat memenuhi kebutuhan pokoknya. Kurangnya pendapatan akan mengakibatkan seseorang kualitas hidupnya rendah. Untuk mengurangi tingkat kemiskinan maka pemerintah Indonesia melalui Kementerian Sosial mengeluarkan Program Keluarga Harapan (PKH) program ini bertujuan untuk mengurangi angka kemiskinan, meningkatkan kualitas sumber daya manusia.

Untuk mengoptimalkan pemberian bantuan sesuai dengan harapan masyarakat yang membutuhkan, dengan adanya pengelompokan desa yang dimana tingkat kemiskinannya tinggi di Kabupaten Indragiri Hilir dalam penelitian ini menggunakan Metode Clustering, data-data masyarakat miskin atau data terpadu kesejahteraan sosial diperlukan sebagai acuan untuk melakukan pengelompokan, Data yang sudah dikelompokkan dengan menggunakan metode K-Means diharapkan akan mempermudah tim penyeleksi memberikan bantuan sesuai dengan kriteria yang telah ditentukan apakah layak atau tidak layak menerima bantuan tersebut.

Beberapa penelitian yang menggunakan algoritma *k-means clustering* diantaranya adalah Mar'i & Suprianto (2018) menggunakan algoritma *k-means* dalam pengklasteran *credit card* pembayaran tagihan dengan hasil penelitian didapati pengujian parameter PSO dengan nilai terbaik [1]. Penelitian Mahmuda dkk (2017) algoritma *k-means clustering* dalam pengklasteran pengunjung perpustakaan dengan hasil

penelitian didapati beberapa cluster pengunjung sesuai cluster yang diinginkan [2]. Penelitian Ridho & Kusuma (2018) algoritma k-means clustering dalam pendeteksi jaringan pada akses log dengan hasil yang memiliki keunggulan yang mendominasi untuk tiga cluster pengunjung [3]. Penelitian Putra & Wadisman (2018) algoritma k-means clustering dalam pengimplementasian pelanggan potensial dengan hasil keakuratan untuk kelompok pelanggan yang potensial [4].

Penelitian Widodo & Wahyuni (2017) algoritma k-means clustering dalam mengimplementasikan bidang sekripsi mahasiswa dengan hasil pengelompokan yang baik dan akurat serta mendapatkan cluster yang diinginkan [5]. Penelitian Rosmini dkk (2018) algoritma k-means clustering dalam pengimplementasian data aktivitas kelompok siswa dengan hasil dapat menentukan kelompok aktivitas siswa sesuai cluster yang diharapkan [6]. Penelitian Rustam dkk (2018) algoritma k-means clustering dalam pengoptimalan penyakit menular pada daerah endemik dengan hasil data rekam medis menunjukkan hasil yang lebih akurat dan performance yang lebih baik [7]. Penelitian Afriyanti dkk (2018) algoritma k-means clustering dalam klasifikasi produk retur dengan hasil performa akurasi sebesar 95,6%, precision sebesar 0,943, dan recall sebesar 0,956 baik [8].

Penelitian Dewata dkk (2018) algoritma k-means clustering untuk memprediksi pengambilan jurusan siswa SMA kelas X dengan hasil menunjukkan bahwa nilai-nilai dari mata pelajaran ekstrak dan psikotes akan mempengaruhi untuk menentukan penjurusan siswa baru [9]. Penelitian Fajrin & Maulida (2018) algoritma k-means clustering dalam analisa pembelian konsumen pada transaksi penjualan dengan hasil bahwa dengan penerapan algoritma k-means clustering didapati penjualan spare part mana yang paling banyak terjual [10]. Penelitian Fatmawati & Windarto (2018) algoritma k-means clustering dalam penerapan rapidminer untuk daerah terjangkau daerah terjangkit demam berdarah dengan hasil bahwa terdapat 4 provinsi dengan cluster tingkatan tinggi (C1), 13 provinsi dengan cluster tingkatan sedang (C2), dan 17 provinsi dengan cluster tingkatan rendah (C3) [11]. Penelitian Gustientiedina dkk (2019) algoritma k-means clustering dalam clustering data obat-obatan pada RSUD Pekan Baru dengan hasil kelompok obat yang termasuk pemakaian sedikit rata-rata permintaan obat setiap tahunnya kurang dari 18000 buah, dan obat yang termasuk pemakaian sedang rata-rata permintaan obat setiap tahunnya diantara 18000-70000 buah, sedangkan obat yang masuk kedalam kelompok obat yang pemakaian tinggi rata-rata permintaan obat setiap tahunnya diatas 70000 buah [12].

Penelitian Maulida (2018) algoritma k-means clustering dalam pengelompokan kunjungan wisata dengan hasil pengelompokan C1 = Taman Impian Jaya Ancol, C2 = Taman Mini Indonesia Indah dan Kebon Binatang Ragunan dan C3 = Monumen Nasional, Museum

Nasional, Museum Satria Mandala, Museum Sejarah Jakarta dan Pelabuhan Sunda Kelapa. Hasil pengelompokan C3 menjadi catatan bagi pemerintah Provinsi DKI Jakarta [13]. Penelitian Silalahi (2018) algoritma k-means clustering untuk penjualan produk pada PT Batamas Niaga Jaya dengan hasil Kelompok C0 produk yang tidak laris terdiri dari 657 produk dan kelompok C1 terdiri dari 70 produk yang merupakan produk yang laris [14]. Penelitian Nur dkk (2017) algoritma k-means clustering untuk clustering jurusan siswa baru dengan hasil diperoleh tiga kelompok yaitu kelompok tidak lulus, kelompok rekayasa perangkat lunak dan kelompok teknik komputer jaringan. Terdapat pusat cluster dengan cluster-1 = 1.4;2.2;2.2, cluster2 = 2.28;1.64;4 dan cluster-3 = 5;2;6 [15]. Penelitian Parlina dkk (2018) algoritma k-means clustering dalam menentukan pegawai yang layak mengikuti assessment center dengan hasil dapat diambil pengelompokan dengan rata-rata data program SDP yang dapat melakukan assessment center lanjutan adalah yang lolos dan hasil klasifikasi program SDP yang hampir lolos harus memperbaiki administrasi seperti kedisiplinan dari bulan juni sampai bulan oktober agar dapat mengikuti assessment center lanjutan, sedangkan hasil klasifikasi daftar data program SDP yang Tidak lolos harus memperbaiki data kedisiplinannya selama 1 Tahun [16].

2. Metodologi Penelitian

Objek penelitian ini menggunakan data yang diperoleh dari laporan tahunan data perkebunan pada Dinas Sosial Kabupaten Indragiri Hilir tahun 2019.

Knowledge Discovery in Database (KDD) merupakan proses non-trivial yang berguna untuk mencari dan mengidentifikasi pola di dalam data, dimana pola yang ditemukan tersebut bersifat benar, agar dapat bermanfaat dan dapat dipahami. KDD juga dapat diartikan sebagai pengorganisasian proses yang bersifat otomatis untuk mengidentifikasi yang benar, yang mana proses tersebut dapat berguna dan dapat menemukan pola dari kumpulan data yang besar (big data) dan data kompleks [17]. Data mining dapat diartikan sebagai teknologi baru yang berguna untuk membantu pihak perusahaan untuk menemukan informasi yang sangat penting dari gudang data mereka. Beberapa aplikasi data mining berfokus pada suatu prediksi, meramalkan apa yang akan terjadi pada situasi baru dari data yang menerangkan kejadian masa lalu.

Clustering merupakan sebuah metode pengelompokan data berdasarkan informasi yang didapatkan dari data yang menguraikan objek tersebut serta jalinan diantaranya yang digunakan untuk mengklasifikasi data dengan cara otomatis untuk menghasilkan beberapa kelompok yang diukur menggunakan asosiasi, agar menghasilkan data yang serupa berada pada kelompok data yang sama serta data yang berbeda berada pada kelompok data yang berbeda pula. Data masukan untuk *cluster* merupakan perangkat data yang memiliki kesamaan atau perbedaan ukuran antara dua data.

Sedangkan untuk data keluaran dari data *cluster* adalah sejumlah kelompok data berbentuk partisi atau kelompok struktur partisi dari sekumpulan data. Hasil tambahan dari data *cluster* adalah gambaran umum dari setiap kelompok yang memiliki peran penting.

Algoritma *k-means* merupakan suatu algoritma yang mengelompokkan iteratif dengan tujuan untuk mempartisi set data pada sejumlah *K clustering* yang telah ditetapkan pada awal pengolahan. Algoritma *k-means* bersifat sederhana dalam pengimplementasian dan menjalankan, pengerjaan relatif cepat, dapat juga dapat beradaptasi dan sudah sering digunakan. Langkah-langkah yang terdapat pada Algoritma *K-means* yaitu dengan menentukan nilai *K* sebagai jumlah *cluster* yang akan dibuat, selanjutnya menentukan titik pusat *cluster* (*K centroid*) awal secara acak, setelah mendapatkan titik pusat *cluster* langkah selanjutnya ialah menghitung jarak setiap data ke masing-masing *centroid* dengan memanfaatkan rumus korelasi antar dua objek, seperti yang dilihat pada rumus *Euclidean Distance* berikut ini.

$$D_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + \dots (n_i - n_j)^2} \quad (1)$$

Dimana:

d_{ij} = Jarak antara *i* dan *j*;

x_i = Koordinat *x* objek;

x_j = Koordinat *x* pusat;

y_i = Koordinat *y* objek;

y_j = Koordinat *y* pusat.

Setelah mendapatkan hasil dari perhitungan jarak setiap data, langkah selanjutnya menggabungkan setiap data yang dilihat berdasarkan jarak terdekat antara data dan *cluster*. Langkah selanjutnya yaitu menentukan tempat

centroid baru dengan menghitung nilai rata-rata berdasarkan data pada *cluster* yang serupa, untuk menentukan tempat *centroid* baru untuk menghitung jarak antara data dan *centroid* dapat menggunakan rumus berikut ini.

$$C_i = (x_1 + x_2 + x_3 + \dots + x_n) / (\sum x) \quad (2)$$

Dimana:

x_1 = nilai record data ke 1;

x_2 = nilai record data ke 2;

$\sum x$ = jumlah record data.

Algoritma *k-means* merupakan algoritma yang dilakukan secara berulang-ulang hingga menghasilkan *cluster* sesuai kelompok pada data-data yang sama, dan yang berbeda di tempatkan pada kelompok yang berbeda.

Adapun kelebihan yang terdapat pada algoritma *K-means* adalah sebagai berikut.

- Sangat mudah dalam pengimplementasian dan proses menjalankan.
- Kebutuhan waktu yang sangat singkat dalam menjalankan pembelajaran.
- Mudah untuk beradaptasi.
- Sangat umum digunakan.

3. Hasil dan Pembahasan

Data yang diambil merupakan data penyaluran dana bantuan yang didapat dari Dinas Sosial Kabupaten Indragiri Hilir melalui bidang sekretariat. Data disajikan pada Tabel 1.

Tabel 1. Data Uji Penyaluran Dana Bantuan Program Keluarga Harapan

No	Kecamatan/Kelurahan	Nama Desa	Tahap Penyaluran				Grand Total
			Tahap 1	Tahap 2	Tahap 3	Tahap 4	
1		Gemilang Jaya	75	74	74	70	293
2		Kuala Sebatu	111	111	111	107	440
3		Pasir Emas	26	26	26	26	104
4		Sialang Jaya	43	41	41	41	166
5	Batang Tuaka	Simpang Jaya	83	83	79	77	322
6		Sungai Dusun	80	78	78	77	313
7		Sungai Junjangan	147	147	144	142	580
8		Sungai Luar	126	126	122	122	496
9		Sungai Piring	207	207	205	196	815
10		Sungai Rawa	156	157	157	158	628

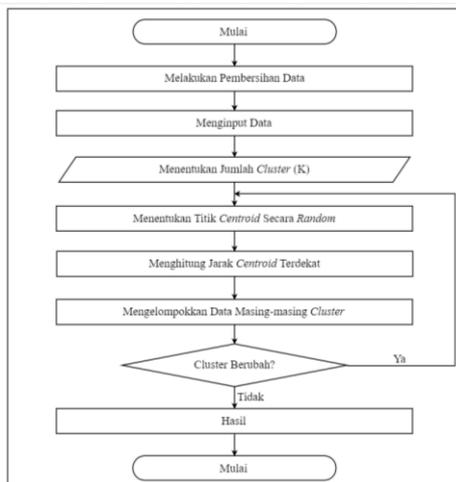
Pada penelitian ini peneliti menggunakan analisa sistem dengan metode *k-means* untuk klaterisasi penyaluran dana bantuan Program Keluarga Harapan (PKH) berdasarkan empat tahapan dengan menjadi 3 kelompok, yaitu kelompok C1 merupakan kelompok kategori Rumah Tangga Hampir Miskin (RTHM), C2 merupakan kelompok kategori Rumah Tangga Miskin (RTM), dan C3 merupakan kelompok kategori Rumah Tangga Sangat Miskin (RTSM). Berikut ini merupakan Kesimpulan langkah-langkah yang ditetapkan pada sistem kerja algoritma *k-means*, yaitu:

- Melakukan pembersihan data;

- Menginput data;
- Menentukan jumlah *cluster* (*K*);
- Menentukan titik *centroid* secara random;
- Menghitung jarak *centroid* terdekat;
- Mengelompokkan data masing-masing *cluster*;
- Hasil *cluster*.

Dalam penerapan algoritma *k-means* yang diterapkan pada penelitian ini, maka akan digambarkan dalam bentuk *flowchart* agar terlihat lebih jelas dan terarah.

Flowchart untuk analisa sistem pada penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Flowchart Proses K-means

3.1. Pembersihan Data

Tahapan ini merupakan suatu proses pengoreksian, pengurangan data sebelumnya atau membuang sebagian data yang ganda, menghilangkan data yang kotor, serta memverifikasi data yang dianggap tidak konsisten dan memverifikasi kesalahan yang ada pada data. Hasil pembersihan dapat dilihat pada Tabel 2.

Tabel 2. Sample Pembersihan Data

No	Nama Desa	Tahap 1	Tahap 2	Tahap 3	Tahap 4
1	Gemilang Jaya	75	74	74	70
2	Kuala Sebatu	111	111	111	107
3	Pasir Emas	26	26	26	26
4	Sialang Jaya	43	41	41	41
5	Simpang Jaya	83	83	79	77
6	Sungai Dusun	80	78	78	77
7	Sungai Junjangan	147	147	144	142
8	Sungai Luar	126	126	122	122
9	Sungai Piring	207	207	205	196
10	Sungai Rawa	156	157	157	158

3.2. Menginput

Proses penginputan data akan dilaksanakan kedalam sistem setelah melakukan pembersihan data (*data cleaning*). Data yang digunakan dalam penginputan pada penelitian ini adalah data traning yang sudah dilakukan pembersihan yang diperoleh dari Dinas Sosial Kabupaten Indragiri Hilir dengan jumlah sampel yang akan digunakan hanya sebanyak 10 data.

3.3. Menentukan Jumlah Cluster

Proses penerapan metode *k-means* harus menentukan terlebih dahulu berapa banyak jumlah *cluster* yang diinginkan, dimana nantinya *cluster* tersebut akan menjadi acuan pada hasil pengambilan keputusan. Pada proses penelitian ini jumlah *cluster* akan dibagi menjadi 3 *cluster* dimana nantinya ke-3 *cluster* ini akan menentukan hasil untuk penelitian mengenai daerah kategori Rumah Tangga Hampir Miskin (RTHM), C2 merupakan kelompok kategori Rumah Tangga Miskin (RTM), dan C3 merupakan kelompok kategori Rumah Tangga Sangat Miskin (RTSM) untuk penyaluran dana bantuan di daerah Kabupaten Indragiri Hilir.

3.4. Menentukan Titik Centroid Acak

Pada penentuan *centroid* awal dalam perhitungan pada penelitian ini dilaksanakan dengan cara random, dengan data *centroid* disajikan pada Tabel 3.

Tabel 3. Sample Centroid Acak

Centroid	Tahap 1	Tahap 2	Tahap 3	Tahap 4
C1	75	74	74	70
C2	126	126	122	122
C3	207	207	205	196

3.5. Menghitung Jarak Centroid terdekat

Rumus yang digunakan untuk menghitung jarak setiap data ke masing-masing *centroid* menggunakan rumus korelasi antara tiga objek (*euclidean distance*). Sehingga dengan rumus ini akan ditemukan jarak yang paling terdekat dari setiap data dengan masing-masing *centroid*. Proses perhitungan jarak masing-masing data ke titik pusat *cluster* berikut ini:

1. Proses Iterasi 1

a. Proses menghitung cluster 1 (C1)

$$D1 = \sqrt{(75-75)^2 + (74-74)^2 + (74-74)^2 + (70-70)^2}$$

$$= 0$$

b. Proses menghitung cluster 2 (C2)

$$D1 = \sqrt{(75-126)^2 + (74-126)^2 + (74-122)^2 + (70-70)^2}$$

$$= 102$$

c. Proses menghitung cluster 3 (C3)

$$D1 = \sqrt{(75-207)^2 + (74-207)^2 + (74-205)^2 + (70-196)^2}$$

$$= 261$$

3.6. Pengelompokan Data Masing-Masing Cluster

Tabel 4. Pengelompokan Data Uji

No	Nama Desa	Tahap1	Tahap2	Tahap3	Tahap4	C1	C2	C3	C0	C1	C2
1	Gemilang Jaya	75	74	74	70	0	102	261	1		
2	Kuala Sebatu	111	111	111	107	74	28	188		1	
3	Pasir Emas	26	26	26	26	95	196	356	1		
4	Sialang Jaya	43	41	41	41	64	165	325	1		
5	Simpang Jaya	83	83	79	77	15	87	247	1		
6	Sungai Dusun	80	78	78	77	10	92	251	1		
7	Sungai Junjangan	147	147	144	142	144	42	118		1	
8	Sungai Luar	126	126	122	122	102	0	160		1	
9	Sungai Piring	207	207	205	196	261	160	0			1
10	Gemilang Jaya	75	74	74	70	168	66	94		1	
Total									5	4	1

Tabel 4 diatas merupakan hasil *clusterisasi* dari masing-masing data sample penelitian, dimana untuk *cluster 1* berjumlah 5 data, *cluster 2* berjumlah 4 data, dan *cluster 3* berjumlah 1 data.

Selanjutnya untuk menghitung iterasi ke-2 harus melalui perhitungan dari menghitung rata-rata hasil data yang sama pada *cluster* yang posisi sama, untuk perhitungan mencari *centroid* baru sebagai berikut:

a. *Cluster 1* (C1) berjumlah 5 data

$$D1 = (75+26+43+83+80)/5 = 61,4$$

$$D2 = (74+26+41+83+78)/5 = 60,4$$

$$D3 = (74+26+41+79+78)/5 = 59,6$$

$$D4 = (70+26+41+77+77)/5 = 58,2$$

b. *Cluster 2* (C2) berjumlah 4 data

$$D1 = (111+147+126+156)/5 = 135$$

$$D2 = (111+147+126+157)/5 = 135,25$$

$$D3 = (111+144+122+157)/5 = 133,5$$

$$D4 = (107+142+122+158)/5 = 132,25$$

c. *Cluster 3* (C3) berjumlah 1 data

$$D1 = (207)/1 = 207$$

$$D2 = (207)/1 = 207$$

3.8. Pengelompokan Data Masing-Masing Cluster

$$D3 = (205)/1$$

$$= 205$$

$$D4 = (196)/1$$

$$= 196$$

Setelah perhitungan selesai, didapat *centroid* baru untuk perhitungan iterasi selanjutnya, proses ini bertujuan untuk memastikan hasil perhitungan sebelumnya masih terdapat perubahan posisi data *cluster* atau tidak, dimana data *centroid* baru yang disajikan pada Tabel 5.

Tabel 5. *Centroid* Baru

<i>Centroid</i>	Tahap 1	Tahap 2	Tahap 3	Tahap 4
C1	61,4	60,40	59,6	58,20
C2	135,0	135,25	133,5	132,25
C3	207,0	207,00	205	196,00

3.7. Menghitung Jarak *Centroid* Baru

Rumus yang digunakan untuk menghitung jarak setiap data ke masing-masing *centroid* menggunakan rumus korelasi antara tiga objek (*euclidean distance*). Sehingga dengan rumus ini akan ditemukan jarak yang paling terdekat dari setiap data dengan masing-masing *centroid*. Data yang dirunakan untuk *centroid* baru didapat dari hasil perhitungan sebelumnya. Proses perhitungan jarak masing-masing data ke titik pusat *cluster* berikut ini:

2. Proses Iterasi 2

a. Proses menghitung *cluster 1* (C1)

$$D1 = \sqrt{(75-61,4)^2 + (74-61,4)^2 + (74-59,6)^2 + (70-58,2)^2} = 27$$

b. Proses menghitung *cluster 2* (C2)

$$D1 = \sqrt{(75-135)^2 + (74-135)^2 + (74-133)^2 + (70-132)^2} = 122$$

c. Proses menghitung *cluster 3* (C3)

$$D1 = \sqrt{(75-207)^2 + (74-207)^2 + (74-205)^2 + (70-196)^2} = 261$$

Setelah proses perhitungan dilakukan pada tahap sebelumnya, maka jarak dari masing-masing data *cluster* 1 (C1), *cluster* 2 (C2), *cluster* 3 (C3) dapat ditampilkan. Hasil perhitungan seluruh data sample penelitian disajikan pada Tabel 6.

Tabel 6. Sample Pengelompokan Data

No	Nama Desa	Tahap1	Tahap2	Tahap3	Tahap4	C1	C2	C3	C0	C1	C2
1	Gemilang Jaya	75	74	74	70	27	122	261	1		
2	Kuala Sebatu	111	111	111	107	100	54	225		1	
3	Pasir Emas	26	26	26	26	68	199	290	1		
4	Sialang Jaya	43	41	41	41	37	188	289	1		
5	Simpang Jaya	83	83	79	77	41	132	253	1		
6	Sungai Dusun	80	78	78	77	37	137	256	1		
7	Sungai Junjangan	147	147	144	142	170	42	182		1	
8	Sungai Luar	126	126	122	122	128	20	226		1	
9	Sungai Piring	207	207	205	196	288	195	84			1
10	Gemilang Jaya	75	74	74	70	194	73	142		1	
Total									5	4	1

Tabel 6 diatas merupakan hasil dari pengelompokan yang dilakukan pada proses iterasi ke-2 dimana dari hasil yang diperoleh dapat disimpulkan bahwa data tersebut tidak didapati perubahan untuk setiap *cluster*. Dengan demikian proses iterasi hanya dilakukan sampai iterasi ke-2 dengan hasil untuk *cluster* 1 berjumlah 5 data desa, *cluster* 2 berjumlah 4 data desa, dan untuk *cluster* 3 berjumlah 1 data desa, dengan demikian proses iterasi dihentikan. Hasil untuk setiap *cluster* dari penelitian ini dapat dilihat berikut:

Tabel 7. Hasil *cluster* 1

No	Nama Desa	Tahap 1	Tahap 2	Tahap 3	Tahap 4
1	Gemilang Jaya	75	74	74	70
2	Pasir Emas	26	26	26	26
3	Sialang Jaya	43	41	41	41
4	Simpang Jaya	83	83	79	77
5	Sungai Dusun	80	78	78	77

Tabel 7 diatas menunjukkan hasil pengelompokan untuk kategori Rumah Tangga Hampir Miskin (RTHM) di Kabupaten Indragiri Hilir pada *cluster* 1 berjumlah 5 Desa.

Tabel 8. Hasil *cluster* 2

No	Nama Desa	Tahap 1	Tahap 2	Tahap 3	Tahap 4
1	Kuala Sebatu	111	111	111	107
2	Sungai Junjangan	147	147	144	142
3	Sungai Luar	126	126	122	122
4	Sungai Rawa	156	157	157	158

Tabel 8 diatas menunjukkan hasil pengelompokan untuk kategori Rumah Tangga Miskin (RTM) di Kabupaten Indragiri Hilir pada *cluster* 2 berjumlah 4 Desa.

Tabel 9. Hasil *cluster* 3

No	Nama Desa	Tahap 1	Tahap 2	Tahap 3	Tahap 4
1	Sungai Piring	207	207	205	196

Tabel 9 diatas menunjukkan hasil pengelompokan untuk kategori Rumah Tangga Sangat Miskin (RTSM) di Kabupaten Indragiri Hilir pada *cluster* 3 berjumlah 1 Desa.

4. Kesimpulan

Berdasarkan hasil dari penelitian mengenai klasterisasi dana bantuan pada Program Keluarga Harapan (PKH) menggunakan *K-means clustering* menghasilkan kanperhitungan yang dilakukan dalam pengelompokan menggunakan 10 data desa menunjukkan bahwa untuk (C1) sebanyak 5 data desa, untuk (C2) sebanyak 4 data desa, dan untuk (C3) sebanyak 1 data desa. Sehingga penelitian ini sangat membantu dalam pengelompokan secara tepat.

Daftar Rujukan

- [1] Mar'i, F., & Supianto, A. A. (2018). Clustering Credit Card Holder Berdasarkan Pembayaran Tagihan Menggunakan Improved K-means dengan Particle Swarm Optimization. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 5(6), 737-744. DOI: <http://dx.doi.org/10.25126/jtiik.201856858> .
- [2] Mahmuda, F., Sitorus, M. A. R., Widyastuti, H., & Kurniawan, D. E. (2017). Clustering Profil Pengunjung Perpustakaan Menggunakan Algoritma K-means. *Journal of Applied Informatics and Computing*, 1(1), 14-21. DOI: <https://doi.org/10.30871/jaic.v1i1.476> .
- [3] Ridho, F., & Kusuma, A. A. (2019). Deteksi Intrusi Jaringan dengan *K-means Clustering* Pada Akses Log dengan Teknik Pengolahan Big Data. *Jurnal Aplikasi Statistika dan Komputasi Statistik*, 10(1), 53-66. DOI: <https://doi.org/10.34123/jurnalasks.v10i1.202> .
- [4] Putra, R. R., & Wadisman, C. (2018). Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K Means. *INTECOMS: Journal of Information Technology and Computer Science*, 1(1), 72-77. DOI: <https://doi.org/10.31539/intecom.v1i1.141> .
- [5] Widodo, W., & Wahyuni, D. (2017). Implementasi Algoritma *K-Means Clustering* untuk Mengetahui Bidang Skripsi Mahasiswa Multimedia Pendidikan Teknik Informatika dan Komputer Universitas Negeri Jakarta. *PINTER: Jurnal Pendidikan Teknik Informatika dan Komputer*, 1(2), 157-166. DOI: <https://doi.org/10.21009/pinter.1.2.10> .
- [6] Rosmini, R., Fadlil, A., & Sunardi, S. (2018). Implementasi Metode *K-means* dalam Pemetaan Kelompok Mahasiswa Melalui Data Aktivitas Kuliah. *IT Journal Research and Development*, 3(1), 22-31. DOI: [https://doi.org/10.25299/itjrd.2018.vol3\(1\).1773](https://doi.org/10.25299/itjrd.2018.vol3(1).1773) .
- [7] Rustam, S., Santoso, H. A., & Supriyanto, C. (2018). Optimasi *K-Means Clustering* untuk Identifikasi Daerah Endemik Penyakit Menular dengan Algoritma Particle Swarm Optimization di Kota Semarang. *ILKOM Jurnal Ilmiah*, 10(3), 251-259. DOI: <https://dx.doi.org/10.33096/ilkom.v10i3.342.251-259> .

- [8] Arifiyanti, A. A., Pradana, R. M., & Novian, I. F. (2018). Klasifikasi Produk Retur dengan Algoritma Pohon Keputusan C4. 5. *Jurnal IPTEK*, 22(1), 79-86. DOI: <https://doi.org/10.31284/j.iptek.2018.v22i1.243> .
- [9] Fajrin, A. A., & Maulana, A. (2018). Penerapan Data Mining untuk Analisis Pola Pembelian Konsumen dengan Algoritma Fp-Growth Pada Data Transaksi Penjualan Spare Part Motor. *Kumpulan Jurnal Ilmu Komputer (KLIK)*, 5(1), 27-36.
- [10] Fatmawati, K., & Windarto, A. P. (2018). Data Mining: Penerapan Rapidminer Dengan *K-means Cluster* Pada Daerah Terjangkit Demam Berdarah Dengue (DBD) Berdasarkan Provinsi. *Journal of Computer Engineering, System and Science*, 3(2), 173-178. DOI: <https://doi.org/10.24114/cess.v3i2.9661> .
- [11] Gustientiedina, G., Adiya, M. H., & Desnelita, Y. (2019). Penerapan Algoritma *K-means* Untuk *Clustering* Data Obat-Obatan. *Jurnal Nasional Teknologi dan Sistem Informasi*, 5(1), 17-24. DOI: <https://doi.org/10.25077/TEKNOSL.v5i1.2019.17-24>
- [12] Maulida, L. (2018). Penerapan Data Mining dalam Mengelompokkan Kunjungan Wisatawan ke Objek Wisata Unggulan di Prov. DKI Jakarta dengan *K-means*. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 2(3), 167-174. DOI: <http://dx.doi.org/10.14421/jiska.2018.23-06> .
- [13] Yahya, K. B., & Mahpuz, M. (2019). Penggunaan Algoritma *K-Means* Untuk Menganalisis Pelanggan Potensial Pada Dealer SPS Motor Honda Lombok Timur Nusa Tenggara Barat. *Infotek: Jurnal Informatika dan Teknologi*, 2(2), 109-118. DOI: <https://dx.doi.org/10.29408/jit.v2i2.1447> .
- [14] Nur, F., Zarlis, M., & Nasution, B. B. (2017). Penerapan Algoritma *K-Means* Pada Siswa Baru Sekolah Menengah Kejuruan Untuk *Clustering* Jurusan. *InfoTekJar: Jurnal Nasional Informatika dan Teknologi Jaringan*, 1(2), 100-105. DOI: <https://doi.org/10.30743/infotekjar.v1i2.70> .
- [15] Parlina, I., Windarto, A. P., Wanto, A., & Lubis, M. R. (2018). Memanfaatkan Algoritma *K-Means* dalam Menentukan Pegawai yang Layak Mengikuti Asessment Center untuk *Clustering* Program SDP. *Journal of Computer Engineering, System and Science*, 3(1), 87-93. DOI: <https://doi.org/10.24114/cess.v3i1.8192> .
- [16] Santoso, S., & Nurmalina, R. (2017). Perencanaan dan Pengembangan Aplikasi Absensi Mahasiswa Menggunakan Smart Card Guna Pengembangan Kampus Cerdas. *Jurnal Integrasi*, 9(1), 84-91. DOI: <https://doi.org/10.30871/ji.v9i1.288>
- [17] Ayu, F., & Permatasari, N. (2018). Perancangan Sistem Informasi Pengolahan Data PKL (Praktek Kerja Lapangan) Di Devisi Humas Pada PT Pegadaian. *Jurnal Intra Tech*, 2(2), 12-26.