

# PENERAPAN *CLUSTERING* DATA KURANG MAMPU DI DESA SITUMEKAR MENGGUNAKAN ALGORITMA *K-MEANS*

Feni Andrianti<sup>1</sup>, Ricky Firmansyah<sup>2</sup>

<sup>1</sup>Universitas Adhirajasa Reswara Sanjaya  
Jl. Sekolah Internasional No. 1-2 Antapani, Bandung, 022-7100124  
e-mail: [feniandrian5@gmail.com](mailto:feniandrian5@gmail.com)

<sup>2</sup>Universitas Adhirajasa Reswara Sanjaya  
Jl. Sekolah Internasional No. 1-2 Antapani, Bandung, 022-7100124  
e-mail: [rickyfirmanars@gmail.com](mailto:rickyfirmanars@gmail.com)

## Abstrak

Kemiskinan seringkali ditandai dengan tingginya tingkat pengangguran dan keterbelakangan. Salah satu aspek penting untuk mendukung strategi penanggulangan kemiskinan adalah dengan tersedianya data kemiskinan yang akurat. Untuk mencegah kesalahan pihak desa dalam menentukan penduduk yang berhak menerima bantuan maka dapat dilakukan pencegahan dengan menggunakan teknik *data mining*. *Data mining* merupakan proses informasi yang menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. *Clustering* merupakan metode salah satu teknik *data mining* yang berfungsi mengelompokkan data secara otomatis tanpa memerlukan label pada data latih. Algoritma yang digunakan untuk pembentukan *cluster* adalah Algoritma *K-Means*. *K-Means* adalah salah satu metode pengelompokkan data nonhierarki (sekatan) yang berusaha mempartisi data yang ada dalam ke dalam bentuk dua atau lebih kelompok. Data yang digunakan berjumlah 137 data, dengan atribut dusun, umur dan jumlah keluarga. *Cluster* yang terbentuk berjumlah 3 *cluster*, dengan *cluster* pertama berjumlah 29 penduduk, *cluster* kedua berjumlah 72 penduduk, dan *cluster* ketiga berjumlah 36 penduduk. Hasil dari penelitian dapat dijadikan sebagai salah satu referensi bagi pihak desa dalam mendukung keputusan untuk melakukan pengelompokkan penduduk untuk menerima bantuan.

**Kata Kunci** : Kemiskinan, *Data Mining*, *Clustering*, *K-Means*

## Abstract

Poverty is often characterized by high levels of unemployment and underdevelopment. One important aspect of supporting poverty reduction strategies is the availability of accurate poverty data. To prevent mistakes from the village in determining which people are entitled to receive assistance, prevention can be done using *data mining* techniques. *Data mining* is the process of interesting information in selected data using certain techniques or methods. *Clustering* is a method of *data mining* techniques that functions to group data automatically without requiring labels on training data. The *algorithm* used for *cluster* formation is the *K-Means Algorithm*. *K-Means* is a non-hierarchical data grouping method that seeks to partition existing data into two or more groups. The data used are 137 data, with the attributes of the hamlet, age and number of families. The *cluster* formed consists of 3 clusters, with the first cluster totaling 29 residents, the second *cluster* totaling 72 residents, and the third *cluster* totaling 36 residents. The results of the research can be used as a reference for the village in supporting the decision to group the population to receive assistance.

**Keywords**: Poverty, *Data Mining*, *Clustering*, *K-Means*

## 1. Pendahuluan

Kemiskinan merupakan permasalahan bangsa yang mendesak dan memerlukan langkah – langkah penanganan

dan pendekatan yang sistematis, terpadu dan menyeluruh dengan. Kemiskinan seringkali ditandai dengan tingginya tingkat pengangguran dan keterbelakangan. Salah satu aspek penting untuk mendukung

Strategi Penanggulangan kemiskinan dengan adanya data yang akurat. Pengukuran ini dipercaya menjadi instrumen tangguh bagi pengambil kebijakan memfokuskan perhatian pada kondisi hidup masyarakat. Data ini bisa digunakan untuk mengevaluasi kebijakan pemerintah terhadap kemiskinan, membandingkan kemiskinan antar waktu dan daerah dan menentukan target penduduk dengan tujuan memperbaiki kondisi mereka (Statistik, 2018). Upaya ini terdapat dua strategi yang harus dijalani. Pertama, melindungi keluarga dan kelompok masyarakat melalui pemenuhan kebutuhan mereka dari berbagai bidang. Kedua, melakukan pelatihan agar memiliki kemampuan untuk melakukan usaha. Upaya ini dilakukan agar mewujudkan cita-cita bangsa yaitu, terciptanya masyarakat yang adil dan makmur. Dengan penyajian data jumlah dan persentase penduduk miskin, informasi ini sangat dibutuhkan oleh kebijakan untuk penanganan masalah kemiskinan (Ferezagia, 2018). Secara umum, hasil penelitian menunjukkan bahwa desa yang memiliki tingkat perkembangan wilayah yang tinggi mempunyai tingkat kemiskinan yang rendah. Sementara pertumbuhan ekonomi tidak berpengaruh terhadap kemiskinan, hal ini menunjukkan bahwa pertumbuhan yang ada kurang berkualitas sehingga tidak mempengaruhi kemiskinan (Zuhdiyati & Kaluge, 2015).

Desa Situmekar adalah sebuah desa yang pemekaran dengan Desa Cisu pada tahun 1982. Lalu dimekarkan pusat pemerintahan Desa Situmekar tidak berpindah dari Dusun Corenda yang berada di perlintasan jalan Provinsi yang menghubungkan Sumedang dan Garut. Desa Situmekar berada di wilayah Kecamatan Cisu, di mana kantor kecamatan berada di Desa Situmekar. Desa situmekar salah satu desa yang sebagian masyarakatnya hidup dalam kondisi kurang mampu dan sebagian besar pekerjaan mereka adalah sebagai buruh tani.

*K-Means* adalah metode data yang mempartisi data ke dalam bentuk satu atau lebih *cluster* sehingga data yang mempunyai karakteristik yang sama dikelompokkan ke dalam satu *cluster* yang sama yang memiliki karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain (Metisen & Sari, 2015). *K-Means clustering* termasuk salah satu metode *clustering* data mining yang menggunakan konsep *descriptive*

model. Metode *K-means* digunakan untuk menjelaskan suatu algoritma dalam penentuan suatu objek kedalam klaster tertentu berdasarkan rataan terdekat. Algoritma *K-Means* dimulai dengan pemilihan secara acak *K*, *K* disini merupakan banyaknya *cluster* yang ingin dibentuk. Lalu tetapkan nilai *K* secara acak, sementara nilai itu menjadi pusat dari *cluster* atau biasa disebut dengan centroid, mean atau "*means*". Hitung jarak tiap data yang ada, masing – masing centroid menggunakan rumus *Euclidean* hingga ditemukan jarak yang paling dekat dari setiap data dengan centroid. Klasifikasikan berdasarkan kedekatannya dengan centroid (*distance space*) (Z & Sarjono, 2016).

*Data Mining* merupakan proses mencari informasi yang menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. *Data Mining* merupakan beberapa ilmu yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, *database* dan visualisasi untuk penanganan permasalahan pengambilan informasi dari database yang besar (Mardi, 2017).

Ada beberapa penelitian sebelumnya yang telah dilakukan oleh peneliti lain mengenai bantuan kurang mampu akan tetapi menggunakan algoritma yang lain seperti pada penelitian (Annur, 2018) dengan metode Algoritma Naive Bayes, lalu pada penelitian yang menggunakan algoritma *K-Means Clustering* dan dataset yang sama yaitu penduduk kurang mampu, akan pada penelitian (Febianto & Palasara, 2019). Sedangkan penelitian yang menggunakan algoritma yang sama tetapi dengan dataset yang digunakan berbeda seperti algoritma *K-Means Clustering* pada penelitian (Sari, Wanto, & Windarto, 2018). Algoritma *Fuzzy C-Means* dan *K-Means* pada penelitian (Febrianti, Hafiyusholeh, & Asyhar, 2016). Algoritma *K-Means Clustering* pada penelitian (Metisen & Sari, 2015).

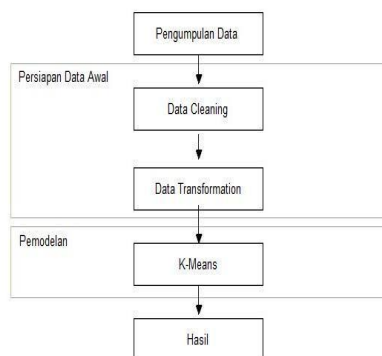
Berdasarkan latar belakang di atas maka penulis tertarik untuk membuat penelitian yang berjudul "**Penerapan Clustering Data Penduduk Kurang Mampu Di Desa Situmekar Menggunakan Algoritma K Means**".

## 2. Metodologi Penelitian

### 2.1. Desain Penelitian

Pada bagian ini, akan dijelaskan gambaran mengenai metodologi penelitian secara keseluruhan. Berikut adalah tahapan

yang dilakukan peneliti dalam melakukan penelitian.



**Gambar 1 Desain Penelitian**

1. **Pengumpulan Data**  
Sebuah langkah yang strategis dalam penelitian, karena memiliki tujuan utama dalam mendapatkan data.
2. **Data Cleaning**  
Sebelum proses *data mining* perlu dilakukan proses *cleaning* pada data yang menjadi focus KDD. Proses *cleaning* antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (*tipografi*). Lalu dilakukan proses *enrichment* yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD. Seperti data atau informasi eksternal.
3. **Data Transformation**  
proses tranformasi pada data yang telah dipilih sehingga data tersebut sesuai untuk proses *data mining*. Proses coding merupakan proses kreatif dan bergantung pada jenis atau informasi yang akan dicari dalam basis data.
4. **K-Means**  
*K-means* merupakan algoritma yang mempartisi data ke dalam bentuk satu atau lebih *cluster* sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu *cluster* yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain.

## 2.2. Instrumen Penelitian

Instrumen Penelitian adalah sebuah alat digunakan untuk mengumpulkan data atau mengukur objek dari suatu variabel penelitian. Untuk mendapatkan data yang benar dan sesuai dengan keadaan

sebenarnya, maka dibutuhkan suatu instrumen yang valid dan konsisten dalam memberikan data hasil penelitian (reliabel) (Yusup, 2018). Pada penelitian ini akan digunakan beberapa penelitian, antara lain sebagai berikut :

1. **Perangkat Lunak (Software)**  
Perangkat lunak yang akan digunakan dalam penelitian ini adalah *software rapidminer* versi 9.6
2. **Dataset**  
Penelitian ini menggunakan *dataset* data Penduduk Kurang Mampu Di Desa Situmekar tahun 2020 sebanyak 137 penduduk yang telah melalui proses *cleaning* dan *transformation* terdiri dari atribut nama penduduk, umur, jumlah keluarga, dusun.

### 2.3. Metode Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data *public* yang diambil langsung dari Desa Situmekar dengan cara melakukan riset kepada bagian sekretaris desa.

#### A. Sampel Data

Sampel data yang digunakan untuk penelitian ini adalah data penduduk Desa Situmekar sebanyak 137 data penduduk terdiri dari atribut Nama Penduduk, Umur, Jumlah Keluarga, Dusun.

#### Tabel Atribut Data

Nama	Nama Penduduk Desa Situmekar
Umur	Umur Penduduk Desa Situmekar
Jumlah Keluarga	Jumlah Keluarga Penduduk
Dusun	Tempat Tinggal Penduduk

### 3.4. Metode Analisa Data

#### A. Pengolahan Data Awal

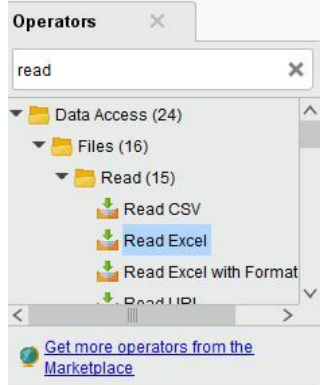
Pada tahap ini merupakan tahap untuk memastikan data penduduk yang dipilih telah layak untuk dilakukan proses pengolahan.

#### B. Permodelan

Pada tahap permodelan ini dilakukan dengan menggunakan software *rapidminer* 9.6 dengan langkah sebagai berikut :

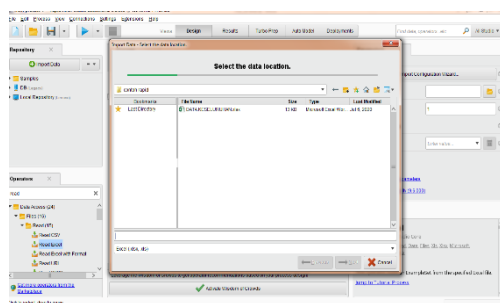
1. Buka *software rapidminer*
2. Lalu pilih menu *file* kemudian *new proses*
3. Masuk kedalam panel *operator* seperti yang terlihat pda gambar 1 kemudian ketik *read excel*, lalu *double* klik atau

drag pada *read excel* agar operator tersebut muncul di panel *process* dan bisa digunakan untuk memasukkan data *excel* yang telah melalui tahap *data cleansing* dan *data transformation*.



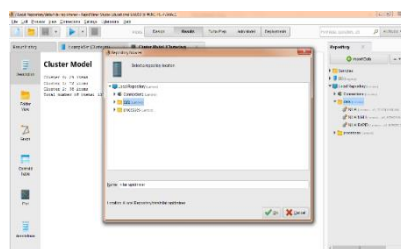
**Gambar 1 Pilih Read Excel**

4. Pilih *button import configuration wizard* agar bisa memilih file *dataset* yang akan diproses di *rapidminer*



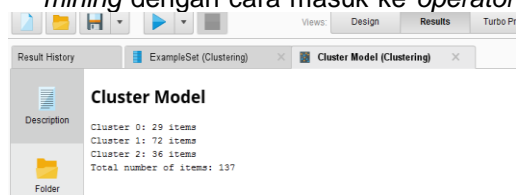
**Gambar 2 Memilih Dataset**

5. Kemudian simpan pada proses dengan nama nilai seperti pada gambar 3

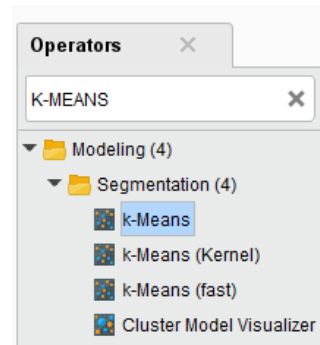


**Gambar 3 Simpan Dataset**

6. Langkah selanjutnya adalah memasukkan model yang akan digunakan untuk proses *clustering data mining* dengan cara masuk ke operator

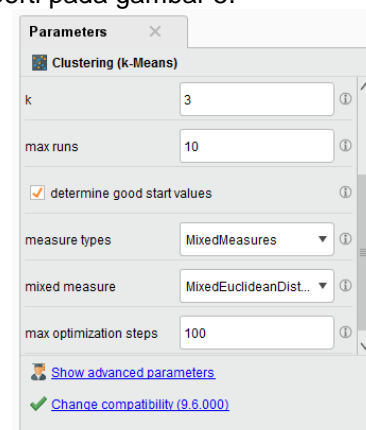


kemudian pilih *K-Means* seperti pada gambar 4



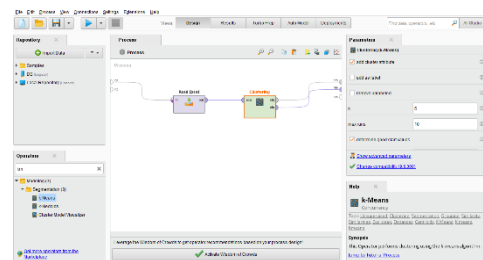
**Gambar 4 Pilih Permodelan**

7. Pada proses *clustering* tentukan terlebih dahulu jumlah klasternya menjadi  $K=3$  dan *mixed measure* menjadi *mixed Euclidean distance* seperti pada gambar 5.



**Gambar 5 Menentukan Jumlah K**

8. Langkah selanjutnya adalah menghubungkan *conector* masing –



masing proses pada *main process* seperti pada gambar 6 selanjutnya klik *button run*.

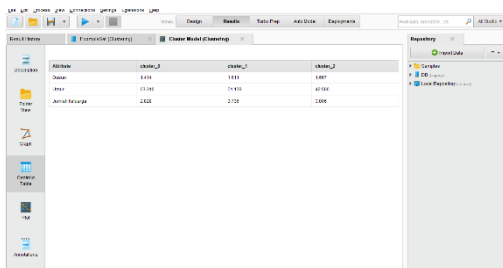
**Gambar 6 Menggabungkan Conector**

9. Pada gambar 7 hasil pengujian *dataset* yang berjumlah 108 data menggunakan *software rapidminer 9.6* terbentuk 3 *clster*. Pada *cluster 0* (*cluster pertama*)

terdapat 29 *items*, pada *cluster 1* (*cluster* kedua) terdapat 72 *items*, dan pada *cluster 2* (*cluster* ketiga) terdapat 36 *items*

**Gambar 7 Hasil Cluster**

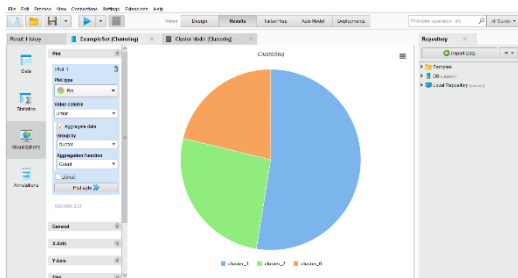
10. Dari 3 *cluster* yang terbentuk kemudian *centroid* akhir untuk setiap *cluster*, yaitu *cluster 0* = 1.414, 63.310, 2.828. Untuk *cluster 1* = 1.611, 51.139, 3.736. Dan *cluster 2* = 1.667, 42.500, 3.806. Seperti yang bisa dilihat pada gambar seperti pada gambar 8



Cluster	Jumlah Item	Jumlah Item	Jumlah Item
Cluster 0	1.414	63.310	2.828
Cluster 1	1.611	51.139	3.736
Cluster 2	1.667	42.500	3.806

**Gambar 8 Centroid Tabel**

11. Kemudian bisa dilihat *chart* dari setiap jumlah *cluster* seperti pada gambar 9

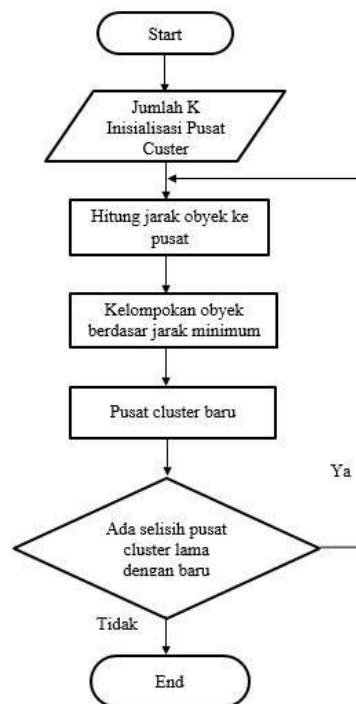


**Gambar 9 Piechart Cluster**

### 3. Hasil Penelitian dan Pembahasan

#### 3.1. Hasil Perhitungan

##### A. Flowchat Pengolahan Data



**Gambar IV. 1 Flowchart K-Means**

Untuk melakukan perhitungan manual seperti pada Flowchart diatas, maka penjelasannya sebagai berikut :

1. Siapkan data terlebih dahulu agar data bisa diolah/dihitung.
2. Tentukan jumlah K Inisialisasi Pusat *Cluster* sesuai yang dibutuhkan.
3. Setelah jumlah K Inisialisasi telah ditentukan maka hitung jarak ke pusat agar mengetahui hasil data yang akan dikelompokan pada setiap *cluster*
4. Setelah jarak objek ke pusat telah diketahui hasilnya, maka dapat dihitung jarak minimum data untuk dikelompokan.
5. Setelah data dikelompokan berdasarkan jarak minimum dan *cluster*, maka akan diketahui pusat *cluster* baru yang akan dihitung kembali
6. Setelah pusat *cluster* baru diketahui dihitung kembali, jika perhitungan antara *cluster* baru dan *cluster* lama ada selisih (berubah) data maka perhitungan diulang kembali, tetapi jika perhitungan antara *cluster* baru dan *cluster* lama sudah tidak ada selisih (tidak berubah) perhitungan bisa berhenti.

##### B. Perhitungan K-Means

Untuk menentukan banyaknya *cluster* k dapat dilakukan dengan beberapa pertimbangan seperti pertimbangan terioritis

dan konseptual yang mungkin diusulkan untuk menentukan berapa banyak *cluster*. Penetapan jumlah *cluster*  $k$  pada penelitian ini yaitu berjumlah 3 *cluster*.

#### Titik Awal Tiap *Cluster*

Titik Pusat Awal	Dusun	Umur	Jumlah Keluarga
Cluster 1	1	50	3
Cluster 2	2	55	4
Cluster 3	2	65	3

Hitung jarak setiap data ke pusat antara objek ke centroid dengan perhitungan jarak *Euclidean*. Perhitungan yang digunakan adalah:

$(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + (p_3 - q_3)^2 + (p_4 - q_4)^2 + (p_5 - q_5)^2}$   
 Persamaan ini digunakan karena atribut yang digunakan berjumlah 3.

Sebagai contoh, akan dihitung jarak dari data siswa pertama ke pusat cluster pertama dengan persamaan:

$$\begin{aligned} (1,1) &= \sqrt{(1-1)^2 + (54-55)^2 + (4-3)^2} \\ &= \sqrt{(0)^2 + (-1)^2 + (1)^2} \\ &= \sqrt{0+1+1} \\ &= \sqrt{2} \\ &= 1.41 \end{aligned}$$

Dari hasil perhitungan di atas maka di dapatkan hasil bahwa jarak data nilai siswa pertama pada *cluster* pertama adalah 1.41. Pada data nilai siswa kedua dapat dihitung jarak pada setiap *cluster* dari nilai data siswa agar mengetahui hasil persamaan:

$$\begin{aligned} (2,1) &= \sqrt{(1-1)^2 + (50-55)^2 + (3-3)^2} \\ &= \sqrt{(0)^2 + (-5)^2 + (0)^2} \\ &= \sqrt{0+25+0} \\ &= \sqrt{25} \\ &= 5 \end{aligned}$$

Maka dari hasil perhitungan di atas di dapatkan hasil jarak bahwa jarak data nilai siswa kedua pada *cluster* pertama adalah 5. Pada data nilai siswa ketiga dapat dihitung jarak pada setiap *cluster* dari nilai data siswa agar mengetahui hasil persamaan:

$$\begin{aligned} (3,1) &= \sqrt{(1-1)^2 + (55-55)^2 + (4-3)^2} \\ &= \sqrt{(0)^2 + (0)^2 + (1)^2} \\ &= \sqrt{0+0+1} \\ &= \sqrt{1} \\ &= 1 \end{aligned}$$

Maka dari hasil perhitungan di atas di dapatkan hasil jarak bahwa jarak data nilai siswa ketiga pada *cluster* pertama adalah 1.

#### Contoh Hasil Perhitungan Data ke Setiap *Cluster*

No	Dusun	Umur	Jumlah Keluarga	Jarak			Cluster
				C1	C2	C3	
1.	1	54	4	1.41	4.12	11.09	1
2.	1	50	3	5	1.41	15.03	2
3.	1	55	4	1	5.10	10.10	1
4.	2	50	4	5.20	0	15.03	2
5.	2	65	3	10.05	15.03	0	3
6.	2	50	4	5.20	0	15.03	2
7.	2	50	4	5.20	0	15.03	2
8.	3	50	3	5.39	1.41	15.03	2
9.	3	65	2	10.25	15.17	1.41	3
10.	3	50	3	5.39	1.41	15.03	2

Kelompokan data ke dalam *cluster* dengan jarak terdekat (minimal).

#### Tabel Dengan *Centroid* Awal

No	Cluster 1	Cluster 2	Cluster 3
1	1	0	0
2	0	1	0
3	1	0	0
4	0	1	0
5	0	0	1
6	0	1	0
7	0	1	0
8	0	1	0
9	0	0	1
10	0	1	0

Setelah semua data dimasukkan kedalam *cluster* sesuai dengan jarak dan tempatnya, kemudian hitung kembali pusat *cluster* yang baru berdasarkan rata-rata anggota yang ada pada *cluster* tersebut sepusat dari setiap *cluster* tidak berubah lagi dan tidak ada lagi data yang berpindah dari *cluster* ke *cluster* yang lain

#### 3.2. Hasil Penelitian dan Pembahasan

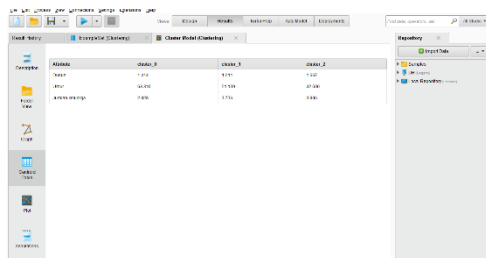
Hasil penelitian yang telah di uji ini maka hasil antara *Software Rapidminer* dan *Microsoft Excel* yaitu mendapatkan hasil yang sama. Pada eksperimen data di *Microsoft Excel* menemukan hasil yang tidak berubah yaitu pada iterasi 7 dengan 3

*cluster* dan beberapa atribut seperti Dusun, Umur dan Jumlah Keluarga.

**Tabel Centroid Baru**

Titik Pusat Awal	Dusun	Umur	Jumlah Keluarga
Cluster 1	1.42	62.9	2.87
Cluster 2	1.61	50.97	3.74
Cluster 3	1.67	42.5	3.81

Untuk memastikan hasil perhitungan manual di *Microsoft Excel* dan perhitungan *software rapidminer* 9.6 tidak jauh berbeda maka dibuat perbandingan seperti berikut :



**Gambar IV. 2 Hasil Software Rapidminer**

**Hasil Microsoft Excel**

Titik Pusat	Dusun	Umur	Jumlah Keluarga
Cluster 1	1.419	62.903	2.871
Cluster 2	1.614	50.971	3.743
Cluster 3	1.667	42.500	3.806

Berdasarkan hasil dari perhitungan yang telah dibuat, 137 data penduduk kurang mampu dibagi ke dalam 3 *cluster* berdasarkan Dusun, Umur, Jumlah Keluarga. Pada *cluster* 1 menghasilkan jumlah 29 penduduk dengan hasil 33.069. Pada *cluster* 2 menghasilkan jumlah 72 penduduk dengan hasil 27.438. Pada *cluster* 3 menghasilkan jumlah 36 penduduk dengan hasil 23.153.

Dari 3 *cluster* yang terbentuk, *cluster* 2 adalah *cluster* yang terbanyak mendapatkan bantuan dengan jumlah 72 penduduk. Pada *cluster* 3 dengan jumlah 36 penduduk yang berhak mendapatkan

bantuan, sedangkan pada *cluster* 1 adalah *cluster* yang paling rendah mendapatkan bantuan dengan jumlah 29 penduduk.

#### 4. Kesimpulan

Berdasarkan hasil dari penelitian yang telah diuji eksperimen data menggunakan *Microsoft Excel* dan *Software Rapidminer* dengan Algoritma *K-Means Clustering* dalam penelitian ini, dapat disimpulkan sebagai berikut:

1. Telah diterapkan Metode Algoritma *K-Means Clustering* untuk mengelompokkan data penduduk kurang mampu yang berhak mendapatkan bantuan.
2. Pengolahan data yang dilakukan menghasilkan kelompok penerima bantuan, sehingga pihak desa bisa lebih mudah dalam menentukan penerima bantuan dengan teknik *Data Mining*.
3. Berdasarkan hasil pengolahan data yang telah melakukan eksperimen dapat diketahui 3 *cluster* yang dibutuhkan oleh Desa Situmekar. Pada *cluster* 1 adalah kelompok penduduk yang paling rendah menerima bantuan, sedangkan *cluster* 2 kelompok penduduk yang paling terbanyak menerima bantuan, dan *cluster* 3 yaitu kelompok kedua yang menerima bantuan.

#### 5. Saran

Berikut ini adalah saran untuk pengembangan pada penelitian selanjutnya antara lain:

1. Untuk penelitian selanjutnya, bisa menambahkan atribut penghasilan kepala keluarga agar lebih tepat dalam menerima bantuan tersebut.
2. Hasil dari penelitian dapat dijadikan sebagai salah satu referensi bagi pihak sekolah dalam mendukung keputusan untuk melakukan pengelompokan penduduk untuk menerima bantuan.

#### Referensi

- Annur, H. (2018). Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes. *ILKOM Jurnal Ilmiah*, 160-165.
- Febianto, N. I., & Palasara, N. (2019). Analisa Clustering K-Means Pada Data Informasi Kemiskinan Di Jawa Barat Tahun 2018. *Jurnal Sisfokom*

- 
- ( *Sistem Informasi dan Komputer*),  
130.
- Febrianti, F., Hafiyusholeh, M., & Asyhar, A. H. (2016). Perbandingan Pengklusteran Data Iris Menggunakan Metode K-Means Dan Fuzzy C-Means. *Jurnal Matematika "MANTIK"*, 7.
- Ferezagia, V. D. (2018). analisis tingkat kemiskinan di indonesia. *jurnal sosial humaniora terapan*, 1-6.
- Mardi, Y. (2017). Data Mining : Klasifikasi Menggunakan algoritma C4.5. *jurnal Edik Informatika*, 213-219.
- Metisen, B. M., & Sari, H. L. (2015). Analisis clustering menggunakan metode K-Means dalam pengelompokkan penjualan produk pada Swalayan Fadhila. *Journal Media Infotama*, 110-118.
- Sari, R. W., Wanto, A., & Windarto, A. P. (2018). Implementasi Rapidminer Dengan Metode K-Means (Study Kasus: Imunisasi Campak Pada Balita Berdasarkan Provinsi). *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, 224-230.
- Statistik, B. P. (2018). DATA DAN INFORMASI KEMISKINAN KABUPATEN / KOTA TAHUN 2018.
- Yusup, F. (2018). UJI VALIDITAS DAN RELIABILITAS INSTRUMEN PENELITIAN KUANTITATIF. *Jurnal Ilmu Pendidikan*, 17-23.
- Z, A. Z., & Sarjono. (2016). Analisis Data Mining Untuk Menentukan Kelompok Prioritas Penerima Bantuan Bedah Rumah Menggunakan Metode Clustering K-Means( Studi Kasus: Kantor Kecamatan Bahar Utara). *jurnal manajemen sistem informasi*, 159-170.
- Zuhdiyati, N., & Kaluge, D. (2015). KEMISKINAN DI INDONESIA SELAMA LIMA TAHUN TERAKHIR ( Studi Kasus Pada 33 Provinsi ). 27-31.