

Analisis Algoritma K-Means untuk Klustering Penerima Bantuan Sosial Covid-19

Analysis of K-Means Algorithm for *Clustering* of Covid-19 Social Assistance Recipients

Sri Rahmayani¹, S Sumarno², Zulia Almada Siregar³

^{1,2,3} STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

Article Info

Genesis Artikel:

Diterima, 3 Maret 2022

Direvisi, 4 Maret 2022

Disetujui, 13 Maret 2022

Kata Kunci:

Data Mining

K-Means

Cluster

Covid-19

Bantuan Sosial

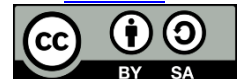
ABSTRAK

Pada masa pandemi Covid-19 ini, pemerintah memberikan bantuan yang di disalurkan melalui masing-masing Kelurahan di seluruh Provinsi Indonesia, salah satunya adalah Kelurahan Pahlawan yang ada di Kecamatan Siantar Timur kota Pematangsiantar. Selama ini bantuan yang dilakukan oleh Kelurahan Pahlawan masih secara manual sehingga tidak menutup kemungkinan kesalahan dalam pendataan dan pembagian bantuan akan terjadi. Untuk mengatasi masalah tersebut dilakukan penelitian dengan menerapkan algoritma K-Means untuk menentukan kluster kelayakan penerima bantuan Covid-19, yang dilakukan dengan cara mengumpulkan data penduduk sesuai atribut yang sudah ditentukan. Kemudian data penduduk tersebut akan di kluster menggunakan algoritma K-Means, dan dilakukan pengujian menggunakan aplikasi Rapid Miner. Hasil klustering yang diperoleh yakni : *cluster* 0 terdiri dari 26 data, *cluster* 1 terdiri dari 24 data. Penerima bantuan sosial Covid-19 dengan menggunakan algoritma K-Means menunjukkan bahwa yang berhak menerima bantuan adalah masyarakat yang lansia (lanjut usia). Berdasarkan hal ini dapat disimpulkan bahwa Algoritma K-Means dapat di terapkan untuk menghasilkan informasi yang lebih efektif dalam menentukan siapa saja yang berhak menerima bantuan.

ABSTRACT

During the Covid-19 pandemic, the government provided assistance distributed through each sub-district throughout the province of Indonesia, one of which was the Pahlawan Village in the East Siantar District Pematangsiantar City. So far, the assistance provided by Kelurahan Pahlawan is still done manually, so errors in data collection and distribution of aid may occur. To overcome this problem, a study was carried out by applying the K-Means algorithm to determine the eligibility cluster of Covid-19 beneficiaries, which was carried out by collecting population data according to predetermined attributes. Then the population data will be clustered using the K-Means algorithm and tested using the Rapid Miner application. The clustering results obtained are that cluster 0 consists of 26 data and that cluster 1 consists of 24 data. The recipients of Covid-19 social assistance using the K-Means algorithm show that those entitled to receive the gift are the elderly (elderly). Based on this, it can be concluded that the K-Means Algorithm can be applied to produce more practical information in determining who is entitled to receive assistance.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Penulis Korespondensi:

Sri Rahmayani,
Program Studi Teknik Informatika,
STIKOM Tunas Bangsa, Pematangsiantar, indonesia
Email: sriramayani69@gmail.com

1. PENDAHULUAN

Data mining merupakan proses pengekstraksian fakta berdasarkan sekumpulan data yang sangat besar melalui penggunaan prosedur pemecahan dan teknik penarikan pada bidang statistik, pembelajaran mesin dan sistem manajemen basis data [1]. Data mining terdapat metode-metode yang dapat digunakan yaitu klasifikasi, *clustering*, regresi, seleksi variabel, dan market basket analisis. K-Means merupakan algoritma heuristik yang memisah perpaduan data ke dalam kluster K dengan meminimalkan jumlah jarak kuadrat pada setiap kluster [2]. Data mining adalah serangkaian proses untuk menggali nilai tambah berdasarkan suatu formasi data berupa pengetahuan yang selama ini tidak diketahui secara manual [3]. Algoritma K-Means dimulai menggunakan pembentukan partisi kluster di awal lalu secara interaktif partisi *cluster* ini diperbaiki sampai tidak terjadi perubahan yang signifikan dalam partisi *cluster*.

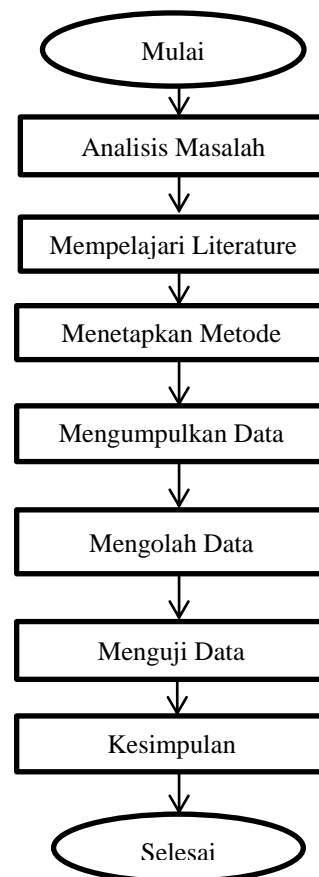
Covid-19 telah merubah perekonomian dunia, termasuk Indonesia dibuktikan adanya pengalihan-pengalihan anggaran dalam rangka penanganannya, termasuk di desa-desa. Penamaan 2019-nCoV disematkan *World Health Organization* (WHO) untuk *coronavirus disease that was discovered in 2019* pada tanggal 11 Februari 2020, dengan sebutan Covid-19. Covid-19 memiliki dampak yang sangat luar biasa di seluruh dunia tidak ketinggalan juga dengan desa-desa yang ada di Indonesia. Desa-desa terpengaruh terutama dalam bidang ekonomi dan pembangunan yang selama ini banyak mengandalkan bantuan dari pemerintah. Begitu juga pada perekonomian masyarakat, terutama masyarakat Kota Pematangsiantar yang kurang mampu. Seperti yang kita ketahui saat ini pemerintah juga memberikan bantuan sosial Covid-19 kepada masyarakat. Dalam penentuan penerima bantuan sosial Covid-19 pada Kantor Kelurahan Pahlawan Kecamatan Siantar Timur juga telah menentukan kriteria untuk masyarakat yang mendapat bantuan.

Beberapa penulis penelitian terdahulu yang telah ditulis dalam sebuah jurnal atau karya ilmiah tentang keputusan penerima bantuan Covid-19 [4], penulis membahas Program bantuan sosial yang diberikan pemerintah untuk penanggulangan pandemi COVID-19 masih kurang efektif karena tidak tepat sasaran.

2. METODE PENELITIAN

2.1. Rancangan Penelitian

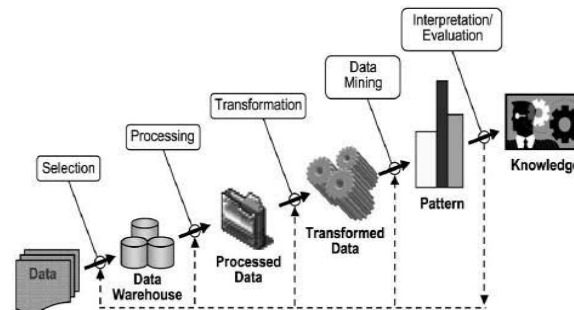
Rancangan terhadap penelitian ini dilakukan dengan melakukan sebuah pengamatan selanjutnya mengumpulkan data, setelah data tersebut dimasukkan ke Microsoft excel lalu data akan diolah melalui proses perhitungan dan mengikuti langkah-langkah metode K-Means. Hasil perhitungan tersebut dapat diaplikasikan ke *Rapid Miner* untuk melihat hasil yang akurat. Gambar dari perancangan penelitian dapat di lihat seperti pada gambar 1 berikut.



Gambar 1. Rancangan Penelitian

2.2. Data Mining

Data Mining merupakan proses pengekstraksian fakta berdasarkan sekumpulan data yang sangat besar melalui penggunaan prosedur pemecahan dan teknik penarikan pada bidang statistik, pembelajaran mesin dan sistem manajemen basis data [5]. Data Mining merupakan proses yang memakai teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi fakta yang berguna dan pengetahuan yang terkait berdasarkan *database* yang besar [6].



Gambar 2. Tahapan Proses Data Mining

Gambar 2 merupakan langkah untuk membentuk proses data yang akan telah disiapkan. Untuk proses penggalian dari sebuah interaksi data.

2.3. Clustering

Analisis Pengelompokan/*Clustering* merupakan proses membagi data pada suatu himpunan ke dalam beberapa grup yang kecenderungan datanya pada suatu grup lebih besar dari pada kecenderungan data tersebut dengan data pada grup lain. Potensi *clustering* adalah bisa dipakai untuk mengetahui struktur pada data yang bisa digunakan lebih lanjut pada aneka macam aplikasi secara luas misalnya klasifikasi, pengolahan gambar, dan sosialisasi pola [7]

2.4. K-Means

K-Means adalah algoritma *heuristik* yang memisah perpaduan data ke dalam kluster K dengan meminimalkan jumlah jarak kuadrat pada setiap kluster [8]. Algoritma K-Means dimulai menggunakan pembentukan partisi *cluster* di awal lalu secara interaktif partisi *cluster* ini diperbaiki sampai tidak terjadi perubahan yang signifikan dalam partisi *cluster*. K-Means memiliki kemampuan mengelompokkan data pada jumlah yang relatif besar menggunakan waktu komputasi yang cukup cepat dan efisien. Namun, *K-Means* memiliki kelemahan yang diakibatkan oleh penentuan pusat awal *cluster*. Hasil *cluster* yang terbentuk berdasarkan metode *K-Means* ini sangatlah tergantung dalam inisiasi nilai pusat awal *cluster* yang diberikan [9].

Berikut ini merupakan langkah-langkah pada algoritma *K-means* [10] :

1. Tentukan jumlah *cluster* (*k*) dalam data set.
2. Tentukan nilai pusat (*centroid*) Penentuan nilai *centroid* dalam tahap awal dilakukan secara acak menggunakan rumus menentukan target awal *k-means*, rumus tadi dipakai untuk menerima target data atau jarak antara kelompok, yaitu titik pusat awal untuk menghitung algoritma *k-means* perulangan 0.

$$\frac{\text{Jumlah data}}{\text{Jumlah class} + 1} \quad (1)$$

Keterangan:

Jumlah data = Jumlah data yang akan digunakan

Jumlah class = Jumlah kelompok yang sudah ditentukan sebelumnya seperti sangat tinggi, tinggi, normal, rendah, sangat rendah.

Sedangkan dalam tahap perulangan dipakai rumus rata-rata, pada rumus tersebut dilakukan perhitungan untuk mencari nilai rata-rata seperti berikut ini :

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj} \quad (2)$$

Keterangan:

V_{ij} = *centroid* rata-rata *cluster* ke-*i* untuk variabel ke-*j*

N_i = jumlah anggota *cluster* ke-*i*

i, k = indeks dari *cluster*

j = indeks dari variabel

X_{kj} = nilai data ke-*k* variabel ke-*j* dalam *cluster* tersebut

3. Pada masing-masing *record*, hitung jarak terdekat menggunakan *centroid*

Jarak *centroid* yang dipakai merupakan *Euclidean Distance*, menggunakan rumus seperti berikut ini :

$$De = \sqrt{(X_i - y_i)^2 + (y_i - t_i)^2} \quad (3)$$

Keterangan:

De = *Euclidean Distance*
 I = banyaknya objek
 (x, y) = koordinat objek

3. HASIL DAN ANALISIS

Hasil akhir dalam penelitian ini menyajikan proses pengolahan data yang dibagi menjadi dua tahap yaitu perhitungan manual menggunakan Algoritma K-Means *Clustering* dan menyesuaikan hasil dari perhitungan manual dengan pengujian menggunakan software Rapid Miner 5.3.

3.1. Pengolahan Data

1. Menentukan data yang akan di *cluster*.

Sampel data yang akan digunakan dalam *clustering* dalam menentukan penerima bantuan sosial Covid-19. Tabel penggunaan data dapat dilihat pada tabel 1.

Tabel 1. Data Kelurahan Pahlawan

No.	Nama Kepala Keluarga	Usia	Jumlah Tanggungan
1	Abdul Rahim Damanik	74	2
2	Affandi	42	5
3	Agus Leonard Hutapea	41	4
4	Amady Muchlis Azhary	28	2
5	Alasen Kaban	56	2
6	Ani Pardede	63	1
7	Anwar Fuady Azhary Hasibuan	31	4
8	Arjuna Pardede	60	2
9	Asmah Siagian	55	2
10	Asnik Marpaung	32	1
11	Astator Gultom	55	5
12	Bachtiar Sinaga	75	3
13	Baharuddin Tampubolon	55	3
14	Betteria Nababan	62	1
15	Bawadi Siahaan	45	2
16	Bintara Bayu	33	1
17	Budi	42	5
18	Budiman Tampubolon	70	2
19	Chairy Azhari	33	3
20	Chandra Sinaga	36	1
21	Dahrís	76	2
22	Dapit Hutapea	52	6
23	Darwin	38	4
24	Difpo Afriansyah	31	2
25	Effendi	49	5
26	Elfian	42	3
27	Erwinsyah Saragih	46	4
28	Ferry Jan Hotman Purba	41	4
29	Fran Dedi	40	4
30	Frans Rudi Purba	29	2
31	Hamin Siagian	46	2
32	Hartaty Tuah	55	2
33	Hasudungan Panggabean	53	5
34	Hendry Harianto Purba	35	4
35	Henry Hamonangan Siahaan	41	4
36	Heru Setiawan	28	1
37	Hotman Arianto Silalahi	31	5
38	Irnawati	53	4
39	Irwan Gunawan	49	3
40	Jahot Sinaga	63	5
41	Jannes	76	2
42	Jan Wissen Purba	69	2
43	Jani Marpaung	64	3
44	Jawasen Sinaga	52	6
45	Jensfrida Purba	58	4
46	Jhon Henri Saragih	57	3
47	Jomson Sidabutar	59	5
48	Joni Asalman Jingga	60	1
49	Kondarius Ambarita	68	4
50	Kontanim Saragih	70	1

2. Menentukan nilai K jumlah *cluster*, Jumlah *cluster* sebanyak 2 *cluster*. *Cluster* yang dibentuk yaitu *cluster* yang menerima (C1), *cluster* yang tidak menerima (C2).

3. Menentukan nilai centroid (pusat *cluster*).

Penentuan pusat *cluster* awal ditentukan secara random yang diambil dari data yang ada dalam range. *Cluster* yang menerima (*cluster* 1) diambil dari usia 50 tahun ke atas pada tabel 4.1. Dan *cluster* yang tidak menerima (*cluster* 2) diambil dari usia di bawah 50 tahun terdapat pada table 1. Berikut daftar tabel centroid data dapat dilihat pada tabel 2 berikut.

Tabel 2. Centoroid Data Awal

Cluster	Nama Kepala Keluarga	Usia	Jumlah Tanggungan
C1	Abdul Rahim Damanik	74	2
C2	Heru Setiawan	28	1

4. Menghitung jarak setiap data terhadap centroid (pusat *cluster*)

Setelah data nilai pusat *cluster* awal ditentukan, maka langkah selanjutnya adalah menghitung jarak masing-masing data terhadap pusat *cluster*. Proses pencarian jarak terpendek pada iterasi 1 dapat dilihat pada perhitungan dan tabel 3 berikut.

$$\text{Abdul Rahim Damanik, C1} = \sqrt{(74 - 74)^2 + (2 - 2)^2} = 0$$

$$\text{Abdul Rahim Damanik, C1} = \sqrt{(74 - 28)^2 + (2 - 1)^2} = 46.01087$$

Tabel 3. Hasil Perhitungan Jarak Terdekat Pusat *Cluster* Iterasi 1

Nama Kepala Keluarga	Usia	Jumlah Tanggungan	C1	C2	Jarak Terdekat
Abdul Rahim Damanik	74	2	0	46.01087	0
Affandi	42	5	32.14031736	14.56022	14.56021978
Agus Leonard Hutapea	41	4	33.06055051	13.34166	13.34166406
Amady Muchlis Azhary	28	2	46	1	1
Alasen Kaban	56	2	18	28.01785	18
Ani Pardede	63	1	11.04536102	35	11.04536102
Anwar Fuady Azhary Hasibuan	31	4	43.0464865	4.242641	4.242640687
Arjuna Pardede	60	2	14	32.01562	14
Asmah Siagian	55	2	19	27.01851	19
Asnik Marpaung	32	1	42.01190308	4	4
Astator Gultom	55	5	19.23538406	27.29469	19.23538406
Bachtar Sinaga	75	3	1.414213562	47.04253	1.414213562
Baharuddin Tampubolon	55	3	19.02629759	27.07397	19.02629759
Betteria Nababan	62	1	12.04159458	34	12.04159458
Bawadi Siahaan	45	2	29	17.02939	17.02938637
Bintara Bayu	33	1	41.01219331	5	5
Budi	42	5	32.14031736	14.56022	14.56021978
Budiman Tampubolon	70	2	4	42.0119	4
Chairy Azhari	33	3	41.01219331	5.385165	5.385164807
Chandra Sinaga	36	1	38.01315562	8	8
Dahrir	76	2	2	48.01042	2
Dapit Hutapea	52	6	22.36067977	24.5153	22.36067977
Darwin	38	4	36.0551275	10.44031	10.44030651
Difpo Afriansyah	31	2	43	3.162278	3.16227766
Effendi	49	5	25.17935662	21.37756	21.37755833
Elfian	42	3	32.01562119	14.14214	14.14213562
Erwinsyah Saragih	46	4	28.0713377	18.24829	18.24828759
Ferry Jan Hotman Purba	41	4	33.06055051	13.34166	13.34166406
Fran Dedi	40	4	34.05877273	12.36932	12.36931688
Frans Rudi Purba	29	2	45	1.414214	1.414213562
Hamin Siagian	46	2	28	18.02776	18.02775638
Hartaty Tuah	55	2	19	27.01851	19
Hasudungan Panggabean	53	5	21.21320344	25.31798	21.21320344
Hendry Harianto Purba	35	4	39.05124838	7.615773	7.615773106
Henry Hamonangan Siahaan	41	4	33.06055051	13.34166	13.34166406
Heru Setiawan	28	1	46.01086828	0	0
Hotman Arianto Silalahi	31	5	43.10452412	5	5
Irnowati	53	4	21.09502311	25.17936	21.09502311
Irwan Gunawan	49	3	25.01999201	21.09502	21.09502311
Jahot Sinaga	63	5	11.40175425	35.22783	11.40175425
Jannes	76	2	2	48.01042	2
Jan Wissen Purba	69	2	5	41.01219	5
Jani Marpaung	64	3	10.04987562	36.05551	10.04987562
Jawasen Sinaga	52	6	22.36067977	24.5153	22.36067977
Jensfrida Purba	58	4	16.1245155	30.14963	16.1245155
Jhon Henri Saragih	57	3	17.02938637	29.06888	17.02938637
Jomson Sidabutar	59	5	15.29705854	31.257	15.29705854
Joni Asalman Jingga	60	1	14.03566885	32	14.03566885
Kondarius Ambarita	68	4	6.32455532	40.11234	6.32455532
Kontanin Saragih	70	1	4.123105626	42	4.123105626

5. Menghitung *centroid* baru menghitung *centroid* baru menggunakan hasil dari setiap anggota pada masing-masing *cluster*. Setelah didapatkan hasil jarak dari setiap objek pada iterasi ke-1 maka lanjut ke iterasi ke-2 pada perhitungan berikut.

$$\begin{aligned}
 &74+56+63+60+55+55+75+55+62+70+ \\
 &76+52+55+53+53+63+76+69+64+52 \\
 D1, C1 &= \frac{+58+57+59+60+68+70}{26} = 61.93208 \\
 &2+2+1+2+2+5+3+3+1+2+2+6+2+5+4+ \\
 D1, C2 &= \frac{5+4+5+2+2+3+6+4+3+5+1+4+1}{26} = 3.25 \\
 &42+41+28+31+32+45+33+42+33+36+ \\
 D1, C2 &= \frac{38+31+49+42+46+41+40+29+46+35+41+28+31+49}{24} = 37.875 \\
 &5+4+2+4+1+2+1+5+3+1+4 \\
 D2, C2 &= \frac{+2+5+3+4+4+4+2+2+4+4+1+5+3}{24} = 3.125
 \end{aligned}$$

Jika tahap iterasi telah mencapai hasil yang sama tanpa dan tidak mengalami perubahan lagi. Maka perhitungan dihentikan.

Perhitungan manual pada data diatas didapatkan hasil akhir yang dimana pada iterasi 1 dan iterasi 2 pengelompokan data yang dilakukan terhadap 2 *cluster* didapatkan hasil yang sama. Hasil dari kedua iterasi tersebut bernilai C1 =26, C2 =24 pada posisi data tiap *cluster*. Sehingga posisi *cluster* pada data tersebut tidak mengalami perubahan lagi maka proses iterasi berhenti sampai iterasi 2.

Table 4. Hasil Pengelompokan Iterasi ke-2

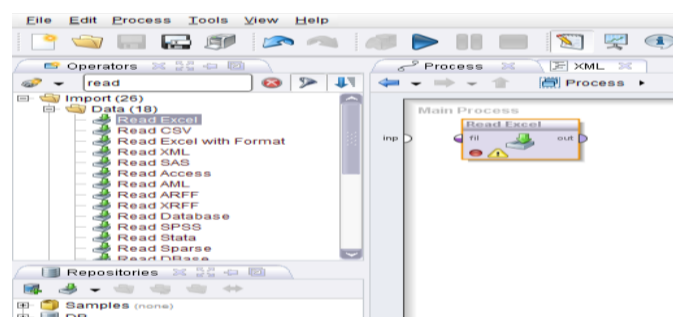
Cluster	Nama-Nama Penerima Bantuan	Jumlah
C1	Abdul Rahim Damanik, Alasen Kaban, Ani Pardede, Arjuna Pardede, Asmah Siagian, Astator Gultom, Bachtar Sinaga, Baharuddin Tampubolon, Betteria Nababan, Budiman Tampubolon, Dahris, Dapit Hutapea, Hartaty Tuah, Hasudungan Panggabean, Irmawati, Jahot Sinaga, Jan Wissen Purba, Jani Marpaung, Jawasen Sinaga, Jensfrida Purba, Jhon Henri Saragih, Jomson Sidabutar, Joni Asalman Jingga, Kondarius Ambarita, kontanim Saragih.	26
C2	Affandi, Agus Leonard Hutapea, Amady Muchlis Azhary, Anwar Fuady Azhary Hasibuan, Asnik Marpaung, Bawadi Siahaan, Bintara Bayu, Chairy Azhari, Chandra Sinaga, Darwin, Difpo Afriansyah, Effendi, Elfian, Erwinsyah Saragih, Ferry Jan Hotman Purba, Fran Dedi, Frans Rudi Purba, Hamin Siagian, Hendry Harianto Purba, Henry Hamonangan Siahaan, Heru Setiawan, Hotman Arianto Silalahi,, Irwan Gunawan, Jannes.	24

3.2. Implementasi pada Rapid Miner

Implementasi pada aplikasi *Rapid Miner* dapat dilihat sebagai berikut :

1. *Import data ke dalam Rapid Miner dalam bentuk Sheet Excel.*

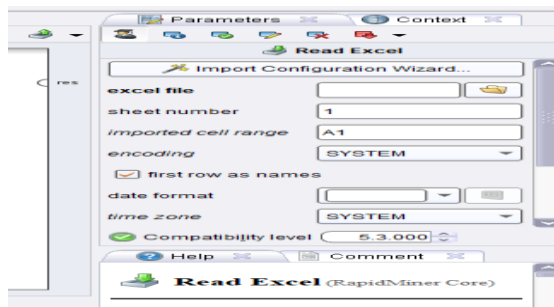
Sistem menjelaskan cara memasukkan data baru yang akan dieksekusi lebih lanjut, pada hal ini data yang akan di eksekusi berupa data *excel*. Klik pada bagian kiri bawah tab *repositories* lalu pilih "*Import Read Excel*". Kemudian akan muncul tampilan seperti gambar di bawah ini.



Gambar 3. *Import data*

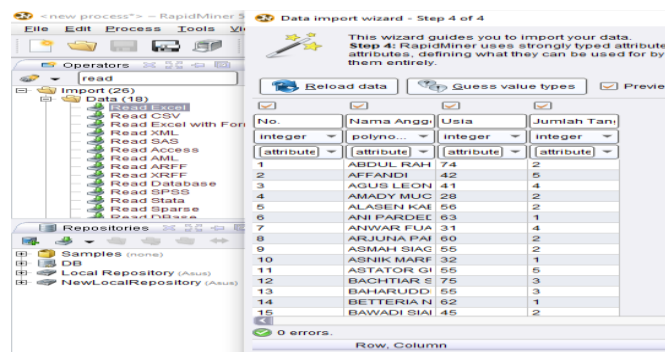
2. *Import Configurasi Wizard*

Kemudian akan muncul data *import configuration wizard* kemudian pilih tempat kita menyimpan data yang akan digunakan. Selanjutnya pilih *file name* data yang akan digunakan. Lalu klik *next* pada bagian kanan bawah seperti dapat dilihat pada gambar berikut.



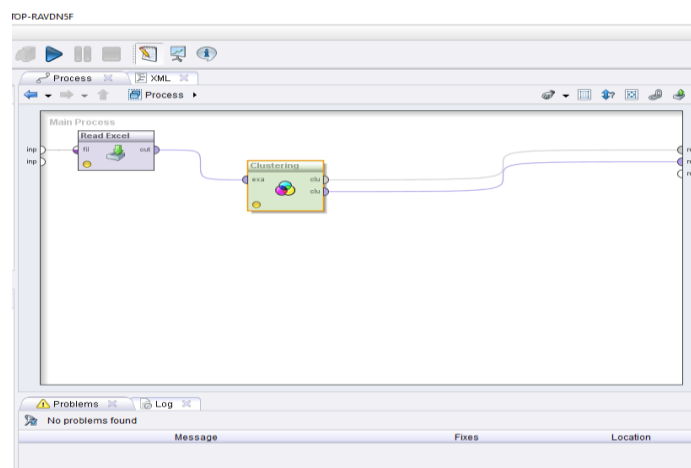
Gambar 4. Konfigurasi wizard

3. Setelah konfigurasi wizard, pilih file sesuai dengan atribut dan tipe data
 Pada tahap ini dilakukan pemilihan tipe data dimana pada bagian attribute yang di awal diubah tipe *label*, *attribute*.



Gambar 5. Penyesuaian Atribut Dan Tipe Data

4. Pemrosesan sistem *Rapid Miner*
 Pada tahap ini akan dijelaskan tahapan-tahapan proses penggunaan *k-means* di dalam *Rapid Miner* data yang telah di impor. Tahapan pertama dengan meng klik *clustering and segmentation* lalu pilih *k-means*, lalu hubungkan antara *Read excel* dengan *clustering* seperti gambar berikut.

Gambar 6. Pemrosesan *Rapid Miner*

5. Hasil pengelompokan dalam *Rapid Miner*
 Maka akan menampilkan hasil akhir serta langkah terakhir dalam penggunaan *tools Rapid Miner* ini. Dapat dilihat pada gambar berikut.

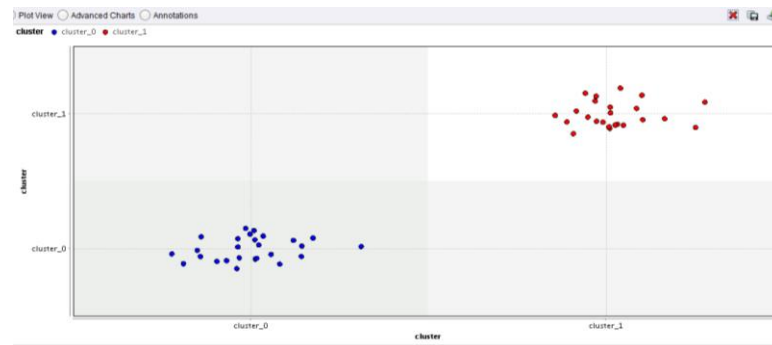
Cluster Model

Cluster 0: 26 items
 Cluster 1: 24 items
 Total number of items: 50

Gambar 7. Hasil Pengelompokan

6. Tampilan *Plot View*

Berdasarkan pada gambar di bawah dapat diketahui bahwa yang menerima bantuan memiliki *node* biru yaitu 26, sedangkan yang tidak menerima bantuan memiliki *node* merah yaitu 24.



Gambar 8. Hasil Tampilan Plot View

Node biru merupakan hasil *cluster* yang menerima. Sedangkan Node merah merupakan hasil *cluster* tidak menerima.

4. KESIMPULAN

Perhitungan algoritma K-Means dengan pengujian *tools Rapid Miner 5.3* diperoleh hasil yang sama dengan perhitungan manual. Hasil *clustering* yang diperoleh dimana *cluster* 0 terdiri dari 26 data, *cluster* 1 terdiri dari 24 data. Penerima bantuan sosial Covid-19 dengan menggunakan algoritma K-Means menunjukkan bahwa yang berhak menerima bantuan adalah masyarakat yang lansia (lanjut usia). Algoritma K-Means metode *Clustering* dapat di terapkan untuk menghasilkan informasi yang dapat diusulkan kepada kantor Lurah Pahlawan untuk kedepannya agar lebih efektif dalam menentukan penerima bantuan.

REFERENSI

- [1] Saleh, A. (2015). Penerapan Data Mining Dengan Metode Klasifikasi Naive Bayes Untuk Memprediksi Kelulusan Mahasiswa Dalam Mengikuti English Proficiency Test (Studi Kasus : Universitas Potensi Utama). Konferensi Nasional Sistem Informasi, At Universitas Klabat, Manado, Indonesia, June, 1–6.
- [2] Mardi, Y. (2017). Data Mining : Klasifikasi Menggunakan Algoritma C4.5. Edik Informatika, 2 (2), 213–219. <https://doi.org/10.22202/ei.2016.v2i2.1465>
- [3] Aditya, A., Jovian, I., & Sari, B. N. (2020). Implementasi K-Means *Clustering* Ujian Nasional Sekolah Menengah Pertama di Indonesia Tahun 2018/2019. Jurnal Media Informatika Budidarma, 4(1), 51. <https://doi.org/10.30865/mib.v4i1.1784>
- [4] Sembiring, F., Fauzi, M. T., Khalifah, S., Khotimah, A. K., & Rubiati, Y. (2020). Sistem Pendukung Keputusan Penerima Bantuan Covid 19 menggunakan Metode Simple Additive Weighting (SAW) (Studi Kasus : Desa Sundawenang). Explore: Jurnal Sistem Informasi Dan Telematika, 11(2), 97. <https://doi.org/10.36448/jsit.v11i2.1563>
- [5] Saleh, A. (2015). Penerapan Data Mining Dengan Metode Klasifikasi Naive Bayes Untuk Memprediksi Kelulusan Mahasiswa Dalam Mengikuti English Proficiency Test (Studi Kasus : Universitas Potensi Utama). Konferensi Nasional Sistem Informasi, At Universitas Klabat, Manado, Indonesia, June, 1–6.
- [6] Handoko, K. (2016). Penerapan Data Mining Dalam Meningkatkan Mutu Pembelajaran Pada Instansi Perguruan Tinggi Menggunakan Metode K-Means *Clustering* (Studi Kasus Di Program Studi Tkj Akademi Komunitas Solok Selatan). Jurnal Teknologi Dan Sistem Informasi, 02(03), 31–40. <http://teknosi.fti.unand.id/index.php/teknosi/article/view/70>
- [7] Sari, R. W., Wanto, A., & Windarto, A. P. (2018). Implementasi *Rapid Miner* Dengan Metode K-Means (Study Kasus: Imunisasi Campak Pada Balita Berdasarkan Provinsi). KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer), 2(1), 224–230. <https://doi.org/10.30865/komik.v2i1.930>
- [8] Aditya, A., Jovian, I., & Sari, B. N. (2020). Implementasi K-Means *Clustering* Ujian Nasional Sekolah Menengah Pertama di Indonesia Tahun 2018/2019. Jurnal Media Informatika Budidarma, 4(1), 51. <https://doi.org/10.30865/mib.v4i1.1784>
- [9] Handoko, K. (2016). Penerapan Data Mining Dalam Meningkatkan Mutu Pembelajaran Pada Instansi Perguruan Tinggi Menggunakan Metode K-Means *Clustering* (Studi Kasus Di Program Studi Tkj Akademi Komunitas Solok Selatan). Jurnal Teknologi Dan Sistem Informasi, 02(03), 31–40. <http://teknosi.fti.unand.id/index.php/teknosi/article/view/70>
- [10] Abdillah, G., Putra, F. A., & Renaldi, F. (2016). Penerapan Data Mining Pemakaian Air Pelanggan Baru Di Pdam Tirta Raharja Menggunakan Algoritma K-Means. Sentika 2016, Sentika, 18–19.