

Penerapan Algoritma K-Means Untuk Pemetaan Penerimaan Bantuan Kesejahteraan Masyarakat di Kota Batam

Nurul Azwanti^{a,*}, Narti Eka Putria^b

^{ab}Universitas Nagoya Indonesia, Kota Batam

*nurulazw21@gmail.com, nartiekaputria@gmail.com

Abstract

2020 is a very difficult year for humans in this world, with a virus outbreak that has swept across the country, this outbreak is called Coronavirus Disease 2019 (Covid-19). Including in Indonesia with cases that continue to increase every day, until now there is no sign of this outbreak ending, with this incident the government issued a policy for large-scale social restrictions or known as PSPB in various regions that fall into the category of large case zones and According to data from the Riau Islands Covid-19 cluster, in Batam City itself, the number of Covid-19 cases continued to increase so that it exceeded the number of 2000. As a result of this protracted disaster, of course, human life has been disrupted, many jobs have been neglected, so that various types of businesses have had to go out of business because their turnover has been much reduced and this has caused business voters to have to lay off their employees. This situation received rapid attention from the government through the social service agency to provide welfare assistance for people affected by COVID-19, but the problem here is from observations while researchers are still many of the assistance that is not right on target, causing polemics for the community. This research will utilize one of the datamining techniques with the K means algorithm to determine the mapping or segmentation of government assistance to communities affected by COVID-19 in order to accelerate people's purchasing power with the aim that research results can be used by community leaders such as RT, RW to record citizens who need this assistance. The results of the analysis show that the clustering segmentation with the data analyzed is 30 data stating that there are priority recipients which are divided into 5 clusters with predetermined criteria.

Keywords : Covid-19;Welfare Assistance;Data Mining;Clustering;K-Means.

Abstrak

2020 merupakan tahun yang sangat berat dihadapi oleh manusia didunia ini, dengan adanya wabah virus yang melanda sampai kepenjuru negeri, wabah ini disebut dengan Coronavirus Disease 2019 (Covid-19). Termasuk di indonesia dengan kasus yang terus meningkat setiap harinya, hingga saat ini belum ada tanda-tanda wabah ini akan berakhir, dengan kejadian ini pemerintah mengeluarkan kebijakan untuk pembatasan sosial berskala besar atau yang dikenal dengan PSPB diberbagai daerah yang masuk kedalam kategori zona kasus yang besardan menetapkan protokol kesehatan, Dikota Batam sendiri kasus Covid-19 terus bertambah sehingga menembus angka 2000, menurut data gugus covid-19 kepulauan Riau. Akibat musibah yang berlarut ini tentunya kehidupan manusia menjadi terganggu, banyak pekerjaan yang terbengkalai hingga berbagai jenis usaha harus gulung tikar karena omset nya jauh berkurang dan menyebabkan pemilih usaha harus memberhentikan karyawannya. Situasi ini mendapat perhatian yang cepat pemerintah melalui dinas sosial memberikan bantuan kesejahteraan bagimasyarakat yang terdampak covid-19 tetapi permasalahannya disini dari hasil pengamatan sementara peneliti masih banyak bantuan tersebut yang tidak tepat sasaran sehingga menimbulkan polemik bagi masyarakat. Penelitian ini akan memanfaatkan salah satu teknik datamining dengan algoritma Kmeans untuk menentukan pemetaan atau segmentasi bantuan pemerintah terhadap masyarakat terdampak covid-19 guna percepatan daya beli masyarakat dengan tujuan hasil penelitian dapat dimanfaatkan oleh pemuka masyarakat seperti RT,RW untuk mendata warganya yang perlu bantuan ini, Hasil analisis menunjukkan segmentasi clustering dengan data yang di analisis sebanyak 30 data menyatakan terdapat prioritas penerima yang di bagi kedalam 5 cluster dengan kriteria yang telah ditentukan.

Kata Kunci : Covid-19;Bantuan Kesejahteraan;Datamining;Clustering;K-Means.

1. Pendahuluan

Tahun 2020 di hebohkan dengan ada fenomena penyebaran wabah asal wuhan salah satu kota di tiangkong yang dikenal dengan Coronavirus Disease 2019 , Infection ini merupakan kelompok yang menyebabkan

beberapa penyakit mulai dari influenza hingga penyakit yang berat seperti center east respiratory syndrome (MERS-CoV) dan extreme intense respiratory syndrome (SARS-CoV).

Dikota Batam sendiri kasus Covid-19 terus bertambah sehingga menembus angka 2000, menurut data gugus menunjukkan bahwa pasien yang dinyatakan sembuh bertambah menjadi 13 orang menjadi total 1.668 orang sedangkan jumlah pasien yang meninggal total 58 orang merupakan data dari 15 Oktober 2020. Sampai saat ini kasus belum menunjukkan berakhir, tentunya hal ini sangat berdampak kepada kehidupan, banyak pekerjaan yang terbengkalai, usaha yang macet, sehingga perekonomian warga menjadi terganggu tidak sedikit dari masyarakat yang gulung tikar usahanya dan PHK dari pekerjaannya karena situasi ini menjadi pelik untuk menyambung kehidupan.

Dengan kejadian ini pemerintah Indonesia telah langsung turun tangan untuk memberikan bantuan bagi warga negara yang terdampak covid-19 baik dari sektor perindustrian, perdagangan dan lain-lain, Meski pemerintah telah mengeluarkan kebijakan bantuan sosial (bansos) dan boost bagi masyarakat namun penerapannya masih belum maksimal. Sebagai contoh, bansos pemerintah masih belum diberikan secara merata kepada masyarakat yang membutuhkan. Belum lagi, system Kartu Pra-Kerja pemerintah dianggap tidak efektif mengantisipasi gelombang pemutusan hubungan kerja (PHK). Menurutny, masyarakat golongan rentan dan hampir miskin ini umumnya bekerja di sektor casual dan banyak yang sangat bergantung pada bantuan-bantuan pemerintah. Masalah ini akan dipecahkan melalui sebuah riset dengan tujuan untuk menganalisa pemetaan atau pembagian bantuan kesejahteraan agar tepat sasaran dengan mengkaji beberapa faktor syarat penerima bantuan, solusi riset ini akan membantu dalam pendataan bantuan terutama para pemuka masyarakat untuk mengelola dan mengorganisir warganya yang benar-benar terdampak covid-19. Metode yang digunakan adalah teknik datamining dengan algoritma klasterisasi K-means yang cocok untuk segmentasi data.

2. Kajian Literatur

2.1 Covid-19

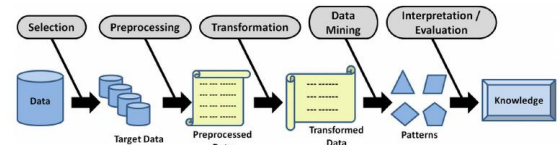
Fenomena ini sangat mengguncang dunia hingga terdampak kesegala penjurus dunia tidak hanya kesehatan manusia yang terganggu tetapi hingga kepada perekonomian dunia ikut merasakan sehingga banyak negara yang menyatakan bahwa mereka mengalami resesi ekonomi termasuk Indonesia, covid-19 ini dinyatakan pada sebuah penelitian bahwasanya *Coronavirus Disease (COVID-19)* merupakan penyakit jenis baru yang belum pernah teridentifikasi pada manusia. Sejak muncul kejadian di Wuhan, China pada

Desember 2019 kemudian diberi nama *Serve Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2)*. COVID-19 disebabkan oleh SARS-CoV-2 yang termasuk dalam keluarga besar corona virus yang sama dengan penyebab SARS pada tahun 2003, hanya berbeda jenis virusnya [1].

2.2 Knowledge Discovery In Database

Penggalan data atau yang sering disebut juga dengan KDD berkali-kali dimanfaatkan dengan cara bergantian dengan tujuan untuk menerangkan proses dari penambangan data dalam suatu tumpukan data dalam jumlah yang sangat besar [2].

Kedua maksud istilah tersebut mempunyai teori yang atau konsep yang berbeda, akan tetapi memiliki keterkaitan satu dengan yang lainnya, dari semua proses KDD tersebut terdapatlah datamining, keterangan untuk memperjelas dapat dilihat pada gambar berikut:



Gambar 1. Knowledge Discovery In Database

2.3 Datamining

Datamining atau teknik penambangan data dalam jumlah yang lebih banyak sudah banyak dimanfaatkan orang dalam berbagai bidang guna mencari informasi baru salah satunya pada penelitian berikut yang mengatakan dalam teorinya Menurut Larose (2005), data mining didefinisikan sebagai sebuah proses untuk menemukan hubungan, pola dan trend baru yang bermakna dengan menyaring data yang sangat besar dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika [3].

Pada penelitian yang pernah dilakukan oleh peneliti sebelumnya tentang analisa faktor penyebab kecelakaan kerja, pada penelitian ini penulis mengambil teori *datamining* merupakan penambangan pengetahuan baru yang akan dijadikan gagasan informasi baru yang diambil dari gunungan data yang berkapasitas banyak atau yang dikenal dengan *database* [4].

Datamining memiliki karakteristik dalam penelitian lainnya yaitu:

1. Mencari pengetahuan dalam pola hubungan antar data dalam *database*.
2. Dimanfaatkan pada jumlah data yang besar agar memperoleh hasil yang efektif dan efisien Sangat bermaaf sekali dalam pengaruh pengambilan keputusan perusahaan [5].

Dengan demikian, datamining membantu mengidentifikasi pola tersembunyi, tren, dan hubungan penting dalam data yang sangat besar, yang pada gilirannya dapat memberikan

wawasan baru yang bernilai dalam berbagai penelitian dan analisis, seperti dalam studi tentang penyebab kecelakaan kerja.

2.4 Klasterisasi

Klasterisasi adalah suatu teknik statistik multivariat yang dirancang untuk menyatakan pengelompokan observasi. Klasterisasi termasuk teknik data mining yang bertujuan mengelompokkan objek data ke dalam klaster yang berbeda sehingga objek data dalam satu klaster lebih mirip dibandingkan objek data pada klaster yang lain, Klasterisasi dapat digunakan untuk mengklasifikasikan daerah-daerah yang padat, menemukan pola-pola distribusi, dan menemukan keterkaitan antara setiap data [6].

Klasterisasi atau *clustering* dalam sebuah penelitian sebelumnya mengatakan *clustering* memegang peranan penting dalam aplikasi data mining, misalnya eksplorasi data ilmu pengetahuan, pengaksesan informasi dan text mining, aplikasi basis data spasial, dan analisis web. *Clustering* diterapkan dalam mesin pencari di Internet. Web mesin pencari akan mencari ratusan dokumen yang cocok dengan kata kunci yang dimasukkan. Dokumen-dokumen tersebut dikelompokkan dalam *cluster-cluster* sesuai dengan kata-kata yang digunakan [7].

Klasterisasi adalah teknik statistik multivariat yang bertujuan untuk mengelompokkan observasi ke dalam kelompok-kelompok. Ini merupakan salah satu teknik dalam data mining yang bertujuan untuk mengelompokkan objek data ke dalam klaster yang berbeda, di mana objek data dalam satu klaster memiliki kesamaan yang lebih tinggi daripada dengan objek data dari klaster lainnya. Klasterisasi dapat digunakan untuk melakukan klasifikasi daerah-daerah padat, menemukan pola distribusi, dan menemukan hubungan antara data-data.

Dalam penelitian sebelumnya, telah diungkapkan bahwa klasterisasi memiliki peran penting dalam aplikasi data mining, seperti eksplorasi data ilmiah, pengaksesan informasi dan text mining, basis data spasial, serta analisis web. Contoh konkret adalah penggunaan klasterisasi dalam mesin pencari internet. Mesin pencari internet akan mencari dan mengelompokkan ratusan dokumen yang relevan dengan kata kunci yang dimasukkan oleh pengguna. Dokumen-dokumen ini dikelompokkan ke dalam klaster sesuai dengan kata-kata yang digunakan.

2.5 K-Means

K-means adalah salah satu metode pengelompokan data nonhierarki yang berusaha untuk mempartisi data yang ada menjadi dua atau lebih kelompok. Metode ini melakukan partisi data ke dalam kelompok

sehingga data dengan karakteristik serupa ditempatkan dalam satu kelompok yang sama, sementara data dengan karakteristik yang berbeda ditempatkan dalam kelompok yang berbeda. Tujuan dari pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang diatur dalam proses pengelompokan.

Dengan demikian, klasterisasi adalah alat penting untuk mengenali pola dan relasi dalam data yang kompleks, dan memiliki berbagai aplikasi yang signifikan dalam berbagai bidang.

Tujuan ini umumnya berfokus pada upaya untuk mengurangi variasi di dalam satu kelompok dan memaksimalkan variasi antar kelompok. [8].

Berikut ini adalah rumus untuk menentukan jumlah *cluster*:

$$K \approx \sqrt{n/2} \quad (1)$$

Berikut rumus pengukuran jarak:

$$d(x,y) = \|x-y\|^2 = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2)$$

Adapun rumus perhitungan jarak lainnya didefinisikan sebagai berikut:

$$d_{(x,y)} = \sqrt{(x_i - y_i)^2 + (x_i - y_i)^2} \quad (3)$$

Di mana:

d = titik dokumen

x = *data record*

y = *data centroid*

Jarak yang terpendek antara *centroid* dengan dokumen menentukan posisi *cluster* suatu dokumen. Misalnya dokumen A mempunyai jarak paling pendek ke *centroid 1* dibanding ke yang lain, maka dokumen A masuk ke *group 1*. Hitung kembali posisi *centroid* baru untuk tiap-tiap *centroid* dengan mengambil rata-rata dokumen yang masuk pada *cluster* awal. Iterasi dilakukan terus hingga posisi *group* tidak berubah.

Berikut rumus dari penentuan *centroid* :

$$C(i) = \frac{1}{|G_i|} \sum_{x \in G_i} x \quad (4)$$

rumus iterasi lainnya didefinisikan sebagai berikut:

$$C(i) = \frac{x_1 + x_2 + x_n + x_{...}}{\sum x} \quad (5)$$

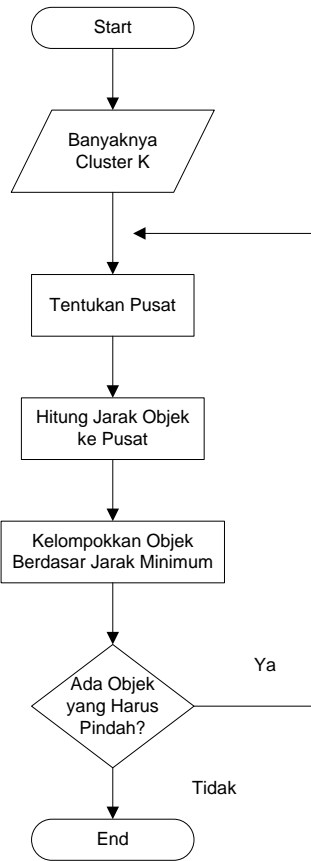
Di mana:

x_1 = nilai data record ke-1

x_2 = nilai data record ke-2

$\sum x$ = jumlah data record

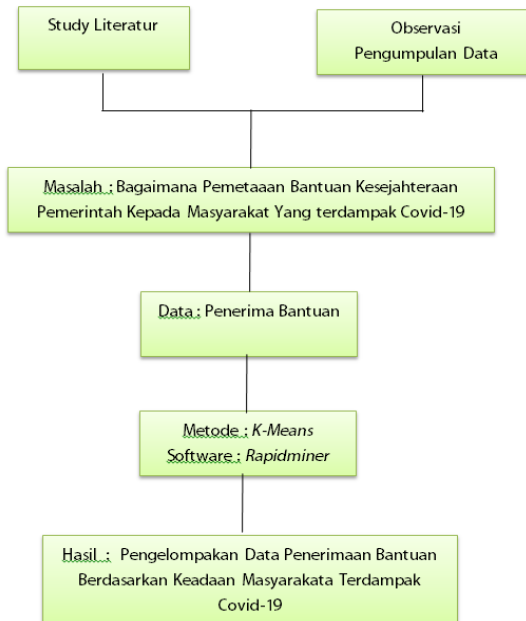
Berikut adalah diagram alir dari proses algoritma *K-means* :



Gambar 2. Diagram Alir K-Means

3. Metode Penelitian

Penelitian ini akan di laksanakan sesuai dengan kerangka dibawah ini :



Gambar 3. Kerangka Penelitian

4. Hasil

4.1 Proses Clustering Menggunakan Algoritma K-Means

Data yang sudah ditetapkan akan dilakukan pengolahan data dengan proses clustering

dengan menggunakan algoritma *K-Means*. Langkah-langkahnya adalah sebagai berikut ini:

1. Penentuan jumlah cluster

Penentuan jumlah cluster dilakukan untuk segmentasi pemakaian material. Maka dalam penelitian ini jumlah cluster yang digunakan adalah sebanyak 5 cluster ($k=5$), sehingga nanti akan diketahui cluster terbaik dalam melakukan penelitian ini.

2. Menentukan centroid

Pusat awal cluster atau centroid ditentukan secara random atau acak. Peneliti menentukan nilai cluster 1 diambil dari data ke-3, nilai cluster 2 pada data ke-9, nilai cluster 3 pada data ke-17, nilai cluster 4 pada data ke-23 dan nilai cluster 5 pada data ke-29. Berikut adalah nilai centroid awal pada segmentasi bantuan kesejahteraan sosial terdampak covid-19 :

Cluster 1 : (2;2;2;1;2)

Cluster 2 : (6;14;3;2;2)

Cluster 3 : (5;14;3;2;3)

Cluster 4 : (3;13;4;3;2)

Cluster 5 : (5;13;3;4;2)

3. Menghitung jarak dari centroid

Menghitung jarak antara titik centroid dengan titik tiap objek dengan menggunakan *Euclidian Distance*.

$$\text{Centroid 1} = \sqrt{(1-2)^2 + (10-2)^2 + (1-2)^2 + (2-1)^2 + (3-2)^2}$$

$$= \sqrt{68,0}$$

$$= 10,12$$

$$\text{Centroid 2} = \sqrt{(1-6)^2 + (10-14)^2 + (1-3)^2 + (2-2)^2 + (3-2)^2}$$

$$= \sqrt{46,0}$$

$$= 7,7$$

$$\text{Centroid 3} = \sqrt{(1-5)^2 + (10-14)^2 + (1-3)^2 + (2-2)^2 + (3-3)^2}$$

$$= \sqrt{36,0}$$

$$= 6,0$$

$$\text{Centroid 4} = \sqrt{(1-3)^2 + (10-13)^2 + (1-4)^2 + (2-3)^2 + (3-2)^2}$$

$$= \sqrt{24,0}$$

$$= 6,7$$

$$\text{Centroid 5} = \sqrt{(1-5)^2 + (10-13)^2 + (1-3)^2 + (2-4)^2 + (3-2)^2}$$

$$= \sqrt{34,0}$$

$$= 10,4$$

4. Alokasikan masing-masing objek ke centroid terdekat, untuk melakukan mengalokasikan objek ke dalam masing-masing cluster dengan cara mengelompokkan berdasarkan jarak minimum objek ke pusat cluster dengan memberikan kode "1" jika hasil cluster mendekati nol. Perhitungan dilakukan terus sampai data ke-30 terhadap pusat cluster. Setelah dilakukan proses perhitungan maka akan didapatkan data seperti tabel.

Tabel 1. Hasil Perhitungan Iterasi Ke-1

	C2	C3	C4	C5	Jarak Terdekat	Kode			
10,1	7,7	6,0	6,7	10,4	6,0			1	
1,0	12,7	13,4	14,2	19,5	1,0	1			
-	13,7	14,4	15,2	20,4	-	1			
8,2	7,7	10,3	7,1	10,4	7,1				1
8,2	7,7	10,3	7,1	10,4	7,1				1
12,3	5,5	6,1	3,3	4,2	3,3				1
13,8	1,0	2,4	4,2	5,7	1,0		1		
26,6	9,4	6,7	9,1	6,4	6,4				1
13,7	-	2,0	4,3	5,4	-		1		
20,2	7,8	6,2	3,7	2,4	2,4				1
8,6	11,9	10,5	9,1	10,6	8,6				1
7,1	9,1	7,6	8,2	11,7	7,1	1			
8,1	9,9	10,5	7,1	8,6	7,1				1
12,1	6,1	6,6	3,6	4,7	3,6				1
15,2	3,4	3,7	2,2	2,4	2,2				1
13,2	3,4	5,7	4,2	6,4	3,4		1		
14,4	2,0	-	4,4	6,0	-			1	
13,5	1,4	2,0	3,2	5,4	1,4		1		
22,9	6,2	7,4	4,7	3,0	3,0				1
20,4	8,0	6,4	3,4	2,7	2,7				1
17,3	8,5	9,1	4,3	3,2	3,2				1
25,5	8,2	7,3	7,4	5,2	5,2				1
15,2	4,3	4,4	-	3,2	-				1
15,2	3,2	3,4	1,4	2,0	1,4				1
3,0	13,6	12,4	13,1	16,4	3,0	1			
1,0	14,7	13,4	16,2	21,4	1,0	1			
24,2	7,6	8,2	5,2	4,1	4,1				1
26,6	9,0	6,4	8,7	6,2	6,2				1
20,4	5,4	6,0	3,2	-	-				1
13,5	4,2	5,0	3,0	3,0	3,0				1

5. Setelah dilakukan iterasi ke-1 maka lakukan iterasi ke-2 untuk membandingkan letak posisi hasil *cluster* 1, *cluster*, *cluster* 3, *Cluster* 4 dan *Cluster* 5. Jika posisinya tidak berubah maka iterasi berikutnya tidak perlu dilakukan lagi. Untuk melakukan iterasi ke-2 maka tentukan *cluster* baru terlebih dahulu. Didapatkan hasil *cluster* baru sebagai berikut :
- Cluster 1 : (2;3;2;2,4;1,4;2,6)
 Cluster 2 : (5;25;13,75;3,25;2;1,75)
 Cluster 3 : (3;12;2;2;3)
 Cluster 4 : (3;11;9,56;3;2,78;1,89)
 Cluster 5 : (5;13,70;3,10;3,90;2,70)
6. Setelah didapatkan titik pusat yang baru dari setiap *cluster*, lakukan kembali dari langkah ketiga hingga titik pusat dari setiap *cluster* tidak berubah lagi dan tidak ada lagi data yang berpindah dari satu *cluster* ke *cluster* yang lain.

Didapatkan hasil *cluster* baru sebagai berikut :

- ✓ Cluster 1 : (2;3;5;2,7;1,7;2,7)
- ✓ Cluster 2 : (4,7;13,6;3,3;2,3;2,0)
- ✓ Cluster 3 : (1;10;1;2;3)
- ✓ Cluster 4 : (3,4;9;2,9;2,9;1,7)
- ✓ Cluster 5 : (5;14,3;3;3,9;2,8)

7. Setelah didapatkan titik pusat yang baru dari setiap *cluster*, lakukan kembali dari langkah ketiga hingga titik pusat dari setiap *cluster* tidak berubah lagi dan tidak ada lagi data yang berpindah dari satu *cluster* ke *cluster* yang lain. Didapatkan hasil *cluster* baru sebagai berikut :
- ✓ Cluster 1 : (2;3;5;2,67;1,67;2,67)
 - ✓ Cluster 2 : (4,63;13,50;3,13;2,38;2)
 - ✓ Cluster 3 : (1;10;1;2;3)
 - ✓ Cluster 4 : (3,4;9;2,8;2,8;1,7)
 - ✓ Cluster 5 : (5,13;14,50;3,13;4,2,8)

Tabel 2. Hasil Perhitungan Iterasi Ke-4

C1	C2	C3	C4	C5	Jarak Terdekat	Kode			
7,0	6,6	-	5,6	10,5	-			1	
1,7	12,8	9,1	9,8	21,7	1,7	1			
2,5	13,7	10,1	10,7	22,7	2,5	1			
7,5	6,9	7,5	2,3	14,4	2,3				1
7,5	6,9	7,5	2,3	14,4	2,3				1
9,0	3,9	5,6	1,3	6,4	1,3				1
11,9	1,8	8,1	6,6	6,1	1,8		1		
19,4	9,0	12,1	13,1	2,8	2,8				1
11,8	1,6	7,7	6,4	5,8	1,6		1		
14,3	5,3	8,1	6,0	2,8	2,8				1
3,9	10,3	6,9	6,1	11,1	3,9	1			
3,738	8,2	3,742	4,8	12,1	3,7	1			
4,7	8,3	6,6	3,4	10,8	3,4				1
8,8	4,4	4,2	1,5	6,9	1,5				1
12,0	1,8	6,4	4,2	4,0	1,8				
12,6	2,5	8,4	5,4	9,7	2,5		1		
11,1	1,8	6,0	7,6	4,5	1,8		1		
11,6	1,2	7,4	6,2	5,8	1,2		1		
18,3	5,4	13,1	8,3	2,9	2,9				1
14,4	5,5	8,7	6,2	2,9	2,9				1
12,7	6,2	8,6	2,5	5,4	2,5				1
19,7	7,6	12,7	11,4	2,9	2,9				1
11,9	2,3	6,7	4,3	4,5	2,3		1		
11,9	1,2	6,7	4,1	3,6	1,2		1		
1,8	12,9	8,3	9,5	16,9	1,8	1			
2,2	14,7	9,1	12,3	21,9	2,2	1			
19,5	6,2	12,9	9,4	3,5	3,5				1
19,4	8,7	12,3	13,1	2,3	2,3				1
15,9	3,3	10,4	5,7	2,3	2,3				1
10,3	2,9	6,6	2,6	5,3	2,6				1

Pada tabel 1 dan 2 letak posisi hasil *cluster* 1, *cluster* 2, *cluster* 3, *cluster* 4 dan *cluster* 5 telah sama. Maka, perhitungan selesai sampai pada iterasi ke 4.

4.2 Segmentasi Prioritas Bantuan Kesejahteraan

Dari hasil perhitungan menggunakan algoritma *K-Means* untuk segmentasi pemakaian material dengan 3 *cluster* didapat pengelompokan seperti pada tabel.

Tabel 3. Hasil Akhir Segmentasi

Id	Pendidikan	Pekerjaan	Jumlah Tanggungan	Pendapatan	Kepemilikan Rumah
<i>Cluster 1 : Prioritas Layak 1</i>					
2	2	3	2	1	2
3	2	2	2	1	2
10	2	5	4	3	3
11	2	7	3	2	3
25	2	2	3	2	3

Id	Pendidikan	Pekerjaan	Jumlah Tanggungan	Pendapatan	Kepemilikan Rumah
26	2	2	2	1	3
<i>Cluster 2 : Prioritas Layak 2</i>					
6	6	14	4	2	2
9	6	14	3	2	2
15	4	13	2	3	2
16	4	13	2	2	1
17	5	14	3	2	3
18	5	14	4	2	2
23	3	13	4	3	2
24	4	13	3	3	2
<i>Cluster 3 : Prioritas Layak 3</i>					
1	1	10	1	2	3
<i>Cluster 4 : Prioritas Layak 4</i>					
4	3	8	3	2	1
5	3	8	3	2	1
6	4	10	3	3	2

Id	Pendidikan	Pekerjaan	Jumlah Tanggungan	Pendapatan	Kepemilikan Rumah
13	2	6	3	3	2
14	3	10	2	3	2
21	4	10	3	4	2
30	5	11	3	3	2
<i>Cluster 5 : Prioritas Layak 5</i>					
8	6	15	2	4	4
10	4	12	3	4	3
19	6	15	5	4	2
20	4	12	4	4	3
22	6	17	2	4	3
27	4	17	3	4	2
28	6	15	3	4	4
29	5	13	3	4	2

5. Simpulan

- Penerapan pendekatan data mining dengan algoritma K-Means *clustering* mampu untuk mengelompokkan data penduduk penerima bantuan kesejahteraan terdampak covid-19 dan memiliki tingkat akurasi yang cukup tinggi.
- Hasil analisis menunjukkan segmentasi clustering dengan data yang di analisis sebanyak 30 data menyatakan terdapat prioritas penerima yang di bagi kedalam 5 *cluster* dengan kriteria yang telah ditentukan. cluster 1 terdapat 6 data cluster 2 : 8 data, cluster 3 : 1 data, cluster 4 : 7 data dan cluster 5 : 8 data , hasil tersebut dapat dijadikan prioritas bagi penerima bantuan kesejahteraan.

DAFTAR PUSTAKA

- A. Mahmudan, "Clustering of District or City in Central Java Based COVID-19 Case Using K-Means Clustering," *J. Mat. Stat. dan Komputasi*, vol. 17, no. 1, pp. 1–13, 2020, doi: 10.20956/jmsk.v17i1.10727.
- Y. S. Luvia, A. P. Windarto, S. Solikhun, and D. Hartama, "Penerapan Algoritma C4.5 Untuk Klasifikasi Predikat Keberhasilan Mahasiswa Di Amik Tunas Bangsa," *Jurasik (Jurnal Ris. Sist. Inf. dan Tek. Inform.)*, vol. 1, no. 1, p. 75, 2018, doi: 10.30645/jurasik.v1i1.12.
- R. H. Pambudi, B. D. Setiawan, and Indriati, "Penerapan Algoritma C4 . 5 Untuk Memprediksi Nilai Kelulusan Siswa Sekolah Menengah Berdasarkan Faktor Eksternal," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 7, pp. 2637–2643, 2018, [Online]. Available: <http://j-ptiik.ub.ac.id>.
- E. Elisa, "Analisa dan Penerapan Algoritma C4.5 Dalam Data Mining Untuk

Mengidentifikasi Faktor-Faktor Penyebab Kecelakaan Kerja Kontruksi PT.Arupadhatu Adisesanti," *J. Online Inform.*, vol. 2, no. 1, p. 36, 2017, doi: 10.15575/join.v2i1.71.

- Meilani, B. Dwi, and W. Azmuri, "Penentuan Pola Yang Sering Muncul Untuk Penerima Kartu Jaminan Kesehatan Masyarakat," *Semin. Nas. "Inovasi dalam Desain dan Teknol.*, vol. x, no. x, pp. 424–431, 2015.
- G. D. Rembulan, T. Wijaya, D. Palullungan, K. N. Alfina, and M. Qurthuby, "Kebijakan Pemerintah Mengenai Coronavirus Disease (COVID-19) di Setiap Provinsi di Indonesia Berdasarkan Analisis Klaster," *JIEMS (Journal Ind. Eng. Manag. Syst.)*, vol. 13, no. 2, 2020, doi: 10.30813/jiems.v13i2.2280.
- R. A. Asroni, "Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang," *Ilm. Semesta Tek.*, vol. 18, no. 1, pp. 76–82, 2015.
- L. Maulida, "Penerapan Datamining Dalam Mengelompokkan Kunjungan Wisatawan Ke Objek Wisata Unggulan Di Prov. DKI Jakarta Dengan K-Means," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 2, no. 3, p. 167, 2018, doi: 10.14421/jiska.2018.23-06.