

# Penerapan Normalisasi Data Dalam Mengelompokkan Data Mahasiswa Dengan Menggunakan Metode K-Means Untuk Menentukan Prioritas Bantuan Uang Kuliah Tunggal

Muhammad Rafli Kusnaldi, Timotius Gulo, Soeb Aripin\*

Fakultas Ilmu Komputer Dan Teknologi Informasi, Teknologi Informasi, Universitas Budi Darma, Medan, Indonesia

Email: <sup>1</sup>rafli.kusnaldi@gmail.com, <sup>2</sup>timotiusgulo@gmail.com, <sup>3,\*</sup>soebaripin@gmail.com

Submitted: 15/08/2022; Accepted: 24/08/2022; Published: 30/08/2022

**Abstrak**-Pada Universitas Budi Darma terdapat kendala dalam memberikan bantuan UKT dimana labat dan kurang sasaran bagi mahasiswa yang mendapatkan. Hal itu terjadi dikarenakan yang layak mendapatkan bantuan ini adalah mahasiswa yang kesulitan biaya oleh sebab itu diperlukan suatu cara dengan melakukan pengelompokan data mahasiswa berdasarkan jenjang sosial yang dimiliki. Dalam menentukan mahasiswa/mahasiswi yang layak mendapatkan bantuan dapat memanfaatkan data-data mahasiswa yang sedang menjalani studi belajarnya di Universitas Budi Darma. Dengan menggali informasi berdasarkan data-data mahasiswa. Agar data tersebut dapat digunakan terlebih dahulu maka dilakukan normalisasi data agar memperoleh data yang lebih akurat. Dimana data mahasiswa dapat dikelompokkan dengan tepat, maka harus dilakukan normalisasi data. Salah satu metode normalisasi yang sering digunakan dalam melakukan normalisasi data adalah Metode decimal scaling yang merupakan metode transformasi data dengan normalisasi untuk menyamakan rentang nilai pada setiap atribut dengan skala tertentu dengan menggerakkan nilai desimal dari data ke arah yang diinginkan. Setelah data ternormalisasi maka proses selanjutnya menggali informasi data mahasiswa dengan menerapkan data mining. Penerapan data mining dilakukan untuk mendapatkan informasi berupa kelompok data mahasiswa yang digunakan sebagai prioritas dalam mendapatkan bantuan UKT. Adapun metode yang digunakan dalam mengelompokkan data mahasiswa dengan menggunakan algoritma K-Means. Adapun pengujian metode secara manual yaitu terdapat 3 cluster dimana jumlah cluster 0 cluster 1 dan cluster 3 sama dengan pengujian aplikasi data mining yaitu rapidminer sehingga yang layak menjadi prioritas mendapatkan bantuan uang kuliah berdasarkan sampel yaitu cluster/pengelompokan 0 yang terdiri dari 22 orang. Penelitian ini bertujuan untuk melihat pengaruh penerapan normalisasi data dalam metode K-Means untuk mengelompokkan data mahasiswa yang digunakan sebagai rekomendasi dalam pemilihan pemberian bantuan UKT

**Kata Kunci:** Normalisasi; Data; Pengelompokan; Prioritas; UKT; K-Means

**Abstract**-At Budi Darma University there are obstacles in providing UKT rocks where it is profitable and less targeted for students who get it. This happened because those who deserved this assistance were students who had difficulty in costs, therefore we needed a way by grouping student data based on their social level. In determining the students who deserve to get the rock, they can use the data of students who are undergoing their studies at Budi Darma University. By digging up information based on student data. So that the data can be used first, the data normalization is carried out in order to obtain more accurate data. Where student data can be grouped correctly, data normalization must be carried out. One of the normalization methods that are often used in normalizing data is the decimal scaling method which is a data transformation method with normalization to equalize the range of values on each attribute with a certain scale by moving the decimal value from data in the desired direction. After the data is normalized, the next process is to explore student data information by applying data mining. The application of data mining is carried out to obtain information in the form of student data groups that are used as a priority in obtaining UKT assistance. The method used in classifying student data is using the K-Means algorithm. The manual testing method is that there are 3 clusters where the number of clusters 0 cluster 1 and cluster 3 is the same as testing data mining applications, namely rapidminer so that those who deserve to be prioritized get tuition assistance based on the sample, namely cluster / grouping 0 which consists of 22 people. This study aims to see the effect of applying data normalization in the K-Means method to classify student data which is used as a recommendation in the selection of UKT assistance.

**Keywords:** Normalization; Data; Grouping; Priority; UKT; K-Means

## 1. PENDAHULUAN

Uang Kuliah Tunggal (UKT) merupakan suatu sistem yang diterapkan dalam melakukan pembiayaan perkuliahan yang ditanggung oleh mahasiswa. Perguruan Tinggi Negeri yang berada di bawah Kementerian Riset, Teknologi, dan Pendidikan Tinggi (Kemristek Dikti). Dengan tujuan yaitu membantu orang tua mahasiswa yang pekerjaannya terkena dampak Covid-19 sehingga mahasiswa tersebut tetap dapat melanjutkan studinya tanpa terhalangi kendala biaya [1]. Bantuan Uang Kuliah Tunggal (UKT) ini di distribusikan kepada masyarakat melalui perguruan tinggi, dimana Pemerintah memberikan tanggung jawab kepada pihak suatu Perguruan tinggi untuk dapat memilih mahasiswa yang layak untuk mendapat bantuan Uang Kuliah Tunggal (UKT) tersebut salah satu perguruan tinggi yang mendapatkan bantuan UKT adalah Universitas Budi Darma Medan.

Pada Universitas Budi Darma terdapat kendala dalam memberikan bantuan UKT dimana labat dan kurang sasaran bagi mahasiswa yang mendapatkan. Hal itu terjadi dikarenakan yang layak mendapatkan bantuan ini adalah mahasiswa yang kesulitan biaya oleh sebab itu diperlukan suatu cara dengan melakukan pengelompokan data mahasiswa berdasarkan jenjang sosial yang dimiliki. Dalam menentukan mahasiswa/mahasiswi yang layak mendapatkan bantuan dapat memanfaatkan data-data mahasiswa yang sedang menjalani studi belajarnya di Universitas Budi Darma. Dengan menggali informasi berdasarkan data-data mahasiswa. Agar data tersebut dapat

digunakan terlebih dahulu maka dilakukan normalisasi data agar memperoleh data yang lebih akurat. Diman data mahasiswa dapat dikelompokkan dengan tepat, maka harus dilakukan normalisasi data.

Metode normalisasi data adalah proses membuat beberapa variabel memiliki rentang nilai yang sama, tidak ada yang terlalu besar maupun terlalu kecil sehingga dapat membuat analisis statistik menjadi lebih mudah. Adapun metode dalam melakukan normalisasi data yaitu Min-Max normalization, Z-score Normalization, dan Decimal Scaling Normalization[2]. Salah satu metode normalisasi yang sering digunakan dalam melakukan normalisasi data adalah Metode decimal scaling yang merupakan metode transformasi data dengan normalisasi untuk menyamakan rentang nilai pada setiap atribut dengan skala tertentu dengan menggerakkan nilai desimal dari data ke arah yang diinginkan. Agar dapat menghasilkan data yang lebih baik. Berdasarkan penelitian terdahulu menyebutkan bahwapembedaan rentang nilai pada setiap atribut dalam proses transformasi datamenyebabkan tidak berfungsinya atribut yang memiliki nilai jauh lebih kecil dibandingkan dengan atribut-atribut lainnya, sehingga data yang telah dinormalisasi dengan metode decimal scaling lebih baik hasilnya secara signifikan daripada klasifikasi tanpa melakukan transformasi data[3].

Setelah data ternormalisasi maka proses selanjutnya menggali informasi data mahasiswa dengan menerapkan datamining. Penerapan datamining dilakukan untuk mendapatkan informasi berupa kelompok data mahasiswa yang digunakan sebagai prioritas dalam mendapatkan bantuan UKT. Adapun metode yang digunakan dalam mengelompokkan data mahasiswa dengan menggunakan algoritma K-Means. Algoritma K-Means adalah algoritma yang digunakan pada algoritma Clustering dimana penerapan algoritma ini dilakukan dengan mengelompokkan data kedalam beberapa kelompok/partisi dengan tujuan untuk menggali informasi pada setiap kelompok/partisi[4]. Adapun penelitian terkait dengan problem permasalahan dan metode yang digunakan dalam penelitian ini yang dijadikan sebagai bahan acuan.

Penelitian yang dilakukan oleh Tuti Hartati, Dkk pada tahun 2021 dimana penerapan algoritma K-Means dalam melakukan pengelompokan wilayah promosi kampus berdasarkan sampel data set sebanyak 675 dengan hasil jumlah cluster sebanyak 30 nilai paling mendekati 0 adalah  $k = 29$  dengan Davies bouldin: 0.070, dengan persebaran anggota cluster terbanyak ada pada cluster 16 berisi anggota Cluster 115 items[5]. Penelitian yang dilakukan Abu salam, Dkk pada tahun 2020 tentang penerapan Algoritma K-Means dalam pengklasteran untuk rekomendasi penerimaan beasiswa PPA di uдинus, algoritma K-Means dapat menghasilkan rekomendasi penerima beasiswa Peningkatan Prestasi Akademik (PPA) dengan melibatkan 7 atribut dan dengan jumlah dataset yang banyak dengan jumlah data set 441, sebanyak 154 mahasiswa direkomendasikan mendapatkan beasiswa PPA sedangkan 287 lainnya tidak mendapatkan beasiswa[6]. Penelitian yang dilakukan oleh Nurul Rohmawati, Dkk pada tahun 2015 tentang implementasi algoritma K-Means dalam pengklasteran mahasiswa pelamar beasiswa, Dimana dengan perbandingan perbandingan nilai purity measure pada hasil clustering dari algoritma k-means, Diketahui nilai purity pada dataset data kodifikasi sebagian untuk hasil cluster algoritma k-means sebesar 0.611 atau 61.11%. Pada dataset kodifikasi keseluruhan nilai purity hasil cluster algoritma k-means sebesar 0.806 atau sebesar 80.56%. Untuk dataset data asli nilai purity hasil cluster algoritma k-means sebesar 0.750 atau 75%. Maka dapat disimpulkan bahwa tingkat akurasi clustering dari hasil cluster algoritma k-means berdasarkan nilai purity measure, dataset yang dikodifikasi keseluruhan lebih baik dari pada dataset yang di kodifikasi sebagian dan dataset data asli[7].

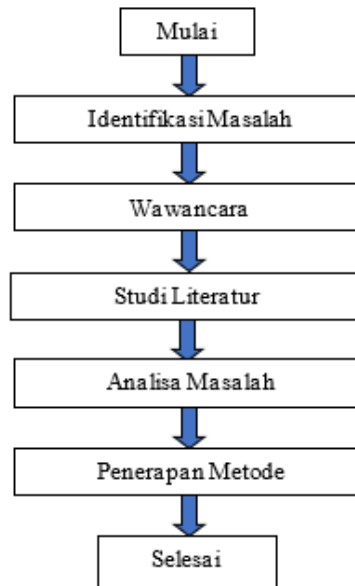
Berdasarkan permasalahan dan penelitian terdahulu yang dijadikan acuan maka penelitian ini bertujuan untuk melihat pengaruh normalisasi data dalam penerapan algoritma K-Means untuk mengelompokkan data mahasiswa yang digunakan sebagai rekomendasi dalam pemilihan pemberian bantuan UKT.

## **2. METODOLOGI PENELITIAN**

### **2.1 Tahapan Penelitian**

Pada bagian tahapan penelitian dalam pengelompokan data mahasiswa penerima bantuan Uang Kuliah Tunggal dengan menggunakan metode K-Means dilakukan beberapa tahapan penelitian seperti dibawah ini:

1. Identifikasi masalah, pada tahapan ini dilakukan untuk mengetahui permasalahan dan metode yang digunakan pada penelitian ini.
2. Wawancara (*interview*), ditahap ini melakukan konsultasi atau tanya jawab secara langsung kepada pihak Perguruan Tinggi Swasta
3. Studi Kepustakaan adalah tahapan dimana peneliti melakukan salah satu cara atau teknik dalam mencari dan menemukan fakta-fakta atau informasi dengan cara melakukan tinjauan pustaka dari berbagai sumber atau referensi yang nyata.
4. Analisa masalah, pada tahap analisa ini penulis melakukan pengumpulan data, mempelajari dan melakukan perumusan demi mendukung penelitian ini dalam melakukan proses pengolahan data.
5. Penerapan Metode, pada tahap ini merupakan tahap normalisasi data dengan pengelompokan data mahasiswa dengan metode K-Medoids yang dimana hasil akhir dari penelitian ini menghasilkan Kelompok data siswa berdasarkan sosial ekonomi orang tua setiap mahasiswa.



Gambar 1. Tahapan Penelitian

## 2.2 Bantuan Uang Kuliah Tunggal

Bantuan Uang kuliah Tunggal (UKT) diperuntukan bagi mahasiswa aktif yang mengalami kondisi keuangan sehingga memerlukan bantuan bukan mahasiswa yang telah mendapatkan Beasiswa baik dari Pemerintah maupun Swasta. Yang diatur dalam Ketentuan Kementerian Pendidikan dan Kebudayaan Nomor 25 Tahun 2020 tentang Ketentuan Penyesuaian UKT, Dana Bantuan UKT Mahasiswa, BOS Afiriasi dan BOS Kinerja[8].

## 2.3 Data Mining

Data mining adalah proses untuk menemukan pola, korelasi dan tren baru yang memiliki bermakna yang bertujuan untuk mendapatkan informasi baru dari kumpulan data yang besar[1]. Banyak orang mengartikan data mining sebagai persamaan dari istilah yang populer seperti Knowledge Discovery Data (KDD), dan yang lain mengartikan data mining hanya sebagai langkah penting dalam proses penemuan pengetahuan[9].

## 2.4 Normalisasi Data

Merupakan suatu kumpulan file data yang saling terhubung dan juga dikendalikan secara terpusat. Database mengkonsolidasikan banyak records yang dahulu disimpan di dalam file yang terpisah-pisah ke dalam satu penampungan records, sehingga bisa digunakan oleh berbagai users dan aplikasi pengolah data.[10]

Decimal scalling merupakan metode normalisasi dengan menggerakkan nilai decimal dari data ke arah yang diinginkan[2]. Metode decimal scalling bekerja dengan cara menggeser titik desimal pada nilai data. Banyaknya penggeseran titik desimal bergantung pada nilai maksimum absolut dari masing-masing fitur data atau variabel[3]. Decimal scalling digunakan ketika nilai berada dalam rentang logaritmik, misalnya rentang suatu nilai adalah antara 0 dan 1 sedangkan nilai lain pada rentang 0 dan 1000[11]. Adapun rumus decimal scalling adalah sebagai berikut :

$$\text{newdata} = \frac{\text{data}}{i} \tag{1}$$

Keterangan :

Newdata = Data hasil normalisasi

i = Adalah nilai max dari kriteria

## 2.5 Metode K-Means

Metode K-Means adalah salah satu teknik dalam data mining untuk mengelompokan (Clustering) data kedalam beberapa kelompok berdasarkan jarak, kriteria, kondisi atau karakteristik. Data dalam satu kelompok harus memiliki jarak terpendek, kriteria, kondisi atau karakteristik yang sama atau hampir sama antara satu dengan lainnya. Algoritma K-Means dapat mengelompokan objek yang memiliki kemiripan[6]. Adapun langkah-langkah dalam pengelompokan data dengan Algoritma K-Means adalah sebagai berikut:

1. Menyiapkan sampling data
2. Menentukan jumlah cluster
3. Menentukan nilai centroid/titik pusat
4. Menghitung jarak masing-masing centroid
5. Mengelompokan data berdasarkan jarak terpendek

Untuk Menentukan jarak masing-masing *centroid* berdasarkan jarak terpendek, kriteria, kondisi atau karakteristis yang sama atau hampir sama, menggunakan model Euclidean yang dirumuskan sebagai berikut:

$$d_{ij} = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + \dots + (x_{ki} - x_{kj})^2} \tag{2}$$

Keterangan:

$d_{ij}$  = Jarak dari data ke i ke pusat cluster j

$x_{ki}$  = Data dari ke - i pada *attribute* data ke - k

$x_{kj}$  = Data dari ke - j pada *attribute* data ke - k

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Pembahasan

Dalam menentukan prioritas mahasiswa yang layak mendapatkan rekomendasi bantuan UKT dilakukan dengan cara mengelompokkan data penerima bantuan uang kuliah tunggal bagi mahasiswa terdampak Covid-19. Sampel data yang digunakan sebanyak 100 mahasiswa yang diambil secara acak. Adapun sampel data yang digunakan dapat dilihat pada tabel dibawah ini:

Tabel 1. Data Sampel

No	Nama	Pekerjaan Orang Tua	Status Kepemilikan Rumah	Penghasilan Orang Tua	
1	Atira Nabila	Nelayan	Sewa	2.500.000	
2	Sawitri Meliana Sabet	Wiraswasta	Sewa	18.000.00	
3	Tambunan	Petani	Milik Sendiri	2.000.000	
4	Desi Novria Siregar	Petani	Milik Sendiri	1.800.000	
5	Handayani Simanjuntak	Wiraswasta	Milik Sendiri	2.500.000	
6	Mardiana Pasaribu	Petani	Sewa	18.000.00	
7	Yamolala Hulu	PNS	Milik Sendiri	5.000.000	
8	David Ridwan Hakim Misael Oktavianda	Wiraswasta	Sewa	3.000.000	
9	Harefa	Petani	Milik Sendiri	1.400.000	
10	Muhammad Ali	Petani	Milik Sendiri	2.000.000	
....	.....	.....	.....	.....	
....	.....	.....	.....	.....	
10	0	Siti Rugun Manurung	Petani	Milik Sendiri	2.000.000

Maka proses selanjutnya melakukan transformasi data agar data sampel data dilakukan normalisasi data. Adapun hasil proses transformasi data sebagai berikut:

Tabel 2. Data Sampel Hasil Transformasi

No	Nama	Pekerjaan Orang Tua	Status Kepemilikan Rumah	Penghasilan Orang Tua	
1	Atira Nabila	1	1	2.500.000	
2	Sawitri Meliana Sabet	2	1	18.000.00	
3	Tambunan	1	2	2.000.000	
4	Desi Novria Siregar	1	2	1.800.000	
5	Handayani Simanjuntak	2	2	2.500.000	
6	Mardiana Pasaribu	1	1	18.000.00	
7	Yamolala Hulu	3	2	5.000.000	
8	David Ridwan Hakim Misael Oktavianda	2	1	3.000.000	
9	Harefa	2	2	1.400.000	
10	Muhammad Ali	1	2	2.000.000	
....	.....	.....	.....	.....	
....	.....	.....	.....	.....	
10	0	Siti Rugun Manurung	1	2	2.000.000

**Keterangan:**

1. Pekerjaan Orang Tua
  - a. Variabel Nelayan dan Petani =1
  - b. Variabel Wiraswasta =2
  - c. Variabel PNS = 3
2. Status Kepemilikan Rumah
  - a. Variabel Sewa =1
  - b. Variabel Milik Sendiri =2

Adapun proses perhitungan metode desimal scaling untuk normalisasi data mahasiswa pada Universitas Budi Darma sebagai berikut:

1. Normalisasi kriteria Pekerjaan Orang Tua

$$\text{Newdata}_1 = 1/3 = 0,33$$

2. Normalisasi Status Kepemilikan Rumah

$$\text{Newdata}_1 = 1/2 = 0,5$$

3. Normalisasi Penghasilan Orang Tua

$$\text{Newdata}_1 = 2.500.000/5000.000 = 25000$$

Untuk melihat hasil normalisasi pada setiap kriteria yang digunakan dapat dilihat pada tabel dibawah ini:

**Tabel 3.** Data Sampel Hasil Normalisasi Data

No	Nama	Pekerjaan Orang Tua	Status Kepemilikan Rumah	Penghasilan Orang Tua
1	Atira Nabila	0,333333333	0,5	0,5
2	Sawitri	0,666666667	0,5	0,36
	Meliana Sabet	0,333333333	1	0,4
3	Tambunan			
4	Desi Novria Siregar	0,333333333	1	0,36
5	Handayani Simanjuntak	0,666666667	1	0,5
6	Mardiana Pasaribu	0,333333333	0,5	0,36
7	Yamolala Hulu	1	1	1
8	David Ridwan Hakim	0,666666667	0,5	0,6
	Misael Oktavianda	0,666666667	1	0,28
9	Harefa			
10	Muhammad Ali	0,333333333	1	0,4
....	.....	.....	.....	.....
....	.....	.....	.....	.....
10		0,333333333	1	0,5
0	Siti Rugun Manurung			

Setelah data sampel dilakukan normalisasi data maka proses selajutnya melakukan pengelompokan data dengan menggunakan metode K-Means. Adapun proses penerapan metode K-Means dalam pengelompokan data mahasiswa berdasarkan sampel data yang telah normalisasikan dengan metode desimal scaling sebagai berikut:

1. Menyiapkan sampling data  
Data sampel yang telah dinormalisasikan pada tabel 3 diatas
2. Menentukan jumlah cluster  
Jumlah cluster / kelompok yaitu 3 sesuai dengan jumlah attribute kriteria
3. Menentukan nilai centroid/titik pusat  
Menentukan nilai centroid awal / titik pusat awal, dengan 3 (tiga) record data sesuai dengan jumlah attribute kriteria. Nilai centroid awal / titik pusat awal diperoleh dari data mahasiswa secara acak.

**Tabel 4.** Nilai Centroid Awal

Pusat	Nama	Pekerjaan Orang Tua	Status Kepemilikan Rumah	Penghasilan Orang Tua
Pusat C0	Muhammad Ali	0,333333333	1	0,4
Pusat C1	Siti Nurhalizah	1	1	1
Pusat C2	Senila Zebua	0,666666667	1	0,5

4. Menghitung jarak masing-masing centroid

a. Cluster C1:

$$d(x_1, c_1) = \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2}$$

$$= \sqrt{(0,333333333 - 0,333333333)^2 + (0,5 - 1)^2 + (0,5 - 0,4)^2} = 25000000$$

b. Cluster C2:

$$d(x_1, c_2) = \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2}$$

$$= \sqrt{(0,333333333 - 1)^2 + (0,5 - 1)^2 + (0,5 - 1)^2} = 25000000$$

c. Cluster C3:

$$d(x_1, c_3) = \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2}$$

$$= \sqrt{(0,333333333 - 0,666666667)^2 + (0,5 - 1)^2 + (0,5 - 0,5)^2} = 25000000$$

Untuk hasil pada masing cluster dapat dilihat pada tabel dibawah ini:

Tabel 5. Hasil Masing-Masing Centroid

Nama	Pekerjaan Orang Tua	Status Kepemilikan Rumah	Penghasilan Orang Tua	DC1	DC2	DC3
Atira Nabila	0,3333	0,5	0,5	0,2600	0,9444	0,3611
Sawitri	0,6667	0,5	0,36	0,3627	0,7707	0,2696
Meliana Sabet		1	0,4	0,0000	0,8044	0,1211
Tambunan	0,3333					
Desi Novria Siregar	0,3333	1	0,36	0,0016	0,8540	0,1307
Handayani		1	0,5	0,1211	0,3611	0,0000
Simanjuntak	0,6667					
Mardiana Pasaribu	0,3333	0,5	0,36	0,2516	1,1040	0,3807
Yamolala Hulu	1,0000	1	1	0,8044	0,0000	0,3611
David Ridwan		0,5	0,6	0,4011	0,5211	0,2600
Hakim	0,6667					
Misael Oktavianda		1	0,28	0,1255	0,6295	0,0484
Harefa	0,6667					
Muhammad Ali	0,3333	1	0,4	0,2600	0,9444	0,3611
.....	.....	.....	.....	.....	.....	.....
.....	.....	.....	.....	.....	.....	.....
Siti Rugun Manurung	0,3333	1	0,5	1,4444	0,1600	0,9211

5. Mengelompokan data berdasarkan jarak terpendek

Aturan yang digunakan untuk mengelompokan / cluster data berdasarkan hasil perhitungan jarak adalah

- a. Jika  $dc1 < dc2$  dan  $dc1 < dc3$  maka cluster = 1
- b. Jika  $dc2 < dc1$  dan  $dc2 < dc3$  maka cluster = 2
- c. Jika  $dc3 < dc1$  dan  $dc3 < dc2$  maka cluster = 3

Tabel 5. Hasil Cluster

Nama	Pekerjaan Orang Tua	Status Kepemilikan an Rumah	Penghasil an Orang Tua	DC1	DC2	DC3	Cluster
Atira Nabila	0,3333	0,5	0,5	0,2600	0,9444	0,3611	C2
Sawitri	0,6667	0,5	0,36	0,3627	0,7707	0,2696	C0
Meliana Sabet		1	0,4	0,0000	0,8044	0,1211	C0
Tambunan	0,3333						
Desi Novria Siregar		1	0,36	0,0016	0,8540	0,1307	C0
Handayani	0,3333	1	0,5	0,1211	0,3611	0,0000	C2
Simanjuntak	0,6667						
Mardiana Pasaribu		0,5	0,36	0,2516	1,1040	0,3807	C0
Yamolala Hulu	0,3333	1	1	0,8044	0,0000	0,3611	C1

Nama	Pekerjaan Orang Tua	Status Kepemilikan Rumah	Penghasilan Orang Tua	DC1	DC2	DC3	Cluster
David		0,5	0,6	0,4011	0,5211	0,2600	C0
Ridwan							
Hakim	0,6667						
Misael		1	0,28	0,1255	0,6295	0,0484	C0
Oktavianda							
Harefa	0,6667						
Muhammad		1	0,4	0,2600	0,9444	0,3611	C0
Ali	0,3333						
.....	.....	.....	.....	.....	.....	.....	
.....	.....	.....	.....	.....	.....	.....	
Siti Rugun	0,3333	1	0,5	1,4444	0,1600	0,9211	C2
Manurung							

Proses selanjutnya melakukan iterasi ke dua dengan pencarian nilai centroid / titik pusat awal pada iterasi ke-2 diperoleh dari nilai rata-rata penjumlahan anggota masing-masing cluster, setelah itu dibagi dengan jumlah anggota masing-masing cluster. Adapun hasil yang didapat pada tabel dibawah ini:

Tabel 6. Nilai Centroid Pada Iterasi Ke-2

Pusat Iterasi Ke2	Pekerjaan Orang Tua	Status Kepemilikan Rumah	Penghasilan Orang Tua
Pusat C0	0,513888889	0,833333333	0,385
Pusat C1	0,944444444	1	1,033333333
Pusat C2	0,55	0,875	0,565

Maka proses selanjutnya lakukan hal yang sama dari langkah 4 dan 5 sampai ditemukan tidak terjadi perubahan cluster antara cluster sebelumnya dengan cluster selanjutnya dalam penelitian ini proses pencarian itersai berhenti pada iterasi ke 5 dimana hasil cluster iterasi 5 memiliki hasil yang sama dengan hasil cluster iterasi 4 seperti tabel hasil dibawah

Tabel 7. Hasil Cluster

Nama	Pekerjaan Orang Tua	Status Kepemilikan Rumah	Penghasilan Orang Tua	DC1	DC2	DC3	Cluster
Atira Nabila	0,3333	0,5	0,5	0,0115	0,0115	0,0187	C1
Sawitri	0,6667	0,5	0,36	0,0230	0,0115	0,0135	C0
Meliana Sabet	0,3333	1	0,4	0,0115	0,0230	0,0150	C1
Tambunan							
Desi Novria	0,3333	1	0,36	0,0115	0,0230	0,0135	C1
Siregar							
Handayani	0,6667	1	0,5	0,0230	0,0230	0,0187	C0
Simanjuntak							
Mardiana	0,3333	0,5	0,36	0,0115	0,0115	0,0135	C1
Pasaribu							
Yamolala Hulu	1,0000	1	1	0,0345	0,0230	0,0374	C2
David Ridwan	0,6667	0,5	0,6	0,0230	0,0115	0,0224	C0
Hakim							
Misael	0,6667	1	0,28	0,0230	0,0230	0,0105	C0
Oktavianda							
Harefa							
Muhammad Ali	0,3333	1	0,4	0,0115	0,0230	0,0150	C1
.....	.....	.....	.....	.....	.....	.....	
.....	.....	.....	.....	.....	.....	.....	
.....	.....	.....	.....	.....	.....	.....	
Siti Rugun	0,3333	1	0,5	0,0115	0,0230	0,0187	C2
Manurung							

Berdasarkan data pengujian itersai 5 didapatkan kelompok cluster C0 sebanyak 22 orang, C1 sebanyak 23 orang dan C1 sebayak 5 orang. Sehingga dapat disimpulkan yang mendapatkan prioritas bantuan uang kuliah tunggal sebanyak 22 orang yang berada pada kelompok cluster C0.

**Tabel 6.** Hasil Cluster/Pengelompokan 1

No	Nama
1	Sawitri
2	Handayani Simanjuntak
3	David Ridwan Hakim
4	Misael Oktavianda Harefa
5	Misael Oktavianda
6	Rohan Kristini Purba
7	Jepri Saprianto Sitorus
8	Muhammad Karel
9	Khamsah Anugrah Parma Nasution
10	Dira Amalia
11	Nelly Irawani Manurung
12	Devi Lestari Lase
13	Indah Cahaya Gea
14	Dwina Pri Indini
15	Zaza Mutiara Arini
16	Lusiana Misimani Giawa
17	Rohan Kristini Purba
18	Jepri Saprianto Sitorus
19	Wanda Anisa
20	Sartika R. Siburian

### 3.2 Hasil Pengujian

Adapun hasil pengujian dalam menentukan pengelompokan data mahasiswa yang menjadi prioritas penerima bantuan Uang Kuliah Tunggal bagi mahasiswa terdampak Covid-19 dengan menggunakan aplikasi data Mining yaitu rapidminer dengan hasil Cluster/Pengelompokan dan bentuk Cluster Model berdasarkan hasil normalisasi desimal scaling dan pengelompokan menggunakan metode K-Means dapat dilihat pada gambar dibawah ini:

Row No.	id	cluster	Pekerjaan Orang Tua	Status Kepemilika...	Penghasilan...
1	1	cluster_1	0.011	0.011	0.019
2	2	cluster_0	0.023	0.011	0.013
3	3	cluster_1	0.011	0.023	0.015
4	4	cluster_1	0.011	0.023	0.013
5	5	cluster_0	0.023	0.023	0.019
6	6	cluster_1	0.011	0.011	0.013
7	7	cluster_2	0.034	0.023	0.037
8	8	cluster_0	0.023	0.011	0.022
9	9	cluster_0	0.023	0.023	0.010
10	10	cluster_1	0.011	0.023	0.015
11	11	cluster_0	0.023	0.011	0.019
12	12	cluster_2	0.034	0.023	0.052
13	13	cluster_1	0.011	0.023	0.019
14	14	cluster_1	0.011	0.023	0.015
15	15	cluster_0	0.023	0.023	0.030
16	16	cluster_0	0.034	0.023	0.015
17	17	cluster_0	0.034	0.023	0.010

**Gambar 2.** Hasil Cluster/Pengelompokan Dengan Aplikasi Rapidminer

**Cluster Model**

Description

Cluster 0: 22 items  
 Cluster 1: 23 items  
 Cluster 2: 5 items  
 Total number of items: 100

**Gambar 3.** Hasil Cluster Model Dengan Aplikasi Rapidminer

## 4. KESIMPULAN

Berdasarkan hasil dan pembahasan dalam penelitian ini tentang penerapan normalisasi data dalam mengelompokkan data mahasiswa dengan menggunakan metode K-Means untuk menentukan prioritas bantuan



Uang Kuliah Tunggal. Dimana metode yang digunakan dalam normalisasi data yaitu Desimal Scaling dan metode pengelompokan data yaitu K-Means dengan hasil pengujian metode secara manual yaitu terdapat 3 cluster dimana jumlah cluster 0 cluster 1 dan cluster 3 sama dengan pengujian aplikasi data mining yaitu rapidminer sehingga yang layak menjadi prioritas mendapatkan bantuan uang kuliah berdasarkan sampel yaitu cluster/pengelompokan 0 yang terdiri dari 22 orang.

## REFERENCES

- [1] T. H. B. Aviani dan A. T. Hidayat, "Sistem Pendukung Keputusan Seleksi Pemberian Uang Kuliah Tunggal Menerapkan Metode WASPAS," *J. Sist. Komput. dan Inform.*, vol. 2, no. 1, hal. 102–119, 2020.
- [2] D. A. Nasution, H. H. Khotimah, dan N. Chamidah, "PERBANDINGAN NORMALISASI DATA UNTUK KLASIFIKASI WINE MENGGUNAKAN ALGORITMA K-NN," *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 4, no. 1, hal. 78–82, 2019.
- [3] A. R. Aziz, B. Warsito, dan A. Prahutama, "Pengaruh Transformasi Data Pada Metode Learning Vector Quantization Terhadap Akurasi Klasifikasi Diagnosis Penyakit Jantung," *J. Gaussian*, vol. 10, no. 1, hal. 21–30, 2021, doi: 10.14710/j.gauss.v10i1.30933.
- [4] Z. Nabila, A. R. Isnain, Permata, dan Z. Abidin, "ANALISIS DATA MINING UNTUK CLUSTERING KASUS COVID-19 DI PROVINSI LAMPUNG DENGAN ALGORITMA K-MEANS," *J. Teknol. dan Sist. Inf.*, vol. 2, no. 2, hal. 100–108, 2021.
- [5] H. Tuti, N. Odi, dan W. Eko, "ANALISIS DAN PENERAPAN ALGORITMA K-MEANS DALAM STRATEGI PROMOSI KAMPUS AKADEMI MARITIM SUAKA BAHARI," *J. Sains Teknol. Transp. Marit.*, vol. 3, no. 1, hal. 1–7, 2021.
- [6] S. Defiyanti, M. Jajuli, dan N. Rohmawati, "Optimalisasi K-MEDOID dalam Pengklasteraan Mahasiswa Pelamar Beasiswa dengan CUBIC CLUSTERING CRITERION," *J. Nas. Teknol. dan Sist. Inf.*, vol. 3, no. 1, hal. 211–218, 2017, doi: 10.25077/teknosi.v3i1.2017.211-218.
- [7] mohamad jajuli nurul rohmawati, sofi defiyanti, "Implementasi Algoritma K-Means Dalam Pengklasteraan Mahasiswa Pelamar Beasiswa," *Jitter 2015*, vol. I, no. 2, hal. 62–68, 2015.
- [8] Rianda Dirkareshza, D. M. Azura, dan R. Pradana, "Government Policy During The Covid-19 Pandemi: Between Welfare State and Healthy State," *J. Mercat.*, vol. 14, no. 1, hal. 46–55, 2021.
- [9] "Data Mining \_ Algoritma dan Implementasi - Anjar Wanto, Muhammad Noor Hasan Siregar, Agus Perdana Windarto, Dedy Hartama, Ni Luh Wiwik."
- [10] M. Iskandar, C. C. Henry, dan A. Aulia, "Perancangan Database Sistem Informasi Akuntansi menggunakan Kombinasi REA Model, ERD, dan Normalisasi Data," *Bina Ekon. Maj. Ilmiah Fak. Ekon. Unpar*, vol. Volume 15, hal. 18, 2011.
- [11] N. Chamidah, Wiharto, dan U. Salamah, "Pengaruh Normalisasi Data pada Jaringan Syaraf Tiruan Backpropagasi Gradient Descent Adaptive Gain (BPGDAG) untuk Klasifikasi," *J. Teknol. Inf. ITS smart*, vol. 1, no. 1, hal. 28, 2016, doi: 10.20961/its.v1i1.582.