

## BAB IV

### HASIL DAN PEMBAHASAN

#### 4.1 Hasil

##### a. Proses Clustering

Pada tahap ini akan di lakukan proses utama yaitu segmentasi atau pengelompokan data penerima dana bantuan siswa miskin. Berikut ini merupakan penerapan algoritma K-Means dengan asumsi bahwa parameter input adalah jumlah dataset sebanyak n data dan jumlah inisialisasi centroid  $k = 3$  sesuai dengan penelitian. Data yang diambil untuk penelitian berjumlah 1044 data siswa untuk dijadikan penerapan ke dalam algoritma K-Means. Percobaan dilakukan dengan menggunakan parameter-parameter berikut :

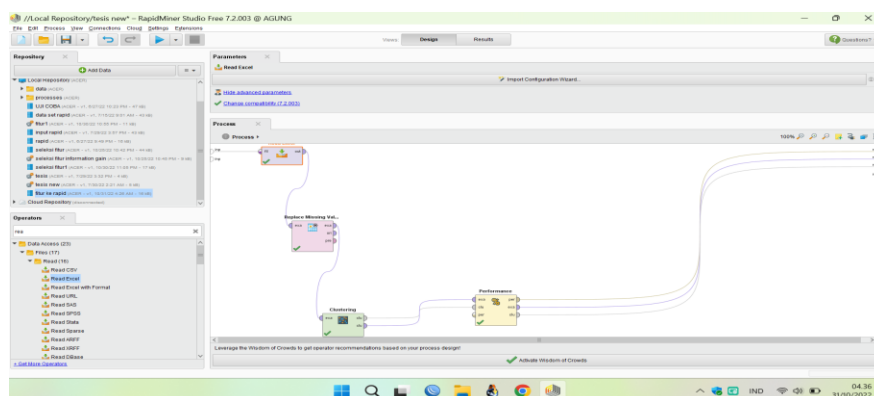
Jumlah cluster : 3

Jumlah data : 1044

Jumlah atribut :4

##### b. Pengujian rapidminer

Pada penelitian ini penulis menggunakan tool rapid miner sebagai alat pengujian data set. Adapun tahapan pengujian yang dilakukan yaitu sebagai berikut :



Gambar 4.1 Design Proses

Gambar 4.1 merupakan design proses tahapan yang dilakukan melalui 4 proses yaitu :

a. Read excel

Tahapan ini dilakukan operasi penginputan dataset berupa file berekstensi.xls

Data siswa calon penerima dana bantuan siswa miskin.

b. Replace missing value

Tahapan ini dilakukan operasi pengisian nilai yang hilang dengan nilai maksimal.

c. Clustering

Tahapan ini dilakukan operasi clustering sebagai algoritma yang digunakan padapenelitian ini.

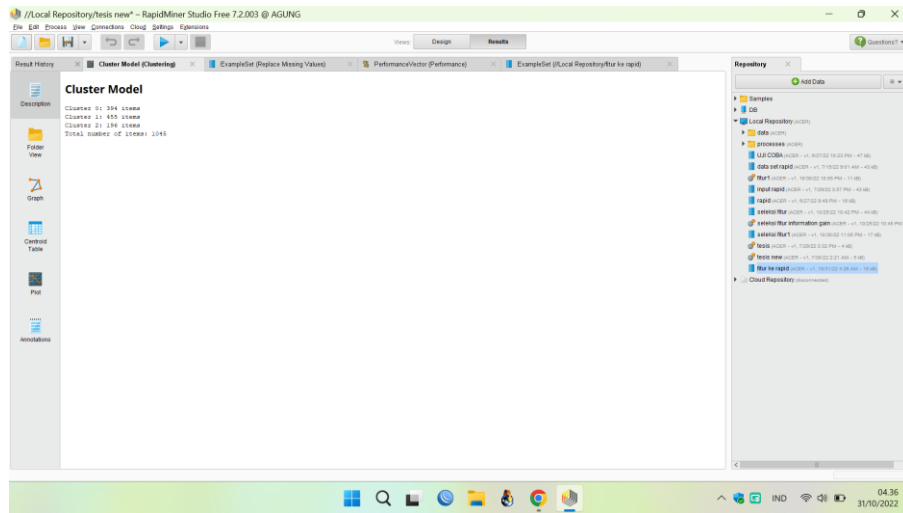
d. Performance

Tahapan ini dilakukan operasi pencarian nilai davies bouldin index.

Row No.	id	cluster	nilai siswa	tanggungan	jarak rumah	penghasilan
1	1	cluster_2	4	1	3	
2	2	cluster_0	4	4	4	
3	3	cluster_0	4	4	4	
4	4	cluster_0	4	5	4	
5	5	cluster_1	5	2	5	
6	6	cluster_0	4	1	3	3
7	7	cluster_0	4	1	5	5
8	8	cluster_1	4	2	2	4
9	9	cluster_1	4	1	1	4
10	10	cluster_0	4	1	4	5
11	11	cluster_2	5	1	3	
12	12	cluster_0	4	1	5	3
13	13	cluster_0	4	1	5	5
14	14	cluster_2	3	2	1	1
15	15	cluster_2	5	2	1	3
16	16	cluster_2	4	1	1	3
17	17	cluster_1	4	2	1	4
18	18	cluster_0	5	1	5	3
19	19	cluster_1	4	1	2	5
20	20	cluster_0	4	2	5	5
21	21	cluster_2	5	1	1	3
22	22	cluster_0	4	1	3	5
23	23	cluster_1	4	1	1	5
24	24	cluster_1	5	1	1	4
25	25	cluster_0	4	1	4	3
26	26	cluster_0	4	1	5	5

Gambar 4.2 Example set result

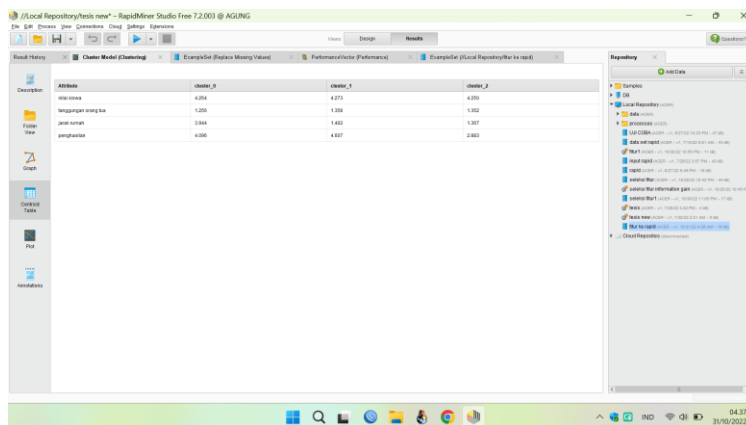
Gambar 4.2 merupakan example set result dimana Pada tahapan ini ditampilkan hasil dari klusterisasi data . Label cluster terbagi menjadi tiga kelompok yaitu cluster 0, cluster 1, cluster 2. Pembagian ini berdasarkan hasil kedekatan tiap masing-masing data dengan jarak terdekat (k).



Gambar 4.3 Cluster Model

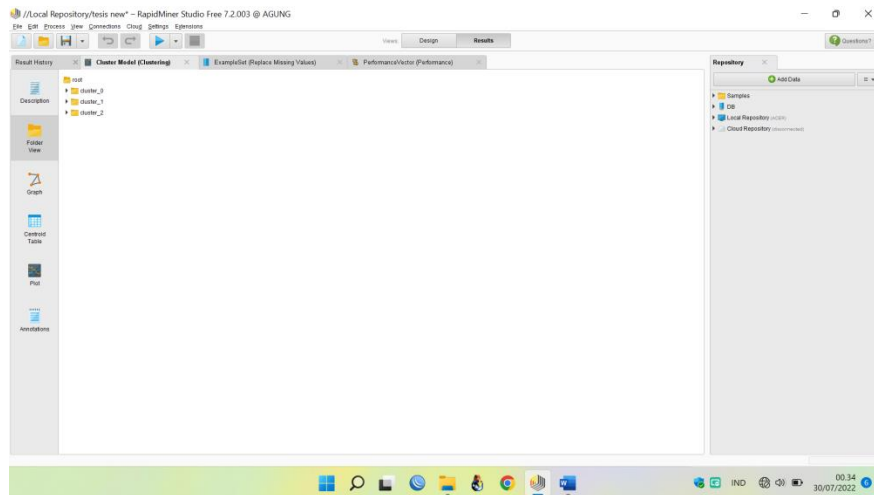
Gambar 4.3 merupakan cluster model dimana Pada tahapan ini ditampilkan hasil pembagian data terhadap tiap cluster.

Cluster 0 memiliki 394 anggota, Cluster 1 memiliki 455 anggota, Cluster 2 memiliki 196 anggota dari total 1044 dataset yang di uji.



Gambar 4.4 Centroid table

Gambar 4.4 merupakan centroid table dimana Pada tahapan ini ditampilkan nilai titik pusat pada tiap cluster. Nilai tersebut menjadikan acuan perhitungan pada tiap dataset dengan cara mengukur kedekatan nilai dengan masing-masing titik pusat cluster.



Gambar 4.5 Folder view

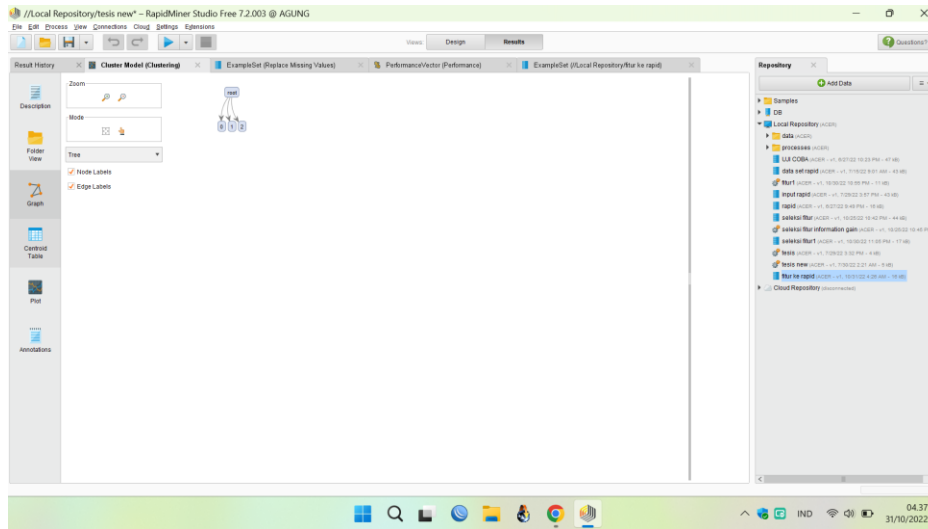
Gambar 4.5 merupakan gambar folder view dimana Pada tahapan ini ditampilkan dari masing -masing anggota cluster.

Name	Type	Missing	Statistics	Filter (0 of 16 attributes)
id	Integer	0	1 1045 523	
cluster	Normal	0	cluster_2 (196) cluster_1 (455) cluster_1 (455), cluster_0 (204) ... [1 more]	
nilai siswa	Integer	0	3 5 4.265	
langganan orang tua	Integer	0	1 0 1.319	
jarak rumah	Integer	0	1 0 2.354	
penghasilan	Integer	0	1 0 4.091	

Gambar 4.6 Example set statistic

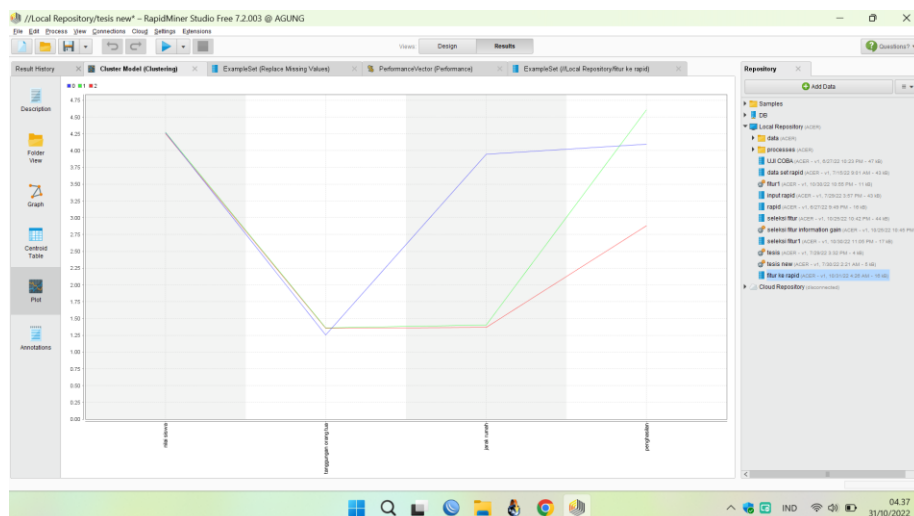
Gambar 4.6 merupakan example set statistic dimana Pada tahapan ini ditampilkan hasil statistik dari data yang sudah di uji, Pada tabel penghasilan orang tua,

tanggung, jarak, nilai terdapat 3 atribut yaitu min sebagai nilai terendah pada tabel dataset, max sebagai nilai tertinggi pada tabel dataset dan average sebagai nilai rata-rata dari penjumlahan dataset tersebut.



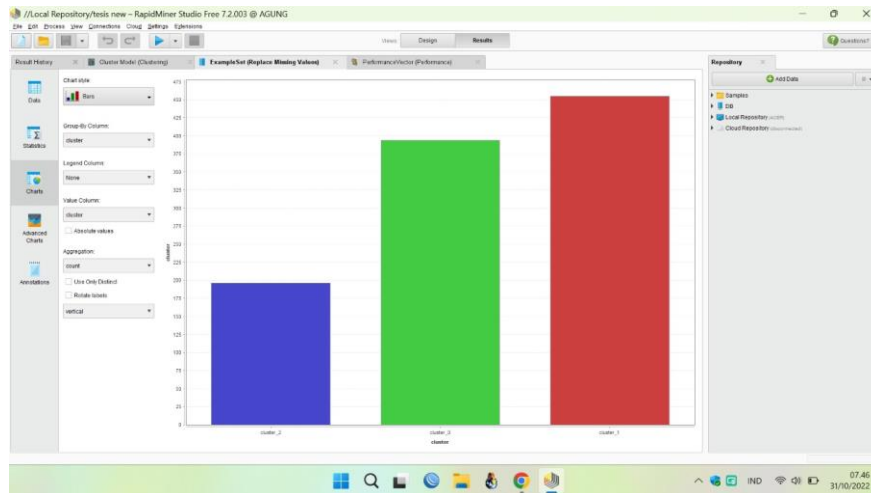
Gambar 4.7 Graph result

Gambar 4.7 merupakan graph result dimana Pada tahapan ini ditampilkan hasil pembagian dari tiga kelompok, dimana setiap kelompok berisikan anggota dari masing-masing cluster.



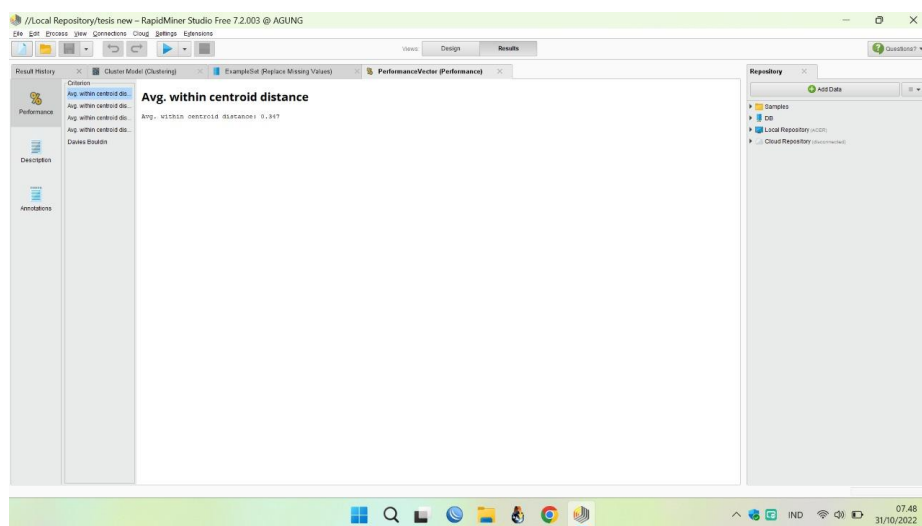
Gambar 4.8 Plot result

Gambar 4.8 merupakan ploth result dimana Pada tahapan ini ditampilkan hasil plot dari hasil pengujian. Pada bagian sebelah kiri menunjukkan angka pendapatan dan bagian sebelah kanan ditampilkan sebagai gambaran rata-ran nilai anggota.



Gambar 4.9 Chart clustering

Gambar 4.9 merupakan chart clustering dimana Pada tahapan ini ditampilkan hasil pengelompokan data dalam bentuk diagram dengan warna. Warna merah mengartikan cluster 0, warna biru mengartikan cluster 1, dan warna hijau mengartikan cluster 2.



Gambar 4.10 Performance Vector

Gambar 4.10 merupakan performance vector dimana Pada tahapan ini adalah menghitung rata-rata jarak dari tiap-tiap cluster .

## 4.2 Pembahasan

### a. Hasil Clustering K-Means

Setelah dilakukan pengujian dengan tool rapid miner , maka dapat di simpulkan sebagai berikut :

- Cluster 0 memiliki 394 anggota,
- Cluster 1 memiliki 455 anggota,
- Cluster 2 memiliki 196 anggota
- dari total 1044 dataset yang di uji

Setelah cluster terbentuk , maka di peroleh kesimpulan sebagai berikut :

Tabel 4.1 Deskripsi data berdasarkan dengan cluster

Cluster	Deskripsi
0	Anggota cluster layak menerima bantuan
1	Anggota cluster dapat dipertimbangkan menerima bantuan
2	Anggota cluster tidak layak menerima bantuan

Table 4.1 merupakan deskripsi data berdasarkan dengan cluster dimana Setelah proses clustering selesai selanjutnya dilakukan operasi performance untuk mengetahui nilai dari Davies bouldin indeks yang bertujuan untuk memaksimalkan pengukuran jarak antar cluster dan meminimalkan jarak antar anggota dalam suatu cluster.

Tabel 4.2 Hasil operasi performance vector

PerformanceVector	Value
Avg. within centroid distance	0.347

Avg. within centroid distance_cluster_0	0.481
Avg. within centroid distance_cluster_1	0.272
Avg. within centroid distance_cluster_2	0.249
Davies Bouldin	0.259

Table 4.2 merupakan Hasil operasi performance vector dimana Evaluasi hasil dari average within centroid distance mendekati angka 0 mengartikan bahwa masing-masing anggota di dalam cluster berada dalam jarak yang berdekatan. Evaluasi menggunakan davies bouldin indeks memiliki skema internal cluster yang dilihat dari kuantitas dan kedekatan antar hasil cluster.

Semakin kecil nilai davies bouldin indeks yang diperoleh (non-negatif) $\geq 0$ ), maka semakin baik cluster yang diperoleh dari pengelompokan menggunakan metode clustering. Hasil perhitungan menggunakan algoritma K-Means menunjukkan nilai 0,259. Angka tersebut memiliki arti masing-masing objek dalam cluster tersebut memiliki kesamaan yang cukup baik karena mendekati angka 0.