

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)

Koko Handoko
Universitas Putera Batam
(cooresponding author) kokohandoko01@gmail.com

Abstract— This research applies data mining using clustering methods to improve the quality of learning in Higher Education Institutions in the Program TKJ Community College South Solok. The algorithm used is K-Means Clustering is a process of grouping a number of data or object into a cluster (group) so that each cluster will contain the data that is as similar as possible and different from the objects in other clusters. Testing is done with RapidMiner 5.3 applications that generate clusters in improving the quality of learning. The samples used were taken from the data tables of students who have ditrasformasi. Where the variables are defined as the first test four variables, namely: IP students, distance students, attendance and parental income. Where the students will present data with the quality of teaching is very good, good, good enough, and less good.

Keyword— Data to Improve Learning, Data Mining, K-Means Clustreing, RapidMiner.

Intisari— Penelitian ini menerapkan Data Mining dengan menggunakan metode clustering untuk Meningkatkan mutu pembelajaran pada Instansi Perguruan Tinggi di Program Studi TKJ Akademi Komunitas Solok Selatan. Algoritma yang digunakan yaitu K-Means Clustering berupa proses pengelompokan sejumlah data atau objek ke dalam cluster (group) sehingga setiap dalam cluster tersebut akan berisi data yang semirip mungkin dan berbeda dengan objek dalam cluster yang lainnya. Pengujian dilakukan dengan aplikasi RapidMiner 5.3 sehingga menghasilkan cluster-cluster dalam meningkatkan mutu pembelajaran. Sampel yang digunakan diambil dari tabel data mahasiswa yang telah ditrasformasi. Di mana variabel yang pengujian pertama ditentukan sebanyak 4 variabel, yaitu : IP mahasiswa, jarak tempuh mahasiswa, jumlah kehadiran dan penghasilan orang tua. Di mana akan mempresentasikan data mahasiswa dengan mutu pembelajaran sangat baik, baik, cukup baik, dan kurang baik.

Kata Kunci— Data Mutu Pembelajaran, Data Mining, K-Means Clustreing, RapidMiner.

I. PENDAHULUAN

Pendidikan merupakan hal penting, dan akan terus berkembang seiring berjalannya zaman, di Indonesia agak sulit mengimbangi perkembangan itu, karena di Indonesia mutu pendidikannya pun masih tergolong rendah. Peningkatan mutu pembelajaran di perguruan tinggi merupakan urgensi yang mendesak untuk segera dilakukan perbaikan. Peningkatan mutu itu pada dasarnya dapat dilakukan dengan strategi merubah salah satu dari subsistem : manusia, struktur, teknologi, dan proses organisasi.

Bagaimana ekstraksi *Data Mining* yang dihasilkan dapat memberikan sebuah pengetahuan baru terhadap instansi Akademi Komunitas Solok Selatan dalam meningkatkan mutu pembelajaran. *Data Mining* dilakukan untuk menggali informasi dari sejumlah data. Keluaran dari *Data Mining* ini bisa dipakai untuk memperbaiki pengambilan keputusan di masa depan [1]. Salah satu teknik yang dikenal dalam *Data Mining* yaitu *clustering* [3]. Analisis *cluster* berguna dalam meringkas data atau sejumlah variabel untuk menjadi lebih sedikit. Dalam melakukan proses meringkas data ini dapat dilakukan dengan mengelompokkan objek-objek berdasarkan kesamaan karakteristik tertentu diantara objek-objek yang hendak diteliti. Sehingga mengetahui karakteristik pengelompokan data meningkatkan mutu pembelajaran dan mengimplementasikan Algoritma *K-Means Clustering* dalam menentukan data yang diambil dari Program Studi TKJ Akademi Komunitas Solok Selatan.

II. KAJIAN LITERATUR

A. Knowledge Discovery in Database

Saat ini, konsep Data Mining semakin dikenal sebagai *tools* penting dalam manajemen informasi karena jumlah informasi yang semakin besar jumlahnya. Data Mining sendiri sering disebut sebagai *Knowledge Discovery in Database* (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola hubungan dalam set data berukuran besar [3].

B. Data Mining

Data Mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari *database* yang besar [2].

Data Mining merupakan suatu rangkaian proses untuk menggali nilai tambah dari sekumpulan data yang berupa pengetahuan yang selama ini tidak diketahui secara manual. *Data Mining* sebenarnya memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (*artificial intelligent*), *machine learning*, statistik dan basis data. Beberapa teknik yang sering disebut-sebut dalam *literatur Data Mining* antara lain : *clustering*, *classification*, *association rule mining*, *neural network*, dan *genetic algorithm* [4].

C. Clustering

Salah satu teknik yang dikenal dalam *Data Mining* yaitu *clustering*. Pengertian *clustering* keilmuan dalam *Data Mining* adalah pengelompokan sejumlah data atau objek ke dalam *cluster* (*group*) sehingga setiap dalam *cluster* tersebut akan berisi data yang semirip mungkin dan berbeda dengan objek dalam *cluster* yang lainnya [3].

Metode *clustering* yang paling banyak digunakan ialah metode *K-Means clustering*. Kelemahan utama dari metode ini adalah hasil yang sensitif terhadap pemilihan pusat *cluster* awal dan perhitungan solusi lokal untuk mencapai kondisi optimal. Analisis *Cluster* merupakan teknik multivariat yang mempunyai tujuan utama untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya. Analisis *Cluster* mengklasifikasi objek sehingga setiap objek yang paling dekat kesamaannya dengan objek lain berada dalam *cluster* yang sama [6].

D. Algoritma K-Means

K-Means mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang relatif cepat dan efisien. Namun, *K-Means* mempunyai kelemahan yang diakibatkan oleh penentuan pusat awal *cluster*. Hasil *cluster* yang terbentuk dari metode *K-Means* ini sangatlah tergantung pada inisiasi nilai pusat awal *cluster* yang diberikan [3].

K-Means merupakan teknik pengelompokan yang bekerja berdasarkan *Partitioned Clustering*. Prinsip kerja dari pengelompokan *Hierarchical Clustering* dilakukan secara bertahap [1].

Beberapa teknik klastering yang paling sederhana dan umum adalah klastering *K-means*. Secara detail teknik ini menggunakan ukuran ketidakmiripan untuk mengelompokkan obyek. Ketidakmiripan dapat diterjemahkan dalam konsep jarak. Dua obyek dikatakan mirip jika jarak dua objek tersebut dekat. Semakin tinggi nilai jarak, semakin tinggi nilai ketidakmiripannya [5].

K-Means Clustering merupakan metode yang termasuk ke dalam golongan algoritma *Partitioning Clustering*.

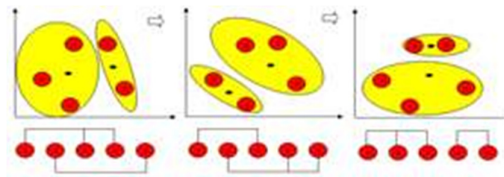
Langkah-langkah dari metode *K-Means* adalah sebagai berikut : [1]

1. Tentukan nilai k sebagai jumlah *cluster* yang ingin dibentuk.
2. Bangkitkan k *centroid* (titik pusat *cluster*) awal secara acak.
3. Hitung jarak setiap data ke masing-masing *centroid* menggunakan rumus korelasi antar dua objek (*Euclidean Distance*).
4. Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan *centroid*nya.
5. Tentukan posisi *centroid* baru (k C) dengan cara menghitung nilai rata-rata dari data yang ada pada *centroid* yang sama.

$$C_k = \left(\frac{1}{n_k} \right) \sum_i$$

Dimana n_k adalah jumlah dokumen dalam *cluster* k dan d_j adalah dokumen dalam *cluster* k .

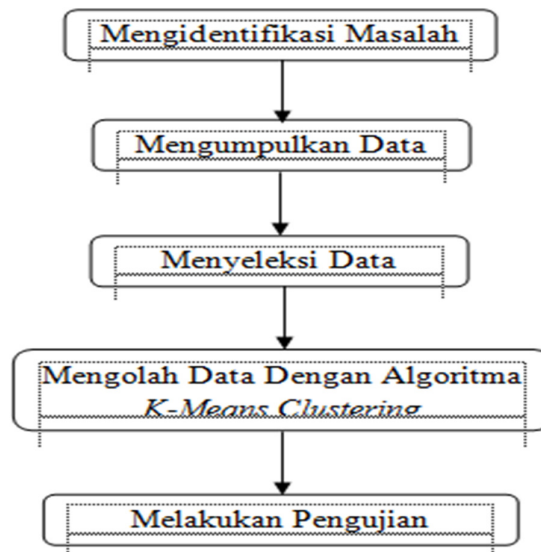
6. Kembali ke langkah 3 jika posisi *centroid* baru dengan *centroid* lama, tidak sama.



Gambar 1 Langkah-langkah metode K-Means

III. METODE PENELITIAN

Kerangka kerja yang akan digunakan dalam penelitian ini dapat dilihat pada Gambar 2.



Gambar 2 Kerangka Kerja Penelitian

IV. PEMBAHASAN

Bagian ini membahas mengenai analisis data dan metode yang digunakan untuk menyelesaikan permasalahan penelitian. Data yang digunakan adalah data mutu pembelajaran Akademi Studi Kasus Di Program Studi TKJ Akademi Komunitas Solok Selatan.

Proses *clustering* dengan menggunakan Algoritma *K-Means*, sebagai berikut :

1. Menentukan Jumlah *Cluster*

Jumlah cluster yang ditentukan untuk mengelompokan data pada penelitian ini sebanyak 4 *cluster*.

2. Menentukan Centroid

Pusat awal *cluster* (*centroid*) ditentukan secara *random* atau acak.

3. Menghitung Jarak dari *Centroid*

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

4. Alokasikan masing-masing objek ke *centroid* terdekat.

Dan metode yang digunakan untuk mengekstraksi data tersebut adalah metode *Clustering* dengan menggunakan Algoritma *K-Means*. Adapun sampel data mutu pembelajaran yaitu:

Tabel 1 Sampel Data Mutu pembelajaran

No	Nama Mahasiswa	IP Mahasiswa	Jarak Tempuh Mahasiswa (KM)	Jumlah Kehadiran	Penghasilan Orang Tua
k1	Muhammad Hafiz	2,98	2	90	3
k2	Gusreni Fitri	3,53	8	86	2
k3	Riri Indriani	3,03	7	79	1
k4	Nofriza Dewita	3,18	6	84	2
k5	Rika Nofrianti	2,73	7	77	1
k6	Parisal Syaputra	2,9	8	85	2
k7	Destuty Fransiska	2,9	2	83	2
k8	Zalti Syafliani	2,95	9	76	2
k9	Yosi Oktaviani	3,38	1	82	2
k10	Lusi Fikadilla	3,48	12	92	1
k11	Kevin Valerto	3,63	15	83	1
k12	Nurkarmila	3,33	15	82	3
k13	Irwan Yusrianto	3,08	15	85	2
k14	Prima Kurniawan	3,16	16	89	3
k15	Satria Dian Putra	2,95	13	88	2
k16	Ilham Desri Yetto	2,9	4	78	1
k17	Arif Budiman	3,38	4	76	1
k18	Khairur Afrar	2,83	4	73	3
k19	Mista Nurjana	3,73	8	88	2
k20	Nofrianti	3,18	2	88	1
k21	Agiliya Rahmaya	3,48	2	88	1
k22	Dhio Pratama	3	4	85	2
k23	Fitri Rahmadani	3,83	7	91	3
k24	Wiwit Syafma Yuni	3,48	6	92	2
k25	Noperawati	3,23	7	91	2

Untuk mengalokasikan objek ke dalam masing-masing *cluster* dengan cara mengelompokkan berdasarkan jarak minimum objek ke pusat *cluster*.

Nilai *centroid* awal pada penelitian ini dilakukan pemilihan secara acak, di mana jumlah *centroid* awal dilakukan sebanyak empat *centroid* awal, yaitu:

Tabel 2 Menentukan *Centroid* (titik pusat)

C1	Ilham Desri Yetto	2,9	4	78	1
C2	Yosi Oktaviani	3,38	1	82	2
C3	Kevin Valerto	3,63	15	83	1
C4	Fitri Rahmadani	3,83	7	91	3

Dalam mengalokasikan kembali objek ke dalam masing-masing *cluster* didasarkan pada perbandingan jarak antara data dengan *centroid* setiap *cluster* yang ada, objek dialokasikan secara tegas ke dalam *cluster* yang mempunyai jarak ke *centroid* terdekat dengan data tersebut. Berikut ini adalah merupakan hal perbandingan jarak antara data dengan

centroid setiap *cluster* yang ada. Perhitungan dilakukan terus sampai data ke-25 terhadap pusat *cluster*. Setelah dilakukan proses perhitungan maka akan didapatkan data selengkapnya adalah:

Tabel 3 Posisi *Cluster* Pada Posisi Pertama

No	C1	C2	C3	C4	C1	C2	C3	C4
1	12,329	8,134	14,914	5,169	0	0	0	1
2	9,022	8,064	7,682	5,205	0	0	0	1
3	3,165	6,791	8,964	12,192	1	0	0	0
4	6,409	5,389	9,122	7,171	0	1	0	0
5	3,167	7,901	10,040	14,185	1	0	0	0
6	8,124	7,631	7,385	6,234	0	0	0	1
7	1,493	5,477	13,059	9,532	1	0	0	0
8	5,477	10,009	9,299	15,191	1	0	0	0
9	5,122	0,000	14,073	10,872	0	1	0	0
10	16,135	14,900	9,488	5,488	0	0	0	1
11	12,105	14,073	0,000	11,491	0	0	1	0
12	11,882	14,036	2,256	12,052	0	0	1	0
13	13,078	14,321	2,303	10,078	0	0	1	0
14	16,403	16,585	6,420	9,244	0	0	1	0
15	13,491	13,423	5,519	6,839	0	0	1	0
16	0,000	5,122	12,105	13,523	1	0	0	0
17	2,057	6,782	13,041	15,434	1	0	0	0
18	5,386	9,555	15,021	18,276	1	0	0	0
19	10,848	9,226	8,661	3,318	0	0	0	1
20	10,202	6,168	13,936	6,199	0	1	0	0
21	10,215	6,165	13,929	6,174	0	1	0	0
22	7,072	4,260	11,243	6,833	0	1	0	0
23	13,523	10,872	11,491	0,000	0	0	0	1
24	14,189	11,181	12,768	1,767	0	0	0	1
25	13,383	10,818	11,365	1,166	0	0	0	1
					7	5	5	8

Jumlah anggota cluster :

C1 berjumlah 7 = 3,5,7,8,16,17,18.

C2 berjumlah 5 = 4,9,20,21,22.

C3 berjumlah 5 = 11,12,13,14,15.

C4 berjumlah 8 = 1,2,6,10,19,23,24,25.

Kemudian kita tentukan lagi pusat *cluster* dari data yang baru, caranya dengan menjumlahkan nilai semua fakultas yang merupakan anggota dari *cluster* dan dibagi total jumlah anggota *cluster* :

Tabel 4 Mengulang Langkah Kedua Sehingga Posisi Data Tidak Mengalami Perubahan

Cluster	IP mahasiswa	Jarak tempuh KM	Jumlah Kehadiran	Penghasilan Ortu
C1	2,96	5,28571	77,4286	1,57
C2	3,244	3	85,4	1,6
C3	3,23	14,8	85,4	2,2
C4	3,395	7,25	89,375	2,13

Setelah melakukan perhitungan kembali, dapatlah iterasi ke dua pada tabel 5:

Tabel 5 Posisi Cluster Pada Posisi Kedua

No	C1	C2	C3	C4	C1	C2	C3	C4
1	13,072	4,9183	13,6273	5,37503	0	1	0	0
2	9,01916	5,05982	6,83593	3,46222	0	0	0	1
3	2,39574	7,57402	10,1627	10,4452	1	0	0	0
4	6,62767	3,33528	8,91305	5,52404	0	1	0	0
5	1,87133	9,33725	11,5365	12,4463	1	0	0	0
6	8,05488	5,04364	6,82268	4,46808	0	0	0	1
7	6,48259	2,65299	13,0288	8,27428	0	1	0	0
8	4,00256	11,1627	11,0507	13,4969	1	0	0	0
9	6,29487	3,96718	14,2149	9,66794	0	1	0	0
10	16,0625	11,1793	7,27341	5,5431	0	0	0	1
11	11,2331	12,2584	2,72029	10,1007	0		1	0
12	10,8371	12,551	3,5	10,7342	0	0	1	0
13	12,3244	12,0144	0,51235	8,90606	0	0	1	0
14	15,8359	13,562	3,87878	8,80477	0	0	1	0
15	13,0939	10,3444	3,18094	5,93016	0	0	1	0
16	1,51978	7,49922	13,151	11,8939	1	0	0	0
17	2,04861	9,47304	14,3688	13,8101	1	0	0	0
18	4,82939	12,5256	16,4682	16,7269	1	0	0	0
19	10,9498	5,67064	7,3	1,60654	0	0	0	1
20	11,0872	2,85028	13,1165	5,54662	0	1	0	0
21	11,0972	2,85932	13,1188	5,5431	0	1	0	0
22	7,69187	1,17454	10,8117	5,46578	0	1	0	0
23	13,7812	7,04723	9,65401	1,91258	0	0	0	1
24	14,6045	7,26469	11,0047	2,91135	0	0	0	1
25	13,6886	6,89349	9,60417	1,6571	0	0	0	1
					6	7	5	7

Jumlah anggota cluster :

C1 berjumlah 6 = 3,5,8,16,17,18.

C2 berjumlah 7 = 1,4,7,9,20,21,22.

C3 berjumlah 5 = 11,12,13,14,15.

C4 berjumlah 7 = 2,6,10,19,23,24,25.

Karena *cluster* ke 2 ini belum sama dengan *cluster* 1, kemudian kita tentukan lagi *centroid* (titik pusat) yang baru, dengan cara menjumlahkan nilai semua data yang merupakan anggota dari *cluster* dan dibagi total jumlah anggota *cluster* :

Tabel 6 Mengulang Langkah Kedua Sehingga Posisi Data Tidak Mengalami Perubahan

Cluster	IP mahasiswa	Jarak tempuh KM	Jumlah Kehadiran	Penghasilan Ortu
C1	2,97	5,83333	76,5	1,5
C2	3,15714	2,71429	85,7143	1,86
C3	3,23	14,8	85,4	2,2
C4	3,45429	8	89,2857	2

Berdasarkan proses iterasi 2 dan 3, terlihat pada titik pusat cluster yang diproses tetap sama, dan tidak ada lagi data yang berpindah dari satu cluster yang lain, sehingga proses iterasi dihentikan pada iterasi ke-3. Jadi data mutu pembelajaran ini untuk C1 yaitu IP mahasiswa memiliki 6 kelompok, C2 yaitu Jarak Tempuh Mahasiswa (KM) memiliki 7 kelompok, C3 yaitu jumlah kehadiran memiliki 5 kelompok, dan C4 yaitu penghasilan orang tua memiliki 7 kelompok. Hasil iterasi ke dua pada tabel 7.

Tabel 7 Posisi *Cluster* Pada Posisi Ketiga

No	C1	C2	C3	C4	C1	C2	C3	C4
1	14,1136	4,49612	13,6273	6,14289	0	1	0	0
2	9,77282	5,30847	6,83593	3,28659	0	0	0	1
3	2,80441	8,01248	10,1627	10,3911	1	0	0	0
4	7,52143	3,70886	8,91305	5,65809	0	1	0	0
5	1,38518	9,75824	11,5365	12,388	1	0	0	0
6	8,78632	5,34186	6,82268	4,32141	0	0	0	1
7	7,56302	2,82207	13,0288	8,70732	0	1	0	0
8	3,24472	11,5733	11,0507	13,3328	1	0	0	0
9	7,35046	4,09936	14,2149	10,1038	0	1	0	0
10	16,6969	11,2505	7,27341	4,9364	0	0	0	1
11	11,2678	12,62	2,72029	9,46261	0	0	1	0
12	10,8008	12,8868	3,5	10,1537	0	0	1	0
13	12,5116	12,3075	0,51235	8,21629	0	0	1	0
14	16,1832	13,7336	3,87878	8,07268	0	0	1	0
15	13,5596	10,5396	3,18094	5,18723	0	0	1	0
16	2,42198	7,87173	13,151	12,0281	1	0	0	0
17	2,00729	9,83894	14,3688	13,911	1	0	0	0
18	4,22856	12,8343	16,4682	16,8111	1	0	0	0
19	11,7376	5,78894	7,3	1,31494	0	0	0	1
20	12,1342	2,5436	13,1165	6,22321	0	1	0	0
21	12,1431	2,56391	13,1188	6,21721	0	1	0	0
22	8,70988	1,48606	10,8117	5,87994	0	1	0	0
23	14,6493	6,93289	9,65401	2,25387	0	0	0	1
24	15,5173	7,10146	11,0047	3,37165	0	0	0	1
25	14,5578	6,80675	9,60417	1,99727	0	0	0	1
					6	7	5	7

Jumlah anggota cluster :

C1 berjumlah 6 = 3,5,8,16,17,18.

C2 berjumlah 7 = 1,4,7,9,20,21,22.

C3 berjumlah 5 = 11,12,13,14,15.

C4 berjumlah 7 = 2,6,10,19,23,24,25.

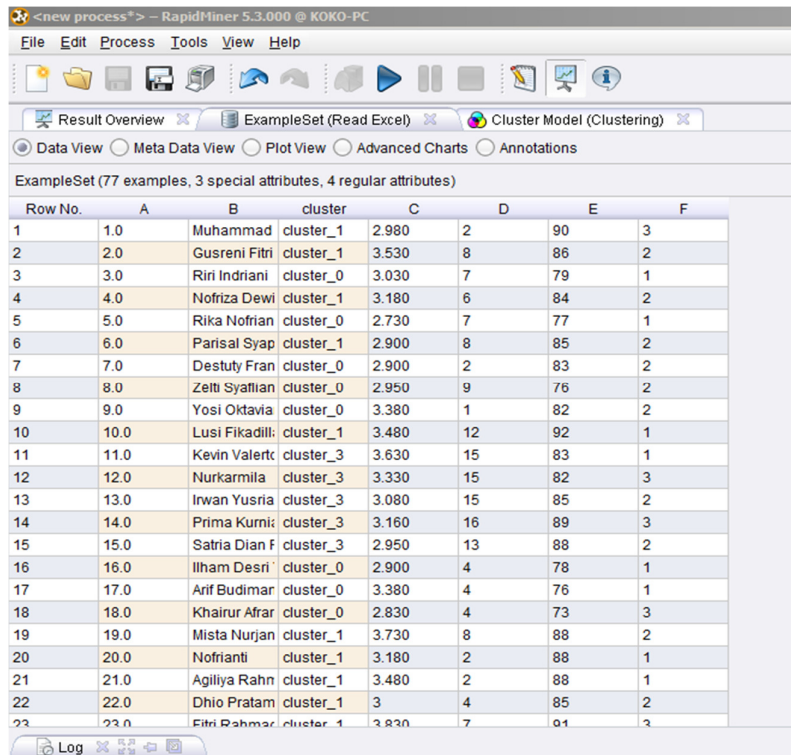
Berdasarkan proses iterasi 2 dan 3, terlihat pada titik pusat *cluster* yang diproses tetap sama, dan tidak ada lagi data yang berpindah dari satu *cluster* yang lain, sehingga proses iterasi dihentikan pada iterasi ke-3. Jadi data mutu pembelajaran ini untuk C1 yaitu IP mahasiswa memiliki 6 kelompok, C2 yaitu Jarak Tempuh Mahasiswa (KM) memiliki 7 kelompok, C3 yaitu jumlah kehadiran memiliki 5 kelompok, dan C4 yaitu penghasilan orang tua memiliki 7 kelompok.

V. HASIL CLUSTERING DENGAN RAPIDMINER

Pada tahap pengujian algoritma ini untuk membuktikan kebenaran pada tahap analisa sebelumnya dan pengujian secara manual, maka perlu dilakukan pengujian lagi untuk pengelompokan data meningkatkan mutu pembelajaran menggunakan algoritma *K-Means*. Selanjutnya setelah *database* berhasil diproses, maka dapat dilihat hasil pengelompokan data Mahasiswa Program Studi TKJ Akademi Komunitas Solok Selatan. Untuk melihat bukti bahwa *database* berhasil diproses :

1. DATA VIEW

Dilihat dari *example set (read excel)* dapat dilihat pada gambar 3:



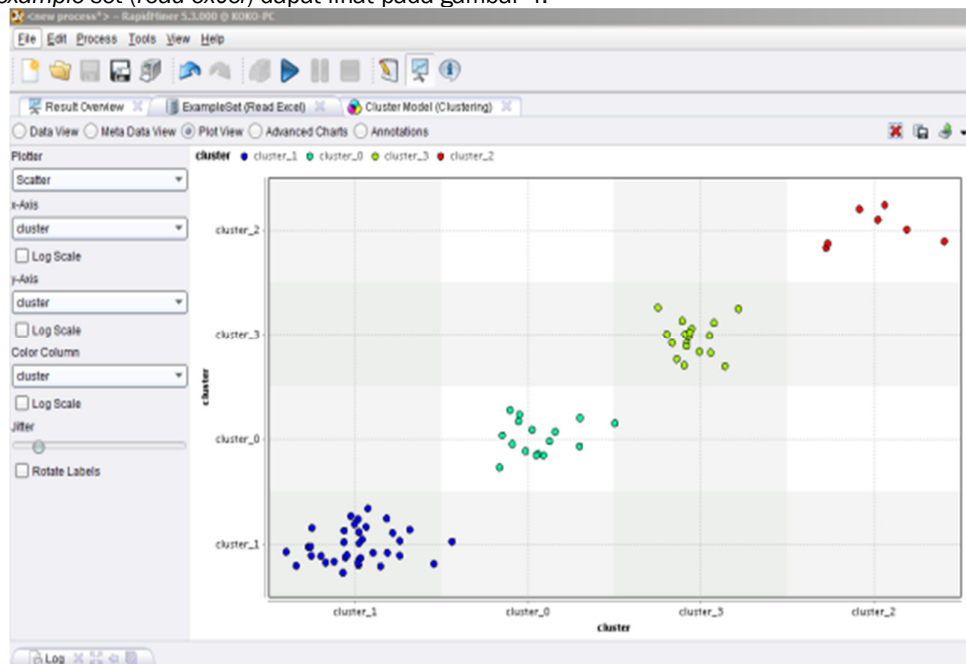
Row No.	A	B	cluster	C	D	E	F
1	1.0	Muhammad	cluster_1	2.980	2	90	3
2	2.0	Gusreni Fitri	cluster_1	3.530	8	86	2
3	3.0	Riri Indriani	cluster_0	3.030	7	79	1
4	4.0	Nofriza Dewi	cluster_1	3.180	6	84	2
5	5.0	Rika Nofrian	cluster_0	2.730	7	77	1
6	6.0	Parisal Syap	cluster_1	2.900	8	85	2
7	7.0	Destuty Fran	cluster_0	2.900	2	83	2
8	8.0	Zelti Syafian	cluster_0	2.950	9	76	2
9	9.0	Yosi Oktavia	cluster_0	3.380	1	82	2
10	10.0	Lusi Fikadilli	cluster_1	3.480	12	92	1
11	11.0	Kevin Valert	cluster_3	3.630	15	83	1
12	12.0	Nurkarmila	cluster_3	3.330	15	82	3
13	13.0	Irwan Yusria	cluster_3	3.080	15	85	2
14	14.0	Prima Kurni	cluster_3	3.160	16	89	3
15	15.0	Satria Dian F	cluster_3	2.950	13	88	2
16	16.0	Ilham Desri	cluster_0	2.900	4	78	1
17	17.0	Arif Budimar	cluster_0	3.380	4	76	1
18	18.0	Khairur Afrar	cluster_0	2.830	4	73	3
19	19.0	Mista Nurjan	cluster_1	3.730	8	88	2
20	20.0	Nofrianti	cluster_1	3.180	2	88	1
21	21.0	Agiliya Rahn	cluster_1	3.480	2	88	1
22	22.0	Dhio Pratam	cluster_1	3	4	85	2
23	23.0	Fitri Rahma	cluster_1	3.830	7	81	2

Gambar 3 Tampilan *ExampleSet (Read Excel) Data View*

Gambar 3 merupakan *sheet* untuk menampilkan *database* yang telah diolah secara keseluruhan lengkap dengan *clusternya* berjumlah 77 data.

2. FLOT VIEW

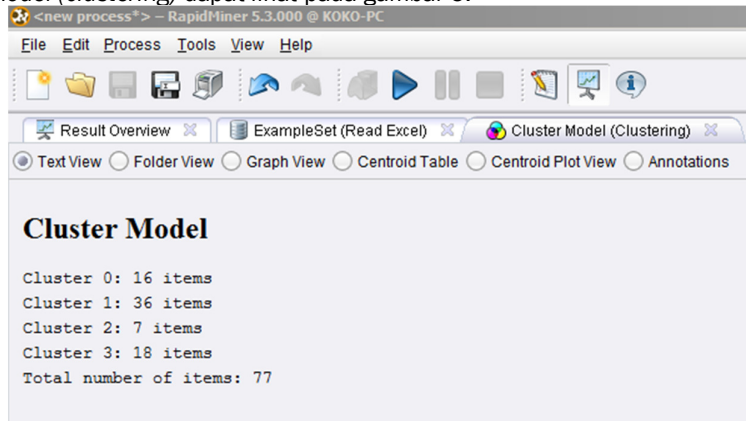
dilihat dari *example set (read excel)* dapat lihat pada gambar 4.

Gambar 4 Tampilan *ExampleSet (Read Excel) Flot View*

Merupakan *sheet* untuk menampilkan *database* yang telah diolah secara keseluruhan lengkap dengan *clusternya*. Urutan *clusternya* yaitu *cluster_1*, *cluster_0*, *cluster_3*, *cluster_2*.

3. TEXT VIEW

Dilihat dari *cluster model (clustering)* dapat lihat pada gambar 5:

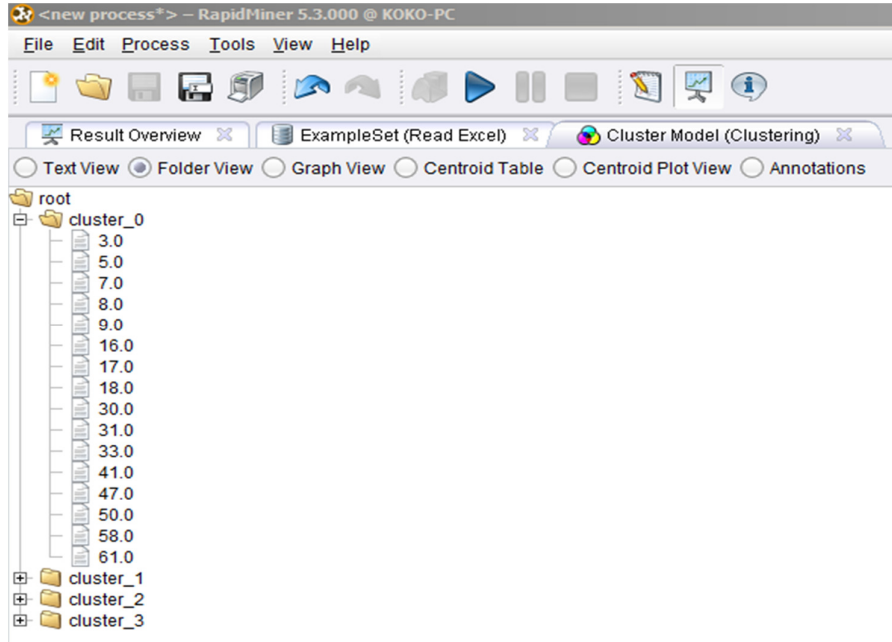


Gambar 5 Tampilan Cluster Model (Clustering) Text View

Gambar 5 merupakan *sheet* untuk menampilkan *database* yang telah diolah secara keseluruhan lengkap dengan *clusternya*. Hasil dari *data* 77 dimana C0 memiliki 16 *items*, C1 memiliki 36 *items*, C2 memiliki 7 *items*, C3 memiliki 18 *items*.

4. FOLDER VIEW

Dilihat dari *cluster model (clustering)* dapat lihat pada gambar 6:



Gambar 6 Tampilan Cluster Model (Clustering) Folder View

Pada Gambar 6 menampilkan *database* yang telah diolah secara keseluruhan lengkap dengan *clusternya*.

Jumlah anggota cluster :

C1 berjumlah 16 = 3,5,7,8,9,16,17,18,30,31,33,41,47,50,58,61

C2 berjumlah 36 = 1,2,4,6,10,19,20,21,22,23,24,25,29,34,35,36,45,48,49,52,53,54,55,57,60,63,64,65,68,69,70,71,72,73,75,77

C3 berjumlah 7 = 26,38,39,40,66,67,74.

C4 berjumlah 18 = 11,12,13,14,15,27,28,32,37,42,43,44,46,51,56,59,62,76

VI. KESIMPULAN

Dari uraian telah dikemukakan pada bab-bab sebelumnya maka dapat ditarik beberapa kesimpulan sebagai berikut

1. Algoritma *K-Means clustering* dapat diterapkan pada data mahasiswa TKJ Akademi Komunitas Solok Selatan, sehingga metode ini sangat membantu dalam mengelompokkan data mutu pembelajaran.
2. Menentukan *centroid* (titik pusat) pada tahap awal algoritma *K-Means* sangat berpengaruh pada hasil *cluster* seperti pada hasil pengujian yang dilakukan dengan menggunakan 25 *dataset* dengan *centroid* yang berbeda menghasilkan hasil *cluster* yang berbeda pula.
3. Data mutu pembelajaran Mahasiswa Program Studi TKJ Akademi Komunitas Solok Selatan bisa kita lihat dari *cluster* yaitu IP mahasiswa tergolong sedang, karena IP mahasiswa dari C1 ini lebih banya nilainya berkisar 2,50 s/d 3,00. Jarak tempuh mahasiswa tergolong jauh, karena jarak dari C2 tempat tinggal mahasiswa dengan kampus banyak yang 5 KM ke atas. Jumlah kehadiran tergolong sedang, karena jumlah kehadiran dari C3 berjumlah 85 s/d 90 kehadiran baik yang alfa maupun izin. Penghasilan orang tua tergolong rendah karena penghasilan orang tua dari C4 berkisar RP 0 s/d RP 1.500.000.

REFERENSI

- [1] R. Handoyo, "Perbandingan Metode Clustering Menggunakan Metode Single Linkage Dan K_Means Pada Pengelompokan Dokumen," Vols. 15, NO 2, OKTTOBER , no. 1412-0100, p. 1, 2015.
- [2] F. R. A. T. Tacbir Hendro Pudjiantoro, "Penerapan Data Mining Untuk Menganalisis Kemungkinan Pengunduran Diri Dari Mahasiswa Baru," no. KNS&111-009, p. 1, 2011.
- [3] B. S. d. A. R. B. Tahta Alfina, "Analisa Perbandingan Metode Hierarchical Clustering, K-Means Dan Gabungan Keduanya Dalam Cluster Data (studi kasus : problerm kerja praktek jurusan teknik industri ITS)," Vols. 1, (sept, 2012), no. 2301-9271, p. 2, 2012.
- [4] Lindawati, "Data Mining Dengan Teknik Clustering Dalam Pengklasipikasian Data Mahasiswa Studi Kasus Prediksi Lama Studi Mahasiswa Universitas Bina Nusantara," no. 1979-2328, p. 2, 2008.
- [5] Suprihatin, "Klastering K-means untuk Penentuan," Vols. JUSI Vol 1, No. 1, no. ISSN 2087-8737, p. 1, Februari 2011.
- [6] M. N. M. N. S. Ediyanto, "pengklasifikasikan karakteristik dengan metode K-Means cluster analysis," Vols. 02, No.2, hal 133 - 136, 2013.