

Penerapan Data Mining Untuk Pengelompokan Siswa Berdasarkan Nilai Akademik dengan Algoritma K-Means

Penda Sudarto Hasugian¹, Jijon Raphita Sagala^{2,*}

¹Teknologi Rekayasa Komputer Jaringan, STMIK Pelita Nusantara, Medan, Indonesia

²Teknik Informatika, STMIK Pelita Nusantara, Medan, Indonesia

Email: ¹penda.hasugian@gmail.com, ^{2,*}sisagala@gmail.com

Email Penulis Korespondensi: sisagala@gmail.com

Abstrak—Proses Data mining dengan menerapkan algoritma K-Means dilakukan untuk mengelompokkan data kedalam bentuk satu atau lebih kelompok, dimana data yang memiliki representasi persamaan dikelompokkan dalam satu kelompok dan data yang memiliki perbedaan masuk kedalam kelompok yang lain. Pengelompokan data siswa dilakukan untuk memudahkan sekolah dalam memfasilitasi siswa berdasarkan perbedaan kemampuannya dalam belajar dan mengikuti pembelajaran yang terdiri dari kelompok atau kelas siswa unggulan, kelompok sedang dan rendah. Aplikasi Data yang digunakan untuk proses perhitungan yaitu data siswa berdasarkan penilaian yang terpusat dalam menyajikan laporan hasil belajar siswa menggunakan hasil nilai rapor yaitu rapid miner. Penilaian tersebut menjadi dasar atribut yang digunakan dalam proses perhitungan untuk menentukan siswa kelas unggulan, sedang dan rendah. Tujuan penelitian ini adalah untuk mengelola data penilaian yang terpusat dalam menyajikan laporan hasil belajar siswa dan pengelompokan siswa pada kelas unggulan dengan menerapkan algoritma K-means dan melakukan pengujian dengan menggunakan aplikasi rapidminer. Sehingga data siswa dapat dikelola dan dikelompokkan dengan efektif dan efisien.

Kata Kunci: Datamining; Algoritma K-means; Pengelompokan Data; Rapid Miner

Abstract—The data mining process by applying the K-Means algorithm is carried out to group data into one or more groups, where data that has representative similarities is grouped into one group and data that has differences is included in another group. Grouping student data is done to facilitate schools in facilitating students based on differences in their ability to learn and participate in learning which consists of groups or classes of superior students, medium and low groups. The data application used for the calculation process is student data based on a centralized assessment in presenting reports on student learning outcomes using the results of report cards, namely the rapid miner. This assessment forms the basis of the attributes used in the calculation process to determine superior, medium and low class students. The purpose of this study is to manage centralized assessment data in presenting reports on student learning outcomes and grouping students in superior classes by implementing the K-means algorithm and conducting tests using the rapidminer application. So that student data can be managed and grouped effectively and efficiently.

Keywords: Datamining; K-means Algorithm; Clustering; Rapid Miner

1. PENDAHULUAN

Perkembangan teknologi informasi semakin pesat dan berpengaruh terhadap perkembangan dan kemajuan berbagai bidang usaha, seperti bidang pendidikan. perkembangan teknologi informasi mendorong untuk terus berinovasi, salah satu bentuk inovasi dalam bidang pendidikan, adalah mengembangkan dan pemanfaatan aplikasi baik dalam proses pembelajaran maupun administratif sekolah. Sebagai upaya meningkatkan mutu pendidikan melalui layanan pembelajaran dilakukan pembentukan kelompok belajar siswa dalam kelas sesuai dengan kriteria-kriteria siswa. Data siswa dikelompokkan menjadi kelompok siswa kelas unggulan. Penentuan masuk kelas unggulan sangatlah penting dimana informasi yang dihasilkan dapat dijadikan sebagai acuan bahan pertimbangan bagi pihak sekolah untuk melakukan langkah secara terstruktur dalam upaya meningkatkan kualitas dan standar mutu pendidikan di SMP Swasta Agia Sophia. Langkah-langkah menentukan siswa masuk ke kelas unggulan di dapat hasil penilain. Dari hasil identifikasi data nilai rapor, menunjukan atribut penilaian beragam hasil. Secara garis besar memiliki persamaan nilai memiliki variasi kriteria jumlah yang cenderung setara antar individu siswa. Selain itu, penempatan kapasitas jumlah siswa kelas unggulan setiap tahun ajaran dapat berubah-ubah seiring dengan pengembangan sekolah dan jumlah penerimaan siswa. Pemecahan permasalahan pengelompokan siswa dengan data yang semakin banyak menjadi kurang efisien dan dibutuhkan pengelompokan siswa yang juga memiliki prestasi untuk masuk kelas unggulan telah ditetapkan. Pengelompokan siswa pada kelompok yang tepat akan mampu menghasikan kualitas pembelajaran menjadi lebih baik.

Data mining dikenal sejak tahun 1990-an, ketika adanya suatu pekerjaan yang memanfaatkan data menjadi suatu hal yang lebih penting dalam berbagai bidang, seperti marketing dan bisnis, sains dan teknik, serta seni dan hiburan. Sebagian ahli menyatakan bahwa data mining merupakan suatu langkah untuk menganalisis pengetahuan dalam basis data atau biasa disebut Knowledge Discovery in Database (KDD).[1]. Data mining atau knowledge discovery in database (KDD) adalah proses resourcing dan penggunaan data untuk menemukan pola atau hubungan dari sekumpulan data berukuran besar[2]. Hasil dari proses data mining dapat digunakan sebagai evaluasi pengambilan keputusan dimasa depan. [3] Data Mining adalah suatu proses pengolahan data yang bertujuan untuk mendapatkan, menggali dan menemukan pengetahuan yang tersembunyi dari sebuah dataset atau sekumpulan data yang berukuran sangat besar[4].

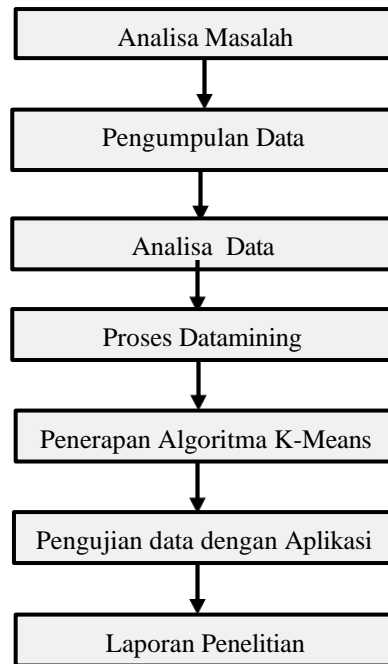
Salah satu metode yang dapat digunakan dalam Data Mining adalah metode Clustering dengan algoritma KMeans[5]. Algoritma K-Means Clustering adalah teknik dalam Data Mining yang mempartisi data yang ada ke dalam beberapa cluster sehingga data yang memiliki karakteristik yang sama akan dikelompokkan ke dalam satu cluster sedangkan data dengan karakteristik yang berbeda akan dikelompokkan ke dalam cluster lain.[6] Algoritma K-Means melakukan dua proses yaitu proses penentuan pusat cluster (centroid) dan proses pencarian anggota dari tiap-tiap cluster.

[7]. K-Means adalah merupakan salah satu metode dalam data mining yang dapat mengelompokkan data atau Clustering sebuah data kedalam bentuk satu cluster atau lebih cluster sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama dan data dengan karakteristik yang berbeda dikelompokkan ke dalam kelompok berbeda yang lainnya. [8]

2. METODOLOGI PENELITIAN

2.1 Kerangka Kerja Penelitian

Kerangka kerja penelitian merupakan tahapan-tahapan yang di lakukan oleh peneliti dalam menyelesaikan penelitian. Adapun langkah-langkah yang dilakukan dalam penelitian ini dapat dilihat pada gambar 1 berikut ini:



Gambar 1. Kerangka Kerja Penelitian

2.2 Uraian Kerangka Kerja Penelitian

Uraian dari kerangka kerja penelitian dilakukan berdasarkan gambar.1 sebagai berikut:

a. Analisa Masalah

Berdasarkan permasalahan yang telah di temukan maka akan dilakukan analisa, tahapan dalam proses analisa masalah dilakukan untuk memahami masalah yang telah di tentukan. Dengan menganalisa permasalahan yang telah di tentukan tersebut, maka di harapkan masalah tersebut dapat dipahami dengan baik. Pada tahap ini penulis menentukan data apa saja yang akan di kumpulkan yang berhubungan dengan penelitian ini.

b. Pengumpulan Data

Pengumpulan data di dalam penelitian dilakukan dengan cara :

1. Observasi adalah mengamati secara langsung terhadap objek yang diteliti dengan mengamati kegiatan-kegiatan yang ada disekolah mengenai hal-hal yang berkaitan dengan proses belajar mengajar disekolah SMP Swasta Agia Sophia
2. Wawancara adalah suatu cara dalam mengumpulkan data dengan cara menanyakan secara langsung kepada pihak sekolah serta siswa yang bersekolah di SMP Swasta Agia Sophia
3. Studi pustaka yaitu suatu data teori-teori yang ada hubungannya dengan masalah atau hal yang akan diteliti dan di butuhkan di dalam penelitian penulis yang nantinya akan di dibandingkan dengan kenyataan yang terjadi dan yang di peroleh penulis di lapangan.

c. Analisis Data

Pada saat melakukan tahap penganalisaan data di mana data yang telah di peroleh pada tahap proses pengumpulan data, data disusun dan di kelompokkan dalam bentuk tabel yang mudah di pahami dan di mengerti. Kemudian data yang di gunakan di pastikan data yang akurat, jelas, dan sesuai dengan kebutuhan penelitian.

d. Proses Data Mining

Pada tahapan ini dilakukan prose pembersihan data yang bertujuan untuk membuang data yang tidak konsisten dan noise selanjutnya melakukan Integrasi data atau penggabungan data dari beberapa sumber data, melakukan transformasi data yaitu bertujuan untuk menguiah data menjadi bentuk yang sesuai untuk di-mining.

e. Penerapan Algoritma K-Means

- Melakukan perhitungan data yang telah diperoleh dengan menerapkan rumus atau langkah-langkah algoritma k means.
- f. Pengujian dengan aplikasi rapidminer
Dari hasil perhitungan yang telah dilakukan, maka pada tahapan ini dilakukan pengujian dengan menggunakan aplikasi pengolahan data yaitu rapid miner untuk menguji dan memastikan hasil penerapan algoritma k-means.
 - g. Penyusunan Laporan Penelitian
Hasil akhir penelitian dibuat berbentuk laporan dan juga publikasi di jurnal ilmiah.

3. HASIL DAN PEMBAHASAN

3.1 Analisa Data

Data yang digunakan bersumber dari SMP swasta Agia Sophia serta adapn data yang akan digunakan daam proses perhitngan dengan algoritma k-means seperti pada pada tabel 1 berikut.

Tabel. 1 Data Awal Sebagai data Primer

No	Nama	Nilai Raport	Nilai Tes Ujian
1	Siswa001	80	85
2	Siswa002	90	87
3	Siswa003	95	90
4	Siswa004	75	78
5	Siswa005	70	89
6	Siswa006	80	70
7	Siswa007	65	75
8	Siswa008	90	71
9	Siswa009	91	75
10	Siswa010	85	70

3.1.1 Penerapan Algoritma K-Means

Proses penerapan agoritma K-means bertujuan untuk mengelompokkan data sesuai dengan keompok yang akan dibentuk[9]. Berikut ini tahapan dalam penerapan algoritma k-means[10]:

- a. Menentukan jumlah kelompok data atau jumlah Cluster
Jumlah cluster yang dibentuk sebanyak 3 cluster, sehingga nilai k=3. Yaitu:
C1 = Unggul, C2 = Sedang, C3 = Rendah.
- b. Menentukan titik pusat cluster
Titik pusat cluster diambil dari data primer yaitu data siswa secara acak. Sebanyak 3 cluster. Yaitu data pada nomor 3, nomor 6 dan nomor 7 Seperti pada tabel 2.

Tabel 2. Pusat Centroid Awal

Centroid	Nama	Nilai Raport	Nilai Tes Ujian	Keadaan
Centroid 1	Siswa003	95	90	Unggul
Centroid 2	Siswa006	80	70	Sedang
Centroid 3	Siswa007	65	75	Rendah

- c. Menghitung jarak dari setiap data terhadap pusat cluster dengan menggunakan rumus [11]:

$$d_{ij} = \sqrt{\sum_{j=1}^m (x_{ij} - c_{kj})^2} \quad (1)$$

Dimana Jarak data pertama pada pusat klaster pertama:

$$d_{11} = \sqrt{(80 - 95)^2 + (85 - 90)^2} = 15,811$$

Jarak data pertama pada pusat klaster kedua adalah:

$$d_{12} = \sqrt{(80 - 90)^2 + (85 - 70)^2} = 15$$

Jarak data pertama pada pusat klaster ke tiga adaah

$$d_{13} = \sqrt{(80 - 65)^2 + (85 - 75)^2} = 18,027.$$

Hasil perhitngan jarak untuk setiap data terhadap masing masing pusat klaster seperti pada tabel 3.

Tabel 3. Perhitungan data terhadap setiap pusat kluster

No	Nama	Nilai Raport	Nilai Tes Ujian	C1	C2	C3	Jarak Terdekat
1	Siswa001	80	85	15,81139	15	18,02776	15

2	Siswa002	90	87	5,830952	19,72308	27,73085	5,830951895
3	Siswa003	95	90	0	25	33,54102	0
4	Siswa004	75	78	23,32381	9,433981	10,44031	9,433981132
5	Siswa005	70	89	25,01999	21,47091	14,86607	14,86606875
6	Siswa006	80	70	25	0	15,81139	0
7	Siswa007	65	75	33,54102	15,81139	0	0
8	Siswa008	90	71	19,64688	10,04988	25,31798	10,04987562
9	Siswa009	91	75	15,52417	12,08305	26	12,08304597
10	Siswa010	85	70	22,36068	5	20,61553	5

Dari hasil perhitungan jarak yang dilakukan maka diperoleh kelompok data untuk masing masing kluster yang diambil berdasarkan jarak terdekat atau terecil yaitu:

Anggota untuk kluster 1 adalah data ke 2 dan 3, anggota kluster 2 adalah data 1,4,6,8,9 dan 10. Sedangkan untuk kluster ke 3 adalah data ke 5 dan 7. Seperti pada tabel 4.

Tabel 4. Anggota setiap cluster

No	Nama	Nilai Raport	Nilai Tes Ujian	C1	C2	C3
1	Siswa001	80	85		*	
2	Siswa002	90	87	*		
3	Siswa003	95	90	*		
4	Siswa004	75	78		*	
5	Siswa005	70	89			*
6	Siswa006	80	70		*	
7	Siswa007	65	75			*
8	Siswa008	90	71		*	
9	Siswa009	91	75		*	
10	Siswa010	85	70		*	

d. Hitung Pusat Cluster Baru

Untuk menghitung pusat cluster yang baru dimana data yang digunakan adalah data dari masing masing cluster. Dengan menggunakan rumus [12]:

$$c_{kj} = \frac{\sum_{h=1}^p y_{hj}}{p}; y_{hj} = x_{hj} \in clusterke - k \quad (2)$$

Untuk cluster 1 (C1) ada 2 data yaitu data ke 2 dan 3 dengan nilai (90,87) dan (95,90). Dimana:

$$C_{11} = \frac{90 + 95}{2} = 92,5$$

$$C_{12} = \frac{87 + 90}{2} = 88,5$$

Untuk Kluster ke 2 (C2) ada 6 data yaitu data ke 1,4,6,8,9,dan 10 dimana:

$$C_{21} = \frac{80 + 75 + 80 + 90 + 91 + 85}{6} = 83,5$$

$$C_{22} = \frac{85 + 78 + 70 + 71 + 75 + 70}{6} = 74,83$$

Untuk Kluster ke 3 (C3) ada 2 data yaitu data ke 5 dan 7 dengan dimana:

$$C_{31} = \frac{70 + 65}{2} = 67,5$$

$$C_{32} = \frac{89 + 75}{2} = 82$$

Sehingga diperoleh pusat kluster yang baru adalah seperti pada tabel 5.

Tabel 5. Pusat Kluster Baru

K1	92,5	88,5
K2	83,5	74,83
K3	67,5	82

e. Menghitung kembali jarak data terhadap masing masing pusat kluster yang baru.[13]

Dimana Jarak data pertama pada pusat klaster pertama:

$$d_{11} = \sqrt{(80 - 92,5)^2 + (85 - 88,5)^2} = 12,98$$

Jarak data pertama pada pusat kluster kedua adalah:

$$d_{12} = \sqrt{(80 - 83,5)^2 + (85 - 74,83)^2} = 10,75$$

Jarak data pertama pada pusat kluster ke tiga adaah

$$d_{13} = \sqrt{(80 - 67,5)^2 + (85 - 82)^2} = 12,85$$

Berikut hasil perhitungan jarak untuk setiap data terhadap masing masing pusat kluster yang baru seperti pada tabel 6

Tabel 6. Hasil Perhitungan jarak dengan pusat cluster baru iterasi 1

No	Nama	Nilai Raport	Nilai Tes Ujian	C1	C2	C3	Jarak Terdekat
1	Siswa001	80	85	12,98075	10,75226	12,85496	10,75226074
2	Siswa002	90	87	2,915476	13,79412	23,04886	2,915475947
3	Siswa003	95	90	2,915476	19,0336	28,64001	2,915475947
4	Siswa004	75	78	20,40833	9,07071	8,5	8,5
5	Siswa005	70	89	22,50555	19,56897	7,433034	7,433034374
6	Siswa006	80	70	22,32711	5,967505	17,32772	5,967504597
7	Siswa007	65	75	30,63495	18,50075	7,433034	7,433034374
8	Siswa008	90	71	17,67767	7,546154	25,04496	7,546154282
9	Siswa009	91	75	13,58308	7,501852	24,5204	7,501851623
10	Siswa010	85	70	19,96246	5,060742	21,2191	5,06074215

Sehingga diperoleh jarak terdekat untuk setiap data terhadap pusat Cluster yaitu pada cluster 1 (C1) ada 2 data, yaitu data ke 2 dan 3. Pada Cluster 2(C2) ada 5 data yaitu data ke 1, 6, 8, 9 dan 10. Sedangkan pada Custer ke 3 (C3) ada 3 data yaitu data ke 4, 5 dan 7.

Perhitungan jarak data ntuk setiap pusat cluster dialkukan kembali sampai setiap anggota cluster tidak berubah posisi. Seperti pada tabel 7.

Tabel 7. Posisi Data untuk setiap cluster pada iterasi 2

No	Nama	Nilai Raport	Nilai Tes Ujian	C1	C2	C3
1	Siswa001	80	85			*
2	Siswa002	90	87	*		
3	Siswa003	95	90	*		
4	Siswa004	75	78			*
5	Siswa005	70	89			*
6	Siswa006	80	70		*	
7	Siswa007	65	75			*
8	Siswa008	90	71		*	
9	Siswa009	91	75		*	
10	Siswa010	85	70		*	

Tabel 8. Posisi Data untuk setiap cluster pada iterasi 3

No	Nama	Nilai Raport	Nilai Tes Ujian	C1	C2	C3
1	Siswa001	80	85			*
2	Siswa002	90	87	*		
3	Siswa003	95	90	*		
4	Siswa004	75	78			*
5	Siswa005	70	89			*
6	Siswa006	80	70		*	
7	Siswa007	65	75			*
8	Siswa008	90	71		*	
9	Siswa009	91	75		*	
10	Siswa010	85	70		*	

Pada iterasi ke 3 posisi data untuk masing masing cluster sama dengan posisi data pada iterasi ke 2. Sehingga poses perhitungan jarak dihentikan. Sehingga diperoleh hasil akhir dari 3 cluster yang dibentuk yaitu:

- Cluster 1 (C1) dengan pusat cluster (92,5 dan 88,5) yang diartikan cluster Unggul yaitu ada 2 siswa, Siswa002 dan siswa003.
- Cluster 2 (C2) dengan pusat cluster (86,5 dan 71,5) yang merupakakn cluster Sedang, dimana ada 4 data yang termasuk didalam cluster ini yaitu Siswa006, Siswa008, Siswa009, dan Siswa010.

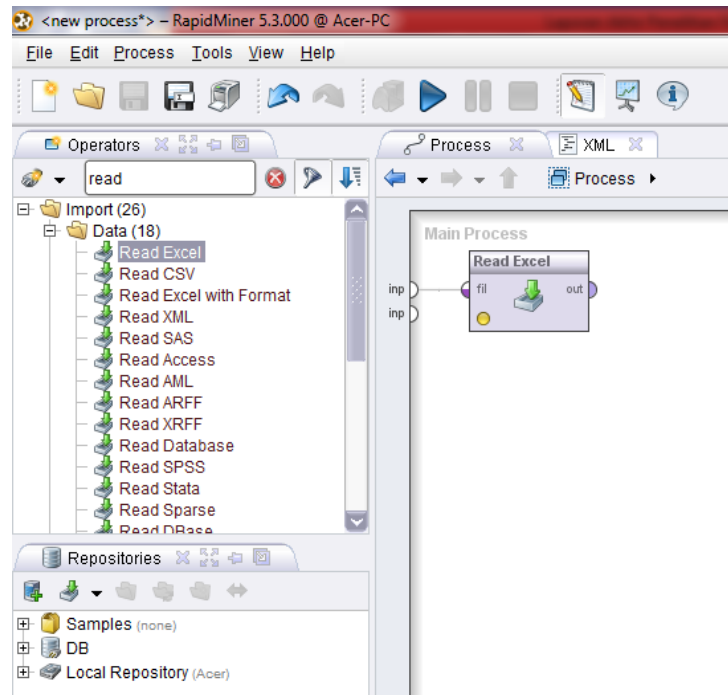
- c. Cluster 3 (C3) dengan pusat cluster (72,5 dan 81,75) merupakan cluster Rendah dimana ada 4 data yang tergolong kedalam cluster ini yaitu Siswa001, Siswa004, Siswa005 dan Siswa007.

3.2 Pengujian dengan Aplikasi Rapid Miner

Pengujian dilakukan untuk menyesuaikan perhitungan data yang dengan menerapkan tahapan dan rumus k-means, berikut tahapan pengujian yang dilakukan[14]:

a. Pengolahan data

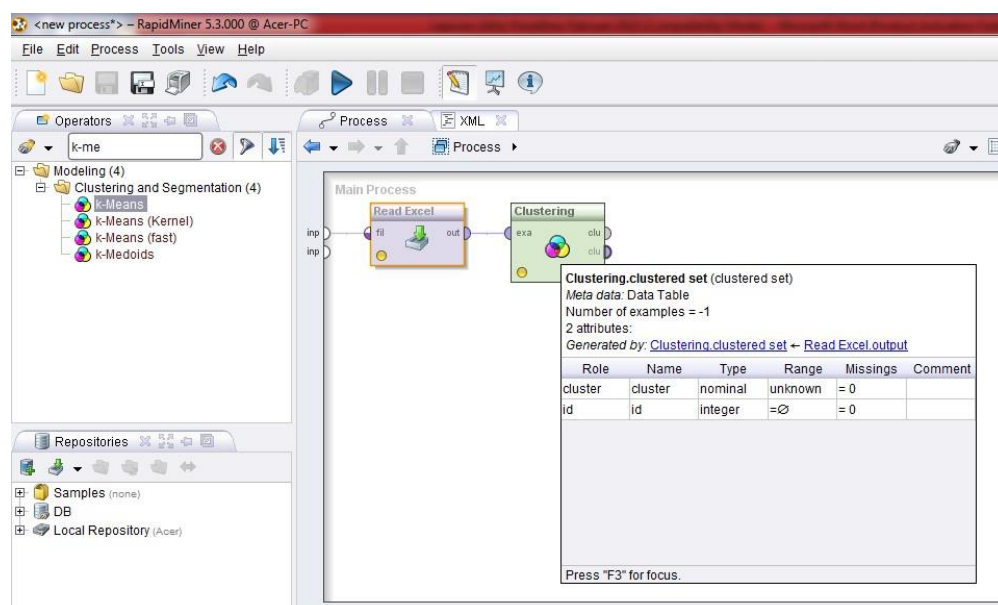
Data yang akan diuji diinputkan kedalam rapidminer dengan format data Excel seperti pada gambar 2:



Gambar 2. Input data Awal

b. Proses data dengan K-Means

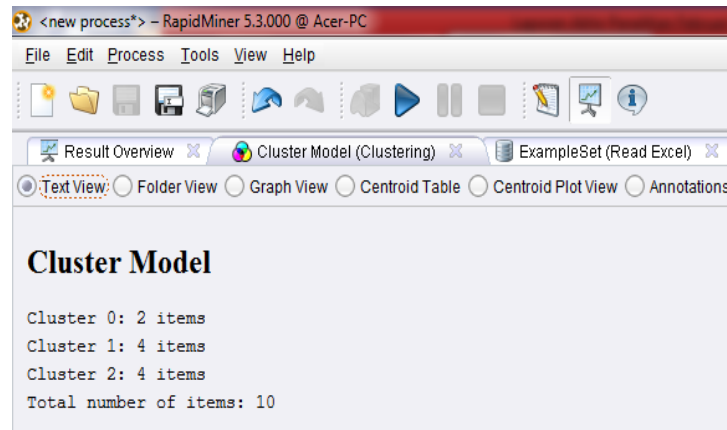
Tahapan berikutnya yaitu memilih metode k-means dimana perlu dilakukan menghubungkan data read excel dengan clustering serta menentukan jumlah cluster yang akan dibentuk[15]. Seperti pada gambar 3



Gambar 3. Memilih metode K-Means

c. Hasil pengujian dengan aplikasi rapid miner

Hasil pengujian memperoleh 3 cluster yaitu cluster 0 ada 2 data, cluster 1 ada 4 data dan cluster 2 ada 4 data. Sesuai dengan perhitungan yang dilakukan diatas. Berikut cluster data yang terbentuk seperti pada gambar 4.



Gambar 4. Hasil Cluster Data dengan Rapidminer

4. KESIMPULAN

Kesimpulan yang dapat diambil dari proses penelitian dengan menerapkan tahapan data mining dengan algoritma k-means clustering untuk melakukan pengelompokan data atau clustering terhadap data siswa yang dikelompokkan menjadi 3 cluster yaitu cluster siswa unggul, cluster sedang dan cluster rendah berdasarkan nilai akademik. Hasil penerapan metode k-means hitung manual dengan pengujian menggunakan aplikasi rapid miner dari data nilai siswa adaah sama dimana Cluster 1 (C1) cluster unggul terdapat 2 siswa, Siswa002 dan siswa003. Cluster 2 (C2) yang merupakan cluster Sedang, dimana ada 4 data yang termasuk didalam cluster ini yaitu Siswa006, Siswa008, Siswa009, dan Siswa010. Cluster 3 (C3) merupakan cluster rendah dimana ada 4 data yang tergolong kedalam cluster ini yaitu Siswa001, Siswa004, Siswa005 dan Siswa007. Proses klasterisasi memberikan hasil klasifikasi pengelompokan data yang efektif. Sehingga dapat menghemat waktu dalam melakukan klasterisasi kelas siswa..

REFERENCES

- [1] Yuli Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4 . 5 Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database (KDD) . Jurnal Edik Informatika," J. Edik Inform., vol. 2, no. 2, pp. 213–219, 2019.
- [2] P. M. S. Tarigan, J. T. Hardinata, H. Qurniawan, M. Safii, and R. Winanjaya, "Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang," J. Janitra Inform. dan Sist. Inf., vol. 2, no. 1, pp. 9–19, 2022, doi: 10.25008/janitra.v2i1.142.
- [3] P. N. Harahap and S. Sulindawaty, "Implementasi Data Mining Dalam Memprediksi Transaksi Penjualan Menggunakan Algoritma Apriori (Studi Kasus PT.Arma Anugerah Abadi Cabang Sei Rampah)," Matics, vol. 11, no. 2, p. 46, 2020, doi: 10.18860/mat.v11i2.7821.
- [4] A. Asroni and R. Adrian, "Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang," Semesta Tek., vol. 18, no. 1, pp. 76–82, 2016, doi: 10.18196/st.v18i1.708.
- [5] Y. A. Wijaya et al., "K-Means Di Sekolah Menengah Kejuruan Wahidin Kota Cirebon," vol. 6, no. 2, pp. 552–559, 2022.
- [6] W. Sirait, S. Defit, and G. W. Nurcahyo, "Algoritma K-Means Untuk Klasterisasi Tugas Akhir Mahasiswa Berdasarkan Keahlian," J. Sistim Inf. dan Teknol., vol. 1, no. 3, pp. 25–30, 2019, doi: 10.35134/jsisfotek.v1i3.6.
- [7] W. Dhuhita, "Clustering Menggunakan Metode K-Mean Untuk Menentukan Status Gizi Balita," J. Inform. Darmajaya, vol. 15, no. 2, pp. 160–174, 2015.
- [8] R. K. Dinata, S. Safwandi, N. Hasdyna, and N. Azizah, "Analisis K-Means Clustering pada Data Sepeda Motor," INFORMAL Informatics J., vol. 5, no. 1, p. 10, 2020, doi: 10.19184/isj.v5i1.17071.
- [9] B. Harahap, "Penerapan Algoritma K-Means Untuk Menentukan Bahan Bangunan Laris (Studi Kasus Pada UD. Toko Bangunan YD Indarung)," Reg. Dev. Ind. Heal. Sci. Technol. Art Life, pp. 394–403, 2019, [Online]. Available: <https://ptki.ac.id/jurnal/index.php/readystar/article/view/82>
- [10] Y. R. Sari, A. Sudewa, D. A. Lestari, and T. I. Jaya, "Penerapan Algoritma K-Means Untuk Clustering Data Kemiskinan Provinsi Banten Menggunakan Rapidminer," CESS (Journal Comput. Eng. Syst. Sci., vol. 5, no. 2, p. 192, 2020, doi: 10.24114/cess.v5i2.18519.
- [11] M. N. Sutoyo, "Algoritma K-Means," no. 1, pp. 1–7, 2019.
- [12] I. Sumadikarta and E. Abeiza, "PENERAPAN ALGORITMA K-MEANS PADA DATA MINING UNTUK MEMILIH PRODUK DAN PELANGGAN POTENSIAL (Studi Kasus : PT Mega Arvia Utama)," J. Satya Inform., no. 1, pp. 1–12, 2014.
- [13] Z. Zulham and B. S. Hasugian, "Pengelompokan Siswa Dalam Menentukan Penerima Beasiswa Berdasarkan Prestasi Akademik Dengan Algoritma K-Means," War. Dharmawangsa, vol. 16, no. 3, pp. 231–241, 2022, doi: 10.46576/wdw.v16i3.2220.
- [14] S. N. Br Sembiring, H. Winata, and S. Kusnasari, "Pengelompokan Prestasi Siswa Menggunakan Algoritma K-Means," J. Sist. Inf. Triguna Dharma (JURSI TGD), vol. 1, no. 1, p. 31, 2022, doi: 10.53513/jursi.v1i1.4784.
- [15] F. Rini, N. Kahar, and Juliana, "Penerapan Algoritma K-Means Pada Pengelompokan Data Siswa Baru Berdasarkan Jurusan Di Smk Negeri 1 Kota Jambi Berbasis Web," Semin. Nas. APTIKOM, pp. 94–99, 2016.