

Data Mining dalam Mengukur Tingkat Keaktifan Siswa dalam Mengikuti Proses Belajar pada SMP IT Andalas Cendekia dengan Menggunakan Metode *K-Means Clustering*

Melissa Triandini^{1✉}, Sarjon Defit², Gunadi Widi Nurcahyo³

¹²³Universitas Putra Indonesia YPTK Padang
melisat.a91@gmail.com

Abstract

The learning process is essentially to develop the activities and creativity of students through various interactions and learning experiences. The teacher is the most important factor in the process of improving the quality of education. In addition, student learning activeness is also an important basic element for the success of the learning process. The quality and activeness of students in learning at school has a lot of diversity which makes students have different levels of understanding, this needs to be a concern for the school, especially teachers as teachers and educators of students in schools. The purpose of this study is to measure the extent to which students' ability to undergo the learning process as well as a reference and evaluation material for the school in the success of educators when carrying out the teaching and learning process. In this study the data were sourced from the Integrated Islamic Junior High School Andalas Cendekia Dharmasraya which consisted of several variables, namely the presence of student data, Academic value (knowledge), Psychomotor value (skills), Affective value (spiritual and social). In grouping the data, the appropriate method in this study is the Clustering method with the K-Means Algorithm. The results of this study obtained 3 groupings of students, namely students who are very active, students who are active and students who are less active. This research is used as a guideline for teachers in the field of study in selecting students to participate in competitions and Olympics, and can be used as a benchmark for schools of the ability of educators in the teaching and learning process.

Keywords: Data Mining; Clustering; K-Means; PBM; Student Activity.

Abstrak

Proses pembelajaran pada hakekatnya untuk mengembangkan aktivitas dan kreatifitas peserta didik melalui berbagai interaksi dan pengalaman belajar. Guru merupakan faktor terpenting dalam proses peningkatan mutu Pendidikan. Selain itu, keaktifan belajar siswa juga merupakan unsur dasar yang penting bagi keberhasilan proses pembelajaran. Kualitas dan keaktifan siswa dalam belajar di sekolah memiliki banyak keragaman yang membuat siswa memiliki tingkat pemahaman yang berbeda-beda, ini perlu menjadi perhatian bagi pihak sekolah khususnya guru selaku pengajar dan pendidik siswa di sekolah. Adapun tujuan dari penelitian ini adalah untuk mengukur sampai sejauh mana kemampuan siswa dalam menjalani proses pembelajaran serta menjadi acuan dan bahan evaluasi bagi pihak sekolah dalam keberhasilan para pendidik saat melaksanakan proses belajar mengajar. Pada penelitian ini data bersumber dari Sekolah Menengah Pertama Islam Terpadu Andalas Cendekia Dharmasraya yang terdiri dari beberapa variabel, yaitu kehadiran data siswa, nilai Akademik (pengatahuan), nilai Psikomotor (keterampilan), nilai Afektif (spiritual dan sosial). Dalam pengelompokan data tersebut metode yang tepat dalam penelitian ini adalah metode Clustering dengan Algoritma K-Means. Hasil dari penelitian ini mendapatkan 3 pengelompokan siswa, yaitu siswa yang sangat aktif, siswa aktif dan siswa yang kurang aktif. Penelitian ini digunakan sebagai pedoman bagi guru bidang studi dalam pemilihan siswa untuk diikutsertakan pada perlombaan maupun olimpiade, serta dapat dijadikan sebagai tolak ukur bagi sekolah terhadap kemampuan para pendidik dalam proses belajar mengajar.

Kata kunci: *Data Mining*; *Clustering*; *K-Means*; *PBM*; Keaktifan Siswa.

© 2021 JIdT

1. Pendahuluan

Salah satu prioritas kebijakan umum pembangunan di Indonesia adalah peningkatan mutu pendidikan. Pendidikan merupakan suatu cara pembelajaran untuk meningkatkan pengetahuan, sikap dan keterampilan seseorang [1]. Sekolah menjadi sarana berlangsungnya pendidikan secara langsung dengan bertemunya guru dan siswa. Guru merupakan faktor terpenting dalam proses peningkatan mutu pendidikan. Selain itu, keaktifan belajar siswa juga merupakan unsur dasar yang penting bagi keberhasilan proses pembelajaran [2]. Kualitas dan keaktifan siswa dalam belajar di

sekolah memiliki banyak keragaman yang membuat siswa memiliki tingkat pemahaman yang berbeda-beda, ini perlu menjadi perhatian bagi pihak sekolah khususnya guru selaku pengajar dan pendidik siswa di sekolah.

Dalam penentuan keaktifan siswa dalam proses belajar mengajar pada sekolah SMP IT Andalas Cendekia ini adalah dengan penerapan *Data Mining*. Adapun tahapan pada *Data Mining* adalah Pembersihan data (*data cleaning*), Integrasi data (*data integration*), Seleksi data (*data selection*), Tranformasi data (*data transformation*), Proses *mining* Evaluasi pola yang

ditemukan (untuk menemukan informasi dan pengetahuan yang menarik) dan yang terakhir Proses Pola (*pattern evaluation*) [3]. Teknik Data Mining dapat mengolah data yang berlimpah menjadi informasi yang penting biasanya disebut knowledge discovery database (KDD) [4]. KDD adalah metode yang digunakan untuk dapat memperoleh pengetahuan yang berasal dari database yang ada. Hasil pengetahuan yang diperoleh dapat dimanfaatkan untuk basis pengetahuan (*knowledge base*) yang digunakan dalam keperluan mengambil keputusan [5]. Adapun metode yang digunakannya adalah clustering dengan Algoritma K-Means [6] [7].

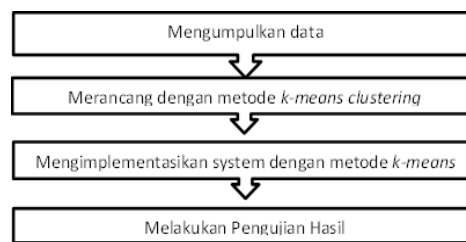
Clustering adalah sebuah proses untuk mengelompokkan data ke dalam beberapa cluster atau kelompok sehingga data dalam satu cluster memiliki tingkat kemiripan yang maksimum dan data antar cluster memiliki kemiripan yang minimum [8] [9]. K-Means merupakan algoritma pengelompokan secara iteratif yang melakukan partisipasi untuk mengklasifikasikan atau mengelompokkan sejumlah besar objek [10]. K-Means merupakan salah satu metode data clustering non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster/kelompok [11]. Kelebihan dari penerapan K-Means adalah mampu mengelompokkan objek besar dan mempercepat proses pengelompokan [12]. Tujuan algoritma ini adalah untuk membagi data menjadi beberapa kelompok.

Dalam penelitian melakukan pemetaan tingkat keaktifan siswa dalam menunjang untuk peningkatan proses belajar mengajar. Tenaga pendidik akan terbantu dalam pemilihan siswa yang akan direkomendasikan untuk mengikuti perlombaan atau penerimaan beasiswa. Dalam penelitian ini dapat melakukan prediksi terhadap tingkat keaktifan dan korelasi antara nilai dengan tingkat keaktifan.

2. Metodologi Penelitian

Dalam penelitian ini metode pengumpulan data yang digunakan adalah:

- Sumber data utama dari nilai rapor siswa sekolah menengah pertama islam terpadu andalas cendekia semester satu. Dilakukan dengan cara tanya jawab secara langsung dengan pihak sekolah atau guru.
- Studi Pustaka dilakukan dengan cara mempelajari dan mencari referensi pada jurnal maupun literature lain yang berhubungan dengan penelitian yang dilakukan yaitu pengelompokkan siswa menggunakan algoritma K-Means.
- Dokumentasi dan mengumpulkan data yang berhubungan dengan penelitian misalnya foto maupun data lain yang mendukung untuk melakukan analisis dalam menentukan keaktifan siswa dalam proses belajar mengajar.



Gambar 1. Metodologi Penelitian

Berdasarkan kerangka kerja pada Gambar 1 maka masing-masing langkahnya dapat diuraikan seperti di bawah ini:

- Mengumpulkan Data yaitu data yang digunakan dalam penelitian ini adalah data nilai rapor siswa SMP IT Andalas Cendekia Kab. Dharmasraya. Data tersebut terdiri dari nilai kognitif (pengetahuan), nilai psikomotorik (keterampilan), nilai afektif (sikap dan spiritual dan nilai sikap sosial).

Berikut sample data nilai rapor siswa SMP IT Andalas Cendekia sebanyak 39 siswa yang disajikan pada Tabel 1.

Tabel 1. Data Nilai Siswa Kelas 7 Semester 1

No	Nama	Peng	Ketr	Spi	Sos
1	Allunan Kasih Ananda	82.4	85.6	82.1	82.1
2	Alya Ramadhani	80.3	84.8	81.7	81.7
3	Amila Okta Fizela	77.7	85.1	81.8	81.8
4	Aqiela Putri Kusuma	73.1	83.0	78.7	77.0
5	Aqiilah Zahra Afif	84.1	83.1	67.0	80.5
6	Arifa Khairinnisw	73.0	83.3	71.8	79.0
7	Assyahirha Intana Alfent	87.5	88.9	82.4	83.3
8	Azizah Izati	81.3	86.1	81.7	81.6
9	Azzahra Febriani	85.8	89.1	75.8	83.3
10	Carissa Firtri Wardani	78.8	85.1	71.9	82.4
11	Chelshi Isra Anwar	83.2	86.6	79.2	82.3
12	Dhifa Dwi Harnesya	81.2	82.9	72.3	81.3
13	Dita Ariska	77.5	82.3	70.2	80.1
14	Dzakiya Qawlan Sadida	82.4	86.1	74.7	80.6
15	Ericha Oktavia Andini	85.1	86.8	83.9	83.2
16	Gadiza Rahma Luthfilah	83.3	88.2	75.0	83.0
17	Hilyah Mufidah Shafwah	84.5	86.7	76.6	81.0
18	Indira Puti Refhalino	88.2	88.7	75.5	81.2
19	Intan Khairunnisa	83.4	87.2	73.8	81.3
20	Irmacaty Ferikesya	84.5	86.6	81.9	81.9
21	Irsyadiyah Ratu Purnama	75.7	81.2	80.5	81.2
22	Keyla Marsha	82.7	86.7	71.6	83.5
23	Khansya Haura Janedis	84.2	86.7	81.9	81.9
24	Lova Hafiz Zahra	70.8	79.0	67.3	77.6
25	Monika Larasati	81.2	85.4	81.7	81.6
26	Mozza Musdalifah	74.7	81.8	72.9	79.5
27	Mutiara Islami	79.9	85.2	81.7	81.6
28	Nafiza Dwi Rahmadan	79.3	84.6	67.2	79.5
29	Najwatul Fajriah	79.2	82.6	66.6	78.4
30	Neva Zhafira Jupri	81.2	87.5	81.3	82.7
31	Pelangi Efriasa	81.8	85.1	75.9	82.6
32	Reni Edni Putri	79.9	85.8	83.2	82.7
33	Salma Khaira Ramadhani	85.2	88.4	74.8	83.1
34	Shelvira Rahmadhani	80.2	86.0	74.4	82.8
35	Shevia Sarche	80.5	85.3	81.6	81.5
36	Shakilla	80.9	85.7	83.6	82.4
37	Ulfa Salsabila	76.5	84.6	81.2	81.1
38	Zahra Adzkie Novandi	78.8	83.9	70.7	81.4
39	Zhandry Ristaviani	76.0	84.4	77.1	79.4

Berdasarkan Tabel 1 di atas, maka data yang akan diolah menggunakan metode *k-means clustering* sebanyak 39 data yang terdiri dari nilai peng (pengetahuan), ketr (keterampilan), spi (spiritual) dan

sos (sosial). Nilai tersebut diperoleh dari nilai rapor siswa kelas 7 semester satu.

b) Perancangan yaitu peneliti yang akan merancang berdasarkan data yang telah dikumpulkan dengan menggunakan algoritma *k-means clustering* [13]. Dengan cara mengelompokkan nilai siswa dari beberapa kelompok kedalam bentuk cluster kemudian dilakukan analisa terhadap pola pembentukan cluster kemudian menetapkan nilai centroid secara random, selanjutnya mengestimasi jarak pada data ke setiap centroid dengan menerapkan metode korelasi antara dua buah objek yang terdapat pada rumus *Euclidean Distance* pada Persamaan 1 [14] [15] [16].

$$D_{(i,j)} = \sqrt{(x_{i1} - x_{1j})^2 + (x_{i2} - x_{2j})^2 + \dots + (x_{ik} - x_{kj})^2} \quad (1)$$

di mana $D_{(i,j)}$ adalah Jarak data ke i ke pusat cluster j, X_{ki} adalah data ke i pada atribut data ke k, X_{kj} adalah titik pusat ke j pada atribut ke k.

Adapun algoritma yang digunakan pada metode *k-means clustering* ini dapat dijelaskan pada Gambar 2 [17] [18].

Algoritma *k-means clustering*:

- Tentukan sebanyak K kelompok yang kita inginkan
- Inisiasi centroid untuk setiap kelompok. Centroid ibarat seperti "ketua kelompok", yang mempresentasikan kelompok
- Hitung kedekatan suatu data terhadap centroid, kemudian masukkan data tersebut ke kelompok yang centroid-nya memiliki sifat terdekat dengan dirinya.
- Pilih kembali centroid baru untuk masing-masing kelompok, yaitu dari anggota kelompok itu sendiri dan lakukan perhitungan kedekatan suatu data terhadap centroid
- Ulangi langkah-langkah sebelumnya sampai tidak ada perubahan anggota untuk semua kelompok.

Gambar 2. Algoritma *k-means clustering*

- Menentukan jumlah *cluster*, $k = 3$
- Inisiasi centroid untuk setiap kelompok, adapun centroid yang dipilih secara random dapat di tampilkan [19] pada Tabel 2.

Tabel 2. Nilai Cluster Awal

K	Peng	Ketr	Spi	Sos
C0	83.3	88.2	75.1	83.0
C1	81.8	85.2	75.9	82.6
C2	82.4	85.6	82.2	82.1

- Menghitung jarak pusat *cluster* dengan *Euclidean Distance* [20]

$$D(1,1) = \sqrt{((82.35-83.27)^2 + (85.60-88.18)^2 + (82.12-75.01)^2 + (82.08-83.00)^2)} = 7.67$$

$$D(1,2) = \sqrt{((82.35-81.79)^2 + (85.60-85.13)^2 + (82.12-75.85)^2 + (82.08-82.55)^2)} = 6.33$$

$$D(1,3) = \sqrt{((82.35-82.35)^2 + (85.60-85.60)^2 + (82.12-82.12)^2 + (82.08-82.08)^2)} = 0$$

Dan seterusnya untuk data ke-2 sampai data ke-39 dengan menggunakan cluster awal, setelah dilakukan perhitungan maka di dapat hasil yang disajikan pada Tabel 3.

Tabel 3. Hasil Perhitungan Nilai Iterasi ke-1 Cluster 0

C0	C1	C2	Jarak Terdekat	Kelompok
2.77	5.65	8.07	2.77	Cluster 0
0.00	3.52	7.67	0.00	Cluster 0
3.18	3.59	6.18	3.18	Cluster 0
5.28	7.43	9.37	5.28	Cluster 0
2.30	3.50	8.55	2.30	Cluster 0
3.85	4.74	10.72	3.85	Cluster 0
1.99	4.89	8.41	1.99	Cluster 0

Berdasarkan Tabel 3 di atas, hasil perhitungan pada iterasi ke-1 untuk cluster 0 sebanyak 7 data yang disajikan pada Tabel 5.

Tabel 4. Hasil Perhitungan Nilai Iterasi ke-1 Cluster 1

C0	C1	C2	Jarak Terdekat	Kelompok
13.41	10.90	11.40	10.90	Cluster 1
9.85	9.59	15.52	9.59	Cluster 1
12.47	10.47	14.43	10.47	Cluster 1
6.27	4.92	10.81	4.92	Cluster 1
6.49	4.42	10.31	4.42	Cluster 1
10.00	8.06	13.47	8.06	Cluster 1
3.27	2.51	7.58	2.51	Cluster 1
18.15	16.03	20.44	16.03	Cluster 1
11.45	8.92	12.84	8.92	Cluster 1
10.09	9.51	15.46	9.51	Cluster 1
11.87	10.82	16.59	10.82	Cluster 1
3.52	0.00	6.33	0.00	Cluster 1
3.82	2.32	8.09	2.32	Cluster 1
7.66	6.14	12.06	6.14	Cluster 1
9.17	6.74	8.62	6.74	Cluster 1

Berdasarkan tabel 4 di atas, hasil perhitungan pada iterasi ke-1 untuk cluster 1 sebanyak 15 data yang disajikan pada Tabel 4.

.Tabel 4. Hasil Perhitungan Nilai Iterasi ke-1 Cluster 1

C0	C1	C2	Jarak Terdekat	Kelompok
7.67	6.33	0.00	0.00	Cluster 2
8.20	6.14	2.27	2.27	Cluster 2
9.38	7.27	4.66	4.66	Cluster 2
8.53	9.50	6.27	6.27	Cluster 2
7.40	5.99	1.35	1.35	Cluster 2
4.53	3.93	3.21	3.21	Cluster 2
9.21	8.91	3.67	3.67	Cluster 2
7.26	6.82	2.39	2.39	Cluster 2
11.81	8.72	8.23	8.23	Cluster 2
7.19	6.71	2.13	2.13	Cluster 2
7.65	5.93	1.36	1.36	Cluster 2
8.14	6.18	2.53	2.53	Cluster 2
6.63	5.94	2.44	2.44	Cluster 2
9.12	7.58	2.72	2.72	Cluster 2
7.85	5.97	2.08	2.08	Cluster 2
9.28	7.82	2.12	2.12	Cluster 2
10.05	7.69	6.14	6.14	Cluster 2

- Menentukan cluster baru dari itersai ke-1

$$C0 = (85.75 + 83.27 + 84.52 + 88.19 + 83.37 + 82.65 + 85.24) / 7 = 84.71$$

$$= (89.09 + 88.18 + 86.72 + 88.65 + 87.15 + 86.68 + 88.39) / 7 = 87.84$$

$$= (75.8 + 75.01 + 76.55 + 75.52 + 73.81 + 71.55 + 74.8) / 7 = 74.72$$

$$= (83.3 + 83 + 80.99 + 81.22 + 81.33 + 83.45 + 83.05) / 7 = 82.33$$

$$C1 = (73.09 + 84.06 + 72.99 + 78.84 + 81.24 + 77.47 + 82.4 + 70.76 + 74.65 + 79.31 + 79.17 + 81.79 + 80.24 + 78.82 + 76.01) / 15 = 78.06$$

$$= (83 + 83.10 + 83.34 + 85.07 + 82.93 + 82.31 + 86.11 + 79.02 + 81.84 + 84.56 + 82.59 + 85.13 + 85.96 + 83.94 + 84.42) / 15 = 83.55$$

$$= (78.69 + 66.98 + 71.84 + 71.91 + 72.25 + 70.16 + 74.7 + 67.3 + 72.91 + 67.22 + 66.56 + 75.85 + 74.35 + 70.73 + 77.08) / 15 = 71.90$$

= (77.02 + 80.54 + 78.95 + 82.39 + 81.34 + 80.11 + 80.64 + 77.55 + 79.52 + 79.48 + 78.36 + 82.55 + 82.8 + 81.41 + 79.39)/15 = 80.14

C2 = (82.35 + 80.3 + 77.73 + 87.53 + 81.26 + 83.2 + 85.05 + 84.51 + 75.66 + 84.16 + 81.17 + 79.94 + 81.21 + 79.92 + 80.45 + 80.86 + 76.45)/17 = 81.28

= (85.6 + 84.8 + 85.14 + 88.9 + 86.08 + 86.61 + 86.82 + 86.58 + 81.2 + 86.67 + 85.39 + 85.2 + 87.49 + 85.83 + 85.28 + 85.67 + 84.57)/17 = 85.76

= (82.12 + 81.73 + 81.83 + 82.36 + 81.67 + 79.2 + 83.94 + 81.89 + 80.45 + 81.89 + 81.67 + 81.67 + 81.27 + 83.16 + 81.58 + 83.59 + 81.17)/17 = 81.83

= (82.08 + 81.67 + 81.78 + 83.34 + 81.62 + 82.28 + 83.23 + 81.85 + 81.2 + 81.85 + 81.62 + 81.62 + 82.66 + 82.7 + 81.53 + 82.43 + 81.12)/17 = 82.03

Dari perhitungan diatas di peroleh nilai centroid baru yang disajikan pada Tabel 5.

Tabel 5. Nilai Centroid Baru

K-New	Peng	Ketr	Spi	Sos
C0	84.71	87.84	74.72	82.33
C1	78.06	83.55	71.90	80.14
C2	81.28	85.76	81.83	82.03

Setelah ada centroid baru maka di lakukan kembali perhitungan pada langkah no.2 dan di peroleh 5 iterasi sampai data cluster tidak berubah. Berikut tampilan iterasi ke-6.

Tabel 6. Hasil Perhitungan Iterasi Ke-5

No	C0	C1	C2	Jarak Terdekat
1	7.63	12.64	1.12	1.12
2	8.13	11.36	1.42	1.42
3	9.49	11.12	3.61	3.61
4	13.17	9.00	10.47	9.00
5	8.99	8.07	15.42	8.07
6	12.25	4.34	13.56	4.34
7	8.70	16.76	7.14	7.14
8	7.42	11.86	0.55	0.55
9	3.12	12.04	8.31	3.12
10	6.08	3.86	10.25	3.86
11	4.48	10.99	3.38	3.38
12	5.64	4.51	10.02	4.51
13	9.45	1.31	12.90	1.31
14	2.36	7.09	7.36	2.36
15	9.30	15.90	4.61	4.61
16	1.34	9.47	7.57	1.34
17	2.34	9.92	6.36	2.34
18	4.84	13.20	9.83	4.84
19	1.39	8.07	8.44	1.39
20	7.18	13.66	3.34	3.34
21	11.60	9.73	7.41	7.41
22	3.67	7.59	10.51	3.67
23	7.14	13.50	3.03	3.03
24	17.71	8.75	19.68	8.75
25	7.59	11.67	0.59	0.59
26	11.06	3.32	12.05	3.32
27	8.13	11.27	1.52	1.52
28	9.57	4.71	15.01	4.71
29	11.18	5.17	16.16	5.17
30	6.97	12.16	1.93	1.93
31	3.06	7.49	6.06	3.06
32	9.31	13.00	2.01	2.01
33	2.08	10.88	8.55	2.08
34	3.79	6.17	7.60	3.79
35	7.82	11.32	1.11	1.11
36	9.38	13.53	1.85	1.85
37	10.10	10.28	5.10	5.10
38	7.22	2.58	11.53	2.58
39	8.99	6.23	7.69	6.23

Dari tabel di atas dapat dikelompokkan untuk setiap cluster sebagai berikut yang disajikan pada Tabel 7 sampai Tabel 9.

Tabel 7. Hasil Perhitungan Iterasi Ke-5 Cluster 0

No	Nama Siswa	Kelompok
9	Azzahra Febriani	Cluster 0
14	Dzakiya Qawlan Sadida	Cluster 0
16	Gadiza Rahma Luthfilah	Cluster 0
17	Hilyah Mufidah Shafwah	Cluster 0
18	Indira Puti Refhalino	Cluster 0
19	Intan Khairunnisa	Cluster 0
22	Keyla Marsha	Cluster 0
31	Pelangi Efriasa	Cluster 0
33	Salma Khaira Ramadhani	Cluster 0
34	Shelvira Rahmadhani	Cluster 0

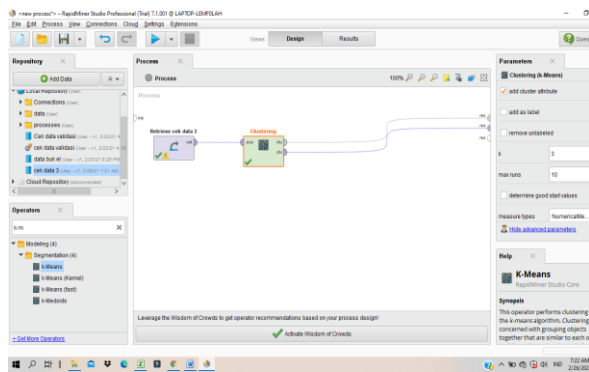
Tabel 8. Hasil Perhitungan Iterasi Ke-5 Cluster 1

No	Nama Siswa	Kelompok
4	Aqiela Putri Kusuma	Cluster 1
5	Aqillah Zahra Afif	Cluster 1
6	Arifa Khairinniswara	Cluster 1
10	Carissa Firtri Wardani	Cluster 1
12	Dhifa Dwi Harnesya	Cluster 1
13	Dita Ariska	Cluster 1
24	Lova Hafiz Zahra	Cluster 1
26	Mozza Musdalifah	Cluster 1
28	Nafiza Dwi Ramadhan	Cluster 1
29	Najwatul Fajriah	Cluster 1
38	Zahra Adzkia Novandi	Cluster 1
39	Zhandry Ristaviani	Cluster 1

Tabel 9. Hasil Perhitungan Iterasi Ke-5 Cluster 2

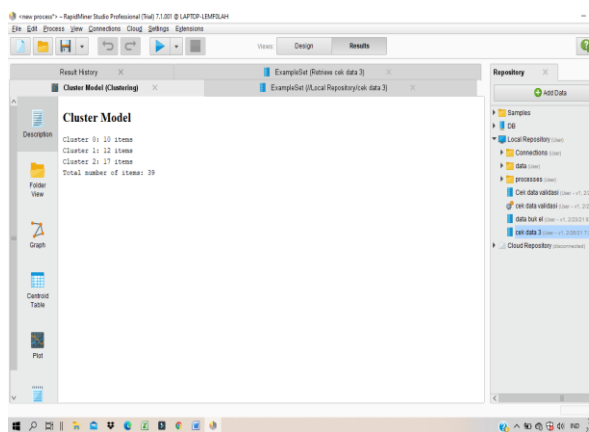
No	Nama Siswa	Kelompok
1	Allunan Kasih Ananda	Cluster 2
2	Alya Ramadhani	Cluster 2
3	Amila Okta Fizela	Cluster 2
7	Assyahirha Intana Alfent	Cluster 2
8	Azizah Izati	Cluster 2
11	Chelshi Isra Anwar	Cluster 2
15	Ericha Oktavia Andini	Cluster 2
20	Imacaty Ferikesya	Cluster 2
21	Irsyadiha Ratu Purnama	Cluster 2
23	Khansya Haura Janedis	Cluster 2
25	Monika Larasati	Cluster 2
27	Mutiara Islami	Cluster 2
30	Neva Zhafira Jupri	Cluster 2
32	Reni Edni Putri	Cluster 2
35	Shevia Sarche	Cluster 2
36	Shakilla	Cluster 2
37	Ulfa Salsabila	Cluster 2

- c) Implementasi : Pada tahap implementasi ini data yang telah diolah menggunakan metode *k-means clustering* akan diimplementasikan kembali menggunakan aplikasi yang mendukung data mining, adapun software yang digunakan adalah *Rapidminer*.
- d) Pengujian : Pada tahap ini peneliti akan menguji hasil dari perhitungan algoritma k-means clustering secara manual dengan software yang ada. Data yang diinputkan merupakan data berupa excel yang disajikan pada Gambar 2.



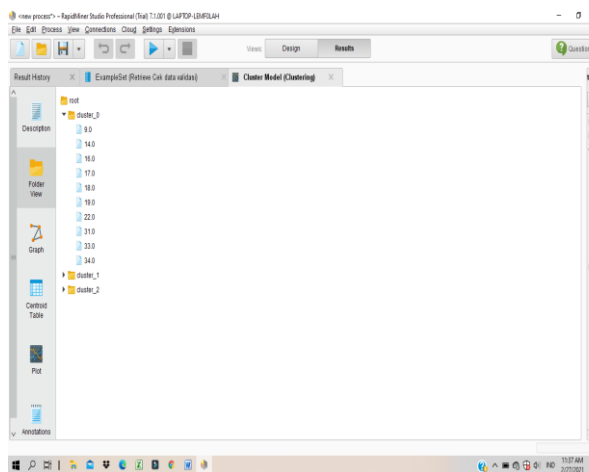
Gambar 2. Tampilan Input Data Excel

Setelah dilakukan input data maka akan diconecttikan dengan operator k-means lalu di proses, berikut hasil dari data proses tersebut yang disajikan pada Gambar 3.

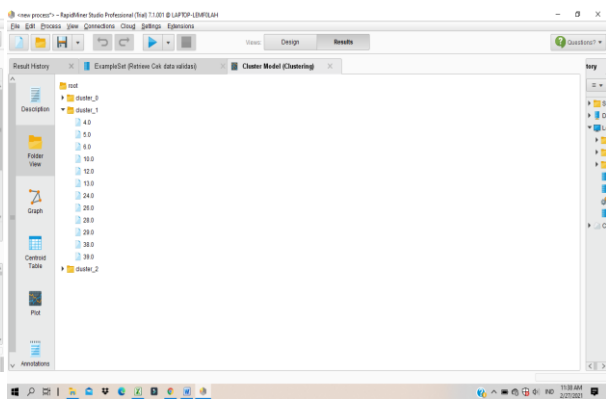


Gambar 3. Hasil Dari Proses K-Means Hasil dan Pembahasan

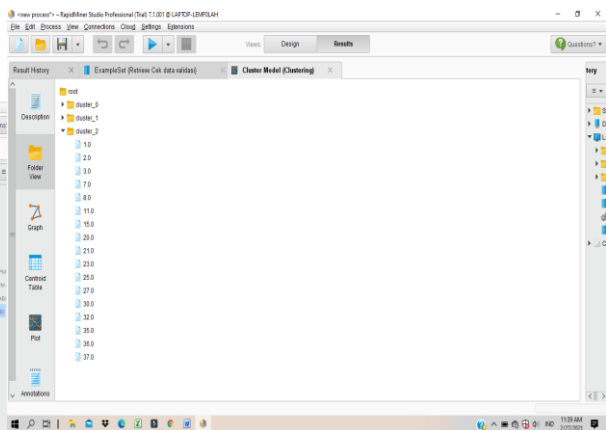
Dari hasil perhitungan manual dan pengujian dengan menggunakan software rapidminer maka diperoleh hasil kelompok yang terdiri 3 Cluster yang jumlah anggotanya sama dengan hasil perhitungan manual. Berikut anggota pada cluster 1 hasil dari pengujian.



Gambar 4. Hasil Cluster 0 Pada RapidMiner



Gambar 5. Hasil Cluster 1 Pada Rapid Miner



Gambar 6. Hasil Cluster 2 Pada Rapid Miner

Dengan performace vektornya adalah *Performance Vector*:

Avg. within centroid distance: -17.690

Avg. within centroid distance_cluster_0: -8.936

Avg. within centroid distance_cluster_1: -32.000

Avg. within centroid distance_cluster_2: -12.739

Davies Bouldin: -0.842

Berdasarkan hasil performance vektor diatas untuk mengukur tingkat keaktifan cluster menggunakan *Operator Cluster Distance Performance* mengambil model cluster centroid dan mengatur cluster sebagai input dan mengukur tingkat keaktifan berdasarkan centroid cluster. Adapun jarak yang terdekat dengan *Avg. within centroid distance* merupakan nilai yang sangat aktif dan itu terdapat pada cluster 2, serta yang jarak yang terjauh dengan *Avg. within centroid distance* bernilai kurang aktif dan itu terdapat pada cluster 1. Dan untuk cluster 0 bernilai aktif. Semakin kecil nilai davies bouldin index yang diperoleh (non-negatif ≥ 0), maka semakin baik cluster yang diperoleh dari pengelompokan menggunakan metode clustering. Hasil perhitungan menggunakan metode k-means menunjukan nilai -0.842. Angka tersebut memiliki arti masing-masing objek dalam cluster tersebut memiliki kesamaan yang cukup baik karena mendekati angka 0.

3. Kesimpulan

Dari hasil dan pembahasan dapat disimpulkan bahwa data dapat dikelompokkan menjadi 3 cluster yaitu, cluster 1 yang bernilai kurang aktif sebanyak 12 Siswa cluster 0 yang bernilai Aktif sebanyak 10 siswa, dan cluster 2 yang bernilai sangat aktif sebanyak 17 siswa. dari kesimpulan pengelompokkan tersebut dapat digunakan sebagai acuan dalam proses belajar mengajar bagi staff pengajar maupun bagi pihak sekolah yang mana bagi siswa yang kurang aktif agar dapat penguatan dalam hal belajar baik itu penguatan afektif, kognitif maupun psikomotorik di sekolah dan untuk siswa yang sangat aktif bisa untuk direkomendasikan dalam kegiatan-kegiatan perlombaan sekolah yang mana mampu untuk menunjang kemampuan siswanya.

Daftar Rujukan

- [1] Aranski, A. W., & Handoko, K. (2019). (2019). Data Mining dalam Pengelompokan Nilai IQ Siswa. *Jurnal Teknologi dan Open Source*, 2(2), 13-22. DOI: <https://ejournal.uniks.ac.id/index.php/JTOS/article/view/347> .
- [2] Wibowo, N. (2016). Upaya Peningkatan Keaktifan Siswa Melalui Pembelajaran Berdasarkan Gaya Belajar di SMK Negeri 1 Saptosari. *Jurnal Electronics, Informatics, and Vocational Education (ELINVO)*, 1(2), 128-139. DOI: <http://journal.uny.ac.id/index.php/elinvo/article/viewFile/10621/8996> .
- [3] Vhallah, I., Sumijan, S., & Santony, J. (2018). Pengelompokan Mahasiswa Potensial Drop Out Menggunakan Metode Clustering K-Means. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 2(2), 572-577. DOI: <https://jurnal.iaii.or.id/index.php/RESTI/article/view/308> .
- [4] Asroni, A., Fitri, H., & Prasetyo, E. (2018). Penerapan Metode Clustering dengan Algoritma K-Means pada Pengelompokan Data Calon Mahasiswa Baru di Universitas Muhammadiyah Yogyakarta (Studi Kasus: Fakultas Kedokteran dan Ilmu Kesehatan, dan Fakultas Ilmu Sosial dan Ilmu Politik). *Semesta Teknika*, 21(1), 60-64. DOI: <https://doi.org/10.18196/st.211211> .
- [5] Gustientiedina, G., Adiya, M. H., & Desnelita, Y. (2019). Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan. *Jurnal Nasional Teknologi dan Sistem Informasi*, 5(1), 17 – 24. DOI: <https://doi.org/10.25077/TEKNOSI.v5i1.2019.17-24> .
- [6] Dewi, S., Defit, S., & Yunus, Y. (2021). Akurasi Pemetaan Kelompok Belajar Siswa Menuju Prestasi Menggunakan Metode K-Means (Studi Kasus SMP Pembangunan Laboratorium UNP). *Jurnal Sistem Informasi dan Teknologi*, 3(1), 28-33. DOI: <https://doi.org/10.37034/jsisfotek.v3i1.98> .
- [7] Nainggolan, R., & Purba, E. (2020). Cluster Analisis of Online Shop Product Reviews Using K-Means Clustering. *International Journal of Entrepreneurship and Business Development*, 3(2), 142-15. DOI: <https://doi.org/10.29138/ijebd.v3i02.977> .
- [8] Binus Nusantara School of Computer science. (2017). Clustering [Online] Available at : <https://socs.binus.ac.id/2017/03/09/clustering/> [Accessed 27 Februari 2021]
- [9] Bastian, A., Sujadi, H., & Febrianto, G. (2018). Penerapan Algoritma K-Means Clustering Analisis Pada Penyakit Menular Manusia (Studi Kasus Kabupaten Majalengka). *Jurnal Sistem Informasi* 14(1), 26-32.
- [10] Oktarian, S., Defit, S., & Sumijan. (2020). Klasterisasi Penentuan Minat Siswa dalam Pemilihan Sekolah Menggunakan Metode Algoritma K-Means Clustering. *Jurnal Informasi dan Teknologi*, 2(3), 68-75. DOI: <https://doi.org/10.37034/jidt.v2i3.65> .
- [11] Triyansyah, D., & Fitriana, D. (2018). Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing. *Jurnal Telekomunikasi dan Komputer*, 8(3), 163-182. DOI: <http://dx.doi.org/10.22441/incomtech.v8i3.4174> .
- [12] Parlambang, B., & Fauziah. (2020). Implementasi Algoritma K-Means dalam Proses Penilaian Kuesioner Kepada Dosen Guna Mendukung Kepuasan Mahasiswa Terhadap Dosen. *Jurnal Ilmiah Teknologi dan Rekayasa*, 25(2), 161-173. DOI: <http://dx.doi.org/10.35760/tr.2020.v25i2.2719> .
- [13] Informatikalogi. (2016). Algoritma K-Means Clustering [Online] Available at: <https://informatikalogi.com/algoritma-k-means-clustering/> [Accessed 27 Februari 2021]
- [14] Yudiarta, N. G., Sudarma, M., & Ariastina, W. G. (2018). Penerapan Metode Clustering Text Mining Untuk Pengelompokan Berita Pada Unstructured Textual Data. *Majalah Ilmiah Teknologi Elektro*, 17(3), 339-344. DOI: <https://doi.org/10.24843/MITTE.2018.v17i03.P06> .
- [15] Armianti, R. I., Gaffar, A. F. O., & Putra, A. B. W. (2020). Penerapan K-Means Clustering Untuk Seleksi Frame Dominan Berbasis Ntsc Pada Obyek Bergerak. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 7(4), 745-754. DOI: <http://dx.doi.org/10.25126/jtiik.2020742184> .
- [16] Oktarina, C., Notodiputro, K. A., & Indahwati, I. (2020). Comparison of K-Means Clustering Method and K-Medoids On Twitter Data. *Indonesian Journal of Statistics and Its Applications*, 4(1), 189-202. DOI: <https://doi.org/10.29244/ijisa.v4i1.599> .
- [17] Mahmudan, A. (2020). Clustering of District or City in Central Java Based COVID-19 Case Using K-Means Clustering. *Jurnal Matematika, Statistika & Komputasi*, 17(1), 1-13. DOI: <https://doi.org/10.20956/jmsk.v17i1.10727> .
- [18] Irawan, Y. (2019). Implementation of Data Mining For Determining Majors Using K-Means Algorithm In Students of SMA Negeri 1 Pangkalan Kerinci. *Journal of Applied Engineering and Technological Science*, 1(1), 17-29. DOI: <https://doi.org/10.37385/jaets.v1i1.18> .
- [19] Nurdian, O., & Pratama, F. A. (2019). Implementasi Algoritma K-Means dalam Penentuan Prioritas Rehabilitasi Daerah Aliran Sungai Cipunagara. *Jurnal Nasional Informatika dan Teknologi Jaringan* 4(1), 76-81. DOI: <https://doi.org/10.30743/infotekjar.v4i1.1633> .
- [20] Adhitama, R., Burhanuddin, A., & Ananda, R. (2020). Penentuan Jumlah Cluster Ideal SMK di Jawa Tengah dengan Metode X-Means Clustering dan K-Means Clustering. *JIKO (Jurnal Informatika dan Komputer)*, 3(1), 1-5. DOI: <http://dx.doi.org/10.33387/jiko.v3i1.1635> .