

SENASTIKA Universitas Malikussaleh

Identifikasi Siswa Kelas Unggulan SMK Padang Panjang Dengan Menggunakan Algoritma k-Means Clustering

M Wanda*¹, Azna Alaiya², Azizi Jabbar³, Alayda Zaielamy Putri⁴
Munirul Ula⁵

^{1,2,3,4} Teknik Informatika, Universitas Malikussaleh, Lhokseumawe, Aceh
Email: ¹wanda.210170041@mhs.unimal.ac.id,
²azna.210170050@mhs.unimal.ac.id, ³azizi.210170070@mhs.unimal.ac.id,
⁴alayda.210170047@mhs.unimal.ac.id, ⁵munirulula@unimal.ac.id

Abstrak

Penelitian ini bertujuan untuk menerapkan algoritma K-Means dalam pengelompokan siswa berdasarkan nilai semester 1 dan semester 2 untuk menentukan kelas unggulan di SMK Padang Panjang. Data yang digunakan terdiri dari 388 siswa, di mana proses clustering dilakukan dengan menghitung jarak antara data dan centroid menggunakan rumus Euclidean. Algoritma K-Means dipilih karena kemampuannya yang sederhana dan efektif untuk mengelola data dalam jumlah besar. Setelah melalui empat iterasi, hasil clustering mencapai kestabilan pada iterasi ke-3 dan ke-4, dengan pembagian cluster dan nilai centroid yang tidak mengalami perubahan. Hal ini menunjukkan bahwa algoritma K-Means berhasil mengelompokkan siswa secara optimal berdasarkan kriteria yang telah ditentukan, sehingga dapat mendukung pengambilan keputusan terkait penentuan kelas unggulan.

Keywords: *K-Means, clustering, data mining, pengelompokan siswa*

1. PENDAHULUAN

Metode pembelajaran yang bertujuan untuk meningkatkan pengetahuan, sikap, dan keterampilan seseorang dikenal sebagai pendidikan. Sekolah telah berubah menjadi sarana berlangsungnya pendidikan secara langsung, di mana guru dan siswa bertemu[1]. Kualitas siswa dalam belajar di sekolah sangat beragam, yang menyebabkan tingkat pemahaman siswa yang berbeda. Sekolah harus memperhatikan hal ini, terutama guru sebagai pendidik dan pendidik siswa. Salah satunya dengan membuat semua siswa memiliki kesempatan untuk berprestasi dalam kelompok belajar yang efektif. Dalam belajar kelompok, lebih dari satu orang saling melengkapi dan berbagi ide[2]. Dimungkinkan untuk mengurangi tingkat kesulitan yang dihadapi siswa saat belajar di sekolah dengan menerapkan metode belajar kelompok[3].

Tingginya prestasi siswa dan rendahnya jumlah siswa yang tidak berprestasi mencerminkan keunggulan sektor Pendidikan[4]. Saat ini, sektor pendidikan diharapkan mampu bersaing dengan memanfaatkan kemajuan teknologi informasi (TI), yang dapat meningkatkan daya saing serta mendukung operasional sehari-hari dan pengambilan keputusan strategis[5]. Secara umum, keberhasilan siswa dievaluasi berdasarkan penilaian terhadap pelajaran teoretis dan praktis, serta kehadiran dan ketidakhadiran yang mempengaruhi nilai mereka[6].

Jumlah data siswa terus meningkat setiap tahunnya[7]. Hal ini menyebabkan penumpukan data yang belum diolah dengan baik, yang seharusnya dapat digunakan untuk mengungkap pengetahuan dan informasi baru berdasarkan pola yang terbentuk dari akumulasi data tersebut[8]. Peningkatan volume pemrosesan data ini membutuhkan penerapan strategi dan metode agar data tersebut dapat diubah menjadi informasi dan pengetahuan yang berguna bagi pendidik dalam pembuatan kebijakan[9]. Hal ini menjadi tantangan bagi guru dan wali kelas, yaitu bagaimana cara menentukan siswa kelas yang unggul.

Clustering adalah teknik Data Mining. Salah satu metode analisis data yang dapat digunakan untuk memecahkan masalah pengelompokan data adalah metode K-means, yang dapat mengelola sejumlah besar data dalam waktu yang relatif cepat dan efisien[10].

Tujuan dari penelitian ini yaitu untuk mengelompokkan siswa berdasarkan kriteria performa akademik, Menggunakan algoritma k-Means untuk mengelompokkan siswa SMK berdasarkan nilai akademik mereka, sehingga dapat diidentifikasi kelompok siswa yang berpotensi untuk ditempatkan di kelas unggulan.

2. METODE PENELITIAN

Peneliti melakukan berbagai proses, dan setiap proses dapat membantu penelitian ini. Dalam penelitian, tahapan penelitian menjelaskan proses yang diambil mulai dari menemukan masalah hingga menemukan solusinya. Tahap penelitiannya sebagai berikut :

1. Metode Pengumpulan data

Dilakukan dengan cara mengambil data melalui kaggle, Dengan jumlah keseluruhan data 388 data siswa, nilai – nilai yang di ambil dari siswa tersebut adalah nilai semester 1 dan semester 2. Link kaggle: <https://www.kaggle.com/datasets/laychristian/data-siswa-smk>.

2. Identifikasi Masalah

Tahap mengidentifikasi masalah yang terjadi, yaitu kurangnya informasi yang tepat dan rinci tentang menggabungkan data nilai siswa berdasarkan hasil penilaian proses pembelajaran siswa yang dilakukan setiap semester.

3. Pendekatan

Metode clustering mengumpulkan data nilai siswa berdasarkan hasil penilaian proses pembelajaran untuk menentukan kelas unggulan siswa berdasarkan nilai semester, yang berdampak pada data evaluasi siswa.

4. Pengolahan Data

Pengolahan data terdiri dari menyiapkan data nilai siswa berdasarkan nilai rata-rata dari nilai semester 1 dan 2 yang akan digunakan, menentukan jumlah cluster yang akan digunakan dan menghitung masing-masing kriteria cluster berdasarkan nilai rata-rata semester 1 dan 2.

5. Clustering

Pengelompokan objek berdasarkan hubungan dan kemiripan dalam kelas kelas yang sama dikenal sebagai clustering. Mengelompokkan keunggulan siswa berdasarkan nilai rata-rata semester 1 dan 2, maka implementasi algoritma k-Means Clustering akan difokuskan pada variabel nilai akademik siswa dari kedua semester.

6. K- Means

Algoritma K-means sederhana dan efektif untuk kumpulan data yang besar jika dibandingkan dengan pengelompokan hierarkis, algoritma ini digunakan untuk inisialisasi parameter. Berikut ini adalah Langkah-langkah algoritma K-Means, sebagai berikut :

- 1) Tentukan jumlah cluster yang mau dibentuk.
- 2) Tentukan pusat cluster sebanyak jumlah cluster yang mau dibentuk secara acak.
- 3) Hitung jarak masing-masing data terhadap masing-masing pusat cluster.

Rumus untuk menghitung jarak menggunakan rumus Euclidean Distance

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

- 4) Tentukan cluster yang terbentuk, lalu tentukan pusat cluster baru, dan hitung Kembali jaraknya.
- 5) Mengulangi langkah sebelumnya sampai hasil pengelompokkan tidak mengalami perubahan [10].

3. HASIL DAN PEMBAHASAN

Untuk menentukan siswa kelas unggulan dengan mengambil 388 data siswa sesuai dengan kriteria perhitungan. Kriteria yang digunakan untuk melakukan perhitungan K-Means ini ditentukan sesuai dengan kebutuhan yaitu kriteria yang diambil dari nilai semester 1 dan semester 2. Adapun daftar kriteria semester 1 dan 2 yang digunakan dapat dilihat dari tabel 1 berikut :

Tabel 1. Kriteria yang digunakan

No	Semester	Inisialisasi
1.	Semester 1	S1
2.	Semester 2	S2

Data awal sebelum dilakukan perhitungan untuk menentukan siswa kelas unggulan dengan penilaian semester 1 dan 2 yang menjadi kriteria, data -data tersebut dapat dilihat pada tabel 2 berikut :

Tabel 2. Sampel data nilai siswa

Nama Siswa	S1	S2
Siswa 1	64,6	63,6
Siswa 2	65,8	65
Siswa 3	62,4	72,2
Siswa 4	72,4	72,8
Siswa 5	64,6	65,2
Siswa 6	65	64,2
Siswa 7	62,6	64,4
Siswa 8	65,6	66,8
Siswa 9	64,6	66,2
Siswa 10	78,6	81,1
...
Siswa 379	70,4	72,6

Siswa 380	67,4	73,6
Siswa 381	75,4	76,8
Siswa 382	57,2	68,2
Siswa 383	73,4	74
Siswa 384	72,2	74,4
Siswa 385	71,4	77
Siswa 386	67,4	74,8
Siswa 387	72,2	74,2
Siswa 388	74	74,8

Selanjutnya menentukan jumlah cluster, jumlah cluster yang mau dibentuk yaitu 2 cluster, dimana pusat clusternya (Centroid) yaitu pada siswa 1 (C1) dan siswa 2 (C2). Pusat awal cluster terdapat pada tabel 3 berikut:

Tabel 3. Pusat awal cluster

Centroid	S1	S2
C1	64,6	63,6
C2	65,8	65

Hitung jarak setiap data yang ada terhadap setiap pusat cluster. Misalkan untuk perhitungan jarak data siswa pertama dengan pusat cluster pertama dengan rumus euclidian distance.

Perhitungan jarak data pertama dengan pusat cluster pertama adalah :

$$d_{(Siswa1. C1)} = \sqrt{(64,6 - 64,6)^2 + (63,6 - 63,6)^2} = \sqrt{0 + 0} = 0$$

Perhitungan jarak data pertama dengan pusat cluster kedua adalah :

$$d_{(Siswa1. C2)} = \sqrt{(64,6 - 65,8)^2 + (63,6 - 65,0)^2} = 3,4 = 1,84 = \sqrt{3,4} = 1,8439$$

Berikut adalah hasil perhitungan lengkap dari perhitungan clustering:

Tabel 4. Hasil Perhitungan Iterasi Ke-1

Nama Siswa	S1	S2	C1	C2	Jarak Terdekat
Siswa 1	64,6	63,6	0,000	1,844	C1
Siswa 2	65,8	65	1,844	0,000	C2
Siswa 3	62,4	72,2	8,877	7,962	C2
Siswa 4	72,4	72,8	12,062	10,218	C2
Siswa 5	64,6	65,2	1,600	1,217	C2
Siswa 6	65	64,2	0,721	1,131	C1
Siswa 7	62,6	64,4	2,154	3,256	C1
Siswa 8	65,6	66,8	3,353	1,811	C2
Siswa 9	64,6	66,2	2,600	1,697	C2
Siswa 10	78,6	81,1	22,411	20,568	C2
...
Siswa 379	70,4	72,6	10,707	8,884	C2
Siswa 380	67,4	73,6	10,385	8,748	C2
Siswa 381	75,4	76,8	17,055	15,212	C2
Siswa 382	57,2	68,2	8,713	9,176	C1
Siswa 383	73,4	74	13,624	11,780	C2
Siswa 384	72,2	74,4	13,206	11,372	C2
Siswa 385	71,4	77	15,027	13,242	C2
Siswa 386	67,4	74,8	11,545	9,930	C2
Siswa 387	72,2	74,2	13,043	11,207	C2
Siswa 388	74	74,8	14,622	12,778	C2

Berasarkan tabel 4 di atas, data kemudian dikelompokkan menjadi dua cluster dengan rincian sebagai berikut.

Tabel 5. Pengelompokkan Iterasi Ke-1

Cluster	Pengelompokkan	Jumlah Data
C1	1,6,7,25,40,41,49,72,77,97,100,104,116,117,121,124,146,149,172,212,217,225,258,260,261,288,293,338,371,382	30
C2	2,3,4,5,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,26,27,28,29,30,31,32,33,34,35,36,37,38,39,42,43,44,45,46,47,48,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,73,74,75,76,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,98,99,101,102,103,105,106,107,108,109,110,111,112,113,114,	358

4 SENASTIKA 2024, Jurusan Informatika Universitas Malikussaleh

115,118,119,20,122,123,125,126,127,128,129,130,131,132,133,134,135,136,137,138,139,140,141,142,143,144,145,147,148,150,151,152,153,154,155,156,157,158,159,160,161,162,163,164,165,166,167,168,169,170,171,173,174,175,176,177,178,179,180,181,182,183,184,185,186,187,188,189,190,191,192,193,194,195,196,197,198,199,200,201,202,203,204,205,206,207,208,209,210,211,213,214,215,216,218,219,220,221,222,223,224,226,227,228,229,230,231,232,233,234,235,236,237,238,239,240,241,242,243,244,245,246,247,248,249,250,251,252,253,254,256,257,259,262,263,264,265,266,267,268,269,270,271,272,273,274,275,276,277,178,279,280,281,282,283,284,285,286,287,289,290,291,292,294,295,296,297,298,299,300,301,302,303,304,305,306,307,308,309,310,311,312,313,314,315,316,317,318,319,320,321,322,323,324,325,326,327,328,329,330,331,332,333,334,335,336,337,339,340,341,342,343,344,345,346,347,348,349,350,351,352,353,354,355,356,357,358,359,360,361,362,363,364,365,366,367,368,369,370,372,373,374,375,376,377,378,379,380,381,383,384,385,386,387,388
--

Berikut ini merupakan hasil rata-rata yang diperoleh pada iterasi ke-1 untuk setiap cluster.

Tabel 6. Rata-Rata Iterasi Ke-1

CLUSTER	S1	S2
C1	57,573	66,660
C2	69,973	73,185

Melanjutkan proses dengan menghitung kembali jarak data ke centroid menggunakan rumus Euclidean untuk memastikan pembaruan kelompok secara akurat (tidak ada perubahan).

Tabel 7. Hasil Perhitungan Iterasi Ke-2

Nama Siswa	S1	S2	C1	C2	Jarak Terdekat
Siswa 1	64,6	63,6	5,646	10,988	C1
Siswa 2	65,8	65	6,172	9,187	C1
Siswa 3	62,4	72,2	6,097	7,637	C1
Siswa 4	72,4	72,8	13,967	2,457	C2
Siswa 5	64,6	65,2	4,965	9,624	C1
Siswa 6	65	64,2	5,703	10,269	C1
Siswa 7	62,6	64,4	3,556	11,469	C1
Siswa 8	65,6	66,8	5,747	7,739	C2
Siswa 9	64,6	66,2	4,767	8,812	C1
Siswa 10	78,6	81,1	23,635	11,678	C2
...
Siswa 379	70,4	72,6	12,103	0,724	C2
Siswa 380	67,4	73,6	10,251	2,606	C2
Siswa 381	75,4	76,8	18,560	6,521	C2
Siswa 382	57,2	68,2	3,069	13,711	C1
Siswa 383	73,4	74	15,406	3,523	C2
Siswa 384	72,2	74,4	14,571	2,537	C2
Siswa 385	71,4	77	15,498	4,073	C2
Siswa 386	67,4	74,8	11,099	3,038	C2
Siswa 387	72,2	74,2	14,465	2,447	C2
Siswa 388	74	74,8	16,320	4,339	C2

Mengacu pada tabel 7 di atas, data selanjutnya dibagi ke dalam dua cluster dengan detail sebagai berikut.

Tabel 8. Pengelompokkan Iterasi Ke-2

Cluster	Pengelompokkan	Jumlah Data
C1	1,2,3,5,6,7,9,25,29,38,40,41,47,49,55,56,72,77,80,91,95,97,98,100,102,104,114,116,117,121,124,134,143,146,149,172,177,183,197,201,208,209,212,217,225,243,253,258,259,260,261,264,267,269,270,287,288,293,301,308,338,356,370,371,376,382	66
C2	4,8,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,26,27,28,30,31,32,33,34,35,36,37,39,42,43,44,45,46,48,50,51,52,53,54,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,73,74,75,76,78,79,81,82,83,84,85,86,87,88,89,90,92,93,94,96,98,99,101,103,105,106,107,108,109,110,111,112,113,115,118,119,120,122,123,125,126,127,128,129,130,131,132,133,135,136,137,138,139,140,141,142,144,145,147,148,150,151,152,153,154,155,156,157,158,159,160,161,162,163,164,165,166,167,168,169,170,171,173,174,175,176,178,179,180,181,182,184,185,186,187,188,189,190,191,192,193,194,195,196,198,199,200,202,203,204,205,206,207,210,211,213,214,215,216,218,219,220,221,222,223,224,226,227,228,229,230,231,232,233,234,235,236,237,238,239,240,241,242,244,245,246,247,248,249,250,251,252,254,255,256,257,262,263,265,266,268,269,271,272,273,274,275,276,277,278,279,280,281,282,283,284,285,286,289,290,291,292,294,295,296,297,298,299,300,302,303,304,305,306,307,309,310,311,312,313,314,315,316,317,318,319,320,321,322,323,324,325,326,327,328,329,330,331,332,333,334,335,336,337,339,340,341,342,343,344,345,346,347,348,3	322

49,350,351,352,353,354,355,357,358,359,360,361,362,363,364,365,366,367,368,369,,372,373,374,375,377,378,379,380,381,383,384,385,386,387,388

Hasil rata-rata untuk setiap cluster pada iterasi ke-2 disajikan sebagai berikut.

Tabel 9. Rata-Rata Iterasi Ke-2

CLUSTER	S1	S2
C1	59,855	67,809
C2	63,763	66,270

Langkah berikutnya adalah menghitung ulang jarak antara data dan centroid menggunakan rumus Euclidean, guna memastikan pembaruan cluster dilakukan dengan tepat (tidak ada perubahan).

Tabel 10. Hasil Perhitungan Iterasi Ke-3

Nama Siswa	S1	S2	C1	C2	Jarak Terdekat
Siswa 1	64,6	63,6	6,343	2,798	C2
Siswa 2	65,8	65	6,575	2,400	C2
Siswa 3	62,4	72,2	5,075	6,085	C1
Siswa 4	72,4	72,8	13,501	10,828	C2
Siswa 5	64,6	65,2	5,415	1,358	C2
Siswa 6	65	64,2	6,285	2,411	C2
Siswa 7	62,6	64,4	4,377	2,202	C2
Siswa 8	65,6	66,8	5,833	1,912	C2
Siswa 9	64,6	66,2	5,010	0,840	C2
Siswa 10	78,6	81,1	22,953	20,947	C2
...
Siswa 379	70,4	72,6	11,582	9,172	C2
Siswa 380	67,4	73,6	9,511	8,183	C2
Siswa 381	75,4	76,8	17,958	15,694	C2
Siswa 382	57,2	68,2	2,684	6,841	C1
Siswa 383	73,4	74	14,893	12,354	C2
Siswa 384	72,2	74,4	13,994	11,717	C2
Siswa 385	71,4	77	14,757	13,170	C2
Siswa 386	67,4	74,8	10,286	9,273	C2
Siswa 387	72,2	74,2	13,901	11,579	C2
Siswa 388	74	74,8	15,778	13,325	C2

Mengacu pada tabel 10 di atas, data selanjutnya dibagi ke dalam dua cluster dengan detail sebagai berikut.

Tabel 11. Pengelompokan Iterasi Ke-3

Cluster	Pengelompokan	Jumlah Data
C1	3,47,49,72,80,97,100,102,114,116,121,134,146,149,177,208,212,217,224,243,253,256,257,258,259,260,287,293,301,308,338,371,378,382	32
C2	1,2,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,48,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,73,74,75,76,77,78,79,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,98,99,101,103,104,105,106,107,108,109,110,111,112,113,115,117,118,119,120,122,123,124,125,126,127,128,129,130,131,132,133,135,136,137,138,139,140,141,142,143,144,145,147,148,150,151,152,153,154,155,156,157,158,159,160,161,162,163,164,165,166,167,168,169,170,171,172,173,174,175,176,178,179,180,181,182,183,184,185,186,187,188,189,190,191,192,193,194,195,196,197,198,199,200,201,202,203,204,205,206,207,209,210,211,213,214,215,216,218,219,220,221,222,223,225,226,227,228,229,230,231,232,233,234,235,236,237,238,239,240,241,242,244,245,246,247,248,249,250,251,252,254,255,261,262,263,264,265,266,267,268,269,270,271,272,273,274,275,276,277,278,279,280,281,282,283,284,285,286,288,289,290,291,292,294,295,296,297,298,299,300,302,303,304,305,306,307,308,309,310,311,312,313,314,315,316,317,318,319,320,321,322,323,324,325,326,327,328,329,330,331,332,333,334,335,336,337,339,340,341,342,343,344,345,346,347,348,349,350,351,352,353,354,355,356,357,358,359,360,361,362,363,364,365,366,367,368,369,370,372,373,374,375,376,377,379,380,381,383,384,385,386,387,388	356

6 SENASTIKA 2024, Jurusan Informatika Universitas Malikussaleh

Rata-rata hasil yang diperoleh dari iterasi ke-3 untuk masing-masing cluster ditampilkan di bawah ini.

Tabel 12. Rata-Rata Iterasi Ke-3

CLUSTER	S1	S2
C1	59,375	75,788
C2	69,881	72,401

Proses dilanjutkan dengan menghitung kembali jarak antara setiap data dan centroid menggunakan rumus Euclidean, memastikan pembagian cluster diperbarui secara akurat.

Tabel 13. Hasil Perhitungan Iterasi Ke-4

Nama Siswa	S1	S2	C1	C2	Jarak Terdekat
Siswa 1	64,6	63,6	13,261	10,264	C2
Siswa 2	65,8	65	12,556	8,452	C2
Siswa 3	62,4	72,2	4,693	7,484	C1
Siswa 4	72,4	72,8	13,363	2,550	C2
Siswa 5	64,6	65,2	11,807	8,930	C2
Siswa 6	65	64,2	12,881	9,544	C2
Siswa 7	62,6	64,4	11,836	10,818	C2
Siswa 8	65,6	66,8	10,933	7,050	C2
Siswa 9	64,6	66,2	10,919	8,145	C2
Siswa 10	78,6	81,1	19,934	12,285	C2
...
Siswa 379	70,4	72,6	11,477	0,556	C2
Siswa 380	67,4	73,6	8,318	2,755	C2
Siswa 381	75,4	76,8	16,057	7,057	C2
Siswa 382	57,2	68,2	7,894	13,359	C1
Siswa 383	73,4	74	14,139	3,865	C2
Siswa 384	72,2	74,4	12,900	3,061	C2
Siswa 385	71,4	77	12,086	4,843	C2
Siswa 386	67,4	74,8	8,086	3,451	C2
Siswa 387	72,2	74,2	12,923	2,935	C2
Siswa 388	74	74,8	14,658	4,766	C2

Mengacu pada tabel 13 di atas, data selanjutnya dibagi ke dalam dua cluster dengan detail sebagai berikut.

Tabel 14. Pengelompokan Iterasi Ke-4

Cluster	Pengelompokan	Jumlah Data
C1	3,47,49,72,80,97,100,102,114,116,121,134,146,149,177,208,212,217,224,243,253,256,257,258,259,260,287,293,301,308,338,371,378,382	32
C2	1,2,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,48,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,73,74,75,76,77,78,79,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,98,99,101,103,104,105,106,107,108,109,110,111,112,113,115,117,118,119,120,122,123,124,125,126,127,128,129,130,131,132,133,135,136,137,138,139,140,141,142,143,144,145,147,148,150,151,152,153,154,155,156,157,158,159,160,161,162,163,164,165,166,167,168,169,170,171,172,173,174,175,176,178,179,180,181,182,183,184,185,186,187,188,189,190,191,192,193,194,195,196,197,198,199,200,201,202,203,204,205,206,207,209,210,211,213,214,215,216,218,219,220,221,222,223,225,226,227,228,229,230,231,232,233,234,235,236,237,238,239,240,241,242,244,245,246,247,248,249,250,251,252,254,255,261,262,263,264,265,266,267,268,269,270,271,272,273,274,275,276,277,278,279,280,281,282,283,284,285,286,288,289,290,291,292,294,295,296,297,298,299,300,302,303,304,305,306,307,308,309,310,311,312,313,314,315,316,317,318,319,320,321,322,323,324,325,326,327,328,329,330,331,332,333,334,335,336,337,339,340,341,342,343,344,345,346,347,348,349,350,351,352,353,354,355,356,357,358,359,360,361,362,363,364,365,366,367,368,369,370,372,373,374,375,376,377,379,380,381,383,384,385,386,387,388	356

Rata-rata hasil yang diperoleh dari iterasi ke-4 untuk masing-masing cluster ditampilkan di bawah ini.

Tabel 15. Rata-Rata Iterasi Ke-4

CLUSTER	S1	S2
C1	59,375	75,788

C2	69,881	72,401
----	--------	--------

Setelah dilakukan sebanyak 4 iterasi, terlihat bahwa pada iterasi ke-3 dan ke-4 memiliki anggota yang sama, dan nilai centroidnya tidak mengalami perubahan. Oleh karena itu, iterasi lebih lanjut tidak diperlukan karena hasil clustering telah mencapai kestabilan dan konvergen.

Siswa kelas unggulan adalah yang masuk ke dalam Cluster C2, karena cluster ini memiliki nilai rata-rata Semester 1 dan 2 yang lebih tinggi dibandingkan Cluster C1. Jumlah siswa kelas unggulan dari total 388 siswa yang dikelompokkan, 356 siswa masuk ke dalam Cluster C2, yang dapat dikategorikan sebagai siswa kelas unggulan.

5. DISKUSI

Identifikasi Siswa Kelas Unggulan SMK Padang Panjang Dengan Menggunakan Algoritma k-Means Clustering memberikan gambaran lengkap tentang proses penerapan k-Means Clustering untuk identifikasi siswa kelas unggulan di SMK Padang Panjang dan bagaimana hasilnya bisa dimanfaatkan untuk keperluan pendidikan.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa algoritma K-Means berhasil digunakan untuk mengelompokkan siswa ke dalam dua cluster berdasarkan nilai semester 1 dan semester 2 dari 388 siswa. Proses clustering dilakukan dengan menghitung jarak antara data dan centroid menggunakan rumus Euclidean. Setelah melalui empat iterasi, hasil clustering mencapai kestabilan pada iterasi ke-3 dan ke-4, di mana pembagian cluster dan nilai centroid tidak mengalami perubahan. Oleh karena itu, iterasi lebih lanjut tidak diperlukan, menunjukkan bahwa algoritma K-Means telah berhasil mengelompokkan siswa secara optimal sesuai kriteria yang digunakan.

DAFTAR PUSTAKA

- [1] S. Dewi, S. Defit, and Y. Yuhandri, "Akurasi Pemetaan Kelompok Belajar Siswa Menuju Prestasi Menggunakan Metode K-Means," *J. Sistim Inf. dan Teknol.*, vol. 3, pp. 28–33, 2021, doi: 10.37034/jsisfotek.v3i1.40.
- [2] I. D. Setiawan and A. Triayudi, "Penerapan Data Mining Dengan Menggunakan Algoritma Clustering K-Means Untuk Pembagian Jurusan Pada Sekolah Menengah Atas," *J. Comput. Syst. Informatics*, vol. 5, no. 2, pp. 380–392, 2024, doi: 10.47065/josyc.v5i2.4970.
- [3] T. Hidayat, "Klasifikasi Data Jamaah Umroh Menggunakan Metode K-Means Clustering," *J. Sistim Inf. dan Teknol.*, vol. 4, pp. 19–24, 2022, doi: 10.37034/jsisfotek.v4i1.115.
- [4] M. S. Said and Y. Yusti, "Penerapan Algoritma K-Means Dalam Penentuan Jurusan Siswa Sman 05 Bombana," *Simtek J. Sist. Inf. dan Tek. Komput.*, vol. 5, no. 2, pp. 114–122, 2020, doi: 10.51876/simtek.v5i2.87.
- [5] I. Savitri, H. 1, S. Defit, and G. W. Nurcahyo, "Jurnal Sistim Informasi dan Teknologi <https://jsisfotek.org/index.php> Simulasi dalam Optimalisasi Pengadaan Barang menggunakan Metode K-Mean Clustering," 2021, doi: 10.37034/jsisfotek.v3i3.79.
- [6] A. Yudhistira and R. Andika, "Pengelompokan Data Nilai Siswa Menggunakan Metode K-Means Clustering," *J. Artif. Intell. Technol. Inf.*, vol. 1, no. 1, pp. 20–28, Feb. 2023, doi: 10.58602/jaiti.v1i1.22.
- [7] A. Sulistiyawati and E. Supriyanto, "Implementasi Algoritma K-means Clustring dalam Penentuan Siswa Kelas Unggulan," *J. Tekno Kompak*, vol. 15, no. 2, p. 25, 2021, doi: 10.33365/jtk.v15i2.1162.
- [8] J. Hutagalung, "Pemetaan Siswa Kelas Unggulan Menggunakan Algoritma K-Means Clustering," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 1, pp. 606–620, 2022, doi: 10.35957/jatisi.v9i1.1516.
- [9] N. R. Wahyudi, Y. Rahmawati, and S. Supriyanto, "Implementasi Rapidminer Untuk Menentukan Siswa Unggulan Menggunakan Metode K-Means," *SKANIKA Sist. Komput. dan Tek. Inform.*, vol. 7, no. 2, pp. 109–118, 2024, doi: 10.36080/skanika.v7i2.3173.
- [10] R. P. Primanda, A. Alwi, and D. Mustikasari, "DATA MINING SELEKSI SISWA BERPRESTASI UNTUK MENENTUKAN KELAS UNGGULAN MENGGUNAKAN METODE K-MEANS CLUSTERING (Studi Kasus di MTS Darul Fikri)," *Komputek*, vol. 5, no. 1, p. 88, 2021, doi: 10.24269/jkt.v5i1.686.