

Penerapan Data Mining Algoritma K-Medoids untuk Rekomendasi Kelompok Mata Pelajaran Pilihan

^{1*}Dea Avrillia Arba'a, ²Risa Helilintar, ³Intan Nur Farida

Teknik Informatika, Universitas Nusantara PGRI Kediri

E-mail: *ldcnzzcode8@gmail.com, risa.helilintar@gmail.com, in.nfarida@gmail.com

Penulis Korespondens : Dea Avrillia Arba'a

Abstrak— Pemilihan kelompok minat belajar siswa di SMA sering kali masih dilakukan secara manual dan subjektif, sehingga menyebabkan ketidaktepatan dalam penempatan siswa. Penelitian ini bertujuan untuk mengelompokkan data siswa berdasarkan minat dan nilai akademik guna membantu guru dalam memberikan arahan yang tepat. Metode yang digunakan adalah algoritma *K-Medoids* dengan perhitungan jarak *Euclidean* untuk proses *clustering*. Pengumpulan data dilakukan pada siswa SMA Pawayatan Daha Kediri dengan atribut nilai IPA, nilai IPS, dan nilai tes peminatan. Prosedur penelitian meliputi *Requirement, Design, Implementation, Testing, dan Maintenance*. Hasil penelitian menunjukkan bahwa metode *K-Medoids* efektif dalam mengelompokkan siswa ke dalam beberapa klaster, yaitu kelompok yang cenderung memilih kelompok IPA atau kelompok IPS, dengan nilai *silhouette* yang menunjukkan kualitas clustering cukup baik yaitu 0,34. Sistem yang dibangun mampu membantu guru dalam mengambil keputusan yang lebih objektif terkait pengelompokkan siswa.

Kata Kunci—*Clustering, Euclidean Distance, K-Medoids, Pengelompokan Siswa, Silhouette Coefficient*

Abstract— *The selection of students' academic interest groups in high schools is often conducted manually and subjectively, which can lead to inaccurate placement. This study aims to cluster student data based on their interests and academic scores to assist teachers in providing appropriate guidance. The method used is the K-Medoids algorithm with Euclidean distance calculation for the clustering process. Data were collected from students at SMA Pawayatan Daha Kediri using attributes such as science scores, social science scores, and interest test scores. The research procedure includes Requirement, Design, Implementation, Testing, and Maintenance stages. The results show that the K-Medoids method is effective in grouping students into several clusters, namely those inclined toward Science or Social Science groups, with a silhouette score of 0.34 indicating a fairly good clustering quality. The developed system helps teachers make more objective decisions regarding student group placement.*

Keywords— *Clustering, Euclidean Distance, K-Medoids, Student Grouping, Silhouette Coefficient*

This is an open access article under the CC BY-SA License.



I. PENDAHULUAN

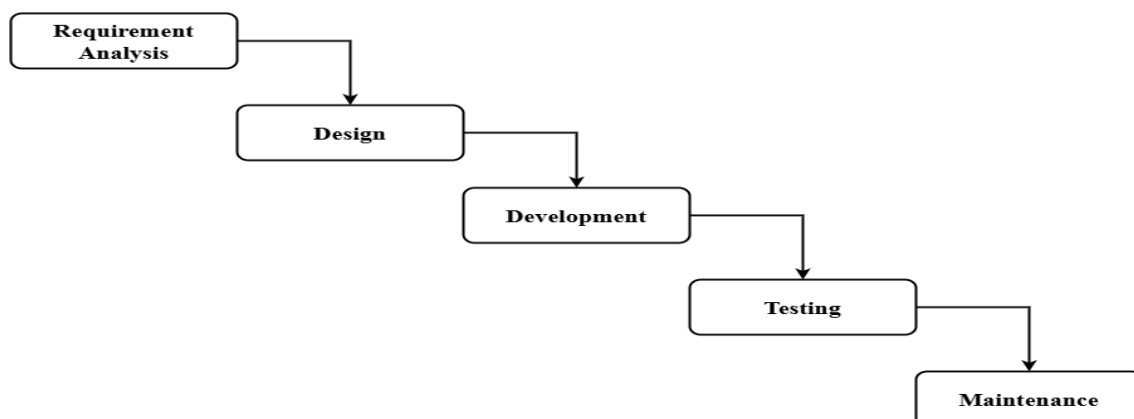
Pemilihan mata pelajaran pilihan dalam dunia pendidikan menengah kini memegang peranan penting dalam menentukan arah dan minat siswa pada jenjang pendidikan selanjutnya. SMA Pawayatan Daha Kediri sebagai institusi yang menerapkan Kurikulum Merdeka dengan KOSP (Kurikulum Operasional Satuan Pendidikan) berupaya memberikan keleluasaan kepada siswa dalam menentukan mata pelajaran yang sesuai dengan minat, bakat, dan kemampuan akademiknya. Namun, pelaksanaan pemilihan tersebut masih banyak dilakukan secara manual sehingga rentan terhadap subjektivitas dan ketidaktepatan penempatan siswa ke kelas pilihan [1].

Metode berbasis teknologi informasi telah banyak dikembangkan untuk membantu proses pengambilan keputusan, termasuk dalam pemetaan data siswa menggunakan pendekatan *data mining* [2] [3]. Salah satu metode klastering yang efektif digunakan untuk membagi kelompok siswa dengan karakteristik yang serupa adalah *K-Medoids*. Metode ini mampu meminimalisir total jarak antara anggota klaster dengan pusat *medoid*, sehingga lebih *robust* terhadap data pencilan dibanding metode lain seperti *K-Means* [4] [5]. Sistem rekomendasi berbasis klastering telah terbukti mampu meningkatkan objektivitas dalam proses seleksi atau rekomendasi, terutama pada konteks pendidikan [6]. Dengan menerapkan *K-Medoids*, siswa dapat diklasifikasikan ke dalam kelompok yang homogen, yang nantinya menjadi dasar rekomendasi mata pelajaran pilihan. Selain itu, algoritma ini mendukung validasi menggunakan *Silhouette Score* untuk mengukur kekompakan antar anggota dalam satu klaster [7] [8].

Pada penelitian sebelumnya, pendekatan berbasis klaster telah diaplikasikan dalam pembagian jurusan IPA dan IPS di tingkat sekolah menengah, dengan hasil yang memuaskan [9]. Namun, belum banyak penelitian yang mengintegrasikan proses ini ke dalam sistem *web* yang dapat langsung diakses oleh guru bimbingan konseling dan siswa secara *real-time* [10]. Oleh karena itu, penelitian ini bertujuan mengembangkan aplikasi berbasis *web* dengan algoritma *K-Medoids* untuk membantu proses rekomendasi mata pelajaran pilihan di SMA Pawyatan Daha Kediri. Penelitian ini mengintegrasikan teknologi *Python* dengan *framework Flask*, *database MySQL*, dan visualisasi hasil klaster. Diharapkan aplikasi yang dikembangkan tidak hanya mempermudah proses rekomendasi tetapi juga memberikan transparansi terhadap proses penentuan pilihan siswa, sehingga mampu meningkatkan kepuasan dan kepercayaan seluruh pihak yang terlibat [11]. Evaluasi sistem dilakukan dengan mengukur rata-rata *Silhouette Score* sebagai indikator kualitas pembagian klaster [12]. Secara keseluruhan, penerapan algoritma *K-Medoids* dalam pengelompokan data siswa diyakini dapat membantu institusi pendidikan dalam memetakan potensi siswa serta merekomendasikan mata pelajaran pilihan secara lebih akurat dan efisien [13].

II. METODE

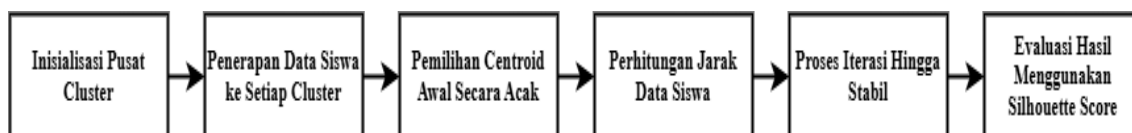
Penelitian ini menggunakan metode *Waterfall* yang merupakan salah satu pendekatan pengembangan perangkat lunak yang menggunakan alur kerja linear dan terstruktur. Setiap tahap dalam metode ini harus diselesaikan secara berurutan sebelum melanjutkan ke tahap berikutnya [14]. Metode ini terdiri dari beberapa tahapan sebagai berikut:



Gambar 1. Prosedur Penelitian

Dalam penelitian ini, penulis melakukan identifikasi kebutuhan sistem melalui studi literatur dan analisis data. Data yang dikumpulkan meliputi nilai akademik dan minat siswa, dalam proses pengelompokkan mata pelajaran pilihan. Penulis menyusun spesifikasi kebutuhan sistem baik dari sisi fungsional maupun non-fungsional. Setelah kebutuhan sistem ditentukan, tahap selanjutnya adalah merancang alur proses sistem. Penulis membuat diagram *use case*, diagram *activity*, dan diagram *sequence*. Proses perancangan juga mencakup metode *clustering* yaitu *K-Medoids* dan perhitungan jarak *Euclidean* sebagai dasar pengelompokkan.

Tahapan *Implementation*, tahap ini merupakan proses pembangunan sistem berbasis *web* menggunakan bahasa pemrograman *Python* dan *framework Flask*. Fungsi-fungsi utama yang dikembangkan mencakup *input* data siswa, proses *clustering K-Medoids*, dan penampilan hasil *Silhouette Score*. Data disimpan dalam *database MySQL*. Sistem diuji fungsionalitasnya serta keakuratan proses *clustering*. Pengujian dilakukan dengan menggunakan nilai *Silhouette Score* untuk mengevaluasi hasil klasterisasi. Nilai tertinggi diperoleh pada saat jumlah klaster dengan *Silhouette Score* yang menunjukkan kualitas pengelompokkan cukup baik. Setelah sistem diuji dan dinyatakan berjalan sesuai kebutuhan, sistem dapat digunakan oleh guru dalam proses pengelompokkan siswa. Tahap ini juga mencakup dokumentasi sistem dan perencanaan pemeliharaan untuk pengembangan lebih lanjut di masa depan.



Gambar 2. Tahapan Pengelompokkan

Proses pengelompokkan siswa dilakukan menggunakan metode *clustering* dengan pendekatan *Silhouette Score* untuk menilai kualitas hasil klasterisasi. Proses dimulai dengan inisialisasi pusat klaster (*centroid*) yang mewakili kelompok masing-masing kelas XI. Setiap klaster memiliki karakteristik masing-masing, seperti kelas XI-1 dan 2 yang berfokus pada bidang IPA, dan XI-3 dan 4 pada bidang IPS. Data siswa yang terdiri dari nilai IPS, nilai IPA, dan nilai tes peminatan dimasukkan ke dalam sistem. Sistem melakukan pemilihan *centroid* awal secara acak dari sekumpulan data siswa yang ada, kemudian digunakan sebagai acuan dalam proses pengelompokkan.

Proses berikutnya adalah perhitungan jarak antara data siswa dengan setiap pusat klaster menggunakan metode *Euclidean Distance*. Penghitungan ini bertujuan untuk menentukan sejauh mana kemiripan setiap siswa terhadap masing-masing klaster. Berdasarkan hasil jarak tersebut, setiap siswa ditempatkan ke dalam klaster yang paling dekat atau paling sesuai. Proses ini dilakukan secara iteratif, di mana pusat klaster akan diperbarui hingga tidak terjadi lagi perubahan signifikan dalam pengelompokkan siswa. Iterasi akan berhenti apabila posisi *centroid* sudah stabil dan seluruh data siswa telah tergolong dalam kelompoknya masing-masing. Sebagai tahap akhir, dilakukan evaluasi hasil *clustering* menggunakan *Silhouette Score*. Nilai *Silhouette Score* memberikan gambaran sejauh mana siswa dalam satu klaster memiliki kedekatan yang kuat satu sama lain, dan sejauh mana mereka berbeda dari siswa di klaster lainnya. Semakin tinggi nilai *silhouette*, semakin baik kualitas pengelompokkan yang dihasilkan. Dengan demikian, metode ini memastikan bahwa pembagian kelas benar-benar optimal dan mencerminkan kemampuan, minat, dan kecenderungan siswa secara objektif.

III. HASIL DAN PEMBAHASAN

Dalam penelitian ini, penulis mengambil data melalui beberapa metode. Pertama, wawancara dilakukan dengan guru Bimbingan Konseling (BK), lalu dengan Wakil Kepala Sekolah (Wakasek), setelah itu dengan guru Kurikulum di SMA Pawyatan Daha Kediri untuk memahami kebutuhan, kriteria pemilihan mata pelajaran pilihan, dan kendala yang dihadapi dalam sistem manual. Kedua, observasi langsung dilakukan terhadap proses pemilihan mata pelajaran pilihan siswa yang sedang berlangsung untuk mendokumentasikan alur kerja sistem manual dan mengidentifikasi permasalahan yang terjadi. Ketiga, data siswa diperoleh melalui dokumentasi dari pihak sekolah, yang mencakup nilai IPA, nilai IPS, dan tes minat, siswa yang dinyatakan lulus pada kelas X. Dapat dilihat pada tabel sebagai berikut:

Tabel 1. Data Siswa Kelas X

No.	Tabel Data Siswa Kelas X				
	Nama	Kelas Asal	IPS	IPA	Nilai Tes Minat
1	Brian	X1	35	37,5	52
2	Defina	X1	27,5	27,5	50
3	Jiyon	X1	30	35	59
4	Erlan	X2	37,5	37,5	84
5	Khoirunnisa	X2	52,5	42,5	37
6	Ahmad	X3	30	42,5	67
7	Ervina	X3	20	50	58
8	Fransiska	X3	37,5	55	79
9	Davino	X2	42,5	55	68
10	Dheanita	X2	32,5	62,5	50

Tabel 1 menyajikan informasi mengenai 10 siswa beserta data mereka yaitu, nilai akademik dalam dua kategori yaitu IPS, dan IPA, serta nilai tes peminatan. Data ini kemudian diolah dan dianalisis menggunakan metode *clustering* untuk mengelompokkan siswa berdasarkan kemiripan karakteristik nilai mereka dengan tahapan sebagai berikut:

A. Menentukan Nilai *Centroid* Awal

Sebelum dilakukan proses iterasi dalam metode *clustering*, terlebih dahulu diketahui nilai *centroid* awal sebagai titik pusat dari masing-masing klaster yang akan digunakan untuk mengelompokkan data siswa berdasarkan nilai yang telah ditentukan, adapun nilai *centroid* awal dapat dilihat pada tabel 2 sebagai berikut:

Tabel 2. Nilai *Centroid* Awal

No.	Tabel Nilai <i>Centroid</i> Awal			
	Nama	Kelas Asal	IPS	IPA
1	Brian	X1	35	37,5
7	Ervina	X3	20	50
5	Khoirunnisa	X2	52,5	42,5
10	Dheanita	X2	32,5	62,5

Setelah diketahui nilai *centroid* yang akan digunakan pada proses perhitungan jarak *Euclidean*, langkah berikutnya adalah menghitung jarak antara setiap data dengan *centroid* yang telah ditetapkan. Hasil perhitungan jarak kemudian menjadi dasar untuk menentukan ke *cluster* mana setiap data akan dikategorikan, yaitu dengan memilih *cluster* yang memiliki jarak paling dekat.

Proses ini terus berlanjut secara iteratif dimana *centroid* akan diperbarui berdasarkan rata-rata posisi data dalam masing-masing *cluster*, hingga diperoleh hasil *clustering* yang stabil dan optimal. Adapun tahapan perhitungan tersebut dapat dilihat sebagai berikut:

B. Perhitungan Jarak Data Siswa

Tabel 3. Nilai Jarak

Tabel Nilai Jarak			
Iterasi 1		Iterasi 2	
Dc1	Dc2	Dc1	Dc2
0	20,42	23,58	25,20
12,65	25,02	31,92	35,35
8,95	18,05	32,34	29,04
32,09	33,74	49,58	42,49
23,58	39,41	0	31,12
16,58	15,40	37,5	26,36
20,42	0	39,41	19,40
32,27	27,78	46,31	30,36
24,86	25,12	34,88	21,91
25,20	19,40	31,12	0

Tabel 3 menunjukkan perhitungan jarak antara titik data siswa Dc1 (jarak *centroid* 1) dan Dc2 (jarak *centroid* 2) terhadap dua iterasi berbeda dalam proses *clustering*. Tabel ini digunakan untuk mengilustrasikan tahap perkiraan jarak data siswa sebagai bagian dari analisis *clustering*.

C. Penentuan Cluster

Setelah dilakukannya proses perhitungan jarak, maka selanjutnya yaitu menentukan *cluster*. Adapun hasil perhitungan jarak dan penentuan *cluster* dapat dilihat tabel 4 sebagai berikut:

Tabel 4. Hasil Penentuan Cluster

Tabel Hasil Penentuan Cluster					
Iterasi 1			Iterasi 2		
Dc1	Dc2	Cluster	Dc1	Dc2	Cluster
0	20,42	1	23,58	25,20	1
12,65	25,02	1	31,92	35,35	1
8,95	18,05	1	32,34	29,04	2
32,09	33,74	1	49,58	42,49	2
23,58	39,41	1	0	31,12	1
16,58	15,40	2	37,5	26,36	2
20,42	0	2	39,41	19,40	2
32,27	27,78	2	46,31	30,36	2
24,86	25,12	1	34,88	21,91	2
25,20	19,40	2	31,12	0	2

Tabel 4 menampilkan hasil pengelompokkan data siswa ke dalam klaster berdasarkan perhitungan jarak yang telah dilakukan pada iterasi sebelumnya. Tabel ini membagi hasil menjadi dua bagian yaitu, iterasi 1 dan iterasi 2 yang menunjukkan perubahan pengelompokkan setelah pembaruan *centroid*.

D. Proses Iterasi

Setelah proses penentuan *cluster* yang diawali dengan menentukan jarak terdekat antara data dan pusat *cluster* yang tersedia. Lalu, terdapat dua nilai yang dibandingkan yaitu, $Dc1$ dan $Dc2$. Apabila nilai $Dc1$ lebih kecil dibandingkan dengan $Dc2$, maka data tersebut akan dikategorikan ke dalam *cluster* pertama yaitu, $C1$, karena jaraknya lebih dekat ke pusat *cluster* $C1$. Sebaliknya, jika $Dc2$ memiliki nilai yang lebih kecil daripada $Dc1$, maka data tersebut akan dimasukkan ke dalam *cluster* kedua yaitu, $C2$, karena menunjukkan bahwa jarak terdekatnya adalah ke pusat *cluster* $C2$. Setelah seluruh data berhasil dikelompokkan ke dalam *cluster* yang sesuai berdasarkan perbandingan jarak, tahap berikutnya adalah melakukan perhitungan terhadap nilai *cost*. Nilai *cost* yang dihitung berdasarkan total dari jarak antara data dengan pusat *cluster* tempatnya tergabung, sehingga dapat digunakan untuk mengevaluasi seberapa optimal proses pembentukan *cluster* yang telah dilakukan. Adapun perhitungan nilai *cost* dapat dilihat sebagai berikut:

Tabel 5. Hasil Nilai *Cost*

Tabel Hasil Nilai <i>Cost</i>			
Iterasi 1		Iterasi 2	
<i>Cost</i>	164,765	<i>Cost</i>	225,101

Setelah nilai *cost* berhasil dihitung, langkah selanjutnya adalah melakukan evaluasi terhadap hasil *clustering* yang telah terbentuk. Jika nilai *cost* yang diperoleh masih dapat diminimalkan, maka proses iterasi akan dilanjutkan dengan memperbarui posisi pusat *cluster* berdasarkan data yang telah dikelompokkan sebelumnya. Proses ini akan terus berulang hingga tercapai kondisi dimana nilai *cost* berada pada titik minimum atau perubahan yang terjadi sudah sangat kecil, sehingga *clustering* dianggap telah mencapai hasil yang optimal.

Perhitungan Manual K-Medoids

Iterasi 1

Medoid:

* [0.4583333333333333, 0.3181818181818182, 0.5571428571428572]
* [0.5416666666666666, 0.8181818181818182, 0.11428571428571428]

Rincian Cost per Titik:

Cluster 0

INDEKS TITIK KANDIDAT	JARAK KE TITIK LAIN (PAIR DISTANCES)	TOTAL COST
0	[0.0000, 0.2172, 0.1414, 0.4043, 0.3062, 0.3929, 0.3600, 0.4594, 0.4219, 0.4347, 0.2472, 0.3537, 0.4857, 0.4990, 0.1572, 0.1681, 0.2153, 0.5312, 0.3124, 0.7007, 0.1815, 0.2568, 0.4812, 0.3929, 0.3675, 0.5223, 0.4638, 0.4383, 0.3949, 0.4201, 0.7928, 0.5150, 0.2103, 0.4013, 0.0476, 0.5362, 0.5044, 0.4959, 0.5316, 0.5679, 0.3457, 0.4591, 0.3367, 0.2649, 0.3285, 0.3762, 0.5749, 0.3995, 0.4508, 0.4932, 0.3452, 0.3341, 0.2567, 0.5376, 0.1681, 0.4124, 0.4049, 0.4887, 0.7445, 0.1414, 0.3610, 0.6315, 0.1742, 0.4485, 0.5050, 0.5379, 0.5617, 0.2419, 0.4497, 0.5312, 0.4994, 0.4228, 0.4744, 0.4725, 0.3651, 0.6145, 0.3324, 0.3361, 0.4299, 0.5847]	32.1685
	[0.2172, 0.0000, 0.1850, 0.4243, 0.3452, 0.4485, 0.4299, 0.5448, 0.4111, 0.6391, 0.3515, 0.4151, 0.6440, 0.4949, 0.0617, 0.2172, 0.0000, 0.1850, 0.4243, 0.3452, 0.4485, 0.4299, 0.5448, 0.4111, 0.6391, 0.3515, 0.4151, 0.6440, 0.4949, 0.0617, 0.1414, 0.4043, 0.3062, 0.3929, 0.3600, 0.4594, 0.4219, 0.4347, 0.2472, 0.3537, 0.4857, 0.4990, 0.1572, 0.1681, 0.2153, 0.5312, 0.3124, 0.7007, 0.1815, 0.2568, 0.4812, 0.3929, 0.3675, 0.5223, 0.4638, 0.4383, 0.3949, 0.4201, 0.7928, 0.5150, 0.2103, 0.4013, 0.0476, 0.5362, 0.5044, 0.4959, 0.5316, 0.5679, 0.3457, 0.4591, 0.3367, 0.2649, 0.3285, 0.3762, 0.5749, 0.3995, 0.4508, 0.4932, 0.3452, 0.3341, 0.2567, 0.5376, 0.1681, 0.4124, 0.4049, 0.4887, 0.7445, 0.1414, 0.3610, 0.6315, 0.1742, 0.4485, 0.5050, 0.5379, 0.5617, 0.2419, 0.4497, 0.5312, 0.4994, 0.4228, 0.4744, 0.4725, 0.3651, 0.6145, 0.3324, 0.3361, 0.4299, 0.5847]	

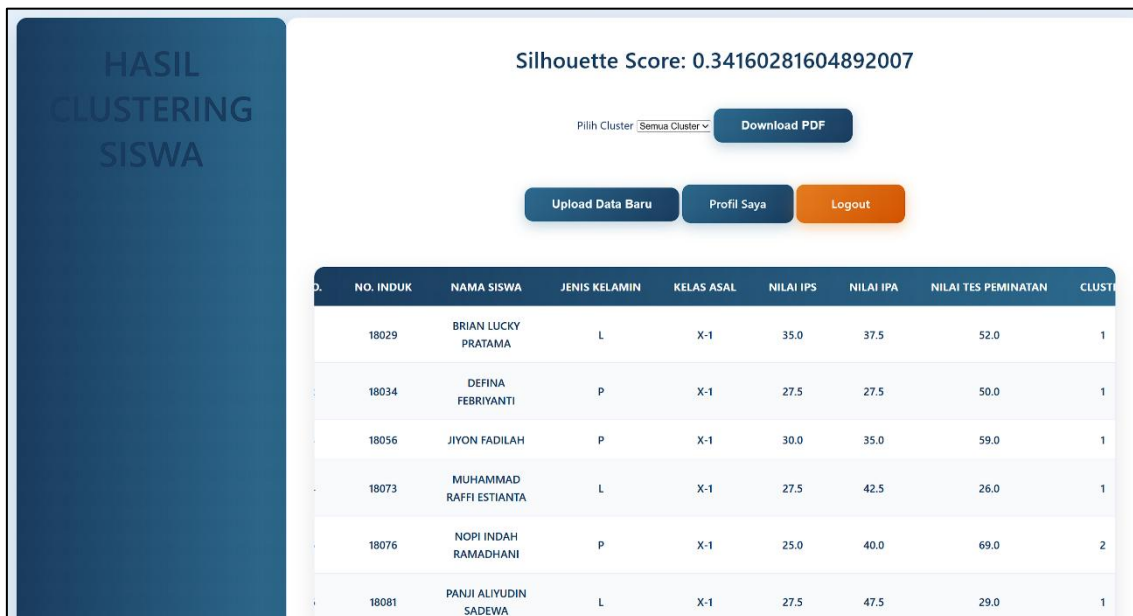
Gambar 3. Perhitungan Iterasi

Pada gambar 3 menampilkan proses perhitungan manual algoritma *K-Medoids* pada iterasi pertama, yang merupakan bagian dari tahap awal dalam proses *clustering*. Pada iterasi ini, algoritma secara acak memilih sejumlah data siswa sebagai *medoid* awal, yang ditunjukkan dalam

bentuk vektor koordinat (berisi nilai-nilai gabungan dari atribut seperti nilai IPA, IPS, dan nilai tes peminatan). Selanjutnya, dilakukan proses perhitungan jarak antar titik data terhadap *medoid* yang telah dipilih. Jarak ini dihitung menggunakan metode *Euclidean Distance*, dan hasilnya ditampilkan dalam bentuk daftar jarak dari setiap titik kandidat ke titik-titik lain dalam kluster. Pada bagian “Rincian Cost per Titik”, ditampilkan hasil perhitungan total biaya (total cost) untuk setiap kandidat sebagai *medoid*. *Cost* ini merupakan jumlah dari seluruh jarak antara *medoid* dan anggota klasternya.

E. Evaluasi Hasil *Silhouette Score*

Evaluasi digunakan pada penelitian ini menggunakan *Silhouette Score*. Dimana metrik yang akan digunakan untuk menilai seberapa data yang telah dikelompokkan dalam analisis *clustering*. Mengukur seberapa mirip sebuah objek dengan *cluster*-nya sendiri dibandingkan dengan *cluster* lain. Pengujian ini dilakukan agar sistem dapat berjalan. Tahap pengujian ini berisi hasil eksekusi program dan penjelasan tentang program yang telah dibuat. Tampilan ini memuat seluruh data siswa di SMA Pawyatan Daha Kediri. Adapun hasil yang dapat dilihat pada Gambar 4 dan 5 sebagai berikut:



NO. INDIK	NAMA SISWA	JENIS KELAMIN	KELAS ASAL	NILAI IPS	NILAI IPA	NILAI TES PEMINATAN	CLUSTER
18029	BRIAN LUCKY PRATAMA	L	X-1	35.0	37.5	52.0	1
18034	DEFINA FEBRIYANTI	P	X-1	27.5	27.5	50.0	1
18056	JIVON FADILAH	P	X-1	30.0	35.0	59.0	1
18073	MUHAMMAD RAFFI ESTIANTA	L	X-1	27.5	42.5	26.0	1
18076	NOPI INDAH RAMADHANI	P	X-1	25.0	40.0	69.0	2
18081	PANJI ALIYUDIN SADEWA	L	X-1	27.5	47.5	29.0	1

Gambar 4. Evaluasi Hasil *Silhouette Score*

Gambar 4 menyajikan hasil evaluasi *silhouette score* bernilai 0.34, yang mengindikasikan bahwa klusterisasi yang dilakukan cenderung lemah. Hal ini disebabkan oleh jarak antar-kluster tidak cukup besar dibandingkan dengan jarak intra-kluster, sehingga pengelompokan siswa berdasarkan nilai IPS dan IPA menjadi kurang optimal. Selain itu, adanya variasi nilai yang tidak konsisten, seperti perbedaan kecil antara nilai IPS dan IPA pada beberapa siswa.

Data Siswa - Semua Klaster

No.	No. Induk	Nama Siswa	Jenis Kelamin	Kelas Asal	Nilai IPS	Nilai IPA	Nilai Tes Peminatan	Cluster
1	18017	ABDILAH AHMAD RABBANI SUGIARTO	L	X-1	15.0	7.5	43.0	1
2	18018	AHMAD FADILA IN'AM NAJIB	L	X-3	30.0	42.5	67.0	2
3	18019	ALEA DILA MAHESWARI	P	X-2	40.0	45.0	38.0	1
4	18020	ALFINA DWI JAYANTI	P	X-3	22.5	17.5	58.0	1
5	18021	ALVINO ADITYA RAMADHANI	L	X-1	35.0	42.5	35.0	1
6	18022	AMELIA SEPTIHARANI	P	X-3	30.0	32.5	25.0	1
7	18023	AMELIA WIJIYANTI	P	X-2	22.0	7.5	63.0	2
8	18024	ANINDYA PARAMESTI ANANTA	P	X-3	15.0	47.5	77.0	2
9	18025	APRILIA REGITA CAHYANI	P	X-2	30.0	20.0	30.0	1
10	18026	ARDELIA KALEB KRISTIARI	P	X-1	60.0	30.0	21.0	1
11	18027	ARFINA KAMELIA AGUSTIN	P	X-2	37.5	12.5	33.0	1
12	18028	ARSITA AYU WAHYUNINGSIH	P	X-1	27.5	40.0	45.0	1
13	18029	BRIAN LUCKY PRATAMA	L	X-1	35.0	37.5	52.0	1
14	18030	CHAVIV ARDHANNA	L	X-2	30.0	22.5	29.0	1
15	18031	CITRA DWI FEBRIYANTI	P	X-1	35.0	32.5	20.0	1
16	18032	DAVINO MARCO FIRSYANDO ZHAFIRO	L	X-2	42.5	55.0	68.0	2
17	18033	DE'A SELVIANA	P	X-3	30.0	40.0	59.0	1
18	18034	DEFINA FEBRIYANTI	P	X-1	27.5	27.5	50.0	1
19	18035	DHEANITA PUSPITA DEWI	P	X-2	32.5	62.5	50.0	1
20	18036	DHYNDA FACCA RAHMADHANY	P	X-2	22.5	35.0	81.0	2
21	18037	DINDA AYUNDA SARI	P	X-2	47.5	37.5	75.0	2
22	18038	DINDA PUSPITASARI	P	X-1	27.5	25.0	25.0	1
23	18039	ERISA FITRIA RAHMAWATI	P	X-2	37.5	17.5	31.0	1
24	18040	ERLAN BANA PUTRA RAHMAWAN	L	X-2	37.5	37.5	84.0	2
25	18041	ERLIS NAILA FARIKHA AGUSTIN	P	X-3	27.5	22.5	64.0	2
26	18042	ERVINA NOVELIA YOSSAM	P	X-3	20.0	50.0	58.0	1

Gambar 5. Hasil *Output* Data Siswa

Gambar 5 menyajikan hasil *output* data siswa kelas X yang telah berhasil diklasterkan ke dalam dua klaster berdasarkan nilai masing-masing siswa, dan data ini dapat dicetak dalam bentuk PDF untuk keperluan dokumentasi.

IV. KESIMPULAN

Penelitian ini membuktikan bahwa algoritma *K-Medoids* efektif untuk merekomendasikan mata pelajaran pilihan siswa berdasarkan atribut data nilai IPA, nilai IPS, dan nilai tes peminatan. *K-Medoids* mampu membentuk dua kelompok siswa yang representatif (IPA dan IPS), sehingga setiap klaster dapat diberikan rekomendasi mata pelajaran lanjutan yang sesuai. Validasi menggunakan nilai *Silhouette* memperlihatkan bahwa hasil *clustering* cukup baik dan seimbang. Keuntungan sistem ini adalah pengelompokan yang lebih objektif dan berbasis data, mengurangi subjektivitas dalam proses manual. Dengan demikian, penerapan *data mining K-Medoids* dapat meningkatkan akurasi dan efisiensi dalam pemilihan mata pelajaran pilihan di lingkungan Kurikulum Merdeka.

DAFTAR PUSTAKA

- [1] S. Mahaly, J. O. Papilaya, and Jumail, "Analisis Pemilihan Minat Mata Pelajaran Pilihan Siswa Sma Laboratorium Universitas Pattimura," PEDAGOGIKA: Jurnal Pedagogik dan Dinamika Pendidikan, vol. 12, no. 1, pp. 101–108, Apr. 2024, doi: <https://doi.org/10.30598/pedagogikavol12issue1page101-108>

- [2] C. Zai and T. Komputer, “Implementasi *Data Mining* Sebagai Pengolahan Data,” *Portaldata.org*, vol. 2, no. 3, 2022.
- [3] Y. Syahra, “Penerapan Data Mining Dalam Pengelompokkan Data Nilai Siswa Untuk Penentuan Jurusan Siswa Pada SMA Tamora Menggunakan Algoritma K-Means Clustering,” *j.sains manaj. inform. dan komp’ut.*, vol. 17, no. 2, hlm. 228, Agu 2018, doi: 10.53513/jis.v17i2.70.
- [4] F. Harahap, “Perbandingan Algoritma K Means dan K Medoids Untuk Clustering Kelas Siswa Tunagrahita,” *TIN: Terapan Informatika Nusantara*. vol. 2, no. 4, 2021.
- [5] N. T. Luchia, H. Handayani, F. S. Hamdi, D. Erlangga, dan S. F. Octavia, “Perbandingan K-Means dan K-Medoids Pada Pengelompokan Data Miskin di Indonesia: Comparison of K-Means and K-Medoids on Poor Data Clustering in Indonesia,” *MALCOM*, vol. 2, no. 2, hlm. 35–41, Sep 2022, doi: 10.57152/malcom.v2i2.422.
- [6] R. Adrianto dan A. Fahmi, “Penerapan Metode Clustering Dengan Algoritma K-Means Untuk Rekomendasi Pemilihan Jalur Peminatan Sesuai Kemampuan Pada Progam Studi Teknik Informatika - S1 Universitas Dian Nuswantoro”, doi: <https://doi.org/10.33633/joins.v1i2.1302>.
- [7] A. R. Nst, D. Irmayani, dan A. P. Juledi, “Clustering Of Algortima K-Means Based National Exam Score Data With Elbow And Silhouette Optimization,” *INFOKUM*, vol. 10, 2022.
- [8] R. Saputra dan I. Purnama, “Optimizing K-Means Algorithm With Elbow And Silhouette Methods For National Exam Score Data Clustering”, *Jurnal Ilmu Komputer Ruru*, vol 1 no. 1, 2024.
- [9] J. Faran dan R. T. Aldisa, “Penerapan Data Mining Untuk Penjurusan Kelas dengan Menggunakan Algoritma K-Medoids,” *bits*, vol. 5, no. 2, Sep 2023, doi: 10.47065/bits.v5i2.4313.
- [10] E. Fitriani, D. Firmansyah, R. Aryanti, dan W. Walim, “Implementasi Model Waterfall Pada Sistem Informasi Akademik Berbasis Web Pada Smk Pertanian Karawang,” *Techno*, vol. 15, no. 2, hlm. 137, Sep 2018, doi: 10.33480/techno.v15i2.923.
- [11] A. L. Ramdani dan H. B. Firmansyah, “Clustering Application for UKT Determination Using Pillar K-Means Clustering Algorithm and Flask Web Framework,” *IJAIDM*, vol. 1, no. 2, hlm. 53, Agu 2018, doi: 10.24014/ijaidm.v1i2.5126.
- [12] B. N. Yulisasih, H. Herman, S. Sunardi, dan H. Yuliansyah, “Evaluation of K-Means Clustering Using Silhouette Score Method on Customer Segmentation,” *Ilk. J. Ilm.*, vol. 16, no. 3, hlm. 330–342, Des 2024, doi: 10.33096/ilkom.v16i3.2325.330-342.

- [13] H. Sariangseh, W. Wanayumini, dan R. Rosnelly, “Penentuan Kelas Menggunakan Algoritma K Medoids Untuk Clustering Siswa Tunagrahita,” *mib*, vol. 5, no. 1, hlm. 83, Jan 2021, doi: 10.30865/mib.v5i1.2547.
- [14] A. Voutama dan E. Novalia, “Perancangan Sistem Informasi Plakat Wisuda Berbasis Web Menggunakan UML dan Model Waterfall,” 15 Jun 2022, vol. Vol 11 No 01 (2022), doi: 10.35706/syji.v11i01.6412.