

KLASTERISASI DATA SISWA BERDASARKAN NILAI AKADEMIK DAN PERILAKU MENGGUNAKAN K-MEANS CLUSTERING BERBASIS WEB DI SDN CIPONDOH 2

Farah Dyah Salsabila^{1*}, Imelda Imelda²

^{1,2}Teknik Informatika, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta, Indonesia.

Email: ^{1*}2011501901@student.budiluhur.ac.id, ²imelda@budiluhur.ac.id

(* : corresponding author)

Abstrak- Variasi kemampuan akademik dan perilaku siswa di SDN Cipondoh 2 merupakan tantangan dalam menyediakan layanan pendidikan yang merata. Guru sering kesulitan memberikan perhatian yang seimbang kepada semua siswa, sehingga menghambat perkembangan mereka. Untuk mengatasi hal ini, metode K-Means Clustering digunakan untuk mengelompokkan siswa berdasarkan nilai akademik dan perilaku. Metode ini membagi siswa ke dalam kelompok-kelompok dengan karakteristik serupa, memudahkan guru dalam menyampaikan pembelajaran yang sesuai dengan kebutuhan setiap kelompok. Penelitian ini melibatkan analisis data akademik dan perilaku siswa yang diolah menggunakan K-Means Clustering melalui aplikasi berbasis web menggunakan mysql dan pemrograman PHP. Aplikasi ini membantu guru dan pengelola sekolah dalam memahami data untuk meningkatkan kualitas pendidikan. Hasil pengujian menunjukkan bahwa dari 231 data siswa, terbentuk tiga cluster: 72 siswa dalam kelompok dengan nilai tertinggi (pemahaman cepat), 80 siswa dalam kelompok nilai menengah (pemahaman lambat), dan 79 siswa dalam kelompok nilai terendah (membutuhkan perhatian lebih). Evaluasi pengelompokan dilakukan dengan Davies-Bouldin Index (DBI), untuk menunjukkan efektivitas metode ini dalam membentuk kelompok siswa dengan karakteristik yang mirip dan terpisah dengan baik. Evaluasi pengelompokan dilakukan dengan Davies-Bouldin Index (DBI), untuk menunjukkan efektivitas metode ini dalam membentuk kelompok siswa dengan karakteristik yang mirip dan terpisah dengan baik. Evaluasi hasil clustering menggunakan DBI (Davies Bouldin Index) menghasilkan nilai 0.93837 artinya hasil klasterisasi cukup baik, karena mendekati nilai 0. Dengan adanya pengelompokan ini, sekolah dapat merancang metode pengajaran yang lebih tepat sasaran dan sesuai dengan kebutuhan masing-masing kelompok siswa, yang diharapkan dapat meningkatkan kualitas pendidikan secara keseluruhan.

Kata Kunci: K-Means Clustering, Pengelompokan Siswa, Perilaku Siswa.

CLUSTERING STUDENT DATA BASED ON ACADEMIC GRADES AND BEHAVIOR USING WEB-BASED K-MEANS CLUSTERING AT SDN CIPONDOH 2

Abstract- *The variation in academic abilities and behavior of students at SDN Cipondoh 2 poses a challenge in providing equitable educational services. Teachers often struggle to give balanced attention to all students, which hinders their development. To address this, the K-Means Clustering method is used to group students based on academic performance and behavior. This method divides students into groups with similar characteristics, making it easier for teachers to deliver lessons that meet the needs of each group. This research involves the analysis of academic data and student behavior processed using K-Means Clustering through a web-based application utilizing MySQL and PHP programming. This application assists teachers and school administrators in understanding data to enhance the quality of education. The test results show that out of 231 student data, three clusters were formed: 72 students in the highest score group (quick understanding), 80 students in the middle score group (slow understanding), and 79 students in the lowest score group. (membutuhkan perhatian lebih). Clustering evaluation was conducted using the Davies-Bouldin Index (DBI) to demonstrate the effectiveness of this method in forming groups of students with similar characteristics that are well-separated. The evaluation of clustering results using the DBI (Davies-Bouldin Index) yielded a value of 0.93837, indicating that the clustering results are quite good, as they approach a value of 0. With this grouping, schools can design teaching methods that are more targeted and suited to the needs of each student group, which is expected to improve the overall quality of education.*

Keywords: *K-Means Clustering, Student Clustering, Student Behavior.*

1. PENDAHULUAN

Perkembangan teknologi informasi di Indonesia yang pesat telah mempengaruhi berbagai aspek kehidupan, termasuk pendidikan. Pendidikan merupakan elemen vital dalam membentuk karakter dan kualitas sumber daya manusia, sehingga memonitor dan menganalisis data siswa secara rutin menjadi sangat penting untuk memastikan efektivitas proses pembelajaran. Tantangan utama dalam pendidikan adalah memberikan pembelajaran yang merata kepada siswa dengan variasi daya tangkap yang sangat beragam. Di SDN Cipondoh 2 Beberapa siswa memahami materi dengan cepat, sementara yang lain membutuhkan waktu lebih lama dan perhatian ekstra, yang menyulitkan guru dalam memberikan penjelasan yang efektif kepada seluruh siswa. Karena setiap siswa berhak mendapatkan kesempatan belajar yang sama, namun kesenjangan dalam daya tangkap menyebabkan beberapa siswa tidak mendapatkan perhatian yang cukup sehingga tidak dapat memaksimalkan potensinya.

Oleh karena itu, penting untuk mengelompokkan siswa dan memberikan kelas tambahan agar semua siswa dapat memperoleh pemahaman yang setara. Dalam penelitian ini data yang digunakan sebanyak 231 data siswa yang diambil dari rata-rata nilai akademik ujian tengah semester sebanyak 8 mata pelajaran dan nilai perilaku siswa selama disekolah berdasarkan skala penilaian. Penelitian ini difokuskan pada penerapan teknologi data mining dengan metode k-means dalam klasterisasi data siswa di SDN Cipondoh 2. Penerapan teknologi berupa aplikasi pengelompokan siswa dapat memfasilitasi penilaian yang lebih objektif dan sistematis terhadap performa akademis dan perilaku siswa, memungkinkan sekolah untuk mengelompokkan siswa dengan lebih efektif.

Data mining adalah serangkaian proses yang bertujuan untuk mencari pola, hubungan dan nilai tambah dari data dan informasi berukuran besar, sehingga dapat diubah menjadi bentuk pengetahuan yang bermanfaat.[1] Teknologi data mining dapat digunakan untuk mengolah dan mengelompokkan data siswa secara cepat dan akurat. Salah satu teknik yang dapat digunakan adalah clustering, *Clustering* adalah teknik penting dalam data mining yang digunakan untuk mengelompokkan data menjadi subset berdasarkan karakteristik atau pola tertentu.[2] Algoritma K-Means adalah salah satu algoritma *clustering* yang paling umum dan sederhana. Algoritma K-Means adalah salah satu algoritma data mining yang digunakan untuk mengelompokkan data.[3] Metode ini mengelompokkan data ke dalam kelompok yang memiliki karakteristik serupa, sementara data dengan karakteristik yang berbeda akan dimasukkan ke dalam kelompok lain. Metode ini telah terbukti efektif dalam berbagai penelitian sebelumnya. Misalnya, penelitian oleh Erdi Amos Saputra & Yessica Nataliani (2021)[4] yang mengelompokkan data nilai siswa untuk menentukan siswa berprestasi, dan penelitian oleh Juniar Hutagalung, Yopi Hendro Syahputra, dan Zohana Pertiwi Tanjung (2022)[5] yang memetakan siswa unggulan berdasarkan nilai rapor dan perilaku. Penelitian-penelitian ini menunjukkan bahwa dengan menggunakan metode k-means clustering, pengelompokan siswa dapat dilakukan dengan lebih objektif dan efektif, sehingga membantu guru dalam memberikan perhatian yang lebih tepat sasaran kepada setiap siswa[5].

Penelitian ini menggunakan evaluasi pengelompokkan Davies Bouldin Index DBI yang digunakan untuk menilai seberapa baik klaster yang dihasilkan terpisah satu sama lain dan seberapa homogen anggota klaster tersebut. DBI sangat berguna dalam mengevaluasi metode klasterisasi seperti K-Means untuk memastikan bahwa kelompok yang terbentuk memiliki karakteristik yang serupa di dalam klaster dan berbeda dengan klaster lainnya[6]. Berbeda dengan penelitian sebelumnya oleh Erdi Amos Saputra dan Yessica Nataliani yang berjudul “Analisis Pengelompokkan Data Nilai Siswa untuk Menentukan Siswa Berprestasi Menggunakan Metode *Clustering* K-Means” [4] Penelitian ini menekankan pengumpulan data yang mencakup nilai akademik dan perilaku siswa dengan tujuan membantu guru agar lebih efektif dalam mengatasi variasi daya tangkap siswa. Selain itu, hasil klasterisasi dari algoritma K-Means dievaluasi menggunakan metode *Davies Bouldin Indeks*. Evaluasi ini memungkinkan analisis hasil klasterisasi dengan berbagai jumlah klaster, sehingga dapat menilai kualitas klasterisasi secara lebih mendalam.

Dengan demikian, penerapan teknologi data mining dengan metode k-means clustering diharapkan dapat membantu SDN Cipondoh 2 dalam mengatasi kesenjangan dalam pembelajaran, memastikan bahwa setiap siswa mendapatkan kesempatan belajar yang sama, serta meningkatkan kualitas pendidikan secara keseluruhan. Model pengelompokan ini juga dapat diterapkan di sekolah-sekolah lain dengan masalah serupa, menjadikannya solusi yang bermanfaat secara luas.

2. METODE PENELITIAN

2.1 Data Penelitian

Data penelitian adalah kumpulan informasi yang dikumpulkan dan dianalisis dalam konteks kegiatan penelitian untuk menjawab pertanyaan penelitian atau mencapai tujuan penelitian. Dalam pembuatan data mining dengan klasterisasi K-Means ini, tujuan utamanya adalah memudahkan pengelompokan siswa berdasarkan

kebutuhan pendidikan mereka. Klasterisasi data mining ini memberikan banyak keuntungan dari segi waktu dan efektivitas.

Data yang digunakan dalam penelitian ini diperoleh dari SDN Cipondoh 2 sebanyak 231 data siswa yang mencakup rata-rata dari nilai ujian tengah semester ganjil dan penilaian sikap/perilaku murid tahun ajaran 2024/2025 dalam bentuk file *Microsoft Excel*. Adapun data yang diberikan oleh SDN Cipondoh 2 meliputi atribut-atribut berikut ini :

Tabel 1. Data Atribut

Atribut	Keterangan
Nama	Nama Siswa
PKN	Pendidikan Kewarganegaraan
B.IND	Bahasa Indonesia
MTK	Matematika
SBDP	Seni dan Budaya
B.ING	Bahasa Inggris
BP	Budi Pekerti
AGM	Agama
PJOK	Pendidikan Jasmani Olahraga dan Kesehatan
Rata – rata	Rata – rata nilai Ujian Tengah Semester Siswa
Sikap	Nilai perilaku/sikap siswa

Dalam penilaian sikap siswa menggunakan acuan skala penilaian yang bersumber dari buku ahli yang berjudul “Skala Pengukuran variabel-variabel penelitian” yang ditulis oleh Dr. Riduwan, M.B.A. Skala kriteria penilaian dapat dilihat pada tabel dibawah ini :

Tabel 2. Skala Penilaian

Tingkat Persentase Aspek	Kriteria
81 - 100	Sangat Baik
61 - 80	Baik
41 - 60	Cukup Baik
21 - 40	Kurang
0 - 20	Sangat kurang

2.2 Tahap Pengumpulan Data

Dalam melaksanakan penelitian, penulis memerlukan data yang relevan dengan topik yang dibahas. Jenis pengumpulan data yang digunakan dalam penelitian ini meliputi :

2.2.1 Wawancara

Tahapan pengumpulan data dalam penelitian ini dilakukan dengan bertanya langsung kepada wali dari setiap kelas di SDN Cipondoh 2 untuk memperoleh data dan informasi mengenai kriteria penilaian, jumlah mata pelajaran, serta sistem perhitungan nilai akademik dan perilaku setiap siswa. Wawancara ini bertujuan untuk mendapatkan wawasan mendalam tentang sistem penilaian yang digunakan dan pemahaman guru mengenai perilaku serta karakteristik siswa yang akan menjadi acuan dalam pembuatan aplikasi. Dalam penilaian sikap siswa menggunakan acuan skala penilaian yang bersumber dari buku ahli yang berjudul “Skala Pengukuran variabel-variabel penelitian” yang ditulis oleh Dr. Riduwan, M.B.A. Kriteria penilaian dapat dilihat pada tabel.

2.2.2 Studi Pustaka

Tahapan pengumpulan data dalam penelitian ini bertujuan untuk memperoleh informasi yang lebih mendalam mengenai data mining dan K-Means *Clustering* melalui jurnal-jurnal yang akan digunakan sebagai referensi dalam pelaksanaan penelitian.

2.3 Tahapan *Knowledge Discovery in Database* (KDD)

2.3.1 Pemilihan Data

Pada tahap ini, data yang relevan untuk penelitian diidentifikasi, yaitu data nilai akademik dan perilaku siswa dari SDN Cipondoh 2. Data tersebut dikumpulkan dari nilai ujian serta catatan guru atau wali kelas. Tahap ini bertujuan untuk memilih atribut yang relevan dalam data sehingga dapat dilakukan analisis. File excel dari setiap kelas digabungkan dan atribut “kelas” ditambahkan untuk memberikan informasi tentang kelas setiap siswa. Dari beberapa atribut yang telah disebutkan sebelumnya, dipilih 5 atribut sebagai berikut:

- 1) Nama
- 2) Kelas
- 3) Jenis Kelamin
- 4) Rata-rata
- 5) Sikap

2.3.2 Pra-pemrosesan Data

Data yang telah dikumpulkan dibersihkan dari kesalahan, duplikasi dan ketidakkonsistenan. Penanganan data yang hilang dilakukan seperti mengisi nilai yang hilang atau menghapus data yang tidak lengkap. Selanjutnya, data nilai akademik dan perilaku dinormalisasi untuk memastikan skala yang konsisten.

2.3.3 Transformasi Data

Data yang telah dibersihkan diubah ke dalam format yang sesuai untuk proses klusterisasi. Jika diperlukan, atribut baru ditambahkan seperti rata-rata nilai akademik atau nilai perilaku. Data kemudian diubah ke dalam bentuk tabel atau matriks yang dapat diolah oleh algoritma *K-Means Clustering*.

2.3.4 Data Mining

Algoritma *K-Means Clustering* digunakan untuk mengelompokkan siswa berdasarkan nilai akademik dan perilaku. Algoritma *K-Means* adalah salah satu dalam algoritma data mining yang digunakan untuk mengelompokkan data. Algoritma ini dapat membagi data jadi dua atau lebih kelompok sesuai pengelompokan data nonhierarki (partisi). Metode ini dapat mengelompokkan data ke kelompok yang memiliki karakteristik serupa, sementara data dengan karakteristik yang berbeda akan dimasukkan ke dalam kelompok lain.[7]

1. menentukan jumlah klaster (k) dan menetapkan pusat klaster secara acak/random.
2. menghitung jarak tiap data ke dalam pusat klaster.
3. Kelompokkan data ke klaster yang memiliki jarak terdekat.
4. Hitung ulang pusat klaster.
5. Ulangi langkah pada langkah ke-2 sampai ke-4 hingga tidak ada data yang pindah klaster lagi.

tahap clustering diawali dengan mengidentifikasi data yang akan dikelompokkan. Dapat menggunakan rumus Euclidean Distance seperti yang tertera pada rumus persamaan dibawah ini :

$$d_{ij} = \sqrt{\sum_{k=1}^m (X_{ij} - C_{jk})^2}$$

Dimana D_{ij} mewakili jarak objek antar nilai data dan nilai ke pusat cluster, m dirumus adalah jumlah dimensi dalam data, X_{ij} adalah nilai dalam data pada dimensi ke- k , dan C_{jk} adalah nilai pada pusat cluster dimensi ke- k .

Untuk menentukan centroid yang baru, dapat ditetapkan rumus persamaan dibawah ini :

$$C = \frac{\sum m}{n}$$

Dimana C adalah centroid dari data, m merupakan anggota data yang berada di dalam centroid tertentu dan menunjukkan jumlah data yang termasuk dalam centroid tersebut. [8]

2.3.5 Evaluasi Pola

Hasil clustering dievaluasi untuk memastikan kelompok yang terbentuk sesuai dengan tujuan penelitian. Metrik validitas seperti *Davies-Bouldin Index* digunakan untuk mengevaluasi kualitas klaster. *Davies-Bouldin*

Indeks (DBI) adalah salah satu teknik untuk mengevaluasi kluster yang digunakan untuk meningkatkan kinerja metode, dengan tujuan menentukan jumlah kluster yang optimal. Sum of Square Within Cluster (SSW) adalah persamaan yang digunakan untuk menghitung matriks kohesi dalam suatu kluster ke-i dan dapat dirumuskan sebagai berikut ini [9]:

$$SSW_i = \frac{1}{m_i} \sum_{j=1}^{m_i} d(x_i, c_i)$$

Keterangan :

m = jumlah suatu data dalam kluster ke-i

c = centroid dari cluster ke-i

d = jarak antar titik

Untuk menghitung pemisahan antara kluster, dapat digunakan Sum of Square Between Cluster (SSB), yang diuraikan sebagai berikut ini [10] :

$$SSB_{ij} = d(c_i, c_j)$$

Keterangan :

d = jarak antar titik

c = centroid dari cluster ke-i

Setelah mendapat nilai kohesi dan pemisahan, langkah selanjutnya adalah mengukur rasio (R_{ij}) untuk menilai perbandingan antara cluster ke-i dan ke-j. Kluster memiliki nilai kohesi terkecil dan nilai pemisah terbesar dianggap sebagai kluster yang baik. Berikut adalah persamaan untuk menghitung nilai rasio tersebut [10] :

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_{ij})$$

K adalah jumlah kluster yang dipakai, semakin kecil nilai DBI yang didapat (yang non-negatif dan ≥ 0), semakin optimal kluster yang didapat.

2.3.6 Interpretasi dan Presentasi Pengetahuan

Hasil pengelompokan disajikan dalam bentuk visualisasi, seperti grafik atau diagram yang menunjukkan kelompok siswa berdasarkan nilai akademik dan perilaku. Aplikasi web dikembangkan untuk memungkinkan guru dan sekolah melihat dan menganalisis hasil clustering. Laporan yang merinci temuan penelitian disusun dan memberikan rekomendasi untuk intervensi pendidikan berdasarkan kelompok yang terbentuk. Dengan langkah-langkah ini, proses KDD dilaksanakan secara sistematis untuk menghasilkan pengetahuan yang bermanfaat bagi peningkatan kualitas pendidikan di SDN Cipondoh 2.

3. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan metode klusterisasi dengan algoritma K-Means yang meliputi beberapa langkah utama. Pertama, menentukan jumlah kluster yang diinginkan, diikuti oleh inisialisasi pusat kluster. Selanjutnya, iterasi dilakukan untuk menemukan kluster yang paling optimal. Proses ini melibatkan pengelompokan data ke dalam kluster berdasarkan kedekatannya dengan pusat kluster terdekat, memperbarui pusat kluster berdasarkan rata-rata data dalam kluster dan mengulangi langkah-langkah tersebut sampai mencapai kestabilan. Hasil akhir dari proses ini adalah pembentukan kelompok kluster yang menggambarkan karakteristik dalam data yang kemudian dianalisis lebih lanjut untuk memudahkan pemahaman hasil klusterisasi dan visualisasi yang digunakan.

3.1 Pemahaman Data

Langkah awal dalam penerapan metode ini adalah mengumpulkan data. Dalam penelitian ini, data yang digunakan untuk pemodelan clustering adalah nilai ujian tengah semester ganjil dari kelas 1A, 1B, 2A, 2B, 3A, 5A dan 5B dengan total 231 baris data. Berikut contoh sampel dataset yang digunakan dalam perhitungan k-means clustering. Nilai Rata-rata diambil dari hasil rata-rata dari nilai ujian tengah semester siswa dan nilai sikap diambil dari hasil wawancara dengan wali kelas berdasarkan skala kriteria penilaian menurut ahli. Detail dari data sampel dapat dilihat pada Tabel 3 :

Tabel 3. Data Sampel

NO	NAMA	KELAS	RATA2	SIKAP
1	ABRAHAM GHAZALI TACHIR	2B	92,5	80
2	AFLAH	2B	86,1	80
3	AISYAH AL ZAHRA ZOHANA	2B	80,5	80
4	ALAYYA AUDIA INARA	2B	86,6	70
5	ALESHA PUTRI HAFIZH	2B	79,9	80
6	ALIKA NURFARANISA AZNI	2B	80,3	90
7	AQILA ANINDYA SULISTIYANTO	2B	91,4	80
8	ARKA SYAKIR RAMADHAN	2B	67,1	65
9	ATHAYA NABIHAN ZAYAN NIDIE	2B	81,1	70
10	AZIZAH RAMADHANIA MAULANA	2B	75,4	75
11	AZKA ALDRINATA	2B	88,5	80
12	AZZAHRA SHAVIRA SETIAWATI	2B	74,6	75
13	BARLY ALDRIC RAILA	2B	90,3	80
14	CITTAHARIN KAVYA WARDANA	2B	87,5	75
15	FAKHRUL ATHAYA PUTROWAR	2B	88,1	80
16	GADIS ADEEVA DZAKIRA	2B	80,5	80
17	GIBRAN RADIKA GALIH	2B	77,8	70
18	HUBBU KHIAR MAKTALDINATA	2B	83,1	90
19	KHALIQA SHANUM ISKANDAR	2B	68	85
20	KHANSA ARSYFA SALSABILA	2B	64,9	75

3.2 Perhitungan Algoritma K-Means

Langkah awal dalam perhitungan algoritma K-means adalah menetapkan jumlah kelompok yang diinginkan. Pada perhitungan ini, telah ditetapkan sebanyak 3 kelompok. Setelah jumlah kelompok ditentukan, langkah berikutnya adalah memilih secara acak 3 data yang berfungsi sebagai centroid awal. Nilai centroid awal yang dipilih secara acak dapat dilihat pada Tabel 4 berikut ini:

Tabel 4. Inisialisasi Centroid Awal

Centroid	Rata-rata	Sikap/Perilaku
1	68	60
2	78	95
3	94	80

Langkah berikutnya adalah menghitung jarak antara setiap data dengan masing-masing centroid menggunakan rumus Euclidean Distance. Proses perhitungan ini dilakukan dan dihasilkan sebagai berikut :

Tabel 5. Iterasi Ke-1

Centroid	Rata-rata	Sikap
1	68	60
2	78	95
3	94	80
Hasil Cluster Iterasi Ke-1		
Jumlah C1		69 Siswa
Jumlah C2		81 Siswa
Jumlah C3		81 Siswa

Tabel 6. Perhitungan Iterasi Ke-1

NO	NAMA	KELAS	RATA	SIKAP	C1	C2	C3	HASIL CLUSTER
1	ABRAHAM GHAZALI TACHIR	2B	93	80	32,015	21,213	1,000	3
2	AFLAH	2B	86	80	26,907	17,000	8,000	3
...
230	ZAHRA ZHAFIRA ZOHHASYAH	5A	76	88	29,120	7,280	19,697	2
231	M. SADIQ	5A	77	85	26,570	10,049	17,720	2

Selanjutnya, perlu ditentukan nilai centroid baru untuk iterasi berikutnya menggunakan rumus persamaan sebagai berikut :

$$C = \frac{\sum m}{n}$$

Pada setiap iterasi, dilakukan perhitungan ulang *centroid* berdasarkan data yang diperbarui. Mulai dari iterasi kedua dan seterusnya. Nilai *centroid* baru dan hasil Iterasi ke-2 dapat dilihat pada tabel dibawah ini :

Tabel 7. Iterasi Ke-2

Centroid Baru	Rata-rata	Sikap
1	74,0144	68,492
2	76,037	84,061
3	85,148	79,197
Hasil Cluster Iterasi Ke-2		
Jumlah C1		77 Siswa
Jumlah C2		80 Siswa
Jumlah C3		74 Siswa

Tabel 8. Perhitungan Iterasi Ke-2

NO	NAMA	KELAS	RATA	SIKAP	C1	C2	C3	HASIL CLUSTER
1	ABRAHAM GHAZALI TACHIR	2B	93	80	22,200	17,442	7,892	3
2	AFLAH	2B	86	80	16,615	10,759	1,170	3
...
230	ZAHRA ZHAFIRA ZOHHASYAH	5A	76	88	19,608	3,938	12,695	2
231	M. SADIQ	5A	77	85	16,775	1,344	10,003	2

Jika hasil clustering masih mengalami perubahan, maka perlu ditentukan nilai centroid baru untuk iterasi berikutnya. Langkah tersebut diulangi hingga centroid terakhir. Hasil centroid baru dan hasil Iterasi ke-3 dapat dilihat pada tabel dibawah ini :

Tabel 9. Iterasi Ke-3

Centroid Baru	Rata-rata	Sikap
1	74,445	68,729
2	75,562	83,55
3	85,831	80,259
Hasil Cluster Iterasi Ke-3		
Jumlah C1		77 Siswa
Jumlah C2		81 Siswa
Jumlah C3		73 Siswa

Tabel 10. Perhitungan Iterasi Ke-3

NO	NAMA	KELAS	RATA	SIKAP	C1	C2	C3	HASIL CLUSTER
1	ABRAHAM GHAZALI TACHIR	2B	93	80	21,708	17,795	7,173	3
2	AFLAH	2B	86	80	16,140	11,024	0,309	3
...
230	ZAHRA ZHAFIRA ZOHHASYAH	5A	76	88	19,332	4,471	12,512	2
231	M. SADIQ	5A	77	85	16,469	2,041	10,022	2

Pada iterasi keenam, proses *clustering* telah selesai, karena semua data telah dikelompokkan secara konsisten dan tidak mengalami perubahan. Oleh karena itu, iterasi tambahan tidak lagi dilakukan. Hasil akhir dari perhitungan clustering menunjukkan bahwa 72 siswa dalam kelompok dengan nilai tertinggi (pemahaman cepat) yaitu Cluster 1, 80 siswa dalam kelompok nilai menengah (pemahaman lambat) yaitu Cluster 2, dan 79 siswa dalam kelompok nilai terendah (membutuhkan perhatian lebih), yaitu Cluster 3.

Tabel 11. Iterasi Ke-6

Centroid Baru	Rata-rata	Sikap
1	75	68
2	75	82
3	86	81
Hasil Cluster Iterasi Ke-6		
Jumlah C1		79 Siswa
Jumlah C2		80 Siswa
Jumlah C3		72 Siswa

3.3 Sum of Square Within Cluster (SSW)

Menghitung hasil *Sum of Square Within Cluster* (SSW) adalah langkah awal dalam menghitung evaluasi hasil clustering. Berikut adalah nilai SSW yang telah diperoleh dari hasil clustering diatas.

Tabel 12. Sum of Square Within Cluster (SSW)

SSW	Cluster 1	Cluster 2	Cluster 3
Nilai	7.478	4.903	5.736

3.4 Sum of Square Between Cluster (SSB)

Selanjutnya menghitung *Sum of Square Between Cluster* (SSB). Berikut adalah nilai SSB yang telah diperoleh dari hasil clustering diatas.

Tabel 13. Sum of Square Between Cluster (SSB)

SSB	Cluster 1	Cluster 2	Cluster 3
Cluster 1	0	14.015	17.072
Cluster 2	14.015	0	11.055
Cluster 3	17.072	11.055	0

3.5 Davies Bouldien Index (DBI)

Langkah terakhir pada tahap evaluasi hasil clustering adalah dengan menggunakan metode Davies-Bouldien Index dengan menghitung rata-rata nilai dari setiap titik dalam himpunan data yang dapat membantu mengevaluasi seberapa baik klaster. Berikut adalah perhitungan yang dihasilkan :

Tabel 14. Davies-Bouldien Index (DBI)

R	Cluster 1	Cluster 2	Cluster 3	R Max
Cluster 1	0	0.88543	0.77400	0.88543
Cluster 2	0.88543	0	0.96484	0.96484
Cluster 3	0.77400	0.96484	0	0.96484

Dari nilai Rasio Maksimal yang diperoleh kita dapat mencari nilai DBI dengan perhitungan sebagai berikut:

$$DBI = 1/3(0.88543 + 0.96484 + 0.96484) = 0.93837$$

Evaluasi hasil *clustering* menggunakan *Davies Bouldien Index* (DBI) menghasilkan nilai 0.93837 artinya cukup baik, karena semakin dekat nilai DBI dengan 0, semakin baik hasil *clustering* yang diperoleh.

4. KESIMPULAN

Penelitian ini mengidentifikasi segmentasi karakteristik yang menjadi masalah utama di SDN Cipondoh 2, yaitu variasi yang signifikan dalam kemampuan akademik dan perilaku siswa. Perbedaan ini menyebabkan tantangan bagi guru dalam memberikan pembelajaran yang merata dan sesuai dengan kebutuhan setiap siswa. Melalui metode K-Means Clustering, siswa dikelompokkan ke dalam tiga cluster berdasarkan nilai akademik dan perilaku mereka. Cluster pertama terdiri dari 72 siswa dengan pemahaman cepat dan nilai akademik tertinggi, cluster kedua terdiri dari 80 yang mencakup siswa dengan pemahaman lambat dan nilai akademik menengah, sementara cluster ketiga terdiri dari 79 yang mencakup siswa yang membutuhkan perhatian lebih dengan nilai akademik terendah. Selain perbedaan kemampuan akademik, perilaku siswa yang bervariasi juga mempengaruhi proses pembelajaran dan interaksi di kelas. Masalah utama yang muncul dari segmentasi ini adalah kesenjangan dalam perhatian yang diberikan oleh guru, di mana siswa yang memerlukan dukungan lebih sering kali tidak mendapatkan perhatian yang memadai, sehingga menghambat perkembangan potensi mereka secara optimal. Implementasi teknologi data mining dengan metode K-Means Clustering terbukti efektif dalam mengelompokkan siswa berdasarkan kebutuhan pendidikan mereka, memisahkan siswa ke dalam beberapa kelompok berdasarkan karakteristik nilai akademik dan perilaku. Aplikasi berbasis web yang dikembangkan memfasilitasi pengelompokan siswa secara objektif, membantu guru dalam merancang strategi pengajaran yang lebih tepat sasaran sesuai dengan kebutuhan masing-masing kelompok siswa.

DAFTAR PUSTAKA

- [1] H. Syukron, "Perbandingan K-Means K-Medoids dan Fuzzy C-Means untuk Pengelompokan Data Pelanggan dengan Model LRFM", *MALCON: Indonesian Journal of Machine Learning and Computer Science*, vol. 2, no. 2, pp. 76-83, 2022.
- [2] N. Hendrastuty, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa," *Jurnal Ilmiah Informatika dan Ilmu Komputer (Jima-Ilkom)*, vol. 3, no. 1, pp. 46-56, 2024, doi: 10.58602/jima-ilkom.v3i1.26.
- [3] S. Nurdiani, S. Linawati, R. Ade Safitri, E. Panca Saputra, and S. Nusa Mandiri Jakarta, "Pengelompokan Perilaku Mahasiswa Pada Perkuliahan E-Learning dengan K-Means Clustering," *Generation Journal*, vol. 19, no. 2, pp. 81-92, 2019.
- [4] E. A. Saputra and Y. Nataliani, "Analisis Pengelompokan Data Nilai Siswa untuk Menentukan Siswa Berprestasi Menggunakan Metode Clustering K-Means," *Journal of Information Systems and Informatics*, vol. 3, no. 3, pp. 429-439, 2021, [Online]. Available: <http://journal-isi.org/index.php/isi>
- [5] J. Hutagalung, Y. Hendro Syahputra, Z. Pertiwi Tanjung, S. Triguna Dharma, and J. I. Pintu Air, "Pemetaan Siswa Kelas Unggulan Menggunakan Algoritma K-Means Clustering," *Hal AH Nasution*, vol. 9, no. 1, 2022, [Online]. Available: <http://jurnal.mdp.ac.id>
- [6] I. T. Umagapi, et al, "Uji Kinerja K-Means Clustering Menggunakan Davies-Bouldin Index Pada Pengelompokan Data Prestasi Siswa," *Prosiding Seminar Nasional Sistem Informasi dan Teknologi (SISFOTEK) VIII*, vol. 7, no. 1, 2023, pp. 303-308.
- [7] I. R. Mahartika and A. Wibowo, "Data Mining Klasterisasi dengan Algoritme K-Means untuk Pengelompokan Provinsi Berdasarkan Konsumsi Bahan Bakar Minyak Nasional," *Prosiding Seminar Nasional Sistem Informasi dan Teknologi (SISFOTEK)*, 2019, vol. 3, no. 1, pp. 87-91.

- [8] S. D. G. W. N. Weri Sirait, “Algoritma K-Means Untuk Klasterisasi Tugas Akhir Mahasiswa Berdasarkan Keahlian,” *Jurnal Sistem Infromasi dan Teknologi*, vol. 1, no. 3, pp. 24-29, 2019.
- [9] M. K. Anam, B. Burhanuddin, S. Hadisaputra, and M. Muntari, “Analisis Hubungan Antara Sikap dan Prestasi Belajar Kuliah Online pada Mata Kuliah Pemisahan Analitik Mahasiswa Pendidikan Kimia FKIP Universitas Mataram,” *Chemistry Education Practice*, vol. 6, no. 1, pp. 52–56, 2023, doi: 10.29303/cep.v6i1.2927.
- [10] S. Kurniawan, A. M. Siregar, and H. Y. Novita, “Penerapan Algoritma K-Means dan Fuzzy C-Means Dalam Mengelompokan Prestasi Siswa Berdasarkan Nilai Akademik,” *Scientific Student Journal for Information, Technology and Science*, vol. 4, no. 1, pp.73-81, 2023.