

Klusterisasi Prestasi Siswa Menggunakan Algoritma K-Means Pada UPTD SD Negeri Kecil Lembah Harapan Mamuju Tengah

Yokebet^{1,*}, Eka Berby², Dikwan Moeis³

^{1*}, ² Sistem Informasi, STMIK Profesional Makassar

³ Ilmu Komputer, STMIK Profesional Makassar

¹yokhebetartatitandi039@gmail.com

Abstract

This study aims to assist SD Negeri Mamuju Tengah in building student clustering based on merit to monitor and improve student achievement. Student achievement clustering is one approach that can be used to group students based on their achievement. This study uses indicators of academic scores and attitude values in order to combine intellectual abilities and characteristics for each student. It is known that the achievement of SD Negeri Mamuju Tengah students is formed in 3 clusters, namely a cluster consisting of 5 students, a medium cluster consisting of 2 students, while a cluster consisting of less than 5 students. At SD Negeri Mamuju Tengah, special attention is still needed related to improving achievement. The suggestion in the next study is to use datasets from primary sources in larger quantities, and use hybrid algorithms in order to obtain better results.

Keyword: achievement, K-means, clustering, students.

Abstrak

Penelitian ini bertujuan untuk membantu pihak SD Negeri Mamuju Tengah dalam membangun klusterisasi siswa berdasarkan prestasi untuk memantau dan meningkatkan prestasi siswa. Klusterisasi prestasi siswa merupakan salah satu pendekatan yang dapat digunakan untuk mengelompokkan siswa berdasarkan prestasi mereka. Penelitian ini menggunakan indikator nilai akademik dan nilai sikap agar dapat dipadukan kemampuan intelektual dan karakteristik terhadap setiap siswa. Diketahui bahwa prestasi siswa SD negeri mamuju tengah terbentuk dalam 3 kluster yaitu kluster banyak terdiri dari 5 siswa, kluster sedang terdiri dari 2 siswa, sedangkan kluster kurang terdiri dari 5 siswa. Pada SD Negeri Mamuju Tengah masih dibutuhkan perhatian khusus terkait peningkatan prestasi. Saran pada penelitian selanjutnya adalah menggunakan dataset dari sumber utama dalam jumlah yang lebih besar, serta menggunakan hybrid algoritma agar dapat diperoleh hasil yang lebih baik.

Kata kunci: prestasi, k-means, klusterisasi, siswa.

PENDAHULUAN

Prestasi siswa di tingkat Sekolah Dasar (SD) merupakan indikator penting dalam mengevaluasi kualitas pendidikan. Namun, setiap siswa memiliki keunikan dan perbedaan dalam prestasinya. Mengidentifikasi pola dan karakteristik dalam prestasi siswa dapat memberikan wawasan yang berharga bagi pendidik dan pihak terkait dalam mengembangkan strategi pembelajaran yang lebih efektif.

Klusterisasi prestasi siswa merupakan salah satu pendekatan yang dapat digunakan untuk mengelompokkan siswa berdasarkan prestasi mereka. Dengan menggunakan algoritma klusterisasi seperti K-means, siswa dapat dikelompokkan dalam kelompok prestasi yang serupa. Hal ini dapat membantu pendidik untuk menyusun strategi pembelajaran yang lebih spesifik dan individualisasi pendekatan pembelajaran.

Namun, hingga saat ini, penelitian mengenai klusterisasi prestasi siswa SD masih terbatas, beberapa penelitian yang telah dilakukan cenderung berfokus pada aspek tertentu, seperti nilai akademik. Penelitian oleh dengan judul

Pengelompokan Hasil Belajar Siswa Dengan Metode *Klusterisasi* K-Means, bertujuan untuk mengetahui keefektifan *klusterisasi* dalam mengelompokan siswa dan meningkatkan hasil belajar siswa (Anwar, 2022). Terdapat kebutuhan yang mendesak untuk lebih memahami variasi prestasi siswa secara komprehensif dan mengidentifikasi faktor-faktor yang mempengaruhi prestasi mereka. Oleh karena itu pada penelitian ini digunakan indikator nilai akademik dan nilai sikap sehingga dapat dipadukan kemampuan intelektual dan karakteristik terhadap setiap siswa. Selanjutnya penelitian berjudul penentuan siswa berprestasi menggunakan metode k-means klusterisasi di smp takhassus al qur'an sadamiyyah (Ismaya et al., 2022), dilakukan untuk menentukan rekomendasi siswa berprestasi untuk menyelesaikan masalah ketepatan waktu sekolah.

Metode K-Means bertujuan mengelompokkan data yang ada kedalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain (Putra et al., 2021). Metode data mining dalam menangani data *outlier* adalah klusterisasi, dimana klusterisasi mendefinisikan *outlier* yang tidak ada di kluster (Nurmawati et al., 2022). Data *mining* menganalisa sejumlah besar kumpulan data observasi, menemukan suatu hubungan tidak terduga serta dapat merangkum data dengan cara yang baru sehingga dapat berguna dan dimengerti bagi pengguna (Gustientiedina et al., 2019).

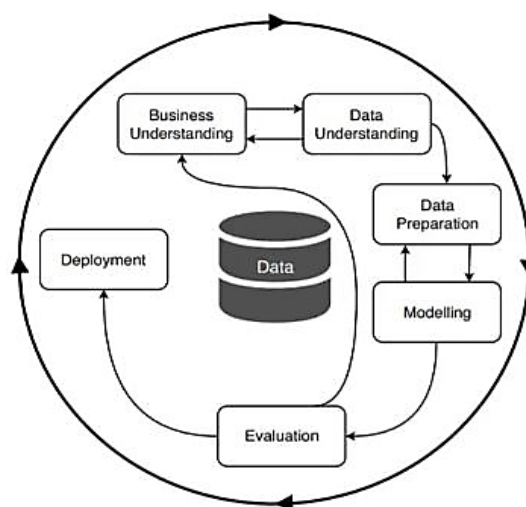
Dalam konteks tersebut, penelitian ini bertujuan untuk melakukan klusterisasi prestasi siswa SD menggunakan algoritma K-Means. Penelitian ini diharapkan dapat memberikan wawasan tentang kelompok prestasi siswa yang serupa, serta mengidentifikasi pola dan karakteristik yang membedakan antar siswa. Hasil penelitian ini diharapkan dapat memberikan kontribusi bagi pengembangan strategi pembelajaran yang lebih adaptif dan efektif di tingkat SD.

METODE PENELITIAN

Pada bagian ini menjelaskan prosedur sistematis yang melibatkan langkah-langkah yang terorganisir dan terstruktur untuk memperoleh jawaban terhadap pertanyaan penelitian atau mencapai tujuan penelitian yang telah ditetapkan.

Data Mining

Data *mining* adalah mengekstraksi informasi akurat yang diinginkan dan bermakna dari kumpulan data melalui metode dan algoritma untuk membantu proses pengelompokan data (Durugkar et al., 2022). Data *mining* dapat diterapkan untuk menemukan pola pengetahuan dari profil pasien dan data riwayat kesehatannya, dimana Pengetahuan yang diperoleh dapat digunakan analisis dan pengambilan keputusan, antara lain untuk memprediksi jenis penyakit, mengetahui pola penyebaran penyakit, dan melihat efektifitas pengobatan. Gambar 1.



Gambar 1. Proses model data *mining* (Martinez-Plumed et al., 2021)

Pada gambar 1 ditampilkan tahapan pada proses data *mining* dimana tahapan dalam data mining ini dapat berulang dan iteratif, di mana setiap tahap dapat disesuaikan atau diperbaiki berdasarkan hasil evaluasi dan kebutuhan yang muncul selama proses.

Metode K-Means

Clustering adalah salah satu metode penambangan data yang mempartisi data berukuran besar menjadi subkelompok sesuai dengan kesamaannya (Ay et al., 2023). Salah satu prosedur *klusterisasi* pada *non-hirarki* adalah dengan menggunakan metode K-Means yang merupakan metode *klusterisasi* yang paling terkenal dan banyak digunakan di berbagai bidang karena sederhana dan mudah diimplementasikan (Chu et al., 2020). Metode K-Means merupakan metode pengelompokan yang bertujuan mengelompokkan obyek sedemikian sehingga jarak tiap-tiap obyek ke pusat kelompok di dalam satu kelompok adalah minimum. Dasar algoritma K-Means adalah sebagai berikut :

- a. Diberikan nilai k sebagai jumlah kluster yang ingin dibentuk
- b. Bangkitkan k centroid (titik pusat kluster) awal secara random
- c. Hitung jarak setiap data ke masing-masing pusat kluster yaitu menggunakan *Euclidean Distance*, dimana jarak *Euclidean* adalah besarnya jarak suatu garis lurus yang menghubungkan antar objek diukur dengan rumus:

$$d_{(x,y)} = \sqrt{\sum_{j=1}^p (x_j - y_j)^2} \quad (1)$$

- d. Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan pusatnya.
- e. Tentukan posisi pusat kluster baru (C_{kj}) dengan cara menghitung nilai rata-rata dari data-data yang ada pada pusat kluster yang sama

$$C_{kj} = \frac{x_{1j} + x_{2j} + \dots + x_{aj}}{a}, j = 1, 2, 3, \quad (2)$$

C_{kj} = pusat kluster ke-k pada variabel ke- j

a = banyak data pada kluster ke-k

- f. Dalam mendeskripsikan karakteristik masing-masing kluster, digunakan persamaan berikut :

$$X = \mu + Z \cdot \sigma \quad (3)$$

Dimana:

X = rata-rata sampel (variabel dalam kluster)

μ = rata-rata populasi

Z = nilai standarisasi

σ = standar deviasi

Metode K-Means merupakan algoritma tradisional yang terdiri dari sejumlah besar data berdimensi tinggi dalam memilih pusat pengelompokan awal dan ketidakpastian nilai K (Zhao, 2021).

HASIL DAN PEMBAHASAN

Penelitian ini menggunakan data siswa yang bersekolah pada SD Negeri Mamuju Tengah pada tahun ajaran 2021-2022. Algoritma pengelompokan Kmeans banyak digunakan untuk keunggulan kesederhanaan dan operasi yang efisien (Gao et al., 2022). Penerapan metode K-Means dijelaskan berdasarkan tahapan yang telah ditentukan.

1. Penentuan jumlah kluster
Jumlah kluster yang ditetapkan adalah 3 sebagai representasi tingkat prestasi tinggi, sedang, rendah.
2. Penentuan Jumlah *dataset*
Jumlah *dataset* yang ditetapkan sebanyak 100, yang mewakili data tahun ajaran 2021-2022. *Dataset* ditampilkan

Pada tabel 1.

Tabel 1. *Dataset* Nilai Siswa

NO	SISWA	NILAI RAPOR	NILAI SIKAP
1	ABDULLAH	80	89
2	ANDI ARIANSYAH	84	90
3	ARFAN SETIAWAN	86	92
4	ARHAM	84	89
5	ASMAUL HUSNA. R	91	92
6	CITRA	85	92
7	DIANA PUTRI	85	90
8	DIKA	75	90
9	FIKRAM	88	92
10	HYIRUDDIN	77	89
11	IRA AMALIYA	81	89
12	JULITA KAMIL	85	90
13	MUTMAINNA	88	84
14	AMIRUDDIN	82	85
15	NURJANNAH	83	85
16	ALIF	88	83
17	MAHARANI	87	83
18	MAHMUDIN	83	82
19	JOKOWI	87	90
20	IMAM NURDIN	88	86
21	AXEL	81	89
22	BERNADETTA SETIAWAN	89	90
23	BOBY APRIADI	82	83
24	CHERYL MESA BARLIANTO	88	89
25	DONNY DOZEN	83	92
26	DUFAN	85	89
27	ERIKA VERIANA	87	83
28	FELICIANA WIDYANATA	87	87
29	FRENDY ANTONIUS	82	87
30	HANSEN	86	89
31	HENRY MALIANTO	80	88
32	HERLIAN STELLA	90	92
33	JESSICA SHAN	86	84
34	JULIUS	82	89
35	KEVIN PERMANA	90	92
36	LENY MARSELINA	90	84
37	MELY FILADELVIA	90	87
38	MERRY CRISTINE	84	91
39	MICHAEL WIJAYA	82	84
40	NOVIA APRILIA	84	84
41	NOVITA CHEN CHEN	84	92
42	PANGERAN YOEL H.SIMORANGKIR	87	87
43	RICHI AMRICSON	90	84
44	RIMIKA CHANIA	80	92
45	ROSALINA	83	90
46	RUYANDI	83	92
47	SELVIANA	90	92
48	SHARON MULAN	80	83
49	SHEN KUANG	77	89
50	SISILIA PUTRI SANJAYA	84	89
51	STEVEN	87	90
52	VENNY CHRISTINE	78	84
53	WONGSO IRVONI	87	82

54	YOUKHARI PUTRI CENJAYA	90	85
55	YULIANTO	85	87
56	ADELLIA EKKA PUTRI	83	92
57	ANDREAS TRIMURNI	85	91
58	CHEN ANDI	85	87
59	CHRISTIAN	85	86
60	CLAVENIA ARIVIAN TI	84	90
61	CRISTINE	77	89
62	DANIEL	88	89
63	ERICKSON NGUILANDA	83	84
64	ERIKA LIU	81	91
65	ERWIN FERNANDO	89	84
66	FELICYA CINDY L	87	91
67	FENA CYNTHIA	88	89
68	HARJIUS	84	82
69	JULIUS OSCAR	86	91
70	KEVIN FAM	87	84
71	LEONARDO	88	89
72	LIE SIAH	80	89
73	LUSIANA	87	83
74	MELISSA SANTOSO	81	83
75	MICHAEL WONG	77	87
76	MICO WIJAYA	84	90
77	NOSIA ROSA MELTA	90	90
78	OKTAVIANUS	83	83
79	RAKA PATRICIA	87	87
80	REYNOLD ANDIKA	82	84
81	RICKY CHANG	83	88
82	ROSANNIE PRISCYLIA	87	86
83	SHINTIA TAMARA	88	85
84	STEVEN SUTANTO	86	83
85	SUPHIN	86	86
86	TJHAI PO PHEN	81	82
87	VALLIAN JIUVAN	76	87
88	VANESSA WINATA	83	86
89	VELIA VALENTINA	80	88
90	VICENNI	81	88
91	AGREGORIUS AGUNG	86	88
92	AILEEN KRISTAL	89	83
93	ALFON	79	87
94	ANGEL	89	91
95	ARIES	88	83
96	BELVA CHRISTOPHER	89	88
97	CHANDRA GUNAWAN	78	90
98	CUNG LIM KONG	88	84
99	DEDDY	80	86
100	DESSI	82	90

3. Penetapan *attribute*

Jumlah atribut yang digunakan sebanyak 2, terdiri dari Nilai Rapor dan Nilai Sikap.

4. Penentuan pusat kluster secara acak

Pada tahap awal ditentukan pusat kluster pada

Data ke-8 sebagai pusat kluster ke-1 terdiri dari nilai 75 dan 90.

Data ke-7 sebagai pusat kluster ke-2 terdiri dari nilai 85 dan 90.

Data ke-2 sebagai pusat kluster ke-3 terdiri dari 84 dan 90.

5. Perhitungan jarak pusat

Proses perhitungan jarak pusat kluster menggunakan rumus 1. Hasil perhitungan ditampilkan pada tabel 2.

Tabel 2. Perhitungan Jarak Pusat Kluster

NO	SISWA	NILAI RAPOR	NILAI SIKAP	C1	C2	C3	JARAK TERDEKAT
1	ABDULLAH	80	89	4.476753546	5.277427	3.771357911	3.771357911
2	ANDI	84	90	0.727272727	8.818182	0	0
3	ARFAN	86	92	2.472994638	11.18034	2.959785563	2.472994638
4	ARHAM	84	89	1.139087644	9.055385	1.016394535	1.016394535
5	ASMAUL HUSNA.	91	92	7.19274198	16.57565	7.893924853	7.19274198
6	CITRA	85	92	2.098799342	10.37639	2.420641265	2.098799342
7	DIANA PUTRI	85	90	0	9.545455	0.727272727	0
8	DIKA	75	90	9.545454545	0	8.818181818	0
9	FIKRAM	88	92	4.070656941	13.24281	4.717647544	4.070656941
10	HYRUDDIN	77	89	7.521319561	2.317736	6.801190951	2.317736143
11	IRA AMALIYA	81	89	3.422096404	6.351937	2.734837992	2.734837992
12	JULITA KAMIL	85	90	0.818181818	10.36364	1.545454545	0.818181818
13	MUTMAINNA	88	84	6.923430096	14.31782	7.313521949	6.923430096
14	AMIRUDDIN	82	85	5.610645136	8.602325	5.320318141	5.320318141
15	NURJANNAH	83	85	5.233395623	9.433981	5.066499925	5.066499925
16	ALIF	88	83	7.806015904	14.76482	8.15399309	7.806015904
17	MAHARANI	87	83	7.417869869	13.89244	7.689211074	7.417869869
18	MAHMUDIN	83	82	8.147909533	11.31371	8.041730006	8.041730006
19	JOKOWI	87	90	2.454545455	12	3.181818182	2.454545455
20	IMAM NURDIN	88	86	5.285251583	13.60147	5.786847441	5.285251583
21	AXEL	81	89	3.683781744	6.082763	2.990342582	2.990342582
22	BERNADETTA	89	90	4.454545455	14	5.181818182	4.454545455
23	BOBY APRIADI	82	83	7.448445398	9.899495	7.232273856	7.232273856
24	CHERYL MESA	88	89	3.59637099	13.0384	4.299721306	3.59637099
25	DONNY DOZEN	83	92	2.527534323	8.246211	2.160884423	2.160884423
26	DUFAN	85	89	1.098458725	10.04988	1.548126033	1.098458725
27	ERIKA VERIANA	87	83	7.417869869	13.89244	7.689211074	7.417869869
28	FELICIANA	87	87	3.876182837	12.36932	4.373095808	3.876182837
29	FRENDY	82	87	3.934379092	7.615773	3.507960251	3.507960251
30	HANSEN	86	89	1.765135258	11.04536	2.40006887	1.765135258
31	HENRY	80	88	4.966000103	5.385165	4.310279851	4.310279851
32	HERLIAN	90	92	5.80965284	15.13275	6.497297595	5.80965284
33	JESSICA SHAN	86	84	6.17379158	12.52996	6.384381769	6.17379158
34	JULIUS	82	89	2.734837992	7.071068	2.075038584	2.075038584
35	KEVIN	90	92	5.80965284	15.13275	6.497297595	5.80965284
36	LENY	90	84	8.108764771	16.15549	8.614805629	8.108764771
37	MELY	90	87	6.225115751	15.29706	6.871308175	6.225115751
38	MERRY	84	91	1.139087644	9.055385	1.016394535	1.016394535
39	MICHAEL	82	84	6.517617574	9.219544	6.2694326	6.2694326
40	NOVIA APRILIA	84	84	6.024742373	10.81665	6.002754189	6.002754189
41	NOVITA CHEN	84	92	2.073046227	9.219544	2.008247458	2.008247458
42	PANGERAN	87	87	3.876182837	12.36932	4.373095808	3.876182837
43	RICHI	90	84	8.108764771	16.15549	8.614805629	8.108764771
44	RIMIKA CHANIA	80	92	4.966000103	5.385165	4.310279851	4.310279851
45	ROSALINA	83	90	1.545454545	8	0.818181818	0.818181818
46	RUYANDI	83	92	2.527534323	8.246211	2.160884423	2.160884423
47	SELVIANA	90	92	5.80965284	15.13275	6.497297595	5.80965284
48	SHARON MULAN	80	83	8.34632596	8.602325	7.973613509	7.973613509
49	SHEN KUANG	77	89	7.61143116	2.236068	6.891124967	2.236067977
50	SISILIA PUTRI	84	89	1.139087644	9.055385	1.016394535	1.016394535
51	STEVEN	87	90	2.454545455	12	3.181818182	2.454545455
52	VENNY	78	84	8.879356689	6.708204	8.357705407	6.708203932
53	WONGSO	87	82	8.368081823	14.42221	8.609527684	8.368081823
54	YOUKHARI	90	85	7.399463907	15.81139	7.950778329	7.399463907

55	YULIANTO	85	87	3.034239867	10.44031	3.224390518	3.034239867
56	ADELLIA EKKKA	83	92	2.527534323	8.246211	2.160884423	2.160884423
57	ANDREAS	85	91	1.098458725	10.04988	1.548126033	1.098458725
58	CHEN ANDI	85	87	3.034239867	10.44031	3.224390518	3.034239867
59	CHRISTIAN	85	86	4.025743605	10.77033	4.170934453	4.025743605
60	CLAVENIA	84	90	0.545454545	9	0.181818182	0.181818182
61	CRISTINE	77	89	7.61143116	2.236068	6.891124967	2.236067977
62	DANIEL	88	89	3.59637099	13.0384	4.299721306	3.59637099
63	ERICKSON	83	84	6.195839713	10	6.055528176	6.055528176
64	ERIKA LIU	81	91	3.683781744	6.082763	2.990342582	2.990342582
65	ERWIN	89	84	7.472815748	15.23155	7.927877375	7.472815748
66	FELICYA CINDY	87	91	2.650432679	12.04159	3.335261151	2.650432679
67	FENA CYNTHIA	88	89	3.59637099	13.0384	4.299721306	3.59637099
68	HARJIUS	84	82	8.01857348	12.04159	8.002065849	8.002065849
69	JULIUS OSCAR	86	91	1.765135258	11.04536	2.40006887	1.765135258
70	KEVIN FAM	87	84	6.482653268	13.41641	6.791462798	6.482653268
71	LEONARDO	88	89	3.59637099	13.0384	4.299721306	3.59637099
72	LIE SIAH	80	89	4.654154813	5.09902	3.946962427	3.946962427
73	LUSIANA	87	83	7.417869869	13.89244	7.689211074	7.417869869
74	MELISSA	81	83	7.846671137	9.219544	7.546002171	7.546002171
75	MICHAEL WONG	77	87	8.119968245	3.605551	7.449000155	3.605551275
76	MICO WIJAYA	84	90	0.545454545	9	0.181818182	0.181818182
77	NOSIA ROSA	90	90	5.454545455	15	6.181818182	5.454545455
78	OKTAVIANUS	83	83	7.168572365	10.63015	7.047653616	7.047653616
79	RAKA PATRICIA	87	87	3.876182837	12.36932	4.373095808	3.876182837
80	REYNOLD	82	84	6.517617574	9.219544	6.2694326	6.2694326
81	RICKY CHANG	83	88	2.527534323	8.246211	2.160884423	2.160884423
82	ROSANNIE	87	86	4.693058	12.64911	5.111161017	4.693058
83	SHINTIA	88	85	6.077325423	13.92839	6.518251553	6.077325423
84	STEVEN	86	83	7.149524633	13.0384	7.332143655	7.149524633
85	SUPHIN	86	86	4.256254513	11.7047	4.556350577	4.256254513
86	TJHAI PO PHEN	81	82	8.750442728	10	8.481871772	8.481871772
87	VALLIAN	76	87	9.056754021	3.162278	8.37400543	3.16227766
88	VANESSA	83	86	4.288173242	8.944272	4.082820286	4.082820286
89	VELIA	80	88	4.966000103	5.385165	4.310279851	4.310279851
90	VICENNI	81	88	4.070656941	6.324555	3.45574142	3.45574142
91	AGREGORIUS	86	88	2.472994638	11.18034	2.959785563	2.472994638
92	AILEEN KRISTAL	89	83	8.297166698	15.65248	8.709261718	8.297166698
93	ALFON	79	87	6.304923958	5	5.67581501	5
94	ANGEL	89	91	4.565410738	14.03567	5.277427372	4.565410738
95	ARIES	88	83	7.806015904	14.76482	8.15399309	7.806015904
96	BELVA	89	88	4.88292691	14.14214	5.554389226	4.88292691
97	CHANDRA	78	90	6.545454545	3	5.818181818	3
98	CUNG LIM	88	84	6.923430096	14.31782	7.313521949	6.923430096
99	DEDDY	80	86	6.054845747	6.403124	5.529784118	5.529784118
100	DESSI	82	90	2.545454545	7	1.818181818	1.818181818

Pada tabel 2. Ditampilkan nilai jarak terdekat dari pusat kluster, ditulis dalam contoh perolehan nilai pada C3:

$$3.77 = \sqrt{((80 - 86.8)^2) + (84 - 90)^2}$$

Seterusnya dilakukan terhadap setiap *dataset*.

6. Pengelompokan data.

K-berarti dapat menghilangkan sebanyak mungkin data noise untuk meminimalkan inferensi pada pengelompokan sementara serta mampu mencapai pengelompokan optimal lokal yang lebih baik(Wei et al., 2020). Pengelompokan ditentukan berdasarkan jarak terdekat dari pusat kluster menggunakan rumus 2. Pada iterasi pertama diperoleh hasil pengelompokan seperti pada tabel 3.

Tabel 3. Hasil Pengelompokan data Iterasi ke-1.

No.	<i>Custer-1</i>	<i>Custer-2</i>	<i>Custer-3</i>
1			1
2			1
3	1		
4			1
5	1		
6	1		
7	1		
8		1	
9	1		
10		1	
11			1
12	1		
13	1		
14			1
15			1
16	1		
17	1		
18			1
19	1		
20	1		
21			1
22	1		
23			1
24	1		
25			1
26	1		
27	1		
28	1		
29			1
30	1		
31			1
32	1		
33	1		
34			1
35	1		
36	1		
37	1		
38			1
39			1
40			1
41			1
42	1		
43	1		
44			1
45			1
46			1
47	1		
48			1
49		1	
50			1
51	1		
52		1	
53	1		
54	1		
55	1		
56			1

57	1	
58	1	
59	1	
60		1
61		1
62	1	
63		1
64		1
65	1	
66	1	
67	1	
68		1
69	1	
70	1	
71	1	
72		1
73	1	
74		1
75		1
76		1
77	1	
78		1
79	1	
80		1
81		1
82	1	
83	1	
84	1	
85	1	
86		1
87		1
88		1
89		1
90		1
91	1	
92	1	
93		1
94	1	
95	1	
96	1	
97		1
98	1	
99		1
100		1

Pada Tabel 3, ditampilkan *cell* yang berisi nilai 1 merupakan kluster terhadap *dataset*.

7. Penentuan Kluster Baru

Algoritma Kmeans adalah algoritma yang banyak digunakan untuk segmentasi pasar, biasanya nilai k dari algoritma Kmeans dipilih secara acak(Durojaye & Obunadike, 2022). Pada Iterasi kedua dan seterusnya dilakukan pembaruan pusat kluster agar setiap *dataset* dapat lebih terseleksi nilai kedekatannya dengan pusat kluster menggunakan rumus 3. Data pada pusat kluster baru pada iterasi kedua ditampilkan pada Tabel 4.

Tabel 4. Pusat kluster baru

Kluster Prestasi Tinggi	86.8	76.0
Kluster Prestasi Sedang	91.3	90.5
Kluster Prestasi Rendah	86.8	76.0

Kmeans adalah salah satu algoritma partisi sederhana dan efisien yang digunakan di banyak solusi cluster (Abdulnassar & Nair, 2023). Berdasarkan tabel 4, jelaskan nilai pusat kluster baru diperoleh berdasarkan total nilai *dataset* pada setiap kluster lalu dibagi jumlah *dataset* terpilih., ditulis pada contoh :

$$86.8 = (86 + 91 + 85 + 85 + 88 + 85) / 6$$

$$76.0 = (92 + 92 + 92 + 90 + 92 + 90) / 6$$

Maka diperoleh hasil perhitungan jarak terdekat dari pusat kluster pada tabel 5.

Tabel 5. Hasil Pengelompokan data Iterasi ke-2

NO	SISWA	NILAI RAPOR	NILAI SIKAP	C1	C2	C3	Jarak Terpendek
1	ABDULLAH	80	89	6.99	4.40	2.15	2.15
2	ANDI ARIANSYAH	84	90	3.24	7.79	1.68	1.68
3	ARFAN SETIAWAN	86	92	1.02	10.07	4.60	1.02
4	ARHAM	84	89	3.62	8.09	1.70	1.70
5	ASMAUL HUSNA. R	91	92	4.73	15.48	9.54	4.73
6	CITRA	85	92	1.72	9.26	3.97	1.72
7	DIANA PUTRI	85	90	2.60	8.51	2.35	2.35
8	DIKA	75	90	11.85	1.16	7.36	1.16
9	FIKRAM	88	92	1.48	12.14	6.39	1.48
10	HYIRUDDIN	77	89	9.96	1.83	5.23	1.83
11	IRA AMALIYA	81	89	5.97	5.44	1.07	1.07
12	JULITA KAMIL	85	90	1.94	9.33	3.14	1.94
13	MUTMAINNA	88	84	7.44	13.61	7.74	7.44
14	AMIRUDDIN	82	85	7.93	8.11	4.26	4.26
15	NURJANNAH	83	85	7.37	8.87	4.30	4.30
16	ALIF	88	83	8.42	14.11	8.45	8.42
17	MAHARANI	87	83	8.34	13.28	7.81	7.81
18	MAHMUDIN	83	82	10.07	10.98	7.28	7.28
19	JOKOWI	87	90	1.35	10.97	4.74	1.35
20	IMAM NURDIN	88	86	5.47	12.77	6.55	5.47
21	Axel	81	89	6.23	5.18	1.34	1.34
22	Bernadetta Setiawan	89	90	2.60	12.96	6.72	2.60
23	Boby Apriadi	82	83	9.60	9.58	6.26	6.26
24	Cheryl Mesa Barlianto	88	89	2.64	12.05	5.69	2.64
25	Donny Dozen	83	92	3.83	7.11	2.83	2.83
26	Dufan	85	89	2.93	9.08	2.69	2.69
27	Erika Veriana	87	83	8.34	13.28	7.81	7.81
28	Feliciana Widyana	87	87	4.34	11.50	5.19	4.34
29	Frendy Antonius	82	87	6.45	6.91	2.27	2.27
30	Hansen	86	89	2.46	10.07	3.69	2.46
31	Henry Malianto	80	88	7.55	4.68	2.63	2.63
32	Herlian Stella	90	92	3.30	14.03	8.16	3.30
33	Jessica Shan	86	84	7.37	11.89	6.41	6.41
34	Julius	82	89	5.31	6.14	0.40	0.40
35	Kevin Permana	90	92	3.30	14.03	8.16	3.30
36	Leny Marselina	90	84	8.01	15.39	9.30	8.01
37	Mely Filadelfia	90	87	5.40	14.39	8.00	5.40
38	Merry Cristine	84	91	2.79	7.97	2.43	2.43
39	Michael Wijaya	82	84	8.75	8.82	5.26	5.26

40	Novia Aprilia	84	84	7.84	10.27	5.51	5.51
41	Novita Chen Chen	84	92	2.85	8.09	3.22	2.85
42	Pangeran Yoel	87	87	4.34	11.50	5.19	4.34
43	Richi Amricson	90	84	8.01	15.39	9.30	8.01
44	Rimika Chania	80	92	6.81	4.23	3.60	3.60
45	Rosalina	83	90	4.00	6.97	1.01	1.01
46	Ruyandi	83	92	3.83	7.11	2.83	2.83
47	Selviana	90	92	3.30	14.03	8.16	3.30
48	Sharon Mulan	80	83	10.74	8.48	6.67	6.67
49	Shen Kuang	77	89	10.05	1.78	5.32	1.78
50	Sisilia Putri Sanjaya	84	89	3.62	8.09	1.70	1.70
51	Steven	87	90	1.35	10.97	4.74	1.35
52	Venny Christine	78	84	11.43	6.79	6.80	6.79
53	Wongso Irvoni	87	82	9.34	13.87	8.63	8.63
54	Youkhari Putri Cenjaya	90	85	7.11	15.00	8.78	7.11
55	Yulianto	85	87	4.68	9.61	3.50	3.50
56	Adellia Ekka Putri	83	92	3.83	7.11	2.83	2.83
57	Andreas Trimurni	85	91	1.80	8.97	3.20	1.80
58	Chen Andi	85	87	4.68	9.61	3.50	3.50
59	Christian	85	86	5.62	10.02	4.21	4.21
60	Clavenia Arivianti	84	90	3.08	7.97	1.84	1.84
61	Cristine	77	89	10.05	1.78	5.32	1.78
62	Daniel	88	89	2.64	12.05	5.69	2.64
63	Erickson Ngulanda	83	84	8.25	9.52	5.29	5.29
64	Erika Liu	81	91	5.78	4.98	2.19	2.19
65	Erwin Fernando	89	84	7.66	14.49	8.50	7.66
66	Felicya Cindy L	87	91	0.40	10.97	5.00	0.40
67	Fena Cynthia	88	89	2.64	12.05	5.69	2.64
68	Harjius	84	82	9.74	11.64	7.44	7.44
69	Julius Oscar	86	91	0.84	9.97	4.08	0.84
70	Kevin Fam	87	84	7.34	12.74	7.03	7.03
71	Leonardo	88	89	2.64	12.05	5.69	2.64
72	Lie Siah	80	89	7.16	4.23	2.33	2.33
73	Lusiana	87	83	8.34	13.28	7.81	7.81
74	Melissa Santoso	81	83	10.14	8.99	6.39	6.39
75	Michael Wong	77	87	10.69	3.63	5.77	3.63
76	Mico Wijaya	84	90	3.08	7.97	1.84	1.84
77	Nosia Rosa Melta	90	90	3.49	13.96	7.72	3.49
78	Oktavianus	83	83	9.15	10.23	6.29	6.29
79	Raka Patricia	87	87	4.34	11.50	5.19	4.34
80	Reynold Andika	82	84	8.75	8.82	5.26	5.26
81	Ricky Chang	83	88	5.03	7.39	1.42	1.42
82	Rosannie Priscylia	87	86	5.34	11.84	5.70	5.34
83	Shintia Tamara	88	85	6.45	13.16	7.10	6.45
84	Steven Sutanto	86	83	8.37	12.46	7.25	7.25
85	Suphin	86	86	5.39	10.92	4.91	4.91
86	Tjhai Po Phen	81	82	10.97	9.84	7.37	7.37
87	Vallian Jiuvan	76	87	11.61	3.50	6.71	3.50
88	Vanessa Winata	83	86	6.53	8.28	3.32	3.32
89	Velia Valentina	80	88	7.55	4.68	2.63	2.63
90	Vicenni	81	88	6.67	5.55	1.82	1.82
91	Agregorius Agung	86	88	3.42	10.26	3.89	3.42
92	Aileen Kristal	89	83	8.63	14.97	9.15	8.63

93	Alfon	79	87	8.90	4.58	4.01	4.01
94	Angel	89	91	2.25	12.96	6.91	2.25
95	Aries	88	83	8.42	14.11	8.45	8.42
96	Belva Christopher	89	88	4.01	13.19	6.80	4.01
97	Chandra Gunawan	78	90	8.87	2.02	4.38	2.02
98	Cung Lim Kong	88	84	7.44	13.61	7.74	7.44
99	Deddy	80	86	8.62	5.99	3.99	3.99
100	Dessi	82	90	4.96	5.98	0.81	0.81

Berdasarkan data pada tabel 5, maka diperoleh hasil pengelompokan baru yang ditampilkan pada tabel 6.

Tabel 6. Hasil Pengelompokan data Iterasi ke-2

NO.	CUSTER -1	CUSTER-2	CUSTER-3
1			1
2			1
3	1		
4			1
5	1		
6	1		
7			1
8		1	
9	1		
10		1	
11			1
12	1		
13	1		
14			1
15			1
16	1		
17			1
18			1
19	1		
20	1		
21			1
22	1		
23			1
24	1		
25			1
26			1
27			1
28	1		
29			1
30	1		
31			1
32	1		
33			1
34			1
35	1		
36	1		
37	1		
38			1
39			1
40			1
41	1		

42	1	
43	1	
44		1
45		1
46		1
47	1	
48		1
49		1
50		1
51	1	
52		1
53		1
54	1	
55		1
56		1
57	1	
58		1
59		1
60		1
61		1
62	1	
63		1
64		1
65	1	
66	1	
67	1	
68		1
69	1	
70		1
71	1	
72		1
73		1
74		1
75		1
76		1
77	1	
78		1
79	1	
80		1
81		1
82	1	
83	1	
84		1
85		1
86		1
87		1
88		1
89		1
90		1
91	1	
92	1	
93		1
94	1	
95	1	
96	1	
97		1
98	1	

99	1
100	1

Proses iterasi dilakukan hingga hasil kluster tidak berubah. Pada penelitian ini proses iterasi hanya dilakukan sebanyak 2 kali, karena pada iterasi ke 3 tidak terjadi perubahan hasil kluster. Sebagaimana dikatakan pada penelitian bahwa Pengelompokan Kmeans adalah metode pengelompokan yang populer untuk mengelompokkan titik data yang serupa berdasarkan kesamaannya (Huang et al., 2021).

Berdasarkan data pengelompokan pada iterasi terakhir diketahui bahwa prestasi siswa SD negeri mamuju tengah terbentuk dalam 3 kluster yaitu Prestasi Tinggi terdiri dari 40 siswa, Prestasi Sedang terdiri dari 8 siswa, sedangkan Prestasi Kurang terdiri dari 52 siswa.

KESIMPULAN

Kesimpulan yang diperoleh pada penelitian ini adalah siswa SD Negeri Mamuju Tengah terbagi kedalam 3 kluster yaitu kluster Prestasi Tinggi terdiri dari 40 siswa, kluster Prestasi Sedang terdiri dari 8 siswa, sedangkan kluster Prestasi Kurang terdiri dari 52 siswa. Jika dilihat sebaran siswa pada setiap kluster diketahui bahwa pada kluster prestasi rendah terdiri dari 52 siswa yang menggambarkan bahwa pada SD negeri mamuju tengah masih dibutuhkan perhatian khusus terkait peningkatan prestasi. Saran pada penelitian selanjutnya adalah menggunakan *dataset* dari sumber utama dalam jumlah yang lebih besar, serta menggunakan *hybrid* algoritma agar dapat diperoleh hasil yang lebih baik.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada STMIK Profesional yang telah memberikan dukungan hingga penelitian ini dapat terlaksana dengan baik.

DAFTAR PUSTAKA

- Abdulnassar, A. A., & Nair, L. R. (2023). Performance analysis of Kmeans with modified initial centroid selection algorithms and developed Kmeans9+ model. *Measurement: Sensors*, 25, 100666. <https://doi.org/10.1016/j.measen.2023.100666>
- Anwar, S. (2022). PENGELOMPOKKAN HASIL BELAJAR SISWA DENGAN METODE CLUSTERING K-MEANS. *JURISISTEKNI (Jurnal Sistem Informasi Dan Teknologi Informasi)*, 4(2), 60–72.
- Ay, M., Özbakır, L., Kulluk, S., Gülmez, B., Öztürk, G., & Özer, S. (2023). FC-Kmeans: Fixed-centered K-means algorithm. *Expert Systems with Applications*, 211, 118656. <https://doi.org/10.1016/j.eswa.2022.118656>
- Chu, X., Lei, J., Liu, X., & Wang, Z. (2020). KMEANS Algorithm Clustering for Massive AIS Data Based on the Spark Platform. *2020 5th International Conference on Control, Robotics and Cybernetics (CRC)*, 36–39. <https://doi.org/10.1109/CRC51253.2020.9253451>
- Durojaye, D. I., & Obunadike, G. (2022). ANALYSIS AND VISUALIZATION OF MARKET SEGEMENTATION IN BANKING SECTOR USING KMEANS MACHINE LEARNING ALGORITHM. *FUDMA JOURNAL OF SCIENCES*, 6(1), 387–393. <https://doi.org/10.33003/fjs-2022-0601-910>
- Durugkar, S. R., Raja, R., Nagwanshi, K. K., & Kumar, S. (2022). Introduction to data mining. *Data Mining and Machine Learning Applications*, 1–19. <https://doi.org/10.1002/9781119792529.ch1>
- Gao, X., Ding, X., Han, T., & Kang, Y. (2022). Analysis of influencing factors on excellent teachers' professional growth based on DB-Kmeans method. *EURASIP Journal on Advances in Signal Processing*, 2022(1), 117. <https://doi.org/10.1186/s13634-022-00948-2>
- 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.

<https://doi.org/10.25077/TEKNOSI.v5i1.2019.17-24>

- Huang, W., Peng, Y., Ge, Y., & Kong, W. (2021). A new Kmeans clustering model and its generalization achieved by joint spectral embedding and rotation. *PeerJ Computer Science*, 7, e450. <https://doi.org/10.7717/peerj-cs.450>
- Ismaya, N., Zyen, A. K., Studi, P., Informatika, T., Islam, U., Ulama, N., Tengah, J., Artikel, I., Takhassus, M. P., K-means, A., & Davies-, I. (2022). Penentuan Siswa Berprestasi Menggunakan Metode K-Means Clustering Di Smp Takhassus Al Qur ' an. *Jurnal Teknik Informatika*, 1(1), 64–68.
- Martinez-Plumed, F., Contreras-Ochando, L., Ferri, C., Hernandez-Orallo, J., Kull, M., Lachiche, N., Ramirez-Quintana, M. J., & Flach, P. (2021). CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories. *IEEE Transactions on Knowledge and Data Engineering*, 33(8), 3048–3061. <https://doi.org/10.1109/TKDE.2019.2962680>
- Nurmayanti, W. P., Ratnaningsih, D. J., Nisrina, S., Rahim, A., Malthuf, M., & Kusuma, W. (2022). Clustrering of BPJS National Health Insurance Participant Using DBSCAN Algorithm. *Jurnal Varian*, 6(1), 25–34. <https://doi.org/10.30812/varian.v6i1.1886>
- Putra, U. N., Ulil, P., Aulia, F., Saepudin, S., Studi, P., Informasi, S., & Putra, U. N. (2021). *Penerapan Data Mining K-Means Clustering Untuk Mengelompokkan Berbagai Jenis Merk Smartphone*. 241–249.
- Wei, Y., Jang-Jaccard, J., Sabrina, F., & Alavizadeh, H. (2020). Large-Scale Outlier Detection for Low-Cost PM₁₀ Sensors. *IEEE Access*, 8, 229033–229042. <https://doi.org/10.1109/ACCESS.2020.3043421>
- Zhao, H. (2021). Research on Improvement and Parallelization of K-Means Clustering Algorithm. *2021 IEEE 3rd International Conference on Frontiers Technology of Information and Computer (ICFTIC)*, 57–61. <https://doi.org/10.1109/ICFTIC54370.2021.9647385>