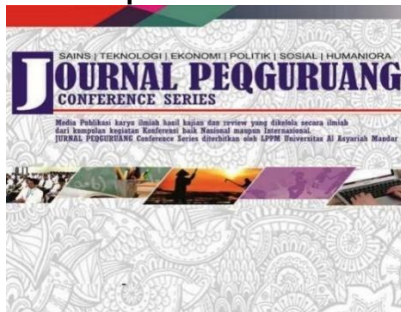


Graphical abstract



CLUSTERING PRESTASI AKADEMIK SISWA MTs MENGGUNAKAN METODE *K-MEANS*

¹Muslihan, ²Yudhi Adhitya, ³Irmayanti.

¹²³ Teknik Informatika. ¹²³ Universitas Al Asyariah
Mandar

*Corresponding author

irmayantiirmayanti965@gmail.com,

yudhiadhitya@gmail.com, ianmuslihan@gmail.com

Abstract

Academic achievement is a crucial element in evaluating the quality of learning and student outcomes in educational institutions, particularly at the Madrasah Tsanawiyah (MTs) level. Routine monitoring of these outcomes is essential for schools to take appropriate action and improve the quality of learning. However, the large number of students and their diverse backgrounds often make manual analysis inefficient and prone to subjectivity. To address this issue, a technology-based approach is needed, one of which is the application of data mining techniques. This study adopted the K-Means clustering method to group students based on academic data such as report card grades, attendance, participation in extracurricular activities, attitude assessments, and final exam results. The K-Means method was chosen because of its ability to automatically group data into clusters with similar characteristics without requiring initial labels. The results of this clustering can be used by schools to more objectively identify student achievement patterns, allowing for more effective learning and mentoring strategies. The system developed in this study is expected to be a tool in academic decision-making and support the improvement of the quality of educational services in MTs. In addition, this system also has the potential to become a foundation for the development of an analytics-based academic information system that drives digital transformation in the world of education and increases the effectiveness of school management as a whole.

Keywords: *Clustering, K-Means Method, Student Academic Achievement, Data Mining*

Abstrak

Prestasi akademik merupakan elemen krusial dalam mengevaluasi mutu pembelajaran dan capaian siswa di institusi pendidikan, khususnya di jenjang Madrasah Tsanawiyah (MTs). Pemantauan terhadap capaian ini perlu dilakukan secara rutin agar sekolah dapat memberikan tindakan yang tepat dan meningkatkan kualitas pembelajaran. Namun, jumlah siswa yang besar serta latar belakang yang beragam sering kali membuat proses analisis manual menjadi tidak efisien dan rawan subjektivitas. Untuk mengatasi permasalahan ini, diperlukan pendekatan berbasis teknologi, salah satunya melalui penerapan teknik *data mining*. Penelitian ini mengadopsi metode *clustering K-Means* guna mengelompokkan siswa berdasarkan data akademik seperti nilai rapor, kehadiran, partisipasi dalam kegiatan ekstrakurikuler, penilaian sikap, serta hasil ujian akhir. Metode *K-Means* dipilih karena kemampuannya mengelompokkan data ke dalam kluster dengan karakteristik serupa secara otomatis tanpa memerlukan label awal. Hasil pengelompokan ini dapat dimanfaatkan oleh pihak sekolah untuk mengidentifikasi pola prestasi siswa secara lebih objektif, sehingga dapat dirancang strategi pembelajaran dan pendampingan yang lebih efektif. Sistem yang dikembangkan dalam penelitian ini diharapkan dapat menjadi alat bantu dalam pengambilan keputusan akademik serta mendukung peningkatan mutu layanan pendidikan di lingkungan MTs. Selain itu, sistem ini juga berpotensi menjadi fondasi bagi pengembangan sistem informasi akademik berbasis analitik yang mendorong transformasi digital dalam dunia pendidikan serta meningkatkan efektivitas manajemen sekolah secara keseluruhan.

Kata kunci: *Clustering, Metode K-Means, Prestasi Akademik Siswa, Data mining*

Article history

DOI: 10.35329/jp.v7i2.6538

Received : 2025-08-05 | Received in revised form : 2025-08-12 | Accepted : 2025-11-26

1. PENDAHULUAN

Pendidikan buat seluruh selalu menjadi bagian integral dari rencana pembangunan berkelanjutan. KTT global tentang Pembangunan Berkelanjutan (WSSD) pada tahun 2002 mengadopsi Rencana Aplikasi Johannesburg (JPOI) yang dalam Bagian X-nya, menegaskan kembali Tujuan Pembangunan Milenium 2 dalam mencapai pendidikan dasar universal di tahun 2015 dan tujuan Kerangka Aksi Dakar dalam Pendidikan untuk Semua buat menghilangkan kesenjangan gender dalam pendidikan dasar dan menengah pada tahun 2005 serta pada seluruh pendidikan pada tahun 2015. JPOI membahas perlunya mengintegrasikan pembangunan berkelanjutan ke dalam pendidikan formal pada seluruh jenjang, serta melalui kesempatan pendidikan informal dan nonformal. (United Nations <https://sdgs.un.org/topics/education>)

Pendidikan merupakan pilar kemajuan sebuah peradaban (masyarakat). Semakin tinggi tingkat pendidikan dalam sebuah masyarakat, maka peradaban masyarakat tersebut akan semakin maju. Sebaliknya, peradaban yang ada akan semakin mundur dan terbelakang apabila masyarakat tidak mampu menjunjung tinggi nilai-nilai pendidikan yang ada (Anwar et al., 2022).

Siswa menggunakan sekolah sebagai sarana untuk melanjutkan pendidikannya. setiap sekolah harus mampu menguasai siswanya untuk membimbing mereka menuju kemajuan. Tingkat keberhasilan yang tinggi dan tingkat kegagalan yang rendah di kalangan siswa merupakan indikator kualitas sistem pendidikan. Saat ini, sektor pendidikan harus mampu bersaing dengan memanfaatkan seluruh sumber daya yang dimilikinya. Menurut Yani Prihati, Suwarno dan Alexander. (Prihati et al., 2019) Tujuan pendidikan itu adalah mencapai standar pendidikan nasional yang meliputi standar isi, proses, kompetensi lulusan, tenaga kependidikan, sarana dan prasarana, pengelolaan, pembiayaan dan penilaian pendidikan. Sekolah MTs Khusnul Khatimah adalah salah satu institusi pendidikan yang berkomitmen untuk memberikan pendidikan yang berkualitas. Untuk mencapai tujuan ini, perlu dilakukan pemantauan dan analisis secara terus-menerus terhadap hasil akademik siswa. Salah satu metode yang dapat digunakan untuk melakukan pengelompokan ini adalah algoritma *K-Means*. (Adam Rifais & Laksana, 2024)

Dalam sektor komersial, kemajuan teknologi sistem informasi mempunyai arti yang sangat penting. Manfaat teknologi sistem informasi yang saat ini berkembang sangat pesat antara lain kemudahan dalam penggunaan, kecepatan pencarian informasi, dan kebebasan dalam pencarian informasi (Al & Mandar, 2025)

Clustering merupakan metode untuk mengelompokkan data berdasarkan tingkat kedekatan (persamaan). Sebuah klaster tidak harus sama; tetapi dikelompokkan berdasarkan kesamaan karakteristik sampel, salah satunya dengan menerapkan rumus jarak Ecclesiastes. Fakhri dan rekan-rekan (2021) menyatakan bahwa pengelompokan merupakan teknik untuk

mengidentifikasi dan mengelompokkan data dengan karakteristik yang sama. Pengelompokan merupakan teknik penambangan data yang tidak memerlukan pengawasan, artinya pendekatan ini digunakan tanpa pelatihan, tanpa instruktur, dan tidak memerlukan hasil yang ditentukan (Ananda, 2024).

Pemanfaatan teknologi data dan informasi kini digunakan dalam berbagai bidang, termasuk pendidikan (Adhitya et al., 2023). Dalam penelitian ini, metode *K-Means Clustering* diterapkan untuk mengelompokkan prestasi akademik siswa secara sistematis, guna mendukung analisis yang lebih akurat dan efisien.

Tujuan dari penelitian ini adalah untuk merancang sebuah sistem cluster dengan menggunakan metode *K-Means* berbasis web menawarkan kemudahan dalam pengolahan data. Dengan sistem berbasis web, data prestasi akademik siswa dapat dikelola, dianalisis, dan disajikan secara real-time. prestasi akademik nilai siswa yang dapat digunakan oleh guru, kepala sekolah, dan pihak terkait dapat dengan mudah mengakses hasil analisis untuk memahami pola prestasi siswa dan merancang program pembelajaran yang lebih sesuai dengan kebutuhan setiap kelompok siswa. Namun, penerapan sistem berbasis web untuk clustering prestasi akademik siswa di lingkungan MTs masih jarang ditemukan.

Oleh karena itu, peneliti mempunyai ide bertujuan untuk merancang sebuah sistem "*Clustering Prestasi Akademik Siswa MTs Menggunakan Metode K-Means*". Sistem ini diharapkan dapat menjadi alat bantu yang efektif bagi sekolah dalam meningkatkan kualitas pembelajaran dan layanan pendidikan.

2. METODE PENELITIAN

K-Means adalah suatu algoritma pengklasteran yang cukup sederhana yang mempartisi database ke dalam beberapa cluster. Algoritma relatif praktis buat diimplementasikan dan dijalankan, relatif cepat, simpel diadaptasi dan banyak digunakan (Wu & Kumar, 2009). Prinsip utama teknik ini ialah menyusun K buah partisi/pusat massa (centroid)/rata-rata (mean) dari sekumpulan data. Algoritma K-Means dimulai menggunakan pembentukan partisi klaster di awal kemudian secara iteratif partisi cluster ini diperbaiki hingga tidak terjadi perubahan yang signifikan di partisi cluster (Sibuea & Safta, 2017)

Rumus yang umum digunakan untuk menghitung jarak antara data x_i dan pusat cluster u_j adalah jarak Euclidean:

1. Menghitung Sum of Square Within-cluster dengan Rumus (1)

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

2. Menghitung Sum of Square Between-cluster dengan Rumus (2).

$$SSB_{i,j} = d(c_i, c_j) \quad (2)$$

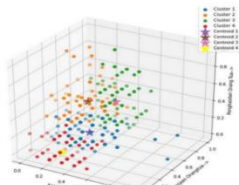
3. Menghitung rasio dengan Rumus (3).

$$R_{i,j} = \frac{SSW_j + SSW_i}{SSB_{i,j}} \quad (3)$$

4. Menghitung indeks DB dengan Rumus (4).

$$\text{Indeks DB} = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_{i,j}) \quad (4)$$

Clustering merupakan teknik dalam penambangan data yang bermanfaat untuk menganalisis data sehingga lebih tepat dalam mengatasi masalah pengelompokan data atau memisahkan kumpulan data menjadi beberapa subset. Tujuan dari pengelompokan adalah untuk mengorganisir data sedemikian rupa sehingga interaksi antara anggota dalam satu kluster menjadi kuat, sementara interaksi antara anggota di kluster yang berbeda menjadi lemah. Objek dalam satu kluster memiliki sifat yang mirip tetapi berbeda dari objek di kluster yang lain. Dengan demikian, pengelompokan sangat bermanfaat untuk mengenali kelompok atau kluster yang tidak dikenal dalam data (Dacwanda & Nataliani, 2021)



Gambar 1 Proses Clustering

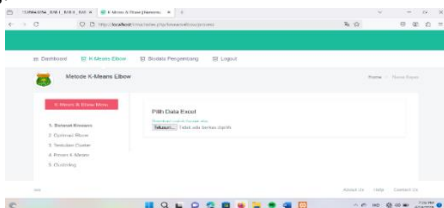
3. HASIL DAN PEMBAHASAN

A. Hasil

Hasil penelitian dari Clustering Prestasi Akademik Siswa MTs Menggunakan Metode K-Means adalah untuk mendapatkan gambaran yang utuh mengenai prestasi akademik siswa MTs Khusnul Khatimah, hasil pengolahan data menggunakan metode K-Means, siswa MTs berhasil dikelompokkan ke dalam tiga kluster berdasarkan kemiripan nilai akademik yang dimiliki. Ketiga kluster tersebut masing-masing mewakili tingkat prestasi siswa, yaitu kluster dengan prestasi tinggi, sedang, dan rendah. Pengelompokan ini bertujuan untuk mengetahui distribusi kemampuan akademik siswa secara lebih sistematis dan terarah.

1. From memasukkan data

Pada menu *K Means Elbow* terdapat tombol telusuri untuk memasukkan dataset yang sesuai format dan fitur yang sama dengan data yang ada pada data testing.

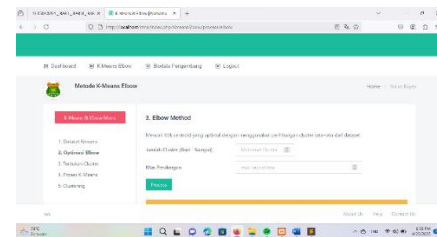


Gambar 2 .Form memasukkan dataset

2. Optimasi Elbow

Algoritma Elbow (metode Elbow) adalah metode bantu untuk menentukan jumlah kluster optimal (nilai K) dalam algoritma *K-Means Cluster*. Dalam *K-Means*, kita harus menentukan berapa banyak kluster (K) yang ingin dibuat. Tapi sering kali kita tidak tahu berapa nilai K yang tepat. Berikut cara kerja algoritma elbow dalam K means:

- 1) Jalankan *K-Means* untuk beberapa nilai K (misalnya dari 1 sampai 10).
- 2) Hitung nilai SSE (Sum of Squared Error) untuk setiap K. SSE adalah total jarak kuadrat antara tiap titik ke centroid-nya.
- 3) Buat grafik nilai SSE vs jumlah kluster (K).
- 4) Lihat titik di mana penurunan SSE tidak terlalu signifikan lagi di situ lah terbentuk "titik siku" (elbow).
- 5) Titik elbow itu dianggap sebagai jumlah kluster yang paling optimal.

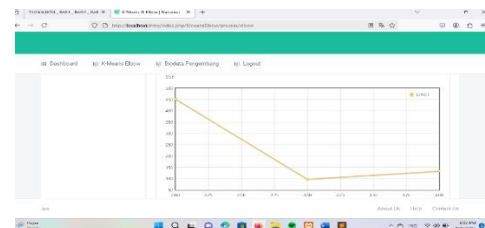


Gambar 2. Menentukan jumlah clustering dan max perulangan

Cluster	SSE	ARI	ARI	ARI	ARI
Cluster 1	481.7671	0.0	0.0000	0.0	0.0000
Cluster 2	481.7671	0.0	0.0000	0.0	0.0000
Cluster 3	481.7671	0.0	0.0000	0.0	0.0000
Cluster 4	481.7671	0.0	0.0000	0.0	0.0000

Cluster	Nilai	Total
2	481.7671	481.7671
3	481.7671	481.7671
4	481.7671	481.7671

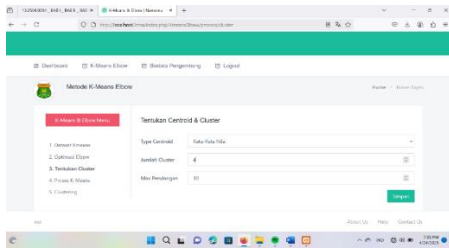
Gambar 3 Hasil optimalisasi elbow panjang nilai dalam bentuk nilai



Gambar 4. Hasil Optimalisasi elbow dari dataset berbentuk garfik SSE

3. Tentukan Cluster

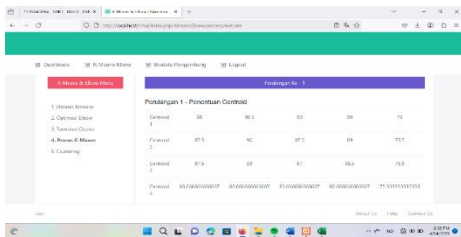
Pada menu tentukan cluster berisi tentang type centroid, jumlah cluster, dan max perulangan. pada jumlah cluster kita memilih berapa kali perulangan dan setelah memilih berapa kali perulangan lalu klik simpan dan max peru langan adalah maximal berapa kali perulangan dalam menentukan *cluster*.



Gambar 5 Penentuan *cluster* pada *K Means*

4. Proses *K-Means*

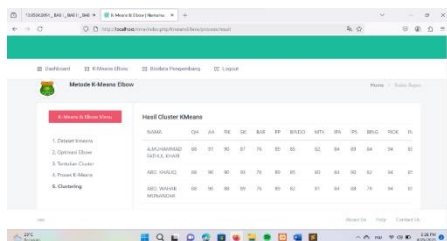
Di dalam proses *K-Means* data di bagi ke dalam kelompok berdasarkan kedudukan jarak antara data, setelah melakukan iterasi (perhitungan) sesuai jumlah yang sudah di tentukan pada penentuan cluster, centroid tidak berubah lagi (*konfergen*) ,atau jumlah iterasi maksimum tercapai.



Gambar 6 Proses perulangan algoritma *K-Means*

5. Clustering

Sesudah iterasi selesai, setiap data sudah tergabung dalam klaster yang paling sesuai. Hasil ini bisa divisualisasikan dalam bentuk graafik klaster berwarna.



Gambar 7 Hasil clustering

Klaster 1 merupakan kelompok siswa dengan prestasi akademik tinggi, ditandai dengan rata-rata nilai yang berkisar antara 85 hingga 100. Siswa dalam klaster ini menunjukkan penguasaan materi yang sangat baik di hampir semua mata pelajaran. Rata-rata nilai mata pelajaran pada klaster ini adalah sebagai berikut: : nilai akademik 89, kehadiran 92, Aktivitas Ekstrakurikuler 90, Nilai Sikap 90, Nilai Ujian 76, Dengan demikian, rata-rata keseluruhan siswa dalam klaster 1 adalah 87,8. Hasil ini menunjukkan bahwa

siswa dalam kelompok ini cenderung aktif, disiplin, dan memiliki motivasi belajar yang tinggi.

Klaster 2 mencerminkan siswa dengan prestasi sedang, dengan rata-rata nilai berada pada rentang 70 hingga 84. Pada kelompok ini, siswa umumnya memiliki pemahaman yang cukup terhadap materi pelajaran, namun belum merata pada semua mata pelajaran. Nilai rata-rata yang diperoleh siswa dalam klaster ini yaitu: nilai akademik 85, kehadiran 80, Aktivitas Ekstrakurikuler 82, Nilai Sikap 80, Nilai Ujian 79. Rata-rata keseluruhan nilai siswa dalam klaster 2 adalah 75,0. Hal ini menunjukkan bahwa siswa pada kelompok ini masih memiliki peluang besar untuk ditingkatkan prestasinya melalui bimbingan belajar atau pengayaan materi.

Klaster 3 terdiri dari siswa dengan prestasi akademik rendah, dengan rata-rata nilai berada di bawah 69. Siswa-siswa dalam cluster ini mengalami kesulitan untuk memahami materi pembelajaran, baik dalam mata pelajaran ekstra maupun sosial. Nilai rata-rata yang diperoleh pada klaster ini adalah: nilai akademik 65, kehadiran 70, Aktivitas Ekstrakurikuler 71, Nilai Sikap 70, Nilai Ujian 60. Rata-rata keseluruhan nilai siswa dalam klaster 3 adalah 65,8.

Hasil clustering ini menunjukkan bahwa metode K-Means efektif dalam mengelompokkan siswa berdasarkan tingkat kemampuan akademiknya, dan dapat menjadi acuan bagi pihak sekolah dalam merancang strategi pembelajaran yang lebih tepat sasaran.

B. Pembahasan

Penelitian ini melalui beberapa tahapan sistematis untuk melakukan proses clustering prestasi akademik siswa menggunakan metode K-Means. adapun tahapan-tahapan tersebut di jelaskan sebagai berikut:

1. Pengumpulan Data

Tahapan awal dalam penelitian ini adalah pengumpulan data prestasi akademik siswa MTs. Data yang dikumpulkan meliputi :

- Nilai ujian
- Nilai sikap
- Kehadiran siswa
- Nilai Akademik
- Aktivitas ekstrakurikuler

2. Pra-pemrosesan Data

Data yang telah dikumpulkan kemudian dipersiapkan sebelum diolah dalam algoritma K-Means . tahapan pra-pemrosesan meliputi :

- Pembersihan data (Data Cleaning) : Menghapus data yang tidak lengkap, duplikat, atau mengandung kesalahan.
- Pembuatan Tabel Excel: Data dimasukkan harus berbentuk Excel yang di sesuaikan dengan data training

3. Pre-processing

Pada tahap ini, metode *K-Means* akan diaplikasikan guna melakukan proses pengelompokan data. Hasil dari proses clustering tersebut akan dimanfaatkan sebagai

dasar dalam mengidentifikasi tingkat kedisiplinan siswa, yang diklasifikasikan ke dalam kategori rendah, sedang, dan tinggi. Penentuan nilai pusat awal (*centroid*) dilakukan dengan memilih data berdasarkan poin tertinggi, terendah, dan poin sedang di dapatkan data :

$$C1 = 9, 8, 9$$

$$C2 = 8, 7, 8$$

$$C3 = 6, 7, 6$$

Langkah selanjutnya yaitu menghitung jarak pusat kluster untuk data peserta didik (siswa), berikut ialah perhitungan jarak kluster untuk pengulangan pertama:

$$C1 = \sqrt{(9-9)^2 + (8-8)^2 + (9-9)^2} = 0$$

$$C2 = \sqrt{(9-8)^2 + (8-7)^2 + (9-8)^2} = 1,73$$

$$C3 = \sqrt{(9-6)^2 + (8-7)^2 + (9-6)^2} = 2,65$$

Hasil perhitungan untuk semua data adalah sebagai berikut.

A.MUHAMMAD FATHUL KHUR	3.7080992435478	9.8488578017961	17.161002301731	23.187640385918
ABD. KHALIQ	3.7080992435478	9.1651513899117	16.7102534973	23.180451534285
ABD. WAHAB MUNANDAR	5.8949130612758	5.56776436283	12.74754878382	18.768294228056
AFNAN ALFARISI	13.444329659749	5.56776436283	11.640446726823	14.730919862656
AHMAD MULIADI	19.202864369672	11.489125293076	7.6485292703892	6.806859285554
ALFINA ANGRAENI	17.168204713389	14.866068747319	7.6485292703892	13.868429375143
DELTA	25.840859118845	19.798989873223	12.509996003197	7.34846492283495

Gambar 8 Hasil Perhitungan Cluster Iterasi 1

Dari gambar 8, kita bisa menentukan jarak cluster terdekat dari setiap data, berikut ialah gambar dari pengelompokkan data berdasar jarak terdekat dari pusat cluster :

A.MUHAMMAD FATHUL KHAIIR	1
ABD. KHALIQ	1
ABD. WAHAB MUNANDAR	2
AFNAN ALFARISI	2
AHMAD MULIADI	4
ALFINA ANGRAENI	3
DELTA	4
DIRHAM	4

Gambar 9 Pengelompokkan data perulangan 1

Dari gambar 9 Kita bisa menetapkan pusat kluster yang baru, berikut ialah perhitungan dalam menentukan pusat kluster yang baru:

C1

$$\frac{(9+9+8+9+9+9)}{6} \Rightarrow \sum = 53$$

$$\frac{(8+8+8+8+9+8)}{6} \Rightarrow \sum = 49$$

$$\frac{(9+8+9+9+8+8)}{6} \Rightarrow \sum = 51$$

$$\text{Total Sigma C1} = 53 + 49 + 51 = 153$$

C2:

$$\frac{(8+8+7+7+8+7+7+7+7+8+8+8+7)}{13} \Rightarrow \sum = 99$$

$$\frac{(7+7+8+7+8+7+7+8+7+8+9+6)}{13} \Rightarrow \sum = 99$$

$$\frac{(8+8+7+7+8+6+7+8+7+8+7+7+8)}{13} \Rightarrow \sum = 97$$

$$\text{Total Sigma C2} = 99 + 99 + 97 = 295$$

C3

$$\frac{(6+7+6+6+7+6)}{16} \Rightarrow \sum = 38$$

$$\frac{(7+6+7+6+7+7)}{16} \Rightarrow \sum = 40$$

$$\frac{(7+7+6+7+6+7)}{16} \Rightarrow \sum = 40$$

$$\text{Total Sigma C3} = 38 + 40 + 40 = 118$$

Rekapitulasi Sigma Total:

a. C1 Sigma: 153

b. C2 Sigma: 295

c. C3 Sigma: 118

$$\text{Total Seluruhnya: } 153 + 295 + 118 = 566$$

Langkah berikutnya ialah menghitung jarak pusat cluster yang baru terhadap data peserta didik (siswa); berikut ini ialah perhitungan jarak cluster baru dalam iterasi 1:

$$C1 = \sqrt{(9-8,83)^2 + (8-8,17)^2 + (9-8,50)^2} = 0,71$$

$$C2 = \sqrt{(9-7,46)^2 + (8-8,23)^2 + (9-7,38)^2} = 1,71$$

$$C3 = \sqrt{(9-6,33)^2 + (8-6,67)^2 + (9-6,67)^2} = 2,52$$

Berikut ini ialah hasil dari seluruh perhitungan data:

Perulangan 2 - Penentuan Centroid													
Centroid 1	88	90.5	90	90	76	89	85	81	84	89.5	83	94	85
Centroid 2	87.5	90	87.5	89	75.5	88	80.5	81.5	82	88.5	79	91.5	85
Centroid 3	89	80	89	87	76	90	85	82	84	86	72	88	85
Centroid 4	86.5	82	84	83	75.25	87.75	78.25	80.25	80	88	71.25	88	85

Gambar 10 Hasil Perhitungan Cluster Iterasi 2

Dari gambar 10 diatas, Kita bisa mengukur jarak cluster terdekat untuk setiap data, berikut ini ialah ilustrasi pengelompokan data berdasarkan jarak terdekat dari pusat cluster yang baru:

A.MUHAMMAD FATHUL KHAIR	1
ABD. KHALIQ	1
ABD. WAHAB MUNANDAR	2
AFNAN ALFARISI	2
AHMAD MULIADI	4
ALFINA ANGRAENI	3
DELTA	4
DIRHAM	4

About Us Help Contact Us

Gambar 11 Pengelompokan Data Perulangan 2

Dari Gambar 11, kita bisa mengidentifikasi pusat klaster yang baru. Berikut ini ialah perhitungan dalam menemukan pusat klaster yang baru:

C1:

$$\sum_{i=1}^6 x \ 1i = 9 + 9 + 8 + 9 + 9 + 9 = 53 \Rightarrow \frac{\sum_{i=1}^6 x \ 1i}{6} = \frac{53}{6} = 8,83$$

$$\sum_{i=1}^6 x \ 2i = 8 + 8 + 8 + 9 + 8 = 49 \Rightarrow \frac{\sum_{i=1}^6 x \ 2i}{6} = \frac{49}{6} = 8,17$$

$$\sum_{i=1}^6 x \ 3i = 9 + 8 + 9 + 9 + 8 + 8 = 51 \Rightarrow \frac{\sum_{i=1}^6 x \ 3i}{6} = \frac{51}{6} = 8,50$$

$$\text{Total } \Sigma \ C1 = 53 + 49 + 51 = 153$$

C2:

$$\sum_{i=1}^{10} x \ 1i = 76 \Rightarrow \frac{\sum_{i=1}^{10} x \ 1i}{10} = \frac{76}{10} = 7,60$$

$$\sum_{i=1}^{10} x \ 2i = 76 \Rightarrow \frac{\sum_{i=1}^{10} x \ 2i}{10} = 7,60$$

$$\sum_{i=1}^{10} x \ 3i = 75 \Rightarrow \frac{75}{10} = 7,50$$

$$\text{Total } \Sigma \ C2 = 76 + 76 + 75 = 227$$

C3:

$$\sum_{i=1}^9 x \ 1i = 59 \Rightarrow \frac{59}{9} = 6,56$$

$$\sum_{i=1}^9 x \ 2i = 61 \Rightarrow \frac{61}{9} = 6,78$$

$$\sum_{i=1}^9 x \ 3i = 61 \Rightarrow \frac{64}{9} = 6,78$$

$$\text{Total } \Sigma \ C3 = 59 + 61 + 61 = 181$$

Total Seluruh Sigma (Σ Semua):

$$\text{Total} = \Sigma \ C1 + \Sigma \ C2 + \Sigma \ C3 = 153 + 227 + 181 = 561$$

Pengulangan perhitungan kluster akan terus dilakukan sampai hasil yang didapat tetap sama atau tidak mengalami perubahan

4. Proses *Clustering K-Means*

Adapun proses tahapan *Clustering* untuk *K-Means* sebagai berikut:

- Input* data baru adalah yang ingin kita klaster kedalam model *K-Means* yang telah dibentuk sebelumnya. Data ini harus memiliki format dan fitur yang sama dengan data yang ada pada contoh data yang format xl
- Tentukan *cluster*. Setelah memasukan dataset, kita memilih tipe centroid yang di sesuaikan dengan kebutuhan data Menentukan jumlah *cluster* dari dataset yang sudah dimasukkan. Tentukan maksimal perulangan untuk mendapatkan hasil yang lebih efisien.
- Proses *K-Mean*
Setelah menentukan tipe centroid, jumlah cluster, dan maksimal perulangan. Maka proses ini dilakukan sampai tidak ada perubahan anggota cluster atau, perubahan sangat kecil atau sudah mencapai hasil maksimum.
- Clustering*. Setelah melalui proses *K-Means* data sudah terkelompok dalam cluster berdasarkan kesamaan nilai atau jarak ke centroid

4. SIMPULAN

Penelitian ini bertujuan untuk mengembangkan sistem pengelompokan data prestasi akademik siswa menggunakan pendekatan data mining dengan algoritma K-Means di MTs Khusnul Khatimah. Dari hasil implementasi sistem, siswa kelas VII berhasil diklasifikasikan ke dalam beberapa kelompok berdasarkan kesamaan nilai akademik mereka. Proses ini menghasilkan tiga kategori utama yang merepresentasikan tingkat prestasi siswa, yakni kategori tinggi, sedang, dan rendah. Siswa pada kategori pertama memiliki rata-rata nilai 87,8 dan menunjukkan penguasaan materi yang sangat kuat di hampir seluruh mata pelajaran. Sementara itu, kategori kedua mencakup siswa dengan nilai rata-rata 75,0, yang secara umum memahami materi dengan baik, meskipun belum sepenuhnya merata di semua mata pelajaran. Kategori ketiga menunjukkan rata-rata 65,8, yang menandakan adanya kesulitan dalam mengikuti pembelajaran, yang menggambarkan bahwa siswa dalam kelompok ini perlu mendapatkan perhatian lebih melalui pendampingan atau program remedial.

DAFTAR PUSTAKA

Adam Rifais, & Laksana, T. G. (2024). Penerapan Algoritma K-Means Clustering Untuk Mengetahui Kemampuan Akademik Siswa Berbasis WEB. *Jurnal Ilmiah Multidisiplin*, 1(2), 157–183. <https://doi.org/10.62282/juilmu.v1i2.157-183>

Adhitya, Y., Mulyani, G. S., Köppen, M., & Leu, J. S. (2023). IoT and Deep Learning-Based Farmer Safety System. *Sensors*, 23(6), 1–18. <https://doi.org/10.3390/s23062951>

Ananda, R. A. (2024). *Clustering Menggunakan Algoritma K-Means*.
Anandardics, M. (2020). *Data mining: Concept, models, methods, and algorithms*.
eans untuk Mengelompokan Data Perjudian Berdasarkan Wilayah di Kota Binjai (Studi Kasus : Pengadilan Negeri Binjai).
4.

Anwar, S., Suprapti, T., Dwilestari, G., & Ali, I. (2022).
Pengelompokan Hasil Belajar Siswa dengan Metode
Clustering K-Means. *JURSISTEKNI (Jurnal Sistem Informasi Dan Teknologi Informasi)*, 4(2), 60–72.

Dacwanda, D. O., & Nataliani, Y. (2021). Implementasi
k-Means Clustering untuk Analisis Nilai Akademik
Siswa Berdasarkan Nilai Pengetahuan dan
Keterampilan. *Aiti*, 18(2), 125–138.
<https://doi.org/10.24246/aiti.v18i2.125-138>

Prihati, Y., Suwarno, & Dharmawan, A. (2019).
Implementasi Algoritma K-Means Untuk Pemetaan
Prestasi Akademik Siswa Disekolah Dasar Terang Bagi
Bangsa Pati. *Kinabalu*, 11(2), 50–57.

Sibuea, M. L., & Safta, A. (2017). Pemetaan Siswa
Berprestasi Menggunakan Metode K-Means Clustering.
Jurteks, 4(1), 85–92.
<https://doi.org/10.33330/jurteks.v4i1.28>