

## Sistem Penyeleksian peserta Didik Baru Menggunakan Algoritma K-Means Clustering

Nuril Miftaql Rozaq<sup>1</sup>, Purnomo Hadi Susilo<sup>2</sup>, Agus Setia Budi<sup>3</sup>,

<sup>1,2,3</sup>Program Studi Sains dan Teknologi Universitas Islam, Jln. Veteran No. 53A, Kelurahan Jetis, Kecamatan Lamongan, Kabupaten Lamongan, Jawa Timur.

email : [nmrkebo08@gmail.com](mailto:nmrkebo08@gmail.com)<sup>1</sup>, [hadyjelak.purnomo@gmail.com](mailto:hadyjelak.purnomo@gmail.com)<sup>2</sup>, [geniusbudi@yahoo.com](mailto:geniusbudi@yahoo.com)<sup>3</sup>

\* Penulis : Nuril Miftaql Rozaq

**Abstract:** The objectives of this study are: 1) to implement the K-Means Clustering algorithm to group data from the selection of prospective new students at SMA Unggulan Karangsawo, and 2) to assess the accuracy level of the resulting clustering in supporting the determination of the Superior Class and Regular Class. This study uses a quantitative approach with the K-Means Clustering method through data mining techniques as the basis for decision making. The research subjects were students of SMA Unggulan Karangsawo, with data in the form of academic test scores used as attributes in the clustering process. The data were then separated into two groups, namely the first cluster for the Superior Class and the second cluster for the Regular Class. The results of the analysis show that K-Means is able to group students according to their academic abilities, so it can help determine the superior class more objectively. Evaluation using the BCV/WCV ratio produced a value of 0.4223, which is included in the good category.

**Keywords:** K-Means; Clustering; Superior Class; BCV/WCV Ratio

**Abstrak:** Tujuan penelitian ini adalah: 1) mengimplementasikan algoritma K-Means Clustering untuk mengelompokkan data hasil seleksi calon peserta didik baru di SMA Unggulan Karangsawo, dan 2) menilai tingkat akurasi pengelompokan yang dihasilkan dalam mendukung penentuan Kelas Unggulan maupun Kelas Reguler. Penelitian ini menggunakan pendekatan kuantitatif dengan metode K-Means Clustering melalui teknik data mining sebagai dasar pengambilan keputusan. Subjek penelitian adalah siswa SMA Unggulan Karangsawo, dengan data berupa nilai tes akademik yang dijadikan atribut dalam proses klasterisasi. Data kemudian dipisahkan ke dalam dua kelompok, yaitu cluster pertama untuk Kelas Unggulan dan cluster kedua untuk Kelas Reguler. Hasil analisis menunjukkan bahwa K-Means mampu mengelompokkan siswa sesuai dengan kemampuan akademiknya, sehingga dapat membantu penentuan kelas unggulan secara lebih objektif. Evaluasi menggunakan rasio BCV/WCV menghasilkan nilai 0,4223, yang termasuk dalam kategori baik.

**Kata kunci:** K-Means; Clustering; Kelas Unggulan; BCV/WCV Ratio

Diterima: 9 September 2025

Direvisi: 6 Oktober 2025

Diterima: 9 Oktober 2026

Diterbitkan: 30 Januari 2026

Versi sekarang: Januari 2026

### 1. Pendahuluan

Kemajuan teknologi informasi yang terus melaju pesat di berbagai aspek kehidupan, khususnya dalam bidang komputer, mendorong lahirnya berbagai inovasi untuk meningkatkan cara pengolahan data menjadi sebuah informasi yang bermanfaat. Perkembangan tersebut menghasilkan volume data yang sangat besar dari beragam sektor, seperti ekonomi, industri, hingga dunia pendidikan. Di ranah pendidikan sendiri, penerapan teknologi informasi mampu menciptakan data yang terus bertambah setiap tahun, baik mengenai peserta didik selama proses belajar maupun data yang diperoleh hingga mereka menyelesaikan studi di sebuah lembaga pendidikan.(Desi et al., 2022). Pendidikan adalah kebutuhan utama yang harus dipenuhi oleh setiap individu untuk bisa bersaing di era globalisasi saat ini.(Andy Setiawan, 2019).

SMA Unggulan Karangsawo di Kecamatan Paciran, Kabupaten Lamongan, menghadapi tantangan seiring meningkatnya jumlah peserta didik setiap tahun. Penumpukan data siswa



Hak cipta: © 2025 oleh penulis.  
Diserahkan untuk kemungkinan  
publikasi akses terbuka  
berdasarkan syarat dan ketentuan  
lisensi Creative Commons  
Attribution (CC BY SA) (  
<https://creativecommons.org/licenses/by-sa/4.0/>)

menyulitkan pengelolaan dan pengambilan keputusan, terutama dalam menentukan kelas unggulan dan reguler. Selama ini, proses pengelompokan masih dilakukan secara manual sehingga memakan waktu, rawan kesalahan, dan kurang objektif. Untuk itu, diperlukan sistem yang lebih efektif dan efisien. Algoritma K-Means Clustering dapat menjadi solusi dengan mengelompokkan siswa berdasarkan kesamaan karakteristik, sehingga proses penentuan kelas lebih akurat, transparan, dan mudah dipantau.

Salah satu teknik yang dianggap efisien dalam pengelompokan data untuk penentuan kelas unggulan adalah metode K-Means Clustering. Metode ini bekerja dengan membagi data ke dalam sejumlah kelompok (cluster), di mana setiap data dalam satu cluster memiliki tingkat kemiripan yang tinggi, sedangkan perbedaan antarcluster relatif lebih besar (Solichin & Khairunnisa, 2020). Metode K-Means berorientasi pada titik pusat (centroid), mampu mengolah data besar secara cepat dan efisien, serta termasuk dalam teknik *unsupervised learning* dengan pendekatan partisi (Bu'ulolo & Purba, 2021). Menurut penelitian (Mauladi & Susilo, 2021), Metode K-Means Clustering telah diaplikasikan dalam berbagai bidang, termasuk pada kasus COVID-19 dan e-learning. Di Kabupaten Lamongan, teknik ini digunakan untuk mengelompokkan wilayah berdasarkan data penyebaran, seperti jumlah ODP, PDP, kasus positif, pasien sembuh, dan angka kematian. Hasil pengelompokan tersebut membentuk zonasi merah, kuning, dan hijau yang berfungsi sebagai dasar pemetaan sekaligus upaya penanggulangan.

Berdasarkan uraian permasalahan yang telah dijelaskan, penulis memilih untuk menggunakan metode data mining dengan algoritma K-Means Clustering dalam mengelompokkan data calon peserta didik baru di SMA Unggulan Karangsawo. Penerapan metode ini bertujuan untuk membagi siswa berdasarkan hasil seleksi, meliputi nilai tes akademik, tes membaca Al-Qur'an, wawancara, rata-rata ijazah, serta total nilai seleksi. Dengan adanya pengelompokan tersebut, proses penentuan Kelas Unggulan dan Kelas Reguler dapat dilakukan lebih mudah. Selain itu, penggunaan K-Means diharapkan mampu meningkatkan efektivitas dan efisiensi dalam pengolahan data.

## 2. Tinjauan Literatur

### 2.1. Peserta Didik

Peserta didik adalah individu yang berperan aktif dalam kegiatan pendidikan dengan tujuan mengembangkan kemampuan diri melalui interaksi dengan lingkungan dan berbagai sumber belajar (Amaliyah & Rahmat, 2021). Definisi peserta didik tidak hanya terbatas pada siswa yang berada di lingkungan sekolah formal, tetapi juga mencakup individu yang mengikuti pendidikan non-formal dan informal (Ahmad et al., 2023). Aspek peserta didik mencakup berbagai hal, antara lain kebutuhan individu, dorongan dalam belajar, pengalaman yang telah diperoleh, keterampilan sosial dan hubungan interpersonal, preferensi metode pembelajaran, serta penguasaan dalam menggunakan teknologi informasi dan komunikasi (Susilo & Rohman, 2019).

### 2.2 kelas Unggulan

Kelas unggulan merupakan salah satu bentuk dari *ability grouping class*. Istilah ini mengacu pada praktik pengelompokan yang dilakukan oleh guru, pihak sekolah, maupun pembuat kebijakan dengan tujuan menempatkan siswa ke dalam kelas atau sekolah tertentu berdasarkan tingkat kemampuan yang dimiliki.

### 2.3 Data Mining

Data mining adalah metode yang digunakan untuk menggali nilai serta pola tersembunyi dari suatu himpunan data. Dalam basis data yang berukuran besar, tersimpan berbagai informasi yang berpotensi untuk diolah menjadi pengetahuan baru. Melalui proses penambangan data, informasi yang relevan dapat diekstraksi secara sistematis dengan memanfaatkan teknik statistik, perhitungan matematis, hingga kecerdasan buatan yang bekerja secara semi-otomatis. Teknologi berbasis artificial intelligence dan machine learning juga berperan penting dalam menganalisis serta mengolah data tersebut. Sebagai cabang ilmu, data mining memberikan kontribusi luas di banyak bidang, seperti industri, prediksi cuaca, sektor keuangan, hingga aplikasi dalam sains dan teknologi yang terkait dengan kehidupan sehari-hari (Aroni & Adrian, 2016).

## 2.4 Clustering

Clustering adalah metode yang berfungsi untuk mengelompokkan data ke dalam beberapa kategori sesuai dengan karakteristik yang dimiliki. Data yang memiliki kesamaan sifat akan ditempatkan dalam cluster yang sama, begitu juga data dengan tingkat kemiripan paling tinggi akan tergabung pada kelompok yang serupa (Iqbal, 2019). Tahap pengelompokan dilakukan dengan tujuan untuk menelaah sekaligus memperoleh pemahaman yang lebih mendetail dan mendalam (Susiliwati, Dharmawan & Budi, 2025).

## 2.5 K-Means

Algoritma K-Means merupakan metode pengelompokan data yang digunakan untuk membagi sekumpulan data ke dalam dua atau lebih cluster (Indriyani & Irfiani, 2019). Clustering menggunakan metode kmeans biasanya dimulai dengan pengumpulan data. Jumlah klaster awal dan *Centroid* dapat dihitung jika data sudah dikumpulkan. Selanjutnya, jarak dari *Centroid* dihitung sampai jarak tidak berubah (Indriyani & Irfiani, 2019).

Proses algoritma *K-Means* melibatkan beberapa langkah utama yaitu :

- Masukkan data
- Menentukan jumlah k
- menginisialisasi *Centroid* secara random.
- Menghitung jarak antara setiap titik data dan *Centroid*.

Perhitungan jarak menggunakan rumus *Euclidean distance* :

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

X1 : data pertama (dari atribut pertama)

Y1 : *cluster Centroid* untuk atribut pertama

X2 : data kedua (dari atribut kedua)

Y2 : *Centroid/ pusat cluster* untuk atribut kedua

- Menentukan *Centroid* baru.

Perhitungan menggunakan rumus :

$$bj(t+1) = \frac{1}{NSj} \sum_{j \in Sj} aj \quad (2)$$

dimana :

$bj(t+1)$  : *Centroid* baru pada iterasi ke (t+1)

$NSj$  : banyak data pada *cluster Sj*.

## 2.6 Evaluasi BCV/WCV Ratio

Setelah algoritma K-Means menghasilkan pembagian data ke dalam beberapa cluster, langkah berikutnya adalah mengevaluasi hasil tersebut. Salah satu metode yang dapat digunakan adalah Sum Squared Error (SSE). Selain itu, kualitas clustering juga dapat dinilai dengan membandingkan nilai Between-Class Variation (BCV) dan Within-Class Variation (WCV) pada iterasi akhir, yang biasanya dinyatakan dalam bentuk rasio. Semakin tinggi nilai rasio yang diperoleh, semakin baik pula kualitas pengelompokan yang dihasilkan (Candra, 2019).

BCV merupakan rata-rata dari *Centroid*, sedangkan WCV adalah SSE itu sendiri. Rumus perhitungannya adalah sebagai berikut :

$$BCV = \sum_{i=1}^K n_i * \|m_i - \bar{m}\|^2 \quad (3)$$

Keterangan :

K = Jumlah *Cluster*

$n_i$  = Jumlah Data(anggota) Dalam *Cluster* ke-i

$m_i$  = *Centroid* Dari *Cluster* Ke-i

$\bar{m}$  = *Overall Means* yaitu Rata-rata seluruh data (tanpa melihat *cluster*)

$\|m_i - \bar{m}\|^2$  : Jarak kuadrat antara *Centroid* *cluster* ke-i dengan overall mean

$$WCV (SSE) = \sum_{j=1}^n \sum_{p \in ci} d(p, m_i)^2 \quad (4)$$

Keterangan :

$p_{ci}$  = Jumlah semua data

$k$  = Jumlah *cluster*

$p$  = *Cluster* jarak terdekat

$m_i$  = Jumlah anggota dari *cluster* ke- $i$

$$\text{Rasio} = \frac{BCV}{WCV}$$

Apabila nilai rasio yang didapat semakin kecil maka semakin bagus pula tingkat hasil dari akurasi *cluster*, menurut (ROUSSEEUW, 1990).

**Tabel 1.** Keterangan Nilai Rasio

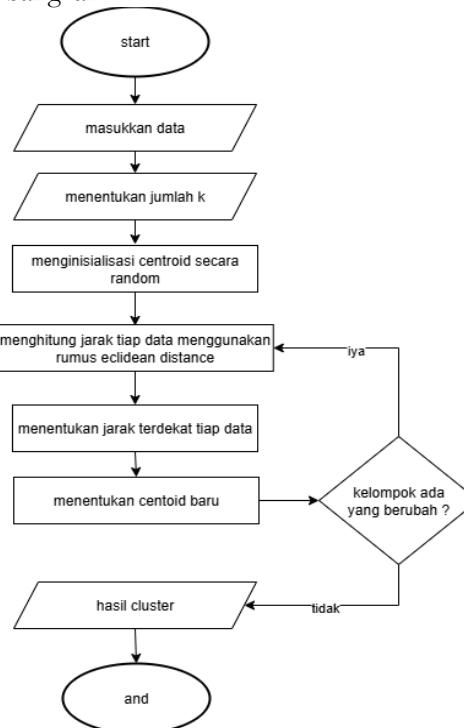
Nilai Rasio	Keterangan
$\leq 0,25$	Sangat baik
0,25 - 0,50	Baik
0,50 - 0,75	Kurang baik
0,75 - 1,00	Buruk

### 3. Metode

Penelitian ini menggunakan pendekatan kuantitatif dengan metode algoritma *K-Means Clustering*. Pendekatan ini digunakan untuk mengelompokkan data hasil seleksi peserta didik baru untuk menentukan kelas unggulan di SMA Unggulan Karangsawo ke dalam dua kelompok (*cluster*), yaitu kelas unggulan dan kelas reguler, berdasarkan data hasil seleksi seperti nilai akademik, nilai baca Al-Qur'an, nilai wawancara, nilai rata-rata ijaza, dan nilai total hasil seleksi. Metode *Clustering K-Means* dipilih karena kemampuannya dalam mengelompokkan data berdasarkan kemiripan atribut, serta tingkat efisiensi dan kecepatan dalam pemrosesan data berukuran besar.

#### 3.1. Flowchart K-Means

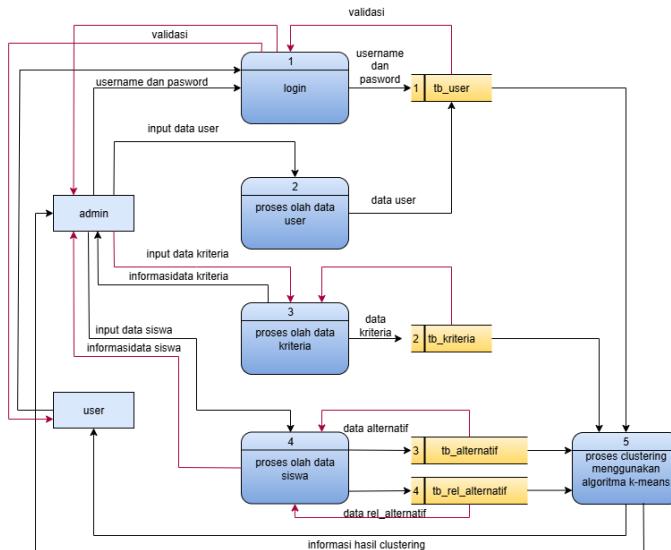
Flowchart metode *K-Means Clustering* digunakan untuk menggambarkan alur proses secara sistematis dan logis, mulai dari inisialisasi *Centroid* hingga konvergensi hasil *clustering*. Flowchart ini penting untuk membantu memahami bagaimana algoritma bekerja di dalam sistem yang dikembangkan.



**Gambar 1.** Flowchart Algoritma K-Means

### 3.2. DFD Level 1

DFD Level 1 menggambarkan proses-proses utama dalam sistem *Clustering Data Peserta Didik Baru Menggunakan Metode K-Means* di SMA Unggulan Karangsawo. Diagram ini menjelaskan bagaimana aliran data terjadi antara entitas eksternal (*User* dan *Admin*), proses-proses sistem, serta penyimpanan data dalam basis data.



**Gambar 2.** DFD Level 1

## 4. Hasil dan Pembahasan

Tahap implementasi merupakan proses realisasi dari hasil perancangan sistem yang telah disusun sebelumnya. Pada tahap ini, sistem mulai dibangun menjadi sebuah aplikasi yang dapat digunakan untuk mengelompokkan peserta didik baru berdasarkan hasil seleksi. Tujuannya adalah untuk membantu penentuan kelas unggulan secara objektif dan sistematis. Implementasi melibatkan penerapan logika algoritma ke dalam kode program, penyusunan antarmuka pengguna, serta integrasi fungsi-fungsi pendukung yang diperlukan.

### 4.1 Hasil Penelitian

Setelah menganalisis masalah dengan data hasil tes calon siswa sebelumnya, data diperlukan untuk menghitung nilai total dan menentukan apakah calon siswa dinyatakan lulus atau tidak. Untuk melakukan ini, metode *clustering K-Means* digunakan untuk menghasilkan kelompok yang dinyatakan lulus dan kelompok yang tidak lulus selama proses seleksi. Data berikut berasal dari daftar calon siswa SMA Unggulan Karangsawo:

**Tabel 2.** Data Siswa

No	Nama Siswa	Nilai Akademik	Nilai Baca Qur'an	Nilai Wawancara	Nilai Rata-Rata Ijazah	Nilai Rata-Rata Kriteria
1	SAFIRA KUSNUL NAFIK	90	80	90	80.86	85.215
2	UMMAROTUS SHOLIKHAH	90	85	85	97.57	89.3925
...	...	...	...	...	...	...
7	TSALITSA HARISMATIN ILFINA	88	85	85	71.68	82.42
8	DIANA PUTRI ANDIKA SARI	86	85	85	95.12	87.78
...	...	...	...	...	...	...
46	KAFA ANANDA MA'ARIF	86	80	70	89.21	81.3025

47	MARSYA NUR AINI	80	85	85	79.04	82.26
48	MOH. FARIKHUL IHSAN	80	85	80	85.08	82.52
49	MOHAMAD AINUR QOHAR	90	80	80	85.85	83.9625
...	...	...	...	...	...	...
52	QURROTUL AIN ARSY	90	70	70	92.48	80.62

Berdasarkan data siswa pada Tabel 1, tahap awal dalam proses clustering dengan metode K Means adalah menentukan pusat klaster (centroid) awal. Penentuan centroid awal dilakukan secara acak atau random dengan memilih beberapa data sebagai titik awal untuk setiap klaster, Pemilihan titik pusat awal dilakukan secara acak.

**Tabel 3.** Centroid

Cluster	Nilai Akademik	Nilai Baca Qu'an	Nilai Wawancara	Nilai Rata-Rata	Nilai Rata-Rata Kriteria
C1	90	80	90	80,86	85,215
C2	77	70	75	72,83	73,7075

Setelah menentukan centroid sebagai titik awal, langkah selanjutnya adalah menghitung jarak antara data dan centroid tersebut. Berdasarkan persamaan 2.1:

Perhitungan jarak dari data ke-1 terhadap pusat *Centroid* 1 adalah

$$d(1,1) = \sqrt{\{(90 - 90)^2 + (80 - 80)^2 + (90 - 90)^2 + (80,86 - 80,85)^2 + (85,215 - 85,215)^2\}} = 0$$

Perhitungan jarak dari data ke-1 terhadap pusat *Centroid* 2 adalah

$$d(1,2) = \sqrt{\{(90 - 77)^2 + (80 - 70)^2 + (90 - 75)^2 + (80,86 - 72,83)^2 + (85,215 - 73,7075)^2\}} = 26.28504$$

**Tabel 4.** Iterasi 1

No	C1	C2	Jarak Terdekat	JCT	JCT^2
1	0	26.28504	C1	0	0
2	18.61923	36.77073	C1	18.61923	346.6756063
...	...	...	...	...	...
7	12.08654	22.87422	C1	12.08654	146.084425
8	16.61105	33.1795	C1	16.61105	275.926825
...	...	...	...	...	...
46	22.3837	23.06487	C1	22.3837	501.0301563
47	12.72967	21.11183	C1	12.72967	162.044425
48	15.81365	22.06179	C1	15.81365	250.071425
49	11.24584	23.84713	C1	11.24584	126.4688563
...	...	...	...	...	...
52	22.3837	23.06487	C1	22.3837	501.0301563

Dari hasil perhitungan sebelumnya, diperoleh centroid terbaru untuk iterasi kedua. Nilai centroid ini ditentukan berdasarkan rata-rata data dalam masing-masing klaster.

**Tabel 5.** Centroid Baru

Cluster	Nilai Akademik	Nilai Baca Qu'an	Nilai Wawancara	Nilai Rata-Rata	Nilai Rata-Rata Kriteria
---------	----------------	------------------	-----------------	-----------------	--------------------------

C1	84.83333	82.55556	82.91667	84.09389	83.59986111
C2	79.75	72.1875	76.875	81.4775	77.5725

Perhitungan iterasi selanjutnya dilakukan menggunakan centroid baru dan mendapatkan hasil iterasi ke 2 sebagaimana di tunjukkan pada tabel berikut

**Tabel 6.** Iterasi 2

No	C1	C2	Jarak Terdekat	JCT	JCT^2
1	9.821692	19.92868	C1	9.821692	96.46563073
2	15.87987	27.09094	C1	15.87987	252.1702668
...	...	...	...	...	...
7	13.26047	20.43838	C1	13.26047	175.8400738
8	12.27703	23.65479	C1	12.27703	150.7253849
...	...	...	...	...	...
46	14.35913	14.86835	C1	14.35913	206.1847141
47	7.811128	16.06712	C1	7.811128	61.01371823
48	6.323123	14.54112	C1	6.323123	39.98187934
49	6.704304	15.35624	C1	6.704304	44.94768628
...	...	...	...	...	...
52	20.74568	16.95451	C2	16.95451	287.4555438

Berdasarkan Hasil perhitungan pada iterasi pertama dan kedua menunjukkan bahwa komposisi klaster masih mengalami perubahan. Dengan demikian, diperlukan iterasi lanjutan menggunakan *centroid* baru dari iterasi ketiga.

Dalam penelitian ini, proses iterasi berlangsung hingga iterasi ke-4. Pada iterasi ke-4 tidak terjadi perubahan atau perpindahan anggota klaster dari hasil iterasi ke-3. Adapun hasil analisis pengelompokan peserta didik baru berdasarkan data seleksi menunjukkan bahwa pada pusat *cluster* pertama (C1) terdapat 29 siswa, sedangkan pada pusat *cluster* kedua (C2) terdapat 23 siswa. Penetapan *cluster* awal didasarkan pada nilai setiap kriteria yang digunakan serta perbedaan rata-rata tertinggi dan terendah dari keseluruhan data yang dianalisis. Berdasarkan klasifikasi tersebut, *cluster* C1 merepresentasikan kelompok dengan rata-rata nilai yang lebih tinggi dan diputuskan sebagai kelas unggulan, sedangkan *cluster* C2 menggambarkan kelompok dengan rata-rata nilai lebih rendah yang kemudian ditempatkan pada kelas reguler.

#### 4.2 Evaluasi BCV/WCV Ratio

Setelah melakukan pengelompokan dengan algoritma *K-Means*, langkah selanjutnya adalah melakukan evaluasi terhadap hasil pengelompokan. Ini penting untuk mengetahui seberapa baik kualitas hasil pengelompokan.

Berikut adalah hasil perhitungan rasio BCV/WCV pada iterasi terakhir:

Rasio Besaran Antara BCV (Between Cluster Variation) dengan WCV (Within Cluster Variation).

Mencari Nilai BCV dengan persamaan 2.2 :

$$n_i \cdot \|m_i - \bar{m}\|^2 = \sqrt{n_i \cdot \{(m_{i1} - \bar{m}_1)^2 + (m_{i2} - \bar{m}_2)^2 + \dots + (m_{in} - \bar{m}_n)^2\}}$$

Overall mean dari kriteria seluruh data adalah (83.26923077, 79.36538462, 81.05769231, 83.28884615, 81.74528846)

Jumlah anggota C1 = 29, C2 = 23

*Centroid* dalam iterasi pertama yaitu :

$$C1 = (85, 84.09677, 83.22581, 84.17871, 84.12532258)$$

$$C2 = (80.714286, 72.38095, 77.85714, 81.97524, 78.23190476)$$

$$C1(n_i \cdot \|m_i - \bar{m}\|^2) = \sqrt{29 \cdot \{(85 - 83.26923077)^2 + (83.22581 - 79.36538462)^2 + (84.17871 - 81.05769231)^2 + (84.12532258 - 83.28884615)^2 + (81.97524 - 81.74528846)^2\}} = 1015.6237$$

$$C2(n_i, \|m_i, \bar{m}\|^2) = \sqrt{23 \cdot \{(80.714286 - 83.26923077)^2 + (83.22581 - 79.36538462)^2 + (77.85714 - 81.05769231)^2 + (81.97524 - 83.28884615)^2 + (78.23190476 - 81.74528846)^2\}} = 1831.3287$$

$$\text{BCV} = C1 + C2 = 1015.6237 + 1831.3287 = 2890.9524$$

$$\text{WCV} = 6845.502281$$

$$\text{Besar Rasio} = \text{BCV}/\text{WCV} = 2890.9524 / 6845.502281 = 0.4223141$$

$$\text{Rasio sekarang} = 0.4223$$

$$\text{Rasio sebelumnya} = 0.3949$$

**Tabel 7.** BCV/WCV Ratio

iterasi	WCV	BCV	Nilai Ra-tio	Keterangan
1	12477.7074	10573.171	0.8474	Rasio awal tinggi
2	7011.5443	2769.1711	0.3949	Rasio Menurun
3	6845.5023	2890.9524	0.4223	Rasio sedikit meningkat
4	6845.5023	2890.9524	0.4223	Rasio stabil

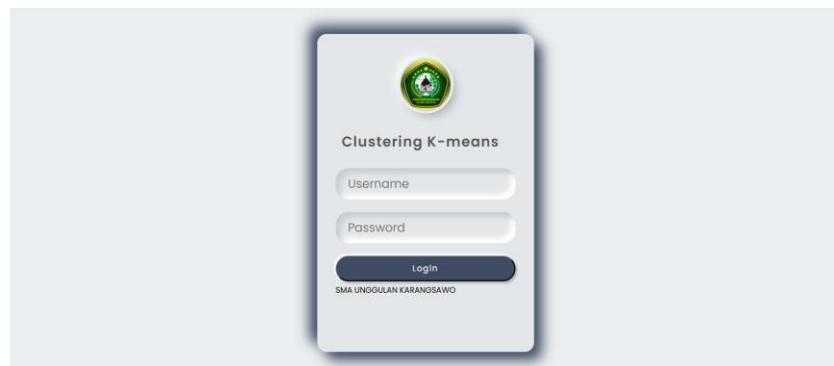
Catatan: Iterasi ke-4 tidak mengalami perubahan dari iterasi sebelumnya, menandakan bahwa algoritma telah konvergen.

Berdasarkan tabel 2.1 tersebut diketahui bahwa terdapat perbandingan antara BCV dan WCV yang mendapatkan nilai rasio akhir (0.4223) berada dalam kategori baik. Hal ini menunjukkan bahwa hasil *clustering* pada iterasi terakhir cukup baik dalam mengelompokkan data peserta didik baru berdasarkan kriteria yang digunakan, seperti nilai tes akademik, TPA, wancara, dan rata-rata ijazah.

### 4.3 Implementasi Halaman

#### 1. Halaman Login

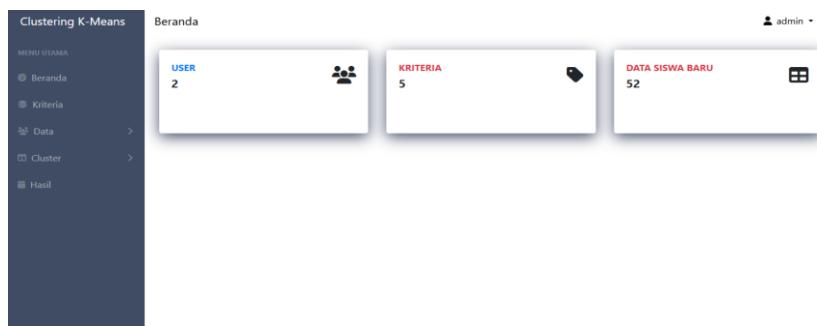
Halaman Login merupakan akses utama ke dalam sistem, di mana pengguna harus memasukkan username dan password untuk autentikasi. Fitur ini memastikan hanya akun sah yang dapat menggunakan sistem. Setelah berhasil login, pengguna akan diarahkan ke dashboard. Tampilan halaman Login ditunjukkan pada gambar berikut.



**Gambar 3.** Halaman Login

#### 2. Halaman Beranda

Halaman beranda merupakan halaman utama yang pertama diakses pengguna setelah login. Pada halaman ini ditampilkan ringkasan mengenai metode K-Means. Tampilan halaman beranda ditunjukkan pada gambar berikut



Gambar 4. Halaman Beranda

### 3. Halaman Perhitungan

Halaman ini menunjukkan proses perhitungan menggunakan algoritma k. Ini berarti bahwa halaman akan menampilkan perhitungan untuk semua data siswa dari iterasi pertama sampai iterasi terakhir.

Pusat Centroid pada Iterasi 1						
Nama Cluster	Nilai Akademik	Nilai Baca Al-Quran	Nilai Wawancara	Nilai Rata-Rata Ijazah	Nilai Rata-Rata Kriteria	
Cluster 1	74.0000	85.0000	80.0000	83.2300	80.5975	
Cluster 2	90.0000	85.0000	85.0000	97.5700	89.3925	

Alternatif	Nilai Akademik	Nilai Baca Al-Quran	Nilai Wawancara	Nilai Rata-Rata Ijazah	Nilai Rata-Rata Kriteria	Centroid Cluster 1	Centroid Cluster 2	Cluster
A001	90.0000	80.0000	90.0000	80.8600	85.2150	20.2067	18.6192	Cluster 2
A002	90.0000	85.0000	85.0000	97.5700	89.3925	23.7633	0.0000	Cluster 2
A003	89.0000	85.0000	85.0000	91.2500	87.5575	19.0526	6.6755	Cluster 2
A004	84.0000	85.0000	90.0000	87.3600	88.5800	15.9200	13.8567	Cluster 2

Gambar 5. Halaman Perhitungan

## 4.4 Pembahasan

Penelitian ini berfokus pada penerapan algoritma K-Means Clustering untuk membantu pengelompokan calon peserta didik baru di SMA Unggulan Karangsawo. Sistem berbasis web yang dikembangkan mengelompokkan siswa ke dalam dua kategori, yaitu Kelas Unggulan dan Kelas Reguler, dengan mempertimbangkan empat kriteria utama: nilai tes akademik, kemampuan membaca Al-Qur'an, hasil wawancara, serta rata-rata nilai ijazah. Proses klasterisasi dilakukan secara bertahap menggunakan perhitungan jarak Euclidean, di mana posisi centroid terus diperbarui hingga mencapai titik konvergensi.

Hasil akhir menunjukkan terbentuknya dua kelompok utama. Cluster 1 (C1) terdiri dari 29 siswa dengan rata-rata nilai lebih tinggi, sehingga digolongkan sebagai Kelas Unggulan. Sementara itu, Cluster 2 (C2) berisi 23 siswa dengan nilai rata-rata lebih rendah dan ditempatkan pada Kelas Reguler. Temuan ini menggambarkan metode yang objektif dalam mengukur performa akademik berdasarkan indikator kuantitatif.

Jika dibandingkan dengan penelitian (Widiastuti & Azzat 2022) yang memperoleh rasio evaluasi BCV/WCV sebesar 0,0017 dengan kualitas data yang baik, penelitian ini menghasilkan rasio 0,4233. Nilai tersebut menunjukkan kualitas klaster yang cukup tinggi serta klasifikasi yang lebih stabil dan logis. Dengan demikian, hasil penelitian dapat dijadikan dasar yang kuat untuk mendukung keputusan penempatan kelas di lingkungan sekolah.

## 5. Kesimpulan Dan Saran

Berdasarkan hasil penelitian, dapat disimpulkan bahwa implementasi algoritma K-Means Clustering dalam pengelompokan data seleksi peserta didik baru di SMA Unggulan Karangsawo berhasil dilakukan dengan baik. Sistem berbasis web mampu mengelompokkan 52 calon peserta didik menjadi dua kelompok, yaitu 29 siswa Kelas Unggulan dan 23 siswa Kelas Reguler, melalui perhitungan jarak Euclidean hingga pusat klaster stabil. Hasil klasterisasi menunjukkan tingkat keakuratan yang baik dengan rasio BCV/WCV sebesar

0,4223, yang menandakan pemisahan kedua kelompok cukup jelas dan sesuai karakteristik data.

Penelitian selanjutnya diharapkan dapat menerapkan pendekatan dan teknik yang berbeda sehingga memungkinkan adanya perbandingan efektivitas antar metode untuk mengetahui hasil pengelompokan yang paling optimal. Selain itu, penelitian di masa mendatang juga disarankan menggunakan jumlah data yang lebih besar agar hasil klasterisasi dapat memberikan gambaran yang lebih representatif serta meningkatkan keandalan analisis.

## Referensi

- [1] Amaliyah, A., & Rahmat, A. (2021). Pengembangan potensi diri peserta didik melalui proses pendidikan. *Attadib: Journal of Elementary Education*, 5(1), 28–37. <https://doi.org/10.32507/attadib.v5i1.926>
- [2] Ahmad, A., Tiara Sari, A. J., Wardana, A. H., Rosyid, M. N. I., Widianto, E., & Rasyad, A. (2023). Tren perkembangan pendidikan non-formal. *Jurnal Pendidikan (Teori dan Praktik)*, 7(2), 76–82. <https://doi.org/10.26740/jp.v7n2.p76-82>
- [3] Andy Setiawan, M. (2019). Sistem pendukung keputusan penerimaan siswa baru di SMK Negeri 2 Blitar menggunakan metode TOPSIS berbasis web. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 3(2), 53–58. <https://doi.org/10.36040/jati.v3i2.859>
- [4] Asroni, A., & Adrian, R. (2016). Penerapan metode K-Means untuk clustering mahasiswa berdasarkan nilai akademik dengan Weka interface studi kasus pada Jurusan Teknik Informatika UMM Magelang. *Semesta Teknika*, 18(1), 76–82. <https://doi.org/10.18196/st.v18i1.708>
- [5] Bu'ulolo, E., & Purba, B. (2021). Algoritma clustering untuk membentuk cluster zona penyebaran Covid-19. *Digital Zone: Jurnal Teknologi Informasi dan Komunikasi*, 12(1), 59–67. <https://doi.org/10.31849/digitalzone.v12i1.6572>
- [6] Candra, F. (2019). *Penerapan algoritma clustering K-Means untuk menentukan potensi wilayah pendistribusian barang* [Skripsi, Universitas Islam Negeri Sultan Syarif Kasim Riau].
- [7] Desi, E., Aliyah, S., Lestari, S., & Dari, W. (2022). Implementasi algoritma K-Means untuk penerimaan siswa baru di SMANPAS berdasarkan nilai rapor dan hasil tes. *IT (Informatic Technique) Journal*, 10(1), 1–10. <https://doi.org/10.22303/it.10.1.2022.01-10>
- [8] Indriyani, F., & Irfiani, E. (2019). Clustering data penjualan pada toko perlengkapan outdoor menggunakan metode K-Means. *JUITA: Jurnal Informatika*, 7(2), 109–115. <https://doi.org/10.30595/juita.v7i2.5529>
- [9] Iqbal, M. (2019). Klasterisasi data jamaah umroh pada Auliya Tour & Travel menggunakan metode K-Means clustering. *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, 5(2), 97–104. <https://doi.org/10.33330/jurteksi.v5i2.352>
- [10] Mauladi, K. F., & Susilo, P. H. (2021). Klasterisasi virus Covid-19 di wilayah Kabupaten Lamongan dengan metode K-Means clustering. *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, 6(2), 325–335. <https://doi.org/10.29100/jipi.v6i2.1999>
- [11] Rousseeuw, L. K., & P. J. (1990). *Finding groups in data: An introduction to cluster analysis*. Wiley.
- [12] Solichin, A., & Khairunnisa, K. (2020). Klasterisasi persebaran virus Corona (Covid-19) di DKI Jakarta menggunakan metode K-Means. *Fountain of Informatics Journal*, 5(2), 52–60. <https://doi.org/10.21111/fij.v5i2.4905>
- [14] Susilo, P. H., & Rohman, M. G. (2019, September 19). *Efektivitas sistem pembelajaran online sebagai media pembelajaran berbasis aplikasi web di era milenial*. Seminar Nasional Sistem Informasi 2019, Fakultas Teknologi Informasi – UNMER Malang, 2038–2044. ISSN 2598-0076.
- [15] Susiliwati, A. G., Dharmawan, J., & Budi, A. S. (2025). Evaluasi Kinerja Dosen Berbasis Kepuasan Mahasiswa dengan Metode K-Means. *Jurnal teknika*, 17(1), 51-58
- [16] Widiastuti, N. A., & Azzat, N. N. (2022). Evaluasi Jarak Centroid Distance On High School Mapping For Location Determination of New Students Promotion Using K-Means Algorithm. *Jurnal Disprotek*, 53-64.