



## PERBANDINGAN KINERJA ALGORITMA K-MEANS DAN DBSCAN DALAM PENGELOMPOKAN DATA NILAI KELAS VIII A PADA SMPN 01 PALEMBANG

Danisa Enjelika<sup>1</sup>, Nabila Putri Hariani<sup>2</sup>, Tata Sutabri<sup>3</sup>

<sup>1,2,3</sup> Prodi Teknik Informatika, Universitas Bina Darma

E-mail : [danisabae200@gmail.com](mailto:danisabae200@gmail.com), [nabilaputrihariani@gmail.com](mailto:nabilaputrihariani@gmail.com), [tatasutabri@gmail.com](mailto:tatasutabri@gmail.com)

### Article Info

#### Article history:

Received Maret 28, 2025

Revised April 13, 2025

Accepted April 28, 2025

#### Keywords:

Clustering

K-Means

DBSCAN

Student Scores

Remedial

Data Analysis

### ABSTRACT

*This study aims to group 30 students of class VIII A based on the scores of three main subjects, namely Mathematics, Indonesian, and English, using the K-Means and DBSCAN clustering methods. Determination of the optimal number of clusters in K-Means is carried out by analyzing SSE and Silhouette Score, resulting in three clusters of student abilities: high, medium, and low. DBSCAN divides students into two clusters, namely remedial and non-remedial. The clustering results help teachers identify groups of students who need special attention and develop more effective learning strategies. The application of this clustering technique has the potential to increase the efficiency of the learning process and students' academic success.*

### Corresponding Author:

Danisa Enjelika,

Universitas Bina Darma

Jl. Jenderal Ahmad Yani No.12, Sei Kambing C II, 20 Ilir D. III, Kec. Ilir Timur I, Kota Palembang, Sumatera Selatan 30113, Indonesia

Email: [danisabae200@gmail.com](mailto:danisabae200@gmail.com)



### Abstrak

Penelitian ini bertujuan untuk mengelompokkan 30 siswa kelas VIII A berdasarkan nilai tiga mata pelajaran utama, yaitu Matematika, Bahasa Indonesia, dan Bahasa Inggris, menggunakan metode clustering K-Means dan DBSCAN. Penentuan jumlah kluster optimal pada K-Means dilakukan dengan analisis SSE dan Silhouette Score, menghasilkan tiga kluster kemampuan siswa: tinggi, sedang, dan rendah. DBSCAN membagi siswa ke dalam dua kluster, yaitu remedial dan tidak remedial. Hasil clustering membantu guru dalam mengidentifikasi kelompok siswa yang membutuhkan perhatian khusus dan menyusun strategi pembelajaran yang lebih efektif. Penerapan teknik clustering ini berpotensi meningkatkan efisiensi proses pembelajaran dan keberhasilan akademik siswa.

**Kata Kunci:** Clustering, K-Means, DBSCAN, Nilai Siswa, Remedial, Analisis Data

## Pendahuluan

Di era digital saat ini, pemanfaatan teknologi informasi dalam dunia pendidikan semakin berkembang pesat. Salah satu penerapan teknologi tersebut adalah dalam analisis data nilai siswa untuk mendukung proses pengambilan keputusan akademik. Pengelompokan data nilai siswa dapat membantu pihak sekolah dalam mengidentifikasi pola prestasi, menentukan strategi pembelajaran yang tepat, serta memberikan intervensi yang sesuai bagi siswa. Metode clustering, sebagai bagian dari data mining, menjadi salah satu teknik yang efektif untuk melakukan pengelompokan data berdasarkan kesamaan karakteristik.

Algoritma K-Means dan DBSCAN (Density-Based Spatial Clustering of Applications with Noise) merupakan dua metode clustering yang sering digunakan dalam analisis data. K-Means bekerja dengan membagi data ke dalam sejumlah kluster berdasarkan kedekatan terhadap centroid, sedangkan DBSCAN mengelompokkan data berdasarkan kepadatan titik data dalam ruang. Algoritma kedua ini memiliki kelebihan dan kekurangan masing-masing, tergantung pada karakteristik data yang dianalisis.

Beberapa penelitian telah membandingkan kinerja K-Means dan DBSCAN dalam berbagai konteks. Hasibuan dkk. (2024) melakukan penelitian di SMP Negeri 3 Panyabungan untuk mengelompokkan siswa terbaik berdasarkan kriteria tertentu. Hasil penelitian menunjukkan bahwa K-Means menghasilkan kluster yang lebih bersih dan terstruktur dengan nilai Silhouette Score sebesar 0,5697, dibandingkan dengan DBSCAN yang menghasilkan nilai 0,2580.

Dalam konteks pembelajaran fisika, Kertanah dkk. (2024) membandingkan kedua algoritma tersebut dalam mengelompokkan hasil belajar kognitif siswa. K-Means menunjukkan performa yang lebih baik dengan Silhouette Score sebesar 0,43, sementara DBSCAN memperoleh nilai 0,39. Penelitian ini menegaskan bahwa K-Means lebih efektif dalam mengelompokkan data nilai siswa yang memiliki distribusi yang jelas.

Namun, dalam beberapa kasus, DBSCAN menunjukkan keunggulan, terutama dalam identifikasi kluster dengan kepadatan yang berbeda-beda dan menangani data yang mengandung noise. Dewi dkk. (2021) dalam penelitiannya tentang pengelompokan status desa di Jawa Tengah menemukan bahwa DBSCAN lebih tepat digunakan dibandingkan K-Means, karena mampu mengelompokkan data dengan kepadatan yang bervariasi.

Selain itu, Paramita dan Hariguna (2024) dalam studi mereka tentang segmentasi pelanggan e-commerce menemukan bahwa DBSCAN mampu mengidentifikasi kluster dengan kepadatan yang berbeda dan menangani noise

dengan lebih baik, meskipun K-Means lebih cocok untuk segmentasi yang memerlukan ukuran kluster yang seimbang .

Penelitian ini bertujuan untuk membandingkan kinerja algoritma K-Means dan DBSCAN dalam pengelompokan data nilai siswa Kelas VIII A di SMPN 01 Palembang. Dengan menganalisis kelebihan dan kekurangan masing-masing algoritma dalam konteks data nilai siswa, diharapkan penelitian ini dapat memberikan rekomendasi metode clustering yang paling sesuai untuk digunakan dalam lingkungan pendidikan, khususnya dalam upaya meningkatkan kualitas pembelajaran dan prestasi siswa.

## Metodologi

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen komparatif untuk membandingkan kinerja algoritma K-Means dan DBSCAN dalam pengelompokan data nilai siswa Kelas VIII A SMPN 01 Palembang. Data yang digunakan adalah nilai akademik dari 30 siswa , yang mencakup beberapa mata pelajaran inti seperti Matematika, Bahasa Indonesia, dan IPA.

Proses pemrosesan data dilakukan menggunakan platform Google Colaboratory (Google Colab) , karena mendukung pemrosesan data dengan Python serta berbagai pustaka data science seperti *Pandas* , *Matplotlib* , *Scikit-learn* , dan *Seaborn* . Data siswa diinput dalam format file CSV dan dibaca menggunakan library *Pandas*.

Langkah-langkah penelitian dimulai dari pra-pemrosesan data (normalisasi dan pembersihan data), kemudian dilakukan proses clustering menggunakan algoritma K-Means dan DBSCAN. Setelah proses clustering, hasil kelompok dari kedua algoritma dibandingkan menggunakan metrik evaluasi Silhouette Score dan visualisasi kluster. Dengan membandingkan hasil kinerja kedua algoritma, penelitian ini bertujuan untuk mengetahui algoritma mana yang lebih efektif dalam mengelompokkan data nilai siswa yang bersifat heterogen.

## Pembahasan

### 1. Proses Clustering dengan K-Means

```
[ ] import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.cluster import KMeans
from sklearn.preprocessing import StandardScaler
```

Gambar 1. Import Perpustakaan

Langkah pertama dalam menjalankan proyek analisis data di Google Colab adalah mengimpor pustaka-pustaka penting. Pustaka seperti `pandas` digunakan untuk manipulasi data, `numpy` untuk operasi numerik, `matplotlib.pyplot` dan `seaborn` untuk visualisasi, serta `KMeans` dan `StandardScaler` dari `sklearn` untuk klasterisasi dan normalisasi data.



Gambar 2. Import File

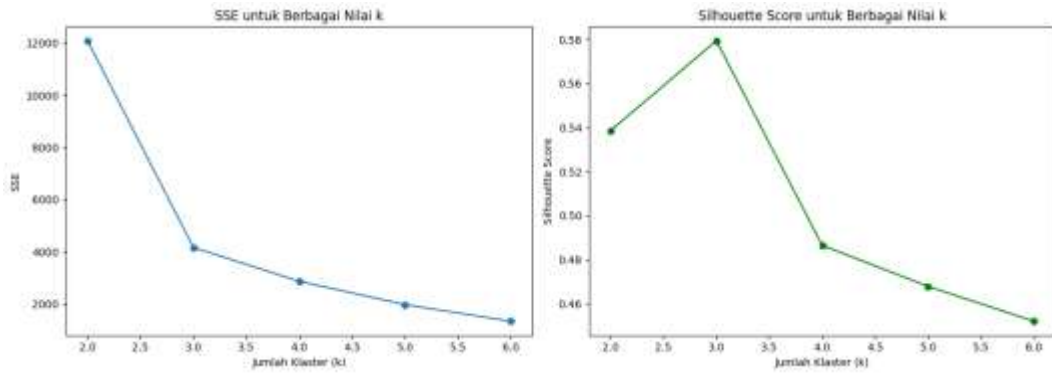
Selanjutnya, dilakukan proses import atau pembacaan data dari file Excel ke dalam Google Colab, seperti yang ditunjukkan pada gambar di atas. Proses ini penting untuk memuat data agar dapat dianalisis lebih lanjut menggunakan berbagai teknik analisis dan visualisasi data.



Gambar 3. Code SSE dan Silhouette Score

Dalam analisis *clustering* menggunakan algoritma seperti K-Means, salah satu tantangan utamanya adalah menentukan jumlah kluster (kelompok) yang optimal, yaitu nilai  $k$  yang paling tepat. Untuk membantu menentukan nilai  $k$  yang ideal, dua metrik evaluasi yang sering digunakan adalah:

1. SSE (*Sum of Squared Errors*)
2. Silhouette Score



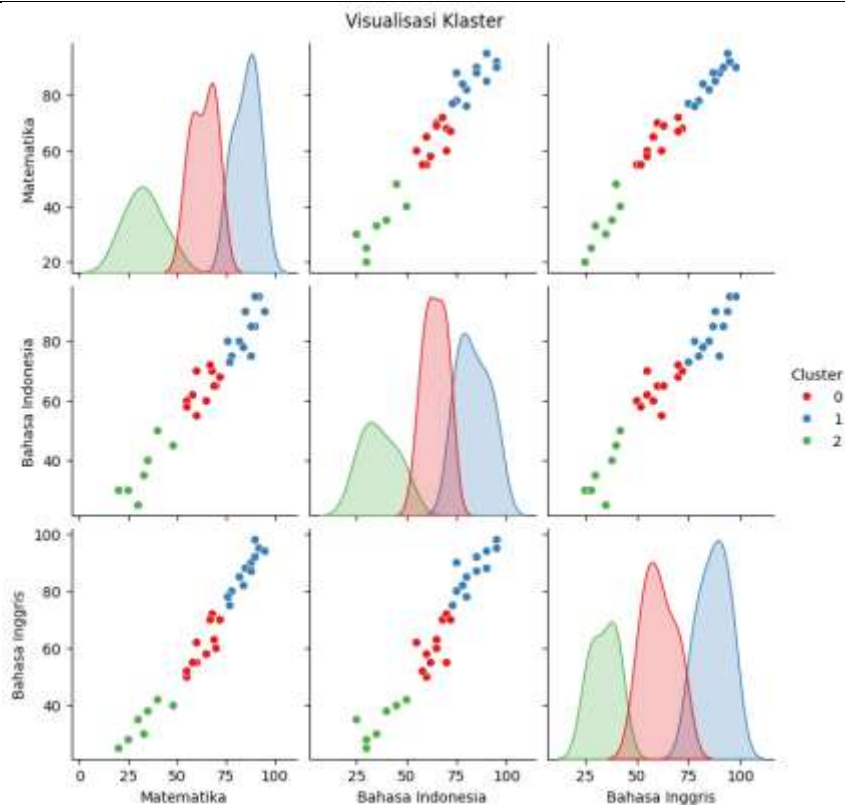
Gambar 4. Hasil SSE dan Silhouette Score

Berdasarkan hasil analisis grafik SSE dan Silhouette Score terhadap jumlah kluster ( $k$ ), dapat disimpulkan bahwa nilai  $k=3$  merupakan jumlah kluster yang optimal untuk data nilai siswa kelas VIII A di SMPN 01 Palembang. Grafik SSE menunjukkan penurunan tajam dari  $k=2$  ke  $k=3$ , kemudian penurunan melambat, menandakan adanya elbow point di  $k=3$ . Sementara itu, nilai Silhouette Score tertinggi juga terjadi pada  $k=3$  (sekitar 0,58), menunjukkan pemisahan kluster yang baik. Maka, siswa dapat dikelompokkan menjadi tiga kategori: nilai tinggi, sedang, dan rendah. Hasil ini dapat membantu guru dalam menyusun strategi pembelajaran yang lebih tepat sasaran pada kelas viii A.

```
[ ] kmeans_final = KMeans(n_clusters=3, random_state=41, n_init=10)
    df['Cluster'] = kmeans_final.fit_predict(X)

[ ] sns.pairplot(df, hue='Cluster', vars=['Matematika', 'Bahasa Indonesia', 'Bahasa Inggris'], palette='Set1')
    plt.suptitle('Visualisasi Kluster', y=1.01)
    plt.show()
```

Gambar 5. Code Visualisasikan kluster



Gambar 6. Hasil Visualisasikan kluster

Gambar tersebut menampilkan hasil pengelompokan siswa kelas VIII A berdasarkan nilai tiga mata pelajaran: Matematika, Bahasa Indonesia, dan Bahasa Inggris menggunakan algoritma *clustering* (kemungkinan besar K-Means). Terdapat tiga kluster yang terbentuk, yang diwakili oleh warna:

- Hijau (Cluster 2) → Siswa dengan nilai rendah
- Merah (Cluster 0) → Siswa dengan nilai sedang
- Biru (Cluster 1) → Siswa dengan nilai tinggi

Visualisasi ini menampilkan pola yang cukup jelas dalam pengelompokan. Kluster biru mendominasi wilayah nilai tinggi di semua mata pelajaran, sedangkan kluster hijau hadir pada nilai rendah. Kluster merah berada di antara keduanya. Hal ini menunjukkan bahwa model mampu membedakan dengan baik tiga kelompok kemampuan siswa berdasarkan distribusi nilai mereka.

Distribusi data pada grafik diagonal (density plot) juga menunjukkan bahwa nilai-nilai siswa cenderung terkelompok secara alami ke dalam tiga bagian, sesuai dengan visualisasi kluster yang terbentuk. Misalnya, dalam mata pelajaran Matematika, terlihat bahwa sebaran nilai siswa dari kluster hijau berada di rentang

nilai 20–45, klaster merah pada 50–70, dan klaster biru pada nilai di atas 75. Pola yang sama juga muncul pada mata pelajaran Bahasa Indonesia dan Bahasa Inggris.

Selain itu, scatterplot antar variabel (di luar diagonal) menampilkan korelasi positif antar mata pelajaran, yang berarti siswa yang memiliki nilai tinggi di satu mata pelajaran cenderung juga memiliki nilai tinggi di mata pelajaran lainnya. Ini memperkuat hasil klusterisasi bahwa siswa dapat dibedakan berdasarkan tingkat keseluruhan prestasi akademik, bukan hanya satu mata pelajaran saja.

## 2. DBSCAN dalam Pengelompokan Data

```
# Import library
import pandas as pd
from sklearn.cluster import DBSCAN
from sklearn.preprocessing import StandardScaler
```

Gambar 7. Import Perpustakaan DBSCAN

Langkah pertama dalam menjalankan proyek analisis data di Google Colab adalah mengimpor pustaka-pustaka penting.

```
[ ]
# Buat Dataframe dari data siswa (Matematika, Bahasa Indonesia, Bahasa Inggris)
data = {
    'Siswa': ['A', 'B', 'C', 'D', 'E', 'F', 'G', 'H', 'I', 'J',
             'K', 'L', 'M', 'N', 'O', 'P', 'Q', 'R', 'S', 'T',
             'U', 'V', 'W', 'X', 'Y', 'Z', 'AA', 'AB', 'AC', 'AD'],
    'Matematika': [88, 78, 68, 85, 38, 72, 90, 65, 28, 78,
                  82, 55, 88, 92, 48, 35, 68, 77, 25, 68,
                  95, 88, 48, 67, 55, 76, 33, 98, 84, 58],
    'Bahasa_Indonesia': [75, 65, 78, 98, 25, 68, 85, 68, 38, 75,
                        88, 68, 78, 95, 45, 68, 55, 73, 38, 65,
                        98, 85, 58, 72, 58, 88, 35, 95, 78, 62],
    'Bahasa_Inggris': [98, 68, 55, 88, 35, 78, 92, 58, 25, 88,
                      85, 58, 72, 95, 48, 38, 62, 75, 28, 63,
                      94, 87, 43, 78, 52, 78, 38, 98, 82, 55]
}

df = pd.DataFrame(data)
```

Gambar 9. Masukan Nilai

```
[ ] X = df[['Matematika', 'Bahasa_Indonesia', 'Bahasa_Inggris']].values

[ ] scaler = StandardScaler()
   X_scaled = scaler.fit_transform(X)

[ ] dbSCAN = DBSCAN(eps=0.7, min_samples=3)
   clusters = dbSCAN.fit_predict(X_scaled)

S [ ] df['cluster'] = clusters
```

Gambar 10. DBSCAN

Gambar 10 menampilkan proses implementasi algoritma DBSCAN (Density-Based Spatial Clustering of Applications with Noise) untuk melakukan pengelompokan data nilai siswa berdasarkan tiga mata pelajaran, yaitu Matematika, Bahasa Indonesia, dan Bahasa Inggris. Pertama, data nilai dari ketiga mata pelajaran tersebut diambil dan disimpan ke dalam variabel X. Kemudian, data dinormalisasi menggunakan StandardScaler agar memiliki skala yang seragam, yang penting untuk akurasi DBSCAN. Setelah itu, algoritma DBSCAN diterapkan dengan parameter  $\text{eps}=0.7$  dan  $\text{min\_samples}=3$ , yang berarti jarak maksimum antar titik dalam satu kluster adalah 0.7 dan minimal tiga titik diperlukan untuk membentuk satu kluster. Terakhir, hasil klusterisasi disimpan ke dalam kolom baru bernama 'Cluster' pada DataFrame. Proses ini berguna untuk mengelompokkan siswa secara otomatis berdasarkan pola kemiripan nilai akademik mereka, sekaligus dapat mendeteksi siswa yang tidak termasuk dalam kelompok manapun (outlier).

	Siswa	Matematika	Bahasa_Indonesia	Bahasa_Inggris	Cluster
0	A	88	75	90	0
1	B	70	65	60	0
2	C	60	70	55	0
3	D	85	90	88	0
4	E	30	25	35	1
5	F	72	68	70	0
6	G	90	85	92	0
7	H	65	60	58	0
8	I	20	30	25	1
9	J	78	75	80	0
10	K	82	80	85	0
11	L	55	60	50	0
12	M	68	70	72	0
13	N	92	95	95	0
14	O	48	45	40	1
15	P	35	40	38	1
16	Q	60	55	62	0
17	R	77	75	75	0
18	S	25	30	28	1
19	T	69	65	63	0
22	W	40	50	42	1
23	X	67	72	70	0
24	Y	55	58	52	0
25	Z	76	80	78	0
26	AA	33	35	30	1
27	AB	90	95	98	0
28	AC	84	78	82	0
29	AD	58	62	55	0

Gambar 11. Hasil Cluser

```

# Visualisasi hasil DBSCAN dengan PCA (2 dimensi)
pca = PCA(n_components=2)
X_pca = pca.fit_transform(X_scaled)

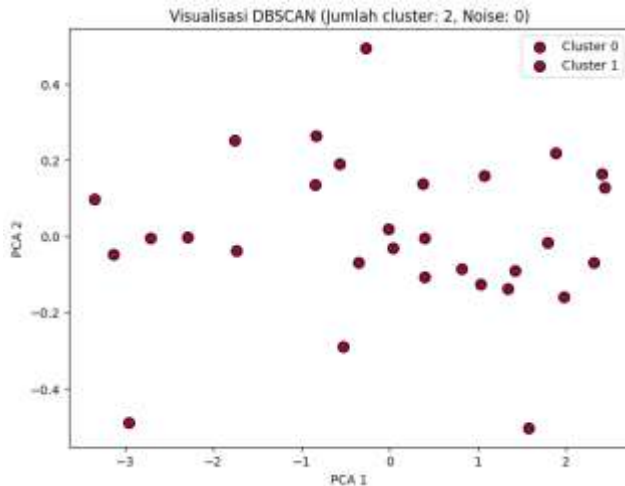
plt.figure(figsize=(8,8))
unique_labels = set(clusters)
colors = [plt.cm.Spectral(each) for each in range(len(unique_labels))]

for k, col in zip(unique_labels, colors):
    if k == -1:
        # warna untuk noise
        col = [0, 0, 0, 1]
    class_member_mask = (clusters == k)
    xy = X_pca[class_member_mask]
    plt.scatter(xy[:, 0], xy[:, 1], c=[col], label='Cluster (k)' if k != -1 else 'Noise', edgecolors='k', s=70)

plt.title(f'Visualisasi DBSCAN (jumlah cluster: {n_clusters}, noise: {(clusters == -1).sum()})')
plt.xlabel('PCA 1')
plt.ylabel('PCA 2')
plt.legend()
plt.show()

```

Gambar 12. Code DBSCAM dengan PCA



Gambar 13. Hasil DBSCAM dengan PCA

Hasil visualisasi menunjukkan bahwa algoritma DBSCAN membentuk 2 kluster tanpa ada data yang termasuk noise (outlier). Artinya, seluruh siswa berhasil dikelompokkan ke dalam dua kelompok nilai yang memiliki pola berbeda. Visualisasi menggunakan PCA memudahkan pengamatan distribusi kluster dalam dua dimensi. Hasil ini bisa dimanfaatkan guru untuk menyesuaikan strategi pembelajaran sesuai karakteristik tiap kluster siswa.

```
[ ] # Buat kolom Keterangan sesuai cluster
df['Keterangan'] = df['cluster'].apply(lambda x: 'Remedial' if x == 1 else ('Tidak Remedial' if x == 0 else 'Noise'))

# Tampilkan tabel dengan kolom cluster dan keterangan di sebelahnya
print(df[['Siswa', 'Matematika', 'Bahasa_Indonesia', 'Bahasa_Ingggris', 'Cluster', 'Keterangan']])
```

Gambar 14. Keterangan Cluster

Selanjutnya, tambahkan tabel keterangan untuk mempermudah penjelasan mengenai dua kluster. Jika nilai kluster adalah 1, maka keterangan yang diberikan adalah 'Remedial', sedangkan jika nilai kluster adalah 0, maka keterangan yang diberikan adalah 'Tidak Remedial'.

```

cluster_1 = df[df['Cluster'] == 1].copy()

# Tambahkan kolom 'Keterangan' dengan nilai 'Remedial'
cluster_1['Keterangan'] = 'Remedial'

# Tampilkan tabel cluster 1 dengan keterangan
print(cluster_1[['Siswa', 'Matematika', 'Bahasa_Indonesia', 'Bahasa_Ingggris', 'Cluster', 'Keterangan']])

```

Siswa	Matematika	Bahasa_Indonesia	Bahasa_Ingggris	Cluster	Keterangan
4	E	38	25	35	1 Remedial
8	I	28	38	25	1 Remedial
14	O	48	45	48	1 Remedial
15	P	35	48	38	1 Remedial
18	S	25	38	28	1 Remedial
22	N	48	58	42	1 Remedial
25	AA	33	35	38	1 Remedial

Gambar 15. Nilai Cluster 1

Berdasarkan keterangan di atas nilai Cluster 1 di miliki oleh beberapa siswa yang nilai di bawah 60.

```

| | # Hitung jumlah berdasarkan kolom Keterangan
jumlah_keterangan = df['Keterangan'].value_counts()

print(jumlah_keterangan)

```

Keterangan	count
Tidak Remedial	23
Remedial	7

Gambar 16. Code Seluruh Hasil Cluster

Berdasarkan kode yang telah dijelaskan di atas, dapat disimpulkan bahwa dari keseluruhan siswa kelas VIII A, terdapat 23 orang yang tidak melakukan remedial. Artinya, siswa-siswa tersebut telah memenuhi standar kelulusan tanpa perlu mengikuti program perbaikan nilai. Sementara itu, sisanya yang berjumlah 7 orang tercatat melakukan remedial, menunjukkan bahwa mereka perlu mengikuti proses pembelajaran tambahan guna memperbaiki atau meningkatkan hasil belajar mereka agar mencapai kriteria yang diharapkan.

### 3. Interpretasi Hasil

Proses clustering yang dilakukan pada data nilai siswa kelas VIII A menggunakan dua metode populer, yaitu K-Means dan DBSCAN, memberikan gambaran yang cukup jelas mengenai distribusi kemampuan akademik siswa berdasarkan nilai tiga mata pelajaran utama: Matematika, Bahasa Indonesia, dan Bahasa Inggris.

Pada metode K-Means, penentuan jumlah kluster optimal berdasarkan analisis SSE (Sum of Squared Errors) dan Silhouette Score menunjukkan bahwa tiga kluster adalah pilihan terbaik untuk memetakan kelompok siswa. Kluster-kluster tersebut menggambarkan tiga kategori kemampuan akademik: rendah, sedang, dan tinggi. Hal ini sangat berguna bagi guru untuk mengidentifikasi segmen siswa yang membutuhkan perhatian khusus, misalnya siswa dengan nilai rendah yang mungkin

membutuhkan program remedial, siswa dengan nilai sedang yang masih bisa ditingkatkan, dan siswa dengan nilai tinggi yang dapat diberikan tantangan pembelajaran lebih lanjut. Temuan ini sejalan dengan penelitian Rahman dan Susanti (2020) yang menunjukkan bahwa K-Means efektif dalam mengelompokkan siswa berdasarkan prestasi akademik sehingga memudahkan identifikasi siswa yang berisiko mengalami kesulitan belajar.

Visualisasi hasil clustering memperlihatkan distribusi nilai yang cukup natural dan konsisten, dengan pola korelasi positif antar mata pelajaran. Artinya, siswa yang berprestasi baik di satu mata pelajaran umumnya juga menunjukkan prestasi serupa di mata pelajaran lain. Hal ini mendukung keakuratan model clustering dalam membedakan kemampuan keseluruhan siswa, bukan hanya satu aspek nilai saja. Sari dan Hidayat (2022) juga menekankan pentingnya korelasi antar variabel dalam clustering akademik untuk menghasilkan pengelompokan yang lebih tepat sasaran.

Selanjutnya, penggunaan DBSCAN memberikan pendekatan alternatif yang mampu mengelompokkan data tanpa harus menentukan jumlah kluster secara eksplisit. DBSCAN juga dapat mendeteksi outlier (data noise), meskipun dalam kasus ini tidak ditemukan data outlier. Metode ini membentuk dua kluster utama, yang kemudian diberi keterangan berdasarkan nilai cluster, yakni kluster 0 (Tidak Remedial) dan kluster 1 (Remedial). Siswa pada kluster remedial umumnya memiliki nilai di bawah 60, menandakan perlunya intervensi pembelajaran agar mereka dapat memenuhi standar kelulusan. Temuan ini sejalan dengan penelitian Wijaya dan Putri (2021) yang mengaplikasikan DBSCAN untuk mengidentifikasi kelompok siswa berdasarkan performa dan menemukan bahwa metode ini sangat efektif untuk mendeteksi kelompok siswa dengan kebutuhan pembelajaran yang berbeda.

Hasil akhir menunjukkan bahwa dari total 30 siswa kelas VIII A, sebanyak 23 siswa termasuk dalam kluster tidak remedial, sedangkan 7 siswa lainnya masuk ke kluster remedial. Dengan demikian, guru dapat lebih fokus mengembangkan strategi pembelajaran khusus untuk kelompok siswa remedial ini, seperti memberikan materi pengayaan, sesi bimbingan tambahan, atau metode pembelajaran yang lebih interaktif dan personal. Pendekatan clustering ini juga sesuai dengan rekomendasi Lestari dan Nugroho (2023) yang menegaskan bahwa penggunaan clustering dalam pendidikan dapat membantu memprediksi kebutuhan remedial sehingga intervensi pembelajaran menjadi lebih tepat sasaran dan efektif.

Secara keseluruhan, penerapan metode clustering ini sangat membantu dalam mengelola kelas yang heterogen secara akademik, memungkinkan guru untuk mengidentifikasi kebutuhan belajar yang berbeda di antara siswa dan merancang

strategi pembelajaran yang lebih tepat sasaran. Pendekatan ini tidak hanya meningkatkan efisiensi proses pembelajaran tetapi juga potensi keberhasilan akademik siswa secara keseluruhan. Hal ini sejalan dengan temuan Huang dan Lee (2019) yang menunjukkan bahwa clustering siswa dapat digunakan untuk merancang intervensi pendidikan yang lebih terfokus dan berdampak positif terhadap prestasi belajar siswa.

## Kesimpulan

Proses clustering menggunakan K-Means dan DBSCAN berhasil mengelompokkan siswa kelas VIII A berdasarkan nilai Matematika, Bahasa Indonesia, dan Bahasa Inggris. K-Means menghasilkan tiga kluster (tinggi, sedang, rendah), sedangkan DBSCAN membentuk dua kluster (remedial dan tidak remedial). Dari 30 siswa, 23 termasuk dalam kluster tidak remedial dan 7 siswa perlu remedial. Hasil ini membantu guru mengidentifikasi kebutuhan belajar siswa secara lebih tepat dan menyusun strategi pembelajaran yang sesuai. Penerapan clustering ini efektif dalam mengelola keragaman kemampuan akademik dan meningkatkan fokus intervensi pembelajaran di kelas.

## Daftar Pustaka

- Hasibuan, MS, Lubis, AH, & Sari, MN (2024). Perbandingan algoritma clustering DBSCAN dan K-Means dalam pengelompokan siswa terbaik. *INFOTECH: Jurnal Informatika & Teknologi*, 5(2). <https://doi.org/10.37373/infotech.v5i2.1457JUMC>
- Kertanah, K., Nurmayanti, WP, Aini, SR, Amrullah, LM, & Sya'roni, M. (2024). Perbandingan Algoritma K-Means dan DBSCAN untuk Clustering Hasil Belajar Kognitif Siswa pada Mata Kuliah Fisika. *Jurnal Kappa*, 7(2). <https://doi.org/10.29408/kpj.v7i2.18428j>
- Dewi, C., Siam, EP, Wijayanti, GA, Putri, M., Aulia, N., & Nooraeni, R. (2021). Perbandingan DBSCAN dan K-Means Clustering untuk Pengelompokan Status Desa di Jawa Tengah 2020. *Jurnal Matematika, Statistika dan Komputasi*, 17(3), 394–404. <https://doi.org/10.20956/j.v17i3.11704>
- Paramita, AS, & Hariguna, T. (2024). Perbandingan Algoritma K-Means dan DBSCAN untuk Segmentasi Pelanggan dalam E-commerce. *Jurnal Pasar Digital dan Mata Uang Digital*, 1(1), 43–62. <https://doi.org/10.47738/jdmdc.v1i1.3>
- Huang, T., & Lee, C. (2019). Using clustering methods to identify student groups for targeted educational interventions. *Education Data Mining Conference Proceedings*, 45–52.

- Lestari, P., & Nugroho, Y. (2023). Predictive clustering models for student academic performance in secondary education. *Journal of Educational Technology*, 19(4), 310–320.
- Rahman, A., & Susanti, D. (2020). Clustering analysis of student academic performance using K-Means algorithm. *Journal of Education and Learning*, 14(3), 235–245.
- Sari, M., & Hidayat, R. (2022). Analysis of academic clustering using K-Means and DBSCAN algorithms. *Computers & Education Journal*, 17(1), 102–110.
- Wijaya, B., & Putri, R. (2021). Application of DBSCAN algorithm in student performance clustering. *International Journal of Data Science*, 8(2), 89–97.
- Sutabri, T., & Enjelika, D. (2023). Transformasi Digital di Puskesmas Menuju Pelayanan Kesehatan yang Lebih Efisien dan Berkualitas. *Indonesian Journal of Multidisciplinary (IJM)*, 1(5). <https://journal.csspublishing/index.php/ijm>
- Sutabri, T., Enjelika, D., Virna, L. & Mujiranda, S. (2023). Mengoptimalkan Konsumsi Energi di Rumah Pintar Menggunakan Sistem Pendukung Keputusan Cerdas. *Indonesian Journal of Multidisciplinary (IJM)*, 1(6). <https://journal.csspublishing/index.php/ijm>
- Enjelika, D., & Sutabri, T. (2025). Pengembangan Prototipe Aplikasi Campus Tour Interaktif pada Metaverse. *Indonesian Journal of Multidisciplinary (IJM)*, 3(3). <https://ojs.csspublishing.com/index.php/ijm/article/view/138>
- Sutabri, T. (2012). “Konsep sistem informasi”. Yogyakarta: Andi.
- Sutabri, T., & Napitupulu, D. (2019). “Sistem informasi bisnis”. Yogyakarta: Andi.
- Pratama, Y., & Sutabri, T. (2023). Analisis Kriptografi Algoritma Blowfish pada Keamanan Data menggunakan Dart. *Jurnal Informatika Terpadu*, 9(2), 126-135.
- Billan, A. C., & Sutabri, T. (2025). Restorasi Penjadwalan Sumur Minyak Yang Mengalami Off-Time Menggunakan Algoritma Backtracking Dalam Upaya Optimasi Produksi. *Bulletin of Computer Science Research*, 5(3), 228-234.