

Klaster Kelompok Belajar Siswa Menggunakan K-Means dan Visualisasi *Dashboard* Power BI di SMPN 9 Mandau

Clustering Student Learning Groups Using the K-Means Algorithm and Power BI Dashboard Visualization at SMPN 9 Mandau

AULIA PUSPA^{1*}, LINTANG YUNIAR BANOWOSARI¹, RACHMAT ARNANDA²

Abstrak

Di SMPN 9 Mandau, siswa disusun dalam suatu kelompok belajar berdasarkan kriteria yang acak sehingga kelompok belajar yang terbentuk memiliki perbedaan antara anggota suatu kelompok dengan kelompok lainnya dalam hal kemampuan akademik. Hal ini tidak memberikan keuntungan kepada guru dalam menyusun strategi mengajar yang tepat sesuai dengan minat belajar siswa. Pembuatan kelompok belajar siswa dapat dilakukan dengan memanfaatkan data nilai hasil belajar siswa, sehingga dihasilkan kelompok-kelompok belajar yang anggotanya memiliki kemiripan karakteristik satu sama lain. Hal ini juga dimaksudkan agar guru dapat mengambil keputusan yang tepat dalam proses pembentukan kelompok melalui penerapan pembelajaran yang berpusat kepada siswa. Pada penelitian ini data yang digunakan adalah nilai rapor siswa kelas 9A SMPN 9 Mandau. Proses analisis dilakukan dengan menggunakan *Python*, *SQL Server 2019 Management Studio*, dan *Microsoft Power BI*. Metode *clustering* yang digunakan adalah K-Means Clustering dengan membentuk empat kelompok belajar, dimana klaster 1 terdiri dari sebelas siswa, klaster 2 terdiri dari lima siswa, klaster 3 terdiri dari empat siswa, dan klaster 4 terdiri dari sebelas siswa. Penelitian ini menggunakan pendekatan multi metrik, yaitu *inertia* dengan *elbow method*, *silhouette score*, dan *davies-bouldin index* untuk menentukan jumlah klaster optimal. Hasil evaluasi *user acceptance test* terhadap visualisasi *business intelligence dashboard* adalah 91.3%.

Kata Kunci: *business intelligence dashboard*, *clustering*, kelompok belajar, K-Means

Abstract

At SMPN 9 Mandau, students are typically grouped based on random criteria, resulting in significant disparities in academic ability among study groups. Such an approach constrains teachers' capacity to design effective instructional strategies that accommodate students' learning preferences and needs. Alternatively, forming study groups using learning outcome data allows for clustering students with similar characteristics, thereby supporting more appropriate, student-centered grouping decisions. This study utilizes grade 9a student report card data from SMPN 9 Mandau. The analytical process integrates *Python*, *SQL Server 2019 Management Studio*, and *Microsoft Power BI*. The K-Means clustering algorithm is applied, resulting in four study groups: Cluster 1 consisting of eleven students, Cluster 2 of five students, Cluster 3 of four students, and Cluster 4 of eleven students. A multi-metric approach, including *inertia* with the *elbow method*, *Silhouette Score*, and *Davies Bouldin Index*, is used to determine the optimal number of clusters. The effectiveness of the developed *business intelligence dashboard* is further assessed through a *user acceptance test*, which yields a satisfaction score of 91.3%.

Keywords: *business intelligence dashboard*, *clustering*, K-Means, learning group

PENDAHULUAN

SMPN 9 Mandau adalah salah satu Sekolah Menengah Pertama Negeri di Kota Duri, Provinsi Riau. Sekolah ini didirikan pada tahun 2010. SMPN 9 Mandau telah berkembang dengan cukup pesat dalam pembenahan kualitas pendidikan seperti tenaga pendidik, tata usaha, peserta didik, hingga sarana dan prasarana. Dalam upaya meningkatkan kualitas pembelajaran

¹Departemen Manajemen Sistem Informasi, Program Pasca Sarjana, Universitas Gunadarma, Depok 16424

²Departemen Teknik Mesin, Fakultas Teknik, Politeknik Negeri Jakarta, Depok 16425

*Penulis Korespondensi: Tel/Faks: 0823-8571-4613; Surel: auliapuspa27@gmail.com

siswa, salah satu metode pembelajaran yang dilakukan yaitu pembelajaran yang berpusat pada siswa, sedangkan guru dijadikan sebagai pendamping. Guru bukanlah sumber utama dalam mendapatkan suatu bahan ajar tetapi siswa yang lebih aktif serta mandiri dalam menemukan sumber pelajaran (Pertiwi *et al.* 2022).

Kegiatan belajar mengajar yang menyenangkan dan implementatif dapat diterapkan melalui kegiatan belajar kolaboratif. Pembelajaran kolaboratif memudahkan para siswa dalam belajar karena memberikan latihan berpikir kritis (*critical thinking*) dan interaksi sosial (*social interaction*), sehingga persentase rata-rata skor belajar kolaboratif adalah sebesar 82% yang menunjukkan telah tercapainya batas ideal yang diharapkan (Pattipeilohy dan Wijaya 2020).

Selama penerapan Kurikulum 2013, pembagian kelompok belajar oleh tenaga pendidik umumnya masih dilakukan berdasarkan pendekatan non-analitis, seperti urutan absen, posisi tempat duduk, peringkat akademik, maupun secara acak. Apabila pengelompokan dilakukan menggunakan metode-metode tersebut, pembagian kelompok belajar menjadi tidak merata. Hal ini akan menyebabkan siswa yang memiliki nilai tinggi merasa bosan karena materi yang sudah dipahaminya, atau sebaliknya, siswa yang tertinggal akan kesulitan memahami materi tersebut. Kelompok belajar dapat dibentuk sesuai dengan kemiripan nilai siswa. Hal ini dilakukan agar guru mampu mengambil tindakan yang tepat terkait metode pembelajaran yang sesuai dengan setiap karakteristik kelompok sehingga dapat diketahui di manakah kesulitan yang dialami oleh setiap kelompok belajar yang terbentuk.

Pembentukan kelompok belajar siswa dapat dilakukan dengan menggunakan analisis kluster. *Clustering* adalah metode yang digunakan dalam *data mining* dengan mencari data dan mengelompokkan data yang mempunyai kemiripan karakteristik antara data satu dengan data lainnya yang telah diperoleh (Dinata *et al.* 2020). Salah satu algoritma *clustering* yang sering digunakan adalah K-Means. K-Means adalah salah satu algoritma pembelajaran mesin sederhana yang tidak memerlukan pengawasan (*unsupervised*). K-Means sering digunakan karena memiliki kemampuan untuk mengelompokkan jumlah data yang besar dengan waktu komputasi cepat dan efisien (Ahmar *et al.* 2018). K-Means *clustering* merupakan algoritma yang membutuhkan *input* sebanyak k yang membagi n objek ke dalam k kluster sehingga memiliki tingkat kemiripan yang tinggi antar anggota dalam satu kluster, serta tingkat kemiripan rendah dengan anggota pada kluster lain (Ahmar *et al.* 2018).

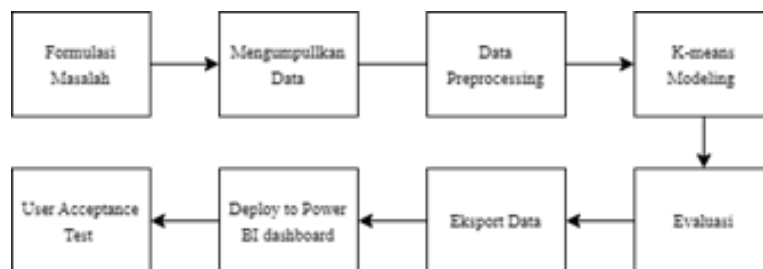
Beberapa penelitian terdahulu telah membahas penerapan algoritma K-Means dalam proses pengelompokan data pendidikan. Sebuah studi menjelaskan bagaimana K-Means dapat digunakan untuk mengelompokkan siswa berdasarkan tingkat pengetahuan yang diperlukan dalam belajar. Algoritma ini diimplementasikan menggunakan *Python* dan diterapkan pada hasil klusterisasi untuk membentuk kelompok belajar heterogen yang terdiri atas siswa dengan tingkat kemampuan yang beragam. Hasil penelitian menunjukkan bahwa capaian belajar siswa mengalami peningkatan, dengan tingkat ketercapaian pembelajaran kolaboratif dan *peer-to-peer learning* masing-masing sebesar 80.95% dan 85.75% (Ikoh dan Catherine 2021). Pada penelitian lainnya, metode *clustering* diterapkan untuk mengelompokkan data nilai siswa menggunakan *software WEKA*. Pengelompokan dengan K-Means membentuk tiga kluster, dengan kluster 0 terdiri dari 72 siswa, kluster 1 terdiri dari 190 siswa, dan kluster 2 berjumlah 133 siswa (Kurniasari *et al.* 2020). Penelitian lain menerapkan analisis kluster menggunakan algoritma K-Means dan Fuzzy C-Means untuk mengelompokkan siswa ke dalam tiga kelompok berdasarkan hasil belajar pada mata kuliah Statistik dan Probabilitas. Pengelompokan dilakukan dengan mempertimbangkan beberapa variabel, yaitu tingkat kehadiran, nilai tugas, nilai ujian tengah semester, dan nilai ujian akhir semester (Sari *et al.* 2021).

Dalam penelitian ini, pembentukan kelompok belajar dilakukan berdasarkan data nilai siswa dengan menggunakan algoritma K-Means untuk menghasilkan pengelompokan yang objektif dan berbasis data. Selanjutnya, hasil analisis tersebut disajikan dalam bentuk *business intelligence dashboard* agar informasi yang dihasilkan dapat divisualisasikan secara jelas, interaktif, dan mudah dipahami, sehingga mendukung proses pengambilan keputusan yang lebih efektif. *Business intelligence dashboard* adalah sebuah dasbor yang berisi visualisasi data

dan alat analitik yang ditampilkan secara visual dalam layar, meliputi *key performance indicators* atau matriks-matriks lainnya yang penting bagi bisnis (Panatagama 2023).

METODE

Tahapan penelitian yang dilakukan pada penelitian ini dapat dilihat pada Gambar 1.



Gambar 1 Tahapan penelitian

Formulasi Masalah

Adapun beberapa masalah yang harus diselesaikan pada penelitian ini yaitu membuat kluster kelompok belajar berdasarkan siswa yang memiliki karakteristik nilai yang serupa, mengidentifikasi apa saja karakteristik pada setiap kluster, *deploy* hasil kluster pada *Power BI dashboard*, dan automasi seluruh proses.

Pengumpulan Data

Data yang diperoleh berasal dari nilai rapor siswa kelas 9A pada semester 5 di SMPN 9 Mandau. Selanjutnya, data tersebut diolah menggunakan implementasi algoritma K-Means untuk membentuk kelompok-kelompok belajar. Pada penelitian ini, jumlah keseluruhan siswa dalam satu kelas akan dijadikan sebagai populasi sehingga tidak melibatkan proses *sampling*. *Microsoft SQL Server Management Studio 19* digunakan sebagai *database*. *SQL Server* digunakan agar dapat dilakukan automasi untuk semua proses pada saat terdapat data baru yang akan diolah. Tahap berikutnya yaitu melakukan *load data* dari *database* ke dalam *Jupyter Notebook* dengan menggunakan *library* yang terdapat pada bahasa pemrograman *Python* untuk diolah dengan cara menghubungkan *SQL Server* dengan memasukkan *driver*, *server*, dan *database* yang digunakan.

Data Preprocessing

Data preprocessing merupakan proses yang dilakukan untuk mengubah data mentah ke dalam bentuk yang lebih mudah dipahami. Pada penelitian ini, tahap *data preprocessing* yang dilakukan yaitu *data exploration*, *data cleaning*, dan *feature engineering*. Langkah *Exploratory Data Analysis* (EDA) pertama yang dilakukan yaitu melakukan investigasi terhadap dimensi dan deskripsi statistik data. Selanjutnya, dilakukan deteksi keberadaan *missing value* dan investigasi elemen yang terdapat dalam setiap *feature*. Hal ini dilakukan agar dapat memahami nilai-nilai unik yang terdapat di dalam *feature* sehingga mempermudah proses selanjutnya. Setelah proses EDA dilakukan, selanjutnya dilakukan *feature engineering* untuk memilih fitur apa saja yang akan diolah ke dalam model, serta melakukan proses normalisasi menggunakan metode *StandardScaler* yang bertujuan untuk mentransformasikan setiap variabel, sehingga memiliki nilai rata-rata (*mean*) sebesar 0 dan standar deviasi sebesar 1. Transformasi ini penting untuk memastikan bahwa seluruh variabel berada pada skala yang sebanding, sehingga tidak ada variabel dengan skala besar yang mendominasi perhitungan jarak dalam algoritma *K-Means* yang berbasis jarak *Euclidean* (Han *et al.* 2012).

Evaluasi

Evaluasi hasil klusterisasi dilakukan dengan menggunakan metrik *inertia*, yaitu jumlah kuadrat jarak antara data dalam satu kluster yang digunakan untuk mengukur tingkat kohesi internal kluster. Semakin kecil nilai *inertia*, maka jarak tiap titik pusat kluster terhadap titik data

akan semakin kecil juga. Jumlah kluster optimal juga didapatkan dari hasil penurunan grafik *inertia* yang membentuk siku-siku (Pradhana *et al.* 2022). Nilai *inertia* tidak digunakan secara langsung sebagai indikator kualitas kluster, namun pola penurunannya diamati menggunakan *elbow method* untuk menemukan titik perubahannya. Untuk memperkuat hasil tersebut, digunakan metrik evaluasi tambahan berupa *silhouette score* dan *davies-bouldin index*.

Ekspor Data

Setelah semua proses pengolahan data selesai dilakukan, langkah selanjutnya yaitu menggabungkan kembali data hasil kluster dengan data sebelumnya untuk diekspor kembali ke *database*.

Deploy ke Power BI Dashboard

Tahap selanjutnya yaitu membuat *dashboard* dan membagikan hasilnya kepada pengguna akhir dengan cara *deploy* atau *publish dashboard* ke situs web atau layanan *Power BI*. Pengguna dapat melihat laporan tersebut secara interaktif. Adapun tahap-tahap pembuatan *dashboard* terlihat pada Gambar 2.



Gambar 2 Tahap pembuatan *dashboard*

Langkah pertama yang harus dilakukan yaitu menghubungkan *database* ke *Power BI*, lalu memasukkan data yang telah diolah untuk divisualisasikan. Selanjutnya, dilakukan perancangan *dashboard* berdasarkan data-data yang akan ditampilkan sesuai dengan kebutuhan. Langkah terakhir yaitu *publish dashboard* ke layanan *web Power BI*.

User Acceptance Test (UAT)

UAT bertujuan untuk melakukan validasi bahwa *dashboard* yang dibuat telah sesuai dengan kebutuhan pengguna. Evaluasi ini dilakukan dengan cara menyebarkan kuesioner yang dibagikan kepada guru-guru SMPN 9 Mandau.

HASIL DAN PEMBAHASAN

Hasil Pengumpulan Data

Pada penelitian ini, data yang diperoleh merupakan hasil nilai akademik siswa dengan jumlah total 31 siswa. Data yang digunakan mencerminkan kondisi aktual populasi dalam satu kelas untuk menemukan variasi kemampuan akademik siswa. Oleh sebab itu, penelitian ini merupakan suatu *case study* yang berfokus pada analisis karakteristik siswa dalam lingkup terbatas, di mana dalam hal ini hanya pada lingkup satu kelas saja. Meskipun dengan jumlah data yang relatif kecil, pendekatan yang digunakan adalah *unsupervised learning* dengan metode K-Means Clustering yang bertujuan untuk menemukan pola serta struktur dalam data berdasarkan kemiripan karakteristik, dan bukan untuk melakukan generalisasi terhadap populasi yang lebih luas (Jain *et al.* 2010).

Dataset terdiri dari nama siswa dan 11 mata pelajaran, yaitu Pendidikan Agama, Bahasa Indonesia, Matematika, IPA, IPS, Bahasa Inggris, Seni Budaya, PJOK, PPKN, Prakarya dan BMR. *Dataset* disimpan ke dalam *dataframe* pada *library pandas* yang disediakan *Python*. Hasil dari pengumpulan data siswa menunjukkan variasi nilai yang cukup beragam pada semua mata pelajaran, seperti Pendidikan Agama dan Bahasa Indonesia yang memperoleh nilai tinggi secara konsisten, serta mata pelajaran tertentu seperti Matematika dan Bahasa Inggris memiliki rentang nilai yang lebih bervariasi. Tampilan lima data teratas dapat dilihat pada Gambar 3.

	nama	pendidikan_agama	ppkn	bahasa_indonesia	matematika	ipa	ips	bahasa_inggris	seni_budaya	pjok	prakarya	bmr
0	ADITIYA RAHMAN	83.0	82.0	82.0	82.0	80.0	80.0	86.0	85.0	84.0	88.0	82.0
1	AHMAD IQBAL INZAKI	93.0	92.0	90.0	90.0	83.0	85.0	93.0	90.0	86.0	86.0	92.0
2	AMEL SEPTYANI MALAU	89.0	88.0	85.0	79.0	82.0	80.0	90.0	85.0	82.0	85.0	85.0
3	AZIMAH	84.0	92.0	85.0	85.0	85.0	80.0	85.0	90.0	82.0	84.0	90.0
4	BIRGITA CLAUDY SITANGGANG	85.0	88.0	82.0	79.0	82.0	80.0	87.0	90.0	82.0	84.0	82.0

Gambar 3 Lima data teratas

Hasil Data Preprocessing

Dataset memiliki dimensi yang terdiri dari 31 baris dan 12 kolom. Tinjauan umum dari *dataset* dilakukan dengan mengamati statistik deskriptif. Hasil dari sebaran statistik menunjukkan rentang nilai siswa yang bervariasi. Nilai rata-rata sebagian besar mata pelajaran berada pada rentang 80-90 yang menjelaskan bahwa secara umum kemampuan akademik siswa cukup baik. Namun terdapat penyebaran nilai yang cukup besar pada standar deviasi sebagian mata pelajaran sehingga proses klasterisasi akan digunakan untuk mengelompokkan siswa berdasarkan kemampuan akademiknya. Rangkuman sebaran statistik *dataset* terlihat pada Gambar 4.

	pendidikan_agama	ppkn	bahasa_indonesia	matematika	ipa	ips	bahasa_inggris	seni_budaya	pjok	prakarya	bmr
count	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000
mean	85.161290	89.032258	84.000000	83.000000	81.741935	82.677419	87.516129	86.709677	83.580645	84.774194	84.129032
std	4.058709	4.423301	3.21455	4.203173	1.590969	4.019843	2.606031	3.951317	1.500538	3.116726	3.810004
min	79.000000	80.000000	79.000000	79.000000	79.000000	79.000000	84.000000	80.000000	80.000000	79.000000	79.000000
25%	83.000000	86.500000	82.000000	79.000000	80.000000	80.000000	85.500000	85.000000	82.500000	84.000000	82.000000
50%	85.000000	88.000000	85.000000	82.000000	82.000000	80.000000	87.000000	85.000000	84.000000	85.000000	83.000000
75%	86.500000	92.000000	85.000000	85.000000	82.500000	85.000000	89.000000	90.000000	84.000000	88.000000	85.000000
max	94.000000	96.000000	90.000000	90.000000	85.000000	90.000000	93.000000	95.000000	86.000000	90.000000	94.000000

Gambar 4 Deskripsi statistik

Identifikasi *missing value* dilakukan agar hasil klaster menjadi lebih tepat. Berdasarkan hasil investigasi pada *dataset* tidak ditemukan *missing value* sehingga tidak perlu dilakukan tindakan lanjutan untuk mengatasinya. Hasil identifikasi *missing value* terlihat pada Gambar 5.

```

nama          0
pendidikan_agama  0
ppkn          0
bahasa_indonesia  0
matematika    0
ipa          0
ips          0
bahasa_inggris  0
seni_budaya   0
pjok         0
prakarya     0
bmr         0
dtype: int64
    
```

Gambar 5 Missing value

Pada *feature engineering*, proses seleksi fitur dilakukan dengan maksud untuk menentukan variabel yang relevan dalam analisis klaster. Hasil seleksi fitur memperlihatkan bahwa dari total 31 baris dan 12 kolom, diperoleh hasil akhir yang akan digunakan pada tahap selanjutnya berupa 31 baris data dengan 11 fitur yang merupakan nilai akademik siswa berupa data numerik. Proses ini bertujuan untuk memastikan hanya variabel relevan saja yang akan digunakan dalam proses *clustering* sehingga dapat meningkatkan kualitas kelompok dan menguraingi *noise* dalam data (Guyon dan Elisseeff 2003). Hasil normalisasi menunjukkan bahwa nilai setiap variabel telah berada pada rentang yang mencakup nilai negatif dan positif. Pola ini menunjukkan bahwa data telah dipusatkan terhadap nilai rata-rata, di mana nilai positif menggambarkan data yang berada di atas rata-rata dan sebaliknya (Han *et al.* 2012). Hasil

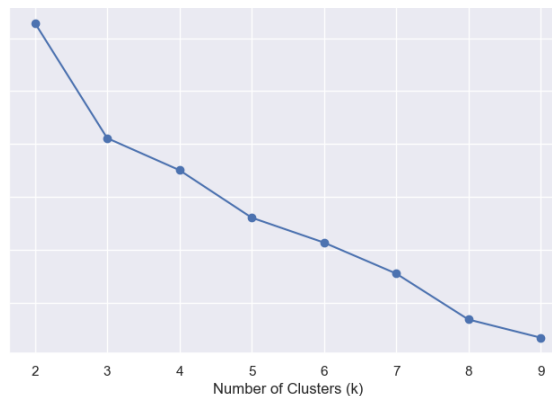
normalisasi menggunakan *StandardScaler* menunjukkan bahwa data telah berhasil ditransformasikan ke dalam distribusi dengan rata-rata mendekati nol dan varians yang seragam. Kondisi ini memastikan bahwa algoritma K-Means dapat bekerja secara optimal tanpa bias skala, sehingga menghasilkan kluster yang lebih valid dan dapat diinterpretasikan secara ilmiah. Hasil dari *feature engineering* dapat dilihat pada Gambar 6.

	pendidikan_agama	ppkn	bahasa_indonesia	matematika	ipa	ips	bahasa_Inggris	seni_budaya	pjok	praktarya	bmr
0	-0.541309	-1.616101	-0.632456	-0.241848	-1.112988	-0.677061	-0.591394	-0.439838	0.284089	1.052107	-0.568038
1	1.963256	0.682024	1.897367	1.692938	0.803825	0.587330	2.139084	0.846480	1.638977	0.399801	2.100018
2	0.961430	-0.237226	0.316228	-0.967393	0.164887	-0.677061	0.968879	-0.439838	-1.070798	0.073647	0.232379
3	-0.290853	0.682024	0.316228	0.483696	2.081700	-0.677061	-0.981462	0.846480	-1.070798	-0.252506	1.566407
4	-0.040396	-0.237226	-0.632456	-0.967393	0.164887	-0.677061	-0.201326	0.846480	-1.070798	-0.252506	-0.568038

Gambar 6 *Feature engineering*

Hasil Evaluasi

Penentuan jumlah kluster dilakukan dengan cara menggabungkan tiga indikator evaluasi yaitu, *inertia* dengan kombinasi elbow point, silhouette score, dan Davies-Bouldin Index (DBI). Pendekatan multi metrik ini diterapkan karena setiap metrik memiliki perspektif evaluasi yang berbeda terhadap struktur kluster (Tan *et al.* 2019). Berdasarkan evaluasi, terdapat penurunan nilai *inertia* yang cukup tajam pada $k-2$ sebesar 225.71 menjadi 182.29 pada $k-3$. Penurunan *inertia* mulai melandai setelah $k-4$ yang menunjukkan bahwa penambahan kluster setelahnya tidak memberikan kontribusi signifikan terhadap model. Titik perubahan gradien mengidentifikasi jumlah kluster yang paling efisien (Thorndike 1953). Berdasarkan *elbow method*, jumlah kluster yang paling efisien berada di rentang $k-3$ hingga $k-5$. Hasil *inertia* dapat dilihat pada Gambar 7.



Gambar 7 Hasil *inertia*

Evaluasi berdasarkan silhouette score memperoleh nilai tertinggi pada $k-2$ sebesar 0.250 namun juga menunjukkan nilai secara keseluruhan tergolong rendah yaitu di bawah 0.3 yang mengindikasikan struktur data tidak terlalu kuat dan memiliki tingkat *overlap* yang cukup tinggi antar kluster (Rousseeuw 1987). Meskipun $k-2$ memberikan nilai terbaik secara numerik, penggunaan dua kluster saja cenderung terlalu menyederhanakan kompleksitas data dan berpotensi mengabaikan variasi penting di dalamnya.

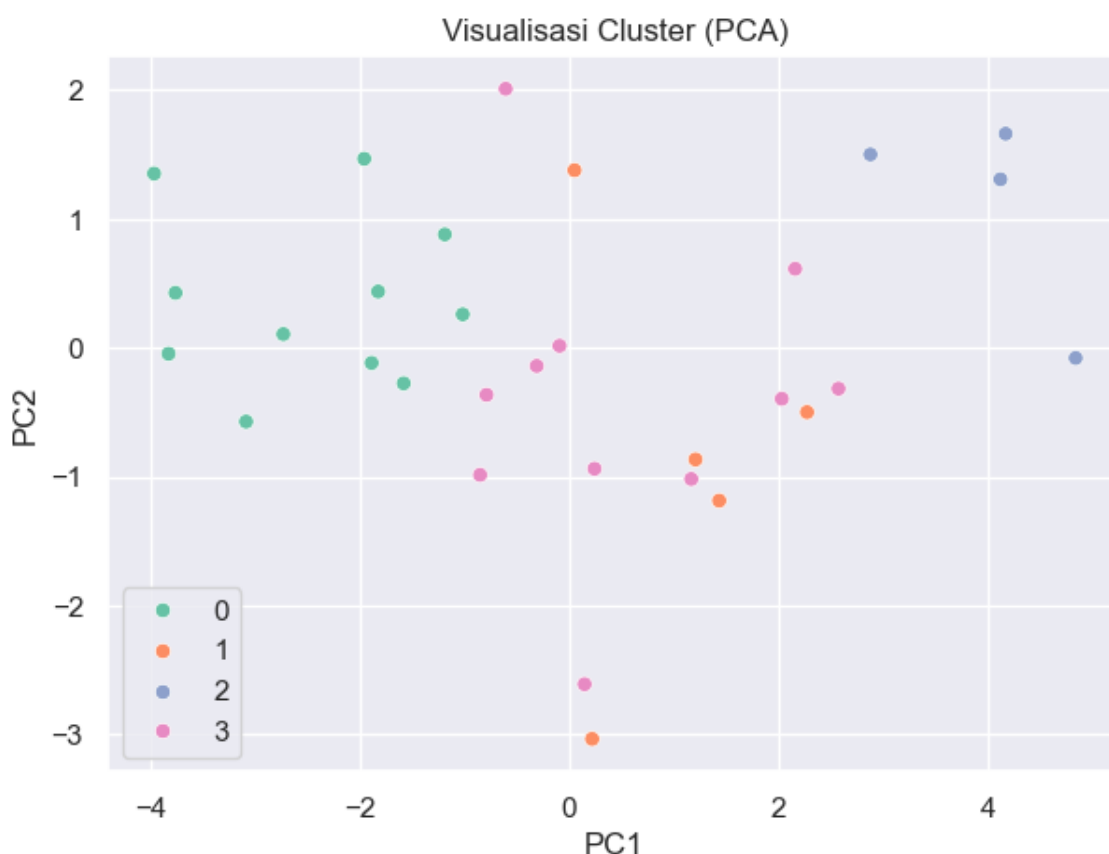
Berbeda dengan silhouette score, DBI menunjukkan bahwa performa terbaik justru diperoleh pada $k-8$ dengan nilai 1.330. Nilai yang lebih rendah menunjukkan bahwa kluster memiliki tingkat kohesi yang tinggi serta separasi yang baik antar kluster (Davies dan Bouldin 1979). Meskipun $k-8$ memberikan hasil terbaik secara matematis, jumlah kluster tersebut berpotensi menghasilkan segmentasi yang terlalu kompleks dan sulit untuk diinterpretasikan secara praktis.

Dengan mempertimbangkan seluruh hasil evaluasi, jumlah kluster yang dipilih dalam penelitian ini adalah $k-4$. Pemilihan ini didasarkan pada prinsip keseimbangan antara kualitas model dan interpretabilitas hasil. Kluster dengan $k-2$ menunjukkan separasi yang terlalu sederhana, sedangkan $k-8$ menunjukkan hasil yang terlalu kompleks dan $k-3$ hingga $k-4$ mampu

memberikan segmentasi yang cukup detail. Keputusan tersebut sejalan dengan pandangan Tan *et al.* (2019) yang menyatakan bahwa penentuan jumlah kluster tidak hanya bergantung pada evaluasi metrik numerik, tetapi juga harus mempertimbangkan konteks data, serta tujuan analisis. Oleh karena itu, pemilihan $k=4$ dalam penelitian ini tidak hanya didasarkan pada hasil evaluasi kuantitatif, tetapi juga pada pertimbangan interpretabilitas dan relevansi hasil terhadap kebutuhan pembentukan kelompok belajar.

Hasil Clustering

Clustering dengan menggunakan metode K-Means telah membentuk kelompok belajar ke dalam empat kluster. Struktur kluster cukup terbentuk, ditandai dengan adanya pengelompokan titik berdasarkan warna. Namun, batas antar-kluster tidak sepenuhnya terpisah, yang terlihat dari adanya *overlap* di beberapa area. Hal ini konsisten dengan hasil evaluasi sebelumnya pada nilai silhouette yang relatif rendah, yang menunjukkan bahwa struktur kluster dalam data bersifat *moderate*. Berdasarkan visualisasi PCA, dapat disimpulkan bahwa hasil *clustering* dengan $k=4$ mampu mengelompokkan data ke dalam beberapa segmen yang cukup jelas. Visualisasi *cluster* dapat dilihat pada Gambar 8.



Gambar 8 Visualisasi cluster

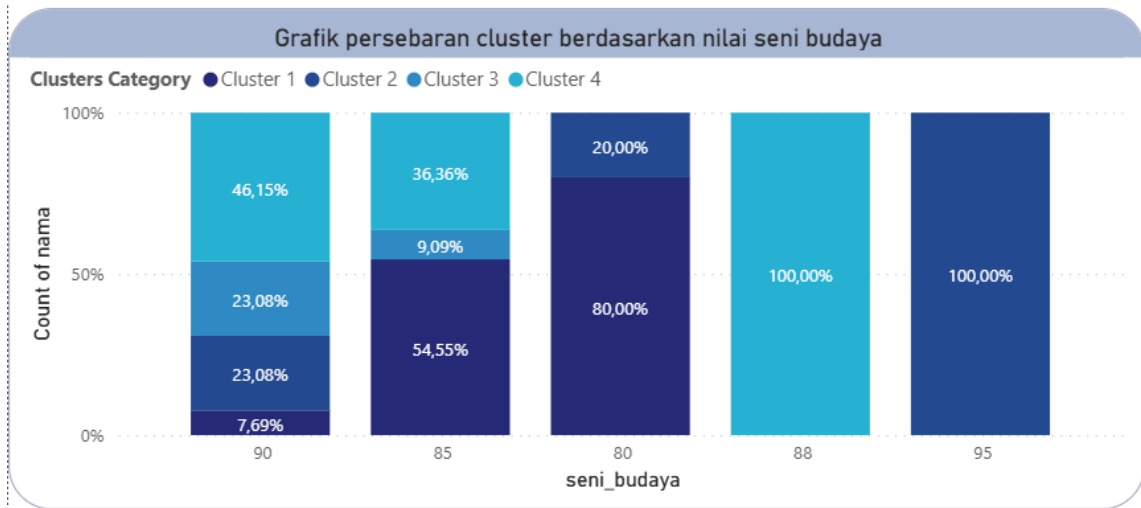
Hasil Visualisasi Power BI dashboard

Hasil dari visualisasi klastering berbentuk matriks tabel dengan *Power BI dashboard* dapat dilihat pada Gambar 9. Berdasarkan data, terdapat 31 jumlah siswa dan 4 kluster yang divisualisasikan menggunakan *matrix*. Kluster 1 terdiri dari sebelas siswa, kluster 2 terdiri dari 5 siswa, kluster 3 terdiri dari empat siswa, dan kluster 4 terdiri dari sebelas siswa.

Hasil dari visualisasi grafik mata pelajaran ditunjukkan pada Gambar 10. Berdasarkan data, terdapat 11 mata pelajaran dan 4 kluster yang divisualisasikan menggunakan *stacked column chart*. Informasi yang diperoleh yaitu berupa persentase siswa atau jumlah siswa pada setiap klasternya pada nilai tertentu. Dari Gambar 10 dapat dilihat bahwa terdapat 11 *chart* serupa sesuai dengan nilai mata pelajaran masing-masing.

Nama Siswa	luster 1	Cluster 2	Cluster 3	Cluster 4	Total
YEMIMA MARTA	1				1
WAWAN KURNIAWAN SITOMPUL	1				1
TIA ELFRENI		1			1
SINTA RAMADHANI			1		1
SILVANI LESTARI				1	1
RISKY RAMADHAN	1				1
RANIA SALSABILA		1			1
RAHMAT GUSTI	1				1
RAFLI ANTONIUS MARPAUNG	1				1
NURCAHAYA SIDABUTAR		1			1
NAISA FADILLAH			1		1
MUHAMMAD KAHFI				1	1
MAULIDINA QORI HAFIZAH				1	1
MARIO CARLOS GULTOM				1	1
MARCELLINO SITOMPUL	1				1
Total	11	5	4	11	31

Gambar 9 Matriks hasil klastering



Gambar 10 Grafik mata pelajaran

Hasil analisis *clustering* menunjukkan bahwa siswa tidak hanya tersegmentasi berdasarkan tingkat capaian akademik (rendah, menengah, tinggi), tetapi juga memperlihatkan variasi pola kemampuan yang lebih kompleks, khususnya antara kemampuan kognitif dan non-kognitif sesuai dengan konsep *multiple intelligences* dimana setiap individu memiliki beragam jenis kecerdasan yang berkembang secara tidak merata (Gardner 1983). Klaster 1 merepresentasikan kelompok siswa dengan performa akademik rendah yang cenderung konsisten pada mata pelajaran seperti matematika, IPA, dan IPS. Dalam sudut pandang pendidikan, pada kelompok ini terdapat keterbatasan pada kecerdasan logis matematis dan analitis. Kondisi ini tidak mengindikasikan kecerdasan yang rendah secara keseluruhan, melainkan bahwa siswa memiliki potensi pada kecerdasan lain namun belum didukung secara optimal (Armstrong 2009). Klaster 2 menunjukkan siswa dengan pencapaian menengah yang relatif stabil di berbagai mata pelajaran tanpa adanya keunggulan yang signifikan. Klaster ini mencerminkan karakteristik siswa dengan kemampuan yang cukup merata, namun belum menunjukkan keunggulan tertentu. Pemberian strategi pembelajaran yang tepat, dapat membantu perkembangan siswa, seperti penguatan konsep dan pembelajaran berbasis pengayaan (Tomlinson 2001). Klaster 3 merepresentasikan siswa dengan performa akademik tinggi yang konsisten, terutama pada mata pelajaran berbasis kognitif seperti matematika, IPA,

dan IPS. Karakteristik kelompok ini didominasi oleh siswa dengan logis matematis dan kemampuan analitis yang kuat. Dalam aspek pendidikan, kelompok ini sering dikategorikan sebagai *high achievers* yang membutuhkan strategi pembelajaran berbasis pengayaan (*enrichment*) agar potensi siswa dapat berkembang secara optimal (Bloom 1984). Klaster 4 memiliki pola yang berbeda, yaitu kelompok siswa yang tidak unggul pada mata pelajaran akademik inti, tetapi memiliki kelebihan pada mata pelajaran seperti seni budaya, PJOK, dan prakarya. Hal ini mengindikasikan dominasi kecerdasan non-kognitif seperti kinestetik, visual-spasial, atau kreatif. Temuan ini menunjukkan bahwa kecerdasan tidak dapat direduksi hanya pada capaian akademik semata, melainkan mencakup berbagai dimensi kemampuan yang beragam (Armstrong 2009). Secara keseluruhan, hasil *clustering* ini memberikan dampak penting bahwa pendekatan pembelajaran yang bersifat seragam tidak begitu relevan untuk mewadahi variasi karakteristik siswa. Setiap klaster menunjukkan kebutuhan strategi pembelajaran yang berbeda, sehingga penerapan strategi *differentiated instruction* sangat diperlukan (Tomlinson 2001). Selain itu, hasil ini juga mempertegas bahwa keberhasilan siswa tidak hanya berfokus pada aspek evaluasi akademik kognitif saja, tetapi juga perlu mendukung potensi non-kognitif yang dimiliki siswa. Karakteristik klaster dapat dilihat pada Tabel 1.

Tabel 1 Karakteristik klaster

Kluster	Kategori	Karakteristik
Klaster 1	Performa akademik rendah	Bahasa Indonesia: 80–85; Bahasa Inggris: 80–86; Seni Budaya: 80–85; IPS: 79–85; Matematika: 75–85; IPA: 79–83; Pendidikan Agama: 82–85; PJOK: 80–85; Prakarya: 80–85; PPKn: 82–86; BMR: 80–85
Klaster 2	Performa akademik sedang	Bahasa Indonesia: 85–90; Bahasa Inggris: 85–91; Seni Budaya: 85–90; IPS: 85–90; Matematika: 80–88; IPA: 80–85; Pendidikan Agama: 83–86; PJOK: 80–90; Prakarya: 82–85; PPKn: 83–90; BMR: 82–88
Klaster 3	Performa akademik tinggi	Bahasa Indonesia: 85–90; Bahasa Inggris: 85–90; Seni Budaya: 85–90; IPS: 83–90; Matematika: 85–92; IPA: 84–90; Pendidikan Agama: 84–90; PJOK: 85–92; Prakarya: 85–90; PPKn: 88–96; BMR: 90–94
Klaster 4	Performa pada non akademik	Bahasa Indonesia: 80–90; Bahasa Inggris: 85–91; Seni Budaya: 85–100; IPS: 80–90; Matematika: 70–88; IPA: 75–85; Pendidikan Agama: 85–95; PJOK: 80–100; Prakarya: 75–90; PPKn: 85–92; BMR: 80–95

Dashboard dengan semua komponen yang sudah lengkap di *publish* ke web *power BI*. *Dashboard* terdiri dari tiga bagian yaitu *cluster summary*, analisis grafik per mata pelajaran, dan analisis karakteristik setiap klaster pada halaman web tersebut. Visualisasi keseluruhan *dashboard* terlihat pada Gambar 11.

Hasil User Acceptance Test

Pengujian dilakukan untuk mengetahui tanggapan pengguna akhir terhadap *dashboard* yang dihasilkan dari penelitian ini dengan memberikan 6 pertanyaan kepada 10 responden yang merupakan guru SMPN 9 Mandau. Tabel 2 adalah skala *likert* yang digunakan sebagai acuan dalam penilaian pada saat memberikan jawaban.

Tabel 2 Skala *likert*

Jawaban	Bobot
Sangat Setuju	5
Setuju	4
Ragu-ragu	3
Tidak Setuju	2
Sangat Tidak Setuju	1

Dengan menggunakan skala *likert* pada Tabel 2, maka disusunlah pertanyaan yang disebarkan menggunakan *google form* kepada para responden. Hasil data kuesioner yang sudah dijumlahkan menurut jawaban yang diberikan responden dapat dilihat pada Tabel 3.

Tabel 4 Hasil perhitungan jawaban responden

Jawaban	Responden	Jumlah skor
Sangat Setuju (SS)	34	170
Setuju (S)	26	104
Rata-rata (R)	0	0
Tidak Setuju (TS)	0	0
Sangat Tidak Setuju (STS)	0	0
Total Skor	60	274

Hasil jawaban tersebut lalu diolah untuk mendapatkan nilai indeks yang menunjukkan persentase penerimaan *dashboard* yang telah dibuat menggunakan persamaan 1 (Y akhir) dan persamaan 2 (Indeks).

$$Y \text{ akhir} = \text{Nilai likert tertinggi} \times \text{Jumlah responden} \times \text{Jumlah pertanyaan} \quad (1)$$

$$\text{Indeks} = (\text{Total skor} / Y \text{ akhir}) \times 100\% \quad (2)$$

Berdasarkan Persamaan 1 dan Persamaan 2, maka perhitungan nilai indeks adalah sebagai berikut:

$$Y \text{ akhir} = 5 \times 10 \times 6 = 300$$

$$\text{Indeks} = (274/300) \times 100\% = 91.33\%$$

Berdasarkan persentase yang diperoleh maka dapat diketahui bahwa tanggapan pengguna terhadap tingkat penerimaan adalah sebesar 91.33%. Hasil persentase tersebut menunjukkan tingkat penerimaan yang sangat kuat berdasarkan kriteria interpretasi skor (Ghifari 2021) pada Tabel 5.

Tabel 5 Kriteria

Keterangan	Persentase
Sangat Lemah	0% - 20%
Lemah	21% - 40%
Cukup	41% - 60%
Kuat	61% - 80%
Sangat Kuat	81% - 100%

SIMPULAN

Klasterisasi kelompok belajar siswa menggunakan algoritma K-Means dan visualisasi *dashboard* dengan *Microsoft Power BI* telah berhasil dilakukan. Berdasarkan data nilai rapor, terbentuk empat klaster dengan karakteristik akademik yang berbeda. Studi ini menentukan jumlah klaster optimal menggunakan pendekatan evaluasi multi-metrik, yang menggabungkan *inertia* (elbow method), silhouette score, dan davies-bouldin index untuk memastikan penilaian komprehensif terhadap kualitas pengelompokan. Hasil pengelompokan menunjukkan bahwa pengelompokan siswa tidak hanya berdasarkan tingkat prestasi akademik tetapi juga mencerminkan beragam pola kemampuan kognitif dan non-kognitif yang selaras dengan teori *multiple intelligences*. Secara spesifik, klaster 1 mewakili siswa dengan prestasi akademik rendah, klaster 2 menunjukkan prestasi sedang dan merata, klaster 3 terdiri dari siswa berprestasi tinggi, dan klaster 4 mencerminkan siswa dengan kekuatan di bidang non-akademik seperti seni dan PJOK.

Dashboard yang dikembangkan menampilkan ringkasan klaster, analisis nilai mata pelajaran, dan karakteristik tiap kelompok, dengan hasil *User Acceptance Test* (UAT) menunjukkan tingkat penerimaan sangat tinggi sebesar 91.33%. Hasil ini membuktikan bahwa *dashboard* dapat membantu guru dalam pengambilan keputusan pembentukan kelompok belajar secara lebih efektif dengan menekankan pentingnya pembelajaran berdiferensiasi, karena pendekatan pembelajaran yang seragam tidak cukup untuk memfasilitasi beragam karakteristik siswa. Hasil penelitian menunjukkan bahwa strategi pendidikan harus memperhatikan potensi kognitif dan non-kognitif untuk mendukung perkembangan siswa. Penelitian ini masih terbatas pada penggunaan algoritma K-Means, sehingga disarankan pada penelitian berikutnya untuk membandingkan dengan algoritma *clustering* lainnya guna memperoleh hasil yang lebih baik.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada pihak SMPN 9 Mandau yang telah memberikan izin dan dukungan dalam pelaksanaan penelitian ini. Ucapan terima kasih juga disampaikan kepada seluruh pihak yang telah membantu, baik secara langsung maupun tidak langsung, khususnya kepada para guru dan staf yang telah berkontribusi dalam penyediaan data dan informasi yang diperlukan. Semoga hasil penelitian ini dapat memberikan manfaat dan kontribusi bagi pengembangan kualitas pendidikan di SMPN 9 Mandau.

DAFTAR PUSTAKA

- Ahmar AS, Napitupulu D, Rahim R, Hidayat R, Sonatha Y, Azmi M. 2018. Using K-Means clustering to cluster provinces in Indonesia. *Journal of Physics: Conference Series*. 1028(1): 0–6. <https://doi.org/10.1088/1742-6596/1028/1/012006>.
- Armstrong T. 2009. Multiple intelligences in the classroom. Ed ke-3. Alexandria (US): ASCD.
- Bloom BS. 1984. The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*. 13(6): 4-16.
- Davies DL, Bouldin DW. 1979. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1(2): 224–227.
- Dinata RK, Safwandi S, Hasdyna N, Azizah N. 2020. Analisis K-Means clustering pada data sepeda motor. *INFORMAL: Informatics Journal*. 5(1): 10. <https://doi.org/10.19184/isj.v5i1.17071>.
- Gardner H. 1983. Frames of mind: the theory of multiple intelligences. New York (US): Basic Books.
- Ghifari AB. 2021. Visualisasi data covid-19 Indonesia menggunakan Microsoft Power BI [skripsi]. Depok(ID): Universitas Gunadarma.
- Guyon I, Elisseeff A. 2003. An introduction to variable and feature selection. *Journal of Machine Learning Research*. 3: 1157-1182.
- Han J, Kamber M, Pei J. 2012. Data mining: Concepts and techniques. Ed ke-3. Burlington (US): Morgan Kaufmann.
- Jain AK, Murty MN, Flynn PJ. 2010. Data clustering: 50 years beyond K-Means . *Pattern Recognition Letters*. 31(8): 651-666.
- Kurniasari D, Asroni A, Kurnianti A. 2020. The implementation of clustering method with K-Means algorithm in grouping data of students' course scores at Universitas Muhammadiyah Yogyakarta. *Emerging Information Science and Technology*. 1(3): 75–83. <https://doi.org/10.18196/eist.v1i3.13172>
- Panatagama. 2023. Kenali business intelligence dashboard dan keuntungannya. *Terralogic [internet]*. <https://terralogiq.com/business-intelligence-dashboard/>
- Pattipeilohy V, Wijaya F. 2020. Implementasi model collaborative learning sebagai landasan untuk membangun keterampilan sosial dan keaktifan mahasiswa. *Maneksi*. 9(1): 283–289.
- Pertiwi AD, Nurfatimah SA, Hasna S. 2022. Menerapkan metode pembelajaran berorientasi student-centered menuju masa transisi Kurikulum Merdeka. *Jurnal Pendidikan Tambusai*. 6(2): 8839–8848.
- Pradhana OJ, Saepudin D, Indwiarti. 2022. Penerapan clustering pada tipe data campuran menggunakan k-prototype pada perusahaan multifinance. *E-Proceeding of Engineering*. 9(6): 1996–2007.
- Rousseeuw PJ. 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*. 20: 53–65.
- Sari IP, Al-Khowarizmi, Batubara IH. 2021. Cluster analysis using K-Means algorithm and Fuzzy C-means clustering for grouping students' abilities in online learning process. *Journal of Computer Science, Information Technology and Telecommunication Engineering (JCoSITTE)*. 2(1): 139–144. <https://doi.org/10.30596/jcositte.v2i1.6504>.

- Tan PN, Steinbach M, Kumar V. 2019. *Introduction to data mining*. Ed ke-2. Boston (US): Pearson.
- Thorndike RL. 1953. Who belongs in the family?. *Psychometrika*. 18(4): 267–276.
- Tomlinson CA. 2001. How to differentiate instruction in mixed-ability classrooms. Ed ke-2. Alexandria (US): ASCD.
- Ikoh MIV, Catherine OD. 2021. Students' group formation using K-Means clustering in combination with a heterogeneous grouping algorithm. *American Journal of Engineering Research*. 10(11): 1–9.