

PENGELOMPOKAN GAYA BELAJAR SISWA MENGGUNAKAN ALGORITMA K-MEANS BERDASARKAN MODEL VAK (STUDI KASUS: SMKN 3 TANGERANG SELATAN)

Sri Yulia, Okta Irawati, Nurhasanah, Syndhe Qumaruw Syty,
Mochammad Rizky, Putri Adela Fadli, Sandi Firmansyah

Teknik Informatika, Universitas Pamulang

Jl. Raya Puspittek, Buaran, Kec. Pamulang, Kota Tangerang Selatan, Banten, Indonesia
dosen02610@unpam.ac.id

ABSTRAK

Perbedaan gaya belajar siswa merupakan tantangan dalam menciptakan pembelajaran yang efektif, khususnya di sekolah menengah kejuruan yang menekankan keseimbangan antara teori dan praktik. Permasalahan yang sering muncul adalah proses identifikasi gaya belajar siswa yang masih dilakukan secara manual, sehingga memerlukan waktu yang lama dan berpotensi menimbulkan subjektivitas. Penelitian ini bertujuan untuk mengelompokkan gaya belajar siswa di SMKN 3 Tangerang Selatan secara objektif menggunakan algoritma *K-Means* berdasarkan model *Visual*, *Auditory*, dan *Kinesthetic* (VAK). Metode penelitian meliputi pengumpulan data melalui kuesioner, pra-proses data, penerapan algoritma *K-Means*, serta penentuan jumlah kluster optimal menggunakan metode evaluasi *Silhouette Score*. Hasil penelitian menunjukkan bahwa jumlah kluster optimal diperoleh pada $k = 3$ berdasarkan nilai *Silhouette Score* tertinggi, yang menghasilkan tiga kluster, yaitu C1 (*Visual-Kinesthetic Moderat*), C2 (*Low Engagement Learner*), dan C3 (*Kinesthetic-Dominant Active Learner*) yang menunjukkan jumlah anggota terbesar. Hasil pengelompokan ini menunjukkan bahwa gaya belajar siswa bersifat multidimensi dan tidak terbatas pada satu kategori VAK tertentu. Temuan penelitian ini diharapkan dapat membantu guru dalam merancang strategi pembelajaran yang lebih terpersonalisasi dan sesuai dengan karakteristik masing-masing kelompok siswa.

Kata kunci : data mining, k-means, gaya belajar VAK, Klastering, orange data mining.

1. PENDAHULUAN

Perkembangan teknologi informasi yang semakin pesat telah mendorong pemanfaatan teknologi komputasi dalam berbagai bidang, termasuk bidang pendidikan. Salah satu pemanfaatan teknologi yang banyak diterapkan adalah *data mining*, yaitu proses penggalian informasi penting dari kumpulan data dalam jumlah besar untuk menemukan pola atau pengetahuan baru yang dapat mendukung pengambilan keputusan. Dalam konteks pendidikan, *data mining* digunakan untuk menganalisis data siswa guna meningkatkan kualitas pembelajaran serta mendukung perencanaan strategi pembelajaran yang lebih efektif [1].

Salah satu permasalahan yang sering dihadapi dalam proses pembelajaran adalah perbedaan karakteristik dan gaya belajar siswa. Setiap siswa memiliki cara yang berbeda dalam menerima, mengolah, dan memahami informasi, sehingga penggunaan metode pembelajaran yang seragam sering kali kurang efektif. Perbedaan gaya belajar ini dapat menyebabkan variasi tingkat pemahaman siswa terhadap materi pembelajaran yang sama, meskipun disampaikan oleh guru dengan metode yang sama [2].

Gaya belajar merupakan faktor penting yang berpengaruh terhadap keberhasilan siswa dalam proses belajar. Pemahaman terhadap gaya belajar siswa dapat membantu guru dalam memilih metode, media, dan strategi pembelajaran yang sesuai dengan karakteristik siswa. Salah satu model gaya belajar yang banyak digunakan dalam dunia pendidikan adalah model

Visual, *Auditory*, dan *Kinesthetic* (VAK). Model VAK mengelompokkan gaya belajar siswa berdasarkan kecenderungan penggunaan indera penglihatan (*visual*), pendengaran (*auditory*), dan aktivitas fisik atau praktik langsung (*kinesthetic*) dalam proses pembelajaran. Penerapan model VAK diharapkan mampu meningkatkan efektivitas pembelajaran karena materi disampaikan sesuai dengan preferensi belajar siswa [3].

Namun, proses identifikasi gaya belajar siswa secara konvensional atau manual membutuhkan waktu yang relatif lama dan berpotensi menimbulkan subjektivitas dalam penilaian. Selain itu, pengolahan data siswa secara manual menjadi kurang efisien apabila jumlah data yang dianalisis cukup besar [4]. Oleh karena itu, diperlukan suatu pendekatan berbasis sistem komputer yang mampu mengelola dan menganalisis data siswa secara objektif, cepat, dan terstruktur menggunakan teknik *data mining* [5].

Salah satu pendekatan yang dapat digunakan untuk mengatasi permasalahan tersebut adalah penerapan teknik *data mining*, khususnya metode klastering [5]. Klastering merupakan teknik pengelompokan data ke dalam beberapa kelompok berdasarkan tingkat kemiripan tertentu tanpa menggunakan label kelas. Dalam bidang pendidikan, klastering banyak digunakan untuk mengelompokkan siswa berdasarkan karakteristik tertentu, seperti prestasi akademik, minat belajar, maupun gaya belajar [5].

Algoritma *K-Means* merupakan salah satu algoritma klastering yang paling populer dan banyak digunakan karena memiliki konsep yang sederhana, efisiensi komputasi yang baik, serta mudah diimplementasikan [6]. Algoritma ini bekerja dengan cara mempartisi data ke dalam sejumlah klaster berdasarkan kedekatan data dengan pusat klaster (*centroid*). Beberapa penelitian sebelumnya menunjukkan bahwa algoritma *K-Means* efektif digunakan dalam pengelompokan data siswa, termasuk pengelompokan berdasarkan gaya belajar [7].

Berdasarkan uraian tersebut, penelitian ini mengusulkan penerapan algoritma *K-Means* untuk melakukan pengelompokan gaya belajar siswa berdasarkan model VAK. Penelitian ini dilakukan dengan studi kasus pada SMKN 3 Tangerang Selatan. Diharapkan hasil penelitian ini dapat membantu pihak sekolah dan guru dalam memahami karakteristik gaya belajar siswa secara lebih objektif dan sistematis, serta menjadi dasar dalam penyusunan strategi pembelajaran yang lebih efektif, adaptif, dan sesuai dengan kebutuhan siswa.

2. TINJAUAN PUSTAKA

2.1. Tinjauan Studi

Memahami dasar dari pengelompokan siswa yang dilakukan dalam penelitian ini, diperlukan pemahaman mendalam mengenai konsep belajar dan klasifikasi gaya belajar. Gaya belajar merupakan salah satu faktor internal yang mempengaruhi keberhasilan siswa dalam menyerap informasi selama proses pembelajaran.

2.2. Pengertian Belajar

Belajar merupakan suatu proses usaha sadar yang dilakukan oleh individu untuk memperoleh perubahan tingkah laku yang baru secara keseluruhan, sebagai hasil dari pengalamannya sendiri dalam interaksi dengan lingkungannya. Belajar bukan hanya sekadar mengingat fakta, tetapi juga melibatkan proses berpikir dan pemahaman untuk memecahkan masalah. Hasil dari proses belajar dapat terlihat dalam bentuk perubahan pengetahuan, pemahaman, keterampilan, serta nilai dan sikap yang lebih positif [8].

2.3. Model Gaya Belajar VAK

Setiap siswa memiliki preferensi yang berbeda dalam menyerap, mengatur, dan mengolah informasi. Preferensi inilah yang disebut sebagai gaya belajar. Salah satu model gaya belajar yang paling umum digunakan adalah model VAK (*Visual, Auditory, Kinesthetic*) yang dikembangkan oleh DePorter dan Hernacki. Model ini mengklasifikasikan cara belajar manusia ke dalam tiga modalitas utama, yaitu:

Siswa dengan gaya belajar *visual* lebih mudah menyerap informasi melalui indra penglihatan. Mereka cenderung lebih suka membaca, melihat grafik, diagram, atau demonstrasi *visual* daripada mendengarkan penjelasan lisan. Karakteristik utama siswa *visual* adalah rapi, teratur, dan sering kali

membutuhkan gambaran menyeluruh sebelum mempelajari detail [3].

Siswa dengan gaya belajar auditori mengandalkan pendengaran sebagai sarana utama dalam menyerap informasi. Mereka lebih mudah memahami materi melalui ceramah guru, diskusi kelompok, atau mendengarkan rekaman. Siswa tipe ini sering kali membaca dengan suara keras atau menggerakkan bibir saat membaca dan mudah terganggu oleh keributan [3].

Siswa kinestetik belajar paling baik melalui aktivitas fisik dan pengalaman langsung. Mereka sulit untuk duduk diam dalam waktu yang lama dan lebih suka memanipulasi objek atau melakukan praktik langsung (*learning by doing*) untuk memahami suatu konsep. Ciri khas siswa ini adalah sering menggunakan isyarat tubuh saat berbicara dan menghafal dengan cara berjalan [3].

2.4. Data Mining

Data mining merupakan bidang ilmu yang berkembang pesat seiring dengan melonjaknya jumlah data di berbagai sektor. Secara sederhana, *data mining* dapat dipahami sebagai upaya untuk mendapatkan informasi yang berharga dari gudang data yang sangat besar.

2.5. Pengertian Data Mining

Data mining merupakan proses analisis data dalam skala besar yang bertujuan untuk mengekstraksi pola, hubungan, dan informasi tersembunyi yang tidak dapat diperoleh melalui pengolahan data secara konvensional. Proses ini sering dikaitkan dengan *Knowledge Discovery in Database (KDD)*, yaitu kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola, atau hubungan dalam set data yang berukuran besar. Keluaran dari *data mining* dapat dimanfaatkan untuk memperbaiki pengambilan keputusan di masa depan [9].

Proses KDD terdiri dari tujuh tahapan utama [9]. Tahap pertama adalah *Data Cleaning* untuk membersihkan *noise* dan inkonsistensi data. Selanjutnya adalah *Data Integration* yang menggabungkan data dari berbagai sumber, diikuti oleh *Data Selection* untuk memilih data yang relevan. Tahap keempat adalah *Data Transformation* untuk mengubah format data agar siap diolah. Proses ini dilakukan pada tahap *Data Mining* untuk mengekstraksi pola, yang kemudian dilanjutkan dengan *Pattern Evaluation* untuk mengidentifikasi pola yang valid. Tahap terakhir adalah *Knowledge Presentation* untuk menyajikan pengetahuan kepada pengguna.

2.6. Klastering

Klastering adalah metode analisis data yang berfungsi untuk mengelompokkan objek ke dalam sejumlah klaster berdasarkan karakteristik kemiripan, tanpa memerlukan informasi label sebelumnya.

Berbeda dengan klasifikasi yang memiliki label kelas yang sudah ditentukan sebelumnya, klastering bertujuan untuk mengelompokkan objek-objek data hanya berdasarkan informasi yang terdapat pada data itu sendiri [10].

Prinsip utama dari klastering adalah memaksimalkan kemiripan (*similarity*) antar anggota dalam satu klaster dan meminimalkan kemiripan antar anggota klaster yang berbeda. Dengan demikian, data yang berada dalam satu kelompok akan memiliki karakteristik yang homogen, sedangkan data antar kelompok yang berbeda akan bersifat heterogen. Metode ini sangat berguna untuk mengenali struktur data yang kompleks dan belum memiliki label yang jelas [10].

2.7. Preprocessing Data

Data yang diperoleh dari lapangan seringkali tidak lengkap, mengandung *noise*, dan tidak konsisten. Oleh karena itu, diperlukan tahapan *data preprocessing* untuk mengubah data mentah menjadi format yang berkualitas.

2.8. Pembersihan Data (Data Cleaning)

Pembersihan data (*data cleaning*) adalah proses mendeteksi dan mengoreksi catatan yang rusak atau tidak akurat dari sebuah set data. Tujuan utama dari tahap ini adalah untuk mengisi nilai yang hilang (*missing values*), menghaluskan *noisy data*, serta menyelesaikan inkonsistensi data yang mungkin terjadi akibat kesalahan penginputan [9].

2.9. Normalisasi Data

Sebelum proses klastering dilakukan, data gaya belajar siswa terlebih dahulu melalui tahap normalisasi untuk menyamakan skala antar atribut *Visual*, *Auditory*, dan *Kinesthetic*. Tahap ini penting karena perbedaan rentang nilai antar atribut dapat menyebabkan salah satu atribut mendominasi perhitungan jarak pada algoritma *K-Means*.

Pada penelitian ini, normalisasi data dilakukan menggunakan metode *Min-Max Normalization*, yang mentransformasikan nilai data ke dalam rentang 0 hingga 1. Metode ini dipilih karena sederhana, mudah diinterpretasikan, serta sesuai untuk algoritma klastering berbasis jarak seperti *K-Means*.

Rumus *Min-Max Normalization* dinyatakan sebagai berikut [9]:

$$V' = \frac{v - \min_A}{\max_A - \min_A}$$

Keterangan:

- V' : Nilai data baru hasil normalisasi.
- v : Nilai data asli.
- $\min_A$: Nilai minimum atribut A.
- $\max_A$: Nilai maksimum atribut A.

Proses normalisasi data dilakukan secara otomatis menggunakan widget *Normalize* pada perangkat lunak *Orange Data Mining*, sehingga setiap atribut gaya belajar memiliki kontribusi yang

seimbang dalam proses perhitungan jarak Euclidean. Data hasil normalisasi selanjutnya digunakan sebagai input pada proses klastering menggunakan algoritma *K-Means*.

2.10. Algoritma K-Means

Algoritma *K-Means* melakukan pengelompokan data dengan membagi objek ke dalam sejumlah klaster, di mana setiap data ditempatkan pada klaster dengan pusat (*centroid*) terdekat berdasarkan perhitungan jarak. Proses *K-Means* diawali dengan penentuan jumlah klaster (k), kemudian dilakukan perhitungan jarak setiap data ke *centroid* untuk menentukan keanggotaan klaster[6].

Perhitungan jarak pada algoritma *K-Means* umumnya menggunakan metode *Euclidean Distance*, yang dirumuskan sebagai berikut[10]:

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

Keterangan:

- $d(x, y)$: Jarak antara data x dan *centroid* y .
- x_i : Data ke- i .
- y_i : *Centroid* klaster ke- i .
- n : Jumlah atribut yang digunakan.

Setelah seluruh data dikelompokkan, nilai *centroid* akan diperbarui dengan menghitung rata-rata dari seluruh data dalam satu klaster. Proses ini akan terus diulang hingga tidak terjadi perubahan signifikan pada keanggotaan klaster [6].

2.11. Penelitian Terdahulu

Penelitian mengenai pengelompokan gaya belajar siswa telah banyak dilakukan dengan berbagai pendekatan dan metode analisis. Andriani dan Nugraheni [2] menganalisis karakteristik gaya belajar siswa dalam pembelajaran berdiferensiasi dan menegaskan bahwa perbedaan gaya belajar berpengaruh terhadap keterlibatan dan pemahaman siswa dalam proses pembelajaran. Hasil penelitian tersebut menunjukkan pentingnya identifikasi gaya belajar sebagai dasar perancangan strategi pembelajaran yang adaptif.

Penelitian lain yang dilakukan oleh Azis et al. [3] mengkaji gaya belajar *Visual*, *Auditory*, dan *Kinesthetic* (VAK) pada siswa sekolah menengah dan menemukan bahwa siswa cenderung memiliki kombinasi gaya belajar, bukan hanya satu modalitas dominan. Temuan ini mengindikasikan bahwa pendekatan pembelajaran yang fleksibel dan berbasis karakteristik siswa sangat diperlukan untuk meningkatkan efektivitas pembelajaran.

Dalam konteks penerapan teknik komputasi, Jatmiko et al. [4] melakukan analisis kecenderungan gaya belajar VAK menggunakan metode klasifikasi dan menunjukkan bahwa pendekatan berbasis algoritma mampu mengurangi subjektivitas dalam penentuan gaya belajar siswa. Penelitian tersebut

memperkuat argumen bahwa pemanfaatan metode komputasi dapat meningkatkan akurasi analisis karakteristik belajar siswa.

Selanjutnya, Hendrastuty [5] menerapkan algoritma *K-Means* dalam evaluasi hasil pembelajaran siswa dan membuktikan bahwa metode klastering efektif dalam mengelompokkan siswa berdasarkan kemiripan karakteristik data. Hasil serupa juga ditunjukkan oleh Yudistira dan Andika [6] serta Setiawan dan Mailoa [7], yang menyatakan bahwa algoritma *K-Means* mampu menghasilkan pengelompokan data siswa yang cukup baik dan mudah diinterpretasikan, termasuk dalam konteks pengelompokan gaya belajar.

Berdasarkan penelitian-penelitian terdahulu tersebut, dapat disimpulkan bahwa pengelompokan gaya belajar siswa menggunakan teknik *data mining*, khususnya algoritma *K-Means*, memiliki potensi besar dalam mendukung pengambilan keputusan di bidang pendidikan. Namun, penelitian yang secara khusus mengkaji pengelompokan gaya belajar VAK pada konteks pendidikan vokasi di SMK masih terbatas. Oleh karena itu, penelitian ini dilakukan untuk mengisi celah tersebut dengan menerapkan algoritma *K-Means* dalam pengelompokan gaya belajar siswa SMK sebagai dasar perancangan strategi pembelajaran yang lebih adaptif.

3. METODE PENELITIAN

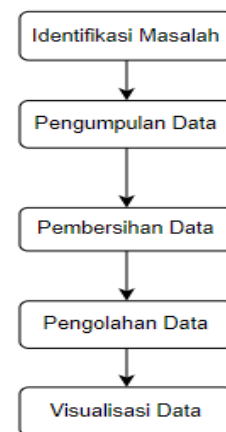
Penelitian ini menggunakan pendekatan kuantitatif eksploratif dengan metode *data mining*, khususnya teknik klastering menggunakan algoritma *K-Means*, untuk mengidentifikasi pola pengelompokan gaya belajar siswa berdasarkan model *Visual, Auditory, dan Kinesthetic* (VAK). Pendekatan eksploratif dipilih karena penelitian ini bertujuan untuk menggali struktur alami data gaya belajar siswa tanpa menggunakan label kelas tertentu, sehingga memungkinkan ditemukannya pola atau karakteristik kelompok yang muncul secara objektif dari data.

Populasi penelitian ini adalah seluruh siswa dalam satu kelas jurusan Teknik dan Bisnis Sepeda Motor (TBSM) di SMKN 3 Tangerang Selatan pada periode akademik 2025, dengan jumlah total 20 siswa. Teknik pengambilan sampel yang digunakan adalah *purposive sampling*, dengan kriteria inklusi berupa siswa aktif yang tidak sedang cuti akademik, serta kriteria eksklusi berupa data responden yang tidak lengkap lebih dari 20%. Berdasarkan kriteria tersebut, diperoleh sebanyak 17 siswa sebagai sampel penelitian, dengan tingkat respons sebesar 85%, yang dinilai telah mewakili populasi kelas secara proporsional.

Data yang digunakan dalam penelitian ini merupakan data primer yang dikumpulkan pada bulan November 2025 melalui penyebaran kuesioner daring menggunakan Google Form. Instrumen penelitian disusun berdasarkan indikator gaya belajar *Visual, Auditory, dan Kinesthetic*, serta menggunakan skala Likert lima tingkat, yaitu Sangat Tidak Setuju hingga

Sangat Setuju. Setiap jawaban responden dikonversi ke dalam nilai numerik dengan rentang 1 sampai 5 untuk mendukung proses analisis kuantitatif. Hasil pengisian kuesioner selanjutnya direkap dan diolah menjadi tiga variabel utama, yaitu *Visual* (V), *Auditory* (A), dan *Kinesthetic* (K), yang masing-masing merepresentasikan kecenderungan gaya belajar siswa.

Untuk memastikan proses penelitian berjalan secara sistematis dan terstruktur, alur penelitian dirancang dalam beberapa tahapan utama sebagaimana ditunjukkan pada Gambar 1, yang meliputi identifikasi masalah, pengumpulan data, pembersihan data, pengolahan data, dan visualisasi data.



Gambar 1. Metode Penelitian

Sebelum dilakukan proses klastering, data terlebih dahulu melalui tahap pra-proses (*preprocessing*) untuk meningkatkan kualitas dan konsistensi data. Tahapan pra-proses meliputi pembersihan data dengan menghapus data yang tidak lengkap atau duplikat, serta pemeriksaan konsistensi nilai untuk menghindari adanya nilai ekstrem yang berpotensi memengaruhi hasil klastering. Selanjutnya dilakukan seleksi atribut dengan menetapkan variabel *Visual, Auditory, dan Kinesthetic* sebagai fitur utama yang digunakan dalam proses analisis.

Tahap berikutnya adalah normalisasi data menggunakan metode *Min-Max Normalization*, yang bertujuan untuk menyamakan skala antar atribut ke dalam rentang nilai 0 hingga 1. Normalisasi ini penting karena algoritma *K-Means* bersifat sensitif terhadap perbedaan skala data dan menggunakan perhitungan jarak Euclidean dalam menentukan kedekatan antar data. Seluruh proses pra-proses dan normalisasi data dilakukan menggunakan perangkat lunak *Orange Data Mining* untuk memastikan proses analisis berjalan secara sistematis dan objektif.

Proses klastering dilakukan menggunakan algoritma *K-Means* dengan menetapkan variasi jumlah kluster (k) sebagai parameter awal. Algoritma diawali dengan inisialisasi *centroid* awal, kemudian menghitung jarak antara setiap data dan *centroid* menggunakan *Euclidean Distance*. Setiap data akan dikelompokkan ke dalam kluster dengan jarak terdekat, dan nilai *centroid* akan diperbarui

berdasarkan rata-rata data pada masing-masing kluster. Proses ini dilakukan secara iteratif hingga mencapai kondisi konvergen, yaitu ketika tidak terjadi perubahan signifikan pada keanggotaan kluster.

Untuk menentukan jumlah kluster yang optimal, penelitian ini menggunakan metode evaluasi *Silhouette Score*, dengan pengujian nilai k dari 2 hingga 5. Nilai *Silhouette Score* digunakan untuk menilai kualitas hasil klastering berdasarkan tingkat kohesi dalam kluster dan separasi antar kluster. Jumlah kluster dengan nilai *Silhouette Score* tertinggi dipilih sebagai konfigurasi terbaik yang merepresentasikan struktur data gaya belajar siswa.

Tahap akhir penelitian adalah visualisasi dan interpretasi hasil klastering, yang bertujuan untuk mempermudah pemahaman terhadap distribusi dan karakteristik setiap kluster gaya belajar. Visualisasi dilakukan menggunakan fitur visualisasi pada *Orange Data Mining*, dan hasil interpretasi digunakan sebagai dasar dalam merumuskan implikasi strategi pembelajaran yang lebih adaptif dan sesuai dengan karakteristik masing-masing kelompok siswa.

4. HASIL DAN PEMBAHASAN

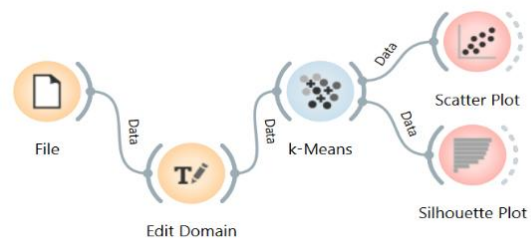
4.1. Hasil Pengumpulan dan Pra-proses Data

Data penelitian merupakan data primer yang diperoleh melalui penyebaran kuesioner kepada siswa SMKN 3 Tangerang Selatan menggunakan Google Form. Instrumen penelitian disusun menggunakan skala Likert lima tingkat, yaitu Sangat Tidak Setuju (STS), Tidak Setuju (TS), Cukup Setuju (CS), Setuju (S), dan Sangat Setuju (SS). Untuk keperluan analisis *data mining*, setiap kategori jawaban ditransformasikan ke dalam bentuk numerik dengan rentang nilai 1 hingga 5.

Berdasarkan hasil pengisian kuesioner, data kemudian dikelompokkan ke dalam tiga variabel utama yang merepresentasikan gaya belajar VAK, yaitu:

- Variabel *Visual* (V), yang merupakan akumulasi skor dari indikator pemahaman melalui gambar, diagram, dan media *visual*.
- Variabel *Auditory* (A), yang merupakan akumulasi skor dari indikator pemahaman melalui penjelasan lisan, diskusi, dan media audio.
- Variabel *Kinesthetic* (K), yang merupakan akumulasi skor dari indikator pembelajaran melalui praktik langsung, aktivitas fisik, dan eksperimen.

Pada tahap pra-proses data menggunakan *Orange Data Mining*, widget Edit Domain digunakan untuk memastikan kesesuaian tipe data, sedangkan widget Select Columns digunakan untuk menetapkan variabel *Visual*, *Auditory*, dan *Kinesthetic* sebagai fitur utama. Data yang telah melalui tahap pembersihan dan seleksi selanjutnya digunakan sebagai input dalam proses klastering menggunakan algoritma *K-Means*.



Gambar 2. Workflow Proses Klastering di *Orange Data Mining*

4.2. Penentuan Jumlah Kluster optimal

Penentuan jumlah kluster optimal (k) dilakukan secara objektif dengan memanfaatkan widget *K-Means* yang dihubungkan dengan widget *Silhouette Plot* pada *Orange Data Mining*. Pengujian dilakukan dengan variasi jumlah kluster dari $k = 2$ hingga $k = 5$.

Hasil evaluasi menunjukkan bahwa nilai *Silhouette Score* tertinggi diperoleh pada $k = 3$, yaitu sebesar 0,35. Nilai ini mengindikasikan bahwa struktur kluster yang terbentuk memiliki tingkat kohesi dan separasi yang cukup baik, sehingga pemisahan antar kluster dapat dikatakan optimal untuk data gaya belajar VAK siswa SMKN 3 Tangerang Selatan.

Meskipun nilai *Silhouette Score* berada pada kategori sedang, nilai ini masih menunjukkan struktur kluster yang dapat diterima untuk data sosial dan pendidikan yang bersifat heterogen.

4.3. Hasil Klastering *K-Means*

Berdasarkan hasil pengolahan data menggunakan algoritma *K-Means* dengan jumlah kluster = 3, diperoleh tiga kluster dengan nilai *centroid* sebagaimana ditunjukkan pada Tabel 1.

Nilai *centroid* yang ditampilkan merupakan hasil agregasi dan transformasi data dalam proses klastering, sehingga tidak merepresentasikan rentang hasil normalisasi mentah (0–1), melainkan digunakan untuk membandingkan kecenderungan relatif antar kluster.

Tabel 1. Nilai Rata-rata *Centroid* Tiap Kluster

Kluster	Centroid			Interpretasi Gaya Belajar
	Visual	Auditory	Kinesthetic	
C1	1.19	1.00	1.29	Visual–Kinesthetic (Mixed)
C2	0.84	0.96	0.75	Low Engagement (Passive)
C3	1.33	2.00	2.55	Active Kinesthetic Dominant

Berdasarkan nilai *centroid* tersebut, setiap kluster menunjukkan karakteristik gaya belajar yang berbeda. Namun demikian, hasil klastering juga menunjukkan bahwa gaya belajar siswa tidak bersifat eksklusif pada satu aspek VAK tertentu, melainkan cenderung bersifat multidimensional.

4.4. Klaster C1 (*Visual–Kinesthetic Moderat*).

Klaster C1 memiliki nilai *centroid* yang relatif seimbang antara aspek *visual* (1.19) dan *kinesthetic* (1.29), dengan nilai *auditory* yang sedikit lebih rendah (1.00). Kondisi ini menunjukkan bahwa siswa dalam klaster ini cenderung memahami materi pembelajaran melalui kombinasi tampilan *visual* dan aktivitas praktik, sehingga dapat dikategorikan sebagai kelompok dengan kecenderungan *visual–kinesthetic*. Siswa pada klaster ini umumnya lebih mudah memahami materi ketika pembelajaran disertai media *visual* dan diikuti dengan aktivitas langsung.

4.5. Klaster C2 (*Low Engagement Learner*).

Klaster C2 memiliki nilai *centroid* yang relatif rendah dan merata pada ketiga aspek VAK, yaitu *visual* (0.84), *auditory* (0.96), dan *kinesthetic* (0.75). Pola ini mengindikasikan bahwa siswa dalam klaster ini tidak menunjukkan dominasi gaya belajar tertentu. Hal ini dapat diinterpretasikan sebagai kelompok siswa dengan tingkat keterlibatan belajar yang relatif rendah atau belum memiliki preferensi gaya belajar yang kuat. Oleh karena itu, siswa pada klaster ini memerlukan pendekatan pembelajaran yang lebih bervariasi dan intensif untuk meningkatkan keterlibatan mereka dalam proses belajar.

4.6. Klaster C3 (*Kinesthetic-Dominant Active Learner*)

Klaster C3 menunjukkan nilai *centroid* tertinggi dibandingkan klaster lainnya, khususnya pada aspek *kinesthetic* (2.55), diikuti oleh *auditory* (2.00) dan *visual* (1.33). Pola ini mengindikasikan bahwa siswa dalam klaster ini memiliki kecenderungan gaya belajar *kinesthetic* yang dominan, namun tetap didukung oleh kemampuan *auditory* dan *visual* yang cukup baik. Siswa dalam klaster ini dapat dikategorikan sebagai pembelajar aktif yang sangat responsif terhadap pembelajaran berbasis praktik, diskusi, dan pengalaman langsung.

4.7. Implikasi Strategi Pembelajaran Berdasarkan Klaster

Berdasarkan karakteristik masing-masing klaster strategi pembelajaran yang dapat diterapkan adalah sebagai berikut:

- Klaster C1 (*Visual–Kinesthetic Moderat*). Disarankan menggunakan media pembelajaran *visual* seperti slide presentasi, diagram, dan video pembelajaran yang dikombinasikan dengan latihan praktik sederhana.
- Klaster C2 (*Low Engagement Learner*). Diperlukan pendekatan pembelajaran yang lebih interaktif dan bervariasi, seperti diskusi kelompok kecil, pembelajaran kolaboratif, serta pendampingan lebih intensif untuk meningkatkan motivasi belajar.
- Klaster C3 (*Kinesthetic-Dominant Active Learner*). Disarankan memperbanyak pembelajaran berbasis

praktik, simulasi, eksperimen, dan project-based learning untuk mengoptimalkan potensi belajar siswa.

4.8. Kebaruan dan Kontribusi Penelitian

Kebaruan penelitian ini terletak pada penerapan klastering gaya belajar VAK berbasis *data mining* pada konteks SMK dengan karakteristik pendidikan vokasi, yang menekankan dominasi pembelajaran kinestetik.

Hasil penelitian ini menunjukkan bahwa pendekatan *data mining* berpotensi mendukung personalisasi pembelajaran di sekolah kejuruan, sehingga guru dapat merancang strategi pembelajaran yang lebih sesuai dengan karakteristik dan kebutuhan belajar siswa.

5. KESIMPULAN DAN SARAN

Hasil penelitian menunjukkan bahwa penerapan algoritma *K-Means* menggunakan *Orange Data Mining* mampu mengidentifikasi pola pengelompokan gaya belajar siswa secara sistematis. Pengelompokan ini menghasilkan tiga klaster utama ($k=3$) dengan validasi *Silhouette Score* sebesar 0,35, yang terdiri dari klaster C1 (*Visual–Kinesthetic Moderat*), klaster C2 (*Low Engagement*), dan klaster C3 (*Active Kinesthetic Dominant*) sebagai kelompok mayoritas yang mencerminkan karakteristik pendidikan vokasi. Temuan ini mengimplikasikan pentingnya personalisasi strategi pembelajaran, di mana disarankan bagi guru untuk menerapkan metode *Project-Based Learning* dan simulasi praktik bagi kelompok dominan kinestetik serta melakukan pendampingan personal bagi kelompok dengan keterlibatan rendah (C2), sementara bagi penelitian selanjutnya disarankan untuk mengkomparasikan algoritma *K-Means* dengan metode lain seperti *Hierarchical Clustering* atau *K-Medoids* dengan jumlah sampel yang lebih besar untuk meningkatkan akurasi pengelompokan.

DAFTAR PUSTAKA

- [1] P. W. Rahayu et al., *Buku Ajar Data Mining*, 1st ed. Jambi: PT. Sonpedia Publishing Indonesia, 2024.
- [2] F. Andriani and N. Nugraheni, “Analisis Karakteristik Gaya Belajar Siswa dalam Pembelajaran Berdiferensiasi,” *Jurnal Riset Pendidikan Dasar (JRPD)*, vol. 5, no. 1, p. 33, Mar. 2024, doi: 10.30595/jrpd.v5i1.16067.
- [3] S. Azis, A. Yurni Ulfa, F. Akbar, and H. Mutiah, “Analisis Gaya Belajar *Visual*, Auditori, dan Kinestetik (VAK) pada Pembelajaran Biologi Siswa SMAN 8 Bulukumba,” *Jurnal Bioshell: Jurnal Pendidikan Biologi, Biologi, dan Pendidikan IPA*, vol. 11, no. 2, pp. 90–99, Oct. 2022, doi: 10.56013/bio.v11i2.1684.
- [4] A. T. Jatmiko, W. Sukmo Wardhono, and S. H. Wijoyo, “Analisis Komparasi Algoritme C4.5 dan Naive Bayes dalam Kasus Klasifikasi

- Kecenderungan Gaya Belajar *Visual Auditori Kinestetik* (VAK),” vol. 7, no. 7, pp. 3507–3516, Jul. 2023, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [5] N. Hendrastuty, “Penerapan *Data Mining* Menggunakan Algoritma *K-Means* Klastering Dalam Evaluasi Hasil Pembelajaran Siswa,” *Jurnal Ilmiah Informatika dan Ilmu Komputer (JIMA-ILKOM)*, vol. 3, no. 1, pp. 46–56, Mar. 2024, doi: 10.58602/jima-ilkom.v3i1.26.
- [6] A. Yudistira and R. Andika, “Pengelompokan Data Nilai Siswa Menggunakan Metode *K-Means* Klastering,” *Journal of Artificial Intelligence and Technology Information (JAITI)*, vol. 1, no. 1, pp. 20–28, Mar. 2023, doi: 10.58602/jaiti.v1i1.22.
- [7] K. Setiawan and E. Mailoa, “Pengelompokan Gaya Belajar Siswa Untuk Mendukung Efektifitas Pembelajaran Dengan *K-Means*,” *JTIK (Jurnal Teknologi Informasi dan Komunikasi)*, vol. 16, no. 2, pp. 26–35, Sep. 2025, [Online]. Available: <http://ejurnal.provisi.ac.id/index.php/JTIKP>
- [8] D. Lestariwati, Q. Mushafanah, and Kiswiyo, “Analisis Gaya Belajar Siswa Berprestasi Kelas V di SD Negeri Bancak 01 Kecamatan Gunungwungkal Kabupaten Pati,” *Dwijaloka Jurnal Pendidikan Dasar & Menengah*, vol. 2, no. 4, pp. 464–475, Dec. 2021.
- [9] M. Arhani and M. Nasir, *Data Mining: Algoritma dan Implementasi*, 1st ed. Yogyakarta: Andi, 2020.
- [10] F. Matheos Sarimole and L. Hakim, “Klasifikasi Barang Menggunakan Metode Klastering *K-Means* Dalam Penentuan Prediksi Stok Barang,” *Jurnal Sains dan Teknologi*, vol. 5, no. 3, pp. 846–854, Feb. 2024, doi: 10.55338/saintek.v5i1.2709