

PENERAPAN K-MEANS CLUSTERING UNTUK PENGELOMPOKAN KINERJA SISWA DALAM MATA PELAJARAN PRODUKTIF

Haula Hasna Nabila*¹, Maulana Fansyuri²

^{1,2}Universitas Pamulang; Jl. Raya Puspitek, Buaran. Kec. Pamulang, Kota Tangerang Selatan, Banten 15310, (021) 7412566

Jurusan Teknik Informatika, Fakultas Teknik, Universitas Pamulang

e-mail: *nabilahaula5@gmail.com, ²dosen02359@unpam.ac.id

Abstrak

Peningkatan kualitas pembelajaran di Sekolah Menengah Kejuruan (SMK) membutuhkan pemahaman yang lebih mendalam mengenai variasi kinerja siswa, khususnya pada mata pelajaran produktif yang berperan langsung dalam membentuk kompetensi kerja. Pada praktiknya, SMK IQRO masih melakukan pengelompokan kinerja siswa secara manual melalui observasi guru dan penilaian nilai akhir. Metode tersebut cenderung membutuhkan waktu, bersifat subjektif, dan kurang efisien ketika jumlah siswa semakin banyak. Seiring berkembangnya teknologi, analisis berbasis data menjadi salah satu solusi untuk mendapatkan gambaran kinerja siswa secara objektif dan terukur. Penelitian ini menerapkan algoritma K-Means Clustering sebagai metode pengelompokan siswa berdasarkan nilai mata pelajaran produktif, dengan tujuan menghasilkan pembagian kelompok kinerja yang lebih akurat dan dapat mendukung perencanaan pembelajaran.

Algoritma K-Means digunakan untuk memetakan siswa ke dalam beberapa cluster berdasarkan kemiripan nilai mereka. Data yang diolah meliputi nilai hasil belajar pada mata pelajaran produktif yang diperoleh dalam satu semester. Hasil penelitian menunjukkan bahwa metode K-Means mampu membentuk cluster yang merepresentasikan kelompok siswa berkinerja tinggi, sedang, dan rendah. Proses clustering menghasilkan nilai Sum of Squared Error (SSE) sebesar 583,33 dengan waktu pemrosesan kurang dari 1 detik, yang menunjukkan bahwa metode yang diterapkan memiliki performa komputasi yang efisien dan stabil. Pengelompokan ini memberikan manfaat bagi guru untuk menyusun strategi pembelajaran yang lebih tepat sasaran, seperti pemberian program remedial, pendampingan tambahan, maupun pengayaan bagi siswa yang berprestasi. Selain itu, hasil penelitian ini juga dapat menjadi referensi bagi sekolah dalam pengambilan keputusan akademik berbasis data, sekaligus mendukung peningkatan efektivitas proses pembelajaran di mata pelajaran produktif.

Kata kunci : K-Means Clustering, SMK IQRO, Pengelompokan Siswa, Kinerja Siswa, Mata Pelajaran Produktif

I. PENDAHULUAN

Pendidikan memiliki peranan penting dalam meningkatkan kualitas sumber daya manusia yang mampu bersaing di era globalisasi. Di Indonesia, pendidikan kejuruan berfungsi untuk menyiapkan tenaga kerja terampil yang sesuai dengan kebutuhan dunia industri. SMK sebagai bagian dari sistem pendidikan nasional dirancang untuk membekali siswa dengan keterampilan praktis, terutama melalui mata

pelajaran produktif yang berfokus pada pengembangan kompetensi khusus yang dibutuhkan di dunia kerja.

Dalam pelaksanaan pembelajaran, khususnya pada mata pelajaran produktif, sering ditemukan variasi kinerja siswa yang cukup beragam. Sebagian siswa dapat memahami materi dengan cepat, sementara yang lain memerlukan bimbingan lebih intensif untuk mencapai hasil belajar yang diharapkan. Variasi ini dipengaruhi oleh beberapa faktor, seperti latar belakang akademik, motivasi belajar, serta kemampuan individual. Oleh sebab itu, mengidentifikasi dan memahami perbedaan kinerja ini menjadi penting agar

guru dapat menerapkan pendekatan pembelajaran yang lebih sesuai dengan kebutuhan masing-masing siswa.

Pengelompokan kinerja siswa merupakan langkah strategis untuk meningkatkan efektivitas pembelajaran. Melalui pengelompokan ini, sekolah dapat mengenali siswa dengan kinerja tinggi, sedang, dan rendah, sehingga guru dapat menyusun strategi pengajaran yang lebih tepat sasaran. Namun, pengelompokan di SMK IQRO masih dilakukan secara manual berdasarkan nilai dan observasi, sehingga membutuhkan waktu lama dan rentan terhadap subjektivitas dalam penilaian.

Dengan perkembangan teknologi, metode pengelompokan berbasis data semakin banyak digunakan dalam dunia pendidikan. Data mining menjadi salah satu pendekatan yang dapat membantu menemukan pola tersembunyi yang mendukung pengambilan keputusan (Zaki & Meira, 2020). Salah satu algoritma data mining yang populer adalah *K-Means Clustering*, yaitu metode *unsupervised learning* yang mampu membagi data menjadi beberapa cluster berdasarkan tingkat kemiripan. Dalam dunia pendidikan, algoritma ini dapat digunakan untuk mengelompokkan siswa berdasarkan pola kinerjanya. Penelitian sebelumnya menunjukkan bahwa *K-Means* efektif untuk membantu guru menganalisis capaian akademik siswa (Yudhistira & Andika, 2023).

Penerapan *K-Means Clustering* di SMK IQRO diharapkan dapat menghasilkan pengelompokan siswa yang lebih objektif dan akurat. Hasil cluster ini dapat menjadi acuan bagi guru dalam menentukan intervensi pembelajaran, seperti pemberian remedial bagi siswa yang mengalami kesulitan atau pengayaan bagi siswa yang berprestasi. Pendekatan ini tidak hanya meningkatkan efektivitas pembelajaran, tetapi juga dapat mendorong motivasi belajar siswa karena pembelajaran lebih sesuai dengan kebutuhan mereka.

Hasil pengelompokan ini juga dapat mendukung sekolah dalam merumuskan kebijakan akademik yang lebih tepat sasaran, seperti penyediaan program pendampingan atau pelatihan tambahan sesuai kebutuhan kelompok siswa. Dengan demikian, penggunaan metode *K-Means Clustering* dapat membantu SMK IQRO mewujudkan tujuan pendidikan kejuruan dalam menghasilkan siswa yang kompeten dan siap bersaing di dunia industri.

Berbeda dengan penelitian sebelumnya yang umumnya hanya membahas penerapan *K-Means Clustering* pada dataset umum atau simulasi, penelitian ini menggunakan data akademik riil siswa SMK serta menekankan aspek implementasi praktis melalui sistem berbasis web. Penelitian ini tidak hanya berfokus pada hasil pengelompokan, tetapi juga mempertimbangkan

performa komputasi metode clustering berdasarkan indikator evaluasi yang terukur. Kebaruan penelitian ini terletak pada integrasi metode clustering dengan sistem pengolahan nilai siswa yang mampu memberikan hasil analisis secara cepat, objektif, dan efisien untuk mendukung evaluasi pembelajaran.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk:

1. Mengelompokan kinerja siswa berdasarkan data nilai mata pelajaran produktif.
2. Menganalisis efektivitas algoritma *K-Means Clustering* dalam proses pengelompokan siswa.
3. Mengukur performa komputasi metode clustering dalam sistem yang dikembangkan.

II. METODE PELAKSANAAN

Penelitian ini menggunakan pendekatan data mining dengan menerapkan algoritma *K-Means Clustering* untuk mengelompokkan kinerja siswa berdasarkan nilai mata pelajaran produktif. Tahapan penelitian meliputi pengumpulan data, praproses data, proses clustering, serta evaluasi hasil. Data yang digunakan berupa nilai akademik siswa yang diperoleh dalam satu semester.

Proses clustering dilakukan dengan menentukan jumlah cluster (K), menghitung jarak antara data dan centroid, serta melakukan iterasi hingga mencapai kondisi konvergen. Evaluasi kualitas clustering dilakukan menggunakan parameter *Sum of Squared Error* (SSE), yang digunakan untuk mengukur tingkat kesalahan dan homogenitas dalam setiap cluster.

1. Tahap Pengumpulan Data

a. Observasi

Observasi dilakukan dengan mengamati secara langsung kegiatan pembelajaran mata Pelajaran produktif di SMK IQRO. Metode ini bertujuan untuk memperoleh gambaran umum mengenai kondisi pembelajaran serta perilaku belajar siswa yang dapat memengaruhi kinerja siswa.

b. Wawancara

Wawancara dilakukan dengan pihak-pihak terkait, seperti guru mata Pelajaran produktif dan wali kelas, untuk mendapatkan informasi tambahan mengenai faktor-faktor yang memengaruhi kinerja siswa serta kendala apa yang dihadapi dalam proses pembelajaran.

c. Dokumentasi Sekolah

Penulis akan mengumpulkan dokumen-dokumen pendukung dari pihak sekolah berupa data nilai hasil ujian siswa pada mata Pelajaran produktif. Data ini

akan dijadikan sebagai bahan utama dalam analisis untuk mengelompokkan siswa berdasarkan kinerja mereka dengan menggunakan metode *K-Means Clustering*.

2. Tahap Perancangan Sistem

Berikut merupakan tahap analisis yang dilakukan:

- a. Merancang arsitektur sistem berbasis web yang mencakup alur komunikasi antara pengguna, server, dan proses pengolahan data clustering.
- b. Menyusun rancangan basis data menggunakan Entity Relationship Diagram (ERD) dan Logical Record Structure (LRS) agar struktur data lebih jelas dan sesuai kebutuhan sistem.
- a. Membuat diagram UML yang meliputi *use case diagram* untuk menggambarkan kebutuhan fungsional sistem, *activity diagram* untuk menjelaskan alur proses dari setiap fungsi utama, *sequence diagram* untuk menunjukkan interaksi antar objek selama proses berlangsung, dan *class diagram* untuk merepresentasikan struktur kelas, atribut, dan relasi dalam sistem.
- c. Mendesain tampilan antarmuka (UI) yang sederhana, informatif, dan mudah digunakan sehingga memudahkan pengguna dalam menginput data, menjalankan proses clustering, serta melihat hasil pengelompokan siswa.

3. Tahap Implementasi

Berikut adalah tahap implementasi:

- a. Mengimplementasikan desain sistem ke dalam aplikasi berbasis web menggunakan bahasa pemrograman dan teknologi yang telah dipilih.
- b. Membangun dan mengonfigurasi database sesuai rancangan ERD dan LRS untuk menyimpan data siswa, nilai, serta hasil clustering.
- c. Mengintegrasikan algoritma K-Means Clustering ke dalam sistem, termasuk proses inialisasi centroid, perhitungan jarak, dan pembentukan cluster akhir.
- d. Mengembangkan modul input data, modul proses clustering, dan modul tampilan hasil agar seluruh fungsi utama sistem dapat digunakan dengan baik.
- e. Mengintegrasikan seluruh komponen dan antarmuka pengguna (UI) sehingga sistem berjalan secara terhubung, responsif, dan sesuai kebutuhan pengguna.

4. Tahap Pengujian

a. White Box Testing

Pengujian ini dilakukan untuk memeriksa logika internal dan alur proses dalam kode program.

White box testing memastikan bahwa setiap fungsi, perhitungan algoritma K-Means, proses iterasi, perhitungan jarak, serta pembentukan cluster telah berjalan sesuai dengan struktur logika yang dirancang. Pengujian dilakukan pada level modul untuk memastikan tidak ada kesalahan perhitungan atau proses loop yang mengakibatkan hasil cluster tidak sesuai.

b. Black Box Testing

Pengujian ini dilakukan untuk memastikan sistem berjalan dengan benar dari sisi pengguna tanpa melihat kode program. Black box testing menguji seluruh fitur, seperti input data siswa, proses clustering, menampilkan hasil cluster, dan penyimpanan data. Pengujian dilakukan dengan memberikan berbagai skenario input untuk melihat apakah sistem memberikan output yang sesuai dengan yang diharapkan.

5. Tahap Evaluasi dan Penyusunan

- a. Melakukan evaluasi terhadap kinerja sistem dan memastikan seluruh fungsi berjalan sesuai tujuan, termasuk akurasi hasil clustering dan kemudahan penggunaan.
- b. Mengidentifikasi kekurangan serta menerima masukan dari pengguna untuk mengetahui aspek-aspek yang perlu diperbaiki atau disempurnakan.
- c. Menyusun laporan penelitian yang memuat seluruh proses mulai dari analisis, perancangan, implementasi, pengujian, hingga hasil evaluasi sebagai bagian dari dokumentasi akhir.

4. Algoritma K-Means

K-Means Clustering adalah metode *unsupervised learning* dalam data mining yang digunakan untuk mengelompokkan data ke dalam beberapa cluster berdasarkan tingkat kemiripan karakteristik. Prinsip dasarnya adalah meminimalkan jarak antara data dan pusat cluster (centroid), sehingga data dalam satu cluster memiliki kemiripan lebih tinggi dibandingkan data pada cluster lain. Metode ini membagi data ke dalam cluster yang memiliki karakteristik serupa dan berbeda dari cluster lainnya (Mardalius, 2018). Algoritma K-Means bekerja berdasarkan centroid dengan tiga parameter utama: jumlah cluster, inialisasi awal, dan perhitungan jarak (Susliansyah, 2019). Menurut Pratiwi (2022), K-Means merupakan algoritma clustering yang sederhana dan paling banyak digunakan. Diperkenalkan oleh Stuart Lloyd pada 1984, algoritma ini populer karena mudah diimplementasikan,

cepat, dan efisien untuk data besar. Selain itu, K-Means dapat dimodifikasi atau digabungkan dengan algoritma lain seperti Genetic Algorithm untuk meningkatkan akurasi (Hussein, 2018).

Dalam penelitian ini, K-Means digunakan untuk mengelompokkan kinerja siswa pada mata pelajaran produktif di SMK IQRO, dengan data berupa nilai ulangan, nilai praktik, kehadiran, atau indikator kinerja lainnya. Hasil cluster membagi siswa ke dalam kategori seperti tinggi, sedang, dan rendah, sehingga dapat menjadi acuan guru dalam merancang strategi pembelajaran yang sesuai. Proses K-Means berfokus pada pengelompokan data melalui perbandingan jarak dengan centroid dan memperbarui posisi centroid secara berulang hingga mencapai kondisi optimal. Karakteristik utama K-Means meliputi similarity-based grouping, centroid representation, dan iterative refinement.

Proses algoritma K-Means meliputi: menentukan jumlah cluster (k), memilih titik centroid awal secara acak, menghitung centroid menggunakan rumus:

$$V_{ij} = \frac{\sum_{k=0}^{N_i} x_{kj}}{N_i}$$

Keterangan:

V_{ij} : Centroid atau nilai rata-rata cluster indeks ke-I pada variable indeks ke-j

N_i : Jumlah data cluster indeks ke-i

i, k : Indeks atau penanda pada cluster

j : Indeks atau penanda pada variable

x_{kj} : Nilai data indeks k variable indeks ke-j pada cluster tersebut.

Untuk setiap row data, hitung jarak minimum data tersebut dengan centroid. Terdapat beberapa metode perhitungan yang dapat digunakan dalam mengukur jarak minimum data ke centroid, diantaranya adalah metode *Euclidean Distance*, *Manhattan/City Block*, serta *Minkowsky*. Dalam penelitian ini, metode perhitungan yang digunakan adalah metode *Euclidean Distance*, rumus perhitungan yang dipakai seperti di bawah ini:

$$\text{Euclidean Distance } \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Keterangan:

n : Jumlah elemen

i : Indeks untuk menyatakan elemen ke-i

x_i : Koordinat titik x ke-i atau titik pertama

y_i : Koordinat titik y ke-i atau titik kedua

III. HASIL DAN PEMBAHASAN

1. Hasil Penelitian

Penelitian ini menggunakan data nilai siswa pada tiga mata pelajaran produktif, yaitu OOP, Basis Data, dan Web & Mobile. Total data yang digunakan dalam perhitungan manual adalah 24 siswa kelas XI RPL, sementara sistem keseluruhan digunakan pada 211 siswa dari berbagai kelas di SMK IQRO

Nilai akhir setiap siswa dihitung menggunakan formula:

Nilai Akhir = 0.3 × Tugas + 0.3 × Praktik + 0.4 × Ujian

Dataset kemudian diolah menggunakan algoritma K-Means Clustering dengan parameter berikut:

- K = 3 cluster (Tinggi, Sedang, Rendah)
- Metode inisialisasi = K-Means++
- Seed terbaik = 12345
- Target validasi SSE = 583.33

Proses Perhitungan Manual K-Means

Proses K-Means dilakukan dalam dua iterasi hingga centroid konvergen. Pergerakan centroid ditunjukkan pada tabel berikut:

Cluster	OOP	Basis Data	Web & Mobile
C1	77.2938	78.1875	76.6500
C2	43.9000	43.0600	41.4600
C3	93.5000	92.3000	93.4000

Dapat dilihat bahwa:

- C1 merepresentasikan kelompok kinerja sedang
 - C2 merepresentasikan kelompok kinerja rendah
 - C3 merepresentasikan kelompok kinerja tinggi
- Hasil perhitungan manual juga menunjukkan bahwa SSE konvergen pada **583.33**, sesuai dengan hasil sistem.

Hasil Implementasi Sistem

Sistem clustering yang dibangun menampilkan berbagai fitur:

- Halaman Proses Clustering
 - Menampilkan jumlah siswa, kelas, serta data yang siap diproses.
 - Admin dapat memilih jumlah cluster dan melakukan eksekusi analisis.
- Hasil Riwayat Clustering

- 1) Sistem menyimpan riwayat tiap proses clustering
 - 2) Rata-rata iterasi: 5,1 kali
 - 3) Waktu proses: 0,00 detik (sangat cepat)
- c. Laporan Clustering
- 1) Sistem menghasilkan laporan Excel yang memuat hasil cluster, nilai siswa, dan rekapitulasi.

Hasil clustering pada implementasi penuh (211 siswa) juga membagi siswa menjadi 3 level kinerja yang konsisten dengan hasil perhitungan manual.

a. Analisis Karakteristik Cluster

Hasil clustering menunjukkan bahwa data siswa dapat dikelompokkan ke dalam tiga cluster utama yang merepresentasikan variasi kinerja akademik. Cluster kinerja tinggi didominasi oleh siswa dengan capaian nilai yang relatif stabil dan konsisten pada seluruh mata pelajaran produktif. Hal ini mengindikasikan tingkat penguasaan materi yang baik serta kemampuan mempertahankan performa akademik secara berkelanjutan.

Cluster kinerja sedang merepresentasikan siswa dengan capaian nilai yang cukup baik, namun masih menunjukkan variasi pada beberapa mata pelajaran. Kondisi ini mencerminkan adanya perbedaan tingkat pemahaman atau konsistensi belajar. Sementara itu, cluster kinerja rendah terdiri dari siswa dengan nilai relatif lebih rendah dibanding kelompok lainnya, yang mengindikasikan perlunya perhatian akademik dan strategi pembelajaran yang lebih adaptif.

b. Makna Akademik dari Hasil Clustering

Pengelompokan yang dihasilkan melalui K-Means Clustering memberikan gambaran objektif mengenai distribusi kinerja siswa. Pendekatan ini memungkinkan identifikasi pola capaian akademik yang tidak mudah diamati melalui metode manual. Dengan clustering, evaluasi kinerja siswa tidak hanya didasarkan pada nilai individual, tetapi juga mempertimbangkan kemiripan karakteristik performa secara keseluruhan.

Hasil clustering menunjukkan bahwa variasi kinerja siswa merupakan fenomena yang terstruktur dan dapat dianalisis secara kuantitatif. Hal ini memperkuat pemanfaatan teknik data mining sebagai alat bantu analisis dalam mendukung proses evaluasi pembelajaran.

c. Analisis Pemilihan Jumlah Cluster

Penentuan jumlah cluster ($K=3$) didasarkan pada kebutuhan praktis evaluasi akademik di sekolah, yang secara umum mengklasifikasikan kinerja siswa ke dalam kategori tinggi, sedang, dan rendah. Pemilihan ini tidak hanya relevan secara konseptual, tetapi juga mempermudah interpretasi hasil clustering oleh pengguna.

Validitas jumlah cluster didukung oleh hasil evaluasi yang menunjukkan bahwa model clustering mampu membagi data secara representatif. Pembagian cluster yang terbentuk mencerminkan variasi kinerja siswa yang dapat digunakan sebagai dasar analisis akademik.

d. Evaluasi Kualitas Clustering

Evaluasi kualitas clustering dilakukan menggunakan parameter Sum of Squared Error (SSE). SSE digunakan untuk mengukur tingkat kedekatan data terhadap centroid cluster, sehingga nilai yang dihasilkan mencerminkan homogenitas data dalam setiap cluster. Nilai SSE yang diperoleh menunjukkan bahwa algoritma K-Means mampu menghasilkan pengelompokan data yang stabil.

Stabilitas hasil clustering mengindikasikan bahwa distribusi data siswa memiliki pola yang dapat dipetakan secara sistematis. Hal ini menegaskan bahwa metode clustering yang digunakan layak diterapkan dalam analisis kinerja akademik siswa.

Selain menunjukkan tingkat homogenitas data, nilai SSE juga merepresentasikan kualitas pemisahan antar cluster. Semakin kecil nilai SSE, semakin dekat jarak data terhadap centroid, yang menunjukkan bahwa variasi dalam cluster semakin rendah. Kondisi ini mengindikasikan bahwa anggota dalam cluster memiliki karakteristik yang relatif seragam dan konsisten.

Stabilitas nilai SSE dalam penelitian ini menunjukkan bahwa proses iterasi algoritma K-Means telah mencapai kondisi konvergen. Hal ini berarti centroid cluster tidak lagi mengalami perubahan signifikan, sehingga hasil pengelompokan dapat dianggap representatif. Dengan demikian, evaluasi clustering tidak hanya menunjukkan validitas komputasional, tetapi juga memperkuat relevansi hasil terhadap analisis kinerja akademik siswa.

e. Implikasi terhadap Strategi Pembelajaran

Informasi cluster yang dihasilkan dapat dimanfaatkan sebagai dasar penyusunan strategi pembelajaran yang lebih adaptif. Siswa dalam cluster

kinerja tinggi dapat memperoleh program pengayaan, cluster kinerja sedang dapat menjadi fokus penguatan materi, sedangkan cluster kinerja rendah dapat diberikan pendampingan tambahan atau remedial.

Pendekatan ini memungkinkan proses pembelajaran yang lebih terdiferensiasi sesuai kebutuhan siswa. Dengan demikian, clustering tidak hanya berfungsi sebagai metode analisis data, tetapi juga sebagai pendukung peningkatan efektivitas pembelajaran.

Pendekatan pembelajaran berbasis cluster memberikan peluang bagi guru untuk menerapkan strategi pengajaran yang lebih terdiferensiasi. Siswa dalam cluster kinerja rendah, misalnya, tidak hanya memerlukan penguatan materi, tetapi juga pendekatan pembelajaran yang lebih adaptif. Strategi seperti pendampingan bertahap, pemberian latihan terstruktur, serta variasi metode penyampaian materi dapat membantu meningkatkan pemahaman siswa secara gradual.

Di sisi lain, siswa dalam cluster kinerja tinggi memerlukan stimulasi akademik yang sesuai dengan tingkat kemampuannya. Program pengayaan, tugas berbasis proyek, maupun tantangan konseptual yang lebih kompleks dapat digunakan untuk mempertahankan motivasi belajar dan mendorong perkembangan kompetensi. Hal ini menunjukkan bahwa hasil clustering dapat dimanfaatkan sebagai dasar dalam merancang intervensi pembelajaran yang lebih efektif.

Dengan adanya pemetaan kinerja siswa yang terstruktur, proses evaluasi pembelajaran juga menjadi lebih sistematis. Guru dapat mengidentifikasi kebutuhan akademik setiap kelompok siswa secara lebih objektif, sehingga pengambilan keputusan terkait metode pengajaran tidak semata-mata bergantung pada observasi subjektif. Pendekatan ini berpotensi meningkatkan kualitas proses belajar mengajar serta efektivitas pencapaian kompetensi siswa.

f. Dukungan terhadap Pengambilan Keputusan Akademik

Hasil clustering dapat digunakan sebagai referensi dalam mendukung pengambilan keputusan akademik di tingkat sekolah. Informasi distribusi kinerja siswa membantu sekolah dalam merancang program pembelajaran, evaluasi capaian akademik, serta penyusunan kebijakan pendidikan berbasis data. Pendekatan berbasis data memberikan keuntungan dalam meningkatkan objektivitas dan akurasi proses evaluasi akademik.

Pemanfaatan hasil clustering memungkinkan pihak sekolah mengidentifikasi pola performa siswa secara lebih komprehensif. Informasi tersebut dapat digunakan untuk menentukan prioritas intervensi akademik, seperti program remedial, pengayaan, maupun pendampingan belajar. Dengan demikian, keputusan akademik tidak hanya didasarkan pada penilaian individual, tetapi juga mempertimbangkan pola distribusi kinerja siswa secara keseluruhan.

Selain itu, pendekatan berbasis data mendukung konsistensi dalam proses evaluasi akademik. Analisis yang dihasilkan sistem membantu meminimalkan potensi subjektivitas serta meningkatkan transparansi dalam pengambilan keputusan. Hal ini menunjukkan bahwa clustering memiliki peran strategis dalam mendukung manajemen akademik di lingkungan sekolah.

g. Penguatan Diskusi Validitas Clustering

Proses clustering yang dilakukan dalam penelitian ini menunjukkan bahwa algoritma K-Means mampu mengelompokkan data siswa secara konsisten. Konsistensi hasil clustering mengindikasikan bahwa distribusi data memiliki pola yang cukup jelas untuk dipetakan menggunakan pendekatan berbasis jarak. Hal ini memperlihatkan bahwa metode yang digunakan sesuai dengan karakteristik data akademik yang bersifat numerik.

Selain itu, stabilitas centroid yang diperoleh pada iterasi akhir menunjukkan bahwa algoritma telah mencapai kondisi konvergen. Kondisi tersebut memperkuat bahwa hasil clustering yang dihasilkan tidak bersifat acak, melainkan representatif terhadap struktur data. Dengan demikian, pendekatan clustering yang diterapkan dapat digunakan sebagai dasar analisis akademik yang objektif.

2. Pembahasan

Efektivitas K-Means dalam Mengelompokkan Kinerja Siswa

Berdasarkan hasil manual dan sistem, algoritma K-Means mampu:

- Menghasilkan cluster yang stabil dan konsisten
- Mengelompokkan nilai mata pelajaran produktif secara objektif
- Menampilkan pembagian kelompok yang jelas: tinggi, sedang, rendah

Keberhasilan ini diperkuat oleh validasi SSE yang sama antara perhitungan manual dan sistem, yang menunjukkan bahwa algoritma berjalan dengan benar.

Analisis Variasi Kinerja Siswa

Pengelompokan menunjukkan tiga pola utama:

- a. Cluster Tinggi: Siswa dengan nilai di atas 90 (contoh: Afifah, Keyla)
- b. Cluster Sedang: Siswa berada di sekitar nilai 75–80
- c. Cluster Rendah: Siswa seperti Chacha, Chika, dan Maitsah yang konsisten berada di rentang nilai 38–45

Perbedaan kinerja ini menunjukkan variasi kemampuan yang signifikan dan membutuhkan pendekatan pembelajaran berbeda.

Pemanfaatan Hasil Bagi Guru dan Sekolah

Hasil clustering telah terbukti bermanfaat untuk:

- a. Menentukan siswa yang membutuhkan remedial
 - b. Memberikan pembelajaran diferensiasi (differentiated learning)
 - c. Menentukan prioritas bimbingan bagi siswa berkinerja rendah
 - d. Mendukung perencanaan akademik berbasis data
- Guru dapat melihat pola kinerja per kelompok dengan lebih cepat tanpa penilaian manual.

Efisiensi Sistem Clustering

- a. Hasil implementasi sistem menunjukkan:
- b. Proses clustering berlangsung hampir instan (0,00 detik)
- c. Pengolahan otomatis menghilangkan kesalahan manual
- d. Sistem mampu memproses ratusan siswa sekaligus

Ini membuktikan bahwa sistem sangat efisien dan layak digunakan secara berkelanjutan oleh sekolah.

a. Analisis Konsistensi Hasil Clustering

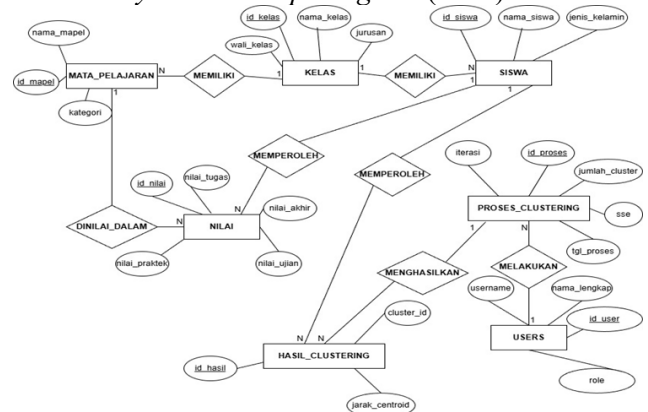
Hasil clustering yang diperoleh menunjukkan konsistensi antara perhitungan manual dan implementasi sistem. Kesamaan nilai SSE mengindikasikan bahwa algoritma K-Means berjalan sesuai dengan prinsip komputasi yang diharapkan. Konsistensi ini memperkuat validitas metode clustering dalam mengelompokkan data akademik siswa.

Stabilitas hasil clustering juga menunjukkan bahwa distribusi data siswa memiliki pola yang cukup jelas. Hal ini menegaskan bahwa pendekatan unsupervised learning mampu mengidentifikasi struktur data tanpa memerlukan label kelas sebelumnya. Dengan demikian, metode clustering yang digunakan tidak hanya akurat secara komputasional, tetapi juga relevan secara akademik.

A. Perancangan Basis Data

Perancangan Basis Data menjelaskan bagaimana struktur data disusun untuk mendukung kebutuhan sistem secara efektif dan teratur. Pada tahap ini, seluruh kebutuhan informasi yang telah dianalisis diterjemahkan menjadi kumpulan tabel beserta atributnya, lengkap dengan penentuan *primary key* dan *foreign key* yang mengatur hubungan antar tabel

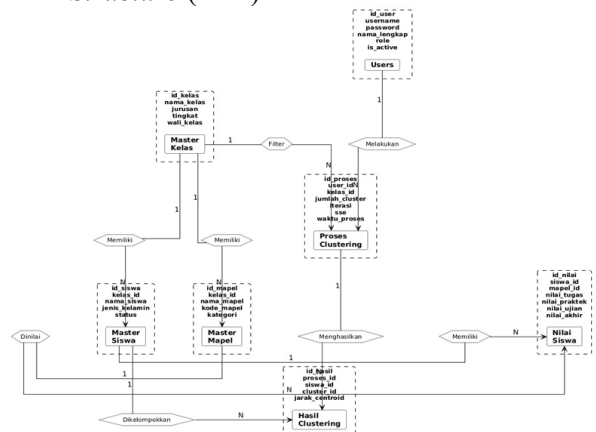
1. Entity Relationship Diagram (ERD)



Gambar 1. Entity Relationship Diagram

ERD tersebut memperlihatkan hubungan data dari pendataan siswa, kelas, dan mata pelajaran hingga nilai yang digunakan sebagai dasar proses clustering. Nilai tugas, praktik, ujian, dan nilai akhir tersimpan dalam entitas Nilai sebagai dataset analisis. Proses clustering dicatat dalam entitas proses_clustering bersama parameter penting seperti jumlah cluster, sedangkan hasil pengelompokan disimpan dalam entitas hasil_clustering yang menunjukkan cluster tiap siswa. Secara keseluruhan, ERD ini menggambarkan alur data akademik yang terintegrasi hingga tahap analisis K-Means.

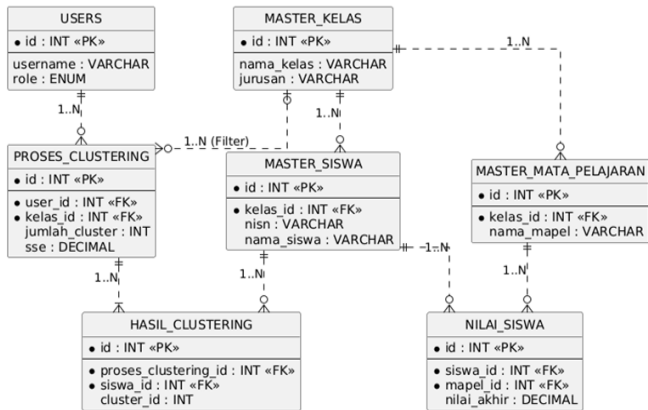
2. Transformasi ERD ke Logical Record Structure (LRS)



Gambar 2. Transformasi ERD ke Logical Record Structure (LRS)

LRS pada gambar tersebut menunjukkan hasil penerjemahan entitas dan relasi pada ERD menjadi struktur tabel yang siap diimplementasikan dalam basis data. Setiap entitas diubah menjadi tabel lengkap dengan atribut, primary key, dan foreign key, seperti *master_siswa*, *master_kelas*, *master_mata_pelajaran*, *nilai_siswa*, *proses_clustering*, dan *hasil_clustering*. Bentuk ini menampilkan susunan data secara teknis sehingga hubungan logis antar tabel terlihat jelas dan siap digunakan dalam proses pengolahan serta pengelompokan data menggunakan metode K-Means.

3. Logical Record Structure (LRS)



Gambar 3. Logical Record Structure (LRS)

Diagram diatas menggambarkan bahwa tabel Users berfungsi sebagai sumber informasi akun yang menjalankan proses clustering dan terhubung langsung dengan Proses_Clustering untuk mencatat pengguna, kelas yang dianalisis, serta parameter seperti jumlah cluster dan SSE. Proses_Clustering juga berkaitan dengan Master_Kelas, sementara entitas Master_Kelas, Master_Siswa, dan Master_Mata_Pelajaran merepresentasikan data akademik yang saling terhubung sebagai dasar penyusunan dataset. Nilai siswa tersimpan dalam tabel Nilai_Siswa dan digunakan dalam proses analisis, sedangkan hasil pengelompokan dicatat dalam Hasil_Clustering yang menghubungkan setiap siswa dengan cluster yang diperolehnya.

B. Pengujian Sistem

Pengujian sistem dilakukan untuk memastikan bahwa algoritma K-Means Clustering dan seluruh fungsi aplikasi berjalan sesuai dengan perancangan.

Pengujian mencakup validasi logika algoritma serta pengujian fungsionalitas sistem dari sisi pengguna.

Jenis Pengujian	Fokus Pengujian	Hasil
White Box Testing	Logika algoritma K-Means dan proses clustering	Valid
Black Box Testing	Fungsionalitas sistem dan kesesuaian output	Valid

White Box Testing dilakukan untuk memverifikasi bahwa proses perhitungan jarak, iterasi, dan pembentukan cluster berjalan dengan benar. Black Box Testing memastikan bahwa fitur sistem, seperti input data, proses clustering, dan penyajian hasil, dapat digunakan dengan baik. Hasil pengujian menunjukkan bahwa sistem berfungsi sesuai dengan kebutuhan dan tidak ditemukan kesalahan proses.

IV. KESIMPULAN

Kesimpulan Penelitian ini membuktikan bahwa penerapan algoritma K-Means Clustering dapat memberikan hasil pengelompokan kinerja siswa yang lebih objektif, akurat, dan konsisten dibandingkan metode manual yang sebelumnya digunakan di SMK IQRO. Sistem berbasis web yang dikembangkan mampu mengotomatisasi seluruh proses mulai dari penginputan nilai, pengolahan data, hingga penyajian hasil clustering, sehingga mempercepat evaluasi dan meminimalkan potensi kesalahan manusia. Hasil pengelompokan yang terbagi menjadi tiga kategori kinerja tinggi, sedang, dan rendah. Memberikan informasi yang lebih jelas mengenai variasi kemampuan siswa dan dapat dimanfaatkan guru untuk merancang strategi pembelajaran yang sesuai kebutuhan masing-masing kelompok, seperti remedial maupun materi pengayaan. Secara keseluruhan, penelitian ini menunjukkan bahwa integrasi metode K-Means dalam proses evaluasi akademik mampu meningkatkan efisiensi, mendukung pengambilan keputusan, dan berkontribusi pada peningkatan kualitas pembelajaran di sekolah.

DAFTAR PUSTAKA

- Adrian, T. &. (2023). Implementasi Data Mining untuk Mengklasifikasi Hasil Kelulusan Madrasah Menggunakan Algoritma Naive Bayes. Jakarta: Abadi.
- Alfares, H. K., & Duffuaa, S. O. (2020). Data Mining dan Penerapannya. Jakarta: Salemba Empat.
- Anggraeni, L., & Widiastuti, M. (2019). Implementasi K-Means untuk Pengelompokan Nilai Akademik. *Jurnal Teknologi Informasi*, 8(2), 45–52.
- Fariq Syafiq Sujana, Y. H. (2022). Rancang Bangun Sistem Informasi Pengelolaan Penyewaan Kamera Berbasis Web.
- Fathansyah. (2015). Basis Data Edisi Revisi. Bandung: Informatika.
- Firmansyah. (2020a). Sistem Informasi Pengaduan Warga Berbasis Website (Studi Kasus: Kelurahan Siantan Tengah, Pontianak Utara).
- Fitriani, A. M. (2023). Analysis and Design of Employee Management Information System Using the Web Based Waterfall Method. *JITSI: Jurnal Ilmiah Teknologi Sistem Informasi*.
- Hussein, H. (2018). Implementation of Genetic Algorithm for Improving K-Means Clustering Accuracy. *Journal of Computer Science*.
- Hussein, M. (2018). Optimization of K-Means Using Genetic Algorithm. *International Journal of Computer Applications*, 179(25), 1–6.
- Hutagalung, I. (2020). Pemetaan Siswa Kelas Unggulan Menggunakan Algoritma K-Means Clustering. *Jurnal Tek. Inform. dan Sist. Informasi (JATISI)*, 9(1).
- Lathifah, A. (2023). Peran Data Mining dalam Menghasilkan Wawasan dari Data Besar. *Jurnal Ilmu Komputer dan Bisnis*.
- Lestari, D. &. (2020). Penerapan Black Box Testing pada Pengujian Aplikasi Berbasis Web Menggunakan Teknik Equivalence Partitioning. *Jurnal Sistem Informasi dan Teknologi*, 8(3).
- Mardalius. (2018). Analisis Clustering Menggunakan Metode K-Means. *Jurnal Sistem Informasi*, 14(1), 12–19.
- Praniffa, A. C. (2023). Pengujian Black Box dan White Box Sistem Informasi Parkir Berbasis Web.
- Prasetyo, D. &. (2017). Implementasi Sistem Informasi Akademik Berbasis Web dengan Menggunakan PHP dan MySQLJu. *Jurnal Teknologi Informasi*, 10(1).
- Prasetyo, E. (2014). Data Mining: Mengolah Data Menjadi Informasi dengan Matlab. Yogyakarta: ANDI.
- Pratiwi, D. (2022). Penerapan Algoritma K-Means untuk Pengelompokan Data Akademik. *Jurnal Ilmiah Informatika*, 6(1), 30–37.
- Susliansyah. (2019). Penerapan Jarak Euclidean dalam Pengelompokan Data. *Jurnal Sains dan Teknologi*, 4(2), 55–63.
- Yudhistira, R., & Andika, P. (2023). Analisis Kinerja Akademik Menggunakan K-Means Clustering. *Jurnal Teknologi Pendidikan*, 11(1), 22–31.
- Zaki, M., & Meira, W. (2020). Data Mining and Machine Learning: Fundamental Concepts and Algorithms. Cambridge University Press.