**Analisis Cluster Mahasiswa Berdasarkan Faktor Gaya Belajar Dan Nilai Akademik Menggunakan Algoritma K-means**

**Rico Herdianti1, Risky Fauzan Gumilang2, Muhammad Syahrevi Perdana3, Nurul Isnayni4 , Cindi Deby5**

1,2,3,4,5 Universitas Bina Sarana Informatika

e-mail: 1xxxx@xxxx.xxx, 2xxxx@xxxx.xxx

3Institusi/Afiliasi

e-mail: xxxx@xxxx.xxx

|  |  |  |
| --- | --- | --- |
| Diterima  01-01-2020 | Direvisi  01-02-2020 | Disetujui  01-03-2020 |

**Abstrak** - Kemajuan teknologi kecerdasan buatan telah memberikan kontribusi signifikan dalam dunia pendidikan, terutama dalam mendukung pengembangan sistem pembelajaran yang adaptif dan personal. Penelitian ini bertujuan untuk mengidentifikasi pola gaya belajar mahasiswa serta keterkaitannya dengan performa akademik melalui penerapan algoritma K-Means sebagai metode klasterisasi. Data yang digunakan berasal dari dataset "Student Performance in Exams" yang diperoleh dari platform Kaggle, dengan total 10.000 entri mahasiswa. Atribut utama yang dianalisis mencakup gaya belajar pilihan (preferred learning style) dan nilai akhir (final grade), serta sejumlah variabel perilaku seperti jam tidur, tingkat stres, dan penggunaan media sosial. Analisis data dilakukan menggunakan perangkat lunak Weka dengan metode percentage split sebesar 50% untuk proses pelatihan dan pengujian model. Hasil klasterisasi menunjukkan bahwa mahasiswa dengan gaya belajar Visual dan Auditory cenderung memiliki performa akademik yang lebih tinggi dibandingkan gaya belajar lainnya. Temuan ini menunjukkan bahwa pemahaman terhadap gaya belajar mahasiswa dapat menjadi dasar dalam merancang strategi pembelajaran yang lebih efektif dan responsif. Selain itu, variabel perilaku juga berkontribusi terhadap hasil akademik, sehingga pendekatan pembelajaran sebaiknya disesuaikan dengan karakteristik individu secara menyeluruh.

Kata Kunci: K-Means, Gaya Belajar, Klasterisasi,

**PENDAHULUAN**

Kemajuan teknologi digital telah membawa perubahan fundamental dalam berbagai sektor kehidupan, termasuk di bidang pendidikan. Dalam beberapa dekade terakhir, digitalisasi telah menjadi kekuatan pendorong utama dalam transformasi sistem pendidikan, mulai dari metode pengajaran, sistem penilaian, hingga cara institusi mengevaluasi dan memahami perilaku belajar peserta didik. Salah satu kontribusi utama dari kemajuan teknologi adalah kemudahan akses terhadap informasi secara cepat dan luas. Tidak hanya itu, digitalisasi juga membuka peluang besar untuk mengimplementasikan pendekatan pembelajaran yang lebih personal dan adaptif, sehingga mampu menjawab kebutuhan belajar yang beragam dari masing-masing individu.

# Konsep AI dan ML secara fundamental memungkinkan sistem komputer untuk belajar dari data dan meningkatkan kemampuannya secara otomatis, tanpa perlu diprogram ulang secara eksplisit. Kemampuan ini membuka peluang besar dalam menciptakan teknologi cerdas yang dapat meniru bahkan melampaui kemampuan manusia dalam menyelesaikan permasalahan yang kompleks. Dalam implementasinya, teknologi AI dan ML tidak hanya terbatas pada industri tertentu, melainkan telah merambah ke berbagai sektor, seperti pendidikan, kesehatan, keuangan, manufaktur, dan lainnya.

# Seiring dengan meningkatnya jumlah data yang tersedia secara digital, pendekatan tradisional dalam menganalisis data tidak lagi memadai. Pengolahan data secara manual cenderung tidak efisien dan kurang akurat, terutama dalam menghadapi big data. Oleh karena itu, dibutuhkan metode yang canggih dan adaptif untuk mengekstrak informasi yang tersembunyi dalam kumpulan data tersebut. Salah satu pendekatan yang mampu mengatasi permasalahan ini adalah machine learning, yaitu metode pembelajaran mesin yang merupakan perpaduan antara artificial intelligence dan computer science (Müller & Guido, 2016).

# Dalam machine learning, terdapat berbagai teknik yang dapat digunakan untuk menganalisis data, salah satunya adalah teknik clustering. Teknik ini termasuk dalam kategori unsupervised learning, di mana proses pembelajaran dilakukan berdasarkan data yang tidak berlabel (unlabeled data). Salah satu algoritma clustering yang paling populer dan banyak digunakan dalam penelitian dan praktik industri adalah algoritma K-Means. Algoritma ini bekerja Dengan cara mengelompokkan data ke dalam sejumlah cluster (kelompok) yang telah ditentukan sebelumnya, dengan tujuan meminimalkan jarak antar data di dalam cluster yang sama dan memaksimalkan jarak antar cluster yang berbeda.

Di tengah pesatnya transformasi digital, kecerdasan buatan *(Artificial Intelligence/AI)* muncul sebagai teknologi kunci yang semakin banyak diadopsi oleh institusi pendidikan. *AI* memungkinkan komputer untuk melakukan tugas-tugas kompleks seperti menganalisis data, mengenali pola, hingga membuat keputusan layaknya manusia. Keunggulan ini memberikan peluang besar bagi dunia pendidikan untuk mengembangkan sistem pembelajaran yang lebih efisien, terukur, dan berbasis data. Salah satu implementasi konkret dari *AI* dalam pendidikan adalah kemampuannya untuk membantu pengajar dan institusi dalam memahami perilaku belajar siswa secara lebih akurat, memberikan rekomendasi pembelajaran individual, serta memprediksi performa akademik siswa di masa mendatang.

Salah satu cabang *AI* yang paling banyak digunakan adalah *machine learning* (pembelajaran mesin). Menurut Han dan Kamber (2006), *machine learning* adalah proses di mana sistem komputer belajar dari data historis dan meningkatkan kemampuannya secara otomatis tanpa harus diprogram ulang secara eksplisit. Dalam pendidikan*, machine learning* telah digunakan dalam berbagai aplikasi seperti sistem rekomendasi materi pembelajaran, klasifikasi tingkat pemahaman siswa, hingga prediksi *dropout rate*. Salah satu pendekatan utama dalam *machine learning* adalah teknik *unsupervised learning,* di mana sistem mempelajari struktur data tanpa label, dan salah satu metode paling umum dalam pendekatan ini adalah teknik *clustering* atau pengelompokan data.

*Clustering* merupakan teknik yang digunakan untuk mengelompokkan data ke dalam kelompok-kelompok *(klaster)* yang memiliki kemiripan karakteristik satu sama lain. Tujuan utama dari clustering adalah untuk menemukan struktur tersembunyi di dalam data, terutama ketika data tidak memiliki label atau kategori yang telah ditentukan sebelumnya. Dalam dunia pendidikan, teknik ini memiliki potensi besar, terutama dalam konteks eksplorasi perilaku siswa. Sebagai contoh, *clustering* dapat digunakan untuk mengidentifikasi kelompok siswa dengan preferensi belajar yang serupa, tingkat stres akademik yang sebanding, atau bahkan respon yang mirip terhadap metode pengajaran tertentu. Dengan pendekatan ini, institusi pendidikan dapat menyusun strategi pembelajaran yang lebih tepat sasaran dan efisien.

Salah satu algoritma clustering yang paling populer adalah *K-Means*. Algoritma ini bekerja dengan menentukan sejumlah pusat klaster *(centroid)* secara acak, kemudian mengelompokkan data berdasarkan kedekatan nilai terhadap *centroid* tersebut. Proses ini dilakukan secara iteratif, di mana posisi *centroid* akan diperbarui terus-menerus hingga stabil atau konvergen. Keunggulan *K-Means* terletak pada kesederhanaan implementasinya, kecepatan prosesnya, serta kemampuannya dalam menangani dataset berukuran besar. Tidak heran jika algoritma ini banyak digunakan dalam berbagai bidang, termasuk pemasaran, kesehatan, dan pendidikan. Studi oleh Hastuti, Saptono, dan Suryani (2012) menunjukkan bahwa *K-Means* efektif dalam mengelompokkan calon penerima beasiswa berdasarkan data akademik dan sosial ekonomi secara akurat, menegaskan aplikabilitas algoritma ini dalam konteks pendidikan.

# Algoritma K-Means dikenal karena kesederhanaannya dalam implementasi serta efisiensinya dalam mengolah data berukuran besar. Selain itu, algoritma ini bersifat iteratif, di mana proses pengelompokan dilakukan secara berulang hingga diperoleh hasil clustering yang optimal. Dalam konteks pendidikan, algoritma K-Means dapat diterapkan untuk mengelompokkan siswa berdasarkan performa akademik mereka, sehingga

Penelitian ini mengambil pendekatan berbeda dengan menerapkan algoritma *K-Means* untuk mengelompokkan mahasiswa berdasarkan gaya belajar dan performa akademik mereka. Gaya belajar merupakan aspek penting yang memengaruhi efektivitas proses belajar, karena setiap mahasiswa memiliki cara unik dalam menyerap, mengolah, dan menyimpan informasi. Tipe gaya belajar yang umum dikenal antara lain *visual, auditori*, dan *kinestetik,* dan masing-masing membutuhkan strategi pembelajaran yang berbeda. Selain gaya belajar, performa akademik yang direpresentasikan melalui nilai akhir juga menjadi fokus karena mencerminkan hasil nyata dari proses pembelajaran. Kombinasi antara kedua atribut ini dapat memberikan gambaran yang komprehensif mengenai karakteristik mahasiswa, yang sangat berguna dalam menyusun pendekatan pembelajaran yang lebih personal dan *adaptif.*

**METODE PENELITIAN**

Penelitian ini menggunakan pendekatan ***kuantitatif*** karena bertujuan untuk menganalisis data numerik secara *objektif* dan sistematis guna memperoleh pemahaman yang mendalam terkait pengelompokan mahasiswa berdasarkan gaya belajar dan performa akademik mereka. Pendekatan ini dipilih untuk memastikan hasil penelitian dapat diukur, diuji ulang, serta digeneralisasikan pada populasi yang lebih luas. Sugiyono (2013) menyebutkan bahwa penelitian kuantitatif bertujuan untuk menguji *hipotesis* atau menjawab pertanyaan penelitian melalui data yang dikumpulkan dengan instrumen terstandar dan dianalisis menggunakan prosedur statistik. Sejalan dengan itu, Creswell (2014) menekankan bahwa metode kuantitatif memungkinkan peneliti untuk bekerja dengan data yang dapat diuji secara objektif dan direplikasi.

Data yang digunakan dalam penelitian ini bersumber dari situs ***Kaggle,*** platform populer yang menyediakan berbagai dataset untuk keperluan penelitian data science dan machine learning. Dataset yang digunakan berjudul “Students Performance in Exams”, dan dapat diakses secara publik melalui tautan: <https://www.kaggle.com/datasets/spscientist/students-performance-in-exams>. Dataset ini terdiri dari **10.000 entri** data mahasiswa dengan beragam atribut multidimensi yang mencerminkan karakteristik demografis, kebiasaan belajar, dan hasil akademik. Beberapa atribut yang tersedia antara lain: jenis kelamin, usia, jam belajar per minggu, gaya belajar *(Preferred Learning Style),* tingkat penyelesaian tugas, nilai ujian, kehadiran, penggunaan teknologi pendidikan, tingkat stres, penggunaan media sosial, jam tidur, serta nilai akhir *(Final Grade).*

Dalam penelitian ini, peneliti memfokuskan analisis pada dua atribut utama, yaitu ***Preferred Learning Style*** dan ***Final Grade.*** Kedua atribut ini dipilih karena memiliki relevansi tinggi dengan tujuan penelitian, yaitu untuk mengidentifikasi klaster mahasiswa berdasarkan kecenderungan belajar dan performa akademik mereka. Atribut gaya belajar digunakan untuk memahami preferensi masing-masing individu dalam menyerap informasi, sedangkan nilai akhir menjadi indikator konkret dari keberhasilan pembelajaran.

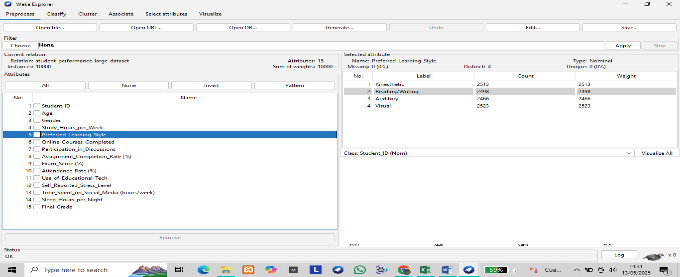
Proses pengolahan dan analisis data dilakukan menggunakan perangkat lunak ***Weka*** (Waikato Environment for Knowledge Analysis), sebuah tool open source yang banyak digunakan dalam implementasi algoritma machine learning. Weka menyediakan antarmuka grafis yang ramah pengguna serta mendukung berbagai algoritma, ***termasuk K-Means,*** yang digunakan untuk proses klasterisasi dalam penelitian ini. Proses klasterisasi dilakukan dengan menentukan jumlah klaster yang diinginkan, kemudian sistem akan mengelompokkan data mahasiswa berdasarkan kemiripan atribut.

Untuk mengevaluasi model klasterisasi yang dihasilkan, digunakan metode validasi ***percentage split*** sebesar **50%.** Artinya, dataset dibagi menjadi dua bagian: 50% untuk pelatihan model *(training set)* dan 50% lainnya untuk pengujian (*testing set).* Metode ini digunakan untuk menilai seberapa baik model dapat menggeneralisasi terhadap data baru yang belum pernah dilihat sebelumnya. Validasi ini penting untuk memastikan bahwa model tidak hanya bekerja baik pada data latih, tetapi juga memiliki performa yang konsisten pada data uji.

Hasil klasterisasi divisualisasikan dalam bentuk grafik melalui antarmuka Weka, yang membantu dalam menginterpretasi pola-pola yang muncul dari masing-masing kelompok. Melalui visualisasi ini, peneliti dapat mengidentifikasi karakteristik khas tiap klaster, seperti kelompok mahasiswa dengan gaya belajar visual dan performa akademik tinggi, atau sebaliknya. Temuan-temuan ini diharapkan dapat memberikan dasar yang kuat untuk menyusun strategi pembelajaran adaptif yang disesuaikan dengan profil masing-masing kelompok mahasiswa.

Dengan demikian, penelitian ini tidak hanya memberikan kontribusi terhadap pengembangan sistem pembelajaran yang lebih responsif dan berbasis *AI,* tetapi juga memperkuat pentingnya pendekatan *data-driven* dalam pengambilan keputusan pendidikan. Strategi yang dibangun berdasarkan hasil klasterisasi ini berpotensi meningkatkan efektivitas pembelajaran, memaksimalkan potensi mahasiswa, serta menciptakan lingkungan belajar yang lebih inklusif dan adaptif terhadap keragaman gaya belajar.

**HASIL DAN PEMBAHASAN**

**

**Gambar 1.** Visualisasi atribut Preferred\_Learning\_Style pada dataset mahasiswa menggunakan Weka

Berdasarkan **Gambar 1**, yang merupakan hasil visualisasi analisis atribut *Preferred\_Learning\_Style* menggunakan perangkat lunak *Weka*, dapat diketahui bahwa mahasiswa dalam dataset ini terbagi ke dalam empat kategori utama gaya belajar, yaitu: *Kinesthetic, Reading/Writing, Auditory*, dan *Visual.* Dari total 10.000 entri data yang dianalisis, distribusi masing-masing gaya belajar adalah sebagai berikut:

* *Kinesthetic*: 2.513 mahasiswa
* *Reading/Writing:* 2.498 mahasiswa
* *Auditory:* 2.466 mahasiswa
* *Visual:* 2.523 mahasiswa

Distribusi ini menunjukkan komposisi yang relatif seimbang antar gaya belajar, di mana tidak terdapat dominasi yang mencolok dari satu gaya belajar terhadap yang lainnya. Selisih jumlah mahasiswa di tiap kategori sangat kecil, hanya berkisar antara 10 hingga 50 individu, sehingga dapat dikatakan bahwa preferensi belajar mahasiswa tersebar secara proporsional di seluruh kategori.

Temuan ini penting karena mencerminkan keragaman preferensi belajar individu dalam populasi mahasiswa yang dianalisis. Keberagaman ini menegaskan bahwa tidak ada pendekatan tunggal yang dapat memenuhi kebutuhan semua mahasiswa, sehingga sangat diperlukan strategi pembelajaran yang bersifat fleksibel dan mampu menyesuaikan dengan berbagai gaya belajar yang ada. Dalam konteks pendidikan tinggi, hal ini menjadi tantangan sekaligus peluang bagi institusi untuk mengembangkan sistem pembelajaran yang bersifat personal dan adaptif.

Menurut Gilakjani (2012), gaya belajar merupakan karakteristik individu yang menunjukkan cara paling efektif bagi seseorang dalam menerima, mengolah, dan menyimpan informasi. Gaya belajar ini tidak hanya bersifat kognitif, tetapi juga afektif dan psikologis, serta memengaruhi tingkat keterlibatan mahasiswa dalam proses pembelajaran. Secara umum:

* Mahasiswa dengan gaya belajar *Visual* cenderung menyukai informasi dalam bentuk gambar, diagram, dan visualisasi data.
* Mereka yang memiliki gaya belajar *Auditory* lebih efektif ketika belajar melalui diskusi, ceramah, atau media audio.
* Gaya belajar *Reading/Writing* menunjukkan preferensi terhadap materi dalam bentuk teks tertulis, baik berupa catatan maupun buku.
* Sedangkan gaya belajar *Kinesthetic* cenderung mengutamakan pengalaman langsung, praktik lapangan, dan pendekatan berbasis gerakan atau simulasi.

Distribusi gaya belajar yang berimbang seperti ini memberikan dasar yang kuat untuk mengembangkan model pembelajaran *hybrid*, yang memadukan berbagai media dan metode pengajaran agar dapat menjangkau seluruh *spektrum preferensi* belajar mahasiswa. Hal ini sejalan dengan gagasan pendidikan *diferensial*, yaitu pendekatan yang menyesuaikan metode pengajaran berdasarkan kebutuhan, kemampuan, dan karakteristik unik peserta didik.

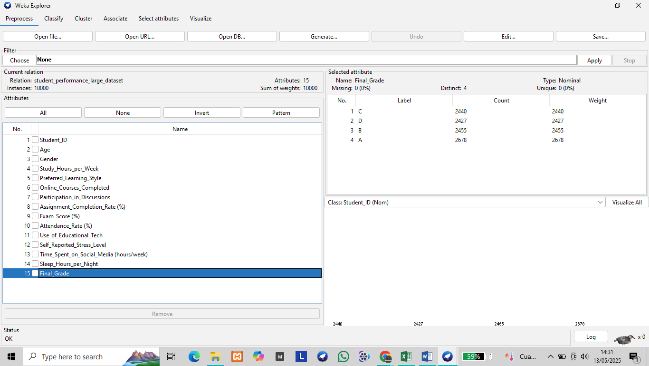
Lebih lanjut, penelitian-penelitian sebelumnya, seperti yang diulas oleh Pashler et al. (2009), menunjukkan bahwa meskipun konsep gaya belajar banyak digunakan, efektivitas penyesuaian metode pengajaran dengan gaya belajar belum sepenuhnya didukung oleh bukti empiris yang kuat. Kritik utama dari pendekatan ini terletak pada risiko simplifikasi berlebihan terhadap kompleksitas proses belajar. Artinya, meskipun penting untuk mengenali gaya belajar, pengajaran yang efektif tetap harus mempertimbangkan faktor lain seperti motivasi, konteks materi, dan keterlibatan aktif siswa dalam proses pembelajaran.

Namun demikian, data tentang preferensi gaya belajar tetap sangat berguna, terutama dalam pengembangan sistem pembelajaran berbasis teknologi yang bersifat adaptif. Dengan bantuan *machine learning*, data seperti ini dapat digunakan untuk:

* Mengelompokkan mahasiswa ke dalam klaster dengan karakteristik serupa *(clustering).*
* Memprediksi gaya belajar mahasiswa baru berdasarkan pola yang telah terbentuk.
* Menyesuaikan materi ajar atau metode evaluasi secara otomatis sesuai dengan klaster yang terbentuk.

Dalam konteks penelitian ini, analisis awal terhadap distribusi gaya belajar memberikan landasan *konseptual* dan *empiris* bagi proses klasterisasi menggunakan algoritma K-Means. Klaster yang terbentuk nantinya diharapkan dapat membantu dalam merancang strategi pembelajaran personal yang lebih efektif, baik secara manual oleh pengajar, maupun secara otomatis melalui sistem pembelajaran cerdas (intelligent learning system).

Dengan demikian, analisis distribusi gaya belajar tidak hanya memberikan informasi statistik semata, tetapi juga kontribusi *strategis* bagi pengembangan pendidikan yang lebih inklusif, adaptif, dan berbasis data. Data ini dapat menjadi referensi penting bagi lembaga pendidikan dalam merumuskan kebijakan, merancang kurikulum, hingga memilih teknologi pembelajaran yang tepat sesuai dengan kebutuhan mahasiswanya.

****

**Gambar 2.** Visualisasi atribut Final\_Grade pada dataset mahasiswa menggunakan Weka

Selanjutnya, sebagaimana ditunjukkan pada **Gambar 2**, atribut *Final\_Grade* (nilai akhir) dari seluruh 10.000 mahasiswa dalam dataset dianalisis guna mengetahui distribusi performa akademik mahasiswa berdasarkan kategori nilai. Atribut ini telah diklasifikasikan ke dalam empat kategori utama, yaitu A, B, C, dan D, yang masing-masing merepresentasikan tingkat pencapaian akademik dari yang tertinggi hingga terendah. Distribusi dari keempat kategori nilai tersebut adalah sebagai berikut:

* Nilai A: 2.678 mahasiswa
* Nilai B: 2.455 mahasiswa
* Nilai C: 2.440 mahasiswa
* Nilai D: 2.427 mahasiswa

Data ini menunjukkan bahwa distribusi nilai akhir bersifat relatif merata, dengan rentang selisih jumlah mahasiswa di tiap kategori tidak terlalu signifikan. Meskipun demikian, terdapat kecenderungan bahwa mayoritas mahasiswa mampu mencapai nilai akademik yang tergolong tinggi, ditunjukkan oleh dominasi kategori A sebagai kelompok dengan jumlah terbesar. Secara statistik, lebih dari 26% dari total populasi berhasil memperoleh nilai A, suatu pencapaian yang dapat dikatakan sangat positif dalam konteks performa akademik.

Distribusi nilai ini mencerminkan bahwa secara umum, tingkat akademik mahasiswa dalam dataset berada dalam kategori baik hingga sangat baik. Fakta ini membuka kemungkinan adanya faktor-faktor pendukung yang turut memengaruhi capaian akademik mereka. Faktor tersebut bisa bersifat internal, seperti tingkat motivasi, kemampuan manajemen waktu, gaya belajar yang sesuai, serta kebiasaan belajar yang konsisten. Di sisi lain, faktor eksternal juga bisa berperan penting, seperti dukungan lingkungan belajar yang kondusif, akses terhadap sumber belajar digital, kualitas pengajaran, hingga intervensi teknologi dalam proses pendidikan.

Namun demikian, distribusi yang tampak proporsional ini belum secara langsung menjelaskan apakah ada hubungan antara gaya belajar danperforma akademik. Dengan kata lain, walaupun kategori nilai tampak seimbang dan dominasi nilai A cukup mencolok, tidak serta-merta dapat disimpulkan bahwa gaya belajar tertentu berkontribusi langsung terhadap capaian tersebut. Oleh karena itu, dibutuhkan analisis lanjutan yang lebih mendalam dan sistematis, seperti:

* *Analisis asosiasi* untuk mengukur kekuatan hubungan antara gaya belajar dan nilai akhir.
* *Klasifikasi supervised learning* untuk memprediksi nilai akhir berdasarkan atribut gaya belajar dan variabel lainnya.
* *Clustering* untuk mengidentifikasi kelompok mahasiswa dengan kecenderungan performa akademik yang serupa, yang kemudian dapat dianalisis lebih lanjut secara deskriptif.

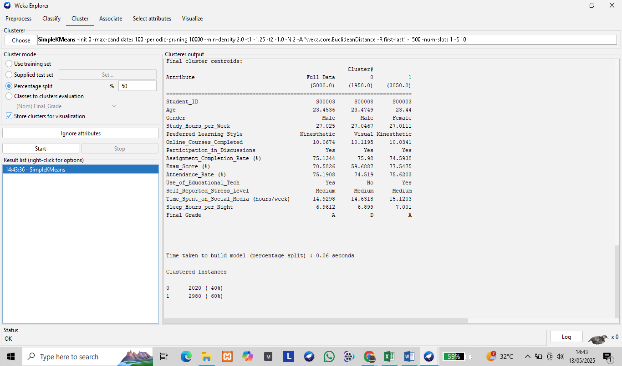
Salah satu pendekatan yang digunakan dalam penelitian ini adalah algoritma *K-Means*, yang memungkinkan pengelompokan mahasiswa berdasarkan kemiripan atribut gaya belajar (*Preferred\_Learning\_Style*) dan nilai akhir (*Final\_Grade*). Proses klasterisasi ini tidak hanya bertujuan untuk menemukan struktur atau pola tersembunyi dalam data, tetapi juga untuk membentuk dasar bagi pengembangan sistem pembelajaran adaptif yang mampu menyajikan strategi belajar berdasarkan kelompok karakteristik yang serupa.

Hal ini sejalan dengan pandangan dari Dunn dan Dunn (1992), yang menyatakan bahwa ketika strategi pengajaran disesuaikan dengan gaya belajar dominan peserta didik, maka proses pembelajaran cenderung menjadi lebih efektif dan hasil akademik akan meningkat. Mahasiswa yang menerima pendekatan belajar sesuai dengan preferensinya cenderung menunjukkan tingkat pemahaman yang lebih baik, antusiasme yang lebih tinggi, serta keterlibatan yang lebih aktif dalam kegiatan pembelajaran. Penyesuaian ini mencakup tidak hanya metode penyampaian materi, tetapi juga bentuk penugasan, aktivitas praktikum, hingga pendekatan evaluasi.

Selain itu, dalam era pendidikan berbasis teknologi dan data *(data-driven education*), distribusi nilai akhir seperti ini dapat menjadi indikator kinerja sistem pembelajaran yang ada. Jika digabungkan dengan data gaya belajar dan variabel pendukung lainnya, maka informasi ini dapat digunakan sebagai landasan dalam pengembangan sistem rekomendasi pembelajaran individual (*personalized learning systems*). Sistem seperti ini memungkinkan:

* Penyusunan materi ajar berbasis profil klaster mahasiswa.
* Rekomendasi metode pembelajaran (misalnya video interaktif untuk visual learner, forum diskusi untuk auditory learner).
* Intervensi dini untuk mahasiswa yang menunjukkan performa rendah, dengan memberikan bimbingan tambahan yang sesuai gaya belajarnya.
* Penyesuaian waktu dan jenis evaluasi akademik, agar lebih sesuai dengan karakteristik belajar individu.

Dengan demikian, analisis terhadap atribut *Final\_Grade* tidak hanya menggambarkan capaian akademik mahasiswa secara umum, tetapi juga memiliki nilai strategis dalam perancangan pendekatan pedagogis yang lebih *adaptif, inklusif,* dan berbasis bukti empiris. Hasil ini memperkuat urgensi integrasi antara *data analytics,* *machine learning,* dan *pedagogi* modern dalam upaya peningkatan mutu pendidikan di era digital.



**Gambar 3.** Hasil Klasifikasi menggunakan *cluster* dengan logika *K-means*

Proses klasterisasi dalam penelitian ini dilakukan dengan menerapkan algoritma *K-Means* yang diimplementasikan menggunakan perangkat lunak *Weka.* Dataset yang digunakan mencakup sebanyak 10.000 entri data mahasiswa, dengan berbagai atribut yang relevan terhadap performa akademik, seperti *Exam Score*, *Assignment Completion Rate*, *Final Grade*, *Study Hours per Week*, dan *Preferred Learning Style*. Untuk mengevaluasi performa model serta menghindari overfitting, digunakan teknik validasi *percentage split* **s**ebesar 50**%**, yaitu membagi data menjadi 50% untuk pelatihan (*training set*) dan 50% sisanya untuk pengujian (*testing set*). Pendekatan ini dipilih karena mampu mengukur kemampuan model dalam mengenali pola pada data baru yang belum pernah digunakan dalam proses pelatihan, sekaligus menilai tingkat generalisasi dari hasil klasterisasi yang dihasilkan.

Hasil dari proses klasterisasi menunjukkan bahwa mahasiswa terbagi ke dalam dua klaster utama, masing-masing terdiri dari sekitar 2.980 mahasiswa *(Cluster 1)* dan 2.020 mahasiswa *(Cluster 0).* Visualisasi hasil klasterisasi disajikan pada Gambar **3**, yang memperlihatkan adanya pemisahan yang cukup jelas antara kedua kelompok berdasarkan atribut utama yang digunakan dalam pemodelan. Pemisahan ini memberikan indikasi awal bahwa terdapat pola tertentu dalam data yang mampu mengelompokkan mahasiswa berdasarkan kesamaan karakteristik akademik dan gaya belajar mereka.

Secara rinci, *Cluster 1* menunjukkan karakteristik mahasiswa dengan performa akademik tinggi, ditandai dengan nilai rata-rata *Exam Score* sebesar 77,5%, *Assignment Completion Rate* sebesar 74,5%, serta dominasi nilai akhir (*Final Grade*) pada kategori A. Sebaliknya, Cluster 0 mencerminkan kelompok mahasiswa dengan performa akademik lebih rendah, dengan nilai rata-rata *Exam Score* sebesar 59,6%, *Assignment Completion Rate* sebesar 75,1%, dan dominasi *Final Grade* pada kategori D. Perlu dicatat bahwa tingkat penyelesaian tugas pada kedua klaster relatif sebanding, namun perbedaan mencolok terlihat pada capaian nilai ujian dan nilai akhir secara keseluruhan. Hal ini mengindikasikan bahwa penyelesaian tugas semata belum cukup menjamin pencapaian akademik yang optimal, dan bahwa faktor lain turut memengaruhi perbedaan tersebut.

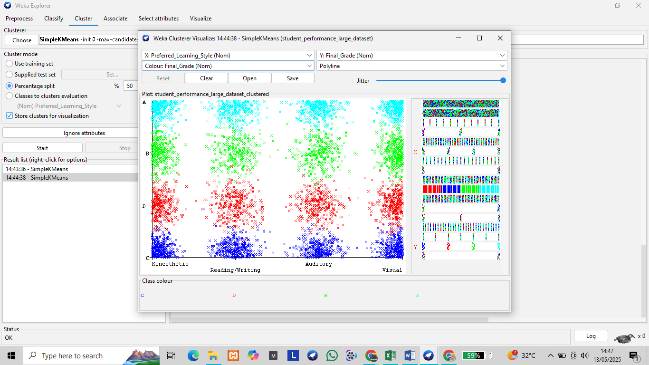
Salah satu temuan penting dalam analisis ini adalah bahwa jumlah jam belajar per minggu *(Study Hours per Week)* antara kedua klaster tidak menunjukkan perbedaan yang signifikan. Dengan demikian, durasi belajar bukan merupakan satu-satunya variabel yang menentukan keberhasilan akademik. Temuan ini mendukung pandangan bahwa efektivitas strategi belajar, kualitas fokus, serta metode yang digunakan dalam belajar memiliki pengaruh yang lebih besar terhadap hasil akhir dibandingkan sekadar lamanya waktu yang dihabiskan untuk belajar. Hal ini menunjukkan pentingnya peran gaya belajar dan pendekatan pedagogis dalam mendukung keberhasilan akademik.

Terkait atribut *Preferred Learning Style,* hasil klasterisasi menunjukkan bahwa *Cluster 1* didominasi oleh mahasiswa dengan gaya belajar Visual dan Kinestetik, sedangkan *Cluster 0* lebih banyak terdiri dari mahasiswa dengan gaya belajar Kinestetik dan *Reading/Writing.* Meskipun gaya belajar *Kinestetik* muncul pada kedua *klaster,* kombinasi dengan gaya *Visual* tampaknya memberikan kontribusi yang lebih positif terhadap performa akademik. Ini mengindikasikan bahwa penggunaan media visual, pendekatan pembelajaran berbasis pengamatan, dan representasi grafis informasi dapat membantu memperkuat pemahaman dan hasil belajar mahasiswa secara signifikan.

Temuan ini sejalan dengan pandangan dari Gilakjani (2012) dan Dunn & Dunn (1992), yang menyatakan bahwa penyesuaian strategi pengajaran dengan gaya belajar siswa dapat memberikan dampak signifikan terhadap hasil belajar. Ketika metode pembelajaran sesuai dengan cara siswa merasa nyaman dalam menyerap informasi, maka efektivitas pembelajaran cenderung meningkat. Namun demikian, penting untuk dicatat bahwa gaya belajar bukan satu-satunya determinan keberhasilan akademik. Kombinasi dengan dukungan lingkungan belajar, motivasi intrinsik, kualitas pengajaran, dan penggunaan teknologi pendidikan yang tepat juga memainkan peran yang sangat penting.

Dengan memahami komposisi klaster berdasarkan kombinasi antara gaya belajar dan performa akademik, institusi pendidikan dapat merancang berbagai intervensi pembelajaran yang lebih terarah dan berbasis data. Misalnya, mahasiswa dalam klaster performa tinggi dapat dijadikan acuan dalam menyusun pendekatan pembelajaran berbasis best practices, sementara mahasiswa dalam klaster performa rendah dapat diberikan bimbingan khusus, pemetaan ke kelas dengan metode pengajaran yang lebih sesuai, atau akses ke sumber belajar yang lebih interaktif. Hasil klasterisasi ini juga dapat digunakan sebagai dasar dalam pengembangan sistem rekomendasi pembelajaran adaptif (adaptive learning systems) yang menyesuaikan materi, metode, dan evaluasi berdasarkan profil belajar mahasiswa secara individual.

Secara keseluruhan, implementasi algoritma *K-Means* dalam penelitian ini memberikan kontribusi penting dalam mengidentifikasi pola tersembunyi dalam data pendidikan, serta memberikan gambaran awal tentang bagaimana pendekatan machine learning dapat diterapkan untuk mendukung pengambilan keputusan berbasis data dalam konteks pendidikan tinggi. Analisis ini tidak hanya memberikan pemahaman yang lebih baik mengenai keberagaman karakteristik mahasiswa, tetapi juga membuka peluang untuk merancang strategi pembelajaran yang lebih personal, efektif, dan inklusif.

**

**Gambar 4.** Hasil Visualisi Klasifikasi menggunakan cluster dengan logika *K-means*

Visualisasi yang ditampilkan pada Gambar **4** memberikan wawasan yang lebih mendalam mengenai hubungan antara *Preferred Learning Style* (gaya belajar) dan *Final Grade* (nilai akhir) mahasiswa. Distribusi yang tergambar menunjukkan adanya pola yang cukup konsisten, di mana mahasiswa dengan gaya belajar *Visual* dan *Auditory* secara proporsional lebih banyak memperoleh nilai A dan B, dibandingkan dengan mahasiswa yang memiliki gaya belajar *Kinestetik* dan *Reading/Writing,* yang cenderung tersebar pada kategori nilai C dan D. Pola ini memberikan indikasi awal bahwa kesesuaian antara gaya belajar dengan pendekatan pengajaran yang digunakan berkontribusi terhadap pencapaian akademik mahasiswa.

Secara khusus, mahasiswa dengan gaya belajar Visual tampak lebih mampu beradaptasi dengan lingkungan pembelajaran berbasis digital yang saat ini menjadi ciri utama sistem pendidikan tinggi modern. Konten pembelajaran berbentuk grafik, diagram, ilustrasi visual, hingga animasi interaktif yang umum ditemukan pada platform e-learning merupakan media yang sangat mendukung cara belajar visual. Temuan ini sejalan dengan pernyataan Gilakjani (2012) yang menyebutkan bahwa siswa dengan preferensi belajar visual cenderung mengalami peningkatan hasil akademik apabila materi disampaikan melalui media pembelajaran *visual* yang sesuai dengan cara mereka memproses informasi.

Demikian pula, gaya belajar *Auditory* menunjukkan korelasi positif terhadap capaian nilai tinggi, terutama pada mahasiswa yang terlibat dalam proses pembelajaran berbasis diskusi aktif, presentasi verbal, forum daring, maupun media audio seperti podcast dan rekaman kuliah. Hal ini menunjukkan bahwa pembelajaran yang mengakomodasi interaksi lisan dan dialog dua arah juga mampu mendukung pencapaian akademik yang lebih baik.

Namun demikian, penting untuk dicatat bahwa gaya belajar bukanlah satu-satunya faktor yang memengaruhi performa akademik. Hasil klasterisasi sebelumnya menunjukkan bahwa sejumlah faktor perilaku lainnya seperti tingkat stres *(stress level),* durasi tidur *(sleep hours),* dan intensitas penggunaan media sosial *(social media usage)* juga turut memberikan kontribusi terhadap perbedaan hasil akademik antar mahasiswa. Mahasiswa yang tergolong dalam klaster performa tinggi umumnya menunjukkan pola tidur yang lebih konsisten, tingkat stres yang moderat, dan penggunaan media sosial yang seimbang, yang secara tidak langsung berkontribusi terhadap kesiapan kognitif dan afektif dalam menghadapi proses pembelajaran.

Temuan ini menggaris bawahi pentingnya pendekatan holistik dalam meningkatkan performa akademik mahasiswa. Tidak hanya aspek *kognitif* seperti gaya belajar yang perlu diperhatikan, tetapi juga aspek psikologis, sosial, dan gaya hidup yang membentuk konteks belajar mahasiswa secara menyeluruh. Dengan kata lain, pengembangan strategi pembelajaran yang efektif tidak hanya mengandalkan pemetaan gaya belajar, tetapi juga perlu mempertimbangkan keseimbangan antara beban akademik, kondisi emosional, serta kebiasaan hidup sehat.

Dari perspektif *pedagogis* dan *institusional,* hasil visualisasi ini memiliki implikasi yang signifikan, terutama dalam konteks pembelajaran daring (*online learning*) dan *blended learning* yang semakin banyak diterapkan di lingkungan pendidikan tinggi di Indonesia. Pemahaman terhadap karakteristik gaya belajar dan perilaku mahasiswa dapat menjadi landasan dalam menyusun kurikulum yang lebih adaptif, mengembangkan materi ajar berbasis media majemuk, serta menyediakan dukungan pembelajaran dan *psikososial* yang bersifat personal. Layanan bimbingan akademik, pelatihan manajemen stres, dan dukungan emosional berbasis teknologi menjadi elemen penting dalam ekosistem pendidikan tinggi yang responsif terhadap kebutuhan mahasiswa masa kini.

Dengan demikian, visualisasi pada Gambar **4** tidak hanya menyajikan informasi deskriptif, tetapi juga menegaskan adanya korelasi bermakna antara preferensi kognitif, pola perilaku belajar, dan capaian akademik. Temuan ini dapat menjadi dasar yang kuat dalam menyusun kebijakan pendidikan berbasis data *(data-driven education policy),* yang mendukung terciptanya proses pembelajaran yang lebih inklusif, adaptif, dan efektif dalam mencapai tujuan pembelajaran jangka panjang di era digital.

**KESIMPULAN**

Penelitian ini bertujuan untuk mengidentifikasi pola gaya belajar mahasiswa serta keterkaitannya dengan performa akademik melalui penerapan algoritma *K-Means* dengan bantuan perangkat lunak *Weka.* Berdasarkan hasil analisis terhadap 10.000 data mahasiswa, tujuan utama penelitian yakni mengelompokkan mahasiswa berdasarkan karakteristik gaya belajar dan capaian nilai akhir telah tercapai dengan baik. Proses klasterisasi menghasilkan beberapa kelompok mahasiswa yang memiliki karakteristik serupa, baik dari segi gaya belajar dominan maupun tingkat performa akademik.

Secara umum, hasil klasterisasi menunjukkan adanya hubungan yang signifikan antara gaya belajar tertentu dengan capaian akademik mahasiswa. Mahasiswa dengan gaya belajar *Visual* dan *Auditory* menunjukkan kecenderungan memperoleh nilai akhir yang lebih tinggi dibandingkan dengan mereka yang memiliki gaya belajar Kinestetik dan *Reading/Writing*. Temuan ini memperkuat pandangan bahwa kesesuaian antara gaya belajar dengan pendekatan pengajaran yang digunakan dalam pembelajaran dapat meningkatkan efektivitas proses belajar dan hasil akhir akademik.

Dalam konteks pendidikan modern yang semakin mengandalkan teknologi, keberhasilan mahasiswa dengan gaya belajar *Visual* dan *Auditory* dapat dijelaskan melalui kecocokan antara gaya belajar mereka dengan media pembelajaran digital seperti video pembelajaran, simulasi interaktif, podcast, dan diskusi daring. Dengan kata lain, pemanfaatan teknologi tidak hanya berperan sebagai alat bantu belajar, tetapi juga sebagai medium yang dapat dioptimalkan untuk menyesuaikan pendekatan pembelajaran dengan kebutuhan kognitif mahasiswa.

Lebih jauh, hasil penelitian juga mengungkap bahwa faktor perilaku lainnya, seperti tingkat stres, durasi tidur, dan penggunaan media sosial, turut memengaruhi performa akademik. Mahasiswa dalam klaster dengan capaian akademik tinggi umumnya menunjukkan pola hidup yang lebih sehat dan teratur, termasuk jam tidur yang cukup, tingkat stres yang terkendali, serta penggunaan media sosial secara bijak. Hal ini menegaskan bahwa performa akademik yang optimal tidak hanya ditentukan oleh kecocokan gaya belajar, tetapi juga oleh keseimbangan dalam aspek psikologis dan gaya hidup mahasiswa.

Dengan demikian, hasil penelitian ini menekankan pentingnya pendekatan pembelajaran yang adaptif, personal, dan berbasis data. Informasi yang diperoleh dari proses klasterisasi dapat dimanfaatkan oleh dosen, perancang kurikulum, pengembang e-learning, hingga institusi pendidikan tinggi sebagai dasar untuk merancang strategi pembelajaran yang lebih relevan dan efektif. Pendekatan semacam ini sangat penting dalam menghadapi tantangan pembelajaran di era digital, di mana kebutuhan dan karakteristik belajar mahasiswa sangat beragam. Penelitian ini membuka berbagai peluang pengembangan di masa mendatang. Salah satu prospek penting adalah pengembangan sistem pembelajaran cerdas berbasis machine learning, yang mampu secara otomatis menyesuaikan konten, media, dan strategi pembelajaran sesuai dengan profil belajar individual mahasiswa. Konsep ini sejalan dengan visi menuju pendidikan yang *truly personalized*, di mana setiap mahasiswa mendapatkan pengalaman belajar yang unik, efektif, dan sesuai dengan preferensi serta kebutuhannya.

Untuk meningkatkan kualitas analisis dan generalisasi hasil, penelitian lanjutan disarankan untuk menggunakan dataset yang lebih kompleks dan kontekstual, misalnya yang mencakup data interaksi dalam platform pembelajaran digital, log aktivitas belajar, tingkat motivasi, hingga *self-efficacy* mahasiswa. Dengan pendekatan ini, dimungkinkan untuk memperoleh wawasan yang lebih holistik mengenai faktor-faktor yang memengaruhi keberhasilan belajar mahasiswa.

Selain algoritma *K-Means*, penerapan algoritma klasterisasi lain seperti *DBSCAN* atau *Hierarchical Clustering* juga layak untuk dieksplorasi*. DBSCAN*, misalnya, cocok untuk mendeteksi outlier dan pola non-linier, sedangkan Hierarchical Clustering dapat memberikan gambaran bertingkat mengenai struktur data yang lebih kompleks. Penggunaan algoritma yang bervariasi ini dapat memperkaya perspektif dalam segmentasi data mahasiswa.

Di samping itu, arah pengembangan berikutnya juga dapat difokuskan pada pembuatan sistem rekomendasi pembelajaran berbasis *AI,* yang mampu memberikan saran mengenai materi, metode, dan waktu belajar yang optimal. Dengan mengintegrasikan hasil-hasil klasterisasi ke dalam sistem pembelajaran digital, diharapkan akan tercipta ekosistem pendidikan tinggi yang inklusif, adaptif, dan berorientasi pada pembelajaran berbasis data, sesuai dengan semangat *Revolusi Industri 4.0 dan Society 5.0.*

**REFERENSI**

Hairani, H. (2022). Pelatihan Implementasi Machine Learning pada Bidang Pendidikan. *ADMA: Jurnal Pengabdian dan Pemberdayaan Masyarakat*, *2*(2), 305-310.Han, J., & Kamber, M. (2006). *Data Mining: Concepts and Techniques*. *Soft Computing* (Vol. 54). https://doi.org/10.1007/978-3-642-19721-5

Saputra, R. F., & Pranoto, Y. A. (2021). IMPLEMENTASI METODE K-MEANS CLUSTERING PADA TES PSIKOLOGI UNTUK MENENTUKAN KELOMPOK BELAJAR SISWA BERBASIS MOBILE. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *5*(1), 328-333.Liza, Y., & Yupinti. (2012). Sistem Informasi persediaan barang pada PT.Surya Nusa Bhaktindo Bengkulu. *Media Infotama*, *8*(1), 90–117.

Sulistiyawati, A., & Supriyanto, E. (2021). Implementasi Algoritma K-means Clustring dalam Penetuan Siswa Kelas Unggulan. *J. Tekno Kompak*, *15*(2), 25.

Pratama, Y. A., Hariyanti, U., & Suharsono, A. (2024). Identifikasi Karakteristik Peserta Didik dalam Mata Pelajaran Informatika menggunakan Algoritma K-Means berdasarkan Hasil Belajar: Studi Kasus pada Kelas X TKJ SMKN 3 Malang. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, *8*(9).

Mustofa, M. (2019). Penerapan Algoritma K-Means Clustering pada Karakter Permainan Multiplayer Online Battle Arena. *Jurnal Informatika*, *6*(2), 246-254

Ananto, D. T., Mahardewantoro, D. D., Mustafa, F., Ardianto, M. G., Rafi, M. M., Zein, R. A., ... & Adharani, Y. (2023, October). Edukasi dan pelatihan pengenalan machine learning dan computer vision untuk mengeksplorasi potensi visual. In *Prosiding Seminar Nasional Pengabdian Masyarakat LPPM UMJ* (Vol. 1, No. 1).Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological science in the public interest*, *9*(3), 105-119.

Retnoningsih, E., & Pramudita, R. (2020). Mengenal machine learning dengan teknik supervised dan unsupervised learning menggunakan python. *Bina Insani Ict Journal*, *7*(2), 156-165.

Kuantitatif, P. P. (2016). Metode Penelitian Kunatitatif Kualitatif dan R&D. *Alfabeta, Bandung*.

Sulastri, H., Mubarok, H., & Iasha, S. S. (2021). Implementasi Algoritma Machine Learning Untuk Penentuan Cluster Status Gizi Balita. *Jurnal Rekayasa Teknologi Informasi (JURTI)*, *5*(2), 184-191.

Kaggle. (n.d.). Students performance in exams. Retrieved from <https://www.kaggle.com/datasets/spscientist/students-performance-in-exams>