VOL. 3. NO. 03 MEI 2026

JOURNAL OF COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE (JCSAI)

PENERAPAN ALGORITMA X-MEANS UNTUK MENINGKATKAN MODEL PENGELOMPOKAN PRESTASI SISWA BERDASARKAN NILAI AKADEMIK DI MAN 2 KOTA CIREBON

Bertio Fajar Pratama[1], Ade Irma Purnamasari[2], Denni Pratama[3], Fatihanursari Dikananda[4]

Program Studi Teknik Informatika[124]
Program Studi Komputerisasi Akuntansi[3]

STMIK IKMI Cirebon
https://ikmi.ac.id/page/18/?lang=de
fajarfaj89@gmail.com

(*) Corresponding Author : fajarfaj89@gmail.com
Published : 30 Mei 2026

Abstract—*This study aims to apply the X-MEANS algorithm for Clustering student achievement based on academic scores at MAN 2 Kota Cirebon. A key challenge in academic Clustering is the need to determine the number of clusters manually, as in the K-Means algorithm, which can lead to unstable groupings and dependence on researcher subjectivity. To address this issue, the present study employs the X-MEANS algorithm, which is able to automatically adjust the number of clusters according to the underlying data structure. The dataset consists of academic scores from 263 students across eight subjects, which first undergo preprocessing steps including data cleaning, attribute selection, encoding, and normalization to meet the requirements of the modeling process. The Clustering procedure is carried out using X-MEANS in Jupyter Notebook, while the quality of the resulting clusters is evaluated using the Davies–Bouldin Index (DBI). The findings show that X-MEANS automatically produces five optimal clusters with a DBI value of 0.6294, indicating that the clusters formed are sufficiently compact and well separated. Based on centroid analysis and cluster mapping, these five clusters are further grouped into three achievement categories, namely 84 high-achieving students, 125 moderately achieving students, and 54 students requiring academic support, thus providing a more representative, measurable, and interpretable picture of students' academic profiles for the school. These results suggest that X-MEANS is adaptive and relevant in the context of Educational Data Mining (EDM), and can be utilized as a basis for more targeted academic decision-making, such as planning remedial programs, enrichment activities, and continuous monitoring of student achievement at the secondary school level.*

Keywords: *X-MEANS, Clustering, Educational Data mining, Davies-Bouldin Index, Student Performance.*

Abstrak—Penelitian ini bertujuan untuk menerapkan algoritma *X-MEANS* dalam pengelompokan prestasi siswa berdasarkan nilai akademik di MAN 2 Kota Cirebon. Permasalahan utama dalam klastering akademik adalah kebutuhan menentukan jumlah klaster secara manual sebagaimana pada algoritma *K-Means*, yang berpotensi menghasilkan pengelompokan kurang stabil dan bergantung pada subjektivitas peneliti. Untuk mengatasi hal tersebut, penelitian ini menggunakan algoritma *X-MEANS* yang mampu menyesuaikan jumlah klaster secara otomatis berdasarkan struktur data. Dataset yang digunakan terdiri dari nilai akademik 263 siswa pada delapan mata pelajaran, yang terlebih dahulu melalui tahapan pra-pemrosesan berupa pembersihan data, seleksi atribut, pengkodean, dan normalisasi agar sesuai dengan kebutuhan pemodelan. Proses klastering dilakukan menggunakan *X-MEANS* di *Jupyter Notebook*, sedangkan kualitas hasil pengelompokan dievaluasi menggunakan *Davies–Bouldin Index (DBI)*. Hasil penelitian menunjukkan bahwa algoritma *X-MEANS* secara otomatis menghasilkan lima klaster optimal dengan nilai DBI sebesar 0,6294, yang mengindikasikan bahwa klaster yang terbentuk cukup kompak dan terpisah dengan baik. Melalui analisis centroid dan pemetaan klaster, kelima klaster tersebut kemudian dikelompokkan ke dalam tiga kategori prestasi, yaitu 84 siswa berprestasi, 125 siswa cukup berprestasi, dan 54 siswa yang memerlukan penguatan akademik, sehingga memberikan gambaran pola akademik yang lebih representatif, terukur, dan mudah diinterpretasikan oleh pihak sekolah. Temuan ini menunjukkan bahwa *X-MEANS* adaptif dan relevan dalam konteks *Educational Data Mining (EDM)*, serta dapat dimanfaatkan sebagai dasar pengambilan

keputusan akademik yang lebih tepat sasaran, seperti perencanaan program remedial, pengayaan, dan pemantauan perkembangan prestasi siswa secara berkelanjutan di tingkat sekolah menengah.

Kata Kunci: *X-MEANS*, *Clustering*, *Educational Data mining*, Davies-Bouldin Index, Prestasi Siswa.

INTRODUCTION

Perkembangan teknologi informasi telah memberikan dampak besar dalam berbagai aspek kehidupan, termasuk bidang pendidikan. Pemanfaatan teknologi untuk menganalisis data akademik siswa menjadi salah satu cara efektif dalam mendukung proses pengambilan keputusan berbasis data. Salah satu bidang yang berkembang dalam hal ini adalah *Educational Data Mining (EDM)*, yaitu penerapan teknik penambangan data (*data mining*) untuk menemukan pola, hubungan, dan pengetahuan baru dari data pendidikan. Menurut Yağcı, (2022), *EDM* membantu sekolah memahami perilaku belajar siswa, memprediksi prestasi akademik, serta meningkatkan strategi pembelajaran yang lebih tepat sasaran.

Meskipun memiliki potensi besar, penerapan *EDM* menghadapi sejumlah tantangan. Yağcı, (2022) menyebutkan bahwa kualitas dan heterogenitas data akademik sering kali menghambat keakuratan *model*. Data siswa yang tidak lengkap atau tidak konsisten dapat menyebabkan hasil analisis yang bias. Selain itu, Raftopoulos et al. (2025) menjelaskan bahwa banyak algoritma pembelajaran mesin menghasilkan *model* yang sulit diinterpretasikan, sehingga guru sulit memahami dasar keputusan yang diberikan sistem. Tantangan lainnya adalah menjaga keadilan (*fairness*) dan privasi data siswa yang menjadi perhatian penting dalam dunia pendidikan modern.

Dalam konteks pengelompokan prestasi siswa (*student achievement grouping*), metode *Clustering* digunakan untuk mengelompokkan siswa berdasarkan kemiripan nilai atau kemampuan akademik. Salah satu algoritma *Clustering* yang populer adalah *K-Means*, karena sederhana dan efisien dalam memproses data besar. Namun, penelitian sebelumnya menunjukkan bahwa *K-Means* memiliki keterbatasan utama yaitu harus menentukan jumlah klaster (K) di awal dan sensitif terhadap pemilihan titik pusat (*centroid*) awal [3]. Hal ini dapat menghasilkan hasil pengelompokan yang tidak stabil dan sulit diulang ketika data berubah.

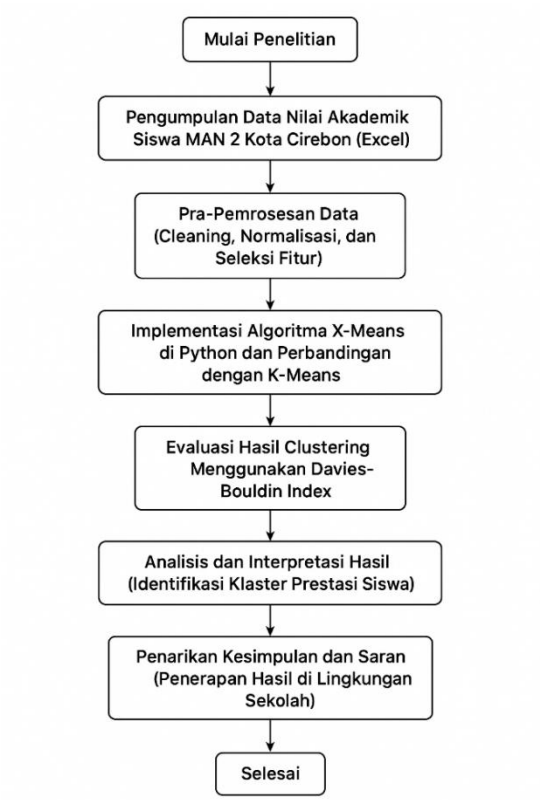
Penelitian terdahulu telah membuktikan bahwa metode *Clustering* dapat membantu mengidentifikasi kelompok prestasi akademik siswa. Izzati et al. (2023) menggunakan *K-Means* untuk membagi mahasiswa menjadi tiga kelompok: berprestasi tinggi, sedang, dan rendah. Hasil penelitian tersebut menunjukkan bahwa pendekatan pengelompokan mampu meningkatkan akurasi analisis akademik. Mohamed Nafuri et al. (2022) membandingkan beberapa metode *Clustering* seperti *K-Means, BIRCH*, dan *DBSCAN*, dan menemukan bahwa metode dengan penyesuaian parameter tertentu memberikan hasil pengelompokan yang lebih baik. Namun, metode-metode tersebut masih memerlukan penentuan jumlah klaster secara manual.

Untuk mengatasi keterbatasan *K-Means* yang memerlukan jumlah klaster ditentukan secara manual, penelitian ini menggunakan algoritma *X-MEANS* yang dapat menentukan jumlah klaster optimal secara otomatis menggunakan *Davies-Bouldin Index (DBI)* dan *Minimum Description Length (MDL)***.** Penelitian oleh Ramdani et al. (2023) menunjukkan bahwa *X-MEANS* menghasilkan nilai *DBI* lebih rendah dibanding *K-Medoids*, sehingga klaster yang terbentuk lebih akurat. Hasil serupa juga ditemukan oleh Wulandari et al. (2024), yang membuktikan bahwa *X-MEANS* mampu menghasilkan pengelompokan data yang lebih stabil dan representatif.

Di MAN 2 Kota Cirebon, pengelompokan prestasi siswa masih dilakukan secara manual berdasarkan rata-rata nilai sehingga belum menggambarkan pola akademik secara menyeluruh. Untuk itu, penelitian ini menerapkan alsgoritma *X-Mean* guna meningkatkan akurasi pengelompokan prestasi siswa [8]. Penelitian Raihanah & Wahyudi, (2024) membuktikan bahwa metode *Clustering* mampu menghasilkan pemetaan akademik yang lebih *objektif* dan membantu sekolah dalam proses pengambilan keputusan berbasis data.

MATERIALS AND METHODS

Berikut ilustrasi alur desain penelitian yang menggambarkan hubungan antar tahap eksperimen



Gambar 1. Alur Penelitian

Diagram alur penelitian tersebut menggambarkan proses yang sistematis dalam menganalisis data nilai akademik siswa untuk menemukan pola prestasi. Penelitian dimulai dari tahap pengumpulan data nilai akademik siswa MAN 2 Kota Cirebon yang diperoleh dalam format Excel. Selanjutnya, data yang telah dikumpulkan melalui tahap pra-pemrosesan, yang meliputi pembersihan data (cleaning) untuk mengatasi data yang tidak lengkap atau inkonsisten, normalisasi untuk menyamakan skala data, serta seleksi fitur guna memilih variabel yang paling relevan dalam analisis. Setelah data siap, dilakukan implementasi algoritma X-Means menggunakan Python, yang kemudian dibandingkan dengan algoritma K-Means untuk melihat performa dan efektivitas dalam menentukan jumlah klaster secara otomatis. Hasil clustering yang diperoleh kemudian dievaluasi menggunakan metode Davies-Bouldin Index untuk mengukur kualitas pemisahan antar klaster. Tahap berikutnya adalah analisis dan interpretasi hasil clustering guna mengidentifikasi kelompok prestasi siswa berdasarkan karakteristik tertentu. Terakhir, penelitian ditutup dengan penarikan kesimpulan dan penyusunan saran yang dapat diterapkan di lingkungan sekolah sebagai bahan pertimbangan dalam pengambilan keputusan, sebelum akhirnya penelitian dinyatakan selesai.

RESULTS AND DISCUSSION

Pada tahap implementasi algoritma *X-MEANS*, data nilai siswa yang sebelumnya sudah melalui proses pembersihan, seleksi atribut, dan normalisasi kemudian digunakan sebagai input utama untuk proses pengelompokan. Data tersebut berisi 263 siswa dengan delapan mata pelajaran yang dijadikan dasar klastering, yaitu Bahasa Indonesia, Bahasa Inggris, PJOK, Seni Budaya, Pendidikan Agama Islam, Ilmu Pengetahuan Sosial, Matematika, dan Pendidikan Kewarganegaraan, ditambah satu kolom penanda berupa kode numerik nama siswa (Nama_Encoded) agar setiap siswa tetap dapat dikenali setelah proses klastering. Seluruh fitur ini digabungkan menjadi satu himpunan data dan dijalankan di *Jupyter Notebook* menggunakan pustaka *pyClustering*. Pada tahap awal, sistem diberi batas jumlah klaster awal yang *relatif* kecil, lalu algoritma *X-MEANS* bekerja secara otomatis untuk membentuk, membelah, dan menata klaster berdasarkan pola kedekatan nilai antar siswa. Pusat klaster awal dipilih secara acak, kemudian algoritma mengelompokkan siswa yang memiliki pola nilai yang mirip ke dalam kelompok yang sama, mengecek apakah klaster yang terbentuk sudah cukup baik, dan jika belum, klaster tersebut dapat dipecah lagi hingga tercapai susunan klaster yang dianggap paling sesuai dengan struktur data.

Proses ini diringkas dalam tabel parameter dan hasil utama algoritma *X-MEANS*. Tabel tersebut menjelaskan secara singkat karakteristik data (jumlah siswa dan jumlah mata pelajaran yang digunakan), cara pembentukan fitur klastering, algoritma yang dipakai, cara pemilihan pusat klaster awal, serta metode pengukuran kedekatan antar data. Di bagian akhir, tabel memuat informasi inti berupa jumlah klaster yang dihasilkan sebagai solusi terbaik dan nilai *indeks Davies–Bouldin* yang digunakan untuk menilai kualitas klaster. Indeks ini menunjukkan seberapa rapat data di dalam satu klaster dan seberapa jelas pemisahan antar klaster semakin kecil nilainya, semakin baik kualitas pengelompokan yang terbentuk. Dengan melihat satu tabel ini saja, pembaca sudah bisa menangkap gambaran besar data apa yang digunakan, bagaimana pengaturannya, dan seberapa bagus hasil pengelompokan yang dihasilkan oleh algoritma.

Hasil klastering kemudian dijelaskan lebih lanjut melalui tabel jumlah anggota tiap klaster yang menampilkan berapa banyak siswa yang masuk ke setiap kelompok hasil pengelompokan. Tabel ini biasanya berisi *label* klaster dan jumlah siswa di dalamnya, sehingga memudahkan pembaca untuk melihat apakah sebaran siswa di setiap

klaster seimbang atau ada klaster tertentu yang jumlah anggotanya lebih dominan. Klaster dengan jumlah anggota yang lebih besar umumnya menggambarkan pola nilai yang lebih umum dimiliki banyak siswa, sedangkan klaster dengan anggota lebih sedikit menunjukkan kelompok siswa dengan karakteristik nilai yang lebih spesifik. Dalam konteks penelitian ini, klaster-klaster tersebut kemudian diinterpretasikan sebagai kelompok siswa berprestasi dan kelompok siswa yang nilainya *relatif* lebih rendah. Dengan demikian, rangkaian hasil yang disajikan melalui implementasi algoritma *X-MEANS*, tabel parameter dan hasil utama, serta tabel jumlah anggota tiap klaster memberikan dasar yang kuat untuk menganalisis pola prestasi siswa dan menyusun rekomendasi pada bagian pembahasan berikutnya.

Tabel 1 Parameter Dan Hasil Utama Algoritma *X-MEANS*

Parameter	Nilai
Jumlah data (n)	263 siswa
Jumlah fitur klastering	9 mata pelajaran
Algoritma	*X-MEANS*
Jumlah klaster minimum (K_min)	2
Metode inisialisasi pusat awal	Pemilihan acak (random)
Metrik jarak	Euclidean
Jumlah klaster optimal (K)	5
Davies–Bouldin Index (DBI)	0,6294

Tabel 1 menyajikan ringkasan parameter dan hasil utama penerapan algoritma *X-MEANS* pada data nilai siswa MAN 2 Kota Cirebon. Di dalam tabel, bagian pertama merangkum karakteristik data yang digunakan, yaitu jumlah sampel sebanyak 263 siswa dengan delapan mata pelajaran sebagai dasar pengelompokan, yang kemudian diolah menjadi sembilan fitur klastering (delapan nilai mata pelajaran dan satu fitur Nama_Encoded).

Bagian berikutnya menjelaskan konfigurasi *model*, antara lain algoritma yang dipakai adalah *X-MEANS* dengan jumlah klaster minimum Kmin=2K_{\text{min}} = 2Kmin=2, inisialisasi pusat klaster dilakukan secara acak, dan perhitungan jarak antar data menggunakan metrik Euclidean. Pada bagian akhir, tabel menampilkan hasil inti dari proses klastering, yaitu jumlah klaster optimal yang diperoleh sebanyak Lima klaster serta nilai *Davies–Bouldin Index (DBI)* sebesar 0.6294 yang menggambarkan kualitas pemisahan klaster yang cukup baik (semakin kecil DBI berarti klaster semakin kompak dan saling berjauhan). Dengan demikian, Tabel 4.2 memberikan gambaran singkat namun lengkap mengenai kondisi data, setting algoritma, dan kualitas hasil klastering, sehingga menjadi acuan utama dalam interpretasi pola kelompok prestasi siswa pada subbab analisis berikutnya.

Tabel 2 Jumlah Anggota Tiap Klaster

Klaster	Jumlah Anggota
0	55
1	51
2	49
3	53
4	55

Berikut Tabel 2 Menampilkan berapa banyak siswa yang masuk ke setiap klaster hasil pengelompokan *X-MEANS*. Tabel ini biasanya cuma punya dua kolom: kolom Klaster (misalnya Klaster 0 dan Klaster 1) dan kolom Jumlah Anggota yang menunjukkan berapa siswa di tiap klaster. Jadi, dari tabel ini pembaca bisa langsung tahu sebaran siswa: klaster mana yang anggotanya lebih banyak dan klaster mana yang anggotanya lebih sedikit.

Secara makna, klaster dengan jumlah anggota lebih besar bisa dianggap mewakili pola nilai yang lebih umum dialami banyak siswa, sedangkan klaster dengan anggota lebih sedikit menunjukkan kelompok siswa dengan pola nilai yang agak berbeda atau lebih khusus. Dalam konteks penelitian saya, kedua klaster ini kemudian diinterpretasikan sebagai Siswa Berprestasi dan Siswa Kurang Berprestasi**,** sehingga tabel jumlah anggota klaster membantu menggambarkan berapa banyak siswa yang termasuk ke tiap kategori tersebut.

Pada tahap evaluasi hasil *Clustering*, kualitas pengelompokan siswa dinilai dengan melihat beberapa output utama yang dihasilkan dari eksperimen di *Jupyter Notebook*. Hasil pertama yang diperhatikan adalah nilai *Davies–Bouldin Index (DBI)* dan ringkasan jumlah anggota tiap klaster. Dari output terlihat bahwa algoritma *X-MEANS* menghasilkan lima klaster dengan nilai DBI sebesar 0,6294. Nilai ini menunjukkan bahwa secara keseluruhan klaster yang terbentuk sudah cukup baik, karena setiap klaster memiliki sebaran data yang *relatif* rapat dan masih terpisah dengan jelas dari klaster lain. Selain itu, jumlah anggota di setiap klaster juga cukup seimbang, yaitu berkisar antara empat puluh satu hingga empat puluh tujuh siswa per klaster. Sebaran yang merata ini mengindikasikan bahwa tidak ada klaster yang terlalu kecil (hanya berisi sedikit siswa) maupun

terlalu besar, sehingga semua klaster tetap relevan untuk dianalisis lebih lanjut.

Evaluasi kemudian diperdalam dengan melihat tabel pusat klaster (*centroid*s) yang memuat rata-rata nilai delapan mata pelajaran pada masing-masing klaster. Setiap baris pada tabel tersebut merepresentasikan satu klaster, sedangkan setiap kolom menunjukkan rata-rata nilai suatu mata pelajaran, seperti Bahasa Indonesia, Bahasa Inggris, PJOK, Seni Budaya, Pendidikan Agama Islam, Ilmu Pengetahuan Sosial, Matematika, dan Pendidikan Kewarganegaraan.

Melalui tabel ini, peneliti dapat membandingkan pola nilai antar klaster, misalnya mengidentifikasi klaster yang rata-rata nilainya lebih tinggi di hampir semua mata pelajaran atau klaster yang memiliki kelemahan di mata pelajaran tertentu. Informasi ini menjadi dasar untuk memberi makna pada tiap klaster, apakah lebih dekat dengan kategori siswa berprestasi, sedang, atau perlu pendampingan. Untuk memastikan hasil *Clustering* benar-benar terhubung dengan data asli, ditampilkan juga contoh tabel data siswa yang sudah dilengkapi kolom "*Cluster*", sehingga dapat dilihat langsung siswa mana yang masuk ke klaster tertentu beserta nilai setiap mata pelajarannya.

Secara keseluruhan, kombinasi antara nilai DBI, distribusi jumlah anggota klaster, tabel pusat klaster, dan contoh data yang sudah ber*label* memberikan gambaran yang jelas bahwa hasil *Clustering* yang diperoleh cukup stabil dan dapat dipercaya. Evaluasi ini menunjukkan bahwa algoritma *X-MEANS* tidak hanya mampu membagi siswa ke dalam enam kelompok yang terstruktur, tetapi juga menghasilkan klaster yang bermakna dan bisa diinterpretasikan lebih lanjut pada subbab analisis dan interpretasi hasil, misalnya untuk memetakan kelompok siswa berprestasi dan kelompok siswa yang membutuhkan intervensi akademik.

Tabel 3 Nilai *Davies–Bouldin Index (DBI)* Untuk Berbagai Jumlah Klaster

No	Jumlah Klaster (K)	Nilai DBI
1	2	0,7634
2	3	0,6735
3	4	0,7632
4	5	0,6294
5	6	0,7230
.....	.....	.....
14	15	1,1891

Tabel 3 menyajikan hasil perhitungan nilai *Davies–Bouldin Index (DBI)* untuk berbagai jumlah klaster yang diuji pada algoritma *X-MEANS*. Setiap baris menampilkan nomor urut, jumlah klaster yang digunakan, dan nilai DBI yang dihasilkan pada konfigurasi tersebut. Melalui tabel ini, pembaca bisa melihat bagaimana kualitas pengelompokan berubah ketika jumlah klaster ditambah, sehingga tidak hanya melihat hasil akhir, tetapi juga proses pemilihan jumlah klaster yang paling sesuai.

Jika diperhatikan, nilai DBI pada Tabel 3 tidak bersifat konstan, tetapi naik turun seiring perubahan jumlah klaster. Nilai DBI terkecil muncul ketika jumlah klaster adalah lima, dengan nilai sebesar 0,6294. Pada jumlah klaster yang lebih sedikit ataupun lebih banyak, nilai DBI justru lebih besar. Karena DBI yang lebih kecil menunjukkan klaster yang lebih kompak di dalam dan cukup terpisah antar kelompok, maka lima klaster dipilih sebagai jumlah klaster yang paling optimal dalam penelitian ini.

CONCLUSION

Berdasarkan hasil penelitian dan analisis terhadap data nilai akademik 263 siswa MAN 2 Kota Cirebon tahun ajaran 2024/2025 menggunakan algoritma *X-MEANS Clustering*, dapat disimpulkan beberapa hal berikut.

Algoritma *X-MEANS* mampu menentukan jumlah klaster secara otomatis dan membentuk struktur pengelompokan prestasi siswa yang cukup baik dan stabil. Dari hasil eksperimen diperoleh lima klaster optimal dengan nilai *Davies–Bouldin Index (DBI)* sebesar 0,6294. Nilai ini menunjukkan bahwa klaster yang terbentuk memiliki tingkat kekompakan internal yang baik dan pemisahan antar-klaster yang masih jelas, sehingga *X-MEANS* dapat diandalkan untuk menganalisis pola nilai akademik siswa tanpa perlu menentukan jumlah klaster secara manual.

Hasil pengelompokan yang diperoleh kemudian dianalisis melalui nilai rata-rata delapan mata pelajaran pada setiap klaster dan dipetakan menjadi tiga kategori utama, yaitu berprestasi, cukup berprestasi, dan perlu penguatan akademik. Berdasarkan pemetaan tersebut, Klaster 2 dan Klaster 3 dikategorikan sebagai kelompok berprestasi dengan total 84 siswa, Klaster 0 dan Klaster 4 termasuk dalam kategori cukup berprestasi dengan total 125 siswa, sedangkan Klaster 1 masuk dalam kategori perlu penguatan akademik dengan jumlah 54 siswa. Hal ini menunjukkan bahwa algoritma *X-MEANS* efektif dalam mendeteksi variasi prestasi akademik siswa dan mengelompokkan mereka ke dalam kategori yang bermakna berdasarkan pola distribusi nilai yang dimiliki.

Penerapan algoritma *X-MEANS* pada data akademik memberikan dampak positif dalam konteks *Educational* Data Mining (EDM). Model pengelompokan yang dihasilkan dapat dimanfaatkan oleh pihak sekolah untuk mengidentifikasi kelompok siswa yang membutuhkan pendampingan akademik lebih awal, merancang program pengayaan bagi siswa berprestasi, serta mendukung perencanaan strategi pembelajaran yang lebih terarah dan berbasis data. Dengan demikian, algoritma *X-MEANS* berpotensi menjadi salah satu alat bantu analitik yang mendukung pengambilan keputusan di bidang pendidikan, khususnya dalam pemetaan dan pemantauan prestasi akademik siswa secara lebih sistematis.

REFERENCE

[1] M. Yağcı, "Educational data mining: prediction of students' academic performance using machine learning algorithms," *Smart Learning Environments*, vol. 9, no. 1, 2022, doi: 10.1186/s40561-022-00192-z.

[2] G. Raftopoulos, G. Davrazos, and S. Kotsiantis, "Evaluating Fairness Strategies in Educational Data Mining : A Comparative Study of Bias Mitigation Techniques," 2025.

[3] Z. Wang, "Higher Education Management and Student Achievement," vol. 2022, 2022.

[4] N. Izzati, M. Talib, N. Aini, A. Majid, and S. Sahran, "applied sciences Identification of Student Behavioral Patterns in Higher Education Using K-Means Clustering and Support Vector Machine," 2023.

[5] A. F. Mohamed Nafuri, N. S. Sani, N. F. A. Zainudin, A. H. A. Rahman, and M. Aliff, "Clustering Analysis for Classifying Student Academic Performance in Higher Education," *Applied Sciences (Switzerland)*, vol. 12, no. 19, 2022, doi: 10.3390/app12199467.

[6] C. Ramdani, I. Artikel, and K. Kunci, "Komparasi Algoritme X-Means dan K-Medoids pada Klasterisasi Data Akademik Mahasiswa," 2023.

[7] V. Wulandari, Y. Syarif, and Z. Alfian, "Comparison of Density-Based Spatial Clustering of Applications with Noise (DBSCAN), K-Means and X-Means Algorithms on Shopping Trends Data," vol. 1, no. February, pp. 1–8, 2024.

[8] S. A. Mukhsyi, A. I. Purnamaari, and A. Bahtiar, "Improving Student Achievement Clustering Model Using K-Means Algorithm in Pasundan Majalaya Vocational School," vol. 4, no. 2, 2025.

[9] S. Raihanah and T. Wahyudi, "K-Means Clustering of Social Studies Performance at Junior High School," vol. 13, no. 2, pp. 460–472, 2024, doi: 10.14421/ijid.2024.4632.