

Analisis Nilai dan Kehadiran Mahasiswa untuk Pengembangan Pengetahuan Akademik Menggunakan Clustering K-Means dan Statistik

¹Wetina Hulu, ²Rani Rosalinda, ³Ronita Olive Angelie, ⁴Sardo Sipayung

^{1,2,3,4}Universitas Katolik Santo Thomas, Medan, Indonesia,

¹wetinahulu@gmail.com, ²lindarani786@gmail.com, ³ronitasitanggung589@gmail.com,

⁴pinsarsiphom@gmail.com

Submit : 25 Des 2025 | Diterima : 18 Jan 2026 | Terbit : 20 Jan 2026

ABSTRAK

Nilai akademik dan tingkat kehadiran mahasiswa merupakan dua indikator utama yang sering digunakan untuk menilai perkembangan akademik di perguruan tinggi. Perbedaan capaian nilai dan kehadiran antar mahasiswa kerap terjadi, namun belum seluruhnya dianalisis secara terstruktur untuk menghasilkan informasi yang mendukung pengambilan keputusan akademik. Penelitian ini bertujuan untuk menganalisis data nilai akademik dan kehadiran mahasiswa dalam mengidentifikasi pola perkembangan akademik melalui pendekatan kuantitatif. Metode yang digunakan meliputi analisis statistik deskriptif dan algoritma clustering K-Means. Statistik deskriptif diterapkan untuk menggambarkan karakteristik data, seperti nilai rata-rata data, capaian terendah, serta capaian tertinggi mahasiswa persentase kehadiran mahasiswa. Selanjutnya, algoritma K-Means digunakan untuk mengelompokkan mahasiswa ke dalam beberapa cluster berdasarkan kesamaan nilai akademik dan tingkat kehadiran. Data penelitian diperoleh dari mahasiswa aktif pada satu program studi dalam satu periode akademik. Berdasarkan hasil analisis, mahasiswa dapat dikelompokkan ke dalam beberapa kategori dengan ciri akademik yang beragam, yaitu kelompok dengan nilai dan kehadiran tinggi, kelompok sedang, serta kelompok dengan nilai dan kehadiran rendah. Temuan ini menunjukkan bahwa hasil pengelompokan dapat memberikan gambaran kondisi akademik mahasiswa secara sistematis dan menjadi dasar pertimbangan bagi pihak perguruan tinggi dalam meningkatkan kualitas proses pembelajaran.

Kata Kunci: *Clustering*; Kehadiran Mahasiswa; *K-Means*; Nilai Akademik

PENDAHULUAN

Nilai akademik dan kehadiran mahasiswa merupakan dua indikator penting dalam menilai keberhasilan proses pembelajaran di perguruan tinggi. Nilai akademik menggambarkan tingkat penguasaan mahasiswa terhadap materi perkuliahan, sedangkan kehadiran menunjukkan partisipasi serta kedisiplinan mahasiswa dalam mengikuti kegiatan akademik. Hubungan antara kedua aspek tersebut telah banyak dikaji dalam penelitian sebelumnya. Astuti dan Pratama (2019) menyebutkan bahwa tingkat pemahaman mahasiswa memiliki keterkaitan erat dengan capaian nilai akademik, sementara Siregar dan Nasution (2020) menjelaskan bahwa kehadiran yang tinggi berpengaruh positif terhadap prestasi belajar mahasiswa.

Perkembangan sistem informasi akademik saat ini memungkinkan perguruan tinggi menyimpan data nilai dan kehadiran mahasiswa dalam jumlah besar. Namun, pemanfaatan data tersebut umumnya masih terbatas pada kebutuhan administratif, seperti penyusunan laporan akademik dan pengisian Kartu Hasil Studi (KHS). Padahal, data yang tersedia memiliki potensi besar untuk dianalisis lebih lanjut guna mendukung evaluasi pembelajaran dan perumusan kebijakan akademik. Rahmawati, Putra, dan Lestari (2021) menegaskan bahwa analisis statistik deskriptif dapat memberikan gambaran awal mengenai kondisi akademik mahasiswa, meskipun belum mampu mengelompokkan mahasiswa berdasarkan karakteristik yang lebih spesifik.

Untuk memperoleh pemahaman yang lebih mendalam, diperlukan pendekatan analisis lanjutan melalui teknik data mining, khususnya metode clustering. Han, Kamber, dan Pei (2012)

menjelaskan bahwa Clustering adalah metode analisis data yang digunakan untuk mengelompokkan objek berdasarkan kesamaan karakteristik tertentu sehingga pola-pola tertentu dapat diidentifikasi secara lebih jelas. Salah satu algoritma yang banyak digunakan adalah K-Means karena kemudahannya dalam implementasi serta kemampuannya mengelola data dalam jumlah besar. Hidayat dan Prasetyo (2022) membuktikan bahwa penerapan algoritma K-Means pada data nilai dan kehadiran mahasiswa mampu menghasilkan pengelompokan yang objektif dan membantu pihak akademik dalam memahami variasi capaian belajar mahasiswa.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk menganalisis data nilai akademik dan kehadiran mahasiswa menggunakan analisis statistik deskriptif serta melakukan pengelompokan mahasiswa dengan metode K-Means. Hasil penelitian ini diharapkan dapat memberikan gambaran kondisi akademik mahasiswa secara lebih jelas dan terstruktur, serta menjadi bahan pertimbangan bagi perguruan tinggi dalam meningkatkan kualitas pembelajaran dan strategi pembinaan akademik.

TINJAUAN PUSTAKA

Nilai Akademik

Menurut Bloom (1956), hasil belajar merupakan pencapaian sasaran pembelajaran yang meliputi ranah kognitif dan afektif, dan psikomotorik. Nilai akademik menjadi salah satu indikator utama dalam menilai keberhasilan mahasiswa dalam proses pembelajaran. Astuti dan Pratama (2019) menjelaskan bahwa nilai akademik memiliki hubungan yang erat dengan tingkat pemahaman mahasiswa terhadap materi perkuliahan. Semakin baik pemahaman mahasiswa, semakin tinggi pula capaian nilai akademik yang diperoleh.

Kehadiran Mahasiswa

Kehadiran mahasiswa dalam kegiatan perkuliahan mencerminkan tingkat kedisiplinan dan partisipasi aktif dalam proses pembelajaran. Siregar dan Nasution (2020) menyatakan bahwa tingkat kehadiran memiliki pengaruh yang signifikan terhadap prestasi akademik mahasiswa. Mahasiswa dengan tingkat kehadiran yang tinggi cenderung lebih mudah memahami materi perkuliahan karena memperoleh penjelasan langsung dari dosen serta dapat berinteraksi secara aktif dalam kegiatan belajar mengajar.

Analisis Statistik Deskriptif

Analisis statistik deskriptif merupakan metode yang digunakan untuk menggambarkan karakteristik suatu data dengan menghitung nilai rata-rata, capaian minimum, capaian maksimum, dan distribusi data secara keseluruhan. Rahmawati, Putra, dan Lestari (2021) menjelaskan bahwa statistik deskriptif dapat memberikan gambaran awal mengenai kondisi akademik mahasiswa. Namun, metode ini masih memiliki keterbatasan karena belum mampu mengelompokkan mahasiswa berdasarkan kesamaan karakteristik akademik yang dimiliki.

Clustering dan Metode K-Means

Clustering merupakan salah satu teknik dalam data mining yang digunakan untuk mengelompokkan data berdasarkan tingkat kemiripan tertentu. Han, Kamber, dan Pei (2012) menyatakan bahwa clustering sangat bermanfaat dalam menemukan pola tersembunyi pada data berukuran besar, termasuk data akademik mahasiswa. Salah satu algoritma clustering yang banyak digunakan adalah K-Means karena kemudahannya dalam implementasi serta efisiensinya dalam mengolah data numerik. Hidayat dan Prasetyo (2022) menjelaskan bahwa penerapan algoritma K-Means dalam pengelompokan mahasiswa berdasarkan nilai dan kehadiran mampu menghasilkan kelompok mahasiswa dengan karakteristik prestasi yang berbeda, sehingga dapat membantu pihak perguruan tinggi dalam melakukan evaluasi dan pembinaan akademik secara lebih terarah.

METODE PENELITIAN

Jenis dan Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan fokus pada analisis data numerik berupa nilai akademik dan persentase kehadiran mahasiswa. Pendekatan kuantitatif dipilih karena data yang digunakan dapat diolah secara statistik dan matematis untuk menghasilkan informasi yang objektif. Selain analisis statistik deskriptif, penelitian ini menerapkan metode clustering untuk mengelompokkan mahasiswa berdasarkan kemiripan karakteristik akademik. Algoritma K-Means digunakan karena efektif dalam mengolah data numerik dan telah banyak diaplikasikan dalam penelitian analisis data pendidikan.

Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari sistem informasi akademik perguruan tinggi. Data bersumber dari basis data akademik resmi yang mencatat nilai akhir mata kuliah dan persentase kehadiran mahasiswa selama satu periode perkuliahan. Pemanfaatan data sekunder dipilih karena data telah terdokumentasi secara sistematis, bersifat objektif, serta mencerminkan kondisi akademik mahasiswa yang sebenarnya. Dataset penelitian terdiri dari data akademik 100 mahasiswa yang memiliki data nilai dan kehadiran lengkap dan valid, sehingga layak digunakan untuk proses analisis statistik dan clustering.

Jenis Data

Jenis data yang digunakan dalam penelitian ini adalah data kuantitatif. Data kuantitatif dipilih karena seluruh variabel penelitian dinyatakan dalam bentuk angka dan dapat dianalisis menggunakan pendekatan statistik serta algoritma clustering. Data numerik memungkinkan proses pengukuran, perhitungan, dan pengelompokan dilakukan secara objektif dan terstruktur.

Tabel 1. Variabel Penelitian

No	Variabel Penelitian	Deskripsi	Skala Data
1	Nilai Akademik	Nilai akhir mahasiswa pada mata kuliah dalam satu periode perkuliahan	Numerik
2	Kehadiran Mahasiswa	Persentase kehadiran mahasiswa selama mengikuti perkuliahan	Numerik

Kedua variabel tersebut dipilih karena dinilai mampu menggambarkan kondisi akademik mahasiswa secara kuantitatif. Nilai akademik mencerminkan tingkat penguasaan mahasiswa terhadap materi pembelajaran, sedangkan kehadiran mahasiswa menunjukkan tingkat partisipasi, kedisiplinan, dan keterlibatan mahasiswa dalam proses perkuliahan.

Tahapan Pengolahan Data

Pengolahan data dilakukan melalui beberapa tahapan yang sistematis untuk memastikan data yang dianalisis memiliki kualitas yang baik dan layak digunakan. Tahapan pengolahan data dalam penelitian ini disajikan pada Tabel 2.

Tabel 2. Tahapan Pengolahan Data

No	Tahapan	Uraian
1	Pemeriksaan Data	Memastikan kelengkapan dan konsistensi data
2	Pembersihan Data	Menghapus data yang tidak lengkap, kosong, atau tidak valid
3	Normalisasi Data	Menyamakan skala antar variabel agar tidak terjadi bias
4	Analisis Statistik	Menghitung statistik deskriptif data
5	Clustering K-Means	Mengelompokkan mahasiswa berdasarkan nilai dan kehadiran
6	Analisis dan Interpretasi	Menafsirkan hasil cluster yang terbentuk

Analisis Statistik Deskriptif

Analisis statistik deskriptif digunakan untuk memberikan gambaran awal mengenai karakteristik data nilai akademik dan kehadiran mahasiswa. Statistik deskriptif yang digunakan dalam penelitian ini meliputi nilai rata-rata (mean), nilai minimum, dan nilai maksimum.

Rumus Rata-rata (Mean):

Rata-rata (Mean)

$$\tilde{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Keterangan:

\tilde{x} = nilai rata – rata

x_i = nilai data ke – i

n = jumlah dat

Nilai Minimum dan Maksimum

Nilai minimum merupakan nilai terendah yang terdapat dalam dataset, sedangkan nilai maksimum adalah nilai tertinggi yang diperoleh. Informasi mengenai nilai minimum dan maksimum digunakan untuk mengetahui rentang data serta variasi awal sebelum dilakukan proses pengelompokan menggunakan metode clustering.

Normalisasi Data

Normalisasi data dilakukan untuk menyamakan skala antara variabel nilai akademik dan kehadiran mahasiswa, sehingga tidak terjadi dominasi salah satu variabel dalam proses clustering. Metode normalisasi yang digunakan dalam penelitian ini adalah Min-Max Normalization.

Rumus Normalisasi Min-Max:

$$x^l = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Keterangan:

x^l = nilai hasil normalisasi

x = nilai asli

x_{min} = nilai minimum

x_{max} = nilai maksimum

Clustering Menggunakan Metode K-Means

Metode K-Means digunakan untuk mengelompokkan mahasiswa ke dalam beberapa cluster berdasarkan tingkat kemiripan nilai akademik dan persentase kehadiran. Algoritma ini bekerja dengan cara meminimalkan jarak antara data dan pusat cluster (centroid).

Tahapan Algoritma K-Means:

- Menentukan jumlah cluster (k)
- Menentukan centroid awal secara acak
- Menghitung jarak setiap data terhadap centroid
- Mengelompokkan data ke centroid terdekat
- Memperbarui nilai centroid
- Mengulangi proses hingga centroid tidak mengalami perubahan signifikan

Rumus Jarak Euclidean

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

Keterangan:

$d(x, y)$ = jarak antara data dan centroid

x_i = nilai data ke – i

y_i = nilai centroid ke - i

n = jumlah variabel

Analisis dan Interpretasi Cluster

Hasil clustering yang diperoleh selanjutnya dianalisis untuk mengidentifikasi karakteristik masing-masing cluster. Analisis dilakukan dengan membandingkan nilai rata-rata nilai akademik dan persentase kehadiran mahasiswa pada setiap cluster.

Tabel 3. Interpretasi Cluster

Cluster	Nilai Akademik	Kehadiran	Interpretasi
1	Tinggi	Tinggi	Mahasiswa berprestasi
2	Sedang	Sedang	Mahasiswa berprestasi sedang
3	Rendah	Rendah	Mahasiswa memerlukan pembinaan

Hasil interpretasi cluster ini digunakan sebagai dasar evaluasi akademik dan rekomendasi bagi perguruan tinggi dalam meningkatkan kualitas pembelajaran serta program pendampingan mahasiswa.

HASIL DAN PEMBAHASAN

Hasil penelitian diperoleh melalui analisis statistik deskriptif dan proses clustering terhadap data nilai akademik dan kehadiran mahasiswa. Analisis statistik deskriptif bertujuan untuk memberikan gambaran awal mengenai sebaran dan karakteristik data sebelum dilakukan pengelompokan. Ringkasan hasil analisis statistik deskriptif disajikan pada Tabel 4.

Berdasarkan Tabel 4, nilai akademik mahasiswa menunjukkan variasi yang cukup lebar dengan nilai terendah sebesar 45,0 dan nilai tertinggi mencapai 95,0. Nilai rata-rata akademik sebesar 72,5 mengindikasikan bahwa secara umum capaian akademik mahasiswa berada pada kategori cukup baik. Nilai standar deviasi sebesar 12,8 menunjukkan adanya perbedaan tingkat pencapaian akademik antar mahasiswa yang relatif signifikan.

Hasil Analisis Statistik Deskriptif

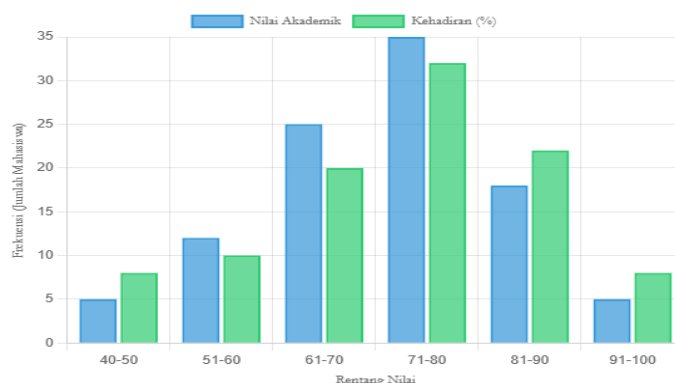
Analisis statistik deskriptif dilakukan terhadap data akademik yang terdiri atas dua variabel utama, yaitu nilai akademik dan persentase kehadiran mahasiswa. Data yang dianalisis berjumlah 100 mahasiswa yang telah melalui proses pemeriksaan, pembersihan, dan normalisasi data sebelum digunakan dalam analisis lanjutan serta proses clustering menggunakan metode K-Means.

Untuk variabel kehadiran mahasiswa, persentase kehadiran minimum tercatat sebesar 40%, sedangkan nilai maksimum mencapai 100%. Nilai rata-rata kehadiran mahasiswa sebesar 78,3% menunjukkan bahwa tingkat partisipasi mahasiswa dalam kegiatan perkuliahan tergolong baik. Standar deviasi kehadiran sebesar 15,2% mengindikasikan adanya variasi tingkat kedisiplinan mahasiswa dalam mengikuti perkuliahan.

Tabel 4. Hasil Analisis Statistik Deskriptif

Variabel	Mean	Min	Max	Std. Deviation
Nilai Akademik	72.5	45.0	95.0	12.8
Kehadiran (%)	78.3	40.0	100.0	15.2

Sebaran nilai akademik dan persentase kehadiran mahasiswa divisualisasikan menggunakan histogram seperti yang ditampilkan pada Gambar 1. Berdasarkan visualisasi tersebut, sebagian besar mahasiswa memiliki nilai akademik pada rentang 65–80 dengan pola distribusi yang mendekati normal. Sementara itu, data kehadiran menunjukkan bahwa mayoritas mahasiswa memiliki tingkat kehadiran di atas 70%, meskipun masih terdapat sejumlah mahasiswa dengan tingkat kehadiran di bawah 60%.



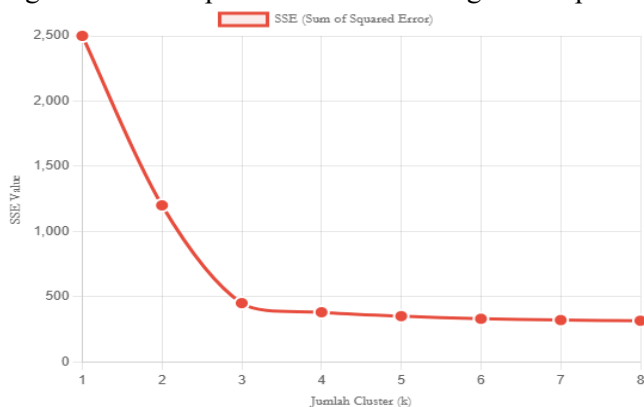
Gambar 1. Distribusi Nilai Akademik dan Kehadiran Mahasiswa

Hasil Clustering K-Means

Proses clustering dilakukan setelah seluruh data dinormalisasi menggunakan metode Min-Max Normalization untuk memastikan kesetaraan skala antar variabel. Penentuan jumlah cluster optimal dilakukan dengan metode Elbow, yang menghasilkan tiga cluster sebagai jumlah paling representatif untuk mengelompokkan mahasiswa berdasarkan kesamaan capaian akademik dan tingkat kehadiran.

Penentuan Jumlah Cluster

Metode Elbow digunakan untuk menentukan jumlah cluster optimal dengan menganalisis perubahan nilai Sum of Squared Error (SSE) terhadap berbagai nilai k. Hasil analisis yang ditunjukkan pada Gambar 2 memperlihatkan penurunan nilai SSE yang cukup tajam hingga k = 3, kemudian penurunan cenderung melandai. Berdasarkan pola tersebut dengan demikian, jumlah cluster yang dianggap paling sesuai dalam penelitian ini adalah tiga kelompok.



Gambar 2. Grafik Elbow Method untuk Penentuan Jumlah Cluster

Hasil Pengelompokan Mahasiswa

Hasil akhir proses K-Means menghasilkan tiga cluster mahasiswa dengan distribusi anggota yang berbeda. Distribusi jumlah mahasiswa pada masing-masing cluster disajikan pada Tabel 5.

Berdasarkan Tabel 5, Cluster 1 merupakan cluster dengan jumlah anggota terbanyak, yaitu 45 mahasiswa (45%). Cluster 2 terdiri dari 35 mahasiswa (35%), sedangkan Cluster 3 beranggotakan 20 mahasiswa (20%). Distribusi ini menunjukkan bahwa mayoritas mahasiswa berada pada kelompok tertentu dengan karakteristik akademik yang relatif serupa.

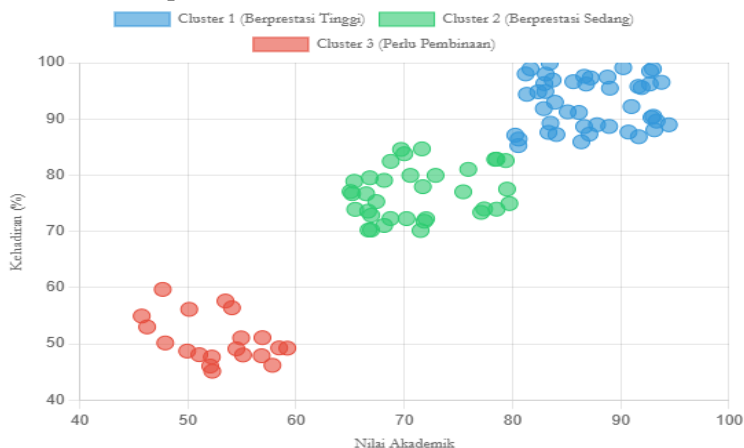
Tabel 5. Distribusi Mahasiswa pada Setiap Cluster

Cluster	Jumlah Mahasiswa	Persentase
Cluster 1	45	45%
Cluster 2	35	35%

Cluster	Jumlah Mahasiswa	Persentase
Cluster 3	20	20%
Total	100	100%

Visualisasi Hasil Clustering Mahasiswa Berdasarkan Nilai Akademik dan Kehadiran

Visualisasi hasil clustering ditampilkan dalam bentuk scatter plot seperti pada Gambar 3. Visualisasi tersebut menunjukkan bahwa setiap cluster memiliki area sebaran yang relatif terpisah, yang menandakan bahwa proses clustering berhasil mengelompokkan mahasiswa berdasarkan kemiripan nilai akademik dan tingkat kehadiran. Cluster 1 berada pada area nilai tinggi, Cluster 2 pada area menengah, dan Cluster 3 pada area rendah.



Gambar 3. Visualisasi Hasil Clustering Mahasiswa Berdasarkan Nilai Akademik dan Kehadiran

Karakteristik Setiap Cluster

Karakteristik masing-masing cluster dianalisis dengan menghitung nilai rata-rata (centroid) untuk variabel nilai akademik dan kehadiran mahasiswa. Hasil analisis karakteristik cluster disajikan pada Tabel 6.

Tabel 6. Karakteristik Setiap Cluster Berdasarkan Nilai Akademik dan Kehadiran

Cluster	Rata-rata Nilai Akademik	Rata-rata Kehadiran (%)	Interpretasi
Cluster 1	85.2	92.5	Mahasiswa Berprestasi Tinggi
Cluster 2	72.8	78.0	Mahasiswa Berprestasi Sedang
Cluster 3	55.4	55.2	Mahasiswa Perlu Pembinaan

Berdasarkan Tabel 6, dapat dijelaskan karakteristik masing-masing cluster sebagai berikut:

Cluster 1 mencakup mahasiswa yang memiliki prestasi akademik tinggi serta tingkat kehadiran yang optimal. Kelompok ini menunjukkan konsistensi antara kedisiplinan dan prestasi akademik, sehingga dapat dijadikan contoh bagi mahasiswa lainnya.

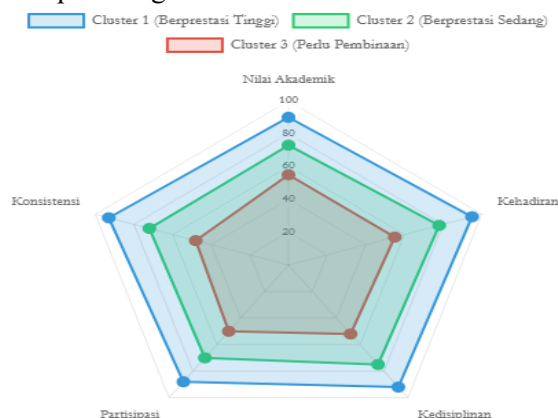
Cluster 2 mencakup mahasiswa dengan prestasi akademik dan kehadiran pada tingkat menengah. Kelompok ini masih memiliki potensi untuk meningkatkan prestasi melalui peningkatan kedisiplinan.

Cluster 3 terdiri dari mahasiswa dengan nilai akademik dan kehadiran yang relatif rendah. Kondisi ini menunjukkan perlunya perhatian dan pembinaan khusus dari pihak akademik.

Profil Karakteristik Setiap Cluster

Profil karakteristik setiap cluster divisualisasikan menggunakan radar chart seperti pada Gambar 4. Visualisasi tersebut memperlihatkan perbedaan karakteristik antar cluster secara jelas,

di mana Cluster 1 memiliki nilai tertinggi pada kedua variabel, Cluster 2 berada pada tingkat sedang, dan Cluster 3 berada pada tingkat rendah.



Gambar 4. Profil Karakteristik Setiap Cluster

Hasil ini menunjukkan bahwa metode K-Means efektif dalam mengelompokkan mahasiswa berdasarkan nilai akademik dan kehadiran, sehingga memberikan gambaran kondisi akademik mahasiswa secara lebih terstruktur.

KESIMPULAN

Berdasarkan hasil analisis dan pembahasan, dapat disimpulkan bahwa nilai akademik mahasiswa memiliki rata-rata sebesar 72,5 dengan rata-rata kehadiran 78,3%, yang menunjukkan adanya variasi capaian akademik dan tingkat kedisiplinan mahasiswa. Penerapan metode K-Means clustering berhasil mengelompokkan mahasiswa ke dalam tiga cluster dengan karakteristik yang berbeda, yaitu mahasiswa berprestasi tinggi (45%), berprestasi sedang (35%), dan mahasiswa yang memerlukan pembinaan (20%).

Hasil penelitian menunjukkan adanya hubungan positif antara tingkat kehadiran dan prestasi akademik mahasiswa. Pengelompokan yang dihasilkan dapat dimanfaatkan sebagai sistem pendukung pengambilan keputusan akademik, khususnya dalam merancang strategi pembinaan dan pendampingan mahasiswa yang lebih tepat sasaran.

Untuk penelitian selanjutnya, disarankan agar variabel tambahan seperti motivasi belajar dan latar belakang sosial ekonomi turut dipertimbangkan. Selain itu, penggunaan algoritma clustering lain serta analisis longitudinal dapat dilakukan untuk memperoleh gambaran perkembangan akademik mahasiswa secara lebih komprehensif.

REFERENSI

- Astuti R, Pratama D. Analisis hubungan nilai akademik dengan tingkat pemahaman mahasiswa. *Jurnal Pendidikan Tinggi*. 2019;13(2):85–94.
- Bloom BS. *Taxonomy of Educational Objectives: The Classification of Educational Goals*. New York: Longmans, Green and Co.; 1956.
- Han J, Kamber M, Pei J. *Data Mining: Concepts and Techniques*. 3rd ed. San Francisco: Morgan Kaufmann; 2012.
- Hidayat A, Prasetyo E. Penerapan algoritma K-Means untuk pengelompokan mahasiswa berdasarkan nilai dan kehadiran. *Jurnal Teknologi Informasi dan Pendidikan*. 2022;15(1):45–54.
- Jain AK. Data clustering: 50 years beyond K-Means. *Pattern Recognition Letters*. 2010;31(8):651–666.
- Moore MG, Kearsley G. *Distance Education: A Systems View of Online Learning*. 3rd ed. Belmont (CA): Wadsworth; 2012.
- Rahmawati D, Putra A, Lestari S. Analisis statistik deskriptif data akademik mahasiswa sebagai dasar evaluasi pembelajaran. *Jurnal Pendidikan dan Statistik*. 2021;9(2):101–110.

-
- Siregar H, Nasution R. Pengaruh kehadiran terhadap prestasi akademik mahasiswa. *Jurnal Ilmu Pendidikan*. 2020;8(1):25–33.
- Sugiyono. *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta; 2017.
- Wijaya F, Putri NA. Penerapan data mining clustering untuk pengelompokan mahasiswa berdasarkan prestasi akademik. *Jurnal Sistem Informasi*. 2020;14(2):67–76.
- Romero C, Ventura S. Educational data mining: A review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*. 2010;40(6):601–618. doi:10.1109/TSMCC.2010.2053532.
- Baker RSJD, Inventado PS. Educational data mining and learning analytics. In: Larusson JA, White B, editors. *Learning Analytics: From Research to Practice*. New York: Springer; 2014. p. 61–75. doi:10.1007/978-1-4614-3305-7_4.
- Hair JF, Black WC, Babin BJ, Anderson RE. *Multivariate Data Analysis*. 7th ed. Upper Saddle River (NJ): Pearson Education; 2014.
- Aggarwal CC, Reddy CK, editors. *Data Clustering: Algorithms and Applications*. Boca Raton (FL): CRC Press; 2014.
- Tinto V. Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*. 1975;45(1):89–125. doi:10.3102/00346543045001089.