

Deteksi Dini Permasalahan Akademik Mahasiswa Menggunakan Algoritma K-Means

Muhardi*, Hasriani, Thabrani, Heriadi

Universitas Dipa Makassar; Jalan Perintis Kemerdekaan Km. 9 Makassar, Telp. (0411) 587194 – Fax. (0411) 588284

e-mail: *¹muhardi23dipa@undipa.ac.id, ²hasriani@undipa.ac.id, ³thabrani@undipa.ac.id, ⁴heriadi@undipa.ac.id

Abstrak

Penelitian ini bertujuan menganalisis pola akademik mahasiswa menggunakan metode clustering berdasarkan atribut IPK, nilai mata kuliah inti, jumlah SKS, mata kuliah diulang, status cuti, dan semester. Penelitian ini juga mengembangkan sistem informasi berbasis web menggunakan framework Laravel sebagai media implementasi algoritma K-Means sehingga proses pengolahan dan visualisasi data dapat dilakukan secara terintegrasi. Data akademik sebanyak 100 mahasiswa diproses melalui tahap prapemrosesan berupa normalisasi Min-Max, kemudian dikelompokkan menggunakan algoritma K-Means dengan empat cluster. Hasil pengelompokan menunjukkan bahwa setiap cluster memiliki karakteristik akademik yang berbeda, yaitu mahasiswa dengan IPK rendah namun nilai inti tinggi, mahasiswa dengan capaian akademik sedang, mahasiswa yang sering mengambil cuti, serta mahasiswa dengan kecenderungan mengulang mata kuliah dan cuti tinggi. Atribut IPK, nilai inti, dan mata kuliah diulang menjadi indikator utama dalam membedakan kelompok mahasiswa berdasarkan tingkat risiko akademik. Sistem berbasis web yang dibangun mampu menyajikan hasil clustering secara visual dan informatif sehingga dapat dimanfaatkan sebagai alat bantu bimbingan akademik dan deteksi dini permasalahan akademik mahasiswa.

Kata kunci: K-Means Clustering, Data Akademik, Sistem Berbasis Web, Laravel.

Abstract

This study aims to analyze students' academic patterns using clustering based on Grade Point Average (GPA), core course grades, number of completed credits, repeated courses, leave status, and semester. This research also develops a web-based information system using the Laravel framework to implement the K-Means algorithm, enabling integrated data processing and visualization. Academic data from 100 students were processed through a preprocessing stage using Min-Max normalization and then clustered using the K-Means algorithm with four clusters. The results show that each cluster has distinct academic characteristics, including students with low GPA but high core course grades, students with moderate academic performance, students who frequently take academic leave, and students with high levels of course repetition and leave status. GPA, core course grades, and repeated courses are identified as the main indicators for distinguishing student groups based on academic risk levels. The developed web-based system is able to present clustering results in a visual and informative manner, supporting academic advising and early detection of students at risk of academic problems.

Keywords: K-Means Clustering, Academic Data, Web-Based System, Laravel.

1. PENDAHULUAN

Peningkatan kualitas pendidikan tinggi sangat bergantung pada kemampuan institusi dalam memantau dan mengevaluasi kinerja akademik mahasiswa secara efektif. Dalam praktiknya, akademik mahasiswa sering kali dinilai berdasarkan Indeks Prestasi Kumulatif (IPK), nilai mata kuliah, jumlah SKS yang lulus, riwayat cuti, dan lama studi. Data data ini apabila tidak dianalisis secara sistematis dapat menyulitkan pihak akademik dalam mengambil keputusan terkait intervensi atau bimbingan yang tepat sasaran. Oleh karena itu, dibutuhkan pendekatan data driven yang mampu mengolah data akademik menjadi informasi yang bermakna dan mendukung tindakan proaktif untuk mengidentifikasi mahasiswa yang berpotensi mengalami permasalahan akademik.

Salah satu pendekatan yang banyak digunakan dalam analisis data pendidikan adalah clustering, khususnya algoritma K Means, yang berfungsi untuk mengelompokkan data berdasarkan kemiripan karakteristiknya tanpa menggunakan label terlebih dahulu. Metode K Means telah banyak diterapkan dalam

konteks pendidikan untuk mengelompokkan mahasiswa berdasarkan prestasi akademik atau nilai, sehingga masing-masing cluster merepresentasikan kelompok dengan pola akademik yang berbeda dan memudahkan strategi pembimbingan [1], [2].

Penelitian terdahulu algoritma K-Means untuk melakukan clustering terhadap nilai mahasiswa menunjukkan bahwa metode K-Means mampu mengelompokkan mahasiswa berdasarkan kesamaan karakteristik nilai akademik, sehingga memudahkan pihak akademik dalam mengidentifikasi kelompok mahasiswa dengan capaian belajar yang berbeda [3], [4].

Meskipun demikian, sebagian penelitian cenderung hanya menggunakan dua atau tiga atribut sederhana seperti IPK dan nilai saja, seperti yang telah dilakukan oleh Anas, dkk. dengan judul Clustering Nilai Mahasiswa Menggunakan Algoritma K Means. Pada penelitian ini menggabungkan beragam atribut akademik (IPK per semester, nilai mata kuliah inti, jumlah SKS lulus, riwayat cuti, lama studi, dan frekuensi mengulang) untuk mendapatkan model clustering yang lebih komprehensif dan akurat dalam mengidentifikasi mahasiswa dengan risiko akademik tinggi maupun rendah.

2. METODOLOGI PENELITIAN

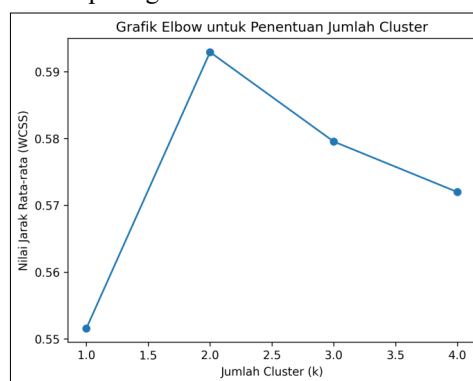
2.1 Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dan pengembangan sistem dengan metode deskriptif analisis dan implementasi berbasis web. Pendekatan kuantitatif dipakai untuk menganalisis data akademik mahasiswa, sedangkan pengembangan sistem berbasis web menggunakan Laravel dipilih agar aplikasi dapat diakses secara multiplatform melalui browser. Sistem ini bertujuan untuk memfasilitasi analisis akademik mahasiswa dengan fitur clustering K Means untuk mendeteksi kelompok mahasiswa berdasarkan karakteristik akademik (IPK, nilai mata kuliah inti, jumlah SKS, mata kuliah diulang, cuti, dan semester) [5], [6].

2.2 Langkah-langkah Penelitian

Penelitian ini dilaksanakan melalui beberapa tahapan sistematis untuk memperoleh hasil clustering mahasiswa yang representatif. Adapun langkah-langkah penelitian yang dilakukan adalah sebagai berikut:

1. Pengumpulan Data
Data penelitian diperoleh dari sistem akademik perguruan tinggi yang mencakup atribut IPK, Rata-rata nilai mata kuliah, jumlah SKS lulus, jumlah mata kuliah yang diulang, status cuti, dan semester aktif mahasiswa.
2. Seleksi dan Pembersihan Data (Preprocessing)
Data yang terkumpul diseleksi untuk memastikan kelengkapan dan konsistensi atribut. Data yang tidak lengkap, tidak valid, atau mengandung nilai kosong (missing value) dieliminasi.
3. Transformasi dan Normalisasi Data
Data akademik memiliki rentang nilai yang berbeda-beda, sehingga dilakukan normalisasi menggunakan metode Min-Max Normalization agar seluruh atribut berada pada skala yang sama, yaitu antara 0 sampai 1.
4. Penentuan Jumlah Cluster (k)
Penentuan jumlah cluster pada penelitian ini dilakukan menggunakan metode Elbow dengan menganalisis nilai jarak rata-rata (Within Cluster Sum of Squares/WCSS) untuk beberapa nilai k. Gambar Grafik Elbow menunjukkan hubungan antara jumlah cluster (k) dan nilai jarak rata-rata hasil clustering yang dapat dilihat pada gambar 1.



Gambar 1. Grafik Elbow Hubungan Cluster.

Berdasarkan grafik, terlihat bahwa penurunan nilai jarak rata-rata mulai melandai pada $k = 4$. Hal ini menunjukkan bahwa penambahan jumlah cluster setelah $k = 4$ tidak memberikan penurunan jarak yang signifikan. Oleh karena itu, jumlah cluster optimal yang digunakan dalam penelitian ini adalah sebanyak 4 cluster, karena telah mampu merepresentasikan pola data akademik mahasiswa secara

- efektif dengan tingkat kesalahan yang relatif kecil.
5. Proses Clustering Menggunakan Algoritma K-Means
Algoritma K-Means diterapkan dengan langkah-langkah sebagai berikut:
 - a. Menentukan centroid awal secara acak dari data yang telah dinormalisasi.
 - b. Menghitung jarak setiap data ke masing-masing centroid menggunakan rumus Euclidean Distance.
 - c. Mengelompokkan data ke cluster dengan jarak terdekat.
 - d. Menghitung ulang centroid baru berdasarkan rata-rata nilai setiap atribut pada masing-masing cluster.
 - e. Mengulangi proses hingga centroid tidak berubah atau iterasi maksimum tercapai.
 6. Penyimpanan Hasil Clustering
Hasil pengelompokan mahasiswa ke dalam cluster disimpan ke dalam basis data yang mencakup informasi cluster, jarak ke centroid, dan iterasi akhir.
 7. Analisis Karakteristik Cluster
Setiap cluster dianalisis dengan menghitung nilai rata-rata (mean) setiap atribut sehingga diperoleh centroid akhir yang merepresentasikan karakteristik umum mahasiswa dalam cluster tersebut. Selanjutnya, centroid diterjemahkan ke dalam kategori kualitatif (rendah, sedang, tinggi, dan sangat tinggi) untuk memudahkan interpretasi.
 8. Implementasi Sistem Berbasis Web
Seluruh proses clustering diimplementasikan dalam aplikasi berbasis web menggunakan framework Laravel. Sistem menyediakan fitur input data, proses clustering secara otomatis, serta visualisasi hasil dalam bentuk tabel dan grafik untuk mendukung analisis akademik.
 9. Evaluasi Hasil
Evaluasi dilakukan dengan menganalisis distribusi jumlah mahasiswa pada setiap cluster serta nilai rata-rata jarak ke centroid untuk menilai tingkat kekompakan cluster. Hasil evaluasi digunakan sebagai dasar interpretasi dan pembahasan penelitian.

2.3 Data Penelitian

Populasi dalam penelitian ini adalah mahasiswa Universitas Dipa Makassar angkatan 2019 sampai 2022, yang pernah tercatat sebagai mahasiswa bimbingan akademik peneliti dan memiliki data akademik lengkap selama masa studi. Sampel penelitian diambil dari populasi tersebut dengan teknik simple random sampling, yaitu pemilihan sampel secara acak dari daftar mahasiswa dengan jumlah sampel yang digunakan dalam penelitian ini sebanyak 100 mahasiswa. Variabel yang digunakan meliputi:

1. IPK (Indeks Prestasi Kumulatif)
2. Rata-rata nilai huruf mahasiswa, kemudian nilai huruf dikonversi ke angka menjadi: A = 4, B = 3, C = 2, D = 1, dan E = 0.
3. Total SKS dilulusi
4. jumlah mata kuliah yang diulang
5. Riwayat cuti akademik
6. Jumlah semester yang telah ditempuh.

Data tersebut digunakan sebagai dasar dalam proses analisis dan pengolahan data menggunakan algoritma K-Means. Data sampel dituangkan dalam bentuk tabel sebagai berikut..

Tabel 1. Data Sampel Mahasiswa

NIM	NAMA	IPK	Rata-rata Nilai	SKS Dilulusi	Pernah Cuti	Jumlah Semester	Jumlah MK Diulang
201198	ROBI LUKIN	3,88	2	109	1	7	5
201207	EFRAIM PRIGITO SOREAN	2,03	3,75	47	1	5	5
201217	ALVINO ALVARENZA	2,26	2	105	1	9	5
221042	SINTYA ANASTASYA MARGARETA	2,07	3,75	52	1	7	1
221043	NAJWA RIZQIYAH MUNIR	2,16	3	68	0	9	2
221057	ILHAM NURFAJRI	3,88	2,75	53	0	6	2
221058	ARTIKA YUDIA	3,3	3	146	1	7	5
221060	ASKA QUL NUR AISYAH	2,55	3,5	86	1	7	2

NIM	NAMA	IPK	Rata-rata Nilai	SKS Dilulusi	Pernah Cuti	Jumlah Semester	Jumlah MK Diulang
...				...		...	
....						...	
221061	AINUL HIKMAH	3,34	3	134	1	8	6
221062	TASYA AMALIA PUTRI	3,32	3,75	68	1	7	5
221063	I GEDE ARTA YASA	2,93	2	76	0	5	4
221064	DENDY QULBY	3,75	2	47	0	4	6

Tabel di atas menampilkan data sampel mahasiswa yang digunakan dalam penelitian ini. Data ini digunakan sebagai variabel pendukung dalam proses analisis penelitian. Sebelum dianalisis, data dibersihkan dari missing values dan outlier, kemudian dinormalisasi menggunakan metode Min-Max Normalization [7].

2.4 Metode Analisis

Metode analisis data yang digunakan dalam penelitian ini adalah algoritma K-Means Clustering, yang bertujuan untuk mengelompokkan mahasiswa berdasarkan kemiripan karakteristik akademik yang dimiliki [8]. Proses analisis dilakukan melalui beberapa tahapan sebagai berikut:

1. Pra-pemrosesan Data

Membersihkan dan menormalisasi data agar semua variabel memiliki skala yang sama. Metode yang digunakan adalah Min-Max Normalization. Setiap atribut ke dalam rentang 0–1 dengan menggunakan persamaan:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

IPK = $(3,88 - 2,03) / (4 - 2,03) = 1,85 / 1,97 = 0,9391$

Rata-rata Nilai = $(2 - 2) / (3,75 - 2) = 0 / 1,75 = 0,0000$

Jumlah SKS Dilulusi = $(109 - 40) / (150 - 40) = 69 / 110 = 0,6273$

Cuti Akademik = $(1 - 0) / (1 - 0) = 1 / 1 = 1,0000$

Jumlah Semester = $(7 - 0) / (11 - 0) = 7 / 11 = 0,6364$

Jumlah Mata Kuliah Diulang = $(5 - 0) / (6 - 0) = 5 / 6 = 0,8333$

Hasil normalisasi data disajikan dalam bentuk tabel sebagai berikut.

Tabel 2. Normalisasi Data

ID	IPK	Rata-rata Nilai	SKS	MK Diulang	Cuti Akademik	Semester
1	0,9391	0,0000	0,6273	0,8333	1,0000	0,4286
2	0,0000	1,0000	0,0636	0,8333	1,0000	0,1429
3	0,1168	0,0000	0,5909	0,8333	1,0000	0,7143
4	0,0203	1,0000	0,1091	0,1667	1,0000	0,4286
5	0,0660	0,5714	0,2545	0,3333	0,0000	0,7143
6	0,4822	0,4286	0,7545	0,3333	1,0000	0,1429
7	0,2132	0,2857	0,9364	0,5000	0,0000	0,4286
8	0,4873	0,0000	0,5091	0,8333	0,0000	0,5714
9	0,1827	0,4286	0,5182	0,1667	0,0000	0,7143
...						
....						
96	0,2640	0,8571	0,4182	0,3333	1,0000	0,4286
97	0,6650	0,5714	0,8545	1,0000	1,0000	0,5714
98	0,6548	1,0000	0,2545	0,8333	1,0000	0,4286
99	0,4569	0,0000	0,3273	0,6667	0,0000	0,1429
100	0,8731	0,0000	0,0636	1,0000	0,0000	0,0000

2. Analisis Cluster

Setelah proses clustering menggunakan algoritma K-Means selesai, tahap selanjutnya adalah melakukan analisis cluster dengan menghitung nilai rata-rata (mean) dari setiap atribut pada masing-masing cluster [9]. Perhitungan rata-rata ini bertujuan untuk menggambarkan karakteristik umum mahasiswa yang tergabung dalam setiap cluster. Berikut ini persamaan yang digunakan:

$$x_c = \frac{1}{n_c} \sum_{i=1}^{n_c} x_i \quad (2)$$

IPK C1 = $0,12 + 0,08 + 0,15 + 0,10 + 0,05 / 5$

IPK C1 = $0,50 / 5 = 0,10$

MK C1 = $0,83 + 0,67 + 1,00 + 0,83 + 0,67 / 5$

MK C1 = $4,00 / 5 = 0,80$

Dan seterusnya untuk setiap atribut yang digunakan. Berikut ini nilai rata-rata atribut setiap cluster yang disajikan dalam bentuk tabel:

Tabel 3. Rata-rata Atribut Tipa Cluster

Cluster	IPK	Rata-rata Nilai	SKS Dilulusi	MK Diulang	Cuti Akademik	Semester
C1	0,3149	0,7083	0,6648	0,5069	0,0000	0,5179
C2	0,5831	0,2095	0,4215	0,5222	0,0000	0,3333
C3	0,4718	0,4286	0,3795	0,2167	1,0000	0,4929
C4	0,6039	0,5330	0,4594	0,7821	1,0000	0,4560

3. Kategori Cluster

Label “aman”, “berisiko”, dan “sangat berisiko” yang digunakan dalam penelitian ini bersifat interpretatif dan diperoleh berdasarkan hasil analisis data menggunakan algoritma K-Means Clustering terhadap atribut akademik mahasiswa. Pemberian label tersebut dilakukan untuk memudahkan pemahaman karakteristik setiap cluster berdasarkan nilai centroid dan pola atribut yang terbentuk.

Label tersebut bukan merupakan diagnosis akademik resmi, melainkan hanya representasi hasil pengelompokan data secara statistik. Oleh karena itu, hasil clustering ini tidak dimaksudkan sebagai penentuan status akademik mahasiswa secara formal, melainkan sebagai alat bantu analisis bagi pihak akademik dalam mengidentifikasi kelompok mahasiswa yang memiliki karakteristik akademik tertentu sehingga dapat dijadikan dasar pertimbangan dalam pengambilan keputusan atau pemberian intervensi akademik.

Proses kategorisasi dilakukan dengan mengelompokkan nilai rata-rata setiap atribut ke dalam tingkat rendah, sedang, dan tinggi, serta pembagian khusus untuk atribut semester menjadi awal, tengah, dan akhir [10]. Atribut IPK, nilai mata kuliah inti, dan jumlah SKS lulus dikategorikan menjadi rendah, sedang, dan tinggi untuk menggambarkan capaian akademik mahasiswa. Untuk atribut semester aktif, kategorisasi dibagi menjadi tiga kelompok, yaitu semester awal, semester tengah, dan semester akhir. Pembagian ini bertujuan untuk menunjukkan posisi mahasiswa dalam masa studi serta keterkaitannya dengan capaian akademik dan risiko keterlambatan studi. Berdasarkan hasil kategorisasi tersebut, setiap cluster kemudian diberikan label evaluatif berupa label positif atau label negatif. Label positif diberikan kepada cluster yang menunjukkan kinerja akademik baik, yang ditandai dengan IPK sedang hingga tinggi, nilai mata kuliah inti yang baik, jumlah SKS lulus yang tinggi, frekuensi pengulangan mata kuliah yang rendah, serta riwayat cuti akademik yang minimal. Cluster dengan karakteristik ini mencerminkan mahasiswa yang berada dalam kondisi akademik relatif aman.

Sebaliknya, label negatif diberikan kepada cluster yang memiliki kinerja akademik rendah, ditandai dengan IPK rendah, nilai mata kuliah inti rendah, jumlah mata kuliah yang diulang tinggi, serta riwayat cuti akademik yang lebih sering. Cluster ini menunjukkan kelompok mahasiswa dengan tingkat risiko akademik yang lebih tinggi dan memerlukan perhatian serta pendampingan akademik lebih intensif. Dengan adanya kategorisasi dan pemberian label pada setiap cluster, hasil clustering K-Means menjadi lebih mudah dipahami dan dapat dimanfaatkan sebagai dasar pengambilan keputusan oleh dosen penasehat akademik dalam merancang strategi pembinaan dan intervensi akademik yang tepat sasaran.

3. HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan suatu sistem yang mampu mengelompokkan data mahasiswa menggunakan algoritma K-Means dengan jumlah cluster sebanyak empat ($k=4$). Proses clustering dilakukan terhadap data akademik mahasiswa yang telah melalui tahapan praproses data, meliputi pembersihan data, transformasi atribut, serta normalisasi menggunakan metode Min–Max normalization. Atribut yang digunakan dalam proses ini terdiri atas IPK, nilai mata kuliah, jumlah SKS lulus, jumlah mata kuliah diulang, status cuti, dan semester aktif.

Hasil pengelompokan menunjukkan bahwa seluruh data mahasiswa berhasil diklasifikasikan ke

dalam empat cluster tanpa adanya data yang tidak terkelompokkan. Ringkasan hasil clustering yang mencakup jumlah mahasiswa dan nilai rata-rata jarak terhadap centroid pada setiap cluster disajikan pada gambar 1.

Cluster	Jumlah	Rata-Rata Jarak
C1	24	0.5516
C2	30	0.5929
C3	20	0.5795
C4	26	0.5720

Gambar 1. Output Ringkasan Hasil Cluster

Berdasarkan gambar 1, terlihat bahwa Cluster C2 memiliki jumlah anggota terbanyak, sedangkan Cluster C3 memiliki jumlah anggota paling sedikit. Nilai rata-rata jarak antar data terhadap centroid pada masing-masing cluster relatif seimbang, yang menunjukkan bahwa proses clustering telah mencapai kondisi konvergen dan pembentukan cluster berlangsung secara optimal. Ringkasan cluster diinterpretasikan menggunakan OpenAI dapat dilihat pada gambar 2.

Hasil dari proses K-Means Clustering menunjukkan adanya empat cluster yang mengindikasikan tingkat risiko permasalahan akademik mahasiswa di kampus Undipa Makassar. Cluster 1, yang terdiri dari 24 mahasiswa, berada dalam kategori sangat berisiko dengan rata-rata jarak 0.55, sementara cluster 2 yang berisiko mencakup 30 mahasiswa dengan rata-rata jarak 0.59. Di sisi lain, cluster 3 yang cukup aman memiliki 20 mahasiswa dengan rata-rata jarak 0.58, dan cluster 4 yang aman berisi 26 mahasiswa dengan rata-rata jarak 0.57. Penempatan mahasiswa dalam kategori risiko ini mengindikasikan perlunya perhatian khusus untuk kelompok dalam cluster 1 dan 2 guna mencegah masalah akademik lebih lanjut, sedangkan mahasiswa dalam cluster 3 dan 4 dapat dianggap relatif aman.

Gambar 2. Output Interpretasi Ringkasan Hasil Cluster

Nilai centroid merepresentasikan rata-rata atribut dari seluruh anggota dalam satu cluster dan digunakan untuk menggambarkan karakteristik umum mahasiswa pada cluster tersebut [11]. Nilai centroid diperoleh dari hasil perhitungan terhadap data yang telah dinormalisasi. Nilai rata-rata centroid untuk setiap cluster ditampilkan pada gambar 3.

Cluster	IPK	Rata-Rata Nilai	SKS Dilulusi	MK Diulang	Cuti Akademik	Jml SMSTR
C1	0.3149	0.7083	0.6648	0.5069	0.0000	0.5179
C2	0.5831	0.2095	0.4215	0.5222	0.0000	0.3333
C3	0.4718	0.4285	0.3795	0.2167	1.0000	0.4929
C4	0.6039	0.5330	0.4594	0.7821	1.0000	0.4560

Gambar 3. Output Rata-rata Nilai Centroid Tiap Cluster

Hasil tersebut menunjukkan adanya perbedaan pola akademik yang cukup jelas antar cluster, khususnya pada atribut nilai mata kuliah, jumlah mata kuliah diulang, dan status cuti. Perbedaan ini mengindikasikan bahwa algoritma K-Means mampu mengidentifikasi karakteristik akademik mahasiswa secara efektif. Interpretasi nilai centroid tiap cluster dapat dilihat pada gambar 4.

Hasil analisis menggunakan K-Means Clustering menunjukkan bahwa mahasiswa di Undipa Makassar terdistribusi dalam empat cluster yang mencerminkan tingkat risiko permasalahan akademik. Cluster 1, yang teridentifikasi sebagai Sangat Berisiko, ditandai dengan IPK rendah, nilai inti tinggi, dan SKS tinggi, namun mengalami masalah dalam mata kuliah, cuti, dan semester yang sedang. Cluster 2, berisiko, menunjukkan IPK dan nilai inti dalam kategori sedang dengan SKS sedang pula, sembari mengalami tantangan dalam mata kuliah dan cuti yang rendah. Cluster 3, cukup aman, memiliki IPK, nilai inti, dan SKS sedang, dengan isu mata kuliah yang rendah serta cuti dan semester yang tinggi, sedangkan Cluster 4, yang dikategorikan aman, menunjukkan IPK, nilai inti, dan SKS dalam kategori sedang dengan indikasi mata kuliah yang tinggi dan cuti serta semester yang tinggi. Cluster-cluster ini memungkinkan identifikasi dini untuk menawarkan intervensi yang sesuai terhadap mahasiswa dengan potensi permasalahan akademik.

Gambar 4. Output Interpretasi Rata-rata Nilai Centroid Tiap Cluster

Untuk mempermudah interpretasi hasil clustering, nilai centroid kemudian dikonversi ke dalam kategori kualitatif, yaitu rendah, sedang, dan tinggi. Kategorisasi ini bertujuan untuk memberikan gambaran kondisi akademik mahasiswa yang lebih mudah dipahami oleh pengguna sistem, khususnya pihak akademik. Hasil kategorisasi karakteristik akademik setiap cluster disajikan pada gambar 5.

Cluster	IPK	Rata-Rata Nilai	SKS Dilulusi	MK Diulang	Cuti Akademik	Jml SMSTR
C1	Rendah	Tinggi	Tinggi	Sedang	Rendah	Sedang
C2	Sedang	Rendah	Sedang	Sedang	Rendah	Sedang
C3	Sedang	Sedang	Sedang	Rendah	Tinggi	Sedang
C4	Sedang	Sedang	Sedang	Tinggi	Tinggi	Sedang

Gambar 5. Output Kategori Risiko Akademik Mahasiswa Berdasarkan Cluster

Berdasarkan hasil tersebut, setiap cluster memiliki karakteristik akademik yang berbeda. Cluster C1 menunjukkan capaian akademik yang relatif baik pada nilai inti dan jumlah SKS, meskipun memiliki IPK yang rendah. Cluster C2 ditandai oleh rendahnya nilai mata kuliah inti. Cluster C3 dan C4 memiliki tingkat cuti yang tinggi, dengan Cluster C4 juga ditandai oleh tingginya jumlah mata kuliah yang diulang. Interpretasi hasil risiko akademik mahasiswa berdasarkan cluster dapat dilihat pada gambar 6.

Hasil penelitian menunjukkan bahwa mahasiswa di Kampus Undipa Makassar terbagi ke dalam empat kategori risiko akademik berdasarkan cluster K-Means. Cluster 1, yang berstatus sangat berisiko, ditandai oleh mahasiswa dengan IPK rendah namun nilai inti dan SKS yang tinggi, serta kondisi mata kuliah, cuti, dan semester dalam kategori sedang atau rendah, mengindikasikan potensi masalah serius meskipun terdapat beberapa nilai positif. Cluster 2, berisiko, menunjukkan mahasiswa dengan IPK sedang, nilai inti rendah, dan SKS sedang, diiringi faktor negatif seperti kondisi mata kuliah dalam kategori sedang dan cuti rendah, menandakan ada tantangan yang perlu diatasi. Cluster 3, cukup aman, menggambarkan mahasiswa dengan IPK dan nilai inti serta SKS sedang, namun dengan potensi masalah pada mata kuliah yang rendah dan cuti tinggi, menggambarkan kebutuhan perhatian lebih. Sementara itu, cluster 4, yang aman, menunjukkan mahasiswa dengan kondisi IPK, nilai inti, dan SKS yang sedang, serta faktor negatif yang membaik, memberikan indikasi bahwa mereka berada dalam jalur yang lebih stabil.

Gambar 6. Output Interpretasi Hasil Kategori Risiko Akademik Mahasiswa

Sebagai bentuk tindak lanjut dari hasil clustering, setiap cluster diberikan label risiko akademik yang diklasifikasikan ke dalam kategori positif dan negatif. Penentuan label ini didasarkan pada kombinasi atribut akademik yang mencerminkan kondisi mahasiswa dalam masing-masing cluster dapat dilihat pada gambar 7.

Cluster	Karakteristik Dominan	Label
C1	Nilai inti dan SKS tinggi	Positif
C2	Nilai inti rendah	Negatif
C3	Tingkat cuti tinggi	Negatif
C4	MK diulang dan cuti tinggi	Negatif

Gambar 7. Output Label Risiko Akademik Mahasiswa Setiap Cluster

Cluster dengan label negatif menunjukkan kelompok mahasiswa yang berpotensi mengalami permasalahan akademik dan memerlukan perhatian serta intervensi lebih lanjut dari pihak akademik.

Untuk mengevaluasi kualitas hasil clustering, penelitian ini menggunakan Silhouette Coefficient sebagai metrik evaluasi internal. Silhouette Coefficient mengukur tingkat kesesuaian suatu data terhadap cluster tempat ia berada dibandingkan dengan cluster lain di sekitarnya. Nilai Silhouette berada pada rentang -1 hingga 1, di mana nilai mendekati 1 menunjukkan bahwa data berada pada cluster yang sesuai, nilai mendekati 0 menunjukkan adanya tumpang tindih antar cluster, dan nilai negatif menunjukkan kemungkinan kesalahan pengelompokan. Hasil evaluasi cluster dapat dilihat pada tabel 4.

Tabel 4. Hasil Evaluasi Cluster Silhouette Coefficient

Data ke-i	a(i)	b(i)	s(i)
1	0,42	0,71	0,55
2	0,38	0,70	0,56
3	0,40	0,75	0,57

Data ke-i	a(i)	b(i)	s(i)
4	0,36	0,68	0,55
5	0,41	0,74	0,56
6	0,39	0,72	0,55
7	0,37	0,69	0,56
8	0,43	0,76	0,57
9	0,35	0,67	0,55
10	0,40	0,73	0,56
...	...	...	...
...	...	...	...
96	0,40	0,73	0,56
97	0,39	0,72	0,57
98	0,36	0,68	0,55
100	0,43	0,76	0,57
Rata-rata			0,56

Kriteria umum evaluasi cluster Silhouette Coefficient yang digunakan $0,71 - 1,00 =$ Sangat Baik, $0,51 - 0,70 =$ Baik, $0,26 - 0,50 =$ Cukup, $\leq 0,25 =$ Lemah. Berdasarkan hasil perhitungan, diperoleh nilai Silhouette Coefficient rata-rata sebesar 0,56, yang menunjukkan bahwa struktur cluster yang terbentuk berada pada kategori Baik dan memiliki tingkat pemisahan antar cluster yang memadai. Dengan demikian, penentuan jumlah cluster $k=4$ tidak hanya didasarkan pada metode Elbow, tetapi juga diperkuat melalui evaluasi kualitas cluster menggunakan Silhouette Coefficient.

Berbeda dengan penelitian-penelitian sebelumnya yang umumnya hanya berfokus pada penerapan algoritma K-Means untuk mengelompokkan data akademik mahasiswa secara statistik, penelitian ini menekankan integrasi antara proses analisis clustering dengan implementasi sistem berbasis web. Sistem dikembangkan menggunakan framework Laravel sehingga memungkinkan proses pengolahan data, visualisasi hasil clustering, serta analisis karakteristik setiap cluster dilakukan secara otomatis dan dapat diakses melalui browser. Selain itu, penelitian ini tidak hanya menghasilkan kelompok mahasiswa, tetapi juga melakukan interpretasi karakteristik cluster berdasarkan atribut IPK, nilai mata kuliah inti, jumlah SKS, mata kuliah diulang, cuti, dan semester. Dengan demikian, kontribusi utama penelitian ini terletak pada penyediaan sistem pendukung analisis akademik berbasis web yang menggabungkan metode K-Means dengan penyajian hasil yang informatif dan mudah digunakan sebagai dasar pengambilan keputusan dalam pembinaan akademik mahasiswa.

Berdasarkan hasil penelitian, dapat disimpulkan bahwa algoritma K-Means mampu mengelompokkan mahasiswa ke dalam beberapa cluster dengan karakteristik akademik yang berbeda secara efektif. Implementasi sistem clustering berbasis web menggunakan Laravel memungkinkan proses analisis data dilakukan secara otomatis, efisien, dan mudah diakses. Dengan adanya pengelompokan ini, pihak akademik dapat merancang strategi pembinaan dan pendampingan yang lebih tepat sasaran sesuai dengan karakteristik masing-masing cluster.

4. KESIMPULAN

Penelitian ini berhasil menerapkan algoritma K-Means untuk mengelompokkan mahasiswa berdasarkan enam atribut akademik utama, yaitu IPK, nilai mata kuliah inti, jumlah SKS lulus, jumlah mata kuliah diulang, status cuti, dan semester aktif. Dengan jumlah cluster sebanyak empat, sistem mampu menghasilkan pengelompokan mahasiswa yang memiliki karakteristik akademik yang berbeda dan dapat diinterpretasikan secara jelas, sehingga mendukung pemetaan kondisi akademik mahasiswa secara objektif serta memungkinkan identifikasi dini terhadap kelompok mahasiswa yang berpotensi mengalami permasalahan akademik. Implementasi sistem berbasis web menggunakan framework Laravel terbukti efektif dalam mengelola proses normalisasi data, perhitungan clustering, serta penyajian hasil analisis secara interaktif. Sistem yang dibangun dapat dimanfaatkan sebagai alat bantu pengambilan keputusan bagi pihak akademik dalam melakukan evaluasi dan pemantauan kinerja mahasiswa secara berkelanjutan, sehingga berkontribusi pada peningkatan kualitas pengelolaan akademik berbasis data.

5. SARAN

Berdasarkan hasil penelitian yang telah dilakukan, disarankan agar sistem clustering mahasiswa ini dikembangkan lebih lanjut dengan menambahkan variabel akademik maupun non-akademik, seperti kehadiran, aktivitas organisasi, atau riwayat beasiswa, sehingga hasil pengelompokan dapat merepresentasikan kondisi mahasiswa secara lebih komprehensif. Selain itu, penerapan metode clustering lain, seperti K-Medoids, Hierarchical Clustering, atau pendekatan berbasis machine learning supervised,

dapat digunakan sebagai pembanding untuk meningkatkan akurasi dan validitas hasil pengelompokan.

Pengembangan selanjutnya juga dapat diarahkan pada integrasi sistem dengan Sistem Informasi Akademik (SIKA) agar proses analisis dapat dilakukan secara otomatis dan berkelanjutan. Selain itu, hasil clustering dapat dimanfaatkan sebagai dasar penyusunan kebijakan akademik, seperti program pembinaan, pendampingan akademik, atau peringatan dini bagi mahasiswa berisiko, sehingga sistem ini tidak hanya bersifat informatif, tetapi juga aplikatif dalam mendukung pengambilan keputusan di lingkungan perguruan tinggi.

DAFTAR PUSTAKA

- [1] T. Simanjuntak, J. A. Gulo, S. P. Sipayung, and H. P. Hia, "Pengelompokan mahasiswa berdasarkan nilai dan kehadiran menggunakan K-means," *Jurnal Minfo Polgan*, vol. 14, no. 2, pp. 3525–3532, 2026.
- [2] W. Hulu, R. Rosalinda, R. O. Angelie, and S. Sipayung, "Analisis nilai dan kehadiran mahasiswa untuk pengembangan pengetahuan akademik menggunakan clustering K-means dan statistik," *Jurnal Minfo Polgan*, vol. 14, no. 2, pp. 3516–3524, 2026.
- [3] A. Anas, A. J. Zebua, and Akhmadi, "Clustering nilai mahasiswa menggunakan algoritma K-means," *Jurnal Ilmiah Media Sisfo*, vol. 19, no. 1, pp. 7–14, 2025.
- [4] R. Ishak and A. Bengnga, "Clustering prestasi akademik lulusan menggunakan metode K-means," *Jambura Journal of Electrical and Electronics Engineering*, 2025.
- [5] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Waltham, MA, USA: Morgan Kaufmann, 2012.
- [6] D. Wahyudi and K. Handoko, "Penerapan algoritma K-means untuk pengelompokan prestasi akademik mahasiswa," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, vol. 5, no. 4, pp. 433–440, 2018.
- [7] R. Hidayat and A. Nugroho, "Penerapan metode K-means clustering untuk pengelompokan mahasiswa berdasarkan prestasi akademik," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, vol. 5, no. 2, pp. 211–218, 2018.
- [8] D. R. Sari and A. Wibowo, "Analisis pengelompokan mahasiswa berpotensi bermasalah akademik menggunakan algoritma K-means," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 3, no. 1, pp. 38–45, 2019.
- [9] E. Rahmawati and S. H. Pramono, "Analisis hasil clustering menggunakan metode K-means untuk pengelompokan data akademik mahasiswa," *Jurnal Informatika*, vol. 10, no. 2, pp. 120–128, 2016.
- [10] E. Sutoyo and H. Al Fatta, "Analisis karakteristik cluster hasil K-means menggunakan nilai centroid untuk pengambilan keputusan," *Jurnal Teknologi Informasi dan Komunikasi*, vol. 7, no. 1, pp. 15–22, 2018.
- [11] F. Rahmadani and A. Setiawan, "Analisis centroid pada algoritma K-means untuk mengetahui karakteristik kelompok data akademik mahasiswa," *Jurnal Teknologi Informasi*, vol. 13, no. 2, pp. 85–92, 2019.