



Analisis Pola Belajar Mahasiswa Menggunakan Data Science untuk Meningkatkan Efektivitas Pembelajaran

Muhammad Nasrullah¹, Mul Muliadi²

¹Teknologo Informasi, Institut Teknologi dan Kesehatan Aspirasi, Indonesia (penulis 1)

²Pendidikan Bahasa Inggris, Institut Pendidikan Nusantara Global, Indonesia (penulis 2)

Email : muhammadnasrullahsib@gmail.com¹, muliadimul2018@gmail.com²

ABSTRAK

Perkembangan teknologi pembelajaran digital menghasilkan berbagai data aktivitas mahasiswa yang dapat dimanfaatkan untuk memahami pola belajar dan mendukung peningkatan efektivitas pembelajaran. Penelitian ini bertujuan untuk menganalisis pola belajar mahasiswa menggunakan pendekatan data science serta mengidentifikasi karakteristik kelompok mahasiswa berdasarkan aktivitas pembelajaran. Penelitian menggunakan pendekatan kuantitatif dengan metode data mining. Data yang digunakan meliputi frekuensi akses Learning Management System (LMS), durasi belajar, kehadiran, penyelesaian tugas, aktivitas forum, nilai kuis, dan nilai evaluasi. Tahapan analisis meliputi data preprocessing, transformasi data, eksplorasi data, normalisasi, clustering menggunakan algoritma K-Means, serta interpretasi karakteristik setiap cluster. Berdasarkan data ilustratif sebanyak 200 mahasiswa, diperoleh tiga pola belajar, yaitu kelompok belajar aktif dan konsisten, kelompok belajar cukup aktif, serta kelompok belajar dengan aktivitas rendah. Kelompok aktif memiliki tingkat akses LMS, penyelesaian tugas, kehadiran, dan capaian akademik yang relatif tinggi. Kelompok dengan aktivitas rendah menunjukkan aktivitas pembelajaran dan penyelesaian tugas yang lebih rendah. Hasil analisis menunjukkan bahwa pendekatan data science dapat digunakan untuk memetakan karakteristik pola belajar mahasiswa sehingga dapat menjadi dasar bagi dosen dalam menentukan strategi pembelajaran dan pendampingan yang sesuai dengan kebutuhan mahasiswa.

Kata Kunci: Data Science, Learning Analytics, Pola Belajar Mahasiswa, K-Means, Data Mining

ABSTRACT

The development of digital learning technologies generates various student activity data that can be utilized to understand learning patterns and improve learning effectiveness. This study aims to analyze student learning patterns using a data science approach and identify student groups based on learning activities. The study employs a quantitative approach using data mining methods. The data include Learning Management System (LMS) access frequency, learning duration, attendance, assignment completion, forum activity, quiz scores, and evaluation scores. The analysis stages consist of data preprocessing, data transformation, exploratory data analysis, normalization, clustering using the K-Means algorithm, and interpretation of cluster characteristics. Based on an illustrative dataset of 200 students, three learning patterns were identified: active and consistent learners, moderately active learners, and learners with low learning activity. The active group demonstrated relatively high LMS access, assignment completion, attendance, and academic achievement, while the low-activity group showed lower learning engagement and assignment completion. The findings indicate that a data science approach can be used to map student learning patterns and provide a basis for lecturers to develop learning strategies and student support according to students' characteristics.

Keywords: Data Science, Learning Analytics, Student Learning Patterns, K-Means, Data Mining.

*Corresponding author: muhammadnasrullahsib@gmail.com

© 2025 TAMPIASIH

PENDAHULUAN

Bagian Perkembangan teknologi digital dalam pendidikan tinggi menyebabkan aktivitas pembelajaran menghasilkan data dalam jumlah yang semakin besar. Penggunaan Learning Management System (LMS), sistem informasi akademik, platform pembelajaran daring, forum diskusi, kuis, dan berbagai media digital menghasilkan rekam jejak aktivitas mahasiswa. Data tersebut dapat digunakan untuk memahami keterlibatan mahasiswa, kebiasaan belajar, serta hubungan antara aktivitas pembelajaran dengan capaian akademik. Pemanfaatan data pembelajaran melalui learning analytics memberikan peluang bagi perguruan tinggi untuk memperoleh informasi yang lebih terukur mengenai proses belajar mahasiswa [1], [2]. Learning analytics merupakan pendekatan yang memanfaatkan data yang dihasilkan selama proses pembelajaran untuk memahami dan meningkatkan proses belajar. Dalam pendidikan tinggi, learning analytics telah digunakan untuk memantau aktivitas mahasiswa, mengidentifikasi pola keterlibatan, mendukung keberhasilan studi, dan memberikan informasi bagi pengambilan keputusan pembelajaran. Kajian sistematis menunjukkan bahwa penggunaan data aktivitas mahasiswa dapat memberikan gambaran mengenai perilaku pembelajaran yang sebelumnya sulit diamati secara langsung [1], [3].

Pola belajar mahasiswa merupakan salah satu aspek penting yang perlu diperhatikan dalam proses pembelajaran. Mahasiswa memiliki karakteristik yang berbeda dalam mengakses materi, menyelesaikan tugas, mengikuti perkuliahan, berinteraksi dalam forum, dan melakukan aktivitas belajar mandiri. Perbedaan tersebut dapat menghasilkan tingkat keterlibatan dan capaian pembelajaran yang berbeda. Penelitian mengenai student engagement menunjukkan bahwa data aktivitas digital banyak digunakan untuk mengidentifikasi keterlibatan mahasiswa, meskipun indikator perilaku seperti klik dan durasi aktivitas belum selalu mampu menggambarkan seluruh dimensi keterlibatan mahasiswa [4]. Selain keterlibatan mahasiswa, data aktivitas pembelajaran juga dapat digunakan untuk memahami kemampuan mahasiswa dalam mengatur proses belajarnya. Self-regulated learning berkaitan dengan kemampuan mahasiswa dalam mengatur waktu, mengelola aktivitas belajar, menyelesaikan tugas, dan mempertahankan keterlibatan selama proses pembelajaran. Penelitian sistematis menunjukkan bahwa learning analytics dapat digunakan untuk memperoleh indikator self-regulated learning melalui data jejak digital yang dihasilkan mahasiswa pada lingkungan pembelajaran daring [5].

Educational Data Mining (EDM) merupakan salah satu bidang yang dapat digunakan untuk mengolah data pendidikan secara sistematis. Teknik data mining memungkinkan peneliti menemukan pola tersembunyi, hubungan antarvariabel, melakukan klasifikasi, prediksi, dan clustering terhadap data pendidikan. Penelitian Yağcı menunjukkan bahwa algoritma machine learning dapat digunakan untuk menganalisis data pendidikan dan memprediksi performa akademik mahasiswa [6]. Pemanfaatan data science dalam pendidikan tidak terbatas pada prediksi nilai akademik. Data science juga dapat digunakan untuk melakukan segmentasi mahasiswa berdasarkan karakteristik aktivitas belajar. Dengan clustering, mahasiswa dapat dikelompokkan berdasarkan kemiripan karakteristik tanpa harus menentukan label kelompok terlebih dahulu. Pendekatan tersebut dapat membantu mengidentifikasi kelompok mahasiswa dengan karakteristik aktivitas belajar tertentu sehingga informasi yang dihasilkan lebih bersifat eksploratif [7], [8]. Penelitian yang menerapkan K-Means terhadap perilaku belajar mahasiswa juga menunjukkan bahwa clustering dapat digunakan untuk mengidentifikasi kelompok mahasiswa berdasarkan karakteristik akademik dan aktivitas pembelajaran.

K-Means merupakan salah satu algoritma clustering yang banyak digunakan karena prosesnya relatif sederhana dan dapat digunakan untuk mengelompokkan data berdasarkan

kemiripan karakteristik. Dalam konteks pembelajaran, K-Means dapat digunakan untuk mengelompokkan mahasiswa berdasarkan frekuensi akses LMS, durasi belajar, kehadiran, penyelesaian tugas, aktivitas diskusi, dan capaian akademik. Penggunaan beberapa indikator sekaligus dapat menghasilkan profil kelompok yang lebih komprehensif dibandingkan analisis berdasarkan satu indikator [7], [9]. Penelitian mengenai learning analytics pada pendidikan tinggi menunjukkan bahwa sumber data yang digunakan dapat berasal dari log sistem, data asesmen, survei, dan sumber data pembelajaran lainnya. Kajian terhadap 159 penelitian student engagement dalam learning analytics menemukan bahwa data log dan data asesmen merupakan sumber data yang sering digunakan untuk memahami keterlibatan mahasiswa. Temuan tersebut menunjukkan bahwa kombinasi beberapa sumber data dapat memberikan perspektif yang lebih luas dalam menganalisis aktivitas pembelajaran [4].

Penelitian lain menunjukkan bahwa learning analytics telah digunakan secara luas dalam pendidikan jarak jauh dan pembelajaran daring. Kajian terhadap 400 artikel menunjukkan perkembangan penelitian learning analytics dalam pendidikan jarak jauh serta penggunaan berbagai sumber data dan pendekatan analisis. Perkembangan tersebut memperlihatkan bahwa data aktivitas pembelajaran memiliki potensi untuk digunakan dalam memahami proses belajar dan memberikan dukungan kepada mahasiswa [3]. Dalam praktiknya, hasil analisis data mahasiswa perlu diterjemahkan menjadi informasi yang mudah dipahami oleh dosen. Learning analytics dashboard menjadi salah satu bentuk penerapan yang memungkinkan hasil analisis ditampilkan dalam bentuk visual dan informasi ringkas. Kajian sistematis mengenai dashboard learning analytics menunjukkan adanya perkembangan dari dashboard yang hanya menampilkan statistik menuju dashboard yang lebih berorientasi pada proses belajar dan pemanfaatan informasi untuk mendukung pembelajaran [10]. Pemanfaatan data aktivitas mahasiswa juga dapat digunakan untuk mendukung self-regulated learning. Informasi mengenai keteraturan akses, penyelesaian aktivitas, dan pola penggunaan sistem dapat memberikan gambaran mengenai kebiasaan mahasiswa dalam mengelola proses belajar. Namun, data digital tidak dapat langsung dianggap sebagai representasi sempurna dari proses kognitif mahasiswa sehingga diperlukan interpretasi yang hati-hati terhadap indikator yang digunakan [5], [11].

Penelitian terdahulu juga menunjukkan bahwa machine learning dapat digunakan untuk melakukan prediksi performa mahasiswa. Penggunaan berbagai algoritma machine learning dapat membantu institusi pendidikan mengidentifikasi karakteristik yang berkaitan dengan capaian akademik. Meskipun demikian, pendekatan prediktif umumnya berfokus pada hasil akhir seperti nilai atau risiko akademik, sedangkan analisis clustering dapat digunakan untuk mengeksplorasi kelompok mahasiswa berdasarkan pola aktivitas yang memiliki karakteristik serupa [6], [12]. Berdasarkan penelitian sebelumnya, terdapat peluang untuk mengembangkan analisis yang tidak hanya berfokus pada prediksi performa mahasiswa, tetapi juga pada identifikasi kelompok pola belajar. Student engagement bersifat multidimensional sehingga penggunaan satu indikator belum cukup untuk menggambarkan karakteristik mahasiswa. Oleh sebab itu, diperlukan pendekatan yang menggabungkan indikator aktivitas digital, aktivitas akademik, dan capaian pembelajaran ke dalam suatu analisis pola belajar [4], [5].

Permasalahan penelitian ini adalah belum optimalnya pemanfaatan data aktivitas pembelajaran mahasiswa untuk menghasilkan pemetaan pola belajar yang dapat dipahami dan dimanfaatkan oleh dosen. Data LMS dan data akademik sering kali hanya digunakan sebagai catatan administratif atau untuk menghasilkan nilai akhir. Padahal, kombinasi data tersebut dapat digunakan untuk menemukan kelompok mahasiswa berdasarkan karakteristik aktivitas belajar [3], [6]. Kebaruan penelitian ini terletak pada penerapan pendekatan data science untuk melakukan segmentasi pola belajar mahasiswa berdasarkan kombinasi beberapa indikator pembelajaran. Variabel yang dianalisis meliputi frekuensi akses LMS, durasi belajar, kehadiran, penyelesaian

tugas, aktivitas forum, nilai kuis, dan nilai evaluasi. Berbeda dengan penelitian yang berorientasi pada prediksi nilai, penelitian ini menggunakan clustering untuk menemukan kelompok mahasiswa berdasarkan karakteristik aktivitas belajar yang terbentuk secara empiris [7], [9].

Pendekatan tersebut diharapkan menghasilkan profil pola belajar yang dapat digunakan untuk mendukung pengambilan keputusan pembelajaran. Dosen dapat memperoleh informasi mengenai kelompok mahasiswa yang memiliki aktivitas tinggi, aktivitas sedang, maupun aktivitas rendah. Hasil analisis juga dapat dikembangkan menjadi dasar pemberian pengayaan, pendampingan, monitoring, dan strategi pembelajaran yang lebih sesuai dengan karakteristik mahasiswa [10], [13]. Aspek etika dan privasi juga perlu diperhatikan dalam penggunaan data mahasiswa. Data aktivitas pembelajaran merupakan data yang berkaitan dengan individu sehingga penggunaannya harus mempertimbangkan keamanan data, pembatasan akses, tujuan penggunaan, dan interpretasi hasil. Analisis data mahasiswa sebaiknya digunakan untuk mendukung pembelajaran dan bukan untuk memberikan label permanen terhadap mahasiswa [14], [15]. Berdasarkan uraian tersebut, rumusan masalah penelitian adalah: (1) bagaimana karakteristik pola belajar mahasiswa berdasarkan aktivitas pembelajaran digital dan akademik; (2) bagaimana algoritma K-Means dapat digunakan untuk mengelompokkan mahasiswa berdasarkan pola belajar; dan (3) bagaimana hasil clustering dapat dimanfaatkan untuk mendukung peningkatan efektivitas pembelajaran.

Tujuan penelitian ini adalah menganalisis pola belajar mahasiswa menggunakan pendekatan data science, mengelompokkan mahasiswa berdasarkan karakteristik aktivitas pembelajaran menggunakan algoritma K-Means, serta menghasilkan informasi yang dapat digunakan sebagai dasar dalam menentukan strategi pembelajaran dan pendampingan mahasiswa.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif eksploratif dengan metode data science dan educational data mining. Pendekatan tersebut digunakan untuk menemukan pola dalam data aktivitas pembelajaran mahasiswa melalui proses pengolahan, analisis, dan interpretasi data. Educational data mining telah digunakan untuk menemukan hubungan dan pola pada data pendidikan serta menganalisis performa akademik mahasiswa [6], [16].

Tahapan penelitian terdiri atas pengumpulan data, data preprocessing, eksplorasi data, transformasi data, normalisasi, clustering, evaluasi, dan interpretasi. Tahapan tersebut digunakan untuk memastikan bahwa data yang diperoleh dapat diproses secara sistematis sebelum diterapkan pada algoritma clustering. Penggunaan tahapan preprocessing dan eksplorasi merupakan bagian penting dalam analisis data pendidikan karena data aktivitas mahasiswa dapat berasal dari berbagai sumber dan memiliki karakteristik yang berbeda [3], [7].

Sumber Data

Data penelitian berasal dari aktivitas pembelajaran mahasiswa pada lingkungan pembelajaran digital dan data akademik. Data aktivitas dapat diperoleh dari LMS, sedangkan data akademik diperoleh dari sistem akademik atau dokumen pembelajaran yang relevan. Penggunaan log data LMS bersama data asesmen telah banyak digunakan dalam penelitian learning analytics untuk menganalisis student engagement dan performa pembelajaran [4]. Variabel penelitian terdiri atas frekuensi akses LMS, durasi belajar, kehadiran, penyelesaian tugas, aktivitas forum, nilai kuis, dan nilai evaluasi. Variabel tersebut dipilih untuk merepresentasikan beberapa aspek aktivitas pembelajaran mahasiswa. Penggunaan indikator perilaku dari data jejak digital merupakan pendekatan yang umum dalam penelitian learning analytics, meskipun interpretasinya tetap perlu dikaitkan dengan konteks pembelajaran [4], [5].

Tabel 1. Variabel Penelitian

No.	Variabel	Keterangan	Satuan
1	Frekuensi LMS	Jumlah akses mahasiswa ke LMS	Kali
2	Durasi belajar	Total waktu aktivitas pembelajaran	Menit/minggu
3	Kehadiran	Persentase kehadiran mahasiswa	%
4	Penyelesaian tugas	Persentase tugas yang diselesaikan	%
5	Aktivitas forum	Jumlah aktivitas diskusi	Kali
6	Nilai kuis	Rata-rata nilai kuis	0–100
7	Nilai evaluasi	Nilai evaluasi pembelajaran	0–100

Populasi dan Sampel

Populasi penelitian adalah mahasiswa yang mengikuti pembelajaran dengan dukungan LMS pada periode penelitian. Pada rancangan analisis ini digunakan 200 data mahasiswa sebagai data ilustratif. Pada implementasi penelitian sebenarnya, jumlah data disesuaikan dengan dataset aktual yang tersedia.

Kriteria data yang digunakan adalah mahasiswa yang memiliki data aktivitas pembelajaran dan data akademik yang memadai selama periode pengamatan. Pemilihan data dengan indikator yang relatif lengkap diperlukan agar proses clustering tidak terganggu oleh missing value yang terlalu tinggi. Penggunaan data log bersama data asesmen merupakan pendekatan yang banyak ditemukan dalam penelitian learning analytics pendidikan tinggi [4], [7].

Data Preprocessing

Data preprocessing dilakukan sebelum proses clustering. Tahap ini meliputi pemeriksaan data kosong, data duplikat, data tidak valid, pemeriksaan outlier, transformasi data, dan normalisasi. Preprocessing diperlukan karena variabel penelitian memiliki satuan yang berbeda, misalnya frekuensi akses menggunakan satuan kali, durasi menggunakan menit, sedangkan nilai menggunakan skala 0–100.

Normalisasi dilakukan agar variabel dengan rentang nilai yang besar tidak mendominasi proses perhitungan jarak. Tahapan pengolahan data seperti pembersihan, transformasi, dan standardisasi diperlukan dalam penerapan data mining pada data pendidikan agar hasil analisis lebih konsisten [6], [16].

Exploratory Data Analysis

Exploratory Data Analysis (EDA) dilakukan untuk mengetahui distribusi setiap variabel, nilai minimum, maksimum, rata-rata, dan variasi data. EDA juga digunakan untuk mengetahui hubungan awal antarvariabel sebelum proses clustering.

Analisis eksploratif diperlukan karena data aktivitas pembelajaran dapat memiliki distribusi yang berbeda dan tidak selalu menunjukkan hubungan linear. Analisis awal terhadap data juga dapat membantu menentukan variabel yang relevan dan mendeteksi karakteristik data yang berpotensi memengaruhi hasil clustering [4], [6].

Algoritma K-Means

Algoritma K-Means digunakan untuk mengelompokkan mahasiswa berdasarkan kemiripan karakteristik aktivitas belajar. K-Means bekerja dengan menentukan jumlah cluster, memilih centroid awal, menghitung jarak setiap data terhadap centroid, mengelompokkan data berdasarkan centroid terdekat, dan memperbarui centroid sampai proses mencapai kondisi konvergen.

K-Means dipilih karena sesuai untuk proses segmentasi data berdasarkan kemiripan karakteristik dan dapat digunakan untuk menemukan kelompok mahasiswa yang memiliki pola aktivitas belajar serupa. Penerapan clustering pada data akademik juga telah digunakan untuk mengidentifikasi pola perilaku belajar mahasiswa [9], [17].

Penentuan Jumlah Cluster

Jumlah cluster ditentukan menggunakan Elbow Method. Metode tersebut digunakan dengan membandingkan nilai Sum of Squared Error (SSE) pada beberapa kemungkinan jumlah cluster. Titik perubahan penurunan SSE digunakan sebagai pertimbangan untuk menentukan jumlah cluster yang sesuai.

Selain Elbow Method, hasil clustering dievaluasi menggunakan Silhouette Score. Penggunaan evaluasi cluster diperlukan untuk memastikan bahwa kelompok yang terbentuk memiliki tingkat pemisahan dan kedekatan internal yang memadai. Pendekatan evaluasi seperti ini juga digunakan pada penelitian clustering learning analytics untuk menentukan struktur kelompok yang sesuai [9], [17].

Evaluasi Clustering

Evaluasi clustering dilakukan menggunakan Silhouette Score. Nilai silhouette mengukur seberapa dekat suatu data dengan anggota cluster-nya dibandingkan dengan cluster lainnya. Nilai yang semakin tinggi menunjukkan pemisahan cluster yang semakin baik.

Selain evaluasi menggunakan nilai silhouette, setiap cluster dianalisis berdasarkan nilai rata-rata variabel. Interpretasi karakteristik cluster diperlukan agar hasil clustering tidak hanya berupa kelompok numerik, tetapi dapat diterjemahkan menjadi profil pola belajar mahasiswa yang dapat dipahami oleh dosen [4], [10].

Interpretasi Hasil

Tahap akhir penelitian adalah menginterpretasikan karakteristik setiap cluster. Interpretasi dilakukan dengan membandingkan rata-rata frekuensi akses LMS, durasi belajar, kehadiran, penyelesaian tugas, aktivitas forum, nilai kuis, dan nilai evaluasi pada setiap kelompok.

Hasil clustering kemudian digunakan untuk menyusun profil pola belajar mahasiswa. Pendekatan ini memungkinkan hasil analisis data science diterjemahkan menjadi informasi yang dapat digunakan untuk mendukung pembelajaran, seperti identifikasi mahasiswa dengan aktivitas tinggi, mahasiswa dengan aktivitas sedang, dan mahasiswa dengan aktivitas rendah [10], [13].

HASIL

Karakteristik Data

Data ilustratif terdiri atas 200 mahasiswa dengan tujuh variabel utama. Hasil statistik deskriptif ditunjukkan pada Tabel 2.

Tabel 2. Statistik Deskriptif Data Aktivitas Belajar Mahasiswa

Variabel	Minimum	Maksimum	Rata-rata
Frekuensi LMS	5	68	31,42
Durasi belajar (menit/minggu)	35	520	246,85
Kehadiran (%)	55	100	86,40
Penyelesaian tugas (%)	40	100	82,35
Aktivitas forum	0	45	14,72
Nilai kuis	45	98	77,63
Nilai evaluasi	48	96	78,41

Data menunjukkan adanya variasi aktivitas belajar mahasiswa. Frekuensi akses LMS memiliki rentang yang cukup besar, demikian pula durasi belajar, aktivitas forum, dan penyelesaian tugas.

Variasi tersebut menunjukkan bahwa mahasiswa memiliki karakteristik aktivitas pembelajaran yang berbeda. Perbedaan tersebut kemudian dianalisis menggunakan metode clustering untuk menemukan kelompok mahasiswa dengan karakteristik yang serupa.

Penentuan Jumlah Cluster

Pengujian menggunakan Elbow Method menghasilkan perubahan nilai SSE pada beberapa jumlah cluster. Berdasarkan pola penurunan SSE, tiga cluster dipilih sebagai struktur yang digunakan dalam analisis. Berdasarkan hasil Elbow Method, jumlah cluster yang digunakan adalah $K = 3$.

Hasil Clustering

Hasil clustering terhadap data ilustratif menghasilkan tiga kelompok mahasiswa.

Tabel 3. Distribusi Mahasiswa Berdasarkan Cluster

Cluster	Jumlah Mahasiswa	Persentase
Cluster 1	76	38,0%
Cluster 2	81	40,5%
Cluster 3	43	21,5%
Total	200	100%

Cluster 2 merupakan kelompok dengan jumlah mahasiswa terbesar, yaitu 81 mahasiswa atau 40,5%. Cluster 3 memiliki jumlah mahasiswa paling sedikit, yaitu 43 mahasiswa atau 21,5%.

Profil Pola Belajar

Tabel 4. Karakteristik Setiap Cluster

Indikator	Cluster 1	Cluster 2	Cluster 3
Frekuensi LMS	48,6	31,7	15,2
Durasi belajar	365,4	238,6	128,3
Kehadiran (%)	94,2	86,1	70,8
Penyelesaian tugas (%)	95,1	82,4	59,7
Aktivitas forum	25,8	13,6	3,8
Nilai kuis	87,5	76,8	63,4
Nilai evaluasi	88,7	77,4	64,1

Cluster 1 memiliki nilai rata-rata tertinggi pada hampir seluruh indikator. Kelompok ini menunjukkan tingkat aktivitas pembelajaran yang tinggi dan konsisten.

Cluster 2 memiliki nilai aktivitas pada tingkat menengah. Kelompok ini menunjukkan keterlibatan pembelajaran yang cukup tetapi belum setinggi Cluster 1.

Cluster 3 memiliki nilai terendah pada sebagian besar indikator. Kelompok ini menunjukkan tingkat aktivitas pembelajaran yang relatif rendah dibandingkan kelompok lainnya.

Evaluasi Clustering

Tabel 5. Nilai Silhouette Berdasarkan Jumlah Cluster

Jumlah Cluster	Silhouette Score
2	0,48
3	0,57
4	0,51

5	0,47
---	------

Nilai ilustratif menunjukkan bahwa tiga cluster menghasilkan nilai silhouette sebesar 0,57. Berdasarkan nilai tersebut, konfigurasi tiga cluster digunakan untuk proses interpretasi pola belajar.

PEMBAHASAN

Hasil analisis menunjukkan bahwa pola belajar mahasiswa memiliki karakteristik yang beragam. Perbedaan terlihat pada frekuensi akses LMS, durasi belajar, kehadiran, penyelesaian tugas, aktivitas forum, serta capaian akademik.

Cluster 1 menunjukkan pola belajar aktif dan konsisten. Mahasiswa pada kelompok ini memiliki aktivitas LMS yang tinggi, durasi belajar yang lebih panjang, tingkat kehadiran yang tinggi, serta penyelesaian tugas yang relatif baik. Aktivitas forum yang tinggi juga menunjukkan bahwa mahasiswa pada kelompok ini memiliki interaksi pembelajaran yang lebih aktif. Cluster 2 menunjukkan pola belajar cukup aktif. Mahasiswa pada kelompok ini masih mengikuti proses pembelajaran secara cukup konsisten, tetapi tingkat aktivitasnya berada di bawah Cluster 1. Kelompok ini dapat menjadi sasaran penguatan agar aktivitas pembelajaran menjadi lebih konsisten. Cluster 3 menunjukkan pola aktivitas belajar yang relatif rendah. Rendahnya akses LMS, durasi belajar, kehadiran, penyelesaian tugas, dan aktivitas forum diikuti oleh capaian nilai yang relatif lebih rendah. Kondisi tersebut menunjukkan bahwa pemantauan aktivitas belajar dapat menjadi salah satu cara untuk memperoleh informasi awal mengenai keterlibatan mahasiswa.

Hasil clustering memberikan keuntungan karena dosen tidak harus melihat setiap indikator mahasiswa secara terpisah. Mahasiswa dapat dilihat berdasarkan kelompok karakteristik yang memiliki pola serupa. Dengan demikian, informasi yang dihasilkan menjadi lebih ringkas dan mudah digunakan dalam perencanaan strategi pembelajaran. Kelompok mahasiswa aktif dapat diberikan aktivitas pengayaan yang lebih kompleks. Kelompok mahasiswa cukup aktif dapat diberikan penguatan terhadap konsistensi belajar. Sementara itu, kelompok mahasiswa dengan aktivitas rendah dapat memperoleh pendampingan, pengingat tugas, dan monitoring yang lebih intensif. Hasil analisis juga menunjukkan bahwa penggunaan beberapa indikator lebih informatif dibandingkan penggunaan satu indikator saja. Nilai akademik dapat menunjukkan hasil belajar, tetapi belum menjelaskan bagaimana mahasiswa menjalani proses pembelajaran. Sebaliknya, data aktivitas LMS dapat menunjukkan perilaku belajar, tetapi tidak secara langsung menunjukkan tingkat penguasaan materi. Oleh karena itu, kombinasi data aktivitas dan data akademik dapat memberikan gambaran yang lebih menyeluruh mengenai pola belajar mahasiswa. Pendekatan ini dapat dikembangkan lebih lanjut menjadi sistem learning analytics yang menampilkan perubahan pola belajar secara berkala.

Hasil clustering juga dapat digunakan sebagai dasar pengembangan dashboard pembelajaran. Dashboard dapat menampilkan distribusi mahasiswa pada setiap cluster, perkembangan aktivitas mahasiswa, penyelesaian tugas, kehadiran, dan perubahan capaian akademik. Namun demikian, hasil cluster tidak seharusnya diperlakukan sebagai label permanen terhadap mahasiswa. Pola belajar dapat berubah dari waktu ke waktu karena mahasiswa dapat meningkatkan atau menurunkan aktivitas belajarnya. Oleh karena itu, analisis sebaiknya dilakukan secara berkala.

Pembahasan dalam artikel bertujuan untuk: (1) menjawab rumusan masalah dan pertanyaan-pertanyaan penelitian; (2) menunjukkan bagaimana temuan-temuan itu diperoleh; (3) menginterpretasi/menafsirkan temuan-temuan; (4) mengaitkan hasil temuan penelitian dengan struktur pengetahuan yang telah mapan; dan (5) memunculkan teori-teori baru atau modifikasi teori yang telah ada. Dalam menjawab rumusan masalah dan pertanyaan-pertanyaan penelitian, hasil penelitian harus disimpulkan secara eksplisit. Penafsiran terhadap temuan dilakukan dengan menggunakan logika dan teori-teori yang ada. Temuan berupa kenyataan di lapangan diintegrasikan/ dikaitkan dengan hasil-hasil penelitian.

KESIMPULAN

Penelitian ini menunjukkan bahwa pendekatan data science dapat digunakan untuk menganalisis pola belajar mahasiswa berdasarkan data aktivitas pembelajaran dan data akademik. Algoritma K-Means dapat digunakan untuk mengelompokkan mahasiswa berdasarkan kemiripan karakteristik aktivitas belajar. Berdasarkan data ilustratif sebanyak 200 mahasiswa, diperoleh tiga kelompok pola belajar, yaitu kelompok mahasiswa aktif dan konsisten, kelompok mahasiswa cukup aktif, dan kelompok mahasiswa dengan aktivitas belajar rendah. Kelompok mahasiswa aktif menunjukkan tingkat akses LMS, durasi belajar, kehadiran, penyelesaian tugas, aktivitas forum, dan capaian akademik yang relatif tinggi. Kelompok mahasiswa dengan aktivitas rendah menunjukkan nilai yang relatif lebih rendah pada sebagian besar indikator.

Hasil penelitian dapat menjadi dasar untuk mengembangkan strategi pembelajaran yang berbeda sesuai dengan karakteristik mahasiswa. Pendekatan data science juga dapat dikembangkan menjadi sistem learning analytics atau dashboard yang mampu memantau perubahan pola belajar mahasiswa secara berkala. Penelitian selanjutnya dapat menggunakan dataset aktual dengan periode pengamatan yang lebih panjang, membandingkan beberapa algoritma clustering, serta mengembangkan model prediksi untuk mengetahui perubahan pola belajar mahasiswa dan hubungannya dengan keberhasilan akademik.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada pihak perguruan tinggi, dosen, mahasiswa, dan seluruh pihak yang telah memberikan dukungan dalam penyediaan data, pelaksanaan penelitian, pengolahan data, dan penyusunan artikel ini.

DAFTAR PUSTAKA

- [1] D. Ifenthaler and J. Y.-K. Yau, "Utilising learning analytics to support study success in higher education: A systematic review," *Educational Technology Research and Development*, vol. 68, pp. 1961–1990, 2020, doi: 10.1007/s11423-020-09788-z.
- [2] N. Johar, S. N. Kew, Z. Tasir, and E. Koh, "Learning analytics on student engagement to enhance students' learning performance: A systematic review," *Sustainability*, vol. 15, no. 10, p. 7849, 2023, doi: 10.3390/su15107849.
- [3] A. Palanci, R. M. Yilmaz, and Z. Turan, "Learning analytics in distance education: A systematic review study," *Education and Information Technologies*, vol. 29, pp. 22629–22650, 2024, doi: 10.1007/s10639-024-12737-5.

- [4] N. Bergdahl, M. Bond, J. Sjöberg, M. Dougherty, and E. Oxley, "Unpacking student engagement in higher education learning analytics: A systematic review," *International Journal of Educational Technology in Higher Education*, vol. 21, no. 63, 2024, doi: 10.1186/s41239-024-00493-y.
- [5] S. Alhazbi, A. Al-ali, A. Tabassum, A. Al-Ali, A. Al-Emadi, T. Khattab, and M. A. Hasan, "Using learning analytics to measure self-regulated learning: A systematic review of empirical studies in higher education," *Journal of Computer Assisted Learning*, vol. 40, no. 4, pp. 1658–1674, 2024, doi: 10.1111/jcal.12982.
- [6] M. Yağcı, "Educational data mining: Prediction of students' academic performance using machine learning algorithms," *Smart Learning Environments*, vol. 9, no. 11, 2022, doi: 10.1186/s40561-022-00192-z.
- [7] F. Rizqiaty and A. Fauka, "Identifying student learning behavior patterns using K-Means clustering: A case study at STMIK Mardira Indonesia," *Jurnal Computech & Bisnis*, vol. 20, no. 1, 2026, doi: 10.56447/jcb.v20i1.02.
- [8] D. S. Pramdhana, E. Nur, I. Sumadikarta, and Y. T. Aprillia, "Learning analytics berbasis Kahoot menggunakan algoritma K-Means clustering untuk mengidentifikasi pola keterlibatan mahasiswa dalam pembelajaran," *Jurnal Teknologi Informasi*, vol. 12, no. 1, 2026.
- [9] A. Palanci, R. M. Yilmaz, and Z. Turan, "Learning analytics in distance education: A systematic review study," *Education and Information Technologies*, vol. 29, pp. 22629–22650, 2024, doi: 10.1007/s10639-024-12737-5.
- [10] L. Paulsen and E. Lindsay, "Learning analytics dashboards are increasingly becoming about learning and not just analytics: A systematic review," *Education and Information Technologies*, vol. 29, pp. 14279–14308, 2024, doi: 10.1007/s10639-023-12401-4.
- [11] D. Heikkinen, M. Saqr, J. Malmberg, and M. Tedre, "Supporting self-regulated learning with learning analytics interventions: A systematic literature review," *Education and Information Technologies*, vol. 28, pp. 3059–3088, 2023, doi: 10.1007/s10639-022-11281-4.
- [12] Y. S. Balcioğlu and M. Artar, "Predicting academic performance of students with machine learning," *Information Development*, 2025, doi: 10.1177/02666669231213023.
- [13] G. Ramaswami, T. Susnjak, and A. Mathrani, "Effectiveness of a learning analytics dashboard for increasing student engagement levels," *Journal of Learning Analytics*, vol. 10, no. 3, pp. 115–134, 2023, doi: 10.18608/jla.2023.7935.
- [14] A. Stojanov and B. K. Daniel, "A decade of research into the application of big data and analytics in higher education: A systematic review of the literature," *Education and Information Technologies*, vol. 29, pp. 5807–5831, 2024, doi: 10.1007/s10639-023-12033-8.
- [15] L. Márquez, V. Henríquez, H. Chevreux, E. Scheihing, and J. Guerra, "Adoption of learning analytics in higher education institutions: A systematic literature review," *British Journal of Educational Technology*, vol. 55, no. 2, pp. 439–459, 2024, doi: 10.1111/bjet.13385.
- [16] F. Ouyang, M. Wu, L. Zheng, L. Zhang, and P. Jiao, "Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering

course,” *International Journal of Educational Technology in Higher Education*, vol. 20, no. 4, 2023, doi: 10.1186/s41239-022-00372-4.

- [17] M. Zhu, A. R. Sari, and M. M. Lee, “Trends and issues in MOOC learning analytics empirical research: A systematic literature review (2011–2021),” *Education and Information Technologies*, vol. 27, pp. 10135–10160, 2022, doi: 10.1007/s10639-022-11031-6.