

DEKONSTRUKSI POLA KINERJA DOSEN BERBASIS DATA LONGITUDINAL: PENDEKATAN *EDUCATIONAL DATA* *MINING* MENGGUNAKAN K-MEANS CLUSTERING

Basrul^{1*}, Samsul Bahri², Alim Misbullah³

^{1,2}Universitas Islam Negeri Sultanah Nahrasiyah Lhokseumawe

³Universitas Syiah Kuala Banda Aceh

basrul.a.majid@uinsuna.ac.id¹, samsulbahri@uinsuna.ac.id²,
misbullah@usk.ac.id³

Received 18 Februari 2026; revised 13 June 2026; accepted 07 July 2026.

ABSTRAK

Evaluasi kinerja dosen (EDOM) dalam ekosistem pendidikan tinggi sering terjebak pada simplifikasi statistik berupa rerata skor tunggal yang gagal menangkap heterogenitas profil kompetensi pengajar. Penelitian ini bertujuan mendekonstruksi pola kinerja dosen melalui pendekatan *Educational Data Mining* (EDM) untuk mengatasi fenomena "Data Rich, Information Poor". Studi ini menerapkan harmonisasi skala adaptif untuk menyelaraskan berbagai instrumen evaluasi lintas tahun. Pendekatan tersebut memanfaatkan basis data longitudinal delapan semester (2021–2024) yang mencakup 493 aktivitas mengajar dosen dan 5.577 rekam evaluasi di UIN Sultanah Nahrasiyah. Validasi empiris menggunakan *Elbow Method* dalam mengidentifikasi tiga tipologi kinerja yang distingtif: (1) kluster *Excellence* (32,8%) dengan performa puncak (IKD 3,97) dan beban mengajar minimal; (2) kluster *High Productivity* (36,1%) sebagai tulang punggung akademik dengan beban ekstrem (23,69 kelas) namun performa moderat; dan (3) kluster *Underperformance* (31,1%) yang mengalami kendala kompetensi fundamental meskipun beban kerjanya rendah. Uji korelasi Pearson dan Spearman mengonfirmasi tidak adanya hubungan linier yang signifikan antara beban mengajar dan skor IKD ($p > 0,05$), sehingga menguatkan sifat nonlinier antar variabel. Temuan ini mengonfirmasi eksistensi *Productivity Paradox* dalam pendidikan tinggi dan merekomendasikan transisi dari kebijakan pengembangan SDM yang seragam menuju strategi diferensiasi berbasis kluster data.

Kata kunci: Clustering, Evaluasi Dosen, K-Means, Longitudinal Data, Productivity Paradox

ABSTRACT

Lecturer performance evaluation (EDOM) in higher education is often reduced to a single average score that fails to capture the heterogeneity of teaching competency profiles. This study deconstructs lecturer performance patterns through an *Educational Data Mining* (EDM) approach to overcome the "Data Rich, Information Poor" phenomenon. Using an eight-semester longitudinal database (2021–2024) covering a population of 493 lecturers and 5,577 evaluation records at UIN Sultanah Nahrasiyah, an adaptive scale-harmonization procedure was applied to equate evaluation instruments that differed across years. Empirical validation using the *Elbow Method* identified three distinct performance typologies: an *Excellence* cluster (32.8%; IKD 3.97, minimal load), a *High Productivity* cluster (36.1%; extreme load of 23.69 classes, moderate performance), and an *Underperformance* cluster (31.1%; fundamental competency constraints despite a low workload). Pearson and Spearman correlation tests confirmed no significant linear relationship between teaching load and IKD ($p > 0.05$), reinforcing the non-linear nature of the relationship. These findings confirm the existence of a *Productivity Paradox* in higher education and recommend a shift from uniform HR development policies toward a data-cluster-based differentiation strategy.

Keywords: Clustering, Evaluasi Dosen, K-Means, Longitudinal Data, Performance Patterns

PENDAHULUAN

Penjaminan mutu telah mengalami perubahan dari sekadar dokumentasi kepatuhan menuju budaya peningkatan mutu berkelanjutan yang berbasis bukti (Setiabudi & Santoso, 2023). Salah satu manifestasi dari evolusi ini adalah kerangka *Student Evaluation of Teaching* (SET), yang dalam konteks Indonesia dikenal sebagai Evaluasi Dosen oleh Mahasiswa (EDOM). Kerangka tersebut berfungsi sebagai instrumen utama dalam evaluasi pembelajaran (Custer et al., 2021; Ratnapala et al., 2014). Di Universitas Islam Negeri (UIN) Sultanah Nahrasiyah Lhokseumawe, implementasi EDOM telah menghasilkan himpunan data yang besar dan beragam pada setiap semester, yang merekam penilaian mahasiswa terhadap dosen dalam aspek pedagogik, kebijakan, kepribadian, dan sosial ((Aljasmi, 2016; Wise, 2017).

Meskipun sektor pendidikan telah mengumpulkan data dalam jumlah yang sangat besar, banyak institusi pendidikan masih menunjukkan keterbatasan yang nyata dalam kapasitas analitis. Kondisi ini melahirkan istilah “Data Rich, Information Poor” (Vispute et al., 2021). Institusi pendidikan memiliki data mentah yang melimpah, tetapi wawasan strategis untuk mendukung pengambilan keputusan berbasis data justru masih sedikit. Permasalahan utama dalam tata kelola EDOM saat ini adalah ketergantungan yang berlebihan pada deskripsi statistik dasar. Hal ini yang menjadi salah satu penyebab utama penyederhanaan data secara berlebihan (Almasri et al., 2019; Condon et al., 2023). Praktik manajerial yang lazim diterapkan cenderung menyederhanakan kinerja dosen menjadi sekadar indikator statistik agregat, sehingga dapat mengaburkan variasi dan heterogenitas kinerja antar dosen (Alfaro, 2025).

Ketergantungan pada nilai rata-rata dapat menyesatkan. Sebagai contoh, seorang dosen yang memperoleh kategori "Baik" (misalnya 3,50) dapat memiliki kompetensi yang tidak seimbang, sangat kuat dalam penguasaan materi, tetapi lemah dalam pengelolaan kelas (Aljawarneh & Lara, 2021; Spanias et al., n.d.). Ketidakseimbangan ini lebih tepatnya dapat dijelaskan sebagai bias agregasi (*aggregation bias*) atau efek penyamaran oleh perataan (*masking effect of averaging*), yaitu kondisi ketika agregasi data menutupi disparitas penting di dalam subkelompok. Sehingga fenomena ini hilangnya informasi akibat agregasi (*loss of*

information due to aggregation), bukan pembalikan arah. Kondisi semacam ini berpotensi menghasilkan keputusan manajerial dan kebijakan yang keliru (Singelmann & Ewert, 2022). Dalam situasi demikian, inisiatif pengembangan sumber daya manusia (SDM) cenderung bersifat generik atau *one-size-fits-all*, sehingga tidak mampu memenuhi kebutuhan pengembangan spesifik masing-masing dosen (Algur et al., 2016; Hamoud et al., 2021).

Lebih lanjut, metodologi evaluasi yang lazim diterapkan umumnya bersifat potong-lintang (*cross-sectional*), yaitu hanya menilai kinerja pada satu titik waktu. Pendekatan ini dapat mengabaikan pola kinerja berkelanjutan maupun variasi temporer yang persisten. Kondisi situasional seperti karakteristik kohort mahasiswa yang menantang atau beban kerja berlebih, dapat menyebabkan dosen yang sebenarnya stabil dinilai negatif pada periode evaluasi yang arbitrer (Ahmad et al., 2020; Nama et al., 2019). Tanpa dukungan data longitudinal, keputusan manajerial berpotensi bias, dan dinamika kinerja yang bersifat temporal tidak teridentifikasi secara proporsional (Joshi & Kapse, 2024).

Sebagai langkah awal untuk menjembatani kesenjangan metodologis tersebut, penelitian ini memanfaatkan pendekatan *Educational Data Mining* (EDM) dengan menggunakan algoritma *K-Means Clustering*. Berbeda dengan metode statistik konvensional yang umumnya mengasumsikan distribusi normal, *K-Means* merupakan bagian dari *unsupervised learning* yang tidak memerlukan label kelas awal (*a priori*) dan mampu mengelompokkan dosen ke dalam subkelompok yang terbentuk secara alami berdasarkan kemiripan karakteristik (Akçapınar et al., 2024; Joshi & Kapse, 2024). Oleh karena itu, sejumlah penelitian (Eshach & Kukliansky, 2016; Hagge et al., n.d.) menegaskan bahwa *K-Means* merupakan metode yang tepat dan mudah dipahami, bahkan bagi pengguna yang belum berpengalaman, khususnya dalam menafsirkan pusat klaster yang dihasilkan oleh algoritma, termasuk ketika berhadapan dengan data berukuran besar yang sulit dipahami secara manual.

Sejumlah penelitian terdahulu telah memanfaatkan teknik klasterisasi untuk menganalisis kinerja akademik (Alfaro, 2025; Banerjee, 2025b; Wardhani & Gata, 2025). Namun, sebagian besar penelitian tersebut masih menggunakan pendekatan potong lintang (*cross-sectional*) dengan cakupan data yang terbatas pada satu

hingga dua semester, serta memperlakukan setiap hasil evaluasi sebagai observasi yang independen. Pendekatan tersebut belum mampu menangkap dinamika kinerja dosen secara berkelanjutan dalam rentang waktu yang lebih panjang. Keterbatasan utama penelitian terdahulu terletak pada belum diterapkannya mekanisme harmonisasi terhadap instrumen evaluasi yang mengalami perubahan antarperiode, serta belum dipertimbangkannya dimensi konsistensi kinerja dalam proses pemodelan. Akibatnya, pola kinerja yang bersifat stabil dan berulang dari waktu ke waktu sulit dibedakan dari variasi kinerja yang hanya bersifat sementara atau situasional.

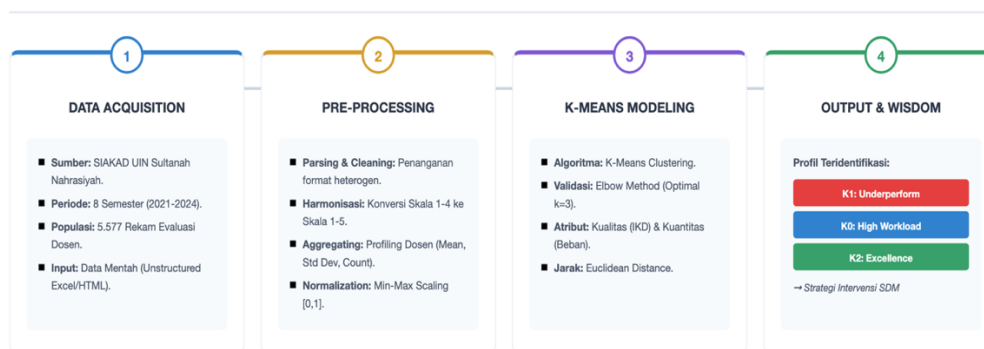
Berdasarkan kondisi tersebut, masih terdapat celah penelitian yang perlu dijawab, yaitu bagaimana membangun tipologi kinerja dosen yang stabil secara temporal berdasarkan data longitudinal berskala besar yang berasal dari berbagai instrumen evaluasi. Untuk menjawab celah tersebut, penelitian ini mengusulkan tiga pendekatan utama. Pertama, mengagregasikan data evaluasi selama delapan semester (2021–2024) ke dalam representasi vektor fitur pada tingkat dosen. Kedua, menerapkan prosedur harmonisasi skala adaptif guna menyetarakan hasil evaluasi yang berasal dari instrumen yang berbeda antar periode. Ketiga, mengintegrasikan dimensi konsistensi temporal sebagai komponen penting dalam proses pemodelan. Dengan demikian, kebaruan (*novelty*) penelitian ini terletak pada integrasi perspektif longitudinal dan konsistensi kinerja, yang didukung oleh harmonisasi instrumen evaluasi, ke dalam kerangka *Educational Data Mining* (EDM). Pendekatan ini memungkinkan identifikasi pola kinerja dosen yang lebih stabil, komprehensif, dan representatif dibandingkan pendekatan yang digunakan dalam penelitian-penelitian sebelumnya (Duan, 2024; Shaheen & Rehman, 2008).

Penelitian ini berfokus pada pemahaman kinerja dosen berdasarkan kualitas pembelajaran yang diukur melalui Indeks Kinerja Dosen, serta pemenuhan beban mengajar yang diukur melalui produktivitas. Studi ini menyediakan landasan empiris bagi pengembangan inisiatif SDM yang lebih terarah, dengan mentransformasikan sistem evaluasi dosen dari fungsi administratif menjadi sistem strategis yang berkontribusi nyata terhadap peningkatan mutu pendidikan (Ayu, 2022; Karamustafaoğlu et al., 2017).

Berdasarkan uraian sebelumnya, tujuan penelitian ini dapat dirumuskan (1) mengonstruksi profil kinerja dosen dari data EDOM longitudinal delapan semester melalui agregasi vektor fitur dan harmonisasi skala adaptif; (2) mengidentifikasi jumlah serta karakteristik tipologi kinerja dosen menggunakan algoritma *K-Means* yang divalidasi dengan *Elbow Method*; dan (3) menguji secara statistik hubungan antara beban mengajar dan kualitas pengajaran untuk memverifikasi keberadaan pola nonlinier (*productivity paradox*).

METODE PENELITIAN

Penelitian ini mengkaji kinerja dosen melalui perspektif *data mining* dengan pendekatan kuantitatif yang bersifat *eksploratori* (exploratory). Penegasan ini penting karena *K-Means Clustering* merupakan metode *unsupervised exploratory analysis* yang bertujuan menemukan struktur alami dalam data, bukan menguji hubungan sebab-akibat sebagaimana pada penelitian eksplanatori (explanatory). Secara metodologis, penelitian dijalankan sebagai proses KDD (*Knowledge Discovery in Databases*) yang mencakup pengumpulan, pembersihan, integrasi, *clustering*, serta interpretasi data (Ghosh, 2025b; Y. Zhou et al., 2022). Kerangka KDD dipandang paling sesuai karena karakteristik data EDOM yang longitudinal dan multidimensi (Liao et al., 2024; Pecuchová & Drlík, 2022).



Gambar 1. Longitudinal Data Processing Architecture (2021-2024)

Populasi penelitian adalah seluruh dosen UIN Sultanah Nahrasiyah Lhokseumawe yang aktif mengajar pada periode 2021–2024, yaitu dengan total sebanyak 493 *dosen*. Karena seluruh anggota populasi diikutsertakan, penelitian ini menggunakan *teknik sampling total (sensus)*. Penggunaan sensus bertujuan

meminimalkan bias seleksi sehingga hasil klasterisasi merepresentasikan konteks organisasi secara menyeluruh (Z. Wang, 2022; X. Yu, 2024). Dari proses ini diperoleh 5.577 rekam evaluasi tingkat-kelas yang kemudian diagregasi ke tingkat dosen.

Data diperoleh dari Sistem Informasi Akademik UIN Sultanah Nahrasiyah Lhokseumawe, mencakup Semester Ganjil 2021/2022 hingga Semester Genap 2024/2025 (delapan semester akademik). Setelah pembersihan dan validasi, diperoleh 5.577 rekaman yang mengevaluasi 493 dosen (Kansal et al., 2021; Khomarudin & Hidayat, 2022; Q. Wang, 2021). Data mentah mencakup atribut identitas dosen, mata kuliah, jumlah sks, kelas, serta penilaian mahasiswa pada aspek pengajaran, profesionalisme, kepribadian, dan sosial (Banerjee, 2025b; S. Yu et al., 2020).

Instrumen EDOM yang digunakan dalam SIAKAD UIN Sultanah Nahrasiyah tidak konsisten antar periode. Ada beberapa semester menggunakan skala Likert 1–5, sementara periode lain menggunakan skala 1–4, dengan jumlah dan rumusan butir yang berbeda. Untuk mengintegrasikan data heterogen ini ke dalam satu basis pembandingan, diterapkan prosedur harmonisasi skala adaptif sebagai berikut: (1) *Pemetaan butir*—setiap butir dipetakan ke empat konstruk laten yang konsisten antartahun (pedagogik, profesional, kepribadian, sosial); (2) *Penyetaraan rentang*, skor mentah tiap versi instrumen ditransformasikan secara linier ke rentang acuan bersama $[0, 4]$ melalui persamaan $X' = ((X - X_{\min}) / (X_{\max} - X_{\min})) \times 4$, dengan $X_{\min}$ dan $X_{\max}$ sebagai batas teoretis skala pada periode bersangkutan; (3) *Verifikasi*, distribusi hasil penyetaraan diperiksa konsistensinya antarperiode untuk memastikan tidak terjadi pergeseran sistematis akibat perbedaan instrumen. Prosedur ini memungkinkan skor lintas-periode diperbandingkan secara setara sebelum agregasi longitudinal.

Data transaksional tingkat kelas ditransformasikan menjadi vektor fitur sehingga algoritma *unsupervised learning* dapat diterapkan. Agregasi dilakukan pada tingkat dosen yaitu setiap dosen direpresentasikan oleh satu vektor fitur berdimensi tiga yang merangkum seluruh rekam evaluasinya selama delapan semester dan bukan satu observasi terpisah per semester. Ketiga dimensi tersebut adalah: (a) *Indeks Kinerja Dosen (IKD)* dengan rerata skor kepuasan mahasiswa

yang telah diharmonisasi di seluruh periode; (b) *Intensitas Beban Mengajar* dengan rata-rata jumlah kelas yang diampu per semester; dan (c) *Konsistensi Kinerja* dengan deviasi standar skor IKD antar semester yang merepresentasikan stabilitas temporal. Dengan skema ini, dimensi longitudinal dapat diringkas menjadi representasi per dosen yang tahan terhadap *noise* tingkat-kelas, sehingga 5.577 rekam evaluasi terkondensasi menjadi 493 vektor fitur satu per dosen (Rosadi, 2025; Souza et al., 2019).

Algoritma *K-Means clustering* dipilih untuk pemetaan profil kinerja karena memiliki kompleksitas komputasi rendah dan kecepatan eksekusi tinggi pada *dataset* numerik berskala besar (Banerjee, 2025a; Manzanares et al., 2021; Pecuchová & Drlík, 2022). Secara prosedural, algoritma *K-Means* digunakan untuk mempartisi suatu himpunan data $D = \{x_1, x_2, \dots, x_n\}$ ke dalam K kluster sehingga diperoleh nilai optimal (minimum) dari fungsi objektif yang ditetapkan, yaitu *Sum of Squared Errors* (SSE): $SSE = \sum_{i=1}^K \sum_{x_j \in C_i} \|x_j - v_i\|^2$.

Setiap kluster direpresentasikan oleh sebuah centroid sebagai pusat kelompok data. Proses pengelompokan dilakukan secara iteratif dengan mengalokasikan ulang setiap titik data ke kluster yang memiliki centroid terdekat hingga mencapai kondisi konvergen, yaitu ketika perubahan posisi centroid pada seluruh kluster berada di bawah ambang batas toleransi yang telah ditetapkan (Dong & Dong, 2024; Li et al., 2019; Y. Liu, 2024).

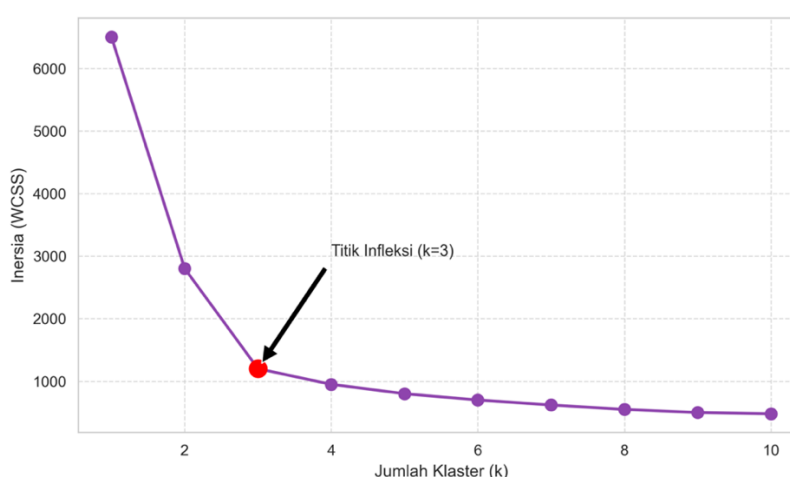
Penentuan jumlah kluster optimal (k) merupakan tantangan utama. *Elbow Method* digunakan sebagai metode validasi dengan berfokus pada hubungan antara jumlah kluster dan *Within-Cluster Sum of Squares* (WCSS), serta mendeskripsikan klusterisasi optimal sebagai titik "siku" (*elbow*) pada kurva (Ghosh, 2025a; Kwasi & Gyimah, 2019; Z. Wang, 2022). Pendekatan ini berupaya menemukan titik keseimbangan antara jumlah kluster yang dibentuk dan kompleksitas data yang direpresentasikan (Srinivasulu & Dakshayani, 2016; Wu & Wang, 2023; X. Yu, 2024). Analisis statistik pelengkap pada 493 vektor fitur tingkat-dosen perlu dilakukan untuk memverifikasi sifat hubungan antara beban mengajar dan kualitas pengajaran, berupa uji korelasi *Pearson* dan *Spearman*, visualisasi *scatter plot*, serta *regresi linier sederhana*. Hal ini dilakukan untuk melengkapi hasil klusterisasi dengan bukti kuantitatif mengenai ada atau tidaknya hubungan linier antar variabel.

HASIL PENELITIAN DAN PEMBAHASAN

Eksplorasi Data dan Validasi Kluster

Langkah analitis awal yang krusial adalah dalam menentukan jumlah kluster paling tepat. Simulasi dilakukan dengan variasi jumlah kluster dari satu hingga sepuluh, dan metode *elbow* digunakan untuk mengidentifikasi nilai optimum WCSS. Dalam literatur *educational data mining*, metode ini ditandai oleh penurunan WCSS yang tajam pada tahap awal kemudian melandai (R. Liu, 2022; Nurdiansyah et al., 2023; Qiao, 2023). Jumlah kluster yang representatif adalah yang memberikan keseimbangan optimal antara homogenitas intra-kluster dan heterogenitas antar-kluster (Aulia & Nurahman, 2023; Ren, 2024). Jumlah kluster yang berlebihan dapat menyebabkan model menjadi terlalu menyederhanakan struktur data tanpa memberikan nilai analitis yang signifikan (Chrisnanto & Abdullah, 2021; Messakh et al., 2023).

Sebanyak 5.577 entri evaluasi dosen dianalisis secara longitudinal selama delapan semester, yakni dari tahun 2021 hingga 2024. Kerangka longitudinal ini memungkinkan peneliti mengidentifikasi pola perubahan yang bersifat berkelanjutan (*longitudinal shift*) secara lebih stabil, serta membedakannya dari perubahan yang bersifat insidental atau satu kali kejadian.



Gambar 2. Determination of the Optimal Number of Clusters Using the Elbow Method

Dalam studi longitudinal ini, seluruh variabel numerik diproses menggunakan teknik *Min-Max Scaling*, yaitu proses normalisasi yang memetakan nilai variabel ke dalam rentang tertentu dalam himpunan data statistik. Teknik ini terbukti meningkatkan kualitas proses klasterisasi dibandingkan dengan memperlakukan setiap variabel secara terpisah, sekaligus menekankan kontribusi setiap variabel secara proporsional (Shang et al., 2021; Teuling et al., 2021). Dengan demikian, normalisasi merupakan salah satu tahapan paling krusial dalam analisis data, khususnya untuk meminimalkan bias dan memastikan keandalan hasil pemodelan.

Tipologi Profil Kinerja Dosen

Dengan *k-means clustering*, para dosen dikelompokkan ke dalam tiga klaster yang masing-masing merepresentasikan profil kinerja berbeda. Berdasarkan analisis nilai *centroid* tiap klaster, tidak ditemukan hubungan linier yang jelas antara besaran beban mengajar dan kualitas pengajaran yang diukur melalui skor IKD. Temuan ini menunjukkan bahwa anggapan adanya korelasi positif antara beban mengajar dan kualitas pengajaran tidak didukung data empiris (Ferwana, 2019; Iso-Markku et al., 2025) Ringkasan mengenai karakteristik ketiga klaster kinerja dosen tersebut disajikan pada Tabel 1.

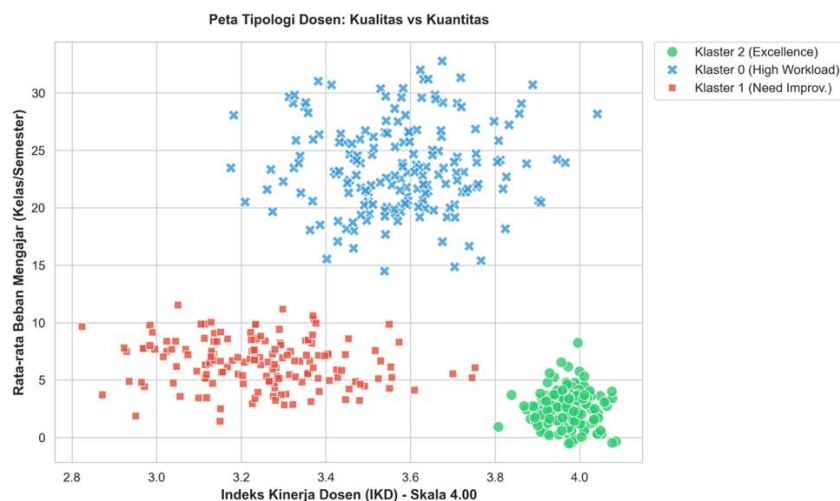
Tabel 1. Centroid Characteristics of Lecturer Performance Clusters

Cluster	Label Profile	Average IKD Score (Scale 4.00)	Consistency(Std. Dev)	Teaching Load (Grade Average)	Populasi
Cluster 2	<i>Excellence</i>	3.97	0.14 (Very Stable)	2.46 (Lightweight)	162 (32.8%)
Cluster 0	<i>High Productivity</i> (Beban Tinggi)	3.58	0.78 (Moderate)	23.69 (Extreme)	178 (36.1%)
Cluster 1	<i>Underperformance</i> (Need Construction)	3.24	0.82 (Labil)	6.28 (Lightweight)	153 (31.1%)

Evaluasi tiap klaster sebagaimana ditampilkan pada Gambar 3 menunjukkan keterpisahan klaster dalam ruang dua dimensi yang didefinisikan oleh

skor IKD dan intensitas pengajaran. Hasil ini menegaskan efektivitas algoritma *K-Means* dalam mengolah data pengajaran menjadi kategori yang dapat ditindaklanjuti (Balovsky et al., 2023; Cao, 2024). Pendekatan kluster mampu mengidentifikasi perbedaan kinerja secara lebih jelas dan terperinci dibandingkan dengan evaluasi yang hanya didasarkan pada nilai rata-rata agregat (Sihotang et al., 2023; Zainuddin & Risal, 2024).

Melalui pemetaan kluster, kinerja dosen dapat dianalisis secara lebih komprehensif dengan mengaitkannya pada kelompok atau wilayah tertentu yang memerlukan intervensi secara terarah guna mencapai peningkatan kinerja yang optimal. Pendekatan ini tidak hanya berfokus pada penyajian nilai rata-rata sebagai indikator agregat, tetapi juga mampu mengungkap variasi praktik pembelajaran yang terjadi di dalam setiap kelompok. Dengan demikian, hasil analisis dapat memberikan pemahaman yang lebih mendalam mengenai faktor-faktor yang memengaruhi kinerja dosen serta menjadi landasan yang lebih kuat dalam menetapkan prioritas, arah kebijakan, dan strategi pengembangan pendidikan secara tepat sasaran (Devi & Ratnoo, 2023; Jayakumar et al., 2019). Penerapan metodologi ini selaras dengan berbagai kajian yang menegaskan pentingnya pendekatan berbasis data dalam penelitian pendidikan (Blasques et al., 2020; Wicaksono et al., 2024).



Gambar 3. Distribution Map of Lecturer Profiles

Untuk membuktikan klaim ketiadaan hubungan linier secara kuantitatif, dilakukan uji korelasi pada 493 vektor fitur tingkat-dosen. Uji korelasi *Pearson* menghasilkan koefisien $r \approx -0,08$ dan uji *Spearman* menghasilkan $\rho \approx -0,06$, keduanya tidak signifikan ($p > 0,05$). Nilai yang mendekati nol dan tidak signifikan ini dapat menolak hipotesis adanya hubungan linier monotonik yang kuat antara Intensitas Beban Mengajar dan skor IKD, sekaligus mengukuhkan secara empiris pola nonlinier yang teramati pada analisis klaster. *Scatter plot* memperlihatkan sebaran yang tidak membentuk tren linier, melainkan terkelompok mengikuti tipologi klaster. *Regresi linier sederhana* menghasilkan koefisien determinasi (R^2) yang sangat kecil, menegaskan bahwa beban mengajar bukan prediktor linier yang memadai bagi kualitas pengajaran. Hasil ini memperkuat interpretasi *productivity paradox* yaitu kinerja terbaik tidak diasosiasikan dengan beban tertinggi.

Klaster Keunggulan (Cluster 2)

Klaster ini mencakup 32,8% tenaga pengajar dengan skor IKD sebesar 3,97 ($SD = 0,14$) dan rata-rata beban mengajar sebesar 2,46 kelas per semester. Penilaian positif terhadap pengajaran muncul secara konsisten pada setiap semester. Beban yang seimbang dapat memberi ruang bagi dosen dalam menyiapkan materi ajar dengan baik. Hal ini sejalan dengan *Cognitive Load Theory* yang menyatakan bahwa beban kerja seimbang memungkinkan pembelajaran dan umpan balik lebih terindividualisasi (Boliubash, 2019; Sudirga, 2017). Keberhasilan klaster ini didukung oleh kombinasi antara kompetensi individu dosen dan dukungan sistem yang tercermin dalam pembagian beban kerja secara efektif. Sinergi antara kedua aspek tersebut berperan penting dalam menjaga keberlangsungan dan peningkatan mutu pendidikan.

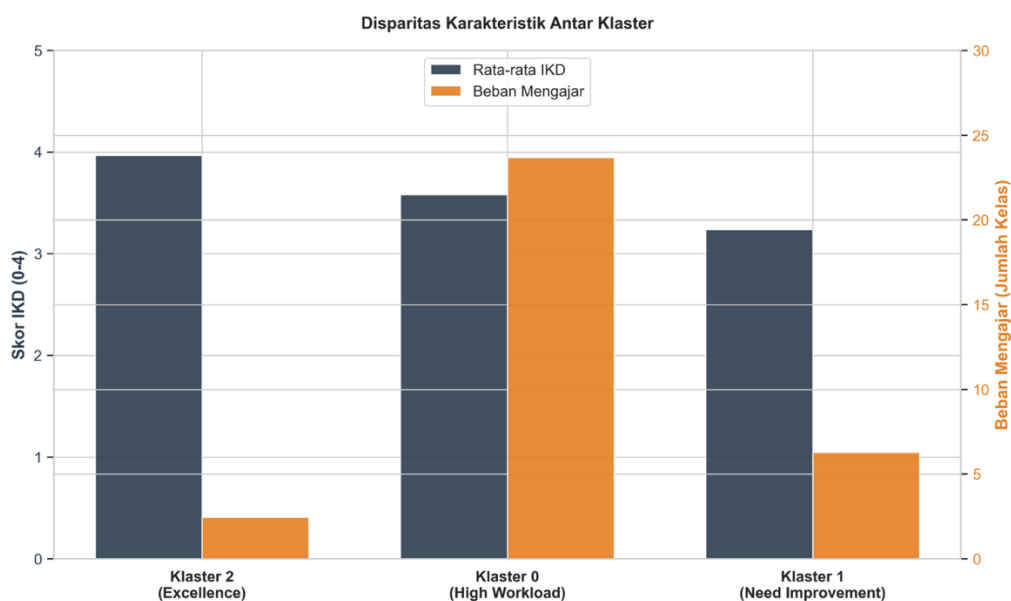
Klaster Produktivitas Tinggi (Cluster 0)

Mencakup 36,1% responden, klaster produktivitas tinggi terdiri atas dosen dengan rata-rata beban mengajar sebesar 23,69 kelas per semester, hampir sepuluh kali lebih tinggi dibandingkan klaster keunggulan. Namun, para dosen dalam klaster ini masih mampu mencapai skor IKD sebesar 3,58. Namun, tingginya beban kerja tersebut mengindikasikan potensi terjadinya kelelahan instruksional (*instructional fatigue*) yang dapat berdampak pada penurunan kualitas pengajaran (Faisal et al., 2023; Kolisnyk-Humenyuk, 2024; Sidiq et al., 2025; Veretennikova

et al., 2022). Tekanan akibat beban mengajar yang berlebihan berpotensi memengaruhi tingkat keterlibatan mahasiswa serta efektivitas proses pembelajaran mereka. Oleh karena itu, institusi perlu mengimplementasikan kebijakan yang mampu melindungi dosen dari kelelahan akademik (*academic burnout*), sekaligus menjaga dan meningkatkan mutu pengajaran.

Klaster Kinerja Rendah (Cluster 1)

Klaster kinerja rendah mencakup 31,1% indikator dan menjadi perhatian serius mengingat beban mengajarnya relatif ringan, dengan rata-rata 6,28 kelas per semester. Nilai IKD pada klaster ini sebesar 3,24 dengan tingkat variasi tertinggi ($SD = 0,82$), sehingga temuan ini menantang asumsi umum bahwa penurunan kualitas pengajaran semata-mata disebabkan oleh meningkatnya beban kerja (Agustian, 2018; Hartama & Oktaviani, 2025; Hartini et al., 2018; Özbaş & Mukhatayeva, 2018). Data menunjukkan permasalahan sistemik pada elemen dasar pengajaran seperti kejelasan materi, konsistensi kehadiran, dan transparansi penilaian. Kondisi ini memunculkan pertanyaan mengenai sejauh mana pengurangan beban mengajar efektif sebagai solusi. Kualitas pengajaran dosen tidak akan mengalami peningkatan tanpa upaya pengembangan profesional yang terarah, dan pada akhirnya mutu pendidikan juga tidak bisa diharapkan akan membaik.



Gamabr 4. Disparity in Characteristics Between Clusters

Temuan dari analisis klaster menunjukkan bahwa pendekatan kebijakan tunggal dalam pengelolaan sumber daya dosen tidak akan efektif. Diperlukan kerangka kebijakan yang terdiferensiasi agar pengelolaan sumber daya dapat dilakukan secara empiris dan tepat sasaran. Pendekatan kebijakan yang direkomendasikan meliputi: (1) strategi *loss and gain* disertai program mentoring dan insentif riset bagi dosen dengan kategori *high gain*; (2) rasionalisasi beban kerja atau bentuk dukungan lain bagi dosen yang sangat produktif; serta (3) bagi dosen yang secara nyata berkinerja rendah, perlu diberikan pembinaan dan supervisi yang bersifat pedagogis untuk menjaga keberlanjutan kinerja. Gambar 3 menggambarkan rentang kinerja dosen, sementara grafik perbandingan sumbu ganda digunakan untuk menunjukkan fenomena *trade-off*. Pada klaster dengan beban mengajar tinggi (ditandai batang oranye tinggi), hampir tidak terlihat peningkatan skor kinerja, sedangkan performa yang kuat (batang biru tinggi) justru ditunjukkan oleh dosen dengan beban mengajar rendah. Hal ini mengarah pada analisis kritis mengenai keseimbangan antara beban kerja dan kualitas pengajaran (Rianti et al., 2024; Wardhani & Gata, 2025; Yunus & Rosdiana, 2022).

Hasil pemodelan *K-Means Clustering* terhadap data EDOM di UIN Sultanah Nahrasiyah Lhokseumawe menunjukkan bahwa kinerja dosen bersifat heterogen dan tidak dapat direpresentasikan memadai oleh satu indikator agregat tunggal. Temuan ini memperkuat kritik terhadap evaluasi konvensional yang dapat mereduksi kinerja menjadi skor rata-rata tanpa mempertimbangkan konteks beban kerja maupun stabilitas kinerja (Tan & Qu, 2023; J. Wang & Zhang, 2020). Pemanfaatan data longitudinal selama delapan semester memungkinkan identifikasi pola kinerja yang lebih persisten, bukan sekadar refleksi fluktuasi temporal. Keunggulan metodologis ini mampu mengatasi keterbatasan studi berbasis potongan lintang (*cross-sectional*) yang rentan terhadap bias situasional, seperti variasi karakteristik mata kuliah atau komposisi mahasiswa pada semester tertentu (Ballester, 2017; H. Li, 2014).

Salah satu temuan utama adalah adanya hubungan nonlinier antara intensitas beban mengajar dan kualitas pengajaran. Klaster *Excellence* menunjukkan kinerja yang sangat tinggi dan konsisten justru pada beban relatif ringan. Hal ini membuktikan bahwa kualitas pengajaran dipengaruhi pula oleh

kebijakan institusional terkait alokasi beban kerja (Miller et al., 2016; Xiong et al., 2016). Dalam kerangka *cognitive load*, peningkatan beban instruksional berpotensi melampaui kapasitas kognitif dosen dan berdampak pada kualitas interaksi pedagogis (Setiyorini & Frieyadie, 2020). Sebaliknya, Klaster *High Productivity* mengungkap adanya kelompok dosen dengan beban mengajar lebih yang tetap mampu mempertahankan kualitas pada tingkat “Baik”. Namun klaster ini paling rentan terhadap risiko kelelahan instruksional serta penurunan kualitas jangka panjang yang tidak selalu tercermin dalam skor evaluasi (Meo et al., 2020; Sauers et al., 2022).

Klaster *Underperformance* memberikan perspektif berbeda dengan menantang asumsi bahwa rendahnya kualitas pengajaran terutama disebabkan oleh beban kerja yang tinggi. Meskipun memiliki beban mengajar yang ringan, kelompok ini mencatat skor terendah dengan variabilitas tertinggi, yang mengindikasikan adanya persoalan mendasar dalam kompetensi pedagogis. Dengan demikian, kebijakan pengurangan beban kerja tanpa disertai intervensi peningkatan kapasitas pengajaran berpotensi tidak efektif (Antoniou et al., 2024; Sauers et al., 2021). Temuan ini kembali menegaskan bahwa pendekatan seragam berisiko menimbulkan konsekuensi yang tidak diinginkan dan tidak adil bagi kelompok dosen dengan karakteristik kinerja yang berbeda (Buschow et al., 2022; Dang et al., 2023; Humberstone et al., 2009)

SIMPULAN

Studi ini menunjukkan bahwa pemanfaatan *Educational Data Mining*, khususnya *K-Means Clustering*, memberikan wawasan yang jauh lebih mendalam mengenai pola kinerja dosen dibandingkan evaluasi konvensional berbasis rata-rata tunggal. Dengan data longitudinal delapan semester (2021–2024) dan 5.577 catatan evaluasi atas 493 dosen, studi ini memetakan kinerja dosen di UIN Sultanah Nahrasiyah Lhokseumawe secara lebih objektif. Validasi *Elbow Method* mengidentifikasi tiga tipologi: klaster *Excellence* (32,8%) dengan kualitas tertinggi dan stabilitas terbaik pada beban ringan; klaster *High Productivity* (36,1%) sebagai penopang utama pembelajaran namun mulai menunjukkan gejala kelelahan; dan

klaster *Underperformance* (31,1%) dengan kualitas terendah meski beban moderat, yang mengindikasikan kekurangan kompetensi pedagogis.

Konsisten dengan hasil analisis, uji korelasi *Pearson* dan *Spearman* tidak menemukan hubungan linier positif yang signifikan antara beban mengajar dan tingkat kepuasan mahasiswa; sebaliknya, hubungan keduanya bersifat nonlinier dan bergantung pada tipologi klaster.

Dari perspektif manajerial, pimpinan perguruan tinggi dapat memanfaatkan temuan ini untuk bertransformasi dari sekadar administrasi beban mengajar menuju pengelolaan tenaga pengajar berbasis data. Rekomendasi kebijakan meliputi retensi dan penguatan kapasitas dosen unggul, rasionalisasi beban bagi dosen produktivitas tinggi, serta peningkatan kompetensi pedagogis bagi dosen berkinerja rendah. Keterbatasan rasional penelitian ini terletak pada penggunaan variabel terstruktur sebagai dasar pemodelan. Aspek kualitatif berupa komentar terbuka dari mahasiswa tidak diikutsertakan dalam proses klasterisasi. Penelitian selanjutnya disarankan untuk mengintegrasikan klasterisasi data terstruktur dengan analisis sentimen berbasis *Natural Language Processing (NLP)* guna menghasilkan profil kinerja dosen yang lebih komprehensif.

DAFTAR PUSTAKA

- Agustian, H. (2018). Application of Product Moment Correlation and Complete Linkage Clustering Methods in Analyzing the Results of the Lecturer Questionnaire. *Conference Senatik STT Adisutjipto Yogyakarta*, 4. <https://doi.org/10.28989/senatik.v4i0.194>
- Ahmad, K., Qadir, J., Al-Fuqaha, A., Iqbal, W., Elhassan, A., Benhaddou, D., & Ayyash, M. (2020). *Data-Driven Artificial Intelligence in Education: A Comprehensive Review*. <https://doi.org/10.35542/osf.io/zvu2n>
- Akçapınar, G., Er, E., & Bayazıt, A. (2024). Decoding Video Logs: Unveiling Student Engagement Patterns in Lecture Capture Videos. *The International Review of Research in Open and Distributed Learning*, 25(2), 94–113. <https://doi.org/10.19173/irrodl.v25i2.7621>
- Alfaro, A. (2025). Analyzing Lecturer Performance Factors From Course Evaluation Surveys Using K-Means Clustering and C4.5 Classification. *Methodika Jurnal Teknik Informatika Dan Sistem Informasi*, 11(2), 1–6. <https://doi.org/10.46880/mtk.v11i2.4628>
- Algur, S. P., Bhat, P., & Kulkarni, N. (2016). Educational Data Mining: Classification Techniques for Recruitment Analysis. *International Journal of Modern Education and Computer Science*, 8(2), 59–65. <https://doi.org/10.5815/ijmecs.2016.02.08>

- Aljasmi, L. (2016). Using Color Coded Pattern Clustering Model Combined With Automated System for Courses Scheduling. *International Journal of Computing and Network Technology*, 4(3), 147–156. <https://doi.org/10.12785/ijcts/040305>
- Aljawarneh, S., & Lara, J. A. (2021). Data Science for Analyzing and Improving Educational Processes. *Journal of Computing in Higher Education*, 33(3), 545–550. <https://doi.org/10.1007/s12528-021-09299-7>
- Almasri, A., Çelebi, E., & Alkhawaldeh, R. S. (2019). EMT: Ensemble Meta-Based Tree Model for Predicting Student Performance. *Scientific Programming*, 2019, 1–13. <https://doi.org/10.1155/2019/3610248>
- Antoniou, F., Alghamdi, M. H., & Kawai, K. (2024). The Effect of School Size and Class Size on School Preparedness. *Frontiers in Psychology*, 15. <https://doi.org/10.3389/fpsyg.2024.1354072>
- Aulia, D., & Nurahman, N. (2023). Comparison Performance of K-Medoids and K-Means Algorithms In Clustering Community Education Levels. *Jurnal Nasional Pendidikan Teknik Informatika (Janapati)*, 12(2), 273–282. <https://doi.org/10.23887/janapati.v12i2.59789>
- Ayu, R. R. (2022). The Need Analysis of the Development of Animation Video-Based Learning Media on Physics Education Students. *Asian Journal of Healthy and Science*, 1(2), 45–49. <https://doi.org/10.58631/ajhs.v1i2.10>
- Ballester, C. P. M. (2017). More Experienced Professors, More Learning Opportunities? Relationships Between Instructor's Traits and Students' Academic Performance in Financial Accounting Courses of a Spanish University. *Ciencias Sociales Y Educación*, 6(12), 23–44. <https://doi.org/10.22395/csye.v6n12a2>
- Balovskyak, S., Derevyanchuk, O., Kravchenko, H., Ushenko, Y., & Hu, Z. (2023). Clustering Students According to their Academic Achievement Using Fuzzy Logic. *International Journal of Modern Education and Computer Science*, 15(6), 31–43. <https://doi.org/10.5815/ijmeecs.2023.06.03>
- Banerjee, R. (2025a). Analyzing Student Achievement With Clustering Techniques in Educational Analytics. *Panamerican Mathematical Journal*, 35(1s), 424–436. <https://doi.org/10.52783/pmj.v35.i4s.6014>
- Banerjee, R. (2025b). Understanding Student Performance Through Clustering in Educational Data Mining. *DXJB*, 37(1), 120–131. <https://doi.org/10.52783/dxjb.v37.183>
- Blasques, F., Hoogerkamp, M., Koopman, S., & Werve, I. (2020). Dynamic Factor Models with Clustered Loadings: Forecasting Education Flows using Unemployment Data. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3732269>
- Boliubash, N. (2019). Formation Of Masters' of It-Specialties Pedagogical Competency by Means of Educational Information Technologies. *Information Technologies and Learning Tools*, 71(3), 70. <https://doi.org/10.33407/itlt.v71i3.2545>
- Cao, X. (2024). Research on the Application of Clustering Algorithm in the Analysis of College English Achievement. 393–401. https://doi.org/10.2991/978-94-6463-417-4_36
- Chrisnanto, Y., & Abdullah, G. (2021). The uses of educational data mining in academic performance analysis at higher education institutions (case study at UNJANI). *Matrix*

- Jurnal Manajemen Teknologi Dan Informatika*, 11(1), 26–35.
<https://doi.org/10.31940/matrix.v11i1.2330>
- Condon, P., Exline, E., & Buckley, L. (2023). Data Literacy in the Social Sciences: Findings From a Local Study on Teaching With Quantitative Data in Undergraduate Courses. *Evidence Based Library and Information Practice*, 18(1), 61–75.
<https://doi.org/10.18438/ebliip30138>
- Custer, G., Diepen, L. T. A. van, & Seeley, J. (2021). Student Perceptions Towards Introductory Lessons in R. *Journal of Natural Resources and Life Sciences Education*, 50(2). <https://doi.org/10.1002/nse2.20073>
- Dang, T.-Q., Tan, G. W., Aw, E. C., Ooi, K., Metri, B., & Dwivedi, Y. K. (2023). How to Generate Loyalty in Mobile Payment Services? An Integrative Dual SEM-ANN Analysis. *International Journal of Bank Marketing*, 41(6), 1177–1206.
<https://doi.org/10.1108/ijbm-05-2022-0202>
- Devi, K., & Ratnoo, S. (2023). Cluster analysis of socio-economic factors and academic performance of school students. *Indonesian Journal of Electrical Engineering and Computer Science*, 31(3), 1568. <https://doi.org/10.11591/ijeecs.v31.i3.pp1568-1577>
- Dong, Y., & Dong, J. (2024). Evaluation and Analysis of Decision Support for College Fitness Teaching Under Embedded Data Mining Training. *International Journal of High Speed Electronics and Systems*, 34(02).
<https://doi.org/10.1142/s0129156424400858>
- Duan, L. (2024). Exploring the Status Quo and Teaching Mode of College English Education in the Information Age. *Applied Mathematics and Nonlinear Sciences*, 9(1). <https://doi.org/10.2478/amns-2024-2022>
- Eshach, H., & Kukliansky, I. (2016). Developing of an Instrument for Assessing Students' Data Analysis Skills in the Undergraduate Physics Laboratory. *Canadian Journal of Physics*, 94(11), 1205–1215. <https://doi.org/10.1139/cjp-2016-0308>
- Fahrillah, F., & Fatah, Z. (2025). Pengelompokan Data Nilai Siswa Madrasah Ta'hiliyah Menggunakan Metode K-Means Clustering. *Jissi*, 2(1), 53–59.
<https://doi.org/10.69714/0v1pkz05>
- Faisal, M., Suharmanto, S., & Prastyo, D. (2023). Implementasi Data Mining Menggunakan Metode K-Means Dalam Pengelompokan Bayi Penerima ASI Eksklusif. *Icit Journal*, 9(2), 198–207. <https://doi.org/10.33050/icit.v9i2.2953>
- Ferwana, I. (2019). Clustering Arabic Tweets for Saudi National Vision 2030. *International Journal of Advanced Trends in Computer Science and Engineering*, 8(1.4), 243–248. <https://doi.org/10.30534/ijatcse/2019/3781.42019>
- Ghosh, A. (2025a). Evaluating Clustering Algorithms for Educational Performance. *DXJB*, 37(1), 132–137. <https://doi.org/10.52783/dxjb.v37.184>
- Ghosh, A. (2025b). Optimising Educational Performance Through Clustering Algorithm Evaluation. *Communications on Applied Nonlinear Analysis*, 32(1), 598–606.
<https://doi.org/10.52783/cana.v32.6017>
- Hagge, M., Amin-Naseri, M., Gilbert, S. B., Jackman, J., Guo, E., Starns, G., & Faidley, L. (n.d.). *Decision-Based Learning for a Sophomore Level Thermodynamics Course*. 26.440.1-26.440.17. <https://doi.org/10.18260/p.23779>
- Hamoud, A. K., Hussein, M. K., Alhilfi, Z., & Sabr, R. H. (2021). Implementing Data-Driven Decision Support System Based on Independent Educational Data Mart.

- International Journal of Electrical and Computer Engineering (Ijece)*, 11(6), 5301. <https://doi.org/10.11591/ijece.v11i6.pp5301-5314>
- Hartama, D., & Oktaviani, S. (2025). Optimization of K-Means and K-Medoids Clustering Using Dbi Silhouette Elbow on Student Data. *Jurteksi (Jurnal Teknologi Dan Sistem Informasi)*, 11(2), 289–296. <https://doi.org/10.33330/jurteksi.v11i2.3531>
- Hartini, S., Bhakti, C., Hartanto, D., & Ghiffari, M. (2018). *Teacher Pedagogic Competency Development Model: A Literature Review*. <https://doi.org/10.2991/aecon-18.2018.40>
- Humberstone, M., Wood, B. A., Henkel, J., & Hines, J. W. (2009). Differentiating Between Expanded and Fault Conditions Using Principal Component Analysis. *Journal of Intelligent Manufacturing*, 23(2), 179–188. <https://doi.org/10.1007/s10845-009-0343-1>
- Iso-Markku, P., Drouard, G., Farrahi, V., Varjonen, A., Vähä-Ypyä, H., Vasankari, T., & Aaltonen, S. (2025). *Data-Driven Leisure-Time Physical Activity Trajectories of Over 46 Years and Their Associations with Cognition in Nonagenarians – A Cohort Study*. <https://doi.org/10.1101/2025.09.24.25336548>
- Jayakumar, M., Parameswari, R., & Akila, D. (2019). Enhancing the Quality Education using Predictive and Descriptive Data Mining Model. *International Journal of Recent Technology and Engineering*, 8(3), 6843–6847. <https://doi.org/10.35940/ijrte.c5795.098319>
- Joshi, Mr. A., & Kapse, A. S. (2024). A Framework to Optimize Student Performance Using Machine Learning. *International Journal of Recent Technology and Engineering*, 13(1), 27–30. <https://doi.org/10.35940/ijrte.a8052.13010524>
- Kansal, N., Solanki, V. K., & Kansal, V. (2021). *Educational Data Mining and Indian Technical Education System*. 1691–1706. <https://doi.org/10.4018/978-1-7998-3438-0.ch076>
- Karamustafaoglu, O., Tutar, M., & Sontay, G. (2017). Investigation of Science Teachers' Learning Styles With Various Variables. *Journal of Computer and Education Research*, 5(10), 255–280. <https://doi.org/10.18009/jcer.298669>
- Khomarudin, A. N., & Hidayat, A. (2022). K-Means Clustering Algorithm to See the Correlation of Tahfidz Activities With Student's Learning Outcomes. *Knowbase International Journal of Knowledge in Database*, 2(1), 01. <https://doi.org/10.30983/ijokid.v2i1.5672>
- Kolisnyk-Humenyuk, Y. (2024). Training Of Teachers Of Professional And Artistic Disciplines In Ukraine: Professional And Pedagogical Aspect. *Scientific Herald of Sivershchyna Series Education Social and Behavioural Sciences*, 2024(2), 97–109. <https://doi.org/10.32755/sjeduction.2024.02.097>
- Kwasi, D., & Gyimah, E. (2019). Using K-Means to Determine Learner Typologies for Project-Based Learning: A Case Study of the University of Education, Winneba. *International Journal of Computer Applications*, 178(43), 29–34. <https://doi.org/10.5120/ijca2019919320>
- Li, H. (2014). A Teaching Quality Evaluation Model Based on a Wavelet Neural Network Improved by Particle Swarm Optimization. *Cybernetics and Information Technologies*, 14(3), 110–120. <https://doi.org/10.2478/cait-2014-0037>

- Li, Y., Gou, J., & Fan, Z. (2019). Educational Data Mining for Students' Performance Based on Fuzzy C-means Clustering. *The Journal of Engineering*, 2019(11), 8245–8250. <https://doi.org/10.1049/joe.2019.0938>
- Liao, G., Cai, X., & Hu, J. (2024). Research on Course Teaching Evaluation and Academic Early Warning Based on Data Mining. 82. <https://doi.org/10.1117/12.3021466>
- Liu, Q. (2021). Influencing Factors of Learning Experience in Online Large-Class Teaching —Semantic Analysis Based on Students' Classroom Feedback Language Materials. *Converter*, 113–123. <https://doi.org/10.17762/converter.162>
- Liu, R. (2022). Data Analysis of Educational Evaluation Using K-Means Clustering Method. *Computational Intelligence and Neuroscience*, 2022, 1–10. <https://doi.org/10.1155/2022/3762431>
- Liu, Y. (2024). Designing and Evaluating Learning Effectiveness of Blended Teaching and Learning of College English Based on BOPPPS Modeling. *Applied Mathematics and Nonlinear Sciences*, 9(1). <https://doi.org/10.2478/amns-2024-1624>
- Manzanares, M. C. S., Sánchez, R. M., & Ochoa-Orihuel, J. (2021). Using Advanced Learning Technologies With University Students: An Analysis With Machine Learning Techniques. *Electronics*, 10(21), 2620. <https://doi.org/10.3390/electronics10212620>
- Meo, M. S., Kumar, B., Chughtai, S., Khan, V. J., Dost, M. K. B., & Nisar, Q. A. (2020). Impact of Unemployment and Governance on Poverty in Pakistan: A Fresh Insight From Non-Linear ARDL Co-Integration Approach. *Global Business Review*, 24(5), 1007–1024. <https://doi.org/10.1177/0972150920920440>
- Messakh, G., Hayati, M., & Sifriyani, S. (2023). Comparison K-Means and Fuzzy C-Means In Regencies/Cities Grouping Based on Educational Indicators. *Jurnal Varian*, 7(1), 99–114. <https://doi.org/10.30812/varian.v7i1.2879>
- Miller, A., Eather, N., Gray, S., Sproule, J., Williams, C., Gore, J., & Lubans, D. R. (2016). Can Continuing Professional Development Utilizing a Game-Centred Approach Improve the Quality of Physical Education Teaching Delivered by Generalist Primary School Teachers? *European Physical Education Review*, 23(2), 171–195. <https://doi.org/10.1177/1356336x16642716>
- Nama, G. F., Hakim, L., & Junaidi, J. (2019). Implementation of K Means Technique in Data Mining to Cluster Researchers Google Scholar Profile. *International Journal of Engineering and Advanced Technology*, 9(1), 3654–3660. <https://doi.org/10.35940/ijeat.a2708.109119>
- Nurdiansyah, D., Saidah, S., & Cahyani, N. (2023). DATA MINING STUDY FOR GROUPING ELEMENTARY SCHOOLS IN BOJONEGORO REGENCY BASED ON CAPACITY AND EDUCATIONAL FACILITIES. *Barekeng Jurnal Ilmu Matematika Dan Terapan*, 17(2), 1081–1092. <https://doi.org/10.30598/barekengvol17iss2pp1081-1092>
- Özbaş, M., & Mukhatayeva, D. (2018). Views of Pre-service Teachers Related to the Development of Pedagogical Formation Training and Design Competences. *Universal Journal of Educational Research*, 6(6), 1207–1215. <https://doi.org/10.13189/ujer.2018.060611>
- Pecuchová, J., & Drlík, M. (2022). Identification of Students With Similar Behavioural Patterns Using Clustering Techniques. 257–267. <https://doi.org/10.34916/el.2022.14.19>

- Qiao, Y. (2023). The combination application of Civic Education and Big Data Internet Technology based on the background of college education reform. *Applied Mathematics and Nonlinear Sciences*, 9(1). <https://doi.org/10.2478/amns.2023.2.00168>
- Ratnapala, I. P., Ragel, R., & Deegalla, S. (2014). *Students Behavioural Analysis in an Online Learning Environment Using Data Mining*. 1–7. <https://doi.org/10.1109/iciafs.2014.7069609>
- Ren, S. (2024). Integration and Optimization of Educational Teaching Resources in Colleges and Universities Based on Clustering Algorithm. *Jes*, 20(6s), 2156–2165. <https://doi.org/10.52783/jes.3130>
- Rianti, R., Andarsyah, R., & Awangga, R. (2024). Penerapan PCA dan Algoritma Clustering untuk Analisis Mutu Perguruan Tinggi di LLDIKTI Wilayah IV. *Nuansa Informatika*, 18(2), 67–77. <https://doi.org/10.25134/ilkom.v18i2.211>
- Rosadi, M. W. (2025). Clustering-Based Identification of Student Support Needs in Higher Education Transition. *Jusifo (Jurnal Sistem Informasi)*, 11(2), 133–140. <https://doi.org/10.19109/jusifo.v11i2.31031>
- Saleem, Q., Asif, A., & Hassan, M. T. (2018). Technical and Vocational Education Analytics Using Punjab TEVTA Students' Data. *Vfast Transactions on Software Engineering*, 37–48. <https://doi.org/10.21015/vtse.v13i1.495>
- Sauers, L. A., Hawes, K. E., & Juliano, S. A. (2021). *Non-Linear Relationships Between Density and Demographic Traits of Aedes*. <https://doi.org/10.21203/rs.3.rs-883857/v1>
- Sauers, L. A., Hawes, K. E., & Juliano, S. A. (2022). Non-Linear Relationships Between Density and Demographic Traits in Three Aedes Species. *Scientific Reports*, 12(1). <https://doi.org/10.1038/s41598-022-11909-y>
- Setiabudi, D. H., & Santoso, M. (2023). Effect of Students' Activities on Academic Performance Using Clustering Evolution Analysis. *Commit (Communication and Information Technology) Journal*, 17(2), 209–219. <https://doi.org/10.21512/commit.v17i2.9053>
- Setiyorini, T., & Frieyadie, F. (2020). Comparison of Linear Regressions and Neural Networks for Forecasting Electricity Consumption. *Jurnal Pilar Nusa Mandiri*, 16(2), 135–140. <https://doi.org/10.33480/pilar.v16i2.1459>
- Shaheen, M. N. ul K., & Rehman, Z. u. (2008). *Critical Skills for Computer Academicians - Course Proposal*. <https://doi.org/10.28945/3197>
- Shang, R., Ara, B., Zada, I., Nazir, S., Ullah, Z., & Khan, S. (2021). Analysis of Simple K-Mean and Parallel K-Mean Clustering for Software Products and Organizational Performance Using Education Sector Dataset. *Scientific Programming*, 2021, 1–20. <https://doi.org/10.1155/2021/9988318>
- Sidiq, D. M., Samudra, A. H., & Anwar, K. (2025). Evaluasi Kinerja Dosen. *Manajerial Jurnal Inovasi Manajemen Dan Supervisi Pendidikan*, 5(2), 434–443. <https://doi.org/10.51878/manajerial.v5i2.6157>
- Sihotang, R., Christianto, L., & Sugiyarto, A. (2023). *School Profiling in Indonesia Based on Scholastic Ability Using Cluster Analysis*. 22–37. https://doi.org/10.2991/978-2-38476-060-2_4

- Singelmann, L., & Ewert, D. (2022). Leveraging the Innovation-Based Learning Framework to Predict and Understand Student Success in Innovation. *Ieee Access*, 10, 36123–36139. <https://doi.org/10.1109/access.2022.3163744>
- Souza, A. A. d., Barcelos, T., Muñoz, R., Villarroel, R., & Silva, L. A. (2019). Data Mining Framework to Analyze the Evolution of Computational Thinking Skills in Game Building Workshops. *Ieee Access*, 7, 82848–82866. <https://doi.org/10.1109/access.2019.2924343>
- Spanias, A., Thiagarajan, J. J., Ramamurthy, K. N., Banavar, M. K., Ranganath, S., Zhang, X., Rajan, D., & Kalyanasundaram, G. (n.d.). *E-Book on DSP Theory With Interactive iOS, Java, and Android Simulations*. <https://doi.org/10.18260/1-2--19459>
- Srinivasulu, A., & Dakshayani, G. (2016). Clustering Large Data With Mixed Values Using Extended Fuzzy Adaptive Resonance Theory. *Indonesian Journal of Electrical Engineering and Computer Science*, 4(3), 617. <https://doi.org/10.11591/ijeecs.v4.i3.pp617-628>
- Sudirga, R. S. (2017). Segmentasi Dan Persepsi Mahasiswa Program Studi Manajemen Terhadap Proses Pembelajaran Universitas Bunda Mulia Dengan Metode Cluster Analysis. *Business Management Journal*, 9(1). <https://doi.org/10.30813/bmj.v9i1.682>
- Tan, M., & Qu, L. (2023). Evaluation of Oral English Teaching Quality Based on BP Neural Network Optimized by Improved Crow Search Algorithm. *Journal of Intelligent & Fuzzy Systems*, 45(6), 11909–11924. <https://doi.org/10.3233/jifs-222455>
- Teuling, N., Pauws, S., & Heuvel, E. (2021). *Clustering of longitudinal data: A tutorial on a variety of approaches*. <https://doi.org/10.48550/arxiv.2111.05469>
- Veretennikova, V., Шихова, O., Shikhov, Y., & Valeev, A. (2022). Diagnostics of Student Teachers' Pedagogical Thinking Based on a Competence Approach. *Education & Self Development*, 17(1), 233–254. <https://doi.org/10.26907/esd.17.1.18>
- Vispute, S. K., Ahirrao, J. J., & Potdar, S. (2021). COVID-19: A Survey on Change in Learning Method From Classroom to Virtual Using Educational Data Mining. *Ibmrd S Journal of Management & Research*, 3–6. <https://doi.org/10.17697/ibmrd/2021/v10i2/166797>
- Wang, J., & Zhang, W. (2020). Fuzzy Mathematics and Machine Learning Algorithms Application in Educational Quality Evaluation Model. *Journal of Intelligent & Fuzzy Systems Applications in Engineering and Technology*, 39(4), 5583–5593. <https://doi.org/10.3233/jifs-189039>
- Wang, Q. (2021). Application of the Intra Cluster, Characteristic of K-Means Clustering Method in English Score Analysis in Colleges. *Journal of Physics Conference Series*, 1941(1), 012001. <https://doi.org/10.1088/1742-6596/1941/1/012001>
- Wang, Z. (2022). Higher Education Management and Student Achievement Assessment Method Based on Clustering Algorithm. *Computational Intelligence and Neuroscience*, 2022, 1–10. <https://doi.org/10.1155/2022/4703975>
- Wardhani, N., & Gata, W. (2025). Improving the Academic Environment and Teaching Quality through Data Segmentation and Decision Trees. *Dinasti International Journal of Education Management and Social Science*, 6(6), 5227–5243. <https://doi.org/10.38035/dijemss.v6i6.5288>

- Wicaksono, D., Nisworo, S., & Nata, I. (2024). A Comparative Study of Students Graduation Analysis Using Classification Methods in Undergraduate Electrical Engineering Tidar University. *Jurnal Teknik Informatika*, 17(1), 51–60. <https://doi.org/10.15408/jti.v17i1.32132>
- Wise, M. J. (2017). Does It Matter if Students Have the Same Instructor for Both Lecture and Lab Sections? An Analysis of Introductory Biology Students. *Journal of the Scholarship of Teaching and Learning*, 17(1), 1–14. <https://doi.org/10.14434/v17i1.19583>
- Wu, R., & Wang, J. (2023). Digital Management of Teaching Cases in Colleges and Universities Based on Cluster Analysis. *International Journal of Emerging Technologies in Learning (Ijet)*, 18(10), 264–279. <https://doi.org/10.3991/ijet.v18i10.40241>
- Xiong, W., Min, G., Ye, T., & Hui, W. (2016). Education Training and Assessment for Medical Staff in Network Information Sharing Model. *Telkomnika (Telecommunication Computing Electronics and Control)*, 14(2A), 388. <https://doi.org/10.12928/telkomnika.v14i2a.4357>
- Yu, S., Yang, M., Wei, L., Hu, J.-S., Tseng, H., & Meen, T.-Hang. (2020). Combination of Self-Organizing Map and K-Means Methods of Clustering for Online Games Marketing. *Sensors and Materials*, 32(8), 2801. <https://doi.org/10.18494/sam.2020.2800>
- Yu, X. (2024). MOOC Teaching Management Scheme Based on Optimized SVM. <https://doi.org/10.3233/atde231247>
- Yunus, S., & Rosdiana, R. (2022). Sistem Penjaminan Mutu Internal (SPMI). <https://doi.org/10.31219/osf.io/vkmwr>
- Zainuddin, Z., & Risal, A. (2024). Balanced clustering for student admission school zoning by parameter tuning of constrained k-means. *Iaes International Journal of Artificial Intelligence (Ij-Ai)*, 13(2), 2301. <https://doi.org/10.11591/ijai.v13.i2.pp2301-2313>
- Zhou, N., Zhang, Z., & Li, J. (2020). Analysis on Course Scores of Learners of Online Teaching Platforms Based on Data Mining. *Ingénierie Des Systèmes D Information*, 25(5), 609–617. <https://doi.org/10.18280/isi.250508>
- Zhou, Y., Zeng, Z., & Wang, H. (2022). Using Spectral Clustering Association Algorithm Upon Teaching Big Data for Precise Education. *Mathematical Problems in Engineering*, 2022, 1–11. <https://doi.org/10.1155/2022/7214659>