

# Optimasi Pemilihan Metode Pengajaran Dosen Menggunakan Data Mining, dan Algoritma K-Means dalam Proses Bisnis Pendidikan

De Ali Farizal<sup>1</sup>, Franciskus Antonius Alijoyo<sup>2</sup>

<sup>1,2</sup> School of business and Information Technology, STMIK LIKMI Bandung, 40132 - Indonesia  
e-mail: alifarizal98@gmail.com<sup>\*1</sup>, Antonius.alijoyo@gmail.com<sup>\*2</sup>

## INFORMASI ARTIKEL

### Sejarah Artikel:

Diterima Redaksi : 03 Agustus 2024

Revisi Akhir : 05 November 2024

Diterbitkan Online : 30 November 2024

### Kata Kunci :

**Perguruan Tinggi, Metode Pembelajaran, Data Mining, K-Means.**

### Korespondensi :

Telepon / Hp : +62 (0265) 272727

E-mail : [email@afiliasi.ac.id](mailto:email@afiliasi.ac.id)

## A B S T R A K

Perguruan tinggi di Indonesia dihadapkan pada tantangan tingginya jumlah lulusan yang kurang memiliki kompetensi sesuai dengan kebutuhan dunia kerja, menciptakan kesenjangan antara lingkungan pendidikan dan pasar kerja. Hal ini menimbulkan kekhawatiran akan hambatan kemajuan bangsa. Salah satu penyebabnya adalah metode pembelajaran tradisional yang tidak mengakomodasi berbagai gaya belajar dan kebutuhan individual mahasiswa. Teknologi informasi dan komunikasi (TIK), khususnya data mining, menawarkan solusi untuk meningkatkan kualitas pendidikan dengan menganalisis karakteristik mahasiswa agar metode pembelajaran dapat disesuaikan dan materi lebih mudah dipahami. Data mining, atau Knowledge Discovery in Database atau yang disebut (KDD), adalah proses otomatis mengumpulkan dan menganalisis data untuk mengidentifikasi pola dengan teknik seperti pengelompokan, asosiasi, dan prediksi. Metode clustering K-Means, yang mengelompokkan data berdasarkan karakteristik serupa, dapat digunakan untuk menentukan metode pembelajaran paling efektif. Universitas Cipasung Tasikmalaya (Uncip) dapat memanfaatkan data mining dan algoritma K-Means untuk mengklasifikasikan dan memilih metode pengajaran yang optimal guna meningkatkan kompetensi lulusannya sesuai dengan kebutuhan dunia kerja

## 1. PENDAHULUAN

Pada saat ini pendidikan di perguruan tinggi di Negara Indonesia, dihadapkan dengan berbagai tantangan, salah satunya ialah tingginya angka lulusan yang kurang mempunyai kompetensi yang sesuai pada kebutuhan dunia kerja[1]. Hal tersebut menimbulkan kekhawatiran akan kesenjangan antara lingkungan pendidikan dan lingkungan kerja, yang dapat menghambat kemajuan bangsa[1]. Bambang Brodjonegoro, Menteri PPN/Kepala Bappenas, menyatakan bahwa tenaga kerja lulusan perguruan tinggi di Indonesia sama dengan tenaga kerja lulusan Sekolah Menengah Atas (SMA) di Negara Denmark, yang menunjukkan adanya kesenjangan antara kompetensi lulusan perguruan tinggi di Indonesia dengan kebutuhan pasar kerja internasional. (<https://indonesiadevelopmentforum.com/>)

Permasalahan di atas banyak faktor yang mempengaruhi kurangnya kompetensi yang di miliki lulusan perguruan tinggi, salah satunya kurang pasnya metode pembelajaran yang digunakan dalam pembelajaran di perguruan tinggi. Metode pembelajaran yang digunakan pada zaman dulu yang berfokus pada ceramah dan hafalan seringkali tidak sanggup mengakomodasi keragaman berbagai gaya belajar dan kebutuhan individual mahasiswa[2].

Pada pemilihan metode pembelajaran yang tepat diperlukan analisis yang kuat terhadap karakteristik setiap mahasiswa. Untuk itu, di Era yang sangat canggih ini, teknologi informasi dan komunikasi (TIK) menawarkan peluang untuk meningkatkan kualitas pendidikan tinggi Salah satunya penggunaan teknologi

data mining[3]. Data Mining dapat digunakan untuk menganalisis karakteristik setiap mahasiswa agar metode pembelajaran yang diterapkan dapat sesuai dan materi yang disampaikan dapat mudah dipahami[4].

Tujuan utama dari data mining adalah proses mendapatkan dan menganalisis informasi dari data dalam jumlah besar, serta data mining ini untuk menghasilkan informasi baru. Selain itu, data mining juga bisa disebut dengan Knowledge Discovery in Database (KDD), Sebab Data mining ini adalah proses yang secara otomatis mengumpulkan data dari sejumlah besar ruang memori untuk mengidentifikasi pola menggunakan teknik seperti pengelompokan, asosiasi dan prediksi[5]. Metode yang tepat untuk menganalisis karakteristik mahasiswa yaitu metode clustering. Salah satu algoritma yang bagus dalam melakukan clustering ialah Metode K-Means.

Metode K-Means merupakan metode pengelompokan data secara partisi dalam data mining, proses pemodelan tanpa supervisi. Metode ini mengumpulkan data menjadi lebih dari satu kelompok, setiap kelompok mempunyai karakteristik yang sama atau serupa satu dan yang lainnya, akan tetapi kelompok lain memiliki karakteristik yang berbeda. Metode K-Means dipergunakan untuk tujuan meningkatkan perbedaan antara kelompok data masing-masing serta meningkatkan perbedaan antara kelompok tersebut dengan kelompok yang lain[6].

Universitas Cipasung Tasikmalaya yang dikenal dengan nama Uncip adalah sebuah perguruan tinggi yang terletak di kabupaten tasikmalaya dan menjadi satu satunya universitas yang ada di kabupaten tasikmalaya,

Uncip masih belum mengetahui metode pengajaran yang tepat untuk menunjang pengajaran yang optimal dalam penyerapan ilmu sehingga mahasiswa memiliki kompetensi yang dibutuhkan dunia kerja.

Data mining dan algoritma k-means dapat digunakan untuk mengklasifikasikan metode pembelajaran mana yang paling efektif digunakan dengan menggunakan algoritma yang ada di dalam data mining[6].

## 2. LANDASAN TEORI

### 2.1 Metode Pengajaran

Mengajar adalah proses yang dilakukan oleh seorang guru atau dosen untuk menciptakan kegiatan belajar bagi siswanya. Metode pembelajaran adalah kumpulan elemen yang dikombinasikan secara efektif untuk meningkatkan kualitas pembelajaran[2].

Jika guru atau dosen dapat menemukan dan menerapkan pendekatan pengajaran yang tepat untuk menyampaikan materi tertentu, maka proses belajar mengajar akan lebih efektif. Pendidik tidak dapat menyelesaikan berbagai masalah belajar dengan menggunakan satu metode belajar-mengajar. Ini karena tidak ada satu pendekatan atau metode pembelajaran yang benar-benar cocok untuk semua aspek pembelajaran, setiap saat, dan dengan setiap mahasiswa[2].

### 2.2 Data Mining

Data mining dapat menggali data dan pengetahuan penting untuk menurunkan biaya dan meningkatkan keuntungan[7]. Data mining menggunakan berbagai teknik seperti matematika, statistik, kecerdasan buatan, dan pembelajaran mesin. Salah satu tugasnya adalah mengekstrak dan menemukan informasi penting yang saling terhubung dari berbagai database. Data mining atau penambangan data menghasilkan sesuatu yang dapat digunakan sebagai dasar untuk tindakan selanjutnya [7].

### 2.3 Algoritma K-Means Clustering

K-Means diklasifikasikan melalui analisis kelompok yang difokuskan pada pemartisian [8]. Kelompok K, juga disebut cluster, terdiri dari N objek pengamatan, dan setiap objek pengamatan memiliki kelompok yang sebanding dengan rata-rata atau mean. Untuk mengelompokkan data menjadi beberapa kelompok, algoritma K-Means digunakan dalam data mining clustering. Setelah kriteria ditetapkan untuk kelompok-kelompok yang dibentuk ini, data dikumpulkan menjadi satu cluster yang sesuai dengan kelompok tersebut. Setiap cluster memiliki centroid, juga dikenal sebagai titik pusat[8]. Adapun tahapan algoritma K-Means sebagai berikut :

- Menentukan jumlah cluster yang diharapkan dalam dataset.
- Pilih Centroid secara acak.
- Temukan jarak terpendek antara setiap data dan Centroid dengan menggunakan rumus

jarak geometris (d). Metode jarak euklide ini dapat digunakan untuk mengukur dua (dua) titik yang tidak sama. Rumus-rumus ini adalah:  $d_e = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2}$

Keterangan :

(x,y) = koordinat objek

(s, t) = koordinat Centroid

i = banyaknya objek

- Menghitung ulang titik Cluster menggunakan anggota Cluster terbaru. Pusat cluster rata-rata berisi semua data yang ada di dalamnya. Metode yang digunakan untuk menghitungnya adalah:

$$V_{ij} = \frac{1}{N_i} = \sum_{k=0}^{N_i} X_{kj}$$

Keterangan :

Vij = Centroid rata-rata dari Cluster ke - i untuk variabel ke - j

Ni = Jumlah anggota Cluster ke - i

i, k = Indeks dari Cluster

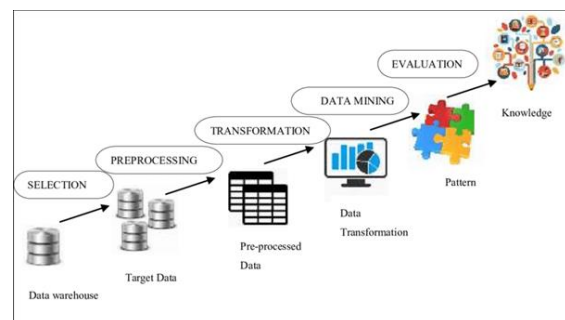
j = Indeks variabel

Xkj = Nilai data ke - k variabel ke - j untuk Cluster tersebut

## 3. METODE PENELITIAN

Pada penelitian ini, peneliti menggunakan metode pengumpulan data dengan Teknik kuesioner untuk memperoleh sumber data primer[4]. Selain Teknik kuesioner, peneliti juga melakukan tinjauan literature untuk memperoleh informasi mengenai penelitian penelitian terdahulu dan metode-metode yang digunakan untuk menunjang pada penelitian ini[9]. Berdasarkan hasil tinjauan literature yang telah dilakukan, metode clustering yang akan digunakan pada penelitian ini ialah metode K-Means Clustering.

Tahapan penelitian yang akan dilakukan pada penelitian ini mengacu pada methodology Knowledge Discovery in Database (KDD)[10]. Berikut gambar tahapan Methodology KDD dapat dilihat pada gambar 1.



Gambar 1 Methodology Knowledge Discovery in Database (KDD)  
(Sumber: www.sis.binus.ac.id)

### 1. Selection

Seleksi dalam proses KDD (Knowledge Discovery in Databases) adalah tahap di mana data yang relevan dan penting dipilih dari kumpulan data yang lebih besar untuk dianalisis[11]. Tujuan utama seleksi ini adalah untuk mengidentifikasi subset data yang paling relevan dan sesuai dengan tujuan analisis atau pemodelan yang

sedang dilakukan. Tahap seleksi ini penting untuk mengurangi kompleksitas dan memfokuskan upaya pada informasi yang paling berharga dalam proses KDD[11].

## 2. Preprocessing

Preprocessing adalah tahap awal di mana data mentah dari sumbernya diubah, dibersihkan, dan diubah agar siap untuk analisis lebih lanjut. Tujuannya adalah meningkatkan kualitas pada data, mengurangi suara, dan mempersiapkan data untuk digunakan dalam analisis. Preprocessing terdiri dari normalisasi atau standarisasi nilai-nilai data, penghapusan data yang tidak lengkap atau tidak relevan, dan transformasi data untuk mempersiapkan data untuk tahapan selanjutnya dalam proses KDD, seperti pemodelan dan evaluasi[10].

## 3. Transformation

Proses mengubah data mentah dari format aslinya menjadi format yang lebih sesuai untuk pemodelan atau analisis lebih lanjut disebut transformasi dalam KDD (Knowledge Discovery in Databases). Tujuan transformasi adalah untuk meningkatkan representasi data atau mengungkapkan pola yang tidak terlihat sebelumnya. Contoh transformasi termasuk menggabungkan data dari berbagai sumber menjadi format yang lebih seragam, mengubah representasi data untuk memenuhi persyaratan teknik analisis tertentu, atau normalisasi data untuk memperbaiki skala atau distribusi[10].

## 4. Data Mining

Data mining (Knowledge Discovery in Databases) adalah proses mengekstraksi pola atau pengetahuan penting dan mungkin tersembunyi dari data yang besar dan kompleks. Tujuan dari data mining adalah untuk menemukan hubungan, tren, atau pola yang tidak terlihat sebelumnya yang mampu membantu didalam menentukan keputusan atau prediksi di masa yang akan datang[12]. Berbagai metode data mining diterapkan. Pada penelitian ini, metode data mining yang akan digunakan untuk mengekstraksi data ialah algoritma K-Means Clustering[8].

## 5. Evaluation

Evaluasi adalah tahap di mana hasil proses pengolahan data dinilai untuk memastikan keakuratan, kegunaan, dan relevansinya untuk mencapai tujuan analisis[13]. Tujuan evaluasi adalah untuk mengevaluasi kualitas model atau pengetahuan yang ditemukan, mengevaluasi seberapa baik model dapat digeneralisasi ke data baru, dan menentukan apakah hasil dari proses KDD dapat menghasilkan nilai tambah yang signifikan. Evaluasi melibatkan penggunaan metrik kinerja yang sesuai dengan jenis model atau analisis yang digunakan. Ini juga dapat mencakup pengujian ulang data yang tidak terlihat sebelumnya atau penggunaan metode validasi silang untuk mengevaluasi kehandalan model[13].

## 4. HASIL DAN PEMBAHASAN

### 3.1 Selection

Berdasarkan hasil pengumpulan data yang dilakukan oleh peneliti melalui kuesioner, diperoleh data sebanyak 1480 dari total populasi 2029 Mahasiswa aktif di Universitas Cipasung Tasikmalaya. Adapun rincian data yang diperoleh dapat dilihat pada tabel 1.

Tabel 1. Rekap Data Responden

No.	Prodi	Jumlah Responden
1	Akuntansi	368
2	Bisnis Digital	136
3	Kewirausahaan	132
4	Manajemen	543
5	Pendidikan Biologi	51
6	Pendidikan Matematika	23
7	Rekayasa Perangkat Lunak	39
8	Sains Aktuaria	19
9	Sistem Informasi	169

Tahap selanjutnya dilakukan seleksi atribut pada data yang dikumpulkan. Adapun atribut yang digunakan yaitu Program Studi, Umur, Jenis Kelamin, Gaya Belajar.

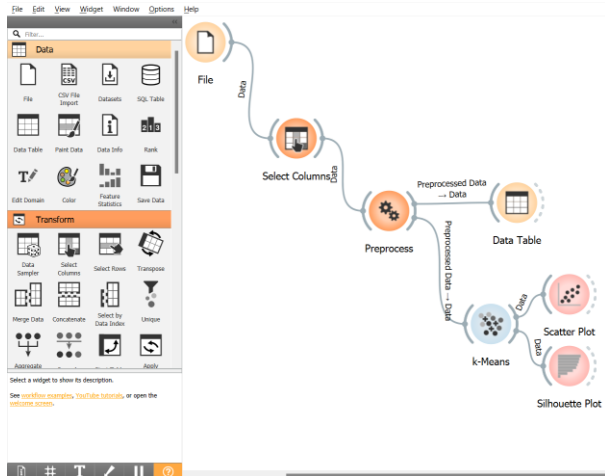
### 3.2 Preprocessing dan Transformation

Pada tahap selanjutnya dilakukan preprocessing data sebelum dilakukan pemodelan data mining[10]. Tahap preprocessing data ditransformasi. Transformasi data digunakan untuk menormalisasi data sehingga data siap untuk diolah pada tahap selanjutnya. Berikut dapat dilihat hasil transformasi data pada gambar 2.

Gambar 2. Normalisasi Data

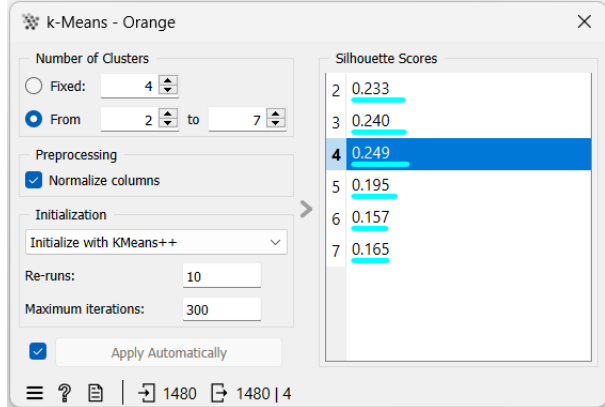
### 3.3 Data Mining

Pada tahap berikutnya, data yang telah siap diolah, dilakukan pemodelan dengan tools Orange Data Mining[14]. Berikut desain model yang dibangun pada Orange Data Mining dapat dilihat pada gambar 2.



Gambar 3. Desain Model pada Orange

Adapun nilai K yang digunakan sebagai jumlah cluster yang dibuat ditentukan berdasarkan nilai silhouette tertinggi rekomendasi dari tools Orange Data Mining. Berikut dapat dilihat pada gambar 4.

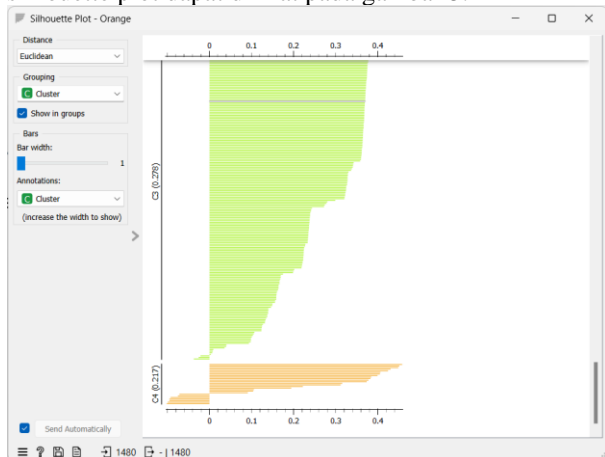


Gambar 4. Silhouette Score

Berdasarkan gambar 4 maka cluster yang dibangun terdiri dari 4 cluster yang menjadi nilai K.

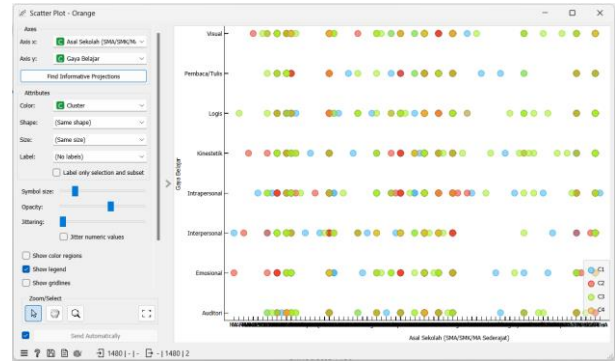
### 3.4 Evaluation

Pada tahap evaluasi, peneliti menggunakan Scatter Plot sebagai visual persebaran data dan Silhouette Plot untuk melihat grafik cluster yang dibangun. Berikut tampilan silhouette plot dapat dilihat pada gambar 5.



Gambar 5. Silhouette Plot

Dilihat pada gambar 5, cluster yang dibangun sudah mewakili data yang di analisis, hal ini dapat diidentifikasi dari grafik silhouette plot yang menunjukkan hanya sebagian kecil yang bernilai negatif. Sedangkan untuk sebaran data cluster dapat dilihat melalui scatter plot. Berikut gambar scatter plot dari cluster yang berhasil dibangun dapat dilihat pada gambar 6.



Gambar 6. Scatter Plot

Berdasarkan pada gambar 6, sebaran data gaya belajar merata dan menyeluruh sehingga cluster yang dibangun tidak kelihatan terlalu signifikan. Untuk lebih jelasnya, dapat dilihat pada tabel 2 anggota dari masing-masing cluster.

Tabel 2. Anggota Cluster

Cluster	Anggota	Asal Sekolah	Jenis Kelamin	Prodi
C1	25	MA = 7 SMA = 5 SMK = 13	L = 13 P = 12	AK = 10 BDG = 2 KWU = 2 MJ = 9 P.BIO = 1 SI = 1
C2	226	MA = 62 SMA = 30 SMK = 134	L = 124 P = 102	RPL = 39 AKT = 19 SI = 168
C3	521	MA = 126 SMA = 99 SMK = 296	P = 521	AK = 154 BDG = 57 KWU = 55 MJ = 216 P.BIO = 27 P. MTK = 12
C4	708	MA = 156 SMA = 104 SMK = 448	L = 708	AK = 204 BDG = 77 KWU = 75 MJ = 318 P.BIO = 23 P. MTK = 11

Adapun sebaran data cluster berdasarkan gaya belajar dapat dilihat pada gambar pivot tabel.



	Count	Cluster				Total
		C1	C2	C3	C4	
Auditori	2.0	30.0	62.0	76.0	170.0	
Emosional	3.0	25.0	66.0	103.0	197.0	
Interpersonal	2.0	35.0	61.0	92.0	190.0	
Intrapersonal	1.0	31.0	74.0	86.0	192.0	
Kinestetik	5.0	28.0	64.0	100.0	197.0	
Logis	1.0	25.0	69.0	90.0	185.0	
Pembaca/Tulis	6.0	27.0	56.0	69.0	158.0	
Visual	5.0	25.0	69.0	92.0	191.0	
<b>Total</b>	<b>25.0</b>	<b>226.0</b>	<b>521.0</b>	<b>708.0</b>	<b>1480.0</b>	

Gambar 7. Pivot Tabel

## 5. KESIMPULAN

Hasil penelitian mengarah pada kesimpulan bahwa algoritma K-Means dapat digunakan dalam analisis clustering data. Setelah dilakukan clustering diperoleh 4 cluster yang terbentuk. Sebaran data cluster cukup menyeluruh sehingga tidak terlihat signifikan. Dari 8 gaya belajar, 5 gaya belajar paling dominan yaitu Emosional, Kinestetik, Intrapersonal, Visual dan Interpersonal. Berdasarkan identifikasi tersebut maka metode pengajaran yang diusulkan yaitu kinestetik.

## 6. SUMBER PUSTAKA/RUJUKAN

- [1] G. S. Santyadiputra and K. Agustini, "Survey Kualitas Lulusan Jurusan Pendidikan Teknik Informatika," *J. Pendidik. Teknol. dan Kejur.*, vol. 13, no. 1, pp. 13–22, 2016, doi: 10.23887/jptk.v13i1.6843.
- [2] L. Baradja and A. A. Oktaviani, "Pengaruh Kecerdasan Emosional (Eq) Mahasiswa Dan Metode Pengajaran Dosen Terhadap Tingkat Pemahaman Akuntansi," *J. Akunt. Trisakti*, vol. 8, no. 1, pp. 41–50, 2021, doi: 10.25105/jat.v8i1.8669.
- [3] F. Antonius and A. Saepudin, "The Impact of IoT on The Storing Process of Leather Raw Material," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 7, no. 3, pp. 422–429, 2023, doi: 10.29207/resti.v7i3.4427.
- [4] E. Sabna and M. Muhandi, "Penerapan Data Mining Untuk Memprediksi Prestasi Akademik Mahasiswa Berdasarkan Dosen, Motivasi, Kedisiplinan, Ekonomi, dan Hasil Belajar," *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 2, no. 2, p. 41, 2016, doi: 10.24014/coreit.v2i2.2392.
- [5] R. Muliono and Z. Sembiring, "Data Mining Clustering Menggunakan Algoritma K-Means Untuk Klasterisasi Tingkat Tridarma Pengajaran Dosen," *CESS (Journal Comput. Eng. Syst. Sci.*, vol. 4, no. 2, pp. 2502–714, 2019.
- [6] L. Fimawahib and E. Rouza, "Penerapan K-Means Clustering pada Penentuan Jenis Pembelajaran di Universitas Pasir Pengaraian," *INOVTEK Polbeng - Seri Inform.*, vol. 6, no. 2, p. 234, 2021, doi: 10.35314/isi.v6i2.2096.
- [7] R. S. J. D. Baker and K. Yacef, "The State of Educational Data Mining in 2009: A Review and Future Visions," *J. Educ. Data Min.*, vol. 1, no. 1, pp. 3–16, 2009.
- [8] S. Huang, Z. Kang, Z. Xu, and Q. Liu, "Robust deep k-means: An effective and simple method for data clustering," *Pattern Recognit.*, vol. 117, p. 107996, 2021, doi: 10.1016/j.patcog.2021.107996.
- [9] S. Anastassia Amellia Kharis and A. Haqqi Anna Zili, "Learning Analytics dan Educational Data Mining pada Data Pendidikan," *J. Ris. Pembelajaran Mat. Sekol.*, vol. 6, pp. 12–20, 2022.
- [10] N. L. P. P. Dewi, I. N. Purnama, and N. W. Utami, "Penerapan Data Mining Untuk Clustering Penilaian Kinerja Dosen Menggunakan Algoritma K-Means (Studi Kasus: STMIK Primakara)," *J. Ilm. Teknol. Inf. Asia*, vol. 16, no. 2, p. 105, 2022, doi: 10.32815/jitika.v16i2.761.
- [11] M. Zubair, M. A. Iqbal, A. Shil, M. J. M. Chowdhury, M. A. Moni, and I. H. Sarker, "An Improved K-means Clustering Algorithm Towards an Efficient Data-Driven Modeling," *Ann. Data Sci.*, 2022, doi: 10.1007/s40745-022-00428-2.
- [12] R. Panthong and T. Wongkanthiya, "Analysis of clustering and association using data mining technique for elderly health condition dataset," *IAES Int. J. Artif. Intell.*, vol. 12, no. 4, pp. 1774–1783, 2023, doi: 10.11591/ijai.v12.i4.pp1774-1783.
- [13] A. Firdiansyah, I. Al Ikrom, M. Khamdanni, and W. C. Utomo, "Pemanfaatan Data Mining Untuk Memprediksi Kelulusan Mata Kuliah dan Referensi Strategi Pembelajaran," *Semin. Nas. Teknol. Sains*, vol. 3, no. 1, pp. 338–344, 2024, doi: 10.29407/stains.v3i1.4338.
- [14] R. Ratra and P. Gulia, "Experimental evaluation of open source data mining tools (WEKA and orange)," *Int. J. Eng. Trends Technol.*, vol. 68, no. 8, pp. 30–35, 2020, doi: 10.14445/22315381/IJETT-V68I8P206S.