

## ANALISIS TRACER STUDY MENGGUNAKAN K-MEANS CLUSTERING UNTUK EVALUASI KURIKULUM

Berta Erwin SLAM<sup>1</sup>, Feri Irawan<sup>2</sup>, Nolan Efranda<sup>3</sup>,  
Rifaldi Herikson<sup>4</sup>, Siti Syifa Alifah<sup>5</sup>

<sup>1,2,3,4</sup> Universitas Maritim Raja Ali Haji, Kepulauan Riau

<sup>5</sup>SMA Negeri 1 Bengkulu Utara, Bengkulu Utara

e-mail: <sup>1</sup>bertaerwinslam@umrah.ac.id, <sup>2</sup>feriirawan@umrah.ac.id,

<sup>3</sup>nolanefranda@umrah.ac.id, <sup>4</sup>rifaldiherikson@umrah.ac.id,

<sup>5</sup>siti.syifa1042@guru.sma.belajar.id

**Abstract:** *Tracer study is an important instrument for evaluating graduate quality and curriculum relevance to labor market needs. This study aims to analyze tracer study data using the K-Means Clustering method to group alumni characteristics and support curriculum evaluation. The dataset consists of 100 records with a relatively balanced composition, including variables such as waiting time for employment, job-field relevance, initial salary, and graduate competencies. The analysis process includes preprocessing, clustering, and evaluation using Silhouette Score, Davies-Bouldin Index, and Calinski-Harabasz Index. The results show that the data are grouped into three clusters representing different levels of job readiness. A Silhouette Score of 0.17, Davies-Bouldin Index of 1.736, and Calinski-Harabasz Index of 23.5 indicate low to moderate clustering quality due to data heterogeneity and overlap between clusters. The findings highlight the need for curriculum improvement, particularly in enhancing technical and soft skills. Future research is recommended to use larger and more balanced datasets, include more diverse variables, and compare clustering methods to improve analysis quality.*

**Keywords:** *K-Means Clustering, Tracer Study, Curriculum Evaluation, Data Mining.*

**Abstrak:** Tracer study merupakan instrumen penting dalam mengevaluasi kualitas lulusan dan relevansi kurikulum terhadap kebutuhan dunia kerja. Penelitian ini bertujuan untuk menganalisis data tracer study menggunakan metode K-Means Clustering untuk mengelompokkan karakteristik alumni serta mendukung evaluasi kurikulum. Data yang digunakan sebanyak 100 data dengan komposisi relatif seimbang, meliputi variabel waktu tunggu kerja, kesesuaian bidang kerja, pendapatan awal, serta kompetensi lulusan. Proses analisis meliputi preprocessing, clustering, dan evaluasi menggunakan Silhouette Score, Davies-Bouldin Index, dan Calinski-Harabasz Index. Hasil penelitian menunjukkan bahwa data terbagi menjadi tiga cluster dengan tingkat kesiapan kerja yang berbeda. Nilai Silhouette Score sebesar 0,17, Davies-Bouldin Index sebesar 1,736, dan Calinski-Harabasz Index sebesar 23,5 menunjukkan kualitas clustering rendah hingga sedang, yang disebabkan oleh sifat data yang heterogen dan adanya overlap antar cluster. Hasil ini mengindikasikan perlunya perbaikan kurikulum, khususnya pada peningkatan kompetensi teknis dan soft skills. Penelitian selanjutnya disarankan menggunakan data yang lebih banyak dan lebih seimbang, menambahkan variabel yang lebih beragam, serta membandingkan metode clustering untuk meningkatkan kualitas analisis.

**Kata Kunci:** K-Means Clustering, Tracer Study, Evaluasi Kurikulum, Data Mining

### PENDAHULUAN

Tracer study merupakan salah satu

instrumen penting yang digunakan oleh institusi pendidikan untuk mengetahui kondisi lulusan setelah menyelesaikan

pendidikan, baik dalam hal keterserapan kerja, kesesuaian bidang pekerjaan, maupun perkembangan karier alumni. Informasi yang diperoleh dari tracer study dapat menjadi dasar dalam mengevaluasi kualitas lulusan serta relevansi kurikulum terhadap kebutuhan dunia kerja. Selain itu, tracer study juga berfungsi sebagai sarana evaluasi dalam meningkatkan mutu pendidikan, memperbaiki proses pembelajaran, serta mendukung proses akreditasi institusi pendidikan (Afriyadi et al., 2025; Mashuri & Assegaff, 2023; Septiana et al., 2023).

Namun demikian, implementasi tracer study pada berbagai institusi pendidikan masih menghadapi berbagai kendala, terutama dalam hal pengelolaan dan pemanfaatan data alumni. Banyak institusi yang masih mengandalkan sistem konvensional seperti kuesioner manual atau penggunaan platform sederhana yang belum terintegrasi secara optimal (Putra & Elmunsyah, 2025; Jannah, 2025). Pengembangan sistem informasi tracer study berbasis web telah dilakukan untuk mengatasi permasalahan tersebut, yang terbukti mampu meningkatkan efisiensi dalam pengumpulan, pengolahan, serta penyajian data alumni (Lutfi et al., 2024; Sipayung et al., 2025; Hamid et al., 2021). Meskipun demikian, sebagian besar sistem yang dikembangkan masih berfokus pada penyimpanan dan penyajian data secara deskriptif, sehingga belum mampu memberikan analisis yang mendalam untuk mendukung pengambilan keputusan akademik.

Seiring dengan perkembangan teknologi informasi, pendekatan data mining mulai banyak digunakan untuk mengolah data dalam jumlah besar dan menghasilkan informasi yang lebih bermakna. Salah satu metode yang sering digunakan dalam data mining adalah K-Means Clustering, yaitu metode pengelompokan data berdasarkan kemiripan karakteristik sehingga data yang memiliki pola serupa akan berada dalam satu kelompok yang sama (Febrianti et al., 2021). Metode ini telah banyak diterapkan dalam berbagai

penelitian, termasuk dalam pengelompokan data alumni untuk mengetahui pola karakteristik lulusan serta tingkat kepuasan terhadap layanan pendidikan (Ramadhan et al., 2026). Selain itu, K-Means juga digunakan dalam analisis keterkaitan antara pemilihan mata kuliah dengan karier alumni untuk menghasilkan rekomendasi akademik yang lebih tepat (Maulana & Fathoni, 2025).

Lebih lanjut, penelitian lain menunjukkan bahwa penerapan K-Means Clustering dapat membantu dalam mengidentifikasi pola data alumni secara lebih sistematis dengan memanfaatkan teknik evaluasi seperti Elbow Method dan Silhouette Score untuk menentukan jumlah cluster yang optimal (Maulana & Fathoni, 2025; Ramadhan et al., 2026). Pendekatan ini memungkinkan pengelompokan data alumni berdasarkan berbagai indikator, seperti kompetensi, kepuasan, dan kesiapan kerja, sehingga menghasilkan informasi yang lebih terstruktur dibandingkan dengan analisis deskriptif konvensional.

Meskipun demikian, pemanfaatan metode K-Means dalam penelitian tracer study masih memiliki keterbatasan. Sebagian besar penelitian hanya berfokus pada proses pengelompokan data tanpa mengaitkan hasil cluster dengan pengambilan keputusan yang lebih luas, khususnya dalam evaluasi kurikulum. Padahal, data tracer study yang telah dikelompokkan dapat memberikan informasi penting mengenai pola karakteristik lulusan, seperti kesesuaian bidang pekerjaan, waktu tunggu kerja, serta kompetensi yang dimiliki, yang seharusnya dapat dimanfaatkan sebagai dasar dalam pengembangan dan perbaikan kurikulum. Dengan demikian, terdapat kesenjangan antara hasil analisis data dengan implementasinya dalam konteks evaluasi akademik.

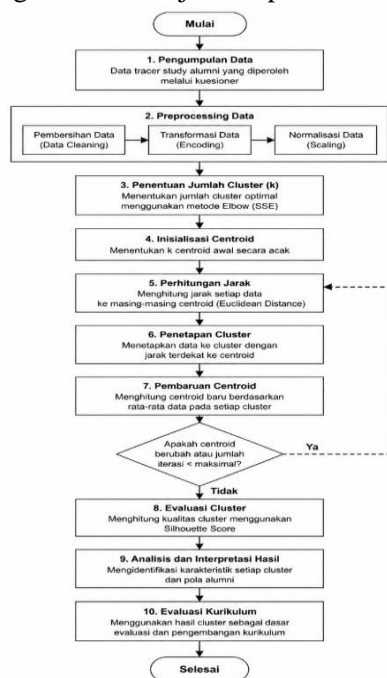
Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk menganalisis data tracer study menggunakan metode K-Means Clustering guna mengelompokkan

karakteristik alumni berdasarkan variabel tertentu. Hasil pengelompokan tersebut kemudian diinterpretasikan sebagai profil alumni yang dapat digunakan sebagai dasar dalam evaluasi kurikulum. Dengan pendekatan ini, penelitian tidak hanya berfokus pada proses clustering, tetapi juga pada pemanfaatan hasil analisis untuk mendukung pengambilan keputusan dalam pengembangan kurikulum yang lebih relevan dengan kebutuhan dunia kerja.

## METODE

Metode penelitian yang digunakan dalam penelitian ini adalah pendekatan data mining dengan algoritma K-Means Clustering untuk mengelompokkan data tracer study alumni berdasarkan kemiripan karakteristik (Febrianti et al., 2021; Ramadhan et al., 2026). Pendekatan ini bertujuan untuk menghasilkan pola data yang dapat digunakan dalam evaluasi kurikulum.

Alur penelitian dilakukan secara sistematis mulai dari pengumpulan data hingga interpretasi hasil clustering, sebagaimana ditunjukkan pada Gambar 1.



Gambar 1 Flowchart Metode Penelitian

Data yang digunakan merupakan data tracer study alumni yang diperoleh melalui kuesioner. Variabel yang digunakan meliputi waktu tunggu kerja, kesesuaian bidang kerja, pendapatan awal, serta kompetensi lulusan seperti kemampuan teknis, komunikasi, kerja sama, dan pemecahan masalah (Afriyadi et al., 2025). Data yang diperoleh kemudian melalui tahap preprocessing yang meliputi pembersihan data (data cleaning), transformasi data (encoding), serta normalisasi data (scaling) untuk memastikan kualitas data sebelum dilakukan proses clustering (Maulana & Fathoni, 2025).

Penentuan jumlah cluster (k) dilakukan menggunakan metode Elbow dengan menghitung nilai Sum of Squared Error (SSE) untuk setiap jumlah cluster. Selanjutnya, proses K-Means Clustering dilakukan secara iteratif dengan tahapan inisialisasi centroid, perhitungan jarak data terhadap centroid menggunakan Euclidean Distance, pengelompokan data ke cluster terdekat, serta pembaruan centroid hingga mencapai kondisi konvergen (Febrianti et al., 2021).

Fungsi objektif K-Means dirumuskan sebagai berikut:

$$J = \sum_{i=1}^k \sum_{j=1}^n \|x_j - c_i\|^2 \quad (1)$$

Perhitungan jarak menggunakan Euclidean Distance dirumuskan sebagai berikut:

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

Pembaruan centroid dilakukan dengan menghitung rata-rata data dalam setiap cluster menggunakan rumus:

$$c_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_j \quad (3)$$

Penentuan jumlah cluster optimal menggunakan metode Elbow dihitung dengan rumus:

$$SEE = \sum_{i=1}^k \sum_{x \in c_i} \|x - c_i\|^2 \quad (4)$$

Untuk mengevaluasi kualitas hasil clustering, digunakan Silhouette Score yang dirumuskan sebagai berikut:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (5)$$

Berdasarkan Gambar 1, setelah proses clustering selesai, dilakukan evaluasi cluster menggunakan Silhouette Score untuk menilai kualitas pengelompokan. Tahap akhir adalah analisis dan interpretasi hasil clustering untuk mengidentifikasi karakteristik masing-masing cluster. Hasil analisis ini digunakan sebagai dasar dalam evaluasi kurikulum dengan melihat pola distribusi alumni dan mengidentifikasi aspek pembelajaran yang perlu diperbaiki (Ramadhan et al., 2026).

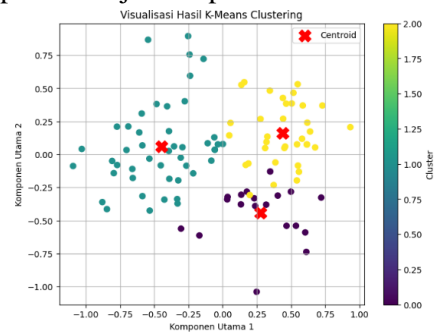
## HASIL DAN PEMBAHASAN

Data yang digunakan dalam penelitian ini merupakan data tracer study alumni sebanyak 100 data. Variabel yang digunakan meliputi waktu tunggu kerja, kesesuaian bidang kerja, pendapatan awal, serta kompetensi lulusan yang terdiri dari kompetensi teknis, komunikasi, kerja sama, dan kemampuan problem solving.

Berdasarkan karakteristik data, terlihat adanya variasi nilai antar alumni, terutama pada waktu tunggu kerja yang berkisar antara cepat hingga cukup lama, serta tingkat kompetensi yang bervariasi dari rendah hingga tinggi. Variasi ini menunjukkan bahwa data memiliki potensi untuk dikelompokkan ke dalam beberapa kategori berdasarkan tingkat kesiapan kerja lulusan.

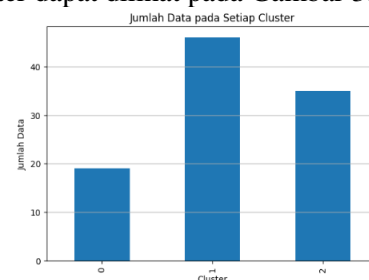
Hasil pengolahan data menggunakan metode K-Means Clustering dengan jumlah cluster sebanyak 3 menghasilkan distribusi data yang tidak sepenuhnya seimbang, yaitu 19 data pada cluster 0, 46 data pada cluster 1, dan 35 data pada cluster 2. Untuk memperjelas hasil pengelompokan,

dilakukan visualisasi menggunakan Principal Component Analysis (PCA) seperti ditunjukkan pada Gambar 2.



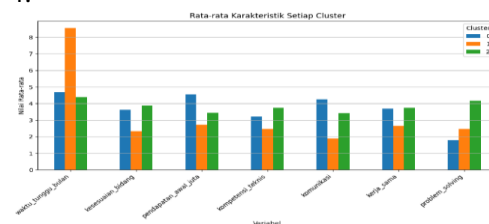
**Gambar 2 Visualisasi Hasil K-Means Clustering Menggunakan PCA**

Berdasarkan Gambar 2, terlihat bahwa data membentuk tiga kelompok, namun terdapat beberapa titik yang saling berdekatan antar cluster. Hal ini menunjukkan bahwa pemisahan cluster tidak sepenuhnya jelas dan terdapat overlap antar kelompok, yang mencerminkan kompleksitas data tracer study. Distribusi jumlah data pada setiap cluster dapat dilihat pada Gambar 3.



**Gambar 3 Distribusi Jumlah Data pada Setiap Cluster**

Berdasarkan Gambar 3, terlihat bahwa cluster 1 memiliki jumlah data paling banyak dibandingkan cluster lainnya. Untuk melihat perbedaan karakteristik antar cluster, ditampilkan rata-rata nilai tiap variabel pada Gambar 4.



**Gambar 4 Rata-rata Karakteristik pada Setiap Cluster**

Berdasarkan Gambar 4, setiap cluster menunjukkan karakteristik yang berbeda. Cluster 1 memiliki waktu tunggu kerja paling lama ( $\pm 8,54$  bulan), kesesuaian bidang rendah ( $\pm 2,33$ ), serta pendapatan dan kompetensi yang relatif rendah. Cluster 0 menunjukkan kondisi cukup baik dengan waktu tunggu sedang ( $\pm 4,68$  bulan), kesesuaian bidang cukup tinggi ( $\pm 3,63$ ), dan pendapatan relatif tinggi ( $\pm 4,54$  juta), namun masih memiliki kelemahan pada aspek tertentu seperti problem solving. Sementara itu, cluster 2 menunjukkan kondisi yang lebih stabil dengan kesesuaian bidang tinggi ( $\pm 3,86$ ) dan kompetensi yang baik, meskipun waktu tunggu kerja masih berada pada tingkat sedang ( $\pm 4,40$  bulan).

Evaluasi clustering menunjukkan nilai Silhouette Score sebesar 0,17, Davies-Bouldin Index sebesar 1,736, dan Calinski-Harabasz Index sebesar 23,5. Nilai tersebut mengindikasikan bahwa kualitas clustering berada pada kategori rendah hingga sedang, yang menunjukkan adanya overlap antar cluster. Kondisi ini disebabkan oleh karakteristik data yang bersifat heterogen serta adanya kemiripan nilai antar variabel, sehingga beberapa data memiliki kedekatan dengan lebih dari satu cluster. Selain itu, rentang nilai antar variabel yang tidak terlalu kontras menyebabkan batas antar cluster menjadi kurang tegas. Hal ini wajar terjadi pada data tracer study yang merepresentasikan kondisi nyata lulusan, di mana karakteristiknya tidak selalu dapat dipisahkan secara jelas. Meskipun demikian, hasil clustering tetap mampu menggambarkan pola umum dalam data dan memberikan informasi yang relevan untuk analisis lebih lanjut.

Berdasarkan hasil tersebut, cluster 1 menunjukkan perlunya perbaikan kurikulum, khususnya pada peningkatan kompetensi teknis dan soft skills. Cluster 0 menunjukkan perlunya penyempurnaan agar kompetensi lebih merata, sedangkan cluster 2 menunjukkan bahwa kurikulum sudah cukup relevan namun masih perlu peningkatan pada aspek kesiapan kerja. Dengan demikian, K-Means Clustering

dapat digunakan sebagai dasar evaluasi kurikulum berbasis data.

## SIMPULAN

Berdasarkan hasil penelitian, metode K-Means Clustering mampu mengelompokkan data tracer study alumni ke dalam tiga cluster dengan karakteristik tingkat kesiapan kerja yang berbeda, yaitu rendah, sedang, dan relatif baik. Hasil evaluasi menunjukkan nilai Silhouette Score sebesar 0,17, Davies-Bouldin Index sebesar 1,736, dan Calinski-Harabasz Index sebesar 23,5, yang mengindikasikan bahwa kualitas clustering berada pada kategori rendah hingga sedang. Kondisi ini disebabkan oleh karakteristik data yang bersifat heterogen serta adanya kemiripan nilai antar variabel sehingga terjadi overlap antar cluster dan batas antar kelompok menjadi kurang tegas. Meskipun demikian, hasil pengelompokan tetap dapat digunakan sebagai dasar dalam evaluasi kurikulum, di mana masih terdapat lulusan dengan kesiapan kerja yang belum optimal, sehingga diperlukan perbaikan dan penyempurnaan kurikulum, khususnya pada peningkatan kompetensi teknis dan soft skills. Untuk penelitian selanjutnya, disarankan menggunakan data dengan jumlah yang lebih banyak dan komposisi yang lebih seimbang, menambahkan variabel yang lebih beragam, serta membandingkan dengan metode clustering lain atau mengembangkan pendekatan berbasis kecerdasan buatan guna menghasilkan analisis yang lebih komprehensif dan rekomendasi kurikulum yang lebih adaptif.

## DAFTAR PUSTAKA

Afriyadi, M. R., Putra, F., & Purnomo, W. A. (2025). Web-based tracer study system design for the Faculty of Computer Science, University of Dharmas Indonesia. *Jurnal Ilmiah*

- Sistem Informasi (JUSI), 4(2), 308–319.  
<https://doi.org/10.51903/phmead65>
- Febrianti, S., Munawir, M., & Fitria, L. (2021). Penerapan metode K-means clustering terhadap alumni berdasarkan kuesioner tracer study. *Journal of Informatics and Computer Science*, 7(2), 117–122.
- Hamid, M. A., Aribowo, D., & Anggraini, R. (2021). Design and development of alumni career information system using PHP MySQL. *ELINVO (Electronics, Informatics, and Vocational Education)*, 6(1), 81–89.  
<https://doi.org/10.21831/elinvo.v6i1.30200>
- Jannah, W. (2025). Sistem informasi tracer study alumni sekolah menengah kejuruan (SMK) Mandiri. *ADA Journal of Information System Research*, 2(3), 111–120.  
<https://doi.org/10.64366/adajisr.v2i3.91>
- Lutfi, N. L., Lutfi, A., & Baijuri, A. (2024). Sistem informasi penelusuran alumni (tracer study) berbasis web pada Fakultas Sains dan Teknologi Universitas Ibrahimy. *Seminastika*, 5(1), 61–70.  
<https://doi.org/10.47002/seminastika.v5i1.793>
- Mashuri, M., & Assegaff, S. (2023). Sistem informasi alumni (tracer study) berbasis web pada SMA Negeri 2 Bungo. *Manajemen Sistem Informasi*, 8(3), 449–456.
- Maulana, R., & Fathoni, A. (2025). Analisis clustering rekomendasi mata kuliah peminatan berdasarkan karir alumni menggunakan K-means. *JOISIE (Journal of Information Systems and Informatics Engineering)*, 9(1), 109–122.  
<https://doi.org/10.35145/joisie.v9i1.4954>
- Putra, D. M., & Elmunsyah, H. (2025). Design of an alumni tracer study information system for vocational high schools. *Letters in Information Technology Education (LITE)*, 8(2), 80–89.
- Ramadhan, M. I., Nazir, A., Irsyad, M., Sanjaya, S., & Syafria, F. (2026). Clustering analysis using the K-means method to identify alumni satisfaction pattern. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 6(1), 31–41.  
<https://doi.org/10.57152/malcom.v6i1.2401>
- Septiana, Y., Fitriani, L., Hawariyan, F., Kurniawati, R., & Ulfa, R. L. (2023). Rancang bangun sistem informasi tracer study alumni berbasis website. *Jurnal Algoritma*, 20(1), 11–21.
- Sipayung, A., Utari, C. T., & Sabila, P. C. (2025). Sistem informasi tracer study berbasis web bagi alumni Fakultas Sains dan Teknologi. *Jurnal Persegi Bulat*, 3(2), 36–45.