Journal of Technology and Informatics (JoTI)
Vol. 8, No. 1, Bulan April Tahun 2026
P-ISSN 2721-4842
E-ISSN 2686-6102

K-Means Algorithm Application for Clustering Recent University Graduates According to Work Readiness Indicators

Wigananda Firdaus Putra Aditya[1*], Agussalim[2], Rizky Parlika[3],
[1*,2,3]Master of Information Technology, Universitas Pembangunan Nasional "Veteran" Jawa Timur, Surabaya, Indonesia
e-mail: 23066020021@student.upnjatim.ac.id[1*], agussalim.si@upnjatim.ac.id[2], rizkyparlika.if@upnjatim.ac.id[3]

Article Information

Article History:
Received : 7 November 2025
Revised : 14 November 2025
Accepted : 28 November 2025
Published : 27 April 2026

***Correspondence:**
23066020021@student.upnjatim.ac.id

Keywords:
K-Means, Work-readiness, graduate employability, Davies-Bouldin Index, Elbow method, Machine Learning

10.37802/joti.v8i1.1281

Journal of Technology and Informatics (JoTI)
P-ISSN 2721-4842
E-ISSN 2686-6102
https://e-journals.dinamika.ac.id/index.php/joti

***Abstract*:**
Graduate work-readiness segmentation is essential for data-driven career services in universities. This study applies K-Means clustering to tracer-study data using four input indicators: GPA (IPK), TOEFL, soft-skill points (SSKM), and study duration, while employment status and waiting time are treated as external outcomes. Records from 669 graduates (2020–2023) were preprocessed via deduplication, range checks, and z-score standardization. The number of clusters was determined data-driven over K=2–10 using the Elbow Method (SSE) and Davies–Bouldin Index; the optimal K=9 was selected at the DBI minimum. PCA visualization indicated a distinguishable cluster structure. Clusters C0, C3, C5, and C7 exhibited faster transitions (median waiting time 2 months) with high employment proportions (up to ~90%), whereas C2 and C8 showed longer waiting times (≥4 months). Cluster C4 was characterized by the longest study duration and a comparatively lower employment proportion. These results demonstrate that unsupervised learning can reveal actionable readiness segments, supporting targeted interventions (e.g., CV/portfolio clinics, interview practice, structured internships) and providing a foundation for subsequent predictive modeling of graduate outcomes.

INTRODUCTION

A survey conducted by Statistics Indonesia (Badan Pusat Statistik/BPS) found that the open unemployment rate among higher education graduates was approximately 5.25% in August 2024 [1]. This figure underscores the existence of a gap in the transition from university to the labour market. In this context, work readiness can be understood as a component of employability, referring to the level of skills and attitudes possessed by individuals [2]. Another definition of work readiness is the extent to which individuals possess the attitudes and characteristics required for success in the workplace [3]. Indicators of work readiness include a sense of responsibility, critical thinking, a range of relevant skills, the ability to communicate effectively, and the capacity for self-evaluation [4]. Other literature conceptualises work readiness as comprising four main aspects: skills, knowledge, understanding, and personal attributes [5]. Previous studies further conclude that several indicators influencing the work readiness of fresh graduates include TOEFL scores, soft skills (SSKM), and grade point average (GPA) [6], [7], [8]. GPA reflects graduates' academic competence, TOEFL scores represent their foreign-language communication ability, whereas soft skills (SSKM) capture non-technical attributes that support the performance of fresh graduates and influence their work readiness [9] Another study reports that the length of study does not affect the waiting time until graduates obtain their first job [10]. Prior research in higher education contexts also concludes that the key employability factors for university graduates are soft skills, problem-solving skills, functional skills, and academic reputation [11].

Tracer studies of university graduates are conducted as follow-up investigations that provide information on the linkage between higher education and the professional labour market, including the transition process from graduation to employment[12], [13]. Tracer study instruments generally capture graduates' post-graduation profiles, including employment status, waiting time to first job, income, competencies acquired, and other aspects used to assess graduates' work readiness [13]. Data from tracer studies typically show heterogeneous patterns of work readiness rather than a single uniform type [14]. A previous tracer study employing the K-Means algorithm concluded that computer skills, communication ability, and integrity were among the dimensions rated as very strong [15]. Therefore, a data-driven segmentation of graduate profiles is needed so that universities can enhance work readiness in a more targeted manner, focusing on groups of graduates with relatively homogeneous characteristics through clustering.

Machine Learning (ML) is a field of study that focuses on developing programs capable of inferring new information from existing knowledge [16]. ML has emerged as a new frontier in education and is recognized as one of the most powerful emerging technologies [17]. In other literature, ML is defined as the process of extracting knowledge from data [18]. Clustering, in this context, refers to the grouping of objects based on information contained in the data such that members within the same group share similarities [19]. The purpose of clustering is to classify data by identifying groupings within a dataset whose structure is initially unknown [20]. Another perspective describes clustering as the organization of data records or observations into classes of objects that exhibit similarity [21]. In addition, clustering can be understood as a method for assigning objects into one or more groups with the aim of analysing and automatically extracting knowledge [22].

The K-Means algorithm is a non-hierarchical clustering method that initiates the grouping process by selecting initial centroids from a subset of the population.[21] Another definition describes K-Means as a clustering algorithm that partitions data based on their proximity to centroid values [23]. Previous clustering studies have shown that K-Means can effectively map patterns in un labelled (unsupervised) data and generate operational insights for decision-makers, for example in analyzing library borrowing trends and managing services based on user needs [24]. The number of clusters can also be determined objectively using the Elbow method and the Davies-Bouldin Index (DBI), as demonstrated in local studies to ensure well-separated clusters [25]. Other K-Means research highlights that combining K-Means with subsequent classification can clearly delineate subpopulations and support decision-making [26]. In the context of fresh graduates, the application of K-Means to academic indicators has shown that clusters with higher average GPAs tend to be associated with shorter

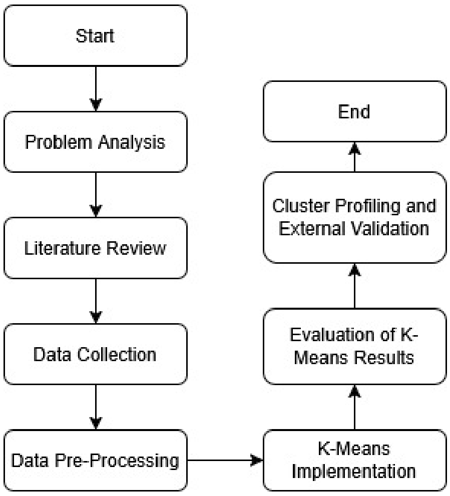

Figure 1. Research Procedure for the K-Means Clustering

waiting times for obtaining the first job, thereby providing actionable information for curriculum development and career guidance programs [27].

Aligned with these previous studies, the present research applies the K-Means algorithm to cluster fresh graduates based on work readiness indicators derived from tracer study data, namely GPA, TOEFL score, and soft skills (SSKM) as input variables. Employment status and waiting time are treated as external variables used to interpret and validate the resulting clusters.

METHOD

This study employs a quantitative approach with an unsupervised clustering method using the K-Means algorithm, with the following objectives: (1) to segment recent graduates based on work readiness indicators, (2) to determine the optimal value of K, and (3) to evaluate the quality of the segmentation produced by K-Means. The data are sourced from the tracer study of XYZ University, and the research procedure follows the workflow illustrated in Figure 1.

1. Data Collection

The data used in this study are drawn from the tracer study of XYZ University, comprising 669 respondents from graduating cohorts in the period 2020–2023. Each row in the dataset represents one graduate, with the following variables: GPA, length of study, TOEFL score, SSKM points, employment status, and waiting time to first job.

2. Pre-Processing Data

Data pre-processing is an initial stage carried out before the data are subjected to further analysis [28]. In this study, pre-processing is performed to ensure that the data are suitable for analysis using K-Means, through the following steps

1. Initial data cleaning, which includes detecting and removing duplicate records and checking whether the values fall within reasonable ranges, such as GPA (0.00–4.00), length of study (6–16 semesters), TOEFL (310–677), and SSKM score (≥ 112).
2. Data standardization, by applying the z-score transformation to selected variables so that they are on a comparable scale, according to:

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j} \quad (1)$$

3. K-Means Implementation

Machine learning with the K-Means algorithm is a non-hierarchical clustering method that initiates the grouping process by selecting initial centroids from a subset of the population [21]. The stages of K-Means implementation in this study are as follows:

1. Determining the number of clusters (K).
   To determine the number of clusters, this study employs the Elbow method and the Davies-Bouldin Index (DBI). The Elbow method identifies the optimal number of clusters (K) by locating the "elbow point" in the evaluation curve, i.e., the point at which adding more clusters no longer produces a substantial reduction in error [29]. DBI is a metric for evaluating clustering quality by measuring how well separated and how compact the formed clusters are [30]. The objective of DBI evaluation is to obtain clusters that minimize within-cluster dispersion while maintaining adequate separation between clusters [31].
2. Initialization of centroids, by randomly selecting K observations from the standardized data as initial centroids.
3. Assigning observations to the nearest cluster, using the Euclidean distance:

$$D(x_i, \mu_j) = \sqrt{\sum_{p=1}^{m} \left(x_{ip} - \mu_{jp}\right)^2} \quad (2)$$

   m is the number of variables/features.

4. Updating centroids for each cluster i, the centroid is updated as the mean of all observations assigned to that cluster.

$$\overline{v_{ij}} = \frac{1}{N_i} \sum_{k=1}^{N_i} x_{kj} \quad (3)$$

   where:
   $\overline{v_{ij}}$: coordinate of the centroid of cluster i on the variable j;
   N_i : number of members in cluster i;
   x_{kj} : value of observation k on variable j in cluster i.

5. Iteration until convergence. The assignment, and update steps are repeated, reallocating all observations to the nearest updated centroid, until there is no change in cluster membership or the changes are negligible. At this point, the algorithm is deemed to have converged and the clustering process is completed.

4. Evaluation of K-Means Results

The results of K-Means are evaluated by analyzing the SSE curve to identify the elbow point as the optimal K, which is then confirmed by the minimum value produced by the DBI [32].

5. Cluster Profiling and External Validation

External validation is conducted after the final clustering solution has been selected by comparing the clustering results with information/variables that were not used during the clustering process [33]. In this study, once K has been selected based on internal evaluation using SSE and DBI, the clustering solution is examined externally by linking the cluster labels to outcomes excluded from the modeling stage, namely employment status and median waiting time. Cluster profiles are then presented by inspecting the centroid values of each indicator in every cluster, so that the dominant characteristics and differences among clusters can be clearly identified.

RESULT AND DISCUSSION

K-Means clustering on the work readiness variables yielded an optimal solution of nine clusters. The best value of K was obtained at K = 9, as indicated in Figure 2.

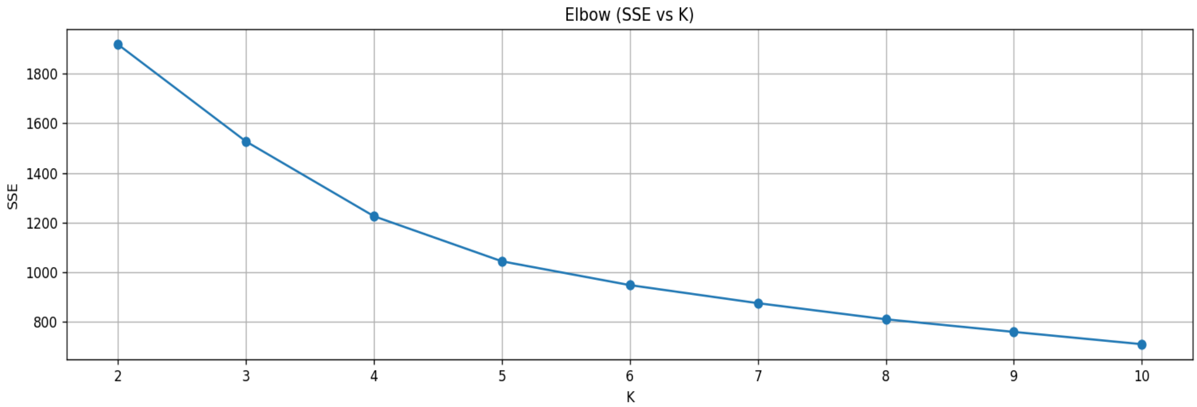

Figure 2. Number of clusters determined using the Elbow Method and SSE

The profile of each cluster is described using the centroids of GPA, TOEFL, SSKM, and length of study, and is further evaluated using the external variables of employment status and waiting time. In general, clusters characterized by higher GPA/TOEFL/SSKM and shorter study duration tend to exhibit a higher proportion of employed graduates and a shorter median waiting time. Table 1, which presents the centroids of the input variables, shows that several clusters are characterized by higher GPA/TOEFL/SSKM values combined with shorter study duration, while other clusters exhibit more moderate averages with longer study periods. This pattern confirms the presence of segmented variation among graduates across clusters.

Table 1. Centroids of input variables *(K=9)*

	GPA	**TOEFL SCORE**	**SSKM**	**Length of Study**
C0	3,744	630,833	280,3	8,1
C1	3,294	486,806	167,66	9,326
C2	3,135	628,333	163,889	9,289
C3	3,683	488,943	169,903	7,943
C4	3,185	526,93	168,684	14,421
C5	3,634	616,25	171,28	7,96
C6	3,703	474,054	258,486	7,824
C7	3,837	504,5	481,2	7,5
C8	3,351	477,778	294,111	11,5

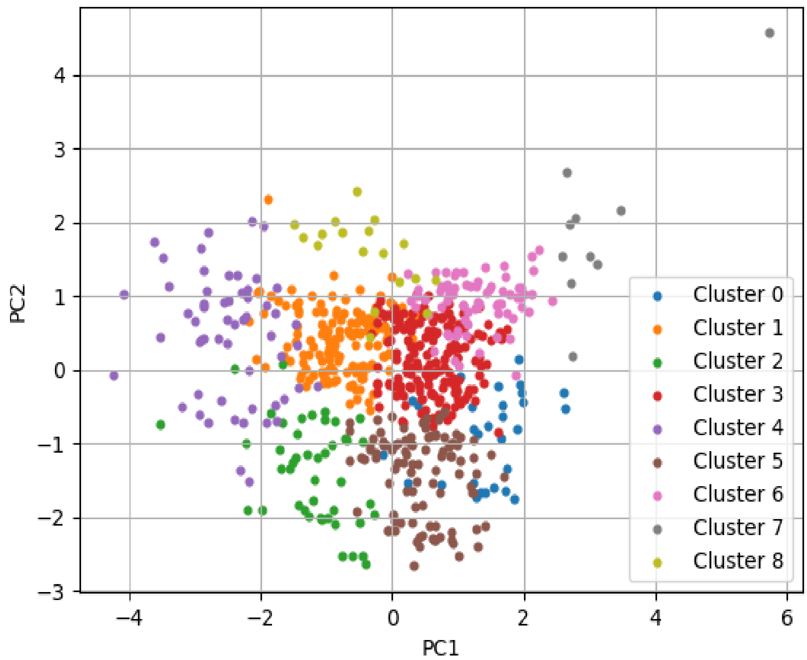

Figure 3. 2D PCA Visualization

The cluster distribution in Figure 3 is displayed as a two-dimensional projection of the Principal Component Analysis (PCA) results on the PC1 and PC2 axes. The visualization shows that several clusters are concentrated around the origin, with a degree of overlap still observable. This pattern indicates that K-Means is able to produce a meaningful segmentation of recent graduates, even though not all cluster boundaries are perfectly separated in the two-dimensional projection.

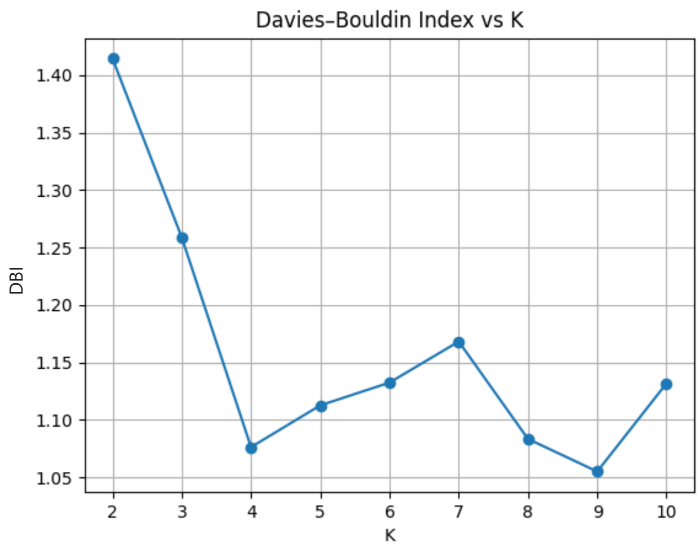

Figure 4. DBI Visualization

The DBI plot in Figure 4 shows a sharp decrease from K = 2 to K = 4, followed by fluctuations with a local minimum at K = 9 before increasing again at K = 10. Because lower DBI values indicate better clustering quality, these results suggest that the best candidate values for K lie in the range of 8–9, with K = 9 yielding the lowest DBI. The final choice of K is determined jointly with the SSE results.

Table 2. Cluster Average Data

Cluster	GPA	TOEFL	SSKM	STUDY LENGTH	Total	Median MT (month)	Employment rate
C0	3,744	630,833	280,300	8,1	30	2,0	0,833
C1	3,294	486,806	167,660	9,326	144	3,0	0,694
C2	3,135	628,333	163,889	9,289	45	4,0	0,689
C3	3,683	488,943	169,903	7,943	175	2,0	0,771
C4	3,185	526,930	168,684	14,421	57	3,0	0,632
C5	3,634	616,250	171,280	7,960	100	2,0	0,790
C6	3,703	474,054	258,486	7,824	74	3,0	0,716
C7	3,837	504,500	481,200	7,5	10	2,0	0,900
C8	3,351	477,778	294,111	11,5	18	4,5	0,778

Table 2. Cluster Averages Data summaries the results of the K-Means clustering. Cluster 7 represents recent graduates with high GPAs, shorter study duration, and high SSKM scores, with an employment proportion of around 90% and a median waiting time of two months. Cluster 0 is characterized by a high TOEFL score but a lower employment rate than Cluster 7. This indicates that English proficiency alone is not sufficient, and still needs to be complemented by other work-readiness factors such as GPA, SSKM, and timely completion of study. Cluster 4 has the lowest employment proportion (63%), suggesting an association between longer study duration and both lower employment rates and longer waiting times. Clusters C0, C3, C5, and C7 exhibit a relatively fast transition, with a median waiting time of two months and high employment proportions, whereas C8 and C2 have the longest waiting times, exceeding four months. Cluster C4 stands out due to its longer study duration. Taken together, these patterns suggest that higher GPA, TOEFL, SSKM, and on-time graduation are associated with more favorable employment outcomes. A chi-square test showed a statistically significant association between cluster membership and employment status, $\chi^2(8) = 10.82$, $p = 0.048$, indicating that the proportion of employed graduates differs across clusters. The clustering results, therefore, indicate the presence of groups of graduates with relatively higher academic achievement and soft skills, which are associated with higher employment proportions and shorter median waiting times. These findings are in line with previous studies that emphasize the influence of academic indicators and soft skills on graduate employability.

CONCLUSION AND RECOMMENDATIONS

Based on the findings of this study, the application of K-Means is effective for clustering graduates' work readiness using indicators such as GPA, TOEFL score, SSKM, and length of study. In this research, the number of clusters was determined in a data-driven manner, with the minimum DBI obtained at K = 9, resulting in nine segments with distinct profiles. Several clusters exhibit the fastest transition to employment with a median waiting time of two months, whereas other clusters have longer waiting times, and one cluster is particularly salient with the longest study duration and a relatively low employment proportion. These results indicate that an unlabeled (unsupervised) machine learning approach can provide actionable information to support more targeted career preparation decisions. The cluster profiles generated in this study can serve as a basis for universities to design policies and programs related to graduate career preparation. For example, clusters with high GPA/TOEFL/SSKM values can be prioritized for direct placement or linkage with industry partners to shorten their waiting time, while clusters with lower scores may be targeted for additional support, such as CV clinics, interview training, or preparatory courses before entering the labor market. This study remains limited in terms of the number of input variables. Future research is recommended to incorporate additional variables such as internship experience, professional certifications, and portfolios. Subsequent studies may also include alternative comparison methods and conduct external statistical validation to obtain more accurate and practically useful segmentations. In addition, tracer study data could be further explored using predictive machine learning models to analyze relationships between internal variables (e.g., TOEFL, GPA, length of study, SSKM) and external outcomes such as graduates' waiting time to first employment.

REFERENCES

[1] Direktorat Statistik Kependudukan dan Ketenagakerjaan, "Keadaan Angkatan Kerja di Indonesia Agustus 2024," 2024.

[2] K. Peersia, N. A. Rappa, dan L. B. Perry, "Work readiness: definitions and conceptualisations," *High. Educ. Res. Dev.*, vol. 43, no. 8, hal. 1830–1845, Nov 2024, doi: 10.1080/07294360.2024.2366322.

[3] S. Sultoni, I. Gunawan, A. A. Firawati, dan A. Cahyaningtyas, "Validity and Reliability of the Student Work Readiness Scale," *JAMP J. Adm. dan Manaj. Pendidik.*, vol. 5, no. 2, hal. 176–185, 2022, doi: 10.17977/um027v5i22022p176.

[4] A. Prianto, W. Winardi, dan U. N. Qomariyah, "The Effect of the Implementation of Teaching Factory and Its Learning Involvement toward Work Readiness of Vocational School Graduates," *Int. J. Instr.*, vol. 14, no. 1, hal. 283–302, Jan 2021, doi: 10.29333/iji.2021.14117a.

[5] B. Sudarsono, F. Tentama, dan F. A. Ghozali, "Employability Analysis of Students in Yogyakarta: Confirmatory Factor Analysis," *AL-ISHLAH J. Pendidik.*, vol. 14, no. 2, hal. 1451–1462, 2022, doi: 10.35445/alishlah.v14i2.1782.

[6] Arina Hafadhotul Husna, "PERSEPSI MAHASISWA (NON – ENGLISH) TENTANG KEBIJAKAN TOEFL SEBAGAI SYARAT KELULUSAN," *JISPENDIORA J. Ilmu Sos. Pendidik. Dan Hum.*, vol. 1, no. 1, hal. 63–72, Jul 2022, doi: 10.56910/jispendiora.v1i1.67.

[7] K. A. C. Labiro dan W. Widjaja, "Pengaruh Keaktifan Berorganisasi, Prestasi Akademik, Lingkungan Belajar terhadap Kesiapan Kerja Universitas Tangerang," *Stud. Ilmu Manaj. dan Organ.*, vol. 5, no. 1, hal. 27–42, Apr 2024, doi: 10.35912/simo.v5i1.2958.

[8] I. Siswanto, I. W. N. Arifah, dan F. E. N. Ramadhan, "Pengaruh Keaktifan di Organisasi dan IPK terhadap Softskills dan Kesiapan Menjadi Guru Mahasiswa Pendidikan Teknik Otomotif," *TAMAN VOKASI*, vol. 7, no. 2, hal. 106, Des 2019, doi: 10.30738/jtv.v7i2.6314.

[9] D. N. Azizah, S. Muslim, dan M. Cholik, "The correlation of industrial work experience and soft skills on work readiness of graduated of vocational high school," *Int. J. Educ. Vocat. Stud.*, vol. 3, no. 4, hal. 248, Agu 2021, doi: 10.29103/ijevs.v3i4.3018.

[10] T. P. Kiri dan A. Atti, "Pengaruh Ipk Dan Masa Studi Terhadap Waktu Tunggu Mendapatkan Pekerjaan," *J. Difer.*, vol. 3, no. 1, hal. 1–12, 2021, doi: 10.35508/jd.v3i1.3998.

[11] S. Sahudin, N. che maideen, R. Wahab, dan N. Shuib, "Literature Review on the Factors Affecting Employability of Engineering Graduates," hal. 13–22, Jun 2022.

[12] H. Schomburg, *Handbook for Graduate Tracer Studies*, 2 ed. Moenchebergstrasse, Germany: International Centre for Higher Education Research, 2003.

[13] S. Syafiq, Ahmad;Fikawati, *Metodologi dan Manajemen Tracer Study*, 1 ed. Depok:

Rajawali Pers, 2016.
[14] PLKA UNDIKA, "Laporan Tracer Study 2020-2023," Surabaya, 2024.
[15] S. Febrianti, M. Munawir, dan L. Fitria, "Penerapan Metode K-Means Clustering Terhadap Alumni Berdasarkan Kuesioner Tracer Study Application of Method K-Means Clustering To Alumni Based on Tracer Study Questionnaire," *J. Informatics Comput. Sci.*, vol. 7, no. 2, hal. 117–122, 2021.
[16] N. Dinata, Rozzi Kesuma; Hasdyna, *Machine Learning*, 1 ed. Lhoksumawe: Unimal Press, 2020.
[17] R. Diana, H. Warni, dan T. Sutabri, "PENGGUNAAN TEKNOLOGI MACHINE LEARNING UNTUK PELAYANAN MONITORING KEGIATAN BELAJAR MENGAJAR PADA SMK BINA SRIWIJAYA PALEMBANG," *JUTEKIN (Jurnal Tek. Inform.*, vol. 11, no. 1, Jun 2023, doi: 10.51530/jutekin.v11i1.709.
[18] B. Raharjo, *Pembelajaran Mesin*, 1 ed. Semarang: Yayasan Prima Agus Teknik, 2021.
[19] R. Astuti dan K. Ukar, "Implementasi Data Mining dengan Metode Clustering Algoritma K-Means untuk Pengelompokan Data Tilang di Instansi Pemerintah," *Media Inform.*, vol. 20, no. 2, hal. 109–121, Jul 2021, doi: 10.37595/mediainfo.v20i2.77.
[20] S. Sindrawati, D. Syaripudin, dan A. R. Raharja, "Penerapan Algoritma K-Means Clustering pada Data Nilai Siswa untuk Menentukan Kelompok Penerima Beasiswa," *SisInfo*, vol. 6, no. 2, hal. 47–55, 2024, doi: 10.37278/sisinfo.v6i2.900.
[21] R. Wahyudi, Mochamad;Masitha;Saragih, *Data Mining: Penerapan Algoritma K-Means Clustering dan K-Medoids Clustering*, 1 ed. Jakarta: Yayasan Kita Menulis, 2020.
[22] T. Ikhsan, "Clustering Data Penduduk Menggunakan Algoritma K-Means," *TIN Terap. Inform. Nusant.*, vol. 5, no. 12, hal. 955–963, 2025, doi: 10.47065/tin.v5i12.7328.
[23] A. Fatkhudin, A. Khambali, F. A. Artanto, N. A. Putra Zade, dan U. Muhammadiyah Pekajangan Pekalongan, "Implementasi Algoritma Clustering K-Means Dalam Pengelompokan Mahasiswa Studi Kasus (Prodi Manajemen Informatika)," *J. Minfo Polgan*, vol. 12, no. 2, hal. 777–783, 2023, [Daring]. Tersedia pada: https://doi.org/10.33395/jmp.v12i2.12494
[24] L. Rosiana dan I. Yuadi, "K-Means Clustering untuk Analisis Tren Peminjaman Buku di Perpustakaan," *J. Technol. Informatics*, vol. 7, no. 1, hal. 1–10, Apr 2025, doi: 10.37802/joti.v7i1.933.
[25] A. Salsabiela, A. P. Kuncoro, P. Subarkah, dan P. Arsi, "Rekomendasi Restock Barang di Toko Pojok UMKM Menggunakan Algoritma K-Means Clustering," *J. Technol. Informatics*, vol. 5, no. 2, hal. 87–92, Apr 2024, doi: 10.37802/joti.v5i2.554.
[26] A. Aryanti, F. D. Wulandari, dan M. Rafiif, "Segmentasi dan Klasifikasi Risiko Perdarahan dengan Algoritma K-Means dan Naive Bayes Berdasarkan Data Klinis dan Transfusi," vol. 7, no. 2, hal. 211–219, 2025.
[27] F. A. Fernaldy, A. A. Arifiyanti, dan D. S. Y. Kartika, "KLASTERISASI TRACER STUDY ALUMNI UNIVERSITAS XYZ MENGGUNAKAN ALGORITMA K-MEANS," *J. Inform. dan Tek. Elektro Terap.*, vol. 13, no. 1, Jan 2025, doi: 10.23960/jitet.v13i1.5581.
[28] H. G. Costa, M. H. T. da Silva, G. N. Santos, A. Bonamigo, dan R. D. Callado, "Clustering Brazilian Public Emergency Healthcare Units," *IFAC-PapersOnLine*, vol. 55, no. 10, hal. 566–571, 2022, doi: 10.1016/j.ifacol.2022.09.454.
[29] E. Schubert, "Stop using the elbow criterion for k-means and how to choose the number of clusters instead," *ACM SIGKDD Explor. Newsl.*, vol. 25, no. 1, hal. 36–42, Jun 2023, doi: 10.1145/3606274.3606278.
[30] S. Ramadhani, D. Azzahra, dan T. Z, "Comparison of K-Means and K-Medoids Algorithms

in Text Mining based on Davies Bouldin Index Testing for Classification of Student's Thesis," *Digit. Zo. J. Teknol. Inf. dan Komun.*, vol. 13, no. 1, hal. 24–33, Mei 2022, doi: 10.31849/digitalzone.v13i1.9292.

[31] F. Ros, R. Riad, dan S. Guillaume, "PDBI: A partitioning Davies-Bouldin index for clustering evaluation," *Neurocomputing*, vol. 528, hal. 178–199, Apr 2023, doi: 10.1016/j.neucom.2023.01.043.

[32] A. Kassambara, *Practical Guide to Cluster Analysis in R: Unsupervised Machine Learning*, vol. 1. STHDA, 2017.

[33] C. Hennig, M. Meila, F. Murtagh, dan R. Rocci, Ed., *Handbook of Cluster Analysis*. Chapman and Hall/CRC, 2015. doi: 10.1201/b19706.