Techno Nusa Mandiri: Journal of Computing and Information Technology
Vol. 22, No. 1 March 2025 | DOI: https://doi.org/10.33480/techno.v21i1.XXXX

ANALYSIS OF K-MEANS CLUSTERING OF GLOBAL ARTIFICIAL INTELLIGENCE STARTUP FUNDING ACTIVITIES WITH ELBOW METHOD

Benedikta Icha Nurcahaya Gea [1*]; Rachman Komarudin [2]

Information System[1]
Universitas Nusa Mandiri, Jakarta, Indonesia[1,2]
www.nusamandiri.ac.id[1,2]
benediktaicha@gmail.com[1*] rachman.rck@nusamandiri.ac.id[2]
(*) Corresponding Author



The creation is distributed under the Creative Commons Attribution-NonCommercial 4.0 International License.

Abstract— *The rapid growth of Artificial Intelligence (AI) startups has significantly increased global investment activities. The large amount of funding data makes it difficult to identify the characteristics of startup funding activities through manual analysis. Therefore, an appropriate clustering method is required to group funding activities based on the similarity of their funding characteristics. This study aims to implement the K-Means Clustering algorithm to classify funding activities of Artificial Intelligence startups based on global funding characteristics and to determine the optimal number of clusters using the Elbow Method. This research employed a quantitative approach using the AI Company & Startup Funding Database obtained from Kaggle. The research stages consisted of data preprocessing, data normalization using Min-Max Scaling, determining the optimal number of clusters through the Elbow Method, and performing the clustering process using the K-Means algorithm. The results indicate that the optimal number of clusters is five (K = 5). The clustering process produced five groups with different characteristics based on centroid values, representing variations in funding amount, pre-money valuation, post-money valuation, number of employees, annual revenue, and funding stage. The findings demonstrate that the combination of K-Means Clustering and the Elbow Method effectively groups funding activities of Artificial Intelligence startups according to their global funding characteristics and provides valuable information to support data analysis and decision-making.*

Keywords*: artificial intelligence startup, data mining, elbow method, global funding, k-means clustering*

Intisari—Perkembangan startup Artificial Intelligence (AI) yang semakin pesat telah meningkatkan aktivitas pendanaan dari berbagai investor di tingkat global. Banyaknya data pendanaan yang tersedia menyebabkan proses identifikasi karakteristik aktivitas pendanaan startup menjadi lebih kompleks apabila dilakukan secara manual, sehingga diperlukan metode yang mampu mengelompokkan aktivitas pendanaan berdasarkan tingkat kemiripan karakteristik pendanaan. Penelitian ini bertujuan menerapkan metode K-Means Clustering untuk mengelompokkan aktivitas pendanaan startup Artificial Intelligence berdasarkan karakteristik pendanaan global serta menentukan jumlah cluster optimal menggunakan evaluasi Elbow Method. Penelitian menggunakan pendekatan kuantitatif dengan memanfaatkan dataset AI Startup Funding Database yang diperoleh dari Kaggle, meliputi tahapan data preprocessing, normalisasi menggunakan Min-Max Scaling, penentuan jumlah cluster menggunakan Elbow Method, serta proses clustering menggunakan algoritma K-Means. Hasil penelitian menunjukkan bahwa jumlah cluster optimal adalah lima kelompok (K = 5), yang menghasilkan lima kelompok dengan karakteristik berbeda berdasarkan nilai centroid, mencerminkan perbedaan jumlah pendanaan, valuasi perusahaan sebelum dan sesudah pendanaan, jumlah karyawan, pendapatan tahunan, serta tahap pendanaan. Hasil penelitian menunjukkan bahwa kombinasi metode K-Means Clustering dan Elbow Method mampu mengelompokkan aktivitas pendanaan startup Artificial Intelligence berdasarkan karakteristik pendanaan global sehingga dapat mendukung proses analisis data dan pengambilan keputusan.

Kata Kunci: artificial intelligence startup, data mining, elbow method, global funding, k-means clustering

Techno Nusa Mandiri: Journal of Computing and Information Technology
Vol. 22, No. 1 March 2025 | DOI https://doi.org/10.33480/techno.v21i1.XXXX

INTRODUCTION

Rapid advances in digital technology have driven the growth of *Artificial Intelligence* (AI)-based *startups* across various industries. As the number of AI *startups* increases, *funding activities* for these *startups* have also grown globally, with diverse funding characteristics such as investment amounts, funding stages, and the geographical distribution of investors (Laksono, Syahidin, & Yunengsih, 2024). This condition has generated large, diverse, and complex AI *startup* funding data, requiring an analytical approach capable of identifying funding patterns and characteristics in greater depth (Schwab-McCoy, Baker, & Gasper, 2021).

Available *startup* funding data can generally be developed beyond descriptive analysis to discover the patterns contained within it. Large and complex datasets offer the potential to identify segmentation or grouping of *startups* based on their funding characteristics; therefore, a *data mining* approach is needed to process and group the data systematically (Dewi, Nurwati, & Rahayu, 2022). *Data mining* is a process of processing and analyzing data to discover useful patterns, relationships, trends, and hidden information, with *clustering* as one of its main functions (Han, Kamber, & Pei, 2016).

One widely used *clustering* method is *K-Means Clustering*, an *unsupervised learning* technique that groups data based on similarity of characteristics using proximity to *centroid*s. This method has been applied in various fields, including grouping student learning outcomes (Hendrastuty, 2024), earthquake data analysis (Prasetio, Makmun, & Dwi, 2023), inpatient data *clustering* (Laksono et al., 2024), and investor-*startup* network analysis (Carniel, Halloy, & Dalle, 2023) . One of the main challenges of K-Means is determining the optimal number of *cluster*s, so an evaluation method such as the *Elbow Method* is needed to determine the number of *cluster*s based on changes in *Within Cluster Sum of Squares* (*WCSS*) (Maori, 2023; Wijaya, 2020).

The development of Artificial Intelligence has also been accompanied by a significant increase in investment in the sector globally (Montanaro, Croce, & Ughetto, 2024). *Artificial Intelligence startups utilize AI technology as an important part of developing products, services, and business models* (Russels Stuart & Norvig Peter, 2022). and generally require staged funding from Pre-Seed and Seed to subsequent Series rounds to support business growth (Arnold, Claveres, & Frie, n.d.). Previous research shows that startup characteristics can be considered in investment decisions (Kartini & Candya CALLISTA, 2021). Machine learning approaches have also been used to support startup investment decisions (Setty, Elovici, & Schwartz, 2024), while the influence of investors has been examined in the funding success of generative AI startups (Siddik, Li, & Du, 2024).

Based on the identification results, the *AI Startup Funding Database 2014–2025* dataset obtained from Kaggle contains 38 *funding activities* (*funding deals*) with 27 attributes, where each row represents one funding activity, so a company may appear more than once if it has *funding activities* in different investment rounds. Previous studies have shown that combining *K-Means Clustering* and the *Elbow Method* can produce effective grouping across various research objects. However, the application of these two methods to *Artificial Intelligence startup funding activities* based on global funding characteristics has not been widely conducted. Therefore, this study does not position the use of *K-Means Clustering* and the *Elbow Method* as a methodological novelty, but focuses on a different object and data characteristics, namely *Artificial Intelligence startup funding activities* based on a global funding dataset.

Based on the description above, this study aims to: (1) apply the *K-Means Clustering* method to group *Artificial Intelligence startup funding activities* based on similarities in funding characteristics; (2) determine the optimal number of *cluster*s using the *Elbow Method* based on *WCSS* values; (3) analyze the patterns and characteristics of *funding activities* in each resulting *cluster*; and (4) determine the *cluster* for new AI *startup* funding data based on its proximity to the *centroid*s produced by *K-Means Clustering*, which is subsequently represented through a simple *prototype*.

MATERIALS AND METHODS

This study uses a quantitative research method with a *data mining* approach. This approach is used to analyze *Artificial Intelligence startup funding activities* based on the funding characteristics contained in the research dataset. A *clustering* technique is used because the study does not use predetermined group labels. The algorithm used is *K-Means Clustering*, while the number of *cluster*s is determined using the *Elbow Method* based on *Within Cluster Sum of Squares* (*WCSS*) values.

The data used are secondary data obtained from the Kaggle platform, namely the *AI Startup Funding Database 2014–2025* dataset compiled by Prajit Datta (Schwab-McCoy et al., 2021). The dataset contains 38 *funding activities* (*funding deals*) and 27 attributes concerning the *funding activities* of companies and *startups* operating in the

field of *Artificial Intelligence*. Because a company may appear more than once when it receives funding in different rounds, the unit of analysis in this study is the funding activity (funding deal), not the number of unique companies.

Of the 27 available attributes, this study uses six numerical attributes as variables in the *clustering* process because these attributes represent the funding characteristics and company conditions that are the focus of the study. Identity attributes, geographical information, and categorical attributes are not used as *clustering* variables because they are not directly used in distance calculations. The research variables are presented in Table 1.

Table 1. Research Variables Used

No	Atribut	Keterangan
1	amount_usd_millions	Nilai pendanaan startup (USD juta)
2	pre_money_valuation_millions	Valuasi sebelum pendanaan (USD juta)
3	post_money_valuation_millions	Valuasi setelah pendanaan (USD juta)
4	employees_approx	Estimasi jumlah karyawan
5	annual_revenue_millions	Estimasi pendapatan tahunan (USD juta)
6	stage_number	Urutan numerik tahap pendanaan

Source: AI Startup Funding Database 2014–2025 (processed, 2026).

The research stages were conducted systematically, beginning with dataset collection, variable selection, *data preprocessing*, data normalization, determination of the number of *cluster*s using the *Elbow Method*, application of *K-Means Clustering*, and analysis and interpretation of the resulting *cluster* characteristics. A summary of the research flow is presented in Figure 1.



Source: Author's design (2026).

Figure 1. Research Flow Diagram

The *data preprocessing* stage was conducted to examine and handle *missing value*s, duplicate data, and data type suitability for the six selected attributes, resulting in a dataset ready for analysis. Next, normalization was performed using *Min-Max Scaling* so that all attributes were within the range of 0 to 1 and no attribute dominated the distance calculation, as expressed in the following equation:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$$

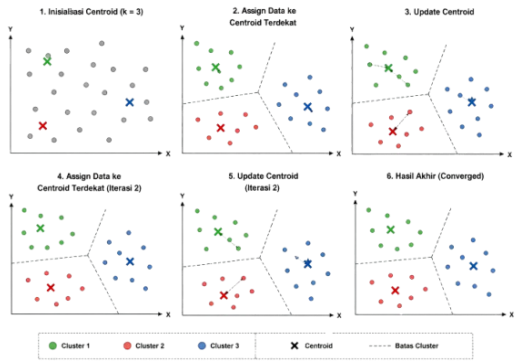
where *x* is the initial data value, *x*min and *x*max are the minimum and maximum values of the attribute, respectively, and"x′" is the normalized value.

The optimal number of clusters was determined using the Elbow Method by calculating WCSS values over a range of to , then selecting the point at which the decrease in WCSS began to flatten (the elbow). After the optimal number of cluster was established, the K-Means Clustering algorithm was applied by measuring the proximity of each data point to the centroid using Euclidean Distance, as shown in Equation (2):

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

Techno Nusa Mandiri: Journal of Computing and Information Technology
Vol. 22, No. 1 March 2025 | DOI https://doi.org/10.33480/techno.v21i1.XXXX

where *d(x,y)* is the distance between the data point and the *centroid*, *x*i is the data value for the *i*, *y*i is the *centroid* value for the *i*-th attribute, and *n* is the number of attributes. The process was performed iteratively until the position of *centroid* no longer changed significantly, as illustrated in Figure 2.



Source: Adapted by the author from (Nur et al., 2022) (2026)

Fusing 2. Illustration of the K-Means Clustering Process

All data processing steps, from *data preprocessing* to the application of *K-Means Clustering*, were performed using the Python programming language in the *Google Colab* environment with the *Scikit-learn library*. (Poolman, Townsend-Nicholson, & Cain, 2022) The *clustering* results were then analyzed based on *centroid* values and the number of members in each *cluster*, and subsequently used as the basis for determining the *cluster* of new funding data through *Euclidean Distance* calculations against the *centroids*, represented in a simple *prototype*.

RESULTS AND DISCUSSION

The research dataset is the *AI Startup Funding Database 2014–2025* obtained from Kaggle, consisting of 38 *funding activities* (*funding deals*) and 27 attributes. Of all available attributes, six numerical attributes were used as *clustering* variables, as described in Table 1. A summary of the research dataset is presented in Table 2.

Table 2. Characteristics of the Research Object

Komponen	Keterangan
Nama Dataset	AI Startup Funding Database 2014–2025
Sumber Data	Kaggle
Jenis Data	Data Sekunder
Jumlah Data	38 aktivitas pendanaan (funding deals)
Jumlah Atribut	27 atribut (6 digunakan sebagai variabel clustering)
Unit Analisis	Aktivitas pendanaan (funding deal)
Metode	K-Means Clustering
Evaluasi	Elbow Method

Source: *AI Startup Funding Database 2014–2025*, processed by the author (2026).

Table 2 shows that the research dataset meets the study requirements because it provides numerical attributes that can be used in the *K-Means Clustering* process. The *data preprocessing* examination of the six attributes showed no *missing value*s (0 out of 38 data points across all attributes), no duplicate data (0 out of 38 data points), and all attributes had appropriate numerical data types, namely Float64 and Int64. Thus, the final dataset consisting of 38 *funding activities* and six numerical attributes was ready for the normalization stage.

Normalization using *Min-Max Scaling* transformed the values of the six attributes into the range of 0 to 1 without changing the relative relationships among the data. Examination of the minimum and maximum values in the normalized data showed that all attributes had a minimum value of 0 and a maximum value of 1, so the normalized dataset was ready for determining the number of *cluster*s using the *Elbow Method*.

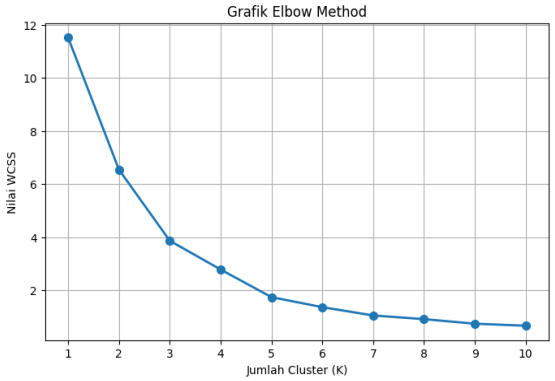
The optimal number of *cluster*s was determined by calculating WCSS values over a range of $K = 1$ hingga $K = 10$ using the K-Means algorithm on the normalized data. The WCSS calculation results are presented in Table 3.

Table 3. *Within Cluster Sum of Squares* (*WCSS*) Values for K = 1 to K = 10

Jumlah Cluster (K)	Nilai WCSS
1	11.5224
2	6.5491
3	3.8700
4	2.7859
5	1.7464
6	1.3692
7	1.0566
8	0.9182
9	0.7460
10	0.6703

Surce: Data processing results using *Google Colab* (2026).

Table 3 shows that WCSS values continuously decrease as the number of *cluster*. The decrease in WCSS is relatively large from $K = 1$ (11,5224) to $K = 5$ (1,7464), whereas after $K = 5$ the decrease in WCSS becomes relatively smaller, namely 1.3692 at $K = 6$ to 0,6703 at $K = 10$. This decreasing pattern is visualized in Figure 3.



Source: Data processing results using *Google Colab* (2026).

Figure 3. *Elbow Method* Graph

Figure 3 shows that the *WCSS* curve decreases sharply up $K = 5$, and then begins to flatten after that point. The change in the slope of the curve indicates the *elbow point* at $K = 5$. The selection of the number of *cluster*s was not based on the smallest *WCSS* value, because *WCSS* continues to decrease as the number of *cluster*s increases, but rather on the point at which the decrease in *WCSS* begins to flatten. Therefore, this study sets $K = 5$ as the optimal number of *cluster*s used in the *K-Means Clustering* implementation.

The *K-Means Clustering* algorithm was applied using the *Scikit-learn library* with the parameters *random_state* = 42 and *n_init* = 10 to 38 *funding activities* based on six numerical attributes. The *clustering* process produced five *clusters*, namely *cluster* 0 to *cluster* 4, with the distribution of members presented in Table 4.

Table 4. Number of *Funding Activities*

Cluster	Jumlah Aktivitas Pendanaan
0	8
1	21
2	3
3	5
4	1

Source: Data processing results using *Google Colab* (2026)

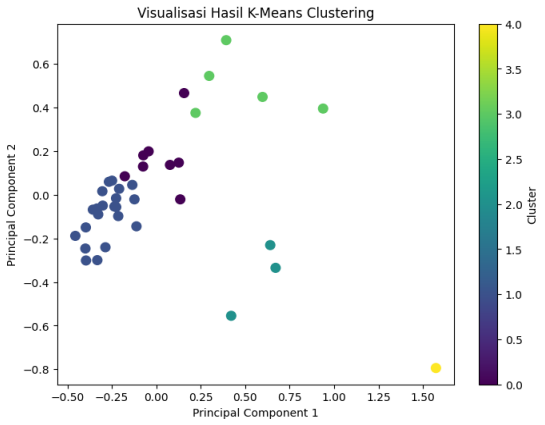
Table 4 shows that the distribution of 38 *funding activities* across the five *cluster*s is uneven. *Cluster* 1 has the largest number of members (21 *funding activities*), while *cluster* 4 contains only 1 funding activity. The characteristics of each *cluster* were then analyzed based on the *centroid* values of the six research attributes, as presented in Table 5.

Table 5. Centroid Values for Each Cluster

Cluster	Pendanaan	Val. Sebelum	Val. Setelah	Karyawan	Pendapatan	Tahap
0	0,1135	0,0844	0,0860	0,1238	0,0954	0,7625
1	0,0359	0,0463	0,0446	0,0490	0,0518	0,3190
2	0,2467	0,6527	0,5990	0,2443	0,2937	0,5667
3	0,0980	0,1848	0,1824	0,7850	0,1454	0,9200
4	1,0000	1,0000	1,0000	0,5270	1,0000	0,7000

Source: Data processing results using *Google Colab* (2026)

Table 5 shows that the *centroid* values across the six attributes differ for each *cluster*, indicating variations in the characteristics of *funding activities* between groups. For easier observation, the *clustering* results were visualized using *Principal Component Analysis* (*PCA*), which reduced the six research attributes to two principal components, as shown in Figure 4. *PCA* was used only as a visualization aid and was not used to form the *cluster*s.



Source: Data processing results using Google Colab (2026).

Gambar 4. Visualization of *K-Means Clustering* Results

Figure 4 shows that *cluster* 1 forms the densest group because it has the largest number of members, *cluster* 0 is relatively close to *cluster* 1, while *cluster*s 2 and 3 have fewer members with relatively different positions. *Cluster* 4 consists of only one funding activity and therefore appears separated from most of the data.

Based on the *centroid* values in Table 5, the characteristics of each *cluster* can be described as follows. ***Cluster* 0** (8 *funding activities*) has

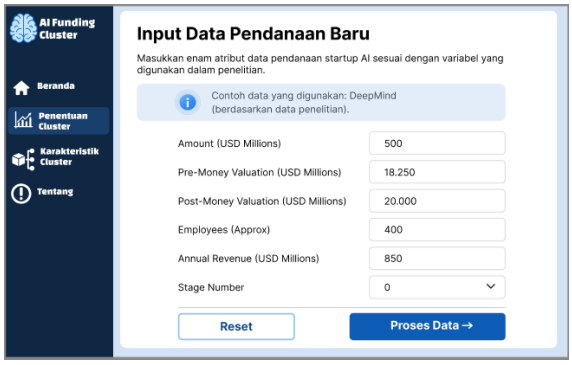
Techno Nusa Mandiri: Journal of Computing and Information Technology
Vol. 22, No. 1 March 2025 | DOI https://doi.org/10.33480/techno.v21i1.XXXX

relatively low funding, valuation, number of employees, and annual revenue values, but a relatively high *stage_number* value (0.7625), indicating *funding activities* at a relatively advanced stage despite their relatively low financial values. ***Cluster* 1** is the largest group (21 *funding activities*), with relatively low values across almost all attributes, including the funding stage (0.3190), reflecting the early-stage funding pattern most frequently found in the dataset.

***Cluster* 2** (3 *funding activities*) has relatively high pre-money valuation (0.6527) and post-money valuation (0.5990) compared with most other *cluster*s, representing a group with high valuation characteristics. ***Cluster* 3** (5 *funding activities*) is prominent in the *employees_approx* (0.7850) and *stage_number* (0.9200) attributes, indicating a group of larger companies with later funding stages, although its funding and valuation values are lower than those of *cluster* 2 and *cluster* 4. ***Cluster* 4** consists of only 1 funding activity, with the values of *amount_usd_millions*, *pre_money_valuation_millions*, *post_money_valuation_millions*, and *annual_revenue_millions* each reaching 1.0000, the maximum normalized value, representing a funding activity with the highest financial characteristics in the dataset.

These results are consistent with previous studies that applied *K-Means Clustering* to various objects, such as grouping student learning outcomes (Hendrastuty, 2024), earthquake data analysis (Prasetio et al., 2023), and the use of the *Elbow Method* based on *WCSS* values to help determine the optimal number of *cluster*s (Anggreani, Nurmisba, Setiawan, & Lukman, 2024; Maori, 2023; Safitri Juanita & Cahyono, 2024; Wijaya, 2020). The difference between this study and those studies lies in the object and data characteristics analyzed, namely *Artificial Intelligence startup funding activities* based on six numerical attributes representing funding characteristics and company conditions. Thus, this study extends the application of *K-Means Clustering* and the *Elbow Method* to the context of global AI *startup* funding analysis.

The resulting *cluster*s were subsequently applied to determine the *cluster* of new *startup* funding data through a simple *prototype*. The new data use the same six attributes as the research variables and are processed through preprocessing and *Min-Max Scaling* normalization using the same parameters as the research data, before their *Euclidean Distance* to the five established *centroid*s is calculated. The new data are assigned to the *centroid* with the smallest distance. In a demonstration using DeepMind data, the data were placed in *Cluster* 1 because they had the smallest distance compared with the distances to the other *cluster*s, as shown in the *prototype* interface in Figure 5.



Source: Prototype design by the author (2026).

Figure 5. Prototype Interface Design for New Funding Data Input

These results show that the *prototype* mechanism can represent the application of *K-Means Clustering* results in determining the *cluster* of new funding data based on proximity to the *centroid*. Overall, the *clustering* results provide an overview of variations in the characteristics of *Artificial Intelligence startup funding activities* that can serve as initial information for understanding the position of a funding activity based on its group characteristics, both for data analysis and for further research development using more diverse data and variables.

CONCLUSION

This study applies *K-Means Clustering* to group 38 *Artificial Intelligence startup funding activities* based on six numerical attributes, namely funding amount, pre- and post-funding valuation, number of employees, annual revenue, and funding stage. The optimal number of *cluster*s determined using the *Elbow Method* based on *WCSS* values over $K = 1$ to $K = 10$ is five ($K = 5$), because the *WCSS* decrease begins to flatten after that point. The *clustering* results produced five *cluster*s with different characteristics: *cluster* 0 (8 *funding activities*) has low to medium financial values with a relatively high funding stage; *cluster* 1 (21 *funding activities*) is the largest group with relatively low values across almost all attributes; *cluster* 2 (3 *funding activities*) has relatively high valuation; *cluster* 3 (5 *funding activities*) is prominent in number of employees and funding stage; and *cluster* 4 (1 funding activity) has funding, valuation, and annual revenue values at the maximum normalized level.

The *K-Means Clustering* results were subsequently applied to determine the *cluster* of new *startup* funding data based on proximity to the *centroid*s, represented through a simple *prototype*.

In a demonstration using DeepMind data, the data were successfully assigned to *Cluster* 1 based on the smallest *Euclidean Distance* to that *cluster centroid*. Thus, the combination of *K-Means Clustering* and the *Elbow Method* can be used to group *Artificial Intelligence startups* based on global funding characteristics and support data analysis and decision-making.

Future research is recommended to use a dataset with a larger number of *funding activities* and a longer observation period, consider adding other relevant numerical attributes, use additional evaluation methods such as *Silhouette Score* to complement *cluster* quality evaluation, and develop the *prototype* into an application fully integrated with the calculation process so that users can enter data directly and obtain *cluster* assignments automatically.

REFERENCE

Anggreani, D., Nurmisba, N., Setiawan, D., & Lukman, L. (2024). Optimization of K-Means Clustering Method by Using Elbow Method in Predicting Blood Requirement of Pelamonia Hospital Makassar. *Internet of Things and Artificial Intelligence Journal*, *4*(3), 541–550. https://doi.org/10.31763/iota.v4i3.755

Arnold, N., Claveres, G., & Frie, J. (n.d.). *IMF WORKING PAPERS Stepping Up Venture Capital to Finance Innovation in Europe Stepping Up Venture Capital to Finance Innovation in Europe*.

Carniel, T., Halloy, J., & Dalle, J. M. (2023). A novel clustering approach to bipartite investor-startup networks. *PLoS ONE*, *18*(1 January). https://doi.org/10.1371/journal.pone.0279780

Dewi, S. P., Nurwati, N., & Rahayu, E. (2022). Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor. *Building of Informatics, Technology and Science (BITS)*, *3*(4), 639–648. https://doi.org/10.47065/bits.v3i4.1408

Han, J., Kamber, M., & Pei, J. (2016). *Data Mining. Concepts and Techniques, 3rd Edition (The Morgan Kaufmann Series in Data Management Systems)*.

Hendrastuty, N. (2024). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa. *Jurnal Ilmiah Informatika Dan Ilmu Komputer (JIMA-ILKOM)*, *3*(1), 46–56. https://doi.org/10.58602/jima-ilkom.v3i1.26

Kartini, K., & Candya CALLISTA, G. (2021). The Influence of Startup Business Characteristics on Investment Decisions of Business Angels: A Case Study in Indonesia. *Journal of Asian Finance*, *8*(6), 931–0938. https://doi.org/10.13106/jafeb.2021.vol8.no6.0931

Laksono, B., Syahidin, Y., & Yunengsih, Y. (2024). Implementasi Data Mining Klasterisasi Data Pasien Rawat Inap dengan Algoritma K-Means Clustering. *Jurnal Teknologi Sistem Informasi Dan Aplikasi*, *7*(2), 621–627. https://doi.org/10.32493/jtsi.v7i2.39354

Maori, N. A. (2023). METODE ELBOW DALAM OPTIMASI JUMLAH CLUSTER PADA K-MEANS CLUSTERING. *Jurnal SIMETRIS*, *14*.

Montanaro, B., Croce, A., & Ughetto, E. (2024). Venture capital investments in artificial intelligence. *Journal of Evolutionary Economics*, *34*(1), 1–28. https://doi.org/10.1007/s00191-024-00857-7

Nur, E., Diah, A. |, Septi, A., Ismayantika, F. |, Puspasari, D., Retno, E., … Rahman, F. (2022). *Tranformasi Bisnis Digital* (M. Si. , gradDipMgt,MBus Dr. Agus Hermawan & S. E. M. Si. AK. C. Dr. Nurika Restuningdiah, Eds.). Fakultas Ekonomi Universitas Nusantara PGRI Kediri.

Poolman, T. M., Townsend-Nicholson, A., & Cain, A. (2022). Teaching genomics to life science undergraduates using cloud computing platforms with open datasets. *Biochemistry and Molecular Biology Education*, *50*(5), 446–449. https://doi.org/10.1002/bmb.21646

Prasetio, A., Makmun, M., & Dwi, M. N. (2023). Analisis Gempa Bumi Di Indonesia Dengan Metode Clustering. *Bulletin of Information Technology (BIT)*, *4*(2), 338–343. https://doi.org/10.47065/bit.v3i1

Russels Stuart, & Norvig Peter. (2022). *GLOBAL EDITION*.

Safitri Juanita, & Cahyono, R. D. (2024). K-MEANS CLUSTERING WITH COMPARISON OF ELBOW AND SILHOUETTE METHODS FOR MEDICINES CLUSTERING BASED ON USER REVIEWS. *Jurnal Teknik Informatika (Jutif)*, *5*(1), 283–289. https://doi.org/10.52436/1.jutif.2024.5.1.1349

Schwab-McCoy, A., Baker, C. M., & Gasper, R. E. (2021). Data Science in 2020: Computing, Curricula, and Challenges for the Next 10 Years. *Journal of Statistics and Data Science Education*, *29*(S1), S40–S50. https://doi.org/10.1080/10691898.2020.1851159

Setty, R., Elovici, Y., & Schwartz, D. (2024). Cost-sensitive machine learning to support startup investment decisions. *Intelligent Systems in*

Techno Nusa Mandiri: Journal of Computing and Information Technology
Vol. 22, No. 1 March 2025 | DOI https://doi.org/10.33480/techno.v21i1.XXXX

Accounting, Finance and Management, *31*(1). https://doi.org/10.1002/isaf.1548

Siddik, A. B., Li, Y., & Du, A. M. (2024). Unlocking funding success for generative AI startups: The crucial role of investor influence. *Finance Research Letters*, *69*. https://doi.org/10.1016/j.frl.2024.106203

Wijaya, B. (2020). CLUSTERING MENGGUNAKAN METODE K-MEANS DENGAN BANTUAN METODE ELBOW. *Jurnal Komputer Dan Informatika*, *15*, 258–264.