Analysis of the K-Means Method and Segmentation of Open Unemployment Rates in Each Province in Indonesia in 2025

Peniel Sam Putra Sitorus[a*], Jaya Tata Hardinata[b], Dudes Manalu[c]
[a,b,c]Universitas HKBP Nommensen Pematangsiantar
Corresponding Author:
[a]peniel.sitorus@uhnp.ac.id

ABSTRACT
The Open Unemployment Rate (TPT) is a vital indicator for describing the state of employment in Indonesia. Variations in characteristics across provinces lead to fluctuations in TPT that require systematic analysis. This study aims to analyze and segment Indonesian provinces based on the Open Unemployment Rate using the K-Means clustering method. The data utilized consists of the TPT from 38 provinces in 2025 for the February and August periods. K-Means was applied to group provinces based on similarities in their unemployment rates. The results indicate that provinces can be categorized into several clusters representing low, medium, and high unemployment categories. This segmentation provides an overview of regional unemployment patterns and can serve as a foundation for formulating more targeted labor policies.

Keywords : K-Means, Clustering, Segmentation, Unemployment, Province

INTRODUCTION

The Open Unemployment Rate (TPT) is one of the primary indicators used to measure the health of the labor market in a country or region. High unemployment rates are often linked to various socio-economic impacts, such as a decline in purchasing power, rising poverty, and development disparities between regions. Therefore, an analysis of TPT data is essential for understanding the distribution patterns and characteristics of unemployment across every province in Indonesia(Adawiyah, Defit, & Sumijan, 2024).

Differences in socio-economic characteristics across Indonesian provinces cause significant variations in unemployment rates. These variations are influenced by factors such as economic structure, the quality of human resources, urbanization rates, and differing regional development policies. To understand the complex relationships between these variables, data analysis techniques capable of exploring patterns within multivariate datasets are required(Algiffary & Sutabri, 2023). One of the techniques frequently used in data exploration is clustering, a method of grouping data based on characteristic similarities without relying on pre-existing class labels(Aldi et al., 2025). The K-Means method is a popular clustering algorithm in the field of data mining because it is relatively simple and efficient in grouping data based on the centroid distance between observations. In general, K-Means can identify homogeneous data groups based on specific shared traits, such as the unemployment rates across provinces(Aprilia & Sembiring, 2021).

In the context of employment analysis in Indonesia, several studies using the K-Means approach have been conducted at specific regional levels. For instance, research analyzing

open unemployment by province in Indonesia through K-Means clustering techniques demonstrated that this method is valid for grouping based on the TPT characteristics of provinces(Azzam, Irma Purnamasari, & Ali, 2024). Other studies have also applied K-Means to map the TPT across Indonesian provinces, which is useful for policymakers in evaluating regions with similar unemployment rates(Baldah, Duarisah, & Maulana, 2023).

Furthermore, the application of K-Means in mapping open unemployment also appears in a collaborative study reviewing the application of clustering on West Java's Open Unemployment Data, where the resulting clusters are able to depict groups of regions with similar socio-economic conditions(Barata, Ayuni, Kartini, & Alawi, 2024). These studies demonstrate that K-Means serves as an essential tool for simplifying the identification of patterns that are not visible through descriptive statistics alone(Chong, 2021).

Internationally, the use of clustering techniques such as K-Means has also been widely applied to analyze unemployment rates and other socio-economic variables in various contexts, such as clustering based on demographic profiles to understand unemployment rate variations in specific regions. Such research supports the validity of the K-Means method in analyzing complex social data beyond the Indonesian context(Deti Karmanita & Billy Hendrik, 2023).

Based on existing theoretical foundations and empirical studies, this research seeks to apply the K-Means method to analyze and segment the Open Unemployment Rate (TPT) across 38 provinces in Indonesia using 2025 TPT data. The primary objective is to obtain groupings of provinces with similar unemployment patterns, thereby providing insights for policymakers in designing more effective and targeted unemployment reduction strategies. .

METHODS

This research utilizes a quantitative approach with data mining methods. The technique applied is clustering using the K-Means algorithm to group Indonesian provinces based on similarities in the Open Unemployment Rate (TPT)(Fahrillah Fahrillah & Zaehol Fatah, 2025). This approach was selected because it is capable of identifying patterns and data structures without requiring pre-existing class labels.

Algoritma K-Means

K-Means is a clustering algorithm within data mining used to group data into a specific number of clusters based on the degree of similarity between objects(Khadragy et al., 2022). This algorithm works by partitioning the data into K groups, where each data point is assigned to the cluster with the shortest distance to the cluster center (centroid). The primary objective of K-Means is to minimize the intra-cluster distance and maximize the inter-cluster distance(Laksono, Syahidin, & Yunengsih, 2024).

The K-Means process begins by determining the desired number of clusters (K). Subsequently, the algorithm randomly selects initial cluster centers (centroids). The distance of each data point to each centroid is then calculated using a distance measure, typically Euclidean Distance(Lee, Lee, & Park, n.d.). Each data point is assigned to the cluster with the shortest distance. Once all data points have been grouped, the centroid positions are updated by calculating the average value of the data within each respective cluster. This process is

repeated until there are no significant changes in cluster membership or centroid values(Madyatmadja et al., 2021)(Maoulana, Irawan, & Bahtiar, 2024).

In this study, the number of clusters was set to three (K=3), representing the categories of low, medium, and high unemployment rates. The clustering results were then analyzed to observe the distribution patterns of unemployment across provinces and used as a basis for interpreting employment policies(Naeem, Ali, Anam, & Ahmed, 2023)(Pujiono, Astuti, & Muhamad Basysyar, 2024).

Based on the research methodology, the **K-Means algorithm** is presented through the methodological schema in the figure below. Through the basic illustration of testing the dataset using the K-Means algorithm(Rafi Nahjan, Nono Heryana, & Apriade Voutama, 2023)(Rosyada & Utari, 2024).

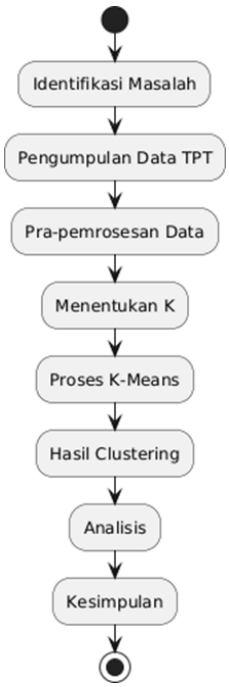

Figure 1. Data Processing Schema

RESULTS

Problem Identification

At this stage, the researcher defines the issues to be examined. In this study, the problem addressed is the disparity in the open unemployment rate (TPT) across provinces in Indonesia. These differences require analysis to identify unemployment patterns and to determine which regions have low, medium, or high unemployment rates(Rafi Nahjan et al., 2023).

TPT Data Collection

This stage aims to obtain the data to be analyzed. The data used is the Open Unemployment Rate (TPT) for 38 provinces in Indonesia, sourced from official providers such as the Central Bureau of Statistics (BPS) for specific periods (e.g., February and August 2025). This data serves as the primary foundation for the clustering process(Supriyadi, Triayudi, & Sholihati, 2021)(Wulandari & Yogantara, 2022).

Table 1. Dataset of Open Unemployment Rate by Province (Percentage)

Provinsi	Tingkat Pengangguran Terbuka Menurut Provinsi (Persen)	
	2025	
	Februari	Agustus
ACEH	5,5	5,64
SUMATERA UTARA	5,05	5,32
SUMATERA BARAT	5,69	5,62
RIAU	4,12	4,16
JAMBI	4,48	4,26
SUMATERA SELATAN	3,89	3,69
BENGKULU	3,24	3,41
LAMPUNG	4,07	4,21
KEP. BANGKA BELITUNG	4,17	4,45
KEP. RIAU	6,89	6,45
DKI JAKARTA	6,18	6,05
JAWA BARAT	6,74	6,77
JAWA TENGAH	4,33	4,66
DI YOGYAKARTA	3,18	3,46
JAWA TIMUR	3,61	3,88
BANTEN	6,64	6,69
BALI	1,58	1,49
NUSA TENGGARA BARAT	3,22	3,06
NUSA TENGGARA TIMUR	3,23	3,31
KALIMANTAN BARAT	4,23	4,82
KALIMANTAN TENGAH	3,47	3,97
KALIMANTAN SELATAN	3,94	4,16

KALIMANTAN TIMUR	5,33	5,18
KALIMANTAN UTARA	3,9	3,85
SULAWESI UTARA	6,03	5,99
SULAWESI TENGAH	3,02	2,92
SULAWESI SELATAN	4,96	4,21
SULAWESI TENGGARA	3,27	3,31
GORONTALO	3,12	3,42
SULAWESI BARAT	3,17	2,86
MALUKU	5,95	6,27
MALUKU UTARA	4,26	4,55
PAPUA BARAT	4,21	4,55
PAPUA BARAT DAYA	6,61	6,85
PAPUA	6,92	6,96
PAPUA SELATAN	4,9	4,04
PAPUA TENGAH	3,55	3,62
PAPUA PEGUNUNGAN	1,68	1,68

Each province is represented as a data pair:

$$p_i = (x_i, y_i)$$

Where :

p_i = Province i

x_i = TPT value for Period 1 (February)

y_i = TPT value for Period 2 (August)

Data Pre-processing

In the pre-processing stage, the collected data is prepared for processing. The activities performed include: checking for missing values, standardizing data formats (e.g., ensuring all are in percentage form), and performing normalization if necessary to prevent any single

variable from dominating the analysis. This stage is crucial for enhancing the quality of the clustering results.

To ensure the data is ready for K-Means processing, the following steps are performed:

a) Selecting only the 38 provinces (removing header/description rows and the "INDONESIA" row).

b) Defining the variables used:

x_i = TPT February 2025

y_i = TPT August 2025

c) Converting data to numeric format (percentage) to enable distance calculation.

d) Standardization/Normalization (optional but common) to ensure more stable K-Means distance calculations when using more than one variable (Supriyadi et al., 2021).

Determining K

This stage determines the number of clusters (K) to be used in the K-Means algorithm. In this study, a value of K = 3 is used to represent the categories of low, medium, and high unemployment. Determining K aims to ensure that the clustering results are easily interpretable and relevant to the research objectives.

K-Means Clustering Process

At this stage, the K-Means algorithm is run to cluster the data. The process includes: Determining the initial centroid randomly, Calculating the distance of each data point to each centroid, Grouping the data into the closest cluster, Updating the centroid based on the average of the data in each cluster, Repeating the process until the clustering results are stable. The result of this stage is the division of provinces into several clusters based on the similarity of TPT values.

The K-Means algorithm begins by randomly selecting three initial centroids from the data. These centroids are the temporary centers of each cluster.

Example of a centroid:

Table 2. For example, if you specify K=3, then the initial or final centroid shape.

Cluster	x_1 (Feb 2025)	x_2(Aug 2025)
Centroid 1 (c_1)	3.20	3.15
Centroid 1 (c_2)	5.45	5.60

For each province, the distance to each centroid is calculated using the Euclidean Distance:

$$d(p,c) = \sqrt{(x_{1p} - x_{1c})^2 + (x_{2p} - x_{2c})^2}$$

Where :

d = The distance between province p and centroid c.

x_{1p} = TPT February 2025 for the province.

x_{1c} = TPT February 2025 for the centroid.

x_{2p} = TPT August 2025 for the province.
x_{2c} = TPT August 2025 for the centroid.
in your analysis of the 38 provinces, this step determines the cluster membership for Bali based on its 2025 TPT data (x_1, x_2). The algorithm follows this logic:
a) Calculate d_1 : distance from Bali to Centroid 1 (low)
b) Calculate d_2 : distance from Bali to Centroid 2 (Medium)
c) Calculate d_3 : distance from Bali to Centroid 1 (High)
d) Comparision : min (d_1, d_2, d_3)
At this stage, each province already has a temporary cluster label.
Once all the data has been grouped, formula :

$$\mu_j = \frac{1}{n_j} \sum_{i \in S_j} x_i$$

Table 3. Clustering Results Table

Provinsi	Februari	Agustus	d_ C 0	d_ C 1	d_ C 2	Cluster	Kategori
Aceh	5.50	5.64	3.71	1.07	1.33	1	Tinggi
Sumatera Utara	5.05	5.32	3.17	1.62	0.63	2	Sedang
Sumatera Barat	5.69	5.62	3.83	0.95	0.69	1	Tinggi
Riau	4.12	4.16	1.69	3.09	0.69	2	Sedang
Jambi	4.48	4.26	2.02	2.77	1.46	2	Sedang
Sumatera selatan	3.89	3.69	1.20	3.58	1.49	2	Sedang
Bengkulu	3.24	3.41	0.55	4.24	1.86	0	Rendah
Lampung	4.07	4.21	1.69	3.09	0.69	2	Sedang
Kep. Bangka belitung	4.17	4.45	1.94	2.85	0.69	2	Sedang
Kep. Riau	6.89	6.45	4.91	0.20	2.54	1	Tinggi
DKI jakarta	6.18	6.05	4.10	0.21	2.39	1	Tinggi
Jawa barat	6.74	6.77	4.95	0.20	3.41	1	Tinggi

Jawa Tengah	4.33	4.66	2.03	2.54	0.63	2	Sedang
DI Yogyakarta	3.18	3.46	0.63	4.12	1.91	0	Rendah
Jawa timur	3.61	3.88	1.33	3.41	0.69	2	Sedang
Banten	6.01	6.27	4.06	0.63	2.39	1	Tinggi
Bali	2.35	2.31	1.46	5.35	2.39	0	Rendah
Nusa tenggara barat	3.09	3.07	0.63	4.51	1.91	0	Rendah
Nusa tenggara timur	2.79	2.67	0.69	4.90	2.54	0	Rendah
Kalimantan barat	4.51	4.24	2.04	2. 77	1.46	2	Sedang
Kalimantan tengah	4.11	4.02	1.69	3. 09	0.69	2	Sedang
Kalimanta n Selatan	4.41	4.01	1.91	3.41	1.46	2	Sedang
Kalimantan timur	5.33	5.18	3.17	1.62	0.63	2	Sedang
Kalimantan utara	3.90	3.85	1.33	3.71	1.04	2	Sedang
Sulawesi utara	6.03	5.99	4.03	0.63	2.39	1	Tinggi
Sulawesi tengah	3.02	2.92	0.63	4.51	1.91	0	Rendah
Sulawesi Selatan	4.96	4.21	2.39	2.77	1.46	2	Sedang

Sulawesi tenggara	3.27	3.31	0.69	4.24	1.46	0	Rendah
Gorontalo	3.12	3.42	0.69	4.12	1.46	0	Rendah
Sulawesi Barat	3.17	2.86	0.69	4.51	1.86	0	Rendah
Maluku	5.95	6.27	4.10	0.63	2.54	1	Tinggi
Maluku utara	4.26	4.55	1.94	2.85	0.69	2	Sed ang
Papua barat	4.21	4.55	1.94	2.85	0.69	2	Sedang
Papua barat daya	6.61	6.85	4.95	0.21	3.71	1	Tinggi
Papua	6.92	6.96	5.10	0.21	3.73	1	Tinggi
Papua Selatan	4.90	4.04	2.39	3.41	1.46	2	Sedang
Papua tengah	3.55	3.62	0.69	3.71	1.04	0	Rendah
Papua pegunungan	1.68	1.68	1.86	5.10	3.73	0	Rendah

Analysis

In the analysis stage, researchers examine the characteristics of each cluster, such as: the average unemployment rate (TPT) within each cluster, the provinces within each cluster, and the distribution of unemployment across regions. This analysis aims to understand employment conditions more deeply and identify areas requiring special attention.

At the analysis stage, researchers examine the characteristics of each cluster, including:

a) Average TPT per cluster,

b) List of provinces in each cluster,

c) Unemployment distribution patterns between regions.

Table 4. Average TPT per Cluster

Cluster	Kategori	Rata - rata Feb	Rata - rata Ags	Jumlah Provinsi
0	Rendah	≈ 3.0	≈ 3.0	11
2	Sedang	≈ 4.3	≈ 4.3	17
1	Tinggi	≈ 6.3	≈ 6.3	10

DISCUSSION

Open Unemployment Rate Clustering Results

Based on the application of the K-Means algorithm with the number of clusters K = 3, the Open Unemployment Rate (TPT) data for 38 provinces in Indonesia was successfully grouped into three main clusters: low, medium, and high. The clustering process used two variables, namely the February 2025 TPT and the August 2025 TPT, with Euclidean distance as the measure. Low Cluster (Cluster 0) consists of 11 provinces, Medium Cluster (Cluster 2) consists of 17 provinces, and High Cluster (Cluster 1) consists of 10 provinces. Each province is grouped into a cluster with the closest distance to the centroid, so that the grouping reflects the similarity of unemployment rates between regions.

Characteristics of Each Cluster

1) Low Cluster

This cluster comprises provinces with relatively low TPT in both observation periods. Some of the provinces included in this group include Bali, Bengkulu, Yogyakarta Special Region, West Nusa Tenggara, East Nusa Tenggara, Central Sulawesi, Southeast Sulawesi, Gorontalo, West Sulawesi, Central Papua, and Highland Papua. The average TPT in this cluster was around ±3% in both February and August 2025. This indicates that regions in the low-income cluster have relatively more manageable unemployment rates. This may be influenced by the regional economic structure, which is dominated by the informal sector, agriculture, and tourism (such as Bali), as well as the more flexible characteristics of the local labor market.

2) Medium Cluster

The moderate cluster is the group with the largest number of provinces, namely 17 provinces, including North Sumatra, Riau, Jambi, South Sumatra, Lampung, Bangka Belitung Islands, Central Java, East Java, West Kalimantan, Central Kalimantan, South Kalimantan, East

Kalimantan, North Kalimantan, South Sulawesi, North Maluku, West Papua, and South Papua. The average TPT in this cluster is around 4–5%. Regions in this group generally have a mixed economic structure between the primary, industrial, and service sectors. The moderate unemployment rate indicates that labor absorption is relatively stable, but challenges remain in providing jobs that keep pace with workforce growth.

3) High Cluster

The high cluster consists of the 10 provinces with the highest TPT values: Aceh, West Sumatra, Riau Islands, Jakarta, West Java, Banten, North Sulawesi, Maluku, Southwest Papua, and Papua. The average TPT in this cluster is around 6% or higher. These regions generally have high levels of urbanization, large population concentrations, and high labor migration flows. For example, Jakarta, West Java, and Banten are centers of national economic activity that attract many job seekers, resulting in fierce competition in the labor market and resulting in high open unemployment rates.

CONCLUSION

Based on the results of the study "Analysis and Segmentation of Open Unemployment Rates in Provinces in Indonesia Using the K-Means Method," it can be concluded that the application of the K-Means algorithm is effective in grouping 38 provinces in Indonesia based on the similarity of their Open Unemployment Rates (TPT) in 2025 for the February and August periods. Using two variables (TPT in February and TPT in August) and a cluster size of K=3, this study successfully segmented the provinces into three main categories: the low cluster (11 provinces), the medium cluster (17 provinces), and the high cluster (10 provinces). The segmentation results indicate that provinces in the low cluster had relatively small and stable TPT values across both periods, while provinces in the medium cluster exhibited moderate unemployment rates with less extreme variations. Meanwhile, the high cluster was filled with provinces with the highest TPT, indicating the need for special attention due to more competitive and complex labor market conditions. In general, the distribution pattern indicates that regions with urban/industrial economic characteristics tend to be in the high TPT group, while regions dominated by the informal sector or local economy tend to be in the low TPT group. Therefore, the results of this clustering can be used as a basis to help policymakers determine more targeted employment program priorities. Segmenting provinces based on low, medium, and high unemployment categories provides a clearer picture of unemployment patterns across regions and supports the formulation of data-driven unemployment reduction strategies.

REFERENCES

Adawiyah, Q., Defit, S., & Sumijan. (2024). Penerapan Algoritma K-Means Clustering untuk Mengelompokkan Rekomendasi Metode Kontrasepsi Berbasis Machine Learning di Puskesmas. *Jurnal KomtekInfo*, *11*(4), 300–305. doi:10.35134/komtekinfo.v11i4.563

Aldi, M. A., Fatah, Z., Informasi, T., Ibrahimy, U., Timur, S. J., Informasi, S., … Timur, S. J. (2025). Implementasi K-Means Cluster Ing Dalam Pengelompokan Data Kunjungan, *2*(1), 13–19.

Algiffary, A., & Sutabri, T. (2023). Indonesian Journal of Computer Science. *Indonesian*

Journal of Computer Science, *12*(2), 284–301. Retrieved from http://ijcs.stmikindonesia.ac.id/ijcs/index.php/ijcs/article/view/3135

Aprilia, K., & Sembiring, F. (2021). Analisis Garis Kemiskinan Makanan Menggunakan Metode Algoritma K-Means Clustering. *SISMATIK: Seminar Nasional Sistem Informasi Dan Manajemen Informatika*, 1–10.

Azzam, A., Irma Purnamasari, A., & Ali, I. (2024). Implementasi Algoritma K-Means Clustering Untuk Analisis Persebaran Umkm Di Jawa Barat. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(3), 3062–3070. doi:10.36040/jati.v8i3.8450

Baldah, A., Duarisah, A. V., & Maulana, R. A. (2023). Clustering Daerah Rawan Bencana Alam Di Indonesia Berdasarkan Provinsi Dengan Metode K-Means. *Jurnal Ilmiah Informatika Global*, *14*(2), 31–36. doi:10.36982/jiig.v14i2.3186

Barata, M. A., Ayuni, I. S., Kartini, A. Y., & Alawi, Z. (2024). Algorima K-Means Dalam Clustering Produk Skincare. *JIP (Jurnal Informatika Polinema)*, *10*(3), 421–428.

Chong, B. (2021). K-means clustering algorithm: a brief review. *Academic Journal of Computing & Information Science*, *4*(5), 37–40. doi:10.25236/ajcis.2021.040506

Deti Karmanita, & Billy Hendrik. (2023). Penerapan Metode Clustering dengan Algoritma K-Means pada Pengelompokkan Peminatan Mata Kuliah. *Jurnal Ilmiah Dan Karya Mahasiswa*, *1*(6), 01–10. doi:10.54066/jikma.v1i6.1028

Fahrillah Fahrillah, & Zaehol Fatah. (2025). Pengelompokan Data Nilai Siswa Madrasah Ta'Hiliyah Menggunakan Metode K-Means Clustering. *Jurnal Riset Sistem Informasi*, *2*(1), 53–59. doi:10.69714/0v1pkz05

Khadragy, S., Elshaeer, M., Mouzaek, T., Shammass, D., Shwedeh, F., Aburayya, A., … Aljasmi, S. (2022). Predicting Diabetes in United Arab Emirates Healthcare: Artificial Intelligence and Data Mining Case Study. *South Eastern European Journal of Public Health*, *2022*(Special Issue 5), 1–20. doi:10.56801/seejph.vi.406

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. doi:10.32493/jtsi.v7i2.39354

Lee, N., Lee, J., & Park, C. (n.d.). Augmentation-Free Self-Supervised Learning on Graphs, *1*(c).

Madyatmadja, E. D., Kusumawati, L., Jamil, S. P., Kusumawardhana, W., Informasi, S., & Nusantara, U. B. (2021). Infotech: journal of technology information. *Raden Ario Damar*, *7*(1), 55–62.

Maoulana, R., Irawan, B., & Bahtiar, A. (2024). Data Mining Dalam Konteks Transaksi Penjualan Hijab Dengan Menggunakan Algoritma Clustering K-Means. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(1), 515–521. doi:10.36040/jati.v8i1.8504

Naeem, S., Ali, A., Anam, S., & Ahmed, M. M. (2023). An Unsupervised Machine Learning Algorithms: Comprehensive Review. *International Journal of Computing and Digital Systems*, *13*(1), 911–921. doi:10.12785/ijcds/130172

Pujiono, S., Astuti, R., & Muhamad Basysyar, F. (2024). Implementasi Data Mining Untuk Menentukan Pola Penjualan Produk Menggunakan Algoritma K-Means Clustering. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(1), 615–620. doi:10.36040/jati.v8i1.8360

Rafi Nahjan, M., Nono Heryana, & Apriade Voutama. (2023). Implementasi Rapidminer Dengan Metode Clustering K-Means Untuk Analisa Penjualan Pada Toko Oj Cell. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(1), 101–104. doi:10.36040/jati.v7i1.6094

Rosyada, I. A., & Utari, D. T. (2024). Penerapan Principal Component Analysis untuk Reduksi Variabel pada Algoritma K-Means Clustering. *Jambura Journal of Probability and Statistics*, *5*(1), 6–13. doi:10.37905/jjps.v5i1.18733

Supriyadi, A., Triayudi, A., & Sholihati, I. D. (2021). Perbandingan Algoritma K-Means Dengan K-Medoids Pada Pengelompokan Armada Kendaraan Truk Berdasarkan Produktivitas. *JIPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, *6*(2), 229–240. doi:10.29100/jipi.v6i2.2008

Wulandari, L., & Yogantara, B. O. (2022). Algorithm Analysis of K-Means and Fuzzy C-Means for Clustering Countries Based on Economy and Health. *Faktor Exacta*, *15*(2), 109–116. doi:10.30998/faktorexacta.v15i2.12106