

Exploring Regional Development Patterns using Machine Learning: A Python-based Clustering Analysis of Human Development Index

Kartika Mariskhana^{1)*}, Ita Dewi Sintawati²⁾, Widiarina³⁾

^{1,2,3)} Universitas Bina Sarana Informatika, Indonesia

¹⁾ kartika.kma@bsi.ac.id, ²⁾ ita.ids@bsi.ac.id, ³⁾ widiarina.wda@bsi.ac.id

Submitted : Mar 5, 2024 | **Accepted** : Mar 18, 2024 | **Published** : Apr 1, 2024

Abstract: Human development has become the main focus for many local governments in efforts to improve the welfare and quality of life of their communities. In West Java, one of the most populous provinces in Indonesia, a deep understanding of human development patterns in various districts/cities is essential for formulating effective development policies. In this study, we explore regional development patterns using machine learning techniques, especially clustering analysis, taking the Human Development Index (HDI) as the main indicator as an example. The scope of this study includes HDI analysis for 27 districts/cities in West Java from 2017 to 2022. The main objective is to identify groups of districts/cities with human development characteristics, the first group consists of districts/cities with the highest level of human development, such as Bandung City and Cirebon City. The second group includes areas with a moderate level of development, while the third group consists of areas with human development that still needs to be improved, such as Ciamis and Cianjur. This clustering analysis provides valuable insights for policy makers in formulating more effective and inclusive development strategies. By understanding the differences in development characteristics in various regions, local governments can allocate resources more precisely and plan more sustainable development programs. By understanding the differences in development characteristics in various regions, local governments can allocate resources more precisely and plan more sustainable development programs.

Keywords: Human Development Index (HDI); machine learning; Clustering analysis; Python; Regional development

INTRODUCTION

Many local governments have placed human development as a top priority in efforts to improve people's welfare and quality of life (Dira, 2023). To create effective development policies in West Java, one of Indonesia's most populous provinces, it is important to understand human development patterns in each district (Iis Sandra Yanti, 2020). The Human Development Index (HDI), the research's main indicator, was evaluated using methods inspired by machine learning theory, in particular clustering analysis.

Machine learning is a branch of artificial intelligence that allows computers to learn from data without needing to be explicitly programmed (Arifin, 2021). In the context of clustering analysis, machine learning algorithms allow us to group data into groups that have similar patterns or characteristics. One of the machine learning algorithms that is often used for clustering analysis is K-Means (Febrianty, 2023).

The scope of this study includes HDI analysis for 27 districts/cities in West Java from 2017 to 2022. The main objective is to identify groups of districts/cities with similar human development characteristics, as well as analyze the differences that exist between them. By using the K-Means algorithm for clustering and the Python programming language as the main tool, we present a solution that enables better mapping and understanding of human development patterns in West Java.

In this context, this research aims to fill the knowledge gap by applying the K-means clustering algorithm to analyze the HDI of West Java Province from 2015 to 2022. Research problems include: human development patterns in West Java Province can be identified using clustering analysis, factors -factors that influence human

*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

development patterns in the region, the results of clustering analysis can provide useful insights for policy makers in designing more effective development programs in West Java Province. The objectives of this research are: Applying the K-means clustering algorithm to identify human development patterns in West Java Province, Analyzing the factors that influence human development patterns in the region, Providing useful insights for policy makers in designing better development programs effective in West Java Province.

Previous studies have highlighted the importance of clustering analysis in understanding human development patterns at the regional level (Basalamah, 2023). For example, research by (name of researcher) uses the K-means clustering algorithm to analyze HDI in several provinces in Indonesia. However, there has been no research that specifically focuses on HDI clustering analysis in West Java Province.

Data taken from the West Java Province Central Statistics Agency (BPS) is official data provided by government agencies responsible for collecting, processing and providing statistical data at the provincial level. The following is some information about data sources taken from BPS West Java Province:

1. **Data Reliability:** Data provided by BPS West Java Province is considered to have a high level of reliability because it is the result of an official survey and a standardized data collection process.
2. **Data Collection Methodology:** West Java Province BPS usually uses time-tested survey and data collection methods to collect information related to the human development index (HDI). This methodology may involve field surveys, interviews, data collection from relevant agencies, and other data sources.
3. **Consistency and Validity:** BPS West Java Province ensures the consistency and validity of data by carrying out data verification and validation processes before the data is published. This aims to ensure that the data provided is reliable and representative.
4. **Data Accessibility:** West Java Province BPS provides data openly through their official website or other platforms. This allows researchers, observers and the general public to freely access the data and use the data for analysis and research.
5. **Data Legality:** Data provided by BPS West Java Province has strong legality because it is official data issued by a legitimate government agency.

By obtaining data from reliable sources such as BPS West Java Province, researchers can ensure that the data used in their research is of good quality and can be relied upon for further analysis.

LITERATURE REVIEW

Previous studies that the author uses as references include the following:

1. Research conducted by Ferista Wahyu Saputri, et al, 2023. This research compares the K-Means, K-Medoids, and DBSCAN clustering methods for clustering provinces in Indonesia based on community welfare indicators using SNE dimensionality reduction data. The results of the clustering evaluation are based on the highest Silhouette coefficient and The lowest Davies-Bouldin index obtained by the best clustering method is K-Means and DBSCAN with parameters per plexity=1, minPts=2 and epsilon=9. Both obtained the same results, namely eight clusters were formed. Cluster 1 and cluster 7 are provinces that have the lowest level of welfare compared to other clusters, so they require more serious and intensive improvement efforts. Provinces in these clusters include the provinces of Aceh, Maluku, West Nusa Tenggara, East Nusa Tenggara, Central Sulawesi, Gorontalo , West Sulawesi, West Papua and Papua. Meanwhile, the provinces in the gang only need improvement efforts in certain aspects(Ferista Wahyu Saputri, 2023).
2. Research conducted by Amelia, et al, 2023. From the research conducted, it was found that there were three best district/city clusters using the elbow method in West Java province. Cluster 1 (districts/cities with quite high levels of poverty), cluster 2 (districts/cities with low levels of poverty), cluster 3 (districts/cities with high levels of poverty). Districts/cities that fall into the category with high levels of poverty are Sukabumi, Cianjur, Garut, Tasikmalaya, Ciamis, Kuningan, Cirebon, Majalengka, Indramayu, Subang, West Bandung, and Pangandaran(Nur, 2023).
3. Research conducted by Yogiek Indra Kurniawan, et al, 2023. This research aims to carry out clustering to group the priorities of countries that need assistance. The method used in grouping countries uses the K-Means algorithm with the Elbow and Silhouette methods. The tool used in grouping is Python. Clustering and Silhouette Coefficient search were carried out using Orange Tools. The dataset used includes information about countries around the world. The results of this research are clustering of countries included in groups C5 to C1, with the highest priority needs in C5 and the lowest in C(Kurniawan, 2023).

The Elbow method is a technique commonly used to determine the optimal number of clusters in a population. This is one of the important techniques in clustering analysis, especially in the K-Means algorithm. The program uses the Elbow Method technique to determine the optimal number of clusters in Regency/City population data in West Java. The Elbow method plots the inertia value against the number of different clusters, then looks for the

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

point where the decrease in inertia is no longer significant (elbow point), which indicates the most optimal number of clusters for clustering analysis. After the optimal number of clusters is determined, the program performs clustering using the K-Means algorithm. This algorithm helps group data into clusters that have similar characteristics. In this way, the program can provide valuable insight in understanding patterns in West Java Regency/City HDI data during the 2017-2022 period. The following are details of the elbow method theory (Priyambadha, 2020).

1. **Inertia:** Inertia, also known as Within-Cluster Sum of Squares (WCSS) (Laurence, 2021), is a metric used in the K-Means algorithm to measure how dense the resulting clusters are. Mathematically, inertia is the sum of the squared distances between each data point in a cluster and the cluster center.
2. **Goal:** Our goal in clustering analysis is to minimize inertia. The lower the inertia, the denser the resulting clusters, and the better our clustering mode.
3. **Determining the Number of Clusters:** One of the challenges in the K-Means algorithm is determining the optimal number of clusters (Adhitama, 2020). The Elbow method helps in solving this challenge by trying several values of the number of clusters and plotting the inertia for each value..

Elbow Method Steps:

1. **Select Cluster Value Range:** Start by determining the value range for the number of clusters you want to test. Typically, this value range starts from 1 and ends at the maximum number of clusters that is practically possible.
2. **Calculate Inertia:** For each value of the number of clusters in the selected range, run the K-Means algorithm and calculate the inertia for the resulting clusters.
3. **Elbow Curve Plot:** Make a plot of the inertia value against the number of clusters. You will notice that as the number of clusters increases, the inertia will tend to decrease. At some point, this decline will slow significantly. This point is called the "elbow" of the curve.
4. **The Elbow formula** in the context of clustering analysis with the K-Means algorithm refers to the calculation of inertia (Within-Cluster Sum of Squares, WCSS) for each value of the number of clusters tested. Inertia is a measure of cluster density, which is the sum of the squares of the distance between each data point in a cluster and its cluster center (Rahmat, 2019).

The general formula for calculating inertia is as follows:

$$\text{Inertia} = \sum_{i=1}^N \sum_{j=1}^k w_{ij} \|x_i - \mu_j\|^2 \tag{1}$$

Information:

1. N is the total number of data points.
2. k is the number of clusters tested.
3. w_{ij} is a function indicator that shows whether point x_i is included in cluster j or not (1 if i is included in cluster j , and 0 if not).
4. $\|x_i - \mu_j\|^2$ is the squared distance between the data point x_i and the cluster center μ_j

In the context of the Elbow method, we calculate inertia for various values of k (number of clusters), then look for "elbow" points in the plot of inertia against the number of clusters (Mutawalli, 2023). This point is the point at which the decrease in inertia begins to slow down significantly, indicating that adding more clusters does not provide a significant decrease in cluster density. That is the optimal number of clusters selected.

METHOD

The stages in this research include research steps. The framework in this research is described as follows:

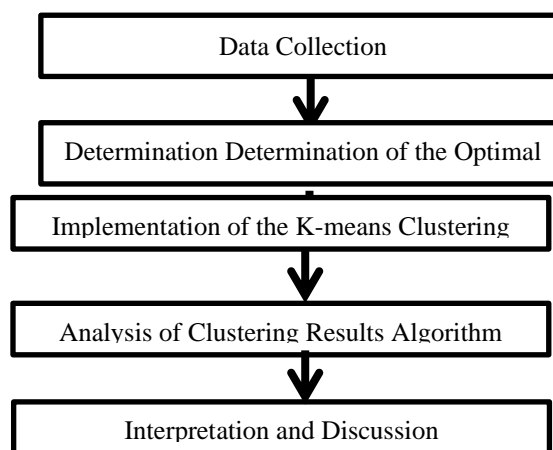


Figure 1. Research Framework

*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

In Figure 1. Data Collection Download West Java Province human development index (HDI) data from the official website of the West Java Province Central Statistics Agency (BPS). Organize data into a format suitable for analysis, for example in the form of a CSV or Excel file (Provinsi Jawa Barat, 2024). Data Preparation: Selects the year column from the dataset for clustering analysis. Standardize data where necessary to ensure that each feature is of similar scale (Alam, 2024). Determination of the Optimal Number of Clusters: Use the Elbow method or other methods to determine the optimal number of clusters. Analyze inertia or other metrics to find the point where adding clusters is no longer significant in reducing data variance. Implementation of the K-means Clustering Algorithm: Using Python libraries or modules such as scikit-learn to implement the K-means clustering algorithm. Specifies a predefined number of clusters. Carrying out a clustering process on West Java Province HDI data. Analysis of Clustering Results: Analyzing clustering results to identify patterns of human development in West Java Province. Examine the characteristics of each cluster, including the cluster center and the distribution of data within it. Interpretation and Discussion: Interpret the results of clustering analysis to explore factors that might influence human development in each cluster. Discusses the implications of research results for understanding human development in West Java Province. Comparing research results with previous findings or existing theories.

RESULT

On the elbow graph, the analysis you do on the graph. Based on the pseudocode program as follows:

1. Read data from CSV file.
2. Convert data to numeric type and replace commas with periods.
3. Select the district/city column and human development index for 2017-2022.
4. Standardize features using StandardScaler.
5. Determine the optimal number of clusters using the elbow method:
 - a. Initialize an empty list for inertia.
 - b. Loop over a range of cluster numbers from 1 to 10:
 - i. Instantiate a KMeans object with the current cluster number.
 - ii. Fit the scaled data to the KMeans object.
 - iii. Append the inertia (intra-cluster sum of squares) to the inertia list.
6. Plot the inertia values against the number of clusters to visualize the elbow method.
7. Select the optimal number of clusters based on the elbow point.
8. Perform clustering using KMeans with the optimal number of clusters:
 - a. Instantiate a KMeans object with the selected number of clusters.
 - b. Fit the scaled data to the KMeans object.
9. Add cluster labels to the data.
10. Display the clustering results by printing the district/city names and their corresponding clusters.
11. Visualize the clustering results:
 - a. Create a figure for the plot.
 - b. Loop over each cluster:
 - i. Filter data for the current cluster.
 - ii. Scatter plot the data points for each year within the cluster.
 - c. Set axis labels, title, legend, and grid for the plot.
 - d. Show the plot.

Elbow Graph Output Results:

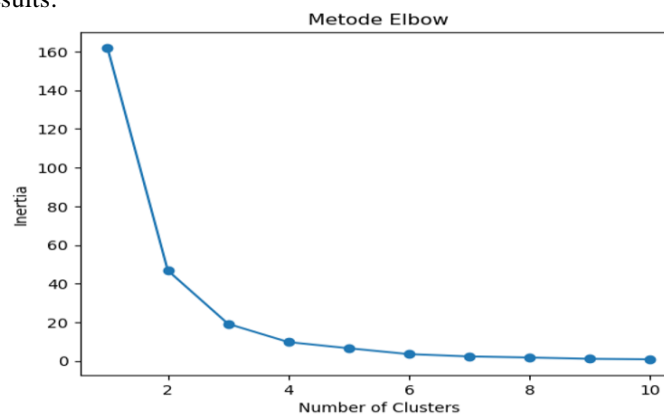


Figure 2. Elbow graph

*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

In Figure 2. X-axis (Number of Clusters): The X-axis represents the number of clusters considered in the analysis. Here, the range of number of clusters is from 1 to 10. Y-axis (Inertia): Y-axis represents the inertia value for each number of clusters. Inertia is a measure of how far the data points in a cluster are from the cluster center. From the graph, we see that the inertia value decreases significantly as the number of clusters increases from 1 to around 3 or 4. After reaching this point (called the "elbow" of the graph), the decrease in the inertia value becomes more gradual as additional clusters are added.

This "elbow" point indicates where adding clusters no longer provides a significant reduction in inertia. Therefore, the number of clusters around the elbow point is the optimal number of clusters for the dataset being analyzed.

Output Results:

	Kabupaten/Kota	Cluster
0	Bogor	0
1	Sukabumi	2
2	Cianjur	2
3	Bandung	0
4	Garut	2
5	Tasikmalaya	2
6	Ciamis	0
7	Kuningan	2
8	Cirebon	2
9	Majalengka	2
10	Sumedang	0
11	Indramayu	2
12	Subang	2
13	Purwakarta	0
14	Karawang	0
15	Bekasi	0
16	Bandung Barat	2
17	Pangandaran	2
18	Kota Bogor	1
19	Kota Sukabumi	0
20	Kota Bandung	1
21	Kota Cirebon	0
22	Kota Bekasi	1
23	Kota Depok	1
24	Kota Cimahi	1
25	Kota Tasikmalaya	0
26	Kota Banjar	0

Figure 3. Cluster

In Figure 3. From the elbow graph, we can see that there is a significant decrease in the inertia value when the number of clusters is increased from 1 to around 3 or 4. After reaching this number of clusters, the decrease in the inertia value becomes more gentle. The point where the graph shows an elbow is around cluster number 3 or 4. This shows that additional clusters after the elbow point do not provide a significant decrease in inertia values. Therefore, the optimal number of clusters for such a dataset is around 3 or 4. Thus, this elbow graph provides a strong indication of the optimal number of clusters for the data being analyzed. This provides useful guidance in selecting the number of clusters for further clustering analysis. Based on the pseudocode program as follows :

1. Create a plot figure with a specific size (12 inches width and 8 inches height).
2. Iterate over each cluster from 0 to n_clusters-1:
 - a. Filter the data to get cluster_data where 'Cluster' column is equal to the current cluster number.
 - b. Scatter plot the data points for each year within the cluster_data:
 - i. Scatter plot for '2017' vs '2018' with label 'Cluster {cluster}'.
 - ii. Scatter plot for '2018' vs '2019' with label 'Cluster {cluster}'.
 - iii. Scatter plot for '2019' vs '2020' with label 'Cluster {cluster}'.
 - iv. Scatter plot for '2020' vs '2021' with label 'Cluster {cluster}'.
 - v. Scatter plot for '2021' vs '2022' with label 'Cluster {cluster}'.
3. Set the x-axis label as 'Indeks Pembangunan Manusia Tahun X'.
4. Set the y-axis label as 'Indeks Pembangunan Manusia Tahun X+1'.
5. Set the title of the plot as 'Hasil Clustering KMeans'.
6. Add a legend to the plot.
7. Show grid lines on the plot.

8. *Display the plot.* Visualization of Clustering Results: This program produces visualization of clustering results using scatter plots. Each cluster is displayed in a different color and symbol. A for loop is used to iterate through each cluster and display the data points on the plot for each pair of years 2017-2022. After clustering, the program provides an output that presents information about each cluster. For each cluster (0, 1, and 2), the program lists the districts/cities with the highest and lowest human development index. These output examples provide insight into the differences in the index human development among its clusters.

Output Results:

```
Cluster 0:
Kabupaten/Kota with the highest human development index: Kota Cirebon - 74.0
Kabupaten/Kota with the lowest human development index: Ciamis - 68.87
Cluster 1:
Kabupaten/Kota with the highest human development index: Kota Bandung - 80.31
Kabupaten/Kota with the lowest human development index: Kota Bogor - 75.16
Cluster 2:
Kabupaten/Kota with the highest human development index: Kuningan - 67.78
Kabupaten/Kota with the lowest human development index: Cianjur - 63.7
```

Figure 4. Cluster Grouping Interpretasi

*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

In Figure 4. Cluster Grouping, as follows ut;

1. Cluster 0 has Cirebon City as the district/city with the highest human development index and Ciamis as the lowest.
2. Cluster 1 has Bandung City as the district/city with the highest human development index and Bogor City as the lowest.
3. Cluster 2 has Kuningan as the district/city with the highest human development index and Cianjur as the lowest.

The results of this program provide useful information about dividing districts/cities into clusters based on their human development index, as well as visualizing these data patterns graphically. This helps in understanding the differences and patterns that may exist between these groups of districts/cities, as well as providing useful insights for policy makers in development planning at the regional level. Based on the pseudocode program as follows:

1. Iterate over each cluster from 0 to n_clusters-1:
 - a. Filter the data to get cluster_data where 'Cluster' column is equal to the current cluster number.
 - b. Scatter plot the data points for each year within the cluster_data:
 - i. Scatter plot for '2017' vs '2018' with label 'Cluster {cluster}'.
 - ii. Scatter plot for '2018' vs '2019' with label 'Cluster {cluster}'.
 - iii. Scatter plot for '2019' vs '2020' with label 'Cluster {cluster}'.
 - iv. Scatter plot for '2020' vs '2021' with label 'Cluster {cluster}'.
 - v. Scatter plot for '2021' vs '2022' with label 'Cluster {cluster}'.
2. Set the x-axis label as 'Year'.
3. Set the y-axis label as 'Human Development Index'.
4. Set the title of the plot as 'KMeans Clustering Results'.
5. Add a legend to the plot.
6. Show grid lines on the plot.
7. Display the plot.

Output results:

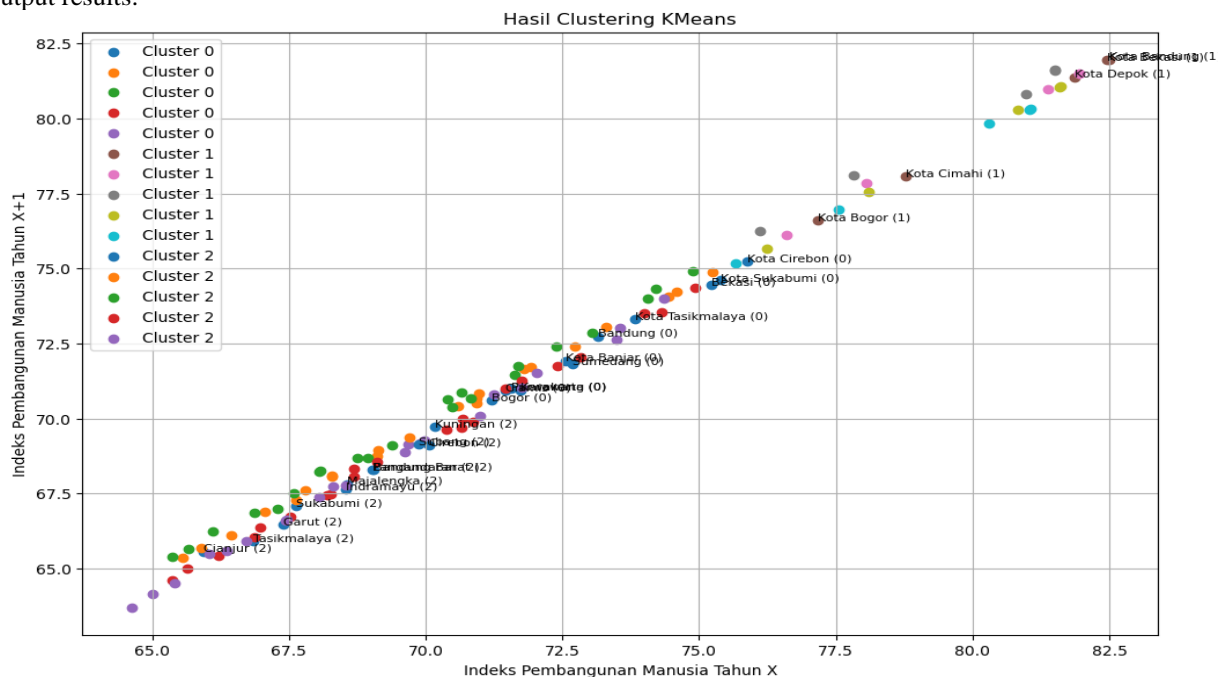


Figure 5. K-means Clustering Graph

In Figure 5. The graph displays a scatter plot of the clustering results using the K-Means algorithm. Each point on the plot represents one district/city in West Java. This graph is divided into three different clusters, represented by a different color for each cluster.

1. Suma y (vertical) shows the human development index from 2017 to 2022.
2. Each cluster is shown in a different color, with each dot representing one district/city.
3. The points on the graph represent the relative position of each district/city in the feature space based on their human development index values over the observed time period.

By looking at this graph, you can identify patterns or trends in human development between clusters, as well as differences in the level of development between districts/cities in each cluster. This provides useful visual insight into data distribution and cluster formation in clustering analysis. This graph helps in understanding the data

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

structure and interpreting clustering results in a more intuitive way, making it easier to make decisions and policy planning at the regional level.

DISCUSSIONS

Based on the results of clustering using the K-Means algorithm on human development index (HDI) data for districts/cities in West Java from 2017 to 2022, the results of the clustering model research are as follows:

1. Differences in Levels of Human Development: There are significant variations in levels of human development among districts/cities in West Java. Clustering identifies groups of districts/cities with similar human development characteristics, as well as highlighting the differences that exist between them.
2. Cluster Identification: Through clustering, we can identify three main groups of districts/cities that have similar human development patterns. Each cluster may have unique characteristics, such as a relatively high or low level of human development.
3. Importance of Clustering Analysis: Clustering analysis is a useful tool for understanding variations in human development in a region. By dividing data into groups based on existing patterns, we can identify areas with similar needs and design more effective and efficient development policies.
4. Policy Considerations: Clustering results can provide valuable insights for policymakers in planning and allocating development resources. By understanding the differences and similarities in human development between clusters, policies can be designed to better suit the needs of each region.

Thus, clustering analysis becomes an important tool in formulating more effective and efficient development policies, and allows the government to better respond to community needs.

CONCLUSION

Research studies have revealed interesting human development patterns in West Java through clustering analysis using machine learning. By utilizing the Human Development Index (HDI) as the main indicator, we succeeded in identifying three main groups of districts/cities with different human development characteristics. First, the first group consists of districts/cities with the highest level of human development, such as Bandung City and Cirebon City. The second group includes areas with a moderate level of development, while the third group consists of areas with human development that still needs to be improved, such as Ciamis and Cianjur. This clustering analysis provides valuable insights for policy makers in formulating more effective and inclusive development strategies. By understanding the differences in development characteristics in various regions, local governments can allocate resources more precisely and plan more sustainable development programs. In conclusion, the machine learning-based clustering analysis approach has proven to be a powerful tool in understanding regional development patterns in West Java and has the potential to be applied in other regional development contexts.

REFERENCES

- Adhitama, R. (2020). Penentuan Jumlah Cluster Ideal Smk Di Jawa Tengah Dengan Metode X-Means Clustering Dan K-Means Clustering Determining Vocational Ideal Cluster Number in Central Java With X-Means Clustering and K-Means Clustering Methods. *Jurnal Informatika Dan Komputer) Akreditasi KEMENRISTEKDIKTI*, 3(1), 1–5. <https://doi.org/10.33387/jiko>
- Alam, R. G. P. (2024). Optimasi K-Means Dengan Particle Swarm Optimization Dalam Penentuan Titik Awal Pusat Klaster Data Telekomunikasi. *Techno.COM*, 23(1), 96–111.
- Arifin, I. (2021). Penerapan Computer Vision Menggunakan Metode Deep Learning pada Perspektif Generasi Ulul Albab. *Jurnal Teknologi Terpadu*, 7(2), 98–107. <https://doi.org/10.54914/jtt.v7i2.436>
- Basalamah, A. T. (2023). Penerapan Algoritma K-Means Clustering Pada Tingkat Penyelesaian Pendidikan Di Provinsi Indonesia. *Jurnal Informatika Dan Teknologi Komputer*, 4(2), 114–121. <https://ejournalunsam.id/index.php/jicom/>
- Dira, A. F. (2023). Pengaruh Investasi dan IPM terhadap Pertumbuhan Ekonomi Hijau di Provinsi Kalimantan Timur. *EKOMBIS REVIEW: Jurnal Ilmiah Ekonomi Dan Bisnis*, 11(2), 1437–1446–1437–1446. <https://jurnal.unived.ac.id/index.php/er/article/view/4181>
- Febrianty, E. (2023). Optimalisasi Strategi Pemasaran dengan Segmentasi Pelanggan Menggunakan Penerapan K-Means Clustering pada Transaksi Online Retail Optimizing Marketing Strategies with Customer Segmentation Using K-Means Clustering on Online Retail Transactions. *Jurnal Teknologi Dan Informasi (JATI)*, 13(September), 122–137. <https://doi.org/10.34010/jati.v13i2>
- Ferista Wahyu Saputri. (2023). Perbandingan Performa Algoritma K-Means, K-Medoids, Dan DbSCAN Dalam Penggerombolan Provinsi Di Indonesia Berdasarkan Indikator Kesejahteraan Masyarakat. *Jurnal Teknologi Informasi: Jurnal Keilmuan Dan Aplikasi Bidang Teknik Informatika*, 7(2), 138–151. <https://doi.org/10.47111/jti.v7i2.9558>
- Iis Sandra Yanti. (2020). Faktor-Faktor Yang Menyebabkan Target Indeks Pembangunan Manusia Sulit Tercapai:

*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

- Sebuah Analisis Terhadap Lingkungan Strategis. *Administrasi Pemerintahan Daerah*, 4(1), 1–23.
- Kurniawan, Y. I. (2023). Pengelompokan Prioritas Negara Yang Membutuhkan Bantuan Menggunakan Clustering K-Means dengan Elbow dan Silhouette. *Jurnal Pendidikan Dan Teknologi Indonesia (JPTI)*, 3(10), 455–463. <https://doi.org/10.52436/1.jpti.343>
- Laurence. (2021). Analisis Peramalan dan Pengelompokan Jumlah Turis ke Jepang. *Journal of Integrated System*, 4(2), 150–167. <https://doi.org/10.28932/jis.v4i2.3164>
- Mutawalli, L. (2023). Komparasi Metode Perhitungan Jarak K-Means Paling Baik Terhadap Pembentukan Pola Kunjungan Wisatawan Mancanegara. *Journal of Information System Research (JOSH)*, 5(1), 159–166. <https://doi.org/10.47065/josh.v5i1.4377>
- Nur, M. (2023). Pengelompokan Tingkat Kemiskinan di Provinsi Jawa Barat dengan Metode K-Means Clustering Info Artikel. *Journal of Data Insights*, 1(2), 51–61. <http://journalnew.unimus.ac.id/index.php/jodi>
- Priyambadha, B. (2020). Penggalan Perilaku Pemain dalam Penentuan Tipe Permainan pada E-Learning Pemrograman Berbasis Gamification. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 7(4), 765–772. <https://doi.org/10.25126/jtiik.2020701295>
- Provinsi Jawa Barat. (2024). *jabar.bps.go.id*. <https://jabar.bps.go.id/indicator/26/123/1/indeks-pembangunan-manusia.html>
- Rahmat, A. (2019). Jurnal Teknik Mesin : Vol . 08 , No . 1 , Februari 2019. *Teknik Mesin*, 08(1), 25–29.

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.