

Clustering of Child Stunting Data in Tangerang Regency Using Comparison of K-Means, Hierarchical Clustering and DBSCAN Methods

Muhammad Azzam Robbani, Gerry Firmansyah, Agung Mulyo Widodo, Budi Tjahjono

Universitas Esa Unggul, Indonesia

E-mail: azzamrobbani23@student.esaunggul.ac.id, gerry@esaunggul.ac.id,
agung.mulyo@esaunggul.ac.id, budi.tjahjono@esaunggul.ac.id

*Correspondence: gerry@esaunggul.ac.id

KEYWORDS

Stunting, k-means;
hierarchical clustering;
DBSCAN; silhouette
score; nutritional status;
preventive actions;
curative measures

ABSTRACT

This study aims to analyze stunting in children in Tangerang Regency using clustering methods such as k-means, Hierarchical Clustering with Agglomerative Nesting, and Density-Based Spatial Clustering of Applications with Noise (DBSCAN). Stunting is a significant health issue affecting child growth due to chronic malnutrition and recurrent infections. The research revealed that k-means produced the best clustering results with a Silhouette Score of 0.52, indicating its effectiveness in categorizing children based on age, nutritional status, and stunting risk. The k-means method identified three clusters: Cluster 0 (ages 46-55 months, good nutrition, no stunting), Cluster 1 (ages 9-18 months, varied nutritional status, high stunting risk), and Cluster 2 (ages 27-36 months, good nutrition, no stunting). The study suggests preventive actions such as balanced nutrition education, regular health monitoring, complete immunizations, and physical activity, alongside curative measures like nutritional consultations and supplements. The findings provide a framework for targeted preventive and curative interventions, enabling Tangerang Regency's health department to effectively address and reduce stunting rates.

Attribution- Share Alike 4.0 International (CC BY-SA 4.0)



Introduction

Indonesia Emas is the hope of the current government in the future. There are several aspects needed to consider between human resources, economy, and culture to achieve. In human resources, the government has two main aspects that need to be considered and are related with each other, namely education and public health (Fahrudin, 2018). Education and public health are a unity that must be achieved. One of the steps that the government can take is to pay attention to the health of future generations, namely Indonesian children (Agapito et al., 2022). One of the problems related to the health of Indonesian children is stunting. Stunting according to WHO (2015) is a disorder of

children's growth and development due to chronic malnutrition and recurrent infections which is marked by their length or height being below the standard which will ultimately have a long-term impact on the country's generation both in the form of health impacts and economic impacts (H Witten et al., 2017). In article 141 of Law no. 36 of 2009 concerning health, all forms of efforts to improve community nutrition have been regulated by using reference standards used for determining nutritional status by anthropometry based on the Decree of the Minister of Health No.1995/Menkes/SK/XII/2010 with a standard reference of the World Health Organization-National Center for Health Statistics (WHO-NCHS) (Yudoprakoso, 2019). The government has also been serious about handling and solving this problem with the existence of Presidential Regulation No. 42 of 2013 concerning the National Movement to Accelerate Nutrition Improvement. The government under its leadership has set strategic measures to improve community nutrition, including stunting prevention and handling (Kusumawati, 2008).

Stunting is a condition of growth failure in children (body and brain growth) due to malnutrition for a long time. So that the child is shorter than a normal child of his age and has a delay in thinking. Malnutrition for a long time occurs from the fetus in the womb to the beginning of the child's life (the first 100 days) (Prendergast & Humphrey, 2014). Statistically, stunting has decreased from 29% in 2015 to 27.6% last year. However, this figure is still above the limit set by the World Health Organization (WHO) which is 20% (Ministry of Health, 2018). The problem of stunting is a nutritional problem faced by the world, especially poor and developing countries. Several factors cause the high incidence of stunting in toddlers and many people are not aware of stunting as a problem compared to other malnutrition problems (Mitra, 2015).

In Indonesia, in general, there are 2 types of interventions that are common in stunting prevention efforts, namely preventive and curative (Black et al., 2013). Preventive efforts such as health education, mentoring and counselling for pregnant women directly or through applications while the second curative effort is the provision of additional nutrients, vitamins, protein intake, reduction of MDA levels and increase of hemoglobin to increase the weight and body weight of stunted children (Nadirawati et al., 2023). Stunting is a significant health problem in Indonesia, this problem occurs in almost every region, including in Tangerang Regency. In Tangerang Regency itself, the number of stunted children under five can be said to be quite large with a total of 5,391 (<https://opendata.tangerangkab.go.id/stunting>, 2024) so this is still a special concern for the Tangerang Regency government which is committed to reducing the stunting rate to 0% which is monitored and evaluated periodically (Shalev-Shwartz & Ben-David, 2014). Efforts to reduce stunting carried out by the government such as monitoring the nutritional status of the community, empowering postnatal health care and the DASHAT (Healthy Kitchen for Stunting) program are in the form of providing nutritious food to children at high risk of stunting (Syukron et al., 2022). Programs or interventions carried out for stunting prevention must pay close attention to aspects of each input and process in order to get optimal output (Huljannah & Rochmah, 2022). Therefore, for each stunting program and control, it is necessary to monitor and evaluate to ensure the effectiveness of this program because the stunting problem is flexible and static that continues to move which is influenced by various factors, especially for Tangerang Regency itself has 29 sub-districts with a total population of approximately 3.5 million people (Hammond et al., 2015).

Therefore, looking at the background that has been described earlier. It is hoped

Clustering of Child Stunting Data in Tangerang Regency Using Comparison of K-Means, Hierarchical Clustering and DBSCAN Methods

that this research will be able to provide analysis and provide input or suggestions that can be used as action and evaluation materials for the Tangerang Regency government in an effort to reduce stunting rates in areas that are the target of the health office's work program. Therefore, the title of this study is "Clustering of Child Stunting Data in Tangerang Regency Using a Comparison of K-Means, Hierarchical Clustering and DBSCAN Methods".

Research Methods

The methodology used in this study is data mining to find patterns and relationships with previous data collection and use. Data mining includes a blend of techniques from various fields such as statistics, databases and data warehouses, machine learning, information retrieval, signal processing and analyzed spatial or temporal data. Another term for this process is commonly known as Knowledge Data Discovery (KDD) which in general prediction and description is a technique in data mining. In this method, clustering is carried out. There are many pentodes for clustering such as K-means which have the advantage of being effective for data with a relatively uniform distribution and a known or estimable number of clusters.

Results and Discussions

Based on each clustering method, namely k-means, *Hierarchical Clustering with Agglomerative Nesting* and *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN), clusters have been formed that according to the calculations of each method and provide a good overview and understanding related to the *stunting* problem in Tangerang Regency. However, it is necessary to conduct testing using the best method that is suitable to be applied to the Tangerang Regency Health Office. in this study to get the best method using the *Silhouette Score calculation* (Fadliana, 2015). The results obtained are as follows:

```
46 from sklearn.metrics import silhouette_score
47 # Menghitung Silhouette Score
48 silhouette_avg = silhouette_score(X, kmeans.labels_)
49 print(f'Silhouette Score: {silhouette_avg}')
50 |
```

PROBLEMS OUTPUT TERMINAL PORTS DEBUG CONSOLE

```
~/Documents/Information Tech/python$ cd k-means/
~/Documents/Information Tech/python/k-means$ python3 kmeans_clustering.py
Silhouette Score: 0.5202817531537228
```

Figure 1 Code Silhouette Score k-means

```
58 from sklearn.metrics import silhouette_score
59 # Menghitung Silhouette Score
60 silhouette_avg = silhouette_score(data_scaled, label_cluster)
61 print(f"Silhouette Score: {silhouette_avg}")
```

PROBLEMS OUTPUT TERMINAL PORTS DEBUG CONSOLE

```
~/Documents/Information Tech/python$ python3 kmeans_clustering.py
Silhouette Score: 0.3022879679963386
```

Figure 2 Code Silhouette Score Hierarchical Clustering with Agglomerative Nesting

```

60 from sklearn.metrics import silhouette_score
61 # Calculate silhouette score
62 if len(set(dbSCAN_labels)) > 1:
63     silhouette_avg = silhouette_score(scaled_features, dbSCAN_labels)
64     print(f'Silhouette Score: {silhouette_avg}')
65 else:
66     print('Silhouette Score cannot be calculated with less than 2 clusters.')
67
PROBLEMS OUTPUT TERMINAL PORTS DEBUG CONSOLE
~/Documents/Information Tech/python/dbSCAN$ python
Silhouette Score: 0.39469401474678745
    
```

Figure 3 Code Silhouette Score DBSCAN Methods
Table 1 Comparison of Cluster Formation Methods

| Method | Number of Clusters | Silhouette Score |
|---|--------------------|------------------|
| <i>Clustering K-Means</i> | 3 | 0.5202 |
| <i>Hierarchical Clustering with Agglomerative Nesting</i> | 3 | 0.3022 |
| <i>Density-Based Spatial Clustering of Applications with Noise (DBSCAN)</i> | 3 | 0.394 |

Based on table 1, it can be seen that the best method to be applied related to *stunting* problems in Tangerang Regency is *Clustering K-Means*. The results of the method produced can be considered in the decision-making of the Tangerang Regency Health Office related to *the problem of stunting* (Balbaa & Abdurashidova, n.d.). In this method, 3 *clusters* are formed. The Preventive and Curative recommendations related to child nutrition in Tangerang Regency can be given as follows:

Table 2 Preventive and Curative Recommendations for Hierarchical Clustering with Agglomerative Nesting

| Cluster 0 | Cluster 1 | Cluster 2 |
|--|--|---|
| Preventive Measures: | Preventive Measures: | Preventive Measures: |
| 1 Nutrition Education: <ul style="list-style-type: none"> Educate parents about the importance of balanced nutrition for pre-school children, including adequate intake of protein, vitamins, and minerals. | 1 Breastfeeding: <ul style="list-style-type: none"> Promote exclusive breastfeeding for up to 6 months and continue until age 2 years or more, with nutritious complementary foods. | 1 Growth and Development Monitoring: <ul style="list-style-type: none"> Monitor growth and development regularly to ensure that children grow optimally. |

Clustering of Child Stunting Data in Tangerang Regency Using Comparison of K-Means, Hierarchical Clustering and DBSCAN Methods

| | | | | | |
|-------------------------|--|-------------------------|---|-------------------------|---|
| 2 | Physical Activity: <ul style="list-style-type: none"> Promote age-appropriate physical activities to support the growth and development of children. | 2 | Food Education: <ul style="list-style-type: none"> Educate mothers about providing complementary foods for breast milk that are rich in iron, protein, and vitamins. | 2 | Nutrition Education: <ul style="list-style-type: none"> Educate parents about the importance of providing nutritious and balanced food according to the needs of children aged 2-5 years. |
| 3 | Periodic Inspections: <ul style="list-style-type: none"> Conduct periodic health check-ups to monitor the growth and development of children. | 3 | Healthy Environment: <ul style="list-style-type: none"> Make sure the home environment is clean and free from infections that can interfere with children's growth. | 3 | Environmental Health: <ul style="list-style-type: none"> Make sure your child gets a clean and healthy environment, free from infectious diseases that can affect growth. |
| Curative Action: | | Curative Action: | | Curative Action: | |
| 1 | Nutrition Consultation: <ul style="list-style-type: none"> If the child's weight or height is still below the standard, it is necessary to consult a nutritionist for a more appropriate diet. | 1 | Nutrition Interventions: <ul style="list-style-type: none"> Provide additional foods rich in energy and protein as well as nutritional supplements if needed to improve the child's nutritional status. | 1 | Nutritionist Consultation: <ul style="list-style-type: none"> If your child's growth is still not optimal, consult a nutritionist to develop a better meal plan. |
| 2 | Nutrition Interventions: <ul style="list-style-type: none"> Give nutritional supplements if needed, especially if certain deficiencies are found based on medical examination. | 2 | Medical References: <ul style="list-style-type: none"> Refer to a healthcare center for further examination and medical treatment if any serious health conditions are found. | 2 | Medical Interventions: <ul style="list-style-type: none"> Provide medical intervention if any disease or condition is found that affects the child's growth. |
| 3 | Monitoring: <ul style="list-style-type: none"> Monitor your child's development more intensively, especially if there are nutritional or health issues that require special attention. | 3 | Rehabilitation Programs: <ul style="list-style-type: none"> Involve your child in a nutritional rehabilitation program available at the health center or hospital for growth recovery. | 3 | Psychosocial Support: <ul style="list-style-type: none"> Make sure your child gets good psychosocial support, such as proper play stimulation and positive social interactions. |

Based on table 2, we can see the recommendations for preventive and curative activities that will be applied to each *cluster* that has been formed in the child data table in Tangerang Regency by performing the necessary fillers based on data analysis that has

Clustering of Child Stunting Data in Tangerang Regency Using Comparison of K-Means, Hierarchical Clustering and DBSCAN Methods

Based on the data that can be filtered in the draw, child nutrition data has been collected in 3 *clusters*. It is hoped that with this data, the Tangerang Regency Health Office can take preventive measures and suppress the growth of *stunting* rates, for example filtering by *name* or by *address* on each child.

Conclusion

Research on stunting analysis in children in Tangerang Regency using the k-means clustering method, Hierarchical Clustering with Agglomerative Nesting, and Density-Based Spatial Clustering of Application with Noise (DBSCAN) showed that k-means produced 3 clusters with cluster 0 aged 46-55 months, the majority of which were well nourished and not stunted; cluster 1 is 9-18 months old with most of the nutrition is good, but there are proportions that need more attention related to nutrition and stunting risk; Cluster 2 is 27-36 months old with the majority of them being well nourished and not stunted. Hierarchical Clustering produced cluster 0 aged 37-53 months, good nutrition, no stunting; cluster 1 is 23-24 months old, good nutrition, not stunted; cluster 2 aged 0-40 months with a z-score TB/U -5.58 indicates most severe stunting. DBSCAN produces cluster-1 aged 26-52 months, good nutrition, no stunting; cluster 0 aged 39-54 months, good nutrition, not stunted; cluster 1 aged 12-25 months with a low z-score TB/U indicates a high probability of stunting. The results of clustering with Silhouette Score show k-means (0.52) as the best method for stunting analysis. Based on the results of k-means, preventive measures such as providing balanced nutrition, health monitoring, immunization, physical activity, and parental education, as well as curative measures such as nutritionist consultation, supplementation, and follow-up examinations can be carried out. The Tangerang Regency Health Office can utilize the results of this analysis for preventive and curative actions that are right on target.

References

- Agapito, G., Milano, M., & Cannataro, M. (2022). A python clustering analysis protocol of genes expression data sets. *Genes*, *13*(10), 1839.
- Balbaa, M., & Abdurashidova, M. (n.d.). *The Impact of Artificial Intelligence in Decision Making: A Comprehensive*.
- Black, R. E., Victora, C. G., Walker, S. P., Bhutta, Z. A., Christian, P., De Onis, M., Ezzati, M., Grantham-McGregor, S., Katz, J., & Martorell, R. (2013). Maternal and child undernutrition and overweight in low-income and middle-income countries. *The lancet*, *382*(9890), 427–451.
- Fadliana, A. (2015). *Penerapan metode Agglomerative Hierarchical Clustering untuk klasifikasi Kabupaten/Kota di Provinsi Jawa Timur berdasarkan kualitas pelayanan keluarga berencana*. Universitas Islam Negeri Maulana Malik Ibrahim.
- Fahrudin, N. (2018). Penerapan Metode Finite State Machine Pada Game Adventure “Franco”. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *2*(1), 446–453.
- H Witten, I., Frank, E., A Hall, M., & J Pal, C. (2017). *Data mining practical machine learning tools and techniques*. Elsevier Inc.
- Hammond, J. S., Keeney, R. L., & Raiffa, H. (2015). *Smart choices: A practical guide to making better decisions*. Harvard Business Review Press.
- Huljannah, N., & Rochmah, T. N. (2022). Program pencegahan stunting di indonesia: A systematic review. *Media Gizi Indonesia*, *17*(3), 281–292.
- Kusumawati, R. (2008). Kecerdasan Buatan Manusia (Artificial Intelligence); Teknologi Impian Masa Depan. *ULULALBAB Jurnal Studi Islam*, *9*(2), 257–274.
- Mitra, M. (2015). Permasalahan Anak Pendek (Stunting) dan Intervensi untuk Mencegah Terjadinya Stunting (Suatu Kajian Kepustakaan). *Jurnal Kesehatan Komunitas*, *2*(6), 254–261.
- Nadirawati, N., Suryaningsih, C., Bangun, A. V., & Rosa, R. D. Dela. (2023). Prevention and Intervention of Stunting in Indonesia: a Scoping Review. *Lux Mensana: Journal of Scientific Health*, 14–28.
- Prendergast, A. J., & Humphrey, J. H. (2014). The stunting syndrome in developing countries. *Paediatrics and international child health*, *34*(4), 250–265.
- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*. Cambridge university press.
- Syukron, H., Fayyad, M. F., Fauzan, F. J., Ikhsani, Y., & Gurning, U. R. (2022). Perbandingan K-Means K-Medoids dan Fuzzy C-Means untuk Pengelompokan Data Pelanggan dengan Model LRFM: Comparison K-Means K-Medoids and Fuzzy C-Means for Clustering Customer Data with LRFM Model. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *2*(2), 76–83.
- Yudoprakoso, P. W. (2019). Kecerdasan Buatan (Artificial Intelligence) Sebagai Alat Bantu Proses Penyusunan Undang-Undang Dalam Upaya Menghadapi Revolusi Industri 4.0 Di Indonesia. *Simposium Hukum Indonesia*, *1*(1), 450–461.