



K-Means Algorithm for District/City Classification in West Java Based on Stunting Toddler Data

Ade Bani Riyan, Ikhsan Nendi

Politeknik Siber Cerdika Internasional, Indonesia

Email: adebani@polteksci.ac.id, nendi@polteksci.ac.id

Abstract

Health issues, especially related to stunting, remain a major concern in Indonesia. Stunting is a condition where toddlers have a height that is too short for their age, which is a sign of chronic malnutrition. The purpose of this study is to group the number of stunting by district/city in West Java. The dataset was obtained from data on the number of stunting toddlers by district/city in West Java Province from 2014 to 2022. This dataset related to the topic of Health is produced by the health Office which is issued every 1 year. The results showed that there are two main groups in the classification of stunting numbers in West Java. The first group (cluster 0) includes Ciamis Regency, Kuningan Regency, Majalengka Regency, Sumedang Regency, Subang Regency, Purwakarta Regency, Bekasi Regency, West Bandung Regency, Pangandaran Regency, Bogor City, Sukabumi City, Bandung City, Cirebon City, Depok City, Cimahi City, Tasikmalaya City, Banjar City. The second group (cluster 1) consists of Bogor Regency, Sukabumi Regency, Cianjur Regency, Bandung Regency, Garut Regency, Tasikmalaya Regency, Cirebon Regency, Indramayu Regency, Karawang Regency, Bekasi City

keywords: Algoritma K-Means, Data Mining, Stunting

INTRODUCTION

Stunting is a condition in which the growth of children under five years old is stunted due to chronic malnutrition, causing them to have a shorter height than they should for their age (Ruaida, 2018). This nutritional deficiency can begin to occur since the fetus is in the womb and continues in the early days after the baby is born (Aprilia, 2020). However, stunting is only seen after the child reaches the age of two. Stunted toddlers have a body length (PB) or height (TB) that is lower than the standard set by WHO-MGRS (Multicentre Growth Reference Study) 2006 for their age, or according to the definition of the Ministry of Health (MoH), stunting occurs when the z-score of children under five is less than -2SD (stunted) and even less than -3SD (severely stunted) (Kemiskinan, 2017) (Grantz et al., 2018). The age of 24-59 months is considered a critical period in the effort to ensure quality human resources, especially since the first two years of a child's life are a golden period for optimal brain growth and development (Nurhayati, 2020). Therefore, serious attention is required in this period. The problem of stunting, also known as short children, is one of the most serious challenges (Widya et al., 2023), especially in countries with low levels of

K-Means Algorithm for District/City Classification in West Java Based on Stunting Toddler Data

poverty and development. According to a report from the United Nations Children's Emergency Fund (UNICEF), more than half of stunted children, about 56%, live in Asia, while more than a third, or about 37%, come from the African continent. Indonesia itself still faces major challenges related to child nutrition and growth problems (Hanifa & Mon, 2021)(Izani, 2021). UNICEF data shows that around 80% of stunted children are spread across 24 developing countries in Asia and Africa. Indonesia ranks fifth in the list of countries with the highest stunting prevalence, after India, China, Nigeria, and Pakistan. Currently, the prevalence of stunting in children under the age of 5 in the South Asia region reaches around 38%.

Nutrition issues remain a major focus in Indonesia, especially in the context of nutrition in toddlers. The health condition and nutritional status of toddlers are important indicators of overall public health. This is due to the impact caused by cases of malnutrition, undernutrition, stunting (growth delay), and other nutritional problems that are a burden for families, communities, and countries (Unicef, 2012). Some factors that are suspected of causing stunting include the mother's pregnancy history, including the mother's short posture, too close pregnancy distance, too many births, the age of the mother during pregnancy who is too old or too young (under 20 years), and insufficient nutritional intake during pregnancy. In addition, factors such as non-implementation of Early Breastfeeding Initiation (IMD), failure of exclusive breastfeeding, and early weaning process also play a role. Economic and sanitation factors also have a correlation with the incidence of stunting (Pangestuti et al., 2023). The impact of stunting includes cognitive, motor, and verbal development that is not optimal in children, increased morbidity and mortality, body posture that is not optimal in adulthood (shorter than average), and learning capacity and performance that is less than optimal at school (Organization, 2020). Cluster analysis is one of the important methods in the field of Data Mining. Data mining is a process that uses various statistical, mathematical, artificial intelligence, and machine-learning techniques to extract and identify valuable information from various large databases (Sitepu et al., 2011). Cluster analysis in data mining is a method used to group a series of data into groups based on predetermined similarities (Matdoan & Van Delsen, 2020). Among the various cluster analysis methods available, two of them are K-Means and K-Medoids Clustering. Both of these methods are types of partitioning clustering methods that have interrelated algorithms. Such methods tend to be faster than hierarchical methods and more advantageous especially when the number of data objects is very large.

The K-Means clustering algorithm plays an important role in the data mining domain and is relatively simple to implement and execute. The K-Means algorithm is a distance-based clustering method that divides data into clusters, and it works primarily on numerical attributes. In practice, this algorithm is often used because of its relative speed and ability to adapt easily (Sangga, 2018).

METHODS

The dataset was obtained from data on the number of stunting toddlers by district/city in West Java Province from 2014 to 2022. This dataset related to the topic of Health is produced by the Health Office which is issued every 1 year.

Table 1 Research Variables

Variabel	Information
X1	Data Stunting 2015
X2	Data Stunting 2016
X3	Data Stunting 2017
X4	Data Stunting 2018
X5	Data Stunting 2019
X6	Data Stunting 2020
X7	Data Stunting 2021
X8	Data Stunting 2022

The stages of research that have been carried out can be seen in Figure 1 below. The image provides a visual picture of the steps taken in conducting this research (Ahmed et al., 2020).

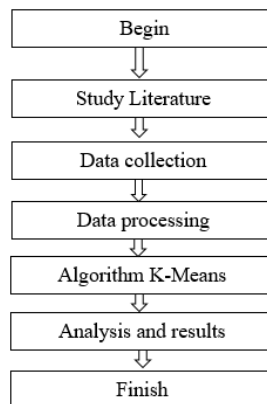


Figure 1. Research Flow

RESULTS AND DISCUSSION

The K-Means algorithm is a non-hierarchical method in data clustering that aims to partition data into one or more clusters or clusters. The main purpose is to group data that has similar characteristics into one cluster, while data that has different characteristics are grouped into different clusters.

K-Means Algorithm for District/City Classification in West Java Based on Stunting Toddler Data

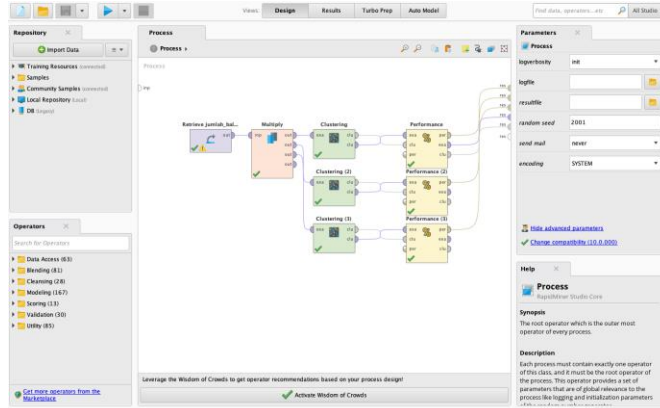


Figure 2. The process of clustering using the K-Means algorithm in RapidMiner Studio.

Based on the results of Data Mining testing using the K-Means Clustering algorithm in the RapidMiner 10 software application, the conclusion that can be drawn is that the centroid value obtained is different due to differences in the amount of data, but the value in the resulting cluster remains the same (Sangga, 2018). Details of the final centroid values can be seen in Figure 2 below:

Attribute	cluster_0	cluster_1
TAHLUN 2015	6984.294	20593.200
TAHLUN 2016	6154.647	23131
TAHLUN 2017	4704.176	18400.100
TAHLUN 2018	4730	14441.500
TAHLUN 2019	4881.176	14345.600
TAHLUN 2020	6080.118	17270.700
TAHLUN 2021	4348.294	13259.300
TAHLUN 2022	3358.412	12294.900

Figure 3. Centroid Value

Cluster Model

Cluster 0: 17 items

Cluster 1: 10 items

Total number of items: 27

Figure 4. K-Means Clusterization

Based on Figure 3, it can be seen that the classification of districts/cities based on the number of stunting in West Java Province is divided into two clusters. Cluster 0 consists of 17 districts/cities, and cluster 1 consists of 10 districts/cities. More detailed information can be found in Table 2 below.

Table 2. Cluster List

Group	Number of Provinces	List of District
Cluster 0	17	Ciamis Regency, Kuningan Regency, Majalengka Regency, Sumedang Regency, Subang Regency, Purwakarta Regency, Bekasi Regency, West Bandung Regency, Pangandaran Regency, Bogor City, Sukabumi City, Bandung City, Cirebon City, Depok City, Cimahi City, Tasikmalaya City, Banjar City
Cluster 1	10	Kabupaten Bogor, Kabupaten Sukabumi, Kabupaten Cianjur, Kabupaten Bandung, Kabupaten Garut, Kabupaten Tasikmalaya, Kabupaten Cirebon, Kabupaten Indramayu, Kabupaten Karawang, Kota Bekasi

CONCLUSION

Based on the findings and analysis in this study, it can be concluded that there are two main groups in the classification of stunting numbers in West Java. The first group (cluster 0) includes Ciamis Regency, Kuningan Regency, Majalengka Regency, Sumedang Regency, Subang Regency, Purwakarta Regency, Bekasi Regency, West Bandung Regency, Pangandaran Regency, Bogor City, Sukabumi City, Bandung City, Cirebon City, Depok City, Cimahi City, Tasikmalaya City, Banjar City. The second group (cluster 1) consists of Bogor Regency, Sukabumi Regency, Cianjur Regency, Bandung Regency, Garut Regency, Tasikmalaya Regency, Cirebon Regency, Indramayu Regency, Karawang Regency, Bekasi City.

REFERENCES

- Ahmed, M., Seraj, R., & Islam, S. M. S. (2020). The k-means algorithm: A comprehensive survey and performance evaluation. *Electronics*, 9(8), 1295.
- Aprilia, W. (2020). Perkembangan pada masa pranatal dan kelahiran. *Yaa Bunayya: Jurnal Pendidikan Anak Usia Dini*, 4(1), 39–56.
- Grantz, K. L., Hediger, M. L., Liu, D., & Louis, G. M. B. (2018). Fetal growth standards: the NICHD fetal growth study approach in context with INTERGROWTH-21st and the World Health Organization Multicentre Growth Reference Study. *American Journal of Obstetrics and Gynecology*, 218(2), S641–S655.
- Hanifa, F., & Mon, N. I. (2021). Hubungan Sanitasi Lingkungan, Berat Lahir dan Panjang Lahir dengan Stunting pada Anak Usia 25-72 Bulan. *Jurnal Ilmiah Kebidanan Indonesia*, 11(03), 163–170.
- Izani, A. R. (2021). United Nations International Children's Emergency Fund (Unicef) Roles In Overcoming Child Trafficking In China. Universitas Muhammadiyah Yogyakarta.

K-Means Algorithm for District/City Classification in West Java Based on Stunting Toddler Data

- Kemiskinan, T. N. P. P. (2017). *Pemilihan 10 Desa Prioritas di 100 Kabupaten/Kota Prioritas Penanganan Kemiskinan dan Stunting*. Jakarta: TNPPK.
- Matdoan, M. Y., & Van Delsen, M. S. N. (2020). Penerapan Analisis Cluster Dengan Metode Hierarki Untuk Klasifikasi Kabupaten/Kota Di Provinsi Maluku Berdasarkan Indikator Indeks Pembangunan Manusia. *Statmat J. Stat. Dan Mat*, 2(2), 20.
- Nurhayati, R. (2020). Pendidikan Anak Usia Dini Menurut Undang–Undang No, 20 Tahun 2003 Dan Sistem Pendidikan Islam. *Al-Afkar, Journal For Islamic Studies*, 57–87.
- Organization, W. H. (2020). *World health statistics 2020*.
- Pangestuti, M., Khomsan, A., & Ekayanti, I. (2023). Determinants of stunting in children aged 6-24 months in rural areas: Case control study. *AcTion: Aceh Nutrition Journal*, 8(3), 318–330.
- Ruaida, N. (2018). Gerakan 1000 hari Pertama Kehidupan Mencegah Terjadinya Stunting (Gizi Pendek) di Indonesia. *Global Health Science*, 3(2), 139–151.
- Sangga, V. A. P. (2018). Perbandingan algoritma K-Means dan algoritma K-Medoids dalam pengelompokan komoditas peternakan di provinsi Jawa Tengah tahun 2015.
- Sitepu, R., Irmeilyana, I., & Gultom, B. (2011). Analisis cluster terhadap tingkat pencemaran udara pada sektor industri di Sumatera Selatan. *Jurnal Penelitian Sains*, 14(3).
- Unicef. (2012). *The state of the world's children 2012: children in an urban world*. Esocialsciences.
- Widya, R., Setyaningrum, S., & Siregar, D. J. S. (2023). *Revitalisasi Gizi Anak: Produk Ternak Yang Asuh Sebagai Solusi Stunting*. PT. Sonpedia Publishing Indonesia.

Copyright holder:

Ade Bani Riyan, Ikhsan Nendi (2024)

First publication right:

Journal Transnational Universal Studies (JTUS)

This article is licensed under:

