

---

**Implementasi Data Mining Untuk Klustering Stunting Gizi Pada Balita  
Dipuskesmas Sigambal Menggunakan Metode K-Medoids Dan K-Means**

Melisa<sup>1</sup>, Syaiful Zuhri Harahap<sup>2</sup>, Masrizal<sup>3</sup>

Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Labuhanbatu<sup>1,2,3</sup>

Email : [meliaja260@gmail.com](mailto:meliaja260@gmail.com)<sup>1</sup>, [syaifulzuhriharahap@gmail.com](mailto:syaifulzuhriharahap@gmail.com)<sup>2</sup>,  
[masrizal120405@gmail.com](mailto:masrizal120405@gmail.com)<sup>3</sup>

Corresponding Author : [meliaja260@gmail.com](mailto:meliaja260@gmail.com)

**Abstract**

*The aim of this study was to identify and understand the different characteristics of toddlers in the context of factors that contribute to nutritional stunting. By using the clustering method, this study aims to group toddlers into several groups based on the similarity of their characteristics, so that more targeted interventions can be designed in dealing with stunting problems. Through this approach, it is hoped that significant patterns and risk factors can be found that distinguish stunted toddlers from toddlers who grow normally, and provide insights that can be used by policy makers and health practitioners to improve the quality of life of children. The method used in this study involves the application of two clustering techniques, namely K-Means and K-Medoids to Group sample data of 116 toddlers. The clustering process is carried out by measuring the distance between the toddler data and the centroid or medoid to determine which group is most suitable. The Data were analyzed to find patterns identifying unique characteristics of each cluster, reflecting differences in nutritional stunting-related risk factors. This process helps in differentiating groups of toddlers who are prone to stunting from those who are not, so that the analysis can be focused on the groups most in need of intervention. The results of clustering analysis showed that as many as 48 toddlers entered the C1 cluster, while the other 68 toddlers entered the C2 cluster. Each cluster describes two groups of toddlers with different characteristics in the context of nutritional stunting risk factors. The findings provide deep insight into the significant differences between the two groups, allowing researchers to identify specific patterns and risk factors. This information is then used to design more specific and effective interventions in addressing nutritional stunting in toddlers, taking into account the unique characteristics of each cluster that has been identified.*

**Keywords:** Metode K-Means, Metode K-Medoids, Clustering.

**I. Pendahuluan**

Balita, atau anak usia bawah lima tahun, adalah periode penting dalam perkembangan seorang anak. Pada tahap ini, balita mengalami pertumbuhan fisik yang cepat, seperti

peningkatan tinggi badan dan berat badan, serta perkembangan motorik kasar dan halus. Selain itu, perkembangan kognitif dan bahasa juga signifikan. Balita mulai mengembangkan keterampilan dasar

seperti berjalan, berbicara, dan berinteraksi sosial. Nutrisi yang tepat dan stimulasi lingkungan yang mendukung sangat penting untuk memastikan tumbuh kembang yang optimal pada masa ini. Selain pertumbuhan fisik dan kognitif, perkembangan emosional dan sosial juga menjadi fokus utama selama masa balita. Anak-anak mulai menunjukkan rasa ingin tahu yang tinggi dan keinginan untuk mengeksplorasi dunia sekitar mereka. Interaksi dengan orang tua, keluarga, dan teman sebaya membantu mereka belajar tentang hubungan sosial dan membangun rasa percaya diri. Pada masa ini, penting bagi orang tua dan pengasuh untuk memberikan perhatian, kasih sayang, dan bimbingan yang konsisten, sehingga balita dapat mengembangkan keterampilan yang diperlukan untuk menjadi individu yang sehat dan bahagia. Stunting adalah kondisi gagal tumbuh pada balita akibat kekurangan gizi kronis, terutama selama 1.000 hari pertama kehidupan, yang dimulai sejak konsepsi hingga anak berusia dua tahun. Stunting gizi pada balita adalah salah satu masalah kesehatan masyarakat yang serius, terutama di negara berkembang. Stunting terjadi akibat kekurangan gizi kronis dalam jangka waktu yang panjang, terutama selama 1.000 hari pertama kehidupan, yang mencakup masa kehamilan hingga anak berusia dua tahun. Imunitas, dan kesehatan secara keseluruhan. Balita yang mengalami stunting cenderung memiliki sistem kekebalan tubuh yang lebih lemah, membuat mereka lebih rentan terhadap infeksi dan penyakit, serta berpotensi mengalami masalah kesehatan di kemudian hari. Masalah stunting pada balita sangat kompleks

dan dipengaruhi oleh berbagai faktor, termasuk asupan gizi yang tidak mencukupi, pola makan yang buruk, dan infeksi berulang. Selain itu, kondisi lingkungan yang tidak higienis, kurangnya akses terhadap layanan kesehatan, serta praktik pengasuhan yang tidak memadai juga berkontribusi terhadap tingginya angka stunting. Masalah stunting pada balita merupakan isu kesehatan yang mendesak dan memerlukan perhatian khusus. Kekurangan gizi kronis yang menyebabkan stunting tidak hanya berdampak pada pertumbuhan fisik, tetapi juga perkembangan kognitif dan kesehatan jangka panjang anak. Untuk memahami lebih dalam tentang prevalensi dan faktor-faktor yang berkontribusi terhadap stunting, penulis ingin melakukan sebuah penelitian yang fokus pada analisis stunting pada balita. Penelitian ini bertujuan untuk mengidentifikasi kelompok-kelompok balita yang mengalami stunting berdasarkan berbagai variabel seperti asupan gizi, kondisi kesehatan, dan lingkungan. Dalam penelitian ini, data balita yang mengalami stunting akan dikumpulkan dan dianalisis menggunakan metode clustering untuk mengelompokkan mereka ke dalam kategori tertentu. Hasil dari penelitian ini diharapkan dapat memberikan wawasan yang lebih mendalam tentang stunting pada balita dan menjadi dasar bagi pengambilan keputusan dalam program-program pencegahan dan penanggulangan stunting di masyarakat.

Dalam penelitian ini, penulis akan melakukan pengelompokan data stunting pada balita menggunakan dua metode model clustering yaitu metode K-means dan K-medoids. Metode K-means adalah salah satu teknik

*clustering* yang populer, di mana data dibagi ke dalam sejumlah cluster berdasarkan jarak Euclidean terdekat dari centroid masing-masing cluster. K-means bekerja dengan menginisiasi centroid secara acak dan kemudian memperbarui posisi centroid iteratif hingga konvergensi tercapai. K-medoids menggunakan medoid, yaitu titik data nyata yang berada di dalam cluster dan memiliki jarak total terkecil terhadap titik-titik lain dalam cluster tersebut. Pendekatan ini membuat K-medoids lebih robust terhadap data yang memiliki outlier atau distribusi yang tidak normal. Dengan membandingkan hasil dari kedua metode ini, penulis berharap dapat menemukan pola dan hubungan yang lebih mendalam dalam data stunting balita, sehingga memungkinkan identifikasi kelompok-kelompok risiko dengan lebih akurat dan penyusunan strategi intervensi yang lebih tepat sasaran.

## II. Landasann Teori

### Metode K-Medoids

Metode K-Medoids merupakan salah satu teknik clustering dalam data mining yang digunakan untuk mengelompokkan data menjadi beberapa kelompok atau klaster berdasarkan kesamaan antara data dalam klaster yang sama. K-Medoids merupakan variasi dari algoritma K-Means, di mana K-Medoids lebih robust terhadap outliers atau data yang memiliki nilai ekstrem. Proses K-Medoids dimulai dengan pemilihan titik medoid awal secara acak di dalam dataset. Data mining biasanya digunakan untuk proses pengumpulan data yang besar dari sekumpulan fakta yang terekam untuk mengetahui pola yang ada untuk menghasilkan sebuah

keluaran yang berupa informasi atau pengetahuan. Metode-metode yang terdapat pada Data mining untuk Clustering salah satunya yaitu metode K-Medoids. Data Mining terdapat 2 metode yang dapat digunakan yaitu metode K-Medoids dan metode K-Means.

K-Means merupakan metode penganalisaan data pada Data mining dimana proses pemodelan tanpa supervisi dan merupakan salah satu metode yang mengelompokkan data secara partisi. Pada metode K-Means data dikelompokkan menjadi beberapa kelompok dimana setiap kelompok mempunyai karakteristik yang mirip atau sama dengan lainnya namun dengan kelompok lainnya memiliki karakteristik yang berbeda. Metode ini meminimalisasi perbedaan antara data di dalam satu Cluster seta memaksimalkan perbedaan dengan Cluster yang lain.

## III. Metode Penelitian

Dalam upaya untuk mengelompokkan data stunting gizi pada balita, metode K-Means dan K-Medoids digunakan sebagai teknik clustering yang efektif. Tahapan penelitian ini mengikuti proses Knowledge Discovery in Database (KDD), yang dimulai dengan pengumpulan data dari sumber-sumber yang relevan. Setelah data dikumpulkan, dilakukan tahap pra-pemrosesan untuk membersihkan dan mengubah data agar siap digunakan dalam analisis. Selanjutnya, data dipilih dan ditransformasikan ke dalam format yang sesuai untuk diolah oleh algoritma clustering. Dengan menggunakan kedua metode ini, penelitian dapat mengevaluasi dan membandingkan

hasil clustering, memastikan kelompok balita yang mengalami stunting diidentifikasi secara akurat. Hasil akhir dari pengelompokan ini akan memberikan wawasan yang mendalam tentang faktor-faktor yang berkontribusi terhadap stunting gizi pada balita, yang dapat digunakan untuk merancang intervensi yang lebih tepat dan efektif dalam penanganan masalah stunting di masyarakat.

#### IV. Hasil dan Pembahasan *Selection*

Untuk data yang digunakan pada penelitian ini yaitu data balita yang diperoleh dari Puskesmas Sigambal, Kecamatan Rantau Selatan, kabupaten Labuhanbatu. Data yang diperoleh sebanyak 116 data balita. Data ini yang nantinya akan digunakan pada Penelitian ini. Data sampel ini yang nantinya akan diolah pada data mining dengan menggunakan metode K-Means dan metode K-Medoids.

**Tabel 1. Data Balita**

Nama Balita	Bulan	Berat Badan	Tinggi Badan
Agus Pratama Santoso	53	13.5	93.4
Ahmad Fauzi Rambe	47	12	89.9
Aisyah Putri Siregar	50	13.7	92.8
Aldi Rian Nasution	13	12	87.9
Angga Pratama Setiawan	36	9.8	84
Arga Pratama Wijaya	39	11.5	91.5
Bella Salsabila Aulia	10	9.8	78.2
Beri Prasetyo Adi	0	9.5	81.4
Bima Aditya Putra	13	11.2	81.8
Bintang Aditya Siregar	10	12.3	90.1
Budi Santoso Rambe	43	12	89
Bunga Melati Nasution	12	11.8	92.1
Calhyani Puspita Rambe	11	12.1	90.1
Chandra Prasetyo Nugroho	20	7.8	74
Cindy Larasati Hasibuan	27	10.7	84.1
Cinta Anggraini Utami	46	12	89.9
Citra Anggraini Dewi	49	13.7	92.8
Citra Rahayu Siregar	44	12	93.7
Dani Ramadhan Pratama	45	13.7	92.8
Daniel Pratama Siregar	31	10.7	88

Untuk tabel diatas merupakan data sampel yang digunakan pada Penelitian ini. Data yang digunakan sdebanyak 116 data sampel, tetapi untuk data yang ada pada table diatas hanya penulis tampilkan sebanyak 20 data sampel balita. Untuk data sampel nya semua akan ditampilkan nanti pada tahapan hasil dari pengelompokkan atau clustering data.

#### *Preprocessing*

Pada tahapan ini dilakukan untuk menyeleksi data yang sudah diperoleh. Hal ini dilakukan karna tidak semua data dapat digunakan pada penelitian ini. Jadi tahap preprocessing adalah tahapan yang dilakukan untuk memilih dan menentukan data yang layak untuk digunakan. Ketentuan data nya yaitu berdasarkan dengan kebutuhan data pada Penelitian ini.

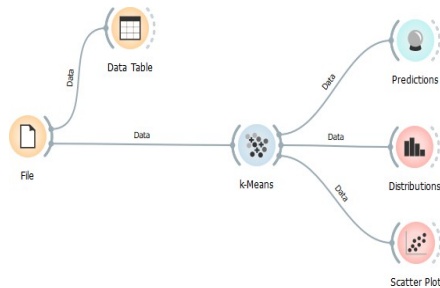
#### *Transformation*

Pada tahapan transformasi adalah tahapan untuk merubah bentuk dan format dari data yang digunakan. Jadi untuk data yang dibutuhkan pada Penelitian ini adalah data yang disusun dengan format Sampel.xlsx. Maka untuk data yang sudah diperoleh akan disusun dalam bentuk dan format sampel.xlsx.

#### **DataMining**

Tahapan data mining merupakan tahapan yang dilakukan untuk mengelompokkan data sampel menggunakan metode K-means dan metode K-Medoids. Tetapi sebelum pengelompokkan, perlu dirancang sebuah model clustering yang

nantinya dapat digunakan untuk melakukan pengelompokan data balita.



**Gambar 1. Model Perancangan Cluster**

Untuk gambar diatas merupakan model clustering yang dirancang untuk dapat digunakan melakukan clustering data balita. Untuk hasil clustering data nantinya yaitu terdapat C1 dan C2. Untuk hasil nya diperoleh dengan menggunakan metode K-Means dan metode K-Medoids.

### Hasil Clustering

Hasil clustering dari data sampel sebanyak 116 balita menunjukkan bahwa sebanyak 48 balita masuk ke dalam cluster C1, sementara sisanya, yaitu 68 balita, masuk ke dalam cluster C2. Cluster C1 dan C2 ini menggambarkan dua kelompok balita dengan karakteristik berbeda berdasarkan faktor-faktor yang berkontribusi terhadap stunting gizi. K-Means dan K-Medoids digunakan untuk mengelompokkan data berdasarkan jarak terdekat dari centroid atau medoid, yang membantu mengidentifikasi pola-pola dan faktor risiko yang berbeda antara kedua cluster tersebut. Analisis ini memberikan wawasan mendalam tentang perbedaan signifikan antara kedua kelompok, memungkinkan peneliti untuk merancang intervensi yang lebih spesifik dan efektif dalam menangani stunting gizi pada balita,

dengan mempertimbangkan karakteristik unik dari setiap cluster.

## V. Kesimpulan dan Saran

### Kesimpulan

Kesimpulan dari analisis clustering yang dilakukan pada 116 sampel balita menunjukkan bahwa terdapat dua kelompok utama, yaitu Cluster C1 yang terdiri dari 48 balita dan Cluster C2 yang terdiri dari 68 balita. Penggunaan metode K-Means dan K-Medoids berhasil mengelompokkan balita berdasarkan faktor-faktor yang berkontribusi terhadap stunting gizi, dengan memperhatikan jarak terdekat dari centroid atau medoid. Analisis ini mengungkap perbedaan signifikan antara kedua cluster, memberikan pemahaman mendalam mengenai pola dan faktor risiko yang mempengaruhi stunting gizi. Dengan demikian, hasil ini dapat menjadi dasar untuk merancang intervensi yang lebih tepat dan efektif, sesuai dengan karakteristik unik dari masing-masing kelompok, guna mengatasi permasalahan stunting gizi pada balita secara lebih optimal.

### Saran

Saran dari analisis ini adalah agar intervensi pencegahan stunting gizi difokuskan berdasarkan karakteristik unik dari masing-masing cluster. Untuk balita dalam Cluster C1, yang menunjukkan faktor risiko lebih tinggi, diperlukan intervensi yang lebih intensif dan cepat. Sedangkan untuk balita dalam Cluster C2, yang cenderung memiliki faktor risiko lebih rendah, program pemantauan dan

pengecahan yang berkelanjutan dapat diterapkan. Pendekatan berbasis cluster ini akan membantu mengoptimalkan alokasi sumber daya dan meningkatkan efektivitas program penanganan stunting gizi pada balita.

## VI. Daftar Pustaka

- A. Kurnia, "Perbandingan Algoritma K-Means dan Fuzzy C-Means Untuk Clustering Puskesmas Berdasarkan Gizi Balita Surabaya," *J. Process.*, vol. 18, no. 1, pp. 83–88, 2023, doi: 10.33398/processor.2023.18.1.696.
- S. A. Sakti, "Pengaruh Stunting pada Tumbuh Kembang Anak Periode Golden Age," *J. Ilm.Fak.Kegur.danIlmuPendidik*, vol. 6, no. 1, pp. 169–175, 2020, [Online]. Available: <http://ejournal.unsub.ac.id/index.php/FKIP>
- H. A. Karim, F. Fitritanti, and Y. Yakub, "Peningkatan Produktifitas Tanaman Sawi Melalui Penambahan Pupuk Kandang Ayam dan NPK 16:16:16," *JAMI J. Ahli Muda Indones.*, vol. 1, no. 1, pp. 65–72, 2020, doi: 10.46510/jami.v1i1.19.
- K. Kondisi Gizi Balita Dan Anak Khairul Zannah, Z. Masruro Nasution, I. Parlina, I. Purnama Sari, J. Teknik Informatika, and S. Tunas Bangsa, "Model Clustering Algoritma K-Mean Dalam Menentukan," *J. Din. Inform.*, vol. 11, no. 1, pp. 34–44, 2022.
- I. C. Indah, M. N. Sari, and M. H. Dar, "Application of the K-Means Clustering Algorithm to Group Train Passengers in Labuhanbatu," *Sinkron*, vol. 8, no. 2, pp. 825–837, 2023, doi: 10.33395/sinkron.v8i2.12260.
- R. N. Juliadi and Y. Puspitarani, "Supervised Model for Sentiment Analysis Based on Hotel Review Clusters using Rapid Miner," *Sinkron*, vol. 7, no. 3, pp. 1059–1066, 2022, doi: 10.33395/sinkron.v7i3.11564.
- F. Paquin, J. Rivnay, A. Salleo, N. Stingelin, and C. Silva, "Multi-phase semicrystalline microstructures drive exciton dissociation in neat plastic semiconductors," *J. Mater. Chem. C*, vol. 3, pp. 10715–10722, 2015, doi: 10.1039/b000000x.
- A. Ali, "Clustering Data Antropometri Balita Untuk Menentukan Status Gizi Balita Di Kelurahan Jumpat Rejo Sukodono Sidoarjo," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 7, no. 3, pp. 395–407, 2020, doi: 10.35957/jatisi.v7i3.530.
- M. Shutaywi and N. N. Kachouie, "Silhouette analysis for performance evaluation in machine learning with application to clustering," *Entropy*, vol. 23, no. 6, pp. 1–17, 2021, doi: 10.3390/e23060759.
- Y. S. Su and S. Y. Wu, "Applying data mining techniques to explore user behaviors and

- watching video patterns in converged IT environments,” *J. Ambient Intell. Humaniz. Comput.*, no. 0123456789, 2021, doi: 10.1007/s12652-020-02712-6.
- R. W. Sembiring Brahmama, F. A. Mohammed, and K. Chairuang, “Customer Segmentation Based on RFM Model Using K-Means, K-Medoids, and DBSCAN Methods,” *Lontar Komput.J. Ilm. Teknol. Inf.*, vol. 11, no. 1, p. 32, 2020, doi: 10.24843/lkjiti.2020.v11.i01.p04.
- P. Cv and W. Menggunakan, “Implementasi Algoritma K-Means Dalam Pengelompokan Data,” vol. 2, no. 1, pp. 188–196, 2023.
- E.Muningsih, “Kombinasi Metode k-Means Dan Decision Tree Dengan Perbandingan Kriteria Dan Split Data,” vol. 16, no. 1, pp. 113–118, 2022.
- W. Sudrajat, I. Cholid, and J. Petrus, “Wahyu Sudrajat et al, Penerapan Algoritma K-Means Untuk ,” p. 27, 2022.