

## Pemanfaatan Data Mining Dengan Algoritma *K-Means* Untuk Analisis Penyebaran Stunting di Kota Samarinda

(*Utilization of Data Mining with the K-Means Algorithm to Analyze the Distribution of Stunting in Samarinda City*)

**Dody Novandi**

Badan Pendapatan Daerah Kota Samarinda  
Komplek Kantor Walikota Samarinda  
Email: [dodysamarinda@gmail.com](mailto:dodysamarinda@gmail.com)

*Diterima: 17 April 2024; Direvisi: 13 Juni 2024; Disetujui: 26 Juni 2024*

### **ABSTRACT**

*The reason this research was conducted is because accelerating stunting reduction is the most crucial program because it is related to the formation of Indonesian people, especially the next generation in Samarinda City, who will be increasingly competitive in the next period of time. This research aims to provide knowledge about the distribution of stunting cases so that the Samarinda City Government can make appropriate policies to optimize the reduction in stunting prevalence rates in Samarinda City. The type of research in this research is quantitative research, with quantitative descriptive methods and the analytical tool used is the Cross-Industry Standard Process Model for Data Mining or CRISP-DM. With the implementation of the K-Means algorithm, the results obtained in this study were the prevalence value in cluster 0 was 4.50%, cluster 1 was 16%, cluster 2 was 6.94%, cluster 3 was 10.89%, cluster 4 of 19%, cluster 5 of 13.93%, cluster 6 of 18.67%, cluster 7 of 20.51%, cluster 8 of 5.27%, and cluster 9 of 12.89%. Sempaja Utara sub-district, which is a member of cluster 7, has the highest prevalence of stunting and is also an area that has the first extreme poverty score.*

*Keywords: Stunting, Data Mining, CRISP-DM, K-Means*

### **ABSTRAK**

Penelitian ini dilakukan karena akselerasi pengurangan stunting merupakan program yang paling krusial karena berkaitan dengan pembentukan manusia Indonesia khususnya generasi penerus di Kota Samarinda yang kian berdaya saing di periode waktu selanjutnya. Penelitian ini bertujuan memberikan pengetahuan terhadap pengelompokan penyebaran kasus stunting sehingga Pemerintah Kota Samarinda dapat membuat kebijakan yang tepat guna optimalisasi penurunan angka prevalensi stunting di Kota Samarinda. Jenis penelitian dalam penelitian ini yaitu penelitian kuantitatif, dengan metode deskriptif kuantitatif dan alat analisis digunakan adalah *Cross-Industry Standard Process Model for Data Mining* atau *CRISP-DM*. Dengan adanya implementasi algoritma *K-Means*, hasil yang didapatkan dalam penelitian ini adalah nilai prevalensi pada *cluster 0* sebesar 4,50%, *cluster 1* sebesar 16%, *cluster 2* sebesar 6,94%, *cluster 3* sebesar 10,89%, *cluster 4* sebesar 19%, *cluster 5* sebesar 13,93%, *cluster 6* sebesar 18,67%, *cluster 7* sebesar 20,51%, *cluster 8* sebesar 5,27%, dan *cluster 9* sebesar 12,89%. Kelurahan Sempaja Utara yang masuk dalam anggota *cluster 7* memiliki prevalensi stunting paling tinggi dan juga merupakan wilayah yang memiliki nilai kemiskinan ekstrem dengan urutan pertama.

**Kata kunci:** Stunting, Data Mining, *CRISP-DM*, *K-Means*

### **PENDAHULUAN**

Percepatan pengurangan stunting terhadap anak dengan kategori Balita merupakan program prioritas Pemerintah sebagaimana tercantum pada RPJMN 2020-2024 (<https://stunting.go.id>). Akselerasi pengurangan stunting merupakan program yang paling

krusial karena berkaitan pembentukan manusia Indonesia khususnya generasi penerus di Kota Samarinda yang kian berdaya saing di periode waktu selanjutnya. Dengan program yang sudah ditentukan, kejadian stunting menggambarkan salah satu permasalahan secara domestik yang mendapatkan perhatian besar dalam penanggulangannya.

Kasus stunting yang terjadi pada anak menggambarkan kondisi adanya kegagalan pertumbuhan anak dengan usia kurang dari 5 (lima) tahun karena ketidakcukupan gizi akut, sehingga menyebabkan anak dengan situasi terlalu pendek untuk usia pada saat itu. Keadaan ini dapat terjadi saat seorang ibu sedang menjalani proses kehamilan serta sampai proses pertumbuhan bayi dengan usia 2 (dua) tahun (Tasman, dkk, 2020).

Beberapa faktor pemicu balita masuk dalam kategori stunting adalah: pemberian makanan bergizi yang didapatkan bagi ibu hamil sampai menjalani proses persalinan sangat kurang, kurang maksimalnya jangkauan kepada sarana dan prasarana kesehatan, cakupan air bersih dan sanitasi sangat minim (Laili & Andriani, 2019). Berdasarkan kajian yang tertuang dalam dokumen Strategi Nasional Percepatan Pencegahan Anak Kerdil (Stunting) ada beberapa faktor secara langsung dan tidak langsung balita mengalami stunting meliputi penerimaan pendapatan dan ketidaksetaraan ekonomi, perdagangan, urbanisasi, globalisasi, sistem pangan, jaminan sosial, sistem kesehatan, pembangunan pertanian, dan pemberdayaan perempuan (Dokumen Strategi Nasional Percepatan Pencegahan Anak Kerdil (Stunting) Periode-2018-2024, 2019)

Merujuk pada informasi data sebaran stunting yang disajikan melalui portal *website* <https://aksi.bangda.kemendagri.go.id> yang dikelola oleh Ditjen Bina Pembangunan Daerah - Kementerian Dalam Negeri, bahwa pada tahun 2023 kasus kejadian stunting yang dialami anak-anak yang pada 38 Provinsi di Indonesia sebanyak 1.172.051 kasus dengan rincian 884.003 kasus dengan kategori stunting pendek dan 288.048 kasus kategori stunting sangat pendek dengan nilai prevalensinya sebesar 7.1%. Dari sudut pandang wilayah Provinsi Kalimantan Timur, bahwa kasus stunting pada tahun 2023 sebanyak 13.305 anak dengan kategori stunting pendek dan 4.422 anak untuk kategori stunting sangat pendek dengan nilai prevalensinya sebesar 10.2%. Untuk Kota Samarinda pada tahun 2023 kejadian kasus stunting sebanyak 1.430 anak menderita dengan kategori stunting pendek dan 566 anak menderita dengan kategori stunting sangat pendek dengan nilai prevalensinya sebesar 9.7%.

Persoalan kasus stunting berimplikasi terhadap pertumbuhan trend daya cipta generasi penerus pada sebuah wilayah termasuk di Kota Samarinda. Anak-anak dengan kondisi prima atau sehat, berpengetahuan serta produktif menjadi aset sangat berharga karena menggambarkan sebagai generasi penerus guna menyokong kemajuan pembangunan di Kota Samarinda. Apabila di Kota Samarinda anak-anak dilahirkan dan tumbuh dalam fase ketiadaan gizi akut, maka dapat mengakibatkan para generasi penerus tersebut akan menderita stunting.

Pemerintah Kota Samarinda perlu melakukan pemetaan faktor pemicu paling dominan yang mendorong tumbuhnya kasus stunting pada setiap kelurahan di Kota Samarinda dengan memperhatikan parameter-parameter penyebab stunting yang sudah ditetapkan dalam dokumen Strategi Nasional Percepatan Pencegahan Anak Kerdil (Stunting). Guna mendukung perumusan kegiatan penanganan stunting yang tepat sasaran perlunya melakukan analisis *cluster* kasus stunting dengan memanfaatkan data untuk mendapatkan pola-pola, sehingga stakeholder mampu mendesain kebijakan dan program berdasarkan wawasan yang dihasilkan.

Seiring dengan adanya perkembangan ilmu pengetahuan dan *disrupsi* teknologi informasi telah melahirkan beberapa ilmu pengetahuan yang berhubungan dengan analisis data, salah satunya adalah Data Mining. Kegiatan penambangan data merupakan mekanisme dalam menemukan pola-pola yang unik dengan menggunakan beberapa *tools* dan algoritma. Tujuan

akhir dari proses ini adalah untuk menemukan pengetahuan dari data. Dengan memanfaatkan teknologi dan algoritma yang tepat maka akan didapatkan wawasan yang dicapai.

Data mining bisa mendukung instansi pemerintah guna melakukan proses eksplorasi data yang potensial serta analisis data dan dapat membantu dalam pengambilan kebijakan yang tepat sasaran dengan merujuk pada pengetahuan yang didapatkan. Ada berbagai teknik yang dapat diimplementasikan pada data mining yaitu salah satunya adalah *clustering*. Guna mendukung *clustering* pada data mining, ada salah satu algoritma yang dapat digunakan yaitu *K-Means*.

Implementasi algoritma *K-Means* bertujuan guna mengelompokan objek yang memiliki tingkat kemiripan maksimum dalam suatu area tertentu. Secara universal, algoritma *K-Means* memiliki beberapa keunggulan diantaranya mudah diterapkan, durasi pembelajaran lebih efisiensi dan mudah untuk diadaptasi (Wahyono, 2018) serta merupakan sebuah algoritma clustering yang banyak penerapannya pada berbagai bidang (Heryadi, Wahyono, 2020).

Dengan melihat potret kasus stunting di Kota Samarinda pada tahun 2023 serta penjelasan permasalahan diatas tersebut, maka peneliti terdorong untuk membuat sebuah jurnal penelitian yang berjudul "Pemanfaatan Data Mining Dengan Algoritma *K-Means* Untuk Analisis Penyebaran Stunting di Kota Samarinda". Penelitian ini bertujuan memberikan pengetahuan terhadap pengelompokan penyebaran kasus stunting sehingga Pemerintah Kota Samarinda dapat membuat kebijakan yang tepat guna optimalisasi penurunan angka prevalensi stunting di Kota Samarinda.

Jenis penelitian yang dilakukan merupakan penelitian kuantitatif. Metode yang digunakan adalah deskriptif kuantitatif dan alat analisisnya yaitu *Cross-Industry Standard Process Model for Data Mining* atau *CRISP-DM* dengan memanfaatkan aplikasi Rapid Miner sebagai tools untuk melakukan proses data mining.

## TINJAUAN PUSTAKA

### **Stunting**

Stunting merupakan gangguan pertumbuhan yang dialami oleh balita yang mengakibatkan keterlambatan pertumbuhan anak yang tidak sesuai dengan standarnya sehingga mengakibatkan dampak baik jangka pendek maupun jangka panjang (Apriyani, Dikananda, Ali, 2023). Stunting adalah sebuah situasi bahwa anak dengan kategori balita mengalami perkembangan sangat tidak relevan (pendek atau kerdil), yang disebabkan adanya kemunduran gizi kronis, dimana kondisi ini terjadi karena buruknya volume dan mutu gizi saat diterima balita dalam jeda periode yang agak panjang (Laili & Andriani, 2019).

### **Data Mining**

Data Mining merupakan kegiatan penambangan pola yang bermanfaat terhadap jumlah data yang sangat luas dikenal sebagai data mining (Apriyani, Dikananda, Ali, 2023). Menurut Fayyed et all (1996) dalam Suyanto (2019), Data Mining adalah langkah analisi terhadap proses penemuan pengetahuan di dalam basis data atau *knowledge discovery in database* atau disebut dengan istilah KDD.

### **Unsupervised Learning**

Sebuah metode yang digunakan oleh sebuah sistem yang akan mempelajari unlabeled data berdasarkan karakteristik- karakteristik dari kumpulan data (Wahyono, 2018).

### Clustering

*Clustering* merupakan sebuah algoritma yang bisa melakukan pengelompokan data-data ke dalam sejumlah kumpulan (*cluster*) menurut kesamaan karakter masing-masing data pada kelompok-kelompok yang ada (Fitriyani, 2021). Sedangkan menurut Darmansyah dan Wardani (2021), *clustering* adalah suatu instrumen yang akan membantu kegiatan data mining untuk mengelompokkan objek-objek kedalam *cluster-cluster*.

### K-Means

*K-Means* merupakan sebuah teknik dalam penambangan data yang berfungsi guna pengelompokan data yang tidak mempunyai kedudukan atau hanya berupa partisi (Nurohmah, Mayasari, Sari, 2023). Algoritma *K-Means* adalah sebuah metode yang dapat mengelompokan data-data ke dalam kelompok, dimana kelompok tersebut dibentuk dengan meminimalkan jumlah dari *euclidean distances* diantara data dengan titik pusat (*centroid*) yang menggambarkan *cluster* tersebut (Wahyono, 2018)

### Cross-Industry Standard Process Model for Data Mining (CRISP-DM)

Menurut Larose (2014) dalam Feblian dan Daihani (2017), *CRISP-DM* adalah pedoman baku yang digunakan proses data mining sebagai kebijakan penyelesaian masalah secara universal dari bisnis atau bagian penelitian.

### Davies-bouldin index (DBI)

*DBI* merupakan salah satu prosedur penilaian internal yang menghitung evaluasi cluster pada suatu cara pengelompokan yang berdasarkan terhadap skor kohesi dan separasi (Nabila, dkk, 2021), dengan nilai *DBI* yang dihasilkan merupakan nilai paling kecil (non negatif  $\geq 0$ ), maka cluster yang dihasilkan merupakan pengelompokan yang paling baik diimplementasikan dengan algoritma *K-Means* (Apriyani, Dikananda, Ali, 2023).

### Prevalensi Stunting

Prevalensi stunting merujuk nilai persentase kuantitas balita pada suatu wilayah yang terjadi kasus *stunting* dalam perkembangan fisik balita (Fitriani, 2023). Formula yang digunakan untuk menghitung angka prevalensi poin adalah : jumlah penderita lama dan baru dalam satu waktu ÷ jumlah penduduk x 100% (Yuri, 2023).

### METODE

Bentuk data yang digunakan dalam oleh peneliti merupakan data sekunder yang bersumber dari portal *website* <https://aksi.bangda.kemendagri.go.id> tentang Dashboard Sebaran Stunting yang di kelola oleh Ditjen Bina Pembangunan Daerah - Kementerian Dalam Negeri. Pengambilan data sekunder yang digunakan dalam oleh peneliti dilakukan pada bulan Maret 2024. Data sekunder berupa potret sebaran kasus stunting di setiap kelurahan dengan kategori pendek dan sangat pendek beserta nilai prevalensinya di Kota Samarinda tahun 2023 dapat dilihat pada *display* data pada tabel 1.

Tabel 1.  
Tabel Sebaran Kasus Stunting di Kota Samarinda Tahun 2023

No	Nama Kelurahan	Jumlah Balita		Stunting		Prevalensi %
		(Anak)	Pendek	Sangat Pendek		
1	Simpang Pasir	144	25	16		28.5

2	Handil Bakti	212	62	21	39.2
3	Bantuas	261	28	4	12.3
4	Bukuan	145	7	3	6.9
5	Rawa Makmur	382	30	9	10.2
6	Selili	439	8	2	2.3
7	Sungai Dama	236	3	5	3.4
8	Sidodamai	349	4	2	1.7
9	Sidomulyo	575	9	1	1.7
10	Pelita	363	8	2	2.8
11	Bugis	233	7	3	4.3
12	Pasar Pagi	63	4	0	6.3
13	Pelabuhan	118	2	2	3.4
14	Sungai Pinang Luar	276	16	7	8.3
15	Karang Mumus	97	2	0	2.1
16	Pulau Atas	120	9	6	12.5
17	Sindang Sari	190	12	2	7.4
18	Makroman	460	24	9	7.2
19	Sambutan	1,311	33	58	6.9
20	Sungai Kapih	283	13	7	7.1
21	Mesjid	188	29	19	25.5
22	Baqa	56	9	3	21.4
23	Sungai Keledang	385	65	15	20.8
24	Mangkupalas	139	19	11	21.6
25	Tenun	44	4	10	31.8
26	Gunung Panjang	89	8	3	12.4
27	Sengkotek	346	23	10	9.5
28	Simpang Tiga	560	29	10	7
29	Tani Aman	545	41	10	9.4
30	Harapan Baru	516	67	19	16.7
31	Rapak Dalam	351	41	15	16
32	Loa Buah	425	26	9	8.2
33	Loa Bakung	773	12	2	1.8
34	Karang Asam Ulu	265	22	4	9.8
35	Teluk Lerong Ulu	274	16	25	15
36	Lok Bahu	436	51	11	14.2
37	Karang Asam Ilir	221	25	4	13.1
38	Karang Anyar	329	18	25	13.1
39	Teluk Lerong Ilir	475	57	7	13.5
40	Jawa	425	66	2	16
41	Dadi Mulya	409	18	3	5.1
42	Sidodadi	529	30	7	7
43	Gunung Kelua	486	30	15	9.3
44	Air Hitam	713	76	16	12.9
45	Air Putih	634	12	1	2.1
46	Bukit Pinang	332	16	5	6.3
47	Lempake	137	6	13	13.9
48	Sempaja Selatan	266	31	8	14.7

49	Sungai Siring	252	15	0	6
50	Tanah Merah	548	16	0	2.9
51	Sempaja Utara	507	56	48	20.5
52	Sempaja Timur	663	78	48	19
53	Sempaja Barat	147	22	2	16.3
54	Budaya Pampang	126	5	0	4
55	Temindung Permai	430	6	7	3
56	Bandara	237	14	3	7.2
57	Sungai Pinang Dalam	387	31	11	10.9
58	Mugirejo	374	19	5	6.4
59	Gunung Lingai	289	15	1	5.5
	<b>TOTAL</b>	20,565	1,430	566	9.7

Sumber : <https://aksi.bangda.kemendagri.go.id/emonev/DashPrev/index/5> (2024)

Dataset yang digunakan dalam penelitian ini merupakan data yang tidak memiliki variabel target sehingga masuk dalam kategori *unsupervised learning*, sehingga algoritma yang digunakan adalah *K-Means*. Implementasi metode ini bertujuan guna mengelompokan objek yang memiliki tingkat kemiripan maksimum dalam suatu area tertentu.

Secara universal, algoritma *K-Means* mempunyai beberapa keunggulan diantaranya mudah diterapkan, durasi pembelajaran lebih efisiensi dan mudah untuk diadaptasi (Wahyono, 2018) serta merupakan sebuah algoritma clustering yang banyak penerapannya pada berbagai bidang (Heryadi, Wahyono, 2020)

Hierarki standar mekanisme data mining dalam penelitian ini menggunakan *Cross-Industry Standard Process Model for Data Mining* atau *CRISP-DM* dan teknik analisis yang digunakan adalah deskriptif kuantitatif. Menurut Larose (2014) dalam Feblian dan Daihani (2017), beberapa tahapan *CRISP-DM* yang akan diimplementasikan dalam penelitian ini adalah sebagai berikut (Feblian & Daihani, 2017):

a. *Business Understanding*, merupakan persepsi terhadap substansi dari proses data mining yang akan dilakukan sehingga menghasilkan output yang akan dihasilkan.

Output terhadap proses *clustering* ini akan memperoleh wawasan yang bisa dijadikan sebagai landasan Pemerintah Kota Samarinda untuk merumuskan kebijakan skala prioritas yang tepat dalam penanganan kasus stunting di Kota Samarinda.

b. *Data Understanding*, merupakan proses penghimpunan data, menelaah data agar dapat menafsirkan data yang akan dipakai dalam kegiatan penelitian, mengenali permasalahan yang berhubungan dengan data

*Dataset* yang digunakan dalam penelitian ini adalah data kasus stunting pada setiap kelurahan di Kota Samarinda tahun 2023 yang didapatkan melalui portal website <https://aksi.bangda.kemendagri.go.id/emonev/DashPrev/index/5>. Setiap atribut-atribut pada dataset yang digunakan mempunyai nilai sebagai berikut:

- 1) Nama Kelurahan merupakan kolom untuk nama kelurahan.
- 2) Jumlah Balita (Anak) merupakan kolom untuk jumlah anak yang masuk dalam kategori balita.
- 3) Stunting Pendek merupakan kolom untuk jumlah anak yang menderita stunting kategori pendek
- 4) Stunting Sangat Pendek merupakan kolom untuk jumlah anak yang menderita stunting kategori sangat pendek
- 5) Prevalensi (%) merupakan kolom nilai prevalensi

c. *Data Preparation*, pada langkah ini susunan basis data akan dipersiapkan sehingga kegiatan data mining akan menjadi mudah

Pada tahapan ini akan dilakukan proses *preprocessing* dengan menggunakan Microsoft Excel untuk pembersihan, mereduksi dan melakukan transformasi data serta aplikasi RapiMiner 10.1 untuk melakukan pengecekan *missing value*. Prosedur ini dilakukan guna mempertahankan kualitas data dengan melakukan beberapa tahapan, yaitu

1. Melakukan seleksi atribut data yang akan digunakan dalam proses clustering. Adapun atribut yang digunakan yaitu: Nama Kelurahan, Stunting Pendek, dan Stunting Sangat Pendek. Jumlah data yang digunakan dalam penelitian ini sebanyak 59 data.
2. Transformasi pada *dataset*. Dilakukan perubahan atribut pada dataset yaitu: Stunting Pendek dirubah menjadi Pendek, dan Stunting Sangat Pendek dirubah menjadi Sangat Pendek.
3. Melakukan proses pemeriksaan untuk memastikan apakah sebuah dataset memiliki *missing value* atau tidak, apabila terdapat *missing value* maka dilakukan *cleansing dataset*.
4. Apabila ada atribut pada *dataset* memiliki karakteristik nominal maka akan dilakukan proses inisiasi, terkecuali atribut data sudah dalam bentuk numerik maka tidak diperlukan inisiasi. Mengingat sifat atribut Nama Kelurahan adalah nominal maka akan dilakukan perubahan status atribut saat proses import dataset pada aplikasi RapidMiner.

d. *Modeling Phase*, sebuah bagian merumuskan teknik data mining yang akan diimplementasikan, memilih tools dan algoritma data mining, serta memilih parameter dengan nilai yang optimal.

Metode yang digunakan dalam proses data *mining* ini yaitu *clustering*. Untuk mendukung semua kegiatan penambangan data penelitian ini menggunakan aplikasi RapidMiner 10.1 untuk proses pengolahan data *clustering*, dan algoritma *K-Means*.

Prosedur yang akan dilakukan saat proses *clustering* dengan menggunakan algoritma *K-Means* adalah sebagai berikut (Dwitri, dkk, 2020) :

- 1) Terlebih dahulu menentukan jumlah *cluster* dan  $k$  buah data secara acak sebagai titik awal *centroid*.
- 2) Menghitung jarak terdekat dari setiap data dengan setiap titik *centroid* dengan menerapkan rumus *Euclidean Distance*

$$de = \sqrt{(xi - si)^2 + (yi - ti)^2}$$

Keterangan Rumus:

(x,y) = adalah koordinat objek

(s,t) = adalah koordinat *centroid*

i = adalah banyaknya objek

- 3) Kemudian data dikelompokkan kedalam *cluster* yang mempunyai titik pusat *cluster* terdekat dengan data tersebut.
- 4) Lakukan perhitungan kembali pusat *cluster*.

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj}$$

Keterangan Rumus:

$V_{ij}$  = adalah *centroid* rata-rata pada cluster ke  $i$  untuk variabel ke  $j$

$N_i$  = adalah jumlah anggota cluster ke  $i$

$i,k$  = adalah indeks dari *cluster*

$j$  = adalah indeks variabel

$X_{kj}$  = adalah nilai data ke  $k$  variabel ke  $j$  untuk cluster tersebut

- 5) Lakukan perhitungan kembali terhadap setiap data yang memakai *centroid* baru. Apabila *centroid* setiap *cluster* tidak mengalami perubahan maka prosedur clustering telah selesai.

e. *Evaluation Phase*,

Pada tahap ini akan dilakukan proses interpretasi terhadap hasil data *mining* yang sudah dihasilkan. *Evaluation Phase* dilakukan agar prosedur pada tahapan *modeling phase* sinkron dengan tahapan *business understanding*.

f. *Deployment Phase*

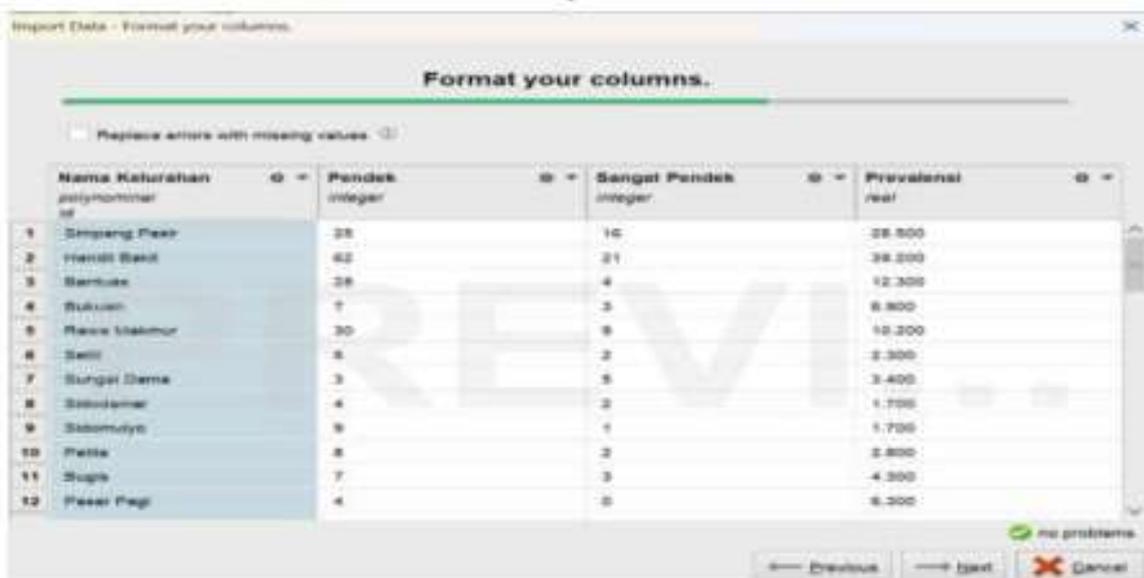
Setelah tahapan *Evaluation Phase* selesai dilakukan, maka prosedur selanjutnya akan dilakukan penyajian laporan tentang pengetahuan yang dihasilkan dari proses pengelompokan sebaran kasus stunting di Kota Samarinda tahun 2023.

## HASIL DAN PEMBAHASAN

### Penerapan Algoritma K-Means Dengan Aplikasi RapidMiner

Berdasarkan metodologi yang sudah dirumuskan, maka prosedur pembentukan model untuk fungsi *clustering* dengan menggunakan algoritma *K-Means* adalah sebagai berikut:

- a. Mengimport *dataset* kedalam Rapid Miner
- b. Menentukan role atribut sesuai kebutuhan, mengingat tipe data atribut Nama Kelurahan adalah nominal, maka atribut tersebut sebagai *id*



Gambar 1.  
Penentuan Role Atribut  
Sumber: Data diolah

- c. Melakukan pemeriksaan *missing value* pada dataset. Berdasarkan hasil output pada RapidMiner, dataset yang digunakan dalam penelitian ini tidak memiliki *missing value*

Name	Type	Missing	Statistics	Filter (3 / 3 attributes)
✓ Nama Kelurahan	Nominal	0	Count Terindung Permai (1)	Word Air Hitam (1)
✓ Pendek	Integer	0	Min 2	Max 76
✓ Sangat Pendek	Integer	0	Min 0	Max 56

Gambar 2.  
Proses Pemeriksaan Missing Value  
Sumber : Data diolah

- d. Untuk menentukan jumlah cluster yang ideal perlu memperhatikan hasil nilai *DBI* yang paling rendah. Sesuai ketentuan dari RapiMiner percobaan iterasi dilakukan sebanyak 10 kali dimulai dengan nilai *k* yang di input dari 2 sampai dengan 10. Setelah dilakukan pengujian nilai *Davies-bouldin index (DBI)*, maka nilai *k* yang paling baik adalah 10.

Tabel 2.  
*Davies-bouldin index (DBI)*

k	<i>Davies-bouldin index (DBI)</i>
2	0,655
3	0,657
4	0,665
5	0,561
6	0,672
7	0,567
8	0,670
9	0,590
10	0,444

Sumber: Data diolah

- e. Dengan adanya penentuan nilai *k* sebesar 10, kemudian dilakukan proses clustering dengan algoritma *K-Means* pada aplikasi RapidMiner, maka didapatkan hasil pengelompokannya sebanyak 10 *cluster* yang dimulai dari *cluster\_0* sampai dengan *cluster\_9*. Untuk *cluster 0* memiliki 17 anggota, *cluster 1* memiliki 1 anggota, *cluster 2* memiliki 1 anggota, *cluster 3* memiliki 14 anggota, *cluster 4* memiliki 1 anggota, *cluster 5* memiliki 2 anggota, *cluster 6* memiliki 4 anggota, *cluster 7* memiliki 1 anggota, *cluster 8* memiliki 14 anggota, dan *cluster 9* memiliki 4.

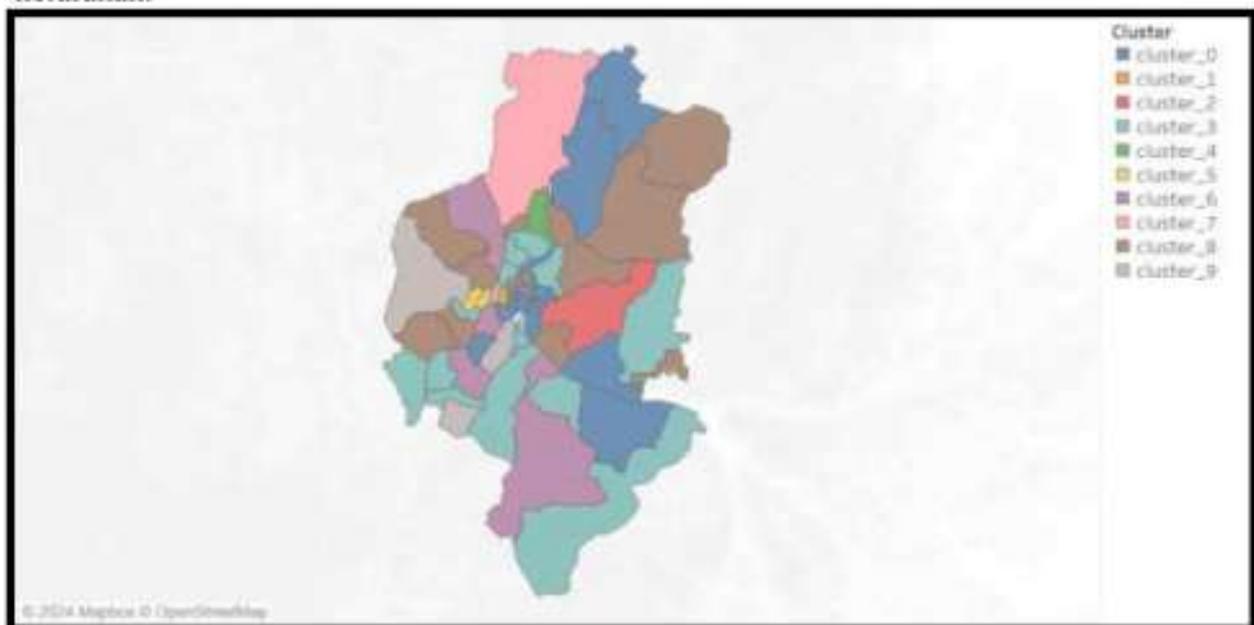
Tabel 3.  
Hasil *Clustering*

Nama Kelurahan	Cluster	Nama Kelurahan	Cluster	Nama Kelurahan	Cluster
Bukuan	cluster_0	Simpang Pasir	cluster_3	Sempaja Utara	cluster_7
Selili	cluster_0	Bantuas	cluster_3	Sungai Pinang Luar	cluster_8
Sungai Dama	cluster_0	Rawa Makmur	cluster_3	Sindang Sari	cluster_8
Sidodamai	cluster_0	Makroman	cluster_3	Sungai Kapih	cluster_8
Sidornulyo	cluster_0	Mesjid	cluster_3	Loa Bakung	cluster_8
Pelita	cluster_0	MANGKUPALAS	cluster_3	Karang Asam Ulu	cluster_8
Bugis	cluster_0	Sengkotek	cluster_3	Dadi Mulya	cluster_8
Pasar Pagi	cluster_0	Simpang Tiga	cluster_3	Air Putih	cluster_8
Pelabuhan	cluster_0	Loa Buah	cluster_3	Bukit Pinang	cluster_8
Karang Mumus	cluster_0	Karang Asam Ilir	cluster_3	Sungai Siring	cluster_8
Pulau Atas	cluster_0	Sidodadi	cluster_3	Tanah Merah	cluster_8
Baqa	cluster_0	Gunung Kelua	cluster_3	SEMPAJA BARAT	cluster_8
TENUN	cluster_0	Sempaja Selatan	cluster_3	Bandara	cluster_8
GUNUNG PANJANG	cluster_0	Sungai Pinang Dalam	cluster_3	Mugirejo	cluster_8
Lempake	cluster_0	SEMPAJA TIMUR	cluster_4	Gunung Lingai	cluster_8
BUDAYA PAMPANG	cluster_0	Teluk Lerong Ulu	cluster_5	Tani Aman	cluster_9
Temindung Permai	cluster_0	Karang Anyar	cluster_5	Rapak Dalam	cluster_9
Jawa	cluster_1	Handil Bakti	cluster_6	Lok Bahu	cluster_9
Sambutan	cluster_2	Sungai Keledang	cluster_6	Teluk Lerong Ilir	cluster_9
		Harapan Baru	cluster_6		
		Air Hitam	cluster_6		

Sumber: Data diolah

### Pengelompokan Sebaran Kasus Stunting

Dengan dilakukannya pengelompokan, maka beberapa wilayah kelurahan di Kota Samarinda terbagi menjadi beberapa anggota *cluster*. Sebanyak 17 kelurahan yang masuk dalam *cluster* 0, 1 kelurahan berada dalam *cluster* 1, 1 kelurahan adalah anggota *cluster* 2, sebanyak 14 kelurahan masuk dalam *cluster* 3, 1 kelurahan adalah anggota *cluster* 4, 2 kelurahan posisinya ada pada *cluster* 5, 4 kelurahan adalah anggota *cluster* 6, hanya 1 kelurahan ada pada *cluster* 7, 14 kelurahan yang masuk dalam *cluster* 8, dan *cluster* 9 memiliki anggota sebanyak 4 kelurahan.



Gambar 1.  
Peta Sebaran Pengelompokan Kasus Stunting di Kota Samarinda Tahun 2023

Ditinjau dari sudut pandang hasil *clustering*, bahwa di Kota Samarinda pada tahun 2023 telah terjadi kasus stunting untuk kategori pendek pada *cluster* 0 sebanyak 97 balita, *cluster* 1 sebanyak 66 balita, *cluster* 2 sebanyak 33 balita, *cluster* 3 sebanyak 380 balita, *cluster* 4 sebanyak 78 balita, *cluster* 5 sebanyak 34 balita, *cluster* 6 sebanyak 270 balita, *cluster* 7 sebanyak 56 balita, *cluster* 8 sebanyak 222 balita, dan *cluster* 9 sebanyak 190 balita.

Sedangkan untuk jumlah kasus stunting dengan kategori sangat pendek ada pada *cluster* 0 sebanyak 52 balita, *cluster* 1 sebanyak 2 balita, *cluster* 2 sebanyak 58 balita, *cluster* 3 sebanyak 142 balita, *cluster* 4 sebanyak 48 balita, *cluster* 5 sebanyak 50 balita, *cluster* 6 sebanyak 71 balita, *cluster* 7 sebanyak 48 balita, *cluster* 8 sebanyak 42 balita, dan *cluster* 9 sebanyak 43 balita



Gambar 2.  
Kasus Stunting di Kota Samarinda Tahun 2023

### Wawasan Hasil Proses *Clustering* Kasus Stunting

Setelah berhasil membentuk model *clustering*, maka wawasan yang dihasilkan adalah sebagai berikut:

1. Berdasarkan formula perhitungan prevalensi stunting, maka didapatkan hasil nilai prevalensi pada *cluster* 0 sebesar 4,50%, *cluster* 1 sebesar 16%, *cluster* 2 sebesar 6,94%, *cluster* 3 sebesar 10,89%, *cluster* 4 sebesar 19%, *cluster* 5 sebesar 13,93%, *cluster* 6 sebesar 18,67%, *cluster* 7 sebesar 20,51%, *cluster* 8 sebesar 5,27%, dan *cluster* 9 sebesar 12,89%
2. Nilai prevalensi stunting yang paling tinggi ada di *cluster* 7 sebesar 20,51%.
3. Dengan melihat hasil kerja *clustering*, maka wilayah di Kota Samarinda yang masuk dalam anggota *cluster* 7 adalah kelurahan Sempaja Utara.
4. Berdasarkan laporan data kemiskinan ekstrem yang dikelola oleh Dinas Sosial dan Pemberdayaan Masyarakat Kota Samarinda sampai pada periode tahun 2023, bahwa kelurahan Sempaja Utara merupakan wilayah di Kota Samarinda dengan status kemiskinan ekstrem yang paling tinggi dimana pada wilayah tersebut memiliki nilai prevalensi stuntingnya menduduki urutan pertama.

Dengan melihat wawasan yang didapatkan, maka ada beberapa strategi kebijakan yang dapat direncanakan oleh Pemerintah Kota Samarinda dalam penanganan kasus stunting di Kota Samarinda, antara lain:

1. Pemerintah Kota Samarinda dapat membuat skala prioritas penanganan stunting berdasarkan rangking nilai prevalensi hasil *clustering*, dimana prioritas paling utama untuk penurunan angka prevalensi adalah wilayah yang masuk dalam *cluster 7* yaitu Kelurahan Sempaja Utara.
2. Pemerintah Kota Samarinda perlu melakukan evaluasi lebih mendalam, apakah dinamika kemiskinan ekstrem yang terjadi di Kota Samarinda memiliki pengaruh langsung dengan perkembangan prevalensi stunting.

## KESIMPULAN

Berdasarkan hasil dan pembahasan, maka penulis mengambil beberapa kesimpulan sebagai berikut:

1. Algoritma *K-Means* dapat digunakan dalam melakukan proses *clustering* terhadap dataset kasus stunting di Kota Samarinda tahun 2023.
2. Ditinjau dari sudut pandang hasil *clustering*, bahwa kelurahan Sempaja Utara yang masuk dalam anggota *cluster 7* memiliki prevalensi stunting paling tinggi dan juga merupakan wilayah yang memiliki nilai kemiskinan ekstrem dengan urutan pertama.

Sebagai langkah akselerasi pencapaian tujuan penurunan prevalensi stunting sebaiknya Pemerintah Kota Samarinda membentuk sebuah kelompok kerja untuk melakukan evaluasi apakah kemiskinan ekstrem memiliki korelasi terhadap perkembangan kasus stunting di Kota Samarinda.

## UCAPAN TERIMAKASIH

Saya selaku penulis mengucapkan terima kasih kepada beberapa pihak yang sudah mendukung dalam penyusunan jurnal penelitian ini, yaitu:

1. Kepala Badan Pendapatan Daerah Kota Samarinda
2. Kepala Dinas Sosial dan Pemberdayaan Masyarakat Kota Samarinda

## DAFTAR PUSTAKA

- Apriyani, P., Dikananda, A. R., & Ali, I. (2023). Penerapan Algoritma K-Means dalam Klasterisasi Kasus Stunting Balita Desa Tegalwangi. *Jurnal Ilmu Komputer*, 2(1), 20–33.
- Darmansah, D. D., & Wardani, N. W. (2021). Analisis Pesebaran Penularan Virus Corona di Provinsi Jawa Tengah Menggunakan Metode *K-Means Clustering*. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 8(1), 105–117.
- Ditjen Bina Pembangunan Daerah - Kementerian Dalam Negeri, (2024). *Monitoring Pelaksanaan 8 Aksi Konvergensi Intervensi Penurunan Stunting Terintegrasi* <https://aksi.bangda.kemendagri.go.id/emonev/DashPrev/index/5>, diakses tanggal 9 Maret 2024
- Dwitri, N., Tampubolon, J. A., Prayoga, S., Ilmi Zer, F., & Hartama, D. (2020). Penerapan Algoritma *K-Means* Dalam Menentukan Tingkat Penyebaran Pandemi *Covid-19* Di Indonesia. *Jti (Jurnal Teknologi Informasi)*, 4(1), 101–105.

- Feblian, D., & Daihani, D. U. (2017). Implementasi Model Crisp-Dm Untuk Menentukan Sales Pipeline Pada PT. X. *Jurnal Teknik Industri*, 6(1), 1-12
- Fitriani, N. (2023). *Data terbaru! Prevalensi stunting di Jabar menurun 4,3%, pencapaian target WHO semakin dekat.* <https://opendata.jabarprov.go.id/id/artikel/data-terbaru-prevalensi-stunting-di-jabar-menurun-43-pencapaian-target-who-semakin-dekat>, diakses tanggal 8 Maret 2024
- Fitriyani, V. (2021). Analisis Clustering Provinsi Indonesia Berdasarkan Persebaran Virus Corona ( Covid-19 ) Menggunakan Algoritma K-Means. *Prosiding Pendidikan Matematika Dan Matematika*
- Kementerian Koordinator Bidang PMK, & Sekretariat Wakil Presiden RI. (2021). *StranasPercepatan Pencegahan Anak Kerdil 2018-2024*, Retrieved from Dokumen Penting:[https://tp-pkk.rokanhulukab.go.id/web-content/uploads/stranas\\_percepatan\\_pencegahan\\_anak\\_kerdil\\_2018\\_2024.pdf](https://tp-pkk.rokanhulukab.go.id/web-content/uploads/stranas_percepatan_pencegahan_anak_kerdil_2018_2024.pdf), diakses pada tanggal 24 Juni 2024
- Laili, U., & Andriani, R. A. (2019). Pemberdayaan Masyarakat Dalam Pencegahan Stunting. *Jurnal Pengabdian Masyarakat Ipteks*, 8-12
- Nabila, Z., Rahman Isnain, A., & Abidin, Z. (2021). Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means. *Jurnal Teknologi dan Sistem Informasi (JTSI)*, 2(2), 100-108.
- Nurohmah, Y., Mayasari, R., & Nurina Sari, B. (2023). Optimalisasi Performa *K-Means Clustering* Dengan Pca Dalam Analisis Tingkat Kemiskinan Di Jawa Barat. *Jurnal Mahasiswa Teknik Informatika*, 7(3), 1657-1665.
- Suyanto. (2019). *Data Mining Untuk Klasifikasi dan Klasterisasi Data, Edisi Revisi*. Bandung: Penerbit Informatika.
- Tasman, R., Wibowo, A., Indawati, R., Dwi Elisanti, A. (2020). Analisis Kluster Kejadian Stunting pada Balita di Provinsi Kalimantan Timur. *Jurnal Kesehatan*, 8(3), 143-150.
- Wahyono, T. (2018). *Fundamental Of Python For Machine Learning (Dasar-Dasar Pemrograman Python Untuk Machine Learning dan Kecerdasan Buatan)*, Cetakan Pertama. Yogyakarta: Penerbit Gava Media
- Wahyono, T., Heryadi, Y. (2020). *Machine Learning (Konsep dan Implementasi)*, Cetakan Pertama. Yogyakarta: Penerbit Gava Media
- Yuri, A. (2023). *Sering Disebut, Apa Arti Prevalensi, Cara Mengukur dan Fungsinya?*. <https://soloabadi.com/sering-disebut-apa-arti-prevalensi-cara-mengukur-dan-fungsinya/>, diakses tanggal 7 Maret 2024.