

OPTIMASI HASIL CLUSTERING DATA REKAM MEDIS BALITA DI DESA JUMPUT REJO DENGAN METODE ELBOW DALAM MENUNJANG PROGRAM PEMERINTAH MENGATASI STUNTING

Amir Ali¹, Sedy Ayu Mitra Uktutias²

^{1,2}Program Studi Rekam Medis & Informasi Kesehatan, Stikes Yayasan RS Dr. Soetomo
Jl. Kalidami No.14-16, Mojo, Kec. Gubeng, Kota SBY, Jawa Timur 60285
(031) 59181757

E-mail: amir_ali@stikes-yrsds.ac.id, sedy@stikes-yrsds.ac.id

ABSTRAK

Negara berkembang seperti Indonesia, masih mengalami masalah gizi buruk. Sesuai dengan Perpres RI no 72 tahun 2021 maka diperlukan suatu penguatan dan pengembangan baik itu sistem, data, informasi, riset dan inovasi. Tujuan dari penelitian ini yaitu melakukan visualisasi data terhadap data rekam medis balita yaitu umur dan berat badan pada balita di suatu wilayah tertentu tepatnya di desa jumput rejo kelurahan jumput rejo kecamatan sukodono kabupaten sidoarjo jawa timur. Visualisasi data ini dalam bentuk kelompok klaster-klaster data antropometri balita berdasarkan data umur dan data berat badan balita. Penelitian ini menggunakan metode teknik clustering dengan algoritma *K-Means* yang dioptimasi pada penentuan jumlah clusternya dengan metode *elbow* dalam menentukan nilai *K*(kluster) yang tepat dengan menggunakan metric *WCSS* (*Within Cluster Sum of Square*). Hasil dari penelitian ini didapatkan informasi optimasi jumlah kluster yang dihasilkan berjumlah 4 cluster data balita. Setelah dilakukan proses clustering maka didapatkan hasil dimana cluster 1 berjumlah 216 data balita, cluster 2 berjumlah 182 data balita, cluster 3 berjumlah 153 data balita dan cluster 4 berjumlah 29 data balita. Kesimpulan dari penelitian ini menghasilkan informasi bahwasannya hasil optimasi cluster didapatkan sejumlah 4 cluster yang akan dijadikan input penentuan cluster yang terbentuk pada proses pengelompokan data menggunakan algoritma *K-Means*

Kata Kunci: *Stunting, K-Means, elbow, Within Cluster Sum of Square, Cluster*

ABSTRACT

Developing countries like Indonesia, are still experiencing malnutrition problems. In accordance with Presidential Decree No. 72 of 2021, it is necessary to strengthen and develop systems, data, information, research and innovation. The purpose of this study is to visualize data on toddler medical record data, namely age and weight in toddlers in a certain area, precisely in Jumput Rejo Village, Jumput Rejo Village, Sukodono District, Sidoarjo Regency, East Java. Visualization of this data is in the form of clusters of toddler anthropometric data based on age data and toddler weight data. This study uses a clustering technique method with the K-Means algorithm which is optimized in determining the number of clusters with the elbow method in determining the correct K(cluster) value using the WCSS (Within Cluster Sum of Square) metric. The results of this study obtained information on optimizing the number of clusters produced totaling 4 toddler data clusters. After the clustering process is carried out, the results are obtained where cluster 1 has 216 toddler data, cluster 2 has 182 toddler data, cluster 3 has 153 toddler data and cluster 4 has 29 toddler data. The conclusion of this study yields information that the cluster optimization results obtained a number of 4 clusters which will be used as input for determining the clusters formed in the data grouping process using the K-Means algorithm

Keywords : *Stunting, K-Means, elbow, Within Cluster Sum of Square, Cluster*

1. PENDAHULUAN

1.1 Latar Belakang

Sesuai dengan Perpres RI no 72 tahun 2021 tentang percepatan penurunan stunting dimana dinyatakan bahwa terdapat salah satu strategi nasional yang dilakukan yaitu menurunkan nilai prevalensi stunting. Angka prevalensi adalah angka jumlah kasus stunting pada balita dalam suatu populasi tertentu. Menurut data kemenkes RI tahun 2019 terkait status gizi, nilai prevalensi balita stunting

di Indonesia tahun 2019 mencapai 27,7%. Secara umum angka prevalensi balita stunting provinsi di Indonesia berada pada kisaran 20-30%, jawa timur angka prevalensinya berada pada kisaran 26,9%. Diperlukan aksi terkait penurunan angka prevalensi stunting pada balita dalam rangka mencapai target nilai prevalensi stunting sebesar 14% pada tahun 2024 yang telah ditetapkan oleh pemerintah.

Kaitannya dengan target penurunan angka prevalensi stunting ini, maka sesuai dengan Perpres

RI no 72 tahun 2021 maka diperlukan suatu penguatan dan pengembangan baik itu sistem, data, informasi, riset dan inovasi. Pengukuran data antropometri balita diperlukan dalam melakukan monitoring terhadap pertumbuhan fisik balita. Data dari hasil pengukuran antropometri balita yang didapat dari data rekam medisnya misalnya berat badan menurut umur (BB/U) dimana data ini dapat digunakan untuk mengkluster data.

K-Means merupakan metode yang digunakan untuk mengelompokkan data. Metode ini mengelompokkan data ke dalam klaster yang mempunyai tingkat kemiripan yang sama (P. Sari et al., 2017). Optimasi jumlah clusternya dilakukan dengan menggunakan metode elbow dalam menentukan nilai K(klaster) yang tepat dengan bantuan metric WCSS (Within Cluster Sum of Square).

Hasil dari klusterisasi dalam bentuk visualisasi data ini dapat dijadikan sebagai dasar informasi bagi tenaga kesehatan untuk memberikan penyuluhan terkait gizi balita kepada pihak terkait dalam mendukung program pemerintah dalam percepatan penurunan stunting.

1.2 Tujuan Penelitian

Tujuan penelitian ini yaitu melakukan visualisasi data balita di suatu wilayah tertentu tepatnya di desa jumpat rejo kelurahan jumpat rejo kecamatan sukodono kabupaten sidoarjo jawa timur. Visualisasi data ini dalam bentuk kelompok klaster-klaster data informasi gizi balita berdasarkan data antropometri balita menggunakan variabel umur dan berat badan.

1.3 Referensi

1.3.1 Rekam Medis

Puskesmas merupakan pusat kesehatan masyarakat. Setiap kali masyarakat berobat tentu catatan hasil pemeriksaannya dicatat dan direkam dalam dokumen rekam medis pasien. Rekam medis mempunyai 2 bagian. Bagian pertama yaitu bagian tentang individu yang berisi suatu informasi tentang kondisi kesehatan dan penyakit pasien. Bagian yang kedua berisi tentang manajemen yang mana berisi informasi tentang pertanggungjawaban dari segi manajemen maupun keuangan dari kondisi kesehatan dan penyakit pasien (Mauladi, 2017)

1.3.2 Indeks Antropometri

Indeks antropometri adalah kombinasi pada beberapa variabel antropometri untuk menilai status gizi. Berat badan menurut umur (BB/U) adalah salah satu indeks antropometri. Berat badan adalah salah satu variabel yang memberikan gambaran massa tubuh. Massa tubuh sangat sensitif terhadap perubahan-perubahan yang mendadak, misalnya karena terserang penyakit infeksi, menurunnya nafsu

makan atau menurunnya jumlah makanan yang dikonsumsi (P. Sari et al., 2017)

1.3.3 Data Mining

Data mining merupakan proses penggalian data dengan pola tertentu dari data yang tersimpan dalam database. Data mining mencakup pengumpulan data, ekstraksi data, analisis data dan statistik data. Data mining juga dikenal sebagai knowledge discovery, knowledge extraction, data/pattern analysis, information harvesting dan lain-lain (Arhani, Muhammad. Nasir, 2020).

Istilah lain untuk data mining adalah Knowledge Discovery in Database (KDD). Knowledge Discovery in Database (KDD) adalah cara non-trivial dalam upaya menemukan pola pada sebuah data yang bersifat baru, dapat dimengerti dan berguna. Terdapat proses dalam tahapan KDD yaitu adanya analisa dari sejumlah data yang besar (Wandana, Jeri. Defit, 2020)

1.3.4 Cluster

Cluster merupakan group atau kelompok. Mengcluster berarti melakukan pengelompokan data. Pada proses pengelompokan data hal yang diperhatikan adalah kemiripan data pada masing-masing anggota *cluster* (Mujilahwati & Wardhani, 2021)

Prinsip dasar dalam melakukan analisis *cluster* adalah mengelompokkan objek pada suatu *cluster* yang memiliki kemiripan yang sangat besar dengan objek lain dalam *cluster* yang sama, tetapi tidak mirip dengan objek lain pada *cluster* yang berbeda (Hidayat et al., 2017)

1.3.5 K-Means

K-Means Clustering adalah metode pengelompokan data ke dalam satu atau lebih cluster/kelompok (Syaputri et al., 2022). Dalam K-Means, untuk setiap *cluster* memiliki titik pusat yang disebut dengan *centroid* (Juliana et al., 2021). Pada K-Means akan dilakukan perhitungan jarak antara data dengan titik pusat (*centroid*) untuk mengukur tingkat kemiripan (Rohmah et al., 2021)

K-Means merupakan salah satu algoritma dalam data mining yang dapat digunakan untuk melakukan pengelompokan data (Dinata et al., 2020). Algoritma K-Means termasuk dalam teknik dalam machine learning yaitu termasuk dalam teknik unsupervised learning (D. N. P. Sari & Sukestiyarno, 2021)

Adapun langkah-langkah dalam algoritma K-Means:

1. Tentukan berapa cluster yang akan dibentuk
2. Bangkitkan initial centroid secara acak dimana banyaknya sesuai dengan banyak cluster yang akan dibentuk

3. Hitung jarak masing-masing data dengan centroid dan pilih yang terdekat
4. Tentukan centroid baru
5. Jika centroid baru posisinya sama dengan centroid lama, maka proses clustering selesai, jika tidak kembali ke tahap 3

1.3.6 Optimasi K-Means menggunakan metode Elbow

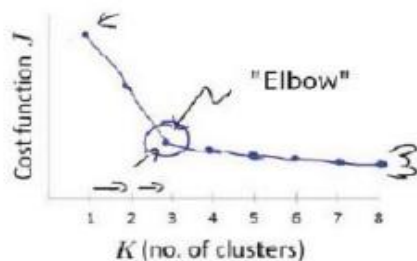
Optimasi K-Means dapat ditambahkan fungsi optimasi yang akan memilih jumlah awal cluster secara tepat. Fungsi optimasi ini menggunakan metode Elbow. Metode Elbow akan menunjukkan jumlah cluster terbaik yang dihasilkan, yang ditunjukkan dengan pembentukan siku pada suatu titik (Muningsih, Elly. Kiswati, 2018). MetricWCSS (Within Cluster Sum of Squares) akan membantu dalam proses optimasi jumlah cluster yang terbentuk dengan menggunakan metode elbow.

Algoritma Metode Elbow dalam menentukan nilai K pada K-Means

1. Mulai
2. Menginisialisasi awal nilai K
3. Menaikkan nilai K
4. Menghitung hasil sum of square error dari tiap nilai K
5. Melihat hasil sum of square error dari nilai K yang turun secara drastis
6. Menetapkan nilai K yang berberbentuk siku
7. Selesai

Rumus SSE pada K-Means :

$$SSE = \sum_{K=1}^K \sum_{X_i \in S_K} \| X_i - C_k \|^2 \quad (1)$$

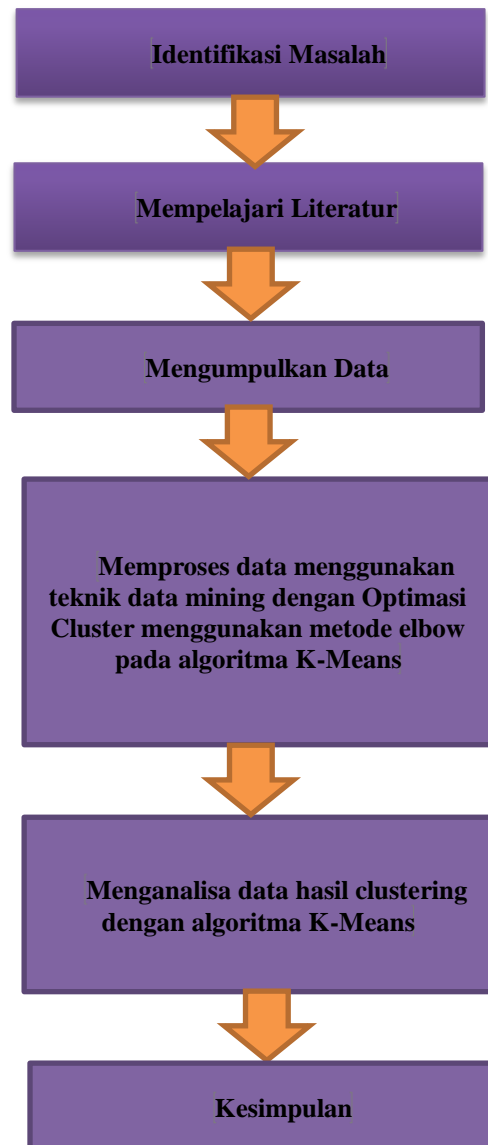


Gambar 1. Grafik nilai k

Terlihat di gambar 2 diatas, bahwasannya nilai K mengalami penurunan dan terlihat pembentukan titik siku terlihat dari nilai cluster = 3, maka jumlah cluster yang optimal adalah K=3.

2. HASIL DAN PEMBAHASAN

Data yang digunakan adalah data antropometri balita. Jumlah data yang di proses sejumlah 1070 data. Penggunaan metode elbow pada penelitian ini digunakan untuk menentukan jumlah cluster yang dijadikan sebagai input pada penentuan jumlah cluster yang akan terbentuk pada proses didalam algoritma K-Means. Data rekam medis balita yang digunakan yaitu umur dan berat badan balita.



Gambar 2. Tahapan penelitian

Dalam memproses data menggunakan teknik data mining dengan menggunakan algoritma K-Means dan optimasi cluster dengan metode elbow, perlu dipersiapkan langkah-langkah sebagai berikut:

2.1 Preprocessing Data

Sebelum dilakukan proses clustering pada data balita tersebut, maka preprocessing data perlu dilakukan. Preprocessing data dilakukan dengan mengambil data transaksi dengan cara purposive

sampling dengan dilakukan pembersihan data yang tidak lengkap (Muningsih, Elly. Kiswati, 2018)

No	Posyandu	Nama	Kelamin	Umur	BB	TB	
0	NaN	NaN	NaN	L=1, P=2	(Bln)	(Kg) (Cm)	
1	1.0	CIRO 1	ARFABIAN IZZAM	1	28	11	87
2	2.0	CIRO 1	M. REZA	1	59	16.4	110
3	3.0	CIRO 1	M. ZAKI	1	61	0	0
4	4.0	CIRO 1	QAMELA PUTRI	2	32	12.6	95
...
1066	1066.0	SURYA ASRI 2B	RAANIYAH	2	51	13.3	98
1067	1067.0	SURYA ASRI 2B	ALFIZ ARASHKA	1	15	7.7	75
1068	1068.0	SURYA ASRI 2B	BERLIAN	2	2	0	0
1069	1069.0	SURYA ASRI 2B	LAVANIA ELSHANUM	2	15	0	0
1070	1070.0	SURYA ASRI 2B	NAYLA NURSALSABILA	2	-1	0	0

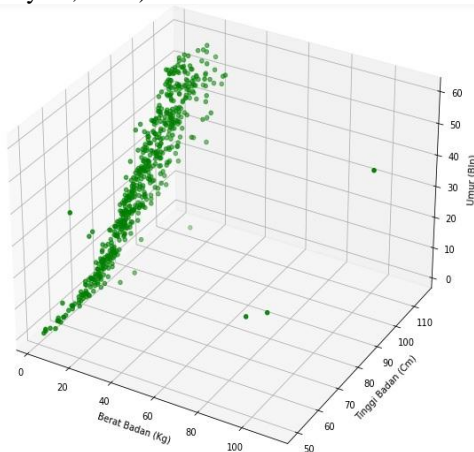
1071 rows x 7 columns

Gambar 1. Data awal penelitian

Pembersihan data dilakukan terhadap data yang bernilai 0. Dari 1070 data, maka didapatkan data sebanyak 583 data yang lengkap, yang terdiri dari 305 data balita laki-laki dan 278 data balita perempuan.

2.2 Perhitungan Normalisasi Data

Normalisasi data ini menggunakan proses scaling data. Normalisasi data merupakan teknik mengubah nilai numerik dalam dataset ke skala umum (Ambarwari, Agus. Andrian Qadhli Jafar, Herdiyeni, 2020).



Gambar 2. Persebaran data antropometri balita

Terdapat perbedaan selisih angka yang cukup jauh antara variabel tinggi badan, variabel berat badan dan variabel umur. Untuk memperkecil selisih angka tersebut dilakukan dengan menghitung nilai normalisasi data. dengan menggunakan rumus berikut (Syaputri et al., 2022):

$$\text{Nilai normalisasi} = \frac{(\text{nilai awal} - \text{nilai minimum})}{(\text{nilai maksimal} - \text{nilai minimum})} \dots (2)$$

Nilai dari variabel yang digunakan akan dinormalisasi ke dalam rentang 0-1. Proses normalisasinya sebagai berikut adalah

1. Mencari nilai maksimum dan minimum
 - a. Variabel berat badan (X)

Didapatkan nilai maksimum (Xmaks) = 112,3 Kg

Didapatkan nilai Minimum (Xmin) = 2,8 Kg
 - b. Variabel tinggi badan (Y)

Didapatkan nilai maksimum (Ymaks) = 115 Cm

Didapatkan nilai Minimum (Ymin) = 50 Cm
 - c. Variabel Umur (U)

Didapatkan nilai maksimum (Umaks) = 60 Bulan

Didapatkan nilai Minimum (Umin) = 1 Bulan

2. Menghitung nilai normalisasi

Normalisasi data dilakukan dengan menggunakan rumus...2) diatas pada 583 data balita. Berikut hasil normalisasi data:

normalisasi_umur	normalisasi_berat_badan	normalisasi_tinggi_badan
0.427119	0.074886	0.569231
0.952542	0.124201	0.923077
0.494915	0.089498	0.692308
0.494915	0.091324	0.600000
0.664407	0.208219	0.846154
...
0.037288	0.035616	0.223077
0.867797	0.193607	0.930769
0.816949	0.079452	0.707692
0.816949	0.095890	0.738462
0.206780	0.044749	0.384615

Gambar 3. Hasil normalisasi data

2.3 Clustering Data

Proses clustering yang dimulai dari penentuan jumlah clusternya dapat menggunakan metode elbow. Penentuan jumlah cluster akan menentukan hasil akhir dari pengelompokan datanya (K-means et al., 2021), lalu diproses menggunakan algoritma K-Means.

Rumus euclidean distance digunakan untuk menghitung jarak dari titik pusat cluster ke titik data masing-masing. Rumusnya sebagai berikut:

$$d = \sqrt{(x1 - x2)^2 + (y1 - y2)^2} \dots (3)$$

Dari hasil perhitungan diatas, maka data dapat dikelompokkan berdasarkan jarak terpendek ke pusat cluster. Pada iterasi berikutnya, lakukan proses perhitungan kembali titik pusat cluster yang baru dengan rumus sebagai berikut:

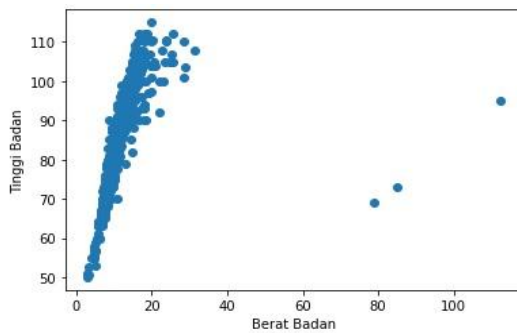
$$C_i = \frac{1}{M} \sum_{j=1}^M X_j \dots (4)$$

Dimana: titik pusat cluster yang baru diperoleh dari penjumlahan data yang menjadi anggota clusternya dibagi sebanyak data anggota clusternya

Jika titik pusat cluster yang baru posisinya sama dengan centroid lama, maka poses clustering selesai, jika tidak kembali pada perhitungan jarak masing-masing data dengan centroidnya seperti pada proses di awal

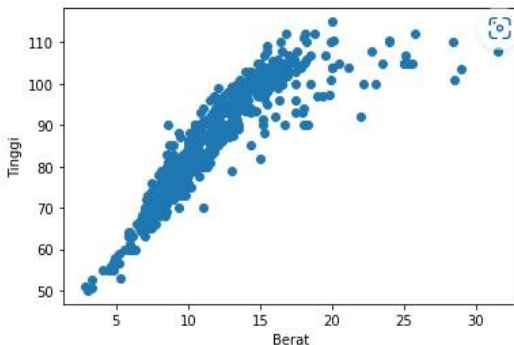
2.3.1 Menghilangkan Data Pencilan(Outlier)

Keberadaan data pencilan (outlier) ini akan menyebabkan penyimpangan terhadap hasil analisis data. Pencilan merupakan suatu pengamatan yang titik pengamatannya menyimpang jauh dari pola data. (P. P. Sari et al., 2021). Oleh karena itu data pencilan (outlier) tersebut perlu dihilangkan.



Gambar 4. Visualisasi data outlier

Dari gambar diatas diperoleh informasi yaitu 3 data yang posisinya berada jauh dari data lainnya. 3 data ini menyimpang dari pola data dan merupakan pencilan (outlier). Dengan proses tertentu maka akan di proses menghilangkan data pencilan (outlier).



Gambar 5. Visualisasi data setelah dihilangkan data outliernya

Dari gambar diperoleh informasi, tidak ada data yang menyimpang jauh dari pola datanya. Sehingga setelah menghilangkan data pencilan (outlier) maka datanya berjumlah 580 data balita

2.3.2 Scaling Data

Data scaling merupakan teknik normalisasi data (Ambarwari, Agus. Andrian Qadhli Jafar, Herdiyeni, 2020). Proses scaling data pada penelitian

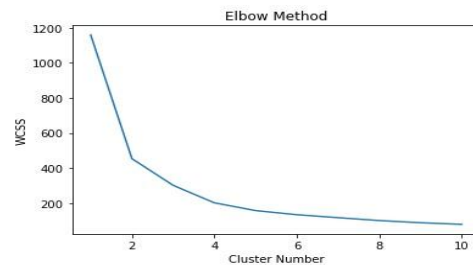
ini hanya menggunakan variabel berat badan menurut umur (BB/U) dimana variabel yang digunakan yaitu umur dan berat badan.

Berikut hasil data scaling untuk variabel umur dan berat badan :

```
array([[ -0.1206623 , -0.26730887],
       [ 1.7975589 ,  1.04038438],
       [ 0.12685011,  0.12015558 ],
       ...,
       [ 1.30253408, -0.14622616],
       [ 1.30253408,  0.28967159],
       [-0.92507765, -1.06645474]])
```

2.3.3 Penentuan Jumlah Cluster Dengan Menggunakan metode elbow

Penentuan jumlah cluster merupakan hal yang penting sebelum memproses datanya dengan algoritma K-Means. Banyaknya cluster akan menentukan hasil akhir dari proses pengelompokan data menggunakan algoritma K-Means. Pada penelitian ini menggunakan pendekatan metode elbow untuk menentukan jumlah cluster yang akan terbentuk.



Gambar 6. Grafik jumlah cluster metode elbow

Dari gambar diatas, diperoleh bahwasannya jumlah cluster yang optimum adalah berjumlah 4 cluster.

2.4 Hasil Clustering Data

Clustering data menghasilkan 4 cluster, dimana :

1. Cluster 1 mempunyai jumlah anggota cluster sebanyak 216 balita, dimana cluster 1 ini mempunyai rentang nilai pada umur dan berat badan yang relatif lebih kecil daripada cluster lain. Jumlah data cluster 1 untuk jenis kelamin laki-laki sebanyak 109 data, sedangkan yang berjenis kelamin perempuan sebanyak 107 data. Dengan rincian anggota per posyandu sebagai berikut :

Tabel 1 . Jumlah balita anggota cluster 1

Nama Posyandu	Jumlah Balita
Ciro 1	34
Ciro 2	28
Citra surya mas	15
Jumput wetan	15
Jumput rejo indah	25

Keling	34
Puri sejahtera 3	6
Surya asri 2b	14
Citra gading	22
Kedung 1	23

2. Cluster 2 mempunyai jumlah anggota cluster sebanyak 182 balita, dimana cluster 2 ini mempunyai rentang nilai pada umur dan berat badan yang relatif lebih tinggi dari pada cluster 1 dan cluster 3. Jumlah data pada cluster 2 untuk jenis kelamin laki-laki sebanyak 91 data, sedangkan untuk jenis kelamin perempuan sebanyak 91 data. Dengan Rincian anggota per posyandu sebagai berikut :

Tabel 2 . Jumlah Balita Anggota Cluster 2

Nama Posyandu	Jumlah balita
Ciro 1	11
Ciro 2	27
Citra surya mas	14
Jumput wetan	17
Jumputrejo indah	24
Keling	19
Puri sejahtera 3	8
Surya asri 2b	15
Citra gading	28
Kedung 1	19

3. Cluster 3 mempunyai jumlah anggota cluster sebanyak 153 balita, dimana cluster 3 ini mempunyai rentang nilai pada umur dan berat badan yang relatif lebih tinggi dari pada cluster 1 dan relatif lebih rendah daripada cluster 2 dan cluster 4. Jumlah data pada cluster 3 untuk jenis kelamin laki-laki sebanyak 88 data, sedangkan untuk jenis kelamin perempuan sebanyak 65 data. Dengan Rincian anggota per posyandu sebagai berikut :

Tabel 3 . Jumlah balita anggota cluster 3

Nama Posyandu	Jumlah Balita
Ciro 1	27
Ciro 2	19
Citra surya mas	7
Jumput wetan	10
Jumputrejo indah	14
Keling	26
Puri sejahtera 3	6
Surya asri 2b	10
Citra gading	13
Kedung 1	21

4. Cluster 4 mempunyai jumlah anggota cluster sebanyak 29 balita, dimana cluster 4 ini mempunyai rentang nilai pada umur dan berat

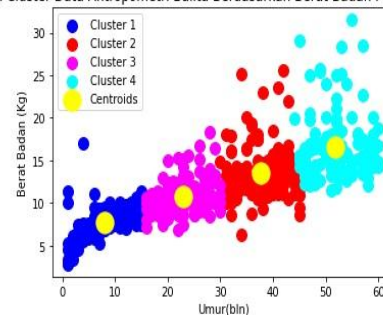
badan yang relatif lebih tinggi dari pada cluster lainnya. Jumlah data pada cluster 4 untuk jenis kelamin laki-laki sebanyak 16, sedangkan untuk jenis kelamin perempuan sebanyak 13. Dengan Rincian anggota per posyandu sebagai berikut :

Tabel 4 . Jumlah Balita Anggota Cluster 4

Nama Posyandu	Jumlah Balita
Ciro 1	3
Ciro 2	3
Jumput wetan	6
Jumputrejo indah	2
Keling	4
Puri Sejahtera 3	2
Surya asri 2b	3
Citra gading	3
Kedung 1	3

Hasil visualisasi data pengelompokan data balita sebagai berikut:

Visualisasi Cluster Data Antropometri Balita Berdasarkan Berat Badan Menurut Umur (BB/U)



Gambar 7. Visualisasi hasil clustering data antropometri balita

3. KESIMPULAN

Sebanyak 580 data antropometri balita berhasil diidentifikasi untuk dikelompokkan berdasarkan tingkat kemiripannya. Pada proses pengelompokan datanya digunakan algoritma K-Means dengan optimasi jumlah cluster menggunakan metode elbow.

Hasil optimasi cluster didapatkan sejumlah 4 cluster yang akan dijadikan input penentuan cluster yang terbentuk pada proses pengelompokan data menggunakan algoritma K-Means. Informasi baru akan dihasilkan dari proses clustering balita di desa jumput rejo sukodono sidoarjo dari sisi umur dan berat badan

Khusus pada cluster 4, ini mempunyai rentang nilai pada umur dan berat badan yang relatif lebih tinggi dari pada cluster lainnya. Cluster ini mempunyai jumlah anggota cluster sebanyak 29 balita. Jumlah data pada cluster 4 untuk jenis kelamin laki-laki sebanyak 16, sedangkan untuk jenis kelamin perempuan sebanyak 13.

Pada penelitian berikutnya perlu dikembangkan metode optimasi lain untuk mendapatkan jumlah cluster yang optimal yang dapat digunakan sebagai input jumlah cluster yang terbentuk pada proses pengelompokan data menggunakan algoritma K-Means

PUSTAKA

- Ambarwari, Agus. Andrian Qadhli Jafar, Herdiyeni, Y. (2020). Analisis Pengaruh Data Scaling Terhadap Performa Algoritme Machine Learning untuk Identifikasi Tanaman. *Resti*, 4, 117–122.
- Arhani, Muhammad. Nasir, M. (2020). *Data Mining*. Dinata, R. K., Safwandi, S., Hasdyna, N., & Azizah, N. (2020). Analisis K-Means Clustering pada Data Sepeda Motor. *INFORMAL: Informatics Journal*, 5(1), 10. <https://doi.org/10.19184/isj.v5i1.17071>
- Hidayat, R., Wasono, R., & Darsyah, M. Y. (2017). Pengelompokan Kabupaten/Kota Di Jawa Tengah Menggunakan Metode K-Means Dan Fuzzy C-Means. *Prosiding Seminar Nasional & Internasional*, 240–250. <https://jurnal.unimus.ac.id/index.php/psn12012010/article/view/3017/2932>
- Juliana, E., Aleyda, V. N., & Yuliana, Y. (2021). Penerapan Metode Clustering K-Means Untuk Membantu Menentukan Tingkatan Status Daerah Dampak Covid-19. *Jurnal MediaTIK*, 4(3), 112. <https://doi.org/10.26858/jmtik.v4i3.23698>
- K-means, M., Purwaningrum, O., Putra, Y. Y., & Arifiyanti, A. A. (2021). *Penentuan Kelompok Status Gizi Balita dengan Menggunakan*. 15(2), 129–136.
- Mauladi, K. F. (2017). Informasi Data Rekam Medis Pasien Di Rs Bedah Mitra Kabupaten Lamongan Dengan Metode Naive Bayes Dan Clustering Dalam Penentuan Penyakit Hypertensi. *Joutica*, 2(1), 2–7. <https://doi.org/10.30736/jti.v2i1.24>
- Mujilahwati, S., & Wardhani, R. (2021). Implementasi Fuzzy C-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Masuk Perguruan Tinggi. *Joutica*, 6(1), 448. <https://doi.org/10.30736/jti.v6i1.582>
- Muningsih, Elly. Kiswati, S. (2018). SISTEM APLIKASI BERBASIS OPTIMASI METODE ELBOW UNTUK PENENTUAN CLUSTERING PELANGGAN. *JOUTICA*, 3.
- Rohmah, A., Sembiring, F., & ... (2021). Implementasi Algoritma K-Means Clustering Analysis Untuk Menentukan Hambatan Pembelajaran Daring (Studi Kasus: Smk Yaspim *Sistem Informasi Dan ...*, 290–298. <https://sismatik.nusaputra.ac.id/index.php/sismatik/article/view/32>
- Sari, D. N. P., & Sukestiyarno, Y. L. (2021). Analisis Cluster dengan Metode K-Means pada Persebaran Kasus Covid-19 Berdasarkan Provinsi di Indonesia. *PRISMA, Prosiding Seminar Nasional Matematika*, 4, 602–610. <https://journal.unnes.ac.id/sju/index.php/prisma/>
- Sari, P. P., Herdiani, E. T., & Sunusi, N. (2021). Outlier Detection Using Minimum Vector Variance Algorithm with Depth Function and Mahalanobis Distance. *Jurnal Matematika, Statistika Dan Komputasi*, 17(3), 418–427. <https://doi.org/10.20956/j.v17i3.12629>
- Sari, P., Pramono, B., & Sagala, L. ode H. S. (2017). Improve K-Means Terhadap Status Nilai Gizi Pada Balita. *SemanTIK*, 3(1), 143–148. <https://doi.org/10.1063/1.2957900>
- Syaputri, V., Hartama, D., Anggraini, F., Safii, M., & Dewi, R. (2022). Penerapan Algoritma K-Means Untuk Menentukan Status Gizi Balita (Studi Kasus: Puskesmas Kecamatan Jawa Maraja Bah Jambi). *JATI (Jurnal Mahasiswa Teknik Informatika)*, 6(1), 94–102. <https://doi.org/10.36040/jati.v6i1.4630>
- Wandana, Jeri. Defit, S. S. (2020). Klasterisasi Data Rekam Medis Pasien Pengguna Layanan BPJS Kesehatan Menggunakan Metode K-Means. *Informasi Dan Teknologi*, 2.