

Performansi Algoritma *Clustering K-Means* untuk Penentuan Status Malnutrisi pada Balita

Siska Narulita ¹, Prihati Prihati ², Andreas Tigor Oktaga ³, Aditya Eka Widyantoro ⁴
^{1,2,3,4} Program Studi Sistem dan Teknologi Informasi, Institut Teknologi dan Bisnis Semarang
Email: siskanarulita84@gmail.com

ABSTRAK

Malnutrisi masih menjadi permasalahan global dan menjadi faktor penyebab kematian yang tertinggi bagi anak-anak di negara berkembang. Status gizi merupakan tolak ukur keberhasilan dalam pemenuhan kebutuhan gizi pada balita atau anak yang dapat dilihat pada pencapaian berat badan balita atau anak terhadap umur. Dampak gizi buruk pada anak balita dalam jangka pendek, yaitu terganggunya pertumbuhan fisik pada anak, gangguan metabolisme, gangguan perkembangan otak, dan gangguan kecerdasan. Dampak pada jangka panjangnya, gizi buruk akan berakibat pada penurunan prestasi belajar dan kemampuan kognitif, penurunan kekebalan tubuh, beresiko mengalami diabetes, pembuluh darah dan jantung, obesitas, stroke, kanker, serta disabilitas di usia lanjut. Untuk mengetahui pertumbuhan dan perkembangan balita, khususnya terkait dengan perkembangan nilai gizi balita, pengukuran terhadap tinggi dan berat badan balita dilakukan secara rutin melalui kegiatan posyandu. Data tinggi dan berat badan balita akan dikelompokkan sesuai dengan karakteristik gizi dari masing-masing balita. Dalam penelitian ini akan dilakukan pengelompokan balita sesuai dengan karakteristik gizinya memakai metode *clustering* dengan algoritma *K-Means*. Dari hasil evaluasi, jumlah *cluster* 3 lebih optimal dari jumlah *cluster* 5 pada *dataset* balita yang digunakan pada penelitian. Hasil penelitian yang diperoleh diharapkan bisa dipakai sebagai dasar atau bahan rekomendasi untuk posyandu di seluruh Indonesia dalam penentuan pengelompokan status gizi balita, sehingga penanganannya lebih tepat dan dapat membantu proses pencegahan malnutrisi pada anak balita.

Kata kunci: Malnutrisi, *K-Means*; *Clustering*; *Data mining*

ABSTRACT

Malnutrition has always been a global problem and is the main cause of child mortality in developing countries. Nutritional status is a measure of success in meeting the nutritional needs of toddlers which can be seen from the achievement of toddler's weight. The short-term effects of malnutrition on toddlers are physical growth disorders in children, metabolic disorders, brain development disorders, and intelligence disorders. In the long term, malnutrition will cause a decrease in learning outcomes and cognitive abilities, decreased immunity, the risk of developing diabetes, heart and blood vessel disease, obesity, stroke, cancer, and disability in old age. To find out about the growth and development of early childhood, especially in relation to the development of early childhood nutritional values, regular measurements of the height and weight of toddlers from an early age are carried out through Posyandu activities. Toddler height and weight data will be grouped according to the nutritional characteristics of each toddler. In this study, infants will be clustered based on their nutritional characteristics using the K-Means algorithm clustering method. Based on the evaluation results, the number of clusters 3 is more optimal than the number of clusters 5 in the toddler dataset used in the study. It is hoped that the research results obtained can serve as a basis or recommendation material for posyandu throughout Indonesia to determine the nutritional status of the under-five age group, so that the handling is more appropriate and can help prevent malnutrition in under-fives.

Keywords: *Malnutrition*, *K-Means*; *Clustering*; *Data mining*

PENDAHULUAN

Pengerdikan (*stunting*) yang terjadi pada anak adalah contoh dari kasus malnutrisi atau gizi buruk yang saat ini masih ada di Indonesia ("Dinas Kesehatan Kabupaten Lima Puluh Kota, 2015). Pengertian *stunting* adalah gejala dari malnutrisi kronis (Dinas Kesehatan Kabupaten Lima Puluh Kota, 2015). Malnutrisi masih menjadi permasalahan global dan menjadi faktor penyebab kematian yang tertinggi bagi anak-anak di negara berkembang (Aryani & Riyandry, 2019). Status gizi merupakan tolak ukur keberhasilan dalam pemenuhan kebutuhan gizi pada anak balita yang dapat dilihat pada pencapaian berat badan balita terhadap umur (Sulistyawati, 2019). Pengertian dari malnutrisi itu sendiri adalah dimana nutrisi makanan yang dibutuhkan tidak sesuai yang dapat berakibat pada terjadinya gangguan biologi (Perdana *et al.*, 2020). Malnutrisi dibedakan menjadi 2 (dua), yaitu gizi kurang (*marasmus*, *kwashiorkor*, dan *marasmus-kwashiorkor*) dan gizi lebih (obesitas) (Perdana *et al.*, 2020).

Dampak jangka pendek dari gizi buruk terhadap balita, yaitu terjadi gangguan pertumbuhan fisik pada anak, gangguan metabolisme, gangguan perkembangan otak, dan gangguan kecerdasan. Dampak jangka panjangnya, gizi buruk akan dapat berakibat pada penurunan prestasi belajar dan kemampuan kognitif, penurunan kekebalan tubuh, resiko terkena diabetes, penyakit pembuluh darah dan jantung, obesitas, stroke, kanker, serta disabilitas di usia lanjut (Kementerian Kesehatan Republik Indonesia, 2019). Berdasarkan "Peraturan Menteri Kesehatan Republik Indonesia Nomor 2 Tahun 2020 tentang Standar Antropometri Anak, standar antropometri anak didasarkan pada parameter berat badan dan panjang/tinggi badan yang terdiri atas 4 (empat) indeks, yaitu (Peraturan Menteri Kesehatan Republik Indonesia Nomor 2 Tahun 2020 Tentang Standar Antropometri Anak, 2020):

- a. Berat badan menurut umur (BB/U);
- b. Panjang/tinggi badan menurut umur (PB/U atau TB/U);
- c. Berat badan menurut panjang/tinggi badan (BB/PB atau BB/TB); dan
- d. Indeks massa tubuh menurut umur (IMT/U)."

Pertumbuhan yang sehat, organ yang berfungsi dengan baik, dan optimalnya perkembangan kognitif pada balita dapat terjamin jika pemenuhan kebutuhan gizinya tercukupi (Fitriyanto & Mahfudz, 2020). Kekurangan gizi dapat berdampak buruk bagi balita bahkan menimbulkan kematian (Fitriyanto & Mahfudz, 2020). Untuk mengetahui pertumbuhan dan perkembangan balita, khususnya terkait dengan perkembangan nilai gizi balita, pengukuran tinggi dan berat badan dilakukan secara rutin melalui kegiatan posyandu (Irfiani & Rani, 2018). Data dari tinggi dan berat badan balita tersebut dikelompokkan sesuai karakteristik gizi dari masing-masing balita (Irfiani & Rani, 2018). Dalam penelitian ini akan dilakukan pengelompokan balita sesuai dengan karakteristik gizinya memakai metode *clustering* dengan algoritma *K-Means*.

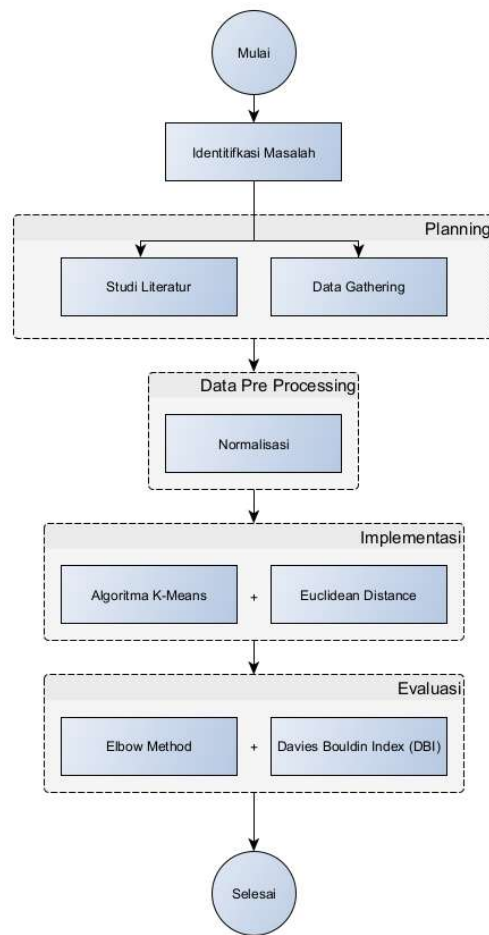
Beberapa penelitian sebelumnya yang melakukan penelitian terkait dengan pengelompokan balita sesuai dengan karakteristik gizi memakai metode *clustering* dengan algoritma *K-Means*, yaitu dilakukan oleh Irfiani dan Rani (2018) yang berjudul "Algoritma *K-Means Clustering* untuk Menentukan Nilai Gizi Balita" (Irfiani & Rani, 2018). Pada penelitiannya digunakan 2 (dua) kriteria, tinggi (TB) dan berat badan (BB), dan pembagian 5 (lima) *cluster*. Dari hasil penelitiannya disimpulkan bahwa masih

terdapat 30% balita dengan kategori balita obesitas dan 11% dengan kategori balita kurang gizi. Penelitian lainnya oleh Sulastrri, Mubarak, dan Iasha (2021) yang diberi judul “Implementasi Algoritma *Machine Learning* untuk Penentuan *Cluster* Status Gizi Balita” (Sulastrri *et al.*, 2021). Dalam penelitiannya, digunakan kriteria tinggi badan (TB), lingkaran kepala (LK), serta berat badan (BB) yang dikelompokkan menjadi 5 (lima) *cluster*. Algoritma yang digunakan adalah *K-Means*. Penelitiannya menunjukkan hasil bahwa dengan menggunakan algoritma *clustering K-Means* berhasil dilakukan pengelompokan data gizi balita, yaitu 27,59% balita masuk ke dalam kelompok gizi buruk, 3,45% balita masuk ke dalam kelompok gizi kurang, 20,69% balita masuk ke dalam kelompok gizi baik, 27,59% balita masuk ke dalam kelompok gizi lebih, dan 20,69% balita masuk ke dalam kelompok obesitas.

Algoritma *data mining clustering K-Means* banyak digunakan karena mempunyai beberapa kelebihan, diantaranya kemudahan dalam implementasi, proses yang dibutuhkan relatif cepat, dan mudah untuk diadaptasi (Nalendra, 2018). Selain itu, penelitian-penelitian terdahulu menyatakan bahwa algoritma *K-Means* mempunyai kemampuan yang baik dalam pengelompokan data baik dalam skala kecil maupun besar (Chandra *et al.*, 2021). Evaluasi dilakukan dengan melakukan pengukuran performansi algoritma *K-Means* yang digunakan dalam pengelompokan status malnutrisi balita. *RapidMiner Studio Version 9.10* merupakan *tool* yang dipakai penelitian ini. Pada tahap evaluasi, juga digunakan *Elbow method* untuk penentuan jumlah *cluster* atau nilai *k* yang paling baik atau optimal terhadap *dataset* yang dipakai pada penelitian. Sedangkan untuk mengukur kualitas hasil *clustering* menggunakan metode evaluasi *cluster internal criterion Davies Bouldin Index* (DBI). Dimana, semakin kecil hasil perhitungan nilai DBI (non-negatif ≥ 0), berarti *cluster* yang dihasilkan dari metode *clustering* tersebut semakin baik. Hasil dari penelitian diharapkan bisa menjadi dasar atau rekomendasi untuk posyandu di seluruh Indonesia dalam penentuan pengelompokan status gizi balita, sehingga penanganannya lebih tepat dan dapat membantu proses pencegahan malnutrisi pada balita.

METODE PENELITIAN

Metode penelitian yang dipakai yaitu *case studies research*. *Case studies* atau studi kasus adalah penelitian dengan fokus terhadap kasus tertentu baik secara individu atau kelompok dipergunakan sebagai bahan penelitian atau studi (Hasibuan, 2007). Pemakaian penelitian studi kasus berfokus pada proses penggalian (*mining*) dan pengumpulan data yang lebih dalam dari suatu objek penelitian untuk menjawab permasalahan yang terjadi (Hasibuan, 2007). Dapat dikatakan, penelitian studi kasus bersifat eksploratif dan deskriptif (Hasibuan, 2007). Kerangka pikir penelitian ditunjukkan pada gambar berikut:



Gambar 1. Kerangka Pikir

A. Identifikasi Masalah

Pada tahap identifikasi masalah dilakukan pendefinisian masalah dan membuat definisi masalah tersebut menjadi lebih terstruktur sebagai suatu langkah awal penelitian. Dengan kata lain, proses identifikasi masalah adalah mendefinisikan masalah penelitian. Masalah penelitian dapat bersumber dari (Tuckman, 1978):

1. Literatur yang relevan (*related theory*);
2. Deduksi atau pembuktian dari teori (*deduction from theory*);
3. Sumber lain bukan pendidikan (*non-educational sources*);
4. Pengalaman (*experiences*).

Langkah-langkah yang dapat dilakukan oleh peneliti dalam mengidentifikasi masalah, yaitu (Mukhtar, 2007):

1. Membaca banyak literatur dan kritis terhadap literatur yang dibaca;
2. Mengikuti ceramah profesional atau kuliah;
3. Mengamati kejadian atau situasi di lingkungan sesuai profesinya;
4. Berpikir terhadap segala kemungkinan permasalahan yang muncul dari materi kuliah;

5. Mengadakan penelitian kecil, kemudian mencatat hasil temuannya;
6. Mengikuti seminar hasil penelitian;
7. Memanfaatkan perpustakaan guna pencarian topik penelitian;
8. Membaca majalah atau jurnal yang ada kaitannya dengan masalah yang hendak diteliti; serta
9. Melakukan pengumpulan bahan yang ada kaitannya dengan masalah yang hendak diteliti.

B. *Planning/Perencanaan*

Pada tahapan *planning* atau perencanaan, peneliti akan melakukan proses:

1. Studi literatur

Studi literatur atau pustaka pada beberapa buku atau jurnal dilakukan oleh peneliti yang ada kaitannya dengan masalah penelitian yang dilakukan.

2. *Data gathering*

Penelitian yang dilakukan termasuk dalam kategori penelitian non-reaktif atau sering disebut sebagai penelitian *unobtrusive*. Peneliti tidak melakukan interaksi langsung pada subjek penelitian, subjek yang dimaksud adalah anak usia balita. Data yang dipergunakan merupakan data sekunder, yaitu data didapat tidak langsung atau melalui pihak lain.

C. *Data Preprocessing*

Data balita yang telah diperoleh, sebelum dipergunakan dalam proses perhitungan metode *clustering* harus diseragamkan terlebih dahulu, sehingga diperlukan proses normalisasi data, hal ini dikarenakan adanya selisih besaran angkat atau jarak yang cukup jauh dapat mempersulit proses pengelompokan. Teknik normalisasi min-max digunakan dalam penelitian ini. Transformasi linear terhadap data asli dilakukan pada teknik normalisasi min-max. Hasil nilai yang telah dinormalisasi akan berada pada *range* atau kisaran tertentu. Keunggulan teknik normalisasi min-max yaitu menyeimbangkan perbandingan nilai antar data, sebelum maupun sesudah dilakukan proses normalisasi. Data diketahui nilai maksimum dan minimumnya. Teknik normalisasi min-max banyak dipakai peneliti untuk melakukan proses normalisasi, hal tersebut dikarenakan teknik ini mudah dilakukan dan hasilnya tidak bias sehingga mempermudah proses perhitungan normalisasi data serta lebih efisien.

D. Implementasi

Implementasi metode *clustering* menggunakan algoritma *K-Means* yang dikombinasikan dengan rumus perhitungan jarak *Euclidean distance*. Nilai k atau jumlah *cluster* ditentukan sejumlah 5 (lima), yaitu gizi buruk, gizi kurang, gizi baik, gizi lebih, dan obesitas.

E. Evaluasi

Dalam penentuan jumlah *cluster* atau nilai *k* yang paling optimal atau maksimal terhadap *dataset* yang digunakan pada penelitian ini digunakan *Elbow method*. Sedangkan untuk mengukur kualitas hasil *clustering* menggunakan metode evaluasi *cluster internal criterion Davies Bouldin Index* (DBI). Dimana nilai dari DBI yang semakin kecil (non-negatif ≥ 0), maka metode *clustering* tersebut menghasilkan *cluster* yang semakin baik.

HASIL DAN PEMBAHASAN

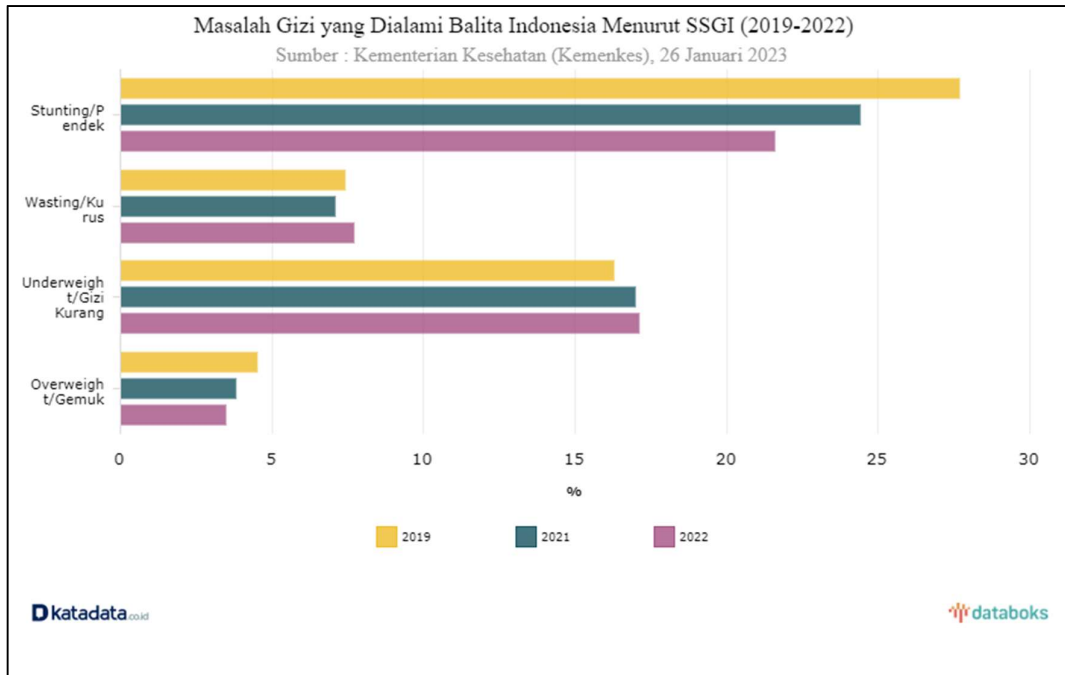
A. Identifikasi Masalah

Masalah penelitian ditemukan oleh peneliti melalui beberapa literatur yang relevan. Misalnya pada media massa *online* yang *update* tentang suatu peristiwa atau topik tertentu. Pada peringatan Hari Gizi Nasional 2023 tanggal 25 Januari 2023 lalu, Kementerian Kesehatan Republik Indonesia menyatakan bahwa hingga sekarang ini, jumlah anak kekurangan gizi di Indonesia masih tinggi (Puspapertiwi, 2023). UNICEF (Juni 2020) menyebutkan bahwa terdapat lebih dari tujuh juta anak yang berusia 5 (lima) tahun mengalami *stunting* dan lebih dari dua juta anak menderita gizi buruk (Puspapertiwi, 2023). *United Nations Children's Fund* (UNICEF), mengemukakan bahwa tindakan yang tepat harus segera diupayakan guna penurunan angka kasus malnutrisi anak di Indonesia atau jumlahnya akan mengalami peningkatan sebanyak 15% (Puspapertiwi, 2023).

Hasil yang diperoleh dari Survei Status Gizi Indonesia (SSGI) yang dilakukan Kementerian Kesehatan menunjukkan bahwa terdapat 4 (empat) permasalahan terkait gizi pada balita di Indonesia, antara lain *overweight*, *underweight*, *wasting*, dan *stunting* (Annur, 2023). Pengertian *stunting* yaitu salah satu permasalahan gizi atau nutrien yang masih menjadi perhatian Pemerintah hingga saat ini dikarenakan nilai prevalensinya yang masih cukup tinggi, yaitu mencapai angka 21,6% pada tahun 2022 (Annur, 2023). Angka tersebut melampaui atau melebihi nilai ambang batas yang ditetapkan atau ditentukan oleh *World Health Organization* (WHO) sebesar 20% (Annur, 2023). Hal ini menandakan bahwa kasus *stunting* di Indonesia masih termasuk dalam golongan kronis (Annur, 2023).

Tabel 1. Permasalahan Gizi yang Tengah Dialami oleh Balita Indonesia di Menurut SSGI (2019-2022)

No	Nama	2019 / %	2021 / %	2022 / %
1	Stunting/Pendek	27,7	24,4	21,6
2	Wasting/Kurus	7,4	7,1	7,7
3	Underweight/Gizi Kurang	16,3	17	17,1
4	Overweight/Gemuk	4,5	3,8	3,5



Gambar 2. Permasalahan Gizi yang Tengah Dialami oleh Balita Indonesia di Menurut SSGI (2019-2022)

Berdasarkan literatur-literatur tersebut, peneliti menentukan permasalahan penelitian tentang malnutrisi atau kurang gizi pada balita.

B. *Planning/Perencanaan*

Pada tahapan *planning* atau perencanaan, peneliti melakukan proses:

1. Studi literatur

Dari proses studi literatur atau pustaka yang dilakukan, peneliti menemukan beberapa literatur terkait kriteria yang digunakan dalam penelitian, yaitu:

Tabel 1. Studi Literatur untuk Penentuan Kriteria

No.	Judul	Tahun	Kriteria	Hasil
1.	"Penerapan Metode <i>K-Means Clustering</i> untuk Menentukan Status Gizi dan Gizi Buruk pada Balita (Studi Kasus Kabupaten Rokan Hulu)" (Dona & Rifqi, 2022)	2022	1) Tinggi badan 2) Berat badan	1) Dihilangkan 2 <i>cluster</i> , gizi buruk dan gizi baik 2) Algoritma <i>K-Means</i> menghasilkan pengetahuan tentang data balita 3) Algoritma <i>K-Means</i> menghasilkan pengetahuan untuk pengambilan keputusan yang lebih optimal

No.	Judul	Tahun	Kriteria	Hasil
2.	"Implementasi Algoritma <i>Machine Learning</i> untuk Penentuan <i>Cluster</i> Status Gizi Balita" (Sulastri et al., 2021)	2021	1) Berat badan 2) Lingkar kepala 3) Tinggi badan	1) Dengan algoritma <i>K-Means</i> berhasil melakukan pengelompokan data gizi balita menjadi 5 kelompok, yaitu obesitas, gizi lebih, gizi baik, gizi kurang, dan gizi buruk 2) Penentuan kelompok status gizi balita dapat membantu pihak Posyandu dalam menangani dan mencegah malnutrisi pada balita
3.	"Model <i>Clustering</i> Menggunakan Algoritma <i>K-Means</i> dalam Menentukan Kriteria Kondisi Gizi Balita dan Anak" (Zannah et al., 2022)	2022	1) Berat badan 2) Tinggi badan	<i>Clustering</i> status gizi balita dengan algoritma <i>K-Means</i> secara umum dapat digunakan sebagai upaya mencegah dan menanggulangi kasus gizi buruk dan obesitas sejak dini bagi para kader Posyandu
4.	"Implementasi Algoritma <i>K-Means Clustering</i> Status Gizi Balita" (Octaviyani et al., 2022)	2022	1) Berat badan 2) Tinggi badan	1) Berdasarkan pengujian, <i>clustering</i> terhadap pengelompokan status gizi balita dipandang cukup baik 2) Visualisasi data dalam bentuk <i>scatter</i> 3D bisa memperlihatkan tingkat akurasi hasil <i>clustering</i>
5.	"Penerapan <i>K-Means</i> untuk <i>Clustering</i> Kondisi Gizi Balita pada Posyandu" (Rahmat et al., 2023)	2023	1) Berat badan 2) Lingkar kepala 3) Tinggi badan	1) Perhitungan manual algoritma <i>K-Means</i> untuk pengelompokan kondisi gizi balita mempunyai hasil yang sama dengan perhitungan menggunakan <i>tools</i> WEKA 2) Diperoleh 3 <i>cluster</i> pengelompokan kondisi gizi balita, yaitu <i>cluster</i> gizi buruk, <i>cluster</i> gizi baik, dan <i>cluster</i> obesitas

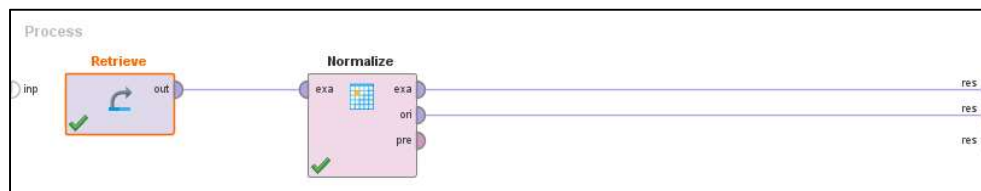
Berdasarkan literatur di atas, peneliti menggunakan kriteria tinggi dan berat badan (BB) dalam penelitian.

2. *Data gathering*

Sumber data sekunder yang dipergunakan di dalam penelitian yaitu rekap data dari hasil penilaian standar antropometri gizi balita di kabupaten Wonosobo. Sampel yang digunakan sebanyak 100 data.

C. Data Preprocessing

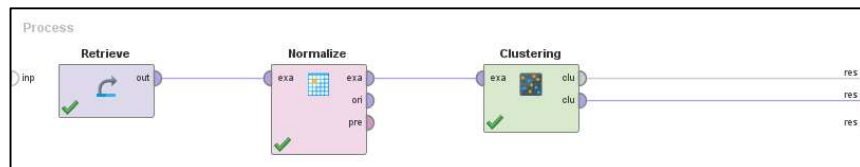
Pada penelitian ini, *data preprocessing* atau pengolahan data awal dilakukan dengan teknik atau metode normalisasi min-max/max-min dalam *tools RapidMiner*, normalisasi min-max/max-min menggunakan *method range-transformation*. Proses normalisasi menggunakan *tools RapidMiner* ditunjukkan sebagai berikut:



Gambar 3. Proses Normalisasi pada *RapidMiner*

D. Implementasi

Setelah data balita tersebut dinormalisasi, selanjutnya dilakukan proses perhitungan memakai metode *clustering* dengan algoritma *K-Means* dengan rumus jarak *Euclidean distance*. Adapun proses *clustering* menggunakan algoritma *K-Means* dan *Euclidean distance* pada *tool RapidMiner* ditunjukkan sebagai berikut:



Gambar 4. *Clustering* dengan Algoritma *K-Means* pada *RapidMiner*

Centroid tiap *cluster* dapat dilihat pada tabel berikut:

Tabel 2. *Centroid* Tiap *Cluster* (k = 5)

Attribute	cluster_0 (1)	cluster_1 (2)	cluster_2 (3)	cluster_3 (4)	cluster_4 (5)
Berat Badan (kg)	0,481	0,589	0,119	0,331	0,745
Tinggi Badan (cm)	0,594	0,873	0,114	0,348	0,789

Jumlah balita pada tiap *cluster* adalah:

Tabel 3. Jumlah Balita pada Tiap *Cluster* (k = 5)

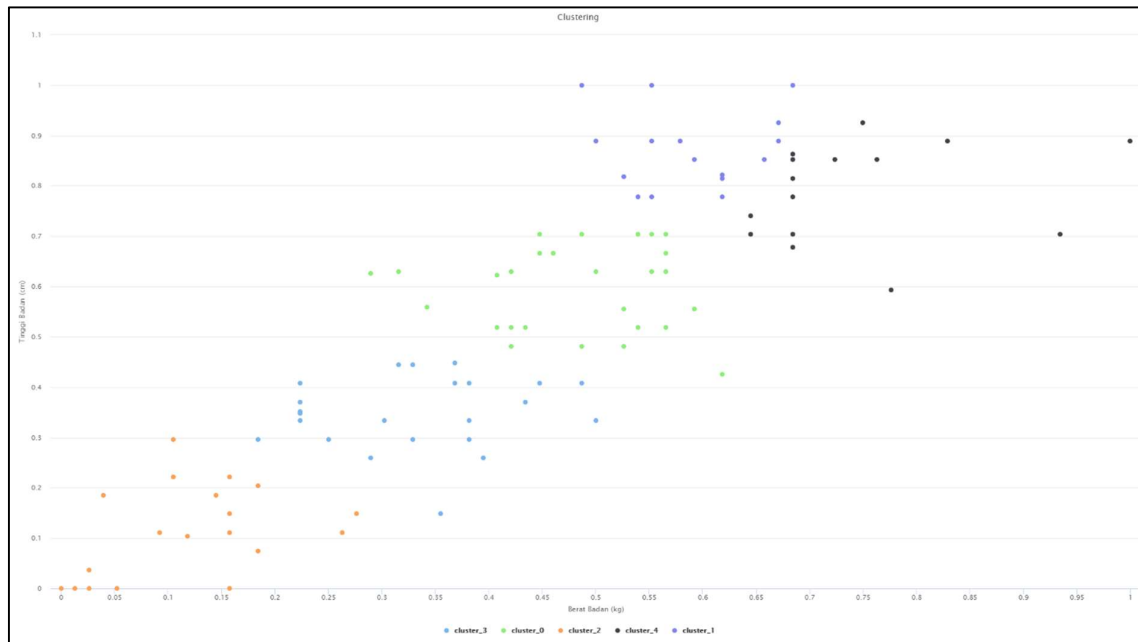
Cluster			Jumlah Balita
0	1	Gizi baik	27
1	2	Gizi lebih	16
2	3	Gizi buruk	19
3	4	Gizi kurang	23
4	5	Obesitas	15
Total			100

Jika diurutkan sebagai berikut:

Tabel 4. Urutan *Cluster* (k = 5)

Cluster	Jumlah Balita
Gizi buruk	19
Gizi kurang	23
Gizi baik	27
Gizi lebih	16
Obesitas	15
Total	100

Visualisasi hasil *clustering* diperlihatkan gambar berikut:

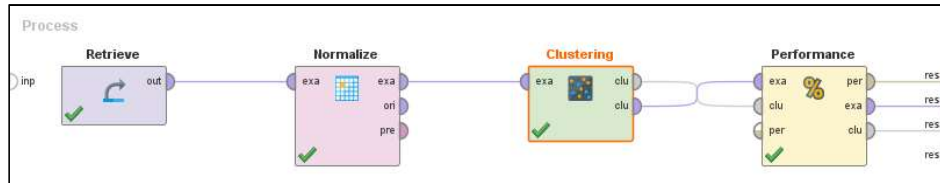


Gambar 5. Visualisasi Hasil *Clustering* (k = 5)

Nilai DBI yang pada *RapidMiner* adalah 0,914.

E. Evaluasi

Untuk mengetahui penentuan nilai k atau jumlah *cluster* sudah optimal atau belum, maka pada *RapidMiner* dilakukan pengukuran kualitas *cluster* menggunakan *cluster distance performance*. Prosesnya diperlihatkan gambar berikut:



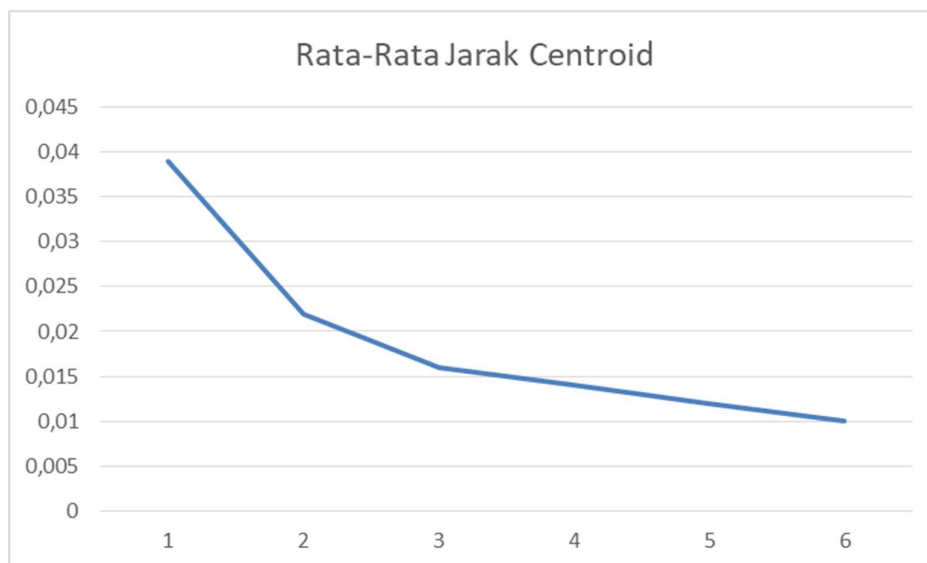
Gambar 6. Proses Pengukuran Kualitas *Cluster* pada *RapidMiner*

Dari beberapa kali percobaan dengan cara mengubah jumlah *cluster* atau nilai k, maka diperoleh nilai rata-rata jarak *centroid* pada tiap *cluster* sebagai berikut:

Tabel 5. Rata-Rata Jarak *Centroid* Berdasarkan Jumlah *Cluster*

Jumlah <i>Cluster</i> (k)	Rata-Rata Jarak <i>Centroid</i>
2	0,039
3	0,022
4	0,016
5	0,014
6	0,012
7	0,010

Dengan *Elbow method*, akan menentukan jumlah *cluster* atau nilai k yang paling bagus, yaitu dengan memvisualisasikan data rata-rata jarak *centroid* pada kurva x dan y sebagai berikut:



Gambar 7. Visualisasi Kurva *Elbow*

Dari gambar kurva di atas, tentukan kurva yang bentuknya paling siku, maka diperoleh *elbow* (siku) pada nilai $k = 3$. Sehingga, nilai k atau jumlah *cluster* yang paling optimal untuk *dataset* balita yang dipergunakan adalah $k = 3$.

Tabel 6. *Centroid* Tiap *Cluster* ($k = 3$)

Attribute	cluster_0 (1)	cluster_1 (2)	cluster_2 (3)
Berat Badan (kg)	0,183	0,652	0,448
Tinggi Badan (cm)	0,192	0,817	0,523

Tabel 7. Jumlah Balita pada Tiap *Cluster* ($k = 3$)

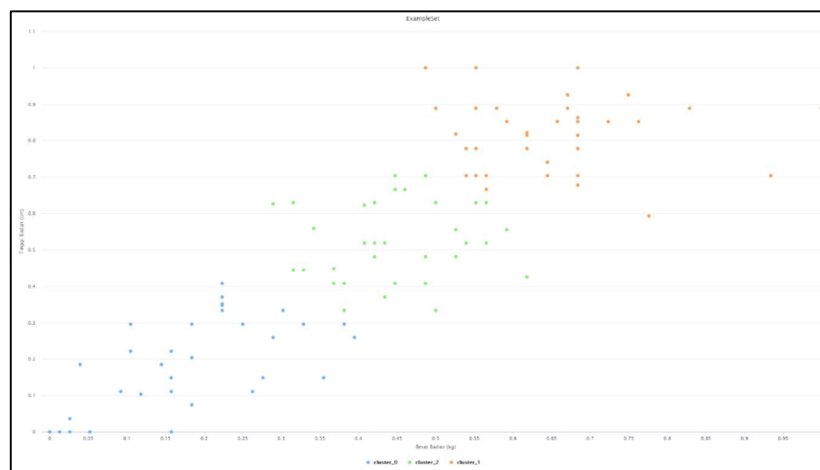
Cluster			Jumlah Balita
0	1	Gizi buruk	32
1	2	Obesitas	35
2	3	Gizi baik	33
Total			100

Jika diurutkan sebagai berikut:

Tabel 8. Urutan *Cluster* ($k = 3$)

Cluster	Jumlah Balita
Gizi buruk	32
Gizi baik	33
Obesitas	35
Total	100

Visualisasi hasil *clustering* dengan nilai $k = 3$ diperlihatkan gambar berikut:



Gambar 8. Visualisasi Hasil *Clustering* ($k = 3$)

Nilai DBI dari *RapidMiner* adalah 0,022.

KESIMPULAN

Sesuai dengan hasil penelitian dan pembahasan, diperoleh kesimpulan nilai k atau jumlah *cluster* yang optimal untuk *dataset* balita yang dipakai penelitian adalah $k = 3$, dimana pembagian *cluster* atau status gizinya adalah *cluster* gizi buruk sebanyak 32 balita (32%), gizi baik sebanyak 33 balita (33%), dan obesitas sebanyak 35 balita (35%).

Berdasar hasil dan pembahasan tersebut, dapat diberikan beberapa saran bagi penelitian yang akan selanjutnya, yaitu:

1. Penggunaan *dataset* yang berbeda dengan tetap menggunakan algoritma *clustering*.
2. Untuk penelitian selanjutnya dapat ditambahkan kriteria atau atribut lainnya.
3. Penggunaan metode *clustering* dengan kombinasi pengukuran jarak yang berbeda.

UCAPAN TERIMA KASIH

Peneliti mengucapkan terima kasih kepada para donatur pendanaan penelitian serta berbagai pihak yang telah membantu terlaksananya penelitian ini.

DAFTAR PUSTAKA

- Andayani, S. (2016). *Pembentukan Cluster dalam Knowledge Discovery in Database dengan Algoritma K-Means*. Seminar Nasional Matematika Dan Pendidikan Matematika 2007, 1–10.
- Annur, C. M. (2023). *Selain Stunting, Ini Deretan Masalah Gizi yang Kerap Dialami Balita di Indonesia*. 3 Februari 2023. <https://databoks.katadata.co.id/datapublish/2023/02/03/selain-stunting-ini-deretan-masalah-gizi-yang-kerap-dialami-balita-di-indonesia>
- Apriliyaningsih, I. D., & Istiawan, D. (2017). *Penerapan Seleksi Atribut Berdasarkan Koefisien Variansi dan Korelasi untuk Inisialisasi Pusat Awal Klaster pada Algoritma K-Means dalam Pemetaan E-Government Tahun 2016*. Jurnal University Research Colloquium 2017 Universitas Muhammadiyah Magelang, 245–250.
- Aryani, L. D., & Riyandry, M. A. (2019). *Vitamin D sebagai Terapi Potensial Anak Gizi Buruk*. Jurnal Penelitian Perawat Profesional, 1(1), 61–70.
- Atika, P. D., & Priatna, W. (2020). *Modul Data Mining*. Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
- Berkhin, P. (2003). *Survey of Clustering Data Mining Techniques*. Accrue Software.
- Chandra, M. D., Irawan, E., Saragih, I. S., Windarto, A. P., & Suhendro, D. (2021). *Penerapan Algoritma K-Means dalam Mengelompokkan Balita yang Mengalami Gizi Buruk Menurut Provinsi*. BIOS: Jurnal Teknologi Informasi Dan Rekayasa Komputer, 2(1), 30–38.
- Dennis Aprilla C, Donny Aji Baskoro, Ambarwati, L., & Wicaksana, I. W. S. (2013). *Belajar Data Mining dengan Rapid Miner (Remi Sanjaya (ed.))*. Open Content Model.
- Dhuhita, W. M. P. (2015). *Clustering Menggunakan Metode K-Means untuk Menentukan Status Gizi Balita*. Jurnal Informatika, 15(2), 160–174.
- Dinas Kesehatan Kabupaten Lima Puluh Kota. (2015). *Malnutrisi, RI bisa Kehilangan Potensi Ekonomi 3 Persen PDB*. <https://dinkes.limapuluhkotakab.go.id>
- Dona, & Rifqi, M. (2022). *Penerapan Metode K-Means Clustering untuk Menentukan Status Gizi dan Gizi Buruk pada Balita (Studi Kasus Kabupaten Rokan Hulu)*. RABIT: Jurnal Teknologi Dan Sistem Informasi Univrab, 7(2), 179–191.
- Everitt, B. (1980). *Cluster Analysis*. Gower Pub Co.

- Fitriyanto, R. E., & Mahfudz, S. (2020). *Management of Severe Malnutrition of Under Five Years Old Patients in RSUD Wonosari*. AJIE - Asian Journal of Innovation and Entrepreneurship, 5(1), 20–26.
- Garcia-Molina, H., Ullman, J. D., & Widom, J. D. (2001). *Database Systems the Complete Book (Pearson Education Inc. (ed.); Second Edi)*. Pearson Prentice Hall.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining, Concepts and Techniques (Third Edit)*. Morgan Kaufmann Publishers.
- Harsemadi, I. G., & Sudarma, I. M. (2017). *Penggolongan Musik Terhadap Suasana Hati Menggunakan Metode K-Means*. Jurnal Konferensi Nasional Sistem Dan Informatika 2017, 49–54.
- Hartanti, N. T. (2020). *Metode Elbow dan K-Means Guna Mengukur Kesiapan Siswa SMK dalam Ujian Nasional*. Jurnal Nasional Teknologi Dan Sistem Informasi, 6(2), 82–89.
- Hasibuan, Z. A. (2007). *Metodologi Penelitian pada Bidang Ilmu Komputer dan Teknologi Informasi, Konsep, Teknik dan Aplikasi*. In *Desain Penelitian* (p. 81). Fakultas Ilmu Komputer Universitas Indonesia.
- Irfiani, E., & Rani, S. S. (2018). *Algoritma K-Means Clustering untuk Menentukan Nilai Gizi Balita*. JUSTIN: Jurnal Sistem Dan Teknologi Informasi, 6(4), 161–168.
- John, C. C., Black, M. M., & Nelson, C. A. (2017). *Neurodevelopment: The Impact of Nutrition and Inflammation during Early to Middle Childhood in Low-Resource Settings*. Pediatrics, 139(1), 59–71.
- Kementrian Kesehatan Republik Indonesia. (2019). *Kemenkes Tingkatkan Status Gizi Masyarakat*. <https://www.kemkes.go.id/article/view/19081600004/kemenkes-tingkatkan-status-gizi-masyarakat.html>
- Lailani, F. K., Yuliana, & Yulastri, A. (2022). *Masalah Terkait Malnutrisi: Penyebab, Akibat, dan Penanggulangannya*. JGK: Jurnal Gizi Dan Kesehatan, 2(2), 129–138.
- Larose, D. T. (2006). *Data Mining Methods and Models*. John Wiley & Sons, Inc.
- Martiana, E. (2003). *Data Preprocessing*. In *Data Preprocessing* (pp. 1–13). Soft Computation Research Group, EEPIS - ITS.
- Mukhtar, M. (2007). *Pedoman Penulisan Karya Ilmiah*. STAIN.
- Nalendra, A. K. (2018). *Pengukuran Keakuratan Metode K-Means untuk Menentukan Status Gizi Balita*. Jurnal Ekonomi Dan Teknik Informatika, 6(2), 48–54.
- Ning Tan Pang, Steinbach, M., & Vipin Kumar. (2004). *Data Mining Cluster Analysis: Basic Concepts and Algorithms* (pp. 1–52). <https://doi.org/10.21803/penamer.6.10.137>
- Octaviyani, N. R., Mayasari, R., & Susilawati. (2022). *Implementasi Algoritma K-Means Clustering Status Gizi Balita*. Jurnal Ilmiah Wahana Pendidikan, 8(13), 370–381.
- Papotot, G. S., Rompies, R., & Salendu, P. M. (2021). *Pengaruh Kekurangan Nutrisi terhadap Perkembangan Sistem Saraf Anak*. Jurnal Biomedik, 13(3), 266–273.
- Peraturan Menteri Kesehatan Republik Indonesia Nomor 2 Tahun 2020 tentang Standar Antropometri Anak, (2020) (testimony of Menteri Kesehatan Republik Indonesia). http://hukor.kemkes.go.id/uploads/produk_hukum/PMK_No__2_Th_2020_ttg_Standar_Antropometri_Anak.pdf
- Perdana, H. M., Darmawansyah, & Faradillah, A. (2020). *Gambaran Faktor Risiko Malnutrisi pada Anak Balita di Wilayah Kecamatan Tamalanrea Kota Makassar Tahun 2019*. UMI Medical Journal, 5(1), 50–56.
- Prasetyo, E. (2012). *Data Mining Konsep dan Aplikasi Menggunakan Matlab*. Andi Yogyakarta.
- Puspapertiwi, E. R. (2023). *Hari Gizi Nasional 2023, Kenali Ciri dan Akibat Anak Kekurangan Gizi*. 25 Januari 2023. <https://www.kompas.com/tren/read/2023/01/25/161500865/hari-gizi-nasional-2023-kenali-ciri-dan-akibat-anak-kekurangan-gizi?page=all>

- Rahmat, C. A., Permatasari, H., Rasywir, E., & Pratama, Y. (2023). Penerapan K-Means untuk Clustering Kondisi Gizi Balita pada Posyandu. *Jurnal Media Informatika Budidarma*, 7(1), 207–213.
- Riandari, F., & Simangunsong, A. (2019). *Penerapan Algoritma C4.5 untuk Mengukur Tingkat Kepuasan Mahasiswa*. *Jurnal Mantik Penusa*, 3(2), 1–7.
- Riskesdas. (2010). *Riset Kesehatan Dasar Tahun 2010*.
- Saranya, C., & Manikandan, G. (2013). *A Study on Normalization Techniques for Privacy Preserving Data Mining*. *International Journal of Engineering and Technology (IJET)*, 5(3), 2701–2704.
- Sulastri, H., & Gufroni, A. I. (2017). *Penerapan Data Mining dalam Pengelompokan Penderita Thalassaemia*. *Jurnal Teknologi Dan Sistem Informasi*, 3(2), 299–305.
- Sulastri, H., Mubarak, H., & Iasha, S. S. (2021). *Implementasi Algoritma Machine Learning untuk Penentuan Cluster Status Gizi Balita*. *JURTI: Jurnal Rekayasa Teknologi Informasi*, 5(2), 184–191.
- Sulistiyawati, A. (2019). *Faktor Risiko Kejadian Gizi Buruk pada Balita di Dusun Teruman Bantul*. *Jurnal Kesehatan Madani Medika*, 10(1), 13–19.
- Syaripul, N. A., & Bachtiar, A. M. (2016). *Visualisasi Data Interaktif Data Terbuka Pemerintah Provinsi DKI Jakarta: Topik Ekonomi dan Keuangan Daerah*. *Jurnal Sistem Informasi (Journal of Information Systems)*, 12(2), 82–89.
- Tuckman, B. W. (1978). *Conducting Educational Research*. Harcourt Brace Jovanovich.
- Uppada, S. K. (2014). *Centroid Based Clustering Algorithms - A Clarion Study*. *International Journal of Computer Science and Information Technologies*, 5(6), 7309–7313. <http://ijcsit.com/docs/Volume 5/vol5issue06/ijcsit2014050688.pdf>
- Wahyuni, I., Auliya, Y. A., Rahmi, A., & Mahmudy, W. F. (2016). *Clustering Nasabah Bank Berdasarkan Tingkat Likuiditas Menggunakan Clustering Nasabah Bank Berdasarkan Tingkat Likuiditas Menggunakan Hybrid Particle Swarm Optimization dengan K-Means*. *Jurnal Ilmiah Teknologi Dan Informasi Asia (JITIKA)*, 10, 24–33.
- Wijanarko, T. S. (2017). *Algoritma K-Means dan K-Medoids Berbasis Measure Distance Metode untuk Pengelompokan Status Desa*. Universitas Dian Nuswantoro Semarang.
- Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data Mining Third Edition*. Elsevier Inc.
- Xu, R., & Wunsch, D. (2008). *Clustering*. Wiley-IEEE Press.
- Zannah, K., Sumarno, Nasution, Z. M., Parlina, I., & Sari, I. P. (2022). *Model Clustering Menggunakan Algoritma K-Means dalam Menentukan Kriteria Kondisi Gizi Balita dan Anak*. *JUTEKIN: Jurnal Teknik Informatika*, 10(1), 25–34.