

ANALISA PERBANDINGAN ALGORITMA *CLUSTERING* UNTUK PEMETAAN STATUS GIZI BALITA DI PUSKESMAS PASIR JAYA

Nenden Siti Fatonah¹, Trisnansita Kintan Pancarani²

¹Fakultas Ilmu Komputer, Universitas Esa Unggul

²Teknik Informatika, Ilmu Komputer, Universitas Mercu Buana

E-mail: nenden.siti@esaunggul.ac.id^{1*}

ABSTRACT

The problem of nutritional status in toddlers is still a major problem that needs attention, one of which is malnutrition. Malnutrition problems are directly caused by inadequate food intake, inadequate family economic factors, and caused by underlying diseases such as tuberculosis, heart disease or birth defects. The purpose of this study is to implement the clustering method with the K-Means and Fuzzy C-Means algorithms which aims to evaluate the nutritional status of children under five in general so that it can be used as a basis for early prevention for health workers at the puskesmas to overcome malnutrition. This study uses 4 parameters, namely height, weight, gender and age and uses a silhouette index validation calculation to measure the resulting cluster cohesiveness. From the result of the cluster analysis, the K-Means algorithm gets a validation result of 0.79 with 4 clusters and the Fuzzy C-Means algorithm gets a validation result of 0.78 with 4 clusters. From the two comparison algorithm, the K-Means algorithm gets the best validation because get the highest validation compared to the Fuzzy C-Means algorithm.

Keywords: *K-means, Fuzzy C-Means, Clustering, Nutrition of children under five*

ABSTRAK

Masalah status gizi pada balita masih menjadi masalah utama yang perlu diperhatikan, salah satunya adalah Gizi buruk. Masalah Gizi buruk secara langsung di sebabkan oleh asupan makanan yang gizinya tidak tercukupi, factor ekonomi keluarga yang tidak memadai, dan disebabkan oleh penyakit yang mendasari seperti TBC, Jantung atau kelainan cacat pada saat lahir. Tujuan dari penelitian ini yaitu, Mengimplementasikan metode klasterisasi dengan algoritma K-Means dan Fuzzy C-Means yang bertujuan untuk mengevaluasi status gizi balita secara umum sehingga dapat dijadikan sebagai dasar pencegahan dini bagi petugas kesehatan di puskesmas untuk menanggulangi gizi buruk. Penelitian ini menggunakan empat parameter yaitu tinggi badan, Berat badan, jenis kelamin dan umur dan menggunakan perhitungan validasi *silhouette index* untuk mengukur kekompakan cluster yang di hasilkan. Dari hasil dari analisis cluster, algoritma K-Means mendapatkan hasil validasi sebesar 0.79 dengan cluster sebanyak 4 dan algoritma Fuzzy C-Means mendapatkan validasi hasil validasi sebesar 0.78 dengan cluster sebanyak 4. Dari kedua perbandingan algoritma tersebut, algoritma K-Means mendapatkan validasi terbaik karena mendapatkan validasi tertinggi dibandingkan dengan Algoritma Fuzzy CMeans.

Kata Kunci: K-Means, Fuzzy C-Means, Clustering, Gizi balita

1. Pendahuluan

Nilai gizi ialah satu dari sekian banyak indikator kesehatan anak umur lima tahun merupakan masa penting saat anak memerlukan nutrisi yang cukup sebagai

penunjang pertumbuhan fisiknya. Nilai gizi balita bisa dihitung dengan menggunakan indikator tinggi badan berlandaskan umur (TB/U), berat badan belandaskan umur (BB/TB), serta berat badan berdasarkan

tinggi badan (BB/U) dengan memasukkan nilai tinggi badan, berat badan dan umur. Cara pengukuran tubuh manusia untuk mengetahui nilai gizi anak adalah melalui antropometri. Cara antropometri umumnya didasarkan pada tinggi badan berat badan dan umur. Antropometri biasa digunakan untuk mengukur nilai gizi balita karena aman dan sederhana dalam tahap pengukuran serta tidak memerlukan tenaga ahli [1].

Rendahnya pendidikan dan perekonomian seringkali menjadi sebab orang tua tidak bisa mengawasi asupan makanan anak-anak mereka. Tidak sehatnya lingkungan sekitar, Rendahnya kesadaran masyarakat akan lingkungan yang kurang bersih di berbagai kota di Indonesia sebagai penyebab balita dan anak-anak rentan terkena berbagai macam penyakit. Balita yang rentan terkena penyakit bisa dikatakan sebagai balita yang mempunyai gizi yang kurang jika dibandingkan oleh balita yang jarang terkena penyakit [2].

Terdapat 2 faktor yang berdampak pada nilai gizi antara lain faktor langsung dan faktor tidak langsung. Faktor langsungnya adalah penyakit infeksi, yaitu jenis makanan yang biasa dimakan baik kualitas maupun kuantitasnya. Faktor tidak langsungnya meliputi: faktor sosial ekonomi, pendidikan, pendapatan, pola asuh yang tidak memadai, kekurangan pangan, dan perilaku terkait pelayanan kesehatan. Kepedulian aparat desa (dalam kasus ini yaitu petugas puskesmas) dan orang tua sangat penting untuk mengontrol pola asupan makanan anak usia dini. Malnutrisi pada bayi dibawah umur 5 tahun tidak terjadi secara tiba-tiba layaknya penyakit biasanya. Ciri-ciri kekurangan berat badan dan pertumbuhan terhambat bisa menjadi indikator awal kekurangan gizi pada anak-anak [3]–[5].

Penyampaian informasi nilai gizi balita yang dapat diberikan secara cepat dan akurat akan membantu mencegah terjadinya kekurangan gizi balita. Penerapan data mining yang menggunakan metode statistik, artificial intelligence (AI) dan pembelajaran mesin dari berbagai database besar selama

pemrosesan, untuk mengekstrak dan mengidentifikasi informasi, merupakan pendekatan yang perlu pada permasalahan ini [6],[7]. Data mining digambarkan sebagai tahapan dalam menemukan koneksi atau motif dari ratusan sampai ribuan field dalam database relasional yang besar [8].

Salah satu metode dalam data mining adalah analisa cluster yang merupakan metode multi-dimensi pengelompokan objek dengan karakteristik yang sama [9], [10]. Proses *clustering* bertujuan untuk meminimalkan munculnya fungsi tujuan yang sudah ditetapkan pada saat proses *clustering*, kebanyakan dikenakan untuk meminimalkan perubahan dalam sebuah *cluster* dan memaksimalkan perubahan antar *cluster* [11], [12]. Dengan memakai *clustering*, kita bisa mengklasifikasikan daerah yang padat, menemukan pola distribusi umum, dan menemukan hubungan menarik diantara atribut data [13], [14].

Metode clustering dibedakan menjadi dua yaitu *partitioning* dan *hierarchical* [15]. Yang termasuk dalam *partitioning* diantaranya adalah *k-means clustering* dan *fuzzy c-means clustering*. Beberapa aplikasi dari metode *clustering* adalah untuk analisa segmentasi pelanggan dengan menggunakan metode Kmeans [16], pengelompokan jumlah daerah yang terjangkit demam berdarah dengue (DBD) berdasarkan provinsi [17], clustering data capaian Ujian Nasional Sekolah Menengah Pertama pada tahun 2018/2019 [18], segmentasi perilaku pembelian pelanggan berdasarkan model RFM [19].

Pada penelitian ini dilakukan pemetaan status gizi balita di Puskesmas Pasir Jaya dan metode yang akan digunakan yaitu metode *clustering* K-Means dan Fuzzy C-Means. Klasterisasi ini bertujuan untuk mengelompokkan nilai gizi balita agar bisa digunakan sebagai dasar pencegahan untuk para petugas kesehatan di puskesmas untuk menanggulangi gizi buruk dan Membantu pihak puskesmas dalam memetakan dan mengidentifikasi kelompok balita berdasarkan status gizi nya.

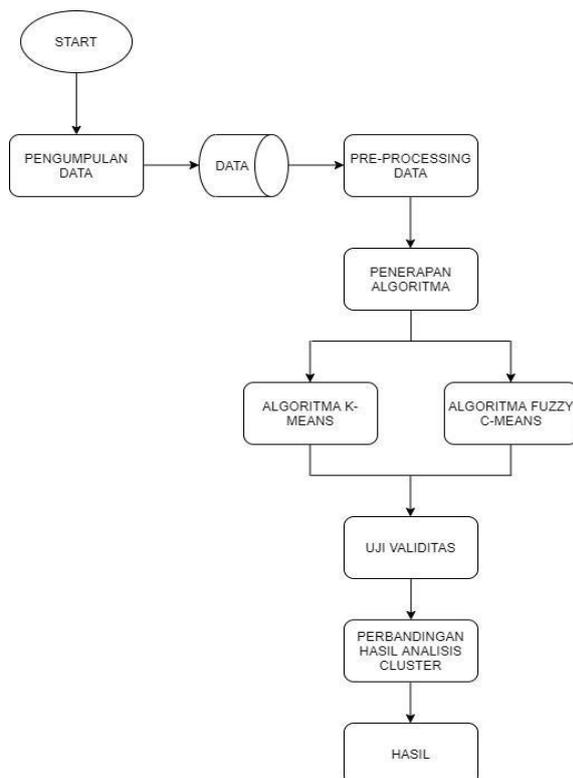
2. Metode

2.1. Data Penelitian

Penelitian ini menggunakan data kuantitatif (data sekunder) yaitu data status gizi anak yang merupakan data statistik dari Puskesmas Pasir Jaya, Cikupa pada Tahun 2019 sebanyak 1562 data. Data ditampilkan berdasarkan tipe data dari setiap parameter yang ada. Data awal yang diperoleh masih memerlukan *preprocessing* untuk selanjutnya diproses menggunakan dua algoritma *clustering*. Dalam tahap *preprocessing* dilakukan pemilihan parameter yang cocok dalam ruang lingkup penelitian. Total parameter yang dipakai untuk penelitian ini sebanyak 4 parameter dari 17 parameter yang ada.

2.2. Tahap Penelitian

Pada penelitian ini terdiri dari beberapa tahapan penelitian antara lain; pengumpulan data, data, *preprocessing* data, penerapan algoritma k-means dan fuzzy c-means, validasi hasil, analisis cluster hasil, perbandingan hasil. Tahapan penelitian dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

A. Tahap Pengumpulan Data

Pada pengumpulan data ini ialah tahap utama pada saat akan melakukan penelitian, karena pada dasarnya penelitian bertujuan untuk memperoleh data. Berikut ini adalah tahap pengumpulan data yang dilakukan:

- Pengamatan

Pengamatan atau observasi ialah salah satu cara pengumpulan data yang efektif. Untuk memperoleh data yang berhubungan dengan status gizi balita, penulis melakukan observasi langsung di Puskesmas Pasir Jaya agar dapat mengetahui secara langsung apa penyebab dan akibat yang akan ada hubungannya dengan penelitian yang ingin dilaksanakan.

- Wawancara

Wawancara atau biasa disebut juga dengan *interview* adalah cara mengumpulkan data sekaligus informasi dengan langsung bertanya kepada pihak yang ada hubungannya dengan penelitian ini. Dalam penelitian ini, penulis melakukan wawancara langsung kepada petugas kesehatan yaitu bidan yang sering menangani kasus masalah gizi pada anak.

- Studi Literatur

Studi literatur disini merupakan pengumpulan bahan-bahan atau materi-materi yang merujuk pada jurnal-jurnal yang sudah di terbitkan dan juga sudah terindeks. Studi literatur ini dilakukan untuk mendapatkan data atau referensi terkait dengan judul yang akan dilakukan pada penelitian ini.

Pada tahap ini data diperoleh langsung dari Puskesmas Pasir Jaya. Data yang digunakan sebanyak 1562. Algoritma Fuzzy C-Means dan K-Means digunakan untuk menyeleksi dan mengolah data yang telah terverifikasi oleh Puskesmas Pasir Jaya. Data bisa dilihat pada Gambar 2.

JENIS KELAMIN	Berat Badan	Tinggi Badan	Umur	
0	2	15	104	4
1	1	16	106	4
2	1	12	90	3
3	1	12	92	3
4	2	11	85	2

Gambar 2. Data

B. Data

Setelah melakukan pengumpulan data, data akan diolah dengan menggunakan cara dimasukkan kedalam excel supaya lebih mudah dalam proses pengolahan datanya.

C. Preprocessing

Pada tahap ini data yang telah didapatkan akan dilakukan *preprocessing* menggunakan *MinMax Normalization*. *MinMax Normalization* ialah metode normalisasi dengan menggunakan transformasi dengan skala 0 hingga 1 [20], [21]. Hasil *preprocessing* bisa diamati pada Gambar 3.

JENIS KELAMIN	Berat Badan	Tinggi Badan	Umur	
0	0.269805	0.308327	0.635448	0.903079
1	0.236110	0.349038	0.618135	0.063432
2	0.491990	0.324107	0.466789	0.091510
3	0.461780	0.287404	0.459603	0.103185
4	0.571275	0.338405	0.281251	0.948340
...
1557	0.966305	0.396843	0.368145	0.081257
1558	0.683699	0.307657	0.522431	0.916175
1559	0.634903	0.332230	0.510453	0.911992
1560	0.936095	0.360140	0.360959	0.092932
1561	0.254700	0.289975	0.631855	0.908916

1562 rows x 4 columns

Gambar 3. Pre-processing Data

D. Penerapan Algoritma K-Means.

Prinsip kerja K-Means ialah membagi objek-objek yang ada ke dalam kelompok-kelompok atau biasa disebut segmen, sehingga objek-objek dalam setiap kelompok lebih mirip daripada objek-objek dalam kelompok yang berbeda. K-Means ialah salah

satu metode data mining dimana cara kerja pemodelannya tanpa supervisi dan salah satu metode pengelompokan data secara partisi. K-Means menggunakan ukuran jarak Euclidean dan secara iteratif menentukan setiap record dari cluster asli [18]. Tahap-tahap algoritma K-Means adalah [22]:

1. Tentukan jumlah k (*cluster*)
2. Pilih titik secara acak dimana titik akan menjadi pusat (*centroid*) dari masing-masing *cluster*
3. Hitunglah *distance* atau jarak dan alokasikan masing-masing data ke *centroid* terdekat. Persamaan *Euclidian Distance*:

$$d(x_i, \mu_j) = \sqrt{\sum (x_i - \mu_j)^2} \quad (1)$$

Dimana x_i : kriteria data dan μ_j : *centroid* pada *cluster* ke- j .

4. Memperbaharui nilai *centroid*. Nilai *centroid* baru di dapat dari rata-rata *cluster* yang berkaitan dengan menggunakan rumus:

$$\mu_j(t+1) = \frac{1}{N_{sj}} \sum_{j \in S_j} x_j \quad (2)$$

Dimana $\mu_j(t+1)$: *centroid* baru pada iterasi ke $(t+1)$ dan N_{sj} : banyak data pada *cluster* S_j .

5. Lakukan secara berulang dari tahap 2 sampai 5, sehingga anggota tiap *cluster* tidak ada yang berganti.
6. Bila langkah ke-6 sudah terlaksana, lalu nilai pusat *cluster* (μ_j) pada perulangan terakhir hendak dipakai sebagai parameter untuk menentukan klasifikasi data.

E. Penerapan Algoritma Fuzzy C-Means.

Fuzzy C-Means yaitu teknik klasterisasi yang kerap diimplementasikan dalam metode *clustering*. Fuzzy C-Means menggunakan penggabungan Fuzzy, dimana setiap data bisa dipunyai oleh sebagian *cluster*, dan setiap *cluster* memiliki derajat keanggotaan yang berbeda. Algoritma Fuzzy C-Means adalah algoritma berulang yang mengiterasi proses *clustering* data. Penggunaan algoritma fuzzy C-Means bertujuan untuk mencari pusat titik cluster,

kemudian digunakan sebagai pencarian data yang masuk kedalam *cluster* [23]. Berikut ini ialah konsep Fuzzy C-means [24]:

- Titik pusat *cluster* yang ditentukan merupakan tanda letak rata-rata setiap *cluster*, dengan kedudukan pertama tidak konsisten atau kurang benar.
- Setiap data mempunyai level keanggotaan untuk tiap-tiap *cluster*.
- Setiap iterasi yang didasari pada minimasi fungsi objektif, nilai keanggotaan dan pusat *cluster* dikoreksi sehingga letak *cluster* dapat berada diposisi yang benar.

Cara kerja dari algoritma Fuzzy C-Means yaitu [25]:

1. Identifikasi data yang ingin di kelompokkan dalam bentuk matriks dengan ukuran $n \times m$ (x_{ij} adalah data ilustrasi ke- i ($i = 1, 2, \dots, n$) dan atribut ke- j ($j = 1, 2, \dots, m$))
2. Tentukanlah banyaknya *cluster* (c), pangkat (w), iterasi paling banyak, error paling sedikit yang diinginkan (ϵ), fungsi objektif awal ($p_0 = 0$) dan iterasi awal ($t = 1$)
3. Hasilkanlah angka acak μ_{ik} ($i = 1, 2, \dots, n$) dan ($k = 1, 2, \dots, c$) sebagai elemen dari matriks partisi utama u .

4. Hitunglah jumlah tiap kolom dengan menggunakan persamaan:

$$q_j = \sum_k^c \mu_{ik} = 1 \quad (3)$$

5. Hitunglah nilai elemen dari matriks partisi anggota himpunan u menggunakan persamaan:

$$\mu_{ik} = \frac{\mu_{ik}}{q_j} \quad (4)$$

6. Hitunglah nilai pusat *cluster* ke- k (v_{kj}) dimana $k=1, 2, \dots, c$ dan $j=1, 2, \dots, m$ dengan persamaan:

$$v_{kj} = (\mu_{ik})^w \sum_{j=1}^m (\mu_{ik})^m \quad (5)$$

7. Hitunglah fungsi objektif pada iterasi ke- t (p_t):

$$p_t = \sum_{i=1}^n \sum_{k=1}^c \left[\left(\sum_{j=1}^m (x_{ij} - v_{kj})^2 \right) (\mu_{ik})^w \right] \quad (6)$$

8. Hitunglah perubahan matriks partisi u :

$$\mu_{ik} = \frac{[\sum_{j=1}^m (x_{ij} - v_{kj})^2]_{w-1}^{-1}}{\sum_{k=1}^c [\sum_{j=1}^m (x_{ij} - v_{kj})^2]_{w-1}^{-1}} \quad (7)$$

9. Periksa kondisi berhenti:

- a. Bila $(p_t - p_{t-1}) < \epsilon$ atau ($t < \text{maxiter}$), itu berarti iterasi berhenti
- b. Bila $t = t + 1$, ulang kembali langkah ke-6 sampai langkah ke-9.

F. Uji Validitas

Pada tahap ini, akan dilakukan percobaan dari $k=2$ sampai $k=4$ untuk mengetahui hasil *cluster* yang akan didapat, kemudian membandingkan hasil validasi sehingga memperoleh hasil validasi tertinggi dari masing-masing algoritma tersebut. Analisis hasil perbandingan, pada penelitian ini ialah hasil dari perbandingan kedua algoritma K-means dan fuzzy C-means dengan percobaan *cluster* $k=2$ sampai dengan $k=4$. Uji validitas menggunakan *Silhouette Index*.

Silhouette Index ialah metode intepretasi untuk memvalidasi kelompok pada suatu objek. Metode ini menghasilkan representasi grafis yang akurat mengenai seberapa baiknya setiap objek yang berada dalam kelompoknya. *Silhouette Index* pertama kali dikembangkan oleh Rousseeuw pada tahun 1986 [19].

Metode ini banyak digunakan untuk memvalidasi *cluster* yang menggabungkan nilai kohesi dan separasi. Untuk menghitung nilai *Silhouette Index* dari sebuah data ke- i , ada 2 komponen yaitu a_i dan b_i . a_i adalah rata-rata jarak data ke- i terhadap semua data lainnya dalam satu *cluster*, sedangkan b_i didapatkan dengan menghitung rata-rata jarak data ke- i terhadap semua data dari *cluster* yang lain tidak dalam satu *cluster* dengan data ke- i , kemudian diambil yang terkecil.

3. Hasil dan Pembahasan

Pada penerapan algoritma K-Means dan Fuzzy C-Means dari $k=2$ hingga $k=4$ diperoleh hasil validasi dari tiap-tiap algoritma tersebut. Hasil validasi bisa diamati pada Tabel 1 & Tabel 2.

Tabel 1. Hasil Validasi
Algoritma K-Means

K	Silhouette Index
2	0.76
3	0.78
4	0.79

Tabel 2. Hasil Validasi
Algoritma Fuzzy C-Means

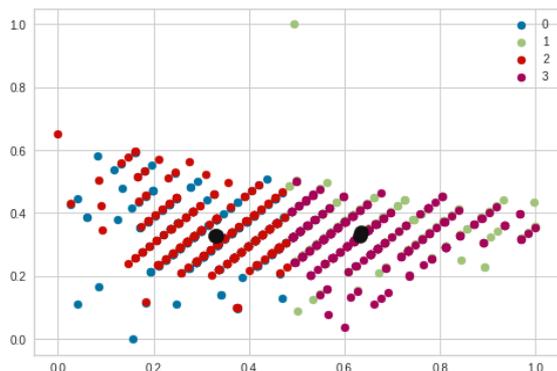
K	Silhouette Index
2	0.76
3	0.76
4	0.78

Bisa dilihat pada Tabel 1 dan Tabel 2, bahwa dengan menggunakan perhitungan validasi *Silhouette Index* pada kedua algoritma tersebut telah berada dikelompok yang tepat. Algoritma K-Means terdapat pada $k=4$ dan algoritma Fuzzy C-Means terdapat pada $k=4$, validasi terbaik dapat dilihat pada Tabel 3.

Tabel 3. Hasil Validasi Terbaik

Algoritma	K	Hasil Validasi Terbaik
K-Means	4	0.79
Fuzzy C-Means	4	0.78

Dapat dilihat pada Tabel 3 bahwa algoritma K-Means mendapatkan validasi tertinggi dibandingkan algoritma Fuzzy C-Means yaitu sebesar 0.79.

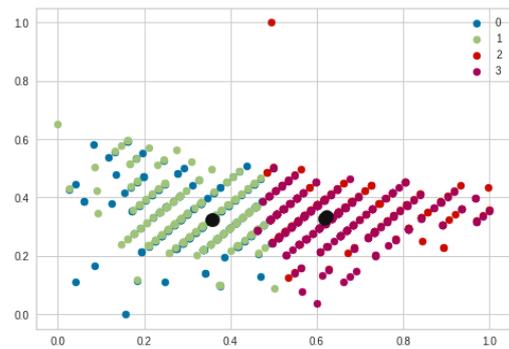


Gambar 4. Hasil Grafik K-Means

Pada Gambar 4, terlihat bahwa *cluster* 0 terdapat jumlah balita sebanyak 380 balita

berjenis kelamin laki-laki, *cluster* 1 terdapat jumlah balita sebanyak 399 balita berjenis kelamin perempuan, *cluster* 2 terdapat jumlah balita sebanyak 352 balita berjenis kelamin perempuan, dan *cluster* 3 terdapat jumlah balita sebanyak 431 balita berjenis kelamin laki-laki.

Hasil grafik yang diperoleh menggunakan $k=4$ pada algoritma Fuzzy C-Means dapat dilihat pada Gambar 5.



Gambar 5. Hasil Grafik Fuzzy C-Means

Pada Gambar 5, terlihat bahwa *cluster* 0 terdapat jumlah balita sebanyak 387 balita berjenis kelamin laki-laki, *cluster* 1 terdapat jumlah balita sebanyak 360 balita berjenis kelamin perempuan, *cluster* 2 terdapat jumlah balita sebanyak 391 balita berjenis kelamin perempuan, dan *cluster* 3 terdapat jumlah balita sebanyak 424 balita berjenis kelamin laki-laki.

Dari perbandingan hasil validasi terbaik yang diperoleh, perbandingan hasil validasi tertinggi serta perbandingan kedua hasil grafik dari algoritma Fuzzy C-Means dan K-Means diperoleh bahwa algoritma K-Means memiliki nilai akurasi tertinggi senilai 0.79 dengan nilai k sebesar 4 dan memiliki hasil validasi lebih tinggi dibandingkan dengan algoritma Fuzzy C-Means.

Hasil analisa *cluster* yang diperoleh menggunakan algoritma K-Means pada penelitian ini ditampilkan pada Tabel 4 berikut ini

Tabel 4. Hasil Analisa Cluster

CLUSTER	JUMLAH BALITA	JENIS KELAMIN
0	380	Laki-Laki
1	399	Perempuan
2	352	Perempuan
3	431	Laki-Laki

Analisa cluster yang dihasilkan yaitu *Cluster 0* merupakan kumpulan balita yang memiliki status gizi kurang. *Cluster 1* merupakan kumpulan balita yang memiliki status gizi obesitas. *Cluster 2* merupakan kumpulan balita yang memiliki status gizi buruk dan *Cluster 3* merupakan kumpulan balita yang memiliki status gizi baik.

4. Simpulan

Dari hasil penelitian yang mengacu kepada beberapa parameter seperti Berat Badan, Tinggi Badan, Posyandu dan Umur diperoleh hasil perhitungan algoritma K-Means dengan validasi terbaik sebesar 0.79 dan algoritma Fuzzy C-Means mendapatkan validasi sebesar 0.78. Dari kedua algoritma tersebut, algoritma K-Means lebih dapat menggambarkan nilai gizi balita pada Puskesmas Pasir Jaya karena mendapatkan hasil Validasi tertinggi yaitu sebesar 0.79 dengan jumlah cluster yang dihasilkan sebanyak 4 cluster.

Pada penelitian ini didapatkan kesimpulan bahwa *Cluster 0* merupakan kumpulan balita yang memiliki status gizi kurang. *Cluster 1* merupakan kumpulan balita yang memiliki status gizi obesitas. *Cluster 2* merupakan kumpulan balita yang memiliki status gizi buruk dan *Cluster 3* merupakan kumpulan balita yang memiliki status gizi baik.

Untuk pengembangan selanjutnya, sebaiknya data yang digunakan adalah data yang baru sehingga informasi yang dihasilkan juga termasuk kedalam informasi yang terbaru. Penelitian yang akan dilakukan berikutnya ada baiknya dibandingkan dengan metode *clustering* lainnya agar bisa menghasilkan penelitian yang lebih baik.

Daftar Pustaka

- [1] E. Irfiani and S. S. Rani, "Algoritma K-Means Clustering untuk Menentukan Nilai Gizi Balita," *Jurnal Sistem dan Teknologi Informasi (JUSTIN)*, vol. 6, no. 4, p. 161, 2018.
- [2] W. M. P. Dhuhita, "Clustering Menggunakan Metode K-Means Untuk," *Jurnal Informatika*, vol. 15, no. 2, pp. 160--174, 2016.
- [3] J. Muningsgar, "Profil Tumbuh Kembang Anak Balita dengan Malnutrisi di Desa Kesongo Kecamatan Tuntang," *Magistrorum et Scholarium: Jurnal Pengabdian Masyarakat*, vol. 1, no. 2, pp. 171--185, 2021.
- [4] S. R. Patty and F. Nugroho, "Kemiskinan Dan Malnutrisi Pada Anak Balita Dalam Keluarga Nelayan Di Wilayah Pesisir Kota Serang," *EMPATI: Jurnal Ilmu Kesejahteraan Sosial*, vol. 8, no. 2, pp. 109--125, 2020.
- [5] N. Nazilia and M. Iqbal, "Peningkatan Pengetahuan Ibu Tentang Gizi untuk Mengatasi Gizi Buruk pada Anak Balita dengan Aplikasi ' Anak Sehat Makan Sehat (Emas),'" *Jurnal Pangan Kesehatan dan Gizi*, vol. 1, no. 1, pp. 46--53, 2020.
- [6] B. S. Ashari, S. C. Otniel, and Rianto, "Perbandingan Kinerja K-Means Dengan DSCAN Untuk Metode Clustering Data Penjualan Online Retail," *Jurnal Siliwangi*, vol. 5, no. 2, pp. 72--77, 2019.
- [7] A. Maulana and A. A. Fajrin, "Penerapan Data Mining Untuk Analisis Pola Pembelian Konsumen Dengan Algoritma Fp-Growth Pada Data Transaksi Penjualan Spare Part Motor," *Klik - Kumpulan Jurnal Ilmu Komputer*, vol. 5, no. 1, p. 27, 2018.
- [8] F. Yunita, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Pada Penerimaan Mahasiswa Baru," *Sistemasi*, vol. 7, no. 3, p. 238, 2018.

- [9] D. R. Ningrat, D. A. I. Maruddani, and T. Wuryandari, "Analisis Cluster Dengan Algoritma K-Means Dan Fuzzy C-Means Clustering Untuk Pengelompokan Data Obligasi Korporasi," *Jurnal Gaussian*, vol. 5, no. 4, pp. 641–650, 2016.
- [10] M. G. Sadewo, A. P. Windarto, and A. Wanto, "Penerapan Algoritma Clustering Dalam Mengelompokkan Banyaknya Desa/Kelurahan Menurut Upaya Antisipasi/ Mitigasi Bencana Alam Menurut Provinsi Dengan K-Means," *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, vol. 2, no. 1, pp. 311–319, 2018.
- [11] F. E. M. Agustin, "Implementasi Algoritma K-Means Untuk Menentukan Kelompok Pengayaan Materi Mata Pelajaran Ujian Nasional (Studi Kasus: Smp Negeri 101 Jakarta)," *Jurnal Teknik Informatika*, vol. 8, no. 1, pp. 73–78, 2015.
- [12] L. Rahmah, "Perbandingan Hasil Penggerombolan K-Means, Fuzzy K-Means, Dan Two Step Clustering," *Jurnal Pendidikan Matematika*, vol. 2, no. 1, p. 39, 2017.
- [13] E. Muningsih and S. Kiswati, "Penerapan Metode K-Means Untuk Clustering Produk Online Shop Dalam Penentuan Stok Barang," *Jurnal Bianglala Informatika*, vol. 3, no. 1, pp. 10–17, 2015.
- [14] B. M. Metisen and H. L. Sari, "Analisis clustering menggunakan metode K-Means dalam pengelompokkan penjualan produk pada Swalayan Fadhila," *Jurnal Media Infotama*, vol. 11, no. 2, pp. 110–118, 2015.
- [15] E. Triandini, F. A. Hermawati, and I. K. P. Suniantara, "Hierarchical Clustering For Functionalities E-Commerce Adoption," *Jurnal Ilmiah Kursor*, vol. 10, no. 3, pp. 111–118, 2020.
- [16] E. T. Puteri, G. Kusnanto, and C. J. Thomas, "Penerapan K-Means Clustering Untuk Segmentasi Pelanggan Pada Sistem Customer Relationship Management Di Pt. Unichem Candi Indonesia," *Konvergensi*, vol. 15, no. 2, 2020.
- [17] K. Fatmawati and A. P. Windarto, "Data Mining: Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi," *Computer Engineering, Science and System Journal*, vol. 3, no. 2, p. 173, 2018.
- [18] A. Aditya, I. Jovian, and B. N. Sari, "Implementasi K-Means Clustering Ujian Nasional Sekolah Menengah Pertama di Indonesia Tahun 2018/2019," *Jurnal Media Informatika Budidarma*, vol. 4, no. 1, p. 51, 2020.
- [19] M. Siti, "Segmentasi Perilaku Pembelian Pelanggan Berdasarkan Model RFM dengan Metode K-Means," *Jurnal Sistem Informasi*, vol. 2, no. 1, pp. 9–15, 2018.
- [20] S. Saifullah, M. Zarlis, Z. Zakaria, and R. W. Sembiring, "Analisa Terhadap Perbandingan Algoritma Decision Tree Dengan Algoritma Random Tree Untuk Pre-Processing Data," *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, vol. 1, no. 2, p. 180, 2017.
- [21] A. Amalia, M. S. Lydia, S. D. Fadilla, and M. Huda, "Perbandingan Metode Klaster dan Preprocessing Untuk Dokumen Berbahasa Indonesia," *Jurnal Rekayasa ElektriKa*, vol. 14, no. 1, pp. 35–42, 2018.
- [22] N. Rohmawati, S. Defiyant, and M. Jajuli, "Implementasi Algoritma K-Means Dalam Pengklasteran Mahasiswa Pelamar Beasiswa," *Jitter 2015*, vol. I, no. 2, pp. 62–68, 2015.
- [23] C. L. Simbolon, N. Kusumastuti, and B. Irawan, "Clustering lulusan mahasiswa matematika fmipa untan pontianak menggunakan algoritma fuzzy c - means," *Buletin Ilmiah Mat.*

- Stat. Dan Terapannya (Bimaster)*, vol. 02, no. 1, pp. 21–26, 2013.
- [24] A. Suryadi, “Sistem Pengenalan Wajah Menggunakan Metode Principal Component Analysis (PCA) Dengan Algoritma Fuzzy C-Means (FCM),” *Jurnal Pendidikan Matematika*, vol. 4, no. 2, pp. 58–65, 2015.
- [25] N. Shofiani, “Segmentasi Supplier Menggunakan Metode K- Means Clustering (Studi Kasus : Ptpn X Pg Meritjan),” Institut Teknologi Sepuluh Nopember Surabaya, 2017.