

Implementasi Data Mining Clustering Menggunakan Metode K-Means (Studi Kasus: Status Gizi Balita di Jawa Tengah)

Sri Wulandari¹⁾, Dimas Mustarsyid²⁾, Keke Dyah Arlita³⁾

¹Politeknik Indonusa Surakarta, ²Politeknik Indonusa Surakarta, ³ Politeknik Indonusa Surakarta
¹²³Jl. K.H Samanhudi No.31, Bumi, Kec. Laweyan, Kota Surakarta
sriwulandari@poltekindonusa.ac.id, b21001@poltekindonusa.ac.id,
b21002@poltekindonusa.ac.id

Abstrak

Kesehatan dan status gizi merupakan salah satu tolak ukur yang dapat menggambarkan keadaan gizi seseorang. Permasalahan terkait gizi pada balita di Indonesia merupakan masalah serius yang mempengaruhi kesehatan balita tersebut. Maka dari itu pemantauan status gizi balita sangat penting untuk mendeteksi secara dini masalah gizi dan memberikan intervensi yang tepat. Penelitian ini bertujuan untuk memprediksi status gizi balita menggunakan algoritma K-Means dengan data dari tiga Puskesmas di Kecamatan Sapuran. Algoritma K-Means digunakan karena kemampuannya dalam mengelompokkan data secara cepat dan efisien berdasarkan kesamaan atribut, serta fleksibilitasnya dalam menangani berbagai tipe data yang berbeda, termasuk data seperti berat badan dan umur. Data yang digunakan meliputi umur, tinggi badan, dan berat badan, dengan pengelompokan menjadi 4 cluster yaitu sangat kurang, kurang, normal, dan risiko lebih. Penelitian ini mengembangkan sistem prediksi status gizi balita berbasis web menggunakan algoritma K-Means. Hasil penelitian menunjukkan bahwa algoritma K-Means mampu memprediksi status gizi balita. Dari hasil dari penelitian menunjukkan bahwa terdapat 22.64% balita yang memiliki status gizi kurang. Balita yang memiliki status gizi normal sebanyak 22.91%. Balita yang memiliki status gizi resiko lebih sebanyak 28.14%. Balita yang memiliki status gizi sangat kurang sebanyak 26.32%.

Kata kunci: K-Means, Prediksi, Status Gizi Balita.

Abstract

Health and nutritional status are key indicators that can describe a person's nutritional condition. Nutritional issues among toddlers in Indonesia are a serious problem that affects their overall health. Therefore, monitoring the nutritional status of toddlers is crucial for early detection of nutritional problems and for providing appropriate interventions. This study aims to predict the nutritional status of toddlers using the K-Means algorithm with data from three community health centers (Puskesmas) in Sapuran District. The K-Means algorithm is used due to its ability to quickly and efficiently group data based on attribute similarities, as well as its flexibility in handling various types of data, including data such as weight and age. The data used includes age, height, and weight, which are clustered into four categories: severely malnourished, malnourished, normal, and at risk of being overweight. This research developed a web-based toddler nutritional status prediction system using the K-Means algorithm. The results show that the K-Means algorithm is capable of predicting the nutritional status of toddlers. The study revealed that 22.64% of toddlers were classified as malnourished, 22.91% were classified as having normal nutritional status, 28.14% were at risk of being overweight, and 26.32% were classified as severely malnourished.

Keywords: K-Means, Prediction, Toddler Nutritional Status

1. PENDAHULUAN

Permasalahan terkait dengan gizi pada anak di Indonesia sangat sering ditemui. Menurut laporan kesehatan dari Organisasi Kesehatan Dunia (WHO) dan pusat kesehatan balita, beberapa masalah umum yang mempengaruhi kesehatan balita terkait dengan status gizi, seperti kekurangan protein dan energi, obesitas, kekurangan vitamin A, gangguan karena kekurangan yodium, serta anemia atau kekurangan zat besi (fe) [1][2]. Indonesia, sebagai negara berkembang, masih menghadapi permasalahan gizi balita yang signifikan. Berdasarkan data dari Kementerian Kesehatan Republik Indonesia, prevalensi balita dengan status gizi kurang dan gizi buruk masih tinggi, meskipun berbagai upaya telah dilakukan untuk mengurangi angka tersebut. Salah satu tantangan dalam penanganan permasalahan gizi adalah kurangnya deteksi dini status gizi pada balita [3]. Deteksi dini sangat penting agar intervensi yang tepat dapat segera diberikan sebelum kondisi gizi memburuk.

Saat ini, kemajuan teknologi komputer berkembang sangat cepat, sehingga semua kebutuhan diintegrasikan dalam sistem untuk mempermudah pengguna melakukan aktivitasnya. Berbagai kegiatan dapat dipermudah dengan adanya sistem, termasuk dalam memprediksi status gizi. Dalam penentuan status gizi pada balita dibutuhkan metode yang tepat untuk melakukan prediksi status gizi.

Penerapan algoritma K-Means dapat digunakan untuk mengelompokkan status gizi balita yang diperoleh dari kondisi balita tersebut. K-Means adalah algoritma yang digunakan dalam kurang, sangat kurang, resiko lebih lebih, dan normal.

Penelitian ini bertujuan untuk mengimplementasikan algoritma K-Means dalam memprediksi status gizi balita pada tiga Puskesmas di Kecamatan Sapuran. Sistem prediksi status gizi balita ini dapat memudahkan pihak puskesmas dalam proses pengelompokan dan analisis data balita secara otomatis.

2. TINJAUAN PUSTAKA

A. Data Mining

Data mining merupakan proses pengolahan data dalam jumlah besar

untuk menemukan pola, hubungan, atau informasi yang sebelumnya tidak diketahui. Proses ini memanfaatkan teknik statistik, matematika, dan kecerdasan buatan untuk menghasilkan pengetahuan yang berguna dari data yang tersedia. Data mining sering digunakan dalam berbagai bidang seperti kesehatan, pendidikan, bisnis, dan pemerintahan untuk membantu proses pengambilan keputusan.

Dalam bidang kesehatan, data mining dapat dimanfaatkan untuk menganalisis data pasien, memprediksi penyakit, serta mengelompokkan kondisi kesehatan berdasarkan karakteristik tertentu. Dengan menggunakan teknik data mining, data kesehatan yang sebelumnya hanya berupa data mentah dapat diolah menjadi informasi yang lebih bermakna sehingga dapat membantu dalam menentukan kebijakan atau tindakan medis yang tepat.

B. Clustering

Clustering merupakan salah satu teknik dalam data mining yang digunakan untuk mengelompokkan data berdasarkan tingkat kemiripan karakteristik antar data. Metode ini termasuk dalam teknik unsupervised learning karena tidak memerlukan label atau kelas yang telah ditentukan sebelumnya. Data yang memiliki karakteristik yang mirip akan dikelompokkan ke dalam satu cluster, sedangkan data yang berbeda akan ditempatkan pada cluster yang berbeda.

Tujuan utama dari clustering adalah untuk menemukan pola atau struktur yang tersembunyi dalam suatu kumpulan data. Dalam bidang kesehatan, metode clustering dapat digunakan untuk mengelompokkan data pasien, mengidentifikasi pola penyakit, serta membantu analisis kondisi kesehatan berdasarkan atribut tertentu seperti usia, berat badan, dan tinggi badan.

C. Algoritma K-Means

K-Means merupakan salah satu algoritma clustering yang paling populer dalam data mining. Algoritma ini digunakan untuk membagi data ke dalam sejumlah cluster yang telah ditentukan sebelumnya. Huruf K pada K-Means menunjukkan jumlah cluster yang ingin dibentuk, sedangkan Means mengacu pada nilai rata-rata dari data dalam suatu cluster yang disebut sebagai centroid. Proses kerja algoritma K-Means dimulai dengan menentukan jumlah cluster yang diinginkan, kemudian memilih centroid awal secara acak. Setelah itu, setiap data akan dihitung jaraknya terhadap masing-masing centroid menggunakan metode perhitungan jarak, biasanya menggunakan Euclidean Distance. Data kemudian ditempatkan pada cluster dengan jarak terdekat terhadap centroid. Selanjutnya centroid baru dihitung berdasarkan rata-rata data pada setiap cluster. Proses ini dilakukan secara berulang hingga tidak terjadi perubahan pada anggota cluster. Algoritma K-Means memiliki beberapa kelebihan, antara lain mudah diimplementasikan, memiliki proses perhitungan yang relatif cepat, serta mampu menangani data dalam jumlah besar. Oleh karena itu, algoritma ini banyak digunakan dalam berbagai penelitian yang berkaitan dengan pengelompokan data.

D. Normalisasi Data

Normalisasi data merupakan proses penyesuaian skala nilai atribut agar berada dalam rentang yang sama. Proses ini penting dilakukan terutama ketika atribut memiliki rentang nilai yang berbeda-beda. Tanpa normalisasi, atribut dengan nilai yang lebih besar dapat mendominasi proses perhitungan jarak dalam algoritma clustering. Salah satu metode normalisasi yang umum digunakan adalah Min-Max Scaling, yaitu metode yang mengubah

nilai data ke dalam rentang antara 0 hingga 1. Dengan normalisasi, setiap atribut memiliki pengaruh yang seimbang dalam proses pengelompokan data sehingga hasil clustering menjadi lebih akurat.

E. Status Gizi Balita

Status gizi balita merupakan kondisi kesehatan balita yang dipengaruhi oleh asupan makanan dan kondisi kesehatan secara umum. Penilaian status gizi biasanya dilakukan berdasarkan indikator antropometri seperti berat badan, tinggi badan, dan umur. Data tersebut kemudian dibandingkan dengan standar pertumbuhan yang telah ditetapkan. Masalah gizi pada balita masih menjadi salah satu permasalahan kesehatan yang cukup serius di Indonesia. Beberapa kondisi yang sering ditemukan antara lain gizi kurang, gizi buruk, gizi normal, dan risiko kelebihan berat badan. Pemantauan status gizi secara rutin sangat penting untuk mendeteksi secara dini adanya masalah gizi sehingga tindakan pencegahan atau penanganan dapat segera dilakukan. Dengan memanfaatkan teknologi informasi dan metode data mining seperti K-Means, proses analisis status gizi balita dapat dilakukan secara lebih cepat dan sistematis. Hal ini dapat membantu tenaga kesehatan dalam mengelompokkan kondisi gizi balita serta mempermudah proses pengambilan keputusan terkait penanganan masalah gizi.

F. Penelitian Terdahulu

Beberapa penelitian sebelumnya telah menerapkan algoritma K-Means dalam berbagai bidang analisis data. Penelitian yang dilakukan oleh Irtawaty menggunakan metode K-Means untuk klasifikasi penyakit ginjal dan memperoleh tingkat akurasi mencapai 90%. Penelitian lain yang dilakukan oleh Al-Rizki menggunakan K-Means untuk memprediksi penyebaran penyakit TBC dan menghasilkan tingkat

akurasi sebesar 84,13%. Selain itu, penelitian oleh Ariawan yang menggunakan metode K-Means untuk memprediksi nilai akhir mahasiswa menunjukkan tingkat akurasi sebesar 93%.

Hasil dari berbagai penelitian tersebut menunjukkan bahwa algoritma K-Means memiliki kemampuan yang baik dalam melakukan pengelompokan data dan menghasilkan tingkat akurasi yang cukup tinggi. Oleh karena itu, metode ini dianggap tepat untuk digunakan dalam penelitian yang berkaitan dengan pengelompokan status gizi balita berdasarkan data antropometri.

3. METODE PENELITIAN

A. Metode K-Means

Definisi dari K-Means adalah, K yang merupakan jumlah cluster yang diinginkan, sedangkan Means merujuk pada nilai rata-rata dari suatu kelompok data yang dalam konteks ini disebut sebagai cluster [9][10]. Sebelum melakukan clustering dengan K-Means, perlu normalisasi data agar nilai-nilai dari atribut yang berbeda dalam rentang yang sama dengan menggunakan rumus berikut [11].

pengelompokan data non-hierarki, yang membagi data menjadi dua kelompok atau lebih [4][5].

$$\text{Hasil } X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

dari pengelompokan ini dapat digunakan untuk memprediksi kategori status gizi anak balita baru berdasarkan pola yang telah diidentifikasi dari data historis. Untuk memprediksi status gizi balita baru, atribut balita akan dianalisis dan dibandingkan dengan centroid dari cluster yang telah terbentuk. Balita tersebut kemudian akan dimasukkan ke dalam cluster dengan centroid terdekat, dan label dari klaster tersebut akan digunakan sebagai prediksi status gizi balita tersebut.

Berdasarkan referensi beberapa penelitian sebelumnya yang menggunakan algoritma K-Means,

algoritma ini diketahui memiliki tingkat ketepatan yang baik. Pada penelitian yang dilakukan oleh Irtawaty [6] yaitu klasifikasi penyakit ginjal dengan algoritma K-Means menunjukkan tingkat akurasi mencapai 90%. Penelitian yang dilakukan oleh Al-Rizki [7] yaitu prediksi penyebaran penyakit TBC 3 menggunakan metode K-Means Clustering menunjukkan tingkat akurasi sebesar 84,13%. Penelitian yang dilakukan oleh Ariawan [8] yaitu prediksi nilai akhir matakuliah mahasiswa menggunakan metode K-Means Clustering menunjukkan tingkat akurasi sebesar 93%. Berdasarkan penelitian terdahulu maka metode ini dapat digunakan untuk klasterisasi status gizi balita yang ada pada tiga puskesmas di Kecamatan Sapuran.

Data yang digunakan dalam penelitian ini diambil dari tiga Puskesmas yang ada di Kecamatan Sapuran, yaitu Puskesmas Gunungpati Kulon, Puskesmas Kaliwiro, dan Puskesmas Batur. Hasil dengan X_{norm} adalah nilai yang sudah dinormalisasi, X adalah nilai asli dari data, X_{min} adalah nilai minimum dari atribut yang sedang dinormalisasi, dan X_{max} adalah nilai maksimum dari atribut yang sedang dinormalisasi.

Setelah data dinormalisasi dilakukan clustering menggunakan *K-Means*. Berikut merupakan langkah-langkah pada algoritma *K-Means* clustering [12]:

- A. Menetapkan jumlah k sebagai cluster yang akan dibentuk.
- B. Menentukan titik pusat dari tiap cluster (centroid awal) secara acak.
- C. Menghitung jarak objek ke centroid serta jarak objek dengan centroid dilakukan menggunakan rumus matematika euclidean distance. Berikut merupakan rumusnya [13]:

$$d(x, c) = \sqrt{\sum_{i=1}^n (x_i - c_i)^2}$$

- D. Berdasarkan hasil perhitungan jarak euclidean, setiap objek dikelompokkan ke dalam cluster yang memiliki jarak terdekat ke centroid.
- E. Setelah data dikelompokkan, hitung centroid baru untuk setiap cluster. Centroid baru dihitung sebagai rata-rata dari semua objek yang ada dalam cluster tersebut dengan menggunakan rumus persamaan berikut [13]: dari penelitian ini nantinya berupa status gizi balita $V = \frac{1}{N_i} \sum X_{kj}$ (3) menggunakan standar acuan berat badan terhadap $ijNik=1kj$ jumlah (BB/U). Untuk memprediksi status gizi dibutuhkan beberapa variabel yaitu variabel umur (bulan) dan berat badan (kg). Klasifikasi dari status gizi balita (BB/U) pada penelitian ini yaitu status gizi dengan V_{ij} adalah centroid baru cluster ke-i untuk atribut ke-j, N_i adalah jumlah objek dalam cluster ke-i, dan X_{kj} adalah nilai atribut ke-j dari objek ke-k dalam cluster ke-i.
- F. Lakukan kembali langkah ke 3 sampai tidak ada anggota cluster yang berpindah ke cluster lain atau nilai centroid tidak berubah.

B. Perancangan Sistem

Dalam penelitian ini, metode perancangan sistem yang digunakan adalah prototyping, yaitu metode pengembangan perangkat lunak yang menghasilkan model fisik kerja sistem yang berfungsi sebagai versi awal dari sistem [14]. Langkah-langkah dalam prototyping adalah sebagai berikut [15]:

- 1) Pengumpulan kebutuhan. Tahap pertama adalah mengidentifikasi dan memahami kebutuhan pengguna melalui diskusi untuk mendapatkan informasi yang tepat tentang apa saja yang diinginkan pengguna dari sistem yang akan dibangun.

- 2) Proses Desain, pada tahap ini dilakukan pengembangan sistem yang melibatkan desain dan proses yang akan terjadi, serta identifikasi input (masukan) dan output (keluaran) yang relevan untuk sistem yang dikembangkan.
- 3) Membangun prototype, pada tahap ini pembuatan prototipe dilakukan berdasarkan desain yang telah disepakati dengan pengguna.
- 4) Evaluasi, pada tahap ini dilakukan evaluasi dan verifikasi terhadap prototype sistem yang telah dikembangkan, serta mengevaluasi apakah prototype sistem tersebut telah sesuai dengan yang diharapkan.

4. HASIL DAN PEMBAHASAN

A. Penerapan Metode K-Means

Pada penelitian ini data yang diperoleh untuk diproses pada penelitian penentuan status gizi balita ini didapatkan dari tiga Puskesmas yang ada di Kecamatan Sapuran, yaitu Puskesmas Gunungpati (Tabel 1), Puskesmas Kaliwiro dan Puskesmas Batur (Tabel 2) yang merupakan puskesmas pembantu puskesmas Tuk-Tuk. Data yang digunakan sebanyak 955. Pengelompokan data status gizi balita dibagi menjadi empat nilai status gizi yaitu gizi kurang, sangat kurang, resiko lebih, dan normal [16].

Tabel 1. Data balita pada puskesmas Gunungpati

7	Alvaro Gultom	41	12.2
8	Amelia Siregar	18	9.5
9	Anggira Purba	48	14.5
10	Anju Napitu	48	15.5

511	Saly Lucia Simamora	9	8.5
512	Uli Arta Sinurat	10	7.9
513	Yohana Manik	49	17
514	Butet Turnip	0	3.4

Tabel 2. Data balita pada puskesmas Kaliwiro dan Batur

No	Nama	Umur (Bulan)	Berat Badan (Kg)
1	Tristan Pakpahan	56	13.8
2	Berliana Sidabutar	54	14.2
3	Sarah Manurung	48	14.7
4	Samuel Sidabutar	45	13.3
5	Kirana Turnip	39	12.9
6	Refalisa Simarmata	30	9.7
7	Butet Nainggolan	18	9.5
8	Ezra Manalu	33	10.7
9	Revan Jonaris Sinaga	52	15.4
10	Butet Simarmata	26	10.6

438	Ucok Sidabutar	0	4
439	Butet Silalahi	0	4.1
440	Artaulina Sinaga	53	15.2
441	Butet Siahaan	14	10.1

Normalisasi data dilakukan agar setiap atribut memiliki skala yang sama. Metode normalisasi yang digunakan adalah *Min-Max Scaling*. Hasil normalisasi data dapat dilihat pada Tabel 3 dan Tabel

4. Berikut ini merupakan contoh perhitungan normalisasi data pertama:

Umur

$$X' = \frac{34 - 0}{59 - 0} = 0.5763$$

Berat Badan

$$X' = \frac{15.5 - 2.7}{24.6 - 2.7} = 0.5845$$

Tabel 3. Data balita puskesmas Gunungpati setelah normalisasi

No	Nama	Umur (Bulan)	Berat Badan (Kg)
1	Abraham Silalahi	0.576271186	0.584474886
2	Adelia Karmisa Silalahi	0.694915254	0.415525114
3	Adriel Imanuel Sinaga	0.627118644	0.470319635
4	Aldo Silalahi	0.762711864	0.488584475
5	Alfrido Davinci Silalahi	0.169491525	0.287671233
6	Alpon J Silalahi	0.372881356	0.470319635
7	Alvaro Gultom	0.694915254	0.433789954
8	Amelia Siregar	0.305084746	0.310502283
9	Anggira Purba	0.813559322	0.538812785
10	Anju Napitu	0.813559322	0.584474886

511	Saly Lucia Simamora	0.152542373	0.264840183

512	Uli Arta Sinurat	0.169491525	0.237442922
513	Yohana Manik	0.830508475	0.652968037
514	Butet Turnip	0	0.03196347

Tabel 4. Data baita puskesmas Kaliwiro dan Batur setelah normalisasi

No	Nama	Umur (Bulan)	Berat Badan (Kg)
1	Tristan Pakpahan	0.949152542	0.591954023
2	Berliana Sidabutar	0.915254237	0.614942529
3	Sarah Manurung	0.813559322	0.643678161
4	Samuel Sidabutar	0.762711864	0.563218391
5	Kirana Turnip	0.661016949	0.540229885
6	Refalisa Simarmata	0.508474576	0.356321839
7	Butet Nainggolan	0.305084746	0.344827586
8	Ezra Manalu	0.559322034	0.413793103
9	Revan Jonaris Sinaga	0.881355932	0.683908046
10	Butet Simarmata	0.440677966	0.408045977

438	Ucok Sidabutar	0	0.028735632
439	Butet Silalahi	0	0.034482759
440	Artaulina Sinaga	0.898305085	0.672413793
441	Butet Siahaan	0.237288136	0.379310345

Setelah normalisasi data dilakukan proses penentuan nilai pusat cluster (centroid) dari data yang digunakan. Tabel 5 dan Tabel 6 berikut adalah hasil penentuan centroid secara acak dari data balita yang akan digunakan:

Tabel 5. Centroid awal puskesmas Gunungpati

Cluster	Umur (Bulan)	Berat Badan (Kg)
C1	0.576271186	0.584474886
C2	0.694915254	0.415525114
C3	0.627118644	0.470319635
C4	0.762711864	0.488584475

Tabel 6. Centroid awal puskesmas Kaliwiro dan Batur

Cluster	Umur (Bulan)	Berat Badan (Kg)
C1	0.627118644	0.862068966
C2	0.610169492	0.574712644
C3	0.576271186	0.517241379
C4	0.559322034	0.66091954

Selanjutnya, dilakukan perhitungan jarak antara setiap data dengan setiap pusat cluster menggunakan rumus jarak euclidean. Dengan demikian, ditemukan jarak terdekat antara setiap data dengan centroid. Cluster awal ditunjukkan pada Tabel 7 dan Tabel 8. Berikut

merupakan contoh perhitungan jarak data pertama ke centroid:

$$d(x, C_1) = \sqrt{(0.5763 - 0.5763)^2 + (0.5845 - 0.5845)^2}$$

$$d(x, C_1) = \sqrt{0 + 0}$$

$$d(x, C_1) = 0$$

$$d(x, C_2) = \sqrt{(0.5763 - 0.6949)^2 + (0.5845 - 0.4155)^2}$$

$$d(x, C_2) = \sqrt{(-0.1186)^2 + (0.1690)^2}$$

$$d(x, C_2) = \sqrt{0.01407 + 0.02856}$$

$$d(x, C_2) = \sqrt{0.04263}$$

$$d(x, C_2) = 0.2065$$

$$d(x, C_3) = \sqrt{(0.5763 - 0.6271)^2 + (0.5845 - 0.4703)^2}$$

$$d(x, C_3) = \sqrt{(-0.0508)^2 + (0.1142)^2}$$

$$d(x, C_3) = \sqrt{0.00258 + 0.01304}$$

$$d(x, C_3) = \sqrt{0.01562}$$

$$d(x, C_3) = 0.1250$$

$$d(x, C_4) = \sqrt{(0.5763 - 0.7627)^2 + (0.5845 - 0.4886)^2}$$

$$d(x, C_4) = \sqrt{(-0.1864)^2 + (0.0959)^2}$$

$$d(x, C_4) = \sqrt{0.03474 + 0.00920}$$

$$d(x, C_4) = \sqrt{0.04394}$$

$$d(x, C_4) = 0.2097$$

Tabel 7. Clustering awal puskesmas Gunungpati

Nama	C1	C2	C3	C4	Cluster
Abraham Silalahi	2.221	0.206	0.12	0.21	C1
Adelia Karnisa Silalahi	0.206	0	0.09	0.1	C2
Adriel Imanuel Sinaga	0.125	0.087	0	0.137	C3
Aldo Silalahi	0.21	0.1	0.14	0	C4
Alfrido Davinci Silalahi	0.504	0.541	0.49	0.626	C3
Alpon J Silalahi	0.233	0.327	0.25	0.39	C1
Alvaro Gultom	0.192	0.018	0.08	0.087	C2
Amelia Siregar	0.385	0.404	0.36	0.491	C3
Anggira Purba	0.242	0.171	0.2	0.071	C4
nju Napitu	0.237	0.206	0.22	0.109	C4
...

Nama	C1	C2	C3	C4	Cluster
Saly Lucia Simamora	0.531	0.563	0.52	0.65	C3
Uli Arta Sinurat	0.535	0.555	0.51	0.644	C3
Yohana Manik	0.263	0.273	0.27	0.178	C4
Butet Turnip	0.798	0.794	0.77	0.889	C3

Tabel 8. Clustering awal puskesmas Kaliwiro dan Batur

Nama	C1	C2	C3	C4	Cluster
Tristan Pakpahan	0.42	0.339	0.38	0.396	C2
Berliana Sidabutar	0.38	0.308	0.353	0.359	C2
Sarah Manurung	0.287	0.215	0.269	0.255	C2
Samuel Sidabutar	0.328	0.153	0.192	0.226	C2
Kirana Turnip	0.324	0.061	0.088	0.158	C2
Refalisa Simarmata	0.519	0.241	0.175	0.309	C3
Butet Nainggolan	0.609	0.382	0.321	0.406	C3
Ezra Manalu	0.453	0.169	0.105	0.247	C3
Revan Jonaris S	0.31	0.292	0.348	0.323	C2
Butet Simarmata	0.491	0.238	0.174	0.279	C3
...
Ucok Sidabutar	1.043	0.819	0.755	0.844	C3
Butet Silalahi	1.038	0.815	0.752	0.84	C3
Artaulina Sinaga	0.331	0.304	0.357	0.339	C2
Butet Siahaan	0.621	0.421	0.366	0.428	C3

Dari hasil perhitungan data, diperoleh centroid baru yang terletak pada C1, C2, C3, dan C4. Data yang telah terkelompok dalam masing-masing cluster kemudian dijumlahkan, dan rata-rata dari data tersebut akan menjadi centroid yang baru (Tabel 9 dan Tabel 10). Berikut ini merupakan contoh perhitungan centroid baru.

$$C_{1,Umur} = \frac{0.5763 + 0.3729 + 0.7288 + 0.2373 + \dots + 0.2542}{58}$$

$$C_{1,Umur} = 0.4793$$

$$C_{1,Berat\ Badan} = \frac{0.5845 + 0.4703 + 0.6530 + 0.4018 + \dots + 0.4064}{58}$$

$$C_{1,Berat\ Badan} = 0.5316$$

Tabel 9. Centroid baru puskesmas Gunungpati iterasi ke-2

Cluster	Umur (Bulan)	Berat Badan (Kg)
C1	0.479251899	0.531648559
C2	0.690677966	0.419710807
C3	0.325204658	0.319588109

Tabel 10. Centroid baru puskesmas Kaliwiro dan Batur iterasi ke-2

Cluster	Umur (Bulan)	Berat Badan (Kg)
C1	0.878787879	0.801985371
C2	0.78782172	0.603448276
C3	0.303275943	0.346281405
C4	0.57748184	0.649835796

Selanjutnya hitung kembali jarak antar data ke centroid baru hingga tiap cluster tidak lagi berpindah. Hasil clustering dapat dilihat pada Tabel 11 dan Tabel 12.

Tabel 11. Hasil clustering pada puskesmas Gunungpati

Nama	C1	C2	C3	C4	Cluster
Abraham Silalahi	0.286	0.102	0.598	0.301	C2
Adelia Karnisa Silalahi	0.318	0.107	0.617	0.259	C2
Adriel Imanuel Sinaga	0.265	0.023	0.577	0.281	C2
Aldo Silalahi	0.399	0.142	0.707	0.159	C2
Alfrido Davinci Silalahi	0.228	0.495	0.1	0.773	C3
Alpon J Silalahi	0.094	0.248	0.371	0.52	C1
Alvaro Gultom	0.321	0.095	0.624	0.246	C2
Amelia Siregar	0.099	0.364	0.218	0.641	C1
Anggira Purba	0.464	0.199	0.775	0.088	C4
Anju Napitu	0.482	0.214	0.795	0.065	C4
...
Saly Lucia Simamora	0.253	0.52	0.071	0.798	C3
Uli Arta Sinurat	0.252	0.518	0.064	0.795	C3
Yohana Manik	0.529	0.265	0.845	0.071	C4
Butet Turnip	0.512	0.772	0.207	1.045	C3

Tabel 12. Hasil clustering pada puskesmas Batur

Nama	C1	C2	C3	C4	Cluster
Tristan Pakpahan	0.155	0.271	0.905	0.604	C1
Berliana Sidabutar	0.125	0.239	0.884	0.58	C1
Sarah Manurung	0.13	0.149	0.81	0.499	C1
Samuel Sidabutar	0.224	0.085	0.725	0.419	C2
Kirana Turnip	0.312	0.043	0.625	0.316	C2
Refalisa Simarmata	0.549	0.28	0.408	0.147	C4
Butet Nainggolan	0.715	0.441	0.22	0.095	C4

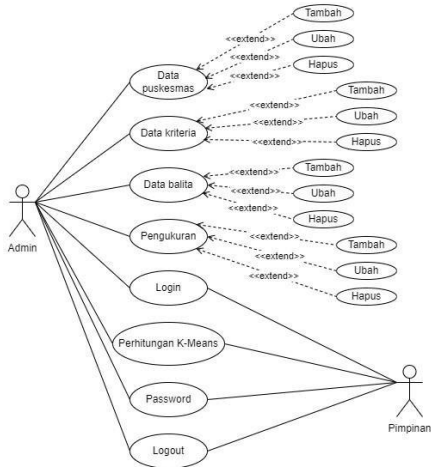
Ezra Manalu	0.472	0.204	0.477	0.187	C4
Revan Jonaris Sinaga	0.059	0.228	0.888	0.577	C1
Butet Simarmata	0.567	0.293	0.368	0.069	C4
...
Ucok Sidabutar	1.148	0.874	0.228	0.534	C3
Butet Silalahi	1.144	0.87	0.223	0.53	C3
Artaulina Sinaga	0.067	0.239	0.897	0.587	C1
Butet Siahaan	0.755	0.485	0.196	0.139	C4

Hasil akhir dari proses pengelompokan status gizi balita di puskesmas Gunungpati, iterasi berhenti pada iterasi ke-18 dimana terdapat 23.73% balita mengalami gizi kurang. Balita dengan status gizi normal terdapat 23.36%. Balita yang memiliki risiko lebih berjumlah 26.07%. Sedangkan balita dengan gizi sangat kurang terdapat 26.84%. Pada pengelompokan status gizi balita di puskesmas Kaliwiro dan batur, iterasi berhenti pada iterasi ke-15 dimana terdapat 21.54% balita mengalami gizi kurang. Balita dengan status gizi normal terdapat 22.45%. Balita yang memiliki risiko lebih berjumlah 30.21%. Sedangkan balita dengan gizi sangat kurang terdapat 25.8%. Hasil penelitian ini menunjukkan bahwa algoritma K-Means mampu mengelompokkan data balita.

B. Perancangan dan Desain Sistem

1) Use case Diagram

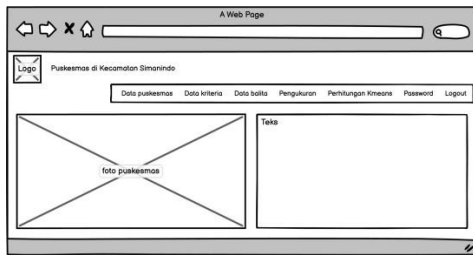
Use case dibawah ini (Gambar 1) memaparkan 2 aktor yang sedang menggunakan sistem yaitu admin dan pimpinan. Pada sistem ini terdapat 8 interaksi diantaranya *login*, data puskesmas, data kriteria, data balita, pengukuran, perhitungan *K-Means*, *password*, dan *logout*.



Gambar 1. Use case diagram

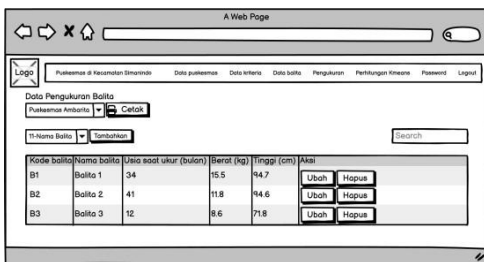
2) Desain Antarmuka

Berikut ini merupakan desain antarmuka sistem (Gambar 2).



Gambar 2. Desain antarmuka halaman utama

Gambar 3 berikut ini merupakan desain antarmuka halaman pengukuran.

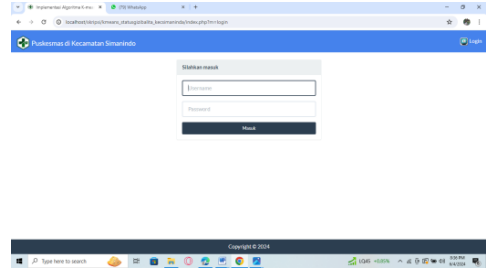


Gambar 3. Desain antarmuka halaman data pengukuran

Gambar 4 berikut ini merupakan desain antarmuka halaman pengukuran.

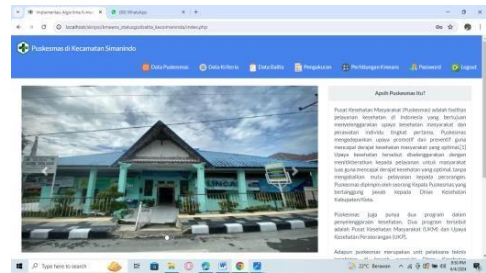
C. Hasil Website

- 1) Implementasi tampil halaman login
Halaman login (Gambar 5) menyediakan hak akses bagi admin dan pimpinan untuk menggunakan sistem. Form login ini mencakup kolom username, password, serta tombol masuk.



Gambar 5. Implementasi tampil halaman login

- 2) Implementasi tampil halaman utama
Setelah login, pengguna akan diarahkan ke halaman utama (Gambar 6). Pada halaman ini, hanya terdapat menu yang memuat informasi tentang puskesmas. Pada tampilan ini tersedia 7 interaksi yaitu: data puskesmas, data kriteria, data balita, pengukuran, perhitungan K-Means, password, dan logout.



Gambar 6. Implementasi tampil halaman utama

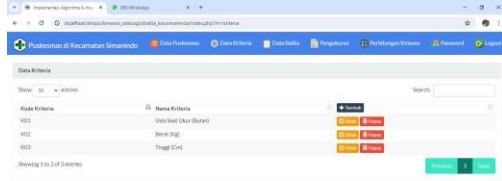
- 3) Implementasi tampil halaman data puskesmas
Halaman data puskesmas (Gambar 7) berisikan seluruh data puskesmas yang sudah di inputkan oleh admin. Admin juga dapat menambah, mengubah, menghapus, dan mencari data puskesmas.



Gambar 7. Implementasi tampil halaman data puskesmas

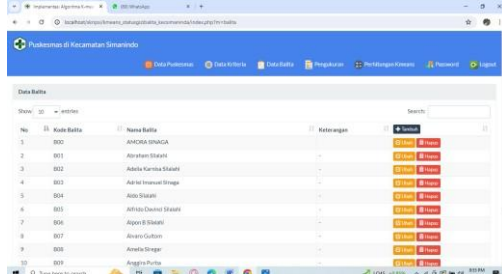
- 4) Implementasi tampil halaman data kriteria
Halaman data kriteria (Gambar 8) berisikan seluruh data kriteria yang

sudah di inputkan oleh admin. Admin juga dapat menambah, mengubah, menghapus, dan mencari data kriteria.



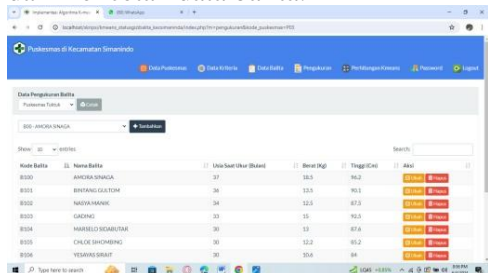
Gambar 8. Implementasi tampil halaman data kriteria

- 5) Implementasi tampil halaman data balita
Halaman data balita (Gambar 9) berisikan seluruh data balita yang sudah di inputkan oleh admin. Admin juga dapat mengubah, menambah, menghapus, dan mencari data balita.



Gambar 9. Implementasi tampil halaman data balita

- 6) Implementasi tampil halaman data pengukuran
Pada halaman ini (Gambar 10), admin terlebih dahulu memilih puskesmas yang ingin diolah. Setelah memilih puskesmas maka sistem akan menampilkan seluruh data pengukuran balita. Admin juga dapat menghapus, mengubah, menambah, mencari data, dan mencetak data balita.

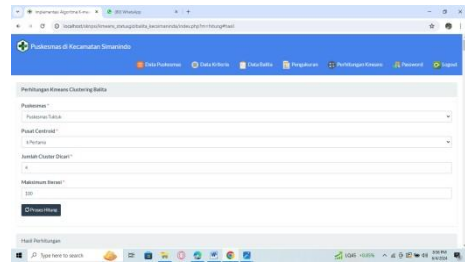


Gambar 10. Implementasi tampil halaman pengukuran

- 7) Implementasi tampil halaman perhitungan *K-Means*

Pada halaman ini (Gambar 11) admin maupun pimpinan terlebih dahulu form agar dapat melakukan perhitungan. Form tersebut terdiri dari puskesmas yang ingin diolah, menentukan centroid, mengisi jumlah cluster, dan mengisi

maksimum iterasi.



Gambar 11. Implementasi tampil halaman perhitungan *K-Means*

- 8) Implementasi tampil halaman *password*
Halaman password (Gambar 12) digunakan jika admin maupun pimpinan ingin mengganti password. Form password terdiri dari password baru, password lama, dan konfirmasi password baru.



Gambar 12. Implementasi tampil halaman *password*

9) PENUTUP

Berdasarkan hasil pengelompokan status gizi balita menggunakan algoritma *K-Means*, diperoleh bahwa pada Puskesmas Kaliwiro dan Batur terdapat 21.54% balita dengan gizi kurang, 22.45% gizi normal, 30.21% berisiko lebih, dan 25.8% gizi sangat kurang. Sementara pada Puskesmas Gunungpati terdapat 23.73% balita dengan gizi kurang, 23.36% gizi normal, 26.07% berisiko lebih, dan 26.84% gizi sangat kurang. Hasil ini menunjukkan bahwa algoritma *K-Means* dapat digunakan untuk mengelompokkan status gizi balita secara efektif. Penelitian selanjutnya diharapkan dapat mengembangkan metode ini atau mengombinasikannya dengan metode lain untuk memperoleh hasil analisis yang lebih optimal.

10) REFERENSI

- A. S. Hutasoit, P. Tarigan, and E. R. Siagian, "Implementasi Data Mining Klasifikasi Status Gizi Balita Pada Posyandu Medan Timur Dengan Menggunakan Metode C 4.5," *J. Pelita Inform.*, vol. 7, no. 2, pp. 120–125, 2018.
B. Sutomo and D. Y. Anggraini, *Menu Sehat Alami untuk Batita & Balita*, Edisi 1.

- Jakarta: Demedia, 2010.
- Y. Sunardi and S. Sunardi, Sehat Dan Pintar, Panduan Nutrisi Bagi Buah Hati, Edisi 1. Yogyakarta: Andi Publisher, 2013.
- A. M. Siregar and A. Puspabhuana, DATA MINING: Pengolahan Data Menjadi Informasi dengan RapidMiner, 1st ed. Surakarta: CV Kekata Group, 2017.
- E. M. Fitri, R. R. Suryono, and A. Wantoro, “Klasterisasi Data Penjualan Berdasarkan Wilayah Menggunakan Metode K-Means Pada Pt Xyz,” J. Komputasi, vol. 11, no. 2, pp. 157–168, 2023, doi: 10.23960/komputasi.v11i2.12582.
- A. S. Irtawaty, “Klasifikasi Penyakit Ginjal dengan Metode K-Means,” JTT (Jurnal Teknol. Terpadu), vol. 5, no. 1, pp. 49–53, 2017, doi: 10.32487/jtt.v5i1.241.
- M. F. I. Al-Rizki, I. Widaningrum, and G. A. Buntoro, “Prediksi Penyebaran Penyakit TBC dengan Metode K-Means Clustering Menggunakan Aplikasi Rapidminer,” JTERA (Jurnal Teknol. Rekayasa), vol. 5, no. 1, pp. 1–10, 2020, doi: 10.31544/jtera.v5.i1.2019.1-10.
- M. P. A. Ariawan, I. B. A. Peling, and G. B. Subiksa, “Prediksi Nilai Akhir Matakuliah Mahasiswa Menggunakan Metode K-Means Clustering (Studi Kasus: Matakuliah Pemrograman Dasar),” J. Nas. Teknol. dan Sist. Inf., vol. 9, no. 2, pp. 122–131, 2023, doi: 10.25077/teknosi.v9i2.2023.122-131.
- R. Helilintar and I. N. Farida, “Penerapan Algoritma K-Means Clustering Untuk Prediksi Prestasi Nilai Akademik Mahasiswa,” J. Sains dan Inform., vol. 4, no. 2, pp. 80–87, 2018.
- A. Febriyanto, S. Achmadi, and A. P. Sasmito, “Penerapan Metode K-Means Untuk Clustering Pengunjung Perpustakaan Itn Malang,” J. Mhs. Tek. Inform., vol. 5, no. 1, pp. 61–70, 2021.
- P. Alga Vredizon, H. Firmansyah, N. Shafira Salsabila, and W. Eko Nugroho, “Penerapan Algoritma K-Means Untuk Mengelompokkan Makanan Berdasarkan Nilai Nutrisi,” J. Technol. Informatics, vol. 5, no. 2, pp. 108–115, 2024, doi: 10.37802/joti.v5i2.577.
- H. Priyatman, F. Sajid, and D. Haldivany, “Klasterisasi Menggunakan Algoritma K-Means Clustering untuk Memprediksi Waktu Kelulusan,” JEPIN, vol. 5, no. 1, pp. 62–66, 2019.
- Silfia, R. Kurniawan, S. N. Harahap, E. Budianita, F. Syafria, and I. Iskandar, “Sistem Penentuan Penceramah Masjid Paripurna Kota Pekanbaru Menggunakan Algoritma Pengklasteran K-Means,” J. Tek. Inform., vol. 14, no. 2, pp. 162–170, 2021, [Online]. Available: <http://journal.uinjkt.ac.id/index.php/ti>
- E. W. Fridayanthie, Haryanto, and T. Tsabitah, “Penerapan Metode Prototype Pada Perancangan Sistem Informasi Penggajian Karyawan (Persis Gawan) Berbasis Web,” Paradig. - J. Komput. dan Inform., vol. 23, no. 2, pp. 151–157, 2021, doi: 10.31294/p.v23i2.10998.
- D. Purnomo, “Model Prototyping Pada Pengembangan Sistem Informasi,” JIMP-Jurnal Inform. Merdeka Pasuruan, vol. 2, no. 2, pp. 54–61, 2017.
- Kementerian Kesehatan RI, “Peraturan Menteri Kesehatan Republik Indonesia Nomor 2 Tahun 2020 tentang Standar Antropometri Anak.” Jakarta, 2020. doi: 10.1016/j.fcr.2017.06.020.