

Kombinasi K-Nearest Neighbor dengan K-Means Clustering Klasifikasi Stunting pada Bayi Berbasis Website

Muhammad Aprilsyah*, Raissa Amanda Putri

Sains dan Teknologi, Sistem Informasi, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia

Email: ^{1,*}Muhammadaprilisyah@gmail.com, ²raissa.ap@uinsu.ac.id

Email Penulis Korespondensi: Muhammadaprilisyah@gmail.com

Submitted 13-03-2025; Accepted 08-04-2025; Published 30-04-2025

Abstrak

Stunting merupakan masalah Kesehatan serius yang terjadi akibat kurangnya asupan nutrisi yang memadai dalam jangka waktu panjang, terutama pada balita. Penelitian ini bertujuan mengembangkan sistem klasifikasi data stunting berbasis web menggunakan kombinasi algoritma K-Means Clustering dan K-Nearest Neighbors (K-NN). Dataset yang digunakan berasal dari dinas Kesehatan kota medan tahun 2021-2024 dengan total 683 data. Tahapan penelitian mencakup identifikasi masalah, Data dikumpulkan melalui peninjauan dan interview, preprocessing informasi yang diterapkan dengan StandartScaler, serta pengelompokkan data menjadi 70% untuk data latih dan 30% untuk data uji. Algoritma K-Means diimplementasikan untuk mengelompokkan data berdasarkan nilai z-score. Hasil clustering digunakan sebagai label untuk klasifikasi menggunakan K-NN. Implementasi sistem menunjukkan klasifikasi dengan rasio hasil 6,9% stunting ringan, 25,8% stunting sedang, dan 67,3% stunting berat. Hasil penelitian ini mengidikasikan kombinasi K-Means dan K-NN menghasilkan klasifikasi yang lebih akurat dibandingkan metode Tunggal. Melalui Studi ini, diharapkan dapat memberikan manfaat bagi Dinas Kesehatan Kota Medan dalam menganalisis data stunting secara lebih efisien dan memberikan kontribusi pada pengembangan sistem klasifikasi stunting di masa depan.

Kata Kunci: Stunting; K-Means Clustering; K-Nearest Neighbors; Klasifikasi; Data Mining

Abstract

Stunting is a serious health issue caused by insufficient nutrition over an extended period, especially in young children. This study aims to develop a web-based stunting data classification system using a combination of K-Means Clustering and K-Nearest Neighbors (K-NN) algorithms. The dataset used is sourced from the Health Department of Medan City in 2021-2024, consisting of 683 data entries. The research process includes problem identification, data gathering conducted through observations and interviews, data preprocessing using StandardScaler, and splitting the dataset into 70% training and 30% testing datasets. The K-Means technique is utilized for data segmentation based on z-score values. The clustering results are then used as labels for classification with K-NN. The system implementation shows a classification result with a distribution of 6.9% for mild stunting, 25.8% for moderate stunting, and 67.3% for severe stunting. The results indicate that the combination of K-Means and K-NN produces more accurate classification compared to using a single method. This study is expected to assist the Health Department of Medan City in analyzing stunting data more efficiently and contribute to the future development of stunting classification systems.

Keywords: Stunting; K-Means Clustering; K-Nearest Neighbors; Classification; Data Mining

1. PENDAHULUAN

Era Society 5.0 diidentifikasi lewat penyatuan teknologi modern, seperti kecerdasan buatan (AI), robotika, blockchain, dan Sistem Perangkat Pintar, ke dalam berbagai dimensi kehidupan masyarakat. Teknologi ini tidak hanya digunakan untuk meningkatkan efisiensi dalam sektor industri dan ekonomi, tetapi juga memainkan peran penting dalam bidang kesehatan. Salah satu permasalahan kesehatan yang memerlukan perhatian serius di Indonesia adalah stunting. Stunting adalah kondisi gagla tumbuh akibat kurangnya asupan nutrisi yang memadai dalam jangka waktu lama, terutama pada masa kehamilan hingga usia 24 bulan [1]. Pemanfaatan teknologi dinilai dapat meningkatkan layanan publik yang dikelola secara menyeluruh akan menciptakan manajemen yang optimal, transparan, serta akuntabel. Ini menunjukkan bahwa pemanfaatan teknologi dalam bidang kesehatan memiliki peran dalam meningkatkan efisiensi layanan kesehatan. Teknologi berbasis data science, seperti data mining dan machine learning, berpotensi memberikan solusi inovatif untuk analisis data kesehatan. Dengan mengolah data secara sistematis, teknologi ini dapat membantu menghasilkan informasi yang bermanfaat bagi pembuatan kebijakan dan pengambilan keputusan yang lebih terarah. Data science memungkinkan pengolahan dan analisis kumpulan informasi dalam jumlah besar yang kemudian dapat menemukan pola-pola tersembunyi [2]. Data yang terkumpul kemudian diolah menjadi sebuah informasi maupun pengetahuan yang bermanfaat, memfasilitasi dalam penentuan keputusan dan pembuatan kebijakan di kemudian hari [2].

Pada studi ini menerapkan kombinasi K-Nearest Neighbors dan K-Means Clustering. Kombinasi ini memungkinkan pengklasifikasian data stunting secara efisien dan akurat, terutama dalam pengelompokkan data sesuai dengan atribut tertentu. Dengan mengkombinasikan K-NN dan K-Means bisa menjadi sebuah pendekatan yang efektif untuk klasifikasi berbasis klaster dan meningkatkan akurasi pada data yang mungkin memiliki pola cluster alami [3]. K-Nearest Neighbors adalah algoritma yang digunakan dalam pengklasifikasian terhadap data sebuah objek berdasarkan data (training) yang memiliki jarak terdekat pada objek tersebut [4]. Algoritma KNN juga merupakan algoritma yang sering menghasilkan hasil yang kompetitif dan signifikan [5]. Klasifikasi sendiri merupakan sebuah proses untuk menemukan model yang menggambarkan dan membedakan kelas dari konsep data [6]. K-Means clustering merupakan algoritma pengelompokkan [6]. Algoritma K-Means Clustering sering digunakan untuk mengelompokkan data ke dalam beberapa kluster berdasarkan kesamaan pola-pola tertentu [7]. Untuk melakukan klastering nilai k pada K-Means harus

ditentukan terlebih dahulu [8]. Teknik clustering dipilih karena mampu mengelompokkan data berdasarkan pola dan karakteristik yang serupa tanpa memerlukan label awal, sehingga sangat cocok untuk dataset yang belum terstruktur[9].

Dengan menerapkan kombinasi algoritma K-NN dan K-Means Clustering untuk pengklasifikasian data stunting. Mengklasifikasikan data dengan mengacu pada nilai z-score dari tiga indikator utama (BB/U), (TB/U) dan (BB/TB) yang kemudian score-score dari anak-anak tersebut akan diklasterkan, menggunakan K-Means. Dimana akan dipecah menjadi 3 kluster. Penentuan kluster dari tiap kelas diperoleh berdasarkan nilai nilai z-score rata-rata dari centroid masing-masing kluster. Selanjutnya setelah didapatkan rata-rata dari nilai z-score, maka kluster tersebut akan diberi label kelas. Setelah semua data latih memiliki kelas. Data latih tersebut kemudian dapat digunakan untuk proses klasifikasi data latih menggunakan K-NN untuk mengklasifikasikan data uji baru.

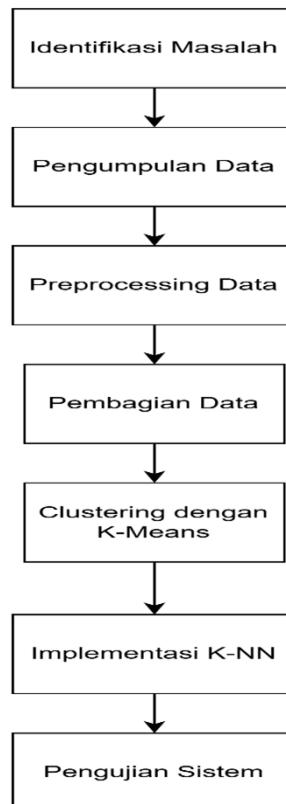
Adapun penelitian sebelumnya yang membahas k-means, k-nearest neighbors, dan stunting, penelitian dilakukan oleh Suhari Rustam dan Haditsa Annur [3] mengelompokkan kelas mata kuliah konsentrasi mahasiswa semester akhir dengan K-Means dan K-Nearest Neighbors mendapatkan akurasi 1. Tidak disebutkan persentase penggunaan data training dan data testing. Selain itu, pada penelitian lain oleh Ragil Nike Pratista dan Budhi Kristianto [10] Klasterisasi kasus stunting pada balita menggunakan K-Means mendapatkan nilai DBI (Davies-boildin Index) sebesar -0,673 mendekati 0 yang menunjukkan evaluasi cluster yang baik. Ada juga penelitian lain yang dilakukan oleh Sri Widia Pebrianti, Rini Astuti, Fadhil M Basysyar [11] dengan menggunakan K-Nearest Neighbors dalam klasifikasi status stunting balita diperoleh nilai akurasi 92% menunjukkan performa yang tinggi dalam mengklasifikasi status gizi balita. Hasil akurasi ini didapat dari pembagian data 90:10 dari jumlah data yang digunakan sebanyak 503 data.

Walaupun telah tersedia penelitian sebelumnya, namun pada penelitian sebelumnya masih memiliki kekurangan seperti cenderung menggunakan metode klasifikasi (K-NN) atau clustering (K-Means) secara terpisah. Namun, belum ada pendekatan kombinasi k-means untuk membentuk kluster data dan K-NN untuk klasifikasi data uji baru. diperlukan peningkatan pada metode klasifikasi agar lebih efisien. Dalam studi ini, kesenjangan yang ada dapat diatasi melalui penerapan metode kombinasi algoritma K-Means dan K-Nearest Neighbors. Dengan adanya kajian ini diharapkan mampu menghasilkan solusi terhadap keterbatasan data berlabel, meningkatkan efisiensi proses klasifikasi, dan mendukung intervensi gizi balita stunting secara lebih efektif. Hasil kajian ini dapat dijadikan pedoman dalam membangun sistem klasifikasi data stunting.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Adapun tahapan penelitian yang dilakukan pada penelitian ini sebagai berikut:



Gambar 1. Tahapan Penelitian

Pada gambar 1 diatas dijelaskan tahap-tahap pada penelitian ini menggunakan kajian yang mengadopsi metode penelitian kuantitatif. Penelitian kuantitatif adalah penelitian yang banyak menggunakan angka-angka mulai dari pengumpulan data, penafsiran, sampai pada hasil atau penarikan kesimpulannya. Penelitian ini dilaksanakan melalui beberapa tahapan yang terstruktur dan sistematis guna memastikan bahwa setiap langkah yang dilakukan akan menghasilkan data dan temuan yang akurat serta dapat digunakan. Tahapan-tahapan ini dirancang dengan cermat untuk menjawab tujuan penelitian secara efektif, dimulai dari perencanaan penelitian, studi literatur, analisis data, perancangan sistem, pengujian. Adapun penjelasan pada Gambar 1 mengenai metode penelitian yang digunakan pada penelitian ini sebagai berikut :

a. Identifikasi Masalah

Pada tahapan identifikasi masalah peneliti mengamati dan mengobservasi ke tempat objek penelitian. Dimana ditemukan permasalahan tentang stunting pada bayi yang merupakan masalah kesehatan yang serius yang membutuhkan identifikasi dan klasifikasi dini untuk mencegah dampak jangka panjang. Oleh karena itu, diperlukan sistem berbasis web yang memanfaatkan kombinasi algoritma K-Means dan K-NN untuk membantu dalam menganalisis data dan menentukan status stunting secara efektif.

b. Study Literatur

Penulis dalam bagian ini melakukan serangkaian proses yang mencakup metode pengumpulan data dan referensi pustaka untuk mendukung penelitian. Upaya ini melibatkan pencarian berbagai sumber tertulis yang sudah ada sebelumnya serta analisis teori yang relevan dengan permasalahan dan metode yang digunakan.

c. Pengumpulan Data

Dalam mengumpulkan data, peneliti melaksanakan 2 tahapan, dengan melakukan peninjauan dan interview. Melalui peninjauan, peneliti berhasil memperoleh data yang tepat dan relevan. selain itu, interview dengan staf bagian gizi dilakukan guna mendapatkan informasi mendetail tentang data stunting yang nanti akan di implementasikan sebagai dasar landasan penelitian ini.

d. Preprocessing Data

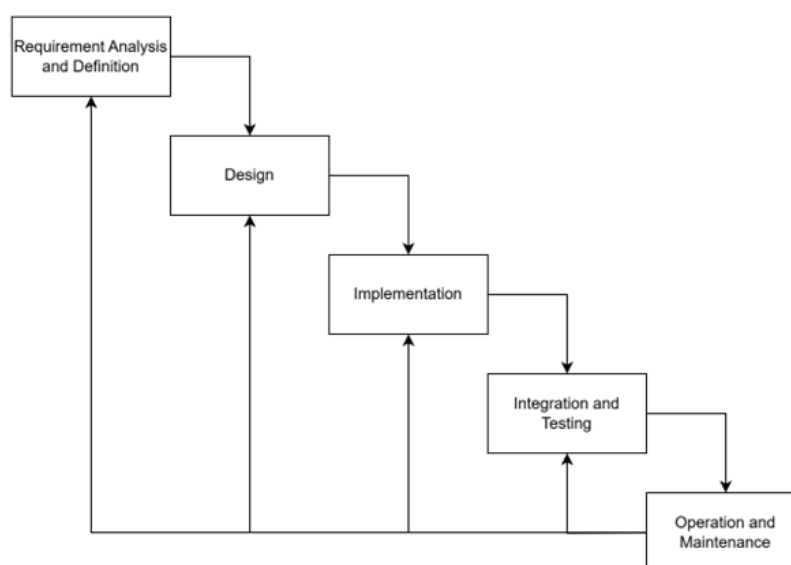
Setelah data berhasil didapatkan, langkah selanjutnya adalah melakukan pengolahan data asli sehingga menghasilkan sebuah data baru. Data asli yang diambil sering kali ditemukan ketidaksesuaian data, gangguan informasi, noise, serta beberapa nilai yang tidak tersedia. Preprocessing data merupakan proses awal yang melibatkan pembersihan dan persiapan data mentah agar sesuai dengan tahapan analisis selanjutnya[5].

Tujuan dari preprocessing dalam teknik data mining yaitu mengubah data mentah menjadi format yang lebih efisien untuk diproses dan lebih efektif untuk memenuhi kebutuhan pengguna. Setelah data asli diolah, tahap selanjutnya adalah pengolahan data lebih lanjut.

e. Pembagian Data

Pembagian data dilakukan dengan maksud untuk mengevaluasi performa model secara objektif dan mengevaluasi kemampuan model agar dapat berfungsi secara optimal pada data yang telah tersedia maupun data yang belum pernah digunakan sebelumnya. Dalam penelitian ini, data dibagi menjadi data menjadi data pelatihan dan data pengujian. Tujuan pembagian data ini agar model yang dibuat tidak hanya baik dalam mengenali pola pada data latih. Tetapi juga dapat memberikan hasil yang akurat saat digunakan untuk data baru di dunia nyata. Pada penelitian ini pembagian dilakukan dengan rasio 70:30 (70% data training, 30% data testing).

2.2 Metode Pengembangan Sistem



Gambar 2. Metode Waterfall

Pada gambar 2 diatas dijelaskan bahwa Penelitian ini mengimplementasikan model *waterfall* sebagai pendekatan yang diterapkan dalam proses pengembangannya, Metode *waterfall* didefinisikan sebagai pendekatan dalam pengembangan perangkat lunak yang dilakukan secara terstruktur, dimana setiap tahap harus diselesaikan sepenuhnya sebelum melangkah ke fase berikutnya. Tujuan utama dari pendekatan *waterfall* yaitu untuk memberikan pendekatan alur hidup perangkat lunak sekuensial atau terurut yang dimulai dengan analisis, desain, pengkodean, pengujian, dan pendukung (support). Berikut penjelasan metode pengembangan sistem *Waterfall* pada gambar 2 diatas yaitu :

a. Analisis Kebutuhan

Fase ini berfungsi sebagai langkah awal dalam proses perancangan sistem. Pada fase ini, dilakukan observasi penelitian pada dinas kesehatan kota medan. Dalam tahap ini, peneliti mengidentifikasi permasalahan serta isu-isu yang berkaitan dengan data stunting, serta menganalisa kebutuhan sistem klasifikasi stunting.

b. Design

Pada tahap ini, fokus utama adalah merancang sistem dan alur kerja yang akan dikembangkan. Setelah data yang dibutuhkan terkumpul, langkah berikutnya meliputi perancangan database, antarmuka, serta alur sistem yang akan dibangun

c. Implementation

Pada tahap ini mengimplementasikan hasil desain yang telah dibuat ke dalam website. Dalam hal ini implementasikan yang diterapkan peneliti, visual studio code sebagai media editor kode. Peneliti juga menggunakan Xampp sebagai penyimpanan dan mengelola basis data. Sistem mulai dikembangkan sesuai rancangan yang telah dibuat. Hal ini diawali dengan pengkodean program dengan menggunakan (html,php, dan css) dan pembuatan database (mysql) [12]

d. Integration and Testing

Fase ini akan menguji sistem, Efektivitas sistem atau aplikasi dalam menemukan keteraturan atau pola-pola klasifikasi data stunting pada Dinas Kesehatan Kota Medan.

e. Operation and Maintenance

Pada tahap ini, mungkin bisa terjadi sebuah pemeliharaan dari sistem yang akan dibuat, namun dalam aspek penelitian ini mungkin tidak dibutuhkan. Proses ini akan terealisasi apabila terjadi kesepakatan dengan pihak Dinas Kesehatan Kota Medan.

2.3 K-Means

Fase ini berfokus pada penerapan algoritma K-Means Clustering. Algoritma K-Means juga dapat diartikan sebagai algoritma mengklasifikasikan sejumlah data ke dalam kelompok-kelompok tertentu (cluster) [13], K-Means termasuk dalam algoritma yang memerlukan jumlah klaster (k)[14], menerapkan algoritma k-means dalam pengelompokkan data bayi ke kategori awal serta menentukan jumlah kluster (k), dan menghasilkan cluster untuk digunakan sebagai acuan pada klasifikasi selanjutnya. Algoritma K-Means adalah metode clustering yang mengelompokkan data ke dalam beberapa cluster berdasarkan kedekatan jarak dan algoritma ini beraktivitas pada attribute numerik[15][16]. K-Means akan memilih beberapa komponen dari populasi sebagai pusat awal cluster. Pada tahap ini, cluster akan ditentukan secara acak dari himpunan populasi data. Clustering juga dapat digunakan untuk mengelompokkan data yang kelasnya belum diketahui dan berfungsi dalam memberikan label pada data yang belum memiliki kelompok tertentu[17]. Algoritma K-Means akan menguji setiap elemen dalam populasi data dan mengelompokkan ke cluster yang paling sesuai berdasarkan jarak minimum pusat cluster [15][18]. Perhitungan jarak antara objek dan masing-masing centroid dapat dilakukan dengan metode Euclidean Distance [19][20]. Rumus mencari Euclidian Distance dapat dilihat dibawah ini :

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Keterangan :

d (x,y) : Jarak data dari x dan y
x : titik data pertama
y : titik data kedua
n : jumlah atribut dari data

2.4 K-Nearest Neighbors

Pada tahap ini merupakan tahapan dalam mengklasifikasi data bayi berdasarkan kategori cluster dari K-Means. Melatih model K-NN menggunakan data cluster yang dihasilkan K-Means, dan mengklasifikasi data bayi baru ke kategori ringan, sedang, berat. K-Nearest Neighbors merupakan algoritma dalam supervised learning yang berfungsi untuk tugas klasifikasi dan regresi [21]. Pendekatan K-NN dilakukan dengan mencari jarak minimum antara data yang diuji dengan sejumlah data terdekat dalam dataset pelatihan[5]. Algoritma K-NN memiliki kelebihan di antaranya ketahanan pada data training yang memiliki banyak noise dan data dalam jumlah yang besar[22]. Pengukuran jarak yang paling sering digunakan adalah Euclidean Distance[22][23]. Rumus dalam pengukuran jarak Euclidean Distance dapat dilihat dibawah ini:

$$d_i = \sqrt{\sum_{i=1}^p (x_1 - x_2)^2} \quad (2)$$

Keterangan :

d : jarak
i : variable data
p : dimensi data
x1 : sample data
x2 : data uji

Proses penghitungan nilai akurasi dapat dilakukan dengan menggunakan persamaan dibawah ini:

$$akurasi = \frac{Jumlah\ klasifikasi\ benar}{Jumlah\ data\ uji} \times 100\% \quad (3)$$

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Sumber data yang digunakan dalam penelitian ini berasal dari hasil observasi dan wawancara terhadap staff gizi dinas kesehatan kota medan. Dengan data yang didapatkan berjumlah 683 data stunting, penulis hanya memilih kolom yang sesuai dengan atribut yang akan dibutuhkan, atribut yang dipilih meliputi jenis kelamin, usia, BB, TB, TB/U. data ini merupakan data yang didapatkan berdasarkan data stunting 2021-2024. Adapun data yang didapatkan disajikan dalam tabel 1 berikut.

Tabel 1. Data Stunting 2021-2024

No	Jenis Kelamin	Usia	BB	TB	TB/U
1	L	15	6,6	68	SANGAT PENDEK
2	L	9	5,8	64	SANGAT PENDEK
3	P	17	8,2	74	PENDEK
4	L	16	5,6	65	SANGAT PENDEK
5	P	15	5,9	67	SANGAT PENDEK
6	P	22	7,1	73,5	SANGAT PENDEK
7	P	19	7,4	71,5	SANGAT PENDEK
8	P	16	6	68,7	SANGAT PENDEK
9	P	16	5,6	63	PENDEK
10	P	21	7,8	76	SANGAT PENDEK
..	..	..	..	..	..
..	..	..	..	..	..
683	P	22	10	78,5	PENDEK

Tabel 1 menunjukkan hasil dari data stunting 2021-2024. Data diatas merupakan data stunting yang ada pada dinas Kesehatan kota medan dari tahun 2021-2024

3.1 Pre Processing Dataset

Dari data yang sudah diperoleh dari Dinas Kesehatan Kota Medan, selanjutnya akan dilakukan pre processing dataset, standartscaler.

a. StandartScaler

Preprocessing data akan menggunakan standartscaler, Dimana skala nilai dalam dataset akan disesuaikan. langkah ini dilakukan bertujuan untuk menyelaraskan skala nilai dalam dataset agar tidak ada variable dengan skala besar yang mendominasi perhitungan. Standartscaler dinilai perlu dilakukan guna mencukupi hipotesis dan memperkuat efisiensi algoritma. Berikut adalah data sebelum dan setelah diproses menggunakan standartscaler.

Tabel 2. Hasil StandartScaler

Atribut	Nilai Asli	Nilai Standaartscaler
Umur	15	-1,49469
BB	6,6	-0,79565
TB	5,8	-1,11144

Dari tabel 2 didapatkan Hasil standartscaler atau nilai asli dataset, terdapat perbedaan nilai yang cukup signifikan, sehingga diperlukan penggunaan StandartScaler untuk mencegah dominasi nilai besar dalam perhitungan, yang akan mengakibatkan hasil kurang akurat. Langkah pertama yang dapat dilakukan adalah memahami dataset secara menyeluruh agar menghasilkan pemahaman yang baik mengenai variable-variable yang digunakan. Selanjutnya, dilakukan perhitungan rata-rata (mean) serta standar deviasi (Standard Deviation) pada setiap fitur dalam dataset.

b. Pembagian data latih dan data uji

Dalam studi ini, dataset dipisahkan menjadi dua bagian, yaitu data training dan data testing dengan komposisi 70% untuk data training dan 30% untuk data testing. Pemisahan ini bertujuan untuk menghasilkan klasifikasi stunting pada balita secara optimal.

c. Clustering K-Means

Tabel 3. Data Hasil K-Means

No	Nama	Z-BB	Z-TB	Z-IMT	Kelas
1	Cluster-1	-2.93	-7.34	0.11	Stunting Berat
2	Cluster-2	-2.45	-4.42	0.12	Stunting Sedang
3	Cluster-3	7.76	-7.17	0.25	Stunting Ringan
4	Cluster-4	14.54	-61.64	2.60	Stunting Berat
5	Cluster-5	-4.50	-12.58	0.10	Stunting Berat

Dari tabel 3 didapatkan Hasil uji coba K-Means, Jumlah kluster yang digunakan adalah 5 (lima) yaitu K=5, dimana didapatkan attribute cluster Cluster-1, Cluster-2, Cluster-3, Cluster-4, Cluster-5. Merupakan attribute klustering yang dimiliki. Dataset yang digunakan sebagai data latih berjumlah 159, pada C1 berjumlah = 82, pada C2 berjumlah = 41, pada C3 berjumlah = 11, pada C4 berjumlah = 2, dan pada C5 berjumlah = 23. Nilai tiap cluster akan berubah tergantung ketentuan jumlah kluster yang digunakan. Pada penelitian ini menetapkan penggunaan kluster sebanyak 5. Hasil penilaian ini berdasarkan rumus Z-Score yang digunakan berdasarkan WHO (World Health Organization).

d. Klasifikasi K-NN

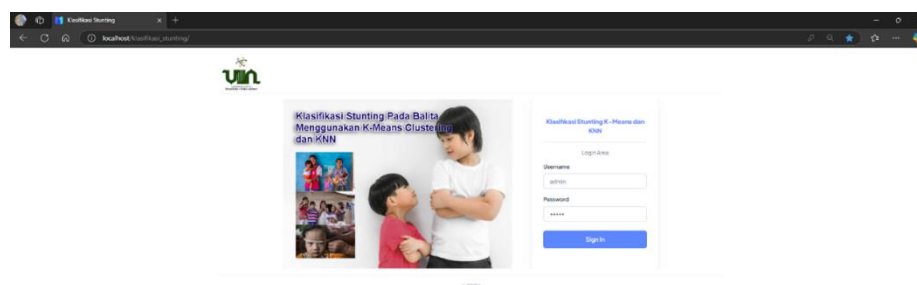
Tabel 4. Data Hasil Klasifikasi K-NN

No	Nama	Jenis Kelamin	Usia	BB	TB	Z-BB	Z-TB	Z-IMT	Kelas
1	Teguh Arya Syahputra	L	15	6,6	68	-2.90	-7.9	0.09	Stunting Berat
2	Yuna	P	17	8,2	74	-1,5	-4	0.11	Stunting Sedang
3	Zahra Fatimah Maulana	P	21	21	70	9.54	-9.49	0.3	Stunting Ringan
4	Michael Jordan	L	36	8,6	11	-4	-59.5	0.78	Stunting Berat
5	Khairi Zahfran	L	16	5,6	65	-4.1	-10.7	0.08	Stunting Berat

Dari tabel 4 didapatkan hasil Klasifikasi K-NN dengan Data yang sudah dicluster sebelumnya oleh K-Means dengan ketentuan K=5, maka didapatkan hasil pengklasifikasian sesuai dengan Tabel 4. Dimana hasil yang tertera diambil dari sample yang sudah dicluster kan. Dimana ketentuan jumlah K Tetangga yang digunakan K=5 juga seperti pada saat pengclusturan. Sehingga akan menampilkan hasil yang sesuai dengan cluster yang ada. Dari 5 data yang ada tiap data akan diklasifikasikan sesuai dengan C1,C2,C3,C4,C5 sesuai dengan status kelas nya.

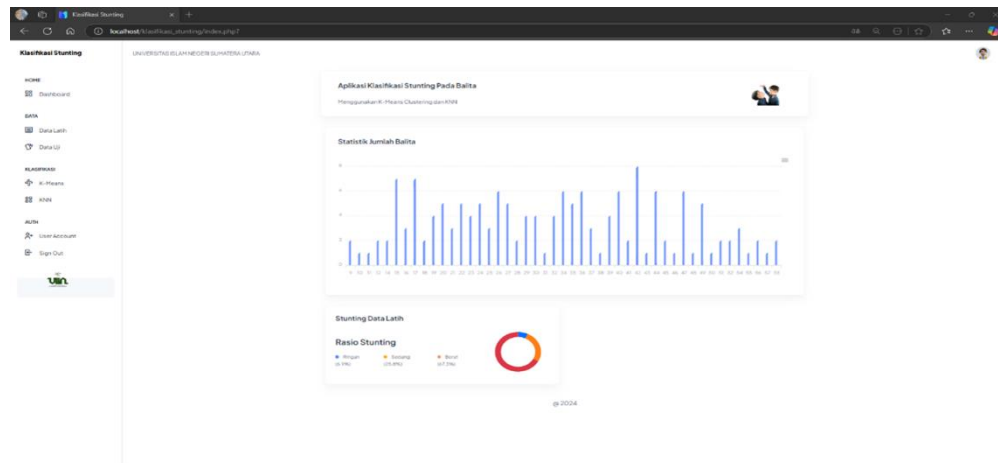
3.3 Implementasi

Setelah dilakukan preprocessing data, pemisahan data latih dan data uji, serta Clustering K-Means dan Klasifikasi K-NN. tahap selanjutnya mengimplementasikan ke dalam sistem yang dibangun. Berikut hasil pengimplementasian dapat diamati pada ilustrasi dibawah ini :



Gambar 3. Visualisasi Login

Pada gambar 3 visualisasi login, user masuk menggunakan akun yang telah terdaftar sebelumnya. Disini user diharapkan dapat masuk menggunakan nama pengguna dan kata sandi yang telah tersimpan dalam database.



Gambar 4. Visualisasi Dashboard

Gambar 4. menampilkan setelah User berhasil login. Hal pertama yang akan ditampilkan pada website merupakan tampilan dashboard. Dimana pada tampilan ini akan menampilkan hasil dari rasio stunting yang sudah ada berdasarkan data uji dan data latih yang sudah pernah dilakukan.

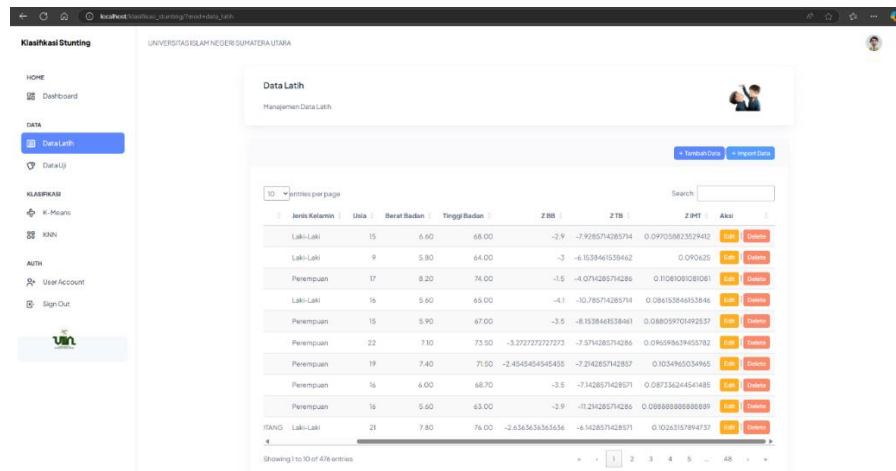
No	Nama	Jenis Kelamin	Usia	Berat Badan	Tinggi Badan	Z Score	Z TB
1	Teguh Aya Syahputra	Laki-Laki	15	6.60	68.00	-2.9	-7.928574285714
2	Rahmad Alfa Rizky	Laki-Laki	9	5.80	64.00	-3	-6.153846153846
3	Yuna	Perempuan	17	8.20	74.00	-1.5	-4.074285714286
4	Khairi Zuhman	Laki-Laki	16	5.60	65.00	-4.1	-10.785714285714
5	Arum Naasha Rizeta	Perempuan	15	5.90	67.00	-3.5	-8.153846153846
6	Dania Ramadhani Nur	Perempuan	22	7.10	73.50	-3.727272727273	-7.574285714286
7	Iqbal Nur Zulakha	Perempuan	19	7.60	71.50	-2.454545454545	-7.214285714286
8	Syafika Salamuddin	Perempuan	16	6.00	68.70	-3.5	-7.142857142857
9	Judeha	Perempuan	16	5.60	63.00	-3.9	-11.214285714286
10	FALSA GICMARSEL SHOTRANG	Laki-Laki	21	7.80	76.00	-2.636363636364	-6.142857142857

Gambar 5. Visualisasi Data Latih

Gambar 5. memperlihatkan visualisasi data latih, akan menampilkan data latih berdasarkan data yang sudah di input. Pada data latih akan terlihat button tambah data. Dimana Ketika button tambah data akan menampilkan tampilan untuk menambahkan data stunting. Seperti gambar dibawah ini:

Gambar 6. Visualisasi Input Data Latih

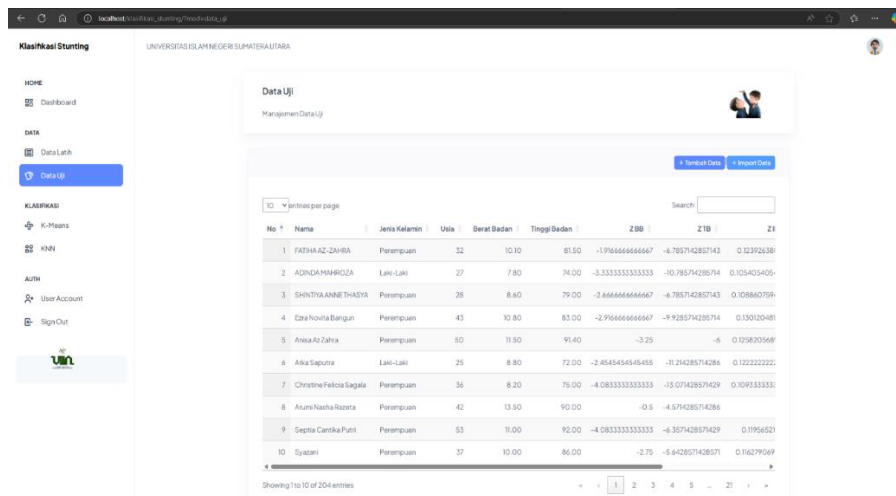
Gambar 6 menampilkan pada Visualisasi input data latih, user bisa mengisi pada kolom nama, jenis kelamin, usia, berat badan, tinggi badan. Setelah user mengisi nama, jenis kelamin, usia, berat badan, dan tinggi badan. Selanjutnya akan tersimpan pada database jika menekan tombol simpan. Namun jika ingin Kembali bisa menekan tombol Kembali, dan akan Kembali pada tampilan data latih seperti gambar 3.



No	Nama	Jenis Kelamin	Usia	Berat Badan	Tinggi Badan	Z Skor	T Skor	Aksi
1	Laki-Laki	15	6.60	68.00	-2.9	-7.928574285714	0.097058623029402	Edit Delete
2	Laki-Laki	9	5.80	64.00	-3	-4.1538461538462	0.090425	Edit Delete
3	Perempuan	17	8.20	76.00	-1.5	-4.07428574286	0.1081081081081	Edit Delete
4	Laki-Laki	16	5.60	65.00	-4.1	-10.78574285714	0.08452846835944	Edit Delete
5	Perempuan	15	5.90	67.00	-3.5	-8.1538461538461	0.088059701492537	Edit Delete
6	Perempuan	22	7.10	73.50	-3.2727272727273	-7.57428574286	0.096598479453782	Edit Delete
7	Perempuan	19	7.40	71.50	-2.4545454545455	-7.24285742857	0.1034965034965	Edit Delete
8	Perempuan	16	6.00	68.70	-3.5	-7.142857428571	0.08735244544485	Edit Delete
9	Perempuan	16	5.60	63.00	-3.9	-11.21428574286	0.088888888888889	Edit Delete
10	Laki-Laki	21	7.80	76.00	-2.4545454545455	-4.142857428571	0.10235157694733	Edit Delete

Gambar 7. Visualisasi Edit & Delete

Gambar 7 menampilkan pada Visualisasi data latih, disediakan juga tombol untuk mengedit data dan menghapus data. Pada tampilan mengedit data dipergunakan apabila data yang sudah di input kedalam database ingin dirubah. Hal ini bisa dilakukan jika menekan button edit, dan kemudian untuk fitur delete bisa digunakan apabila data yang sudah ada tidak lagi dibutuhkan atau ingin dihapus. Bisa menekan button delete.



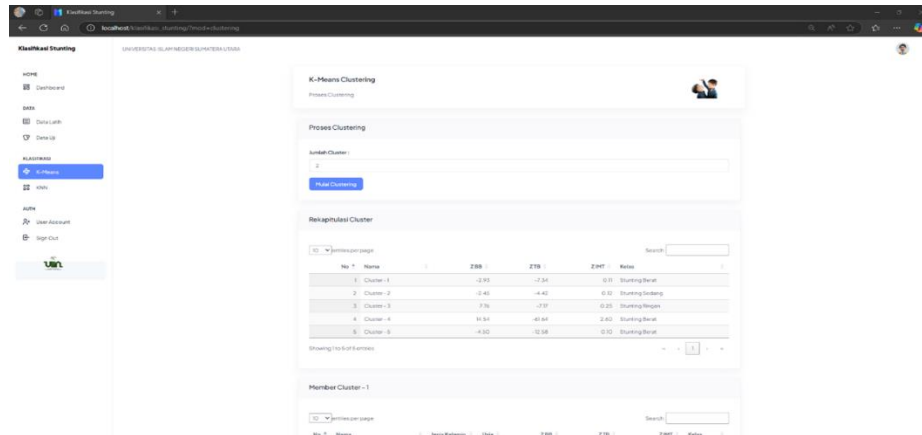
No	Nama	Jenis Kelamin	Usia	Berat Badan	Tinggi Badan	Z Skor	T Skor	Aksi
1	FATHA AZ-ZAHRA	Perempuan	32	10.10	81.50	-1.9166666666667	-6.785742857143	0.12392638
2	ADINDA NARHIZA	Laki-Laki	27	7.80	74.00	-3.3333333333333	-10.78574285714	0.103401405
3	SHAKTIVA ANNE THADYA	Perempuan	28	8.40	79.00	-2.6666666666667	-6.785742857143	0.108660756
4	Eta Novila Bangun	Perempuan	43	10.80	83.00	-2.9166666666667	-9.928574285714	0.130100481
5	Azka Az Zahra	Perempuan	50	11.50	91.40	-3.25	-6	0.1258205969
6	Azka Saputra	Laki-Laki	25	8.80	72.00	-2.4545454545455	-11.21428574286	0.122222222
7	Christine Felicia Sagala	Perempuan	36	8.20	75.00	-4.0833333333333	-13.07428574286	0.109333333
8	Acuna Nasha Razata	Perempuan	42	13.50	90.00	-0.5	-4.57428574286	
9	Septia Cernika Putri	Perempuan	53	11.00	92.00	-4.0833333333333	-6.357428574286	0.11954521
10	Syazani	Perempuan	37	10.00	86.00	-2.75	-5.42857428571	0.116279569

Gambar 8. Visualisasi Data Uji

Gambar 8 menampilkan pada visualisasi data uji, akan menampilkan data uji berdasarkan data yang sudah di input. Pada data uji akan terlihat button tambah data. Dimana Ketika button tambah data akan menampilkan tampilan untuk menambahkan data stunting.

Pada visualisasi input data latih, user bisa mengisi pada kolom nama, jenis kelamin, usia, berat badan, tinggi badan. Setelah user mengisi nama, jenis kelamin, usia, berat badan, dan tinggi badan. Selanjutnya akan tersimpan pada database jika menekan tombol simpan. Namun jika ingin Kembali bisa menekan tombol Kembali, dan akan Kembali pada tampilan data uji seperti pada data latih.

Pada visualisasi data uji, disediakan juga tombol untuk mengedit data dan menghapus data. Pada tampilan mengedit data dipergunakan apabila data yang sudah di input kedalam database ingin dirubah. Hal ini bisa dilakukan jika menekan button edit, dan kemudian untuk fitur delete bisa digunakan apabila data yang sudah ada tidak lagi dibutuhkan atau ingin dihapus. Bisa menekan button delete.



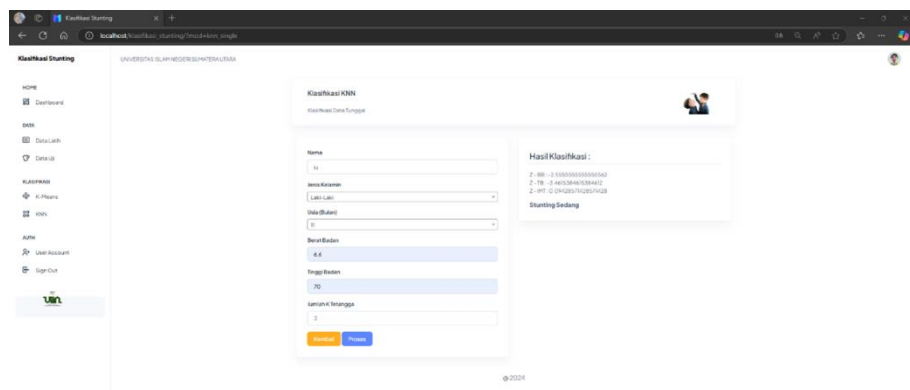
Gambar 9. Visualisasi K-Means

Gambar 9 menampilkan pada visualisasi K-Means, disediakan fitur untuk memasukkan jumlah cluster yang ingin digunakan, pada penelitian ini, ditentukan penggunaan jumlah cluster 5, sehingga didapatkan rekapitulasi cluster sebanyak 5. Dimana hasil rekapitulasi ini berdasarkan z-score yang ada pada WHO, dan akan ditampilkan pada table dibawah ini

Tabel 5. Rekapitulasi Cluster

No	Nama	Z-BB	Z-TB	Z-IMT	KELAS
1	Cluster – 1	-2.93	-7.34	0.11	Stunting Berat
2	Cluster – 2	-2.45	-4.42	0.12	Stunting Sedang
3	Cluster – 3	7.76	-7.17	0.25	Stunting Ringan
4	Cluster – 4	14.54	-61.64	2.60	Stunting Berat
5	Cluster - 5	-4.50	-12.58	0.10	Stunting Berat

Hasil pada tabel 5 Didapatkan berdasarkan jumlah data yang ada yang kemudian akan dihitung berdasarkan Z-Score masing-masing kolom yang digunakan, seperti Z-BB, Z-TB, Z-IMT, kemudian akan dicluster kan kedalam kelas Stunting ringan, Stunting Sedang, Stunting Berat. Data pada Rekapitulasi Cluster dapat berubah tergantung penggunaan jumlah Cluster, tentu hal ini akan mempengaruhi dan merubah Z-BB, Z-TB, Z-IMT, dan kelas tiap stunting nya.



Gambar 10. Visualisasi K-NN

Pada Visualisasi K-NN akan menggunakan data uji, Dimana data uji ini akan mengklasifikasikan data yang sebelumnya sudah dicluster pada data uji. Kemudian user bisa menggunakan data uji untuk menentukan klasifikasi stunting. Pada tampilan K-NN user bisa mengisi Nama, Jenis Kelamin, Usia (Bulan), Berat Badan, Tinggi Badan, Jumlah K Tetangga.

3.4 Hasil

Pada penelitian yang telah dilakukan, Algoritma K-Means Clustering dan K-Nearest Neighbors mampu dikombinasikan sehingga menghasilkan klasifikasi yang lebih baik. Dari hasil ini didapatkan nilai rasio stunting berdasarkan jumlah kluster yang digunakan sebanyak 5. Didapatkan hasil ratio 6,9% (stunting ringan), 25,8% (stunting sedang), dan 67,3% (stunting berat). Dengan begitu penerapan kombinasi algoritma K-Nearest Neighbors dan K-Means Clustering lebih baik dibandingkan tanpa menerapkan kombinasi.

Batasan penelitian ini melibatkan dataset yang digunakan berupa data yang didapatkan dari Dinas Kesehatan Kota Medan, data tersebut merupakan data stunting 2021-2024. Dimana dataset yang digunakan melibatkan ukuran dataset 683, attribute yang digunakan, Serta aspek lain yang dapat memengaruhi akurasi klasifikasi data. Kendala ini harus

diperhitungkan dalam menafsirkan hasil. Penelitian ini berkontribusi terhadap pengembangan sistem klasifikasi data stunting balita. Dampaknya dapat dimanfaatkan dalam proses klasifikasi data stunting, terutama di wilayah kota medan

4. KESIMPULAN

Dari pengujian yang telah dilakukan, menunjukkan bahwa algoritma K-Nearest Neighbors dan K-Means Clustering memiliki hasil yang baik. Dimana hasil dari clustering dapat di implementasikan pada pengklasifikasian K-NN, sehingga data sebelumnya yang belum dilabel, dapat diberi label oleh K-Means berdasarkan centroid. Kemudian akan di klasifikasikan berdasarkan data yang sudah diberi label oleh K-NN. Dengan pengujian ditentukan kluster sebanyak 5, maka didapatkan hasil ratio 6,9%(stunting ringan), 25,8%(stunting sedang), dan 67,3%(stunting berat). Dengan begitu penerapan kombinasi K-Means dan K-Nearest Neighbors dalam klasifikasi data stunting mampu menghasilkan output yang lebih optimal. Melalui penelitian ini berpotensi memberikan pengetahuan bagi pengembangan riset penelitian dimasa mendatang yang akan menggunakan kombinasi algoritma ini sebagai refrensi dalam proses klasifikasi. Selain itu, penelitian ini juga diharapkan dapat membantu dinas kesehatan dalam mengklasifikasi data stunting yang ada dan memberikan efisiensi dalam pengklasifikasiannya.

REFERENCES

- [1] Hartati, Asrida, Leli, and A. Jariyah, "Penyuluhan Upaya Pencegahan Stunting Pada Anak," *J. Hum. Educ.*, vol. 4, no. 3, pp. 213–218, 2024.
- [2] B. Kabupaten *et al.*, "Website: <https://j-innovative.org/index.php/Innovative> Implementasi Algoritma K-Means Untuk Klasterisasi Data Stunting," *Innov. J. Soc. Sci. Res.*, vol. 4, pp. 363–373, 2024.
- [3] S. Rustam and H. Annur, "Akademik Data Mining (Adm) K-Means Dan K-Means K-Nn Untuk Mengelompokan Kelas Mata Kuliah Kosentrasi Mahasiswa Semester Akhir," *Ilk. J. Ilm.*, vol. 11, no. 3, pp. 260–268, 2019.
- [4] D. T. Worung, S. R. U. A. Sompie, and A. Jacobus, "Implementasi K-Means dan K-NN pada Pengklasifikasian Citra Bunga," *J. Tek. Inform.*, vol. 15, no. 3, pp. 217–222, 2020.
- [5] Nikmatun, I. Alvi, Waspada, and Indra, "Implementasi Data Mining Untuk Klasifikasi Masa Studi Mahasiswa Menggunakan Algoritma K-Nearest Neighbor," *J. SIMETRIS*, vol. 10, no. 2, pp. 421–432, 2019.
- [6] T. Widyanti, S. S. Hilabi, A. Hananto, Tukino, and E. Novalia, "Implementasi K-Means dan K-Nearest Neighbors pada Kategori Siswa Berprestasi," *J. Inf. dan Teknol.*, vol. 5, no. 1, pp. 75–82, 2023.
- [7] M. Silalahi, "Analisis Clustering Menggunakan Algoritma K-Means Terhadap Penjualan Produk Padapt Batamas Niaga Jaya," *Comput. Based Inf. Syst. J.*, vol. 6, no. 2, pp. 20–35, 2018.
- [8] N. T. Luchia, H. Handayani, F. S. Hamdi, D. Erlangga, and S. F. Octavia, "Perbandingan K-Means dan K-Medoids Pada Pengelompokan Data Miskin di Indonesia," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 2, no. 2, pp. 35–41, 2022.
- [9] L. G. Rady Putra and A. Anggrawan, "Pengelompokan Penerima Bantuan Sosial Masyarakat dengan Metode K-Means," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 21, no. 1, pp. 205–214, 2021.
- [10] R. N. Pratistha and B. Kristianto, "Implementasi Algoritma K-Means dalam Klasterisasi Kasus Stunting pada Balita di Desa Randudongkal," *J. Indones. Manaj. Inform. dan Komun.*, vol. 5, no. 2, pp. 1193–1205, 2024.
- [11] S. Widia Pebrianti, R. Astuti, and F. M. Basysyar, "Penerapan Algoritma K-Nearest Neighbor Dalam Klasifikasi Status Stunting Balita Di Desa Bojongemas," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 2, pp. 2479–2488, 2024.
- [12] I. R. Rohaturohman, A. Faizin, and T. Y. Prawira, "Perancangan Dan Implementasi Sistem Informasi Pemesanan Sepatu Di Toko Tasik Salem Berbasis Web," *J. Tek. Inform. Dan Sist. Inf.*, vol. 2, no. 2, pp. 79–86, 2022.
- [13] A. Yahya and R. Kurniawan, "Implementation of K-Means Algorithm for Clustering Sales Data Based on Sales Patterns Implementasi Algoritma K-Means untuk Pengelompokan Data Penjualan Berdasarkan Pola Penjualan," vol. 5, no. January, pp. 350–358, 2025.
- [14] R. Kurniawan, M. S. Hasibuan, and R. Hasibuan, "Klasterisasi Wilayah Prioritas Vaksin Menggunakan Algoritma K-Means Clustering," *Media Online*, vol. 4, no. 3, pp. 1585–1592, 2023.
- [15] G. Sonia and R. A. Putri, "Penerapan Metode K-Means Clustering Untuk Mengelompokkan Data Kelayakan Penerima Bantuan Renovasi Rumah," *Build. Informatics, Technol. Sci.*, vol. 5, no. 2, pp. 442–455, 2023.
- [16] R. Kurniawan and R. Dewi, "Penerapan Algoritma K-Means Clustering Dalam Persentase Merokok Pada Penduduk Umur Di Atas 15 Tahun Menurut Provinsi," *J. Sist. Komput. dan Inform. Hal*, vol. 2, no. 2, pp. 178–186, 2021.
- [17] P. Apriyani, A. R. Dikananda, and I. Ali, "Penerapan Algoritma K-Means dalam Klasterisasi Kasus Stunting Balita Desa Tegalwangi," *Hello World J. Ilmu Komput.*, vol. 2, no. 1, pp. 20–33, 2023.
- [18] Ni Komang Sri Julyantari, I. K. Budiarta, and N. M. D. K. Putri, "Implementasi K-Means Untuk Pengelompokan Status Gizi Balita (Studi Kasus Banjar Titih)," *J. Janitra Inform. dan Sist. Inf.*, vol. 1, no. 2, pp. 92–101, 2021.
- [19] H. Hartatik and R. Rosyid, "Pengaruh User Profiling Pada Rekomendasi Sistem Menggunakan K Means Dan Knn," *J. Inf. Syst. Manag.*, vol. 2, no. 1, pp. 13–18, 2020.
- [20] D. T. Saputra, "Comparison of Stunting Clusters for Each Province in Indonesia in 2019 and 2020 With 2021 and 2022 Using the K-Means Method," *J. Indones. Sos. Teknol.*, vol. 5, no. 2, pp. 384–393, 2024.
- [21] Kemal Musthafa Rajabi, W. Witanti, and Rezki Yuniarti, "Penerapan Algoritma K-Nearest Neighbor (KNN) Dengan Fitur Relief-F Dalam Penentuan Status Stunting," *Innov. J. Soc. Sci. Res.*, vol. 3, pp. 3555–3568, 2023.
- [22] S. R. Cholil, T. Handayani, R. Prathivi, and T. Ardianita, "Implementasi Algoritma Klasifikasi K-Nearest Neighbor (KNN) Untuk Klasifikasi Seleksi Penerima Beasiswa," *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 6, no. 2, pp. 118–127, 2021.
- [23] S. K. P. Loka and A. Marsal, "Perbandingan Algoritma K-Nearest Neighbor dan Naïve Bayes Classifier untuk Klasifikasi Status Gizi Pada Balita," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 3, no. 1, pp. 8–14, 2023.