

**IMPLEMENTASI CLUSTERING USING REPRESENTATIVES (CURE)
DALAM PENGELOMPOKAN STATUS MALNUTRISI PADA LANSIA**

SKRIPSI

**NAYATA SANDRA CLAUDIA NASUTION
211401012**



**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS SUMATERA UTARA
2025**

**IMPLEMENTASI CLUSTERING USING REPRESENTATIVES (CURE)
DALAM PENGELOMPOKAN STATUS MALNUTRISI PADA LANSIA**

SKRIPSI

**Diajukan untuk melengkapi tugas dan memenuhi syarat memperoleh ijazah
Sarjana Ilmu Komputer**

**NAYATA SANDRA CLAUDIA NASUTION
211401012**



**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS SUMATERA UTARA
2025**

PERSETUJUAN

Judul : IMPLEMENTASI *CLUSTERING USING REPRESENTATIVES (CURE)* DALAM PENGELOMPOKAN STATUS MALNUTRISI PADA LANSIA

Kategori : SKRIPSI

Nama : NAYATA SANDRA CLAUDIA NASUTION

Nomor Induk Mahasiswa : 211401012

Program Studi : SARJANA (S-1) ILMU KOMPUTER

Fakultas : ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

Tanggal Sidang : 9 Januari 2025

Komisi Pembimbing :

Pembimbing II

Fuzy Yustika Manik, S.Kom., M.Kom.

NIP. 198710152019032010

Pembimbing I

Dewi Sartika Br. Ginting, S.Kom., M.Kom.

NIP. 199005042019032023

Diketahui/Disetujui oleh

Program Studi S-1 Ilmu Komputer

Ketua



Dr. Amalia, S.T., M.T.

NIP. 197812212014042001

PERNYATAAN

**IMPLEMENTASI CLUSTERING USING REPRESENTATIVES
(CURE) DALAM PENGELOMPOKAN STATUS
MALNUTRISI PADA LANSIA**



Medan, 9 Januari 2025

Nayata Sandra Claudia Nasution
211401012

PENGHARGAAN

Segala puji dan syukur penulis panjatkan kehadirat Allah SWT atas limpahan rahmat, hidayah, serta karunia-Nya yang tak terhingga, sehingga penulis dapat menuntaskan tugas akhir ini sebagai salah satu syarat kelulusan pada Program Studi Ilmu Komputer, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Sumatera Utara.

Dalam proses penyusunan dan penyelesaian penelitian ini, penulis menyadari sepenuhnya bahwa keberhasilan ini tidak lepas dari bantuan, doa, dan dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih yang mendalam kepada:

1. Bapak Dr. Muryanto Amin, S.Sos., M.Si. selaku Rektor Universitas Sumatera Utara.
2. Ibu Dr. Maya Silvi Lydia B.Sc., M.Sc. selaku Dekan Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Sumatera Utara.
3. Ibu Dr. Amalia S.T., M.T. selaku Ketua Prodi S-1 Ilmu Komputer, Universitas Sumatra Utara.
4. Ibu Dewi Sartika Br. Ginting, S.Kom., M.Kom. selaku Dosen Pembimbing I yang dengan penuh kesabaran telah banyak memberikan bimbingan, arahan, serta saran yang sangat berharga selama penyusunan skripsi ini.
5. Ibu Fuzy Yustika Manik, S.Kom., M.Kom. selaku Dosen Pembimbing II yang telah memberikan waktu dan pemikiran, serta turut memberikan banyak masukan dan arahan dalam proses penyusunan skripsi ini.
6. Bapak Dr. Eng Ade Candra, S.T., M.Kom. selaku Dosen Penasihat Akademik yang selalu memberikan dukungan dan nasihat akademis selama masa perkuliahan, serta motivasi yang sangat membantu penulis dalam menyelesaikan tugas akhir ini.
7. Kak Rita Kartika Simatupang, A.M.G. dan Kak Yanti Frasasty Saragih A.M.K. selaku pakar yang sudah memberikan wawasan dan ilmu untuk menunjang penelitian ini.
8. Kedua orang tua dan saudari-saudari saya, Ibunda Indrawati, Ayahanda M. Riza Fahrozi Nasution, Kakak Shaila Fitri Hasyim Nasution dan Adik

Rasyqa Anggi Qasthari Nasution yang tak pernah berhenti memberikan perhatian, doa yang tulus, motivasi, serta dukungan yang luar biasa berharga bagi penulis sepanjang proses ini.

9. Orang-orang terdekat saya, Nurhalimah, Meysha Sabrina Az-zahra, Dea Oktavia, Andhira Dianty Sukma dan Indah Ramadhani yang telah memberikan dukungan serta bantuan tanpa kenal lelah kepada penulis.
10. Teman-teman Stambuk 2021, khususnya Kom C, yang selalu memberikan semangat dan dukungan kepada penulis.
11. Serta setiap pihak yang turut berkontribusi, baik secara langsung maupun tidak langsung, namun tidak dapat disebutkan satu per satu.



ABSTRAK

Malnutrisi adalah masalah serius yang sering terjadi pada lansia, yang dapat menyebabkan dampak negatif dalam aspek kesehatan, sosial, dan perawatan. Lansia rentan mengalami malnutrisi akibat penurunan fisiologis, keterbatasan akses makanan bergizi, dan penyakit penyerta. Gejala malnutrisi sering tidak terdeteksi dini, sehingga diperlukan pendekatan penilaian gizi yang terintegrasi untuk mengidentifikasi risiko dengan lebih tepat. Deteksi awal malnutrisi penting untuk mencegah komplikasi serius, seperti infeksi dan penurunan kualitas hidup. Dalam beberapa tahun terakhir, algoritma *machine learning* memiliki kemampuan untuk mengelompokkan data mengenai gizi. Algoritma CURE memungkinkan pengelompokan data kompleks dengan distribusi tidak merata, meningkatkan akurasi identifikasi kelompok risiko. Metode penelitian ini meliputi beberapa tahapan: pengumpulan data lansia, *pre-processing* data untuk mengatasi inkonsistensi, penerapan algoritma CURE untuk pengelompokan, serta evaluasi hasil menggunakan *Davies-Bouldin Index* dan *Silhouette Analysis*. Penelitian ini diharapkan dapat memberikan panduan yang lebih tepat dalam menentukan intervensi gizi sesuai kebutuhan kelompok lansia yang berisiko malnutrisi. Penelitian ini menggunakan dataset yang berisi 517 baris data lansia. Berdasarkan hasil analisis menggunakan metode *Silhouette Analysis*, dataset tersebut menghasilkan *cluster* optimal berjumlah 3. *Cluster* yang terbentuk adalah status gizi normal, malnutrisi ringan, dan malnutrisi berat, yang mencerminkan variasi tingkat risiko pada lansia. Algoritma CURE yang diterapkan pada dataset menghasilkan nilai *Davies-Bouldin Index (DBI)* sebesar 0.85, menunjukkan kualitas pengelompokan yang baik. Selain itu, perbandingan antara hasil deteksi sistem dan diagnosis pakar gizi menunjukkan bahwa akurasi sistem mencapai 88%, menunjukkan efektivitas sistem dalam mengidentifikasi status gizi lansia.

Kata Kunci: *clustering*, CURE, lansia, malnutrisi.

**IMPLEMENTATION OF CLUSTERING USING REPRESENTATIVES
(CURE) FOR MALNUTRITION STATUS IN THE ELDERLY**

ABSTRACT

Malnutrition is a serious problem that often occurs in the elderly, which can cause negative impacts in health, social and care aspects. The elderly are vulnerable to malnutrition due to physiological decline, limited access to nutritious food, and comorbidities. Symptoms of malnutrition are often not detected early, so an integrated nutrition assessment approach is needed to identify risks more precisely. Early detection of malnutrition is important to prevent serious complications, such as infections and reduced quality of life. In recent years, machine learning algorithms have gained the ability to cluster nutrition data. The CURE algorithm enables clustering of complex data with uneven distribution, improving the accuracy of risk group identification. The research method includes several stages: collecting data on the elderly, pre-processing the data to resolve inconsistencies, applying the CURE algorithm for clustering, and evaluating the results using the Davies-Bouldin Index and Silhouette Analysis. This research is expected to provide more precise guidance in determining nutritional interventions according to the needs of elderly groups at risk of malnutrition. This research uses a dataset containing 517 rows of elderly data.. Based on the results of the analysis using the Silhouette Analysis method, the dataset produces an optimal cluster of 3. The clusters formed are good nutritional status, mild malnutrition, and severe malnutrition, which reflect variations in the level of risk in the elderly. The CURE algorithm applied to the dataset resulted in a Davies-Bouldin Index (DBI) value of 0.85, indicating good clustering quality. In addition, a comparison between the system's detection results and the nutrition expert's diagnosis showed that the system's accuracy reached 88%, demonstrating the system's effectiveness in identifying the nutritional status of the elderly.

Keywords: clustering, CURE, elderly, malnutrition.

DAFTAR ISI

PERSETUJUAN.....	iii
PERNYATAAN.....	iv
PENGHARGAAN.....	v
ABSTRAK.....	vii
ABSTRACT.....	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL.....	xiii
BAB 1 PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	2
1.3. Batasan Masalah.....	3
1.4. Tujuan Penelitian.....	3
1.5. Manfaat Penelitian.....	4
1.6. Metodologi Penelitian.....	4
1.7. Penelitian Relevan.....	6
1.8. Sistematika Penulisan.....	8
BAB 2 LANDASAN TEORI.....	9
2.1 <i>Data Mining</i>	9
2.2 <i>Machine Learning</i>	10
2.2.1. <i>Supervised Learning</i>	10
2.2.2. <i>Unsupervised Learning</i>	11
2.2.3. <i>Semi-supervised Learning</i>	12
2.3 <i>Clustering</i>	12
2.4 <i>Clustering Using Representatives (CURE)</i>	16
2.5 Gizi.....	22
2.6 Malnutrisi.....	22
2.7 Lansia.....	25
BAB 3 ANALISIS DAN PERANCANGAN.....	26
3.1. Analisis Sistem.....	26
3.1.1. Analisis Masalah.....	27
3.1.2. Analisis Kebutuhan.....	27
3.1.3.1. Analisis Kebutuhan Fungsional.....	27
3.1.3.2. Analisis Kebutuhan Non-Fungsional.....	27
3.1.3. Analisis Data.....	28
3.1.4. Analisis Proses.....	39
3.2. Diagram Umum Sistem.....	48

3.3. Pemodelan Sistem.....	48
3.2.1. <i>Use Case Diagram</i>	49
3.2.2. <i>Activity Diagram</i>	49
3.2.3. <i>Sequence Diagram</i>	51
3.2.4. <i>Flowchart</i>	51
3.2.4.1. <i>Flowchart</i> Sistem.....	52
3.2.4.2. <i>Flowchart</i> Model CURE.....	53
3.4. Perancangan <i>Interface</i>	54
3.4.1. Halaman Beranda.....	54
3.4.2. Halaman Deteksi Malnutrisi dan Intervensi.....	55
3.4.3. Halaman Informasi.....	56
3.4.4. Halaman Visualisasi.....	56
BAB 4 IMPLEMENTASI DAN PENGUJIAN.....	57
4.1. Implementasi Sistem.....	57
4.1.1. Tampilan Halaman Beranda.....	57
4.1.2. Tampilan Halaman Deteksi dan Intervensi.....	58
4.1.3. Tampilan Halaman Informasi.....	59
4.1.4. Tampilan Halaman Visualisasi.....	60
4.2. Pengujian Sistem.....	61
4.3. Hasil Pengujian.....	71
4.3.1. Pengujian Jumlah <i>Cluster</i> Optimal.....	71
4.3.2. Pengujian Kualitas <i>Cluster</i>	72
BAB 5 PENUTUP.....	74
5.1. Kesimpulan.....	74
5.2. Saran.....	74
DAFTAR PUSTAKA.....	75

DAFTAR GAMBAR

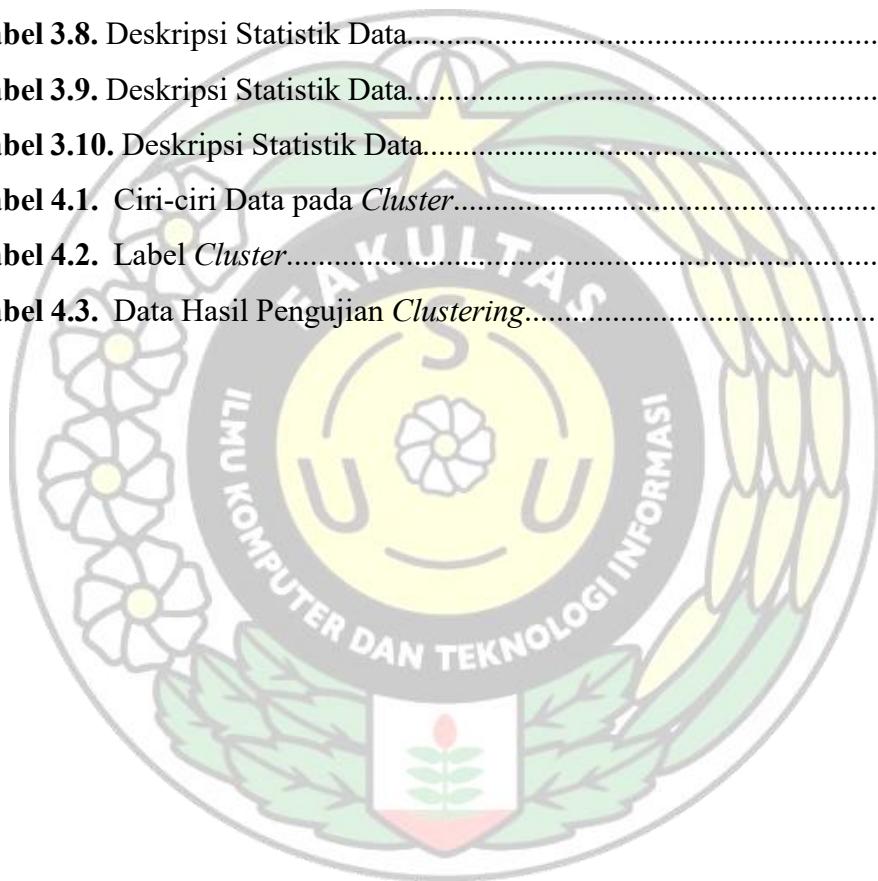
Gambar 2.1. <i>Data Mining</i>	9
Gambar 2.2. <i>Clustering</i>	13
Gambar 2.3. Bentuk <i>Cluster</i> yang Kompleks.....	16
Gambar 2.4. Tahapan Algoritma CURE.....	18
Gambar 2.5. Visualisasi data pada tahapan algoritma CURE.....	20
Gambar 3.1. Diagram <i>Ishikawa/fishbone</i>	26
Gambar 3.2. Histogram Distribusi Umur.....	33
Gambar 3.3. Histogram Distribusi Berat Badan.....	33
Gambar 3.4. Histogram Distribusi Tinggi Badan.....	34
Gambar 3.5. Diagram Umum Sistem.....	48
Gambar 3.6. <i>Use Case Diagram</i>	49
Gambar 3.7. <i>Activity Diagram</i>	50
Gambar 3.8. <i>Sequence Diagram</i>	51
Gambar 3.9. <i>Flowchart</i> Sistem.....	52
Gambar 3.10. <i>Flowchart</i> CURE.....	53
Gambar 3.11. Halaman Beranda.....	54
Gambar 3.12. Halaman Deteksi Malnutrisi.....	55
Gambar 3.13. Halaman Intervensi.....	55
Gambar 3.14. Halaman Informasi.....	56
Gambar 3.15. Halaman Visualisasi.....	56
Gambar 4.1. Tampilan Halaman Beranda.....	57
Gambar 4.2. Tampilan Input Data <i>User</i> pada Halaman Deteksi.....	58
Gambar 4.3. Tampilan Hasil Deteksi dan Intervensi.....	59
Gambar 4.4. Tampilan Halaman Informasi.....	59
Gambar 4.5. Tampilan Halaman Visualisasi.....	60
Gambar 4.6. Visualisasi <i>Clustering</i> dengan PCA.....	61
Gambar 4.7. Distribusi Data di Setiap <i>Cluster</i>	62
Gambar 4.8. Uji Deteksi <i>Cluster</i> Gizi Normal.....	64

Gambar 4.9. Uji Deteksi <i>Cluster</i> Malnutrisi Berat.....	65
Gambar 4.10. Implementasi Metode <i>Silhouette Analysis</i>	71



DAFTAR TABEL

Tabel 3.1. Atribut Data dan Deskripsi.....	28
Tabel 3.2. Data Lansia.....	29
Tabel 3.3. Data Lansia.....	31
Tabel 3.4. Nilai <i>Encoding</i> Atribut.....	34
Tabel 3.5. Data Lansia Setelah <i>Pre-Processing</i>	36
Tabel 3.6. Data Lansia Setelah <i>Pre-Processing</i>	36
Tabel 3.7. Data Lansia Setelah <i>Pre-Processing</i>	37
Tabel 3.8. Deskripsi Statistik Data.....	38
Tabel 3.9. Deskripsi Statistik Data.....	38
Tabel 3.10. Deskripsi Statistik Data.....	39
Tabel 4.1. Ciri-ciri Data pada <i>Cluster</i>	63
Tabel 4.2. Label <i>Cluster</i>	64
Tabel 4.3. Data Hasil Pengujian <i>Clustering</i>	66



BAB 1

PENDAHULUAN

1.1. Latar Belakang

Malnutrisi ialah kondisi yang sering terjadi pada lansia dan menimbulkan masalah serius, baik dalam aspek kesehatan, sosial, maupun perawatan. Lansia rentan mengalami malnutrisi akibat berbagai faktor, seperti penurunan fisiologis terkait usia, keterbatasan akses terhadap makanan bergizi, serta penyakit penyerta (komorbiditas).

Berdasarkan kelompok dari WHO, malnutrisi terbagi menjadi tiga tingkat keparahan, yaitu malnutrisi berat, sedang, dan ringan. Pembagian ini didasarkan pada penilaian terhadap indikator-indikator yang berkaitan dengan gizi individu. Tiap tingkat keparahan memiliki manifestasi klinis dan dampak yang berbeda-beda terhadap kesehatan lansia.

Malnutrisi berat ditandai dengan penurunan berat badan yang signifikan, kelemahan otot yang parah, dan sering kali disertai dengan komplikasi medis yang serius. Malnutrisi sedang ditandai dengan penurunan berat badan moderat, kelemahan otot, dan risiko yang lebih tinggi untuk mengalami infeksi. Sementara itu, malnutrisi ringan ditandai dengan penurunan berat badan yang ringan dan dapat menyebabkan penurunan fungsi tubuh secara bertahap.

Setiap tingkat keparahan membutuhkan intervensi yang berbeda. Malnutrisi berat memerlukan penanganan segera dan intensif, sedangkan malnutrisi ringan mungkin dapat diatasi dengan perbaikan pola makan dan suplementasi nutrisi.

Kondisi malnutrisi pada lansia sering tidak terdeteksi secara dini karena gejala yang muncul cenderung tidak spesifik dan sering kali disalahartikan sebagai tanda-tanda penuaan normal. Gejala seperti kehilangan berat badan secara tidak terencana dan kelemahan otot menjadi tanda-tanda umum yang sering diabaikan. Malnutrisi pada lansia sering terabaikan karena adanya variasi dalam faktor risiko, sehingga menunjukkan perlunya pendekatan penilaian gizi yang terintegrasi untuk mengidentifikasi risiko secara lebih tepat.

Jika malnutrisi tidak segera diatasi, komplikasi serius dapat terjadi. Penurunan status gizi berpotensi memperburuk kondisi kesehatan kronis yang telah ada, meningkatkan risiko infeksi, dan memperlambat penyembuhan luka. Oleh karena itu, deteksi dini sangat penting agar intervensi dapat dilakukan sebelum kondisinya semakin memburuk. Deteksi awal memungkinkan diambilnya langkah-langkah pencegahan yang dapat meningkatkan kualitas hidup lansia. Deteksi dini dan intervensi malnutrisi sangat penting untuk mencegah komplikasi lebih lanjut, termasuk penurunan daya tahan tubuh dan kualitas hidup.

Dalam beberapa tahun terakhir, perkembangan teknologi informasi dan analisis data telah membuka peluang baru dalam pengelompokan risiko malnutrisi pada lansia menggunakan algoritma *machine learning*. Salah satu metode yang relevan adalah *Clustering Using Representatives* (CURE). Algoritma ini memungkinkan pengelompokan data yang kompleks, termasuk data dengan distribusi tidak merata. Algoritma CURE bekerja dengan memilih sejumlah representatif dari setiap *cluster* yang dihasilkan, kemudian secara iteratif menggerakkan representatif ini menuju pusat *cluster* untuk menangkap bentuk asli *cluster*.

CURE memungkinkan identifikasi yang lebih akurat dalam proses pengelompokan dibandingkan metode *clustering* tradisional. Penelitian Manjul dan Nandakumar (2018) menunjukkan bahwa algoritma CURE sangat efektif dalam mengelompokkan data yang kompleks dan menangani *outlier*, sementara penelitian Salman (2023) menemukan kelebihan signifikan dari CURE dalam pengelompokan data berbentuk grafis tiga dimensi. Dengan demikian, penerapan CURE dalam penelitian ini diharapkan dapat menghasilkan pengelompokan lansia berdasarkan status gizi mereka dengan kualitas clustering yang tinggi.

1.2. Rumusan Masalah

Deteksi dini malnutrisi pada lansia sangat penting karena gejala malnutrisi sering tidak terdeteksi dengan jelas akibat kemunculannya yang samar dan sering dianggap sebagai bagian dari proses penuaan normal. Hal ini menyebabkan

penanganan yang terlambat, yang berpotensi memperburuk kondisi kesehatan lansia secara signifikan. Pengelompokan status malnutrisi pada lansia bertujuan untuk mempermudah identifikasi risiko dan memberikan panduan yang lebih akurat dalam merancang intervensi gizi yang tepat sesuai dengan tingkat risiko yang teridentifikasi pada setiap lansia.

1.3. Batasan Masalah

Berikut ini adalah batasan masalah yang diterapkan dalam penelitian ini:

- 1) Fokus penelitian pada variabel berat badan, tinggi badan, penurunan berat badan, asupan makanan, asupan protein, kemampuan mobilitas, kondisi kesehatan, dan fungsi kognitif. Pemilihan variabel ini didasarkan pada standar *Mini Nutritional Assessment* (MNA), yang digunakan untuk mengevaluasi status gizi lansia.
- 2) Data yang digunakan merupakan data yang diperoleh dari UPTD. Pelayanan Sosial Lanjut Usia Binjai dan dilengkapi dengan penyebaran formulir secara online kepada keluarga lansia, dengan total keseluruhan data sebanyak 517 baris.
- 3) Menggunakan algoritma *Clustering Using Representatives* (CURE) untuk proses pengelompokan data lansia berdasarkan status gizi.
- 4) Program dirancang menggunakan bahasa *Python* dan berbasis *website* dengan menggunakan *framework Python*, yaitu *Streamlit*.

1.4. Tujuan Penelitian

Penelitian ini bertujuan untuk mengidentifikasi dan mengelompokkan status malnutrisi pada lansia berdasarkan data yang dikumpulkan melalui metode *clustering* menggunakan algoritma CURE. Melalui deteksi dini risiko malnutrisi, yang sering kali tidak terdeteksi akibat gejala yang samar dan sering dianggap sebagai bagian dari proses penuaan normal, hasil pengelompokan ini diharapkan dapat memberikan panduan yang lebih efektif dalam menentukan intervensi gizi yang sesuai, sehingga dapat membantu dalam memberikan perhatian dan perawatan yang tepat kepada lansia sesuai dengan tingkat risiko mereka.

1.5. Manfaat Penelitian

Adapun manfaat yang diperoleh dari penelitian ini antara lain:

1. Deteksi Dini Malnutrisi pada Lansia: Penelitian ini berkontribusi dalam mendeteksi risiko malnutrisi pada lansia secara lebih awal. Dengan mendeteksi malnutrisi lebih cepat, diharapkan intervensi dapat dilakukan lebih cepat untuk mencegah komplikasi kesehatan yang serius, sehingga meningkatkan kesehatan dan kesejahteraan lansia di masyarakat.
2. Panduan Intervensi yang Lebih Tepat: Hasil penelitian memberikan panduan praktis bagi keluarga, perawat, dan tenaga kesehatan untuk menyesuaikan intervensi gizi berdasarkan kelompok risiko malnutrisi. Penelitian ini dapat membantu masyarakat dalam merancang program gizi yang lebih efektif, meningkatkan kualitas hidup lansia, dan mengurangi beban pada sistem kesehatan.
3. Peningkatan Kesadaran Masyarakat: Dengan penelitian ini, masyarakat dapat lebih memahami pentingnya deteksi dini malnutrisi pada lansia. Kesadaran ini dapat mendorong tindakan proaktif dalam menjaga kesehatan lansia di lingkungan sekitar, serta meningkatkan perhatian masyarakat terhadap masalah gizi lansia.

1.6. Metodologi Penelitian

Metode yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Studi Pustaka

Penelitian ini diawali dengan mengumpulkan referensi dari sumber-sumber yang kredibel, serta melakukan kajian literatur melalui berbagai jurnal, artikel ilmiah, dan *website* yang relevan dengan topik malnutrisi pada lansia dan penerapan algoritma CURE. Referensi ini digunakan untuk memperoleh wawasan lebih mendalam tentang tantangan dalam deteksi dini malnutrisi pada lansia serta penerapan algoritma CURE dalam mengelompokkan risiko malnutrisi secara lebih efektif.

2. Analisa dan Perancangan

Berdasarkan fokus penelitian, penulis melakukan analisis mendalam terhadap data yang telah diperoleh, seperti data terkait status gizi lansia. Data ini terdiri dari informasi personal seperti nama, usia, jenis kelamin, berat badan, dan tinggi badan, serta beberapa faktor lain yang mempengaruhi status gizi seperti penurunan berat badan, asupan makanan, asupan protein, kemampuan mobilitas, kemampuan mengunyah, kondisi kesehatan, dan fungsi kognitif.

Langkah-langkah proses *clustering* menggunakan algoritma CURE kemudian dirancang berdasarkan data tersebut dalam bentuk diagram alir (*flowchart*), yang menggambarkan alur kerja sistem mulai dari *preprocessing* data untuk memastikan data dalam format yang sesuai, implementasi algoritma CURE untuk pengelompokan data lansia berdasarkan faktor risiko malnutrisi, hingga evaluasi hasil *clustering* guna menilai efektivitas pengelompokan yang dihasilkan.

3. Implementasi

Algoritma CURE diimplementasikan menggunakan bahasa pemrograman *Python*. Proses ini mengikuti alur yang telah dirancang, dimulai dari inisialisasi *cluster*, penggabungan *cluster* berdasarkan jarak titik representatif terdekat, hingga pembentukan *cluster* akhir. Tujuan dari sistem ini adalah untuk mengelompokkan data lansia ke dalam beberapa kategori berdasarkan status malnutrisi.

4. Pengujian

Pada tahap ini, sistem diuji dengan menggunakan data lansia yang telah diproses. Hasil *clustering* dievaluasi untuk menilai kualitas pengelompokan dengan algoritma CURE dalam mengidentifikasi risiko malnutrisi, baik dalam kategori berat, ringan, maupun kondisi gizi normal. Pengujian dilakukan dengan menggunakan metrik evaluasi seperti *Davies-Bouldin Index* untuk mengukur kualitas *clustering* secara kuantitatif. Selain itu, dilakukan juga analisis untuk menentukan jumlah *cluster* yang optimal, menggunakan metode seperti *Silhouette Analysis* untuk memastikan pengelompokan yang

lebih akurat.

5. Dokumentasi

Seluruh rangkaian penelitian didokumentasikan, mulai dari analisis, perancangan, implementasi, hingga pengujian, yang dituangkan dalam bentuk skripsi. Dokumentasi tersebut mencakup penjelasan mengenai metode yang diterapkan, hasil serta analisis *clustering*, dan evaluasi kinerja sistem dalam pengelompokan status malnutrisi pada lansia. Tujuan dari hal ini adalah untuk menyajikan penjelasan yang terperinci dan terstruktur mengenai langkah-langkah penelitian yang telah dilakukan.

1.7. Penelitian Relevan

1. Berdasarkan penelitian yang dilakukan oleh Norman *et al.* (2021) dengan judul “*Malnutrition in Older Adults—Recent Advances and Remaining Challenges*”, yang menghasilkan kesimpulan bahwa penyebab malnutrisi pada lansia sangat kompleks dan multifaktorial, termasuk faktor penuaan yang memperburuk kemampuan tubuh untuk mengatasi stres fisik, kognitif, dan mental. Selain itu, penyakit terkait usia, seperti anoreksia akibat penuaan, serta proses inflamasi yang meningkat dan gangguan kognitif, berperan penting dalam mempengaruhi status gizi lansia.
2. Berdasarkan penelitian yang dilakukan oleh Kaiser *et al.* (2009) dengan judul “*Validation of the Mini Nutritional Assessment short-form (MNA®-SF): A practical tool for identification of nutritional status*”, yang menghasilkan kesimpulan bahwa alat penilaian gizi yang komprehensif seperti *Mini Nutritional Assessment* (MNA) lebih efektif dalam mendekripsi malnutrisi pada lansia.
3. Berdasarkan penelitian yang dilakukan oleh Kumble, A., & Tewari, V. (2020) dengan judul “*Cure Clustering Algorithm and Its Various Implementations*”, yang menghasilkan kesimpulan bahwa algoritma CURE mampu menangani pengelompokan data yang memiliki *outlier* dan distribusi tidak merata secara lebih efektif dibandingkan metode

clustering tradisional. Penelitian ini menunjukkan bahwa CURE dapat diterapkan pada data kesehatan yang kompleks seperti data status gizi lansia.

4. Berdasarkan penelitian yang dilakukan oleh Manjul, V., & Nandakumar, A. N. (2018) dengan judul "*An Effective Cure Clustering Algorithm in Education Data Mining Techniques to Valuate Student's Performance*", yang menghasilkan kesimpulan bahwa algoritma CURE dapat digunakan untuk mengelompokkan data besar dengan hasil yang lebih akurat. Algoritma ini cocok digunakan untuk pengelompokan risiko malnutrisi pada lansia berdasarkan data status gizi yang beragam.
5. Berdasarkan penelitian yang dilakukan oleh Sharma *et al.* (2020) dengan judul "*Malnutrition, Health, and the Role of Machine Learning in Clinical Setting*", penelitian ini menyimpulkan bahwa algoritma *machine learning*, termasuk *unsupervised learning*, dapat membantu dalam mendekripsi dini risiko malnutrisi pada pasien lanjut usia. Dengan menggunakan data rekam medis elektronik yang mencakup berbagai parameter kesehatan seperti BMI dan komposisi tubuh, algoritma ini mampu mengidentifikasi pola yang mengarah pada risiko malnutrisi dan menghasilkan rekomendasi intervensi yang lebih cepat dan efisien.
6. Berdasarkan penelitian yang dilakukan oleh Tomasiewicz *et al.* (2024) dalam artikel berjudul "*Advancing the Understanding of Malnutrition in the Elderly Population: Current Insights and Future Directions*", penelitian ini menyimpulkan bahwa alat diagnostik saat ini masih belum cukup akurat untuk mendekripsi malnutrisi pada lansia. Mereka merekomendasikan pendekatan multidisipliner yang memadukan nutrisi, aktivitas fisik, dan dukungan psikososial. Penelitian ini juga menyarankan perlunya metode deteksi yang lebih akurat, termasuk penggunaan *machine learning* untuk personalisasi intervensi berdasarkan biomarker dan kondisi kesehatan spesifik.

1.8. Sistematika Penulisan

Sistematika penulisan skripsi ini terdiri dari beberapa bagian, seperti yang dijelaskan berikut ini.

BAB 1 PENDAHULUAN

Bab ini membahas mengenai latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, metode penelitian, penelitian relevan serta sistematika penulisan yang akan digunakan dalam penulisan skripsi ini.

BAB 2 LANDASAN TEORI

Bab ini membahas konsep-konsep dasar mengenai *Data Mining*, *Machine Learning*, *Clustering*, serta metode *Clustering Using Representatives* (CURE), yang dikaitkan dengan topik gizi, malnutrisi, dan lansia.

BAB 3 ANALISIS DAN PERANCANGAN

Bab ini menjelaskan tentang analisis kebutuhan dan desain sistem yang akan dikembangkan, yang disesuaikan dengan rumusan dan batasan masalah yang telah ditetapkan dalam penelitian ini.

BAB 4 IMPLEMENTASI DAN PENGUJIAN

Bab ini menjelaskan tentang penerapan algoritma CURE untuk pengelompokan status malnutrisi pada lansia, yang sebelumnya telah dianalisis dan dirancang, serta dilakukan evaluasi untuk menguji efektivitas sistem.

BAB 5 KESIMPULAN DAN SARAN

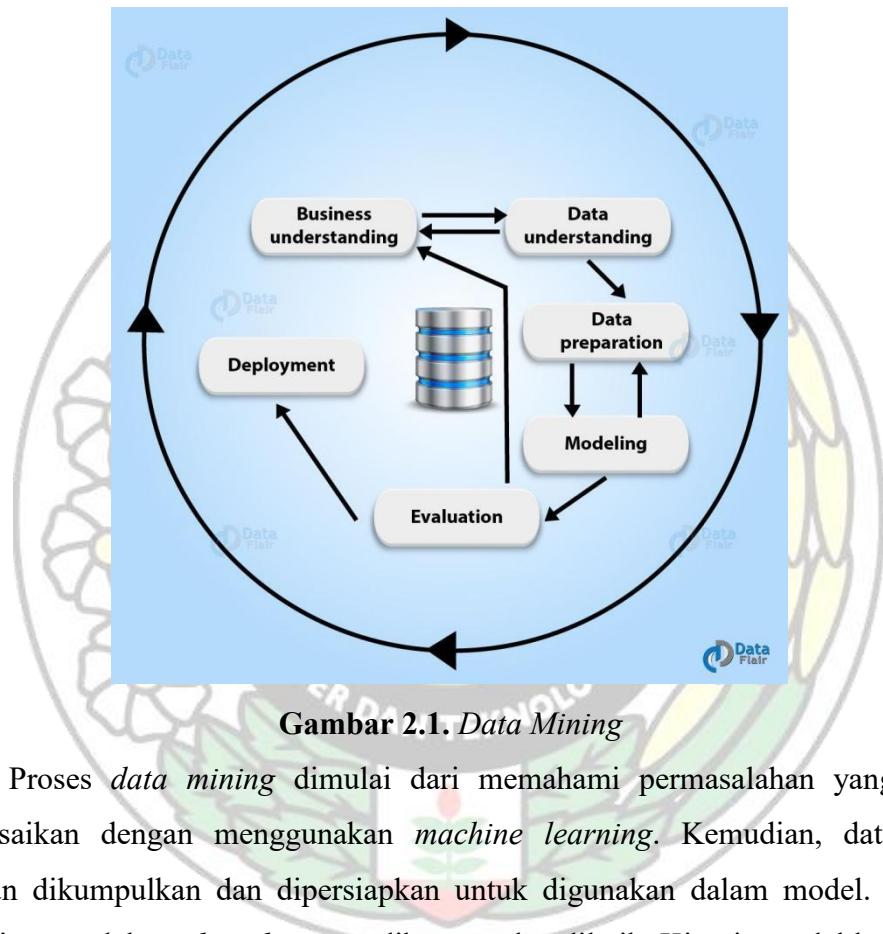
Bab ini merangkum kesimpulan, hasil penelitian, serta saran untuk pengembangan penelitian selanjutnya.

BAB 2

LANDASAN TEORI

2.1 Data Mining

Data mining adalah proses mengekstraksi informasi berharga dari kumpulan data yang besar melalui penggunaan teknik analisis dan pemodelan statistik serta kecerdasan buatan.



Gambar 2.1. *Data Mining*

Proses *data mining* dimulai dari memahami permasalahan yang ingin diselesaikan dengan menggunakan *machine learning*. Kemudian, data yang relevan dikumpulkan dan dipersiapkan untuk digunakan dalam model. Setelah data siap, model *machine learning* dibangun dan dilatih. Kinerja model kemudian dievaluasi untuk memastikan model tersebut akurat dan dapat diandalkan. Terakhir, model yang sudah jadi diterapkan atau di-deploy ke dalam sistem.

Dalam penelitian ini, *data mining* digunakan untuk mengidentifikasi pola dan karakteristik risiko malnutrisi pada lansia. Dengan metode ini, sistem dapat mengelompokkan data lansia ke dalam berbagai kategori malnutrisi, termasuk malnutrisi berat, ringan, dan gizi normal, berdasarkan berbagai parameter yang mempengaruhi status gizi lansia, seperti berat badan, frekuensi makan, dan kondisi kesehatan.

Data mining memainkan peran penting dalam konteks penelitian ini karena mampu menangani volume data yang besar dan variabilitas dari setiap individu lansia. Teknik yang digunakan, seperti *clustering*, membantu menemukan pola tersembunyi dalam data lansia yang berbeda-beda. Algoritma CURE (*Clustering Using Representatives*), sebagai bagian dari teknik *data mining*, digunakan untuk membagi data ke dalam beberapa kelompok berdasarkan risiko malnutrisi pada data kesehatan lansia. Dengan algoritma ini, penanganan terhadap lansia bisa lebih akurat karena pengelompokan didasarkan pada analisis yang mendalam dari data yang kompleks.

Data mining dalam penelitian kesehatan juga dapat digunakan untuk melakukan diagnosis dini dan memberikan rekomendasi intervensi berbasis data. Dalam penelitian ini, teknik *data mining* juga membantu merancang sistem rekomendasi yang memberikan panduan intervensi gizi yang sesuai bagi kelompok lansia berdasarkan hasil pengelompokan risiko malnutrisi yang dihasilkan.

2.2 *Machine Learning*

Machine learning ialah bidang dalam kecerdasan buatan yang bertujuan mengembangkan algoritma yang memungkinkan komputer untuk mempelajari pola dari data dan mengambil keputusan berdasarkan temuan tersebut.

Penerapan *machine learning* dalam bidang kesehatan semakin luas, terutama dalam mengembangkan model prediksi ataupun deteksi yang membantu dokter dalam diagnosis, perawatan pasien, serta manajemen data medis. Algoritma *machine learning* dapat digunakan untuk analisis data besar yang tidak bisa diolah secara manual, memungkinkan deteksi dini dari pola-pola yang tidak terlihat oleh manusia.

2.2.1. *Supervised Learning*

Supervised learning ialah teknik pembelajaran mesin di mana model dilatih menggunakan data yang sudah dilabeli dengan output yang diinginkan. Metode ini melibatkan pembelajaran dari data yang sudah diketahui outputnya untuk kemudian digunakan dalam prediksi data baru. Misalnya, pada prediksi malnutrisi, data input seperti berat badan, tinggi badan, dan asupan

nutrisi harian digunakan untuk melatih model guna mengklasifikasikan individu apakah berisiko malnutrisi atau tidak.

Beberapa algoritma yang umum digunakan dalam *supervised learning* adalah *Support Vector Machine* (SVM), *Decision Tree*, dan *Random Forest*. Dalam konteks malnutrisi, *supervised learning* dapat membantu memprediksi status gizi seseorang berdasarkan parameter kesehatan dan faktor risiko lainnya, serta dapat membantu merancang intervensi medis yang lebih terfokus.

2.2.2. *Unsupervised Learning*

Unsupervised learning ialah jenis *machine learning* yang berfungsi tanpa memerlukan data berlabel, sehingga cocok untuk menganalisis dataset yang tidak memiliki informasi klasifikasi yang jelas. Dalam pendekatan ini, algoritma berusaha untuk menemukan pola atau struktur tersembunyi dalam data yang tidak memiliki label atau kategori yang telah ditentukan sebelumnya. Hal ini memungkinkan peneliti untuk mengidentifikasi hubungan dan pola yang tidak terduga, serta mengelompokkan data berdasarkan kesamaan fitur.

Dalam konteks penelitian tentang malnutrisi pada lansia, *unsupervised learning* memiliki aplikasi yang sangat relevan. Metode *clustering* dapat dimanfaatkan untuk mendeteksi pola-pola dalam asupan gizi dan status kesehatan lansia. *Clustering* dapat mengelompokkan lansia berdasarkan risiko malnutrisi, yang membantu dalam merancang program intervensi gizi yang lebih tepat sasaran.

Penelitian oleh *Medina et al. (2021)* menunjukkan penggunaan *unsupervised learning* untuk menganalisis pola makan. Pada penelitian ini algoritma *clustering* digunakan untuk mengelompokkan data berdasarkan asupan makanan harian dan status kesehatan. Temuan penelitian ini menekankan bahwa pendekatan ini tidak hanya membantu mengidentifikasi kelompok dengan risiko tinggi malnutrisi, tetapi juga memberikan gambaran yang lebih jelas tentang kebutuhan gizi spesifik dari kelompok-kelompok tersebut.

Dengan demikian, *unsupervised learning*, khususnya melalui metode *clustering*, terbukti sangat berguna dalam analisis data kesehatan, memungkinkan deteksi dini masalah gizi serta mendukung perancangan intervensi yang lebih efisien dan efektif untuk populasi lansia.

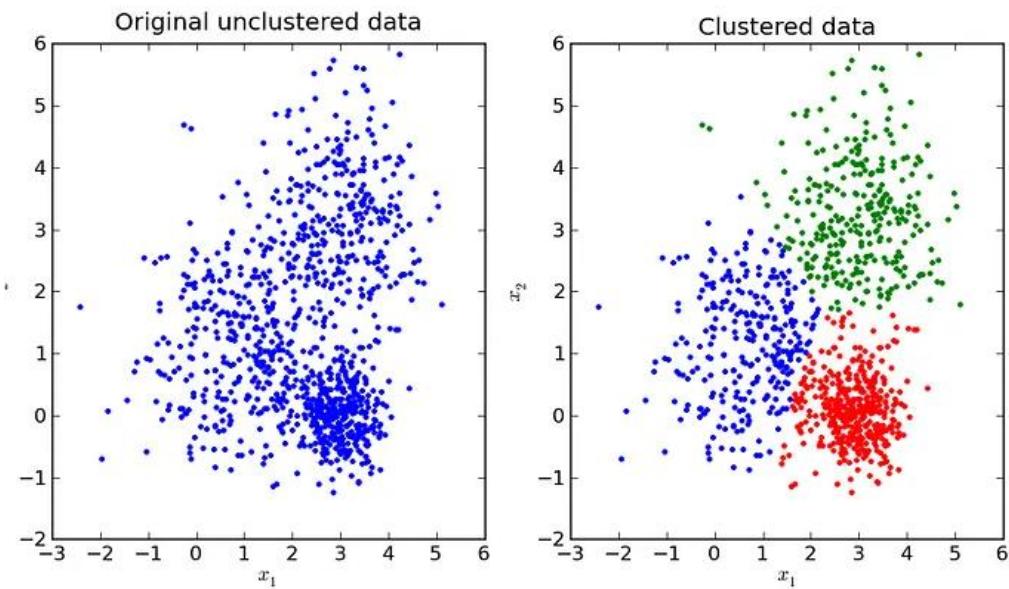
2.2.3. *Semi-supervised Learning*

Semi-supervised learning menggabungkan *supervised* dan *unsupervised learning*, dengan memanfaatkan sejumlah kecil data berlabel dan data tanpa label yang lebih besar untuk pelatihan model. Ini sangat berguna dalam situasi di mana pelabelan data mahal atau memakan waktu, seperti dalam analisis data kesehatan lansia.

Dalam pengelompokan malnutrisi, *semi-supervised learning* dapat digunakan untuk mengidentifikasi pola-pola dari data lansia yang tidak berlabel sepenuhnya, sehingga memungkinkan pengambilan keputusan yang lebih baik tanpa memerlukan pelabelan lengkap pada semua data. Pendekatan *semi-supervised* ini menjadi semakin penting dalam analisis kesehatan lansia, di mana hanya sebagian data yang mungkin dilengkapi dengan diagnosis atau status gizi yang jelas, sementara sisanya dapat dikategorikan melalui pola-pola yang ditemukan oleh algoritma.

2.3 *Clustering*

Clustering adalah teknik dalam analisis data yang bertujuan untuk membagi data ke dalam kelompok-kelompok (*clusters*) yang memiliki karakteristik serupa, berdasarkan kedekatan atau kesamaan atribut di antara elemen-elemen data tersebut. Proses ini melibatkan identifikasi dan pengelompokan data yang memiliki karakteristik serupa, sehingga data dalam satu kelompok memiliki kesamaan yang lebih tinggi dibandingkan dengan data di kelompok lain. *Clustering* menjadi alat yang sangat penting dalam analisis data eksploratif karena membantu dalam memahami struktur data, mengidentifikasi *outlier*, dan mendeteksi pola tersembunyi dalam dataset yang besar.



Gambar 2.2. Clustering

Proses *clustering* terdiri dari beberapa tahap utama yang memungkinkan data dapat dikelompokkan sesuai dengan karakteristik tertentu. Tahap pertama adalah pemilihan fitur, di mana variabel atau atribut penting dalam dataset dipilih untuk mendukung keakuratan proses pengelompokan. Selanjutnya, dilakukan perhitungan jarak antar data menggunakan metrik tertentu seperti jarak *Euclidean* atau *Manhattan* untuk mengukur seberapa mirip atau berbeda data satu dengan yang lain.

Jarak *Euclidean* dapat diperoleh dengan menerapkan rumus berikut:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.1)$$

Jarak *Manhattan* dapat dihitung dengan menggunakan persamaan berikut:

$$d(x, y) = \sqrt{\sum_{i=1}^n |x_i - y_i|} \quad (2.2)$$

Di mana:

- i. x dan y adalah dua titik dalam ruang n -dimensi.
- ii. n adalah jumlah dimensi fitur.

Setelah jarak dihitung, jika algoritma memerlukan, *centroid* atau pusat *cluster* ditentukan sebagai titik acuan awal yang merepresentasikan *cluster*. Algoritma kemudian mengelompokkan data ke *centroid* terdekat dan memperbarui posisi *centroid* pada setiap iterasi. Pembaruan *centroid* ini dilakukan

dengan menghitung rata-rata posisi data yang berada dalam *cluster* tersebut, sehingga *centroid* bergerak mendekati titik yang merepresentasikan *cluster* dengan lebih akurat. Proses ini berlanjut hingga posisi *centroid* stabil atau mencapai kriteria penghentian yang ditentukan.

Centroid C dari *cluster* yang terdiri dari titik x_1, x_2, \dots, x_n dapat dihitung dengan menggunakan rumus ini:

$$C = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.3)$$

Pada saat penggabungan antara *cluster* terdekat, maka terdapat pembaruan centorid yang dapat dihitung dengan menggunakan rumus ini:

$$\text{centroid baru} = \frac{(self.c \times totalPoints_1) + (clust.cr \times totalPoints_2)}{totalPoints_1 + totalPoints2} \quad (2.4)$$

Pada tahap penggabungan *cluster*, terdapat beragam metode yang dapat digunakan untuk menentukan pasangan *cluster* mana yang akan digabungkan terlebih dahulu. Berikut adalah beberapa pendekatan yang dapat diterapkan untuk menggabungkan kedua *cluster*.

Metode *Single Linkage*:

$$d(A, B) = \min_{a \in A, b \in B} d(a, b) \quad (2.5)$$

Metode *Complate inkage*:

$$d(A, B) = \max_{a \in A, b \in B} d(a, b) \quad (2.6)$$

Metode *Average Linkage*:

$$d(A, B) = \frac{1}{|A| \times |B|} \sum_{a \in A} \sum_{b \in B} d(a, b) \quad (2.7)$$

Di mana:

- i. $d(A, B)$ adalah jarak antara dua *cluster*.
- ii. \min adalah nilai terkecil dari semua kemungkinan jarak.
- iii. \max adalah nilai terbesar dari semua kemungkinan jarak.
- iv. $d(a, b)$ adalah jarak antara dua titik data, yaitu titik data a yang berasal dari *cluster* A dan titik data b yang berasal dari *cluster* B.
- v. $a \in A, b \in B$ menunjukkan bahwa titik a adalah anggota dari *cluster* A dan titik b adalah anggota dari *cluster* B.

Setiap metode ini memiliki karakteristik dan asumsi yang berbeda, sehingga pemilihan metode yang tepat sangat bergantung pada sifat data dan tujuan analisis.

Pada tahapan *clustering*, terdapat tahap evaluasi yang berfungsi untuk menilai kualitas dan efektivitas pembagian data ke dalam *cluster*. Evaluasi dapat dilakukan dengan membuat visualisasi hasil *clustering*, seperti grafik atau plot, untuk melihat sebaran data dalam setiap *cluster* dan memahami apakah pola pembagian data sesuai. Pada tahap ini juga diperlukan penggunaan metrik evaluasi seperti *Davies-Bouldin Index* memberikan nilai kuantitatif mengenai kualitas *cluster*. Hal ini dapat membantu untuk mengidentifikasi apakah hasil *clustering* sudah optimal.

Rumus *Davies-Bouldin Index (DBI)*:

$$DBI = \frac{1}{n} \sum_{i=1}^n \left(\frac{s_i + s_j}{d_{ij}} \right) \quad (2.8)$$

di mana:

- i. n adalah jumlah *cluster* yang terbentuk,
- ii. s_i adalah rata-rata jarak semua titik dalam *cluster* i ke *centroid*-nya (menunjukkan ukuran sebaran dalam *cluster*),
- iii. s_j adalah rata-rata jarak semua titik dalam *cluster* j ke *centroid*-nya (menunjukkan ukuran sebaran dalam *cluster*),
- iv. d_{ij} adalah jarak antara *centroid* *cluster* i dan *centroid* *cluster* j ,
- v. $\max_{i \neq j}$ mencari nilai maksimum dari rasio jarak antar *cluster* untuk tiap pasangan *cluster* i dan j .

Selain itu, untuk menentukan jumlah *cluster* yang optimal, metode seperti *Silhouette Analysis* dapat digunakan. *Silhouette Analysis* mengukur seberapa dekat setiap titik data dalam suatu *cluster* dengan titik data di *cluster* lain.

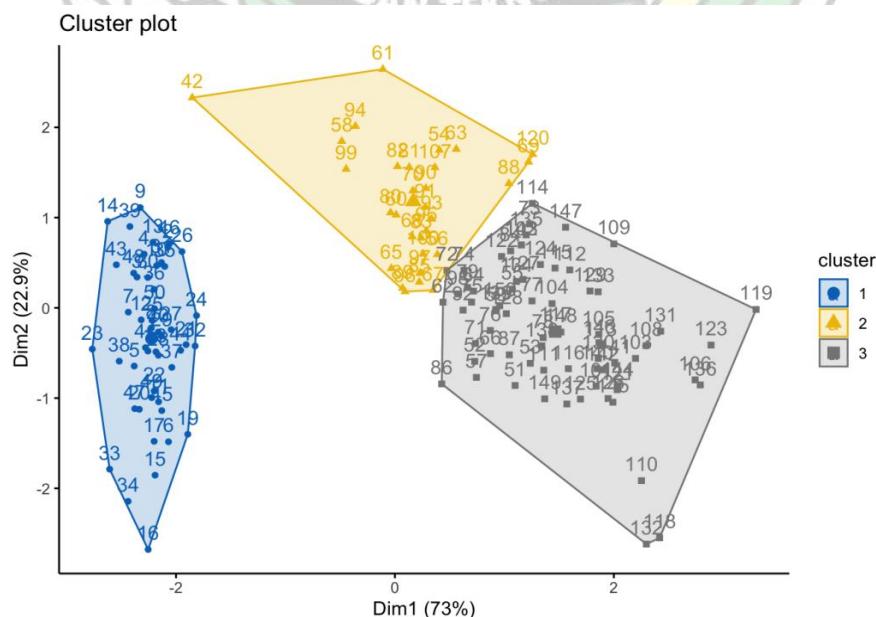
Dalam konteks kesehatan, *clustering* sering digunakan untuk mengidentifikasi kelompok populasi yang memiliki karakteristik serupa, seperti risiko kesehatan atau status gizi. Metode *clustering* dapat digunakan untuk menganalisis status gizi pada lansia, mengidentifikasi kelompok-kelompok dengan pola makan yang berbeda dan risiko malnutrisi. Dengan ini diharapkan bahwa teknik *clustering* dapat membedakan antara lansia yang memiliki risiko

berat dan ringan terhadap malnutrisi, sehingga memungkinkan intervensi yang lebih terfokus.

Clustering dapat dieksplorasi untuk mengelompokkan lansia berdasarkan status kesehatan dan kebiasaan makan mereka. Dengan pendekatan ini, pola yang tidak terlihat sebelumnya dapat membantu dalam merancang program kesehatan yang lebih efektif dan intervensi gizi yang disesuaikan dengan kebutuhan masing-masing kelompok. Secara keseluruhan, *clustering* memainkan peran penting dalam memahami kompleksitas data kesehatan dan gizi, khususnya dalam konteks populasi lansia. Melalui identifikasi pola dan kelompok dalam data, teknik ini dapat memberikan wawasan berharga untuk pengembangan strategi intervensi yang lebih tepat sasaran dan berbasis data.

2.4 Clustering Using Representatives (CURE)

CURE (*Clustering Using Representatives*) adalah algoritma *clustering* yang dikembangkan untuk menangani *cluster* yang tidak teratur dan kehadiran *outlier*. CURE memperkenalkan pendekatan yang berbeda dengan merepresentasikan setiap *cluster* menggunakan beberapa titik representatif (*representative points*) alih-alih hanya satu pusat *cluster*. Ini memberikan kemampuan yang lebih baik dalam mencakup bentuk *cluster* yang kompleks dan distribusi data yang beragam.



Gambar 2.3. Bentuk *Cluster* yang Kompleks

CURE telah diterapkan dalam berbagai bidang penelitian, termasuk kesehatan dan analisis data yang kompleks. Sebagai contoh, penelitian oleh Nandakumar et al. (2018) menggunakan CURE untuk menganalisis data genomik, di mana mereka menemukan bahwa CURE lebih efektif dalam mengelompokkan gen berdasarkan ekspresi dibandingkan dengan algoritma *clustering* tradisional. Temuan ini menunjukkan keunggulan CURE dalam mengelola dataset dengan bentuk yang tidak teratur dan distribusi yang kompleks. Kelebihan dari penggunaan CURE dibandingkan algoritma *clustering* tradisional adalah:

- 1) Pengelompokan Data Kompleks: CURE mampu mengelompokkan data yang memiliki distribusi atau bentuk yang tidak reguler. Titik representatif yang digunakan oleh CURE memungkinkan algoritma untuk menangkap karakteristik *cluster* yang lebih rumit, menjadikannya lebih efektif untuk dataset dengan distribusi yang kompleks.
- 2) Fleksibilitas dalam Jumlah *Cluster*: Algoritma ini tidak membatasi jumlah *cluster* yang harus ditentukan di awal, sehingga lebih fleksibel dalam menangani data yang tidak diketahui jumlah *cluster*-nya.
- 3) Kapasitas untuk Data Besar: CURE dapat diimplementasikan untuk dataset besar, yang membuatnya cocok untuk analisis data yang berskala besar dan berdimensi tinggi.
- 4) Kemampuan Menangani *Outlier*: CURE lebih tahan terhadap *outlier* karena titik representatifnya dapat membantu mengurangi pengaruh data ekstrem.

Meskipun memiliki sejumlah keunggulan, algoritma CURE juga memiliki beberapa kekurangan, yaitu sebagai berikut:

- 1) Kompleksitas Komputasi yang Tinggi

Algoritma CURE membutuhkan waktu komputasi yang cukup besar, terutama untuk dataset berdimensi tinggi. Hal ini disebabkan oleh kebutuhan untuk memilih titik representatif dan menghitung jarak antar *cluster*, yang memerlukan waktu lebih lama dibandingkan algoritma *clustering* sederhana.

2) Pemilihan Parameter yang Sulit

CURE membutuhkan pengaturan parameter yang tepat, seperti jumlah titik representatif dan faktor *shrinkage*. Parameter yang tidak dipilih dengan baik dapat menurunkan kualitas *clustering* sehingga menimbulkan hasil yang kurang akurat.

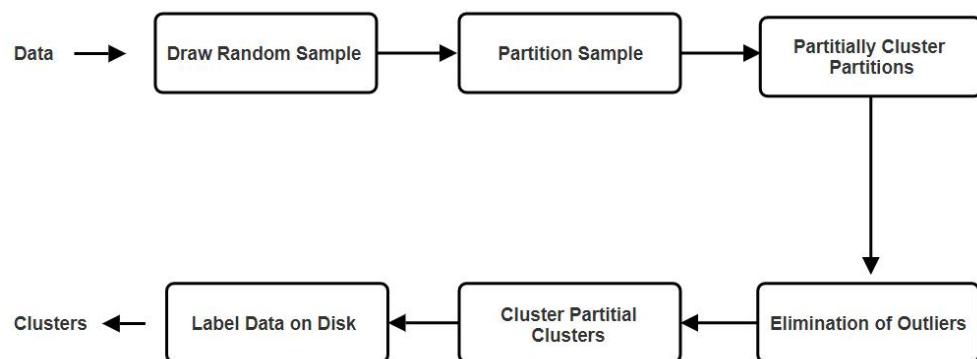
3) Kebutuhan Memori yang Besar

Karena algoritma ini menyimpan beberapa titik representatif untuk setiap *cluster*, CURE memerlukan lebih banyak memori dibandingkan algoritma *clustering* lainnya, khususnya pada data yang berskala besar.

4) Kurang Efisien untuk Data dengan Struktur yang Heterogen

Meskipun fleksibel, CURE masih memiliki keterbatasan dalam menangani data dengan ukuran atau kerapatan *cluster* yang sangat berbeda-beda. Dalam beberapa kasus, algoritma ini kurang efisien untuk data dengan variasi kerapatan atau ukuran *cluster* yang sangat kontras.

Berikut adalah tahapan-tahapan dalam algoritma CURE (*Clustering Using Representatives*) yang secara bertahap dilakukan untuk mengelompokkan data secara lebih akurat:



Gambar 2.4. Tahapan Algoritma CURE

1) *Draw Random Sample* (Pengambilan Sampel Acak):

Algoritma CURE dimulai dengan mengambil sampel acak dari dataset besar. Hal ini dilakukan untuk mengurangi jumlah data yang diproses pada tahap awal tanpa mengorbankan keakuratan hasil *clustering*. *Sampling* ini bertujuan untuk membuat proses lebih efisien.

2) *Partition Sample* (Partisi Sampel):

Setelah mengambil sampel acak, sampel tersebut dipartisi menjadi beberapa subset. Setiap partisi akan diolah secara terpisah untuk mempermudah dan mempercepat proses *clustering*. Partisi ini mengelompokkan data dalam bagian yang lebih kecil.

3) *Partially Cluster Partitions* (*Clustering* Sebagian dari Partisi):

Pada tahap ini, masing-masing partisi dikelompokkan secara terpisah menggunakan titik representatif (*representative points*). Hasil pengelompokan ini bersifat sementara dan belum merupakan hasil akhir, tetapi penting untuk menangkap struktur awal dari data.

4) *Elimination of Outliers* (Penghapusan *Outlier*):

Setelah pembentukan *cluster* sementara, algoritma melakukan deteksi dan penghapusan *outlier*. *Outlier* adalah data yang sangat berbeda dari mayoritas data lainnya dalam *cluster*. Dengan menghapus *outlier*, CURE memastikan bahwa *cluster* yang terbentuk lebih representatif dan tidak terdistorsi oleh data ekstrem.

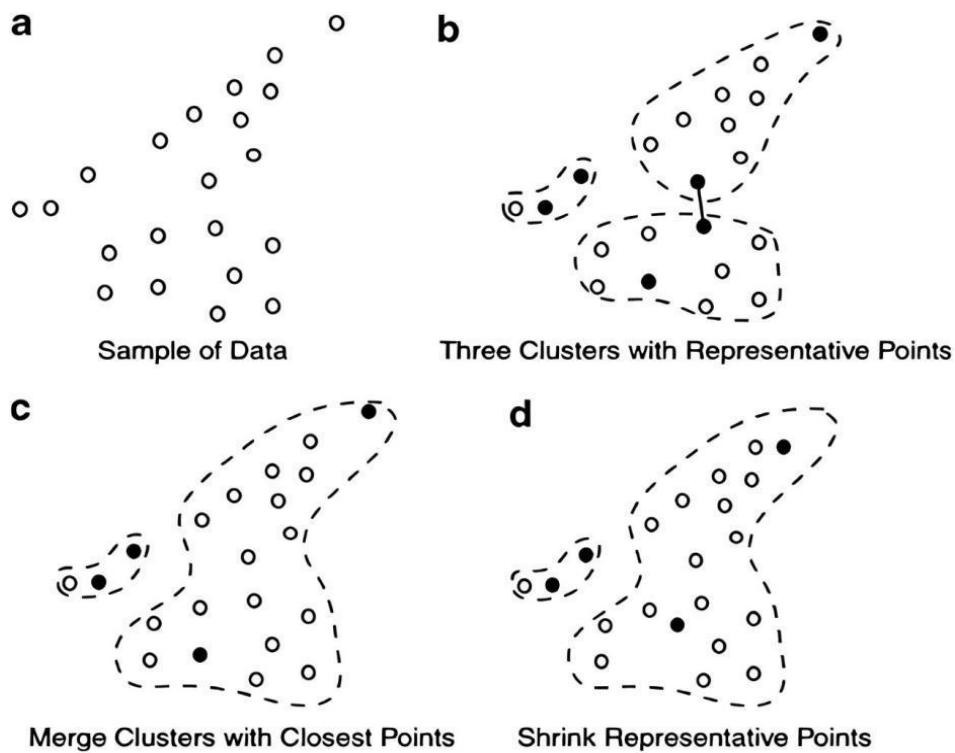
5) *Cluster Partial Clusters* (Penggabungan *Cluster* Parsial):

Setelah menghilangkan *outlier* dan membentuk *cluster* sementara di setiap partisi, *cluster-cluster* ini digabungkan untuk membentuk *cluster* akhir yang lebih besar. Penggabungan dilakukan berdasarkan jarak antara titik-titik representatif dari masing-masing *cluster* parsial.

6) *Label Data on Disk* (Pelabelan Data di *Disk*):

Tahap terakhir adalah melabeli data asli berdasarkan *cluster* yang telah terbentuk. Setelah semua data berhasil dikelompokkan, label diberikan kepada data pada *disk*. Hal ini memungkinkan data diidentifikasi sesuai dengan *cluster* yang relevan, dan proses *clustering* dapat diimplementasikan pada dataset besar.

Berikut adalah contoh visualisasi data yang melalui tahapan-tahapan algoritma CURE:



Gambar 2.5. Visualisasi data pada tahapan algoritma CURE

- Gambar ini menunjukkan sekumpulan data yang tersebar secara acak. Setiap titik mewakili satu data *point*. Tujuan dari *clustering* adalah untuk mengelompokkan titik-titik data ini menjadi beberapa kelompok berdasarkan kesamaan karakteristik mereka.
- Pada gambar ini, data telah dikelompokkan menjadi tiga *cluster* yang berbeda. Setiap *cluster* diwakili oleh sebuah titik hitam yang disebut "representatif". Titik representatif ini merupakan pusat atau pusat massa dari masing-masing *cluster*. Algoritma *clustering* akan berusaha menempatkan titik representatif sedemikian rupa sehingga meminimalkan jarak antara titik-titik data dalam satu *cluster* dengan titik representatifnya.
- Pada tahap ini, algoritma akan mencari dua *cluster* yang jarak antara titik representatifnya paling dekat. Kedua *cluster* tersebut kemudian digabung menjadi satu *cluster* yang lebih besar. Titik representatif baru dari *cluster* gabungan ini akan dihitung ulang. Proses penggabungan

cluster akan terus dilakukan hingga tercapai kriteria penghentian yang telah ditentukan (misalnya, jumlah *cluster* yang diinginkan).

- d) Setelah penggabungan *cluster*, titik-titik representatif dari *cluster* yang baru terbentuk akan "dikerutkan" atau "diperkecil". Proses ini bertujuan untuk menjaga agar *cluster* tidak terlalu besar dan tetap mempertahankan bentuk aslinya. Dengan memperkecil titik representatif, maka *cluster* akan lebih fokus pada daerah yang padat data.

Perhitungan Jarak Antar Titik

Dalam algoritma CURE, jarak antar titik umumnya dihitung menggunakan jarak *Euclidean*. Ini penting untuk menentukan seberapa dekat satu titik data dengan titik lainnya dalam sebuah *cluster*. Rumus jarak *Euclidean* dapat dilihat pada rumus 2.1.

Generate Representative Baru

Dalam CURE, titik representatif baru untuk setiap *cluster* dibentuk dengan menyusutkan titik representatif lama ke arah *centroid* baru berdasarkan faktor *shrinkage* tertentu. Faktor ini biasanya berkisar antara 0 dan 1.

Penggabungan *Cluster* Berdasarkan Titik Representatif

Penggabungan dua *cluster* dalam algoritma CURE dilakukan dengan cara menghitung jarak antara titik-titik representatif yang ada di masing-masing *cluster*. Setiap *cluster* memiliki beberapa titik representatif yang dipilih untuk mencerminkan distribusi data dalam *cluster* tersebut. Jarak antar *cluster* dihitung dengan membandingkan titik representatif dari kedua *cluster* yang ingin digabungkan. Proses ini memastikan bahwa perbedaan signifikan antara *cluster* tetap terjaga meskipun penggabungan dilakukan.

Ketika dua *cluster* dengan titik representatif terdekat ditemukan, kedua *cluster* tersebut digabungkan menjadi satu *cluster*. Penggabungan ini dilakukan dengan cara memperbarui titik representatif yang ada, yaitu dengan

mencari titik representatif baru yang dapat mewakili *cluster* gabungan. Titik representatif baru ini diperoleh dengan memilih titik yang paling jauh dari pusat *cluster* dan menggesernya menggunakan parameter tertentu, seperti *alpha*, untuk menghasilkan titik representatif yang lebih akurat dan terdistribusi.

2.5 Gizi

Gizi merupakan salah satu aspek krusial dalam mempertahankan kesehatan lansia, terutama terkait dengan asupan nutrisi yang memadai. Pada usia lanjut, penurunan nafsu makan dan perubahan fisiologis dapat menyebabkan asupan makanan yang kurang mencukupi, sehingga meningkatkan risiko malnutrisi. Kondisi seperti penurunan fungsi pencernaan, gangguan oral seperti gigi tanggal, serta masalah kesehatan kronis seperti diabetes dan hipertensi dapat memengaruhi pola makan lansia. Selain itu, malnutrisi pada lansia juga dapat menyebabkan berbagai komplikasi, termasuk penurunan massa otot (*sarcopenia*), gangguan sistem kekebalan tubuh, serta peningkatan risiko penyakit infeksi.

2.6 Malnutrisi

Malnutrisi adalah kondisi kesehatan yang serius yang terjadi ketika tubuh tidak menerima asupan nutrisi yang cukup, dan hal ini sering dialami oleh lansia. Kekurangan gizi ini dapat mengakibatkan tubuh tidak memperoleh kalori, protein, vitamin, dan mineral yang diperlukan untuk menjalankan fungsi dasar tubuh. Dampak malnutrisi pada lansia sangat signifikan, termasuk kerusakan jaringan tubuh, gangguan fungsi organ, melemahnya sistem kekebalan tubuh, serta meningkatnya risiko penyakit kronis dan infeksi. Menurut *World Health Organization* (WHO), malnutrisi di kalangan lansia dapat memperburuk kondisi kesehatan yang ada dan berkontribusi pada penurunan kualitas hidup secara keseluruhan. Malnutrisi pada lansia sering kali sulit dikenali karena dapat disalahartikan sebagai bagian dari proses penuaan normal. Tanda-tanda seperti penurunan berat badan yang tidak disengaja, kelemahan otot, kelelahan berkelanjutan, dan penurunan imunitas menjadi umum tetapi sering terabaikan.

Beberapa faktor yang menyebabkan malnutrisi pada lansia meliputi penurunan nafsu makan, masalah pencernaan, keterbatasan mobilitas, dan kondisi medis yang mempengaruhi penyerapan nutrisi. Penelitian oleh Shuremu *et al.* (2023) menunjukkan bahwa lansia yang mengalami masalah kesehatan kronis, seperti diabetes dan hipertensi, memiliki risiko lebih tinggi untuk mengalami kekurangan nutrisi, terutama protein dan mikronutrien yang penting. Kekurangan nutrisi menyebabkan penurunan massa otot, yang meningkatkan risiko jatuh dan cedera.

Lansia yang mengalami malnutrisi juga lebih rentan terhadap infeksi dan memiliki proses pemulihan yang lebih lambat. Penelitian oleh Moloney, L., & Jarrett, B. (2021) mengungkapkan bahwa intervensi nutrisi yang tepat dan deteksi dini dapat secara signifikan mengurangi risiko komplikasi yang dihadapi lansia, sekaligus meningkatkan kualitas hidup mereka .

Menurut *World Health Organization* (WHO), malnutrisi pada lansia dapat dikategorikan menjadi tiga tingkat keparahan, yaitu malnutrisi ringan, sedang, dan berat, berdasarkan tingkat kekurangan gizi yang dialami.

Pengelompokan ini disesuaikan oleh pakar dalam mendeteksi status gizi lansia secara lebih komprehensif, serta menentukan intervensi yang tepat sesuai dengan kondisi mereka. Pengelompokan ini mempertimbangkan penurunan berat badan, perubahan dalam asupan nutrisi, serta dampak terhadap fungsi fisik dan kesehatan secara keseluruhan.

- 1) Malnutrisi Ringan: Pada tahap ini, lansia mengalami sedikit penurunan berat badan atau pengurangan asupan nutrisi, tetapi masih mampu menjalankan aktivitas sehari-hari secara relatif normal. Gejala malnutrisi mungkin belum terlihat jelas, sehingga intervensi yang dibutuhkan bersifat pencegahan untuk mencegah kondisi semakin memburuk.

Intervensi:

- i. Pengaturan pola makan yang lebih seimbang dan bervariasi untuk mencegah malnutrisi lebih lanjut.
- ii. Kegiatan fisik ringan yang teratur untuk menjaga kesehatan fisik dan mobilitas.

- iii. Konseling psikologis ringan, jika diperlukan, untuk membantu manajemen stres.
 - iv. Pengawasan berkala terhadap status gizi untuk memastikan stabilitas kondisi.
- 2) Malnutrisi Sedang: Lansia pada tingkat ini mulai menunjukkan penurunan berat badan yang lebih jelas dan asupan nutrisi yang semakin berkurang. Gejala seperti kelelahan, kelemahan otot, atau penurunan daya tahan tubuh mungkin mulai tampak. Pada tahap ini, diperlukan intervensi gizi dan perhatian medis untuk mencegah dampak kesehatan yang lebih serius.
- Intervensi:
- i. Konsultasi ahli gizi untuk merancang rencana makan yang lebih kaya nutrisi dan bervariasi.
 - ii. Program rehabilitasi fisik yang ringan untuk meningkatkan mobilitas dan kekuatan fisik.
 - iii. Pemberian suplemen gizi, terutama untuk meningkatkan asupan protein dan mikronutrien.
 - iv. Pengawasan berkala oleh tenaga medis untuk mencegah perkembangan ke malnutrisi berat.
- 3) Malnutrisi Berat: Tingkat ini ditandai dengan penurunan berat badan yang signifikan dan kekurangan nutrisi yang parah, menyebabkan lansia mengalami kelemahan fisik, masalah mobilitas, serta peningkatan risiko infeksi. Intervensi gizi segera dan tindakan medis sangat diperlukan untuk mencegah komplikasi yang lebih serius serta meningkatkan kualitas hidup lansia.
- Intervensi:
- i. Pendekatan medis yang intensif, termasuk evaluasi komprehensif oleh dokter, ahli gizi, dan tim medis terkait.
 - ii. Pemberian nutrisi melalui metode enteral atau parenteral sesuai kebutuhan, untuk memastikan asupan nutrisi yang tepat bagi pasien.
 - iii. Terapi fisik dan intervensi mobilitas yang bertahap guna

membantu meningkatkan kekuatan fisik.

- iv. Konseling psikologis untuk mengelola stres, depresi, atau kecemasan yang dapat memperburuk kondisi malnutrisi.

2.7 Lansia

Lansia, atau orang tua, merujuk pada individu yang umumnya berusia di atas 60 tahun, sesuai dengan definisi yang diberikan oleh *World Health Organization* (WHO). Pada fase kehidupan ini, individu sering mengalami perubahan fisiologis yang signifikan yang dapat memengaruhi berbagai aspek kesehatan. Penurunan metabolisme adalah salah satu perubahan utama, yang sering mengakibatkan penurunan massa otot dan kekuatan fisik. Fungsi organ vital, seperti jantung dan ginjal, juga dapat mengalami penurunan, yang berkontribusi pada peningkatan risiko berbagai penyakit.

Selain itu, lansia cenderung lebih rentan terhadap masalah nutrisi. Beberapa faktor yang berkontribusi terhadap masalah ini antara lain berkurangnya nafsu makan, perubahan dalam kemampuan mencerna makanan, dan penurunan kemampuan untuk mengakses makanan bergizi. Masalah kesehatan, seperti gigi yang tanggal atau penyakit kronis, juga dapat mempengaruhi pola makan mereka.

Kekurangan asupan nutrisi pada lansia dapat mempercepat penurunan fungsi tubuh dan memperburuk kondisi kesehatan yang sudah ada, seperti meningkatkan risiko jatuh akibat kelemahan otot. Malnutrisi dapat berkontribusi terhadap peningkatan morbiditas dan mortalitas pada lansia, sehingga perhatian yang cukup terhadap status gizi lansia sangat penting untuk meningkatkan kualitas hidup dan mempertahankan kesehatan mereka di usia lanjut.

BAB 3

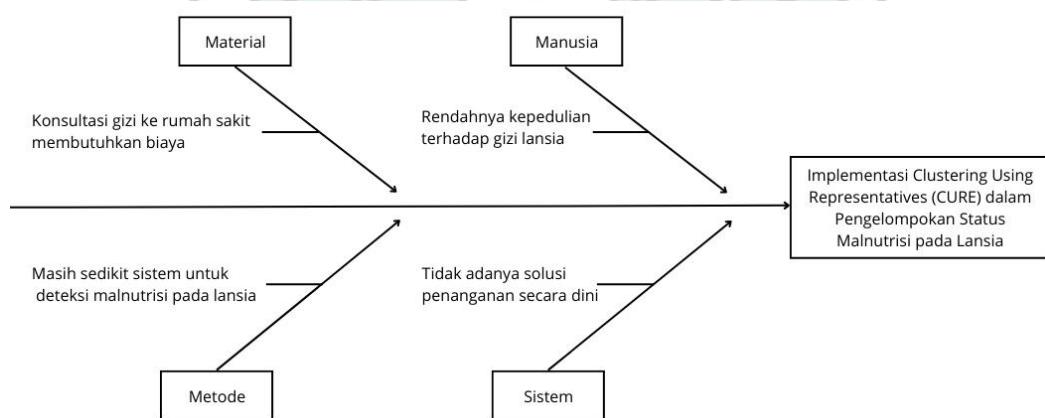
ANALISIS DAN PERANCANGAN

3.1. Analisis Sistem

Analisis sistem merupakan pendekatan pemecahan masalah yang membagi suatu komponen menjadi bagian-bagian lebih kecil untuk memudahkan identifikasi dan analisis, sehingga mengurangi potensi kesalahan dalam perancangan sistem. Proses analisis ini biasanya terdiri dari empat tahap utama: analisis masalah, analisis kebutuhan, analisis data, dan analisis proses. Setiap tahap memainkan peran penting dalam memastikan sistem yang dirancang dapat beroperasi sesuai dengan tujuan serta spesifikasi yang diharapkan.

3.1.1. Analisis Masalah

Langkah analisis masalah merupakan tahap awal dalam mengenali isu-isu yang terdapat pada sistem yang akan dikembangkan. Dalam penelitian ini, fokus masalahnya adalah bagaimana merancang sistem yang mampu mengelompokkan status malnutrisi pada lansia. Sistem tersebut bertujuan untuk menyediakan panduan lebih akurat dalam menentukan intervensi gizi yang sesuai bagi kelompok lansia yang berisiko mengalami malnutrisi. Untuk memahami penyebab masalah secara lebih rinci, metode seperti diagram Ishikawa atau fishbone diagram dapat digunakan. Contoh penerapan diagram Ishikawa/fishbone terlihat pada gambar berikut.



Gambar 3.1. Diagram Ishikawa/Fishbone

Pada Gambar 3.1., terdapat permasalahan-permasalahan dari 4 bagian. Karena beberapa permasalahan tersebut maka diperlukannya sebuah sistem pengelompokan untuk mendeteksi secara dini status malnutrisi pada lansia.

3.1.2. Analisis Kebutuhan

Tahap selanjutnya adalah melakukan identifikasi terhadap kebutuhan sistem. Langkah ini bertujuan untuk merinci komponen-komponen yang diperlukan untuk membangun sistem yang sesuai dengan tujuan yang ingin dicapai. Melalui proses ini, akan ditentukan elemen-elemen yang harus ada agar sistem dapat berfungsi sebagaimana mestinya. Analisis kebutuhan sistem mencakup aspek fungsional dan non-fungsional, dengan memastikan semua persyaratan pengguna dan kebutuhan proses dapat terpenuhi secara efektif dalam tahap perancangan.

3.1.3.1. Analisis Kebutuhan Fungsional

Beberapa hal yang diperlukan secara fungsional pada penelitian ini adalah:

1. Sistem mampu mengelompokkan status malnutrisi lansia berdasarkan data yang dikumpulkan.
2. Sistem mampu menghasilkan hasil pengelompokan status malnutrisi berdasarkan input data yang diberikan oleh pengguna untuk setiap lansia.
3. Sistem dapat memberikan rekomendasi intervensi gizi yang sesuai berdasarkan kelompok malnutrisi lansia.

3.1.3.2. Analisis Kebutuhan Non-Fungsional

Beberapa hal yang diperlukan secara non fungsional pada penelitian ini adalah:

1. Sistem memiliki *interface* yang mudah digunakan oleh *user*.
2. Sistem hanya dapat diakses dengan menggunakan jaringan internet.
3. Sistem dapat diakses secara langsung oleh pengguna tanpa menggunakan perangkat khusus.

3.1.3. Analisis Data

Pada penelitian ini, data yang digunakan dikumpulkan dari UPTD. Pelayanan Sosial Lansia Binjai serta melalui penyebaran formulir online kepada keluarga lansia. Dataset terdiri dari 517 baris data yang mencakup informasi demografi lansia, status kesehatan, asupan nutrisi harian, serta kondisi fisik yang relevan dengan malnutrisi. Data tersebut disimpan dalam format CSV agar memudahkan proses pengolahan menggunakan perangkat lunak seperti *Python*.

Berikut deskripsi mengenai atribut data yang digunakan dalam penelitian ini.

Tabel 3.1. Atribut Data dan Deskripsi

Atribut	Deskripsi
Nama	Nama lansia sebagai identitas data individu.
Umur	Usia lansia dalam tahun, menunjukkan kategori usia atau tingkat lanjutnya usia.
Jenis Kelamin	Jenis kelamin lansia, berisi nilai 'Laki-laki' atau 'Perempuan'.
Berat Badan	Berat badan lansia dalam kilogram (kg), menunjukkan kondisi massa tubuh lansia.
Tinggi Badan	Tinggi badan lansia dalam centimeter (cm).
Penurunan Berat Badan	Indikator apakah lansia mengalami penurunan berat badan signifikan, berkaitan dengan status gizi.
Frekuensi Makan	Seberapa sering lansia makan dalam sehari, sebagai indikator pola makan.
Variasi Makanan	Tingkat variasi jenis makanan yang dikonsumsi lansia, mengukur keragaman nutrisi yang diperoleh.
Asupan Protein	Seberapa sering dalam sehari lansia mengonsumsi makanan yang mengandung protein.
Mobilitas	Kemampuan lansia untuk bergerak secara mandiri, seperti berjalan atau bergerak di tempat tidur.
Aktivitas Sehari-hari	Kemampuan lansia dalam menjalani aktivitas harian, seperti makan, mandi, berpakaian, dll.
Kemampuan Mengunyah	Kondisi kesehatan mulut, mencakup gigi dalam kemampuan untuk mengunyah, yang dapat memengaruhi asupan makanan.
Penyakit Kronis	Adanya penyakit kronis pada lansia yang memengaruhi status gizi.

Atribut	Deskripsi
Stres	Adanya kondisi stres yang dialami lansia, yang dapat berdampak pada pola makan dan kesehatan umum.
Masalah Kognitif	Indikator adanya gangguan kognitif pada lansia.

Berikut ini adalah contoh 10 baris data lansia yang digunakan dalam penelitian ini.

Tabel 3.2. Data Lansia (Atribut Nama - Variasi Makanan)

Nama	Umur	Jenis Kelamin	Berat Badan	Tinggi Badan	Penurunan Berat Badan	Frekuensi Makan	Variasi Makanan
Junaidi	90	Laki-laki	70	166	0	Tiga kali atau lebih	Makanan bervariasi (misalnya daging, sayuran, buah)
Tumini	82	Perempuan	56	145	0	Tiga kali atau lebih	Makanan bervariasi (misalnya daging, sayuran, buah)
Yusraidi Sitompul	74	Laki-laki	52	146	0	Tiga kali atau lebih	Makanan bervariasi (misalnya daging, sayuran, buah)

Nama	Umur	Jenis Kelamin	Berat Badan	Tinggi Badan	Penurunan Berat Badan	Frekuensi Makan	Variasi Makanan
Puspita Itayana	64	Perempuan	48	157	4	Satu kali atau kurang	Makanan sangat terbatas atau monoton
Edwin Gunawan	66	Laki-laki	65	170	5	Satu kali atau kurang	Makanan sangat terbatas atau monoton
Ratna Maulida	67	Perempuan	73	154	4	Satu kali atau kurang	Makanan sangat terbatas atau monoton
Ruzila	65	Perempuan	45	148	2	Dua kali	Makanan terbatas pada beberapa jenis saja
Zaimah	64	Perempuan	53	161	1	Tiga kali atau lebih	Makanan bervariasi (misalnya daging, sayuran, buah)
Ida	65	Perempuan	80	154	2	Dua kali	Makanan terbatas pada beberapa jenis saja

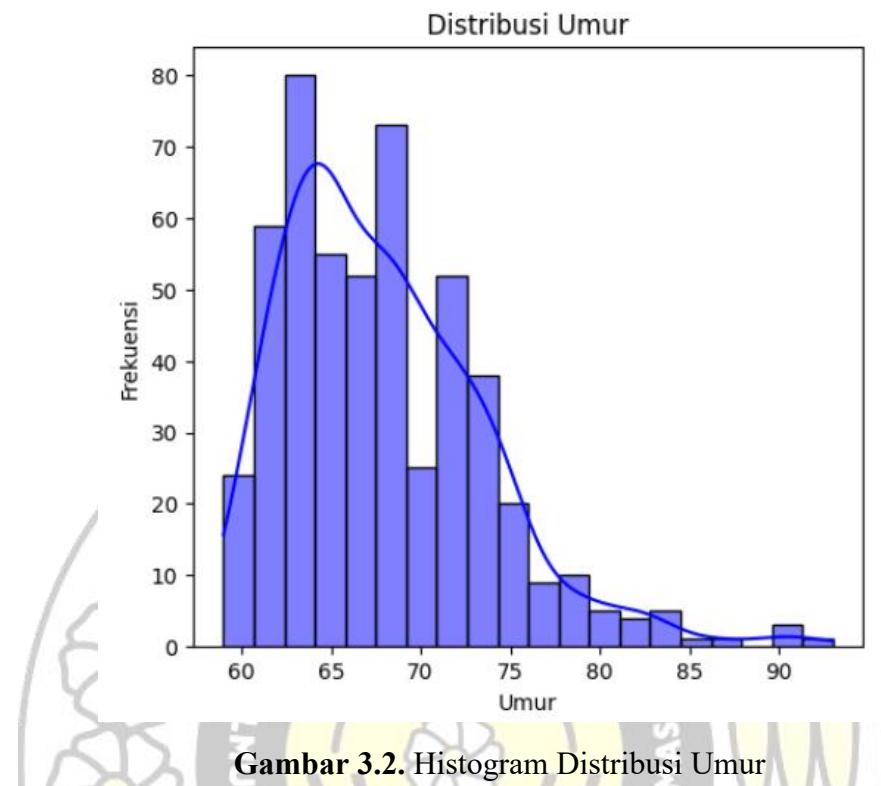
Nama	Umur	Jenis Kelamin	Berat Badan	Tinggi Badan	Penurunan Berat Badan	Frekuensi Makan	Variasi Makanan
Siti Aminah	69	Perempuan	55	154	3	Tiga kali atau lebih	Makanan terbatas pada beberapa jenis saja

Tabel 3.3. Data Lansia (Atribut Asupan Protein - Masalah Kognitif)

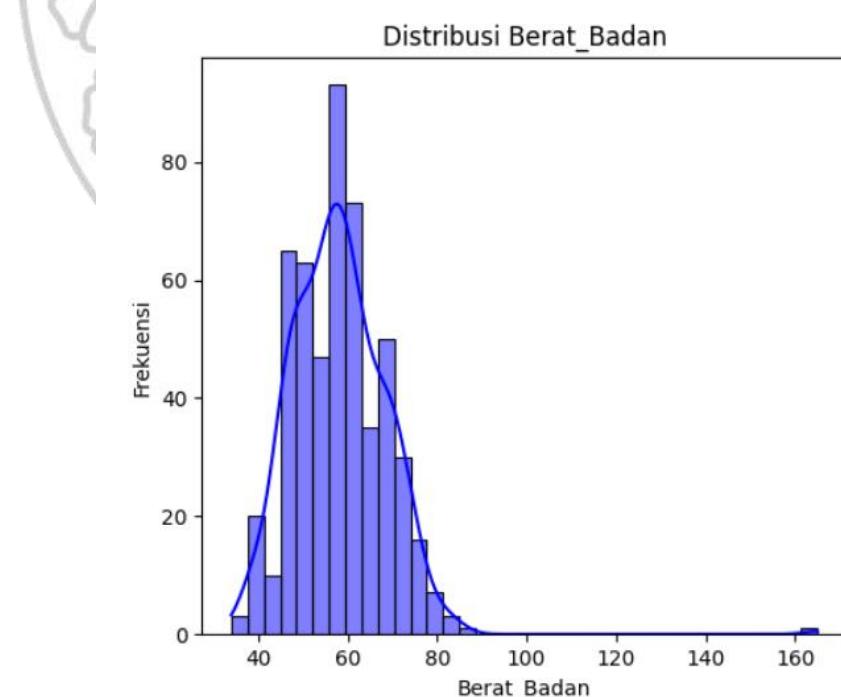
Asupan Protein	Mobilitas	Aktivitas Sehari-hari	Kemampuan Mengunyah	Penyakit Kronis	Stres	Masalah Kognitif
Dua kali sehari atau lebih	Bergerak dengan bebas tanpa bantuan	Tidak ada kesulitan	Tidak ada masalah	Tidak	Tidak	Ada sedikit masalah daya ingat
Dua kali sehari atau lebih	Bergerak dengan bebas tanpa bantuan	Tidak ada kesulitan	Tidak ada masalah	Tidak	Tidak	Ada sedikit masalah daya ingat
Dua kali sehari atau lebih	Bergerak dengan bebas tanpa bantuan	Tidak ada kesulitan	Tidak ada masalah	Tidak	Tidak	Tidak ada masalah daya ingat
Kurang dari sekali sehari	Tidak dapat bergerak tanpa bantuan orang lain	Kesulitan berat	Kesulitan berat	Ya	Ya	Masalah daya ingat yang signifikan

Asupan Protein	Mobilitas	Aktivitas Sehari-hari	Kemampuan Mengunyah	Penyakit Kronis	Stres	Masalah Kognitif
Kurang dari sekali sehari	Tidak dapat bergerak tanpa bantuan orang lain	Kesulitan berat	Kesulitan berat	Ya	Ya	Masalah daya ingat yang signifikan
Sekali sehari	Bergerak dengan bantuan	Kesulitan ringan	Ada kesulitan ringan	Ya	Tidak	Ada sedikit masalah daya ingat
Dua kali sehari atau lebih	Bergerak dengan bebas tanpa bantuan	Tidak ada kesulitan	Ada kesulitan ringan	Ya	Tidak	Tidak ada masalah daya ingat
Dua kali sehari atau lebih	Bergerak dengan bebas tanpa bantuan	Tidak ada kesulitan	Tidak ada masalah	Tidak	Tidak	Ada sedikit masalah daya ingat
Dua kali sehari atau lebih	Tidak dapat bergerak tanpa bantuan orang lain	Kesulitan berat	Tidak ada masalah	Tidak	Ya	Ada sedikit masalah daya ingat
Sekali sehari	Bergerak dengan bantuan	Kesulitan ringan	Ada kesulitan ringan	Ya	Tidak	Ada sedikit masalah daya ingat

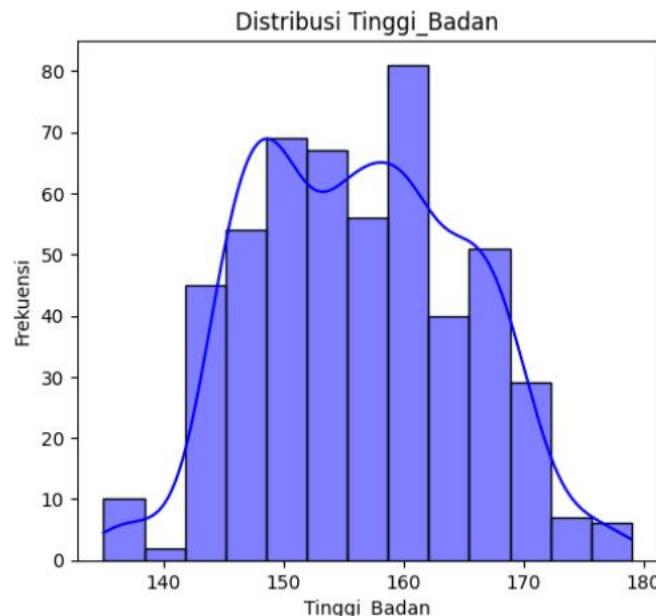
Berikut adalah representasi grafis sebaran data untuk variabel umur, berat badan, dan tinggi badan dari data penelitian.



Gambar 3.2. Histogram Distribusi Umur



Gambar 3.3. Histogram Distribusi Berat Badan



Gambar 3.4. Histogram Distribusi Tinggi Badan

Histogram menunjukkan frekuensi data dalam setiap interval nilai, dengan sumbu X mewakili nilai data dan sumbu Y menunjukkan jumlah kemunculannya. Kurva KDE (*Kernel Density Estimation*) pada histogram di atas memberikan gambaran distribusi data yang lebih halus dan kontinu. Garis kurva tersebut membantu mengidentifikasi pola distribusi seperti simetris atau miring (*skewed*). Dengan visualisasi ini, dapat terlihat karakteristik data, seperti konsentrasi nilai, variasi, dan kemungkinan adanya *outlier*, serta memperoleh wawasan tentang bagaimana data tersebar.

Dataset yang akan digunakan harus melalui tahap *pre-processing* untuk memastikan kualitas data yang optimal. Salah satu langkah penting dalam proses ini adalah *encoding* data kategorik. Encoding ini bertujuan untuk mengonversi data kategorik menjadi format numerik yang dapat diproses oleh algoritma *machine learning*.

Tabel 3.4. Nilai *Encoding* Atribut

Atribut	Nilai	Nilai Angka
Jenis Kelamin	Perempuan	1
	Laki-laki	0
Penyakit Kronis	Ya	1
	Tidak	0

Atribut	Nilai	Nilai Angka
Stres	Ya	1
	Tidak	0
Frekuensi Makan	Tiga kali atau lebih	3
	Dua kali	2
	Satu kali atau kurang	1
Variasi Makanan	Makanan bervariasi (misalnya daging, sayuran, buah)	3
	Makanan terbatas pada beberapa jenis saja	2
	Makanan sangat terbatas atau monoton	1
Asupan Protein	Dua kali sehari atau lebih	3
	Sekali sehari	2
	Kurang dari sekali sehari atau tidak pernah	1
Mobilitas	Bergerak dengan bebas tanpa bantuan	3
	Bergerak dengan bantuan	2
	Tidak dapat bergerak tanpa bantuan orang lain	1
Aktivitas Sehari-hari	Tidak ada kesulitan	3
	Kesulitan ringan	2
	Kesulitan berat	1
Kemampuan Mengunyah	Tidak ada masalah	3
	Ada kesulitan ringan	2
	Kesulitan berat	1
Masalah Kognitif	Tidak ada masalah daya ingat	3
	Ada sedikit masalah daya ingat	2
	Masalah daya ingat yang signifikan	1

Selain encoding, proses standarisasi dilakukan menggunakan *Standard Scaler* untuk mengubah nilai atribut numerik menjadi rentang yang sama. Hal ini memastikan bahwa atribut yang memiliki rentang nilai yang besar tidak mendominasi analisis atau hasil pengelompokan. Proses ini bertujuan untuk menghapus bias yang disebabkan oleh perbedaan skala antar atribut.

Tabel 3.5 Data Lansia Setelah Pre-Processing (Atribut Umur - Penurunan Berat Badan)

Umur	Jenis Kelamin	Berat Badan	Tinggi Badan	Penurunan Berat Badan
39.250.210.869.447.200	-10.618.719.397.652.900	1.110.797.059.103.820	11.051.722.900.826.600	-13.476.449.715.377.300
25.014.953.752.993.800	9.417.331.436.604.600	-1.889.037.352.371.850	-13.071.428.593.800.300	-13.476.449.715.377.300
10.779.696.636.540.500	-10.618.719.397.652.900	-5.602.468.193.346.170	-11.922.707.094.056.200	-13.476.449.715.377.300
-7.014.374.759.026.090	9.417.331.436.604.600	-9.315.899.034.320.490	7.132.294.031.293.520	10.991.766.273.560.500
-34.555.604.799.127.600	-10.618.719.397.652.900	646.618.203.982.037	15.646.608.899.803.200	1.710.882.027.079.500
-16.761.533.403.560.900	9.417.331.436.604.600	1.389.304.372.176.900	-27.329.350.961.030.800	10.991.766.273.560.500
-5.234.967.619.469.420	9.417.331.436.604.600	-12.100.972.165.051.200	-9.625.264.094.567.950	-12.423.417.209.083.500
-7.014.374.759.026.090	9.417.331.436.604.600	-4.674.110.483.102.590	5.308.115.402.105.930	-7.359.395.718.142.830
-5.234.967.619.469.420	9.417.331.436.604.600	2.039.154.769.347.400	-27.329.350.961.030.800	-12.423.417.209.083.500
1.882.660.938.757.230	9.417.331.436.604.600	-2.817.395.062.615.430	-27.329.350.961.030.800	48.747.122.763.261.100
39.250.210.869.447.200	-10.618.719.397.652.900	1.110.797.059.103.820	11.051.722.900.826.600	-13.476.449.715.377.300

Tabel 3.6 Data Lansia Setelah Pre-Processing (Atribut Frekuensi Makan - Aktivitas Sehari-hari)

Frekuensi Makan	Variasi Makanan	Asupan Protein	Mobilitas	Aktivitas Sehari-hari
12.848.916.786.193.000	12.848.916.786.193.000	12.848.916.786.193.000	12.838.677.875.535.600	12.838.677.875.535.600
12.848.916.786.193.000	12.848.916.786.193.000	12.848.916.786.193.000	12.838.677.875.535.600	12.838.677.875.535.600
12.848.916.786.193.000	12.848.916.786.193.000	12.848.916.786.193.000	12.838.677.875.535.600	12.838.677.875.535.600
-15.358.683.971.606.500	-15.358.683.971.606.500	15.358.683.971.606.500	-14.760.475.811.166.800	-14.760.475.811.166.800
-15.358.683.971.606.500	-15.358.683.971.606.500	15.358.683.971.606.500	-14.760.475.811.166.800	-14.760.475.811.166.800
-15.358.683.971.606.500	-15.358.683.971.606.500	15.358.683.971.606.500	-14.760.475.811.166.800	-14.760.475.811.166.800
-12.548.835.927.067.500	-12.548.835.927.067.500	12.548.835.927.067.500	-9.608.989.678.155.640	-9.608.989.678.155.640
12.848.916.786.193.000	12.848.916.786.193.000	12.848.916.786.193.000	12.838.677.875.535.600	12.838.677.875.535.600
-12.548.835.927.067.500	-12.548.835.927.067.500	12.548.835.927.067.500	-9.608.989.678.155.640	-9.608.989.678.155.640

Frekuensi Makan	Variasi Makanan	Asupan Protein	Mobilitas	Aktivitas Sehari-hari
-12.548.835.927.067.500	-12.548.835.927.067.500	-12.548.835.927.067.500	-9.608.989.678.155.640	-9.608.989.678.155.640
12.848.916.786.193.000	12.848.916.786.193.000	12.848.916.786.193.000	12.838.677.875.535.600	12.838.677.875.535.600

Tabel 3.7 Data Lansia Setelah *Pre-Processing* (Atribut Kemampuan Mengunyah - Masalah Kognitif)

Kemampuan Mengunyah	Penyakit Kronis	Stres	Masalah Kognitif
12.838.677.875.535.600	-15.282.289.982.092.800	-5.468.043.296.046.320	12.838.677.875.535.600
12.838.677.875.535.600	-15.282.289.982.092.800	-5.468.043.296.046.320	12.838.677.875.535.600
12.838.677.875.535.600	-15.282.289.982.092.800	-5.468.043.296.046.320	12.838.677.875.535.600
-14.760.475.811.166.800	6.543.521.953.658.520	18.288.077.578.373.400	-14.760.475.811.166.800
-14.760.475.811.166.800	6.543.521.953.658.520	18.288.077.578.373.400	-14.760.475.811.166.800
-14.760.475.811.166.800	6.543.521.953.658.520	18.288.077.578.373.400	-14.760.475.811.166.800
-9.608.989.678.155.640	6.543.521.953.658.520	-5.468.043.296.046.320	-9.608.989.678.155.640
12.838.677.875.535.600	-15.282.289.982.092.800	-5.468.043.296.046.320	12.838.677.875.535.600
-9.608.989.678.155.640	6.543.521.953.658.520	-5.468.043.296.046.320	-9.608.989.678.155.640
-9.608.989.678.155.640	6.543.521.953.658.520	-5.468.043.296.046.320	-9.608.989.678.155.640
12.838.677.875.535.600	-15.282.289.982.092.800	-5.468.043.296.046.320	12.838.677.875.535.600

Berikut adalah uraian rinci mengenai statistik data setelah melalui tahap pre-processing. Statistik deskriptif dari data yang telah diproses, seperti rata-rata (mean), standar deviasi, nilai minimum dan maksimum, serta distribusi data, disajikan untuk memberikan gambaran menyeluruh tentang kualitas dan karakteristik dataset yang akan digunakan dalam analisis lebih lanjut.

Tabel 3.8. Deskripsi Statistik Data setelah *Pre-processing* (Atribut Umur - Penurunan Berat Badan)

	Umur	Jenis Kelamin	Berat Badan	Tinggi Badan	Penurunan Berat Badan
count	517.0	517.0	517.0	517.0	517.0
mean	-8,52E-01	1,58E-01	-1,37E+00	9,21E-01	-5,50E-02
std	10.009.685.232 .294.400	10.009.685.23 2.294.400	1.000.968.523.22 9.430	10.009.685.232.2 94.300	10.009.685.232.2 94.300
min	15.911.410.456 .809.400	10.618.719.39 7.652.900	2.231.290.697.77 3.060	24.558.643.591.2 41.800	13.476.449.715.3 77.300
25%	7.014.374.759. 026.090	10.618.719.39 7.652.900	7.459.183.613.83 3.330	847.654.259.482. 381	7.359.395.718.14 2.830
50%	16.761.533.403 .560.900	9.417.331.436 .604.600	3.232.193.188.46 9.070	4.354.920.966.14 7.920	12.423.417.209.0 83.500
75%	544.147.521.78 7.057	9.417.331.436 .604.600	646.618.203.982. 037	7.605.558.401.59 4.220	48.747.122.763.2 61.100
max	4.458.843.228. 811.720	9.417.331.436 .604.600	9.930.195.306.41 7.830	2.598.510.239.75 0.050	1.710.882.027.07 9.500

Tabel 3.9. Deskripsi Statistik Data setelah *Pre-processing* (Atribut Frekuensi Makan - Aktivitas Sehari-hari)

	Frekuensi Makan	Variasi Makanan	Asupan Protein	Mobilitas	Aktivitas Sehari-hari
count	517.0	517.0	517.0	517.0	517.0
mean	5,50E-02	5,50E-02	5,50E-02	-3,85E-01	-3,85E-01
std	10.009.685.232 .294.300	10.009.685.232 .294.300	10.009.685.232.2 94.300	10.009.685.23 2.294.300	10.009.685.232. 294.300
min	15.358.683.971 .606.500	15.358.683.971 .606.500	15.358.683.971.6 06.500	14.760.475.81 1.166.800	14.760.475.811. 166.800
25%	12.548.835.927 .067.500	12.548.835.927 .067.500	12.548.835.927.0 67.500	9.608.989.678 .155.640	9.608.989.678.1 55.640
50%	12.548.835.927 .067.500	12.548.835.927 .067.500	12.548.835.927.0 67.500	9.608.989.678 .155.640	9.608.989.678.1 55.640
75%	12.848.916.786 .193.000	12.848.916.786 .193.000	12.848.916.786.1 93.000	12.838.677.87 5.535.600	12.838.677.875. 535.600
max	12.848.916.786 .193.000	12.848.916.786 .193.000	12.848.916.786.1 93.000	12.838.677.87 5.535.600	12.838.677.875. 535.600

Tabel 3.10. Deskripsi Statistik Data setelah *Pre-processing* (Atribut Kemampuan Mengunyah - Masalah Kognitif)

	Kemampuan Mengunyah	Penyakit Kronis	Stres	Masalah Kognitif
count	517.0	517.0	517.0	517.0
mean	-3,85E-01	1,10E+00	-2,75E-01	-3,85E-01
std	10.009.685.232.29 4.300	10.009.685.232.29 4.300	10.009.685.232.294.30 0	10.009.685.232.294.3 00
min	14.760.475.811.16 6.800	15.282.289.982.09 2.800	-5.468.043.296.046.320	14.760.475.811.166.8 00
25%	- 9.608.989.678.155. 640	- 15.282.289.982.09 2.800	-5.468.043.296.046.320	- 9.608.989.678.155.64 0
50%	- 9.608.989.678.155. 640	6.543.521.953.658 .520	-5.468.043.296.046.320	- 9.608.989.678.155.64 0
75%	12.838.677.875.53 5.600	6.543.521.953.658 .520	-5.468.043.296.046.320	12.838.677.875.535.6 00
max	12.838.677.875.53 5.600	6.543.521.953.658 .520	18.288.077.578.373.40 0	12.838.677.875.535.6 00

Tahapan analisis data ini bertujuan untuk memastikan bahwa dataset yang digunakan bersih dan siap untuk diolah, sehingga dapat memberikan hasil yang akurat dalam proses pengelompokan malnutrisi pada lansia.

3.1.4. Analisis Proses

Dalam penelitian ini, sistem akan dibangun dengan menerapkan metode *Clustering Using Representatives* (CURE) untuk mengelompokkan status malnutrisi pada lansia. Hasil pengelompokan ini nantinya dapat dijadikan acuan dalam mendiagnosa status malnutrisi lansia. Adapun proses kerja sistem dalam penelitian ini adalah sebagai berikut.

1. Pengumpulan dan Persiapan Data

- a. Data Demografis dan Gizi Lansia: Data ini menjadi input utama untuk analisis. Data ini mencakup nama, usia, jenis kelamin, berat badan, tinggi badan, penurunan berat badan, asupan makanan, asupan protein, kemampuan mobilitas, kondisi kesehatan, dan fungsi kognitif.

b. *Preprocessing*:

- i. *Attribute Removal*: Menghapus atribut yang tidak memiliki korelasi yang kuat dengan variabel target.
- ii. *Encoding*: Mengubah data yang berbentuk kategorikal menjadi numerik.
- iii. Standarisasi: Data numerik distandarisasi agar semua fitur memiliki skala yang seragam, sehingga tidak ada fitur yang lebih dominan dalam proses pengelompokan.

2. Penerapan Model CURE

- a. Algoritma CURE diterapkan pada data yang telah dipilih. Algoritma ini akan mengelompokkan data menjadi beberapa *cluster* berdasarkan kemiripan karakteristik. Dalam konteks ini, *cluster* akan mewakili kelompok-kelompok lansia dengan status gizi yang berbeda (gizi normal, malnutrisi ringan, dan malnutrisi berat).
- b. Proses perhitungan jarak antar cluster dan pembaruan centroid: Sebelum perhitungan jarak antar cluster dilakukan, sebuah matriks jarak dibuat untuk menyimpan jarak antar cluster yang ada. Matriks ini mempermudah proses iterasi dalam mengidentifikasi cluster dengan jarak terkecil. Jarak antar cluster dihitung menggunakan rumus Euclidean, seperti yang ditunjukkan pada rumus 2.1. Pada setiap iterasi, jarak terkecil antara dua cluster diidentifikasi melalui matriks jarak, kemudian kedua cluster tersebut digabungkan. Setelah penggabungan, centroid baru cluster yang terbentuk dihitung ulang menggunakan rumus yang dapat dilihat pada rumus 2.4.

Iterasi 1:

Titik dari Cluster 252 (A):

[-0.70143748, 0.94173314, 0.36811089, 0.87542799, -0.73593957, 1.28489168, 1.28489168, 1.28489168, 1.28386779, 1.28386779, 1.28386779, -1.528229, -0.54680433, 1.28386779]

Titik dari Cluster 224 (B):

[-0.70143748, 0.94173314, 0.36811089, 0.87542799, -
0.73593957, 1.28489168, 1.28489168, 1.28489168, 1.28386779,
1.28386779, 1.28386779, -1.528229, -0.54680433, 1.28386779]

Jarak Antar Titik

$$d(A, B) = \sqrt{\sum_{i=1}^n (A_i - B_i)^2} = \sqrt{(0)^2} = 0.0$$

Pembaruan Centroid

C : [-0.70143748, 0.94173314, 0.36811089, 0.87542799, -
0.73593957, 1.28489168, 1.28489168, 1.28489168, 1.28386779,
1.28386779, 1.28386779, -1.528229, -0.54680433, 1.28386779]

Iterasi 2:

Titik dari Cluster 314 (A):

[-0.52349676, -1.06187194, 0.18243935, 1.56466089, -
0.12423417, -0.12548836, -0.12548836, -0.12548836, -
0.0960899, -0.0960899, -0.0960899, 0.6543522, -0.54680433,
-0.0960899]

Titik dari Cluster 300 (B):

[-0.52349676, -1.06187194, 0.18243935, 1.56466089, -
0.12423417, -0.12548836, -0.12548836, -0.12548836, -0.0960899,
-0.0960899, -0.0960899, 0.6543522, -0.54680433, -0.0960899]

Jarak Antar Titik

$$d(A, B) = \sqrt{\sum_{i=1}^n (A_i - B_i)^2} = \sqrt{(0)^2} = 0.0$$

Pembaruan Centroid

C : [-0.52349676, -1.06187194, 0.18243935, 1.56466089, -
0.12423417, -0.12548836, -0.12548836, -0.12548836, -
0.0960899, -0.0960899, -0.0960899, 0.6543522, -0.54680433, -
0.0960899]

Iterasi 3:

Titik dari Cluster 319(A):

[-1.0573189, 0.94173314, 0.08960358, -0.27329351, -0.12423417, -0.12548836, -0.12548836, -0.12548836, -0.0960899, -0.0960899, -0.0960899, 0.6543522, -0.54680433, -0.0960899]

Titik dari Cluster 271 (B):

[-1.0573189, 0.94173314, 0.08960358, -0.27329351, -0.12423417, -0.12548836, -0.12548836, -0.12548836, -0.0960899, -0.0960899, 0.6543522, -0.54680433, -0.0960899]

Jarak Antar Titik

$$d(A, B) = \sqrt{\sum_{i=1}^n (A_i - B_i)^2} = \sqrt{(0)^2} = 0.0$$

Pembaruan Centroid

C : [-1.0573189, 0.94173314, 0.08960358, -0.27329351, -0.12423417, -0.12548836, -0.12548836, -0.12548836, -0.0960899, -0.0960899, -0.0960899, 0.6543522, -0.54680433, -0.0960899]

Iterasi 4:

Titik dari Cluster 349 (A):

[-0.52349676, 0.94173314, 0.18243935, -0.15842136, -0.12423417, -0.12548836, -0.12548836, -0.12548836, -0.0960899, -0.0960899, -0.0960899, 0.6543522, -0.54680433, -0.0960899]

Titik dari Cluster 294 (B):

[-0.52349676, 0.94173314, 0.18243935, -0.15842136, -0.12423417, -0.12548836, -0.12548836, -0.12548836, -0.0960899, -0.0960899, -0.0960899, 0.6543522, -0.54680433, -0.0960899]

Jarak Antar Titik

$$d(A, B) = \sqrt{\sum_{i=1}^n (A_i - B_i)^2} = \sqrt{(0)^2} = 0.0$$

Pembaruan Centroid

C : [-0.52349676, 0.94173314, 0.18243935, -0.15842136, -0.12423417, -0.12548836, -0.12548836, -0.12548836, -0.0960899, -0.0960899, -0.0960899, 0.6543522, -0.54680433, -0.0960899]

Iterasi 5:

Titik dari Cluster 113 (A):

[-0.70143748, 0.94173314, 0.27527512, -0.73278211, -0.73593957, -0.12548836, -0.12548836, -0.12548836, -0.0960899, -0.0960899, 0.6543522, -0.54680433, -0.0960899]

Titik dari Cluster 13 (B):

[-0.70143748, 0.94173314, 0.36811089, -0.73278211, -0.73593957, -0.12548836, -0.12548836, -0.12548836, -0.0960899, -0.0960899, -0.0960899, 0.6543522, -0.54680433, -0.0960899]

Jarak Antar Titik

$$d(A, B) = \sqrt{\sum_{i=1}^n (A_i - B_i)^2} = \sqrt{(0.09283577^2)} \\ = \sqrt{0.00861848} = 0.09283577$$

Pembaruan Centroid

$$\text{Atribut 3 : } \frac{0.27527512 + 0.36811089}{2} = 0.32169301$$

C : [-0.70143748, 0.94173314, 0.32169301, -0.73278211, -0.73593957, -0.12548836, -0.12548836, -0.12548836, -0.0960899, -0.0960899, -0.0960899, 0.6543522, -0.54680433, -0.0960899]

Iterasi 6:

Titik dari Cluster 326 (A):

[-0.87937819, 0.94173314, 0.08960358, -0.15842136, -0.12423417, -0.12548836, -0.12548836, -0.12548836, -0.0960899]

0.0960899, -0.0960899, -0.0960899, 0.6543522, -0.54680433, -0.0960899]

Titik dari Cluster 264 (B):

[-0.87937819, 0.94173314, 0.18243935, -0.15842136, -0.12423417, -0.12548836, -0.12548836, -0.12548836, -0.0960899, -0.0960899, -0.0960899, 0.6543522, -0.54680433, -0.0960899]

Jarak Antar Titik

$$d(A, B) = \sqrt{\sum_{i=1}^n (A_i - B_i)^2} = \sqrt{(0.09283577)^2} \\ = \sqrt{0.00861848} = 0.09283577$$

Pembaruan Centroid

$$\text{Atribut 3 : } \frac{0.08960358+0.18243935}{2} = 0.136021465$$

C : [-0.87937819, 0.94173314, 0.136021465, -0.15842136, -0.12423417, -0.12548836, -0.12548836, -0.12548836, -0.0960899, -0.0960899, 0.6543522, -0.54680433, -0.0960899]

Iterasi 7:

Titik dari Cluster 360 (A):

[-0.87937819, -1.06187194, -0.28173951, 0.41593939, -0.12423417, -0.12548836, -0.12548836, -0.12548836, -0.0960899, -0.0960899, -0.0960899, 0.6543522, -0.54680433, -0.0960899]

Titik dari Cluster 166 (B):

[-0.87937819, -1.06187194, -0.18890374, 0.41593939, -0.12423417, -0.12548836, -0.12548836, -0.12548836, -0.0960899, -0.0960899, -0.0960899, 0.6543522, -0.54680433, -0.0960899]

Jarak Antar Titik

$$d(A, B) = \sqrt{\sum_{i=1}^n (A_i - B_i)^2} = \sqrt{(0.09283577)^2}$$

$$= \sqrt{0.00861848} = 0.09283577$$

Pembaruan Centroid

$$\text{Atribut 3 : } \frac{(-0.28173951) + (-0.18890374)}{2} = 0.235321625$$

$$C : [-0.87937819, 0.94173314, 0.235321625, -0.15842136, -0.12423417, -0.12548836, -0.12548836, -0.12548836, -0.0960899, -0.0960899, -0.0960899, 0.6543522, -0.54680433, -0.0960899]$$

Iterasi 8:

Titik dari Cluster 139 (A):

$$[1.25591038, 0.94173314, -1.11726145, -0.84765426, -1.34764497, 1.28489168, 1.28489168, 1.28489168, 1.28386779, 1.28386779, 1.28386779, -1.528229, -0.54680433, 1.28386779]$$

Titik dari Cluster 44 (B):

$$[1.25591038, 0.94173314, -1.11726145, -0.96252641, -1.34764497, 1.28489168, 1.28489168, 1.28489168, 1.28386779, 1.28386779, 1.28386779, -1.528229, -0.54680433, 1.28386779]$$

Jarak Antar Titik

$$d(A, B) = \sqrt{\sum_{i=1}^n (A_i - B_i)^2} = \sqrt{(0.11487215)^2} \\ = \sqrt{0.01319966} = 0.11487215$$

Pembaruan Centroid

$$\text{Atribut 4 : } \frac{(-0.84765426) + (-0.96252641)}{2} = -0.90509033$$

$$C : [1.25591038, 0.94173314, -1.11726145, -0.90509033, -1.34764497, 1.28489168, 1.28489168, 1.28489168, 1.28386779, 1.28386779, 1.28386779, -1.528229, -0.54680433, 1.28386779]$$

Iterasi 9:

Titik dari Cluster 294 (A):

$$[-1.0573189, 0.94173314, 1.11079706, 0.30106724, 1.09917663, -1.5358684, -1.5358684, -1.5358684, -1.47604758,$$

-1.47604758, -1.47604758, 0.6543522, 1.82880776, -1.47604758]

Titik dari Cluster 150 (B):

[-1.0573189, 0.94173314, 1.11079706, 0.18619509, 1.09917663, -1.5358684, -1.5358684, -1.5358684, -1.47604758, -1.47604758, 0.6543522, 1.82880776, -1.47604758]

Jarak Antar Titik

$$\begin{aligned} d(A, B) &= \sqrt{\sum_{i=1}^n (A_i - B_i)^2} = \sqrt{(0.11487215^2} \\ &= \sqrt{0.01319966} = 0.11487215 \end{aligned}$$

Pembaruan Centroid

$$\text{Atribut 4 : } \frac{0.30106724 + 0.18619509}{2} = 0.24363117$$

C : [-1.0573189, 0.94173314, 1.11079706, 0.24363117, 1.09917663, -1.5358684, -1.5358684, -1.5358684, -1.47604758, -1.47604758, 0.6543522, 1.82880776, -1.47604758]

Iterasi 10:

Titik dari Cluster 336 (A):

[-0.70143748, 0.94173314, 0.18243935, 0.07132294, -0.12423417, -0.12548836, -0.12548836, -0.12548836, -0.0960899, -0.0960899, -0.0960899, 0.6543522, -0.54680433, -0.0960899]

Titik dari Cluster 317 (B):

[-0.70143748, 0.94173314, 0.18243935, -0.04354921, -0.12423417, -0.12548836, -0.12548836, -0.12548836, -0.0960899, -0.0960899, -0.0960899, 0.6543522, -0.54680433, -0.0960899]

Jarak Antar Titik

$$\begin{aligned} d(A, B) &= \sqrt{\sum_{i=1}^n (A_i - B_i)^2} = \sqrt{(0.11487215^2} \\ &= \sqrt{0.01319966} = 0.11487215 \end{aligned}$$

Pembaruan Centroid

$$\text{Atribut 4 : } \frac{0.07132294 + (-0.04354921)}{2} = 0.01388687$$

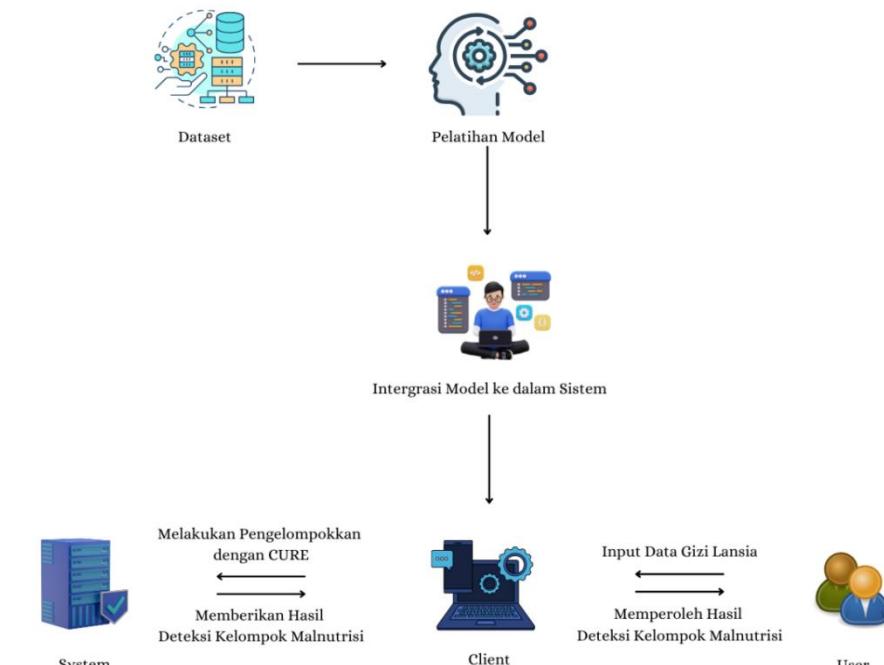
C : [-0.70143748, 0.94173314, 0.18243935, 0.01388687, -0.12423417, -0.12548836, -0.12548836, -0.12548836, -0.0960899, -0.0960899, -0.0960899, 0.6543522, -0.54680433, -0.0960899]

Proses ini diulang secara iteratif untuk menggabungkan cluster dengan jarak terdekat hingga akhirnya terbentuk cluster final yang optimal. Jumlah cluster optimal ini diidentifikasi menggunakan metode *silhouette analysis*, yang didasarkan pada karakteristik dataset yang digunakan pada penelitian.

3. Integrasi ke dalam Sistem *Website*
 - a. *Framework Streamlit*: Model CURE yang telah dilatih akan diintegrasikan ke dalam sebuah file dengan format pkl. *Framework streamlit* ini akan menjadi jembatan antara sistem *website* dan model.
 - b. Input *User*: *User* memasukkan data gizi lansia melalui *website*. Data ini kemudian dikirim ke model.
 - c. Pemrosesan oleh Model: Data input diteruskan kepada model CURE. Model akan memproses data baru ini dan menentukan ke *cluster* mana data tersebut termasuk.
 - d. Output: Hasil deteksi (kelompok malnutrisi) kemudian dikirim kembali ke *website* untuk ditampilkan kepada *user*.
4. Tampilan Hasil di *Website*
 - a. *User Interface*: Sistem akan menampilkan hasil deteksi dalam bentuk yang mudah dipahami oleh *user*, yaitu dalam bentuk label.
 - b. Informasi Tambahan: Selain hasil deteksi, sistem juga menampilkan informasi tambahan berupa intervensi yang sesuai dengan kelompok hasil deteksi.

3.2. Diagram Umum Sistem

Diagram umum sistem merupakan sebuah diagram yang menggambarkan keseluruhan proses, alur, serta seluruh komponen yang saling berinteraksi yang terdapat pada sistem. Rancangan sistem dapat dilihat melalui gambar di bawah ini.



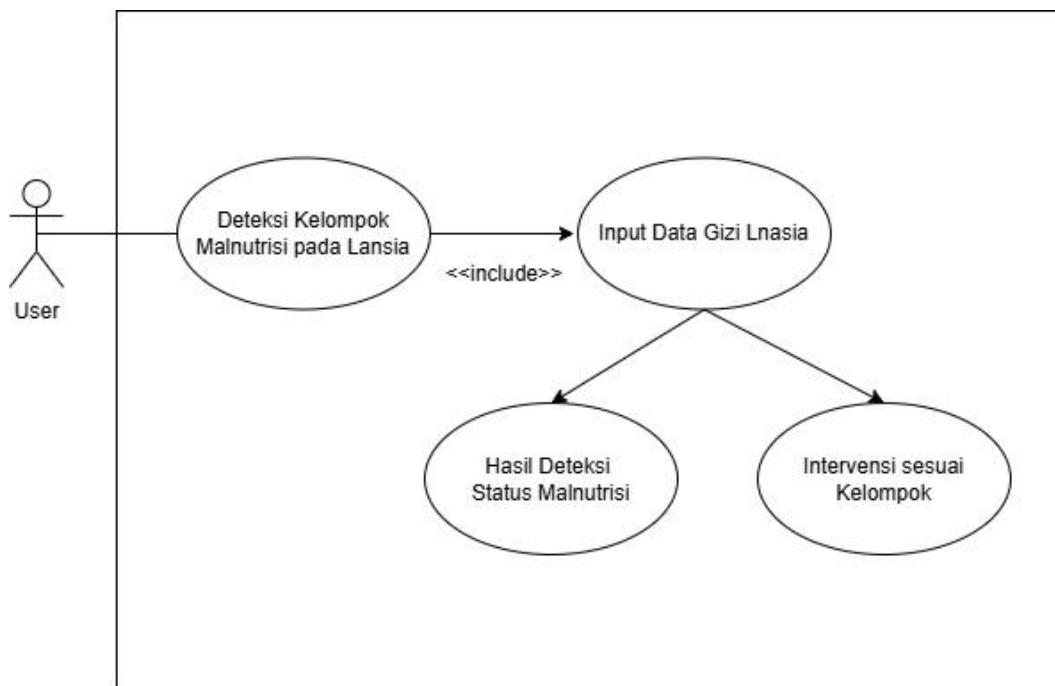
Gambar 3.5. Diagram Umum Sistem

3.3. Pemodelan Sistem

Setelah tahapan analisis sistem, selanjutnya adalah tahapan pemodelan sistem yang bertujuan untuk merepresentasikan bagaimana interaksi antara suatu sistem kepada pengguna yang telah dirancang dalam proses yang terstruktur. Pemodelan sistem umumnya dapat direpresentasikan melalui *Use Case Diagram*, *Activity Diagram*, *Sequence Diagram* serta *Flowchart*.

3.2.1. Use Case Diagram

Use Case Diagram adalah representasi visual yang menunjukkan hubungan antara sistem dan penggunanya. Diagram ini dalam penelitian ini dapat ditemukan pada Gambar 3.6. berikut:

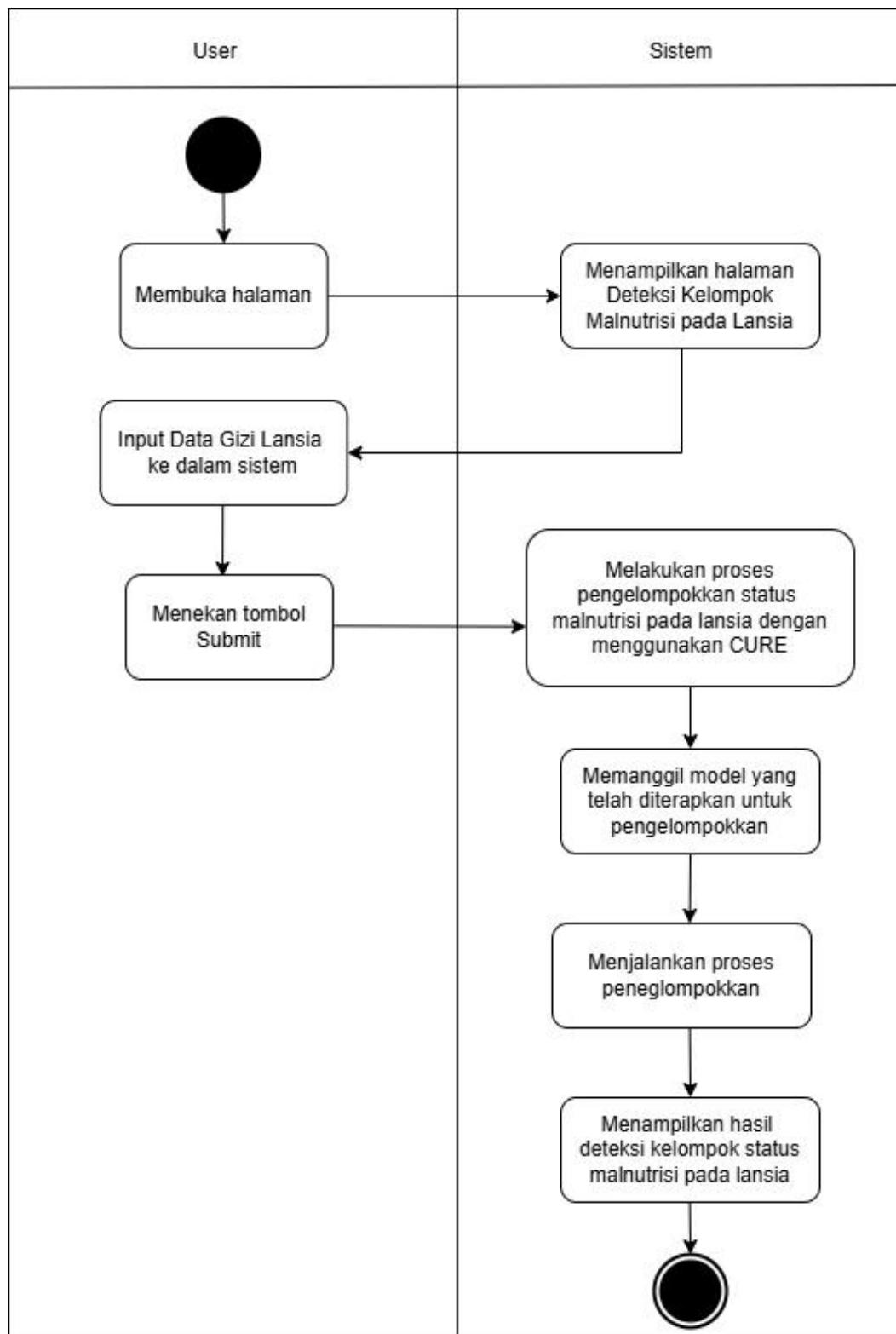


Gambar 3.6. Use Case Diagram

Berdasarkan Gambar 3.6., ditunjukkan bahwa untuk menggunakan sistem pengguna/user harus menginput data terkait gizi lansia terlebih dahulu, lalu akan ditampilkan hasil deteksi kelompok malnutrisi pada user tersebut.

3.2.2. Activity Diagram

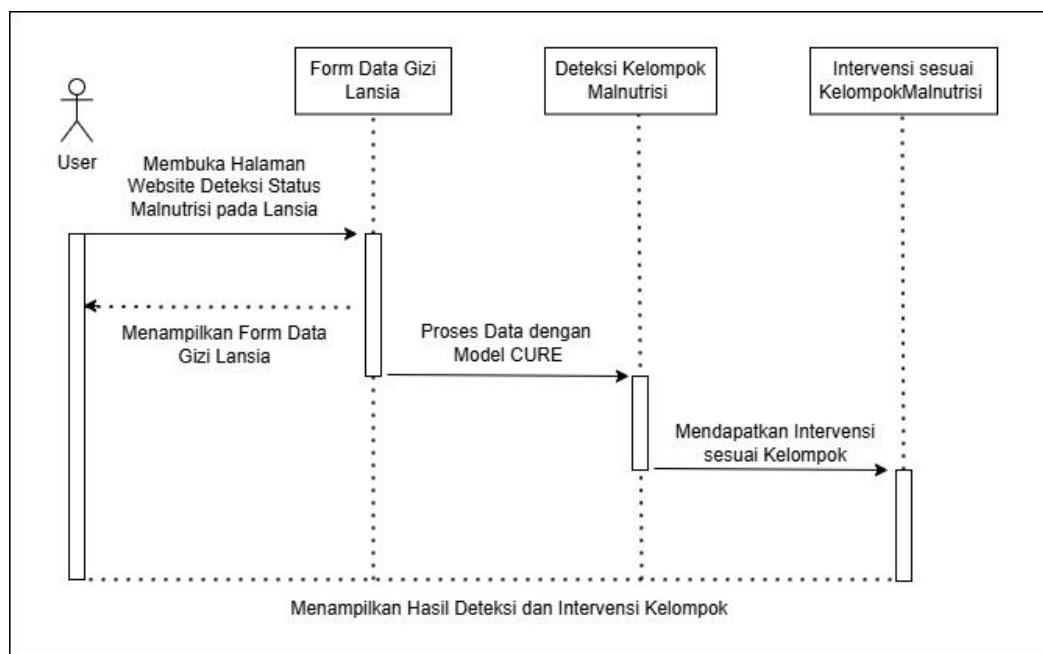
Activity Diagram merupakan sebuah diagram yang menunjukkan rancangan aliran proses kerja suatu sistem. Diagram aktivitas pada penelitian ini dapat ditemukan pada Gambar 3.7.



Gambar 3.7. Activity Diagram

3.2.3. Sequence Diagram

Sequence diagram adalah bagan yang menggambarkan bagaimana interaksi setiap objek yang ada di dalam sistem. *Sequence diagram* sistem untuk mendeteksi status malnutrisi pada lansia beserta intervensi sesuai kelompoknya terdapat pada gambar berikut.

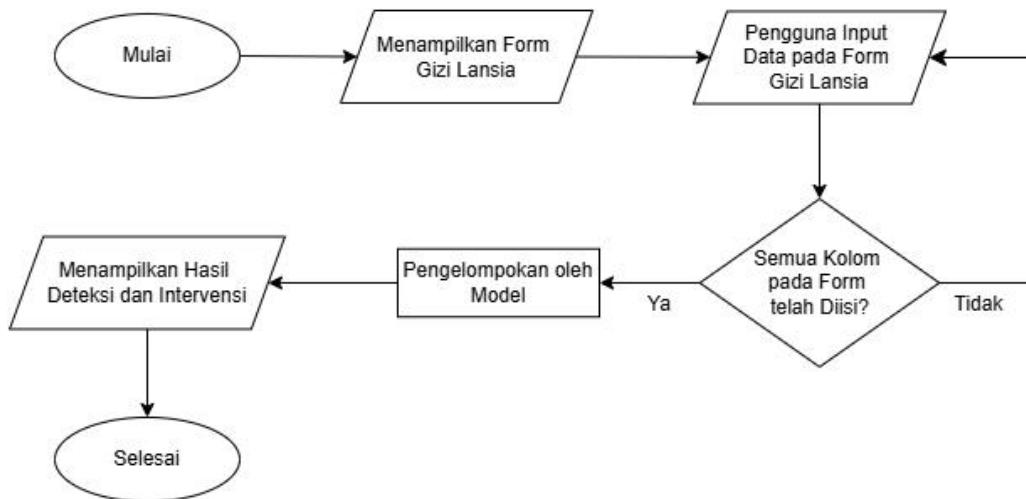


Gambar 3.8. Sequence Diagram

3.2.4. Flowchart

Flowchart adalah gambaran langkah – langkah yang menggambarkan proses dan solusi dari sebuah masalah atau program. Setiap tahap divisualisasikan dalam bentuk diagram yang dihubungkan oleh garis dan panah.

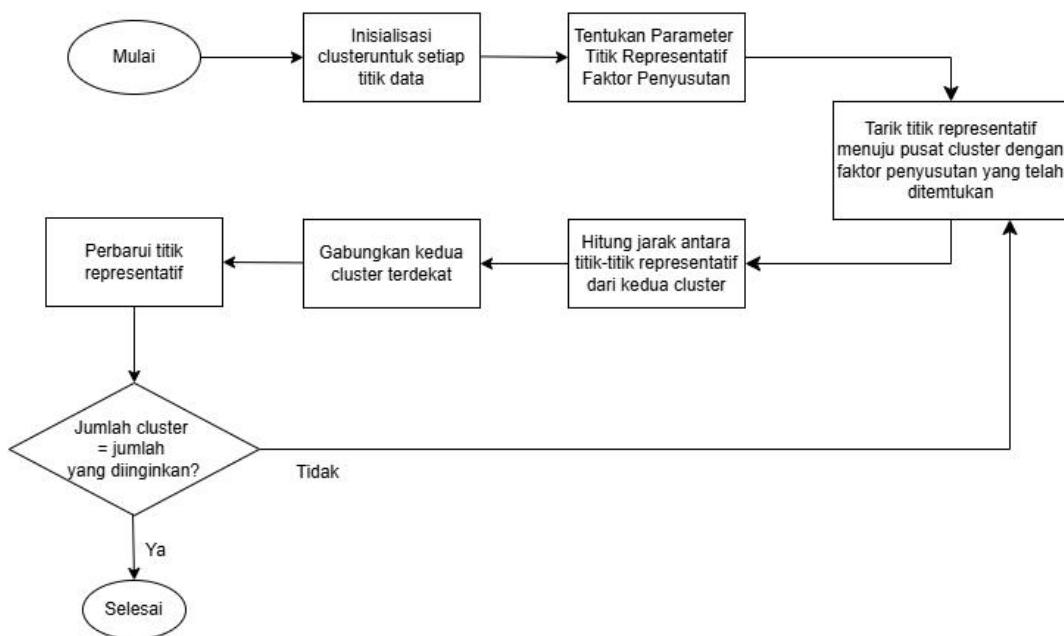
3.2.4.1. Flowchart Sistem



Gambar 3.9. Flowchart Sistem

Berdasarkan *flowchart* pada Gambar 3.9., pengguna/user menginputkan data gizi lansia. Data tersebut kemudian akan mengalami tahap pra-proses yang terdiri atas *encoding data*, pemilihan fitur dan standarisasi data. Data ini kemudian diolah menggunakan model *machine learning* yang akan mengelompokkan lansia ke dalam kategori. Sistem akan memastikan bahwa semua data telah diisi dengan benar sebelum melakukan analisis. Setelah proses analisis selesai, sistem akan memberikan hasil berupa kategori gizi lansia tersebut dan rekomendasi intervensi yang sesuai, seperti saran perubahan pola makan atau konsultasi dengan ahli gizi.

3.2.4.2. Flowchart Model CURE



Gambar 3.10. Flowchart CURE

Flowchart di atas menggambarkan langkah-langkah dalam algoritma CURE, sebuah metode *clustering* hierarkis yang efektif untuk menangani dataset besar dan bentuk *cluster* yang tidak konvensional.

Prosesnya dimulai dengan inisialisasi setiap data *point* sebagai satu *cluster* tunggal. Kemudian, algoritma menentukan beberapa parameter penting, yaitu jumlah *cluster* yang diinginkan, jumlah titik representatif per *cluster*, dan faktor penyusutan. Titik representatif ini akan mewakili *cluster* dan digunakan untuk menghitung jarak antar *cluster*.

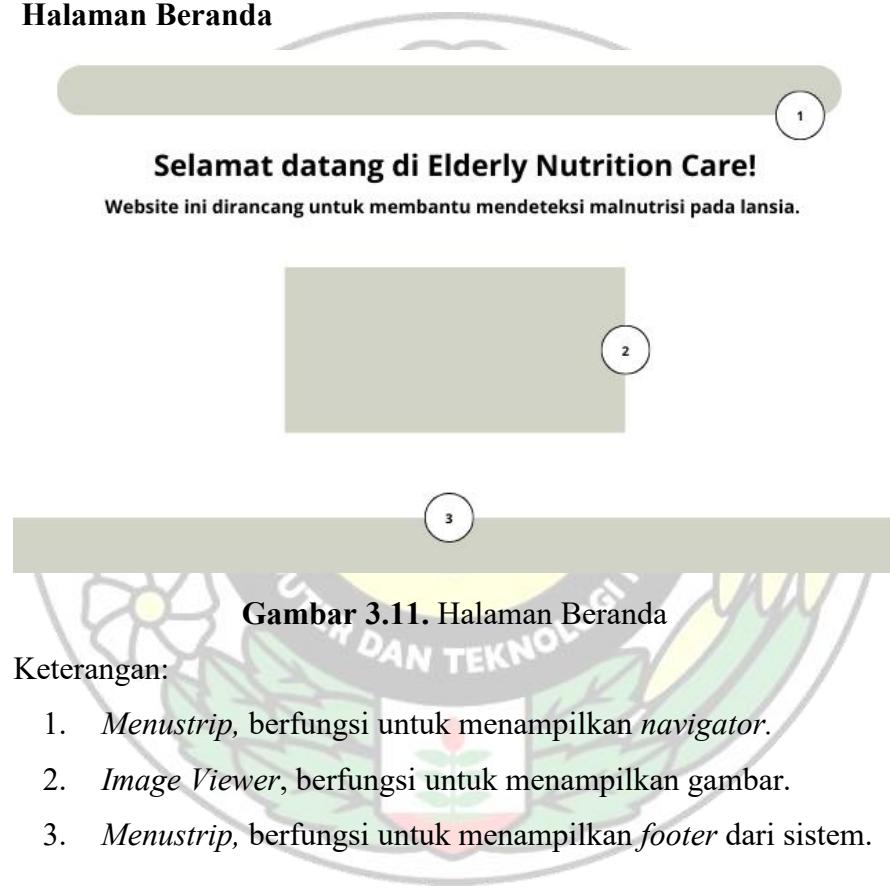
Selanjutnya, algoritma secara iteratif menggabungkan dua *cluster* terdekat. Jarak antara dua *cluster* dihitung berdasarkan jarak antara titik-titik representatifnya. Setelah penggabungan, titik-titik representatif dari *cluster* baru akan ditarik menuju pusat *cluster* dengan faktor penyusutan tertentu. Proses ini bertujuan untuk mengurangi pengaruh *outlier* dan menghasilkan *cluster* yang lebih kompak.

Iterasi penggabungan *cluster* akan terus berlanjut hingga jumlah *cluster* final tercapai. Pada setiap iterasi, jarak antara semua pasangan *cluster* akan diperbarui dan dua *cluster* terdekat akan digabungkan.

3.4. Perancangan *Interface*

Perancangan *interface* merupakan salah satu tahap penting pada saat melakukan pengembangan sistem. Hal ini dikarenakan *interface* merupakan penghubung antara sistem dan *user*, sehingga sebuah *interface* harus bersifat menarik dan juga mudah digunakan oleh *user*. Pada penelitian ini, peneliti memilih untuk menggunakan aplikasi berbasis *website* dikarenakan dapat diakses oleh lebih dari 1 perangkat serta *website* dapat diakses dimana saja tanpa perlu perangkat tambahan secara khusus.

3.4.1. Halaman Beranda

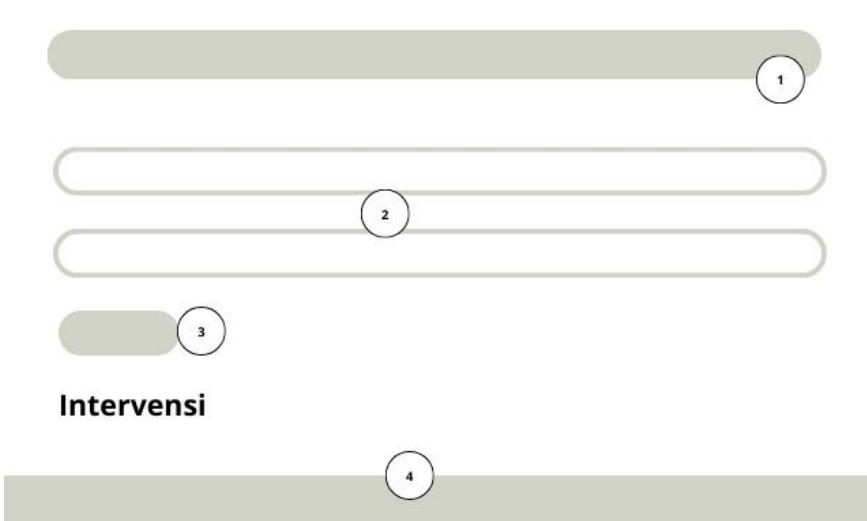
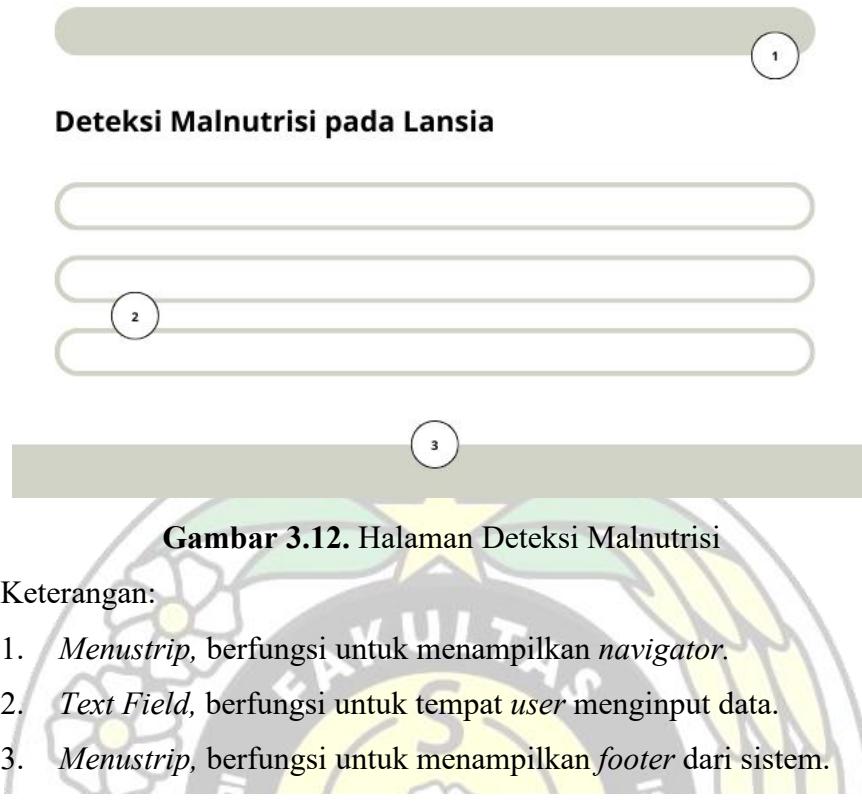


Gambar 3.11. Halaman Beranda

Keterangan:

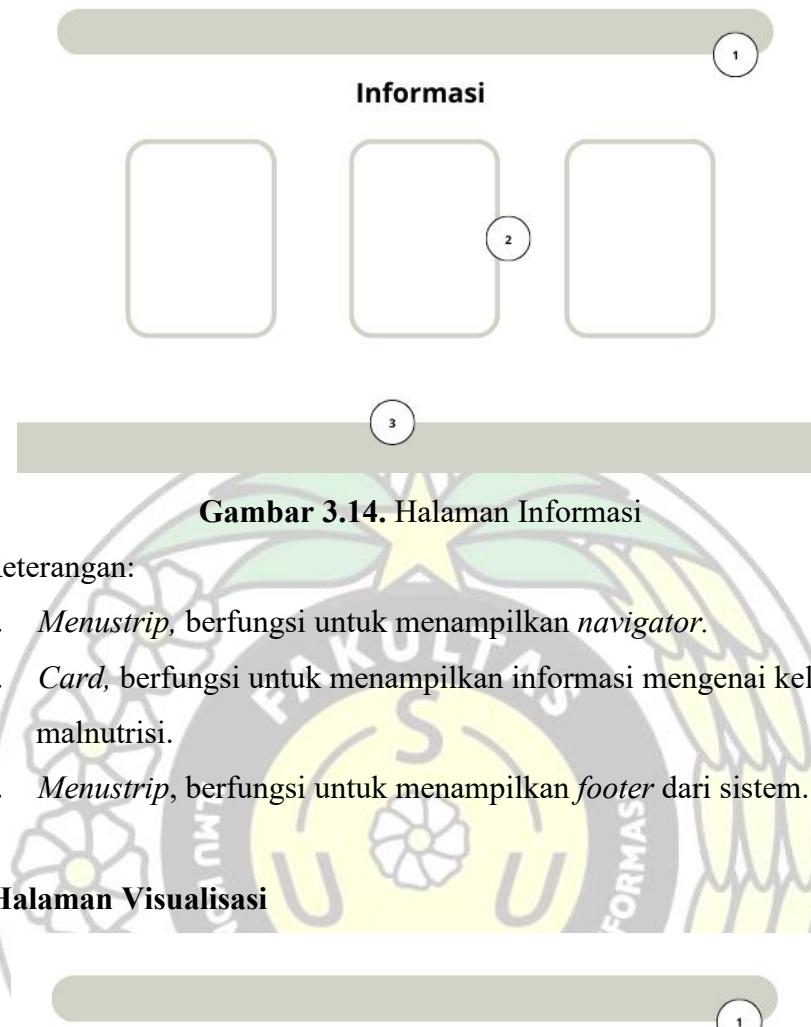
1. *Menustrip*, berfungsi untuk menampilkan *navigator*.
2. *Image Viewer*, berfungsi untuk menampilkan gambar.
3. *Menustrip*, berfungsi untuk menampilkan *footer* dari sistem.

3.4.2. Halaman Deteksi Malnutrisi dan Intervensi



1. *Menustrip*, berfungsi untuk menampilkan *navigator*.
2. *Text Field*, berfungsi untuk tempat *user* menginput data.
3. *Button*, berfungsi untuk mengirimkan inputan data ke model.
4. *Menustrip*, berfungsi untuk menampilkan *footer* dari sistem.

3.4.3. Halaman Informasi



Gambar 3.14. Halaman Informasi

Keterangan:

1. *Menustrip*, berfungsi untuk menampilkan *navigator*.
2. *Card*, berfungsi untuk menampilkan informasi mengenai kelompok malnutrisi.
3. *Menustrip*, berfungsi untuk menampilkan *footer* dari sistem.

3.4.4. Halaman Visualisasi



Gambar 3.15. Halaman Visualisasi

Keterangan:

1. *Menustrip*, berfungsi untuk menampilkan *navigator*.
2. *Image Viewer*, berfungsi untuk menampilkan gambar visualisasi.
3. *Menustrip*, berfungsi untuk menampilkan *footer* dari sistem.

BAB 4

IMPLEMENTASI DAN PENGUJIAN

4.1. Implementasi Sistem

Implementasi dari penelitian ini adalah membangun sistem deteksi dini status malnutrisi pada lansia menggunakan algoritma CURE untuk pengelompokan. Sistem ini dikembangkan menggunakan bahasa pemrograman *Python* dan berbasis website yang dibangun dengan *Framework Streamlit Python*.

4.1.1. Tampilan Halaman Beranda



Gambar 4.1. Tampilan Halaman Beranda

Keterangan:

- a. Judul: "Selamat datang di Elderly Nutrition Care!" - Judul ini menyampaikan tujuan utama website, yaitu memberikan perawatan nutrisi bagi lansia.
- b. Subjudul: "Website ini dirancang untuk membantu mendeteksi malnutrisi pada lansia" - Kalimat ini memperjelas fungsi utama website, yaitu sebagai alat deteksi awal malnutrisi pada lansia.
- c. Ilustrasi:
 - i. Tiga orang lansia: Menggambarkan target pengguna utama website ini, yaitu lansia.
 - ii. Makanan sehat: Menunjukkan fokus pada nutrisi yang baik untuk lansia.

- iii. Sistem: Mengindikasikan bahwa deteksi dini malnutrisi pada lansia dibuat ke dalam sistem berbasis *website*.
- d. Tahun dan keterangan: "2024 Deteksi Malnutrisi pada Lansia" menunjukkan tahun pembuatan *website* dan menegaskan kembali fokus utama *website*.
- e. Menu navigasi: Menu "Beranda", "Deteksi Malnutrisi", "Informasi", dan "Visualisasi" memberikan gambaran umum tentang fitur-fitur yang tersedia di *website*.

4.1.2. Tampilan Halaman Deteksi dan Intervensi

Pada halaman ini, pengguna akan menginputkan data yang berkaitan dengan gizi lansia yang akan dideteksi status malnutrisinya. Setelah hasil deteksi keluar, maka akan diberikan juga intervensi sesuai dengan kelompok status malnutrisi dari hasil deteksi.

The screenshot shows a user interface for a nutritional screening system. At the top, there is a navigation bar with five items: 'Beranda' (Home), 'Deteksi Malnutrisi' (selected and highlighted in red), 'Informasi', and 'Visualisasi'. On the far right of the navigation bar is a 'Deploy' button. Below the navigation bar, the main title is 'Deteksi Malnutrisi pada Lansia'. The form consists of several input fields:

- Nama:** A text input field containing placeholder text.
- Umur:** An input field with the value '0' and increment/decrement buttons.
- Jenis Kelamin:** A dropdown menu currently showing two options.
- Berat Badan (kg):** An input field with the value '0' and increment/decrement buttons.
- Tinggi Badan (cm):** An input field with the value '0' and increment/decrement buttons.

At the bottom of the form, there is a copyright notice: '© 2024 Deteksi Malnutrisi pada Lansia'.

Gambar 4.2. Tampilan Input Data *User* pada Halaman Deteksi

Klaster Cukup Gizi:

Intervensi:

- Pemeliharaan gizi dengan pola makan yang seimbang dan aktivitas fisik teratur.
- Konsultasi berkala dengan ahli gizi untuk memastikan asupan nutrisi tetap optimal.
- Pemantauan kondisi kesehatan secara teratur untuk mendeteksi perubahan status gizi atau kondisi fisik.

© 2024 Deteksi Malnutrisi pada Lansia

Gambar 4.3. Tampilan Hasil Deteksi dan Intervensi pada Halaman Deteksi

Berdasarkan pedoman World Health Organization (WHO), malnutrisi terbagi menjadi tiga kelompok utama:

Malnutrisi Ringan	Malnutrisi Sedang	Malnutrisi Berat
Intervensi:	Intervensi:	Intervensi:
<ul style="list-style-type: none"> • Pengaturan pola makan yang lebih seimbang dan bervariasi. • Kegiatan fisik ringan yang teratur. • Konseling psikologis ringan, jika diperlukan. • Pengawasan berkala terhadap status gizi. 	<ul style="list-style-type: none"> • Konsultasi ahli gizi untuk merancang rencana makan yang lebih kaya nutrisi. • Program rehabilitasi fisik yang ringan. • Pemberian suplemen gizi, terutama protein dan mikronutrien. • Pengawasan berkala oleh tenaga medis. 	<ul style="list-style-type: none"> • Pendekatan medis segera, dengan evaluasi menyeluruh oleh dokter dan ahli gizi. • Pemberian makanan atau nutrisi enteral/parenteral, jika diperlukan. • Terapi fisik untuk meningkatkan kekuatan fisik. • Konseling psikologis untuk mengurangi stres dan kecemasan.

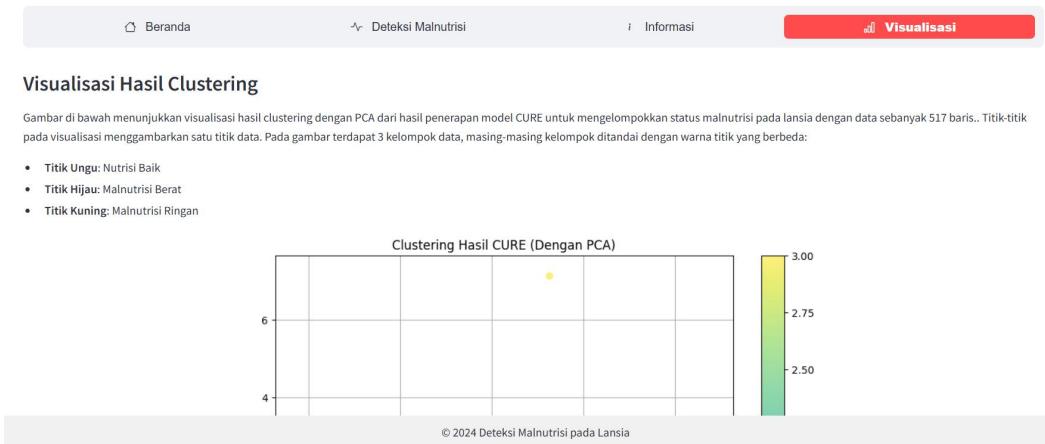
© 2024 Deteksi Malnutrisi pada Lansia

Gambar 4.4. Tampilan Halaman Informasi

Keterangan:

- Judul: "Informasi"
- Subjudul: "Berdasarkan pedoman World Health Organization (WHO), malnutrisi terbagi menjadi tiga kelompok utama" - Memberikan konteks mengenai asal-usul kelompok malnutrisi yang disajikan.
- Kartu Informasi: Informasi mengenai masing-masing tingkatan malnutrisi. Tiap bagian berisi:
 - Tiga Judul Tingkatan: Malnutrisi Ringan, Malnutrisi Sedang, dan Malnutrisi Berat.
 - Intervensi: Informasi mengenai tindakan yang perlu dilakukan untuk mengatasi malnutrisi pada masing-masing tingkatan.

4.1.4. Tampilan Halaman Visualisasi



Gambar 4.5. Tampilan Halaman Visualisasi

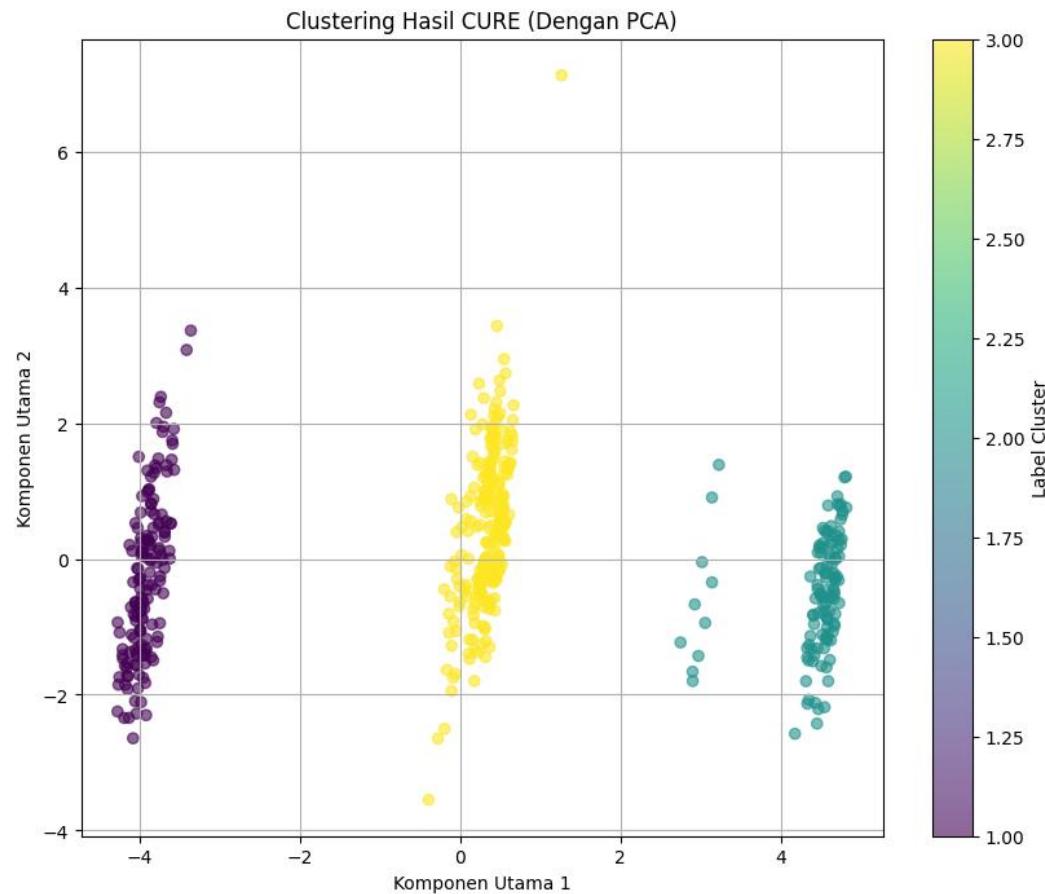
Halaman visualisasi ini menyajikan hasil dari analisis *clustering* data nutrisi lansia. Analisis ini bertujuan untuk mengelompokkan data lansia berdasarkan status gizinya.

Keterangan:

- Judul: "Visualisasi Hasil Clustering"
- Deskripsi Singkat: Dijelaskan bahwa visualisasi ini menggunakan metode PCA (*Principal Component Analysis*) untuk mengurangi dimensi data dan memudahkan visualisasi.
- Visualisasi:
 - Grafik Scatter Plot: Grafik ini menampilkan titik-titik data yang mewakili masing-masing individu.
 - Warna: Setiap kelompok data diberi warna yang berbeda untuk menunjukkan kelompok yang berbeda.
 - Label: Terdapat label yang menjelaskan masing-masing warna:
 - Ungu: Gizi Normal
 - Hijau: Malnutrisi Berat
 - Kuning: Malnutrisi Ringan
- Sumbu: Sumbu x dan y pada grafik mewakili dimensi baru yang dihasilkan dari proses PCA.

4.2. Pengujian Sistem

Implementasi *Clustering Using Representatives* (CURE) dalam pengelompokan status malnutrisi pada lansia menghasilkan visualisasi seperti yang ditunjukkan pada Gambar 4.6. Untuk memudahkan interpretasi, visualisasi ini dilakukan dengan melakukan reduksi dimensi menggunakan *Principal Component Analysis* (PCA).



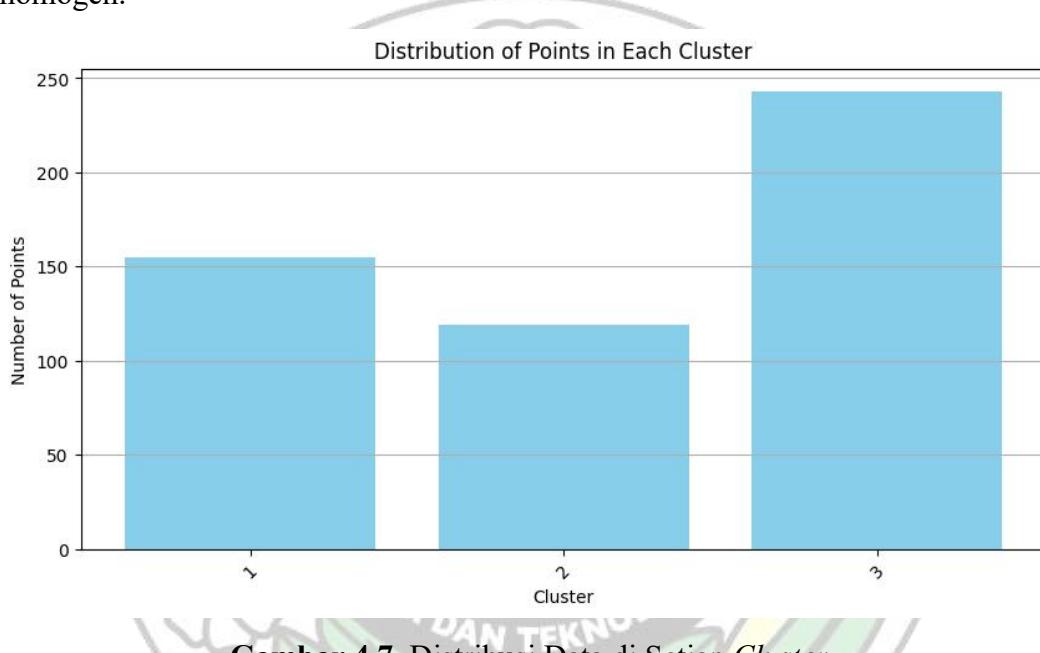
Gambar 4.6. Visualisasi *Clustering* dengan PCA

Data asli yang memiliki berbagai variabel (berat badan, tinggi badan, penurunan berat badan, asupan makanan, asupan protein, kemampuan mobilitas, kondisi kesehatan, dan fungsi kognitif), direduksi menjadi dua komponen utama. Komponen-komponen utama ini menangkap sebagian besar variabilitas dari data asli, memungkinkan visualisasi data dalam ruang dua dimensi tanpa kehilangan terlalu banyak informasi penting.

Pada grafik, setiap titik mewakili satu individu lansia, dan warna berbeda menunjukkan kelompok yang berbeda hasil dari proses *clustering*. Lansia yang

tergabung dalam *cluster* yang sama memiliki karakteristik yang mirip, yang ditandai dengan warna yang sama atau serupa dalam visualisasi. Semakin mirip warnanya, semakin dekat karakteristik data tersebut dalam *cluster* yang sama.

Proses ini menghasilkan beberapa *cluster* yang mengelompokkan lansia berdasarkan kesamaan karakteristik mereka, yang mungkin terkait dengan berbagai faktor seperti asupan nutrisi, status kesehatan umum, tingkat aktivitas fisik, dan faktor lain yang relevan. Dengan demikian, setiap *cluster* merepresentasikan kelompok lansia dengan karakteristik status gizi yang relatif homogen.



Gambar 4.7. Distribusi Data di Setiap Cluster

Gambar 4.7. menunjukkan distribusi jumlah data dalam masing-masing *cluster* hasil pengelompokan. Pada histogram, setiap bar mewakili satu *cluster*, dengan tinggi bar menunjukkan jumlah individu lansia yang tergabung dalam *cluster* tersebut.

Histogram ini memberikan gambaran mengenai seberapa besar setiap *cluster* relatif terhadap *cluster* lainnya. Dengan mengetahui distribusi data setiap *cluster*, maka ketiga *cluster* dapat diberikan label berdasarkan ciri-ciri yang menonjol dari data pada setiap kelompok.

Label-label ini dirancang untuk mencerminkan karakteristik utama yang mengelompokkan individu lansia tersebut, seperti tingkat asupan nutrisi, kondisi kesehatan tertentu, atau aspek lain yang relevan. Untuk memastikan validitas dan

relevansinya, label-label ini telah divalidasi oleh pakar gizi, sehingga setiap cluster dapat lebih mudah diidentifikasi dalam konteks status malnutrisi pada lansia..

Tabel 4.1. Ciri-ciri Data pada *Cluster*

Cluster	Ciri-ciri Utama Data pada Cluster	Label
1	Lansia dalam <i>cluster</i> ini memiliki berat badan yang stabil. Mereka makan tiga kali sehari dengan variasi makanan yang kaya nutrisi, dan umumnya tidak memiliki masalah dalam mengunyah atau mobilitas. <i>Cluster</i> ini mencakup lansia yang bebas dari masalah kesehatan kronis yang berat dan memiliki kondisi kognitif yang baik serta tingkat stres yang sangat rendah.	Gizi Normal
2	Pada <i>cluster</i> ini, umumnya lansia mengalami penurunan berat badan lebih dari 3 kg dalam 3 bulan terakhir. Asupan makan dan variasi makanan sangat terbatas, sering kali kurang dari sekali makan sehari. Lansia di <i>cluster</i> ini memiliki mobilitas yang sangat terbatas, sering kali bergantung penuh pada bantuan. Mereka juga mengalami kesulitan berat dalam mengunyah makanan dan umumnya memiliki penyakit kronis yang serius serta masalah kognitif yang signifikan.	Malnutrisi Berat
3	Lansia dalam cluster ini biasanya mengalami penurunan berat badan sekitar 1-3 kg dalam 3 bulan terakhir. Asupan makanan terbatas, dengan pola makan dua kali sehari dan variasi makanan yang kurang beragam. Mobilitas lansia pada <i>cluster</i> ini umumnya tidak memiliki masalah, namun ada juga beberapa yang membutuhkan alat bantu. Selain itu, lansia memiliki penyakit kronis, meskipun umumnya mereka tidak mengalami stres.	Malnutrisi Ringan

Tabel 4.2. Label Cluster

Cluster	Jumlah Data	Label
1	155	Gizi Normal
2	119	Malnutrisi Berat
3	243	Malnutrisi Ringan

Beranda
Deteksi Malnutrisi
Informasi
Visualisasi

Deteksi Malnutrisi pada Lansia

Nama: Junaidi

Umur: 90

Jenis Kelamin: Laki-laki

Berat Badan (kg): 70

Tinggi Badan (cm): 166

Penurunan Berat Badan: 0

Frekuensi Makan: Tiga kali atau lebih

Variasi Makanan: Makanan bervariasi (misalnya daging, sayuran, buah)

Asupan Protein: Dua kali sehari atau lebih

Mobilitas: Bergerak dengan bebas tanpa bantuan

Aktivitas Sehari-hari: Tidak ada kesulitan

Kesehatan Mulut: Tidak ada masalah

Stres: Tidak

Penyakit Kronis: Tidak

Masalah Kognitif: Tidak ada masalah daya ingat

Deteksi Kluster

Klaster Gizi Normal:

Intervensi:

- Pemeliharaan gizi dengan pola makan yang seimbang dan aktivitas fisik teratur.
- Konsultasi berkala dengan ahli gizi untuk memastikan asupan nutrisi tetap optimal.
- Pemantauan kondisi kesehatan secara teratur untuk mendeteksi perubahan status gizi atau kondisi fisik.

© 2024 Deteksi Malnutrisi pada Lansia

Gambar 4.8. Uji Deteksi Klaster Gizi Normal

Deteksi Malnutrisi pada Lansia

Nama: Puspita

Umur: 64

Jenis Kelamin: Perempuan

Berat Badan (kg): 48

Tinggi Badan (cm): 157

Penurunan Berat Badan: 4

Frekuensi Makan: Satu kali atau kurang

Variasi Makanan: Makanan sangat terbatas atau monoton

Asupan Protein: Kurang dari sekali sehari atau tidak pernah

Mobilitas: Tidak dapat bergerak tanpa bantuan orang lain

Aktivitas Sehari-hari: Kesulitan berat

Kesehatan Mulut: Kesulitan berat

Stres: Ya

Penyakit Kronis: Ya

Masalah Kognitif: Masalah daya ingat yang signifikan

Deteksi Kluster

Klaster Malnutrisi Berat:

Intervensi:

- Pendekatan medis yang intensif, termasuk evaluasi komprehensif oleh dokter, ahli gizi, dan tim medis terkait.
- Pemberian nutrisi melalui metode enteral atau parenteral sesuai kebutuhan, untuk memastikan asupan nutrisi yang tepat bagi pasien.
- Terapi fisik dan intervensi mobilitas yang bertahap guna membantu meningkatkan kekuatan fisik.
- Konseling psikologis untuk mengelola stres, depresi, atau kecemasan yang dapat memperburuk kondisi malnutrisi.

© 2024 Deteksi Malnutrisi pada Lansia

Gambar 4.9. Uji Deteksi Klaster Malnutrisi Berat

Pada tahap pengujian sistem, sistem yang telah dirancang akan diuji dan dianalisis untuk mengetahui apakah sistem dapat berjalan dengan baik. Pada pengujian ini, akan dilakukan uji coba terhadap 25 data uji yang dimana akan dilakukan perbandingan antara hasil diagnosis pakar dengan hasil deteksi sistem.

Tabel 4.3. Data Hasil Pengujian Clustering

Nama	Umur	Jenis Kelamin	Berat Badan	Tinggi Badan	Penurunan Berat Badan	Frekuensi Makan	Variasi Makanan	Asupan Protein	Mobilitas	Aktivitas Sehari hari	Kesehatan Mulut	Penyakit Kronis	Stres	Masalah Kognitif	Sistem	Pakar	Ket
														Gizi Normal			
Junaidi	90	L	70	166	0	Tiga kali atau lebih	Makanan bervariasi (misalnya daging, sayuran, buah)	Dua kali sehari atau lebih	Bergerak dengan bebas tanpa bantuan	Tidak ada kesulitan	Tidak ada masalah	Tidak	Tidak	Ada sedikit masalah daya ingat	Gizi Normal	Gizi Normal	Sesuai
Tumini	82	P	56	145	0	Tiga kali atau lebih	Makanan bervariasi (misalnya daging, sayuran, buah)	Dua kali sehari atau lebih	Bergerak dengan bebas tanpa bantuan	Tidak ada kesulitan	Tidak ada masalah	Tidak	Tidak	Ada sedikit masalah daya ingat	Gizi Normal	Gizi Normal	Sesuai
Yusraldi Sitompul	74	L	52	146	0	Tiga kali atau lebih	Makanan bervariasi (misalnya daging, sayuran, buah)	Dua kali sehari atau lebih	Bergerak dengan bebas tanpa bantuan	Tidak ada kesulitan	Tidak ada masalah	Tidak	Tidak	Tidak ada masalah daya ingat	Gizi Normal	Gizi Normal	Sesuai
Puspita Itayana	64	P	48	157	4	Satu kali atau kurang	Makanan sangat terbatas atau monoton	Kurang dari sekali sehari	Tidak dapat bergerak tanpa bantuan orang lain	Kesulitan berat	Kesulitan berat	Ya	Ya	Masalah daya ingat yang signifikan	Malnutrisi Berat	Malnutrisi Berat	Sesuai
Edwin Gunawan	66	L	65	170	5	Satu kali atau kurang	Makanan sangat terbatas atau monoton	Kurang dari sekali sehari	Tidak dapat bergerak tanpa bantuan orang lain	Kesulitan berat	Kesulitan berat	Ya	Ya	Masalah daya ingat yang signifikan	Malnutrisi Berat	Malnutrisi Berat	Sesuai

Nama	Umur	Jenis Kelamin	Berat Badan	Tinggi Badan	Pemurunan Berat Badan	Frekuensi Makan	Variasi Makanan	Asupan Protein	Mobilitas	Aktivitas Sehari hari	Kesehatan Mulut	Penyakit Kronis	Stres	Masalah Kognitif	Sistem	Pakar	Ket
															Ada sedikit masalah daya ingat	Malnutrisi Berat	Malnutrisi Berat
Ratna Maulida	67	P	73	154	4	Satu kali atau kurang	Makanan sangat terbatas atau monoton	Sekali sehari	Bergerak dengan bantuan	Kesulitan ringan	Ada kesulitan ringan	Ya	Tidak	Ada sedikit masalah daya ingat	Malnutrisi Berat	Malnutrisi Berat	Malnutrisi Sesuai
Ruzila	65	P	45	148	2	Dua kali	Makanan terbatas pada beberapa jenis saja	Dua kali sehari atau lebih	Bergerak dengan bebas tanpa bantuan	Tidak ada kesulitan	Ada kesulitan ringan	Ya	Tidak	Tidak ada masalah daya ingat	Malnutrisi Ringan	Malnutrisi Ringan	Malnutrisi Sesuai
Zaimah	64	P	53	161	1	Tiga kali atau lebih	Makanan bervariasi (misalnya daging, sayuran, buah)	Dua kali sehari atau lebih	Bergerak dengan bebas tanpa bantuan	Tidak ada kesulitan	Tidak ada masalah	Tidak	Tidak	Ada sedikit masalah daya ingat	Gizi Normal	Gizi Normal	Gizi Sesuai
Ida	65	P	80	154	2	Dua kali	Makanan terbatas pada beberapa jenis saja	Dua kali sehari atau lebih	Tidak dapat bergerak tanpa bantuan orang lain	Kesulitan berat	Tidak ada masalah	Tidak	Ya	Ada sedikit masalah daya ingat	Malnutrisi Ringan	Malnutrisi Ringan	Malnutrisi Sesuai
Siti Aminah	69	P	55	154	3	Tiga kali atau lebih	Makanan terbatas pada beberapa jenis saja	Sekali sehari	Bergerak dengan bantuan	Kesulitan ringan	Ada kesulitan ringan	Ya	Tidak	Ada sedikit masalah daya ingat	Malnutrisi Ringan	Malnutrisi Ringan	Malnutrisi Sesuai
Sumiyati	70	P	50	150	0	Tiga kali atau lebih	Makanan bervariasi (misalnya daging, sayuran, buah)	Dua kali sehari atau lebih	Bergerak dengan bebas tanpa bantuan	Tidak ada kesulitan	Ada kesulitan ringan	Tidak	Tidak	Ada sedikit masalah daya ingat	Gizi Normal	Gizi Normal	Gizi Sesuai

Nama	Umur	Jenis Kelamin	Berat Badan	Tinggi Berat Badan	Pemurungan Berat Badan	Frekuensi Makan	Variasi Makanan	Asupan Protein	Mobilitas	Aktivitas Sehari hari	Kesehatan Mulut	Penyakit Kronis	Stres	Masalah Kognitif	Sistem	Pakar	Ket
Suryati	68	P	72	152	6	Satu kali atau kurang	Makanan sangat terbatas atau monoton	Sekali sehari	Bergerak dengan bantuan (misalnya tongkat, walker)	Tidak ada masalah berat	Ya	Ya	Tidak ada masalah daya ingat	Malnutrisi Berat	Malnutrisi Berat	Sesuai	
Hj Sukesih S.Ag	72	P	83	166	5	Tiga kali atau lebih	Makanan bervariasi (misalnya daging, sayuran, buah)	Dua kali sehari atau lebih	Bergerak dengan bebas tanpa bantuan	Tidak ada kesulitan	Tidak ada masalah	Ya	Tidak ada masalah daya ingat	Gizi Normal	Gizi Normal	Sesuai	
Dahrul Nasution	72	L	62	165	0	Tiga kali atau lebih	Makanan terbatas pada beberapa jenis saja	Sekali sehari	Bergerak dengan bebas tanpa bantuan	Ada kesulitan ringan	Tidak ada kesulitan	Tidak	Ada sedikit masalah daya ingat	Gizi Normal	Malnutrisi Ringan	Tidak Sesuai	
Surya	60	L	105	190	0	Tiga kali atau lebih	Makanan bervariasi (misalnya daging, sayuran, buah)	Dua kali sehari atau lebih	Bergerak dengan bebas tanpa bantuan	Kesulitan ringan	Ada kesulitan ringan	Tidak	Ada sedikit masalah daya ingat	Gizi Normal	Gizi Normal	Sesuai	
Sari Narulita	74	P	56	150	0	Dua kali	Makanan bervariasi (misalnya daging, sayuran, buah)	Sekali sehari	Bergerak dengan bebas tanpa bantuan	Tidak ada kesulitan	Tidak ada masalah	Tidak	Ada sedikit masalah daya ingat	Gizi Normal	Gizi Normal	Sesuai	

Nama	Umur	Jenis Kelamin	Berat Badan	Tinggi Berat Badan	Pemurunan Berat Badan	Frekuensi Makan	Variasi Makanan	Asupan Protein	Mobilitas	Aktivitas Sehari hari	Kesehatan Mulut	Penyakit Kronis	Stres	Masalah Kognitif	Sistem	Pakar	Ket		
															Makanan bervariasi (misalnya daging, sayuran, buah)	Dua kali sehari atau lebih	Bergerak dengan bantuan (misalnya tongkat, walker)	Kesulitan ringan	Ada kesulitan ringan
Nuraz-mah	91	P	45	153	0	Tiga kali atau lebih	Makanan bervariasi (misalnya daging, sayuran, buah)	Dua kali sehari atau lebih	Bergerak dengan bantuan (misalnya tongkat, walker)	Kesulitan ringan	Ada kesulitan ringan	Tidak	Tidak	Masalah daya ingat yang signifikan	Malnutrisi Ringan	Malnutrisi Ringan	Gizi Normal	Gizi Normal	Sesuai
Bayu Mulyono	67	L	60	166	0	Tiga kali atau lebih	Makanan terbatas pada beberapa jenis saja	Sekali sehari	Bergerak dengan bebas tanpa bantuan	Kesulitan ringan	Tidak ada masalah	Tidak	Tidak	Ada sedikit masalah daya ingat	Ada sedikit masalah daya ingat	Ada sedikit masalah daya ingat	Gizi Normal	Gizi Normal	Sesuai
Luse Sihombing	75	P	90	145	0	Tiga kali atau lebih	Makanan bervariasi (misalnya daging, sayuran, buah)	Dua kali sehari atau lebih	Bergerak dengan bebas tanpa bantuan	Tidak ada masalah	Ya	Ya	Tidak ada masalah	Ada sedikit masalah daya ingat	Ada sedikit masalah daya ingat	Gizi Normal	Gizi Normal	Sesuai	
Mangara Tua Sihaaan	61	L	65	163	0	Tiga kali atau lebih	Makanan bervariasi (misalnya daging, sayuran, buah)	Dua kali sehari atau lebih	Bergerak dengan bebas tanpa bantuan	Tidak ada masalah	Tidak	Tidak	Tidak ada masalah daya ingat	Tidak ada masalah daya ingat	Tidak ada masalah daya ingat	Gizi Normal	Gizi Normal	Sesuai	
Michael Sihite	58	L	65	169	0	Tiga kali atau lebih	Makanan bervariasi (misalnya daging, sayuran, buah)	Dua kali sehari atau lebih	Bergerak dengan bebas tanpa bantuan	Tidak ada masalah	Tidak	Tidak	Tidak ada masalah daya ingat	Tidak ada masalah daya ingat	Tidak ada masalah daya ingat	Gizi Normal	Gizi Normal	Sesuai	

Nama	Umur	Jenis Kelamin	Berat Badan	Tinggi Berat Badan	Pemurunan Berat Badan	Frekuensi Makan	Variasi Makanan	Asupan Protein	Mobilitas	Aktivitas Sehari hari	Kesehatan Mulut	Penyakit Kronis	Stres	Masalah Kognitif	Sistem	Pakar	Ket
Mariska	71	P	64	160	5	Tiga kali atau lebih	Makanan bervariasi (misalnya daging, sayuran, buah)	Sekali sehari	Bergerak dengan bantuan (misalnya tongkat, walker)	Tidak ada kesulitan	Ada kesulitan ringan	Tidak	Tidak	Ada sedikit masalah daya ingat	Malnutrisi Ringan	Malnutrisi Ringan	Sesuai
Amina	85	P	45	165	2	Tiga kali atau lebih	Makanan terbatas pada beberapa jenis saja	Dua kali sehari atau lebih	Tidak dapat bergerak tanpa bantuan orang lain	Kesulitan berat dalam mengunyah atau makan	Kesulitan berat dalam mengunyah atau makan	Tidak	Tidak	Ada sedikit masalah daya ingat	Malnutrisi Ringan	Malnutrisi Berat	Tidak Sesuai
Mariani	65	P	68	160	3	Tiga kali atau lebih	Makanan bervariasi (misalnya daging, sayuran, buah)	Sekali sehari	Bergerak dengan bebas tanpa bantuan	Tidak ada kesulitan	Tidak ada masalah	Tidak	Tidak	Ada sedikit masalah daya ingat	Gizi Normal	Gizi Normal	Sesuai
willi	65	L	60	180	1	Dua kali	Makanan bervariasi (misalnya daging, sayuran, buah)	Sekali sehari	Bergerak dengan bebas tanpa bantuan	Tidak ada masalah	Tidak	Ya	Ada sedikit masalah daya ingat	Gizi Normal	Malnutrisi Ringan	Tidak Sesuai	

Berdasarkan pengujian terhadap 25 data uji yang tercantum dalam Tabel 4.3, tingkat akurasi sistem dapat dihitung dengan membandingkan hasil deteksi yang dihasilkan oleh sistem dengan diagnosis yang diberikan oleh pakar.

$$\text{Akurasi} = \frac{\text{Banyak data uji yang sesuai}}{\text{Total data uji}} \times 100\%$$

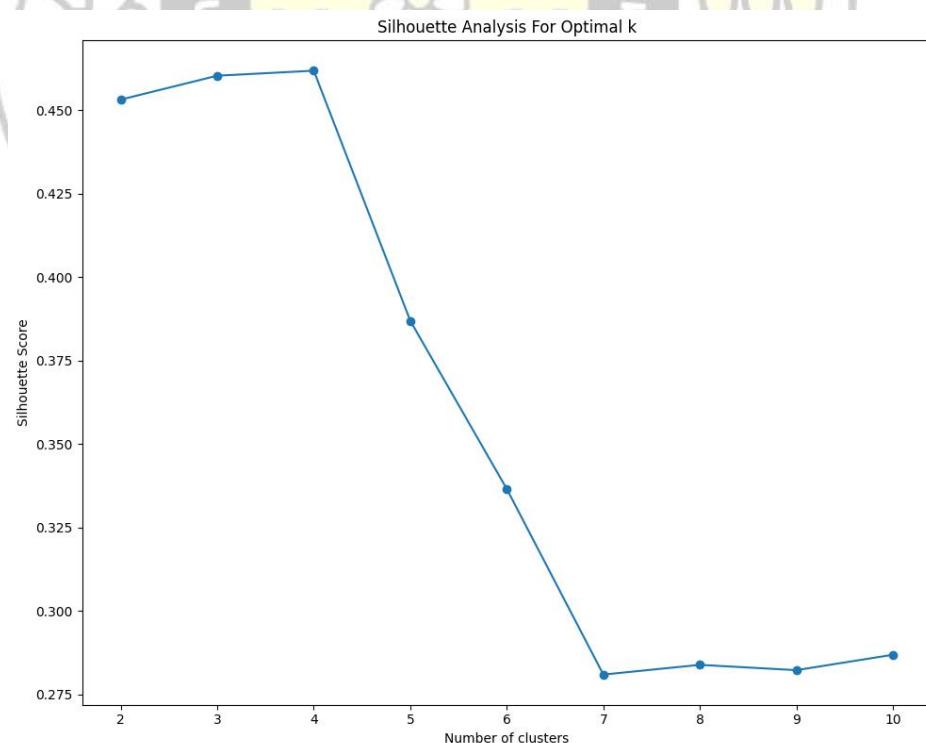
$$\text{Akurasi} = \frac{22}{25} \times 100\% = 88\%$$

Dari hasil perhitungan akurasi dari pengujian sistem deteksi pengelompokan ini, diperoleh tingkat akurasi sebesar 88%. Hal ini menunjukkan bahwa sistem deteksi untuk pengelompokan status malnutrisi pada lansia telah berfungsi dengan baik.

4.3. Hasil Pengujian

4.3.1. Pengujian Jumlah *Cluster* Optimal

Pada penelitian ini, pengujian jumlah *cluster* optimal dilakukan untuk menentukan jumlah *cluster* terbaik dalam mengelompokkan status malnutrisi pada lansia berdasarkan dataset yang digunakan. Proses ini menggunakan metode *Silhouette Analysis*.



Gambar 4.10. Implementasi Metode *Silhouette Analysis*

Hasil grafik menunjukkan bahwa titik tertinggi terdapat pada

angka 3 dan 4. Berdasarkan analisis dataset yang digunakan dalam penelitian ini, cluster optimal yang terbentuk adalah 3 cluster. Oleh karena itu, pengelompokan status malnutrisi pada lansia, yang didasarkan pada dataset yang telah melalui tahap *pre-processing*, menghasilkan 3 cluster yang meliputi cluster dengan status gizi normal, cluster malnutrisi ringan, dan cluster malnutrisi berat.

4.3.2. Pengujian Kualitas *Cluster*

Dalam penelitian ini, *Davies-Bouldin Index (DBI)* digunakan sebagai metrik evaluasi untuk menilai kualitas hasil *clustering* pada pengelompokan data status malnutrisi lansia. DBI merupakan metrik yang mengukur kualitas *clustering* dengan mengevaluasi seberapa kompak setiap *cluster* dan seberapa jauh jarak antar *cluster*. Semakin rendah nilai DBI, semakin baik hasil *clustering*, karena ini menunjukkan bahwa setiap *cluster* memiliki sebaran yang kecil dan terpisah secara jelas dari *cluster* lain. Rumus DBI dapat dilihat pada rumus 2.7.

Dengan menggunakan *Davies-Bouldin Index (DBI)*, diharapkan dapat diperoleh nilai evaluasi yang menunjukkan tingkat homogenitas dan separasi antar *cluster*, sehingga kualitas pengelompokan dapat diukur dengan baik.

$$\begin{aligned}s_1 &= 1.978, s_2 = 1.879, s_3 = 1.874 \\d_{12} &= 4.525, d_{13} = 8.367, d_{23} = 4.507 \\R_{12} &= 0.8525, R_{13} = 0.4605, R_{23} = 0.8329\end{aligned}$$

$$\text{Klaster 1: } \max(R_{12}, R_{13}) = \max(0.8525, 0.4605) = 0.8525$$

$$\text{Klaster 2: } \max(R_{12}, R_{23}) = \max(0.8525, 0.8329) = 0.8525$$

$$\text{Klaster 3: } \max(R_{13}, R_{23}) = \max(0.4605, 0.8329) = 0.8329$$

$$DBI = \frac{1}{3}(0.8525 + 0.8525 + 0.8329)$$

$$DBI = \frac{1}{3}(2.5379) \approx 0.8459$$

Pada penelitian ini, penggunaan DBI menghasilkan nilai evaluasi sebesar 0.85. Hasil DBI yang memiliki nilai di bawah 1 menunjukkan bahwa *cluster* yang terbentuk memiliki tingkat separasi dan homogenitas yang relatif baik. Semakin rendah nilai DBI, semakin baik kualitas pengelompokan karena *cluster* yang terbentuk lebih kompak dan memiliki jarak yang signifikan dari *cluster* lainnya.

Secara keseluruhan, nilai 0.85 pada DBI menunjukkan bahwa metode pengelompokan telah menghasilkan pemisahan *cluster* yang cukup optimal dalam mengelompokkan status malnutrisi pada lansia.



BAB 5

PENUTUP

5.1. Kesimpulan

Berdasarkan temuan dari penelitian yang telah dilakukan, kesimpulan yang dapat diambil adalah sebagai berikut:

1. Algoritma CURE yang diterapkan pada 517 data lansia berhasil menghasilkan pengelompokan dan deteksi status malnutrisi dengan nilai *Davies-Bouldin Index* (DBI) sebesar 0.85.
2. Berdasarkan hasil analisis menggunakan *Silhouette Analysis*, dataset yang digunakan menunjukkan bahwa pembentukan *cluster* optimal adalah sebanyak 3 *cluster*. *Cluster* ini menggambarkan kelompok lansia berdasarkan status gizinya, yaitu *cluster* dengan status gizi normal, *cluster* malnutrisi ringan, dan *cluster* malnutrisi berat.
3. Hasil perbandingan antara deteksi sistem dan diagnosis pakar gizi dengan menggunakan 25 data uji menunjukkan bahwa akurasi sistem mencapai 88%.

5.2. Saran

Saran yang dapat disampaikan oleh penulis untuk penelitian berikutnya yaitu diperlukannya pengembangan metode pengelompokan agar kualitas *clustering* dapat meningkat. Hal ini dapat dilakukan dengan menambah jumlah data, sehingga karakteristik tiap kelompok malnutrisi pada lansia lebih teridentifikasi secara optimal.

DAFTAR PUSTAKA

- Babu, B. H., Chandra, N. S., & Gopal, T. V. (2017). CURE: High dimensional data clustering using clustering using representatives. *Complexity International Journal*, 21(1), 151.
- Dewiasty, E., Agustina, R., Saldi, S. R. F., Pramudita, A., Hinssen, F., Kumaheri, M., de Groot, L. C. P. G. M., & Setiati, S. (2022). Malnutrition prevalence and nutrient intakes of Indonesian community-dwelling older adults: A systematic review of observational studies. *Frontiers in Nutrition*, 9, 780003.
- Donini, L. M., Stephan, B. C. M., Rosano, A., Molfino, A., Poggioigalle, E., Lenzi, A., Siervo, M., & Muscaritoli, M. (2020). What are the risk factors for malnutrition in older-aged institutionalized adults? *Nutrients*, 12(9), 2862.
- Guigoz, Y., & Vellas, B. (2021). Nutritional assessment in older adults: MNA® 25 years of a screening tool & a reference standard for care and research; what next? *Journal of Nutrition, Health & Aging*, 25(4), 528–583.
- Kang, M., Choi, J. Y., Yoo, H. J., Park, S. Y., Kim, Y., Kim, J. Y., Kim, S., Kim, C. H., & Kim, K. I. (2023). Impact of malnutrition evaluated by the mini nutritional assessment on the prognosis of acute hospitalized older adults. *Frontiers in Nutrition*
- Kiesswetter, E., Colombo, M. G., Meisinger, C., Peters, A., Thorand, B., Holle, R., Ladwig, K.-H., Schulz, H., Grill, E., Diekmann, R., Schrader, E., Stehle, P., Sieber, C. C., & Volkert, D. (2019). Malnutrition and related risk factors in older adults from different health-care settings: An enable study. *Public Health Nutrition*, 23(3), 446–456.
- Kostecka, M., & Bojanowska, M. (2021). An evaluation of the nutritional status of elderly with the use of the MNA questionnaire and determination of factors contributing to malnutrition: A pilot study. *Roczniki Państwowej Szkoły Wyższej w Gdyni*, 72(2), 175–183.
- Kumble, N., & Tewari, V. (2020). CURE clustering algorithm and its various implementations. *International Journal For Innovative Research In Multidisciplinary Field*, 6(6).

- Ma, L., & Fan, S. (2017). CURE-SMOTE algorithm and hybrid algorithm for feature selection and parameter optimization based on random forests. *BMC Bioinformatics*, 18, 169.
- Medina-Vadora, M. M., Plaza-Diaz, J., Llorente-Cantarero, F. J., Severi, C., Lecot, C., Ruiz-López, M. D., & Gil, Á. (2024). A clustering study of dietary patterns and physical activity among workers of the Uruguayan State Electrical Company. *Nutrients*, 16(2), 304.
- Moloney, L., & Jarrett, B. (2021). Nutrition assessment and interventions for the prevention and treatment of malnutrition in older adults: An evidence analysis center scoping review. *Journal of the Academy of Nutrition and Dietetics*, 121(10), 2108-2140.e6.
- Nagy, D., Aszalós, L., & Mihálydeák, T. (2019). Finding the representative in a cluster using correlation clustering. *Pollack Periodica: An International Journal for Engineering and Information Sciences*, 14(1), 15–24.
- Norman, K., Haß, U., Pirlich, M., et al. (2021). Malnutrition in older adults: Recent advances and remaining challenges. *Nutrients*, 13(8), 2764.
- Piccoli, G. B., Cederholm, T., Avesani, C. M., Bakker, S. J. L., Bellizzi, V., Cuerda, C., Cupisti, A., Sabatino, A., Schneider, S., Torreggiani, M., Fouque, D., Carrero, J. J., & Barazzoni, R. (2023). Nutritional status and the risk of malnutrition in older adults with chronic kidney disease – Implications for low protein intake and nutritional care: A critical review endorsed by ERN-ERA and ESPEN. *Clinical Nutrition*, 42(4), 443–457.
- Putri Nst, U. R., Sutarman, & Sirait, P. (2020). Analysis of representative values in clustering using the CURE algorithm. *International Journal of Recent Technology and Engineering (IJRTE)*, 9(2). ISSN: 2277-3878.
- Salman, M. (2023). A novel clustering method with consistent data in a three-dimensional graphical format over existing clustering mechanisms. *Information Sciences*, 649.
- Sharma, V., Sharma, V., Khan, A., Wassmer, D. J., Schoenholtz, M. D., Hontecillas, R., Bassaganya-Riera, J., Zand, R., & Abedi, V. (2020). Malnutrition, health, and the role of machine learning in clinical settings. *Frontiers in Nutrition*, 7, Article 44.

- Shuremu, M., Belachew, T., & Hassen, K. (2023). Nutritional status and its associated factors among elderly people in Ilu Aba Bor Zone, Southwest Ethiopia: A community-based cross-sectional study. *BMJ Open*, 13(1), e067787.
- Susetyowati, S., Winarti, H., Roselani, A., Handayani, S., Sanubari, N. D. G., & Sholikhati, A. S. (2024). Utilization of Mini Nutrition Assessment-Short Form to identify nutritional status of hospitalized elderly patients at Dr Sardjito General Hospital, Yogyakarta. *Amerta Nutrition*, 8(2), 263–268.
- Tomasiewicz, A., Polański, J., & Tański, W. (2023). Advancing the understanding of malnutrition in the elderly population: Current insights and future directions. *Nutrients*, 16(15), Article 2502.

