E-ISSN 3032-601X & P-ISSN 3032-7105
Vol. 3, No. 1, 2026

MISTER

Journal of Multidisciplinary Inquiry in Science, Technology and Educational Research

Jurnal Penelitian Multidisiplin dalam Ilmu Pengetahuan, Teknologi dan Pendidikan

UNIVERSITAS SERAMBI MEKKAH
KOTA BANDA ACEH
mister@serambimekkah.ac.id

Journal of Multidisciplinary Inquiry in Science Technology and Educational Research

Journal of MISTER

Vol. 3, No. 1, 2026

Pages: 859-867

Klasterisasi Individu Berdasarkan Pola Konsumsi Makanan dan Aktivitas Fisik untuk Penentuan Profil Risiko Obesitas Menggunakan Algoritma K-Means

Nanda Irioni Romadhon, Hasbi Firmansyah, Wahyu Asriyani, Rizki Prasetyo Tulodo

Universitas Pancasakti Tegal

Article in Journal of MISTER

Available at	:	https://jurnal.serambimekkah.ac.id/index.php/mister/index
DOI	:	https://doi.org/10.32672/mister.v3i1. 4002

How to Cite this Article

APA	:	Irioni Romadhon, N., Firmansyah, H., Asriyani, W., & Prasetyo Tulodo, R. (2025). Klasterisasi Individu Berdasarkan Pola Konsumsi Makanan dan Aktivitas Fisik untuk Penentuan Profil Risiko Obesitas Menggunakan Algoritma K-Means. Journal of Multidisciplinary Inquiry in Science, Technology and Educational Research, 3(1), 859 - 867. https://doi.org/10.32672/mister.v3i1.4002
Others Visit	:	https://jurnal.serambimekkah.ac.id/index.php/mister/index

ISSN 3032-7105 ISSN 3032-601X

Journal of Multidisciplinary Inquiry in Science, Technology and Educational Research
UNIVERSITAS SERAMBI MEKKAH
https://jurnal.serambimekkah.ac.id/index.php/mister

e-ISSN3032-601X & p-ISSN3032-7105
Vol. 3 No. 1, Tahun 2026
Doi: 10.32672/mister.v3i1.4002
Hal. 859-867

Klasterisasi Individu Berdasarkan Pola Konsumsi Makanan dan Aktivitas Fisik untuk Penentuan Profil Risiko Obesitas Menggunakan Algoritma K-Means

Nanda Irioni Romadhon[1*], Hasbi Firmansyah [2], Wahyu Asriyani [3], Rizki Prasetyo Tulodo [4]
Program Studi Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Pancasakti Tegal, Tegal, Indonesia[1,2,4]
Program Studi Pendidikan Bahasa dan Sastra Indonesia, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Pancasakti Tegal, Tegal, Indonesia[3]

[*]Email: nandairioni@gmail.com, hasbifirmansyah@upstegal.ac.id, asriyani1409@gmail.com, Rizki.prasetyo.tulodo@gmail.com

Diterima: 08-12-2025 | Disetujui: 18-12-2025 | Diterbitkan: 20-12-2025

ABSTRACT

This study aims to identify patterns of obesity risk based on eating habits and physical condition by applying the K-Means clustering method to the Estimation of Obesity Levels Based on Eating Habits and Physical Condition dataset obtained from the UCI Machine Learning Repository. The research is motivated by the increasing prevalence of obesity and the need for analytical approaches to systematically map its risk factors. The research stages include data preprocessing, BMI attribute construction, feature selection, and categorical data encoding using one-hot encoding and label encoding. The optimal number of clusters was determined using the elbow method prior to implementing the K-Means algorithm. The results indicate that k = 3 provides the best clustering quality, forming three clusters that represent low, moderate, and high obesity risk levels. The resulting clusters show clear differences in dietary patterns and physical activity intensity. These findings demonstrate the effectiveness of clustering analysis in obesity risk mapping and highlight its potential contribution to the development of preventive health decision support systems.
***Keywords**: Clustering, Eating Habits, K-Means, Obesity, Physical Activity.*

ABSTRAK

Penelitian ini bertujuan mengungkap pola risiko obesitas berdasarkan kebiasaan makan dan kondisi fisik dengan menerapkan metode *K-Means clustering* pada dataset *Estimation of Obesity Levels Based on Eating Habits and Physical Condition* dari UCI Machine Learning Repository. Latar belakang penelitian didorong oleh meningkatnya kasus obesitas serta kebutuhan akan pendekatan analitik untuk mengidentifikasi faktor-faktor risiko secara sistematis. Proses penelitian meliputi tahap *data preprocessing*, pembuatan atribut BMI, seleksi atribut, serta pengkodean data kategorikal menggunakan *one-hot encoding* dan *label encoding*. Penentuan jumlah klaster optimal dilakukan dengan metode *elbow* sebelum penerapan algoritma K-Means. Hasil analisis menunjukkan bahwa nilai k = 3 menghasilkan kualitas klaster terbaik, yang merepresentasikan kelompok risiko obesitas rendah, sedang, dan tinggi. Klaster yang terbentuk menunjukkan perbedaan yang jelas pada pola konsumsi makanan dan tingkat aktivitas fisik. Temuan ini menegaskan efektivitas analisis klaster dalam pemetaan risiko obesitas berbasis perilaku dan potensinya untuk mendukung pengembangan sistem pencegahan kesehatan.
Katakunci: Aktivitas Fisik, Kebiasaan Makan, *K-Means*, Klasterisasi, Obesitas.

e-ISSN3032-601X & p-ISSN3032-7105

PENDAHULUAN

Obesitas merupakan masalah kesehatan yang kian mendesak, baik dalam skala global maupun nasional. Di Indonesia, tren peningkatan jumlah penderita obesitas, terutama pada orang dewasa dan remaja, berbanding lurus dengan perubahan perilaku hidup. Pola makan yang tidak sehat dan penurunan aktivitas fisik menjadi pemicu utama, yang pada gilirannya meningkatkan prevalensi berbagai penyakit tidak menular (PTM) seperti diabetes, hipertensi, dan penyakit jantung(Aghniya & Prasetyowati, 2024). Berbagai riset di lapangan telah mengukuhkan adanya hubungan yang signifikan antara pola makan dan aktivitas fisik dengan kejadian obesitas. Misalnya, penelitian menunjukkan bahwa peningkatan konsumsi makanan diiringi pengurangan aktivitas fisik selama periode perubahan gaya hidup, seperti pandemi, secara langsung berhubungan dengan obesitas pada remaja (Mutia et al., 2022). Dampak serupa juga terlihat pada siswa sekolah dasar, di mana frekuensi makan dan tingkat aktivitas fisik memengaruhi status gizi dan obesitas (Octaviani et al., 2018), serta pada remaja sekolah menengah atas dengan hubungan kuat antara kedua faktor tersebut dan kejadian obesitas (Elsa Sari Saputri & Samsudi, 2024).

Kondisi tersebut menekankan pentingnya penelitian mendalam mengenai faktor pola makan, olahraga, dan kondisi fisik sebagai kontributor obesitas (Agus Mulyana et al., 2024). Penelitian lanjutan harus melampaui analisis hubungan linier, melainkan harus mampu membagi populasi menjadi kelompok-kelompok yang memiliki karakteristik risiko spesifik. Metode tradisional seperti survei atau statistik inferensial seringkali berasumsi populasi bersifat homogen, padahal variasi kebiasaan makan dan tingkat aktivitas fisik antar individu sangatlah nyata. Oleh karena itu, penerapan teknik pengolahan data seperti *clustering* menawarkan pendekatan metodologis inovatif untuk mengungkap pola tersembunyi dalam data obesitas.

Pemanfaatan metode komputasional di sektor kesehatan Indonesia telah menunjukkan beberapa upaya yang menjanjikan. Algoritma K-Means, sebagai contoh, telah diterapkan untuk mengelompokkan jenis makanan berdasarkan nilai gizi (Vredizon et al., 2024) dan untuk mengkategorikan status gizi anak balita (Ni Komang Sri Julyantari et al., 2021). Selain itu, metode *data mining* seperti *Random Forest* juga digunakan dalam memprediksi risiko obesitas berdasarkan konsumsi makanan (Rosidah et al., 2025)). Menariknya, salah satu studi bahkan membuktikan bahwa K-Means menghasilkan klaster terbaik berdasarkan indeks Davies-Bouldin dibandingkan metode pengelompokan lainnya (Setiawati et al., 2024). Meskipun demikian, penelitian di Indonesia yang secara spesifik memanfaatkan *dataset* publik internasional untuk pengelompokan kebiasaan makan dan pola aktivitas fisik guna estimasi obesitas masih tergolong jarang.

Berdasarkan latar belakang tersebut, studi ini akan memanfaatkan *dataset* publik internasional "*Estimation of Obesity Levels Based On Eating Habits and Physical Condition*" dari repositori UCI sebagai sumber data empiris. Dengan menerapkan *Metode K-Means Clustering*, penelitian ini bertujuan untuk mengidentifikasi pembagian populasi berdasarkan pola makan dan aktivitas fisik, sehingga karakteristik unik serta tingkat risiko obesitas pada setiap klaster dapat dianalisis. Keunggulan metodologis ini memungkinkan penemuan pola tersembunyi, yang tidak tercakup oleh analisis tradisional, dan menjadi dasar kuat untuk merumuskan intervensi yang lebih fokus dan tepat sasaran. Dengan demikian, studi ini diharapkan memberikan kontribusi akademis dan sekaligus fondasi ilmiah untuk program pencegahan obesitas dan perencanaan kebijakan kesehatan di Indonesia.

e-ISSN3032-601X & p-ISSN3032-7105

METODE PENELITIAN

1. Pendekatan penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan menerapkan teknik data mining berbasis algoritma K-Means. Seluruh proses analisis dilakukan menggunakan perangkat lunak RapidMiner, yang memiliki keunggulan dalam menangani alur preprocessing, transformasi data, dan klasterisasi secara terstruktur. Algoritma K-Means dipilih karena kemampuannya dalam mengelompokkan data numerik dengan cepat, efisien, dan akurat pada dataset berukuran kecil hingga menengah. Efektivitas algoritma ini telah dibuktikan dalam sejumlah penelitian, termasuk studi di bidang kesehatan yang memanfaatkan klasterisasi untuk pengelompokan pola perilaku dan risiko penyakit (Halawa et al., 2023).

2. Sumber Data dan Karakteristik Dataset

Dataset yang digunakan dalam penelitian ini merupakan data sekunder, yaitu *Estimation of Obesity Levels Based on Eating Habits and Physical Condition* yang berasal dari UCI Machine Learning Repository. Karena dataset ini bersifat publik dan telah tersedia secara lengkap, penelitian tidak memerlukan proses pengambilan data primer ataupun sampling. Seluruh individu dalam dataset diasumsikan sebagai populasi studi. Pendekatan populatif ini dinilai tepat ketika jumlah observasi relatif kecil hingga menengah, tersedia secara utuh, dan tidak membutuhkan generalisasi terhadap populasi yang lebih luas (Murni et al., 2018). Dataset tersebut berisi kombinasi variabel kategorikal dan numerik yang menggambarkan kebiasaan makan, aktivitas fisik, serta indikator yang berkaitan dengan potensi risiko obesitas.

3. Tahap Pra-Pemrosesan Data

Tahap *preprocessing* dilakukan untuk memastikan kualitas, konsistensi, dan kompatibilitas data terhadap algoritma yang digunakan. Proses ini meliputi pembersihan data, pengecekan nilai yang tidak konsisten, dan verifikasi format setiap atribut. Variabel kategorikal diubah menjadi numerik menggunakan teknik *one-hot* encoding dan *label encoding*, agar dapat diproses oleh algoritma K-Means yang hanya menerima data numerik. Penggunaan teknik encoding tersebut sejalan dengan praktik umum dalam klasterisasi data yang mengandung atribut kategorikal yang harus disesuaikan ke bentuk numerik (Fadilah & Wijayanto, 2023). Selanjutnya, normalisasi dilakukan menggunakan metode *z-score* untuk memastikan kesetaraan skala antar variabel sehingga perhitungan jarak Euclidean tidak mendistorsikan bobot fitur tertentu. Normalisasi berbasis *z-score* terbukti meningkatkan stabilitas pembentukan klaster pada algoritma berbasis jarak (Saqila et al., 2023).

4. Penentuan Jumlah Klaster Optimal

Penentuan jumlah klaster dilakukan menggunakan pendekatan *Elbow Method*, yakni dengan menguji nilai k dalam rentang 2 hingga 6 klaster. Setiap nilai k dievaluasi berdasarkan *Sum of Squared Error* (SSE), yaitu total jarak kuadrat antara anggota klaster dengan centroid-nya. Nilai k optimal dipilih pada titik ketika penurunan SSE mulai melambat secara signifikan, atau berada pada "titik siku" grafik. Pendekatan ini sesuai dengan rekomendasi dalam studi analisis varian internal klaster yang memprioritaskan keseimbangan antara kompaksi klaster dan efisiensi segmentasi (Refialy et al., 2021). Prosedur ini memastikan bahwa jumlah klaster yang dihasilkan tidak terlalu sedikit sehingga informasi penting hilang, dan tidak terlalu banyak hingga menghasilkan segmentasi berlebihan.

5. Proses Klasterisasi

Setelah menetapkan nilai k optimal, dilakukan proses klasterisasi menggunakan operator *K-Means Clustering* pada RapidMiner. Output utama yang dihasilkan meliputi nilai centroid, ukuran masing-

masing klaster, dan jarak rata-rata antar anggota klaster. Centroid berfungsi sebagai representasi nilai rata-rata dari setiap atribut pada klaster. Hasil klasterisasi kemudian dianalisis untuk melihat kecenderungan pola makan, tingkat aktivitas fisik, serta variabel pendukung lainnya. Analisis deskriptif digunakan untuk menginterpretasikan karakteristik setiap klaster dan mengidentifikasi pola risiko yang muncul. Proses interpretasi ini mengikuti pendekatan pada studi perilaku konsumen yang menggunakan klasterisasi sebagai dasar segmentasi risiko (Fajar et al., 2025).

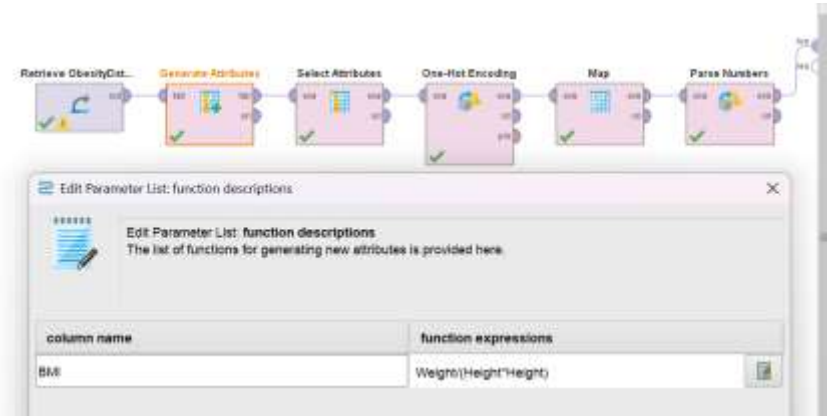
6. Validasi dan Interpretasi Hasil Klaster

Pada tahap akhir, hasil klasterisasi dievaluasi kembali untuk memastikan bahwa pembentukan klaster konsisten dengan pola data. Validasi dilakukan melalui pemeriksaan nilai jarak antar centroid dan distribusi anggota klaster guna memastikan setiap klaster memiliki karakter yang berbeda secara signifikan. Selanjutnya, interpretasi hasil dilakukan dengan menelaah atribut kebiasaan makan dan aktivitas fisik untuk menilai kecenderungan risiko obesitas pada masing-masing klaster. Tahap ini bertujuan memberikan pemahaman yang komprehensif terkait segmentasi risiko obesitas berdasarkan karakteristik perilaku individu.

HASIL DAN PEMBAHASAN

Data Prepocessing

Tahap pertama dari penelitian dimulai dengan pengolahan data untuk memastikan bahwa semua atribut berada dalam format yang seragam dan siap untuk analisis klaster. Langkah pertama adalah menciptakan atribut baru yang disebut Indeks Massa Tubuh (BMI) menggunakan rumus standar, yaitu berat badan dalam kilogram dibagi dengan tinggi badan dalam meter kuadrat. Atribut BMI ini dihasilkan dengan menggunakan fitur *Generate Attributes* di *RapidMiner*, sehingga indeks massa tubuh dapat diperhitungkan sebagai variabel numerik tambahan dalam langkah klasterisasi.



Gambar 1 Proses *Generate Attributes* untuk BMI

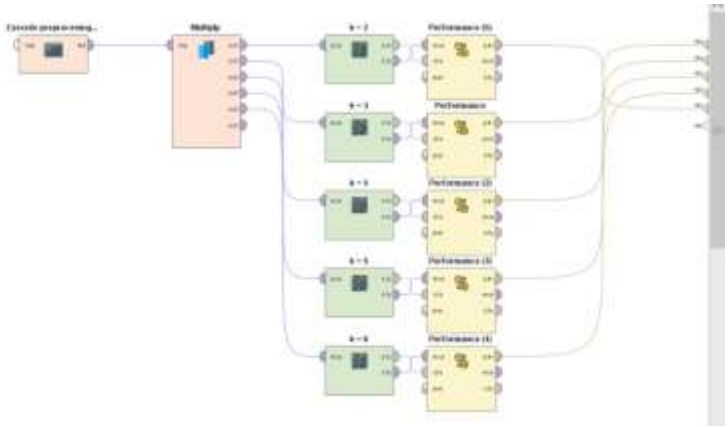
Langkah selanjutnya adalah pemilihan atribut untuk menghapus variabel yang tidak relevan dengan tujuan analisis. Variabel yang diabaikan termasuk Label, SCC, dan *Smoke*, karena tidak memberikan kontribusi langsung dalam segmentasi risiko obesitas atau berpotensi mengganggu pembentukan klaster. Setelah pemilihan atribut, proses dilanjutkan dengan One Hot Encoding untuk mengonversi atribut kategorikal menjadi bentuk numerik. Atribut Gender dan MTRANS diubah menjadi beberapa variabel biner

agar bisa diproses oleh algoritma K-Means yang membutuhkan data dalam format numerik. Di samping itu, atribut kategorikal lainnya yang tidak dapat diubah dengan *One Hot Encoding* diproses melalui *Label Encoding* sebagai alternatif konversi. Seluruh rangkaian proses ini memastikan bahwa dataset telah dalam format numerik, bebas dari inkonsistensi, dan memenuhi syarat teknis untuk analisis klaster berbasis jarak.

Tabel 1 Contoh Data Sebelum dan Sesudah *Preprocessing*

Tahap	Atribut	Contoh Nilai Sebelum	Contoh Nilai Sesudah
Generate BMI	Berat, Tinggi	70 kg; 1.65 m	BMI = 25.71
Select Attributes	Label	Obesity_Type_1	Dihapus
One Hot Encoding	Gender	Male/Female	Gender_Male = 1, Gender_Female = 0
One Hot Encoding	MTRANS	Automobile	MTRANS_Automobile = 1, others = 0
Label Encoding	FAVC	yes/no	1/0

Penentuan Nilai K Menggunakan Metode Elbow



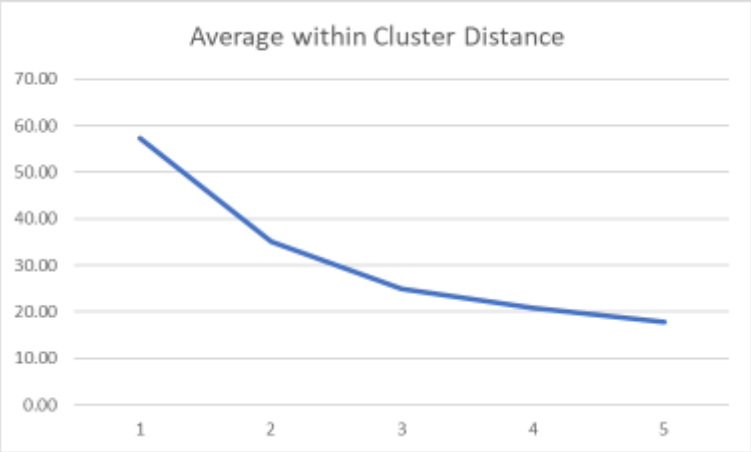
Gambar 2 Proses Perbandingan Nilai *Avg. Within Centroid Distance*

Penentuan jumlah klaster yang paling baik dilakukan dengan memanfaatkan metode *Elbow*, yang merupakan cara untuk menilai pergeseran nilai *Average Within Centroid Distance* terhadap berbagai opsi K. Proses pengujian dilakukan menggunakan fitur *Cluster Distance Performance* di *RapidMiner* dengan memperhatikan tren penurunan nilai kesalahan atau distorsi saat jumlah klaster meningkat.

Tabel 2 Nilai *Average Within Centroid Distance* untuk Beberapa K

K	Average Within Cluster Distance
2	57.41
3	35.067
4	24.982
5	20.846
6	17.887

e-ISSN3032-601X & p-ISSN3032-7105

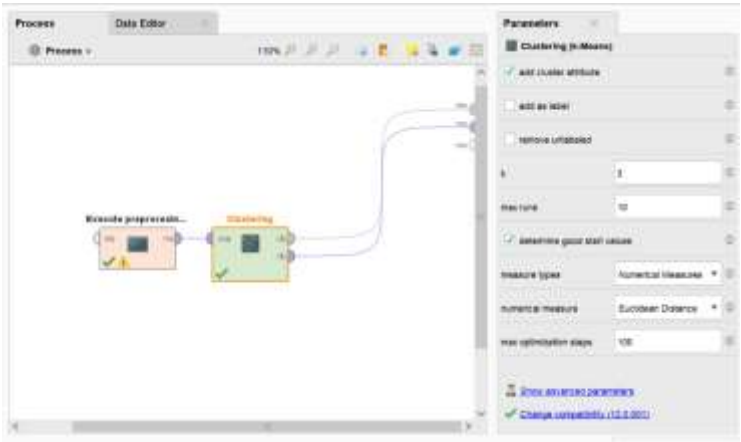


Gambar 3. Grafik *Elbow* Berdasarkan *Average Within Centroid Distance*

Analisa yang dilakukan menunjukkan adanya titik siku yang jelas pada K = 3, ditunjukkan dengan penurunan nilai *within-cluster distance* yang tidak signifikan untuk K di atas angka itu. Oleh karena itu, K = 3 diambil sebagai jumlah klaster yang paling tepat untuk merefleksikan struktur alami dalam set data. Keputusan ini sejalan dengan prinsip utama metode elbow, yaitu memilih K pada titik di mana perubahan terjadi dengan tajam sebelum kurva distorsi mulai datar.

Proses Klastering Menggunakan *K-Means*

Setelah jumlah klaster optimal ditetapkan, proses *clustering* dilakukan menggunakan algoritma *K-Means* dengan nilai K = 3. Algoritma dijalankan melalui operator *K-Means* di RapidMiner yang menghitung jarak Euclidean antara setiap data dengan centroid awal yang ditentukan secara acak. Selanjutnya, centroid diperbarui secara iteratif sampai mencapai kondisi konvergen, yaitu ketika perubahan antar iterasi tidak lagi signifikan.



Gambar 4. Proses *K-Means Clustering* di RapidMiner

Hasil dari pemodelan menunjukkan pembagian dataset ke dalam tiga klaster dengan karakteristik centroid yang berbeda berdasarkan pola gaya hidup, kebiasaan konsumsi makanan, aktivitas fisik, dan nilai BMI. Klaster-klaster tersebut kemudian diinterpretasikan lebih lanjut pada bagian berikutnya untuk

mengetahui kelompok berisiko rendah, sedang, maupun tinggi terhadap obesitas. Visualisasi hasil klaster yang diperoleh menunjukkan bahwa ketiga klaster terbentuk secara jelas dengan pemisahan jarak yang cukup tegas, mengindikasikan bahwa atribut yang digunakan efektif dalam membedakan pola risiko obesitas antar individu.

Gambar 5 Visualisasi Klaster

Tabel 3 Hasil Centroid Akhir untuk K = 3

Atribut	Cluster 0	Cluster 1	Cluster 2
family_history_with_overweight	0.638	0.994	0.915
FAVC	0.798	0.985	0.886
CAEC	1.255	1.026	1.085
CALC	0.701	0.776	0.71
Gender = Male	0.508	0.505	0.502
MTRANS = Public_Transportation	0.83	0.888	0.139
MTRANS = Walking	0.049	0.004	0.016
MTRANS = Automobile	0.108	0.109	0.83
MTRANS = Motorbike	0.008	0	0.009
Age	20.592	24.092	36.36
FCVC	2.34	2.549	2.356

Interpretasi Klaster Berdasarkan Nilai Centroid

Analisis pengelompokan mengungkap tiga kategori risiko obesitas yang berbeda. Kluster 0 menggambarkan kategori Dewasa Muda dengan Risiko Rendah hingga Sedang, yang ditandai oleh angka BMI terendah (23,012) dan usia yang paling muda. Meskipun mereka memiliki kecenderungan genetik yang moderat serta pola makan camilan yang relatif tinggi, kelompok ini masih menjalani gaya hidup yang cukup aktif, ditandai dengan penggunaan transportasi umum yang dominan. Kluster 1 menunjukkan kategori dengan Risiko Obesitas Paling Tinggi, yang terlihat dari BMI tertinggi (37,64). Risiko yang sangat tinggi ini dipengaruhi oleh faktor genetik yang sangat kuat, hampir 100% memiliki riwayat kelebihan berat badan dan pola makan yang buruk (sangat tingginya konsumsi makanan berkalori), yang tidak diimbangi dengan tingkat aktivitas fisik yang memadai. Terakhir, Kluster 2 mewakili kategori Dewasa Lebih Tua dengan Risiko Sedang hingga Tinggi, yang ditandai oleh usia tertinggi dan BMI menengah-tinggi (29,805). Kluster ini menunjukkan pola makan yang tidak sehat serta gaya hidup paling sedentari, dengan aktivitas fisik yang sangat minimal dan dominasi penggunaan kendaraan pribadi, yang menandakan bahwa mobilitas

pasif adalah faktor besar yang berkontribusi terhadap risiko obesitas pada kelompok ini.

Tabel 4 Hasil Interpretasi Klaster Berdasarkan Nilai Centroid

Klaster	Karakteristik Utama	Risiko Obesitas
0	Usia muda, BMI rendah, mobilitas aktif, kebiasaan makan sedang	Rendah
1	BMI sangat tinggi, faktor genetik kuat, makan tidak sehat, aktivitas rendah	Tinggi
2	Usia lebih tua, BMI tinggi sedang, mobilitas sangat pasif (mobil), pola makan kurang sehat	Sedang

KESIMPULAN

Penerapan metode *K-Means clustering* pada dataset *Estimation of Obesity Levels Based on Eating Habits and Physical Condition* sukses menampilkan pola pengelompokan risiko obesitas yang berkaitan dengan karakteristik kebiasaan makanan dan kondisi fisik para responden. Melalui tahap pra-pemrosesan data yang mencakup pembuatan atribut BMI, pemilihan atribut yang relevan, serta penerapan encoding untuk variabel kategorikal, data telah disiapkan dengan baik untuk langkah klasterisasi. Penetapan jumlah klaster dengan metode *elbow* menunjukkan bahwa k = 3 adalah angka yang paling representatif berdasarkan penurunan rata-rata jarak antara titik dan centroid (*Average Within Centroid Distance*). Hasil dari klasterisasi menghasilkan tiga kelompok dengan karakteristik yang berbeda, mencakup variasi risiko obesitas yang terdiri dari kategori rendah, sedang, dan tinggi. Penemuan ini menunjukkan bahwa variabel kebiasaan makan, tingkat aktivitas fisik, dan rutinitas harian berperan penting dalam pembentukan klaster. Secara keseluruhan, studi ini mengindikasikan bahwa metode K-Means dapat berfungsi dengan baik untuk memetakan risiko obesitas dan memberikan landasan analitis untuk pengembangan sistem pendukung keputusan di bidang kesehatan preventif.

DAFTAR PUSTAKA

Aghniya, R., & Prasetyowati, P. (2024). Deteksi Dini dan Pencegahan Penyakit Tidak Menular Melalui Aktivitas Fisik, Edukasi dan Promosi Kesehatan Di UPTD Yosomulyo Kota Metro. *Jurnal Pengabdian Sosial*, *1*(6), 408–413. https://doi.org/10.59837/tpmh3j73

Agus Mulyana, Dela Lestari, Dhilla Pratiwi, Nabila Mufidah Rohmah, Nabila Tri, Neng Nisa Audina Agustina, & Salma Hefty. (2024). Menumbuhkan Gaya Hidup Sehat Sejak Dini Melalui Pendidikan Jasmani, Olahraga, Dan Kesehatan. *Jurnal Bintang Pendidikan Indonesia*, *2*(2), 321–333. https://doi.org/10.55606/jubpi.v2i2.2998

Elsa Sari Saputri, & Samsudi. (2024). Hubungan Pola Makan Dan Aktivitas Fisik Dengan Kejadian Obesitas Pada Remaja di SMA Negeri 1 Abuki . *Jurnal Penelitian Sains Dan Kesehatan Avicenna*, *3*(2 SE-Artikel), 156–164. https://doi.org/10.69677/avicenna.v3i2.86

Fadilah, Z. R., & Wijayanto, A. W. (2023). Perbandingan Metode Klasterisasi Data Bertipe Campuran: One-Hot-Encoding, Gower Distance, dan K-Prototype Berdasarkan Akurasi (Studi Kasus: Chronic Kidney Disease Dataset). *Journal of Applied Informatics and Computing*, *7*(1), 57–67. https://doi.org/10.30871/jaic.v7i1.5857

Fajar, M., Adam, S., Putra, B., Puteri, S. I., Fajrissiddiq, A., & Sani, L. (2025). Eksplorasi dan Analisis Data Mining untuk Prediksi Pola Konsumen Menggunakan Teknik Klasifikasi dan Clustering. *SENTIMETER (Seminar Nasional Teknologi Informasi, Mekatronika Dan Ilmu Komputer) Universitas Nusa Putra, 17 Mei 2025 Eksplorasi*.

Halawa, H., Jaya Harmaja, O., Hulu, W. sandi, & Loi, seriani. (2023). Implementasi Algoritma K-Means Clustering Untuk Pengelompokkan Penyakit Pasien Pada Puskesmas Pulo Brayan. *Jurnal Sains Dan Teknologi*, *5*(1 SE-), 150–157. http://ejournal.sisfokomtek.org/index.php/saintek/article/view/1306

Murni, A. P., Sekolah, A. P., Kelulusan, A., Sekolah, J., Pendahuluan, I., Siswa, J., Kasar, A. P., Partisipasi, A., Sekolah, A. P., Kelulusan, A., Melanjutkan, A., & Sekolah, J. (2018). APLIKASI PEMETAAN KUALITAS PENDIDIKAN DI INDONESIA MENGGUNAKAN METODE K-MEANS Analisis perancangan yang digunakan dalam pembuatan sistem ini menggunakan UML (Unified Modeling Language) dimana setiap aktivitas pada sistem akan dikelompokkan secara sendir. *Aplikasi Pemetaan Kualitas Pendidikan Di Indonesia Menggunakan Metode K-Means*, *17*(2), 13–23.

Mutia, A., Jumiyati, J., & Kusdalinah, K. (2022). Pola makan dan aktivitas fisik terhadap kejadian obesias remaja. *Journal of Nutrition College*, *11*(1), 26–34. http://ejournal3.undip.ac.id/index.php/jnc/

Ni Komang Sri Julyantari, Budiarta, I. K., & Putri, N. M. D. K. (2021). Implementasi K-Means Untuk Pengelompokan Status Gizi Balita (Studi Kasus Banjar Titih). *Jurnal Janitra Informatika Dan Sistem Informasi*, *1*(2), 92–101. https://doi.org/10.25008/janitra.v1i2.134

Octaviani, P., Dody Izhar, M., & Amir, A. (2018). Relation Between Dietary Habit and Physical Activity With Nutritional Status Of Elementary School Students in SD Negeri 47/IV Jambi City. *Jurnal Kesmas Jambi*, *2*(2), 56–66.

Refialy, L. P., Maitimu, H., & Pesulima, M. S. (2021). Perbaikan Kinerja Clustering K-Means pada Data Ekonomi Nelayan dengan Perhitungan Sum of Square Error (SSE) dan Optimasi nilai K cluster. *Techno.Com*, *20*(2), 321–329. https://doi.org/10.33633/tc.v20i2.4572

Rosidah, N., Prathivi, R., & Susanto, S. (2025). OPTIMASI METODE RANDOM FOREST UNTUK KLASIFIKASI RISIKO OBESITAS BERDASARKAN POLA MAKAN. *Jurnal Informatika Teknologi Dan Sains (Jinteks)*, *7*(1 SE-Articles), 56–62. https://doi.org/10.51401/jinteks.v7i1.5065

Saqila, S. E., Ferina, I. P., & Iskandar, A. (2023). Analisis Perbandingan Kinerja Clustering Data Mining Untuk Normalisasi Dataset. *Jurnal Sistem Komputer Dan Informatika (JSON)*, *5*(2), 356. https://doi.org/10.30865/json.v5i2.6919

Setiawati, E., Fernanda, U. D., & Agesti, S. (2024). *Implementation of K-Means , K-Medoid and DBSCAN Algorithms In Obesity Data Clustering*. *1*(February), 23–29.

Vredizon, P. A., Firmansyah, H., Salsabila, N. S., & Nugroho, W. E. (2024). *Penerapan Algoritma K-Means untuk Mengelompokkan Makanan Berdasarkan Nilai Nutrisi*. *5*(2). https://doi.org/10.37802/joti.v5i2.577