



**EVALUASI JUMLAH *CLUSTER* OPTIMAL PADA *K-MEDOIDS*
MENGUNAKAN *SILHOUETTE* DAN *QUANTITATIVE*
DISCRIMINANT METHOD DENGAN VALIDASI *DBI*
UNTUK PENGELOMPOKAN BALITA
BERDASARKAN INDIKATOR STATUS GIZI**

SKRIPSI

**Untuk memenuhi persyaratan dalam
menyelesaikan program sarjana Strata-1 Matematika**

**Oleh:
MUHAMMAD NAUFAL NOR AKMAL
NIM. 2211011210021**

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
2026**

HALAMAN PENGESAHAN

SKRIPSI

EVALUASI JUMLAH *CLUSTER* OPTIMAL PADA *K-MEDOIDS* MENGGUNAKAN *SILHOUETTE* DAN *QUANTITATIVE DISCRIMINANT METHOD* DENGAN VALIDASI *DBI* UNTUK PENGELOMPOKAN BALITA BERDASARKAN INDIKATOR STATUS GIZI

Oleh:

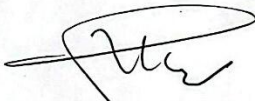
Muhammad Naufal Nor Akmal

NIM.2211011210021

telah dipertahankan di depan Dosen Penguji pada tanggal 30 Juni 2026.

Susunan Dosen Penguji:

Pembimbing I



Drs. Faisal, M.Si.
NIP. 196309021992031001

Dosen Penguji

1. Dr. Muhamad Ahsar Karim, S.Si., M.Sc.
2. Muhammad Afief Balya, S.Mat., M.Si.



Pembimbing II



Oni Soesanto, S.Si., M.Si.
NIP. 197301262005011003

Banjarbaru, 16 Juli 2026
Ketua Jurusan Matematika
FMIPA ULM



Dr. Na'imah Hijriati, S.Si., M.Si.
NIP. 19791122200812013

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Banjarbaru, 16 Juli 2026



Muhammad Naufal Nor Akmal
2211011210021

ABSTRAK

EVALUASI JUMLAH *CLUSTER* OPTIMAL PADA *K-MEDOIDS* MENGGUNAKAN *SILHOUETTE* DAN *QUANTITATIVE DISCRIMINANT METHOD* DENGAN VALIDASI *DBI* UNTUK PENGELOMPOKAN BALITA BERDASARKAN INDIKATOR STATUS GIZI (Oleh: Muhammad Naufal Nor Akmal; Pembimbing: Faisal, Oni Soesanto; 2025; 51 halaman)

Gizi balita merupakan faktor penting yang menentukan kualitas pertumbuhan dan perkembangan anak pada usia dini. Permasalahan gizi, termasuk *stunting* yang masih memiliki prevalensi sekitar 19,8% di Indonesia, menunjukkan perlunya pemantauan status gizi yang lebih efektif. Analisis *clustering* dapat digunakan untuk mengelompokkan balita berdasarkan kesamaan karakteristik gizinya sehingga memudahkan identifikasi kelompok-kelompok yang perlu mendapatkan perhatian khusus. Tujuan penelitian ini adalah untuk menentukan dan mengevaluasi jumlah *cluster* optimal dari pengelompokan gizi balita. Metode yang digunakan pada penelitian ini adalah algoritma *K-Medoids* dengan evaluasi jumlah *cluster* optimal menggunakan metode *Silhouette Coefficient* dan *Quantitative Discriminant Method* (QDM), sedangkan jumlah *cluster* yang dihasilkan dari kedua metode tersebut selanjutnya divalidasi menggunakan *Davies–Bouldin Index* (DBI). Hasil penelitian menunjukkan bahwa pengelompokan gizi balita menggunakan *K-Medoids* dengan metode evaluasi *Silhouette Coefficient* dan QDM menghasilkan jumlah *cluster* optimal pada $K=4$, dengan nilai *Silhouette* sebesar 0,605 yang juga konsisten dengan nilai sudut interseksi terkecil pada QDM. Validasi DBI menghasilkan nilai 0,547 yang menandakan kualitas *cluster* baik. *Cluster* yang dihasilkan menunjukkan bahwa dua *cluster* merepresentasikan balita dengan gizi kurang dan dua *cluster* lainnya merepresentasikan balita dengan gizi normal.

Kata kunci: *K-Medoids*, *Silhouette*, QDM, DBI, Gizi Balita

ABSTRACT

EVALUATION OF THE OPTIMAL NUMBER OF CLUSTERS IN K-MEDOIDS USING SILHOUETTE AND QUANTITATIVE DISCRIMINANT METHODS WITH DBI VALIDATION FOR GROUPING TODDLERS BASED ON NUTRITIONAL STATUS INDICATORS (By: Muhammad Naufal Nor Akmal; Supervisors: Faisal, Oni Soesanto; 2025; 51 pages)

Toddler nutrition is an important factor that determines the quality of children's growth and development at an early age. Nutritional problems, including stunting which still has a prevalence of approximately 19.8% in Indonesia, indicate the need for more effective monitoring of nutritional status. Clustering analysis can be used to group toddlers based on similarities in their nutritional characteristics, thereby facilitating the identification of groups that need special attention. The objective of this study is to determine and evaluate the optimal number of clusters for grouping toddler nutrition. The method used in this study is the K-Medoids algorithm with the evaluation of the optimal number of clusters using the Silhouette Coefficient and the Quantitative Discriminant Method (QDM), while the number of clusters produced by both methods will be validated using the Davies–Bouldin Index (DBI). The results show that the grouping of toddler nutrition using K-Medoids with the Silhouette Coefficient and QDM evaluation methods yields an optimal number of clusters at $K = 4$, with a Silhouette value of 0.605, which is also consistent with the smallest intersection angle value obtained from the QDM. DBI validation produces a value of 0.547, indicating good cluster quality. The resulting clusters show that two clusters represent underweight toddlers and the other two clusters represent toddlers with normal nutritional status.

Keywords: K-Medoids, Silhouette, QDM, DBI, Toddler Nutrition

PRAKATA

Puji syukur penulis panjatkan ke hadirat Allah Subhanahu Wa Ta'ala yang telah melimpahkan rahmat, taufik, dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Evaluasi Jumlah *Cluster* Optimal Pada *K-Medoids* Menggunakan *Silhouette* Dan *Quantitative Discriminant Method* Dengan Validasi DBI Untuk Pengelompokan Balita Berdasarkan Indikator Status Gizi”. Penyusunan skripsi ini bertujuan untuk memenuhi persyaratan dalam menyelesaikan program sarjana Strata-1 Matematika di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lambung Mangkurat.

Penulis menyadari bahwa penyusunan skripsi ini tidak akan berhasil tanpa bantuan, arahan, serta motivasi dari berbagai pihak yang terlibat. Oleh karena itu, penulis menyampaikan rasa hormat dan terima kasih kepada:

1. Bapak Prof. Ir. Sunardi, S.Si., M.Sc., Ph.D. selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat.
2. Ibu Dr. Nai'mah Hijriati, S.Si., M.Si., selaku Ketua Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat.
3. Bapak Akhmad Yusuf, S.Si., M.Kom., dan Bapak Muhammad Afief Balya, S.Mat., M.Si., selaku dosen pembimbing akademik yang selalu memberikan arahan selama masa perkuliahan.
4. Bapak Drs. Faisal, M.Si., selaku dosen pembimbing I dan Bapak Oni Soesanto, S.Si., M.Si., selaku dosen pembimbing II yang telah memberikan bantuan, bimbingan, nasihat, petunjuk, serta motivasi dalam pelaksanaan penelitian dan penyusunan skripsi ini.
5. Seluruh dosen dan staf Jurusan Matematika FMIPA ULM atas ilmu dan bantuan yang diberikan selama ini yang sangat bermanfaat.
6. Sahabat penulis yaitu Abdul Hadi, Ghaitsa Zahira Shofa, Siti Rahmah, Sabna, Arsella Putri Nabila dan Annisa Kamilah yang selalu mengingatkan dalam berbagai hal serta selalu mendukung selama masa perkuliahan.

7. Seluruh rekan mahasiswa Jurusan Matematika FMIPA ULM, khususnya teman-teman Matrix angkatan 2022 yang telah memberikan dukungan, bantuan dan motivasi kepada penulis selama penyusunan skripsi.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih banyak kekurangan dan kesalahan, sehingga saran dan kritik yang membangun sangat penulis harapkan. Akhir kata penulis berharap skripsi ini dapat bermanfaat bagi pembaca dan semua pihak.

Banjarbaru, 16 Juli 2026



Muhammad Naufal Nor Akmal
2211011210021

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
PERNYATAAN	iii
ABSTRAK	iv
ABSTRACT	v
PRAKATA	vi
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xii
ARTI LAMBANG DAN SINGKATAN	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan Penelitian	3
1.3 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	4
2.1 Data Mining	4
2.2 <i>Cross-Industry Standard Process for Data Mining (CRISP-DM)</i>	6
2.3 <i>Clustering</i>	8
2.4 <i>Outlier</i>	9
2.5 Algoritma <i>K-Medoids</i>	10
2.6 <i>Silhouette Coefficient</i>	12
2.7 <i>Quantitative Discriminant Method (QDM)</i>	13
2.8 <i>Davies-Bouldin Index (DBI)</i>	15
2.9 Gizi Balita	16
2.10 Standar Antropometri Anak	18
BAB III METODOLOGI PENELITIAN	20
3.1 Metode Analisis	20
3.2 Data Penelitian	20
3.3 Prosedur Penelitian	21
3.4 Diagram Alur Penelitian	22

BAB IV HASIL DAN PEMBAHASAN	23
4.1 Persiapan Data	23
4.2 Preprocessing Data.....	23
4.3 Clusterisasi Data	27
4.4 Evaluasi Jumlah Cluster Optimal.....	32
4.4.1 <i>Silhouette Coefficient</i>	33
4.4.2 <i>Quantitative Discriminant Method (QDM)</i>	36
4.5 Validasi Jumlah Cluster Optimal	44
4.6 Interpretasi Hasil Cluster	47
BAB V PENUTUP	49
5.1 Kesimpulan	49
5.2 Saran	49
DAFTAR PUSTAKA	50
LAMPIRAN	53

DAFTAR TABEL

	Halaman
Tabel 2.1 Kategori dan Ambang Batas Status Gizi Balita	19
Tabel 4.1 Data Balita.....	23
Tabel 4.2 Hasil Transformasi Atribut Jenis Kelamin.....	25
Tabel 4.3 Data Balita setelah Winsorizing.....	26
Tabel 4.4 Data Balita setelah Normalisasi Min-Max.....	27
Tabel 4.5 Medoid Awal.....	28
Tabel 4.6 Hasil Perhitungan Jarak untuk $K=4$	29
Tabel 4.7 Hasil Pengelompokan Data untuk $K=4$	29
Tabel 4.8 Kandidat Medoid Baru untuk $K=4$	30
Tabel 4.9 Hasil Pengelompokan Data Baru untuk $K=4$	30
Tabel 4.10 Kandidat Medoid Baru ke-2 untuk $K=4$	31
Tabel 4.11 Hasil Pengelompokan Data Baru untuk $K=4$	31
Tabel 4.12 Final Medoid untuk $K=4$	32
Tabel 4.13 Hasil Cluster dari $K=2$ sampai $K=10$	32
Tabel 4.14 Hasil Perhitungan Silhouette Coefficient untuk $K=4$	35
Tabel 4.15 Nilai Silhouette Coefficient untuk $K=2$ sampai $K=10$	35
Tabel 4.16 Data Anggota Cluster 1 untuk $K=4$	37
Tabel 4.17 Data Anggota Cluster 2 untuk $K=4$	38
Tabel 4.18 Data Anggota Cluster 3 untuk $K=4$	39
Tabel 4.19 Data Anggota Cluster 4 untuk $K=4$	40
Tabel 4.20 Nilai SSE untuk $K=1$ sampai $K=10$	41
Tabel 4.21 Nilai Mean Distortion untuk $K=1$ sampai $K=10$	41
Tabel 4.22 Nilai Normalisasi MD untuk $K=1$ sampai $K=10$	42
Tabel 4.23 Nilai Jarak Euclidian E_{ij} , E_{jk} , dan E_{ik}	42
Tabel 4.24 Nilai Sudut Interseksi untuk $K=2$ sampai $K=9$	43
Tabel 4.25 Nilai DBI untuk $K=2$ sampai $K=10$	46
Tabel 4.26 Final Medoid $K=4$	46

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Proses <i>Knowledge Discovery in Database</i>	5
Gambar 2.2 Tahapan Bagian CRISP-DM	7
Gambar 2.3 Skema Identifikasi Outlier dengan Boxplot	10
Gambar 3.1 Diagram Alur Penelitian.....	22
Gambar 4.1 Hasil Pengecekan <i>Missing Value</i>	24
Gambar 4.2 Boxplot Data Balita	25
Gambar 4.3 Boxplot Data Balita setelah Winsorizing	26
Gambar 4.4 Grafik Perhitungan Silhouette K=2 sampai K=10.....	36
Gambar 4.5 Grafik Perhitungan QDM K=2 sampai K=9.....	44

DAFTAR LAMPIRAN

- Lampiran 1.** Tabel Data Balita yang digunakan dalam Penelitian.
- Lampiran 2.** Tabel Transformasi Atribut Jenis Kelamin.
- Lampiran 3.** Tabel Data Balita Setelah Winsorizing
- Lampiran 4.** Tabel Data Balita setelah Normalisasi Min-Max
- Lampiran 5** Tabel Hasil Perhitungan Jarak untuk $K=4$
- Lampiran 6.** Tabel Hasil Pengelompokan Data untuk $K=4$
- Lampiran 7.** Tabel Hasil Pengelompokan Data Baru untuk $K=4$
- Lampiran 8.** Tabel Hasil Pengelompokan Data Baru untuk $K=4$
- Lampiran 9.** Tabel Hasil Perhitungan Silhouette Coefficient untuk $K=4$
- Lampiran 10.** Tabel Data Anggota Cluster 1 untuk $K=4$
- Lampiran 11.** Tabel Data Anggota Cluster 2 untuk $K=4$
- Lampiran 12.** Tabel Data Anggota Cluster 3 untuk $K=4$
- Lampiran 13.** Tabel Data Anggota Cluster 4 untuk $K=4$
- Lampiran 14.** Tabel Standar Berat Badan menurut Tinggi Badan Balita Laki-Laki
- Lampiran 15.** Tabel Standar Berat Badan menurut Tinggi Badan Balita Perempuan
- Lampiran 16.** Script Algoritma Google Colab K-Medoids
- Lampiran 17.** Script Algoritma Google Colab Silhouette
- Lampiran 18.** Script Algoritma Google Colab QDM
- Lampiran 19.** Script Algoritma Google Colab DBI