

**PERBANDINGAN METODE *K-MEDOIDS* DAN *FUZZY POSSIBILISTIC C-MEANS* DALAM MENGELOMPOKKAN
TINGKAT KESEHATAN BALITA BERDASARKAN
PENANGANAN KASUS STUNTING DI INDONESIA**

Skripsi

**Disusun untuk memenuhi salah satu syarat
memperoleh gelar Sarjana Matematika**



Adam Ramadhan

1305621002

**PROGRAM STUDI MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI JAKARTA**

2025

LEMBAR PERSETUJUAN HASIL SIDANG SKRIPSI

PERBANDINGAN METODE K-MEDOIDS DAN FUZZY

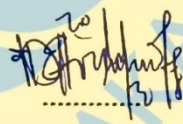
POSSIBILISTIC C-MEANS DALAM MENGELOMPOKKAN TINGKAT

KESEHATAN BALITA BERDASARKAN PENANGANAN KASUS

STUNTING DI INDONESIA

Nama : Adam Ramadhan

No. Registrasi : 1305621002

	Nama	Tanda Tangan	Tanggal
Penanggung Jawab			
Dekan	: <u>Dr. Hadi Nasbey, S.Pd., M.Si.</u> NIP. 197909162005011004		<u>29-07-2025</u>
Wakil Penanggung Jawab			
Wakil Dekan I	: <u>Dr. Meiliasari, S.Pd., M.Sc.</u> NIP. 197905042009122002		<u>29-07-2025</u>
Ketua	: <u>Dr. Lukita Ambarwati, S.Pd., M.Si.</u> NIP. 197210262001122001		<u>22-07-2025</u>
Sekretaris	: <u>Devi Eka Wardani M., S.Pd., M.Si.</u> NIP. 199005162019032014		<u>18-07-2025</u>
Penguji Ahli	: <u>Ibnu Hadi, M.Si.</u> NIP. 198107182008011017		<u>21-07-2025</u>
Pembimbing I	: <u>Dr. Eti Dwi Wiraningsih, S.Pd., M.Si.</u> NIP. 198102032006042001		<u>22-07-2025</u>
Pembimbing II	: <u>Dr. Yudi Mahatma, M.Si.</u> NIP. 197610202008121001		<u>22-07-2025</u>

Dinyatakan lulus ujian skripsi tanggal 16 Juli 2025

LEMBAR PERNYATAAN

LEMBAR PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan di bawah ini, mahasiswa Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta:

Nama : Adam Ramadhan
No Registrasi : 1305621002
Program Studi : Matematika

Dengan ini menyatakan bahwa skripsi yang saya buat dengan judul *"Perbandingan Metode K-Medoids dan Fuzzy Possibilistic C-Means dalam Mengelompokkan Tingkat Kesehatan Balita Berdasarkan Penanganan Kasus Stunting di Indonesia* adalah:

1. Dibuat sendiri, mengadopsi hasil kuliah, buku-buku, dan referensi acuan yang tertera di dalam referensi pada skripsi saya.
2. Bukan merupakan hasil duplikasi skripsi yang telah dipublikasikan atau pernah dipakai untuk mendapatkan gelar sarjana di Universitas lain kecuali pada bagian-bagian sumber informasi dicantumkan berdasarkan tata cara referensi yang semestinya.

Pernyataan ini dibuat dengan sesungguhnya dan saya bersedia menanggung segala akibat yang timbul jika pernyataan saya tidak benar.

Jakarta, 29 Juni 2025



Adam Ramadhan



KEMENTERIAN PENDIDIKAN TINGGI, SAINS DAN TEKNOLOGI
UNIVERSITAS NEGERI JAKARTA
UPT PERPUSTAKAAN

Jalan Rawamangun Muka Jakarta 13220
Telepon/Faksimili: 021-4894221
Laman: lib.unj.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika Universitas Negeri Jakarta, yang bertanda tangan di bawah ini, saya:

Nama : Adam Ramadhan
NIM : 1305621002
Fakultas/Prodi : Matematika dan Ilmu Pengetahuan Alam / Matematika
Alamat email : radam2536@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada UPT Perpustakaan Universitas Negeri Jakarta, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah:

☒ Skripsi ☐ Tesis ☐ Disertasi ☐ Lain-lain (.....)

yang berjudul :

Perbandingan Metode K-Medoids dan Fuzzy Possibilistic C-Means dalam Mengelompokkan
Tingkat Kesehatan Balita Berdasarkan Penanganan Kasus Stunting di Indonesia

Dengan Hak Bebas Royalti Non-Eksklusif ini UPT Perpustakaan Universitas Negeri Jakarta berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya, dan menampilkan/mempublikasikannya di internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan Universitas Negeri Jakarta, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 30 Juli 2025

Penulis

(Adam Ramadhan)
nama dan tanda tangan

ABSTRAK

ADAM RAMADHAN. Perbandingan Metode *K-Medoids* dan *Fuzzy Possibilistic C-Means* dalam Mengelompokkan Tingkat Kesehatan Balita Berdasarkan Penanganan Kasus Stunting di Indonesia. Skripsi, Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta. Juni 2025.

Clustering adalah teknik pengelompokan data yang pada dasarnya dilakukan atas suatu kesamaan atau karakteristik tertentu, seperti jarak, kemiripan dan lain sebagainya. Dalam proses *clustering* terdapat beberapa metode yang biasa digunakan, diantaranya adalah *K-Medoids* dan *Fuzzy Possibilistic C-Means*. Untuk *K-Medoids* dalam proses perhitungannya akan memerlukan *distance metric* berupa jarak *euclidean* dan menggunakan metode *elbow* serta *silhouette* untuk menentukan jumlah *cluster* (c) optimal. Sedangkan, untuk *Fuzzy Possibilistic C-Means* dalam proses perhitungannya akan memerlukan *distance metric* berupa jarak *square euclidean* dan menggunakan nilai PE, PC serta MPC untuk menentukan jumlah *cluster* (c) optimal. Penelitian ini bertujuan untuk mengelompokkan tingkat kesehatan balita dari 34 provinsi di Indonesia ditinjau berdasarkan 11 indikator variabel penanganan kasus stunting dengan $c = 2, 3, \dots, 6$. Menggunakan *K-Medoids* berdasarkan metode *elbow* diperoleh bahwa c optimal adalah sebanyak 3 *cluster*, sedangkan berdasarkan metode *silhouette* diperoleh bahwa c optimal adalah sebanyak 4 *cluster*. Menggunakan *Fuzzy Possibilistic C-Means* berdasarkan nilai PE, PC dan MPC diperoleh bahwa c optimal adalah sebanyak 2 *cluster*. Berdasarkan perbandingan menggunakan nilai *davies-bouldin index* (DBI) diperoleh bahwa untuk *K-Medoids* dengan $c = 3$ nilai DBI-nya yaitu 1.662832, sedangkan untuk $c = 4$ nilai DBI-nya yaitu 1.265332. Dengan menggunakan *Fuzzy Possibilistic C-Means* dengan $c = 2$ diperoleh nilai DBI-nya yaitu 1.762145. Jadi, pengelompokan tingkat kesehatan balita berdasarkan penanganan kasus stunting di Indonesia akan lebih baik menggunakan *K-Medoids* dengan jumlah *cluster* (c) optimal ditentukan berdasarkan metode *silhouette* dimana $c = 4$.

Kata Kunci: *Clustering, K-Medoids, Fuzzy Possibilistic C-Means, Davies-Bouldin Index, Stunting*

ABSTRACT

ADAM RAMADHAN. Comparison of K-Medoids and Fuzzy Possibilistic C-Means Methods in Classifying the Health Level of Toddlers Based on the Handling of Stunting Cases in Indonesia. Mini Thesis, Department of Mathematics, Faculty of Mathematics and Natural Sciences, State University of Jakarta. June 2025.

Clustering is a data grouping technique that is basically performed based on certain similarities or characteristics, such as distance, similarity, and so on. In the clustering process, there are several methods that are commonly used, including K-Medoids and Fuzzy Possibilistic C-Means. For K-Medoids, the calculation process requires an Euclidean distance metric and uses the elbow and silhouette methods to determine the optimal number of clusters (c). Meanwhile, for Fuzzy Possibilistic C-Means, the calculation process requires a square Euclidean distance metric and uses the PE, PC, and MPC values to determine the optimal number of clusters (c). This study aims to group the health levels of toddlers from 34 provinces in Indonesia based on 11 indicators of stunting case management with $c = 2, 3, \dots, 6$. Using K-Medoids based on the elbow method, it was found that the optimal c is 3 clusters, while based on the silhouette method, it was found that the optimal c is 4 clusters. Using Fuzzy Possibilistic C-Means based on the PE, PC, and MPC values, it was found that the optimal c is 2 clusters. Based on a comparison using the Davies-Bouldin Index (DBI), it was found that for K-Medoids with $c = 3$, the DBI value is 1.662832, while for $c = 4$, the DBI value is 1.265332. Using Fuzzy Possibilistic C-Means with $c = 2$, the DBI value is 1.762145. Therefore, clustering the health status of infants based on stunting case management in Indonesia is better achieved using K-Medoids with the optimal number of clusters (c) determined using the silhouette method, where $c = 4$.

Keywords: *Clustering, K-Medoids, Fuzzy Possibilistic C-Means, Davies-Bouldin Index, Stunting*

KATA PENGANTAR

Pertama-tama puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas berkat segala rahmat dan hidayah-Nya membuat penulis dapat menyelesaikan penyusunan skripsi ini dengan judul "Perbandingan Metode *K-Medoids* dan *Fuzzy Possibilistic C-Means* dalam Mengelompokkan Tingkat Kesehatan Balita Berdasarkan Penanganan Kasus Stunting di Indonesia" dengan baik sebagai salah satu syarat untuk memperoleh gelar Sarjana Matematika selama studi di program studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta.

Skripsi ini dapat terselesaikan tentunya tidak terlepas dari doa, arahan, bimbingan, dukungan dan bantuan dari berbagai pihak. Untuk itu pada kesempatan ini penulis ingin mengucapkan terima kasih kepada:

1. Ibu Dr. Eti Dwi Wiraningsih, M.Si. selaku dosen pembimbing I dan Bapak Dr. Yudi Mahatma, M.Si. selaku dosen pembimbing II atas segala bimbingan, saran dan masukan yang senantiasa diberikan kepada penulis sehingga skripsi ini dapat diselesaikan dengan baik.
2. Bapak/Ibu dosen program studi Matematika yang telah mengajarkan penulis dan memberikan ilmu pengetahuan yang sangat bermanfaat dalam proses penyusunan skripsi ini.
3. Kedua orang tua dan keluarga penulis yang selalu mendoakan serta memberikan dukungan selama penyusunan skripsi dengan penuh rasa kasih sayang.
4. Seluruh teman-teman dan semua pihak lainnya yang tidak tercantumkan, terima kasih atas segala doa dan bantuan yang telah diberikan.

Semoga Tuhan Yang Maha Esa memberikan anugerah dan rahmat-Nya kepada semua pihak yang telah memberikan bantuan dalam penyelesaian skripsi ini. Besar harapan penulis bahwa informasi pada skripsi ini dapat memberikan kebermanfaatan bagi lingkungan akademik program studi

Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta, maupun pihak lain yang membutuhkan serta juga bagi penulis secara pribadi.

Tidak lupa penulis juga menyadari bahwa pembuatan skripsi ini tidak akan bisa lepas dari berbagai kekurangan dan jauh dari kesempurnaan, oleh karena itu penulis mengharapkan kritik dan saran yang membangun dari pembaca demi perbaikan di masa mendatang.

Jakarta, 29 Juni 2025

Adam Ramadhan



DAFTAR ISI

HALAMAN PERSETUJUAN	i
HALAMAN PERNYATAAN	ii
HALAMAN PUBLIKASI	iii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR	vi
DAFTAR ISI	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	7
1.3 Tujuan Penelitian	7
1.4 Batasan Penelitian	8
1.5 Manfaat Penelitian	9
BAB II KAJIAN PUSTAKA	10
2.1 <i>Data Mining</i>	10

2.2	<i>Clustering</i>	11
2.3	<i>Preprocessing Data</i>	14
2.3.1	Deteksi <i>Missing Data</i>	14
2.3.2	Normalisasi Data	16
2.3.3	Deteksi Multikolinearitas	19
2.4	Ruang Vektor <i>Euclidean</i>	21
2.4.1	Norma Vektor dan Jarak di $\mathbb{R}^n$	21
2.4.2	Jarak <i>Euclidean</i>	23
2.4.3	Jarak <i>Square Euclidean</i>	25
2.5	Metode <i>K-Medoids</i>	25
2.6	Metode <i>Fuzzy Possibilistic C-Means</i>	30
2.7	Penentuan Jumlah <i>Cluster (c)</i> Optimal untuk Metode <i>K-Medoids</i>	39
2.7.1	Metode <i>Elbow</i>	39
2.7.2	Metode <i>Silhouette</i>	43
2.8	Penentuan Jumlah <i>Cluster (c)</i> Optimal untuk Metode <i>Fuzzy Possibilistic C-Means</i>	49
2.8.1	<i>Partition Entropy (PE)</i>	49
2.8.2	<i>Partition Coefficient (PC)</i>	50
2.8.3	<i>Modified Partition Coefficient (MPC)</i>	52
2.9	<i>Davies-Bouldin Index (DBI)</i>	53
2.10	Tingkat Kesehatan Balita	59
2.11	Indeks Khusus Penanganan Stunting (IKPS)	60
BAB III	METODOLOGI PENELITIAN	67
3.1	Data dan Sumber Data	67
3.2	Variabel Penelitian	67
3.3	Struktur Data Penelitian	68
3.4	Langkah Penelitian	69
BAB IV	HASIL DAN PEMBAHASAN	71
4.1	<i>Preprocessing Data</i>	71
4.2	Analisis Statistik Deskriptif	76

4.3	<i>Clustering</i> dengan Metode <i>K-Medoids</i>	77
4.3.1	Perhitungan Jarak <i>Euclidean</i>	77
4.3.2	<i>K-Medoids Clustering</i>	79
4.4	Jumlah <i>Cluster (c)</i> Optimal untuk Metode <i>K-Medoids</i>	84
4.4.1	Perhitungan dengan Metode <i>Elbow</i>	84
4.4.2	Perhitungan dengan Metode <i>Silhouette</i>	88
4.5	<i>Clustering</i> dengan Metode <i>Fuzzy Possibilistic C-Means</i>	96
4.5.1	Perhitungan Jarak <i>Square Euclidean</i>	96
4.5.2	<i>Fuzzy Possibilistic C-Means Clustering</i>	97
4.6	Jumlah <i>Cluster (c)</i> Optimal untuk Metode <i>Fuzzy Possibilistic C-Means</i>	107
4.6.1	Perhitungan berdasarkan Nilai <i>Partition Entropy (PE)</i>	108
4.6.2	Perhitungan berdasarkan Nilai <i>Partition Coefficient (PC)</i>	109
4.6.3	Perhitungan berdasarkan Nilai <i>Modified Partition Co-efficient (MPC)</i>	110
4.7	Interpretasi Hasil <i>Clustering</i> menggunakan Metode <i>K-Medoids</i>	111
4.8	Interpretasi Hasil <i>Clustering</i> menggunakan Metode <i>Fuzzy Possibilistic C-Means</i>	117
4.9	Perbandingan Hasil <i>Clustering</i> berdasarkan Nilai <i>Davies-Bouldin index</i>	119
BAB V	KESIMPULAN DAN SARAN	131
5.1	Kesimpulan	131
5.2	Saran	133
	DAFTAR PUSTAKA	135
	LAMPIRAN	141
	RIWAYAT HIDUP	152

DAFTAR TABEL

Tabel 2.1	Tabel <i>missing</i> data	15
Tabel 2.2	Tabel data	18
Tabel 2.3	Tabel data setelah dinormalisasi	19
Tabel 2.4	Tabel pusat <i>cluster</i>	24
Tabel 2.5	Tabel hasil jarak <i>euclidean</i>	25
Tabel 2.6	Tabel pusat <i>cluster</i>	28
Tabel 2.7	Tabel total jarak	28
Tabel 2.8	Tabel pusat <i>cluster</i>	28
Tabel 2.9	Tabel total jarak	29
Tabel 2.10	Tabel hasil <i>clustering</i>	30
Tabel 2.11	Tabel hasil <i>clustering</i>	39
Tabel 2.12	Tabel nilai SSE	41
Tabel 2.13	Tabel nilai besar sudut	42
Tabel 2.14	Tabel nilai SC	48
Tabel 2.15	Tabel nilai PE	50
Tabel 2.16	Tabel nilai PE dan PC	52
Tabel 2.17	Tabel nilai PE, PC dan MPC	53
Tabel 2.18	Tabel hasil <i>clustering Fuzzy Possibilistic C-Means</i>	56
Tabel 2.19	Tabel pusat <i>cluster Fuzzy Possibilistic C-Means</i>	56
Tabel 2.20	Tabel nilai DBI	58
Tabel 3.1	Tabel variabel penelitian	68
Tabel 3.2	Tabel struktur data penelitian	68
Tabel 4.1	Tabel data	71
Tabel 4.2	Tabel pendeteksian <i>missing</i> data	72

Tabel 4.3	Tabel nilai VIF	75
Tabel 4.4	Tabel statistik deskriptif	76
Tabel 4.5	Tabel hasil jarak <i>euclidean</i>	79
Tabel 4.6	Tabel pusat <i>cluster</i>	80
Tabel 4.7	Tabel total jarak	80
Tabel 4.8	Tabel pusat <i>cluster</i>	81
Tabel 4.9	Tabel total jarak	82
Tabel 4.10	Tabel hasil <i>clustering</i>	83
Tabel 4.11	Tabel hasil <i>clustering</i>	85
Tabel 4.12	Tabel pusat <i>cluster</i>	85
Tabel 4.13	Tabel nilai SSE	86
Tabel 4.14	Tabel nilai besar sudut	88
Tabel 4.15	Tabel hasil <i>clustering</i>	89
Tabel 4.16	Tabel nilai SC	95
Tabel 4.17	Tabel hasil <i>clustering</i>	106
Tabel 4.18	Tabel hasil <i>clustering</i>	106
Tabel 4.19	Tabel nilai PE	109
Tabel 4.20	Tabel nilai PE dan PC	110
Tabel 4.21	Tabel nilai PE, PC dan MPC	111
Tabel 4.22	Tabel <i>clustering</i> dengan $c = 3$ berdasarkan metode <i>elbow</i> . . .	113
Tabel 4.23	Tabel <i>clustering</i> dengan $c = 4$ berdasarkan metode <i>silhouette</i> .	115
Tabel 4.24	Tabel <i>clustering</i> dengan $c = 2$ berdasarkan nilai PE, PC dan MPC	117
Tabel 4.25	Tabel hasil <i>clustering K-Medoids</i> berdasarkan metode <i>elbow</i> . .	119
Tabel 4.26	Tabel pusat <i>cluster K-Medoids</i> berdasarkan metode <i>elbow</i> . . .	119
Tabel 4.27	Tabel hasil <i>clustering K-Medoids</i> berdasarkan metode <i>silhouette</i>	120
Tabel 4.28	Tabel pusat <i>cluster K-Medoids</i> berdasarkan metode <i>silhouette</i> .	120
Tabel 4.29	Tabel hasil <i>clustering FPCM</i> berdasarkan nilai PE, PC dan MPC	120
Tabel 4.30	Tabel pusat <i>cluster FPCM</i> berdasarkan nilai PE, PC dan MPC	121
Tabel 4.31	Tabel nilai DBI	127
Tabel 4.32	Hasil akhir <i>clustering</i> dan interpretasinya	127
Tabel 4.33	Hasil akhir <i>clustering</i> dan interpretasinya	128

DAFTAR GAMBAR

Gambar 2.1 Visual nilai multikolinearitas variabel	21
Gambar 2.2 Grafik metode <i>elbow</i>	41
Gambar 2.3 Besar sudut pada grafik metode <i>elbow</i>	43
Gambar 3.1 Diagram alir	70
Gambar 4.1 Visual nilai multikolinearitas variabel	75
Gambar 4.2 Grafik metode <i>elbow</i>	86
Gambar 4.3 Besar sudut pada grafik metode <i>elbow</i>	88

DAFTAR LAMPIRAN

Lampiran 1:	Data Penelitian	141
Lampiran 2:	Statistik Deskriptif	142
Lampiran 3:	Tabel Hasil Jarak <i>Euclidean</i> untuk $c = 2$	142
Lampiran 4:	Tabel Total Jarak untuk $c = 2$ ($v_1 =$ Riau; $v_2 =$ Sulawesi Tengah)	143
Lampiran 5:	Tabel Total Jarak untuk $c = 2$ ($v_1 =$ Aceh; $v_2 =$ Sumatera Utara)	144
Lampiran 6:	Tabel Hasil <i>Clustering</i> metode <i>K-Medoids</i> untuk $c = 2$	145
Lampiran 7:	Hasil Jarak Euclidean dari Setiap Data	146
Lampiran 8:	Tabel Hasil <i>Clustering</i> metode <i>K-Medoids</i> untuk $c = 3$	147
Lampiran 9:	Tabel Hasil <i>Clustering</i> metode <i>K-Medoids</i> untuk $c = 4$	148
Lampiran 10:	Tabel Hasil <i>Clustering</i> metode FPCM untuk $c = 2$	149
Lampiran 11:	Hasil Akhir Matriks Kekhasan Relatif (U) dan Matriks Kekhasan Absolut (T)	150
Lampiran 12:	<i>Syntax Code</i> yang digunakan pada <i>R Studio</i>	151