



SKRIPSI

**KLASTERISASI TINGKAT KEMISKINAN
MENGUNAKAN *FUZZY C-MEANS* DENGAN
OPTIMALISASI *ENTROPY WEIGHT* DAN PCA**

JIHAN OCTAVIA SALSABILLAH HIDAYAT
NPM 21081010034

DOSEN PEMBIMBING

Dr. Rr. Ani Dijah Rahajoe, ST, M.Cs.

Muhammad Muharrom Al Haromainy, S.Kom., M.Kom.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2025**



SKRIPSI

KLASTERISASI TINGKAT KEMISKINAN MENGUNAKAN *FUZZY C-MEANS* DENGAN OPTIMALISASI *ENTROPY WEIGHT* DAN PCA

JIHAN OCTAVIA SALSABILLAH HIDAYAT
NPM 21081010034

DOSEN PEMBIMBING

Dr. Rr. Ani Dijah Rahajoe, ST, M.Cs.

Muhammad Muharrom Al Haromainy, S.Kom., M.Kom.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2025**

Halaman ini sengaja dikosongkan

LEMBAR PENGESAHAN

KLASTERISASI TINGKAT KEMISKINAN MENGUNAKAN *FUZZY C-MEANS* DENGAN OPTIMALISASI *ENTROPY WEIGHT* DAN PCA

Oleh :

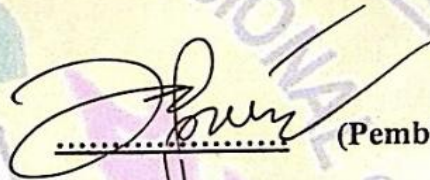
JIHAN OCTAVIA SALSABILLAH HIDAYAT

NPM. 21081010034

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal 7 Maret 2025

Dr. Rr. Ani Dijah Rahajoe, ST, M.Cs.

NIP. 19730512 200501 2003

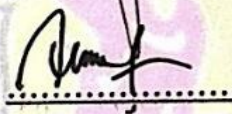


(Pembimbing I)

Muhammad Muharrom Al Haromainy,

S.Kom., M.Kom.

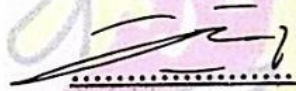
NIP. 19950601 202203 1 006



(Pembimbing II)

Fawwaz Ali Akbar, S.Kom, M.Kom

NIP. 19920317 201803 1 002



(Ketua Penguji)

Firza Prima Aditiawan, S.Kom., MTI

NIP. 19860523 202121 1 003



(Anggota Penguji)

Mengetahui,

Dekan Fakultas Ilmu Komputer



Prof. Dr. Ir. Novirina Hendrasarie, MT.

NIP. 19681126 199403 2 001

Halaman ini sengaja dikosongkan

LEMBAR PERSETUJUAN

KLASTERISASI TINGKAT KEMISKINAN MENGUNAKAN *FUZZY C-MEANS* DENGAN OPTIMALISASI ENTROPY WEIGHT DAN PCA

Oleh:

JIHAN OCTAVIA SALSABILLAH HIDAYAT
NPM. 21081010034

Menyetujui,

**Koordinator Program Studi Informatika
Fakultas Ilmu Komputer**



Fetty Tri Anggraeny, S.Kom., M.Kom.
NIP. 19820211 2021212 005

Halaman ini sengaja dikosongkan

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan dibawah ini :

Nama : Jihan Octavia Salsabillah Hidayat
NPM : 21081010034
Program : Sarjana(S1)
Program Studi : Informatika
Fakultas : Fakultas Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila dikemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.



Surabaya, 23 Mei 2025
Yang Membuat pernyataan



Jihan Octavia Salsabillah Hidayat
NPM. 21081010034

Halaman ini sengaja dikosongkan

ABSTRAK

Nama Mahasiswa / NPM : Jihan Octavia Salsabillah Hidayat / 21081010034
Judul Skripsi : Klasterisasi Tingkat Kemiskinan Menggunakan *Fuzzy C-Means* Dengan Optimalisasi *Entropy weight* Dan PCA
Dosen Pembimbing : 1. Dr. Rr. Ani Dijah Rahajoe, ST, M.Cs.
2. Muhammad Muharrom Al Haromainy, S.Kom., M.Kom.

Kemiskinan merupakan salah satu permasalahan utama yang dihadapi oleh Provinsi Jawa Timur, dengan tingkat kemiskinan tertinggi ketiga di Indonesia. Untuk memahami pola dan karakteristik kemiskinan di wilayah ini, diperlukan metode analisis yang mampu mengelompokkan daerah berdasarkan indikator sosial-ekonomi secara lebih akurat. Penelitian ini bertujuan untuk melakukan *clustering* tingkat kemiskinan di kabupaten/kota di Jawa Timur dengan menggunakan metode *Fuzzy C-Means* dengan optimalisasi *entropy weight* dan *Principal Component Analysis* (PCA). Proses analisis diawali dengan seleksi fitur menggunakan metode *entropy weight*, yang menentukan bobot objektif dari setiap indikator kemiskinan berdasarkan variasinya. Dari 14 variabel awal, diperoleh 6 variabel utama, yaitu Jumlah Penduduk Miskin, Persentase Penduduk Miskin, P1 (Indeks Kedalaman Kemiskinan), P2 (Indeks Keparahan Kemiskinan), Garis Kemiskinan, dan Pengeluaran per Kapita. Selanjutnya, dilakukan reduksi dimensi dengan PCA, yang menghasilkan dua komponen utama (PC1 dan PC2) yang mampu menjelaskan 86,01% dari total informasi dalam dataset. Tahap *clustering* menggunakan *Fuzzy C-Means* (FCM) menentukan jumlah *cluster* optimal berdasarkan *silhouette score*, dengan hasil 3 *cluster* dan nilai *silhouette score* sebesar 0,553. *cluster* 1 terdiri dari 10 kabupaten/kota dengan tingkat kemiskinan terendah, *cluster* 2 mencakup 21 kabupaten/kota dengan ekonomi menengah, dan *cluster* 3 terdiri dari 7 Kabupaten/Kota dengan tingkat kemiskinan tertinggi. Evaluasi model menggunakan *silhouette score* dengan hasil 0,55, yang menunjukkan bahwa hasil *clustering* telah layak atau sesuai, dengan *cluster* 1 memiliki distribusi keanggotaan yang lebih seragam, sedangkan *cluster* 2 dan 3 memiliki penyebaran yang lebih luas dalam sosial-ekonomi.

Kata kunci: *Fuzzy C-Means, Entropy weight, Principal Component Analysis, Klasterisasi, Kemiskinan, Jawa Timur.*

ABSTRACT

Student Name / NPM : Jihan Octavia Salsabillah Hidayat / 21081010034
Thesis Title : *Poverty Level Clustering Using Fuzzy C-Means with Entropy weight Optimization and PCA*
Advisor : 1. Dr. Rr. Ani Dijah Rahajoe, ST, M.Cs.
2. Muhammad Muharrom Al Haromainy, S.Kom., M.Kom.

Poverty is one of the major issues faced by East Java Province, ranking as the third-highest in Indonesia in terms of poverty rate. To understand the patterns and characteristics of poverty in this region, an analytical method capable of accurately clustering areas based on socioeconomic indicators is required. This study aims to cluster the poverty levels in districts/cities in East Java using the Fuzzy C-Means (FCM), optimized with Entropy weight and Principal Component Analysis (PCA).

The analysis process begins with feature selection using the entropy weight method, which determines the objective weight of each poverty indicator based on its variation. Out of 14 initial variables, six key variables were selected: Number of Poor Population, Percentage of Poor Population, P1 (Poverty Depth Index), P2 (Poverty Severity Index), Poverty Line, and Per Capita Expenditure. Subsequently, dimensionality reduction using PCA was performed, resulting in two main components (PC1 and PC2) that explain 86.01% of the total information in the dataset.

The clustering stage using Fuzzy C-Means (FCM) determines the optimal number of clusters based on the Silhouette score, with results showing three clusters and a Silhouette score of 0.55. Cluster 1 consists of 10 districts/cities with the lowest poverty rate, Cluster 2 includes 21 districts/cities with a middle-income economy, and Cluster 3 consists of 7 districts/cities with the highest poverty rate.

The model evaluation using Silhouette score resulted in a score of 0.55, indicating that the clustering results are adequate and appropriate. Cluster 1 has a more compact membership distribution, while clusters 2 and 3 show a broader socioeconomic spread.

Keywords: *Fuzzy C-Means, Entropy weight, Principal Component Analysis, Clustering, Poverty, East Java.*

KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas segala rahmat, hidayah dan karunia-Nya kepada penulis sehingga skripsi dengan judul **“Klasterisasi Tingkat Kemiskinan Menggunakan *Fuzzy C-Means* Dengan Optimalisasi *Entropy Weight* Dan PCA”** dapat terselesaikan dengan baik.

Penelitian ini disusun sebagai salah satu persyaratan kelulusan dalam menyelesaikan program Strata Satu (S1) di Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur.

Penulis menyadari bahwa penyusunan skripsi ini tidak terlepas dari dukungan berbagai pihak yang telah memberikan bimbingan, motivasi, serta bantuan, sehingga laporan ini dapat diselesaikan dengan baik. Oleh karena itu penulis mengucapkan terima kasih kepada:

1. Mama Farida Yuliatin dan Papa Rahmad Hidayat selaku orang tua penulis, yang selalu mengusahakan segalanya demi keberhasilan penulis, baik dalam pendidikan maupun kehidupan. Doa, kasih sayang, serta pengorbanan yang tiada henti menjadi kekuatan terbesar bagi penulis dalam menghadapi setiap tantangan. Tidak hanya sebagai orang tua, tetapi juga sebagai sumber motivasi, tempat bersandar, serta alasan utama bagi penulis untuk terus berjuang dan meraih impian. Kehadiran serta dukungan tanpa syarat dari kedua orang tua merupakan anugerah terbesar dalam hidup, menjadikan setiap pencapaian terasa lebih bermakna dan penuh kebahagiaan. Terima kasih atas setiap doa yang dipanjatkan, setiap usaha yang dilakukan, dan setiap kasih sayang yang diberikan tanpa batas. Segala pencapaian ini tidak akan terwujud tanpa cinta dan dukungan yang tulus dari mama dan papa.
2. Ibu Prof. Dr. Ir. Novirina Hendrasarie, MT selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Ibu Fetty Try Anggraeny, S.Kom, M.Kom selaku Ketua Program Studi Informatika Fakultas Ilmu Sosial Dan Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. Ibu Dr. Rr. Ani Dijah Rahajoe, ST, M.Cs. dan Bapak Muhammad Muharrom Al Haromainy, S.Kom., M.Kom. selaku Dosen Pembimbing 1 dan 2 penulis.
5. Muhammad Hafiddien Rois Hidayat selaku kakak penulis, yang telah menjadi bagian dari perjalanan hidup. Terima kasih atas setiap interaksi dan pelajaran yang pernah ada. Meskipun tidak semua hal berjalan dengan mudah, kehadiranmu tetap

memiliki peran dalam membentuk sebagian dari pandangan serta pengalaman hidup penulis hingga saat ini. Semoga apa yang pernah dijalani dapat menjadi pembelajaran berharga bagi kita. Penulis berharap yang terbaik untuk setiap langkah hidup ke depannya.

6. Sabrina Rahma Maulida Hidayat selaku adik penulis, yang selalu menjadi tempat berbagi cerita dalam segala keadaan, baik dalam kebahagiaan maupun kesedihan. Kehadirannya membawa kebahagiaan dan ketenangan, menjadikan setiap momen bersama penuh makna. Terima kasih atas setiap pelukan, tawa, serta dukungan tulus yang selalu diberikan. Memiliki adik sepertinya adalah anugerah yang sangat berharga, yang membuat perjalanan hidup semakin berarti dan penuh warna. Semangat dan kasih sayangnnya mendorong penulis untuk terus berusaha menjadi kakak yang lebih baik.
7. Cintya Juanita Elizabeth Kolimon dan Fioletta Augustine Rumaseb selaku sahabat yang tidak hanya menemani perjalanan akademik, tetapi juga kehidupan penulis secara keseluruhan. Persahabatan yang terjalin telah menjadi bagian penting dalam hidup penulis, menghadirkan kebahagiaan melalui tawa, dukungan, dan kebersamaan dalam setiap momen. Keberadaan sahabat yang selalu memahami dan menerima penulis apa adanya adalah salah satu hal yang paling disyukuri. Setiap cerita yang dibagikan, setiap pelukan di saat sulit, serta setiap kenangan indah yang tercipta merupakan bagian dari kebahagiaan yang tak ternilai. Terima kasih atas setiap momen kebersamaan, mulai dari tawa dalam suka hingga bahu yang selalu siap menopang dalam duka. Persahabatan ini mengajarkan bahwa keluarga tidak selalu terbentuk dari hubungan darah, tetapi juga dari ikatan yang kuat di hati.
8. Mochammad Fahmi Yasir selaku seseorang yang telah mendukung perjalanan akademis dan pribadi penulis dari awal kuliah hingga saat ini. Terima kasih atas kesabaran dalam menghadapi penulis, atas setiap nasihat yang diberikan saat penulis merasa bimbang, serta atas setiap usaha, baik kecil maupun besar, yang dilakukan demi kebahagiaan penulis. Dukungan dan kehadirannya telah menjadi bagian penting dalam perjalanan ini, menunjukkan bahwa kasih sayang dan kepedulian mampu memberikan kekuatan. Memiliki seseorang yang selalu percaya pada kemampuan penulis serta tak henti mendorong untuk meraih impian adalah anugerah yang sangat berharga.

9. Treasure, terima kasih telah menjadi teman yang menyenangkan selama perjalanan hidup penulis. Lagu dan konten yang mereka hadirkan membantu penulis untuk tetap semangat, terutama saat merasa lelah dan jenuh. Kehadiran dan energi positif mereka menjadi pengingat bahwa tidak apa merasa lelah, selama tidak berhenti. Terima kasih sudah ikut menemani hari-hari penuh perjuangan dengan cara yang sederhana tapi berdampak besar.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih terdapat berbagai kekurangan. Oleh karena itu, kritik dan saran yang membangun dari berbagai pihak sangat penulis harapkan demi perbaikan dan penyempurnaan di masa mendatang. Penulis berharap laporan ini dapat memberikan manfaat, baik bagi pembaca secara umum maupun bagi penulis.

Surabaya, 4 Maret 2025

Jihan Octavia Salsabillah Hidayat

DAFTAR ISI

LEMBAR PENGESAHAN	i
LEMBAR PERSETUJUAN	iii
SURAT PERNYATAAN BEBAS PLAGIASI.....	v
ABSTRAK	vii
KATA PENGANTAR.....	ix
DAFTAR ISI.....	xii
DAFTAR GAMBAR.....	xiv
DAFTAR TABEL	xv
DAFTAR KODE	xvi
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	5
1.3. Tujuan Penelitian.....	6
1.4. Manfaat Penelitian.....	6
1.5. Batasan Masalah.....	7
BAB II TINJAUAN PUSTAKA	8
2.1. Penelitian Terdahulu	8
2.2. Kemiskinan Provinsi Jawa Timur	12
2.3. <i>Data Mining</i>	12
2.4. Normalisasi Data	13
2.5. <i>Entropy weight</i>	14
2.6. <i>Principal Component Analysis (PCA)</i>	15
2.7. <i>Clustering</i>	17
2.8. Logika Fuzzy	17
2.9. <i>Fuzzy C-Means</i>	18
5.1 <i>Silhouette score</i>	19
BAB III METODOLOGI	21
3.1. Tahap Penelitian	21
3.2. Studi Literatur	21
3.3. Pengumpulan Data	22
3.4. <i>Preprocessing Data</i>	23

3.4.1	<i>Data Cleaning</i>	23
3.4.2	Normalisasi Data	24
3.4.3	<i>Feature Selection</i>	26
3.4.4	<i>Dimensionality Reduction</i>	28
3.5.	<i>Fuzzy C-Means</i>	31
3.6.	Evaluasi Model	32
BAB IV HASIL DAN PEMBAHASAN		34
4.1	Pengumpulan Data	34
4.2	<i>Library</i>	35
4.3	<i>Preprocessing Data</i>	37
4.3.1	<i>Data Cleaning</i>	37
4.3.2	Normalisasi Data	38
4.3.3	<i>Feature Selection</i>	39
4.3.4	<i>Dimensionality Reduction</i>	41
4.4	Penentuan Jumlah <i>Cluster</i>	44
4.5	<i>Fuzzy C-Means</i>	46
4.5.1	Parameter dalam <i>Fuzzy C-Means</i>	47
4.5.2	Matriks Partisi Awal	47
4.5.3	Pusat <i>Cluster</i> ke-i	48
4.5.4	Fungsi Objektif	49
4.5.5	Perubahan Matriks Partisi	50
4.5.6	Pusat <i>Cluster</i> Akhir	51
4.5.7	Menambahkan Kolom Kab/Kota dalam Matriks Keanggotaan	51
4.5.8	Karakteristik <i>Cluster</i>	52
4.6	Evaluasi Model	55
4.7	Hasil <i>Clustering</i>	57
4.8	Antarmuka Pengguna (GUI)	62
BAB V KESIMPULAN DAN SARAN		68
5.1	Kesimpulan	68
5.2	Saran	69
DAFTAR PUSTAKA		70

DAFTAR GAMBAR

Gambar 2. 1 Tahap penyelesaian <i>Data Mining</i>	13
Gambar 3. 1 Tahap Penelitian.....	21
Gambar 3. 2 Diagram Tahapan <i>Preprocessing Data</i>	23
Gambar 3. 3 Alur <i>Feature Selection</i>	26
Gambar 3. 4 Alur <i>Dimensionality Reduction</i>	28
Gambar 3. 5 Alur <i>Fuzzy C-Means</i>	31
Gambar 4. 1 Dataset Kemiskinan Jawa Timur.....	34
Gambar 4. 2 Dataset setelah <i>Data Cleaning</i>	38
Gambar 4. 3 Dataset setelah Normalisasi	39
Gambar 4. 4 Hasil <i>Feature Selection</i>	41
Gambar 4. 5 Hasil Perhitungan Eigenvalue	42
Gambar 4. 6 Hasil Perhitungan Eigenvector.....	43
Gambar 4. 7 Variabel yang paling Berpengaruh.....	44
Gambar 4. 8 Grafik <i>Silhouette score</i>	46
Gambar 4. 9 Jumlah <i>Cluster</i> Optimal	46
Gambar 4. 10 Hasil Fungsi Objektif	50
Gambar 4. 11 Hasil Karakteristik Setiap <i>Cluster</i>	54
Gambar 4. 12 <i>Heatmap</i> Karakteristik Setiap <i>Cluster</i>	54
Gambar 4. 13 Hasil Evaluasi Model	55
Gambar 4. 14 Visualisasi <i>Silhouette score</i>	57
Gambar 4. 15 Hasil <i>Clustering</i> dengan PC	60
Gambar 4. 16 <i>Scatter Plot Clustering</i>	62
Gambar 4. 17 Tampilan Awal GUI.....	63
Gambar 4. 18 Tampilan Setelah Upload Data	64
Gambar 4. 19 Tampilan Akhir GUI	65

DAFTAR TABEL

Tabel 2. 1 Review Peneliti Terdahulu.....	8
Tabel 3. 1 Contoh Data Kemiskinan	22
Tabel 3. 2 Tabel kriteria <i>Silhouette Coefficient</i>	33
Tabel 4. 1 Parameter dan Deskripsi dari Dataset	35
Tabel 4. 2 Bobot Setiap Variabel	41
Tabel 4. 3 Nilai Silhouette untuk Setiap Nilai Jumlah K <i>Cluster</i>	45
Tabel 4. 4 Matriks Partisi Awal	48
Tabel 4. 5 Pusat <i>Cluster</i> ke-i	49
Tabel 4. 6 Perubahan Matriks Partisi	50
Tabel 4. 7 Pusat <i>Cluster</i> Akhir	51
Tabel 4. 8 Hasil <i>Clustering</i>	58

DAFTAR KODE

Kode 4. 1 <i>Library</i> yang digunakan	35
Kode 4. 2 <i>Feature Selection</i>	39
Kode 4. 3 Perhitungan Eigenvalue	41
Kode 4. 4 Perhitungan Eigenvector	42
Kode 4. 5 Menentukan Variabel yang paling Berpengaruh	43
Kode 4. 6 Menentukan Jumlah <i>Cluster</i> Optimal.....	44
Kode 4. 7 Pemilihan Jumlah <i>Cluster</i> Optimal	46
Kode 4. 8 Parameter <i>Fuzzy C-Means</i>	47
Kode 4. 9 Matriks Partisi Awal.....	48
Kode 4. 10 Pusat <i>Cluster</i> ke-i	49
Kode 4. 11 Fungsi Objektif	49
Kode 4. 12 Total Iterasi dan Perubahan Matriks Partisi	50
Kode 4. 13 Pusat <i>Cluster</i> Akhir	51
Kode 4. 14 Menyimpan Hasil <i>Clustering</i>	51
Kode 4. 15 Penggabungan Derajat Keanggotaan.....	51
Kode 4. 16 Karakteristik Setiap <i>Cluster</i>	52
Kode 4. 17 Evaluasi Model.....	55
Kode 4. 18 Hasil di simpan ke Excel	58
Kode 4. 19 Hasil dengan <i>Principal Components</i> disimpan ke Excel	60
Kode 4. 20 Visualisasi <i>Scatter Plot</i>	61
Kode 4. 21 Tampilan GUI.....	65
Kode 4. 22 Upload <i>File Excel</i>	66
Kode 4. 23 Proses Hasil <i>Clustering</i>	66
Kode 4. 24 Menampilkan Tabel.....	66