



SKRIPSI

IMPLEMENTASI *SPECTRAL CLUSTERING* DENGAN OPTIMASI PSO UNTUK PENGELOMPOKAN WILAYAH BERDASARKAN TINGKAT KEMISKINAN DI JAWA TIMUR

RAHMALIA ANINDYA HERDIANTI
NPM 21083010084

DOSEN PEMBIMBING

Wahyu Syaifullah Jauharis Saputra., S.Kom., M.Kom.
Dr. Eng. Ir. Dwi Arman Prasetya, S.T., M.T., IPU., Asean, Eng.

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



SKRIPSI

**IMPLEMENTASI *SPECTRAL CLUSTERING*
DENGAN OPTIMASI PSO UNTUK
PENGELOMPOKAN WILAYAH BERDASARKAN
TINGKAT KEMISKINAN DI JAWA TIMUR**

RAHMALIA ANINDYA HERDIANTI
NPM 21083010084

DOSEN PEMBIMBING

Wahyu Syaifullah Jauharis Saputra., S.Kom., M.Kom.
Dr. Eng. Ir. Dwi Arman Prasetya, S.T., M.T., IPU., Asean, Eng.

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



SKRIPSI

IMPLEMENTASI *SPECTRAL CLUSTERING* DENGAN OPTIMASI PSO UNTUK PENGELOMPOKAN WILAYAH BERDASARKAN TINGKAT KEMISKINAN DI JAWA TIMUR

RAHMALIA ANINDYA HERDIANTI
NPM 21083010084

DOSEN PEMBIMBING

Wahyu Syaifullah Jauharis Saputra., S.Kom., M.Kom.
Dr. Eng. Ir. Dwi Arman Prasetya, S.T., M.T., IPU., Asean, Eng.

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

LEMBAR PENGESAHAN

IMPLEMENTASI *SPECTRAL CLUSTERING* DENGAN OPTIMASI PSO UNTUK PENGELOMPOKAN WILAYAH BERDASARKAN TINGKAT KEMISKINAN DI JAWA TIMUR

Oleh:
RAHMALIA ANINDYA HERDIANTI
NPM. 21083010084

Telah dipertahankan di hadapan dan diterima oleh Tim Penguji Sidang Skripsi Program Studi Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal 16 Juni 2025:

Menyetujui,

Wahyu Syaifullah Jauharis S., S.Kom.,
M.Kom.

NIP. 19860825 202121 1 003

Dr. Eng. Ir. Dwi Arman Prasetya, S.T.,
M.T., IPU., Asean, Eng.

NIP. 19801205 200501 1 002

Dr. Ir. Mohammad Idhom, S.P., S.Kom
M.T.

NIP. 19830310 202121 1 006

Amri Muhaimin, S.Stat., M.Stat., M.S.

NIP. 19950723 202406 1 002

(Pembimbing I)

(Pembimbing II)

(Ketua Penguji)

(Penguji I)

Mengetahui,

Dekan Fakultas Ilmu Komputer

Prof. Dr. Ir. Novirina Hendrasarie, MT.
NIP. 19681126 199403 2 001

LEMBAR PERSETUJUAN

IMPLEMENTASI *SPECTRAL CLUSTERING* DENGAN OPTIMASI PSO UNTUK PENGELOMPOKAN WILAYAH BERDASARKAN TINGKAT KEMISKINAN DI JAWA TIMUR

Oleh:
RAHMALIA ANINDYA HERDIANTI
NPM. 21083010084

Telah disetujui untuk mengikuti Ujian Skripsi

Menyetujui,

**Koordinator Program Studi Sains Data
Fakultas Ilmu Komputer**

Dr. Eng. Ir. Dwi Arman Prasetya, S.T., M.T., IPU., Asean, Eng.
NIP. 19801205 200501 1 002

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan di bawah ini:

Nama Mahasiswa : Rahmalia Anindya Herdianti
NPM : 21083010084
Program : Sarjana (S1)
Program Studi : Sains Data
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 di kemudian 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, 16 Juni 2025
Yang Membuat Pernyataan,


RAHMALIA ANINDYA HERDIANTI
NPM. 21083010084

ABSTRAK

Nama Mahasiswa / NPM : Rahmalia Anindya Herdianti / 21083010084
Judul Skripsi : Implementasi *Spectral Clustering* dengan Optimasi PSO untuk Pengelompokan Wilayah Berdasarkan Tingkat Kemiskinan di Jawa Timur
Dosen Pembimbing : 1. Wahyu Syaifullah Jauharis Saputra., S.Kom., M.Kom.
2. Dr. Eng. Ir. Dwi Arman Prasetya, S.T., M.T., IPU., Asean, Eng.

Tingkat kemiskinan yang tinggi di Jawa Timur merupakan masalah yang memerlukan perhatian khusus. Tujuan dari penelitian ini adalah melakukan pemetaan tingkat kemiskinan pada setiap kabupaten/kota di Jawa Timur menggunakan algoritma *Spectral Clustering* dengan optimasi *Particle Swarm Optimization* (PSO). *Spectral Clustering* digunakan untuk mengelompokkan kabupaten/kota, sementara PSO untuk melakukan peningkatan terhadap nilai parameter sehingga akurasi pada *clustering* mengalami peningkatan. Jumlah penduduk miskin, garis kemiskinan, rata – rata lama sekolah, dan persentase penduduk miskin merupakan beberapa faktor utama yang mempengaruhi tingkat kemiskinan dalam penelitian ini. Dari hasil penelitian dapat diidentifikasi terdapat dua *cluster* yaitu, wilayah dengan tingkat kemiskinan tinggi (31 kabupaten/kota) dan tingkat kemiskinan rendah (7 kabupaten/kota). Evaluasi terhadap kualitas *clustering* menunjukkan bahwa nilai *Davies-Bouldin Index* (DBI) mengalami penurunan dari 0.4401 menjadi 0.0943, sedangkan nilai *Silhouette Score* mengalami peningkatan dari 0.6655 menjadi 0.9315 setelah dilakukan optimasi. Hal ini mengindikasikan peningkatan signifikan dalam kualitas pemisahan dan kohesi *cluster* yang terbentuk. Penelitian ini diharapkan dapat memberikan wawasan yang lebih dalam tentang pola kemiskinan di Jawa Timur dan menjadi dasar untuk kebijakan pemerintah yang lebih tepat dan efektif untuk melakukan penanganan pada permasalahan kemiskinan.

Kata kunci : Jawa Timur, Kemiskinan, Pengelompokan, Particle Swarm Optimization (PSO), *Spectral Clustering*.

Halaman ini sengaja dikosongkan

ABSTRACT

Student Name / NPM : Rahmalia Anindya Herdianti / 21083010084
Thesis Title : *Implementation of Spectral Clustering with PSO Optimization for Regional Grouping Based on Poverty Level in East Java*
Advisor : 1. Wahyu Syaifullah Jauharis Saputra., S.Kom., M.Kom.
2. Dr. Eng. Ir. Dwi Arman Prasetya, S.T., M.T., IPU., Asean, Eng.

ABSTRACT

The high poverty rate in East Java is a problem that requires special attention. The purpose of this research is to map the poverty level in each district/city in East Java using the Spectral Clustering algorithm with Particle Swarm Optimization (PSO) optimization. Spectral Clustering is used to group districts/municipalities, while PSO is used to increase the parameter values so that the accuracy of the clustering increases. The number of poor people, poverty line, average years of schooling, and percentage of poor people are some of the main factors that affect the poverty rate in this study. From the results of the research, two clusters can be identified, namely, areas with high poverty rates (31 districts/municipalities) and low poverty rates (7 districts/municipalities). Evaluation of the clustering quality shows that the Davies-Bouldin Index (DBI) value decreased from 0.4401 to 0.0943, while the Silhouette Score value increased from 0.6655 to 0.9315 after optimization. This indicates a significant improvement in the quality of separation and cohesion of the clusters formed. This research is expected to provide a deeper insight into the pattern of poverty in East Java and become the basis for more appropriate and effective government policies to deal with the problem of poverty.

Keywords: *East Java, Poverty, Clustering, Particle Swarm Optimization (PSO), Spectral Clustering.*

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Puji syukur kehadirat Allah SWT atas segala rahmat, hidayah dan karunia-Nya kepada penulis sehingga skripsi dengan judul **“Implementasi *Spectral Clustering* dengan Optimasi PSO untuk Pengelompokan Wilayah Berdasarkan Tingkat Kemiskinan di Jawa Timur”** dapat terselesaikan dengan baik.

Penulis mengucapkan terima kasih kepada Bapak Wahyu Syaifullah Jauharis Saputra., S.Kom., M.Kom. selaku Dosen Pembimbing utama yang telah meluangkan waktunya untuk memberikan bimbingan, nasehat serta motivasi kepada penulis. Dan penulis juga banyak menerima bantuan dari berbagai pihak, baik itu berupa moril, spiritual maupun materiil. Untuk itu penulis mengucapkan terima kasih kepada:

1. Ibu Prof. Dr. Ir. Novirina Hendrasarie, MT selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Bapak Dr. Eng. Ir. Dwi Arman Prasetya, S.T., M.T., IPU., Asean, Eng. selaku Ketua Program Studi Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Seluruh dosen Program Studi Sains Data yang telah memberikan ilmu dan pengalamannya kepada penulis.
4. Kedua orang tua, saudara-saudara, dan seluruh keluarga yang senantiasa memberikan dukungan moril dan doa.
5. Semua pihak yang telah membantu penulis dalam menyelesaikan skripsi ini.

Penulis menyadari bahwa di dalam penyusunan skripsi ini banyak terdapat kekurangan. Untuk itu kritik dan saran yang membangun dari semua pihak sangat diharapkan demi kesempurnaan penulisan skripsi ini. Akhirnya, dengan segala keterbatasan yang penulis miliki semoga laporan ini dapat bermanfaat bagi semua pihak umumnya dan penulis pada khususnya.

Surabaya, 16 Juni 2025

Penulis

Halaman ini sengaja dikosongkan

DAFTAR ISI

LEMBAR PENGESAHAN	ii
LEMBAR PERSETUJUAN	iii
SURAT PERNYATAAN BEBAS PLAGIASI.....	iv
ABSTRAK	v
ABSTRACT	vii
KATA PENGANTAR.....	ix
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL	xv
DAFTAR LAMPIRAN	xvii
DAFTAR NOTASI.....	xix
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Batasan Masalah	4
1.4. Tujuan Penelitian	4
1.5. Manfaat Penelitian	4
BAB II TINJAUAN PUSTAKA	5
2.1. Penelitian Terdahulu	5
2.2. Dasar Teori.....	8
2.2.1 Kemiskinan	9
2.2.2 Data Mining	9
2.2.3 Robust Scaler	10
2.2.4 Metode <i>Clustering</i>	10
2.2.5 Algoritma <i>Spectral Clustering</i>	11
2.2.6 Metode Optimasi.....	13
2.2.7 <i>Particle Swarm Optimization</i> (PSO).....	13
2.2.8 Evaluasi <i>Davies-Bouldin Index</i>	16
2.2.9 Evaluasi <i>Silhouette Score</i>	17
2.2.10 Aplikasi <i>Grafical User Interface</i> (GUI)	18

BAB III.....	33
DESAIN DAN IMPLEMENTASI SISTEM.....	33
3.1 Variabel Penelitian dan Sumber Data	33
3.1.1. Variabel Penelitian.....	33
3.1.2. Sumber Data.....	34
3.2. Langkah Analisis	35
3.2.1. Pengumpulan Data	36
3.2.2. Eksplorasi Data Analisis	37
3.2.3. Preprocessing Data.....	37
3.2.4. Evaluasi Jumlah <i>Cluster</i> Optimal	37
3.2.5. Penerapan <i>Spectral Clustering</i>	38
3.2.6. Penerapan Optimasi Parameter Menggunakan PSO.....	39
3.2.7. Pemetaan Hasil Cluster	39
3.2.8. Analisis Hasil	39
3.2.9. Pembuatan Antarmuka Pengguna (<i>Graphical User Interface</i>).....	40
3.3. Desain Sistem.....	40
BAB IV HASIL DAN PEMBAHASAN	33
4.1. Pengumpulan Data	33
4.2. <i>Preprocessing</i> Data.....	33
4.3. Eksplorasi Data Analisis	34
4.4. Evaluasi Jumlah Cluster Optimal.....	44
4.5. Penerapan <i>Spectral Clustering</i>	47
4.6. Penerapan Optimasi Parameter pada <i>Spectral Clustering</i>	51
4.7. Pemetaan Hasil Klaster	63
4.8. Analisis Hasil	66
4.9. Implementasi GUI Berdasarkan Hasil Pengelompokan Menggunakan <i>Spectral Clustering</i> dengan Optimasi PSO.....	70
BAB V PENUTUP.....	77
5.1. Kesimpulan	77
5.2. Saran	78
DAFTAR PUSTAKA.....	79
LAMPIRAN.....	83

DAFTAR GAMBAR

Gambar 1.1 Persentase Penduduk Miskin di Jawa Timur 2013-2023	1
Gambar 2.1 Ilustrasi Persebaran Metode Cluster	11
Gambar 2.2 Ilustrasi Persebaran Particle Swarm Optimization	13
Gambar 3.1 Diagram Alir Langkah Analisis Penelitian.....	35
Gambar 3.2 Diagram Alir Langkah Analisis Penggunaan GUI.....	41
Gambar 3.3 Tampilan Desain Sistem GUI.....	42
Gambar 4.1 Histogram distribusi variabel Persentase Penduduk Miskin (%)	35
Gambar 4.2 Histogram distribusi variabel Jumlah Penduduk Miskin (%).....	36
Gambar 4.3 Histogram distribusi variabel Harapan Lama Sekolah	37
Gambar 4.4 Histogram distribusi variabel Rata-Rata Lama Sekolah.....	37
Gambar 4.5 Histogram distribusi variabel Tingkat Pengangguran Terbuka (%).....	38
Gambar 4.6 Histogram distribusi variabel Tingkat Partisipasi Angkatan Kerja (%).....	39
Gambar 4.7 Histogram distribusi variabel Angka Harapan Hidup	39
Gambar 4.8 Histogram distribusi variabel Garis Kemiskinan.....	40
Gambar 4.9 Histogram distribusi variabel Indeks Pembangunan Manusia	41
Gambar 4.10 Histogram distribusi variabel Rata-Rata Upah Pekerja Berdasarkan Lapangan Pekerjaan Utama	41
Gambar 4.11 Histogram distribusi variabel Rata – Rata Pendapatan Pekerja Berdasarkan Pendidikan Tertinggi - Jumlah.....	42
Gambar 4.12 Visualisasi Heatmap Korelasi Fitur.....	44
Gambar 4.13 Visualisasi Evaluasi Jumlah Kluster Optimal.....	46
Gambar 4.14 Visualisasi Persebaran Kluster <i>Spectral Clustering</i>	51
Gambar 4.15 Visualisasi Persebaran Kluster <i>Spectral Clustering</i> dengan Optimasi PSO	60
Gambar 4.16 Diagram Importance Variabel dalam Analisis <i>Clustering</i> Kemiskinan di Jawa Timur	69
Gambar 4.17 Visualisasi Persebaran Kluster Berdasarkan 2 Variabel.....	70
Gambar 4.18 Tampilan Awal Antar Muka Pengguna Grafis	71
Gambar 4.19 Tampilan Menu Upload Data oleh Pengguna.....	72
Gambar 4.20 Tampilan Sistem Setelah Pengguna Melakukan Input Data	72
Gambar 4.21 Antarmuka Sistem Saat Melakukan EDA	73

Gambar 4.22 Penerapan Metode RobustScaler pada Dataset	74
Gambar 4.23 Evaluasi Jumlah Klaster Optimal Menggunakan Silhouette Score dan Davies-Bouldin Index (DBI).....	74
Gambar 4.24 Penerapan PSO untuk Menentukan Nilai Gamma Optimal pada <i>Spectral Clustering</i>	75
Gambar 4.25 Tampilan Antarmuka GUI pada Tahapan Akhir sebagai Rangkuman Hasil Analisis dan Pengelompokan	75

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	5
Tabel 2.2 Parameter <i>Spectral Clustering</i>	13
Tabel 3.1 Variabel Penelitian	33
Tabel 3.2 Dataset Penelitian	34
Tabel 4.1 Dataset Indikator Kemiskinan	33
Tabel 4.2 Hasil Scaling Dataset	34
Tabel 4.3 Statistika Deskriptif.....	43
Tabel 4.4 Lanjutan Statistika Deskriptif.....	43
Tabel 4.5 Evaluasi Jumlah Kluster Optimal.....	45
Tabel 4.6 Rentang Interpretasi Silhouette Score	46
Tabel 4.7 Evaluasi Nilai Gamma	47
Tabel 4.8 Hasil Metrik Evaluasi <i>Spectral Clustering</i>	50
Tabel 4.9 Percobaan 2 Sample Optimasi Manual	54
Tabel 4.10 Fitur 2 Sample Optimasi Manual	55
Tabel 4.11 Hasil Metrik Evaluasi <i>Spectral Clustering</i> dengan Optimasi PSO.....	58
Tabel 4.12 Eigenvector Cluster	61
Tabel 4.13 Hasil Perbandingan Metrik Evaluasi <i>Spectral Clustering</i> dan <i>Spectral Clustering</i> dengan Optimasi PSO	62
Tabel 4.14 Hasil <i>Clustering</i> Kabupaten/Kota di Jawa Timur Menggunakan Metode <i>Spectral Clustering</i> dengan Optimasi PSO.....	64
Tabel 4.15 Lanjutan Hasil <i>Clustering</i> Kabupaten/Kota di Jawa Timur Menggunakan Metode <i>Spectral Clustering</i> dengan Optimasi PSO.....	65
Tabel 4.16 Perbandingan Rata-Rata Variabel Berdasarkan Cluster.....	66
Tabel 4.17 Bobot Importance Variabel dalam Analisis <i>Clustering</i> Kemiskinan di Jawa Timur.....	68

Halaman ini sengaja dikosongkan

DAFTAR LAMPIRAN

Lampiran 1. Data Indikator Kemiskinan.....	83
Lampiran 2. Source Code Model	84
Lampiran 3. Source Code GUI Streamlit.....	85
Lampiran 4. Letter of Acceptance (LOA) Jurnal INSTEK.....	86

Halaman ini sengaja dikosongkan

DAFTAR NOTASI

$s(x_i, x_j)$	:	Similarity/ Kemiripan antara dua poin data
$\ x_i - x_j\ ^2$	:	Kuadrat dari jarak Euclidean
σ^2	:	Variansi dari distribusi
W	:	Matriks Similarity
d_i	:	Degree dari titik data ke-i
L_{sym}	:	Matriks simetris yang ternormalisasi
$D^{-\frac{1}{2}}$	:	Invers dari akar kuadrat matriks degree
I	:	Matriks identitas
X_i	:	Nilai data individu ke-i
n	:	Jumlah total data/ Ukuran sampel
$d(x, y)$	:	Jarak Euclidean antara dua titik x dan y
C_{kj}	:	<i>Rata-rata himpunan data indeks k dan j</i>
$\sum_{i=1}^p X_{ij}$	:	Penjumlahan dari p nilai
p	:	Jumlah sampel pada satu kelompok
V^k	:	Kecepatan dari partikel pada iterasi ke-k
ω	:	Faktor inersia
P_{pbest}^{k-1}	:	Posisi terbaik yang pernah dicapai partikel
P_{gbest}^{k-1}	:	Posisi terbaik yang pernah dicapai seluruh populasi partikel
P^k	:	Posisi baru partikel pada iterasi ke-k
v	:	Nilai rata-rata dari sekumpulan data
DBI	:	Mengukur kualitas <i>clustering</i> berdasarkan hubungan antara kohesi (kedekatan data dalam cluster) dan separasi (jarak antar cluster).
$R_{i,j}$	:	Rasio antara jumlah dispersi (kepadatan) dalam cluster i dan j dengan jarak antara centroid dari cluster i dan j.
S_i	:	Silhouette Coefficient untuk objek i