

**ANALISIS *CLUSTERING* BERDASARKAN
FREKUENSI KEJADIAN DAN DAMPAK
BENCANA HIDROMETEOROLOGI
DI INDONESIA MENGGUNAKAN
SELF ORGANIZING MAPS DAN K-MEDOIDS**

Skripsi

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



**Gloria Kristin Natalia Sitanggang
1314622001**

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

2026

LEMBAR PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi saya dengan judul **"Analisis *Clustering* Berdasarkan Frekuensi Kejadian Dan Dampak Bencana Hidrometeorologi Di Indonesia Menggunakan *Self Organizing Maps* Dan *K-Medoids*"** yang disusun sebagai syarat untuk memperoleh gelar Sarjana Statistika dari Program Studi Statistika Universitas Negeri Jakarta adalah karya ilmiah saya dengan arahan dosen pembimbing.

Sumber informasi yang diperoleh dari penulisan lain yang telah dipublikasikan yang disebutkan dalam teks ini, telah dicantumkan dalam Daftar Pustaka sesuai dengan norma, kaidah, dan etika penulisan ilmiah.

Jika dikemudian hari ditemukan sebagian besar skripsi ini bukan hasil karya saya sendiri, saya bersedia menerima sanksi pencabutan gelar akademik yang saya sanding dan sanksi-sanksi lainnya sesuai dengan peraturan perundang-undangan yang berlaku.

Jakarta, 17 Juli 2026



Gloria Kristin Natalia Sitanggang

ABSTRAK

GLORIA KRISTIN NATALIA SITANGGANG. Analisis *Clustering* Berdasarkan Frekuensi Kejadian Dan Dampak Bencana Hidrometeorologi Di Indonesia Menggunakan *Self Organizing Maps* Dan K-Medoids. Skripsi, Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta. Juli 2026.

Bencana hidrometeorologi merupakan jenis bencana yang paling dominan terjadi di Indonesia dan menimbulkan dampak yang berbeda pada setiap provinsi. Perbedaan karakteristik tersebut menunjukkan perlunya pengelompokan provinsi agar pola kemiripan berdasarkan frekuensi kejadian dan dampak bencana dapat diidentifikasi. Penelitian ini bertujuan untuk mengidentifikasi karakteristik, pola kemiripan, serta mengelompokkan provinsi di Indonesia berdasarkan frekuensi kejadian dan dampak bencana hidrometeorologi menggunakan pendekatan *two-stage clustering* melalui *Self Organizing Maps* (SOM) dan K-Medoids. Data yang digunakan merupakan data sekunder Badan Nasional Penanggulangan Bencana (BNPB) tahun 2025 yang mencakup 38 provinsi dengan tujuh variabel, yaitu jumlah kejadian, jumlah fasilitas umum rusak, jumlah rumah terendam, jumlah rumah rusak, jumlah orang terluka, jumlah meninggal, dan jumlah orang hilang. Hasil penelitian menunjukkan bahwa visualisasi SOM mampu menggambarkan karakteristik dan pola kemiripan antarprovinsi berdasarkan kombinasi ketujuh variabel tersebut. Selanjutnya, pendekatan *two-stage clustering* menghasilkan empat kelompok provinsi dengan karakteristik frekuensi kejadian dan dampak bencana hidrometeorologi yang berbeda. Hasil pengelompokan ini diharapkan dapat menjadi salah satu referensi dalam memahami karakteristik bencana hidrometeorologi antarprovinsi di Indonesia.

Kata kunci: *Self Organizing Maps*, K-Medoids, *Two-stage Clustering*, Bencana Hidrometeorologi, *Clustering*.

ABSTRACT

GLORIA KRISTIN NATALIA SITANGGANG. *Clustering Analysis Based on the Frequency of Occurrence and Impact of Hydrometeorological Disasters in Indonesia Using Self Organizing Maps and K-Medoids. Thesis, Statistics Study Program, Faculty of Mathematics and Natural Sciences, Jakarta State University. July 2026.*

Hydrometeorological disasters are the most dominant type of disaster in Indonesia and have different impacts on each province. These differences in characteristics indicate the need for provincial groupings to identify similarity patterns based on the frequency of occurrence and disaster impacts. This study aims to identify the characteristics, similarity patterns, and group provinces in Indonesia based on the frequency of occurrence and impact of hydrometeorological disasters using a two-stage clustering approach through Self Organizing Maps (SOM) and K-Medoids. The data used is secondary data from the National Disaster Management Agency (BNPB) in 2025, covering 38 provinces with seven variables: number of incidents, number of damaged public facilities, number of submerged houses, number of damaged houses, number of injured people, number of deaths, and number of missing people. The results of the study indicate that the SOM visualization is able to describe the characteristics and similarity patterns between provinces based on the combination of these seven variables. Furthermore, the two-stage clustering approach produces four groups of provinces with different characteristics of the frequency of occurrence and impact of hydrometeorological disasters. The results of this grouping are expected to serve as a reference in understanding the characteristics of hydrometeorological disasters between provinces in Indonesia.

Keywords: *Self Organizing Maps, K-Medoids, Two-stage Clustering, Hydrometeorological Disasters, Clustering.*

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas kelimpahan kasih karunia, penyertaan, dan hikmat-Nya yang tak terbatas. Hanya oleh anugerah dan pertolongan-Nya, penulis dianugerahi kekuatan serta keteguhan hati untuk menyelesaikan skripsi yang berjudul “Analisis *Clustering* Berdasarkan Frekuensi Kejadian dan Dampak Bencana Hidrometeorologi di Indonesia Menggunakan *Self Organizing Maps* dan K-Medoids”. Karya ilmiah ini disusun sebagai salah satu syarat untuk meraih gelar Sarjana Statistika pada Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta.

Terselesaikannya skripsi ini tidak lepas dari bantuan, dukungan, serta bimbingan dari berbagai pihak. Oleh karena itu, dengan kerendahan hati penulis menyampaikan terima kasih dan apresiasi yang sebesar-besarnya kepada:

1. Kedua Orang Tua, Abang, dan Adik Penulis, serta Keluarga Besar, atas doa yang tak putus, kasih sayang, serta dukungan moril maupun materil yang luar biasa sepanjang perjalanan akademik penulis.
2. Prof. Dr. Dian Handayani, M.Si. selaku Koordinator dan Dosen Pembimbing Akademik Prodi Statistika FMIPA Universitas Negeri Jakarta, atas ilmu, dukungan, serta bimbingan akademis yang berharga dari masa perkuliahan hingga kelulusan.
3. Prof. Dr. Ir. Bagus Sumargo, M.Si. selaku Dosen Pembimbing I, atas bimbingan, arahan, dan waktu yang senantiasa diluangkan dalam mengoreksi serta membimbing jalannya penyusunan skripsi ini dengan penuh kesabaran dan ketulusan.
4. Ibu Faroh Ladayya, M.Si. selaku Dosen Pembimbing II, atas bimbingan, arahan, motivasi, serta diskusi yang sangat membantu penulis dalam menyelesaikan seluruh rangkaian tugas akhir ini.
5. Seluruh Bapak/Ibu Dosen dan Staf Administrasi Prodi Statistika FMIPA Universitas Negeri Jakarta, atas ilmu pengetahuan dan bantuan administratif yang diberikan selama masa perkuliahan.

6. Zain Sausan Alayya selaku sahabat sekaligus saudara angkat sejak masa mahasiswa baru, Agustina Dwi Wulandari selaku sahabat di prodi sekaligus partner berkembang, serta seluruh Dhasana Artara's People, atas kebersamaan, solidaritas, dan dukungan yang diberikan sejak awal perkuliahan.
7. Zakiyah Rodja selaku sahabat sejak SMP, serta Bela Fitriani dan Kak Inge Veronica selaku sahabat Keluarga Mahasiswa Katolik (KMK), atas doa, dukungan, dan persahabatan tulus yang selalu ada menguatkan penulis dalam berbagai situasi.
8. Rekan-rekan di Tempat Magang (Bokem 20): Kak Nadiya, Naia, Kak Shafa, Kak Syifa 1, Kak Tania, Kak Bestari, Kak Syifa 2, Anya, Ali, Kian, Azri, Reyhan, Yudha, Arif, Kak Achi, Alsa, Ahnaf, Ikhsan, dan Ratih, atas pengalaman berharga, kerja sama luar biasa, serta keceriaan selama masa magang.
9. Miu dan Theta, beserta seluruh anaknya yang merupakan kucing kesayangan penulis atas kesetiaannya menemani dan menjadi pereda stres terbaik sepanjang penyusunan skripsi ini.
10. Diri Penulis Sendiri (Gloria), terima kasih atas setiap usaha, air mata, dan keluh kesah yang berhasil dilalui dengan luar biasa. Terima kasih telah membuktikan bahwa meski harus berjuang membagi waktu dengan bekerja, kamu mampu mengalahkan rasa lelah, tetap bertahan, bahkan terus mengeksplorasi potensi diri hingga titik ini. Kamu hebat dan teramat berharga.

Penulis menyadari bahwa penulisan skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, penulis terbuka menerima segala masukan, kritik, dan saran yang membangun demi pengembangan pengetahuan di masa depan. Semoga skripsi ini memberikan manfaat bagi pembaca.

Jakarta, 17 Juli 2026



Gloria Kristin Natalia Sitanggang

DAFTAR ISI

ABSTRAK.....	i
ABSTRACT.....	ii
KATA PENGANTAR	iii
DAFTAR ISI.....	v
DAFTAR TABEL.....	ix
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN.....	xi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	8
1.3 Batasan Masalah	8
1.4 Tujuan Penelitian	9
1.5 Manfaat Penelitian	9
BAB II KAJIAN PUSTAKA.....	11
2.1 Bencana Hidrometeorologi	11
2.2 Dampak Bencana	12
2.3 Machine Learning	12
2.3.1 Supervised Learning	13
2.3.2 Unsupervised Learning	14
2.3.3 Reinforcement Learning	15
2.4 Analisis Cluster.....	15
2.5 Artificial Neural Network.....	16
2.6 Competitive Learning.....	21
2.7 Self Organizing Maps (SOM)	24
2.8 Normalisasi Data.....	27
2.9 Hyperparamater SOM	28
2.10 Algoritma SOM.....	32
2.10.1 Competitive Process	33
2.10.2 Cooperative Process	35

2.10.3	<i>Adaptive Process</i>	37
2.10.4	Nilai <i>Learning Rate</i>	38
2.10.5	Proses Iteratif dan Konvergensi	38
2.11	Evaluasi Model SOM	39
2.11.1	<i>Quantization Error (QE)</i>	39
2.11.2	<i>Topographic Error (TE)</i>	41
2.12	K-Medoids	42
2.13	Evaluasi <i>Clustering</i>	48
2.13.1	<i>Silhouette Coefficient</i>	48
2.13.2	<i>Davies Bouldin Index</i>	49
2.13.3	<i>Dunn Index</i>	50
2.14	Visualisasi SOM	51
2.14.1	<i>Counts Plot</i>	52
2.14.2	<i>U-Matrix</i>	53
2.14.3	<i>Component Planes</i>	54
2.14.4	<i>Codes Plot</i>	56
2.14.5	<i>Mapping Plot</i>	57
BAB III METODE PENELITIAN		58
3.1	Data Penelitian	58
3.2	Variabel Penelitian	58
3.3	Tahapan Analisis Data	59
3.4	Diagram Alir	65
BAB IV HASIL DAN PEMBAHASAN		67
4.1	Analisis Data Eksploratif	67
4.1.1	Statistik Deskriptif	67
4.1.2	Distribusi Data	69
4.1.3	Analisis Korelasi Antarvariabel	72
4.1.4	Normalisasi Data Menggunakan Standardisasi <i>Z-score</i>	74
4.2	Penentuan <i>Hyperparameter</i> SOM	75
4.2.1	Penentuan Struktur <i>Grid</i> SOM	75
4.2.2	Penentuan Parameter Pelatihan SOM	75

4.3	Pembentukan Model SOM.....	76
4.4	Evaluasi Model SOM.....	76
4.4.1	<i>Quantization Error</i>	77
4.4.2	<i>Topographic Error</i>	77
4.4.3	Pemilihan <i>Grid</i> SOM Terbaik.....	78
4.5	Analisis Hasil SOM.....	79
4.5.1	Ilustrasi Perhitungan SOM.....	79
4.5.2	<i>Counts Plot</i>	88
4.5.3	<i>Unified Distance Matrix (U-Matrix)</i>	89
4.5.4	<i>Component Planes</i>	90
4.5.5	<i>Codes Plot</i>	91
4.5.6	<i>Mapping Plot</i>	93
4.6	Pengelompokan Neuron SOM Menggunakan K-Medoids.....	95
4.6.1	Pembentukan Matriks Jarak Antar <i>Codebook Vector</i>	95
4.6.2	Pengelompokan Neuron Menggunakan K-Medoids.....	97
4.7	Evaluasi <i>Cluster</i>	98
4.7.1	Hasil Evaluasi <i>Cluster</i>	98
4.7.2	Pemilihan Jumlah <i>Cluster</i>	99
4.8	Pembentukan <i>Cluster</i> Final.....	100
4.9	Interpretasi Karakteristik <i>Cluster</i>	102
4.9.1	Karakteristik <i>Cluster</i> 1.....	102
4.9.2	Karakteristik <i>Cluster</i> 2.....	103
4.9.3	Karakteristik <i>Cluster</i> 3.....	104
4.9.4	Karakteristik <i>Cluster</i> 4.....	106
4.9.5	Perbandingan Karakteristik Antar <i>cluster</i>	107
4.9.6	Implikasi Hasil <i>Clustering</i> terhadap Mitigasi Bencana.....	108
BAB V KESIMPULAN DAN SARAN.....		111
5.1	Kesimpulan.....	111
5.2	Saran.....	112
DAFTAR PUSTAKA.....		114
LAMPIRAN.....		118

RIWAYAT HIDUP.....	123
--------------------	-----



DAFTAR TABEL

Tabel 3. 1 Variabel Penelitian	59
Tabel 4. 1 Statistik Deskriptif Variabel Penelitian.....	67
Tabel 4. 2 Konfigurasi <i>Grid</i> SOM.....	75
Tabel 4. 3 Parameter Pelatihan SOM	75
Tabel 4. 4 Nilai <i>Quantization Error</i>	77
Tabel 4. 5 Nilai <i>Topographic error</i>	78
Tabel 4. 6 Hasil Evaluasi Model SOM.....	78
Tabel 4. 7 Bobot Awal Setiap Neuron.....	81
Tabel 4. 8 Hasil Perhitungan Jarak <i>Euclidean</i> Vektor Input	84
Tabel 4. 9 Pemetaan Provinsi pada Neuron SOM.....	94
Tabel 4. 10 Matriks Jarak Antar <i>Codebook Vector</i>	96
Tabel 4. 11 Hasil Evaluasi <i>Cluster</i> Menggunakan <i>Sillhouette Coefficient</i> , <i>Dunn Index</i> , dan <i>Davies-Bouldin Index</i>	98
Tabel 4. 12 Penentuan Keanggotaan <i>Cluster</i> Provinsi	100
Tabel 4. 13 Karakteristik Rata-rata Variabel pada <i>Cluster</i> 1	102
Tabel 4. 14 Karakteristik Rata-rata Variabel pada <i>Cluster</i> 2	103
Tabel 4. 15 Karakteristik Rata-rata Variabel pada <i>Cluster</i> 3	104
Tabel 4. 16 Karakteristik Rata-rata Variabel pada <i>Cluster</i> 4	106
Tabel 4. 17 Karakteristik Utama <i>Cluster</i>	107

DAFTAR GAMBAR

Gambar 2. 1 Jenis-jenis <i>Machine Learning</i>	13
Gambar 2. 2 Ilustrasi Neuron Biologis sebagai Dasar Pengembangan	16
Gambar 2. 3 Struktur <i>Artificial Neural Network (Multilayer Perceptron)</i>	17
Gambar 2. 4 Arsitektur <i>Competitive Learning</i>	21
Gambar 2. 5 Ilustrasi Pergerakan Bobot pada <i>Competitive Learning</i>	24
Gambar 2. 6 Arsitektur SOM	25
Gambar 2. 7 Ilustrasi Pemetaan Data ke Dalam <i>Grid</i> SOM.....	26
Gambar 2. 8 Fungsi <i>Neighborhood (Sombrero Function)</i> pada SOM	29
Gambar 2. 9 Ilustrasi Pembangunan Bobot pada SOM.....	32
Gambar 2. 10 <i>Counts Plot</i> pada SOM	52
Gambar 2. 11 <i>U-Matrix</i> pada SOM	53
Gambar 2. 12 <i>Component Planes</i> pada SOM.....	55
Gambar 2. 13 <i>Codes Plot</i> pada SOM	56
Gambar 2. 14 <i>Mapping Plot</i> pada SOM	57
Gambar 3. 1 Diagram Alir Penelitian	66
Gambar 4. 1 Histogram pada Data Variabel Penelitian.....	69
Gambar 4. 2 <i>Boxplot</i> Variabel Penelitian	70
Gambar 4. 3 Matriks Korelasi <i>Pearson</i> Antarvariabel	73
Gambar 4. 4 Perubahan <i>Learning Rate</i> Selama Proses Pelatihan SOM.....	87
Gambar 4. 5 <i>Counts Plot</i> Hasil SOM	88
Gambar 4. 6 <i>U-Matrix</i> Hasil SOM	89
Gambar 4. 7 <i>Component Plane</i> Hasil SOM	90
Gambar 4. 8 <i>Codes Plot (Fan Plot)</i> Hasil SOM.....	92
Gambar 4. 9 <i>Mapping</i> Provinsi pada Neuron SOM	93
Gambar 4. 10 Persebaran <i>Cluster</i> Provinsi di Indonesia	100
Gambar 4. 11 <i>Heatmap</i> Karakteristik Setiap <i>Cluster</i>	102

DAFTAR LAMPIRAN

Lampiran 1. Data Frekuensi Kejadian dan Dampak Bencana Hidrometeorologi Menurut Provinsi.....	118
Lampiran 2. Hasil Standardisasi <i>Z-Score</i> Data Bencana Hidrometeorologi	119
Lampiran 3. <i>Codebook Vector</i> pada Neuron yang Terisi	120
Lampiran 4. Matriks Jarak <i>Euclidean</i> Antar <i>Codebook Vector</i> pada Neuron yang Terisi	120





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

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 : Gloria Kristin Natalia Sitanggang
NIM : 1314622001
Fakultas/Prodi : Fakultas Matematika dan Ilmu Pengetahuan Alam / Statistika
Alamat email : gloriakrstnn@gmail.com

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

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

yang berjudul :

Analisis Clustering Berdasarkan Frekuensi Kejadian Dan Dampak Bencana Hidrometeorologi
Di Indonesia Menggunakan Self Organizing Maps Dan K-Medoids

Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan dan Kearsipan 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.

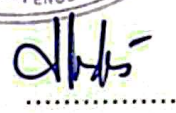
Jakarta

Penulis

(Gloria Kristin Natalia.S)
nama dan tanda tangan

LEMBAR PERSETUJUAN HASIL SIDANG SKRIPSI
ANALISIS *CLUSTERING* BERDASARKAN
FREKUENSI KEJADIAN DAN DAMPAK
BENCANA HIDROMETEOROLOGI
DI INDONESIA MENGGUNAKAN
***SELF ORGANIZING MAPS* DAN K-MEDOIDS**

Nama : Gloria Kristin Natalia Sitanggang
 NIM : 1314622001

	Nama	Tanda Tangan	Tanggal
Penanggung Jawab			
Dekan	: <u>Dr. Hadi Nasbey, S.Pd., M.Si.</u> NIP. 197909162005011004		05/08 2026
Wakil Penanggung Jawab			
Wakil Dekan I	: <u>Dr. Meiliasari, S.Pd., M.Sc.</u> NIP. 197905042009122002		05/08 2026
Ketua Penguji	: <u>Prof. Dr. Dian Handayani, M.Si.</u> NIP. 197404151998032001		27/07 2026
Sekretaris	: <u>Nilam Novita Sari, S.Stat., M.Stat.</u> NIP. 199703102024062003		28/07/2026
Anggota			
Pembimbing I	: <u>Prof. Dr. Ir. Bagus Sumargo, M.Si.</u> NIP. 196309221986011001		28/07 2026
Pembimbing II	: <u>Faroh Ladayya, M.Si.</u> NIP. 199401282020122018		27/07 2026
Penguji Ahli	: <u>Dr. Vera Maya Santi, S.Si., M.Si.</u> NIP. 197905312005012006		29/07 2026
Dinyatakan lulus ujian skripsi tanggal: 20 Juli 2026			

