

**SKRIPSI**  
**PENERAPAN ALGORITMA *SELF-ORGANIZING MAPS* (SOM)**  
**UNTUK KLASTERISASI INDIKATOR INDEKS**  
**PEMBANGUNAN MANUSIA (IPM) DI PROVINSI SULAWESI**  
**SELATAN**



**FATRISON**

**E0221526**

**PROGRAM STUDI STATISTIKA**  
**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**  
**UNIVERSITAS SULAWESI BARAT**  
**2025**

## SURAT PERNYATAAN

Yang bertanda tangan dibawa ini:

Nama : Fatrison  
Tempat/Tgl Lahir : Maranatha, 03 Maret 2003  
NIM : E0221526  
Program Studi : Statistika

Menyatakan bahwa tugas akhir dengan judul “Penerapan Algoritma *Self-Organizing Maps* (SOM) Untuk Klasterisasi Indikator Indeks Pembangunan Manusia (IPM) Di Provinsi Sulawesi Selatan” disusun berdasarkan prosedur ilmiah yang telah melalui pembimbingan dan bukan merupakan plagiat dari karya ilmiah/naskah yang lain. Apabila di kemudian hari terbukti bahwa pernyataan ini tidak benar, maka saya bersedia menerima sanksi sesuai yang berlaku.

Matene, 29 Oktober 2025

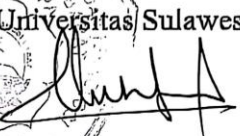


## HALAMAN PENGESAHAN

Skripsi ini diajukan oleh:

Nama : Fatrison  
NIM : E0221526  
Judul : Penerapan Algoritma *Self-Organizing Maps* (SOM)  
Untuk Klasterisasi Indikator Indeks Pembangunan  
Manusia (IPM) Di Provinsi Sulawesi Selatan.

Telah berhasil dipertahankan di depan Tim Penguji (SK Nomor: 128/UN55.7/HK.04/2025, tanggal 02 Oktober 2025) dan diterima sebagai bagian persyaratan memperoleh gelar sarjana S1 pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sulawesi Barat

Disahkan oleh:  
Dekan FMIPA  
Universitas Sulawesi Barat  
  
Musafira, S.Si., M.Sc.  
NIP. 197709112006042002

Tim Penguji:

Ketua Penguji : Musafira, S.Si., M.Sc.

Sekretaris : Putri Indi Rahayu, S.Si., M.Stat.

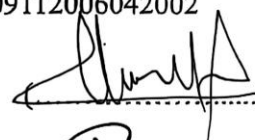
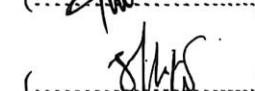
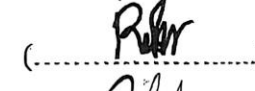
Pembimbing 1 : Muhammad Hidayatullah, S.Pd., M.Kom.

Pembimbing 2 : Putri Indi Rahayu, S.Si., M.Stat.

Penguji 1 : Reski Wahyu Yanti, S.Si., M.Si.

Penguji 2 : Afif Budi Andy B, S.Si., M.Si.

Penguji 3 : Irwan Usman, S.Pd., M.Stat.


## ABSTRAK

Indeks Pembangunan Manusia (IPM) merupakan indikator komposit yang menggambarkan kualitas pembangunan manusia melalui dimensi kesehatan, pendidikan, dan ekonomi, sehingga penting digunakan untuk menilai ketimpangan pembangunan antarwilayah. Penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di Provinsi Sulawesi Selatan berdasarkan indikator Indeks Pembangunan Manusia (IPM) menggunakan algoritma *Self-Organizing Maps* (SOM). Data yang digunakan meliputi empat indikator IPM, yaitu Umur Harapan Hidup (UHH), Rata-rata Lama Sekolah (RLS), Harapan Lama Sekolah (HLS), dan Pengeluaran per Kapita (PPK). Metode SOM diterapkan karena mampu memvisualisasikan data berdimensi tinggi ke dalam bentuk dua dimensi sehingga pola hubungan antarwilayah dapat diidentifikasi secara lebih jelas. Hasil analisis menunjukkan bahwa jumlah kluster optimal yang terbentuk adalah dua kluster, dengan nilai *Silhouette Coefficient* sebesar 0,5878 dan *Davies–Bouldin Index* (DBI) sebesar 0,6059. Kluster pertama terdiri atas Kota Makassar, Kota Parepare, dan Kota Palopo yang memiliki nilai IPM tinggi, sedangkan kluster kedua mencakup sebagian besar kabupaten dengan capaian IPM yang lebih rendah. Perbedaan utama antarwilayah terletak pada indikator Rata-rata Lama Sekolah (RLS) dan Pengeluaran per Kapita (PPK), yang menunjukkan bahwa faktor pendidikan dan kondisi ekonomi menjadi penentu utama ketimpangan pembangunan manusia di Provinsi Sulawesi Selatan.

**Kata kunci:** IPM, Klasterisasi, *Self-Organizing Maps* (SOM), Sulawesi Selatan.

## **ABSTRACT**

*The Human Development Index (HDI) is a composite indicator that illustrates the quality of human development across dimensions of health, education, and economy, making it important for assessing development disparities between regions. This study aims to group regencies/cities in South Sulawesi Province based on the Human Development Index (HDI) indicators using the Self-Organizing Maps (SOM) algorithm. The data used consist of four HDI indicators: Life Expectancy (LE), Mean Years of Schooling (MYS), Expected Years of Schooling (EYS), and Expenditure per Capita (EPC). The SOM method was applied because it can visualize high-dimensional data into a two-dimensional map, enabling clearer identification of interregional patterns. The analysis results show that the optimal number of clusters formed is two, with a Silhouette Coefficient value of 0.5878 and a Davies–Bouldin Index (DBI) value of 0.6059. The first cluster consists of Makassar City, Parepare City, and Palopo City, which have high HDI achievements, while the second cluster includes most districts with lower HDI levels. The main differences among regions lie in the Mean Years of Schooling (MYS) and Expenditure per Capita (EPC) indicators, indicating that education and economic conditions are the primary factors influencing human development disparities in South Sulawesi Province.*

**Keywords:** *Clustering, HDI, Self-Organizing Maps (SOM), South Sulawesi.*

# **BAB I**

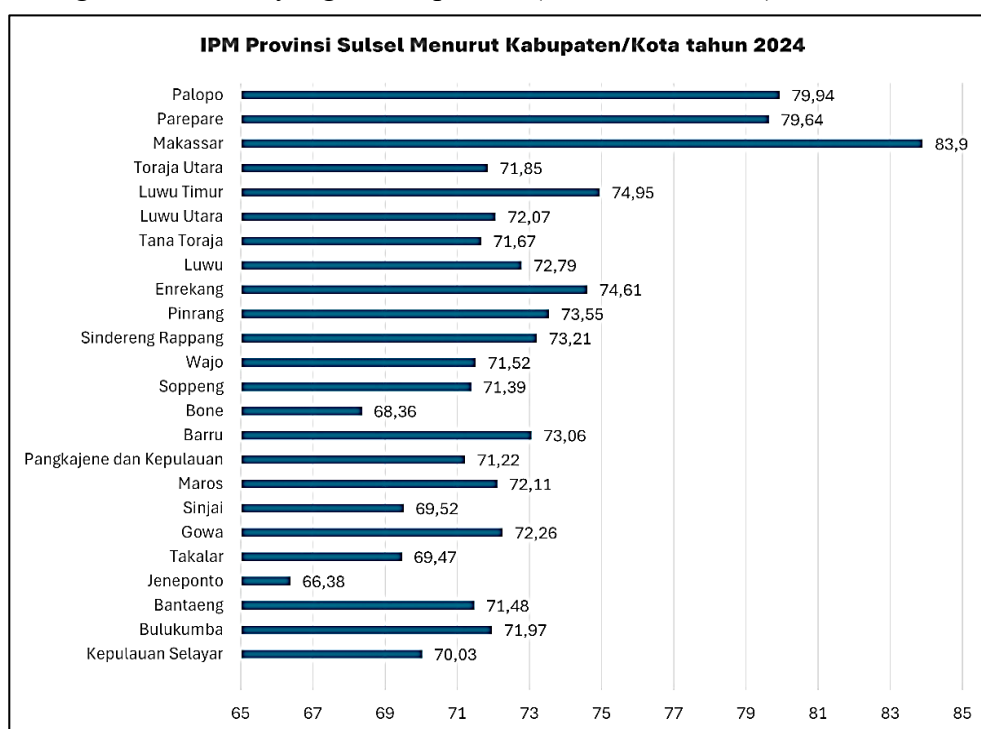
## **PENDAHULUAN**

### **1.1 Latar Belakang**

Indeks Pembangunan Manusia (IPM) adalah ukuran statistik yang digunakan untuk mengevaluasi kemajuan dan kualitas hidup manusia di suatu negara (Santoso et al., 2024). IPM digunakan untuk menilai kinerja pembangunan suatu wilayah berdasarkan tiga dimensi utama, yaitu harapan hidup dan kesehatan, pengetahuan dan pendidikan, serta Pengeluaran per Kapita yang disesuaikan. Secara geografis, Indonesia merupakan negara kepulauan yang terdiri dari berbagai provinsi yang tersebar di berbagai pulau dan dipisahkan oleh laut. Kondisi geografis ini menyebabkan tingginya disparitas pembangunan manusia antar provinsi, karena akses terhadap layanan pendidikan, kesehatan, dan ekonomi tidak merata (Hermawan & Hasugian, 2022). Sejak tahun 2016, Indonesia telah mencapai status IPM tinggi (di atas 70) dan terus mengalami peningkatan dari 72,81 pada 2020 menjadi 74,39 pada 2023 (BPS, 2024). IPM juga menunjukkan korelasi erat dengan indikator kesejahteraan lainnya seperti Pengeluaran per Kapita dan tingkat kemiskinan (Nasri et al., 2024). Meskipun disparitas antar provinsi masih tinggi, IPM nasional tetap menunjukkan peningkatan karena dihitung sebagai rata-rata dari seluruh provinsi. Dengan kata lain, meskipun terdapat daerah dengan capaian IPM rendah seperti Papua pegunungan, provinsi dengan IPM tinggi seperti DKI Jakarta dan beberapa provinsi maju lainnya berkontribusi mendorong kenaikan indeks komposit secara nasional. Oleh karena itu, analisis dan pemetaan IPM secara regional menjadi penting sebagai dasar penyusunan kebijakan pembangunan yang tepat sasaran.

Sulawesi Selatan sebagai provinsi terluas kedua di Pulau Sulawesi dengan wilayah 45.330,55 km<sup>2</sup>, terbagi dalam 21 kabupaten dan 3 kota, dengan kabupaten Luwu Utara sebagai wilayah terbesar dan Kota Parepare sebagai yang terkecil (BPS Sulsel, 2025). Keberagaman geografis dan demografis di provinsi ini berdampak langsung terhadap perbedaan capaian pembangunan manusia, akses pendidikan, layanan kesehatan, serta peluang ekonomi yang pada akhirnya mempengaruhi

capaian IPM masing-masing wilayah. Pada tahun 2023, IPM Sulawesi Selatan mencapai angka 74,60 dengan rincian angka harapan hidup sebesar 73,63 tahun, rata-rata lama sekolah 8,76 tahun, harapan lama sekolah 13,54 tahun, dan pengeluaran per Kapita sebesar Rp11,84 juta per tahun. Dalam empat tahun terakhir (2020–2023), IPM provinsi ini meningkat sebesar 1,52 poin, dengan tren percepatan pertumbuhan yang menunjukkan kemajuan menuju tingkat pembangunan manusia yang lebih optimal. (BPS Sulsel, 2024).



**Gambar 1. 1 IPM Provinsi Sulsel Menurut Kabupaten/Kota Tahun 2024**

Gambar 1.1 memperlihatkan capaian Indeks Pembangunan Manusia (IPM) di seluruh kabupaten/kota Provinsi Sulawesi Selatan pada tahun 2024. Data tersebut menunjukkan variasi yang cukup mencolok antarwilayah, yang mencerminkan tingkat pembangunan manusia yang belum merata. Kota Makassar mencatatkan nilai IPM tertinggi sebesar 83,90, diikuti oleh Kota Palopo (79,94) dan Kota Parepare (79,64). Ketiga wilayah ini merupakan kota administratif yang umumnya memiliki infrastruktur layanan dasar yang lebih baik dibandingkan kabupaten. Di sisi lain, Kabupaten Jeneponto (66,38) dan Kabupaten Bone (68,36) menunjukkan

capaian IPM terendah, mengindikasikan masih adanya keterbatasan dalam akses pendidikan, kesehatan, maupun standar hidup layak.

Kisaran IPM di Provinsi Sulawesi Selatan berada antara 66,38 hingga 83,90 dengan selisih sebesar 17,52 poin. Rentang ini mencerminkan adanya ketimpangan pembangunan manusia antarwilayah yang cukup lebar. Oleh karena itu, diperlukan strategi pembangunan yang lebih inklusif dan merata, khususnya di wilayah dengan capaian IPM yang masih rendah, agar target pembangunan manusia di tingkat provinsi dapat tercapai secara lebih adil. Namun demikian, perlu dicermati bahwa capaian IPM antar kabupaten/kota di Sulawesi Selatan masih menunjukkan perbedaan. Untuk memahami karakteristik pembangunan manusia secara lebih mendalam, dibutuhkan pendekatan analisis yang mampu mengelompokkan wilayah-wilayah berdasarkan kesamaan indikator IPM. Salah satu metode yang relevan untuk tujuan tersebut adalah teknik analisis *Clustering*.

*Clustering* merupakan suatu proses untuk membagi sekumpulan data ke dalam beberapa kelompok atau subset yang disebut sebagai klaster. Setiap objek dalam satu klaster memiliki karakteristik yang relatif serupa, namun berbeda dengan objek yang berada di klaster lainnya. Berbeda dengan klasifikasi yang membutuhkan variabel target, *Clustering* bersifat *unsupervised* karena tidak memerlukan variabel target dalam proses pengelompokannya. Salah satu metode *Clustering* yang dapat digunakan untuk memahami struktur data seperti Indeks Pembangunan Manusia (IPM) adalah algoritma *Self-Organizing Maps* (SOM), yang mampu mengelompokkan data berdasarkan kesamaan pola antar objek (Fauzi et al., 2024).

*Self-Organizing Maps* (SOM) merupakan salah satu metode yang termasuk ke dalam kategori *unsupervised learning*. Algoritma pada jenis pembelajaran ini bersifat deskriptif, yang berfungsi untuk melakukan pengelompokan atau pengkategorian terhadap data tanpa memerlukan label atau target variabel (Fauzi et al., 2024). SOM pertama kali diperkenalkan oleh *Kohonen*, yang mengembangkan metode ini sebagai bagian dari jaringan syaraf tiruan (JST) dengan pendekatan *winner-take-all*, di mana hanya neuron yang menang (paling dekat dengan *input*) yang mengalami pembaruan bobot (Imani et al., 2023). Metode SOM memiliki

keunggulan dalam memetakan data berdimensi tinggi ke dalam peta berdimensi rendah secara visual, sehingga memudahkan dalam memahami pola atau struktur data yang kompleks. Selain itu, SOM juga dikenal memiliki fleksibilitas yang tinggi terhadap proses paralelisasi, yang memungkinkan peningkatan kecepatan transfer data secara efektif (*throughput*) dalam aplikasi, serta mendukung efisiensi konsumsi daya yang rendah (Midyanti & Bahri, 2023).

Beberapa penelitian sebelumnya yang telah membahas penerapan metode SOM antara lain, penelitian yang dilakukan oleh Hastuti et al. (2023) menerapkan metode *Clustering* SOM dan *K-Affinity Propagation* (K-AP) untuk membandingkan hasil pengelompokan keduanya. Hasilnya, kedua metode membentuk tiga klaster, namun SOM menunjukkan nilai rasio yang lebih baik sehingga dinilai lebih unggul.

Penelitian oleh Zahra (2022), membandingkan metode SOM dan *Fuzzy C-Means* (FCM) dalam proses pengelompokan. Hasil pengujian menunjukkan bahwa SOM memiliki akurasi lebih tinggi, yaitu 100%, sedangkan FCM hanya mencapai 80% dengan 20% data tidak sesuai, sehingga SOM dinilai lebih efektif dalam melakukan pengelompokan. Selain itu, Iyohu et al. (2023) membandingkan SOM dengan *K-Means Clustering* dan memperoleh hasil bahwa SOM memiliki nilai DBI yang lebih baik, menunjukkan performa klasterisasi yang lebih optimal.

Berdasarkan Penelitian sebelumnya, dapat disimpulkan bahwa metode SOM terbukti efektif dalam memberikan hasil pengelompokan yang lebih optimal dibandingkan metode lain pada berbagai kasus. Dengan landasan tersebut, penelitian ini diarahkan untuk menerapkan algoritma *Self-Organizing Maps* (SOM) dalam proses klasterisasi indikator Indeks Pembangunan Manusia (IPM) di Provinsi Sulawesi Selatan, serta mengevaluasi hasil pengelompokannya untuk memperoleh struktur klaster yang optimal.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah yang akan dibahas dalam penelitian ini adalah sebagai berikut :

1. Bagaimana hasil pengelompokan kabupaten/kota di Provinsi Sulawesi Selatan berdasarkan indikator IPM, menggunakan metode *Self-Organizing Maps* (SOM)?
2. Bagaimana karakteristik klaster yang terbentuk di Provinsi Sulawesi Selatan berdasarkan indikator IPM menggunakan metode *Self-Organizing Maps* (SOM)?

## 1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan, tujuan dari penelitian ini adalah sebagai berikut :

1. Mengetahui hasil pengelompokan kabupaten/kota di Provinsi Sulawesi Selatan berdasarkan indikator IPM, menggunakan metode *Self-Organizing Maps* (SOM).
2. Mengetahui karakteristik klaster pengelompokan kabupaten/kota di Provinsi Sulawesi Selatan berdasarkan indikator IPM, menggunakan metode *Self-Organizing Maps* (SOM).

## 1.4 Batasan Masalah

1. Data yang digunakan dalam penelitian ini diperoleh dari situs resmi Badan Pusat Statistik (BPS) Sulawesi Selatan.
2. Penelitian menggunakan metode *Self-Organizing Maps* (SOM) untuk menganalisis pola pengelompokan kabupaten/kota berdasarkan indikator IPM.
3. Atribut yang digunakan dalam analisis adalah indikator IPM yang terdiri dari UHH (Umur Harapan Hidup), HLS (Harapan Lama Sekolah), RLS (Rata-rata Lama Sekolah), dan PPK (Pengeluaran per Kapita) yang disesuaikan per tahun.
4. Data yang dianalisis adalah data Indikator IPM Provinsi Sulawesi Selatan tahun 2024.

### **1.5 Manfaat Penelitian**

Penelitian ini diharapkan dapat memberikan manfaat bagi berbagai pihak, antara lain:

1. Menambah wawasan dan pemahaman tentang metode *Self-Organizing Maps* (SOM) dalam analisis data IPM.
2. Memberikan informasi mengenai pola pembangunan manusia di kabupaten/kota di Provinsi Sulawesi Selatan berdasarkan indikator IPM.
3. Mendukung pengambilan keputusan berbasis data, khususnya dalam perumusan kebijakan peningkatan IPM di daerah-daerah tertinggal.

## **BAB II**

### **KAJIAN PUSTAKA**

#### **4.1 Indeks Pembangunan Manusia (IPM)**

Indeks Pembangunan Manusia (IPM) merupakan alat ukur yang membandingkan tingkat harapan hidup, pendidikan, dan standar hidup suatu negara, termasuk Indonesia. IPM digunakan untuk mengklasifikasikan tingkat kemajuan suatu negara, apakah termasuk negara maju, berkembang, atau tertinggal, serta untuk menilai sejauh mana kemampuan ekonomi memengaruhi kualitas hidup masyarakat (Radityana et al., 2023). IPM juga merefleksikan keberhasilan pembangunan sumber daya manusia di suatu wilayah dengan tujuan menciptakan masyarakat yang hidup lebih lama, sehat, dan produktif. IPM mencakup tiga dimensi utama, yaitu kesehatan yang diukur melalui Umur Harapan Hidup (UHH), pendidikan diukur melalui Harapan Lama Sekolah (HLS) dan Rata-rata Lama Sekolah (RLS), serta ekonomi yang dilihat dari Pengeluaran per Kapita (PPK) yang disesuaikan per tahun (Irwan et al., 2024). Semakin tinggi nilai IPM, maka semakin tinggi pula pendapatan masyarakat yang diperoleh. Peningkatan ini berdampak langsung pada pengurangan angka kemiskinan, sehingga IPM menjadi indikator penting dalam menilai keberhasilan suatu daerah dalam membangun kualitas hidup penduduknya (Nasri et al., 2024).

#### **4.2 Data Mining**

Data Mining merupakan suatu proses yang digunakan untuk menemukan pola, struktur tersembunyi, atau informasi bernilai dari kumpulan data dalam jumlah besar. Proses ini dapat diterapkan pada berbagai sumber data seperti basis data, data *warehouse*, situs web, serta sistem informasi yang dinamis dan terus berkembang (Aziz & Mustakim, 2022). Tujuan utama dari Data Mining adalah mengekstraksi pengetahuan yang berguna dari data historis untuk mendukung pengambilan keputusan. Dewi dan Pramita (2019) menjelaskan bahwa melalui penerapan Data Mining, pengguna dapat mengidentifikasi tren dan pola penting yang selama ini tersembunyi, sehingga dapat dimanfaatkan untuk berbagai keperluan analisis maupun pengembangan sistem lebih lanjut.

Data *Mining* mengintegrasikan beragam teknik dari disiplin ilmu seperti statistika, matematika, kecerdasan buatan (AI), serta pembelajaran mesin (*Machine Learning*). Kamila et al. (2019) menekankan bahwa salah satu metode utama dalam Data *Mining* adalah *Clustering*, yang berfungsi untuk mengelompokkan data berdasarkan kesamaan karakteristik guna mengeksplorasi data lebih mendalam, menemukan anomali, dan menyusun prediksi. Menurut Nazeriandy et al. (2021), Data *Mining* juga dapat dipahami sebagai suatu upaya untuk menggali nilai tambah dari data yang semula tidak teridentifikasi secara manual. Pengetahuan yang dihasilkan memiliki manfaat praktis dalam berbagai bidang. Meskipun bukan bidang baru, Data *Mining* merupakan pengembangan dari ilmu statistik dan sangat erat kaitannya. Tantangan utamanya adalah sifatnya yang multidisipliner, karena menggabungkan banyak teknik dan pendekatan dari berbagai cabang ilmu.

Dalam penelitian ini, Data *Mining* digunakan untuk menganalisis faktor-faktor yang berkontribusi terhadap pembangunan manusia di wilayah Provinsi Sulawesi Selatan. Teknik ini berperan dalam mengidentifikasi pola-pola yang signifikan dari indikator IPM sehingga dapat digunakan sebagai dasar dalam pengelompokan wilayah berdasarkan karakteristik pembangunan manusianya. Data *Mining* menjadi kerangka utama dalam penelitian ini, di mana metode *Clustering* berbasis SOM diterapkan. Melalui pendekatan ini, penelitian bertujuan untuk mengidentifikasi wilayah dengan karakteristik pembangunan manusia yang serupa guna mendukung perencanaan kebijakan pembangunan yang lebih tepat sasaran dan berkeadilan.

#### 4.3 *Machine Learning*

*Machine Learning* merupakan salah satu cabang dari Data *Mining* yang terus berkembang dengan berbagai inovasi dan teknik yang dirancang untuk menyelesaikan permasalahan data yang kompleks. Dua pendekatan utama yang paling umum digunakan dalam *Machine Learning* adalah *supervised learning* dan *unsupervised learning* dimana pendekatan tersebut merupakan yang paling umum dan sangat masif penggunaannya dibandingkan pendekatan lainnya (Parulian & Butar, 2023). Sebagai bagian dari kecerdasan buatan (*Artificial Intelligence*), *Machine Learning* berfungsi meniru perilaku manusia dalam proses pengambilan

keputusan berdasarkan pengalaman dari data. Dalam pendekatan *supervised learning*, model dilatih menggunakan data yang telah diberi label, sedangkan *unsupervised learning* digunakan untuk menemukan pola dari data tanpa label (Fahmi, 2023).

Fokus utama dari penerapan *Machine Learning* adalah mengembangkan sistem yang mampu belajar secara mandiri dan beradaptasi terhadap perubahan tanpa pemrograman ulang. Dengan menganalisis data dalam jumlah besar (*big data*), *Machine Learning* dapat menemukan pola atau tren yang tersembunyi dan bernilai untuk proses prediksi maupun pengambilan keputusan otomatis (Fahmi, 2023).

#### **4.4 Analisis Clustering**

*Clustering* merupakan teknik dalam analisis data yang bertujuan untuk mengelompokkan individu atau objek ke dalam beberapa partisi berdasarkan tingkat kemiripan tertentu. Setiap objek dalam satu kluster memiliki karakteristik yang relatif homogen, sementara antar kluster bersifat heterogen. (Talakua et al., 2017). Dalam bidang *Data Mining*, *Clustering* termasuk ke dalam pendekatan *unsupervised learning* yang populer digunakan untuk mengidentifikasi pola dari data tanpa label. Objek-objek yang dikelompokkan ke dalam kluster memiliki tingkat kesamaan yang tinggi dalam satu kelompok, dan perbedaan mencolok dibandingkan dengan kelompok lainnya (Aziz & Mustakim, 2022; Alfina et al., 2012).

Berbeda dengan metode klasifikasi yang membutuhkan variabel target yang telah ditentukan sebelumnya, *Clustering* bekerja tanpa variabel target, karena prosesnya dilakukan murni berdasarkan pola yang terdapat dalam data. Artinya, data dikelompokkan murni berdasarkan pola atau kemiripan karakteristik yang ditemukan di dalamnya. Pengelompokan data dilakukan untuk memahami struktur tersembunyi dalam data IPM serta mengidentifikasi kesamaan karakteristik antar wilayah (Fauzi et al., 2024). Metode *Clustering* yang umum digunakan meliputi *K-Means*, *K-Medoids*, *Hierarchical Clustering*, dan *SOM* (Alfina et al., 2012; Dewi & Pramita, 2019).

#### 4.5 Jaringan Syaraf Tiruan

Salah satu teknik komputasi yang termasuk dalam kecerdasan buatan (*Artificial Intelligence*) adalah Jaringan Syaraf Tiruan (*Artificial Neural Network*). ANN merupakan sistem pemrosesan yang dirancang dan dilatih agar memiliki kemampuan seperti manusia dalam menyelesaikan permasalahan kompleks melalui proses belajar yang melibatkan perubahan bobot sinapsis. Jaringan ini mensimulasikan struktur dan fungsi otak biologis, sehingga mampu mengenali pola-pola yang rumit serta belajar dari pengalaman masa lalu (Ginting, 2013). Dalam bidang data *science*, ANN berperan sebagai metode analitik yang memiliki kemampuan dalam klasifikasi, asosiasi, *self-organizing*, dan optimasi (Ismail, 2021). Dalam penelitian ini, SOM dipilih karena mampu memetakan data berdimensi tinggi ke dalam bentuk visual yang lebih mudah dianalisis, serta efektif dalam mengelompokkan wilayah berdasarkan indikator pembangunan manusia (Fauzi et al., 2024).

#### 4.6 *Self-Organizing Maps* (SOM)

*Self-Organizing Maps* (SOM) adalah salah satu jenis jaringan saraf tiruan yang dikembangkan oleh *Teuvo Kohonen* pada tahun 1982. Metode ini berfungsi untuk memetakan data berdimensi tinggi ke dalam ruang dua dimensi (biasanya dalam bentuk *grid*), sehingga mempermudah proses visualisasi dan interpretasi hasil pengelompokan (Fauzi et al., 2024). SOM digunakan untuk pembelajaran mesin tanpa pengawasan (*unsupervised learning*), di mana data dikelompokkan berdasarkan kesamaan karakteristik tanpa memerlukan label. Proses kerjanya melibatkan pemilihan neuron pemenang, yakni neuron dengan bobot paling dekat terhadap *input*, diikuti dengan pembaruan bobot neuron tersebut dan neuron tetangganya berdasarkan radius jarak tertentu dalam jaringan (Syamhuri et al., 2022). *Input* dalam teknik ini berupa data hasil observasi dengan banyak dimensi, sedangkan lapisan *output* terdiri dari jaringan neuron yang memiliki bobot masing-masing dan tersusun dalam bentuk kisi dua dimensi, baik persegi maupun *hexagonal*. Berbeda dengan jaringan saraf tiruan lainnya, SOM tidak memiliki *hidden layer*. Bobot pada neuron-neuron dalam jaringan akan diperbarui selama

proses pelatihan agar semakin menyerupai pola data *input*. Namun, proses pelatihan SOM dapat memakan waktu relatif lama apabila dimensi *input* yang digunakan terlalu besar. Hasil pelatihan SOM dapat dimanfaatkan untuk berbagai tujuan, antara lain memvisualisasikan distribusi tiap vektor *input*, mengidentifikasi area topik yang paling relevan, serta membentuk kluster. Salah satu teknik visualisasi yang populer adalah *U-Matrix (Unified Distance Matrix)*, yang mampu menggambarkan keterkaitan antar vektor dan mempermudah proses pengelompokan data (Pangestu et al., 2017).

Dalam SOM, salah satu komponen penting yang digunakan adalah *learning rate*. Parameter ini berfungsi untuk mengatur tingkat adaptasi bobot terhadap data masukan selama proses pelatihan. Nilai *learning rate* dinyatakan sebagai fungsi skalar dengan rentang  $0 < \alpha(t) < 1$ . Semakin besar nilai  $\alpha(t)$ , semakin cepat bobot jaringan menyesuaikan diri terhadap vektor *input*, sehingga perubahan bobot menjadi lebih signifikan. Sebaliknya, seiring bertambahnya iterasi, nilai  $\alpha(t)$  akan menurun mendekati nol, yang menyebabkan perubahan bobot semakin kecil dan memungkinkan pemetaan vektor *input* menjadi lebih stabil. Selain *learning rate*, SOM juga menggunakan fungsi lingkungan (*neighborhood function*) yang dilambangkan dengan  $h(t)$ , dengan rentang  $0 \leq h(t) \leq 1$ . Fungsi ini berperan dalam menentukan seberapa besar pengaruh neuron pemenang (*Best Matching Unit/BMU*) terhadap neuron-neuron di sekitarnya, berdasarkan jarak dalam radius lingkungan tertentu. Beberapa bentuk fungsi lingkungan yang umum digunakan adalah *Gaussian neighborhood* dan *bubble neighborhood*. *Gaussian neighborhood* dianggap lebih andal karena menghasilkan peta yang konsisten meskipun inisialisasi berbeda, sedangkan *bubble neighborhood* cenderung memberikan tingkat kesalahan kuantisasi yang lebih rendah dan lebih efisien secara komputasi (Kania et al., 2019).

SOM memiliki tiga karakteristik utama, yaitu kompetisi, di mana setiap vektor bobot saling bersaing untuk menjadi neuron pemenang; kooperasi, ketika neuron pemenang berinteraksi dengan neuron-neuron di sekitarnya; serta adaptasi, yaitu proses penyesuaian bobot pada neuron pemenang beserta tetangganya (Imani et al., 2023). Keunggulan utama SOM terletak pada kemampuannya dalam

mereduksi dimensi data yang kompleks menjadi dua dimensi yang lebih mudah dianalisis secara visual (Pangestu et al., 2017). Selain itu, SOM juga efektif digunakan dalam mengidentifikasi kemiripan antar data dalam proses *Clustering* karena mampu mengenali pola kemiripan antar data secara otomatis (Fauzi et al., 2024).

Adapun langkah-langkah algoritma SOM adalah sebagai berikut:

1. Normalisasi data

Normalisasi dilakukan agar setiap variabel berada pada skala yang sama. Pada penelitian ini, metode normalisasi yang diterapkan adalah *Robust Scaler*, yaitu teknik penskalaan data yang memperhitungkan keberadaan *outlier*. Rumus *Robust Scaler* ditunjukkan pada persamaan berikut:

$$x_{scaled} = \frac{(x - Q_2(x))}{(Q_3(x) - Q_1(x))} \quad (2.1)$$

Dengan,

$x$  : Data asli dari fitur atau variabel

$Q_2(x)$  : Median (kuartil kedua) dari fitur  $x$

$Q_1(x)$  : Kuartil pertama dari fitur  $x$

$Q_3(x)$  : Kuartil ketiga dari fitur  $x$

Persamaan (2.1) menghitung selisih data terhadap median, lalu membaginya dengan rentang antar kuartil (IQR), yakni selisih antara kuartil ketiga dan pertama. Dengan cara ini, pengaruh nilai ekstrim (*outlier*) dapat diminimalkan dalam proses normalisasi.

2. Inisialisasi vektor data input  $X_i$  secara acak dari dataset.
3. Inisialisasi bobot awal secara acak dengan rentang nilai 0 – 1.
4. Menghitung jarak *Euclidean* antara vektor input  $X_i$  dan setiap vektor bobot  $W_j$
5. Penentuan *Best Matching Unit* (BMU) yaitu memilih *neuron* dengan jarak terkecil antara vektor input.
6. Menentukan bobot terbaru pada setiap neuron.

Perbarui bobot BMU dan neuron tetangganya menggunakan rumus:

$$W_{ij}(t + 1) = W_{ij}(t) + \alpha \cdot h(t) \cdot (X_i - W_{ij}(t)) \quad (2.2)$$

Dengan,

$\alpha$  : *Learning rate* (laju pembelajaran) yang menurun seiring iterasi

$h(t)$  : Fungsi tetangga yang menurun terhadap waktu dan jarak dari neuron pemenang (BMU)

$i$  : Menunjukkan indeks fitur atau variabel ke- $i$

$j$  : Menunjukkan indeks neuron ke- $j$  pada lapisan *output*

$X_i$  : Vektor *input* ke- $i$

$X_i - W_{ij}(t)$  : Selisih antara *input* dan bobot lama (arah pergeseran)

7. Perbaharui *learning rate* ( $\alpha$ ) dengan rumus:

$$\alpha(t) = \alpha_i \left(1 - \frac{t}{t_{max}}\right) \quad (2.3)$$

Dengan,

$\alpha(t)$  : *Learning rate* pada iterasi ke- $t$

$\alpha_i$  : *Learning rate* awal, (biasanya 0,1 sampai 0,5)

$t$  : Iterasi ke- $t$  saat pelatihan berlangsung

$t_{max}$  : Jumlah iterasi maksimum yang telah ditentukan sebelumnya

Kurangi nilai *learning rate*  $\alpha(t)$  dan radius fungsi tetangga  $h(t)$  seiring iterasi berjalan.

8. Iterasi

Ulangi langkah 2–6 hingga jumlah iterasi maksimum tercapai atau perubahan bobot sangat kecil.

#### 4.7 *Euclidean Distance*

Dalam metode SOM, ukuran jarak yang digunakan untuk mengelompokkan data adalah jarak *Euclidean*. Jarak *Euclidean* sendiri merupakan salah satu metode pengukuran yang digunakan untuk menentukan seberapa jauh dua titik berada dalam ruang *Euclidean*, baik itu dalam dua dimensi, tiga dimensi, maupun dimensi yang lebih tinggi. Metode ini umum digunakan untuk menilai tingkat kemiripan antar data, dengan menghitung jarak langsung antar titik berdasarkan koordinatnya (Nishom, 2019). Untuk mengukur tingkat kemiripan antar data, digunakan rumus *Euclidean distance* berikut.

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^N (x_i - y_i)^2} \quad (2.4)$$

Dengan,

$d$  : Jarak antara dua titik data  $x$  dan  $y$

$\mathbf{x}$  : Nilai pusat kluster

$\mathbf{y}$  : Nilai dari masing-masing atribut pada data

$i$  : Indeks untuk setiap data

$N$  : Total jumlah data

$x_i$  : Nilai pada pusat kluster ke- $i$

$y_i$  : Nilai dari data ke- $i$

#### 4.8 *Silhouette Coefficient*

*Silhouette coefficient* merupakan salah satu metode evaluasi untuk menilai kualitas hasil pengelompokan data pada pendekatan *unsupervised learning*. Nilai koefisien ini berada pada rentang -1 hingga 1, dan hasil klasterisasi dianggap baik apabila nilai *silhouette* mendekati 1 (Syamhuri et al., 2022). Metode ini bekerja dengan cara mengukur sejauh mana kesesuaian suatu objek terhadap kluster tempat ia berada, dibandingkan dengan kluster lain yang paling dekat. Nilai koefisien yang tinggi mengindikasikan objek berada pada kluster yang tepat, sedangkan nilai rendah atau negatif menunjukkan kemungkinan salah kluster. *Silhouette* menggabungkan dua konsep penting dalam *Clustering*, yaitu *kohesi* (kemiripan dalam kluster) dan *separasi* (perbedaan antar kluster) (Fauzi et al., 2024). Perhitungan jarak antar data dalam *silhouette coefficient* dapat menggunakan ukuran seperti *Euclidean Distance* maupun *Manhattan Distance*. Untuk menilai kualitas hasil klasterisasi, nilai *silhouette* dapat dihitung dari setiap kluster, termasuk kluster yang dihasilkan oleh algoritma *Clustering* yang digunakan (Pribadi et al., 2022).

Untuk mengetahui kualitas hasil pengelompokan, nilai *silhouette* dapat dihitung dari setiap kluster, termasuk dari kluster yang dihasilkan oleh algoritma *Clustering* yang digunakan. Tahapan-tahapan dalam metode ini dapat diuraikan sebagai berikut (Pribadi et al., 2022):

1. Menghitung rata-rata jarak antara satu data, misalnya data  $i$ , dengan seluruh data lain yang berada dalam kluster yang sama.

$$a_i = \frac{1}{|A| - 1} \sum_{j \in A, i \neq j} d(i, j) \quad (2.5)$$

Dengan,

$a_i$  : Rata-rata jarak antara data ke- $i$  dengan seluruh data lain dalam kluster yang sama

$|A|$  : Jumlah data dalam kluster  $A$

$i, j$  : Indeks dari data (di mana  $i$  adalah data utama dan  $j$  adalah data pembandingan)

$d(i, j)$  : Jarak antara data  $i$  dengan data  $j$  dalam kluster yang sama

2. Menghitung rata-rata jarak antara data  $i$  dengan seluruh data dari kluster lainnya, kemudian diambil nilai yang paling kecil dari hasil perhitungan tersebut.

$$d(i, C) = \frac{1}{|C|} \sum_{j \in C} d(i, j) \quad (2.6)$$

Dengan,

$d(i, C)$  : Jarak rata-rata antara data ke- $i$  dengan semua objek pada kluster lain ( $C$ ) dimana  $C \neq A$

$|C|$  : Jumlah anggota dalam kluster  $C$

$d(i, j)$  : Jarak antara data  $i$  dan data  $j$

langkah selanjutnya adalah menentukan jarak minimum dari data  $i$  terhadap semua kluster yang bukan merupakan klasternya sendiri. Nilai terkecil dari jarak rata-rata ini kemudian disebut sebagai  $b(i)$ , yang merepresentasikan jarak rata-rata terdekat data  $i$  terhadap kluster lain.

$$b(i) = \min_{C \neq A} d(i, C) \quad (2.7)$$

Dengan,

$b(i)$  : Nilai jarak rata-rata terkecil dari data ke- $i$  terhadap kluster lain terdekat

- Setelah diperoleh nilai jarak minimum, langkah selanjutnya adalah menghitung nilai *Silhouette Coefficient* untuk data  $i$  menggunakan rumus berikut.

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (2.8)$$

#### 4.9 Davies-Bouldin Index (DBI)

*Davies-Bouldin Indeks* (DBI) adalah metode validasi kuantitatif yang digunakan untuk menilai kinerja *Clustering*. DBI mengevaluasi rasio antara jarak antar-klaster dan sebaran dalam klaster. Semakin kecil nilai DBI, maka semakin baik kualitas klaster yang terbentuk. *Davies-Bouldin Indeks* (DBI) juga dapat membantu dalam mengidentifikasi klaster yang memiliki penyimpangan besar (*outlier*) karena nilai DBI yang tinggi menandakan tumpang tindih antar klaster (Aziz & Mustakim, 2022).

*Davies-Bouldin Index* (DBI) didasarkan pada dua komponen utama, yaitu *kohesi*, yang mengukur kedekatan data terhadap pusat klaster yang menaunginya, dan *separasi*, yang didasarkan pada jarak antar pusat klaster (*centroid*) yang terbentuk (Dinata et al., 2020). Dalam *Clustering*, *kohesi* merujuk pada tingkat kedekatan data dengan pusat klaster tempat data tersebut berada, sementara *separasi* menggambarkan sejauh mana pemisahan antar klaster yang diukur berdasarkan jarak antar *centroid* masing-masing klaster (Musid et al., 2023).

Tahapan perhitungan nilai *Davies-Bouldin Index* (DBI) dapat dijelaskan sebagai berikut:

- Persamaan yang digunakan untuk menghitung nilai *kohesi* dalam klaster ke- $i$  dikenal sebagai *Sum of Square Within Cluster* (SSW), yang dinyatakan melalui rumus berikut (Musid et al., 2023).

$$SSW_i = \frac{1}{m_i} \sum_{j=i}^{m_i} d(x_i, c_i) \quad (2.9)$$

Dengan,

$m$  : Banyaknya data yang terdapat pada klaster ke- $i$

$c$  : Titik pusat (*centroid*) dari klaster ke- $i$

$d$  : Nilai jarak antara dua titik data

2. Untuk menghitung nilai separasi antar klaster, digunakan *Sum of Square Between Cluster* (SSB), yaitu sebuah persamaan yang menggambarkan tingkat pemisahan antar klaster (Musid et al., 2023):

$$SSB_{ij} = d(c_i, c_j) \quad (2.10)$$

Dengan,

$d$  : Nilai jarak antar titik data

$c$  : Titik pusat (*centroid*) dari klaster ke- $i$

3. Setelah nilai kohesi dan separasi diperoleh, dilakukan perhitungan rasio  $R_{ij}$  guna membandingkan klaster ke- $i$  dengan klaster ke- $j$ . Klaster terbaik ditandai oleh *kohesi* yang kecil dan *separasi* yang besar. Rasio ini dihitung menggunakan rumus tertentu (Musid et al., 2023):

$$R_{ij} = \frac{SSW_i + SSW_j}{SSB_{ij}} \quad (2.11)$$

4. Nilai rasio tersebut kemudian digunakan untuk menghitung nilai *Davies-Bouldin Index* (DBI) dengan menggunakan rumus yang telah ditentukan (Dinata et al., 2020).

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_{i,j}) \quad (2.12)$$

Adapun semakin kecil nilai *Davies-Bouldin Index* (DBI) yang diperoleh (non-negatif  $\geq 0$ ), maka semakin baik pula kualitas klaster yang terbentuk. Nilai DBI mencerminkan rasio antara sebaran data dalam suatu klaster dan jarak antar klaster. Jumlah klaster  $k$  yang digunakan dalam perhitungan DBI disesuaikan dengan jumlah klaster yang dihasilkan oleh metode SOM.

## **BAB V**

### **PENUTUP**

#### **5.1 Kesimpulan**

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan menggunakan Algoritma *Self-Organizing Maps* (SOM) dalam melakukan klusterisasi indikator IPM Provinsi Sulawesi Selatan tahun 2024, dapat disimpulkan beberapa hal sebagai berikut.

1. Hasil pengelompokan wilayah Provinsi Sulawesi Selatan menghasilkan dua klaster utama. Klaster pertama terdiri atas Kota Makassar, Kota Parepare, dan Kota Palopo, yang memiliki capaian indikator IPM relatif tinggi, terutama pada aspek pendidikan dan ekonomi. Sedangkan klaster kedua mencakup sebagian besar kabupaten di Provinsi Sulawesi Selatan yang memiliki capaian IPM lebih rendah. Hal ini menunjukkan adanya ketimpangan pembangunan manusia antara wilayah perkotaan dan perdesaan.
2. Karakteristik klaster yang terbentuk menunjukkan bahwa perbedaan utama antarwilayah terletak pada indikator Rata-rata Lama Sekolah (RLS) dan Pengeluaran Per Kapita (PPK). Klaster pertama memiliki nilai RLS dan PPK yang lebih tinggi dibandingkan klaster kedua, yang menandakan bahwa aspek pendidikan dan ekonomi merupakan faktor pembeda paling dominan dalam pola pembangunan manusia di Sulawesi Selatan. Oleh karena itu, peningkatan akses pendidikan dan kesejahteraan ekonomi menjadi kunci untuk mengurangi disparitas pembangunan antarwilayah.

#### **5.2 Saran**

Berdasarkan hasil dan pembahasan yang telah diperoleh, sehingga saran yang penulis berikan adalah sebagai berikut.

1. Pemerintah daerah diharapkan dapat memanfaatkan hasil penelitian ini sebagai dasar dalam perumusan kebijakan pembangunan manusia yang lebih terarah, terutama bagi kabupaten yang tergolong dalam klaster dengan nilai IPM rendah.

2. Fokus peningkatan perlu diarahkan pada pemerataan akses pendidikan dan peningkatan kesejahteraan ekonomi masyarakat, mengingat kedua aspek tersebut menjadi faktor dominan dalam perbedaan tingkat pembangunan manusia antarwilayah.
3. Bagi peneliti selanjutnya, disarankan untuk mengembangkan penelitian ini dengan menggunakan periode waktu yang lebih panjang atau menambahkan variabel sosial ekonomi lainnya, seperti tingkat kemiskinan, tingkat pengangguran terbuka, dan tingkat partisipasi angkatan kerja. Dengan demikian, hasil analisis yang diperoleh akan lebih komprehensif dan mampu memberikan gambaran yang lebih menyeluruh mengenai faktor-faktor yang memengaruhi ketimpangan pembangunan manusia antarwilayah di Provinsi Sulawesi Selatan.

## DAFTAR PUSTAKA

- Alfina, T., Santosa, B., & Barakbah, A. R. (2012). Analisa Perbandingan Metode Hierarchical Clustering, K-means dan Gabungan Keduanya dalam Cluster Data. *Teknik Its*, 1(1), 521-525. <http://www.ejurnal.its.ac.id/index.php/teknik/article/view/1794>
- Aziz, S., & Mustakim, M. (2022). Implementasi Algoritma Self Organizing Map untuk Identifikasi Pola Pengelompokan Tingkat Kesejahteraan Keluarga Kabupaten Siak. *Indonesian Journal of Informatic Research and Software Engineering (IJIRSE)*, 2(2), 82–92. <https://doi.org/10.57152/ijirse.v2i2.431>
- Badan Pusat Statistik (BPS). (2024). Indeks Pembangunan Manusia Provinsi Sulawesi Selatan 2023 (Volume 9). Badan Pusat Statistik.
- Badan Pusat Statistik Provinsi Sulawesi Selatan. (2025). Provinsi Sulawesi Selatan dalam Angka 2025.
- Dewi, D. A. I. C., & Pramita, D. A. K. (2019). Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali. *Matrix: Jurnal Manajemen Teknologi Dan Informatika*, 9(3), 102–109. <https://doi.org/10.31940/matrix.v9i3.1662>
- Dinata, R. K., Novriando, H., Hasdyna, N., & Retno, S. (2020). Reduksi Atribut Menggunakan Information Gain. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 6(1), 48–53.
- Fahmi, M. N. (2023). Implementasi Mechine Learning menggunakan Python Library : Scikit-Learn (Supervised dan Unsupervised Learning). *Sains Data Jurnal Studi Matematika Dan Teknologi*, 1(2), 87–96. <https://doi.org/10.52620/sainsdata.v1i2.31>
- Fauzi, A., Harli, E., & Kusmanto, T. H. (2024). Penerapan Algoritma Self Orginizing Mapping untuk Analisis Clustering Indeks Pembangunan Manusia di Jawa Barat. 2(3), 408–418.
- Ginting, R. (2013). Analisis Penggunaan Algoritma Kohonen Pada Jaringan Syaraf Tiruan Backpropagation Dalam Pengenalan Pola Penyakit Paru. Tesis, 01, 1–78.
- Hastuti, S. H., Nurmayanti, W. P., & Saputri, A. A. (2023). Penerapan Metode Clustering Self Organizing Maps (Som) Dan K-Affinity Propagation (K-Ap) Dalam Mengelompokkan Nilai Tukar Petani Di Indonesia 2022. *VARIANCE: Journal of Statistics and Its Applications*, 5(1), 79–88. <https://doi.org/10.30598/variancevol5iss1page79-88>

- Hermawan, & Hasugian, H. (2022). Penerapan Data Mining Untuk Clustering Indeks Pembangunan Manusia Berdasarkan Provinsi Di Indonesia. *Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, 1(1), 525–532.
- Imani, N., Alfassa, A. I., & Yolanda, A. M. (2023). Analisis Cluster Terhadap Indikator Data Sosial Di Provinsi Nusa Tenggara Timur Menggunakan Metode Self Organizing Map (SOM). *Jurnal Gaussian*, 11(3), 458–467. <https://doi.org/10.14710/j.gauss.11.3.458-467>
- Irwan, Sanusi, W., & Hasanah, A. (2024). Perbandingan Analisis Cluster Metode Complete Linkage dan Metode Ward dalam Pengelompokan Indeks Pembangunan Manusia di Sulawesi Selatan. 7(1), 75–86.
- Ismail, I. (2021). Penerapan Data Science Menggunakan Artificial Neural Network (ANN) Metode Self Organizing Mapping (SOM) untuk Klasifikasi Industri. *Warta Akab*, 45(2), 66–70. <https://doi.org/10.55075/wa.v45i2.62>
- Iyohu, L. R., Djakaria, I., & Nashar, L. O. (2023). Perbandingan Metode K-Means Clustering dengan Self-Organizing Maps (SOM) untuk Pengelompokan Provinsi di Indonesia Berdasarkan Data Potensi Desa. *Jurnal Statistika Dan Aplikasinya*, 7(2), 195–206. <https://doi.org/10.21009/jsa.07208>
- Kamila, I., Khairunnisa, U., & Mustakim, M. (2019). Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau. *Jurnal Ilmiah Rekayasa Dan Manajemen Sistem Informasi*, 5(1), 119. <https://doi.org/10.24014/rmsi.v5i1.7381>
- Kania, S., Rachmatin, D. and Dahlan, J.A. (2019) ‘Program Aplikasi Pengelompokan Objek Dengan Metode Self Organizing Map Menggunakan Bahasa R’, *Jurnal EurekaMatika*, 7(2), pp. 17–29.
- Midyanti, D. M., & Bahri, S. (2023). Implementasi Self Organizing Maps Untuk Pengelompokan Kabupaten / Kota Berdasarkan Indeks Pembangunan Manusia Implementation of Self-Organizing Maps for District / City Grouping Based on the Human Development Index. 10(6), 1265–1272. <https://doi.org/10.25126/jtiik.2023107647>
- Musid, F. N., Dikananda A.R., & Fathurrohman. (2023). Implementasi Algoritma K-Means Clustering Dalam Pengelompokan Data Jumlah Kerusakan Rumah Berdasarkan Kondisi Di Jawa Barat. *Journal of Student Research*, 1(3), 101–114. <https://doi.org/10.55606/jsr.v1i3.1155>
- Nasri, M., Nujum, S., & Arfah, A. (2024). Pengaruh Pendidikan, Upah Minimum dan Indeks Pembangunan Manusia terhadap Tingkat Kemiskinan di Sulawesi Selatan. *SEIKO : Journal of Management & Business*, 7(2), 543–555.

- Nazeriandy, N., Syahra, Y., & Syaifudin, M. (2021). Penerapan Data Mining Untuk Memprediksi Penggunaan Daya Listrik Pada PT.PLN (Persero) Rayon Medan Selatan Dengan Menggunakan Metode Regresi Linier Berganda. *Jurnal SAINTIKOM (Jurnal Sains Manajemen Informatika Dan Komputer)*, 20(1), 20. <https://doi.org/10.53513/jis.v20i1.2431>
- Nishom, M. (2019). Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square. *Jurnal Informatika: Jurnal Pengembangan IT*, 4(1), 20–24. <https://doi.org/10.30591/jpit.v4i1.1253>
- Pangestu, B., Fatichah, C., & Purwitasari, D. (2017). Kartografi Menggunakan Self-Organizing Maps
- Pangestu, B., Purwitasari, D., & Fatichah, C. (2017). Visualisasi Similaritas Topik Penelitian dengan Pendekatan Kartografi Menggunakan Self-Organizing Maps (SOM). *Jurnal Teknik ITS*, 6(2). <https://doi.org/10.12962/j23373539.v6i2.23706>
- Parulian, R., & Butar, B. (2023). Analisis Hierarchical Dan Non-Hierarchical Clustering Untuk Pengelompokkan Potensi Ekonomi Kelautan Indonesia 2021 Hierarchical and Non-Hierarchical Clustering Analysis for Classifying Indonesia ' s Sea Economic Potential in 2021. 11(3), 543–553. <https://doi.org/10.26418/justin.v11i3.67283>
- Pribadi, W. W., Yunus, A., & Wiguna, A. S. (2022). Perbandingan Metode K-Means Euclidean Distance Dan Manhattan Distance Pada Penentuan Zonasi Covid-19 Di Kabupaten Malang. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 6(2), 493–500. <https://doi.org/10.36040/jati.v6i2.4808>.
- Radityana, I. D. G. M., Djayastra, I. K., Danendra A. A. N. B., & Murthi, N. W. (2023). Pengaruh Upah Minimum, Indeks Pembangunan Manusia dan Pengangguran Terbuka terhadap Jumlah Kemiskinan di Provinsi Bali. *Jurnal Kajian Ekonomi Dan Manajemen Indonesia (JKEMI)*, 1(1), 16–24. <https://doi.org/10.61079/jkemi.v1i1.3>
- Santoso, D. H., Arsyi, F., Clarissa, A. A., Setiawan, I. N., Kurniati, E., & Delyana, S. (2024). Indeks Pembangunan Manusia 2023. 18, 1–282.
- Syamhuri, W. B., Furqon, M. T., & Dewi, C. (2022). Pengelompokan Film Berdasarkan Alur Cerita menggunakan Metode Self Organizing Maps dan Silhouette Coefficient. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 6(12), 5898–5904.
- Talakua, M. W., Leleury, Z. A., & Talluta, A. W. (2017). Analisis Cluster Dengan Menggunakan Metode Provinsi Maluku Berdasarkan Indikator Indeks Pembangunan Manusia Tahun 2014 Cluster Analysis By Using K-Means

Method for Grouping of District / City in Maluku Province Industrial Based on Indicators of Maluku Dev. Jurnal Ilmu Matematika Dan Terapan, 11(2), 119–128.

Zahra, R. R. A. (2022). Analisis perbandingan metode Self Organizing Map dan metode Fuzzy C-Means pada pengelompokkan permintaan jurusan di Sekolah Menengah Kejuruan. Jurnal Ilmiah SINUS, 20(1), 45–56. <https://doi.org/10.30646/sinus.v20i1.562>.