

# BAB I PENDAHULUAN

## A. Latar Belakang

Kabupaten Lingga memiliki pola curah hujan yang unik, dipengaruhi oleh faktor-faktor lokal seperti kondisi geografis, atmosfer, dan angin. Menurut data dari Pusat Statistik Kabupaten Lingga Tahun 2024, curah hujan Kabupaten Lingga menunjukkan variasi musiman yang signifikan, dengan bulan Desember sebagai bulan dengan curah hujan tertinggi rata-rata 417 milimeter.

Salah satu sektor yang paling terpengaruh adalah pertanian dan perikanan. Bagi para petani, pola curah hujan yang tidak menentu dapat menyebabkan gagal panen akibat kelebihan air atau kekeringan pada musim tanam tertentu. Sementara bagi nelayan, perubahan pola hujan yang disertai peningkatan kecepatan angin dan gelombang laut dapat menghambat aktivitas melaut, sehingga menurunkan produktivitas dan pendapatan. Oleh karena itu, memahami pola curah hujan secara lebih mendalam sangat penting untuk mendukung perencanaan pertanian dan aktivitas ekonomi masyarakat pesisir di Kabupaten Lingga.

Menganalisis pola curah hujan memerlukan pendekatan yang tepat untuk memahami data yang multivariat, seperti kelembapan, temperatur, kecepatan angin, dan lain-lain. Data cuaca sering kali terdiri dari berbagai faktor yang saling terkait, dan memerlukan pemahaman mendalam untuk melihat hubungan antara faktor-faktor tersebut [1]. Menggunakan metode statistik konvensional atau analisis sederhana sering kali tidak cukup untuk menangkap dinamika yang terkandung dalam data tersebut. Analisis curah hujan harus memperhitungkan interaksi antar variabel-variabel cuaca yang bersifat dinamis dan multivariat untuk memberikan hasil yang lebih akurat dan representatif. Penelitian ini juga menekankan bahwa metode yang lebih canggih diperlukan untuk menggali kompleksitas pola cuaca di berbagai wilayah, termasuk Kabupaten Lingga, yang memiliki kondisi atmosfer yang sangat variatif [2].

Metode tradisional dalam analisis curah hujan, seperti regresi linear atau analisis statistik klasik, sering kali memiliki keterbatasan dalam menangkap pola non-linear yang terdapat dalam data cuaca. Pola curah hujan dapat berubah secara tidak terduga seiring waktu, sehingga tidak selalu dapat dijelaskan menggunakan

model linier [3]. Interaksi antara berbagai variabel cuaca juga cenderung kompleks, sehingga pendekatan klasik kurang efektif dalam memahami hubungan tersebut. Pendekatan yang lebih modern, seperti *machine learning* atau algoritma klusterisasi, lebih sesuai untuk menangani ketidakpastian dan pola non-linear dalam analisis curah hujan. Metode ini memungkinkan pemodelan hubungan yang lebih dinamis antar variabel yang memengaruhi fenomena curah hujan [4], termasuk di wilayah Kabupaten Lingga.

Dengan semakin berkembangnya teknik analisis seperti *Self-Organizing Map* (SOM) atau algoritma berbasis kecerdasan buatan, masalah dalam analisis pola curah hujan dapat diatasi dengan lebih efektif. Metode-metode ini mampu mengidentifikasi pola tersembunyi dalam data yang tidak dapat dideteksi dengan metode tradisional [5]. Menggunakan teknik yang lebih modern memungkinkan untuk menemukan kluster atau grup curah hujan berdasarkan berbagai variabel cuaca yang ada, serta menangkap hubungan yang lebih kompleks antar faktor-faktor tersebut.

*Particle Swarm Optimization* (PSO) adalah algoritma optimasi berbasis populasi yang terinspirasi dari perilaku kawanan hewan, seperti burung atau ikan, dalam mencari sumber daya [6]. *Particle Swarm Optimization* (PSO) digunakan untuk menemukan solusi optimal dalam ruang pencarian dengan mengevaluasi setiap partikel dalam populasi. Dalam konteks analisis data, *Particle Swarm Optimization* (PSO) dapat digunakan untuk mengoptimalkan parameter dalam algoritma pembelajaran mesin, termasuk *Self-Organizing Map* (SOM) [7].

*Self-Organizing Map* (SOM) adalah metode klusterisasi yang efektif dalam menangani data berdimensi tinggi dan besar. Namun, kinerja *Self-Organizing Map* (SOM) dapat dipengaruhi oleh pemilihan bobot awal dan parameter lainnya. Oleh karena itu, mengintegrasikan *Particle Swarm Optimization* (PSO) dengan *Self-Organizing Map* (SOM) dapat meningkatkan akurasi klusterisasi dengan mengoptimalkan bobot dan parameter SOM secara efektif [7]. Dalam Jurnal Eksponensial menunjukkan bahwa penerapan PSO pada SOM dalam pengelompokan desa/kelurahan tertinggal di Kabupaten Kutai Kartanegara mampu meningkatkan akurasi klusterisasi. Penelitian ini mengindikasikan bahwa kombinasi PSO dan SOM dapat menghasilkan kluster yang lebih representatif dan

akurat dalam analisis data kompleks. Selain itu, penelitian oleh Wahyuni, 2022, dalam Jurnal *Syntax Literate* membahas penggunaan SOM dalam pengelompokan dan pemetaan karakteristik kemiskinan di Provinsi Nusa Tenggara Barat. Meskipun tidak secara langsung mengintegrasikan PSO dengan SOM, studi ini menunjukkan efektivitas SOM dalam analisis data sosial-ekonomi yang kompleks. [8]

*Gradient descent* merupakan salah satu metode optimasi yang digunakan untuk menemukan nilai minimum dari suatu fungsi biaya (*loss function*). Metode ini bekerja secara iteratif dengan cara menyesuaikan parameter ke arah negatif gradien dari fungsi biaya agar nilai fungsi biaya berkurang pada setiap iterasi. *Gradient Descent* banyak di gunakan dalam berbagai bidang, seperti *machine learning*, pengolahan citra, serta optimasi sistem nonlinear, karena kesederhanaan dan efektivitasnya dalam menemukan minimum fungsi biaya. [9]

Integrasi kedua metode ini menawarkan pendekatan yang lebih efektif dalam analisis data yang kompleks dan berdimensi tinggi, seperti data cuaca atau data sosial-ekonomi. Pendekatan ini memungkinkan identifikasi pola tersembunyi dalam data yang tidak dapat dideteksi dengan metode tradisional, sehingga meningkatkan akurasi dan efisiensi dalam proses klusterisasi.

Dalam proses klusterisasi, terdapat beberapa parameter utama yang disesuaikan untuk meningkatkan performa, yaitu learning rate, yang mengatur seberapa cepat model belajar dari data, serta radius tetangga, yang menggambarkan sejauh mana pengaruh satu neuron terhadap neuron lainnya dalam proses pembelajaran.

Untuk mengatasi tantangan tersebut, metode *Self-Organizing Map* (SOM) digunakan untuk melakukan klusterisasi data curah hujan karena kemampuannya dalam pengelompokan data tanpa supervisi. Namun, performa SOM sangat bergantung pada pemilihan parameter yang optimal dan penyesuaian bobot neuron. Dalam proses klusterisasi, terdapat beberapa parameter utama yang memengaruhi performa *Self-Organizing Map* (SOM), yaitu learning rate, yang menentukan seberapa cepat model belajar dari data, dan radius tetangga, yang menggambarkan sejauh mana pengaruh satu neuron terhadap neuron lainnya selama pembelajaran.

Untuk mengoptimalkan parameter-parameter ini, penelitian ini membandingkan dua metode optimasi, yaitu *Particle Swarm Optimization* (PSO) dan *Gradient Descent* (GD). PSO digunakan untuk menemukan kombinasi

parameter awal yang optimal berdasarkan evaluasi populasi partikel dalam ruang pencarian, sedangkan GD secara iteratif memperbarui parameter dengan meminimalkan fungsi kesalahan (*quantization error*) menggunakan turunan gradien. Dengan membandingkan kedua metode ini, penelitian bertujuan menilai efektivitas masing-masing pendekatan dalam meningkatkan akurasi klusterisasi dan konvergensi SOM, serta menentukan metode yang lebih tepat untuk mengoptimalkan *learning rate* dan radius tetangga.

## **B. Perumusan Masalah**

Dalam penelitian ini rumusan masalah yang di kemukakan berdasarkan pemaparan yang telah disampaikan pada latar belakang yaitu bagaimana perbandingan metode Particle Swarm Optimization (PSO) dan Gradient Descent (GD) dalam mengoptimalkan parameter Self-Organizing Map (SOM), *learning rate* dan radius tetangga, untuk meningkatkan performa klusterisasi data curah hujan?

## **C. Batasan Penelitian**

Penelitian ini terdiri dari beberapa batasan masalah yang di gunakan sebagai berikut :

1. Penelitian ini hanya menggunakan data harian cuaca dari BMKG Dabo Singkep dari tahun 2000 sampai 2024 untuk analisis curah hujan. Data yang di gunakan sebanyak 9.373 data.
2. Data yang digunakan mencakup variabel kelembapan rata-rata, temperatur minimum, temperatur maksimum, temperatur rata-rata, kecepatan angin rata-rata, arah angin terbanyak, lamanya penyinaran matahari, dan curah hujan.
3. Algoritma yang di gunakan, *Self organizing map*, *particle swarm optimization* dan *gradient descent*
4. Klusterisasi dibagi menjadi kategori sangat ringan, ringan, sedang, lebat, sangat lebat.
5. Parameter yang di optimasi *learning rate*, dan radius tetangga.
6. Data BMKG yang digunakan telah mengalami proses pengolahan oleh pihak ketiga, sehingga berpotensi mengandung bias dan ketidaktepatan,

penelitian ini tidak memiliki kendali terhadap tahapan pengumpulan maupun validasi data tersebut, sehingga akurasi hasil prediksi tidak dapat sepenuhnya dianggap sebagai representatif kondisi aktual.

7. Waktu komputasi yang diperoleh pada penelitian ini tidak bersifat absolut karena dipengaruhi oleh kondisi dan alokasi sumber daya pada server *Google Colaboratory* saat proses pelatihan model dilakukan.

#### **D. Tujuan Penelitian**

Adapun tujuan dari dilakukannya penelitian ini adalah membandingkan metode *Particle Swarm Optimization* (PSO) dan *Gradient Descent* (GD) dalam mengoptimalkan parameter Self-Organizing Map (SOM), *learning rate* dan radius tetangga, untuk meningkatkan performa klusterisasi data curah hujan di Kabupaten Lingga.

#### **E. Manfaat Penelitian**

Penelitian ini diharapkan memberikan kontribusi dalam bidang analisis data, adapun manfaat yang diperoleh dari penelitian ini antara lain:

1. Menjadi referensi ilmiah dalam penerapan metode klusterisasi untuk data curah hujan.
2. Memberikan kontribusi dalam pengembangan teknik analisis data meteorologi menggunakan pendekatan *mechine learning*.
3. Mendorong pemanfaatan algoritma optimasi dalam meningkatkan performa metode klusterisasi
4. Memperluas wawasan dan pemahaman tentang integrasi metode klusterisasi dan optimasi.

#### **F. Keaslian Penelitian**

Penelitian ini mengusulkan pendekatan baru dalam klusterisasi data curah hujan dengan memanfaatkan Self-Organizing Map (SOM) yang dioptimalkan menggunakan metode Particle Swarm Optimization (PSO) dan Gradient Descent (GD). Berbeda dengan penelitian sebelumnya yang hanya berfokus pada optimasi bobot SOM, penelitian ini menitikberatkan optimasi pada parameter utama SOM, yaitu *learning rate* dan radius tetangga, sekaligus membandingkan efektivitas kedua metode optimasi tersebut dalam meningkatkan performa klusterisasi. Data

yang digunakan mencakup delapan variabel meteorologi, yaitu kelembaban rata – rata, temperatur minimum, temperatur maksimum, temperatur rata – rata, kecepatan angin rata – rata, arah angin terbanyak, lamanya penyinaran matahari, dan curah hujan.

Pembeda penelitian dan penelitian sebelumnya pada tabel di bawah:

Tabel 1. Pembeda Penelitian Sebelumnya Dan Penelitian Yang Diusulkan

Aspek	Penelitian Sebelumnya	Penelitian yang di Usulkan
Fokus	Optimasi <i>Self Organizing Map</i> (SOM) Menggunakan <i>Particle Swarm Optimization</i> (PSO) Untuk Mengelompokkan Desa/Kelurahan Tertinggal Di Kabupaten Kutai Kartanegara, Provinsi Kalimantan Timur	Optimasi <i>Self Organizing Map</i> (SOM) Menggunakan <i>Particle Swarm Optimization</i> (PSO) dan <i>Gradient Descent</i> untuk <i>Clustering</i> Data Curah Hujan
Variabel yang Digunakan	19 variabel sosial-ekonomi (kepadatan penduduk, jumlah sarana pendidikan, kesehatan, jumlah bencana alam, dll.).	8 variabel meteorologi (kelembapan rata-rata, temperatur minimum, temperatur maksimum, temperatur rata-rata, kecepatan angin rata-rata, arah angin terbanyak, lamanya penyinaran matahari, dan curah hujan)
Tujuan Penelitian	Mengidentifikasi desa/kelurahan tertinggal untuk membantu pemerintah dalam menyalurkan bantuan.	Penelitian diharapkan dapat meningkatkan akurasi klusterisasi tren curah hujan
Metode Analisis	SOM yang dioptimalkan dengan PSO untuk mengelompokkan desa berdasarkan data sosial-ekonomi	SOM yang dioptimalkan dengan PSO untuk mengelompokkan data curah hujan
Parameter yang di Optimasi	Bobot pada <i>Self Organizing Map</i>	Parameter <i>Self Organizing Map</i> , yaitu <i>learning rate</i> dan radius tetangga