

BAB I

PENDAHULUAN

1.1 Latar Belakang

Bencana hidrometeorologi merupakan bencana alam yang berkaitan dengan proses atmosfer dan hidrologi serta dipicu oleh dinamika cuaca dan iklim (BNPB, 2025). Dalam beberapa dekade terakhir, frekuensi dan intensitas bencana hidrometeorologi di berbagai belahan dunia menunjukkan kecenderungan meningkat seiring dengan perubahan iklim global. Perubahan iklim memicu terjadinya peningkatan kejadian cuaca ekstrem, seperti curah hujan dengan intensitas tinggi, kekeringan berkepanjangan, serta gangguan terhadap siklus hidrologi yang berdampak pada meningkatnya risiko bencana hidrometeorologi. Kondisi tersebut menjadikan perubahan iklim sebagai salah satu isu strategis di tingkat global karena memberikan dampak yang luas terhadap keberlanjutan kehidupan manusia, keseimbangan ekosistem, serta berbagai sektor pembangunan (Idrus & Usi, 2024). Peningkatan kejadian bencana hidrometeorologi tersebut menunjukkan bahwa risiko bencana tidak hanya dipengaruhi oleh faktor alam, tetapi juga berkaitan erat dengan perubahan iklim serta aktivitas manusia yang berkontribusi terhadap meningkatnya tingkat kerentanan suatu wilayah.

Peningkatan frekuensi dan intensitas bencana hidrometeorologi sebagai dampak perubahan iklim mendorong perlunya upaya yang lebih sistematis dalam meningkatkan ketahanan dan kapasitas adaptasi terhadap bencana. Komitmen tersebut diwujudkan melalui Agenda 2030 *for Sustainable Development* yang disepakati oleh negara-negara anggota Perserikatan Bangsa-Bangsa pada tahun 2015 dan memuat 17 tujuan *Sustainable Development Goals (SDGs)* sebagai kerangka pembangunan berkelanjutan di tingkat global (United Nations, 2015). Salah satu tujuan yang berkaitan langsung dengan isu perubahan iklim adalah *SDG ke-13 (Climate Action)*, yang menekankan pentingnya tindakan untuk mengatasi perubahan iklim beserta dampaknya. Secara lebih spesifik, target 13.1 menekankan penguatan kapasitas ketahanan (*resilience*) dan kemampuan adaptasi terhadap bahaya yang berkaitan dengan perubahan iklim serta bencana alam di

seluruh negara (Jia dkk., 2021). Upaya pencapaian target tersebut memerlukan informasi yang mampu menggambarkan karakteristik dampak bencana pada setiap wilayah sehingga penyusunan strategi mitigasi, adaptasi, dan penanggulangan bencana dapat dilakukan secara lebih efektif dan tepat sasaran.

Kondisi geografis dan karakteristik iklim menjadikan Indonesia sebagai salah satu negara yang memiliki tingkat kerawanan bencana hidrometeorologi yang tinggi. Banjir, tanah longsor, kekeringan, dan cuaca ekstrem merupakan jenis bencana yang paling sering terjadi di Indonesia (Baldah dkk., 2023). Dokumen Rencana Pembangunan Jangka Menengah Nasional (RPJMN) menyebutkan bahwa lebih dari 95% kejadian bencana di Indonesia termasuk ke dalam kelompok bencana hidrometeorologi yang berkaitan erat dengan dinamika iklim (Tarumingkeng, 2025). Tingginya dominasi bencana hidrometeorologi tersebut menunjukkan bahwa perubahan dan variabilitas iklim memberikan kontribusi yang signifikan terhadap peningkatan risiko bencana di berbagai wilayah Indonesia. Kondisi tersebut mendorong perlunya penyelenggaraan penanggulangan bencana yang dilakukan secara terencana, terpadu, terkoordinasi, dan menyeluruh sebagaimana diamanatkan dalam Undang-Undang Nomor 24 Tahun 2007 tentang Penanggulangan Bencana. Pasal 35 dan Pasal 41 pada peraturan tersebut menegaskan bahwa analisis risiko bencana merupakan salah satu dasar dalam penyusunan rencana penanggulangan bencana serta perencanaan pembangunan. Ketentuan tersebut menunjukkan bahwa ketersediaan informasi kebencanaan yang akurat dan komprehensif memiliki peranan penting dalam mendukung penyusunan strategi mitigasi, adaptasi, dan penanggulangan bencana di Indonesia (Republik Indonesia, 2007).

Informasi kebencanaan yang akurat tidak hanya diperlukan untuk mengetahui frekuensi kejadian bencana, tetapi juga untuk memahami karakteristik dampak yang ditimbulkan pada setiap wilayah. Menurut *United Nations Office for Disaster Risk Reduction* (UNDRR, 2025), dampak bencana mencakup berbagai aspek, antara lain korban jiwa, kerusakan aset, gangguan terhadap layanan publik, hilangnya mata pencaharian, perpindahan penduduk, serta terhambatnya proses pembangunan. Karakteristik dampak bencana hidrometeorologi di Indonesia

menunjukkan variasi yang cukup besar antarprovinsi, baik ditinjau dari jumlah kejadian bencana, jumlah korban terdampak, kerusakan rumah, rumah terendam, maupun kerusakan fasilitas umum. Variasi tersebut mengindikasikan bahwa setiap provinsi memiliki karakteristik dampak bencana yang berbeda sehingga memerlukan penanganan yang disesuaikan dengan kondisi masing-masing wilayah. Identifikasi pola kemiripan karakteristik dampak bencana menjadi penting untuk memperoleh informasi mengenai kelompok wilayah yang memiliki karakteristik serupa sebagai dasar dalam penyusunan strategi mitigasi, adaptasi, dan penanggulangan bencana yang lebih efektif. Karakteristik tersebut sulit diamati secara langsung hanya melalui penyajian statistik deskriptif sehingga diperlukan pendekatan analisis yang mampu mengelompokkan wilayah berdasarkan tingkat kemiripan karakteristik dampak bencana. Salah satu pendekatan yang banyak digunakan untuk tujuan tersebut adalah analisis *clustering*, yaitu teknik analisis yang bertujuan membentuk kelompok objek dengan tingkat kemiripan yang tinggi di dalam kelompok (*within-cluster similarity*) dan tingkat perbedaan yang tinggi antar kelompok (*between-cluster dissimilarity*).

Perkembangan *Machine Learning* memungkinkan proses pengenalan pola dan ekstraksi informasi dari data dilakukan secara otomatis melalui penerapan berbagai algoritma pembelajaran. Secara umum, *Machine Learning* terdiri atas tiga paradigma utama, yaitu *supervised learning*, *unsupervised learning*, dan *reinforcement learning*. *Supervised learning* digunakan untuk membangun model berdasarkan data yang telah memiliki label, *unsupervised learning* bertujuan menemukan pola atau struktur tersembunyi pada data tanpa label, sedangkan *reinforcement learning* mempelajari tindakan terbaik melalui interaksi dengan lingkungan berdasarkan mekanisme *reward* (Bishop, 2006). Salah satu teknik yang paling banyak digunakan dalam *unsupervised learning* adalah *clustering*, yaitu metode yang bertujuan mengelompokkan objek ke dalam kelompok-kelompok berdasarkan tingkat kemiripan karakteristiknya, sehingga objek dalam kelompok yang sama memiliki tingkat kemiripan yang lebih tinggi dibandingkan dengan objek pada kelompok lain. Berdasarkan mekanisme pembentukannya,

metode *clustering* dapat diklasifikasikan ke dalam beberapa pendekatan, antara lain *partitioning methods*, *hierarchical methods*, *density-based methods*, dan *grid-based methods* (Han dkk., 2011).

Berbagai penelitian telah menunjukkan bahwa metode *clustering* konvensional mampu dimanfaatkan untuk mengidentifikasi pola dan karakteristik wilayah berdasarkan data kebencanaan. Baldah dkk. (2023) menerapkan algoritma *K-Means* untuk mengelompokkan daerah rawan bencana di Indonesia dan memperoleh dua *cluster* optimal dengan nilai *Silhouette Coefficient* sebesar 0,744. Penelitian lain yang dilakukan oleh Pambudhi dkk. (2024) membandingkan algoritma *K-Means* dan *K-Medoids* dalam klasterisasi wilayah rawan bencana di Kabupaten Situbondo, di mana *K-Means* menghasilkan kualitas *cluster* yang lebih baik berdasarkan nilai *Davies-Bouldin index (DBI)* sebesar 0,853 dibandingkan *K-Medoids* sebesar 1,388. Setiawan & Oktafianto (2025) menggunakan metode *Fuzzy C-Means* untuk mengelompokkan tingkat kerawanan bencana di Provinsi Jawa Timur dan memperoleh tiga *cluster* dengan nilai *Silhouette Coefficient* sebesar 0,2807. Sementara itu, Boer dkk. (2023) menerapkan *DBSCAN* dalam pengelompokan daerah rawan bencana di Pulau Sumatera dan menghasilkan dua *cluster* serta 44 data *noise* dengan nilai *Silhouette Coefficient* sebesar 0,46. Penelitian yang dilakukan oleh Sudrajat dkk. (2025) juga menunjukkan bahwa metode *K-Means* mampu digunakan untuk memetakan wilayah rawan bencana di Kota Cirebon dengan menghasilkan nilai *Silhouette Coefficient* sebesar 0,6804 dan *purity score* sebesar 1,0. Temuan-temuan tersebut menunjukkan bahwa berbagai metode *clustering* konvensional telah berhasil diterapkan untuk mengidentifikasi pola pada data kebencanaan. Meskipun demikian, sebagian besar metode tersebut masih berfokus pada pembentukan kelompok berdasarkan kedekatan objek sehingga belum mampu merepresentasikan hubungan topologi antarobjek secara eksplisit pada data multidimensi.

Meskipun berbagai metode *clustering* konvensional telah berhasil diterapkan dalam analisis data kebencanaan, masing-masing metode masih memiliki keterbatasan ketika digunakan untuk menganalisis data multidimensi yang kompleks. Metode berbasis *partitioning*, seperti *K-Means*, menentukan

pusat *cluster* menggunakan *centroid* yang diperoleh dari nilai rata-rata setiap kelompok sehingga keberadaan observasi dengan nilai ekstrem (*outlier*) dapat memengaruhi posisi pusat *cluster* dan menurunkan kualitas hasil pengelompokan (Han dkk., 2011). Selain itu, sebagian besar metode *clustering* konvensional berorientasi pada pembentukan kelompok berdasarkan kedekatan jarak antarobjek tanpa mempertimbangkan hubungan topologi di dalam data. Akibatnya, objek yang memiliki karakteristik serupa belum tentu dipetakan pada posisi yang berdekatan sehingga struktur hubungan pada data multidimensi belum dapat direpresentasikan secara optimal (Kohonen, 2001). Karakteristik data dampak bencana hidrometeorologi yang melibatkan berbagai indikator dengan hubungan yang kompleks menunjukkan perlunya pendekatan analisis yang tidak hanya mampu mengelompokkan objek berdasarkan kemiripan karakteristik, tetapi juga mempertahankan struktur hubungan antarobjek agar pola dalam data dapat direpresentasikan secara lebih komprehensif.

Salah satu pendekatan yang mampu mengatasi keterbatasan tersebut adalah *artificial neural network* (ANN), yaitu metode komputasi yang dirancang untuk meniru mekanisme kerja jaringan saraf biologis dalam mempelajari pola dari data. Salah satu mekanisme pembelajaran pada ANN adalah *competitive learning*, yaitu proses pembelajaran tanpa pengawasan (*unsupervised learning*) di mana neuron-neuron saling berkompetisi untuk menjadi representasi terbaik dari data masukan. Konsep *competitive learning* tersebut menjadi dasar pengembangan metode *Self Organizing Maps* (SOM) yang diperkenalkan oleh Kohonen. Melalui mekanisme tersebut, SOM mampu memetakan data berdimensi tinggi ke dalam representasi dua dimensi berdasarkan tingkat kemiripan karakteristik objek dengan tetap mempertahankan hubungan topologi antarobjek sehingga pola kemiripan dalam data menjadi lebih mudah diamati dan diinterpretasikan (Kohonen, 2001).

Berbagai penelitian menunjukkan bahwa SOM merupakan salah satu metode *unsupervised learning* yang memiliki kemampuan baik dalam menganalisis data multidimensi, termasuk pada bidang kebencanaan.. Simamora dkk. (2024) menunjukkan bahwa SOM mampu menghasilkan distribusi *cluster*

yang lebih merata sekaligus mempertahankan hubungan topologi data dibandingkan metode *K-Means* pada analisis percepatan tanah maksimum akibat gempa bumi di Indonesia. Penelitian yang dilakukan oleh Dermawan dkk. (2023) juga menunjukkan bahwa SOM mampu mengelompokkan provinsi di Indonesia berdasarkan karakteristik dampak bencana gempa bumi dengan nilai *silhouette coefficient* sebesar 0,65. Hasil serupa dilaporkan oleh Risky dkk. (2023), yang menunjukkan bahwa SOM mampu memvisualisasikan pola kemiripan antarwilayah secara lebih interpretatif dalam pengelompokan dampak bencana tanah longsor dengan nilai *silhouette coefficient* sebesar 0,7066. Selain itu, Hartatik & Cahya (2020) memperoleh tingkat akurasi sebesar 96,20% dalam pengelompokan kerusakan akibat gempa bumi menggunakan SOM. Temuan-temuan tersebut menunjukkan bahwa SOM tidak hanya mampu mengelompokkan objek berdasarkan karakteristik yang serupa, tetapi juga mempertahankan hubungan topologi serta menyajikan visualisasi yang memudahkan interpretasi pola kemiripan antarobjek pada data multidimensi.

Meskipun berbagai penelitian menunjukkan bahwa SOM memiliki kemampuan yang baik dalam memetakan pola kemiripan data multidimensi, hasil yang diperoleh pada dasarnya masih berupa representasi topologi dalam bentuk neuron dan *codebook vectors*. Representasi tersebut menggambarkan hubungan kedekatan antarobjek berdasarkan kemiripan karakteristiknya, tetapi belum secara langsung menghasilkan partisi kelompok yang tegas (*crisp partition*) sehingga batas antar *cluster* masih perlu ditentukan lebih lanjut (Vesanto dkk., 2000). Kondisi tersebut menyebabkan hasil pemetaan SOM lebih sesuai digunakan sebagai representasi struktur data dibandingkan sebagai hasil akhir pengelompokan objek. Oleh karena itu, diperlukan metode *clustering* lanjutan yang mampu mengelompokkan neuron-neuron hasil SOM menjadi kelompok yang lebih jelas sehingga interpretasi terhadap karakteristik setiap *cluster* dapat dilakukan secara lebih objektif dan sistematis (Vesanto dkk., 2000).

Pendekatan *two-stage clustering* yang diperkenalkan oleh Vesanto dkk. (2000) merupakan salah satu pendekatan yang dapat digunakan untuk mengatasi keterbatasan SOM dalam menghasilkan partisi *cluster* yang tegas. Pada

pendekatan tersebut, SOM digunakan untuk membentuk *prototype vectors* pada tahap pertama, kemudian *codebook vectors* hasil pemetaan dikelompokkan kembali menggunakan metode *clustering* pada tahap kedua, seperti *hierarchical clustering* maupun K-Means. Pendekatan ini tidak membatasi penggunaan metode tertentu pada tahap pengelompokan *codebook vectors*, sehingga metode *clustering* lanjutan dapat dipilih sesuai dengan karakteristik data yang dianalisis (Vesanto dkk., 2000). Dalam penelitian ini, metode K-Medoids dipilih sebagai tahap kedua karena menggunakan medoid, yaitu objek representatif yang dipilih langsung dari data, sebagai pusat setiap *cluster* (Kaufman & Rousseeuw, 1990). Berbeda dengan K-Means yang menggunakan nilai rata-rata (*centroid*) sebagai pusat *cluster*, K-Medoids relatif lebih tahan terhadap pengaruh observasi bernilai ekstrem (*outlier*) (Han dkk., 2011). Karakteristik tersebut dinilai sesuai dengan data dampak bencana hidrometeorologi yang umumnya memiliki sebaran tidak seimbang serta mengandung observasi dengan nilai yang relatif ekstrem. Selain itu, penerapan K-Medoids sebagai metode pengelompokan *codebook vectors* hasil SOM pada analisis data kebencanaan masih relatif jarang ditemukan sehingga pendekatan ini diharapkan dapat memberikan alternatif dalam menghasilkan partisi *cluster* yang lebih representatif tanpa menghilangkan struktur topologi yang telah dibentuk pada tahap pemetaan menggunakan SOM.

Berdasarkan uraian yang telah dipaparkan, penelitian ini menerapkan pendekatan *two-stage clustering* dengan mengombinasikan SOM dan K-Medoids untuk mengelompokkan provinsi di Indonesia berdasarkan frekuensi kejadian dan dampak bencana hidrometeorologi. SOM digunakan pada tahap awal untuk memetakan pola kemiripan data multidimensi sekaligus mempertahankan hubungan topologi antarobjek, sedangkan K-Medoids diterapkan pada *codebook vectors* hasil SOM untuk membentuk partisi *cluster* yang lebih jelas sesuai dengan konsep *two-stage clustering* yang dikemukakan oleh (Vesanto dkk., 2000). Penelitian ini menawarkan penerapan K-Medoids sebagai metode pengelompokan *codebook vectors* hasil SOM pada analisis dampak bencana hidrometeorologi antarprovinsi di Indonesia. Pendekatan tersebut masih relatif jarang diterapkan pada penelitian kebencanaan sehingga diharapkan dapat menjadi alternatif dalam

memperoleh hasil pengelompokan yang lebih representatif. Kualitas hasil pengelompokan dievaluasi menggunakan *silhouette coefficient*, *dunn index*, dan *davies-bouldin index* untuk menentukan konfigurasi *cluster* yang paling representatif. Hasil penelitian diharapkan memberikan informasi mengenai karakteristik kelompok provinsi berdasarkan frekuensi kejadian dan dampak bencana hidrometeorologi yang dapat digunakan sebagai bahan pertimbangan dalam penyusunan strategi mitigasi, penanggulangan bencana, serta perumusan kebijakan pembangunan yang lebih tepat sasaran.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan sebelumnya, maka rumusan masalah untuk penelitian ini adalah sebagai berikut:

1. Bagaimana karakteristik frekuensi kejadian dan dampak bencana hidrometeorologi antarprovinsi di Indonesia?
2. Bagaimana pola kemiripan antarprovinsi berdasarkan frekuensi kejadian dan dampak bencana hidrometeorologi yang dihasilkan menggunakan metode SOM?
3. Bagaimana hasil pengelompokan provinsi di Indonesia berdasarkan frekuensi kejadian dan dampak bencana hidrometeorologi menggunakan pendekatan *two-stage clustering* melalui SOM dan *K-Medoids*?
4. Bagaimana karakteristik masing-masing *cluster* yang dihasilkan dari pengelompokan provinsi berdasarkan frekuensi kejadian dan dampak bencana hidrometeorologi di Indonesia?

1.3 Batasan Masalah

Penelitian ini masih terdapat beberapa batasan masalah diantaranya adalah sebagai berikut:

1. Pembentukan *cluster* dilakukan menggunakan pendekatan *two-stage clustering*, yaitu algoritma SOM sebagai proses pemetaan data dan algoritma K-Medoids untuk mengelompokkan *codebook vectors* hasil SOM.

2. Penelitian ini menggunakan *Euclidean Distance* sebagai ukuran jarak pada proses pembentukan *Best Matching Unit* (BMU) dalam algoritma SOM sesuai dengan algoritma standar yang dikembangkan oleh Kohonen.
3. Penentuan *hyperparameter* pada SOM dilakukan berdasarkan hasil eksperimen dalam penelitian ini, yaitu menggunakan *learning rate* awal sebesar 0,05 dan *learning rate* akhir sebesar 0,01, jumlah iterasi (*rlen*) sebanyak 500, serta topologi heksagonal (*hexagonal topology*).

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah dan batasan masalah diantaranya adalah sebagai berikut:

1. Mendeskripsikan karakteristik frekuensi kejadian dan dampak bencana hidrometeorologi antarprovinsi di Indonesia.
2. Mengidentifikasi pola kemiripan antarprovinsi berdasarkan frekuensi kejadian dan dampak bencana hidrometeorologi menggunakan metode SOM.
3. Mengelompokkan provinsi di Indonesia berdasarkan frekuensi kejadian dan dampak bencana hidrometeorologi menggunakan pendekatan *two-stage clustering* melalui metode SOM dan *K-Medoids*.
4. Mendeskripsikan karakteristik masing-masing *cluster* hasil pengelompokan provinsi berdasarkan frekuensi kejadian dan dampak bencana hidrometeorologi di Indonesia.

1.5 Manfaat Penelitian

Hasil penelitian ini diharapkan dapat memberikan kontribusi dalam memperkaya kajian penerapan metode *unsupervised learning*, khususnya melalui pendekatan *two-stage clustering* yang SOM dan *K-Medoids* dalam analisis data kebencanaan. Penelitian ini juga diharapkan dapat menjadi referensi bagi penelitian selanjutnya yang berkaitan dengan penerapan metode *Machine Learning* dan analisis *clustering* untuk mengidentifikasi pola kemiripan data multidimensi, khususnya pada bidang kebencanaan. Selain itu, hasil penelitian ini diharapkan

dapat memberikan informasi mengenai karakteristik kelompok provinsi berdasarkan frekuensi kejadian dan dampak bencana hidrometeorologi di Indonesia sehingga dapat menjadi bahan pertimbangan bagi pemerintah, Badan Nasional Penanggulangan Bencana (BNPB), pemerintah daerah, dan pemangku kepentingan lainnya dalam menyusun strategi mitigasi, adaptasi, serta penanggulangan bencana yang lebih tepat sasaran sesuai dengan karakteristik masing-masing kelompok wilayah. Hasil penelitian ini juga diharapkan dapat mendukung penyediaan informasi berbasis data dalam perencanaan kebijakan penanggulangan bencana serta berkontribusi terhadap upaya pencapaian *Sustainable Development Goals* (SDGs), khususnya Tujuan 13 (*Climate Action*) target 13.1 yang berfokus pada penguatan ketahanan (*resilience*) dan kapasitas adaptasi terhadap bencana yang berkaitan dengan perubahan iklim.

