

# BAB I

## PENDAHULUAN

Bab ini berisikan langkah awal dalam penyusunan skripsi yang meliputi latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, metode penelitian, dan sistematika penulisan.

### 1.1 Latar Belakang Masalah

Data menjadi satu acuan penting dalam menentukan suatu keputusan. Keputusan yang bersumber dari data ini seringkali dianggap sebagai keputusan yang tepat karena terdapat perhitungan matematis yang menyertainya. Sehingga tidak heran jika berbagai bidang telah menyadari manfaat akan hal tersebut. Akibatnya, banyak orang mulai bermunculan dalam mengeksplorasi mencari solusi terbaik, mengembangkan ilmu yang lebih efektif, dan melakukan pembaharuan teknik-teknik tertentu dalam menganalisis data.

Salah satu analisis data yang biasa diterapkan dalam mengelompokkan berdasarkan ukuran kemiripan yaitu analisis pengelompokan atau *cluster*. Metode analisis *cluster* atau pengelompokan memiliki beragam jenis metode sesuai dengan pendekatannya masing-masing. Umumnya metode ini terbagi menjadi 2 bagian yaitu metode hierarki dan metode non-hierarki. Salah satu dari jenis metode non-hierarki yang dapat melakukan pembelajaran tanpa pengawasan atau algoritma harus mencari pola untuk menyimpulkan hasil akhir dari data yang tidak diberi label sebelumnya adalah metode *Self Organizing Maps*.

Gagasan pada metode ini terletak pada pemetaan topologi data berdimensi tinggi ke unit peta melalui pembelajaran mandiri dimana jarak antara titik data dipertahankan. *Output* visualisasi yang mudah dan menarik menjadi salah satu keunggulan dalam metode ini. Metode yang ditemukan oleh Profesor Teuvo Kohonen pada tahun 1980-an sampai saat ini masih eksis dipakai dan terus dikaji lebih lanjut. Informasi yang didapatkan lebih dari 5000 publikasi telah dilaporkan

di literatur terbuka, dan banyak proyek komersial menggunakan *Self Organizing Maps* sebagai alat untuk memecahkan masalah dunia nyata yang sulit [1].

Melakukan analisis data tentu tak luput dari tantangan yang dihadapinya. Salah satu tantangan dalam menganalisis sebuah data yaitu menangani adanya permasalahan pencilan. Metode *Self Organizing Maps* tidak mampu menunjukkan kinerja yang terbaik dalam melakukan analisis *cluster* [2], [3], [4]. Pendekatan melalui metode untuk menangani permasalahan ini telah dikembangkan dalam beberapa tahun terakhir termasuk *Self Organizing Maps* (SOM). Namun, dalam literatur hanya sedikit yang membahas masalah pencilan di SOM padahal penggunaan metode SOM yang efektif sangat aplikatif untuk mengatasi permasalahan dalam berbagai bidang misalnya, pengenalan pola, robotika, pemrosesan informasi semantik, analisis dan pengendalian industri, telekomunikasi, analisis biomedis, dan keuangan.

Untuk itu pendekatan menanggulangi efek negatif dari pencilan data dalam menangani masalah ini ditawarkan oleh jurnal [5] melalui metode *Smoothed Self Organizing Maps* sehingga menghasilkan cluster yang *robust*. Karena metode ini merupakan metode baru, maka perlu ditinjau lebih lanjut terkait akibat atau pengaruh dari model tersebut. Hal ini dapat dilihat dengan cara mengevaluasi kinerja dari metode *Smoothed Self Organizing Maps* dan meyakinkan diri benarkah tujuan yang diinginkan dari *cluster* yang *robust* dapat tercapai atau tidak.

## 1.2 Rumusan Masalah

Penelitian ini berfokus pada analisis *cluster* menggunakan metode *Self Organizing Maps*. Alasan metode tersebut digunakan karena metode ini sedikit berbeda dengan metode analisis *cluster* yang lainnya. Hasil pengelompokkan dari metode *Self Organizing Maps* ini berupa *grid* yang sangat memudahkan dalam mengamati kemiripan data. Sayangnya metode ini tidak begitu baik jika terdapat pencilan di dalamnya. Hal ini dapat diperlihatkan melalui nilai kesalahan yang tinggi apabila terdapat pencilan. Akibatnya hasil pengelompokkan atau akurasinya menjadi tidak valid jika memiliki kesalahan terlalu tinggi.

Pendekatan menggunakan metode *Smoothed Self Organizing Maps* yang hadir pada tahun 2020 diharapkan mampu menjadi alternatif solusi dari permasalahan tersebut. Oleh karenanya, permasalahan penelitian yang akan dijawab dalam penelitian ini dirumuskan dalam pertanyaan penelitian sebagai berikut:

1. Bagaimana data pencilan dapat mempengaruhi analisis *cluster*?
2. Bagaimana metode *Self Organizing Maps* dapat dikatakan memiliki kinerja yang buruk dalam melakukan analisis *cluster* yang mengandung data pencilan?
3. Seberapa jauh metode *Smoothed Self Organizing Maps* dapat memiliki kinerja yang lebih baik dibanding dengan metode *Self Organizing Maps* dalam melakukan analisis *cluster* yang mengandung data pencilan?

### 1.3 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Data yang digunakan dalam penelitian ini adalah data simulasi yang memenuhi asumsi adanya pencilan dan melakukan penerapan menggunakan data covid-19 di Jawa Barat.
2. Jenis pencilan yang dipergunakan yaitu pencilan *cluster* dan radial
3. Menggunakan pendekatan jarak *euclid*
4. Inisialisasi bobot yang acak
5. Memakai fungsi gauss dalam pendekatan fungsi tetangga
6. Angka pembelajaran menggunakan fungsi eksponen
7. Tidak menjelaskan optimum *cluster* terbaik

### 1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Menganalisis pengaruh data pencilan terhadap analisis cluster
2. Mengkaji metode *Self Organizing Maps* dan memperlihatkan kinerja dalam melakukan analisis *cluster* yang mengandung data pencilan

3. Mengkaji metode *Smoothed Self Organizing Maps* dan memperlihatkan kinerja dalam melakukan analisis *cluster* yang mengandung data pencilan
4. Memperlihatkan perbedaan hasil kinerja kedua metode dalam melakukan analisis *cluster* yang mengandung data pencilan

### 1.5 Metode Penelitian

Metode penelitian yang dilakukan dalam penelitian ini yaitu sebagai berikut:

1. Melakukan studi literatur, yaitu proses dimana penulis akan mengevaluasi lebih lanjut metode *Smoothed Self Organizing Maps* dalam mengatasi data pencilan pada analisis *cluster* dengan mengkaji sumber pustaka berupa buku, jurnal dan skripsi yang berkaitan.
2. Melakukan simulasi menggunakan bahasa R.

### 1.6 Sistematika Penulisan

Sistematika penulisan dalam Skripsi ini terdiri atas 5 bab, daftar pustaka, riwayat hidup, dan lampiran, dimana pada setiap bab terdapat beberapa sub bab. Lebih jelasnya terangkum sebagai berikut:

**BAB I : PENDAHULUAN**

Bab ini berisikan latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, metode penelitian, dan sistematika penulisan.

**BAB II : LANDASAN TEORI**

Bab ini berisikan teori-teori yang menjadi dasar dalam penelitian ini. Secara keseluruhan dalam bab ini menjelaskan terkait data pencilan, data multivariat, analisis cluster, kuantisasi vektor, validasi analisis cluster, dan bahasa R

**BAB III : METODE SMOOTHED SELF ORGANIZING MAPS  
DALAM ANALISIS CLUSTER YANG  
MENGANDUNG DATA PENCILAN.**

Bab ini berisikan inti permasalahan yang akan dibahas meliputi, metode *self organizing maps*, metode *smoothed self organizing maps*, pengaruh data pencilan dalam analisis *cluster*, metode *self organizing maps* dalam data pencilan, dan algoritma metode dalam bahasa R.

**BAB IV : STUDI KASUS DAN ANALISIS**

Bab ini berisikan simulasi dan penerapan kasus *real* dari metode *Self Organizing Maps* dan *Smoothed Self Organizing Maps*.

**BAB V : KESIMPULAN DAN SARAN**

Bab ini berisikan kesimpulan yang menjawab dari apa yang menjadi tujuan dalam penulisan studi literatur dan saran sebagai langkah pengembangan penelitian ke depan atau sebagai pembandingan terhadap apa yang sudah didapat.