

BAB I

PENDAHULUAN

1.1. Latar Belakang

Kesejahteraan merupakan titik ukur bagi suatu masyarakat bahwa telah berada pada kondisi yang mapan, Kesejahteraan tersebut dapat diukur dari kesehatan, keadaan ekonomi, kebahagiaan dan kualitas hidup rakyat[1]. Kesejahteraan identik dengan kemiskinan karena kemampuan sebuah keluarga dalam memenuhi kebutuhan pokok dalam kehidupan sehari-harinya. berdasarkan data BPS (2023), Kemiskinan pada Jawa Timur masih terdapat sekitar 4.188.810 orang penduduk di Jawa Timur tergolong sebagai penduduk miskin. kemiskinan menjadi suatu permasalahan yang sulit diatasi karena permasalahan tersebut saling berkaitan satu sama lain. Tidak hanya berkaitan dengan masalah rendahnya tingkat pendapatan dan konsumsi, kemiskinan juga berkaitan dengan kesehatan dan tidak berdaya masyarakat untuk berpartisipasi dalam pembangunan, rendahnya tingkat pendidikan, serta berbagai masalah lain yang berkaitan dengan pembangunan manusia. Oleh karena itu, kemiskinan dapat berdampak pada kondisi kesejahteraan masyarakat secara langsung.

Dalam upaya meningkatkan kesejahteraan dan mengentaskan kemiskinan masyarakat, pemerintah berfokus pada pembangunan ekonomi, peningkatan infrastruktur, dan pemberdayaan sumber daya untuk menciptakan peluang kerja dan kesejahteraan pada kalangan masyarakat. Memahami tingkat kesejahteraan sosial dapat membantu pemerintah dalam merumuskan keputusan dan strategi sebagai landasan meningkatkan kualitas hidup masyarakat. Untuk membuat keputusan yang lebih tepat dan terarah maka harus dilakukan *clustering* wilayah berdasarkan tingkat kesejahteraan. Hal tersebut menjadi kunci dalam meningkatkan pembangunan serta alokasi sumber daya yang tepat sesuai dengan kebutuhan dan karakteristik sebuah daerah. Salah satu teknik yang dapat digunakan dalam *clustering* daerah adalah *clustering*.

Clustering merupakan proses untuk mengelompokkan data sehingga objek yang berada dalam klaster-klaster yang memiliki karakteristik yang mirip [2]. Dalam pengklasteran sendiri banyak metode yang dapat digunakan, salah satunya adalah

Self-Organizing Map (SOM). *Self-Organizing Map* (SOM) salah satu metode pengklasteran *unsupervised learning* dengan konsep *Artificial Neural Network* dengan hasil yang cukup valid. Algoritma SOM memiliki kelebihan dalam memetakan data dengan dimensi yang tinggi ke dalam dimensi yang lebih rendah menjadi 2 dimensi. Selain itu SOM mampu menangani relasi non-linear dalam data, yang seringkali tidak dapat diatasi dengan metode *clustering* tradisional seperti K-Means.

Setelah *cluster* telah didapatkan, Hasil *clustering* yang didapat harus diketahui kualitasnya agar analisis lebih tepat dan mengetahui kualitas *cluster* yang terbentuk. Salah satu *cluster* yang dapat digunakan dalam mengukur hasil klaster adalah *Silhouette coefficient*. Dimana *Silhouette Coefficient* akan mengevaluasi penempatan setiap objek dalam setiap klaster dengan membandingkan jarak rata-rata objek dalam satu klaster dan jarak antara objek dengan klaster yang lainnya. Nilai yang tinggi menunjukkan bahwa objek tersebut cocok dengan *cluster*, sementara nilai rendah menunjukkan objek hasil *cluster* yang kurang baik [3].

Penelitian terhadap proses *clustering* dengan metode SOM, data indikator kesejahteraan masyarakat dan algoritma *Self-Organizing Map* telah banyak dilakukan seperti penelitian yang dilakukan oleh Nurul Iman Dkk dalam melakukan *clustering* kabupaten/kota berdasarkan data indikator kesejahteraan rakyat provinsi Nusa Tenggara Timur menggunakan algoritma *self-organizing map*. Berdasarkan hasil penelitian *clustering* menghasilkan 4 *cluster* yang memiliki karakteristik berbeda-beda. *Cluster* 1 terdiri dari 16 kabupaten dengan aspek menonjol seperti umur harapan hidup, jumlah penduduk, tingkat pengangguran terbuka, dan angka buta huruf. *Cluster* 2, dengan 3 kabupaten, mirip dengan *Cluster* 1 tetapi memiliki nilai lebih rendah pada aspek-aspek tersebut. *Cluster* 3, terdiri dari 2 kabupaten, menunjukkan potensi perkembangan menjadi kota di masa depan dengan variabel yang hampir sama besar tetapi tidak setinggi *Cluster* 4. *Cluster* 4 hanya terdiri dari Kota Kupang yang memiliki nilai dan indeks tinggi di hampir semua variabel penelitian[4].

Penelitian selanjutnya telah dilakukan oleh Irwan dkk. tentang penerapan algoritma *Self-Organizing Map* dalam mengelompokkan kabupaten berdasarkan

indikator kesejahteraan masyarakat. Variabel pada penelitian ini sebanyak 30 variabel dengan bentuk hasil survei yang dilakukan oleh badan pusat statistik. *Pre-processing* data peneliti menggunakan normalisasi *Min-Max* agar data ada pada rentang yang tidak terlalu jauh. Dari hasil pemrosesan *clustering* data dengan algoritma *Self-Organizing Map* dengan 5 *cluster* menunjukkan pembentukan *cluster* yang baik dengan hasil skor pada pengujian DBI yang sangat baik dengan nilai 0,17[5].

Penelitian terkait lainnya diteliti oleh Syarifuddin N dkk yang membahas penerapan algoritma *Self-Organizing Map* untuk mengelompokkan sekolah berdasarkan mutu pendidikan berdasarkan 31 atribut penilaian dengan 50 data. Hasil yang didapat pada penelitian ini hasil *cluster* sebanyak 4, dengan nilai *Davies Boldin Index* pada jenjang SD adalah 1,667, *learning rate* 0,3 dan MSE 0,00000067. *Cluster* 1 terdiri dari 42 sekolah, *cluster* 2 sebanyak 9 sekolah, *cluster* 3 sebanyak 5 sekolah dan *cluster* 4 sebanyak 4 sekolah[6].

Penelitian selanjutnya dilakukan oleh Mujiati Dwi Kartika Sari dalam mengelompokkan provinsi berdasarkan konsumsi pangan. Dari penelitian yang telah dilakukan, analisis data konsumsi pangan di Indonesia menunjukkan bahwa jumlah *cluster* optimum adalah 4, ditentukan berdasarkan nilai *Davies-Bouldin Index* terkecil. Penggunaan algoritma *Self-Organizing Map* mengelompokkan wilayah menjadi *cluster* 1 dengan 22 provinsi, *cluster* 2 dengan 10 provinsi, dan *cluster* 3 serta *cluster* 4 masing-masing dengan 1 provinsi. Hasil *clustering* ini mengungkapkan bahwa provinsi dalam *cluster* 2, 3, dan 4 memiliki rata-rata kecukupan kalori penduduk yang masih di bawah standar permenkes No. 75 Tahun 2013 [7].

Berdasarkan tinjauan penelitian yang telah dilakukan dapat disimpulkan bahwa algoritma *Self-Organizing Map (SOM)* merupakan metode yang efektif dalam proses klasterisasi data dengan dimensi yang cukup tinggi, terutama ketika dievaluasi menggunakan *Davies-Bouldin Index (DBI)*. Nilai *DBI* yang rendah dari hasil pengujian dalam berbagai studi menunjukkan bahwa *SOM* mampu menghasilkan klaster dengan kualitas tinggi, dimana klaster tersebut memiliki pemisahan yang baik dan struktur internal yang kompak. Namun pada peninjauan

penelitian terdapat kekurangan yaitu data yang digunakan pada penelitian sebagian tidak mengalami tahapan *pre-processing* terutama normalisasi nilai dan juga uji *outlier*. Padahal *pre-processing* data nantinya akan berpengaruh terhadap kualitas hasil pengklasteran. Pengujian pada penelitian juga sangat bervariasi namun peneliti rasa validasi metode yang ada perlu adanya teknik pengujian yang dalam memvalidasi hasil *cluster*. Maka dari itu, peneliti mengusulkan validasi atau uji hasil *cluster* dengan *Silhouette Coefficient* dalam penerapan algoritma *Self-Organizing Map (SOM)* untuk mendapatkan hasil *cluster* yang lebih optimal.

Pada penelitian ini terdapat beberapa indikator yang digunakan dalam proses *clustering* kabupaten/kota. Indikator tersebut meliputi kependudukan, kesehatan, pendidikan, ketenagakerjaan, konsumsi masyarakat, perumahan dan lingkungan, dan kemiskinan. Masing-masing indikator akan dipilih variabel yang berpengaruh dalam indikator kesejahteraan sosial. Beberapa variabel diantaranya adalah Jumlah penduduk, Presentase pemilik jaminan kesehatan, presentase angka melek huruf, rata-rata lama sekolah, angkatan kerja, tingkat pengangguran terbuka, pengeluaran per kapita untuk makanan, akses terhadap air yang layak, akses terhadap sanitasi layak, presentase penduduk miskin. Dengan penerapan metode *Self Organizing Map (SOM)* dan pemanfaatan berbagai indikator kesejahteraan, penelitian ini diharapkan dapat mengidentifikasi dan *clustering* wilayah di Provinsi Jawa Timur berdasarkan tingkat kesejahteraan dengan algoritma *Self Organizing Map (SOM)* untuk mengetahui performa algoritma tersebut dengan *Silhouette coefficient*. Penelitian ini juga diharapkan dapat memberikan rekomendasi data yang jelas atas kondisi sosial-ekonomi di berbagai wilayah, sehingga mempermudah Dinas Sosial Provinsi Jawa Timur dalam alokasi sumber daya, serta program pemberdayaan masyarakat ke wilayah-wilayah yang membutuhkan penanganan prioritas.

1.2. Rumusan Masalah

1.2.1. Permasalahan

Provinsi Jawa Timur, dengan keragaman karakteristik sosial dan ekonomi di setiap kabupaten/kota, menghadapi tantangan dalam *clustering* wilayah berdasarkan indikator kesejahteraan rakyat. Kompleksitas ini menjadi perhatian

utama dalam penyusunan kebijakan yang lebih tepat sasaran untuk meningkatkan kualitas hidup penduduk di berbagai wilayah. Oleh karena itu, diperlukan metode yang efektif untuk mengidentifikasi perbedaan karakteristik kesejahteraan antar daerah guna mendukung pengambilan keputusan yang lebih terarah dan tepat.

1.2.2. Metode Usulan

Metode yang diusulkan dalam penelitian ini yaitu *Self-Organizing Map* dengan validasi hasil *cluster* menggunakan *Silhouette Coefficient* yang dianggap sesuai dalam penyelesaian masalah dalam megelompokkan kabupaten/kota berdasarkan indikator kesejahteraan sosial.

1.2.3. Pertanyaan Penelitian

Berdasarkan latar belakang, maka rumusan masalah yang akan di kaji pada penelitian ini yakni :

1. Berapa jumlah *cluster* yang optimal untuk mengelompokkan wilayah berdasarkan indikator kesejahteraan rakyat menggunakan algoritma *Self-Organizing Map* ?
2. Berapa hasil *Silhouette Coefficient* yang optimal dan parameter optimal dari penerapan algoritma *Self-Organizing Map* untuk *clustering* pada data kabupaten/kota berdasarkan indikator kesejahteraan rakyat ?

1.3. Tujuan dan Manfaat

1.3.1. Tujuan Penelitian

Berdasarkan pada pertanyaan dari penelitian, maka tujuan dari penelitian ini yakni :

1. Untuk mengetahui jumlah *cluster* yang optimal yang dihasilkan oleh algoritma *Self-Organizing Map* berdasarkan data indikator kesejahteraan rakyat Provinsi Jawa Timur tahun 2021-2023.
2. Untuk mengetahui performa dari algoritma *Self-Organizing Map cluster* dalam mengelompokkan Kabupaten/Kota berdasarkan indikator kesejahteraan rakyat dengan *Silhouette coefficient*.

1.3.2. Manfaat Penelitian

Berdasarkan uraian pada permasalahan dan tujuan penelitian, maka manfaat yang akan didapat penelitian ini yaitu :

1. Peneliti dapat menambah wawasan mengenai penyelesaian *clustering* daerah dengan metode *Self-Organizing Map*.
2. Peneliti mengetahui performa dari algoritma *Self-Organizing Map* menggunakan *Silhouette coefficient* pada data indikator kesejahteraan rakyat tahun 2021-2023.
3. Membantu dinas terkait dalam memetakan kabupaten/kota berdasarkan tingkat kesejahteraan penduduknya.
4. Memberikan informasi yang berguna untuk dinas Sosial sebagai dasar pengambilan keputusan dalam mengoptimalkan peningkatan kesejahteraan masyarakat Jawa Timur.

1.4. Batasan Masalah

Beberapa batasan pada penelitian ini yaitu :

1. Dataset yang digunakan dalam 3 periode yaitu tahun 2021-2023 dengan jumlah kabupaten/kota pada provinsi Jawa Timur yang terdiri dari 38 kabupaten/kota
2. Data yang digunakan dalam penelitian ini diambil laman web data publikasi tahunan yang telah dilakukan oleh Badan Pusat Statistik (BPS) provinsi Jawa Timur dengan judul “Jawa Timur Dalam Angka” dan “statistik kesejahteraan rakyat provinsi Jawa Timur”. <https://jatim.bps.go.id/>
3. Variabel yang digunakan pada penelitian ini yaitu jumlah penduduk, presentase pemilik jaminan kesehatan, presentase angka melek huruf, rata-rata lama sekolah, angkatan kerja, tingkat pengangguran terbuka, pengeluaran per kapita untuk makanan, akses terhadap air minum yang layak, akses terhadap sanitasi layak, presentase penduduk miskin.

1.5. Sistematika Penulisan

Sistematika penulisan skripsi ini dibagi menjadi lima bab, yang masing-masing bab telah dirancang untuk tujuan tertentu, berikut penjelasan masing-masing bab:

BAB I : PENDAHULUAN

Pada bab ini yang membahas tentang latar belakang permasalahan kesejahteraan dan, perumusan masalah, tujuan dan manfaat serta batasan masalah.

BAB II : KAJIAN PUSTAKA

Bab ini memaparkan teori-teori penunjang yang berkaitan dengan, *data mining*, *Self-organizing map*, *shilhouete Coeficient* kesejahteraan sosial, serta penelitian-penelitian terdahulu yang berhubungan dengan permasalahan yang diteliti dan menjadi acuan konseptual.

BAB III : METODE PENELITIAN

Pada bab ini mendeskripsikan mengenai metodologi penelitian seperti dataset yang digunakan, serta alur kerja sistem.

BAB IV : HASIL DAN PEMBAHASAN

Bab ini merupakan implementasi rancangan langkah-langkah yang telah dibuat pada bab 3 yaitu penerapan algoritma *Self-Organizing Map* dengan *pre-processing* dan evaluasi hasil atau skenario untuk melihat performa kode program yang telah dibuat

BAB IV : KESIMPULAN DAN SARAN

Bab ini menjelaskan temuan penelitian yang telah dilakukan pada bab 4 dan menjawab pertanyaan penelitian, memberikan informasi berdasarkan permasalahan yang ada dan memberikan saran atas penelitian yang telah dilakukan untuk kualitas penelitian yang lebih baik kedepannya.

DAFTAR PUSTAKA

Pada bab ini berisi tentang keterangan mengenai sumber rujukan yang digunakan dalam penyusunan tugas akhir. Pada penelitian ini daftar pustaka disusun secara otomatis dengan menggunakan aplikasi *mendeley*