

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kesehatan merupakan aspek penting bagi setiap orang, karena setiap orang berisiko mengalami gangguan kesehatan, terutama yang rentan terhadap berbagai penyakit. Menurut Fertman & Allensworth (2010), kesehatan adalah keadaan sejahtera secara fisik, mental, dan sosial secara utuh dan bukan hanya sekadar terbebas dari penyakit, kelemahan, atau cacat. Kesehatan adalah salah satu hal yang diinginkan oleh semua orang, karena tidak ada yang menginginkan hidup dalam kondisi tidak sehat. Penyakit sering kali datang tanpa diduga. Bahkan, seseorang yang terlihat sehat pun dapat mengidap penyakit secara tiba-tiba. Saat seseorang mengidap atau terserang suatu penyakit maka akan menghambat aktivitas yang positif, dan pada saat itulah seseorang semakin menyadari pentingnya menjaga kesehatan. Sebaliknya, ketika tubuh tidak menunjukkan tanda-tanda adanya kelelahan dan penyakit, biasanya seseorang cenderung mengabaikan kesehatannya dengan tidak memberikan waktu untuk istirahat yang cukup bagi tubuhnya.

Indonesia masih menghadapi tantangan besar dalam mengatasi penyakit menular yang ditandai dengan seringnya terjadi Kejadian Luar Biasa (KLB). Salah satu penyakit menular yang sering merenggut nyawa manusia adalah Tuberkulosis (TBC). Indonesia merupakan salah satu negara dengan angka penderita TBC yang tinggi, bahkan menempati peringkat kedua di dunia setelah India (Kementerian Kesehatan RI, 2024).

Masalah kesehatan di Indonesia memiliki tantangan besar karena penyakit menular tetap menjadi masalah yang belum terselesaikan. Permasalahan kesehatan

mengenai penyakit menular menunjukkan rendahnya pemahaman masyarakat Indonesia tentang pentingnya menjaga kesehatan. Jika tidak terkendali dengan baik, kasus penyakit menular berpotensi memicu Kejadian Luar Biasa (KLB). Untuk mengatasi hal ini, pemerintah telah melakukan berbagai upaya pencegahan penyakit menular, mulai dari peningkatan fasilitas kesehatan hingga pemberdayaan masyarakat melalui program edukasi kesehatan (Kementerian Pertahanan RI, 2014). Oleh karena itu, penyakit menular merupakan masalah yang sangat mendesak dan harus segera diatasi. Salah satu pendekatan yang dapat diterapkan adalah mengelompokkan provinsi di Indonesia berdasarkan karakteristik kesehatan yang hampir sama atau kemiripan. Pengelompokan ini bertujuan untuk memberikan informasi lebih spesifik mengenai kondisi kesehatan di setiap provinsi, sehingga strategi intervensi dapat disesuaikan dengan kebutuhan masing-masing wilayah. Salah satu cara yang dapat dilakukan dalam pengelompokan ini yaitu menggunakan analisis *cluster*.

Analisis *cluster* merupakan suatu teknik analisis statistik yang bertujuan untuk mengelompokkan sekumpulan objek ke dalam dua kelompok atau lebih berdasarkan perbedaan karakteristik yang dimiliki objek tersebut (Simamora, 2005). Analisis *cluster* dapat diterapkan dengan beberapa metode, yaitu metode analisis *cluster* hierarki dan non-hierarki. Metode *cluster* hierarki mengelompokkan observasi berdasarkan kesamaan dan karakteristiknya secara teratur. Dalam metode ini, jumlah kelompok belum ditentukan di awal melainkan terbentuk secara alami berdasarkan tingkat kesamaan antar objek (Mattjik & Sumertajaya, 2011). Metode *cluster* hierarki meliputi *Single Linkage*, *Complete Linkage*, *Average Linkage*, *Ward* dan *Centroid*. Metode non-hierarki mengelompokkan objek dimana jumlah

kelompok yang akan dibentuk dapat ditentukan terlebih dahulu sebagai bagian dari prosedur pengelompokan (Johnson & Wichern, 2007). Metode non-hierarki yaitu terdapat metode *K-Means* dan *Fuzzy C-Means*.

Beragamnya metode hierarki dan non-hierarki tentu akan menghasilkan solusi kelompok yang beragam pula, sehingga akan sulit menentukan solusi pengelompokan yang tepat dalam kondisi tersebut. Kesulitan dalam menentukan solusi kelompok yang tepat dari beragam metode hierarki dan non-hierarki dapat diatasi dengan menggabungkan solusi kelompok yang diperoleh dengan menggunakan metode *Cluster Ensemble*. *Cluster Ensemble* merupakan suatu metode pengelompokan objek dengan cara menggabungkan hasil pengelompokan dari beberapa metode yang berbeda untuk memperoleh suatu solusi gabungan sebagai solusi akhir (Strehl & Ghosh, 2002). Ide dasar dari *Cluster Ensemble* adalah menggabungkan sekumpulan hasil kelompok yang dibentuk atas dasar metode yang biasa digunakan, yaitu metode hierarki dan non-hierarki (Hu & Yoo, 2004).

Penelitian relevan yang dilakukan oleh Ristiyanti (2017) membandingkan kinerja metode *Cluster Hierarki*, *K-Means* dan *Cluster Ensemble* pada studi kasus Indikator Pelayanan Kesehatan Ibu Hamil. Berdasarkan penelitian ini, ditemukan hasil bahwa *Cluster Ensemble* merupakan metode yang paling tepat dibandingkan *Cluster Hierarki* dan *K-Means*. Solusi akhir yang dihasilkan dari *Cluster Ensemble* membentuk 3 kelompok dengan 34 objek provinsi di Indonesia. Selanjutnya pada penelitian lainnya yang dilakukan oleh Amelia (2023) membandingkan metode *Clustering Ensemble* dan *K-Means* pada pengelompokan provinsi Sumatera Barat berdasarkan Indeks Pembangunan Manusia. Berdasarkan penelitian tersebut, didapatkan hasil bahwa pengelompokan dengan menggunakan *Cluster Ensemble*

lebih optimal dibandingkan dengan *K-Means* untuk mengelompokan provinsi Sumatera Barat berdasarkan Indeks Pembangunan Manusia yang menghasilkan 2 kelompok dengan 19 objek kabupaten/kota di provinsi Sumatera Barat. Penelitian lainnya dilakukan oleh Yusfar et al. (2020) membandingkan kinerja metode Hierarki, Non-Hierakri dan *Cluster Ensemble* pada Indikator Kinerja Pembangunan Ekonomi Daerah. Berdasarkan penelitian tersebut didapatkan hasil bahwa *Cluster Ensemble* memiliki pembentukan kelompok yang lebih baik dari pada metode *Clustering* Hierarki dan Non-Hierarki. Solusi akhir yang dihasilkan dari *Cluster Ensemble* membentuk 3 kelompok dengan objek 24 kabupaten/kota di provinsi Sulawesi Selatan. Selain itu, penelitian yang dilakukan oleh Veriani (2020) tentang pengelompokan provinsi di Indonesia berdasarkan Variabel Penyakit Menular menggunakan metode *Complete Linkage*, *Average Linkage* dan *Ward*. Berdasarkan penelitian tersebut didapatkan hasil 6 kelompok. Penelitian lainnya yang dilakukan oleh Kuncheva et al (2006) membandingkan 24 metode *clustering* dengan metode *Cluster Ensemble* pada 24 dataset, yang mencakup *artificial* dan *real* dataset. Berdasarkan penelitian tersebut didapatkan hasil bahwa *Cluster Ensemble* memiliki pembentukan kelompok yang lebih baik dan akurat dari pada 24 metode *clustering* lainnya.

Berdasarkan pada uraian tersebut, penulis tertarik dalam melakukan penelitian mengenai analisis *cluster* dengan mengambil kasus penyakit yang sering menyerang kesehatan di Indonesia diantaranya Tuberkulosis, HIV, AIDS, Kusta, Campak, DBD, dan Malaria. Judul dalam penelitian ini adalah “Penerapan *Cluster Ensemble* dalam Pengelompokan Provinsi di Indonesia Berdasarkan Variabel Penyakit Menular”. Penelitian ini menerapkan metode *Cluster Ensemble* dan

evaluasi kualitas *cluster* dilihat dari nilai *Davies-Bouldin* (DB) yang minimum dan nilai *Dunn* yang paling besar.

1.2 Rumusan Masalah

Rumusan masalah yang dapat diidentifikasi berdasarkan latar belakang yang telah dijelaskan sebelumnya, yaitu sebagai berikut:

1. Bagaimana mendapatkan *cluster* terbaik dari metode *Cluster Ensemble*, *Single Linkage*, *Complete Linkage*, *Average Linkage*, *Centroid*, *Ward*, *K-Means*, dan *Fuzzy C-Means* berdasarkan nilai *Davies Bouldin Index* dan *Dunn Index*?
2. Bagaimana karakteristik *cluster* yang dihasilkan dari pengelompokan provinsi di Indonesia berdasarkan variabel penyakit menular dengan metode *Cluster Ensemble*?

1.3 Batasan Masalah

Batasan masalah yang digunakan pada penelitian ini yaitu:

1. Data yang digunakan pada penilitan ini difokuskan pada kasus penyakit menular tahun 2023 pada provinsi di Indonesia.
2. Metode statistika yang akan digunakan yakni metode *Cluster Ensemble*, *Single Linkage*, *Complete Linkage*, *Average Linkage*, *Centroid*, *Ward*, *K-Means*, dan *Fuzzy C-Means*.
3. Pada penelitian ini digunakan *cluster* sebanyak 2, 3, dan 4 *cluster*, karena untuk jumlah *cluster* > 4 akan menghasilkan ukuran *Binary Association Matrix* yang besar.

1.4 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah:

1. Mendapatkan hasil *cluster* terbaik dari metode *Cluster Ensemble*, *Single Linkage*, *Complete Linkage*, *Average Linkage*, *Centroid*, *Ward*, *K-Means*, dan *Fuzzy C-Means* berdasarkan nilai *Davies Bouldin Index* dan *Dunn Index*.
2. Menganalisis karakteristik *cluster* yang dihasilkan dari pengelompokan provinsi di Indonesia berdasarkan variabel penyakit menular dengan metode *Cluster Ensemble*.