

BAB I

PENDAHULUAN

1.1 Latar Belakang

HIV/AIDS merupakan penyakit menular yang disebabkan oleh infeksi virus HIV (*Human Immunodeficiency Virus*) yang menyerang sistem kekebalan tubuh. Infeksi tersebut menyebabkan penderita mengalami penurunan ketahanan tubuh sehingga sangat mudah untuk terinfeksi berbagai macam penyakit lain yang disebut dengan AIDS (*Acquired Immunodeficiency Syndrome*) (Nafsiah Mboi, 2017). Berdasarkan Laporan Perkembangan HIV AIDS pada tahun 2020, Provinsi Jawa Barat merupakan salah satu provinsi yang memiliki jumlah kasus HIV tertinggi ke-2 pada tahun 2020 periode januari-desember dengan jumlah ODHA 41.987 yang ditemukan setelah provinsi Jawa Timur (7.157), Jawa Barat (5.666), Jawa Tengah (5.531) dan DKI Jakarta (4.931), jumlah ODHA sebagian besar terdapat pada kelompok umur 25 - 49 tahun (68,7%) dan berjenis kelamin laki-laki (68%), lalu berdasarkan faktor resiko sebanyak 24,9% homoseksual yang merupakan kelompok populasi LSL (23,6%) dan Waria (1,3%). Jumlah kumulatif ODHA ditemukan (kasus HIV) yang dilaporkan sampai dengan Desember 2020 sebanyak 419.551 orang, sedangkan jumlah kumulatif kasus AIDS yang dilaporkan sampai dengan Desember 2020 sebanyak 129.740. Dari laporan perkembangan HIV diatas adanya peningkatan jumlah HIV/AIDS dibandingkan dengan laporan triwulan IV pada tahun 2019 dari jumlah 121.101 orang menjadi 129.740 orang (Laporan Triwulan IV, 2020).

Secara nasional, Jawa Barat termasuk provinsi kedua dengan jumlah kasus HIV/AIDS tertinggi pada tahun 2020-2021. Hal ini menunjukkan bahwa HIV/AIDS telah menjadi ancaman bagi masyarakat luas karena penyakit ini mengancam nyawa penderitanya yang menimbulkan dampak negatif berupa masalah fisik, psikologis, social, spiritual yang membuat pengidap HIV menjalani kehidupan yang penuh tekanan. Jumlah angka kematian HIV/AIDS di Jawa Barat dari tahun 2019 sampai 2020 dengan jumlah 0,59%, lalu dari tahun 2020 sampai 2021 dengan jumlah persentase 0,79%.

Berdasarkan data persentase diatas, maka diperlukan teknik komputer *machine learning* untuk mengekstrak informasi penting pada data yang disebut dengan data mining. Data *mining* memiliki metode berbasis data tersendiri yaitu *scoring*, *clustering*, prediksi, klasifikasi dan asosiasi. Salah satu teknik yang dapat digunakan untuk memetakan data adalah teknik *clustering*. *Clustering* adalah teknik yang digunakan dalam data mining yang mengelompokkan data ke dalam *cluster* berdasarkan kesamaan. Terdapat beberapa algoritma yang tersedia, salah satunya adalah *k-means* yang digunakan untuk mengelompokkan suatu *dataset*. (Murniasih et al., 2020).

Mengenai penelitian *clustering* data, antara lain Hayati., et al (2021) tentang penerapan algoritma analisis *clustering K-Means* pada pasien HIV/AIDS studi kasus di Kabupaten Banjar. Hasil penelitian menunjukkan bahwa algoritma *K-Means* merupakan metode *clustering* data *non-hierarki* yang bertujuan untuk membagi data yang ada menjadi satu atau lebih cluster, dengan *cluster* pertama terdiri dari 90% laki-laki dan sisanya 10% adalah perempuan. Oleh karena itu, dengan menggunakan metode *k-means clustering* dapat diperoleh gambaran masing-masing kecamatan di Kabupaten Banjar untuk perumusan kebijakan dan strategi ke depan. Penelitian yang lain dilakukan oleh Yoga Nugraha (2022) tentang penerapan algoritma *K-Means clustering* pada penyakit menular pada manusia pada studi kasus di Kabupaten Karawang menggunakan algoritma *K-Means* dengan metode *elbow*. Hasil penelitian yang dilakukan menunjukkan hasil algoritma *clustering K-Means* dengan metode *Elbow* dapat menghasilkan 3 *cluster* optimal dalam pengelompokan penyakit menular pada manusia di Kabupaten Karawang pada tahun 2020. Hasil pengelompokan dibagi menjadi 2 *cluster*, dimana 3 cluster merupakan *cluster* yang nilai dropnya paling signifikan atau paling tinggi, sehingga 3 cluster merupakan *cluster* terbaik.

Berdasarkan riwayat penelitian sebelumnya menunjukkan bahwa algoritma *k-means* dapat diimplementasikan untuk mengelompokkan data menurut kategori yang sama ke dalam *cluster* yang sama dan kategori yang berbeda ke dalam *cluster* lainnya. Evaluasinya menggunakan metode siku, yaitu menentukan jumlah *cluster* terbaik dengan melihat persentase hasil

perbandingan antara jumlah *cluster* yang membentuk siku pada titik (Irwanto, et al 2012). Berdasarkan masalah di atas, peneliti perlu melakukan penelitian untuk mengetahui kelompok penularan HIV/AIDS salah satu upaya yang dapat membantu pemerintah dalam menurunkan jumlah kasus HIV/AIDS di Provinsi Jawa Barat. Tujuan dari penelitian ini adalah untuk mengetahui pemetaan tingkatan *cluster* kasus HIV/AIDS di Jawa Barat berdasarkan parameter tahun dengan menggunakan *tools* pengolahan data *RapidMiner* pada *cluster* kabupaten/kota, sehingga dapat menjadi salah satu upaya membantu pemerintah menurunkan jumlah kasus HIV/AIDS di Jawa Barat.

1.2 Rumusan Masalah

Berdasarkan penjabaran latar belakang di atas, maka rumusan masalah yang dibuat penulis adalah :

1. Bagaimana penerapan algoritma *K-Means* dengan evaluasi metode *elbow* untuk *clustering* penyakit HIV/AIDS di Jawa Barat?
2. Bagaimana gambaran tingkatan *cluster* mengenai HIV/AIDS di Jawa Barat berdasarkan parameter tahun dengan menggunakan *tools Rapidminer*?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah tersebut, maka tujuan penelitian ini adalah:

1. Menerapkan algoritma *K-Means* dan mengetahui hasil pengklasteran dari penerapan metode *elbow* untuk pengklasteran kasus HIV di Jawa Barat.
2. Mengetahui tingkatan *cluster* berdasarkan parameter tahun pada kasus HIV/AIDS di Jawa Barat.

1.4 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut :

1. Teoritis
 - a. Mampu mengetahui proses penyelesaian masalah penyakit menular HIV dengan menggunakan algoritma *K-Means clustering* agar menemukan pengelompokan *cluster* yang relevan.

- b. Mampu menganalisis dan membahas bagaimana cara kerja algoritma K-Means clustering yang diusulkan.
- c. Menggambarkan hasil cluster dengan menggunakan *python* dan metode *Elbow* dalam pengklasteran kasus HIV di Jawa Barat.

2. Praktis

- a. Bagi masyarakat dan lingkungan, meningkatkan wawasan dan pengetahuan masyarakat tentang deteksi dini penyakit HIV yang nantinya diharapkan masyarakat dapat kooperatif dan mengetahui hal apa saja yang harus diperhatikan dalam deteksi dini penyakit HIV.
- b. Bagi pemerintah, gagasan ini dapat membantu pemerintah melakukan penanganan yang lebih efektif terhadap penyakit HIV/AIDS. Pemerintah nantinya dapat lebih fokus terlebih dahulu ke daerah-daerah yang memiliki kasus HIV tingkat tinggi, sehingga jumlah kasus HIV di Provinsi Jawa Barat dapat semakin berkurang.
- c. Bagi pembaca dan penulis, gagasan ini dapat menambah pengetahuan mengenai algoritma *K-Means Clustering* mengenai pengelompokan daerah yang terkena kasus HIV di Jawa Barat.

