

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada 31 Desember 2019, WHO China Country Office melaporkan kasus pneumonia yang tidak diketahui penyebabnya di Kota Wuhan, Provinsi Hubei, China. Pada tanggal 7 Januari 2020, China mengidentifikasi pneumonia yang tidak diketahui penyebabnya tersebut sebagai jenis baru coronavirus (novel coronavirus) (Safrizal *et al.*, 2020). Coronavirus disebut dengan virus zoonotik yaitu virus yang ditransmisikan dari hewan ke manusia. Berdasarkan penelitian dari 59 kasus yang diduga, 41 pasien di Wuhan dipastikan terinfeksi coronavirus (2019-nCoV) (Yuliana, 2020).

World Health Organization (WHO) memberi nama virus baru tersebut *Severe Acute Respiratory Syndrome Coronavirus-2* (SARS-CoV-2) dan resmi menetapkan penyakit tersebut dengan nama Novel Coronavirus pada manusia dengan sebutan Coronavirus Diseases 2019 atau yang lebih dikenal dengan sebutan Covid-19 (Safrizal *et al.*, 2020). Ketika menyerang manusia, virus ini menyebabkan infeksi saluran pernafasan seperti flu, hingga penyakit yang lebih berat seperti *Middle East Respiratory Syndrome* (MERS) dan *Severe Acute Respiratory Syndrome* (SARS) (Setiawan, 2020). Sampai saat ini sudah dipastikan 222 Negara yang telah terjangkit, dengan kasus konfirmasi 127.877.462 dengan jumlah kematian 2.796.561 (WHO, 2021).

Covid-19 pertama dilaporkan masuk ke Indonesia pada tanggal 2 Maret 2020 dengan jumlah terkonfirmasi positif sebanyak 2 kasus. Menurut data dari Kementerian Kesehatan Republik Indonesia per tanggal 31 Maret 2021 jumlah yang terkonfirmasi positif Covid-19 yaitu 1.511.712 dengan jumlah kematian

40,858 dan sembuh 1.348.330. Negara Indonesia merupakan Negara dengan kasus konfirmasi tertinggi di ASEAN (Kemenkes, 2021). Sedangkan, kasus Covid-19 di wilayah Provinsi Lampung pertama dilaporkan masuk pada tanggal 18 Maret 2020 dengan jumlah terkonfirmasi positif sebanyak 1 kasus. Di Provinsi Lampung sendiri jumlah yang terkonfirmasi positif Covid-19 per tanggal 31 Maret 2021 yaitu 14.019 dengan jumlah kematian 759 dan sembuh 12.749 (Lampung, 2021).

Dinas Kesehatan Provinsi Lampung dan Badan Perencanaan Pembangunan Daerah Provinsi Lampung sebelumnya sudah merincikan lima kriteria zonasi berdasarkan warna. Namun belum adanya perhitungan secara umum dan secara detail hanya berdasarkan pengumpulan data dan kajian maupun analisis dari tim pakar Gugus Tugas Nasional penentuan zona menggunakan indikator yang secara total terdapat 15 indikator utama, diantaranya adalah indikator kesehatan masyarakat, yang terbagi menjadi 11 indikator epidemiologi, 2 indikator surveilans kesehatan masyarakat dan 2 indikator pelayanan kesehatan (Tim Pakar Gugus Tugas, 2020).

Pada penelitian ini penulis ingin menerapkan data mining pada data kasus Covid-19 di Provinsi Lampung menggunakan algoritma K-Means Clustering. Data mining adalah proses menemukan pola dan pengetahuan yang menarik dari suatu data dalam jumlah yang besar. Data mining menggambarkan sebuah pengumpulan teknik-teknik dengan tujuan untuk menemukan pola-pola yang tidak diketahui pada data yang telah dikumpulkan (Han, Kamber and Oei, 2012). Algoritma K-Means merupakan salah satu algoritma Clustering yang masuk ke dalam kelompok *Unsupervised learning*

yang digunakan untuk mengelompokkan data ke dalam beberapa kelompok dengan sistem partisi (Dwitri *et al.*, 2020).

Berdasarkan uraian latar belakang tersebut, maka peneliti mencoba menganalisis data kasus Covid-19 di Provinsi Lampung agar dapat mengetahui pengelompokkan pada masalah kasus Covid-19 di Provinsi Lampung, serta dapat memberikan informasi secara detail mengenai pengelompokkan data kasus Covid-19 di Provinsi Lampung. Sehingga peneliti mengangkat penelitian dengan judul **“Analisis Data Mining untuk Clustering Kasus Covid-19 di Provinsi Lampung dengan Algoritma K-Means”**.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, maka rumusan masalah dalam penelitian ini adalah:

1. Bagaimana menganalisis data kasus Covid-19 agar dapat mengetahui pengelompokkan pada masalah kasus Covid-19 di Provinsi Lampung?
2. Bagaimana menerapkan data mining untuk *Clustering* data kasus Covid-19 di Provinsi Lampung dan hasil pengujian validasi menggunakan algoritma K-Means?

1.3 Batasan Masalah

Agar penelitian ini lebih terarah, penulis memberikan batasan masalah agar menghindari pembahasan menjadi terlalu luas. Adapun batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Data yang digunakan adalah data kasus Covid-19 di Provinsi Lampung dari bulan Maret 2020 sampai dengan bulan Maret 2021.

2. Atribut yang akan digunakan adalah Kabupaten/Kota, Suspek, Probable, Konfirmasi Positif, Selesai Isolasi dan Kematian.
3. Dikelompokkan menjadi 4 *cluster* yaitu C0 sebagai Zona Merah, C1 sebagai Zona Orange, C2 sebagai Zona Kuning dan C3 sebagai Zona Hijau.
4. Analisis menggunakan metode Clustering dengan algoritma K-Means.
5. *Tools* yang digunakan dalam penelitian ini adalah *RapidMiner 9.8*.

1.4 Tujuan Penelitian

Dari rumusan masalah yang telah diuraikan, maka tujuan penelitian ini adalah:

1. Untuk menganalisis data kasus Covid-19 agar dapat mengetahui pengelompokkan pada masalah kasus Covid-19 di Provinsi Lampung.
2. Untuk memberikan informasi secara detail dan perhitungan secara umum menggunakan algoritma *K-Means* mengenai pengelompokkan data kasus Covid-19 di Provinsi Lampung.

1.5 Manfaat Penelitian

Adapun manfaat yang diharapkan dalam penelitian ini adalah dapat memberikan informasi perhitungan secara umum dan secara detail mengenai pengelompokkan kasus Covid-19 pada masing-masing Kabupaten/Kota yang ada di Provinsi Lampung.