

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pada puncak tahun 2019, dunia digemparkan dengan adanya penemuan virus jenis baru. Virus jenis baru ini sangat membahayakan, sehingga berakibat pada kematian. Virus ini teridentifikasi pertama kali di daerah Wuhan provinsi Hubei, Tiongkok[1]. Adanya paparan pasar grosir makanan laut di Wuhan yang banyak menjual spesies hewan hidup diduga sebagai munculnya virus jenis baru ini. Penyakit atau virus jenis baru ini dengan cepat menyebar di dalam negeri bagian ke bagian lain China[2].

Virus corona atau yang dikenal dengan COVID-19 merupakan penyakit jenis baru yang ditemukan tahun 2019 yang belum pernah teridentifikasi menyerang manusia. Virus covid-19 dapat menyebabkan infeksi saluran pernapasan atas ringan hingga sedang, seperti halnya penyakit flu.

Dengan adanya *2019-nCoV* cukup menarik perhatian global. Pada 30 Januari 2020, WHO telah menyatakan *covid-19* sebagai darurat kesehatan masyarakat yang menjadi perhatian internasional[2]. WHO (*World Health Organization*) secara resmi mendeklarasikan virus *corona (covid-19)* sebagai pandemic pada tanggal 9 Maret 2020[3]. Pada tanggal 14 Maret 2020, pemerintah telah menetapkan wabah *coronavirus* atau *covid-19* sebagai Bencana Nasional[4].

Penyebaran kasus *covid-19* yang merata di seluruh daerah Indonesia merupakan penyebaran yang cukup cepat dan pesat. Cepatnya penyebaran virus *covid-19* ini sangat berdampak negatif pada semua aspek atau bidang dalam

kehidupan[5]. Jumlah kasus *covid-19* di Indonesia saat ini terus meningkat pesat, hingga Oktober 2021 sebanyak 4.270.794 kasus terkonfirmasi[6].

Bersumber pada bukti ilmiah, *covid-19* dapat menular dari manusia ke manusia melalui percikan batuk / bersin (*droplet*). Virus *covid-19* berisiko tertular, terkhusus orang yang kontak erat atau langsung dengan pasien *covid-19* termasuk yang merawat pasien *covid-19*[4].

Banyaknya warga Negara Indonesia, menyulitkan pemerintah untuk mengedukasi masyarakat berkenaan pencegahan penularan *covid-19* serta pelayanan kesehatan untuk masyarakat. Luasnya wilayah Indonesia, memungkinkan untuk pengelompokkan bagian-bagian berdasarkan wilayah Indonesia. Pengelompokkan ini akan menghasilkan titik-titik pusat penyebaran kasus *covid-19*.

*Data mining* atau *Knowledge Discovery in Database (KDD)* adalah ekstraksi informasi atau pola penting atau menarik dari data yang ada di database yang besar. *Data mining* memiliki beberapa metode, salah satunya adalah *clustering*. *Clustering* adalah salah satu metode dalam *data mining* untuk mengelompokkan data ke dalam beberapa *cluster* sehingga data dalam satu *cluster* memiliki tingkat kemiripan yang maksimum dan data antar *cluster* memiliki kemiripan yang minimum.

*Data mining* dalam fungsi *cluster* analisis bertujuan untuk menghasilkan pengelompokkan objek yang mirip antara satu sama lain dalam kelompok-kelompok. Semakin besar kemiripan objek dalam suatu *cluster* dan semakin besar perbedaan tiap *cluster* maka kualitas analisis *cluster* semakin baik. Algoritma *K-means* merupakan algoritma *clustering* yang termasuk dalam kelompok

*unsupervised learning*, yang digunakan untuk mengelompokkan data ke dalam beberapa kelompok dengan sistem partisi[7].

Berdasarkan uraian di atas dapat disimpulkan bahwa Indonesia masih kekurangan informasi mengenai perkembangan *covid-19*. Dengan cara mengelompokkan penyebaran penyakit *covid-19*, sehingga masyarakat dan pemerintah dapat mengetahui kondisi penyebaran *covid-19* di Indonesia dan pemilihan metode menggunakan metode *k-means*, karena metode *k-means* sudah terbukti dapat mengelompokkan penyebaran *covid-19*. Maka diangkatlah sebuah tema skripsi dengan judul "Implementasi Algoritma *K-means* Dalam Menentukan Tingkat Penyebaran *Covid-19* di Indonesia".

## 1.2 Rumusan Masalah

Dari uraian latar belakang masalah di atas dapat dirumuskan masalah penelitian yaitu Bagaimana implementasi algoritma *k-means* dalam menentukan tingkat penyebaran *covid-19* di Indonesia?

## 1.3 Batasan Masalah

Karena luasnya permasalahan yang ada, maka perlu adanya pembatasan terhadap pembahasan beberapa masalah diantaranya:

1. Metode yang digunakan adalah *clustering* yaitu *k-means*.
2. Penelitian berfokus terhadap *clustering* tingkat penyebaran kasus *covid-19* di Indonesia.
3. Sumber dataset yang digunakan berdasarkan [kawalcovid19.id](https://kawalcovid19.id)

4. Data yang digunakan merupakan data *covid-19* dari 34 provinsi di Indonesia.
5. Data yang digunakan dalam periode 1 Juni 2021 hingga 30 September 2021.
6. Metode evaluasi menggunakan *Elbow Method*, *Davies Bouldin Index*, dan *Silhouette Coefficient*.

#### 1.4 Maksud dan Tujuan Penelitian

Setiap penelitian yang akan dilakukan pastilah memiliki tujuan. Penelitian ini dilakukan dengan beberapa maksud dan tujuan. Adapun maksud dan tujuan dari penelitian ini adalah:

1. Mengetahui tingkat penyebaran kasus *covid-19* di Indonesia.
2. Menerapkan metode *k-means* sebagai konsep pengetahuan di dalam pengelompokan tingkat penyebaran *covid-19*.
3. Pengelompokan provinsi berdasarkan tingkat penyebaran *covid-19* di Indonesia menggunakan algoritma *k-means*.
4. Hasil *clustering* metode *k-means* diharapkan dapat mengelompokkan penyebaran kasus *covid-19*.

### 1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberi manfaat bagi pihak-pihak yang terkait. Manfaat tersebut diantaranya adalah:

1. Bagi Peneliti

Diharapkan melalui penelitian ini, peneliti dapat memperkaya pengetahuan dan wawasan dalam disiplin ilmu yang ditekuni.

2. Bagi Universitas Amikom Yogyakarta

Diharapkan hasil penelitian ini dapat digunakan untuk menambah pengetahuan, wawasan, referensi, acuan, pembandingan, dan bahan masukan dalam penelitian selanjutnya, terutama penelitian yang berkaitan dengan tingkat penyebaran *covid-19*, dan dapat menyempurnakan kelemahan dalam penelitian ini.

3. Bagi Peneliti Selanjutnya

Diharapkan penelitian ini dapat digunakan sebagai acuan dalam pengembangan penelitian selanjutnya. Penggunaan variabel sampel yang digunakan sebisa mungkin di daerah lainnya, sehingga dapat memperkaya informasi mengenai tingkat penyebaran *covid-19* pada daerah lainnya.

## 1.6 Metode Penelitian

Penelitian menggunakan metode penelitian sebagai acuan dasar dalam penyusunan laporan dan kerangka kerja. Adapun metode-metode yang digunakan antara lain seperti yang akan dibahas dibawah ini:

### 1.6.1 Identifikasi Masalah

Pada tahap ini akan mempelajari masalah yang terjadi dengan cara melakukan observasi melalui berita ataupun data-data yang mendukung. Dengan begitu mendapatkan permasalahan berdasarkan penelitian terbaru maupun berdasarkan pengamatan keadaan sekarang.

### 1.6.2 Metode Pengumpulan Data

Data yang digunakan merupakan *dataset* publik yang berisi kumpulan data-data *covid-19* yang dikelompokkan ke dalam bentuk format file *Microsoft Excel* (*xlsx*). *Dataset* yang digunakan merupakan rujukan dari penelitian yang sebelumnya telah dilakukan. Data yang digunakan sebagai pengujian terkait *clusteriasi data covid-19 di Indonesia*

### 1.6.3 Metode Analisis

Metode ini dilakukan untuk mengidentifikasi masalah menggunakan metode *K-Means*. Analisis dilaksanakan dengan pengelompokkan data sehingga dapat mengetahui tingkat persebaran *covid-19* di Indonesia. Analisis dengan menggunakan studi literatur dilakukan dengan mencari sumber pustaka yang relevan dengan masalah yang dapat dijadikan topik sehingga dapat dianalisis untuk

mendapatkan pemahaman mengenai penerapan metode *k-means* dalam mengelompokkan data dalam penyebaran virus *covid-19*.

### **1.7 Sistematika Penulisan**

Sistematika merupakan kerangka dan pedoman penulisan skripsi ditujukan agar mempermudah baca dan mengetahui pembahasan yang ada pada skripsi ini secara menyeluruh. Adapun sistematika penulisannya adalah sebagai berikut:

#### **BAB I PENDAHULUAN**

Bab ini menghasilkan latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, metode penelitian, dan sistematika penulisan skripsi.

#### **BAB II LANDASAN TEORI**

Bab ini menjelaskan tentang tinjauan pustaka yang digunakan sebagai referensi dan rujukan dari penelitian sebelumnya. Bab ini juga memuat tentang dasar-dasar teori dari objek penelitian dan metode pendekatan yang digunakan pada penelitian kali ini.

#### **BAB III METODE PENELITIAN**

Bab ini memuat uraian tentang metode yang digunakan pada penelitian ini, uraian tentang bagaimana penelitian dilakukan, serta alur bagaimana pendekatan yang diusulkan akan mengatasi masalah yang diangkat pada penelitian ini.

#### **BAB IV HASIL DAN PEMBAHASAN**

Bab ini merupakan bagian utama pada penelitian ini, berisi implementasi pendekatan yang diusulkan. Analisis dari hasil yang didapatkan pada pengujian akan diulas tuntas pada bab ini. Bab ini memuat hasil dari pengujian yang telah diukur dengan berbagai skenario pengujian yang telah dilakukan.

## BAB V PENUTUP

Bab ini memuat jawaban dari tujuan penelitian yang dilakukan. Apakah hasil akhir dari rangkaian proses penelitian yang telah dilakukan mampu menjawab tujuan penelitian untuk mengatasi masalah yang diangkat. Bab ini juga memuat segala kekurangan dari penelitian sebagai saran yang dapat dilakukan kajian dan pengembangan oleh peneliti berikutnya.

