

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

COVID-19 (*Coronavirus Disease 2019*) adalah penyakit menular yang disebabkan oleh virus SARS-CoV2 (*Severe Acute Respiratory Syndrome Coronavirus 2*) yang dapat menyerang sistem pernapasan manusia. Virus SARS - CoV-2 disebarkan melalui droplet respiratorik yang terjadi saat bersin dan batuk serta secara tidak langsung ditemukan pada benda dan permukaan yang terkontaminasi. Orang yang terinfeksi COVID-19 memiliki gejala umum sesak napas akut, seperti demam, batuk, dan sesak napas.

Menurut Tiodora (2020), setiap orang berisiko terinfeksi COVID-19, namun ada beberapa kelompok orang yang berisiko tinggi terinfeksi COVID-19 dan bisa berakibat fatal. Untuk meningkatkan kesadaran dan kewaspadaan masyarakat terhadap pencegahan COVID-19, terdapat banyak penelitian yang telah dilakukan untuk mengetahui kelompok orang yang berisiko COVID-19 lebih rentan. Dalam penelitian yang dilakukan Tiodora (2020) dengan menggunakan metode *Discourse Network Analysis*, ia sampai pada kesimpulan bahwa kelompok yang rentan terpapar COVID-19 yaitu kelompok lansia, penderita penyakit kronis, perokok, pria, dan orang dengan golongan darah A.

Studi yang dilakukan ini akan berfokus pada salah satu kelompok berisiko yaitu orang dengan penyakit kronis atau penyakit penyerta yang ketika terinfeksi COVID-19, semakin merusak sistem kardiovaskular dalam tubuh. Diketahui bahwa penyakit kardiovaskular dikenal sebagai penyebab utama kematian di dunia. Menurut dr Hari Yusti Laksono, Sp.JP dalam artikel kesehatan RSST (2020), penyakit kardiovaskular dapat dilakukan pencegahan di luar beberapa faktor risiko seperti usia, genetik, kecenderungan muncul pada pria, dan beberapa faktor risiko lainnya. Faktor risiko tekanan darah tinggi, dislipidemia, diabetes, obesitas, dan merokok dapat dicegah dengan mengubah pola hidup sehari-hari. Menurut Herick, dkk. (2019), pada pasien COVID-19, komplikasi kardiovaskular memiliki risiko

morbiditas (derajat penyakit) dan mortalitas (pentingnya mortalitas pada populasi wilayah tertentu) yang lebih tinggi terjadi pada pasien penderita komorbid dibandingkan pasien tanpa penyakit penyerta. Infeksi COVID-19 yang berat pada pasien dengan komorbid dapat menyebabkan komplikasi akut dengan peningkatan risiko kardiovaskular jangka panjang.

Jurnal penelitian sebelumnya yang dilakukan oleh Rizkiana *et al.* (2020), melakukan klusterisasi pasien COVID-19 dengan variabel jumlah kasus COVID-19, kasus sembuh, kasus meninggal dunia, dan CFR berdasarkan tiap provinsi di Indonesia. Hasil penelitian tersebut menghasilkan tingkat risiko kematian pasien COVID-19 dengan kategori rendah, sedang, dan tinggi untuk tiap provinsi.

Sebuah lembaga penelitian Covid Analytics melakukan analisa terhadap 160 lebih data riset medis dari jurnal kesehatan terbitan beberapa Negara untuk melakukan analisa berbagai pengetahuan baru mengenai COVID-19. Peneliti Covid Analytics membagikan set data yang berisikan kasus medis pasien terinfeksi COVID-19 dari berbagai rumah sakit untuk bisa dilakukan penelitian lebih jauh dengan metode penelitian yang lebih luas lagi.

Pada kebijakan penggunaannya, Covid Analytics menyatakan bahwa dataset klinis yang disediakan tidak dapat digunakan untuk memprediksi tingkat kematian pada populasi umum melihat dari kohort hasil laboratoriumnya. Selain itu, dijelaskan pula bahwa dataset klinis tersebut memungkinkan untuk memperoleh perkiraan yang cukup akurat dari jumlahnya jika dengan (a) memperhitungkan prevalensi pasien tanpa gejala, dan (b) hanya memasukkan studi yang representatif.

Maka dari itu penelitian ini hanya akan berfokus pada informasi prevalensi komorbiditasnya serta nilai tengah usia pada sebuah populasi dikarenakan cukup representatif untuk menggolongkan kelompok pasien berkomorbid dengan tiga kategori, yaitu berderajat ringan, berderajat sedang, dan berderajat berat.

Berdasarkan latar belakang tersebut, dapat diimplementasikan metode K-Means *Clustering* dengan memanfaatkan pengetahuan *Machine learning* melalui tahapan analisis data, penentuan pola dan pembuatan model pengujiannya. Algoritma K-Means *Clustering* cukup efisien dan mudah diterapkan untuk pengklasteran sebuah dataset yang besar karena kompleksitas perhitungannya yang bersifat linier.

Dari penguraian di atas, penulis tertarik untuk melakukan penelitian mengenai pengelompokan pasien COVID-19 berdasarkan komorbiditas dengan judul **“ANALISIS DAN PERBANDINGAN KELOMPOK PASIEN COVID-19 BERDASARKAN KOMORBIDITAS MENGGUNAKAN *K-MEANS CLUSTERING*”**

## 1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka dapat disimpulkan rumusan masalah sebagai berikut:

- a. Bagaimana mengelompokkan pasien COVID-19 dan prevalensi penyakit penyerta-nya dengan metode *K-Means Clustering*?
- b. Bagaimana hasil dari proses *clustering* yang akan didapatkan?

## 1.3 Batasan Masalah

Batasan masalah dari proposal ini di antaranya:

- a. Model untuk *clustering* pasien COVID-19 dengan komorbiditas berdasarkan nilai tengah usia dan persentase komorbid pada ukuran populasinya.
- b. Data yang di olah di dapatkan dari website Covid Analytics ([www.covidanalytics.io/dataset](http://www.covidanalytics.io/dataset)) yang merupakan rangkuman data riset kesehatan pada pasien COVID-19 dari beberapa populasi yang terintegrasi sejak bulan Desember 2019 sampai April 2020.
- c. Jumlah cluster yang akan dihasilkan pada proses *clustering* sebanyak 3 cluster yang didapatkan dari perhitungan metode Elbow terhadap data untuk menemukan nilai k yang optimal.
- d. Metode *K-Means Clustering* untuk mengelompokkan pasien COVID-19 dengan komorbiditas yang meliputi Hipertensi, Diabetes, Penyakit Jantung Koroner dan Kardiovaskular, Penyakit Paru Obstruktif Kronik, dan Gagal Ginjal Akut. Atribut data yang digunakan berupa: *Median* (nilai tengah) Usia dan Persentase komorbid yang disesuaikan dengan ukuran Populasi-nya.

## 1.4 Tujuan Penelitian

Tujuan penelitian yang ingin dicapai yaitu melakukan analisis data dengan implementasi metode *K-Means Clustering* untuk menghasilkan kelompok pasien COVID-19 berkomorbid dengan prevalensi atau derajat ringan-berat berdasarkan rentang usia dan populasinya.

## 1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini yaitu:

- a. Hasil *clustering* dapat menghasilkan kelompok risiko rentan sesuai dengan data pasien berkomorbid dan populasinya.
- b. Menambah pengetahuan dan keahlian penulis dalam melakukan analisa dan pemodelan dari atribut data yang diperoleh
- c. Dapat membantu proses pengambilan keputusan untuk melihat tingkat prevalensi yang akan di alami pada kelompok orang berkomorbid tertentu.

## 1.6 Luaran yang diharapkan

Luaran yang diharapkan dari penelitian ini adalah dapat menghasilkan sebuah sistem untuk menampilkan hasil dari pemodelan *K-means Clustering* berupa ukuran tingkat prevalensi komorbiditas dari pasien COVID-19 dalam visualisasi yang mudah untuk dipahami beserta fitur untuk memprediksi cluster berdasarkan data yang dimasukkan.

## 1.7 Sistematika Penulisan

Dalam penulisan, penelitian ini menggunakan sistematika penulisan sebagai berikut.

### **BAB I PENDAHULUAN**

Bab ini berisikan Latar Belakang, Rumusan Masalah, Batasan Masalah, Tujuan Penelitian, Manfaat Penelitian, Luaran yang Diharapkan, dan Sistematika Penulisan.

### **BAB II LANDASAN TEORI**

Bab ini berisikan landasan teori yang menjadi dasar dari penyusunan proposal, serta uraian penelitian terdahulu yang relevan dengan judul penelitian ini.

### **BAB III METODOLOGI PENELITIAN**

Bab ini berisikan beberapa metode penelitian yang akan digunakan oleh penulis seperti Kerangka Berpikir, Alat Bantu Penelitian, dan Jadwal Penelitian.

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

Bab ini berisikan proses pemodelan algoritma *K-Means Clustering* untuk mengelompokkan kelompok pasien COVID-19 berkomorbiditas serat dilakukan pembahasan hasil yang didapatkan serta gambaran sistem yang dihasilkan.

### **BAB V PENUTUP**

Bab ini berisi kesimpulan dari keseluruhan pembahasan bab sebelumnya dan saran yang diberikan untuk penelitian lanjutan yang berkaitan.

### **DAFTAR PUSTAKA**

### **RIWAYAT HIDUP**

### **LAMPIRAN**