

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Virus Covid-19 adalah penyakit jenis baru dan belum pernah ditemukan menyerang manusia. Kasus pertama covid-19 ditemukan di Provinsi Wuhan, China pada tahun 2019 [1]. Pneumonia dengan gejala umum mirip flu diduga sebagai wabah pertama virus covid-19. Namun, berbeda dengan flu, virus Covid-19 dapat berkembang dengan cepat dan menyebabkan infeksi serius bahkan kematian. Karena penularan virus yang sangat cepat, WHO menyatakan virus Covid-19 sebagai pandemi pada 11 Maret 2020. Status pandemi menunjukkan bahwa penyebaran Covid-19 sedang berlangsung [2]. Siapapun bisa terkena virus ini, dari anak kecil hingga orang tua. Pemerintah sudah melakukan cara untuk mencegah agar virus covid-19 tidak menyebar dengan cepat di Indonesia. Bukan hanya pemerintah namun seluruh pihak melakukan berbagai macam cara agar dapat memberhentikan virus covid-19, karena virus ini banyak masyarakat yang terpapar dikarenakan kasus semakin tinggi setiap harinya [3].

Namun berbagai cara telah dilakukan, namun kasus covid-19 terus meningkat. Oleh karena itu, untuk membatasi penyebaran covid-19, perlu dilakukan klasifikasi daerah-daerah yang memiliki tingkat kerawanan tinggi terhadap penyebaran covid-19[4]. Setiap wilayah yang terkena virus covid-19 memiliki penyebaran yang berbeda-beda, mulai dari jumlah kasus positif, jumlah kasus sembuh, dan jumlah kasus yang dirawat hingga meninggal. Kondisi yang berbeda memerlukan penanganan yang berbeda dari setiap daerah [5]. Oleh karena itu, diperlukan klasifikasi untuk mengelompokkan jumlah pasien yang terdampak Covid-19 berdasarkan usia dan wilayah dengan menggunakan teknik pengolahan data mining. Data mining merupakan suatu metode untuk pengolahan data dalam skala besar, data yang diolah dengan teknik ini akan menghasilkan sebuah pengetahuan yang baru yang digunakan untuk pengambilan keputusan [6]. Salah satu algoritma data mining yang digunakan adalah algoritma K-Means. K-Means adalah algoritma unsupervised learning yang digunakan untuk mengelompokkan data menjadi beberapa partisi. Data yang dikelompokkan menggunakan algoritma k-means memiliki sifat yang sama, tetapi kelompok yang lain memiliki sifat yang berbeda [7]. K-means memiliki kelebihan dibandingkan dengan algoritma yang lainnya yaitu sederhana, mudah diimplementasikan, tidak lambat dalam pengujian dan mudah disesuaikan. K-means juga biasa digunakan dalam pengolahan data mining [8]. Algoritma k-means digunakan dalam penelitian yang dilakukan oleh (Lilis Suriani, 2020) untuk mengelompokkan data kriminal di Poldasu dalam mengetahui pola rawan dalam aktivitas kriminal. Karena algoritma K-Means merupakan algoritma yang sederhana dan efektif untuk menemukan kelompok dalam data [9].

Oleh karena itu, penelitian ini menggunakan algoritma k-means untuk mengetahui tingkat penyebaran covid-19 berdasarkan usia dan wilayah. Diharapkan dengan adanya penelitian ini dapat membantu pemerintah dalam mengambil keputusan untuk menanggulangi penyebaran covid-19. Sehingga keputusan yang diambil dapat memutuskan rantai penyebaran covid-19.

1.2. Rumusan Masalah

Dengan latar belakang diatas, maka rumusan masalah yang akan diselesaikan yaitu untuk mengetahui penyebaran pasien covid-19 berdasarkan usia dan wilayah.

1.3. Tujuan dan Manfaat

Tujuan dari penelitian ini adalah untuk mengetahui penyebaran Covid-19 berdasarkan usia dan wilayah. Sehingga nanti akan dapat menjadi solusi bagi pemerintah untuk cepat menanggulangi daerah yang memiliki kasus Covid-19 terbanyak.

Manfaat penelitian ini, yaitu:

1. Untuk membantu pemerintah dalam memberikan penanggulangan yang baik untuk setiap daerah.
2. Untuk mengetahui tingkat penyebaran Covid-19 berdasarkan usia dan wilayah.

1.4. Keterbaruan

Penelitian yang dilakukan oleh Merinda Lestandy dan Lailis Syafa'ah (2020) dalam penelitiannya menggunakan algoritma KNN untuk memprediksi kasus aktif Covid-19. Penelitian ini menggunakan 260 data dengan 13 parameter dan data training sebesar 80% dan data testing 20%. Hasil dari pengujian ini menghasilkan nilai akurasi 72,3337% dan MSE sebesar 0,007 [10].

Penelitian yang dilakukan Nayuni Dwitri dkk (2020), pada penelitiannya menggunakan algoritma K-Means untuk mengetahui tingkat penyebaran Covid-19 pada Indonesia. Penelitian ini menggunakan data jumlah penyebaran virus covid-19 pada tanggal 9 Mei 2020 yang diperoleh dari Kementerian Kesehatan Republik Indonesia. Pengujian dilakukan dengan menggunakan rapidminer untuk menghasilkan 3 cluster. Cluster 1 mempunyai 5056 kasus pasien positif dan 427 kasus meninggal, untuk cluster 2 mempunyai 4525 kasus positif dan 348 kasus meninggal, sedangkan cluster 3 mempunyai 4043 kasus positif dan 184 kasus meninggal [11].

Penelitian yang dilakukan Gunaawan dkk (2020) dalam penelitiannya menggunakan algoritma K-Medoids untuk mengklasterisasi provinsi di Indonesia yang terkena virus Covid-19. Pada penelitiannya menggunakan data pasien dari tanggal 02 Maret 2020 sampai 30 Juni 2020 dengan menggunakan 3 variabel yaitu terkonfirmasi, meninggal dan sembuh. Hasil pengujian didapatkan 3 cluster, cluster 1 memiliki jumlah kasus terkonfirmasi 12259, kasus meninggal 793 dan kasus sembuh 5631. Untuk cluste 2 memiliki jumlah kasus terkonfirmasi 2632, kasus meninggal 108 dan kasus sembuh 1077. Sedangkan cluster 3 memiliki jumlah kasus terkonfirmasi 388, kasus meninggal 1 dan kasus sembuh 210 [12].

Penelitian yang dilakukan oleh Kristin dkk (2020) tentang penerapan algoritma k-means untuk mengetahui tingkat kepuasan pembelajaran online selama Covid-19. Dalam penelitiannya menggunakan data yang diisi melalui link kuesioner, hasil pengujian yang dilakukan menghasilkan 3 cluster yaitu cluster pertama yang menyatakan setuju untuk pembelajaran online termasuk rendah dan yang menyatakan tidak setuju termasuk tinggi, cluster kedua yang menyatakan setuju untuk pembelajaran online termasuk sedang dan yang menyatakan tidak setuju termasuk sedang, sedangkan cluster 3 yang menyatakan setuju pada pembelajaran online termasuk tinggi dan yang menyatakan tidak setuju termasuk rendah [13].

Penelitian yang dilakukan oleh Dina Sunia dkk (2017) menggunakan algoritma k-means dalam penelitiannya untuk mencari data tentang masyarakat miskin. Penelitian menggunakan data yang diperoleh dari BPS Kota Jambi pada bulan Maret 2017 dengan jumlah masyarakat miskin sebanyak 286,55 ribu jiwa. Hasil perhitungan menghasilkan 512 sampel data dengan jumlah cluster sebanyak 5. Cluster pertama terdapat 13 penduduk, Cluster kedua terdapat 153 penduduk, cluster ketiga terdapat 129 penduduk, Cluster keempat terdapat 138 penduduk, dan cluster kelima sebanyak 79 penduduk [14].

Penelitian yang dilakukan oleh Hasyrif SY dkk (2019) dalam penelitiannya menggunakan algoritma k-means untuk pengelompokan penyebaran diare di Kota Makassar. Penelitian ini menggunakan data kesehatan pada tahun 2016. Hasil pengujian mendapatkan 2 cluster. Pada cluster kedua merupakan cluster dengan daerah tertinggi [15].

Penelitian yang dilakukan oleh Abdullah D dkk (2021) dalam penelitiannya menggunakan algoritma k-means untuk mengelompokkan provinsi yang terdampak pandemi covid-19. Data yang digunakan dalam penelitian ini menggunakan data yang diperoleh pada tanggal 19 April 2020 dari gugus tugas covid-19. Pengujian dalam penelitian ini menghasilkan 3 kelompok yaitu kelompok 1 ada 4 Provinsi, kelompok 2 ada 28 Provinsi sedangkan pada kelompok 3 terdapat 1 Provinsi [16].