

BAB I PENDAHULUAN

I.1 Latar Belakang

Maraknya penyebaran penyakit yang diakibatkan oleh virus COVID-19 yang telah ditetapkan sebagai pandemi oleh WHO pada tanggal 12 Maret 2020. Akibat dari virus COVID-19 ini banyak pasien yang mengalami kematian (Ciotti et al., 2020). COVID-19 pertama kali terdeteksi pada manusia dengan gejala peradangan paru-paru (pneumonia) di Wuhan, Cina tanggal 30 Desember 2019. Pada awal penyebab munculnya virus ini dapat dikaitkan dengan kelelawar. Namun prediksi tersebut masih belum dapat dipastikan kebenarannya karena COVID-19 telah menjadi masalah di seluruh dunia. Meskipun tingkat kematian akibat virus ini dianggap relatif rendah namun penularan virus ini dianggap begitu cepat melalui hidung, mulut, dan tetesan air liur melalui bersin atau batuk (Triana et al., 2021). Pemerintah Indonesia dan juga negara-negara di dunia tengah berupaya mengembangkan dan menghadirkan Vaksin COVID-19. Namun sebagian besar mengutarakan kemungkinan efek samping dari vaksin ini menjadi faktor kekhawatiran utama, sehingga meningkatkan keraguan dan kurangnya kesadaran pada masyarakat (Iskak et al., 2021). Faktor lainnya terdapat masyarakat pada kelurahan tertentu yang belum terjangkau terhadap program vaksinasi di daerahnya akibat distribusi vaksin yang belum merata. Namun demikian, pemerintah juga harus memastikan bahwa program vaksinasi ini sudah berjalan dengan baik dan melaksanakan program vaksinasi lebih merata (Sejati & Dharmaningtias, 2021).

Program vaksinasi COVID-19 dimulai pada tanggal 13 Januari 2021. Untuk gelombang pertama, vaksin tersebut diberikan ke tenaga kesehatan, petugas publik, dan lansia. Pada gelombang kedua, sasaran vaksinasi adalah masyarakat kelompok rentan dan masyarakat umum lainnya. Jenis vaksin yang telah dan akan digunakan di Indonesia adalah AstraZeneca, Moderna, Pfizer, Sinopharm dan Sinovac. Kelima jenis vaksin tersebut memiliki fungsi yang berbeda-beda berdasarkan uji klinis yang telah dilakukan (Nasir et al., 2021). Masyarakat dapat melihat siapa saja yang sudah menerima vaksin berdasarkan golongan usia dan

wilayah melalui *Open government Data* (OGD) yang telah disediakan oleh pemerintah.

Open government Data yaitu sebuah doktrin pemerintahan yang terbuka mengenai data-data yang dimilikinya untuk hal-hal tertentu sehingga publik dapat dengan bebas menggunakan data tersebut. Keterbukaan mengenai data ini selain sebagai bentuk transparansi dan akuntabilitas tetapi juga harapannya dapat digunakan oleh akademisi, pebisnis, birokrat, profesional dan kalangan lainnya untuk mengembangkan keahlian dibidang masing-masing sehingga dapat berdiskusi tentang kebijakan-kebijakan pemerintah (Yudan & Arief Virgy, 2021). Secara umum *Open government Data* adalah data yang dipublikasikan di Internet dengan cara yang sedemikian rupa sehingga siapapun dapat dengan bebas menggunakannya kembali, dan mendistribusikan ulang data ini. Riwayat file vaksinasi dapat di buat berbentuk *Open Governement Data* yang kini sudah diterapkan oleh kota kota besar yang berada di Indonesia salah satunya yaitu DKI Jakarta.

DKI Jakarta telah meluncurkan portal *Open data* Riwayat file vaksinasi yang di bagi menjadi dua basis yang pertama berbasis kelurahan dan kecamatan dan yang kedua berbasis RT yang dapat diakses melalui laman website vaksinasi DKI Jakarta¹. Website ini berupa portal *Open data* yang bisa diakses oleh seluruh masyarakat untuk mendapatkan berbagai macam informasi mengenai penerima vaksinasi berdasarkan usia dan wilayahnya. Data yang diberikan memiliki manfaat yang sangat penting guna menentukan tindakan yang tepat sesuai dengan kondisi wilayah. Berdasarkan data Badan Pusat Statistik (BPS) permasalahan yang terjadi adalah terjadi kesulitan dalam membagi wilayah penerima vaksinasi. Hal ini dikarenakan jumlah wilayah yang banyak dengan kondisi yang berbeda-beda setiap wilayahnya. Salah satu upaya untuk membantu proses mengolah data vaksinasi yang besar adalah teknik data *mining*.

Secara teknis, data *mining* dapat disebut sebagai proses untuk menemukan korelasi atau pola dari ratusan atau ribuan field dari sebuah relasional database

¹ Laman website vaksinasi DKI Jakarta: <https://riwayat-file-vaksinasi-dki-jakarta-jakartagis.hub.arcgis.com/>

yang besar (Noviyanto, 2020). Penggalan data atau dalam bahasa Inggrisnya yaitu *Data Mining* adalah ekstraksi pola yang menarik dari data dalam jumlah besar. Penggalan data memiliki beberapa nama alternatif, meskipun definisi eksaknya berbeda, seperti KDD (*Knowledge Discovery in Database*). *Data mining* sangat perlu dilakukan terutama dalam mengelola data yang sangat besar untuk memudahkan aktivitas recording suatu transaksi dan untuk proses data *warehousing* agar dapat memberikan informasi yang akurat bagi penggunanya (Arianto, 2019). Dari beberapa pemaparan di atas dapat disimpulkan bahwa *data mining* merupakan serangkaian proses untuk menggali nilai tambah berupa informasi yang diperoleh dengan cara mengenali pola yang penting dari data yang terdapat pada basis data (Sundari et al., 2019). Berbagai penelitian telah melakukan pemanfaatan metode *Data Mining* khususnya *Clustering* dalam melakukan pengelompokan data. Pada penelitian ini, peneliti menggunakan teknik *Clustering* untuk pengelompokan data wilayah (kelurahan) penerima vaksinasi di DKI Jakarta.

Clustering untuk vaksinasi di DKI Jakarta dilakukan dengan pengelompokan wilayah dengan kondisi yang sama. Berdasarkan hasil perekapan data vaksinasi DKI Jakarta² sejak awal vaksinasi pada tanggal 13 Januari 2021 hingga tanggal pengambilan data yaitu 19 April 2022, terdapat 1.471.464 dari 8.941.211 warga belum menerima vaksinasi sama sekali atau sekitar 16,38%. Dengan persentase yang relatif tinggi pada warga yang belum menerima vaksinasi selama kurun waktu 15 bulan penyebaran vaksin di Indonesia, maka dapat dikatakan bahwa pemberian vaksinasi masih sangat perlu diperhatikan dengan pemerataan sasaran di setiap wilayah DKI Jakarta yang dapat dilakukan dengan proses *clustering*. *Clustering* sendiri baru diterapkan pada pemetaan wilayah penyebaran COVID-19 dan belum dilakukan untuk pemetaan program vaksinasi di DKI Jakarta. Hal ini penting untuk dilakukan mengingat DKI Jakarta merupakan salah satu provinsi dengan tingkat penyebaran COVID-19 tertinggi di Indonesia³. Terlebih lagi DKI Jakarta saat ini mulai melakukan aktivitas-aktivitas WFO dan pembelajaran tatap

² Laman website vaksinasi DKI Jakarta: <https://riwayat-file-vaksinasi-dki-jakarta-jakartagis.hub.arcgis.com/>

³ Peta persebaran COVID-19 di Indonesia | <https://covid19.go.id/>

muka yang dikhawatirkan dapat menimbulkan COVID-19 gelombang baru bila penyebaran vaksinasi belum dilakukan secara menyeluruh.

Clustering merupakan suatu proses pengelompokan record, observasi, atau mengelompokkan kelas yang memiliki kesamaan objek. Perbedaan *clustering* dengan klasifikasi yaitu tidak adanya variabel target dalam melakukan pengelompokan pada proses sering dilakukan sebagai langkah awal dalam proses data mining. Terdapat banyak algoritma *clustering* yang telah digunakan oleh peneliti sebelumnya seperti *K-means*, *Improved K-means*, *Fuzzy C-Means*, *DBSCAN*, *K-medoids* (PAM), *CLARANS* dan *Fuzzy Subtractive*. Setiap algoritma memiliki kelebihan dan kekurangan masing-masing, tetapi memiliki prinsip algoritma yang sama, yaitu mengelompokkan data sesuai dengan karakteristik dan mengukur jarak kemiripan antar data dalam satu kelompok (Arianto, 2019). Seiring perkembangan zaman banyak teknologi yang telah banyak membantu manusia untuk menyelesaikan suatu masalah (Ningsih et al., 2019). Hal inilah yang mendorong adanya algoritma-algoritma dari Data Mining. Salah satunya adalah Algoritma *K-medoids* yang merupakan metode dalam teknik *clustering* (Kumar & Rathee, 2011). Dalam hal ini peneliti memilih untuk menggunakan algoritma *K-medoids* karena metode *K-medoids* dapat representatif dalam meminimalkan jumlah ketidaksamaan data. Metode Algoritma *K-medoids* digunakan dalam pembuatan suatu sistem *cluster* untuk mengelompokkan wilayah yang terdampak COVID-19 sehingga memberikan hasil pemetaan berupa kluster wilayah atau kelurahan yang terdampak dan dapat diperoleh pengetahuan berupa informasi prioritas wilayah yang perlu diberikan vaksinasi (Park & Jun, 2009).

K-medoids adalah salah satu metode partisi, karena menggunakan objek yang paling terpusat (*medoids*) di *cluster* menjadi pusat *cluster* dari nilai rata-rata objek dalam sebuah *cluster*. Metode *K-medoids* lebih cocok untuk mengelompokkan data dibandingkan metode *K-means* (Marlina et al., 2018). Dibandingkan dengan *K-Means*, *K-medoids* lebih kuat untuk menangani kebisingan (*noise*) dan pencilan (*outlier*) karena meminimalkan beberapa *dissimilarities* berpasangan, bukan jumlah kuadrat jarak Euclidean. Sebuah *medoid* dapat diartikan sebagai objek *cluster* yang rata-rata perbedaannya untuk semua objek dalam *cluster* minimal

yaitu titik paling berlokasi di *cluster* (Syahrul dkk, 2004). Alasan peneliti menggunakan metode ini karena metode *K-medoids* representative dalam meminimalkan jumlah ketidaksamaan data, alasan lainnya yaitu data yang digunakan dalam penelitian ini dalam kriteria data kecil atau sedikit.

Penelitian sebelumnya yang berjudul “Analisis *K-medoids Clustering* Dalam Pengelompokan Data Imunisasi Campak Balita di Indonesia” mendapatkan kesimpulan bahwa telah digunakan jumlah record sebanyak 34 provinsi yang menghasilkan 3 *cluster* yakni *cluster* rendah sebanyak 2 provinsi, *cluster* sedang sebanyak 30 provinsi, dan *cluster* tinggi sebanyak 2 provinsi (Sundari et al., 2019). Penelitian tersebut menunjukkan hasil yang sama dengan analisis perhitungan metode dimana diperoleh 2 provinsi *cluster* rendah yang menjadi pusat perhatian bagi pemerintah dalam melakukan sosialisasi dan pemerataan dalam pemberian imunisasi campak balita pada provinsi tersebut.

Penelitian lainnya yang berjudul “Algoritma *K-medoids* Untuk Penentuan Strategi Pemasaran Produk” menyimpulkan bahwa berdasarkan dataset transaksi penjualan pada sebuah supermarket menghasilkan 5 *cluster* dengan *cluster* pertama terdiri dari 909 records, *cluster* kedua terdiri dari 166 records, *cluster* ketiga terdiri dari 66 records, *cluster* keempat terdiri dari 132 records, dan *cluster* kelima terdiri dari 87 records. Sehingga dalam hal strategi pemasaran produk dapat difokuskan untuk melakukan promosi pada *cluster* kelima yang memiliki kombinasi jumlah barang dibeli yang paling tinggi, yakni sekitar 49-303 barang (Triyanto, 2015) .

Berdasarkan permasalahan dan paparan penelitian sebelumnya, peneliti membuat sebuah penelitian yang berjudul “Pengelompokan Wilayah Penerima Vaksinasi COVID-19 di DKI Jakarta Dengan Metode *Clustering* Menggunakan Algoritma *K-medoids*”. Penelitian ini dapat menjadi sebuah solusi agar memudahkan dalam mengetahui wilayah mana yang paling banyak terjadi pada masyarakat yang belum melakukan vaksinasi COVID-19, sehingga pemerintah dapat melakukan penambahan dosis vaksin di wilayah yang masih sedikit menerima vaksinasi COVID-19 tersebut.

I.2 Perumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan di atas, maka perumusan masalah dalam penelitian ini adalah sebagai berikut :

1. Bagaimana menentukan jumlah *cluster* optimal dalam proses algoritma *K-medoids* untuk mengelompokan wilayah penerima vaksinasi COVID-19 di DKI Jakarta ?
2. Bagaimana interpretasi *cluster* optimal atau *cluster* yang dipilih sebagai kelompok wilayah dengan penerima vaksinasi COVID-19 yang paling banyak ?

I.3 Tujuan Penelitian

Berdasarkan rumusan masalah sebelumnya, Adapun tujuan dari penelitian ini adalah sebagai berikut :

1. Untuk mengetahui cara mengimplementasikan Algoritma *K-medoids* pada pengelompokan (*cluster*) wilayah DKI Jakarta yang telah menerima vaksin COVID-19.
2. Untuk mengetahui proses Algoritma *K-medoids* untuk mengklaster data pada wilayah yang paling banyak menerima vaksin COVID-19 yang di ambil melalui portal web yang di telah di sediakan oleh pemerintah DKI Jakarta.

I.4 Batasan Penelitian

Beberapa batasan masalah dalam penelitian ini adalah sebagai berikut :

1. Penelitian ini hanya berfokus pada implementasi algoritma *K-medoids* terhadap program vaksinasi COVID-19 di DKI Jakarta.
2. Berdasarkan Algoritma *K-medoids* yang berfokuskan melalui *Open government Data* DKI Jakarta.
3. Parameter atribut yang di gunakan pada penelitian ini terdiri dari : kelurahan, lansia total vaksin di berikan, pelayanan public total vaksin di berikan, gotong royong total vaksin di berikan, tenaga kesehatan total vaksin di berikan dan remaja total vaksin di berikan. `

I.5 Manfaat Penelitian

Adapun manfaat penelitian dalam penelitian ini adalah sebagai berikut :

1. Bagi Pemerintah

Dengan pengelompokan (*cluster*) wilayah DKI Jakarta berdasarkan warga yang menerima vaksin COVID-19 maka dapat di ketahui wilayah atau daerah mana saja yang memiliki karakteristik yang sama sehingga penelitian ini dapat menjadi referensi dalam pengambilan kebijakan untuk melaksanakan program vaksinasi yang lebih merata pada kelurahan tertentu.

2. Bagi Peneliti

Mengetahui kondisi penyebaran vaksinasi DKI Jakarta dengan metode clustering dan penerapan algoritma K-Medoids.

3. Bagi Kontribusi Keilmuan

Diharapkan penelitian ini dapat menjadi rujukan, sumber informasi dan bahan referensi penelitian selanjutnya dalam pengembangan ilmu teknologi, khususnya pada sistem informasi.

I.6 Sistematika Penelitian

Penyusunan laporan skripsi pada penelitian ini diuraikan dalam beberapa bab dengan sistematika penulisan sebagai berikut:

Bab I Pendahuluan

Pada bab ini berisi uraian mengenai latar belakang penelitian, perumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah dan sistematika penulisan.

Bab II Tinjauan Pustaka

Pada bab ini berisi literatur yang relevan dengan permasalahan yang diteliti dan dibahas pula hasil-hasil penelitian terdahulu yang mempunyai keterkaitan dengan penelitian yang sedang berlangsung.

Bab III Metodologi Penelitian

Pada bab ini dijelaskan langkah-langkah penelitian secara rinci meliputi apa saja yang dilakukan pada saat penelitian dan terdapat konseptual yang dimodelkan dan sistematika suatu penulisan sehingga dapat digambarkan lebih jelas tentang apa saja yang dilakukan.

Bab IV Pengumpulan dan Pengolahan Data

Bab ini menjelaskan mengenai gambaran umum objek penelitian meliputi fungsi algoritma *K-medoids* yang terlibat dalam penelitian ini.

Bab V Analisis dan Pembahasan

Pada bab Analisis dan Pembahasan akan membahas hasil dari sebuah rancangan Algoritma *K-medoids* dan Hasil Pengujiannya.

Bab VI Kesimpulan dan Saran

Pada bab ini dijelaskan kesimpulan dari hasil penelitian yang dilakukan serta saran terhadap penelitian yang telah dilakukan.