

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Demam Berdarah (DBD) adalah penyakit menular yang disebabkan oleh virus *dengue* dari genus *Flavivirus*, yang ditularkan melalui gigitan nyamuk *Aedes aegypti*. Penyakit ini umumnya menyerang anak-anak dan dapat menimbulkan gejala berat seperti demam tinggi, pendarahan, hingga risiko kematian. Di Indonesia, DBD merupakan masalah kesehatan masyarakat yang serius, dengan jumlah kasus yang meningkat setiap tahun (Roffa, 2024).

Kasus Demam Berdarah (DBD) di Kota Medan menunjukkan tren peningkatan yang signifikan dalam beberapa tahun terakhir. Sepanjang tahun 2023, Kota Medan mencatat 965 kasus DBD dari total 4.687 kasus di Sumatera Utara, menjadikannya penyumbang terbesar di provinsi tersebut. Tren ini berlanjut pada tahun 2024, di mana dalam periode Januari hingga Juli, telah tercatat 558 kasus DBD di Kota Medan. Bahkan, pada Januari 2024 saja, Sumatera Utara melaporkan 405 kasus DBD dengan Kota Medan sebagai kontributor terbesar. Data ini mengindikasikan tingginya angka penyebaran DBD di Kota Medan dan perlunya perhatian lebih dalam upaya pencegahan dan pengendalian penyakit ini (Manurung, 2024).

Kenaikan ini dipengaruhi oleh berbagai faktor, seperti populasi vektor yang besar, tingkat virulensi virus, kekebalan komunitas, demografi, kepadatan penduduk, mobilitas pasien, kemampuan virus bereplikasi dalam tubuh nyamuk, serta perilaku manusia. Interaksi antara inang, vektor, dan kondisi lingkungan berperan penting dalam penyebaran dan perkembangan DBD, di mana faktor-faktor lingkungan juga memengaruhi tingkat kesehatan masyarakat secara keseluruhan (Hamid et al., 2023).

Salah satu strategi kunci dalam manajemen epidemiologi untuk DBD adalah pemetaan wilayah penanganan, yang mencakup identifikasi zona berisiko tinggi yang memerlukan intervensi intensif. Namun, salah satu permasalahan utama dalam penanganan demam berdarah di Kota Medan adalah kurangnya informasi yang

terstruktur mengenai persebaran wilayah yang membutuhkan prioritas penanganan. Data yang tersedia seringkali hanya berupa laporan kasus tanpa adanya analisis yang mendalam untuk mengidentifikasi pola atau klaster daerah rawan. Kondisi ini menghambat pihak berwenang dalam menentukan langkah-langkah strategis untuk pengendalian dan pencegahan. Salah satu pendekatan yang dapat digunakan untuk mengatasi permasalahan di atas adalah teknik data *mining* dengan metode *clustering*. Data *mining* adalah proses menggali, menemukan, dan memperoleh informasi penting atau wawasan tersembunyi dari kumpulan data yang besar. Salah satu tujuan utama data *mining* adalah menemukan pola atau tren tertentu yang dapat dimanfaatkan untuk mendukung prediksi dan pengambilan keputusan di masa depan. Metode *clustering*, dalam hal ini, merupakan teknik yang memisahkan titik-titik data ke dalam kelompok atau klaster tertentu. Data dengan karakteristik serupa akan dikelompokkan ke dalam satu klaster, sedangkan data dengan perbedaan karakteristik akan ditempatkan di klaster yang berbeda (Supriyadi et al., 2021). Dalam hal ini, dua algoritma klaster yang sering digunakan adalah *K-Means* dan *K-Medoids*.

Algoritma *K-Means* adalah metode analisis klaster yang bertujuan untuk membagi objek ke dalam  $k$  *cluster* dengan menjadikan nilai rata-rata (*mean*) sebagai pusat dari setiap klaster. Sebaliknya, algoritma *K-Medoids* menggunakan objek yang paling representatif (*medoid*) sebagai pusat klaster untuk masing-masing kelompok (Riyanto dalam Syamfithriani et al., 2023).

Terdapat beberapa penelitian telah dilakukan terkait penerapan algoritma *K-Means* dan *K-Medoids*, di antaranya penelitian yang dilakukan oleh (Syamfithriani et al., 2023) dalam evaluasi algoritma *K-Means* dan *K-Medoids* untuk pemetaan wilayah penanganan kasus diare pada balita di Kabupaten Kuningan, hasil analisis menggunakan metode *Silhouette* dan *Elbow* menunjukkan bahwa algoritma *K-Means* memiliki kinerja yang lebih baik dibandingkan algoritma *K-Medoids*. Sedangkan penelitian menurut (Kamila et al., 2019) Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau. Menunjukkan bahwa algoritma *K-Means* lebih cepat dalam pengolahan data, membutuhkan waktu rata-rata hanya 1 detik, sementara *K-Medoids* membutuhkan waktu sekitar 1 menit 38 detik. Nilai

DBI juga menunjukkan bahwa *K-Means* lebih unggul dengan nilai DBI yang lebih rendah.

Berdasarkan penelitian sebelumnya, penentuan jumlah kluster optimal sebagian besar dilakukan dengan metode DBI pada tahap evaluasi, sementara beberapa penelitian lain menggunakan salah satu dari metode *Silhouette Coefficient* atau metode *Elbow*. Oleh karena itu, penelitian ini menggunakan dua metode untuk menentukan jumlah kluster optimal, yaitu metode *Silhouette Coefficient* dan *Elbow*. Dengan menggunakan kedua metode ini, diharapkan hasil klusterisasi menggunakan algoritma *K-Means* dan *K-Medoids* dapat menghasilkan skema pemetaan yang lebih akurat untuk wilayah prioritas dalam upaya pencegahan dan pengendalian penyebaran penyakit demam berdarah dengue.

Dari hasil Analisa di atas, maka penulis tertarik untuk melakukan penelitian dengan judul “Perbandingan Algoritma *K-Means* dan *K-Medoids* Untuk Pemetaan Daerah penanganan Demam Berdarah di Kota Medan”.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, rumusan masalah yang diambil adalah:

1. Bagaimana cara menerapkan algoritma *K-Means* dan *K-Medoids* untuk melakukan klusterisasi data kasus demam berdarah di Kota Medan?
2. Bagaimana hasil klusterisasi yang dihasilkan oleh algoritma *K-Means* dibandingkan dengan *K-Medoids* dalam pemetaan wilayah berdasarkan tingkat kasus demam berdarah?
3. Algoritma mana yang memiliki performa lebih optimal antara *K-Means* dan *K-Medoids* dalam melakukan klusterisasi data kasus demam berdarah di Kota Medan?

## 1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Menerapkan algoritma *K-Means* dan *K-Medoids* untuk melakukan klusterisasi data kasus demam berdarah di Kota Medan.

2. Membandingkan hasil klasterisasi antara algoritma *K-Means* dan *K-Medoids* dalam pemetaan wilayah berdasarkan tingkat kasus demam berdarah.
3. Mengevaluasi performa kedua algoritma untuk menentukan algoritma yang lebih optimal.

#### 1.4 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut:

1. Memberikan kontribusi pada pengembangan ilmu pengetahuan, khususnya dalam bidang data *mining*, terkait penerapan algoritma *K-Means* dan *K-Medoids* untuk klasterisasi data.
2. Membantu instansi pemerintah, seperti Dinas Kesehatan Kota Medan, dalam memetakan wilayah berdasarkan tingkat kasus demam berdarah sehingga dapat menjadi referensi dalam pengambilan keputusan.
3. Memudahkan identifikasi daerah prioritas untuk intervensi dan alokasi sumber daya kesehatan, seperti penyemprotan *fogging* atau edukasi masyarakat.

#### 1.5 Ruang Lingkup dan Batasan Penelitian

Pada penelitian ini, terdapat sejumlah batasan yang telah ditentukan, yaitu sebagai berikut:

1. Penelitian ini hanya menggunakan data kasus demam berdarah di Dinas Kesehatan Kota Medan tahun 2024 dengan jumlah kasus 1102.
2. Penelitian ini terbatas pada dua algoritma, yaitu *K-Means* dan *K-Medoids*, untuk klasterisasi data.
3. Penelitian ini hanya memetakan wilayah persebaran kasus demam berdarah berdasarkan hasil klasterisasi tanpa memberikan analisis lebih lanjut terkait penyebab utama atau rekomendasi penanganan spesifik