

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Dalam dunia kesehatan, pemanfaatan teknologi informasi sangat berperan penting dalam pengelolaan data medis, termasuk data jenis penyakit pasien. Rumah sakit kini semakin banyak yang memanfaatkan sistem digital untuk mencatat, mengolah, dan menganalisis informasi kesehatan. Rumah Sakit Arun Lhokseumawe yang memiliki dengan luas 9450 m<sup>2</sup> dan 150 kamar untuk pasien, adalah salah satu rumah sakit besar di Kota Lhokseumawe, yang menjadikannya pusat layanan kesehatan penting bagi masyarakat (Putra, M. I. 2024).

Jumlah pasien di Rumah Sakit Arun terus bertambah setiap harinya, dengan berbagai jenis penyakit yang berbeda. Laporan yang dikenal sebagai data rekam medis dapat digunakan untuk melihat peningkatan jumlah pasien ini. Sayangnya, sebagian besar data tersebut masih sebatas laporan jumlah kunjungan dan jenis penyakit yang diderita pasien, tanpa diolah lebih lanjut agar memperoleh informasi yang lebih mendalam. Padahal, pihak rumah sakit memerlukan informasi tambahan yang lebih terstruktur, seperti pengelompokan jenis penyakit yang paling sering diderita pasien. Data ini bisa mendukung tenaga medis dalam menentukan langkah-langkah yang tepat untuk mencegah dan mengobati masalah yang ada (Andini & Arifin, 2020).

Seiring dengan perkembangan teknologi, penerapan data mining menjadi salah satu solusi agar dapat menunjang dalam pengelompokan penyakit pasien. Salah satu metode yang sering dipakai dalam data mining data adalah pengelompokan, yang memungkinkan pengelompokan jenis penyakit pasien rawat inap. Namun, pemilihan algoritma clustering yang tepat masih menjadi tantangan, sebab setiap metode mempunyai kelebihan serta keterbatasan tersendiri (Maulia et al., 2021). Penelitian terdahulu terkait dengan topik penelitian ini yaitu, penelitian yang dilakukan di Puskesmas Tlogomulyo berjudul “Clustering Data Pasien Berdasarkan Usia di Puskesmas Menerapkan Metode K-Means” yang membahas

tentang pengelompokan penyakit berdasarkan usia pasien. Hasilnya menyoroti dua kelompok utama, skabies pada usia 10-20 tahun dan ISPA pada kelompok usia 50-55 tahun (Herlin Lutfiannisa et al., 2024). Penelitian lainnya yang berjudul “Implementasi Algoritma K-Means Clustering Untuk Pengelompokan Penyakit Pasien Pada Puskesmas Pulo Brayan” yang berfokus dalam mengelompokkan jenis penyakit berdasarkan usia, jenis kelamin, dan diagnosis. Analisis ini menghasilkan tiga cluster dengan tingkat risiko berbeda, di mana cluster risiko tinggi didominasi oleh pasien laki-laki berusia di atas 40 tahun dengan diagnosis utama ISPA (Okta Jaya Harmaja1, 2023). Selanjutnya penelitian yang berjudul “Implementasi Algoritma K-Medoids untuk Klasterisasi Data Penyakit Pasien di RSUD Kota Bandung” Studi ini mengkaji pengelompokan penyakit yang dialami oleh pasien dibagi menjadi 3 kelompok berdasarkan kamar dan jumlah pasien. Cluster 0 mencakup penyakit di klinik gawat darurat dengan total 18 data. Cluster 1 adalah penyakit di klinik kemuning dan klinik gawat darurat. Sedangkan cluster 3 berisi 20 data terkait klinik rehabilitasi medis (Andini & Arifin, 2020). Ketiga penelitian ini menunjukkan bagaimana clustering dapat membantu mengidentifikasi pola penyakit dan karakteristik pasien yang relevan, mendukung pengambilan keputusan yang lebih baik dalam perencanaan dan penyediaan layanan kesehatan.

Berdasarkan beberapa penelitian terdahulu penulis menemukan bahwa penelitian ini memiliki perbedaan utama dibandingkan penelitian terdahulu, yaitu dengan membandingkan dua algoritma klasterisasi untuk menentukan metode yang paling optimal dalam mengelompokkan jenis penyakit yang sering diderita pasien rawat inap. Penelitian ini mempertimbangkan aspek validasi hasil klasterisasi dengan menggunakan metrik evaluasi *Silhouette Coefficient*. Matrik ini digunakan untuk menilai seberapa baik struktur cluster yang dihasilkan oleh masing-masing algoritma (Dbscan & Hasan, 2024).

Dengan demikian, penelitian ini tidak hanya berfokus pada hasil pengelompokan penyakit, tetapi juga memperhatikan kualitas cluster yang terbentuk, termasuk koherensi internal dan pemisahan yang jelas antar kelompok. Dengan mengevaluasi efektivitas kedua algoritma yang digunakan, penelitian ini

diharapkan dapat memberikan rekomendasi metode klasterisasi yang lebih optimal dalam mengelompokkan jenis penyakit yang rentan diderita pasien rawat inap. Di sisi lain, hasil penelitian ini juga diharapkan mampu menyumbangkan wawasan bagi tenaga medis dan pengelola rumah sakit mengenai pola distribusi penyakit yang paling dominan diderita oleh pasien rawat inap di Rumah Sakit Arun. Informasi tersebut dapat dimanfaatkan dalam perencanaan pelayanan kesehatan yang lebih efisien

## **1.2 Rumusan Masalah**

Dari latar belakang masalah diatas, maka rumusan masalah yang hendak diulas dalam penelitian ini, yaitu.

1. Bagaimana cara mengelompokkan jenis penyakit pasien rawat inap dengan menggunakan metode *Agglomerative Hierarchical Clustering* dan *K-Medoids*?
2. Metode klasterisasi mana yang lebih akurat dan optimal dalam mengelompokkan jenis penyakit pasien rawat inap?

## **1.3 Tujuan Penelitian**

Adapula tujuan yang hendak dicapai dalam penelitian ini adalah:

1. Mengelompokkan jenis penyakit pasien rawat inap dengan menggunakan metode *Agglomerative Hierarchical Clustering* dan *K-Medoids*.
2. Mengevaluasi kualitas klasterisasi yang terbentuk menggunakan matrik evaluasi *Silhouette Coefficient* untuk menentukan algoritma yang paling optimal dalam mengelompokkan jenis penyakit pasien rawat inap.

## **1.4 Ruang Lingkup dan Batasan Masalah**

Supaya penelitian ini menjadi lebih terfokus dan hasil yang diperoleh lebih valid, terdapat beberapa batasan yang diterapkan dalam penelitian ini, yaitu:

1. Data yang digunakan dalam studi ini merupakan data jenis penyakit pasien yang dirawat di Rumah Sakit Arun.

2. Adapun variabel yang digunakan dalam penelitian ini mencakup nama penyakit, jumlah pasien laki-laki, jumlah pasien perempuan, dan umur pasien.
3. Penelitian ini berfokus pada perbandingan kinerja dua metode klasterisasi, yaitu metode *Agglomerative Hierarchical Clustering* dengan pendekatan *Average Linkage* yang akan dibandingkan dengan metode *K-Medoids*.
4. Pengelompokan dilakukan ke dalam tiga cluster utama, yaitu cluster tinggi, cluster sedang, dan cluster rendah.

### **1.5 Manfaat Penelitian**

Penelitian ini diharapkan mampu menyumbangkan manfaat, adapula manfaat dari penelitian ini, yaitu:

1. Memberikan wawasan bagi tenaga medis dan pengelola rumah sakit mengenai jenis penyakit yang paling dominan, sehingga dapat membantu dalam pengadaan fasilitas medis dan obat-obatan yang lebih tepat sasaran.
2. Sebagai sarana bagi penulis untuk mengimplementasikan pengetahuan yang diperoleh selama menempuh studi.
3. Menjadi acuan dan dasar pengembangan untuk penelitian serupa, baik dalam konteks kesehatan pasien maupun penerapan algoritma klasterisasi dalam masalah medis lainnya.