

BAB III

METODOLOGI PENELITIAN

3.1. Metode Penelitian

Metode yang digunakan dalam penelitian ini adalah *K-Means*, sebuah algoritma *Clustering* yang berfungsi untuk mengelompokkan data ke dalam beberapa *Cluster* berdasarkan kemiripan karakteristik. *K-Means* termasuk dalam metode unsupervised learning yang sering digunakan dalam analisis data untuk menemukan pola tersembunyi dalam kumpulan data yang besar. Dengan cara ini, data yang memiliki karakteristik serupa akan dikelompokkan dalam satu *Cluster*, sementara data yang berbeda akan dimasukkan ke dalam *Cluster* lain. Proses ini sangat berguna dalam berbagai bidang, termasuk dalam analisis layanan kesehatan, di mana pasien dapat dikategorikan berdasarkan pola perawatan yang diterima.

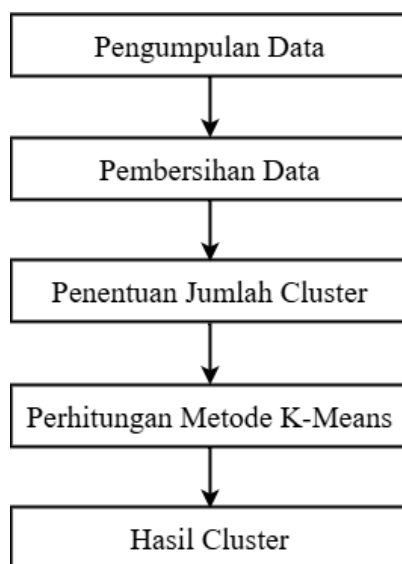
K-Means bekerja dengan menentukan sejumlah centroid secara acak sebagai titik awal *Cluster*. Setiap data kemudian diklasifikasikan ke dalam *Cluster* berdasarkan jarak terdekatnya ke centroid tersebut. Setelah semua data terklasifikasi, centroid akan diperbarui dengan menghitung rata-rata posisi seluruh data dalam setiap *Cluster*. Proses ini terus berulang hingga tidak ada lagi perubahan signifikan dalam pembagian *Cluster* atau telah mencapai jumlah iterasi maksimum yang telah ditentukan. Algoritma ini sangat efisien dalam mengolah data dalam jumlah besar dan mampu memberikan hasil yang lebih cepat dibandingkan metode *Clustering* lainnya.

Dalam penelitian ini, *K-Means* digunakan untuk menganalisis data rekam

medis pasien di RSUD Rantau Prapat. Dengan menerapkan metode ini, pasien dapat dikelompokkan berdasarkan pola layanan kesehatan yang mereka terima, seperti jenis penyakit, frekuensi kunjungan, dan kategori perawatan. Hasil dari *Clustering* ini diharapkan dapat membantu rumah sakit dalam memahami karakteristik pasien dengan lebih baik serta meningkatkan efektivitas pelayanan kesehatan.

3.2. Kerangka Penelitian

Pada penelitian ini terdapat kerangka kerja yang dirancang dalam bentuk *Flowchart* yaitu sebagai berikut.

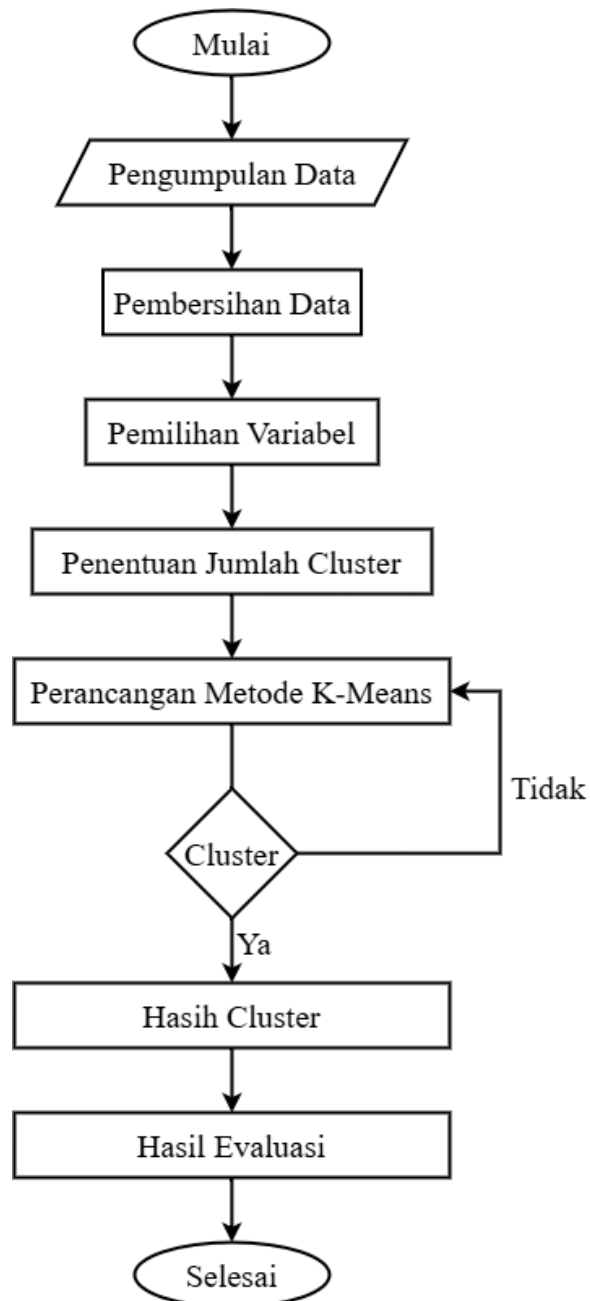


Gambar 3. 1. Kerangka Kerja Penelitian

Pada tahapan diatas merupakan tahapan yang akan dilakukan dalam melakukan perancangan metode *K-Means* untuk merancang perhitungannya. Untuk penjelasan dari setiap poin diatas yaitu sebagai berikut.

- 1) Pengumpulan data – Data rekam medis pasien RSUD Rantau Prapat dikumpulkan sebagai sumber utama dalam penelitian ini untuk dianalisis lebih lanjut.

- 2) Pembersihan data – Data yang telah dikumpulkan diperiksa, difilter, dan dibersihkan dari duplikasi, data kosong, atau nilai yang tidak valid agar hasil analisis lebih akurat.
- 3) Penentuan jumlah kluster – Jumlah kluster optimal ditentukan menggunakan metode evaluasi seperti Elbow Method atau Silhouette Score untuk mendapatkan hasil klustering yang lebih akurat.
- 4) Perhitungan metode *K-Means* – Algoritma *K-Means* diterapkan dengan memilih centroid awal secara acak, menghitung jarak setiap data ke centroid, dan memperbarui posisi centroid hingga mencapai kondisi optimal.
- 5) Hasil kluster – Data pasien dikelompokkan ke dalam kluster tertentu berdasarkan pola layanan kesehatan mereka, yang kemudian dianalisis untuk meningkatkan efektivitas layanan di RSUD Rantau Prapat.



Gambar 3. 2. *Flowchart Metode K-Means*

Pada penelitian ini, terdapat *Flowchart* untuk melakukan penerapan metode *K-Means*. Untuk penjelasan tiap poinnya yaitu sebagai berikut:

- 1) Pengumpulan Data – Data rekam medis pasien RSUD Rantau Prapat dikumpulkan sebagai dasar untuk proses klastering menggunakan metode *K-Means*.
- 2) Pembersihan Data – Data diperiksa dan difilter untuk menghapus duplikasi, nilai yang hilang, atau data yang tidak relevan agar meningkatkan akurasi analisis.
- 3) Pemilihan Variabel – Variabel yang berpengaruh dalam pengelompokan pasien, seperti jenis penyakit, frekuensi kunjungan, dan layanan yang diterima, dipilih untuk diproses lebih lanjut.
- 4) Penentuan Jumlah *Cluster* – Jumlah klaster optimal ditentukan menggunakan metode evaluasi seperti Elbow Method atau Silhouette Score untuk mendapatkan hasil yang lebih representatif.
- 5) Perancangan Model *Cluster* – Model klaster dibuat dengan menginisialisasi centroid secara acak dan menyesuaikannya berdasarkan data yang telah dipilih.
- 6) Klaster – Proses klastering dilakukan dengan mengelompokkan pasien ke dalam beberapa klaster berdasarkan kemiripan karakteristik yang dianalisis.
- 7) Hasil *Cluster* – Data pasien telah terbagi ke dalam klaster tertentu yang mencerminkan pola layanan kesehatan yang mereka terima.
- 8) Hasil Evaluasi – Hasil klastering dievaluasi untuk memastikan efektivitas pengelompokan dan memberikan rekomendasi perbaikan dalam layanan kesehatan RSUD Rantau Prapat.

3.3. Lokasi dan Waktu Penelitian

3.3.1. Lokasi

Penelitian ini dilakukan di RSUD Rantau Prapat, sebuah rumah sakit umum daerah yang berlokasi di Kabupaten Labuhanbatu, Sumatera Utara. Sebagai fasilitas kesehatan rujukan utama di wilayah tersebut, RSUD Rantau Prapat melayani berbagai jenis pasien dengan beragam kebutuhan medis, mulai dari layanan rawat jalan hingga perawatan intensif. Dengan jumlah pasien yang terus meningkat, rumah sakit ini menghadapi tantangan dalam mengelola data rekam medis secara efisien.

3.3.2. Waktu

Pada penelitian ini, untuk waktu penelitian akan dipaparkan dalam bentuk tabel yaitu sebagai berikut.

Tabel 3. 1. Waktu Penelitian

No	Nama Kegiatan	Bulan																			
		Februari				Maret				April				Mei				Juni			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	Identifikasi dan ekstraksi data pasien																				
2	Pembersihan dan standarisasi data																				
3	Seleksi fitur utama untuk <i>Clusterisasi</i>																				
4	Penentuan jumlah <i>Cluster</i> optimal																				
5	Implementasi algoritma <i>K-Means</i>																				

mengidentifikasi pola tertentu dalam layanan kesehatan. Dengan bantuan RapidMiner, analisis dapat dilakukan secara lebih efisien, memungkinkan visualisasi hasil *Clusterisasi* yang memudahkan interpretasi data. Hasil dari analisis ini diharapkan dapat memberikan wawasan yang lebih jelas bagi rumah sakit dalam meningkatkan efektivitas dan kualitas pelayanan kesehatan berdasarkan pola pasien yang terbentuk.

3.6. Perhitungan Data

Perhitungan data dalam penelitian ini dilakukan untuk mengolah informasi dari data rekam medis pasien yang telah dikumpulkan, seperti menghitung total ID Pasien, jenis layanan, cara bayar dan status pasien. Proses ini bertujuan untuk memperoleh gambaran awal mengenai pola data sebelum dilakukan proses klasterisasi menggunakan metode *K-Means*. Dengan perhitungan yang tepat, data menjadi lebih siap untuk dianalisis dan menghasilkan klaster yang akurat.

3.6.1. Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan melalui teknik dokumentasi, yaitu dengan meminta data rekam medis pasien secara resmi kepada pihak RSUD Rantau Prapat. Data yang dikumpulkan mencakup informasi penting seperti ID Pasien, jenis layanan, cara bayar dan status pasien. Teknik ini dipilih karena data rekam medis merupakan sumber informasi yang valid dan relevan untuk dianalisis guna mendukung tujuan penelitian, yaitu meningkatkan kualitas layanan kesehatan melalui proses klasterisasi data pasien.

Tabel 3. 2. Variabel Penelitian

No	Nama Variable	Tipe Variabel	Deskripsi
1	ID Pasien	Numerik	Demi menjaga keamanan data pasien maka nama pasien diganti dengan ID pasien. Variable ini tidak digunakan dalam proses perhitungan.
1	Jenis Layanan	Kualitatif (numerik)	Tipe layanan medis yang diterima pasien, merepresentasikan jenis perawatan yang dijalani.
2	Cara Bayar	Kualitatif (dikodekan ke numerik)	Metode pembayaran yang digunakan oleh pasien untuk jenis bayar.
3	Status Pasien	Kualitatif (dikodekan ke numerik)	Kondisi pasien selama atau setelah perawatan, yang menggambarkan status kesehatan pasien.

Dalam penelitian ini, variabel yang digunakan untuk pengelompokan pasien rekam medis adalah ID Pasien, jenis layanan, cara bayar dan status pasien. Variabel Jenis Layanan menggambarkan tipe layanan medis yang diterima oleh pasien selama menjalani perawatan di rumah sakit, seperti rawat inap, rawat jalan, rehabilitasi, dan tindakan bedah. Variabel ini penting untuk mengidentifikasi karakteristik layanan yang dipilih pasien dan pola pemanfaatan layanan kesehatan. Selanjutnya, variabel Cara Bayar merepresentasikan metode pembayaran yang digunakan pasien dalam membiayai perawatan medisnya, yang dapat berupa asuransi, pembayaran mandiri, BPJS, atau layanan gratis. Variabel ini membantu dalam memahami aspek pembiayaan yang dapat berpengaruh terhadap pengelolaan dan pemetaan pasien berdasarkan kemampuan dan opsi pembayaran mereka. Terakhir, variabel Status Pasien menunjukkan kondisi atau status kesehatan pasien selama atau setelah perawatan, yang meliputi pasien yang sembuh, melakukan kontrol rutin, masih dalam perawatan, atau sedang dirujuk ke fasilitas lain. Ketiga variabel ini diubah ke dalam format numerik agar dapat diolah menggunakan

algoritma *K-Means*, yang bekerja berdasarkan pengukuran jarak antar titik data. Dengan menggunakan kombinasi variabel ini, diharapkan pengelompokan pasien dapat mencerminkan kelompok pasien dengan karakteristik layanan, pembiayaan, dan kondisi kesehatan yang serupa, sehingga dapat meningkatkan efisiensi dan mutu layanan kesehatan di RSUD Rantau Prapat.

Tabel 3. 3. Dataset Pasien Rekam Medis

ID Pasien	Jenis Layanan	Cara Bayar	Status Pasien
P001	Rawat Jalan	BPJS	Kontrol Rutin
P002	Rawat Inap	BPJS	Sembuh
P003	Rawat Jalan	Mandiri	Sembuh
P004	Rawat Inap	Gratis	Dirujuk
P005	Rehabilitasi	Gratis	Dalam Perawatan
P006	Rawat Jalan	Gratis	Kontrol Rutin
P007	Rehabilitasi	BPJS	Kontrol Rutin
P008	Rehabilitasi	Gratis	Kontrol Rutin
P009	Rawat Jalan	Gratis	Dirujuk
P010	Rawat Inap	Asuransi	Kontrol Rutin
.....			
.....			
P090	Bedah	Gratis	Sembuh
P091	Bedah	Gratis	Kontrol Rutin
P092	Rehabilitasi	BPJS	Dirujuk
P093	Rawat Jalan	Mandiri	Dirujuk
P094	Rawat Jalan	Asuransi	Dirujuk
P095	Rawat Inap	Asuransi	Dalam Perawatan
P096	Rawat Jalan	Mandiri	Dalam Perawatan
P097	Rawat Jalan	Asuransi	Kontrol Rutin
P098	Rehabilitasi	BPJS	Sembuh
P099	Rehabilitasi	BPJS	Kontrol Rutin
P100	Rawat Jalan	Gratis	Sembuh

Tabel tersebut berisi data sampel penelitian yang terdiri dari 100 pasien rekam medis di RSUD Rantau Prapat. Data ini mencakup informasi penting seperti ID Pasien, jenis layanan, cara bayar dan status pasien. Ketiga variabel ini diubah ke

dalam format numerik agar dapat diolah menggunakan algoritma *K-Means*, yang bekerja berdasarkan pengukuran jarak antar titik data. Dengan menggunakan kombinasi variabel ini, diharapkan pengelompokan pasien dapat mencerminkan kelompok pasien dengan karakteristik layanan, pembiayaan, dan kondisi kesehatan yang serupa, sehingga dapat meningkatkan efisiensi dan mutu layanan kesehatan di RSUD Rantau Prapat.

Pembersihan Data

Preprocessing data adalah tahap awal dalam pengolahan data yang bertujuan untuk membersihkan, memformat, dan menyiapkan data agar siap dianalisis. Proses ini penting untuk memastikan data berada dalam bentuk yang sesuai dengan metode yang akan digunakan. Pada penelitian ini, *preprocessing* data dilakukan dengan mengubah data sampel rekam medis yang bersifat kategorikal menjadi data numerikal agar dapat diproses menggunakan metode.

Tabel 3.4. Transformasi data rekam medis

Atribut	Partisi	Nilai
Jenis Layanan	Rawat Inap	1
	Rawat Jalan	2
	Rehabilitasi	3
	Bedah	4
Jenis bayar	Asuransi	11
	Mandiri	12
	BPJS	13
	Gratis	14
Status Pasien	Sembuh	21
	Kontrol Rutin	22
	Dalam Perawatan	23
	Dirujuk	24

Pada tabel di atas ditampilkan hasil transformasi data kategorik menjadi data numerik yang digunakan dalam proses klasterisasi dengan metode *K-Means*. Atribut seperti Jenis Layanan, Jenis bayar, dan Status Pasien telah dipartisi dan diberi nilai numerik untuk memudahkan proses perhitungan dan pengelompokan

data. Sebagai contoh, Jenis Layanan seperti Rawat Inap diberi nilai 1, Bedah diberi nilai 2, dan Rehabilitasi diberi nilai 3 sesuai dengan tingkat intensitas dan jenis penanganannya. Jenis bayar dikonversi langsung ke nilai representatif, sedangkan Status Pasien seperti Sembuh, Kontrol Rutin, Dalam Perawatan, dan Dirujuk masing-masing diberi bobot antara 21 hingga 24. Transformasi ini penting agar algoritma *K-Means* dapat mengelompokkan data berdasarkan karakteristik numerik yang seragam dan relevan.

3.6.2. Penentuan Jumlah *Cluster*

Penentuan jumlah klaster pada penelitian ini ditetapkan sebanyak 3 klaster agar data pasien dapat dikelompokkan secara optimal berdasarkan kemiripan karakteristik. Pemilihan ini mempertimbangkan variasi dalam jenis layanan, jenis bayar, dan status pasien, sehingga tiga klaster dinilai cukup representatif untuk menggambarkan pola layanan kesehatan di RSUD Rantau Prapat. Dengan jumlah klaster ini, hasil analisis diharapkan dapat memberikan informasi yang lebih terfokus dan mudah diinterpretasikan.