

## BAB III

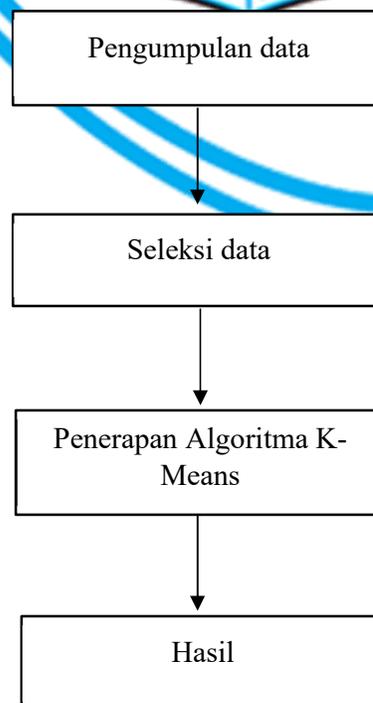
### METODE PENELITIAN

#### 1.1. Gambaran umum

Berdasarkan data yang digunakan adalah data jumlah penderita penyakit filariasis berdasarkan provinsi di Indonesia yang didapat secara *online*. Data tersebut akan diolah dengan teknik data mining dengan menerapkan metode clustering. Salah satu metode *clustering* yaitu menggunakan Algoritma k-means untuk mengelompokkan daerah penderita filariasis. Pengelompokan penderita filariasis diantaranya kelompok daerah tinggi, sedang atau rendah berdasarkan provinsi. Kemudian dilakukannya pengujian dengan menggunakan *tools* yaitu RapidMiner.

#### 1.2. Metode penelitian

Metode penelitian yang dilakukan yaitu pengumpulan data, seleksi data, penerapan Algoritma K-means, dan hasil.



Gambar 3.1 Tahapan penelitian

### 1.2.1. Pengumpulan data

Tahap pengumpulan data sangat penting untuk mendukung penelitian ini. Pengumpulan data dilakukan untuk mendapatkan informasi dan data. Data yang akan digunakan diunduh dari *situs web* BPS (Badan Pusat Statistik). Alamat *situs web* dari <https://data.go.id/dataset/penderita-kasus-filariasis-per-provinsi> pada tanggal 21 januari 2019 pada pukul 09.15 WIB. Data yang didapat dari data penderita filariasis dari tahun 2008 sampai dengan tahun 2012. Terdapat 33 provinsi di Indonesia terdapat atribut nama provinsi, tahun dan jumlah penderita filariasis berdasarkan tahun.

### 1.2.2. Seleksi data

Data jumlah penderita filariasis diolah terlebih dahulu agar memiliki *sampling error* yang kecil. Sehingga data tersebut akan menjadi *dataset* sesuai kebutuhan. Dari data yang didapat yaitu data penderita filariasis semua data selama 5 tahun digunakan untuk melakukan perhitungan secara manual. Selanjutnya pengujian menggunakan rapidminer.

### 1.2.3. Penerapan Algoritma K-means

Tahap ini melakukan pengolahan data sesuai dengan tahapan - tahapan Algoritma K-means:

1. Pengelompokan akan dibentuk menjadi 3 *cluster* yang tinggi, sedang, rendah
2. Menentukan titik pusat awal dari setiap *cluster* (*centroid 1*, *centroid 2* dan *centroid 3*) dari data jumlah penderita filariasis yang nantinya dipilih secara random.
3. Menghitung setiap data jumlah penderita filariasis ke centroid masing-masing (*centroid 1*, *centroid 2*, *centroid 3*).
4. Setiap data jumlah penderita filariasis memilih *centroid* terdekat dengan memasukan nilai awal jumlah penderita filariasis.
5. Menentukan posisi centroid yang baru dengan cara menghitung nilai rata-rata dari data-data yang terletak pada *centroid* yang sama.

6. Kembali ke langkah ke-3 sampai posisi centroid baru dengan yang lama tidak sama.

#### 1.2.4. Hasil

Hasil pengelompokan ini nantinya terdapat 3 kelompok yaitu tinggi, sedang dan rendah. Dari 33 provinsi nanti akan terlihat jumlah provinsi berdasarkan *cluster*. Kemudian akan diketahui provinsi mana saja yang terdapat penderita filariasis termasuk tinggi, sedang dan rendah. Hasil penelitian ini dapat digunakan sebagai penunjang keputusan penderita filariasis berdasarkan provinsi. Selain itu berharap mejadi solusi dalam menentukan tindakan. Tindakan berupa pencegahan atau penanggulangan bagi daerah yang memiliki penderita filariasis berdasarkan kondisi tinggi, sedang, rendahnya kasus filariasis.



