

## **BAB III**

### **METODOLOGI PENELITIAN**

#### **3.1 Gambaran Umum Penelitian**

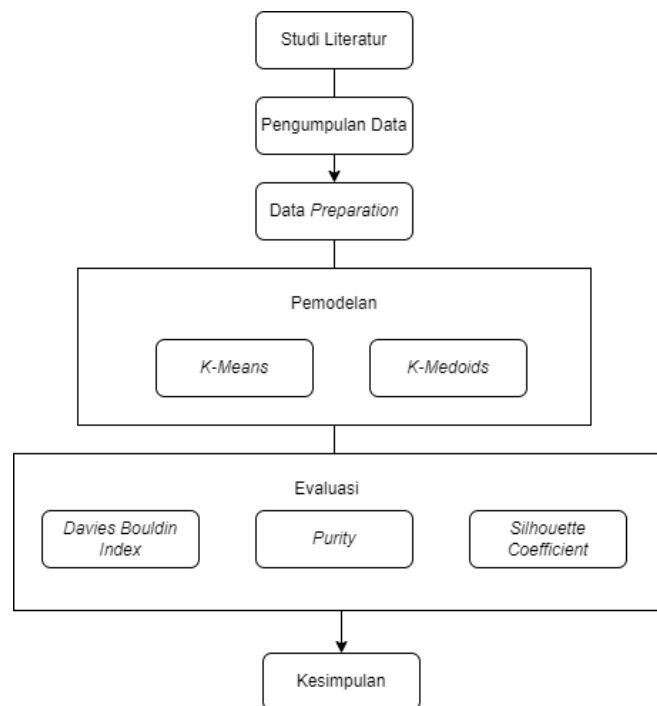
Penelitian ini dilakukan untuk mengelompokkan *cluster* jumlah positif *Covid-19* di Jawa Barat menggunakan metode *K-Means* dan *K-Medoids* dengan evaluasi *cluster* menggunakan *Davies-Bouldin Index* (DBI), *Purity* dan *Silhouette Coefficient*. Penelitian ini bertujuan untuk mengetahui algoritma *Clustering* terbaik antara *K-Means* dan *K-Medoids* dalam pengelompokan penyebaran *Covid-19* di Jawa Barat dengan evaluasi menggunakan *Davies-Bouldin Index* (DBI), *Purity* dan *Silhouette Coefficient* dimana hasilnya akan mengetahui efisiensi dari algoritma tersebut. Selain itu tujuan dari penelitian ini diharapkan dapat membantu sebagai alat bagi pemerintah atau pemangku kepentingan lainnya dalam menangani penderita positif *Covid-19* di Jawa Barat supaya mengurangi penyebaran *Covid-19* dan resiko kematian akibat *Covid-19* ataupun kasus serupa lainnya dijadikan pengetahuan atau sebagai acuan untuk penelitian selanjutnya.

#### **3.2 Sumber Data**

Data yang digunakan pada penelitian ini adalah data sekunder yang berasal dari Open Data Jawa Barat Satuan Tugas Penanganan *Corona Virus Disease* 2019 (SATGAS *COVID-19*). Data yang digunakan yaitu data *Covid-19* di Jawa Barat dari 02 Maret 2020 sampai dengan 30 Juni 2022.

### 3.3 Tahapan Penelitian

Pada bagian ini menjelaskan tahapan-tahapan sistematis yang akan dijadikan acuan sebagai kerangka penelitian. Tahapan penelitian terdiri dari studi literatur, pengumpulan data, data *preparation*, pemodelan *clustering* dengan algoritma *K-Means* dan *K-Medoids*, evaluasi dengan metode *Davies-Bouldin Index* (DBI), *Purity*, *Silhouette Coefficient* dan kesimpulan. Tahapan dari penelitian ini dapat dilihat pada gambar 3.1 berikut.



Gambar 3. 1 Tahapan Penelitian

#### 1. Studi Literatur

Studi literatur dilakukan dengan mencari beberapa penelitian yang terkait dengan mempelajari dari beberapa sumber seperti buku, jurnal, serta laporan penelitian yang berkaitan dengan penyebaran *Covid-19*, algoritma *K-Means*, dan Algoritma *K-Medoids* serta evaluasi *Davies-Bouldin Index* (DBI), *Purity*

dan *Silhouette Coefficient*.

## 2. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data *Covid-19* di Jawa Barat dari 02 Maret 2020 sampai dengan 30 Juni 2022 yang diambil dari Satuan Tugas Penanganan *Corona Virus Disease 2019* (SATGAS *COVID-19*).

## 3. Data Preparation

Aktivitas yang dilakukan pada tahap ini meliputi pemilihan data yang akan digunakan untuk proses modeling, proses data *preprocessing* seperti mengatasi *outlier*, *noisy data*, *inconsistent* dan *missing value* yang ada dan data *transformation*, melakukan normalisasi data. Hasil dari tahap ini akan digunakan untuk proses pemodelan.

## 4. Pemodelan

Dalam penelitian ini diperlukan proses pengujian model yang diusulkan. Proses pengujian model/metode menggunakan bagian dari data yang ada. Semua data kemudian diuji menggunakan metode *K-Means* dan *K-Medoids* dengan bantuan aplikasi Rapid Miner, Ms. Excel dan lain-lain.

## 5. Evaluasi

Setelah selesai pada proses pemodelan menggunakan algoritma *K-Means* dan *K-Medoids*. Tahap selanjutnya dilakukan evaluasi terhadap model yang telah dibuat menggunakan *Davies-Bouldin Index* (DBI), *Purity* dan *Silhouette Coefficient* yang bertujuan untuk mengetahui kualitas dari suatu objek pada suatu *cluster*. Dalam melakukan pengukuran pada kualitas *cluster* dengan menggunakan pengujian yang berpengaruh pada jumlah nilai *cluster*. Sehingga

hasil akhirnya akan mendapatkan nilai efisiensi terbaik dari 2 metode algoritma yang digunakan.

#### 6. Kesimpulan

Tahap terakhir yaitu menarik kesimpulan dari hasil membandingkan 2 metode atau algoritma *Clustering* dengan 3 metode evaluasi yaitu *Davies-Bouldin Index (DBI)*, *Purity* dan *Silhouette Coefficient* sehingga dapat diketahui metode mana yang paling terbaik dengan nilai efisiensi terbaik dalam mengelompokkan jumlah positif *Covid-19* di Jawa Barat.