BAB III
METODE PENELITIAN

1.1 Objek Penelitian

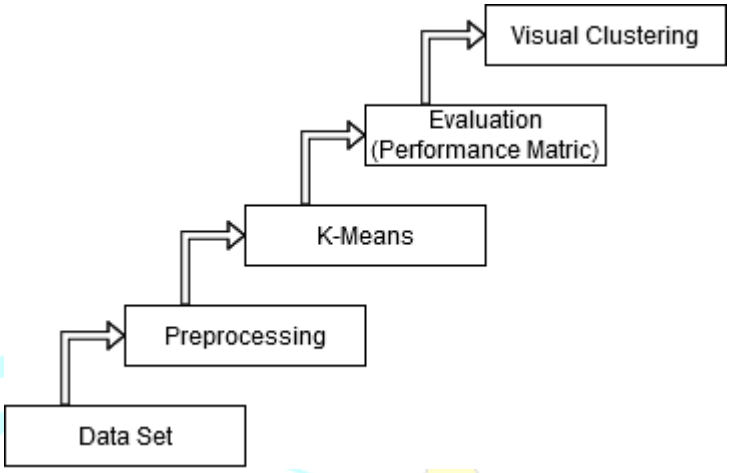
Objek penelitian ini adalah tingkat kemiskinan pada tingkat provinsi di Indonesia yang dianalisis melalui data sekunder dari Badan Pusat Statistik (BPS). Data didapat dari open data jabar yang diperoleh dari link https://opendata.jabarprov.go.id/. Objek ini dipilih karena kemiskinan merupakan salah satu indikator utama dalam mengukur keberhasilan pembangunan, sekaligus sebagai isu strategis nasional yang membutuhkan perhatian serius dari berbagai pihak, khususnya pemerintah pusat dan daerah. Objek penelitian difokuskan pada data kuantitatif yang merepresentasikan kondisi kemiskinan di 38 provinsi di Indonesia selama periode tahun 2010 hingga 2024. Indikator kemiskinan yang digunakan meliputi: persentase penduduk miskin, jumlah penduduk miskin, indeks kedalaman kemiskinan, indeks keparahan kemiskinan

Seluruh indikator ini diolah menjadi representasi rata-rata tahunan untuk masing-masing provinsi. Dengan mengelompokkan provinsi berdasarkan kemiripan indikator-indikator tersebut, penelitian ini bertujuan untuk mengidentifikasi pola regional kemiskinan yang dapat mendukung perumusan kebijakan pembangunan berbasis data.

3.2 Prosedur Penelitian

Dalam penelitian ini, proses analisis data dilakukan dengan pendekatan *clustering K-Means*. Untuk memastikan hasil analisis yang valid dan dapat diinterpretasikan, langkah-langkah sistematis diikuti mulai dari pengolahan data awal hingga evaluasi performa model. Diagram alur penelitian *K-Means clustering* tersebut dapat di lihat pada tabel 2. Dibawah ini

Tabel 1 Diagram alur penelitian K-Means Clustering



Dari gambar tersebut menggambarkan secara keseluruhan tahapan-tahapan yang dilakukan dalam penelitian ini. Proses dimulai dengan dataset, dalam konteks *clustering* pada data *mining*, dataset biasanya berisi sekumpulan data mentah yang memiliki sejumlah fitur atau atribut yang mewakili objek yang akan dikelompokkan yang menjadi input utama untuk analisis (Chaudhry et al., 2023b). Dataset ini kemudian melalui proses *preprocessing*, Tahap *preprocessing* dalam analisis *clustering* adalah langkah penting untuk memastikan data bersih, relevan, dan siap digunakan oleh algoritma (Simon et al., 2020). Proses ini dimulai dengan memahami karakteristik dataset melalui analisis deskriptif untuk mengidentifikasi distribusi data, *missing values*, dan *outlier*. Langkah berikutnya adalah pembersihan data, termasuk imputasi nilai yang hilang, penghapusan data duplikat, serta perbaikan kesalahan format (Winkler, 2020).

Setelah *preprocessing* selesai, data dianalisis algoritma *clustering K-Means*, tahap modeling dalam penelitian ini bertujuan untuk mengelompokkan provinsi-provinsi di Indonesia berdasarkan tingkat kemiskinan menggunakan algoritma *tersebut.* Algoritma *K-Means clustering* bekerja dengan membagi data ke dalam sejumlah kelompok *(cluster)* berdasarkan kedekatan jarak antar data dalam ruang multidimensi. Jumlah *cluster* ditentukan melalui metode evaluasi seperti *Elbow Method* untuk menentukan jumlah *cluster* optimal dengan melihat grafik inertia (dalam hal ini, total jarak antar data dan *centroid*nya). Penurunan yang signifikan pada grafik menunjukkan jumlah cluster yang tepat. Titik di mana penurunan

melambat disebut sebagai *"elbow"* dan menjadi indikasi jumlah *cluster* yang optimal (Amit, 2021).

Unutk mengevaluasi metode K-Means digunakan *Silhouette Score. Silhouette Score* mengukur seberapa baik objek dikelompokkan dalam *cluster* mereka, dengan nilai antara -1 dan +1. Semakin tinggi nilai *Silhouette*, semakin baik pengelompokan yang dilakukan. Nilai *Silhouette* yang optimal memberikan indikator kuat untuk jumlah *cluster* yang tepat. Setelah proses *clustering* dan evaluasi selesai, hasil analisis akan divisualisasikan untuk memberikan gambaran yang jelas tentang distribusi provinsi berdasarkan tingkat kemiskinan. Untuk visualisasi ini, digunakan plot *scatter* yang menunjukkan pengelompokan provinsi dengan menggunakan warna yang berbeda untuk setiap *cluster*. Dalam visualisasi hasil *clustering*, peneliti akan menganalisis karakteristik masing-masing *cluster* untuk mengidentifikasi provinsi-provinsi dengan pola kemiskinan serupa. Hal ini akan memberikan wawasan mengenai ketimpangan kemiskinan antarprovinsi dan memungkinkan peneliti untuk memberikan rekomendasi kebijakan yang lebih tepat sasaran. Misalnya, provinsi dengan tingkat kemiskinan tinggi dan karakteristik serupa dapat diarahkan untuk menerima kebijakan pengentasan kemiskinan yang lebih terfokus dan berbasis wilayah.