

BAB III

METODE PENELITIAN

3.1 Objek Penelitian

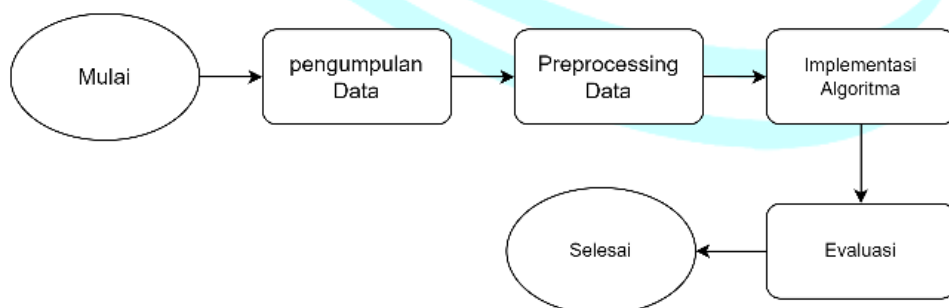
Objek dalam penelitian ini adalah kasus gizi buruk yang terjadi di Provinsi Jawa Barat. Penelitian ini berfokus pada data kejadian gizi buruk yang dilaporkan di tingkat kabupaten/kota, yang mencerminkan kondisi kesehatan masyarakat, terutama pada kelompok usia rentan seperti balita. Pemilihan objek ini didasarkan pada pentingnya isu gizi buruk sebagai salah satu indikator kesejahteraan dan kualitas hidup masyarakat.

Data yang digunakan diperoleh dari situs resmi Pemerintah Provinsi Jawa Barat, yaitu melalui portal <https://opendata.jabarprov.go.id/id/dataset/jumlah-kejadian-balita-gizi-buruk-berdasarkan-desakelurahan-di-jawa-barat>. Pelaksanaan penelitian dilakukan di Universitas Buana Perjuangan Karawang.

Dengan menjadikan kasus gizi buruk di Jawa Barat sebagai objek penelitian, diharapkan hasil yang diperoleh dapat memberikan gambaran mengenai distribusi dan pola kasus, serta menjadi dasar untuk pengambilan keputusan yang lebih tepat dalam upaya pencegahan dan penanggulangan masalah gizi buruk di daerah tersebut.

3.2 Prosedur Penelitian

Prosedur pada penelitian ini dapat dilihat pada Gambar 3.1



Gambar 3.1 Prosedur Penelitian

3.2.1 Pengumpulan Data

Sumber data diperoleh dari website <https://opendata.jabarprov.go.id/data/set/jumlah-kejadian-balita-gizi-buruk-berdasarkan-desakelurahan-di-jawa-barat> pada tanggal 01 April 2025 mengenai data balita gizi buruk berdasarkan kategori di Jawa Barat dari tahun 2019-2024. Pengumpulan data yang diperoleh berjumlah 486 baris dalam data gizi buruk berdasarkan kategori gizi buruk dan tahun, yang berisi Id, kode provinsi, nama provinsi, kode kabupaten, nama kabupaten_kota, kategori gizi_buruk, jumlah balita, satuan, dan tahun. Berikut adalah dataset yang digunakan pada penelitian ini.

Gambar 3.2 Data Gizi Buruk

Id	Kode provinsi	Nama provinsi	Kode kabupaten	Nama kabupaten_kota	Kategori gizi_buruk	Jumlah balita	Satuan	Tahun
1	32	Jawa Barat	3201	Kab. Bogor	Balita gizi kurang	20999	Orang	2019
2	32	Jawa Barat	3201	Kab. Bogor	Balita gizi pendek	86706	Orang	2019
3	32	Jawa Barat	3201	Kab. Bogor	Balita gizi kurus	16179	Orang	2019
4	32	Jawa Barat	3202	Kab. Sukabumi	Balita gizi kurang	8016	Orang	2019
5	32	Jawa Barat	3202	Kab. Sukabumi	Balita gizi pendek	11352	Orang	2019
...	...	...	...	...	...	...	...	...
482	32	Jawa Barat	3278	Kota Tasikmalaya	Balita gizi pendek	5633	Orang	2024
483	32	Jawa Barat	3278	Kota Tasikmalaya	Balita gizi kurus	4677	Orang	2024
484	32	Jawa Barat	3279	Kota Banjar	Balita gizi kurang	810	Orang	2024
485	32	Jawa Barat	3279	Kota Banjar	Balita gizi pendek	846	Orang	2024
486	32	Jawa Barat	3279	Kota Banjar	Balita gizi kurus	1273	Orang	2024

3.2.2 Preprocessing

Tahapan *preprocessing* data dilakukan sebelum proses pengklasteran dengan tujuan agar data yang digunakan dapat menghasilkan hasil klaster yang optimal. Proses ini meliputi pembersihan data, seleksi data, transformasi data, serta normalisasi.

1 Pembersihan Data

Langkah awal dalam proses *preprocessing* adalah melakukan pembersihan data. Pada penelitian ini, pembersihan data dilakukan dengan cara menghapus nilai yang hilang (*missing value*) serta mengeliminasi data yang duplikat.

2 Seleksi Data

Seleksi data bertujuan untuk menghapus atribut-atribut yang tidak diperlukan. Hal ini dilakukan karena terdapat beberapa atribut yang dianggap tidak relevan atau kurang dibutuhkan dalam proses analisis.

1 Transformasi Data

Transformasi data digunakan untuk merubah bentuk suatu data dari sekumpulan data yang akan dibutuhkan dalam dataset. Data di transformasikan kedalam bentuk yang sudah ditambahkan dari tahun 2019-2024 untuk mempermudah pada proses pengolahan data.

3.2.3 Implementasi Algoritma

Tahap selanjutnya adalah implementasi algoritma *K-Means*, dilanjut dengan *Hierarchical Clustering* menggunakan *Python*. Dalam proses ini, digunakan dua algoritma clustering, yaitu *K-Means* dan *Hierarchical Clustering*, yang bertujuan untuk melihat dan membandingkan hasil pengelompokan dari masing-masing metode. Berikut merupakan pseudocode untuk Algoritma *K-Means* (Cevin et al., n.d.).

Tabel 3.1 Pseudocode Algoritma *K-Means*

Input:
Data
Jumlah kluster
max_iter: Jumlah iterasi maksimum (opsional)
Output:
Centroid
Kluster
Metode:
1. Inisialisasi:
a. Pilih secara acak k centroid awal dari data.
b. Tetapkan max_iter, jika perlu.
2. Ulangi langkah berikut sampai konvergensi atau mencapai max_iter:
a. Assign Step
Untuk setiap data point di dalam data:
1. Hitung jarak dari data point ke masing-masing centroid.
2. Tetapkan data kedalam kluster dengan centroid terdekat.
b. Update Step
Untuk setiap kluster:
1. Hitung centroid baru.
c. Konvergensi
1. Jika centroid tidak berubah atau perubahan sangat kecil, hentikan iterasi.
3. Output:
1. Return centroid dari setiap kluster.
2. Return assignment data point ke masing-masing kluster.

Setelah melakukan perhitungan menggunakan *K-Means*, kemudian melakukan perhitungan menggunakan *Hierarchical Clustering*. Berikut adalah pseudocode *Hierarchical Clustering* (Nafilah et al., 2024)

Tabel 3.2 Pseudocode Algoritma *Hierarchical Clustering***Input:**

1. Data points $D = \{x_1, x_2, \dots, x_n\}$
2. Jumlah cluster yang diinginkan: K
3. Metode pengukuran jarak: (misalnya Euclidean Distance)
4. Linkage method: (misalnya Ward, Single, Complete, Average)

Output:

K buah cluster hasil pengelompokan

1. Inisialisasi:

Setiap data point dianggap sebagai satu cluster $\rightarrow$ total ada n cluster.

2. Hitung jarak antar semua pasangan cluster (menggunakan metode linkage yang dipilih).

3. Ulangi langkah berikut hingga jumlah cluster tersisa = K :

- a. Temukan dua cluster yang memiliki jarak terkecil.
- b. Gabungkan dua cluster tersebut menjadi satu cluster baru.
- c. Perbarui matriks jarak antar cluster (dengan metode linkage yang dipilih).

4. Kembalikan hasil cluster yang terbentuk.

3.2.4 Evaluasi

Dalam penelitian ini, proses evaluasi dilakukan untuk menilai kualitas objek yang dikelompokkan ke dalam klaster hasil dari penerapan algoritma K-Means dan Hierarchical Clustering, dengan memanfaatkan metode Silhouette Coefficient. Metode Silhouette Coefficient sendiri merupakan kombinasi dari dua konsep, yaitu cohesion dan separation.

1. Cohesion berfungsi untuk mengukur kedekatan antar objek dalam satu klaster.
2. Sedangkan separation digunakan untuk menilai seberapa jauh suatu klaster terpisah dari klaster lainnya.

Nilai silhouette coefficient yang ideal adalah nilai rata-rata yang paling tinggi atau mendekati angka 1, karena menunjukkan bahwa objek-objek dalam klaster saling berdekatan dan cukup terpisah dari klaster lain (Simanjuntak & Khaira, 2021).

