

BAB III

METODE PENELITIAN

3.1. Objek Penelitian

Penelitian ini menggunakan data kriminalitas yang diambil dari Polres Kabupaten Karawang, data merupakan laporan kasus kriminalitas pada tahun 2020 sampai tahun 2022. Jumlah data yang di pakai dalam penelitian ini sebanyak 5539 data dengan jumlah atribut 6, diantaranya variabel TGL LP (Tanggal Lapor), TGL.KEJ (Tanggal Kejadian), Perkara, TKP, Kecamatan dan Wilkum Sek.

3.2. Lokasi dan Waktu

Lokasi penelitian ini dilakukan di Polres Kabupaten Karawang sejak bulan Oktober 2022 sampai bulan Januari 2023. Secara rinci ditunjukkan pada Tabel 3.1

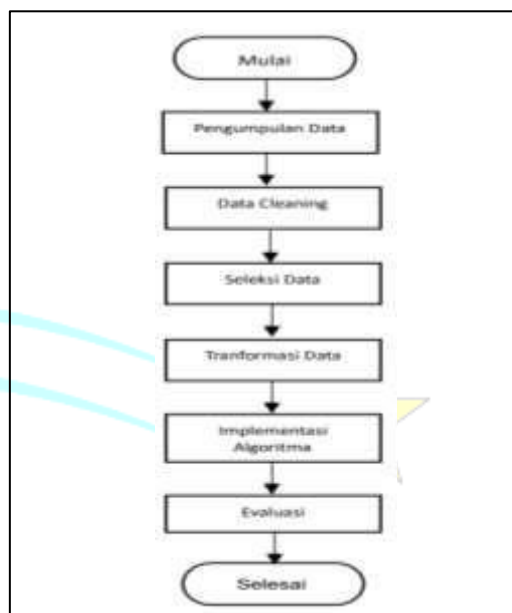
Tabel 2. 1 Waktu Penelitian

No	Kegiatan	2023												2024											
		11				12				1				2				3				4			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	Pengumpulan Data																								
2	Analisis Data																								
3	Pengolahan dan Seleksi Data																								
4	Implementasi Algoritma																								
6	Evaluasi																								

3.3. Prosedur Penelitian

Penelitian ini terdiri dari empat tahap yang dilakukan secara berurutan, yaitu pengumpulan data, *preprocessing*, implementasi algoritma serta evaluasi dengan

menguji validasi *cluster* yang dihasilkan oleh kedua algoritma tersebut, flowchart prosedur penelitian dapat dilihat pada gambar berikut.



Gambar 3. 1 Flowcart Prosedur Penelitian

3.3.1. Pengumpulan Data

Pada penelitian ini mengambil data dari polres Kabupaten Karawang, data yang diperoleh merupakan data tindakan kriminalitas di Kabupaten Karawang dari tahun 2020-2022 yang terdiri dari 2539 data dan 6 variabel, diantaranya variabel TGL LP (Tanggal Laporan), TGL.KEJ (Tanggal Kejadian), Perkara, TKP, Kecamatan dan Wilkum Sek.

Tabel 3. 1 Data Kriminalitas

No	TGL LP	TGL.KEJ	PERKARA	KEC	LOKASI	Wilkum Sek
1	24/12/2021	24/12/2021	CURANMOR	PURWASARI	PEMUKIMAN	KARAWANG
2	24/12/2021	15/12/2021	KEKERASAN	TELUKJAMBE TIMUR	PEMUKIMAN	KARAWANG
3	24/12/2021	24/12/2021	CURANMOR	TELUKJAMBE TIMUR	PERKANTORAN	KARAWANG
4	24/12/2021	24/12/2021	KEKERASAN	TELUKJAMBE TIMUR	PEMUKIMAN	KARAWANG
5	25/12/2021	15/12/2021	CURANMOR	KARAWANG BARAT	PEMUKIMAN	KARAWANG
6	26/12/2021	23/12/2021	KEKERASAN	JATISARI	JALAN UMUM	KARAWANG
7	27/12/2021	26/12/2021	CURANMOR	KARAWANG TIMUR	JALAN UMUM	KARAWANG
8	28/12/2021	01/12/2021	KEKERASAN	CIBUAYA	PEMUKIMAN	KARAWANG
9	28/12/2021	28/12/2021	CURANMOR	CILAMAYA WETAN	PEMUKIMAN	KARAWANG
10	24/12/2021	27/12/2021	KEKERASAN	KARAWANG TIMUR	PERKANTORAN	KARAWANG

2537	23/12/2021	20/12/2021	CURANMOR	TEGALWARU	PERKANTORAN	TEGALWARU
2538	23/12/2021	19/12/2021	CURANMOR	TEGALWARU	JALAN UMUM	TEGALWARU

3.3.2. Preprocessing Data

Setelah melakukan pengumpulan data, langkah selanjutnya yaitu melakukan *preprocessing* data dengan 3 tahapan yaitu data *Cleaning*, seleksi data dan transformasi data. Berikut adalah tahapan dari proses *preprocessing* data.

1. Data Cleaning

Pada proses pembersihan data (*data Cleaning*) dilakukan pengecekan data yang tidak konsisten, noise dan nilai yang hilang, pada proses ini dilakukan pengecekan terhadap keseluruhan data. Setelah dilakukan pengecekan pada data, tidak terdapat data noise dan nilai yang hilang. Data hasil *Cleaning* dapat dilihat pada berikut.

Tabel 3. 2 Hasil Data *Cleaning*

Coloumn	Missing_Values	Percent_Missing %
TGL LP	0	0.0
TGL.KEJ	0	0.0
PERKARA	0	0.0
KEC	0	0.0
LOKASI	0	0.0
Wilkum Sek	0	0.0

2. Seleksi Data

Setelah proses pembersihan data tidak ditemukan nilai yang hilang atau noise, maka dilanjutkan ke proses seleksi data. Pada proses ini menghilangkan variabel yang tidak di perlukan saat proses penelitian. Variabel yang tidak dipakai dan dihilangkan pada penelitian ini adalah variabel TGL LP, TGL. KEJ, LOKASI dan WILKUM SEK, sedangkan variabel yang digunakan hanya variabel KECAMATAN dan PERKARA. Berikut hasil data setelah di seleksi.

Tabel 3. 3 Hasil Seleksi Data

No	PERKARA	KECAMATAN
1	CURANMOR	PURWASARI
2	UU_FIDUSIA	TELUKJAMBE TIMUR

3	UU_FIDUSIA	TELUKJAMBE TIMUR
:	:	:
2538	CURANMOR	TIRTAMULYA
2539	CURANMOR	KLARI

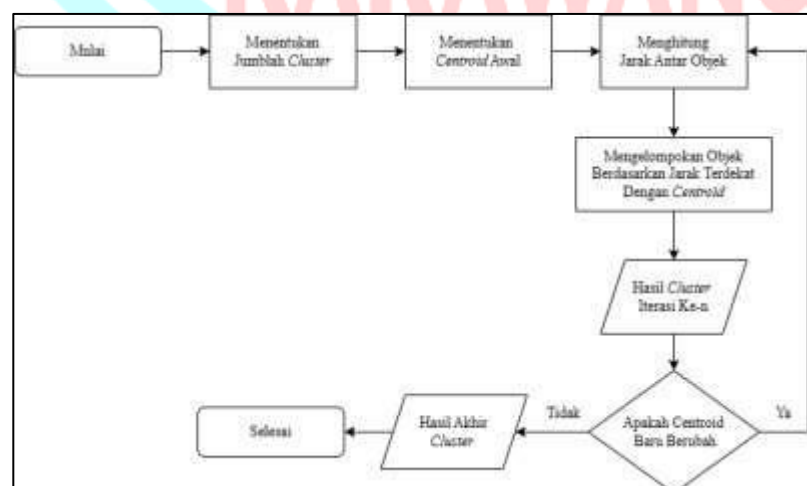
3. Transformasi Data

Pada proses transformasi data yaitu merubah data menjadi numerik dengan cara merekap data, pada proses ini adalah mengubah atribut Perkara dengan merekap jumlah kasus kejahatan di setiap kecamatan, berikut variabel yang akan digunakan pada penelitian ini; Abors, Cabul, Curanmor, Kekerasan, Pemalsuan, Pembunuhan, Pemasaran, Penculikan, Pencurian, Pengeroyokan, Pengrusakan dan Penipuan.

3.3.3. Implementasi Algoritma

1. Algoritma *K-Means*

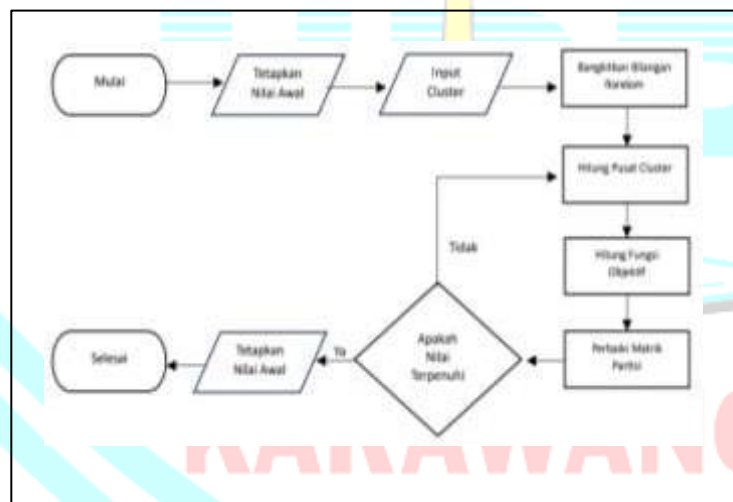
Cara kerja algoritma *K-Means Clustering* yaitu dengan mempartisi data ke dalam beberapa *cluster* dengan berbasis jarak. Algoritma *K-Means* ini mempartisi data berdasarkan karakteristik yang dimiliki tiap data. Data yang digunakan adalah pengelompokan data masing-masing jenis kriminalitas dari tiap kecamatan berdasarkan jumlah kasus. Kemudian dilakukan *clustering* dengan menentukan jumlah *K cluster* pada dataset, menginisialisasi *K* pusat *cluster* secara acak dan untuk menentukan pusat *cluster* (*centroid*). Alur algoritma *K-Means* secara rinci dapat di lihat pada Gambar 3.2



Gambar 3. 2 Flowchart Cara Kerja Algoritma *K-Means*

2. Algoritma *Fuzzy C-Means*

Cara kerja algoritma *Fuzzy C-Means* yaitu dengan membuat model pengelompokan data berdasarkan derajat keanggotaan yang berbeda antara 0 hingga 1. Data yang digunakan adalah data yang di kelompokkan tiap kecamatan berdasarkan jumlah data seluruh kejadian, informasi kasus (jumlah tiap kasus dan keterangan kasus) yang sudah di lakukan perkalian sesuai dengan bobotnya masing-masing. Kemudian diolah dengan Menentukan terlebih dahulu nilai perhitungan awal, menentukan kembali pusat *cluster* yang baru dengan cara menghitung rata-rata dari keanggotaan *cluster* yang sekarang, pengalokasian objek ke *cluster* terdekat dengan menggunakan *centroid* yang baru dan Cek kondisi berhenti sampai iterasi lebih dari 100 atau nilai kesalahan terkecil lebih kecil dari 0,0001. Alur algoritma *Fuzzy C-Means* dapat dilihat pada Gambar 3.3



Gambar 3. 3 Flowchart Cara Kerja Algoritma *Fuzzy C-Means*

3.3.4. Evaluasi

Tahap evaluasi ini diambil dari hasil akurasi algoritma *K-Means* dan *Fuzzy C-Means*, menggunakan metode evaluasi kluster *Koefisien silhouette*. Kemudian algoritma yang mempunyai nilai akurasi yang tinggi merupakan algoritma terbaik pada proses Klustering ini. *Silhouette Coefficient* digunakan untuk mengevaluasi sejauh mana suatu objek ditempatkan dengan baik dalam suatu *cluster*.