

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

1. Hasil Penelitian Andrian Dwi Putra (2020)

Penelitian Andrian Dwi Putra (2020) yang berjudul "Faktor-Faktor yang Memengaruhi Tingkat Kriminalitas di Indonesia Tahun 2018" menggunakan metode deskriptif kualitatif. Penelitian ini bertujuan untuk mengidentifikasi faktor-faktor yang mempengaruhi tingkat kriminalitas di Indonesia pada tahun 2018 serta mengukur besarnya pengaruh masing-masing faktor.

Hasil penelitian perhitungan koefisien jalur model struktural kemiskinan (X4) didapatkan pengaruh langsung dari variabel jumlah penduduk sebesar -0,213, variabel pendidikan sebesar -0,449, dan variabel pengangguran sebesar 0,073. Hasil uji signifikansi menunjukkan nilai-p untuk pengaruh kemiskinan oleh variabel jumlah penduduk sebesar 0,255 (nilai-p \geq 0,05), variabel pendidikan sebesar 0,038 (nilai-p $<$ 0,05), dan variabel pengangguran sebesar 0,732 (nilai-p \geq 0,05). Dapat disimpulkan bahwa variabel pendidikan (X2) signifikan mempengaruhi kemiskinan (X4), sedangkan variabel jumlah penduduk (X1) dan pengangguran (X3) tidak signifikan mempengaruhi kemiskinan (X4) di Indonesia tahun 2018. Model yang dihasilkan memiliki nilai R-Square sebesar 0,178 yang artinya bahwa variabel eksogen (variabel bebas) dalam penelitian dapat menjelaskan variabel kemiskinan sebanyak 17,8% ,dan

82,2% lainnya dijelaskan oleh variabel lainnya, serta dihasilkan variabel sisanya (e_4) sebesar 0,907.

Kesimpulan penelitian menunjukkan bahwa jumlah penduduk dan kemiskinan memiliki pengaruh signifikan terhadap tingkat kriminalitas di Indonesia, sementara pendidikan dan pengangguran tidak menunjukkan pengaruh signifikan secara parsial. Model yang dihasilkan menunjukkan nilai R-Square yang rendah, sehingga disarankan untuk memodifikasi model atau menambahkan variabel penjelas dalam penelitian selanjutnya untuk mendapatkan hasil yang lebih baik.

2. Hasil Penelitian Insiyah (2021)

Penelitian Insiyah (2021) yang berjudul "Penerapan Metode Ward Clustering untuk Pengelompokan Daerah Rawan Kriminalitas di Jawa Timur Tahun 2021" menggunakan metode deskriptif kualitatif.

Diperoleh hasil Cluster 1 beranggotakan 31 Kabupaten/kota yaitu Pacitan, Ponorogo, Trenggalek, Tulungagung, Blitar, Kediri, Malang, Situbondo, Probolinggo, Pasuruan, Sidoarjo, Mojokerto, Jombang, Nganjuk, Madiun, Magetan, Ngawi, Bojonegoro, Tuban, Lamongan, Sampang, Pamekasan, Sumenep, Kota Kediri, Kota Blitar, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, kota Surabaya, dan kota Batu. Cluster 2 beranggotakan 5 kabupaten/kota yaitu Lumajang, Banyuwangi, Bondowoso, Bangkalan, dan Kota Malang. Cluster 3 beranggotakan 2 kabupaten/kota yaitu Jember dan Gresik..

Kesimpulan Diperoleh 3 cluster kecamatan yang memiliki kemiripan karakteristik dari studi clustering dengan metode Ward untuk kabupaten/kota yang rawan kriminalitas di Jawa Timur Tahun 2021.

3. Hasil Penelitian Siti Nur Aisah (2022)

Penelitian Siti Nur Aisah (2022) yang berjudul "Implementasi Fuzzy C-Means Clustering (FCM) pada Pemetaan Daerah Potensi Transmigrasi di Jawa Timur" menggunakan metode deskriptif kualitatif. Penelitian ini mengklasterkan daerah kota/kabupaten di Provinsi Jawa Timur dengan metode Fuzzy C-Means menjadi tiga cluster: daerah berpotensi transmigrasi, daerah kurang berpotensi transmigrasi, dan daerah tidak berpotensi transmigrasi. Parameter yang digunakan meliputi angka penduduk kurang mampu, pasangan usia subur, dan kepadatan penduduk.

Hasil analisis untuk tahun 2017 menunjukkan 12 daerah berpotensi transmigrasi, 22 daerah kurang berpotensi, dan 4 daerah tidak berpotensi. Pada tahun 2018, terdapat 4 daerah berpotensi transmigrasi, 12 daerah kurang berpotensi, dan 22 daerah tidak berpotensi. Untuk tahun 2019, jumlah daerah berpotensi transmigrasi meningkat menjadi 22, dengan 12 daerah kurang berpotensi dan 4 daerah tidak berpotensi. Sementara itu, pada tahun 2020, terdapat 14 daerah berpotensi transmigrasi, 4 daerah kurang berpotensi, dan 20 daerah tidak berpotensi.

Kesimpulan pengklasteran daerah kota/kabupaten di Provinsi Jawa Timur dengan menggunakan metode Fuzzy C-Means diperoleh tiga cluster yaitu cluster 1 berpotensi transmigrasi, cluster 2 Kurang berpotensi transmigrasi

dan cluster 3 tidak berpotensi transmigrasi. Dengan beberapa parameter yang digunakan berupa angka penduduk kurang mampu, pasangan usia subur, dan kepadatan penduduk.

4. Hasil Penelitian Yoni Aswan (2021)

Penelitian Yoni Aswan (2021) yang berjudul "Algoritma K-Means Clustering dalam Mengklasifikasi Data Daerah Rawan Tindak Kriminalitas (Polres Kepulauan Mentawai)" membahas tantangan Polres Kepulauan Mentawai dalam mengklasifikasikan daerah rawan kriminalitas, khususnya pencurian kendaraan bermotor, ke dalam kategori sangat rawan, cukup rawan, dan tidak rawan. Penelitian ini dilakukan di Polres Kepulauan Mentawai, yang bertugas memberikan keamanan di Kabupaten Kepulauan Mentawai, terdiri dari empat pulau besar dan 10 kecamatan, di mana tingkat kriminalitas terus meningkat setiap tahun. Peneliti menggunakan metode Algoritma K-Means Clustering, sebuah metode pengelompokan data non-hirarkis, untuk menganalisis kasus pencurian kendaraan bermotor dari tahun 2016 hingga 2020.

Hasil penelitian ini menggunakan Algoritma K-Means Clustering untuk memprediksi angka kejahatan di Kabupaten Kepulauan Mentawai, khususnya dalam mengidentifikasi wilayah yang rawan tindak pidana pencurian kendaraan bermotor. Data yang digunakan adalah data kasus pencurian kendaraan bermotor selama 5 tahun terakhir (2016–2020). Hasil pengujian menunjukkan bahwa Kecamatan Sipora Selatan merupakan wilayah yang paling rawan terhadap tindak pidana pencurian kendaraan bermotor, dibandingkan dengan kecamatan lainnya.

Kesimpulan berdasarkan hasil penelitian, Algoritma K-Means Clustering terbukti efektif dalam mempartisi wilayah berdasarkan tingkat kerawanan tindak pidana. Kecamatan Sipora Selatan teridentifikasi sebagai wilayah yang paling rentan terhadap kasus pencurian kendaraan bermotor,

sehingga dapat dijadikan fokus utama bagi pihak kepolisian untuk mengantisipasi lonjakan kejahatan yang akan datang. Implementasi sistem prediksi ini diharapkan dapat membantu Kepolisian Daerah Kepulauan Mentawai dalam mengklasifikasikan wilayah berdasarkan tingkat kerawanan, sehingga langkah pencegahan dan pengamanan dapat dilakukan lebih efektif.

5. Hasil Penelitian Hana Sugiastu Firdaus (2021)

Penelitian Hana Sugiastu Firdaus (2021) yang berjudul "Perbandingan Metode Fuzzy C-Means dan K-Means untuk Pemetaan Daerah Rawan Kriminalitas di Kota Semarang" membahas permasalahan kriminalitas di Kota Semarang, yang selama ini hanya dicatat tanpa visualisasi spasial. Penelitian ini bertujuan untuk membandingkan efektivitas dua metode clustering, yaitu Fuzzy C-Means dan K-Means, dalam memetakan daerah rawan kriminalitas. Fuzzy C-Means mengelompokkan data berdasarkan derajat keanggotaan, sementara K-Means menggunakan centroid kejadian kriminalitas.

Hasil Penelitian ini membandingkan metode Fuzzy C-Means dan K-Means dalam memetakan wilayah rawan kriminalitas di Kota Semarang berdasarkan data kejadian kriminalitas pada bulan Januari-Maret 2019, dengan total 164 kasus yang tersebar di 16 kecamatan. Hasil clustering dari kedua metode tersebut diverifikasi menggunakan data kejadian kriminalitas faktual. Dari uji verifikasi, metode Fuzzy C-Means menunjukkan akurasi rata-rata 71,23%, sedangkan metode K-Means memiliki akurasi rata-rata 65,45%. Dengan demikian, dapat disimpulkan bahwa Fuzzy C-Means menghasilkan klasifikasi wilayah rawan kriminalitas yang lebih akurat dibandingkan dengan metode K-Means untuk kasus di Kota Semarang.

Kesimpulan data dari 1.965 kasus kriminalitas selama tahun 2016-2018 menunjukkan bahwa kedua metode menghasilkan peta kerawanan yang berbeda. Uji Partition Coefficient Index (PCI) untuk Fuzzy C-Means menghasilkan nilai 0,818, sedangkan uji Silhouette Index (SI) untuk K-Means

menghasilkan nilai 0,569. Verifikasi data untuk tahun 2019 menunjukkan bahwa metode Fuzzy C-Means lebih unggul, dengan akurasi mencapai 71,23%.

6. Resti Noor Fahmi (2021)

Penelitian Resti Noor Fahmi (2021) yang berjudul "Analisis Pemetaan Tingkat Kriminalitas di Kabupaten Karawang Menggunakan Algoritma K-Means" bertujuan untuk mengevaluasi pengaruh faktor-faktor penyebab kriminalitas di Kabupaten Karawang dengan menggunakan algoritma K-Means. Penelitian ini menyusun paradigma penelitian, menguji koefisien jalur, dan mencari koefisien korelasi untuk memahami hubungan kausal antara variabel eksogen dan endogen, serta mengidentifikasi faktor-faktor yang paling berpengaruh terhadap tingkat kriminalitas. Data yang digunakan diperoleh dari Badan Pusat Statistik dan Kepolisian Resort Kota Manado, mencakup periode tahun 1990 hingga 2008.

Hasil penelitian ini menggunakan metode K-Means Clustering untuk memetakan daerah rawan kriminalitas di Kabupaten Karawang pada tahun 2019 dan 2020, serta divisualisasikan menggunakan QGIS. Pada tahun 2019, hasil clustering menunjukkan 23 kecamatan masuk kategori tidak rawan, 3 kecamatan kategori rawan, dan 4 kecamatan kategori sangat rawan. Pada tahun 2020, terdapat sedikit perubahan dengan 22 kecamatan dalam kategori tidak rawan, 4 kecamatan kategori rawan, dan 4 kecamatan kategori sangat rawan. Evaluasi clustering menggunakan silhouette coefficient menghasilkan skor 0,52 pada tahun 2019 dan 0,54 pada tahun 2020, yang termasuk dalam kategori medium structure

Kesimpulan analisis menunjukkan bahwa jumlah penduduk dan tingkat pengangguran memiliki pengaruh langsung terhadap tingkat kriminalitas, sedangkan jumlah industri dan tingkat kemiskinan mempengaruhi secara tidak langsung.

Tabel 2.1 Penelitian terdahulu

Nama	Judul	Variable	Metode Analisis	Hasil Analisis
Andrian Dwi Putra, 2020	Faktor-Faktor yang Memengaruhi Tingkat Kriminalitas di Indonesia Tahun 2018	Jumlah penduduk, Pendidikan, Pengangguran, Kemiskinan	Deskriptif, Kualitatif	<p>Penelitian Andrian Dwi Putra (2020) yang berjudul "Faktor-Faktor yang Memengaruhi Tingkat Kriminalitas di Indonesia Tahun 2018" menggunakan metode deskriptif kualitatif. Penelitian ini bertujuan untuk mengidentifikasi faktor-faktor yang mempengaruhi tingkat kriminalitas di Indonesia pada tahun 2018 serta mengukur besarnya pengaruh masing-masing faktor.</p> <p>Hasil penelitian menunjukkan bahwa jumlah penduduk dan kemiskinan memiliki pengaruh signifikan terhadap tingkat kriminalitas di Indonesia, sementara pendidikan dan pengangguran tidak menunjukkan pengaruh signifikan secara parsial. Model yang dihasilkan menunjukkan nilai R-Square yang rendah, sehingga disarankan untuk memodifikasi model atau menambahkan variabel penjelas dalam penelitian selanjutnya untuk mendapatkan hasil yang lebih baik</p>
Insiyah, 2021	Penerapan Metode Ward Clustering Untuk Pengelompokan Daerah Rawan Kriminalitas Di Jawa Timur Tahun 2021	Kriminalitas	Deskriptif, Kualitatif	<p>Penelitian Insiyah (2021) yang berjudul "Penerapan Metode Ward Clustering untuk Pengelompokan Daerah Rawan Kriminalitas di Jawa Timur Tahun 2021" menggunakan metode deskriptif kualitatif. Hasil penelitian menunjukkan pembagian daerah rawan kriminalitas ke dalam tiga cluster sebagai berikut cluster 1 beranggotakan 31 Anggota, cluster 2 beranggotakan 5 cluster 3 beranggotakan 2 Anggota</p>
Siti Nur Aisah, 2022	Implementasi Fuzzy C-Means Clustering (FCM) pada Pemetaan Daerah Potensi	Transmigrasi Penduduk	Deskriptif, Kualitatif	<p>Penelitian Siti Nur Aisah (2022) yang berjudul "Implementasi Fuzzy C-Means Clustering (FCM) pada Pemetaan Daerah Potensi Transmigrasi di Jawa Timur" menggunakan metode deskriptif kualitatif. Penelitian ini mengklasterkan daerah kota/kabupaten di Provinsi Jawa Timur dengan metode Fuzzy C-Means menjadi tiga cluster: daerah berpotensi transmigrasi, daerah kurang berpotensi transmigrasi, dan daerah tidak berpotensi transmigrasi. Parameter yang digunakan meliputi angka penduduk kurang mampu, pasangan usia subur, dan kepadatan penduduk. Hasil analisis untuk tahun 2017 menunjukkan 12 daerah</p>

Transmigrasi di
Jawa Timur

berpotensi transmigrasi, 22 daerah kurang berpotensi, dan 4 daerah tidak berpotensi. Pada tahun 2018, terdapat 4 daerah berpotensi transmigrasi, 12 daerah kurang berpotensi, dan 22 daerah tidak berpotensi. Untuk tahun 2019, jumlah daerah berpotensi transmigrasi meningkat menjadi 22, dengan 12 daerah kurang berpotensi dan 4 daerah tidak berpotensi. Sementara itu, pada tahun 2020, terdapat 14 daerah berpotensi transmigrasi, 4 daerah kurang berpotensi, dan 20 daerah tidak berpotensi.

2.2 Teori Terkait

2.2.1. Pengertian Klustering

Pengertian dari klustering adalah suatu proses yang digunakan dalam mengolah sebuah data termasuk dalam kluster mana data tersebut akan ditempatkan (Pratama, 2022). Klustering digunakan untuk mengelompokkan dan mengidentifikasi data atau informasi yang memiliki karakteristik tertentu. Teknik ini bertujuan untuk mengelompokkan data ke dalam grup-grup yang memiliki kesamaan sehingga pola atau hubungan yang mendasarinya bisa lebih mudah diidentifikasi. (Firman, Halik, dan Septiana 2022).

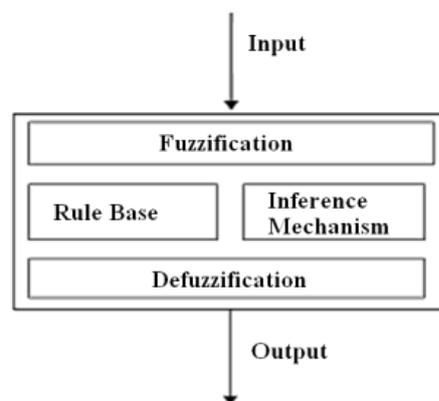
2.2.2. Metode Fuzzy

Fuzzy memiliki derajat keanggotaan yang diketahui dengan nilai berkisar dari 0 sampai 1. Fuzzy secara bahasa didefinisikan sebagai tidak jelas atau samar-samar. Nilainya bisa besar atau salah pada saat yang bersamaan. Logika fuzzy dianggap mampu menghubungkan input dan output tanpa mengabaikan faktor-faktor lain. Logika fuzzy memiliki akses ke data yang ada dan dianggap sangat fleksibel. Bentuk yang biasanya digunakan untuk membentuk suatu sistem yang penalarannya serupa dengan firasat atau perasaan seseorang adalah salah satu model aturan fuzzy. Meskipun model ini memberikan akurasi tinggi, proses perhitungannya cukup rumit dan membutuhkan waktu yang lama (Mursalin 2020). Berikut logika fuzzy:

1. Fuzzifikasi input adalah proses mengubah nilai crisp menjadi nilai keanggotaan menggunakan fungsi keanggotaan. Proses ini memungkinkan nilai input yang spesifik diubah menjadi derajat

keanggotaan dalam himpunan fuzzy, yang mempermudah penanganan data dalam sistem logika fuzzy.

2. Pengaplikasian aturan fuzzy melibatkan penghubungan antara input dan output dengan menggunakan aturan fuzzy. Aturan-aturan ini mendefinisikan bagaimana nilai keanggotaan dari input mempengaruhi hasil output dalam sistem logika fuzzy.
3. Inferensi fuzzy adalah proses menghitung nilai output fuzzy berdasarkan aturan fuzzy yang telah ditetapkan. Selama tahap ini, sistem menggunakan aturan-aturan fuzzy untuk menggabungkan nilai keanggotaan input dan menghasilkan nilai keanggotaan output yang sesuai.
4. Defuzzifikasi output adalah proses mengubah nilai output fuzzy kembali menjadi nilai crisp, atau nilai yang konkret dan terukur. Ini dilakukan untuk mengembalikan hasil dari bentuk fuzzy ke bentuk yang dapat diterapkan secara langsung dalam sistem, seperti nilai numerik yang dapat digunakan dalam aplikasi praktis.



Gambar 2. 1 Gambar alur proses logika fuzzy

2.2.3. Metode Silhouette Coefficient

Hasil dari algoritma Hierarchical Clustering yang menghasilkan sejumlah cluster perlu dilakukan evaluasi. Evaluasi dilakukan guna mengetahui seberapa mirip atau seberapa dekat jarak antar objek dalam suatu cluster dan seberapa jauh jarak cluster terpisah dengan cluster lain. Semakin mirip atau semakin dekat jarak antar objek dalam suatu cluster dan semakin jauh jarak objek terhadap objek di cluster lain maka semakin baik kualitas Klasterisasi tersebut.

Silhouette Coefficient	Interpretation
0.71 – 1.0	A strong structure is found
0.51 – 0.7	A reasonable structure is found
0.26 – 0.50	Structure may be artificial
< 0.26	No substantial structure found

Menghitung jarak rata-rata dari suatu data misalkan i dengan semua data lain yang berada dalam klaster yang sama (a_i).

$$a(i) = \frac{1}{|A|-1} \sum_{j \in A, j \neq i} d(i, j)$$

Dengan j adalah data lain dalam satu cluster A dan $d(i,j)$ adalah jarak antara data i dengan j .

Menghitung rata-rata jarak dari data i tersebut dengan semua data di klaster lain, lalu ambil nilai terkecilnya.

$$d(i, C) = \frac{1}{|A|-1} \sum_{j \in C} d(i, j)$$

Dengan $d(I,C)$ adalah jarak rata-rata data I dengan semua objek pada cluster lain C
dimana $A \neq C$

$$b(i) = \min_{C \neq A} d(i,C)$$

Rumus Silhouette coefficient adalah:

$$s(i) = (b(i) - a(i)) / \max(a(i), b(i))$$

Dengan $s(i)$ adalah semua rata-rata pada semua kumpulan data.

2.2.4. Metode Fuzzy C-Means (FCM)

Metode Fuzzy C-Means (FCM) adalah teknik pengelompokan data yang menentukan keterlibatan setiap titik data dalam sebuah cluster berdasarkan derajat keanggotaan. Algoritma ini, yang diperkenalkan oleh J.C. Bezdek pada tahun 1981, merupakan metode pengelompokan yang memerlukan penetapan jumlah cluster terlebih dahulu. Metode Fuzzy C-Means (FCM) adalah teknik pengelompokan data yang menentukan keterlibatan setiap titik data dalam sebuah cluster berdasarkan derajat keanggotaan. Algoritma ini, yang diperkenalkan oleh J.C. Bezdek pada tahun 1981, merupakan metode pengelompokan yang memerlukan penetapan jumlah cluster terlebih dahulu (Sandhika Jaya 2020).

1. Masukkan data yang akan peneliti gunakan , yang akan di *cluster* ,kedalam sebuah matriks berukuran a x b, dengan a adalah jumlah data yang akan di kelompokkan dan b sebagai atribut dari a.
2. Tentukan

- a. Jumlah cluster = c;
- b. Pangkat/pembobot = w;
- c. Maksimum iterasi = Maks_Iterasi;
- d. Error yang diharapkan = ξ ;
- e. Fungsi Objektif awal = $P_0 = 0$;
- f. Iterasi awal = $t = 1$;

3. Tentukan bilangan acak untuk μ_{ik} (dengan $i=1,2,\dots,m$ dan $k=1,2,\dots,c$) sebagai elemen matriks partisi awal U, dengan X_i adalah data ke-I, sesuai dengan persamaan 2.2.

$$U = \begin{bmatrix} \mu_{11}(X_1) & \mu_{21}(X_1) & \dots & \mu_{c1}(X_1) \\ \vdots & \dots & \dots & \vdots \\ \mu_{1i}(X_i) & \dots & \dots & \mu_{ci}(X_i) \end{bmatrix} \quad (2.1)$$

Dengan jumlah setiap kolom dalam satu baris adalah 1 (satu). sesuai dengan persamaan 2.3.

$$\sum_{i=1}^c \mu_{ci} = 1 \quad (2.2)$$

4. Hitung pusat cluster ke-k : V_{kj} , sesuai dengan persamaan 2.4 dengan $k=1,2,\dots,c$ dan $j = 1,2,\dots,n$

$$V_{kj} = \frac{\sum_{i=1}^n (U_{ik})^w * X_{ij}}{\sum_{i=1}^n (U_{ik})^w} \quad (2.3)$$

5. Hitung fungsi objektif pada iterasi ke-t, P_t : sesuai dengan persamaan 2.5.

$$P_t = \sum_{i=1}^m \sum_{k=1}^c \left(\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] \right) (U_{ik})^w \quad (2.4)$$

6. Hitung perubahan derajat keanggotaan setiap data pada setiap cluster (memperbaiki matriks partisi U) dengan : sesuai dengan persamaan 2.6.

$$U_{ik} = \frac{\left[\sum_{j=1}^n (x_{ij} - V_{kj})^2 \right]^{\frac{-1}{w-1}}}{\sum_{j=1}^n \left[\sum_{j=1}^n (x_{ij} - V_{kj})^2 \right]^{\frac{-1}{w-1}}} \quad (2.5)$$

dengan : $i = 1, 2, \dots, n$ dan $k = 1, 2, \dots, c$.

7. Cek kondisi berhenti : sesuai dengan persamaan 2.7.

Jika : $(|P_t - P_{t-1}| < \xi)$ atau $(t > \text{Maks_Iterasi})$ maka (2. 6)

berhenti ;

Jika tidak : $t = t+1$, ulangi langkah 4.